
Design, Synthesis and Evaluation of Inhibitors of POT1-DNA Interactions

Adnan Malik, BSc.

**Thesis submitted to the University of Nottingham
for the degree of Doctor of Philosophy**

July 2013

Abstract

The unlimited replicative potential of cells is one of the hallmarks of cancer. Telomeres, DNA structures found at the ends of chromosomes have attracted a great deal of interest in recent years as potential anti-cancer drug targets since they play an important role in cancer cell immortality.

The repetitive TTAGGG sequences of telomeres are complexed to a group of six indispensable proteins, one of which is the protection of telomeres 1 (POT1) protein. This specialised protein binds to a ten nucleotide single stranded DNA sequence at the ends of chromosomes and plays an important role in telomere capping and length regulation. It has recently been proposed that the key function of POT1 is to suppress a potent DNA damage response at telomeres thereby protecting chromosome tips from being recognised as sites of DNA damage. Deletion of POT1 from telomeres in a variety of organisms including humans results in cytogenetic aberrations, senescence and cell death. These results indicate that POT1 is an integral telomere end-protection protein which is necessary for continued cellular proliferation and therefore POT1 is becoming a promising new target in cancer.

Using a structure-based approach, several small molecule inhibitors of POT1 have been designed to affect telomere integrity by disrupting the binding interaction of human POT1 with its target DNA sequence thereby driving cancer cells into senescence/apoptosis. Using a range of computational tools, a suitable drug binding pocket in POT1 has been identified and the *de novo* design of a specific class of POT1 inhibitor was completed. Using this novel scaffold, a small focussed library of

hit-like compounds were synthesised and screened in a new POT1 fluorescence polarisation displacement assay developed by scientists at the University of Nottingham. In total, over 90 small molecule inhibitors based on two different scaffolds: pyrido[1,2-*a*]pyrimidines and sulfathiazoles have been synthesized with some inhibitors effectively decreasing POT1-DNA binding between 10-54% at 100 μ M ligand concentration. The biological results have established that electron-withdrawing substituents on the pendent phenyl ring of the pyrimidine core are essential for strong binding. These results have the potential to guide future development of improved lead compounds as therapeutics for the treatment of cancer.

Acknowledgements

First of all, I owe my deepest gratitude to my supervisors Dr Charles Laughton (University of Nottingham) and Dr Marc Hummersone (Pharminox Ltd) for their encouragement, guidance and support.

I would also like to acknowledge all those who have directly contributed to this thesis: Dr Lodwyk Dekker and his postdoctoral research associates Dr Prudence Mutowo and Dr Lucy Johnson for the POT1 assay development and for the biological compound screening without which this project would never had been possible in the first place. I am also grateful to Dr Mark Frigerio who has made available his support in a number of ways throughout my 3 years as a mentor and someone who I can always turn to for advice regarding the project, Lucy who trained me in using the POT1 assay and Dr Ian Hutchinson for his much help and needed advice in the initial stage of chemical synthesis.

I also wish to express my warmest thanks to all my colleagues at the University of Nottingham and Pharminox Ltd especially Julie Roy, Yazhou Li and Dr Tummala Reddy for sharing their knowledge and expertise and whom it has been a pleasure to work with.

Lastly, I offer my regards and blessings to my loving parents and family for their love and support. All that I have achieved so far is actually their achievement. The words are insufficient to express my feelings for them. They really did a lot for me in these 3 years.

This work was supported by EPSRC and Pharminox Ltd.

Contents

Abstract i

Acknowledgements..... iii

Contents iv

Abbreviations viii

Chapter 1: Cancer and Telomeres

1.1 Introduction 1

1.1.1 Self-sufficiency in growth signals 2

1.1.2 Insensitivity to growth-inhibitory signals 2

1.1.3 Evasion of programmed cell death (apoptosis) 3

1.1.4 Limitless replicative potential 4

1.1.5 Sustained angiogenesis 5

1.1.6 Tissue invasion and metastasis 5

1.1.7 Hallmarks in cancer therapy 6

1.2 Telomeres 8

1.2.1 Shelterin complex 10

1.2.1.1 TRF1 and TRF2 11

1.2.1.2 RAP1 13

1.2.1.3 TIN2 13

1.2.1.4 TPP1 14

1.2.1.5 POT1 15

1.2.1.6 Shelterin accessory factors 16

1.3 End replication problem 17

1.4 Telomere shortening triggers DDR 19

1.5 DNA damage response (DDR) 20

1.6 Dysfunctional telomeres and cancer 24

1.7 Telomerase 27

1.7.1 Telomerase activity in cells 27

1.7.2 Structure and function of telomerase 28

| | |
|---|----|
| 1.7.3 Telomerase inhibition..... | 32 |
| 1.7.3.1 Reverse transcriptase inhibitors..... | 32 |
| 1.7.3.2 Targeting telomerase RNA | 34 |
| 1.7.4 Limitations of telomerase inhibition..... | 36 |

Chapter 2: Computer aided drug design (CADD)

| | |
|--|----|
| 2.1 Introduction..... | 38 |
| 2.2 CADD strategies in drug discovery | 39 |
| 2.2.1 Structure-based drug design..... | 39 |
| 2.2.2 Ligand-based drug design | 40 |
| 2.3 Binding pockets | 41 |
| 2.3.1 Geometric algorithms..... | 42 |
| 2.3.2 Probe/energy based algorithms | 42 |
| 2.4 Molecular docking..... | 43 |
| 2.4.1 Search algorithm..... | 44 |
| 2.4.1.1 Monte Carlo methods..... | 45 |
| 2.4.1.2 Genetic algorithms..... | 45 |
| 2.4.1.3 Distance geometry methods | 46 |
| 2.4.1.4 Incremental construction methods | 47 |
| 2.4.2 Scoring functions | 47 |
| 2.4.2.1 Force field-based methods | 48 |
| 2.4.2.2 Empirical free energy methods | 49 |
| 2.4.2.3 Knowledge based methods..... | 50 |
| 2.4.3 GOLD | 50 |
| 2.4.3.1 Genetic algorithm..... | 51 |
| 2.4.3.2 Validation and comparison..... | 51 |
| 2.4.3.3 The fitness function..... | 52 |
| 2.4.3.3.1 GoldScore..... | 53 |
| 2.4.3.3.2 ChemScore..... | 54 |

Chapter 3: Basis of present investigation

| | |
|--|----|
| 3.1 Introduction | 55 |
| 3.2 Human POT1 as a viable anti-cancer target..... | 58 |
| 3.3 Human POT1 function | 60 |
| 3.4 Human POT1 crystal structure..... | 61 |

Chapter 4: Protein-DNA interaction analysis

| | |
|--|----|
| 4.1 Introduction | 66 |
| 4.2 Filter binding assay | 66 |
| 4.3 EMSA assay | 67 |
| 4.4 ITC..... | 67 |
| 4.5 SPR assay | 68 |
| 4.6 Fluorescence Polarisation assay (FP) | 69 |
| 4.6.1 Introduction | 69 |
| 4.6.2 Theory..... | 69 |

Chapter 5: POT1 binding site characterisation

| | |
|--|----|
| 5.1 Introduction | 73 |
| 5.1.1 POT1 potential binding sites | 73 |
| 5.1.2 Favourable POT1 binding pocket | 74 |
| 5.1.3 Conclusion | 81 |

Chapter 6: Ligand design and Hit identification

| | |
|--|-----|
| 6.1 Introduction | 83 |
| 6.2 Dock and design of POT1 inhibitor | 83 |
| 6.3 Synthetic accessibility dilemma | 89 |
| 6.4 Synthetic route for pyrimidine analogues | 91 |
| 6.5 Screening results and discussion | 94 |
| 6.6 Structural modifications of hit compounds | 105 |
| 6.6.1 Modifications of compound 18e | 105 |
| 6.6.2 Modifications of compound 18m | 108 |
| 6.6.3 Probing electron withdrawing substituents..... | 112 |

| | |
|--|-----|
| 6.6.4 Modification of ligand 32e | 116 |
| 6.6.5 Modification of N-acyl group | 117 |
| 6.6.6 SAR summary..... | 121 |

Chapter 7: Virtual compound libraries

| | |
|----------------------------------|-----|
| 7.1 Introduction | 122 |
| 7.2 Library design | 122 |
| 7.3 Library chemistry | 125 |
| 7.4 Results and discussion | 126 |

Chapter 8: Sulfathiazole analogues

| | |
|---|-----|
| 8.1 Introduction | 130 |
| 8.2 Selection of sulfathiazole-based inhibitor..... | 130 |
| 8.3 Sulfathiazole modifications | 131 |
| 8.4 Succinylsulfathiazole fragments..... | 136 |
| 8.5 Further modifications | 138 |

| | |
|---|------------|
| Chapter 9 Conclusions and Future work..... | 141 |
|---|------------|

| | |
|--|------------|
| Chapter 10: Experimental methods..... | 145 |
|--|------------|

| | |
|-------------------------|------------|
| Appendices | 232 |
|-------------------------|------------|

| | |
|-------------------------|------------|
| References | 236 |
|-------------------------|------------|

Abbreviations

| | |
|--------|--|
| μl | microlitre |
| μM | micromolar |
| °C | degree celsius |
| A | adenine |
| Å | angstrom |
| ALT | alternative lengthening of telomeres |
| aq. | aqueous |
| ARG | arginine |
| ASN | asparagine |
| ASP | aspartic acid |
| ATM | ataxia telangiectasia mutated |
| ATP | adenosine-5'-triphosphate |
| ATR | ataxia telangiectasia and Rad3 related |
| AZT | azidothymidine |
| BFB | breakage-fusion-bridge cycles |
| C | cytosine |
| CADD | computer aided drug design |
| CCDC | Cambridge crystallographic data centre |
| Cdc25 | cell division cycle 25 |
| cdk | cyclin dependent kinase |
| CDMT | 2-chloro-4-6,dimethoxy-1,3,5-triazine |
| Chk1/2 | checkpoint 1/2 |

| | |
|---------------------------------|---|
| CS ₂ CO ₃ | cesium carbonate |
| D- loop | displacement loop |
| Da | dalton(s) |
| DCC | dicyclohexylcarbodiimide |
| DCM | dichloromethane |
| DMA | dimethylacetamide |
| DDR | DNA damage response |
| DMF | dimethylformamide |
| DMSO | dimethyl sulfoxide |
| dec | decomposed |
| DNA | deoxyribonucleic acid |
| E | energy |
| EMSA | electrophoresis mobility shift assay |
| eq | equation |
| EtOAc | ethyl acetate |
| EtOH | ethanol |
| FP | fluorescence polarisation |
| G | guanine |
| GA | genetic algorithm |
| GOLD | genetic optimisation for ligand docking |
| GS | goldscore |
| h | hour(s) |
| H-bond | hydrogen bond |
| HCl | hydrochloric acid |
| HIV | human immunodeficiency virus |

| | |
|--------------------|--|
| HPLC | high performance liquid chromatography |
| HR | homologous recombination |
| Hsp90 | heat shock protein 90 |
| hTERC | human telomerase RNA component |
| hTERT | human telomerase reverse transcriptase |
| hTR | human telomerase RNA |
| Hz | hertz |
| IC ₅₀ | half maximal inhibitory concentration |
| ITC | isothermal titration calorimetry |
| kDa | kilo dalton(s) |
| LBDD | ligand based drug design |
| LEU | leucine |
| LYS | lysine |
| m.p. | melting point |
| M1/2 | mortality stage 1/2 |
| mg | milligram |
| MgSO ₄ | magnesium sulphate |
| min | minute(s) |
| mM | millimolar |
| mRNA | messenger RNA |
| NaHCO ₃ | sodium hydrogen carbonate |
| NHEJ | non-homologous end joining |
| nM | nanomolar |
| NMM | N-methylmorpholine |
| NMR | nuclear magnetic resonance |

| | |
|------------------------------------|---|
| OB | oligonucleotide/oligosaccharide binding |
| p | protein |
| P14 ^{ARF} | protein 14 alternate reading frame |
| P16 ^{INK4a} | protein 16 cyclin-dependent kinase inhibitor 4A |
| PAGE | polyacrylamide gel electrophoresis |
| PCR | polymerase chain reaction |
| Pd(PPh ₃) ₄ | tetrakis triphenyl phosphine palladium (0) |
| PDB | protein data bank |
| PHE | phenylalanine |
| PNA | peptide nucleic acid |
| POT1 | protection of telomeres 1 |
| ppm | parts per million |
| pRb | retinoblastoma protein |
| QSAR | quantitative structure activity relationship |
| RAP1 | repressor activator protein 1 |
| RM | reaction mixture |
| RMSD | root mean square deviation |
| RNA | ribonucleic acid |
| RPA | replication protein A |
| RPM | rotations per minute |
| RT | room temperature |
| s | second |
| SAR | structure activity relationship |
| SBDD | structure based drug design |
| SER | serine |

| | |
|-------------------|---|
| SOCl ₂ | thionyl chloride |
| Sp1 | specificity protein 1 |
| SPR | surface plasmon resonance |
| SV40 | simian virus 40 |
| T | thymine / thymidine |
| t | time |
| t _R | retention time |
| TEBP | telomeric end binding protein |
| TFA | trifluoroacetic acid |
| TGF α/β | transforming growth factor α/β |
| THR | threonine |
| TIF | telomere dysfunction induced foci |
| TIN2 | TRF interacting nuclear protein 2 |
| TLC | thin layer chromatography |
| T-loop | terminal loop |
| TRAP | telomeric repeat amplification protocol |
| TRF1/2 | telomeric repeat factor 1/2 |
| Tris | tris(hydroxymethyl)aminoethane |
| TYR | tyrosine |
| USD | united states dollar |
| vdW | van der Waals |
| VEGF | vascular endothelial growth factor |
| vHTS | virtual high throughput screening |
| w | watt(s) |

Chapter 1: Cancer and Telomeres

1.1 Introduction

Cancer is a term used to describe a group of diseases characterised by uncontrolled cellular growth. There are currently more than 100 distinct types and subtypes of cancer (Hanahan and Weinberg, 2000). Based on GLOBOCAN 2008 estimates, around 12.7 million new cancer cases were diagnosed and 7.6 million deaths occurred worldwide in 2008 (Ferlay *et al*, 2010). The World Health Organization estimates that annual global cancer death rate could rise to 15 million by 2020 (Rastogi *et al*, 2004).

In humans, cancer development is a multistep process in which several genetic and epigenetic alterations drive the progressive transformation of normal cells into highly malignant cancerous cells (Hanahan and Weinberg, 2000; You and Jones, 2012). These malignant cells contain defects in regulatory mechanisms which govern normal cell proliferation and homeostasis. In 2000, Hanahan and Weinberg proposed six common 'hallmarks' that nearly all cancers share which govern the transformation of normal cells to induce malignant growth:

- Self-sufficiency in growth signals
- Insensitivity to growth-inhibitory (antigrowth) signals
- Evasion of programmed cell death (apoptosis)
- Limitless replicative potential (immortal)
- Sustained angiogenesis
- Tissue invasion and metastasis

1.1.1 Self-sufficiency in growth signals

Cell growth and division is governed by various growth signals a cell receives from its surrounding cells. The most important of these signals come from hormones called growth factors which activate protein kinase receptors in the cell membranes. These receptors subsequently induce a signal transduction program leading to the transcription of proteins required for cell growth and division (Hanahan and Weinberg, 2000).

The majority of cancers suffer defects in this signalling process which allows them to divide continuously. Many cancer cells show a greatly reduced dependence on this exogenous growth stimulation and have the potential to generate their own growth factors and therefore reduce their dependence on stimulation from the local tissue microenvironment. For example, glioblastomas and sarcomas produce the platelet-derived growth factor (PDGF) and transforming growth factor α (TGF- α) for growth development. In some instances, growth factor receptors can be overexpressed in cancer cells allowing the cells to become sensitive to low levels of growth factors. For example, in breast, brain and stomach cancers, the epidermal growth factor receptor (EGFR) is upregulated. Moreover, the human epidermal growth factor receptor 2 (HER2) is overexpressed in stomach and mammary carcinomas. Some cancer cells can also produce abnormal receptors which are constantly switched on despite the lack of growth factors (Hanahan and Weinberg, 2000).

1.1.2 Insensitivity to growth-inhibitory signals

In normal cells, multiple anti-proliferative signals exist in the form of soluble growth inhibitors and immobilised inhibitors embedded on the surface of neighbouring cells

and in the extracellular matrix which maintain normal tissue homeostasis. In order to divide, cancer cells must overcome these antigrowth signals (Hanahan and Weinberg, 2000).

The circuitry which allows normal cells to respond to growth inhibitory signals is associated with cell cycle regulation. At the molecular level, most of these signals are dependent on the retinoblastoma protein (pRb) which is a critical gatekeeper of cell cycle progression along with its family members, p107 and p130. In its hypophosphorylated form, pRb halts cell proliferation and progression through the cell cycle. It does this by inhibiting the E2F family of transcription factors which control the expression of genes necessary for progression from G1 to S phase of the cell cycle. Evidence suggests that cancer cells suffer defects in the pRb pathway allowing E2F transcription factors to be released which activate factors called cyclins which move the cell through the cell cycle via the activation of cyclin-dependent kinases. This allows cells to continue proliferation and renders cells insensitive to antigrowth factors such as transforming growth factor β (TGF β) which normally operates along this pathway to prevent the phosphorylation which inactivates pRb (Hanahan and Weinberg, 2000).

1.1.3 Evasion of programmed cell death (apoptosis)

When normal cells become damaged, they undergo programmed cell death (apoptosis) via two distinct pathways: an extrinsic route which involves proteins called death activator proteins which bind to death receptors on cell membranes initiating a signalling process leading to apoptosis (Hanahan and Weinberg, 2000). T-lymphocytes generated by the immune system can also perforate the cell membrane

of damaged cells releasing an enzyme called granzyme which triggers apoptotic cell death (Elmore, 2007).

DNA damage resulting from exposure to chemicals, drugs or oxidative stress can initiate apoptosis via an alternative route: the intrinsic pathway involving the tumour suppressor protein p53 which upregulates the expression of Bax. Both the extrinsic and intrinsic signals which elicit apoptosis stimulate the mitochondria to release cytochrome C which in turn activates protease enzymes called caspases which execute the death program. Apoptosis is regulated by a family of proteins such as Bad and Bax which promote apoptosis and Bcl-2 and Bcl-X which suppress it. Cancer cells acquire resistance towards apoptosis via several strategies. In 50% of cancers, the gene coding for p53 is mutated hence the lack of apoptosis inducing p53 protein. The genes coding for apoptosis suppressors Bcl-2 and Bcl-X are also overexpressed in several tumour types (Hanahan and Weinberg, 2000).

1.1.4 Limitless replicative potential

Telomeres, repetitive DNA sequences at the terminal ends of chromosomes play an important role in the immortalisation of cancer cells. The function of telomeres is to protect and stabilize the ends of chromosomes and its length acts as a mitotic clock which dictates the number of times a cell can divide. In normal cells, the inability of DNA polymerase to completely replicate the 3' ends of chromosomes results in a 50-200 base pairs of telomeric DNA being lost from the telomere in each replication cycle (Levy *et al*, 1991). As the cells age, the telomeres continue to shorten and eventually reach a critical stage where they lose their function and a DNA damage

response is triggered resulting in senescence or apoptosis (Hanahan and Weinberg, 2000).

In cancer cells, telomere length is maintained above the critical threshold by the upregulation of telomerase, an enzyme with reverse transcriptase activity which rebuilds and maintains telomere length. This allows cancer cells to divide continuously without undergoing telomere shortening and therefore prevents cells from undergoing senescence and apoptosis. In most normal cells, the telomerase dependent pathway is strongly suppressed preventing normal cells from achieving limitless replicative potential (Hanahan and Weinberg, 2000).

1.1.5 Sustained angiogenesis

As tumours grow in size, they require continual supply of oxygen and nutrients for cell function and survival. They do this by forming new blood vessels via a process known as angiogenesis. Tumour cells release growth factors such as vascular endothelial growth factor (VEGF) and fibroblast growth factors (FGF-1/2) which interact with receptors on endothelial cells of nearby blood vessels and stimulate these cells to divide, leading to the branching and extension of existing capillaries (Hanahan and Weinberg, 2000).

1.1.6 Tissue invasion and metastasis

Not all tumours are malignant in nature. Benign tumours are growths which remain restricted to the place of its origin and do not spread to other areas of the body. These tumours are non-malignant and do not develop into a cancer. Malignant tumours are cancerous since they begin as a small growth and increase in size. The

most important feature of malignant tumours is that malignant cancer cells can grow and invade adjacent surrounding tissue and can break away from their origin and spread (metastasise) to distant sites to form secondary growths. This stage is critical since 90% of human cancer deaths occur due to metastasis (Hanahan and Weinberg, 2000).

Normal cells contain adhesion molecules which belong to the calcium-dependent cadherin family which ensure that cells of the same character adhere to each other. Adhesion to the extracellular matrix is also important and necessary for cell survival and usually involves molecules called integrins. When a cell becomes detached, it stops growing and apoptosis is triggered. Cell adhesion molecules in metastasised cancer cells are missing as a consequence of gene mutations thus allowing cells to break away from the primary tumour and spread (Hanahan and Weinberg, 2000).

1.1.7 Hallmarks in cancer therapy

Each of the hallmarks of cancer represents an alternative route to cancer therapy and most drugs which have been developed target specific proteins within key hallmark-associated signalling pathways. It is proposed that targeting multiple pathways with a combination of targeted agents may be more beneficial (Hanahan and Weinberg, 2011). Some prominent examples of marketed therapies which target the various hallmarks and drugs which are currently in the clinical phase are listed in Table 1.

| Hallmark | Drug and target description | Company |
|--|---|--|
| Self-sufficiency in growth signals | <ul style="list-style-type: none"> EGFR inhibitor: gefitinib (Iressa) HER2 monoclonal antibody: trastuzumab (Herceptin) | AstraZeneca /Teva Genetech |
| Insensitivity to antigrowth signals | <ul style="list-style-type: none"> Cyclin-dependent kinase inhibitor: flavopiridol (Alvocidib) / Phase II | Sanofi Aventis |
| Evasion of apoptosis | <ul style="list-style-type: none"> Small molecule drug which activates p53: CX-5461 / Phase I China has approved the first gene therapy approach of introducing TP53 gene into humans for the treatment of head and neck cancers (Peng, 2005). Drug which stimulates apoptosis in cancer cells by inhibiting multiple Bcl-2 family proteins: AT-101 / Phase II | Cylene Pharmaceuticals Ascenta Therapeutics |
| Limitless replicative potential | <ul style="list-style-type: none"> Telomerase inhibitor: Imetelsat (GRN163L) / Phase II | Geron |
| Sustained angiogenesis | <ul style="list-style-type: none"> Anti-VEGF monoclonal antibody: bevacizumab (Avastin) VEGF inhibitor: axitinib (Inlyta) | Genetech/Roche Pfizer |
| Tissue invasion and metastasis | <ul style="list-style-type: none"> Inhibitor of c-MET which plays multiple roles in tissue invasion and metastasis: tivantinib (ARQ 197) / Phase III | Arqule |

Table 1: Therapeutic targeting of cancer hallmarks.

1.2 Telomeres

Telomeres have attracted a great deal of interest in recent years as potential anti-cancer drug targets since they play an important role in the immortality of cancer cells. As mentioned previously, telomeres are specialised DNA structures which cap the ends of linear eukaryotic chromosomes, first discovered in the 1930s by two independent geneticists Hermann Müller and Barbara McClintock (Figure 1). The name 'telomere' was derived from the Greek words 'telos' meaning end and 'meros' meaning part (Müller, 1938; McClintock, 1941).

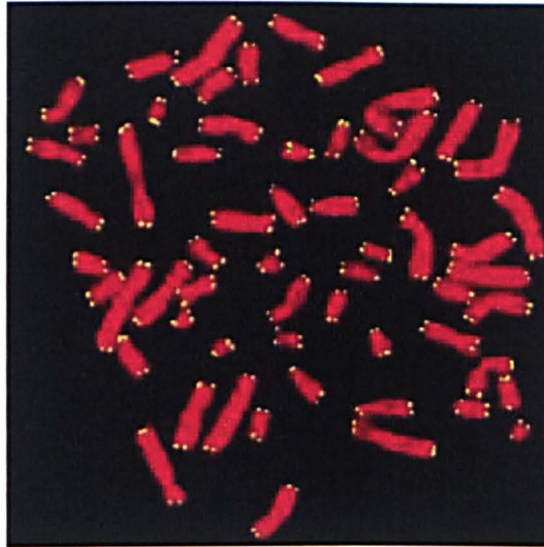


Figure 1: Human Chromosomes (red) capped by telomeres (yellow). Image provided by the courtesy of Professor Jerry Shay, UT Southwestern Medical Center.

Human telomeres are composed of 15-25 kilobase pairs of short repetitive guanine rich sequences (Moyzis *et al*, 1988) containing the hexanucleotide repeat TTAGGG which are highly conserved in vertebrates (Meyne *et al*, 1989). Alternative sequences have been found in other species such as TTGGGG in *Tetrahymena thermophila* (Blackburn and Gall, 1978) and TTTTGGGG in *Oxytricha nova* (Klobutcher *et al*, 1981).

Whilst most of the telomeric DNA is double stranded, the G-rich strand runs in a 5' to 3' direction towards the chromosome end and terminates into a short 100-200 nucleotide 3' single-stranded overhang (Makarov *et al*, 1997). The precise mechanism of how G-overhangs are generated remains unclear but it is thought that the combined action of nucleases which resect the C-strand, the inability of the lagging strand synthesis machinery to position the final RNA primers at the very end of the chromosome and elongation of the G-strand by telomerase all contribute to G-overhang formation (Dai *et al*, 2010).

Telomeres are critical to chromosome integrity and allow the ends of linear DNA to be replicated completely. Telomeres also have a multifunctional role to distinguish natural chromosome ends from being recognised as double strand DNA breaks and protect chromosome ends from all aspects of the DNA damage response (DDR). It is well known that double strand DNA breaks at chromosome ends are catastrophic and if not repaired, the chromosome ends can undergo degradation resulting in the loss of genetic information and cell death (Campisi *et al*, 2001).

Inappropriate repair of double-strand DNA breaks can also lead to chromosomal end-to-end fusions resulting in dicentric as well as multicentric chromosomes and ultimately genomic instability. Therefore without the protective telomere caps, chromosome ends are at risk from degradation and recombination by the DNA repair machinery (Counter *et al*, 1992).

Telomere shortening acts as a tumour suppressor mechanism and limits the proliferation of cells by inducing replicative senescence or apoptosis. However recent studies also point to the critical role of telomeres in cellular ageing. Growing evidence

suggests that telomere shortening limits stem cell function, tissue regeneration, and organ maintenance and therefore may promote cellular aging. Since telomere length acts as a mitotic clock which limits the number of cell divisions, it is suggested that telomere length may also prove to be an important biomarker of ageing (Jiang *et al*, 2007).

1.2.1 Shelterin complex

Human TTAGGG repeats are anchor sites for an array of telomere binding proteins which forms a protective complex at the ends of chromosomes known as shelterin (Figure 2). The function of the shelterin complex is to provide protection to the chromosome ends by capping the ends and preventing them from eliciting a DDR in addition to regulating telomere maintenance by telomerase (Palm and de Lange, 2008).

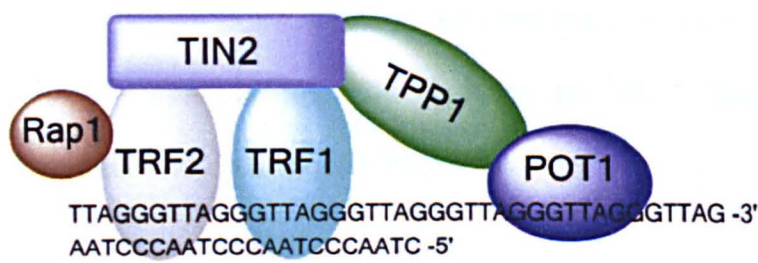


Figure 2: Mammalian telomeres are composed of TTAGGG repeats and protected by the shelterin complex which consists of six proteins: TRF1, TRF2, RAP1, TIN2, TPP1 and POT1 (Adapted from Palm and de Lange, 2008).

In its most abundant form, the complex consists of six proteins, three of which bind directly to telomeric DNA including the double-strand DNA binding proteins TRF1 and TRF2 and the single-strand DNA binding protein POT1. The TRF1 and TRF2

proteins subsequently recruit three bridging proteins: RAP1, TIN2 and TPP1 to the telomeres via protein-protein interactions. It is suggested that the shelterin proteins function only at the telomeres and are present at telomeres throughout the cell cycle (Palm and de Lange, 2008).

1.2.1.1 TRF1 and TRF2

TRF1 and TRF2 (Telomeric Repeat binding Factor 1 and 2) bind to double stranded telomeric DNA through their C-terminal SANT/Myb domains which are highly specific for the sequence TAGGGTT (Court *et al*, 2005). Both proteins bind to DNA as homodimers via homotypic interactions in their TRF homology (TRFH) domain. However both proteins do not interact directly with each other (Broccoli *et al*, 1997). The TRFH domain of both proteins recruits several other proteins to the telomeres (Chen *et al*, 2008). Both proteins adopt a multimeric binding mode and can act as architectural proteins which can modify the overall conformation of the telomeric tract (Bianchi *et al*, 1999). The overexpression of TRF1 results in the gradual shortening of telomeres whereas a dominant-negative mutant of TRF1 leads to telomere elongation (van Steensel and de Lange, 1997).

Tankyrase 1, a poly(ADP-ribose) polymerase is known to regulate TRF1 function at telomeres. TRF1 ADP-ribosylation by Tankyrase 1 results in the inhibition of TRF1-DNA binding (Smith *et al*, 1998). Overexpression of Tankyrase 1 results in the loss of TRF1 from telomeres (Smith and de Lange, 2000) and subsequently TRF1 degradation and telomere elongation (Chang *et al*, 2003). The overexpression of TRF2 also results in telomere shortening hence both proteins negatively regulate telomere length (Smogorzewska *et al*, 2000).

Recent electron microscopy studies suggest that TRF2 has the ability to form t-loop structures *in vitro* in which the 3' single stranded overhang loops back into the duplex part of the telomeric repeat array to form a lariat- or lasso-like structure referred to as the terminal loop (t-loop) (Figure 3).

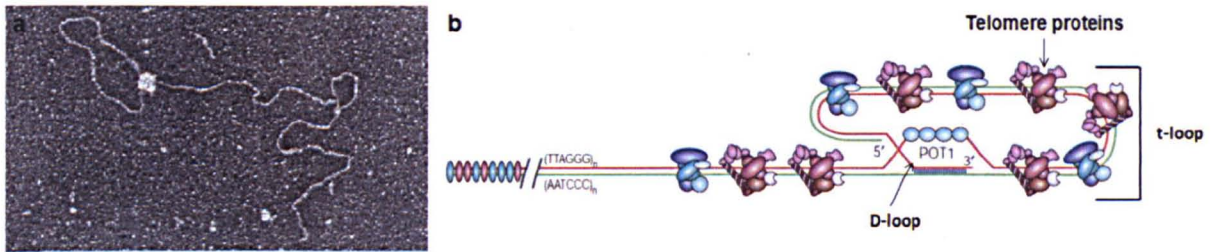


Figure 3: (a) Electron micrograph of a t-loop generated *in vitro* by the human TRF2 protein (Taken from Griffith *et al*, 1999). (b) Diagrammatic structure of the t-loop complexed to telomere binding proteins in which the overhang invades the double stranded repeat array forming the D-loop which is bound by POT1 (Taken from de Lange, 2004).

A displacement loop (D-loop) is also formed (Figure 3b) in which the single stranded G-rich overhang invades a homologous double stranded region and base pairs with the C-rich strand protecting the terminus and creating a structure which is distinct from a broken DNA end. In association with TRF2, the shelterin protein POT1 also participates in the protection of telomeres by binding to the D-loop of the t-loop configuration. T-loop size is dependent on telomere length; long telomeres tend to form larger t-loops than short ones. *In vitro* t-loops have been observed in humans, mice (Griffith *et al*, 1999) *Oxytricha fallax* (Murti and Prescott, 1999) and *Trypanosoma brucei* (Jordan *et al*, 2001).

It is not yet clear whether all telomeres exist in the t-loop configuration throughout the cell cycle. The exact role of t-loops *in vivo* is unknown but it may possibly play an

important role in telomere protection by hiding the G-overhang from inappropriate DNA damage response pathways. This could potentially be the mechanism by which the cell caps and protects natural chromosome ends from being recognised as DNA breaks (Griffith *et al*, 1999).

1.2.1.2 RAP1

RAP1 (Repressor/Activator Protein 1) is a poorly characterised component of the shelterin complex and its role in telomere biology has remained largely uncharacterised to date. Human RAP1 contains 3 domains: a Myb domain, a BRCT domain and a C-terminal region which interacts with TRF2 (Palm and de Lange, 2008). RAP1 does not directly bind to DNA and depends on TRF2 for telomere localisation and stability (Li *et al*, 2000). It binds TRF2 as a 1:1 complex (Zhu *et al*, 2000) and most of it is depleted upon TRF2 deletion (Celli and de Lange, 2005). Inhibition of RAP1 or using a dominant negative mutant of RAP1 in human cells leads to telomere elongation and loss of telomere heterogeneity (O'Connor *et al*, 2003; Li and de Lange, 2003).

1.2.1.3 TIN2

TIN2 (TRF1 and TRF2 Interacting Nuclear Factor 2) binds to the double stranded telomere binding proteins, TRF1 and TRF2 (Ye *et al*, 2004). TIN2 is a negative regulator of telomere length (Kim *et al*, 1999) and connects the DNA binding proteins within the shelterin complex (Palm and de Lange, 2008). It stabilizes TRF1 at telomeres and protects it from tankyrase driven degradation (Ye and de Lange, 2004). TIN2 mutants which can no longer bind to TRF1 and TRF2 have a

destabilising effect on shelterin (Kim *et al*, 2004). The protein also recruits TPP1 and therefore POT1 to telomeres. Little is known whether TIN2 binds to all three proteins at once or switches between TRF1, TRF2 or TPP1 (Palm and de Lange, 2008).

1.2.1.4 TPP1

TPP1 (also known as TINT1, PTOP or PIP1) interacts with TIN2 and recruits POT1 to the six-member complex thereby providing a link which connects single-stranded DNA binding proteins to double-stranded DNA binding proteins (Hockemeyer *et al*, 2007). TPP1 has also been shown to enhance POT1-DNA binding activity suggesting that POT1 binds to DNA as a heterodimer with TPP1. The heterodimer appears to enhance the activity and processivity of telomerase at chromosome ends (Wang *et al*, 2007).

It is suggested that the oligonucleotide/oligosaccharide binding (OB) domain of TPP1 interacts with telomerase and TPP1 is involved in telomerase recruitment and regulation at telomeres (Abreu *et al*, 2010). Depletion of TPP1 results in the loss of POT1 from telomeres (Liu *et al*, 2004) and knockdown of TPP1 reduces the amount of POT1 in the nucleus (Chen *et al*, 2007). Furthermore, low levels of TPP1 leads to a telomere elongation phenotype that is consistent with POT1 depletion (Ye *et al*, 2004). Surprisingly, a number of studies suggest that mutant POT1 which can no longer bind to TPP1 can still localise to telomeres but the exact mechanism of how this recruitment occurs is still unclear (Colgin *et al*, 2003; He *et al*, 2006).

1.2.1.5 POT1

POT1 (Protection Of Telomeres 1) is a single-stranded telomere binding protein found in a variety of organisms ranging from the fission yeast *Schizosaccharomyces pombe* (Baumann and Cech, 2001), plants (Shakirov *et al*, 2009), animals (Palm *et al*, 2009) and humans (Baumann *et al*, 2002). Homologues of POT1 have also been found in hypotrichous ciliates as telomeric end binding protein (TEBP) (Horvath *et al*, 1998) and in the budding yeast *Saccharomyces cerevisiae* as Cdc13 (Nugent *et al*, 1996).

Human POT1 was initially identified by its sequence similarity to the α -subunit of the telomeric end binding protein α/β complex in *Oxytricha nova* (Horvath *et al*, 1998). POT1 binds to telomeres with high sequence specificity for the G-rich single stranded DNA overhang. However POT1 does not bind to double stranded telomeric DNA or to the complementary C-rich strand (Baumann *et al*, 2002).

Human POT1 has two important domains: an N-terminal OB region which is necessary for DNA binding and a C-terminal region which interacts with TPP1. Surprisingly, a truncated POT1 mutant lacking the OB domain can still bind to telomeres via TRF1 (Loayza and de Lange, 2003).

In comparison to mice which carry two POT1 proteins, humans contain a single POT1 gene which encodes for five splice variants including the full-length POT1, NH₂-terminally truncated POT1 and COOH-terminally truncated POT1 whose functions are poorly understood (Yang *et al*, 2007). In human cells, full length POT1 is involved in telomere length regulation and plays dual roles as positive or negative regulator of telomerase activity. The protein inhibits telomerase activity at the 3' end

of DNA by controlling the accessibility of the telomeric single-stranded DNA substrate to telomerase. The exact role of the DNA binding activity of POT1 is unknown, however the crystal structure of human POT1 suggests that POT1 physically caps the end of chromosomes and buries the last guanine base of DNA into a hydrophobic pocket thus rendering it inaccessible to telomerase. The physical capping of chromosome ends by POT1 may also prevent the ends from eliciting a DNA damage response (Lei *et al*, 2004).

The expression of human POT1 mutant which lacks the DNA binding domain results in the extension of telomeres in telomerase positive cells suggesting that the C-terminal region of POT1 is not essential for DNA binding (Kelleher *et al*, 2005). POT1 also acts as a positive regulator of telomere length. In some settings, overexpression of full length POT1 in telomerase positive cells leads to the lengthening of telomeres (Colgin *et al*, 2003). The G-rich overhang can be extended by telomerase if POT1 is bound to a site more internal within the telomere leaving at least an eight nucleotide free overhang. An overhang with less than eight nucleotides cannot be extended due to steric hindrance between POT1 and telomerase (Lei *et al*, 2005). It is possible therefore that POT1 may have dual roles at telomeres; one is to promote telomere elongation and the other being inhibitory to telomerase (Loayza *et al*, 2004).

1.2.1.6 Shelterin accessory factors

In addition to the six shelterin components, mammalian telomeres bind to additional proteins which are thought to be important for chromosome maintenance and protection. Most of these proteins have non-telomeric functions and bind only transiently and non-exclusively to telomeres and are known as shelterin accessory

factors. The majority of these factors have prominent roles in general genome maintenance and found less abundant at telomeres compared to the shelterin components (Palm and de Lange, 2008). Examples include the DNA damage sensor, MRN, a complex consisting of Mre11, Rad50 and Nbs1 (Zhu *et al*, 2000), DNA repair proteins involved in non-homologous end joining Ku70/80 (Hsu *et al*, 2000), the DNA excision repair proteins ERCC1/XPF (Zhu *et al*, 2003), an exonuclease Apollo (Lenain *et al*, 2006), a protein involved in homologous recombination, RAD51D (Tarsounas *et al*, 2004), WRN and BLM RecQ helicases (Opresko *et al*, 2002) and tankyrases (Smith *et al*, 1998).

It is surprising that one of the major roles of telomeres is to prevent the activation of the DDR yet proteins involved in the DDR are found abundantly at telomeres. The exact function of some of these factors at telomeres is not yet clear but they presumably function to either sensitize the cell to damage to the telomeres or prevent telomeres from being recognised as DNA breaks (Palm and de Lange, 2008). The majority of these proteins associate with telomeres by interacting with the shelterin proteins TRF1, TRF2 and RAP1 (Chen *et al*, 2008).

1.3 End replication problem

In 1971, the Russian biologist, Alexey Olovnikov first recognised that chromosome ends could not be fully replicated to the end of the base sequence (Olovnikov, 1971). In an independent study a year later, James Watson realised that chromosomes lose DNA from the ends in each round of cell division because DNA polymerase cannot replicate the linear end of the DNA molecule, he termed this as the 'end replication problem' (Watson, 1972).

According to the semi-conservative model of DNA replication (Figure 4), the two DNA strands in the double helix run in opposite orientations, one in the 5'→3' direction and the other in the 3'→5' direction which act as templates for the synthesis of new DNA strands. DNA polymerase requires short RNA primers to initiate replication and can catalyse the growth of the DNA chain towards the replication fork in a 5' → 3' direction by adding deoxyribonucleotides to the 3' hydroxyl end of the polynucleotide chain.

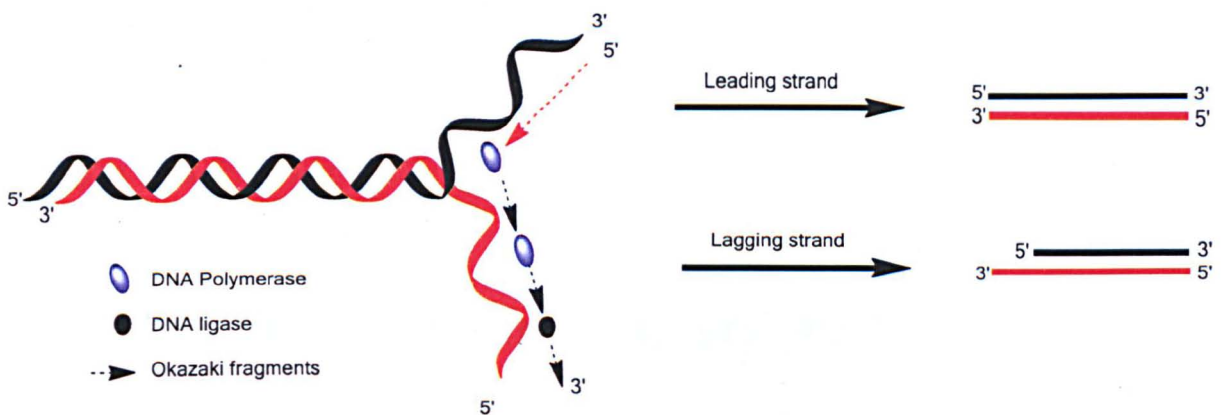


Figure 4: A diagram to illustrate the end-replication problem. The leading strand is replicated completely whereas the lagging strand is incomplete.

Replication differs between the two strands. The leading strand is usually replicated continuously and completely to the end of the telomere in the direction of the fork movement to generate a blunt-ended DNA. However, the lagging strand is replicated discontinuously in steps from several RNA primers which are elongated by DNA polymerase to form short chains of DNA called Okazaki fragments. As DNA synthesis continues, the Okazaki fragments are processed whereby the RNA primers are removed and replaced by DNA sequences and the Okazaki fragments are subsequently ligated by DNA ligase to form a continuous DNA strand.

When the last RNA primer attaches on the lagging strand, it cannot be replaced by DNA sequences since it requires a DNA sequence in front of the primer. The RNA primer is subsequently destroyed and as a result the lagging strand is not fully synthesised to the end. Consequently the new DNA molecule is shorter than the parent by at least the length of one RNA primer or more depending on where that last primer was placed with respect to the end of the chromosome. This results in telomere shortening and the loss of 50-200 base pairs of telomeric DNA at the chromosome end in each round of cell division (Levy *et al*, 1991). The progressive shortening of telomeres through successive cycles of replication eventually limits their ability to protect the chromosome ends and a DDR is triggered (d'Adda di Fagagna *et al*, 2003).

1.4 Telomere shortening triggers DDR

In 1961, Leonard Hayflick noticed that normal human diploid fibroblasts are not immortal and stopped dividing after undergoing a limited number of cell divisions in culture. Such cells cease to divide after they reach the 'Hayflick limit', usually after 30-50 doublings and undergo an irreversible arrest of cell division, a state referred to as replicative senescence, (Hayflick and Moorhead, 1961).

In 1990, Harley and colleagues made the initial connection between telomere shortening and replicative senescence and showed that the telomeres of human fibroblasts shorten as they reach the 'Hayflick limit' (Harley *et al*, 1990). Telomere length was shown to be like a 'mitotic clock' which limited the number of cell divisions a cell can undergo by acting as a molecular counting mechanism against unlimited proliferation. It is now known that when telomeres reach a critical length, usually 4

kilo bases, a DNA damage response is triggered via the activation of p53 and p16/pRb pathways (Levy *et al*, 1991). The cells subsequently enter a stage referred to as replicative senescence also known as Mortality stage 1 (M1) (Harley *et al*, 1990; Hayflick, 1965).

Replicative senescence acts as a potent barrier to limit cellular proliferation in normal cells (Wright and Shay, 1992). Senescent cells remain metabolically active and arrest in the G1 phase of the cell cycle. Senescent cells also have altered levels of gene expression and show distinctive characteristics in cell morphology. The cells are large and flat in shape and have increased granularity and β -galactosidase activity (Roninson, 2003; Shelton *et al*, 1999). Senescence can also be activated in response to various types of stress including DNA damage, ionizing radiation and oxidative stress, where it is called accelerated or stress induced senescence (Lleonart *et al*, 2009).

1.5 DNA damage response (DDR)

Telomeres can also trigger a DDR in the absence of telomere shortening when proteins of the shelterin complex are perturbed. Loss of POT1 results in cell death, cell cycle arrest, telomere fusions and G-overhang extension (Churikov *et al*, 2006). TPP1 knockdown results in a p53 dependent growth arrest (Guo *et al*, 2007). Loss of TIN2 causes telomere uncapping, telomere dysfunction, growth arrest and cell death (Kim *et al*, 2008). TRF1 deletion results in rapid cell senescence (Martinez *et al*, 2009). Loss of TRF2 leads to a strong DDR response, degradation of the G-rich overhang and chromosomal end-to-end fusions. This data shows how shelterin is able to suppress different aspects of the DDR pathway (Celli and de Lange, 2005).

The DDR triggered as a result of telomere shortening may be caused due to the loss of shelterin bound proteins from telomeric DNA. It is suggested that when telomeres reach a critical length, not enough shelterin is recruited to protect chromosome ends hence telomeres become dysfunctional and resemble double-strand DNA breaks. Such DNA lesions activate the canonical DDR pathway which engages p53 to induce senescence or apoptosis depending on cell type (Reviewed in Shiloh, 2003; Palm and de Lange, 2008).

The phosphatidylinositol-3-kinase related protein kinases, ATM (ataxia-telangiectasia mutated) and ATR (ataxia-telangiectasia and Rad3 related) play an important role in response to DNA lesions (Shay and Wright, 2004). ATM becomes activated as a result of double-strand DNA breaks and ATR becomes activated due to lesions in single stranded DNA (Palm and de Lange, 2008).

The initial step in ATM activation involves recognition of double-strand DNA breaks by the DNA damage sensor MRN, a trimeric complex consisting of (Mre11, Rad50, Nbs1) which recognises double-strand DNA breaks as sites of DNA damage and recruits ATM to the site. The ATM kinase subsequently phosphorylates the histone variant H2AX on serine 139 to produce γ -H2AX which promotes the localisation of several DNA damage factors (53BP1, MDC1 and Mre11) to the damaged site (Palm and de Lange, 2008). This results in the accumulation of DNA damage foci which span several kilobases from the site of lesion (Takai *et al*, 2003). The exact function of these DNA foci is poorly understood but they may have important roles in signal amplification and DNA repair (Palm and de Lange, 2008).

Lesions in single-stranded DNA become recognised as sites of DNA damage by replication protein A (RPA) which recruits ATR and its partner protein ATRIP to the damaged site. Once activated, both ATM and ATR phosphorylate downstream kinases, Chk1 and Chk2 which induce G1/S and G2/M arrest and halt cell cycle progression. Chk1 and Chk2 do this by preventing the activation of cyclin-dependent kinases (cdk's) through the inhibition of cdc25 phosphatases (Palm and de Lange, 2008).

ATM and ATR can also activate the tumour suppressor protein p53 which causes cell cycle arrest by upregulating the cdk inhibitor, p21. After the establishment of growth arrest, the levels of p53 and p21 decrease whereby the cdk inhibitor p16^{Ink4A} becomes upregulated to maintain growth arrest (Shay and Roninson, 2004; Roninson, 2003).

It is suggested that cell proliferation is incompatible with ATM and ATR activation. Telomeres ensure that the ATM and ATR signalling pathways remain dormant to enable cell proliferation and survival (Palm and de Lange, 2008). The ATM and ATR pathways are suppressed in an independent and specific manner at telomeres (Denchi and de Lange, 2007).

The inhibitions of individual shelterin proteins have shown how telomeres prevent ATM and ATR activation. A dominant-negative mutant of TRF2 results in ATM activation, degradation of the 3'-overhang, chromosomal fusions and p53 dependent apoptosis and senescence (van Steensel *et al*, 1998; Karlseder *et al*, 1999). The ATM-dependent DNA damage foci are also found at telomeres and are called telomere dysfunction-induced foci (TIFs) which contain the same factors which are induced by double-strand DNA breaks (Celli and de Lange, 2005).

The mechanism used by TRF2 to inhibit ATM activation is not clear. However two models have been proposed: one in which TRF2 through its ability to interact directly with ATM might inhibit its action at telomeres (Karlseder *et al*, 2004). Secondly, TRF2 has the ability to form t-loops which can sequester the telomere ends from being recognised as sites of DNA damage (Griffith *et al*, 1999).

The ATR pathway on the other hand is shown to be suppressed by POT1. Inhibition of POT1 triggers ATR activation via the induction of Chk1 and Chk2 and in the accumulation of TIFs at telomeres which contain the DNA damage factors γ -H2AX, MDC1 and 53BP1 (Denchi and de Lange, 2007). Knockdown of POT1 leads to telomere elongation, growth arrest and cell death (Veldman *et al*, 2004).

It is suggested that POT1 suppresses the ATR pathway by inhibiting the binding of replication protein A (RPA) to the G-rich overhang. Although RPA is a more abundant protein, the interaction of POT1 with TPP1 could enhance the ability of POT1 to act as an effective competitor. The binding of RPA upon POT1 delocalisation from telomeres results in the recruitment of ATR to the telomere by RPA (Takai *et al*, 2011). Figure 5 summarises how ATM and ATR induce senescence.

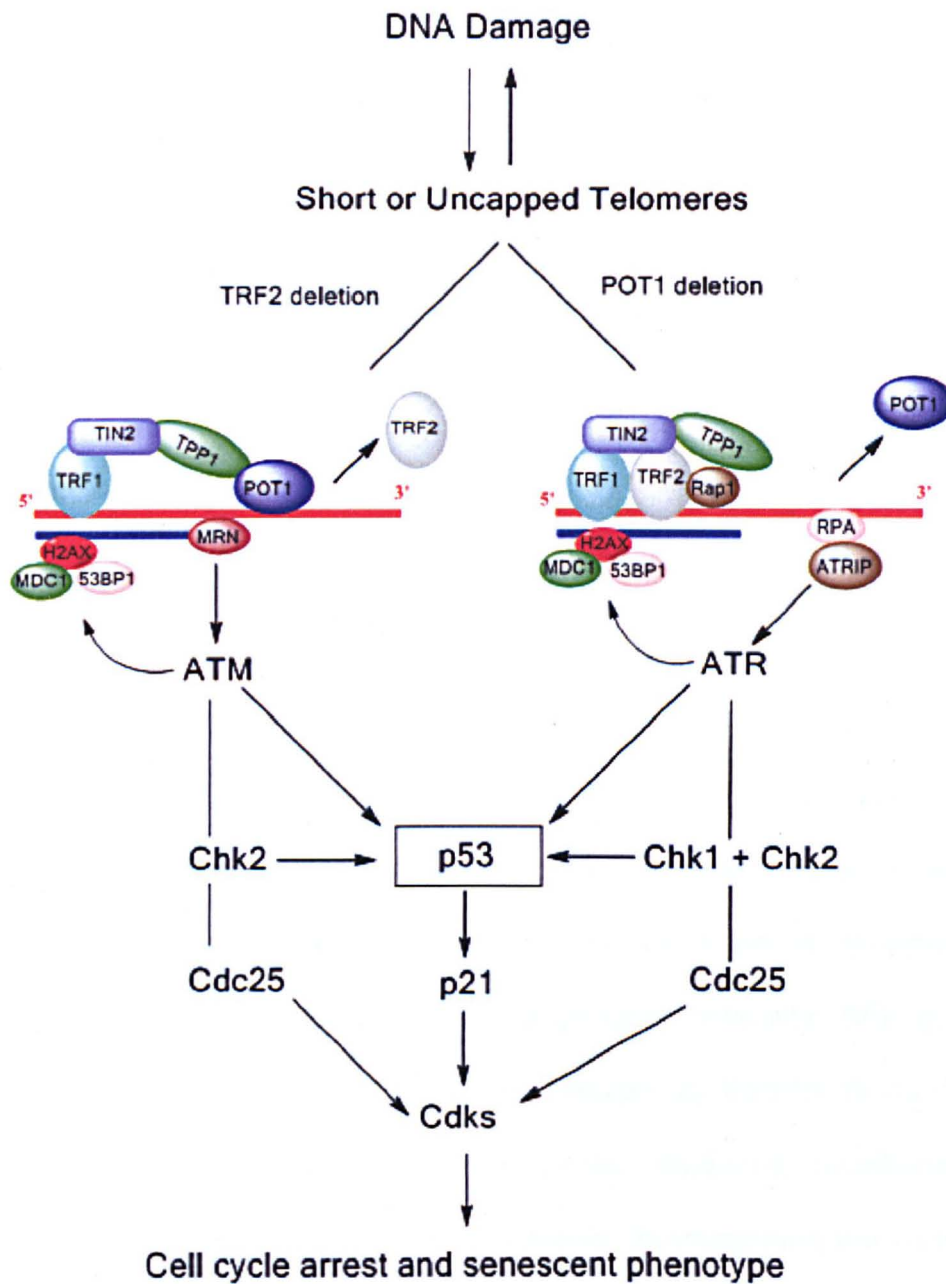


Figure 5: Senescence program in normal cells (Adapted from Palm and de Lange, 2008).

1.6 Dysfunctional telomeres and cancer

Human cells where pathways responsible for growth arrest have been disabled, for instance the inactivation of p53 and pRb via the introduction of viral oncogenes (Ozer

et al, 1996; Dyson *et al*, 1989; Scheffner *et al*, 1990) or anti-sense oligonucleotides allow cells to bypass the senescence checkpoint. Cells continue to divide and their telomeres continue to erode further until they reach a second proliferative block known as crisis or mortality stage 2 (M2), a state where cell division is balanced by cell death (Hara *et al*, 1991).

At crisis, short telomeres lead to the deprotection of chromosome ends. The ends subsequently become substrates for repair activities, resulting in chromosomal end-to-end fusions mediated by non-homologous end joining (NHEJ) and homologous recombination (HR) pathways. The fusion between chromosome ends results in a bridge. The dicentric chromosomes are then broken during anaphase and the uncapped extremity will then fuse with another uncapped extremity, initiating breakage-fusion-bridge (BFB) cycles. At the beginning only a few short telomeres will initiate BFB cycles, eventually the number of deprotected chromosome ends available for fusion will increase leading to genome instability. BFB cycles induce gross chromosome rearrangements via non-reciprocal translocations. BFB cycles also cause chromosome arm gains and losses, deletions, mutations and gene amplifications. At the end the number of unstable chromosomes will be too high, the cells will enter mitotic catastrophe and die (Vallejo, 2008).

A cell will not survive crisis unless it develops a mechanism to maintain its telomere length, a rare cell (1 in 10^7) will emerge from crisis and become immortalised by exhibiting some form of telomere stabilization usually by reactivating telomerase (Wright *et al*, 1989; Counter *et al*, 1992), an enzyme which replenishes the telomeres or by engaging the alternative lengthening of telomeres (ALT) pathway (Dunham *et al*, 2000). Telomere stabilization is therefore essential for the continued cellular

proliferation of cancer cells. The relationship between telomere shortening, senescence and cancer is depicted in Figure 6.

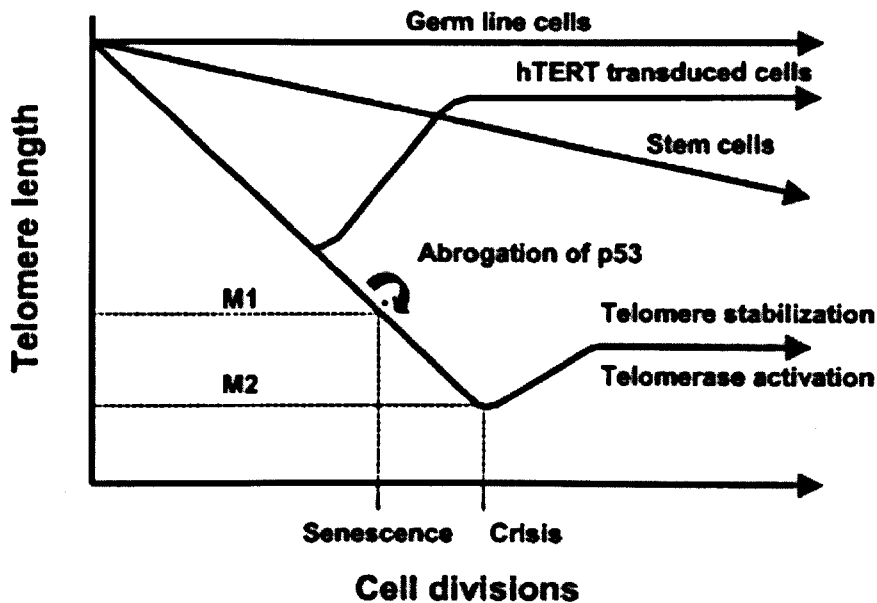


Figure 6: A diagram showing the link between cellular senescence and immortalisation. Most human somatic cells are telomerase-negative and therefore with each cell division they experience telomere shortening. Germ cell telomere lengths are maintained by telomerase. Telomere length in stem cells is not maintained and therefore stem cell telomeres also shorten but at slower rates compared to telomerase-negative cells. Cells with critically short telomeres enter senescence at the Hayflick limit, or M1. The inactivation of p53 or pRb/p16 allows cells to bypass the M1 proliferative checkpoint. The telomeres of such cells will continue to shorten and the cells will ultimately enter crisis or M2 which is characterised by widespread cell death. Rare cells emerge from crisis and acquire unlimited proliferative potential through stabilisation of telomere length almost universally via telomerase activation (Taken from Cong et al, 2002).

1.7 Telomerase

In 1985, Elizabeth Blackburn and Carol Greider working on the ciliate *Tetrahymena thermophila* found out that a terminal transferase was adding tandem TTGGGG repeats to the ends of the synthetic telomere primers. It is now known that this terminal transferase is a reverse transcriptase enzyme known as 'Telomerase' which compensates for telomere loss due to the end replication problem (Greider and Blackburn, 1985).

In the absence of telomerase activity, normal human cells in culture have a finite lifespan and undergo senescence as a result of telomere shortening. Telomerase activity is therefore necessary for normal tissue renewal and also for the survival of cancer cells since it restores and maintains telomere length (Blackburn, 1992).

1.7.1 Telomerase activity in cells

The development of a PCR-based assay known as TRAP (Telomeric Repeat Amplification Protocol) has allowed the detection of telomerase in a large number of different cell types (Kim *et al*, 1994).

In humans, telomerase activity is not detectable in most normal somatic cells but germ line cells (Wright *et al*, 1996), hematopoietic cells, lymphocytes (Hiyama *et al*, 1995), stem cells (Hiyama and Hiyama, 2007), specific epithelial cells of the breast (Hiyama *et al*, 1996), cells of the endometrium (Kyo *et al*, 1997), hair follicles, epidermis of the skin (Harle-Bachor and Boukamp, 1996; Ramirez *et al*, 1997) and intestinal crypt cells (Hiyama *et al*, 1996) continue to express low levels of telomerase activity. In addition, the majority of tumour-derived cell lines possess

telomerase activity signifying that the maintenance of telomere length is necessary for the unlimited replicative potential of human cells (Kim *et al*, 1994). It is now known that telomerase maintains telomere length stability in 90% of human cancers and the enzyme's activity is not found in adjacent normal cells (Kim *et al*, 1994; Shay and Bacchetti, 1997). This makes telomerase a very attractive drug target for cancer therapy.

1.7.2 Structure and function of telomerase

Human telomerase is a 1000 kiloDalton ribonucleoprotein complex consisting of two major components: the functional human telomerase RNA component (hTERC), also known as the human telomerase RNA (hTR) which serves as a template for the *de novo* addition of TTAGGG repeats to the 3' end of chromosomes and a catalytic protein component with reverse transcriptase activity known as human telomerase reverse transcriptase (hTERT) which adds the telomeric repeats onto the end of chromosomes (Feng *et al*, 1995).

The hTR component is transcribed by RNA polymerase II and its 3' end is subsequently processed to generate a mature transcript containing 451 nucleotides. The 5' end of the molecule contains the template for reverse transcription which is 11-base pairs long with the sequence 3'-CAAUCCCAAUC-5' (Cong *et al*, 2002; Feng *et al*, 1995).

The majority of tissues express the hTR component regardless of telomerase activity but cancer cells have a five-fold higher expression compared to normal cells (Avillion *et al*, 1996; Yi *et al*, 1999). It is suggested that the primary determinant of telomerase activity is the hTERT subunit since it is upregulated in cancer cells but suppressed in

normal cells and its expression is closely correlated with telomerase activity (Counter *et al*, 1998).

There is compelling evidence to suggest that the active telomerase complex is a dimer which consists of two copies of each of hTERT, hTR and dyskerin (Cohen *et al*, 2007). It is suggested that the telomerase dimer may simultaneously extend the telomeres belonging to sister chromatids in parallel (Simonsson, 2003).

Recent findings demonstrate that telomerase expression in several cell types allow these cells to escape crisis and become immortal (Cong *et al*, 2002). The expression of telomerase in telomerase-negative epithelial cells and fibroblasts leads to the bypass of senescence and extension of life-span (Bodnar *et al*, 1998). Moreover, the expression of telomerase in conjunction with SV40 large T antigen and an oncogenic allele of H-ras have shown to induce the tumorigenic conversion of normal human cells (Hahn *et al*, 1999).

How telomerase becomes activated during immortalization is currently under intensive investigation. Several transcription factors have been identified which participate in hTERT gene expression. Wang *et al* (1998) have shown that c-Myc (a proto-oncogene product with transcriptional activity) induces hTERT expression and telomerase activity in normal human primary fibroblasts and mammary epithelial cells. Another transcription factor, sp1 can also co-operate with c-myc to activate telomerase (Kyo *et al*, 2000). Telomerase can also be activated by the human papillomavirus 16 E6 protein (Klingelhutz *et al*, 1996).

Telomere elongation by telomerase is a multi-step process which is precisely regulated. Cong *et al* (2002) suggests that several processes must occur before

telomerase is recruited to the telomeres including gene transcription, splicing of the mRNA, maturation and posttranslational modifications of hTR and hTERT, transport and subcellular localization of both components and assembly into an active enzyme. Several accessory proteins have been identified which interact with hTR and hTERT and may play important roles in the regulation of telomerase (Holt *et al*, 1999; Seimiya *et al*, 2000; Harrington *et al*, 1997).

When telomerase is recruited to the telomere end, it undergoes progressive cycles in which it elongates the telomere (Figure 7). It is suggested that in each cell cycle, telomerase may display an increasing preference for telomeres as their length declines and therefore the enzyme may not act on every telomere. This suggests that telomeres may switch between extendible and non-extendible states (Teixeira *et al*, 2004).

In humans, the shelterin complex negatively regulates telomerase by sequestering its DNA substrate into a closed/capped conformation. Since the amount of shelterin proteins bound to a telomere is proportional to the telomere length, longer telomeres are proposed to have a greater ability of inhibiting telomerase. The progressive shortening of telomeres causes the gradual loss of telomere bound shelterin thus allowing telomere elongation by telomerase (de Lange, 2005).

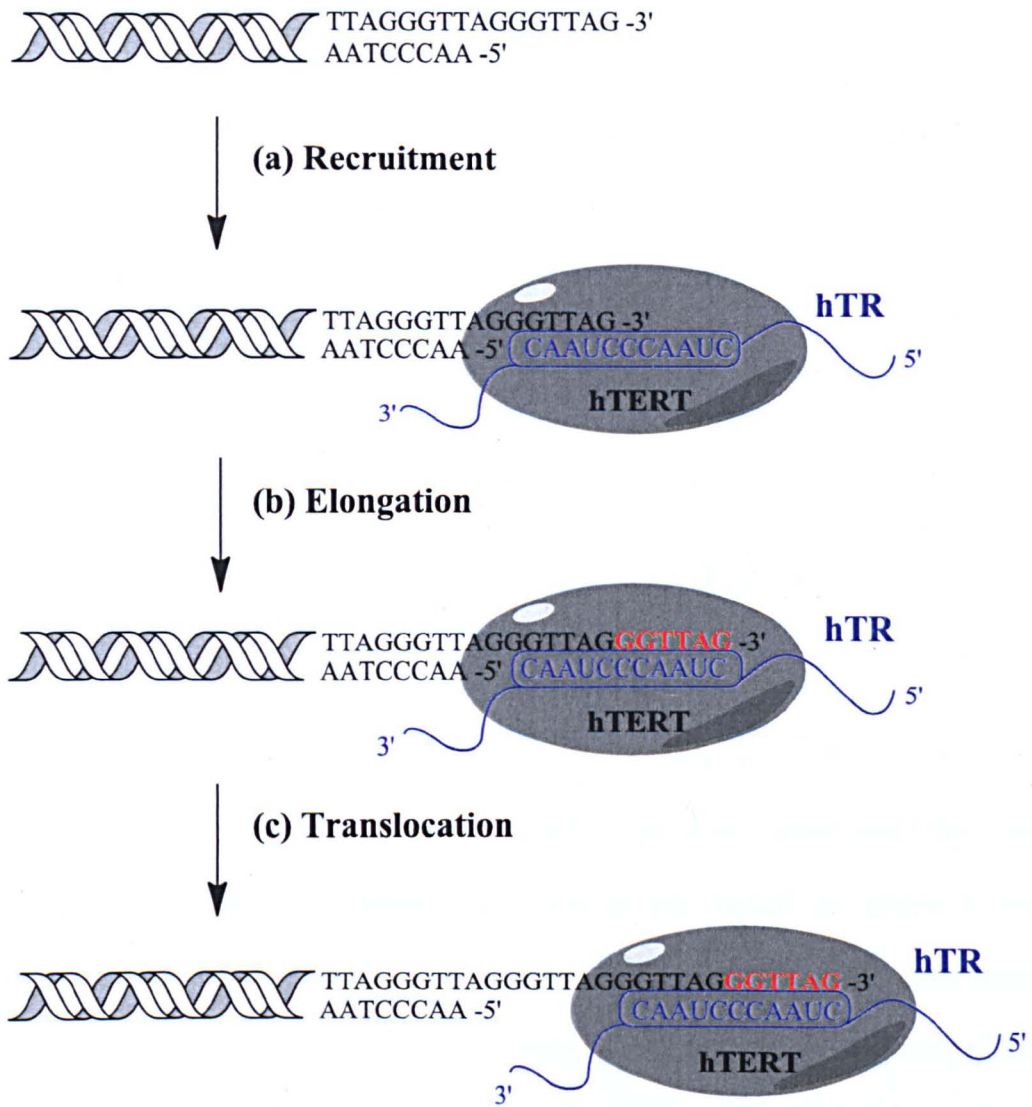


Figure 7: Schematic representation of how telomerase maintains telomeric DNA ends. (a) Telomere extension begins when telomerase is recruited to the telomere end and the template region of its RNA component partially anneals with the telomeric ssDNA overhang. (b) The region of telomerase RNA template which does not base pair with the telomeric DNA is subsequently reverse transcribed where a GGTAG repeat is added to the end. (c) The telomerase enzyme then translocates to the end where it adds another GGTAG repeat. The cycle is repeated several times before telomerase dissociates from the end. The conventional DNA replication machinery subsequently fills in the C-rich strand, thus preventing any loss of DNA (Adapted from Simmonson, 2003).

1.7.3 Telomerase Inhibition

The inhibition of telomerase by its dominant-negative mutant or by antisense oligonucleotides induced telomere shortening, senescence and apoptosis in telomerase-positive cancer cells (Zhang *et al*, 1999; Kondo *et al*, 1998). This has been taken as a proof of principle that induction of telomere shortening is a viable therapeutic strategy.

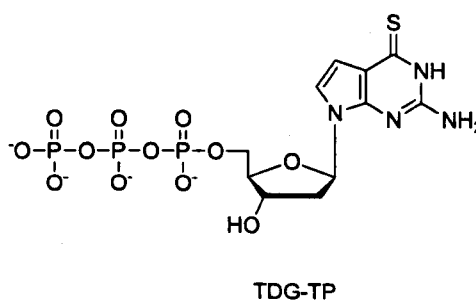
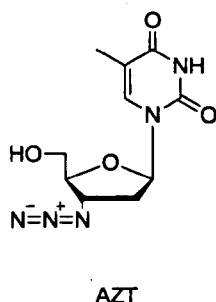
1.7.3.1 Reverse transcriptase inhibitors

Since telomerase is a reverse transcriptase, several nucleoside based analogues have been tested on telomerase (Strahl and Blackburn, 1994). Many of these nucleoside inhibitors cause telomere shortening in several cancer cell lines. Although nucleoside compounds were among the first to be tested as chain terminating inhibitors. Evidence suggests that such compounds exert their telomere shortening effects not by chain termination as would be expected but rather the compounds bind and compete for the nucleoside triphosphate binding site in telomerase (Strahl and Blackburn, 1996).

Azidothymidine (AZT) (Figure 8) also used in the treatment of HIV resulted in telomere shortening, increased p14^{ARF} expression and telomerase inhibition (Datta *et al*, 2006). AZT also triggered apoptosis and inhibited the cell growth of human parathyroid cancer cells in culture (Falchetti *et al*, 2005). Another potent nucleoside inhibitor of telomerase, 6-thio-7-deaza-2'-deoxyguanosine 5'-triphosphate (TDG-TP) inhibits telomerase activity with an IC₅₀ of 60nM (Figure 8) (Fletcher *et al*, 2001).

Several non-nucleoside drugs have also been developed and inhibit telomerase activity by targeting the hTERT component with some compounds achieving IC_{50} in the submicromolar range. Examples include bisindole derivatives (Sasaki *et al*, 2001), tea catechin epigallocatechin gallate (EGCG) (Figure 8) (Naasani *et al*, 1998) and its derivative MST-312 (Seimiya *et al*, 2002). A small synthetic inhibitor BIBR1532 (2-[(E)-3-naphtalen-2-yl-but-2-enoylamino]-benzoic acid) is also worth mentioning (Figure 8). It is a mixed type non-competitive and selective inhibitor of telomerase (Pascolo *et al*, 2002). BIBR1532 inhibits telomerase activity with an IC_{50} of 93nM. In several cancer cells, BIBR1532 induced telomere attrition with no acute cytotoxicity. The compound triggered senescence and cell cycle arrest after a characteristic lag period (Damm *et al*, 2001).

Nucleoside Inhibitors



Non-Nucleoside Inhibitors

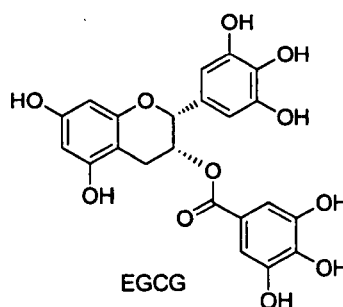
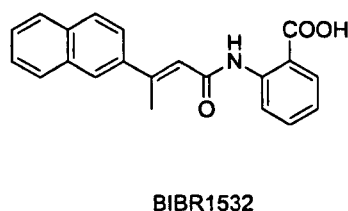


Figure 8: Chemical structures of nucleosides and non-nucleoside telomerase inhibitors.

1.7.3.2 Targeting telomerase RNA

The template region of hTR seems to be an attractive target for complementary antisense oligonucleotides. Typical antisense oligonucleotides have DNA bases which can form complexes with mRNA. RNase H then recognises these complexes, cleaves them and leads to mRNA destruction thus reducing protein expression (Baker and Monia, 1999). Such approaches have been used in the inhibition of telomerase (Mukai *et al*, 2000). However cleavage of hTR is not necessary and oligonucleotides can bind to hTR and act as competitive inhibitors (Correy, 2002).

Conventional chemically unmodified DNA oligomers are too unstable for *in vivo* applications since they are readily degraded by nucleases (Cian *et al*, 2008). Various modifications have been undertaken by several groups to enhance the stability, cellular uptake, potency and half-life of telomerase oligonucleotide inhibitors (Correy, 2002).

Peptide nucleic acids (PNAs) were among the first oligomers tested with some effectively inhibiting telomerase in the low nanomolar range (Norton *et al*, 1996). Since then several modified oligonucleotides have been tested as telomerase inhibitors and are presented in Figure 9 including locked nucleic acids (LNA) (Elayadi *et al*, 2002), 2'-O-methyl RNA (Pitts and Correy, 1998), 2'-O-methoxyethyl RNA (Chen *et al*, 2003), phosphorothioate (Matthes and Lehmann, 1999) and phosphoramidate oligodeoxynucleotides (Asai *et al*, 2003).

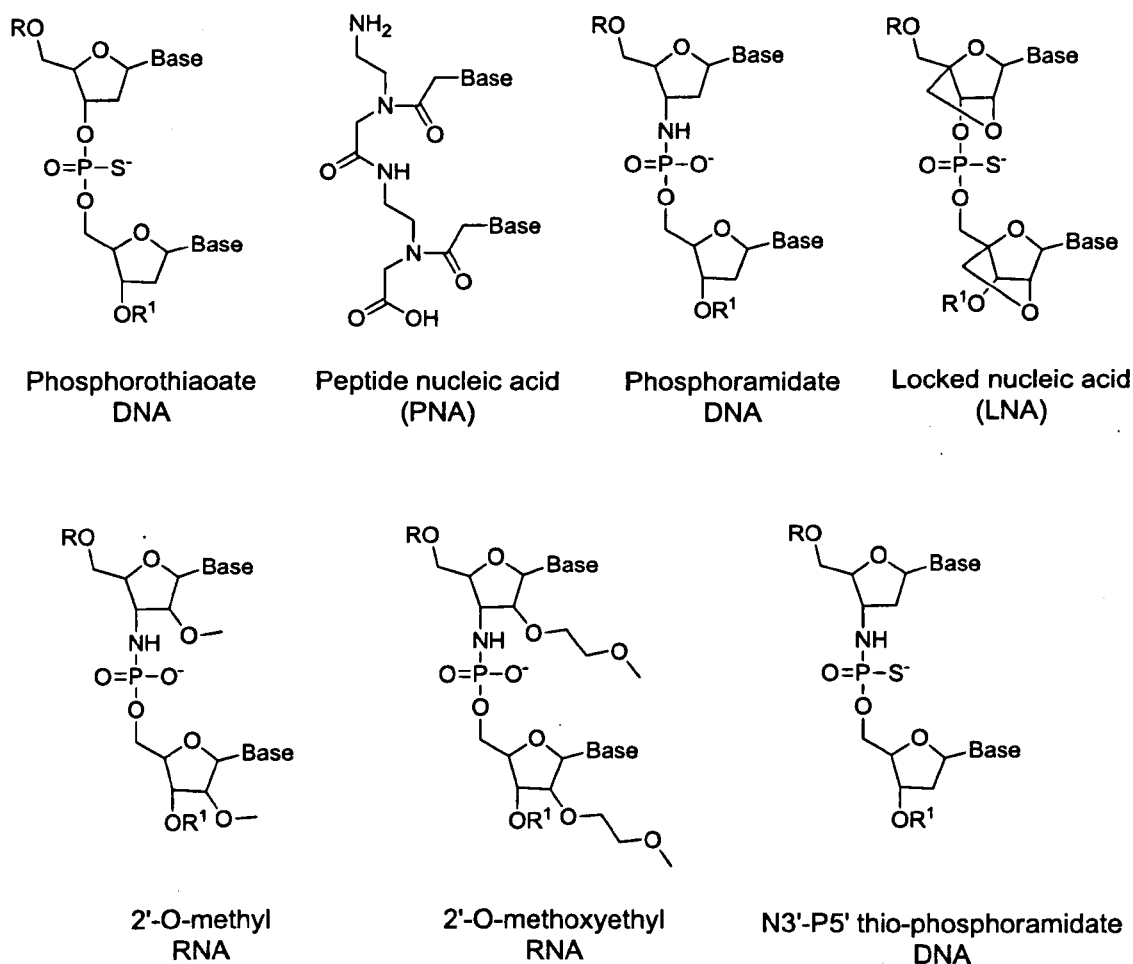


Figure 9: Chemical structures of modified oligonucleotides.

GRN163 is a 13-mer N3'-P5' thio-phosphoramidate oligonucleotide containing the sequence TAGGGTTAGACAA. GRN163 has recently shown promising anti-cancer activity and is a template antagonist which binds to the complementary 13-nucleotide long region which overlaps and extends four nucleotides beyond the 5' boundary of the hTR template. It inhibits telomerase activity with IC₅₀ of 0.14nM in the TRAP assay and induces telomere shortening, growth arrest and cell death in multiple cancer cell lines as well as tumour regression in prostate cancer xenografts. Its lipid-modified counterpart GRN163L which is attached to a C16 (Palmitoyl) lipid moiety has now entered phase 2 clinical trials (Geron Corporation) for the treatment of solid

tumors and hematological malignancies and could become the first telomerase inhibitor to be available for cancer treatment (Asai *et al*, 2003).

1.7.4 Limitations of telomerase inhibition

There have been several concerns regarding telomerase as an anticancer target. Firstly, direct inhibition of telomerase would have detrimental effects on human cells which express telomerase. Although telomerase is not detected in most normal tissues, the enzyme is expressed in hematopoietic progenitor cells, germline cells and other regenerative tissues as mentioned earlier in section 1.7.1. It is thought, however, that side effects would be minimal since telomerase expression is lower and infrequent in these cells compared to cancer cells. Furthermore, these telomerase positive-normal cells are less susceptible to telomerase inhibition since they maintain much longer telomeres compared to cancer cells and have a lower proliferation rate and as a result will incur less telomeric attrition (Shay and Wright, 2002). Despite this, correlations have been observed between short telomeres and human diseases associated with premature ageing syndromes (Vulliamy and Dokal, 2008), bone marrow failure (Calado, 2009), atherosclerosis (Benetos *et al*, 2004), Alzheimer's disease (Panossian *et al*, 2003) and liver cirrhosis (Wiemann *et al*, 2002). However it is not fully understood how telomere shortening contributes to disease development in affected individuals.

Telomerase inhibitors might incur a significant lag time before telomere shortening induces cell senescence or apoptosis. The lag time will depend on the initial telomere length. It is suggested that cells which have much longer telomeres will require more replications to erode their telomeric sequences compared to cells with short

telomeres which suggests that telomerase inhibition may only be effective against tumours with short telomeres (Shay and Wright, 2002).

Alternative mechanisms for telomere maintenance have also been reported. About 10-15% of cancers maintain their telomeres via the alternative lengthening of telomeres (ALT) pathway (Neumann and Reddel, 2002). The ALT pathway involves homologous recombination amongst telomeres but the mechanism by which it occurs is not fully understood. A distinguishing feature of ALT cells is the presence of exceptionally long and heterogeneous telomeres. Nuclear promyelocytic leukemia bodies which contain recombination proteins also tend to associate with ALT cells (Morrish and Greider, 2009). Pure telomerase inhibitors will therefore have no effect on cancer cells which maintain telomeres via ALT and may result in telomerase independent cancer cells which are drug resistant (Shay and Wright, 2002).

Chapter 2: Computer aided drug design (CADD)

2.1 Introduction

Over the last two decades, due to rapid developments in structural biology and computer capabilities, computer-aided drug design (CADD) has evolved very quickly and is dramatically changing the strategy and pipeline for drug discovery (Tang *et al*, 2006). The drug development process is challenging, time consuming and expensive. It takes 10-15 years and on average between USD 0.8-1.5 billion for a drug to enter market from concept (Tamimi and Ellis, 2009). CADD is becoming indispensable for major pharmaceutical companies and offers an *in silico* alternative to medicinal chemistry techniques for investigating the structure and predicting the biological activity of drugs. For this reason CADD is showing promise and is quickly being recognised globally with regard to its advantages of both speed and low cost (Huang *et al*, 2010).

There are several successful examples of marketed drugs whose development has been dependent on CADD techniques. In 2006, 5 out of 20 approved drugs on the market were derived using CADD techniques (Hegde and Schmidt, 2007). Early successes included zanamivir (Relenza) which was developed against neuraminidase as treatment for influenza (Varghese, 1999) and raltitrexed (Tomudex) which was developed against thymidylate synthase for the treatment of cancer (Rutenber and Stroud, 1996). Recent examples include imatinib (Gleevec) which inhibits the BCR-ABL kinase for the treatment of leukaemia and darunavir (Prezista), a protease inhibitor used for the treatment of HIV. However many more examples exist, reviewed in Congreve and Marshall (2010).

2.2 CADD strategies in drug discovery

In the design of novel drugs, strategies for CADD vary depending on the information available. CADD techniques fall into two broad categories, namely structure-based drug design (SBDD) and ligand-based drug design (LBDD). In SBDD, a detailed three-dimensional structure of the target is known from x-ray crystallography, NMR or homology modelling allowing the design of novel therapeutic agents and characterisation of key interactions between ligand and target. LBDD on the other hand relies on knowledge of known active ligands which interact with the target of interest. This allows scanning of databases for ligands that are similar on the basis of shape, structure and physicochemical properties (Meek *et al*, 2006).

2.2.1 Structure-based drug design

When a detailed structure of the target is known and the pocket of interest has been identified, two approaches to drug design can be adopted: virtual high throughput screening (vHTS) and *de novo* drug design. In vHTS also known as *in silico* screening, collections of compound libraries are screened against the target of interest. Molecules which tend to bind strongly to the target are then extracted for further testing. With computing power becoming cheaper, several million compounds can be screened in a matter of days on large clustered computer systems (Dutta *et al*, 2010).

Many compound databases are not structurally novel and are biased towards particular classes of compounds (Hert *et al*, 2009). For this reason *de novo* drug design is becoming increasingly important for the discovery of novel compounds. In *de novo* design, the 3D structure of the target is used to design new molecules from

scratch (Kalyaanamoorthy and Chen, 2011). Complete molecules can be designed from building blocks to fill in the binding site of the target protein. Building blocks can be atoms, fragments, functional groups or small molecules. Complete molecules can be constructed in two major ways within the binding pocket. In the fragment linking approach, entire molecules can be constructed by linking building blocks together by first mapping the binding site to identify possible anchor points for functional groups and finally joining these groups together using linkers to form complete molecules. In the growing approach method, the molecule is grown in the binding pocket from an 'embryo' molecule under the control of an appropriate search algorithm which assesses the growing possibility using a scoring function (Tang *et al*, 2006). Examples of *de novo* design programs include BUILDER (Roe and Kuntz, 1995) and LUDI (Böhm, 1992).

2.2.2 Ligand-based drug design

In some cases where a 3D structure of a target protein is not available, LBDD can be used which utilizes known ligands of the target protein as the starting point. LBDD relies on the general approach of elucidating the relationship of a compound's structure as well as its physicochemical attributes to its biological activity. This structure activity relationship (SAR) may then provide the foundation to predict new compounds with enhanced biological activities (Shim and MacKerell, 2011).

LBDD techniques range from pharmacophore models which identify essential structural features of a molecule which are required for biological activity, similarity searches which screen databases of compounds to find similar compounds to quantitative structure-activity relationships (QSAR) which provides quantitative

estimates of activities based on physicochemical properties (Shim and MacKerell, 2011).

2.3 Binding pockets

The identification and visualization of protein binding pockets is the starting point for many SBDD applications. Location of binding pockets is a fundamental step in order to investigate the molecular recognition mechanism and function of proteins and prerequisite for the docking of small molecule ligands (An *et al*, 2005). A major issue with binding site analysis is the lack of definition of what constitutes a binding pocket since it is not easy to exactly define where a binding pocket ends and free space begins (Perot *et al*, 2010; Hendlick *et al*, 1997).

Ligands tend to bind to pockets (clefts, grooves) on the surface of proteins, often some of the residues of ligand binding pockets are both solvent accessible and deep (Tan *et al*, 2011). Ligand binding sites vary widely in shape and size from being spherical in shape to curved grooves composed of several interconnecting sub pockets (Perot *et al*, 2010). Empirical studies show that drugs usually tend to bind to the largest pocket in the proteins surface since a large pocket provides a large surface area thereby increasing the opportunity for the protein to form interactions with the ligand (Laskowski *et al*, 1996).

Visual examination of protein surfaces is a remarkably effective approach for identifying binding pockets but due to the large increase in the number of protein structures becoming available, computer methods to predict ligand binding sites are becoming increasingly important (Harris *et al*, 2008). A variety of computational methods for identifying and characterising binding sites have been reported. Binding

pocket detection algorithms can be divided into two categories: geometric algorithms and probe/energy based algorithms.

2.3.1 Geometric algorithms

Geometric pocket detection algorithms are fast and simple. Surfnet (Laskowski, 1995) delineates gap regions as cavities on the proteins surface by fitting spheres into spaces between atoms. Depth (Tan *et al*, 2011) is a web server to predict small molecule ligand binding pockets in proteins. Since residues lining a ligand binding pocket are solvent exposed and deep, Depth measures the extent of residue burial within a protein by measuring the distance of a protein atom to its nearest water molecule from bulk solvent. Ligsite (Hendlich *et al*, 1997) is an improved algorithm to Pocket (Levitt and Banaszak, 1992) and is sufficiently fast for processing large number of proteins. It places a Cartesian grid around the biomolecule of interest and scans along the x, y and z axis and the cubic diagonals for areas which are enclosed on both sides of the protein. PASS (Brady and Stouten, 2000) identifies binding cavities using a set of spheres called active site points. The program coats the protein surface with a layer of spheres and filters out spheres which clash with the protein and which are not sufficiently buried.

2.3.2 Probe/energy based algorithms

Probe mapping and energy based algorithms incorporate some level of protein physics. These methods estimate the interaction energy between a specific probe molecule and the protein of interest. GRID (Goodford, 1984) places a grid around the protein of interest and measures the interaction between different probe groups and

protein residues. Qsitefinder (Alasdair *et al*, 2005) uses an energy based approach to locate binding pockets. It calculates the interaction energy between the protein and a methyl probe. Individual probes are subsequently clustered according to their spatial proximity and energetically favourable binding sites are then located by calculating the total interaction energy of the probes within each cluster. AutoLigand (Harris *et al*, 2008) uses a grid-based representation of the binding affinity potential to identify binding sites of maximum affinity. The program generates affinity potentials for hydrogen, carbon, nitrogen, oxygen and sulphur probes and the best binding site within the energy grid is calculated. AutoLigand has been tested on a set of 187 protein-ligand complexes and is found to be successful in 73% of cases in predicting the location as well as the approximate volume of the binding site.

2.4 Molecular docking

The ability to accurately predict how a small molecule drug binds to its target binding pocket is of great importance in rational drug design. In general, docking consists of two parts, first the computer program predicts the orientation (pose) of the ligand within the binding site of the target protein and secondly it estimates the binding free energy (affinity) of the complex formed, a process known as scoring (Kontoyianni *et al*, 2004). Generally a molecule with a good score is potentially a good binder (Kroemer, 2007).

Docking protocols consist of two components, a search algorithm which generates optimum number of ligand conformations which includes the experimentally determined binding mode and a scoring function which serves multiple purposes, first to differentiate the experimental binding mode from all other binding modes obtained

by the search algorithm and then to rank the different ligands according to their binding affinity (Taylor *et al*, 2002). Currently, there are more than 60 different docking programs available which have different search algorithms and scoring functions (Reviewed in Moitessier *et al*, 2008). Some commonly used docking programs include GOLD (Jones *et al*, 1997), AutoDock (Morris *et al*, 1998), FlexX (Rarey *et al*, 1996) and Glide (Friesner *et al*, 2004).

2.4.1 Search algorithm

The search algorithm determines the correct binding mode of a ligand which involves finding the correct placement, orientation and conformation of the ligand within the target binding site. An accurate search algorithm will find all possible binding modes between the ligand and receptor and treat both as flexible by exploring the six degrees of translational and rotational freedom of the ligand as well as exploring the internal degrees of freedom of both ligand and protein. Due to the large size of the search space, this is impractical and only a limited amount of the total conformational space can be explored. Hence a balance is required between the search space examined and computational expense (Taylor *et al*, 2002).

Early docking programs such as DOCK (Kuntz *et al*, 1982) treated both the ligand and target as rigid bodies. However most current docking algorithms treat the receptor as rigid and take into account the flexibility of the ligand with the exception of a few programs which also apply partial flexibility to side chains of the proteins (Jones *et al*, 1997). Several search algorithms used in molecular docking have been developed including Monte Carlo methods, Genetic algorithms, Distance geometry methods and Incremental construction methods.

2.4.1.1 Monte Carlo methods

Monte Carlo (MC) simulation methods are based on random sampling and are among the most established and widely used stochastic optimisation technique and is often used in combination with simulated annealing. The method uses a technique called importance sampling to generate states of low energy (Leach, 2001). A standard Metropolis MC method involves applying random Cartesian moves to the ligand such that the internal conformation of the ligand is changed through the rotation of a bond. Molecular mechanics subsequently calculates the energy of the ligand within the binding site and the move is either accepted or rejected based on a Boltzmann probability (Taylor *et al*, 2002). Programs which use MC methods include ICM (Abagyan *et al*, 1994), DOCK (Ewing and Kuntz, 1997) and MCDOCK (Liu and Wang, 1999).

2.4.1.2 Genetic algorithms

Genetic algorithms (GA) use ideas based on biological evolution. A population of potential solutions is first created by randomly generating conformations of a molecule. The population is represented as a chromosome where each member of the population encodes a possible solution (i.e. a potential ligand conformation within the protein binding pocket) to the docking problem. The docked conformation is then assigned a fitness score based on the relative merit of the solution (Jones *et al*, 1995).

The chromosome codes for the values of the torsion angles of rotatable bonds in the molecule as well as the information about the mapping of ligand H-bond atoms onto complementary H-bond atoms and mapping of ligand hydrophobic points onto protein

hydrophobic points. This information is usually stored as a linear string of bits of 0 and 1s. After decoding each chromosome, a least squares fitting procedure is used to position the ligand within the binding pocket. A fitness score for each member of the population is generated. As the population of chromosomes is iteratively optimised, a new population is generated. This new population is subjected to genetic operators. For example, a mutation may occur in a chromosome at each step or two chromosomes may possibly mate to give a child. The population will eventually evolve towards better solutions i.e. chromosomes which correspond to ligand dockings with good fitness scores since the selection of the parent chromosomes is biased towards fitter members of the population (Jones *et al*, 1995).

2.4.1.3 Distance geometry methods

Distance geometry methods explore the conformational space of a molecule in terms of distances between all pairs of atoms. They randomly generate many distance matrices which are then converted into conformations in Cartesian space. However it is not always possible to obtain low energy ligand conformations by this method since it is not always possible to arbitrarily assign values to the inter-atomic distances in a molecule. The inter-atomic distances are closely inter-related and many combinations of distances are geometrically impossible. Although these methods allow fast sampling of the conformational space, they do not always give better results (Leach, 2001).

2.4.1.4 Incremental construction methods

Incremental construction based methods are becoming popular due to lower computational expense. These methods work by identifying a base or anchor fragment in the ligand to construct the conformation of a molecule within the binding site. The anchor fragment is usually a rigid moiety such as a ring system which is placed into the binding pocket independently of the rest of the ligand. The anchor fragment is first docked into the binding pocket and the docked orientation of this fragment then represents the starting point for conformational analysis of the remaining ligand. Incremental construction methods have mainly been used in *de novo* ligand design, however these methods are also gaining widespread attention for use in docking. FlexX is a program which uses an incremental construction based method to generate ligand conformations in protein binding pockets (Rarey *et al*, 1996).

2.4.2 Scoring functions

The second component of molecular docking is the scoring function. The purpose of the scoring function is to differentiate the experimental binding mode from all other binding modes explored through the search algorithm (Taylor *et al*, 2002). The two most important aspects of a good scoring function is speed and accuracy and an ideal scoring function is computationally reliable and efficient (Huang and Zou, 2010).

There are three general classes of scoring functions which include: Force field-based methods, Empirical free energy scoring functions and Knowledge based functions.

2.4.2.1 Force field-based methods

Molecular mechanics force field scoring functions are based on a simple model of the interactions within a system. Force fields methods usually measure the sum of two energy values, the receptor-ligand interaction energy which is described using van der Waals and electrostatic energy terms and the internal ligand energy which is usually a measure of the steric strain induced by binding from processes such as stretching of bonds, the opening and closing of angles and the rotation about single bonds. The majority of force field methods consider only a single protein conformation hence the calculation of the internal protein energy is disregarded to simplify the scoring (Kitchen *et al*, 2004; Leach, 2001). The basic functional form of a force field consists of both bonded and non-bonded energies (eq.1):

$$\text{Energy (E)}_{\text{Total}} = E_{\text{bonded}} + E_{\text{non-bonded}} \quad (\text{eq.1})$$

The components of the bonded and non-bonded interactions are given by the following summations (eq.2 and eq.3). More sophisticated force fields may contain additional terms but they always contain these five components.

$$E_{\text{bonded}} = E_{\text{bond}} + E_{\text{angle}} + E_{\text{torsions}} \quad (\text{eq.2})$$

$$E_{\text{non-bonded}} = E_{\text{electrostatic}} + E_{\text{vanderWaals}} \quad (\text{eq.3})$$

Where E is termed as the steric energy, E_{bond} is the energy for bond stretching, E_{angle} is the energy for angle bending, E_{torsions} is the torsional energy due to twisting about bonds, $E_{\text{electrostatic}}$ is the energy due to electrostatic interactions usually modelled using Coulomb's potential and $E_{\text{vanderWaals}}$ is the energy due to van der Waals interactions usually modelled using a Lennard-Jones potential.

In molecular mechanics, atoms are treated as balls with a defined radius connected by bonds which are treated as springs. The bond and angle terms are usually modelled as harmonic oscillators by applying Hooke's Law (Leach, 2001). Two of the most popular force fields used are AMBER (Cornell *et al*, 1995) and CHARMM (Brooks *et al*, 1983).

2.4.2.2 Empirical free energy methods

Empirical scoring functions were developed from the pioneering work of Böhm (1994) to predict binding affinities between ligands and proteins. The scoring function is computationally efficient and estimates protein-ligand receptor binding by summing up interaction terms which are derived from weighted structural parameters of protein-ligand complexes. These weights are usually the free energy contributions from hydrogen bonding, hydrophobic and ionic interactions and are assigned by linear regression methods which fit predicted and experimental binding affinities against a given set of training complexes. Entropic contributions are also taken into consideration. The individual interactions reflect the way in which a medicinal chemist would break down the contributions to binding between protein and ligand (Gohlke and Klebe, 2001).

Several empirical scoring functions have been developed which differ in the number and nature of terms used to make up the equation, including X-Score (Wang *et al*, 2002) and the ChemScore function implemented in GOLD (Eldridge *et al*, 1997).

2.4.2.3 Knowledge based methods

Knowledge based scoring functions were developed from statistical analysis of protein-ligand crystal structures and are based on empirical knowledge. The scoring function uses structural information stored in databases to derive interaction potentials for protein-ligand atom pairs. These interaction potentials are also known as potentials of mean force (PMF). It is suggested that binding modes which fit to the maxima of distributions of occurrence frequencies among inter-atomic interactions between specific atom pairs in experimentally determined structures are the most favourable. Based on this notion, the scoring function should be able to rank high all ligand conformations which are geometrically similar to the native pose (Gohlke *et al*, 2000).

The crystal structure represents the optimum placement of the ligand atoms relative to the protein atoms and there are millions of observed distances between ligand and protein atoms when one considers hundreds of such protein-ligand complexes. This observed distance distribution of specific atom type interactions are converted in pairwise potentials. The final score is calculated by summing up all the interaction pairs between each ligand and protein atom lying within a certain distance. Examples of knowledge based methods include ASP score (Mooij and Verdonk, 2005) and Drug score (Gohlke *et al*, 2000).

2.4.3 GOLD

GOLD (Genetic Optimisation for Ligand Docking) is an automated docking program which is highly regarded for its accuracy and reliability within the molecular modelling community. GOLD takes into account the full conformational flexibility of the ligand

and partial flexibility of the protein in all dockings. During docking, GOLD specifically rotates the torsion angles of hydroxyl groups, especially those which belong to serine, threonine and tyrosine in order to optimise the hydrogen bonding interactions of these residues with ligands. Lysine NH_3^+ groups position is also optimised during docking (Jones *et al*, 1997).

GOLD uses a method based on fitting points to place the ligand in the binding site. The program first adds fitting points to hydrogen bonding groups on the protein and ligand. It then maps donor points on the ligand onto acceptor points in the protein and vice versa. Hydrophobic fitting points in the protein are also generated by GOLD onto which ligand CH groups are mapped (Verdonk *et al*, 2003).

2.4.3.1 Genetic algorithm

GOLD uses a Genetic Algorithm (GA) as described in section 2.4.1.2 to explore potential ligand binding modes. The ligand ring geometries and all dihedrals of ligand rotatable bonds and protein NH_3^+ and OH groups are modified by the GA (Verdonk *et al*, 2004). The conformation information is encoded by binary strings where each byte in the string encodes an angle of rotation about a rotatable bond. Two binary strings are used, one for the ligand and one for the protein. Each torsion is rotated between -180° and 180° in step size of 1.4° (Jones *et al*, 1997).

2.4.3.2 Validation and comparison

The ability to reproduce the experimental ligand binding mode is one of the key characteristics of a good docking program. A ligand is taken out of the protein-ligand x-ray crystal structure and docked back into the same binding site. The docked

binding mode is subsequently compared to the experimental binding mode and if the root-mean square deviation (RMSD) between the two structures is below 2.0 Å, it is regarded as successful (Verdonk *et al*, 2003). In this regard GOLD has been validated on a large set of 305 protein-ligand complexes from the CCDC/Astex test set and achieved an overall 68% success rate (Nissink *et al*, 2002).

The performance of any docking program will not only depend on the core algorithm but also in the time invested in parameterisation and optimisation of the methodology (Taylor *et al*, 2002). Several case studies have compared GOLD to other docking programs and have concluded GOLD to be marginally the best performing docking program in terms of accuracy and in giving the best enrichment rates (Blisantz *et al*, 2000; Jenkins *et al*, 2003; Li *et al*, 2010; Kirtay *et al*, 2007). Enrichment measures the quantity of active compounds which the program has placed at the top of the ranked list (Hawkins *et al*, 2008).

2.4.3.3 The fitness function

GOLD offers a choice of scoring functions, the two major fitness functions used in GOLD are GoldScore and ChemScore. Both scoring functions are as equally reliable however in certain cases one may perform better than the other. When screening large libraries of compounds one may rescore the docking poses with an alternative scoring function to improve the overall rank ordering of the ligands (Verdonk *et al*, 2003).

2.4.3.3.1 GoldScore

The GoldScore fitness function is a force field based default scoring function in GOLD and is made up of four terms: protein-ligand hydrogen bond energy (external H-bond), protein-ligand van der Waals energy (external vdW), ligand internal vdW energy (internal vdW) and ligand torsional strain energy (internal torsion). The ligand intra-molecular hydrogen bond energy (Internal H-bond) can also be added as a fifth component. A covalent term can also be included within the scoring function when docking covalently bound ligands. The final fitness score is taken as the sum of negative energy terms such that larger fitness scores are better (eq. 4). An empirical correction is also made when the total fitness score is computed to encourage protein-ligand hydrophobic contact by multiplying the external vdW score by a factor of 1.375.

$$\text{GoldScore} = S(\text{hb_ext}) + 1.375 * S(\text{vdw_ext}) + S(\text{hb_int}) + S(\text{vdw_int}) + S(\text{tors}) \quad (\text{eq.4})$$

Where $S(\text{hb_ext})$ is the protein-ligand hydrogen bond score, $S(\text{vdw_ext})$ is the protein-ligand van der Waals score, $S(\text{hb_int})$ is the score from intramolecular hydrogen bond in the ligand, $S(\text{vdw_int})$ is the score from intramolecular strain in the ligand and $S(\text{tors})$ is the score from the ligand torsional strain.

The GoldScore fitness function has been specifically optimised for the prediction of ligand binding positions, however some correlation with the prediction of binding affinities have been observed (Jones *et al*, 1997).

2.4.3.3.2 ChemScore

ChemScore is an empirical scoring function which was originally developed by (Eldridge *et al*, 1997) expanding on the work of others and has since then been incorporated into GOLD. The fitness function has been empirically derived and trained by regression against measured binding affinity data from a set of 82 protein-ligand complexes. However there is no clear evidence that it is superior to GoldScore in predicting affinities.

ChemScore uses simple contact terms in order to estimate hydrogen bonding interactions, lipophilic and metal-ligand binding interactions. ChemScore estimates the total free energy change which occurs when ligands bind to proteins as shown in the equation given below (eq. 5).

$$\Delta G_{\text{binding}} = \Delta G_0 + \Delta G_{\text{hbond}} + \Delta G_{\text{metal}} + \Delta G_{\text{lipo}} + \Delta G_{\text{rot}} \quad (\text{eq. 5})$$

Each component in the equation is the result of a term which depends on the magnitude of a specific physical contribution to free energy (e.g. hydrogen bonding) and a scale factor which is obtained by regression. To prevent against poor internal conformations and close contacts in docking, an internal torsion term and a clash penalty is added to the equation when the final ChemScore is calculated (Jones *et al*, 1997).

Chapter 3: Basis of present investigation

3.1 Introduction

Under physiological ionic conditions, telomeric G-rich sequences can adopt other secondary structures apart from the Watson and Crick duplex known as G-quadruplexes. Although the exact function of these structures is unknown, they are thought to play important roles in key biological processes (Balasubramanian and Neidle, 2009).

G-quadruplexes are formed from G-quartets where each quartet consists of four guanine bases held together in a cyclic arrangement by Hoogsteen hydrogen bonding. The G-quartets stack on top of each other in a helical fashion to form the G-quadruplex structure (Figure 10). The formation of G-quadruplexes is dependent on the presence of monovalent cations such as potassium (K^+) and sodium (Na^+) ions which sit in a central channel between each pair of G-quartets providing stability to the structure (Phan, 2010).

G-quadruplex structures can fold in a variety of ways and can be formed by a single strand (intramolecular) or by several strands (intermolecular) to form monomeric, dimeric and tetrameric structures by using one, two or four separate G-rich strands (Phan, 2010). The relative orientations of the strands and glycosidic conformations (syn/anti) of the guanines give rise to a variety of G-quadruplex structures (Figure 10).

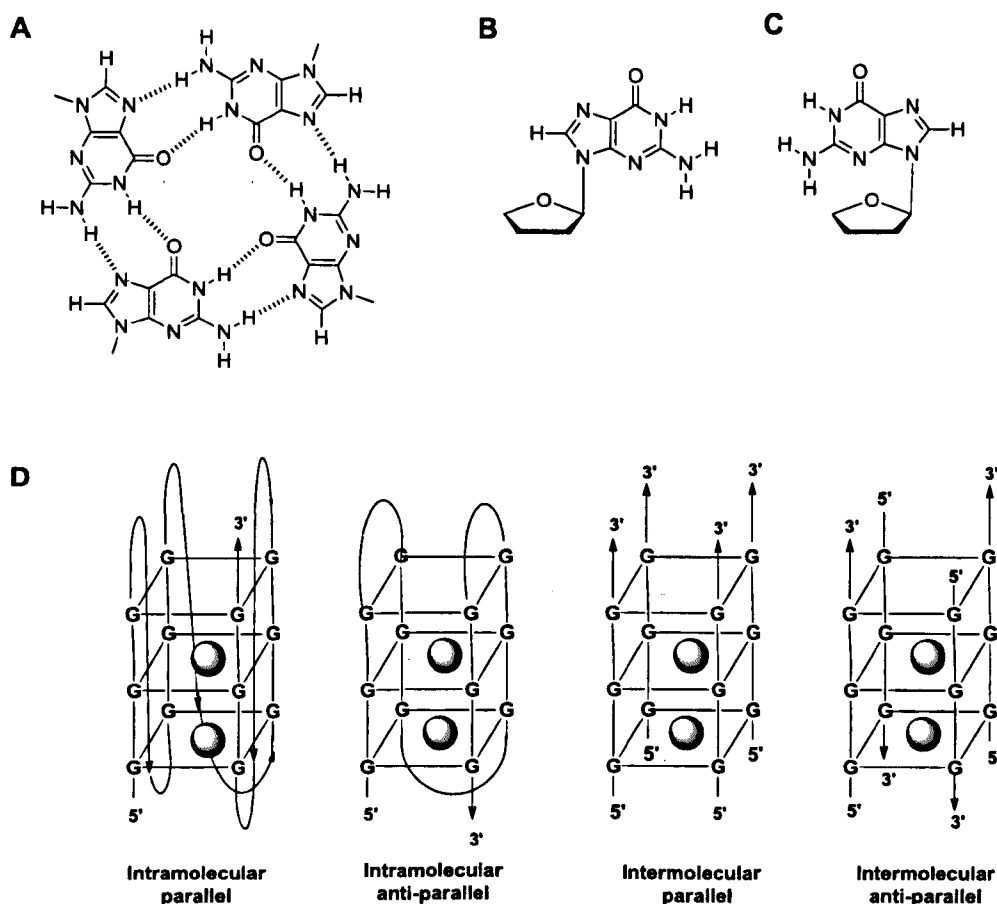


Figure 10: (A) Schematic diagram of a G-tetrad (B, C) Guanine in (B) anti and (C) syn glycosidic conformations (D) Four different G-quadruplex structures. Arrows indicate the strand orientation from 5' to 3' direction (Adapted from Phan, 2010).

The presence of G-quadruplex structures at telomeres has attracted a great deal of interest in recent years since telomerase requires an unfolded and therefore accessible single stranded DNA substrate for telomere elongation. The folding of the 3' single stranded overhang into a G-quadruplex structure will prevent the effective recognition of the telomere overhang by the telomerase RNA template domain since it requires it to be single stranded thereby inhibiting telomere elongation by telomerase. Ligands which can stabilise G-quadruplex structures therefore have the potential to inhibit telomerase activity at telomeres (Zahler *et al*, 1991).

The ability of small ligands to inhibit telomerase activity through G-quadruplex stabilization was first demonstrated using a disubstituted anthraquinone derivative (Sun *et al*, 1997). Since then several telomeric G-quadruplex interactive agents have been reported (Cian *et al*, 2008).

Telomeres are not the only DNA sequences in the human genome which can form G-quadruplex structures. A growing body of evidence suggests that G-quadruplexes can be found in gene promoter sequences of specific oncogenes such as c-myc, VEGF, c-kit and bcl2 (Qin and Hurley, 2008) suggesting that G-quadruplexes may be directly involved in gene regulation at the transcription level (Figure 11). This hypothesis offers attractive opportunities for the discovery of small molecules which can target disease-related genes by interacting with such DNA G-quadruplexes and modulate the expression of the gene regulated by the promoter sequences. In support of this hypothesis, genetic point mutations which destabilised the c-myc G-quadruplex structure led to an increase in transcriptional activity whereas TMPyP4, a small-molecule G-quadruplex stabiliser was found to suppress c-myc transcriptional activity (Jain *et al*, 2002).

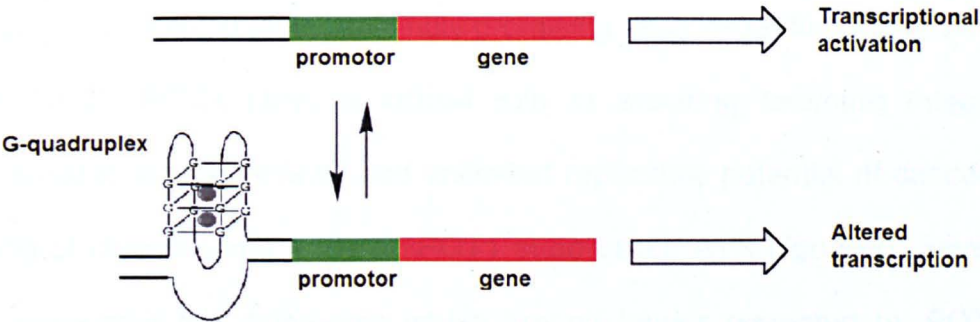


Figure 11: G-quadruplex formation in the gene promoters alters transcription (Adapted from Huppert and Balasubramanian, 2007).

It was originally hypothesized that G-quadruplex ligands were simple telomerase inhibitors which would block telomerase access to the G-overhang resulting in the reduction of telomeric DNA and induction of senescence in cancer cells and would require an extended time scale before antitumor effects became apparent (Bilsland *et al*, 2011). Emerging evidence now suggests that G-quadruplex ligands elicit their effects at least in part by affecting the telomere capping ability of shelterin proteins and produce more rapid toxic effects leading to rapid senescence and cell death (Bilsland *et al*, 2011).

G-quadruplex ligands therefore have the ability to inhibit telomerase-positive cancer cells as well as ALT dependent cancer cells. However due to the wide existence of G-quadruplex structures in the human genome, there is a concern regarding the selectivity and mechanism of these ligands. A new and rational approach is to target the shelterin proteins directly which may afford more selective ligands (Bilsland *et al*, 2011).

3.2 Human POT1 as a viable anti-cancer target

POT1, the shelterin protein which binds and caps the 3' single stranded telomeric overhang is becoming an increasingly promising drug target for cancer treatment (Lei *et al*, 2004). POT1 plays a critical role in ensuring telomere integrity and is indispensable for the survival and unlimited replicative potential of cancer cells. The capping of chromosome ends by POT1 suppresses ATR signaling. Several studies have suggested that telomeres which are no longer protected by POT1 become dysfunctional and activate the ATR DNA damage response pathway which triggers senescence, including p53-dependent apoptosis (Denchi and de Lange, 2007).

Growing evidence also suggests that the anticancer activity of several G-quadruplex ligands including Telomestatin (Gomez *et al*, 2006), RHPS4 (Salvati *et al*, 2007), BRACO-19 (Gunaratnam *et al*, 2007) and BRACO-19 dimers (Fu *et al*, 2009) is through the uncapping of POT1 from telomeric DNA. These G-quadruplex ligands cause telomere dysfunction not only by telomere shortening as would be expected from telomerase inhibition, but also cause telomere injury by disrupting the binding of POT1 to telomeric DNA which consequently results in a potent DDR (Salvati *et al*, 2007).

Telomestatin induces apoptosis in cancer cells via a mechanism which involves the uncapping of POT1 from telomeres. However the ligand has no effect on TRF2 binding (Gomez *et al*, 2006). RHPS4 induced short term apoptosis as well as senescence in melanoma cell lines when used at high concentrations. However no telomere shortening was observed in these cells and the apoptosis phenotype correlated with increased incidence of telomere fusions due to the liberation of POT1 from telomeres. RHPS4 treatment in mice xenografts significantly reduced tumour weight due to apoptosis but no telomere erosion was observed. Tumours which overexpressed POT1 were resistant to RHPS4 treatment (Salvati *et al*, 2007).

POT1 is expressed in a wide range of tissues (Baumann and Cech, 2001); however recent studies on RHPS4 suggest that RHPS4 does not trigger telomere damage in normal human fibroblasts. The lack of telomere damage in normal cells correlates with the inability of RHPS4 to liberate POT1 from telomeres. It is proposed that the shelterin protein composition of cancerous and normal cell telomeres may differ (Salvati *et al*, 2007). One such possibility is that the shelterin proteins in normal cells

are more strongly bound to telomeric DNA than in cancer cells, thereby providing an opportunity to selectively target cancer cells without effecting normal cells.

3.3 Human POT1 function

The precise role of POT1 varies considerably between organisms. In fission yeast, POT1 deletion leads to rapid loss of telomeric sequences and widespread cell death (Baumann and Cech, 2001). In mice, two POT1 genes have been found which encode two POT1 proteins, POT1a and POT1b. It is thought that the two proteins have distinct functions at the telomeres. POT1a suppresses a DNA damage signal at telomeres whereas POT1b regulates the amount of single-stranded DNA at chromosome ends (Hockemeyer *et al*, 2006).

Removal of POT1 in vertebrates results in an ATR mediated checkpoint response and rapid cell cycle arrest and cell death (Churikov *et al*, 2006). Several phenotypes have been observed when human POT1 levels are depleted (Yang *et al*, 2005). Knockdown of human POT1 in HeLa cells using siRNA results in telomere instability, apoptosis, decrease in cell viability and telomere elongation. In HT1080 and IMR90 cells, a senescent-like growth arrest is induced (Veldman *et al*, 2004).

The inhibition of POT1 using antisense oligonucleotides results in telomere shortening and increase in the frequency of anaphase bridges, a telltale sign of telomere dysfunction (Kondo *et al*, 2004). In another study, reduction of POT1 by RNA interference leads to the loss of single strand telomere overhangs, increased p53 expression, chromosomal instability, senescence and apoptosis in cancer cells (Yang *et al*, 2005). The expression of POT1 in gastric cancers is also significantly higher compared to adjacent normal tissue and its expression is correlated with

tumour stage being higher in late-stage tumours (Gao *et al*, 2011). These results suggest that POT1 may be a viable anti-cancer target.

3.4 Human POT1 crystal structure

The crystal structure of the N-terminal DNA binding domain of human POT1 bound to a telomeric single stranded DNA decamer TTAGGGTTAG (Figure 12) has been solved at a resolution of 1.73Å (Lei *et al*, 2004).

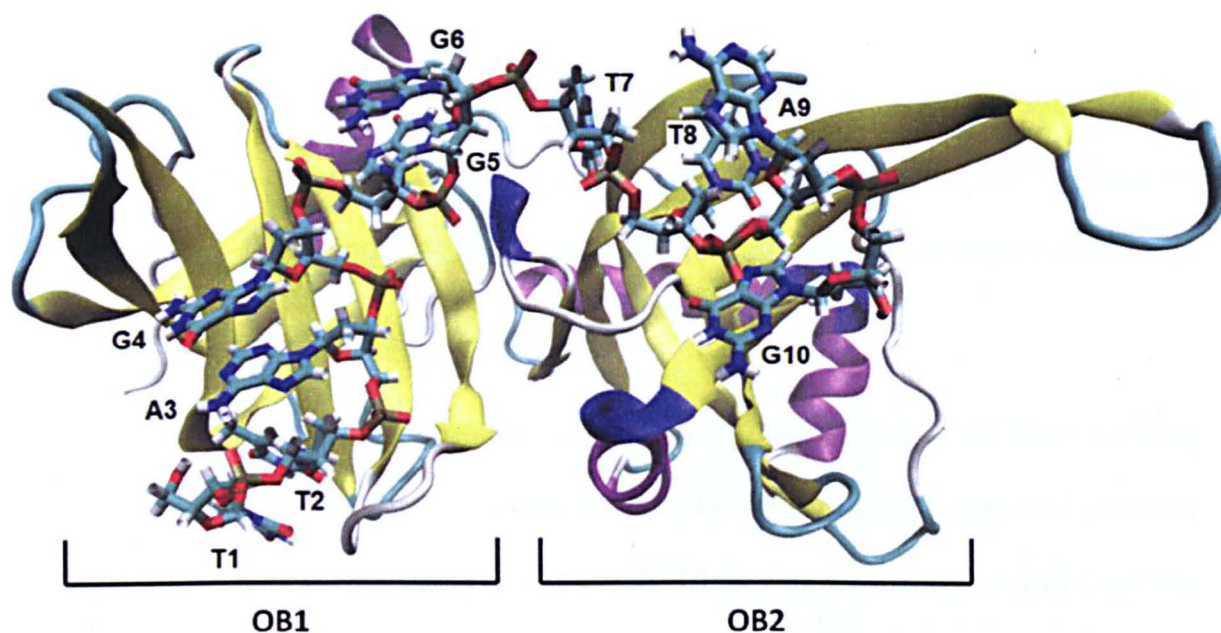


Figure 12: Crystal structure of human POT1 (solid ribbon) bound to telomeric DNA (CPK sticks) discovered using x-ray crystallography. The N-terminal region of POT1 consists of two domains: OB1 and OB2. The β -sheets are displayed as yellow ribbons and the α -helices are displayed as purple coils.

The N-terminal half of human POT1 contains two oligonucleotide binding (OB1 & OB2) domains which adopt an extended conformation connected by a short linker. The DNA adopts an irregular and extended conformation and binds in a continuous

basic concave groove spanning both OB1 and OB2 domains. The DNA bases are completely or partially buried in a total solvent excluded contact area of 900Å with its backbone exposed to solvent (Lei *et al*, 2004).

In total, 31 H-bonds are observed in the bound complex, with OB1 making a much more extensive contact with bases T1-G6 contributing to a total of 22 hydrogen bonds with OB1 residues. A total of 9 H-bonds are observed between OB2 residues and bases T7-G10, out of which eight H-bonds are contributed by T7 and G10. Base A9 makes no H-bonding interaction with POT1. The DNA is governed by several stacking interactions. T1 stacks with T2, A3 with G4, G5 with G6 and T8 with A9. The T7 base is sandwiched in a stack of aromatic side chains and the last base G10 at the 3' end of DNA is buried within the protein with its edge surrounded by four H-bond interactions thus rendering it inaccessible for telomerase extension (Lei *et al*, 2004).

POT1 has the ability to accommodate additional 3' nucleotides in its DNA binding cleft since the terminal G10 base has its 3'-hydroxyl group solvent exposed which is consistent with the idea that several human POT1 proteins could bind and coat the entire single stranded overhang. The protein can therefore remain in phase with the telomeric sequence with a proportion of POT1 in complex with TPP1 (Lei *et al*, 2004).

The presence of a high resolution crystal structure of the N-terminal domain of human POT1 provides an insight into the mechanism by which POT1 recognises, protects and binds telomeric DNA and therefore provides a good starting point for structure-based drug design.

Disruption of protein-DNA interactions for a long time has been considered extremely challenging. However considerable progress in recent years to identify inhibitors which disrupt protein-DNA interactions has been achieved.

Chan *et al* (2008) have discovered inhibitors of the AIF-DNA interaction with the best inhibitor achieving an IC_{50} of $23\mu M$ using a photonic crystal based assay. Kong *et al* (2005) have developed compounds to inhibit the interaction between the transcription factor HIF-1 and its DNA binding partner. Using an EMSA based assay, a potent inhibitor which completely inhibited HIF1-DNA binding at concentrations as low as $320nM$ was discovered. Nehar *et al* (2010) have developed inhibitors against the xeroderma pigmentosum group A (XPA) protein-DNA interaction and have identified an inhibitor which inhibits XPA-DNA interactions at IC_{50} of $20\mu M$ using an ELISA based assay. Andrews and Turchi (2004) have discovered several inhibitors which perturb RPA-DNA binding using a fluorescence based assay. Several compounds have been identified which have IC_{50} in sub micromolar range. Overall this data suggests that the approach which is being undertaken to target POT1-DNA interactions is feasible.

According to the literature, no research has been done to identify inhibitors of POT1-DNA interactions. The aim of this project is to design one of the first novel small molecule inhibitors of human POT1 which may have the potential to disrupt the binding interaction of POT1 with its single stranded DNA binding partner thereby driving cancer cells into senescence/apoptosis.

The following guiding principles will be employed to discover POT1 inhibitors:

- Identify a drug binding site in human POT1 using active site molecular modeling techniques.
- To carry out the *de novo* design of a specific class of POT1 inhibitor that targets the predicted binding site.
- To design focused virtual compound libraries based on the designed inhibitor and carry out docking studies to evaluate top scored structures.
- Synthesize and characterize compounds using NMR, MS and HPLC.
- To carry out biological screening of compounds and identify hits which can be further optimized to enhance activity.
- To carry out optimization and build a complete structure-activity relationship (SAR) profile of hits in order to identify a potent inhibitor of POT1-DNA interaction for further development into an ideal drug candidate.

The research plan is illustrated in Figure 13.

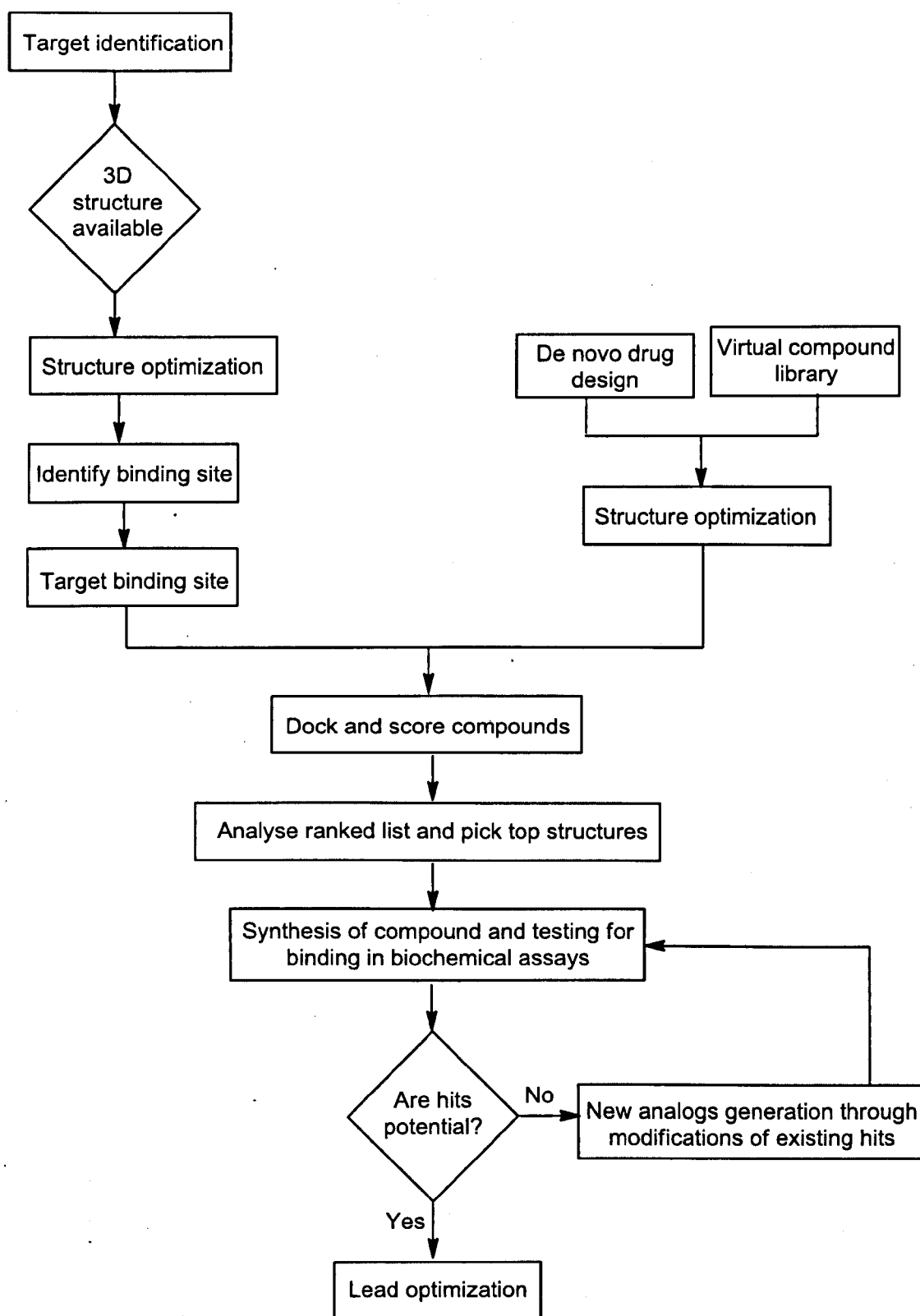


Figure 13: Illustration of research plan. The plan includes protein structure preparation, binding-site identification, drug design, compound library preparation, docking of compounds against the protein binding pocket and analyses.

Chapter 4: Protein-DNA interaction analysis

4.1 Introduction

Understanding the nature of protein-DNA interactions is important in drug discovery. Alongside the theoretical studies and methods described above, there exist a wide range of experimental methods to study this phenomenon. In this chapter, an introduction to some of these techniques will be presented.

4.2 Filter binding assay

Filter binding assays were developed in the early 1970's and provide a rapid and simple way of measuring protein-DNA interactions. The assay measures the affinity between a protein and DNA and is based on a simple process. In a typical experiment, both the protein and DNA are mixed and incubated together. The mixture is then subjected to electrophoresis separation and blotted onto a negatively charged nitrocellulose filter. The protein will bind to the nitrocellulose since it has a net positive charge whereas the DNA will not remain on the filter since it contains negatively charged phosphate groups. However DNA which is bound to the protein is also retained on the nitrocellulose and can be quantified by measuring the radioactivity of the label which was introduced to the DNA prior to the incubation with the protein (Helwa and Hoheisel, 2010).

4.3 EMSA assay

Electrophoresis mobility shift assay (EMSA) measures the electrophoretic mobility of the nucleic acid. The assay is based on the observation that molecules will move through the gel at different speeds according to their size and charge hence protein-DNA complexes will migrate more slowly than that of the free DNA fragment. The assay is simple and robust to accommodate a range of binding conditions. In a typical experiment, the nucleic acid is labelled with a radioisotope such as ^{32}P and combined with the protein. Fluorescent and biotin nucleic acid labels are also widely used. The mixture is subjected to electrophoresis through either a polyacrylamide or agarose gel under native conditions. After completion, the free DNA and DNA which is bound to the protein is determined by autoradiography of the labelled nucleic acid (Hellman and Fried, 2007).

4.4 ITC

Isothermal titration calorimetry (ITC) is a powerful tool for measuring binding affinities and provides the thermodynamics of binding between protein and DNA. In a typical ITC experiment, a syringe containing DNA is titrated into a sample cell containing solution of the protein. As the two binding partners interact, the heat released upon their interaction can be measured over time. As successive amounts of DNA are titrated into the ITC cell, heat is either released or absorbed. The heat is proposed to be in direct proportion to the amount of protein-DNA binding. The protein in the cell will eventually become saturated with the added DNA and the heat signal will diminish until only the background noise is observed due to the heat of dilution. Integration of the heat measurements allows the determination of binding affinity (K_d),

stoichiometry (n), binding enthalpy (ΔH), entropy (ΔS), heat capacity (ΔC) and the change in free energy (ΔG) upon protein-DNA binding (Liang, 2008).

4.5 SPR assay

Surface plasmon resonance (SPR) has emerged as a powerful optical detection technique for studying label-free protein-DNA interactions in real-time. In a typical SPR experiment, a light source is passed through a prism and is reflected off the backside of a sensor chip and into a detector. The sensor chip is usually a metal surface made from a thin layer of gold or silver. At a certain incident angle, light is absorbed by the electrons of the metal surface causing them to resonate. These resonating electrons are also called surface plasmons. The resonance effect causes the reflected light from the sensor chip to lose intensity which can be observed as an SPR minimum in reflectivity (Ritzefeld and Sewald, 2011).

The surface plasmons are sensitive to changes in the surrounding environment. In order to measure protein-DNA interactions, the DNA is immobilised on the surface of the sensor chip. The interacting protein is injected into the running buffer that passes the surface at a constant flow. The mass on the sensor chip increases when protein molecules bind to DNA. Since the incident angle at which resonance occurs is dependent on the refractive index. Protein binding to the other side of the sensor chip causes a change in the refractive index resulting in the shift of the SPR minimum which can be measured to quantify the amount of binding (Ritzefeld and Sewald, 2011).

4.6 Fluorescence Polarisation assay (FP)

4.6.1 Introduction

Fluorescence polarisation (FP) is an indispensable technique for studying molecular interactions and is widely used to study protein-DNA and protein-protein interactions (Moerke, 2009). It has been successfully used to study a range of targets including nuclear receptors, G-protein coupled receptors, proteases and kinases. FP is based on the principle that when a fluorescently labelled molecule becomes excited by plane polarised light, the molecule will emit light to a certain degree of polarisation which is inversely proportional to the rate of molecular rotation. This property of FP can be used to give a direct, nearly instantaneous measure of the interaction between a small labelled ligand and a large protein and provides the foundation for competition binding assays (Moerke, 2009).

As an assay technology, FP is widely used in high throughput screening applications and has several key advantages. It allows molecular binding events in solution and permits true equilibrium analysis in the low picomolar range. Since FP measurements are taken in 'real time', kinetic experiments are also widely conducted. FP is readily adaptable to low volumes (i.e. in the order of 10 μ l) and is a non-radioactive technique. Due to the truly homogeneous nature of FP, it does not require separation of bound and free ligand (Moerke, 2009).

4.6.2 Theory

The theory of FP was first described in 1926 by Perrin and is based on the concept that when a small fluorescently labelled molecule (typically <1500 Da) is excited by

plane-polarised light, the emitted light will largely be depolarised since the molecule is rapidly tumbling in solution during its fluorescence lifetime (the time between excitation and emission) which is caused by the Brownian molecular rotation of the labelled ligand. However, when the labelled ligand is bound to a much larger protein (typically >10kDa), it will tumble less rapidly in solution and reorients to a much smaller degree thus allowing the emitted light to be polarised to a significant degree. The basic principle of FP is presented in Figure 14.

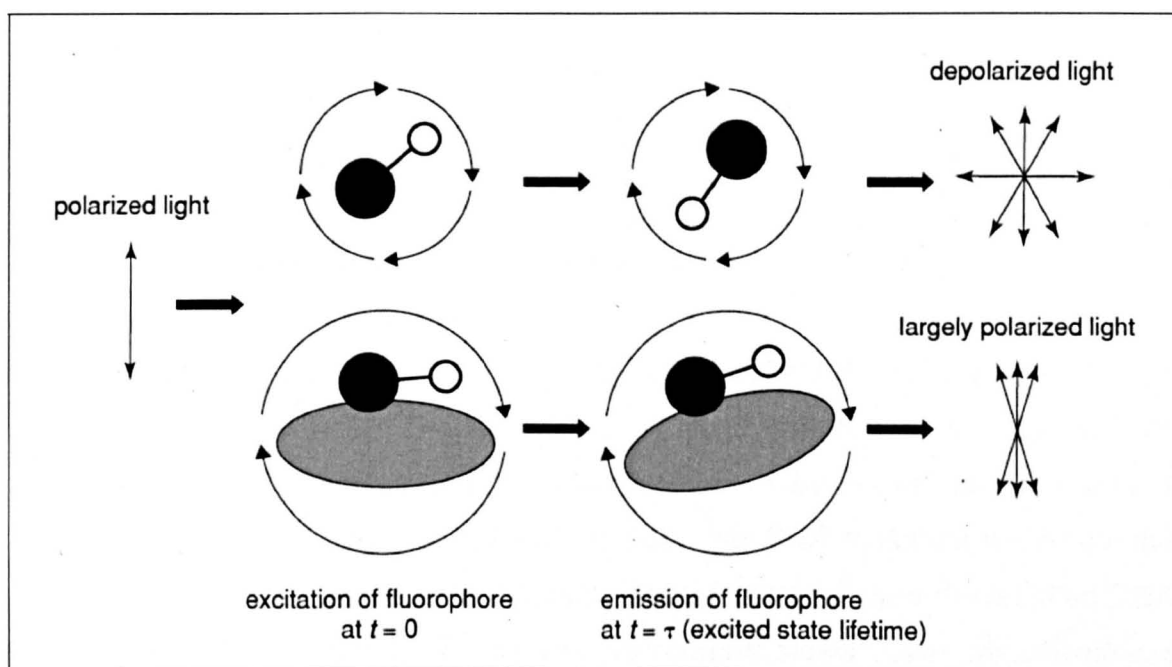


Figure 14: Schematic diagram illustrating the fundamental principle of FP. When a small ligand such as a nucleic acid (dark circle) is attached to a fluorescent label (white circle) becomes excited by polarised light at the excitation wavelength of the fluorophore, the nucleic acid reorients to a considerable degree as a result of molecular tumbling during the excited state lifetime of the fluorophore. This tumbling effect causes depolarization of the emitted light. However when the nucleic acid is bound to a large protein (Gray ellipse), the resulting complex will tumble at a slower rate and therefore the emitted light retains its polarisation (Taken from Moerke, 2009).

To test the ability of compounds to inhibit the POT1-DNA interaction, an FP displacement assay was developed by Dr Lodewyk Dekker and his team at the University of Nottingham. The basic principle of the POT1 FP assay is depicted in Figure 15 and the assay methodology is presented in Chapter 10.3.

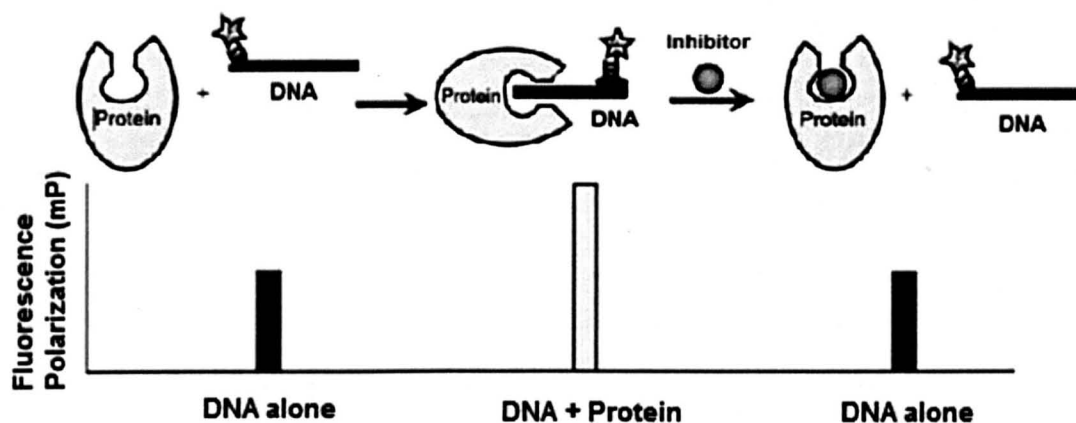


Figure 15: Schematic graph depicting how the POT1 FP assay works. A 15mer oligonucleotide sequence TTAGGGTTAGGGTTA is attached to fluorescein. When the fluorescent DNA molecule in solution is excited by plane-polarised light, the emitted light will largely be depolarised because the DNA molecule will be tumbling and rotating rapidly in solution (Low polarization). However when the labelled DNA is bound to the much larger POT1 protein, its rotation slows down considerably, the emitted light will thus be polarized to a significant degree (Polarization changes from low to high polarization). When a non-fluorescent inhibitor binds to POT1 it will displace the DNA into the solution where it will tumble more rapidly thus decreasing the FP of the emitted light. By measuring the decrease in the FP signal upon inhibitor binding, a competition binding assay can be established to measure the decrease in POT1-DNA binding.

The rest of the thesis is arranged as follows: Chapter 5 describes how the binding site in POT1 was identified. Chapter 6 describes ligand design, hit identification, modifications and biological screening results for inhibitors based on the pyrido[1,2-a]pyrimidin-2-one template. The POT1 library of compounds and their biological screening results are presented in Chapter 7. The inhibitors based on the sulfathiazole template and their biological results are presented in Chapter 8. In Chapter 9, a general conclusion and future work are presented and the key methodologies used for this work are described in Chapter 10.

Chapter 5: POT1 binding site characterisation

5.1 Introduction

Since no co-crystallised structure of human POT1 bound to any inhibitors yet exists, the first task was to identify potential binding sites in POT1 since the ligand binding site is not known. In this section the different computational tools which were used to locate and characterise the ideal binding pocket in POT1 are discussed.

5.1.1 POT1 potential binding sites

A tool in SYBYL[®] named Molcad Multichannel Surfaces was used to identify potential binding pockets (solvent accessible regions) prior to ligand design (Figure 16).

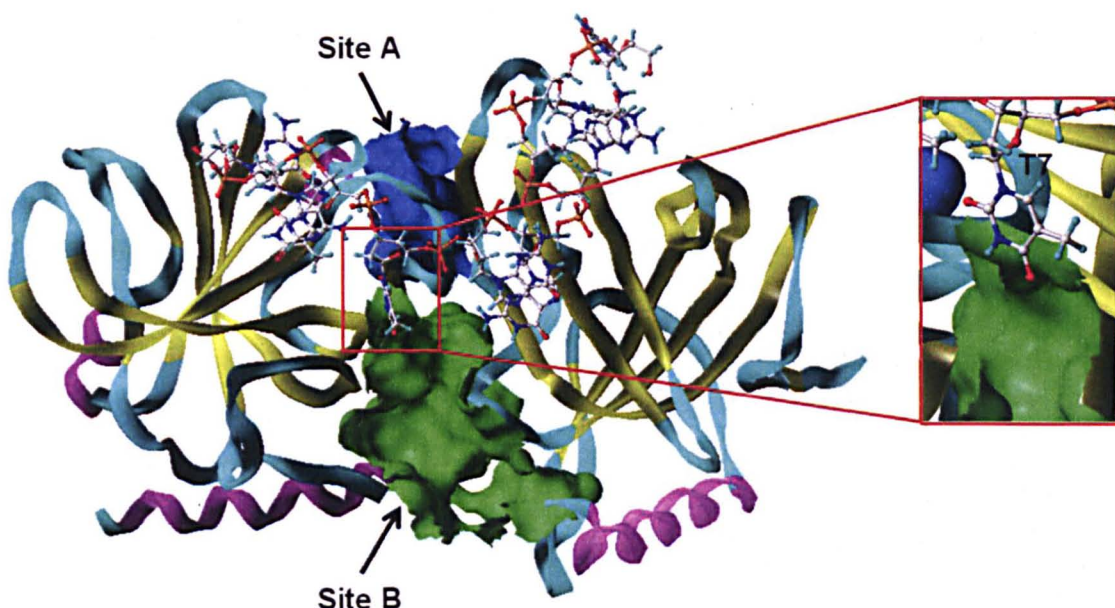


Figure 16: Structure of POT1 (solid ribbon) bound to telomeric DNA (CPK ball and stick). The magenta, yellow and cyan colours in the figure represent helix, β -strand and loop regions, respectively. Two binding clefts are identified: Site A (Blue region) has a surface area of 304.87 \AA^2 and cavity volume of 273.57 \AA^3 and Site B (green region) contains a surface area of 1156.84 \AA^2 and cavity volume of 1148.67 \AA^3 .

From Figure 16, it can be seen that SYBYL[®] has identified two major solvent accessible clefts, both lying at the OB1-OB2 interface in POT1. Site A (blue region) does not make any contact with DNA and lies away from where the DNA binds. A drug targeting site A will not have a direct inhibitory effect on DNA binding but could potentially act in an allosteric manner. Site B (green region) has attracted particular interest as a potential drug target, not only is it the largest of the two binding clefts but is also located near the region of space where the DNA binds to POT1. The upper surface of Site B is occupied by the thymidine-7 base of DNA. It is suggested that ligands which can disrupt the binding of thymidine-7 may inhibit POT1-DNA binding.

5.1.2 Favourable POT1 binding pocket

Proteins function by interacting with other bio-molecules through their interfaces. It has been found that the binding energies are not uniformly distributed at protein interfaces since some critical residues known as hot spots comprising a small fraction of the interface account for the majority of the binding energy (Tuncbag *et al*, 2009). Typically a hot spot contributes more than 1-2 kcal/mol to the free energy of binding (Recio, 2011).

The distribution of amino acids in hot spots vary, tryptophan, arginine and tyrosine residues are more likely to be found in these regions whereas certain residues are found rarely in hot spots including leucine, methionine, serine, threonine and valine

(Bogan and Thorn, 1998). Hot spots have therefore become important targets for rational drug design and can be found experimentally by evaluating the change in binding free energy by mutating a certain residue to an alanine. Since these mutagenesis studies require significant experimental effort, there is a need for accurate and reliable computational methods to predict hot spots (Tuncbag *et al*, 2010).

The ten nucleotide DNA sequence of human POT1 binds in an irregular and extended conformation along the POT1 surface forming several key interactions with several amino acids. It is suggested that only a few residues in the POT1-DNA complex may contribute to the binding affinity and therefore are absolutely essential for interaction. These hot spot residues may act as potential drug targets such that small molecules could be specifically designed to target these regions and have a greater ability to disrupt the complex. The aim is to target regions where the interaction between POT1 and DNA are strong to cause maximum effect of liberating DNA from POT1.

One program capable of identifying potential hot spots in protein-DNA complexes is Anal, an energy analysis module of AMBER 7 (Case *et al*, 2002). Anal carries out the energetic analysis of individual structures. The program decomposes the energy amongst the different groups of atoms in order to find the interaction energies between different parts of the structure. Amino acids having the most favourable interaction energy with the DNA, and DNA bases having favourable interaction energy with POT1 are calculated (Figure 17). The two most important terms that Anal calculates are van der Waals and electrostatic interactions between POT1 and DNA.

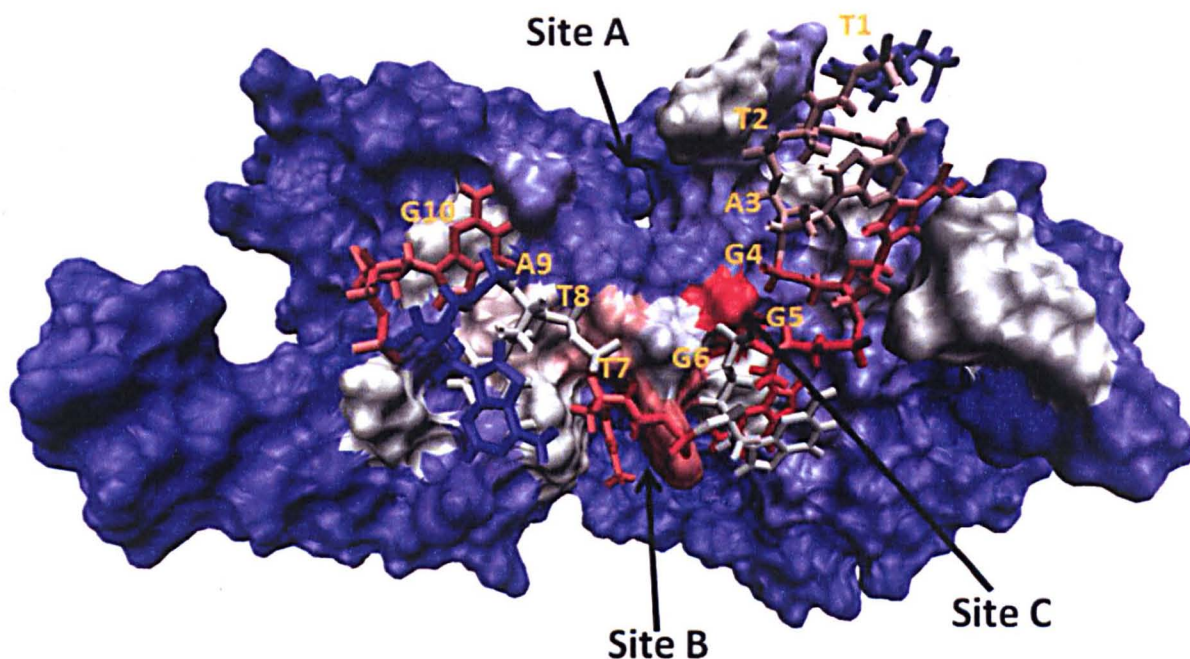


Figure 17: Hot spot map showing total energy (sum of van der Waals and electrostatic energies) between POT1 (CPK surface) and DNA (CPK sticks). The residues predicted to be hot spots are shown in red which represent regions where interaction between POT1 and DNA are strong in terms of total energy. White to blue regions represent moderate to weak interaction between POT1 and DNA.

The hot spot map of POT1 identified two potential drug binding sites (red regions in Figure 17: [site B and C]) where interaction between POT1 and DNA are strongest. Site A, the solvent accessible pocket which was previously found using SYBYL[®] is not identified as a hot spot and makes a very weak interaction with DNA since it does not lie close to the DNA binding surface. By examining the hot spot map of POT1, it can be seen that certain residues in the POT1-DNA complex contribute more to the binding affinity than others. The interaction energies for the top residues in the POT1-DNA complex are listed in Appendix A.

Four DNA nucleotides (G4, G5, T7 and G10) appear to make particularly strong interactions with POT1. However it can be seen that the amino acids of POT1 do not make an equally strong interaction with G4 and G10 bases of DNA in these regions. The POT1 amino acids: LYS-28, TYR-266 and ARG-268 are the major contributors to the total energy of the complex. Two of these residues, TYR-266 and ARG-268 are found at Site B which also corresponds to the largest solvent accessible pocket previously identified using SYBYL[®]. Table 2 lists the amino acids which are found at the 3 different binding sites.

| Site A | Site B | Site C |
|---------------|---------------|---------------|
| TYR-43 | PHE-27 | LYS-28 |
| TYR-66 | ASN-70 | TYR-31 |
| HIS-228 | TYR-156 | THR-264 |
| SER-232 | ASP-158 | |
| | SER-238 | |
| | TYR-266 | |
| | ARG-268 | |

Table 2: List of amino acids which are found at the three binding sites.

Site B contains a deep pocket which has the potential to accommodate a reasonable sized ligand. In the original crystal structure, the T7 base of DNA binds at site B where the T7 base is sandwiched in a stack of aromatic side chains (Figure 18). The major interaction at Site C is a strong electrostatic interaction between LYS-28 and

the phosphate of guanosine-6. Site C is flat and does not contain any solvent accessible cleft thereby making it difficult to design small molecules to target this specific region.

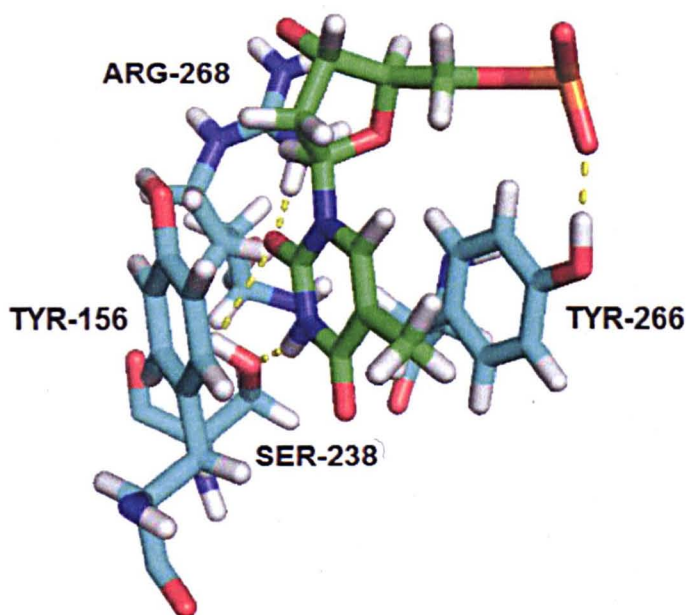


Figure 18: Detailed interaction between T7 base of DNA (green CPK sticks) and its interacting protein residues (cyan CPK sticks). H-bonds are depicted as yellow dotted lines.

To determine which of the 3 binding sites was more energetically favourable, the computational tool GRID (Goodford, 1985) was used. The algorithm uses small probe groups which may be an integral part of a larger ligand i.e. a drug with distinct chemical entities such as amine nitrogen, carbonyl oxygen and many others which move through a regular grid of points around the target POT1 protein (Wade *et al*,

1993). At each point around the grid, the interaction energy between the probe group and POT1 is calculated using an empirical energy function (eq. 6).

$$E_{\text{(Energy)}} = \sum E_{\text{vanderWaals}} + \sum E_{\text{electrostatic}} + \sum E_{\text{hydrogen-bond}} \quad (\text{eq.6})$$

The energies calculated at various positions around POT1 may then be displayed by computer graphics as three dimensional contours (Figure 19). Energetically favourable binding sites can be found by examining contours at large negative energies which correspond to regions of attraction between a particular probe and POT1. Contours at large positive energies indicate regions of repulsion between probe and POT1 (Wade *et al*, 1993).

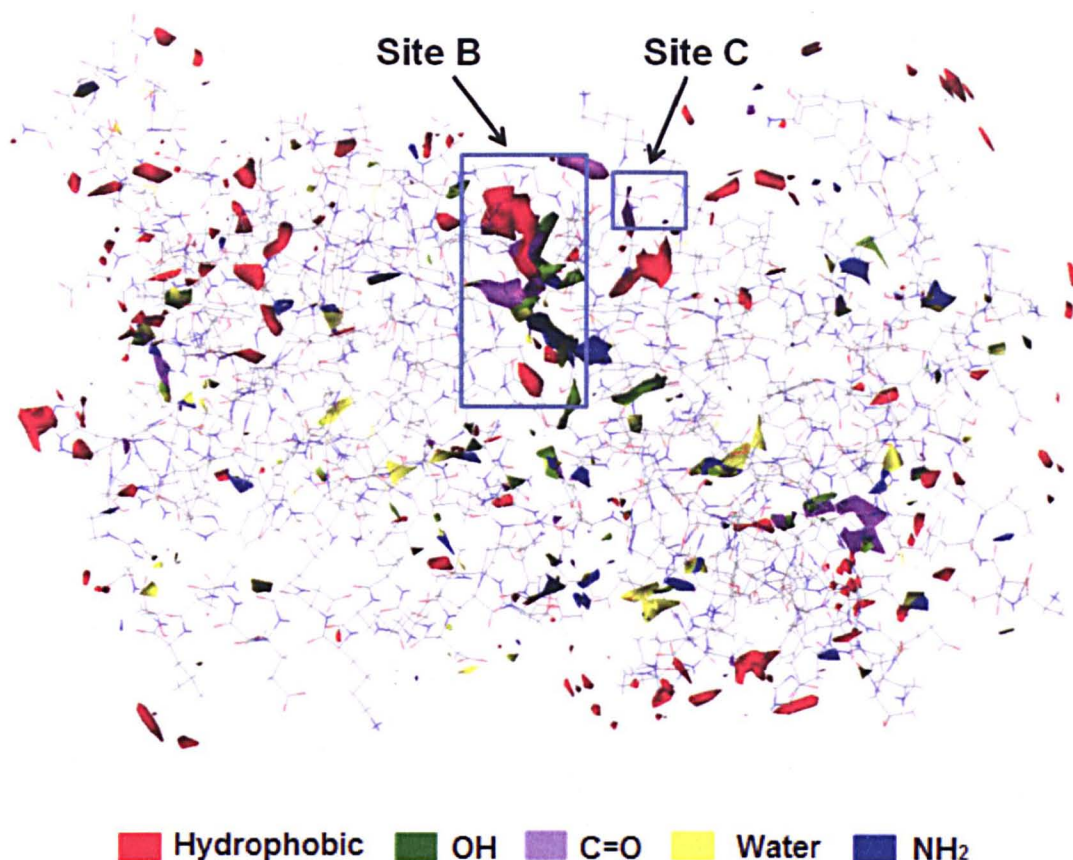


Figure 19: GRID contour map of POT1. Several probe groups were used: hydrophobic, hydroxyl (OH), carbonyl oxygen (C=O), water (H₂O) and neutral amino (NH₂). The contours are shown between -10 Kcal/mol and -13 Kcal/mol.

By analysing the contour map of POT1 (Figure 19), it can be seen that the originally-identified hotspot Site B is particularly rich in sites that have the possibility of favourable interactions with a range of functional groups. Site A (not shown in the figure since it lies slightly on the opposite side to Site B) and Site C however do not seem to make favourable interactions with the different probe groups.

Visual examination of site B highlights areas which are not yet exploited by the T7 base of DNA. Ligands which can target this region have the potential to be extended deeper within the pocket to explore the unoccupied space and to pick up additional

interactions which the T7 base of DNA cannot. The hydrophobic probe identifies a very strong hydrophobic interaction between TYR-156 and TYR-266. Strong hydrogen bonding and electrostatic interactions are also observed between hydroxyl and amino probes and ASP-158 and ASN-70. A strong H-bond interaction is also observed between the carbonyl probe and TYR-156 and TYR-266. However limited interaction with the water probe is observed at site B indicating that the site is quite hydrophobic and prefers to bind to hydrophobic functional groups.

5.1.3 Conclusion

In conclusion, site B contains a deep hydrophobic pocket with a large size and volume and the presence of a rich array of functional groups. It was therefore decided that site B should be chosen as a potential druggable binding pocket in POT1. The important amino acids at site B which may be essential for ligand binding have been identified (Figure 20). The aim is to design a small molecule ligand which can bind at site B and stack in between TYR-156 and TYR-266 thereby prevent the T7 base of DNA from binding. It is suggested that blocking this single nucleotide from binding to POT1 may be sufficient to prevent the entire DNA molecule from binding to POT1.

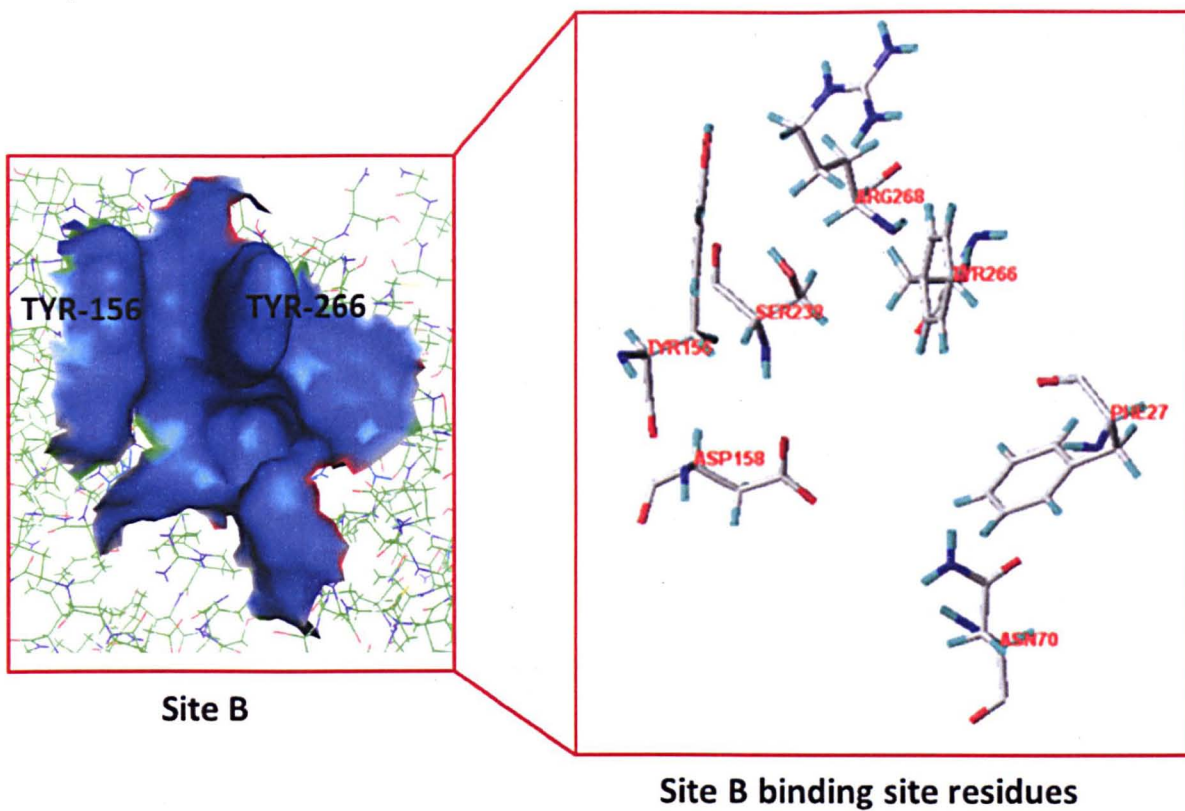


Figure 20: Site B residues (CPK sticks) which may be essential for ligand binding include PHE-27, ASN-70, TYR-156, ASP-158, SER-238, TYR-266 and ARG-268.

Chapter 6: Ligand design and hit identification

6.1 Introduction

By means of *ab initio* drug design and docking, the aim of the project was to design a specific class of POT1 inhibitor through a series of modifications to an initial core fragment. The size of the pocket, lipophilicity and distribution of hydrogen bond donors and acceptors within the pocket were used to design a novel ligand which targets site B (see Figure 20) in POT1. Care was taken so that the designed ligands did not violate Lipinski's rule of five (Lipinski *et al*, 2001) which is a rule of thumb for drug likeness. The rule states that in general an orally active drug has no more than one violation of the following criteria:

- No more than 5 H-bond donors
- No more than 10 H-bond acceptors
- A molecular mass less than 500 daltons
- A log P value of less than 5

6.2 Dock and design of POT1 inhibitor

Several small lipophilic and planar fragments were initially selected manually and designed using the sketch molecule tool in SYBYL[®] (Table 3). The fragments were designed such that they can bind near the surface of the pocket and stack between the π systems of TYR-156 and TYR-266 thereby preventing T7 of DNA from binding and stacking into this region. The majority of the fragments were chosen randomly based on having a planar ring system and/or the presence of some H-bond donor or

acceptor functional groups in order to exploit H-bonding interactions within the binding pocket. Fragment **8a** on the other hand was discovered using the similarity search tool in the Sigma Aldrich compound collection to identify structures similar to thymine. The nine fragments (**1-9**) were subsequently docked into the predicted POT1 binding pocket using GOLD and their binding modes and fitness scores were evaluated (Table 3). The computational methodology of how the ligands and the POT1 structure were prepared prior to docking is presented in section 10.2.


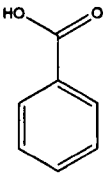
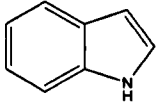
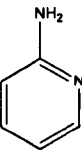
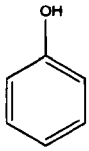
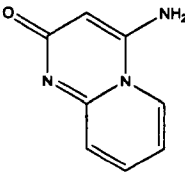
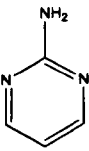
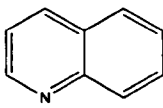
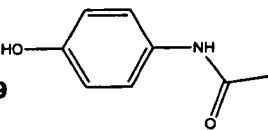
| Fragment | GoldScore | Fragment | GoldScore | Fragment | GoldScore |
|--|-----------|--|-----------|---|-----------|
| 1  | 25.34 | 4  | 33.07 | 7  | 32.54 |
| 2  | 28.50 | 5  | 34.48 | 8a  | 40.58 |
| 3  | 28.43 | 6  | 35.12 | 9  | 38.40 |

Table 3: Structures of several small fragments with their corresponding GoldScores.

The majority of these fragments intercalated in between TYR-156 and TYR-266 and underwent π - π stacking interaction with them. However based on the obtained GoldScores for the nine fragments, fragment **8a** ranked the top and was selected for further modification. Fragment **8a** contains a pyrido[1,2-a]pyrimidin-2-one core with similar functional groups as thymine and therefore acted as a good mimic. The carbonyl group of fragment **8a** participated in H-bond interactions with ARG-268 and SER-238 (Figure 21).

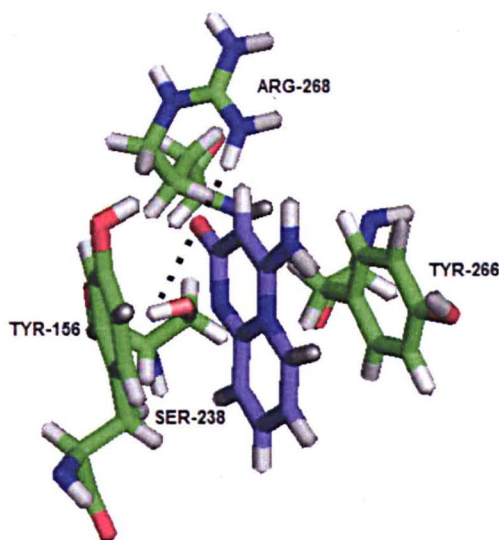


Figure 21: Predicted binding mode of fragment **8a** (blue CPK sticks) within site B. The protein side chains are displayed as green CPK sticks. H-bonds are depicted as black dotted lines.

To explore areas deeper within the cleft where thymidine-7 could not bind and to exploit additional interactions within the binding pocket, a series of structural modifications (shown in Figure 22) were carried out on fragment **8a**. The first task was to increase the binding affinity of fragment **8a** for the predicted pocket. The lipophilic nature of fragment **8a** was augmented by fusing an additional aromatic ring to it with the intention of increasing the hydrophobic interaction with the binding pocket. The redesigned fragment **10** was redocked into the POT1 binding pocket which resulted in a higher GoldScore of 46.28.

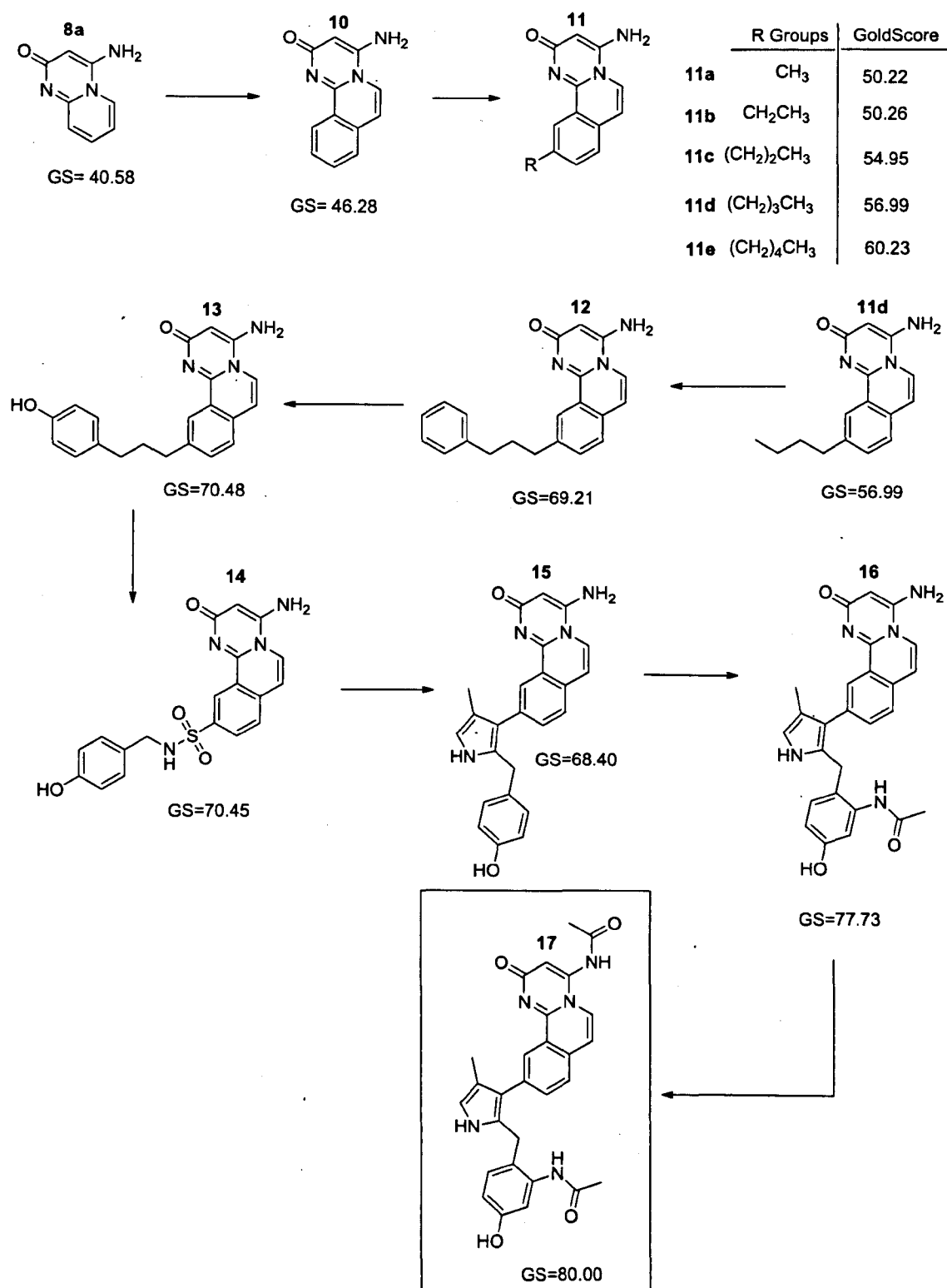


Figure 22: A simplified diagram showing the series of modifications (highlighted in blue) carried out on fragment **8a** to produce ligand **17** (shown in the red box) with a high predicted affinity for the POT1 pocket.

The binding of fragment **10** was further enhanced by incorporating a carbon alkyl chain to the fragment. It was anticipated that alkyl chains of different lengths would provide a means for the ligand to dig deeper within the hydrophobic pocket exploiting lipophilic interactions with the surrounding residues. This was indeed observed and the GoldScores for fragments **11a-e** increased with increasing carbon chain length. A butyl chain was identified as the ideal length resulting in fragment **11d**. Although additional CH₂ groups to the alkyl chain via the design of a homologous series increased the GoldScore further. It was rationalised that sufficient space should be left for the positioning of a phenyl ring to fill unoccupied space within the current pocket, thereby resulting in fragment **12** whose GoldScore increased to 69.21.

To investigate H-bonding interactions between ligand and binding pocket. A hydroxyl group was incorporated at the para position of the pendent phenyl ring in order to interact with ASN-70 and PHE-27 thereby giving fragment **13** with a GoldScore of 70.48 when docked into the protein. To optimise the H-bonding interaction with ASP-158, the propyl chain in fragment **13** was replaced by a sulphonamide group to give fragment **14** where the NH group of the sulphonamide was designed to exploit a H-bond with ASP-158. Since the oxygens of the sulphonamide group did not pick up any H-bonding interactions, the sulphonamide group was replaced by a heterocyclic methyl pyrrole (Fragment **15**, GoldScore= 68.40). This moiety occupied space within the hydrophobic pocket whilst still retaining the H-bonding interaction with ASP-158. An additional H-bond interaction with ASN-70 was also considered. An amide group was incorporated at the ortho position of the phenol ring resulting in fragment **16** with GoldScore of 77.73.

The 4-amino group of the pyrido[1,2-a]pyrimidin-2-one core did not pick up any interactions with the binding site residues pointing outwards towards the solvent and away from the binding pocket. Pleasingly, replacement of the amino group with an amide resulted in an H-bond interaction with the hydroxyl group of TYR-156. Hence through an iterative process of ligand modification and GOLD-based redocking and scoring, ligand **17** was obtained which had a high predicted affinity for the POT1 pocket.

The binding mode of **17** (Figure 23) indicates that the structure is capable of forming 7 hydrogen bonds with the binding site residues at site B. The carbonyl of the pyrimidine core forms 2 H-bonds with ARG-268 and SER-238. The carbonyl of the 4-acetamide group on the pyrimidine formed H-bond with the hydroxyl group of TYR-156. The NH group of pyrrole formed H-bond with ASP-158, while the hydroxyl group of the phenol participates in a bidentate interaction with the carbonyl group of ASN-70 and NH group of PHE-27. The acetamide group of the phenol also shares a H-bond with the amino group of ASN-70.

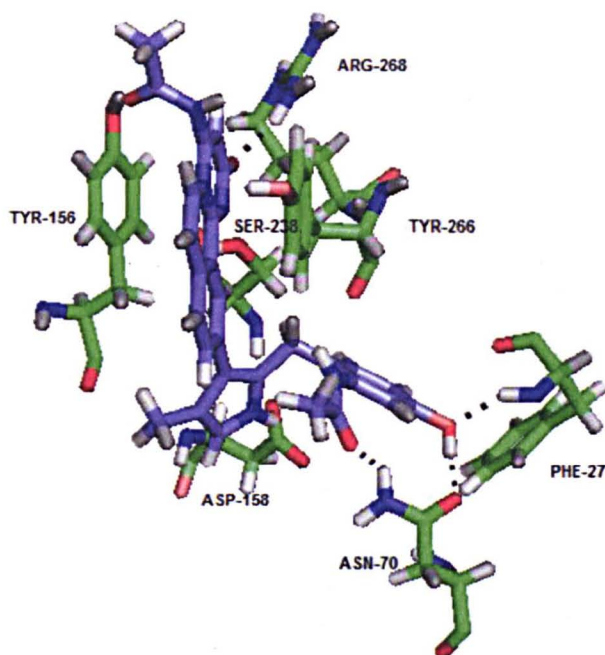


Figure 23: Binding mode of ligand **17** (blue CPK sticks). The protein side chains are displayed as green CPK sticks. H-bonds are depicted as black dotted lines.

6.3 Synthetic accessibility dilemma

There is limited literature information regarding the synthesis of the pyrimido[2,1-*a*]isoquinoline acetamide structure of ligand **17** highlighted in blue (Figure 24a). In addition, retrosynthetic analysis suggests that there could be significant issues of difficulty, time and cost involved in the synthesis of **17**. Therefore a new ligand structure **18** (represented in green, Figure 24b) was proposed where the phenyl pyrido[1,2-*a*]pyrimidine acetamide core could readily be synthesized by a procedure reported by Dorokhov *et al* (1990). Moreover, the new structure contained as many elements of the originally designed scaffold.

Ligand **18** was docked into the POT1 pocket and found to mimic the binding mode of the tricyclic pyrimidine ring of **17** (Figure 24c). However a slight twist in the C-C bond between the pyrimidine core and the pendent phenyl ring of **18** was observed. An emphasis was placed on the novelty of the new designed chemotype which was used to generate a focussed library of synthetically feasible ligands of POT1 that have never been described in patents or in the published literature.

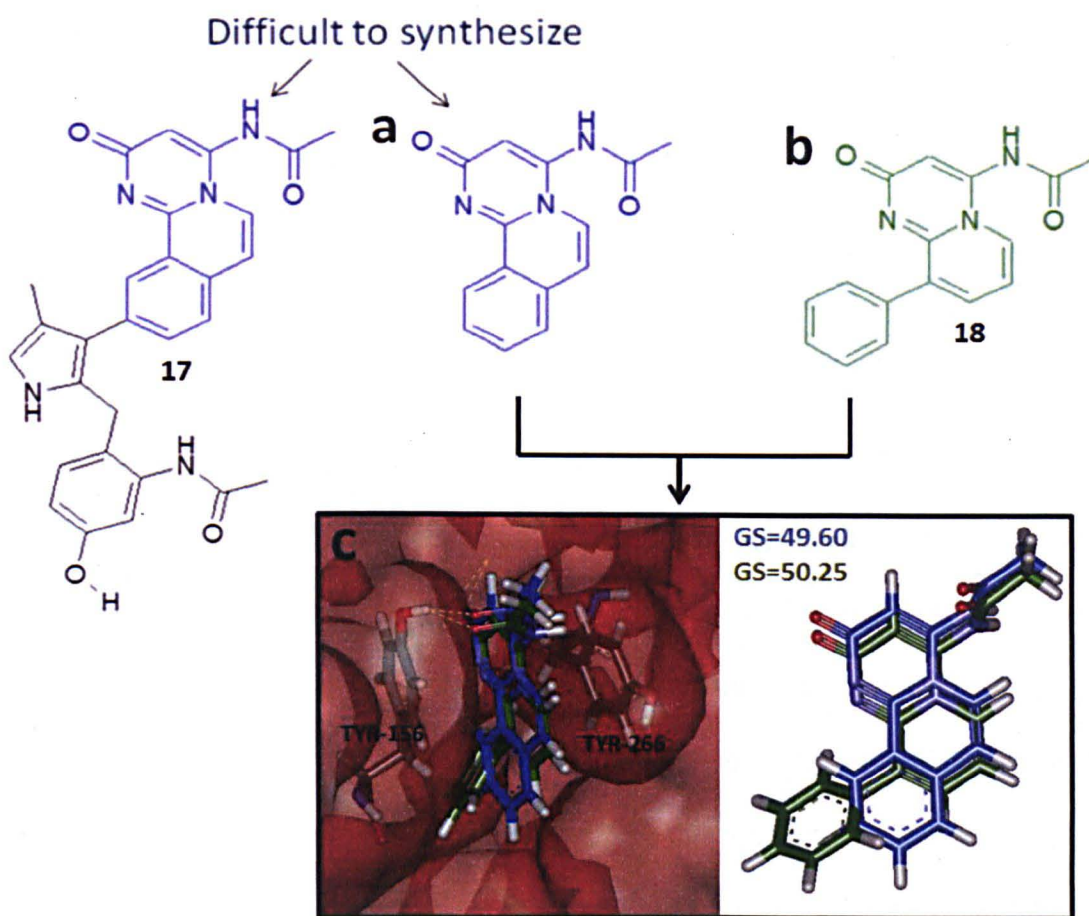
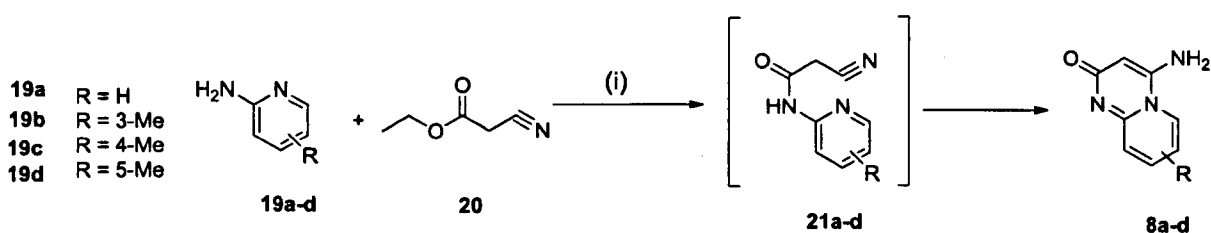


Figure 24: (a) Pyrimido[2,1-a]isoquinoline acetamide structure of ligand **17**. (b) Newly designed structure **18**. (c) Superimposed docked binding modes of **17** and **18** with their corresponding GoldScores. H-bonds are depicted as yellow dotted lines.

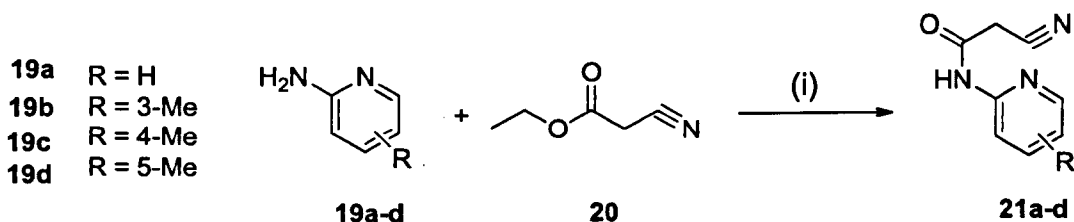
6.4 Synthetic route for pyrimidine analogues

The synthesis of pyrido[1,2-*a*]pyrimidine derivatives have been reported by Dorokhov *et al* (1990). They can be directly obtained by reacting various 2-aminopyridine derivatives (**19a-d**) with ethyl cyanoacetate (**20**) at high pressure in the presence of benzene resulting in derivatives of 4-amino-2*H*-pyrido[1,2-*a*]pyrimidines (**8a-d**) in yields of 40-85% (Scheme 1).



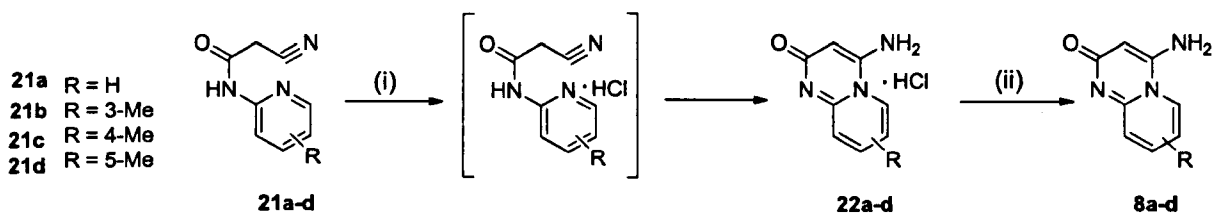
Scheme 1: Synthetic scheme for the synthesis of **8a-d**. Reagents and conditions (i) C_6H_6 , 80-100°C, 14kbar.

It is suggested that cyanoacetamides (**21a-d**) are intermediately formed and further cyclise into (**8a-d**). Initial attempts to synthesize **8a** using similar conditions with 2-aminopyridine (**19a**) as the starting reagent resulted in failure. The synthesis of cyanoacetamides (**21a-d**) can also be carried out by reacting derivatives of 2-amino pyridine (**19a-d**) with ethyl cyanoacetate (**20**) in the absence of high pressures as reported by Dorokhov *et al* (1990). The reaction is carried out at 135-165°C with the simultaneous distillation of ethanol resulting in cyanoacetamides (**21a-d**) in 50% yield (Scheme 2).



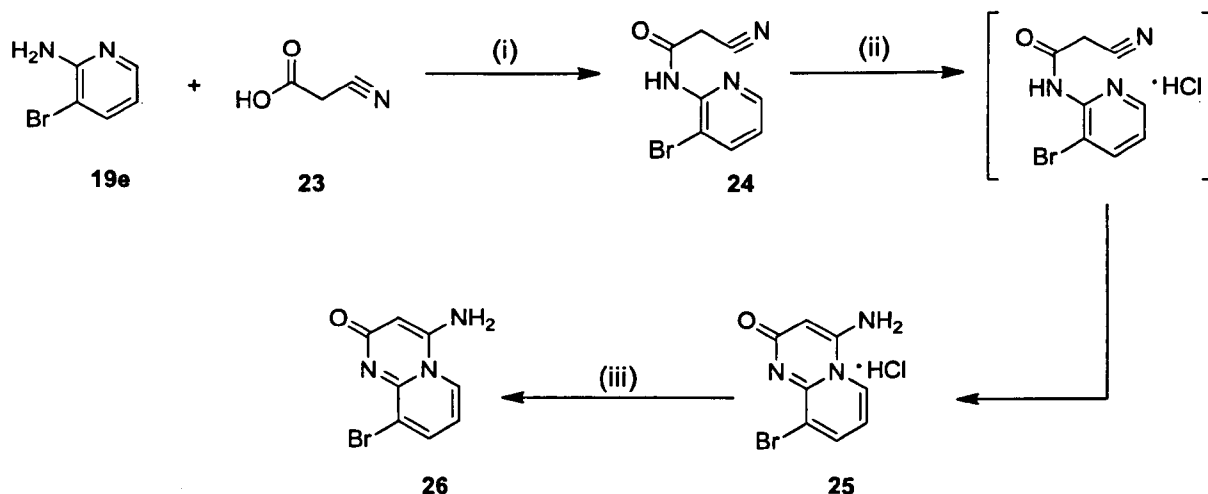
*Scheme 2: Synthetic scheme for the synthesis of **21a-d** derivatives in the absence of high pressures. Reagents and conditions (i) 135-165°C.*

The cyanoacetamides (**21a-d**) then undergo treatment with a solution of hydrogen chloride in ethanol under high temperatures to form the hydrochlorides (**22a-d**) which subsequently convert into free bases (**8a-d**) by the action of aqueous sodium hydrogen carbonate solution (Scheme 3).



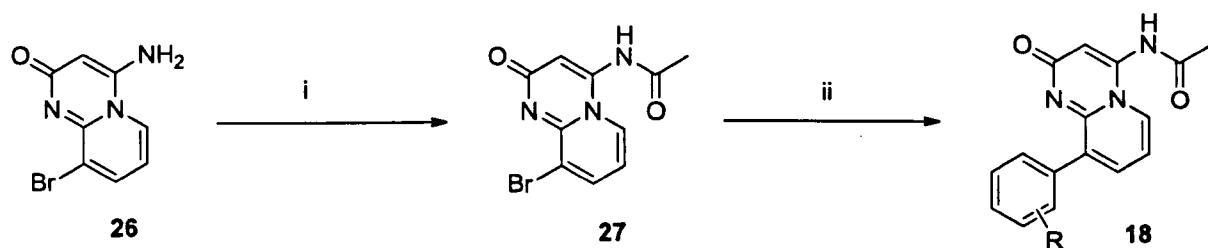
*Scheme 3: Synthetic scheme for the synthesis of **8a-d** derivatives. Reagents and conditions (i) HCl in EtOH, 110°C, 1h (ii) NaHCO₃(aq).*

Initial attempt to synthesize **8a** in the absence of high pressures was successful, producing **8a** in 16% yield. The reaction was subsequently repeated using 2-amino-3-bromo pyridine instead of 2-aminopyridine (**19a**) with ethyl cyanoacetate (**20**) to enable the synthesis of derivatives of **8**. The resulting reaction failed and the reaction scheme was subsequently modified. Ethyl cyanoacetate (**20**) was replaced by cyanoacetic acid (**23**) and the reaction was carried out in the presence of dicyclohexylcarbodiimide (DCC) coupling agent to give 4-amino-9-bromo-2H-pyrido[1,2-a]pyrimidin-2-one (**26**) in 60% yield (Scheme 4).



*Scheme 4: Synthetic scheme for the synthesis of **26**. Reagents and conditions (i) DCC, DCM, reflux, 2h (ii) 1.25M HCl in EtOH, reflux, 2h (iii) NaHCO₃(aq).*

To produce derivatives of **18**, acetylation of the 4-amino terminus of **26** was carried out using acetic anhydride in the presence of anhydrous pyridine to give **27** in 78% yield (Scheme 5). Compound **27** then underwent a Suzuki coupling (Miyaura and Suzuki, 1995) carbon-carbon bond forming reaction with a range of boronic acids to give products of **18** in yields between 8-50%.



*Scheme 5: Synthetic scheme for the synthesis of derivatives of **18**. Reaction and conditions (i) Ac₂O, C₆H₅N, 110°C, 2h (ii) Pd(PPh₃)₄, Cs₂CO₃, 1,4-Dioxane:DMA (10:1), boronic acids, 160°C, 10min, 300w in microwave .*

6.5 Screening results and discussion

To determine whether the pyrido[1,2-*a*]pyrimidin-2-one scaffold had the potential to inhibit the interaction between POT1 and its DNA binding partner, four fragments were initially synthesised and tested in the POT1 FP assay (Figure 25). The FP results are presented in Figure 26.

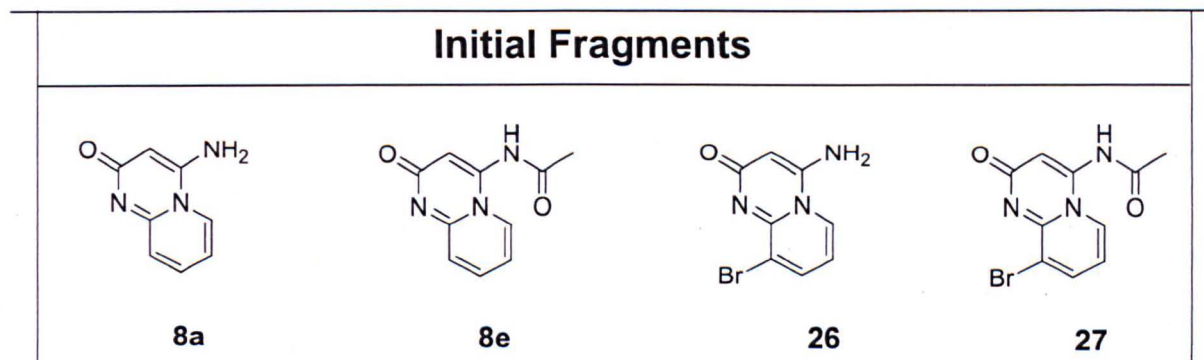


Figure 25: Chemical structures of four pyrido[1,2-*a*]pyrimidin-2-one fragments.

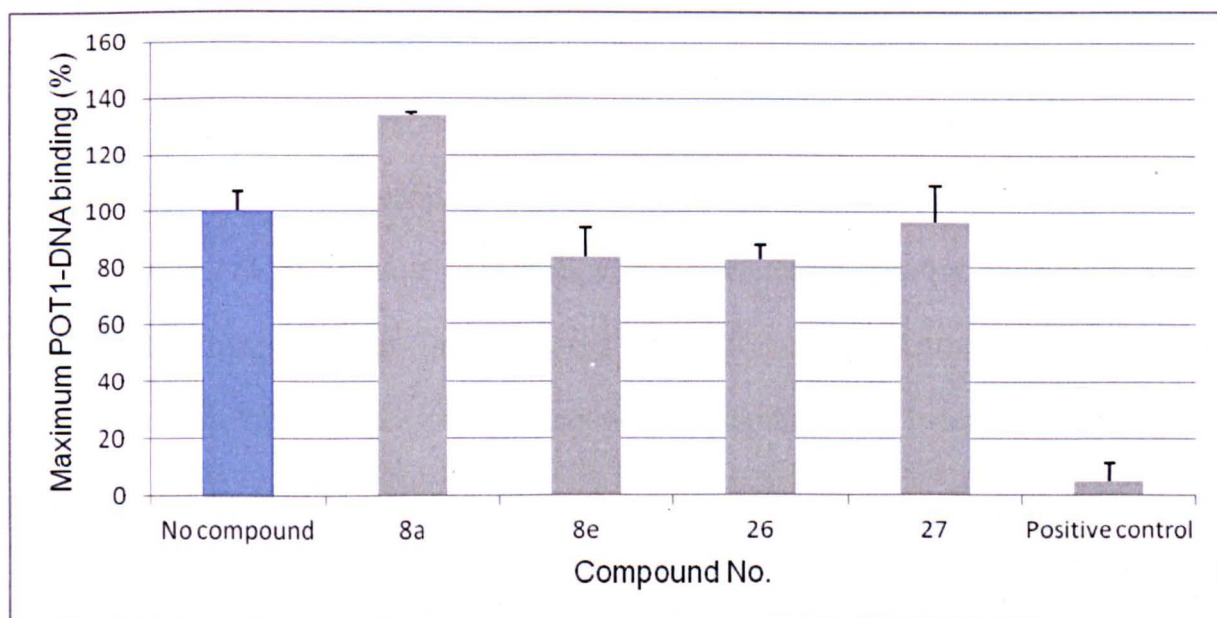
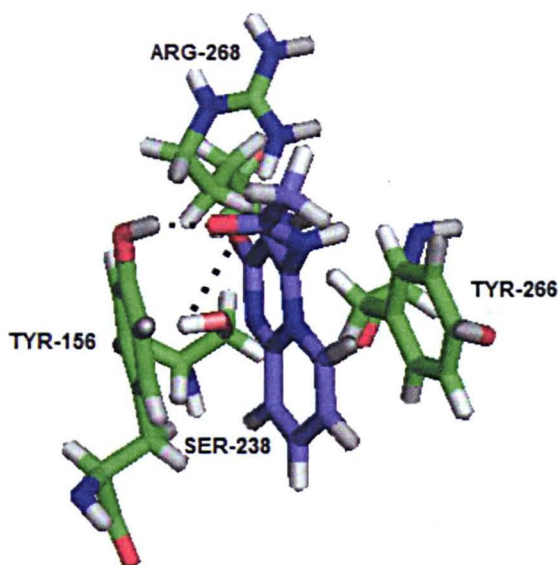


Figure 26: FP results for the four initial fragments. Percentage remaining signal at 100 μ M compound relative to no compound signal (mean \pm standard error) of 3 determinations.

All compounds herein were initially tested at a single dose of 100 μ M in triplicates to detect even the weakest inhibitory activity unless otherwise stated. In Figure 26, no compound is given the value of 100% which represents maximum POT1-DNA binding in the absence of ligand.

Potential inhibitors of POT1 will tend to decrease the binding of DNA to POT1 since they will compete with the DNA for the POT1 binding pocket. Therefore these potential inhibitors will reduce the maximum POT1-DNA signal. A positive control was used in the assay which is a non-fluorescently labelled oligonucleotide with the same sequence 'TTAGGGTTAGGGTTA' as the fluorescently labelled oligonucleotide. When the non-fluorescent DNA binds to POT1, it tends to decrease the maximum POT1-DNA binding signal back down to the base level at 100 μ M (IC_{50} of competitive non-fluorescent DNA = 200 ± 15 nM).

The first round of screening identified two hits: **8e** and **26** (Figure 26) which showed some weak activity in disrupting POT1-DNA interaction. Compound **8a** with the 4-amino group was found to be inactive whereas its derivative **8e** with a 4-amide group decreased POT1-DNA binding by 16% at 100 μ M ligand concentration. The predicted binding modes of **8e** at site B suggests that the fragment has the ability to stack in between the π -systems of TYR-156 and TYR-266 and the carbonyl group of the amide may be important for binding since it participates in a H-bond with TYR-156 (Figure 27).

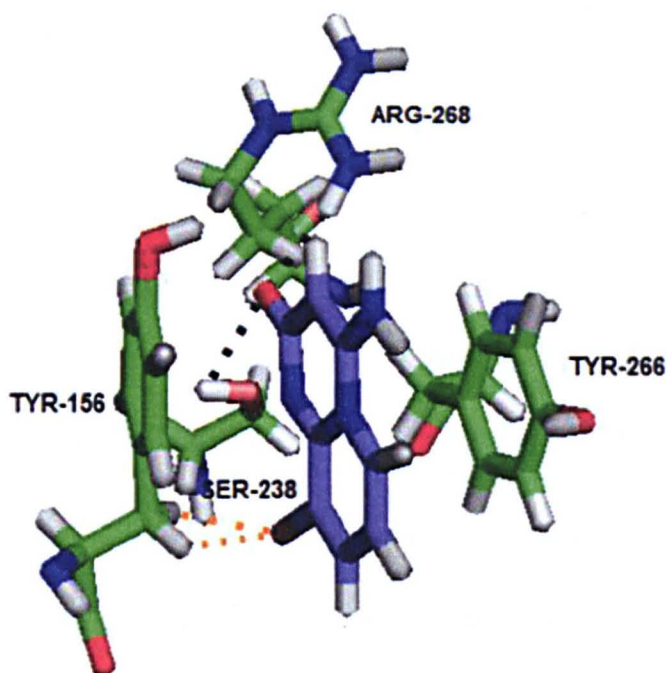


*Figure 27: Predicted binding mode of **8e** (blue CPK sticks). The protein side chains are displayed as green CPK sticks. H-bonds are depicted as black dotted lines.*

The limited activity which is observed for fragment **8a** may be due to the structure not being able to form the extra H-bond with the hydroxyl group of TYR-156. It is proposed that this extra bonding interaction may be necessary for enhanced inhibitory activity. However it is currently unclear why **8a** is showing more than 100% maximum POT1-DNA binding. It is suggested that ligands which increase the signal beyond 100% POT1-DNA binding could (1) cause light scattering and as a result the scattered light is 100% polarised, (2) the ligand is potentially interacting with the fluorescein tag on the DNA and therefore quenching the fluorescence of the tag, (3) the ligand is binding to an allosteric site on POT1 thereby causing a conformational change which results in a more stable POT1-DNA complex.

It was anticipated that fragment **27** would show better activity in comparison to fragment **26** since the structure contains an amide functional group which could potentially exploit the additional H-bond interaction with TYR-156. However **27** was

found to be inactive in the assay. Compound **26** on the other hand displayed some inhibitory activity against POT1 and decreased POT1-DNA binding by 17%. The docked binding mode of **26** is presented in Figure 28. It can be seen that the amino group of **26** is not involved in any interaction within the binding site residues. The structure forms 2 H-bonds with SER-238 and ARG-268 and the bromine atom participates in vdW interactions with TYR-156.



*Figure 28: Binding mode of **26** (blue CPK sticks). The protein side chains are displayed as green CPK sticks. H-bonds are depicted as black dotted lines and vdW contacts are displayed as orange dotted lines.*

The results suggest that fragments **8e** and **26** have some marginal activity in disrupting POT1-DNA interactions and since the fragments are small, they can potentially be modified to improve inhibitory activity. Ligand **18** (Table 4) with a pendent unsubstituted phenyl ring was next synthesized to further investigate the

scaffold as a potential POT1 inhibitor. Ligand **18** was found to be inactive in the assay. It was suggested that the pendent phenyl ring may not be picking up any interaction within the binding pocket. In order to improve the affinity of **18** for POT1, various substituents were positioned around the pendent phenyl ring. This was done to determine whether such modifications can enhance the structure's activity and participate in additional interactions with the binding site residues.

When the synthesis of a large number of compounds is difficult and when biological testing is readily available, the optimum substitution on the pendent phenyl ring of ligand **18** can be found using the Topliss method (Topliss, 1977). The Topliss method has been successfully used for the efficient optimisation of the potency of compounds with the minimization of the number of compounds that need to be synthesized. The only pre-requirement to use this technique is that the compound must contain an unfused benzene ring.

The primary function of the Topliss scheme is to synthesize a series of analogues where various substituent groups are positioned around the unfused benzene ring. The substituents differ in their physicochemical properties. Commonly used property descriptors include the substituent hydrophobicity constant (π) which measures the hydrophobicity of a substituent relative to hydrogen, the Hammett substituent constant (σ) which measures the electron withdrawing and donating properties of a substituent and the Taft's steric factor (E_s) which measures the steric effects of a substituent. In order to maximise the chances of finding the most potent compounds in the series as early as possible, the Topliss scheme was implemented and 16 additional ligands were synthesized and screened in the FP assay (Table 4).



| Compound NO. | R ₂ | R ₃ | R ₄ | Maximum POT1-DNA binding (%) |
|--------------|----------------|-----------------|--|------------------------------|
| 18 | H | H | H | 108±5 |
| 18b | H | H | OMe | 35±4 |
| 18c | H | H | Cl | 81±6 |
| 18d | H | NH ₂ | H | 86±3 |
| 18e | H | H | CON(CH ₂ CH ₃) ₂ | 72±2 |
| 18f | H | Cl | Cl | 59±3 |
| 18g | H | H | NH ₂ | 93±9 |
| 18h | H | H | COOH | 100±14 |
| 18i | H | Cl | H | 85±10 |
| 18j | H | OMe | OMe | 94±1 |
| 18k | H | H | CONH ₂ | 108±18 |
| 18l | H | H | CH ₃ | 119±5 |
| 18m | H | H | F | 46±9 |
| 18n | OMe | H | H | 132±7 |
| 18o | H | OMe | H | 119±8 |
| 18p | Cl | H | Cl | 103±3 |
| 18q | H | COOH | H | 101±3 |

Table 4: FP results for derivatives of **18**. Percentage remaining signal at 100μM compound relative to no compound signal (mean ± standard error) of 3 determinations. The top structures **18m** and **18e** were screened twice in the assay and their mean calculated. Compounds highlighted in the red box are marked as false positives.

Examining the data in Table 4 suggests that the substitution pattern on the phenyl ring is important for optimum binding. The majority of ligands (**18d**, **18g-l** & **18n-q**) showed very little or no effect in disrupting POT1-DNA interaction. Ligands **18b** and **18f** displayed some moderate inhibitory activity but were identified as false positives since their fluorescence characteristics tended to interfere with the assay.

The replacement of the hydrogen substituent (**18**) by a chlorine substituent at the 4-position of the phenyl ring as in **18c** decreased maximum POT1-DNA binding by 19%. The chlorine substituent is more lipophilic ($+\pi$) and electron withdrawing ($+\sigma$) than hydrogen. Therefore the activity of the substituent can be attributed to either a $+\pi$ effect, a $+\sigma$ effect or a combination of both. However when the 3,4-dichloro compound (**18f**) was tested which has a larger $+\pi$ and $+\sigma$ effect, the ligand had no effect on POT1-DNA binding and tended to interfere with the assay readout. The results suggest that further increase in either $+\pi$ or $+\sigma$ is unfavourable.

However the replacement of chlorine with a more electronegative fluorine substituent (**18m**) which has a smaller $+\pi$ effect than chlorine increased activity more than 2-fold and decreased POT1-DNA binding by 54% at 100 μ M ligand concentration. It is suggested that the electron withdrawing properties of these halogens substituents may be driving the activity of the ligands. Subsequent modelling studies suggest that the halogen atoms in both **18c** and **18m** interact with the binding site backbone fragment (H-C $_{\alpha}$ -C=O) of PHE-157. The chlorine and fluorine substituents, both form a short orthogonal multipolar C-X \cdots C=O and C-X \cdots H-C $_{\alpha}$ where (X= Cl or F) interactions with PHE-157 within the predicted pocket. The predicted binding mode of **18m** is presented in Figure 29.

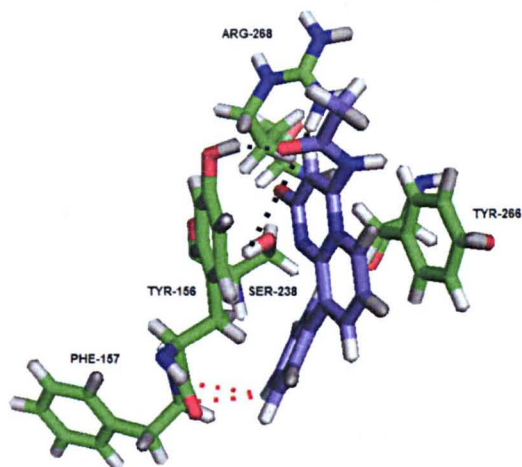


Figure 29: Binding mode of **18m** (blue CPK sticks). The protein side chains are displayed as green CPK sticks. H-bonds are depicted as black dotted lines and multipolar interactions are displayed as red dotted lines.

It is suggested that since fluorine is more electronegative than chlorine, ligand (**18m**) which contains the fluorine substituent is more potent since it forms a stronger orthogonal multipolar interaction with PHE-157 backbone.

Orthogonal multipolar interactions have recently gained attention and have been found in numerous protein-ligand PDB complexes. Multiple SAR examples have shown that the presence of such interactions can dramatically improve binding affinity (Müller *et al*, 2007). However it cannot be assumed that the increase in binding affinity could be ascribed to the interaction by fluorine alone. Large components may also be due to changes in residual mobility and desolvation (Bissantz *et al*, 2010).

Ligand **18k** with an amide substituent positioned at the 4-position of the phenyl ring resulted in no activity. Replacement of the amide group by a diethylacetamide group as in ligand **18e** decreased POT1-DNA binding by 28%. These results suggest that

the lipophilicity of the diethyl chains in **18e** may be driving the activity of the structure. The predicted binding mode of **18e** is presented in Figure 30. The diethylacetamide group in **18e** participates in a H-bond with the backbone of ASP-158 and a multipolar interaction with the backbone of PHE-157. Furthermore the two ethyl chains in **18e** extend outwards and deeper within the hydrophobic pocket.

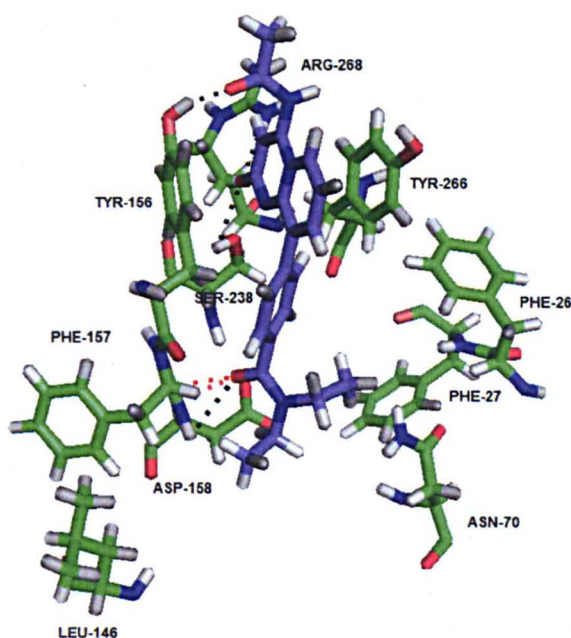


Figure 30: Binding mode of **18e** (blue CPK sticks). The protein side chains are displayed as green CPK sticks. H-bonds are depicted as black dotted lines and multipolar interactions are displayed as red dotted lines.

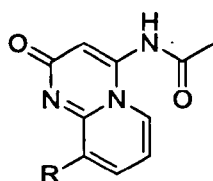
Ligands which contained the acid, amino and amide substituents such as those which are found in **18d**, **18g**, **18h**, **18k** and **18q** were synthesised to exploit H-bonds and electrostatic interactions within the binding pocket. The docked binding modes of these ligands suggested that their substituents are capable of picking up H-bonds and electrostatic interactions with ASP-158 within the binding site. However they did

seem to produce the expected increase in POT1-DNA inhibition despite binding in the predicted manner. It is proposed that due to large desolvation penalties incurred by polar and charged groups on binding in aqueous solution (Freire, 2008), the interaction between such ligands and POT1 may not be favourable and strong enough to compensate for the unfavourable binding enthalpy associated with desolvation. Therefore these ligands do not necessarily contribute to binding affinity.

Ligands with hydrophobic substituents such as the 4-methyl group in **18l** as well as the bulky methoxy substituents in **18j**, **18n** and **18o** were found to be inactive. The predicted binding modes of these structures suggest that these ligands tend to bind deeper within the hydrophobic pocket away from TYR-156 and TYR-266. It is thought that such structures may be act more like POT1-DNA stabilizers rather than destabilizers since some of them show more than 100% POT1-DNA binding. By binding deeper within site B, these structures may be causing a conformational change which allows the DNA to bind more strongly to POT1.

An attempt was made to obtain dose-response data in order to predict IC_{50} values for the two potent compounds: **18e** and **18m**. Both compounds were tested at 5 different concentrations of 100 μ M, 10 μ M, 1 μ M, 0.1 μ M and 0.01 μ M. Unfortunately no dose-dependent behaviour was observed over this concentration range. It was therefore decided that further compounds should be synthesized to identify ligands with improved percentage inhibition at 100 μ M ligand concentration.

To investigate whether the pendent phenyl ring of **18** was necessary for binding, two additional ligands were synthesized in which the pendent phenyl ring was replaced by a pyridine moiety and a pyrazole heterocyclic ring (Table 5).



| Compound No. | R | Maximum POT1-DNA binding (%) |
|--------------|---|---------------------------------|
| 28 | | 121±3 |
| 29 | | 82±4 |

Table 5: FP results for heterocyclic modifications. Percentage remaining signal at 100μM compound relative to no compound signal (mean ± standard error) of 3 determinations.

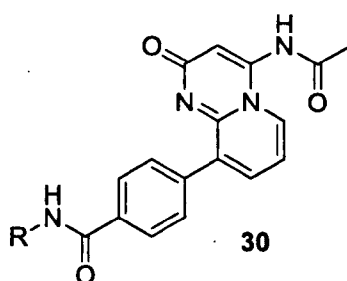
The replacement of the unsubstituted phenyl ring in **18** with pyridine (**28**) had no effect on POT1-DNA binding. However replacement with a pyrazole heterocyclic structure (**29**) improved POT1-DNA inhibitory activity whereby the ligand decreased maximum POT1-DNA binding by 18% at 100μM ligand concentration. The docked binding modes of these compounds suggest that the pyridine ring in **28** does not participate in any interaction with the binding site residues. However the hydrogen atom of the pyrazole ring participates in a short vdW interaction with the methylene group of TYR-156.

6.6 Structural modifications of hit compounds

Since modifications on the pendent phenyl ring were more feasible and since ligands **18e** and **18m** achieved the highest inhibitory activity, they were chosen as initial hits for further optimisation.

6.6.1 Modifications of compound **18e**

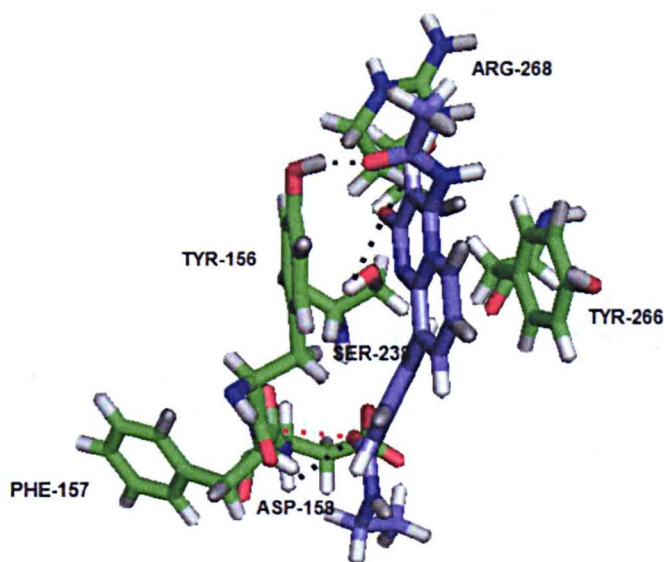
In order to explore ligand **18e** further, the diethyl chains in **18e** were replaced by a single ethyl chain as in **30a** (Table 6).



| Compound NO. | R | Maximum POT1-DNA binding (%) |
|--------------|---|------------------------------|
| 30a | CH ₂ CH ₃ | 71±2 |
| 30b | (CH ₂) ₃ CH ₃ | 82±13 |
| 30c | (CH ₂) ₅ CH ₃ | 85±5 |
| 30d | (CH ₂) ₆ CH ₃ | 71±3 |
| 30e | (CH ₂) ₇ CH ₃ | 91±5 |
| 30f | (CH ₂) ₈ CH ₃ | 88±8 |
| 30g | (CH ₂) ₉ CH ₃ | 94±7 |

Table 6: FP results for analogues of ligand **30**. Percentage remaining signal at 100µM compound relative to no compound signal (mean ± standard error) of 3 determinations.

Substituting the diethyl chains in **18e** with a single ethyl chain (**30a**) decreased POT1-DNA binding by 29%. Subsequent modelling studies suggest that ligand **30a** binds to the pocket in a similar way to **18e**. The carbonyl of the ethylamide in **30a** participates in a H-bond with the backbone of ASP-158 and a dipole-dipole interaction with the carbonyl carbon of PHE-157 (Figure 31).



*Figure 31: Binding mode of **30a** (blue CPK sticks). The protein side chains are displayed as green CPK sticks. H-bonds are depicted as black dotted lines and the dipole-dipole interaction is displayed as a red dotted line.*

This data indicates that two ethyl chains are probably not necessary for binding and a single ethyl chain is sufficient to retain the same binding activity. A series of analogues (**30b-g**) were synthesized where the ethyl carbon chain was increased incrementally (Table 6). It was anticipated that by increasing the alkyl chain length in a new cycle of synthesis and evaluation, a structure with an optimum chain length and enhanced potency than **18e** would be achieved. This would then allow the placement of various H-bond acceptors and donor substituents at the end of the alkyl

chain to exploit additional H-bonds with the binding pocket and further enhance ligand activity.

The data in Table 6 suggests that incrementally increasing the carbon chain length results in a decrease in ligand inhibitory activity. Ligands (**30b**) with a butyl chain and ligand (**30c**) with a hexyl chain decreased POT1-DNA binding by 18% and 15% respectively. Replacement by a heptyl chain (**30d**) resulted in similar inhibitory activity to **30a** where POT1-DNA binding decreased by 29%. Further increase in the chain length to octyl (**30e**), nonyl (**30f**) and decyl (**30g**) had very little or no effect on POT1-DNA binding.

The predicted binding model suggests that the heptyl chain length in **30d** is optimal within the binding pocket (Figure 32a). It is thought that further increase in the chain length leads to a steric clash within the binding pocket. In support of this, the docking studies suggest that **30e** (Figure 30b) as well as **30f** and **30g** adopt a completely different binding mode within the pocket. The pyrimidine core of these ligands tends to bind deeper within the pocket and away from TYR-156 and TYR-266. Therefore such structures have very little effect in displacing DNA from POT1.

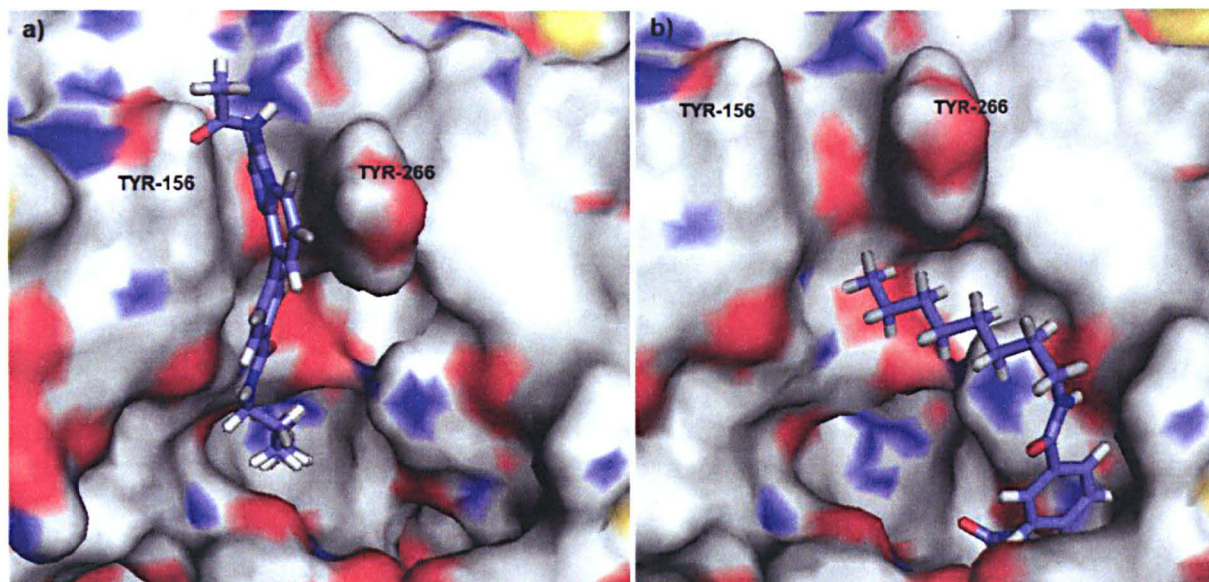
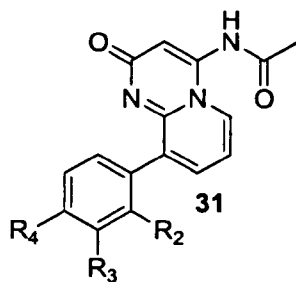


Figure 32: Predicted binding modes of (a) **30d** and (b) **30e**. The protein is displayed as a grey CPK surface and ligands are displayed as blue CPK sticks.

6.6.2 Modification of compound **18m**

To probe the electron withdrawing effect of the fluorine substituent in **18m**, several additional compounds were synthesized where the electronegative fluorine was replaced by a range of different electron-withdrawing substituents. This was done to establish whether the electron-withdrawing property of fluorine was important for binding (Table 7).



| Compound NO. | R ₂ | R ₃ | R ₄ | Maximum POT1-DNA binding (%) |
|--------------|----------------|----------------|---------------------------------|------------------------------|
| 18m | H | H | F | 46±9 |
| 31a | H | H | CF ₃ | 69±8 |
| 31b | H | H | CN | 70±7 |
| 31c | H | H | SO ₂ CH ₃ | 74±18 |
| 31d | H | H | NO ₂ | 81±16 |
| 31e | H | H | OCF ₃ | 68±15 |
| 31f | H | H | COCH ₃ | 79±8 |
| 31g | F | H | H | 65±5 |
| 31h | H | F | H | 117±63 |

Table 7: FP results for derivatives of 31. Percentage remaining signal at 100μM compound relative to no compound signal (mean ± standard error) of 3 determinations.

Electron withdrawing substituents on the phenyl ring generally performed better in decreasing POT1-DNA binding compared to long flexible alkyl chain derivatives of **18e**. The replacement of the 4-fluorine substituent (**18m**) with a more electronegative 4-trifluoromethyl substituent (**31a**) decreased POT1-DNA binding by 31%. The 4-trifluoromethoxy substituent (**31e**) also showed comparable activity to **31a**. Replacement of fluorine by other electron withdrawing groups such as 4-CN, 4-SO₂CH₃, 4-NO₂ and 4-COCH₃ substituents also decreased POT1-DNA binding

between 19-30%. It is proposed that the high inhibitory activity observed for **18m** may be due to the smaller size of the fluorine substituent which allows the ligand to bind better with the PHE-157 backbone as compared to other electron-withdrawing substituents.

Modelling studies suggest that the 4- CF_3 substituent (**31a**) also forms a short orthogonal multipolar interaction with the backbone of PHE-157 (Figure 33).

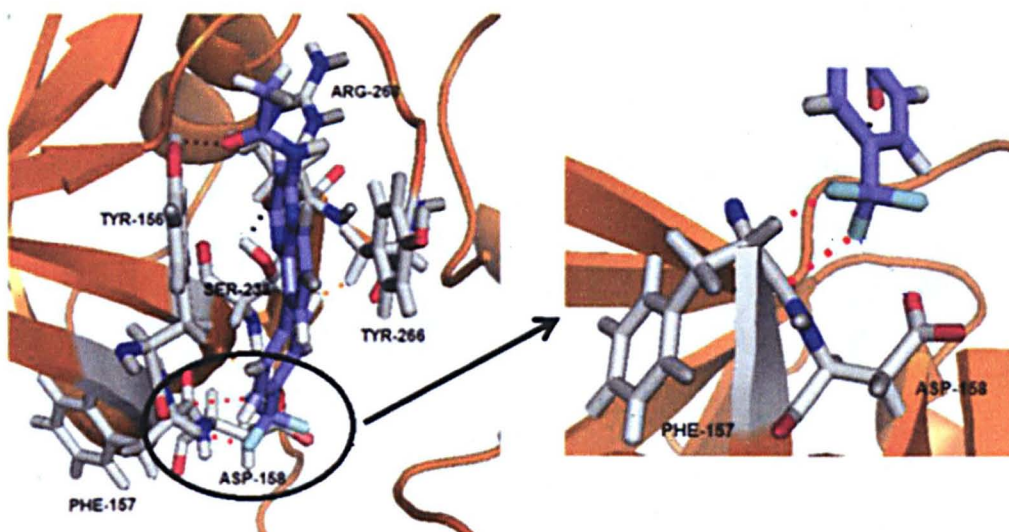
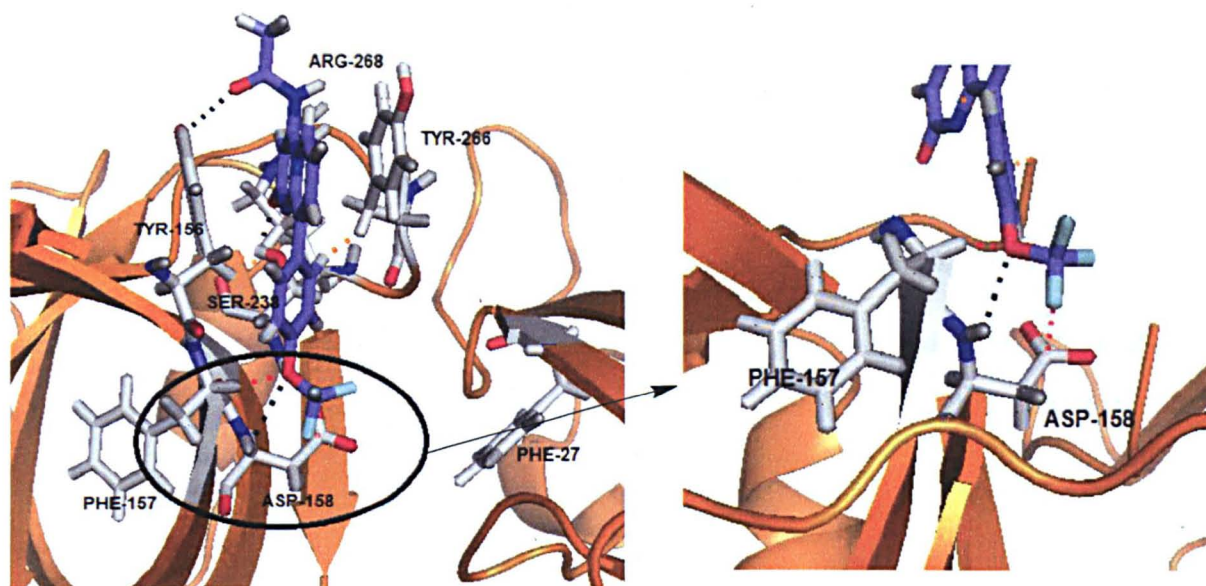


Figure 33: Binding mode of **31a** (blue CPK sticks). POT1 is displayed as an orange solid ribbon. The protein side chains are displayed as grey CPK sticks. H-bonds are depicted as black dotted lines and multipolar interactions are depicted as red dotted lines.

Examining the ligand binding modes of **31c** and **31f** suggest that they are also capable of undergoing orthogonal multipolar interactions with the carbonyl carbon and H-C_α backbone of PHE-157. These results indicate that the interaction with PHE-157 is essential for activity. The binding mode of the nitro substituent (**31d**) suggests that it also forms an interaction with the H-C_α backbone of PHE-157. However the

interaction with the C=O group of PHE-157 is lost. The nitrile substituent (**31b**) forms a H-bond with the NH group of ASP-158 whereas the trifluoromethoxy substituent (**31e**) forms a H-bond with the NH group of ASP-158 and one of the fluorine atoms in the trifluoro group also participates in a dipole-dipole interaction with the carbonyl carbon of ASP-158 (Figure 34).



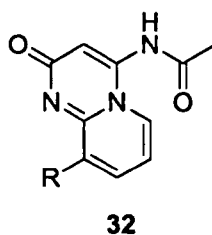
*Figure 34: Binding mode of **31e** (blue CPK sticks). POT1 is displayed as an orange solid ribbon. The protein side chains are displayed as grey CPK sticks. H-bonds are depicted as black dotted lines, vdW contacts are displayed as orange dotted lines and dipole-dipole interactions are depicted as red dotted lines.*

The results indicate that substituents which are capable of undergoing multipolar interactions with PHE-157 and ASP-158 may be essential for strong ligand binding. In order to identify the optimal position for the fluorine substituent, the substituent was moved to the 2- and 3- positions around the pendent phenyl ring resulting in ligands **31g** and **31h** (Table 7). Moving the fluorine substituent to the 2-position (**31g**) was well tolerated whereby POT1-DNA binding decreased by 35%. The fluorine

substituent at this position forms a lipophilic interaction with the methylene CH₂ group of TYR-156. However, moving fluorine to the 3-position (**31h**) resulted in no effect on POT1-DNA binding. The binding mode suggests that the fluorine in **31h** is also capable of a lipophilic interaction with the methylene CH₂ group of SER-238. However due to the large standard error in the mean, it cannot be ruled out that **31h** may also have the ability to decrease POT1-DNA binding to the same extent as **31g**. Currently **31h** seems to be an outlier and needs to be re-tested in the FP assay.

6.6.3 Probing electron withdrawing substituents

The top electron withdrawing analogues were elaborated synthetically to enable additional interactions within the binding pocket resulting in the development of new ligands (Table 8).



| Compound No. | R | Maximum POT1-DNA binding (%) |
|--------------|---|---------------------------------|
| 32a | | 85±7 |
| 32b | | 75±5 |

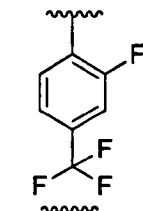
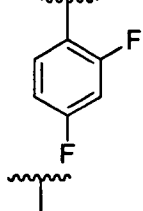
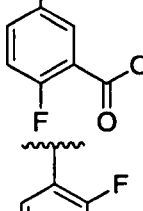
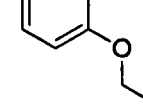
| | | |
|-----|---|-------|
| 32c |  | 85±7 |
| 32d |  | 86±14 |
| 32e |  | 84±11 |
| 32f |  | 80±10 |

Table 8: FP results for electron-withdrawing substituents. Percentage remaining signal at 100μM compound relative to no compound signal (mean ± standard error) of 3 determinations.

It was expected that an increase in binding affinity would be observed when an additional fluorine substituent was placed at the ortho position of the phenyl ring as in **32a**. On the contrary, the results indicated more than 2-fold decrease in POT1 inhibitory activity compared to **31g** where a single fluorine substituent is used. According to the binding model (Figure 35), the positioning of the extra fluorine results in a change in the conformation of the ligand. This conformational change causes the pyrimidine amide to lose its H-bond with TYR-156.

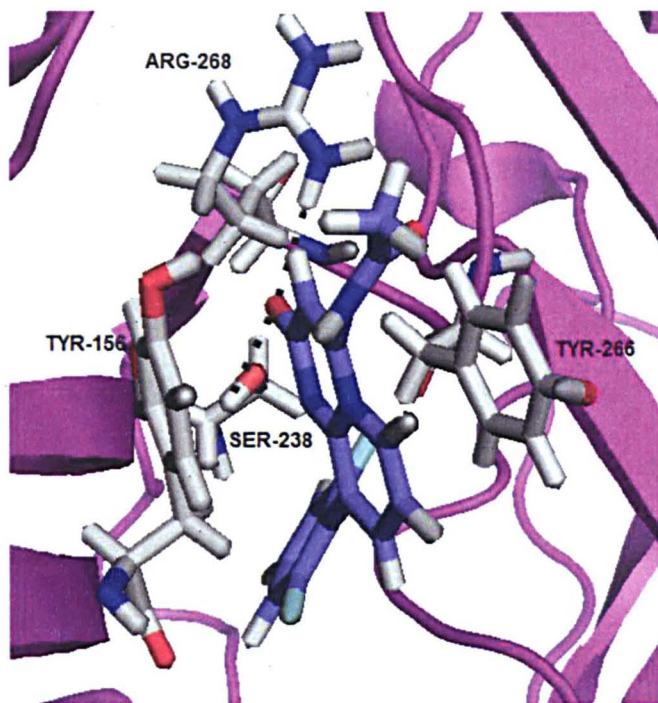


Figure 35: Binding mode of **32a** (blue CPK sticks). POT1 is displayed as a magenta solid ribbon. The protein side chains are displayed as grey CPK sticks. H-bonds are depicted as black dotted lines.

The change in ligand conformation was further supported when the 2-fluorine substituent in **31g** was replaced with a more electron-withdrawing 2-trifluoromethyl group (**32b**) which also resulted in a decrease in activity. However **32b** was still more active than **32a** whereby POT1-DNA binding decreased by 25% as compared to 15% for **32a**. The decrease in activity is probably due to the steric bulk of the trifluoromethyl group which causes a change in the ligands conformation and thus the carbonyl of the amide loses its H-bond interaction with TYR-156.

Structures **32c** and **32d** were synthesized where the top substituents, 2- and 4-fluorine substituents and the 4-trifluoromethyl substituent were combined into a

single structure. It was anticipated that ligands with electron withdrawing groups at both the ortho and para positions on the phenyl ring would show better activity than the ligands with substituents on just a single position. However **32c** and **32d** only decreased POT1-DNA binding by 15% and 14% respectively. The binding modes of these ligands suggested that the reduction in activity is most likely due to the loss of the H-bond which the amide forms with TYR-156. These results suggest that placing extra substituents on the phenyl ring has a dramatic effect on the conformation of the ligand which causes the important H-bond with TYR-156 to be lost and thereby results in ligands with reduced POT1-DNA inhibitory activity.

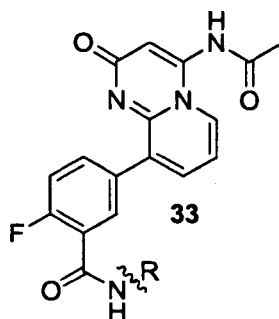
Ligand **32e** (Table 8) was synthesized to enable the fluorine to maintain its multipolar interaction with PHE-157 backbone and the acid group to exploit additional interactions within the binding pocket. However no increase in binding activity was observed compared to **18m** and the ligand only decreased POT1-DNA binding by 16%. Ligand **32f** (Table 8) was also synthesized so that ethyl ether moiety at the 3-position could bind and exploit lipophilic interactions deeper within the pocket. However when tested, the compound decreased POT1-DNA binding by 20%.

Subsequent modelling studies suggest that the placement of the acid group and ethyl ether group at the 3- position in **32e** and **32f** respectively, results in the ligands adopting an alternative conformation. According to the docked binding modes, the hydroxyl group of the acid moiety in **32e** participates in a H-bond with the amino group of ASN-70. Docking studies also suggest that the ethyl ether moiety in **32f** extends deeper with the pocket as anticipated. However it is suggested that the gain in POT1 binding affinity is being reduced due to the additional placement of functional groups on the pendent phenyl ring. These additional groups cause the

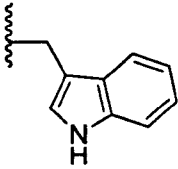
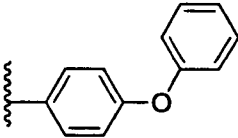

ligands to adopt slightly different binding modes within the pocket. Although extra interactions are formed by these groups, the change in the conformation causes other vital interactions such as the H-bond interaction between the amide group and TYR-156 to be lost thus attenuating inhibitory activity.

6.6.4 Modification of ligand 32e

To improve the affinity of **32e** (Table 8) for the POT1 pocket, the structure was elaborated synthetically resulting in ligands **33a-e** (Table 9). It was anticipated that fluorine would retain its multipolar interaction with the backbone of PHE-157 and the various R groups at the 3-position on the phenyl ring would allow the ligands to exploit lipophilic interactions deeper within the pocket.



| Compound | R | Maximum POT1-DNA |
|------------|---|------------------|
| NO. | | binding (%) |
| 33a | | 147±15 |
| 33b | | 112±4 |

| | | |
|------------|---|--------|
| 33c |  | 151±57 |
| 33d |  | 110±6 |
| 33e |  | 153±28 |

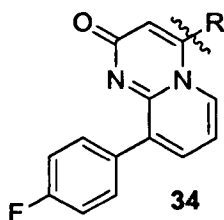
*Table 9: FP results for derivatives of **33**. Percentage remaining signal at 100μM compound relative to no compound signal (mean ± standard error) of 3 determinations.*

The FP data concludes that modifications at the 3-position on the pendent phenyl ring results in loss of activity. It is suggested that large bulky modifications leads to steric hindrance at the binding pocket and therefore such ligands do not bind to POT1. The majority of the ligands show more than 100% POT1-DNA binding suggesting that these compounds may be interfering with the assay. One concern is that these compounds may be insoluble and therefore likely to cause light scattering resulting in the emitted light to be 100% polarised.

6.6.5 Modification of N-acyl group

Since further modifications to the pendent phenyl ring were unsuccessful to identify superior POT1 inhibitors. A more feasible approach was undertaken to modify the N-acyl terminal of **18m**. According to the binding model, the N-acyl group pointed outwards towards the solvent. The carbonyl functional group of the amide was important for binding and participated in a H-bond interaction with the hydroxyl group

of TYR-156. Analogues of **34** (Table 10) were subsequently synthesized where the R group was replaced by various amide and sulphonamide groups to exploit additional interactions with residues near the surface of the binding pocket.

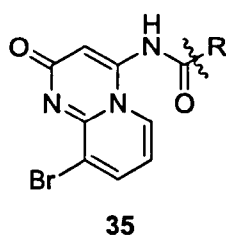


| Compound NO. | R | Maximum POT1-DNA binding (%) |
|--------------|---|------------------------------|
| 18m | | 46±9 |
| 34a | | 88±5 |
| 34b | | 90±7 |
| 34c | | 144±14 |
| 34d | | 97±6 |
| 34e | | 98±2 |
| 34f | | 115±13 |

Table 10: FP results for derivatives of **34**. Percentage remaining signal at 100μM compound relative to no compound signal (mean ± standard error) of 3 determinations.

The majority of phenyl sulphonamides and phenyl amides were found to be inactive and had no effect on POT1-DNA binding. It is suggested that such bulky groups may lead to a steric clash at the binding pocket. However the smaller R groups such as the amino (**34a**) and trifluoroamide (**34b**) groups did display some marginal activity and decreased POT1-DNA binding by 12% and 10% respectively.

Modifications to the N-acyl terminal of the initial starting fragment **27** was also initiated (Table 11).



| Compound NO. | R | Maximum POT1-DNA binding (%) |
|--------------|-----------------|------------------------------|
| 27 | CH ₃ | 96±13 |
| 35a | | 107±6 |
| 35b | | 143±3 |
| 35c | | 64±4 |
| 35d | | 90±7 |

Table 11: FP results for derivatives of **35**. Percentage remaining signal at 100μM compound relative to no compound signal (mean ± standard error) of 3 determinations.

Replacing the methyl group in **27** with a 4-nitro phenyl group (**35a**), a morpholine moiety (**35b**) or a trifluoroamide group (**35d**) had no effect on POT1-DNA binding. Surprisingly, the replacement of the methyl group by a more flexible succinic acid group (**35c**) decreased POT1-DNA binding by 36%. The binding mode of **35c** (Figure 36) suggests that the flexibility of the succinic group allows the carbonyl of the acid to participate in a H-bond with the hydroxyl group of TYR-266. This H-bond with TYR-266 may enhance the structure's bonding affinity for POT1. It is suggested that flexible polar groups capable of exploiting H-bonds with both hydroxyl groups of TYR-156 and TYR-266 may afford more potent POT1 inhibitors.

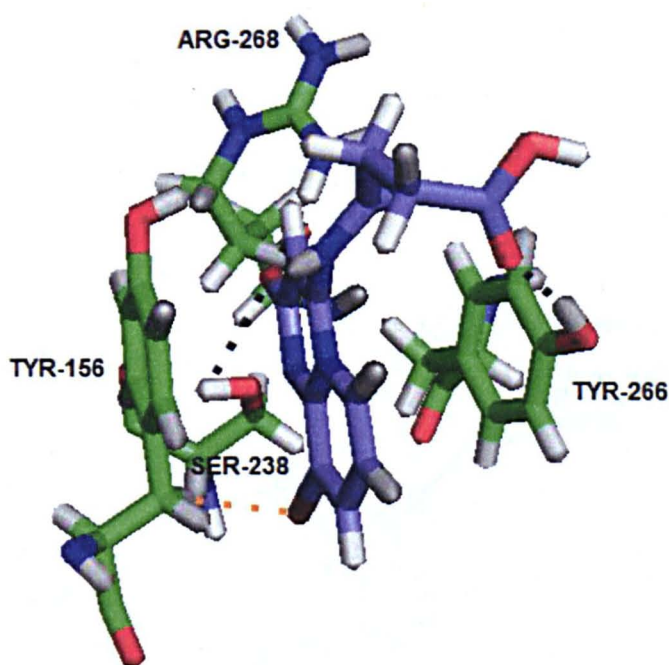


Figure 36: Binding mode of **35c** (blue CPK sticks). The protein side chains are displayed as green CPK sticks. H-bonds are depicted as black dotted lines and vdW interactions depicted as orange dotted lines.

6.6.6 SAR summary

To summarize, it has been found that electron-withdrawing substituents at the 4- or either of the 2- positions on the pendent phenyl ring are prerequisite for binding. However placing electron withdrawing groups at both 2- and 4- positions leads to loss of activity most likely due to a conformational change in the ligands binding mode. Placing polar or charged substituents on the phenyl ring also results in the loss of activity. The amide group on the pyrimidine core is essential for activity since it participates in a H-bond interaction with TYR-156. It is suggested that if this particular H-bond is lost, ligand inhibitory activity decreases. Replacing the 4-amide group with much larger phenyl amides leads to complete loss of activity. Furthermore, large hydrophobic groups placed at the 3- position on the pendent phenyl ring are also not tolerated and leads to loss of activity. Figure 37 summarizes the types of modifications which have been carried out.

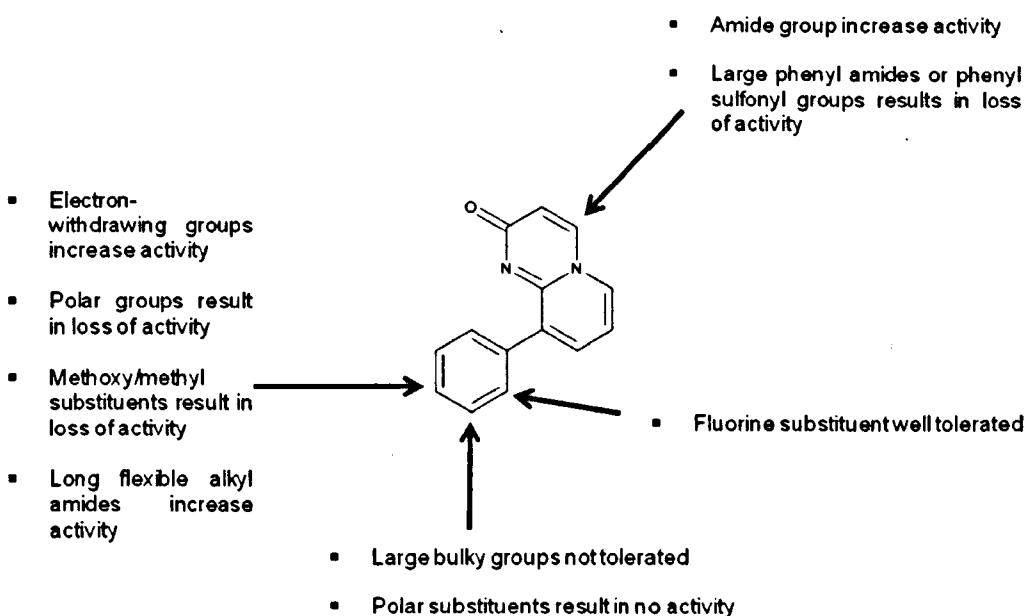


Figure 37: Summary of SAR.

Chapter 7: Virtual compound libraries

7.1 Introduction

Combinatorial chemistry is a powerful approach for the generation and optimisation of lead compounds in medicinal chemistry. It promises the synthesis of a large number of compounds for high throughput screening applications and involves the systematic assembly of a set of 'building blocks' to produce a large library of chemically diverse structures. Computational methods have been developed to represent combinatorial libraries *in silico*. Virtual libraries have the advantages of saving time and money by reducing the number of compounds to be experimentally tested and to eliminate compounds which are unlikely to bind to the target in addition to improving the drug discovery success rate (Leach, 2001). To increase the probability of finding potent hits against POT1, an *in silico* based focussed library was designed.

7.2 Library design

POT1 *in silico* compound libraries were enumerated using SMIRKS in the ChemOffice package (Mills, 2006) by Dr Marc Hummersone and Dr Mark Frigerio, Pharminox Ltd. SMIRKS is a line notation which uses alphanumeric characters to specify reaction transforms and enumeration is the process by which the connection tables for the product structures in a virtual library are produced.

The libraries were based on the pyrido[1,2-*a*]pyrimidine template and were generated using a reaction transform approach which reacts the initial reagent structures according to the rules of synthetic chemistry to produce the end products. An example is illustrated in Figure 38. When the acid functional group of the pyrido[1,2-*a*]pyrimidine template reacts with a set of commercially available amines, the enumeration engine applies a reaction transform approach which results in an amide bond between the two reactants to produce three different end products.

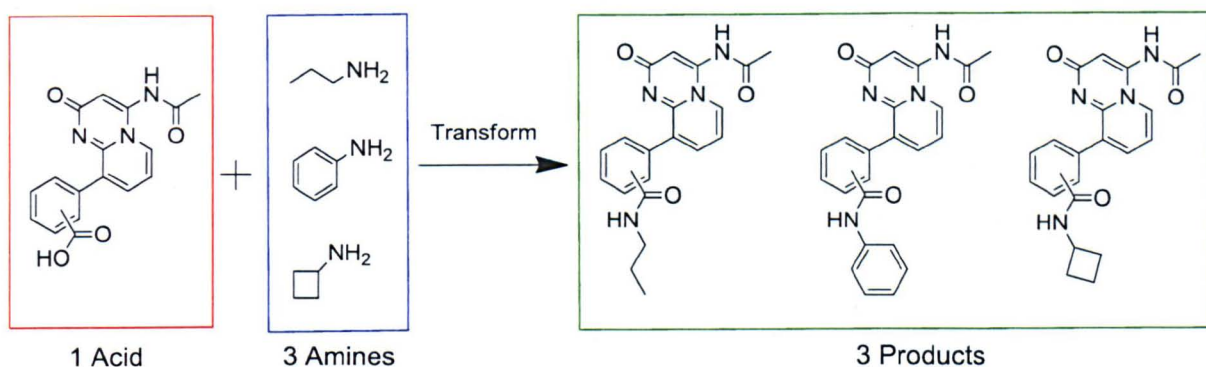


Figure 38: Reaction based enumeration of pyrido[1,2-*a*]pyrimidine template.

Using four different intermediate pyrido[1,2-*a*]pyrimidine templates: **36**, **37**, **38** and **39** (Figure 39). Eight structurally diverse focussed libraries were generated.

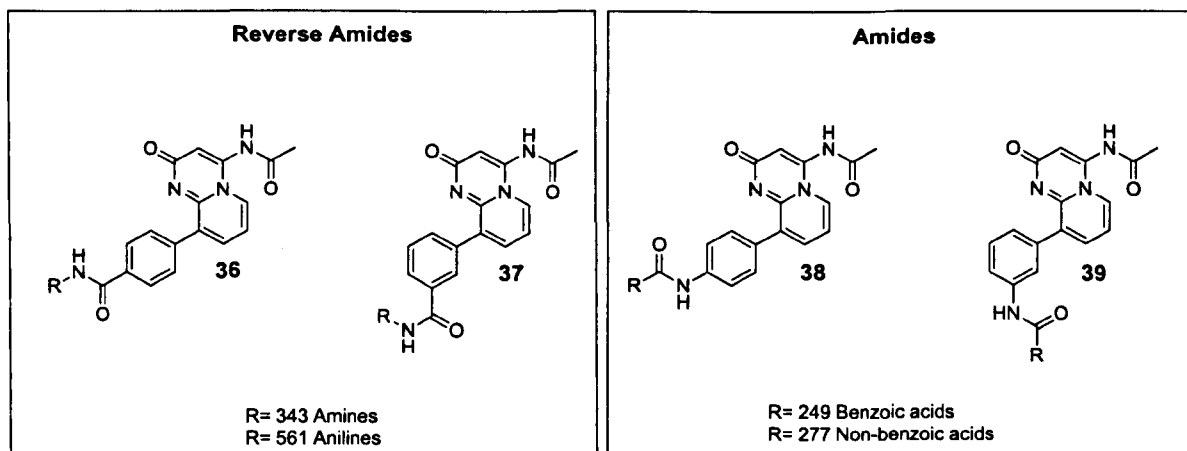


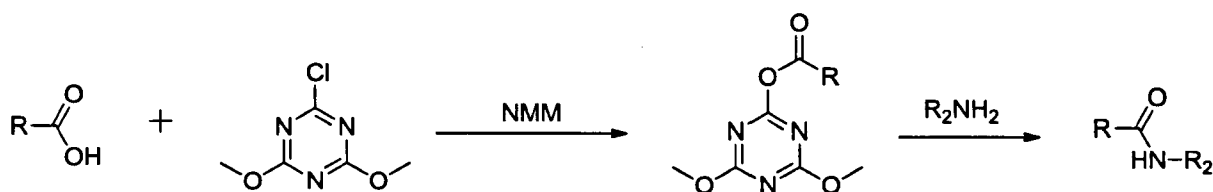
Figure 39: Chemical structures of the four pyrido[1,2-*a*]pyrimidine intermediates. Libraries were classified as amides or reverse amides depending on whether the acid or amine moieties were directly attached to the pendent phenyl ring. The amine and acid groups were placed at the 3- and 4- positions of the phenyl ring which were then reacted with their corresponding commercially available complementary partners: benzoic acids, non-benzoic acids, amines and anilines. Each intermediate structure generated two libraries hence four intermediates generated 8 libraries. For example, ligand **36** formed 2 libraries, one with 343 structures containing amine R groups and the other library containing 561 aniline R groups structures. The reagent structure files were downloaded from the Sigma Aldrich and Acros Organics compound collection.

A total of 2860 ligand structures were generated *in silico* based on eight focussed libraries and docked into predicted POT1 pocket using GOLD. The libraries were ranked according to GoldScore and the top 7 structures from each of the eight libraries were chosen for synthesis and further evaluation giving a combined total of 56 structures. These 56 structures are deposited in Appendix B. Libraries were also created on modifications at the 2-position of the pendent phenyl ring. However due to low GoldScores of these structures, they were not synthesised. It was found that

modifications at the 2-position were unfavourable due to steric hindrance between the ligands and the binding pocket residues.

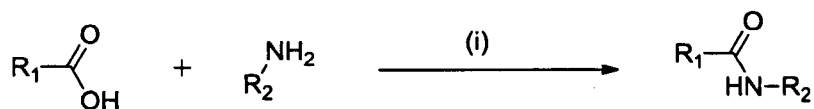
7.3 Library chemistry

The synthesis of POT1 ligand libraries were carried out using a simple peptide coupling reaction involving CDMT (2-chloro-4,6-dimethoxy-1,3,5-triazine). CDMT is used as a coupling agent and the procedure is well documented (Garrett *et al*, 2002). The standard method of making the POT1 amide ligands is to activate the acid using CDMT and a base such as *N*-methylmorpholine (NMM) which generates an active ester. The ester is subsequently reacted with the amine coupling partner in the same pot (Scheme 6).



Scheme 6: Typical procedure for amide bond formation using CDMT.

Scheme 7 describes the reaction conditions used to synthesise the POT1 ligand libraries.



Scheme 7: Synthetic scheme for the synthesis of POT1 inhibitors. Reagents and conditions (i) CDMT, 4-methylmorpholine, DMF, 2h, room temperature.

7.4 Results and discussion

A total of 15 structures were selected for synthesis from the 56 ligands deposited in Appendix B. Compounds were selected from all four pyrido[1,2-a]pyrimidin-2-one library templates. The chemical structures of the 15 ligands are presented in Figure 40 and their FP results are presented in Figure 41.

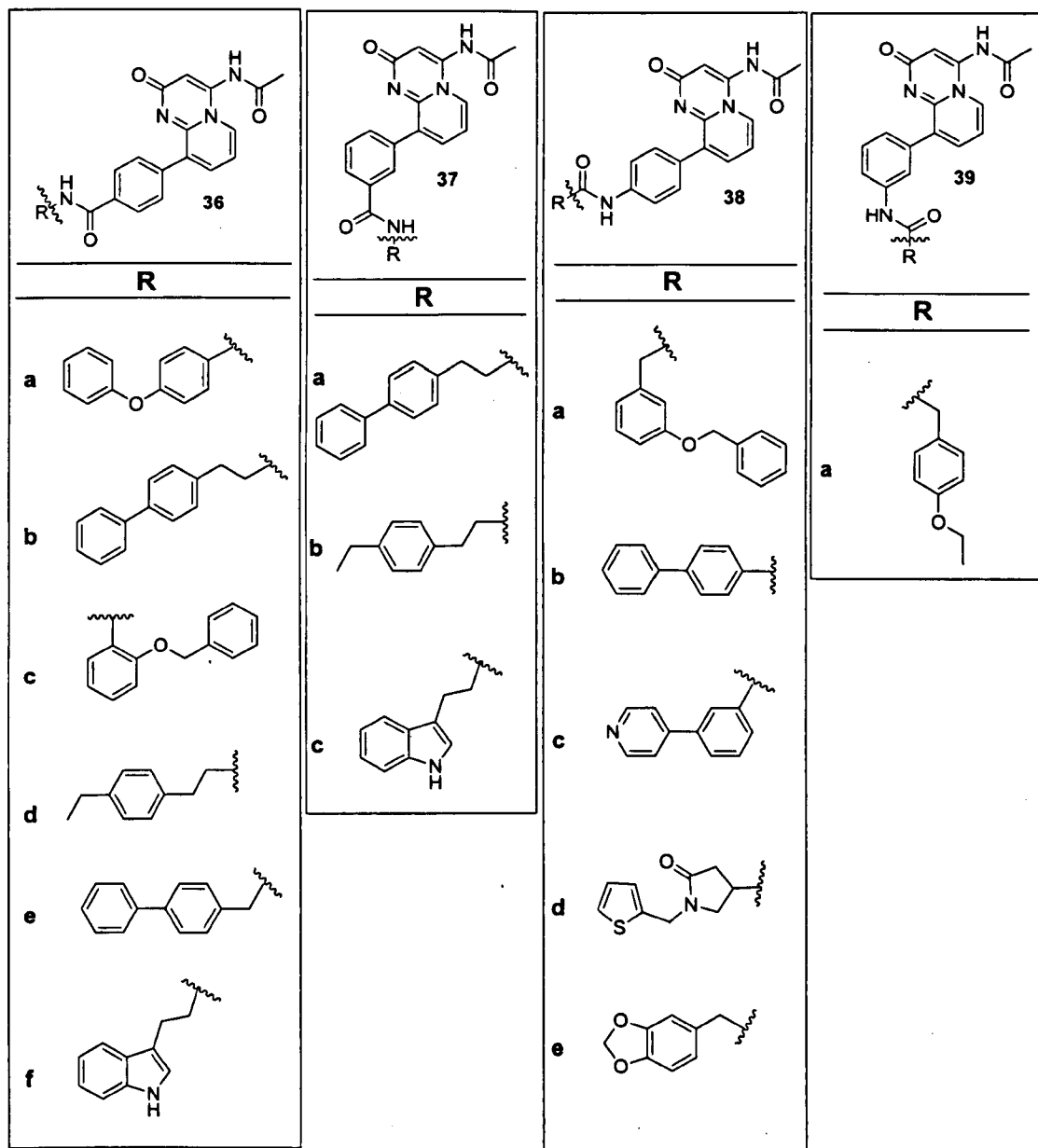


Figure 40: Chemical structures of the 15 ligands selected for synthesis from the library set.

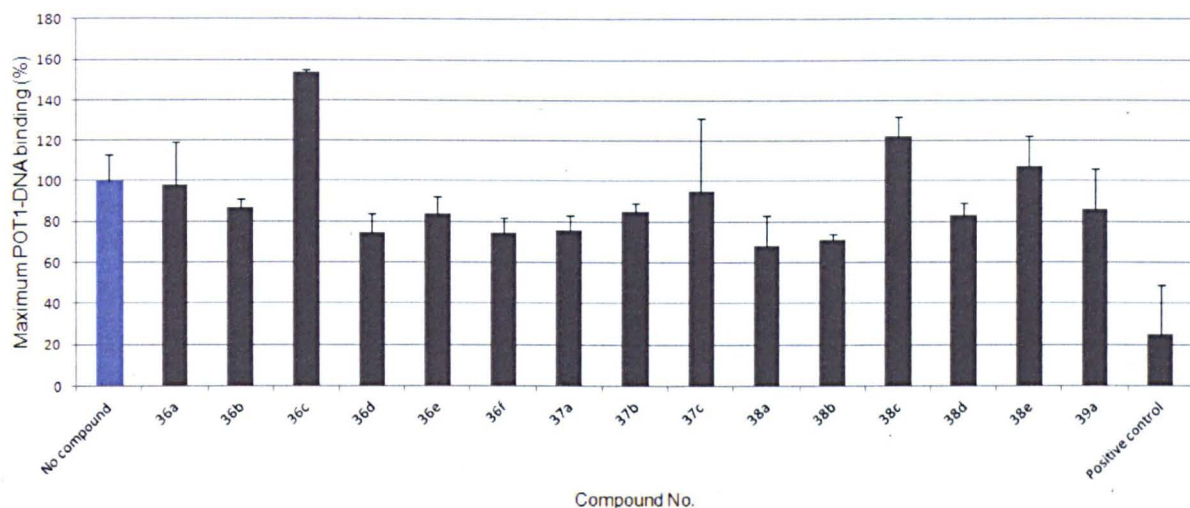


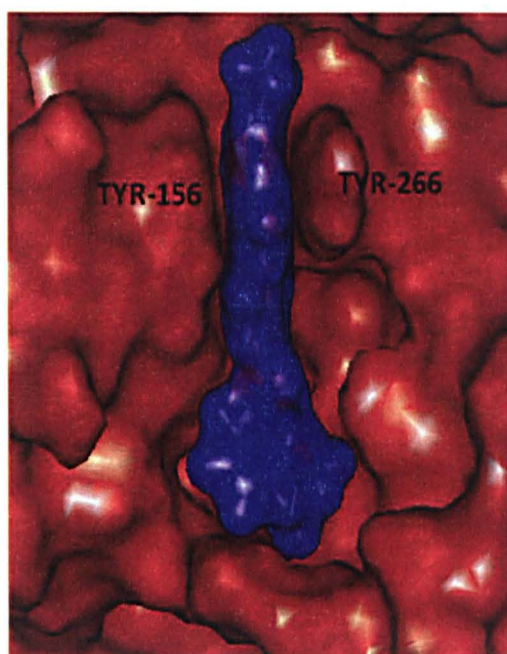
Figure 41: FP screening results for POT1 library ligands. Percentage remaining signal at 100 μ M compound relative to no compound signal (mean \pm standard error) of 3 determinations.

It was anticipated that structurally diverse R group modifications at the 3- and 4-positions on the pendent phenyl ring would discover superior POT1 inhibitors. The majority of compounds which were tested decreased POT1-DNA binding between 13-32%. Six ligands were synthesized from the library based on template **36**. From this library set, two ligands: **36a** and **36c** had no effect on POT1-DNA binding. It is suggested that the high signal received from **36c** may be an artefact. Ligand **36d** decreased POT1-DNA binding by 25% and is predicted to be able to form 3 H-bonds with TYR-156, SER-238 and ARG-268. Ligand **36b** and **36e** decreased POT1-DNA binding by 13% and 16% respectively. Structure **36f** also decreased POT1-DNA binding by 25% and subsequent modelling studies suggest that its pyrimidine core participates in 3 H-bonds with TYR-156, SER-238 and ARG-268 and the NH of tryptamine forms an additional H-bond with ASP-158.

Three structures were synthesized from the library based on template **37**. Structures **37a** and **37b** decreased POT1-DNA binding by 24% and 15% respectively. The

docked binding modes suggest both ligands are able to form 3 H-bonds with nearby protein residues and the R-groups extend deeper into the hydrophobic pocket. However structure **37c** was found to be inactive and had no effect in disrupting POT1-DNA interaction.

Five ligands were synthesized based on template **38**. Structures **38a** and **38b** were found to be the top ligands in terms of inhibitory activity from the entire library set and decreased POT1-DNA binding by 32% and 29% respectively. The modelling results suggest that ligand **38a** is able to participate in 4 H-bond interactions with TYR-156, ARG-268, SER-238 and ASN-70. The docked binding mode of **38a** (Figure 42) suggests that although the ligand is extending deeper into the binding pocket, a large area of the binding pocket is still not exploited and additional modifications to the structure may be necessary for optimal binding.



*Figure 42: Predicted binding mode of **38a** (blue CPK surface). POT1 is displayed as an orange CPK surface.*

Ligand **38b** on the other hand is able to form 3 H-bonds with TYR-156, ARG-268 and SER-238, all of them contributed by the pyrimidine core. The biphenyl ring binds deeper within the pocket exploiting lipophilic interactions with the binding site residues. Ligand **38d** also decreased POT1-DNA binding by 17% and the docking results suggest that it forms 3 H-bonds with TYR-156, SER-238 and the amide of the phenyl also shares a H bond with ASP-158. Ligands **38c** and **38e** did not have any effect in disrupting POT1-DNA interaction.

A single structure was synthesized based on template **39**. Structure **39a** decreased POT1-DNA binding by 14%. The binding mode suggests that the entire ligand forms only one H-bond interaction with the binding site residues.

It is suggested that due to the lack of sufficient H-bond donor and acceptor groups in the majority of the R groups, not enough H-bonds are being formed between the R groups of the ligands and binding site residues. Currently the majority of H-bonds are being contributed by the pyrimidine core of the ligands. Therefore R groups which contain H-bond donor and acceptor functional groups should be synthesized for further investigation.

Chapter 8: Sulfathiazole analogues

8.1 Introduction

In this chapter a slightly different approach to *ab initio* drug design was undertaken to identify additional chemical scaffolds which have the potential to bind to POT1. Instead of designing a molecule from 'scratch', *in silico* screening techniques were used to identify potential structures which could subsequently be synthesized and screened against POT1.

8.2 Selection of sulfathiazole-based inhibitor

Using ZINC (<http://zinc.docking.org/>), a free chemical database containing 21 million commercially available compounds for virtual screening purposes (Irwin *et al*, 2012). A pre-chosen subset of 55000 drug like structures was selected with the following molecular constraints: xLogP = <5, molecular weight = >150 and <500g/mol, H-bond acceptors = <10, rotatable bonds = <8 and polar surface area = <150Å². The structure library was subsequently docked into the predicted POT1 pocket and the top 6 structures with the highest GoldScores were selected for further evaluation. The chemical structures of the top 6 compounds are deposited with their corresponding GoldScores in Appendix C.

On the basis of synthetic accessibility, the sulfathiazole-based compound (ZINC ID= ZINC09223825) was selected and a substructure search was carried out using the Sigma Aldrich compound collection to find similar structures. Succinylsulfathiazole

(40), a marketed antibacterial drug used for the treatment of gastrointestinal infections was highlighted due to its low cost (Figure 43).

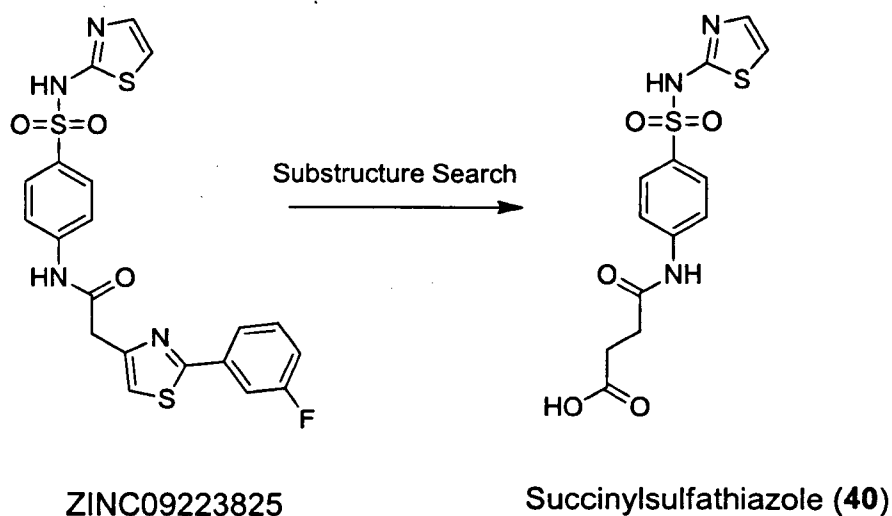


Figure 43: Chemical structures of ZINC09223825 which was initially identified through virtual screening and succinylsulfathiazole (40) which was identified using substructure search.

8.3 Sulfathiazole modifications

The docked binding mode of succinylsulfathiazole (40) (Figure 44) suggests that the sulfathiazole moiety binds in between TYR-156 and TYR-266 and the phenyl amide extends into the pocket where it forms a H-bond with ASP-158 and the hydroxyl group of the acid participates in a H-bond with the backbone of PHE-27.

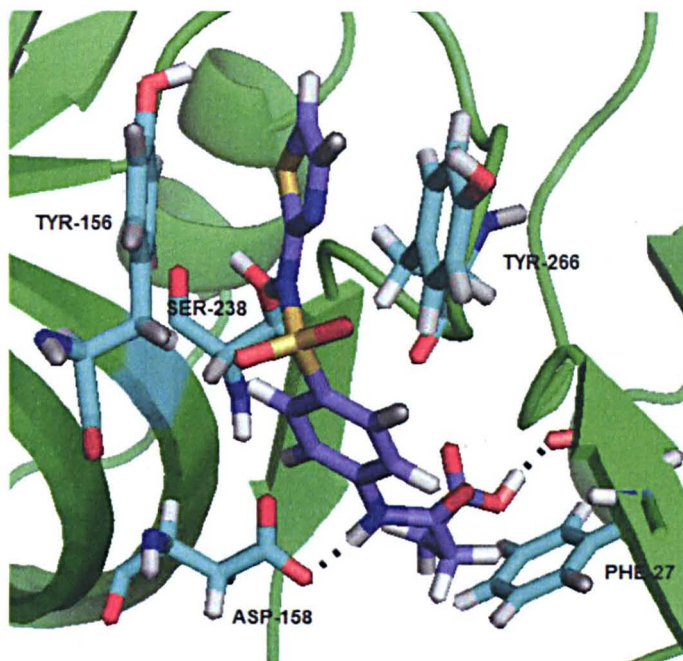


Figure 44: Binding mode of succinylsulfathiazole (**40**) (blue CPK sticks). POT1 is displayed as a green solid ribbon. The protein side chains are displayed as cyan CPK sticks. H-bonds are depicted as black dotted lines.

In order to exploit the π - π stacking interaction with PHE-27, subtle modifications were carried out on **40**. It was decided that simple phenyl amide derivatives should be placed at the end of the succinic group to exploit this π - π stacking interaction resulting in the synthesis of six compounds (Figure 45).

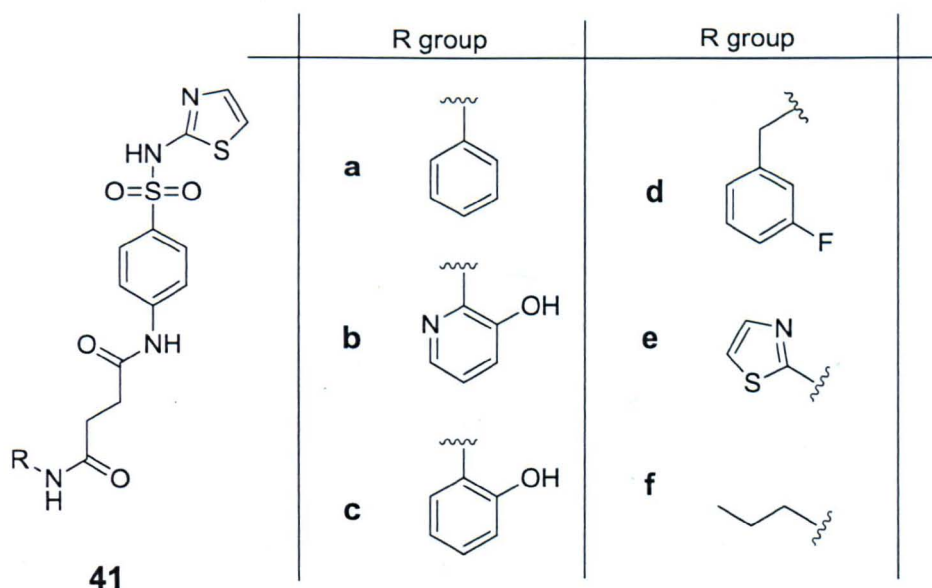


Figure 45: Chemical structures of sulfathiazole analogues.

The FP results of all six analogues as well as succinylsulfathiazoles (**40**) are presented in Figure 46.

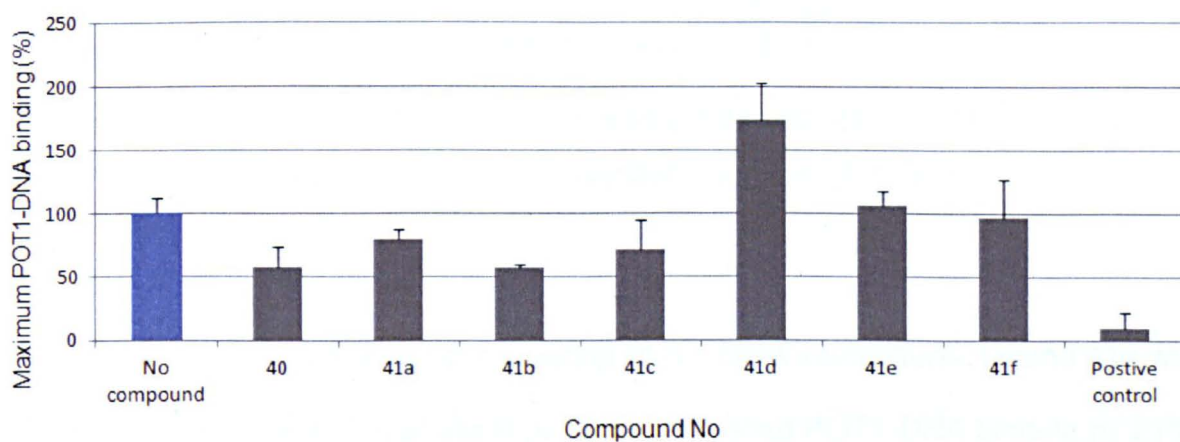


Figure 46: FP results for sulfathiazole analogues. Percentage remaining signal at 100 μ M compound relative to no compound signal (mean \pm standard error) of 3 determinations.

The results in Figure 46 suggest that succinylsulfathiazole (**40**) has the ability to decrease POT1-DNA binding by 43%. However its derivatives **41d**, **41e** and **41f** were found to be inactive and had no effect on POT1-DNA binding. Subsequent modelling studies suggest that these ligands bind away from TYR-156 and TYR-266 and therefore are ineffective in disrupting the binding and stacking of the T7 base of DNA (Figure 47).

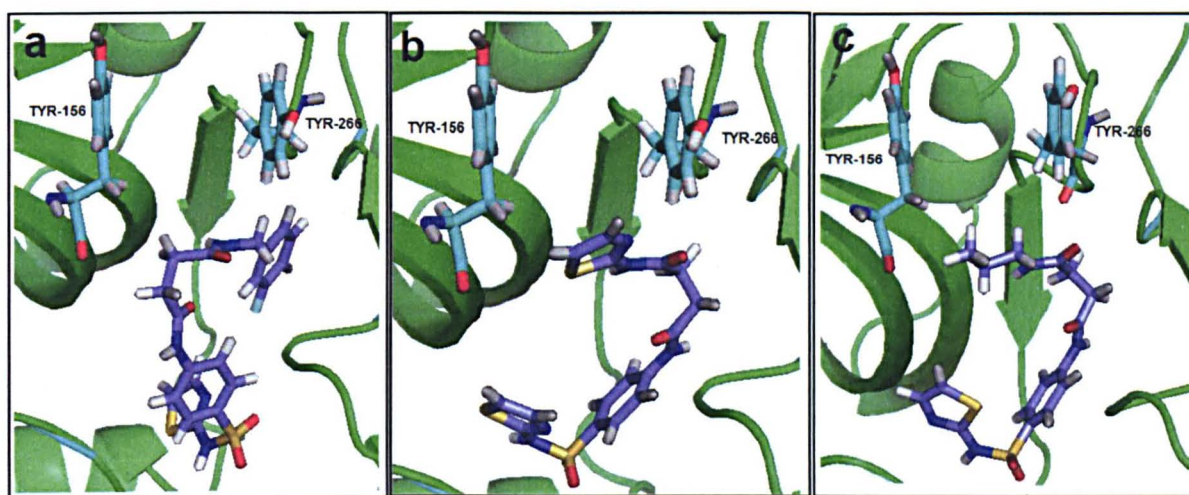


Figure 47: Predicted binding modes of (a) **41d** (b) **41e** (c) **41f**. POT1 is displayed as a green solid ribbon and the ligands are displayed as blue CPK sticks.

Ligands **41a-c** had some effect in disrupting POT1-DNA interactions. Ligand **41a** with an unsubstituted phenyl ring at the R position decreased POT1-DNA binding by 21%. The structure is predicted to form two H-bonds with ASP-158 and THR-144 and its binding mode is different to that of **40**. It was initially suggested that the phenyl rings in **41a-c** would exploit π - π stacking interactions with PHE-27 however this was not observed. The phenyl group instead stacked in between TYR-156 and TYR-266 and the sulfathiazole moiety extended deeper into the pocket (Figure 48).

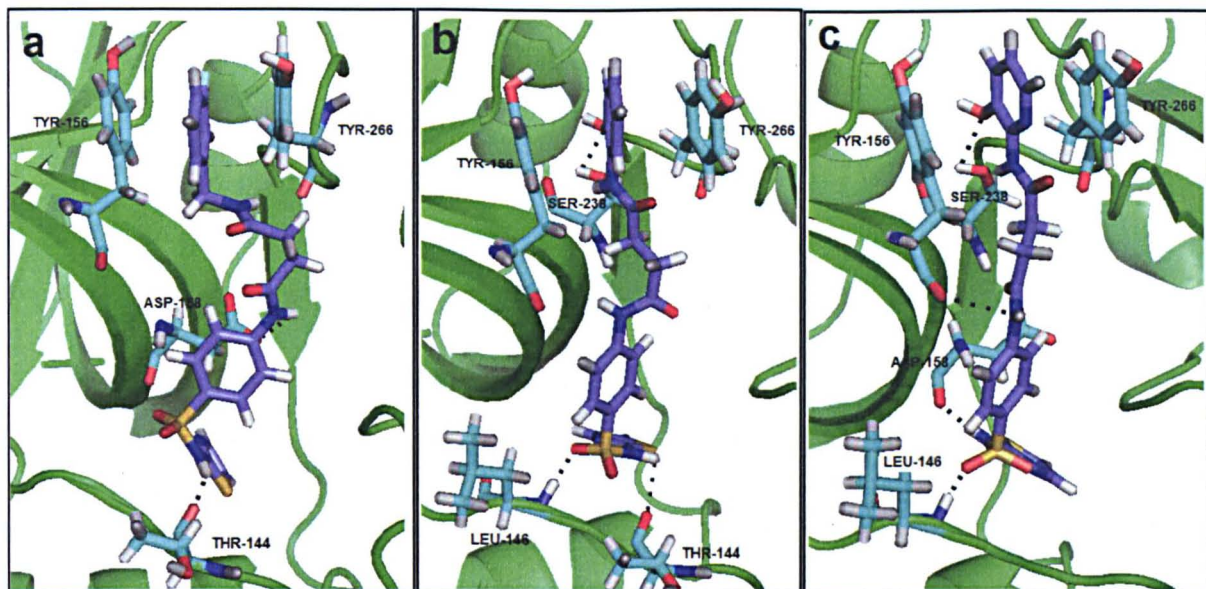


Figure 48: Predicted binding mode of (a) **41a** (b) **41c** (c) **41b**. POT1 is displayed as a green solid ribbon and the ligands are displayed as blue CPK sticks. H-bonds are depicted as black dotted lines.

The replacement of the unsubstituted phenyl ring with a 2-hydroxyphenyl group (**41c**) decreased POT1-DNA binding by 29%. It was suggested that the hydroxyl group may be involved in H-bond interactions with the binding site residues. This was indeed observed (Figure 48b) where the 2-hydroxyl group participated in a H-bond with SER-238 and the amide of the 2-hydroxyphenyl exploited an additional H-bond with SER-238. When the 2-hydroxyphenyl group was replaced by the 2-hydroxypyridine group as in **41b**, a further increase in ligand inhibitory activity was observed whereby POT1-DNA binding decreased by 43%. It was thought that the nitrogen atom in the pyridine ring was potentially interacting with neighbouring residues. However the binding mode suggested that the pyridine nitrogen formed no interactions, instead a minor conformational change in the ligand permitted the ligand to pick up 5 H-bonds with the binding site residues whereas **41c** could only form 4 H-bonds (Figure 48c).

8.4 Succinylsulfathiazole fragments

The structure of succinylsulfathiazole (**40**) was broken down into six small fragments (Figure 49). Since the fragments were commercially available, they were bought from standard commercial suppliers and screened in the POT1 FP assay. The FP results of the fragments are presented in Figure 50. The purpose of this was to determine whether the individual fragments have the ability to retain the activity of **40**. By identifying the fragment which is essential for activity would allow the reconstruction of the ligand to enhance inhibitory activity.

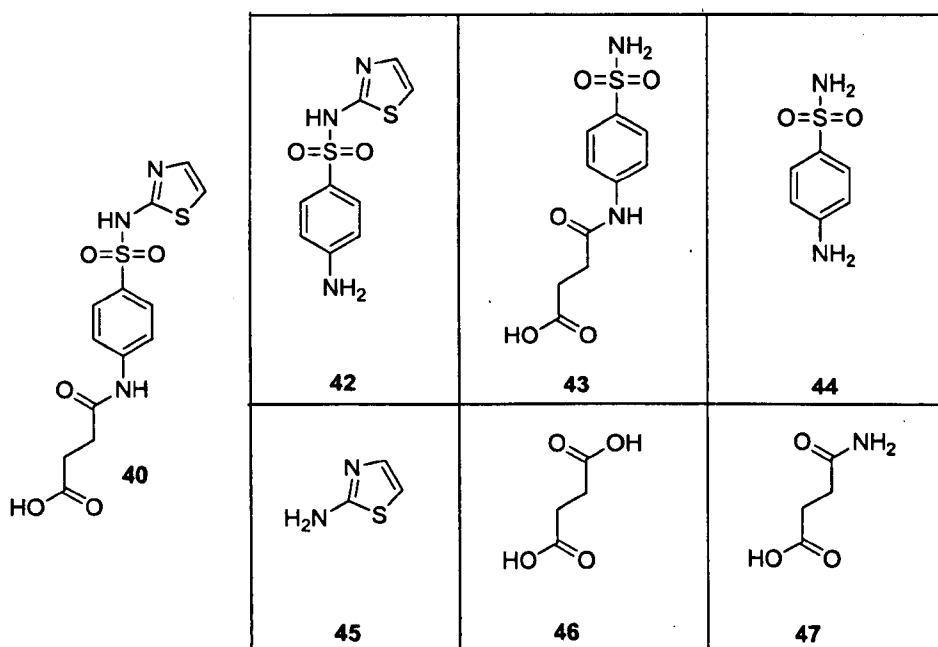


Figure 49: Chemical structures of succinylsulfathiazole fragments.

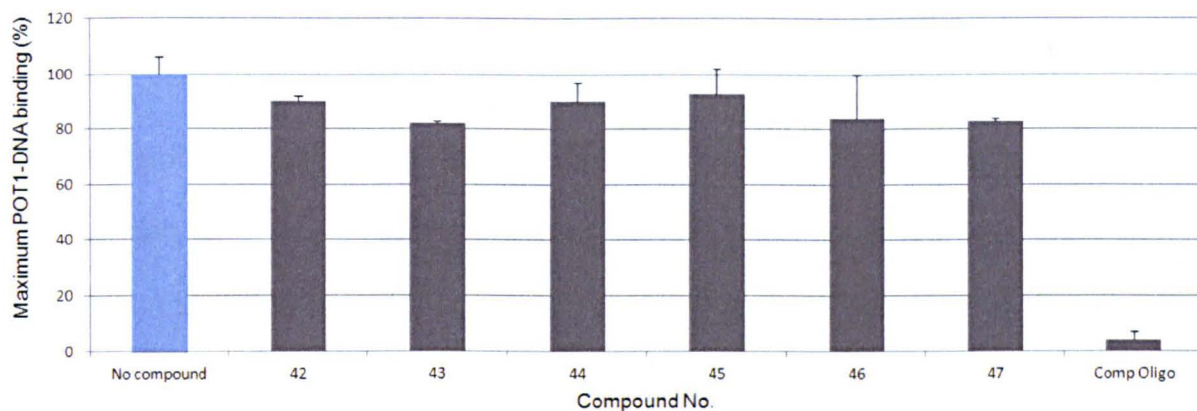


Figure 50: FP results for succinylsulfathiazole fragments. Percentage remaining signal at 100 μ M compound relative to no compound signal (mean \pm standard error) of 3 determinations.

The results suggested that individual fragments were not as active as the complete structure. The majority of the fragments have weak activity in disrupting POT1-DNA interactions. The largest fragment **43** appeared to have the best inhibitory activity whereby it decreased POT1-DNA binding by 18%. The least active fragment was the smaller 2-aminothiazole fragment (**45**) which decreased POT1-DNA binding by 7%. Fragments **42** and **44** decreased POT1-DNA binding by 10% and fragments **46** and **47** decreased POT1-DNA binding by 16% and 17% respectively. Given that the fragments are small, it is suggested that their affinity for POT1 is not strong enough. Furthermore, these small fragments may possibly be occupying different regions of space within the binding pocket as compared to the bound fragments in **40** in which the structure keeps the fragments locked into an ideal position for optimum binding.

8.5 Further modifications

The decision to carry out further modifications on the succinic moiety of **40** was commenced in order to determine whether the succinic group was important for biological activity. A range of ligand structures were synthesized to explore the SAR of this scaffold (Figure 51).

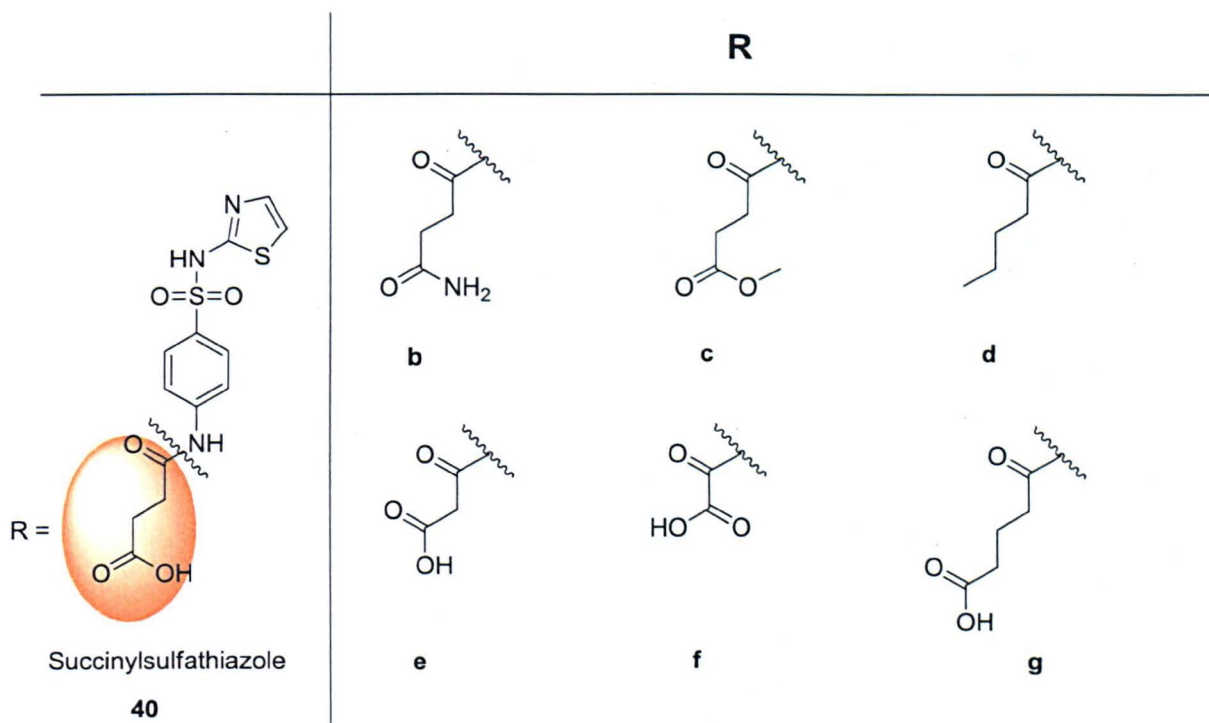


Figure 51: Chemical structures of succinylsulfathiazole modifications.

The carboxyl hydroxyl group in **40** was replaced by an amino group (**40b**). This was done to determine whether the hydroxyl group was important for binding and whether the amino substituent is able to retain similar activity. The hydroxyl group was converted into an ester (**40c**) to determine whether the original proton of the hydroxyl

group is involved as a H-bond donor. The acid functional group in **40** was also removed resulting in **40d** in order to determine whether the acid functional group was essential for binding. Furthermore, the length of the succinic group was incrementally decreased and increased resulting in ligands **40e-g**. This was done to determine the optimum placement of the acid moiety and whether it is important for binding. The FP results for structures **40b-g** are presented in Figure 52.

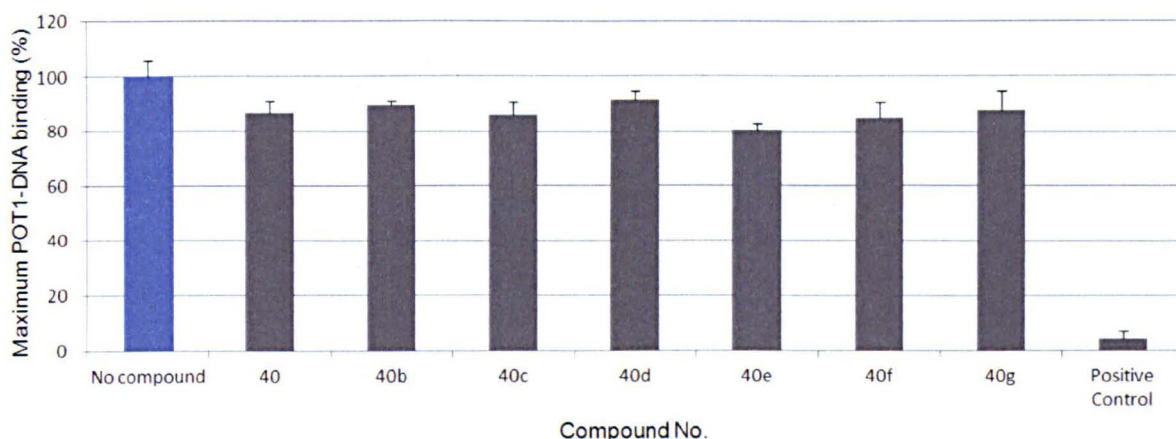


Figure 52: FP results for ligand **40b-g**. Percentage remaining signal at 100 μ M compound relative to no compound signal (mean \pm standard error) of 3 determinations.

The results suggested that the replacement of the hydroxyl group with an amino group or converting it to an ester and removing the acid moiety altogether led to a decrease in inhibitory activity. Furthermore, increasing or decreasing the succinyl group also led to a decrease in activity where POT1-DNA binding decreased between 13-20%. Since the results were unconvincing and inconclusive, succinylsulfathiazole (**40**) was retested again in the FP assay. When tested the second time in the POT1 FP assay, **40** decreased POT1-DNA binding by 13% (Figure 52). This suggests that **40** is not as active against POT1 as originally thought.

There is the possibility that the commercially available compound **40** was not as pure as expected. It could be that the first compound batch that was tested was satisfactory and the compound in the second batch had decomposed. Another possibility is that the material that was tested in the first assay contained an impurity that actually caused the biological activity and this impurity was not present in the second assay.

Chapter 9: Conclusions and Future Work

In conclusion, the structure based design, synthesis and biological evaluation of inhibitors of POT1-DNA interactions have been described. A druggable binding pocket in POT1 was successfully identified using a range of computational tools and the design of POT1 inhibitors was undertaken to target this pocket. More than 90 small molecule ligands have been synthesized based on the pyrido[1,2-*a*]pyrimidin-2-one and sulfathiazole scaffolds. These compounds have been biologically tested in a newly developed POT1 FP displacement assay to confirm whether the structures have the ability to disrupt POT1-DNA interactions.

The majority of compounds tested in the POT1 FP assay decreased POT1-DNA binding to some extent, a general 10-54% decrease in binding at 100 μ M ligand concentration was observed. In the case of the pyrido[1,2-*a*]pyrimidin-2-one ligands, a clear and logical structure activity-relationship between the inhibitor structure and biological activity has been observed. It is found that electron-withdrawing groups at the 2- or either at the 4- positions on the pendent phenyl ring are prerequisite for binding. Placing further electron-withdrawing groups on the pendent phenyl ring leads to a decrease in activity due to a conformational change in the ligands binding mode.

The amide group on the pyrimidine core is also essential for activity since it exploits a H-bond with TYR-156. If this H-bond is lost, ligand activity also decreases. Replacement of the 4-amide group with much larger phenyl groups lead to complete loss of activity. Furthermore, large hydrophobic groups at the 3- position on the pendent phenyl ring are not tolerated and leads to loss of activity. The best inhibitor

which has been obtained is **18m** with a fluorine substituent on the 4- position of the pendent phenyl ring. It is suggested that fluorine forms multipolar interactions with the backbone of PHE-157 which is possibly why this structure as well as other electron-withdrawing groups at the 4- position are showing activity.

The information gained from the FP assay allowed a better understanding of what modifications are necessary to the pyrido[1,2-*a*]pyrimidine scaffold and which modifications are unfavourable. These results have the potential to guide future development of improved lead compounds which could one day be used for the treatment of cancer.

For the sulfathiazole based compounds, the initial succinylsulfathiazole ligand decreased POT1-DNA binding by 43%. However when retested second time in the POT1 assay, the ligand decreased POT1-DNA binding by merely 13%. Further modifications to the succinic acid part of the structure and breaking the structure into its constitutive fragments led to no increase in activity hence no clear SAR was established. Future work will include the synthesis of the commercially available succinylsulfathiazole structure and retesting of the compound to determine biological activity.

Future work on the pyrimidine analogues will include extensive modifications at the N-acyl terminal of **18m**. It is suggested that flexible groups similar to that of succinic acid may be capable of exploiting H-bonds with TYR-156 and TYR-266 thereby resulting in new ligands with better biological activities.

One of the key limitations of this work was the lack of dose-response inhibition data. Although attempts were made to obtain dose-response data at the top ligand concentration of 100 μ M. Concentrations beyond this range needs to be tested in order to acquire IC₅₀ values for the ligands.

The statistical significance of the inhibitory activity of some of the hit compounds is also questionable since the majority of compounds were only tested in triplicates in the assay. Future work will also need to include further compound testing in order to obtain more data repeats and statistical analysis of the data needs to be carried out to determine whether the inhibitory activity of some of the active compounds is statistically significant.

The majority of the compounds synthesized are large lipophilic compounds, hence the solubility of such structures is of concern. Solubility testing needs to be carried out to confirm whether such compounds are soluble at the concentration that they were tested at and whether they are worth pursuing. Solubility is a major issue in FP assays since insoluble compounds tend to cause light scattering and therefore affect the assay readout.

Additional work also needs to be carried out in establishing alternative POT1 biochemical and cell-based assays. One of the limitations of using the FP displacement assay is that one cannot tell whether ligands which decrease POT1-DNA binding are interacting with POT1 or its DNA binding partner. An alternative assay such as ITC needs to be developed to show direct binding of compounds to POT1.

However there is the possibility that the limited activity that is observed for these compounds is linked to an intrinsic limitation in the approach that is being undertaken. It is likely that even if one could discover a ligand that binds strongly to the space occupied normally by the T7 base, the DNA may still bind with reasonable affinity to POT1 by adjusting its conformation. It could be that a significant reduction in POT1-DNA interaction, though not complete may still bring about a significant biological effect.

Chapter 10: Experimental Methods

10.1 Synthesis

10.1.1 General Experimental

All solvents and reagents used in the present study were purchased from standard commercial sources and used without further purification unless otherwise stated.

NMR spectra were acquired on Bruker Avance 400MHz and Bruker Avance (III) 500MHz instruments using deuterated DMSO as solvent. Data analysis for NMR was performed using TopSpin[®] software. Chemical shifts (δ) are reported in parts per million (ppm), downfield from tetramethylsilane (TMS) internal reference. Coupling constants (J) are reported in hertz (Hz). Notations for the ¹H-NMR splitting patterns includes: singlet (s), doublet (d), triplet (t), quartet (q), broad (br), multiplet/overlapping peaks (m), doublet of doublets (dd) and doublet of doublet of doublets (ddd).

Mass spectra were recorded using a Waters 2795 single quadrupole (ESI) spectrometer. TLC was performed using Merck Kieselgel 60 F₂₅₄ plates and spots were visualised under UV light. Purification using preparative TLC was carried out using in house glass backed plates (200mm x 200mm x 1mm) coated with Fluka silica gel 60 F₂₅₄. Flash chromatography was performed using Merck Kieselgel 60 by glass column, Biotage argonaut flash master II or Biotage flash master personal. Melting points were recorded on a Gallenkamp melting point apparatus. Microwave assisted chemistry was carried out using Discover[®] CEM microwave synthesizer.

Analytical HPLC was carried out using Waters 2525 binary gradient module with a Waters 2487 dual λ absorbance UV detector. All retention times (t_R) are quoted in minutes with percentage purity at 254nm. System used: Kromasil C18 column (250mm x 4.6mm, 5 μ M particle size) and gradient elution with acetonitrile/water containing 0.1% TFA (10% to 100% organic over 10-20 min at a flow rate of 1.00 ml/min) unless otherwise stated.

10.1.2 Procedures

This section describes the general procedures for the synthesis of pyrido[1,2-*a*]pyrimidine compounds and sulfathiazole-based compounds. Structures and analytical data are presented in section 10.1.3.

10.1.2.1 General procedure for the synthesis of pyrido[1,2-*a*]pyrimidine compounds

P1: One molar equivalent of *N*-(9-bromo-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-4-yl)acetamide (**27**) was dissolved in 5ml mixture of 1,4-dioxane:dimethylacetamide (10:1) and 1.2 molar equivalents of boronic acid, 0.1 molar equivalents of tetrakis(triphenylphosphine) palladium(0) and 2 molar equivalents of cesium carbonate were added. The mixture was heated at 160°C for 10min with continuous stirring in a CEM microwave reactor (power= 300watts, pressure= 300psi, ramp time= 5min). The resulting mixture was cooled down to room temperature and filtered through Celite. The solvent was removed *in vacuo* and product purified using

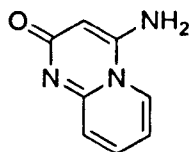
preparative TLC unless otherwise stated using methanol: dichloromethane (1:10) to yield the products as white/cream solids.

10.1.2.2 General procedure for amide coupling reaction using CDMT

P2: One molar equivalent of acid was dissolved in DMF (10ml). 1.5 molar equivalents of 2-chloro,4,6-dimethoxy-1,3,5-triazine (CDMT) and 1.5 molar equivalents of 4-methylmorpholine were subsequently added. The reaction mixture was stirred at room temperature for 1h after which 1.2 molar equivalents of amine was added. The reaction mixture was stirred at room temperature for further 2h. Water (25ml) was added to the mixture. If product precipitated out, the reaction mixture was filtered and purified using preparative TLC. If product did not precipitate out, the product was extracted using ethyl acetate (2x30ml) and subsequently dried over magnesium sulphate and the solvent evaporated *in vacuo*. The compound was subsequently purified by preparative TLC using methanol:dichloromethane (1:10) system.

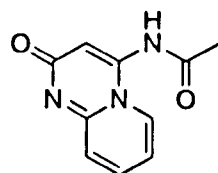
10.1.3 Synthesized compounds

4-amino-2*H*-pyrido[1,2-*a*]pyrimidin-2-one (**8a**):



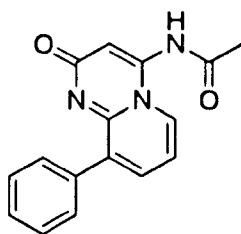
To a stirred solution of ethyl cyanoacetate (2.03g, 0.018mol) in ethanol (15ml) was added 2-aminopyridine (1.41g, 0.015mol). The resulting mixture was heated at 135°C for 2h. The reaction was cooled down to room temperature and product subsequently filtered and washed using ethanol (10ml) and ether (30ml). This solid product was then added to a solution of 1.25M HCl in EtOH (10ml, 0.01mol) in ethanol (20ml). The resulting reaction mixture was heated for 1.5h at 110°C and allowed to cool down to RT. The compound was precipitated out and subsequently filtered and neutralised using NaHCO₃. The resulting product was washed using EtOH (20ml) and ether (20ml) to afford **8a** as a white solid (1.90g, yield 79%) [m.p. dec >150°C]. Anal. HPLC: t_R = 2.8 min, 10-100% MeCN over 10min, purity 100%); ¹H NMR (400 MHz, DMSO-*d*₆): δ/ppm 8.19 (1H, dd, J_a =1.40Hz, J_b =7.08Hz, Ar-H), 7.65 (1H, ddd, J_a =1.40Hz, J_b =6.60Hz, J_c =8.48Hz, Ar-H), 7.11 (1H, dd, J_a =1.46Hz, J_b =8.48Hz, Ar-H), 6.94 (1H, ddd, J_a =1.40Hz, J_b =6.60Hz, J_c =8.48Hz, Ar-H), 6.81 (2H, br, NH₂), 5.83 (1H, s, Ar-H); HRMS (ES) m/z : found 162.0679 (C₈H₈N₃O [M+H]⁺), requires 162.0667.

N-(2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-4-yl)acetamide (**8e**)



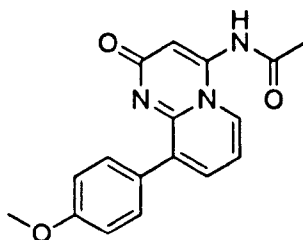
To a stirred solution of acetic anhydride (1014mg, 9.94mmol) in anhydrous pyridine (20ml) was added compound **8a** (200mg, 1.24mmol). The reaction mixture was heated at 110°C for 2h and subsequently cooled down to room temperature. The product was precipitated out and filtered. The resulting product was washed using ether (25ml) to give **8e** as a cream solid (125.5mg, yield 50%) (MeOH:DCM 1:10; R_f = 0.10) [mp. dec >250°C]; Anal. HPLC: t_R = 2.9 min, 10-100% MeCN over 10min, purity 96%); ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 8.97 (1H, dd, J_a =1.38Hz, J_b =7.08Hz, Ar-H), 8.37 (1H, ddd, J_a =1.38Hz, J_b =6.58Hz, J_c =8.50Hz, Ar-H), 7.87 (1H, dd, J_a =1.38Hz, J_b =7.08Hz, Ar-H), 7.58 (1H, ddd, J_a =1.38Hz, J_b =6.58Hz, J_c =8.50Hz, Ar-H), 6.85 (1H, s, Ar-H), 6.29 (1H, br, NH), 2.23 (3H, s, CH₃); ^{13}C NMR (500 MHz, DMSO- d_6): δ /ppm 172.7, 168.8, 154.4, 152.6, 144.5, 135.8, 129.8, 127.4, 112.6, 21.6; HRMS (ES) m/z : found 204.0786 (C₁₀H₁₀N₃O₂ [M+H]⁺), requires 204.0773.

N-(2-oxo-9-phenyl-2*H*-pyrido[1,2-*a*]pyrimidin-4-yl)acetamide (**18**):



Following procedure **P1**, **27** (75mg, 0.27mmol) was dissolved in dioxane:DMA 10:1 (4ml). Benzeneboronic acid (38mg, 0.31 mmol), Pd(PPh₃)₄ (30mg, 0.03mmol) and Cs₂CO₃ (170mg, 0.52mmol) were added. Product **18** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.29) and obtained as a white solid (16.5mg, yield 22%) [mp. dec >250°C]; Anal. HPLC: *t*_R = 6.7 min, 10-100% MeCN over 11min, purity 95%); ¹H NMR (400 MHz, DMSO-*d*₆): δ/ppm 10.63 (1H, br, N-H), 8.31 (1H, d, *J*=7.08Hz, Ar-H), 7.72 (1H, dd, *J*_a=1.52Hz, *J*_b=7.08Hz, Ar-H), 7.58 (2H, dd, *J*_a=1.52Hz, *J*_b= 8.08Hz, Ar-H), 7.44 (3H, m, Ar-H), 7.03 (1H, t, *J*=7.08Hz, Ar-H), 6.33 (1H, s, Ar-H), 2.18 (3H, s, CH₃); ¹³C NMR (500 MHz, DMSO-*d*₆): δ/ppm 170.7, 167.9, 150.8, 142.2, 137.2, 136.9, 134.4, 130.3, 129.2, 128.3, 128.8, 112.5, 110.1, 23.7; HRMS (ES) *m/z*: found 280.1070 (C₁₆H₁₄N₃O₂ [M+H]⁺), requires 280.1086.

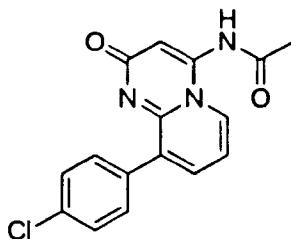
N-(9-(4-methoxyphenyl)-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-4-yl)acetamide (**18b**)



Following procedure **P1**, **27** (75mg, 0.27mmol) was dissolved in dioxane:DMA 10:1 (4ml). 4-methoxybenzeneboronic acid (40mg, 0.27 mmol), Pd(PPh₃)₄ (30mg,

0.03mmol) and Cs_2CO_3 (170mg, 0.52mmol) were added. Product **18b** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.33) and obtained as a white solid (20mg, yield 24%) [mp. dec $>250^\circ\text{C}$]; Anal. HPLC: t_R = 6.9 min, 10-100% MeCN over 11min, purity 98%; ^1H NMR (400 MHz, $\text{DMSO}-d_6$): δ /ppm 10.36 (1H, br, N-H), 8.29 (1H, d, $J=7.08\text{Hz}$, Ar-H), 7.69 (1H, dd, $J_a=1.52\text{Hz}$, $J_b=7.08\text{Hz}$, Ar-H), 7.55 (2H, d, $J=8.84\text{Hz}$, Ar-H), 7.01 (3H, m, Ar-H), 6.31 (1H, s, Ar-H), 3.82 (3H, s, CH_3), 2.19 (3H, s, CH_3); ^{13}C NMR (500 MHz, $\text{DMSO}-d_6$): δ /ppm 170.7, 168.0, 159.5, 150.9, 142.3, 137.3, 136.2, 133.9, 131.6, 129.4, 128.6, 113.7, 112.6, 58.2, 23.4; HRMS (ES) m/z : found 310.1181 ($\text{C}_{17}\text{H}_{16}\text{N}_3\text{O}_3$ $[\text{M}+\text{H}]^+$), requires 310.1192.

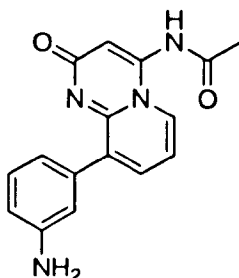
N-(9-(4-chlorophenyl)-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-4-yl)acetamide (**18c**)



Following procedure **P1**, **27** (60mg, 0.21mmol) was dissolved in dioxane:DMA 10:1 (4ml). 4-chlorophenylboronic acid (40mg, 0.26 mmol), $\text{Pd}(\text{PPh}_3)_4$ (25mg, 0.02mmol) and Cs_2CO_3 (140mg, 0.43mmol) were added. Product **18c** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.32) and obtained as a white solid (8.7mg, yield 13%) [mp. dec $>250^\circ\text{C}$]; Anal. HPLC: t_R = 9.4 min, 10-100% MeCN over 12min, purity 100%; ^1H NMR (400 MHz, $\text{DMSO}-d_6$): δ /ppm 10.76 (1H, br, N-H), 8.40 (1H, d, $J=7.08\text{Hz}$, Ar-H), 7.73 (1H, dd, $J_a=1.52\text{Hz}$, $J_b=7.08\text{Hz}$, Ar-H), 7.63 (2H, d, $J=8.59\text{Hz}$, Ar-H), 7.51 (2H, d, $J=8.59\text{Hz}$, Ar-H), 7.00 (1H, t, $J=7.08\text{Hz}$, Ar-H), 6.36 (1H, s, Ar-H),

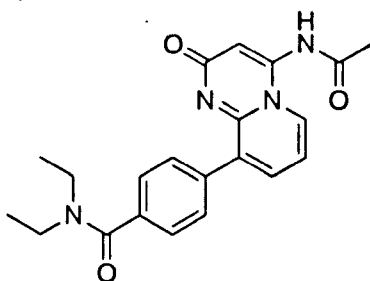
2.16 (3H, s, CH₃); HRMS (ES) *m/z*: found 314.0686 (C₁₆H₁₃ClN₃O₂ [M+H]⁺), requires 314.0696.

N-(9-(3-aminophenyl)-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-4-yl)acetamide (**18d**)



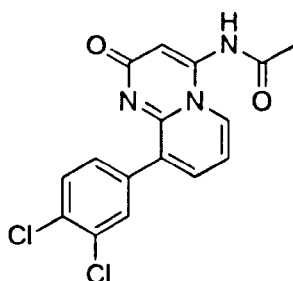
Following procedure **P1**, **27** (500mg, 1.77mmol) was dissolved in dioxane:DMA 10:1 (15ml). 3-aminobenzeneboronic acid (291mg, 2.13mmol), Pd(PPh₃)₄ (205mg, 0.18mmol) and Cs₂CO₃ (1156mg, 3.54mmol) were added. Product **18d** was purified using flash chromatography (MeOH:DCM 1:10; R_f = 0.06) and obtained as a cream solid (170mg, yield 33%) [mp. 201.2-203.3°C]; Anal. HPLC: *t*_R = 3.7 min, 10-100% MeCN over 12min, purity 90%); ¹H NMR (400 MHz, DMSO-*d*₆): δ/ppm 11.52 (1H, br, N-H), 8.86 (2H, br, NH₂), 8.71 (1H, d, *J*=7.08Hz, Ar-H), 7.60 (1H, m, Ar-H), 7.47 (1H, dd, *J*_a=1.52Hz, *J*_b=6.82Hz, Ar-H), 7.06 (1H, d, *J*=8.84Hz, Ar-H), 6.87 (2H, m, Ar-H), 6.61 (1H, s, Ar-H), 6.50 (1H, s, Ar-H), 2.06 (3H, s, CH₃); ¹³C NMR (500 MHz, DMSO-*d*₆): δ/ppm 170.2, 169.4, 150.4, 147.3, 144.4, 136.7, 136.3, 132.1, 129.2, 128.0, 127.9, 123.5, 113.8, 111.5, 111.1, 23.5; HRMS (ES) *m/z*: found 295.0253 (C₁₆H₁₅N₄O₂ [M+H]⁺), requires 295.1195.

4-(4-acetamido-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-9-yl)-*N,N*-diethylbenzamide (**18e**)



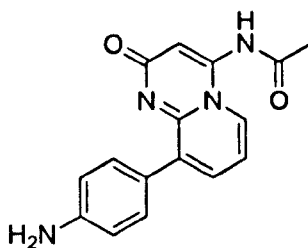
Following procedure **P1**, **27** (60mg, 0.21mmol) was dissolved in dioxane:DMA 10:1 (4ml). 4-(*N,N*-diethylaminocarbonyl)phenylboronic acid (56mg, 0.26mmol), Pd(PPh₃)₄ (26mg, 0.02mmol) and Cs₂CO₃ (139mg, 0.43mmol) were added. Product **18e** was purified using preparative TLC (MeOH:DCM 1:10; *R*_f = 0.33) and obtained as a white solid (18.8mg, yield 23%) [mp. dec >250°C]; Anal. HPLC: *t*_R = 8.4 min, 10-100% MeCN over 11min, purity 96%); ¹H NMR (400 MHz, DMSO-*d*₆): δ/ppm 10.74 (1H, br, N-H), 8.55 (1H, d, *J*=7.08Hz, Ar-H), 7.72 (1H, dd, *J*_a=1.52Hz, *J*_b=7.08Hz, Ar-H), 7.67 (2H, d, *J*=8.34Hz, Ar-H), 7.40 (2H, d, *J*=8.34Hz, Ar-H), 6.98 (1H, t, *J*=7.08Hz, Ar-H), 6.43 (1H, s, Ar-H), 3.45 (4H, br, CH₂), 2.14 (3H, s, CH₃), 1.15 (6H, br, CH₃); ¹³C NMR (500 MHz, DMSO-*d*₆): δ/ppm 171.6, 170.2, 168.4, 161.6, 150.5, 138.1, 137.9, 137.1, 133.4, 130.4, 129.3, 128.9, 128.3, 126.1, 112.3, 109.4, 43.3, 24.4, 13.3; HRMS (ES) *m/z*: found 379.1768 (C₂₁H₂₃N₄O₃ [M+H]⁺), requires 379.1770.

N-(9-(3,4-dichlorophenyl)-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-4-yl)acetamide (**18f**)



Following procedure **P1**, **27** (60mg, 0.21mmol) was dissolved in dioxane:DMA 10:1 (4ml). 3,4-dichlorobenzeneboronic acid (49mg, 0.26 mmol), Pd(PPh₃)₄ (25mg, 0.02mmol) and Cs₂CO₃ (139mg, 0.43mmol) were added. Product **18f** was purified using preparative TLC (MeOH:DCM 1:10; *R*_f = 0.33) and obtained as a white solid (8mg, yield 11%) [mp. dec >230°C]; Anal. HPLC: *t*_R = 7.8 min, 10-100% MeCN over 11min, purity 95%); ¹H NMR (400 MHz, DMSO-*d*₆): δ/ppm 10.52 (1H, br, N-H), 8.57 (1H, d, *J*=7.08Hz, Ar-H), 7.89 (1H, d, *J*=2.02Hz, Ar-H), 7.75 (1H, dd, *J*_a=1.52Hz, *J*_b=7.08Hz, Ar-H), 7.72 (1H, d, *J*=8.34Hz, Ar-H), 7.59 (1H, dd, *J*_a=2.27Hz, *J*_b=8.34Hz, Ar-H), 6.98 (1H, t, *J*= 7.08Hz, Ar-H), 6.44 (1H, s, Ar-H), 2.13 (3H, s, CH₃); HRMS (ES) *m/z*: found 348.0295 (C₁₆H₁₂Cl₂N₃O₂ [M+H]⁺), requires 348.0307.

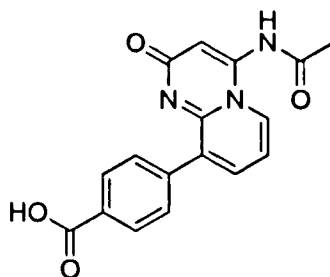
N-(9-(4-aminophenyl)-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-4-yl)acetamide (**18g**)



Compound **27** (1600mg, 5.67mmol) was added to an Ace sealed pressure tube containing dioxane:DMA 10:1 (20ml). 4-aminophenylboronic acid pinacol ester

(1491mg, 6.81mmol), Pd(PPh₃)₄ (655mg, 0.57mmol) and Cs₂CO₃ (3699mg, 11.34mmol) were added. The reaction mixture was stirred at 110°C for 24h. The resulting compound was filtered through Celite and solvent dried *in vacuo*. Product **18g** was purified using flash chromatography (MeOH:DCM 1:10; R_f = 0.08) and obtained as a cream solid (683mg, yield 41%) [mp. 224.3-226.1°C]; Anal. HPLC: t_R = 4.3 min, 10-100% MeCN over 8min, purity 92%); ¹H NMR (400 MHz, DMSO-d₆): δ/ppm 10.88 (1H, br, N-H), 9.98 (2H, br, NH₂), 8.23 (1H, d, J= 7.08Hz, Ar-H), 7.61 (1H, dd, J_a=1.52Hz, J_b=7.08Hz, Ar-H), 7.32 (2H, d, J=8.34Hz, Ar-H), 6.97 (1H, t, J=7.08Hz, Ar-H), 6.61 (2H, d, J=8.34Hz, Ar-H), 6.32 (1H, s, Ar-H), 2.19 (3H, s, CH₃); HRMS (ES) m/z: found 295.0310 (C₁₆H₁₅N₄O₂ [M+H]⁺), requires 295.1195.

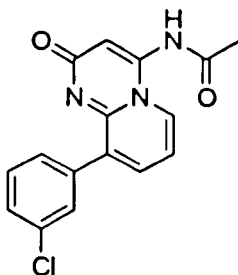
4-(4-acetamido-2-oxo-2H-pyrido[1,2-a]pyrimidin-9-yl)benzoic acid (**18h**)



Following procedure **P1**, **27** (600mg, 2.13mmol) was dissolved in dioxane:DMA 10:1 (15ml). 4-carboxybenzeneboronic acid (424mg, 2.55mmol), Pd(PPh₃)₄ (246mg, 0.21mmol) and Cs₂CO₃ (1387mg, 4.26mmol) were added. Product **18h** was purified using flash chromatography (MeOH:DCM 1:10; R_f = 0.06) and obtained as a cream solid (210mg, yield 31%) [mp. dec >250°C]; Anal. HPLC: t_R = 6.2 min, 10-100% MeCN over 8min, purity 91%); ¹H NMR (400 MHz, DMSO-d₆): δ/ppm 10.22 (2H, br, N-H, COOH), 9.07 (1H, d, J=7.08Hz, Ar-H), 7.94 (2H, d, J=8.34Hz, Ar-H), 7.57 (1H,

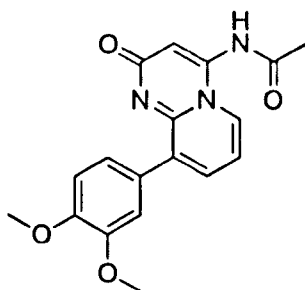
dd, $J_a=1.52\text{Hz}$, $J_b=7.08\text{Hz}$, Ar-H), 7.54 (2H, d, $J=8.34\text{Hz}$, Ar-H), 6.86 (1H, t, $J=7.08\text{Hz}$, Ar-H), 6.67 (1H, s, Ar-H), 2.02 (3H, s, CH_3); HRMS (ES) m/z : found 324.0975 ($\text{C}_{17}\text{H}_{14}\text{N}_3\text{O}_4$ $[\text{M}+\text{H}]^+$), requires 324.0984.

N-(9-(3-chlorophenyl)-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-4-yl)acetamide (**18i**)



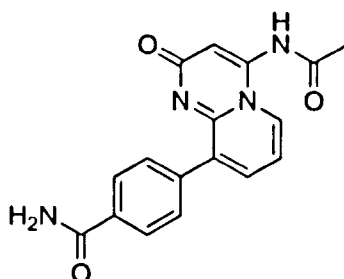
Following procedure **P1**, **27** (60mg, 0.21mmol) was dissolved in dioxane:DMA 10:1 (4ml). 3-chlorobenzeneboronic acid (40mg, 0.26mmol), $\text{Pd}(\text{PPh}_3)_4$ (25mg, 0.02mmol) and Cs_2CO_3 (139mg, 0.43mmol) were added. Product **18i** was purified using preparative TLC (MeOH:DCM 1:10; $R_f=0.35$) and obtained as a white solid (7.4mg, yield 11%) [mp. dec $>250^\circ\text{C}$]; Anal. HPLC: $t_R = 7.3$ min, 10-100% MeCN over 10min, purity 90%); ^1H NMR (400 MHz, $\text{DMSO}-d_6$): δ/ppm 10.58 (1H, br, N-H), 8.35 (1H, d, $J=7.08\text{Hz}$, Ar-H), 7.73 (1H, dd, $J_a=1.52\text{Hz}$, $J_b=7.08\text{Hz}$, Ar-H), 7.46 (3H, m, Ar-H), 7.28 (1H, m, Ar-H), 7.00 (1H, t, $J=7.08\text{Hz}$, Ar-H), 6.35 (1H, s, Ar-H), 2.18 (3H, s, CH_3); HRMS (ES) m/z : found 314.0674 ($\text{C}_{16}\text{H}_{13}\text{ClN}_3\text{O}_2$ $[\text{M}+\text{H}]^+$), requires 314.0696.

N-(9-(3,4-dimethoxyphenyl)-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-4-yl)acetamide (**18j**)



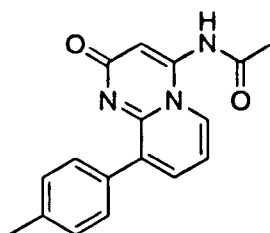
Following procedure **P1**, **27** (60mg, 0.21mmol) was dissolved in dioxane:DMA 10:1 (4ml). 3,4-dimethoxybenzeneboronic acid (46mg, 0.26mmol), Pd(PPh₃)₄ (25mg, 0.02mmol) and Cs₂CO₃ (139mg, 0.43mmol) were added. Product **18j** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.26) and obtained as a brown solid (11.8mg, yield 16%) [mp. dec >250°C]; Anal. HPLC: *t*_R = 6.6 min, 10-100% MeCN over 12min, purity 96%); ¹H NMR (400 MHz, DMSO-*d*₆): δ/ppm 10.16 (1H, br, N-H), 9.28 (1H, dd, *J*_a=1.52Hz, *J*_b=7.33Hz, Ar-H), 7.50 (1H, dd, *J*_a=1.52Hz, *J*_b=6.82Hz, Ar-H), 7.22 (1H, d, *J*=2.02Hz, Ar-H), 7.15 (1H, dd, *J*_a=2.02Hz, *J*_b=8.34Hz, Ar-H), 6.99 (1H, d, *J*=8.34Hz, Ar-H), 6.77 (1H, t, *J*=7.08Hz, Ar-H), 6.76 (1H, s, Ar-H), 3.80 (3H, s, OCH₃), 3.77 (3H, s, OCH₃), 1.95 (3H, s, CH₃); HRMS (ES) *m/z*: found 340.1275 (C₁₈H₁₈N₃O₄ [M+H]⁺), requires 340.1297.

4-(4-acetamido-2-oxo-2H-pyrido[1,2-a]pyrimidin-9-yl)benzamide (**18k**)



Following procedure **P1**, **27** (60mg, 0.21mmol) was dissolved in dioxane:DMA 10:1 (4ml). 4-aminocarbonylbenzeneboronic acid (42mg, 0.26mmol), Pd(PPh₃)₄ (25mg, 0.02mmol) and Cs₂CO₃ (139mg, 0.43mmol) were added. Product **18k** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.10) and obtained as a white solid (12.3mg, yield 18%) [mp. dec >250°C]; Anal. HPLC: *t*_R = 5.8 min, 10-100% MeCN over 12min, purity 90%); ¹H NMR (400 MHz, DMSO-*d*₆): δ/ppm 10.63 (1H, br, N-H), 9.27 (2H, br, NH₂), 8.51 (1H, d, *J*=7.08Hz, Ar-H), 7.68 (1H, dd, *J*_a=1.52Hz, *J*_b=7.08Hz, Ar-H), 7.63 (2H, d, *J*=8.34Hz, Ar-H), 7.45 (2H, d, *J*=8.34Hz, Ar-H), 7.00 (1H, t, *J*=7.08Hz, Ar-H), 6.44 (1H, s, Ar-H), 2.18 (3H, s, CH₃); HRMS (ES) *m/z*: found 323.1133 (C₁₇H₁₅N₄O₃ [M+H]⁺), requires 323.1144.

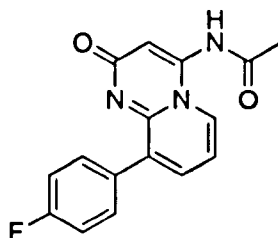
N-(2-oxo-9-(p-tolyl)-2H-pyrido[1,2-a]pyrimidin-4-yl)acetamide (**18l**)



Following procedure **P1**, **27** (60mg, 0.21mmol) was dissolved in dioxane:DMA 10:1 (4ml). 4-methylphenylboronic acid (35mg, 0.26mmol), Pd(PPh₃)₄ (25mg, 0.02mmol)

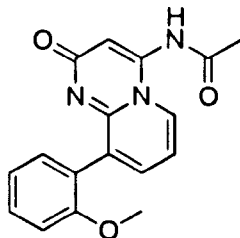
and Cs_2CO_3 (139mg, 0.43mmol) were added. Product **18l** was purified using preparative TLC (MeOH:DCM 1:10; $R_f=0.31$) and obtained as a white solid (19.8mg, yield 32%) [mp. dec $>250^\circ\text{C}$]; Anal. HPLC: $t_R = 13.7$ min, 10-100% MeCN over 20min, purity 100%); ^1H NMR (400 MHz, $\text{DMSO}-d_6$): δ/ppm 10.57 (1H, br, N-H), 8.30 (1H, d, $J=7.08\text{Hz}$, Ar-H), 7.69 (1H, dd, $J_a=1.52\text{Hz}$, $J_b=7.08\text{Hz}$, Ar-H), 7.48 (2H, d, $J=8.08\text{Hz}$, Ar-H), 7.26 (2H, d, $J=8.08\text{Hz}$, Ar-H), 7.00 (1H, t, $J=7.08\text{Hz}$, Ar-H), 6.32 (1H, s, Ar-H), 2.37 (3H, s, CH_3), 2.18 (3H, s, CH_3); HRMS (ES) m/z : found 294.1236 ($\text{C}_{17}\text{H}_{16}\text{N}_3\text{O}_2$ $[\text{M}+\text{H}]^+$), requires 294.1243.

N-(9-(4-fluorophenyl)-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-4-yl)acetamide (**18m**)



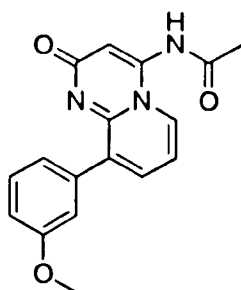
Following procedure **P1**, **27** (60mg, 0.21mmol) was dissolved in dioxane:DMA 10:1 (4ml). 4-fluorophenylboronic acid (36mg, 0.26mmol), $\text{Pd}(\text{PPh}_3)_4$ (25mg, 0.02mmol) and Cs_2CO_3 (139mg, 0.43mmol) were added. Product **18m** was purified using preparative TLC (MeOH:DCM 1:10; $R_f=0.35$) and obtained as a white solid (13mg, yield 21%) [mp. dec $>230^\circ\text{C}$]; Anal. HPLC: $t_R = 6.9$ min, 10-100% MeCN over 12min, purity 97%); ^1H NMR (400 MHz, $\text{DMSO}-d_6$): δ/ppm 10.69 (1H, br, N-H), 8.45 (1H, d, $J=7.08\text{Hz}$, Ar-H), 7.92 (1H, dd, $J_a=1.52\text{Hz}$, $J_b=7.08\text{Hz}$, Ar-H), 7.73 (2H, d, $J=8.36\text{Hz}$, Ar-H), 7.48 (2H, d, $J=8.36\text{Hz}$, Ar-H), 7.02 (1H, t, $J=7.08\text{Hz}$, Ar-H), 6.41 (1H, s, Ar-H), 2.16 (3H, s, CH_3); HRMS (ES) m/z : found 298.0984 ($\text{C}_{16}\text{H}_{13}\text{FN}_3\text{O}_2$ $[\text{M}+\text{H}]^+$), requires 298.0992.

N-(9-(2-methoxyphenyl)-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-4-yl)acetamide (**18n**)



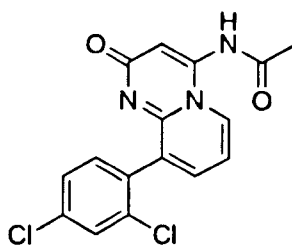
Following procedure **P1**, **27** (75mg, 0.27mmol) was dissolved in dioxane:DMA 10:1 (4ml). 2-methoxybenzeneboronic acid (48mg, 0.32mmol), Pd(PPh₃)₄ (31mg, 0.03mmol) and Cs₂CO₃ (173mg, 0.53mmol) were added. Product **18n** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.29) and obtained as a brown solid (18.5mg, yield 23%) [mp. 220.5-222.1°C]; Anal. HPLC: *t*_R = 6.7 min, 10-100% MeCN over 12min, purity 96%); ¹H NMR (400 MHz, DMSO-*d*₆): δ/ppm 10.78 (1H, br, N-H), 8.86 (1H, d, *J*=7.08Hz, Ar-H), 7.54 (1H, dd, *J*_a=1.52Hz, *J*_b=7.08Hz, Ar-H), 7.39 (1H, m, Ar-H), 7.20 (1H, dd, *J*_a=1.78Hz, *J*_b=7.58Hz, Ar-H), 7.10 (1H, d, *J*=7.58Hz, Ar-H), 6.94 (1H, t, *J*=7.33Hz, Ar-H), 6.71 (1H, t, *J*=7.08Hz, Ar-H), 6.58 (1H, s, Ar-H), 3.69 (3H, s, CH₃), 2.14 (3H, s, CH₃); ¹³C NMR (500 MHz, DMSO-*d*₆): δ/ppm 171.7, 169.4, 157.2, 150.9, 148.9, 139.4, 137.2, 132.6, 129.2, 126.8, 120.5, 117.3, 111.9, 111.1, 108.6, 55.9, 26.7; HRMS (ES) *m/z*: found 309.9998 (C₁₇H₁₆N₃O₃ [M+H]⁺), requires 310.1192.

N-(9-(3-methoxyphenyl)-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-4-yl)acetamide (**18o**)



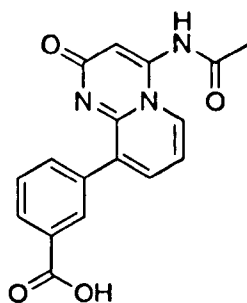
Following procedure **P1**, **27** (75mg, 0.27mmol) was dissolved in dioxane:DMA 10:1 (4ml). 3-methoxybenzeneboronic acid (48mg, 0.32mmol), Pd(PPh₃)₄ (31mg, 0.03mmol) and Cs₂CO₃ (173mg, 0.53mmol) were added. Product **18o** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.29) and obtained as a cream solid (22.2mg, yield 27%) [mp. 245.5-247.1°C]; Anal. HPLC: *t*_R = 6.8 min, 10-100% MeCN over 12min, purity 97%); ¹H NMR (400 MHz, DMSO-*d*₆): δ/ppm 10.54 (1H, br, N-H), 8.65 (1H, d, *J*=7.08Hz, Ar-H), 7.64 (1H, dd, *J*_a=1.52Hz, *J*_b=7.08Hz, Ar-H), 7.34 (1H, t, *J*=8.34Hz, Ar-H), 7.13 (2H, m, Ar-H), 6.96 (1H, m, Ar-H), 6.92 (1H, t, *J*=7.08Hz, Ar-H), 6.47 (1H, s, Ar-H), 3.79 (3H, s, CH₃), 2.09 (3H, s, CH₃); ¹³C NMR (500 MHz, DMSO-*d*₆): δ/ppm 172.3, 168.7, 159.1, 150.6, 145.3, 138.9, 136.3, 134.0, 129.2, 129.0, 127.7, 116.3, 113.5, 111.7, 106.7, 55.6, 25.6; HRMS (ES) *m/z*: found 310.0063 (C₁₇H₁₆N₃O₃ [M+H]⁺), requires 310.1192.

N-(9-(2,4-dichlorophenyl)-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-4-yl)acetamide (**18p**)



Following procedure **P1**, **27** (60mg, 0.21mmol) was dissolved in dioxane:DMA 10:1 (4ml). 2,4-dichlorobenzeneboronic acid (49mg, 0.26mmol), Pd(PPh₃)₄ (25mg, 0.02mmol) and Cs₂CO₃ (139mg, 0.43mmol) were added. Product **18p** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.27) and obtained as a cream solid (7.6mg, yield 10%) [mp. dec >250°C]; Anal. HPLC: *t*_R = 7.6 min, 10-100% MeCN over 12min, purity 99%); ¹H NMR (400 MHz, DMSO-*d*₆): δ/ppm 10.39 (1H, br, N-H), 8.89 (1H, d, *J*=7.08Hz, Ar-H), 7.65 (1H, m, Ar-H), 7.56 (1H, dd, *J*_a=1.52Hz, *J*_b=7.08Hz, Ar-H), 7.50 (1H, m, Ar-H), 7.42 (1H, d, *J*=8.34Hz, Ar-H), 6.93 (1H, t, *J*=7.08Hz, Ar-H), 6.54 (1H, s, Ar-H), 2.09 (3H, s, CH₃); HRMS (ES) *m/z*: found 348.0294 (C₁₆H₁₂Cl₂N₃O₂ [M+H]⁺), requires 348.0307.

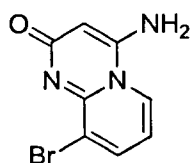
3-(4-acetamido-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-9-yl)benzoic acid (**18q**)



Compound **27** (500mg, 1.77mmol) was added to an Ace sealed pressure tube containing dioxane:DMA 10:1 (15ml). 3-carboxybenzeneboronic acid (353mg,

2.13mmol), Pd(PPh₃)₄ (205mg, 0.18mmol) and Cs₂CO₃ (1156mg, 3.55mmol) were subsequently added. The reaction mixture was stirred at 110°C for 5h and filtered through Celite and solvent dried *in vacuo*. Product **18q** was purified using flash chromatography (MeOH:DCM 1:10; R_f = 0.04) and obtained as a cream solid (150mg, yield 26%) [mp. dec >250°C]; Anal. HPLC: t_R = 6.3 min, 10-100% MeCN over 12min, purity 95%; ¹H NMR (400 MHz, DMSO-d₆): δ/ppm 9.63 (1H, br, COOH), 8.58 (1H, br, N-H), 7.93 (1H, d, J=7.08Hz, Ar-H), 7.66 (2H, m, Ar-H), 7.44 (1H, t, J=7.83Hz, Ar-H), 7.34 (1H, m, Ar-H), 7.35 (1H, s, Ar-H), 6.99 (1H, t, J=7.08Hz, Ar-H), 6.43 (1H, s, Ar-H), 2.16 (3H, s, CH₃); HRMS (ES) m/z: found 324.0976 (C₁₇H₁₄N₃O₄ [M+H]⁺), requires 324.0984.

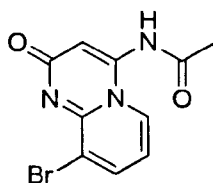
4-amino-9-bromo-2H-pyrido[1,2-a]pyrimidin-2-one (**26**):



To a stirred solution of cyanoacetic acid (12.92g, 0.15mol) in DCM (220ml) was added 2-amino-3-bromopyridine (21.04g, 0.12mol) and dicyclohexylcarbodiimide (31.34g, 0.15mol). The resulting mixture was refluxed at 110°C for 2h. The resulting reaction was cooled down to room temperature and product subsequently filtered and washed using DCM (30ml) and ether (30ml). This solid product was then added to a solution of 1.25M HCl in EtOH (52ml, 0.06mol) in ethanol (85ml). The resulting reaction mixture was heated for 1.5h at 110°C and allowed to cool down to RT. The compound was precipitated out and subsequently filtered and neutralised using NaHCO₃. The resulting product was washed using ethanol (20ml) and ether (20ml) to

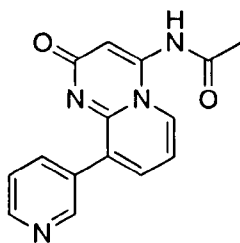
afford **26** as a cream solid (15.2g, yield 52%) [m.p. 263.1-265.2°C] Anal. HPLC: t_R = 7.9 min, 10-100% MeCN over 10min, purity 100%); ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 9.09 (2H, s, NH_2), 8.92 (1H, dd, $J_a=0.76\text{Hz}$, $J_b=7.08\text{Hz}$, Ar-H), 8.51 (1H, dd, $J_a=0.76\text{Hz}$, $J_b=7.33\text{Hz}$, Ar-H), 7.40 (1H, t, $J=7.33\text{Hz}$, Ar-H), 6.44 (1H, s, Ar-H); HRMS (ES) m/z : found 239.9764 ($\text{C}_8\text{H}_7\text{BrN}_3\text{O}$ $[\text{M}+\text{H}]^+$), requires 239.9772.

N-(9-bromo-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-4-yl)acetamide (**27**):



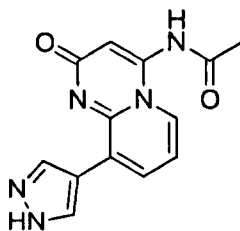
To a stirred solution of acetic anhydride (11.9g, 0.12mol) in anhydrous pyridine (50ml) was added compound **26** (3.5g, 0.015mol). The reaction mixture was heated at 110°C for 2h and subsequently cooled down to room temperature. The product precipitated out and was filtered and washed using ether (25ml) to give **27** as a cream solid (3.2g, yield 78%) (MeOH:DCM 1:10; R_f = 0.19) [mp. dec >270°C]; Anal. HPLC: t_R = 6.2 min, 10-100% MeCN over 10min, purity 98%); ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 11.28 (1H, br, N-H), 8.36 (1H, dd, $J_a=1.26\text{Hz}$, $J_b=7.08\text{Hz}$, Ar-H), 8.19 (1H, dd, $J_a=1.26\text{Hz}$, $J_b=7.08\text{Hz}$, Ar-H), 6.84 (1H, t, $J=7.08\text{Hz}$, Ar-H), 6.34 (1H, s, Ar-H), 2.19 (3H, s, CH_3); ^{13}C NMR (500 MHz, DMSO- d_6): δ /ppm 170.6, 167.9, 148.9, 142.3, 140.4, 129.8, 117.4, 112.4, 111.2, 23.6; HRMS (ES) m/z : found 281.9872 ($\text{C}_{10}\text{H}_9\text{BrN}_3\text{O}_2$ $[\text{M}+\text{H}]^+$), requires 281.9878.

N-(2-oxo-9-(pyridin-3-yl)-2*H*-pyrido[1,2-*a*]pyrimidin-4-yl)acetamide (**28**)



Following procedure **P1**, **27** (60mg, 0.21mmol) was dissolved in dioxane:DMA 10:1 (4ml). Pyridine-3-boronic acid (31mg, 0.26mmol), Pd(PPh₃)₄ (25mg, 0.02mmol) and Cs₂CO₃ (139mg, 0.43mmol) were added. Product **28** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.24) and obtained as a cream solid (6.9mg, yield 12%) [mp. dec >250°C]; Anal. HPLC: *t*_R = 4.8 min, 10-100% MeCN over 12min, purity 100%; ¹H NMR (400 MHz, DMSO-*d*₆): δ/ppm 10.29 (1H, br, N-H), 8.76 (2H, m, Ar-H), 8.57 (1H, dd, *J*_a=1.52Hz, *J*_b=4.80Hz, Ar-H), 8.03 (1H, m, Ar-H), 7.75 (1H, dd, *J*_a=1.26Hz, *J*_b=6.82Hz, Ar-H), 7.47 (1H, m, Ar-H), 6.97 (1H, t, *J*=7.08Hz, Ar-H), 6.52 (1H, s, Ar-H), 2.11 (3H, s, CH₃); ¹³C NMR (500 MHz, DMSO-*d*₆): δ/ppm 173.3, 169.1, 150.5, 148.9, 146.9, 137.9, 133.6, 133.4, 130.8, 129.6, 127.5, 123.3, 111.4, 104.5, 23.5; HRMS (ES) *m/z*: found 281.1030 (C₁₅H₁₃N₄O₂ [M+H]⁺), requires 281.1039.

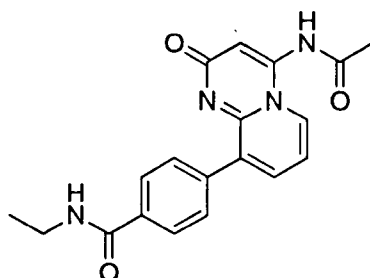
N-(2-oxo-9-(1*H*-pyrazol-4-yl)-2*H*-pyrido[1,2-*a*]pyrimidin-4-yl)acetamide (**29**)



Following procedure **P1**, **27** (60mg, 0.21mmol) was dissolved in dioxane:DMA 10:1 (4ml). 1*H*-pyrazol-3-yl boronic acid (29mg, 0.26mmol), Pd(PPh₃)₄ (25mg, 0.02mmol)

and Cs_2CO_3 (139mg, 0.43mmol) were added. Product **29** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.06) and obtained as a cream solid (9.2mg, yield 16%) [mp. dec $>250^\circ\text{C}$]; Anal. HPLC: t_R = 5.6 min, 10-100% MeCN over 12min, purity 98%); ^1H NMR (400 MHz, DMSO-d_6): δ /ppm 13.08 (1H, br, N-H), 10.57 (1H, br, N-H), 8.52 (3H, m, Ar-H), 8.04 (1H, dd, J_a = 1.52Hz, J_b = 7.08Hz, Ar-H), 6.93 (1H, t, J = 7.08Hz, Ar-H), 6.50 (1H, s, Ar-H), 2.13 (3H, s, CH_3); HRMS (ES) m/z : found 270.0977 ($\text{C}_{13}\text{H}_{12}\text{N}_5\text{O}_2$ $[\text{M}+\text{H}]^+$), requires 270.0991.

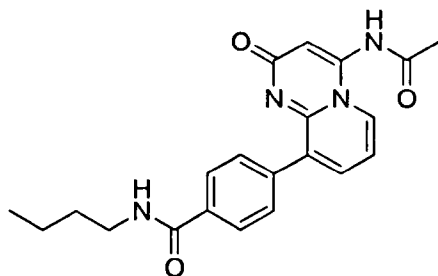
4-(4-acetamido-2-oxo-2H-pyrido[1,2-a]pyrimidin-9-yl)-N-ethylbenzamide (**30a**)



Following procedure **P2**, **18h** (60mg, 0.19mmol) was dissolved in DMF (6ml). Ethylamine (10mg, 0.22mmol), CDMT (49mg, 0.28mmol) and 4-methylmorpholine (28mg, 0.28mmol) were added. Product **30a** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.14) and obtained as a yellow solid (28.7mg, yield 44%) [mp. dec $>270^\circ\text{C}$]; Anal. HPLC: t_R = 6.3 min, 10-100% MeCN over 11min, purity 95%); ^1H NMR (400 MHz, DMSO-d_6): δ /ppm 10.82 (1H, s, N-H), 8.56 (1H, t, J = 5.54Hz, N-H), 8.32 (1H, dd, J_a = 1.02Hz, J_b = 7.08Hz, Ar-H), 7.90 (2H, d, J = 8.44Hz, Ar-H), 7.77 (1H, dd, J_a = 1.12Hz, J_b = 7.08Hz, Ar-H), 7.67 (2H, d, J = 8.44Hz, Ar-H), 7.03 (1H, t, J = 7.08Hz, Ar-H), 6.31 (1H, s, Ar-H), 3.30 (2H, m, CH_2), 2.20 (3H, s, CH_3), 1.15 (3H, t, J = 7.14Hz, CH_3); ^{13}C NMR (500 MHz, DMSO-d_6): δ /ppm 170.8, 167.9, 166.2,

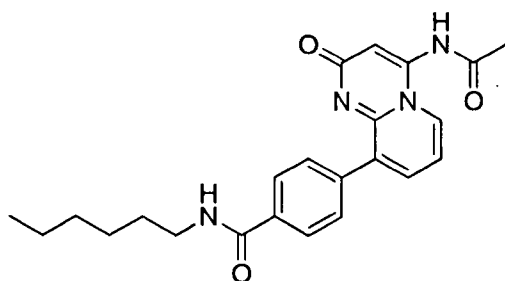
150.7, 142.1, 139.7, 137.2, 134.4, 133.5, 130.2, 129.5, 127.0, 112.6, 111.1, 34.5, 23.6, 15.3; HRMS (ES) m/z : found 351.1473 ($C_{19}H_{19}N_4O_3$ $[M+H]^+$), requires 351.1457.

4-(4-acetamido-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-9-yl)-*N*-butylbenzamide (**30b**)



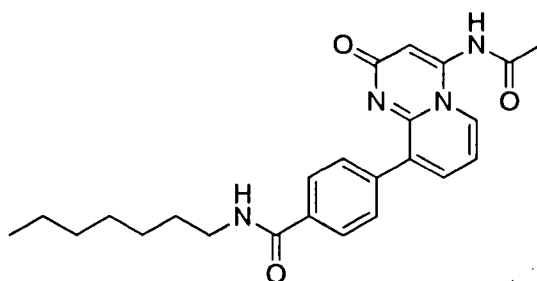
Following procedure **P2**, **18h** (50mg, 0.15mmol) was dissolved in DMF (4ml). Butylamine (14mg, 0.18mmol), CDMT (40.5mg, 0.23mmol) and 4-methylmorpholine (23mg, 0.23mmol) were added. Product **30b** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.16) and obtained as a cream solid (27.8mg, yield 48%) [mp. dec $>250^\circ\text{C}$]; Anal. HPLC: t_R = 6.8 min, 10-100% MeCN over 11min, purity 94%); ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 10.44 (1H, br, N-H), 8.51 (2H, m, N-H, Ar-H), 7.97 (1H, dd, $J_a=1.50\text{Hz}$, $J_b=7.08\text{Hz}$, Ar-H), 7.92 (2H, d, $J=8.48\text{Hz}$, Ar-H), 7.52 (2H, d, $J=8.48\text{Hz}$, Ar-H), 6.96 (1H, t, $J=7.08\text{Hz}$, Ar-H), 6.76 (1H, s, Ar-H), 3.28 (2H, m, CH_2), 2.05 (3H, s, CH_3), 1.52 (2H, m, CH_2), 1.34 (2H, m, CH_2), 0.91 (3H, m, CH_3); ^{13}C NMR (500 MHz, DMSO- d_6): δ /ppm 170.9, 166.3, 166.1, 149.9, 141.4, 138.1, 136.2, 133.8, 133.3, 130.3, 129.1, 128.2, 113.6, 111.4, 40.5, 31.7, 24.9, 20.1, 14.2; HRMS (ES) m/z : found 379.1757 ($C_{21}H_{23}N_4O_3$ $[M+H]^+$), requires 379.1770.

4-(4-acetamido-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-9-yl)-*N*-hexylbenzamide (**30c**)



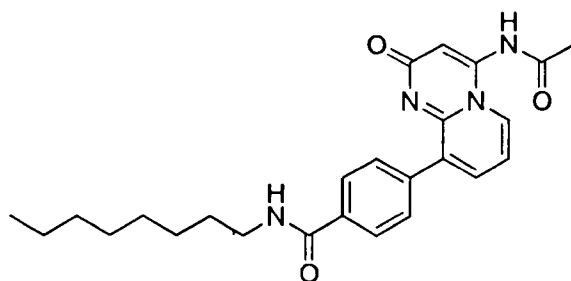
Following procedure **P2**, **18h** (60mg, 0.19mmol) was dissolved in DMF (4ml). Hexylamine (22.5mg, 0.22mmol), CDMT (49mg, 0.28mmol) and 4-methylmorpholine (28mg, 0.28mmol) were added. Product **30c** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.25) and obtained as a cream solid (11.5mg, yield 15%) [mp. dec >250°C]; Anal. HPLC: t_R = 7.9 min, 10-100% MeCN over 11min, purity 100%); ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 10.68 (1H, br, N-H), 8.52 (1H, t, J =5.63Hz, N-H), 8.31 (1H, d, J =7.08Hz, Ar-H), 7.89 (2H, d, J =8.34Hz, Ar-H), 7.77 (1H, dd, J_a =1.50Hz, J_b =7.08Hz, Ar-H), 7.67 (2H, d, J =8.34Hz, Ar-H), 7.03 (1H, t, J =7.08Hz, Ar-H), 6.32 (1H, s, Ar-H), 3.28 (2H, m, CH_2), 2.19 (3H, s, CH_3), 1.54 (2H, m, CH_2), 1.30 (6H, m, CH_2), 0.88 (3H, t, J =6.57Hz, CH_3); HRMS (ES) m/z : found 407.2070 ($\text{C}_{23}\text{H}_{27}\text{N}_4\text{O}_3$ $[\text{M}+\text{H}]^+$), requires 407.2083.

4-(4-acetamido-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-9-yl)-*N*-heptylbenzamide (**30d**)



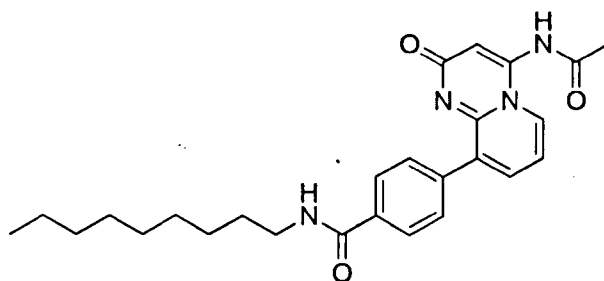
Following procedure **P2**, **18h** (100mg, 0.30mmol) was dissolved in DMF (6ml). Heptylamine (43mg, 0.37mmol), CDMT (81mg, 0.46mmol) and 4-methylmorpholine (47mg, 0.46mmol) were added. Product **30d** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.23) and obtained as a cream solid (5mg, yield 4%) [mp. dec $>250^\circ\text{C}$]; Anal. HPLC: t_R = 8.4 min, 10-100% MeCN over 11min, purity 100%); ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 10.84 (1H, br, N-H), 8.50 (2H, m, N-H, Ar-H), 7.89 (2H, d, $J=8.48\text{Hz}$, Ar-H), 7.73 (1H, dd, $J_a=1.48\text{Hz}$, $J_b=7.08\text{Hz}$, Ar-H), 7.67 (2H, d, $J=8.48\text{Hz}$, Ar-H), 6.99 (1H, t, $J=7.08\text{Hz}$, Ar-H), 6.40 (1H, s, Ar-H), 3.28 (2H, m, CH_2), 2.15 (3H, s, CH_3), 1.54 (2H, m, CH_2), 1.30 (8H, m, CH_2), 0.87 (3H, t, $J=6.92\text{Hz}$, CH_3); ^{13}C NMR (500 MHz, DMSO- d_6): δ /ppm 170.8, 168.3, 166.2, 150.5, 142.2, 139.9, 136.8, 134.4, 133.4, 130.2, 129.4, 127.0, 112.1, 111.1, 40.5, 31.7, 29.6, 28.9, 26.9, 24.6, 22.5, 14.5; HRMS (ES) m/z : found 421.2253 ($\text{C}_{24}\text{H}_{29}\text{N}_4\text{O}_3$ $[\text{M}+\text{H}]^+$), requires 421.2240.

4-(4-acetamido-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-9-yl)-*N*-octylbenzamide (**30e**)



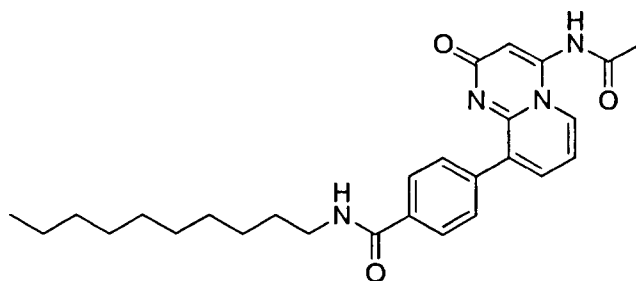
Following procedure **P2**, **18h** (60mg, 0.19mmol) was dissolved in DMF (4ml). Octylamine (29mg, 0.22mmol), CDMT (49mg, 0.28mmol) and 4-methylmorpholine (28mg, 0.28mmol) were added. Product **30e** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.21) and obtained as a cream solid (20.6mg, yield 26%) [mp. dec >250°C]; Anal. HPLC: t_R = 9.1 min, 10-100% MeCN over 11min, purity 92%); ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 10.70 (1H, br, N-H), 8.52 (1H, t, J =5.62Hz, N-H), 8.31 (1H, d, J =7.08Hz, Ar-H), 7.89 (2H, d, J =8.48Hz, Ar-H), 7.77 (1H, dd, J_a =1.50Hz, J_b =7.08Hz, Ar-H), 7.67 (2H, d, J =8.48Hz, Ar-H), 7.03 (1H, t, J =7.08Hz, Ar-H), 6.32 (1H, s, Ar-H), 3.28 (2H, m, CH₂), 2.19 (3H, s, CH₃), 1.54 (2H, m, CH₂), 1.28 (10H, m, CH₂), 0.86 (3H, t, J =6.88Hz, CH₃); ^{13}C NMR (500 MHz, DMSO- d_6): δ /ppm 170.7, 167.9, 166.3, 150.7, 142.1, 139.7, 137.1, 134.5, 133.5, 130.2, 129.5, 127.0, 112.5, 111.1, 40.5, 31.7, 29.6, 29.2, 29.1, 26.9, 23.6, 22.6, 14.5; HRMS (ES) m/z : found 435.2394 (C₂₅H₃₁N₄O₃ [M+H]⁺), requires 435.2396.

4-(4-acetamido-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-9-yl)-*N*-nonylbenzamide (**30f**)



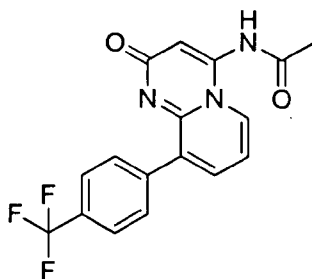
Following procedure **P2**, **18h** (60mg, 0.19mmol) was dissolved in DMF (4ml). Nonylamine (32mg, 0.22mmol), CDMT (49mg, 0.28mmol) and 4-methylmorpholine (28mg, 0.28mmol) were added. Product **30f** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.28) and obtained as a cream solid (14mg, yield 17%) [mp. dec >250°C]; Anal. HPLC: t_R = 9.7 min, 10-100% MeCN over 11min, purity 99%); ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 10.72 (1H, br, N-H), 8.52 (1H, t, J =5.62Hz, N-H), 8.31 (1H, d, J =7.08Hz, Ar-H), 7.89 (2H, d, J =8.48Hz, Ar-H), 7.77 (1H, dd, J_a =1.10Hz, J_b =7.08Hz, Ar-H), 7.67 (2H, d, J =8.48Hz, Ar-H), 7.04 (1H, t, J =7.08Hz, Ar-H), 6.32 (1H, s, Ar-H), 3.28 (2H, m, CH₂), 2.19 (3H, s, CH₃), 1.53 (2H, m, CH₂), 1.26 (12H, m, CH₂), 0.86 (3H, t, J =6.80Hz, CH₃); ^{13}C NMR (500 MHz, DMSO- d_6): δ /ppm 170.7, 167.9, 166.3, 150.7, 142.1, 139.7, 137.1, 134.5, 133.5, 130.2, 129.5, 127.0, 112.5, 111.1, 40.5, 31.7, 29.6, 29.5, 29.3, 29.2, 26.9, 23.6, 22.5, 14.5; HRMS (ES) m/z : found 449.2538 (C₂₆H₃₃N₄O₃ [M+H]⁺), requires 449.2553.

4-(4-acetamido-2-oxo-2H-pyrido[1,2-a]pyrimidin-9-yl)-N-decylbenzamide (**30g**)



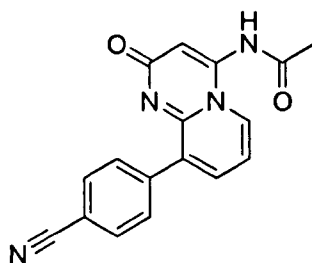
Following procedure **P2**, **18h** (60mg, 0.19mmol) was dissolved in DMF (4ml). Decylamine (35mg, 0.22mmol), CDMT (49mg, 0.28mmol) and 4-methylmorpholine (28mg, 0.28mmol) were added. Product **30g** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.29) and obtained as a cream solid (11.8mg, yield 14%) [mp. dec >250°C]; Anal. HPLC: t_R = 14 min, 10-100% MeCN over 16min, purity 96%); ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 8.85 (1H, br, N-H), 8.51 (1H, t, J =5.62Hz, N-H), 8.44 (1H, d, J =7.08Hz, Ar-H), 7.89 (2H, d, J =8.34Hz, Ar-H), 7.74 (1H, dd, J_a =1.50Hz, J_b =7.08Hz, Ar-H), 7.67 (2H, d, J =8.34Hz, Ar-H), 7.01 (1H, t, J =7.08Hz, Ar-H), 6.38 (1H, s, Ar-H), 3.28 (2H, m, CH₂), 2.16 (3H, s, CH₃), 1.53 (2H, m, CH₂), 1.26 (14H, m, CH₂), 0.86 (3H, t, J =6.58Hz, CH₃); ^{13}C NMR (500 MHz, DMSO- d_6): δ /ppm 170.7, 167.9, 166.2, 150.7, 142.2, 139.7, 137.0, 134.4, 133.5, 130.2, 129.5, 127.0, 112.5, 111.1, 40.9, 31.7, 29.6, 29.5, 29.4, 29.3, 29.2, 26.9, 23.6, 22.5, 14.5; HRMS (ES) m/z : found 463.2716 (C₂₇H₃₅N₄O₃ [M+H]⁺), requires 463.2709.

N-(2-oxo-9-(4-(trifluoromethyl)phenyl)-2*H*-pyrido[1,2-*a*]pyrimidin-4-yl)acetamide (**31a**)



Following procedure **P1**, **27** (100mg, 0.35mmol) was dissolved in dioxane:DMA 10:1 (5ml). 4-trifluoromethylbenzene boronic acid (81mg, 0.43mmol), Pd(PPh₃)₄ (41mg, 0.035mmol) and Cs₂CO₃ (231mg, 0.71mmol) were added. Product **31a** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.21) and obtained as a cream solid (37.8mg, yield 31%) [mp. dec >250°C]; Anal. HPLC: t_R = 7.7 min, 10-100% MeCN over 11min, purity 97%; ¹H NMR (400 MHz, DMSO-*d*₆): δ/ppm 10.73 (1H, br, N-H), 8.55 (1H, d, *J*=7.08Hz, Ar-H), 7.81 (4H, s, Ar-H), 7.76 (1H, dd, *J*_a=1.48Hz, *J*_b=7.08Hz, Ar-H), 7.00 (1H, t, *J*=7.08Hz, Ar-H), 6.42 (1H, s, Ar-H), 2.14 (3H, s, CH₃); HRMS (ES) *m/z*: found 348.0961 (C₁₇H₁₃F₃N₃O₂ [M+H]⁺), requires 348.0960.

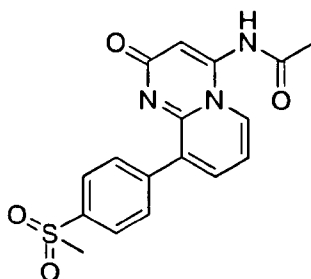
N-(9-(4-cyanophenyl)-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-4-yl)acetamide (**31b**)



Following procedure **P1**, **27** (90mg, 0.32mmol) was dissolved in dioxane:DMA 10:1 (5ml). 4-cyanophenylboronic acid (56mg, 0.38mmol), Pd(PPh₃)₄ (37mg, 0.032mmol)

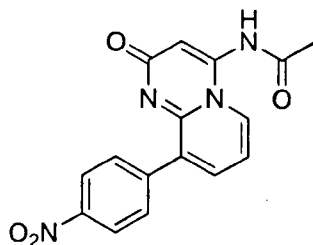
and Cs_2CO_3 (208mg, 0.64mmol) were added. Product **31b** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.27) and obtained as a white solid (9.8mg, yield 9%) [mp. dec $>215^\circ\text{C}$]; Anal. HPLC: t_R = 6.8 min, 10-100% MeCN over 11min, purity 97%); ^1H NMR (400 MHz, $\text{DMSO}-d_6$): δ /ppm 10.21 (1H, br, N-H), 8.37 (1H, d, $J=7.08\text{Hz}$, Ar-H), 7.93 (2H, d, $J=8.48\text{Hz}$, Ar-H), 7.81 (3H, m, Ar-H), 7.05 (1H, t, $J=7.08\text{Hz}$, Ar-H), 6.33 (1H, s, Ar-H), 2.20 (3H, s, CH_3); HRMS (ES) m/z : found 305.1036 ($\text{C}_{17}\text{H}_{13}\text{N}_4\text{O}_2$ $[\text{M}+\text{H}]^+$), requires 305.1039.

N-(9-(4-(methanesulfonyl)phenyl)-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-4-yl)acetamide (**31c**)



Following procedure **P1**, **27** (90mg, 0.32mmol) was dissolved in dioxane:DMA 10:1 (5ml). 4-(methanesulfonyl)phenylboronic acid (77mg, 0.38mmol), $\text{Pd}(\text{PPh}_3)_4$ (37mg, 0.032mmol) and Cs_2CO_3 (208mg, 0.64mmol) were added. Product **31c** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.16) and obtained as a cream solid (48.7mg, yield 43%) [mp. $193.5\text{--}194.2^\circ\text{C}$]; Anal. HPLC: t_R = 6.4 min, 10-100% MeCN over 11min, purity 96%); ^1H NMR (400 MHz, $\text{DMSO}-d_6$): δ /ppm 9.93 (1H, br, N-H), 8.37 (1H, d, $J=7.08\text{Hz}$, Ar-H), 8.00 (2H, d, $J=8.48\text{Hz}$, Ar-H), 7.86 (2H, d, $J=8.48\text{Hz}$, Ar-H), 7.82 (1H, dd, $J_a=1.50\text{Hz}$, $J_b=7.08\text{Hz}$, Ar-H), 7.06 (1H, t, $J=7.08\text{Hz}$, Ar-H), 6.34 (1H, s, Ar-H), 3.29 (3H, s, CH_3), 2.19 (3H, s, CH_3); HRMS (ES) m/z : found 358.0849 ($\text{C}_{17}\text{H}_{16}\text{N}_3\text{O}_4\text{S}$ $[\text{M}+\text{H}]^+$), requires 358.0862.

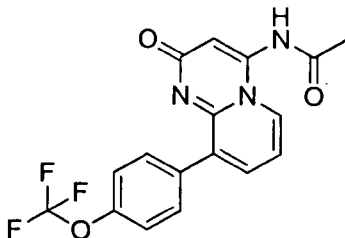
N-(9-(4-nitrophenyl)-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-4-yl)acetamide (**31d**)



Following procedure **P1**, **27** (75mg, 0.27mmol) was dissolved in dioxane:DMA 10:1 (5ml). 4-nitrobenzeneboronic acid (53mg, 0.32mmol), Pd(PPh₃)₄ (31mg, 0.027mmol) and Cs₂CO₃ (173mg, 0.53mmol) were added. Product **31d** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.26) and obtained as a cream solid (8.6mg, yield 10%) [mp. dec >250°C]; Anal. HPLC: *t*_R = 7.0 min, 10-100% MeCN over 11min, purity 96%); ¹H NMR (400 MHz, DMSO-*d*₆): δ/ppm 10.89 (1H, br, N-H), 8.74 (1H, d, *J*=7.08Hz, Ar-H), 8.29 (2H, d, *J*=8.48Hz, Ar-H), 7.89 (2H, d, *J*=8.48Hz, Ar-H), 7.77 (1H, dd, *J*_a=1.50Hz, *J*_b=7.08Hz, Ar-H), 6.98 (1H, t, *J*=7.08Hz, Ar-H), 6.51 (1H, s, Ar-H), 2.10 (3H, s, CH₃); HRMS (ES) *m/z*: found 325.0919 (C₁₆H₁₃N₄O₄ [M+H]⁺), requires 325.0937.

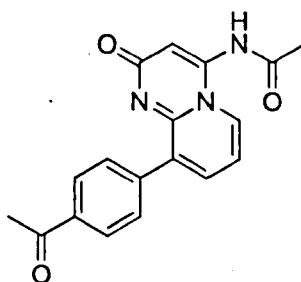
N-(2-oxo-9-(4-(trifluoromethoxy)phenyl)-2*H*-pyrido[1,2-*a*]pyrimidin-4-yl)acetamide

(**31e**)



Following procedure **P1**, **27** (90mg, 0.32mmol) was dissolved in dioxane:DMA 10:1 (5ml). 4-(trifluoromethoxy)phenylboronic acid (79mg, 0.38mmol), Pd(PPh₃)₄ (37mg, 0.032mmol) and Cs₂CO₃ (208mg, 0.64mmol) were added. Product **31e** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.29) and obtained as a white solid (9.2mg, yield 8%) [mp. dec >240°C]; Anal. HPLC: *t*_R = 15.3 min, 10-100% MeCN over 20min, purity 100%); ¹H NMR (400 MHz, DMSO-*d*₆): δ/ppm 10.88 (1H, br, N-H), 8.56 (1H, d, *J*=7.08Hz, Ar-H), 7.72 (3H, m, Ar-H), 7.44 (2H, dd, *J*_a=0.88Hz, *J*_b=8.80Hz, Ar-H), 6.98 (1H, t, *J*=7.08Hz, Ar-H), 6.43 (1H, s, Ar-H), 2.13 (3H, s, CH₃); HRMS (ES) *m/z*: found 364.0920 (C₁₇H₁₃F₃N₃O₃ [M+H]⁺), requires 364.0909.

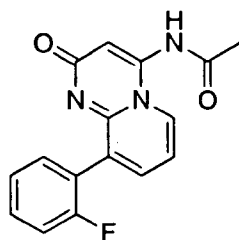
N-(9-(4-acetylphenyl)-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-4-yl)acetamide (**31f**)



Following procedure **P1**, **27** (90mg, 0.32mmol) was dissolved in dioxane:DMA 10:1 (5ml). 4-acetylphenylboronic acid (63mg, 0.38mmol), Pd(PPh₃)₄ (37mg, 0.032mmol)

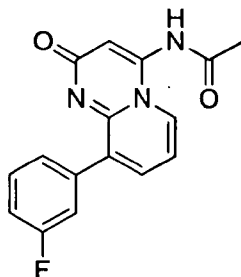
and Cs_2CO_3 (208mg, 0.64mmol) were added. Product **31f** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.22) and obtained as a cream solid (19.6mg, yield 19%) [mp. dec $>250^\circ\text{C}$]; Anal. HPLC: t_R = 6.7 min, 10-100% MeCN over 11min, purity 93%); ^1H NMR (400 MHz, $\text{DMSO}-d_6$): δ /ppm 10.69 (1H, br, N-H), 8.33 (1H, d, $J=7.10\text{Hz}$, Ar-H), 8.03 (2H, d, $J=8.48\text{Hz}$, Ar-H), 7.79 (1H, dd, $J_a=1.50\text{Hz}$, $J_b=7.10\text{Hz}$, Ar-H), 7.75 (2H, d, $J=8.48\text{Hz}$, Ar-H), 7.05 (1H, t, $J=7.10\text{Hz}$, Ar-H), 6.32 (1H, s, Ar-H), 2.64 (3H, s, CH_3), 2.19 (3H, s, CH_3); HRMS (ES) m/z : found 322.1165 ($\text{C}_{18}\text{H}_{16}\text{N}_3\text{O}_3$ $[\text{M}+\text{H}]^+$), requires 322.1192.

N-(9-(2-fluorophenyl)-2-oxo-2H-pyrido[1,2-*a*]pyrimidin-4-yl)acetamide (**31g**)



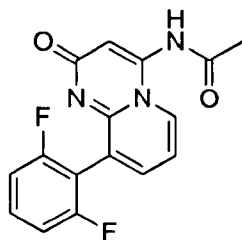
Following procedure **P1**, **27** (75mg, 0.27mmol) was dissolved in dioxane:DMA 10:1 (4ml). 2-fluorobenzenboronic acid (45mg, 0.32mmol), $\text{Pd}(\text{PPh}_3)_4$ (31mg, 0.027mmol) and Cs_2CO_3 (173mg, 0.53mmol) were added. Product **31g** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.26) and obtained as a cream solid (36mg, yield 46%) [mp. dec $>230^\circ\text{C}$]; Anal. HPLC: t_R = 6.8 min, 10-100% MeCN over 11min, purity 97%); ^1H NMR (400 MHz, $\text{DMSO}-d_6$): δ /ppm 10.80 (1H, br, N-H), 8.54 (1H, d, $J=7.08\text{Hz}$, Ar-H), 7.69 (1H, dd, $J_a=1.50\text{Hz}$, $J_b=7.08\text{Hz}$, Ar-H), 7.46 (2H, m, Ar-H), 7.28 (2H, m, Ar-H), 6.99 (1H, t, $J=7.08\text{Hz}$, Ar-H), 6.39 (1H, s, Ar-H), 2.14 (3H, s, CH_3); HRMS (ES) m/z : found 298.0974 ($\text{C}_{16}\text{H}_{13}\text{FN}_3\text{O}_2$ $[\text{M}+\text{H}]^+$), requires 298.0992.

N-(9-(3-fluorophenyl)-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-4-yl)acetamide (**31h**)



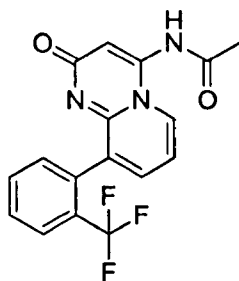
Following procedure **P1**, **27** (75mg, 0.27mmol) was dissolved in dioxane:DMA 10:1 (4ml). 3-fluorobenzenboronic acid (45mg, 0.32mmol), Pd(PPh₃)₄ (31mg, 0.027mmol) and Cs₂CO₃ (173mg, 0.53mmol) were added. Product **31h** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.25) and obtained as a cream solid (22.2mg, yield 28%) [mp. dec >240°C]; Anal. HPLC: *t*_R = 6.9 min, 10-100% MeCN over 11min, purity 97%); ¹H NMR (400 MHz, DMSO-*d*₆): δ/ppm 10.64 (1H, br, N-H), 8.37 (1H, d, *J*=7.08Hz, Ar-H), 7.77 (1H, dd, *J*_a=1.50Hz, *J*_b=7.08Hz, Ar-H), 7.49 (2H, m, Ar-H), 7.24 (2H, m, Ar-H), 7.02 (1H, t, *J*=7.08Hz, Ar-H), 6.34 (1H, s, Ar-H), 2.16 (3H, s, CH₃); HRMS (ES) *m/z*: found 298.0013 (C₁₆H₁₃FN₃O₂ [M+H]⁺), requires 298.0992.

N-(9-(2,6-difluorophenyl)-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-4-yl)acetamide (**32a**)



Following procedure **P1**, **27** (75mg, 0.27mmol) was dissolved in dioxane:DMA 10:1 (4ml). 2,6-difluorobenzeneboronic acid (50mg, 0.32mmol), Pd(PPh₃)₄ (31mg, 0.027mmol) and Cs₂CO₃ (173mg, 0.53mmol) were added. Product **32a** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.18) and obtained as a brown solid (22.3mg, yield 27%) [mp. dec >200°C]; Anal. HPLC: *t*_R = 7.0 min, 10-100% MeCN over 11min, purity 99%; ¹H NMR (400 MHz, DMSO-*d*₆): δ/ppm 10.90 (1H, br, N-H), 8.49 (1H, d, *J*=7.08Hz, Ar-H), 8.15 (1H, d, *J*=7.32Hz, Ar-H), 7.59 (3H, m, Ar-H), 7.05 (1H, t, *J*=7.08Hz, Ar-H), 6.35 (1H, s, Ar-H), 2.13 (3H, s, CH₃); HRMS (ES) *m/z*: found 316.0878 (C₁₆H₁₂F₂N₃O₂ [M+H]⁺), requires 316.0898.

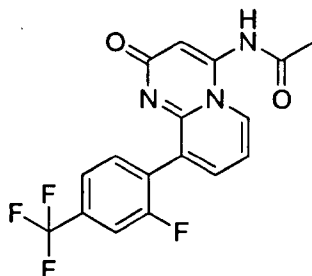
N-(2-oxo-9-(2-(trifluoromethyl)phenyl)-2*H*-pyrido[1,2-*a*]pyrimidin-4-yl)acetamide (**32b**)



Following procedure **P1**, **27** (75mg, 0.27mmol) was dissolved in dioxane:DMA 10:1 (4ml). 2-trifluoromethylbenzeneboronic acid (61mg, 0.32mmol), Pd(PPh₃)₄ (31mg, 0.027mmol) and Cs₂CO₃ (173mg, 0.53mmol) were added. Product **32b** was purified

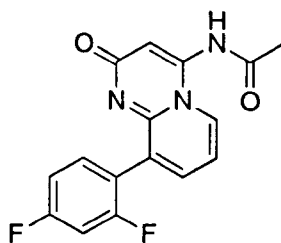
using preparative TLC (MeOH:DCM 1:10; R_f = 0.29) and obtained as a cream solid (13.5mg, yield 15%) [mp. dec >200°C]; Anal. HPLC: t_R = 7.3 min, 10-100% MeCN over 11min, purity 98%); ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 10.64 (1H, br, N-H), 8.37 (1H, d, J =7.08Hz, Ar-H), 7.85 (1H, d, J =7.74Hz, Ar-H), 7.75 (1H, m, Ar-H), 7.65 (2H, m, Ar-H), 7.45 (1H, d, J =7.76Hz, Ar-H), 7.03 (1H, t, J =7.08Hz, Ar-H), 6.29 (1H, s, Ar-H), 2.19 (3H, s, CH₃); HRMS (ES) m/z : found 348.0947 (C₁₇H₁₃F₃N₃O₂ [M+H]⁺), requires 348.0960.

N-(9-(2-fluoro-4-(trifluoromethyl)phenyl)-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-4-yl)acetamide (**32c**)



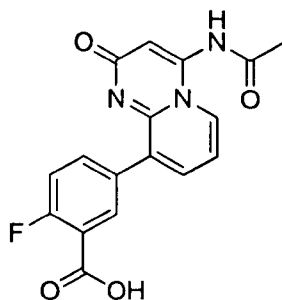
Following procedure **P1**, **27** (75mg, 0.27mmol) was dissolved in dioxane:DMA 10:1 (4ml). 2-fluoro-4-trifluoromethylbenzene boronic acid (66mg, 0.32mmol), Pd(PPh₃)₄ (31mg, 0.027mmol) and Cs₂CO₃ (173mg, 0.53mmol) were added. Product **32c** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.30) and obtained as a blue solid (17.3mg, yield 18%) [mp. dec >200°C]; Anal. HPLC: t_R = 7.8 min, 10-100% MeCN over 11min, purity 93%); ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 10.82 (1H, br, N-H), 8.45 (1H, d, J =7.08Hz, Ar-H), 7.81 (2H, m, Ar-H), 7.71 (2H, m, Ar-H), 7.05 (1H, t, J =7.08Hz, Ar-H), 6.35 (1H, s, Ar-H), 2.18 (3H, s, CH₃); HRMS (ES) m/z : found 366.0847 (C₁₇H₁₂F₄N₃O₂ [M+H]⁺), requires 366.0866.

N-(9-(2,4-difluorophenyl)-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-4-yl)acetamide (**32d**)



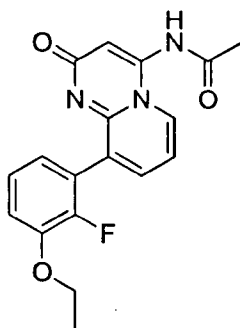
Following procedure **P1**, **27** (75mg, 0.27mmol) was dissolved in dioxane:DMA 10:1 (4ml). 2,4-difluorobenzeneboronic acid (50mg, 0.32mmol), Pd(PPh₃)₄ (31mg, 0.027mmol) and Cs₂CO₃ (173mg, 0.53mmol) were added. Product **32d** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.28) and obtained as a brown solid (18.8mg, yield 22%) [mp. dec >215°C]; Anal. HPLC: *t*_R = 7.0 min, 10-100% MeCN over 11min, purity 96%); ¹H NMR (400 MHz, DMSO-*d*₆): δ/ppm 9.70 (1H, br, N-H), 8.37 (1H, d, *J*=7.08Hz, Ar-H), 7.76 (1H, dd, *J*_a=1.50, *J*_b=7.08Hz, Ar-H), 7.51 (1H, m, Ar-H), 7.36 (1H, ddd, *J*_a=2.53Hz, *J*_b=9.60Hz, *J*_c=19.70Hz, Ar-H), 7.20 (1H, ddd, *J*_a=2.53Hz, *J*_b=8.59Hz, *J*_c=17.43Hz, Ar-H), 7.04 (1H, t, *J*=7.08Hz, Ar-H), 6.32 (1H, s, Ar-H), 2.19 (3H, s, CH₃); HRMS (ES) *m/z*: found 316.0888 (C₁₆H₁₂F₂N₃O₂ [M+H]⁺), requires 316.0898.

5-(4-acetamido-2-oxo-2H-pyrido[1,2-a]pyrimidin-9-yl)-2-fluorobenzoic acid (**32e**)



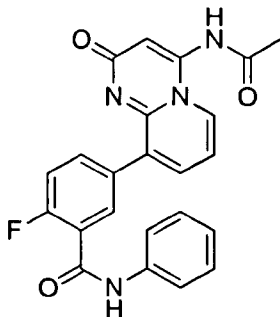
Compound **27** (400mg, 1.42mmol) was added to an Ace sealed pressure tube containing dioxane:DMA 10:1 (15ml). 4-fluoro-3-carboxybenzeneboronic acid (287mg, 1.56mmol), Pd(PPh₃)₄ (164mg, 0.14mmol) and Cs₂CO₃ (925mg, 2.84mmol) were added. The reaction mixture was stirred at 100°C for 24h. The resulting compound was filtered through Celite and solvent dried *in vacuo*. Product **32e** was purified using flash chromatography (MeOH:DCM 1:10; R_f = 0.02) and obtained as a cream solid (392mg, yield 81%) [mp. dec >250°C]; Anal. HPLC: *t*_R = 4.8 min, 10-100% MeCN over 11min, purity 90%); ¹H NMR (400 MHz, DMSO-d₆): δ/ppm 10.78 (1H, br, N-H), 10.20 (1H, br, OH), 9.23 (1H, d, *J*=7.08Hz, Ar-H), 7.75 (1H, m, Ar-H), 7.50 (1H, dd, *J*_a=1.52Hz, *J*_b=7.08Hz, Ar-H), 7.43 (2H, m, Ar-H), 7.04 (1H, t, *J*=7.08Hz, Ar-H), 6.75 (1H, s, Ar-H), 2.09 (3H, s, CH₃); HRMS (ES) *m/z*: found 342.0873 (C₁₇H₁₃FN₃O₄ [M+H]⁺), requires 342.0890.

N-(9-(3-ethoxy-2-fluorophenyl)-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-4-yl)acetamide (**32f**)



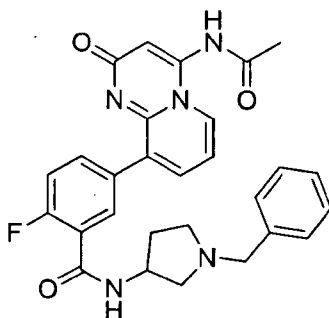
Following procedure **P1**, **27** (60mg, 0.21mmol) was dissolved in dioxane:DMA 10:1 (4ml). 3-ethoxy-2-fluorophenylboronic acid (47mg, 0.26mmol), Pd(PPh₃)₄ (25mg, 0.02mmol) and Cs₂CO₃ (139mg, 0.43mmol) were added. Product **32f** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.26) and obtained as a cream solid (34.6mg, yield 48%) [mp. dec >250°C]; Anal. HPLC: *t*_R = 7.3 min, 10-100% MeCN over 11min, purity 94%); ¹H NMR (400 MHz, DMSO-*d*₆): δ/ppm 10.72 (1H, br, N-H), 8.35 (1H, d, *J*=7.08Hz, Ar-H), 7.72 (1H, dd, *J*_a=1.08Hz, *J*_b=7.08Hz, Ar-H), 7.21 (2H, m, Ar-H), 7.02 (1H, t, *J*=7.08Hz, Ar-H), 6.93 (1H, m, Ar-H), 6.30 (1H, s, Ar-H), 4.15 (2H, q, *J*=7.08Hz, CH₂), 2.19 (3H, s, CH₃), 1.40 (3H, t, *J*=7.08Hz, CH₃); HRMS (ES) *m/z*: found 342.1234 (C₁₈H₁₇FN₃O₃ [M+H]⁺), requires 342.1254.

5-(4-acetamido-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-9-yl)-2-fluoro-*N*-phenylbenzamide
(33a)



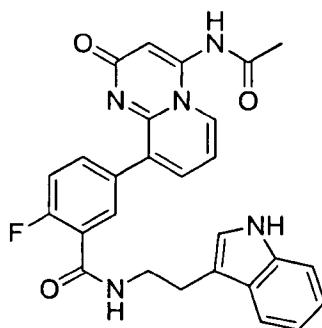
Following procedure **P2**, **32e** (48mg, 0.14mmol) was dissolved in DMF (6ml). Aniline (16mg, 0.17mmol), CDMT (37mg, 0.21mmol) and 4-methylmorpholine (21mg, 0.21mmol) were added. Product **33a** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.23) and obtained as a cream solid (15.5mg, yield 26%) [mp. dec >250°C]; Anal. HPLC: t_R = 7.7 min, 10-100% MeCN over 11min, purity 99%); ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 10.62 (1H, br, N-H), 10.51 (1H, s, N-H), 8.35 (1H, d, J =7.08Hz, Ar-H), 7.83 (3H, m, Ar-H), 7.74 (2H, d, J =7.83Hz, Ar-H), 7.45 (1H, t, J =9.21Hz, Ar-H), 7.36 (2H, t, J =7.83Hz, Ar-H), 7.12 (1H, t, J =7.33Hz, Ar-H), 7.04 (1H, t, J =7.08Hz, Ar-H), 6.34 (1H, s, Ar-H), 2.19 (3H, s, CH₃); HRMS (ES) m/z : found 417.1360 (C₂₃H₁₈FN₄O₃ [M+H]⁺), requires 417.1363.

5-(4-acetamido-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-9-yl)-*N*-(1-benzylpyrrolidin-3-yl)-2-fluorobenzamide (**33b**)



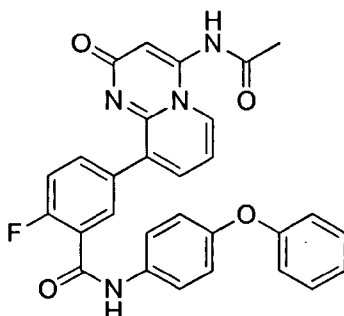
Following procedure **P2**, **32e** (60mg, 0.18mmol) was dissolved in DMF (8ml). 1-benzyl-3-aminopyrrolidine (37mg, 0.21mmol), CDMT (46mg, 0.26mmol) and 4-methylmorpholine (27mg, 0.26mmol) were added. Product **33b** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.21) and obtained as a cream solid (27.8mg, yield 32%) [mp. 190.1-192.6°C]; Anal. HPLC: t_R = 6.8 min, 10-100% MeCN over 11min, purity 96%; ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 10.81 (1H, br, N-H), 8.75 (1H, d, J =5.80Hz, N-H), 8.35 (1H, d, J =7.12Hz, Ar-H), 7.80 (1H, dd, J_a =0.92Hz, J_b =6.97Hz, Ar-H), 7.76 (2H, m, Ar-H), 7.47 (2H, m, Ar-H), 7.35 (4H, m, Ar-H), 7.03 (1H, t, J =7.12Hz, Ar-H), 6.33 (1H, s, Ar-H), 4.16 (1H, m, CH), 3.98 (2H, s, CH₂), 2.77 (2H, m, CH₂), 2.66 (2H, m, CH₂), 2.20 (3H, s, CH₃), 1.82 (2H, m, CH₂); HRMS (ES) m/z : found 500.2082 (C₂₈H₂₇FN₅O₃ [M+H]⁺), requires 500.2098.

N-((1*H*-indol-3-yl)methyl)-5-(4-acetamido-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-9-yl)-2-fluorobenzamide (**33c**)



Following procedure **P2**, **32e** (70mg, 0.21mmol) was dissolved in DMF (8ml). Tryptamine (39mg, 0.25mmol), CDMT (54mg, 0.31mmol) and 4-methylmorpholine (31mg, 0.31mmol) were added. Product **33c** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.19) and obtained as a white solid (9.6mg, yield 10%) [mp. dec >250°C]; Anal. HPLC: t_R = 4.8 min, 10-100% MeCN over 11min, purity 94%); ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 10.73 (1H, br, N-H), 9.79 (1H, br, N-H), 8.47 (1H, t, J =5.58Hz, N-H), 8.32 (1H, d, J =7.08Hz, Ar-H), 7.76 (1H, dd, J_a = 1.48Hz, J_b =7.08Hz, Ar-H), 7.59 (2H, m, Ar-H), 7.36 (3H, m, Ar-H), 7.20 (1H, d, J =7.20Hz, Ar-H), 7.03 (3H, m, Ar-H), 6.33 (1H, s, Ar-H), 3.55 (2H, m, CH₂), 2.90 (2H, t, J =7.07Hz, CH₂), 2.20 (3H, s, CH₃); HRMS (ES) m/z : found 506.1600 (C₂₇H₂₂FN₅NaO₃ [M+Na]⁺), requires 506.1604.

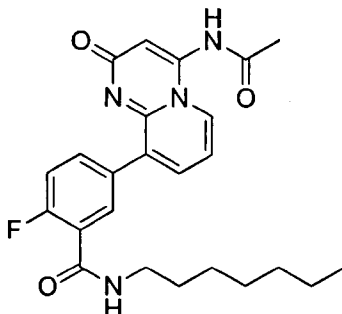
5-(4-acetamido-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-9-yl)-2-fluoro-*N*-(4-phenoxyphenyl) benzamide (**33d**)



Following procedure **P2**, **32e** (70mg, 0.21mmol) was dissolved in DMF (8ml). 4-phenoxyaniline (46mg, 0.25mmol), CDMT (54mg, 0.31mmol) and 4-methylmorpholine (31mg, 0.31mmol) were added. Product **33d** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.25) and obtained as a cream solid (13.2mg, yield 13%) [mp. dec >250°C]; Anal. HPLC: t_R = 9.0 min, 10-100% MeCN over 11min, purity 94%); ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 10.82 (1H, br, N-H), 10.55 (1H, s, N-H), 8.35 (1H, d, J =7.08Hz, Ar-H), 7.83 (3H, m, Ar-H), 7.76 (2H, d, J =9.08Hz, Ar-H), 7.41 (3H, m, Ar-H), 7.12 (1H, t, J =7.40Hz, Ar-H), 7.05 (3H, m, Ar-H), 6.99 (2H, d, J =7.83Hz, Ar-H), 6.34 (1H, s, Ar-H), 2.19 (3H, s, CH_3); HRMS (ES) m/z : found 509.1640 ($\text{C}_{29}\text{H}_{22}\text{FN}_4\text{O}_4$ $[\text{M}+\text{H}]^+$), requires 509.1625.

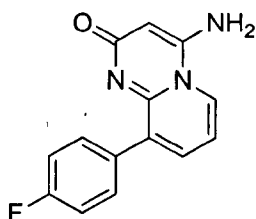
5-(4-acetamido-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-9-yl)-2-fluoro-*N*-heptylbenzamide

(**33e**)



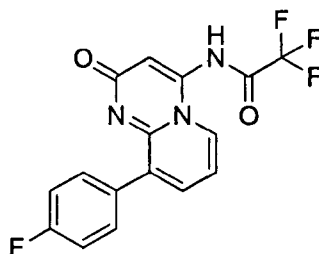
Following procedure **P2**, **32e** (60mg, 0.18mmol) was dissolved in DMF (8ml). Heptylamine (24mg, 0.21mmol), CDMT (46mg, 0.26mmol) and 4-methylmorpholine (27mg, 0.26mmol) were added. Product **33e** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.21) and obtained as a white solid (15.5mg, yield 20%) [mp. dec >250°C]; Anal. HPLC: t_R = 8.9 min, 10-100% MeCN over 11min, purity 96%); ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 10.71 (1H, br, N-H), 8.37 (1H, t, J =5.60Hz, N-H), 8.31 (1H, d, J =7.08Hz, Ar-H), 7.75 (3H, m, Ar-H), 7.36 (1H, m, Ar-H), 7.02 (1H, t, J =7.08Hz, Ar-H), 6.32 (1H, s, Ar-H), 3.25 (2H, q, J =6.78Hz, CH_2), 2.19 (3H, s, CH_3), 1.52 (2H, m, CH_2), 1.29 (8H, m, CH_2), 0.87 (3H, t, J =6.88Hz, CH_3); ^{13}C NMR (500 MHz, DMSO- d_6): δ /ppm 170.7, 167.8, 163.8, 160.2, 158.2, 150.7, 142.2, 137.1, 133.7, 131.6, 129.5, 124.5, 124.3, 115.9, 112.4, 110.9, 39.2, 31.7, 29.4, 28.9, 26.8, 23.7, 22.5, 14.4; HRMS (ES) m/z : found 439.2150 ($\text{C}_{24}\text{H}_{28}\text{FN}_4\text{O}_3$ [$\text{M}+\text{H}$] $^+$), requires 439.2145.

4-amino-9-(4-fluorophenyl)-2H-pyrido[1,2-a]pyrimidin-2-one (**34a**)



Compound **18m** (300mg, 1.01mmol) was added to a stirred solution of hydrochloric acid:methanol 50:50 (20ml). The reaction mixture was stirred at RT for 24h and subsequently neutralised using saturated sodium hydroxide solution. The product was precipitated out, filtered and washed using water (25ml) to afford **34a** as a brown solid (193.6mg, yield 75%) [mp. dec >250°C]; Anal. HPLC: t_R = 4.8 min, 10-100% MeCN over 11min, purity 100%); ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 9.50 (2H, br, NH_2), 8.45 (1H, dd, $J_a=1.77\text{Hz}$, $J_b=4.80\text{Hz}$, Ar-H), 7.84 (1H, dd, $J_a=1.77\text{Hz}$, $J_b=7.83\text{Hz}$, Ar-H), 7.44 (4H, m, Ar-H), 7.26 (1H, t, $J=8.84\text{Hz}$, Ar-H), 6.54 (1H, s, Ar-H); HRMS (ES) m/z : found 256.0883 ($\text{C}_{14}\text{H}_{11}\text{FN}_3\text{O}$ $[\text{M}+\text{H}]^+$), requires 256.0886.

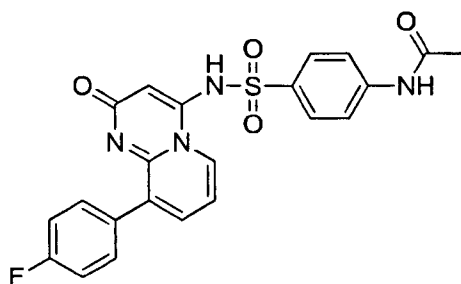
2,2,2-trifluoro-*N*-(9-(4-fluorophenyl)-2-oxo-2H-pyrido[1,2-a]pyrimidin-4-yl)acetamide (**34b**)



To a stirred solution of **34a** (50mg, 0.19mmol) in DMF (10ml) was added sodium hydride (9mg, 0.23mmol). The reaction mixture was stirred at RT for 1h after which

trifluoroacetyl chloride (52mg, 0.39mmol) was added. The reaction mixture was stirred further for 2h at RT. Water (20ml) was added to the reaction mixture. The product was precipitated out, filtered and purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.10). Product **34b** was obtained as a white solid (28.4mg, yield 41%) [mp. >300°C]; Anal. HPLC: t_R = 9.7 min, 10-100% MeCN over 11min, purity 97%); ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 10.04 (1H, br, N-H), 9.08 (1H, d, J =7.08Hz, Ar-H), 7.80 (1H, dd, J_a =1.77Hz, J_b =7.08Hz, Ar-H), 7.70 (4H, m, Ar-H), 7.22 (1H, t, J =7.08Hz, Ar-H), 6.23 (1H, s, Ar-H); HRMS (ES) m/z : found 352.0702 ($\text{C}_{16}\text{H}_{10}\text{F}_4\text{N}_3\text{O}_2$ $[\text{M}+\text{H}]^+$), requires 352.0709.

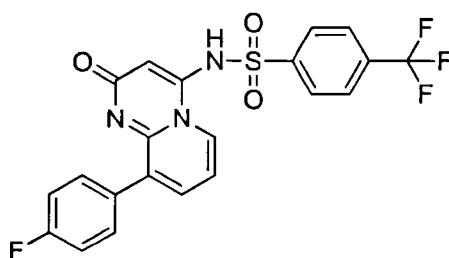
N-(4-(*N*-(9-(4-fluorophenyl)-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-4-yl)sulfamoyl)phenyl)acetamide (**34c**)



To a stirred solution of **34a** (50mg, 0.19mmol) in DMF (10ml) was added sodium hydride (9mg, 0.23mmol). The reaction mixture was stirred at RT for 1h after which 4-acetamidobenzene sulfonyl chloride (46mg, 0.19mmol) was added. The reaction mixture was stirred further for 2h at RT. Water (20ml) was added to the reaction mixture. The product was precipitated out, filtered and purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.15). Product **34c** was obtained as a yellow solid (12.6mg, yield 14%) [mp. 141.2-142.7°C]; Anal. HPLC: t_R = 8.03 min, 10-100% MeCN over

11min, purity 92%); ^1H NMR (400 MHz, DMSO-d_6): δ /ppm 10.24 (1H, s, N-H), 9.88 (1H, br, N-H), 9.15 (1H, d, $J=7.08\text{Hz}$, Ar-H), 7.81 (1H, m, Ar-H), 7.75 (2H, d, $J=8.72\text{Hz}$, Ar-H), 7.67 (2H, d, $J=8.72\text{Hz}$, Ar-H), 7.60 (2H, m, Ar-H), 7.26 (3H, m, Ar-H), 6.07 (1H, s, Ar-H), 2.06 (3H, s, CH_3); HRMS (ES) m/z : found 453.1027 ($\text{C}_{22}\text{H}_{18}\text{FN}_4\text{O}_4\text{S} [\text{M}+\text{H}]^+$), requires 453.1033.

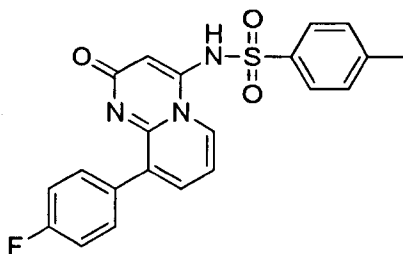
N-(9-(4-fluorophenyl)-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-4-yl)-4-(trifluoromethyl)benzenesulfonamide (**34d**)



To a stirred solution of **34a** (50mg, 0.19mmol) in DMF (10ml) was added sodium hydride (9mg, 0.23mmol). The reaction mixture was stirred at RT for 1h after which 4-trifluoromethylbenzene sulfonyl chloride (48mg, 0.19mmol) was added. The reaction mixture was stirred further for 2h at RT. Water (20ml) was added to the reaction mixture. The product was precipitated out, filtered and purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.38). Product **34d** was obtained as a white solid (8mg, yield 9%) [mp. dec $>250^\circ\text{C}$]; Anal. HPLC: t_R = 10.1 min, 10-100% MeCN over 12min, purity 100%); ^1H NMR (400 MHz, DMSO-d_6): δ /ppm 10.62 (1H, s, N-H), 9.10 (1H, d, $J=4.54\text{Hz}$, Ar-H), 7.89 (1H, dd, $J_a=1.52\text{Hz}$, $J_b=4.55\text{Hz}$, Ar-H), 7.67 (4H, m, Ar-H), 7.15 (1H, t, $J=4.55\text{Hz}$, Ar-H), 7.09 (2H, d, $J=4.29$, Ar-H), (2H, d, $J=4.29\text{Hz}$, Ar-H), 5.38

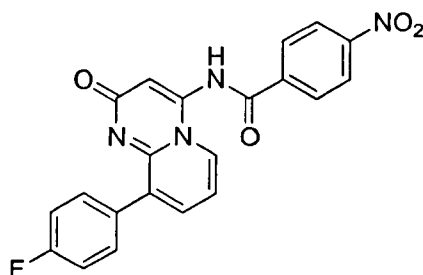
(1H, s, Ar-H); HRMS (ES) m/z : found 464.0701 ($C_{21}H_{14}F_4N_3O_3S$ $[M+H]^+$), requires 464.0692.

N-(9-(4-fluorophenyl)-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-4-yl)-4-methylbenzene sulphonamide (**34e**)



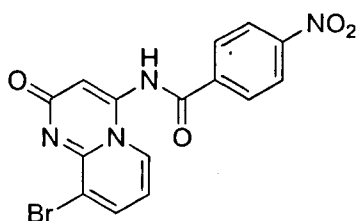
To a stirred solution of **34a** (50mg, 0.19mmol) in DMF (10ml) was added sodium hydride (9mg, 0.23mmol). The reaction mixture was stirred at RT for 1h after which *p*-toluene sulfonyl chloride (37mg, 0.19mmol) was added. The reaction mixture was stirred further for 2h at RT. Water (20ml) was added to the reaction mixture. The product was precipitated out, filtered and purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.43). Product **34e** was obtained as a cream solid (7.5mg, yield 9%) [mp. dec >250°C]; Anal. HPLC: t_R = 9.4 min, 10-100% MeCN over 11min, purity 95%); 1H NMR (400 MHz, DMSO- d_6): δ /ppm 10.24 (1H, br, N-H), 9.10 (1H, d, J =8.34Hz, Ar-H), 7.81 (2H, d, J =8.34Hz, Ar-H), 7.71 (2H, d, J =8.34Hz, Ar-H), 7.60 (1H, m, Ar-H), 7.49 (2H, d, J =8.08Hz, Ar-H), 7.27 (1H, t, J = 8.34, Ar-H), (2H, d, J =8.08Hz, Ar-H), 6.54 (1H, s, Ar-H), 2.29 (3H, s, CH_3); HRMS (ES) m/z : found 410.0964 ($C_{21}H_{17}FN_3O_3S$ $[M+H]^+$), requires 410.0975.

N-(9-(4-fluorophenyl)-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-4-yl)-4-nitrobenzamide (**34f**)



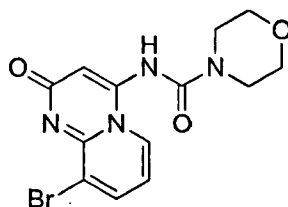
To a stirred solution of **34a** (50mg, 0.19mmol) in ethyl acetate (10ml) was added triethylamine (0.03ml, 0.21mmol). The reaction mixture was stirred at RT for 0.5h after which 4-nitrobenzoyl chloride (38mg, 0.21mmol) was added. The reaction mixture was stirred further for 2h at RT. Water (20ml) was added to the reaction mixture. The product was precipitated out, filtered and purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.34). Product **34f** was obtained as a yellow solid (6.4mg, yield 8%) [mp. dec >250°C]; Anal. HPLC: t_R = 8.9 min, 10-100% MeCN over 11min, purity 100%); ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 12.08 (1H, br, N-H), 9.55 (1H, d, J =7.49Hz, Ar-H), 8.41 (2H, d, J =8.88Hz, Ar-H), 8.26 (2H, d, J =8.88Hz, Ar-H), 7.68 (3H, m, Ar-H), 7.27 (3H, m, Ar-H), 7.03 (1H, s, Ar-H); HRMS (ES) m/z : found 405.0995 ($\text{C}_{21}\text{H}_{14}\text{FN}_4\text{O}_4$ $[\text{M}+\text{H}]^+$), requires 405.0999.

N-(9-bromo-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-4-yl)-4-nitrobenzamide (**35a**)



To a stirred solution of **26** (50mg, 0.21mmol) in ethyl acetate (10ml) was added triethylamine (2-3 drops) and 4-nitrobenzoyl chloride (41mg, 0.22mmol) at 0°C. The reaction mixture was stirred in an ice bath at 0°C for 2h. Water (20ml) was added and the resulting product was extracted using ethyl acetate (3 x 20ml). The organic layers were combined and dried over magnesium sulphate and solvent evaporated *in vacuo* to afford **35a** as a cream solid (11.9mg, yield 15%) (MeOH:DCM 1:10; R_f = 0.13) [mp. dec >250°C]; Anal. HPLC: t_R = 8.0 min, 10-100% MeCN over 11min, purity 100%; ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 12.07 (1H, br, N-H), 8.41 (2H, d, J =8.70Hz, Ar-H), 8.32 (1H, d, J =7.08Hz, Ar-H), 8.27 (2H, d, J =8.70Hz, Ar-H), 8.17 (1H, d, J =7.08Hz, Ar-H), 7.01 (1H, t, J =7.08Hz, Ar-H), 6.54 (1H, s, Ar-H); HRMS (ES) m/z : found 410.9693 ($\text{C}_{15}\text{H}_9\text{BrN}_4\text{NaO}_4$ [$\text{M}+\text{Na}$] $^+$), requires 410.9705.

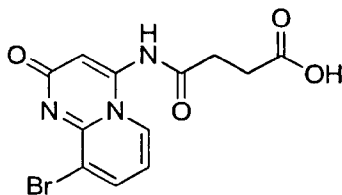
N-(9-bromo-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-4-yl)morpholine-4-carboxamide (**35b**)



To a stirred solution of **26** (100mg, 0.42mmol) in dimethylformamide (10ml) was added sodium hydride (20mg, 0.5mmol). The reaction mixture was stirred at RT for

1h after which 4-morpholinecarbonyl chloride (62mg, 0.42mmol) was added. The reaction mixture was stirred further for 2h at RT. Water (30ml) was added to the reaction mixture. The product subsequently precipitated out, filtered and purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.30). Product **35b** was obtained as a cream solid (23.6mg, yield 16%) [mp. dec >250°C]; Anal. HPLC: t_R = 6.2 min, 10-100% MeCN over 11min, purity 99%); ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 10.62 (1H, br, N-H), 9.21 (1H, d, J =7.20Hz, Ar-H), 8.29 (1H, dd, J_a =1.04Hz, J_b =7.26Hz, Ar-H), 7.05 (1H, t, J =7.26Hz, Ar-H), 6.93 (1H, s, Ar-H), 3.52 (4H, m, CH₂), 3.21 (4H, m, CH₂); HRMS (ES) m/z : found 353.0260 (C₁₃H₁₄BrN₄O₃ [M+H]⁺), requires 353.0249.

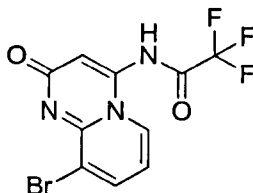
4-((9-bromo-2-oxo-2H-pyrido[1,2-a]pyrimidin-4-yl)amino)-4-oxobutanoic acid (**35c**)



Following procedure **P2**, succinic acid (100mg, 0.85mmol) was dissolved in DMF (15ml). Compound **26** (244mg, 1.02mmol), CDMT (221mg, 1.27mmol) and 4-methylmorpholine (128mg, 1.27mmol) were added. Product **35c** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.06) and obtained as a cream solid (23.6mg, yield 8%) [mp. dec >250°C]; Anal. HPLC: t_R = 4.7 min, 10-100% MeCN over 11min, purity 100%); ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 12.55 (1H, s, O-H), 10.62 (1H, br, N-H), 8.30 (1H, d, J =7.08Hz, Ar-H), 8.02 (1H, d, J =7.08Hz, Ar-H), 6.95 (1H, m, Ar-H), 6.17 (1H, s, Ar-H), 2.64 (2H, t, J =7.12Hz, CH₂), 2.39 (2H, t, J =7.12Hz, CH₂); ^{13}C NMR (500 MHz, DMSO- d_6): δ /ppm 176.0, 175.5, 151.0, 147.4, 141.9,

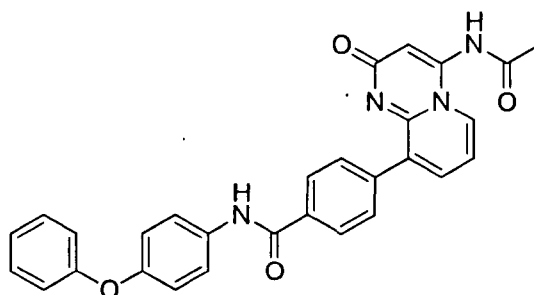
140.5, 124.6, 119.3, 113.9, 110.2, 34.7, 33.2; HRMS (ES) m/z : found 339.9960 ($C_{12}H_{11}BrN_3O_4$ $[M+H]^+$), requires 339.9933.

N-(9-bromo-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-4-yl)-2,2,2-trifluoroacetamide (**35d**)



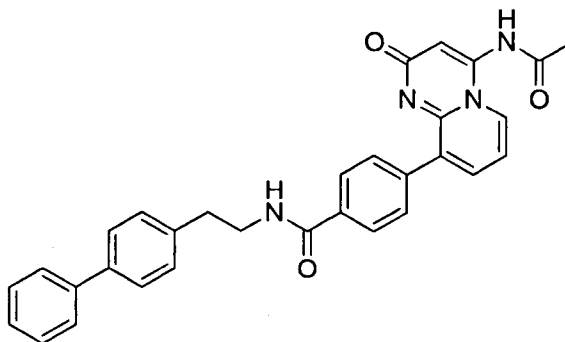
Reaction similar to that of **35b** except that compound **26** (200mg, 0.83mmol), dimethylformamide (15ml), sodium hydride (40mg, 1mmol) and trifluoroacetyl chloride (110mg, 0.83mmol) were used. Product **35d** was purified using column chromatography (MeOH:DCM 1:10; R_f = 0.02) and obtained as a brown solid (106.6mg, yield 38%) [mp. dec $>270^\circ\text{C}$]; Anal. HPLC: t_R = 8.9 min, 10-100% MeCN over 14min, purity 95%); ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 13.40 (1H, br, N-H), 9.57 (1H, dd, $J_a=1.28\text{Hz}$, $J_b=7.20\text{Hz}$, Ar-H), 8.61 (1H, dd, $J_a=1.30\text{Hz}$, $J_b=7.24\text{Hz}$, Ar-H), 7.72 (1H, s, Ar-H), 7.47 (1H, t, $J=7.20\text{Hz}$, Ar-H); ^{13}C NMR (500 MHz, DMSO- d_6): δ /ppm 168.9, 162.5, 156.8, 148.4, 142.9, 129.6, 119.2, 118.6, 117.6, 94.4; HRMS (ES) m/z : found 335.9599 ($C_{10}H_6BrF_3N_3O_2$ $[M+H]^+$), requires 335.9595.

4-(4-acetamido-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-9-yl)-*N*-(4-phenoxyphenyl)
benzamide (**36a**)



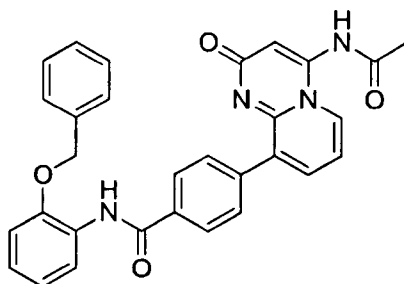
Following procedure **P2**, **18h** (80mg, 0.25mmol) was dissolved in DMF (10ml). 4-phenoxyaniline (55mg, 0.30mmol), CDMT (65mg, 0.37mmol) and 4-methylmorpholine (38mg, 0.37mmol) were added. Product **36a** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.25) and obtained as a cream solid (16mg, yield 13%) [mp. 148.9-150.5°C]; Anal. HPLC: t_R = 12.1 min, 10-100% MeCN over 14min, purity 99%; ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 12.25 (1H, br, N-H), 10.39 (1H, s, N-H), 8.43 (1H, d, J =7.08Hz, Ar-H), 8.02 (2H, d, J =8.44Hz, Ar-H), 7.84 (2H, d, J =9.01Hz, Ar-H), 7.79 (1H, dd, J_a =1.40Hz, J_b =7.08Hz, Ar-H), 7.75 (2H, d, J =8.44Hz, Ar-H), 7.39 (2H, m, Ar-H), 7.12 (1H, m, Ar-H), 7.03 (5H, m, Ar-H), 6.38 (1H, s, Ar-H), 2.18 (3H, s, CH₃); HRMS (ES) m/z : found 491.1690 (C₂₉H₂₃N₄O₄ [M+H]⁺), requires 491.1675.

N-(2-([1,1'-biphenyl]-4-yl)ethyl)-4-(4-acetamido-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-9-yl)benzamide (**36b**)



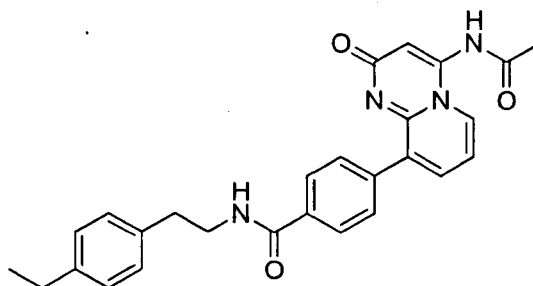
Following procedure **P2**, **18h** (100mg, 0.31mmol) was dissolved in DMF (8ml). 2-(4-biphenyl)ethylamine (73mg, 0.37mmol), CDMT (81mg, 0.46mmol) and 4-methylmorpholine (47mg, 0.46mmol) were added. Product **36b** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.28) and obtained as a cream solid (22.6mg, yield 15%) [mp. dec >250°C]; Anal. HPLC: t_R = 8.8 min, 10-100% MeCN over 11min, purity 97%); ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 8.95 (1H, d, J =7.08Hz, Ar-H), 8.67 (1H, t, J =6.52Hz, N-H), 8.45 (1H, s, N-H), 7.88 (2H, d, J =8.44Hz, Ar-H), 7.65 (7H, m, Ar-H), 7.45 (2H, m, Ar-H), 7.36 (3H, m, Ar-H), 6.90 (1H, t, J =7.08Hz, Ar-H), 6.61 (1H, s, Ar-H), 3.55 (2H, q, J =7.12Hz, CH_2), 2.92 (2H, t, J =7.12Hz, CH_2), 2.05 (3H, s, CH_3); HRMS (ES) m/z : found 503.2086 ($\text{C}_{31}\text{H}_{27}\text{N}_4\text{O}_3$ $[\text{M}+\text{H}]^+$), requires 503.2038.

4-(4-acetamido-2-oxo-2H-pyrido[1,2-a]pyrimidin-9-yl)-N-(2-(benzyloxy)phenyl)benzamide (**36c**)



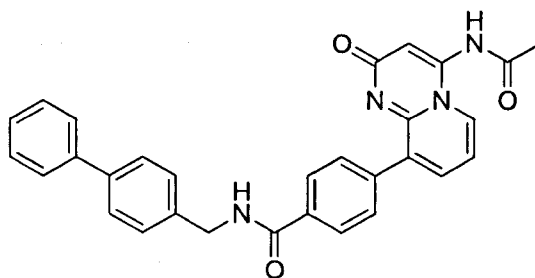
To a stirred solution of HBTU (211mg, 0.56mmol) in dichloromethane (8ml) was added compound **18h** (90mg, 0.28mmol). The reaction mixture was stirred at RT for 1h after which 2-benzyloxyaniline (67mg, 0.33mmol) was added and the mixture was subsequently stirred for a further 1.5h at RT. The solvent was evaporated *in vacuo* and product **36c** was purified using column chromatography (MeOH:DCM 1:10; R_f = 0.29) and obtained as a cream solid (28.7mg, yield 29%) [mp. dec $>140^\circ\text{C}$]; Anal. HPLC: t_R = 6.9 min, 10-100% MeCN over 11min, purity 91%); ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 10.69 (1H, br, N-H), 9.60 (1H, s, N-H), 8.33 (1H, d, $J=7.08\text{Hz}$, Ar-H), 8.03 (2H, d, $J=8.47\text{Hz}$, Ar-H), 7.97 (2H, m, Ar-H), 7.85 (1H, d, $J=8.47\text{Hz}$, Ar-H), 7.79 (1H, m, Ar-H), 7.75 (2H, d, $J=8.52\text{Hz}$, Ar-H), 7.71 (1H, m, Ar-H), 7.54 (2H, m, Ar-H), 7.41 (2H, m, Ar-H), 7.17 (1H, m, Ar-H), 7.04 (1H, t, $J=7.08\text{Hz}$, Ar-H), 6.32 (1H, s, Ar-H), 3.90 (2H, s, CH_2), 2.19 (3H, s, CH_3); HRMS (ES) m/z : found 505.1877 ($\text{C}_{30}\text{H}_{25}\text{N}_4\text{O}_4$ $[\text{M}+\text{H}]^+$), requires 505.1831.

4-(4-acetamido-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-9-yl)-*N*-(4-ethylphenethyl)benzamide
(**36d**)



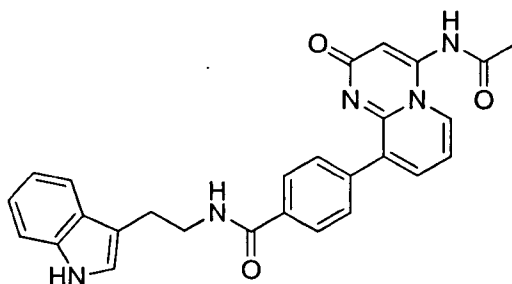
Following procedure **P2**, **18h** (100mg, 0.31mmol) was dissolved in DMF (8ml). 4-ethylphenethylamine (55mg, 0.37mmol), CDMT (81mg, 0.46mmol) and 4-methylmorpholine (47mg, 0.46mmol) were added. Product **36d** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.28) and obtained as a cream solid (32.9mg, yield 23%) [mp. dec >250°C]; Anal. HPLC: t_R = 8.3 min, 10-100% MeCN over 11min, purity 95%); ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 9.73 (1H, br, N-H), 8.65 (1H, t, J =5.58Hz, N-H), 8.40 (1H, d, J =7.10Hz, Ar-H), 7.89 (2H, d, J =8.41Hz, Ar-H), 7.76 (1H, dd, J_a =1.40Hz, J_b =7.10Hz, Ar-H), 7.68 (2H, d, J =8.41Hz, Ar-H), 7.17 (4H, m, Ar-H), 7.02 (1H, t, J =7.10Hz, Ar-H), 6.36 (1H, s, Ar-H), 3.49 (2H, q, J =7.05Hz, CH₂), 2.83 (2H, t, J =7.05Hz, CH₂), 2.57 (2H, q, J =7.60Hz, CH₂), 2.18 (3H, s, CH₃), 1.17 (3H, t, J =7.60Hz, CH₃); HRMS (ES) m/z : found 455.2088 (C₂₇H₂₇N₄O₃ [M+H]⁺), requires 455.2038.

N-([1,1'-biphenyl]-4-ylmethyl)-4-(4-acetamido-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-9-yl) benzamide (**36e**)



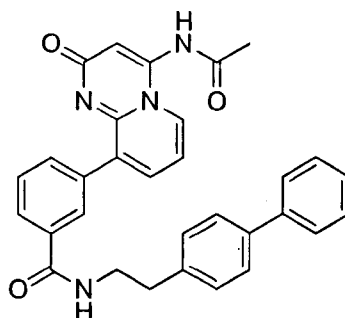
Following procedure **P2**, **18h** (100mg, 0.31mmol) was dissolved in DMF (8ml). 4-phenylbenzylamine (68mg, 0.37mmol), CDMT (81mg, 0.46mmol) and 4-methylmorpholine (47mg, 0.46mmol) were added. Product **36e** was purified using column chromatography (MeOH:DCM 1:10; R_f = 0.20) and obtained as a cream solid (26.9mg, yield 18%) [mp. dec >260°C]; Anal. HPLC: t_R = 8.5 min, 10-100% MeCN over 11min, purity 96%); ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 10.43 (1H, br, N-H), 9.18 (1H, t, J =5.98Hz, N-H), 8.34 (1H, d, J =7.08Hz, Ar-H), 7.99 (2H, d, J =8.44Hz, Ar-H), 7.78 (1H, dd, J_a =1.40Hz, J_b =7.08Hz, Ar-H), 7.71 (2H, d, J =8.44Hz, Ar-H), 7.65 (4H, m, Ar-H), 7.45 (4H, m, Ar-H), 7.36 (1H, m, Ar-H), 7.04 (1H, t, J =7.08Hz, Ar-H), 6.34 (1H, s, Ar-H), 4.57 (2H, d, J =5.98Hz, CH_2), 2.18 (3H, s, CH_3); HRMS (ES) m/z : found 489.1941 ($\text{C}_{30}\text{H}_{25}\text{N}_4\text{O}_3$ [$\text{M}+\text{H}$] $^+$), requires 489.1927.

N-(2-(1*H*-indol-3-yl)ethyl)-4-(4-acetamido-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-9-yl) benzamide (**36f**)



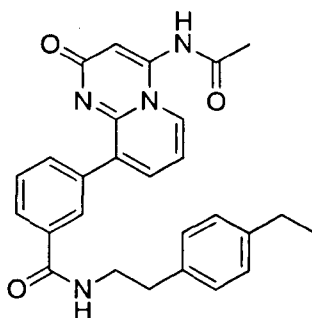
Following procedure **P2**, **18h** (100mg, 0.31mmol) was dissolved in DMF (8ml). Tryptamine (59mg, 0.37mmol), CDMT (81mg, 0.46mmol) and 4-methylmorpholine (47mg, 0.46mmol) were added. Product **36f** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.14) and obtained as a cream solid (35mg, yield 24%) [mp. 108.8-109.4°C]; Anal. HPLC: t_R = 10.0 min, 10-100% MeCN over 12min, purity 96%); ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 10.84 (1H, s, N-H), 9.66 (1H, br, N-H), 8.70 (1H, t, J =5.62Hz, N-H), 8.46 (1H, d, J =7.08Hz, Ar-H), 7.92 (2H, d, J =8.44Hz, Ar-H), 7.76 (1H, dd, J_a =1.40Hz, J_b =7.08Hz, Ar-H), 7.68 (2H, d, J =8.44Hz, Ar-H), 7.61 (1H, d, J =7.84Hz, Ar-H), 7.36 (1H, t, J =8.28Hz, Ar-H), 7.21 (1H, dd, J_a = 2.24Hz, J_b =15.40Hz, Ar-H), 7.08 (1H, m, Ar-H), 7.01 (2H, m, Ar-H), 6.39 (1H, s, Ar-H), 3.58 (2H, q, J =6.60Hz, CH₂), 2.98 (2H, t, J =6.60Hz, CH₂), 2.17 (3H, s, CH₃); HRMS (ES) m/z : found 466.1891 (C₂₇H₂₄N₅O₃ [M+H]⁺), requires 466.1879.

N-(2-([1,1'-biphenyl]-4-yl)ethyl)-3-(2-oxo-4-(prop-1-en-2-ylamino)-2*H*-pyrido[1,2-*a*]pyrimidin-9-yl)benzamide (**37a**)



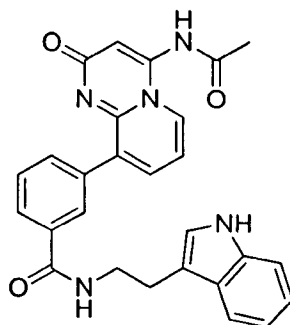
Following procedure **P2**, **18q** (90mg, 0.28mmol) was dissolved in DMF (8ml). 2-(4-biphenyl)ethylamine (66mg, 0.33mmol), CDMT (73mg, 0.42mmol) and 4-methylmorpholine (42mg, 0.42mmol) were added. Product **37a** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.31) and obtained as a white solid (65mg, yield 46%) [mp. dec >260°C]; Anal. HPLC: t_R = 8.8 min, 10-100% MeCN over 20min, purity 91%); ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 10.67 (1H, br, N-H), 8.66 (1H, t, J =5.58Hz, N-H), 8.32 (1H, d, J =7.10Hz, Ar-H), 7.99 (1H, m, Ar-H), 7.87 (1H, m, Ar-H), 7.76 (1H, m, Ar-H), 7.64 (2H, dd, J_a =1.18Hz, J_b =8.28Hz, Ar-H), 7.61 (2H, d, J =8.28Hz, Ar-H), 7.55 (1H, t, J =7.76Hz, Ar-H), 7.45 (2H, t, J =7.62Hz, Ar-H), 7.34 (3H, m, Ar-H), 7.03 (1H, t, J =7.10Hz, Ar-H), 6.33 (1H, s, Ar-H), 3.54 (2H, q, J =7.07Hz, CH_2), 2.91 (2H, t, J =7.07Hz, CH_2), 2.18 (3H, s, CH_3); HRMS (ES) m/z : found 503.2068 ($\text{C}_{31}\text{H}_{27}\text{N}_4\text{O}_3$ $[\text{M}+\text{H}]^+$), requires 503.2083.

N-(4-ethylphenethyl)-3-(2-oxo-4-(prop-1-en-2-ylamino)-2*H*-pyrido[1,2-*a*]pyrimidin-9-yl) benzamide (**37b**)



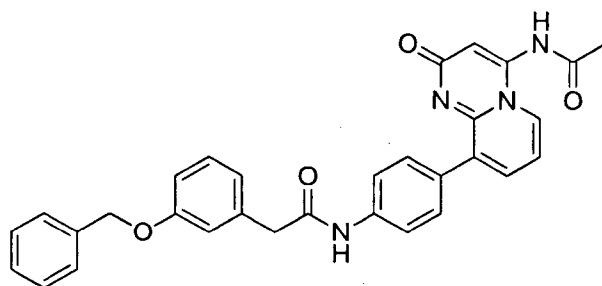
Following procedure **P2**, **18q** (65mg, 0.20mmol) was dissolved in DMF (6ml). 4-ethylphenethylamine (36mg, 0.24mmol), CDMT (53mg, 0.30mmol) and 4-methylmorpholine (30mg, 0.30mmol) were added. Product **37b** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.27) and obtained as a cream solid (50.6mg, yield 55%) [mp. 195.5-197.3°C]; Anal. HPLC: t_R = 11.4 min, 10-100% MeCN over 14min, purity 100%); ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 10.71 (1H, br, N-H), 8.62 (1H, t, J =5.62Hz, N-H), 8.33 (1H, d, J =7.08Hz, Ar-H), 7.98 (1H, m, Ar-H), 7.86 (1H, m, Ar-H), 7.77 (2H, m, Ar-H), 7.54 (1H, t, J =7.58Hz, Ar-H), 7.15 (4H, m, Ar-H), 7.05 (1H, t, J =7.08Hz, Ar-H), 6.33 (1H, s, Ar-H), 3.47 (2H, q, J =7.05Hz, CH_2), 2.82 (2H, t, J =7.05Hz, CH_2), 2.56 (2H, q, J =7.58Hz, CH_2), 2.19 (3H, s, CH_3), 1.17 (3H, t, J =7.58Hz, CH_3); HRMS (ES) m/z : found 455.2075 ($\text{C}_{27}\text{H}_{27}\text{N}_4\text{O}_3$ $[\text{M}+\text{H}]^+$), requires 455.2083.

N-(2-(1H-indol-3-yl)ethyl)-3-(2-oxo-4-(prop-1-en-2-ylamino)-2H-pyrido[1,2-*a*]pyrimidin-9-yl)benzamide (**37c**)



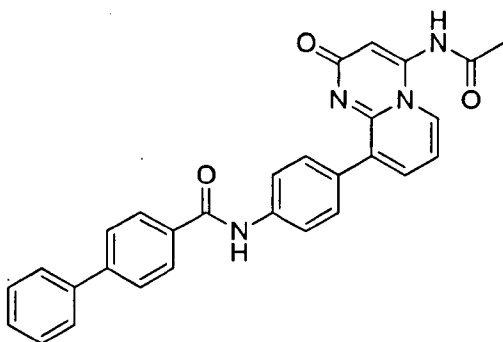
Following procedure **P2**, **18q** (65mg, 0.20mmol) was dissolved in DMF (6ml). Tryptamine (39mg, 0.24mmol), CDMT (53mg, 0.30mmol) and 4-methylmorpholine (30mg, 0.30mmol) were added. Product **37c** was purified using column chromatography (MeOH:DCM 1:10; R_f = 0.22) and obtained as a white solid (17.6mg, yield 19%) [mp. dec >250°C]; ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 10.82 (1H, s, N-H), 10.72 (1H, br, N-H), 8.66 (1H, t, J =5.62Hz, N-H), 8.33 (1H, d, J =7.08Hz, Ar-H), 8.01 (1H, m, Ar-H), 7.88 (1H, m, Ar-H), 7.77 (2H, m, Ar-H), 7.59 (1H, d, J =7.76Hz, Ar-H), 7.55 (1H, t, J =7.70Hz, Ar-H), 7.34 (1H, d, J =8.03Hz, Ar-H), 7.19 (1H, d, J =7.15Hz, Ar-H), 7.02 (3H, m, Ar-H), 6.33 (1H, s, Ar-H), 3.56 (2H, q, J =7.23Hz, CH_2), 2.97 (2H, t, J =7.23Hz, CH_2), 2.19 (3H, s, CH_3); HRMS (ES) m/z : found 466.1871 ($\text{C}_{27}\text{H}_{24}\text{N}_5\text{O}_3$ [$\text{M}+\text{H}$] $^+$), requires 466.1879.

N-(4-(4-acetamido-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-9-yl)phenyl)-3-(benzyloxy)benzamide (**38a**)



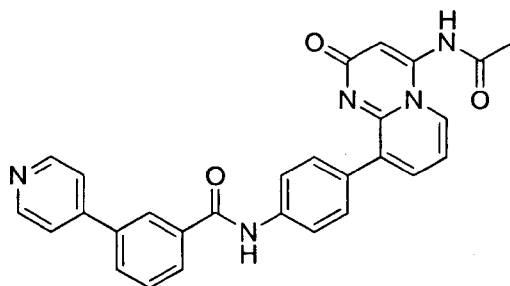
Following procedure **P2**, 3-benzyloxyphenyl)acetic acid (80mg, 0.33mmol) was dissolved in DMF (9ml). **18g** (117mg, 0.39mmol), CDMT (87mg, 0.50mmol) and 4-methylmorpholine (50mg, 0.50mmol) were added. Product **38a** was purified using column chromatography (MeOH:DCM 1:10; R_f = 0.27) and obtained as a cream solid (19mg, yield 11%) [mp. dec >250°C]; Anal. HPLC: t_R = 8.8 min, 10-100% MeCN over 11min, purity 97%; ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 10.44 (1H, br, N-H), 10.29 (1H, s, N-H), 8.31 (1H, d, J =7.08Hz, Ar-H), 7.68 (3H, m, Ar-H), 7.56 (2H, d, J =8.40Hz, Ar-H), 7.46 (2H, d, J =7.10Hz, Ar-H), 7.37 (3H, m, Ar-H), 7.25 (1H, t, J =7.84Hz, Ar-H), 7.03 (1H, s, Ar-H), 7.00 (1H, t, J =7.08Hz, Ar-H), 6.93 (2H, m, Ar-H), 6.33 (1H, s, Ar-H), 5.10 (2H, s, CH_2), 3.65 (2H, s, CH_2), 2.17 (3H, s, CH_3); HRMS (ES) m/z : found 541.1863 ($\text{C}_{31}\text{H}_{26}\text{N}_4\text{NaO}_4$ $[\text{M}+\text{Na}]^+$), requires 541.1852.

N-(4-(4-acetamido-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-9-yl)phenyl)-[1,1'-biphenyl]-4-carboxamide (**38b**)



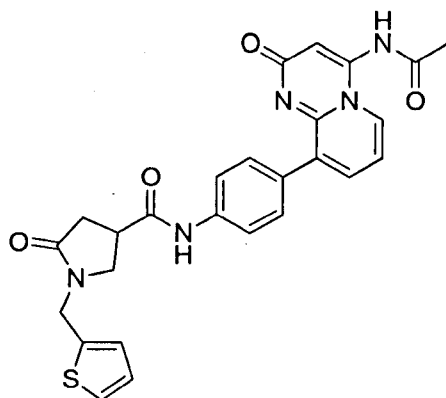
Following procedure **P2**, biphenyl-4-carboxylic acid (80mg, 0.40mmol) was dissolved in DMF (8ml). **18g** (143mg, 0.48mmol), CDMT (106mg, 0.61mmol) and 4-methylmorpholine (61mg, 0.61mmol) were added. Product **38b** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.24) and obtained as a brown solid (15.7mg, yield 8%) [mp. dec >230°C]; Anal. HPLC: t_R = 8.7 min, 10-100% MeCN over 11min, purity 99%); ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 10.70 (1H, br, N-H), 10.45 (1H, s, N-H), 8.29 (1H, d, J =7.08Hz, Ar-H), 8.11 (2H, d, J =8.34Hz, Ar-H), 7.89 (4H, m, Ar-H), 7.77 (3H, m, Ar-H), 7.63 (2H, d, J =8.59Hz, Ar-H), 7.53 (2H, t, J =7.83Hz, Ar-H), 7.44 (1H, m, Ar-H), 7.03 (1H, t, J =7.08Hz, Ar-H), 6.33 (1H, s, Ar-H), 2.19 (3H, s, CH₃); HRMS (ES) m/z : found 475.1768 (C₂₉H₂₃N₄O₃ [M+H]⁺), requires 475.1770.

N-(4-(4-acetamido-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-9-yl)phenyl)-3-(pyridin-4-yl)benzamide (**38c**)



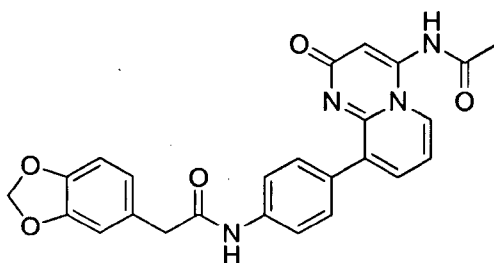
Following procedure **P2**, 3-(pyridin-4-yl)benzoic acid (80mg, 0.40mmol) was dissolved in DMF (8ml). **18g** (130mg, 0.44mmol), CDMT (105mg, 0.60mmol) and 4-methylmorpholine (61mg, 0.60mmol) were added. Product **38c** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.33) and obtained as a cream solid (9.3mg, yield 5%) [mp. dec >250°C]; Anal. HPLC: t_R = 6.4 min, 10-100% MeCN over 11min, purity 99%); ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 10.67 (1H, br, N-H), 10.53 (1H, s, N-H), 8.71 (2H, d, J =6.08Hz, Ar-H), 8.37 (1H, s, Ar-H), 8.29 (1H, d, J =7.08Hz, Ar-H), 8.07 (3H, m, Ar-H), 7.89 (2H, d, J =8.44Hz, Ar-H), 7.86 (1H, d, J =6.12Hz, Ar-H), 7.74 (2H, m, Ar-H), 7.65 (2H, d, J =8.44Hz, Ar-H), 7.02 (1H, t, J =7.08Hz, Ar-H), 6.32 (1H, s, Ar-H), 2.19 (3H, s, CH₃); HRMS (ES) m/z : found 476.1721 (C₂₈H₂₂N₅O₃ [M+H]⁺), requires 476.1723.

N-(4-(4-acetamido-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-9-yl)phenyl)-5-oxo-1-(thiophen-2-ylmethyl)pyrrolidine-3-carboxamide (**38d**)



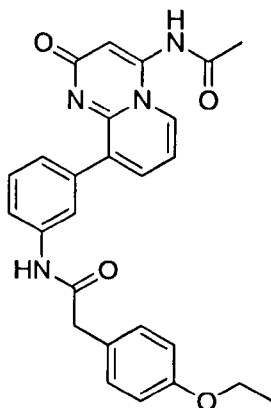
Following procedure **P2**, 3-oxo-1-(2-thienylmethyl)pyrrolidine-3-carboxylic acid (90mg, 0.40mmol) was dissolved in DMF (8ml). **18g** (141mg, 0.48mmol), CDMT (105mg, 0.60mmol) and 4-methylmorpholine (61mg, 0.60mmol) were added. Product **38d** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.13) and obtained as a cream solid (10.7mg, yield 5%) [mp. dec >250°C]; Anal. HPLC: t_R = 7.1 min, 10-100% MeCN over 11min, purity 95%); ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 10.65 (1H, br, N-H), 10.24 (1H, s, N-H), 8.27 (1H, d, J =7.08Hz, Ar-H), 7.70 (1H, dd, J_a =1.40Hz, J_b =7.08Hz, Ar-H), 7.65 (2H, d, J =8.70Hz, Ar-H), 7.56 (2H, d, J =8.70Hz, Ar-H), 7.46 (1H, m, Ar-H), 7.05 (1H, dd, J_a =0.86Hz, J_b =3.48Hz, Ar-H), 7.00 (2H, m, Ar-H), 6.31 (1H, s, Ar-H), 4.59 (2H, s, CH₂), 3.50 (2H, m, CH₂), 2.98 (1H, m, CH), 2.59 (2H, d, J =8.34Hz, CH₂), 2.18 (3H, s, CH₃); HRMS (ES) m/z : found 502.1541 (C₂₆H₂₄N₅O₄S [M+H]⁺), requires 502.1549.

N-(4-(4-acetamido-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-9-yl)phenyl)-2-(benzo[*d*][1,3]dioxol-5-yl)acetamide (**38e**)



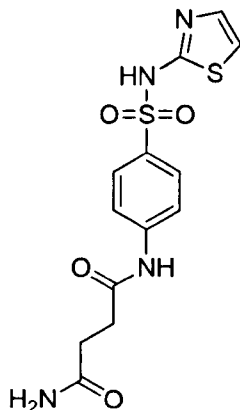
Following procedure **P2**, 3,4-methylenedioxyphenylacetic acid (50mg, 0.28mmol) was dissolved in DMF (8ml). **18g** (98mg, 0.33mmol), CDMT (73mg, 0.42mmol) and 4-methylmorpholine (42mg, 0.42mmol) were added. Product **38e** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.25) and obtained as a cream solid (18.6mg, yield 15%) [mp. dec >250°C]; Anal. HPLC: t_R = 8.2 min, 10-100% MeCN over 11min, purity 97%); ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 10.64 (1H, br, N-H), 10.25 (1H, s, N-H), 8.26 (1H, d, J =7.08Hz, Ar-H), 7.69 (1H, dd, J_a =1.05Hz, J_b =7.08Hz, Ar-H), 7.66 (2H, d, J =8.58Hz, Ar-H), 7.56 (2H, d, J =8.58Hz, Ar-H), 6.99 (1H, t, J =7.08Hz, Ar-H), 6.93 (1H, s, Ar-H), 6.87 (1H, d, J =7.85Hz, Ar-H), 6.81 (1H, dd, J_a =1.50Hz, J_b =7.86Hz, Ar-H), 6.31 (1H, s, Ar-H), 5.99 (2H, s, CH₂), 3.59 (2H, s, CH₂), 2.18 (3H, s, CH₃); HRMS (ES) m/z : found 457.1519 (C₂₅H₂₁N₄O₅ [M+H]⁺), requires 457.1512.

N-(3-(4-acetamido-2-oxo-2*H*-pyrido[1,2-*a*]pyrimidin-9-yl)phenyl)-2-(4-ethoxyphenyl)acetamide (**39a**)



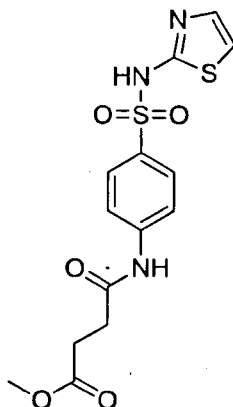
Following procedure **P2**, 4-ethoxyphenylacetic acid (59mg, 0.33mmol) was dissolved in DMF (6ml). **18d** (116mg, 0.39mmol), CDMT (86mg, 0.49mmol) and 4-methylmorpholine (50mg, 0.49mmol) were added. Product **39a** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.25) and obtained as a cream solid (16.8mg, yield 11%) [mp. dec >250°C]; Anal. HPLC: t_R = 7.8 min, 10-100% MeCN over 11min, purity 94%); ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 10.51 (1H, br, N-H), 10.39 (1H, s, N-H), 8.33 (1H, d, J =7.08Hz, Ar-H), 7.77 (1H, m, Ar-H), 7.68 (2H, m, Ar-H), 7.36 (1H, t, J =7.90Hz, Ar-H), 7.25 (2H, d, J =8.60Hz, Ar-H), 7.22 (1H, d, J =7.90Hz, Ar-H), 7.02 (1H, t, J =7.08Hz, Ar-H), 6.86 (2H, d, J =8.60Hz, Ar-H), 6.31 (1H, s, Ar-H), 3.99 (2H, q, J =7.15Hz, CH_2), 3.59 (2H, s, CH_2), 2.20 (3H, s, CH_3), 1.04 (3H, t, J =7.15Hz, CH_3); HRMS (ES) m/z : found 457.1870 ($\text{C}_{26}\text{H}_{25}\text{N}_4\text{O}_4$ $[\text{M}+\text{H}]^+$), requires 457.1876.

*N*¹-(4-(*N*-(thiazol-2-yl)sulfamoyl)phenyl)succinamide (**40b**)



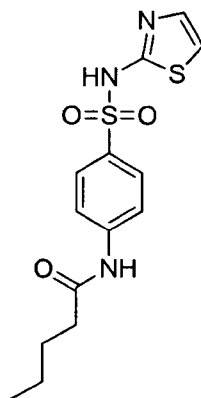
To a stirred solution of succinylsufathiazole (1400mg, 3.94mmol) in anhydrous toluene (30ml) was added thionyl chloride (939mg, 7.89mmol). The resulting mixture was refluxed at 100°C for 2h. Excess toluene was removed in vacuo and the resulting product added to concentrated ammonia solution (20ml) and stirred at 0°C for 1h. The product was precipitated out by neutralization with dilute hydrochloric acid and subsequently filtered and washed using hot ether and dried in vacuo to afford **40b** as a cream powder (39.7mg, yield 5%) [mp. 215.7-216.8°C]; Anal. HPLC: t_R = 7.9 min, 10-100% MeCN over 20min, purity 100%); ¹H NMR (400 MHz, DMSO-*d*₆): δ/ppm 12.75 (1H, br, N-H), 10.45 (1H, s, N-H), 7.72 (4H, m, Ar-H), 7.25 (2H, s, NH₂), 7.23 (1H, d, *J*=4.54Hz, Ar-H), 6.81 (1H, d, *J*=4.54Hz, Ar-H), 2.56 (2H, t, *J*=7.15Hz, CH₂), 2.39 (2H, t, *J*=7.12Hz, CH₂); ¹³C NMR (500 MHz, DMSO-*d*₆): δ/ppm 176.8, 173.7, 171.7, 143.0, 136.4, 127.3, 124.9, 118.8, 108.5, 32.1, 30.3; HRMS (ES) *m/z*: found 355.0502 (C₁₃H₁₅N₄O₄S₂ [M+H]⁺), requires 355.0535.

Methyl 4-oxo-4-((4-(*N*-(thiazol-2-yl)sulfamoyl)phenyl)amino)butanoate (**40c**)



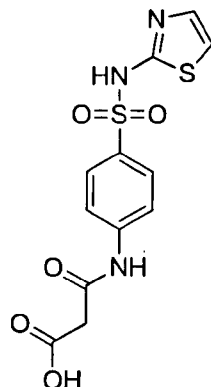
To a stirred solution of succinylsulfathiazole (1000mg, 2.82mmol) in anhydrous methanol (25ml) at 0°C was added concentrated sulphuric acid (276mg, 2.82mmol) over a 5min period. The resulting reaction mixture was stirred at RT for 4h and subsequently filtered. The compound was washed using saturated sodium bicarbonate solution and recrystallised from ethanol to afford **40c** as a white powder (119.1mg, yield 11%) [mp. 201.3-202.5°C]; Anal. HPLC: t_R = 14.5 min, 10-100% MeCN over 20min, purity 99%); ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 12.49 (1H, br, N-H), 10.33 (1H, s, N-H), 7.72 (4H, m, Ar-H), 7.23 (1H, d, $J=4.54\text{Hz}$, Ar-H), 6.79 (1H, d, $J=4.54\text{Hz}$, Ar-H), 3.60 (3H, s, CH_3), 2.63 (4H, m, CH_2); ^{13}C NMR (500 MHz, DMSO- d_6): δ /ppm 173.2, 170.9, 169.2, 142.7, 136.8, 127.4, 125.5, 118.9, 108.5, 51.8, 31.4, 28.8; HRMS (ES) m/z : found 370.0530 ($\text{C}_{14}\text{H}_{16}\text{N}_3\text{O}_5\text{S}_2$ $[\text{M}+\text{H}]^+$), requires 370.0531.

N-(4-(*N*-(thiazol-2-yl)sulfamoyl)phenyl)pentanamide (**40d**)



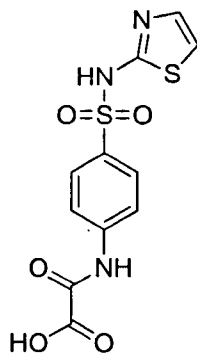
To a stirred solution of sulfathiazole (100mg, 0.39mmol) in DMF (15ml) was added sodium hydride (19mg, 0.47mmol) and the resulting mixture stirred at RT for 1h. Valeryl chloride (57mg, 0.47mmol) was subsequently added and the reaction mixture stirred for a further 2h. Water (40ml) was added to the reaction mixture and compound extracted using ethyl acetate (3 x 50ml). The organic layers were combined and dried over magnesium sulphate and solvent evaporated *in vacuo* to give **40d** as a white powder (73.7mg, yield 55%) [mp. 219.8-221.3°C]; Anal. HPLC: t_R = 15.6 min, 10-100% MeCN over 20min, purity 100%; ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 12.58 (1H, br, N-H), 10.20 (1H, s, N-H), 7.72 (4H, s, Ar-H), 7.24 (1H, d, J =4.54Hz, Ar-H), 6.80 (1H, d, J =4.54Hz, Ar-H), 2.33 (2H, t, J =7.54Hz, CH₂), 1.57 (2H, m, CH₂), 1.32 (2H, m, CH₂), 0.89 (3H, t, J =7.30Hz, CH₃); ^{13}C NMR (500 MHz, DMSO- d_6): δ /ppm 172.3, 169.2, 142.9, 136.6, 127.4, 125.2, 118.9, 108.5, 36.6, 27.6, 22.2, 14.2; HRMS (ES) m/z : found 340.0778 (C₁₄H₁₈N₃O₃S₂ [M+H]⁺), requires 340.0790.

3-oxo-3-((4-(*N*-(thiazol-2-yl)sulfamoyl)phenyl)amino)propanoic acid (**40e**)



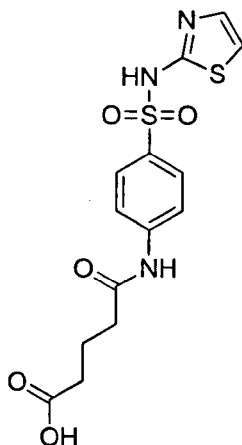
Compound **42** (300mg, 0.81mmol) was added to 2.5% sodium hydroxide solution (10ml). The resulting mixture was stirred at 90°C for 2h. The product was precipitated from the solution by neutralization with dilute hydrochloric acid and subsequently filtered and washed using diethyl ether (20ml) to give **40e** as a white powder (246.6mg, yield 89%) [mp. dec >250°C]; Anal. HPLC: t_R = 10.1 min, 10-100% MeCN over 20min, purity 96%); ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 11.83 (1H, s, N-H), 7.59 (5H, m, Ar-H, N-H), 6.90 (1H, d, J =3.80Hz, Ar-H), 6.42 (1H, d, J =3.80Hz, Ar-H), 4.23 (1H, br, OH), 2.99 (2H, s, CH₂); ^{13}C NMR (500 MHz, DMSO- d_6): δ /ppm 171.6, 170.2, 168.9, 140.9, 140.3, 137.7, 127.5, 118.2, 107.5, 31.2; HRMS (ES) m/z : found 342.0170 (C₁₂H₁₂N₃O₅S₂ [M+H]⁺), requires 342.0218.

2-oxo-2-((4-(*N*-(thiazol-2-yl)sulfamoyl)phenyl)amino)acetic acid (**40f**)



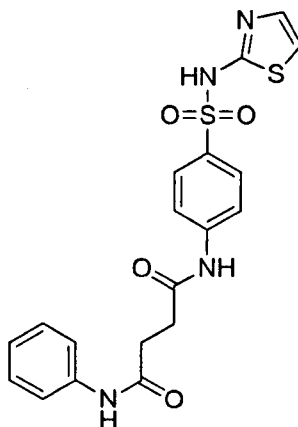
Compound **43** (350mg, 0.99mmol) was added to 2.5% sodium hydroxide solution (10ml). The resulting mixture was stirred at 90°C for 2h. The product was precipitated from the solution by neutralization with dilute hydrochloric acid and subsequently filtered and washed using diethyl ether (20ml) to give **40f** as a white powder (284.9mg, yield 88%) [mp. dec >270°C]; Anal. HPLC: t_R = 1.3 min, 10-100% MeCN over 20min, purity 100%); ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 12.61 (1H, br, N-H), 10.35 (1H, s, N-H), 7.70 (4H, m, Ar-H), 6.90 (1H, d, J =3.74Hz, Ar-H), 6.41 (1H, d, J =3.74Hz, Ar-H), 5.43 (1H, s, OH); HRMS (ES) m/z : found 328.0003 ($\text{C}_{11}\text{H}_{10}\text{N}_3\text{O}_5\text{S}_2$ $[\text{M}+\text{H}]^+$), requires 328.0062.

5-oxo-5-((4-(*N*-(thiazol-2-yl)sulfamoyl)phenyl)amino)pentanoic acid (**40g**)



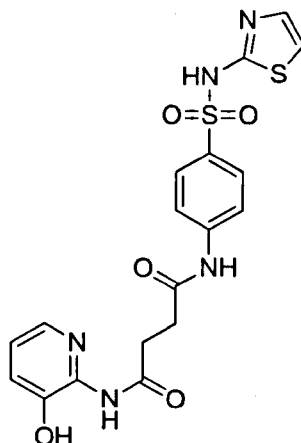
To a stirred solution of sulfathiazole (100mg, 0.39mmol) in anhydrous dichloromethane (20ml) was added glutaric anhydride (54mg, 0.47mmol) and triethylamine (48mg, 0.47mmol). The resulting mixture was refluxed at 80°C for 2h and solvent evaporated *in vacuo*. The resulting product was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.02) and obtained as a cream solid (89.3mg, yield 62%) [mp. dec >250°C]; Anal. HPLC: t_R = 12.2 min, 10-100% MeCN over 20min, purity 95%; ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 10.24 (1H, s, N-H), 7.75 (1H, s, N-H), 7.67 (4H, m, Ar-H), 7.07 (1H, d, J =3.75Hz, Ar-H), 6.62 (1H, d, J =3.75Hz, Ar-H), 5.74 (1H, br, OH), 2.37 (2H, t, J =7.08Hz, CH_2), 1.79 (2H, t, J =7.08Hz, CH_2), 1.70 (2H, m, CH_2); ^{13}C NMR (500 MHz, DMSO- d_6): δ /ppm 175.1, 171.8, 169.7, 142.0, 138.4, 127.3, 125.3, 118.7, 108.1, 36.0, 33.9, 20.9; HRMS (ES) m/z : found 370.0526 ($\text{C}_{14}\text{H}_{16}\text{N}_3\text{O}_5\text{S}_2$ [$\text{M}+\text{H}$] $^+$), requires 370.0531.

*N*¹-phenyl-*N*⁴-(4-(*N*-(thiazol-2-yl)sulfamoyl)phenyl)succinamide (**41a**):



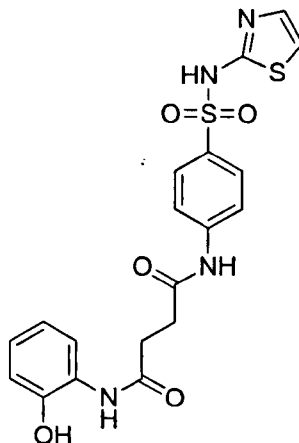
Following procedure **P2**, succinylsulfathiazole (100mg, 0.28mmol) was dissolved in DMF (8ml). Aniline (52mg, 0.56mmol), CDMT (74mg, 0.42mmol) and 4-methylmorpholine (43mg, 0.42mmol) were added. Product **41a** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.49) and obtained as a cream solid (23.9mg, yield 20%) [mp. dec >230°C]; Anal. HPLC: t_R = 15.3 min, 10-100% MeCN over 20min, purity 98%); ¹H NMR (400 MHz, DMSO-*d*₆): δ/ppm 11.93 (1H, br, N-H), 10.36 (1H, s, N-H), 10.01 (1H, s, N-H), 7.72 (4H, s, Ar-H), 7.58 (2H, d, J =7.60Hz, Ar-H), 7.28 (2H, t, J =7.64Hz, Ar-H), 7.23 (1H, d, J =4.56Hz, Ar-H), 7.02 (1H, t, J =7.63Hz, Ar-H), 6.79 (1H, d, J =4.56Hz, Ar-H), 2.67 (4H, t, J =4.48Hz, CH₂); ¹³C NMR (500 MHz, DMSO-*d*₆): δ/ppm 171.5, 170.7, 169.2, 142.8, 139.8, 136.6, 129.1, 127.4, 125.4, 123.4, 119.3, 118.8, 108.5, 31.7, 31.4; HRMS (ES) m/z : found 453.0656 (C₁₉H₁₈N₄NaO₄S₂ [M+Na]⁺), requires 453.0667.

*N*¹-(3-hydroxypyridin-2-yl)-*N*⁴-(4-(*N*-(thiazol-2-yl)sulfamoyl)phenyl)succinamide (**41b**)



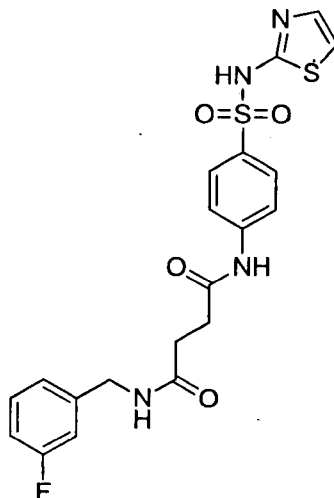
Following procedure **P2**, succinylsulfathiazole (100mg, 0.28mmol) was dissolved in DMF (8ml). 2-amino-3-hydroxypyridine (62mg, 0.56mmol), CDMT (74mg, 0.42mmol) and 4-methylmorpholine (43mg, 0.42mmol) were added. Product **41b** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.31) and obtained as a white solid (42.4mg, yield 34%) [mp. dec >250°C]; Anal. HPLC: t_R = 7.68 min, 10-100% MeCN over 11min, purity 95%); ¹H NMR (400 MHz, DMSO-*d*₆): δ/ppm 10.63 (1H, br, N-H), 10.24 (1H, s, N-H), 9.41 (1H, br, N-H), 7.67 (4H, m, Ar-H), 8.21 (1H, dd, J_a =1.52Hz, J_b =4.55Hz, Ar-H), 7.30 (1H, m, Ar-H), 7.09 (1H, d, J =4.50Hz, Ar-H), 6.82 (1H, dd, J_a =1.52Hz, J_b =7.58Hz, Ar-H), 6.63 (1H, d, J =4.50Hz, Ar-H), 5.38 (1H, br, OH), 2.69 (2H, m, CH₂), 2.56 (2H, m, CH₂); HRMS (ES) m/z : found 448.0736 (C₁₈H₁₈N₅O₅S₂ [M+H]⁺), requires 448.0749.

*N*¹-(2-hydroxyphenyl)-*N*⁴-(4-(*N*-(thiazol-2-yl)sulfamoyl)phenyl)succinamide (**41c**)



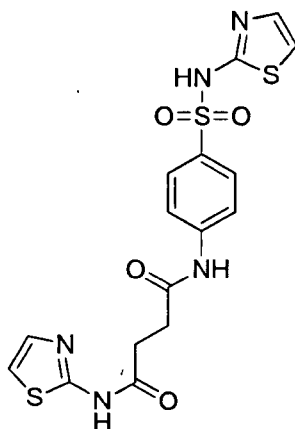
Following procedure **P2**, succinylsulfathiazole (100mg, 0.28mmol) was dissolved in DMF (8ml). 2-amino phenol (61mg, 0.56mmol), CDMT (74mg, 0.42mmol) and 4-methylmorpholine (43mg, 0.42mmol) were added. Product **41c** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.42) and obtained as a cream solid (13.2mg, yield 11%) [mp. dec >250°C] Anal. HPLC: t_R = 15.0 min, 10-100% MeCN over 20min, purity 96%); ¹H NMR (400 MHz, DMSO-*d*₆): δ/ppm 11.19 (1H, br, N-H), 10.30 (1H, s, N-H), 9.72 (1H, s, N-H), 7.70 (4H, s, Ar-H), 7.18 (1H, d, J =4.48Hz, Ar-H), 6.87 (4H, m, Ar-H), 6.73 (1H, d, J =4.48Hz, Ar-H), 3.86 (1H, br, OH), 2.70 (4H, m, CH₂); HRMS (ES) m/z : found 447.0795 (C₁₉H₁₉N₄O₅S₂ [M+H]⁺), requires 447.0797.

*N*¹-(3-fluorobenzyl)-*N*⁴-(4-(*N*-(thiazol-2-yl)sulfamoyl)phenyl)succinamide (**41d**)



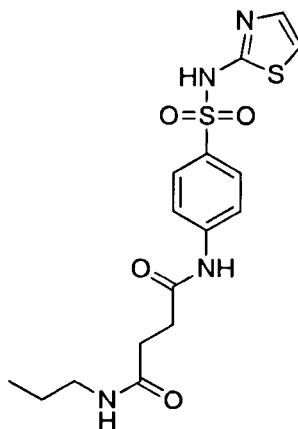
Following procedure **P2**, succinylsulfathiazole (100mg, 0.28mmol) was dissolved in DMF (8ml). 3-fluorobenzylamine (70mg, 0.56mmol), CDMT (74mg, 0.42mmol) and 4-methylmorpholine (43mg, 0.42mmol) were added. Product **41d** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.47) and obtained as a cream solid (23.9mg, yield 18%) [mp. dec >250°C]; Anal. HPLC: t_R = 15.2 min, 10-100% MeCN over 20min, purity 96%); ¹H NMR (400 MHz, DMSO-*d*₆): δ /ppm 11.93 (1H, br, N-H), 10.30 (1H, s, N-H), 8.46 (1H, t, J =5.94Hz, N-H), 7.72 (4H, s, Ar-H), 7.32 (1H, m, Ar-H), 7.22 (1H, d, J =4.54Hz, Ar-H), 7.06 (3H, m, Ar-H), 6.79 (1H, d, J =4.54Hz, Ar-H), 4.28 (2H, d, J =5.94Hz, CH₂), 2.63 (2H, t, J =6.88Hz, CH₂), 2.48 (2H, t, J =6.88Hz, CH₂); ¹³C NMR (500 MHz, DMSO-*d*₆): δ /ppm 171.8, 171.5, 169.8, 163.7, 143.2, 143.1, 142.8, 136.7, 127.4, 125.7, 123.5, 123.4, 118.8, 114.1, 108.5, 42.0, 32.1, 30.5; HRMS (ES) m/z : found 463.0914 (C₂₀H₂₀FN₄O₄S₂ [M+H]⁺), requires 463.0910.

*N*¹-(thiazol-2-yl)-*N*⁴-(4-(*N*-(thiazol-2-yl)sulfamoyl)phenyl)succinamide (**41e**)



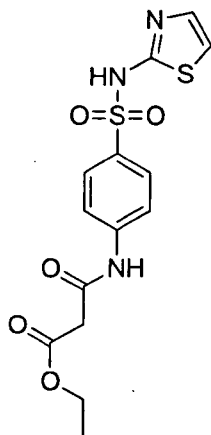
Following procedure **P2**, succinylsulfathiazole (100mg, 0.28mmol) was dissolved in DMF (8ml). 2-aminothiazole (56mg, 0.56mmol), CDMT (74mg, 0.42mmol) and 4-methylmorpholine (43mg, 0.42mmol) were added. Product **41e** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.42) and obtained as a white solid (12.7mg, yield 10%) [mp. dec >250°C]; Anal. HPLC: t_R = 3.8 min, 10-100% MeCN over 20min, purity 100%); ¹H NMR (400 MHz, DMSO-*d*₆): δ /ppm 12.14 (1H, br, N-H), 11.15 (1H, br, N-H), 10.28 (1H, s, N-H), 7.67 (4H, m, Ar-H), 7.45 (1H, d, J =3.56Hz, Ar-H), 7.18 (1H, d, J =3.56Hz, Ar-H), 7.10 (1H, d, J =4.20Hz, Ar-H), 6.64 (1H, d, J =4.20Hz, Ar-H), 2.72 (4H, m, CH₂); HRMS (ES) m/z : found 438.0344 (C₁₆H₁₆N₅O₄S₃ [M+H]⁺), requires 438.0364.

*N*¹-propyl-*N*⁴-(4-(*N*-(thiazol-2-yl)sulfamoyl)phenyl)succinamide (**41f**)



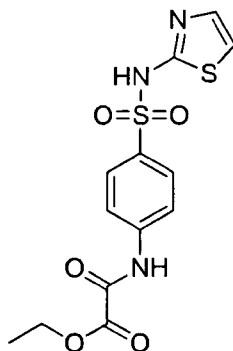
Following procedure **P2**, succinylsulfathiazole (100mg, 0.28mmol) was dissolved in DMF (10ml). Propylamine (20mg, 0.34mmol), CDMT (74mg, 0.42mmol) and 4-methylmorpholine (43mg, 0.42mmol) were added. Product **41f** was purified using preparative TLC (MeOH:DCM 1:10; R_f = 0.42) and obtained as a white solid (15.7mg, yield 14%) [mp. dec >250°C]; Anal. HPLC: t_R = 13.8 min, 10-100% MeCN over 20min, purity 100%; ¹H NMR (400 MHz, DMSO-*d*₆): δ /ppm 12.38 (1H, br, N-H), 10.28 (1H, s, N-H), 7.85 (1H, t, J =5.30Hz, N-H), 7.71 (4H, s, Ar-H), 7.24 (1H, d, J =4.60Hz, Ar-H), 6.81 (1H, d, J =4.60Hz, Ar-H), 2.99 (2H, q, J =6.87Hz, CH₂), 2.57 (2H, t, J =7.10Hz, CH₂), 2.40 (2H, t, J =7.10Hz, CH₂), 1.39 (2H, m, CH₂), 0.83 (3H, t, J =7.40Hz, CH₃); ¹³C NMR (500 MHz, DMSO-*d*₆): δ /ppm 173.3, 171.6, 171.3, 142.9, 136.5, 127.4, 125.4, 118.8, 108.5, 40.8, 32.2, 30.5, 22.9, 11.9; HRMS (ES) m/z : found 397.0996 (C₁₆H₂₁N₄O₄S₂ [M+H]⁺), requires 397.0999.

Ethyl 3-oxo-3-((4-(*N*-(thiazol-2-yl)sulfamoyl)phenyl)amino)propanoate (**42**)



Sulfathiazole (860mg, 3.37mmol) was added to diethyl malonate (1079mg, 6.75mmol) and heated under reflux for 2h at 150°C. After cooling the reaction mixture to RT, the solid product was filtered and washed using diethyl ether (20ml) and dried *in vacuo*. The product was subsequently recrystallised from ethanol to give **42** as a white powder (891.5mg, yield 72%) [mp. 205.3-205.8°C]; Anal. HPLC: t_R = 14.9 min, 10-100% MeCN over 20min, purity 90%); ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 12.70 (1H, br, N-H), 10.51 (1H, s, N-H), 7.73 (4H, m, Ar-H), 7.25 (1H, d, J =4.60Hz, Ar-H), 6.82 (1H, d, J =4.60Hz, Ar-H), 4.12 (2H, q, J =7.10Hz, CH₂), 3.49 (2H, s, CH₂), 1.20 (3H, t, J =7.10Hz, CH₃); HRMS (ES) m/z : found 370.0525 (C₁₄H₁₆N₃O₅S₂ [M+H]⁺), requires 370.0531.

Ethyl 2-oxo-2-((4-(*N*-(thiazol-2-yl)sulfamoyl)phenyl)amino)acetate (**43**)



To a stirred solution of sulfathiazole (1000mg, 3.92mmol) in ethanol (15ml) was added diethyl oxalate (1145mg, 7.84mmol). The resulting mixture was stirred under reflux for 2h at 110°C. After cooling the reaction mixture to RT, the solid product was filtered and washed using diethyl ether (30ml) to give **43** as a cream powder (720.9mg, yield 52%) [mp. 244.9-245.2°C]; Anal. HPLC: t_R = 7.7 min, 10-100% MeCN over 12min, purity 95%; ^1H NMR (400 MHz, DMSO- d_6): δ /ppm 12.73 (1H, br, N-H), 11.07 (1H, s, N-H), 7.85 (4H, m, Ar-H), 7.26 (1H, d, J =4.54Hz, Ar-H), 6.83 (1H, d, J =4.54Hz, Ar-H), 4.32 (2H, q, J =7.10Hz, CH₂), 1.32 (3H, t, J =7.10Hz, CH₃); HRMS (ES) m/z : found 356.0358 (C₁₃H₁₄N₃O₅S₂ [M+H]⁺), requires 356.0375.

10.2 Computational procedure

10.2.1 Docking methodology

The docking methodology is described in three parts, setting up the protein and ligand structures along with setting parameters for the docking procedure.

10.2.1.1 Protein preparation

The crystal structure of human POT1 bound to telomeric single stranded DNA (1xjv) was downloaded from the Brookhaven protein databank (www.pdb.org). The structure contained some missing protein residues (Residues PRO-146, SER-147 and TRP-148). The structure was processed using SWISS-MODEL (Schwede *et al*, 2003), an automated protein structure homology-modelling server to obtain the complete POT1 crystal structure. The missing residues obtained using the SWISS-MODEL were then incorporated back into the original crystal structure pdb file. This was done to ensure that the structure is kept close to the original pdb structural data. The ionisation/tautomeric states of the amino acid side chains were assigned using the WHATIF web server (<http://swift.cmbi.ru.nl/servers/html/index.html>). Hydrogen atoms were added to the protein and the protein structure was subsequently minimised in the presence of DNA using the TIP3P explicit solvent model and Amber ff03 parameters (Duan *et al*, 2003).

10.2.1.2 Ligand preparation

All ligand structures were designed using the 'sketch molecule' tool in SYBYL 8.0 and assigned appropriate atom and bond types. Hydrogen atoms were subsequently added and the full geometry optimisation of the constructed structures was performed using the standard Tripos force field with a distance-dependent dielectric function and a 0.001 kcal/(mol*Å) energy gradient convergence criterion with 10000 iterations. Gasteiger-Huckel charges were assigned to each structure for the calculation of electrostatic interactions. The lowest energy conformation for each structure was obtained and used for docking. For the virtual screening, ligands were selected from the ZINC database. A subset of drug-like structures were selected resulting in a database of 55000 structures with the following molecular constraints: xlogP <5, molecular weight >150 and <500 g/mol, H-bond acceptors <10, rotatable bonds <8 and polar surface area <150Å². The 3D structures of these compounds were generated using the Concord module in SYBYL 8.0.

10.2.1.3 GOLD docking procedure

Docking runs were carried out using GOLD V3.0.1. The ten nucleotide DNA molecule bound to human POT1 was deleted and the binding site was selected using active site radius in GOLD. All residues within 10Å of serine-238 in POT1 were selected as binding site residues which was found to be optimal. In this study, the GOLD default genetic algorithm settings were used where GOLD performs 10 docking runs for each compound, each consisting of 10 000 genetic algorithm operations. GOLD was allowed to terminate early if the top three dockings were within 1.5Å RMSD of each

other. GOLD allowed the full flexibility of the ligands and partial flexibility of POT1 which was limited to the hydroxyl groups of serine, tyrosine and threonine.

10.3 Biological screening procedure

10.3.1 Materials

The materials for the assay were purchased from the following chemical suppliers: The DNA oligonucleotides (15mer) were purchased Sigma Aldrich, United Kingdom (Table 12).

| Oligo name | Sequence | Position of fluorescein label |
|-----------------|----------------------------|-------------------------------|
| 5''Flou Tel seq | FLCNTTAGGGTTAGGGTTA | N-terminal serine residue |
| Telseq Oligo 2 | TTAGGGTTAGGGTTA | Unlabelled |

Table 12: DNA oligonucleotides used in this study.

Pierce® centrifuge columns (5ml) were purchased from Thermo Scientific, glutathione sepharase beads were purchased from GE Healthcare Bio-Sciences, Complete Mini EDTA-free protease inhibitor cocktail tablets were purchased from Roche and glutathione reduced was purchased from Fisher scientific.

10.3.2 Expression and purification of human POT1 DNA (hPOT1-OB) binding domain as a GST fusion protein

The pGEX6p-1OB based expression construct encoding the hPOT1 OB domain containing residues 1-294 with an N-terminal GST tag was generously provided by Dr Lodewyk Dekker, University of Nottingham. The constructed plasmids were transformed into *Escherichia coli* BL21 (DE3) cells (Novagen, USA). Overnight growth of 5ml cultures with Ampicillin (1µl/ml) in LB broth was carried out at 37°C inoculated from a single colony. These overnight cultures were used to inoculate larger 500ml cultures and grown at 37°C until OD₆₀₀ reached 0.6-0.8. The temperature was then subsequently decreased to 25°C. After approximately two hours, protein expression was induced using 1mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) overnight. Cells were harvested by centrifugation at 3000 rotations per minute (RPM) for 30 minutes and thawed pellets were resuspended in 20ml GST binding buffer (25mM Tris-HCl, pH 7.5, 150mM NaCl, 1mM EDTA) containing complete mini EDTA-free protease inhibitor cocktail tablets. Cells were lysed by sonication (20 micron amplitude, 12x30 seconds burst with 1 minute ice rest time in between) and centrifuged at 15000RPM for 50 minutes in a Beckman ultracentrifuge. The clear supernatant was incubated with glutathione sepharase beads (1ml) for 2 hours at 4°C with rotation. The beads were subsequently poured onto a 5ml Pierce® centrifuge column to allow unbound material to run through and washed using GST binding buffer (10ml). The beads were then incubated with 5ml GST elution buffer (50mM Tris-HCl, pH 8, 5% glycerol, 10mM reduced glutathione) for 15 minutes at room temperature with shaking. The

bound protein was subsequently eluted and collected and the glutathione sepharase beads were further washed with 5ml GST elution buffer. The homogeneity of the protein preparation was analysed by SDS-PAGE.

10.3.3 Fluorescence Polarisation assay procedure

The following solutions were prepared: a 200 μ M DNA Oligo stock solution was prepared in distilled water and diluted to 100nM by adding 1 μ l of 200 μ M DNA Oligo to 2ml reaction buffer (50mM Tris, pH 8, 50mM NaCl, 10mM MgCl₂). A 100 μ M flourescein stock solution was prepared in distilled water and diluted to 100nM by adding 5 μ l flourescein (100 μ M) to 250 μ l elution buffer and 250 μ l reaction buffer. The compounds were prepared as 10mM stock solutions in 100% DMSO and subsequently diluted 1 in 10 to make 1mM solutions in elution buffer (50mM Tris-HCl, pH 8, 5% glycerol, 10% DMSO, and 10mM reduced glutathione). For dose-response curves a series of dilutions were carried out to give 100 μ M, 10 μ M, 1 μ M, 0.1 μ M and 0.01 μ M.

Fluorescence polarisation (FP) assays were conducted in elution buffer (50mM Tris-HCl, pH 8, 5% glycerol and 10mM reduced glutathione). The fluorescein labelled DNA Oligo used was **FLCNTTAGGGTTAGGGTTA**. The FP assays were carried out using 10nM fluorescein labelled oligo. For POT1-DNA inhibitor assay, small molecules were pre-incubated with the POT1 protein for 5 minutes. The labelled oligo was then added and incubated for 10 minutes. All FP measurements were carried out in a 384-well low-volume black round-bottom polystyrene NBS microplate using

an EnVision multilabel plate reader. The polarization values are reported in millipolarization units (mP) and were measured at an excitation wavelength of 480nm and an emission wavelength of 535 nm. The unlabelled DNA Oligo with the sequence TTAGGGTTAGGGTTA was used as a positive control. Each test compound was tested in triplicate. The final DMSO concentration was maintained at 1% in all assays.

To calculate the percentage inhibition, the difference between the fluorescence values of reactions containing DNA and POT1 and reactions containing DNA, POT1, and the test compound was divided by the difference between the fluorescence values of reactions with DNA with POT1 and the fluorescence of the DNA alone. This value was then multiplied by 100 to give percentage inhibition.

Appendices

Appendix A

POT1-DNA interaction energy graphs

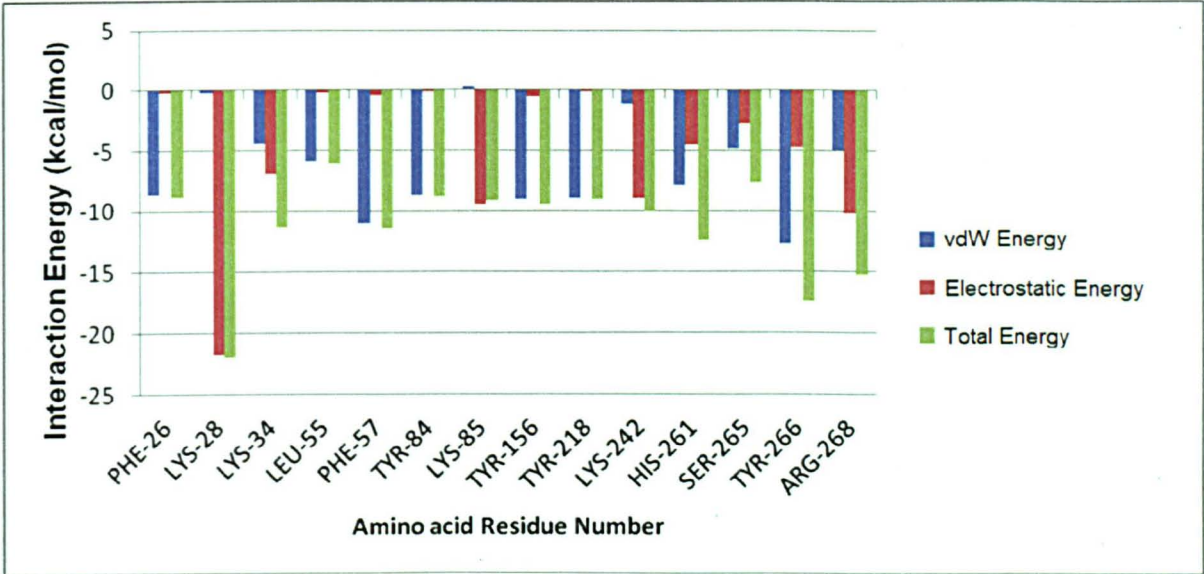


Figure A1. Interaction energies of top 14 POT1 amino acid residues for DNA.

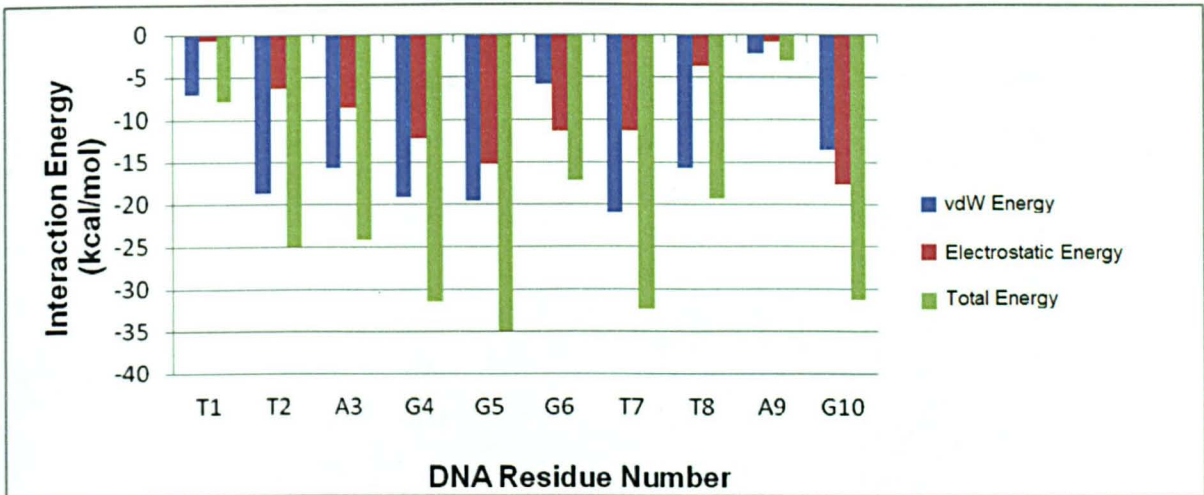
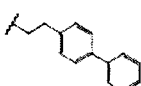
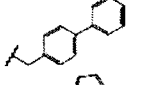
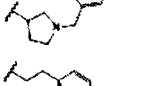
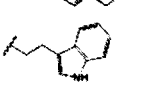
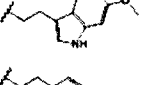
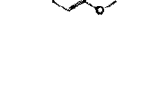
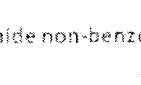


Figure A2: Interaction energies of DNA nucleotides for POT1.

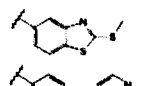
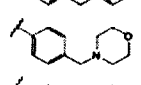
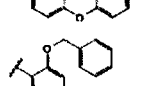
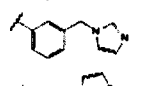
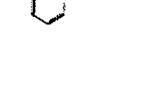

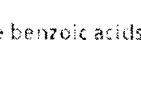
Appendix B

POT1 library chemical structures

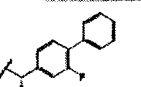
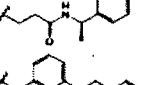
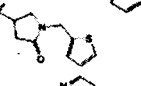
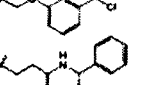
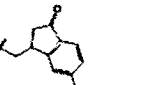
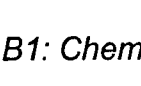

4, N-acyl reverse amide amines

| R | GoldScore |
|---|-----------|
|  | 64.43 |
|  | 64.32 |
|  | 63.70 |
|  | 62.79 |
|  | 62.64 |
|  | 62.22 |
|  | 60.52 |

4, N-acyl reverse amide anilines

| R | GoldScore |
|---|-----------|
|  | 68.85 |
|  | 68.42 |
|  | 68.11 |
|  | 66.36 |
|  | 66.23 |
|  | 65.88 |
|  | 65.83 |

4, N-acyl amide non-benzoic acids

| R | GoldScore |
|---|-----------|
|  | 71.80 |
|  | 69.44 |
|  | 68.44 |
|  | 66.52 |
|  | 66.36 |
|  | 66.10 |
|  | 66.02 |

4, N-acyl amide benzoic acids

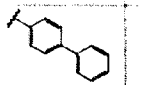
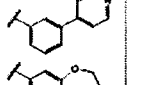
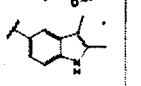
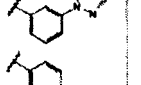
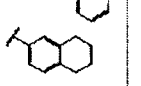
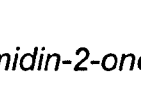



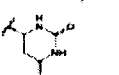
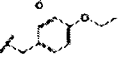
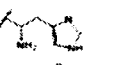
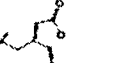
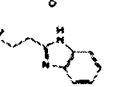







| R | GoldScore |
|---|-----------|
|  | 64.50 |
|  | 62.87 |
|  | 62.80 |
|  | 61.54 |
|  | 60.85 |
|  | 60.13 |
|  | 59.9 |

Figure B1: Chemical structures of pyrido[1,2-a]pyrimidin-2-one library.

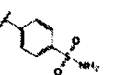
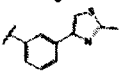

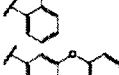
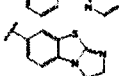
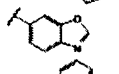
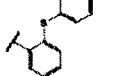
3, N-acyl amide non-benzoic acids

| R | GoldScore |
|---|-----------|
|  | 73.61 |
|  | 71.37 |
|  | 71.04 |
|  | 70.87 |
|  | 70.01 |
|  | 69.34 |
|  | 68.98 |

3, N-acyl amide benzoic acids

| R | GoldScore |
|---|-----------|
|  | 68.21 |
|  | 67.97 |
|  | 67.90 |
|  | 67.61 |
|  | 67.23 |
|  | 66.91 |
|  | 66.76 |

3, N-acyl reverse amide anilines

| R | GoldScore |
|---|-----------|
|  | 66.89 |
|  | 66.57 |
|  | 66.34 |
|  | 65.21 |
|  | 65.19 |
|  | 64.88 |
|  | 64.85 |

3, N-acyl reverse amide amines

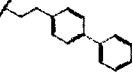
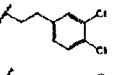
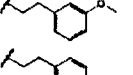
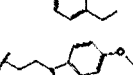
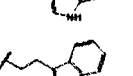
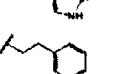

| R | GoldScore |
|---|-----------|
|  | 80.12 |
|  | 68.92 |
|  | 68.70 |
|  | 68.24 |
|  | 67.67 |
|  | 67.25 |
|  | 66.58 |

Figure B2: Chemical structures of pyrido[1,2-a]pyrimidin-2-one library.

Appendix C

ZINC structures

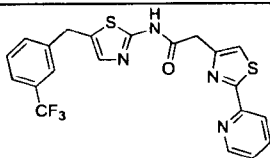
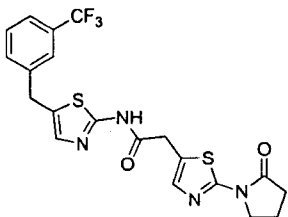
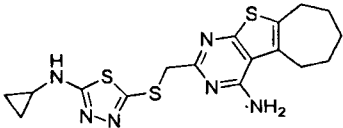
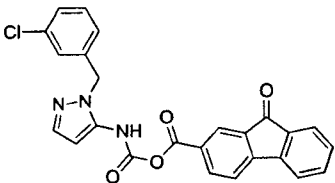
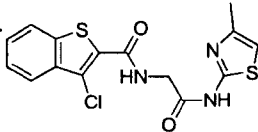
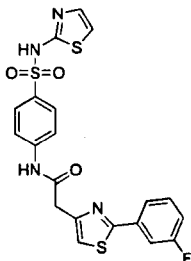
| Structure | Zinc ID | GoldScore |
|---|--------------|-----------|
|  | ZINC10972330 | 73.29 |
|  | ZINC10972460 | 70.74 |
|  | ZINC10536204 | 70.65 |
|  | ZINC10581619 | 70.29 |
|  | ZINC10555040 | 68.39 |
|  | ZINC09223825 | 68.24 |

Figure C1: Top ligands identified from virtual screening with corresponding ZINC identification numbers and GoldScores.

References

1. Abagyan, R., Totrov, M. and Kuznetsov, D. ICM - A new method for protein modeling and design: Applications to docking and structure prediction from the distorted native conformation. *J. Comput. Chem.* **1994**, 15, 488-506.
2. Abreu, E., Aritonovska, E., Reichenbach, P., Cristofari, G., Culp, B., Terns, R.M., Lingner, J. and Terns, M.P. TIN2-tethered TPP1 recruits human telomerase to telomeres *in vivo*. *Mol. Biol. Cell.* **2010**, 30, 2971-2982.
3. Alasdair, T., Laurie, R. and Jackson, R.M. Q-SiteFinder: An energy-based method for the prediction of protein–ligand binding sites. *Struct. Bioinformatics.* **2005**, 21, 1908-1916.
4. An, J., Totrov, M. and Abagyan, R. Pocketome via comprehensive identification and classification of ligand binding envelopes. *Mol. Cell. Proteomics.* **2005**, 4, 752-761.
5. Anantha, N.V., Azam, M. and Sheardy, R.D. Porphyrin binding to quadrupled T4G4. *Biochemistry.* **1998**, 37, 2709-2714.
6. Andrews, B.J. and Turchi, J.J. Development of a high throughput screen for inhibitors of replication protein A and its role in nucleotide excision repair. *Mol. Cancer. Ther.* **2004**, 3, 385-391.
7. Asai, A., Oshima, Y., Yamamoto, Y., Uochi, T-a., Kusaka, H., Akinaga, S., Yamashita, Y., Pongracz, K., Pruzan, R., Wunder, E., Piatyszek, M., Li, S., Chin, A.C., Harley, C.B. and Gryaznov, S. A novel telomerase template antagonist (GRN163) as a potential anticancer agent. *Cancer Res.* **2003**, 63, 3931-3939.
8. Avilion, A.A., Piatyszek, M.A., Gupta, J., Shay, J.W., Bacchetti, S. and Greider, C.W. Human telomerase RNA and telomerase activity in immortal cell lines and tumor tissues. *Cancer Res.* **1996**, 56, 645-650.
9. Baker, B.F. and Monia, B.P. Novel mechanisms for antisense-mediated regulation of gene expression. *Biochim. Biophys. Acta, Gene. Struct. Expr.* **1999**, 1489, 3-18.
10. Balasubramanian, S. and Neidle, S. G-quadruplex nucleic acids as therapeutic targets. *Curr. Opin. Chem. Biol.* **2009**, 13, 1-9.
11. Baumann, P. and Cech, T.R. Pot1, the putative telomere end-binding protein in fission yeast and humans. *Science.* **2001**, 292, 1171-1175.

12. Baumann, P., Podell, E. and Cech, T.R. Human pot1 (protection of telomeres) protein: Cyto-localization, gene structure, and alternative splicing. *Mol. Biol. Cell.* **2002**, 22, 8079-8087.
13. Benetos, A., Gardner, J.P., Zureik, M., Labat, C., Xiaobin, L., Adamopoulos, C., Temmar, M., Bean, K.E., Thomas, F. and Aviv, A. Short telomeres are associated with increased carotid atherosclerosis in hypertensive subjects. *Hypertension.* **2004**, 43, 182-185.
14. Bianchi, A., Stansel, R.M., Fairall, L., Griffith, J.D., Rhodes, D. and de Lange, T. TRF1 binds a bipartite telomeric site with extreme spatial flexibility. *EMBO J.* **1999**, 18, 5735-5744.
15. Bilsland, A.E., Cairney, C.J. and Keith, W.N. Targeting the telomere and shelterin complex for cancer therapy: Current views and future perspectives. *J. Cell. Mol. Med.* **2011**, 15, 179-186.
16. Bissantz, C., Folkers, G. and Rognan, D. Protein-based virtual screening of chemical databases. Evaluation of different docking/scoring combinations. *J. Med. Chem.* **2000**, 43, 4759-4767.
17. Bissantz, C., Kuhn, B. and Stahl, M. A medicinal chemist's guide to molecular interactions. *J. Med. Chem.* **2010**, 53, 5061-5084.
18. Blackburn, E.H. and Gall, J.G. A tandemly repeated sequence at the termini of the extrachromosomal ribosomal RNA genes in *Tetrahymena*. *J. Mol. Biol.* **1978**, 120, 33-53.
19. Blackburn, E.H. Telomerases. *Annu. Rev. Biochem.* **1992**, 61, 113-129.
20. Bodnar, A.G., Ouellette, M., Frolkis, M., Holt, S.E., Chiu, C.P., Morin, G.B., Harley, C.B., Shay, J.W., Lichtsteiner, S. and Wright, W.E. Extension of life-span by introduction of telomerase into normal human cells. *Science.* **1998**, 279, 349-352.
21. Bogan, A.A. and Thorn, K.S. Anatomy of hot spots in protein interfaces. *J. Mol. Biol.* **1998**, 280, 1-9.
22. Böhm, H.J. The computer program LUDI: A new method for the de novo design of enzyme inhibitors. *J. Comput. Aided Mol. Des.* **1992**, 6, 61-78.
23. Böhm, H.J. The development of a simple empirical scoring function to estimate the binding constant for a protein-ligand complex of known three-dimensional structure. *J. Comput. -Aided Mol. Des.* **1994**, 8, 243-256.
24. Brady, G.P. and Stouten, P.F.W. Fast prediction and visualization of protein binding pockets with PASS. *J. Comput.-Aided Mol. Des.* **2000**, 14, 383-401.

25. Broccoli, D., Smogorzewska, A., Chong, L. and de Lange, T. Human telomeres contain two distinct Myb-related proteins, TRF1 and TRF2. *Nat. Genet.* **1997**, 17, 231-235.
26. Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S. and Karplus, M. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **1983**, 4, 187-217.
27. Calado, R.T. Telomeres and marrow failure. *Hematology*. **2009**, 338-343.
28. Campisi, J., Kim, S.H., Lim, C.S. and Rubio, M. Cellular senescence, cancer and aging: the telomere connection. *Exp. Gerontol.* **2001**, 36, 1619-1637.
29. Case, D. A., Pearlman, D. A., Caldwell, J. W., Cheatham III, T. E., Wang, J., Ross, W. S., Simmerling, C. L., Darden, T. A., Merz, K. M., Stanton, R. V., Cheng, A. L., Vincent, J. J., Crowley, M., Tsui, V., Gohlke, H., Radmer, R. J., Duan, Y., Pitera, J., Massova, I., Seibel, G. L., Singh, U. C., Weiner, P. K. and Kollman, P. A. Assisted model building with energy refinement 7 (AMBER 7). **2002**. University of California, San Francisco, CA.
30. Celli, G.B. and de Lange, T. DNA processing is not required for ATM-mediated telomere damage response after TRF2 deletion. *Nat. Cell Biol.* **2005**, 7, 712-718.
31. Chan, L.L., Pineda, M., Heeres, J.T., Hergenrother, P.J. and Cunningham, B.T. A general method for discovering inhibitors of protein-DNA interactions using photonic crystal biosensors. *ACS. Chem. Biol.* **2008**, 3, 437-448.
32. Chang, W., Dynek, J.N. and Smith, S. TRF1 is degraded by ubiquitin-mediated proteolysis after release from telomeres. *Genes Dev.* **2003**, 17, 1328-1333.
33. Chen, L-Y., Liu, D. and Songyang, Z. Telomere maintenance through spatial control of telomeric proteins. *Mol. Biol. Cell.* **2007**, 27, 5898-5909.
34. Chen, Y., Yang, Y., van Overbeek, M., Donigian, J.R., Baciou, P., de Lange, T. and Lei, M. A shared docking motif in TRF1 and TRF2 used for differential recruitment of telomeric proteins. *Science*. **2008**, 319, 1092-1096.
35. Chen, Z., Koenenman, K.S. and Corey, D.R. Consequences of telomerase inhibition and combination treatments for the proliferation of cancer cells. *Cancer Res.* **2003**, 63, 5917-5925.
36. Churikov, D., Wei, C. and Price, C.M. Vertebrate POT1 restricts G-overhang length and prevents activation of a telomeric DNA damage checkpoint but is dispensable for overhang protection. *Mol. Cell. Biol.* **2006**, 26, 6971-6982.

37. Cian, A.D., Lacroix, L., Douarre, C., Temime-Smaali, N., Trentesaux, C., Riou, J-F. and Mergny, J-L. Targeting telomeres and telomerase. *Biochimie*. **2008**, 90, 131-155.
38. Cohen, S.B., Graham, M.E., Lovrecz, G.O., Bache, N., Robinson, P.J. and Reddel, R.R. Protein composition of catalytically active human telomerase from immortal cells. *Science*. **2007**, 315, 1850-1853.
39. Colgin, L.M., Baran, K., Baumann, P., Cech, T.R. and Reddel, R.R. Human POT1 facilitates telomere elongation by telomerase. *Curr. Biol*. **2003**, 13, 942-946.
40. Cong, Y-S., Wright, W.E. and Shay, J.W. Human telomerase and its regulation. *Microbiol. Mol. Biol. Rev.* **2002**, 66, 407-425.
41. Congreve, M. and Marshall, F. Themed section: Molecular pharmacology of G protein-coupled receptors. *Br. J. Pharmacol.* **2010**, 159, 986-996.
42. Cookson, J.C., Dai, F., Smith, V., Heald, R.A., Laughton, C.A., Stevens, M.F.G. and Burger, A.M. Pharmacodynamics of the G-quadruplex-stabilizing telomerase inhibitor 3,11-difluoro-6,8,13-trimethyl-8Hquino[4,3,2-kl]acridinium methosulfate (RHPS4) *in vitro*: Activity in human tumor cells correlates with telomere length and can be enhanced, or antagonized, with cytotoxic agents. *Mol. Pharmacol.* **2005**, 68, 1551-1558.
43. Corey, D.R. Telomerase inhibition, oligonucleotides, and clinical trials. *Oncogene*. **2002**, 21, 631-637.
44. Cornell, W.D., Cieplak, P., Bayly, C.I., Gould, I.R., Merz, K.M., Ferguson, D.M., Spellmeyer, D.C., Fox, T., Caldwell, J.W. and Kollman, P.A. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* **1995**, 117, 5179-5197.
45. Counter, C.M., Avilion, A.A., LeFeuvrel, C.E., Stewart, N.G., Greider, C.W., Harley, C.B. and Bacchetti, S. Telomere shortening associated with chromosome instability is arrested in immortal cells which express telomerase activity. *EMBO J.* **1992**, 11, 1921-1929.
46. Counter, C.M., Hahn, W.C., Wei, W., Caddle, S.D., Beijersbergen, R.L., Lansdorp, P.M., Sedivy, J.M. and Weinberg, R.A. Dissociation among *in vitro* telomerase activity, telomere maintenance, and cellular immortalization. *Proc. Natl. Acad. Sci. USA.* **1998**, 95, 14723-14728.
47. Court, R., Chapman, L., Fairall, L. and Rhodes, D. How the human telomeric proteins TRF1 and TRF2 recognize telomeric DNA: A view from high-resolution crystal structures. *EMBO Rep.* **2005**, 6, 39-45.
48. d'Adda di Fagagna, F., Reaper, P.M., Clay-Farrace, L., Fiegler, H., Carr, P., von Zglinicki, T., Saretzki, G., Carter, N.P. and Jackson, S.P. A DNA damage

- checkpoint response in telomere-initiated senescence. *Nature*. **2003**, 426, 194-198.
49. Dai, X., Huang, C., Bhusari, A., Sampathi, S., Schubert, K. and Chai, W. Molecular steps of G-overhang generation at human telomeres and its function in chromosome end protection. *EMBO J*. **2010**, 29, 2788-2801.
 50. Damm, K., Hemmann, U., Garin-Chesa, P., Huel, N., Kauffmann, I., Priepke, H., Niestroj, C., Daiber, C., Enenkel, B., Guilliard, B., Lauritsch, I., Muller, E., Pascolo, E., Sauter, G., Pantic, M., Martens, U.M., Wenz, C., Lingner, J., Kraut, N., Rettig, W.J. and Schnapp, A. A highly selective telomerase inhibitor limiting human cancer cell proliferation. *EMBO J*. **2001**, 20, 6958-6968.
 51. Datta, A., Bellon, M., Sinha-Datta, U., Bazarbachi, A., Lepelletier, Y., Canioni, D., Waldmann, T.A., Hermine, O. and Nicot, C. Persistent inhibition of telomerase reprograms adult T-cell leukemia to p53-dependent senescence. *Blood J*. **2006**, 108, 1021-1029.
 52. de Lange, T. Shelterin: The protein complex that shapes and safeguards human telomeres. *Genes Dev*. **2005**, 19, 2100-2110.
 53. de Lange, T. T-loops and the origin of telomeres. *Nat. Rev. Mol. Cell Biol*. **2004**, 5, 323-329.
 54. Denchi, E.L. and de Lange, T. Protection of telomeres through independent control of ATM and ATR by TRF2 and POT1. *Nature*. **2007**, 448, 1068-1071.
 55. Dorokhov, V.A., Baranin, S.V., Dib, A., Bogdanov, V.S., Yakovlev, I.P., Stashina, G.A. and Zhulin, V.M. Synthesis of N-(2-pyridyl)cyanoacetamides and 4-amino-2H-pyrido-[1,2-a]pyrimidin-2-ones from ethyl cyanoacetate and 2-aminopyridine. *Translated from Izvestiya Akademii Nauk SSSR*. **1990**, 9, 2107-2113.
 56. Duan, Y., Wu, C., Chowdhury, S., Lee, M.C., Xiong, G., Zhang, W., Yang, R., Cieplak, P., Luo, R., Lee, T., Caldwell, J., Wang, J. and Kollman, P. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem*. **2003**, 24, 1999-2012.
 57. Dunham, M.A., Neumann, A.A., Fasching, C.L. and Reddel, R.R. Telomere maintenance by recombination in human cells. *Nat. Genet*. **2000**, 26, 447-450.
 58. Durrant, J., Amaro, R.E. and McCammon, J.A. AutoGrow: A novel algorithm for protein inhibitor design. *Chem. Biol. Drug. Des*. **2009**, 73, 168-178.
 59. Dutta, S., Sutradhar, S. and Sachan, K. Computer-aided drug design - A new approach in drug design and discovery. *Int. J. Pharm Sci. Rev. Res*. **2010**, 4, 146-151.

60. Dyson, N., Howley, P.M., Münger, K. and Harlow, E. The human papilloma virus-16 E7 oncoprotein is able to bind to the retinoblastoma gene product. *Science*. **1989**, 243, 934-937.
61. Elayadi, A.N., Braasch, D.A. and Corey, D.R. Implications of high-affinity hybridization by locked nucleic acid oligomers for inhibition of human telomerase. *Biochemistry*. **2002**, 41, 9973-9981.
62. Eldridge, M.D., Murray, C.W., Auton, T.R., Paolini, G.V. and Mee, R.P. Empirical scoring functions: The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *J. Comput.-Aided Mol. Des.* **1997**, 11, 425-445.
63. Elmore, S. Apoptosis: A review of programmed cell death. *Toxicol. Pathol.* **2007**, 35, 495-516.
64. Ewing, T.J.A. and Kuntz, I.D. Critical evaluation of search algorithms for automated molecular docking and database screening. *J. Comput. Chem.* **1997**, 18, 1175-1189.
65. Falchetti, A., Franchi, A., Bordi, C., Mavilia, C., Masi, L., Cioppi, F., Recenti, R., Picariello, L., Marini, F., Monte, F.D., Ghinoi, V., Martinetti, V., Tanini, A. and Brandi, M.L. Azidothymidine induces apoptosis and inhibits cell growth and telomerase activity of human parathyroid cancer cells in culture. *J. Bone. Miner. Res.* **2005**, 20, 410-418.
66. Fedoroff, O.Y., Salazar, M., Han, H., Chemeris, V.V., Kerwin, S.M. and Hurley, L.H. NMR-based model of a telomerase-inhibiting compound bound to G-quadruplex DNA. *Biochemistry*. **1998**, 37, 12367-12374.
67. Feng, J., Funk, W.D., Wang, S.S., Weinrich, S.L., Avilion, A.A., Chiu, C.P., Adams, R.R., Chang, E., Allsopp, R.C. and Yu, J. The RNA component of human telomerase. *Science*. **1995**, 269, 1236-1241.
68. Ferlay, J., Shin, H.R., Bray, F., Forman, D., Mathers, C. and Parkin, D.M. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int. J. Cancer*. **2010**, 127, 2893-2917.
69. Fiset, S. and Chabot, B. hnRNP A1 may interact simultaneously with telomeric DNA and the human telomerase RNA *in vitro*. *Nucleic Acids Res.* **2001**, 29, 2268-2275.
70. Fletcher, T.M., Cathers, B.E., Ravikumar, K.S., Mamiya, B.M. and Kerwin, S.M. Inhibition of human telomerase by 7-deaza-2'-deoxyguanosine nucleoside triphosphate analogs: Potent inhibition by 6-thio-7-deaza-2'-deoxyguanosine 5'-triphosphate. *Bioorg. Chem.* **2001**, 29, 36-55.

71. Ford, L.P., Suh, J.M., Wright, W.E. and Shay, J.W. Heterogeneous nuclear ribonucleoproteins C1 and C2 associate with the RNA component of human telomerase. *Mol. Cell. Biol.* **2000**, 20, 9084-9091.
72. Freire, E. Do enthalpy and entropy distinguish first in class from best in class?. *Drug Discov. Today.* **2008**, 13, 869-874.
73. Friesner, R.A., Banks, J.L., Murphy, R.B., Halgren, T.A., Klicic, J.J., Mainz, D.T., Repasky, M.P., Knoll, E.H., Shelley, M., Perry, J.K., Shaw, D.E., Francis, P. and Shenkin, P.S. Glide: A new approach for rapid, accurate docking and scoring. Method and assessment of docking accuracy. *J. Med. Chem.* **2004**, 47, 1739-1749.
74. Fu, Y-T., Keppler, B.R., Soares, J. and Jarstfer, M.B. BRACO19 analog dimers with improved inhibition of telomerase and hPot1. *Bioorg. Med. Chem.* **2009**, 17, 2030-2037.
75. Gao, J., Zhang, J., Long, Y. and Lu, X. Expression of telomere binding proteins in gastric cancer and correlation with clinicopathological parameters. *Asia-Pacific. J. Clin. Oncol.* **2011**, 7, 339-345.
76. Garrett, C.E., Jiang, X., Prasad, K. and Repic, O. New observations on peptide bond formation using CDMT. *Tetrahedron Lett.* **2002**, 43, 4161-4165.
77. Gohlke, H. and Klebe, G. Statistical potentials and scoring functions applied to protein–ligand binding. *Curr. Opin. Struct. Biol.* **2001**, 11, 231-235.
78. Gohlke, H., Hendlich, M. and Klebe, G. Knowledge-based scoring function to predict protein-ligand Interactions. *J. Mol. Biol.* **2000**, 295, 337-356.
79. Gomez, D., O'Donohue, M-F., Wenner, T., Douarre, C., Macadrè, J., Koebel, P., Giraud-Panis, M-J., Kaplan, H., Kolkes, A., Shin-ya, K. and Riou, J-F. The G-quadruplex ligand telomestatin inhibits POT1 binding to telomeric sequences *in vitro* and induces GFP-POT1 dissociation from telomeres in human cells. *Cancer Res.* **2006**, 66, 6908-6912.
80. Goodford, P.J. A computational procedure for determining energetically favourable binding sites on biologically important macromolecules. *J. Med. Chem.* **1985**, 28, 849-857.
81. Gowan, S.M., Harrison, J.R., Patterson, L., Valenti, M., Read, M.A., Neidle, S. and Kelland, L.R. A G-quadruplex-interactive potent small-molecule inhibitor of telomerase exhibiting *in vitro* and *in vivo* antitumor activity. *Mol. Pharmacol.* **2002**, 61, 1154-1162.
82. Gowan, S.M., Heald, R., Stevens, M.F.G. and Kelland, L.R. Potent inhibition of telomerase by small-molecule pentacyclic acridines capable of interacting with G-quadruplexes. *Mol. Pharmacol.* **2001**, 60, 981-988.

83. Greider, C.W. and Blackburn, E.H. Identification of a specific telomere terminal transferase activity in tetrahymena extracts. *Cell*. **1985**, 43, 405-413.
84. Griffith, J.D., Comeau, L., Rosenfield, S., Stansel, R.M., Bianchi, A., Moss, H. and de Lange, T. Mammalian telomeres end in a large duplex loop. *Cell*. **1999**, 97, 503-514.
85. Gunaratnam, M., Greciano, O., Martins, C., Reszka, A.P., Schultes, C.M., Morjani, H., Riou, J-F. and Neidle, S. Mechanism of acridine-based telomerase inhibition and telomere shortening. *Biochem. Pharmacol.* **2007**, 74, 679-689.
86. Guo, X., Deng, Y., Lin, Y., Cosme-Blanco, W., Chan, S., He, H., Yuan, G., Brown, E.J. and Chang, S. Dysfunctional telomeres activate an ATM-ATR-dependent DNA damage response to suppress tumorigenesis. *EMBO J.* **2007**, 26, 4709-4719.
87. Hahn, W.C., Counter, C.M., Lundberg, A.S., Beijersbergen, R.L., Brooks, M.W. and Weinberg, R.A. Creation of human tumour cells with defined genetic elements. *Nature*. **1999**, 400, 464-468.
88. Han, H., Cliff, C.L. and Hurley, L.H. Accelerated assembly of G-quadruplex structures by a small molecule. *Biochemistry*. **1999**, 38, 6981-6986.
89. Hanahan, D. and Weinberg, R.A. Hallmarks of cancer: The next generation. *Cell*. **2011**, 144, 646-674.
90. Hanahan, D. and Weinberg, R.A. The Hallmarks of Cancer. *Cell*. **2000**, 100, 57-70.
91. Hara, E., Tsurui, H., Shinozaki, A., Nakada, S. and Oda, K. Cooperative effect of antisense-Rb and antisense-p53 oligomers on the extension of life span in human diploid fibroblasts, TIG-1. *Biochem. Biophys. Res. Commun.* **1991**, 179, 528-534.
92. Harle-Bachor, C. and Boukamp, P. Telomerase activity in the regenerative basal layer of the epidermis in human skin and in immortal and carcinoma-derived skin keratinocytes. *Proc. Natl. Acad. Sci. USA*. **1996**, 93, 6476-6481.
93. Harley, C.B., Futcher, A.B. and Greider, C.W. Telomeres shorten during ageing of human fibroblasts. *Nature*. **1990**, 345, 458-560.
94. Harrington, L., McPhail, T., Mar, V., Zhou, W., Oulton, R., Program, A., Bass, M.B., Arruda, I. and Robinson, M.O. A mammalian telomerase-associated protein. *Science*. **1997**, 275, 973-977.
95. Harris, R., Olson, A.J. and Goodsell, D.S. Automated prediction of ligand-binding sites in proteins. *Proteins*. **2008**, 70, 1506-1517.

96. Hawkins, P.C.D., Warren, G.L., Skillman, A.G. and Nicholls, A. How to do an evaluation: pitfalls and traps. *J. Comput-Aided. Mol. Des.* **2008**, 22, 179-190.
97. Hayflick, L. and Moorhead, P.S. The serial cultivation of human diploid cell strains. *Exp. Cell Res.* **1961**, 25, 585-621.
98. Hayflick, L. The limited *in vitro* lifetime of human diploid cell strains. *Exp. Cell Res.* **1965**, 37, 614-636.
99. He, H., Multani, A.S., Cosme-Blanco, W., Tahara, H., Ma, J., Pathak, S., Deng, Y. and Chang, S. POT1b protects telomeres from end-to-end chromosomal fusions and aberrant homologous recombination. *EMBO J.* **2006**, 25, 5180-5190.
100. Heald, R.A., Modi, C., Cookson, J.C., Hutchinson, I., Laughton, C.A., Gowan, S.M., Kelland, L.R. and Stevens, M.F.G. Antitumor polycyclic acridines. Synthesis and telomerase-inhibitory activity of methylated pentacyclic acridinium salts. *J. Med. Chem.* **2002**, 45, 590-597.
101. Hegde, S. and Schmidt, M. To market, to market 2006. *Annu. Rep. Med. Chem.* **2007**, 42, 505-554.
102. Hellman, L.M. and Fried, M.G. Electrophoretic mobility shift assay (EMSA) for detecting protein- nucleic acid Interactions. *Nat. Protoc.* **2007**, 2, 1849-1861.
103. Helwa, R. and Hoheisel, J.D. Analysis of DNA-protein interactions: from nitrocellulose filter binding assays to microarray studies. *Anal. Bioanal. Chem.* **2010**, 398, 2551-2561.
104. Hendlich, M., Rippmann, F. and Barnickel, G. LIGSITE: Automatic and efficient detection of potential small molecule binding sites in proteins. *J. Mol. Graphics Modell.* **1997**, 15, 359-363.
105. Hert, J., Irwin, J.J., Laggner, C., Keiser, M.J. and Shoichet, B.K. Quantifying biogenic bias in screening libraries. *Nat. Chem. Biol.* **2009**, 5, 479-483.
106. Hiyama, E. and Hiyama, K. Telomere and telomerase in stem cells. *Brit. J. Cancer.* **2007**, 96, 1020-1024.
107. Hiyama, E., Gollahon, L., Kataoka, T., Kuroi, K., Yokoyama, T., Gazdar, A.F., Hiyama, K., Piatyszek, M.A. and Shay, J.W. Telomerase activity in human breast tumors. *J. Natl. Cancer. Inst.* **1996**, 88, 116-122.
108. Hiyama, E., Hiyama, K., Tatsumoto, N., Kodama, T., Shay, J.W. and Yokoyama, T. Telomerase activity in human intestine. *Int. J. Oncol.* **1996**, 9, 453-458.
109. Hiyama, K., Hirai, Y., Kyoizumi, S., Akiyama, M., Hiyama, E., Piatyszek, M.A., Shay, J.W., Ishioka, S. and Yamakido, M. Activation of telomerase in human

- lymphocytes and hematopoietic progenitor cells. *J. Immunol.* **1995**, 155, 3711-3715.
110. Hockemeyer, D., Daniels, J-P., Takai, H. and de Lange, T. Recent expansion of the telomeric complex in rodents: Two distinct POT1 proteins protect mouse telomeres. *Cell.* **2006**, 126, 63-77.
 111. Hockemeyer, D., Palm, W., Else, T., Daniels, J-P., Takai, K.K., Ye, J.Z.S., Keegan, C.E., de Lange, T. and Hammer, G.D. Telomere protection by mammalian Pot1 requires interaction with Tpp1. *Nat. Struct. Mol. Biol.* **2007**, 14, 754-761.
 112. Holt, S.E., Aisner, D.L., Baur, J., Tesmer, V.M., Dy, M., Ouellette, M., Trager, J.B., Morin, G.B., Toft, D.O., Shay, J.W., Wright, W.E. and White, M.A. Functional requirement of p23 and Hsp90 in telomerase complexes. *Genes Dev.* **1999**, 13, 817-826.
 113. Horvath, M.P., Schweiker, V.L., Bevilacqua, J.M., Ruggles, J.A. and Schultz, S.C. Crystal structure of the *Oxytricha nova* telomere end binding protein complexed with single strand DNA. *Cell.* **1998**, 95, 963-974.
 114. Hsu, H-L., Gilley, D., Galande, S.A., Hande, M.P., Allen, B., Kim, S-H., Li, G.C., Campisi, J., Kohwi-Shigematsu, T. and Chen, D.J. Ku acts in a unique way at the mammalian telomere to prevent end joining. *Genes Dev.* **2000**, 14, 2807-2812.
 115. Huang, H-J., Yu, H.W., Chen, C-Y., Hsu, C-H., Chen, H-Y., Lee, K-J., Tsai, F-J. and Chen, C. Y-C. Current developments of computer-aided drug design. *J. Taiwan. Inst. Chem. Eng.* **2010**, 41, 623-635.
 116. Huang, S-Y. and Zou, X. Advances and challenges in protein-ligand docking. *Int. J. Mol. Sci.* **2010**, 11, 3016-3034.
 117. Humphrey, W., Dalke, A. and Schulten, K. VMD - Visual molecular dynamics. *J. Molec. Graphics.* **1996**, 14, 33-38.
 118. Huppert, J.L. and Balasubramanian, S. G-quadruplexes in promoters throughout the human genome. *Nucleic Acids Res.* **2007**, 35, 406-413.
 119. Incles, C.M., Schultes, C.M., Kempiski, H., Koehler, H., Kelland, L.R. and Neidle, S. A G-quadruplex telomere targeting agent produces p16-associated senescence and chromosomal fusions in human prostate cancer cells. *Mol. Cancer. Ther.* **2004**, 3, 1201-1206.
 120. Irwin, J.J., Sterling, T., Mysinger, M.M., Bolstad, E.S. and Coleman, R.G. ZINC: A free tool to discover chemistry for biology. *J. Chem. Inf. Model.* **2012**, 52, 1757-1768.

121. Izbicka, E., Wheelhouse, R.T., Raymond, E., Davidson, K.K., Lawrence, R.A., Sun, D., Windle, B.E., Hurley, L.H. and Von Hoff, D.D. Effects of cationic porphyrins as G-quadruplex interactive agents in human tumor cells. *Cancer Res.* **1999**, 59, 639-644.
122. Jain, A.S., Grand, C.L., Bearss, D.J. and Hurley, L.H. Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *PNAS.* **2002**, 99, 11593-11598.
123. Jenkins, J.L., Kao, R.Y.T. and Shapiro, R. Virtual screening to enrich hit lists from high-throughput screening: A case study on small-molecule inhibitors of angiogenin. *Proteins: Struct. Funct. Bioinf.* **2003**, 50, 81-93.
124. Jiang, H., Ju, Z. and Rudolph, K.L. Telomere shortening and cancer. *Z. Gerontol. Geriat.* **2007**, 40, 314-324.
125. Jones, G., Willet, P. and Glen, R.C. Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. *J. Mol. Biol.* **1995**, 245, 43-53.
126. Jones, G., Willett, P., Glen, R.C., Leach, A.R. and Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997**, 267, 727-748.
127. Jordan, J.L.M., Cross, G.A.M., de Lange, T. and Griffith, J.D. T-loops at trypanosome telomeres. *EMBO J.* **2001**, 20, 579-588.
128. Kalyaanamoorthy, S. and Chen, Y-P.P. Structure-based drug design to augment hit discovery. *Drug Discov. Today.* **2011**, 16, 831-839.
129. Karlseder, J., Broccoli, D., Dai, Y., Hardy, S. and de Lange, T. p53- and ATM-dependent apoptosis induced by telomeres lacking TRF2. *Science.* **1999**, 283, 1321-1325.
130. Karlseder, J., Hoke, K., Mirzoeva, O.K., Bakkenist, C., Kastan, M.B., Petrini, J.H.J. and de Lange, T. The telomeric protein TRF2 binds the ATM kinase and can inhibit the ATM-dependent DNA damage response. *PLoS Biol.* **2004**, 2, 1150-1156.
131. Kelleher C., Kurth, I. and Lingner, J. Human protection of telomeres 1 (POT1) is a negative regulator of telomerase activity in vitro. *Mol. Cell. Biol.* **2005**, 25, 808-818.
132. Kim, M-Y., Gleason-Guzman, M., Izbicka, E., Nishioka, D. and Hurley, L.H. The different biological effects of telomestatin and TMPyP4 can be attributed to their selectivity for interaction with intramolecular or intermolecular G-quadruplex structures. *Cancer Res.* **2003**, 63, 3247-3256.

133. Kim, N.W., Piatyszek, M.A., Prowse, K.R., Harley, C.B., West, M.D., Ho, P.L.C., Coviello, G.M., Wright, W.E., Weinrich, S.L. and Shay, J.W. Specific association of human telomerase activity with immortal cells and cancer. *Science*. **1994**, 266, 2011-2015.
134. Kim, S., Beausejour, C., Davalos, A.R., Kaminker, P., Heo, S. and Campisi, J. TIN2 mediates functions of TRF2 at human telomeres. *J. Biol. Chem.* **2004**, 279, 43799-43804.
135. Kim, S., Davalos, A.R., Heo, S., Rodier, F., Zou, Y., Beausejour, C., Kaminker, P., Yannone, S.M. and Campisi, J. Telomere dysfunction and cell survival: Roles for distinct TIN2-containing complexes. *J. Cell. Biol.* **2008**, 181, 447-460.
136. Kim, S., Kaminker, P. and Campisi, J. TIN2, a new regulator of telomere length in human cells. *Nat. Genet.* **1999**, 23, 405-412.
137. Kirtay, C.K., Mitchell, J.B.O. and Lumley, J.A. Scoring functions and enrichment: A case study on Hsp90. *BMC Bioinformatics*. **2007**, 8, 1-9.
138. Kitchen, D.B., Decornez, H., Furr, J.R. and Bajorath, J. Docking and scoring in virtual screening for drug discovery: Methods and applications. *Nat. Rev. Drug Discov.* **2004**, 3, 935-949.
139. Klingelhutz, A.J., Foster, S.A. and McDougall, J.K. Telomerase activation by the E6 gene product of human papillomavirus type 16. *Nature*. **1996**, 380, 79-82.
140. Klobuthcher, L.A., Swanton, M.T., Donini, P. and Prescott, D.M. All gene-sized DNA molecules in four species of hypotrichs have the same terminal sequence and an unusual 3' terminus. *Proc. Natl. Acad. Sci. USA*. **1981**, 78, 3015-3019.
141. Kondo, S., Tanaka, Y., Kondo, Y., Hitomi, M., Barnett, G.H., Ishizaka, Y., Liu, J., Haqqi, T., Nishiyama, A., Villeponteau, B., Cowell, J.K. and Barna, B.P. Antisense telomerase treatment: Induction of two distinct pathways, apoptosis and differentiation. *FASEB J.* **1998**, 12, 801-811.
142. Kondo, T., Oue, N., Yoshida, K., Mitani, Y., Naka, K., Nakayama, H. and Yasui, W. Expression of POT1 is associated with tumor stage and telomere length in gastric carcinoma. *Cancer Res.* **2004**, 64, 523-529.
143. Kondo, Y., Kondo, S., Tanaka, Y., Haqqi, T., Barna, B.P. and Cowell, J.K. Inhibition of telomerase increases the susceptibility of human malignant glioblastoma cells to cisplatin-induced apoptosis. *Oncogene*. **1998**, 16, 2243-2248.
144. Kong, D., Park, E.J., Stephen, A.G., Calvani, M., Cardellina, J.H., Monks, A., Fisher, R.J., Shoemaker, R.H. and Melillo, G. Echinomycin, a small-molecule

- inhibitor of hypoxia-inducible factor-1 DNA-binding activity. *Cancer Res.* **2005**, 65, 9047-9055.
145. Kontoyianni, M., McClellan, L.M. and Sokol, G.S. Evaluation of docking performance: Comparative data on docking algorithms. *J. Med. Chem.* **2004**, 47, 558-565.
 146. Kroemer, R.T. Structure-based drug design: Docking and scoring. *Curr. Protein Pept. Sci.* **2007**, 8, 312-328.
 147. Kuntz, I.D., Blaney, J.M., Oatley, S.J., Langridge, R. and Ferrin, T.E. A geometric approach to macromolecule-ligand interactions. *J. Mol. Biol.* **1982**, 161, 269-288.
 148. Kyo, S., Takakura, M., Kohama, T. and Inoue, M. Telomerase activity in human endometrium. *Cancer Res.* **1997**, 57, 610-614.
 149. Kyo, S., Takakura, M., Taira, T., Kanaya, T., Itoh, H., Yutsudo, M., Ariga, H. and Inoue, M. Sp1 cooperates with c-Myc to activate transcription of the human telomerase reverse transcriptase gene (hTERT). *Nucleic Acids Res.* **2000**, 28, 669-677.
 150. Laskowski, R.A. SURFNET: A program for visualizing molecular surfaces, cavities, and intermolecular interactions. *J. Mol. Graphics.* **1995**, 13, 323-330.
 151. Laskowski, R.A., Luscombe, N.M., Swindells, M.B. and Thornton, J.M. Protein clefts in molecular recognition and function. *Protein Sci.* **1996**, 5, 2438-2452.
 152. Leach, A.R. Molecular modelling: Principles and applications. **2001**. Pearson Education Ltd. Second Edition.
 153. Lei, M., Podell, E.R. and Cech, T.R. Structure of human POT1 bound to telomeric single-stranded DNA provides a model for chromosome end protection. *Nat. Struct. Mol. Biol.* **2004**, 11, 1223-1228.
 154. Lei, M., Zaug, A.J., Podell, E.R. and Cech, T.R. Switching human telomerase on and off with hPOT1 protein *in vitro*. *J. Biol. Chem.* **2005**, 280, 20449-20456.
 155. Lenain, C., Bauwens, S., Amiard, S., Brunori, M., Giraud-Panis, M-J. and Gilson, E. The Apollo 5' exonuclease functions together with TRF2 to protect telomeres from DNA repair. *Curr. Biol.* **2006**, 16, 1303-1310.
 156. Leonetti, C., Amodei, S., D'Angelo, C., Rizzo, A., Benassi, B., Antonelli, A., Elli, R., Stevens, M.F.G., D'Incalci, M., Zupi, G. and Biroccio, A. Biological activity of the G-quadruplex ligand RHPS4 (3,11-difluoro-6,8,13-trimethyl-8H-quino[4,3,2-kl]acridinium methosulfate) is associated with telomere capping alteration. *Mol. Pharmacol.* **2004**, 66, 1138-1146.

157. Levitt, D.G. and Banaszak, L.J. POCKET: A computer graphics method for identifying and displaying protein cavities and their surrounding amino acids. *J. Mol. Graph.* **1992**, 10, 229-234.
158. Levy, M.Z., Allsopp, R.C., Futcher, A.B., Greider, C.W. and Harley, C.B. Telomere end-replication problem and cell aging. *J. Mol. Biol.* **1991**, 225, 951-960.
159. Li, B. and de Lange, T. Rap1 affects the length and heterogeneity of human telomeres. *Mol. Biol. Cell.* **2003**, 14, 5060-5068.
160. Li, B., Oestreich, S. and de Lange, T. Identification of human Rap1: Implications for telomere evolution. *Cell.* **2000**, 101, 471-483.
161. Li, X., Li, Y., Cheng, T., Liu, Z. and Wang, R. Evaluation of the performance of four molecular docking programs on a diverse set of protein-ligand complexes. *J. Comput. Chem.* **2010**, 31, 2109-2125.
162. Liang, Y. Applications of isothermal titration calorimetry in protein science. *Acta. Biochim. Biophys. Sin.* **2008**, 40, 565-576.
163. Lipinski, C.A., Lombardo, F., Dominy, B.W. and Feeney, P.J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug. Del. Rev.* **2001**, 46, 3-26.
164. Liu, D., Safari, A., O'Connor, M.S., Chan, D.W., Laegeler, A., Qin, J. and Songyang, Z. PTOP interacts with POT1 and regulates its localization to telomeres. *Nat. Cell Biol.* **2004**, 6, 673-680.
165. Liu, M. and Wang, S. MCDOCK: A Monte Carlo simulation approach to the molecular docking problem. *J. Comput.-Aided. Mol. Des.* **1999**, 13, 435-451.
166. Lleónart, M.E., Artero-Castro, A. and Kondoh, H. Senescence introduction: A possible cancer therapy. *Mol. Cancer.* **2009**, 8, 1-10.
167. Loayza, D. and de Lange, T. POT1 as a terminal transducer of TRF1 telomere length control. *Nature.* **2003**, 423, 1013-1018.
168. Loayza, D., Parsons, H., Donigian, J., Hoke, K. and de Lange, T. DNA binding features of human POT1. *J. Biol. Chem.* **2004**, 279, 13241-13248.
169. Makarov, V.L., Hirose, Y. and Langmore, J.P. Long G tails at both ends of human chromosomes suggest a C strand degradation mechanism for telomere shortening. *Cell.* **1997**, 88, 657-666.
170. Martinez, P., Thanasoula, M., Muñoz, P., Liao, C., Tejera, A., McNees, C., Flores, J.M., Fernández-Capetillo, O., Tarsounas, M. and Blasco, M.A. Increased telomere fragility and fusions resulting from TRF1 deficiency lead to

- degenerative pathologies and increased cancer in mice. *Genes Dev.* **2009**, 23, 2060-2075.
171. Matthes, E. and Lehmann, C. Telomerase protein rather than its RNA is the target of phosphorothioate-modified oligonucleotides. *Nucleic Acids Res.* **1999**, 27, 1152-1158.
 172. McClintock, B. The stability of broken ends of chromosomes in *Zea Mays*. *Genetics.* **1941**, 26, 234-282.
 173. Meek, P.J., Liu, Z., Tian, L., Wang, C.Y., Welsh, W.J. and Zauhar, R.J. Shape signatures: Speeding up computer aided drug discovery. *Drug Discov. Today.* **2006**, 11, 895-904.
 174. Meyne, J., Ratliff, R.L. and Moyzis, R.K. Conservation of the human telomere sequence (TTAGGG)_n among vertebrates. *Proc. Natl. Acad. Sci. USA.* **1989**, 86, 7049-7053.
 175. Mills, N. ChemDraw Ultra 10.0. *J. Am. Chem. Soc.* **2006**, 128, 13649-13650.
 176. Mitchell, J.R., Cheng, J. and Collins, K. A box H/ACA small nucleolar RNA-like domain at the human telomerase RNA 3' end. *Mol. Cell. Biol.* **1999**, 19, 567-576.
 177. Miyaura, N. and Suzuki, A. Palladium-catalyzed cross-coupling reactions of organoboron compounds. *Chem. Rev.* **1995**, 95, 2457-2483.
 178. Moerke, N.J. Fluorescence polarization (FP) assays for monitoring peptide-protein or nucleic acid-protein binding. *Curr. Protoc. Chem. Biol.* **2009**, 1, 1-15.
 179. Moitessier, N., Englebienne, P., Lee, D., Lawandi, J. and Corbeil, C.R. Towards the development of universal, fast and highly accurate docking/scoring methods: A long way to go. *Br. J. Pharmacol.* **2008**, 153, 7-26.
 180. Mooij, W.T. and Verdonk M.L. General and targeted statistical potentials for protein-ligand interactions. *Proteins.* **2005**, 61, 272-287.
 181. Morris, G.M., Goodsell, D.S., Halliday, R.S., Huey, R., Hart, W.E., Belew, R.K. and Olson, A.J. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.* **1998**, 19, 1639-1662.
 182. Morrish, T.A. and Greider, C.W. Short telomeres initiate telomere recombination in primary and tumor cells. *PLoS Genet.* **2009**, 5, 1-15.
 183. Moyzis, R.K., Buckingham, J.M., Cram, L.S., Dani, M., Deaven, L.L., Jones, M.D., Meyne, J., Ratliff, R.L. and Wu, J.R. A highly conserved repetitive DNA

- sequence, (TTAGGG)_n, present at the telomeres of human chromosomes. *Proc. Natl. Acad. Sci. USA*. **1988**, 85, 6622-6626.
184. Mukai, S., Kondo, Y., Koga, S., Komata, T., Barna, B.P. and Kondo, S. 2-5A antisense telomerase RNA therapy for intracranial malignant gliomas. *Cancer Res*. **2000**, 60, 4461-4467.
 185. Muller, H.J. The remaking of chromosomes. *Collecting Net*. **1938**, 8, 182-198.
 186. Müller, K., Faeh, C. and Diederich, F. Fluorine in pharmaceuticals: Looking beyond intuition. *Science*. **2007**, 317, 1881-1886.
 187. Murti, K.G. and Prescott, D.M. Telomeres of polytene chromosomes in a ciliated protozoan terminate in duplex DNA loops. *PNAS*. **1999**, 96, 14436-14439.
 188. Naasani, I., Seimiya, H. and Tsuruo, T. Telomerase inhibition, telomere shortening, and senescence of cancer cells by tea catechins. *Biochem. Biophys. Res. Commun*. **1998**, 249, 391-396.
 189. Nehar, T.M., Shuck, S.C., Liu, J., Zhang, J-T. and Turchi, J.J. Identification of novel small molecule inhibitors of the XPA proteins using *in silico* based screening. *ACS. Chem. Biol*. **2010**, 5, 953-965.
 190. Neumann, A.A. and Reddel, R.R. Telomere maintenance and cancer – look, no telomerase. *Nat. Rev. Cancer*. **2002**, 2, 879-884.
 191. Nissink, J.W.M., Murray, C., Hartshorn, M., Verdonk, M.L., Cole, J.C. and Taylor, R. A new test set for validating predictions of protein–ligand interaction. *Proteins*. **2002**, 49, 457-471.
 192. Norton, J.C., Piatyszek, M.A., Wright, W.E., Shay, J.W. and Correy, D.R. Inhibition of human telomerase activity by peptide nucleic acids. *Nat. Biotechnol*. **1996**, 14, 615-618.
 193. Nugent, C.I., Hughes, T.R., Lue, N.F. and Lundblad, V. Cdc13p: A single-strand telomeric DNA-binding protein with a dual role in yeast telomere maintenance. *Science*. **1996**, 274, 249-252.
 194. O'Connor, M.S., Safari, A., Liu, D., Qin, J. and Songyang, Z. The human Rap1 protein complex and modulation of telomere length. *J. Biol. Chem*. **2003**, 279, 28585–28591.
 195. Olovnikov, A. M. Principle of marginotomy in template synthesis of polynucleotides, *Dokk. Akad. Nauk X.S.S.R*. **1971**, 201, 1496-1499.
 196. Opresko, P.L., von Kobbe, C., Laine, J-P., Harrigan, J., Hickson, I.D. and Bohr, V.A. Telomere-binding protein TRF2 binds to and stimulates the werner and bloom syndrome helicases. *J. Biol. Chem*. **2002**, 277, 41110-41119.

197. Ozer, H.L., Banga, S.S., Dasgupta, T., Houghton, J., Hubbard, K., Jha, K.K., Kim, S.H., Lenahan, M., Pang, Z., Pardinas, J.R. and Patsalis, P.C. SV40-mediated immortalization of human fibroblasts. *Exp. Gerontol.* **1996**, 31, 303-310.
198. Palm, W. and de Lange, T. How shelterin protects mammalian telomeres. *Annu. Rev. Genet.* **2008**, 42, 301-334.
199. Palm, W., Hockemeyer, D., Kibe, T. and de Lange, T. Functional dissection of human and mouse POT1 proteins. *Mol. Biol. Cell.* **2009**, 29, 471-482.
200. Panossian, L.A., Porter, V.R., Valenzuela, H.F., Zhu, X., Reback, E., Masterman, D., Cummings, J.L. and Effros, R.B. Telomere shortening in T cells correlates with Alzheimer's disease status. *Neurobiol. Aging.* **2003**, 24, 77-84.
201. Pascolo, E., Wenz, C., Lingner, J., Hael, N., Priepke, H., Kauffmann, I., Garin-Chesa, P., Rettig, W.J., Damm, K. and Schnapp, A. Mechanism of human telomerase inhibition by BIBR1532, a synthetic, non-nucleosidic drug candidate. *J. Biol. Chem.* **2002**, 277, 15566-15572.
202. Peng, Z. Current status of gendicine in China: Recombinant human Ad-p53 agent for treatment of cancers. *Hum. Gene Ther.* **2005**, 16, 1016-1027.
203. Perot, S., Sperandio, O., Miteva, M.A., Camproux, A-C. and Villoutreix, B.O. Druggable pockets and binding site centric chemical space: A paradigm shift in drug discovery. *Drug Discov. Today.* **2010**, 15, 656-667.
204. Perrin, F. Polarization de la lumière fluorescence. Vie moyenne des molécules dans l'état excité. *J. Phys. Radium.* **1926**, 7, 390-401.
205. Phan, A.T. Human telomeric G-quadruplex: structures of DNA and RNA sequences. *FEBS J.* **2010**, 277, 1107-1117.
206. Pitts, A.E. and Corey, D.R. Inhibition of human telomerase by 2'-O-methyl-RNA. *Proc. Natl. Acad. Sci. USA.* **1998**, 95, 11549-11554.
207. Qin, Y. and Hurley, L.H. Structures, folding patterns, and functions of intramolecular DNA G-quadruplexes found in eukaryotic promoter regions. *Biochimie.* **2008**, 90, 1149-1171.
208. Ramirez, R.D., Wright, W.E., Shay, J.W. and Taylor, R.S. Telomerase activity concentrates in the mitotically active segments of human hair follicles. *J. Invest. Dermatol.* **1997**, 108, 113-117.
209. Rarey, M., Kramer, B., Lengauer, T. and Klebe, G. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* **1996**, 261, 470-489.

210. Rastogi, T., Hildesheim, A. and Sinha, R. Opportunities for cancer epidemiology in developing countries. *Nature Rev.* **2004**, 4, 909-917.
211. Recio, J.F. Prediction of protein binding sites and hot spots. *WIREs Comput. Mol. Sci.* **2011**, 1, 680-698.
212. Ritzefeld, M. and Sewald, N. Real-time analysis of specific protein-DNA interactions with surface plasmon resonance. *J. Amino Acids.* **2011**, 2012, 1-19.
213. Rodriguez, R., Müller, S., Yeoman, J.A., Trentesaux, C., Riou, J-F. and Balasubramanian, S. A novel small molecule that alters shelterin integrity and triggers a DNA-damage response at telomeres. *J. Am. Chem. Soc.* **2008**, 130, 15758-15759.
214. Roe, D.C. and Kuntz, I.D. BUILDER v.2: Improving the chemistry of a de novo design strategy. *J. Comput. Aided Mol. Des.* **1995**, 9, 269-282.
215. Roninson, I.B. Tumour cell senescence in cancer treatment. *Cancer Res.* **2003**, 63, 2705-2715.
216. Rossetti, L., Franceschin, M., Bianco, A., Ortaggi, G. and Savino, M. Perylene diimides with different side chains are selective in inducing different G-quadruplex DNA structures and in inhibiting telomerase. *Bioorg. Med. Chem. Lett.* **2002**, 12, 2527-2533.
217. Rutenber, E.E. and Stroud, R.M. Binding of the anticancer drug ZD1694 to E. coli thymidylate synthase: Assessing specificity and affinity. *Structure.* **1996**, 4, 1317-1324.
218. Salvati, E., Leonetti, C., Rizzo, A., Scarsella, M., Mottolese, M., Galati, R., Sperduti, I., Stevens, M.F.G., D'Incalci, M., Blasco, M., Chiorino, G., Bauwens, S., Horard, B., Gilson, E., Stoppacciaro, A., Zupi, G. and Biroccio, A. Telomere damage induced by the G-quadruplex ligand RHPS4 has an antitumor effect. *J. Clin. Invest.* **2007**, 117, 3236-3246.
219. Sasaki, S., Ehara, T., Sakata, I., Fujino, Y., Harada, N., Kimura, J., Nakamura, H. and Maeda, M. Development of novel telomerase inhibitors based on a bisindole unit. *Bioorg. Med. Chem. Lett.* **2001**, 11, 583-585.
220. Scheffner, M., Werness, B.A., Huibregtse, J.M., Levine, A.J and Howley, P.M. The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53. *Cell.* **1990**, 63, 1129-1136.
221. Schwede, T., Kopp, J., Guex, N. and Peitsch, M.C. SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res.* **2003**, 31, 3381-3385.

222. Seimiya, H., Oh-hara, T., Suzuki, T., Naasani, I., Shimazaki, T., Tsuchiya, K. and Tsuruo, T. Telomere shortening and growth inhibition of human cancer cells by novel synthetic telomerase inhibitors MST-312, MST-295, and MST-199. *Mol. Cancer Ther.* **2002**, 1, 657-665.
223. Seimiya, H., Sawada, H., Muramatsu, Y., Shimizu, M., Ohko, K., Yamane, K. and Tsuruo, T. Involvement of 14-3-3 proteins in nuclear localization of telomerase. *EMBO J.* **2000**, 19, 2652-2661.
224. Shakirov, E.V., Song, X., Joseph, J.A. and Shippen, D.E. POT1 proteins in green algae and land plants: DNA-binding properties and evidence of co-evolution with telomeric DNA. *Nucleic Acids Res.* **2009**, 37, 7455-7467.
225. Shammas, M.A., Reis, R.J.S., Akiyama, M., Koley, H., Chauhan, D., Hideshima, T., Goyal, R.K., Hurley, L.H., Anderson, K.C. and Munshi, N.C. Telomerase inhibition and cell growth arrest by G-quadruplex interactive agent in multiple myeloma. *Mol. Cancer Ther.* **2003**, 2, 825-833.
226. Shammas, M.A., Reis, R.J.S., Li, C., Koley, H., Hurley, L.H., Anderson, K.C. and Munshi, N.C. Telomerase inhibition and cell growth arrest after telomestatin treatment in multiple myeloma. *Clin. Cancer Res.* **2004**, 10, 770-776.
227. Shao, J., Tanner, S.W., Thompson, N. and Cheatham, T.E. Clustering molecular dynamics trajectories: Characterizing the performance of different clustering algorithms. *J. Chem. Theory Comput.* **2007**, 3, 2312-2334.
228. Shay, J.W. and Bacchetti, S.A. Survey of telomerase activity in cancer cells. *Eur. J. Cancer.* **1997**, 33, 787-791.
229. Shay, J.W. and Roninson, I.B. Hallmarks of senescence in carcinogenesis and cancer therapy. *Oncogene.* **2004**, 23, 2919-2933.
230. Shay, J.W. and Wright, W.E. Telomerase: A target for cancer therapeutics. *Cancer Cell.* **2002**, 2, 257-265.
231. Shay, J.W. and Wright, W.E. Telomeres are double-strand DNA breaks hidden from DNA damage responses. *Mol. Cell.* **2004**, 14, 420-421.
232. Shelton, D.N., Chang, E., Whittier, P.S., Choi, D. and Funk, W.D. Microarray analysis of replicative senescence. *Curr. Biol.* **1999**, 9, 939-945.
233. Shiloh, Y. ATM and related protein kinases: safeguarding genome integrity. *Nat. Rev.* **2003**, 3, 155-168.
234. Shim, J. and MacKerell, A.D. Computational ligand-based rational design: Role of conformational sampling and force fields in model development. *Med. Chem. Commun.* **2011**, 2, 356-370.

235. Shin-ya, K., Wierzba, K., Matsuo, K., Ohtani, T., Yamada, Y., Furihata, K., Hayakawa, Y. and Seto, H. Telomestatin, a novel telomerase inhibitor from *Streptomyces anulatus*. *J. Am. Chem. Soc.* **2001**, 123, 1262-1263.
236. Simonsson, T. A substrate for telomerase. *Trends Biochem. Sci.* **2003**, 28, 632-638.
237. Smith, S. and de Lange, T. Tankyrase promotes telomere elongation in human cells. *Curr. Biol.* **2000**, 10, 1299-1302.
238. Smith, S., Gariat, I., Schmitt, A. and de Lange, T. Tankyrase, a poly(ADP-ribose) polymerase at human telomeres. *Science*. **1998**, 282, 1484-1487.
239. Smogorzewska, A., van Steensel, B., Bianchi, A., Oelmann, S., Schaefner, M.R., Schnapp, G. and de Lange, T. Control of human telomere length by TRF1 and TRF2. *Mol. Cell. Biol.* **2000**, 20, 1659-1668.
240. Strahl, C. and Blackburn, E.H. Effects of reverse transcriptase inhibitors on telomere length and telomerase activity in two immortalized human cell lines. *Mol. Cell. Biol.* **1996**, 16, 53-65.
241. Strahl, C. and Blackburn, E.H. The effects of nucleoside analogs on telomerase and telomeres in Tetrahymena. *Nucleic Acids Res.* **1994**, 22, 893-900.
242. Sumi, M., Tauchi, T., Sashida, G., Nakajima, A., Gotoh, A., Shin-Ya, K., Ohyashiki, J.H. and Ohyashiki, K. A G-quadruplex-interactive agent, telomestatin (SOT-095), induces telomere shortening with apoptosis and enhances chemosensitivity in acute myeloid leukemia. *Int. J. Oncol.* **2004**, 24, 1481-1487.
243. Sun, D., Thompson, B., Cathers, B.E., Salazar, M., Kerwin, S.M., Trent, J.O., Jenkins, T.C., Neidle, S. and Hurley, L.H. Inhibition of human telomerase by a G-quadruplex interactive compound. *J. Biol. Chem.* **1997**, 40, 2113-2116.
244. SYBYL 8.0, Tripos International, 1699 South Hanley Rd., St. Louis, Missouri, 63144, USA.
245. Tahara, H., Shin-ya, K., Seimiya, H., Yamada, Y., Tsuruo, T. and Ide, T. G-quadruplex stabilization by telomestatin induces TRF2 protein dissociation from telomeres and anaphase bridge formation accompanied by loss of the 3' telomeric overhang in cancer cells. *Oncogene*. **2006**, 25, 1955-1966.
246. Takai, H., Smogorzewska, A. and de Lange, T. DNA damage foci at dysfunctional telomeres. *Curr. Biol.* **2003**, 13, 1549-1556.

247. Takai, K.K., Kibe, T., Donigian, J.R., Frescas, D. and de Lange, T. Telomere protection by TPP1/POT1 requires tethering to TIN2. *Mol. Cell.* **2011**, 44, 647-659.
248. Tamimi, N.A.M. and Ellis, P. Drug development: From concept to marketing. *Nephron. Clin. Pract.* **2009**, 113, 125-131.
249. Tan, K.P., Varadarajan, R. and Madhusudhan, M.S. DEPTH: A web server to compute depth and predict small-molecule binding cavities in proteins. *Nucleic Acids Res.* **2011**, 39, 242-248.
250. Tang, Y., Zhu, W., Chen, K. and Jiang, H. New technologies in computer-aided drug design: Toward target identification and new chemical entity discovery. *Drug Discov. Today.* **2006**, 3, 307-313.
251. Tarsounas, M., Muñoz, P., Claas, A., Smiraldo, P.G., Pittman, D.L., Blasco, M.A. and West, S.C. Telomere maintenance requires the RAD51D recombination/repair protein. *Cell.* **2004**, 117, 337-347.
252. Tauchi, T., Shin-ya, K., Sashida, G., Sumi, M., Nakajima, A., Shimamoto, T., Ohyashiki, J.H and Ohyashiki, K. Activity of a novel G-quadruplex-interactive telomerase inhibitor, telomestatin (SOT-095), against human leukemia cells: involvement of ATM-dependent DNA damage response pathways. *Oncogene.* **2003**, 22, 5338-5347.
253. Tauchi, T., Shin-ya, K., Sashida, G., Sumi, M., Okabe, S., Ohyashiki, J.H. and Ohyashiki, K. Telomerase inhibition with a novel G-quadruplex-interactive agent, telomestatin: *in vitro* and *in vivo* studies in acute leukemia. *Oncogene.* **2006**, 25, 5719-5725.
254. Taylor, R.D., Jewsbury, P.J. and Essex, J.W. A review of protein-small molecule docking methods. *J. Comput.-Aided Mol. Des.* **2002**, 16, 151-166.
255. Teixeira, M.T., Arneric, M., Sperisen, P. and Lingner, J. Telomere length homeostasis is achieved via a switch between telomerase- extendible and - nonextendible states. *Cell.* **2004**, 117, 323-335.
256. The PyMOL molecular graphics system, version 1.5.0.4 Schrödinger, LLC.
257. Topliss, J.G. A manual method for applying the Hansch approach to drug design. *J. Med. Chem.* **1977**, 20, 463-469.
258. Tuncbag, N., Gursoy, A. and Keskin, O. Identification of computational hot spots in protein interfaces: Combining solvent accessibility and inter-residue potentials improves the accuracy. *Struct. Bioinformatics.* **2009**, 25, 1513-1520.
259. Tuncbag, N., Keskin, O. and Gursoy, A. HotPoint: Hot spot prediction server for protein interfaces. *Nucleic Acids Res.* **2010**, 38, 402-406.

260. Vallejo, J.A.L. Telomere instability and cancer. *Biochimie*. **2008**, 90, 73-82.
261. van Steensel, B. and de Lange, T. Control of telomere length by the human telomeric protein TRF1. *Nature*. **1997**, 385, 740-743.
262. van Steensel, B., Smogorzewska, A. and de Lange, T. TRF2 protects human telomeres from end-to-end fusions. *Cell*. **1998**, 92, 401-413.
263. Varghese, J.N. Development of neuraminidase inhibitors as anti-influenza virus drugs. *Drug Dev. Res.* **1999**, 46, 176-196.
264. Vaziri, H. and Benchimol, S. Reconstitution of telomerase activity in normal human cells leads to elongation of telomeres and extended replicative life span. *Curr. Biol.* **1998**, 8, 279-282.
265. Veldman, T., Etheridge, K.T. and Counter, C.M. Loss of hPot1 function leads to telomere instability and a cut-like phenotype. *Curr. Biol.* **2004**, 14, 2264-2270.
266. Verdonk, M.L., Berdini, V., Hartshorn, M.J., Mooij, W.T.M., Murray, C.W., Taylor, R.D. and Watson, P. Virtual screening using protein-ligand docking: Avoiding artificial enrichments. *J. Chem. Inf. Comput. Sci.* **2004**, 44, 793-806.
267. Verdonk, M.L., Cole, J.C., Hartshorn, M.J., Murray, C.W. and Taylor, R.D. Improved protein-ligand docking using GOLD. *Proteins*. **2003**, 52, 609-623.
268. Vriend, G. WHAT IF: A molecular modeling and drug design program. *J. Mol. Graph.* **1990**, 8, 52-56.
269. Vulliamy, T.J. and Dokal, I. Dyskeratosis congenita: The diverse clinical presentation of mutations in the telomerase complex. *Biochimie*. **2008**, 90, 122-130.
270. Wade, R.C., Clark, K.J. and Goodford, P.J. Further development of hydrogen bond functions for use in determining energetically favourable binding sites on molecules of known structure. 1. Ligand probe groups with the ability to form two hydrogen bonds. *J. Med. Chem.* **1993**, 36, 140-147.
271. Wang, F., Podell, E.R., Zaug, A.J., Yang, Y., Baciou, P., Cech, T.R. and Lei, M. The POT1-TPP1 telomere complex is a telomerase processivity factor. *Nature*. **2007**, 445, 506-510.
272. Wang, J., Xie, L.Y., Allan, S., Beach, D. and Hannon, G.J. Myc activates telomerase. *Genes Dev.* **1998**, 12, 1769-1774.
273. Wang, R., Lai, L. and Wang, S. Further development and validation of empirical scoring functions for structure-based binding affinity prediction. *J. Comput.-Aided Mol. Des.* **2002**, 16, 11-26.

274. Watson, J.D. Origin of concatameric T7 DNA. *Nature New Biol.* **1972**, 239, 197-201.
275. Weiner, P.K. and Kollman, P.A. AMBER: Assisted model building with energy refinement. A general program for modeling molecules and their interactions. *J. Comput. Chem.* **1981**, 2, 287-303.
276. Wheelhouse, R.T., Sun, D., Han, H., Han, F.X and Hurley, L.H. Cationic porphyrins as telomerase inhibitors: The Interaction of tetra-(N-methyl-4-pyridyl)porphine with quadruplex DNA. *J. Am. Chem. Soc.* **1998**, 120, 3261-3262.
277. Wiemann, S.U., Satyanarayana, A., Tsahuridu, M., Tillmann, H.L., Zender, L., Klemonauer, J., Flemming, P., Franco, S., Blasco, M.A., Manns, M.P. and Rudolph, K.L. Hepatocyte telomere shortening and senescence are general markers of human liver cirrhosis. *FASEB J.* **2002**, 16, 935-942.
278. Wlodawer, A. and Vondrasek, J. Inhibitors of HIV-1 protease: A major success of structure-assisted drug design. *Annu. Rev. Biophys. Biomol. Struct.* **1998**, 27, 249-284.
279. Wright, W.E. and Shay, J.W. The two-stage mechanism controlling cellular senescence and immortalization. *Exp. Gerontol.* **1992**, 27, 383-389.
280. Wright, W.E. Pereira-Smith, O.M. and Shay, J.W. Reversible cellular senescence: Implications for immortalization of normal human diploid fibroblasts. *Mol. Cell. Biol.* **1989**, 9, 3088-3092.
281. Wright, W.E., Piatyszek, M.A., Rainey, W.E., Byrd, W. and Shay, J.W. Telomerase activity in human germline and embryonic tissues and cells. *Dev. Genet.* **1996**, 18, 173-179.
282. Yang, Q., Zhang, R., Horikawa, I., Fujita, K., Afshar, Y., Kokko, A., Laiho, P., Aaltonen, L.A. and Harris, C.C. Functional diversity of human protection of telomeres 1 isoforms in telomere protection and cellular senescence. *Cancer Res.* **2007**, 67, 11677-11686.
283. Yang, Q., Zheng, Y-L. and Harris, C.C. POT1 and TRF2 cooperate to maintain telomeric integrity. *Mol. Cell. Biol.* **2005**, 25, 1070-1080.
284. Ye, J.Z.S. and de Lange, T. TIN2 is a tankyrase 1 PARP modulator in the TRF1 telomere length control complex. *Nat. Genet.* **2004**, 36, 618-623.
285. Ye, J.Z.S., Donigian, J.R., van Overbeek, M., Loayza, D., Luo, Y., Krutchinsky, A.N., Chait, B.T. and de Lange, T. TIN2 binds TRF1 and TRF2 simultaneously and stabilizes the TRF2 complex on telomeres. *J. Biol. Chem.* **2004**, 279, 47264-47271.

- 286. Ye, J.Z.S., Hockemeyer, D., Krutchinsky, A.N., Loayza, D., Hooper, S.M., Chait, B.T. and de Lange, T. POT1-interacting protein PIP1: A telomere length regulator that recruits POT1 to the TIN2/TRF1 complex. *Genes Dev.* **2004**, 18, 1649-1654.
- 287. Yi, X., Tesmer, V.M., Savre-Train, I., Shay, J.W. and Wright, W.E. Both transcriptional and posttranscriptional mechanisms regulate human telomerase template RNA levels. *Mol. Cell. Biol.* **1999**, 19, 3989-3997.
- 288. You, J.S. and Jones, P.A. Cancer genetics and epigenetics: Two sides of the same coin. *Cancer Cell.* **2012**, 22, 9-20.
- 289. Zahler, A.M., Williamson, J.R., Cech, T.R. and Prescott, D.M. Inhibition of telomerase by G-quartet DNA structures. *Nature.* **1991**, 350, 718-720.
- 290. Zhang, X., Mar, V., Zhou, W., Harrington, L. and Robinson, M.O. Telomere shortening and apoptosis in telomerase-inhibited human tumor cells. *Genes Dev.* **1999**, 13, 2388-2399.
- 291. Zhu, X-D., Küster, B., Mann, M., Petrini, J.H.J. and de Lange, T. Cell-cycle-regulated association of RAD50/MRE11/NBS1 with TRF2 and human telomeres. *Nat. Genet.* **2000**, 25, 347-352.
- 292. Zhu, X-D., Niedernhofer, L., Kuster, B., Mann, M., Hoeijmakers, J.H.J. and de Lange, T. ERCC1/XPF removes the 3' overhang from uncapped telomeres and represses formation of telomeric DNA-containing double minute chromosomes. *Mol. Cell.* **2003**, 12, 1489-1498.