TOWARDS THE DEVELOPMENT OF FLUORESCENT PROBES TARGETING ALDEHYDE DEHYDROGENASE (ALDH) IN CANCER

Expression and epigenetic modulation of ALDH1A1, ALDH2 and ALDH3A1 in selected in vitro models

Laura COSENTINO

Submitted for the degree of Doctor of Philosophy

Institute of Cancer Therapeutics
University of Bradford

2012
ABSTRACT

The cancer stem cell (CSC) concept is still very controversial; therefore identification and isolation of this specific population remain challenging. A variety of putative markers have been described and measurement of high aldehyde dehydrogenase (ALDH) activity has been defined as a characteristic of stem cells (SCs). In this study, a library of novel small molecules (1,4-disubstituted acetalanthraquinones, AAQs), containing an acetal group as protected aldehyde functionality, was designed with the aim of probing affinity for ALDH metabolism and demonstrating their potential as molecular fluorescent probes to identify CSCs. The AAQs were shown to be subjective to acidic hydrolysis using 2M HCl at 37°C; however compounds containing secondary or tertiary amine functionalities in their sidechain were only partly hydrolysed at 70 °C. Metabolism studies were conducted using cytosolic fractions from rat liver enriched in ALDHs, yeast ALDH and human recombinant ALDH1A1. Some evidence was demonstrated which linked ALDH metabolism with aldehyde functionalities of hydrolysed AAQs (HAAQs). The AAQs were shown to emit far-red fluorescence (600-750 nm). A close relationship between structure modifications and alteration of cellular localisation, with gained specificity for selected sub-cellular compartments were achieved when assessed in A549 and U-2 OS cell lines. Thermal DNA denaturation and chemosensitivity assays were used to obtain information about DNA binding properties and cytotoxicity of AAQs and HAAQ congeners. All compounds were shown to be weak-to-moderately binding to DNA, and symmetrical 1,4-disubstituted compounds were shown to be non-toxic (IC\textsubscript{50} = 100 µM) with non-symmetrical analogues generating IC\textsubscript{50} values in the 1-100 µM range. No fundamental variation in the biological activity was observed when comparing AAQs with HAAQs in the A549 (+ALDH) and MCF7 (-ALDH) cell lines. A pilot investigation revealed that aberrant gene methylation was cell-type dependent for three ALDH isoforms (1A1, 2, 3A1). Decitabine treatment led to enhanced protein expression for ALDH1A1 (A549), ALDH2 (MCF7) and ALDH3A1 (A549). In contrast, the protein level was reduced for ALDH1A1 in HT29 cells after decitabine treatment. ALDH1A1, ALDH2 and ALDH3A1 were highly expressed in prostate cell lines, with expression linked to promoter methylation. In contrast, low levels of DNA methylation were found in primary prostate cancer cells and benign prostatic hyperplasia. Interestingly, ALDH1A1, considered a SC marker, was found to be expressed at low levels in CD133\textsuperscript{+}/α\textsubscript{2}β\textsubscript{1}\textsuperscript{hi} stem cell fraction and upregulated in CD133\textsuperscript{−}/α\textsubscript{2}β\textsubscript{1}\textsuperscript{lo} differentiated prostate cancer cells. In summary, the results in this thesis demonstrate the complexity and tumour type specificity of ALDH expression. This creates challenges for the development of selective probes for CSC isolation, such as the AAQs discussed in this thesis. Although inconclusive results were obtained in regard to AAQs and their potential in targeting ALDHs, selected AAQs were shown to reveal interesting biological features highlighting them as potential non-invasive cytometric probes for tracking molecular interactions in live cells.
TABLE OF CONTENTS

ABSTRACT II

TABLE OF CONTENTS III

LIST OF FIGURES XII

LIST OF TABLES XVII

ACKNOWLEDGEMENTS XIX

ABBREVIATIONS XX

Chapter I - Introduction - 1 -

I.1 Cells of origin in cancer - 2 -
I.2 Identification of stem cell markers - 4 -
I.3 The cancer stem cells plasticity and therapeutic challenges - 7 -
I.4 Epigenetic regulation of gene expression in cancer - 10 -
  I.4.1 Introduction - 10 -
  I.4.2 DNA methylation: health and disease - 11 -
  I.4.3 Gene silencing and cancer - 13 -
  I.4.4 Reactivating silenced genes: the epigenetic therapy - 14 -
  I.4.5 Histone modifications - 18 -
  I.4.6 Dual therapy: DNA methyltransferase inhibitors and histone deacetylase inhibitors - 19 -
I.5 Aldehyde dehydrogenase family - 21 -
  I.5.1 Definition and implications - 21 -
  I.5.2 Aldehyde dehydrogenase activity - 23 -
  I.5.3 Aldehyde dehydrogenase as a functional biomarker - 26 -
  I.5.4 Heterogeneity of ALDH isoforms and biological implications - 28 -
  I.5.5 The use of ALDH activity to isolate stem cell populations - 29 -
  I.5.6 Aldefluor-based cell sorting - 30 -
  I.5.7 Aldefluor and ALDH selectivity - 33 -
I.6 Anthraquinone-based ALDH targeting reagents - 34 -
Chapter II - Design and synthesis of 1,4-di-substituted acetalanthraquinones

II.1 Introduction and Aims

II.1.1 Symmetrical and non-symmetrical 1,4-disubstituted anthraquinones

II.1.2 Rationale for the design of ALDH-targeting anthraquinone-based agents

II.1.3 Aims of the chapter

II.2 Results and Discussion

II.2.1 Synthesis of acetalamino-sidechains

II.2.1.1 Sidechain 1: 5, 5-dimethoxypentan-1-amine

II.2.1.2 Sidechain 2: 3-amino-N-(2,2-dimethoxyethyl)propanamide

II.2.1.3 Sidechain 3: N-(2-Aminoethyl)-3,3-dimethoxypropanamide

II.2.2 Synthesis of chromophores

II.2.3 Synthesis of and symmetrical and non-symmetrical 1,4-disubstituted-acetalanthraquinones

II.3 Experimental Details

II.3.1 Chemicals, Reagents and Instrumentations

II.3.2 Synthesis of 1,4-disubstituted-acetalanthraquinones

II.3.2.1 Synthesis of acetalamino sidechains

II.3.2.2 Synthesis of chromophores

II.3.2.3 Synthesis of symmetrical 1,4-disubstituted anthraquinones

II.3.2.4 Synthesis of non-symmetrical 1,4-disubstituted anthraquinones

Chapter III - Chemical hydrolysis and enzymatic metabolism of novel 1,4-di-substituted acetalanthraquinones (AAQs)
III.1 Introduction and Aims - 82 -

III.2 Materials and Methods - 84 -

III.2.1 Chemical hydrolysis of AAQs - 84 -
  III.2.1.1 Chemicals and Reagents - 84 -
  III.2.1.2 Chemical hydrolysis of AAQs - 84 -
  III.2.1.3 Acidic stability assessment via LC/MS - 85 -

III.2.2 Metabolism of AAQs - 86 -
  III.2.2.1 Chemicals and reagents - 86 -
  III.2.2.2 Isolation of cytosolic fraction from rat liver homogenate - 86 -
  III.2.2.3 Metabolism of AAQs in cytosolic fraction of rat liver homogenate - 87 -
  III.2.2.4 Metabolism of AAQs in yeast ALDH - 88 -
  III.2.2.5 Metabolism of AAQs in human recombinant ALDH1A1 - 89 -
  III.2.2.6 Evaluation of specific activity for enzyme metabolism - 89 -
  III.2.2.7 AAQs metabolite identification by LC/MS - 90 -

III.3 Results - 91 -

III.3.1 Chemical hydrolysis of AAQs - 91 -

III.3.2 Identification of chemical hydrolysis products via mass spectrometry - 96 -

III.3.3 Sensitivity of hydrolysis products to mobile phase composition - 97 -

III.3.4 Metabolism studies - 98 -
  III.3.4.1 Metabolism studies in cytosolic fraction of rat liver homogenate - 98 -
  III.3.4.2 Yeast ALDH metabolism - 103 -
  III.3.4.3 Metabolism of AAQs using human recombinant ALDH1A1 and metabolites identification via LC-MS - 107 -

III.4 Discussion - 111 -

Supplementary Information III.A-D (DVD enclosed)

Chapter IV - Evaluation of the DNA binding properties and cytotoxic effects of 1,4-di-substituted - 116 -
acetalanthraquinones (AAQs) and their corresponding hydrolysed products

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>IV.1 Introduction and Aims</td>
<td>117</td>
</tr>
<tr>
<td>IV.1.1 Anthraquinones as DNA intercalating agents</td>
<td>117</td>
</tr>
<tr>
<td>IV.1.2 Mitoxantrone, a 1,4-di-substituted anthraquinone with significant clinical activity</td>
<td>117</td>
</tr>
<tr>
<td>IV.1.3 DNA-binding affinity via DNA melting studies</td>
<td>119</td>
</tr>
<tr>
<td>IV.1.4 Cytotoxic effect of anthraquinones and resistance mechanism</td>
<td>120</td>
</tr>
<tr>
<td>IV.1.5 Aims of this study</td>
<td>121</td>
</tr>
<tr>
<td>IV.2 Materials and Methods</td>
<td>122</td>
</tr>
<tr>
<td>IV.2.1 AAQs stock solutions</td>
<td>122</td>
</tr>
<tr>
<td>IV.2.2 Melting curve assays for DNA binding analysis</td>
<td>122</td>
</tr>
<tr>
<td>IV.2.3 Mammalian cell lines culture</td>
<td>123</td>
</tr>
<tr>
<td>IV.2.3.1 Maintenance of mammalian cell lines</td>
<td>123</td>
</tr>
<tr>
<td>IV.2.3.2 Mammalian cell passaging</td>
<td>124</td>
</tr>
<tr>
<td>IV.2.3.3 Determination of live cell number</td>
<td>124</td>
</tr>
<tr>
<td>IV.2.3.4 Cryopreservation of mammalian cells</td>
<td>125</td>
</tr>
<tr>
<td>IV.2.3.5 Cellular proliferation assessment using MTT assay</td>
<td>125</td>
</tr>
<tr>
<td>IV.2.3.6 Statistical analysis</td>
<td>126</td>
</tr>
<tr>
<td>IV.3 Results</td>
<td>127</td>
</tr>
<tr>
<td>IV.3.1 DNA melting temperature as an indicator of DNA</td>
<td>127</td>
</tr>
<tr>
<td>IV.3.2 Effect of novel AAQs on cell viability</td>
<td>128</td>
</tr>
<tr>
<td>IV.4 Discussion</td>
<td>135</td>
</tr>
<tr>
<td>Supplementary Information IV.A-B (DVD enclosed)</td>
<td></td>
</tr>
</tbody>
</table>

Chapter V - Evaluation of fluorescent properties of novel 1,4-di-substituted acetalanthraquinones as potential ALDH-targeting probes

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>V.1 Introduction and Aims</td>
<td>140</td>
</tr>
<tr>
<td>V.1.1 The development of fluorescent probes to explore live cells</td>
<td>140</td>
</tr>
</tbody>
</table>
V.1.2 Anthraquinone-based fluorescent probes - seeking a functional readout - 144 -
V.1.3 Cell model systems for novel probe screening - 148 -
V.1.4 Aims of this study - 151 -
V.2 Materials and Methods - 154 -
V.2.1 Probe stock solutions - 154 -
V.2.2 Determination of fluorescence excitation/emission peaks - 154 -
V.2.3 Maintenance of mammalian cell lines - 154 -
V.2.4 Cell seeding - 155 -
V.2.5 Imaging - 155 -
  V.2.5.1 AAQs labelling of fixed cells - 157 -
  V.2.5.2 AAQs loading of live cells - 157 -
  V.2.5.3 Transmission imaging - 157 -
  V.2.5.4 Fluorescence imaging - 158 -
  V.2.5.5 Confocal laser scanning microscopy (CLSM) - 158 -
  V.2.5.6 Quantitation of confocal images using MetaMorph software - 159 -
  V.2.5.7 Time-lapse imaging - 159 -
V.2.6 Flow cytometry - 161 -
  V.2.6.1 Statistical analysis of functional capacity of selected AAQs - 162 -
  V.2.6.2 Aldefluor assay - 162 -
V.3 Results - 164 -
V.3.1 UV–Vis spectral studies - 164 -
V.3.2 Fluorescence spectral studies - 170 -
V.3.3 Screening of AAQs labelling in fixed U-2 OS cells - 175 -
V.3.4 Screening of AAQs loading in U-2 OS live cells - 177 -
V.3.5 Screening of AAQs loading in A549 live cells - 178 -
V.3.6 Confocal laser microscopy analysis in the U-2 OS cell line - 179 -
V.3.7 Uptake of AAQs in A549 cell model and ABCG2-dependent efflux - 186 -

V.3.8 Hydrolysis effect on cellular uptake of LC-AAQ derivatives - 189 -

V.3.9 ALDH Activity by Aldefluor flow cytometry assay in A549 cells - 192 -

V.4 Discussion - 194 -

Supplementary Information V.A - 202 -

Supplementary Information V.B - 204 -

Supplementary Information V.C - 212 -

**Chapter VI - Epigenetic regulation of ALDHs and ABC transporters** - 214 -

VI.1 Introduction and Aims - 215 -

VI.2 Materials and Methods - 219 -

VI.2.1 Mammalian cell lines culture and drug treatments - 219 -

VI.2.1.1 Maintenance of mammalian cell lines - 219 -

VI.2.1.2 Mammalian cell passaging - 219 -

VI.2.1.3 Determination of live cell number - 220 -

VI.2.1.4 Cryopreservation of mammalian cells - 220 -

VI.2.1.5 Drug stock solution - 220 -

VI.2.1.6 Treatment with decitabine - 221 -

VI.2.1.7 Combination treatment: decitabine and trichostatin A - 221 -

VI.2.2 Methylation analysis - 221 -

VI.2.2.1 Isolation of genomic DNA and quantification - 221 -

VI.2.2.2 Bisulfite conversion - 222 -

VI.2.2.3 Control samples - 224 -

VI.2.2.4 Calponin quantification - 224 -

VI.2.2.5 Polymerase chain reaction (PCR) - 225 -

VI.2.2.6 LINE1 analysis: PCR and pyrosequencing - 225 -

VI.2.2.7 Assay design for target genes - 230 -
VI.2.2.8 PCR conditions for designed assay
- 231 -
  VI.2.2.8.1 Primers
- 231 -
  VI.2.2.8.2 PCR amplification conditions
- 232 -
VI.2.2.9 Pyrosequencing
- 232 -
VI.2.2.10 Statistical analysis
- 233 -

VI.2.3 Gene expression analysis using quantitative RT-PCR (qRT-PCR) before and after epigenetic treatment
- 233 -
  VI.2.3.1 RNA extraction and quantification
- 233 -
  VI.2.3.2 Complementary DNA synthesis
- 234 -
  VI.2.3.3 Universal Probe Library technology
- 235 -
  VI.2.3.4 QRT-PCR primers design
- 235 -
  VI.2.3.5 QRT-PCR method
- 236 -
  VI.2.3.6 QRT-PCR efficiency validation
- 237 -
  VI.2.3.7 Data analysis
- 238 -
  VI.2.3.8 Statistical analysis
- 240 -

VI.2.4 Western blot analysis
- 241 -
  VI.2.4.1 Protein extraction
- 241 -
  VI.2.4.2 Determination of protein concentration
- 241 -
  VI.2.4.3 Polyacrylamide gel preparation
- 242 -
  VI.2.4.4 Protein transfer to PVDF membrane
- 242 -
  VI.2.4.5 Immunodetection of electrophoresed proteins after transfer to nitrocellulose membrane
- 243 -
  VI.2.4.6 Enhanced chemiluminescent detection
- 244 -
  VI.2.4.7 Stripping western blots for reprobing with loading control
- 245 -

VI.3 Results
- 246 -
  VI.3.1 Pyrosequencing analysis of the promoter region in LINE-1
- 246 -
  VI.3.2 Promoter methylation status of ALDH-ABC genes in selected cancer cell lines
- 249 -
  VI.3.3 Gene expression profiling of ALDH and ABC target genes
- 257 -
VI.3.4 Western blot analysis of total protein extracts from selected cancer cell lines

VI.4 Discussion

Supplementary Information VI.A

Supplementary Information VI.B

Supplementary Information VI.C-D (DVD enclosed)

Chapter VII - Epigenetic modulation of ALDH expression in prostate cancer

VII.1 Introduction and Aims

VII.1.1 Background

VII.1.2 Prostate cancer and ALDH

VII.1.3 ALDH investigation in human primary cells

VII.1.4 Rationale and aims of the investigation

VII.2 Materials and Methods

VII.2.1 Maintenance of mammalian prostate cell lines

VII.2.2 Treatment with decitabine

VII.2.3 Maintenance of primary cultures

VII.2.4 Irradiation of mouse embryonic fibroblasts

VII.2.5 Selection of subpopulations on primary prostatic cells

VII.2.6 Separation of $\alpha_2\beta_1$ integrin high/low ($\alpha_2\beta_1^{hi/lo}$) prostate cancer cells from primary cultures

VII.2.7 Isolation of CD133$^+$/\(\alpha_2\beta_1^{hi}\) prostate stem cells from primary samples

VII.2.8 Methylation analysis

VII.2.9 Gene expression analysis using quantitative RT-PCR (qRT-PCR) prior and after epigenetic treatment

VII.3 Results

VII.3.1 Expression profiling of ALDH genes prostate cell lines

VII.3.2 Effect of decitabine on ALDH promoter methylation pattern and gene expression in prostate cell lines
VII.3.3 Expression profiling of ALDH genes in primary epithelial cultures from prostatic tissues - 304 -
VII.3.4 Expression profiling of ALDH genes in the prostate epithelial hierarchy - 307 -
VII.4 Discussion - 312 -
Supplementary Information VII.A - 316 -
Supplementary Information VII.B - 317 -
Chapter VIII – Final Conclusions and Future Directions - 319 -
References - 326-
Appendix I – List of suppliers i
Appendix II - Composition of buffers, cell media and stock solutions iii
Appendix III - Chemical structures of final AAQs (Insert enclosed)
Appendix IV- Time-lapse video (DVD enclosed)
Appendix V – Abstracts to attended conferences vii
LIST OF FIGURES

Chapter I

Figure 1.1 Mutations in SCs and/or progenitor cells. - 3 -

Figure 1.2 ATP-dependent mechanism of drug transport. - 8 -

Figure 1.3 DNA methylation patterns. - 12 -

Figure 1.4 DNMT enzymatic reaction and mechanism of DNMT inhibition of DAC. - 15 -

Figure 1.5 Dose-dependent activity of decitabine. - 17 -

Figure 1.6 Interplay between RNA, histone modification and DNA methylation in heritable silencing. - 19 -

Figure 1.7 Ethanol metabolism by ADH and ALDH. - 24 -

Figure 1.8 ALDH proposed mechanism of aldehyde metabolism. - 26 -

Figure 1.9 Proof of principle of BODIPY-DA functionality. - 31 -

Figure 1.10 The basis of the Aldefluor reaction. - 33 -

Chapter II

Figure 2.1 Intercalation models for 1,4-bis[(diethylaminoethyl)amino]anthraquinone. - 41 -

Figure 2.2 1,4-di-substituted anthraquinone structure with possible modification sites. - 42 -

Figure 2.3 Chemical structures of exemplars of 1,4-di-substituted symmetrical and non-symmetrical anthraquinone-based compounds with high DNA affinity. - 43 -

Figure 2.4 1,4-di-substituted acetalanthraquinone LC- and HA-libraries of compounds. - 44 -

Figure 2.5 Structural modifications of the sidechains to insert 1,4-di-substituted acetalanthraquinones. - 46 -

Figure 2.6 Chemical structures of 1,4-difluoro-anthraquinones. - 52 -

Chapter III

Figure 3.1 Proposed mechanism for acetal functionality conversion. - 82 -

Figure 3.2 HPLC spectra of hydrolysis profile of LC-111. - 92 -

Figure 3.3 HPLC analysis of LC-111 short term hydrolysis. - 93 -
| Figure 3.4 | HPLC analysis of HA-218 before and after hydrolysis. |
| Figure 3.5 | MS spectra of LC-111 before and after exposure to acidic environment. |
| Figure 3.6 | Metabolism of propionaldehyde (10 µM) in rat liver cytosol. |
| Figure 3.7 | UV/Vis spectrum of NAD+ supplemented rat liver cytosol in presence of propionaldehyde. |
| Figure 3.8 | Kinetic analyses of several substrates (10-100 µM) in cytosolic fraction of rat liver homogenate. |
| Figure 3.9 | Cytosolic fraction of rat liver homogenate kinetic analysis of LC-111 before/after hydrolysis. |
| Figure 3.10 | Optimisation of yeast ALDH kinetic analysis. |
| Figure 3.11 | Optimisation of yeast ALDH kinetic analysis using propionaldehyde. |
| Figure 3.12 | Kinetic profiles of selected substrates using hrALDH1A1. |
| Figure 3.13 | HPLC spectrum of the metabolic profile of LC-111 following incubation with hrALDH1A1. |
| Figure 3.14 | MS analysis of LC-111 hydrolysed derivative after hrALDH1A1 metabolism. |
| Figure 3.15 | Overview of the structural modification on the AAQ sidechains. |
| Figure 3.16 | Proposed metabolism of hydrolysed LC-111 derivative. |

Chapter IV

| Figure 4.1 | Chemical structure of mitoxantrone. |
| Figure 4.2 | DNA denaturation model. |
| Figure 4.3 | Structure activity relationships with selected LC-AAQ derivatives. |
| Figure 4.4 | Modification of the carbon sidechain skeleton of LC-105, LC-112 and LC-108. |
| Figure 4.5 | Modifications of the secondary amine in the HA-AAQ library. |
| Figure 4.6 | Proposed DNA cleavage via Schiff’s base formation. |
| Figure 4.7 | ALDH-mediated conversion of aldehyde moiety into a carboxylic acid. |
Chapter V

Figure 5.1 DRAQ5™ fluorescent probe. - 146 -

Figure 5.2 DRAQ5 spectral analysis. - 147 -

Figure 5.3 Hoechst 33342 chemical structure and nuclear staining in A549 cells. - 148 -

Figure 5.4 Database survey of 59 human tumour NCI60 panel cell lines. - 150 -

Figure 5.5 A549 SP-detection. - 151 -

Figure 5.6 Typical time-lapse setup for 6 well plates analysis. - 160 -

Figure 5.7 Activation methodology for Aldefluor kit. - 163 -

Figure 5.8 Structural diversities in LC-AAQ probes. - 164 -

Figure 5.9 Absorbance spectra analysis of LC-103, LC-055 and LC-110 derivatives. - 165 -

Figure 5.10 Absorbance spectra analysis of LC-106, LC-112, LC-108 and LC-111. - 166 -

Figure 5.11 Absorbance spectra analysis of LC-111, LC-112, LC-113 and LC-114. - 167 -

Figure 5.12 Structural diversities in HA-AAQ probes. - 168 -

Figure 5.13 Absorbance spectra analysis of HA-AAQs. - 169 -

Figure 5.14 Fluorescence spectra of carbon chain substituted LC-AAQs. - 171 -

Figure 5.15 Fluorescence spectra of LC-AAQs containing an amide bond in the sidechain. - 172 -

Figure 5.16 Fluorescence spectra of non-symmetrical LC-AAQs containing amide and reverse amide bond in the sidechain. - 173 -

Figure 5.17 Fluorescence spectra of HA-AAQs containing a secondary amine on the sidechain. - 174 -

Figure 5.18 LC-055, LC-103 and LC-110 fluorescence analysis via direct confocal imaging. - 181 -

Figure 5.19 LC-112, LC-106, LC-111 and LC-108 fluorescence analysis via direct confocal imaging. - 182 -

Figure 5.20 LC-111, LC-112, LC-113 and LC-114 fluorescence analysis via direct confocal imaging. - 184 -
Figure 5.21 HA-AAQs fluorescence analysis via direct confocal imaging.

Figure 5.22 Flow cytometry setup for LC- and HA-AAQs analysis.

Figure 5.23 Uptake of selected AAQs on A549 cell line by flow cytometry analysis.

Figure 5.24 AAQ uptake profiles in the A549 cell line +/- FTC pre-treatment.

Figure 5.25 Parental/hydrolysed AAQ uptake profiles in the A549 cell line in media vs. Aldefluor buffer.

Figure 5.26 Histograms of Aldefluor uptake +/- DEAB.

Figure 5.27 Overview of the general principle for live cell imaging.

Chapter VI

Figure 6.1 Chemical reaction of bisulfite conversion of cytosine to uracil.

Figure 6.2 Schematic representation of the progress of the enzyme reaction in liquid-phase pyrosequencing.

Figure 6.3 PyroMark Q96 Vacuum Workstation.

Figure 6.4 Order of tips in the tip holder.

Figure 6.5 Exemplar of standard curves (PPIA, ALDH3A1 and ABCG2) derived from a dilution series of A549 cDNA using qRT-PCR (TaqMan Assays).

Figure 6.6 Exemplar of successful sodium bisulfite modifications.

Figure 6.7 Reproducibility of LINE-1 promoter methylation analysis using pyrosequencing.

Figure 6.8 ALDH1A1 promoter methylation analysis in HeLa cells.

Figure 6.9 ALDH2 promoter methylation analysis in MCF7 using pyrosequencing.

Figure 6.10 ALDH3A1 promoter methylation analysis using pyrosequencing.

Figure 6.11 ABCB1 downstream promoter methylation analysis using pyrosequencing.

Figure 6.12 A549 qRT-PCR plot and Ct values.

Figure 6.13 MCF7 qRT-PCR plot and Ct values.
Figure 6.14 HT29 qRT-PCR plot and Ct values. - 260 -
Figure 6.15 HeLa qRT-PCR plot and Ct values. - 261 -
Figure 6.16 Western blot analysis of A549 cell line. - 263 -
Figure 6.17 Western blot analysis of MCF7 cell line. - 264 -
Figure 6.18 Western blot analysis of HT29 and HeLa cell lines. - 266 -

Chapter VII

Figure 7.1 Schematic representation of the architecture of the human normal prostate epithelium. - 282 -
Figure 7.2 The hierarchical pathway of human prostate epithelium. - 283 -
Figure 7.3 Microarray analysis of selected ALDH isoform expression levels in PCa primary cells. - 286 -
Figure 7.4 Microarray analysis of selected ALDH isoform expression levels within ovarian cancer patients. - 288 -
Figure 7.5 ALDH1A1 and 2 expressions before/after chemotherapy. - 289 -
Figure 7.6 Selection markers and differentiation status of subpopulation obtained from basal cells selection. - 293 -
Figure 7.7 Relative ALDH gene expressions in PCa cell lines. - 297 -
Figure 7.8 ALDH1A1 promoter methylation and gene expression analysis in prostate cell lines before/after DAC treatment. - 300 -
Figure 7.9 ALDH3A1 promoter methylation and gene expression analysis in prostate cell lines before/after DAC treatment. - 302 -
Figure 7.10 qRT-PCR analysis of ALDHs gene expression in prostate primary epithelial cultures. - 304 -
Figure 7.11 ALDH1A1 and ALDH3A1 promoter methylation analysis in biopict sample from benign hyperplasia (BPH) and prostate cancer (PCa). - 306 -
Figure 7.12 ALDH1A1 gene expression analysis in basal epithelial cell after subpopulation selection. - 308 -
Figure 7.13 ALDH2 gene expression analysis in basal epithelial cell after subpopulation selection. - 309 -
Figure 7.14 ALDH3A1 gene expression analysis in basal epithelial cell after subpopulation selection. - 310 -
LIST OF TABLES

Chapter I

Table 1.1 Cell surface phenotype of CSCs identified in selected solid tumours. - 7 -

Table 1.2 Human ALDHs: tissue distribution and subcellular location. - 22 -

Chapter II

Table 2.1 Yields obtained from preparation of 1,4-di-substituted anthraquinones. - 58 -

Chapter III

Table 3.1 Kinetic analyses of several substrates using yeast ALDH as source of enzyme. - 106 -

Chapter IV

Table 4.1 Culture of mammalian cell lines. - 124 -

Table 4.2 CT-DNA/AAQs melting temperatures before and after acidic hydrolysis. - 127 -

Table 4.3 Growth inhibition (IC50) of LC-AAQ and HA-AAQ libraries against MCF7 and A549 cancer cell lines. - 129 -

Table 4.4 Growth inhibition (IC50) of hydrolysed AAQs (HAAQs) against MCF7 and A549 cancer cell lines. - 130 -

Chapter V

Table 5.1 Exemplar fluorescent labels commonly used for the functional analysis of living cells. - 143 -

Table 5.2 Culture of mammalian cell lines - 155 -

Table 5.3 CBS camera filter block set-up. - 158 -

Table 5.4 FACS Vantage fluorescence optics specifics. - 161 -

Table 5.5 K-S statistics to validate comparison between different histograms. - 162 -

Chapter VI

Table 6.1 Culture of mammalian cell lines. - 219 -

Table 6.2 Optimised PCR conditions for bisulfite conversion step. - 223 -

Table 6.3 Optimised PCR conditions for calponin amplification. - 225 -
Table 6.4 Optimised PCR conditions for LINE-1 amplification.

Table 6.5 Reagent volumes for pyrosequencing analysis.

Table 6.6 PCR conditions for amplification of target genes.

Table 6.7 Optimised annealing temperature for PCR amplification of target genes.

Table 6.8 Optimised conditions for cDNA synthesis.

Table 6.9 Primer sequences and assay details for qPCR analysis of target genes.

Table 6.10 PCR efficiency value calculated using the equation reported above.

Table 6.11 P values summary and significance.

Table 6.12 List of used primary antibodies and working dilutions.

Table 6.13 List of used secondary antibodies and working dilutions.

Table 6.14 ABCG2 downstream promoter methylation analysis in selected cell lines.

Table 6.15 Densitometry analysis of ABCB1 protein expression in MCF7 cell line.

Chapter VII

Table 7.1 Cell culture conditions of prostate cell lines.

Table 7.2 Ct values of ALDH gene expressions before/after DAC treatment in prostate cell lines.
ACKNOWLEDGEMENTS

I would like to express my gratitude to all those who gave me the possibility to complete this thesis.

Firstly and foremost, I would like to specially thank my supervisor Dr. Klaus Pors for his guidance and inspiring dedication in this research, for his full support, invaluable advice and continued encouragement throughout this work. I will miss our long hours spent talking of “science”, which definitely has brought out the “scientist in me”!

I wish to acknowledge and thank EPSRC and particularly Stefan Ogrodzinski and Biostatus for funding this project.
I like to thank Prof. Laurence Patterson and Dr. Roger Phillips for their irreplaceable help and support throughout the project.

I am grateful to Dr. Hamdy M. Abdel-Rahman for providing a library of compounds for my investigations and Dr. Helen Sheldrake for helping me with the organic chemistry.

I would like to thank Prof. Paul Smith and Dr. Rachel Errington, from Cardiff University, for their excellent expertise and for guiding me in the fluorescence imaging and flow cytometry field with remarkable patience and support. Thanks to Sally Chappell and Marie Wiltshire for their amazing help with probe evaluation.
I am grateful to Dr. Phil Burns, from Leeds Institute of Molecular Medicine, for allowing me to approach the gene expression analysis field and for his stimulating suggestion and encouragement in a very important time of my PhD. I must also say thanks to Dr. Julie Burns for her irreplaceable support on the huge amount of western blot analysis.
I am indebt to Prof. Bob Brown, from Imperial College in London, and his fantastic team for his incomparable expertise on the epigenetic field and for growing in me the interest for this fantastic discipline. A special thank you goes to Nahal and Naina for being such great people to work with and for the lovely time spent together.
I like to say thank you to Prof. Norman Maitland and Dr. Euan Polson, from University of York, for their brilliant expertise and for allow me to produce all the data in prostate cells. Thanks to Dr. Davide Pellacani for his help with pyrosequencing work. It has been a great and “useful” time in York!

From the bottom of my heart I would like to thank all the people at the ICT and “around the world” that made this PhD experience very special and unforgettable. Kelly, Rob, Laura, Antonia, Chris, Qasem, Oscar and Victoria…I will miss you very much! Thank you to my friend Linda for always giving me a place to feel at home. Now to my family and friends, who always supported me every single day of my life! Big thank you to my mum who have always encouraged me to fulfil personal ambitions which I didn't always think were possible. Finally thanks to Jay, who, no matter what, deserves “grazie”.
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>°C</td>
<td>Degrees Celsius</td>
</tr>
<tr>
<td>A</td>
<td>Adenosine</td>
</tr>
<tr>
<td>AAQ</td>
<td>1,4-Di-substituted acetalanthraquinone</td>
</tr>
<tr>
<td>ABC</td>
<td>ATP-binding cassette</td>
</tr>
<tr>
<td>Abs</td>
<td>Absorbance</td>
</tr>
<tr>
<td>ALDH</td>
<td>Aldehyde dehydrogenase</td>
</tr>
<tr>
<td>ANOVA</td>
<td>Analysis of Variance</td>
</tr>
<tr>
<td>APS</td>
<td>Adenosine 5'phosphosulphate</td>
</tr>
<tr>
<td>AQ</td>
<td>Anthraquinone</td>
</tr>
<tr>
<td>arb</td>
<td>Arbitrary</td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosine triphosphate</td>
</tr>
<tr>
<td>au</td>
<td>Arbitrary units</td>
</tr>
<tr>
<td>BODIPY</td>
<td>Boron-dipyrromethene</td>
</tr>
<tr>
<td>bp</td>
<td>Base pair</td>
</tr>
<tr>
<td>BPH</td>
<td>Benign prostatic hyperplasia</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine serum albumin</td>
</tr>
<tr>
<td>C</td>
<td>Cytosine</td>
</tr>
<tr>
<td>cDNA</td>
<td>Complementary DNA</td>
</tr>
<tr>
<td>CGI</td>
<td>CpG island</td>
</tr>
<tr>
<td>CIMP</td>
<td>CpG island methylator phenotype</td>
</tr>
<tr>
<td>Cntr</td>
<td>Control</td>
</tr>
<tr>
<td>CO₂</td>
<td>Carbon dioxide</td>
</tr>
<tr>
<td>CSC</td>
<td>Cancer stem cell</td>
</tr>
<tr>
<td>DAC</td>
<td>Decitabine, 5-aza-2'-deoxycytidine</td>
</tr>
<tr>
<td>DEAB</td>
<td>Diethylaminobenzaldehyde</td>
</tr>
<tr>
<td>dH₂O</td>
<td>Distilled water</td>
</tr>
<tr>
<td>DMEM</td>
<td>Dulbecco's Modified Eagle Medium</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethyl sulfoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>DNMT</td>
<td>DNA methyltransferase</td>
</tr>
<tr>
<td>dNTP</td>
<td>Deoxynucleotide triphosphate</td>
</tr>
<tr>
<td>DRAQ5</td>
<td>1,5-bis [2-(methylamino)ethyl]amino -4,8-dihydroxy anthracene-9,10-dione</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylene diaminetetraacetic acid</td>
</tr>
<tr>
<td>em</td>
<td>Emission</td>
</tr>
<tr>
<td>ER</td>
<td>Endoplasmic reticulum</td>
</tr>
<tr>
<td>ex</td>
<td>Excitation</td>
</tr>
<tr>
<td>FACS</td>
<td>Fluorescence-activated cell sorting</td>
</tr>
<tr>
<td>FBS</td>
<td>Foetal bovine serum</td>
</tr>
<tr>
<td>FCS</td>
<td>Foetal calf serum</td>
</tr>
<tr>
<td>FDA</td>
<td>Food and Drug Administration</td>
</tr>
<tr>
<td>FITC</td>
<td>Fluorescein isothiocyanate</td>
</tr>
<tr>
<td>FTC</td>
<td>Fumitremorgin C</td>
</tr>
<tr>
<td>G</td>
<td>Guanine</td>
</tr>
<tr>
<td>g</td>
<td>Gravity</td>
</tr>
<tr>
<td>g</td>
<td>gram</td>
</tr>
<tr>
<td>GAPDH</td>
<td>Glyceraldehyde-3-phosphate dehydrogenase</td>
</tr>
<tr>
<td>GFP</td>
<td>Green fluorescent protein</td>
</tr>
<tr>
<td>h</td>
<td>Hours</td>
</tr>
<tr>
<td>HAAQ</td>
<td>Hydrolysed AAQ</td>
</tr>
<tr>
<td>HCl</td>
<td>Hydrochloric acid</td>
</tr>
<tr>
<td>HDAC</td>
<td>Histone deacetylase</td>
</tr>
<tr>
<td>HEPES</td>
<td>4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid</td>
</tr>
<tr>
<td>HPLC</td>
<td>High performance liquid chromatography</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Hz</td>
<td>Hertz</td>
</tr>
<tr>
<td>IC&lt;sub&gt;50&lt;/sub&gt;</td>
<td>Half maximal inhibitory concentration</td>
</tr>
<tr>
<td>IVM</td>
<td>In vitro methylated DNA</td>
</tr>
<tr>
<td>k</td>
<td>kilo</td>
</tr>
<tr>
<td>kDa</td>
<td>kiloDalton</td>
</tr>
<tr>
<td>KSFM</td>
<td>Keratinocyte serum-free media</td>
</tr>
<tr>
<td>L</td>
<td>litre</td>
</tr>
<tr>
<td>LC/MS</td>
<td>Liquid chromatography/mass spectrometry</td>
</tr>
<tr>
<td>m</td>
<td>milli</td>
</tr>
<tr>
<td>m</td>
<td>meter</td>
</tr>
<tr>
<td>M</td>
<td>Molar concentration</td>
</tr>
<tr>
<td>MDR</td>
<td>Multidrug resistance</td>
</tr>
<tr>
<td>min</td>
<td>minutes</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger ribonucleic acid</td>
</tr>
<tr>
<td>MTT</td>
<td>3-(4,5- dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide</td>
</tr>
<tr>
<td>n</td>
<td>nano</td>
</tr>
<tr>
<td>NaCl</td>
<td>Sodium chloride</td>
</tr>
<tr>
<td>NAD&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Nicotinamide adenine dinucleotide</td>
</tr>
<tr>
<td>NADH</td>
<td>Nicotinamide adenine dinucleotide reduced</td>
</tr>
<tr>
<td>NCI</td>
<td>National cancer institute</td>
</tr>
<tr>
<td>p value</td>
<td>Probability value</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate buffered saline</td>
</tr>
<tr>
<td>PCa</td>
<td>Prostate cancer</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PPIA</td>
<td>Cyclophilin A</td>
</tr>
<tr>
<td>PVDF</td>
<td>Polyvinylidene fluoride</td>
</tr>
<tr>
<td>qRT-PCR</td>
<td>Quantitative reverse transcriptase PCR</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>RPLP0</td>
<td>60S acidic ribosomal protein P0</td>
</tr>
<tr>
<td>rpm</td>
<td>Rotation per minute</td>
</tr>
<tr>
<td>RPMI</td>
<td>Roswell Park Memorial Institute</td>
</tr>
<tr>
<td>RT</td>
<td>Room temperature</td>
</tr>
<tr>
<td>RT</td>
<td>Reverse transcription</td>
</tr>
<tr>
<td>SAM</td>
<td>S-adenosylmethionine</td>
</tr>
<tr>
<td>SAR</td>
<td>Structure activity relationship</td>
</tr>
<tr>
<td>SC</td>
<td>Stem cell</td>
</tr>
<tr>
<td>SCLC</td>
<td>Small cell lung cancer</td>
</tr>
<tr>
<td>SD</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodium dodecyl sulfate-polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>sec</td>
<td>seconds</td>
</tr>
<tr>
<td>SP</td>
<td>Side population</td>
</tr>
<tr>
<td>T</td>
<td>Thymine</td>
</tr>
<tr>
<td>TFA</td>
<td>Trifluoroacetic acid</td>
</tr>
<tr>
<td>Tris</td>
<td>2-amino-2-(hydroxymethyl)propane-1,3-diol</td>
</tr>
<tr>
<td>TSA</td>
<td>Trichostatin A</td>
</tr>
<tr>
<td>TSG</td>
<td>Tumour suppressor gene</td>
</tr>
<tr>
<td>U</td>
<td>Units</td>
</tr>
<tr>
<td>UTR</td>
<td>Untranslated region</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet</td>
</tr>
<tr>
<td>V</td>
<td>Volts</td>
</tr>
<tr>
<td>W</td>
<td>Watt</td>
</tr>
<tr>
<td>WT</td>
<td>Wild type</td>
</tr>
<tr>
<td>Y</td>
<td>Pyrimidine (thymine or cytosine)</td>
</tr>
<tr>
<td>µ</td>
<td>micro</td>
</tr>
</tbody>
</table>
Chapter I

Introduction
I.1 Cells of origin in cancer

The understanding of the normal cellular hierarchy within a given tissue is an important prerequisite to identifying the cellular origin of cancers. Organ development proceeds in a hierarchical manner from stem cells to committed progenitor cells, which in turn yield differentiated cells that constitute the bulk of the tissue or organ (Figure 1). The most primitive cells, stem cells (SCs), have been defined as cells that have the ability to perpetuate themselves through self-renewal and to generate mature cells of a particular tissue through differentiation. SCs have been favoured candidates for targets of transformation because of their capacity for self-renewal and their longevity, which would allow the sequential accumulation of genetic or epigenetic mutations required for carcinogenesis (Visvader, 2011). Any cell in the hierarchy with proliferative capacity could serve as a cell of origin in cancer, if it acquires mutations that reactivate self-renewal capacity and prevent differentiation to a post-mitotic state (Lin, 2002).

Over the last decade, a distinction has been made between SCs and cancer stem cells (CSCs). One of the proposed differences is their degree of dependence on the SC niche, a specialised microenvironment in which SCs reside. The SC niche in adult somatic tissues plays an essential role in maintaining a balance between proliferation-inhibiting and proliferation-promoting signals (Li and Neaves, 2006). CSCs may arise from an intrinsic mutation, leading to self-sufficient cell proliferation, and/or may also involve deregulation or alteration of the niche by dominant proliferation-promoting signals (Lin, 2002).
The theory of the CSC has generated as much excitement and optimism as any area of cancer research, as it provides opportunities for therapeutic targeting and ways to understand tumour progression and maintenance. Biologically, these cells are distinct from the other cells that form the bulk of a tumour in that they can self-perpetuate and produce progenitor cells in the same way of classical SCs (Matsui et al., 2004). Therefore, CSCs are thought to be responsible of tumourigenesis and possibly also play a key role in metastatic...
disease. However, while several studies (Bjerkvig et al., 2005) have demonstrated the ability of CSCs to initiate and maintain a primary tumour, the functional contribution of CSCs to metastatic behaviour remains poorly understood (Visvader and Lindeman, 2008). CSCs may generate tumours through the SC processes of self-renewal and differentiation into multiple cell types. Such cells are proposed to persist in tumours as a distinct population, which cause relapse and metastasis by giving rise to new tumours (Maenhaut et al., 2010). For example, CSCs may contribute to tumourigenic and chemoresistant sub-fractions in a variety of malignancies, including brain tumours, leukaemias, and breast carcinomas (Croker AK, 2009; Reya et al., 2001) leading to a search for targets for the specific elimination of CSCs (Maenhaut et al., 2010) or the mechanisms by which CSCs evade cytotoxic therapies (Visvader, 2011).

A better understanding of the CSC biology could aid in elucidating whether these cells can grow tumours on their own and what other properties they might have to contribute to clinical outcomes. Importantly for the focus of this thesis, these might include roles in drug resistance or metastasis.

I.2 Identification of stem cell markers

The quest to understand the functioning of normal and cancer SCs have led to an intensified search for a more practical way of defining these populations. “Stemness” markers or genes are badly sought after. The term “stemness” refers to all these common markers as well as the critical biological functions in
terms of the self-renewal potential and the ability to differentiate to various downstream mature progenitors (Bjerkvig et al., 2005).

Attempts to identify such common stemness markers have relied on taking molecular and biological markers of well-defined SC populations, such as embryonic SCs, hematopoietic SCs, and mesenchymal SCs and applying those markers to different tissues (Moreb et al., 2008). SC markers are given shorthand names based on the molecules that bind to the surface receptors. In many cases, a combination of multiple markers is used to identify a particular SC subtype. Therefore, researchers often identify SCs in shorthand by a combination of marker names reflecting the presence (\( + \)) or absence (\( - \)) of them (see Table 1.1).

In breast cancer, stem-like cells have, for example, been isolated from primary tumours based on a CD44\(^+\)CD24\(^-\) phenotype (Al-Hajj et al., 2003). Subsequent experimental studies have also isolated CD44\(^+\)CD24\(^-\) breast cancer cells and demonstrated increased \textit{in vitro} expression of SC markers and enhanced capacity for mammosphere formation, invasion, and resistance to radiation (Croker AK, 2009). Furthermore, clinical studies indicate that CD44\(^+\)CD24\(^-\) tumour-initiating cells (TIC) express an invasive gene signature and may be associated with distant metastases (Hennessy et al., 2009). Additional putative CSC phenotypes have also been identified in other solid cancers, including CD133\(^+\) (brain, colon, pancreatic cancer) (Lee et al., 2005) (Oshima et al., 2007) and CD44\(^+\)CD133\(^+\) (prostate cancer) (Richardson et al., 2004). However, because of the heterogeneous nature of solid tumours, the reliability of using cell surface markers as the sole way to isolate CSCs remains controversial.
(Clarke and Fuller, 2006). Furthermore, human SC sources (including tumours) may contain alternate stem/progenitor cell lineages not efficiently isolated using variably expressed cell surface markers (Holmes C, 2007).

The reason of this dispute is probably caused by the identification of CSCs from studying “normal systems” (Matsui et al., 2004). In brain tumours, for example, early research focused on the cell-surface protein CD133, which had been identified as a marker for normal neural SCs. In addition to brain cancer, CD133 has been used to identify potential CSCs in colon cancer, whereas CD44 has been used to identify breast cancer SCs. Aldehyde dehydrogenase (ALDH), in combination with other markers, has been used to identify CSCs in breast, prostate, and pancreatic cancer (Moreb, 2008) (see Table 1.1).

Fundamental for isolation of the tumour-specific CSC population is to further explore the expression of a distinctive repertoire of surface markers alone and in combination with selected functional enzymes. However, it should be noted that the markers proposed are different for the different type of cancer tissue and different from those of the somatic SCs (Maenhaut et al., 2010). Their potential biological role is little considered and it is often assumed that they mark the same cells.
<table>
<thead>
<tr>
<th>Tumour type</th>
<th>Cell surface marker(s)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breast</td>
<td>CD44&lt;sup&gt;+&lt;/sup&gt;CD24&lt;sup&gt;−/low&lt;/sup&gt;Lineage ESA&lt;sup&gt;+&lt;/sup&gt;</td>
<td>(Al-Hajj et al., 2003)</td>
</tr>
<tr>
<td>CNS</td>
<td>CD133&lt;sup&gt;+&lt;/sup&gt;</td>
<td>(Singh et al., 2003)</td>
</tr>
<tr>
<td>Colon</td>
<td>CD133&lt;sup&gt;+&lt;/sup&gt;</td>
<td>(O’Brien et al., 2007)</td>
</tr>
<tr>
<td>Head and neck</td>
<td>CD44&lt;sup&gt;+&lt;/sup&gt;Lineage&lt;sup&gt;+&lt;/sup&gt;</td>
<td>(Prince et al., 2007)</td>
</tr>
<tr>
<td>Melanoma</td>
<td>ABCB5&lt;sup&gt;+&lt;/sup&gt;CD271&lt;sup&gt;+&lt;/sup&gt;</td>
<td>(Boiko et al., 2010; Schatton et al., 2008)</td>
</tr>
<tr>
<td>Liver</td>
<td>CD90&lt;sup&gt;+&lt;/sup&gt;CD45&lt;sup&gt;-&lt;/sup&gt;(CD44&lt;sup&gt;-&lt;/sup&gt;)</td>
<td>(Yang et al., 2008)</td>
</tr>
<tr>
<td>Ovarian</td>
<td>CD44&lt;sup&gt;+&lt;/sup&gt;CD117&lt;sup&gt;+&lt;/sup&gt;</td>
<td>(Zhang et al., 2008)</td>
</tr>
<tr>
<td>Prostate</td>
<td>CD133&lt;sup&gt;+&lt;/sup&gt;/alpha 2 beta 1 integrin/ CD44&lt;sup&gt;+&lt;/sup&gt;</td>
<td>(Maitland and Collins, 2008)</td>
</tr>
<tr>
<td>Pancreas</td>
<td>CD44&lt;sup&gt;+&lt;/sup&gt;CD24&lt;sup&gt;+&lt;/sup&gt;ESA&lt;sup&gt;+&lt;/sup&gt;, CD133&lt;sup&gt;+&lt;/sup&gt;</td>
<td>(Li et al., 2007)</td>
</tr>
<tr>
<td>Lung</td>
<td>CD133&lt;sup&gt;+&lt;/sup&gt;</td>
<td>(Eramo et al., 2008)</td>
</tr>
</tbody>
</table>

Table 1.1 Cell surface phenotype of CSCs identified in selected solid tumours. CD: cluster of differentiation, ESA: epithelial-specific antigen. Table adapted from Visvader et al. (Visvader and Lindeman, 2008).

I.3 The cancer stem cells plasticity and therapeutic challenges

Multidrug resistance proteins, also known as ABC (ATP-binding cassette) transporters, are transmembrane proteins that utilise the energy of adenosine triphosphate (ATP) to transport a wide variety of substrates across extra- and intracellular membranes, including metabolic products, lipids and sterols, and drugs (Sparreboom et al., 2003). Enhanced drug efflux mediated by ABC transporters is one of several mechanisms of multi-drug resistance potentially linked with the chemoresponse failure of human cancers. Recently, a role of ABC transporters as phenotypic markers of SCs has been uncovered (Alison,
2003; Bunting, 2002). As an example, the novel ABCB5 drug transporter (Huang et al., 2004) is a chemoresistance mediator in human malignant melanoma (Schatton et al., 2008) and has been recognised as a CSC marker (Boiko et al., 2010). ABCB5+ tumour cells, detected in human melanoma patients, show a primitive molecular phenotype and in vivo genetic lineage tracking can demonstrate a specific capacity of ABCB5+ sub-populations for self-renewal and differentiation (Schatton et al., 2008).

**Figure 1.2 ATP-dependent mechanism of drug transport.** These membrane proteins function as pumps that extrude toxins and drugs out of the cell via ATP hydrolysis. TMD: transmembrane domain; ICD: intracellular domain; NBD: nucleotide binding domain. Figure adapted from Dong et al. (Dong et al., 2005).

Both breast cancer resistance protein (BCRP1) and ABCB1 (also known as P-glycoprotein or MDR1, multidrug resistance protein-1) are expressed in SCs, and enforced expression of either confers a typical phenotype in a variety of transduced cells (Zhou et al., 2002).
Some DNA interactive drugs are preferably subject to ATP-dependent efflux by BCRP1, a 655 aminoacids polypeptide designated as ABCG2 (MXR/BCRP/ABCP1) (Doyle and Ross, 2003). ABCG2 also shows an intriguing linkage with the CSC phenotype (Visvader and Lindeman, 2008), given that enhanced ABCG2 transporter expression contributes to the definition of pluripotential "side population" of cells (Goodell, 2002; Goodell et al., 2005). In many mammals, including humans, an adult SC subset termed the "side population" (SP) has been identified (Challen and Little, 2006). SP cells can rapidly exclude the DNA minor groove-binding dye Hoechst 33342 to produce a characteristic profile based on fluorescence-activated flow cytometric analysis (Hirschmann-Jax et al., 2005; Storms et al., 2000). Moreover, it is likely that the physiological function of ABCG2 expression in hematopoietic SCs is to provide protection from cytotoxic xenobiotics, such as mitoxantrone or topotecan (Zhou et al., 2002).

The ABCG2 half-transporter is sensitive to inhibitors (i.e. fumitremorgin C, FTC), which may lead to resistance reversal (Rabindran et al., 2000). However, ABCG2 is yet to have a fully defined role in drug resistance in human cancers due to complex substrate specificity patterns, co-expression of other ATP-binding cassette transporters (as ABCB1/MDR-1), effects of mutation on the spectrum of molecules transported and uncertainty over the impact of only moderate increases in transporter function (Doyle and Ross, 2003; Rabindran et al., 2000).
I.4 Epigenetic regulation of gene expression in cancer

I.4.1 Introduction

Genetics alone cannot explain human variation and disease. The increasingly popular term *epigenetics* embodies a partial explanation of both phenomena. It was first introduced by C.H. Waddington (Haldane and Waddington, 1931; Waddington, 1939) to describe “the causal interactions between genes and their product, which brings the phenotype into being”. More recently, epigenetics has been used to define heritable changes in gene expression which are not due to any alteration in DNA sequence (Esteller, 2009). While the genetic code in an individual is the same in every cell, the epigenetic code can be cell-specific and may change over time with aging, disease or environmental stimuli (e.g., nutrition, lifestyle, toxin exposure). Type of epigenetic modifications includes: methylation, acetylation, phosphorylation, ubiquitination, ADP-ribosylation.

On a molecular level, covalent modifications of cytosine bases and histones, and changes in the positioning of nucleosomes, are commonly regarded as driving epigenetic mechanisms. They are fundamental to the regulation of many cellular processes, including gene and microRNA expression, DNA-protein interactions, suppression of transposable element mobility, cellular differentiation, embryogenesis, X-chromosome inactivation and genomic imprinting (Portela and Esteller, 2010). To date, the cytosine methylation of DNA is the most widely studied epigenetic modification in humans (Jones and Takai, 2001).
An explosion of data indicating the importance of epigenetic processes, especially those resulting in the silencing of key regulatory genes, has led to the realisation that genetics and epigenetics cooperate at all stages of cancer development (Robertson and Wolffe, 2000). Recent advances include the understanding that (i) silencing is part of global epigenomic alterations in cancer, (ii) alteration of pathways are relevant to SC growth and differentiation and (iii) there is a need for the approval of drugs that target these defects in cancer patients (Jones and Baylin, 2007).

**I.4.2 DNA methylation: health and disease**

DNA methylation occurs almost exclusively in the context of CpG dinucleotides. The CpG dinucleotides tend to cluster in regions called CpG islands, where "p" simply indicates that C and G are connected by a phosphodiester bond. According with the definition of Gardiner-Garden (Gardiner-Garden and Frommer, 1987), CpG islands are regions of more than 200 bases with a G+C content of at least 50% and a ratio of observed to statistically expected CpG frequencies of at least 0.6. CpG dinucleotides are usually quite rare in mammalian genomes (~1%). About 60% of human gene promoters are associated with CpG islands and are usually unmethylated in normal cells, although some of them (~6%) become methylated in a tissue-specific manner during early development or in differentiated tissues (Esteller, 2009). DNA methylation plays a key role in genomic imprinting, where hypermethylation at one of the two parental alleles leads to monoallelic expression (Kacem and Feil,
The pattern of DNA methylation changes substantially when cells become cancerous (see Figure 1.3), as a result of two major events.

Firstly, cancer cells show genome-wide hypomethylation, which has been associated with chromosomal instabilities as well as activation of normally silenced repetitive DNA elements (Teodoridis et al., 2004). Secondarily, discrete portions at the promoter region of tumour-suppressor genes (TSGs) undergo intense hypermethylation (Esteller, 2009).

Figure 1.3 DNA methylation patterns. A representation of a region of DNA in non-cancerous (A) and cancerous (B) tissues shows the differences in DNA methylation in the two phenotypes. In non-cancerous tissue, genome wide hypermethylation of CpGs (green circles) and an actively transcribed tumour suppressor gene (TSG) is associated with a hypomethylated CGI (green lines). In cancerous tissue, the opposite is seen with genome wide hypomethylation (red lines) leading to genomic instability, and CGI hypermethylation (red circles) contributing to transcriptional silencing of a TSG.
Loss of DNA cytosine methylation (hypomethylation) results in genome instability, whereas focal hypermethylation of gene promoters causes heritable changes and therefore inactivation of TSGs. A combined effect of these two events contributes to cancer causation in humans, as alteration of homeostasis of epigenetic mechanisms (Baylin S. B., 2006).

I.4.3 Gene silencing and cancer

The initial reports of hypermethylation of the CpG islands in the promoter region of the retinoblastoma (Rb) TSG were followed by the findings that hypermethylation of the CpG island was a mechanism of inactivation of the tumour-suppressor genes VHL (associated with von Hippel-Lindau disease), the cell-cycle inhibitor p16^{INK4a} (Esteller and Herman, 2002), the mismatch DNA repair gene hMLH1 (Herman and Baylin, 2003) and BRCA1 (breast-cancer susceptibility gene) (Esteller, 2008). Therefore, the list of cancer-related genes affected by this disruption is growing steadily.

It has been estimated that in tumours there are on average 600 CpG islands aberrantly methylated compared to normal tissue (Costello et al., 2000), although this can vary widely between tumour types and within different histological fractions (Teodoridis et al., 2004). Moreover, methylation does not occur randomly, as there are CpG islands that are methylated in multiple tumour types, while other CpG islands are methylated only in specific cancer types (Costello et al., 2000). For example, BRCA1 hypermethylation is specific in breast and ovarian tumours but does not occur in other cancers (Strathdee
and Brown, 2002), whereas hMLH1 methylation-mediated silencing is typical of colorectal, gastric and endometrial cancer but it is unmethylated in other solid tumours (Davidson et al., 2001).

Silencing via hypermethylation of the CpG-island promoter can affect genes involved in the cell cycle, DNA repair, the metabolism of carcinogens, cell-to-cell interaction, apoptosis, and angiogenesis, all of which are involved in the development of cancer (Esteller, 2008; Teodoridis et al., 2004). Hypermethylation occurs at different stages in the development of cancer in different cellular networks, and it interacts with genetic lesions. In other words, the normal epigenetic modulation of these genes allows them to prevent stem or precursor cells from becoming immortalised and acquiring infinite cell renewal capacity during periods of chronic stresses and renewal pressures on cell systems. The inappropriate silencing of these genes blocks their activation and allows for abnormal survival and clonal expansion and prevents differentiation (Jones and Baylin, 2007).

I.4.4 Reactivating silenced genes: the epigenetic therapy

The heritable activation of cancer-related genes by altered DNA methylation and chromatin modification has led to the realisation that silenced chromatin may represent a viable therapeutic target. Reversal of aberrant DNA methylation, using agent with hypomethylating activity, results not only in reactivation of epigenetically silenced genes including TSGs, but also in anticancer activity.
Several small molecule inhibitors of DNA methylation that are derivatives of 2'-deoxycytidine are known (Goffin and Eisenhauer, 2002), e.g. 5-aza-2'-deoxycytidine (DAC, *decitabine*, Dacogen) or 5-azacytidine (5-azaC, *Vidaza*). DAC was approved in 2006 by the American Food and Drug Administration (FDA) for the treatment of myelodysplasic syndromes (Baylin S. B., 2006; Esteller, 2008; Jones and Baylin, 2007), as anticancer activity was demonstrated in several clinical trials in hematopoietic malignancies (Esteller, 2009; Jones and Taylor, 1980; Yang et al., 2006).

These small molecules, known as cytosine analogues, are incorporated into DNA in place of the natural nucleoside during DNA replication (see Figure 1.4).

---

**Figure 1.4** DNMT enzymatic reaction and mechanism of DNMT inhibition of DAC. This consists of the covalent addition of a methyl group from the methyl donor *S*-adenosyl methionine (SAM) catalysed by DNMTs. Adapted from (Gowher and Jeltsch, 2004).
Once incorporated into the DNA, these molecules trap the DNA methyltransferases (DNMTs) and target them for degradation (Baylin S. B., 2006), thus inhibiting the restoration of the original methylation pattern in daughter cells. However, the pathway of DNA demethylation by DAC turned out to be more complicated as demonstrated by the findings that the drug has a dual mechanism of action: an immediate cytotoxic effect is followed by induction of hypomethylation of the DNA (see Figure 1.5). The relative strength of both types of responses depends on the dose of the drug (Santi et al., 1983). Covalent attachment of the various DNMTs to DNA might well be responsible for the cytotoxicity of these agents, particularly at high doses (Michalowsky and Jones, 1987).
Figure 1.5 Dose-dependent activity of decitabine. Incorporation of high doses and low doses of decitabine lead to different cellular response due to activation of different biological pathways (Oki et al., 2007).
I.4.5 Histone modifications

Histones are key players in epigenetics. The core histones H2A, H2B, H3 and H4 group into two H2.A-H2.B dimers and one H3-H4 tetramer to form the nucleosome (De Ruijter et al., 2003). During activation of gene transcription, this compact, inaccessible DNA is made available to DNA binding proteins via modification of the nucleosome (Esteller, 2008). This architecture of chromatin is strongly influenced by post-translational modifications of the histones. To date, the acetylation of core histones is probably the best understood type of modification (Egger et al., 2004). Histone acetylation relaxes the normally tight supercoiling of chromatin, enhancing accessibility of DNA-binding transcriptional regulatory proteins to promoter regions. Conversely, HDAC3 maintains chromatin in a transcriptionally silent state. Loss of acetylation is mediated by HDACs, which have been found to be overexpressed or mutated in different tumour types (Zhu et al., 2004). Genes that can cause cell differentiation are normally downregulated by HDAC activity (Alcarraz-Vizan et al., 2009). HDAC inhibitors can induce differentiation, growth arrest and/or apoptosis in transformed cells in culture and in tumours. The driving hypothesis is that accumulation of acetylated proteins, particularly histones, results in the induction of genes and the upregulation of others that have become epigenetically silenced. In particular, the gene encoding p21 (Esteller, 2007), which is a cell-cycle kinase inhibitor, is commonly upregulated in tumour cells treated with these agents in the absence of p53 (Egger et al., 2004).
I.4.6 Dual therapy: DNA methyltransferase inhibitors and histone deacetylase inhibitors

The links between histone modification and DNA methylation have encouraged investigators to think about dual therapies combining DNA methyltransferase (DNMT) inhibitors with histone deacetylase (HDAC) inhibitors (Egger et al., 2004).

![Interplay between RNA, histone modification and DNA methylation in heritable silencing](image)

**Figure 1.6 Interplay between RNA, histone modification and DNA methylation in heritable silencing.** Histone modification can attract DNA methyltransferases to initiate cytosine methylation, which in turn can reinforce histone modification patterns conducive to silencing (Egger et al., 2004).

In 1999, Cameron *et al.* (Cameron et al., 1999) explored in detail the synergy of DNA demethylating agents and HDAC inhibitors. Indeed, the simultaneous inhibition of both processes would be the most efficacious approach to reactivating key genes for therapeutic purposes (Baylin S. B., 2006).
synergy between these two families of compounds (see Figure 1.6) might allow the reduction of individual doses, which could minimise the toxic effect and optimise the gene reactivating response (Esteller, 2009). Some clinical studies concerning the combination of demethylating agents and HDAC inhibitors in the treatment of patients with haematological cancers led to complete or partial response (Gowher and Jeltsch, 2004).

Currently there is considerable interest in the potential of activating silenced genes with epigenetic drugs which can sensitise cells towards the treatment with traditional chemotherapy (Egger et al., 2004). However, reactivation of genes may also results in activation of drug metabolising enzymes and other chemoprotecting protein such as ABC transporter family. This could, for example, impacts choice of combination treatment for clinical purposes.

I.5 Aldehyde dehydrogenase family

I.5.1 Definition and implications

Aldehydes are generated from a wide variety of endogenous and exogenous precursors during numerous physiological processes, including the biotransformation of endogenous compounds such as amino acids, neurotransmitters, carbohydrates, and lipids. While some aldehydes play vital roles in normal physiological processes, including vision, embryonic development, and neurotransmission, many are cytotoxic and carcinogenic (Marchitti et al., 2008). Drugs such as cyclophosphamide and ifosfamide are metabolised to aldehyde intermediates, which are substrates for ALDHs
(Bunting and Townsend, 1996). As such, ALDH protect the organism from potentially harmful xenobiotics that contain aldehyde functional groups or xenobiotics that give rise to aldehydes (Moreb, 2008; Moreb et al., 2008).

Aldehydes are detoxified primarily through reductive and oxidative phase I enzyme-catalysed reactions, including the non-P450 aldehyde reduction enzymes as alcohol dehydrogenase (ADH), aldo-keto reductase (AKR), short-chain dehydrogenase/reductase (SDR), and the aldehyde oxidation enzymes as xanthine oxidase (XO), aldehyde oxidase (AOX) and aldehyde dehydrogenase (ALDH) (Marchitti et al., 2008).

Aldehyde dehydrogenases (ALDHs) are a group of NAD(P)+ dependent enzymes involved in oxidising a wide variety of aldehydes into their corresponding carboxylic acids (Ucar et al., 2009). At present, 19 putatively functional human ALDH genes (Marchitti et al., 2008) with distinct chromosomal locations have been identified (see Table 1.2). Furthermore, many allelic variants within the ALDH gene family have been identified, resulting in pharmacogenetic heterogeneity between individuals, which in most cases, results in distinct phenotypes (Moreb et al., 2008). Broadly, ALDH isoenzymes can be categorised as critical for normal development and/or physiological homeostasis. ALDH proteins are found in all subcellular regions including cytosol, mitochondria and endoplasmic reticulum with several found in more than one compartment. ALDH isoenzymes found in organelles other than cytosol possess leader or signal sequences that allow their translocation to specific subcellular regions (Braun et al., 1987).
<table>
<thead>
<tr>
<th>ALDH</th>
<th>Constitutive tissue distribution</th>
<th>Subcellular location (predominant)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1A1</td>
<td>liver, kidney, erythrocytes, skeletal muscle, lung, breast, lens, stomach mucosa, brain, pancreas, testis, ovary, more</td>
<td>Cytosol</td>
</tr>
<tr>
<td>1A2</td>
<td>testis, small amount in liver, kidney, more</td>
<td>Cytosol</td>
</tr>
<tr>
<td>1A3</td>
<td>kidney, skeletal muscle, lung, breast, testis, ovary, stomach mucosa, salivary glands</td>
<td>Cytosol</td>
</tr>
<tr>
<td>1B1</td>
<td>liver, kidney, heart, erythrocytes, skeletal muscle, lung, brain, prostate, testis, placenta</td>
<td>Mitochondria</td>
</tr>
<tr>
<td>1L1</td>
<td>liver, kidney, skeletal muscle</td>
<td>Cytosol</td>
</tr>
<tr>
<td>1L2</td>
<td>pancreas, heart, and brain.</td>
<td>Mitochondria</td>
</tr>
<tr>
<td>2</td>
<td>liver, kidney, heart, skeletal muscle, lung, lens, stomach mucosa, brain, pancreas, prostate, spleen, more</td>
<td>Mitochondria</td>
</tr>
<tr>
<td>3A1</td>
<td>stomach mucosa, cornea, breast, lung, lens, esophagus, salivary glands, skin, more</td>
<td>Cytosol, partially in nucleus</td>
</tr>
<tr>
<td>3A2</td>
<td>liver, kidney, heart, skeletal muscle, lung, brain, pancreas, placenta, more</td>
<td>Endoplasmic reticulum</td>
</tr>
<tr>
<td>3B1</td>
<td>kidney, lung, pancreas, placenta</td>
<td>Endoplasmic reticulum</td>
</tr>
<tr>
<td>3B2</td>
<td>parotid glands</td>
<td>Endoplasmic reticulum</td>
</tr>
<tr>
<td>4A1</td>
<td>liver, kidney, heart, skeletal muscle, brain, placenta, lung, pancreas, spleen, more</td>
<td>Mitochondria</td>
</tr>
<tr>
<td>5A1</td>
<td>liver, kidney, heart, skeletal muscle, brain</td>
<td>Mitochondria</td>
</tr>
<tr>
<td>6A1</td>
<td>liver, kidney, heart, skeletal muscle</td>
<td>Mitochondria</td>
</tr>
<tr>
<td>7A1</td>
<td>foetal liver, heart, skeletal muscle, more</td>
<td>Cytosol</td>
</tr>
<tr>
<td>8A1</td>
<td>liver, kidney, brain, breast, testis, more</td>
<td>Cytosol</td>
</tr>
<tr>
<td>9A1</td>
<td>liver, kidney, heart, skeletal muscle, brain, pancreas, adrenal gland, spinal cord, more</td>
<td>Cytosol</td>
</tr>
<tr>
<td>16A1</td>
<td>kidney, liver, blood</td>
<td>Potential transmembrane protein</td>
</tr>
<tr>
<td>18A1</td>
<td>kidney, heart, skeletal muscle, pancreas, testis, prostate, spleen, ovary, thymus, more</td>
<td>Mitochondria</td>
</tr>
</tbody>
</table>

Table 1.2 Human ALDHs: tissue distribution and subcellular location. Table adapted from Sladek et al. (Sladek, 2003).
The primary, if not sole, role of these enzymes appears to be catalysing the biotransformation of single endobiotics for which they are relatively specific and where the resultant metabolite is essential to the organism (Sladek, 2003). Most of the human ALDHs for which the relevant information is available fall into this category.

I.5.2 Aldehyde dehydrogenase activity

The activity of mammalian ALDHs was first observed in bovine liver more than 50 years ago (Raker, 1949) and thereafter several types of ALDHs were distinguished on the basis of their physico-chemical characteristics, enzymological properties, subcellular localization, and tissue distribution. Since the primary structures of human cytosolic (Hempel et al., 1984) and mitochondrial (Hempel et al., 1985) ALDHs are known, a structurally based classification has been adopted. The most adopted classification is grouped into three distinct, but related, structural families: class 1 and class 3 ALDHs are cytosolic, while class 2 ALDHs are mitochondrial isoenzymes. Despite this distinction, the common function is the irreversible oxidation of a wide spectrum of endogenous and exogenous aldehydes.

Class 1 and class 2 ALDHs preferentially oxidise small aliphatic aldehydes such as propionaldehyde, acetaldehyde, and retinaldehyde (Vasiliou et al., 2004) and require a NAD$^+$ cofactor (Sladek, 2003). Class 3 (also called “tumour-associated”) ALDHs are cytosolic or microsomal enzymes that appears to prefer aromatic aldehydes such as benzaldehyde as substrate. This class can utilize NADP$^+$ as cofactor in addition to NAD$^+$ (Jelski and Szmitkowski, 2008). Class 3
ALDHs also efficiently oxidise lipid aldehydes which can arise from lipid peroxidation and thus may play a physiological role in the cell under conditions of oxidative stress (Sladek, 2002).

![Ethanol metabolism diagram](image)

**Figure 1.7 Ethanol metabolism by ADH and ALDH.** Ethanol is firstly converted into acetaldehyde via NAD-dependent oxidation mediated by ADH and acetaldehyde is the substrate for NAD-dependent reaction by ALDH to produce acetate, as harmless product.

The role of ALDHs in alcohol metabolism (see Figure 1.7), vitamin A absorption, and resistance against oxazaphosphorines (Sladek, 1999) are some of the most studied aspects of their activities. Although preferred substrates and cofactors have been identified for several ALDHs, most are able to catalyse, albeit less efficiently, the oxidation of a wide range of aldehydes as well as to use the alternate cofactor. The overall reaction catalysed by the ALDHs is:

\[
RCHO + NAD^+(P) + H_2O \rightarrow RCOOH + NADH (P) + H^+ 
\]
In this NAD(P)⁺ dependent reaction, the aldehyde enters the active site through a channel located on the outside of the enzyme. The active site contains a Rossman fold (a protein structural motif found in proteins that bind nucleotides, especially the cofactor NAD) and interactions between the cofactor and the Rossman fold allow the enzyme to keep the active site functional (Liu et al., 1997). The mechanism is elucidated in detail in Figure 1.8.

Figure 1.8 ALDH proposed mechanism of aldehyde metabolism. Cys-302 residue is found to be determinant for the interaction with the substrate (Hempel et al., 1985).
I.5.3 Aldehyde dehydrogenase as a functional biomarker

A biomarker is defined as a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacological responses to a therapeutic intervention (Ruddon, 2007). Identification of selective biomarkers will facilitate new selective strategies for isolation and treatment of normal and cancer SCs in a target-orientated manner to enable further investigation of this precious cellular population. Unfortunately, attempts to establish a SC molecular signature using gene profiling have been hampered by the difficulty of obtaining a pure population of SCs.

A potential strategy for identifying normal or CSC populations involves measurement of ALDH activity (Jiang et al., 2009; Ma and Allan, 2011). The retinoic acid signalling of the ALDH is linked to cellular differentiation during development and plays a role in SC self-protection throughout an organism’s lifespan (Croker AK 2008). Furthermore, the role of ALDHs in drug resistance (Sladek, 1999) and retinoic acid generation is potentially crucial for the protection of SCs against toxic endogenous and exogenous aldehydes and for their ability to differentiate respectively (Bunting and Townsend, 1996). Hematopoietic SCs have been shown to express high levels of ALDHs (Storms et al., 1999) and many studies have demonstrated successful isolation of normal human hematopoietic progenitor cells (Jones et al., 1995), using an ALDH activity-based strategy alone or in combination with cell surface markers (i.e. CD133).

A crucial study by Ginestier’s group has demonstrated that expression of ALDH1 in breast tumours is a predictor of poor clinical outcome (Ginestier et al.,
and that high ALDH activity is present in both normal and tumourigenic human mammary epithelial cells with stem or progenitor properties. Therefore, the use of ALDH activity as a purification strategy allows non-toxic and efficient isolation of subpopulations based on a developmentally conserved stem/progenitor cell function.

I.5.4 Heterogeneity of ALDH isoforms and biological implications

Several ALDH isoenzymes have been identified for their role in the protection of SCs against endogenous and exogenous aldehydes and for their role in differentiation. The most studied ALDH1A1 is a well established marker for hematopoietic SC progenitors (Moreb et al., 2008) and its function in retinoic acid formation has been extensively investigated in dopaminergic neuron differentiation and survival (Jacobs et al., 2007). Although many studies point to ALDH1 as the candidate marker for SCs (Ginestier et al., 2007; Kastan et al., 1990), it is becoming apparent that other isoforms may contribute to the ALDH activity being used to identify SC progenitors. For example, the recent identification of several ALDH1 isoforms (ALDH1A1, ALDH1A2, ALDH1A3) in the SC compartment suggests that the ALDH activity observed in SCs is not exclusively related to the 1A1 isoform. Moreover, retinoic acid synthesis by ALDH1A2 promotes differentiation, cell growth arrest and apoptosis and has been suggested as a candidate TSG in prostate cancer (Kim et al., 2005). ALDH1A3 expression is downregulated in human breast cancer MCF7 cells (Marcato et al., 2011) and a recent study has shown that ALDH1A3 deficiency may play a critical role in cancer and metastasis. Studies into the role of retinoic
acid in granulocyte differentiation of hematopoietic SCs revealed that ALDH1A1 and 1B1, but not ALDH1A2 or 1A3, were expressed in CD34<sup>+</sup> hematopoietic progenitors (Moreb et al., 2008).

ALDH2 is the primary enzyme involved in ethanol metabolism and several mutant ALDH2 alleles have indicated a potential mechanism of DNA damage and cancer development (Matsuda et al., 2006), but not SC specific implication has been found. In vitro, ALDH3A1 has been shown to prevent DNA damage and reduce apoptosis from various toxins (hydrogen peroxide, mitomycin C and etoposide) (Lassen et al., 2006) and over-expression of ALDH3A1 together with ALDH1A1 in cell lines and normal hematopoietic progenitors results in a significant increase in resistance to the active metabolites of cyclophosphamide (Sreerama and Sladek, 2001). Furthermore, it has been proven that down-regulation of ALDH1A1 and 3A1 using antisense RNA results in increased sensitivity of tumour cells to 4-hydroperoxycyclophosphamide, an active derivative of cyclophosphamide (Moreb et al., 2000).

Recently ALDH7A1, also known as antiquitin, has been shown to be a candidate marker for human prostate cancer. It has been identified in the overlap gene profile of different SC populations and it may be required for the acquisition of a metastatic stem or progenitor cell phenotype in prostate cancer (van den Hoogen et al., 2011b).

In summary, with the increasing number of reports on ALDH isoforms, it has become apparent that there is a need for a better understanding of the importance of their expression and function in both normal and cancer tissues.
I.5.5 The use of ALDH activity to isolate stem cell populations

Previous techniques for ALDH detection, including measurement of substrate oxidation in whole cell lysates or reaction of fixed cells with antibodies, have been lethal to the cells being studied (Russo et al., 1989). A technique that can measure cytosolic ALDH in intact, viable cells is of interest as part of a strategy to isolate normal and cancer SCs. Furthermore, isolating viable CSCs by their cytosolic ALDH levels could have important implications for studying drug resistance, including cyclophosphamide (Jones et al., 1995).

I.5.6 Aldefluor-based cell sorting

Multiple studies (see section I.5.4) have been published using ALDH1 as a marker to select normal or cancer SCs (Ginestier et al., 2007; Jones et al., 1995). This has been greatly facilitated by Aldefluor, a functional flow cytometric assay. Aldefluor assay (StemCell Technologies) is based on a fluorescent compound that is able to undergo intracellular metabolism. The active form of this fluorescent dye, known as BODIPY-amino-acetaldehyde (BAAA) (see Figure 1.9), is recognised and metabolised by ALDH. The result is an ALDH-dependent intracellular accumulation of BODIPY-aminoacetate (BAA) with intense fluorescence of viable cells that express high levels of ALDH activity (Storms et al., 1999). Aldefluor is supplied in the form of a protected aldehyde, BODIPY-aminoacetaldehyde diethyl acetal (BAAA-DA), which itself is not a substrate of ALDH1. However, when BAAA-DA is exposed to aqueous acid prior to cell-treatment and converted into BODIPY-amino-acetaldehyde (BAAA),
metabolism by ALDH is possible. BAAA is uncharged and can easily diffuse across the cellular membrane of intact viable cells.

The net negative charge, generated by ALDH via aldehyde oxidation to form the oxidised BODIPY-aminoacetate (BAA) product, does not reduce the ability for transport across the cellular membrane. Indeed, the BODIPY chromophore is a good substrate for the multidrug resistance (MDR) transporter family, and hence successful application of the Aldefluor assay is possible only when samples are co-treated with a MDR inhibitor (i.e. verapamil) (StemCellTechnologies. 2004), which blocks the efflux and allows the detection of fluorescence signal (Pearce and Bonnet, 2007). The laser source to analyse the positive cells is determined by having a control sample that is treated similarly except for addition of diethylaminobenzaldehyde (DEAB), an ALDH inhibitor, which will prevent the formation of fluorescent BAA (see Figure 1.10).

Thus, the Aldefluor assay, for the first time, allows the identification and isolation of viable cells that expresses high ALDH activity (Alison et al., 2010). These cells are referred to as ALDH bright. The assay reaction is detected using a flow cytometer equipped with a 488 nm argon laser for excitation and an optical filter set to detect 515-545 nm emission fluorescence.
Figure 1.9 Proof of principle of BODIPY-DA functionality. Structures of BAAA-DA, BAAA, and BAA are reported. BAAA-DA is converted to BAAA in the presence of 2 M hydrochloric acid for 30 min. ALDH then converts BAAA into BAA, which is trapped intracellularly due to concurrent use of the MDR inhibitor verapamil.
Aldefluor-based cell sorting, alone or in combination with tissue-specific cell surface markers (see Table 1.1) is now an established technique for isolating normal and cancer SC fractions. Nonetheless, as elucidated above, the question is which ALDH isoenzyme(s) contribute to the ALDH activity in SCs and whether these are expressed differently depending on the tissue that is being analysed.

**Figure 1.10 The basis of the Aldefluor reaction.** Cells are incubated with BAAA in the presence of verapamil to inhibit efflux by ABC transporter. In the presence of DEAB, ALDH activity is abolished and no intense green fluorescent subpopulation can be detected. Both plots are theoretical FACS histograms. Adapted from (Alison et al., 2010).
I.5.7 Aldefluor and ALDH selectivity

In the past five years, literature has been intensively populated with studies on SC isolation using the Aldefluor method. Several reports have addressed the question regarding which isozyme(s) contribute to the ALDH activity detected by Aldefluor (Sladek et al., 2002). As stated in the section I.5.4 (Kastan et al., 1990), ALDH1A1 is the main isoform in hematopoietic SCs and in normal and cancerous tissue of the breast while ALDH3A1 is also detected (Sladek et al., 2002). Similar data exist in both normal and malignant colon tissue (Yin et al., 1994). Furthermore, the presence of very high ALDH activity such as in A549 cells exposed a new limitation to the assay. This limitation was most likely related to the limited ability of DEAB to inhibit the activity of large amounts of ALDH1A1 and ALDH3A1. Indeed, according to the manufacturer’s recommendations, Aldefluor staining is optimised in cells with low levels of ALDH activity. Recent studies have shown that the enzymatic activity of ALDH1A2 alone (Yokota et al., 2009) or in combination with ALDH2 can be detected by the Aldefluor assay (Moreb et al., 2012). In addition, ALDH1A3 has been found of primary importance for Aldefluor activity in the isolation of breast CSCs from patient tumour samples (Marcato et al., 2011).

In light of these findings, it is worthwhile to assess the potential regulation of ALDH1A1 and other selected ALDH isoforms, as it may have important prognostic revelations in cancers.
I.6 Anthraquinone-based ALDH targeting reagents

As discussed in section I.5.5, the challenge of developing selective fluorescent probes able to identify SC populations via targeting of the ALDHs is still unaccomplished. As is the case with the use of the Aldefluor assay, the production of metabolites that may be substrates for the ABC transporter family of proteins, is considered as a further marker for SC population identification (Patrawala L, 2005).

The concept of design reporter molecules with distinct and informative fluorescence profiles that can rapidly form equilibrium within nuclear compartments can aid the building of pharmacokinetic models to support drug development and deployment (Njoh KL, 2006). Many synthetic small molecules are currently being explored as sophisticated and biologically safe bioimaging probes to detect specific biomolecules with the purpose of visualising cellular and molecular events in living cells. One class of agent, that includes DRAQ5, is used successfully to stain both fixed and live cells is based on the tricyclic anthraquinone chromophore. DRAQ5 (discussed in further details in section V.1.2) is already in the market as a non-invasive cytometric tool for tracking molecular interactions in live cells. This current work focuses on the potential that an anthraquinone fluorescent chemical probe can be developed to target ALDH expressed in normal or CSCs. A deep understanding of the design and the utility of the anthraquinones as fluorophores or as anticancer agents (discussed in section II.1) provides an opportunity to develop small molecules with improved features over Aldefluor that can be used to probe and identify the CSC (see sections II.1-2 for more details).
Chapter I

I.7 The concept of probe design for a dynamic assay

There are several characteristics of a successful optical molecular probe for live cell imaging, including wavelength, brightness, bio- and photo-stability, and cellular loading including eventual specific or non-specific affinity for cellular components (Kobayashi et al., 2009). Design of a small molecule fluorophore as a functional assay is a combination of these features. Generally, those that are bright, small, and hydrophilic, contain no net charge are better candidates for live imaging. For example, key advantages of DRAQ5 exploited in section I.6 include a simple labelling protocol, spectral compatibility with other markers (i.e. GFP, FITC) and a capacity to rapidly enter and intercalate with the DNA of living cells. However, a major concern for DRAQ5 cell loading is its toxicity, which does not allow recovering the cells after staining. Therefore, a major challenge with molecular probes is synthesing agents that are selective for a cellular target while minimising toxicity by having affinity for an extrusion mechanism to rescue the pool of cells after analysis. Interestingly, low molecular weight fluorophores can be designed to be sensitive to enzymatic catalysis, i.e. ALDH affinic, so that they activate in specific environments. Other conditions such as acidic pH and the presence of singlet oxygen or other reactive species can influence the performance of small molecule fluorophores. In this thesis, all these requirements will be carefully brought together to design a functional probe with specific properties to use successfully within a cell-based assay.
I.8 Aims of the thesis

There is currently a poor knowledge of how the ALDH family of enzymes is expressed and regulated (see section I.5.4). This includes a limited understanding on substrate selectivity of ALDHs and an insight into the competitive extrusion mechanisms of their metabolites. Although the Aldefluor assay is frequently used in isolating cell populations with SC properties by targeting ADH1A1, its utility as a functional assay is somewhat hampered by lack of ALDH isoform selectivity. Accordingly, this project is in part undertaken to better understand the ALDH expression in vitro, with specific focus on potential epigenetic regulation of this class of enzymes (see section I.4.4).

In addition, ABC transporters will also be assessed for potential epigenetic regulation, as these are often responsible for mediating anthraquinone-based efflux outside of cells. Overall, the current lack of knowledge poses difficult challenges in the design and development of effective molecular fluorescent probes targeting specific ALDH isoforms. Therefore, the main goals of this PhD project are (i) to contribute to the understanding of ALDH expression and regulation, and (ii) to carry out synthetic and biological investigations of novel anthraquinone-based small molecules that can be used to explore the possibility of targeting ALDHs. The specific aims are as follows:

1. To design and synthesise a novel library of 1,4-di-substituted acetalanthraquinones (AAQs) (Chapter II);

2. To investigate whether the AAQs are susceptible to acidic hydrolysis and ALDH-dependent metabolism (Chapter III);
3. To assess the DNA binding properties and cytotoxicity of the AAQs and their hydrolysed (HAAQ) counterparts (Chapter IV);

4. To investigate the fluorescent properties, cellular uptake and compartmentalisation of the designed AAQs and selected HAAQs in order to understand their potential as functional assay (Chapter V);

5. To investigate the expression and epigenetic regulation of isoforms selected from the ALDH 1-3 family (ALDH1A1, 2 and 3A1) and two principal ABC transporters (ABCB1 and ABCG2) which are known to pump anthraquinones out of cells (Chapter VI);

6. To investigate the expression and epigenetic regulation of ALDH1A1, 2 and 3A1 in prostate cell lines (normal and malignant) and in clinical specimens containing SC populations.

Each individual chapter in this thesis will further provide pertinent overviews to frame the context of more detailed experimental objectives and provide details of the specific methodologies employed. Finally, overall conclusions and directions for future research will be discussed in Chapter VIII.
Chapter II

Design and synthesis of 1,4-di-substituted acetalanthraquinones
II.1 Introduction and Aims

II.1.1 Symmetrical and non-symmetrical 1,4-di-substituted anthraquinones

Over the last 30 years, a plethora of anthraquinones have been synthesised and biologically evaluated aiming to understand their interaction with DNA and develop applications for subsequent biological application. DNA intercalators, including anthracyclines and anthraquinones, contain an electron-deficient chromophore that stabilises the binding of these planar compounds between the electron rich bases of DNA in a non-covalent manner. The driving force is the stacking and electron donor/acceptor interactions, due to van der Waals and electrostatic forces (Neidle and Abraham, 1984). The interaction between DNA and 1,4-di-substituted anthraquinones has been extensively studied and molecular modelling has been used to aid the understanding of the various binding mode differences (Figure 2.1). The strong DNA binding leads to various biological effects including chromatin compaction, topoisomerase II poisoning and DNA strand breaks (Bailly et al., 1996). In regard to the DNA binding, previous studies (Islam SA, 1985; Lerman, 1961; Murdock K. C., 1979) have confirmed that it is possible to increase or decrease the DNA affinity and the potential anticancer activity by making modifications to the anthraquinone chromophore or to the sidechain positions attached to the chromophore.
Some of the most interesting modifications have included 5,8-dihydroxylation of the aromatic nucleus, which substantially increased biological activity due to a slower dissociation from the DNA macromolecule (Pors et al., 2003; Routier et al., 1996).
Not all chromophore modifications have resulted in potent *in vitro* activity as evidenced by lack of activity from 5,6-dihydroxyl, 6-methyl, 6-carboxyl or 6,7-dichloro substituted anthraquinones. Increased chain length between the key nitrogen atoms reduced (n = 3) or abolished (n = 4, 5) activity and the optimal distance between the two N atoms was established to be 3-carbon linker (n = 2) (Figure 2.2).

![Figure 2.2 1,4-di-substituted anthraquinone structure with possible modification sites.](image)

Terminal N-substituted pyridyl, N-aryl or N-acyl derivatives are often inactive and it has been demonstrated that the N-terminal atoms must be fully basic to bind efficiently to phosphodiester residues of the DNA backbone (Murdock K. C., 1979). Further studies demonstrated that insertion of a methyl group at the chromophore aniline nitrogen led to decreased activity (Krapcho et al., 1990; Krapcho et al., 1991), probably because the substituent is forced out of the plane defined by the aromatic system interfering with the DNA intercalation process. Furthermore, modification of the ethylenediamine sidechains by
insertion of \( N \)-(hydroxyalkyl) substituents led to improved biological activity, suggesting that high hydrophilicity is more important than “mere steric brevity” (Murdock K.C., 1979). Compounds containing the \( N,N \)-dimethylethylenediamine sidechain structure in position 1, such as “AQ4” or “AQ6” derivatives (Figure 2.3) have been shown to be excellent DNA binders and cytotoxic agents (Smith et al., 1997).

![Chemical structures of exemplars of 1,4-di-substituted symmetrical and non-symmetrical anthraquinone-based compounds with high DNA affinity. (A) AQ4 and (B) AQ6.](image)

**Figure 2.3** Chemical structures of exemplars of 1,4-di-substituted symmetrical and non-symmetrical anthraquinone-based compounds with high DNA affinity. (A) AQ4 and (B) AQ6.

**II.1.2 Rationale for the design of ALDH-targeting anthraquinone-based agents**

Based on the existent knowledge of (i) structure-activity relationships (SARs) of 1,4-di-substituted anthraquinones reviewed above, (ii) the understanding of the fluorescent properties of anthraquinones as functional molecular probes (exemplified by DRAQ5, section V.1.2) and (iii) the structural composition of Aldefluor (see section 1.5.6), this project was initially established to synthesise and biologically evaluate novel 1,4-di-substituted symmetrical and non-symmetrical anthraquinone-based compounds as chemical probes with
potential use as ALDH-targeting fluorescent reagents. The design of such chemical probes was focussed on the modification of the chemical functionalities in the β-position to the terminal acetal group as outlined in Figure 2.4.

**Figure 2.4** 1,4-di-substituted acetalanthraquinone LC- and HA- libraries of compounds.
II.1.3 Aims of the chapter

The aim of this chapter is to synthesise novel 1,4-di-substituted acetalanthraquinones (AAQs) with potential of targeting ALDH expressed in SCs. Specifically, acetal-incorporating sidechains were designed to create a range of different DNA binding affinities, which could lead to chemical probes targeting different cellular compartments in the cell. In the present chapter, the "LC-AAQ" library consisting of twelve compounds is discussed in detail and the "HA-AAQ" library consisting of four compounds is presented here, but the synthesis is not discussed, as it was carried out by Dr. Hamdy M. Abdel-Rahman. Both libraries of AAQs are important in regard to deriving meaningful SARs, which will be investigated and discussed in Chapters III-V.
II.2 Results and Discussion

II.2.1 Synthesis of acetalamino-sidechains

Three different sidechains were prepared using synthetic routes comprising 3-5 steps. The sidechains (see Figure 2.5) were subsequently reacted with anthraquinone chromophores to afford a novel library of acetalanthraquinones.

![Figure 2.5 Structural modifications of the sidechains to insert 1,4-di-substituted acetalanthraquinones.](image)

II.2.1.1 Sidechain 1: 5, 5-dimethoxypentan-1-amine

Following a method reported by Sall et coll. (Sall et al., 1997), commercially available 5-aminopentanol was N-protected using phthalic anhydride. The intermediate, isolated in quantitative yield, was oxidised at the primary alcohol with pyridinium chlorochromate (PCC) to afford the respective aldehyde product in 83% yield. The aldehyde group was subsequently protected into dimethyl acetal. Acetal derivatisation was obtained using the method reported by Cameron et al. (Cameron A., 1953). To avoid the usage of harmful dry HCl reported by Cameron et al., the acid catalyst employed was a cationic exchange resin (Amberlite IR-120). The dimethyl acetal was obtained without a need for any purification steps. Hydrazinolysis of the phthalamide protective group
(Sasaki et al., 1978) lead to the desired compound being obtained in quantitative yield (Scheme 2.1).

\[
\begin{align*}
&\text{H}_2\text{N} \quad \text{OH} + \quad \begin{array}{c}
\text{O} \\
\text{O} \\
\text{N} \\
\text{O} \\
\text{O} \\
\text{H}
\end{array} \\
&\xrightarrow{i} \quad \begin{array}{c}
\text{O} \\
\text{O} \\
\text{N} \\
\text{O}
\end{array}
\end{align*}
\]

\[
\begin{align*}
&\text{ii, iii} \quad \text{a} \quad \text{or} \quad \text{iii} \quad \text{b} \\
&\xrightarrow{\text{ii, iii} \quad \text{a} \quad \text{or} \quad \text{iii} \quad \text{b}} \quad \begin{array}{c}
\text{O} \\
\text{O} \\
\text{N} \\
\text{O}
\end{array} \\
&\xrightarrow{\text{iv}} \quad \begin{array}{c}
\text{H}_2\text{N} \\
\text{OH}
\end{array}
\end{align*}
\]

**Scheme 2.1** Route to synthesise 5,5-dimethoxypentan-1-amine. (i) 138 °C 20 h; (ii) CH\textsubscript{2}Cl\textsubscript{2}, MgSO\textsubscript{4}, PCC, RT 1.5 h; (iii\textsubscript{a}) dry HCl 2 min, reflux 15 min; (iii\textsubscript{b}) anh. CH\textsubscript{3}OH, cationic exchange resin, RT 4 days; (iv) CH\textsubscript{3}OH, hydrazine, reflux 4 h.

**II.2.1.2 Sidechain 2: 3-amino-N-(2,2-dimethoxyethyl)propanamide**

Z-Amino protected β-alanine was coupled with commercially available 2,2-dimethoxyethanamine (Scheme 2.2).

\[
\begin{align*}
&\text{O} \\
&\text{O} \\
&\text{N} \\
&\text{O}
\end{align*}
\]

\[
\begin{align*}
&\xrightarrow{i} \quad \begin{array}{c}
\text{O} \\
\text{O} \\
\text{N} \\
\text{O}
\end{array}
\end{align*}
\]

**Scheme 2.2** Adopted synthetic route with Z-β-alanine. (i) Dry DMF, EDC, triethylamine, RT, 36 h.
The coupling reaction was performed using 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide to activate the terminal carboxylic acid towards coupling with the amine terminal group. The reaction was completed after two days at RT. The mechanism of the reaction is shown in Scheme 2.3.

Scheme 2.3 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) reaction mechanism. EDC activates the carboxylic acid towards amide formation.

The advantage of using the Z-β-Alanine was related to the required conditions for the deprotection. Conveniently, the catalytic hydrogenation needed to cleave the Z-group had no effect on the stability of the acetal moiety; following hydrogenation the desired compound was afforded in quantitative yields.

The last step involved deprotection of the amino group using catalytic hydrogenation in the presence of 10% palladium-carbon and hydrogen or cyclohexene at atmospheric pressure as hydrogen donors (Scheme 2.4). Hydrogen was preferred to cyclohexene because the crude product was afforded in quantitative yield and without any need for a purification step.
II.2.1.3 Sidechain 3: N-(2-Aminoethyl)-3,3-dimethoxypropanamide

The N-(2-aminoethyl)-3,3-dimethoxypropanamide sidechain was prepared in two steps. Basic cleavage of 3,3-dimethoxypropanoate using a 1:1 mixture of sodium hydroxide and methanol at RT yielded a colourless oily carboxylic acid as the first intermediate (Scheme 2.5).

![Scheme 2.4 Synthetic route for Z-amino group deprotection. (i) Dry CH₃OH, cyclohexene, 10% Pd-C, reflux 2 h; (ii) dry CH₃OH, H₂, 10% Pd-C, reflux, 2h.]

The second intermediate to prepare the target sidechain needed to be a monoprotected ethylenediamine. In an attempt to obtain this compound, several protecting groups were used to favour the selective reaction of 3,3-dimethoxypropanoic acid with one of the two amino groups on the
ethylenediamine. Carbamate-based protecting groups were chosen according with preferential monoprotection and easiness of deprotection. According to the literature (Greene T. W., 1999), several protective groups are known to be cleaved under basic conditions but a few of them are able to give monoprotected products.

Following the method of Alfred et al. (Alfred J.C., 1996), 2-bromoethylamine was N-protected using di-tert-butyl dicarbonate (BOC₂O) in the presence of triethylamine (Scheme 2.6). Purification by flash chromatography afforded the product in 63% yield.

\[
\begin{align*}
\text{Br} & \quad \text{NH}_2 \\
\text{HBr} & \quad \text{N} \\
\text{H} & \quad \text{Br} \\
\text{O} & \quad \text{N} \\
\text{O} & \quad \text{O} \\
\text{N} & \quad \text{O} \\
\text{O} & \quad \text{NH}_3 \\
\text{Cl} & \quad -\text{Cl}
\end{align*}
\]

Scheme 2.6 Synthetic route for 2-(1,3-dioxoisindolin-2-yl)ethanaminium chloride synthesis. (i) BOC₂O, triethylamine, dry CHCl₃, reflux, 1 h, (ii) potassium phthalamide, dry DMF, 75 °C, 2 days, (iii) 3 M dry HCl/EtOAc, RT, 1 h.

The substitution of the phthalamide group for bromide was performed by reaction of the BOC-protected intermediate in anhydrous DMF with potassium phthalamide salt. The potassium bromide generated with the reaction was removed by filtration and the desired compound was isolated in 80% yield. The
deprotection of the BOC-amino group was achieved by treating the intermediate with a saturated solution of dry HCl in anhydrous ethyl acetate. The phthalamide-protected ethylenediamine was isolated as the HCl salt in 84% yield.

The 3,3-dimethoxypropanoic acid (Scheme 2.5) and the monoprotected ethylenediamine were reacted to form the amide bond using EDC as coupling reagent (see Figure 2.3 for detailed mechanism). The product was obtained after extraction from dichloromethane (46% yield). Deprotection with hydrazine monohydrate yielded the final sidechain in 80% yield (Scheme 2.7).

\[
\begin{align*}
&\text{Scheme 2.7} & \text{Coupling reaction and deprotection of the amino group to yield } N-(2\text{-aminoethyl})-3,3\text{-dimethoxypropanamide. (i) Dry DMF, EDC, triethylamine, RT, 36 h; (ii) dry methanol, hydrazine monohydrate, reflux, 4 h.}
\end{align*}
\]

II.2.2 Synthesis of chromophores

The library of compounds described in detail in this section includes two different chromophores (see Figure 2.10A-B).
Figure 2.6 Chemical structures of 1,4-difluoro-anthraquinones. (A) 1,4-difluoro-anthraquinones; (B) 1,4-difluoro-5,8-dihydroxyanthraquinone.

Chromophore A was supplied by Sigma, but chromophore B was prepared following the procedure described by Krapcho et al. (Krapcho et al., 1990) by reacting 1,4-dihydroxybenzene and 3,6-difluorophthalic anhydride in the presence of a mixture of ground AlCl$_3$ and NaCl at 220 °C (Scheme 2.8).

Scheme 2.8 Preparation of 1,4-difluoro-5,8-dihydroxyanthraquinone. (i) AlCl$_3$, NaCl, 220 °C, 3 h.

After 3 h, the reaction was quenched by pouring the melt onto a mixture of ice and concentrated HCl. Without any purification step, the desired compound was obtained in excellent yield (92%).
Non-symmetrical chromophores were prepared by first reacting N,N-dimethylethlenediamine either with 1,4-difluoroanthraquinone or 1,4-difluoro-5,8-dihydroxyanthraquinone in pyridine (Scheme 2.9).

Scheme 2.9 Formation of non-symmetrical AAQs. (A) Synthesis 1-{2-[N,N-(dimethyl)amino]ethylamino}-4-fluoro-5,8-dihydroxy-anthraquinone. (i) Pyridine, RT, 24 h; (B) Synthesis of 1-{2-[N,N-(dimethyl)amino]ethylamino}-4-fluoro-5,8-dihydroxy-anthraquinone; (i) Pyridine, RT, 24 h.

The intermediate target compounds were isolated in 75% (A) and 67% (B) yield respectively after purification by column chromatography (see Experimental Details at the end of the chapter).

Generally, compounds such as halobenzenes show resistance to nucleophilic attack due to the electron rich nature of the aryl system. However, as reported in the literature (Kikuchi et al., 1982), the insertion of a strong electron-
withdrawing group in the ortho and/or para position to the halogen can favour a substitution with an addition/elimination sequence. As shown in Scheme 2.10, the nucleophilic primary amine forms a bond with the chromophore producing a negatively charged sigma complex (also known as *Meisenheimer* complex).

![Scheme 2.10 Proposed mechanism for nucleophilic aromatic substitution by addition-elimination reaction, via *Meisenheimer* complex.](image)

In this intermediate, the two electron-withdrawing quinones of the anthraquinone structure participate in the electron flow, activating the ring towards nucleophilic attack and the departure of the halide. Strong electron-withdrawing nature of fluorine also helps with this – promotes attack and stabilises intermediate carbanion by inductive effect.
II.2.3 Synthesis of and symmetrical and non-symmetrical 1,4-di-substituted-acetalanthraquinones (AAQ)

The synthesis of the desired compounds was achieved by reaction of the novel sidechains of interest with symmetrical and non-symmetrical anthraquinone chromophores.

Scheme 2.11 Synthesis of symmetrical 1,4-di-substituted-anthraquinones.

The symmetrical 1,4-di-substituted anthraquinones were prepared by reacting the respective chromophores with the aminoalkylacetal sidechains (Schemes 2.11-2.12). The reaction conditions used to generate all target compounds consisted of mixing starting materials and reagents in dry pyridine.
and heating to reflux for 3 h, providing the desired compounds in a range of yields (see Table 2.1).

Scheme 2.12 Synthesis of symmetrical 1,4-di-substituted-5,8-dihydroxy-anthraquinones.

Non-symmetrical anthraquinones were obtained by reacting the intermediate 1-(dimethylamino)ethylamino-4-fluoro-5,8-dihydroxyanthraquinone with the sidechain of interest (Scheme 2.13-2.14) to afford the non symmetrical target compounds in a range of yields (see Table 2.1).
Scheme 2.13 Synthesis of non-symmetrical 1,4-di-substituted-anthraquinones.

Scheme 2.14 Synthesis of non-symmetrical 1,4-di-substituted-5,8-dihydroxy-anthraquinones.
### Table 2.1

Yields obtained from preparation of 1,4-di-substituted anthraquinones.

<table>
<thead>
<tr>
<th>Compound</th>
<th>Scheme reference</th>
<th>Formula</th>
<th>MW [g/mol]</th>
<th>Yield [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>LC-103</td>
<td>2.10</td>
<td>C$<em>{28}$H$</em>{38}$N$_2$O$_6$</td>
<td>498.6</td>
<td>44</td>
</tr>
<tr>
<td>LC-105</td>
<td>2.13</td>
<td>C$<em>{25}$H$</em>{33}$N$_3$O$_6$</td>
<td>471.5</td>
<td>49</td>
</tr>
<tr>
<td>LC-106</td>
<td>2.10</td>
<td>C$<em>{28}$H$</em>{38}$N$_4$O$_6$</td>
<td>556.2</td>
<td>59</td>
</tr>
<tr>
<td>LC-107</td>
<td>2.10</td>
<td>C$<em>{28}$H$</em>{38}$N$_4$O$_6$</td>
<td>556.2</td>
<td>30</td>
</tr>
<tr>
<td>LC-055</td>
<td>2.11</td>
<td>C$<em>{28}$H$</em>{38}$N$_2$O$_6$</td>
<td>530.3</td>
<td>52</td>
</tr>
<tr>
<td>LC-108</td>
<td>2.11</td>
<td>C$<em>{28}$H$</em>{38}$N$<em>4$O$</em>{10}$</td>
<td>588.6</td>
<td>66</td>
</tr>
<tr>
<td>LC-109</td>
<td>2.11</td>
<td>C$<em>{28}$H$</em>{38}$N$<em>4$O$</em>{10}$</td>
<td>588.6</td>
<td>40</td>
</tr>
<tr>
<td>LC-110</td>
<td>2.12</td>
<td>C$<em>{25}$H$</em>{33}$N$_3$O$_4$</td>
<td>439.5</td>
<td>70</td>
</tr>
<tr>
<td>LC-111</td>
<td>2.13</td>
<td>C$<em>{25}$H$</em>{32}$N$_4$O$_7$</td>
<td>500.5</td>
<td>90</td>
</tr>
<tr>
<td>LC-112</td>
<td>2.12</td>
<td>C$<em>{25}$H$</em>{32}$N$_4$O$_5$</td>
<td>468.5</td>
<td>33</td>
</tr>
<tr>
<td>LC-113</td>
<td>2.12</td>
<td>C$<em>{25}$H$</em>{32}$N$_4$O$_5$</td>
<td>468.5</td>
<td>60</td>
</tr>
<tr>
<td>LC-114</td>
<td>2.13</td>
<td>C$<em>{25}$H$</em>{32}$N$_4$O$_7$</td>
<td>500.5</td>
<td>60</td>
</tr>
</tbody>
</table>

To summarise, three novel amino sidechains containing acetal functionalities were synthesised and substituted onto two fluorinated anthraquinone building blocks to give a library of twelve LC-AAQs for biological evaluation. The results of these studies are discussed in Chapters III-V.
II.3 Experimental Details

II.3.1 Chemicals, Reagents and Instrumentations

All reagents were supplied by Sigma. Chromatography column silica: particle size 35-70 µm and 20-35 µm, was supplied by Fisher Scientific. Thin layer chromatography plates (Aluminium backed) supplied by Merck. All solvents were of standard grade. Dichloromethane, petroleum ether 40-60, chloroform, ethanol, methanol, diethylether, hexane, toluene, ethylacetate and acetonitrile were supplied by Fisher Scientific. Anhydrous DMF was purchased from Aldrich. Proton Nuclear Magnetic Resonance ($^1$H NMR) spectra were recorded using Bruker AMX400 (400 MHz) spectrometer. Carbon Nuclear Magnetic Resonance ($^{13}$C and NMR) were performed on the same machines operating at 101 MHz. Deuterated chloroform ($\text{CDCl}_3$) was used as the NMR solvent unless otherwise stated. Chemical shifts are reported in parts per million (δ, ppm). $^1$H NMR chemical shifts are reported relative to an internal reference (tetramethylsilane) or residual proton signals of the solvent. Coupling constants (J) are expressed in Hertz (Hz). The splitting patterns in NMR spectra are reported with the following abbreviations; singlet (s), doublet (d), triplet (t), quartet (q), quintet (quint.), multiplet (m) and broad (br). $^{13}$C NMR chemical shifts are reported relative to the signal of the solvent. Routine mass spectra was carried out by Andrew Healey and were run on a Micromass Quattro Ultima spectrometer in the electron impact (EI), chemical ionisation (CI) or positive (+) electrospray mode as stated. High resolution mass spectrometry (HRMS) was carried out at the ESPRC National mass spectrometry service centre located in Swansea University. Purity of the final LC-compounds was performed by HPLC.
analysis on Agilent Technologies 1200 HPLC system with diode array detection, using C18 reversed phase columns (Agilent Eclipse XDB - analytical: 4.6 x 100 mm). Solvent A: H2O + 0.045% TFA; Solvent B: MeCN 90%; H2O 10% + 0.045% TFA.

II.3.2 Synthesis of 1,4-disubstituted-acetalanthraquinones

II.3.2.1 Synthesis of acetalamino sidechains

2-(5-Hydroxypentyl)isoindoline-1,3-dione

A stirred mixture of 5-aminopentanol (1 g, 9.69 mmol) and phthalic anhydride (1.44 g, 9.75 mmol) was heated to 138 °C for 20 h to yield (quantitative) the title compound as yellow oil. No work-up was performed and the product was directly used for following reaction step. Rf 0.06 (EtOAc:PE, 3:7); δH (CDCl3, 400 MHz) 7.83 (2H, dd, J 3.0, 5.6 Hz), 7.71 (2H, dd, J 3.0, 5.6 Hz), 3.69 (2H, t, J 7.3 Hz), 3.64 (2H, t, J 6.6 Hz), 1.87 (1H, br s, OH), 1.71 (2H, qn, J 7.3 Hz), 1.58-1.65 (2H, m), 1.38-1.45 (2H, m). Results agreed with data reported by Cameron et al. (Cameron A., 1953).
**5-(1,3-Dioxoisoindolin-2-yl)pentanal**

![5-(1,3-Dioxoisoindolin-2-yl)pentanal](image)

To a stirred solution of 2-(5-hydroxypentyl)isoindoline-1,3-dione (1.72 g, 7.37 mmol) in CH₂Cl₂ (35 mL) was added MgSO₄ (3.5 g) and PCC (4.77 g, 22.11 mmol) and the resulting suspension was stirred at RT for 1.5 h. The reaction mixture was filtered through Celite™, concentrated *in vacuo* and purified by flash column chromatography (EtOAc:PE, 2:3) to yield the title compound (1.41 g, 83%) as a colourless oil. R<sub>f</sub> 0.22 (EtOAc:PE, 2:3); δ<sub>H</sub> (CDCl₃, 400 MHz) 9.58 (1H, t, J 1.5, CHO), 7.84 (2H, dd, J 3.0, 5.6 Hz), 7.71 (2H, dd, J 3.0, 5.6), 3.71 (2H, t, J 7.1 Hz), 2.50 (2H, dt, J 1.5, 7.1 Hz), 1.63-1.77 (4H, m); m/z (ES+) 233.2 (M, 17%), 232.2 (M+H, 100).

**2-(5,5-Dimethoxypentyl)isoindoline-1,3-dione**

![2-(5,5-Dimethoxypentyl)isoindoline-1,3-dione](image)

To a stirred solution of 5-(1,3-dioxoisoindolin-2-yl)pentanal (1.98 g, 8.5 mol) at RT in anhydrous CH₃OH (70 mL) was added Amberlite IR-120. After 4 days, the reaction mixture was filtered, washing with MeOH. The filtrate was concentrated *in vacuo* to yield the title compound as a colourless oil (1.4 g, 60%). δ<sub>H</sub> (CDCl₃, 400 MHz) 7.77 (2H, m, J 2.02, 5.0 Hz), 7.63 (2H, dd, J 3.0, 5.5 Hz), 4.27 (1H, t,
5,5-Dimethoxypentan-1-amine

To a stirred solution of 2-(5,5-dimethoxypentyl)isoindolin-1,3-dione (1.4 g, 5.05 mmol) in CH$_2$OH (25 mL) was added hydrazine (1.01 mL, 20.19 mmol) and the reaction mixture was heated at reflux temperature for 4 h. The solvent was removed in vacuo and the reaction mixture was dissolved in 10 mL of CH$_2$Cl$_2$, washed with H$_2$O (3 x 10 mL) and dried over MgSO$_4$. Filtrate was concentrated in vacuo to yield the title compound (580 mg, 78%) as a light yellow oil. δ$_H$(CDCl$_3$, 400 MHz) 4.29 (1H, t, $J$ 5.5, 6.0 Hz), 3.24 (6H, s, O(CH$_3$)$_2$), 2.62 (2H, t, $J$ 7.1 Hz), 1.56-1.51 (2H, m), 1.43-1.36 (2H, m), 1.34-1.27 (2H, m); δ$_C$(CDCl$_3$, 100 MHz) 104.4, 52.6, 41.9, 33.5, 33.2, 24.0; m/z (ES+) 147.9 ($M+H$, 100).

Benzyl 3-(2,2-dimethoxyethylamino)-3-oxopropylcarbamate

Method of Jackson (Jackson A. E., 1976): To a stirred solution of Z-β-Alanine (4.0 g, 20 mmol) in anh. DMF (50 mL) was added 2,2-dimethoxyethanamine (2.83 g, 0.03 mol), ethyl-(N',N'-dimethylamino)propylcarbodiimide hydrochloride (EDC) (6.86 g, 0.04 mol) and triethylamine (4.92 mL, 0.04 mol). The resulting mixture was stirred for 36 h at RT, and then the solvent was evaporated under
reduced pressure. The resulting solid was dissolved in CH$_2$Cl$_2$ (20 mL) and washed with 5% citric acid (2 x 5 mL), 5% NaHCO$_3$ (2 x 5 mL) and brine (1 x 10 mL) before being dried over MgSO$_4$. The solution was filtered and concentrated in vacuo to yield the crude final compound as a colourless oil (3 g, 54%). $\delta$$_H$ (CDCl$_3$, 400 MHz) 7.27 (5H, m), 5.02 (2H, s), 4.28 (1H, t, $J$ 5.1 Hz), 3.4 (2H, t, $J$ 6.1 Hz), 3.32 (2H, $J$ 5.5 Hz), 3.30 (6H, s), 2.35 (2H, t, $J$ 6.1 Hz); $\delta$$_C$ (CDCl$_3$, 100 MHz) 128.5, 128.1, 128.0, 102.5, 54.4, 40.9; m/z (ES+) 311 ($M+H$, 100).

3-Amino-N-(2,2-dimethoxyethyl)propanamide

To a solution of benzyl 3-(2,2-dimethoxyethylamino)-3-oxopropylcarbamate (100 mg, 0.32 mmol) in CH$_3$OH (3 mL) was added cyclohexene (264 mg, 3.22 mmol) and commercial 10% Pd-C catalyst (20 mg). The mixture was heated under reflux for 2 h then the solution was filtered through Celite™. The catalyst was washed with CH$_3$OH and the combined filtrates were concentrated under reduced pressure to give the deprotected crude product (54 mg, 95%). $\delta$$_H$ (CDCl$_3$, 400 MHz) 7.05 (1H, br s, NH), 4.41 (1H, t, $J$ 5.5 Hz), 3.41 (2H, t, $J$ 5.5 Hz), 3.4 (6H, s), 3.03 (2H, t, $J$ 5.5 Hz), 2.38 (2H, t, $J$ 4.0 Hz), 2.19 (2H, br s, NH); $\delta$$_C$ (CDCl$_3$, 100 MHz) 172.4, 102.6, 54.2, 50.7, 40.7, 38.3, 38.1; m/z (ES+) 177 ($M+H$, 100).
3,3-Dimethoxypropanoic acid

\[
\begin{align*}
\text{O} & \quad \text{O} \\
\text{\hspace{0.5cm} O} & \quad \text{\hspace{0.5cm} OH}
\end{align*}
\]

Methanol was added dropwise to a solution of NaOH (2 M) in order to prepare a NaOH-CH\(_3\)OH (1:1) solution. The methyl 3,3-dimethoxypropanoate (1 g, 6.75 mmol) was stirred at RT overnight in 10 mL of the prepared NaOH: CH\(_3\)OH (1:1) solution. The reaction mixture was treated with Amberlite IR-120 until pH reached 6. The resin was washed several times with CH\(_3\)OH and H\(_2\)O and concentrated under vacuum. The remaining H\(_2\)O was removed by lyophilisation to yield the title compound as colourless crude oil (1 g, quantitative). \(\delta_H (\text{CDCl}_3, 400 \text{ MHz}) 6.14 (1\text{H, br s, NH}), 4.79 (1\text{H, t, } J 5.6 \text{ Hz}), 3.33 (6\text{H, s}), 2.58 (2\text{H, d, } J 5.6 \text{ Hz}).\)

**Tert-butyl 2-bromoethylcarbamate**

\[
\begin{align*}
\text{O} & \quad \text{N} \\
\text{\hspace{0.5cm} O} & \quad \text{\hspace{0.5cm} Br}
\end{align*}
\]

To a stirred slurry of 2-bromoethylamine (200 mg, 0.98 mmol) and triethylamine (0.14 mL, 0.976 mmol) in CH\(_3\)Cl (5mL) was added a solution of (BOC)_2O (213mg, 0.98 mmol) in CH\(_3\)Cl (5 mL). After addition, the reaction mixture was heated for 1 h under reflux, cooled to RT and the solution was washed with H\(_2\)O (3 x 15 mL) and dried over Na\(_2\)SO\(_4\). The organic solvent was evaporated under vacuum and the crude compound was purified by column chromatography (CHCl\(_3\):CH\(_3\)OH, 9:1) to yield the title compound (128 mg, 63%). \(\delta_H (\text{CDCl}_3, 400 \text{ MHz}) 6.76 (1\text{H, CONH}), 3.57 (2\text{H, t, } J 4.5 \text{ Hz}), 3.31 (2\text{H, t, } J 4.5 \text{ Hz}), 1.46 (9\text{H,}}\)
δ<sub>C</sub> (CDCl<sub>3</sub>, 100 MHz) 155.7, 83.7, 56.53, 53.20, 48.8, 32.5, 28.0; m/z (ES+) 305.2, 306.2 (M+H, 100).

**Tert-butyl 2-(1,3-dioxoisooindolin-2-yl)ethylcarbamate**

![Chemical Structure](image)

A solution of tert-butyl 2-bromoethylcarbamate (5.6 g, 0.03 mol) in anhydrous DMF (20 mL) was added to a stirred slurry of potassium phtalamide (4.93 g, 0.03 mol) in anhydrous DMF (20 mL), and the mixture was heated at 75 °C for 2 days. After cooling, KBr was filtered off and the DMF was evaporated under vacuum. The residue was dissolved in CH<sub>2</sub>Cl<sub>2</sub> (100 mL) and the solution was washed with H<sub>2</sub>O (3 x 100 mL), dried over Na<sub>2</sub>SO<sub>4</sub> and evaporated under vacuum to yield a white powder (6.16 g, 80%). δ<sub>Η</sub> (CD<sub>3</sub>OD, 400 MHz) 7.70 (4H, m), 6.60 (1H, brs, NH), 3.65 (2H, t, J 5.8 Hz), 3.23 (2H, t, J 5.8 Hz), 1.11 (9H, s).

**2-(1,3-Dioxoisooindolin-2-yl)ethanaminium chloride**

![Chemical Structure](image)

To a 3 M solution of dry HCl in anhydrous ethyl acetate (1 mL) at 0 °C was added tert-butyl 2-(1,3-dioxoisooindolin-2-yl)ethylcarbamate (60 mg, 0.21 mmol). The slurry mixture was allowed to warm up to RT and stirred for 1 h. The solution was filtered and the precipitate was collected and washed with ether to
yield the crude title compound as a white powder (50 mg, quantitative). δ$_\text{H}$ (CD$_3$OD, 400 MHz) 7.86 (4H, s), 3.94 (2H, J 5.55 Hz), 3.45 (2H, J 5.55 Hz). δ$_\text{C}$ (CDCl$_3$, 100 MHz) 167.7, 133.2, 132.0, 126.7, 41.9, 36.7.

**N-(2-(1,3-Dioisoindolin-2-yl)ethyl)-3,3-dimethoxypropanamide**

\[
\text{N}^2\text{(2-(1,3-Dioisoindolin-2-yl)ethyl)-3,3-dimethoxypropanamide}
\]

To a stirred solution of 3,3-dimethoxypropanoic acid (1.54 g, 0.012 mol) in DMF (15 mL) at RT was added 2-(1,3-dioisoindolin-2-yl)ethylammonium chloride (3.84 g, 0.02 mol), EDC (4.49 g, 0.02 mol) and triethylamine (6.6 mL, 0.05 mol). The resulting mixture was stirred for 36 h. The solvent was evaporated under reduced pressure and the resulting solid was dissolved in CH$_2$Cl$_2$ (20 mL) and washed with 5% citric acid (2 x 5 mL), 5% NaHCO$_3$ (2 x 5 mL) and brine (1 x 10 mL) before being dried over MgSO$_4$. The solution was filtered and concentrated *in vacuo* to yield the final compound (1.68 g, 46%). δ$_\text{H}$ (CDCl$_3$, 400 MHz) 7.75 (4H, s), 7.4 (1H, brs, NH), 4.5 (1H, J 5.6 Hz), 3.69 (2H, J 5.6 Hz), 3.37 (2H, t, J 5.6 Hz), 3.21 (2H, J 5.6 Hz), 3.1 (6H, s). δ$_\text{C}$ (CDCl$_3$, 100 MHz) 173.3, 165.4, 132.3, 130.2, 125, 100.2, 53.1, 49.2, 39.6, 36.2; m/z (ES+) 329 (M+H, 100).

**N-(2-Aminoethyl)-3,3-dimethoxypropanamide**

\[
\text{N}^2\text{(2-Aminoethyl)-3,3-dimethoxypropanamide}
\]
To a stirred solution of $N$-(2-(1,3-dioxisoindolin-2-yl)ethyl)-3,3-dimethoxypropanamide (1.68 g, 5 mmol) in CH$_3$OH (60 mL) was added hydrazine monohydrate (1.06 mL, 0.02 mol) and the resulting mixture was stirred under reflux for 4 h. The mixture was cooled to RT and the solvent was evaporated under reduced pressure. The resulting solid was dissolved in CH$_2$Cl$_2$ (20 mL) and the undissolvable phthalazide was removed by filtration. The solution was concentrated in vacuo to yield the title compound (756.5 mg, 80%). $\delta$$_H$(CDCl$_3$, 400 MHz) 6.44 (1H, brs, NH), 4.63 (1H, t, $J$ 5.1 Hz), 3.32 (6H, s), 3.23 (2H, t, $J$ 6.1 Hz), 2.76 (2H, $J$ 4.5 Hz); $\delta$$_C$(CDCl$_3$, 100 MHz) 169.4, 162.5, 102.2, 54.2, 42.1, 41.0, 40.9; m/z (ES+) 177 ($M$+H, 100).

II.3.2.2 Synthesis of chromophores

1,4-Difluoro-5,8-dihydroxyanthraquinone

A mixture of ground AlCl$_3$ (2.95 g, 22.16 mmol), NaCl (432 mg, 7.39 mmol), 1,4-dihydroxybenzene (224 mg, 2.03 mmol) and 3,6-difluorophthalic (340 mg, 1.85 mmol) were mixed in a round bottom flask and heated to 220 °C for 3 h. The oil-bath was removed and the reaction quenched by addition of ice and concentrated hydrochloric acid (10 mL). The final aqueous solution was filtered under suction and the residue was solubilised in acetonitrile followed by freeze-
drying to yield the title compound as a brown solid (470 mg, 92%). $\delta_H$ (CDCl$_3$, 400 MHz) 7.3 (s, 2H), 7.55 (m, 2H), 12.90 (s, 2H); m/z (ES+) 277 (M+H, 100).

1-(2-(Dimethylamino)ethylamino)-4-fluoroanthracene-9,10-dione

To a stirred solution of 1,4-difluoro-anthaquinone (300 mg, 1.23 mmol) in pyridine (1 mL) was added $N,N$-dimethylethane-1,2-diamine (108 mg, 1.23 mmol) and the resulting solution was heated under reflux for 15 min. The solvent was evaporated under reduced pressure and the reaction mixture was purified by flash chromatography using gradient elution (CH$_3$OH:CH$_2$Cl$_2$, 1:99→4:96) to yield the title compound (236 mg, 75%). $\delta_H$ (CDCl$_3$, 400 MHz); 9.90 (1H, s), 8.15 (2H, dd, $J$ 1.51, 7.5 Hz), 8.14 (2H, dd), 7.30 (1H, dd, $J$ 1.51, 7.5 Hz), 3.32 (2H, q, $J$ 6.0 Hz), 2.59 (2H, t, $J$ 6.5 Hz), 2.28 (6H, s). $\delta_C$ (CDCl$_3$, 100 MHz) 182.4, 154.5, 151.9, 148.7, 134.3, 133.0, 131.2, 130.2, 130.0, 126.6, 120.2, 119.9, 111.5, 57.9, 45.6, 42.0, 41.1; m/z (ES+) 313 (M+H, 100).

1-(2-(Dimethylamino)ethylamino)-4-fluoro-5,8-dihydroxyanthracene-9,10-dione
To a stirred solution of 1,4-difluoro-5,8-dihydroxyanthracene-9,10-dione (600 mg, 2.17 mmol) in pyridine (7 mL) was added \( N,N \)-dimethylethane-1,2-diamine (238 mg, 2.17 mmol) and the resulting solution was stirred at 90 °C for 15 min. The mixture was cooled down to RT and the solvent was evaporated under reduced pressure. The reaction mixture was dissolved in \( CH_2Cl_2 \) (30 mL) and washed with saturated NaHCO\(_3\) (2 x 20 mL). The organic layer was collected and dried over MgSO\(_4\). After evaporation of the solvent, the crude compound was purified by flash chromatography using gradient eluent (\( CH_3OH:CH_2Cl_2, 0:100 \rightarrow 6:94 \)) to yield the title compound (300 mg, 40%).

\[ \delta (CDCl_3, 400 MHz); 13.4 \ (2H, s, OH), 10.7 \ (1H, s), 7.4 \ (1H, dd, J 1.5, 7.5 Hz), 6.8 \ (2H, s), 6.7 \ (2H, s), 3.4 \ (2H, q, J 6.5 Hz), 2.7 \ (2H, t, J 6.5 Hz), 2.35 \ (6H, s). \delta_C (CDCl_3, 100 MHz) 182.1, 154.2, 144.2, 124.8, 123.1, 120.3, 119.2, 118.5, 110.8, 59.1, 46.2, 43.2; m/z (ES+) 344 (M+H, 100).

II.3.2.3 Synthesis of symmetrical 1,4-di-substituted anthraquinones

1,4-Bis(5,5-dimethoxypentylamino)anthracene-9,10-dione

To a stirred solution of 1,4-difluoro-anthraquinone (74 mg, 0.29 mmol) in pyridine (1 mL) was added 5,5-dimethoxypentanamine (261 mg, 1.72 mmol)
and the resulting solution was stirred at 70 °C for 3 h. The solvent was evaporated under reduced pressure and the reaction mixture was purified by flash chromatography using gradient eluent (EtOAc:PE, 0:100→3:7) to yield the title compound as a dark blue powder (62 mg, 44%). \( \delta_H (CDCl_3, 400 \text{ MHz}) 8.24 (2H, dd, J 3.5, 6.0 \text{ Hz}), 7.6 (2H, dd, J 3.5, 6.0 \text{ Hz}), 7.14 (2H, s), 4.32 (2H, t, J 5.6), 3.62 (4H, t, J 6.5 \text{ Hz}), 3.36 (4H, t, J 5.5 \text{ Hz}), 3.25 (12H, s, O(CH_3)_2), 1.75-1.45 (12H, m); \delta_C (CDCl_3, 100 \text{ MHz}) 181.3, 145, 133.5, 130.7, 125.3, 122.1, 108.7, 103.4, 51.7, 41.2, 30.8, 28.5, 21.1; m/z (ES+) 499 (M+H, 100).

3,3'-(9,10-Dioxo-9,10-dihydroanthracene-1,4-diyl)bis(azanediyl)bis(N-(2,2-dimethoxyethyl)propanamide)

To a stirred solution of 1,4-difluoro-anthraquinone (80 mg, 0.33 mmol) in pyridine (1 mL) was added 3-aminoo-N-(2,2-dimethoxyethyl)propanamide (345.7 mg, 1.96 mmol) and the resulting solution was stirred at 70°C for 4 h. The solvent was evaporated under reduced pressure and the reaction mixture was purified by flash chromatography using gradient eluent (CH_3OH:CH_2Cl_2, 1:99→1:9 to yield the title compound as a dark blue powder (107 mg, 59%). \( \delta_H (CDCl_3, 400 \text{ MHz}) 8.15 (2H, dd J 3.5, 6.0 \text{ Hz}), 7.6 (2H, dd J 3.5, 6.0 \text{ Hz}), 7.12 (2H, s), 4.32 (2H, t, J 5.6 \text{ Hz}), 3.61 (4H, t, J 6.5 \text{ Hz}), 3.37 (4H, t, J 5.6 \text{ Hz}), 3.28 (12H, s, O(CH_3)_2), 2.53 (4H, t, J 6.5 \text{ Hz}); \delta_C (CDCl_3, 100 \text{ MHz}) 182.2, 170.8,
145.6, 134.2, 132.1, 126.0, 123.3, 110.3, 102.8, 54.2, 41.2, 36.5, 39.1; m/z (ES+) 557 (M+H, 100).

\[N,N'-(2,2'-(9,10-Dioxo-9,10-dihydroanthracene-1,4-
diyl)bis(azanediyl)bis(ethane-2,1-diyl))bis(3,3-dimethoxypropanamide)\]

To a stirred solution of 1,4-difluoro-anthraquinone (80 mg, 0.33 mmol) in pyridine (1 mL) was added \(N\)-(2-aminoethyl)-3,3-dimethoxypropanamide (346 mg, 1.96 mmol) and the resulting solution was stirred at 70 °C for 4 h. The solvent was evaporated under reduced pressure and the reaction mixture was purified by flash chromatography using gradient eluent (CH\(_3\)OH:CH\(_2\)Cl\(_2\), 1:99→1:9). After the chromatographic purification residual acidic aliphatic impurities were removed by washing with 0.1 M NaOH (2 x 10mL), and brine (1 x 10mL) before being dried over MgSO\(_4\). After filtration, the solvent was concentrated in vacuo to yield the title compound as a dark blue powder (25 mg, 29%). \(\delta\) \(\text{H}\) (CDCl\(_3\), 400 MHz) 8.10 (2H, dd, \(J\) 3.5, 6.0 Hz), 7.57 (2H, dd, \(J\) 3.0, 5.5 Hz), 7.06 (1H, br s, NH), 6.98 (2H, s), 5.20 (2H, s), 4.69 (2H, t, \(J\) 5.6 Hz), 3.51 (4H, t, \(J\) 6.0 Hz), 3.25 (12H, s, O(CH\(_3\))\(_2\)), 2.60 (4H, d, \(J\) 5.6 Hz); \(\delta\) \(\text{C}\) (CDCl\(_3\), 100 MHz) 182.2, 169.9, 168.3, 134.2, 132.1, 126.0, 123.2, 109.9, 102.3, 54.3, 42.1, 40.9, 39.2; m/z (ES+) 557 (M+H, 100).
1,4-Bis(5,5-dimethoxypentylamino)-5,8-dihydroxyanthracene-9,10-dione

To a stirred solution of 1,4-difluoro-5,8-dihydroxyanthracene-9,10-dione (47 mg, 0.17 mmol) in pyridine (1 mL) was added 5,5-dimethoxypentan-1-amine (200 mg, 1.36 mmol) and the resulting solution was heated under reflux for 3 h. The reaction mixture was diluted with CH$_2$Cl$_2$ (20 mL) and washed with cold brine (2 x 10 mL). The organic solution was dried with MgSO$_4$, filtered and the remaining solution was concentrated in vacuo. The crude solid was purified by flash chromatography using gradient eluent (EtOAc:PE, 1:9→1:1) to yield the title compound as a dark blue powder (47 mg, 52%). $\delta$$_H$(CDCl$_3$, 400 MHz) 13.14 (2H, s), 10.28 (2H, br s), 6.99 (2H, s), 6.96 (2H, s), 4.33 (2H, t, $J$ 5.6 Hz), 3.27 (12H, s, O(CHO$_3$)$_2$), 1.73-1.44 (8H, m), 1.19 (4H, s); $\delta$$_C$(CDCl$_3$, 100 MHz) 183.8, 154.3, 145.4, 123.5, 122.7, 114.4, 107.6, 103.3, 51.8, 41.8, 28.3, 21.67, 13.1; m/z (ES+) 531 (M+H, 100).
3,3′-(5,8-dihydroxy-9,10-dioxo-9,10-dihydroanthracene-1,4diyl)bis (azanediyl)bis(N-(2,2-dimethoxyethyl)propanamide)

To a stirred solution of 1,4-difluoro-5,8-dihydroxyanthracene-9,10-dione (80 mg, 0.29 mmol) in pyridine (1 mL) was added 3-amino-N-(2,2-dimethoxyethyl) propanamide (306. mg, 1.74 mmol) and the resulting solution was heated under reflux for 3 h. The solvent was evaporated under reduced pressure and the reaction mixture was purified by flash chromatography using gradient eluent (CH$_3$OH:CH$_2$Cl$_2$, 1:99→1:9) to yield the title compound as a dark blue powder (112 mg, 66%). $\delta_H$ (DMSO, 400 MHz) 13.4 (2H, s, OH), 10.6 (2H, br s, NH), 8.15 (2H, br s, NH), 7.47 (2H, s), 7.11 (2H, s), 4.32 (2H, t, J 5.55 Hz), 3.73 (4H, q, J 5.5 Hz), 3.30 (12H, s, O(CH$_3$)$_2$), 3.25 (4H, t, J 5.5 Hz), 2.43 (4H, m); $\delta_C$ (DMSO, 100 MHz) 183.3, 170.2, 154.5, 146.8, 125.4, 124.3, 114.9, 107.2, 102.1, 53.2, 40.5, 38.9, 35.2; m/z (ES+) 589 (M+H, 100).
To a stirred solution of 1,4-difluoro-5,8-dihydroxyanthracene-9,10-dione (71 mg, 0.26 mmol) in pyridine (1 mL) was added \(N\)-(2-aminoethyl)-3,3-dimethoxypropanamide (272 mg, 1.54 mmol) and the resulting solution was heated under reflux for 4 h. The solvent was evaporated under reduced pressure and the reaction mixture was purified by flash chromatography using gradient eluent (\(\text{CH}_3\text{OH}:\text{CH}_2\text{Cl}_2\), 1:99→1:9) to yield the title compound as a dark blue powder (10 mg, 37%). \(\delta\) \(H\) (DMSO, 400 MHz) 13.5 (2H, s, OH), 10.7 (2H, br s, NH), 8.13 (2H, br s, NH), 7.64 (2H, s), 7.21 (2H, s), 4.72 (2H, t, \(J\) 5.5 Hz) 3.53 (4H, q, \(J\) 5.5 Hz), 3.35 (4H, t, \(J\) 5.5 Hz), 3.25 (12H, s, O(\(\text{CH}_3\))\(\text{C}_2\)); \(\delta\) \(C\) (DMSO, 100 MHz) 182.3, 172.2, 154.5, 140.9, 124.4, 124.3, 115.9, 108.5, 103.1, 53.7, 44.5, 38.9, 34.2; \(m/z\) (ES+) 589 (\(M+H\), 100).
II.3.2.4 Synthesis of non-symmetrical 1,4-di-substituted anthraquinones

1-(5,5-Dimethoxypentylamino)-4-(2-(dimethylamino)ethylamino)anthracene-9,10-dione

To a stirred solution of 1-(2-(dimethylamino)ethylamino)-4-fluoroanthracene-9,10-dione (60 mg, 0.19 mmol) in pyridine (1 mL) was added 5,5-dimethoxypentan-1-amine (113 mg, 0.77 mmol) and the resulting solution was stirred at 70 °C for 4 h. The solvent was evaporated under reduced pressure and the reaction mixture was purified by flash chromatography using gradient eluent (CH$_3$OH:CH$_2$Cl$_2$, 1:99→1:9) to yield the title compound as a dark blue powder (59 mg, 70%). $\delta$$_H$(CDCl$_3$, 400 MHz) 10.7 (2H, br s, NH), 8.3 (2H, dd, $J$ 4.0, 5.5 Hz), 7.6 (2H, dd, $J$ 4.0, 5.5 Hz), 7.3 (2H, dd, $J$ 6.0, 15.6 Hz), 4.32 (1H, t, $J$ 5.6 Hz), 3.53 (2H, t, $J$ 5.6 Hz), 3.33 (2H, t, $J$ 6.5 Hz), 3.26 (6H, s), 2.72 (2H, t, $J$ 6.5 Hz), 2.33 (6H, s), 1.7 (2H, m), 1.63 (2H, m), 1.42 (2H, m); $\delta$$_C$(CDCl$_3$, 100 MHz) 182.7, 182.4, 146.6, 145.6, 134.5, 134.4, 132.1, 132.0, 126.2, 126.0, 123.6, 123.5, 110.2, 109.8, 104.4, 58.4, 52.9, 45.4, 45.39, 42.8, 40.7, 32.3, 29.5, 22.3; m/z (ES+) 440 (M$^+$H, 100).
**N-(2,2-Dimethoxyethyl)-3-(4-(2-(dimethylamino)ethylamino)-9,10-dioxo-9,10-dihydroanthracen-1-ylamino)propanamide**

![Chemical Structure](image)

To a stirred solution of 1-(2-(dimethylamino)ethylamino)-4-fluoroanthracene-9,10-dione (60 mg, 0.19 mmol) in pyridine (1 mL) was added 3-amino-N-(2,2-dimethoxyethyl)propanamide (135 mg, 0.77 mmol) and the resulting solution was stirred at 70 °C for 6 h. The solvent was evaporated under reduced pressure and the reaction mixture was purified by flash chromatography using gradient eluent (CH$_3$OH:CH$_2$Cl$_2$, 1:99→6:94) to yield the title compound as a dark blue powder (56 mg, 70%). δ$_H$ (CDCl$_3$, 400 MHz) 10.6 (2H, br s, NH), 8.2 (2H, d, J 6.0 Hz), 7.5 (2H, d, J 5.5 Hz), 5.95 (1H, br s, NH), 4.37 (1H, t, J 5.3 Hz), 3.71, (2H, t, J 5.5 Hz), 3.5 (2H, t, J 5.5 Hz), 3.3 (2H, t, J 6.5 Hz), 3.25 (6H, s), 2.76 (2H, t, J 6.5 Hz), 2.54 (2H, t, J 5.5 Hz), 2.41 (6H, s); δ$_C$ (CDCl$_3$, 100 MHz) 181.7, 181.7, 169.7, 144.7, 144.5, 133.3, 131.1, 125.1, 124.9, 122.5, 122.3, 109.2, 109.2, 101.5, 57.0, 53.4, 44.1, 40.0, 39.2, 38.1, 35.8; m/z (ES+) 469 (M+H, 100).
\textit{N-(2-(4-(2-(dimethylamino)ethylamino)-9,10-dioxo-9,10-dihydroanthracen-1-ylamino)ethyl)-3,3-dimethoxypropanamide}

To a stirred solution of 1-(2-(dimethylamino)ethylamino)-4-fluoroanthracene-9,10-dione (50 mg, 0.16 mmol) in pyridine (1 mL) was added \textit{N-(2-aminoethyl)-3,3-dimethoxypropanamide} (112 mg, 0.64 mmol) and the resulting solution was heated under reflux for 4 h. The solvent was evaporated under reduced pressure and the reaction mixture was purified by flash chromatography using gradient eluent (\(\text{CH}_3\text{OH}:\text{CH}_2\text{Cl}_2\), 1:99→6:94) to yield the title compound as a dark blue powder (45 mg, 60%). \(\delta_H\) (\(\text{CDCl}_3\), 400 MHz); 10.7 (1H, br s, NH), 10.6 (1H, br s, NH), 8.2 (2H, dq, \(J\ 4.5, 9.0\ Hz\)), 7.64 (2H, dd, \(J\ 2.0, 4.0\ Hz\)), 7.2 (2H, dd, \(J\ 9.6, 25.2\ Hz\)), 6.71 (1H, NH br s), 4.6 (1H, t, \(J\ 5.5\ Hz\)), 3.5 (2H, t, \(J\ 5.5\ Hz\)), 3.2 (6H, s), 2.73 (2H, t, \(J\ 6.5\ Hz\)), 2.59 (2H, t, \(J\ 6.5\ Hz\)), 2.43 (2H, t, \(J\ 5.5\ Hz\)), 2.37 (6H, s); \(\delta_C\) (\(\text{CDCl}_3\), 100 MHz) 182.7, 182.4, 169.8, 169.2, 169.1, 146.0, 145.9, 134.4, 134.3, 132.2, 126.0, 125.9, 123.4, 110.0, 102.3, 102.2, 101.9, 58.3, 54.1, 51.4, 45.4, 41.9, 40.7, 39.4, 37.0; \(m/z\) (ES+) 469 (\(M+H\), 100).
1-(5,5-Dimethoxypentylamino)-4-(2-(dimethylamino)ethylamino)-5,8-dihydroxyanthracene-9,10-dione

To a stirred solution of 1-(2-(dimethylamino)ethylamino)-4-fluoro-5,8-dihydroxyanthracene-9,10-dione (60 mg, 0.17 mmol) in pyridine (1 mL) was added 5,5-dimethoxypentan-1-amine (102.61 mg, 0.7 mmol) and the resulting solution was stirred at 70 °C for 6 h. The solvent was evaporated under reduced pressure and the reaction mixture was purified by flash chromatography using gradient eluent (CH$_3$OH:CH$_2$Cl$_2$, 1:99→6:94) to yield the title compound as a dark blue powder (40 mg, 49%). $\delta_{H}$ (CDCl$_3$, 400 MHz) 13.3 (2H, brs, OH), 10.3 (2H, br s, NH), 7.16-6.96 (4H, m), 4.32 (1H, t, $J$ 5.5 Hz), 3.45 (2H, t, $J$ 6.5 Hz), 3.21 (6H, s), 2.74 (2H, t, $J$ 6.5 Hz), 2.45 (6H, s), 1.74-1.42 (6H, m); $\delta_{C}$ (CDCl$_3$, 100 MHz) 185.6, 185.2, 155.5, 155.46, 146.6, 145.9, 125.0, 124.8, 123.9, 123.7, 115.3, 115.2, 109.4, 108.8, 104.4, 57.7, 52.9, 45.0, 42.9, 40.2, 32.2, 29.4, 22.2; m/z (ES+) 472 (M+H, 100).
To a stirred solution of 1-(2-(dimethylamino)ethylamino)-4-fluoro-5,8-dihydroxyanthracene-9,10-dione (60 mg, 0.17 mmol) in pyridine (1 mL) was added 3-amino-N-(2,2-dimethoxyethyl)propanamide (123 mg, 0.7 mmol) and the resulting solution was stirred at 70 °C for 6 h. The solvent was evaporated under reduced pressure and the reaction mixture was purified by flash chromatography using gradient eluent (CH$_3$OH:CH$_2$Cl$_2$, 1:99→6:94) to yield the title compound as a dark blue powder (79 mg, 90%). δ$_H$ (CDCl$_3$, 400 MHz) 13.3 (2H, brs, OH), 10.1 (2H, br s, NH), 6.93 (2H, d, J 6.0 Hz), 6.82 (2H, t, J 6.0 Hz), 6.23 (1H, br s, NH), 4.37 (1H, t, J 5.7 Hz), 3.45 (2H, t, J 6.5 Hz), 3.41 (2H, t, J 5.2 Hz), 3.31 (6H, s), 2.64 (2H, t, J 6.5 Hz), 2.52 (2H, t, J 5.2 Hz), 2.33 (6H, s); δ$_C$ (CDCl$_3$, 100 MHz) 184.9, 184.8, 170.6, 155.3, 155.2, 146.1, 145.9, 124.7, 124.5, 123.6, 123.3, 115.2, 115.1, 108.9, 108.7, 102.5, 58.1, 54.3, 45.5, 41.1, 40.8, 39.2, 36.5; m/z (ES+) 501 (M+H, 100).
To a stirred solution of 1-(2-(dimethylamino)ethylamino)-4-fluoro-5,8-dihydroxyanthracene-9,10-dione (54 mg, 0.16 mmol) in pyridine (1 mL) was added \( N \)-(2-aminoethyl)-3,3-dimethoxypropanamide (111 mg, 0.63 mmol) and the resulting solution was stirred at 70 °C for 6 h. The solvent was evaporated under reduced pressure and the reaction mixture was purified by flash chromatography using gradient eluent (\( CH_3OH:CH_2Cl_2 \), 1:99→5:95) to yield the title compound as a dark blue powder (45 mg, 60%). \( \delta_H \) (CDCl\(_3\), 400 MHz) 13.3 (2H, br s, OH), 10.2 (2H, br s, NH), 6.92 (2H, t, \( J \) 5.9 Hz), 6.82 (2H, t, \( J \) 5.9 Hz), 6.2 (1H, br s, NH), 4.62 (1H, t, \( J \) 5.5 Hz), 3.47 (2H, q, \( J \) 6.4 Hz), 3.41 (2H, t, \( J \) 5.5 Hz), 3.33 (6H, s), 2.63-2.54 (4H, m), 2.52 (2H, t, \( J \) 5.5 Hz), 2.27 (6H, s); \( \delta_C \) (CDCl\(_3\), 100 MHz) 184.8, 184.5, 169.8, 169.2, 155.2, 155.1, 146.2, 124.5, 124.2, 123.6, 123.4, 115.2, 115.1, 108.8, 108.5, 102.2, 58.1, 51.4, 45.5, 40.8, 40.8, 37.0, 36.8; \( m/z \) (ES+) 501 (\( M+H \), 100).
Chapter III

Chemical hydrolysis and enzymatic metabolism of novel 1,4-di-substituted acetalanthraquinones
III.1 Introduction and Aims

Acetal groups are often employed in organic chemistry to protect alcohols (Greene T. W., 1999), but their stability and susceptibility to undergo enzymatic metabolism is not well researched. Acetal groups can also be used to protect aldehyde functionalities, which can be unmasked when exposed to acidic conditions (Greene T. W., 1999). In this context, it is important to understand whether the 1,4-di-substituted acetalanthraquinones (AAQs) synthesised in Chapter II are prone to undergo acidic hydrolysis and generate corresponding aldehyde substrates suitable for ALDH mediated oxidation (see Figure III.1).

![Proposed mechanism for acetal conversion](image)

**Figure 3.1** Proposed mechanism for acetal conversion. ALDH mediated transformation of the acetal into the active aldehyde via the hemiacetal intermediate.
Furthermore, the acetal moiety located at the terminal position of the aliphatic sidechains may be a target for oxidative metabolism by ALDH per se; hence it is important to also investigate their metabolic stability.

As outlined in Figure 3.1 the mechanism is proposed to occur via the formation of a hemiacetal species as an instable intermediate, which rapidly converts into the aldehyde. The development of the AAQs examined in this thesis is based on the principle that after exposure to acidic hydrolysis they may be metabolically transformed into carboxylic acid via ALDH reaction. As stated above, the acetal group can be chemically hydrolysed under acidic conditions, but its susceptibility to undergo this transformation may be subject to functional groups in neighbouring α-δ positions. Accordingly, this chapter is focussed on the investigations concerning the potential selectivity of ALDH towards aldehydes lined to anthraquinones using cell extracts and recombinant proteins as sources of enzyme. Therefore, the specific aims of this chapter are:

1. To assess the acidic stability of the novel AAQs to HCl, in order to identify compounds that can chemically generate an aldehyde moiety;

2. To investigate LC-111 properties as prototype AAQ to understand the metabolic profile generated by acidic hydrolysis and compare the acid hydrolysis of LC- and HA-libraries and to assess the requirements for the hydrolysis step to achieve ALDH selectivity;

3. To determine the metabolic fate of selected LC-AAQs before and after chemical hydrolysis in several ALDH sources and to isolate the potential metabolic products for LC-MS identification.
Chapter III

III.2 Materials and Methods

III.2.1 Chemical hydrolysis of AAQs

III.2.1.1 Chemicals and reagents

All the tested compounds were stored in dry powder form at -20°C until required for investigation. DMSO, methanol, trifluoroacetic acid (TFA) and HCl were all supplied by Sigma. Analytical HPLC-grade acetonitrile and deionised water were supplied by Fisher Scientific. Ammonium formate and formic acid were purchased from VWR international Ltd. Plastic consumables and 96-flat bottomed well plates were purchased from Corning.

III.2.1.2 Chemical hydrolysis of AAQs

The acidic hydrolysis was performed using the same conditions for all the AAQs (see section II.4 for name and structures). The investigational compound was dissolved in DMSO to 10 mM final concentration. Working solutions were prepared by diluting the compound to 100 µM in deionised water to a final volume of 100 µL and the mixture was vortexed for 2 min. Finally, 100 µL of 2 M HCl was added and the mixture was heated in a heating block at 37 °C for 30 min. Specifically for LC-111, fractions were collected at 15 min and thereafter every 30 min up to 2.5 hours total incubation time. Immediately after collection, the fractions were freeze dried at -56 °C until dry powder was obtained (typically 18h) using Alpha 1-2 LQ lyophiliser (Martin Christ). Compounds were used for LC/MS analysis.
III.2.1.3 Acidic stability assessment via LC/MS

Conditions for AAQ analysis using methanol as the organic solvent of the mobile phase:

The samples were analysed by HPLC (High Pressure Liquid Chromatography) in gradient elution mode. Mobile phase A (MPA) consisted of 10% methanol and 90% deionised water containing 0.05% TFA. Mobile phase B (MPB) consisted of 90% methanol and 10% deionised water containing 0.05% TFA. Analysis was performed on a Waters C8 10 cm Acquity column 1.7 µm (10 cm x 2.1 mm) (Waters Ltd.) with a flow rate of 0.3 mL/min. Pumps were run with a gradient of 20% MPB and 80% MPA to 70% MPB and 30% MPA. The total run time was 40 min. The flow rate was maintained and 2 µL of sample was injected using a Waters Acquity Separation Module (Waters Ltd.). Partial loop injections were made and the loop size was 10 µL. Detection was performed using a Waters Quattro Premier MS/MS (Waters Ltd.) in parallel with diode array UV-Vis absorbance detection.

Conditions for AAQ analysis using acetonitrile as organic solvent of the mobile phase:

The samples were analysed by HPLC in gradient elution mode. MPA consisted of 10% acetonitrile and 90% deionised water containing 0.05% TFA. MPB consisted of 90% acetonitrile and 10% deionised water containing 0.05% TFA. Analysis was performed on a Waters C8 10 cm Acquity column 1.7 µm (10 cm x 2.1 mm) (Waters Ltd.) with a flow rate of 0.3 mL/min. Pumps were run with a gradient of 20% MPB and 80% MPA to 70% MPB and 30% MPA. The total run
time was 40 min. The flow rate was maintained and 2 µL of sample was injected using a Waters Acquity Separation Module (Waters Ltd.). Partial loop injections were made and the loop size was 10 µL. Detection was performed using a Waters Quattro Premier MS/MS (Waters Ltd.) in parallel with diode array UV-Vis absorbance detection.

**Mass spectrometry conditions:**

A MRM (multiple reaction monitoring) channel was extracted in electrospray positive (ESI+) mode at 600.0>58.1 m/z. Capillary voltage was set to 3.05 kV, cone 29 V, extraction 3 V, source temperature 150 °C, desolvation temperature 250 °C, collision energy 62 V and collision gas flow 0.35 mL/min.

**III.2.2 Metabolism of AAQs**

**III.2.2.1 Chemicals and reagents**

All the reagents and cofactors were supplied by Sigma unless otherwise stated. Human recombinant ALDH1A1 was purchased from R&D Systems. Plastic consumables were purchased from Corning.

**III.2.2.2 Isolation of cytosolic fraction from rat liver homogenate**

Two adult male Sprague Dawley rats were sacrificed and livers were isolated, rinsed twice with PBS to clean the blood residues and cut into small slices. PBS was added in 2x 50 mL falcon tubes and the tissue slices were transferred using sterile forceps. The mixture was homogenised with a T25 Ultra-Turrax basic homogeniser (IKA) and transferred to ultracentrifuge tubes. The liver
homogenate was centrifuged to produce S9 fraction at 9,000 g for 30 min at 4 °C in a Beckman Ultracentrifuge. The supernatant was transferred in fresh ultracentrifuge tubes and centrifuged at 105,000 g for 70 min at 4 °C. The white lipid layer on top of the supernatant was removed and the supernatant containing the cytosolic fraction was aliquoted in cryovials (Thermo). 10% of glycerol was added to each vial and the samples kept at -80 °C till required. Prior to experimental use, protein quantification was assessed using the Bradford assay (Spector, 1978).

**III.2.2.3 Metabolism of AAQs in cytosolic fraction of rat liver homogenate**

The incubations comprised of 0.1 mM pyrazole, 5 mM glutathione and 4 mM NAD⁺ as co-factors in 1 mL PBS-EDTA pH 7.4. All the co-factor concentrations were optimised to achieve velocity rate comparable to reported literature (Schomburg D, 2004). Cytosolic fraction was added at a final concentration of 1 mg/mL. Control substrates for investigations in these experiments were propionaldehyde and acetaldehyde at concentration range between 100 and 5 µM. Diethylaminobenzaldehyde (DEAB), an ALDH inhibitor (Ma et al., 2010), was used as the negative control. Bromoacetal, LC-111, and its hydrolysed product mixture (after 30 min hydrolysis, see section III.2.1.2) were tested at concentration range between 1 and 100 µM.

Two identical quartz cuvettes (Thermo Fisher Scientific) were set up for each experiment: one cuvette containing the complete incubate as described above with the selected substrate and the other excluding NAD⁺ as control blank. Both incubation mixture were prepared initially without NAD⁺ and pre-incubated at 37 °C for 2 min. The metabolism was then started by addition of NAD⁺ to the
appropriate cuvette. Determination of NAD$^+$ to NADH conversion was assessed spectrophotometrically in a temperature controlled Cary 400-Bio UV-Vis system (Varian), using the “scan” function to read the absorbance spectrum between 200 and 800 nm for 60 min. The enzymatic activity was assessed at 37 °C by reduction of the NAD$^+$ peak at 260 nm and formation of NADH peak at 340 nm. Furthermore, reaction velocity (ΔA/min) was calculated in the “enzyme kinetics” mode by setting the absorbance wavelength at 340 and monitoring the reaction for 30 min.

III.2.2.4 Metabolism of AAQs in yeast ALDH

Lyophilised ALDH enzyme from baker’s yeast (S. Cerevisiae) (Sigma) was dissolved in 50 mM Tris-HCl pH 8.0 with 0.02% BSA. As described in the literature (Wang et al., 2009), the reaction mixture contained 4U/mL of enzyme, 3 mM NAD$^+$, 1 mM EDTA, 200 mM KCl, 1.7% (v/v) acetonitrile (as a solvent carrier for aldehydes), 100 mM mercaptoethanol and varied concentrations of substrates (10-0.01 mM) (Bostian and Betts, 1978; Wang et al., 2009). Propionaldehyde was used as a positive substrate control. Substrate concentrations and pH were optimised (as showed in section III.3.4.2) to a total reaction volume of 100 µL. The reaction mixtures were prepared in a flat-bottomed 96 well plate and incubated at 37 °C for 2 min, before adding NAD$^+$ and starting to measure the kinetic rate of NADH formation. The plate was read in kinetic mode at 340 nm in a Multiskan Spectrum plate reader (Thermo) and temperature was maintained at 37 °C.
III.2.2.5 Metabolism of AAQs in human recombinant ALDH1A1

Recombinant human ALDH1A1 (rh ALDH1A1) (Abnova) was dissolved in 20 µg/mL of deionised water. The reaction mixture was prepared as follows according to the manufacturer protocol: 100 µL 1 M Tris pH 8.5, 50 µL 2 M KCl, 50 µL 20 mM NAD$^+$, 50 µL 40 mM DTT, 200 µL deionised water and the selected substrate at the appropriate concentration [propionaldehyde (10 mM), LC-111 AAQ and LC-111 (HAAQ) (0.1 µM)]. The HAAQ was generated after 30 min hydrolysis (see section III.2.1.2). 50 µL of the reaction mixture was combined with the same volume of enzyme mixture (20 µg/mL final enzyme concentration) directly in the 96 flat-bottomed well plate. A blank containing 50 µL deionised water (instead of the enzyme mixture) and 50 µL of reaction mixture was also added to the plate. The reaction was monitored for 20 min at 340 nm and RT using the kinetic mode in a Multiskan Spectrum plate reader (Thermo).

III.2.2.6 Evaluation of specific activity for enzyme metabolism

The concentration of NADH produced by the reaction was calculated using the Beer-Lambert law (or Beer’s law), which considers the linear relationship between absorbance and concentration of an absorbing species. The general Beer-Lambert law is usually written as: $A = \varepsilon cl$, where $A$ is the measured absorbance, $\varepsilon$ is a wavelength-dependent absorptivity coefficient ($\varepsilon_{340\text{ nm}} = 6,220 \text{ M}^{-1}\text{cm}^{-1}$ for NADH), $l$ is the cell-path length (1 cm in this experimental condition), and $c$ is the analyte concentration.
Specific activity was used to measure the rate of the enzymatic reaction. This term was defined as the amount of substrate the enzyme converts (reaction catalyzed), per mg protein in the enzyme preparation, per unit of time. The specific activity was expressed as units of activity/amount of enzyme and it was calculated with the following formula:

\[
\text{Specific Activity} = \frac{\Delta \text{Abs/min}}{(\text{nmol/min/mg})} \times 10^6
\]

Where \(\Delta \text{Abs/min}\) is the amount of activity, \(\varepsilon_{340\,\text{nm}}\) is the extinction coefficient of NADH (6,220 \(\text{M}^{-1}\text{cm}^{-1}\)) and \(P\) is the amount of enzyme in the reaction (mg).

### III.2.2.7 AAQ metabolite identification by LC/MS

Metabolism was carried out for LCw111 and its hydrolysed product (HAAQ) after 30 min hydrolysis as described in section III.2.4.5. The final concentration of the incubation was: 100 mM Tris pH 8.5, 200 mM KCl, 4 mM NAD\(^+\), 2 mM DTT, 20 \(\mu\text{g/mL}\) rhALDH1A1 and 1 \(\mu\text{M}\) LC-111 or the corresponding HAAQ with a total assay volume of 100 \(\mu\text{L}\) in deionised water.

After the metabolism was monitored spectrophotometrically, protein was precipitated out from the reaction mixture. Mixture was centrifuged at 13,000 rpm for 15 min to sediment the enzyme before the supernatant was injected into the LC/MS system. The mobile phase composition containing acetonitrile and mass spectrometry settings was as described in section III.2.1.3.
III.3 Results

III.3.1 Chemical hydrolysis of AAQs

According to the HPLC analysis (see Figure 3.2), parental AAQs were completely hydrolysed after 60 min of exposure to 2 M HCl. Using HPLC conditions including methanol in the mobile phase composition (see section III.2.1.3), a sharp peak was identified for all the analysed AAQs and retention time values varied according to the chemical features (refers to Figure 3.2A).

In this chapter, LC-111 is used as an exemplar of typical hydrolysis formation (for HPLC spectrum of LC-110, see Supplementary Information III.A.1; for LC-107 HPLC spectrum see III.B).

Analysis of LC-111 revealed one major peak, which was detected at 260 nm with a retention time of 13.13 min (see Figure 3.3). After 30 min of acidic hydrolysis, total conversion of parental AAQ was detected with appearance of a new broad peak visible at 260 nm with a retention time of 9.40 min. The same peak was also observed after 60, 180 and 210 min of exposure to this acidic environment. However, after 60 min of acidic treatment, two secondary peaks were detected at 260 nm at retention times 12.53 and 13.57 min respectively. The intensity of these secondary peaks showed an enhancement after 60 min of acidic treatment, suggesting that exposure to these acidic conditions generate other unwanted by-products with time. The amount of hydrolysed derivative formed after 60 min was 100% of the parental AAQ compound, as measured by disappearance of LC-111 at 260 nm.
Figure 3.2 HPLC spectra of hydrolysis profile of LC-111. Spectra of hydrolysed compounds were collected after incubation at 37°C with 2 M HCl. Analysis of the (C) LC-111 and the corresponding HAAQ products were performed at 260 nm after (B) 30 min and (A) 60 min hydrolysis.

A more detailed analysis of LC-111 was performed to evaluate the hydrolysis products at intermediate time-points. Acidic treatment was monitored via HPLC at 15 and 30 min. The determination was performed using the conditions reported in section III.2.1.3. As shown in Figure 3.3A, LC-111 showed a major peak at 15.8 min (red peak). After 15 min of exposure to these acid conditions, the intensity of the peak at 15.8 min was considerably decreased (see Figure 3.3B) and a new intense peak appeared with a retention time of 6.27 min.
Figure 3.3 HPLC analysis of LC-111 short term hydrolysis. Spectra were collected at 260 nm after treatment with 2 M HCl for (B) 15 min and (C) 30 min. (A) LC-111 (t = 0). The column used on this experiment was different from the one used in Figure 3.2, therefore the retention times were not maintained.
The same peak was confirmed after 30 min (see Figure 3.3C), but some minor peaks, respectively at 4.50 and 3.56 min, seemed to appear with protracted acidic exposure. The results suggest, as stated above for longer exposure times, that LC-111 require 30 min incubation with acid to achieve the maximum rate of hydrolysis conversion.

However, β-positioned secondary and tertiary amines (HA-AAQs, see Table II.3 for structures) revealed an increased resistance to the acidic hydrolysis of the acetal moiety (refers to Supplementary Information III.B.1-2 for HPLC spectra of HA-222 and HA-252). Here HA-218 is reported as an exemplar of the HA-AAQ series. Elution of HA-218 produced a very sharp peak with a retention time of 15.79 min (Figure 3.4A).
Indeed, after 210 min at 37 °C these compounds did not generate any hydrolysis products (data not showed). The treatment was repeated heating the mixture at 75 °C for 210 min and the LC-MS analysis showed only partial conversion of the acetal moiety. As shown in Figure 3.4B, HA-218 was converted into two peaks observed at 5.61 and 28.28 min respectively,
revealing that only 20-30% of the parental compound peak was converted into unidentified hydrolysis products.

**III.3.2 Identification of chemical hydrolysis products via mass spectrometry**

Tandem mass spectrometry (MS) analysis was used to identify the fragmentation pattern for the AAQs and hydrolysed species. In this section only LC-111 MS spectra are shown as exemplar of the hydrolysis formation (see Figure 3.5A-B), but, MS analysis of LC-110 can be found in Supplementary Information III.A.2.

As described in the section III.3.1, HPLC separation of LC-111 revealed a sharp peak at 260 nm with a retention time of 13.13 min (see Figure 3.2). This peak corresponded to a mass of 501 m/z, as shown in the MS in Figure 3.5A. After 30 min of acidic hydrolysis, total conversion of LC-111 was detected with the appearance of a new broad peak visible at 260 nm and with a retention time of 9.40 min. As shown in Figure 3.5B, two intense peaks were related to the broad peak at 9.40 min and MS analysis revealed masses of 487 and 455 m/z, which were correlated with hemiacetal and aldehyde formation respectively. Further analysis was performed on LC-111 in order to obtain a MS spectrum with a unique aldehyde fragmentation peak; however, no separation was achieved. The results suggest that either (i) co-existence of the two species is real or (ii) is an artefact which cannot be resolved using LC-MS due to instability of the aldehyde product under the mobile phase conditions employed.
III.3.3 Sensitivity of hydrolysis products to mobile phase composition

In order to explore whether the hydrolysis products were stable in the methanol/water system used for HPLC chromatography, a further elution system was tested in LC-111 to perform the separation. Methanol/water was
replaced by acetonitrile/water on the basis of the possible implication of methanol in reversing the aldehyde formation, generating the hemiacetal intermediate of acid hydrolysis (see Figure 3.1 for proposed mechanism). As enclosed in the spectra in Supplementary Informations III.C1-2 at the end of the present chapter, no difference was observed on the LC/MS analysis of LC-111. The retention time of LC-111 and the corresponding HAAQ were maintained almost identical in both systems. Moreover, the MS analysis of the eluted peaks in the methanol/water and acetonitrile/water systems revealed the same mass for AAQ (500.87 and 500.79) and the corresponding HAAQ derivatives (454.88 and 454.91) respectively, suggesting that these systems can be interchanged without alteration of the separation outcome. This suggests that the co-existence of hemiacetal and aldehyde is present after acidic hydrolysis and exclude any potential artefacts as explained in section III.1.

III.3.4 Metabolism studies

III.3.4.1 Metabolism studies using cytosolic fraction from rat liver homogenate

The cytosolic fraction from rat liver homogenate was used in this study to evaluate the possible metabolism of AAQs and their corresponding HAAQ products in a multi-enzymatic complex containing several ALDH isoforms (Tottmar et al., 1973). In this study, the cytosolic fraction was isolated from rat liver and an assay was designed and optimised to evaluate whether NAD\(^+\)-dependent conversion of the novel analogues could be detected by a
spectrophotometric analysis. Microsomal fraction was removed via differential centrifugation (see section III.2.2.2) to avoid any interference from metabolism by cytochrome P450s.

The assay was optimised by adding pyrazole to inhibit alcohol dehydrogenase activity (Moon et al., 2005) and glutathione to optimise the protein folding of the ALDHs (Chakravarthi et al., 2006).

![Figure 3.6 Metabolism of propionaldehyde (10 µM) in rat liver cytosol. NADH production was measured at 340 nm as described in section III.2.2.3.](image)

Several substrates were tested at different concentrations. Propionaldehyde (1-100 µM) was used as positive control (Shum and Blair, 1972) and bromoacetal was tested to investigate whether the cytosolic fraction was able to metabolise the acetal moiety per se. Propionaldehyde was a good substrate and the reaction velocity using 10 µM propionaldehyde is presented in Figure 3.6A. An enzyme blank, containing the entire reaction component except for the enzyme, was coupled to each reaction measurement showing no detectable activity (see
Figure 3.6B), suggesting that none of the components of the mixture were interfering with the determination.

Analysis of the full spectrum (200-800 nm) also showed the detection of a characteristic peak at 460 nm (see Figure 3.7) due to the presence of bilirubin, which is a breakdown product of haemoglobin in red cells of rat liver.

![UV/Vis spectrum of NAD\(^+\) supplemented rat liver cytosol in presence of propionaldehyde.](image)

**Figure 3.7** UV/Vis spectrum of NAD\(^+\) supplemented rat liver cytosol in presence of propionaldehyde.

Diethylaminobenzaldehyde (DEAB), a known inhibitor of ALDH activity (Chute et al., 2006), was added to the reaction mixture to establish whether the conversion was ALDH-dependent or whether other cytosolic enzymes such as glutamate dehydrogenase, monoamine oxidase or xanthine oxidase (Tottmar et
al., 1973) were responsible for the NAD$^+$ to NADH formation. Interestingly, addition of DEAB with or without propionaldehyde led to a rapid increase of enzyme activity within the first 2-3 min of incubation and a subsequent steady state was reached (see Supplementary Information III.D for kinetic plots).

Furthermore, bromoacetal was tested to verify the presence of cytosolic enzymes capable of converting the acetal via NAD$^+$ consumption. As reported in Figure 3.8, activity was measured with 100 µM bromoacetal with drop in activity at lower concentrations, suggesting the presence of enzymes in the cytosolic fraction capable of reacting with the acetal moiety in a NAD$^+$-dependent mode.

![Figure 3.8 Kinetic analyses of several substrates (10-100 µM) in the cytosolic fraction of rat liver homogenate. Every determination was coupled with an enzyme blank reading. The enzymatic rate (nmol/mg/min) was calculated using Beer-Lambert law (see section III.2.2.6). The data are expressed as the mean of three independent experiments ±SD.](image-url)
The same experiment was repeated to investigate the potential metabolism of LC-111 and its corresponding HAAQ. Intriguingly, greater activity was reported with the parental AAQ compared to its HAAQ product (see Figures 3.8-3.9).

As reported in Figure 3.8, specific activity values of bromoacetal and LC-111 varied independently from the substrate concentration. The finding that both LC-111 and its corresponding HAAQ product revealed enzymatic activity suggested the possibility of side-reactions occurring that may not be related to ALDH metabolism. As a consequence of this complex metabolism, the use of cytosol fractions from rat liver homogenate was deemed not suitable as a source to study the metabolism of novel AAQs and HAAQs. Therefore, other sources of ALDHs were explored.

![Figure 3.9 Kinetic analysis of LC-111 and its corresponding HAAQ using cytosolic fraction of rat liver homogenate.](image)

Reaction was measured at 340 nm with substrate concentration of 10 \( \mu \)M. Every determination was coupled with an enzyme blank reading.
III.3.4.2 Yeast ALDH metabolism

Yeast ALDH was used as its active region is 95-100% homologous to that of the mammalian ALDHs. Veverka et al. (1997) also used yeast ALDH for studies related to the mechanism of ALDH inhibition by disulfiram (Maninang et al., 2009). As previously described (Wang et al., 2009), yeast ALDH has been shown to oxidise preferentially small aliphatic aldehydes such as acetaldehyde and propionaldehyde. In this study, propionaldehyde was mainly used as control because of the limitations of handling acetaldehyde (volatile liquid). In addition, propionaldehyde has also been shown as the preferred substrate (Wang et al., 2009).

Figure 3.10 Optimisation of yeast ALDH kinetic analysis. Reaction was assessed using acetaldehyde and propionaldehyde at several concentrations in the presence of mercaptoethanol. Reaction was typically run at pH 8.0 although pH 8.5 was also evaluated.
As reported in Figure 3.10, the ALDH enzymatic reaction was initially tested at several pH conditions, but pH 8.0 was found to be the optimal option for both propionaldehyde and acetaldehyde. Furthermore, the role of mercaptoethanol, to stabilise the enzyme conformation on its reduced state, was revealed to be fundamental for obtaining a greater velocity rate. The reaction rate using mercaptoethanol and propionaldehyde as substrate appeared to reach the best rate at 10 mM of substrate concentration (Figure 3.10, green line). The reaction velocity decreased in the absence of mercaptoethanol (Figure 3.10, purple line).

As Maninang’s group reported (Maninang et al., 2009), the presence of 0.02% of BSA improved the ALDH reaction rate, but only up to 5 mM substrate concentration (see Figure 3.11, blue line).

![Figure 3.11 Optimisation of yeast ALDH kinetic analysis using propionaldehyde. Several concentrations were tested (0-25 mM). BSA activity and reaction volume were added in order to evaluate the impact on the reaction rate.](image-url)
In this study, a concentration above 5 mM led to substrate inhibition, suggesting that addition of BSA reduced the optimal substrate concentration. Furthermore, the velocity of the reaction was reduced by scaling down the reaction volume to 50 µL (see Figure 3.11, red line). Although the shape of the kinetic was consistent, use of half of the original volume (100 µL) resulted in reduction of almost a third of the original rate at 5 mM substrate concentration (0.55 vs. 0.18 Abs/min). As reported in Table 3.1, lactones, isocyanate and acetal- derivative were not substrates for ALDH. Aromatic aldehydes, such as tolualdehyde or p-anisaldehyde, showed similar specific activity but catalytic efficiency of aliphatic aldehydes was greater than aromatic substrates. Interestingly, propionate ester gave a specific activity of 4.42 ±0.66 suggesting some esterase activity of the yeast ALDH, which is in agreement with a previous published report (Blackwell et al., 1983; Marchitti et al., 2008).

LC-111 and its corresponding HAAQ were tested using the optimised conditions reported in section III.2.4.4 at concentrations between 0.1-100 µM. Due to the deep blue colour of the AAQs, the final test volume could not contain substrate concentration higher than 100 µM. Several AAQs were tested in this ALDH yeast model, but no NADH formation was observed for either AAQ or its corresponding HAAQ.
<table>
<thead>
<tr>
<th>Name</th>
<th>Structure</th>
<th>rate ((\Delta\text{Abs/min}))</th>
<th>Specific Activity ±SD ((\text{nmol/min/mg protein}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Propionaldehyde</td>
<td><img src="image" alt="Structure" /></td>
<td>0.22</td>
<td>8.87±0.72</td>
</tr>
<tr>
<td>3-Isochromanone</td>
<td><img src="image" alt="Structure" /></td>
<td>no reaction</td>
<td>-</td>
</tr>
<tr>
<td>δ-Valerolactone</td>
<td><img src="image" alt="Structure" /></td>
<td>no reaction</td>
<td>-</td>
</tr>
<tr>
<td>Tolualdehyde</td>
<td><img src="image" alt="Structure" /></td>
<td>0.10</td>
<td>4.02±0.34</td>
</tr>
<tr>
<td>Ethyl-3-bromo-propionate</td>
<td><img src="image" alt="Structure" /></td>
<td>0.11</td>
<td>4.42±0.66</td>
</tr>
<tr>
<td>Ethyl-isocyanate</td>
<td><img src="image" alt="Structure" /></td>
<td>no reaction</td>
<td>-</td>
</tr>
<tr>
<td>Valeraldehyde</td>
<td><img src="image" alt="Structure" /></td>
<td>0.20</td>
<td>8.06±0.89</td>
</tr>
<tr>
<td>p-Anisaldehyde</td>
<td><img src="image" alt="Structure" /></td>
<td>0.14</td>
<td>5.63±0.28</td>
</tr>
<tr>
<td>Bromoacetal</td>
<td><img src="image" alt="Structure" /></td>
<td>no reaction</td>
<td>-</td>
</tr>
</tbody>
</table>

**Table 3.1 Kinetic analyses of several substrates using yeast ALDH as source of enzyme.**

Several concentrations were tested (1-100 µM). The kinetic rate was calculated as Abs/min and specific activity was generated at 10 µM using the Beer-Lambert law (see section III.2.2.6). Specific activity values were generated as average of three experiments ±SD.
III.3.4.3 Metabolism of AAQs using human recombinant ALDH1A1 and metabolites identification via LC-MS

Human recombinant ALDH1A1 (hrALDH1A1) kinetic assay was tested using propionaldehyde as control sample, as recommended by the R&D System rhALDH1A1 manufacturer protocol. LC-111 and its corresponding HAAQ were analysed in duplicate as test substrates at a final concentration of 1 µM. Metabolism was monitored measuring the NADH formation at 340 nm for 5 min., but the results as shown in Figure 3.12, revealed almost undetectable activity for LC-111 and its corresponding HAAQ product.

Figure 3.12 Kinetic profiles of selected substrates using hrALDH1A1. Propionaldehyde, LC-111 parental and the corresponding hydrolysed derivative were tested at 1 µM concentration. All the curves were blank corrected. Sp Act: specific activity, calculated as reported in section III.2.2.6.
Indeed, data showed that under the tested experimental conditions, propionaldehyde was a substrate for human ALDH but LC-111 compounds were poor substrates. After kinetic analysis, the hrALDH1A1 protein was precipitated and both LC-111 and HAAQ metabolites were analysed via LC/MS.
Figure 3.13 HPLC analysis of the metabolic profile of LC-111 following incubation with hrALDH1A1. Analysis was performed at 612 nm. (A) LC-111; (B) hydrolysed LC-111 (HAAQ) before incubation with hrALDH1A1; (C) hydrolysed LC-111 (HAAQ) product after metabolism. The peaks indicated (A, B, C, D) were analysed by MS.

Metabolism of LC-111 produced one major sharp peak at 612 nm (retention time of 13.69 min, see Figure 3.13A) whereas the HAAQ product (hydrolysed for 30 min) produced a single broad peak at 612 with a retention time of 11.76 min (see Figure 3.13B). After incubation of the hydrolysed LC-111 with hrALDH1A1 several peaks were obtained and a very complex mixture of potential metabolites was identified through MS analysis (see Figure 3.13C). Although, several peaks were observed with retention times between 8.0 and 17.0 min, in the present study MS analysis was performed only for most intense signals (see Figure 3.14A-D).
The expected fragmentation pattern of oxidation catalysis, as carboxylic acid, for hydrolysed LC-111 derivative was calculated to be 470.48 m/z. However, none of the produced fragments could be correlated to that expected carboxylic acid mass.

Figure 3.14 MS analysis of LC-111 hydrolysed derivative after hrALDH1A1 metabolism. Spectra A-D refers to the corresponding labelled peak from the HPLC analysis (see Figure 3.13C). The red dotted line represents the expected m/z value for the main carboxylic acid metabolite.
III.4 Discussion

In this study, the sensitivity of novel AAQs to acidic hydrolysis has been investigated. The investigated LC-AAQs were all prone to hydrolysis via hemiacetal formation (refers to Supplementary Information III.B for additional HPLC spectra). The presence of an amide bond in the sidechain structure had no effect on the acetal to aldehyde conversion (see Figure 3.15 for structures).

Figure 3.15 Overview of the structural modification on the AAQ sidechains. Acetal hydrolysis sensitivity was highly dependent on the secondary amine substitution.

Symmetrical and non-symmetrical compounds showed a similar hydrolysis profile. A more detailed stability test was performed on the non-symmetrical derivatives, because of the ease of monitoring conversion of a mono-acetal sidechain. Isolation and identification of the hydrolysis product revealed
production of an aldehyde moiety, as a unique product after 30 min hydrolysis. The HA-library of compounds showed more resistance to acidic hydrolysis than the LC-AAQs, suggesting a protective role of the neighbouring β-positioned tertiary amine with different substituents (such as hydrogen, methyl, and benzyl). However, prolonged acid exposure (up to 60 min) at higher temperatures (75 °C) allowed aldehyde conversion on the HA-derivatives as demonstrated by HPLC analysis. Therefore, our findings suggested that the acetal moiety was a very efficient masking group to deliver an active aldehyde as single product of the hydrolysis reaction; however, as demonstrated for HA-AAQs, the sidechain skeleton plays a crucial role on the sensitivity to acidic hydrolysis, revealing that β-positioned tertiary amine induces resistance to hydrolysis even at high temperature.

We then investigated the potential metabolism of the active aldehyde using several sources of ALDH. Firstly, a preliminary screening was assessed using cytosolic fraction from rat liver as a rich source of target enzyme. Indeed, numerous studies have indicated the presence of several isoforms of hepatic ALDHs in mammals (Deitrich et al., 1975; Shum and Blair, 1972; Tottmar et al., 1973). For example, Shum et coll. demonstrated (Shum and Blair, 1972) that at least two NAD⁺-dependent ALDH isoenzymes were characterised and partially purified from rat liver cytosol. Furthermore, Lindahl et al. extensively investigated rat liver ALDHs (Lindahl, 1980; Lindahl and Evces, 1984, 1987), providing evidence for multiple molecular forms in microsomal and cytosolic fractions.
We analysed spectrophotometrically ALDH-dependent NADH production. Indeed, both NAD$^+$ and NADH strongly absorbed near UV-light due to the aromaticity of the adenine moiety. Specifically, peak absorption of NAD$^+$ was at a wavelength of 260 nm, with an extinction coefficient of 16,900 M$^{-1}$cm$^{-1}$. NADH also absorbed at higher wavelengths, with a second peak detected by UV absorption at 340 nm and with an extinction coefficient of 6,220 M$^{-1}$cm$^{-1}$ (Dawson, 1985). The amount of NADH produced was considered as proportional to the oxidation rate of aldehyde into its respective carboxylic acid (see Figure 3.16)

![Diagram of LC-111 hydrolysed (aldehyde) and Principal metabolite (carboxylic acid)](image)

**Figure 3.16 Proposed metabolism of hydrolysed LC-111 derivative.** The aldehyde was converted into a correspondent carboxylic acid (molecular weight: 470.48) via NAD-dependent ALDH catalysis.

Several compounds were tested using this enzymatic source but no conclusive results were obtained because metabolism was observed with aldehydes as well as with acetal substrates. The results suggest the potential co-existence of other enzymatic entities within the rat liver cytosol that are able to react with acetal groups in a NAD-dependent manner. Indeed, it was presumed that
metabolites were the product of sidechain oxidation, since the anthraquinone chromophore has been shown to be refractory to oxidation (Smyth et al., 1986; Wolf et al., 1986).

From these initial studies, it was therefore of interest to explore a different enzymatic model to determine whether ALDH could be involved in the metabolism of the novel AAQs. Hence, further investigations were attempted using yeast K⁺-activated ALDH, which was selected as a model protein because its catalytic site is 95-100% homologous to that of mammalian ALDHs (Saigal et al., 1991; Devaraj et al., 1997). The validity of yeast ALDH has been demonstrated in studies related to the mechanism of ALDH-inhibition by disulfiram, where high similarity with mammalian proteins was revealed (Veverka et al., 1997). Yeast ALDH, designated mitochondrial ALDH2 according to Wang’s group (Wang et al., 1998), demonstrated high catalytic efficiency with small aliphatic and aromatic aldehydes, such as propionaldehyde, valeraldehyde, tolualdehyde and p-anisaldehyde (see Table 3.1 for structures and kinetics). Interestingly, this study demonstrated esterase activity by yeast ALDH, as proven with the metabolism of ethyl-propionate ester. Ester metabolism by ALDHs is not an unknown phenomenon as some ALDHs have been shown to catalyse ester hydrolysis (Marchitti et al., 2008). Human ALDH2, in particularly, has been demonstrated to also possess the ability to hydrolyse p-nitrophenyl esters (Sidhu and Blair, 1975).

After optimisation of the reaction conditions, a kinetic analysis was performed using LC-111. Unfortunately, no metabolism of LC-111 or its corresponding HAAQ product was observed in a range of concentrations (0.1-100 µM). In vitro
metabolism of aldehydes using this yeast model system was not successful, maybe be due to intense blue colour of the anthraquinone, which prevent high enough concentrations to be used for metabolism to occur. Certainly, the yeast model system has been used in previous studies with demonstration of both qualitative and quantitative differences in substrate specificity between yeast and human ALDH (Wang et al., 2009; Wang et al., 1998).

Further studies were subsequently carried out with hrALDHs to explore whether human ALDH was able to produce a carboxylic acid. Despite the affinity of ALDH1A1 for several aliphatic aldehydes (Gross et al., 2009; Lovold et al., 2006), our findings demonstrated no measurable NAD-dependent activity using LC-111 or its corresponding HAAQ product. However, isolation of the metabolites after hrALDH1A1 incubation and subsequent identification of potential metabolites via LC-MS did demonstrate the disappearance of the aldehyde peak of hydrolysed LC-111. This suggest either that (i) ALDH1A1 metabolism did occur but the corresponding carboxylic acid was not detectable or (ii) secondary reactions unrelated to ALDH1A1 occurred and prevented the desired ALDH-metabolism in being produced.
Chapter IV

Evaluation of the DNA binding properties and cytotoxic effects of 1,4-di-substituted acetalanthraquinones and their corresponding hydrolysed products
IV.1 Introduction and Aims

IV.1.1 Anthraquinones as DNA intercalating agents

In 1961, Lerman et al. (Lerman, 1961) proposed the theory of DNA intercalation based on observations from studying DNA-acridine complexes. DNA intercalators now comprise one of the most significant classes of agent that bind in a non-covalent manner with DNA. DNA intercalators include anticancer drugs such as doxorubicin, epirubicin and mitoxantrone, which possess a planar aromatic chromophore that is able to insert itself between adjacent base pairs within the DNA double helix. Stacking forces between the DNA bases and intercalators provide a stable interaction, which is further strengthened via hydrogen bonding between the basic alkylamino sidechains and the phosphodiester backbone (Islam SA, 1985) (further explained in section II.3).

DNA intercalation is often insufficient to cause cytotoxicity in cells, however long residence time at the DNA often interfere with the DNA strand-breakage-reunion reaction of topoisomerase II. This result in conversion of the enzyme into a poison (Pratt W. B., 1994) and leads to DNA double strand breaks, which are difficult to repair and therefore lethal to the cell.

IV.1.2 Mitoxantrone, a 1,4-di-substituted anthraquinone with significant clinical activity

Anthraquinone-based compounds in anticancer drug discovery are well described in the literature (Cheng et al., 1979; Cheng et al., 1983; Johnson et
al., 1979; Murdock KC et al., 1979; Zee-Cheng and Cheng, 1978). Mitoxantrone (see Figure 4.1), a 1,4-di-substituted anthraquinone anticancer drug (Shenkenberg and Von Hoff, 1986), was the lead molecule discovered in this class of synthetic compounds. Today, it is clinically used in the treatment of haematological malignancies, breast cancer and in advanced hepatic or ovarian carcinoma (Faulds et al., 1991). Further studies have consolidated its role in the treatment of prostate cancer in combination with prednisone (Tannock et al., 2004).

![Figure 4.1 Chemical structure of mitoxantrone](image)

Figure 4.1 Chemical structure of mitoxantrone. This compound contains a 5,8-OH di-substitution and amino-sidechains with a polar ethyl-alcohol feature.

Mitoxantrone binds to DNA through a high-affinity intercalation process, but also a low-affinity binding sustained through electrostatic forces contributes to the overall DNA binding (Pratt W. B., 1994). Mitoxantrone is cytotoxic primarily to cells in late S phase with arrest of cells in G2 due to topoisomerase II inhibition (Faulds et al., 1991). The major resistance mechanisms are associated with drug efflux mediated by P-glycoprotein (ABCB1) and ABCG2 (Doyle and Ross,
2003b), but also changes in topoisomerase II and DNA-repair activity contribute to drug resistance (Pratt W. B., 1994).

**IV.1.3 DNA-binding affinity via DNA melting studies**

1,4-di-substituted alkylamino-anthraquinones are often potent DNA-binding agents due to protonation at physiological pH of their basic secondary or tertiary amine functionalities, which allow stabilisation of the chromophore-DNA interaction (Patterson et al., 1994). The acetalanthraquinones (AAQs) synthesised in this study have been designed to bind weak-to moderately with DNA in order to reduce the cytotoxicity compared with mitoxantrone and improve their potential as “non-toxic” molecular fluorescent probes. Their affinity for DNA can be evaluated in various biochemical assays, but in this study, a DNA thermal denaturation assay is used to assess DNA-AAQ binding. In principal, this assay informs about what occurs when DNA in solution is heated in the absence or presence of DNA-affinic agents, such as 1,4-di-substituted anthraquinones. The denaturation of DNA is measured at 260 nm using a UV-spectrophotometer (Marmur and Doty, 1962). In a typical melting experiment, the increase in UV absorbance corresponds to the unwinding and denaturation of DNA (Figure 4.2), which occur due to the temperature increase. The melting temperature (Tm) of the DNA can be used as an indication of the strength of the interaction between DNA and the compounds to be investigated (Schallon et al., 2011).
Figure 4.2 DNA denaturation model. The scheme depicts the melting process of a DNA double helix. A-T-rich regions melt first, and then additional basepairs melt and unwind the helix. Finally G-C-regions melt to arrive at unwound single-stranded DNA region. Figure adapted from (Fox, 1997).

IV.1.4 Cytotoxic effect of anthraquinones and resistance mechanism

The success of aromatic anthraquinone-based anticancer agents has mainly been associated with their capacity to function as efficient DNA intercalators (see section II.1). The chemosensitivity of anthraquinone-based DNA intercalators are affected by common resistance mechanisms related to the expression of ABCB1 and ABCG2 transporters (Borst et al., 1999), also discussed above with mitoxantrone. As described in section I.3, overexpression of ABC transporters has been reported as a phenotypic marker of SCs and as a selective advantage to obtain cell protection. Therefore, an intimate balance between DNA targeting and cellular xenobiotic extrusion needs to be addressed in order to develop suitable fluorescent anthraquinone-based chemical probes.
IV.1.5 Aims of this study

Based on previous findings as described above and in section II.3, the cytotoxicity of the novel AAQs is expected to be related to their DNA-binding efficacy. However, it is not clear the impact that the incorporation of the acetal or chemically hydrolysed aldehyde modification has on the DNA binding, cellular distribution and cytotoxicity. Accordingly, the aims of this chapter are two-fold:

1. To assess the DNA affinity of the novel AAQs and their hydrolysed counterparts (HAAQs) using calf thymus DNA (CT-DNA) thermal denaturation assay;

2. To evaluate the potential cytotoxicity of AAQs and their hydrolysed counterparts (HAAQs) in the human lung A549 (+ALDH1A1) and breast MCF7 (-ALDH1A1) carcinoma cell lines.
IIV.2 Materials and Methods

IV.2.1 AAQs stock solutions

Compound stock solutions were prepared for use by dissolving 1 mg of AAQ or corresponding hydrolysed HAAQ in 1.2 mL of DMSO (Sigma). The final concentration of the DMSO in the test volume did not exceed 1%, as a concentration greater than 1% DMSO can act to destabilise DNA secondary structures (Doyle and Ross, 2003). Hydrolysis of AAQs was carried out as described in section III.2.1.2.

IV.2.2 Melting curve assays for DNA binding analysis

Calf thymus DNA (CT-DNA, Sigma) solutions were prepared for use by dissolving 21-23 mg in 20 mL of CT-DNA melting buffer (see Appendix II for composition) and stored at 4ºC. Mitoxantrone (Sigma) was solubilised in DMSO, diluted to desired concentrations in 10 mM Tris-HCl buffer (pH 7.4) and used as control. Its concentration was determined spectrophotometrically using a molar extinction coefficient of 8,360 M\(^{-1}\) cm\(^{-1}\) at 608 nm (Enache and Volanschi, 2010).

The CT-DNA thermal melting assay mixtures were prepared by adding sufficient CT-DNA stock solution to CT-DNA melting buffer, to ensure the preparation of a 50 µM CT-DNA solution. AAQ stock solution was then added to the 50 µM CT-DNA solution, to ensure the final concentration of compound in the mixture was 10 µM. All UV absorbance experiments were conducted on a Cary 400Bio Spectrophotometer (Varian, Agilent) equipped with a Peltier Temperature Controller. This was attached to a Cary Thermostatable Multicell Block, to
provide stable, continuous temperature control. A pair of masked quartz 1.2 mL microcells cells with a 4 mm width, and 10 mm path length was used for all absorbance studies. 1.2 mL of CT-DNA melting buffer was added to the reference cuvette and 1.2 mL of DNA-anthraquinone mixture was added to the reaction cuvette. Absorbance versus temperature profiles were measured at 260 nm with melting temperature (Tm) measurements initiated at 40 °C. The temperature of the test solution was increased at 1 °C/min till 95 °C was reached.

Melting curves generated by the computer software were printed out and the melting points were determined using the published graphical method by Fox et al. (Brown and Fox, 1997). The Tm was the mean of two separate determinations within 0.2 °C range. The assay was performed initially by incubating 50 µM of calf thymus DNA (CT-DNA), in the presence of 10 µM of mitoxantrone, to validate the assay and to have a positive control for comparison. 10 µM solutions of AAQ and HAAQs counterparts were incubated with 50 µM CT-DNA.

**IV.2.3 Mammalian cell lines culture**

**IV.2.3.1 Maintenance of mammalian cell lines**

Mammalian cell lines were cultured as outlined in Table 4.1. The two cell lines were obtained from the American Type Culture Collection (ATCC) and grown in a monolayer in R10 media (see Appendix II) supplemented with 1 mM sodium pyruvate (Sigma). All cells were grown at 37 °C, 5% CO₂ and 100% humidity.
Cell culture plastics were from Corning. Cells were kept at subconfluent levels and were typically passaged when 70–80% confluent.

<table>
<thead>
<tr>
<th>Cell line</th>
<th>Origin of cell line</th>
<th>Approx. frequency of subculture</th>
<th>Dilution upon subculture</th>
</tr>
</thead>
<tbody>
<tr>
<td>A549</td>
<td>Human lung adenocarcinoma</td>
<td>4 days</td>
<td>1:5-1:10</td>
</tr>
<tr>
<td>MCF7</td>
<td>Human breast adenocarcinoma</td>
<td>4-5 days</td>
<td>1:10</td>
</tr>
</tbody>
</table>

Table 4.1 Culture of mammalian cell lines.

IV.2.3.2 Mammalian cell passaging

The old media was aspirated carefully and discarded from each 75 cm² flask. The cells were washed with 10 mL Hank’s Balanced Salt solution (Sigma), which was then aspirated and discarded. To break down cellular attachments to the flask, 2 mL trypsin in EDTA (Sigma) (1:10 dilution) was added to cover all cells. The flasks were incubated for 3-5 min, and the contents of the flask were checked under light microscope to ensure that the cells were mobile and not clumped. The trypsin was quenched with 8 mL of media, which was added slowly to avoid the formation of bubbles. The cells were then centrifuged at 1,000 rpm for 5 min. The supernatant was discarded and cells were resuspended in culture media to the appropriate volume.

IV.2.3.3 Determination of live cell number

To determine live cell counts, 0.5 mL 0.4% Trypan Blue Stain (Sigma), 0.3 mL Hank’s balanced salt solution and 0.1 mL cell suspension were assembled in a
tube, mixed vigorously by vortexing and incubated at room temperature for 5 min. Total cell number (non-stained cells and blue cells) and live cell number (non-stained cells) were then determined using a haemocytometer.

**IV.2.3.4 Cryopreservation of mammalian cells**

For storage in liquid nitrogen, mammalian cells were trypsinised, sedimented by centrifugation and suspended in ice-cold standard freezing media (see Appendix II for media).

**IV.2.3.5 Cellular proliferation assessment using MTT assay**

The viability of A549 or MCF7 in response to novel 1,4-AAQ-based compounds (see Appendix III leaflet attached) was determined using the MTT (3-(4,5-dimethylthiazolyl-2)-2, 5-diphenyltetrazolium bromide) assay (Mosmann, 1983). In living cells, the membrane permeable yellow MTT tetrazolium salt is reduced to purple formazan, which is membrane impermeable. MTT reduction is generally attributed to mitochondrial activity, but it has also been related to non-mitochondrial enzymes as well as to endosomes and lysosomes (Berridge and Tan, 1993). Initially, cells were seeded into a rounded-bottom 96 well plate (Corning) at six different cell seeding densities ($0.5 \times 10^4$, $1 \times 10^4$, $2 \times 10^4$, $3 \times 10^4$, $4 \times 10^4$, and $5 \times 10^4$) per mL of R10 (cells/mL) in order to optimise the assay and establish the most suitable seeding assay. Cell number was calculated according to section III.3.1.3. Concentration of $1 \times 10^4$ cells/mL was determined to be the optimum seeding density for both cell lines for 96 h total incubation. A control lane contained supplemented R10 only. The 96 well plates were incubated overnight at 37 °C, 5% CO₂ and 100% humidity to allow cell
adhesion. Cells were then treated respectively with novel AAQs (0.25-100 µM) or doxorubicin (supplied as doxorubicin hydrochloride, Sigma) (0.005-100 µM) as a positive cytotoxic control. The cells were incubated with the AAQs or corresponding HAAQ counterparts for 72 h. All anthraquinones were dissolved in DMSO (Sigma), ensuring that the DMSO concentration was no more than 0.1% per well. A control of 0.1% DMSO was included to make certain that any observed effect was due to only the tested anthraquinone. Following the 72 h incubation period, media containing the respective test compound was removed and replaced with supplemented R10 containing a 1:10 dilution of 5 mg/mL of MTT solution in water. Cells were incubated for a further 4 h. The MTT solution was removed and the resulting purple formazan crystals were resuspended in 150 µl/well of DMSO and absorbance was read at 540 nm. The cytotoxicity was extrapolated by comparing drug treated cell data with an untreated control, and curves were determined as percentage of survival. Finally, IC\textsubscript{50} (concentration that inhibits the control growth by 50%) values were calculated using Microsoft Excel software.

IV.2.3.6 Statistical analysis

Statistical significance of the results was evaluated on data from at least three independent experiments. The means of IC\textsubscript{50} values were calculated and plotted using Microsoft Excel software. R\textsuperscript{2} coefficient determination was used to give information about the accuracy of the obtained results.
IV.3 Results

IV.3.1 DNA melting temperature as an indicator of DNA

In the present study, assessment of variation in Tm due to possible interaction between CT-DNA and anthraquinones was carried out using a thermal DNA denaturation assay (results shown in Table 4.2).

<table>
<thead>
<tr>
<th>Compound</th>
<th>AAQ Tm (ºC)</th>
<th>HAAQ Tm (ºC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mitoxantrone</td>
<td>&gt;95</td>
<td>-</td>
</tr>
<tr>
<td>LC-103</td>
<td>69.0</td>
<td>69.7</td>
</tr>
<tr>
<td>LC-105</td>
<td>74.0</td>
<td>71.9</td>
</tr>
<tr>
<td>LC-106</td>
<td>69.0</td>
<td>70.0</td>
</tr>
<tr>
<td>LC-107</td>
<td>69.4</td>
<td>70.1</td>
</tr>
<tr>
<td>LC-108</td>
<td>69.6</td>
<td>**</td>
</tr>
<tr>
<td>LC-109</td>
<td>69.0</td>
<td>**</td>
</tr>
<tr>
<td>LC-110</td>
<td>72.1</td>
<td>*</td>
</tr>
<tr>
<td>LC-111</td>
<td>76.4</td>
<td>*</td>
</tr>
<tr>
<td>LC-112</td>
<td>72.0</td>
<td>72.9</td>
</tr>
<tr>
<td>LC-113</td>
<td>72.0</td>
<td>71.6</td>
</tr>
<tr>
<td>LC-114</td>
<td>71.6</td>
<td>74.6</td>
</tr>
<tr>
<td>HA-218</td>
<td>74.2</td>
<td>70.4</td>
</tr>
<tr>
<td>HA-222</td>
<td>70.8</td>
<td>*</td>
</tr>
<tr>
<td>HA-236</td>
<td>70.0</td>
<td>70.0</td>
</tr>
<tr>
<td>HA-252</td>
<td>73.0</td>
<td>72.6</td>
</tr>
</tbody>
</table>

Table 4.2 CT-DNA/AAQs melting temperatures before and after acidic hydrolysis.* Excess of interference, hence any calculation of Tm was inaccurate. ** Solubility issues in the working buffer, therefore no experiment was performed. Melting curve plots can be found in Supplementary Information IV.A.
Naked CT-DNA showed a Tm of 69.1 °C and addition of mitoxantrone increased this value to higher than 95°C (ΔTm > 30 °C). In comparison, the effect of the novel AAQs on Tm values was weak to moderate with Tms falling between 69 and 76 °C. Furthermore, hydrolysis of the majority of the LC-AAQs was demonstrated to have adverse or no effect on the Tm of CT-DNA.

The same experiment was performed with the HA-AAQ derivatives and the parent compounds showed a moderate increase in Tm compared with the LC-AAQ congeners. The most DNA-affinic compounds, HA-218 and HA-252, produced ΔTm = 6.06 and 4.86 °C respectively. However, compared with mitoxantrone, this was a moderate effect on CT-DNA. Hydrolysed HA-AAQs were investigated in the same conditions, however no significant effect on the DNA affinity was observed. Furthermore, solubility issues and post-hydrolysis precipitations of some AAQs (noted with * on Table 4.2) created a loss of accuracy and potential artefacts in the absorbance readings.

**IV.3.2 Effect of novel AAQs and HAAQ counterparts on cell viability**

Symmetrical and non-symmetrical AAQs were tested with a concentration range between 0.25 and 100 µM for 72 h and doxorubicin, as positive cytotoxic control, was tested in parallel at the same doses (complete IC50 values list was reported in Tables 4.3-4.4).
<table>
<thead>
<tr>
<th></th>
<th>AAQ</th>
<th>MCF7 (µM)</th>
<th>R² value</th>
<th>A549 (µM)</th>
<th>R² value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Doxorubicin</td>
<td></td>
<td>0.028</td>
<td>0.1831</td>
<td>0.05</td>
<td>0.1759</td>
</tr>
<tr>
<td>Symmetrical</td>
<td></td>
<td>0.750</td>
<td>0.7865</td>
<td>&gt;100</td>
<td>0.9034</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt;100</td>
<td>0.7199</td>
<td>&gt;100</td>
<td>0.6439</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.75</td>
<td>0.6174</td>
<td>&gt;100</td>
<td>0.2725</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt;100*</td>
<td>0.1835</td>
<td>&gt;100*</td>
<td>0.7433</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt;100</td>
<td>0.4418</td>
<td>&gt;100</td>
<td>0.5332</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt;100</td>
<td>0.6994</td>
<td>83.2</td>
<td>0.9444</td>
</tr>
<tr>
<td>Non-symmetrical</td>
<td></td>
<td>11.3</td>
<td>0.5697</td>
<td>11.7</td>
<td>0.568</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.66</td>
<td>0.5541</td>
<td>17.5</td>
<td>0.9475</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.44</td>
<td>0.4437</td>
<td>1.7</td>
<td>0.7563</td>
</tr>
<tr>
<td></td>
<td></td>
<td>15.70</td>
<td>0.7976</td>
<td>22.5</td>
<td>0.7371</td>
</tr>
<tr>
<td></td>
<td></td>
<td>19.5</td>
<td>0.7258</td>
<td>22.3</td>
<td>0.9473</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.8</td>
<td>0.757</td>
<td>38.2</td>
<td>0.9546</td>
</tr>
<tr>
<td>Others</td>
<td></td>
<td>1.73</td>
<td>0.6857</td>
<td>7.5</td>
<td>0.7395</td>
</tr>
<tr>
<td></td>
<td></td>
<td>37.3</td>
<td>0.6787</td>
<td>12.4</td>
<td>0.6857</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt;100</td>
<td>0.9528</td>
<td>&gt;100</td>
<td>0.3755</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.2</td>
<td>0.4871</td>
<td>1.37</td>
<td>0.3569</td>
</tr>
</tbody>
</table>

Table 4.3 Growth inhibition (IC₅₀) of LC-AAQ and HA-AAQ libraries against MCF7 and A549 cancer cell lines. * Solubility problems, blue crystals appeared after 48 h of treatment. R² is a statistical measure of how well the regression line approximates the real data points. R² of 1.0 indicates that the regression line perfectly fits the data.
<table>
<thead>
<tr>
<th>Hydrolysed compound</th>
<th>MCF-7 (µM)</th>
<th>( R^2 ) value</th>
<th>A549 (µM)</th>
<th>( R^2 ) value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Symmetrical</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LC-103</td>
<td>56.73*</td>
<td>0.9434</td>
<td>&gt;100</td>
<td>0.9621</td>
</tr>
<tr>
<td>LC-055</td>
<td>&gt;100</td>
<td>0.9651</td>
<td>&gt;100</td>
<td>0.51</td>
</tr>
<tr>
<td>LC-106</td>
<td>&lt; 0.250*</td>
<td>0.8793</td>
<td>&gt;100*</td>
<td>0.8352</td>
</tr>
<tr>
<td>LC-107</td>
<td>&gt;100*</td>
<td>0.5492</td>
<td>&gt;100*</td>
<td>0.6735</td>
</tr>
<tr>
<td>LC-108</td>
<td>&gt;100*</td>
<td>0.7382</td>
<td>&gt;100*</td>
<td>0.6318</td>
</tr>
<tr>
<td>LC-109</td>
<td>&gt;100</td>
<td>0.7128</td>
<td>&gt;100</td>
<td>0.7946</td>
</tr>
<tr>
<td><strong>Non-symmetrical</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LC-110</td>
<td>40*</td>
<td>0.933</td>
<td>&gt;100*</td>
<td>0.8502</td>
</tr>
<tr>
<td>LC-105</td>
<td>29.51</td>
<td>0.7825</td>
<td>12.7</td>
<td>0.7826</td>
</tr>
<tr>
<td>LC-111</td>
<td>&lt; 0.250</td>
<td>0.3528</td>
<td>1.24</td>
<td>0.6786</td>
</tr>
<tr>
<td>LC-112</td>
<td>2.26</td>
<td>0.5933</td>
<td>34</td>
<td>0.8435</td>
</tr>
<tr>
<td>LC-113</td>
<td>20.11</td>
<td>0.723</td>
<td>45.2</td>
<td>0.8345</td>
</tr>
<tr>
<td>LC-114</td>
<td>8.60</td>
<td>0.6085</td>
<td>32.3</td>
<td>0.6153</td>
</tr>
<tr>
<td><strong>Others</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HA-218</td>
<td>&lt; 0.250</td>
<td>0.7012</td>
<td>1.74</td>
<td>0.7028</td>
</tr>
<tr>
<td>HA-222</td>
<td>&lt; 0.250</td>
<td>0.9027</td>
<td>5.3</td>
<td>0.5976</td>
</tr>
<tr>
<td>HA-236</td>
<td>ND</td>
<td>0.9778</td>
<td>&gt;100</td>
<td>0.727</td>
</tr>
<tr>
<td>HA-252</td>
<td>0.35</td>
<td>0.3797</td>
<td>0.54</td>
<td>0.5141</td>
</tr>
</tbody>
</table>

Table 4.4 Growth inhibition (IC\(_{50}\)) of hydrolysed AAQs (HAAQs) against MCF7 and A549 cancer cell lines. * Solubility problems, blue crystals appeared after 48 h of exposure. ND: not determined. \( R^2 \) is a statistical measure of how well the regression line approximates the real data points. \( R^2 \) of 1.0 indicates that the regression line perfectly fits the data.

Some impact on cell viability was dictated by structural modifications to the chromophore skeleton as pointed out in Figure 4.3 (top panel). The addition of hydroxyl groups at carbon 5 and 8 positions in the chromophore was generally
correlated with improved DNA affinity and cytotoxicity compared with the 5,8 non-hydroxylated derivatives, which is in agreement with other similar studies (Bailly et al., 1996).

Figure 4.3 Structure activity relationships with selected LC-AAQ derivatives. This representation points out two possible key modifications (black box) within the AAQ structure responsible of enhanced cellular cytotoxicity.

The tethering of a basic \( N,N \)-dimethylethylamine sidechain with high affinity for DNA (as exemplified in AQ4 and AQ6, see section II.1.1) to generate non-symmetrical derivatives led to enhanced cytotoxicity in all the LC-AAQ compounds. For example, non-symmetrical configured compounds such as LC-105, LC-111 and LC-113 exhibited enhanced cytotoxic effects when compared with their symmetrical counterparts LC-055, LC-108 and LC-107, probably
because of the stronger interaction of the \( N,N \)-dimethylethylamine moiety with the DNA sugar backbone as well as higher solubility (see Figure 4.3 bottom panel). Moreover, hydrolysis of the symmetrical AAQs generally led to an observable enhanced cytotoxicity, whereas loss of cytotoxic activity was revealed with non-symmetrically configured AAQs. Modifications to the carbon sidechain structure were found to affect the potency (see Figure 4.4 for structural variations).

Figure 4.4 Modification of the carbon sidechain skeleton of LC-105, LC-112 and LC-108. Black boxes represent the modification points.

LC-108 and LC-112, containing amide and “reverse” amide bond respectively, were less potent than the LC-105 congener. The results are in agreement with a previous study concerning incorporation of non-basic heteroatoms into the
sidechain of anthraquinones, which also showed a reduction in chemosensitivity (Krapcho et al., 1990). Incorporation of bulky substituents at the terminal nitrogen of the HA-AAQ library revealed a decrease in cytotoxic effect (see Figure 4.5 for structural modifications): HA-236 (benzyl) < HA-222 (methyl) < HA-218 (H).

As discussed in section III.3.1, HA-AAQs were resistant to hydrolysis at 37 °C, so exposure to 2M HCl at 75 °C was necessary. Hydrolysed HA-AAQ products showed an enhanced activity on cell viability compared with their AAQ counterparts, except for HA-236. Indeed, the benzyl group in HA-236 was

Figure 4.5 Modifications of the secondary amine in the HA-AAQ library.
shown to have no effect under the tested conditions, which is likely to be due
the bulkiness of the benzyl moiety which prevents DNA binding (Murdock et al.,
1979).
IV.4 Discussion

The experiments conducted in this chapter have provided information about how AAQs and their corresponding hydrolysed counterparts bind with DNA and whether this DNA binding can be correlated with their cytotoxic potential.

We have investigated two libraries of AAQs and none of the test compounds were shown to possess high affinity for CT-DNA as measured by a thermal DNA denaturation assay. A few compounds including HA-218, HA-252, LC-105 and LC-111 showed moderate affinity for DNA, but in comparison with mitoxantrone this binding was negligible. The results clearly indicate that although the anthraquinone is an excellent DNA intercalator (Lerman, 1961; Lown et al., 1985), its strength in binding is almost exclusively related to the nature of sidechains inserted onto the anthraquinone chromophore.

It was proposed that a stronger interaction between the aldehyde moiety of the HAAQs and the DNA (see Fig. 3.5 for the activation scheme) would enhance the DNA affinity via potential for generating DNA crosslinks (Hilton, 1984; Kuykendall and Bogdanffy, 1992). The electron-rich guanine or adenine bases are known nucleophiles capable of generating covalent bonds to drugs via Schiff base formation (Kurtz and Lloyd, 2003). However, the results suggested that no increased DNA binding occurred with the HAAQs. Indeed, it may be that the failure to increase ΔTm was due to depurination occurring at elevated temperatures, induced by Schiff base formation (see Figure 4.6) (Goffin et al., 1984).
Depurination has been defined as an alteration of DNA in which the purine base (adenine or guanine) is removed from the deoxyribose sugar by hydrolysis of the beta-N-glycosidic bond (Holmquist, 1979). After depurination, the sugar phosphate backbone remained intact with the sugar ring displaying a hydroxyl (-OH) group in the place of the purine. At elevated temperatures as employed in the thermal denaturation assay, the Schiff base linkage may induce
depurination at the DNA-drug interaction site. This phenomenon could explain the lowering in absorption and Tm values of the HAAQs.

All novel compounds were investigated in A549 (high ALDH expression) (Moreb JS, 2008) and MCF7 (low expression of ALDH) (Bunting KD, 1994). From the cytotoxicity plots (reported in Supplementary Information IV.B) and IC$_{50}$ values (see Tables 4.3-4.4) it is evident that MCF7 generally showed an increased sensitivity to AAQs compared with A549 cell line.

Figure 4.7 ALDH-mediated conversion of aldehyde moiety into a carboxylic acid. The carboxylic acid is known to be a good substrate for efflux pump proteins.

A factor contributing to higher resistance in A549 cell line is possibly related to the higher expression of ALDH. The presence of these chemoprotecting enzymes are likely to convert aldehyde moieties into carboxylic acid, which are poor components for DNA binding as well as better substrate for ABC transporters (see Figure 4.7) (Zhou et al., 2002). Although this loss in activity is explainable for the HAAQs, it is not clear why the parental AAQs are less cytotoxic in A549 cell line. Unless, the acetal moiety is prone to oxidative
metabolism as experiments using cytosolic fractions from rat liver, enriched in ALDHs, may suggest (see section III.3.4.1).

In general, convincing SARs between AAQs and their hydrolysed counterparts are difficult. However, a general classification of the AAQ or HAAQ effect on cell viability may distinguish the compounds as cytotoxic (below 12% cell survival), growth inhibiting (between 25-75% cell survivals) or cytostatic agents (above 75% cell survival). On this basis, the shape of the cytotoxicity curves was analysed. The non-symmetric derivatives with an amide bond in the sidechain (i.e. LC-111, LC-112, LC-113, and LC-114) revealed a relatively higher cytotoxic profile in both cell lines. Similarly, HA-AAQs revealed a cytotoxic response independently from the cell line model, except for HA-236, which had a consistent cytostatic effect. However, these results may be due to lack of cell penetration. An interesting behaviour was revealed with LC-110, which was shown to be a cytostatic compound. However its hydrolysed product was classified as cytotoxic without cell-type discrimination. Furthermore, as reported in Table 4.4, some derivatives revealed solubility issues after hydrolysis, which impacts the AAQ availability in solution and potential to interact with DNA-binding and cell viability.

In summary, the novel AAQs revealed to be weak to moderate DNA binders. Hydrolysed AAQs were not significantly different to the parental AAQs in DNA binding or in generating cytotoxic activity. In comparison with doxorubicin, the anthraquinones were less cytotoxic and in many cases not toxic under the conditions investigated. Accordingly, their utility as molecular fluorescent probes to study live cells can be further studied.
Chapter V

Evaluation of fluorescent properties of novel 1,4-di-substituted acetalanthraquinones as potential ALDH-targeting probes
V.1 Introduction and Aims

V.1.1 The development of fluorescent probes to explore live cells

Since the nineteenth century, optical microscopy has been a leading technology to observe living cells, but the lack of contrast in several tissues soon led to the introduction of dyes, for their ability to stain or label various cell structures with a different grade of selectivity (McCutcheon and Lucre, 1924; Tsien, 2005). It was common that these dyes could not cross intact cell membranes and had to be applied at high concentration because of their low affinity for intracellular targets. Therefore, until 1930 biological microscopy became widely the study of dead or fixed tissues and consequently a relatively static discipline. In 1942, Coons et al. introduced fluorescently labelled antibodies to recognise any target endogenous protein with high affinity and selectivity (Ploem, 1971). This converged with the development of a new microscopy configuration that enabled the attachment of an epi-illumination for the excitation and detection of fluorescent probes.

Fluorescence consists of the ability of reagents not just to absorb light (Williams and Bridges, 1964) but also to re-emit light at longer wavelength (Lakowicz et al., 1992). It is a cyclical process that occurs in a very small percentage of light-absorbing molecules, called fluorophores or fluorescent dyes, as most just convert the light into heat (Tsien, 2005). A fluorescent probe is a fluorophore designed to respond to a specific stimulus or to specifically localize within a biological specimen and detectable with much greater sensitivity than non-fluorescent dyes. A desirable property for such fluorescent probes used to study living cells is that it must not impair normal cell function (Taylor and Wang,
1980). Therefore, the approach of probing live cells to investigate highly
dynamic and transitory cellular mechanisms has presented an active and
exciting research field in its own right (Coling and Kachar, 2001).

Since the early ‘70s, when Waggoner demonstrated the potential dependent
partition of cyanine-dye molecules between the cells and the extracellular
medium in blood cells (Sims et al., 1974; Waggoner and Stryer, 1970), the
concept of using fluorescent compounds to address biological questions started
to become a promising approach. In the same years, Haugland founded
Molecular Probes with the intention to make fluorescent dyes available for
biological research and this had crucial impact on the application of
fluorescence to cellular biology (Haugland, 1999). Moreover, another
paradigm-shift that occurred during the 90s was the cloning and expression of
the jellyfish Green Fluorescent Protein (GFP), which resulted in the expression
of a fluorescent marker as a result of gene expression and protein targeting in
intact cells and organisms (Tsien, 1998; Tsien, 2003).

Thus, these and many other discoveries in molecular tagging, cellular labelling
have resulted in an ongoing grand challenge of developing fluorophores as
small molecules that can be part of a molecule (intrinsic fluorophores) or added
to it (extrinsic fluorophores) and so they can be used as natural indicators to
study the structure, dynamics, and metabolism of living cells. In recent date,
the fluorescence tagging of antibodies has become a routine part of the
characterization of tumour biomarkers for delivering an "image and treat"
concept in patient diagnosis (Ardeshirpour Y, 2012). At the same time, there is
a need for tissue-based multivariate biomarker assays to improve the accuracy of diagnostic, prognostic, and predictive testing (Nederlof et al., 2011).

In this chapter, studies are focussed on the development and assessment of fluorescence-based assays for the analysis of many important live-cell functions, suitable for determining apoptosis, cell adhesion, multi-drug resistance, endocytosis, secretion and signal transduction. As reported in Table 5.1, the number of such probes within a functional assay design has been rapidly growing and these approaches are becoming critical tools for detection of diverse cell processes (Dailey M.E., 2006).

Diverse platforms have been designed and engineered to detect fluorescent probes in cells (Errington et al., 2005). Many of the assays revealed above can be analyzed on a cell-by-cell basis and some are equally suitable for detection with a classic fluorescence microscope, a flow cytometer or a microplate reader with the advantage that the majority of the assays have the capacity for high-throughput analysis (Held et al., 2010).
<table>
<thead>
<tr>
<th>Probe Name</th>
<th>Functional Assay</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Blue-green illumination</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AlamarBlue</td>
<td>Cell health indicator</td>
<td>(Nakayama et al., 1997)</td>
</tr>
<tr>
<td>Aldefluor</td>
<td>ALDH activity</td>
<td>(Jiang et al., 2009)</td>
</tr>
<tr>
<td>BCECF</td>
<td>Cytoplasm pH changes indicator</td>
<td>(Han and Burgess, 2009)</td>
</tr>
<tr>
<td>Carboxy-DCFDA</td>
<td>Hypoxia detector; pH sensor</td>
<td>(Hagar et al., 1996)</td>
</tr>
<tr>
<td>CFSE</td>
<td>Live cell marker Lineage tracker</td>
<td>(Weston and Parish, 1990)</td>
</tr>
<tr>
<td>Fluo-4AM</td>
<td>Calcium indicator</td>
<td>(Ma et al., 2010b)</td>
</tr>
<tr>
<td>JC-1</td>
<td>Mitochondrial depolarization; Cellular apoptosis</td>
<td>(Smiley et al., 1991)</td>
</tr>
<tr>
<td>MitoSOX Red</td>
<td>Mitochondrial superoxide indicator</td>
<td>(Żielonka et al., 2008)</td>
</tr>
<tr>
<td>MitoTracker Green FM</td>
<td>Mitochondrial membrane depolarisation</td>
<td>(Pendergrass et al., 2004)</td>
</tr>
<tr>
<td>RedoxSensor CC-1</td>
<td>Cytosolic redox sensor Hypoxia detector Free radicals detector</td>
<td>(Chen and Gee, 2000) (Oksvold et al., 2002)</td>
</tr>
<tr>
<td>Rhodamine 123</td>
<td>Mitochondrial depolarization Cellular apoptosis ABCB1-mediated exclusion</td>
<td>(Emaus et al., 1986; Johnson et al., 1980; Eytan et al., 1997; Yumoto et al., 1999)</td>
</tr>
<tr>
<td>SNARF-1,-4F,-5F</td>
<td>pH indicators</td>
<td>(Han and Burgess, 2009)</td>
</tr>
<tr>
<td><strong>Red illumination</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CellROX</td>
<td>Cellular oxidative stress Reactive oxygen species measurement</td>
<td>(Beazley et al., 2011)</td>
</tr>
<tr>
<td>DRAQ5</td>
<td>Cell cycle DNA content</td>
<td>(Smith PJ, 2000; Smith et al., 1999)</td>
</tr>
<tr>
<td>MitoTracker Red FM</td>
<td>Mitochondria depolarisation</td>
<td>(Ba et al., 2010; Csiszar et al., 2009; Hailey et al., 2010)</td>
</tr>
<tr>
<td><strong>UV illumination</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CellTrace Violet</td>
<td>Proliferative indicator Lineage tracker</td>
<td>(Jones et al., 2010)</td>
</tr>
<tr>
<td>ER-Tracker Blue-White DPX</td>
<td>Endoplasmic Reticulum labelling</td>
<td>(Abodeely et al., 2009)</td>
</tr>
<tr>
<td>Hoechst 33342</td>
<td>SP discrimination</td>
<td>(Goodell et al., 1996)</td>
</tr>
<tr>
<td>MCB</td>
<td>Glutathione depletion reactive oxygen species measurement</td>
<td>(Abramov et al., 2007)</td>
</tr>
<tr>
<td>ThiolTracker</td>
<td>Glutathione localisation detection cellular redox status</td>
<td>(Mandavilli and Janes, 2001)</td>
</tr>
</tbody>
</table>

Table 5.1 Exemplar fluorescent labels commonly used for the functional analysis of living cells.
V.1.2 Anthraquinone-based fluorescent probes - seeking a functional readout

The concept of fluorescence detection using an anthraquinone quencher dye conjugated to a variety of biologically relevant compounds became of interest in 1970s (Donati et al., 1973; Morsi and Williams, 1977). The chromophores of DNA-binding agents have been exploited extensively in spectroscopic and live cell studies to probe features of the nucleus and to explore the dynamics of ligand binding. Tracking DNA-ligand interactions in live cells also facilitates the design of anticancer agents, since many drugs are auto-fluorescent (Njoh KL, 2006).

Anthracycline antibiotics are an important class of anticancer therapeutics and are known for their ability to bind DNA as intercalating agents and act as topoisomerase II inhibitors (Tewey et al., 1984). 1,4-Dihydroxyanthraquinone is the smallest unit in the anthracycline structures to possess the characteristic of the chromophore: a broad UV-visible absorbance spectrum (360-515 nm). The binding of 1,4-dihydroxyanthraquinone to DNA shifts the absorption spectrum to longer wavelength, as a consequence of the red-light excitation of the chromophore (Eriksson M, 1988). Indeed, nuclear DNA is believed to be the primary target of anti-tumour activity of the anthracyclines, resulting in inhibition of DNA transcription and replication (Cozzarelli and Wang, 1990). In vitro, the anthracycline-DNA interaction mainly occurs by intercalation of the planar ring system between DNA basepairs (see section II.1). In vivo, anthracycline biofunction has been documented by several structure-activity relationship studies and many compounds that have reached the clinical stage, for example,
generally belong to the group with the highest affinity for DNA (Crow and Crothers, 1994). Mitoxantrone is another example of anticancer anthraquinone drug, structurally related to the DNA intercalating anthracycline antibiotics. The auto-fluorescence of such agents, such as anthracyclines, has enabled uptake studies or pharmacokinetic studies to be performed using flow cytometry or fluorescence microscopy (Krishan and Ganapathi, 1980). Critically, both the development of novel drugs with DNA binding properties and target specificity, and molecular probes with desired chromophore performance must address the sequence preferences of ligands and the modes of DNA interaction (Njoh KL, 2006).

Anthraquinone-based fluorescent probes are already in use and been patented by Biostatus Ltd. The lead agent DRAQ5™ (see Figure 5.1), a 1,5-bis[2-(methylamino)ethyl]amino-4,8-dihydroxy anthracene-9,10-dione, and its derivatives were demonstrated with DNA binding capability through intercalation (Smith PJ, 2000; Smith et al., 1999). The 1,5-di-substituted derivatives intercalate with each substituent positioned in each groove and the intercalation is stabilized by electrostatic interactions between the protonated tertiary amino group of the sidechain and the phosphate backbone of the DNA (Njoh KL, 2006). Molecular dynamics simulation of the DNA–DRAQ5 complex (see Figure 5.1B) without any constraints leads to DRAQ5 protrusion into the interface of two A-T (adenine-thymine) pairs by displacing the aromatic rings of two base pairs out of the DNA backbone and DRAQ5 stacking between those aromatic rings (Islam SA, 1985).
Figure 5.1 DRAQ5™ fluorescent probe. An anthraquinone molecule that enters live cells, binds to the nucleus and provides a quantitative fluorescence readout for DNA content: (A) DRAQ5 chemical structure; (B) DNA-DRAQ5 complex indicates DRAQ5 protrusion into the interface of two A-T pairs in the DNA minor groove (figure adapted from Dr Zloh M., University of London); (C) DRAQ-5 nuclear staining using widefield-microscopy (kindly provided by Sally Chappell, Cardiff University).

DRAQ5 achieves good nuclear discrimination (calculated high nucleus/cytoplasm ratio) by its high affinity for DNA and does not show fluorescence enhancement with DNA in free solution (Errington et al., 2005). With an excitation at 647 nm, DRAQ5 has a fluorescence spectrum extending from 665 nm to 780 nm (see spectra in Figure 5.2). This far-red emission shows little overlap with other commonly used green emitting cellular dyes (Smith et
al., 2000). However, mitoxantrone-like chromophores have been found to have very weak fluorescence (Krishan and Ganapathi, 1980).

![Figure 5.2 DRAQ5 spectral analysis. (A) Absorbance and (B) fluorescence spectra.](image)

Although it has been classified as a weak fluorophore (quantum yield ranging from 0.003 to 0.004, demonstrating that most of the absorbed photon energy translates to nonradiative decay) (Njoh KL, 2006), the high DNA binding affinity of this “red-shifted” molecule has ability to identify cell cycle position through DNA content analysis. Moreover, the stoichiometric DNA binding and the high membrane penetrance of this molecular probe results in a valid and robust tool for live cell analysis (Smith et al., 1999). Therefore, DRAQ5 represents an innovation, progressing a fluorescent agent into a fluorescent label suitable for assaying DNA content in a live cell.
V.1.3 Cell model systems for novel probe screening

Two extensively characterised cell lines are selected for the study: the human osteosarcoma cell line U-2 OS and the human lung adenocarcinoma cell line A549.

![Chemical structure and nuclear staining](image)

**Figure 5.3 Hoechst 33342 chemical structure and nuclear staining in A549 cells.** Figure adapted from Motiwala *et al.* (Motiwala *et al.*, 2004). Hoechst 33342 does not bind DNA by intercalation or ionic interaction with the phosphate groups, but it has high specificity for DNA minor groove of double-stranded DNA.

The U-2 OS cell line has been selected as it represents a cell line widely used in biomedical research. Two tumour suppressive genes, *p53* and *pRb*, are functional in U-2 OS cells, whereas in other, more aggressive, osteosarcoma cell lines such as Saos-2, these genes are mutated (Ponten and Saksela, 1967) and they are routinely used for drug screening purposes and high-through-put analysis (Vollmers *et al.*, 2008; Zhuang *et al.*, 2011). Moreover, no pumps and no ALDH isoforms have been detected in U-2 OS cell system (Zhuang *et al.*, 2011).
A database survey of 59 cancer cell lines from the NCI60 panel (kindly donated by collaborator Smith PJ) has been undertaken to evaluate the prevalence and patterns of ABCG2 and ALDH1A1 over-expression in established human tumour lines, and A549 cell line shows high expression of both proteins (see Figure 5.4). In this study, A549 cell line is also chosen as an established cell model with ability to maintain side populations (SPs), identified by a reduced accumulation of Hoechst dye 33342 (see Figure 5.3A for structure) attributable to ABCG2-mediated efflux (Scharenberg et al., 2002). Hoechst 33342 dye has been used as vital probe for nuclear location, however, as a substrate for efflux pumps it is not well retained in many cell types (see Figure 5.3A) (Morgan et al., 1990), and the combination of nuclear staining with ABCG2-dependent discrimination results in a quantitative functional assay for attributes of drug resistance or SC-like properties. The SP phenotype has been characterized by low violet and red fluorescence intensity on a dot-plot displaying dual-wavelength fluorescence of blue versus red detected at 424 and 675 nm (Scharenberg et al., 2002), respectively.
Figure 5.4 Database survey of 59 human tumour NCI60 panel cell lines. The study allocates cell lines into four groups according to ALDH and ABCG2 degree of expression (adapted from Smith PJ; personal communication).

As shown in Figure 5.5, only A549 cells exhibited a significant SP of Hoechst 33342 low cells trailing to the origin of the plot and thus displaying low
fluorescence in both the violet and red channels. SP experiments (Wiltshire M; unpublished data) were performed with Hoechst 33342 that showed significant shifts in uptake with these cells after addition of ABCG2 inhibitor Fumitremorgen C, (FTC).

![Diagram](image)

**Figure 5.5 A549 SP-detection.** A549 cell line sustains a background of SP cells identified by reduced Hoechst 33342 uptake defined at 40 min exposure using flow cytometry (work conducted by Marie Wiltshire, at Cardiff University).

Moreover, the heterogeneity of lung cancer cell line on ALDH expression has been extensively studied using Aldefluor assay (Moreb et al., 2007) and A549 cells have been reported to have very high ALDH activity (Stuelten et al., 2010) and therefore appear to be a valid model system to test the novel AAQs.

**V.1.4 Aims of this study**

The already established labelling and assay development with DRAQ5, together with the understanding of an SP-enrichment technique using Hoechst 33342,
provides the benchmark, non-invasive cytometric tools for tracking molecular interactions in live cells. However, there is an ongoing demand and effort required to develop fluorescent, photostable dyes with high affinity for a target enzyme, as ALDH, in order to discriminate sub-populations in heterogeneous cellular samples. The potential of the anthraquinone chromophore have been already assessed with DRAQ5 (see section V.1.2) and the fundamental concept of ALDH-based cell isolation has been established with the commercially available kit ‘Aldefluor assay’ (see section I.5.6). Here, the overall goal has been to combine both these two features within the same small molecule to obtain an intrinsically fluorescent probe to track a quantifiable biofunctional readout with combined features that brings additional benefits. Furthermore, intent to increase signal to noise ratio and potentially eliminate concurrent use of pump inhibitors, as indispensable feature of Aldefluor kit, could provide an end-point detection by combining together DNA binding and cell retention capacity (refers to section I.7 features of the molecular probes design concept).

In light of the biological investigation of the cytotoxic profile, the DNA labelling affinity (see chapter IV) of the novel AAQ library of compounds (refer to chapter II for synthesis), the experimental objectives of this current chapter were to use the described cellular model to gain insights into:

1. The excitation/emission profile of focussed library of AAQs in order to validate the quality of absorbance-fluorescence spectra;

2. The use of novel probes in fluorescence imaging to investigate the ability of these molecules to penetrate fixed and live cellular systems and the stability of the fluorescence signal in the cellular environment over time;
3. To determine the spatial-temporal cellular localisation of the probe by using confocal laser scanning microscopy and time-lapse microscopy;

4. To determine the ability of the AAQ derivatives (before/after hydrolysis of the acetal moiety) to separate cells on the basis of ALDH activity by establishing a flow cytometric assay based on Aldefluor protocol; specifically, to validate the potential of finding an alternative assay kit and approach to quantify and/or separate cellular subpopulations by targeting ALDH activity.
V.2 Materials and Methods

V.2.1 Probe stock solutions

AAQs (see Table 5.2) were dissolved in DMSO (Sigma) to 20 mM stock concentration. DRAQ5 (Biostatus Ltd) was used as a control agent and dissolved in deionised water. Working solutions were prepared in PBS or HEPES (Sigma) at 20 or 40 µM.

V.2.2 Determination of fluorescence excitation/emission peaks

Probe solutions (20 µM), respectively in PBS or HEPES, were pipetted into a 1 mL quartz cuvette (Thermo Fisher Scientific). After obtaining baseline spectra in PBS to zero the instrument, absorbance measurements were obtained using a Beckmann Coulter DU 800 UV/Visible spectrophotometer. The excitation wavelengths were scanned from 280 to 800 nm. Fluorescent spectral scans were performed using a Perkin Elmer LS-50B Luminescence spectrophotometer. For emission scans, the excitation wavelength was set accordingly to the recorded values for each probe and the emission wavelength were scanned from 400 nm to 900 nm in 1 nm increments. A Xenon flash lamp was used as the light source, with the lamp energy set to high. For each data point ten measurements were acquired with the photomultiplier tube sensitivity set at 125.

V.2.3 Maintenance of mammalian cell lines

Mammalian cell lines were cultured as outlined in Table 5.2.
Table 5.2 Culture of mammalian cell lines. See Appendix II for cell media composition.

The cells were obtained from the American Type Culture Collection (ATCC) and were grown at 37 °C, 5% CO₂ and 100% humidity. Cell culture plastics were from Corning. Cells were kept at subconfluent levels and were typically passaged when 70–80% confluent. Cell passaging, Determination of live cell number and cryopreservation were assessed as previously described on sections IV.3.1.2- IV.3.1.4.

V.2.4 Cell seeding

A549 and U-2 OS cells were counted (see section IV.3.1.3) and seeded at 4 x 10⁴ cells in 1 mL of the respective media per 24-well plate well for cells to be used in experiment after 24 h. Seeded cells were left to adhere and establish in a 37 °C humidified environment of 5% CO₂ in air for overnight after seeding.

V.2.5 Imaging

All camera, shutter and stage were initially controlled by AQM 2000 software (Kinetic Imaging Ltd) before all systems were moved to MetaMorph version 7.5
(Molecular Devices). All image analyses were conducted using MetaMorph Offline version 7.5. Several microscopy platforms have been used:

1. Fluorescence microscope CBS camera system (CBS Cardiff) consisting of an inverted Axiovert S100 TV microscope (Carl Zeiss Inc.) with Hamamatsu camera (model C4742-95-12, Hamamatsu Photonics), fluorescent lamp (model ebx75 isolated, Carl Zeiss Inc.) and shutter Lambda 10-2 (Kinetic Imaging Ltd).

2. Confocal laser microscope consisting of Bio-Rad Radiance 2000MP system (Bio-Rad) linked to a Nikon TE300 inverted microscope (Nikon). Image capture and processing was primarily performed using Bio-Rad LaserSharp 2000 v4 software. Further image processing for cell localisation was performed using MetaMorph software.

3. Time-lapse microscope system composed of a HAL100 microscope (Carl Zeiss Inc.) fitted with temperature (37 °C) regulating incubator system and CO₂ (5%) supply (Solent Scientific). Illumination was controlled by a shutter (Prior Scientific) placed in front of the transmission lamp and a motorized xyz stage (Prior Scientific) permitted multi-field acquisition. Phase transmission images (x10 objective lens) were acquired with a cooled CCD (charge coupled device) camera (model 4920 COHU) every 30 sec.
V.2.5.1 AAQs labelling of fixed cells

Cells were plated in a 24 well plate as reported on section V.2.4. A 4% paraformaldehyde solution in PBS was made from a 16% stock solution (Thermo Fisher Scientific). Cells in 24 well plates were washed gently with 3x 2 mL PBS per well. The PBS was removed and 0.5 mL of 4% paraformaldehyde solution was added in each well and cells were incubated for 30 min. The paraformaldehyde was removed and cells were washed 3x 2 mL PBS per well. 1mL of PBS or HEPES was added to each well and the 24 well plates were treated respectively with 20 µM of AAQ probe or 20 µM of DRAQ5 as control. Before analysis, cells were incubated at 37 °C in the dark for 1 h and for 72 h. Samples were analysed directly without any further washing step.

V.2.5.2 AAQs loading of live cells

After seeding the cells in 24 well plates (see section V.2.4), 20 µM AAQ solutions or 20 µM of DRAQ5 solutions were added respectively in D10 media (see Appendix II). Prior to analysis, cells were incubated at 37 °C in the dark for 30 min and for 18 h. Samples were analysed directly without any further washing step.

V.2.5.3 Transmission imaging

Bright field images from live cells in full culture media were captured on the CBS camera system.
V.2.5.4 Fluorescence imaging

Images from fluorescently-stained slides were captured on the CBS camera system using the suitable filters block (Chroma Technology Corp) (see Table 5.3).

<table>
<thead>
<tr>
<th>Dye</th>
<th>Excitation (nm)</th>
<th>DCLP (nm)</th>
<th>Emission (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DRAQ5/AAQs</td>
<td>ET620/60x</td>
<td>T660LPXR</td>
<td>ET700/75m</td>
</tr>
</tbody>
</table>

Table 5.3 CBS camera filter block set-up.

V.2.5.5 Confocal laser scanning microscopy (CLSM)

CLSM was used in preference to a standard fluorescent widefield-microscopy system as it permitted the acquisition of volume limited optical sections through the cell and media. This is perfect for the situations where the fluorescent probe gave high background fluorescence. Excitation (637 nm) and emission (>665 nm) values for the probe analysis were the optimal conditions for DRAQ5 analysis. These data sets were then processed and assembled into a maximum projection image that represented all optical sections. The system enables both 2D and 3D images to be generated encoding multiple parameters such as the fluorescent intensity, which were later analysed using either the supplied software (Bio-Rad) or exported into alternative programs such as the MetaMorph program for encoding or Adobe Photoshop for presentation.
V.2.5.6 Quantitation of confocal images using MetaMorph software

All images were analysed using the MetaMorph software. Images were thresholded to remove background fluorescence, once set, the level of background threshold was maintained throughout the experiment. Fluorescence intensity was acquired in the nuclear area (pixels) and divided to the fluorescence intensity acquired in the cytoplasm. Results were sorted as nuclear against cytoplasm ratio in an Excel template sheet and plotted as histograms.

V.2.5.7 Time-lapse imaging

Time-lapse microscopy was used to monitor cells growing under normal culture conditions and also following treatment with various agents over a given time period, to monitor their morphological appearance and also to measure their mitotic properties.

U-2 OS cell growth rates and plating densities were measured by plating serial cell densities on 24 well plates (refers to section V.2.4). Static phase contrast images were recorded for 48 h on a microscope mounted CCD camera (Nikon Eclipse TS120). Immediately following treatment (unless otherwise stated) cells set up in 24 well plates, were placed onto the time-lapse microscope (Zeiss Axiosvert 200). The microscope and staging area of this microscope we enclosed in a transparent incubation chamber which maintained cell culture conditions of 37°C at 5% CO₂ (see Figure 5.6). Humidity was maintained by injecting CO₂ through distilled water inside the 24 well plates.
Figure 5.6 Typical time-lapse setup for 6 well plates analysis. Example of a time-lapse microscope incorporating a surround incubator, image adapted from http://www.irp.oist.jp/g0/equipment.php. Cells were seeded in 6 well plates, with corresponding control and treated cells.

The staging area (X, Y axis); focal plane (Z); shutter and illumination control were automated to permit unassisted image capture throughout the experiments (Prior Scientific). All images were captured as previously described by Errington et al. (Errington et al., 2005) and Marquez et al. (Marquez et al., 2003). Following initial setup for region co-ordinates; interval time and duration the software was used to automatically cycle through the experiment. The time-lapse microscope was set to capture images from one region per well (24 wells) containing colonies of cells at the correct confluency. Image capture of phase/contrast images was set to 5 min for cell cycle experiments. Image
capture was controlled through by MetaMorph software (Kinetic Imaging Ltd). Images from time-lapse capture were analysed using MetaMorph image analysis software (Universal Imaging Corp.).

V.2.6 Flow cytometry

FACS Vantage flow cytometer (Becton-Dickinson Systems) equipped with a Coherent Enterprise II laser (Coherent Inc.) simultaneously emitting at multiline UV (351–355 nm range) and 488 nm wavelength with the beams made noncolinear using dichroic separators. The laser power was regulated at 30 mW (monitored on the multiline UV output). Optics used for both systems are listed in Table 5.4.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Excitation (nm)</th>
<th>Parameter &amp; emission filter (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aldefluor</td>
<td>488</td>
<td>FL1 530 / 30</td>
</tr>
<tr>
<td>AAQs</td>
<td>488</td>
<td>FL3 695 LP</td>
</tr>
<tr>
<td>DRAQ5</td>
<td>488</td>
<td>FL3 695 LP</td>
</tr>
<tr>
<td></td>
<td>637</td>
<td>FL4 710 LP</td>
</tr>
</tbody>
</table>

Table 5.4 FACS Vantage fluorescence optics specifics. LP: longpass filter.

All parameters were collected using CellQuest software (Becton-Dickinson Systems). Forward and 90° light scatter were analyzed to exclude any cell debris. Unless otherwise stated data were collected for $1 \times 10^4$ cells using the FSC (forward scatter) parameter as the master signal.
V.2.6.1 Statistical analysis of functional capacity of selected AAQs

The statistical analysis approach used for the current study was carried out using CellQuest software (Becton-Dickinson Systems). The Kolmogorov-Smirnov (K-S) two-sample test was performed to determine if two overlaid histograms came from different populations (Young, 1977). The calculation computed the summation of the curves and obtained the greatest difference between the summation curves. The statistics, shown on Table 5.5, were reported for the portion of the histogram used in the analysis (1000 events).

<table>
<thead>
<tr>
<th>D/s(n)</th>
<th>Index of similarity for the two curves. If D/s(n) = 0, the curves are identical</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>The K-S statistic or the greatest difference between the two curves</td>
</tr>
<tr>
<td>P Value</td>
<td>Probability of D being as large as it is, given that the two selected histograms are from the same population</td>
</tr>
</tbody>
</table>

Table 5.5 K-S statistics to validate comparison between different histograms.

V.2.6.2 Aldefluor assay

The principle of Aldefluor has been extensively discussed in section I.5.6. As reported in Figure 5.7, a vial of dry Aldefluor was solubilised in 25 µL of DMSO and incubated for 1 min at room temperature. After addition of one volume of 2 M HCl, the mixture was incubated for 15 min at room temperature and, subsequently, 360 µL of Aldefluor buffer was added to quench the reaction. A549 cells were harvested and live cell number was calculated as already
reported on section V.2.3. 1x $10^6$ cells were resuspended on 1 mL of Aldefluor buffer or cell media, according to the experiment. Recommended amount of DEAB solution (in 95% ethanol) was added to the “control” tube. Subsequently, activated Aldefluor solution (5 µl per mL of sample) was added to the “test” tube and immediately after mixing 0.5 mL of the mixture were transferred to the DEAB “control” tube. Samples were incubated at 37°C for 30 and 60 min, 30 and 45 min respectively, before analysis. After incubation, cells were centrifuged at 250 g for 5 min and cell pellet was resuspended either in Aldefluor buffer or cell media and kept on ice. Flow cytometry analysis was performed as reported on section V.2.5.7 set up and data acquisition followed the manufacturer recommendations (StemCell Technologies).

**Figure 5.7** Activation methodology for Aldefluor kit. Figure adapted from StemCell Technologies protocol.

**V.3 Results**
V.3.1 UV–Vis spectral studies

UV–Vis absorbance and emission spectra were determined for LC- and HA-libraries of novel AAQs in PBS and HEPES buffer at pH 7.4. In this section only spectra in PBS are reported, but Supplementary Information V.5.A (see end of this chapter) refers to absorbance values summary in PBS and HEPES.

As already presented in section IV.3.2, the tested probes presented three key-sites for chemical modifications playing a role in absorbance-fluorescence properties (see Figure 5.8).

![Figure 5.8 Structural diversities in LC-AAQ probes.](image)

Figure 5.8 Structural diversities in LC-AAQ probes. The coloured panels show the key-sites for chemical modifications. Blue box represents DRAQ5-like side-chain, added in the non-symmetrical derivatives. These three modification points have a significant relevance on DNA binding and cytotoxic effect (see Chapter IV).

As shown in Figure 5.9, the absorbance spectra of the LC-103 compound, a symmetrical derivative containing a carbon sidechain and no OH-groups in 5,8-positions, gave maxima at 610 and 660 nm. Its congener LC-055, OH-substituted in 5,8-positions (5,8-OH), showed a single absorbance peak at 585 nm and a slight reduction on intensity was detected, due to OH insertion.
Interestingly, as reported for LC-110, addition of DRAQ5-like sidechain (see Figure 5.8, blue box) translated into increased peak intensity, compared to the symmetrical LC-103. Moreover, LC-110 absorption spectrum was observed to shift slightly to shorter wavelength, with maxima occurring at 586 and 630 nm.

![Absorbance spectra analysis of LC-103, LC-055 and LC-110 derivatives. LC-103 (black line) with LC-055 (dotted grey line, left plot) and LC-110 (dotted grey line, right plot) [all plots are normalised to DRAQ5 peak absorbance]. The red dotted line represents an arbitrary threshold for comparing peak intensity across the panels.](image)

Chemical insertion of an amide bond within the carbon sidechain (see Figure 5.8, red box) translated into an increased absorbance as reported for the symmetrical LC-106 (containing no 5,8-OH) with maxima at 586 and 630 nm. As shown in Figure 5.10, insertion of DRAQ5-like sidechain in non-symmetrical LC-112 led into loss of peak intensity, however, no peak shift was observed.  

Differently, symmetrical LC-108 and non-symmetrical LC-111 containing both 5,8-OH, had very similar spectral profile. Dramatic loss in peak intensity was
recorded in LC-108, compared to LC-106, its derivative without OH, remarking once again the role of OH in reducing the intensity of the peak.

Figure 5.10 Absorbance spectra analysis of LC-106, LC-112, LC-108 and LC-111. LC-106 (black line, left plot) with LC-112 (dotted grey line, left plot) and LC-108 (black line, right plot) with LC-112 (dotted grey line, right plot) normalised to DRAQ5 maximum absorbance. The red dotted line represents an arbitrary threshold for evaluating peak intensity.

Analysis of LC-111 (5,8-OH) and LC-112 (no 5,8-OH) in Figure 5.11 (top left) showed a red shifted spectra on LC-112. Furthermore, the absorbance ratio at peaks (see Supplementary Information V.A.1 for Abs1/Abs2 values) was >1 in LC-111, OH-substituted, and <1 in LC-112, no-OH substituted, suggesting a key role of this group on determining the spectral shape. Same spectral behaviour was reported in the derivatives containing reverse amide bond within the side chain skeleton. As reported in Figure 5.11 (top right), LC-114 (no 5,8-OH) spectra was red shifted compared to LC-113 (5,8-OH) and peak ratio was >1 in LC-114 and <1 in LC-113.
Figure 5.11 Absorbance spectra analysis of LC-111, LC-112, LC-113 and LC-114. Left plots: absorbance spectra of LC-111 (black line) with LC-112 or LC-114 (dotted grey line). Right plots: absorbance spectra of LC-113 (black line) with LC-114 or LC-112 (dotted grey line). Plots are normalised to DRAQ5 maximum absorbance. Red dotted line represents an arbitrary threshold for evaluating peak intensity.

Other interesting observation was the increased absorbance in LC-113 and 114 compared to LC-111 and 112, suggesting that reverse amide bond substitution had a positive effect on the peak intensity (see Figure 5.11 bottom plots). However, no spectral shift was related to this structural variation. Thus overall for the LC compounds the rules of absorbance of the chromophore were
addition of DRAQ5-like sidechain, together with addition of the amide bond within the second sidechain.

Figure 5.12 Structural diversities in HA-AAQ probes. The coloured panels show the key-sites for chemical modifications.

Diverse structural modifications characterised HA-AAQs as reported in Figure 5.12. Investigation of the symmetrical HA-AAQs established HA-218 as a greater absorber compared to its counterparts (see Figure 5.13).

Indeed, substitution on the secondary amine of the sidechain with a methyl- (HA-222) or a benzyl- (HA-236) led into loss of peak intensity and shape. Moreover, no remarkable spectral differences were achieved on HA-252, containing a secondary amine directly attached to the chromophore and non-symmetrical structure.
Figure 5.13 Absorbance spectra analysis of HA-AAQs. Fluorescence spectra of HA-218 and its derivatives normalised to DRAQ5 maximum absorbance. HA-218 spectrum is reported in black on the top left plot and with a dotted grey line in the other plots. Red dotted line represents an arbitrary threshold for evaluating peak intensity.
V.3.2 Fluorescence spectral studies

Emission spectra were determined for LC- and HA- libraries of novel AAQs in PBS and HEPES buffer at pH 7.4, but only spectra in PBS are reported in this section. For fluorescence values summary in PBS and HEPES refers to Supplementary Information V.5.A. Fluorescence emission were obtained at 540 nm and at the peak 1 absorbance value (see dotted lines on the plots) for each molecule.

As observed in Figure 5.14, symmetrical LC-AAQs containing a carbon sidechain were very weak fluorophores. LC-055, containing 5,8-OH modification, revealed an enhanced fluorescence profile compared to its congener LC-103. Interestingly, the insertion of DRAQ5-like moiety on LC-110 dramatically improved the absorbance and fluorescence properties of the fluorophore. However, LC-110 had an incomplete spectrum above 710 nm.

Insertion of amide bond in the sidechain is showed on Figure 5.15. Spectral analysis of non-symmetrical LC-112 and symmetrical LC-106 showed similar fluorescence intensity, revealing no difference in fluorescence by DRAQ5-like sidechain insertion.
Figure 5.14 Fluorescence spectra of carbon chain substituted LC-AAQs. Compounds were excited at 540 nm (red line) and at excitation peak 1 (blue line) normalised to peak value. The excitation wavelengths are represented by red (540) and blue (peak1, PK1) dotted line. Grey line represents the given absorbance spectra.

Signal was maintained either at 540 or at peak excitation. Same comparison between LC-111 and LC-108, respectively non-symmetrical and symmetrical AAQ-derivatives, but containing 5,8-OH, showed a reduced fluorescence on LC-108 at 540 nm, compared to LC-111.
Figure 5.15 Fluorescence spectra of LC-AAQs containing an amide bond in the sidechain. Compounds were excited at 540 nm (red line) and at excitation peak 1 (blue line) normalised to peak value. The excitation wavelengths are represented by red (540) and blue (peak1, PK1) dotted line. Grey line represents the absorbance spectra.

Moreover, 5,8-OH substitution in LC-112 and LC-108 translated into an incomplete fluorescence emission above 710 nm. Non-symmetrical LC-AAQs (see Figure 5.16) containing amide and reverse amide bond in the sidechain showed a good profile on fluorescence emission.
Figure 5.16 Fluorescence spectra of non-symmetrical LC-AAQs containing amide and reverse amide bond in the sidechain. Compounds were excited at 540 nm (red line) and at excitation peak 1 (blue line) normalised to peak value. The excitation wavelengths are represented by red (540) and blue (peak1, PK1) dotted line. Grey line represents the absorbance spectra.

However, 5,8-OH substitution in LC-111 and LC-114 gave incomplete emission spectra above 710 nm and a noisier profile at 540 nm compared to the no 5,8-substituted LC-112 and LC-113. Ultimately, symmetrical HA-AAQs fluorescence emission analysis revealed very different spectral profile, depending on the degree of substitution on the secondary amine (see Figure 5.12 for structural
modifications). HA-218 revealed the brightest fluorescence signal compared to its derivatives (see Figure 5.17).

![Figure 5.17 Fluorescence spectra of HA-AAAQs containing a secondary amine on the sidechain.](image)

Compounds were excited at 540 nm (red line) and at excitation peak 1 (blue line) normalised to peak value. The excitation wavelength is represented by red (540) and blue (peak1, PK1) dotted line. Grey line represents the absorbance spectra.

Indeed, insertion of methyl group on the sidechain skeleton led to loss absorption and consequently to a very little fluorescence, whereas benzyl group addition led to no fluorescence. Non-symmetrical HA-252, containing a
secondary amine on the sidechain directly linked to the chromophore, demonstrated not enhanced fluorescence properties compared to HA-218.

V.3.3 Screening of AAQs labelling in fixed U-2 OS cells

Wide-field fluorescence microscopy was used to evaluate the cell labelling in fixed U-2 OS cell model, this essentially evaluated the ability of the fluorescent agent to label cellular compartments without active uptake or efflux prohibiting delivery into the given cells. First image collection was performed after 1 h of exposure to probes in PBS, as reported in Supplementary Information V.B.1 at the end of the chapter. The structural counterparts LC-055, LC-103 and LC-110, featuring a carbon sidechain, showed diverse staining when put into cells. LC-055 gave a very dim cytoplasmic fluorescence, whereas LC-103 and LC-110, both not containing 5,8-OH substitution, had a diffuse cytoplasmic staining, with some perinuclear and nucleoli labelling. Moreover, very bright background was observed in LC-110 compared to LC-103.

The second group of LC-derivatives, with amide functionality inserted in the side chain scaffold, included LC-106, LC-108 (symmetrical) and LC-111, LC-112 (non-symmetrical). Both LC-106 and LC-108 revealed high background fluorescence and very little contrast; LC-108 showed some precipitation, as little dark spots. The non symmetric derivatives LC-111 and LC-112, containing a DRAQ5-like sidechain, showed heterogeneity in cytoplasmic staining and positive spots in the nuclei, probably due to DNA binding. LC-112 had better contrast than LC-111. Moreover, LC-113 and LC-114, containing the reverse
amide feature on the sidechain, showed very high background fluorescence, especially in LC-113, with cytoplasmic staining and heterogeneity on labelling of mitotic figures.

Analysis of the HA-AAQ derivatives reported a very bright fluorescence on HA-218 with cytoplasmic staining, good object recognition, and some heterogeneity on nuclear staining. On the other hand, HA-222 and HA-252 had a perinuclear staining and the majority of the cells had no DNA label, whereas HA-236 had gave a very dim staining almost no detectable also with longer exposure time.

After 72 h of incubation with the probes, images were recollected to assess the fluorescence stability in PBS and the degree of labelling over time, as reported in Supplementary Information V.B.2. LC-103 and LC-110 revealed loss of fluorescence, but some labelling was still detectable, whereas LC-055, LC-106 and LC-108 showed no detectable fluorescence. Addition of DRAQ5-like sidechain rescued LC-111, LC-112 and LC-114 cytoplasmic signal, characterised by diffused background fluorescence. However, enhanced nuclear staining was detected in LC-113 compared to 1 h labelling and the high background fluorescence was maintained. HA-AAQs had stable fluorescence, with loss of signal for HA-218, but perinuclear and cytoplasmic labelling was revealed for the other derivatives. Better staining was observed in HA-236 compared to 1 h incubation, whereas HA-222 and HA-252 had the same staining intensity detected after 1 h.
V.3.4 Screening of AAQs loading in U-2 OS live cells

In order to understand whether AAQs were subjected to cellular uptake and possible compartmentalisation, cells were initially incubated with the respective probes for 30 min. All images were collected at the same exposure time (30 ms) to ensure good comparison; these images are shown in Supplementary Information V.B.3.

The counterparts LC-055, LC-103 and LC-110 showed high background and low contrast. In particular, LC-055 had some precipitation, probably due to low solubility in media, whereas LC-103 and LC-110 had cytoplasmic staining with labelling of some mitotic figures and bright perinuclei. LC-106 and LC-108 had no live cell access with bright background fluorescence. LC-108 presented some precipitate in solution. Counterparts with DRAQ5-like sidechain, LC-111 and LC-112 had a similar staining pattern, with heterogeneous staining in nucleus and cytoplasm and background fluorescence, however LC-111 had brighter signal than LC-112. Comparing LC-113 and LC-114 analogues, both revealed high background and nuclear staining, but LC-114 had some precipitate. The HA-derivatives showed less cell loading and predominant cytoplasmic staining compared to LC-AAQs. HA-222 had really bright background fluorescence, no observed for the other HA-probes.

After 18 h the images were recollected to determine whether cell viability was affected by probe loading (bright-field microscopy) and to measure penetration properties over time. The screenings at 18 h are reported in Supplementary Information V.B.4-5.
Image collection in transmission showed healthy and flat cells after exposure to LC-055 and LC-103, however no fluorescence staining was detected on LC-055 but a relative good load was maintained on LC-103. Cells exposed to LC-110 were very rounded and dense, suggesting loss of viability, but stable cytoplasmic staining was still detectable. LC-106 and LC-108 did not affect cell viability because of no live cell penetration, indeed no loading was observed over time. After 18 h staining with LC-111, LC-112, LC-113 and LC-114 cells were no viable and therefore retain of fluorescence was really low. Complete loss was observed in LC-112 together with HA-222, HA-236 and HA-252, as consequence of cell death. HA-218 reported intracellular accumulation and stacking over time, however cells appeared no viable.

V.3.5 Screening of AAQs loading in A549 live cells

A549 cell line was used as a different cell model to understand whether presence of SP cell fraction led heterogeneity on cell penetration and to a consequent variation of the probe loading and localisation.

Cells were incubated for 30 min with the respective probe and then images were collected at 30 ms exposure time (see Supplementary Information V.B.6). LC-055, LC-106 and LC-108 showed no cell penetration and no intracellular fluorescence signal was detected. LC-103 and LC-110 had high background and heterogeneous cytoplasmic staining. The LC-111, LC-112 and LC-113 analogues revealed very bright background fluorescence, whereas LC-114 had a greater contrast and diffuse cellular staining. The HA-derivatives showed
heterogeneity on cytoplasmic load on HA-218, with bright fluorescence signal and staining of mitotic figures, whereas HA-222 had perinuclear accumulation and much dimmer signal intensity. HA-236 had no cell loading and very low loading was detected on HA-252.

After 18 h of continuous incubation with the probes A549 were screened again to ask for cellular viability and retain of fluorescence signal, as reported in Supplementary Information V.B.7-8. Cells, incubated with LC-055, LC-106 and LC-108, were healthy and monolayered as no uptake was shown; furthermore, LC-108 showed lots of precipitation. LC-103 exposure gave healthy and flat cells retaining perinuclear fluorescence over time, whereas cells exposed to LC-110 were almost all dead but fluorescent signal was still retained. LC-111, LC-112, LC-113 and LC-114 produced high toxicity and therefore very rounded and granular cells; however perinuclear accumulation was maintained on LC-112 and LC-113, whereas most of the fluorescence signal was not retained in LC-111 and LC-114. HA-236 exposure showed healthy cells as the dye was no loaded, whereas high cytotoxicity was observed on the HA-218, HA-222 and HA-252 with retain of perinuclear fluorescence.

V.3.6 Confocal laser microscopy analysis in the U-2 OS cell line

Confocal laser scanning microscopy was used to collect novel probe signals after 30 min incubation at different z depth and eliminating the solution background. Nucleus/cytoplasm ratio was calculated to quantify the localisation of the probe per cell.
As reported in Figure 5.18, confocality allowed detection of low levels of probe in the perinuclear region, with probable overload in Golgi body and large vesicles for LC-055 and LC-103. The accumulation area appeared similarly to the MTOC (microtubule organising centre), however LC110, containing DRAQ5-like sidechain, gave additional nuclear signal.

Analysis of the counterparts LC-106 and LC-108 confirmed a very low uptake and intense background fluorescence, as already observed in the live cell screening, whereas its derivatives containing a DRAQ5-like sidechain had a greater nuclear localisation, with some vesicular perinuclear staining (see Figure 5.19). As reported on the nucleus/cytoplasm graph, the localisation was predominantly nuclear and this caused high cellular toxicity, as already shown on the 18 h incubation screening (see Supplementary Information V.B.4-5).
Figure 5.18 LC-055, LC-103 and LC-110 fluorescence analysis via direct confocal imaging. Analysis was performed using excitation (637 nm) and emission (>615 nm) wavelength. Plot reports the nucleus/cytoplasm index to determine DNA binding in live cells. Calibration bar represents 10 µm.
Figure 5.19 LC-112, LC-106, LC-111 and LC-108 fluorescence analysis via direct confocal imaging. Analysis was performed using excitation (637 nm) and emission (>615 nm) wavelength. Plot reports the nucleus/cytoplasm index to determine DNA binding in live cells. Calibration bar represents 10 µm.
As reported in Figure 5.20, LC-derivatives containing amide bond on the sidechain skeleton revealed a strong nuclear uptake and perinuclear staining, with possible ER accumulation. A whole cell staining was a specific feature of LC-113 and LC-112, whereas LC-114 was mainly loaded in the nuclear compartment. Black spots were detected on nuclei after LC-111 staining, due to probe accumulation.

The HA-series, as shown in Figure 5.21, appeared to have a predominant cytoplasmic localisation. However, HA-218 had a nuclear signal and a reticular perinuclear pattern, probably with ER compartmentalisation. As already observed for LC-055 and LC-103, HA222 and HA236 revealed large vesicular accumulation and no nuclear staining. HA218 had a characteristic quenching of signal due to accumulation of the probe, as demonstrated by detection of black perinuclear spots in transmission.
Figure 5.20 LC-111, LC-112, LC-113 and LC-114 fluorescence analysis via direct confocal imaging. Analysis was performed using excitation (637 nm) and emission (>615 nm) wavelength. Plot reports the nucleus/cytoplasm index to determine DNA binding in live cells. Calibration bar represents 10 µm.
Figure 5.21 HA-AAQs fluorescence analysis via direct confocal imaging. Analysis was performed using excitation (637 nm) and emission (>615 nm) wavelength. Plot reports the nucleus/cytoplasm index to determine DNA binding in live cells. HA-218 in transmission reports the same field analysed in fluorescence, black spots represent probe accumulation. Calibration bar represents 10 µm.
V.3.7 Uptake of AAQs in A549 cell model and ABCG2-dependent efflux

A549 uptake profile was investigated for selected AAQs and cells were incubated for 30 min with 20 µM of the selected probe before performing flow cytometry detection of the single cell signal.

* 50% of total cell number

Figure 5.22 Flow cytometry setup for LC- and HA-AAQs analysis. Dynamic range was set on DRAQ5 (red curve) parameters. The inset represents DRAQ5 dot plot distribution, used as positive control. Light blue line represents empty cells and dark blue line represents LC-111 uptake, taken as AAQ exemplar. The dotted back line is the median of the distribution and represents 50% of the total cell population.
As shown in Figure 5.22, DRAQ5 was used as to set the dynamic range and for flow cytometry detection setup for anthraquinone-based fluorophores (FL-3, see section V.2.5.7 for parameters details). However, DRAQ5 and AAQs distribution plot were not comparable because of cellular DNA content-related histogram of DRAQ5 (Smith et al., 2000). Forward- and side-scattered signal was collected for 10,000 cells and were analyzed to exclude any cell debris. Analysis was conducted using the viable and intact fraction of cells, corresponding to ~95% of the total amount under investigation.

![Figure 5.23 Uptake of selected AAQs on A549 cell line by flow cytometry analysis. Histograms represent the geometric mean of fluorescence intensity ± SD. Dynamic range was set on DRAQ5 parameters.](image-url)
As shown in Figure 5.23, the population cellular uptake was obtained for all the selected probes with a median fluorescence between 36.93 and 72.93 (see Supplementary Information V.C.1 for detailed analysis). Uptake above background comparison revealed a median fluorescence for empty cells of 19.27 versus an average of 50 for the cells pre-incubated with the probe. Within selected LC-derivatives, LC-111 showed a lower uptake than its derivatives LC-110 and LC-114 (median fluorescence of 36 vs. 50-70). The low loader sub-fraction of cells was explored and heterogeneity on the probe loading was reported (see 25%ile data on Supplementary Information V.C.1) probably dependent on the size of the cells. Moreover, the scatter properties before/after exposure to the AAQ probes were maintained even if highly dense particles were introduced to the system.

As reported in Figure 5.24, functional profiling of ABCG2 activity was performed in ABCG2-expressing A549 cell line using flow cytometry detection of the uptake/exclusion of the selected AAQs with or without the presence of the ABCG2 inhibitor, fumitremorgin C (FTC, Sigma). During this time the SP experiments were performed with Hoechst 33342 that showed significant shifts in uptake with these cells and FTC batch. In the tested compounds, uptake of LC-110 was higher in cells without FTC compared to the FTC-treated cells (median fluorescence of 72.01 vs. 55.95), revealing that the uptake was less sensitive to FTC addition, suggesting a potential non-ABCG2 mediated efflux. LC-111, instead, had a better cell loading in cells treated with FTC compared to the non-treated cells (median fluorescence of 36.93 vs. 41.32) consistent with blocking the ABCG2-dependent exclusion. K-S statistical analysis was used to
validate the dissimilarity between the paired histograms (refers to V.2.5.7 for details).

<table>
<thead>
<tr>
<th></th>
<th>LC-110</th>
<th>LC-111</th>
<th>HA-218</th>
<th>HA-222</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>D/s(n)</strong></td>
<td>25.59</td>
<td>15.30</td>
<td>3.12</td>
<td>1.49</td>
</tr>
<tr>
<td><strong>D</strong></td>
<td>0.37</td>
<td>0.22</td>
<td>0.04</td>
<td>0.02</td>
</tr>
<tr>
<td><strong>P value</strong></td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.05</td>
</tr>
</tbody>
</table>

* K-S statistic analysis on AAQs uptake +/- FTC pre-incubation

Figure 5.24 AAQ uptake profiles in the A549 cell line +/- FTC pre-treatment. The cells were pre-incubated for 1 h with 10 µM FTC or DMSO, before incubation with 20 µM of the respective AAQ (30 min). The results represent an average of three replicates ± SD. P values represent the significance of the difference between the paired histograms.

V.3.8 Hydrolysis effect on cellular uptake of LC-AAQ derivatives

Flow cytometry analysis was performed to evaluate the effect of pre-treatment with HCl to generate an active aldehyde (refers to section III.3.1 for further details on the chemical hydrolysis) on the cell uptake properties. A549 cells
were incubated for 30 min with selected parental/hydrolysed LC-AAQ (20 µM) prior to analysis. Furthermore, the rate of uptake was tested in A549 cell media (see section V.2.3 for cell media details) and in Aldefluor buffer, containing the pan MDR-blocker verapamil (Pearce and Bonnet, 2007), to evaluate the effect in uptake with/without the Aldefluor assay system. Analysis was conducted in the viable subset of cells, corresponding to ~95% of the total amount under investigation.

As reported in Figure 5.25, no difference in cellular uptake was observed for the parental AAQs with or without addition of verapamil. However, there did seem to be some enhanced efflux of hydrolysed AAQs (HAAQs) in the absence of the ABC pump inhibitor verapamil. Hydrolysed LC-110 and LC-111 showed a statistically significant greater uptake in Aldefluor buffer than in media (median fluorescence of 102.97 vs. 64.9 for LC-110 and 51.51 vs. 61.44 for LC-111) whereas LC-114 had similar uptake level in both conditions (refer to Supplementary Informations V.C.2 for detailed analysis).
**Figure 5.25** Parental/hydrolysed AAQ uptake profiles in the A549 cell line in media vs. Aldefluor buffer. Aldefluor buffer contained verapamil vs. cell media. The cells were incubated with 20 µM of the respective AAQ/HAAQs (30 min) before analysis. The results represent an average of three replicates ± SD. P values represent the significance of the difference between the paired histograms.
V.3.9 ALDH Activity by Aldefluor flow cytometry assay in A549 cells

The amount of Aldefluor positive cells (%) was determined by placing the gate in the region based on the shift of fluorescent cells seen after DEAB (inhibitor of ALDH activity) treatment. As shown in Figure 5.26A-B, high ALDH activity in A549 cells was observed after 30 min of incubation with Aldefluor before analysis (10.71% in Figure 5.26A, 5.99% in Figure 5.26B).

Fluorescence shifting with DEAB addition was maintained and appeared to be more prominent at 30 min (0.47%), compared to 45 and 60 min incubation with Aldefluor reagent. Indeed, shoulder on the histogram without DEAB started disappearing after 45 min (6.4%) and it was almost cleared after 60 min (2.77%) of exposure to Aldefluor.

A)
B)

Figure 5.26 Histograms of Aldefluor uptake +/- DEAB. Plots represent uptake analysis with (grey curves) or without DEAB (black curves) for A549 cells, after incubation with the reagent for 30 and 45 min (A) and 30 and 60 min (B) before flow cytometry analysis.

The low activity was considered as an artefact of very high ALDH expression level in A549, as already reported by Moreb et al. (Moreb et al., 2007). Same experiment was performed using cell media instead of Aldefluor buffer to assess whether the presence of verapamil was crucial for Aldefluor retention within the cells and no fluorescence was detected.
V.4 Discussion

In summary, novel AAQs have been explored as functional probes in order to investigate potential targets within selected cell models. Here, the study focused on determining the structural related rules of the AAQ library previously developed to their respective absorbance/fluorescence properties, the labelling properties in permeabilised/live cellular models, the probe loading in two different cell models with/without active pump, the possible compartmentalisation in specific subcellular locations and the single cell uptake.

Similar studies, conducted on the development of DRAQ5 as functional assay (Smith et al., 2000), revealed that an anthraquinone-based fluorophore had a characteristic absorbance spectrum with a double peak shape and peak ratio (Abs1/Abs2>1 in OH-derivatives and Abs1/Abs2<1 in H-derivatives) depending on the intramolecular hydrogen bonding of the 5,8- OH substitution (Marasinghe and Gillispie, 1989). This chapter elucidated how minimal variations on the molecular skeleton were translated into a consistent alteration of the absorbance profile, as a red-shift due to the 5,8- OH insertion, probably depending on the strength of the hydrogen bond between the hydroxyl groups and the carbonyl groups of AAQs and the solvent molecules (Inoue et al., 1982; Yoshida and Takabayashi, 1968). Interestingly, increased peak intensity was obtained by addition of DRAQ5-like features or amide bond within the sidechain, whereas 5,8- OH substitution reduced the absorbance intensity. These features were crucial to obtain a significant fluorescent emission. Spectral studies on DRAQ5 acknowledged this probe as a “weak” far-red
fluorophore with a low quantum yield (Smith et al., 1999) and comparison with the novel AAQs showed comparable fluorescent emission profile between 650 and >800 nm when excited at the corresponding absorbance peak. However, comparison of the emission spectra revealed that LC-106, LC-112 and LC-113 probes had a quite intense fluorescent signal compared to the other derivatives. Same analysis on HA-AAQ family, containing diverse substitution patterns on the secondary amine of the sidechain, revealed an intense absorbance spectrum only for HA-218, however the fluorescence spectra were quite noisy for all the HA-AAQ counterparts, suggesting that HA-AAQs were less emitting compared to the LC-AAQs. No significant spectral differences were observed by using HEPES as a working buffer instead of PBS, confirming no effect of the interaction with phosphates on the AAQ stability (Jeon et al., 2004). Furthermore, fluorescence measurement of the tested AAQs after 15 months in DMSO solution showed the same intensity of the signal, suggesting a stable shelflife (data not shown; personal communication Sally Chappell) and therefore a good potential as fluorophore.

The chapter reported the screening of the AAQs ability to stain fixed and thus permeabilised cells and to show affinity for a specific cellular compartment at different timepoints. DRAQ5 acted as the benchmark agent. We found that AAQs containing a carbon sidechain produced a localised cytoplasmic staining, and, as elucidated in the spectral studies, greater cellular signal was observed in LC-110 with addition of DRAQ5-like sidechain, revealing once again the “rescuer” role of this modification for the fluorescence signal. These data confirmed the poor affinity for the CT-DNA showed on DNA binding analysis
(Rahban et al., 2010), suggesting a preferred perinuclear compartmentalisation. Moreover, insertion of an amide bond in the symmetrical derivatives produced heterogeneity in the staining and high background, i.e. in LC-106. The high background could have been removed by washing out the dye after incubation; however potential loss of fluorescence signal could be a side effect of a washing step. Combination of amide bond in the sidechain and DRAQ5-like modification (i.e. LC-111 or LC-112) showed a probe that located to the nucleus while maintaining some perinuclear labelling, with heterogeneity on cell compartmentalisation. The structural modifications on HA-AAQs derivatives led to a cytoplasmic labelling; however, DNA-binding data showed that HA-AAQs had a greater affinity for the CT-DNA than some nuclear binders LC-AAQs, suggesting a substantial difference in affinity between highly purified CT-DNA and chromosomal DNA, associated with several different positively charged proteins (histones) (Faulds et al., 1991), and therefore less accessible to the HA-AAQ probes binding. This study demonstrated a stable fluorescence over time for most of the probes tested, however solubility issues occurred with LC-108 and precipitation of unbound dye increased after 72 h incubation in PBS.

It has been well established that observing a biological event as it unfolds in the living cell offers an unique insight into the nature of the phenomenon under study (Tsien, 1998). We found that in live U-2 OS and A549 cell models to evaluate cell penetration and potential loading levels depending on the expression of membrane pumps. In general, as reported in Figure 5.27, several of the fluorescent probes were successful in the labelling of cellular
compartments and this could vary with our chosen two cell-types (Chen H, 2010). U-2 OS cell line uptake revealed a complete exclusion from the cells of some probes, which consequently were not taken forward for further live cell imaging studies, as LC-106 and LC-108.

Figure 5.27 Overview of the general principle for live cell imaging. Probes are constructed to label specific targets; the overall loading and localisation of fluorescence can be detected using fluorescence microscopy (Chen H, 2010).

Less background fluorescence was found in A549 cell line, suggesting that the high expression of membrane transporter (Scharenberg et al., 2002) was helping the probe loading, therefore the cell model needed to be chosen carefully as membrane effluxers could play an important role on the final
biodistribution of the probe (Chen H, 2010). Intrinsic differences on cell morphology between the selected cell models needed to be bared in mind when wide-field screening was performed. Indeed, U-2 OS are typically flat cells, whereas A549 cells are characterised by protruding nuclei and very rounded shape (Hermanns et al., 2004).

The study suggested that nuclear localisation affected cell viability after 18 h for LC-AAQs, as already assessed via cytotoxicity studies. Indeed exposure of A549 to LC-111 showed low µM IC_{50} value after 72 h (refers to Chapter IV) and in the present study it was confirmed that cell shape was affected already at 18 h indicating an alteration on cell viability.

We have determined the spatial-temporal cellular localisation of the probe by using confocal laser scanning microscopy and time lapse microscopy and we quantified the ratio of the mean nucleus/cytoplasmic fluorescence ratio to rank the compounds with the localisation. LC-111,112,113 and 114 reported a nuclear affinity, suggesting a DNA affinity, together with some perinuclear staining, which reflected the high cytotoxic effect on cell viability assessment. Different affinity has been found in live cell tracking of HA-AAQ derivatives; these molecules confirmed the same cytoplasmic affinity shown in fixed cells and a reticular labelling, suggested Golgi compartmentalisation. Intriguingly, we demonstrated on HA-218 analysis how stacking of the probe after cell loading translated into fluorescence signal quenching and interaction with cell viability.

Moreover, we proved how cytoplasmic and Golgi affinity of LC-110 led to comparably high cytotoxicity, suggesting once again that also cytoplasmic
compartimentalisation could be crucial for cell viability. Using time-lapse microscopy, (Dailey M.E., 2006), we analysed the probe redistribution in U-2 OS cell line after mitotic event at 48 h and we confirmed a redistribution and probe segregation to the two daughter cells, even if the probe was highly segregated in subcellular organelle, possibly in endoplasmic reticulum, Golgi or endosomes (Bergeland et al., 2001). Observation of the cell shape obtained using time-lapse microscopy, confirmed the data obtained in section IV.3.2, regarding the AAQs effect on cell health status.

Interestingly, LCw103, a symmetrical derivative with carbon chains, had no detrimental effect on cell viability, as proved by imaging analysis, timelapse detection at 48 h and, already discussed, MTT assessment after 72 h (see Chapter IV). This probe revealed a consistent and unique accumulation in endoplasmic reticulum (ER), suggesting this compound as a potential candidate as ER-tracker.

On the basis of the data collected on the imaging and spectral analysis, we selected five AAQ-derivatives characterised by bright fluorescence, good cell loading and unique cell localisation patterns to explore the single cell loading on A549 cells via flow cytometry analysis. After assessing a positive loading for the panel of AAQs, our studies were directed to better understanding the implication of ABCG2 transporters on the probe uptake, on the basis of the high affinity between BCG2 and anthraquinones (Doyle and Ross, 2003). Fumitremorgin C (FTC), a potent inhibitor of the ABCG2 transporter (Rabindran et al., 2000), provided a selective tool for evaluating the influence of the transporter in AAQs uptake. Experiment limitations were not ideal since the collection conditions
could have been improved to extend the dynamic range giving a more sensitive assay for the effect of ABCG2 inhibition. Similar or slightly increased uptake was demonstrated after addition of FTC, suggesting a potential ABCG2-dependent efflux for the investigated compounds. Surprisingly, LC-110 revealed reduced uptake with FTC addition suggesting the potential prevailing of efflux mediated by different pumps. Indeed, its unique uptake profile was cleared by using verapamil. In this chapter we evaluate the potential to identify unique populations of A549 cells, possibly related to the SP subfraction, using the novel LC-AAQs (before/ after hydrolysis of the acetal moiety) on the basis of ALDH activity - by establishing a flow cytometric assay based on Aldefluor protocol. HA-derivatives were not considered because previous hydrolysis studies revealed resistance to the proposed hydrolysis conditions (refers to section III.2.1.2). The proposed study wanted to specifically validate the potential of finding an alternative assay kit and approach to quantify and/or separate cellular sub-populations by targeting ALDH activity. Our findings revealed no difference in uptake +/- verapamil, a general pump blocker (Pearce and Bonnet, 2007), whereas after hydrolysis the role of the concurrent use of verapamil became crucial for enhancing intensity signal on LC-110 and LC-111. However, no significant difference was reported for LC-114 suggesting less affinity for the exporters for LC-114 or possibly less aldehyde formation after hydrolysis. Although ABCG2-driven SP functionality was not proved, here we demonstrate a potential substrate specificity of the tested AAQs thanks to a greater signal in Aldefluor-like conditions rather than normal cell media.
We have not yet demonstrated a specific affinity for ALDH family; however this study led to the characterisation of interesting novel AAQs, pointing out probe-like features in these molecules (i.e. LC-103, LC-111 and LC-114).

Overall this study provided a solid screening of the potential of these novel compounds and interesting features were reported in selected ones as high cellular penetration, discrete and adjustable compartmentalisation, and ability to maintain cell viability after loading. These molecules allowed the simultaneous or differential labelling of both nuclear and cytoplasmic compartments in live and fixed cells to clearly render the precise location of cell boundaries which may be beneficial for quantitative expression measurements, cell-cell interactions, and most recently compound *in vitro* toxicology testing (Edward, 2012).
Supplementary Information V.A

AAQs UV-Vis absorbance (1) and fluorescent emission (2) values

Peak 1 (λ1) and peak 2 (λ2) were measured. The molecules were excited at the measured UV-absorbance wavelength (Pk1, Pk2) and correspondent emitted fluorescence values were recorded (Pk1, Pk2). DRAQ5 was used as control.

1) UV-absorbance in **PBS** (nm)  |  UV-absorbance in **HEPES** (nm)

<table>
<thead>
<tr>
<th>Compound</th>
<th>λ1 (nm)</th>
<th>λ2 (nm)</th>
<th>Abs1/Abs2</th>
<th>λ1 (nm)</th>
<th>λ2 (nm)</th>
<th>Abs1/Abs2</th>
</tr>
</thead>
<tbody>
<tr>
<td>LC-055</td>
<td>585</td>
<td>-</td>
<td>NA</td>
<td>587</td>
<td>634</td>
<td>1.36</td>
</tr>
<tr>
<td>LC-108</td>
<td>610</td>
<td>665</td>
<td>1.11</td>
<td>574</td>
<td>610</td>
<td>1.09</td>
</tr>
<tr>
<td>LC-111</td>
<td>610</td>
<td>665</td>
<td>1.12</td>
<td>610</td>
<td>660</td>
<td>1.14</td>
</tr>
<tr>
<td>LC-114</td>
<td>610</td>
<td>664</td>
<td>1.2</td>
<td>610</td>
<td>664</td>
<td>1.1</td>
</tr>
<tr>
<td>HA-218</td>
<td>610</td>
<td>670</td>
<td>1.43</td>
<td>607</td>
<td>663</td>
<td>1.38</td>
</tr>
<tr>
<td>HA-222</td>
<td>610</td>
<td>660</td>
<td>1.08</td>
<td>610</td>
<td>664</td>
<td>1.23</td>
</tr>
<tr>
<td>HA-236</td>
<td>650</td>
<td>-</td>
<td>NA</td>
<td>610</td>
<td>-</td>
<td>NA</td>
</tr>
<tr>
<td>HA-252</td>
<td>550</td>
<td>-</td>
<td>1.5</td>
<td>564</td>
<td>606</td>
<td>1.33</td>
</tr>
<tr>
<td>DRAQ5</td>
<td>600</td>
<td>647</td>
<td>1.2</td>
<td>600</td>
<td>647</td>
<td>1.2</td>
</tr>
</tbody>
</table>

- H in 5,8 position

<table>
<thead>
<tr>
<th>Compound</th>
<th>λ1 (nm)</th>
<th>λ2 (nm)</th>
<th>Abs1/Abs2</th>
<th>λ1 (nm)</th>
<th>λ2 (nm)</th>
<th>Abs1/Abs2</th>
</tr>
</thead>
<tbody>
<tr>
<td>LC-103</td>
<td>610</td>
<td>660</td>
<td>0.98</td>
<td>600</td>
<td>650</td>
<td>0.95</td>
</tr>
<tr>
<td>LC-106</td>
<td>586</td>
<td>630</td>
<td>0.81</td>
<td>585</td>
<td>630</td>
<td>0.85</td>
</tr>
<tr>
<td>LC-110</td>
<td>586</td>
<td>630</td>
<td>0.92</td>
<td>586</td>
<td>630</td>
<td>0.94</td>
</tr>
<tr>
<td>LC-112</td>
<td>584</td>
<td>628</td>
<td>0.87</td>
<td>583</td>
<td>627</td>
<td>0.88</td>
</tr>
<tr>
<td>LC-113</td>
<td>583</td>
<td>627</td>
<td>0.92</td>
<td>583</td>
<td>628</td>
<td>0.90</td>
</tr>
</tbody>
</table>
2)

<table>
<thead>
<tr>
<th>Compound</th>
<th>λ1 Ex</th>
<th>540 Ex</th>
<th>λ1 Ex</th>
<th>540 Ex</th>
</tr>
</thead>
<tbody>
<tr>
<td>LC-055</td>
<td>747.5</td>
<td>-</td>
<td>740</td>
<td></td>
</tr>
<tr>
<td>LC-108</td>
<td>684</td>
<td>712</td>
<td>726</td>
<td>NA</td>
</tr>
<tr>
<td>LC-111</td>
<td>685</td>
<td>692</td>
<td>684</td>
<td>-</td>
</tr>
<tr>
<td>LC-114</td>
<td>684</td>
<td>684</td>
<td>684</td>
<td>-</td>
</tr>
<tr>
<td>HA-218</td>
<td>688</td>
<td>690</td>
<td>684</td>
<td>-</td>
</tr>
<tr>
<td>HA-222</td>
<td>690.5</td>
<td>684</td>
<td>684</td>
<td>-</td>
</tr>
<tr>
<td>HA-236</td>
<td>-</td>
<td>-</td>
<td>680</td>
<td>-</td>
</tr>
<tr>
<td>HA-252</td>
<td>678</td>
<td>690</td>
<td>696</td>
<td>-</td>
</tr>
<tr>
<td>DRAQ5</td>
<td>&gt;665</td>
<td>635</td>
<td>&gt;665</td>
<td>-</td>
</tr>
</tbody>
</table>

- OH in 5,8 position

- H in 5,8 position
Supplementary Information V.B

Screening of LC- and HA-AAQs in fixed and in live U-2 OS and A549 cell models
**V.B.2** Labeling of fixed U-2 OS cells in PBS after a 72 h incubation

Screening of LC- and HA-AAP probes for localisation and fluorescence stability over time using low-resolution wide-field fluorescence microscopy. DRAQ5 is used as benchmark. Calibration bar represents 50 μm.
V.B.3 Loading of live U-2 OS cells in media after a 30 min incubation

Screening of LC- and HA-AAA probes for cell uptake properties using low-resolution wide-field fluorescence microscopy. DRAQ5 is used as benchmark. Calibration bar represents 50 μm.
V.B.4 Loading of five U2 OS cells in media after a 18 h incubation

Screening of LC- and HA-AAG probes for cellular viability and shape, using low resolution bright-field microscopy. DRAQ5 is used as benchmark. Calibration bar represents 50 μm.
V.B.5 Loading of live U2 OS cells in media after a 16 h incubation

Screening of LC- and HA-AAQ probes for cellular viability and localisation of the probe, using low resolution wide-field fluorescence microscopy. DRAQ5 is used as benchmark.

Calibration bar represents 50 μm.
V.B.6 Loading of live A549 cells in media after a 30 min incubation

Screening of LC- and HA-AAC probes for cell uptake and localisation of the probe, using low resolution wide-field fluorescence microscopy. DRAQ5 is used as a benchmark.

Calibration bar represents 50 μm.
V.B.7 Loading of live A549 cells in media after a 10 h incubation

Screening of LC- and HA-AAG probes for cellular viability and shape, using low resolution bright-field microscopy. DRAQ5 is used as a benchmark. Calibration bar represents 50 μm.
V.B.9 Loading of Hve A549 cells in media after a 18 h incubation

Screening of LC- and HA-AAQ probes for cellular viability and localization of the probe, using low resolution wide-field fluorescence microscopy. DRAQ5 is used as benchmark. Calibration bar represents 50 μm.
Supplementary Information V.C

V.C.1 Statistical analysis of A549 uptake +/- Fumitremorgin C (FTC). Data were processed using FlowJo software (Tree Star). DRAQ5 was used to set the dynamic range and as control dye and empty cells were used to determine auto-fluorescence.

<table>
<thead>
<tr>
<th>A549 sample</th>
<th>Median</th>
<th>Covariance</th>
<th>SD</th>
<th>% ile (25) *</th>
</tr>
</thead>
<tbody>
<tr>
<td>Empty</td>
<td>19.27</td>
<td>31.07</td>
<td>31.07</td>
<td>16.3</td>
</tr>
<tr>
<td>DRAQ5</td>
<td>192.79</td>
<td>35.73</td>
<td>35.73</td>
<td>162.03</td>
</tr>
<tr>
<td>LC-110</td>
<td>72.01</td>
<td>41.06</td>
<td>41.06</td>
<td>60.2</td>
</tr>
<tr>
<td>LC-110 + FTC</td>
<td>55.95</td>
<td>45.07</td>
<td>45.07</td>
<td>47.27</td>
</tr>
<tr>
<td>LC-111</td>
<td>36.93</td>
<td>30.12</td>
<td>30.12</td>
<td>31.99</td>
</tr>
<tr>
<td>LC-111 + FTC</td>
<td>41.32</td>
<td>31.48</td>
<td>31.48</td>
<td>36.03</td>
</tr>
<tr>
<td>HA-218</td>
<td>54.05</td>
<td>35.94</td>
<td>35.94</td>
<td>46.3</td>
</tr>
<tr>
<td>HA-218 + FTC</td>
<td>52.68</td>
<td>38.16</td>
<td>38.16</td>
<td>45.07</td>
</tr>
<tr>
<td>HA-222</td>
<td>50.43</td>
<td>45.53</td>
<td>45.53</td>
<td>42.35</td>
</tr>
<tr>
<td>HA-222 + FTC</td>
<td>50.76</td>
<td>52.10</td>
<td>52.10</td>
<td>42.18</td>
</tr>
</tbody>
</table>

* Low loader cell population (representative of SP)
V.C.2 Statistical analysis of A549 uptake before/after hydrolysis.

Experiments were performed in cell media vs. Aldefluor buffer. Data were processed using FlowJo software (Tree Star). DRAQ5 was used to set the dynamic range and as control dye and empty cells were used to determine auto-fluorescence.

<table>
<thead>
<tr>
<th>A549 sample</th>
<th>Solution</th>
<th>Median</th>
<th>Covariance</th>
<th>SD</th>
<th>% ile (25)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Empty</td>
<td>media</td>
<td>15.94</td>
<td>31.29</td>
<td>31.29</td>
<td>13.48</td>
</tr>
<tr>
<td>DRAQ5</td>
<td>media</td>
<td>217.4</td>
<td>31.61</td>
<td>31.61</td>
<td>191.27</td>
</tr>
<tr>
<td>LC-110 parental</td>
<td>media</td>
<td>101.48</td>
<td>34.58</td>
<td>34.58</td>
<td>84.64</td>
</tr>
<tr>
<td></td>
<td>buffer</td>
<td>96.13</td>
<td>32.87</td>
<td>32.87</td>
<td>81.66</td>
</tr>
<tr>
<td>LC-110 hydrolysed</td>
<td>media</td>
<td>64.9</td>
<td>35.28</td>
<td>35.28</td>
<td>55.56</td>
</tr>
<tr>
<td></td>
<td>buffer</td>
<td>102.97</td>
<td>31.38</td>
<td>31.38</td>
<td>86.76</td>
</tr>
<tr>
<td>LC-111 parental</td>
<td>media</td>
<td>75.48</td>
<td>28.15</td>
<td>28.15</td>
<td>65.32</td>
</tr>
<tr>
<td></td>
<td>buffer</td>
<td>67.08</td>
<td>32.09</td>
<td>32.09</td>
<td>57.65</td>
</tr>
<tr>
<td>LC-111 hydrolysed</td>
<td>media</td>
<td>51.51</td>
<td>33.5</td>
<td>33.5</td>
<td>44.66</td>
</tr>
<tr>
<td></td>
<td>buffer</td>
<td>61.44</td>
<td>38.88</td>
<td>38.88</td>
<td>51.25</td>
</tr>
<tr>
<td>LC-114 parental</td>
<td>media</td>
<td>72.42</td>
<td>26.78</td>
<td>26.78</td>
<td>62.72</td>
</tr>
<tr>
<td></td>
<td>buffer</td>
<td>74.94</td>
<td>28.85</td>
<td>28.85</td>
<td>64.64</td>
</tr>
<tr>
<td>LC-114 hydrolysed</td>
<td>media</td>
<td>59.33</td>
<td>32.6</td>
<td>32.6</td>
<td>49.93</td>
</tr>
<tr>
<td></td>
<td>buffer</td>
<td>57.89</td>
<td>35.36</td>
<td>35.36</td>
<td>48.48</td>
</tr>
</tbody>
</table>

* Low loader cell population (representative of SP)
Chapter VI

Epigenetic regulation of ALDHs and ABC transporters
VI.1 Introduction and Aims

Extensive studies have shown that aberrant methylation of CpG islands (CGIs) in promoter regions is associated with transcriptional inactivation of genes involved in all aspects of tumour development (Teodoridis et al., 2004). Furthermore, identification of epigenetic control in maintaining SC fate has been well characterised in hematopoietic and neural SCs, demonstrating that genes active in earlier progenitors are gradually silenced at later stages of development whereas other subsets of cell type-specific genes are activated (Lunyak and Rosenfeld, 2008). Transcription factors that are involved in regulating epigenetic information and particularly those that are necessary in reprogramming epigenetic marks in SCs are under intense study.

However, less is known about the aberrant methylation of enzymes associated with drug resistance. Consequently, this study is aimed at exploring the expression of ALDH and ABCs, two different families of proteins that are connected through the metabolism and removal of xenobiotics from the intracellular pool such as Aldefluor and potentially also the novel anthraquinone probe molecules described in this thesis. Specifically, focus is on the expression of isoforms from the ALDH 1-3 sub-families (ALDH1A1, ALDH2, and ALDH3A1) and the major ATP-binding transporter proteins ABCB1 and ABCG2. In line with previous Chapters in this thesis, A549 and MCF7 cell lines will be examined. In addition, HT29 and HeLa, two other cell lines derived from different cancer tissue will be examined in order to explore more broadly if the expression of ALDH and ABC proteins are cell-type specific and under epigenetic control.
A wide variety of methodologies can be applied to obtain DNA methylation data. The detection of this epigenetic alteration of the genome is based on the ability to differentiate between cytosine and 5-methylcytosine in the DNA sequence.

There are three principal approaches to detect this difference in methylation: (1) the use of the chemical modification of DNA with sodium bisulfite (see Figure 6.1), (2) digestion of DNA with either a methylation-sensitive or insensitive restriction enzyme, or (3) immunoprecipitation of 5-methylcytosine to distinguish any methylated fractions.

![Chemical Reaction of Bisulfite Conversion of Cytosine to Uracil](image)

**Figure 6.1 Chemical reaction of bisulfite conversion of cytosine to uracil.** Step 1: sulfonation; Step 2: hydrolytic deamination; Step 3: alkali-desulfonation. Bisulfite conversion is performed under acidic conditions and preferentially deaminates cytosine in a nucleophilic attack whilst the methyl group on 5-methylcytosine is protecting the amino group from the deamination. Cytosine reacts with bisulfite, but not 5-methylcytosine.
Sodium bisulfite conversion relies on the differential deamination of cytosine to uracil without affecting 5-methylcytosine content. This conversion will produce a difference in DNA sequence, which depends on its original methylation status.

Established methods using the bisulfite conversion of DNA described above include methylation-specific PCR (Herman et al., 1996), bisulfite sequencing (Frommer et al., 1992), combined bisulfite restriction analysis (COBRA) (Xiong and Laird, 1997) and, more recently, pyrosequencing (Ronaghi et al., 1998). Pyrosequencing can quantitate CpG methylation at individual sites following bisulfite treatment. As explained in Figure 6.2, the sequencing by synthesis based techniques involves the luminometric detection of pyrophosphate following sequential single nucleotide incorporation (Ronaghi, 2001).

Figure 6.2 Schematic representation of the progress of the enzyme reaction in liquid-phase pyrosequencing. Primed DNA template and four enzymes involved in liquid-phase pyrosequencing are placed in a well of a microtiter plate. The four different nucleotides are added stepwise and incorporation is followed using ATP sulfurylase and luciferase. The nucleotides are continuously degraded by nucleotide-degrading enzyme allowing addition of subsequent nucleotide. (d)XMP indicates one of the four nucleotides (Ronaghi, 2001).
Pyrosequencing allows the detection of 10 successive CpGs in a single sequencing reaction spanning up to 80 nucleotides (Tost and Gut, 2007). The main advantages of this technique over traditional sequencing methods are that it is a fast and an efficient quantitative method; it uses a PCR product to directly obtain information without the requirement of multiple sequencing reactions.

In this study pyrosequencing technology will be used, together with quantitative RT-PCR gene expression profiling to understand whether there is a link between methylation status and gene expression of the ALDH and ABC target genes. Analysis of untreated samples will be compared with decitabine (DAC) treated samples or samples treated concurrently with DAC and TSA (trichostatin A) (refers to sections I.4.4-6 for epigenetic therapy details).

Therefore, the specific aims of this study are:

1. To investigate the methylation level of ALDH1A1, ALDH2, ALDH3A1, ABCB1 and ABCG2 in the four human cancer cell lines A549, MCF7, HeLa and HT29 before and after DAC exposure;

2. To investigate if concurrent use of DAC and TSA alters the DNA methylation pattern compared with DAC alone;

3. To investigate the gene expression of the selected ALDH and ABCs using qRT-PCR and to correlate findings with the methylation status;

4. To quantify protein expression by western blotting and to establish if there is a connection between the methylation status of the selected genes and their gene and protein expression by use of DAC alone or in combination with TSA.
VI.2 Materials and Methods

VI.2.1 Mammalian cell lines culture and drug treatments

VI.2.1.1 Maintenance of mammalian cell lines

Cells were obtained from the American Type Culture Collection (ATCC) and were grown as outlined in Table 6.1 at 37 °C, 5% CO$_2$ at 100% humidity. Cell culture plastics were from Corning. Cells were kept at subconfluent levels and were typically passaged when 70–80% confluent.

<table>
<thead>
<tr>
<th>Cell line</th>
<th>Origin of cell line</th>
<th>Culture Medium</th>
<th>Approx. frequency of subculture</th>
<th>Dilution upon subculture</th>
</tr>
</thead>
<tbody>
<tr>
<td>A549</td>
<td>Human lung adenocarcinoma</td>
<td>R10</td>
<td>4 days</td>
<td>1:5-1:10</td>
</tr>
<tr>
<td>MCF7</td>
<td>Human breast adenocarcinoma</td>
<td>R10</td>
<td>4-5 days</td>
<td>1:10</td>
</tr>
<tr>
<td>HT29</td>
<td>Human colon adenocarcinoma</td>
<td>R10</td>
<td>3-4 days</td>
<td>1:10</td>
</tr>
<tr>
<td>HeLa</td>
<td>Human cervical carcinoma</td>
<td>R10</td>
<td>3 days</td>
<td>1:20</td>
</tr>
</tbody>
</table>

Table 6.1 Culture of mammalian cell lines. See Appendix II for cell media composition.

VI.2.1.2 Mammalian cell passaging

The old media was aspirated carefully and discarded from each 75 cm$^2$ flask. The cells were washed with 10 mL Hank’s Balanced Salt solution (Sigma), which was then aspirated and discarded. To break down cellular attachments to the flask, 2 mL trypsin in EDTA (Sigma) (1:10 dilution) was added to cover all
cells. The flasks were incubated for 3-5 min, and the contents of the flask were checked under light microscope to ensure that the cells were mobile and not clumped. The trypsin was quenched with 8 mL of medium, which was added slowly to avoid the formation of bubbles. The cells were then centrifuged at 1,000 rpm for 5 min. The supernatant was discarded and cells were resuspended in culture medium to the appropriate volume.

**VI.2.1.3 Determination of live cell number**

To determine live cell counts, 0.5 mL 0.4% Trypan Blue Stain (Sigma), 0.3 mL Hank’s balanced salt solution and 0.1 mL cell suspension were assembled in a tube, mixed vigorously by vortexing and incubated at room temperature for 5 min. Total cell number (non-stained cells and blue cells) and live cell number (non-stained cells) were then determined using a haemocytometer.

**VI.2.1.4 Cryopreservation of mammalian cells**

For storage in liquid nitrogen, mammalian cells were trypsinised, sedimented by centrifugation and suspended in ice-cold standard freezing medium.

**VI.2.1.5 Drug stock solution**

Drug stock solutions were prepared by dissolving the target compound (DAC or TSA) in DMSO (Sigma) to 30 mM concentration and kept at -20 °C till required. Working solutions were prepared in cell media immediately before use. Equal amount of DMSO was used to treat the cell control to establish the effect of the carrier on the cell activity.
VI.2.1.6 Treatment with decitabine

Cells were seeded at a concentration of $1 \times 10^5$ cells in 75 cm$^2$ flasks. The next day, treatment of cells with 0, 100 nM, 1µM decitabine (DAC, 5-aza-2'-deoxycytidine) (Sigma) was started, and DAC was removed by changing medium 24 h later. The cells were harvested 6 days after removal of DAC for DNA and RNA extraction. For comparison, two more MCF7 samples which were treated with DMSO or 1 µM DAC every other day, for 7 days were kindly donated by Dr. Flanagan (Imperial College, London).

VI.2.1.7 Combination treatment: decitabine and trichostatin A

Cells were seeded at a concentration of $1 \times 10^5$ cells in 75 cm$^2$ flasks. The next day, treatment of cells with 0 or 100 nM DAC (Sigma) was started, and DAC was removed by changing medium 24 h later. Cells were harvested after 5 days in standard conditions. On day 5, treatment of cells with 100 nM TSA (TCI) was carried out for 24 h. Cells were collected after removal of TSA for DNA-RNA extraction.

VI.2.2 Methylation analysis

VI.2.2.1 Isolation of genomic DNA and quantification

After removing the media, cells were washed twice with Hanks' Balanced Salt solution (Sigma) and detached from the flask using a cell scraper. $3 \times 10^6$ cells were collected in a 1.5 mL microcentrifuge tube and centrifuged for 5 min at maximum speed. The supernatant was removed and the cell pellet was
resuspended in 200 µL PBS (Sigma). No proteinase K was added and total DNA was isolated from cultured cells with QIAamp DNA Mini Kit (QIAGEN) following the manufacturer's protocol. Briefly, 200 µL of buffer AL was added to the sample and mixed by pulse vortexing for 15 sec. Samples were heated at 56°C for 10 min and centrifuged to remove drops from the inside of the lid. In order to lysate the samples, 200 µL of 100% ethanol was added and mixed by vortexing. The mixture was carefully transferred in a QIAamp Mini spin column and centrifuged at 8000 rpm for 1 min. The filtrate was discarded and 500 µL of buffer AW1 (provided with the kit) was added to wash the spin column. After repeating the centrifugation step and discarding the supernatant, 500 µL of buffer AW2 (provided with the kit) was added to wash the spin column and centrifuged at full speed for 3 min and 1 min. The collection tube containing the filtrate was discarded and total DNA eluted adding 200 µL of RNase-free water. The samples were incubated for 1 min at room temperature (25°C) and then centrifuged at 8000 rpm for 1 min. To determine the concentration and purity of the total DNA samples, the absorbance of UV light was quantified using NanoDrop™ 2000 (Thermo Fisher Scientific) and 260/280 ratio was calculated. Samples were stored at -70°C until required.

**VI.2.2.2 Bisulfite conversion**

A total of 1 µg of genomic DNA (gDNA) for each sample was bisulfite treated using Epitect Bisulfite kit (QIAGEN). According with the measured DNA concentration, the DNA solution was mixed in a PCR tube with 85 µL of bisulfite mix (previously dissolved according with manufacturer’s protocol) and 35 µL of DNA protect buffer, adjusting the final reaction volume with RNase-free water to
140 µL. The reaction mixture was mixed thoroughly and the bisulfite DNA conversion was performed using a MJ Research PTC-200 Peltier thermal cycler using the following conditions in Table 6.2.

<table>
<thead>
<tr>
<th>Step</th>
<th>Time</th>
<th>Temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Denaturation</td>
<td>5 min</td>
<td>95°C</td>
</tr>
<tr>
<td>Incubation</td>
<td>25 min</td>
<td>60°C</td>
</tr>
<tr>
<td>Denaturation</td>
<td>5 min</td>
<td>95°C</td>
</tr>
<tr>
<td>Incubation</td>
<td>85 min</td>
<td>60°C</td>
</tr>
<tr>
<td>Denaturation</td>
<td>5 min</td>
<td>95°C</td>
</tr>
<tr>
<td>Incubation</td>
<td>175 min</td>
<td>60°C</td>
</tr>
<tr>
<td>Hold</td>
<td>Indefinite (overnight)</td>
<td>20°C</td>
</tr>
</tbody>
</table>

Table 6.2 Optimised PCR conditions for bisulfite conversion step.

After bisulfite conversion, the PCR tubes containing the bisulfite reactions were briefly centrifuged and then each mixture was transferred into clean microcentrifuge tube. 560 µL of freshly prepared BL buffer (provided with the kit) containing 10 µg/mL of carrier RNA were added to each sample. The mixture was vortexed and transferred into an Epitect spin column (QIAGEN). The column was centrifuged at full speed for 1 min and the supernatant was discarded. 500 µL of washing buffer BW (provided with the kit) was added to each spin column and centrifuged at full speed for 1 min. The supernatant was discarded and the procedure was repeated by adding 500 µL of buffer BD and
incubating for 15 min at room temperature (25°C). The mixture was washed twice with 500 µL of buffer BW and centrifuged at maximum speed to discard the flow-through. The spin columns were heated at 56 °C for 5 min to enable the evaporation of any of any remaining liquid. The purified DNA was eluted by centrifugation for 1 min at 12000 rpm. The DNA samples were stored at -20°C until needed.

VI.2.2.3 Control samples

Positive CpG Universal Methylated Control DNA (Millipore), also known as in vitro methylated DNA (IVM) and human male gDNA (Promega) were bisulfite treated as standards. Negative unmethylated control DNA was amplified from any gDNA sample using whole genome amplification GenomiPhi V2 DNA amplification kit (GE Healthcare) and bisulfite converted with the target samples. Moreover, three additional serial dilutions of 25%, 50%, 75% of IVM and male gDNA were mixed and prepared prior to bisulfite treatment. Water control (WC) was included to detect eventual contaminations.

VI.2.2.4 Calponin quantification

The sequence of modified calponin-specific primers was obtained from (Sriraksa et al., 2009) and used to check bisulfite modified DNA with 333 bp product size. Modified calponin-specific primer sequences were:

5′-GGAAGGTAGTTGAGGTTGTG-3′ (forward primer)
5′-CCCAAACCTCAAAACTCTAACC-3′ (reverse primer)
**VI.2.2.5 Polymerase chain reaction (PCR)**

PCR samples were prepared with a 25 µL final volume using: 1× PCR buffer (16.6 mM ammonium sulfate and 67 mM Tris, pH 8.8), 3 mM MgCl$_2$, 200 µM of dNTPs, 5 pmol of each primer, 150 ng of bisulfite-treated DNA and 1.5 U of Taq polymerase. The PCR was performed using a MJ Research PTC-200 peltier thermal cycler using the conditions in Table 6.3 for 35 cycles. 3 µL of PCR products was separated by 2% agarose gel electrophoresis and visualized under UV illuminator after ethidium bromide staining.

<table>
<thead>
<tr>
<th>Step</th>
<th>Time</th>
<th>Temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial denaturation</td>
<td>5 min</td>
<td>95°C</td>
</tr>
<tr>
<td>Denaturation</td>
<td>30 sec</td>
<td>95°C</td>
</tr>
<tr>
<td>Annealing</td>
<td>30 sec</td>
<td>63°C</td>
</tr>
<tr>
<td>Elongation</td>
<td>30 sec</td>
<td>72°C</td>
</tr>
<tr>
<td><strong>Repeat 34 times from step 2</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Final elongation</td>
<td>10 min</td>
<td>72°C</td>
</tr>
</tbody>
</table>

Table 6.3 Optimised PCR conditions for calponin amplification.

**VI.2.2.6 LINE1 analysis: PCR and pyrosequencing**

Methylation analysis of LINE-1 (long interspersed nucleotide element-1, or L1) promoter (GenBank accession number X58075) was investigated using a
pyrosequencing-based methylation analysis. PCR and pyrosequencing of LINE-1 were performed using the PyroMark Q96 CpG LINE-1 kit (QIAGEN).

The LINE-1 specific primer sequences were:

5'-TTTTTTGAGTTAGGTGTGGG-3' (forward primer)
5'-BIO-TCTCACTAAAAATACCAAAACA-3' (reverse-biotinylated primer)
5'-GGGTGGGAGTGAT-3' (sequencing primer)

Primer stocks were all prepared in RNA-free water at the concentration of 1 µg/µL for forward/reverse and 10 µM for the sequencing primer. Each PCR mixture contained the forward and reverse primer (each 0.2 µmol/L), 0.8 µmol/L of dNTPs, 1.5 mmol/L of MgCl₂, 1x PCR buffer (QIAGEN), 0.64 U of HotStar Taq polymerase (QIAGEN), and 1 µL of bisulfite-converted template DNA in a total volume of 25 µL reaction. The PCR was performed using a MJ Research PTC-200 Peltier thermal cycler and the conditions were as follows:

<table>
<thead>
<tr>
<th>Step</th>
<th>Time</th>
<th>Temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial denaturation</td>
<td>15 min</td>
<td>95°C</td>
</tr>
<tr>
<td>Denaturation</td>
<td>20 sec</td>
<td>95°C</td>
</tr>
<tr>
<td>Annealing</td>
<td>20 sec</td>
<td>50°C</td>
</tr>
<tr>
<td>Elongation</td>
<td>20 sec</td>
<td>72°C</td>
</tr>
<tr>
<td>Repeat 44 times from step 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Final elongation</td>
<td>10 min</td>
<td>72°C</td>
</tr>
</tbody>
</table>

Table 6.4 Optimised PCR conditions for LINE-1 amplification.
After amplification of a 146 kb fragment, samples were stored overnight at 2-8 °C or at -20 °C for longer storage. 3 µL of PCR products was separated by 2% agarose gel electrophoresis and visualized under UV illuminator after ethidium bromide staining to ensure no contamination on the negative control.

A simplex Entry for the PyroMark CpG LINE-1 assay was assessed in PyroMark CpG Software using the following parameters:

TTYGTGGTGYGTYGTTTTTAAAGTYGGTTT (Sequence to Analyze)
ATCAGTGTGTTCAGTCAGTTAGTCTG (Dispensation Order)

The biotinylated PCR product (10 µL) was purified and made single-stranded to act as a template in a pyrosequencing reaction as recommended by the manufacturer using the PyroMark Q96 Vacuum Workstation (QIAGEN).

The DNA immobilization component (80 µL per well) was prepared in a PCR plate with the following volumes:

<table>
<thead>
<tr>
<th>Master mix component</th>
<th>Volume per sample (µL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Streptavidin Sepharose HP beads</td>
<td>2</td>
</tr>
<tr>
<td>PyroMark Binding Buffer</td>
<td>38</td>
</tr>
<tr>
<td>RNase-free water</td>
<td>30</td>
</tr>
<tr>
<td>PCR product</td>
<td>10</td>
</tr>
<tr>
<td>Total volume</td>
<td>80</td>
</tr>
</tbody>
</table>

Table 6.5 Reagent volumes for pyrosequencing analysis.
Briefly, the PCR product that was bound to streptavidin sepharose HP (horseradish peroxidase) and sepharose beads (QIAGEN) containing the immobilized PCR product were purified, washed, denatured using a 0.2 M NaOH solution, and washed again (Estecio et al., 2007). The sequencing primer was diluted in Annealing buffer (QIAGEN) to a final concentration of 0.3 µM and 12 µL of the prepared solution was added to the PyroMark Q96 H sequencing plate. The PCR plate were agitated constantly for 5-10 min at room temperature (15-25°C) using a mixer. The reagents were respectively added in the PyroMark Q96 Vacuum Workstation (QIAGEN) as illustrated in Figure 6.3.

![Figure 6.3 PyroMark Q96 Vacuum Workstation.](image)

The marked positions contain 110 mL 70% ethanol, (1), 90 mL PyroMark Denaturation Solution (2), 110 mL PyroMark Wash Buffer (3), and 110 mL high-purity water (4) (adapted from PyroMark Q96 handbook).

After applying vacuum to the vacuum handset tool (see Figure 6.3), the beads containing immobilised DNA were captured onto the filter probes. The handset tool (P, see Figure 6.3) was transferred first to the trough containing 70%
ethanol (purification), secondarily to the trough containing PyroMark Denaturation Solution (denaturation), and finally to the trough containing Wash Buffer (washing). After the vacuum was switched off, the beads were released into the wells containing sequencing primer by gently shaking the tool in the wells. The PyroMark Q96 H Plate with the samples was heated at 80 °C for 2 min using a heating block. The samples were cooled down to room temperature and PyroMark Q96 H Plate was loaded into the PSQ96 H System (QIAGEN).

After the LINE-1 assay was set up in the PSQ96 H System, the reagent tips and capillary tips were loaded in the tip holder with the appropriate volumes of PyroMark Gold Q96 Reagents (QIAGEN), as shown in Figure 6.4.

![Figure 6.4 Order of tips in the capillary tip holder. E: Enzyme Mixture; S: Substrate Mixture; G: dGTP; C: dCTP; T: dTTP; A: dATP (adapted from PyroMark Q96 handbook).](image)

After the tip holder and the PyroMark Q96 H Plate were respectively loaded in the dispensing unit the run was performed. The analysis results (methylation level) and quality control were assessed using PyroMark Q96 MD Software and methylation values were calculated as an average of all CpG sites within the assay as determined by the Pyro Q-CpG Software. PyroMark CpG LINE-1
assay was repeated in two independent biological replicates to assess the reproducibility of the methylation analysis via pyrosequencing.

**VI.2.2.7 Assay design for target genes**

Gene sequence, transcription start site, CpG island loci and any additional information were established using UCSC genome browser (http://genome.ucsc.edu/). Primer design was performed by PyroMark Assay Design software using the *in silico* bisulfite converted DNA sequence. After the target region was selected and the desired sequence length was set, the automatic assay design was run to generate the primer sets. A biotin tag was placed on either the forward or reverse primer (depending on the strand to be sequenced). The utility indicated if there were repeated bases in the region of interests, potential mispriming and if the pyrosequencing template (bisulfite-treated PCR product) could form 3' template loops. Mispriming and template loops caused background in the pyrosequencing reaction that could render an assay unusable. These problems could be solved by moving one or both of the PCR products to exclude the offending sequence. The utility assigned a score and quality for each sequencing primer that it designed, penalizing a primer for any of the above problems and also for length more than 15 bases. The primer set (forward, reverse and sequencing) with the highest quality and no warnings was selected as final set and saved to be imported and used in other pyrosequencing software. All the final assays for the studied genes are reported at the end of the present chapter in Supplementary Information VI.A.
VI.2.2.8 PCR Conditions for designed assay

VI.2.2.8.1 Primers

The sequence of bisulfite converted specific primers was obtained for every target gene as explained on section VI.2.2.7.

VI.2.2.8.2 Primers Polymerase chain reaction (PCR)

The sequence PCR samples contained a final volume of 25 µL and were prepared using: 1× PCR buffer (16.6 mM ammonium sulfate and 67 mM Tris, pH 8.8), 3 mM MgCl₂, 200 µM of dNTPs, 5 pmol of each primer, 150 ng of bisulfite-treated DNA and 1.5 U of Taq polymerase. The PCR was performed using a MJ Research PTC-200 Peltier thermal cycler using the conditions listed on the Table 6.6.

<table>
<thead>
<tr>
<th>Step</th>
<th>Time</th>
<th>Temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial denaturation</td>
<td>5 min</td>
<td>95 °C</td>
</tr>
<tr>
<td>Denaturation</td>
<td>30 sec</td>
<td>95 °C</td>
</tr>
<tr>
<td>Annealing</td>
<td>30 sec</td>
<td>see table below</td>
</tr>
<tr>
<td>Elongation</td>
<td>30 sec</td>
<td>72 °C</td>
</tr>
<tr>
<td>Repeat 34 times from step 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Final elongation</td>
<td>10 min</td>
<td>72 °C</td>
</tr>
</tbody>
</table>

Table 6.6 PCR conditions for amplification of target genes.
Primer-template annealing temperatures were optimised for every assay using a temperature gradient between 53 and 62°C to avoid PCR dimers by annealing of their 3'-termini or non-specific amplification, which cause the appearance of multiple bands on agarose gels (Rychlik et al., 1990). The final annealing temperature values are listed in Table 6.7.

After the reaction, 3 µL of PCR product was separated by 2% agarose gel electrophoresis and visualized under UV illuminator after ethidium bromide staining. Negative control was added to detect DNA contamination.

<table>
<thead>
<tr>
<th>Target gene</th>
<th>Annealing Temperature (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALDH1A1</td>
<td>59</td>
</tr>
<tr>
<td>ALDH2</td>
<td>56</td>
</tr>
<tr>
<td>ALDH3A1</td>
<td>59</td>
</tr>
<tr>
<td>ABCB1</td>
<td>62</td>
</tr>
<tr>
<td>ABCG2</td>
<td>59</td>
</tr>
</tbody>
</table>

Table 6.7 Optimised annealing temperature for PCR amplification of target genes.

VI.2.2.9 Pyrosequencing

A simplex Entry for the PyroMark CpG assay was assessed in PyroMark CpG Software using the gene-specific parameters listed in Supplementary Information VI.A. The pyrosequencing procedure was performed as already described for the LINE-1 assay (refer to “LINE-1 analysis” section for details).
PyroMark CpG assay was repeated in two PCR replicates to assess the reproducibility of the methylation analysis.

VI.2.2.10 Statistical analysis

Each treatment of the cell lines was performed in duplicate. The two values were then averaged to obtain one value for each culture. From these data, the mean value ± standard error of the mean (SEM) was calculated for each experimental condition. An analysis of variance was carried out on the data for testing the treatment effect. For comparison of two means, Student’s t-test was used for unpaired samples. A difference was considered significant when the P value was <0.05.

VI.2.3 Gene expression analysis using quantitative RT-PCR (qRT-PCR) before and after epigenetic treatment

VI.2.3.1 RNA extraction and quantification

The cell monolayer was washed twice with Hanks’ Balanced Salt solution (Sigma) and detached from the flasks using a cell scraper. 2x10^6 cells were collected in a 1.5 mL microcentrifuge tube and centrifuged for 5 min at maximum speed. The supernatant was removed and the cell pellets were resuspended in 200 µL PBS (Sigma). No DNase treatment was performed and RNA was extracted from the cell pellets using an RNeasy RNA extraction kit (QIAGEN) according to the manufacturer’s instructions. Briefly, cells were disrupted in the lysis buffer and homogenized using a QIA spin shredder column (QIAGEN) or a blunt 20-gauge needle fitted to an RNase-free syringe.
70% ethanol was added to homogenate the lysate and mixed well by pipetting. Supernatant was transferred in a spin column (QIAGEN) and the spin column membrane was washed multiple times to avoid carryover of ethanol. Finally the RNA was eluted by adding 40 µL of RNase-free water and centrifuging for 1 min at maximum speed. Quality and quantity of extracted RNA was analysed by measuring the absorbance of UV light using NanoDrop™ 2000 (Thermo Fisher Scientific) and calculating the 260/280 ratio. Samples were stored at -70°C until required.

**VI.2.3.2 Complementary DNA synthesis**

Complementary DNA (cDNA) single stranded was generated from total RNA in 20 µL reactions using Multiscribe Reverse Transcriptase kit (Applied Biosystems) as follows: 10x reverse transcription buffer, 5.5 mM MgCl₂, 500 µM/dNTP, 2.5 mM random hexamers, 0.5 U/µL RNase inhibitor, 1.25 U/µL multiscribe reverse transcriptase and RNase free water.

<table>
<thead>
<tr>
<th>Step</th>
<th>Time</th>
<th>Temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incubation</td>
<td>10 min</td>
<td>25 °C</td>
</tr>
<tr>
<td>Synthesis</td>
<td>30 min</td>
<td>48 °C</td>
</tr>
<tr>
<td>Enzyme inactivation</td>
<td>5 min</td>
<td>95 °C</td>
</tr>
</tbody>
</table>

*Table 6.8 Optimised conditions for cDNA synthesis.*
2 µL of specimen RNA was added to each reaction. Reactions were carried out in a MJ Research PTC-200 Peltier thermal cycler using the conditions listed on Table 6.8.

**VI.2.3.3 Universal Probe Library technology**

PCR primers were designed using Roche Universal Probe Library (UPL) (Roche Diagnostic Corporation), containing 90 short hydrolysis probes (8-9 nucleotides), labelled at the 5' end with fluorescein (5'-FAM, Ex/Em: 494/522 nm) and at the 3' end with a dark quencher dye. In order to maintain the specificity and melting temperature that hybridizing qRT-PCR probes requires, Locked Nucleic Acids (LNA) were incorporated into the sequence of each UPL probe. These LNAs had the ribose ring “locked” with a methylene bridge by connecting the 2'-O atom with the 4'-C atom. This modification conferred base stacking and increased thermal stability compared to normal DNA nucleotides.

**VI.2.3.4 QRT-PCR primers design**

The Probe Finder software (Roche Diagnostic Corporation) allowed a rapid design of qRT-PCR assays for the targets of choice in three steps: organism selection, gene name or gene accession number insertion, and sequence paste into the appropriate field. The software created a specific qRT-PCR assay by combining a suitable UPL probe with a set of target specific PCR primer pairs. Multiple primer pair options were available and the system ranked them to favour intron spanning amplicons to remove false signals from contaminating gDNA together with a small amplicon size for reproducible and robust assays. Three best ranked assays were chosen for each target gene and the one giving the best amplification profile was kept as definitive. Detailed informations
regarding the final assays are listed in Supplementary Information VI.B at the end of the present chapter.

VI.2.3.5 QRT-PCR method

QRT-PCR was set up in a 96 well plate in triplicates in a UV-irradiated hood on a Stratagene Mx3005P QPCR System (Agilent Technologies). The assay was performed using 25 µL reactions consisting of 1 µL of cDNA, 12.5 µL of 2x TaqMan Universal PCR master mix (Roche Applied Biosystems), 400 nM forward and reverse primers for the target genes and 400 nM forward and reverse primers and for housekeeping genes GAPDH and PPIA. TaqMan probes were used at a concentration of 50 nM for all genes. Optimal primers and probe concentrations were deduced through optimisation reactions for each gene, and amplification efficiencies were obtained using standard curves for each gene primer and probe combination. For each assay, no-template controls were included.

<table>
<thead>
<tr>
<th>Step</th>
<th>Time</th>
<th>Temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enzyme activation</td>
<td>10 min</td>
<td>95 °C</td>
</tr>
<tr>
<td>Denaturation</td>
<td>15 sec</td>
<td>95 °C</td>
</tr>
<tr>
<td>Annealing/Extension</td>
<td>60 sec</td>
<td>60 °C</td>
</tr>
<tr>
<td>Repeat 39 times from step 2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Table 6.9* Primer sequences and assay details for qPCR analysis of target genes.
A commercial kit containing primers and TaqMan probes for the housekeeping genes GAPDH and PPIA respectively (Human Endogenous Control (FAM / MGB Probe, Non-Primer Limited), Roche Applied Biosystems) was used for the normalisation of the reaction. Cycling conditions are listed in Table 6.9.

VI.2.3.6 QRT-PCR efficiency validation

To validate the qRT-PCR efficiency (E) of target gene primers, standard curves were generated from a dilution series of target cDNA (example shown in Figure 6.5). The PCR efficiency was calculated from the slope of the standard curve using the following equation:

\[ E = (10^{-1/slope}) - 1 \]

For example, a slope of -3.32 represents a PCR efficiency of 1 (100%). The PCR efficiency (E) values of the target genes were ranged between 50-85%, as listed in Table 6.10.

<table>
<thead>
<tr>
<th>Target gene</th>
<th>Efficiency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALDH1A1</td>
<td>73</td>
</tr>
<tr>
<td>ALDH2</td>
<td>64</td>
</tr>
<tr>
<td>ALDH3A1</td>
<td>80</td>
</tr>
<tr>
<td>ABCB1</td>
<td>50</td>
</tr>
<tr>
<td>ABCG2</td>
<td>67</td>
</tr>
</tbody>
</table>

Table 6.10 PCR efficiency value calculated using the equation reported above.
Figure 6.5 Standard Curves of PPIA, ALDH3A1, and ABCG2 derived from a dilution series of A549 cDNA using quantitative real-time PCR (TaqMan Assays). The Ct values correlated to the copy number of the gene ($R^2 = 0.992$, $Y = -4.959 \times \text{LOG}(X) + 21.55$, PPIA; $R^2 = 0.991$, $Y = -4.171 \times \text{LOG}(X) + 25.26$, ALDH3A1; $R^2 = 0.962$, $Y = -4.526 \times \text{LOG}(X) + 26.37$, ABCG2).

VI.2.3.7 Data analysis

This assay was based on measuring fluorescence using fluorescent chemistries such as TaqMan probes (see “Universal Probe Library technology” section for details). The fluorescence intensity detected during each cycle was proportionally correlated to the amount of amplicon produced.

MxPro QPCR Software (Agilent Technologies) was used to perform all data analysis. The baseline for the amplification plot was determined above the small changes in fluorescence signals produced by the initial cycles of PCR. As the PCR products accumulates during PCR amplification, the fluorescence intensity increased above the baseline. Consequently, a cycle threshold (Ct or Cts) was
fixed and set above the baseline. The Ct of each PCR cycle represented the time point where fluorescent signals of target amplicons were first determined and then crossed to the fixed threshold (Stewart et al., 2006).

Cts < 29 were strong positive reactions indicative of abundant target nucleic acid in the sample. Cts of 30-37 were positive reactions indicative of moderate amounts of target nucleic acid. Cts of 38-40 were weak reactions indicative of minimal amounts of target nucleic acid (Hunt, 2010).

Relative quantification was used to estimate the ratio between a target mRNA expression and an endogenous reference gene. These ratios were compared between samples being assayed. The endogenous references were housekeeping or maintenance genes whose expressions were relatively constant and not affected by experimental conditions. In this study, GAPDH (Glyceraldehyde 3-phosphate dehydrogenase) or PPIA (peptidylprolyl isomerase A or Cyclophilin A) were used as validated housekeeping genes in the selected cell lines.

$2^{\Delta Ct}$ method was used to calculate the relative expression of target gene (Livak and Schmittgen, 2001). This method was mainly based on Ct values. ΔCt was firstly calculated for each sample using the following equation:

$$\Delta Ct \text{ (sample)} = Ct \text{ target gene} - Ct \text{ reference gene}$$

$2^{\Delta\Delta Ct}$ method was used to compare the fold change of gene expression between chemically treated and untreated samples. The fold changes were obtained by applying the following equation:

$$\Delta Ct \text{ (calibrator or untreated cells)} = Ct \text{ target gene} - Ct \text{ reference gene}$$
Then, \( \Delta \Delta Ct \) value for each sample was calculated by subtracting the \( \Delta Ct \) value of the calibrator from the \( \Delta Ct \) value of the sample:

\[
\Delta \Delta Ct = \Delta Ct \text{ (sample)} - \Delta Ct \text{ (calibrator or untreated cells)}
\]

Finally, the \( \Delta \Delta Ct \) formula was used to estimate the normalized fold differences between treated and untreated cells \( 2^{-\Delta \Delta Ct} \).

VI.2.3.8 Statistical Analysis

The significance of results was assessed through a comparison of means using a paired two-tailed t-test with two way ANOVA (analysis of variance) followed by Bonferroni correction for multiple testing (Bonferroni, 1936). All statistical analysis was performed using GraphPad Prism version 4.00 (GraphPad Software, www.graphpad.com). Results were expressed as the mean ± standard error of the mean unless otherwise stated. P values were calculated to determine statistical significance of the results obtained as outlined in Table 6.11.

<table>
<thead>
<tr>
<th>P value</th>
<th>Wording</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;0.001</td>
<td>Extremely significant</td>
<td>***</td>
</tr>
<tr>
<td>0.001 to 0.01</td>
<td>Very Significant</td>
<td>**</td>
</tr>
<tr>
<td>0.01 to 0.05</td>
<td>Significant</td>
<td>*</td>
</tr>
<tr>
<td>&gt;0.05</td>
<td>Not Significant</td>
<td>n</td>
</tr>
</tbody>
</table>

Table 6.11 P values summary and significance.
VI.2.4 Western blot analysis

DAC was demonstrated to reflect its ability to reactivate methylation-silenced genes in a cell dependent manner. To establish protein expression of the target families of enzymes western blotting analysis was performed in the samples that showed positive correlation between DAC treatment and gene reactivation.

VI.2.4.1 Sample preparation

Mammalian cell lines were cultured and treated as outlined in section VI.2.1. After treatment the cells were harvested and centrifuged at 1000 rpm for 5 min at 4°C. The supernatant was removed and pellets were resuspended in RIPA lysis buffer (see Appendix II for buffers and solutions). Cell suspensions were maintained in constant agitation for 30 min at 4°C. Samples were centrifuged at 12000 rpm for 10 min at 4°C and the supernatant was transferred to a fresh eppendorf tube. The concentration of protein in the supernatant was determined using the Bradford assay.

VI.2.4.2 Determination of protein concentration

Protein standards ranging from 0 mg/mL to 1 mg/mL were prepared by serial dilutions of a bovine serum albumin (BSA) (Sigma) standard in deionised water. The cell lysates were diluted 1:5 in a new 1.5 mL eppendorf tube. Bradford reagent (Bio-Rad) was prepared by diluting one part of dye with four parts of deionised water. The diluted reagent was filtered through Whatman paper to remove particulates. 20 µL of sample or standard dilution were added to 1 mL of diluted dye reagent in a 1.5 mL eppendorf tube in duplicate. The mixture was vortexed and incubated at room temperature for 5 min. The absorbance of
standards and samples was measured at 595 nm using a spectrophotometer and the water blank was subtracted. To allow quantification of sample protein concentration, a standard curve was created using the protein standards. To normalise the protein loading in each well, the volume of sample containing 10 µg of protein was calculated. This volume was transferred to a new eppendorf tube and 1x SDS loading buffer (see Appendix II) was added to give a total volume of 20 µL. Samples were subsequently denaturated at 95°C for 5 min.

**VI.2.4.3 Polyacrylamide gel preparation**

The gel loading assembly (Bio-Rad) was cleaned with 70% ethanol and assembled as directed in the manufacturer instructions. A 10% resolving gel (see Appendix II) was prepared and pipetted into the assembled apparatus. The gel was allowed to set at room temperature for 20 min. A 4% stacking gel (see Appendix II) was prepared and pipetted over the running gel. A gel comb was inserted and the gel let to set for 20 min. The comb was removed and the gel apparatus was transferred to an electrophoresis buffer tank (Bio-Rad) filled with 1x TBS-Tween20 running buffer. The denaturated samples containing 10 µg of protein were loaded in the gel wells along with 10 µL of pre-stained molecular weight marker (Full range rainbow marker 10-250 kDa, Amersham Bioscienceses). The gel was run at 3W for 1 h.

**VI.2.4.4 Protein transfer to PVDF membrane**

Following electrophoresis, wet blotting was carried out using a PVDF (polyvinylidene fluoride) membrane soaked up in transfer buffer. Transfer was carried out at 100V for 1 h.
VI.2.4.5 Immunodetection of electrophoresed proteins after transfer to nitrocellulose membrane

The cassette was dismantled and the blotted PVDF membrane was placed in a Petri dish and rinsed with deionised water. The membrane was placed in 10 mL of 4% blocking solution (see Appendix II for solutions) for 30 min on a rocker. The blocking solution was removed and replaced with 10 mL of 2% blocking solution containing one of the unconjugated primary antibodies listed on Table 6.12.

<table>
<thead>
<tr>
<th>Specificity</th>
<th>Host species</th>
<th>Molecular weight (kD)</th>
<th>Dilution</th>
<th>Manufacturer</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALDH1A1</td>
<td>Rabbit polyclonal</td>
<td>54.9</td>
<td>1:1000</td>
<td>Prestige Antibodies, Sigma</td>
</tr>
<tr>
<td>ALDH2</td>
<td>Mouse monoclonal</td>
<td>50</td>
<td>1:500</td>
<td>Calbiochem</td>
</tr>
<tr>
<td>ALDH3A1</td>
<td>Mouse polyclonal</td>
<td>50</td>
<td>1:500</td>
<td>Abnova</td>
</tr>
<tr>
<td>ABCB1</td>
<td>Mouse monoclonal</td>
<td>36</td>
<td>1:500</td>
<td>Sigma</td>
</tr>
<tr>
<td>ABCG2</td>
<td>Mouse monoclonal</td>
<td>72</td>
<td>1:500</td>
<td>Chemicon, Millipore</td>
</tr>
</tbody>
</table>

Table 6.12 List of used primary antibodies and working dilutions.

This was incubated at 4°C on a rocker for overnight. The membrane was then washed in 1x TBS-Tween20 for 4 x 15 min. The membrane was subsequently incubated for 1 h with 10 mL of secondary antibody (see Table 6.13) solution in 2% blocking solution (see Appendix II). The membrane was washed in TBS/Tween20 for 4 x 15 min and then prepared for detection of the bands.
<table>
<thead>
<tr>
<th>Specificity</th>
<th>Host species</th>
<th>Dilution</th>
<th>Manufacturer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anti-Mouse</td>
<td>Polyclonal goat</td>
<td>1:4000</td>
<td>Bio-Rad</td>
</tr>
<tr>
<td>Anti-Rabbit</td>
<td>Polyclonal goat</td>
<td>1:4000</td>
<td>Southern Biotech</td>
</tr>
</tbody>
</table>

Table 6.13 List of used secondary antibodies and working dilutions.

VI.2.4.6 Enhanced chemiluminescent detection

Membranes were developed using the enhanced chemiluminescence system (ECL Plus™ detection kit, GE Healthcare). The detection solutions were prepared according to the manufacturer instructions. The excess washing buffer was drained off and the prepared detection solution pipetted onto the membrane surface protein side up. This was incubated at room temperature for 5 min on a clean filter paper. The excess of detection reagent was removed and the membrane was placed in a Bio-Rad 170-8126 Universal Hood (Bio-Rad) under epi-white illumination using ChemiDoc XRS Gel documentation system. The exposure time and configurations were adjusted using Quantity One analysis software (Bio-Rad). Band intensity was determined using densitometry and Quantity One analysis software (Bio-Rad).

VI.2.4.7 Stripping western blots for reprobing with loading control

After chemiluminescence detection, the membrane was washed in TBS/Tween20 for 2 x 10 min. Excess of washing buffer was drained off and the membrane was stripped using 10 M urea in 100 mM Tris-HCl pH 6.8. The blot was incubated in hybaid oven at 55 °C for 1 h with shaking. After stripping, the membrane was washed with TBS-Tween20 for 2 x 10 min. The blocking step
with 4% skimmed milk was repeated for 30 min. Anti-GAPDH monoclonal antibody produced in mouse (Calbiochem) or anti-α-tubulin monoclonal antibody produced in mouse (Bio-Rad) was used as control, diluted 1:2000 in 2% blocking solution (see Appendix II) and incubated for 1 h at RT on a rocker. Incubation with secondary antibody was carried out as previously described on section VI.2.4.5.
VI.3 Results

VI.3.1 Pyrosequencing analysis of the promoter region in LINE-1

DNA methylation measured in LINE-1 has been correlated with global DNA methylation (Irahara et al.). Several studies have demonstrated that LINE-1 retrotransposable elements made up about 15% of human genome and DNA methylation played a fundamental role within the promoter region of human LINE-1 elements for maintaining transcriptional inactivation and for inhibiting transposition (Singer, 1982).

In the analysed samples, methylation levels of the LINE-1 repetitive elements were used as a surrogate marker of genome-wide methylation changes (Estecio et al., 2007). Given the abundance of LINE-1 elements in the genome, minimum amounts of DNA were required for their amplification and analysis. Therefore the technique was validated by investigating the methylation level of LINE-1 promoter together with methylated and unmethylated controls. Methylation analysis of LINE-1 promoter was used to evaluate the reproducibility of the data generated via pyrosequencing (see Supplementary Information VI.C at the end of the present chapter for LINE-1 methylation analysis plots).

Prior to pyrosequencing analysis, successful bisulfite modification of the DNA sample was verified using a region of the calponin promoter as described in section VI.2.2. This sequence could only be amplified if the cytosines in the template sequence were converted to uracils (Sriraksa et al., 2009). Samples which did not give a band of similar intensity were considered unmodified or incompletely modified and the modification experiment was repeated for these, as illustrated in Figure 6.6.
Figure 6.6 Examples of successful sodium bisulfite modifications. Calponin amplification product appeared at 333 bp. NC, reaction without template (negative control); 1-5: IVM, *in vitro* methylated DNA and DNA from whole male blood (0-100% matrix); 6-18 successfully modified DNA from tumour samples.

Reproducibility of the methodology was assessed for the entire set of samples by repeating the experimental procedure in two independent pyrosequencing runs with LINE-1 assay (see Figure 6.7). In the analysed cell lines, the DMSO treated controls showed a LINE-1 methylation status of 50-70% and consistency of the values between the two replicates was obtained.

Reproducibility was then assessed after treatment with DAC at 100 nM (D100) or 1 µM (D1) and in combination with TSA (D+T). After a 24 h treatment with 100 nM DAC, average LINE-1 methylation in the analysed cancer cell DNA decreased by 10-20%, showing a generalised effect. No evident difference was shown with higher dose of DAC or the combination with TSA.
Figure 6.7 Reproducibility of LINE-1 promoter methylation analysis using pyrosequencing. The four graphs show data from two independent experiments before/after treatment with DAC or DAC and TSA. Control samples showed higher methylation level compared to treatment with DAC (D100: DAC 100nM; D1: DAC 1µM) or DAC/TSA (D+T: DAC 100nM+ TSA 100nM).

A previous study of reproducibility and calibration of pyrosequencing demonstrated that experimental errors, classified as insertions, deletions, mismatches (substitutions) and ambiguous base calls were estimated between 2 and 4% of the measured methylation value (Huse et al., 2007). On this basis, the accuracy of the pyrosequencing method was maintained on heterogeneous
DNA templates suggesting that the higher error with experimental data was derived from the experimental manipulation of the sequences prior to pyrosequencing.

VI.3.2 Promoter methylation status of ALDH and ABC genes in selected cancer cell lines

The amplification regions were primarily designed to cover CpG islands (CGIs) overlapping with the 5' UTR of the target genes. When no CGI was annotated by the UCSC genome browser in close proximity of TSS, the sequence directly surrounding the 5' UTR was used for primer design, containing a variable number of CG sites. The initial methylation data were filtered to exclude poor quality measurements and adjustment of the design assay was made to exclude any region responsible of pyrosequencing failure. For excluded regions, PCR was identified as the leading cause of reaction failure.

The methylation status of three ALDH isoforms (ALDH1A1, 2 and 3A1) and two ABC transporters (ABCB1 and ABCG2) were examined in four different cell lines of the NCI-60 panel (A549, MCF7, HT29 and HeLa) before and after treatment with the demethylating agent decitabine (DAC) alone and in combination with histone deacetylation inhibitor trichostatin A (TSA). The pyrosequencing assays were designed for each gene as reported in Supplementary Information VI.A. According with the literature (Shaw et al., 2006), promoter methylation value below 10% is considered “unmethylated” and above 10% were considered “methylated”. ALDH1A1 gene did not show annotated CGIs in the promoter region; hence the analysed amplicon
(containing three CpG sites) was selected around the TSS, in a portion of DNA including three CpG sites.

Figure 6.8 ALDH1A1 promoter methylation analysis in HeLa cell line. Analysis was performed in three CpG sites selected around the TSS. Percentage methylation was reported above the respective sample as average value of the CpG sites under analysis. The graph shows the average data from two independent pyrosequencing runs before/after treatment with DAC (D100: DAC 100nM; D1: DAC 1µM) or DAC/TSA (D+T: DAC 100nM+ TSA 100nM). The first two samples (0, 100% Me) represented the DNA control.

The percentage methylation of ALDH1A1 analysis of the control samples was very low (between 4-6%) in A549, MCF7 and HT29, revealing an unmethylated status of the CpG sites around the promoter region (<10%) (see Supplementary Information VI.D.1 for methylation plots). However, DNA from HeLa cells (see Figure 6.8) showed a methylation level of 49,5% whereas DNA from DAC-treated HeLa cells was 29-33% methylated at each CpG site. As expected, a dose-dependent demethylation effect was revealed between 100 nM and 1 µM.
24 h treatment with DAC. Furthermore, combination of DAC with TSA did not show a greater effect on demethylation than DAC alone.

ALDH2 methylation status was investigated within a CGI of 481 bp. The analysed amplicon covered ten CpG sites. The promoter region was unmethylated in the analysed region for A549, HT29 and HeLa control with methylation status below 5 % for the three cell lines (see Supplementary Information VI.D.2 for methylation plots). Indeed, no variation was shown in the treated samples.

* \( p< 0.05 \), *** \( p< 0.001 \)

**Figure 6.9 ALDH2 promoter methylation analysis using pyrosequencing in MCF7 cell line.**
Analysis was performed in six CpG sites selected from a CGI of 481 bp. Percentage methylation was reported above the respective sample as average value of the CpG sites under analysis. The graph shows the average data from two independent pyrosequencing runs before/after treatment with DAC (D100: DAC 100nM; D1: DAC 1µM) or DAC/TSA (D+T: DAC 100nM+ TSA 100nM). The first two samples (0, 100% Me) represented the DNA control. In the MCF7 plot two additional samples showed the DMSO control and DAC treatment for 7 days, with new drug addition every 2 days.
In contrast, the MCF7 investigation (refers to Figure 6.9) showed a high methylated pattern in the control (84%) and a significant demethylation was determined after DAC treatment (60%), suggesting the efficacy of the drug in hypomethylating the analysed gene region.

Demethylation did not appear to be dose-dependent and no greater results were shown via combination treatment with TSA. However, as shown in the last two samples of Figure 6.9, prolonged exposure of MCF7 with multiple doses of 1 µM DAC for 7 days resulted into a further demethylation of around 47%.

Ultimately, ALDH3A1 promoter analysis of eight CpG sites, within a selected CGI upstream of the TSS, revealed a very low percentage methylation of the control samples (between 2-5%) in HT29 and HeLa cells, suggesting that the ALDH3A1 gene promoter is unmethylated in these cancer cell lines (see Supplementary Information VI.D.3 for methylation plots).

A549 (see Figure 6.10A) showed a methylation level above 10% (18%) and according to the pyrosequencing technology, a CGI with a methylation status above 10% is considered as methylated (Ronaghi, 2001).

DAC treatment of this cell line at 100 nM and 1 µM respectively led to demethylation of maximum 6%. A similar response was obtained after combination with TSA, suggesting that no further demethylation could be obtained after combination treatment.

However, DNA from MCF7 cells (see Figure 6.10B) showed a very high methylated pattern (92%) whereas DNA from DAC-treated MCF7 cells was shown to contain 71% methylation after 100 nM DAC treatment, and further
demethylation was observed after 1 µM DAC exposure (60%). This data revealed an interesting dose-dependent demethylation mechanism. Moreover, combination of DAC and TSA (both at 100 nM concentration) showed greater effect (67%) than compared with DAC 100 nM alone (71%). A dramatic demethylation was observed in MCF7 by longer exposure (7 days) to DAC 100 nM: methylation pattern changed from a highly methylated DMSO control (96%) to half of the original methylation level after treatment (45%).

A)

![Graph showing ALDH3A1 in A549](image)

*** p< 0.001
B)

Figure 6.10 ALDH3A1 promoter methylation analysis using pyrosequencing in A549 and MCF7 cell lines. Analysis was performed in eight CpG sites selected on a CGI of 353 bp. Percentage methylation was reported above the respective sample as average value of the CpG sites under analysis. The graph (A: A549, B: MCF7) show the average data from two independent pyrosequencing runs before/after treatment with DAC (D100: DAC 100nM; D1: DAC 1µM) or DAC/TSA (D+T: DAC 100nM+ TSA 100nM). The first two samples (0, 100% Me) represents control DNA. In the MCF7 plots two additional samples showed the DMSO control and DAC treatment for 7 days, with new drug addition every 2 days.

The final assay for ABCB1 included the analysis of only two CpG sites due to failure of the pyrosequencing run and consequent shortening of the initial designed assay. As previously proven by Reed et al. (Reed et al., 2008), contribution of ABCB1 downstream promoter had a fundamental role in the expression of ABCB1 transcripts in the MCF7 cell line. Here, pyrosequencing analysis showed hypermethylation of the ABCB1 downstream promoter in
MCF7 control cells (46%) and a demethylation effect after DAC treatment (35%), with no greater effect at higher DAC dose or in combination with TSA (see Figure 6.11A).

Same analysis was performed in A549 and HeLa cells, but the control samples showed unmethylated status (<10%) in the promoter region of both genes (see Supplementary Information VI.D.4 for methylation plots). A different profile was found in the HT29 cell line (see Figure 6.11B). Indeed, the methylation pattern in the untreated control was above 10% (13%) and the percentage methylation was slightly reduced via DAC treatment, however, perhaps due to the borderline results with an unmethylated pattern, the statistic analysis classified the results as no significant (p>0.05).

A)

* p< 0.05
Figure 6.11 ABCB1 downstream promoter methylation analysis using pyrosequencing.

Analysis is performed in two CpG sites selected on a CGI of 950 bp. Percentage methylation was reported above the respective sample as average value of the CpG sites under analysis. The graph (A: MCF7, B: HT29) show the average data from two independent pyrosequencing runs before/after treatment with DAC (D100: DAC 100 nM; D1: DAC 1 µM) or DAC/TSA (D+T: DAC 100 nM + TSA 100 nM). The first two samples (0, 100% Me) represents control DNA. In the MCF7 plot two additional samples show the DMSO control and DAC treatment for 7 days, with new drug addition every 2 days.

The second isoenzyme explored from the ABC family was ABCG2. The assay for ABCG2 was designed within a CGI of 698 bp and revealed an unmethylated status as a common feature in the four cell lines under analysis.

As shown in Table 6.14, the percentages of methylation were always comprised between 1-2% before and after treatment without cell-to-cell variation (see Supplementary Information VI.D.5 for methylation plots).
Table 6.14 ABCG2 downstream promoter methylation analysis in selected cell lines. Average of four CpG sites methylation status is reported. Experiment was repeated in duplicate, before/after treatment with DAC (D100: DAC 100nM; D1: DAC 1µM) or DAC/TSA (D+T: DAC 100nM+ TSA 100nM).

VI.3.3 Gene expression profiling of ALDH and ABC target genes

CGI methylation within promoter regions of genes have shown an association with transcriptional repression of the reciprocal gene (Bird and Wolffe, 1999). Here, the RNA expression levels of ALDH1A1, ALDH2 and ALDH3A1 together with ABCB1 and ABCG2 were quantified in the panel of selected cell lines A549, MCF7, HT29 and HeLa using qRT-PCR. PPIA or GAPDH housekeeping gene were, respectively, used to normalise the target gene expression and allow for differences in initial RNA concentration. These genes have been widely used as control genes for relative-quantification RT-PCR reactions, and optimisation showed that the selected housekeeping gene displayed consistent levels of expression over the established cell lines.

QRT-PCR was used to quantitatively examine the expression levels by measuring mRNA levels in the four cell lines for both ALDH and ABC selected
isoforms before and after treatment with DAC. Gene profiling was also analysed after combination of DAC with trichostatin A, TSA.

The expression was calculated successively in three individual experiments to verify values obtained, as shown in the graphs represented in Figures 6.12-14. Statistical analysis was performed as referred to in section VI.2.3.

In Figure 6.11, the A549 cell line showed high expression of ALDH1A1, 2 and 3A1 as revealed by the low number of cycles to achieve gene amplifications (Ct values around 25 cycles). No variation after DAC treatment with or without TSA was observed in the expression level of these genes.

* \( p < 0.05 \), ** \( p < 0.01 \), *** \( p < 0.001 \)

**Figure 6.12 A549 qRT-PCR plot.** PPIA was used for normalisation. Analysis was repeated in triplicate before/after treatments and values represent an average of the three independent experiments. P values summary and significance are reported in Table 6.17.
However, ABCB1 and ABCG2 expression in A549 showed no variation after DAC treatment.

The same analysis was performed in MCF7 cells (see Figure 6.13) and the expression level revealed a very pronounced reactivation of all genes after epigenetic treatment. High dose of DAC (1 µM) and the combination experiment with TSA showed a greater effect in gene reactivation (up to 75 fold change for ALDH3A1), revealing a trend in all the performed analysis. Statistical analysis showed high significance (p<0.001) for the majority of the expression values.

![Figure 6.13 MCF7 qRT-PCR plot. PPIA was used for normalisation. Analysis was repeated in triplicate before/after treatments and values represent an average of the three independent experiments. P values summary and significance are reported in Table 6.17.](image-url)
The highly pronounced reactivation after DAC treatment suggests that in some genes, such as ALDH3A1 or ABCB1, that methylation of the promoter region is associated with reduced mRNA expression.

In the HT29 profiling (see Figure 6.14), only ABCB1 showed gene reactivation after epigenetic treatment, perhaps due to the endogenous expression level of the whole panel of analysed genes within this cell model.

![HT29 qRT-PCR plot](image)

* p< 0.05, ** p< 0.01, *** p< 0.001

Figure 6.14 HT29 qRT-PCR plot. PPIA was used for normalisation. Analysis was repeated in triplicate before/after treatments and values represent an average of the three independent experiments. P values summary and significance are reported in Table 6.17.

Indeed, no variation was observed in the ALDH isoforms expression, but up to two fold increased expression was observed on ABCG2.
As shown in Figure 6.15, different results were obtained in analysis of HeLa cells, where ALDH1A1 mRNA level increased up to 8 folds after exposure to DAC at 100 nM, and up to 1000 folds when treated in combination with TSA 100 nM.

**Figure 6.15 HeLa qRT-PCR plot.** GAPDH was used for normalisation. Analysis was repeated in triplicate before/after treatments and values represent an average of the three independent experiments. P values summary and significance are reported in Table 6.17.

Gene reactivation after DAC/TSA combination treatment was shown also for ALDH2, ABCB1 and ABCG2 genes, but not for ALDH3A1.
VI.3.4 Western analysis of total protein extracts from selected cell lines

Immunoblotting detection was performed on the protein extract from the harvested cell lines before and after DAC treatment alone or in combination with TSA. Protein extract from A549 contained a high level of ALDH1A1 (see Figure 6.16A), which is in accordance with previous published data (Ucar et al., 2009) and detectable levels of ALDH2 (see Figure 6.16). However, no difference in the expression was obtained after epigenetic treatment.

ALDH3A1 protein level revealed a slight augmentation (up to 6 fold) after DAC treatment (see Figure 6.16), demonstrating a potential correlation between gene activation and protein translation. The ABCG2 protein was detected but no change was observed after DAC treatment. In contrast, the ABCB1 protein was downregulated after DAC treatment, which may be a result of an indirect “turn-off” signal activated via DAC and DAC/TSA exposure.

As shown in Figure 6.17A, protein extract from untreated MCF7 revealed no expression of ALDH1A1 and ALDH3A1, as already reported in past studies (Marcato et al., 2011; Sreerama and Sladek, 2001). No variation in the ALDH1A1 and 3A1 protein levels were observed after epigenetic treatment. However, DAC treatment led to ALDH2 protein induction and greater signal was detected after combination treatment with TSA (see Figure 6.17).
Figure 6.16 Western blot analysis ALDH and ABC proteins in A549 cell line. Untreated control and expression profiles following treatment for 24 h with DAC alone (100 nM) or in combination with TSA (100 nM) for a further 24 h.
Figure 6.17 Western blot analysis of ALDH and ABC proteins in MCF7 cell line. Untreated control and expression profiles following treatment for 24 h with DAC alone (100 nM) or in combination with TSA (100 nM) for a further 24 h.

The ABCB1 protein (see Figure 6.17) showed a silent status in the control sample, but after epigenetic treatment protein augmentation >2000 folds was observed, as shown by densitometry analysis of the signal intensity (refers to Table 6.15).
Table 6.15 Densitometry analysis of ABCB1 protein expression in MCF7 cell line. All the bands were normalised to GAPDH. Quantification was performed by calculating the ratio between the number of pixels in the treated samples and the untreated control.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Area/GAPDH</th>
<th>Normalised to untreated control</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCF7 Control</td>
<td>(3 \times 10^5)</td>
<td>1,00</td>
</tr>
<tr>
<td>MCF7 DAC</td>
<td>0,07</td>
<td>2216,36</td>
</tr>
<tr>
<td>MCF7 DAC+TSA</td>
<td>0,08</td>
<td>2701,60</td>
</tr>
</tbody>
</table>

As already widely demonstrated, analysis of ABCG2 in the MCF7 cell lysate showed high protein expression in the untreated sample (Kim et al., 2002), but exposure to DAC alone or with TSA led to a loss of protein signal.

The protein levels were also investigated in HT29 and HeLa cell lines (see Figure 6.18). As already proven in previous studies, western blotting analysis in HeLa-control cells did not show expression of ABCB1 (Lazo et al., 2010) or ABCG2 (Seamon et al., 2006). Furthermore, no expression of ALDH isoforms was observed in HeLa before or after drug exposure. In HT29, the only relevant protein level was reported for ALDH1A1 (see Figure 6.18B), where the protein level was downregulated after DAC treatment alone but re-established after combination of DAC/TSA, suggesting a role of TSA in the activation of the protein translation.
Figure 6.18 Western blot analysis of HT29 and HeLa cell lines. Untreated control and expression profiles, following treatment for 24 h with DAC alone (100 nM) or in combination with TSA (100 nM) for further 24 h.

No other endogenous target protein expression was detected in HT29 which is in agreement with the literature on ABCB1- and ABCG2-negative (Oh et al., 2011; Takakura et al., 2010).
VI.4 Discussion

The acquisition of aberrant CGI methylation is now a widely accepted hallmark of cancer (Plass and Smiraglia, 2006; Strathdee and Brown, 2002). It is evident therefore that defined groups of genes can become concordantly methylated in specific cancers which supports the concept of the "CpG island methylator phenotype" (CIMP) and that epigenetic gene silencing in the drug resistance field has recently been defined as polygenic (Glasspool et al., 2006; Segura-Pacheco et al., 2006). Indeed, previous studies showed that concordant promoter hypermethylation of multiple genes, which is known as the CIMP, may exist in gastric and colorectal carcinomas (An et al., 2005; Toyota et al., 2000), in ovarian cancer (Wei et al., 2006) and in neuroblastoma (Abe et al., 2005). The recent surge of interest in epigenetic studies has created high demand for reliable techniques to quantify DNA methylation in cells and tissues. In many instances, DNA methylation is not uniform, meaning that individual CpG sites are not 100% methylated or unmethylated within a cell population or even within individual cells. We selected pyrosequencing technique (Reed et al., 2009; Ronaghi, 2001) to gain insight in the epigenetic regulation of the promoter regions of a selected panel of genes and we combined the qRT-PCR gene profiling to establish relationship between promoter methylation and mRNA expression.

In this “pilot” study, we investigated the methylation status relative to the promoter regions of ALDH1A1, ALDH2 and ALDH3A1 in a heterogeneous panel of cell lines, demonstrating that these ALDH isoforms are controlled via promoter methylation in a tissue-specific manner and that the methylation
pattern of these genes may be important in transcriptional regulation. Till date, very little is known regarding the potential epigenetic control of ALDH enzymes and their susceptibility for reactivation by DAC exposure. The data generated in this study can contribute towards a better understanding of ALDH biology.

The promoter region of ALDH3A1 was demonstrated to be heavily methylated in MCF7 and A549 cell lines, and exposure of these cell lines to DAC led to demethylation and increased mRNA levels. Analysis of the protein status prior and post epigenetic treatment revealed ALDH3A1 protein augmentation in A549 but not in MCF7 cells, suggesting either a (i) tissue-specific related response or (ii) a difference in post-translational pathways. Combination treatment with DAC and TSA also translated into ALDH3A1 protein augmentation, but only in A549 and not in MCF7, indicating synergistic interplay between DNA methylation and histone acetylation. This result is consistent with a previous report that gene silencing conferred by methylated DNA can be also influenced by inhibition of HDAC, facilitating the remodelling of chromatin and transcriptional activation (Irvine et al., 2002; Lee et al., 2008).

It was demonstrated that no CGI within the promoter region of ALDH1A1 existed, therefore investigations in the ALDH1A1 methylation status was carried out in proximity to the transcription start site. The data showed only significant hypermethylation in HeLa cells, which could be modulated by DAC-treatment. A dramatic 1000-fold gene reactivation was observed when low dose DAC (100 nM) was combined with non-toxic dose of TSA (100 nM). This result suggests ALDH1A1 gene silencing can be reactivated via modulation of both the ALDH1A1 gene and histone acetylation. DAC and TSA have been discussed as
synergistic therapies (Esteller, 2009), but as TSA is considered a pan HDAC inhibitor it is at present difficult to assess how this synergy is achieved.

The ALDH2 gene was hypomethylated in all the analysed cell lines with exception of MCF7 cells. An inverse relationship between expression and DNA demethylation after DAC treatment was observed. However, no protein augmentation was detected after gene reactivation of ALDH1A1 and ALDH2, implying that other mechanisms are required for post transcriptional control.

In attempt to understand the regulation of ALDHs and their potential role as relevant biomarkers, this study has concerned epigenetic regulation of isoforms from ALDH1-3 sub-families. Moreover, due to the affinity of many anthraquinones for ABC drug transporters (Doyle and Ross, 2003) it was also desirable to understand whether the principal efflux pump proteins ABCB1 and ABCG2 were affected by epigenetic mechanisms. Increased ABCB1 expression upon DAC-mediated demethylation of its downstream promoter was observed in MCF7 cells. This is at large in agreement with a recent study examining the role of methylation in the ABCB1 promoter region where correlation was made between ABCB1 promoter hypermethylation and acquisition of docetaxel resistance in MCF7 cells (Reed et al., 2008).

Data presented in this chapter revealed that significantly increased ABCB1 gene expression after DAC treatment was also correlated with protein augmentation. In contrast, the ABCG2 gene was not silenced via methylation of the gene promoter in any of the four cell lines investigated.

These findings are in line with other studies on established cancer cell lines from other tissues including pancreatic (Chen et al., 2012), gastric (Hiraki et al., 2010) and renal (To et al., 2006). This could suggest lack of ABCG2 promoter
methylation as a cancer specific phenotype (Hiraki et al., 2010). However, ABCG2 promoter methylation has been observed in multiple myeloma (Turner et al., 2006) and in leukaemia (Bram et al., 2009), resulting in epigenetic silencing. Interestingly, drug selection with topotecan or mitoxantrone has been shown to induce complete demethylation of the ABCG2 promoter in both leukaemia and ovarian carcinoma model cell lines (Bram et al., 2009), suggesting that ABCG2 expression is controlled in part by methylation of its promoter.

In summary, this study provides some insight into how selected ALDH isoforms and ABC transporters may be under epigenetic control and how their expression can be modulated with epigenetic drugs such as DAC and TSA, either alone or in combination. The data may also suggest tissue-specific silencing of both ALDH and ABC proteins, but further studies are needed to support preliminary investigations presented in this chapter. Indeed, the concept of the potential epigenetic control of ALDH family could negatively impact the use of i.e. ALDH1A1 as a biomarker for SCs as well as complicating molecular fluorescent probe design and development. In contrast, this study demonstrates that ABCG2 transporter appears at large not to be under epigenetic control, hence supporting it as valid marker of MDR cancer cells or CSCs.
Supplementary Information VI.A

Final assays for target genes

1. Assay for ALDH1A1

The ALDH1A1 gene was transcribed from the antisense strand. The position of the gene on chromosome 9 according to UCSC was 74705407-74757789. No CGIs were found according to UCSC genome browser. The sequence chosen for the DNA-methylation analysis for primer set was on the antisense strand of the genomic DNA near the transcription start site (TSS). The final assay was designed as shown in the sequence and Table VI.A.I.

Genomic sequence searched:

```
AGGGCCCTTTTCTTCCCCAACACGCCCCTTGTCTGAGACACGGCCCTTCTACTGAGA
ACAAAGTGCCTTCTTATAGCCCTTTTCAATCTCAATCTCTGATTCCAA
GTCTGTACAGAAACACCAAGGTACATAGTACATATCTAAAAAGCATGAGAA
GTCAAAAAATATAACTACCCGCTTCTTTCTTCAATCTGATCCAGTAC
ACTTTATACACGGTTTCGCTTTTGAATATTACATCTGAAAATAGTCA
ACTGTCTGTGAGTACATCTGGACTTTATTTAGTCAATAGGAGTACAGTGGCT
TGACGTTTTTCTTTTTTCTCTCTGAGTTTTGATATCCAGTAC
```

Bisulfite converted sequence:

```
AGGGTTTCTTCTTTTTTTTAATAGTAATTATTGATTTTGGGAGATGGATTT
GATTTTGTGGACAGTTTTGGTACAGATTGTTGAGTTTTTTTATTTGA
GAATAGGTGCTTTTTTTTAGATTTTTTTTTAAAAATTTTTTTGATTT
TAAGTGTGTTAGGAAATTAGAAAAATTATTAGATTTAAAAAGTATAG
AGAAGTAAAAAAAATATATATGTTTTTAGGTTAGAGTAGTTGTTTG
TAATTATTATTATTAGCTTTCCGTTTTGAATTTTTATTTTTTATTTTTG
TAGGTGATTGTTTTTATTGAGTTTTTATTTATTTTTTATTTTTTATTTTT
TAGGTTGATTGTTTTTATTGAGTTTTTATTTATTTTTTATTTTTTATTTTT
```

The biotin-modification was made on the reverse primer. The position of the TSS according to UCSC on the (-) strand was 74757787 (highlighted with red in the sequence above). The
amplicon length was 208 bp and the assay covered three CpG sites, including one CpG site analysed by Illumina (Illumina cg16601861, highlighted in yellow in the sequence above) (Illumina, CA, USA).

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Legend</th>
<th>Length (nb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forward primer 5'-GGTTTTTTAAGGATTTAAGTTAAAAGT - 3' Highlighted in green 27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reverse primer 5'-Biotin-AAAACACAAAATAACTCAACAAATT -3' Highlighted in pink 27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sequencing primer 5'- TTGGTTTTGAGTTTGGTTTTAATTTAAT -3' Highlighted in blue 24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sequence to analyze CCGATCGAGTGTAATAAACTTTAGCCCGTGCAAG Underlined 38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Converted sequence to analyze CCGATCGAGTGTAATAAACTTTAGCCCGTGCAAG Underlined 38</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table VI.A.1 Primers and sequence to analyse (highlighted in the genomic and bisulfite converted sequences) for ALDH1A1 gene.

2. Assay for ALDH2

The ALDH2 gene was transcribed from the sense strand and was positioned on chromosome 12. A CGI with 481 bp was positioned between 112204499-112204979 loci on the gene according to the UCSC, covering 58 CpG counts. The selected CGI included the final assay as shown in the sequence and Table VI.A.II.

Genomic sequence searched:

CGAGCTGGTTGACGGGGGGCAGCTGCGCTGCGCGTTGTGGGGCTCGCGTTTTTCCGGCCCGAGTCCCC
CGCCCCGGGGGAGCAGGGGCACACCGCGCTTTTGGTGCTGTGGCGCCGGAGGTGGCCTCCCTGTTCTGGCTGCAGGCCCCTAGGAAGGCCCCGCGCCGCCGTGGGCCTTAGTG

Forward primer
Seq primer
TSS
Reverse primer
Seq Analyzed
**Chapter VI**

**Bisulfite converted sequence:**

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Legend</th>
<th>Length (nb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forward primer</td>
<td>5'-AGGAATGGTAGAGGTAGT-3'</td>
<td>Highlighted in green</td>
</tr>
<tr>
<td>Reverse primer</td>
<td>5'-Biotin-AACAAAAACTAAAATCTCAAAACCACTA-3'</td>
<td>Highlighted in pink</td>
</tr>
<tr>
<td>Sequencing primer</td>
<td>5'-GAATGGTAGAGGTAGTT-3'</td>
<td>Highlighted in blue</td>
</tr>
<tr>
<td>Sequence to analyze</td>
<td>CGGCCCGCGCCCGCGCCCCCGCCCTTCCA TGGCTGccCGCGGCGGCGGAGCGG GGTcGGCT</td>
<td>Underlined</td>
</tr>
<tr>
<td>Converted sequence to analyze</td>
<td>TCGTTTTCCGTTCCGTTTGGTTTTTTGTTTAGTTCG</td>
<td>Underlined</td>
</tr>
</tbody>
</table>

**Table VI.A.II** Primers and sequence to analyse (highlighted in the genomic and bisulfite converted sequences) for ALDH2 gene. The biotin-modification was made on the reverse primer. The position of the TSS according to UCSC on (+) strand was 89079780 (highlighted in red in the sequence above) and the amplicon length was 109 bp and the assay covered ten CpG sites.

**3. Assay for ALDH3A1**

The ALDH3A1 gene, located on chromosome 17 was transcribed from the antisense strand. A CGI with 353 bp was positioned between 19648140-19648492 loci on the gene according to the
UCSC, covering 30 CpG counts. The selected CGI included the final assay as shown in the sequence and Table VI.A.III.

Genomic sequence searched:

```
GCCGAGCCCAGGGAAGTCCCTTCCTATAGAATTCAGGCAGGGTGGGAGGC
```

*Forward primer*

```
AGGGCGCGCTCGTGCCCCTCAGCCAGCTGCAGGTGCTCTCTGTCCCCAGG
```

*Sequencing primer*

```
CGCCATGAGCAAGATCAGCGAGGCCGTGAAGCGCGCCCGCGCCCTTCA
```

*Seq Analyzed*

```
GCTCGGGCAGGACCCGTCCGCTGCAGTTCCGGATC
```

*Reverse primer*

```
CAGCAGCTGGAGGCG
```

Bisulfite converted sequence:

```
GTCGAGTTTAGGGAAGTTTTTTTTTATAGAATTTAGGTAGGGTGGGAGGT
```

*Forward primer*

```
AGGGCGCGTTCGTGTTTTTTAGTTAGTTAGTTAGTTAGTTTTAGG
```

*Sequencing primer*

```
CGTTATGAGTAAGATTAGCGAGGTCGTGAAGCGCGTTCGCGTCGTTTTTA
```

*Seq Analyzed*

```
TTGTAGCGTTTGATTTAGGAGTAGGAGTAGGAGTTGGTGGGCGCGTTGGT
```

*Reverse primer*

```
CGGGGTTATGCGATGGGGCGTGGGGGGAGGGGTTGGGGAGAGTGCGATTT
```

The bisulfite-modification was made on the reverse primer. The position of the TSS according to UCSC was on (-) strand 19651621 (highlighted in red in the sequence above), the amplicon length was 256 bp and the assay covered 11 CpG sites. Due to failure of the pyrosequencing run, the last three CpG sites were deleted from the assay.
4. **Assay for ABCB1**

The ABCB1 gene, located on chromosome 7 (position 86970884-87180500) was transcribed from the antisense strand. ABCB1 downstream gene promoter sequences were based on Reed et al. study (Reed et al., 2008). The same study described also ABCB1 upstream promoter, but in the analysed samples the upstream assay failed after multiple attempts, so it was excluded from the analysis. A CGI with 950 bp was positioned between 87256959-87258444 loci on the gene according to the UCSC, covering 51 CpG counts. The selected CGI included the final assay as shown in the sequence and Table VI.A.IV.

**Genomic sequence searched:**

```
TTTCAATCCATTATCATCAATAAAGGATGAAACAGATGTAACCTCAGAAA
CTGTCAGGCAATGCTGAAAGGACACACTGCAGAGAAAAATTCTCTAGCCT
TTCTCAAAGGTGTGAGAACAGAAAAAGTATACTTGCAATTGGAGAGGTG
AGTTTTTGATTAAAATGTTTATGAGAATGCAGAACCACTTCTCAGTAGA
ACTACGGCTCTGCTGATTAGGATAATGGAGACTATTATGTGAACCTTTAAG
ACGTTGCTCATATAAGTTGAAAGTGTGCCAATTGGAGAGGTCAGCTGCAG
TCTCTAGCTTCTTTCTAAGGTGAAGGGCAGACGCCGCAGACATCTCCCT
CTGAGAATCTCAACTTCCAGGATGATTGCTGCTTCCCTAGAAATATCTG
ATGCAGGGCGAGGAGCTCAATGCTGACATGGCGGCTCGAGATTAAGCTG
TGCTCTGCAGATTCGGGATGAGCTGAGCAAAGTATGATGGGTGAGAGGA
ACGGCGCCGGGGCGTGGGCTGAGCACAGCCGCTTCGCTCTTTGCCACA
```

**Forward primer**

```
AGTTTTTGTATTAACTGTATTAAATGCGAATCCCGAGAAAATTTCCCTTA
```

**Reverse primer**

```
ACACGCGCCGGGGCGTGGGCTGAGCACAGCCGCTTCGCTCTTTGCCACA
```

**Sequencing primer**

```
TGTGGCTGGCCAGAG
```

**Sequence to analyze**

```
TTCAGCACAGGCCCCCCTGGGCTTTCTATTTCGTGCAGGC
TTCAGCACAGGCCCCCCTGGGCTTTCTATTTCGTGCAGGC
TTCAGCACAGGCCCCCCTGGGCTTTCTATTTCGTGCAGGC
```

**Converted sequence to analyze**

```
TTTAGTTTGGGAGTGGATTCCGGTTGTAGTTCTCGAGATCTCGGATT
```

**Table VI.A.III Primers and sequence to analyse (highlighted in the genomic and bisulfite converted sequences) for ALDH3A1 gene.**
The biotin-modification was made on the forward primer. The position of the TSS according to UCSC was on (-) strand 87180530 (highlighted in red in the sequence above), the amplicon length was 348 bp and the final assay covered only 2 CpG sites, due to failing of the run with the original 6 CpG sites.
Table V.A.IV  Primers and sequence to analyse (highlighted in the genomic and bisulfite converted sequences) for ABCB1 gene.

5. **Assay for ABCG2**

The ABCG2 gene was located on chromosome 4 and transcribed from the antisense strand. A CGI with 698 bp was positioned on between 89079757-89080454 loci on the gene according to the UCSC, covering 77 CpG sites. The selected CGI included the final assay as shown in the sequence and Table VI.A.V.

**Genomic sequence searched:**

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Legend</th>
<th>Length (nb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forward primer</td>
<td>_5’<em>Biotin</em> AGGTGTTAGGAAGTAGAAGGATATA-3’</td>
<td>Highlighted in green 28</td>
</tr>
<tr>
<td>Reverse primer</td>
<td>5’-CTATTCCTACCAACCAATCAACCTCA-3’</td>
<td>Highlighted in pink 27</td>
</tr>
<tr>
<td>Sequencing primer</td>
<td>5’-CAATCAACCTCACCACAA-3’</td>
<td>Highlighted in blue 19</td>
</tr>
<tr>
<td>Sequence to analyze</td>
<td>GAGGAATCAGATTAGTCAATCCCCGGGCGGGGACACTGCTGGTGCTATCCGGGCCCGGCAGTCGGGGCCACGCCTCACCCCCGCCCGCGAACCCCGACCTGGGGAAACCCGGGGCGCTGGGGAGGGGCCACTGCGTTCACTCTGGCGGTCCACAGCCCGAAGCGCGGCTTAGGAAGTTCGTGTCAGCGCTGCCTGAGCTCGTCCCCTGGATGTCCGGGTCTCCCCAGGGGCCACCCG</td>
<td></td>
</tr>
<tr>
<td>Converted sequence to analyze</td>
<td>TAACTACTCCCRACCCRAATTAACTAAATA CTAATTCCTC</td>
<td>Underlined 40</td>
</tr>
</tbody>
</table>

**Sequence Legend Length (nb)**

<table>
<thead>
<tr>
<th>Forward primer</th>
<th>Sequencing primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>_5’<em>Biotin</em> AGGTGTTAGGAAGTAGAAGGATATA-3’</td>
<td>Highlighted in green 28</td>
</tr>
<tr>
<td>5’-CTATTCCTACCAACCAATCAACCTCA-3’</td>
<td>Highlighted in pink 27</td>
</tr>
<tr>
<td>5’-CAATCAACCTCACCACAA-3’</td>
<td>Highlighted in blue 19</td>
</tr>
</tbody>
</table>

**Sequence Legend Length (nb)**

- Forward primer
- Sequencing primer
- Analyzed sequence
- Reverse primer

**TSS**
Bisulfite converted sequence:

CGAGTAGCGTTTGTGATTGGGTAATTTGTGCGTTAGCGTTTTCGGTGTTT
CGGCCGTTTCGGTATAGTGACGGCGATTAATTTTAGTTAGTATTAGCAGAGT
ATGCTTAGTTTAATGACGCTTGGGATTTCGGTATTATTTTGCTAGTTATTA
TTATTTGCTGATTTGCCAGGCGGAGGATTTGGATTGTGTGGGATTTGGCGTTG
TTAGCCTGGGTTTTAGTTAGGTTAGACGAGGTATTGATTAGTTTAATGAGCGTTTGGTGATTTTCGTAGTTAATTATTTTGG
GCGTCTATCGCCTGTCGTAGTCGCGTTACGCTATTATTTGTCGTCG
AATTTCGATTTGGGAAAATTCGGGGCTTGGGGAGGGTTATTGGCTTTA
GTTTGGCGTTATTTATAGTCTGGAGGGTTCTGGGTTTTCTAGCCTGCTG
TTGTTGGAGTTGTTGTTTGGAGTTGCTCTGGGTTTTTTATAGCCGTTATTTC
GCGTTTTATCTGCTGCTGTTTGGAGTTGCTCTGGGTTTTTTATAGCCGTTATTTC
GCGCGGAGGCGTTCGGGCGGCGTGCGGGGTGGGTTGGGTTTTTTTGGGCGGTTAT
GCAGAGGGAGAGGGAGGAGTGTTTGGCTTGTCCCTG

**Forward primer**

**Sequencing primer**

**Reverse primer**

---

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Legend</th>
<th>Length (nb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forward primer</td>
<td>5'-GGTGTTGATGTTGGTAATTTTG-3'</td>
<td>Highlighted in green</td>
</tr>
<tr>
<td>Reverse primer</td>
<td>5'-Biotin-CAAACTAAAATCACCCTACC-3'</td>
<td>Highlighted in pink</td>
</tr>
<tr>
<td>Sequencing primer</td>
<td>5'-GTTGTAATTATTGTTA-3'</td>
<td>Highlighted in blue</td>
</tr>
<tr>
<td>Sequence to analyze</td>
<td>TTGCGTTGATCCGAGGGCG</td>
<td>Underlined</td>
</tr>
<tr>
<td>Converted sequence to analyze</td>
<td>TTTGCGTTGATCCGAGGGCG</td>
<td>Underlined</td>
</tr>
</tbody>
</table>

**Table VI.A.V** Primers and sequence to analyse (highlighted in the genomic and bisulfite converted sequences) for ABCG2 gene.

The biotin-modification was made on the reverse primer. The position of the TSS according to UCSC on (+)-strand was 89079780 (highlighted in red in the sequence above), the amplicon length was 216 bp and the assay covered 4 CpG sites.
### Supplementary Information VI.B

#### qRT-PCR oligonucleotides and cycling conditions

<table>
<thead>
<tr>
<th>Target gene</th>
<th>Probe Number</th>
<th>Forward Primer</th>
<th>Reverse Primer</th>
<th>Amplicon Length (bp)</th>
<th>Tm(°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALDH1A1</td>
<td>81</td>
<td>5'gltgtgtggtggggtggtgg 3'</td>
<td>5'gggctttgggctttgggcttt 3'</td>
<td>91</td>
<td>59-60</td>
</tr>
<tr>
<td>ALDH2</td>
<td>26</td>
<td>5'aagagtccctgctacgtgga 3'</td>
<td>5'ctggccactgttcatgaattt 3'</td>
<td>87</td>
<td>60</td>
</tr>
<tr>
<td>ALDH3A1</td>
<td>85</td>
<td>5'aagagtccctgctacgtgga 3'</td>
<td>5'ctggccactgttcatgaattt 3'</td>
<td>87</td>
<td>60</td>
</tr>
<tr>
<td>ABCB1</td>
<td>90</td>
<td>5'ccatagctcgtgccctctgg 3'</td>
<td>5'agggcttcttggacaacctt 3'</td>
<td>100</td>
<td>60</td>
</tr>
<tr>
<td>ABCG2</td>
<td>29</td>
<td>5'ccatagctcgtgccctctgg 3'</td>
<td>5'agggcttcttggacaacctt 3'</td>
<td>83</td>
<td>59</td>
</tr>
</tbody>
</table>
Chapter VII

Epigenetic modulation of ALDH expression in prostate cancer
VII.1 Introduction and Aims

VII.1.1 Background

To date, the knowledge and understanding of the biochemical and molecular processes implicated in the pathogenesis of prostate cancer are rapidly growing. Clinical benign prostatic hyperplasia (BPH) has been described as a non premalignant condition with no correlation to prostatic cancer (PCa), however it may well give rise to a “transition cancer zone” (Maitland et al., 2011). Furthermore, manifestation of symptomatic BPH is generally recognised in men over the age of 50 years and therefore the aging factor needs to be considered as a risk factor relating to the disease (Kirby et al., 2005). The normal, mature human prostate consists of a high level of cellular organisation (Figures 7.1 and 7,2).

![Figure 7.1 Schematic representation of the architecture of the human normal prostate epithelium.](image)

The human prostate epithelium consists of a basal layer of relatively undifferentiated basal cells and a luminal layer of terminally differentiated secretory luminal cells. Figure adapted from Oldridge EE (Oldridge et al., 2011).
Briefly, three phenotypically and morphologically distinct cell types are included within a two-layered epithelium: columnar secretory, luminal cells, relatively undifferentiated basal cells and rare neuroendocrine cells (Oldridge et al., 2011; Tang et al., 2007).

![Hierarchical Pathway of Human Prostate Epithelium](image)

**Figure 7.2** The hierarchical pathway of human prostate epithelium. SCs are mostly quiescent and generate rapidly proliferating transit-amplifying cells. These cells still maintain a degree of multipotency, but commit to differentiation giving rise to committed basal cells. Through the differentiation process, terminally differentiated secretory luminal cells are then formed. Figure from Oldridge EE (Oldridge et al., 2011).

Importantly, it has been established that PCa is primarily a disease of the luminal secretory epithelium in the prostate gland (Maitland et al., 2011). Indeed, differentiated luminal cells (see Figure 7.2), which constitute the major component of normal and malignant prostate, are the ‘factory’ within the epithelium generating secretory products like prostate-specific antigen (PSA).
and prostatic acid phosphatase (PAP) and can be identified by the expression of androgen receptor (AR) (Maitland and Collins, 2008).

There is a consistent body of evidence that SCs reside in the basal layer (Lang et al., 2009; Signoretti et al., 2000). Within this layer, SCs represent a small subpopulation of quiescent cells with high proliferative potential in vitro with potential for reconstructing functional prostate acinar structures in vivo (Oldridge et al., 2011; Richardson et al., 2004).

### VII.1.2 Prostate cancer and ALDH

PCa is the most commonly diagnosed cancer in men and the second leading cause of death. Although PCa detected at an early stage can be successfully eradicated by radical prostatectomy and radiotherapy, there is a great need for novel therapies against castration-resistant PCa and metastatic disease (Jemal et al., 2009; van den Hoogen et al., 2010). Although it has been estimated that prostate tumours comprise only 0.1% of SCs, it is rationalised that such cells are crucial to the formation of drug-resistant tumours and metastasis, causing patients to relapse (Collins et al., 2005; Kelly and Yin, 2008).

A study from Burger et al. (Burger et al., 2009) demonstrated that murine prostate stem/progenitor cells expressed high levels of ALDH activity and that almost all of these cells co-expressed Sca-1 (stem cell antigen-1). Furthermore, the ALDH\textsuperscript{hi} cells had higher in vitro and in vivo proliferative potential than cells expressing low levels of this enzyme, revealing the potential of ALDH as functional marker in these specific tumour cells. On this basis, van den Hoogen
et al. (van den Hoogen et al., 2010) showed that the ALDH$^{\text{hi}}$ subpopulation of human PCa cells had not only enhanced clonogenicity, migration, and tumorigenicity but also readily form metastases \textit{in vivo}. It is important to note that these observations were derived on the basis of the Aldefluor assay that is proposed to target ALDH1, but may not be selective for the 19 different isoforms identified to date (Marchitti et al., 2008) (refers to section I.5.1 for details). Indeed, in PCa, it has been demonstrated that higher expression of ALDH isoforms other than ALDH1A1 was found in Aldefluor-positive cells, including ALDH3A2, ALDH4A1, ALDH7A1, ALDH9A1 and ALDH18A1 (van den Hoogen et al., 2011). Notably, ALDH7A1 expression was at particularly high levels in cell lines, primary tissue and matched bone metastasis samples, suggesting that at least for PCa (van den Hoogen et al., 2011a) ALDH7A1 is contributing to the Aldefluor activity of these cells.

Aberrant DNA methylation is known to be an early molecular event in PCa development (Baylin, 2005), but only two studies have described the epigenetic modulation of ALDHs in PCa (Ju et al., 2010; Kim et al., 2005). Importantly, one of the studies (Kim et al., 2005) demonstrated dense hypermethylation of the ALDH1A2 gene promoter and reported it as a candidate tumour suppressor gene in PCa. Given the importance of epigenetics in cancer (see section I.6.3), it is vital to understand how ALDH is expressed and regulated under such circumstances, and how such information can be translated into a clinical setting.
VII.1.3 ALDH investigation in human primary cells

Preliminary studies by Prof. Norman Maitland’s group (data not published, but shown in Figure 7.3), using DNA microarray analysis of stem and progenitor cell types isolated from benign and PCA tissues revealed heterogeneity in ALDH expression within these cell populations.

**Figure 7.3** Microarray analysis of selected ALDH isoform expression levels in PCA primary cells. Analysis was performed within primitive SCs (S) and their more differentiated progeny (C) from both clinical PCA (P) and benign (N) samples. ALDH isoforms are shown at the bottom of the graph with the individual probes above. Coloured horizontal bars represent individual patient values with black bar indicating the mean of each data set. The shaded rectangles show the interquartile range (IQR, a measure of statistical spread) for the data. The lower solid red line highlights the background expression limit for the assay (data used with permission from Prof. Norman Maitland, University of York).
Gene expression analysis of all 19 ALDHs in 12 patient samples demonstrated high expression of ALDH2, 3A2, 4A1, 9A1, 16A1 and 18A1, but not ALDH1A1. ALDH1A3 was observed to be the highest expressed isoform, which also has been shown to play an important role in breast cancer (Marcato et al., 2011).

Analysis of the ALDH isoform expression levels between PCa SCs and their differentiated progeny revealed that only changes in expression of ALDH3B2 and ALDH16A1 isoforms (highlighted in red on Figure 7.3) were statistically significant.

Further proof of ALDH1A1 not being the key ALDH is evident from microarray analysis on ovarian cancer bioptic samples (data not published, but kindly provided for the purpose of this thesis by Dr. Phil Burns, Leeds Institute of Molecular Medicine) ALDH1A3 and ALDH9A1 were shown to be the predominant isoforms in the set of analysed samples whereas ALDH1A1 expression was shown to be insignificant (see Figure 7.4). Moreover, ALDH2, as well as ALDH3A2, both revealed higher level than ALDH1A1 but no comparison with ALDH3A1 could be made as it was not present in the analysed chip.
As shown in Figure 7.5, clusterisation of the naïve samples versus the treated ones illustrated a higher expression of ALDH1A1 and ALDH2 isoforms after treatment, which is in agreement with a fundamental role of these enzymes in drug resistance acquisition (Moreb et al., 2000; Sladek et al., 2002).
VII.1.4 Rationale and aims of the investigation

As reported in section I.4, ALDH activity has emerged as an important identification and purification marker for cancer stem cells (CSCs). Despite the plethora of information on ALDH1 as a marker of cells with “stemness” properties (see section I.4.3), there is little literature with regard to the epigenetic control of ALDH (Kim et al., 2005). Given the importance of epigenetics in carcinogenesis and resistance, such information could have a profound impact on the understanding of cancer development and the identification and characterisation of CSCs. Accordingly, this study focussed on the expression and regulation of ALDH genes in cancer cells as well as in benign and malignant prostatic tissues derived from patients. In addition, decitabine (DAC), which modify epigenetic chemical tags (see section I.4.4 for details), was used to determine whether specific ALDH genes are under

Figure 7.5 ALDH1A1 and 2 expressions before/after chemotherapy. Clusterisation of microarray analysis showed on Figure 7.4 for ALDH1A1 and ALDH2 isoform in treated versus drug-naive primary bioptic samples. Y axes represent log2 expression.
epigenetic control, and if so, assess the consequences of modulating such epigenetic signatures. In summary, in order to understand the implications of ALDH regulation in PCa, the aim in this chapter was to use methodologies already optimised and established in Chapter VI to gain insights into:

1. Gene expression level of selected ALDH isoforms (ALDH1A1,2 and 3A1) using qRT-PCR with already optimised primer sets (see section VI.3.3.4) in a wide panel of prostate cell lines and in primary prostate epithelial cultures (benign and malignant bioptic tissue);

2. Pyrosequencing-based methylation analysis of selected ALDH promoter regions in prostate cell lines versus primary prostate epithelial cultures;

3. ALDH gene status in the prostate epithelial hierarchy (benign and malignant) in order to investigate cancer specific features. The analysis was performed in three selected cell subtypes: committed basal (CB), transit amplifying (TA) and stem cells (SC) to explore the cell type-dependent ALDH regulation;

4. ALDH gene expression status before and after DAC treatment as well as potential methylation pattern alterations in cell lines versus primary cultures.
VII.2 Materials and Methods

VII.2.1 Maintenance of mammalian prostate cell lines

Mammalian prostate cell lines were cultured as outlined in table 7.1 and propagated as previously elucidated in section IV.3.1.2. Cells were cryopreserved as explained in section VI.3.1.4.

<table>
<thead>
<tr>
<th>Cell line</th>
<th>Origin of cell line</th>
<th>Culture Media</th>
<th>Approx. frequency of subculture</th>
<th>Dilution upon subculture</th>
</tr>
</thead>
<tbody>
<tr>
<td>PNT1A</td>
<td>benign immortalized prostate epithelial cell</td>
<td>R10</td>
<td>4 days</td>
<td>1:10</td>
</tr>
<tr>
<td>PNT2C2</td>
<td>benign immortalized prostate epithelial cell</td>
<td>R10</td>
<td>4-5 days</td>
<td>1:10</td>
</tr>
<tr>
<td>PC3</td>
<td>Human prostatic adenocarcinoma metastatic site in bone</td>
<td>H7</td>
<td>3-4 days</td>
<td>1:5-1:10</td>
</tr>
<tr>
<td>P4E6</td>
<td>Well-differentiated prostate cancer</td>
<td>K2</td>
<td>3-4 days</td>
<td>1:2-1:4</td>
</tr>
<tr>
<td>RC165</td>
<td>benign immortalized epithelial prostate cell</td>
<td>KSFM</td>
<td>4-5 days</td>
<td>1:10</td>
</tr>
<tr>
<td>DU145</td>
<td>androgen independent prostate cancer cell</td>
<td>R10</td>
<td>3 days</td>
<td>1:10</td>
</tr>
<tr>
<td>LNCaP</td>
<td>Human prostatic adenocarcinoma metastatic site in supraclavicular lymphnode</td>
<td>R10</td>
<td>3-4 days</td>
<td>1:5</td>
</tr>
<tr>
<td>STO</td>
<td>Mouse embryonic fibroblast</td>
<td>D10</td>
<td>3-4 days</td>
<td>1:10-1:20</td>
</tr>
</tbody>
</table>

Table 7.1 Cell culture conditions of prostate cell lines. For media compositions refer to Appendix II.
VII.2.2 Treatment with decitabine

The listed cells were seeded at a concentration of 1x10^5 cells in 75cm^2 flasks. The next day, treatment of cells with 1 µM decitabine (DAC) (Sigma) was started, and fresh drug was added every 24 h for 96 h (total exposure time) before the cells were harvested for DNA and RNA extraction.

VII.2.3 Maintenance of primary cultures

Primary prostate epithelial cells (see Supplementary Information VII.A at the end of the chapter) were cultured on BioCoat™ Collagen I cellware (BD Biosciences) in complete stem cell medium (Chaproniere and McKeehan, 1986; Hoshi and McKeehan, 1984). Primary cells were co-cultured with irradiated STO feeder cells where indicated and were typically sub cultured 1:2 when approximately 80% confluent. Culture medium was typically replenished every second day or when cell debris was present in the culture supernatant.

Primary prostate fibroblasts were cultured in R10 medium supplemented with 100 U/mL of penicillin and 100 µg/mL of streptomycin.

VII.2.4 Irradiation of mouse embryonic fibroblasts

For mitotic inactivation by irradiation, STO embryonic fibroblasts (Sandoz inbred mouse, Thioguanine- and Ouabain-resistant) were trypsinised at approximately 80-90% confluence and were sedimented by centrifugation. The cell pellet was washed in 10 mL PBS per 75 cm^2 culture surface before cells were resuspended in 10 mL complete SC medium per 100 cm^2 of culture surface and transferred into a centrifuge tube. Cells in centrifuge tubes were treated with a
radiation dose of 60 Gy. Following irradiation, cells were stored at 4 °C for up to 7 days before use.

**VII.2.5 Selection of sub-populations of primary prostatic cells**

As elucidated in Figure 7.6, sub-populations were selected from the prostate primary cultures by using CD133 and α2β1 integrin expression in a two step separation protocol (Collins et al., 2005; Collins et al., 2001).

**Figure 7.6** Selection markers and differentiation status of sub-population obtained from basal cells selection.
VII.2.6 Separation of α₂β₁ integrin high/low (α₂β₁^{hi/lo}) prostate cancer cells from primary cultures

Cultures were allowed to reach approximately 80% confluence before selection was performed. Prior to selection, BioCoat™ Collagen culture dishes were blocked with 2 mL BSA blocking buffer for 1 h at 37 °C. Cells were washed twice in 10 mL PBS per 75 cm² culture surface and trypsinised using 2 mL of 1x trypsin/EDTA for 5 min. 3 mL of R10 culture medium were added to the plate and cells were collected in a centrifuge tube. The trypsinisation was repeated by adding 1 mL of 10x trypsin/EDTA for 10 min. The cells were combined in the same centrifuge tube after addition of 3 mL of R10 culture medium to stop the trypsinisation step. Cells were centrifuged at 1500 rpm for 3 min and resuspended in 3 mL of complete stem cell medium. The cells were counted (refer to section IV.3.1.3) and plated out in the BSA-blocked collagen dishes in triplicate. Cells in the BSA-blocked collagen dishes were incubated for 20 min at 37 °C in 5% CO₂ atmosphere to allow the α₂β₁^{hi} cell sub-population to adhere. The medium containing α₂β₁^{lo} cells was collected and centrifuged at 1500 rpm for 10 min. The cell pellets were resuspended in 3 mL of complete stem cell medium and plated out in triplicate.

VII.2.7 Isolation of CD133⁺/α₂β₁^{hi} prostate stem cells from primary samples

Cells adhered to BSA-blocked collagen dishes were collected by adding 1 mL of 1x trypsin/EDTA for 5 min. 3 mL of R10 culture medium were added to the plates and cells were collected in a centrifuge tube. The trypsinisation was repeated by adding 1 mL of 10x trypsin/EDTA for 10 min. The cell suspension was passed through a cell strainer (40 µm pore size, BD Biosciences) to
remove cell clumps and cells were counted before being collected in a centrifuge tube and sedimented at 1500 rpm for 10 min.

After centrifugation, CD133 expressing prostate SCs were isolated from transit amplifying cells with a direct CD133 Cell Isolation Kit (Miltenyi Biotec). Up to $10^8$ cells were suspended in 300 µL (direct CD133 cell isolation kit) of MACS buffer and magnetic cell labelling solution. Cell separation on MACS MS columns were performed according to the manufacturer’s instructions. Following elution from the first column, cells were passed over a second column to increase purity of the selected population. One separation column was used for approximately $6 \times 10^6$ cells for the first round of selection and up to 4 mL eluate for the second round of selection. CD133$^+$ cell fraction was eluted in a final volume of 0.5 mL MACS buffer. The eluate was transferred to a 15 mL centrifuge tube and combined with 14 mL of PBS. After inverting the tube several times, CD133-expressing cells were sedimented and the supernatant was carefully aspirated, leaving approximately 100 µL of liquid above the cells. The cells were resuspended in culture medium and 10 µL of this suspension were used to determine live cell number (for detailed information, see section VI.3.1.3). The remaining cell suspension was combined with an appropriate amount of complete SC media and then plated.
VII.2.8 Methylation analysis

For methylation analysis protocols, see materials and methods section VI.3.2. Isolation of genomic DNA from prostate cell lines, committed basal and transit amplifying cells was performed as elucidated in section VI.3.2.1. DNA extraction from SC-selected population using QIAamp Micro Kit (QIAGEN) was carried out as described in the manufacturer's protocol.

VII.2.9 Gene expression analysis using qRT-PCR prior and after epigenetic treatment

For the qRT-PCR materials and methods refers to the sections VI.3.3.4-8. Isolation of total RNA from prostate cell lines, committed basal and transit amplifying cells was carried out as explained in detail in section VI.3.3.1. For total RNA extraction from SC-selected population total DNA was isolated from cultured cells with RNeasy Micro Kit (QIAGEN) following the manufacturer's protocol. Complementary DNA synthesis was performed as described in section VI.3.3.2. In this analysis, RPLP0 was used as the housekeeping gene as it was already optimised by Prof Norman Maitland’s group (University of York).
VII.3 Results

VII.3.1 Expression profiling of ALDH genes in prostate cell lines

Gene expression levels were analysed in three selected ALDH isoforms (ALDH1A1, 2 and 3A1) in 7 prostate cell lines. This panel comprised of 3 benign epithelial prostatic cell lines and 4 PCa cell lines. As reported in Figure 7.7, qRT-PCR (see section VI.3.3.4 for assay design) revealed no clear distinction between malignant and epithelial cell lines in the ALDH expression pattern. ALDH1A1 was expressed at variable levels in the PCa cell lines but also the benign PNT1A cell line was shown to express ALDH1A1 (Figure 7.6A). ALDH2 showed very high level of expression on PNT1A (up to 1000 folds) compared with the general trend of ALDH2 expression in both benign and cancerous cell lines.

A)
Figure 7.7 Relative ALDH gene expressions in PCa cell lines. (A) ALDH1A1, (B) ALDH2, (C) ALDH3A1 fold change plots. Each set of data is normalised to the smallest expression value. Data are represented as average of three experiments ±SD. RPLP0 was used as control gene (optimised by Prof. Maitland’s group, University of York). Note: difference scale of the y axis.

The same analysis was performed on ALDH3A1 gene expression and PNT1A together with PC3 revealed very high endogenous levels of this gene (above
500 folds) compared to the set under analysis, suggesting once again no correlation between cell-type and expression pattern.

**VII.3.2 Effect of DAC on ALDH promoter methylation pattern and gene expression in prostate cell lines**

The potential epigenetic regulation of the ALDH target genes (ALDH1A1, 2 and 3A1) in PCa was investigated by performing DNA methylation analysis of gene promoters, using already optimised assays in the previous chapter (section VI.3.2.7). Unfortunately, pyrosequencing analysis could not be repeated with ALDH2 due to lack of reproducibility of the pre-designed assay in the new set of cell samples.

As reported in Figure 7.8A, ALDH1A1 promoter was methylated in PNT2C2 control (65.7%) and DAC treatment induced reduction of the methylation level (44.35%). qRT-PCR analysis revealed that DAC exposure led to a 5-fold increase in ALDH1A1 gene expression when compared with the DMSO control (see Figure 7.8B). Similarly, ALDH1A1 was shown to be heavily methylated in benign PNT2C2 and malignant Du145 prostate cells. DAC exposure to these cell lines resulted in 44.35% and 51.40% demethylation of ALDH1A1 respectively.
Figure 7.8 ALDH1A1 promoter methylation and gene expression analysis in prostate cell lines before and after DAC treatment. (A) ALDH1A1 promoter methylation analysis was performed at three CpG sites with average % methylation reported as a red line across the bars. The same experiment was repeated in duplicate. For assay design, see section VI.3.2.7. (B) Relative ALDH1A1 gene expressions. DMSO control samples are set at 1 for all the graphs. Data are represented as average of three experiments ±SD. RPLP0 was used as a control gene (optimised by Prof. Maitland's group, University of York). DAC (1 µM) treatment was repeated every other day for 96 h before RNA extraction.
Only two of the three CpG sites under analysis seemed to have consistent high methylation levels in both cell lines and therefore were more affected by the DAC exposure. No significant ALDH1A1 methylation level was detected in the other cell lines under analysis, hence no difference was observed after epigenetic treatment in gaining demethylation. However, PNT1A and P4E6 cell lines, both expressing ALDH1A1 gene (see section VII.3.1 for relative gene expression data), showed an upregulation after DAC exposure, but not directly related to demethylation of the ALDH1A1 promoter.

The same investigation was performed in the ALDH3A1 promoter and the overall methylation status was above 25% for the entire sample set under analysis with the exception of the P4E6 cell line (see Figure 7.8A). The benign RC165 cell line had a very heavily methylated promoter region (88.18%) with reduction of the methylation level of almost 30% after DAC exposure (61.59%). However, as evident from Figure 7.9B, this demethylating effect did not result in gene reactivation. In terms of DAC-dependent demethylation, a common trend was followed in all the prostate cell lines: small fluctuations in the methylation status after treatment did not led to significant re-expression of the ALDH3A1 gene as revealed via qRT-PCR analysis (see Figure 7.9B). Moreover, the mRNA levels increased after DAC treatment on Du145 and LNCaP cell lines, but independently from demethylation of the ALDH3A1 gene promoter.
Figure 7.9 ALDH3A1 promoter methylation and gene expression analysis in prostate cell lines before and after DAC treatment. (A) ALDH3A1 promoter methylation analysis was performed in eight CpG sites. The same experiment was repeated in duplicate. For assay design, see section VI.3.2.7. (B) Relative ALDH3A1 gene expressions. DMSO control samples are set at 1 for all the graphs. Data are represented as average of three experiments ±SD. RPLP0 was used as a control gene (optimised by Prof. Maitland’s group, University of York). DAC (1 µM) treatment was repeated every other day for 96 h before RNA extraction.
Finally, a direct comparison of the Ct values of ALDH1A1 versus the ALDH3A1 gene revealed that ALDH1A1 mRNA levels were reduced compared with ALDH3A1 endogenous levels, suggesting a quite low abundance of the ALDH1A1 gene in the prostatic cell lines examined in this study.

<table>
<thead>
<tr>
<th>Sample</th>
<th>ALDH1A1 Ct</th>
<th>ALDH2 Ct</th>
<th>ALDH3A1 Ct</th>
<th>RPLP0 Ct</th>
</tr>
</thead>
<tbody>
<tr>
<td>PNT1A DMSO</td>
<td>37.92</td>
<td>28.39</td>
<td>29.91</td>
<td>19.18</td>
</tr>
<tr>
<td>PNT1A DAC</td>
<td>34.01</td>
<td>28.82</td>
<td>30.58</td>
<td>19.09</td>
</tr>
<tr>
<td>PNT2C2 DMSO</td>
<td>39.50</td>
<td>34.34</td>
<td>34.74</td>
<td>19.75</td>
</tr>
<tr>
<td>PNT2C2 DAC</td>
<td>36.01</td>
<td>31.41</td>
<td>33.33</td>
<td>18.70</td>
</tr>
<tr>
<td>RC165 DMSO</td>
<td>N.D</td>
<td>33.61</td>
<td>34.23</td>
<td>21.36</td>
</tr>
<tr>
<td>RC165 DAC</td>
<td>N.D</td>
<td>33.18</td>
<td>34.57</td>
<td>21.36</td>
</tr>
<tr>
<td>P4E6 DMSO</td>
<td>36.98</td>
<td>35.86</td>
<td>31.64</td>
<td>22.30</td>
</tr>
<tr>
<td>P4E6 DAC</td>
<td>30.72</td>
<td>29.93</td>
<td>27.79</td>
<td>20.06</td>
</tr>
<tr>
<td>PC3 DMSO</td>
<td>39.51</td>
<td>32.91</td>
<td>28.77</td>
<td>19.77</td>
</tr>
<tr>
<td>PC3 DAC</td>
<td>37.88</td>
<td>29.86</td>
<td>28.23</td>
<td>19.57</td>
</tr>
<tr>
<td>Du145 DMSO</td>
<td>37.39</td>
<td>32.08</td>
<td>34.35</td>
<td>20.98</td>
</tr>
<tr>
<td>Du145 DAC</td>
<td>35.15</td>
<td>30.03</td>
<td>30.62</td>
<td>20.64</td>
</tr>
<tr>
<td>LNCaP DMSO</td>
<td>37.97</td>
<td>30.41</td>
<td>37.78</td>
<td>20.03</td>
</tr>
<tr>
<td>LNCaP DAC</td>
<td>37.73</td>
<td>31.7</td>
<td>35.36</td>
<td>21.34</td>
</tr>
</tbody>
</table>

**Table 7.2** Ct values of ALDH gene expressions before and after DAC treatment in prostate cell lines. RPLP0 was used as control gene (optimised by Prof. Norman Maitland’s group, University of York). DAC (1 µM) treatment was repeated every other day for 96 h before RNA extraction.
VII.3.3 Expression profiling of ALDH genes in primary epithelial cultures from prostatic tissues

Expression of the selected ALDH isoforms (ALDH1A1, 2 and 3A1) and methylation of their promoter (see Supplementary Information VI.A) were measured in primary epithelial cultures derived from clinical samples of BPH and PCa. The three isoforms were expressed in both BPH and PCa samples and no significant differences in mRNA level were seen between benign and malignant samples (see Figure 7.10). Furthermore, primary cultures were shown to express ALDH isoforms in a comparable manner with established human PCa cell lines.

A)
Figure 7.10 qRT-PCR analysis of ALDHs gene expression in prostate primary epithelial cultures. (A) ALDH1A1, (B) ALDH2, and (C) ALDH3A1 gene expression profiles. RPLP0 was used as normalising gene (optimised by Prof. Norman Maitland’s group, University of York). Cell cultures were generated from specimens of 4 human PCa and 4 non-malignant BPH control tissues.
Next, investigation of the methylation status of ALDH gene promoters was performed via pyrosequencing (see Figure 7.11).

A) ALDH1A1 in primary samples

<table>
<thead>
<tr>
<th></th>
<th>BPH 01</th>
<th>BPH 02</th>
<th>BPH 03</th>
<th>BPH 04</th>
<th>PCa 01</th>
<th>PCa 02</th>
<th>PCa 03</th>
<th>PCa 04</th>
</tr>
</thead>
<tbody>
<tr>
<td>0% Me</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100% Me</td>
<td>5.40%</td>
<td>9.17%</td>
<td>4.85%</td>
<td>8.93%</td>
<td>9.35%</td>
<td>7.83%</td>
<td>10.87%</td>
<td>11.46%</td>
</tr>
</tbody>
</table>

% DNA Methylation

B) ALDH3A1 in primary samples

<table>
<thead>
<tr>
<th></th>
<th>BPH 01</th>
<th>BPH 02</th>
<th>BPH 03</th>
<th>BPH 04</th>
<th>PCa 01</th>
<th>PCa 02</th>
<th>PCa 03</th>
<th>PCa 04</th>
</tr>
</thead>
<tbody>
<tr>
<td>0% Me</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100% Me</td>
<td>5.56%</td>
<td>6.28%</td>
<td>5.26%</td>
<td>4.93%</td>
<td>6.37%</td>
<td>6.55%</td>
<td>6.93%</td>
<td>9.18%</td>
</tr>
</tbody>
</table>

% DNA Methylation

Figure 7.11 ALDH1A1 and ALDH3A1 promoter methylation analysis in bioptic sample from benign hyperplasia (BPH) and prostate cancer (PCa). Cell cultures were generated from specimens of human prostate cancers (n = 4, red bars) and non-malignant BPH control tissues (n = 4, green bars). Analysis was performed at three CpG sites for (A) ALDH1A1 and eight CpG sites for (B) ALDH3A1. The average of % methylation is reported in red above the correspondent sample. For assay design, see section VI.3.2.7.

The DNA methylation levels were shown to be low, with the average methylation less than 12%, in all the samples analysed. No significant
distinction was observed between BPH and PCa samples. In fact, none of the samples contained significantly hypermethylated DNA compared with control samples. Due to the lack of hypermethylation of ALDH gene promoters, it was not possible to epigenetically modulate ALDH gene levels in this small cohort of prostate primary epithelial cells.

**VII.3.4 Expression profiling of ALDH genes in the prostate epithelial hierarchy**

ALDH expression was next analysed by qRT-PCR in stem (SC, CD133+/α2β1hi), transit amplifying (TA, CD133+/α2β1hi) and committed basal (CB, CD133+/α2β1lo) cell populations isolated from low passage (<10) primary prostate epithelial cultures. In ALDH1A1, the expression was consistent in unselected cultures (see section VII.3.4). CB cells had the highest mRNA level compared with TA and SC and this trend was evident in both BPH and PCa cultures (see Figure 7.12). The Ct values in Figure 7.9 revealed a consistent trend, where CB had the lowest Ct score compared with the other sub-populations in the all samples under investigation.
The same analysis, performed for ALDH2 gene expression (see Figure 7.13), resulted in an inverted regulation pattern with high expression level in SCs population compared with the CB cells. This result was confirmed by Ct values as shown in Figure 7.11, which suggested an inverted, but still consistent, trend compared to ALDH1A1 Ct values. Once again no detectable difference was observed between BPH and PCa.
Finally, the same qRT-PCR expression analysis for the ALDH3A1 gene was investigated and similar expression trend resulted in ALDH3A1 and ALDH1A1 genes.
Figure 7.14 ALDH3A1 gene expression analysis in basal epithelial cell after subpopulation selection. SC: stem cells (CD133+/α2β1hi); TA: transit amplifying cells (CD133+/α2β1hi); CB: committed basal cells (CD133+/α2β1lo).

The very low expression of ALDH3A1 at the mRNA level (see Figure 7.14) seemed to be a common feature in the SC population, whereas higher levels were measured in TA and CB respectively.

These results provided evidence that the investigated ALDH isoforms could be regulated at the transcriptional level during differentiation of prostate epithelia from clinical tissues. Accordingly, methylation analysis of the selected sub-
populations was needed to demonstrate whether this differential regulation was sustained by DNA methylation. Pyrosequencing analysis (section VI.3.2.7) was carried out on selected SC, TA and CB cells. However, very low levels of methylation (<12%) were found in all the populations without remarkable differences between BPH and PCa samples.

VII.4 Discussion

Prostate cancer (PCa) has been classified as a molecularly and phenotypically heterogeneous disease and its “chaos of phenotypes” has challenging clinical implications (Roudier et al., 2003), for example in regard to distinguishing and selecting PCa cells with tumour- and metastasis-initiating ability (van den Hoogen et al., 2010). As a consequence, the hypothetical model of hierarchical organization of PCa cells (see Figure 7.2) is crucial in helping to explain how the tremendous heterogeneity associated with PCa can be generated.

Recent studies sustained that identification of cancer SC-like sub-populations in PCa cell lines, stained positive for ALDH, are capable of self-renewal and re-establishment of parental cells (Hellsten et al., 2011). Furthermore, a high expression of ALDH in PCa SCs has been shown to be positively correlated with Gleason score (see Supplementary Information VII.A for Gleason sum definition), pathologic stage, and inversely correlated with overall survival in PCa patients (Li et al., 2009). Moreover, high ALDH activity has successfully been used to identify tumour initiating PCa cells and components of metastases (van den Hoogen et al., 2010).
The data presented in this study demonstrated that ALDH1A1, ALDH2 and ALDH3A1 are expressed in prostate cell lines but with no apparent distinction between benign or malignant cellular phenotype. Interestingly, findings in this study indicated that ALDH1A1 and ALDH3A1 expression in part is controlled via DNA methylation in prostate cell lines, where an inverse correlation between expression and DNA methylation was observed, together with upregulation of mRNA levels after gene demethylation by DAC treatment.

Previous studies have strongly suggested that the basal cell layer of the adult human prostate contain highly undifferentiated progenitor cells (Lang et al., 2009). It has been demonstrated that a high percentage of proliferating cells in the basal layer microenvironment may be important in maintaining fundamental properties of SCs strictly localized to this compartment (Tang et al., 2007). Therefore, it was desirable in the present study to investigate the ALDH expression level and gene promoter methylation status within primary epithelial cells from human prostatic biopsies. The data demonstrated that primary prostate cultures expressed similar mRNA levels of ALDH1A1 and ALDH3A1, whereas low expression levels were observed for ALDH2. Preliminary studies carried out by Prof. Maitland’s group suggested that in basal epithelial cultures ALDH2 is expressed at a higher level than ALDH1A1 and 3A1 and in particular that PCa samples had higher expression than BPH cultures (see Figure 7.3). Furthermore, DNA methylation profiling data obtained from the primary epithelial cultures were in contrast with data obtained from the prostate cell lines. The very low levels of promoter methylation found in these samples indicated that the lack of expression of ALDH1A1 and ALDH3A1 was
independent of promoter methylation, indicating that other mechanisms are important in controlling their expression in prostate tissues and primary cultures. For example, several ALDH isoforms (ALDH1A1, ALDH1A2, ALDH1A3 and ALDH8A1) have been shown to function in retinoic acid cell signalling via oxidation of all-trans-retinal and 9-cis-retinal, which has been linked to the “stemness” characteristics of CSCs (Gudas and Wagner, 2011).

As reported in section VII.1.1, prostate epithelium is highly heterogeneous and formed by a variety of cell populations at different differentiation stages (Collins et al., 2001). When SC (CD133+/α2β1hi), TA (CD133'/α2β1hi) and CB (CD133'/α2β1lo) populations from primary epithelial cultures were analysed separately, gene expression analysis revealed that ALDH1A1 and ALDH3A1 had higher mRNA levels in CB compared to SC sub-populations, whereas an inverted correlation was observed for the ALDH2 expression pattern. This preliminary investigation provided evidence that ALDH2 may be regulated at the transcriptional level during differentiation of prostate epithelia from basal tissues, which is in agreement with Prof. Maitland's unpublished data (see Figure 7.3).

A recent study identified and isolated ALDH1A1+ expressing cells from two human PCa cell lines (PC3 and LNCaP) on the basis of the Aldefluor assay. The study showed such cells to be highly clonogenic and with tumour forming potential in animal models (Li et al., 2009). The results in this chapter suggest that ALDH1A1 expression is not linked with the SC-like phenotype in selected cells from the primary prostate epithelium. Data provided by Dr. Phil Burns (see Figure 7.3) also suggest that ALDH1A1 is not the key enzyme in primary
ovarian samples although this contradicts findings from a recent study (Deng et al., 2010; Silva et al., 2010). Some doubts of using ALDH1A1 as a SC marker has been described in the literature (Ginestier et al., 2007; Moreb, 2008), which further emphasises the need to understand the expression and regulation of this class of enzymes in cancer.

These results suggest differential ALDH expression pattern between cell lines and primary tissues. Furthermore, the results obtained showed no disparate methylation of the ALDH1A1 and ALDH3A1 promoter in individual populations. Regulation of ALDH1A1 and 3A1 gene expression appears to be independent of DNA methylation and may involve retinoid-dependent pathways as described above.

In conclusion, ALDH1A1, ALDH2 and ALDH3A1 expression appears to be differentially expressed in prostate cell lines relative to the primary cultures. This expression is independent of gene promoter methylation, but it remains to be seen whether epigenetic marks on the histones have implications for ALDH regulation and expression. A larger number of clinical samples must also be analysed in order to fully provide conclusions to the possibility of epigenetic modulation of ALDHs not limited to just ALDH1A1, 2 and 3A1. Indeed, the data from Prof. Maitland’s group (see Figure 7.3) suggest that ALDH3B2 and ALDH16A1 isoform may be statistically important, but also showed that ALDH2, 3A2, 4A1, 9A1, 16A1 and 18A1 could be important players in PCa. In addition to these preliminary studies, the recent published studies from van den Hoogen (van den Hoogen et al., 2010; van den Hoogen et al., 2011) suggest that ALDH7A1 is a key enzyme in tumourigenesis. The number of ALDH isoforms
that may be important as highlighted in this chapter suggests a complexity in their regulation and expression. A better understanding of the importance of ALDH in PCa can only be achieved by examining larger number of clinical samples and by carefully assessing their expression within the various cell layers of the prostate. The complexity of ALDH expression also complicates the potential of using a specific ALDH isoform as a biomarker for PCa and indeed the development of a specific molecular fluorescent chemical probe.
Supplementary Information VII.A
Prostatic carcinoma primary samples list

<table>
<thead>
<tr>
<th>Identification code</th>
<th>Sample name</th>
<th>Origin of tissue</th>
</tr>
</thead>
<tbody>
<tr>
<td>BPH 01</td>
<td>125/11</td>
<td>Benign Prostatic Hyperplasia</td>
</tr>
<tr>
<td>PCa 01</td>
<td>120/11 Ra**</td>
<td>Prostate cancer (Gleason 6*)</td>
</tr>
<tr>
<td>BPH 02</td>
<td>68/11</td>
<td>Benign Prostatic Hyperplasia</td>
</tr>
<tr>
<td>PCa 02</td>
<td>116/11 Ra**</td>
<td>Prostate cancer (Gleason 7*)</td>
</tr>
<tr>
<td>PCa 03</td>
<td>116/11 Rb†</td>
<td>Prostate cancer (Gleason 7*)</td>
</tr>
<tr>
<td>PCa 04</td>
<td>50/11 Ra**</td>
<td>Prostate cancer (Gleason 6*)</td>
</tr>
<tr>
<td>BPH 03</td>
<td>60/11</td>
<td>Benign Prostatic Hyperplasia</td>
</tr>
<tr>
<td>BPH 04</td>
<td>71/11</td>
<td>Benign Prostatic Hyperplasia</td>
</tr>
</tbody>
</table>

**Ra**: Right apex
†Rb: Right base

*The Gleason Grading System is used to ensure that physicians are aware of the prostate cancer’s stage and to help evaluate the tumour prognosis (Gleason, 1977). Cancers with a higher Gleason score are more aggressive and have a worse prognosis. The Gleason score may be between 2 to 10. Gleason grades are associated with the following features:

- Scores from 2 to 4 are very low on the cancer aggression scale.
- Scores from 5 to 6 are mildly aggressive.
- A score of 7 indicates that the cancer is moderately aggressive.
- Scores from 8 to 10 indicate that the cancer is highly aggressive.
Supplementary Information VII.B

Relative quantification of ALDH gene expression in selected primary samples: (A) ALDH1A1, (B) ALDH2, (C) ALDH3A1. Histograms represent fold change of the respective ALDH target gene normalised to RPLP0 housekeeping gene. A full list and details of the primary samples under analysis are reported in Supplementary Information VII.A.

A)

![Relative ALDH1A1 expression in selected primary samples](image)

B)

![Relative ALDH2 expression in selected primary samples](image)
C)

Relative ALDH3A1 expression in selected primary samples
Chapter VIII

Final Conclusions and Future Directions
New strategies for investigating SCs are required to fully understand the complex biology behind this cellular phenotype. Indeed, despite many years of intensive research, there is very little consensus about suitable markers to locate and isolate SCs from adult tissues. The focus in the recognition of CSCs by featuring high expression ALDHs has been a hot topic over the past decade (Moreb, 2008; Moreb et al., 2008). Currently, cells expressing high ALDH activity are identified by employing the Aldefluor assay with subsequent sorting of cells by FACS. This strategy has been successful in the isolation of SCs from many human tumours (Ginestier et al., 2007), notably breast cancer (Marcato et al., 2011), prostate (van den Hoogen et al., 2010; van den Hoogen et al., 2011), ovarian (Sun and Wang, 2011) and non-small cell lung cancer (Jiang et al., 2009). However, new findings regarding the heterogeneity of tissue-specific ALDH isoforms underline complexity in identifying a unique SC population by measuring ALDH activity. Because of this limited knowledge, the approach of designing fluorescent small molecules selective for a specific ALDH isoform is a difficult challenge.

The aim of the thesis was to develop a novel library of LC-AAQs focussed on the modification of the chemical functionalities in the β-position to the terminal acetal group. This LC-AAQ library, obtained in a three-five steps synthesis, was combined with a second library of HA-AAQ derivatives, featuring a β-positioned secondary or tertiary amine in proximity to the terminal acetal moiety. Therefore, 12 compounds were synthesised (LC-AAQ library), another 4 compounds provided for this study for biological analysis (HA-AAQ library) and a total of 32 compounds (16 AAQs and 16 hydrolysed counterparts, HAAQs) were
biologically investigated. AAQs were subjected to acidic hydrolysis (with 2M HCl) in order to generate active aldehyde moieties capable of targeting ALDHs. The results obtained in Chapter III demonstrated that LC-AAQ derivatives were prone to acidic hydrolysis, generating the target aldehydes (albeit not in 100% yield), whereas modifications in HA-AAQ library of compounds revealed intrinsic resistance to hydrolysis even at high temperature (70°C). These results show unexpected differences in capacity of acetal moieties to undergo acidic hydrolysis, which at present is unexplainable.

Metabolism studies, as discussed in Chapter III, were conducted using several enzyme sources to evaluate the potential ALDH-dependent conversion of the aldehyde-containing HAAQs into carboxylic acid products. Incubation with HAAQs showed activity in rat liver cytosolic fraction (Shum and Blair, 1972), but not in yeast or human recombinant ALDH. The results from the latter provided inconclusive results and hence it is still possible that ALDH1A1 metabolism did occur. Probable explanations for lack of yeast ALDH activity could be that (i) it mainly contains ALDH2 isoform (Bostian and Betts, 1978a, b; Devaraj et al., 1997) which may not be able to accommodate AAQs in their active sites, or (ii) yeast ALDH only metabolise substrates in the high μM to mM range; as the anthraquinones are dark blue, such high concentrations are not compatible with measuring kinetic reactions due to absorbance interference. Furthermore, human recombinant ALDH1A1 was only preliminary investigated due to cost limitations. In order to investigate it in more detail, future works would include the purification of ALDH1A1 and ALDH3A1 from both human and rat liver and establish optimal conditions for setting up robust enzymatic assays (Lindahl and
Evces, 1984; Sidhu and Blair, 1975), which could be used to test the selectivity
of AAQs. Interestingly, parental AAQs appeared to be metabolised when
incubated with the cytosolic fractions from rat liver, which may suggest the
acetal moieties are prone to oxidative metabolism by other enzymes than
ALDHs present in the cytosol.

Assessment of the DNA affinity of the novel AAQ compounds revealed a weak
DNA interaction with CT-DNA and a general lack of cytotoxicity (Chapter IV),
which is encouraging in regard to the generation of a live cell tracker to be used
in a commercial setting. Furthermore, the chemosensitivity studies revealed that
AAQs before and after hydrolysis had greater impact on cell viability in ALDH-
deficient MCF7 cell line compared to the ALDH-positive A549 cell line,
suggesting that an interplay between ALDH level and cell survival may have
occurred. Assessment of AAQs and HAAQs were carried out in the MCF7 and
A549 cell lines derived from breast and lung cancer tissue respectively. As
such, meaningful data in regard to the involvement of ALDH is difficult.
Therefore, future work should ideally include the establishment of isogenic cell
lines expressing various ALDH isoforms so that more direct comparisons of the
AAQ and HAAQ chemical probes can be made.

A screening of the fluorescence properties of these novel AAQs was assessed
in two fixed and live cell models, U-2 OS and A549, to gain insight into their
potential ability of labelling cells with or without the involvement of membrane
uptake systems (Chapter V). The data generated revealed that the tested AAQs
were stable fluorophores with structure-dependent cellular localisation.
Although, the affinity for ALDH-metabolism is unclear, the presented AAQs
showed potential as non-toxic fluorescent probes, with selective affinity for endoplasmic reticulum, Golgi or nuclei respectively. Indeed, efficient cellular uptake was obtained in live cells with stable fluorescent signal overtime. Future directions in this part of the thesis should concern investigations of the fluorescent properties of AAQs in a larger number of cell lines using time-lapse and confocal microscopy. It would also be desirable to track the specific AAQ location after cell mitosis and to investigate the effect on cell viability in a microscopy-based time course experiment.

In light of these results, a second library of AAQ with different structural modifications could be synthesised and biologically tested. Specific objectives could involve (i) further explorations of chemical modification to the β-position located in proximity to the terminal acetal moiety and (ii) synthesis of other disubstituted anthraquinones including 1,5 or 1,8 or 2,6-functionalised probes to further explore fluorescent properties and cellular localisations.

Due to a relatively poor understanding of the ALDH expression and regulation, it was desirable to aid the current understanding of the genetic and epigenetic regulation of isoforms from the ALDH 1-3 family (ALDH1A1, 2 and 3A1). A “pilot” study was established with the aim of obtaining information in regard to the ALDH expression and regulation in four cell lines (MCF7, A549, HT29 and HeLa) derived from different cancer tissues. The data discussed in Chapter VI revealed that aberrant gene methylation was cell-type dependent for the three ALDH isoforms under investigation. Decitabine treatment led to enhanced protein expression for ALDH1A1 (A549), ALDH2 (MCF7) and ALDH3A1 (A549). In contrast, the protein level was reduced for ALDH1A1 in the HT29 colon
cancer cell line after decitabine treatment, suggesting the involvement of other mechanism than DNA methylation is responsible of gene silencing in this cell line. ALDH1A1, ALDH2 and ALDH3A1 were highly expressed in prostate cell lines, with expression linked to promoter methylation. In contrast, low levels of DNA methylation were found in primary prostate cancer cells and benign prostatic hyperplasia (Chapter VII). Interestingly, ALDH1A1, considered a SC marker, was found to be expressed at low levels in CD133$^+/\alpha_2\beta_1^{hi}$ SC fraction and upregulated in CD133$^+/\alpha_2\beta_1^{lo}$ differentiated prostate cancer cells.

Considering the limited number of analysed primary prostate samples, the present study can only be considered a preliminary screening. However, future directions should take into account the tremendous difference between immortalised cancer cell lines and primary cultures and probably should focus more on detailed investigations on the latter. Furthermore, a larger number of prostatic primary samples should be included in order to apply statistical validity and significance to the preliminary data. Separate studies could also be carried out in other cancer tissues, which would possibly elucidate differential ALDH isoform expression and therefore validation of the potential of ALDHs as biomarkers. In addition, a number of ALDH isoforms (ALDH1A1, ALDH1A2, ALDH1A3 and ALDH8A1) have been shown to function in retinoic acid cell signalling via oxidation of all-trans-retinal and 9-cis-retinal (Marcato et al., 2011; Sladek, 2003). This function in particular has been linked to the “stemness” characteristics of CSCs (Ginestier et al., 2007; Gudas and Wagner, 2011), hence retinoid expression and possible regulation of ALDH activity could also be investigated.
In summary, the results in the presented thesis demonstrate the complexity and tumour type specificity of ALDH expression. This creates challenges for the development of selective probes for CSC isolation, such as the AAQs discussed in this study. Although inconclusive results were obtained in regard to AAQs and their potential in targeting ALDHs, selected AAQs were shown to reveal interesting biological properties, suggesting them as potential non-invasive cytometric probes for tracking molecular interactions in live cells.
References


Hunt, M. (2010). REAL TIME PCR tutorial (University of South Carolina, School of Medicine).


Krapcho, A.P., Getahun, Z., and Avery, K.J. (1990). The synthesis of 1,4-difluoro-5,8-dihydroxyanthracene-9,10-dione and ipso substitutions of the fluorides by diamines leading to 1,4-bis-[(aminoalkyl)amino]-5,8-dihydroxyanthracene-9,10-diones. Synthetic Communications 20, 2139-2146.


Moreb, J.S., Ucar, D., and Han S, A.J., Goldstein A S, Ostmark B, Chang LJ. (2012). The enzymatic activity of human aldehyde dehydrogenases 1A2 and 2 (ALDH1A2 and ALDH2) is detected by Aldefluor, inhibited by diethylaminobenzaldehyde and has significant effects on cell proliferation and drug resistance. Chem Biol Interact 195, 52-60.


References


cyclin-dependent kinase/Aurora kinase inhibitor. Molecular Cancer Therapeutics 5, 2459-2467.


Spector, T. (1978). Refinement of Coomassie Blue method of protein quantification - simple and linear spectrophotometric assay for less than or equal to 0.5 to 50 mg of protein Analytical Biochemistry 86, 142-146.


Sun, S., and Wang, Z. (2011). ALDH(high) adenoid cystic carcinoma cells display cancer stem cell properties and are responsible for mediating metastasis. Biochemical and Biophysical Research Communications 396, 843-848.


Appendix
**Appendix I: List of suppliers**

<table>
<thead>
<tr>
<th>Supplier</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abnova</td>
<td>Peterborough, UK</td>
</tr>
<tr>
<td>AGTC Bioproducts</td>
<td>Hessle, UK</td>
</tr>
<tr>
<td>ATCC</td>
<td>Manassas, VA, USA</td>
</tr>
<tr>
<td>Beckmann Coulter Ltd</td>
<td>High Wycombe, UK</td>
</tr>
<tr>
<td>Becton-Dickinson Systems</td>
<td>San Jose, CA, USA</td>
</tr>
<tr>
<td>Bio-Rad</td>
<td>Hemel Hempstead, UK</td>
</tr>
<tr>
<td>Biostatus Ltd</td>
<td>Leicestershire, U.K.</td>
</tr>
<tr>
<td>Calbiochem</td>
<td>Nottingham, UK</td>
</tr>
<tr>
<td>Carl Zeiss Inc.</td>
<td>Welwyn Garden City, UK</td>
</tr>
<tr>
<td>Chroma Technology Corp</td>
<td>Bellows Falls, VT, USA</td>
</tr>
<tr>
<td>Coherent Inc.</td>
<td>Santa Clara, CA, USA</td>
</tr>
<tr>
<td>COHU</td>
<td>San Diego, CA, USA</td>
</tr>
<tr>
<td>Corning</td>
<td>Loughborough, UK</td>
</tr>
<tr>
<td>First Link</td>
<td>Birmingham, UK</td>
</tr>
<tr>
<td>Fisher Scientific</td>
<td>Loughborough, UK</td>
</tr>
<tr>
<td>GE Healthcare</td>
<td>Little Chalfont, UK</td>
</tr>
<tr>
<td>Hamamatsu Photonics</td>
<td>Welwyn Garden City, UK</td>
</tr>
<tr>
<td>IKA</td>
<td>Staufen, Germany</td>
</tr>
<tr>
<td>Invitrogen</td>
<td>Paisley, UK</td>
</tr>
<tr>
<td>Kinetic Imaging Ltd</td>
<td>Wirral, UK</td>
</tr>
<tr>
<td>LGC Prochem</td>
<td>Teddington, UK</td>
</tr>
<tr>
<td>Lonza Biologics</td>
<td>Cambridge, UK</td>
</tr>
<tr>
<td>Martin Christ</td>
<td>Osterode, Germany</td>
</tr>
<tr>
<td>Merck</td>
<td>Dorset, UK</td>
</tr>
<tr>
<td>Millipore</td>
<td>Watford, UK</td>
</tr>
<tr>
<td>Miltenyi Biotec</td>
<td>Gladbach, Germany</td>
</tr>
<tr>
<td>Company</td>
<td>Location</td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>Molecular Devices</td>
<td>Downingtown, PA, USA</td>
</tr>
<tr>
<td>Nikon</td>
<td>Amstelveen, The Netherlands</td>
</tr>
<tr>
<td>PAA</td>
<td>Sommerset, UK</td>
</tr>
<tr>
<td>Perkin Elmer</td>
<td>Cambridge, UK</td>
</tr>
<tr>
<td>Prior Scientific</td>
<td>Cambridge, UK</td>
</tr>
<tr>
<td>Promega</td>
<td>Southampton, UK</td>
</tr>
<tr>
<td>QIAGEN</td>
<td>Valencia, CA, USA</td>
</tr>
<tr>
<td>R&amp;D Systems</td>
<td>Abingdon, UK</td>
</tr>
<tr>
<td>Roche Applied Biosystems</td>
<td>Burgess Hill, UK</td>
</tr>
<tr>
<td>Sigma</td>
<td>Dorset, UK</td>
</tr>
<tr>
<td>Solent Scientific</td>
<td>Segensworth, UK</td>
</tr>
<tr>
<td>Southern Biotech</td>
<td>Birmingham, AL, USA</td>
</tr>
<tr>
<td>StemCell Technologies</td>
<td>Grenoble, France</td>
</tr>
<tr>
<td>TCI</td>
<td>Portland, OR, USA</td>
</tr>
<tr>
<td>Thermo Fisher Scientific</td>
<td>Loughborough, UK</td>
</tr>
<tr>
<td>Tree Star</td>
<td>San Carlos, CA, USA</td>
</tr>
<tr>
<td>Varian Agilent Tech.</td>
<td>Stockport, UK</td>
</tr>
<tr>
<td>VWR international Ltd</td>
<td>Lutterworth, UK</td>
</tr>
<tr>
<td>Waters Ltd.</td>
<td>Hertfordshire, UK</td>
</tr>
</tbody>
</table>
Appendix II: Composition of buffers, cell media and stock solutions

**DNA-binding assay buffers**

**CT-DNA melting buffer**

CT-DNA melting buffer solutions used for the UV-melting experiments were prepared by making up a 1 L solution A of (10 mM NaH$_2$PO$_4$ and 1 mM Na$_2$EDTA) and a separate 1 L solution B (10 mM Na$_2$HPO$_4$ and 1 mM Na$_2$EDTA). Solution A was added to solution B, until pH 7.0 was achieved.

**Cell culture media**

Ham’s F-12 cell culture media was purchased from Lonza Biologics and Eagle’s Minimal essential media was obtained from LGC Prochem. Foetal Calf serum was purchased from PAA. Hank’s balanced salt solution was obtained from Sigma. All other reagents were purchased from Invitrogen unless otherwise stated.

**Hank’s balanced salt solution**

5.4 mM KCl, 0.3 mM Na$_2$HPO$_4$, 0.4 mM KH$_2$PO$_4$, 4.2 mM NaHCO$_3$, 1.3 mM CaCl$_2$, 0.5 mM MgCl$_2$, 0.6 mM MgSO$_4$, 137 mM NaCl, 5.6 mM D-glucose.

**Standard freezing media**

90% (v/v) complete mammalian cell line culture media, 10% (v/v) DMSO.

**D10 culture media**

Dulbecco’s modified Eagle’s media (DMEM) (Sigma), high glucose, supplemented with 10% (v/v) foetal calf serum and 2 mM L-glutamine.

**H7 culture media**

Ham’s F-12 media supplemented with 7% (v/v) foetal calf serum and 2 mM L-glutamine.
K2 culture media

KSFM (see below) supplemented with 2% (v/v) foetal calf serum.

M10 culture media

McCoy’s 5A Modified Media (Sigma) supplemented with 10% (v/v) foetal calf serum and 2 mM L-glutamine.

R10 culture media

RPMI 1640 media (Sigma) supplemented with 10% (v/v) foetal calf serum and 2 mM L-glutamine.

Keratinocyte serum-free media (KSFM)

KSFM supplemented with bovine pituitary extract (BPE, 50 µg/mL), human recombinant epidermal growth factor (hEGF, 5 ng/mL) and 2 mM L-Glutamine (Sigma).

Complete stem cell media

KSFM, supplemented with leukaemia inhibitory factor (LIF, 2 ng/mL, Millipore), stem cell factor (SCF, 2 ng/mL, First Link), granulocyte macrophage colony-stimulating factor (GM-CSF, 1 ng/mL, First Link) and 100 ng/mL cholera toxin (Sigma).

BSA blocking buffer

0.75 g of BSA (Sigma) were dissolved in 250 mL of PBS and heated for protein denaturation for 5 minutes at 80°C then proceeded to filter sterilisation.

MACS buffer

PBS, pH 7.2, supplemented with 0.5% foetal calf serum and 2 mM EDTA.
Solutions for molecular biology

All reagents were purchased from Sigma unless otherwise indicated.

Western Blotting

1x TAE buffer

40 mM Tris-acetate, 1 mM EDTA, pH 8.0.

RIPA Lysis buffer

150 mM NaCl, 1.0% Triton X-100, 0.5% sodium deoxycholate, 0.1% SDS (GE Healthcare), 50 mM Tris, Aprotinin 8µg/mL, Leupeptin 10µg/mL, pH 8.0.

5x sample loading buffer

250mM Tris-HCl pH 6.8, 1% SDS (GE Healthcare), 30% glycerol (GE Healthcare), 0.02% bromophenol blue, 5% mercaptoethanol, made up to 10 mL with deionised water.

Lower gel buffer

1.2 M Tris in deionised pH adjusted to 8.8.

Upper gel buffer

545 mM Tris in deionised water, pH adjusted to 6.8.

10% Resolving gel:

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower gel buffer</td>
<td>1.87 mL</td>
</tr>
<tr>
<td>10% SDS (GE Healthcare)</td>
<td>0.1 mL</td>
</tr>
<tr>
<td>Protogel (AGTC Bioproducts)</td>
<td>4.7 mL</td>
</tr>
<tr>
<td>10% ammonium persulfate</td>
<td>0.04 mL</td>
</tr>
<tr>
<td>TEMED</td>
<td>0.012 mL</td>
</tr>
<tr>
<td>deionised water</td>
<td>39.7 mL</td>
</tr>
</tbody>
</table>
### 4% Stacking gel:

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper gel buffer</td>
<td>3.15 mL</td>
</tr>
<tr>
<td>10% SDS (GE Healthcare)</td>
<td>0.5 mL</td>
</tr>
<tr>
<td>Protogel (AGTC Bioproducts)</td>
<td>6.65 mL</td>
</tr>
<tr>
<td>deionised water</td>
<td>39.7 mL</td>
</tr>
</tbody>
</table>

Add 0.04 mL 10% ammonium persulfate, 0.012 mL TEMED to 10 mL of 4% stacking gel before use.

### 5x TBS-Tween20 running buffer

150 mM NaCl, 50 mM Tris-HCl pH 7.5, 0.01% Tween20 made up to 1 L with deionised water.

### Transfer buffer

0.024 M Tris, 0.113 M glycine, 20% methanol in 1 L of deionised water.

### 4% Blocking solution

4% Blocking Reagent (GE Healthcare) in 1x TBS-Tween20.

### 2% Blocking solution

2% Blocking Reagent (GE Healthcare) in 1x TBS-Tween20.
Appendix V: Abstracts presented to attended conferences

AACR Annual Meeting, April 2011, Orlando, FL

Instability of ABCG2-associated side population expression in human A549 lung cancer cells and its contribution to drug resistance

Laura Cosentino¹ Marie Wiltshire², Sally Chappell³, Phil Burns⁴, Laurence H Patterson¹, Klaus Pors¹, Rachel J Errington³ and Paul J Smith⁵

¹Institute of Cancer Therapeutics, University of Bradford, West Yorkshire BD7 1DP, UK, ²Department of Pathology and ³Department of Medical Biochemistry and Immunology, School of Medicine, Cardiff University, Heath Park, Cardiff, CF14 4XN, UK, ⁴Section of Pathology & Tumor Biology, Leeds Institute of Molecular Medicine, St James's University Hospital, Leeds, LS9 7TF, UK

Keywords: ABCG2; side-population cells; drug resistance; A549 cells

Dynamic changes in tumour subpopulations may contribute to post-therapeutic repopulation and treatment failure. Human tumours and established cell lines can maintain ‘side’ populations (SPs) - identified by a reduced accumulation of Hoechst dye 33342 attributable to ABCG2-mediated efflux. Progenitors in SP fractions present a challenge for therapies that attempt to target active S phase if they also comprise agents subject to active cellular efflux. The current study addresses the impact of the SP phenotype on drug resistance characteristics with a focus on a predicted resistance to the S-phase targeting anticancer drug DNA topoisomerase I poison topotecan (TPT). The human lung cancer A549 cell line system was used to isolate subclones that initially presented a range of SP expression (< 0.5 - >30 %) and a positive correlation was found between the size of the SP fraction and reduced TPT accumulation associated with ABCG2 expression. Representative clones were established with an initially low SP fraction (<1% SP; A549LowSP) or an initially high SP fraction (>30% SP; A549HighSP) compared with the A549parent (~7% SP). Using RT-PCR, cells were compared for the expression of drug resistance-associated ATP binding cassette transporters (ABCG2, ABCB1). Increased ABCG2, but not ABCB1, expression, correlated with the degree of SP expression. A database survey of 59 human tumor NCI60 panel cell lines suggested that expression of the cancer stem cell marker aldehyde dehydrogenase (ALDH1A1) may augment TPT resistance in cells with low-level expression of ABCG2. However, parental, A549LowSP and A549HighSP cells showed similar co-expression patterns for ALDH genes (ALDH1A1, ALDH2, and ALDH3A1). A549LowSP increased its SP fraction during continued culture while A549HighSP maintained the SP fraction. The extent of the SP fraction reflected the degree of TPT efflux that was sensitive to the ABCG2-inhibitor Fumitremorgin C. To address whether innate SP expression provided an initial advantage for recovery from TPT exposure we studied the dynamic changes in the SP under in vitro selection of A549parent cells. We found that human SPs have a selective advantage, with enrichment of >20-fold, in cell populations recovering from multiple rounds of TPT treatment. Clonal analysis of TPT recovering populations found no evidence of long-term enrichment of SP fraction suggesting innate instability in SP expression.

The data suggest that A549 non-SP cell populations can re-create SP fractions indicating that they are likely to arise by variations that provide a proliferative advantage rather than a unique stem-like progenitor. Innate SP fractions provide the tumor population with a short term selective advantage following TPT exposure related to their enhanced ABCG2 expression, and become enriched under dose fractionation.
Cotreatment with Decitabine and Paclitaxel is Affected by the Re-expression of Drug-Metabolising Enzymes

Laura Cosentino¹, Nahal Masrour², Julie Burns³, Weishuo Fang⁴, Robert Brown², Phil Burns³ and Klaus Pors¹

¹Institute of Cancer Therapeutics, University of Bradford, BD7 1DP, U.K., ²Imperial College London, Hammersmith Hospital Campus, London W12 0NN, U.K., ³Leeds Institute of Molecular Medicine, St James’s Hospital, Leeds LS9 7TF, U.K., ⁴Chemistry of Natural Products, Institute of Materia Medica, Beijing 100050, PR China

The literature describes the role of DNA methylation in regulation of tumour suppressor genes, DNA repair genes, but not much attention has been paid to drug metabolising enzymes (DMEs). Some evidence has been reported and includes methylation of the promoter regions of CYP450 and ABC transporters. Cotreatment with decitabine (DAC) and paclitaxel (PAC) has been reported in several studies, but few of these have reported in a focused manner on the effect on DMEs. As a consequence, the major DMEs involved in the detoxification of PAC (namely ABCB1, CYP2C8 and CYP3A4) were investigated. The aldehyde dehydrogenase enzymes (ALDH1A1, ALDH2 and ALDH3A1) were also explored, as they have recently been shown to play a chemo-protective role against taxanes. Analysis of the CpG island methylation level indicated that all the above-mentioned DMEs were densely methylated and that treatment of MCF-7 cells with DAC for 24 h resulted in a decrease in methylation within the gene promoter region. Furthermore, correlation of decrease in DNA methylation with increase in mRNA level of ABCB1, CYP3A4, ALDH2 and ALDH3A1 was observed, but not for CYP2C8 and ALDH1A1. However, only the ABCB1 protein level was shown to significantly increase. Chemosensitivity of DAC combined with PAC was tested in MCF-7 cells: treatment with DAC for 24 h was followed by PAC exposure on days 4, 6 or 8. Evaluation of the cotreatment on cell survival demonstrated that the cytotoxic potential of PAC was reduced proportionally to the exposure time to DAC before PAC treatment, suggesting that ABCB1 protein confers chemo-protection on the cells. To investigate whether this resistance was applicable to other taxanes with proven efficacy against MDR-resistant cells, the preclinical agent Lx2-32c (azidotaxel) was evaluated. Lx2-32c also lost potency similar to PAC, indicating that ABCB1 is responsible for loss of activity of these two taxanes. The same experiments were also carried out in HeLa cells as they were shown to contain low level of DNA methylation of the above-mentioned DME genes including the ABCB1 gene. DAC treatment did not lead to significant reactivation at protein level and cotreatment of DAC and PAC produced a synergistic effect on cell killing. In conclusion, our data demonstrate that cotreatment of DAC and PAC can produce a synergistic effect, but that it is affected by the re-expression of DMEs. Our observations provide useful information for clinical evaluation of combination strategies comprising DAC and PAC, and chemotherapeutics in general that are affected by DMEs.