Investigating the roles of CDA and DCTD in cancer nucleotide metabolism

A thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

Paolo Spingardi

St Edmund Hall

Supervisors:

Prof. Skirmantas Kriaucionis and Prof. Benedikt Kessler

Ludwig Cancer Research

Nuffield Department of Medicine

University of Oxford

Hilary 2021

Investigating the roles of CDA and DCTD in cancer nucleotide metabolism

Abstract

Nucleotides play essential roles in cellular function, including as components of the nucleic acids DNA and RNA. Cancer cells have elevated proliferation and DNA replication, necessitating a reliable supply of nucleotides from the nucleotide *de novo* synthesis and salvage pathways. These have consequently been exploited as targets for nucleotide analogue chemotherapy. Resistance to chemotherapy can arise from enzymes in the synthesis and salvage pathway, including upregulation of the pyrimidine salvage enzyme cytidine deaminase (CDA). CDA functionally overlaps with the salvage enzyme deoxycytidylate deaminase (DCTD), and it is uncertain whether either enzyme is the principal deaminase and whether they deaminate cytidine nucleotides derived from nucleotide synthesis, in crosstalk between the synthesis and salvage pathways. I investigated these queries by setting up and optimising targeted HPLC-MS/MS systems for nucleic acid composition and nucleotide pool analysis, using them to measure isotopically labelled nucleotide precursor incorporation into cell lines and observe effects of CDA and DCTD siRNA knockdown (KD). The nucleoside analysis platform I developed is capable of sensitive analysis of nearly 20 nucleosides, whereas the free nucleotide analysis platform has analyte sensitivity comparable to or lower than those reported in literature and can detect most nucleotide phosphates without sample purification steps. The KD and pulse-chase experiment provided evidence that CDA and DCTD are mutually redundant in deaminated 2'-deoxycytidine nucleotides and validates previous reports that DCTD has low activity against cytidine monophosphate. Additionally, there was no impact of CDA and DCTD KD on L-glutamine labelling of cytidine nucleotides, suggesting that *de novo* synthesis-derived cytidine nucleotides are not dephosphorylated and deaminated in significant amounts. It was also observed that nucleotide phosphate interconversion occurs at much shorter timeframes than one hour.

Taken together, the studies in this thesis, using targeted and sensitive MS analytical methods, support the theory of redundancy of the CDA and DCTD, with implications for their roles in resistance to cancer chemotherapy as well as in other disease states.

Declaration

This thesis is submitted to the Medical Sciences Division, University of Oxford, in fulfilment of the requirements for the degree of Doctor of Philosophy. This thesis describes my own work and does not constitute part of any other thesis. This work was undertaken at Ludwig Cancer Research Oxford and the Target Discovery Institute, under the supervisions of Prof. Skirmantas Kriaucionis and Prof. Benedikt Kessler. Any work reproduced in this thesis of a third party is mentioned and referenced appropriately.

Paolo Spingardi

Word count: 3129,057983.

Statement on impact of COVID-19 on my research

Due to the COVID-19 pandemic, I was unable to enter the laboratory for experiments between the end of March and the beginning of July. This three-month delay was compensated with an extension to my studies. When I was able to return to the lab, however, time undertaking research and interaction with colleagues were limited to support social distancing measures. Consequently, collaborative experiments I had envisioned for further result validation did not occur.

Acknowledgments

I would first of all like to thank my supervisors, Prof. Skirmantas Kriaucionis and Prof. Benedikt Kessler. From the beginning of my time in their groups, they have been superb guides, mentors and friends; they gave me space to learn, make mistakes and to formulate my own research ideas. They would say that they are paying it forward, that science is at its best as a meeting of many minds and not a centralised initiative, but it is no less appreciated. I am grateful that they collaborated.

I want to thank the Medical Research Council and the Nuffield Department for Medicine for the offering me DPhil studentship, and to Ludwig Cancer Research and the Target Discovery Institute for making me glad to go to work each day. I would also like to thank Agilent Technologies for their support and patience.

I was surrounded by wonderful colleagues; they were ready to give answers to my research questions, often not the answers I wanted to hear, but were full of kindness and support for me and one another. Of past and current members within the Kriaucionis lab, I want to thank Anandhakumar Chandran, Ashvina Segaran, Hiromi Tagoh, Markéta Tomková, Martin Cusack, Michael McClellan, Olena Yavorska, Sophie Kirschner, Wenjun Huang and Ying Bi. The Kessler group has felt like a big family: it is impossible to get to know everyone well, but the get-togethers are wonderful. Many thanks to George Berridge, Zhanru Yu, Marie-L. Thezenas, Hannah Scott and Kathryn Pugh for help, advice and encouragement with mass spectrometry hiccups. Thanks to Andreas Damianou for the excellent advice on siRNA experimental design.

I was very fortunate to have fallen in with the Oxford University Shorinji Kempo Club. Thank you to Chris-Sensei, Steve, Jess, Andy, Mariana and Jai for helping me forget my worries by providing more immediately pressing ones.

I want to thank my wonderful family. While I am not sure whether I managed to explain exactly what it is I have been doing over the last four years, you never gave up asking and have always celebrated my smallest achievements. Thank you to my parents for continually nudging me along the road, to Carlo and Giulia for your friendship and cheerful support, and to my little sisters Helena and Amalia for reminding me how weird and wonderful it is to study things we cannot even see.

Finally, thank you Jane for always being in my corner, even when I was not there myself.

Table of Contents

Table of Contents vi

List of figures x

List of tables xvii

List of abbreviations xviii

1. Introduction 1

1.1 Introduction abstract 1

1.2 The role of nucleotides in cell processes 1

1.3 Nucleotide base pairing underpins nucleic acid replication 5

1.4 Nucleotides are produced through *de novo* synthesis and salvage pathways 7

1.4.1 Purine nucleotide metabolism 9

1.4.2 Pyrimidine nucleotide metabolism 12

1.5 Nucleotide depletion to avoid pool imbalance and to remove damaging nucleotides from pool 16

1.6 Removal of incorrect bases in DNA through repair 18

1.7 Role of nucleotide metabolism in cancer 20

1.7.1 Mechanisms of chemotherapy in cancer nucleotide metabolism 21

1.7.1.1 Salvage and sanitation pathways in chemotherapy resistance 22

1.7.1.2 CDA creates vulnerability to epigenetically modified nucleotides 24

1.8 Roles of pyrimidine deaminases CDA and DCTD 25

1.9 Methods of studying nucleotide metabolism 29

1.9.1 Tools for enzyme manipulation 30

1.9.2 Nucleotide metabolism analysis 31

1.9.2.1 Nucleotide metabolism analysis: the free nucleotide pool 31

1.9.2.2 Nucleotide metabolism analysis: nucleic acid composition 35

1.9.3 Overview of the Agilent 1290-6495a QQQ HPLC-MS/MS 37

1.10 Aims of the project 39

2. Materials and Methods 43

2.1 Cell culture 43

2.2 siRNA transfection 44

2.3 Immunoblotting 45

2.4 Pulse-chase experiment 47

2.4.1 Nucleotide extraction from biological samples 48

2.5 Nucleotide analysis 48

2.5.1 Nucleotides external and internal standards 48

2.6 Nucleotide mass spectrometric conditions 49

2.7 Nucleoside analysis 50
2.7.1 Standards for nucleoside mass spectrometry analysis and for heavy labelled incorporation 50
2.7.2 Nucleic acid isolation 50
2.7.3 Nucleic acid hydrolysis 51
2.7.4 Nucleoside HPLC-MS/MS analysis 51
2.8 Statistical analysis 52
3. Optimising an HPLC-MS/MS method for nucleic acid analysis 54
3.1 Optimising the hydrolysis process of nucleic acids to nucleosides 54
3.2 Choice of HPLC column for optimising nucleoside separation 56
3.3 Optimising the Agilent 6495a QQQ MS for nucleoside signal 62
3.3.1 Choice of QQQ polarity for nucleoside signal 62
3.3.2 Impact of HPLC mobile phase on nucleoside MS sensitivity 63
3.3.2.1 Investigation of 3 mobile phase compositions for nucleoside response 63
3.3.2.2 Changing Solvent B from 40 % ACN to 100 % MeOH increases nucleoside sensitivity 66
3.3.3 MRM transition choice for maximising sensitivity 68
3.3.4 Optimisation of MS source conditions 71
3.4 Use of internal standards for consistency in quantitation 75
3.5 Determining nucleoside limits of detection 77
3.6 Combination of HPLC-UV and MS analysis 79
3.7 Validation of MS sample analysis stability 82
3.8 Discussion 83
3.8.1 Nucleic acid hydrolysis optimisation 83
3.8.2 Optimising HPLC analyte separation 84
3.8.3 Optimising MS for nucleoside signal 84
3.8.4 Stability of nucleoside sample analysis 86
4. Setting up a workflow for nucleotide analysis 87
4.1 Nucleotide pool extraction optimisation 88
4.1.1 Solvent extraction 89
4.1.2 Sample treatment 93
4.1.3 Injection volume 96
4.2 Optimising HPLC for nucleotide analysis 97
4.2.1 Column 97
4.3 Optimising MS/MS for nucleotide analysis 103
4.3.1 Choice of polarity of instrument 103
4.3.2 Solvent composition 109
4.3.3 Solvent composition: salt 109

4.3.3.1 Solvent composition: pH 111
4.3.3.2 Solvent composition: ammonium fluoride 114
4.3.4 HPLC gradient optimisation for new solvent 116
4.3.5 Mass spectrometric conditions 119
4.3.5.1 Collision energy 119
4.3.5.2 EMV 121
4.3.6 Determination of sensitivity 122
4.4 Validation 125
4.5 Discussion 127
4.5.1 Nucleotide pool extraction 127
4.5.2 Choice of column chromatography 128
4.5.3 MS optimisation 128
4.5.4 Validation 130
5. Cell-based tools for investigating nucleotide metabolism 132
5.1 CDA and DCTD siRNA KD 133
5.1.1 Cell line choice 133
5.1.2 Validation of siRNAs for CDA and DCTD knockdown 138
5.2 Choice of heavy labelled tracers for MS 141
5.2.1 Measurement of nucleotide precursors in media 141
5.2.2 Suitable labelled tracers for the pyrimidine synthesis and salvage pathways 143
5.2.2.1 Synthesis pathway 143
5.2.2.2 Salvage pathway 147
5.2.3 Simultaneous discrimination of synthesis and salvage tracer incorporation 148
5.3 Discussion 151
5.3.1 CDA and DCTD KD 151
5.3.2 Heavy labelled tracers for MS 151
6. Role of CDA and DCTD in Nucleotide Pool Metabolism 154
6.1 Introduction 154
6.2 Sample DNA and nucleotide pool harvest 155
6.3 LC-MS/MS signal quality for nucleosides and nucleotides 156
6.4 Observations of nucleotide metabolism from pulse-chase analysis 158
6.4.1 Nucleic acid labelling after 24 hours exposure 159
6.4.2 Free nucleotides have faster turnover than nucleic acids 160
6.4.3 RNA is turned over faster than DNA 164
6.5 Roles of CDA and DCTD in pyrimidine deamination 174
6.5.1 CDA KD reduces rC deamination more than DCTD KD in cell lines 174
6.5.2 CDA KD induces faster rC nucleotide signal decrease than DCTD 180
6.6 CDA KD may create a bottleneck for rC deamination 185

6.7 L-glutamine label incorporation not affected by CDA nor DCTD KD 188
6.8 Discussion 191
6.8.1 General observations of the data 191
6.8.2 CDA and DCTD both contribute to cytidine nucleotide deamination 192
6.8.3 *De novo* synthesis cytidine nucleotides are not metabolised by CDA and DCTD 193
7. General Discussion 195
7.1 Optimisation of HPLC-MS/MS method for nucleoside analysis 195
7.2 Set up of HPLC-MS/MS analysis of cellular nucleotide pool 196
7.3 Cell-based tools for investigating nucleotide metabolism 198
7.4 Role of CDA and DCTD in nucleotide pool metabolism 199
8. References 203
Appendix A 219

List of figures

Figure 1.1 Nucleotide metabolism plays central role in cells. Nucleotides are components in DNA (blue, thymidine monomer pictured); RNA (orange, uridine monomer pictured); enzyme cofactors (red, NAD+ pictured); cell signalling (pink, cAMP pictured); energy transport (green, ATP pictured). 2
Figure 1.2 Illustration of the structures of the nucleobases C, U, T, G, A, X, H and the nucleosides dC and rC (N-glycosidic linkage in red). 5
Figure 1.3 Reading sequence direction of DNA, in the nucleotide triad thymidine-thymidine-cytidine (TTC). 5
Figure 1.4 Base pairing in DNA and RNA. Base pairing through complementary hydrogen bonding between nucleobases provides mechanism for DNA replication and transcription, and RNA translation. **A,** C:G base pairing (RNA and DNA); **B,** T:A base pairing (DNA); **C,** U:A base pairing (RNA). 6
Figure 1.5 Human purine de novo synthesis pathway 10
Figure 1.6 Human purine salvage pathway. 12
Figure 1.7 Human pyrimidine de novo synthesis pathway 14
Figure 1.8 Human pyrimidine salvage pathway. 16
Figure 1.9 DNA Chemical damage to a base can lead to mutation in daughter cells. Deamination of C forms a U base. If not removed, upon the next replication it will pair with an A. This leads to a C->T sequence mutation. 19
Figure 1.10 Deamination of dC to dU by CDA. CDA also deaminates rC, and also cytidine analogues used in chemotherapy such as gemcitabine. 24
Figure 1.11 Structures and active sites for the deaminases. CDA is tetrameric and DCTD is hexameric. Both incorporate zinc as a participating ion. Crystal structures obtained from Protein Data Bank. **A,** structure of CDA; **B,** structure of DCTD; **C,** deamination of dC by CDA, showing active site of CDA bound to dC; **D,** deamination of dCMP by DCTD, showing active site of DCTD bound to dCMP. 26
Figure 1.12 DCTD is more uniformly distributed in human tissues than CDA. Tissue-specific mRNA levels, obtained through BioGPS. **A,** mRNA tissue map of CDA; **B,** mRNA tissue map of DCTD. 27
Figure 1.13 Schematic representation of the Agilent Jet Stream electrospray ionisation (AJS ESI) source. The details of the AJS arrangement in front of the front end of a triple quadrupole mass spectrometry (QQQ) are shown, illustrating the gas flow (red) and the area of enhanced desolvation (blue), reproduced with permission from Agilent Technologies ("Agilent Jet Stream Thermal Gradient Focusing Technology", Mordehai and Fjelsted).[199] 38
Figure 1.14 Schematic of mechanism of MRM mode analysis. Parent ions A and C are isolated from mix A, B, C by quadrupole 1 (Q1), fragmented using nitrogen (N2) gas in Q2, leading to daughter ions A1, A2, C1, C2, which are again isolated for detection in Q3. 39
Figure 1.15 Workflow for analysis of nucleotide pool and nucleic acid composition from treated samples. Treatment: cells are treated for CDA and DCTD KD, and for precursor labelling exposure. **Extraction**: nucleotides are extracted and nucleic acid is isolated from remaining cell pellet. Nucleic acid is hydrolysed to nucleosides by nucleases and phosphatase. **Analysis:** nucleotide and nucleoside labelling is quantified by MS. 41
Figure 3.1 Increasing hydrolysis time does not improve signal of dT by UV. 500 ng of DNA hydrolysed, n=3 technical replicate hydrolysis experiments. 55
Figure 3.2 Increasing PDE concentration in hydrolysis mix does not increase dT signal by UV. 500 ng of DNA hydrolysed, n=3 technical replicate hydrolysis experiments. 56
Figure 3.3 2'-deoxycytidine (dC), with atoms likely to interact favourably with polar solvents in red, and the hydrophobic molecule skeleton in black. N-5 is

not highlighted as its lone pair of electrons partially delocalised in the aromatic heterocycle and may engage less in solvent interactions. 57
Figure 3.4 Column 2 gives best nucleoside resolution. Extracted ion chromatograms (EICs) of 12.5 pmol of nucleoside standards cadC, dC, hmdC, hmdU, dU, mdC, dG, fdC, 8-oxo-dG, dT analysed by LC-MS scan mode. **A,** separation by column 1; **B,** separation by column 2; **C,** separation by column 3. 62
Figure 3.5 Nucleosides can form different ion adducts. Structures of nucleoside positive ion adducts generated by ESI. **A,** protonated dC (m/z 228); **B,** sodiated dC (m/z 250); **C,** structure of sodiated dT (m/z 265). 63
Figure 3.6 Tributylamine suppresses nucleoside signal through saturation. Extracted ion chromatograms (EICs) of tributylamine and dU in solvent system 1. TBA signal is saturated, with no signal from dU**. A,** EIC of dU Na^+, 12.5pmol; **B,** EIC of TBA H^+. 65
Figure 3.7. **MS nucleoside detection sensitivity is solvent-dependent sensitivity of nucleoside detection.** Response of 1.25 pmol of nucleoside standards dU Na+ (green bars), hmdC Na+ (blue bars) and hmdC H+ (red bars) in LC-MS mode using solvent systems 2 (left panel) or system 3 (right panel). n=4, standard deviation is shown. 66
Figure 3.8 Changing Solvent B composition from 40:60 ACN:H_2O to 100 % MeOH increases BrdU signal, along with other nucleosides. The nucleoside signal of 500 fmol standards was compared for solvent B 40:60 ACN:H_2O (magenta) vs solvent B MeOH (yellow). **A,** total ion chromatogram (TIC, summation of all individual EICs) of nucleosides; **B,** EIC of BrdU. 68
Figure 3.9 Search for a suitable parent ion for MRM. A, full scan ion chromatogram of dC ion adducts; **B,** schematic of beta-glycoside fragmentation. 69
Figure 3.10 Ions with the same m/z interfere with each other's quantification. Extracted MRM ion chromatogram of m/z = 251 in 1.25 pmol of standards, showing both dU Na^+ and dC +1 Da Na^+. In biological samples, dC+1 Da Na^+ will have a much higher signal than dU Na^+. 70
Figure 3.11 Nucleoside separation post MS and HPLC optimisation. Extracted dMRM ion chromatograms of nucleosides from a 500 fmol standard injection. 71
Figure 3.12 Varying MS source parameters allows identification for best nucleoside signal. A, dU Na+ response vs capillary voltage; **B,** dU Na+ response vs sheath gas temperature. 73
Figure 3.13 dMRM mode gives more than 10 x higher signal-to-noise ratio than scan mode. EIC of **scan mode** dC H+ (green, 1.25 pmol, dMRM, SNR=1123) and EIC of **dMRM mode** dC H+ (purple, 1.25 pmol, MS, SNR=92.5). 75
Figure 3.14 Nucleosides are liable to changing ion adduct preferences over the course of multiple chromatographic sequences. Comparisons of the EIC of dC ion adducts, at start of worklist (purple) and end of worklist (green), with 500 fmol injections. **A,** dC H^+; **B,** dC Na^+. 76
Figure 3.15 Calibration curves for hmdC with internal standard. Two injections per amount on column, normalised to proportion compared to internal standard injected (62.5 fmol). **A,** hmdC H^+ whole curve; **B**, hmdC H^+ curve at low amounts; **C**, comparison of hmdC H^+ signal (500 fmol) with $^{13}C_1{}^{15}N_2$ hmdC H^+ (62.5 fmol). 79
Figure 3.16 dC MS signal saturates at 1.25 pmol and above, while UV signal does not. Colour keys: 1.25 pmol (blue), 2.5 pmol (green), 12.5 pmol (red) and 25 pmol (purple) injected on the column. **A,** MS signal; **B,** UV absorbance; **C,** internal standard ($^{13}C_1{}^{15}N_2$ dC). The four calibration points were obtained by 5 and 10 µL injections of 250 nM and 2.5 µM standards, so therefore the 10 uL injections have a stronger internal standard signal than the 5 µL injections (twice the amount injected). 81
Figure 3.17 UV Calibration curves for dC, one injection per concentration. **A,** dC whole curve; **B,** dC H curve at low amounts. 82

Figure 3.18 % mdC/dG is stable over different measurements. 500 ng HEK-293 TREX-TET-CD-GFP +Dox, n=4 replicates hydrolysed on different days. 83
Figure 4.1 50 % ACN gave best response for dCMP and dCDP, mixed response from dCTP and dTTP. Effect of extraction mixture choice on nucleotide extracted ion chromatogram signals, (n=3). A, effect of extraction solvent composition on dCMP; B, effect of extraction solvent composition on dCDP; C, effect of extraction solvent composition on dCTP; D, effect of extraction solvent composition on dTTP.92
Figure 4.2 dCMP is second peak of 308→112 EIC. EICs of labelled (red) and unlabelled dCMP H^+ (black) in H1299 cells. heavy labelled dCMP H^+ (37.5 pmol concentration, red) aligns with second peak of transition including dCMP H^+ unlabelled (black). 93
Figure 4.3 Heating samples increases nucleotide signal. Effect of sample treatment on nucleotide signal (n=3). A, effect for dCMP; **B**, effect for dCDP; **C**, effect for dCTP. 95
Figure 4.4 Metabolite extraction process does not cause significant nucleotide triphosphate dephosphorylation. 1 µM of internal standards were spiked into extraction solvent of nucleotides extracted from H1299 cells, with 1 µL of prepared sample injected. Comparison of the EIC of the internal standard of dCMP (320→119, green) with the EIC of the calculated dephosphorylated product of the internal standard of dCTP (317→116, yellow) to assess extent of nucleotide triphosphate degradation. 95
Figure 4.5 dTTP Na^+ signal plateaus at 1 µL injection. Effect of injection volume on dTTP Na^+ signal (n=3). 97
Figure 4.6 The nucleotides dCMP, dCDP and dCTP have similar structure and differ in the number of phosphates. 98
Figure 4.7 Schematic of nucleotide-column interactions with dCMP. A, the PFP column retains polar aromatic compounds such as nucleotides through its dipole-dipole interactions between the fluorine atoms and phosphate hydrogens, as well as aryl-aryl interactions between the phenyl group and the nucleotide heterocycles; **B,** the HILIC-Z column retains polar compounds by dipole-dipole interactions between the bonded zwitterions and the charged nucleotide phosphates. 100
Figure 4.8 Nucleotide separation is better on HILIC-Z column than PFP column. Separation was judged by UV absorbance analysis. The PFP column only showed separation of monophosphate nucleotides, with no signal obtained of multiply phosphorylated nucleotides. The HILIC-Z column resolved the mono-, di- and tri-phosphate nucleotides. **A,** separation of 5 pmol of dCMP, dUMP and dGMP; **B,** separation of 5 pmol dCDP, dCTP, dGDP and dGTP (no detection); **C,** separation of 5 pmol of dCMP, dCDP and dCTP. 103
Figure 4.9 Positive and negative ionisation of dCTP. The dCTP H^+ ion has a protonated nucleobase, whereas dCTP-H^- has a deprotonated phosphate group. 104
Figure 4.10 The ion transitions of positively and negatively charged nucleotides are different. A, the principal ion transition of ATP and dGTP in negative mode (m/z 506→159); **B,** the principal ion transitions of ATP and dGTP in positive mode (m/z 508→136 and 508→152). 106
Figure 4.11 Positive mode gives stronger ATP and dGTP signal than negative mode. A, extracted ion chromatogram (m/z 506→136) in negative mode, 1pmol injected on column; **B,** overlaid extracted ion chromatograms of ATP and dGTP in positive mode (m/z 508→136 and 508→152), 1pmol injected on column. These are stronger than in negative mode and are orthogonal, with no signal interference. .. 107
Figure 4.12 The main transitions in positive mode for dCMP and dTTP. A, schematic of transitions for dTTP in positive mode, as the Na^+ adduct; **B,** the main transitions for 1 pmol dCMP H^+; **C,** the main transitions for 1 pmol dTTP Na^+. 108
Figure 4.13 10 mM ammonium bicarbonate gives the strongest nucleotide signal. 1 pmol of nucleotides was injected onto the column (n=4)**. A,** effect on dCMP H^+; **B,** effect on dUMP H^+; **C,** effect on dTTP H^+ 111

Figure 4.14 Solvent gave strongest signal overall at pH 6. 1 pmol of nucleotides was injected onto the column (n=4). **A,** effect on dCMP; **B,** effect on dTTP; **C,** effect on dUMP. 114
Figure 4.15 Ammonium fluoride reduces nucleotide monophosphate signal and increases nucleotide triphosphate signal. 1 pmol of nucleotides was injected onto the column (n=4). **A,** effect on dCMP; **B,** effect on dUMP; **C,** effect on dTTP.116
Figure 4.16 Percentage solvent B over time of gradients A, B and C. 117
Figure 4.17 Gradient C gave best resolution of dCMP, dCDP and dCTP. 1 pmol of nucleotides was injected onto the column (n=3) **A,** nucleotide resolution of dCMP, dCDP and dCTP with gradient A; **B,** nucleotide resolution of dCMP, dCDP and dCTP with gradient B; **C,** nucleotide resolution of dCMP, dCDP and dCTP with gradient C. 119
Figure 4.18 Best collision energy voltage for nucleotide transitions varies per nucleotide. 1 pmol of nucleotides was injected onto the column per run. **A,** dCMP H^+ (10 V); **B,** dCTP H^+ (30 V); **C,** dTTP Na^+ (20 V). 121
Figure 4.19 Increasing delta EMV increases signal of dCMP. 1 pmol of nucleotides was injected onto the column per run. 122
Figure 4.20 Calibration curves for dCMP H^+ with internal standard. Two injections per amount on column, normalised to proportion compared to internal standard injected (100 fmol). **A,** dCMP H^+ whole curve; **B**, dCMP H^+ curve at low amounts; **C**, comparison of dCMP H^+ signal (100 fmol) with $^{13}C_9$ $^{15}N_3$ dCMP H^+. .. 124
Figure 4.21 Optimised signal and chromatographic separation of all nucleotides. 1 pmol of nucleotides was injected onto the column per run. 124
Figure 4.22 CDA overexpression may reduce intracellular dCTP. Comparison of dCTP H^+ signal from WT H1299 vs H1299 overexpressed CDA (not enough points for statistical analysis). **A,** dCTP H^+ in H1299 WT (n=3) and in overexpressed CDA (**n=1**); **B,** WB showing CDA and β-actin in WT H1299 and H1299 CDA overexpressed. 126
Figure 5.1 Workflow of siRNA and pulse-chase experiments. Cells are seeded and transfected the day after. At t=48 hours after transfection, cells are treated with heavy labelled precursors. The media is changed at t=72 hours after transfection and the cells are regularly harvested for 24 hours, with metabolites and DNA analysed by LC-MS/MS. 133
Figure 5.2 Cells with high CDA and DCTD mRNA levels compared to NCI-60 panel. log_2(FPKM+1) of CDA (blue) and DCTD (red) as a percentage compared to highest FPKM in NCI-60 cell line panel. Data obtained from CellMiner.[236,237] 134
Figure 5.3 PC-3, MDA MB-231 and HOP-92 have detectable levels of CDA and DCTD protein. Immunoblot of CDA and DCTD for all cell lines, with β-actin blot showing loading distribution. **A,** CDA protein levels **B,** DCTD protein levels. 137
Figure 5.4 CDA and DCTD KD is stable over time in the MDA MB-231 cell line. Immunoblot of CDA and DCTD for MDA MB-231, with Ponceau stain showing loading distribution. **A**, CDA KD compared to scrambled siRNA from 72 hrs to 96 hrs post exposure to siRNAs**; B**, DCTD KD compared to scrambled siRNA from 72 hrs to 96 hrs post exposure to siRNAs**.** 139
Figure 5.5 CDA and DCTD KD is stable over time in the PC-3 cell line. Immunoblot of CDA and DCTD for MDA MB-231, with Ponceau stain showing loading distribution. **A**, CDA KD compared to scrambled siRNA from 72 hrs to 96 hrs post exposure to siRNAs**; B**, DCTD KD compared to scrambled siRNA from 72 hrs to 96 hrs post exposure to siRNAs**.** 140
Figure 5.6 Labelled L-glutamine and orotic acid are metabolised to form labelled pyrimidines. $^{13}C_5$ L-glutamine forms $^{13}C_3$ pyrimidine bases and $^{13}C^{15}N_2$ orotic acid forms $^{13}C^{15}N_2$ pyrimidine bases. Heavy atoms are marked with an asterisk. 145
Figure 5.7 Incorporation of heavy labelled synthesis pathway tracers as dC. Cells were treated with labelled tracers for 24 hrs in technical triplicates, mean

values shown, error bars give +-1 standard deviation. **A,** comparison of +3 Da dC signal in untreated cells, MDA MB-231 with 500 μM labelled L-glutamine or 10 μM labelled orotic acid; **B,** comparison of +3 Da dC signal in untreated cells, MDA MB-231 and PC-3, each treated with 500 μM labelled L-glutamine..... 147
Figure 5.8 Labelling of $^{13}C_9$ dC (dC +9 Da) and $^{13}C_5$ rC (rC +5 Da). Heavy atoms are marked with an asterisk 147
Figure 5.9 % incorporation of $^{13}C_9$ dC and $^{13}C_5$ rC. into DNA via incorporation of heavy labelled synthesis pathway tracers as dC. Cells were treated with labelled tracers for 24 hrs in technical triplicates, mean values shown, error bars give +-1 standard deviation. **A,** incorporation of $^{13}C_9$ dC as dC; **B,** incorporation of $^{13}C_5$ rC as dC; **C,** incorporation of $^{13}C_9$ dC as dT; **D,** incorporation of $^{13}C_5$ rC as dT. 148
Figure 5.10 Incorporation of labelled nucleotide synthesis and salvage precursors as dC in nucleic acids. A, peaks of labelled dC compared to unlabelled dC in PC-3 cell line (red); **B,** zoom in to peaks of labelled dC in PC-3 cell line from $^{13}C_5$ L-glutamine (yellow), from $^{13}C_5$ rC (black) and from $^{13}C_9$ dC (magenta); **C,** incorporation fates of heavy labelled L-glutamine (Q), dC and rC. 150
Figure 6.1 Samples harvested at set times over 24 hrs in pulse-chase experiment. After siRNA transfection, the samples were exposed to 1 μM dC +9 Da, 1 μM rC +5 Da and 500 μM L-glutamine +5 Da for 24 hrs. The media was replaced and the samples were harvested at 0, 1, 2, 4, 6, 12 and 24 hrs after replacement of media. Both the nucleotide pool and nucleic acid were extracted from the samples for label incorporation analysis..... 156
Figure 6.2 Precursor labelling is observed in DNA/RNA nucleosides and in the nucleotide pool. Label distributions in MDA MB-231 exposed to scrambled siRNA at t=0 hrs post media replacement. **A,** labelling of dC (dC H^+) in DNA; **B,** labelling of rC (rC H^+) in RNA; **C,** labelling of dT (dT Na^+) in DNA; **D,** labelling of dCMP (dCMP H^+) in nucleotide pool; **E,** labelling of CMP (CMP H^+) in nucleotide pool..... 158
Figure 6.3 rC/rU +5 Da nucleotide fractional turnover is similar across phosphorylation states. +5 Da incorporation levels in cells exposed to scrambled siRNA decay at similar rates for mono- di- and tri-phosphates for both cytidine and uridine ribonucleotides. +5 Da rC proportion in RNA reduces at a slower rate than in nucleotides. 3 biological replicates normalised to t=0 hrs per each biological replicate, standard deviation is shown. **A,** +5 Da proportion reduction for cytidine ribonucleotides and cytidine in RNA in MDA MB-231 cell line; **B,** +5 Da proportion reduction for cytidine ribonucleotides and cytidine in RNA in PC-3 cell line; decay for +5 Da uridine ribonucleotides. **C,** +5 Da proportion reduction for uridine ribonucleotides in MDA MB-231 cell line; **D,** +5 Da proportion reduction for cytidine ribonucleotides in PC-3 cell line. 164
Figure 6.4 DNA labelling decays slower than RNA labelling. The dC +5 Da (DNA) label proportion decreases more slowly that of rC +5 Da (RNA). Labelled incorporation levels in cells exposed to scrambled siRNA, 3 biological replicates normalised to t=0 hrs per each biological replicate, standard deviation is shown. **A,** comparison of the decay of % +5 Da in dC and rC in MDA MB-231 cell line; **B,** comparison of the decay of % +5 Da in dC and rC in PC-3 cell line. 165
Figure 6.5 Labelled dCMP has a longer decay time than labelled CMP. Measurement of decay of proportion of dCMP +9 Da, dCMP +5 Da and CMP +5 Da over 24 hours. Experiment was in cells exposed to scrambled siRNA, 3 biological replicates normalised to t=0 hrs per each biological replicate, standard deviation is shown. **A,** signal decay over time in MDA MB-231 cell line; **B,** signal decay over time in PC-3 cell line. 167
Figure 6.6 Label proportions from different precursors decay at different rates. The percentage incorporation of labelled species varies across the 24 hours going down fastest in labelling derived from dC +9 Da and slowest in labelling derived from L-glutamine (+3 Da). Labelled incorporation levels in cells exposed to scrambled siRNA, 3 biological replicates normalised to t=0 hrs per each biological replicate,

standard deviation is shown. **A,** decay of label proportion in dC in MDA MB-231 cell line; **B,** decay of label proportion in dC in PC-3 cell line; **C,** decay of label proportion in dT in MDA MB-231 cell line; **D,** decay of label proportion in dT in PC-3 cell line; **E,** decay of label proportion in CMP in MDA MB-231 cell line; **F,** labelled % incorporation in CMP in PC-3 cell line; **G,** decay of label proportion in CDP in MDA MB-231 cell line; **H,** decay of label proportion in CDP in PC-3 cell line..... 173
Figure 6.7 KD of CDA affects rC +5 incorporation, while DCTD KD does not. CDA KD increases incorporation in of metabolites of rC +5 in cytidine variants, but the same effect is not observed for DCTD KD. Data from cell lines MDA MB-231 (M in figure) and PC-3 (P in figure) at t= 0 hrs post 24 hrs heavy label pulse. 2 orthogonal siRNAs were used per CDA and DCTD KD and compared to a scrambled siRNA with 3 biological replicates, standard deviation is shown. A, incorporation of +5 Da label in rC; B, incorporation of +5 Da label in dC; C, incorporation of +5 Da label in CMP; D, incorporation of +5 Da label in CDP; E, incorporation of +5 Da label in CTP; F, incorporation of +5 Da label in UMP; G, incorporation of +5 Da label in UDP; H, incorporation of +5 Da label in UTP; I, incorporation of dC +9 as dC..... 180
Figure 6.8 CDA KD correlates with greater reduction of +5 Da signal over time across cytidine metabolites, with effect not seen for +9 Da. 2 orthogonal siRNAs were used per CDA and DCTD KD and compared to a scrambled siRNA with 3 biological replicates normalised to t=0 hrs per each biological replicate, standard deviation is shown and significance is judged at t=12 hrs. **A,** +5 Da labelling of CMP in MDA MB-231 cell line (CDA1 relative to scrambled, p=0.024, CDA 3 relative to scrambled, p=0.0143); **B,** +5 Da labelling of CMP in PC-3 cell line; **C,** +5 Da labelling of CDP in MDA MB-231 cell line (CDA1 relative to scrambled, p=0.037, CDA 3 relative to scrambled, p=0.039); **D,** +5 Da labelling of CDP in PC-3 cell line (CDA1 relative to scrambled, p=0.0057, CDA 3 relative to scrambled, p=0.0083, DCTD 2 relative to scrambled, p=0.042); **E,** +5 Da labelling of CTP in MDA MB-231 cell line; **F,** +5 Da labelling of CTP in PC-3 cell line (CDA 3 relative to scrambled, p=0.019). 185
Figure 6.9 CDA KD increases deaminated dC incorporation as dT in PC-3 cell line, but not MDA MB-231 cell line, at t=0 hrs. Corresponding decrease in dC +9 Da incorporation not seen. 2 orthogonal siRNAs were used per CDA and DCTD KD and compared to a scrambled siRNA with 3 biological replicates, standard deviation is shown. **A,** incorporation of dC +9 as dT; **B,** incorporation of rC +5 as dT. 186
Figure 6.10 Schematic of competition between unlabelled, +5 Da and +9 Da dTTP nucleotide pools. With dT supplementation in media, CDA KD reduces the nucleotide pool size of dTTP derived from rC +5 Da, but not dTTP derived from dC +9 Da. DCTD KD reduces neither. **A,** schematic of dTTP provenance in WT conditions; **B,** schematic of dTTP provenance upon CDA KD; **C,** schematic of dTTP provenance upon DCTD KD..... 187
Figure 6.11 Labelled L-glutamine incorporation does not change upon CDA or DCTD KD, at t=0 hrs. 2 orthogonal siRNAs were used per CDA and DCTD KD and compared to a scrambled siRNA with 3 biological replicates, standard deviation is shown. **A,** incorporation of dC +3 Da; **B,** incorporation of dC +3 Da; **C,** incorporation of CMP +3 Da; **D,** incorporation of CDP +3 Da..... 190
Figure 0.1 PC-3, MDA MB-231 and HOP-92 have detectable levels of CDA and DCTD protein, second replicate. Immunoblot of CDA and DCTD for all cell lines, with β-actin blot showing loading distribution. **A,** CDA protein levels **B,** DCTD protein levels, also showing from DCTD ECL stain. 228
Figure 0.2 CDA and DCTD KD is stable over time in the MDA MB-231 cell line, second replicate. Immunoblot of CDA and DCTD for MDA MB-231, with Ponceau stain showing loading distribution. **A**, CDA KD compared to scrambled siRNA from 72 hrs to 96 hrs post exposure to siRNAs**; B**, DCTD KD compared to scrambled siRNA from 72 hrs to 96 hrs post exposure to siRNAs**.** 229

Figure 0.3 CDA and DCTD KD is stable over time in the PC-3 cell line, second replicate. Immunoblot of CDA and DCTD for MDA MB-231, with Ponceau stain showing loading distribution. **A**, CDA KD compared to scrambled siRNA from 72 hrs to 96 hrs post exposure to siRNAs**; B**, DCTD KD compared to scrambled siRNA from 72 hrs to 96 hrs post exposure to siRNAs**.** 230

List of tables

Table 2.1 Sequences of oligonucleotides used for siRNA KD of CDA and DCTD 44
Table 2.2 Antibodies used for comparing CDA and DCTD levels across samples. 47
Table 3.1 Column types used in literature for nucleoside separation. 58
Table 3.2 Description of columns tested for nucleoside, judging for separation. 59
Table 3.3 List of mobile phase systems tested for nucleoside signal. 64
Table 3.4 MS source parameters optimised for nucleoside signal. 74
Table 4.1 Solvents tested for pool extraction. 89
Table 4.2 Columns tested for nucleotide resolution. 99
Table 4.3 Percentage solvent composition over time of gradient C. 119
Table 5.1 Cell lines chosen for high CDA and DCTD (log_2(FPKM+1)). 135
Table 5.2 Concentrations of nucleosides found in complete media. 142
Table 6.1 Percentage incorporation of heavy labelled precursors in nucleic acid components. 160
Table 13 A list of each nucleoside transition monitored by dMRM with collision energies and retention times. 220
Table 14 Table of nucleoside retention times, LODs and LLOQs for each nucleoside analysed. 223
Table 15 List of nucleotides retention times, their MS transitions and their LLOQs. 225

List of abbreviations

^{13}C	carbon-13
^{15}N	nitrogen-15
8-oxodG	8-hydroxy-2'-deoxyguanosine
8-oxoG	8-hydroxyguanosine
°C	degrees centigrade
µL	microlitre
A	adenine
a	atto
ACN	acetonitrile
ADP	adenosine diphosphate
AMP	adenosine monophosphate
ATP	adenosine triphosphate
bp	base pair
BrdU	5-bromo-2'-deoxyuridine
C	cytosine
cadC	5-carboxy-2'-deoxycytidine
CDA	cytidine deaminase
CDP	cytidine diphosphate
CMP	cytidine monophosphate
CTP	cytidine triphosphate
dA	2'-deoxyadenosine
Da	Dalton
dADP	2'-deoxyadenosine diphosphate
dAMP	2'-deoxyadenosine monophosphate
dATP	2'-deoxyadenosine triphosphate
dC	2'-deoxycytidine
dCDP	2'-deoxycytidine diphosphate
dCMP	2'-deoxycytidine monophosphate
DCTD	deoxycytidylate deaminase
dCTP	2'-deoxycytidine triphosphate
dG	2'-deoxyguanosine
dGDP	2'-deoxyguanosine diphosphate
dGMP	2'-deoxyguanosine monophosphate
dGTP	2'-deoxyguanosine triphosphate
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DNMT	DNA methyltransferase
dT	thymidine
dTDP	thymidine diphosphate
dTMP	thymidine monophosphate
dTTP	thymidine triphosphate
dU	2'-deoxyuridine
dUDP	2'-deoxyuridine diphosphate
dUMP	2'-deoxyuridine monophosphate
dUTP	2'-deoxyuridine triphosphate
EdU	5-ethynyl-2'-deoxyuridine

fdC	5-formyl-2'-deoxycytidine
FPKM	fragments per kilobase per million
G	guanine
GDP	guanosine diphosphate
GMP	guanosine monophosphate
GTP	guanosine triphosphate
H2O	water
hmC	5-hydroxymethylcytosine
hmdC	5-hydroxymethyl-2'-deoxycytidine
hmdU	5-hydroxymethyl-2'-deoxuridine
HPLC	high-performance liquid chromatography
hr	hour(s)
IP	immunoprecipitation
KD	gene knockdown
KO	gene knockout
LLOQ	lower limit of quantitation
LOD	limit of detection
M	molar
m	metre
m6dA	n6-methyl-2'-deoxyadenosine
m6ra	n6-methyladenosine
mC	5-methylcytosine
mdC	5-methyl-2'-deoxycytidine
MeOH	methanol
min	minute
mL	millilitre
mol	mole
MS	mass spectrometry
NMR	nuclear magnetic resonance
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline solution
PDE I	phosphodiesterase I
Q	L-glutamine
rA	adenosine
rC	cytidine
rG	guanosine
RNA	ribonucleic acid
RNA-seq	RNA sequencing
RNR	ribonucleotide reductase
RRM1	ribonucleoside-diphosphate reductase large subunit
rU	uridine
s	seconds
shRNA	short hairpin RNA
siRNA	small interfering RNA
T	thymine
TBST	tris-buffered saline supplemented with Tween-20
TYMS	thymidylate synthase
U	uracil

UDP	uridine diphosphate
UMP	uridine monophosphate
UTP	uridine triphosphate
UV	ultraviolet
WB	western blot
WT	wild type

prefixes

k	kilo
m	milli
µ	micro
n	nano
p	pico
f	femto
a	atto

1. Introduction

1.1 Introduction abstract

Nucleotide metabolism is central to cell survival, with nucleotides playing roles in processes including nucleic acid synthesis and repair, energy transport and cell signalling. Cancer cells have faster cell division with a corresponding increase in DNA replication and in nucleotide requirements. The nucleotide acquisition pathways have therefore been frequent targets for chemotherapeutic agents, with drug resistance often occurs through the same pathways. There are uncertainties as to the complete role of the salvage enzymes cytidine deaminase (CDA) and deoxycytidylate deaminase (DCTD), functionally overlapping enzymes that maintain the cytidine:thymidine and cytidine:uridine ratios in the nucleotide pool and which also have activity against certain chemotherapy drugs.

This introduction describes the function of the nucleotide metabolism pathway related to DNA and RNA synthesis, subsequently focusing on nucleotide metabolism's role in preventing DNA damage as well as the role of enzymes such as CDA and DCTD play in chemotherapy resistance in cancer. Lastly, it discusses different experimental methods developed to study nucleotide metabolism and my strategy for probing the roles of CDA and DCTD.

1.2 The role of nucleotides in cell processes

Nucleotides are organic molecules with essential roles in cell replication, signalling and metabolism (**Figure 1.1**).

Figure 1.1 Nucleotide metabolism plays central role in cells. *Nucleotides are components in DNA (blue, thymidine monomer pictured); RNA (orange, uridine monomer pictured); enzyme cofactors (red, NAD+ pictured); cell signalling (pink, cAMP pictured); energy transport (green, ATP pictured).*

Deoxyribonucleic acid (DNA) encodes all genetic information in a cell and, together with ribonucleic acid (RNA), is responsible for storing and disseminating the information required to make proteins in a cell. DNA and RNA are polymers composed of nucleotides, with human DNA composed of over 6 billion monomers.[1]

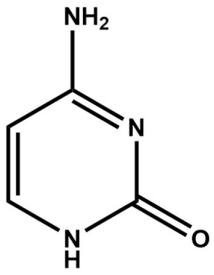
Similarly, cell makes use of nucleotide species. Adenosine and ATP are involved in purinergic signalling, while GTP functions as a second messenger interacting with G-proteins, and the cAMP-dependent pathway involved GTP, ATP and cyclic AMP (cAMP).

Cell metabolism employs nucleotides for a range of essential functions. Adenosine triphosphate (ATP), the most abundant free nucleotide, is the main cellular free energy carrier and is generated through cellular respiration. It drives central processes such as nucleic acid and protein synthesis, and muscle contraction. Guanosine triphosphate (GTP) is a specific

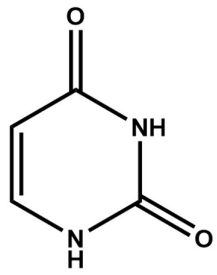
energy carrier for protein synthesis and gluconeogenesis. The metabolism cofactors nicotinamide adenine dinucleotide (NAD) and nicotinamide adenine dinucleotide phosphate (NADP) are involved in redox reactions.

Nucleotides are formed from the linking of a nucleobase, a ribose or deoxyribose sugar, and one to four phosphate groups. Nucleobases are the aromatic nitrogenous heterocycle families purines and pyrimidines (**Figure 1.2**). The pyrimidines are cytosine (C), uracil (U) and thymine (T), and the purines are adenosine (A), guanosine (G), with lesser amounts of the intermediates xanthosine (X) and hypoxanthine (H).

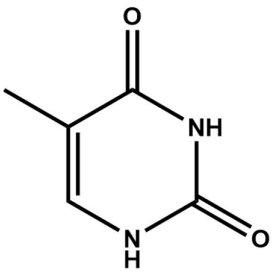
Nucleotides are structured through the attachment of N-9 of a purine, or N-1 of a pyrimidine, to C-1' of the sugar through an N-glycosidic linkage; C-5' of the sugar is phosphorylated. A nucleobase bonded to a sugar and with no phosphorylation is called a nucleoside.



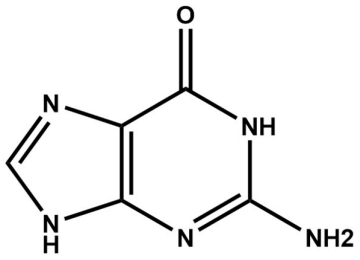
C



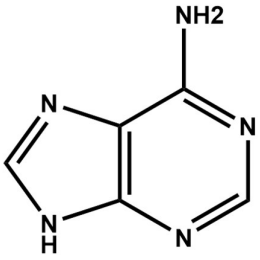
U



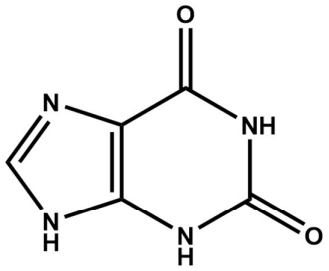
T



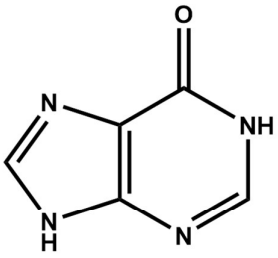
G



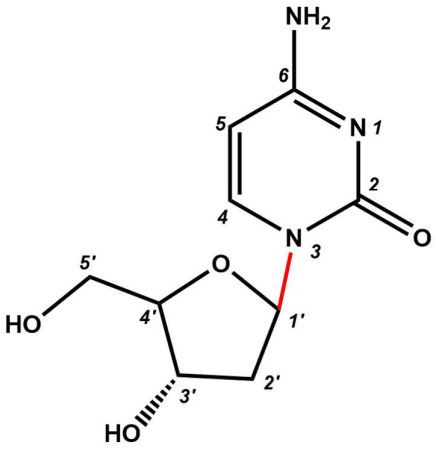
A



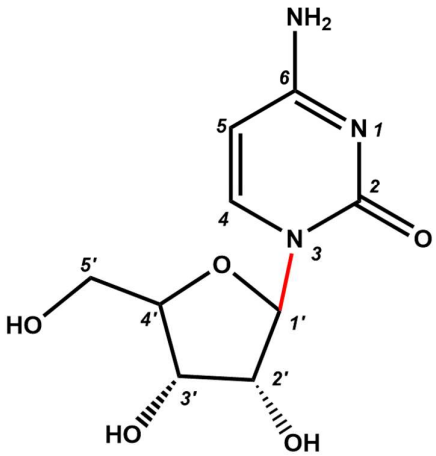
X



H



dC



rC

Figure 1.2 Illustration of the structures of the nucleobases C, U, T, G, A, X, H and the nucleosides dC and rC (N-glycosidic linkage in red).

1.3 Nucleotide base pairing underpins nucleic acid replication

The nucleic acids DNA and RNA are nucleotide polymers, linking the nucleotide monomers through phosphate bridges at the 5' and 3' end of the sugars (**Figure 1.3**).

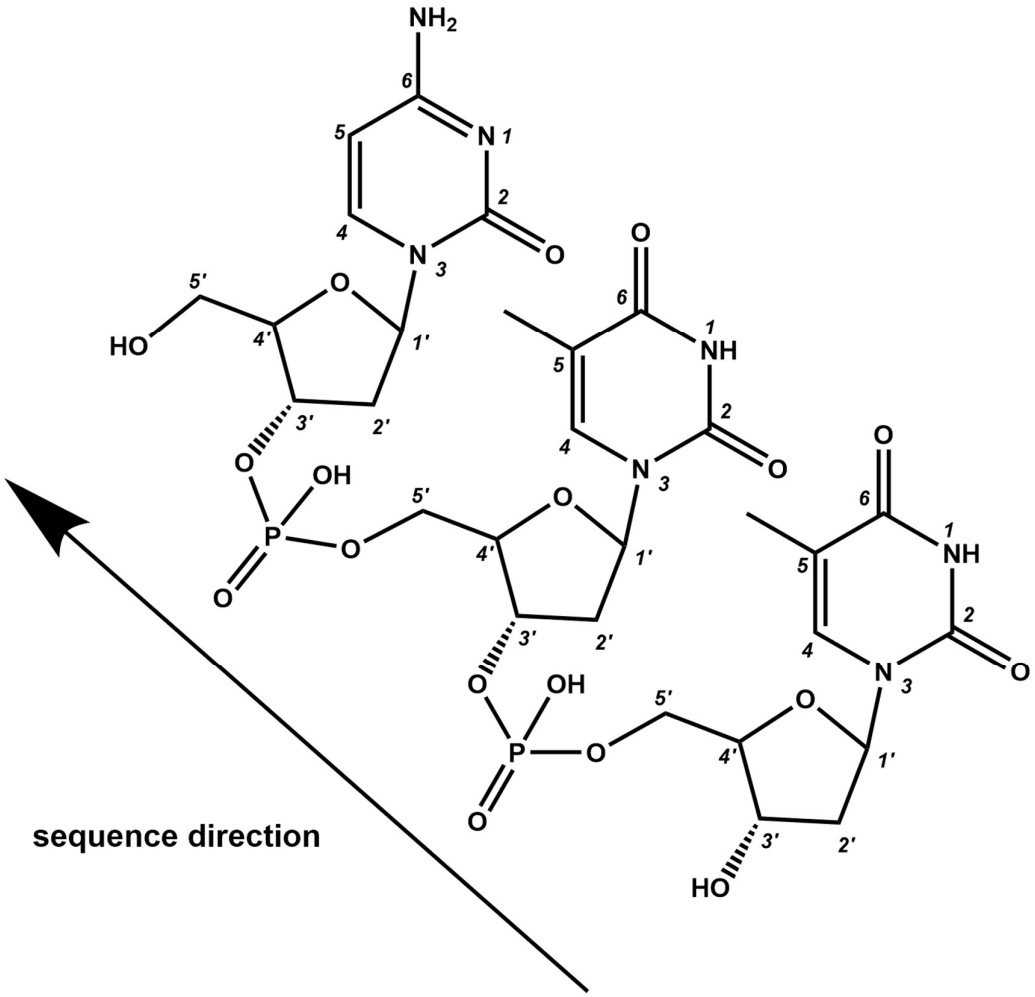


Figure 1.3 Reading sequence direction of DNA, in the nucleotide triad thymidine-thymidine-cytidine (TTC).

DNA is double-stranded, with the strands woven in a double helix through base pairs directed by complementary hydrogen bonding. In DNA, C pairs with G and A pairs with T, whereas U takes the place of T in RNA and pairs with A (**Figure 1.4**).

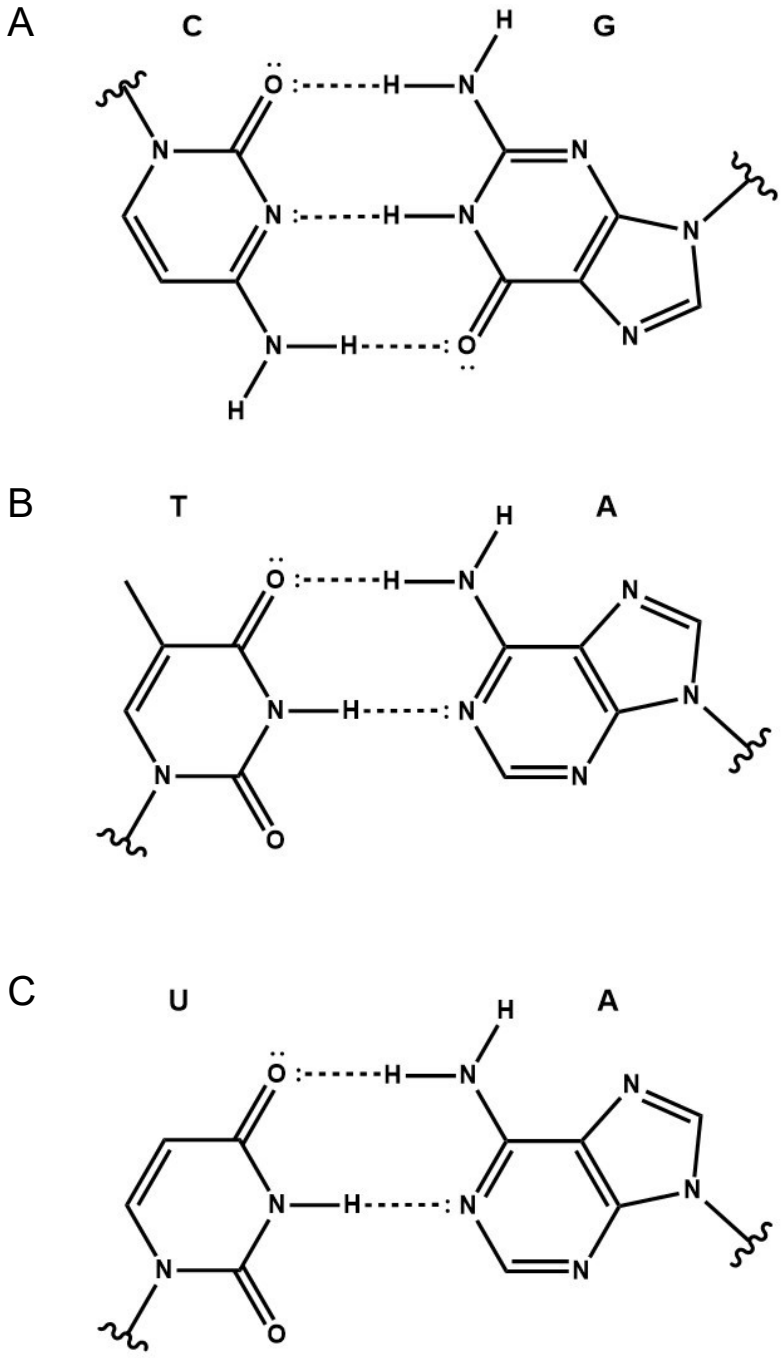


Figure 1.4 Base pairing in DNA and RNA. *Base pairing through complementary hydrogen bonding between nucleobases provides mechanism for DNA replication and transcription, and RNA translation.* ***A,*** *C:G base pairing (RNA and DNA);* ***B,*** *T:A base pairing (DNA);* ***C,*** *U:A base pairing (RNA).*

Base pairing is the key to DNA replication: when undergoing mitosis, the strands separate in a replication fork and are filled by DNA polymerases in a 5'—3' direction with sequential bases

chosen for hydrogen bonding complementarity. This, together with repair of replication errors, ensures that the daughter strands of DNA are near identical to the parent strands. The base pairing also underpins the relationship between DNA, RNA and proteins. The strand of DNA encoding a gene is transcribed into precursor mRNA (messenger RNA) by RNA polymerases in a transcription bubble, creating the nascent RNA from ribonucleotides with complementary base pairing to the DNA strand. The pre-mRNA is processed to mature mRNA and migrates to the ribosome. Here it is translated by tRNA containing complementary bases to the mRNA into an amino acid chain that folds into a protein.

1.4 Nucleotides are produced through *de novo* synthesis and salvage pathways

The nucleotide pool is in continual flux through enzymatic metabolism as well as spontaneous hydrolysis of nucleotides. Due to the greater turnover of RNA compared to DNA (and the roles the ribonucleotides ATP, ADP, GTP, GDP and cAMP play in signalling and metabolism) the ribonucleotide pool is between five and ten times more abundant than the deoxyribonucleotide pool—apart from ATP, which is nearly 100x more concentrated than dATP (2100 µM vs 2.4 µM).[2]

The essential functions of nucleotides require a reliable supply to cells; an imbalanced nucleotide pool can increase nucleotide misincorporation into DNA and RNA by polymerases, which can lead to cell mutation and death.[3–5] Most eukaryotes have two pathways for nucleotide acquisition: *de novo* synthesis and salvage.

In addition to *de novo* synthesis, cells can produce nucleotide triphosphates through the salvage pathway, generating nucleotide triphosphates from catabolised nucleic acids as well as importing nucleosides of digested nucleic acids from food and apoptotic cells. The salvage

pathway is the network of enzymes that enables the regeneration from nucleoside to nucleotide triphosphate and helps maintain the correct base ratio in the nucleotide pool.[6]

Salvage from food-derived nucleic acids begins in the stomach through their acid hydrolysis to oligonucleotides, aided through action of the peptidase enzyme pepsin—a recent and unexpected finding considering the role of pepsin as a protease.[7] The pancreatic enzymes ribonuclease 1 and deoxyribonuclease 1 digest the oligonucleotides to smaller nucleotide chains by the pancreatic enzymes.[8,9] The oligonucleotides are then catabolised to nucleotide monophosphates by the small intestine-specific nucleases and to nucleosides by alkaline phosphatase.[10] Following this, the small intestine villi actively absorb the nucleosides into the bloodstream across the intestinal brush border membrane.[11] The proportion of the nucleotide pool deriving from food sources has been investigated with tracer experiments, ranging from 2-5 % in rodents.[12] Nucleosides enter cells through the action of two classes of transporters: concentrative (CNT) and equilibrative (ENT). CNTs are encoded by the SLC28 family and are active transporters coupled to a transmembrane sodium gradient, creating a unidirectional flow of nucleosides into the cell.[13] The three CNT enzymes are CNT1 (SLC28A1), CNT2 (SLC28A2), and CNT3 (SLC28A3).[14–16] CNT1 has a higher selectivity for pyrimidines, CNT2 has a higher selectivity for purines and CNT3 transports both classes of nucleosides. All CNTs transport uridine.[17] ENTs allow nucleoside diffusion in both directions along concentration gradients. ENT1 (SLC29A1) and ENT2 (SLC29A2) transport a broad range of nucleosides and ENT3 (SLC29A3) is a mitochondrial transporter for intracellular nucleoside flux. ENT4, highly expressed in the brain, only weakly transports adenosine, with a larger role transporting cations.[18] There are other paths for nucleoside transport reported, such as the SLC15 peptide transporter family and the SLC22 organic ion transporter family. However, the CNTs and ENTs have ubiquitous tissue expression, while the SLC15 and the SLC22 families are confined to the liver, intestine and kidney.[19]

1.4.1 Purine nucleotide metabolism

Purine synthesis pathway

Purines are synthesised from the starting point of ribose-5-phosphate (**Figure 1.5**), derived from the pentose phosphate pathway. Ribose-5-posphate is converted to 5-phosphoribosyl-1- pyrophosphate (PRPP) by phosphoribosyl pyrophosphate synthase (PRPS).[20] Subsequently, the enzyme L-glutamine phosphoribosylpyrophosphate amidotransferase (GPAT) substitutes the pyrophosphate group for an amine at the C-1' position to form 5-phosphoribosyl-1-amine.[21] Glycinamide ribonucleotide synthase (GARS) then reacts 5-phosphoribosyl-1-amine with glycine and ATP to produce glycinamide ribonucleotide (GRA), which is then transformed to formylglycineamide phospho-ribonucleotide (FGAR) through the addition of formate and catalysed by GAR transformylase (GART).[22]

The enzyme phosphoribosylformylglycinamidine synthase (PFAS) substitutes the amide carbonyl for an imine with L-glutamine-derived ammonia, producing formylglycineamidine ribonucleotide (FGAM).[23] FGAM then cyclises through a 5-exo-trig cyclisation catalysed by aminoimidazole ribonucleotide synthase (AIRS) to form 5-aminoimidazole ribonucleotide (AIR).[22] The next step is catalysis by phosphoribosylaminoimidazole carboxylase (AIR carboxylase, encoded by PAICS) to cause the activation of bicarbonate with ATP and the imidazole ring attacking it to form carboxyaminoimidazole ribonucleotide (CAIR).

The enzyme phosphoribosylaminoimidazolesuccinocarboxamide synthase (SAICAR synthase encoded by PAICS) catalyses the attack of an aspartate molecule to the imidazole carboxylate to produce 5-aminoimidazole-4-succinylcarboxamide ribonucleotide (SAICAR).[24]

Subsequently, adenylosuccinate lyase cleaves SAICAR to generate fumarate and 5-aminoimidazole-4 carboxamide ribonucleotide (AICAR), then reacted to form 5-

formaminoimidazole-4 carboxamide ribonucleotide (FAICAR) in a process catalysed by 5-aminoimidazole-4-carboxamide ribonucleotide formyltransferase (AICARFT encoded by ATIC).[25]

The first purine nucleotide in the pathway, inositol monophosphate (IMP) is formed as a 6-exo-trig cyclisation product, catalysed by IMP cyclohydrolase (also encoded by ATIC, protein termed Bifunctional purine biosynthesis protein PurH).

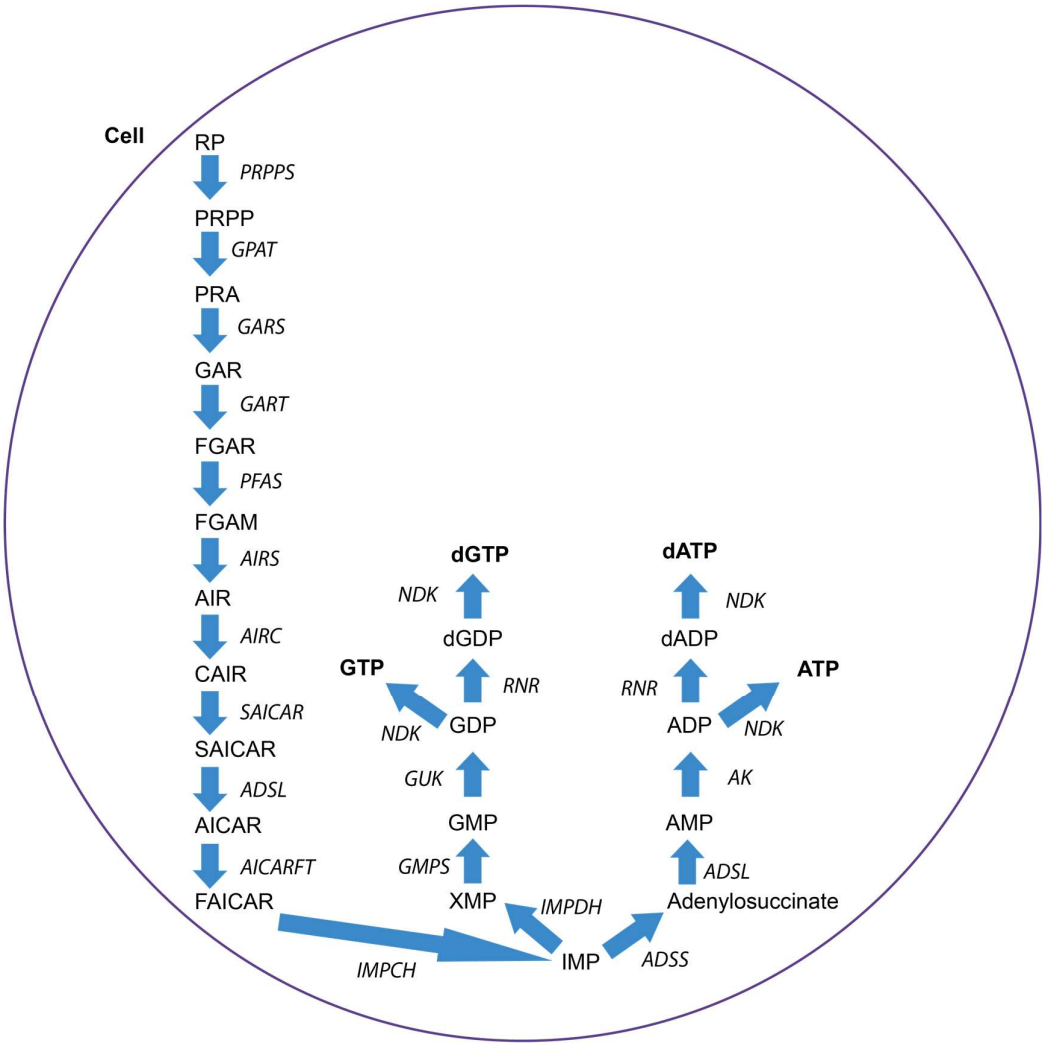


Figure 1.5 Human purine de novo synthesis pathway

IMP can undergo two different chains of reactions to form AMP and GMP. IMP is converted to adenylsuccinate by adenylsuccinate synthase (ADSS) and then into AMP by adenylsuccinate lyase (ADSL) through loss of fumarate.[25,26] To make GMP, IMP dehydrogenase (IMPDH)

oxidises IMP to XMP, which then reacts with ATP and L-glutamine to form GMP in a reaction catalysed by guanine monophosphate synthase (GMPS).[27,28]

AMP and GMP are phosphorylated to ADP and GDP by adenylate kinase (AK) and guanylate kinase (GUK1) respectively.[29,30] Ribonucleotide reductase (RNR) may then reduce them to dADP and dGDP.[31] ADP, GDP, dADP and dGDP are all phosphorylated by NDK to the final triphosphate molecules.[32]

Purine salvage pathway

Deoxyguanosine (dG) is phosphorylated by deoxyguanosine kinase (DGUOK) to dGMP, then phosphorylated by guanidylate kinase 1 (GUK1) to dGDP and finally by NDK to dGTP (**Figure 1.6**).[30,32,33]

Deoxyadenosine is similarly phosphorylated by deoxyadenosine kinase (AdK) to dAMP.[33] From there it is phosphorylated to dADP by the family of adenylate kinases (AK) and to dATP by NDK.[29,32] GMP is phosphorylated by guanidylate kinase 1 (GUK1) to GDP and from there to GTP by NDK.[30,32] AMP is phosphorylated by the AK family and to ATP by NDK.[29,32] As previously mentioned for purine synthesis, ribonucleotide diphosphates are substrates for RNR, a cell cycle dependent enzyme that reduces them to deoxyribonucleotide diphosphates. ADP and GDP are substrates and are reduced to dADP and dGDP, in turn phosphorylated to dATP and dGTP.[31,32]

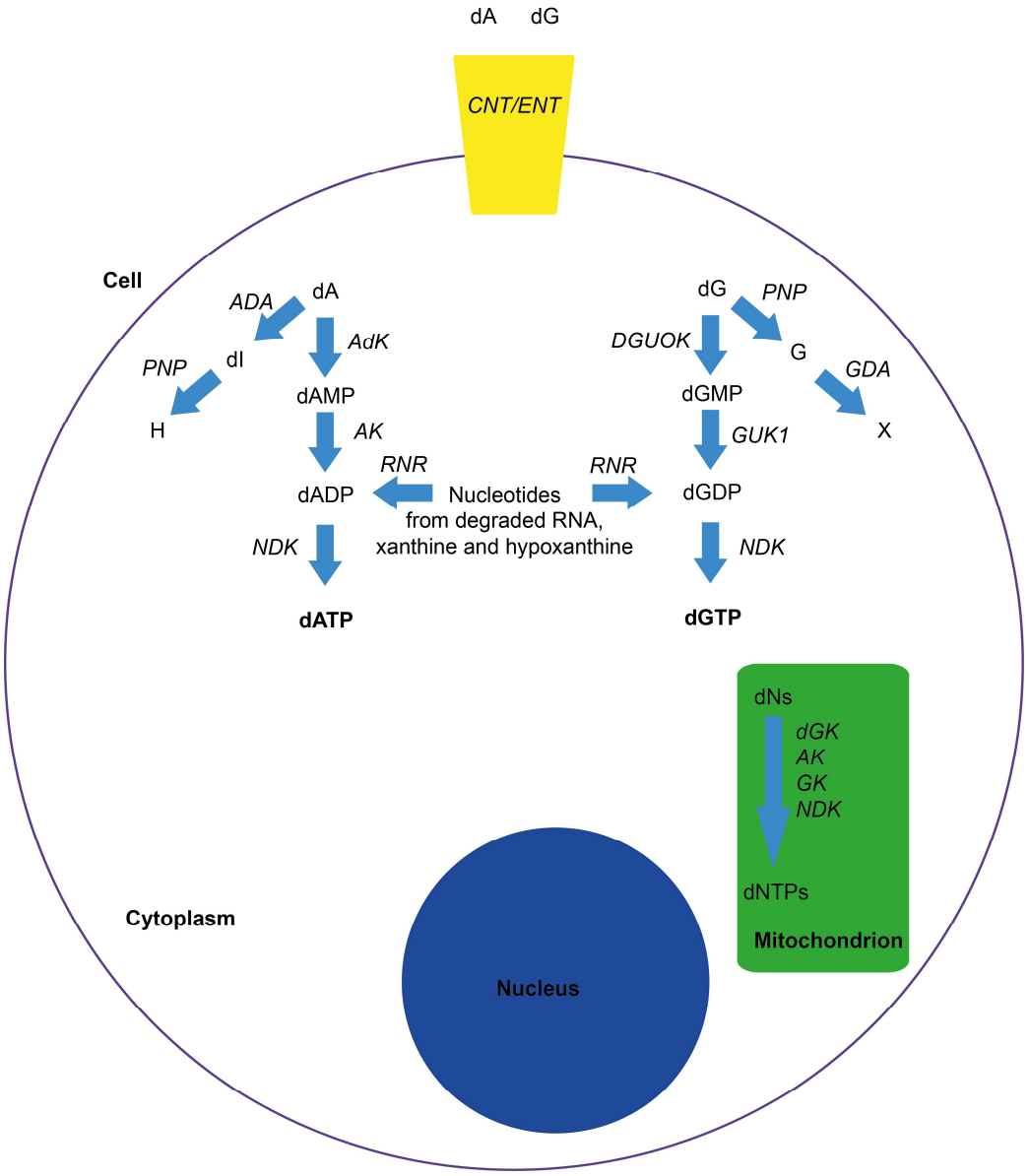


Figure 1.6 Human purine salvage pathway.

1.4.2 Pyrimidine nucleotide metabolism

Pyrimidine synthesis pathway

Pyrimidine *de novo* synthesis begins through the formation of the nucleobase (**Figure 1.7**). The trifunctional enzyme CAD (carbamoyl-phosphate synthase 2, aspartate

carbamoyltransferase, and dihydroorotase) has roles in the first three enzymatic steps of pyrimidine formation. Carbamoyl-phosphate synthase 2 reacts a hydrogencarbonate ion, L-glutamine and two ATP molecules to form carbamoyl phosphate. In parallel, aspartate aminotransferase converts oxaloacetate derived from the Krebs Cycle to an aspartate molecule. Aspartate and carbamoyl phosphate are then fused by aspartate carbamoyltransferase to form carbamoyl aspartate. Dihydroorotase then cyclises the molecule to form dihydroorotate in a condensation reaction.[34] Dihydroorotate is dehydrated to orotate by dihydroorotate dehydrogenase with coenzyme Q.[35] It is then converted to orotidine monophosphate (OMP) by uridine monophosphate synthase (UMPS), which also decarboxylates OMP to uridine monophosphate (UMP).[36]

The immediate pyrimidine precursors to DNA and RNA are deoxycytidine triphosphate (dCTP), thymidine triphosphate (TTP) and uridine triphosphate (UTP), all of which can be synthesised from UMP.

UMP is phosphorylated to UDP by CMPK1.[37] UDP can then be reduced to dUDP by ribonucleotide reductase (RNR) and phosphorylated to dUTP by a nucleotide diphosphate kinase (NDK).[31,32] dUTP is dephosphorylated to dUMP by dUTP pyrophosphatase (DUT)[38] and methylated to thymidine monophosphate (TMP) by thymidylate synthase (TYMS) with 5,10-methylenetetrahydrofolate.[39,40] dTMP is then phosphorylated to thymidine diphosphate (TDP) by deoxythymidylate kinase (DTYMK) and to thymidine triphosphate (TTP) by NDK.[32,41]

If not reduced by RNR, UDP can be phosphorylated to UTP by NDK, and then converted to CTP by the cytidine triphosphate synthases (CTPS 1 & 2).[42] CTP can be dephosphorylated to CDP by NDK, reduced to dCDP by RNR and phosphorylated by NDK to dCTP.[31,32]

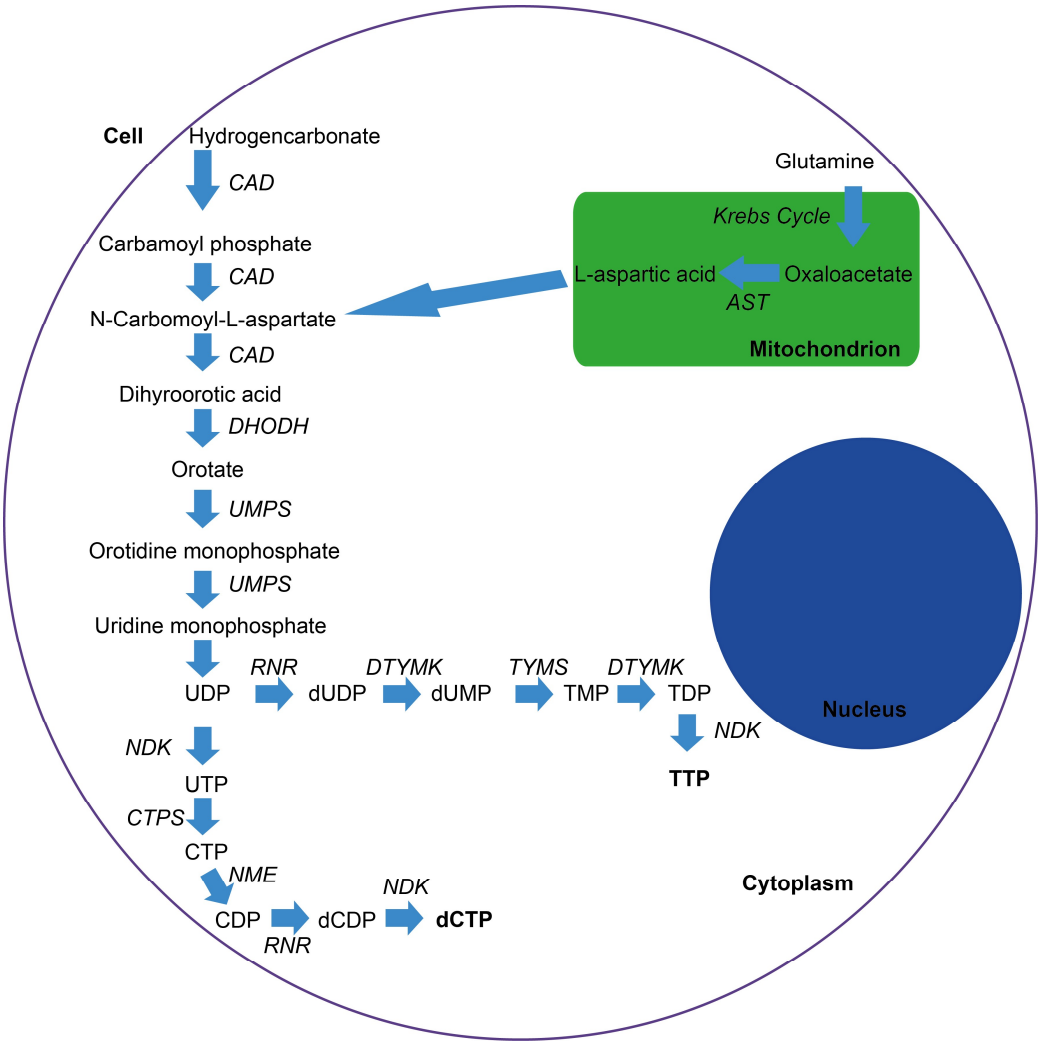


Figure 1.7 Human pyrimidine de novo synthesis pathway

Pyrimidine salvage pathway

Imported thymidine (dT) is salvaged through sequential phosphorylation: thymidine kinase 1 (TK1) phosphorylates dT to dTMP, in turn phosphorylated by thymidylate kinase (DTYMK) to dTDP (**Figure 1.8**). dTDP is phosphorylated to dTTP by an NDK enzyme.[32,37,41]

Deoxycytidine (dC) salvage occurs through two routes. In the first, deoxycytidine kinase (DCK1) phosphorylates dC to dCMP. That is then phosphorylated to dCDP by UMP-CMP kinase (CMPK1) and then to dCTP by NDK.[32,37,43] In the second route, dC and dCMP are deaminated

respectively by cytidine deaminase (CDA) and deoxycytidylate deaminase (DCTD) to deoxyuridine (dU) and deoxyuridine monophosphate (dUMP).[44,45] dU is phosphorylated to dUMP by TK1 and then TYMS methylates it to form dTMP.[33,40,46]

Ribonucleosides follow a similar pathway. Uridine (rU), is phosphorylated by uridine-cytidine kinases 1 and 2 (UCK1, UCK2) to UMP.[47] UMP is phosphorylated to UDP by CMPK1 and to UTP by NDK.[32,37]

Cytidine (rC) nucleotides are a substrate for CDA, though not for DCTD. It can therefore be deaminated by CDA to form rU, or it is phosphorylated by UCK1/2, CMPK1 and NDK to CTP.[32,37,47]

As seen for the ribonucleoside purines, the CDP and UDP can be reduced by RNR to dCDP and dUDP, converting them to deoxyribonucleotides.[31]

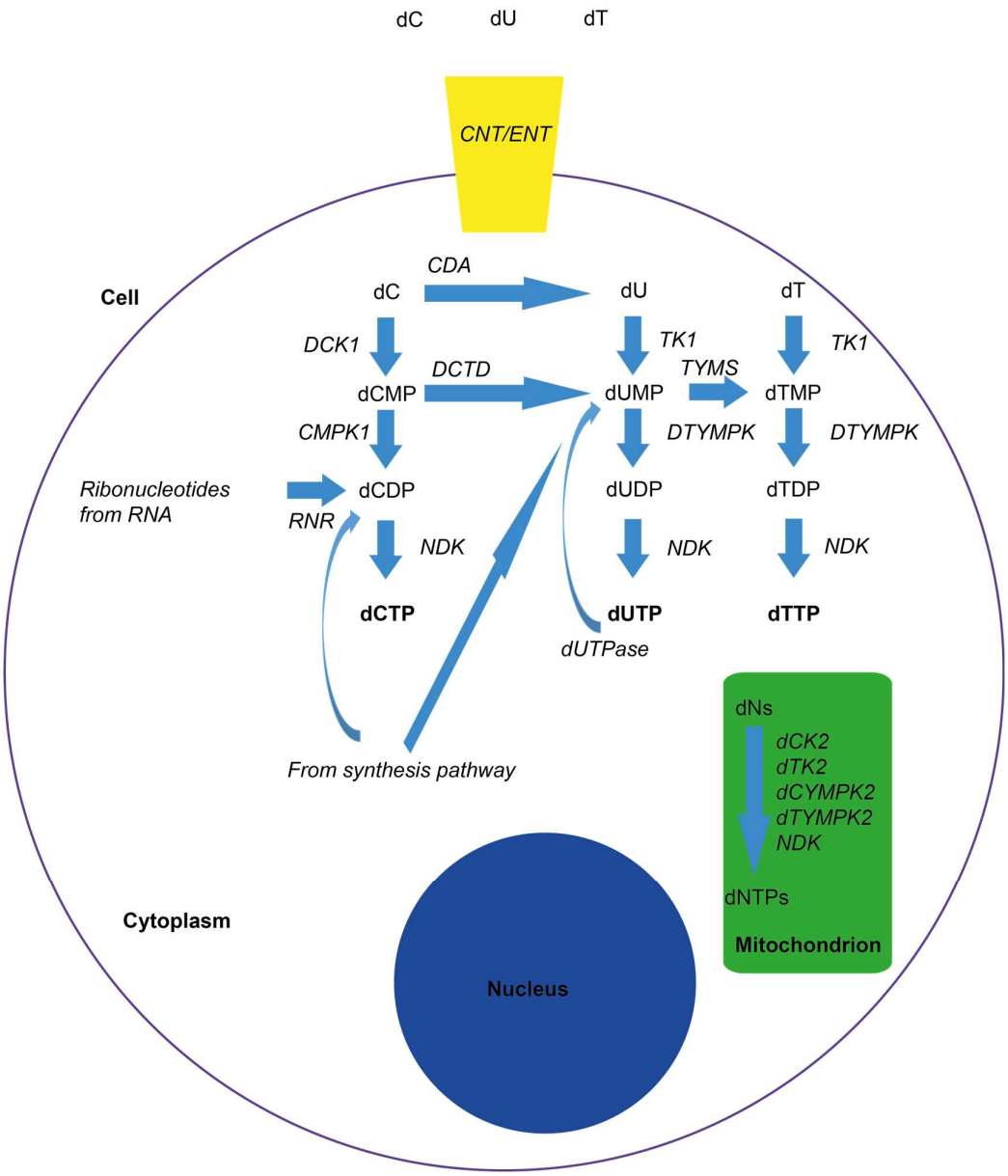


Figure 1.8 Human pyrimidine salvage pathway.

1.5 Nucleotide depletion to avoid pool imbalance and to remove damaging nucleotides from pool

Nucleotide demand is periodic; while there is continual RNA synthesis through the cell cycle, DNA synthesis—the primary consumer of deoxyribonucleotides – only occurs in the

S-phase.[48–50] The abundance of deoxyribonucleotides is controlled through variation in enzyme levels. The protein level of the enzyme complex required for deoxyribonucleotide synthesis, RNR, is periodic with cell cycle and rises in S-phase.[51] RRM2, a component of RNR is degraded during G2 phase to avoid overaccumulation of dNTPs and errors in this regulation of RRM2 are associated with a higher mutational burden due to an overly large deoxyribonucleotide pool.[52] In parallel, it is reported that the levels of SAM domain and HD domain-containing protein 1 (SAMHD1), which converts dNTPs to dNMPs, are high in G1 phase and low in S-phase, causing a periodic change in dNTP pools.[53] This mechanism is used in dendritic and myeloid-derived cells infected with HIV to stop the virus replicating.[54]

Pyrimidines can be degraded and removed from the nucleotide pool by catabolism, starting with the action of dihydropyrimidine dehydrogenase (DPYD) on uracil, an enzymatic chain that leads to the formation of beta-alanine.[55] The purine xanthine, a downstream metabolite of all purines, is catabolised to uric acid by xanthine oxidase (XO).[56]

Enzymes in the nucleotide metabolism pathways play a role in preventing DNA damage from occurring by removing damaging nucleotide triphosphates before they are incorporated into DNA. dUTP, a product of the synthesis and salvage pathways, can be readily incorporated into DNA in place of dTTP opposite dATP. The nucleotide pool sanitation enzyme dUTP diphosphatase (DUT) reduces this potentially limiting dU incorporation by dephosphorylating dUTP to dUMP, concomitantly increasing the precursor for dTTP, whereas loss of the pyrophosphatase dCTPP1 has been reported to increase the dUTP/dTTP ratios and lead to dU incorporation, providing evidence for its degradation of dUTP.[57–59] 2-hydroxy-dATP diphosphatase (MTH1) catalyses the dephosphorylation of the oxidised purine nucleotides 2-hydroxy-dATP and 8-oxo-dGTP to their respective monophosphates.[60] The bases 5-hydroxymethyl-2'-deoxycytidine (hmdC) and 5-formyl-2'-deoxycytidine (fdC), formed through epigenetic processes and obtained from degraded DNA or from food, are

phosphorylated by dCK, but the monophosphates are not further phosphorylated by CMPK1. DCTD has been reported to deaminate hmdCMP into hmdUMP, which is then degraded by DNPH1 (2'-deoxynucleoside 5'-monophosphate N-glycosidase).[61] These processes prevent modified base incorporation into DNA, where they could induce unwanted epigenetic effects.[62]

Damaged nucleobases such as 8-oxoG exhibit different hydrogen bonding and engage in different base pairing to their undamaged parent bases, with mutagenic potential if incorporated.[63] In the case of 8-oxoG, it forms a base pair with A, instead of the natural base pair of G, C. Uracil can also form in DNA through hydrolytic deamination of dC, changing a C:G base pair to U:G. Cytosine deamination to uridine in DNA is a leading cause of genetic lesions and will create potentially lethal or cancerous mutations in the cell.[64,65]

1.6 Removal of incorrect bases in DNA through repair

Enzymes in the nucleotide metabolism pathway such as dUTPase have evolved to minimise the incidence of incorrect nucleotide incorporation into DNA. However, this still occurs, either due to an imbalanced nucleotide pool, or through polymerase errors.

If there are errors in DNA replication without subsequent repair, the genome undergoes mutations to its sequence with consequences for gene transcription, enzyme functionality and cell survival (**Figure 1.9**). Not all mutations are damaging to a cell: the incidence of mutations helpful to survival is what drives natural selection and evolution.

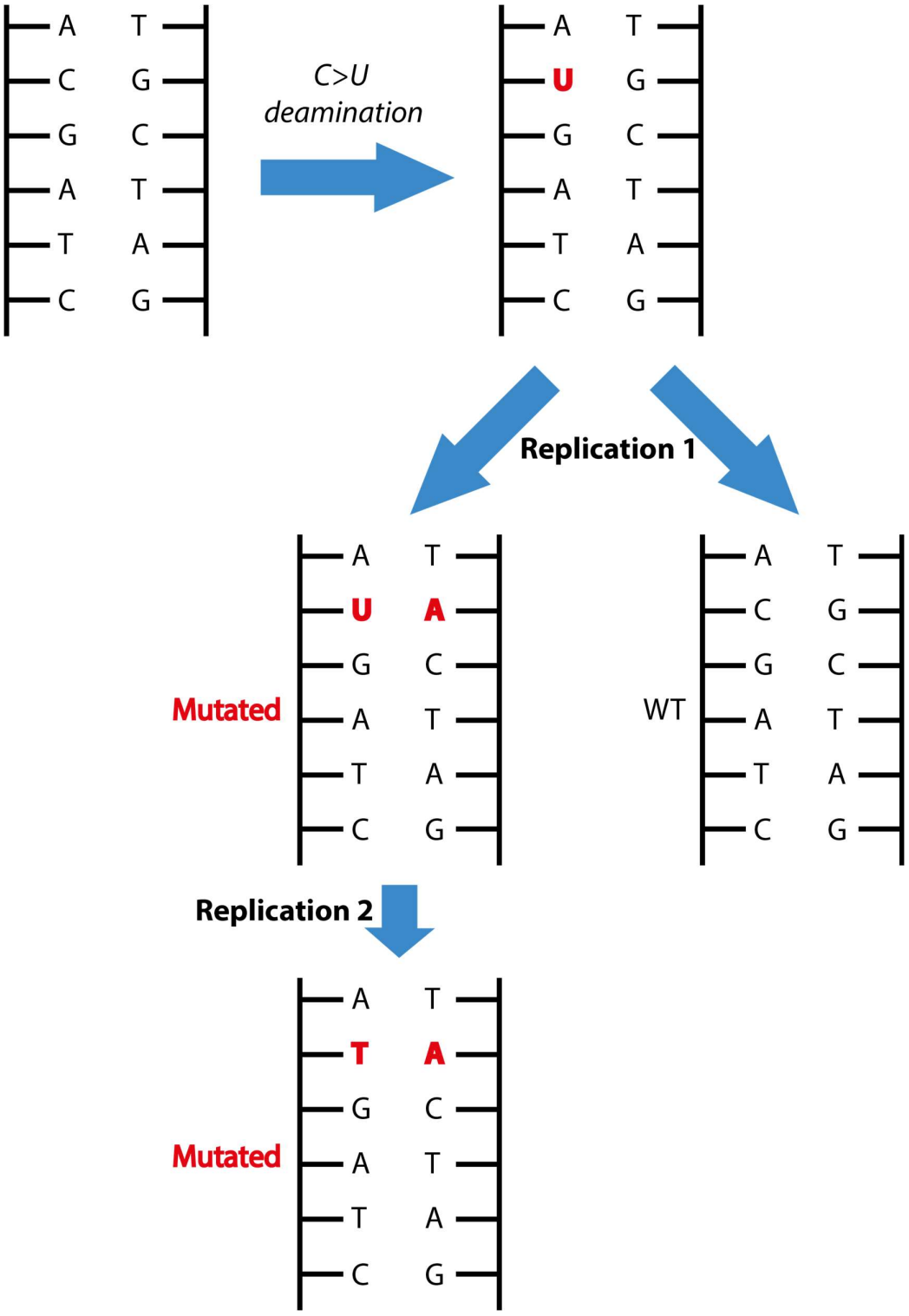


Figure 1.9 DNA Chemical damage to a base can lead to mutation in daughter cells. *Deamination of C forms a U base. If not removed, upon the next replication it will pair with an A. This leads to a C->T sequence mutation.*

High mutations in a particular cell increases the likelihood of deactivating mutations in tumour suppressor genes such as P53 and DNA repair pathways, as well as activating mutations in oncogenes such as the Ras family.[66–68] this could then lead to rapidly dividing and aggressive cell populations and ultimately cancer.

Excessive mutation is avoided through the action of the cell's DNA repair mechanisms that remedy DNA lesions caused either by nucleotide misincorporation or DNA damage.

DNA Mismatch repair (MMR) excises incorrectly incorporated bases from newly formed DNA. This works to remove incorrect bases from the newly synthesised daughter strand that occur through polymerase errors.[69,70]

In addition, base excision repair (BER) removes damaged or mutated bases from DNA that could cause mutations by mismatch; there are several uracil-DNA glycosylases (UDGs) that remove deaminated and oxidised bases from DNA through the creation of abasic sites, with replacement with the correct base enzyme 8-oxoG DNA glycosylase (OGG1) removes 8-oxoG from DNA.[71–74]

1.7 Role of nucleotide metabolism in cancer

Cancer describes the diseases that show uncontrolled cell proliferation, causing organism damage and sometimes death through disruption of critical tissue function. Its causes are various and tissue specific but are often associated with deletion or deactivation of tumour suppressor genes through mutation or epigenetic silencing, or activation of oncogenes that promote uncontrolled growth and metabolism.[66,67]

The increased cell division in cancer leads to a higher DNA and RNA turnover, with a consequently greater demand for nucleotides.[75] Genes upstream of nucleotide metabolism

are often upregulated in cancer cells: the oncogene cMYC, a transcription factor which increases expression of a broad range of genes, upregulates genes in the nucleotide synthesis pathway, while the protein kinase mTor increases pyrimidine synthesis gene expression.[76–78]

Enzymes involved in the MMR pathway are frequently affected in cancer. A heterozygous germline mutation in the mismatch repair enzyme MutL homolog 1 (MLH1) is a risk factor for hereditary nonpolyposis colorectal cancer (HNPCC), while MLH1 has been reported to be inactivated in colorectal cancer through sporadic somatic mutations and epigenetic inactivation.[79–81]

1.7.1 Mechanisms of chemotherapy in cancer nucleotide metabolism

DNA replication, crucial for cancer cell proliferation, requires a balanced pool of deoxyribonucleotide triphosphates;[82] faults in nucleotide metabolism can therefore be lethal to rapidly dividing cells. Consequently, cancer's nucleotide synthesis and salvage pathways have been a frequent target for chemotherapy by varying mechanisms of action.

Disrupting nucleotide metabolism and unbalancing the dNTP pool has been exploited as a treatment strategy. Thymidylate synthase (TYMS) methylates dUMP to dTMP together with the methyl donor methylenetetrahydrofolate, an essential step dTTP synthesis. The uracil analogue 5-fluorouracil (5FU) was developed to inhibit this step. 5FU is salvaged as a uracil analogue and its phosphorylated metabolite stably binds and inhibits TYMS, depleting the dTTP pool.[83,84] Additionally, the dNTP pool can be depleted by inhibition of RNR, the enzyme responsible for reduction of ribonucleotides to deoxyribonucleotides. The drugs gemcitabine (2',2'-difluorodeoxycytidine) and fludarabine (fluorinated adenine arabinoside) are successively phosphorylated to form the respective diphosphates, which inhibit RNR and

diminish the dNTP pool, and also compete with dCTP and dATP for DNA incorporation. They induce a chain termination step and inhibit DNA synthesis, leading to cell death.[85,86]

Nucleotide analogue chemotherapy can also lead to cell death through disruption of DNA synthesis. In addition to nucleotide pool disruption, gemcitabine can be sequentially phosphorylated to gemcitabine triphosphate, after which it is incorporated into DNA and prevents further replication through polymerase stalling, in a similar manner to fludarabine triphosphate.[86,87]

Another angle of therapy is to introduce nucleoside analogues that induce epigenetic changes and consequently disrupt the gene expression levels in cells. The deoxycytidine analogue decitabine is phosphorylated to decitabine triphosphate and incorporated into DNA, where it covalently binds to the epigenetic marker DNMT1. At high concentrations this causes DNA damage, while at low concentrations its mechanism of action is through the reduction of active DNMT1. DNMT1 is responsible for replicating cytosine methylation to the daughter strand and its reduction leads to a reversal of the CpG island hypermethylation seen in many cancers, reactivating tumour suppressor genes and inducing an antineoplastic effect.[88–90]

1.7.1.1 Salvage and sanitation pathways in chemotherapy resistance

Due to its high replication, mutation from errors in polymerisation and differences in epigenetic profile between cells in different microenvironments, the community of cells that form the tumour are highly heterogeneous.[91] As a result, it can exhibit resistance to nucleoside analogue chemotherapy, often occurring through alteration of the nucleotide metabolism pathway. The common modes of resistance occur through several ways.

The nucleotide analogue compound may be biochemically deactivated; the deaminase CDA deaminates the cytidine analogues gemcitabine, decitabine, cytarabine and their

monophosphorylated metabolites to less cytotoxic ones.[92–94] DCTD deaminates cytarabine monophosphate and gemcitabine monophosphate to the uridine analogues. However, the data about DCTD function in cancer resistance mechanisms is conflicting. Some reports state that high DCTD renders the drug innocuous and others that it potentiates the anti-tumour effect, leading to ara-UTP accumulation, while another group reported that DCTD is not predictive of patient response to gemcitabine therapy. [94–98]

Certain nucleotide analogues can also be deactivated through pyrimidine degradation. The enzyme dihydropyrimidine dehydrogenase (DPYD) degrades uracil to dihydrouracil, the rate-limiting step in pyrimidine catabolism. It degrades 5FU and high expression levels of DPYD have been correlated with poor response to 5FU.[99,100]

Another cause of resistance is an increase in the protein level of the enzyme targeted by chemotherapy; it has been reported that in studies of cancer patients, there is an inverse correlation between TYMS expression and protein level, and response to 5FU treatment. The increased protein level may increase the proportion of uninhibited TYMS and negate the effect of 5FU.[83,101,102]

Finally, cell export proteins also play a role in nucleotide analogue chemotherapy resistance, with expression of the nuclear transporter ABCC5 reported to be inversely correlated with the activity of gemcitabine in non-small cell lung cancer cell lines.[103]

In summary, different phenotypes of cancer cells can enable resistance to nucleotide analogues, through routes including the upregulation of target enzymes, upregulation of enzymes involved in the sanitation and degradation of nucleotide analogues, and increased expression of cell export proteins.

1.7.1.2 CDA creates vulnerability to epigenetically modified nucleotides

The deaminase CDA deaminates cytosine nucleoside variants to the uracil equivalent (**Figure 1.10**).[104] As described in **1.7.1.1**, its overexpression is a prominent cause of chemotherapy resistance, deaminating gemcitabine, decitabine and cytarabine to deactivate uridine analogues.

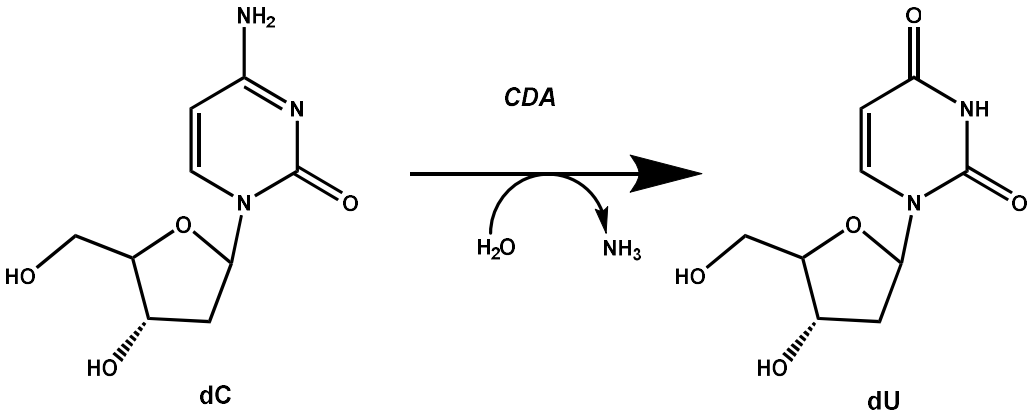


Figure 1.10 Deamination of dC to dU by CDA. *CDA also deaminates rC, and also cytidine analogues used in chemotherapy such as gemcitabine.*

However, high CDA expression makes cells vulnerable to other classes of pyrimidine analogue chemotherapy. For instance, a downstream metabolite of the drug capecitabine (Xeloda) is deaminated by CDA to eventually produce 5FU, an active TYMS inhibitor. Epigenetically modified cytidines are also damaging to a cell with high CDA expression. The epigenetic nucleosides hmdC and fdC can be phosphorylated to hmdCMP and fdCMP by DCK, but these cannot be phosphorylated to hmdCDP and fdCDP by CMPK1. CDA deaminates the nucleosides to hmdU and fdU, which can be phosphorylated to hmdUTP and fdUTP. These are incorporated into DNA and lead to damage once they are recognised by UDGs as uracil and excised, and are lethal to cells with high CDA expression[62].

1.8 Roles of pyrimidine deaminases CDA and DCTD

The vulnerability to nucleotide analogues that CDA overexpression creates and the presence of DCTD as another deaminase enzyme's raise questions as to the evolutionary advantage of both of their activities. CDA DCTD perform mutually redundant functions, deaminating cytidine nucleotides to uridine nucleotides, as well as deaminating cytidine analogue chemotherapeutic agents and the epigenetic nucleoside 5-methyl-2'-deoxycytidine (mdC).[93,105–107] The deamination process has a dual role: the first converts cytosines into uridine/thymine nucleotides (**Figure 1.10**) and the second is to convert cytidines to uridines for pyrimidine degradation by DPYD as a supply of carbon and nitrogen to the cell.[55,104]

CDA and DCTD are evolutionarily conserved enzymes across various kingdoms of life.[108] CDA is the principal source of dUMP for *Trypanosoma Brucei*, a protozoan parasite with no DCTD and a weak *de novo* pyrimidine synthesis pathway[109]. CDA (EC 3.5.4.5) and DCTD (EC 3.5.4.12) are both 15 kDa sized cytoplasmic enzymes in the pyrimidine salvage pathway and their active domain are a CMP (cytidine monophosphate)/dCMP (deoxycytidine monophosphate)-type deaminase domain.[110] CDA is tetrameric and DCTD is hexameric, while both bind a zinc ion at their active sites.[111–113] They have tetrahedral active sites (**Figure 1.11**), where the zinc ion lowers the energetic barrier for nucleophilic attack of a water molecule against the cytosine nucleobase substrate and results in the release of an ammonia molecule and the formation of a uridine nucleotide.[112,114]

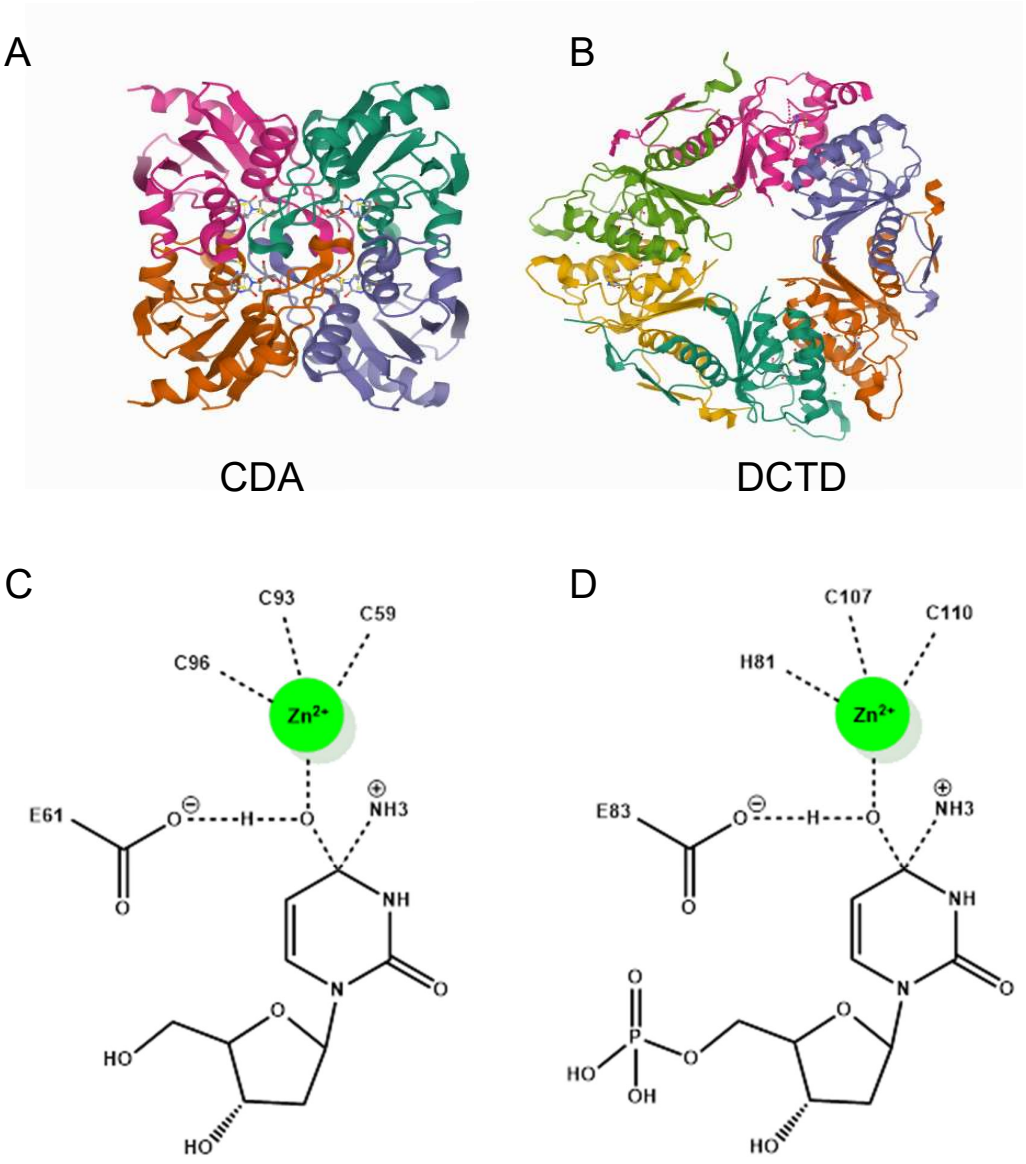


Figure 1.11 Structures and active sites for the deaminases. *CDA is tetrameric and DCTD is hexameric. Both incorporate zinc as a participating ion. Crystal structures obtained from Protein Data Bank.* ***A,*** *structure of CDA;* ***B,*** *structure of DCTD;* ***C,*** *deamination of dC by CDA, showing active site of CDA bound to dC;* ***D,*** *deamination of dCMP by DCTD, showing active site of DCTD bound to dCMP.*

CDA and DCCTD are detectable in all human tissues, with CDA having the highest mRNA levels in blood cells, the spleen and the liver and pancreas, while DCTD expression is more uniform, with high levels in the thyroid and blood (**Figure 1.12**).[115–119]

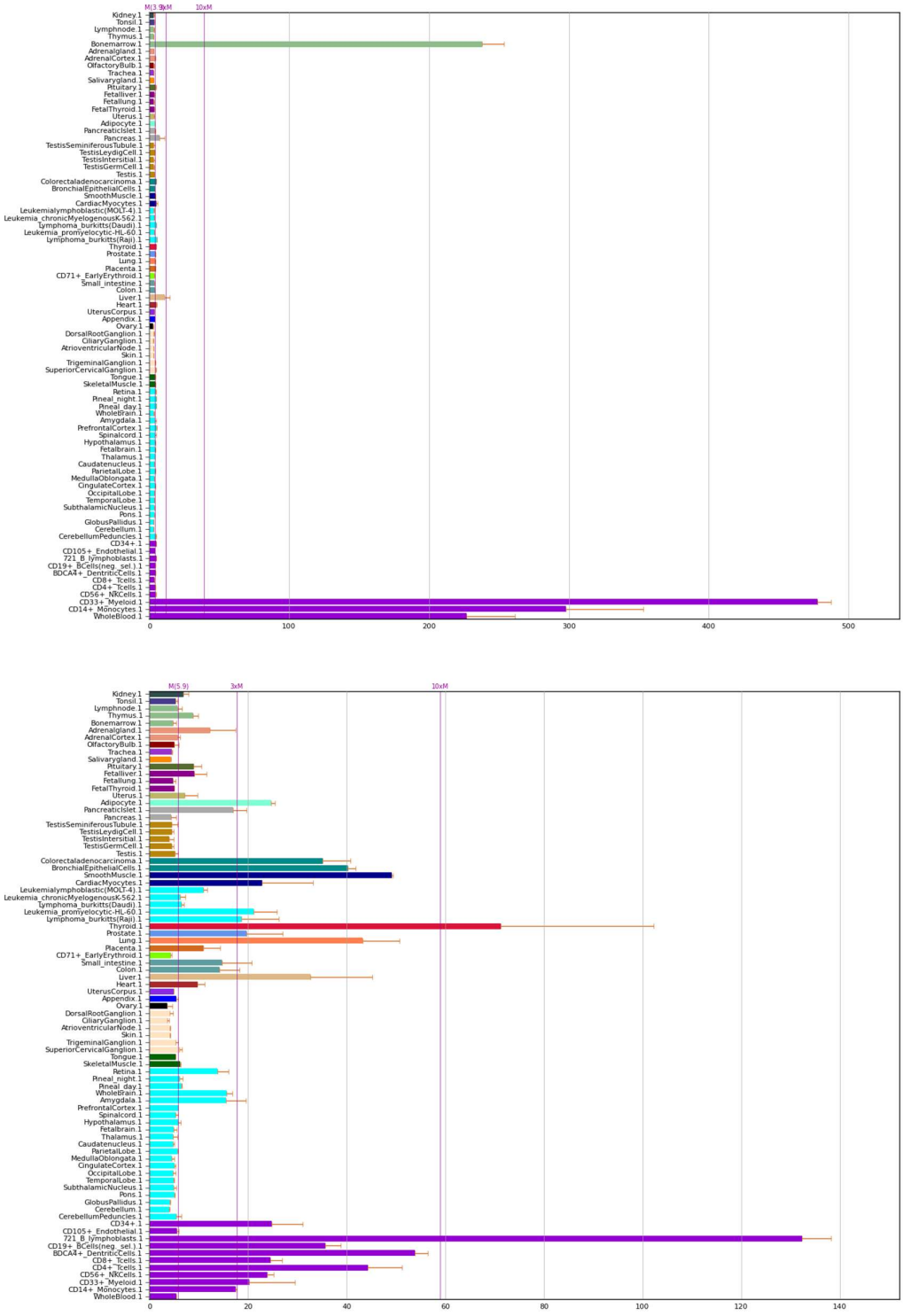


Figure 1.12 DCTD is more uniformly distributed in human tissues than CDA. *Tissue-specific mRNA levels, obtained through BioGPS.* ***A,*** *mRNA tissue map of CDA;* ***B,*** *mRNA tissue map of DCTD.*

In addition to the effects of high CDA expression in cancer resistance and vulnerability to chemotherapy (**1.7.1.1** and **1.7.1.2**), deficiency of functional CDA through downregulation or genetic deletion/inactivation may cause disease. CDA deficiency has been reported in patients with Bloom Syndrome, due to the non-functional BLM DNA helicase.[120] The study claims that the low level of CDA reduces deamination and increases the concentration of cytidine nucleotides, causing the nucleotide pool to imbalance at a mutagenic level. The C:T imbalance triggers nucleotide misincorporation and faulty sister chromatid DNA segregation and the formation of ultra-fine anaphase bridges (UFCs) through PARP-1 inhibition by accumulated dCTP.[82,120]

The wide-ranging CDA and DCTD expression in different clinical presentations has led to research on the enzymes' regulation. DCTD has been shown to be allosterically inhibited by dTTP and activated by dCTP, with a smaller pool of evidence for CDA allosteric regulation.[121,122] Bloom Syndrome patients have downregulated CDA expression through a currently unknown mechanism involving CDA downregulation of the BLM helicase. There have also been transcription factor binding sites reported in CDA's promoter and enhancer regions, as well as microRNA downregulation and upregulation.[123–127]

Small molecule inhibitors have been developed to study CDA and DCTD. CDA is inhibited by the small molecules tetrahydrouridine (THU), zebularine and cedazuridine,[128–130]whereas DCTD is inhibited by the monophosphate of zebularine and weakly by gemcitabine monophosphate.[131–133]A benefit of small molecule inhibitors for CDA and DCTD is that they have the potential to reduce deamination of chemotherapeutic cytidine analogues through resistance or drug metabolism, potentiating the therapy response. For instance, cedazuridine has been combined with decitabine as a treatment for chronic myelomonocytic leukemia (CMML) to increase the bioavailability decitabine, by inhibiting CDA in the gut and liver.[130]

The roles CDA and DCTD mutations in their sequences play in enzymatic activities have been investigated. CDA has two common germline polymorphisms due to a prominent mutation (79A>C, K27Q), giving either a lysine or L-glutamine. The two single nucleotide polymorphisms (SNPs) are both functional, with K27Q found most commonly among white ethnicities.[134] The variant 208G>A, A70T is also predominantly found among Asian and African ethnicities.[135–137] The A70T mutant is less active for cytidine and cytarabine while K27Q is more active for deoxycytidine and cytidine.[135,138] There are conflicting reports as to the differences in activities between the SNPs for cytidine analogues: Micozzi et al. report K27Q less active with cytarabine as a substrate, while Kirch et al. find it to be more active.[138,139] Farell et al. found the K27 variant to give a longer overall survival time to gemcitabine patients, while marginally increasing toxicity.[140] One rare mutation, 435C>T, T145T, has been linked to acute toxicity from gemcitabine in pancreatic cancer patients, pointing to a deactivating role.[141] DCTD has three known SNPs (-47T>C, V116V; 255G>C, Ala/Ala; 315T>C, Val/Val), but they have not been reported to affect overall survival, specifically -47T>C for pancreatic cancer patients undergoing gemcitabine therapy, and 255G>C and 315T>C for breast cancer patients.[142,143]

1.9 Methods of studying nucleotide metabolism

Interest in nucleotide metabolism's role in disease has led to the development of tools for its investigation. Nucleotide metabolism can be probed by disrupting the function of an enzymatic in nucleotide metabolism and observing the resultant change in metabolite composition through analytical techniques.

Additionally, the activity of an enzyme in the nucleotide metabolism pathway can be investigated by exposing an organism to a labelled nucleotide precursor that is then

metabolised by the enzyme. The labelling allows monitoring of the enzyme's effect on the precursor.

1.9.1 Tools for enzyme manipulation

There have been various tools developed to investigate roles of enzymes, which can be applied to nucleotide metabolism. One strategy is through use of small molecule inhibitors, which bind to and disrupt enzymatic activity. The disadvantages of inhibitors is that they may be non-target specific and can produce unwanted side effects that may be lethal or skew metabolism; for instance, the CDA and DCTD inhibitor zebularine also reduces cytosine methylation levels by downregulation of the epigenetic methyl transferases (DNMT).[144,145]

Non-inhibitor tools include protein manipulation techniques include tools to increase protein expression such as through transient or stable upregulation of genes encoding enzymes of interest, tools to suppress gene transcript translation into protein such as short hairpin RNAs (shRNAs) and small interfering RNAs (siRNAs), and the mutation of a gene to induce loss of function in the protein. The advent of precise gene editing through CRISPR Cas technology ("Clustered Regularly Interspaced Palindromic Repeats" and "CRISPR-Associated protein 9") has increased the ease of knocking out or mutating enzymes.[146] Catalytically dead Cas9 (dCas9) has also been used to achieve tuneable suppression and activation through its fusion with domains of transcription factors and repressors.[147–149] The use of conditional expression and knockouts has given information on the rapid dynamic changes in the metabolism from gene manipulation, as well as make experiments possible *in vivo* in tissue-specific contexts where a constitutive enzyme perturbation could be embryonically lethal.[150,151]

Stably implementing genetic changes has drawbacks. For instance, integrating an external DNA cassette into genomic DNA may cause insertional mutagenesis if it is integrated in a

region that will disrupt function of another gene.[152,153] Knocking out genes (KO) through restriction or CRISPR-Cas systems may disrupt through off-target effects or permanently skew nucleotide metabolism through selection of faster growing cells.[154]

By contrast, transient KD through siRNA transfection requires no DNA integration and risk of mutation. siRNAs cause a rapid and transient knockdown of the target gene, with their effect reported to last for several days.[155]

Gene manipulation can be performed in cell lines, cells obtained from organisms and immortalised to replicate indefinitely. Cell lines are derived from primary tissue and have a specific expression of genes influenced by factors including their culture medium, their mutations and the tissue and patient of origin.[156,157] They are useful tools to probe the impact of enzymes, but reliance on a single cell line leaves the research conclusions vulnerable to cell-specific features not related to the enzyme in question. Conversely, it is difficult to experiment on many different cell lines, as it would create handling and sample analysis constraints.

1.9.2 Nucleotide metabolism analysis

Nucleic acid metabolism is probed by measuring the free nucleotide pool or analysing the composition of nucleic acids, which are a product of free nucleotide metabolism.

1.9.2.1 Nucleotide metabolism analysis: the free nucleotide pool

The free nucleotide pool includes every nucleotide variant and is a subset of the collection of small molecules dubbed the metabolome. There has been keen interest in measuring the

metabolome to research its role in disease and development, as not only is it the end product of networks thought of as upstream—the genome, transcriptome and the proteome—but it has been shown to directly impact the previous processes through errors induced by imbalanced nucleotide pools, as well as precursors for epigenetic marks and protein post-translational modifications.[57,69,70,81,82,158]

The challenges in metabolome analysis arise from its size—there are over 8,000 detected endogenous metabolites, with many more predicted.[159] This large number makes it challenging to analyse metabolite changes.

One strategy for analysing a subset of metabolites is through radiolabelling. Radiolabelling involves organism exposure to a precursor containing a radioactive isotope, such as ^{32}P, ^{35}S or ^{3}H. The radiolabelled precursor, being an isotopologue of the unlabelled precursor and chemically the same, is metabolised by the organism into different metabolites and biomolecules. The material of interest is isolated and its radioactivity is measured, from which one can infer the extent of labelling and the activity of a metabolic pathway. For example, Stevenson-Paulik et al. describe a ^{32}P labelling method to study kinase and phosphatase dynamics in inositol phosphate metabolism.[160] The drawbacks of radiolabelling are that only a small subset of the metabolome is visible, one must take personal safety precautions against the radioactivity, and it is possible that the organism under study may suffer from radioactive effects, reducing the validity of the experiment conclusions.[161]

Another technology often used in metabolomics is nuclear magnetic resonance (NMR). NMR makes use of the quantum-derived magnetic moments in certain nuclei such as ^{1}H, ^{13}C and ^{31}P and their behaviour in a strong magnetic field to identify multiple analytes at the same time.[162–164] The advantage of NMR is that it is non-destructive and requires minimal sample preparation. Its drawbacks are that it has low sensitivity and cannot study all metabolites present due to interference of low abundance metabolites by abundant analytes.[165,166]

The principal method used in metabolomics is mass spectrometry (MS). Mass spectrometry discriminates analytes by ionising them and using electrical fields to separate them by mass-to-charge ratio (*m/z*). The force an electrical field applies to a charge is proportional to its charge (z), as per Coulomb's Law.

$$F \propto z$$

(1)

As described by Newton's second law of motion, the acceleration experienced by an object is inversely proportional to its mass.

$$F = ma$$

(2)

Taken together, the acceleration of a mass spectrometry analyte is dependent on both its charge and its mass, employed by mass spectrometry to measure different molecules.

Mass spectrometry is often coupled with a preceding analyte separation step, either capillary electrophoresis (CE) or high-performance liquid chromatography (HPLC). CE and HPLC are themselves analytical tools when after the analyte separation, a UV absorbance detector or a radioactivity counter for radiolabelling measure signal intensity, but their sensitivity is not as high as mass spectrometry; their main strengths are in separating analytes that may have similar *m/z*, reducing the problem of analyte interference for MS, as well as assigning a time of separation to each analyte as another characteristic for identification (known as retention time for HPLC and migration time for CE).

HPLC-MS and CE-MS can measure changes across samples for multiple analytes.[163,167–173] HPLC-MS is used more widely than CE-MS, possibly due to variations in migration times for CE-MS that can impede analyte identification.[174]

One of the challenges of mass spectrometry is that analytes with the same mass and charge (*m/z*) will have interfering signals if they have similar retention times. The problem is reduced

with better chromatographic analyte resolution and the use of standards for aiding identification, as well as through accurate mass instruments that can discriminate between ions of similar *m/z*.[165] Sensitivity increases further if instead of measuring a wide continuous range of *m/z* values (untargeted analysis), the instrument uses targeted analysis; this uses more than one mass analyser, with an analyte fragmentation step in between. Known as tandem mass spectrometry, or MS/MS for two mass analysers, it allows multiple filtrations of analytes and their products based on *m/z* and consequently reduces noise of impurities and other metabolites that may have had the same initial *m/z*.

The drawback of targeted analysis is that the instrument will only analyse a defined subset of analytes, with a narrower ability to observe the impact of the perturbation and reducing the possibility of unexpected discoveries.

In addition to measuring metabolite concentrations, MS and NMR techniques can probe the activity of enzyme networks through stable isotope labelled tracer experiments. In a similar way to radiolabelling, the organism of interest is exposed to a metabolic precursor labelled with a stable isotope such as ^{13}C, ^{2}H and ^{15}N (an isotopologue). The isotopologue is sequentially metabolised by the series of enzymes in a particular pathway to form different labelled metabolites. These can be harvested at different time points to gain insights into metabolism kinetics and pool dynamics. Stable isotope tracing using ^{13}C-labelled has been used to quantify metabolite fluxes (turnover of a molecule pool over time) in cancer cells and *in vivo* to better understand energetic contributions from different metabolites.[175,176] Labelling has also been applied in nucleotide metabolism on the role of enzymes in the pyrimidine biosynthesis pathways as well as measuring the contributions of salvage and synthesis.[177–179]

Labelling experiments are broadly run in two methods. The first, steady state labelling, exposes an organism to a labelled metabolic precursor and the equilibrium labelling is

measured after a period of time.[160] The second is pulse-chase analysis, where organisms are exposed to labelled precursors for a limited time, during which the compounds are metabolised into the downstream molecules of the pathway. Measuring the change in concentration over different time points gives information on the kinetics of the enzymes in the pathway.[105,180]

1.9.2.2 Nucleotide metabolism analysis: nucleic acid composition

As well as studying the free nucleotide pool, groups have also investigated nucleotide metabolism by analysing the composition of nucleic acids, using labelling techniques to track the nucleotide incorporation into DNA and RNA. While it is not possible to analyse the differences between nucleotides of different phosphorylation states through nucleic acid analysis, it still gives the opportunity to measure the impact of enzyme manipulation, replication and nucleic acid turnover.[181–184]

The established methods for measuring nucleotide precursor incorporation into nucleic acids in cells are chemical labelling, radiolabelling and mass spectrometric analysis of isotopic labelling.

Chemical labelling involves the incorporation into nucleic acids molecules against which one can use subsequent chemical derivatisation, such as 5-ethynyl-2'-deoxyuridine (EdU), or through antibody conjugation as in the case of 5-bromo-2'-deoxyuridine (BrdU). Both examples above are thymidine analogues, and the fluorescence is used for evidence of DNA replication.[185–187] With the advent of super resolution microscopy, it has been possible to obtain sub-nuclear spatial resolution of incorporation by EdU.[188] The drawbacks of chemo/immuno-fluorescence for measuring incorporation in nucleic acids are that one cannot definitively determine the proportion of DNA that has been derived from labelled

precursors–only relative change between conditions– and that it is not possible to measure interconversion between nucleosides such as between cytidine and uridine. In addition, BrdU has mutagenic properties[189,190] and EdU induces double-strand breaks[191] (DSBs), creating the potential for mutations in enzymes critical to nucleoside metabolism and thus pathway disruption.

Radiolabelling of nucleic acids has been used for studying nucleotide metabolism: Stasolla et al. exposed tobacco BY-2 cells to ^{14}C-labelled precursors to pyrimidines to study nucleotide metabolism during programmed cell death by measuring radioactivity of the isolated nucleic acid material.[192]

Mass spectrometry avoids exposure to radioactivity encountered through radiolabelling, and the process is not mutagenic to the organism under study as it uses stable isotopologues of nucleotide precursors as tracers, instead of chemically tagged species.[193] It is extensively used in nucleic acids research, notably for quantifying epigenetically modified or damaged nucleosides such as mdC and 8OdG. It is sensitive, with limits of detection of the major and minor bases dC, mdC and hmdC reported to 5 amol or lower using targeted triple quadrupole mass spectrometry (LC-MS/MS).[194,195] The technique has been used for isotope tracing in nucleic acids; by exposing organisms to isotopically tagged nucleotide precursors, hydrolysing the nucleic acids into nucleosides and analysing them by mass spectrometry, one can obtain the global proportion of the precursor incorporation into nucleic acids and its subsequent metabolism through isotope tracing.[105,182,193]

Using MS, one could simultaneously probe incorporation via the synthesis and salvage pyrimidine pathways through the use of differently-labelled precursors and allow comparison of both synthesis and salvage pathway-mediated incorporation.

1.9.3 Overview of the Agilent 1290-6495a QQQ HPLC-MS/MS

For analysis of both the free nucleotide pool and nucleic acid composition, it was decided to use an Agilent 6495a triple quadrupole mass spectrometer (QQQ), coupled to an Agilent 1290 Infinity ultra-high-performance liquid chromatography (UHPLC) instrument to form an HPLC-MS/MS system. This system was chosen given similar instruments' capabilities of sensitive multiple compound analyses of DNA—with nucleoside detection sensitivities reaching below 5 amol—as well as the use of HPLC-QQQ systems for nucleotide analysis.[182,194–197]

The combination of HPLC and MS takes advantage of the quantification power of MS and the separation of analytes to avoid ionisation competition at the MS source and signal interference between analytes of similar *m/z*.

The HPLC separates analytes according to their chemical properties. The sample is injected onto a column containing chemically functionalised material to which the sample components have different affinities. The column is the stationary phase and the solvent, or buffer, passing through it is the mobile phase. The mobile phase is composed of at least two different solvents with different polarities, the relative ratio of which changes throughout a run; some analytes will preferentially adsorb to the column stationary phase while exposed to a mobile phase primarily composed of one solvent, then will desorb and elute from the column when the mobile phase changes composition. This way, analytes are separated by properties like hydrophobicity, or interactions with ion-pairing agents present in the mobile phase.

Once the analytes elute from the column, they pass through a UV absorbance detector for HPLC-UV analysis (if present) and then to the MS source.

The Agilent 6495a source generates ions by a variant of electrospray ionisation (ESI) called Agilent Jet Stream ESI (AJS ESI). ESI is a gentle process that ionises samples by creating charged solvent droplets and drying them to form gaseous analyte ions; AJS ESI is reported to increase

the signal further, confining and heating the droplet spray with superheated nitrogen gas to improve desolvation and focusing the ion beam (**Figure 1.13**).[198]

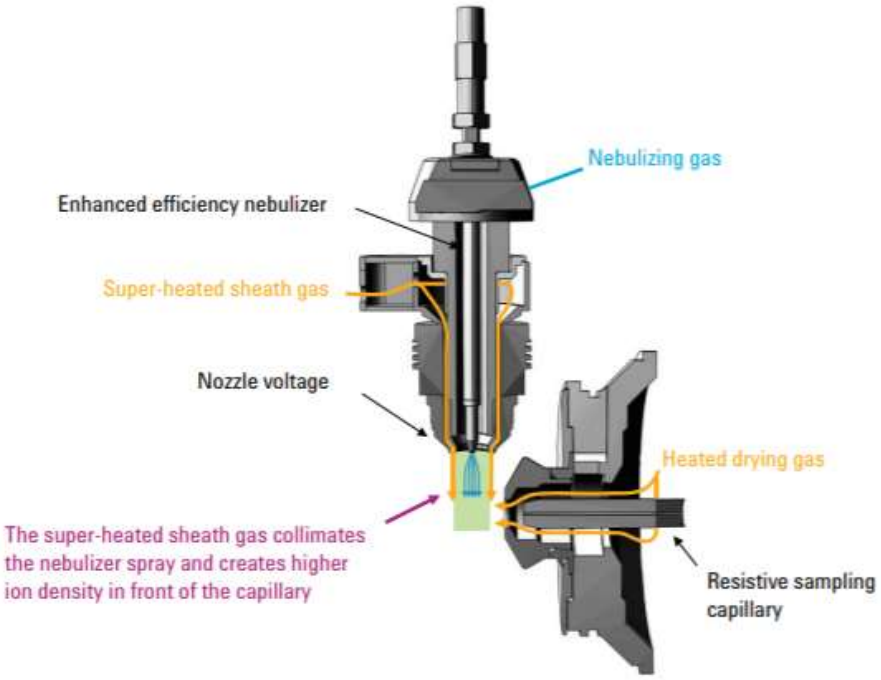


Figure 1.13 Schematic representation of the Agilent Jet Stream electrospray ionisation (AJS ESI) source. *The details of the AJS arrangement in front of the front end of a triple quadrupole mass spectrometry (QQQ) are shown, illustrating the gas flow (red) and the area of enhanced desolvation (blue), reproduced with permission from Agilent Technologies ("Agilent Jet Stream Thermal Gradient Focusing Technology", Mordehai and Fjelsted).*[199]

ESI performed in the AJS source consists of the following steps: i) The sample compounds exit the HPLC and enter the nebuliser; ii) The nebuliser converts the solvent into a fine mist; iii) The nozzle ionises the droplets, which are subjected to desolvation by the sheath gas to leave ions; iv) Ions are attracted into a charged capillary and v) The ions leave the capillary and pass into high and low RF funnels that eliminate most of the remaining neutral gas.

The Agilent 6495a is a QQQ instrument and measures analytes by generating complex electric fields with its two quadrupole mass analysers to filter charged species of certain *m/z*.[200] Once the ions leave the source, they are either analysed in untargeted or targeted mode.

In untargeted mode, the instrument rapidly scans for every possible *m/z* to obtain a broad landscape of ion signals. This is useful for the detection of unknown sample components, but there is no *m/z* filtration and so consequently there is noise present from other analytes and background impurities that impairs sensitivity. The instrument is most sensitive when it uses both mass analysers and is run in targeted Molecular Reaction Monitoring mode (MRM) (**Figure 1.14**). In essence, MRM consists of the following steps: i) The ions (termed precursor ions) reach the first quadrupole, where only the electric fields only allow analytes with the correct *m/z* to pass through; ii) The ions reach the collision cell (the second quadrupole) and are fragmented by a stream of nitrogen gas through collision-induced dissociation to form fragment ions (termed product ions); iii) The product ions are filtered again by *m/z* by the third quadrupole; iv) The product ions reach the detector for quantitation.

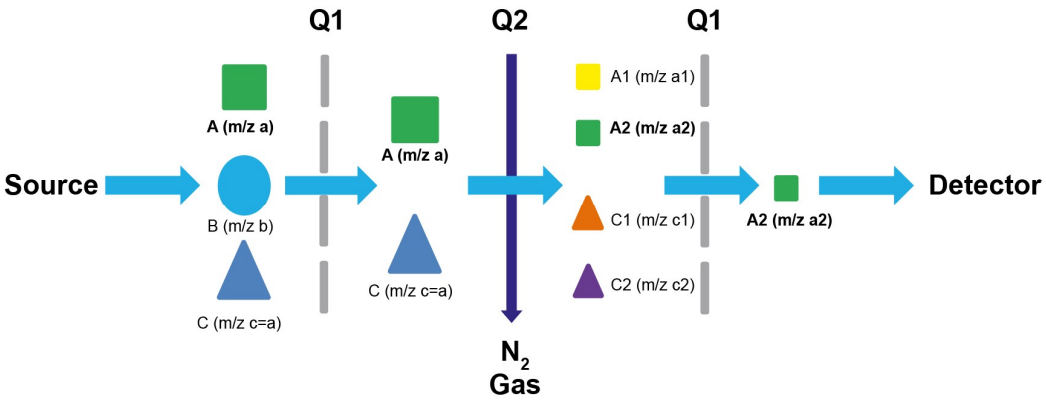


Figure 1.14 Schematic of mechanism of MRM mode analysis. *Parent ions A and C are isolated from mix A, B, C by quadrupole 1 (Q1), fragmented using nitrogen (N2) gas in Q2, leading to daughter ions A1, A2, C1, C2, which are again isolated for detection in Q3.*

1.10 Aims of the project

The enzymes CDA and DCTD are responsible for maintaining the balance between cytidine and thymidine species in the nucleotide pool, necessary for correct DNA replication. The

relevance of CDA is well known in disease conditions like chemotherapy resistant tumours and Bloom Syndrome, but DCTD is less known.

It is unclear why CDA and DCTD have both evolved, or whether either CDA or DCTD is dominant as a deaminase. CDA deaminates both dC and rC, while DCTD is predominantly active against dCMP, with low activity for CMP. It may be that there is an overproduction of deoxycytidine nucleotides or overactivity of RNR for which DCTD is designed to counteract.

Additionally, it is uncertain whether CDA and DCTD act on newly synthesised nucleotides that have been dephosphorylated and are thus also part of the *de novo* synthesis pathway. This would occur through the dephosphorylation by DCTPP1 of dCTP to dCMP[59]. If this does occur, it is also unknown as to whether they are dephosphorylated to nucleosides or just to nucleotide monophosphate. Considering literature suggests that food derived nucleosides, the bulk of extracellular nucleosides, contribute to only up to 5 % incorporation *in vivo* leaves open the possibility of their action on newly derived nucleotides.[12] Evidence of such activity would further intertwine the synthesis and salvage pathways.

My project aims were to develop tools for CDA and DCTD enzyme manipulation and to assess the impact on nucleotide metabolism. Specifically, I chose to investigate whether either CDA or DCTD is the principal deaminase in cells, as well as whether newly synthesised cytidine nucleotides are deaminated by CDA and DCTD. Considering the impact of deaminase expression on cancer prognosis, a deeper understanding of the relative activity and importance of the deaminases could aid decision making in the clinical administration of chemotherapy to cancer patients with high expression of CDA and DCTD.

Using LC-MS/MS to analyse both the free nucleotide pool and nucleic acid composition, I set out my research to perform CDA and DCTD enzyme siRNA KD in cells and to probe their nucleotide metabolism with labelled nucleotide precursors (**Figure 1.15**). The labelling

proportion of the nucleotide pool and nucleic acids would provide the answers to my research questions on impact of CDA and DCTD on pyrimidine nucleotide metabolism.

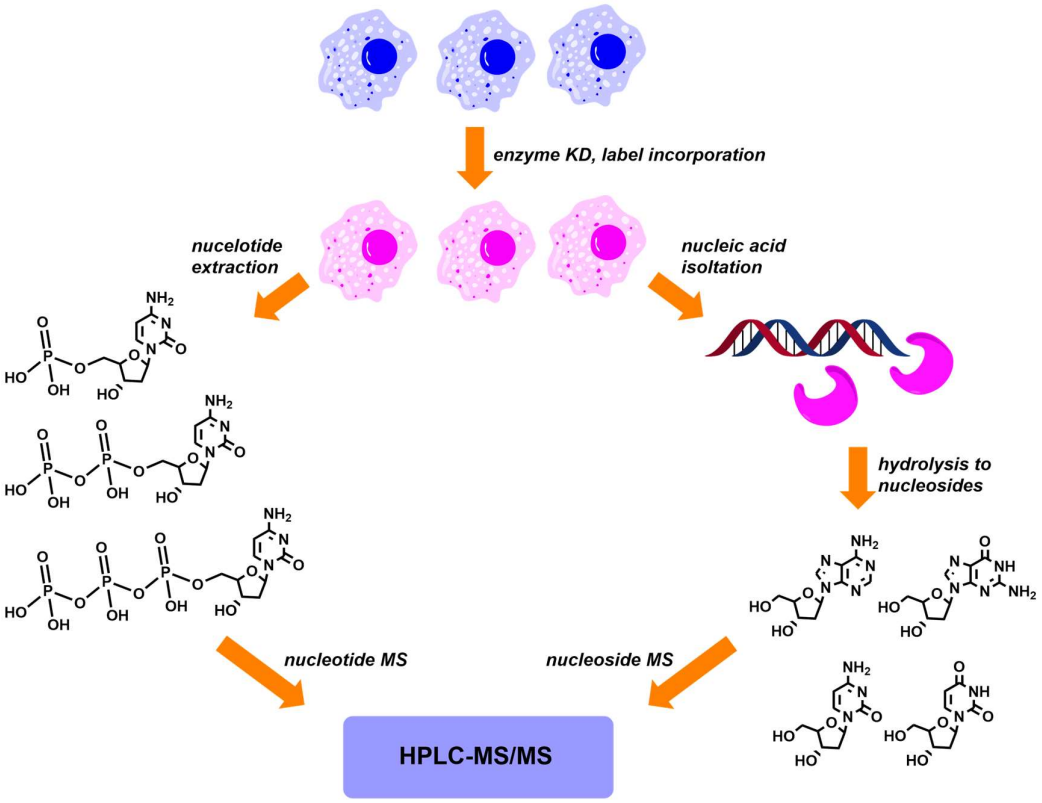


Figure 1.15 Workflow for analysis of nucleotide pool and nucleic acid composition from treated samples. Treatment: *cells are treated for CDA and DCTD KD, and for precursor labelling exposure.* ***Extraction****: nucleotides are extracted and nucleic acid is isolated from remaining cell pellet. Nucleic acid is hydrolysed to nucleosides by nucleases and phosphatase.* ***Analysis:*** *nucleotide and nucleoside labelling is quantified by MS.*

2. Materials and Methods

2.1 Cell culture

MDA MB-231, MDA MB-231 shCDA, H1299 and H1299 dsRedCDA were grown in in DMEM (D5671, Merck) complemented with 10 % fetal bovine serum, 2 mM L-glutamine, 1 mM pyruvate and 1 % penicillin-streptomycin (Lonza).

Hop-92 was grown in RPMI 1640 (21875, Gibco), complemented with 10 % fetal bovine serum and 1 % penicillin-streptomycin. PC-3 was grown in Ham's F-12K (Kaighn's) (21127, Gibco), complemented with 10 % fetal bovine serum and 1 % penicillin-streptomycin.

The cells were split upon confluency to a ratio of 1:4, using PBS for washing and trypsin (25300-054, Gibco) for detaching from the flask. The incubator conditions were 37 °C with a 5 % CO_2 atmosphere.

The cell lines were tested monthly for mycoplasma infection.

When freezing cells for storage, the cells were grown in T75 flask. Each flask was washed in 1 mL PBS and had 1 mL trypsin added, then was placed in the incubator for 2 minutes. 1 mL of media was added to the flask and the cells collected in a falcon tube (227 261, Greiner bio). The tube was centrifuged in a centrifuge (5910R, Eppendorf) at 300 g for 5 minutes. The supernatant was then removed and replaced with 3 mL of 90 % media, 10 % DMSO (D2650, Merck). The cells were aliquoted into 3 different cryotubes and frozen at -80 °C, for later storage at -196 °C.

When thawing and growing cells, the cells were thawed at room temperature and resuspended in a T25 flask. They were split 1:4 the next day.

2.2 siRNA transfection

For MDA MB-231 and PC-3: 50,000 cells were seeded per well in a 12-well plate for transfection the following day. Cells were counted with A TC10 automated cell counter (Biorad). For each reaction, 2 µl RNAiMax transfection reagent (Thermo Scientific) was diluted in 125 µl OptiMEM (Thermo Scientific), and 0.5 µl of each 100 µM siRNA (Mission siRNA, Merck) was diluted in 125 µl Opti-MEM reduced serum medium (31985062, Gibco).

Sequences of siRNAs used shown in **Table 2.1**.

Name used	**Target enzyme**	**Merck siRNA ID**	**siRNA target sequence**	**siRNA antisense sequence**
CDA1	CDA	SASI_Hs01_00229075	CTGGAGAACTTCATAAAGA	UCUUUAUGAAGUUCUCCAGdTdT
CDA2	CDA	SASI_Hs01_00229077	GCACCAACTGGCCCGTGTA	UACACGGGCCAGUUGGUGCdTdT
CDA3	CDA	SASI_Hs02_00332546	CCAGTGACATGCAAGATGA	UCAUCUUGCAUGUCACUGGdTdT
DCTD1	DCTD	SASI_Hs01_00037530	CTAATGCCTTTCATCTTGA	CUAAUGCCUUUCAUCUUGAdTdT
DCTD2	DCTD	SASI_Hs01_00037531	GGATTATCGTCTTCTAAGA	GGAUUAUCGUCUUCUAAGAdTdT
DCTD3	DCTD	SASI_Hs01_00037532	GATGTGAAAGGCTGTAGTA	GAUGUGAAAGGCUGUAGUAdTdT

Table 2.1 Sequences of oligonucleotides used for siRNA KD of CDA and DCTD

After incubation at RT for 5 min, the diluted transfection reagent was mixed into the siRNA dilution and incubated for 20 min at RT. During this time, the medium on the cells was exchanged with 750 µl fresh medium per well. 250 µl of the transfection mixture was added dropwise to each well. The final concentration of individual siRNAs per well was 50 nM. The cells were incubated at 37 °C for 24 h, after which they were expanded to 6-well plates.

2.3 Immunoblotting

Cells were lysed in RIPA buffer (20 mM HEPES at pH 7.5, 300 mM NaCl, 5 mM EDTA, 10 % glycerol, 1 % Triton X-100, supplemented with protease inhibitors (Protease Inhibitor Cocktail Set III, EDTA-Free, Merck). They were agitated on a rotator (Z675164, Merck) at 4 °C for half an hour and then centrifuged at maximum acceleration for 10 minutes at 4 °C (5424R, Eppendorf).

The lysates then quantified for protein concentration with Bradford reagent (B6916, Merck) on a FLUOstar Omega multiplate reader, calibrated with BSA standard (P083, Merck).

20 µg of protein lysate was loaded on a Mini PROTEAN precast 4-20 % acrylamide gel (4568095, Biorad). For marking protein sizes, pre-stained protein marker was used on either end of the gel (Page Ruler Plus, Thermo Scientific).

The gel was electrophoresed with a Mini Protean Tetra cell (1658005EDU, Biorad) in 1x SDS-PAGE Running Buffer (25 mM Tris-base, 192 mM glycine, 0.1 % SDS) for 1hr at 150 V.

After electrophoresis, the sample proteins were transferred to a nitrocellulose membrane (GE10600002, Amersham) in a 2-Gel Tetra Blotting Module (1660827EDU, Biorad) filled with 4 °C 1x Transfer Buffer (25 mM Tris-HCl pH 7.6, 192 mM glycine, 20 % methanol) with an ice

pack for 1 hour at 100 V. To assess the success of the transfer and equal loading of samples, the membrane was stained with Ponceau S (P7170, Merck) and imaged with a ChemiDoc instrument (BioRad). The membranes were then de-stained with 1x TBST (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.1 % Tween-20 (Fisher Chemicals BP337-500)). The membranes were blocked in 5 % milk in TBST (Marvel Original dried skinned milk) for 1h at room temperature. After blocking, the membranes were cut according to protein size bands and were placed within a 50mL Falcon tube containing a primary antibody overnight, rotating and incubating at 4 °C overnight. The membranes were washed in 1x TBST for five minutes three times and then incubated with an HRP-conjugated secondary antibody, and incubated at room temperature for 1 hr. In parallel, the membrane band spanning actin was incubated with HRP-conjugated primary antibody for 1 hr at room temperature. The membranes were then washed in 1x TBST for five minutes three times. The membranes were treated with chemiluminescence reagent (Amersham ECL Select Western Blotting Detection Reagent) and then imaged with the ChemiDoc System (Bio-Rad) with Image Lab software (Bio-Rad, version 5.2).

All antibodies and their dilutions used are presented in **Table 2.2**.

Antibody	Target protein	Source	Dilution	Species	Usage
Anti-CDA (centre) antibody	CDA	SAB1300717, Merck	1:500	Rabbit	Primary
Anti-DCTD	DCTD	16784-1-AP, Proteintech	1:2000	Rabbit	Primary
Recombinant HRP Anti-muscle Actin antibody	Actin	ab185058, Abcam	1:50000	Rabbit	Primary, HRP-conjugated
Goat Anti-Rabbit IgG (H + L)-HRP Conjugate	Rabbit	1706515, Biorad	1:10000	Goat	Secondary, HRP-conjugated

Table 2.2 Antibodies used for comparing CDA and DCTD levels across samples.

2.4 Pulse-chase experiment

After 48 hours post siRNA transfection, the media was replaced with media containing 1 µM $^{13}C_9$ dC (Silantes), 1 µM $^{13}C_5$ rC (Merck) and 500 µM $^{13}C_5$ L-glutamine (Merck). 21 hours later, the media was changed and the cells were harvested at t=69 hrs, t=70 hrs, t=71 hrs, t=73 hrs, t=75 hrs, t=81 hrs, t=93 hrs post transfection.

2.4.1 Nucleotide extraction from biological samples

For adherent cell culture: the cells were washed twice with 1 mL PBS and snap frozen in liquid nitrogen.

The cells were lysed by addition of 1000 µL of 50 % acetonitrile (100029, Merck), and 50 % water (AM9937, Life Technologies Ltd). If no heavy labelled nucleotide precursors were added to the cells, heavy labelled internal standards were added for a final concentration of 1 µM.

The cells were scraped from the dish and the resulting suspension was transferred to an Eppendorf tube. It was heated at 95 °C for 1 minutes and then cooled rapidly in an ice bath. It was then sonicated on a Bioruptor UCD-200TM-EX for one minute and spun down at 4°C at maximum acceleration for ten minutes. 750 µL of the supernatant were transferred to an Eppendorf and dried on SpeedVac.

2.5 Nucleotide analysis

2.5.1 Nucleotides external and internal standards

The nucleotides were bought from the following sources:

dNTPs (Scientific Laboratory Supplies Limited); ADP, CMP, CTP, UMP, UTP, GMP, GTP, dAMP, dADP, dCMP, dCDP, dUMP, dGMP, U-^{13}C UTP, U-^{15}N ATP, U-^{15}N GTP (Merck); AMP (Fluorochem); ATP, dUDP (Jena Bioscience); CDP, dTMP, dTDP (Carbosynth); UDP, dUTP (Alfa Aesar); GDP (Abcam); dGDP (Chemcruz); 8-oxo-dGTP (Trilink); U-^{13}C dNTPs, U-^{13}C,^{15}N dNMPs, U-^{13}C,^{15}N rNMPs, U-^{13}C,^{15}N rNDPs (Silantes GmbH).

2.6 Nucleotide mass spectrometric conditions

The dried nucleotide samples were resuspended in 20 µL of water and 2 uL were injected.

For the analysis by HPLC-QQQ mass spectrometry, a 1290 Infinity UHPLC was fitted with an InfinityLab Poroshell 120 peek-lined HILIC-Z (2.7 µm, 2.1 mm, 150 mm; Agilent Technologies) and coupled to a 6495a Triple Quadrupole mass spectrometer (Agilent Technologies) equipped with a Jetstream ESI-AJS source. The data were acquired in dMRM mode using positive electrospray ionisation (ESI).

The AJS ESI settings were as follows: drying gas temperature 230 °C, the drying gas flow 14 Lmin-1, nebulizer 20 psi, sheath gas temperature 400 °C, sheath gas flow 11 Lmin-1, Vcap 2,000 V and nozzle voltage 0 V. The iFunnel parameters were as follows: high pressure RF 110 V, low pressure RF 80 V. The fragmentor of the QQQ mass spectrometer was set to 380 V and the delta EMV set to +400.

The buffers are as follows. Buffer A: H_2O (115333, Merck), 10 mM ammonium bicarbonate (09830, Merck), pH 6 with glacial acetic acid (10394970, Fisher Scientific) and 5 µM medronic acid (5191-4506, Agilent Technologies).

The gradient used to elute the nucleosides started by a 2-min isocratic gradient composed with 15 % buffer A (and 85 % buffer B (100 % acetonitrile) with a flow rate of 500 µL min^{-1} and was followed by the subsequent steps: 10-12 min, 69.9 % A; 15-16 min, 15 % A. The gradient was followed by a 5min post time to re-equilibrate the column.

The transitions and retention times used for the characterization of nucleotides and their adducts are summarized in Appendix A, **Table 3**.

2.7 Nucleoside analysis

2.7.1 Standards for nucleoside mass spectrometry analysis and for heavy labelled incorporation

The suppliers of the following nucleosides and nucleotide precursors is as follows:

dC, dG, dA, dT, 3mdC, 5mdC, 8-oxo-dG, m^6dA, dU, hmdU, BrdU, EdU, rU, rC, rG, rA, m^6rA, dI, rI, $^{13}C_5$ L-glutamine and $^{13}C_4$ aspartic acid were purchased from Merck.

hmdC was purchased from Cayman Chemical.

fdC and cadC were purchased from Jena Bioscience.

$^{13}C^{15}N_2$ dC, $^{13}C^{15}N_2$ dU, $^{13}C^{15}N_2$ mdC, $^{13}C^{15}N_2$ hmdC, $^{13}C^{15}N_2$ 8-oxo-dG, $^{13}C^{15}N_2$ rU, $^{13}C_5$ dG, $^{13}C_5$ dA, $^{13}C_5$ dT, $^{13}C_5$ rG, $^{13}C_5$ rA, $^{13}C_5$ rC, $^{13}C_5$ rI and $^{13}C_1{}^{15}N_2$ orotic acid were purchased from Toronto Research Chemicals.

$^{13}C_9$ dC was purchased from Silantes GmbH.

2.7.2 Nucleic acid isolation

Nucleic acid was extracted from cultured cells with two kits.

The first kit was the Zymo Research Quick-DNA 96 Kit (D3011), used for isolating DNA obtained from the pulse-chase experiment samples. The manufacturer's instructions were followed. Before adding genomic lysis buffer, the samples were dried of metabolite extraction buffer on a hot plate at 95 °C.

The second kit was the GeneJET Genomic DNA Purification Kit (ThermoFisher Scientific) following the manufacturer's instructions without use of RNase A.

Nucleic acid concentration was determined by Qubit fluorometric quantitation (ThermoFisher Scientific) with the ds DNA HS assay kit, following the manufacturer's instructions.

2.7.3 Nucleic acid hydrolysis

Hydrolysis was carried out in 40 µl reactions that contained 500 ng DNA, 100 mM NaCl, 20 mM $MgCl_2$, 20 mM Tris pH 7.9, 1000 U/mL Benzonase (E8263, Merck), 600 mU/mL Phosphodiesterase I (P3243, Merck), 80 U/mL Alkaline phosphatase (P7923, Merck), 36 µg/mL EHNA hydrochloride and 2.7 mM deferoxamine for 2 hours at 37 °C. Following lyophilisation, nucleosides were resuspended in buffer A (H_2O (115333, Merck) with 10 mM ammonium acetate (10365260, Fisher Scientific), brought to pH 6 with glacial acetic acid (10394970, Fisher Scientific)) with 12.5 nM isotopically labelled internal standards.

2.7.4 Nucleoside HPLC-MS/MS analysis

For the analysis by HPLC– QQQ mass spectrometry, a 1290 Infinity UHPLC was fitted with a Zorbax Eclipse plus C18 column, (1.8 µm, 2.1 mm 150 mm; Agilent Technologies) and coupled to a 6495a Triple Quadrupole mass spectrometer (Agilent Technologies) equipped with a Jetstream ESI-AJS source. The data were acquired in dMRM mode using positive electrospray ionisation (ESI1). Rare DNA/RNA components were quantified by mass spectrometry, whereas abundant components were quantified by HPLC-UV. The gradient used to elute the nucleosides started with a 5 min isocratic gradient composed of 100 % buffer A and 0 % buffer B (100 % methanol—106035, Merck) with a flow rate of 0.4 mL/min and was followed by the subsequent steps: 8–9 min, 94.4 % A; 16 min 86.3 % A; 17– 21 min 0 % A; 24.3–25 min 100 % A. The AJS ESI settings were as follows: drying gas temperature 230 °C, the drying gas flow 14

$Lmin^{-1}$, nebulizer 20 psi, sheath gas temperature 400 °C, sheath gas flow 11 $Lmin^{-1}$, Vcap 2,000 V and nozzle voltage 0 V. The iFunnel parameters were as follows: high pressure RF 110 V, low pressure RF 80 V. The fragmentor of the QQQ mass spectrometer was set to 380 V and the delta EMV set to +200.

The transitions and retention times used for the characterization of nucleosides and their adducts are summarized in Appendix A, **Table 1** and **2**.

2.8 Statistical analysis

The raw mass spectrometry data for nucleotide and nucleoside analysis was analysed by Agilent MassHunter Workstation Quantitative Analysis for QQQ Version 10.0. The sample concentrations were obtained by comparing relative ratios of sample/internal standard responses with those of the calibration curves.

Composition, heavy labelling and epigenetic, of nucleic acids and the nucleotide pool was calculated through Microsoft Excel. p-values were derived through a two-tailed Student's T Test, with significance declared at $p \leq 0.05$.

3. Optimising an HPLC-MS/MS method for nucleic acid analysis

Nucleotide precursors are substrates for the nucleotide metabolism pathway enzymes, including CDA and DCTD, eventually forming the end products DNA and RNA. Using stable isotopologues of nucleotide precursors are used as tracers, it is possible to monitor the effect of CDA and DCTD, as manipulation of these enzymes will affect the incorporation of the heavy labelled tracers into nucleic acids, the nucleoside composition of which can be detected by MS.[182]

In this chapter, I describe the work done to optimise a LC-MS/MS workflow for nucleic acid nucleoside composition analysis on an Agilent 1290-4695a instrument. The chapter begins with the adaptation of nucleic acid hydrolysis to nucleosides for HPLC-UV sample preparation to one suitable for HPLC-MS/MS analysis. It follows on to HPLC considerations, in particular the column choice for optimal separation of nucleosides and then to optimisation of MS parameters to for nucleoside sensitivity. Finally, I discuss the validation of the system's capability to measure DNA composition.

3.1 Optimising the hydrolysis process of nucleic acids to nucleosides

The method of nucleic acid hydrolysis I use is adapted from one previously used for HPLC-UV analysis.[62,201] HPLC-MS/MS is more sensitive than HPLC-UV and so I expected to need less starting material and to use less of the hydrolysis agents.

Using 500 ng of DNA from a cell line derived from HEK-293 in a reaction volume of 40 µL, I investigated the impact of changing the time of hydrolysis and phosphodiesterase I (PDE I) concentration, comparing the UV absorbance of dT between conditions.

Hydrolysis time: The previous hydrolysis time was overnight. I tested hydrolysis for 1, 2, 4 and 24 hrs to see if there is a benefit to a longer hydrolysis. Aside from the high variance for t=1 hr, **Figure 3.1** does not show a difference between the times. I reduced the hydrolysis time to 1 hour, although due to the high variance at t=1 hr, the time of hydrolysis optimisation experiment may benefit from repetition.

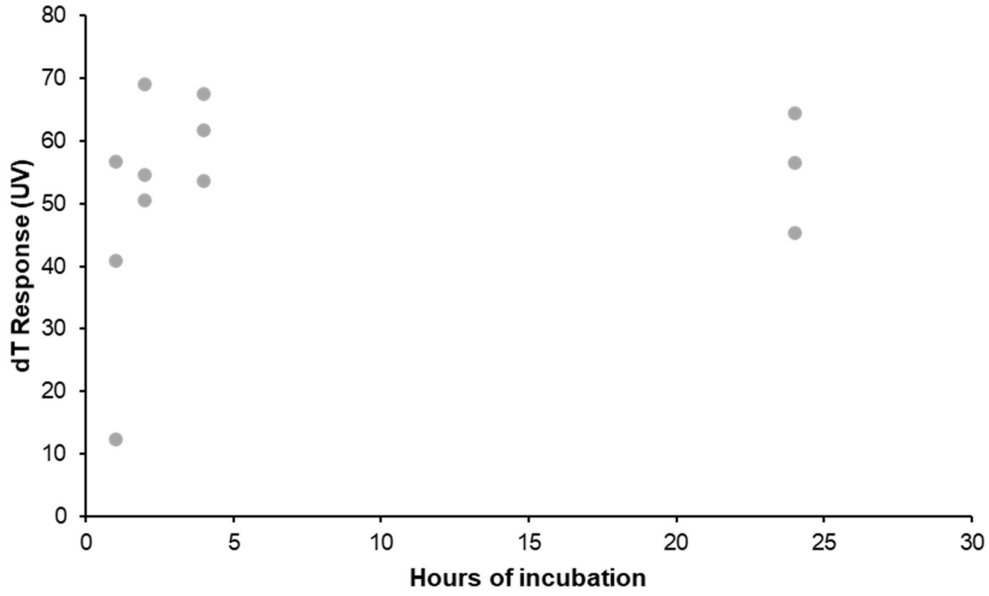


Figure 3.1 Increasing hydrolysis time does not improve signal of dT by UV. *500 ng of DNA hydrolysed, n=3 technical replicate hydrolysis experiments.*

Amount of PDE I: PDE I is purified from snake venom and is likely to have impurities that are batch specific. For this reason, it is best to reduce its use to remove chance of contamination and to prolong the lifetime of the acquired batch. The previous volume was 600 U/mL and I tested 125 to 500 U/mL. The readout of UV absorption response for dT showed no dependence on PDE units over the range tested (**Figure 3.2**). As the lowest concentration gave a comparable dT signal, I reduced the PDE I amount used to 125 U/mL.

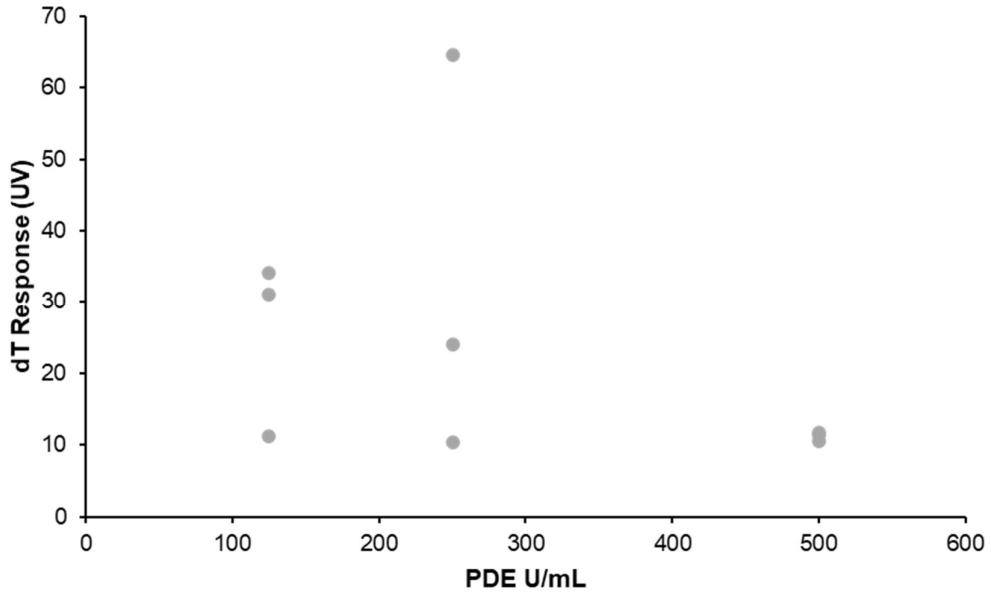


Figure 3.2 Increasing PDE concentration in hydrolysis mix does not increase dT signal by UV. *500 ng of DNA hydrolysed, n=3 technical replicate hydrolysis experiments.*

It may be possible to obtain sufficient hydrolysis with shorter incubation times and lower concentrations of PDE I and other reagents, but this requires further investigation.

3.2 Choice of HPLC column for optimising nucleoside separation

Nucleoside separation is crucial to avoid co-elution of certain nucleosides, such as dC and dU that generate ions which only differ by 1 *m/z* unit. Their natural isotopologues may interfere with each other's quantification if they co-elute.

Different column classes have been used to separate nucleosides for LC-MS analysis, with each class having different solid phase chemistry designed to separate certain classes of analytes. Reversed-phase columns are composed of a silica stationary phase bonded to hydrocarbon chains that retain hydrophobic molecules when the mobile phase is aqueous (predominantly solvent A), until washed off with organic mobile phase (predominantly

solvent B). Hydrophilic interaction chromatography (HILIC) uses a polar stationary phase that retains polar molecules while the mobile phase is organic and then the nucleosides elute when the mobile phase is aqueous.[202]

Nucleosides have polar groups with a carbon-rich and aromatic central molecular structure (**Figure 3.3**). Therefore, nucleosides have the potential to interact both with HILIC and reversed-phase column classes.

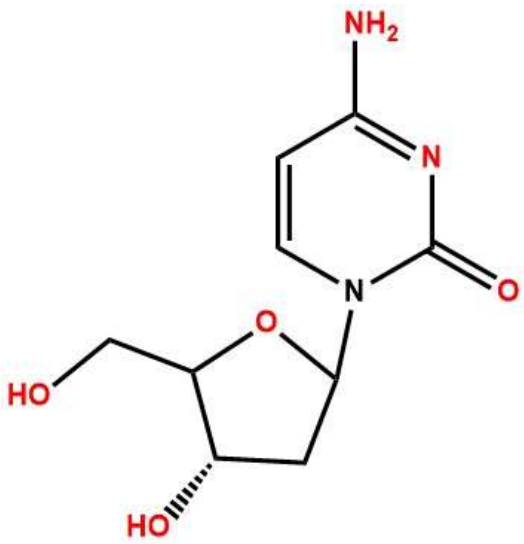


Figure 3.3 2'-deoxycytidine (dC), with atoms likely to interact favourably with polar solvents in red, and the hydrophobic molecule skeleton in black. *N-5 is not highlighted as its lone pair of electrons partially delocalised in the aromatic heterocycle and may engage less in solvent interactions.*

Both systems have been used in literature (**Table 3.1**) but reversed-phase chromatography has been preferred. This may be due to the ease of sample preparation, as the sample is resuspended in aqueous solvent instead of volatile organic solvent. Injecting aqueous sample into an organic mobile phase as found at the beginning of HILIC runs can lead to chromatography disruption, with misshapen peak shapes due to poor analyte partitioning.[203]

Column	Type of column	Chromatography class	Reference
1	Waters BEH HILIC	HILIC	Zhang et al.[194]
2	Agilent Zorbax Eclipse Plus C18	Reversed-phase	Pfaffeneder et al.[182]
3	Agilent Zorbax SB-C18	Reversed-phase	Zauri et al[62], influenced by Pfaffeneder et al.[182]
4	Agilent Zorbax Eclipse Plus C18	Reversed-phase	Amouroux et al.[195]

Table 3.1 Column types used in literature for nucleoside separation.

I aimed to maximise resolution (complete separation) of nucleosides peaks through chromatography, testing three reversed-phase columns (**Table 3.2**). As the system is an LC-MS/MS, unresolved peaks do not affect analyte quantitation unless they are near isobaric, although co-elution of abundant and rare analytes may cause signal suppression of the rare analyte through ionisation competition.

Column	Type of column	Length (mm)	Width (mm)	Particle size (μm)	Chain length
1	Agilent Zorbax SB-C18	150	4.6	1.8	18
2	Agilent Zorbax Eclipse Plus C18	150	2.1	1.8	18
3	Acquity UPLC BEH C8	150	2.1	1.7	8

Table 3.2 Description of columns tested for nucleoside, judging for separation.

Column 1 and 2 have the same carbon chain length of 18, as well as particle size which influences peak resolution. The difference is the smaller column width, which would allow similar chromatography with less solvent consumption.[204]

Column 3 has a shorter carbon chain of 8 instead of 18. Its smaller chain length reduces the column density and the number of hydrophobic interactions experienced by the analytes when passing through the column. This was expected to give shorter nucleoside retention time.[205,206]

Based on the criteria described above, I tested separation on each column at a flowrate of 0.4 mL min^{-1} (a flow rate that did not cause pressures to exceed the HPLC high pressure limit in any of the columns) with a gradient optimised for separation for each column. The solvent composition during column testing was as follows Solvent A: H_2O, 10 mM ammonium acetate, pH to 6 with glacial acetic acid. Solvent B: 40 % ACN, 60 % H_2O.

Figure 3.4 shows the extracted ion chromatograms (EICs, signals obtained by MS) of all nucleosides I then possessed. This excludes dA, which is off view but had good resolution with each column.

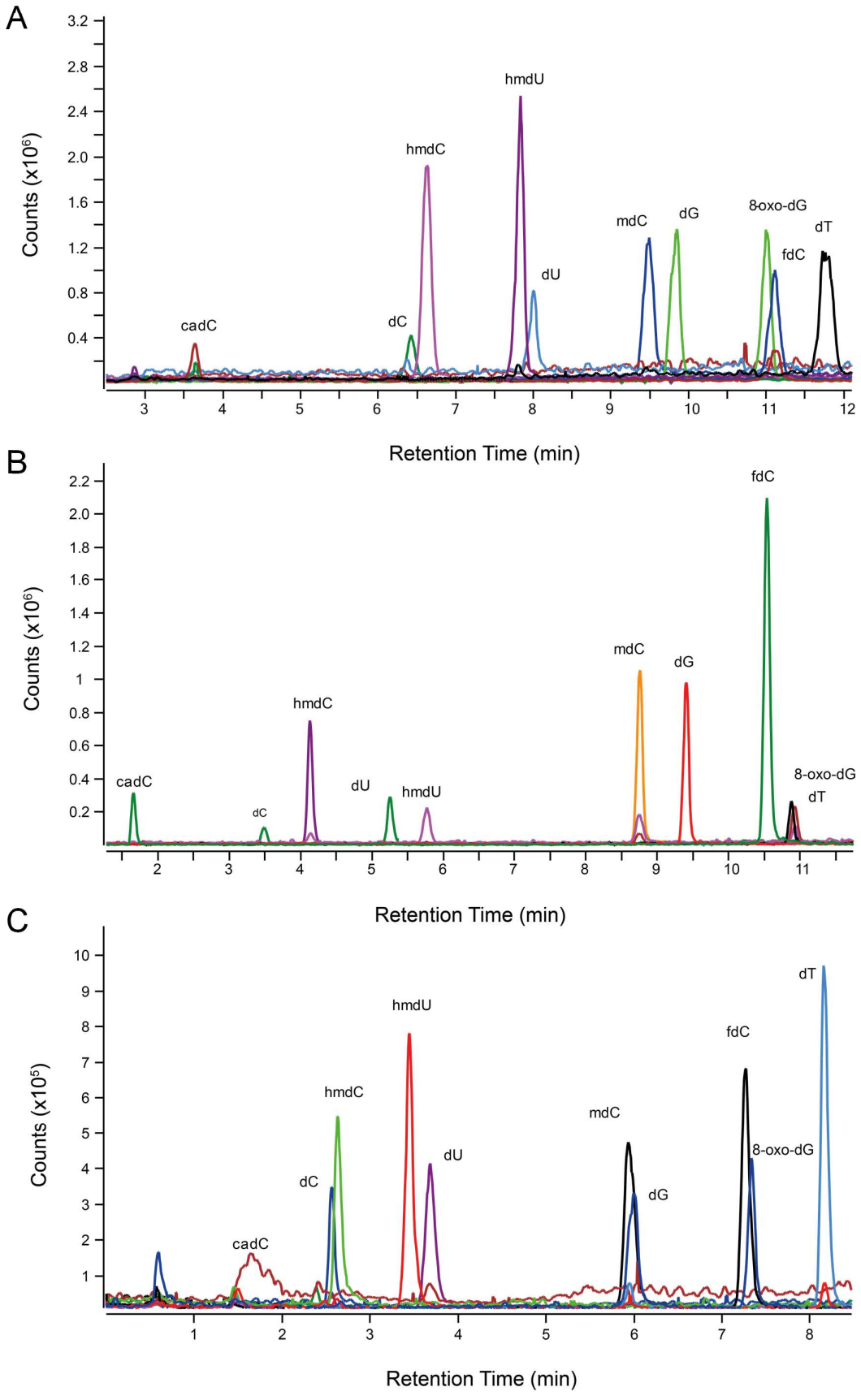

A
Counts (x10⁶)
cadC
dC
hmdC
hmdU
dU
mdC
dG
8-oxo-dG
fdC
dT
Retention Time (min)
B
Counts (x10⁶)
cadC
dC
hmdC
dU
hmdU
mdC
dG
fdC
8-oxo-dG
dT
Retention Time (min)
C
Counts (x10⁵)
cadC
dC
hmdC
hmdU
dU
mdC
dG
fdC
8-oxo-dG
dT
Retention Time (min)

Figure 3.4 Column 2 gives best nucleoside resolution. *Extracted ion chromatograms (EICs) of 12.5 pmol of nucleoside standards cadC, dC, hmdC, hmdU, dU, mdC, dG, fdC, 8-oxo-dG, dT analysed by LC-MS scan mode.* ***A,*** *separation by column 1;* ***B,*** *separation by column 2;* ***C,*** *separation by column 3.*

As expected, column 3 had the shortest nucleoside retention times and had several unresolved nucleosides. Column 2 gave the best resolution of nucleosides, with only the pairs dC, U and dT, 8-oxo-dG unresolved. The use of column 2 compared to column 1 allowed good separation and reduced solvent consumption. The gradient used for column 2 is found in **2.7.4**.

3.3 Optimising the Agilent 6495a QQQ MS for nucleoside signal

3.3.1 Choice of QQQ polarity for nucleoside signal

The QQQ is capable detecting ions in positive and negative mode. The literature on nucleoside analysis mass spectrometry favours the application of positive ion mode.[62,182,194,195] The rationale behind this is that nucleosides have more than one electronegative atom, such as oxygen and nitrogen that can chelate (**Figure 3.5**), and so are likely to produce positively charged ions through protonation or chelation with the group 1 metal ions sodium (Na^+) and potassium (K^+).

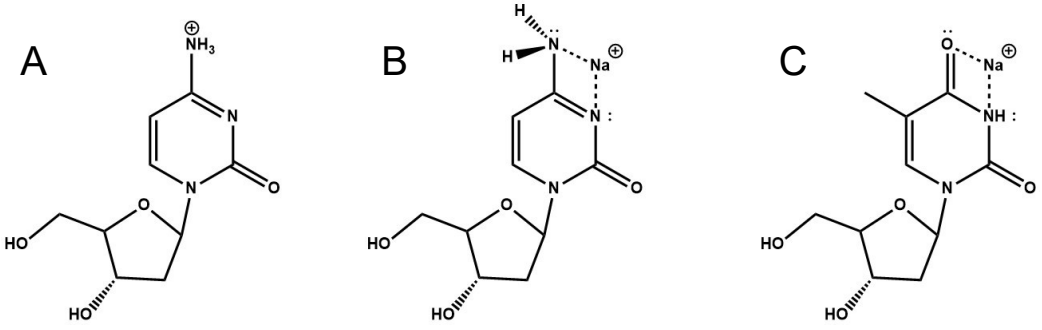


Figure 3.5 Nucleosides can form different ion adducts. *Structures of nucleoside positive ion adducts generated by ESI.* ***A,*** *protonated dC (m/z 228);* ***B,*** *sodiated dC (m/z 250);* ***C,*** *structure of sodiated dT (m/z 265).*

I therefore decided to conduct the mass spectrometry analysis in positive ion mode.

3.3.2 Impact of HPLC mobile phase on nucleoside MS sensitivity

3.3.2.1 Investigation of 3 mobile phase compositions for nucleoside response

The mobile phase of the HPLC is composed of two solvents, A and B, that have different polarities. Nucleosides elute from the column when the mobile phase is a certain composition, depending on their chemical properties. Once eluted, they reach the MS source. The mobile phase composition at that moment influences the chemical environment in which the nucleosides are prior to desolvation: it affects the salt concentrations, pH and solvent molecule type, thus impacting the ionisation process and signal at the detector.

As the MS ionisation process is solvent dependent, I assessed the suitability of three solvent systems for the LC-MS/MS system (**Table 3.3**), prioritising nucleoside signal.

Solvent system classification	Solvent A	Solvent B	Reference
1	97:3 H_2O/MeOH with 10 mM tributylamine + 15 mM acetic acid	MeOH with 10 mM tributylamine + 15 mM acetic acid	Sartain[207]
2	H_2O, 10 mM ammonium acetate, pH 6 with acetic acid	40 % ACN, 60 % H_2O	Zauri et al.[62]
3	H_2O, 10 mM ammonium acetate, pH 4 with acetic acid	40 % ACN, 60 % H_2O	Similar to Zauri et al.[62]

Table 3.3 List of mobile phase systems tested for nucleoside signal.

System 1 was developed by Agilent and is suggested as universal LC-MS/MS C-18 mobile phase suitable for nucleosides among other central carbon metabolism analytes.[207] Its defining feature is the use of tributylamine (TBA) a tertiary amine that is suitable for AJS ESI.

System 2 is a solvent system previously used in the Kessler group for nucleoside analysis and has been reported to be used by other groups.[62] System 3 is the same as system 2, but with a lower pH, which I did to test whether increasing proton availability would increase nucleoside ionisation.

System 1 gave a low sensitivity (**Figure 3.6A**) and a high saturated background count of protonated TBA (*m/z* 186), at concentrations of which would suppress nucleoside ionisation.[208] Having subsequently learned from Agilent technical specialists that this solvent system was best used in negative ion mode (personal communication), I decided not to use

this solvent system and to concentrate on the variants of the established system from the group.

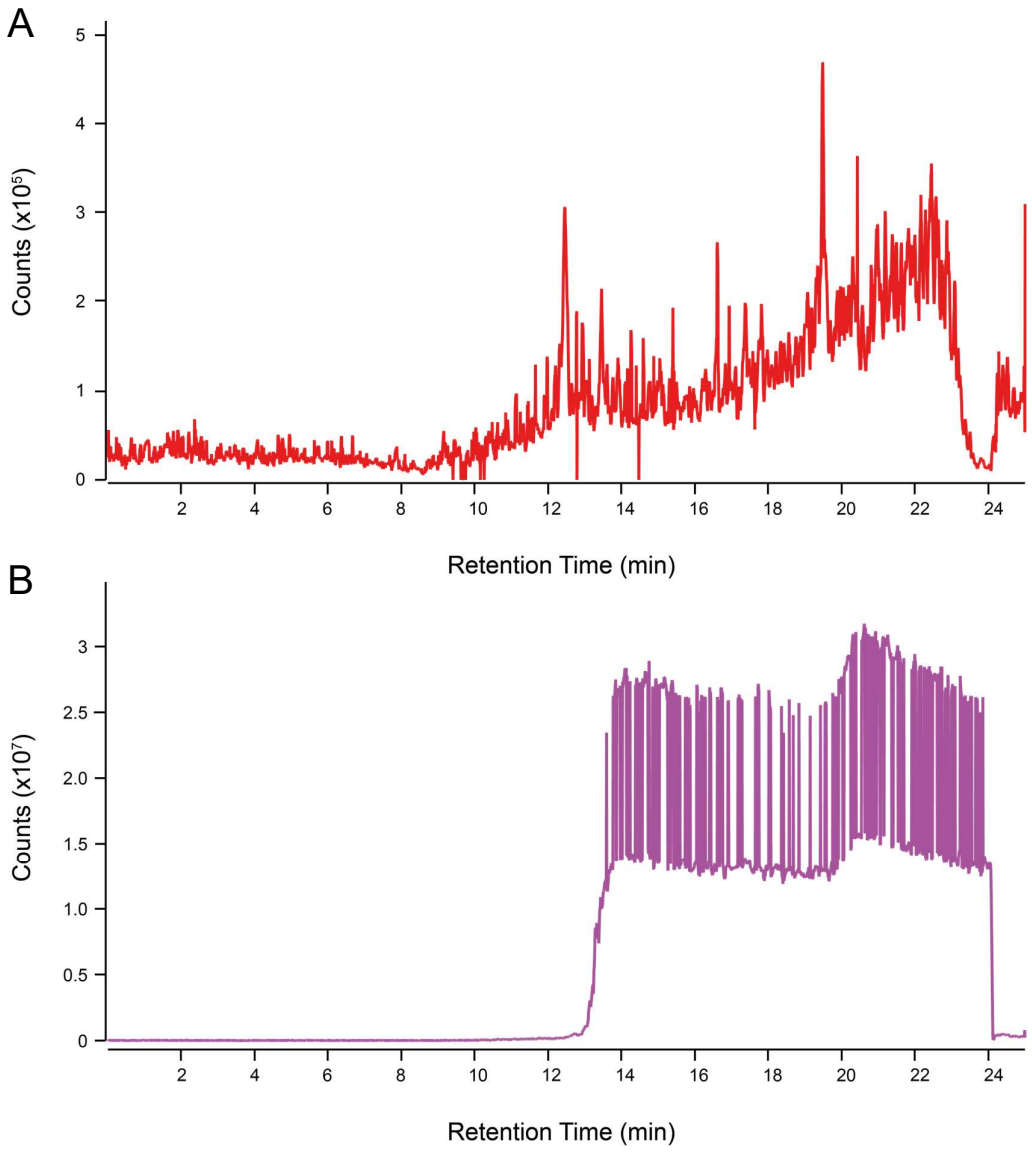


Figure 3.6 Tributylamine suppresses nucleoside signal through saturation. *Extracted ion chromatograms (EICs) of tributylamine and dU in solvent system 1. TBA signal is saturated, with no signal from dU.* ***A,*** *EIC of dU Na^+, 12.5pmol;* ***B,*** *EIC of TBA H^+.*

I tested systems 2 and 3 by injection of 12.5 pmol of each standard in scan mode on a Zorbax SB-C18 column on the gradient previously used in the Kessler group and monitored signal strength of hmdC and dU (**Figure 3.7**), hypothesising that the decreased pH in solvent system

4 would aid protonation of the basic nucleosides and increase signal, with a possible decrease in the sodiated signal due to the increase in free protons.

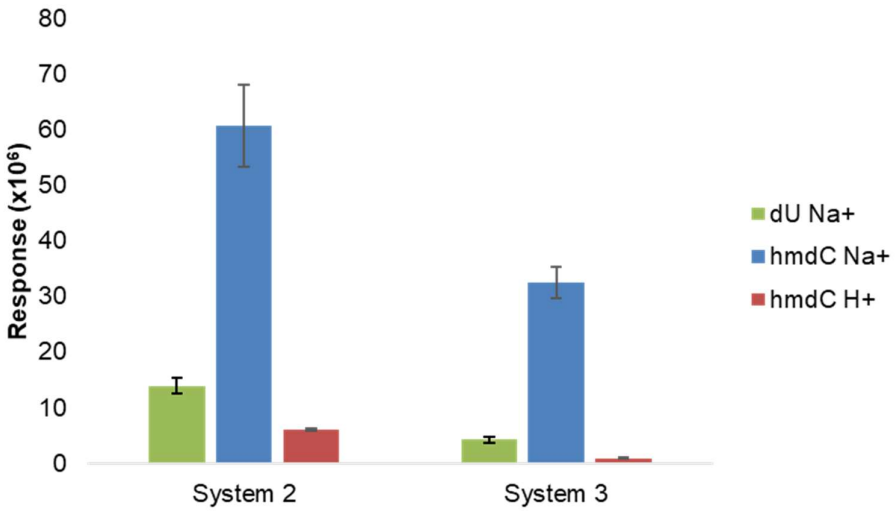


Figure 3.7. MS nucleoside detection sensitivity is solvent-dependent sensitivity of nucleoside detection. *Response of 1.25 pmol of nucleoside standards dU Na+ (green bars), hmdC Na+ (blue bars) and hmdC H+ (red bars) in LC-MS mode using solvent systems 2 (left panel) or system 3 (right panel). n=4, standard deviation is shown.*

However, the strongest response for all examined nucleosides was from system 2 for both protonated and sodiated adducts, and so it was decided to use system 2; it is unclear why the lower pH of system 3 would not increase protonation signal as expected, but it is possible that classical solvent thermodynamics are not applicable when ions are in the gaseous phase upon desolvation.

3.3.2.2 Changing Solvent B from 40 % ACN to 100 % MeOH increases nucleoside sensitivity

I entered a collaboration to provide orthogonal validation for a method using an Oxford Nanopore Technologies MinION to measure BrdU incorporation in DNA. The collaborator was interested in DNA replication dynamics in eukaryotic genomes, using BrdU as a tracer to

identify active DNA replication origins, fork direction, pausing and stalling sites, as well as replication termination sites.

The QQQ instrument sensitivity for BrdU was too low for analysis of the collaborator's samples and so I set out to investigate whether changing the composition of Solvent B would influence BrdU ionisation. I therefore changed Solvent B from 40 % ACN, 60 % H_2O to 100 % MeOH and ran the equivalent gradient for the standard mix. The retention times of all the nucleosides were similar to how they had been with 40 % ACN, 60 % H_2O as solvent B and the signal for most analytes was equal or higher (**Figure 3.8A**). In the case of BrdU, the signal increased twentyfold (**Figure 3.8B**).

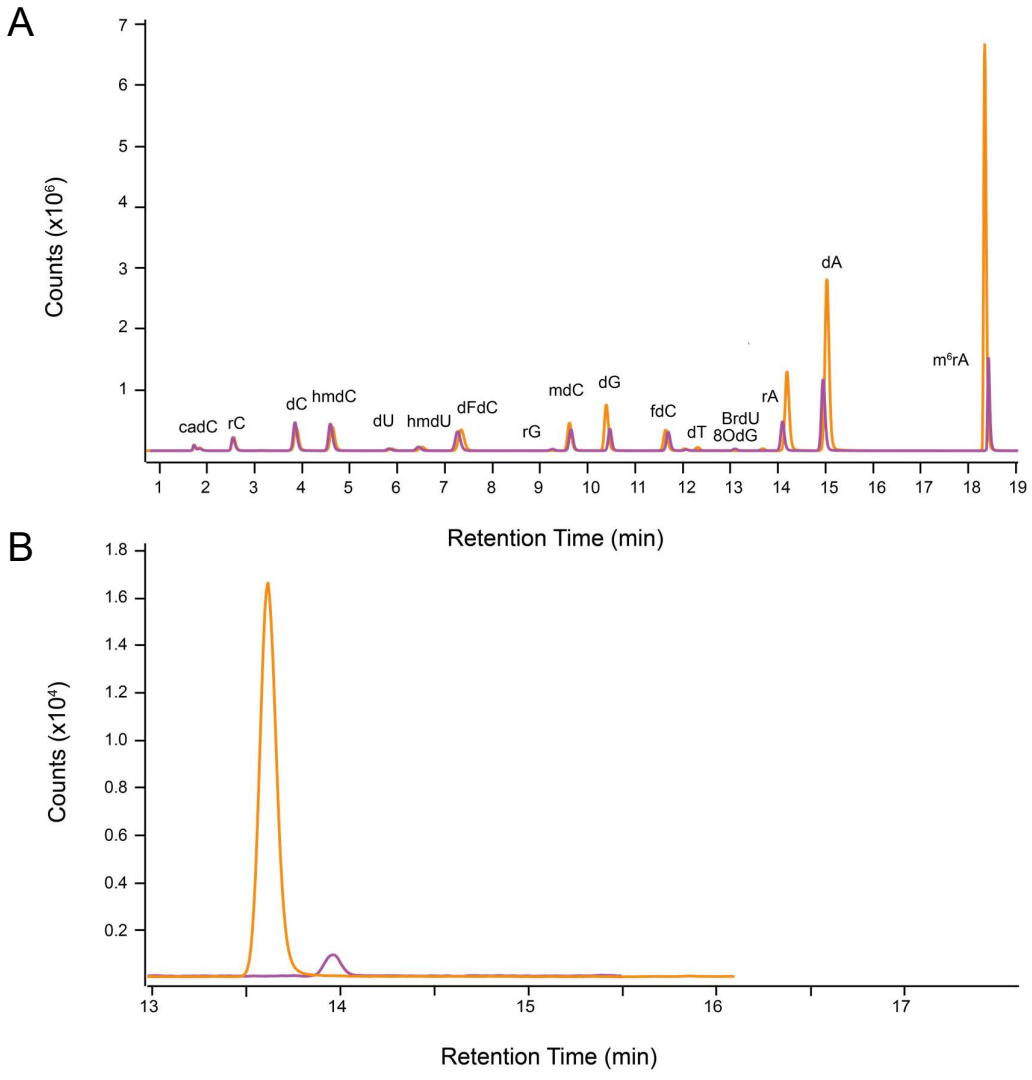


Figure 3.8 Changing Solvent B composition from 40:60 ACN:H_2O to 100 % MeOH increases BrdU signal, along with other nucleosides. *The nucleoside signal of 500 fmol standards was compared for solvent B 40:60 ACN:H_2O (magenta) vs solvent B MeOH (yellow).* ***A,*** *total ion chromatogram (TIC, summation of all individual EICs) of nucleosides;* ***B,*** *EIC of BrdU.*

The increase in BrdU signal allowed composition analysis of the collaborator's samples, with which she validated her sequencing method on the MinION and published in 2019.[209] The increase in signal across most ions led me to permanently change solvent B to 100 % MeOH.

3.3.3 MRM transition choice for maximising sensitivity

To use MRM, I needed to identify and optimise precursor-to-product transitions with, which I could engage in targeted identification and quantitation. I studied the scan spectrum of the nucleosides and found that in addition to the proton, sodium and potassium adducts formed, there were *m/z* ratios that corresponded to nucleoside dimers and the nucleobase (**Figure 3.9A**).

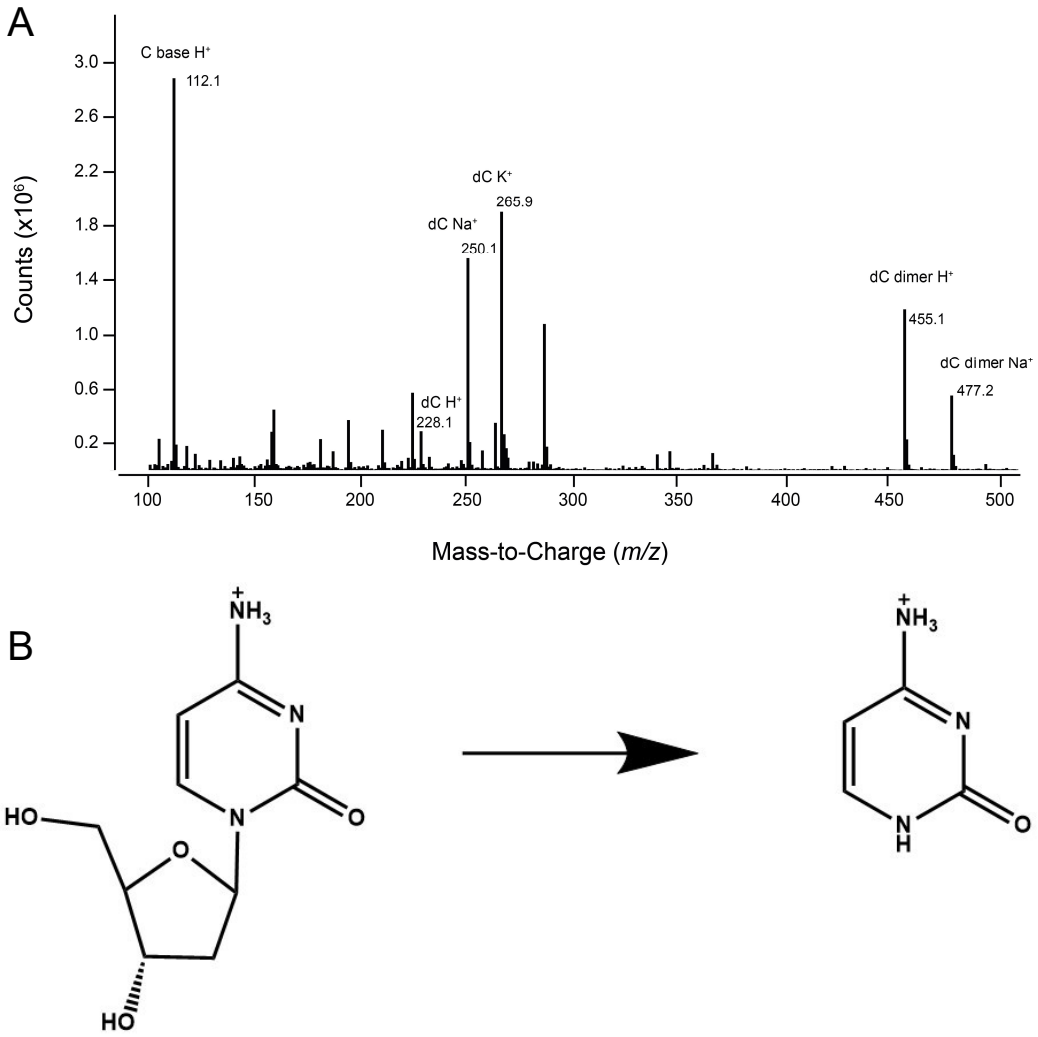


Figure 3.9 Search for a suitable parent ion for MRM. A, *full scan ion chromatogram of dC ion adducts;* ***B,*** *schematic of beta-glycoside fragmentation.*

The nucleobase signal is due to in-source beta-glycosidic linkage fragmentation of the nucleoside (**Figure 3.9B**). Its strong signal implied that it was an easily occurring fragmentation. Furthermore, the nucleoside to nucleobase transition is ubiquitously used in literature for nucleoside MRMs[182,195,210].

I verified this transition's potential for MRM by using the QQQ Optimiser software to determine for each nucleoside which product ions are formed upon nitrogen gas collision and

at which collision energy are the fragment ion signals maximised (data not shown). The software confirmed the applicability of the nucleoside to nucleobase transition.

A potential issue of MRM is the difficulty in quantifying analytes that are isobaric (have the same *m/z* values). The software may only register the analyte with the stronger signal and discard the other analyte. This has implications for nucleosides as some nucleosides have only a difference in *m/z* of 1, such as dC and dU. dU is a rare component of DNA and is orders of magnitude less abundant than dC; the first heavy isotope of dC through natural abundance will give large enough a signal to confound quantification of dU (**Figure 3.10**).

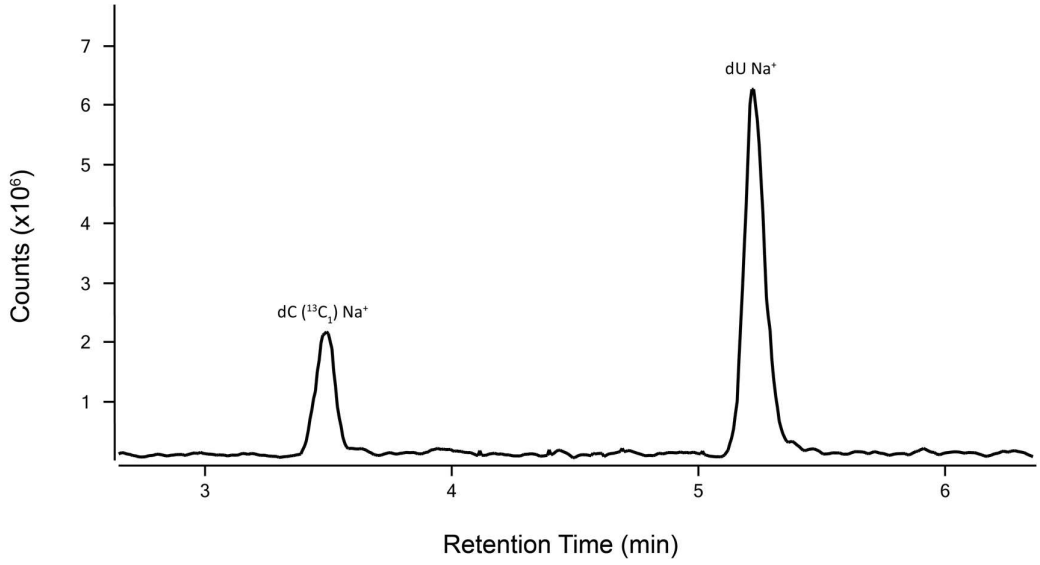


Figure 3.10 Ions with the same m/z interfere with each other's quantification. *Extracted MRM ion chromatogram of m/z = 251 in 1.25 pmol of standards, showing both dU Na^+ and dC +1 Da Na^+. In biological samples, dC+1 Da Na^+ will have a much higher signal than dU Na^+.*

This problem is circumvented by using dynamic MRM mode (dMRM).[211] dMRM sets detection windows for each nucleoside and so if nucleosides with similar *m/z* ratios are not co-eluting, then it is possible to quantify each nucleoside without interference from the other. I therefore created a dMRM method in which I recorded the transitions of the proton and salt adduct nucleoside ions, which are shown in Appendix A, **Table 3**, together with collision energies.

The dMRM EICs of a selection of the nucleosides under analysis are illustrated in **Figure 3.11**.

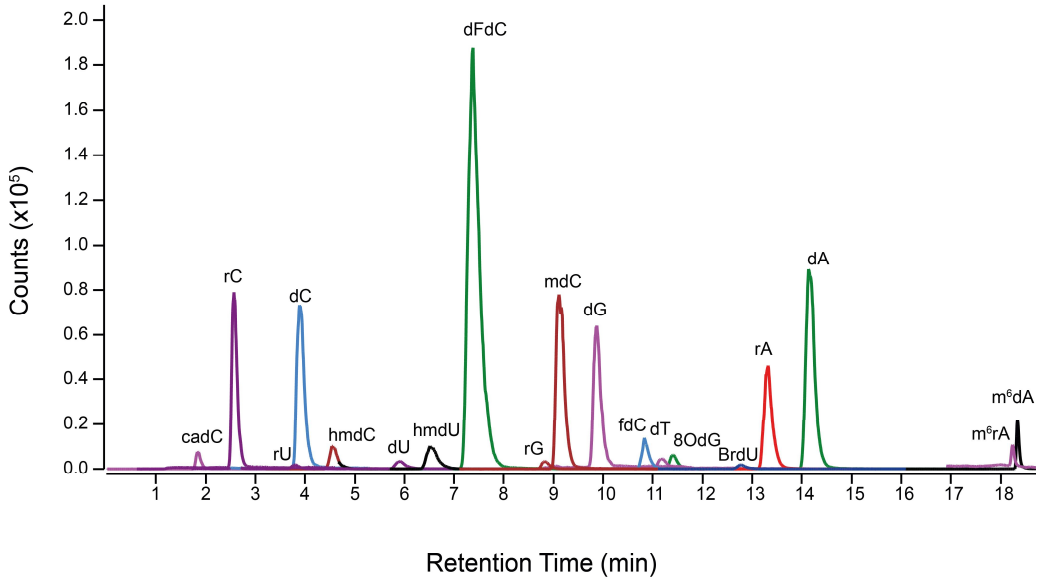


Figure 3.11 Nucleoside separation post MS and HPLC optimisation. *Extracted dMRM ion chromatograms of nucleosides from a 500 fmol standard injection.*

3.3.4 Optimisation of MS source conditions

The remaining MS optimisation was that of QQQ source conditions that affected performance. The MS source is responsible for transforming the analytes into ions and the instrument allows changing adjusting the parameter values.

The parameters that can be changed are:

- Gas Temperature
- Gas Flow
- Nebuliser
- Sheath Gas Temperature
- Sheath Gas Flow

- Capillary Voltage
- Nozzle Voltage
- iFunnel parameters (the ion funnel removes gas from the ion stream by concentrating the ions with RF voltage and passing them through slightly offset funnels.)

I optimised their values with the Agilent Source Optimizer software, thus finding the optimal value for each parameter through variation in successive runs.

The nucleosides behaved similarly with changes in parameters (**Figure 3.12** and **Table 3.4**), but when differences in parameters value preference were apparent, I prioritised the signal of dU, the nucleoside with the lowest sensitivity.

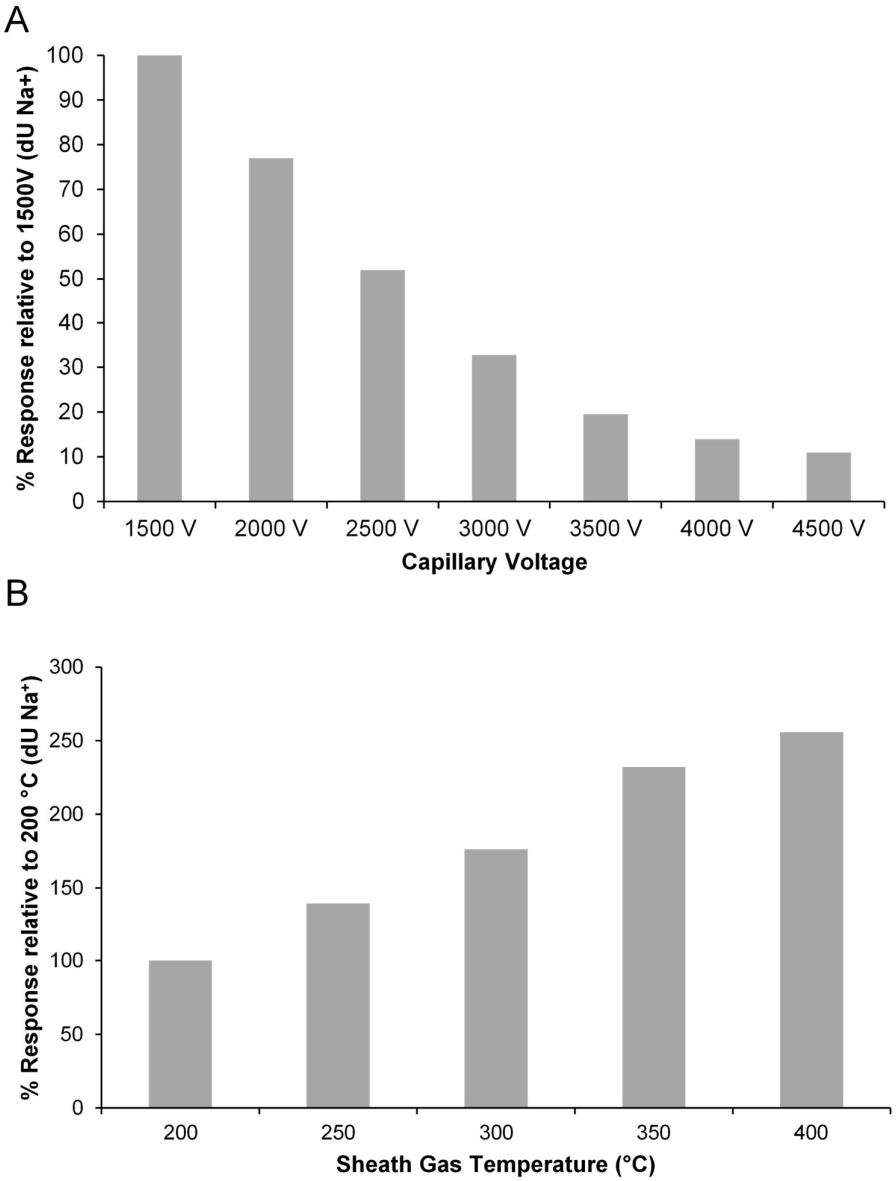


Figure 3.12 Varying MS source parameters allows identification for best nucleoside signal. A, *dU Na+ response vs capillary voltage;* ***B,*** *dU Na+ response vs sheath gas temperature.*

Parameter	Optimised Value
High Pressure RF	110 V
Low Pressure RF	80 V
Gas Temperature	230 °C
Gas Flow	14 L/min
Nebuliser	20 psi
Sheath Gas Temperature	400 °C
Sheath Gas Flow	11 L/min
Capillary Voltage	1500 V
Nozzle Voltage	0 V

Table 3.4 MS source parameters optimised for nucleoside signal.

To measure the sensitivity gained from scan to dMRM mode with optimised source conditions, I ran standards with 1.25 pmol injected onto the column (**Figure 3.13**). For dC, the dMRM peak had a signal-to-noise ratio of 1123, vs 92.5 in scan mode, demonstrating dMRM mode's power even at abundant, picomole amounts on column.

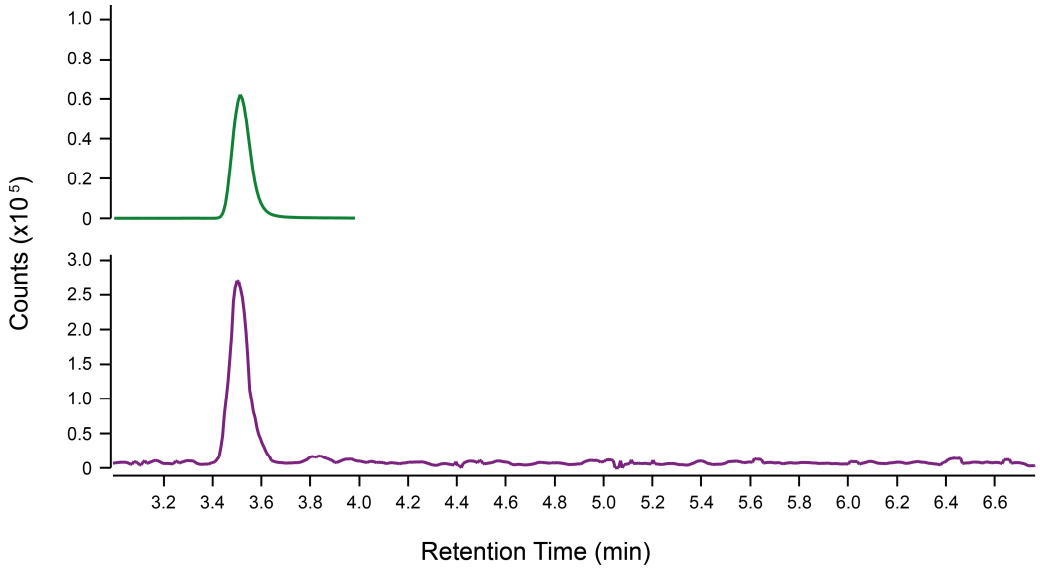


Figure 3.13 dMRM mode gives more than 10 x higher signal-to-noise ratio than scan mode. *EIC of* ***scan mode*** *dC H+ (green, 1.25 pmol, dMRM, SNR=1123) and EIC of* ***dMRM mode*** *dC H+ (purple, 1.25 pmol, MS, SNR=92.5).*

3.4 Use of internal standards for consistency in quantitation

A disadvantage to AJS ESI is that the distribution of ion adducts (protonated, sodiated, potassiated, etc), as well as the ionisation efficiency, depend on parameters including salt concentrations in the solvents, salt and organic material in the samples (matrix effect), salt and organic deposits in the source, MS air flow and external temperature. As a result, the sensitivity of the analytes can change dramatically from the beginning to the end of a run.

This posed an obstacle to nucleoside quantitation; to understand the extent of the problem, I ran external calibration curves at the start and at the end of a sequence of 64 biological sample analyses and compared the responses of the nucleosides as well as the distribution of their adducts. As the concentrations of the samples are determined by comparison against a standard curve, the results could only be valid if the standard nucleoside responses were consistent from the start to the end of the run. From the beginning or the run to the end of the run, during which I injected biological samples, there was a greater than 10 % change in signal for both dC H^+ and dC Na^+ (**Figure 3.14**): an increase of 16.5 % in dC H^+ signal and a

decrease of 10.2 % in dC Na^+ signal. This reduced confidence in quantitation through a standard curve at a different time in the run compared to the sample.

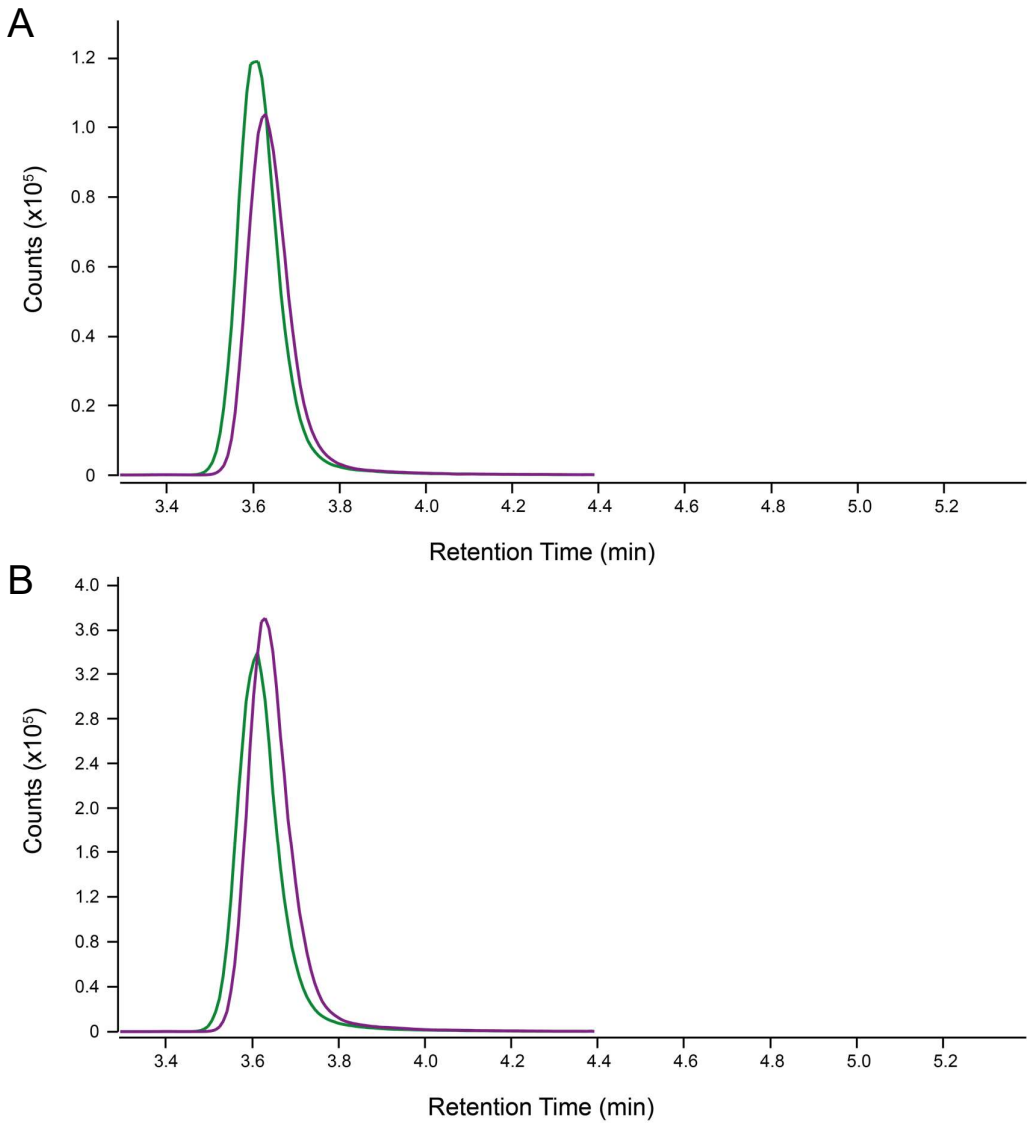


Figure 3.14 Nucleosides are liable to changing ion adduct preferences over the course of multiple chromatographic sequences. *Comparisons of the EIC of dC ion adducts, at start of worklist (purple) and end of worklist (green), with 500 fmol injections.* ***A,*** *dC H^+;* ***B,*** *dC Na^+.*

To adjust for this problem, I used internal standards. Internal standards involve the addition of stable-isotope-labelled nucleosides to the samples and standards used in the calibration curve, a technique previously used in nucleoside analysis.[182,195] The heavy nucleosides should be chemically near-identical to the light nucleosides and so will have the same ionisation

efficiency, adduct distributions and response. By spiking in heavy labelled standards to both the sample and the external calibration standards at the same concentration, it is possible to compare the ratio light against heavy nucleoside in both the calibration standards and the samples. This ensures consistency, removing concerns about changes in instrument performance and the effect of the sample matrix.

I acquired isotopologues for all the nucleosides I needed to quantify for collaborative purposes and added them to my standards at a fixed concentration of 12.5 nM.

3.5 Determining nucleoside limits of detection

Once the system was optimised for sensitivity and separation, I set out to find the nucleoside limit of detection (LOD, signal-to-noise ratio= 3) and subsequent lower limit of quantification (LLOQ, signal-to-noise ratio= 10). Using internal standards, I ran a sequence of standards increasing in amount injected by a factor of $\sqrt{10}$ from 158 amol to 500 fmol to obtain a calibration curve, with a constant internal standard amount injected of 62.5 fmol.

The calibration curve for the most responsive standard hmdC gives the LOD at 158 amol (**Figure 3.15 A** and **B**). Appendix A, **Table 2** gives the complete list of the standards used, their LODs and their LLOQs. I noticed that the internal standard for hmdC had a shorter retention time by around 10 seconds (**Figure 3.15C**) compared to the unlabelled standard. This may be explained by the difference in hydrogen bonding strength between light ^{14}N in the unlabelled standard and ^{15}N in the labelled internal standard, affecting subsequent interactions with the mobile phase and analyte retention time.

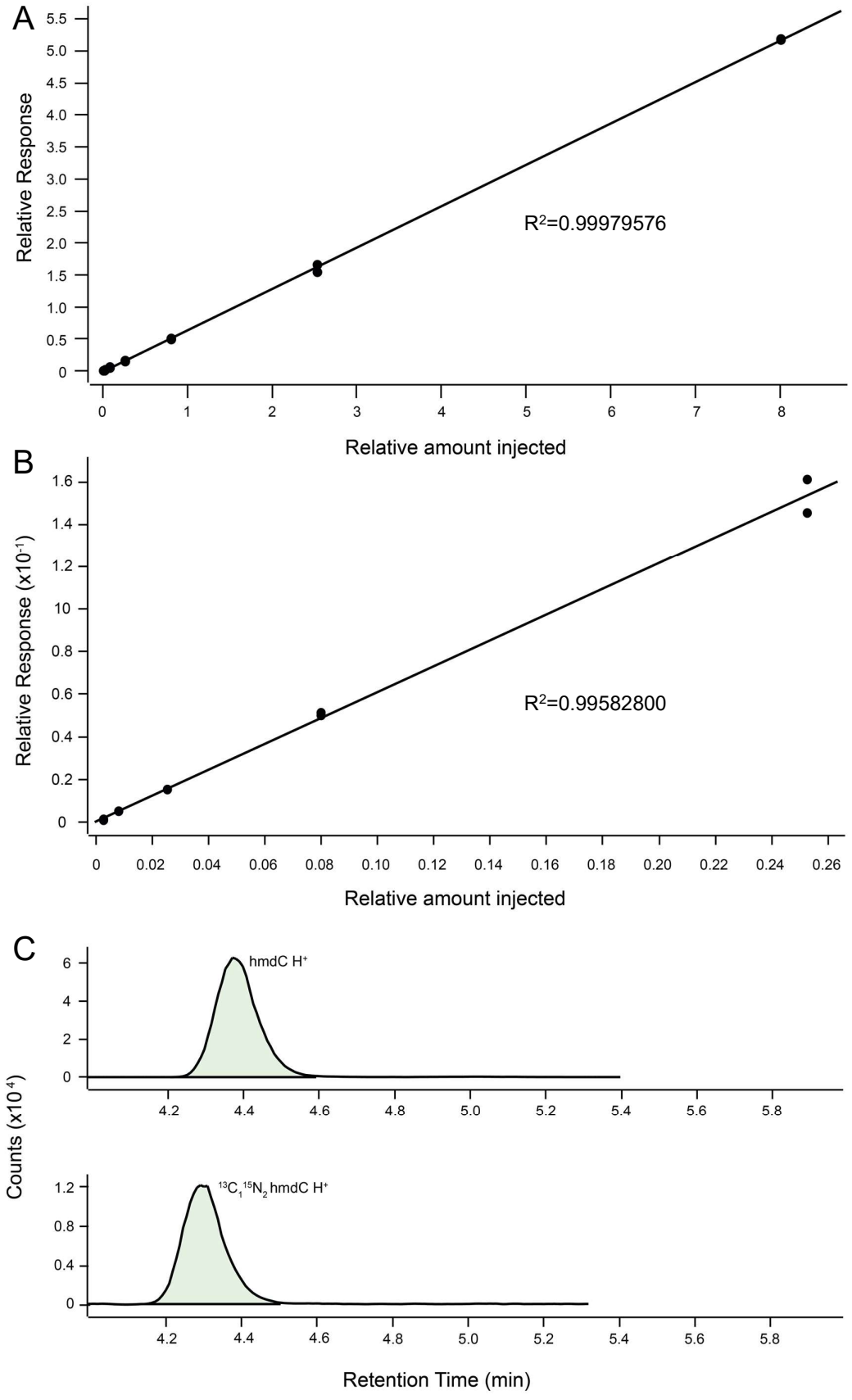

A
Relative Response
R²=0.99979576
Relative amount injected
B
Relative Response (x10⁻¹)
R²=0.99582800
Relative amount injected
C
hmdC H⁺
¹³C₁¹⁵N₂ hmdC H⁺
Counts (x10⁴)
Retention Time (min)

Figure 3.15 Calibration curves for hmdC with internal standard. *Two injections per amount on column, normalised to proportion compared to internal standard injected (62.5 fmol).* ***A,*** *hmdC H^+ whole curve;* ***B****, hmdC H^+ curve at low amounts;* ***C****, comparison of hmdC H^+ signal (500 fmol) with $^{13}C_1{}^{15}N_2$ hmdC H^+ (62.5 fmol).*

3.6 Combination of HPLC-UV and MS analysis

The LC-MS/MS system was optimised for sensitive quantitation of rare nucleosides or heavy nucleoside incorporation; however, for epigenetic analysis, the rare nucleosides are normalised against abundant nucleosides such (dC, dG, dT and dA for DNA). These are not well quantified by MS, as they saturate the ionisation step and cannot be measured precisely (**Figure 3.16A**). By contrast, the corresponding signal from UV does not saturate (**Figure 3.16B**).

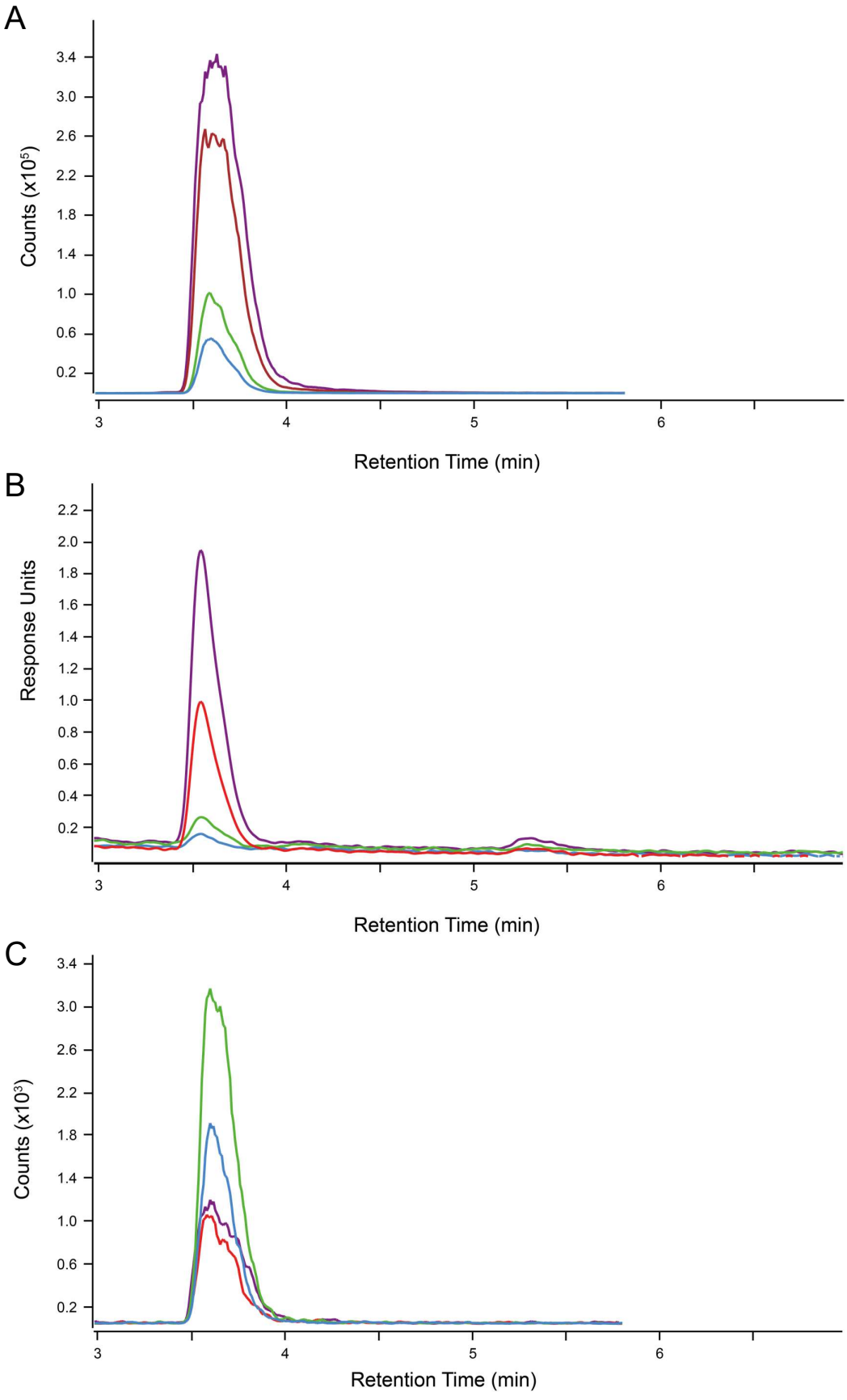

A
Counts (x10^5)
3.4
3.0
2.6
2.2
1.8
1.4
1.0
0.6
0.2
3
4
5
6
Retention Time (min)
B
Response Units
2.2
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
3
4
5
6
Retention Time (min)
C
Counts (x10^3)
3.4
3.0
2.6
2.2
1.8
1.4
1.0
0.6
0.2
3
4
5
6
Retention Time (min)

Figure 3.16 dC MS signal saturates at 1.25 pmol and above, while UV signal does not. *Colour keys: 1.25 pmol (blue), 2.5 pmol (green), 12.5 pmol (red) and 25 pmol (purple) injected on the column.* ***A,*** *MS signal;* ***B,*** *UV absorbance;* ***C,*** *internal standard ($^{13}C_1{}^{15}N_2$ dC). The four calibration points were obtained by 5 and 10 µL injections of 250 nM and 2.5 µM standards, so therefore the 10 uL injections have a stronger internal standard signal than the 5 µL injections (twice the amount injected).*

The saturation is also seen in the response of the internal standards (**Figure 3.16C**). The use of internal standards averts this problem at low pmol amounts injected on column, as the isotopologue will be suppressed in the same proportion to the light nucleoside. However, beyond that amount the peaks lose definition through saturation and therefore the peak integration and resulting calibration curve are no longer ideal.

This problem was resolved by combining MS and UV detection. As described in literature, the rare nucleosides would be analysed through MS and the abundant nucleosides would be analysed by UV.[182]

Before the advent of HPLC-MS/MS, global nucleic acid composition analysis was carried out with HPLC-UV. It is less sensitive but capable of analysis of greater quantities of material on column. The advantage of MS and UV combination for nucleoside analysis are that the sample analysis workflow is unchanged, with no requirement for sample reanalysis at a different concentration. This saves time and avoids potential sources of human error.

Figure 3.17 shows the UV calibration curves of dC and the MS calibration curve of hmdC. The difference in magnitude of the LLOQ of hmdC and the HLOQ of dC gives a theoretical lower limit of proportion of hmdC of dC at 1 ppm.

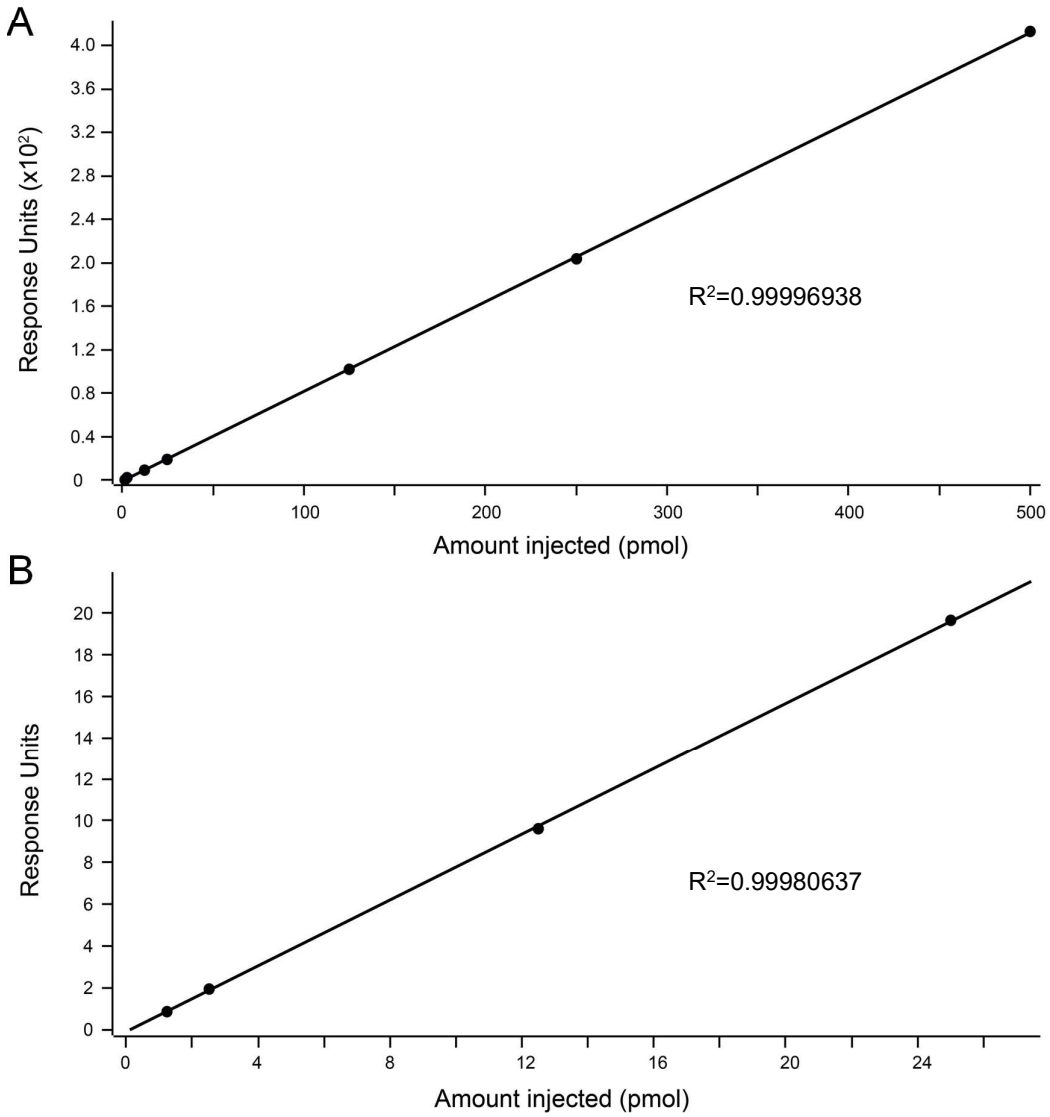


Figure 3.17 UV Calibration curves for dC, one injection per concentration. ***A,*** *dC whole curve;* ***B,*** *dC H curve at low amounts.*

3.7 Validation of MS sample analysis stability

Once the workflow was ready for sample analysis, I validated the analytical system's reproducibility by analysing the global dC methylation (mdC) proportion as a percentage of dG in the DNA of the cell line derived from HEK-293. I performed the hydrolysis on different days over the span of several months (**Figure 3.18**).

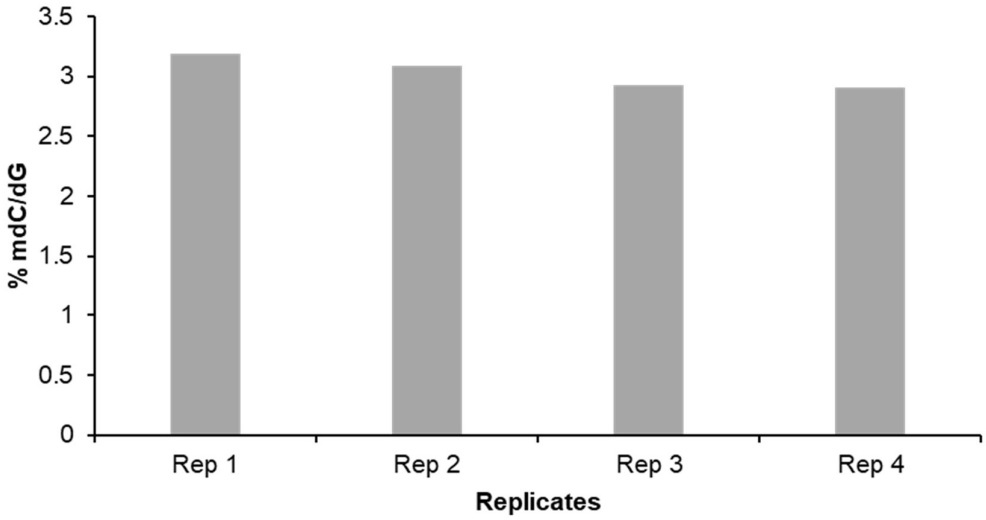


Figure 3.18 % mdC/dG is stable over different measurements. *500 ng HEK-293 TREX-TET-CD-GFP +Dox, n=4 replicates hydrolysed on different days.*

Over four separate runs in the space of months, the % mdC/dG proportion did not change by more than +/-10 %, indicating the instrument gives reproducible measurement of biological samples over time.

3.8 Discussion

In this chapter, I have described the work done to set up an LC-MS/MS method for nucleoside analysis, optimising for sensitivity and nucleoside separation.

3.8.1 Nucleic acid hydrolysis optimisation

I adapted a nucleic acid hydrolysis process suitable for UV analysis and therefore capable of hydrolysis of tens of micrograms of DNA.[62,201] As MS analysis has greater sensitivity, I only hydrolysed nanogram amounts of DNA. I therefore investigated whether nucleic acid

hydrolysis was successful at shorter incubation times and with lower concentrations of phosphodiesterase I (PDE I). I showed that the incubation time could be reduced from 24 hours to 1 hr (**Figure 3.1**) and that PDE I (**Figure 3.2**) can be diluted fourfold without causing issues in hydrolysis.

I have not tested having lower incubation times or using less PDE I or less of any other reagent. The enzymatic reagents remain in the sample and may cause blockages in the HPLC column. It may be possible to reduce the amount of reagent further without sacrificing hydrolysis efficacy, though this will require a retesting of the optimal hydrolysis time in that parameter.

3.8.2 Optimising HPLC analyte separation

Analyte separation was important as certain nucleosides like dC and dU have similar *m/z* and so would interfere with each other's quantification if they elute at similar times. Additionally, abundant nucleosides would suppress the signal of co-eluting rare nucleosides through ionisation competition. To improve nucleoside resolution and minimise analyte co-elution, I tested nucleoside separation with three separate columns, finding that the Agilent Zorbax Eclipse Plus C18 gave most analyte resolution (**Figure 3.4**).

3.8.3 Optimising MS for nucleoside signal

Solvent composition

I investigated the best solvent system for nucleoside analysis in positive ionisation mode, testing the use of the ion pairing agent TBA and varying the pH of solvent A (**Figure 3.6** and **Figure 3.7**). TBA caused signal saturation and corresponding ion suppression, while 10 mM

ammonium acetate gave the best signal at pH 6. The signal was improved further by changing solvent B from 40 % ACN to 100 % MeOH (**Figure 3.8**). There may be further improvement in signal possibly by investigating whether the effect of other solvent A compositions, such as ammonium formate. Additionally, it may be that the addition of matrix derived from hydrolysis mix to the standards could lead to a better calibration and is worth exploring.

MS parameters

The 6495a QQQ is most sensitive in targeted MRM mode, involving the filtration of an analyte parent ion in Q1, its fragmentation with nitrogen gas in Q2 and the filtration of target ions in Q3. A commonly used transition in literature for nucleoside MS analysis is the charged nucleoside fragmenting to the charged nucleobase, so I verified that the charged parent ions adducts were present in the dC peak (**Figure 3.9**) and used them to create a dMRM method together with the analyte retention times.[182,195,210] I optimised the dMRM method by varying the source parameters to obtain the strongest nucleotide signal (**Figure 3.12**).

Optimising for instrument consistency

Over an analysis sequence, the instrument may lose or gain sensitivity due to the accumulation of salts from the solvent bottles and from samples (**Figure 3.14**). To avoid the issue of variability in sensitivity, I added stable isotope labelled internal standards in both samples and external calibration to maintain calibration consistency, a previously reported solution.[195]

At high amounts of sample injected on column, abundant analytes in DNA like dC saturate the MS detector, affecting sample quantification. This problem is resolved by using HPLC-UV analysis for abundant nucleosides and HPLC-MS/MS for rare nucleosides.[182]

The optimised LC-MS/MS method gives analyte LLOQs ranging from 158 amol to 50 fmol (**Figure 3.15**) and the UV analysis increases the proportion of the rare hmdC nucleoside to the abundant dC nucleoside to the potential dilution to a 1,000,000 fold range (**Figure 3.17**).

There are possibilities for improving instrument sensitivity. The greatest signal in scan mode was from the nucleobase (**Figure 3.9A**), which was used as the parent ion by Zhang et al.[212] I did not use this transition and it may increase sensitivity further.

I also did not increase the delta EMV (the electron voltage multiplier) beyond 200 V. Increasing it further could increase the signal-to-noise ratio of the nucleosides, increasing sensitivity.

This notwithstanding, the current signal and separation are suitable for my experiments on the roles of the deaminases on DNA incorporation.

3.8.4 Stability of nucleoside sample analysis

To verify that the LC-MS/MS system gave consistent nucleic acid composition analysis, I hydrolysed HEK-293 cell-derived DNA on four separate days and compared the % mdC/dG (**Figure 3.18**). The proportion was constant, underlining the stability of the LC-MS/MS analysis system.

4. Setting up a workflow for nucleotide analysis

The LC-MS/MS system described in the previous chapter allows sensitive nucleic acid composition analysis for isotopic labelled tracing in DNA/RNA and could aid the investigation of my project enquiries: whether either CDA or DCTD is the primary cytidine nucleotide deaminase and whether newly synthesised cytidine nucleotides are deaminated by them.

It was also desirable to directly analyse the free nucleotide pool, as CDA deaminates dC and rC, whereas and DCTD deaminates dCMP. Analysis of the nucleotide pool would give information on the biological effect of the deaminases on nucleotides with different phosphorylation states. However, nucleotide measurement by mass spectrometry has been experimentally challenging, due to their similar properties impeding chromatographic separation and their chemical instability.[213] Several approaches have been described, measuring the cellular nucleotide pool by polymerase experiments, luminescence-based detection of ATP, NMR and mass spectrometry.[58,164,214–216]

The advantage of mass spectrometry-based detection is the sensitivity and mass discrimination of mass spectrometry measurements, allowing for isotopic labelling experiments in a cellular context. Measuring the labelling of the free nucleotide pool over time in a pulse-chase experiment will give information on the kinetics of dC and dCMP deamination.

LC-MS has been used in both targeted and untargeted modes to analyse the nucleotide pool.[177,196,215,217] Untargeted analysis gives information on the wider metabolome, at the expense of quantification and sensitivity. Targeted MS/MS analysis allows quantification at lower amounts of the restricted set of metabolites that is the nucleotide pool.

Considering that my research questions were limited to the nucleotide pool and the roles of CDA and DCTD, I decided to develop a targeted method optimised for accurate and sensitive nucleotide detection.

This chapter describes the work carried out in this thesis to set up an LC-MS/MS based method to accurately measure nucleotides and their isotope labelled standards using the Agilent 1290 HPLC-6495a QQQ available in the Kessler lab (TDI). It begins with optimisation of the extraction process of the nucleotide pool from cell samples. This is followed by the work done to identify the most appropriate column and mobile phase gradient. Subsequently optimisation of the MS instrument, including charge polarity, signal dependence on solvent composition, and instrument parameters is described. Finally, I illustrate efforts to validate the system for nucleotide analysis in samples.

4.1 Nucleotide pool extraction optimisation

As stated in **1.9.2.1**, there are several thousand metabolites in each cell and many share the same polar properties as nucleotides. The co-extraction of these non-nucleotide metabolites from cells may then suppress the signal of nucleotides through ion competition in the MS source. Extraction methods from literature to remove non-nucleotide metabolites involve significant processing, such as the use of solid phase extraction cartridges.[218,219] The benefit of sample processing is that co-extracted metabolites and salts that would interfere with nucleotide ionisation are removed, but more involved processes reduce sample yield and could potentially dephosphorylate the nucleotides. My aim was to extract the maximum amount of nucleotides from cells, without degrading the multiply phosphorylated nucleotides and disrupting the biological ratios of monophosphates to di- and tri-phosphates.

4.1.1 Solvent extraction

When extracting metabolites from cells, it is important to stop cellular metabolic processes to freeze the state the analytes are in (metabolite quenching), crucial for rapidly metabolised molecules such as nucleotides.[179,220,221]

This is generally done by washing cells with PBS and then snap freezing. The extraction is then achieved through addition of solvents that lyse the cells and dissolve the metabolites of interest. For polar metabolites such as nucleotides, polar solvents such as water and methanol are used, sometimes in combination with other solvents.[165,167]

I tested a selection of the extraction mixtures to optimise nucleotide extraction (***Table 4.1***).

Extraction method	Solvent mixture composition	Use in literature
1	80 % MeOH, 20 % H_2O	Walsby-Tickle et al.[222]
2	60 % MeOH, 20 % H_2O	Wilson et al.[223]
3	50 % ACN, 50 % H_2O	De Val et al.[224]

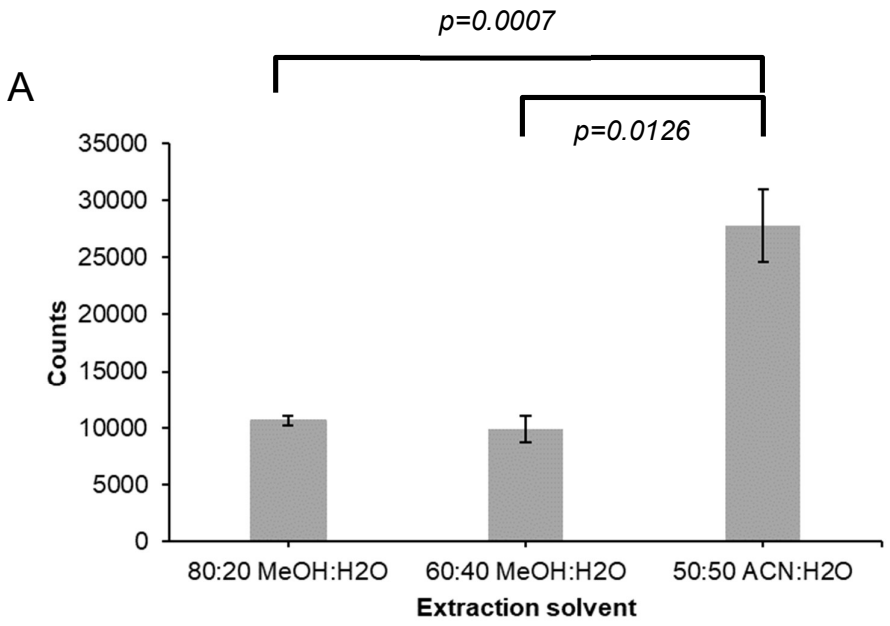
Table 4.1 Solvents tested for pool extraction.

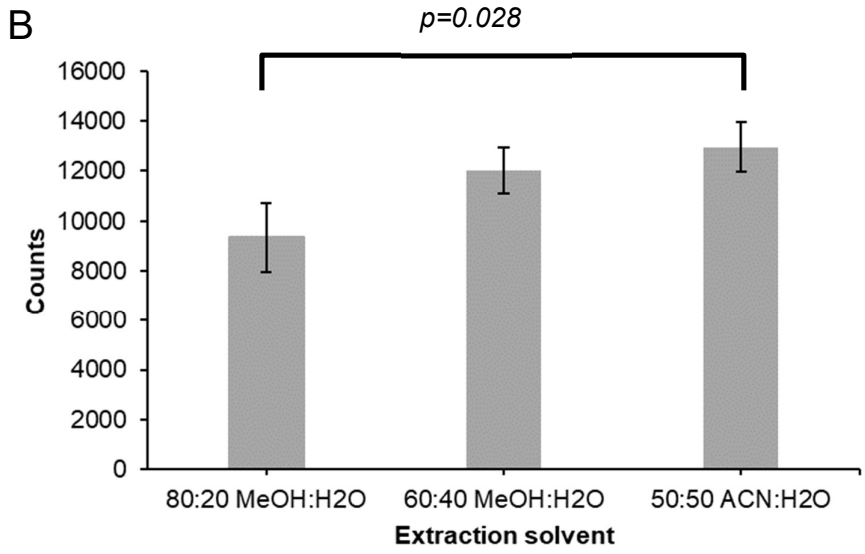
For each solvent mixture, I followed the same workflow:

- Seed 100,000 H1299 cells per well in triplicates in complete DMEM in a 6-well plate.

- 24 hrs later, wash the cells twice in 2x PBS
- Snap freeze them in liquid nitrogen to stop metabolism
- Add 1000 µL extraction solvent to lyse extract metabolites
- Scrape cell material off wells, vortex and centrifuge at 4 °C to pellet insoluble cell material
- Collect 750 µL of the supernatant and dry in speed-vac
- Resuspend in 20 µL of nuclease-free water
- Inject 1 µL onto LC-MS/MS

The analysis was performed in targeted dMRM mode and I judged the efficacy of nucleotide extraction based on MS nucleotide signal. I obtained the best results for 50 % ACN for dCMP and dCDP (**Figure 4.1A** and **B**). dCTP was lower in 50 % ACN than 80 % MeOH (**Figure 4.1C**), but I do not think it is due to degradation as there is no concomitant decrease in dTTP signal (**Figure 4.1D**). I therefore used 50 % ACN as the extraction mixture. It is uncertain as to why certain nucleotides had different extraction efficiencies compared to the others depending on the extraction solvent used, it may be due to their electrochemical properties: the higher phosphorylated nucleotides dCTP and dTTP had a trend towards higher extraction efficiency in methanol and water, possibly due to the properties of their third phosphate.


A
p=0.0007
p=0.0126
35000
30000
25000
20000
15000
10000
5000
0
Counts
80:20 MeOH:H2O
60:40 MeOH:H2O
50:50 ACN:H2O
Extraction solvent


B
p=0.028
16000
14000
12000
10000
8000
6000
4000
2000
0
Counts
80:20 MeOH:H2O
60:40 MeOH:H2O
50:50 ACN:H2O
Extraction solvent

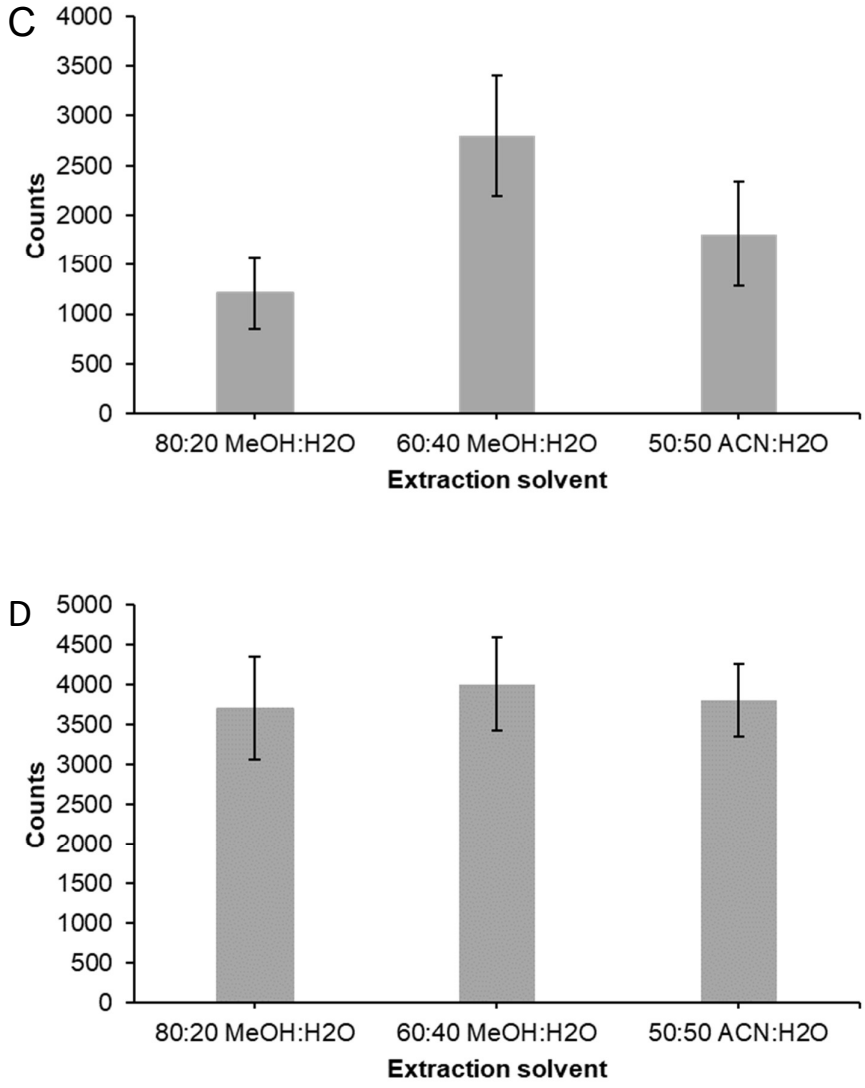


Figure 4.1 50 % ACN gave best response for dCMP and dCDP, mixed response from dCTP and dTTP. *Effect of extraction mixture choice on nucleotide extracted ion chromatogram signals, (n=3). A, effect of extraction solvent composition on dCMP; B, effect of extraction solvent composition on dCDP; C, effect of extraction solvent composition on dCTP; D, effect of extraction solvent composition on dTTP.*

Notably, the transition for dCMP showed a double peak (**Figure 4.1A**). To demonstrate that this was due to another species with the same transition as dCMP H^+ (308→112), I ran samples spiked from the start of solvent extraction with heavy-labelled dCMP (**Figure 4.2**). In that sample, one can see the labelled peak co-elute with the second peak, but not the first. It

is unclear what the non-dCMP peak is. It may be cyclic uridine monophosphate, but it would need to be verified with standards.

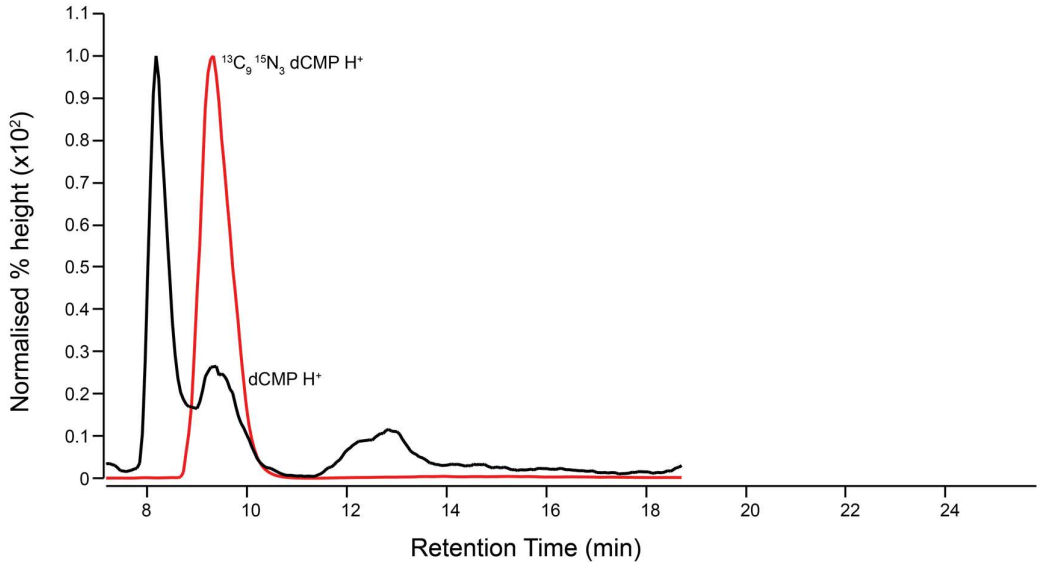
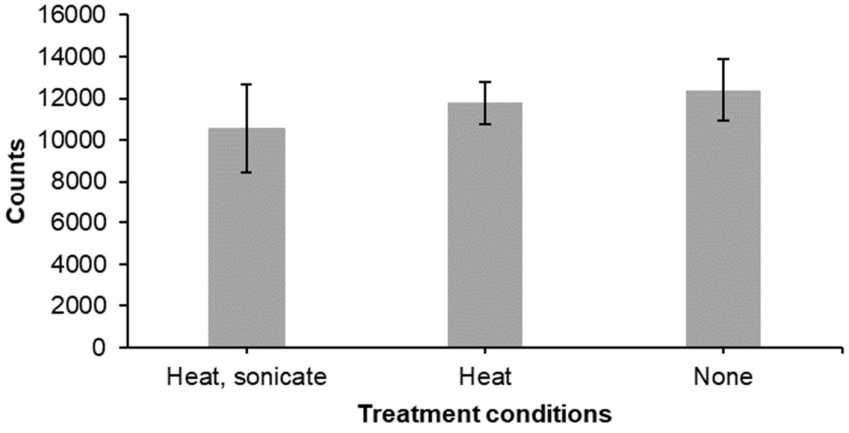


Figure 4.2 dCMP is second peak of 308→112 EIC. *EICs of labelled (red) and unlabelled dCMP H^+ (black) in H1299 cells. heavy labelled dCMP H^+ (37.5 pmol concentration, red) aligns with second peak of transition including dCMP H^+ unlabelled (black).*

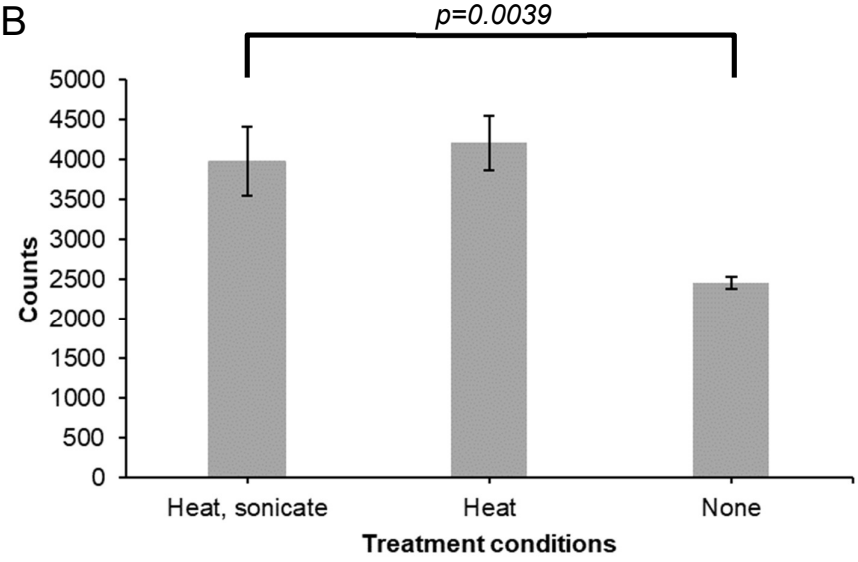
4.1.2 Sample treatment

My goal was to obtain the most nucleotides from my samples. To maximise the amount of nucleotides they extracted, Wilson et al. heated their samples to 95 °C and then sonicated them briefly.[214] I tested whether either would increase nucleotide signal, testing the combinations of sonicating and heating for one minute each (**Error! Reference source not found.**).

A


16000
14000
12000
10000
8000
6000
4000
2000
0
Counts
Heat, sonicate
Heat
None
Treatment conditions

B


p=0.0039
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
Counts
Heat, sonicate
Heat
None
Treatment conditions

C

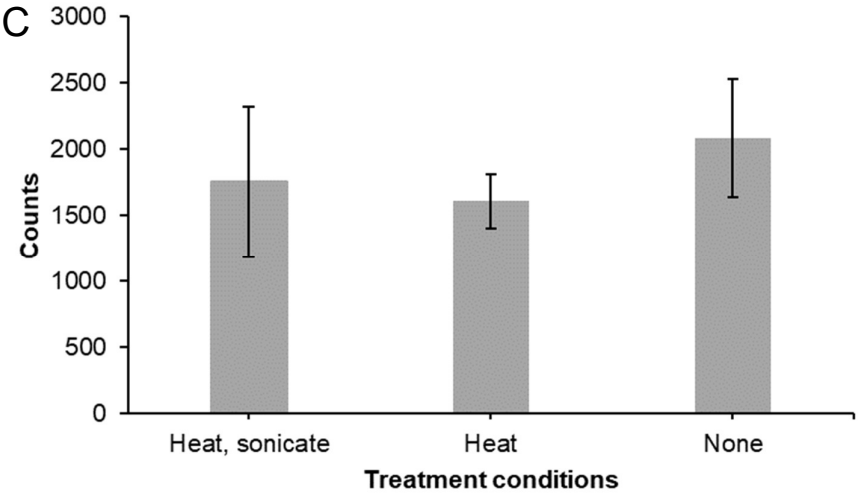

3000
2500
2000
1500
1000
500
0
Counts
Heat, sonicate
Heat
None
Treatment conditions

Figure 4.3 Heating samples increases nucleotide signal. Effect of sample treatment on nucleotide signal (n=3). A, *effect for dCMP;* ***B,*** *effect for dCDP;* ***C,*** *effect for dCTP.*

Heating the sample gave the best signal for dCDP, with no difference noted in either dCMP or dCTP. Sonication neither increased nor decreased dCDP signal; for these experiments I kept it and included heating and sonication to the workflow, although it seems to be unnecessary and can be removed from the workflow.

A concern was that the nucleotide extraction process may cause significant nucleotide triphosphate and diphosphate dephosphorylation and interfere with the signal of lower-phosphorylation nucleotides. To measure the extent of sample dephosphorylation, I spiked 1 µM of heavy labelled nucleotide internal standards into the extraction solvent and extracted the metabolites of H1299 cells. I compared the signal of the dCMP internal standard (320→119) with the dephosphorylated internal standard of dCTP (317→116) (**Figure 4.4**).

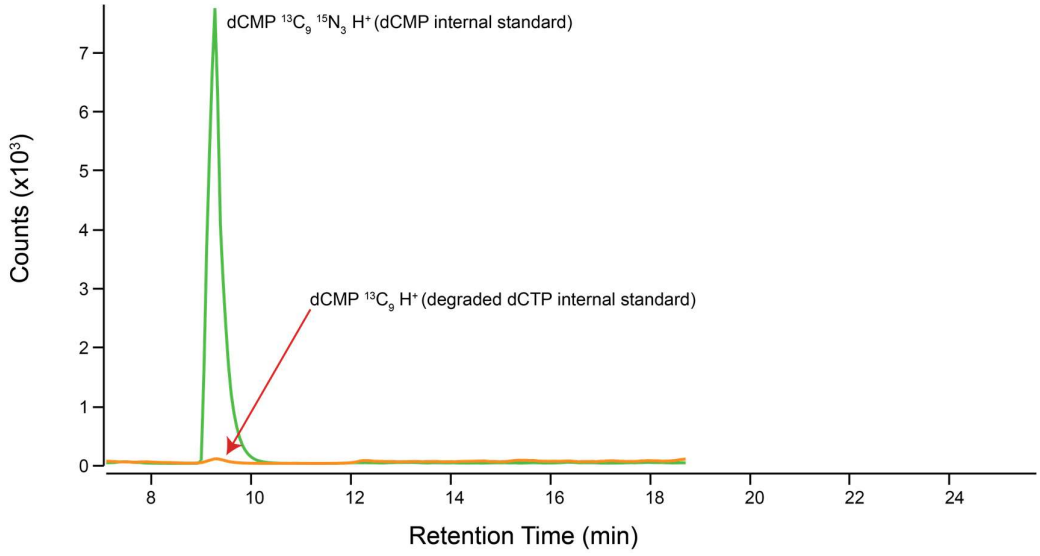


Figure 4.4 Metabolite extraction process does not cause significant nucleotide triphosphate dephosphorylation. *1 µM of internal standards were spiked into extraction solvent of nucleotides extracted from H1299 cells, with 1 µL of prepared sample injected. Comparison of the EIC of the internal standard of dCMP (320→119, green) with the EIC of the calculated dephosphorylated product of the internal standard of dCTP (317→116, yellow) to assess extent of nucleotide triphosphate degradation.*

The signal from dephosphorylated dCTP compared to that of the internal standard of dCMP was not large enough to be challenging for dCMP quantitation, as pyrimidine nucleotide mono-, di- and tri-phosphate concentrations are within an order of magnitude of each other.[2] The transition for the dCDP product derived from dCTP degradation to dCDP (397→116) was unfortunately not measured, not giving information about the extent of triphosphate to diphosphate degradation.

4.1.3 Injection volume

Lastly, I wanted to test the effect of injection volume on nucleotide signal. With nucleoside analysis, increasing the volume of injection increased signal due to the greater amount of analyte injected, until the system saturated. Additionally, I was using HILIC chromatography (stationary phase is hydrophilic and adsorbs polar metabolites), the injection of aqueous solution (necessary for dissolving nucleotides from samples) into an initial primarily organic mobile phase can disrupt the peak shapes and chromatography.[225] It would be preferable to minimise the volume of sample injected onto the HILIC column.

I tested the signal, retention time and peak shape for the nucleotides under different injection volumes, from 0.2 µL to 2 µL, showing the peak shape and response of dTTP as a representative nucleotide (**Figure 4.5**).

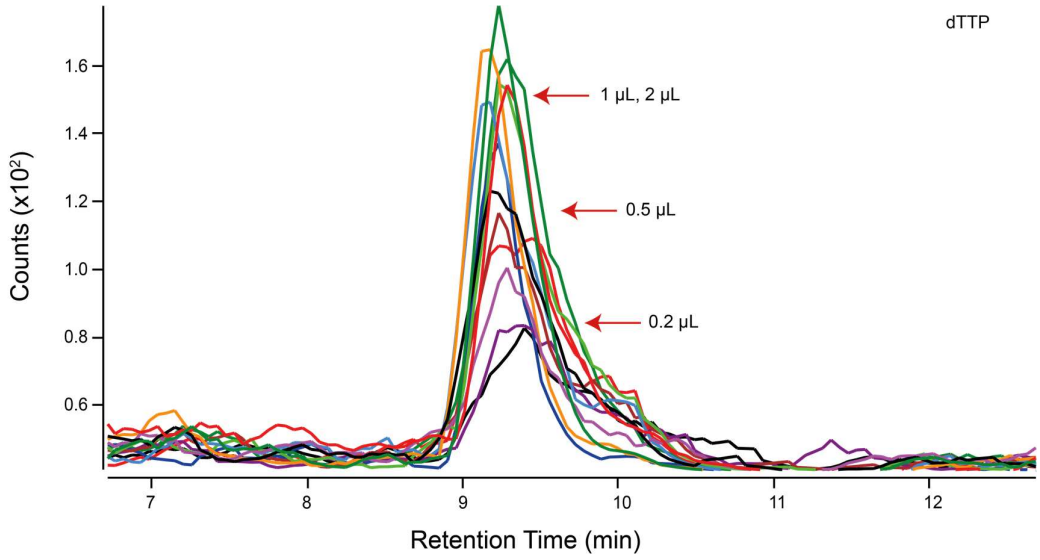


Figure 4.5 dTTP Na^+ signal plateaus at 1 µL injection. *Effect of injection volume on dTTP Na^+ signal (n=3).*

The retention times and peak shapes of the nucleotides did not change upon increasing volume injected. The signal increased from 0.2 µL to 1 µL, but then remained constant from 1 µL to 2 µL suggesting saturation was reached at within that range. I therefore moved forward injecting 1 µL per experiment.

4.2 Optimising HPLC for nucleotide analysis

4.2.1 Column

Due to their phosphate groups and nucleoside backbone, nucleotides are polar molecules with similar properties (**Figure 4.6**).

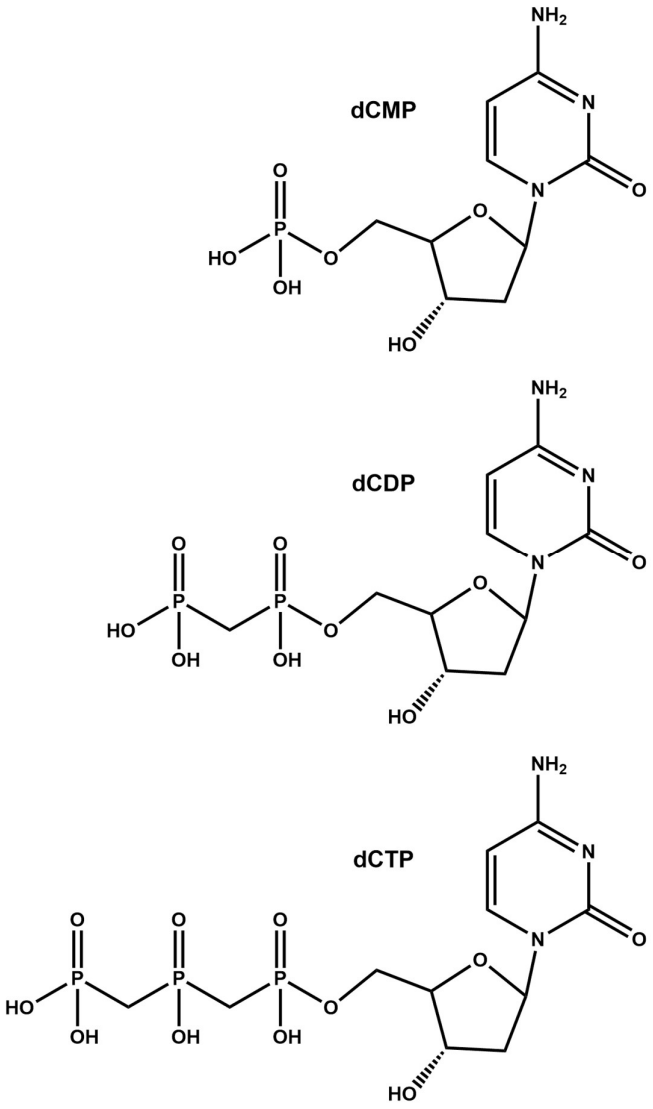


Figure 4.6 The nucleotides dCMP, dCDP and dCTP have similar structure and differ in the number of phosphates.

Their chemical similarity makes them difficult to separate by HPLC. Strategies to design appropriate chromatography for polar compounds include ion-pairing agents on a reversed-phase column, ion-exchange chromatography coupled with an ion suppressor, a functionalised reversed-phase column that interacts with the nucleotide structure, and hydrophilic interaction chromatography (HILIC).[169,196,222]

Ion-pair chromatography uses the interaction of long-chained ions such as tributyl ammonium in the mobile phase with both the hydrophobic stationary phase and polar analytes such as

nucleotides. This interaction increases the analytes' retention times and improves their resolution. However, ion-pairing agents are not suitable for mass spectrometry as they suppress analyte signal through ionisation competition.[226]

The instrument I am using is the same as that for nucleoside analysis, which does not have capacity for an ion suppressor. I therefore tested a reversed-phase column functionalised with pentafluorophenyl groups (PFP) and a zwitterionic HILIC column (**Table 4.2**).

Column	Type of column	Manufacturing Company	Length (mm)	Width (mm)	Particle size (µm)
1	Pentafluorophenyl (PFP) reversed-phase	Agilent Technologies[227]	150	4.6	3
2	HILIC-Z, Peek-lined to reduced adsorption	Agilent Technologies[228]	150	2.1	2.7

Table 4.2 Columns tested for nucleotide resolution.

The PFP column was designed to retain analytes such as nucleotides by polar interactions between its fluoride atoms and the nucleotides' polar phosphate groups, as well as aryl-aryl interactions between the phenyl groups and the pyrimidine and purine nucleobases (**Figure 4.7A**). The HILIC-Z column was designed to retain polar nucleotides through interactions between its bonded zwitterion group (a zwitterion is a molecule that is both positively and negatively charged with an overall charge of zero) and the nucleotides' polar phosphate groups (**Figure 4.7B**).

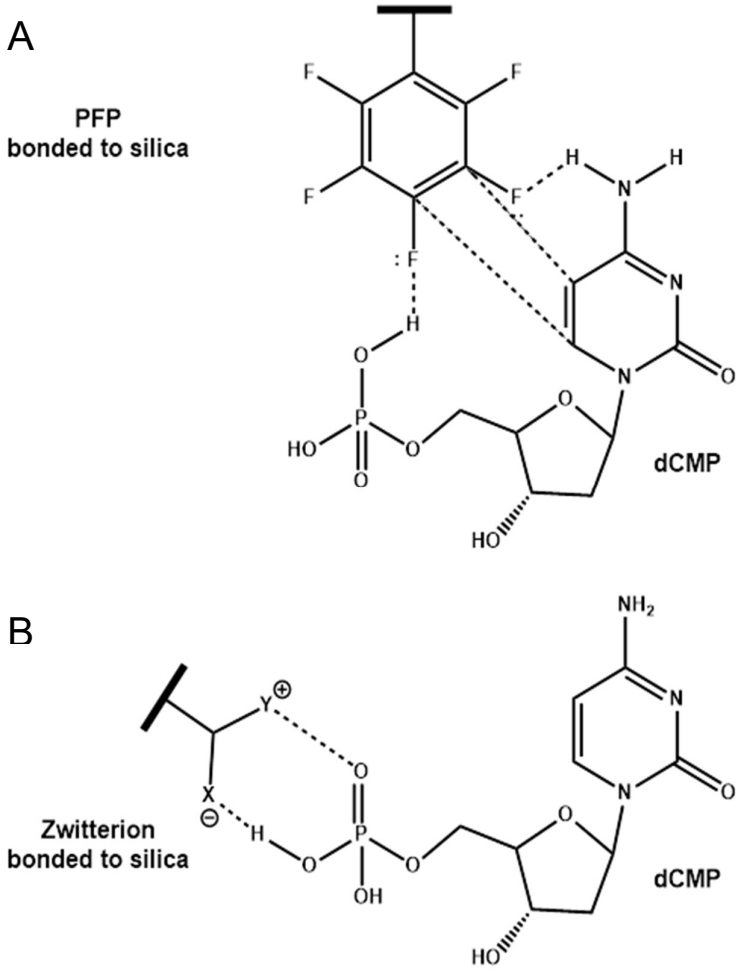


Figure 4.7 Schematic of nucleotide-column interactions with dCMP. A, *the PFP column retains polar aromatic compounds such as nucleotides through its dipole-dipole interactions between the fluorine atoms and phosphate hydrogens, as well as aryl-aryl interactions between the phenyl group and the nucleotide heterocycles;* ***B,*** *the HILIC-Z column retains polar compounds by dipole-dipole interactions between the bonded zwitterions and the charged nucleotide phosphates.*

I investigated nucleotide retention times on both columns, injecting standard mixes on the instrument and using UV absorbance as a readout. The PFP column solvent system was as follows: solvent A, 0.1 % formic acid in H_2O, solvent B, 100 % methanol. The HILIC-Z column solvent system was as follows: solvent A, 10 mM ammonium acetate in H_2O, pH 9 with ammonium hydroxide, 5 µM medronic acid deactivating agent; solvent B, 100 % ACN, with 5 µM medronic acid (methylene diphosphate) deactivating agent. Medronic acid was recommended by Agilent to reduce analyte adsorption onto the HPLC capillaries that were not bio-inert.[229]

The PFP column separated dCMP, dUMP and dGMP, (**Figure 4.8A**), but there was no signal of the multiply phosphorylated dCDP, dCTP and dGDP (**Figure 4.8B**). The only UV signal from the multiply phosphorylated nucleotides eluted at the same time as dGMP and so will have arisen from dGDP dephosphorylation to dGMP. By contrast, the HILIC-Z column showed non-overlapping peaks of dCMP, dCDP and dCTP and at longer retention times (**Figure 4.8C)**. This promised better resolution for the other nucleotides in the pool.

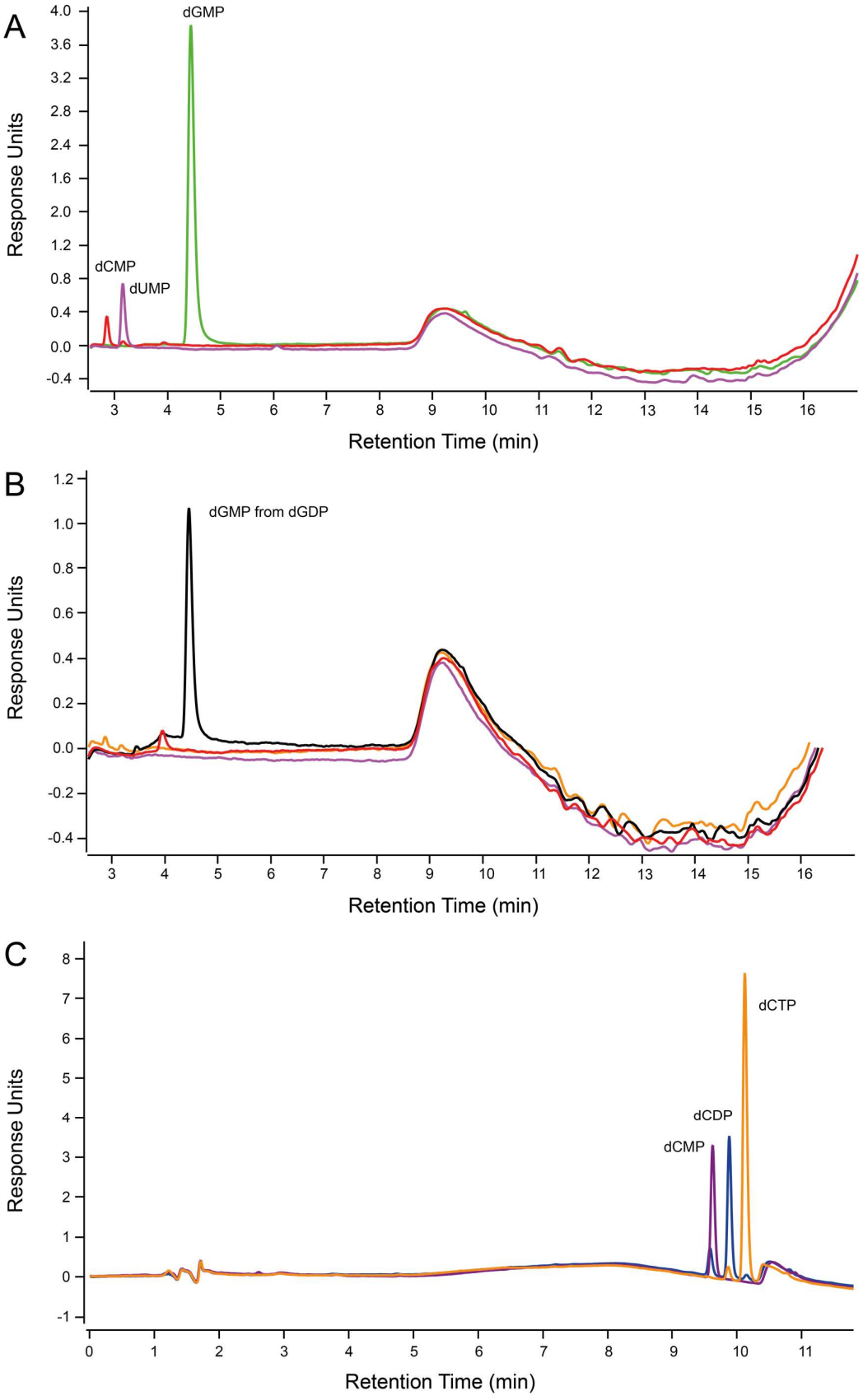

A
dGMP
dCMP
dUMP
Response Units
Retention Time (min)
B
dGMP from dGDP
Response Units
Retention Time (min)
C
dCTP
dCDP
dCMP
Response Units
Retention Time (min)

***Figure 4.8 Nucleotide separation is better on HILIC-Z column than PFP column**. Separation was judged by UV absorbance analysis. The PFP column only showed separation of monophosphate nucleotides, with no signal obtained of multiply phosphorylated nucleotides. The HILIC-Z column resolved the mono-, di- and tri-phosphate nucleotides. **A,** separation of 5 pmol of dCMP, dUMP and dGMP; **B,** separation of 5 pmol dCDP, dCTP, dGDP and dGTP (no detection); **C,** separation of 5 pmol of dCMP, dCDP and dCTP.*

To understand the reason of the absence of signal of multiply phosphorylated nucleotides with the PFP column, I was in correspondence with Agilent Technology. My hypothesis was that the absence of multiply phosphorylated nucleotides was due to the acid solvent system used (pH of 0.1 % formic acid is 2.6). They were of the opinion instead that the multiply phosphorylated nucleotides were too polar for column retention and were eluting in the injection peak.

Considering this and that the HILIC-Z column gave resolution of nucleotide mono-, di- and tri-phosphates, I chose the HILIC-Z column for nucleotide chromatography.

4.3 Optimising MS/MS for nucleotide analysis

4.3.1 Choice of polarity of instrument

The choice of instrument polarity, whether to analyse ions in positive or negative mode, was not immediately clear. Nucleotides have basic amine groups and endocyclic nitrogen atoms that can be protonated, but they have acidic phosphate groups that can be deprotonated (**Figure 4.9**). Due to the basic nitrogen atoms in the nucleobase and the acidic phosphate groups, nucleotides are not basic or acidic (the pKa of ATP is 6.5). There is therefore no clear preference for either positive or negative ionisation. It is true that the electrospray ionisation creates gaseous ions that are not subject to the influence of solvation as is seen in solution,

so it may be that the behaviours in solution of a molecule's chemical groups do not influence the molecule's ionisation in the gaseous phase.

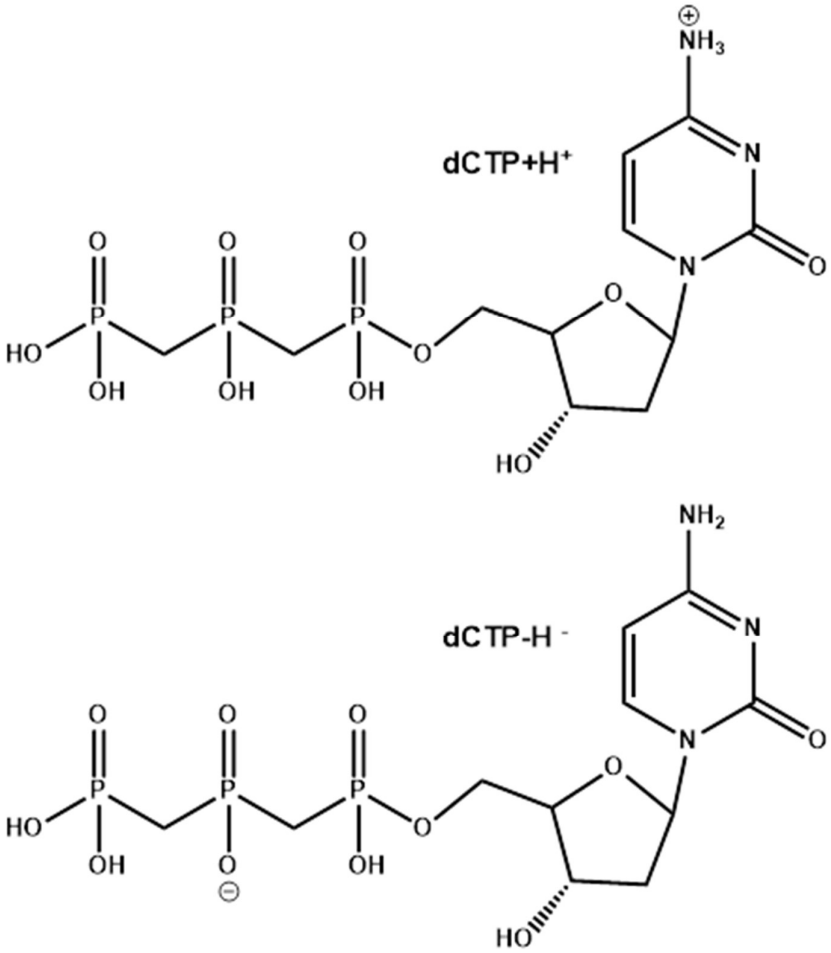


Figure 4.9 Positive and negative ionisation of dCTP. *The dCTP H^+ ion has a protonated nucleobase, whereas dCTP-H^- has a deprotonated phosphate group.*

Given reports in the literature of both positive and negative ionisation mass spectrometry, I investigated both to determine the strongest possible nucleotide signal.[196,197,222,230] **Figure 4.10** shows transition of ATP and its isomer dGTP that have transitions are representative of the other nucleotides.

For negative ionisation mode, I used the transitions reported by Strzelecka et al in their paper on fragmentation patterns of nucleotides in negative ionisation mode (**Figure 4.10A**).[230] For

positive ionisation mode, I used transitions published by three groups that were in agreement (**Figure 4.10B**).[196,197,231] The solvent system I used to test polarity signal: solvent A composed of 10 mM ammonium acetate and medronic acid, acidified to pH 7 with acetic acid and solvent B was pure acetonitrile (ACN).

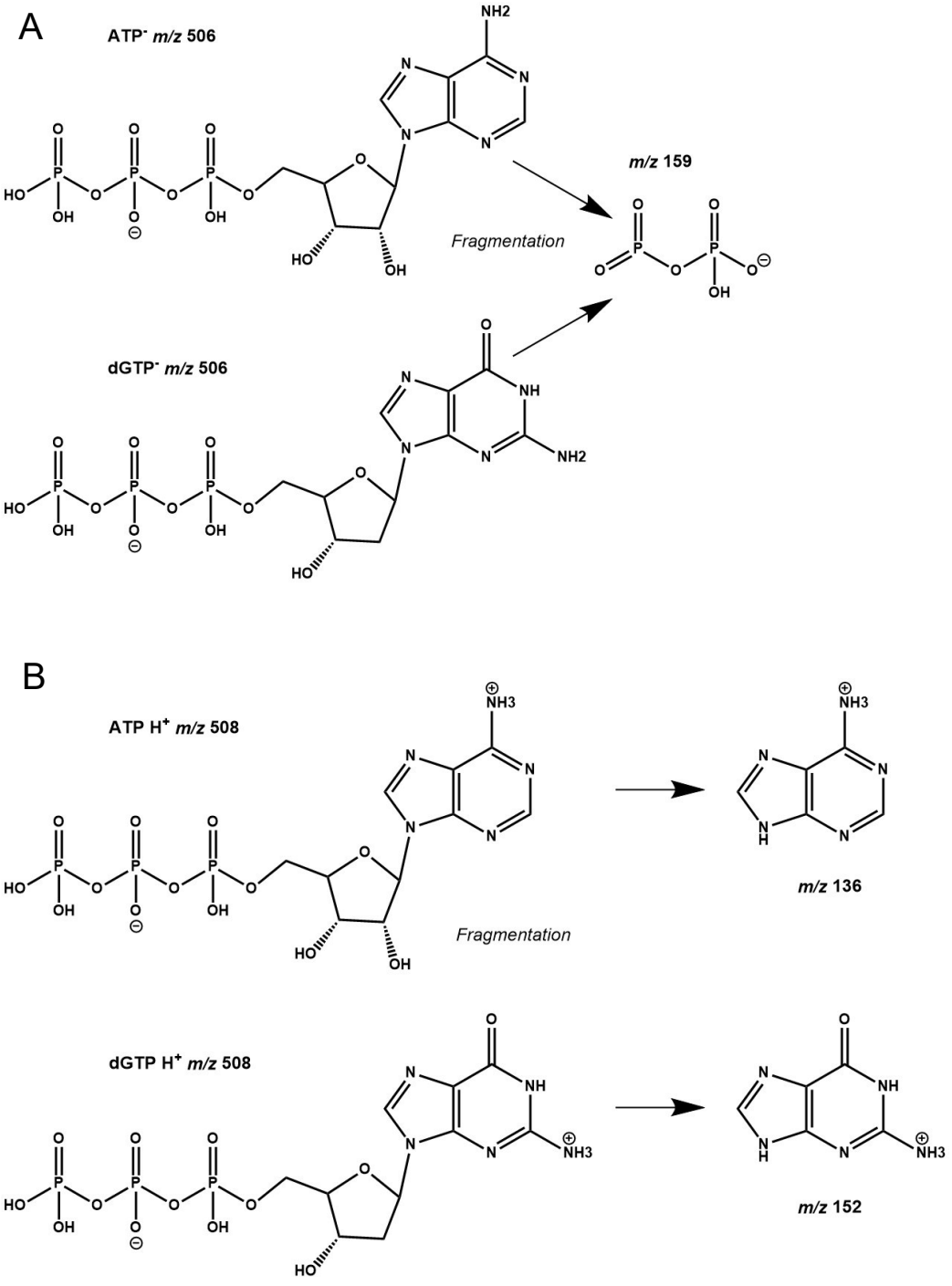


Figure 4.10 The ion transitions of positively and negatively charged nucleotides are different. A, *the principal ion transition of ATP and dGTP in negative mode (m/z 506→159);* ***B,*** *the principal ion transitions of ATP and dGTP in positive mode (m/z 508→136 and 508→152).*

I compared the signal derived from negative and positive ionisation modes (**Figure 4.11**). ATP and dGTP have the same dMRM transitions in negative ionisation modes (**Figure 4.11A**). As ATP is 500 times more abundant in biological samples than dGTP, it would be challenging to identify the signal from dGTP.[2] By contrast, the ions given in positive mode are either base-derived or sugar-derived (**Figure 4.12A**) and so deoxyribonucleotides can be differentiated from isomeric ribonucleotides. Additionally, the signals for ATP and dGTP were 20x and 40x stronger respectively than in negative mode.

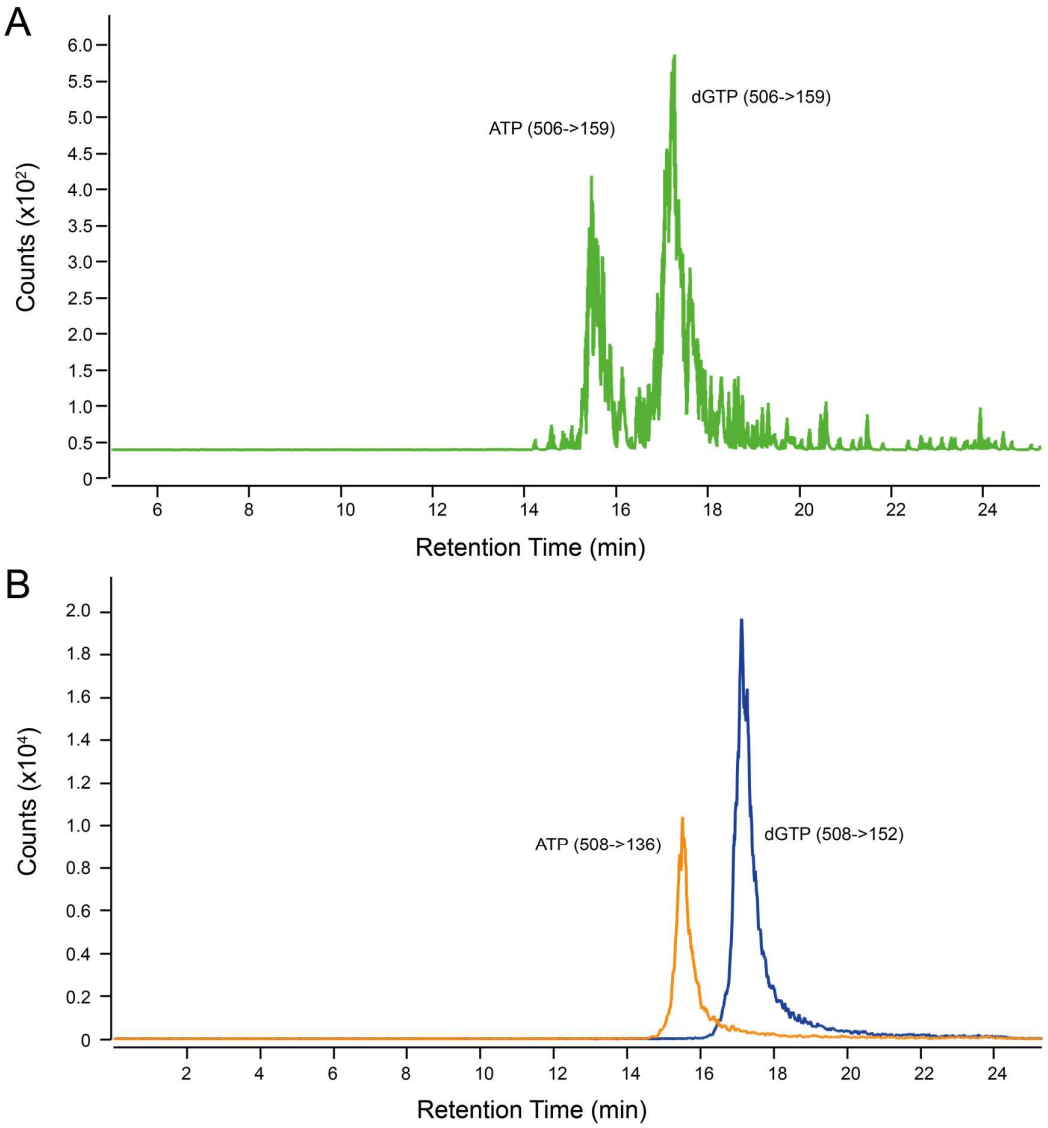


Figure 4.11 Positive mode gives stronger ATP and dGTP signal than negative mode. A, *extracted ion chromatogram (m/z 506→136) in negative mode, 1pmol injected on column;* ***B,*** *overlaid extracted ion chromatograms of ATP and dGTP in positive mode (m/z 508→136 and 508→152), 1pmol injected on column. These are stronger than in negative mode and are orthogonal, with no signal interference.*

As observed for nucleoside analysis, nucleotide-salt adducts had different signal distributions depending on what the nucleobase was: for instance, dCMP had the strongest signal as when protonated, whereas dTTP's strongest adduct was Na^+ (**Figure 4.12**).

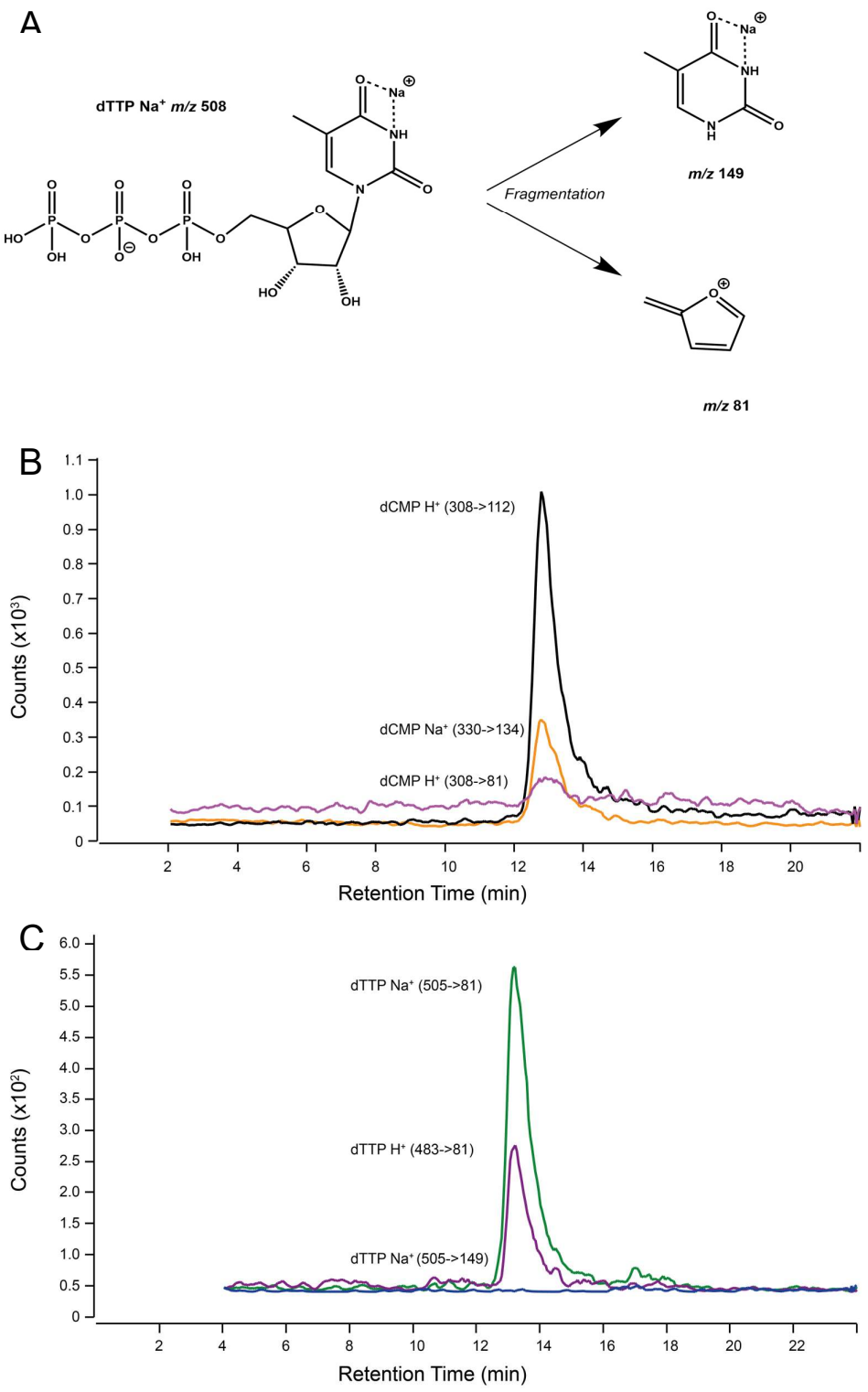


Figure 4.12 The main transitions in positive mode for dCMP and dTTP. A, *schematic of transitions for dTTP in positive mode, as the Na^+ adduct;* ***B,*** *the main transitions for 1 pmol dCMP H^+;* ***C,*** *the main transitions for 1 pmol dTTP Na^+.*

The dominant adduct for dTTP was later dTTP H^+, due to continually changing conditions at the source as a consequence of multiple sample runs. In my final experiments it reverted to being dTTP Na^+.

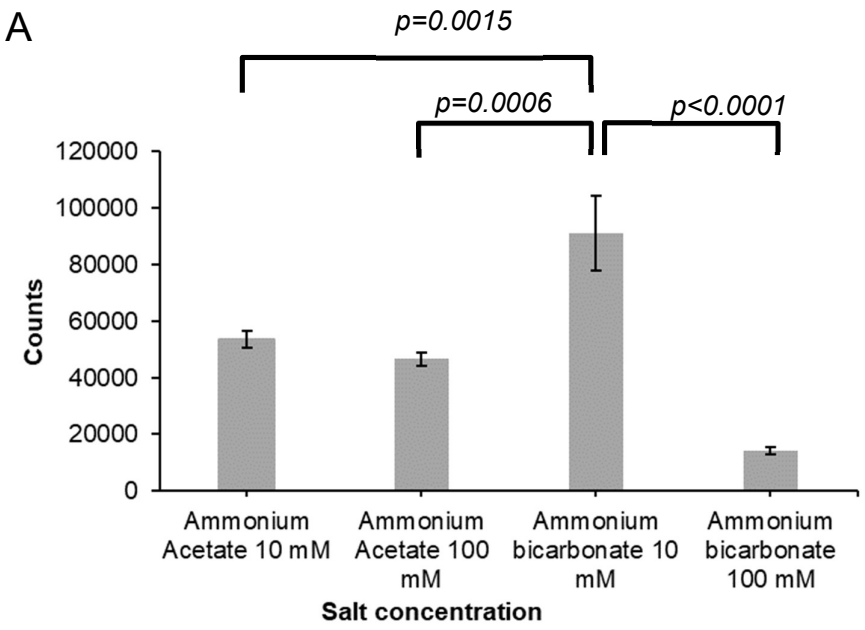
Given the increase in sensitivity and orthogonality of transitions, I chose to analyse the nucleotides in positive ionisation mode.

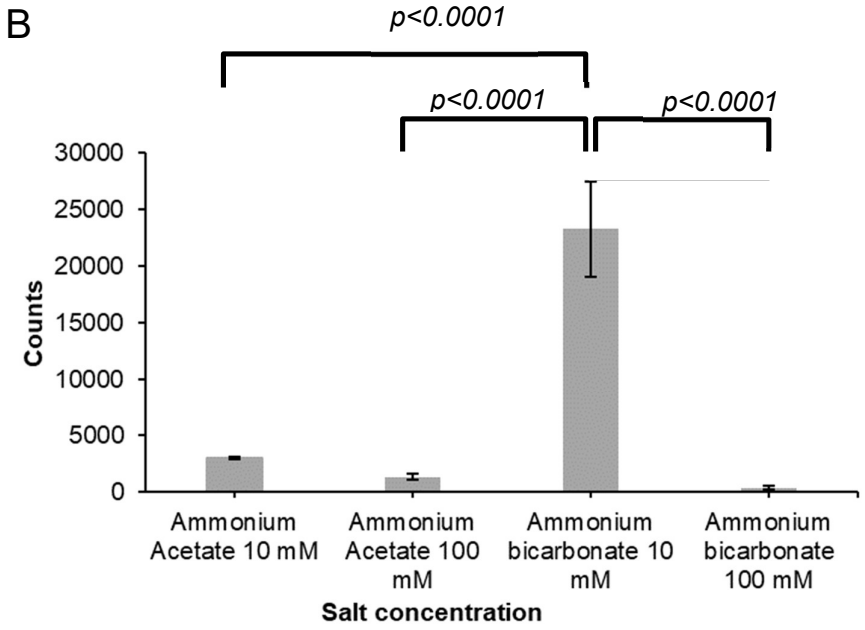
4.3.2 Solvent composition

The solvent composition impacts the desolvation at the source and influences the potential for analyte ionisation. I therefore tested 4 different solvent A salt compositions and a wide pH range to optimise for nucleotide ionisation, keeping solvent B at 100 % ACN.

4.3.3 Solvent composition: salt

The salt in a mobile phase is required to maintain a stable pH in the system for reliable retention times and analyte ionisation in the source. The solvent salt compositions I tested are ammonium bicarbonate (NH_4HCO_3, 10 mM and 100 mM) and ammonium acetate ($NH_4CH_2CO_2$, 10 mM and 100 mM). Ammonium bicarbonate is used by Matsuda et al., while ammonium acetate is used by Kong et al. 10 mM ammonium bicarbonate gave the strongest signal for all nucleotides tested (**Figure 4.13**). The reasons for bicarbonate giving the strongest and weakest signals depending on concentration are unclear. It is not due to an undetected increase in ammonium adducts, as ammonium acetate would also have that effect, when instead 100 mM ammonium acetate gives a similar signal to 10 mM ammonium acetate.


A
p=0.0015
p=0.0006
p<0.0001
120000
100000
80000
60000
40000
20000
0
Counts
Ammonium Acetate 10 mM
Ammonium Acetate 100 mM
Ammonium bicarbonate 10 mM
Ammonium bicarbonate 100 mM
Salt concentration


B
p<0.0001
p<0.0001
p<0.0001
30000
25000
20000
15000
10000
5000
0
Counts
Ammonium Acetate 10 mM
Ammonium Acetate 100 mM
Ammonium bicarbonate 10 mM
Ammonium bicarbonate 100 mM
Salt concentration

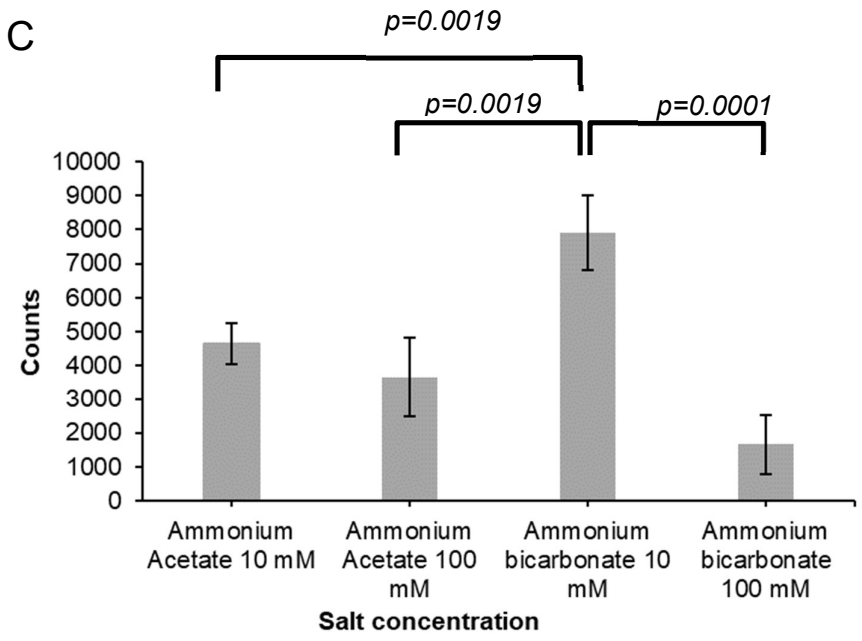


Figure 4.13 10 mM ammonium bicarbonate gives the strongest nucleotide signal. *1 pmol of nucleotides was injected onto the column (n=4).* ***A,*** *effect on dCMP H^+;* ***B,*** *effect on dUMP H^+;* ***C,*** *effect on dTTP H^+.*

4.3.3.1 Solvent composition: pH

pH affects the charge state of nucleotides on the column, influencing the retention time, as well as the electric field strength around the zwitterionic groups on the column. In addition, it influences the ability of nucleotides to ionise at the desolvation step: this could affect the strength of signal and peak width and shape, all three of which can impact sensitivity.

I tested the effect of pH in the range of pH 5 to pH 10 (**Figure 4.14**). pH 8 gave the strongest signals in dCMP and dUMP, but the weakest for dTTP (Na^+). The extreme pHs were unsuitable: pH 10 gave uniformly weak signals while pH 5 gave wide and irregular peak shapes. Although pH 5 gave strong signals for dCMP and dUMP, the width of the peaks would risk worse chromatography and a lower signal-to-noise, especially relevant when analysing biological samples. pH 6 gave consistent and well-formed peaks, with the strongest signal for dTTP. This

was also seen for protonated dTTP (dTTP H^+, data not shown). I therefore chose pH 6 as the pH for solvent A.

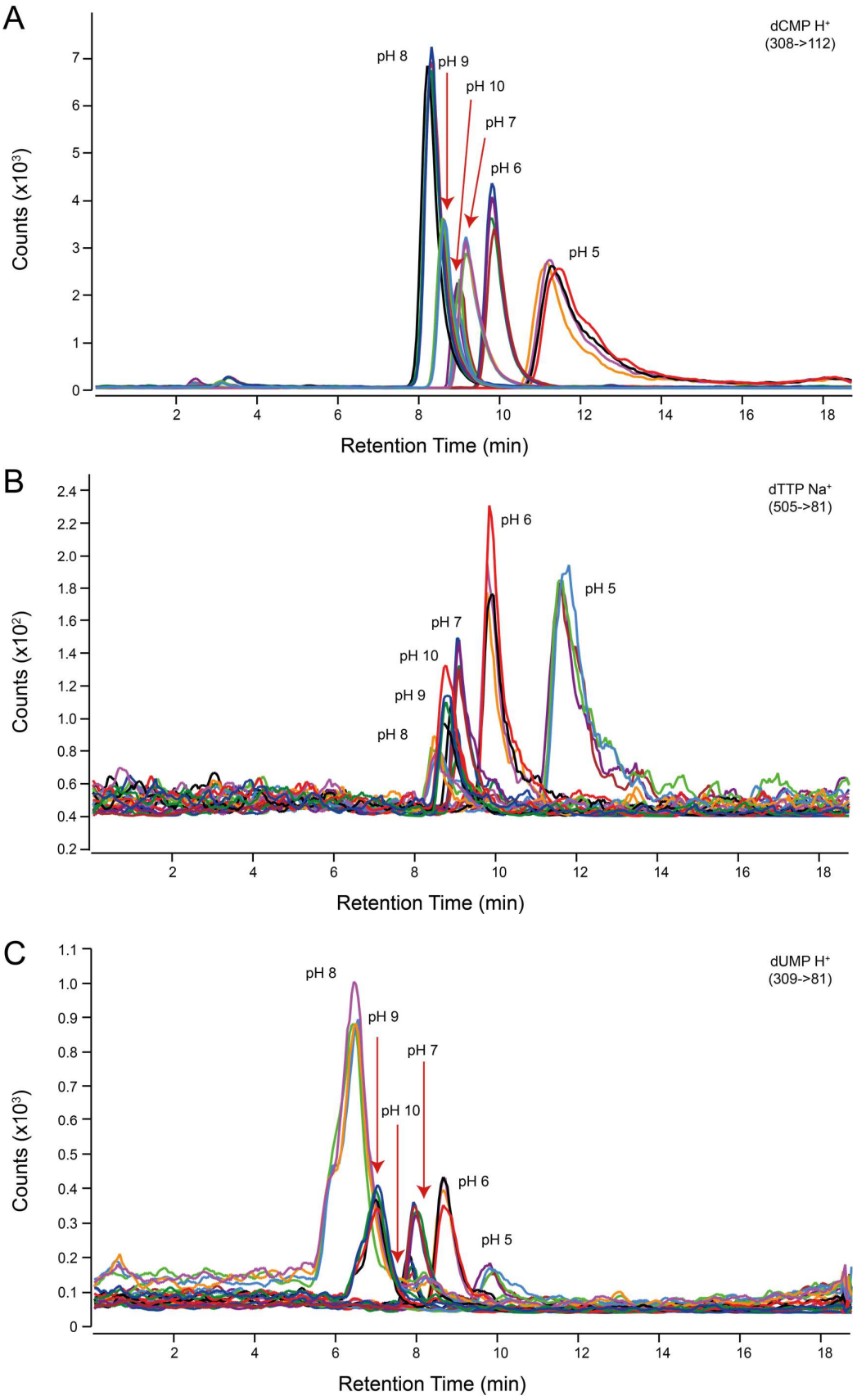
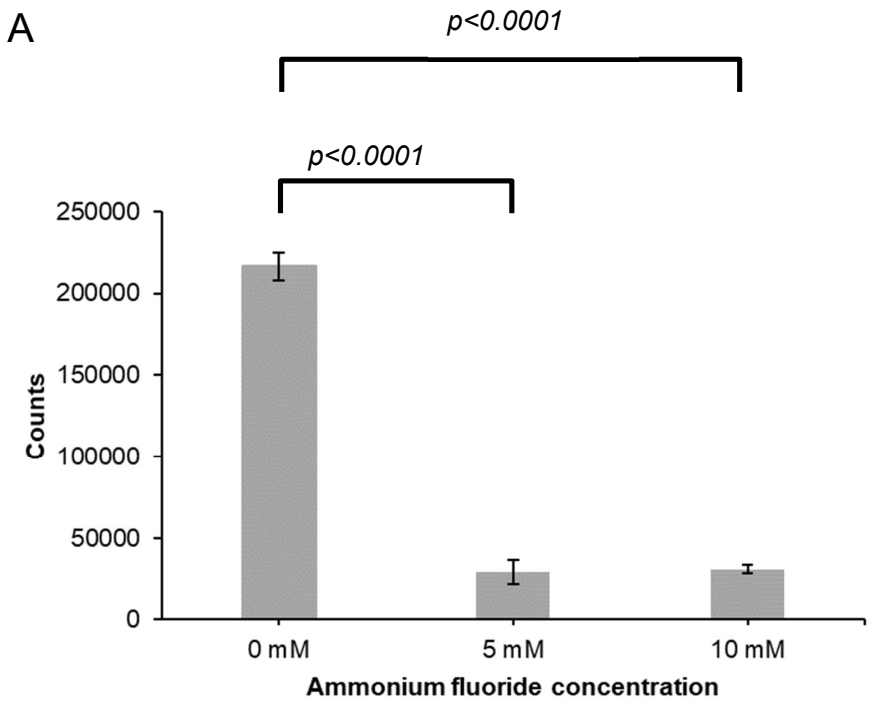

A
dCMP H+
(308->112)
pH 8
pH 9
pH 10
pH 7
pH 6
pH 5
Counts (x10³)
Retention Time (min)
B
dTTP Na+
(505->81)
pH 6
pH 5
pH 7
pH 10
pH 9
pH 8
Counts (x10²)
Retention Time (min)
C
dUMP H+
(309->81)
pH 8
pH 9
pH 7
pH 10
pH 6
pH 5
Counts (x10³)
Retention Time (min)

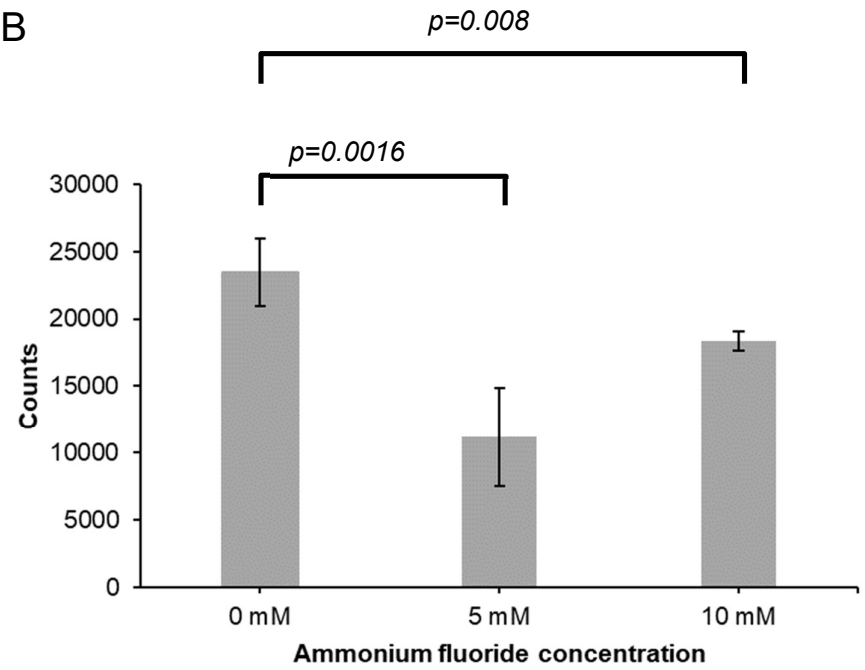
Figure 4.14 Solvent gave strongest signal overall at pH 6. *1 pmol of nucleotides was injected onto the column (n=4).* ***A,*** *effect on dCMP;* ***B,*** *effect on dTTP;* ***C,*** *effect on dUMP.*

4.3.3.2 Solvent composition: ammonium fluoride

The use of ammonium fluoride as a mobile phase additive has been empirically reported to improve the sensitivity of some analytes for LC-ESI-MS/MS.[232] I tested the concentrations of 0 mM, 5 mM and 10 mM to see if it improved nucleotide sensitivity **(Figure 4.15**).



B


p=0.008
p=0.0016
30000
25000
20000
15000
10000
5000
0
Counts
0 mM
5 mM
10 mM
Ammonium fluoride concentration

C

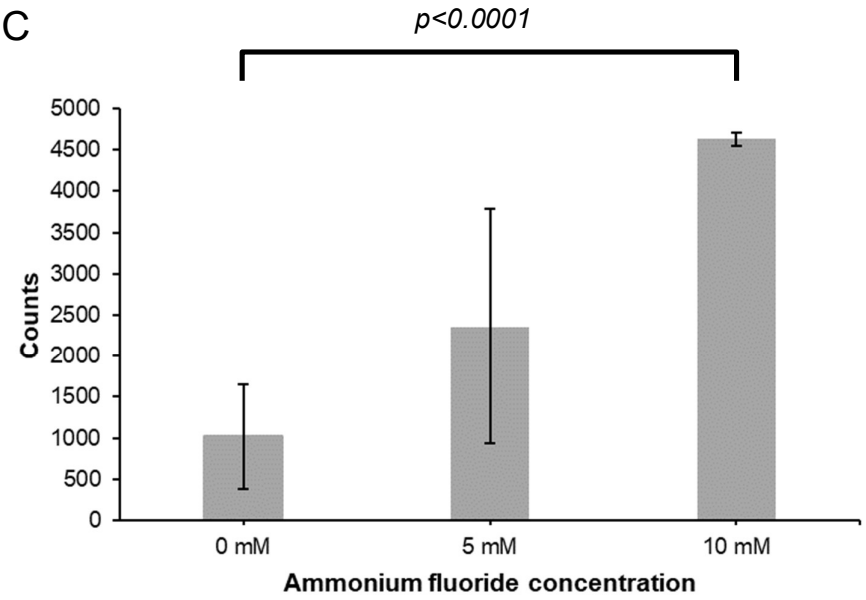

p<0.0001
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
Counts
0 mM
5 mM
10 mM
Ammonium fluoride concentration

Figure 4.15 Ammonium fluoride reduces nucleotide monophosphate signal and increases nucleotide triphosphate signal. *1 pmol of nucleotides was injected onto the column (n=4).* ***A,*** *effect on dCMP;* ***B,*** *effect on dUMP;* ***C,*** *effect on dTTP.*

It seems that the effect of ammonium fluoride is related to the number of phosphate groups of nucleotides; the signal for nucleotide monophosphates was strongest with no ammonium fluoride, illustrated by dCMP and dUMP (**Figure 4.15A and B**). However, the signal for nucleotide triphosphates was strongest with 10 mM ammonium fluoride (**Figure 4.15C**), with no change for nucleotide diphosphates (data not shown).

The reduction in signal for nucleotide monophosphates upon addition of ammonium fluoride is much greater than the increase for nucleotide triphosphates. In addition, I am primarily interested in the monophosphates dCMP and dUMP for investigating the role of dCTD and CDA, so the use of ammonium fluoride was discontinued.

The final solvent composition is as follows. Solvent A is 10 mM ammonium bicarbonate in H_2O, pH to 6 with glacial acetic acid, 5 µM medronic acid; solvent B is 100 % ACN.

4.3.4 HPLC gradient optimisation for new solvent

The solvent composition also influences retention times of nucleotides on the column and their separation, which is important as some nucleotides share MS transitions and so will have problematic quantification if they coelute. I therefore optimised the column gradient to increase differences in nucleotide column retention times.

I tested different solvent gradients to maximise nucleotide separation, from the steepest (A, magenta) to less steep (black) and to the shallowest (red) (**Figure 4.16**). The gradients were all shallow considering the similar properties of the nucleotide analytes.

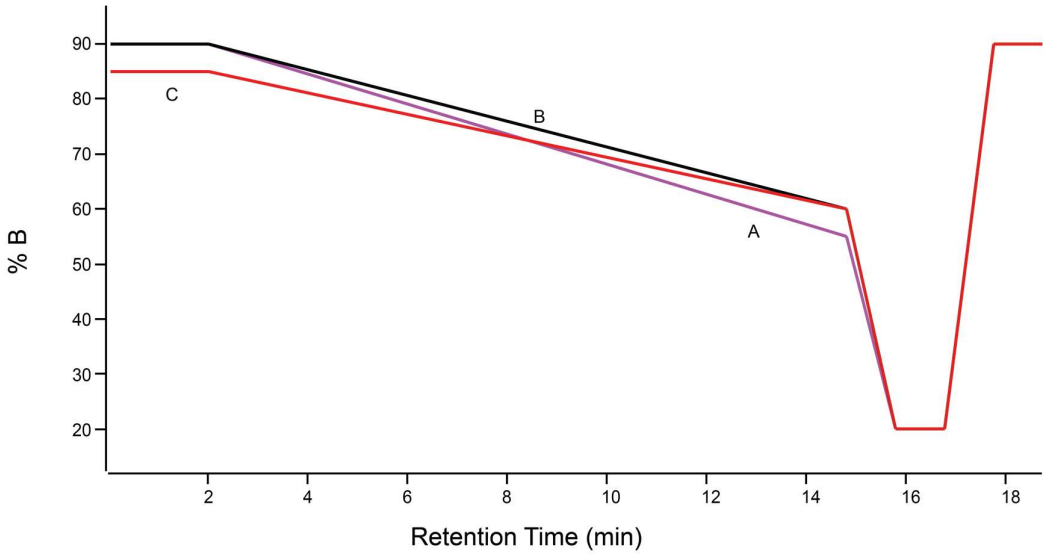


Figure 4.16 Percentage solvent B over time of gradients A, B and C.

I compared the resolution for dCMP, dCDP and dCTP (**Figure 4.17**). It is important to avoid full co-elution in the event of dephosphorylation of multiply phosphorylated nucleotides like dCDP or dCTP in the MS source, creating artefacts that would interfere with quantification of less phosphorylated nucleotides. The gradient is by necessity shallow to minimise nucleotide co-elution and so the nucleotides have wide peak widths. A strategy to narrow the peak widths could have been to increase flow rate, but the column pressure would have increased beyond the allowable upper limit. Despite all the tested gradients giving low resolution peaks, gradient C separated dCMP from dCDP and dCTP most. The peaks are, however, wider and the dCMP signal is lower than seen for the other gradients. This may be because of less favourable solvent composition upon desolvation. I decided that the increase in separation

was worth the loss in signal and used gradient C going forwards.

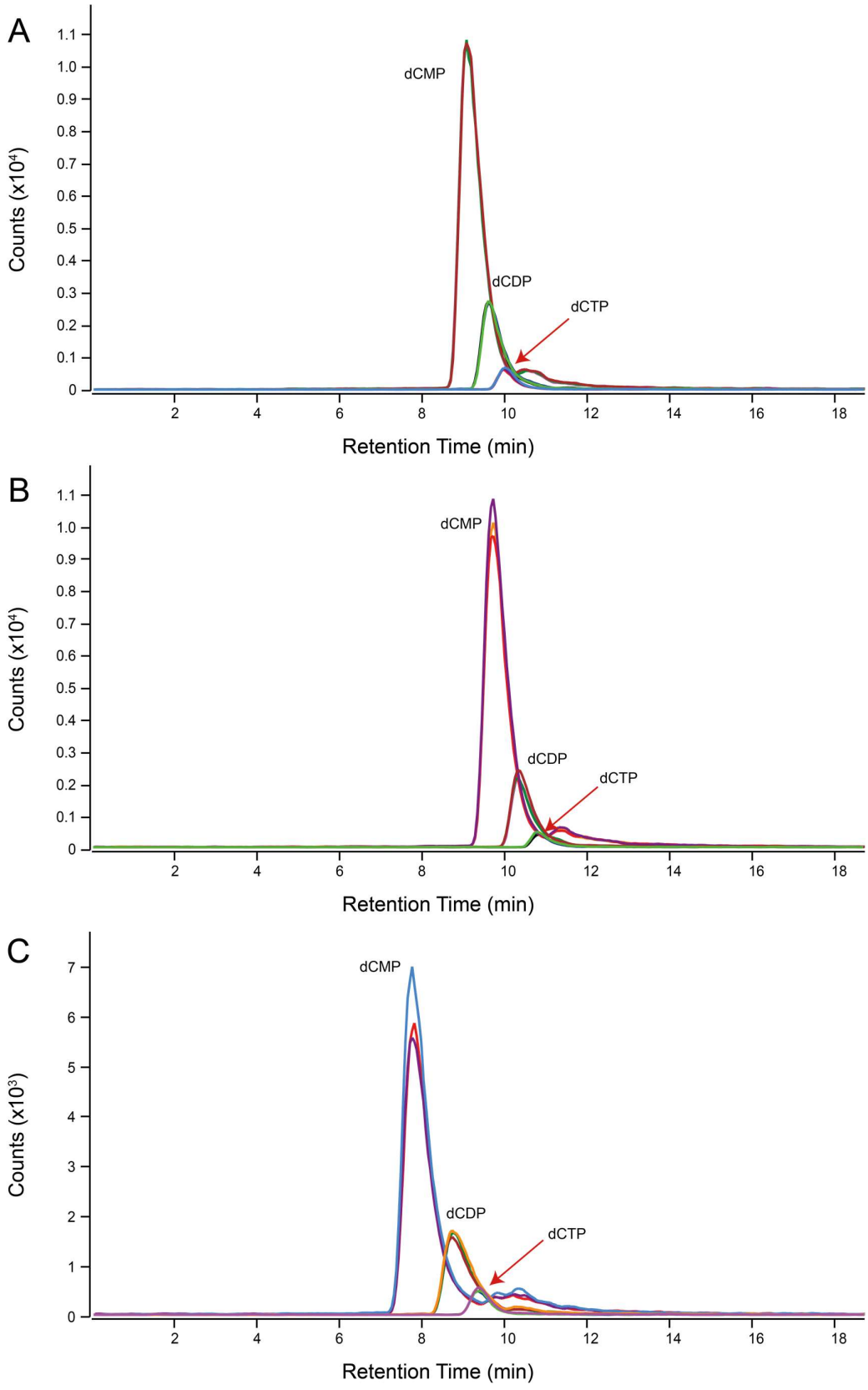


Figure 4.17 Gradient C gave best resolution of dCMP, dCDP and dCTP. *1 pmol of nucleotides was injected onto the column (n=3)* ***A,*** *nucleotide resolution of dCMP, dCDP and dCTP with gradient A;* ***B,*** *nucleotide resolution of dCMP, dCDP and dCTP with gradient B;* ***C,*** *nucleotide resolution of dCMP, dCDP and dCTP with gradient C.*

The gradient that gives the best nucleotide resolution is gradient C (**Table 4.3**).

Time (minutes)	% A	% B
0	15	85
2	15	85
15	85	60
17	60	20
18	10	90
19	10	90
19-24 (5 minute post run wash)	15	85

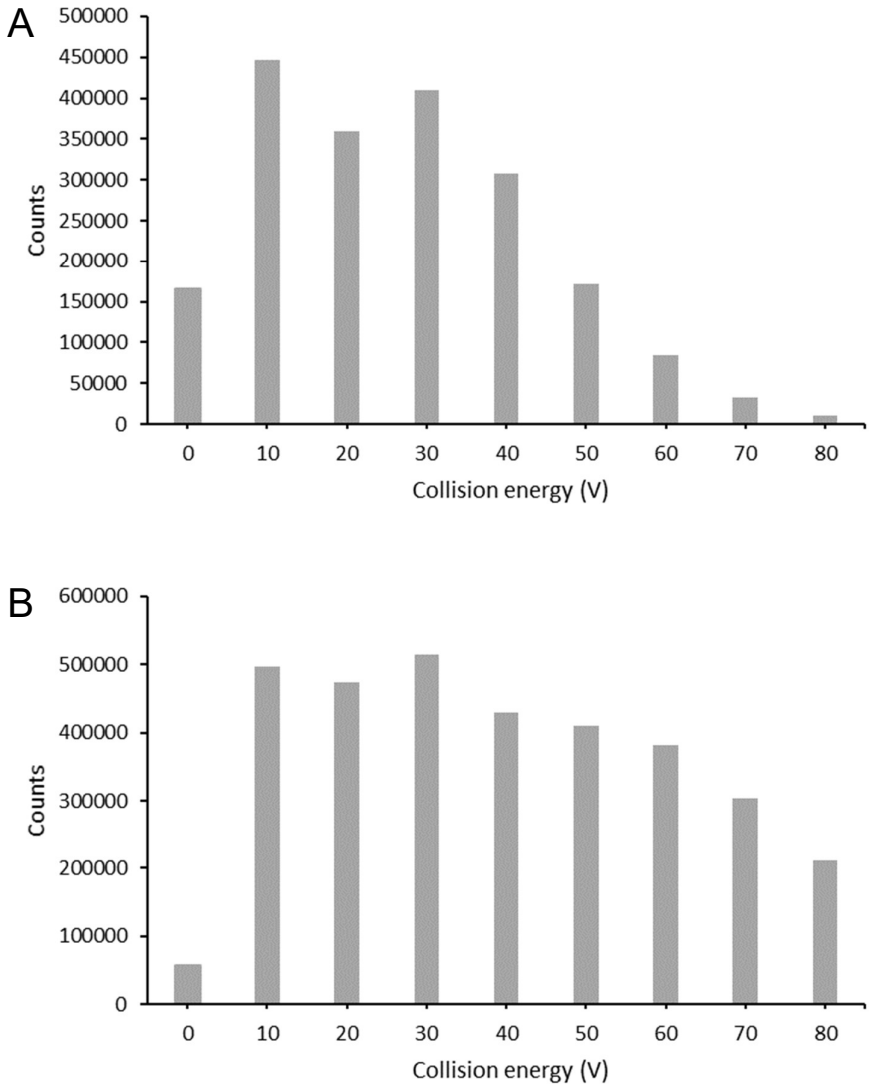
Table 4.3 Percentage solvent composition over time of gradient C.

4.3.5 Mass spectrometric conditions

4.3.5.1 Collision energy

In parallel, I optimised the signal for the ion transitions. Using the transitions from the published papers, I optimised the transition collision energies (the energy with which the

instrument fragments the parent ions into target ions) for the nucleotides, to look for the most collision energy that gave the most signal of target ion. I tested energies from 0 V to 80 V (**Figure 4.18**).



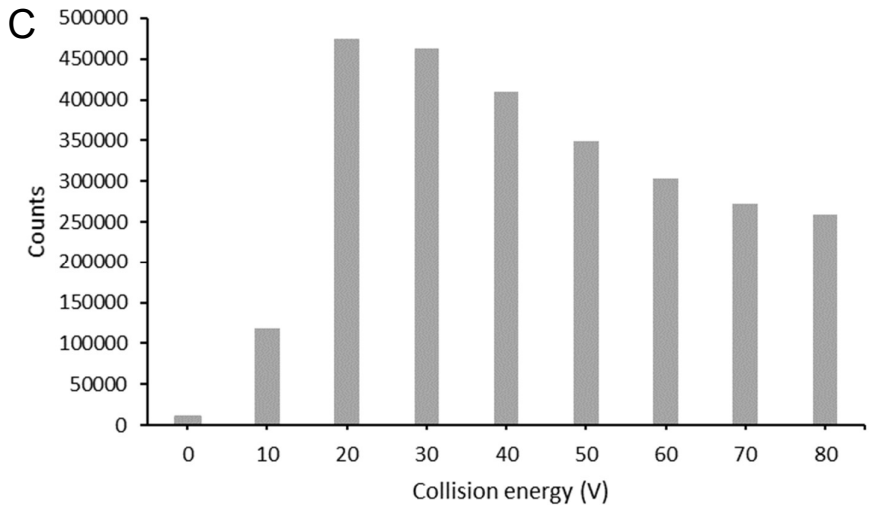


Figure 4.18 Best collision energy voltage for nucleotide transitions varies per nucleotide. *1 pmol of nucleotides was injected onto the column per run.* ***A,*** *dCMP H^+ (10 V);* ***B,*** *dCTP H^+ (30 V);* ***C,*** *dTTP Na^+ (20 V).*

The nucleotides showed a distribution of preferences for collision energies with dCMP, dCTP and dTTP shown as illustrative examples. I updated the dMRM method to use the best collision energy for each nucleotide transition.

4.3.5.2 EMV

To increase the signal-to-noise ratio further, I tested changing the electron multiplier voltage (EMV). This parameter can be increased to raise total signal by multiplying the signal obtained from the detector. Increasing the EMV too much may also increase noise and reduce multiplier lifetime, so I tested the effect of increasing the EMV by up to 800 V, with dCMP as a representative nucleotide shown in **Figure 4.19**.

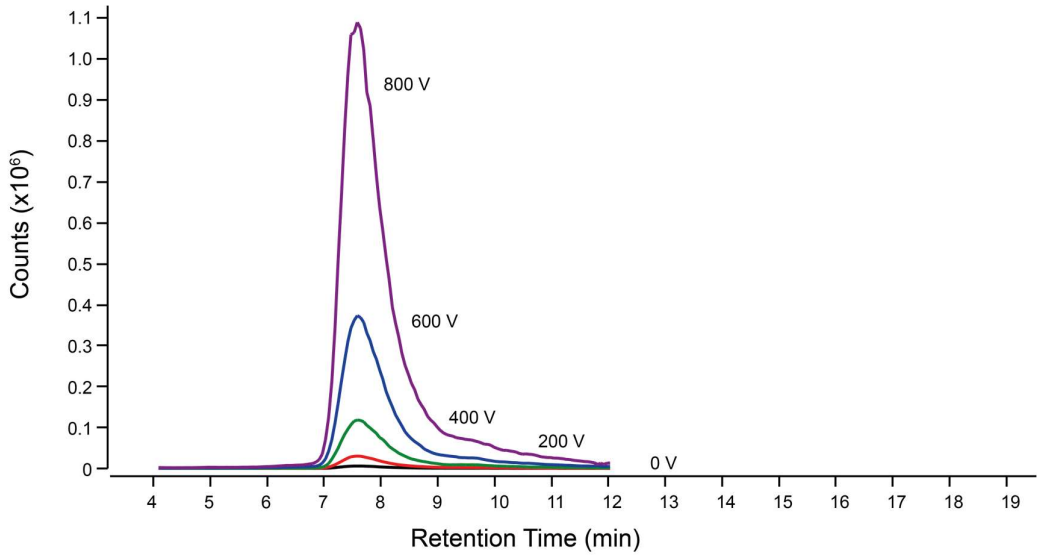


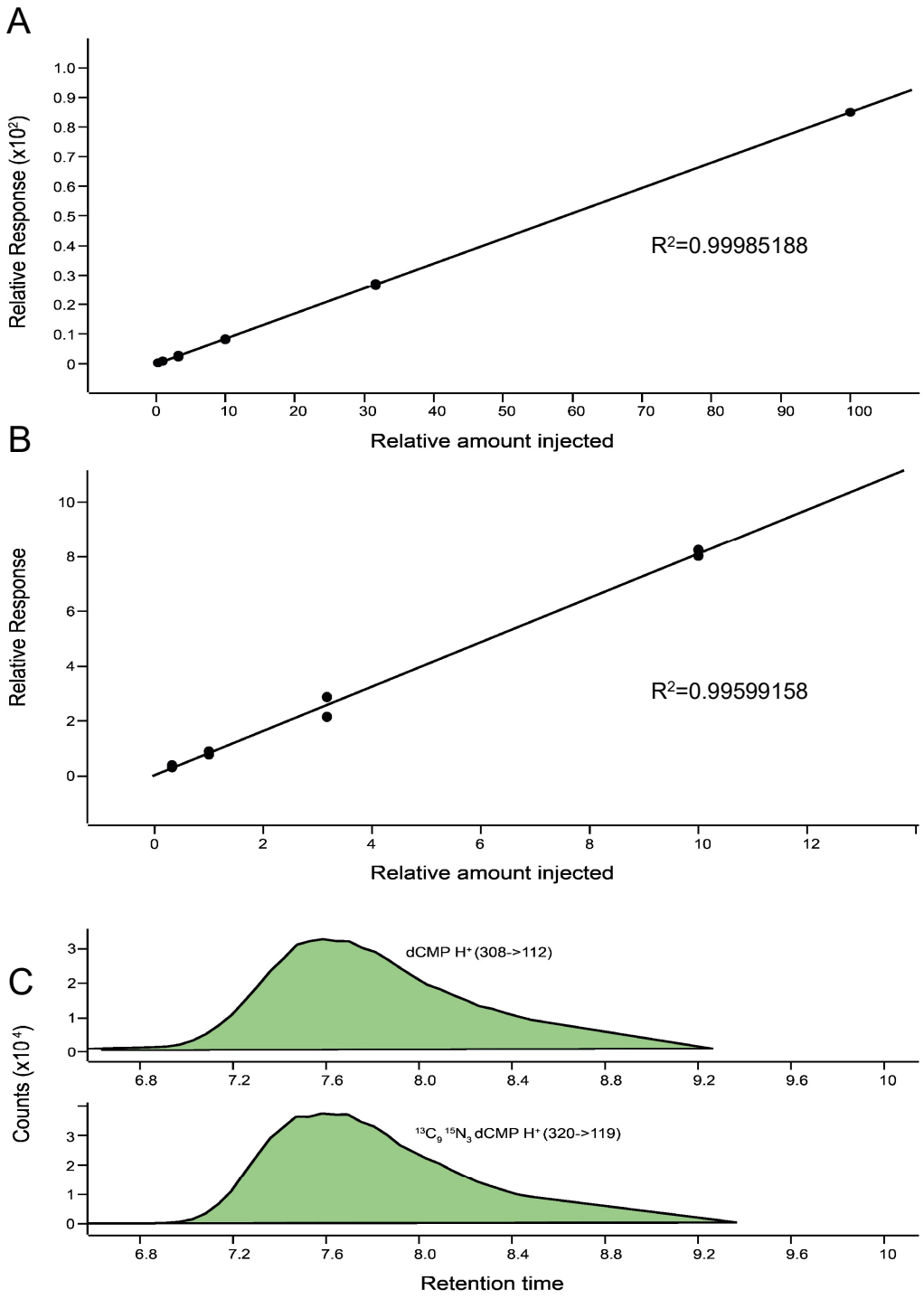
Figure 4.19 Increasing delta EMV increases signal of dCMP. *1 pmol of nucleotides was injected onto the column per run.*

The signal improved for every incremental increase in EMV. However, Agilent recommend increasing the EMV by no more than 400 V to avoid increasing noise more than signal was increasing; I chose to be cautious and limit the increase to 400 V in case the noise would increase in the matrix-affected conditions of samples. Further investigation of delta EMV parameter, including testing EMV in biological samples instead of standards, may yield a stronger nucleotide signal-to-noise ratio.

4.3.6 Determination of sensitivity

Now that the system was ready for sample analysis, I tested its sensitivity and ran calibration curves from 31.6 fmol on column to 10 pmol on column, in increments rising by a factor of $\sqrt{10}$ (**Figure 4.20**). I spiked in the standards with 100 fmol of heavy-labelled internal standards per injection for all nucleotides, apart from the deoxyribonucleotide diphosphates, which was not commercially available. The limit of detection (LOD)—the amount analysed lower than which there is difference to background—can be defined as 3x the baseline noise, and the

lower limit of quantification (LLOQ)—the amount analysed lower than which there is no



proportional change in intensity—is at 10x the baseline noise.[233,234]

Figure 4.20 Calibration curves for dCMP H⁺ with internal standard. *Two injections per amount on column, normalised to proportion compared to internal standard injected (100 fmol).* ***A,*** *dCMP H⁺ whole curve;* ***B****, dCMP H⁺ curve at low amounts;* ***C****, comparison of dCMP H⁺ signal (100 fmol) with* $^{13}C_9$ $^{15}N_3$ *dCMP H⁺.*

The system gives sensitive calibrations for all nucleotides, with most nucleotides having a lower limit of detection (LLOQ) smaller than 31.6 fmol. This is comparable to the results obtained by Matsuda et al. and is extended to all canonical nucleotides, not just the nucleotide triphosphates.[197] As for many nucleotides, the lowest amount on column injected gave a signal-to-noise ratio greater than 10. Further standard dilutions will be able to find the true LLOQ and how sensitive the instrument can be.

The nucleotide response and retention times are shown in **Figure 4.21**. Despite a shallow chromatographic gradient, the nucleotides elute within 6 minutes of each other and selected nucleotides are annotated for illustration below.

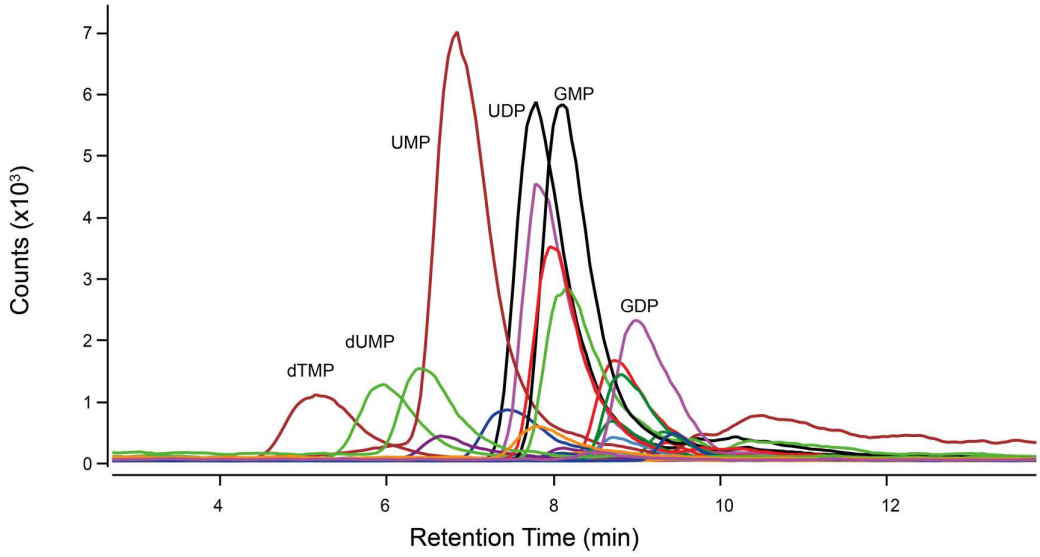


Figure 4.21 Optimised signal and chromatographic separation of all nucleotides. *1 pmol of nucleotides was injected onto the column per run.*

The complete list of nucleotide transitions and LLOQs is given in Appendix A, **Table 3**.

4.4 Validation

Having optimised the LC-MS/MS system and sample preparation for nucleotide analysis, I set out to validate its capability with biologically relevant sample. I extracted metabolites from H1299 WT cells—chosen as they have low CDA expression—and H1299 cells with stable CDA overexpression, which were generated previously in the lab.[62] Low expression of CDA has been reported to lead to a higher dCTP concentration in cells and so an artificial overexpression of CDA may have the opposite effect, considering that CDA overexpression is lethal when upon administration of epigenetically modified cytosines and thus has a noticeable deaminating effect in the nucleotide pool.[62,82]

I would judge the system to be validated if it could detect a difference in signal from the deoxycytidine nucleotides in the sample nucleotide pools due to activity of CDA (**Figure 4.22**).

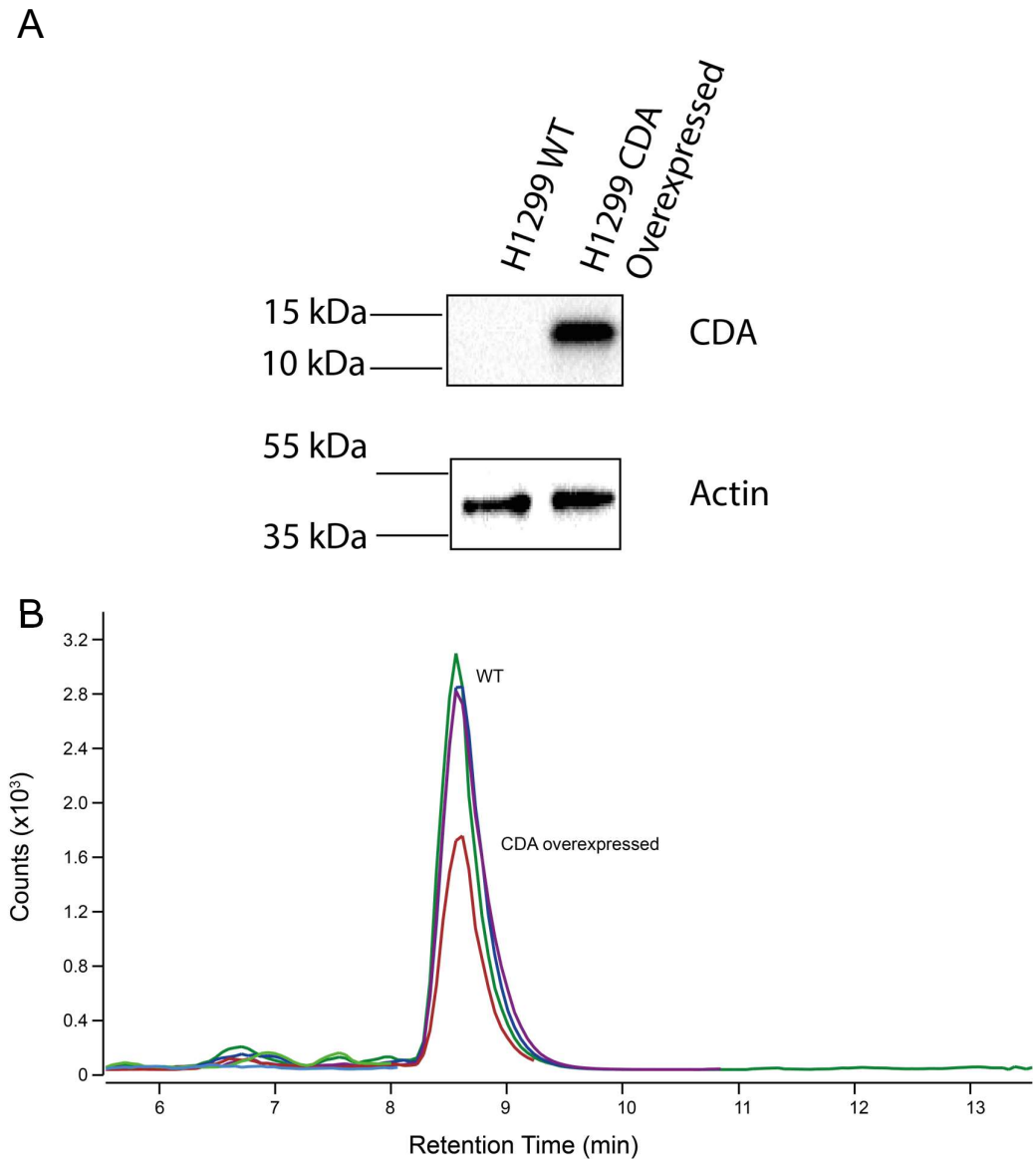


Figure 4.22 CDA overexpression may reduce intracellular dCTP. *Comparison of dCTP H^+ signal from WT H1299 vs H1299 overexpressed CDA (not enough points for statistical analysis).* ***A,*** *dCTP H^+ in H1299 WT (n=3) and in overexpressed CDA (**n=1**);* ***B,*** *WB showing CDA and β-actin in WT H1299 and H1299 CDA overexpressed.*

The level of dCTP in the nucleotide pool appears to be decreased in H1299 cells overexpressing CDA compared to WT. However, only one replicate for H1299 with overexpressed CDA was run successfully and so more replicates would need to be analysed to validate this result. The limitation of this experiment is that there are other enzymes

involved in nucleotide metabolism, notably the NDPK and NME families, as well as RNR. A better experiment for validation would be to measure isotopically labelled cytidine tracing in the cytidine and uridine nucleotide pools across the phosphorylation states.

4.5 Discussion

I have developed a sensitive LC-MS/MS method for analysis of the nucleotide pool, requiring few processing steps and with most nucleotides detectable at concentrations comparable or lower than literature without using ion-pairing agents.[196,197] Its discrimination of isotopically labelled internal standards demonstrates that it can trace labelling in the nucleotide pool and can be used in determining the roles of CDA and DCTD in the pyrimidine salvage pathway.

4.5.1 Nucleotide pool extraction

Metabolite extraction varies according to the metabolites under study, by extraction solvent and also by post-extraction sample treatment.[165,214,215,218,219,222] I tested three extraction solvents to maximise signal of the polar nucleotides, with 50 % ACN giving the strongest signals for most nucleotides of interest (**Figure 4.1**). I investigated whether heating and sonication of the samples increased nucleotide signal. I found an increase in nucleotide signal from heating, with no impact from sonication (**Figure 4.3**). Finally, I tested whether sample volume injection would affect the chromatography on the HILIC-Z column, as well as impact nucleotide signal (**Figure 4.5**). The nucleotide peak shapes were not affected by changing the volume, while the signal did not increase beyond 1 µL, indicating ion saturation beyond that volume due to other metabolites or salt contamination.

4.5.2 Choice of column chromatography

Nucleotides have similar chemical properties, due to their polar phosphate groups and aromatic nucleobase, causing challenges for their separation. I tested a reversed-phase column functionalised with pentafluorophenyl (PFP) groups and a HILIC column functionalised with zwitterionic molecules (HILIC-Z) (**Figure 4.8**). The PFP column did not show UV signal from multiply phosphorylated nucleotides, possibly due to no retention on column or the samples degrading in the solvent conditions. By contrast, the HILIC-Z column had retention of nucleotide mono-, di- and triphosphates.

4.5.3 MS optimisation

Polarity

As nucleotides have basic atoms like nitrogen that favour forming positive ion adducts, and acidic phosphate groups that favour negative ions, it was not clear whether positive or negative instrument polarity would give the strongest signal—both positive and negative ionisation mode have been reported in literature.[196,197,230] I compared nucleotide dMRM transition signals of positive and negative mode (**Figure 4.10** and **Figure 4.11**). Positive mode gave the strongest signal for nucleotides and the dMRM transitions generated had greater differences, aiding sample quantification.

Solvent composition

The solvent influences an analyte's environment in the MS source, impacting drying and ionisation.

I tested the main components of solvent A in previous work on nucleotide analysis in positive ion mode, ammonium bicarbonate[197] and ammonium acetate[196], at concentrations of 10 mM

and 100 mM (**Figure 4.13**). 10 mM ammonium bicarbonate gave the strongest signal. The nature of the anion has a large impact on electrospray analyte ionisation;[235] it is possible that another concentration of ammonium bicarbonate would give stronger signals still, or by using another salt such as ammonium formate.

The impact of the pH of solvent A was investigated, from pH 5 to pH 9 in increments of 1 (**Figure 4.14**). pH 6 gave consistent peaks for all nucleotides, and the strongest signal for dTTP.

Ammonium fluoride was tested as an additive, as it had been reported to increase MS signal of the steroid spironolactone.[232] I compared nucleotide signal using 0 mM, 5 mM and 10 mM ammonium fluoride and found that nucleotide monophosphates had signal suppression with 5 mM and 10 mM ammonium fluoride, whereas nucleotide triphosphate signals were strongest with 10 mM ammonium fluoride (**Figure 4.15**). My focus was on the lower phosphorylation states nucleotides as dCMP is a substrate of DCTD, so I did not use ammonium fluoride.

MS instrument parameters

Changing the QQQ instrument parameters can influence the analyte signal. I tested varying the collision energy (CE) in stepwise manner, obtaining the optimised collision energy signals for each nucleotide (**Figure 4.18**). Additionally, I investigated the effect of increasing the electron multiplier voltage (EMV) to increase the nucleotide signal post-detection (**Figure 4.19**), deciding to use +400 V as a compromise between signal and the chance of noise multiplication.

The sensitivity of the LC-MS/MS system was tested with standards, with most nucleotides having an LLOQ <31.6 fmol (**Figure 4.20**).

There is scope for further gains to be made in the MS optimisation. If the EMV can be increased beyond +400 V without increasing noise in biological samples, signal-to-noise ratio

could be increased severalfold. Absent in this chapter has been discussion on source optimisation, which needed more time for completion. It is simple to do and will be the first thing to be done to progress the setup of this method.

4.5.4 Validation

To verify the capability of the instrument for analysing biological samples, I extracted nucleotides from H1299 cells that have low CDA, as well as H1299 with CDA overexpression (**Figure 4.22**). I would have considered the system validated if the signal from dCTP decreased in the overexpressed CDA cell line, but two biological replicates of the overexpressed CDA nucleotide pool sample did not run correctly. Thus, this experiment must be repeated, as well as through measuring differences in stable isotope-labelling of the free nucleotide pool upon cell exposure of labelled dC.

5. Cell-based tools for investigating nucleotide metabolism

My project objectives were investigate the relative importance of CDA and DCTD in the pyrimidine salvage pathways, and whether they act on nucleotides derived from the *de novo* synthesis pathway.

I carried this out by combining the methods for nucleic acid and nucleotide pool analysis described in the previous chapters with stable isotope-labelled precursor tracing and manipulation of expression of the deaminase proteins.

The protein manipulation method I chose was transient knockdown (KD) by siRNA (short interfering RNA). siRNAs cause a rapid and transient knockdown of the target gene, with their effect reported to last for several days. This would fulfil the time requirement necessary for my pulse-chase experiment.[155]

In this chapter, I will describe my work to validate transient siRNA knockdown of CDA and DCTD in cell lines, as well as to find the most appropriate labelled precursors for the subsequent phase of the experiment, pulse-chase analysis (**Figure 5.1**).

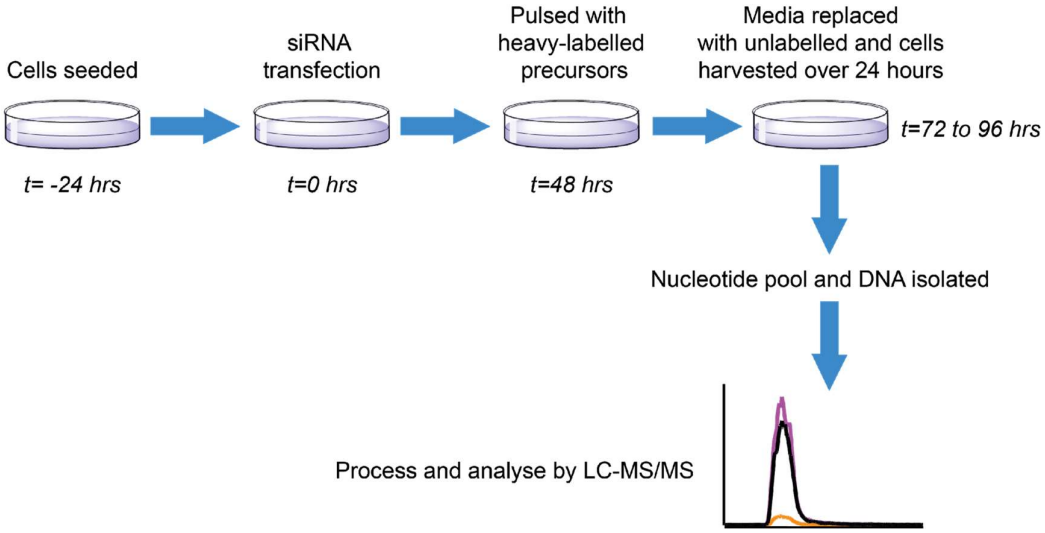


Figure 5.1 Workflow of siRNA and pulse-chase experiments. *Cells are seeded and transfected the day after. At t=48 hours after transfection, cells are treated with heavy labelled precursors. The media is changed at t=72 hours after transfection and the cells are regularly harvested for 24 hours, with metabolites and DNA analysed by LC-MS/MS.*

5.1 CDA and DCTD siRNA KD

5.1.1 Cell line choice

Enzymes are expressed at different levels in cell and so I searched for cell lines with high CDA and DCTD mRNA levels among the NCI-60 panel of cell lines on Cellminer (RNA-seq), theorising that a higher enzyme level change from WT to siRNA KD will have a more pronounced effect.[236,237]

I compared the cell lines as a percentage of the maximum $\log_2$(FPKM+1) per enzyme among the NCI-60 panel, with median $\log_2$(FPKM+1) as a benchmark (**Figure 5.2**).

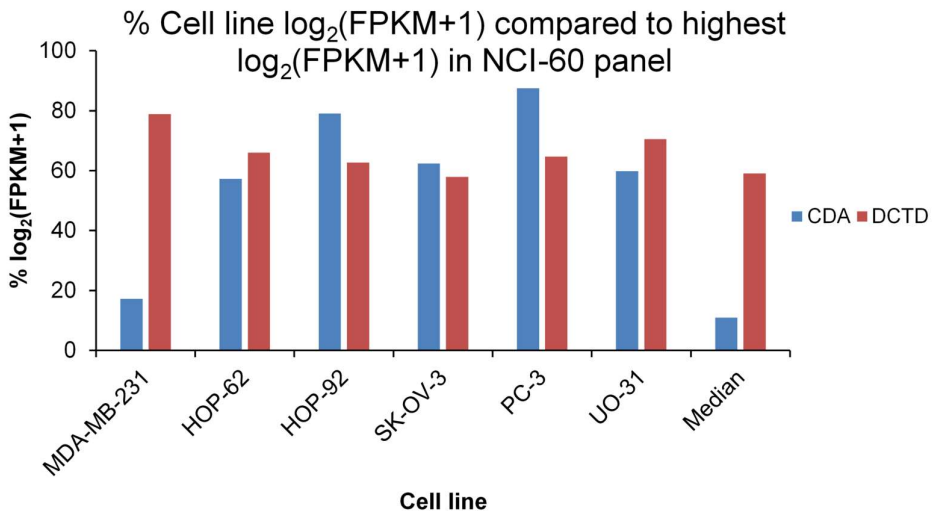


Figure 5.2 Cells with high CDA and DCTD mRNA levels compared to NCI-60 panel. *log_2(FPKM+1) of CDA (blue) and DCTD (red) as a percentage compared to highest FPKM in NCI-60 cell line panel. Data obtained from CellMiner.*[236,237]

The cell lines I shortlisted have high relative values of both CDA and DCTD compared to the NCI-60 panel. MDA MB-231 has notably lower CDA mRNA levels compared to the others, but I included it as it is still higher than the median value and the cell line has previously been studied for the effects of high CDA expression.[62] From the cell lines listed above, I chose three based on availability in the department (**Table 5.1**).

Number	Cell line	Cell line provenance	log_2(FPKM+1) CDA	log_2(FPKM +1) DCTD	Media used
1	PC-3	Prostate cancer, bone metastasis.[238]	6.547	3.982	Ham's F-12K (Kaighn's), 2 mM L-glutamine. Gibco, PN 21127
2	HOP-92	Lung non-small cell carcinoma.[239]	5.911	3.857	RPMI 1640, 2 mM L-glutamine. Gibco, PN 21875
3	MDA MB-231	Breast adenocarcinoma, derived from pleural effusion.[240]	1.288	4.85	DMEM, added L-glutamine to 2 mM and pyruvate to 1 mM. Merck, PN D5671

Table 5.1 Cell lines chosen for high CDA and DCTD (log_2(FPKM+1)).

mRNA is a good predictor of a protein abundance overall, but there are numerous exceptions.[241–243] I therefore validated the protein level of CDA and DCTD by immunoblot (**Figure 5.3;** Appendix A, **Figure 1**), also measuring the CDA level in a negative control cell line—the CDA KD cell line MDA MB-231 shCDA8 previously generated in my group.[62] A second replicate is shown in Appendix A, **Figure 1.**

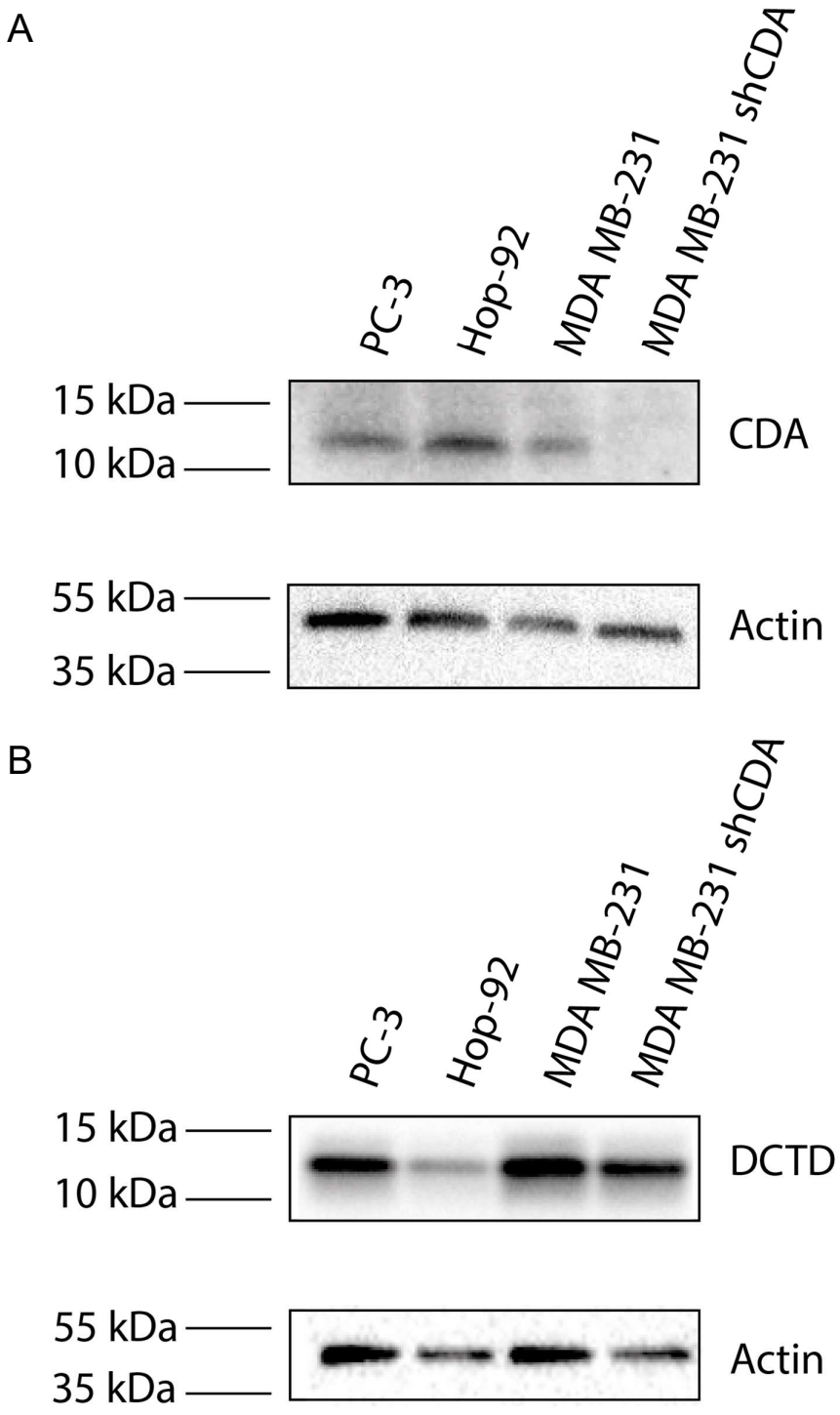


Figure 5.3 PC-3, MDA MB-231 and HOP-92 have detectable levels of CDA and DCTD protein. *Immunoblot of CDA and DCTD for all cell lines, with β-actin blot showing loading distribution.* ***A,*** *CDA protein levels* ***B,*** *DCTD protein levels.*

The cell lines all have detectable protein levels of CDA and DCTD. The cell line HOP-92 was significantly slower growing than PC-3 and MDA MB-231. As the experiment involves different time points, its different growth rate significantly complicated the experimental setup. I therefore stopped its culture and continued with the other two cell lines.

5.1.2 Validation of siRNAs for CDA and DCTD knockdown

siRNA KD is only appropriate as a tool for investigating the pyrimidine synthesis and salvage pathways if its effect lasts during the probing period of the chase phase of the heavy labelled pulse-chase experiment (**Figure 5.1**).

I used Merck Mission siRNAs and validated three orthogonal siRNA per target to avoid my research conclusions being vulnerable to off-target effects of the siRNAs.

I tested the strength of extent of knockdown in the cell lines by performing Immunoblots on samples 72 and 96 hours after transfection (**Figure 5.4** and **Figure 5.5**; Appendix A, **Figure 2** and **Figure 3**). I determined the variation in gel loading of different samples with the Ponceau stain for all the blots because I was not successful in obtaining sufficiently good bands for β-actin.

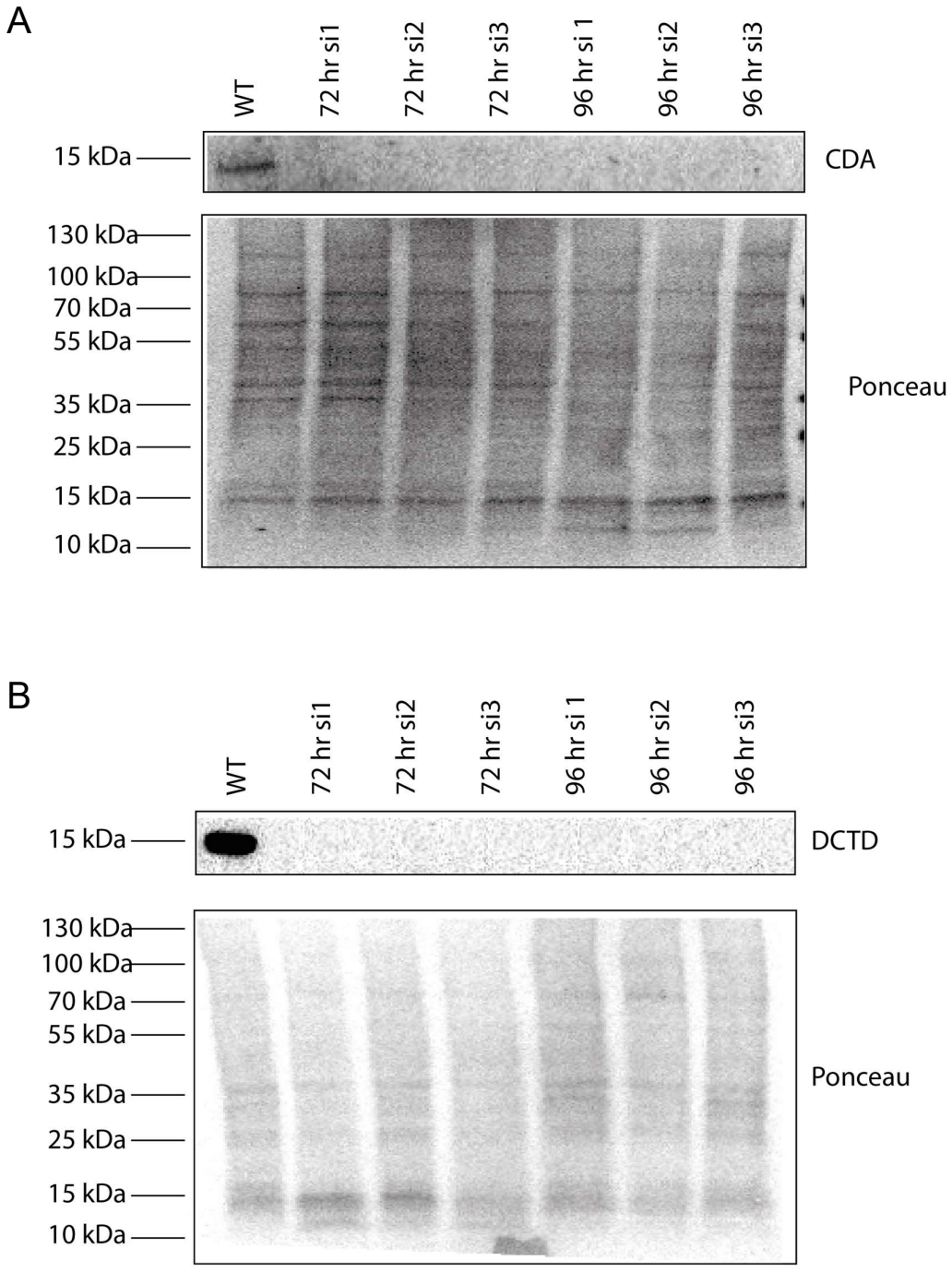


Figure 5.4 CDA and DCTD KD is stable over time in the MDA MB-231 cell line. *Immunoblot of CDA and DCTD for MDA MB-231, with Ponceau stain showing loading distribution.* ***A****, CDA KD compared to scrambled siRNA from 72 hrs to 96 hrs post exposure to siRNAs;* ***B****, DCTD KD compared to scrambled siRNA from 72 hrs to 96 hrs post exposure to siRNAs.*

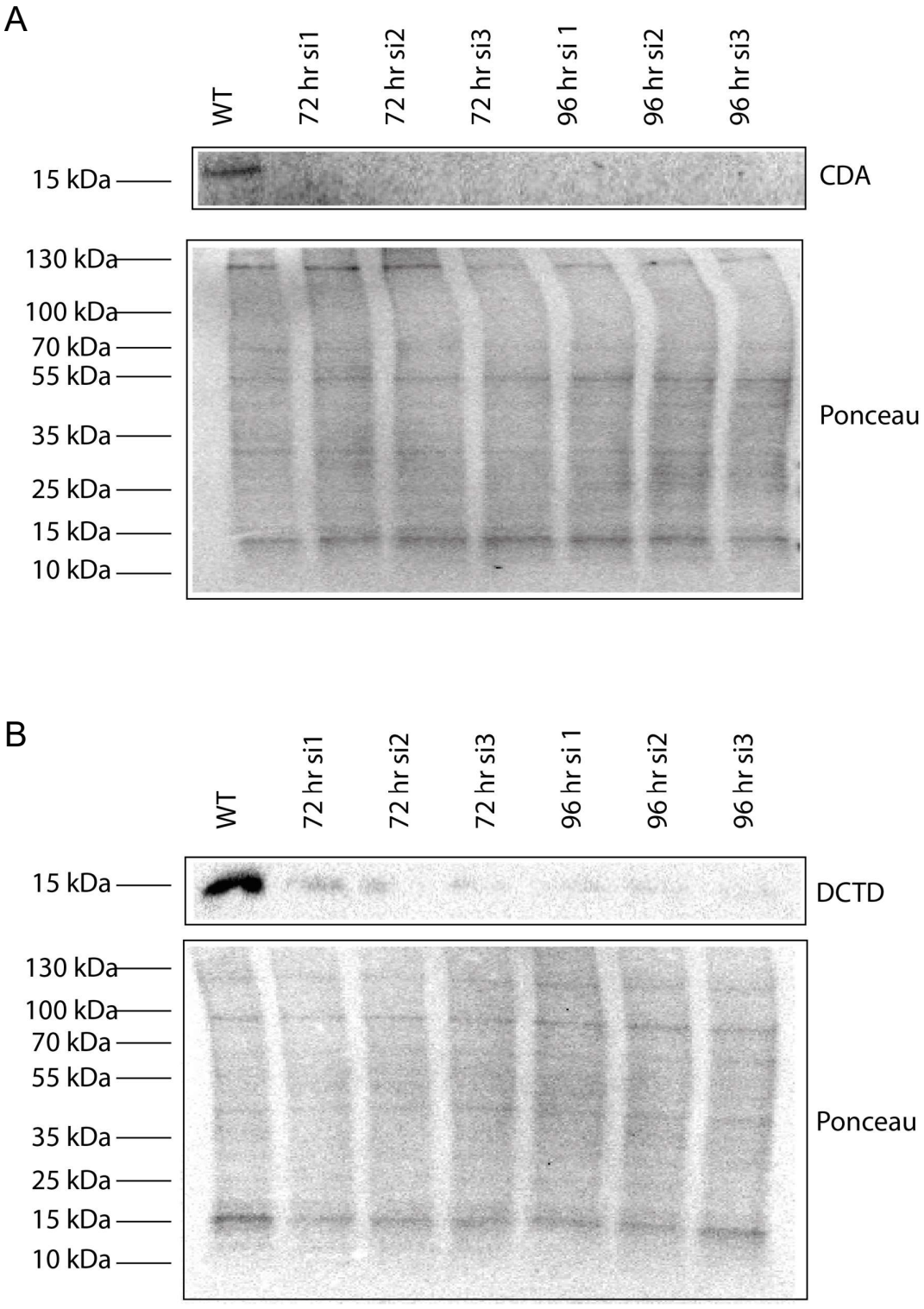


Figure 5.5 CDA and DCTD KD is stable over time in the PC-3 cell line. *Immunoblot of CDA and DCTD for MDA MB-231, with Ponceau stain showing loading distribution.* ***A****, CDA KD compared to scrambled siRNA from 72 hrs to 96 hrs post exposure to siRNAs****; B****, DCTD KD compared to scrambled siRNA from 72 hrs to 96 hrs post exposure to siRNAs****.*

The siRNAs induced knockdown for CDA and DCTD in both cell lines, lasting long enough for the planned pulse-chase experiment.

5.2 Choice of heavy labelled tracers for MS

5.2.1 Measurement of nucleotide precursors in media

Nucleosides are common in blood and extracellular environments, deriving from cell debris and food sources. As the media includes 10 % FBS, I measured the level of nucleosides in the different media of the two cell lines (**Table 5.2**) by LC-MS/MS. Both of their L-glutamine concentrations were 2 mM.

Nucleoside	Concentration (µM) in DMEM +10 % FBS (MDA MB-231)	Concentration (µM) in F-12K +10 % FBS (PC-3)
dC	0.247	0.221
dT	0.242	2.10
dA	Not detected	Not detected
dG	Not detected	Not detected
rC	1.17	1.04
rU	2.04	1.92
rA	Not detected	Not detected
rG	Not detected	Not detected
mdC	0.0127	0.0130

Table 5.2 Concentrations of nucleosides found in complete media.

The media + 10 % FBS only had pyrimidine nucleosides. Of these, rC and rU were in the low micromolar range, while dC was in hundreds of nanomolar range. dT was ten times higher in F-12K media +FBS than in DMEM (2.1 µM vs 2.42 µM), as dT is an ingredient in F-12K media. The purine nucleosides were not detected; it is unclear as to why as, at least in human serum, the pyrimidine and purine bases are both at low micromolar extracellular levels.[2] It is possible

that fetal bovine serum has less extracellular purines than pyrimidines, or that part of the filtration process of FBS manufacture sequesters serum purines.

Another nucleoside present due to FBS was the epigenetic mark mdC, around 20 times more dilute than dC. This is not unexpected as the ratio of mdC to all bases is around 1 % in mammalian tissue.[244] To ensure that changing FBS composition would not skew my experiments, I only used one batch of FBS for all experiments.

5.2.2 Suitable labelled tracers for the pyrimidine synthesis and salvage pathways

As the QQQ functions best as a targeted instrument, I would not be able to track labelling in large numbers of analytes as done for ^{13}C metabolism analysis to obtain fluxes across metabolite classes.[175,176,245] The aim of the labelled pulse-chase analysis was instead to focus on the impact of CDA and DCTD on relative nucleotide and nucleic acid labelling.

5.2.2.1 Synthesis pathway

When choosing which heavy labelled precursors to use, I considered whether the precursors or their downstream metabolites could be substrates for CDA and DCTD.

The synthesis products dCTP and CTP can be dephosphorylated by dCTP pyrophosphatase 1 (DCTPP1) and ectonucleotide pyrophosphatase 1 (ENPP1) to the monophosphates dCMP and CMP, respectively.[59,246,247] dCMP is a substrate for DCTD. dCMP and CMP are dephosphorylated further by 5'-nucleotidase to dC and rC, both of which are substrates for CDA.[10,248] Newly synthesised nucleotides may therefore be dephosphorylated and subsequently deaminated by enzymes classically belonging to the salvage pathway. By

labelling precursors for the first synthesised nucleotide, UMP, it would be possible to study the impact of CDA and DCTD in the synthesis pathway.

I focused on precursors for the pyrimidine bases and not the (deoxy)-ribose group as CDA and DCTD affect the nucleobases and not the sugar. I did not consider precursors that would give single-atom labelling to avoid interference from labelled metabolites with natural abundance isotope levels. The nucleobase precursors for UMP are L-glutamine, oxaloacetate, aspartate and orotate. They were appropriate candidates as their downstream metabolites will retain a greater proportion of labelling.

L-glutamine

L-glutamine, a major exogenous nutrient, is metabolised in the Krebs cycle to oxaloacetate and subsequently to aspartate, after which it is metabolised to orotate by CAD.

Oxaloacetate

Oxaloacetate is a metabolic intermediate and can be formed in the Krebs cycle. It is aminated to aspartate and then converted to orotate by CAD.

Aspartate

Aspartic acid is the substrate of CAD and is formed by amination of oxaloacetate.

Orotate

Orotate acid is produced by CAD and is the substrate for UMPS to form UMP. It is present in human blood plasma at low to mid micromolar concentrations.[249]

L-glutamine and oxaloacetate are active metabolites for a series of processes aside from nucleotide biosynthesis, including protein synthesis, gluconeogenesis and the Krebs cycle.[245,250] Aspartate is involved in protein synthesis and the urea cycle.[251] Given the multiple downstream fates of these molecules, one may consequently need to expose cells

to a high concentration of their stable isotopologues to see significant labelling of downstream nucleotides. Orotate is only active in nucleotide metabolism and so labelled orotate will not be depleted by competing pathways.

L-glutamine is present at 2 mM in media. As L-glutamine is added to media as a preparation step, it is possible to add significant amounts labelled L-glutamine and correspondingly less unlabelled L-glutamine to maintain the total L-glutamine concentration invariant. The other precursor candidates are not added to media and so greater additions of labelled precursors may be disruptive to nucleotide metabolism. I tested $^{13}C_5$ L-glutamine and $^{13}C^{15}N_2$ orotic acid for incorporation in the synthesis pathway (**Figure 5.6**).

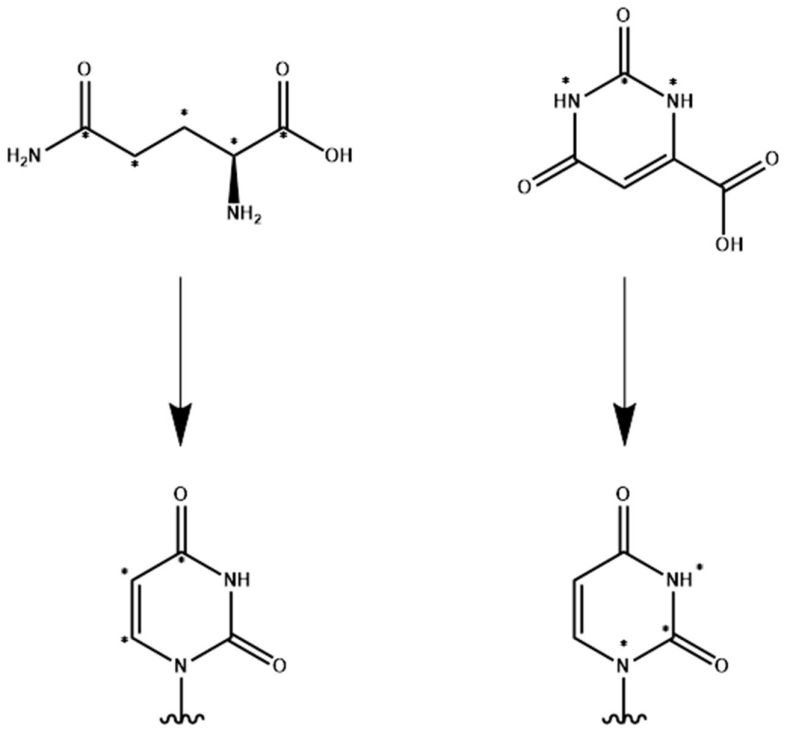


Figure 5.6 **Labelled L-glutamine and orotic acid are metabolised to form labelled pyrimidines.** *$^{13}C_5$ L-glutamine forms $^{13}C_3$ pyrimidine bases and $^{13}C^{15}N_2$ orotic acid forms $^{13}C^{15}N_2$ pyrimidine bases. Heavy atoms are marked with an asterisk.*

They both give labelling of +3 Da. L-glutamine also gives labelling of +2 Da and +1 Da if it is further metabolised in the Krebs cycle, but those labels will experience stronger interference by natural abundance isotopologues.

I exposed 200,000 seeded cells for 24 hours with either labelled L-glutamine at 500 µM (20 % total amount of L-glutamine) or labelled orotic acid at 10 µM and measured their DNA incorporation. I compared the +3 Da signal to that derived from the natural isotopologues of cells present with no heavy labelled treatment (**Figure 5.7A**). The difference in initial concentration is reflective of L-glutamine's high natural concentration and orotic acid's absence. In MDA MB-231, L-glutamine incorporation as dC +3 Da significantly higher than for orotic acid incorporation as dC +3 Da ($p=0.0217$). The exogenous labelled molecule concentrations are different, but it is less disruptive to use 500 µM of labelled L-glutamine than it is to use 10 µM labelled orotic acid as it is possible to reduce the amount of unlabelled L-glutamine in the media so that the total concentration is invariant. L-glutamine incorporation in PC-3 was higher than in MDA MB-231 (**Figure 5.7B**), possibly due to a more active synthesis pathway in PC-3; the rate limiting enzyme for pyrimidine *de novo* synthesis, CAD, is more highly expressed in PC-3 than MDA MB-231 (Cellminer gives $\log_2$(FPKM+1) values of 3.531 and 0.628, respectively).

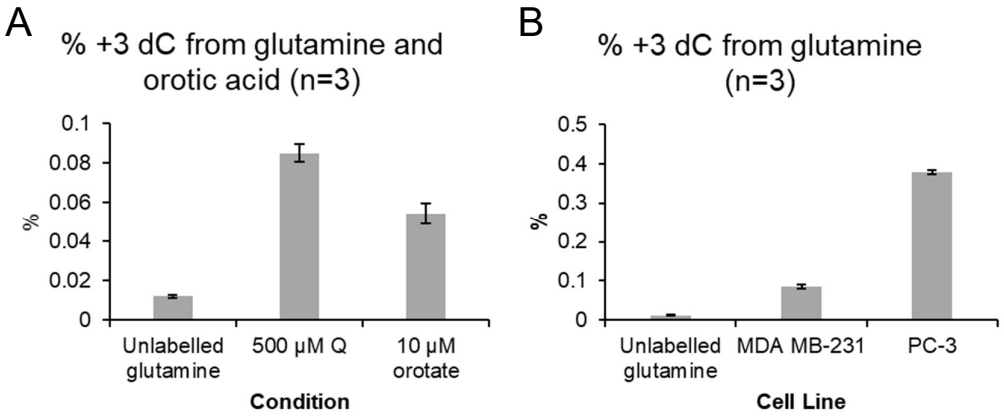


Figure 5.7 Incorporation of heavy labelled synthesis pathway tracers as dC. *Cells were treated with labelled tracers for 24 hrs in technical triplicates, mean values shown, error bars give +-1 standard deviation.* ***A,*** *comparison of +3 Da dC signal in untreated cells, MDA MB-231 with 500 µM labelled L-glutamine or 10 µM labelled orotic acid;* ***B,*** *comparison of +3 Da dC signal in untreated cells, MDA MB-231 and PC-3, each treated with 500 µM labelled L-glutamine.*

5.2.2.2 Salvage pathway

In the salvage pathway, CDA deaminates dC and rC, whereas DCTD deaminates dCMP but not CMP. From rC, all pyrimidine nucleotide triphosphates can be made and dCMP is upstream of all pyrimidine deoxyribonucleotide triphosphates.

Of dC, dCMP and rC, only dC and rC are imported by endogenous transporters. I chose to test incorporation of $^{13}C_9$ dC and $^{13}C_5$ rC (**Figure 5.8**).

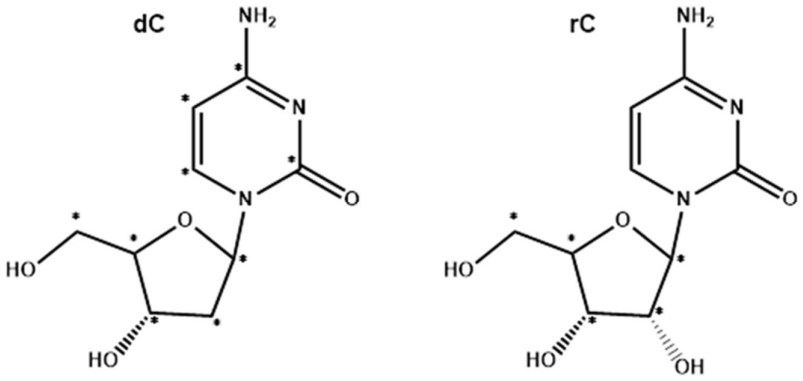


Figure 5.8 Labelling of $^{13}C_9$ dC (dC +9 Da) and $^{13}C_5$ rC (rC +5 Da). Heavy atoms are marked with an asterisk

Given the concentrations of pyrimidines in media are of the order of magnitude of 1 µM (**Table 5.2**), I exposed 200,000 cells to 1 µM dC and rC for 24 hours and measured incorporation (**Figure 5.9**).

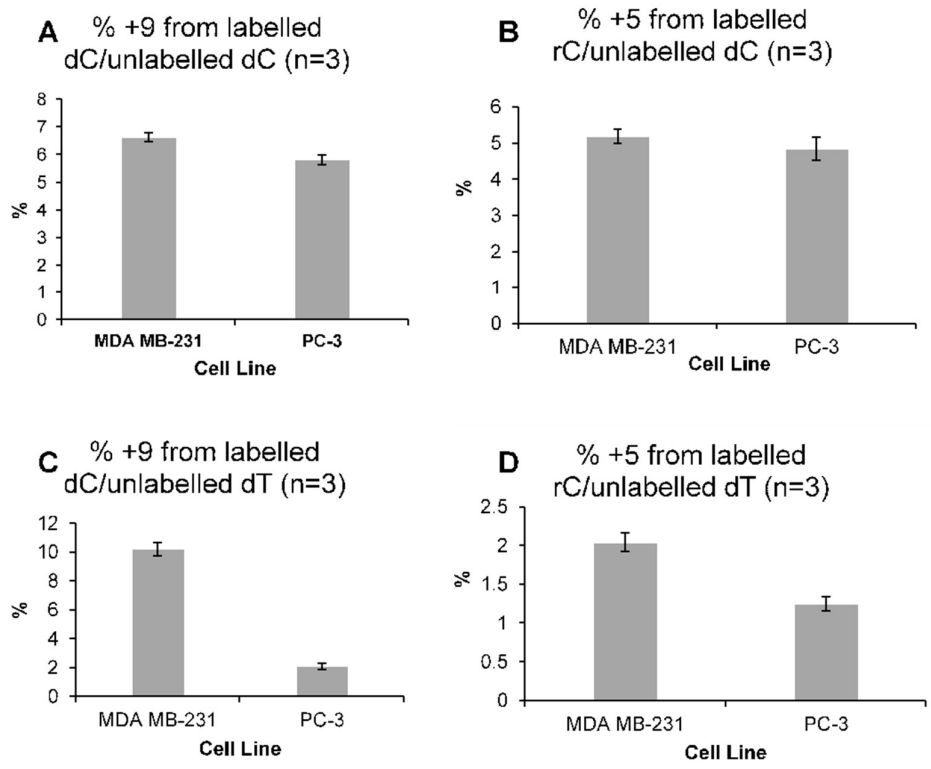
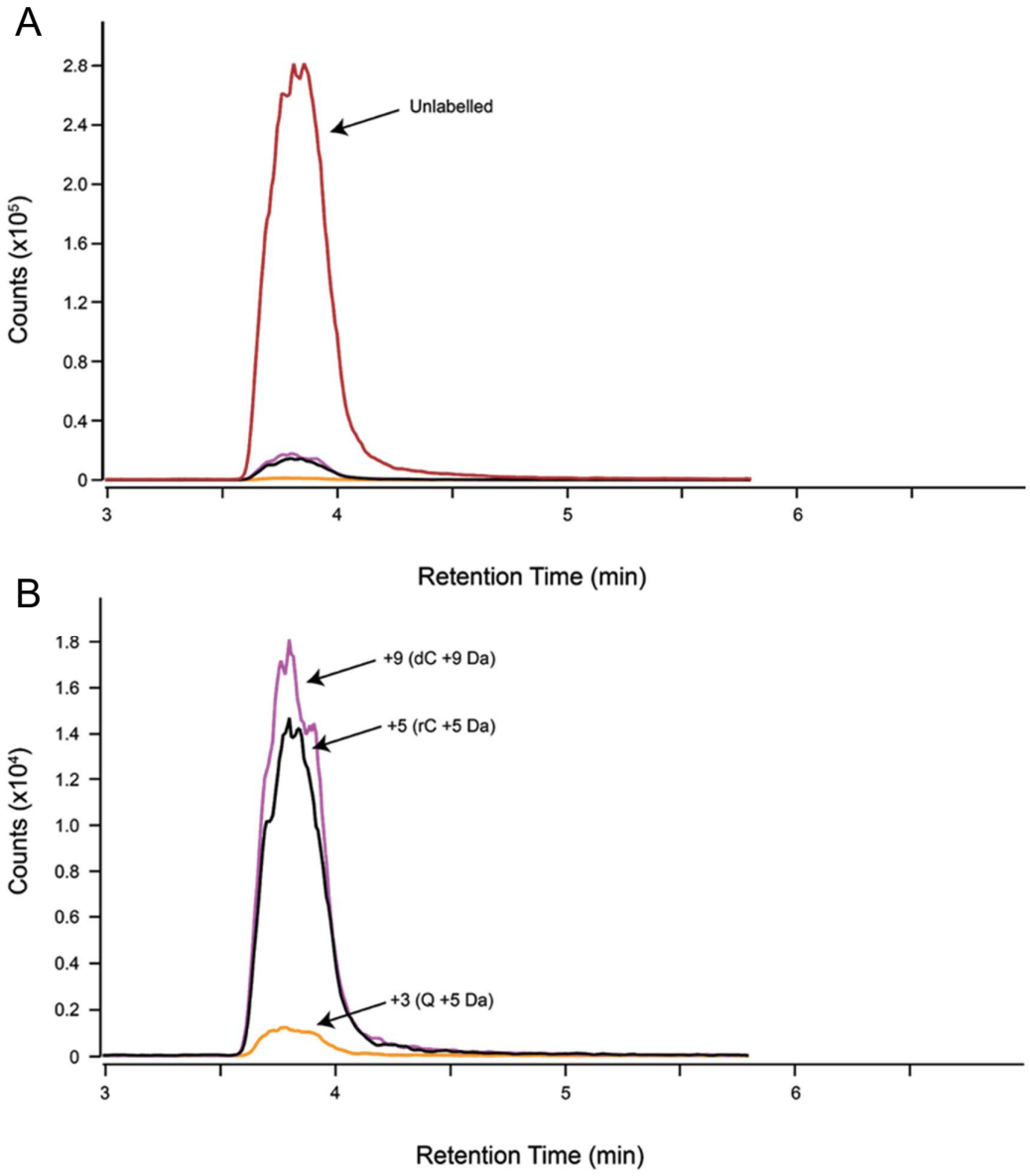


Figure 5.9 % incorporation of $^{13}C_9$ dC and $^{13}C_5$ rC. into DNA via incorporation of heavy labelled synthesis pathway tracers as dC. *Cells were treated with labelled tracers for 24 hrs in technical triplicates, mean values shown, error bars give +-1 standard deviation.* ***A,*** *incorporation of $^{13}C_9$ dC as dC;* ***B,*** *incorporation of $^{13}C_5$ rC as dC;* ***C,*** *incorporation of $^{13}C_9$ dC as dT;* ***D,*** *incorporation of $^{13}C_5$ rC as dT.*

Labelled dC and rC were incorporated in DNA as dC and dT at measurable levels. Incorporation as dT was lower for PC-3 than for MDA MB-231 (**Figure 5.9C**), which I hypothesise is due to the high concentration of dT in F-12K media.

5.2.3 Simultaneous discrimination of synthesis and salvage tracer incorporation

The synthesis and salvage heavy labelled nucleotide precursors have different labelling and so can be added together to a cell line with simultaneous incorporation measurement possible (**Figure 5.10**).


A
Counts (x10⁵)
2.8
2.4
2.0
1.6
1.2
0.8
0.4
0
Unlabelled
3
4
5
6
Retention Time (min)
B
Counts (x10⁴)
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0
+9 (dC +9 Da)
+5 (rC +5 Da)
+3 (Q +5 Da)
3
4
5
6
Retention Time (min)

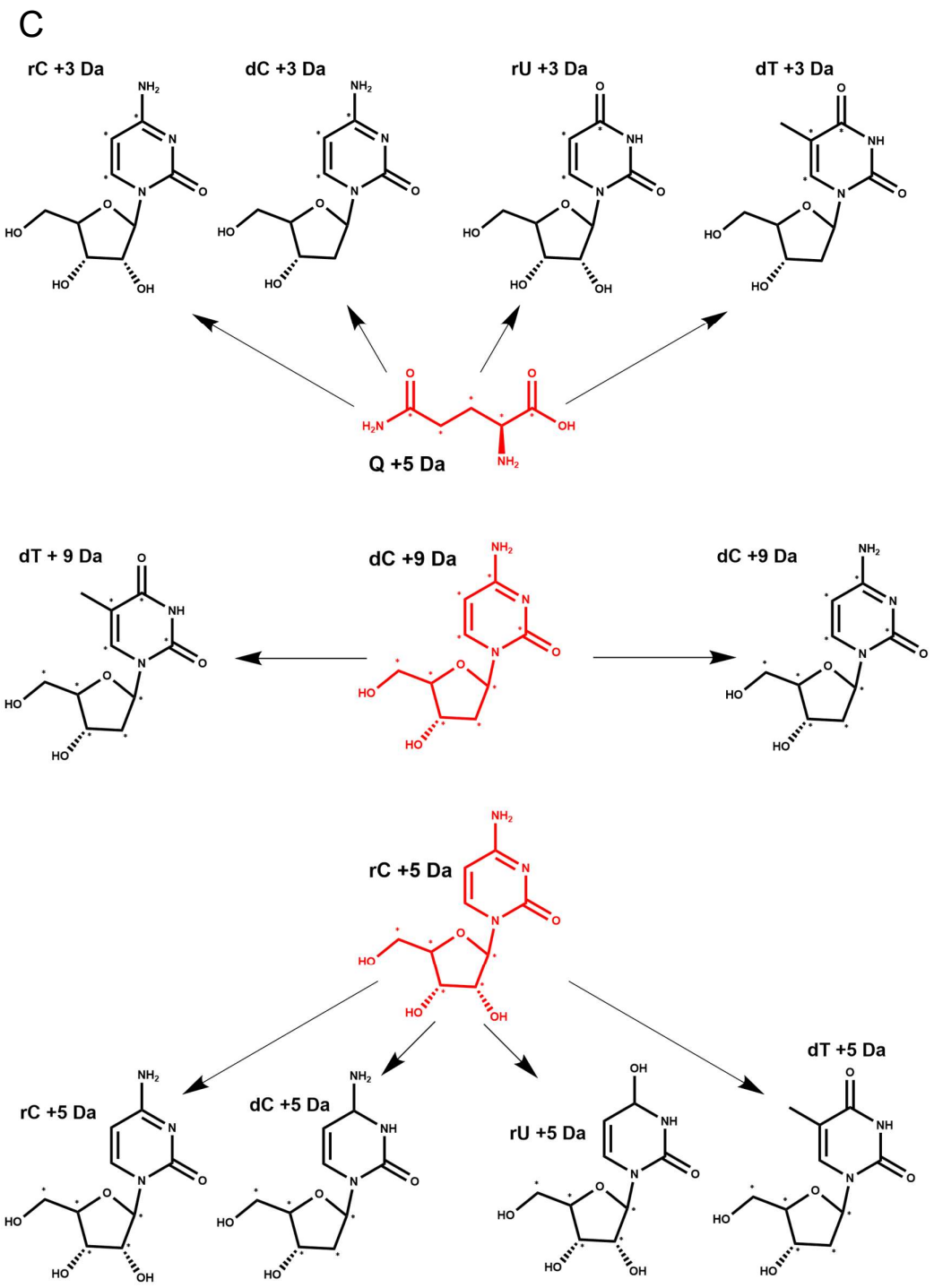


Figure 5.10 Incorporation of labelled nucleotide synthesis and salvage precursors as dC in nucleic acids. A, *peaks of labelled dC compared to unlabelled dC in PC-3 cell line (red);* ***B,*** *zoom in to peaks of labelled dC in PC-3 cell line from* $^{13}C_5$ *L-glutamine (yellow), from* $^{13}C_5$ *rC (black) and from* $^{13}C_9$ *dC (magenta);* ***C,*** *incorporation fates of heavy labelled L-glutamine (Q), dC and rC.*

5.3 Discussion

The validation of siRNA KD of CDA and DCTD, together with the use of isotopically labelled dC, rC and L-glutamine for synthesis and salvage pathway tracing, allowed progression to the final project experiment—of performing pulse-chase labelling and MS analysis of both nucleic acid composition and nucleotide pool to probe the contributions from CDA and DCTD to nucleotide metabolism (**Figure 5.1**).

5.3.1 CDA and DCTD KD

To obtain a large perturbation in CDA and DCTD protein levels by siRNA, I stratified the NCI-60 cell line panel by CDA and DCTD mRNA levels using CellMiner (**Figure 5.2**). Of the high CDA and DCTD mRNA cell lines, I tested the CDA and DCTD protein levels of MDA MB-231, PC-3 and HOP-92 cells by immunoblot and found detectable amounts in all three cell lines (**Figure 5.3**).

siRNA KD is reported to last several days, which was necessary for my experiments of siRNA KD followed by pulse-chase MS analysis.[155] I tested three orthogonal siRNAs per enzyme target against a scrambled siRNA that should give no KD, verifying that the KD lasted up to 96 hrs post transfection with siRNA in MDA MB-231 and PC-3 (**Figure 5.4** and **Figure 5.5**).

5.3.2 Heavy labelled tracers for MS

As the deaminases reduce the nucleotide pool cytidine-uridine/thymidine nucleotide ratio, heavy label nucleotide precursor tracing is suitable for investigation of the role of CDA and DCTD in the pyrimidine synthesis and salvage pathway.[1,18] Use of the appropriate labelled

precursors coupled with individual enzyme level manipulation and MS analysis can shed light on their impact on the nucleotide pool.

To find the most appropriate labelled nucleotide precursors, I analysed the cell media for exogenous nucleoside concentration, with the nucleosides dT, dC and rC found to be in the high nanomolar to low micromolar range (**Table 5.2**).

I then tested the labelling incorporation in DNA, observing that of the *de novo* synthesis precursors, 500 µM L-glutamine +5 Da (+3 Da pyrimidine labelling) gave more incorporation than 10 µM orotic acid +3 Da (**Figure 5.7**). L-glutamine was more appropriate as the amount of unlabelled L-glutamine can be adjusted to keep an invariant total L-glutamine concentration, which cannot be done for orotic acid. The labelled salvage nucleotide precursors dC +9 Da and rC +5 Da showed detectable labelling (**Figure 5.9**) and give complementary information on the substrates of CDA and DCTD.

All labelled nucleotide precursors can be used simultaneously due to their different labels. The concentrations to be used for the pulse-chase experiments were 500 µM $^{13}C_5$ L-glutamine, 1 µM $^{13}C_5$ rC and 1 µM $^{13}C_9$ dC.

6. Role of CDA and DCTD in Nucleotide Pool Metabolism

6.1 Introduction

It has been unreported as to whether CDA or DCTD is the principal deaminase of pyrimidine nucleotides. Despite their similar mechanisms of action, CDA is expressed highly in certain tissues and low in others, while DCTD has low tissue specificity (**Figure 1.12**). Additionally, high CDA expression has been associated with unfavourable outcomes in cancers, while high DCTD expression has been associated with mixed outcomes.[94–97,106,252]

The siRNA KD of CDA and DCTD, coupled with sensitive LC-MS/MS methods for the nucleotide pool and nucleic acid composition analysis and the use of labelled nucleotide precursors allowed me to investigate where either CDA or DCTD is the principal deaminase of pyrimidine nucleotides and whether they deaminate newly synthesised cytidine nucleotides.

As siRNAs can have off-target effects and impacts on cellular function, I employed two orthogonal siRNAs (targeting different gene sequences) for both CDA and DCTD to avoid reliance on only one siRNA. I chose CDA siRNA1 and CDA siRNA3, DCTD siRNA1 and DCTD siRNA2 from the six siRNAs I validated (**5.1.2**). In the following results, I only consider effects seen in both siRNAs for each enzyme as biologically indicative.

All graphs derived from the data in this experiment, shown in the following chapter or not, can be found in Appendix B (sheets highlighted in blue refer to nucleoside analysis, sheets highlighted in orange refer to free nucleotide analysis).

6.2 Sample DNA and nucleotide pool harvest

The nucleotide pool and nucleic acid were extracted from the same cell samples to maximise sample correspondence and to streamline the workflow. To obtain both the free nucleotide pool and nucleic acid from the same sample, I had to modify the extraction process, combining those described in **2.4.1** and **2.7.2**.

As a summary of the experiment workflow, the cells were seeded and transfected the day after (**Figure 5.1**). At t=48 hours after transfection, cells were treated with the heavy-labelled precursors $^{13}C_9$ dC (1 µM), $^{13}C_5$ rC (1 µM), and $^{13}C_5$ L-glutamine (500 µM). The media was changed at t=72 hours after transfection and the cells were harvested at the time points 0, 1, 2, 4, 6, 12 and 24 hours after replacement of media (**Figure 6.1**). Both the nucleotide pool and nucleic acid were extracted from the samples.

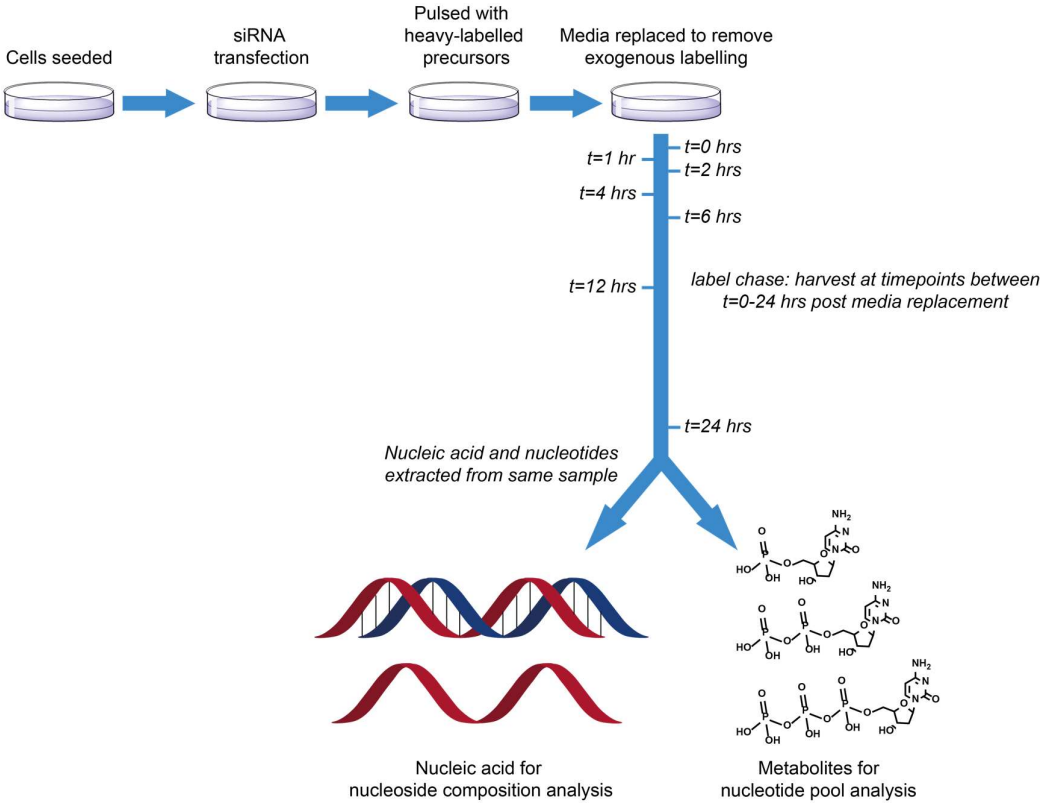


Figure 6.1 Samples harvested at set times over 24 hrs in pulse-chase experiment. *After siRNA transfection, the samples were exposed to 1 µM dC +9 Da, 1 µM rC +5 Da and 500 µM L-glutamine +5 Da for 24 hrs. The media was replaced and the samples were harvested at 0, 1, 2, 4, 6, 12 and 24 hrs after replacement of media. Both the nucleotide pool and nucleic acid were extracted from the samples for label incorporation analysis.*

This was done through washing the cells with PBS and snap freezing in liquid nitrogen. The nucleotide pool was extracted with 50 % ACN. The nucleic acid fraction remained in the pellet of insoluble material. To remove residual extraction solvent, I dried the samples on a heat block at 95 °C. Given the large number of samples to process, I used the Zymo Research Quick-DNA 96 Kit for DNA isolation. I hydrolysed all the DNA sample obtained to nucleosides to maximise nucleoside signal.

6.3 LC-MS/MS signal quality for nucleosides and nucleotides

I created dMRM MS nucleoside and nucleotide methods to measure the proportion of labelling of nucleic acids and the free nucleotide pool across phosphorylation levels, having formed through pyrimidine nucleotide metabolism (**Figure 5.10C**).

I assessed the level of labelled precursor incorporation in nucleosides derived from extracted nucleic acids in the samples at t=0 hrs after replacing the heavy labelled media with the unlabelled media. There was measurable labelling signal from most samples in the deoxyribonucleosides dC, dT, and the ribonucleoside rC (**Figure 6.2A-C**). rU had too low sensitivity to obtain usable data for labelled incorporation. The abundant presence of RNA nucleosides in the extracted nucleic acid section was surprising, as the DNA isolation kit claims to be specific for DNA; it may simply reduce RNA presence instead. Being able to quantify labelling in ribonucleosides was a positive outcome as it allowed analysis of the impact of CDA

and DCTD KD on RNA nucleotide metabolism. It must be noted, however, that there are different classes of RNA, such as messenger RNA (mRNA), transfer RNA (tRNA), ribosomal RNA (rRNA), with different functions, structures and interacting proteins. As RNA classes have different turnover rates, preferable retention on the isolation column for one class of RNA may skew labelled nucleotide incorporation results.[75,253,254] This notwithstanding, there is approximately between 3 and 10 times as much RNA as DNA in human cells, whereas the UV signals of dA and rA in my samples approached equality in signal (data not shown).[255,256] rRNA contributes to 80 % of mammalian RNA and another 15 % is tRNA. As the reduction in ribonucleoside signal through DNA purification was modest, I expect the ribonucleosides observed by MS to still be primarily derived from rRNA and tRNA, not mRNA.[257]

Both ribonucleotides and deoxyribonucleotides were detected by LC-MS/MS analysis of the nucleotide pool, but the deoxyribonucleotide signals were weaker due to their lower cellular concentration (**Figure 6.2D** and **E**) and had high standard deviations of heavy labelled incorporation proportions. The labelled signals of the CMP, CDP, CTP, UMP, UDP and UTP nucleotides had lower variability and could give insights into the roles of CDA and DCTD.

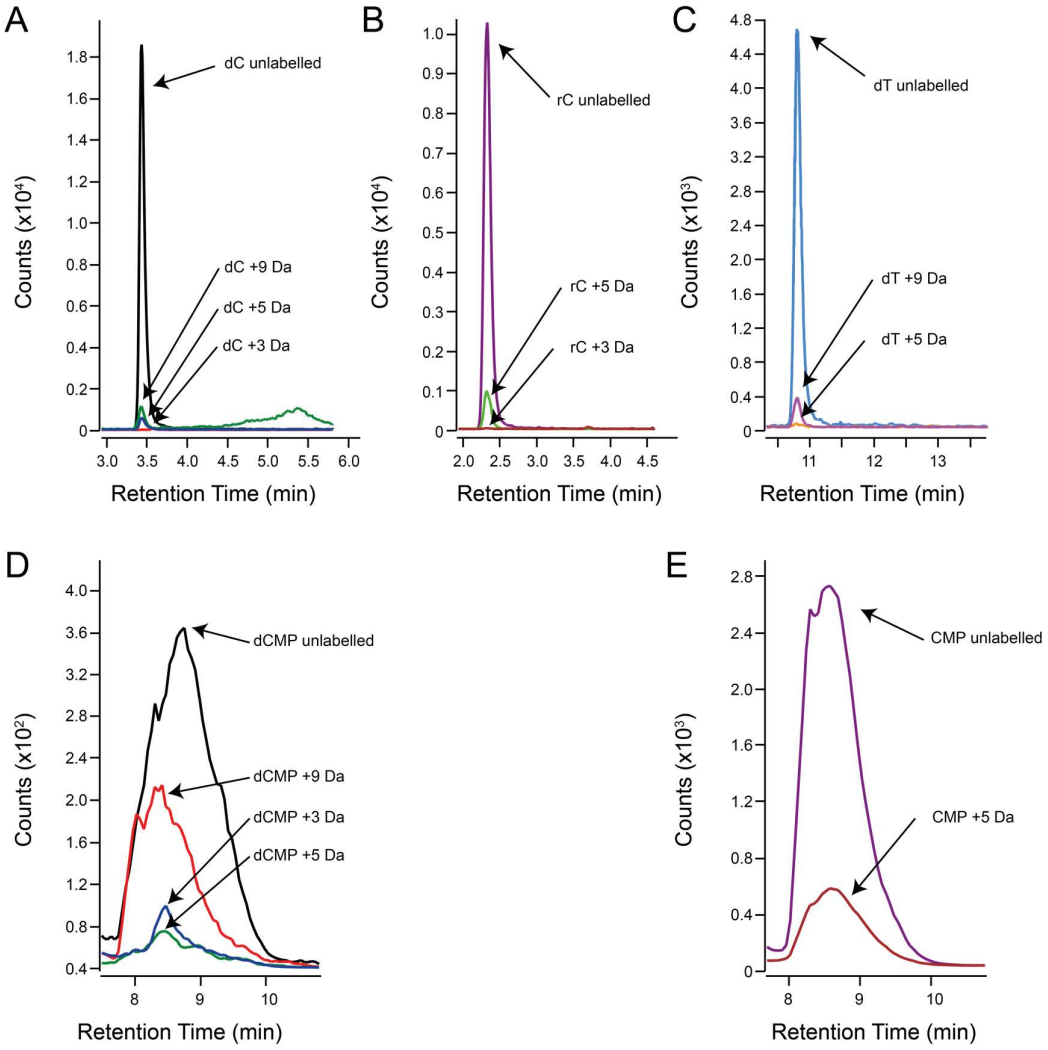


Figure 6.2 Precursor labelling is observed in DNA/RNA nucleosides and in the nucleotide pool. *Label distributions in MDA MB-231 exposed to scrambled siRNA at t=0 hrs post media replacement.* ***A,*** *labelling of dC (dC H^+) in DNA;* ***B,*** *labelling of rC (rC H^+) in RNA;* ***C,*** *labelling of dT (dT Na^+) in DNA;* ***D,*** *labelling of dCMP (dCMP H^+) in nucleotide pool;* ***E,*** *labelling of CMP (CMP H^+) in nucleotide pool.*

6.4 Observations of nucleotide metabolism from pulse-chase analysis

In addition to probing the roles of CDA and DCTD, the pulse-chase analysis provided insights into the rate of metabolism of the free nucleotides pool nucleic acids.

6.4.1 Nucleic acid labelling after 24 hours exposure

First, I examined to what extent the labelled precursors are incorporated in DNA and RNA in both cell lines (**Table 6.1**). After 24 hours of labelled precursor exposure, PC-3 had greater incorporation of the +3 Da label (derived from L-glutamine +5 Da) than MDA MB-231, which may indicate a stronger reliance on the synthesis pathway in nucleotide metabolism. This is supported by the relative incorporation of salvage-derived precursors: MDA MB-231 had higher incorporation of the salvage-derived labels +9 Da and +5 Da (dC +9 Da and rC +5 Da) than did PC-3.

The difference in incorporation of salvage-derived labels was most pronounced for incorporation in dT. This may be because the high concentration of dT present in the media used for PC-3 cell lines (Ham's F-12K (Kaighn's)) competed with the deaminated product of the labelled cytidine nucleotides for incorporation into DNA as dT (**5.2.1**).

In both cell lines, dC and dT had greater derivation from labelled dC compared to labelled rC; while exogenous dC was around 75 % labelled (1 µM labelled dC, 247 nM unlabelled dC present in media), with only around 50 % labelling of rC (1 µM labelled rC, 1.17 µM unlabelled C present in media), there is a greater proportion of labelling of dC and dT in DNA from dC compared to rC per unit concentration of exogenous nucleoside. This may be due the labelled rC being a precursor for both ribonucleotides and deoxyribonucleotides, with consequently less available for forming deoxyribonucleotides compared to labelled dC, or it may be due to faster kinetics for labelled dC to be incorporated as a DNA monomer as there are fewer enzymatic steps required.

The values of heavy labelled precursor incorporation are lower than they were when measured in **5.2.2.1** and **5.2.2.2**; it may be explained by slower cell growth, which could be a consequence of siRNA treatment.

			% incorporation in nucleic acids					
			MDA MB-231 cell line			*PC-3 cell line*		
Labelled precursor	**Shift (Da)**	**Exogenous concentration**	**dC**	**dT**	**rC**	**dC**	**dT**	**rC**
L-glutamine	+3	500 µM (25 % of total L-glutamine)	0.030 +/- 0.028	**ND**	0.14 +/- 0.028	0.17 +/- 0.032	**ND**	0.49 +/- 0.12
rC	+5	1 µM (25 % of total rC)	3.3 +/- 0.24	0.96 +/- 0.18	9.1 +/- 0.66	2.2 +/- 0.12	0.62 +/- 0.095	4.5 +/- 0.26
dC	+9	1 µM (75 % of total dC)	5.4 +/- 0.72	6.4 +/- 0.59	**N/A**	4.3 +/- 0.91	1.1 +/- 0.17	**N/A**

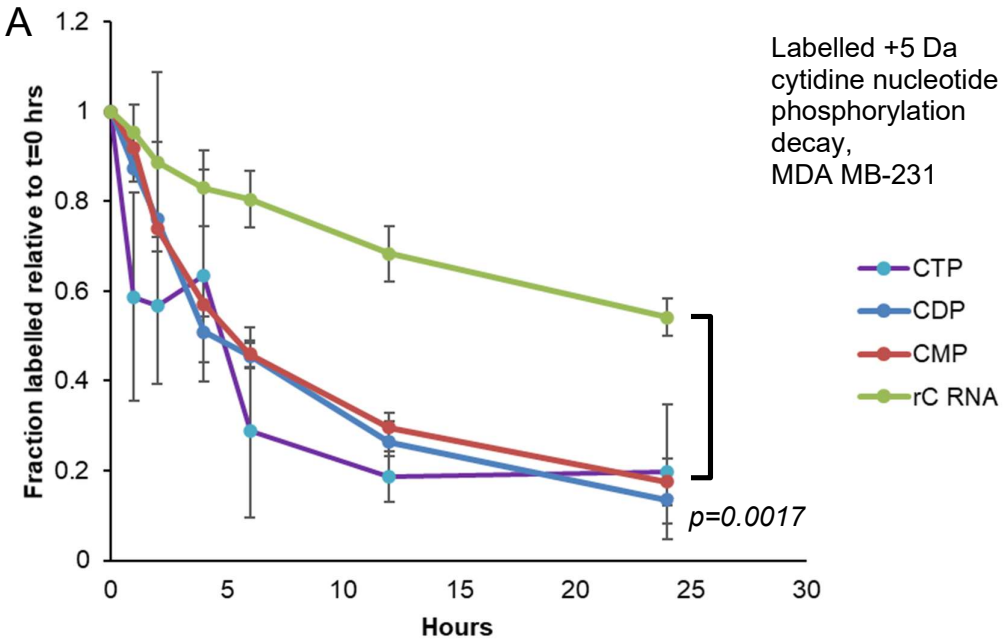
Table 6.1 Percentage incorporation of heavy labelled precursors in nucleic acid components.

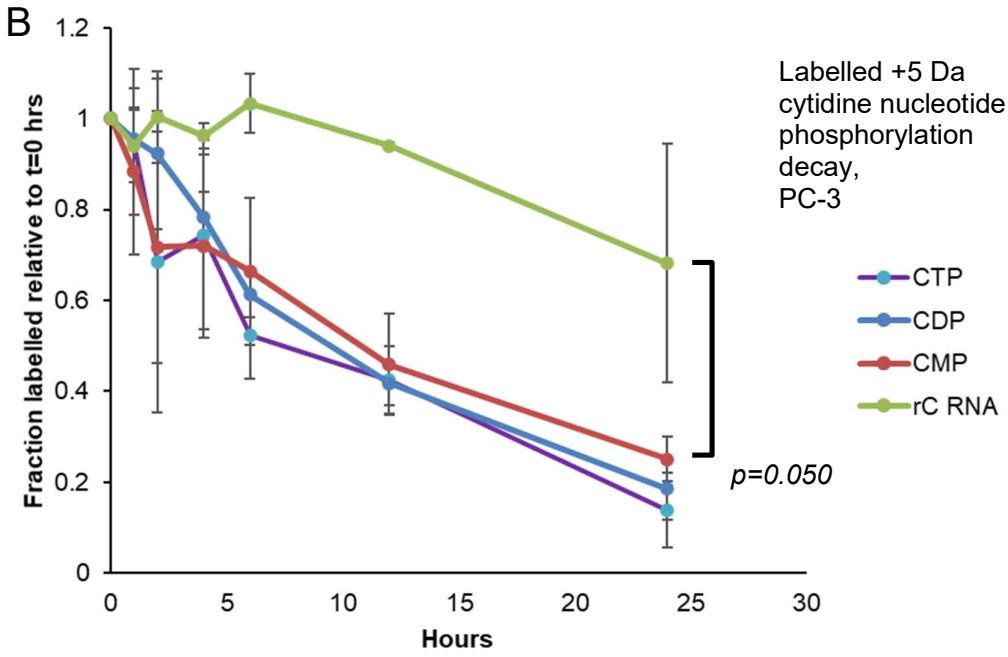
The labelling from L-glutamine (+3 Da), rC (+5 Da) and dC (+9 Da) given as a percentage compared to the total amount of each component, at t= 0 hrs post 24 hrs heavy label pulse in cells exposed to scrambled siRNA. Data is from 3 biological replicates, with +/- 1 standard deviation. ***ND*** *means not detected.* ***N/A*** *means not applicable due to the label not entering certain pathways.*

6.4.2 Free nucleotides have faster turnover than nucleic acids

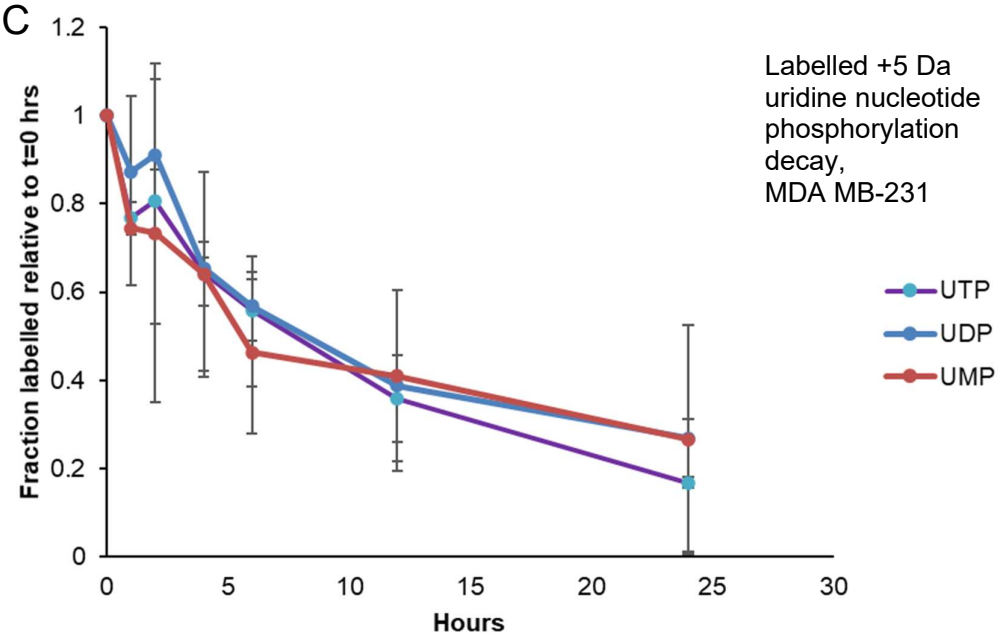
To investigate whether nucleic acids and nucleotides of different phosphorylation states had different rates of turnover, I measured the decay in label incorporation over time in differently phosphorylated nucleotides in the nucleotide pool as well as in nucleosides derived from DNA and RNA. I expected that nucleic acid labelling would have a longer half-life as it isn't a metabolic intermediate in contrast to free nucleotides, but I did not know whether the deamination to make uridine nucleotides would affect the decay rate.

RNA labelled cytidine decayed more than three times more slowly compared to those of cytidine mono-, di- and tri-phosphates in both the MDA MB-231 and PC-3 cell lines (**Figure 6.3A** and **B**). This is in agreement with reports that rRNA and tRNA have half-lives in the range of hours and days.[254] Notably, labelling decay did not vary across the nucleotide phosphorylation states for both cytidine and uridine nucleotides (**Figure 6.3A-D**) in either cell line, with a labelling half-life of around 6 hours. The overlap in values may be explained by the possibility that nucleotides in cells are metabolised too rapidly for the resolution of my time points. It is unlikely to be due to extensive degradation of nucleotide triphosphates into lower phosphate analogues during sample preparation, as I had observed low amounts of nucleotide degradation in the nucleotide extraction process (**Figure 4.4**).


A
Fraction labelled relative to t=0 hrs
1.2
1
0.8
0.6
0.4
0.2
0
0
5
10
15
20
25
30
Hours
Labelled +5 Da cytidine nucleotide phosphorylation decay, MDA MB-231
CTP
CDP
CMP
rC RNA
p=0.0017

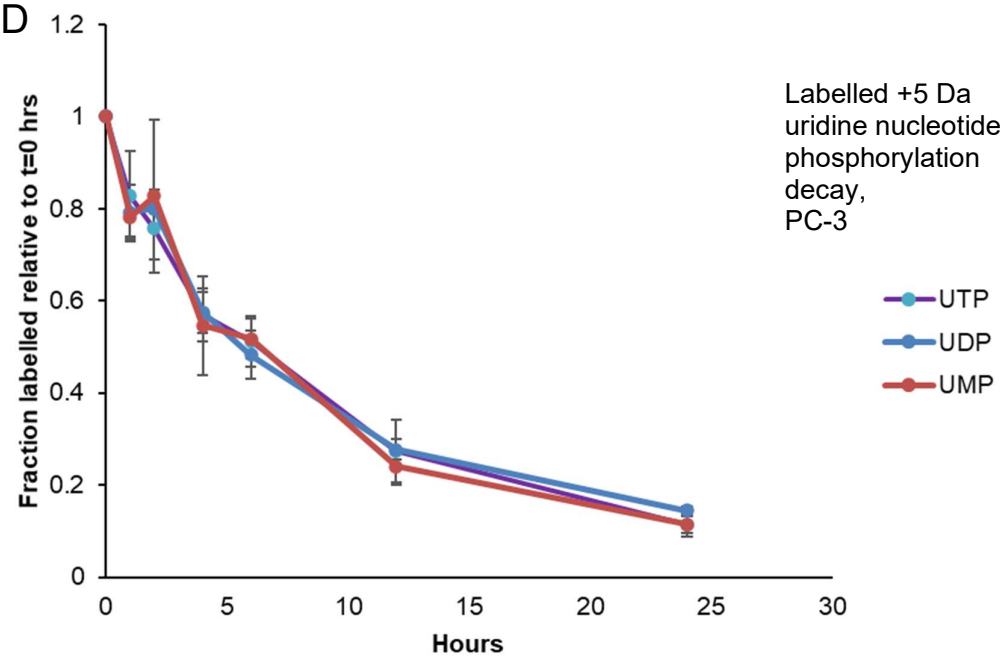

B
Fraction labelled relative to t=0 hrs
1.2
1
0.8
0.6
0.4
0.2
0
0
5
10
15
20
25
30
Hours
Labelled +5 Da cytidine nucleotide phosphorylation decay, PC-3
CTP
CDP
CMP
rC RNA
p=0.050

C



Labelled +5 Da uridine nucleotide phosphorylation decay, MDA MB-231

D



Labelled +5 Da uridine nucleotide phosphorylation decay, PC-3

Figure 6.3 rC/rU +5 Da nucleotide fractional turnover is similar across phosphorylation states. *+5 Da incorporation levels in cells exposed to scrambled siRNA decay at similar rates for mono- di- and tri-phosphates for both cytidine and uridine ribonucleotides. +5 Da rC proportion in RNA reduces at a slower rate than in nucleotides. 3 biological replicates normalised to t=0 hrs per each biological replicate, standard deviation is shown.* ***A,*** *+5 Da proportion reduction for cytidine ribonucleotides and cytidine in RNA in MDA MB-231 cell line;* ***B,*** *+5 Da proportion reduction for cytidine ribonucleotides and cytidine in RNA in PC-3 cell line; decay for +5 Da uridine ribonucleotides.* ***C,*** *+5 Da proportion reduction for uridine ribonucleotides in MDA MB-231 cell line;* ***D,*** *+5 Da proportion reduction for cytidine ribonucleotides in PC-3 cell line.*

6.4.3 RNA is turned over faster than DNA

I compared labelling decay of RNA (**Figure 6.3A** and **B**) with that of DNA, expecting to see a difference in turnover as the majority of new DNA base incorporation is through DNA replication, which occurs once per cell cycle. By contrast, RNA turnover is not solely limited to replication and so was expected to have a higher turnover.[75,253,254] While label incorporation in RNA did decay over time, the label proportion in DNA did not (**Figure 6.4 A** and **B**).

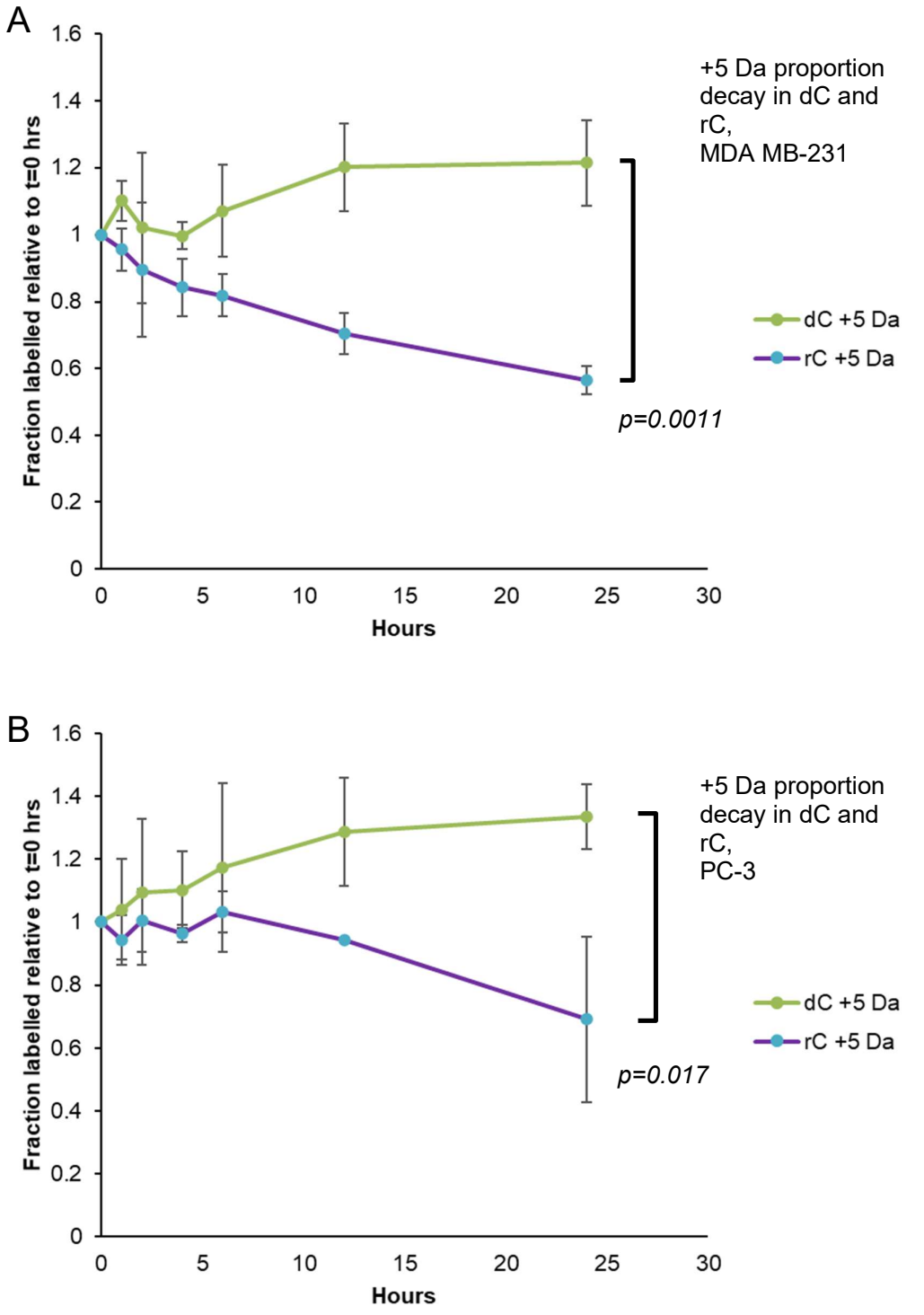


Figure 6.4 DNA labelling decays slower than RNA labelling. *The dC +5 Da (DNA) label proportion decreases more slowly that of rC +5 Da (RNA). Labelled incorporation levels in cells exposed to scrambled siRNA, 3 biological replicates normalised to t=0 hrs per each biological replicate, standard deviation is shown.* ***A,*** *comparison of the decay of % +5 Da in dC and rC in MDA MB-231 cell line;* ***B,*** *comparison of the decay of % +5 Da in dC and rC in PC-3 cell line.*

A similar effect was also seen in the comparison of the labelling decay of the nucleotide monophosphates dCMP and CMP: dCMP derived from labelled rC has a longer half-life than CMP derived from labelled rC in MDA MB-231 (**Figure 6.5A**), with the same trend observed for PC-3 (**Figure 6.5B**). dCMP derived from labelled dC followed a similar trend to that of CMP derived from labelled rC in both cell lines.

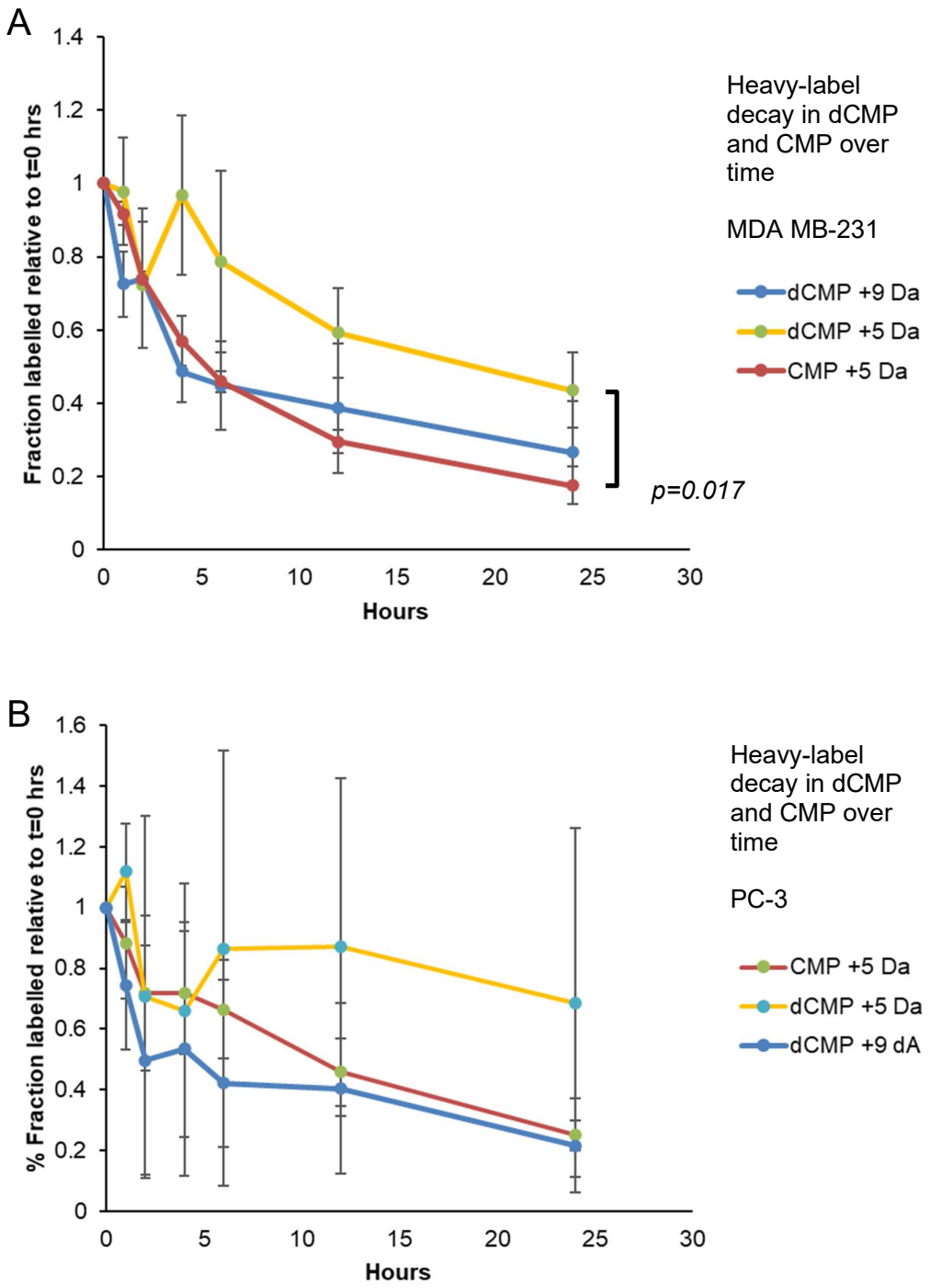
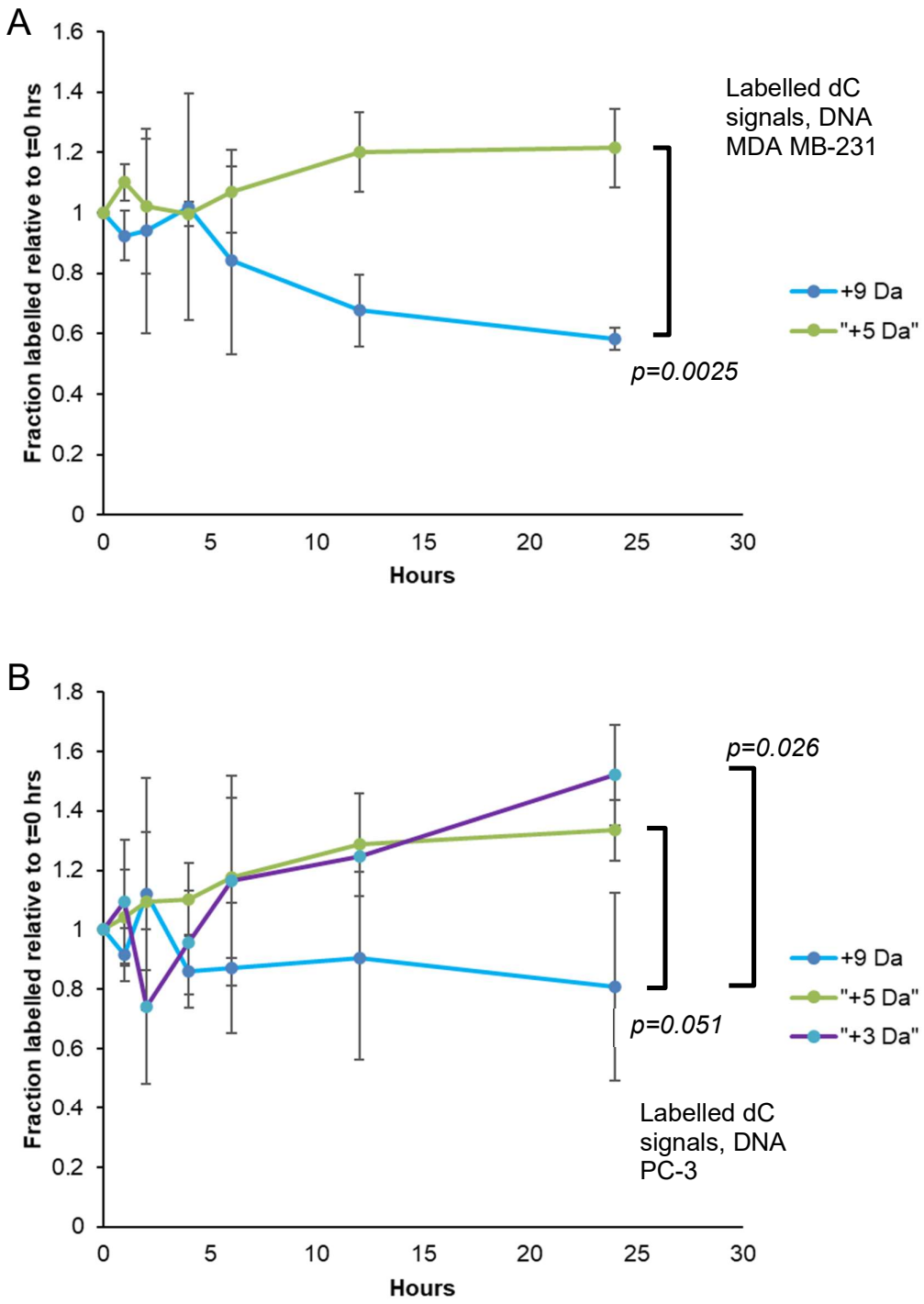
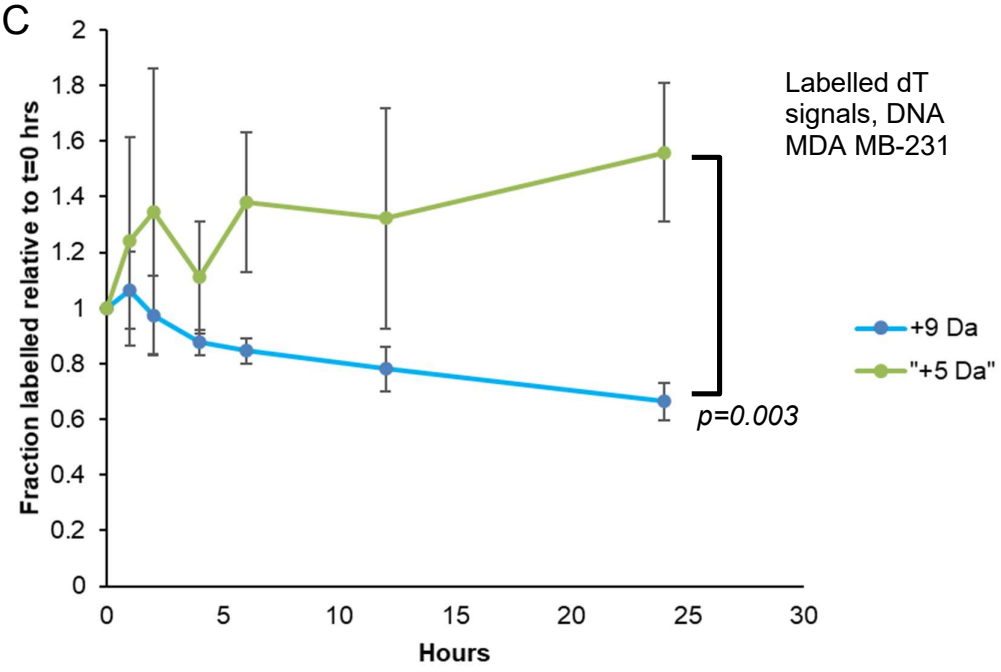


Figure 6.5 Labelled dCMP has a longer decay time than labelled CMP. *Measurement of decay of proportion of dCMP +9 Da, dCMP +5 Da and CMP +5 Da over 24 hours. Experiment was in cells exposed to scrambled siRNA, 3 biological replicates normalised to t=0 hrs per each biological replicate, standard deviation is shown.* ***A,*** *signal decay over time in MDA MB-231 cell line;* ***B,*** *signal decay over time in PC-3 cell line.*

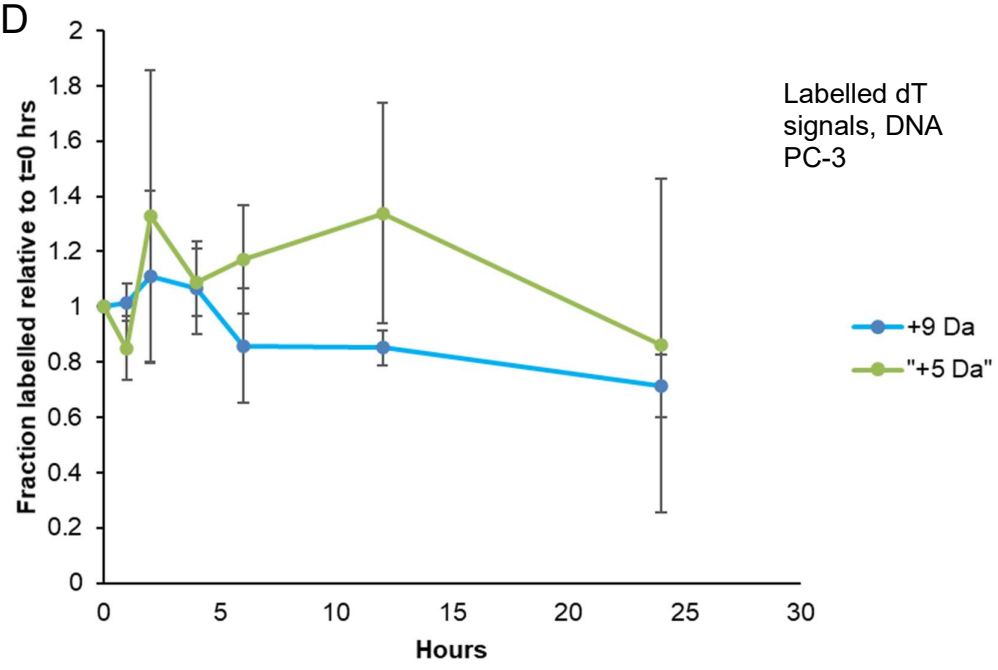
The experiment was over 24 hrs, approximately one doubling time for both MDA MB-231 and PC-3. One would consequently expect there to be a decay in the total labelled DNA proportion through dilution by the synthesis of non-labelled DNA. However, only the +9 Da (dC-derived) showed a trend towards decay in MDA MB-231, whereas the +5 Da (rC-derived) and the +3 Da (L-glutamine-derived) labels showed little change or increased over time (**Figure 6.6A-D**). There was also a trend for the L-glutamine-derived labelling decaying slower than the rC-derived labelling in both nucleic acids and nucleotides (**Figure 6.6 B** and **E-H**).


A
Fraction labelled relative to t=0 hrs
Hours
Labelled dC signals, DNA MDA MB-231
+9 Da
"+5 Da"
p=0.0025
B
Fraction labelled relative to t=0 hrs
Hours
p=0.026
p=0.051
+9 Da
"+5 Da"
"+3 Da"
Labelled dC signals, DNA PC-3

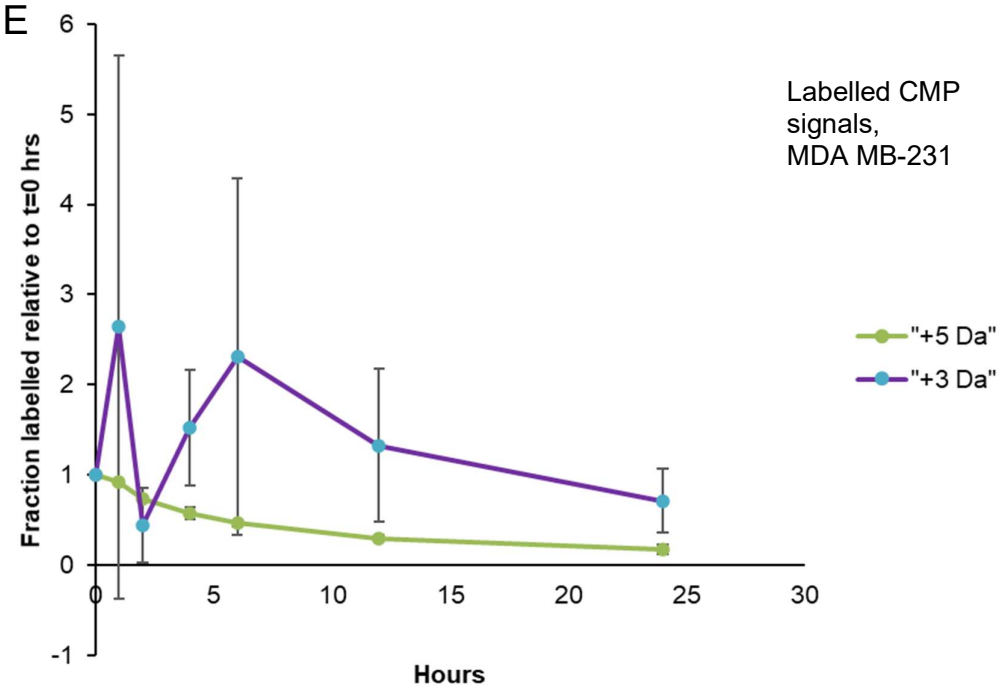
C


Fraction labelled relative to t=0 hrs
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
5
10
15
20
25
30
Hours
Labelled dT signals, DNA MDA MB-231
+9 Da
"+5 Da"
p=0.003

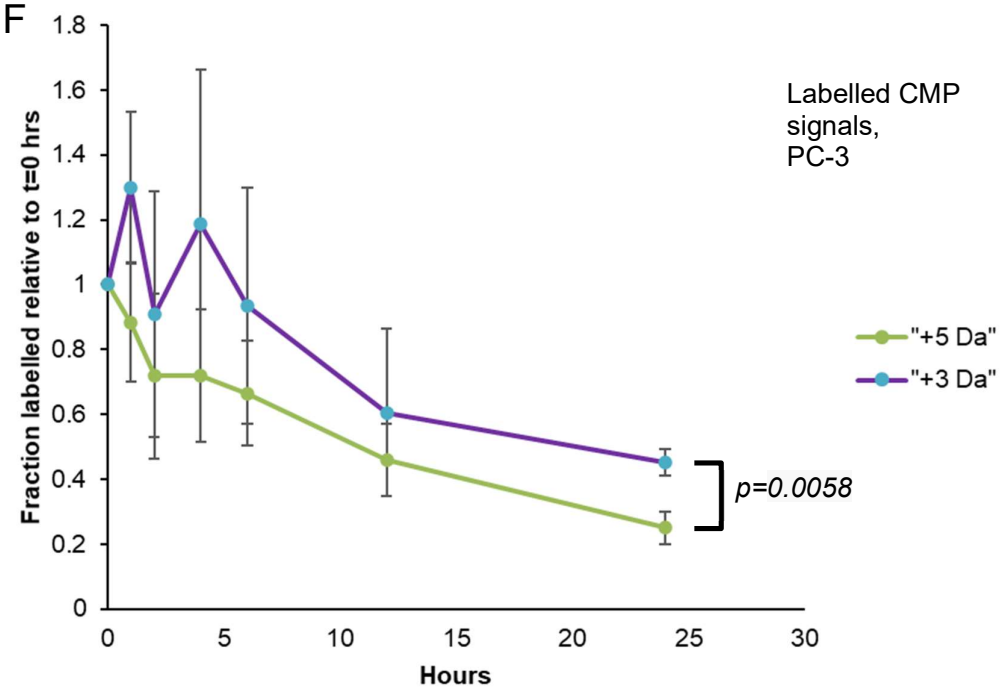
D


Fraction labelled relative to t=0 hrs
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
5
10
15
20
25
30
Hours
Labelled dT signals, DNA PC-3
+9 Da
"+5 Da"

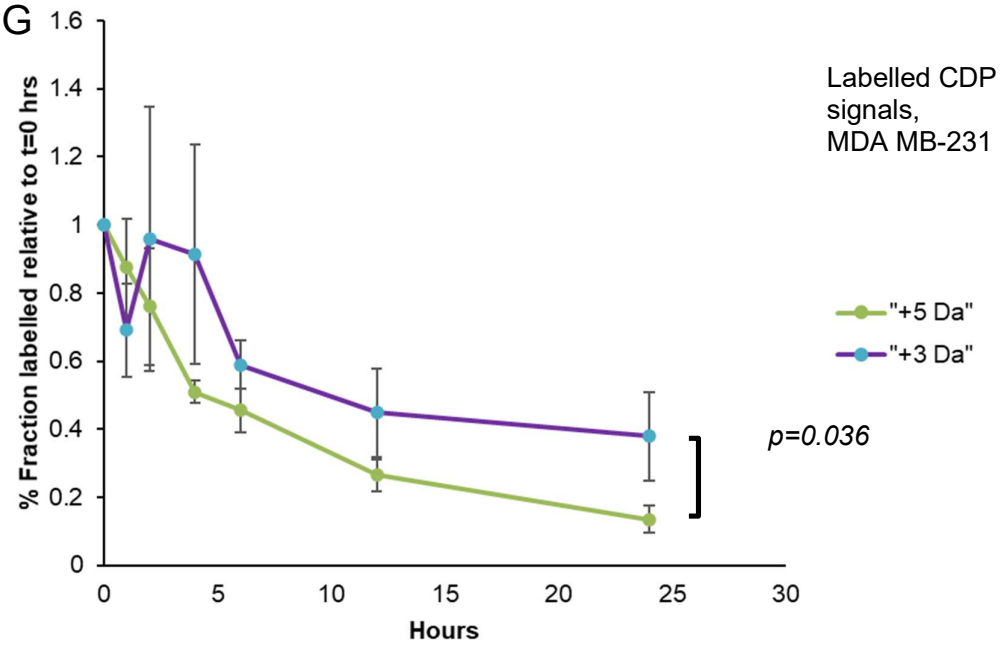
E


Labelled CMP signals, MDA MB-231
"+5 Da"
"+3 Da"
Fraction labelled relative to t=0 hrs
Hours
6
5
4
3
2
1
0
-1
0
5
10
15
20
25
30

F


Labelled CMP signals, PC-3
"+5 Da"
"+3 Da"
p=0.0058
Fraction labelled relative to t=0 hrs
Hours
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
5
10
15
20
25
30

G


% Fraction labelled relative to t=0 hrs
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
5
10
15
20
25
30
Hours
Labelled CDP signals, MDA MB-231
"+5 Da"
"+3 Da"
p=0.036

H

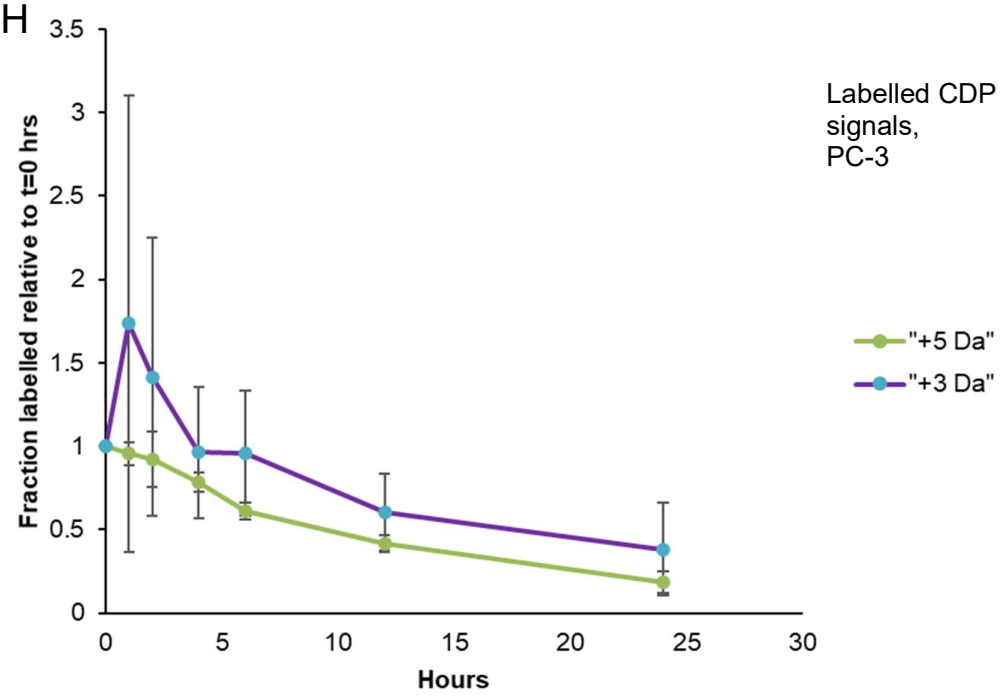

Fraction labelled relative to t=0 hrs
3.5
3
2.5
2
1.5
1
0.5
0
0
5
10
15
20
25
30
Hours
Labelled CDP signals, PC-3
"+5 Da"
"+3 Da"

Figure 6.6 Label proportions from different precursors decay at different rates. *The percentage incorporation of labelled species varies across the 24 hours going down fastest in labelling derived from dC +9 Da and slowest in labelling derived from L-glutamine (+3 Da). Labelled incorporation levels in cells exposed to scrambled siRNA, 3 biological replicates normalised to t=0 hrs per each biological replicate, standard deviation is shown.* ***A,*** *decay of label proportion in dC in MDA MB-231 cell line;* ***B,*** *decay of label proportion in dC in PC-3 cell line;* ***C,*** *decay of label proportion in dT in MDA MB-231 cell line;* ***D,*** *decay of label proportion in dT in PC-3 cell line;* ***E,*** *decay of label proportion in CMP in MDA MB-231 cell line;* ***F,*** *labelled % incorporation in CMP in PC-3 cell line;* ***G,*** *decay of label proportion in CDP in MDA MB-231 cell line;* ***H,*** *decay of label proportion in CDP in PC-3 cell line.*

The observation that DNA labelling decays fastest and that the labelling is source-dependent is not explainable by the presence of larger reservoirs of labelled rC and L-glutamine in cells than labelled dC; otherwise, +5 Da decay would be similar between DNA and RNA (**Figure 6.4**). Furthermore, pure DNA labelling such as BrdU halves in intensity with every cell division.[258] My hypothesis is that labelled RNA is subsequently catabolised in significant amounts and the ribonucleotides are then salvaged, a portion of which are reduced and converted to deoxyribonucleotides. Therefore, even after the media is replaced, there is a reservoir of labelled nucleotides that maintains or increases the labelling proportion in DNA. dC +9 Da is not an upstream metabolite for RNA and so RNA catabolism would not maintain labelled +9 Da availability, in agreement with the steeper decay for +9 Da labelling.

It is unknown what the proportion of nucleotide monomers derived from DNA are salvaged compared to the proportion catabolised to other metabolites. However, most salvaged monomers will have been derived from RNA, having a shorter half-life than DNA. Salvaged nucleotides from rRNA and tRNA would likely have the greatest impact on the flux of nucleotides on the basis of their abundance, with potential contribution from RNAs with high fractional turnover such as mRNA and possibly the non-gene-coding products of pervasive transcription.[259]

Glutamine is also a component of proteins; the low decay rate of +3 Da (from L-glutamine) compared to +5 Da (from rC) may be explained by the degradation of proteins by the lysosome and proteasome into amino acids. Subsequently, previously protein-incorporated labelled L-glutamine could be metabolised by the pyrimidine *de novo* synthesis, sustaining the labelling in nucleotides and nucleic acids after exogenous labelled nucleotide precursors are removed.

6.5 Roles of CDA and DCTD in pyrimidine deamination

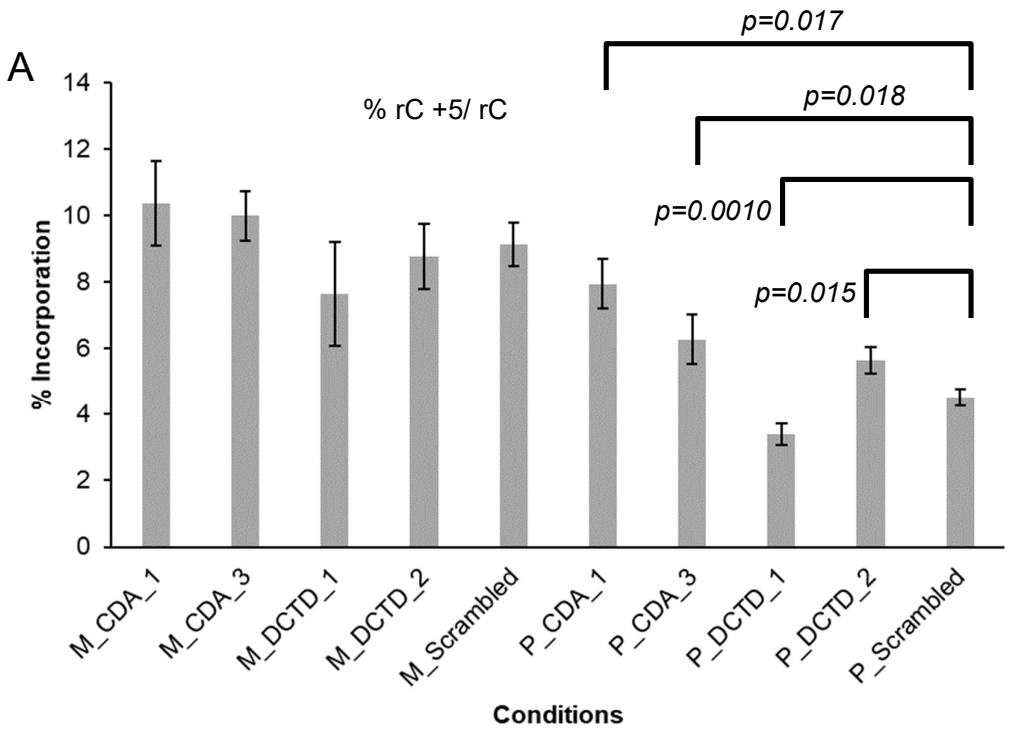
I examined how KD of CDA and DCTD affected label incorporation in the nucleotide pool and in nucleic acids. My prediction was that if one of the deaminases were the major deaminase, the knockdown of the other would not affect incorporation of labelled precursors in cytidine and uridine nucleotides, while KD of the other enzyme would increase cytidine labelling and decrease uridine labelling from salvage precursors.

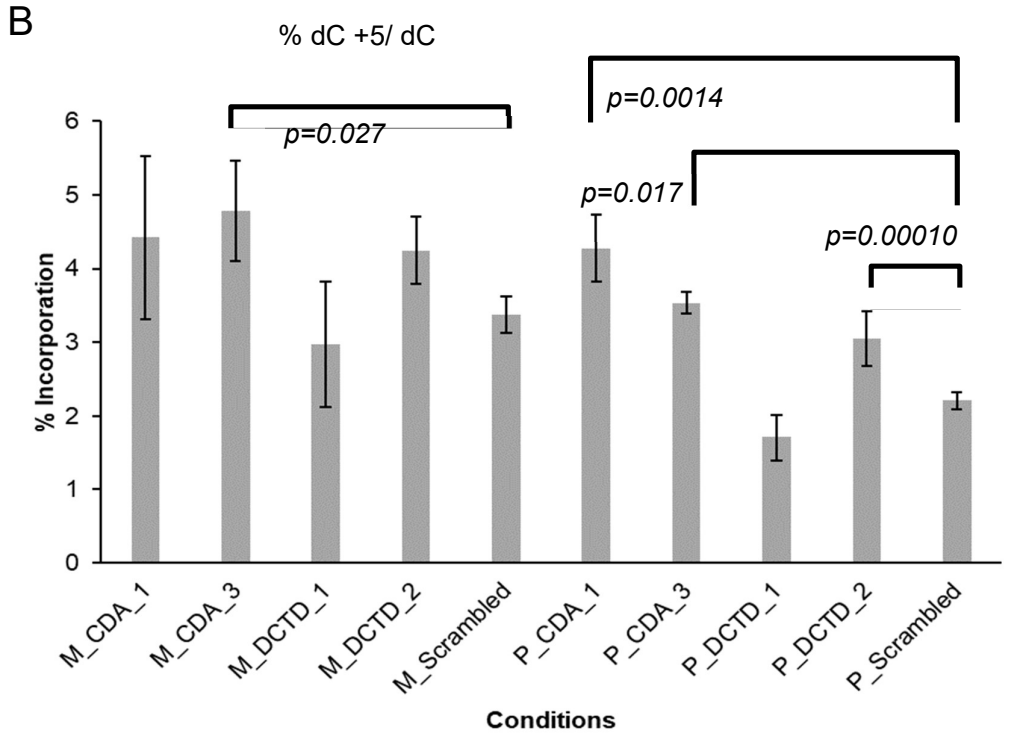
6.5.1 CDA KD reduces rC deamination more than DCTD KD in cell lines

At t=0 hrs, CDA KD induced an increase in rC +5 Da-derived labelling in cytidine nucleotides compared to scrambled siRNA, more so than DCTD KD (**Figure 6.7A-E**); both cell lines followed this trend, although there were more significant results in PC-3. In parallel, CDA KD led to a decrease in uridine nucleotide +5 Da labelling (**Figure 6.7F-H**). This difference, however, is not seen for +9 Da labelling in dC (**Figure 6.7I**). It suggests that CDA and DCTD have mutually redundant roles for dC deamination, but that DCTD does not deaminate rC nucleotides; this trend validates what was previously reported.[260]

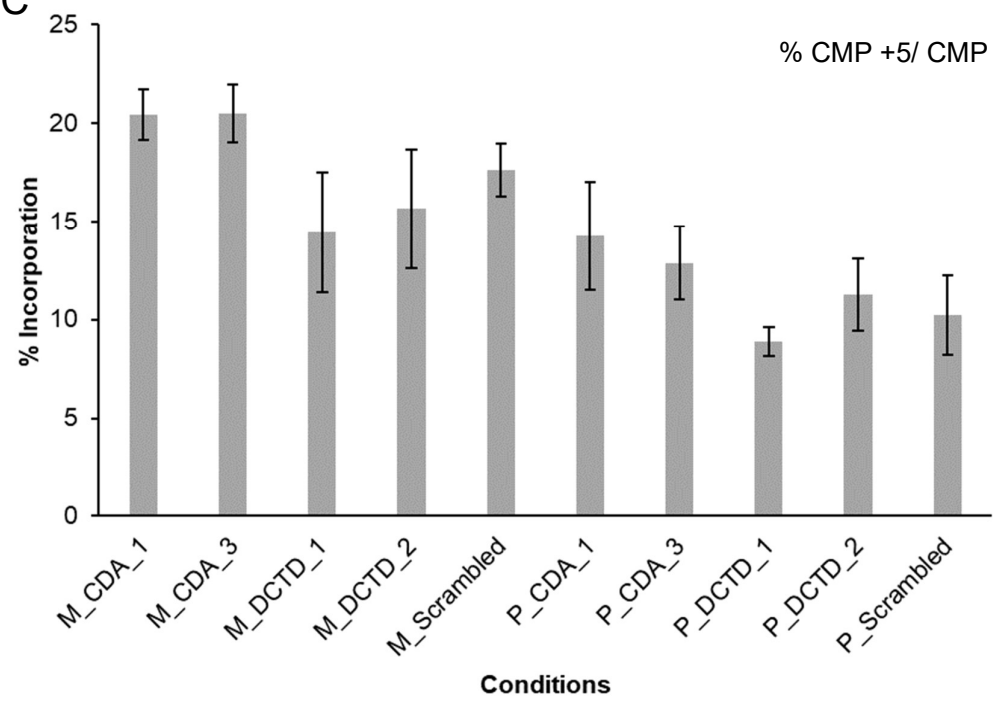
DCTD KD in MDA MB-231 had lower CTP +5 Da labelling than scrambled siRNA, though this wasn't observed in PC-3 (**Figure 6.7E**), with a possible similar trend seen in CMP and CDP (**Figure 6.7Figure 6.7C** and **D)**. This could be indicative of pathway compensation in MDA MB-231 as a reaction to DCTD KD, through increase in activity or protein level of CDA or a corresponding decrease in CTPS1 and CTPS2 activities or protein levels. This could be verified by immunoblot of these enzymes upon DCTD KD.

It was also observed that the samples treated with the siRNA DCTD1 had lower incorporation from the rC +5 Da precursor than when treated with the siRNA DCTD2 (**Figure 6.7A** and **B**). This may be due to an off-target effect of either of the siRNAs and could be clarified through validation with a third siRNA. Similarly, **Figure 6.7Figure 6.7E** shows a difference between the CDA siRNAs dC +9 Da labelling in both MDA MB-231 and PC-3, that could be investigated by using a third siRNA.

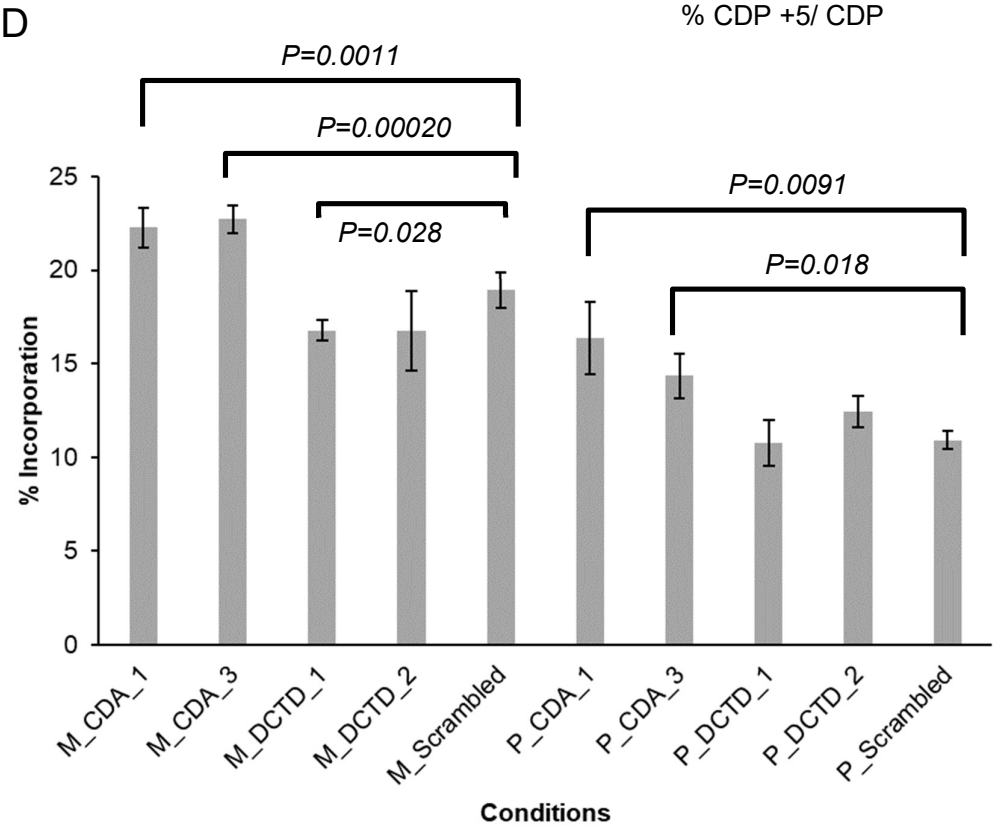

A
% rC +5/ rC
p=0.017
p=0.018
p=0.0010
p=0.015
% Incorporation
0
2
4
6
8
10
12
14
M_CDA_1
M_CDA_3
M_DCTD_1
M_DCTD_2
M_Scrambled
P_CDA_1
P_CDA_3
P_DCTD_1
P_DCTD_2
P_Scrambled
Conditions


B
% dC +5/ dC
p=0.027
p=0.0014
p=0.017
p=0.00010
% Incorporation
0
1
2
3
4
5
6
M_CDA_1
M_CDA_3
M_DCTD_1
M_DCTD_2
M_Scrambled
P_CDA_1
P_CDA_3
P_DCTD_1
P_DCTD_2
P_Scrambled
Conditions

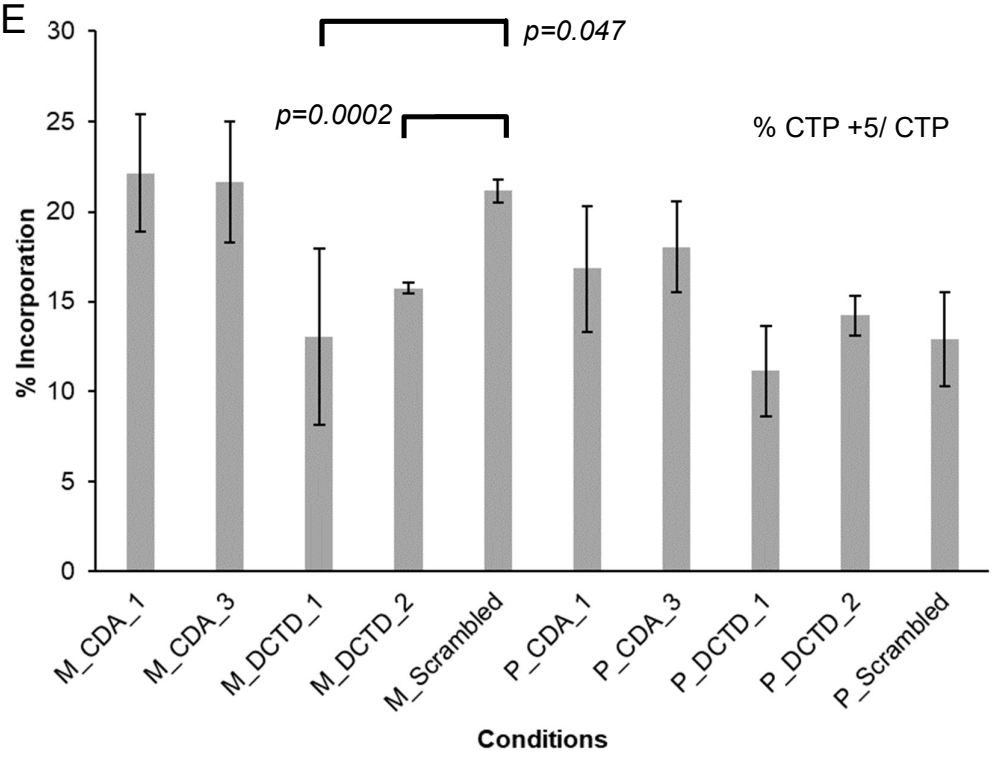
C


% CMP +5/ CMP
% Incorporation
25
20
15
10
5
0
M_CDA_1
M_CDA_3
M_DCTD_1
M_DCTD_2
M_Scrambled
P_CDA_1
P_CDA_3
P_DCTD_1
P_DCTD_2
P_Scrambled
Conditions

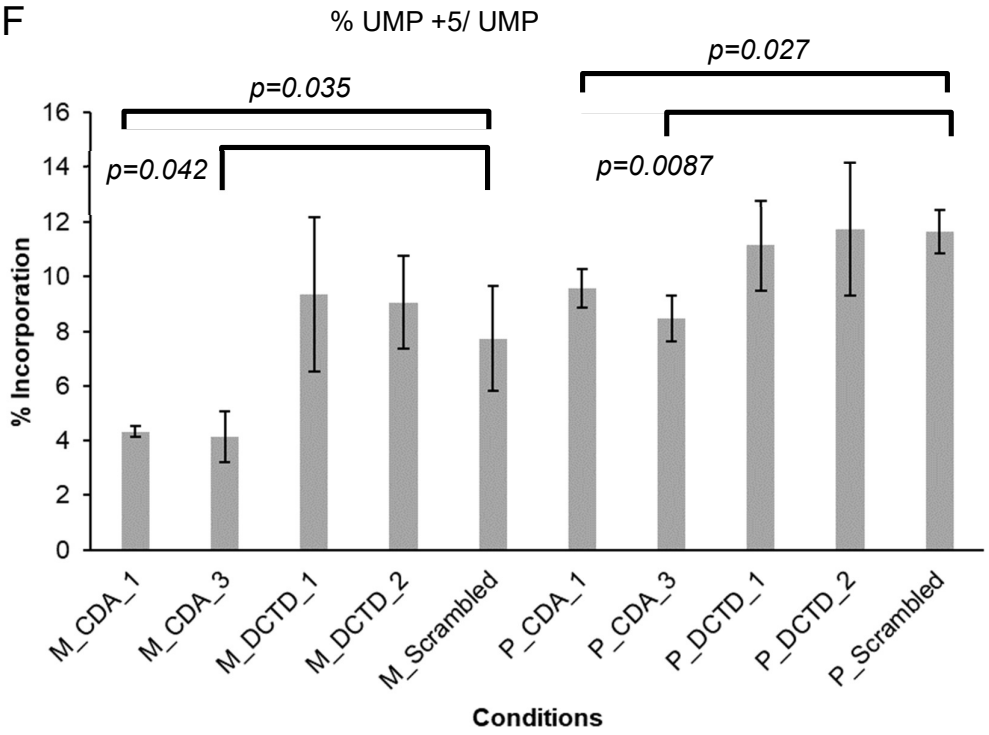
D


% CDP +5/ CDP
P=0.0011
P=0.00020
P=0.028
P=0.0091
P=0.018
% Incorporation
25
20
15
10
5
0
M_CDA_1
M_CDA_3
M_DCTD_1
M_DCTD_2
M_Scrambled
P_CDA_1
P_CDA_3
P_DCTD_1
P_DCTD_2
P_Scrambled
Conditions

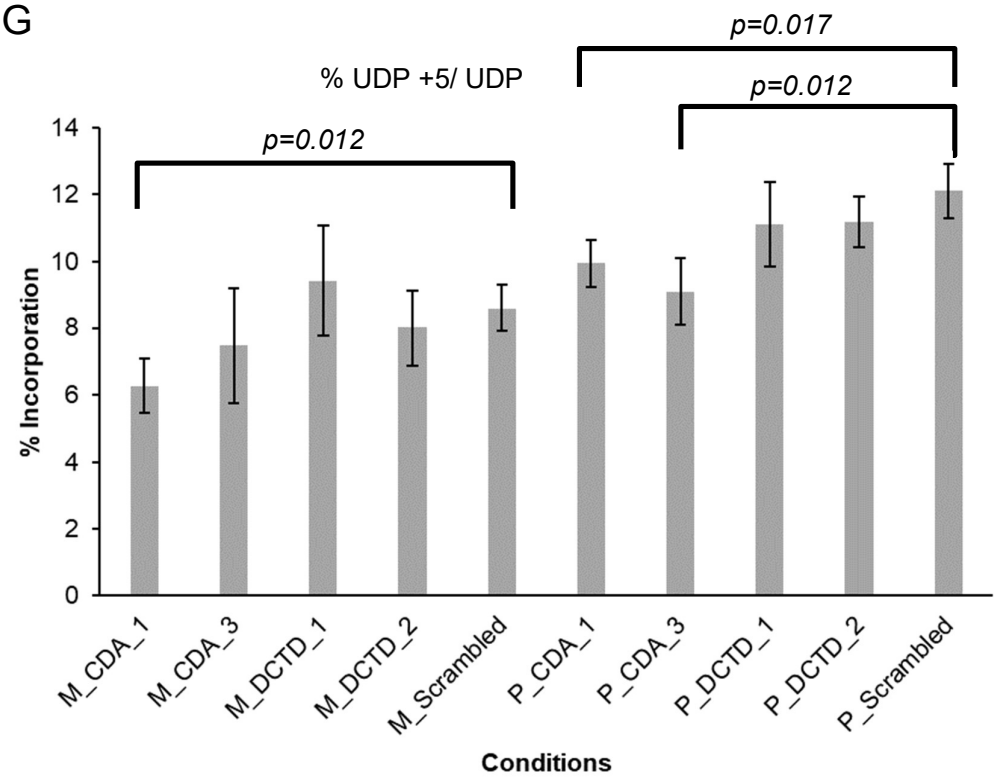
E


p=0.047
p=0.0002
% CTP +5/ CTP
% Incorporation
30
25
20
15
10
5
0
M_CDA_1
M_CDA_3
M_DCTD_1
M_DCTD_2
M_Scrambled
P_CDA_1
P_CDA_3
P_DCTD_1
P_DCTD_2
P_Scrambled
Conditions

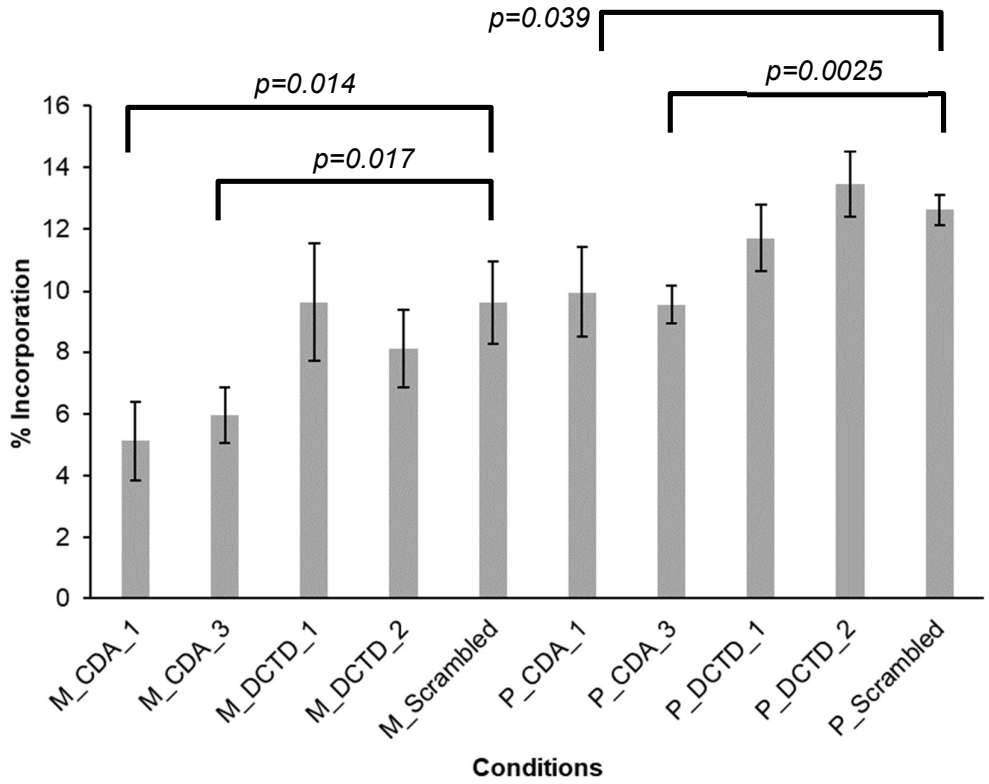
F


% UMP +5/ UMP
p=0.035
p=0.042
p=0.027
p=0.0087
% Incorporation
16
14
12
10
8
6
4
2
0
M_CDA_1
M_CDA_3
M_DCTD_1
M_DCTD_2
M_Scrambled
P_CDA_1
P_CDA_3
P_DCTD_1
P_DCTD_2
P_Scrambled
Conditions

G


% UDP +5/ UDP
p=0.017
p=0.012
p=0.012
% Incorporation
0
2
4
6
8
10
12
14
M_CDA_1
M_CDA_3
M_DCTD_1
M_DCTD_2
M_Scrambled
P_CDA_1
P_CDA_3
P_DCTD_1
P_DCTD_2
P_Scrambled
Conditions

H


% UTP +5/ UTP
p=0.039
p=0.0025
p=0.014
p=0.017
% Incorporation
0
2
4
6
8
10
12
14
16
M_CDA_1
M_CDA_3
M_DCTD_1
M_DCTD_2
M_Scrambled
P_CDA_1
P_CDA_3
P_DCTD_1
P_DCTD_2
P_Scrambled
Conditions

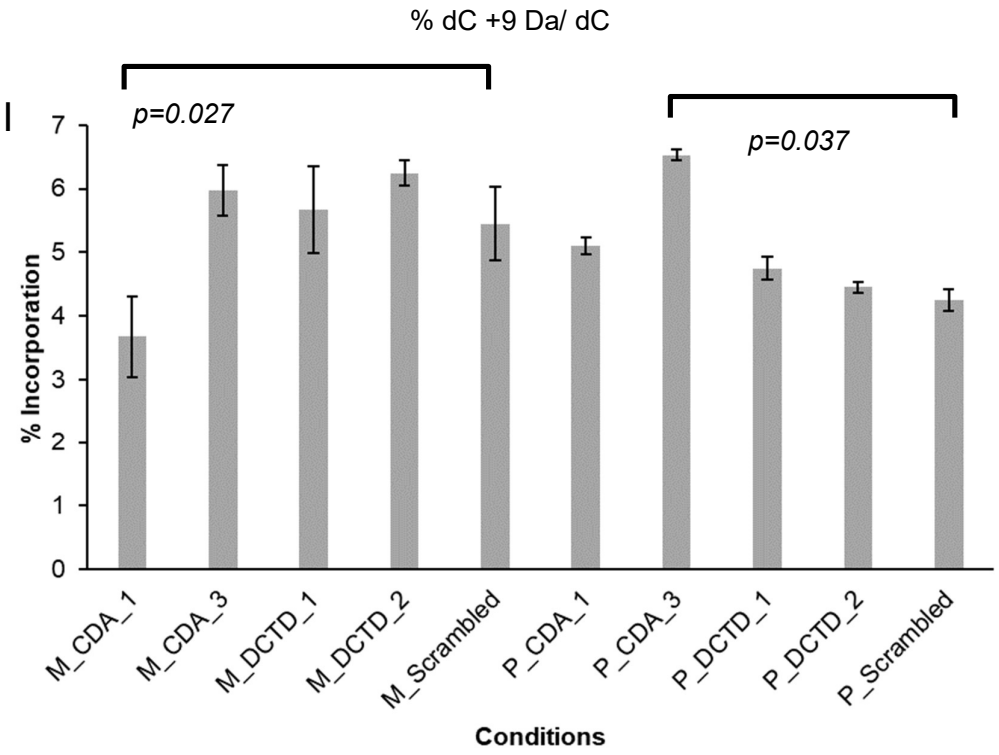
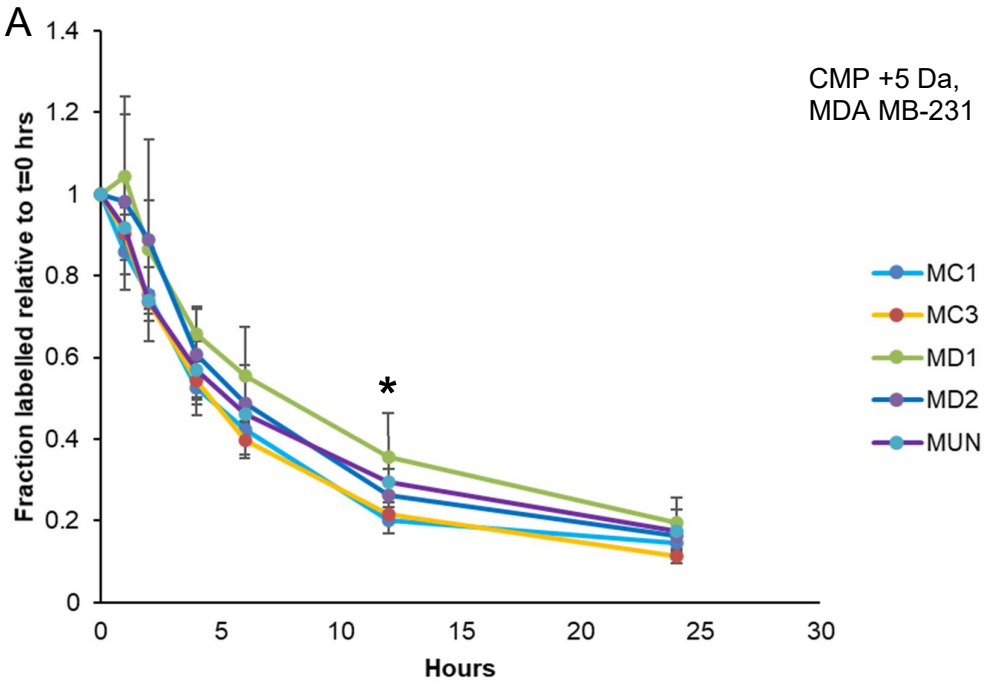


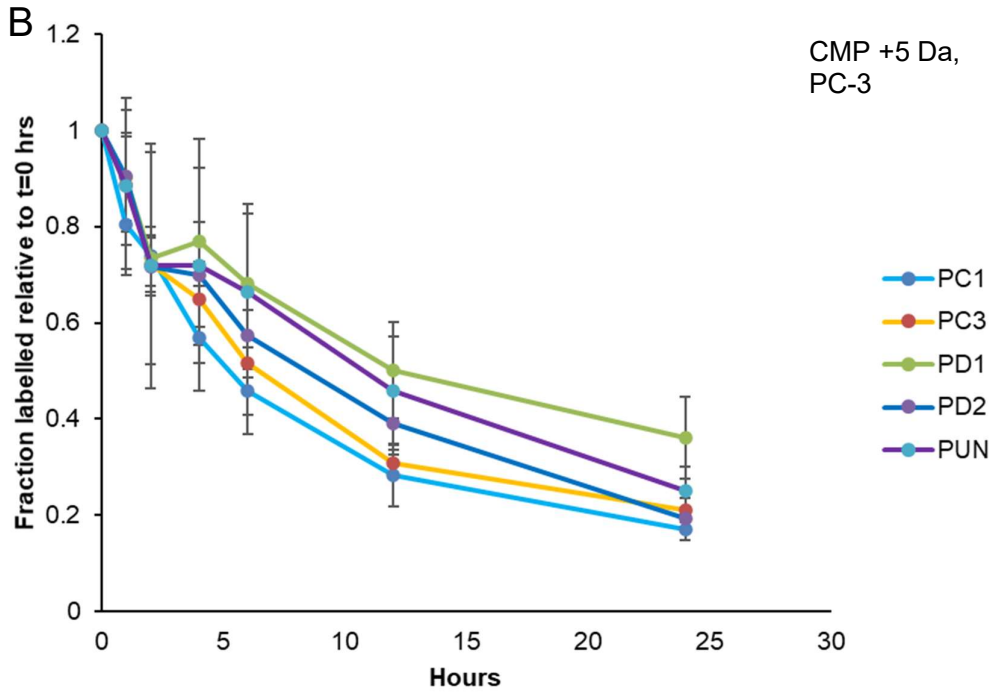
Figure 6.7 KD of CDA affects rC +5 incorporation, while DCTD KD does not. *CDA KD increases incorporation in of metabolites of rC +5 in cytidine variants, but the same effect is not observed for DCTD KD. Data from cell lines MDA MB-231 (M in figure) and PC-3 (P in figure) at t= 0 hrs post 24 hrs heavy label pulse. 2 orthogonal siRNAs were used per CDA and DCTD KD and compared to a scrambled siRNA with 3 biological replicates, standard deviation is shown. A, incorporation of +5 Da label in rC; B, incorporation of +5 Da label in dC; C, incorporation of +5 Da label in CMP; D, incorporation of +5 Da label in CDP; E, incorporation of +5 Da label in CTP; F, incorporation of +5 Da label in UMP; G, incorporation of +5 Da label in UDP; H, incorporation of +5 Da label in UTP; I, incorporation of dC +9 as dC.*

6.5.2 CDA KD induces faster rC nucleotide signal decrease than DCTD

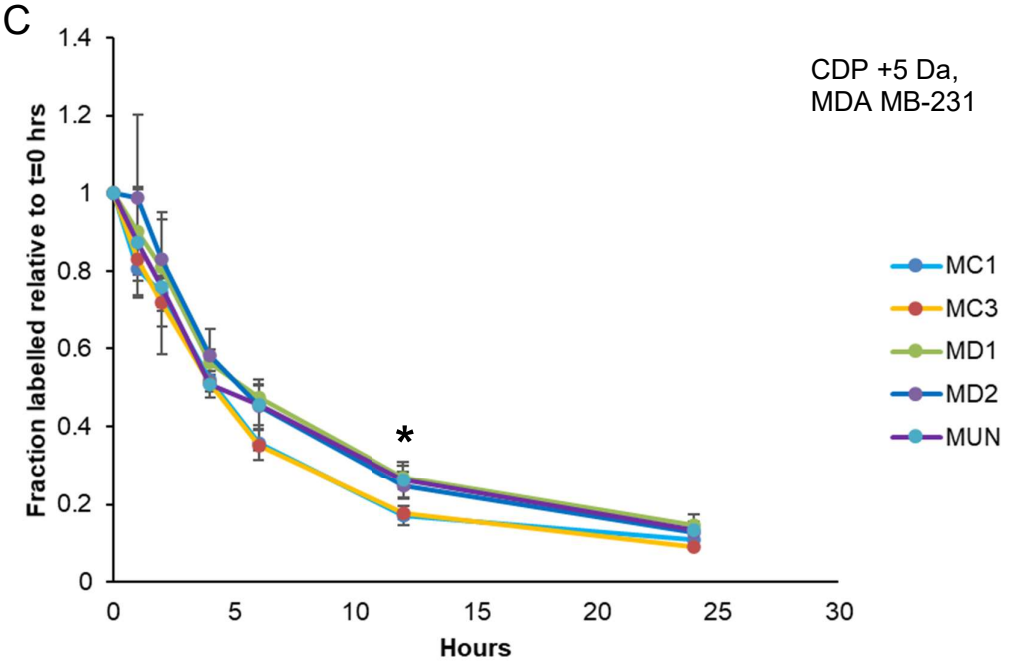
Low activity of DCTD against CMP was further observed when investigating relative signal decay of +5 Da rC nucleotides over time. CDA KD steepened the decay curve of +5 Da rC nucleotides over time compared to scrambled siRNA (**Figure 6.8A-D,** p-values given in legend) in both cell lines; DCTD KD did not alter the curve shape, although there was one statistically

significant point at DCTD siRNA2 at t=12hrs for CDP, where it was less steep than with scrambled siRNA. The decay for CTP had too high variance for equivalent analysis (**Figure 6.8E** and **F)**, apart from for the CDA siRNA3 vs scrambled in PC-3 (**Figure 6.8F).** A proposed explanation for the steeper decay for CDA KD is that rC nucleotide phosphorylation activity is increased, with a corresponding increase in turnover. The initial +5 Da labelling of CDP for CDA KD was greater at t=0 hrs (**Figure 6.7C**) than for the other siRNAs. This may imply that elevated cytidine nucleotides may trigger nucleotide stoichiometry balancing mechanisms, in turn increasing turnover.

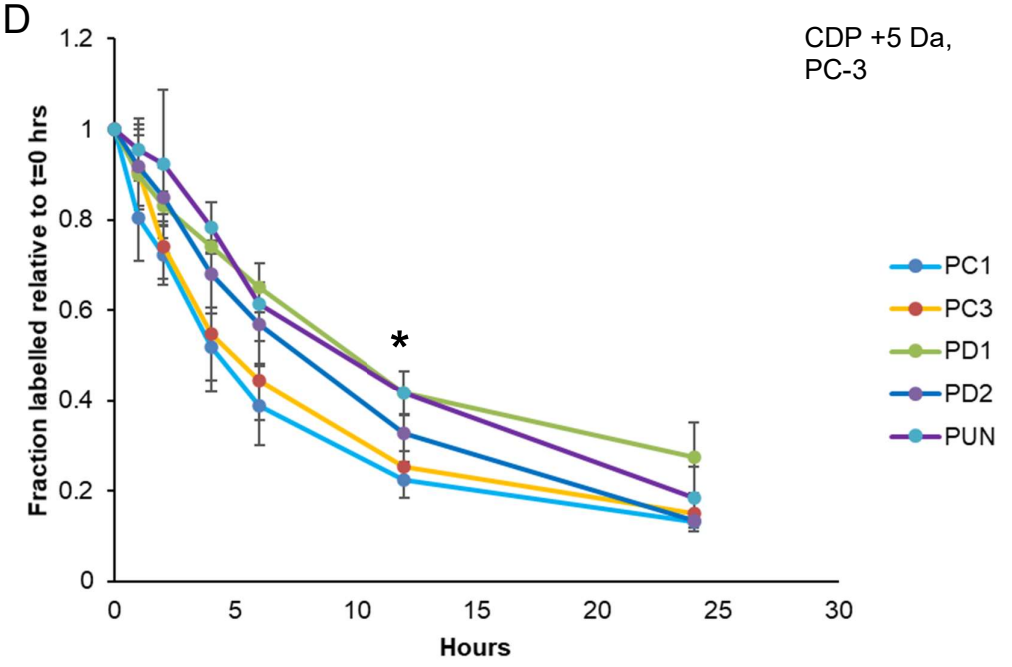

A
CMP +5 Da,
MDA MB-231
Fraction labelled relative to t=0 hrs
Hours
*
MC1
MC3
MD1
MD2
MUN


B
CMP +5 Da,
PC-3
Fraction labelled relative to t=0 hrs
Hours
PC1
PC3
PD1
PD2
PUN

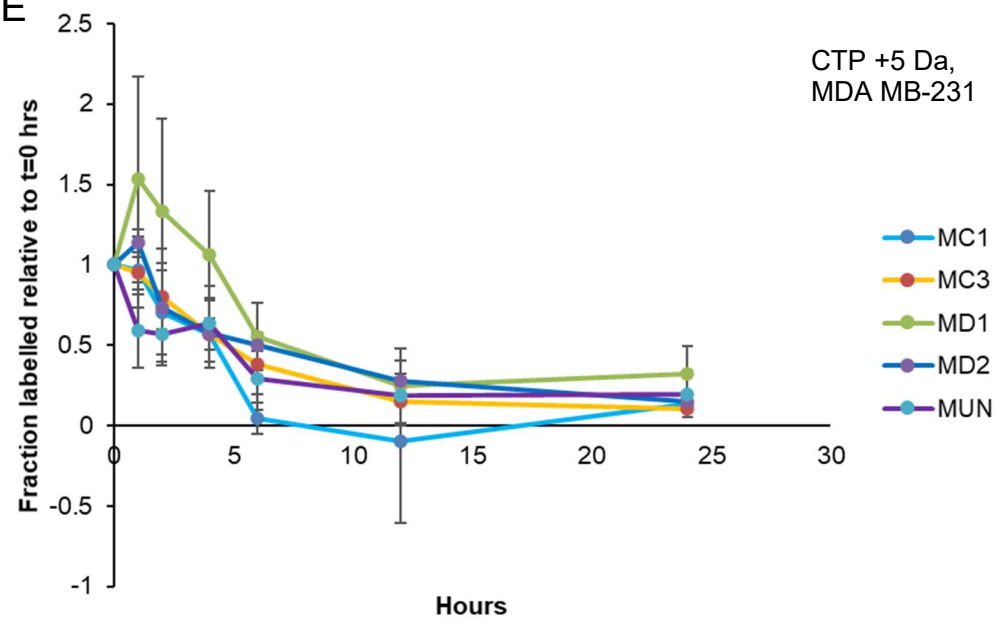
C


CDP +5 Da,
MDA MB-231
Fraction labelled relative to t=0 hrs
Hours
MC1
MC3
MD1
MD2
MUN
*

D


CDP +5 Da,
PC-3
Fraction labelled relative to t=0 hrs
Hours
PC1
PC3
PD1
PD2
PUN
*

E


CTP +5 Da,
MDA MB-231
Fraction labelled relative to t=0 hrs
Hours
MC1
MC3
MD1
MD2
MUN

F

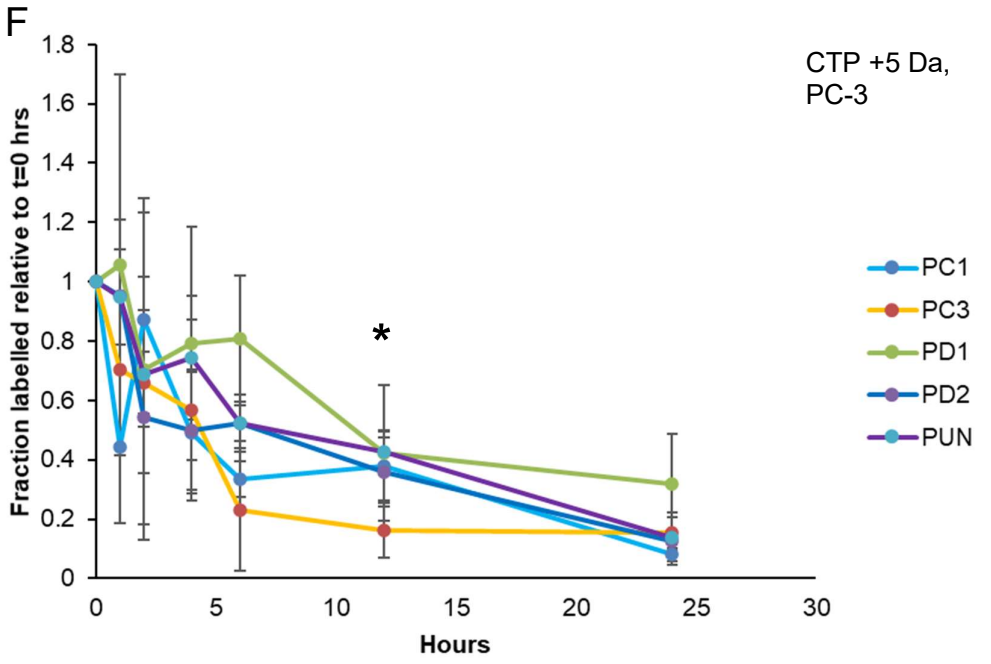

CTP +5 Da,
PC-3
Fraction labelled relative to t=0 hrs
Hours
*
PC1
PC3
PD1
PD2
PUN

Figure 6.8 CDA KD correlates with greater reduction of +5 Da signal over time across cytidine metabolites, with effect not seen for +9 Da. *2 orthogonal siRNAs were used per CDA and DCTD KD and compared to a scrambled siRNA with 3 biological replicates normalised to t=0 hrs per each biological replicate, standard deviation is shown and significance is judged at t=12 hrs.* ***A,*** *+5 Da labelling of CMP in MDA MB-231 cell line (CDA1 relative to scrambled, p=0.024, CDA 3 relative to scrambled, p=0.0143);* ***B,*** *+5 Da labelling of CMP in PC-3 cell line;* ***C,*** *+5 Da labelling of CDP in MDA MB-231 cell line (CDA1 relative to scrambled, p=0.037, CDA 3 relative to scrambled, p=0.039);* ***D,*** *+5 Da labelling of CDP in PC-3 cell line (CDA1 relative to scrambled, p=0.0057, CDA 3 relative to scrambled, p=0.0083, DCTD 2 relative to scrambled, p=0.042);* ***E,*** *+5 Da labelling of CTP in MDA MB-231 cell line;* ***F,*** *+5 Da labelling of CTP in PC-3 cell line (CDA 3 relative to scrambled, p=0.019).*

6.6 CDA KD may create a bottleneck for rC deamination

In the PC-3 cell line, CDA KD leads to an increase in dC +9 Da incorporation as dT (**Figure 6.9A**). This is not seen in MDA MB-231, nor does it have a corresponding change in dC +9 Da proportion (**Figure 6.7E**). This seemingly counterintuitive result could be because the medium of PC-3 has a high dT concentration, which competes with nucleotides derived from dC and rC for incorporation as dT. CDA KD affects both dC +9 Da and rC +5 Da, but DCTD does not efficiently deaminate CMP, meaning that upon CDA KD, there is a smaller dTTP +5 Da pool in the system compared to the pool size of dTTP +9 Da. There is therefore less competition for dTTP +9 Da to be incorporated as dT in DNA (**Figure 6.10**).

There is still dT +5 Da labelling upon CDA KD in PC-3 (**Figure 6.9B**). This could imply that DCTD deaminates some CMP, or that there are further paths for rC +5 Da to be incorporated as dT +5 Da, such as metabolisation to dCDP +5 Da and subsequent dephosphorylation to dCMP +5 Da from which it can be deaminated to dUMP +5 Da and methylated by TYMS to dTMP +5 Da. DCTD KD by DCTD siRNA2 gave higher +5 rC-derived dT in both cell lines, but it is uncertain whether this is due to off-target siRNA effects and the effect would require further investigation with orthogonal siRNAs.

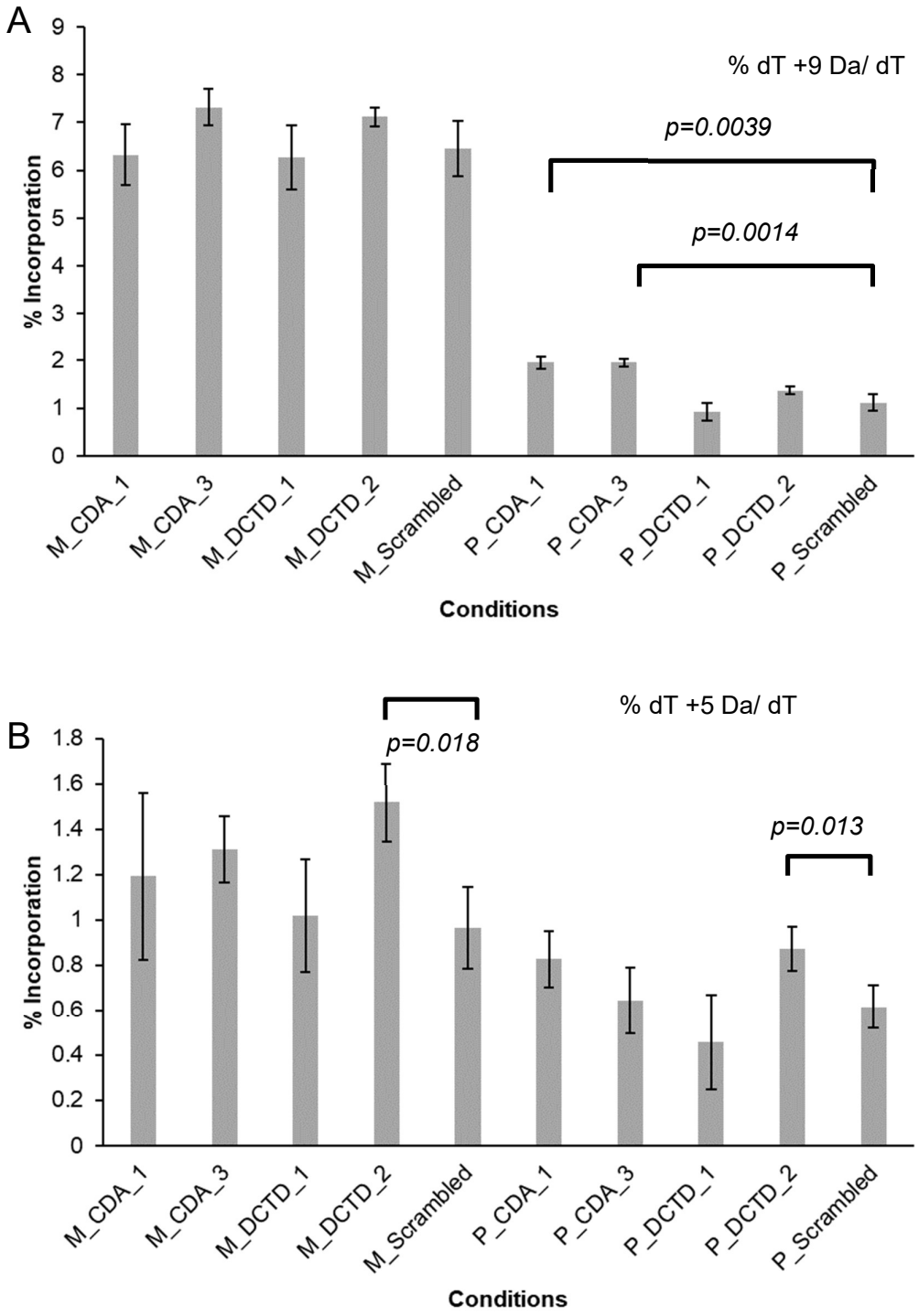


Figure 6.9 CDA KD increases deaminated dC incorporation as dT in PC-3 cell line, but not MDA MB-231 cell line, at t=0 hrs. *Corresponding decrease in dC +9 Da incorporation not seen. 2 orthogonal siRNAs were used per CDA and DCTD KD and compared to a scrambled siRNA with 3 biological replicates, standard deviation is shown.* ***A,*** *incorporation of dC +9 as dT;* ***B,*** *incorporation of rC +5 as dT.*

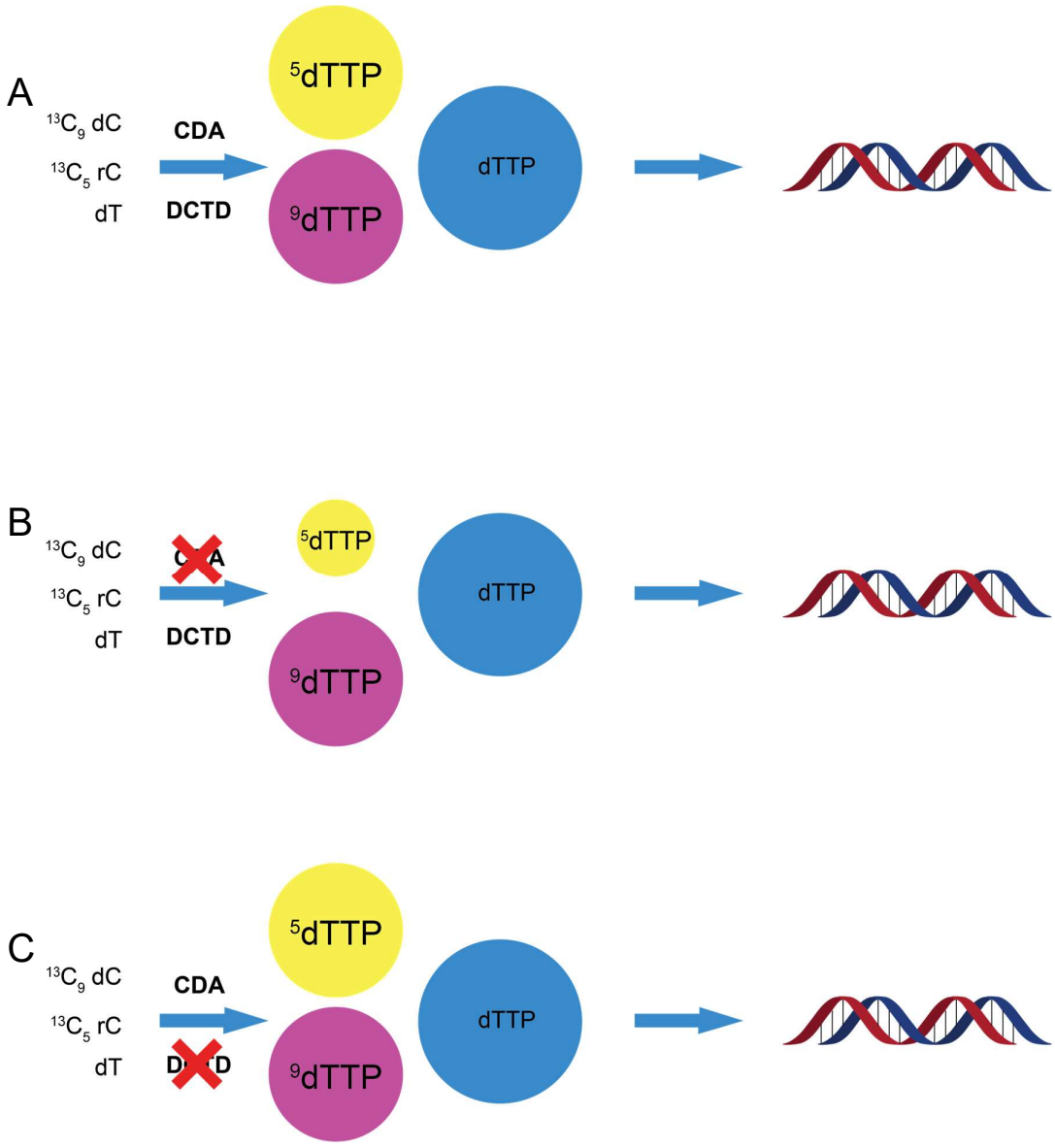
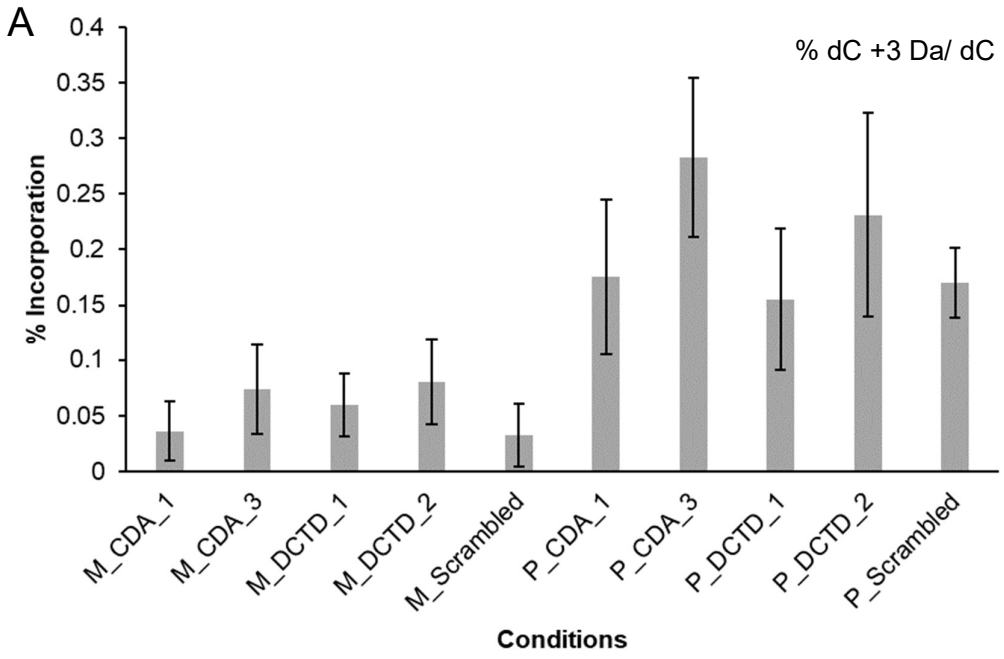


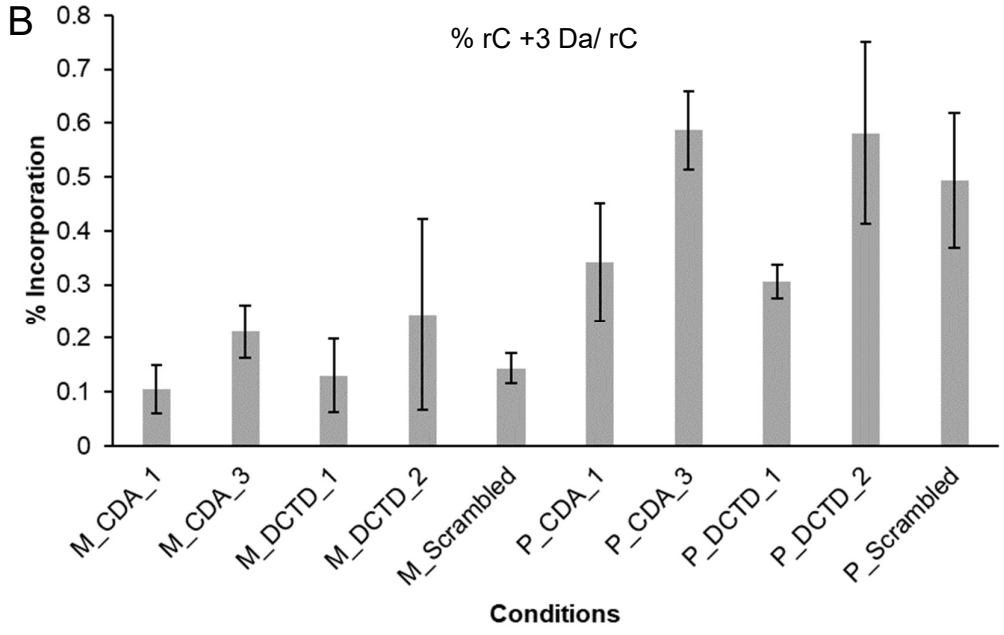
Figure 6.10 Schematic of competition between unlabelled, +5 Da and +9 Da dTTP nucleotide pools. *With dT supplementation in media, CDA KD reduces the nucleotide pool size of dTTP derived from rC +5 Da, but not dTTP derived from dC +9 Da. DCTD KD reduces neither.* ***A,*** *schematic of dTTP provenance in WT conditions;* ***B,*** *schematic of dTTP provenance upon CDA KD;* ***C,*** *schematic of dTTP provenance upon DCTD KD.*

6.7 L-glutamine label incorporation not affected by CDA nor DCTD KD

Finally, I investigated whether CDA or DCTD KD affected the incorporation of synthesis-derived nucleotides. If nucleotides derived from labelled L-glutamine were dephosphorylated and were metabolised by CDA or DCTD, KD of the enzymes would increase the proportion of labelled cytidine nucleotides and would decrease the proportion of labelled uridine/thymidine nucleotides.

The incorporation of the +3 Da label, derived from L-glutamine, was not affected by KD of CDA or of DCTD (**Figure 6.11**).


A
% dC +3 Da/ dC
% Incorporation
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
M_CDA_1
M_CDA_3
M_DCTD_1
M_DCTD_2
M_Scrambled
P_CDA_1
P_CDA_3
P_DCTD_1
P_DCTD_2
P_Scrambled
Conditions


B
% rC +3 Da/ rC
% Incorporation
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
M_CDA_1
M_CDA_3
M_DCTD_1
M_DCTD_2
M_Scrambled
P_CDA_1
P_CDA_3
P_DCTD_1
P_DCTD_2
P_Scrambled
Conditions

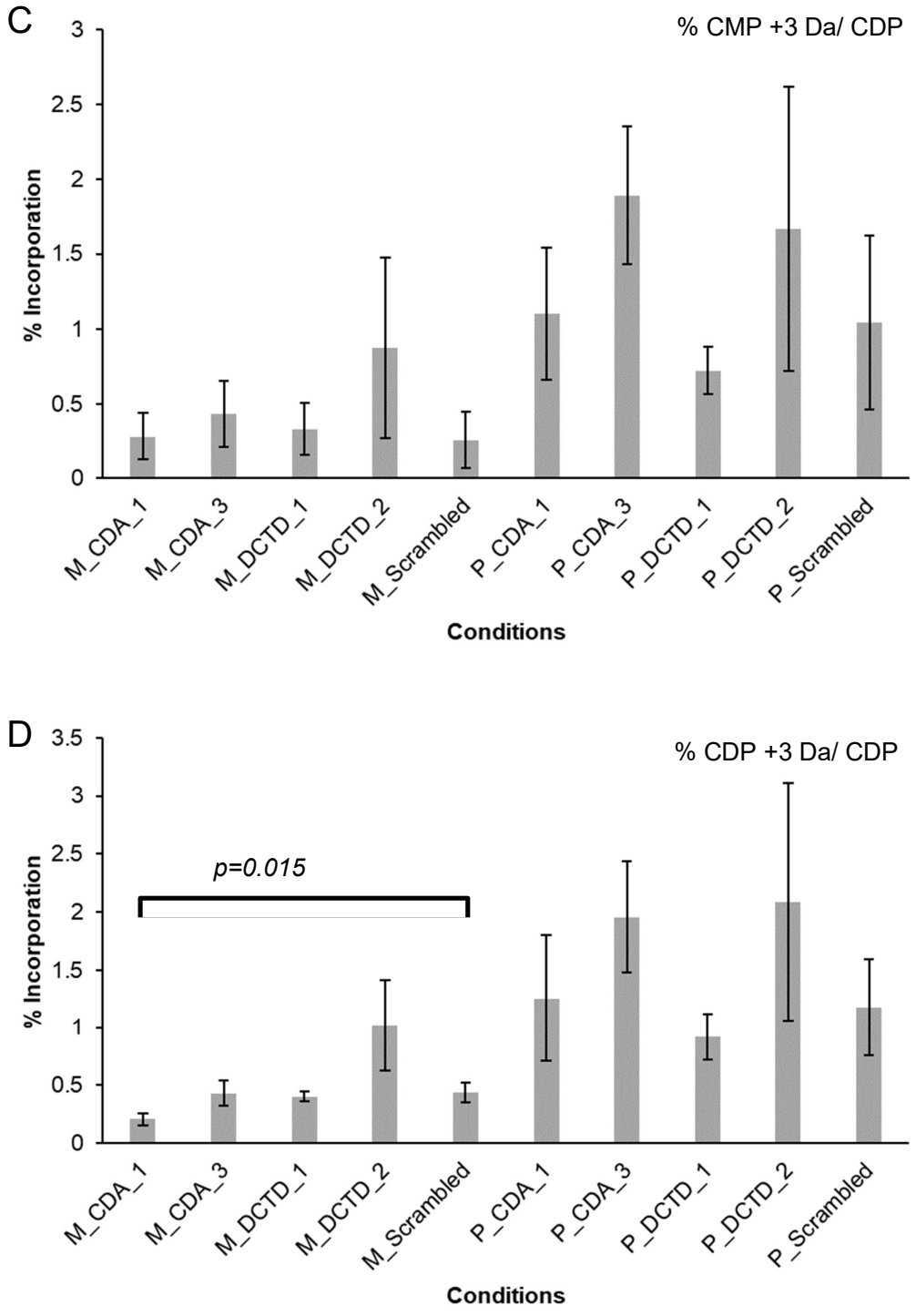


Figure 6.11 Labelled L-glutamine incorporation does not change upon CDA or DCTD KD, at t=0 hrs. *2 orthogonal siRNAs were used per CDA and DCTD KD and compared to a scrambled siRNA with 3 biological replicates, standard deviation is shown.* ***A,*** *incorporation of dC +3 Da;* ***B,*** *incorporation of dC +3 Da;* ***C,*** *incorporation of CMP +3 Da;* ***D,*** *incorporation of CDP +3 Da.*

There was no significant difference seen through KDs, apart from CDA KD through CDA siRNA1 in CDP +3 Da labelling in MDA MB-231 (**Figure 6.11D**). There was a trend of an increase observed for CDA siRNA3 and DCTD siRNA2 (**Figure 6.11A-D**). It is possible that their impact on L-glutamine incorporation, or the absence of impact from the other siRNAs, is due to siRNA off-target effects. This can be verified by testing other orthogonal siRNAs in a simple steady-state labelling experiment (pulse-chase analysis but with no time chase). Another confounding factor is that the +3 Da signal strength is low, possibly too low or varying over 3 biological replicates to detect an effect.

If the results observed stand as correct, synthesised cytidine nucleotides are not dephosphorylated in sufficient quantities in cells for CDA or DCTD to have an impact on +3 Da, suggesting that there is limited recycling of nucleotides from the pyrimidine synthesis pathway in the salvage pathways.

6.8 Discussion

The experiment combined CDA and DCTD knockdown of two cell lines with nucleic acid composition and nucleotide pool MS pulse-chase analyses from the same samples. This gave the opportunity to study the impact of CDA and DCTD on nucleotides from the *de novo* synthesis pathway (L-glutamine-derived) and the salvage pathway (dC- and rC-derived) simultaneously over time.

6.8.1 General observations of the data

In both cell lines, the DNA components dC and dT had greater +9 Da labelling than +5 Da (**Table 6.1**), indicating that dC-nucleotides are more readily incorporated into DNA than rC-

derived nucleotides. This could be due to rC-derived nucleotides being used for RNA synthesis or because the enzymatic pathway to incorporation in DNA is simpler for dC, with no requirement for reduction by RNR.

As the +5 Da labelling did not vary between phosphorylation states over time in cytidine and uridine nucleotides (**Figure 6.3**), the free nucleotide pool turnover may be faster than hourly intervals could measure. RNA labelling proportion declined at a gentler rate, while DNA labelling decayed at an even shallower rate or increased (**Figure 6.4** and **Figure 6.6**). Together with a similar observation that the CMP labelling proportion decays faster than for dCMP (**Figure 6.5**), it supports the hypothesis that nucleotides derived from RNA catabolism form a major part of the deoxyribonucleotides pool and DNA composition.

6.8.2 CDA and DCTD both contribute to cytidine nucleotide deamination

The data obtained did not support CDA or DCTD to be the primary deaminase, as enzyme KD did not affect the dC +9 Da signal decay curve (**Figure 6.8E** and **F**) or change incorporation at t= 0 hrs of the +9 Da label. This supports the hypothesis that the enzymes are mutually redundant deaminases of dC nucleotides. The data did, however, validate an earlier report that DCTD deaminates rC nucleotides less than CDA (**Figure 6.7**).[260] A low activity of DCTD on rC nucleotides may explain the observation that CDA KD increased dC +9 incorporation as dT: knocking down the route of CMP deamination would reduce competition for deaminated dC nucleotides to be incorporated as dT. However, this was only seen in cell lines with high exogenous dT in media, which may exacerbate incorporation competition between deaminated nucleotides.

A complicating factor is that CDA and DCTD likely do not have the same protein concentrations as each other in cells. It is therefore possible that one of the enzymes was not concentrated enough to have a proportionate impact on pyrimidine metabolism. Protein concentration difference could be accounted for by insertion of a reporter to the enzyme RNA through genome editing.[261,262] Alternatively, one could utilise methods for tunable cellular protein concentration and study the impact of different protein concentrations on pyrimidine metabolism.[148] However, my experiment made use of more than one cell line to specifically keep the issue of protein level variation in consideration, both of which had detectable protein levels of CDA and DCTD. Furthermore, my experiment provides insights into the function of CDA and DCTD by knocking down their expression transiently from their natural level with siRNAs, without potentially skewing cell metabolism by stable genomic insertion.[154]

6.8.3 *De novo* synthesis cytidine nucleotides are not metabolised by CDA and DCTD

CDA or DCTD knockdown did not cause changes in +3 Da incorporation, implying that nucleotides synthesised from L-glutamine were not dephosphorylated and deaminated in significant levels (**Figure 6.11**). This may indicate that, despite the cell lines MDA MB-231 and PC-3 having moderately high mRNA levels of the pyrophosphatase DCTPP1 and the phosphohydrolase SAMHD1 compared to the rest of the NCI-60 panel, there is little deoxycytidine triphosphate and cytidine triphosphate degradation in cells.[236,237]

7. General Discussion

Nucleotide metabolism is crucial for DNA replication; the rapid proliferation of cancer cells leads to their high demand for nucleotides and thus the nucleotide metabolism pathway enzymes have been exploited as targets for chemotherapy.[75] Resistance to nucleotide analogue chemotherapy frequently arises through the action of synthesis and salvage pathway enzymes, such as the inactivation of gemcitabine through deamination by CDA.[93,263,264] CDA is one of two cytidine nucleotide deaminases together with DCTD and it is unclear what cellular advantage is obtained by the presence of two functionally overlapping cytidine nucleotide deaminases.

In this thesis, I investigated the roles of CDA and DCTD, specifically probing the following lines of enquiry. The first was to determine whether CDA or DCTD is the principal deaminase in cells, or whether they provide substantial enzymatic redundance. The second was to discover whether newly synthesised cytidine nucleotides were dephosphorylated and deaminated in sufficient quantities by the deaminases, providing crosstalk between the *de novo* synthesis and salvage pathways. To achieve this, I set up and optimised LC-MS/MS methods for sensitive nucleic acid composition and nucleotide pool analysis, as well as siRNA KD of CDA and DCTD in two cell lines, showing sufficient duration for a pulse-chase labelled metabolomics analysis of the roles CDA and DCTD in pyrimidine metabolism.

7.1 Optimisation of HPLC-MS/MS method for nucleoside analysis

Nucleotide metabolism has previously been investigated through analysis of nucleic acid composition, achieved by DNA hydrolysis to nucleosides, with subsequent analysis by LC-MS/MS.[105,182] Nucleic acid composition analysis allows probing of the role of nucleotide

metabolism enzymes in maintaining the nucleotide pool balance as nucleotides are precursors to DNA and RNA. I set up a QQQ LC-MS/MS method capable of nucleoside analysis of samples exposed to labelled tracers to investigate the role of the CDA and DCTD.

Firstly, I adapted a nucleic acid hydrolysis workflow for HPLC-UV DNA analysis to be suitable for the higher sensitivity of LC-MS/MS analysis. I then tested three different HPLC columns to obtain complete resolution for most of 20 nucleosides and subsequently optimised the dMRM mode MS signal by varying solvent composition, parent and transition ions, and MS instrument parameters. Through the use of internal standards and UV analysis for abundant nucleoside analytes, the system shows calibration consistency and the ability to measure different nucleosides in a one millionfold concentration range. Finally, I demonstrated the system's reproducibility in nucleoside quantitation by four separate hydrolyses of the same source of DNA and MS analysis on four separate occasions.

It may be possible to gain further sensitivity through investigation of different solvent systems, parallel ion transitions and experimentation of the effect of changing delta EMV on nucleoside signal. However, my optimisation rendered the system suitable for analysis of the impact of CDA and DCTD on nucleic acid labelling proportions of the pyrimidine nucleosides. Additionally, the system is capable of sensitive nucleoside analysis of almost 20 analytes. It gives complete resolution of most of the nucleosides and has been employed collaboratively for nucleic acid composition analysis across different fields of research.[209,265,266]

7.2 Set up of HPLC-MS/MS analysis of cellular nucleotide pool

Together with analysis of nucleic acid composition, the free cellular nucleotide pool is also a target for analysis in nucleotide metabolism investigations.[58,82,196,197,215] I have developed a

sensitive LC-MS/MS method for nucleotide pool analysis that avoids processing steps that may cause nucleotide degradation.

I optimised the nucleotide extraction process for high nucleotide signal, testing three extraction solvents used in literature for extraction of polar metabolites, as well as measuring the effect of further sample treatment. The nucleotide pool would have been extracted together with non-nucleotide metabolites that could suppress MS signal through ionisation competition. To minimise analyte co-elution, as well as co-elution of isobaric compounds that may have MS signal interference, I tested two chromatographic systems. Of the functionalised reversed phase (PFP column) and the HILIC-Z column, the HILIC-Z column gave both separation of nucleotides and detection of nucleotide di- and tri-phosphates, demonstrating its suitability. I subsequently optimised the QQQ for targeted analysis of the nucleotide pool, with considerations including instrument polarity, ion fragmentation transitions, solvent composition, and delta EMV.

The HPLC-MS/MS system is capable of sensitive analysis of nucleotides and can detect the majority of nucleotides in cell samples, and its limits of nucleotide detection are comparable or lower to those reported in than literature.[196,197] Its ability to discriminate between isotopically labelled and unlabelled standards emphasises its potential for pulse-chase analyses in the nucleotide pool for investigation of the roles of CDA and DCTD in nucleotide metabolism.

There is scope for further improvement of the method. Validating the best increase in EMV in cell-derived samples could increase signal severalfold, as may the optimisation of the source conditions. It is also possible that another buffer composition, such as one containing ammonium formate, may have superior properties to those of the buffers investigated in my research. Another unknown is the extent of nucleotide signal lost due to ion suppression by non-nucleotide polar metabolites such as L-glutamine, as well as salts present in the cells and

in PBS. There have been efforts to clean nucleotide samples through the use of purification cartridges or by using chromatography that preferentially retains aromatic molecules such as porous graphitic carbon columns.[218,267] It is possible that a sample cleaning step to remove salts and non-nucleotide metabolites may reduce ionisation suppression and increase nucleotide signal further.

7.3 Cell-based tools for investigating nucleotide metabolism

To be able to study the roles of CDA and DCTD in cells through pulse-chase analysis, it was necessary to induce CDA and DCTD KD in suitable cell lines and to validate detectable labelled nucleotide precursor incorporation by MS.

I set out to achieve CDA and DCTD KD through siRNAs as they required transient transfection, which would avoid the potential cellular function distortion as is possible for stable transfection. To maximise the perturbation of CDA and DCTD KD, I sorted candidate cell lines from the NCI-60 panel by their mRNA expression of CDA and DCTD, subsequently verifying that there were detectable amounts of both enzymes at the protein levels in three cell lines. In two of those cell lines I validated that the length of siRNA KD duration was adequate for the pulse-chase experiment through immunoblotting. Finally, I exposed the cell lines to isotopically labelled nucleotide precursors from the synthesis and salvage pathways, detecting sufficient incorporation levels in the cellular DNA from the compounds dC +3 Da, rC +5 Da and Q +5 Da, allowing progression to the experiment of enzyme KD, pulse-chase labelling and MS analysis.

7.4 Role of CDA and DCTD in nucleotide pool metabolism

With the methods optimised, I was able to explore the roles of CDA and DCTD in cytidine nucleotide deamination. CDA's role in nucleotide metabolism and disease has been investigated extensively, through clinical observations and genomic manipulations.[62,82,93] In one study, CDA and DCTD were simultaneously knocked out to alter the salvage of mdC as dT, but the individual impact of both CDA and DCTD KO/KD had not been reported.[105]

My aims were to determine whether either CDA or DCTD is the principal cytidine nucleotide deaminase in cells, and whether newly synthesised cytidine nucleotides are deaminated by CDA and DCTD.

To investigate the dynamics of CDA and DCTD activity in cells, I treated cultured cells with siRNA and then exposed them to labelled nucleotide precursors in a pulse-chase experiment. Both nucleic acid and free nucleotides were analysed by LC-MS/MS to investigate what impact CDA and DCTD KD had on nucleotide metabolism.

A general observation is that the +9 Da signal derived from labelled dC decays the fastest, with the +3 Da signal derived from labelled L-glutamine decaying the slowest. My proposed explanation for this is that is as RNA has more rapid turnover than DNA, a significant proportion is catabolised to ribonucleotide monophosphates, some of which are metabolised to deoxyribonucleotides and incorporated into DNA. A part of these nucleotides are labelled, maintaining the labelling proportion in DNA. Alternatively, it could be due to increased dC flux upon nucleotide pool disruption through changing of exogenous media dC concentration. This could be clarified by repeating the pulse-chase experiment with a titration of different exogenous nucleotide concentrations to see whether the decay rate changes.

It was also noted that the decay in labelling does not change between nucleotide mono-, di- and tri-phosphates, indicating that their interconversion occurs at timeframes much faster

than one hour. Further exploration by higher frequency sample harvests in the pulse-chase experiment at the beginning of the chase could provide richer information on nucleotide phosphorylation kinetics.

From my investigation, I could not see evidence to support either CDA or DCTD to be the primary deaminase, as neither enzyme KD affect the dC +9 Da signal decay curve. This indicates that both CDA and DCTD have relatively equal roles in dC nucleotide deamination. The presence of mutually redundant CDA and DCTD in cells may serve to ensure a large dU and dT nucleotide pool, either as precursors for dTTP, or as a feedstock for carbon metabolism through uridine and thymidine catabolism.[268] This result could be complemented with tracer experiments in cells with high expression of only one of the deaminases. This could be achieved through controllable protein concentrations: Cas9 Endonuclease Dead (dCas9) and a suitable guide RNA, co-transfected with a transcription factor fused to a destabilised domain and an RNA aptamer, has been shown to allow tuning the amount of protein produced by the gene targeted by the guide RNA when cells are exposed to a small molecule stabiliser of the destabilised domain.[148] Different destabilised domain-small molecule stabiliser pairs provide orthogonality and the ability to simultaneously tune the concentrations of CDA and DCTD. An alternative approach is to perform siRNA KD and isotope tracing on a wide panel of cell lines to obtain a broader view of the relative impacts of CDA and DCTD on nucleotide metabolism. The benefit of a cell line with a tunable protein level compared to stable protein overexpression is that it maintains the protein reaching levels to physiologically relevant amounts.

With regards to rC nucleotides, KD of DCTD did not have significant impact on rC incorporation proportions, in agreement with previous literature.[260] DCTD's low activity against rC nucleotides fits with the observation that KD of CDA increased dC +9 Da incorporation as dT. Removing the route of CMP deamination may lead to less competition for incorporation of

deaminated dC nucleotides as dT. This was observed in the PC-3 cell line, which has mid micromolar concentrations of dT in its media, but not in the MDA MB-231, which does not have high dT concentrations. The extracellular dT may increase incorporation pressures for newly deaminated nucleotides, a possibility which can be verified by measuring the effect of titrating the amount of dT in different cell lines exposed to heavy labelled dC and rC.

Finally, no link was seen between CDA and DCTD knockdown and a change in +3 Da incorporation in both siRNAs. The absence of impact of deaminase KD suggests that newly synthesised cytidine nucleotides are not deaminated in significant quantities. The reason for this may be that nucleotide triphosphate and diphosphates are not significantly degraded in cells and thus synthesised cytidine nucleotides are not substrates for CDA or DCTD. Alternatively, the effect of CDA and DCTD deamination of nucleotides from *de novo* synthesis may be obscured from CDA and DCTD activity on salvaged nucleotides, which may compete for incorporation with synthesis nucleotides. It is also possible that the absence of evidence for CDA and DCTD impact on synthesised cytidine nucleotides may be due to off-target effects due to the individual siRNA: CDA siRNA3 and DCTD siRNA2 gave increased cytidine +3 Da nucleotide labelling, while the other siRNAs did not. This can be verified by testing other orthogonal siRNAs in a simple steady-state labelling experiment.

8. References

1. Pertea, M. *et al.* CHESS: A new human gene catalog curated from thousands of large-scale RNA sequencing experiments reveals extensive transcriptional noise. *Genome Biol.* **19**, 208 (2018).
2. Traut, T. W. Physiological concentrations of purines and pyrimidines. *Molecular and Cellular Biochemistry* vol. 140 1–22 (1994).
3. Phear, G. & Meuth, M. The genetic consequences of DNA precursor pool imbalance: sequence analysis of mutations induced by excess thymidine at the hamster aptr locus. *Mutat. Res. - Fundam. Mol. Mech. Mutagen.* **214**, 201–206 (1989).
4. Houghton, J. A., Harwood, F. G. & Tillman, D. M. Thymineless death in colon carcinoma cells is mediated via Fas signaling. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 8144–8149 (1997).
5. Kellner, V. & Luke, B. Molecular and physiological consequences of faulty eukaryotic ribonucleotide excision repair. *EMBO J.* **39**, e102309 (2020).
6. Mathews, C. K. Deoxyribonucleotides as genetic and metabolic regulators. *FASEB Journal* vol. 28 (2014).
7. Liu, Y. *et al.* Digestion of nucleic acids starts in the stomach. *Sci. Rep.* **5**, 1–11 (2015).
8. Sorrentino, S. The eight human 'canonical' ribonucleases: Molecular diversity, catalytic properties, and special biological actions of the enzyme proteins. *FEBS Letters* vol. 584 2194–2200 (2010).
9. Takeshita, H. *et al.* Mammalian deoxyribonucleases I are classified into three types: Pancreas, parotid, and pancreas-parotid (mixed), based on differences in their tissue concentrations. *Biochem. Biophys. Res. Commun.* **269**, 481–484 (2000).
10. Bianchi, V. & Spychala, J. Mammalian 5′-Nucleotidases. *Journal of Biological Chemistry* vol. 278 46195–46198 (2003).
11. Theisinger, A., Grenacher, B., Rech, K. S. & Scharrer, E. Nucleosides are efficiently absorbed by Na+-dependent transport across the intestinal brush border membrane in veal calves. *J. Dairy Sci.* **85**, 2308–2314 (2002).
12. Carver, J. D. & Allan Walker, W. The role of nucleotides in human nutrition. *The Journal of Nutritional Biochemistry* vol. 6 58–72 (1995).
13. Pastor-Anglada, M. & Pérez-Torras, S. Emerging Roles of Nucleoside Transporters. *Front. Pharmacol.* **9**, 606 (2018).
14. Ritzel, M. W. L. *et al.* Molecular cloning and functional expression of cDNAs encoding a human Na+-nucleoside cotransporter (hCNT1). *Am. J. Physiol. - Cell Physiol.* **272**, (1997).
15. Wang, J. *et al.* Na $^{+}$ -dependent purine nucleoside transporter from human kidney: cloning and functional characterization. *Am. J. Physiol. Physiol.* **273**, F1058–F1065 (1997).
16. Ritzel, M. W. L. *et al.* Molecular Identification and Characterization of Novel Human and Mouse Concentrative Na+-Nucleoside Cotransporter Proteins

(hCNT3 and mCNT3) Broadly Selective for Purine and Pyrimidine Nucleosides (System cib). *J. Biol. Chem.* **276**, 2914–2927 (2001).

17. Ashraf, T., Kao, A. & Bendayan, R. Functional expression of drug transporters in glial cells: Potential role on drug delivery to the CNS. in *Advances in Pharmacology* vol. 71 45–111 (Academic Press Inc., 2014).
18. Young, J. D., Yao, S. Y. M., Sun, L., Cass, C. E. & Baldwin, S. A. Human equilibrative nucleoside transporter (ENT) family of nucleoside and nucleobase transporter proteins. *Xenobiotica* vol. 38 995–1021 (2008).
19. Cano-Soldado, P. & Pastor-Anglada, M. Transporters that translocate nucleosides and structural similar drugs: structural requirements for substrate recognition. *Med. Res. Rev.* **32**, 428–457 (2012).
20. Fox, I. H. & Kelley, W. N. Human phosphoribosylpyrophosphate synthetase. Distribution, purification, and properties. *J. Biol. Chem.* **246**, (1971).
21. HARTMAN, S. C. PHOSPHORIBOSYL PYROPHOSPHATE AMIDOTRANSFERASE. PURIFICATION AND GENERAL. *J. Biol. Chem.* **238**, (1963).
22. Aimi, J., Qiu, H., Williams, J., Zalkin, H. & Dixon, J. E. De novo purine nucleotide biosynthesis: Cloning of human and avian cDNAs encoding the trifunctional glycinamide ribonucleotide synthetase-aminoimidazole ribonucleotide synthetase-glycinamide ribonucleotide transformylase by functional complementation in E.coli. *Nucleic Acids Res.* **18**, (1990).
23. Barnes, T. S. *et al.* Purification of, Generation of Monoclonal Antibodies to, and Mapping of Phosphoribosy N-Formylglycinamide Amidotransferase. *Biochemistry* **33**, (1994).
24. Schild, D., Brake, A. J., Kiefer, M. C., Young, D. & Barr, P. J. Cloning of three human multifunctional de novo purine biosynthetic genes by functional complementation of yeast mutations. *Proc. Natl. Acad. Sci. U. S. A.* **87**, (1990).
25. Stone, R. L., Zalkin, H. & Dixon, J. E. Expression, purification, and kinetic characterization of recombinant human adenylosuccinate lyase. *J. Biol. Chem.* **268**, (1993).
26. Van Der Weyden, M. B. & Kelly, W. N. Human adenylosuccinate synthetase. Partial purification, kinetic and regulatory properties of the enzyme from placenta. *J. Biol. Chem.* **249**, (1974).
27. Abrams, R. & Bentley, M. Biosynthesis of nucleic acid purines. III. Guanosine 5′-phosphate formation from xanthosine 5′-phosphate and l-glutamine. *Arch. Biochem. Biophys.* **79**, (1959).
28. The Control of Purine Biosynthesis in Cultured Mammalian Cells. *J. Biol. Chem.* **235**, (1960).
29. Panayiotou, C., Solaroli, N. & Karlsson, A. The many isoforms of human adenylate kinases. *International Journal of Biochemistry and Cell Biology* vol. 49 (2014).
30. Agarwal, K. C., Miech, R. P. & Parks, R. E. Guanylate Kinases from Human Erythrocytes, Hog Brain, and Rat Liver. *Methods Enzymol.* **51**, (1978).
31. MOORE, E. C. & REICHARD, P. ENZYMATIC SYNTHESIS OF DEOXYRIBONUCLEOTIDES. VI. THE CYTIDINE DIPHOSPHATE REDUCTASE SYSTEM FROM NOVIKOFF HEPATOMA. *J. Biol. Chem.* **239**, (1964).

32. Mourad, N. & Parks, R. E. Erythrocytic nucleoside diphosphokinase. II. Isolation and kinetics. *J. Biol. Chem.* **241**, (1966).

33. Arnér, E. S. J. & Eriksson, S. Mammalian deoxyribonucleoside kinases. *Pharmacology and Therapeutics* vol. 67 (1995).

34. Coleman, P. F., Suttle, D. P. & Stark, G. R. Purification from hamster cells of the multifunctional protein that initiates de novo synthesis of pyrimidine nucleotides. *J. Biol. Chem.* **252**, (1977).

35. Lu, Z. H., Zhang, R. & Diasio, R. B. Purification and characterization of dihydropyrimidine dehydrogenase from human liver. *J. Biol. Chem.* **267**, (1992).

36. McClard, R. W., Black, M. J., Livingstone, L. R. & Jones, M. E. Isolation and Initial Characterization of the Single Polypeptide That Synthesizes Uridine 5'-Monophosphate from Orotate in Ehrlich Ascites Carcinoma. Purification by Tandem Affinity Chromatography of Uridine-5'-monophosphate Synthase. *Biochemistry* **19**, (1980).

37. Van Rompay, A. R., Johansson, M. & Karlsson, A. Identification of a novel human adenylate kinase cDNA cloning, expression analysis, chromosome localization and characterization of the recombinant protein. *Eur. J. Biochem.* **261**, 509–517 (1999).

38. Ladner, R. D., McNulty, D. E., Carr, S. A., Roberts, G. D. & Caradonna, S. J. Characterization of distinct nuclear and mitochondrial forms of human deoxyuridine triphosphate nucleotidohydrolase. *J. Biol. Chem.* **271**, (1996).

39. Lv, Y. T. *et al.* A novel approach to cloning and expression of human thymidylate synthase. *Asian Pacific J. Cancer Prev.* **14**, (2013).

40. Davisson, V. J., Sirawaraporn, W. & Santi, D. V. Expression of human thymidylate synthase in Escherichia coli. *J. Biol. Chem.* **264**, (1989).

41. Lee, L. S. & Cheng, Y. C. Human thymidylate kinase. Purification, characterization, and kinetic behavior of the thymidylate kinase derived from chronic myelocytic leukemia. *J. Biol. Chem.* **252**, (1977).

42. Formation of Cytidine Nucleotides from Uridine Nucleotides by Soluble Mammalian Enzymes: Requirements for Glutamine and Guanosine Nucleotides. *J. Biol. Chem.* **235**, (1960).

43. Durham, J. P. & Ives, D. H. Deoxycytidine kinase. I. Distribution in normal and neoplastic tissues and interrelationships of deoxycytidine and 1-beta-D-arabinofuranosylcytosine phosphorylation. *Mol. Pharmacol.* **5**, (1969).

44. Maley, G. F., Lobo, A. P. & Maley, F. Properties of an affinity-column-purified human deoxycytidylate deaminase. *Biochim. Biophys. Acta (BBA)/Protein Struct. Mol.* **1162**, 161–170 (1993).

45. Camiener, G. W. Studies of the enzymatic deamination of ara-cytidine-V. Inhibition in vitro and in vivo by tetrahydrouridine and other reduced pyrimidine nucleosides. *Biochem. Pharmacol.* **17**, (1968).

46. Liu, N. *et al.* N6 -methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. *Nature* **518**, 560–564 (2015).

47. Van Rompay, A. R., Norda, A., Lindén, K., Johansson, M. & Karlsson, A. Phosphorylation of uridine and cytidine nucleoside analogs by two human uridine-cytidine kinases. *Mol. Pharmacol.* **59**, 1181–1186 (2001).

48. Takeda, D. Y. & Dutta, A. DNA replication and progression through S phase.

Oncogene vol. 24 2827–2843 (2005).

49. Salem, C., El-Alfy, M. & Leblond, C. P. Changes in the rate of RNA synthesis during the cell cycle. *Anat. Rec.* **250**, 6–12 (1998).
50. Baudrimont, A. *et al.* Multiplexed gene control reveals rapid mRNA turnover. *Sci. Adv.* **3**, e1700006 (2017).
51. Turner, M. K., Abrams, R. & Lieberman, I. Levels of ribonucleotide reductase activity during the division cycle of the L cell. *J. Biol. Chem.* **243**, (1968).
52. D'Angiolella, V. *et al.* Cyclin F-mediated degradation of ribonucleotide reductase M2 controls genome integrity and DNA repair. *Cell* **149**, 1023–1034 (2012).
53. Franzolin, E. *et al.* The deoxynucleotide triphosphohydrolase SAMHD1 is a major regulator of DNA precursor pools in mammalian cells. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 14272–14277 (2013).
54. Laguette, N. *et al.* SAMHD1 is the dendritic- and myeloid-cell-specific HIV-1 restriction factor counteracted by Vpx. *Nature* **474**, 654–657 (2011).
55. Smith, A. E. & Yamada, E. W. Dihydrouracil dehydrogenase of rat liver. Separation of hydrogenase and dehydrogenase activities. *J. Biol. Chem.* **246**, (1971).
56. McAllan, A. B. The degradation of nucleic acids in, and the removal of breakdown products from the small intestines of steers. *Br. J. Nutr.* **44**, 99–112 (1980).
57. Vértessy, B. G. & Tóth, J. Keeping uracil out of DNA: physiological role, structure and catalytic mechanism of dUTPases. *Acc. Chem. Res.* **42**, 97–106 (2009).
58. Requena, C. E., Peŕez-Moreno, G., Ruiz-PÉREZ, L. M., Vidal, A. E. & González-Pacanowska, D. The NTP pyrophosphatase DCTPP1 contributes to the homoeostasis and cleansing of the dNTP pool in human cells. *Biochem. J.* **459**, 171–180 (2014).
59. Martínez-Arribas, B. *et al.* DCTPP1 prevents a mutator phenotype through the modulation of dCTP, dTTP and dUTP pools. *Cell. Mol. Life Sci.* **77**, 1645–1660 (2020).
60. Tsuzuki, T. *et al.* Spontaneous tumorigenesis in mice defective in the MTH1 gene encoding 8-oxo-dGTPase. *Proc. Natl. Acad. Sci. U. S. A.* **98**, (2001).
61. Fugger, K. *et al.* Targeting the nucleotide salvage factor DNPH1 sensitizes BRCA-deficient cells to PARP inhibitors. *Science (80-.).* **372**, 156–165 (2021).
62. Zauri, M. *et al.* CDA directs metabolism of epigenetic nucleosides revealing a therapeutic window in cancer. *Nature* **524**, 114–118 (2015).
63. Cheng, K. C., Cahill, D. S., Kasai, H., Nishimura, S. & Loeb, L. A. 8-Hydroxyguanine, an abundant form of oxidative DNA damage, causes G → T and A → C substitutions. *J. Biol. Chem.* **267**, (1992).
64. Alexandrov, L. B. *et al.* Signatures of mutational processes in human cancer. *Nature* **500**, 415–421 (2013).
65. Krokan, H. E., Drabløs, F. & Slupphaug, G. Uracil in DNA - Occurrence, consequences and repair. *Oncogene* vol. 21 8935–8948 (2002).
66. Hollstein, M., Sidransky, D., Vogelstein, B. & Harris, C. C. p53 mutations in human cancers. *Science (80-.).* **253**, 49–53 (1991).

67. Barbacid, M. ras Genes. *Annu. Rev. Biochem.* **56**, 779–827 (1987).

68. Truninger, K. *et al.* Immunohistochemical analysis reveals high frequency of PMS2 defects in colorectal cancer. *Gastroenterology* **128**, 1160–1171 (2005).

69. Lyer, R. R., Pluciennik, A., Burdett, V. & Modrich, P. L. DNA mismatch repair: Functions and mechanisms. *Chemical Reviews* vol. 106 302–323 (2006).

70. Manhart, C. M. & Alani, E. DNA replication andmismatch repair safeguard againstmetabolic imbalances. *Proceedings of the National Academy of Sciences of the United States of America* vol. 114 5561–5563 (2017).

71. Lu, R., Nash, H. M. & Verdine, G. L. A mammalian DNA repair enzyme that excises oxidatively damaged guanines maps to a locus frequently lost in lung cancer. *Curr. Biol.* **7**, 397–407 (1997).

72. Nash, H. M. *et al.* Cloning of a yeast 8-oxoguanine DNA glycosylase reveals the existence of a base-excision DNA-repair protein superfamily. *Curr. Biol.* **6**, 968–980 (1996).

73. Olsen, L. C., Aasland, R., Wittwer, C. U., Krokan, H. E. & Helland, D. E. Molecular cloning of human uracil-DNA glycosylase, a highly conserved DNA repair enzyme. *EMBO J.* **8**, 3121–3125 (1989).

74. Slupphaug, G. *et al.* Nuclear and mitochondrial forms of human uracil-DNA glycosylase are encoded by the same gene. *Nucleic Acids Res.* **21**, 2579–2584 (1993).

75. Borek, E. *et al.* High Turnover Rate of Transfer RNA in Tumor Tissue. *Cancer Res.* **37**, (1977).

76. Liu, Y. C. *et al.* Global regulation of nucleotide biosynthetic genes by c-myc. *PLoS One* **3**, (2008).

77. Robitaille, A. M. *et al.* Quantitative phosphoproteomics reveal mTORC1 activates de novo pyrimidine synthesis. *Science (80-.).* **339**, 1320–1323 (2013).

78. Ben-Sahra, I., Howell, J. J., Asara, J. M. & Manning, B. D. Stimulation of de novo pyrimidine synthesis by growth signaling through mTOR and S6K1. *Science (80-.).* **339**, 1323–1328 (2013).

79. Vasen, H. F. A. What is hereditary nonpolyposis colorectal cancer (HNPCC). in *Anticancer Research* vol. 14 1613–1615 (Anticancer Res, 1994).

80. Veigl, M. L. *et al.* Biallelic inactivation of hMLH1 by epigenetic gene silencing, a novel mechanism causing human MSI cancers. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 8698–8702 (1998).

81. Liu, B. *et al.* Mismatch repair gene defects in sporadic colorectal cancers with microsatellite instability. *Nat. Genet.* **9**, (1995).

82. Gemble, S. *et al.* Pyrimidine Pool Disequilibrium Induced by a Cytidine Deaminase Deficiency Inhibits PARP-1 Activity, Leading to the Under Replication of DNA. *PLOS Genet.* **11**, e1005384 (2015).

83. Longley, D. B., Harkin, D. P. & Johnston, P. G. 5-Fluorouracil: Mechanisms of action and clinical strategies. *Nature Reviews Cancer* vol. 3 330–338 (2003).

84. Santi, D. V., McHenry, C. S. & Sommer, H. Mechanism of Interaction of Thymidylate Synthetase with 5-Fluorodeoxyuridylate. *Biochemistry* **13**, 471–481 (1974).

85. Cerqueira, N. M. F. S. A., Fernandes, P. A. & Ramos, M. J. Understanding

ribonucleotide reductase inactivation by gemcitabine. *Chem. - A Eur. J.* **13**, 8507–8515 (2007).

86. Gandhi, V. & Plunkett, W. Cellular and clinical pharmacology of fludarabine. *Clinical P1. Gandhi, V. & Plunkett, W. Cellular and clinical pharmacology of fludarabine. Clinical Pharmacokinetics vol. 41 (2002).harmacokinetics* vol. 41 (2002).

87. De Sousa Cavalcante, L. & Monteiro, G. Gemcitabine: Metabolism and molecular mechanisms of action, sensitivity and chemoresistance in pancreatic cancer. *European Journal of Pharmacology* vol. 741 8–16 (2014).

88. Vesely, J. & Cihak, A. 5-Aza-2'-deoxycytidine: preclinical studies in mice. *Neoplasma* **27**, 113–119 (1980).

89. Taylor, S. M. 5-Aza-2′-deoxycytidine: Cell differentiation and DNA methylation. *Leukemia* **7**, 3–8 (1993).

90. Adams, R. L. P., Fulton, J. & Kirk, D. The effect of 5-azadeoxycytidine on cell growth and DNA methylation. *BBA - Gene Struct. Expr.* **697**, 286–294 (1982).

91. Dagogo-Jack, I. & Shaw, A. T. Tumour heterogeneity and resistance to cancer therapies. *Nature Reviews Clinical Oncology* vol. 15 81–94 (2018).

92. Qin, T. *et al.* Mechanisms of Resistance to Decitabine in the Myelodysplastic Syndrome. *PLoS One* **6**, e23372 (2011).

93. Frances, A. & Cordelier, P. The Emerging Role of Cytidine Deaminase in Human Diseases: A New Opportunity for Therapy? *Molecular Therapy* vol. 28 357–366 (2020).

94. Gandhi, V., Xu, Y. Z. & Estey, E. Accumulation of arabinosyluracil 5'-triphosphate during arabinosylcytosine therapy in circulating blasts of patients with acute myelogenous leukemia. *Clin. Cancer Res.* **4**, 1719–1726 (1998).

95. Drake, J. C., Hande, K. R., Fuller, R. W. & Chabner, B. A. Cytidine and deoxycytidylate deaminase inhibition by uridine analogs. *Biochem. Pharmacol.* **29**, 807–811 (1980).

96. Schröder, J. K., Seidelmann, M., Kirch, H. C., Seeber, S. & Schütte, J. Assessment of resistance induction to cytosine arabinoside following transfer and overexpression of the deoxycytidylate deaminase gene in vitro. *Leuk. Res.* **22**, 619–624 (1998).

97. Lamba, J. K. Genetic factors influencing cytarabine therapy. *Pharmacogenomics* vol. 10 1657–1674 (2009).

98. Elander, N. O. *et al.* Intratumoural expression of deoxycytidylate deaminase or ribonuceotide reductase subunit M1 expression are not related to survival in patients with resected pancreatic cancer given adjuvant chemotherapy. *Br. J. Cancer* **118**, 1084–1088 (2018).

99. Jiang, W., Lu, Z., He, Y. & Diasio, R. B. Dihydropyrimidine dehydrogenase activity in hepatocellular carcinoma: implication in 5-fluorouracil-based chemotherapy. *Clin. Cancer Res.* **3**, (1997).

100. Byoung, K. Y. *et al.* Identification of genes conferring resistance to 5-fluorouracil. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 12938–12943 (2009).

101. Johnston, P. G. *et al.* Thymidylate Synthase Gene and Protein Expression Correlate and Are Associated with Response to 5-Fluorouracil in Human Colorectal and Gastric Tumors. *Cancer Res.* **55**, (1995).

102. Lenz, H. J. *et al.* p53 point mutations and thymidylate synthase messenger

RNA levels in disseminated colorectal cancer: an analysis of response and survival. *Clin. Cancer Res.* **4**, (1998).

103. Oguri, T. *et al.* The determinants of sensitivity and acquired resistance to gemcitabine differ in non-small cell lung cancer: A role of ABCC5 in gemcitabine sensitivity. *Mol. Cancer Ther.* **5**, 1800–1806 (2006).
104. Nygaard, P. On the role of cytidine deaminase in cellular metabolism. *Adv. Exp. Med. Biol.* **195 Pt B**, 415–420 (1986).
105. Spada, F. *et al.* Active turnover of genomic methylcytosine in pluripotent cells. *Nat. Chem. Biol.* (2020) doi:10.1038/s41589-020-0621-y.
106. Frese, K. K. *et al.* Nab-paclitaxel potentiates gemcitabine activity by reducing cytidine deaminase levels in a mouse model of pancreatic cancer. *Cancer Discov.* **2**, 260–269 (2012).
107. Morita, T., Matsuzaki, A., Kurokawa, S. & Tokue, A. Forced Expression of Cytidine Deaminase Confers Sensitivity to Capecitabine. *Oncology* **65**, 267–274 (2003).
108. Stelzer, G. *et al.* The GeneCards suite: From gene data mining to disease genome sequence analyses. *Curr. Protoc. Bioinforma.* **2016**, 1.30.1-1.30.33 (2016).
109. Moro-Bulnes, A. *et al.* Contribution of Cytidine Deaminase to Thymidylate Biosynthesis in Trypanosoma brucei : Intracellular Localization and Properties of the Enzyme . *mSphere* **4**, (2019).
110. Johansson, E., Mejlhede, N., Neuhard, J. & Larsen, S. Crystal structure of the tetrameric cytidine deaminase from Bacillus subtilis at 2.0 Å resolution. *Biochemistry* **41**, 2563–2570 (2002).
111. Reizer, J., Buskirk, S., Reizer, A., Saier, M. H. & Bairoch, A. A novel zinc-binding motif found in two ubiquitous deaminase families. *Protein Sci.* **3**, 853–856 (1994).
112. Chung, S. J., Fromme, J. C. & Verdine, G. L. Structure of human cytidine deaminase bound to a potent inhibitor. *J. Med. Chem.* **48**, 658–660 (2005).
113. Sehnal, D., Rose, A., Koča, J., Burley, S. & Velankar, S. Mol*: towards a common library and tools for web molecular graphics. *Proc. Work. Mol. Graph. Vis. Anal. Mol. data* (2018).
114. wwPDB: 2W4L. https://www.wwpdb.org/pdb?id=pdb_00002w4l.
115. Watanabe, S. I. & Uchida, T. Expression of cytidine deaminase in human solid tumors and its regulation by 1 α,25-dihydroxyvitamin D3. *Biochim. Biophys. Acta - Mol. Cell Res.* **1312**, 99–104 (1996).
116. Su, A. I. *et al.* A gene atlas of the mouse and human protein-encoding transcriptomes. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 6062–6067 (2004).
117. Wu, C., Jin, X., Tsueng, G., Afrasiabi, C. & Su, A. I. BioGPS: Building your own mash-up of gene annotations and expression profiles. *Nucleic Acids Res.* **44**, D313–D316 (2016).
118. Wu, C., MacLeod, I. & Su, A. I. BioGPS and MyGene.info: Organizing online, gene-centric information. *Nucleic Acids Res.* **41**, D561–D565 (2013).
119. Wu, C. *et al.* BioGPS: An extensible and customizable portal for querying and organizing gene annotation resources. *Genome Biol.* **10**, R130 (2009).
120. Chabosseau, P. *et al.* Pyrimidine pool imbalance induced by BLM helicase

deficiency contributes to genetic instability in Bloom syndrome. *Nat. Commun.* **2**, 1–6 (2011).

121. Momparler, R. L. *et al.* Kinetic interaction of 5-AZA-2'-deoxycytidine-5'-monophosphate and its 5'-triphosphate with deoxycytidylate deaminase. *Mol. Pharmacol.* **25**, (1984).

122. McFerran, N. V., Smyth, M. & Orsi, B. A. Regulation of cytidine aminohydrolase. *Biochem. J.* **114**, 8P (1969).

123. Ge, Y. *et al.* The Role of Cytidine Deaminase and GATA1 Mutations in the Increased Cytosine Arabinoside Sensitivity of Down Syndrome Myeloblasts and Leukemia Cell Lines. *Cancer Res.* **64**, 728–735 (2004).

124. Demontis, S. *et al.* Isolation and characterization of the gene coding for human cytidine deaminase. *Biochim. Biophys. Acta - Gene Struct. Expr.* **1443**, 323–333 (1998).

125. Binenbaum, Y. *et al.* Transfer of miRNA in macrophage-derived exosomes induces drug resistance in pancreatic adenocarcinoma. *Cancer Res.* **78**, 5287–5299 (2018).

126. Ye, F. G. *et al.* Cytidine deaminase axis modulated by mir-484 differentially regulates cell proliferation and chemoresistance in breast cancer. *Cancer Res.* **75**, 1504–1515 (2015).

127. Rajabpour, A. *et al.* MiR-608 regulating the expression of ribonucleotide reductase M1 and cytidine deaminase is repressed through induced gemcitabine chemoresistance in pancreatic cancer cells. *Cancer Chemother. Pharmacol.* **80**, 765–775 (2017).

128. Eliopoulos, N., Cournoyer, D. & Momparler, R. L. Drug resistance to 5-aza-2'-deoxycytidine, 2',2'-difluorodeoxycytidine, and cytosine arabinoside conferred by retroviral-mediated transfer of human cytidine deaminase cDNA into murine cells. *Cancer Chemother. Pharmacol.* **42**, 373–378 (1998).

129. Lemaire, M., Momparler, L. F., Raynal, N. J. M., Bernstein, M. L. & Momparler, R. L. Inhibition of cytidine deaminase by zebularine enhances the antineoplastic action of 5-aza-2-deoxycytidine. *Cancer Chemother. Pharmacol.* **63**, (2009).

130. Savona, M. R. *et al.* An oral fixed-dose combination of decitabine and cedazuridine in myelodysplastic syndromes: a multicentre, open-label, dose-escalation, phase 1 study. *Lancet Haematol.* **6**, e194–e203 (2019).

131. Laboratories, L. R. *Cellular Elimination of 2',2'-Difluorodeoxycytidine 5'-Triphosphate: A Mechanism of Self-Potentiation1*. *ICANCER RESEARCH* vol. 52 (1992).

132. Gribaudo, G., Riera, L., Caposio, P., Maley, F. & Landolfo, S. Human cytomegalovirus requires cellular deoxycytidylate deaminase for replication in quiescent cells. *J. Gen. Virol.* **84**, 1437–1441 (2003).

133. MALEY, F. & MALEY, G. F. Nucleotide interconversions. II. Elevation of deoxycytidylate deaminase and thymidylate synthetase in regenerating rat liver. *J. Biol. Chem.* **235**, (1960).

134. Kuhn, K., Bertling, W. M. & Emmrich, F. Cloning of a Functional cDNA for Human Cytidine Deaminase (CDD) and Its Use as a Marker of Monocyte/Macrophage Differentiation. *Biochem. Biophys. Res. Commun.* **190**, (1993).

135. Yue, L. *et al.* A functional single-nucleotide polymorphism in the human

cytidine deaminase gene contributing to ara-C sensitivity. *Pharmacogenetics* **13**, (2003).

136. Sugiyama, E. *et al.* Ethnic differences of two non-synonymous single nucleotide polymorphisms in CDA gene. *Drug Metab. Pharmacokinet.* **24**, (2009).
137. Fukunaga, A. K., Marsh, S., Murry, D. J., Hurley, T. D. & McLeod, H. L. Identification and analysis of single-nucleotide polymorphisms in the gemcitabine pharmacologic pathway. *Pharmacogenomics J.* **4**, 307–314 (2004).
138. Micozzi, D. *et al.* Human cytidine deaminase: A biochemical characterization of its naturally occurring variants. *Int. J. Biol. Macromol.* **63**, 64–74 (2014).
139. Kirch, H. C. *et al.* Recombinant gene products of two natural variants of the human cytidine deaminase gene confer different deamination rates of cytarabine in vitro. *Exp. Hematol.* **26**, (1998).
140. Farrell, J. J. *et al.* Cytidine deaminase single-nucleotide polymorphism is predictive of toxicity from gemcitabine in patients with pancreatic cancer: RTOG 9704. *Pharmacogenomics J.* **12**, 395–403 (2012).
141. Hryciuk, B. *et al.* Severe acute toxicity following gemcitabine administration: A report of four cases with cytidine deaminase polymorphisms evaluation. *Oncol. Lett.* **15**, 1912–1916 (2018).
142. Okazaki, T., Javle, M., Tanaka, M., Abbruzzese, J. L. & Li, D. Single nucleotide polymorphisms of gemcitabine metabolic genes and pancreatic cancer survival and drug toxicity. *Clin. Cancer Res.* **16**, (2010).
143. Rha, S. Y. *et al.* An Association Between *RRM1* Haplotype and Gemcitabine-Induced Neutropenia in Breast Cancer Patients. *Oncologist* **12**, 622–630 (2007).
144. Matt, S. M. *et al.* Inhibition of DNA methylation with zebularine alters lipopolysaccharide sickness behavior and neuroinflammation in mice. *Front. Neurosci.* **12**, (2018).
145. Cheng, J. C. *et al.* Inhibition of DNA Methylation and Reactivation of Silenced Genes by Zebularine. *JNCI J. Natl. Cancer Inst.* **95**, 399–409 (2003).
146. Adli, M. The CRISPR tool kit for genome editing and beyond. *Nature Communications* vol. 9 1–13 (2018).
147. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science (80-.).* **337**, 816–821 (2012).
148. Maji, B. *et al.* Multidimensional chemical control of CRISPR-Cas9. *Nat. Chem. Biol.* **13**, 9–11 (2017).
149. Parsi, K. M., Hennessy, E., Kearns, N. & Maehr, R. Using an inducible CRISPR-dCas9-KRAB effector system to dissect transcriptional regulation in human embryonic stem cells. in *Methods in Molecular Biology* vol. 1507 221–233 (Humana Press Inc., 2017).
150. Clarke, A. R. Manipulating the germline: its impact on the study of carcinogenesis. *Carcinogenesis* **21**, 435–441 (2000).
151. Zhang, J. *et al.* Conditional gene manipulation: Cre-ating a new biological era. *Journal of Zhejiang University: Science B* vol. 13 511–524 (2012).
152. Robinson, H. L. *et al.* Cancer induction by insertional mutagenesis: the role of viral genes in avian leukosis virus induced cancers. *Prog. Clin. Biol. Res.* **119**,

37–42 (1983).

153. Varmus, H. E. Using retroviruses as insertional mutagens to identify cellular oncogenes. *Prog. Clin. Biol. Res.* **119**, 23–35 (1983).

154. Fu, Y. *et al.* High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.* **31**, 822–826 (2013).

155. Bartlett, D. W. & Davis, M. E. Effect of siRNA nuclease stability on the in vitro and in vivo kinetics of siRNA-mediated gene silencing. *Biotechnol. Bioeng.* **97**, 909–921 (2007).

156. Kaur, G. & Dufour, J. M. Cell lines. *Spermatogenesis* **2**, 1–5 (2012).

157. Pastor, D. M. *et al.* Primary cell lines: False representation or model system? A comparison of four human colorectal tumors and their coordinately established cell lines. *Int. J. Clin. Exp. Med.* **3**, 69–83 (2010).

158. Mews, P. *et al.* Alcohol metabolism contributes to brain histone acetylation. *Nature* **574**, (2019).

159. Wishart, D. S. *et al.* HMDB: The human metabolome database. *Nucleic Acids Res.* **35**, (2007).

160. Stevenson-Paulik, J. *et al.* Inositol phosphate metabolomics: Merging genetic perturbation with modernized radiolabeling methods. *Methods* **39**, 112–121 (2006).

161. HU, V. W. & HEIKKA, D. S. Radiolabeling revisited: metabolic labeling with 35 S-methionine inhibits cell cycle progression, proliferation, and survival . *FASEB J.* **14**, 448–454 (2000).

162. Mancuso, A. *et al.* Real-time detection of 13C NMR labeling kinetics in perfused EMT6 mouse mammary tumor cells and βHC9 mouse insulinomas. *Biotechnol. Bioeng.* **87**, 835–848 (2004).

163. Beckonert, O. *et al.* Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts. *Nat. Protoc.* **2**, (2007).

164. Fan, T. W. M., Tan, J., McKinney, M. M. & Lane, A. N. Stable isotope resolved metabolomics analysis of ribonucleotide and RNA metabolism in human lung cancer cells. *Metabolomics* **8**, 517–527 (2012).

165. Lu, W. *et al.* Metabolite Measurement: Pitfalls to Avoid and Practices to Follow. *Annu. Rev. Biochem.* **86**, 277–304 (2017).

166. Emwas, A. H. M. The strengths and weaknesses of NMR spectroscopy and mass spectrometry with particular focus on metabolomics research. *Methods Mol. Biol.* **1277**, 161–193 (2015).

167. Jang, C., Chen, L. & Rabinowitz, J. D. Metabolomics and Isotope Tracing. *Cell* vol. 173 (2018).

168. Lai, Z. & Fiehn, O. Mass spectral fragmentation of trimethylsilylated small molecules. *Mass Spectrometry Reviews* vol. 37 (2018).

169. Coulier, L. *et al.* Simultaneous quantitative analysis of metabolites using ion-pair liquid chromatography-electrospray ionization mass spectrometry. *Anal. Chem.* **78**, 6573–6582 (2006).

170. Yuan, M., Breitkopf, S. B., Yang, X. & Asara, J. M. A positive/negative ion-switching, targeted mass spectrometry-based metabolomics platform for bodily fluids, cells, and fresh and fixed tissue. *Nat. Protoc.* **7**, 872–881 (2012).

171. Bajad, S. U. *et al.* Separation and quantitation of water soluble cellular metabolites by hydrophilic interaction chromatography-tandem mass spectrometry. *J. Chromatogr. A* **1125**, 76–88 (2006).

172. Jandera, P. Stationary and mobile phases in hydrophilic interaction chromatography: A review. *Analytica Chimica Acta* vol. 692 1–25 (2011).

173. Soga, T. *et al.* Quantitative Metabolome Analysis Using Capillary Electrophoresis Mass Spectrometry. *J. Proteome Res.* **2**, 488–494 (2003).

174. Ramautar, R. Capillary Electrophoresis-Mass Spectrometry for Clinical Metabolomics. in *Advances in Clinical Chemistry* vol. 74 1–34 (Academic Press Inc., 2016).

175. Antoniewicz, M. R. A guide to 13C metabolic flux analysis for the cancer biologist. *Experimental and Molecular Medicine* vol. 50 19 (2018).

176. Hui, S. *et al.* Glucose feeds the TCA cycle via circulating lactate. *Nat. Publ. Gr.* **551**, (2017).

177. Reaves, M. L., Young, B. D., Hosios, A. M., Xu, Y.-F. & Rabinowitz, J. D. Pyrimidine homeostasis is accomplished by directed overflow metabolism. *Nature* **500**, 237–241 (2013).

178. Strong, J. M., Anderson, L. W., Monks, A., Chisena, C. A. & Cysyk, R. L. A 13C tracer method for quantitating de novo pyrimidine biosynthesis in vitro and in vivo. *Anal. Biochem.* **132**, 243–253 (1983).

179. ZAHAREVITZ, D. W. *et al.* Contribution of de-novo and salvage synthesis to the uracil nucleotide pool in mouse tissues and tumors in vivo. *Eur. J. Biochem.* **210**, 293–296 (1992).

180. Daley, L. S. & Bidwell, R. G. S. Phosphoserine and Phosphohydroxypyruvic Acid. *Plant Physiol.* **60**, 109–114 (1977).

181. Spada, F. *et al.* Oxidative and non-oxidative active turnover of genomic methylcytosine in distinct pluripotent states. *bioRxiv* 846584 (2019) doi:10.1101/846584.

182. Pfaffeneder, T. *et al.* Tet oxidizes thymine to 5-hydroxymethyluracil in mouse embryonic stem cell DNA. *Nat. Chem. Biol.* **10**, 574–581 (2014).

183. Maybaum, J., Kott, M. G., Johnson, N. J., Ensminger, W. D. & Stetson, P. L. Analysis of bromodeoxyuridine incorporation into DNA: Comparison of gas chromatographic/mass spectrometric, CsCl gradient sedimentation, and specific radioactivity methods. *Anal. Biochem.* **161**, 164–171 (1987).

184. Salic, A. & Mitchison, T. J. A chemical method for fast and sensitive detection of DNA synthesis in vivo. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 2415–2420 (2008).

185. Vogel, W., Autenrieth, M. & Mehnert, K. Analysis of chromosome replication by a BrdU antibody technique. *Chromosoma* **98**, 335–341 (1989).

186. Jackson, D. & Cook, P. R. Analyzing DNA replication I: Labeling animals, tissues, and cells with bromodeoxyuridine (BrdU). *Cold Spring Harb. Protoc.* **3**, (2008).

187. Salic, A. & Mitchison, T. J. *A chemical method for fast and sensitive detection of DNA synthesis in vivo*. *PNAS* vol. 105 www.pnas.orgcgidoi10.1073pnas.0712168105 (2008).

188. Pierzyńska-Mach, A. *et al.* Subnuclear localization, rates and effectiveness of UVC-induced unscheduled DNA synthesis visualized by fluorescence

widefield, confocal and super-resolution microscopy. *Cell Cycle* **15**, 1156–1167 (2016).

189. Kaufman, E. R. & Davidson, R. L. Biological and biochemical effects of bromodeoxyuridine and deoxycytidine on Syrian hamster melanoma cells. *Somatic Cell Genet.* **4**, 587–601 (1978).

190. Rodríguez-Reyes, R. & Morales-Ramírez, P. Sister chromatid exchange induction and the course of DNA duplication, two mechanisms of sister chromatid exchange induction by ENU and the role of BrdU. *Mutagenesis* **18**, 65–72 (2003).

191. Kohlmeier, F., Maya-Mendoza, A. & Jackson, D. A. EdU induces DNA damage response and cell death in mESC in culture. *Chromosom. Res.* **21**, 87–100 (2013).

192. Stasolla, C. Alterations in pyrimidine nucleotide metabolism as an early signal during the execution of programmed cell death in tobacco BY-2 cells. *J. Exp. Bot.* **55**, 2513–2522 (2004).

193. Macallan, D. C. *et al.* Measurement of cell proliferation by labeling of DNA with stable isotope-labeled glucose: Studies in vitro, in animals, and in humans. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 708–713 (1998).

194. Zhang, J. J. *et al.* Analysis of global DNA methylation by hydrophilic interaction ultra high-pressure liquid chromatography tandem mass spectrometry. *Anal. Biochem.* **413**, 164–170 (2011).

195. Amouroux, R. *et al.* De novo DNA methylation drives 5hmC accumulation in mouse zygotes. *Nat. Cell Biol.* **18**, 225–233 (2016).

196. Kong, Z. *et al.* Simultaneous determination of ribonucleoside and deoxyribonucleoside triphosphates in biological samples by hydrophilic interaction liquid chromatography coupled with tandem mass spectrometry. *Nucleic Acids Res.* **46**, e66 (2018).

197. Matsuda, S. & Kasahara, T. Simultaneous and absolute quantification of nucleoside triphosphates using liquid chromatography-triple quadrupole tandem mass spectrometry. *Genes Environ.* **40**, (2018).

198. Mordehai, A. & Fjeldsted, J. *Agilent Jet Stream technology enhances ESI-MS sensitivity Agilent Jet Stream Thermal Gradient Focusing Technology*. www.agilent.com/chem/lcms.

199. *Agilent Technologies Agilent 6400 Series Triple Quadrupole LC/MS System Concepts Guide The Big Picture*.

200. Yost, R. A. Triple quadrupole mass spectrometry for direct mixture analysis and structure elucidation. *Analytical Chemistry* vol. 51 (1979).

201. Kriaucionis, S. & Heintz, N. The nuclear DNA base 5-hydroxymethylcytosine is present in purkinje neurons and the brain. *Science (80-.).* **324**, 929–930 (2009).

202. HPLC Separation Modes – Polarity, Phases, & Chromatography in HPLC | Waters. https://www.waters.com/waters/en_US/HPLC-Separation-Modes/nav.htm?cid=10049076&locale=en_US.

203. *Thermo Scientific HILIC Separations Technical Guide HILIC Separations A Practical Guide to HILIC Mechanisms, Method Development and Troubleshooting*.

204. The Effects of Changing Column Inner Diameter-Part 1.

https://phenomenex.blog/2017/07/18/effects-of-changing-column-id-part-1/.

205. Barrett, D. A. *et al. Characterization of a Range of Alkyl-Bonded Silica HPLC Stationary Phases: Chromatographic Behavior of Neutral, Acidic, and Basic Test Solutes*. *Journal of Chromatographic Science* vol. 34 https://academic.oup.com/chromsci/article/34/3/146/291371 (1996).

206. Difference between C8 and C18 Columns Used in HPLC System : Pharmaceutical Guidelines. https://www.pharmaguideline.com/2018/05/difference-between-c8-and-c18-columns.html.

207. Sartain, M. *The Agilent Metabolomics Dynamic MRM Database and Method Author Technical Overview*.

208. Lu, W., Bennett, B. D. & Rabinowitz, J. D. Analytical strategies for LC-MS-based targeted metabolomics. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **871**, 236–242 (2008).

209. Müller, C. A. *et al.* Capturing the dynamics of genome replication on individual ultra-long nanopore sequence reads. *Nat. Methods* **16**, 429–436 (2019).

210. Visnes, T. *et al.* Small-molecule inhibitor of OGG1 suppresses proinflammatory gene expression and inflammation. *Science (80-.).* **362**, 834–839 (2018).

211. Stone, P., Glauner, T., Kuhlmann, F., Schlabach, T. & Miller, K. *New Dynamic MRM Mode Improves Data Quality and Triple Quad Quantification in Complex Analyses*.

212. Zhang, J. J. *et al.* Analysis of global DNA methylation by hydrophilic interaction ultra high-pressure liquid chromatography tandem mass spectrometry. *Anal. Biochem.* **413**, 164–170 (2011).

213. Mičová, K., Friedecký, D. & Adam, T. Mass Spectrometry for the Sensitive Analysis of Intracellular Nucleotides and Analogues. in *Mass Spectrometry* (InTech, 2017). doi:10.5772/68073.

214. Wilson, P. M. *et al.* A novel fluorescence-based assay for the rapid detection and quantification of cellular deoxyribonucleoside triphosphates. *Nucleic Acids Res.* **39**, (2011).

215. Jang, C., Chen, L. & Rabinowitz, J. D. Metabolomics and Isotope Tracing. *Cell* vol. 173 822–837 (2018).

216. Borowski, L. S., Dziembowski, A., Hejnowicz, M. S., Stepien, P. P. & Szczesny, R. J. Human mitochondrial RNA decay mediated by PNPase-hSuv3 complex takes place in distinct foci. *Nucleic Acids Res.* **41**, 1223–1240 (2013).

217. Sharma, S. *et al.* Proofreading deficiency in mitochondrial DNA polymerase does not affect total dNTP pools in mouse embryos. *Nature Metabolism* vol. 2 (2020).

218. Bustamante, S., Gilchrist, R. B. & Richani, D. A sensitive method for the separation and quantification of low-level adenine nucleotides using porous graphitic carbon-based liquid chromatography and tandem mass spectrometry. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **1061**–**1062**, 445–451 (2017).

219. Pabst, M. *et al.* Nucleotide and nucleotide sugar analysis by liquid chromatography- electrospray ionization-mass spectrometry on surface-conditioned porous graphitic carbon. *Anal. Chem.* **82**, 9782–9788 (2010).

220. Vuckovic, D. Current trends and challenges in sample preparation for global metabolomics using liquid chromatography-mass spectrometry. *Analytical and Bioanalytical Chemistry* vol. 403 1523–1548 (2012).

221. Gil, A. *et al.* The degradation of nucleotide triphosphates extracted under boiling ethanol conditions is prevented by the yeast cellular matrix. *Metabolomics* **13**, 1–10 (2017).

222. Walsby-Tickle, J. *et al.* Anion-exchange chromatography mass spectrometry provides extensive coverage of primary metabolic pathways revealing altered metabolism in IDH1 mutant cells. *Commun. Biol.* **3**, (2020).

223. Wilson, M. J., Shivapurkar, N. & Poirier, L. A. Hypomethylation of hepatic nuclear DNA in rats fed with a carcinogenic methyl-deficient diet. *Biochem. J.* **218**, 987–90 (1984).

224. Del Val, I. J., Kyriakopoulos, S., Polizzi, K. M. & Kontoravdi, C. An optimized method for extraction and quantification of nucleotides and nucleotide sugars from mammalian cells. *Anal. Biochem.* **443**, (2013).

225. (No Title). https://www.agilent.com/cs/library/technicaloverviews/public/5991-9271EN_HILIC_method_development_TechOverview.pdf.

226. HolčapekHolˇHolčapek, M. *et al.* Effects of ion-pairing reagents on the electrospray signal suppression of sulphonated dyes and intermediates. *J. MASS Spectrom. J. Mass Spectrom* **39**, 43–50 (2004).

227. A3051150X020 | Agilent. https://www.agilent.com/store/en_US/Prod-A3051150X020/A3051150X020.

228. (No Title). https://www.agilent.com/cs/library/flyers/public/5991-8547EN_hilic_columns_flyer.pdf.

229. *Storage recommendations*. www.agilent.com/chem/infinitylab-additive.

230. Strzelecka, D., Chmielinski, S., Bednarek, S., Jemielity, J. & Kowalska, J. Analysis of mononucleotides by tandem mass spectrometry: Investigation of fragmentation pathways for phosphate- and ribose-modified nucleotide analogues. *Sci. Rep.* **7**, (2017).

231. Curtis, M. *et al.* Direct analysis in real time (DART) mass spectrometry of nucleotides and nucleosides: Elucidation of a novel fragment [C5H5O] + and its in-source adducts. *J. Am. Soc. Mass Spectrom.* **21**, 1371–1381 (2010).

232. Takkis, K. *et al.* Signal Enhancement in the HPLC-ESI-MS/MS analysis of spironolactone and its metabolites using HFIP and NH4F as eluent additives. *Anal. Bioanal. Chem.* **409**, 3145–3151 (2017).

233. (26) Why the signal-to-noise ratio determines the limit of detection and LC method performance | LinkedIn. https://www.linkedin.com/pulse/why-signal-to-noise-ratio-determines-limit-detection-lc-heidorn/.

234. How low can you go? A practical guide to the Limit of Detection – Andy Connelly. https://andyjconnelly.wordpress.com/2017/02/17/how-low-can-you-go-a-practical-guide-to-the-limit-of-detection/.

235. Mirza, U. A. & Chait, B. T. Effects of Anions on the Positive Ion Electrospray Ionization Mass Spectra of Peptides and Proteins. *Anal. Chem.* **66**, 2898–2904 (1994).

236. Reinhold, W. C. *et al.* CellMiner: A web-based suite of genomic and pharmacologic tools to explore transcript and drug patterns in the NCI-60 cell line set. *Cancer Res.* **72**, 3499–3511 (2012).

237. Shankavaram, U. T. *et al.* CellMiner: A relational database and query tool for the NCI-60 cancer cell lines. *BMC Genomics* **10**, 277 (2009).

238. Kaighn, M. E., Narayan, K. S., Ohnuki, Y., Lechner, J. F. & Jones, L. W. Establishment and characterization of a human prostatic carcinoma cell line (PC-3). *Invest. Urol.* **17**, (1979).

239. NCI Thesaurus. https://ncit.nci.nih.gov/ncitbrowser/ConceptReport.jsp?dictionary=NCI_Thesaurus&ns=ncit&code=C117136.

240. Cailleau, R., Young, R., Olivé, M. & Reeves, W. J. Breast tumor cell lines from pleural effusions. *J. Natl. Cancer Inst.* **53**, (1974).

241. Liu, Y., Beyer, A. & Aebersold, R. Leading Edge Review On the Dependency of Cellular Protein Levels on mRNA Abundance. *Cell* **165**, 535–550 (2016).

242. Lichtinghagen, R. *et al.* Different mRNA and protein expression of matrix metalloproteinases 2 and 9 and tissue inhibitor of metalloproteinases 1 in benign and malignant prostate tissue. *Eur. Urol.* **42**, 398–406 (2002).

243. Chen, G. *et al.* Discordant protein and mRNA expression in lung adenocarcinomas. *Mol. Cell. Proteomics* **1**, 304–313 (2002).

244. Ehrlich, M. & Wang, R. Y. H. 5-Methylcytosine in eukaryotic DNA. *Science* vol. 212 1350–1357 (1981).

245. Yoo, H., Antoniewicz, M. R., Stephanopoulos, G. & Kelleher, J. K. Quantifying reductive carboxylation flux of glutamine to lipid in a brown adipocyte cell line. *J. Biol. Chem.* **283**, 20621–20627 (2008).

246. Corson, T. W., Cavga, H., Aberle, N. & Crews, C. M. Triptolide Directly Inhibits dCTP Pyrophosphatase. *ChemBioChem* **12**, 1767–1773 (2011).

247. Kato, K. *et al.* Crystal structure of Enpp1, an extracellular glycoprotein involved in bone mineralization and insulin signaling. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 16876–16881 (2012).

248. KOSHLAND, D. E. & SPRINGHORN, S. S. Mechanism of action of 5'-nucleotidase. *J. Biol. Chem.* **221**, 469–476 (1956).

249. Sass, J. O. & Skladal, D. Plasma concentrations and renal clearance of orotic acid in argininosuccinic acid synthetase deficiency. *Pediatr. Nephrol.* **13**, 912–916 (1999).

250. Stumvoll, M., Perriello, G., Meyer, C. & Gerich, J. Role of glutamine in human carbohydrate metabolism in kidney and other tissues. *Kidney Int.* **55**, 778–792 (1999).

251. Rabinovich, S. *et al.* Diversion of aspartate in ASS1-deficient tumours fosters de novo pyrimidine synthesis. *Nature* **527**, 379–383 (2015).

252. Hu, H. *et al.* Gene Expression and Methylation Analyses Suggest DCTD as a Prognostic Factor in Malignant Glioma. *Sci. Rep.* **7**, 1–9 (2017).

253. Ross, J. mRNA stability in mammalian cells. *Microbiological Reviews* vol. 59 423–450 (1995).

254. Defoiche, J. *et al.* Measurement of Ribosomal RNA Turnover In Vivo by Use of Deuterium-Labeled Glucose. *Clin. Chem.* **55**, 1824–1833 (2009).

255. Macromolecular Components of E. coli and HeLa Cells | Thermo Fisher Scientific - UK. https://www.thermofisher.com/uk/en/home/references/ambion-tech-support/rna-tools-and-calculators/macromolecular-components-of-e.html.

256. How much RNA does a typical mammalian cell contain? - QIAGEN. https://www.qiagen.com/gb/resources/faq?id=06a192c2-e72d-42e8-9b40-3171e1eb4cb8&lang=en.

257. Lodish, H. *et al.* Processing of rRNA and tRNA. (2000).

258. Humphreys, B. D. Cutting to the chase: Taking the pulse of label-retaining cells in kidney. *American Journal of Physiology - Renal Physiology* vol. 308 F29–F30 (2015).

259. Tisseur, M., Kwapisz, M. & Morillon, A. Pervasive transcription - Lessons from yeast. *Biochimie* vol. 93 1889–1896 (2011).

260. Ellims, P. H., Kao, A. Y. & Chabner, B. A. Kinetic behaviour and allosteric regulation of human deoxycytidylate deaminase derived from leukemic cells. *Mol. Cell. Biochem.* **57**, 185–190 (1983).

261. Lo, C.-A. *et al.* Quantification of Protein Levels in Single Living Cells In Brief. *CellReports* **13**, 2634–2644 (2015).

262. Sakamoto, K. M. *et al.* Protacs: Chimeric molecules that target proteins to the Skp1-Cullin-F box complex for ubiquitination and degradation. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 8554–8559 (2001).

263. Giovannetti, E. *et al.* Correlation between cytidine deaminase genotype and gemcitabine deamination in blood samples. in *Nucleosides, Nucleotides and Nucleic Acids* vol. 27 720–725 (Nucleosides Nucleotides Nucleic Acids, 2008).

264. Grindey, G. B., Hertel, L. W. & Plunkett, W. Cytotoxicity and antitumor activity of 2'2'difluorodeoxycytidine (gemcitabine). *Cancer Invest.* **8**, 313–314 (1990).

265. Cusack, M. *et al.* Distinct contributions of DNA methylation and histone acetylation to the genomic occupancy of transcription factors. *Genome Res.* **30**, 1393–1406 (2020).

266. Motnenko, A. *et al.* Identification of UHRF2 as a novel DNA interstrand crosslink sensor protein. *PLoS Genet.* **14**, (2018).

267. Pabst, M. *et al.* Nucleotide and nucleotide sugar analysis by liquid chromatography- electrospray ionization-mass spectrometry on surface-conditioned porous graphitic carbon. *Anal. Chem.* **82**, 9782–9788 (2010).

268. Tabata, S. *et al.* Thymidine Catabolism as a Metabolic Strategy for Cancer Survival. *Cell Rep.* **19**, (2017).

Appendix A

Table of Contents

Investigating the roles of CDA and DCTD in cancer nucleotide metabolism i

Relating to Chapter 3, “Setting up a workflow for nucleoside analysis” 220

Relating to Chapter 4, “Setting up a workflow for nucleotide analysis” 225

Relating to Chapter 5 “Cell-based tools for investigating nucleotide metabolism” 227

Relating to Chapter 3, "Setting up a workflow for nucleoside analysis"

Table 2 A list of each nucleoside transition monitored by dMRM with collision energies and retention times.

Compound	Ion Adduct	Precursor Ion (*m/z*)	Product Ion (*m/z*)	Collision Energy (V)
cadC	cadC H^+	272	156	12
	cadC K^+	310	194	12
	cadC Na^+	294	178	12
rC	rC H^+	244	112	10
	rC K^+	282	150	10
	rC Na^+	266	134	10
dC	dC H^+	228	112	10
	dC K^+	266	150	10
	dC Na^+	250	134	10
rU	rU H^+	245	113	10
	rU K^+	283	151	10
	rU Na^+	267	135	10
hmdC	hmdC H^+	258	142	12
	hmdC K^+	296	180	12
	hmdC Na^+	280	164	12

dU	dU H^+	229	113	10
	dU K^+	267	151	10
	dU Na^+	251	135	10
hmdU	hmdU H^+	259	143	10
	hmdU K^+	297	181	10
	hmdU Na^+	281	165	10
rG	rG H^+	284	152	10
	rG K^+	322	190	10
	rG Na^+	264	174	10
mdC	mdC H^+	242	126	10
	mdC Na^+	264	148	10
	mdC K^+	280	164	10
dG	dG H^+	268	152	10
	dG K^+	306	190	10
	dG Na^+	290	174	10
fdC	fdC H^+	256	140	8
	fdC K^+	294	178	8
	fdC Na^+	278	162	8
8-oxodG	8-oxo-dG K^+	322	206	10

	8-oxo-dG Na^+	306	190	10
	8-oxodG H^+	284	168	10
dT	dT H^+	243	127	10
	dT K^+	281	165	10
	dT Na^+	265	149	10
EdU	EdU H^+	253	137	10
	EdU Na^+	275	159	10
rA	rA H^+	268	136	8
	rA K^+	306	176	8
	rA Na^+	290	158	8
BrdU	BrdU Na^+ 79	329	213	10
	BrdU Na^+ 81	331	215	10
dA	dA H^+	252	136	10
	dA K^+	274	158	10
m6ra	m6A H^+	282	150	20
	m6A K^+	320	188	20
	m6A Na^+	304	172	20
m6dA	m6dA H^+	266	150	20
	m6dA K^+	304	188	20

	m6dA Na	288	172	20

Table 3 Table of nucleoside retention times, LODs and LLOQs for each nucleoside analysed.

Nucleoside	**Retention time**	**LOD (amol)**	**LLOQ (amol)**
cadC	1.76	15800	15800
rC	2.46	500	500
rU	3.6	15800	50000
dC	3.69	5000	15800
hmdC	4.39	158	500
dU	5.53	15800	50000
hmdU	6.12	15800	15800
3mdC	6.74	158	158
rI	7.97	15800	50000
rG	8.46	15800	15800
mdC	8.9	500	500
dI	9.11	50000	50000
dG	9.57	50000	50000
fdC	10.5	5000	5000
dT	11	50000	50000
8-oxo-dG	1.1	5000	5000

BrdU	12.5	50000	50000
rA	12.9	15800	15800
dA	13.6	15800	15800
mrA	18.2	158	500
mdA	18.3	158	158

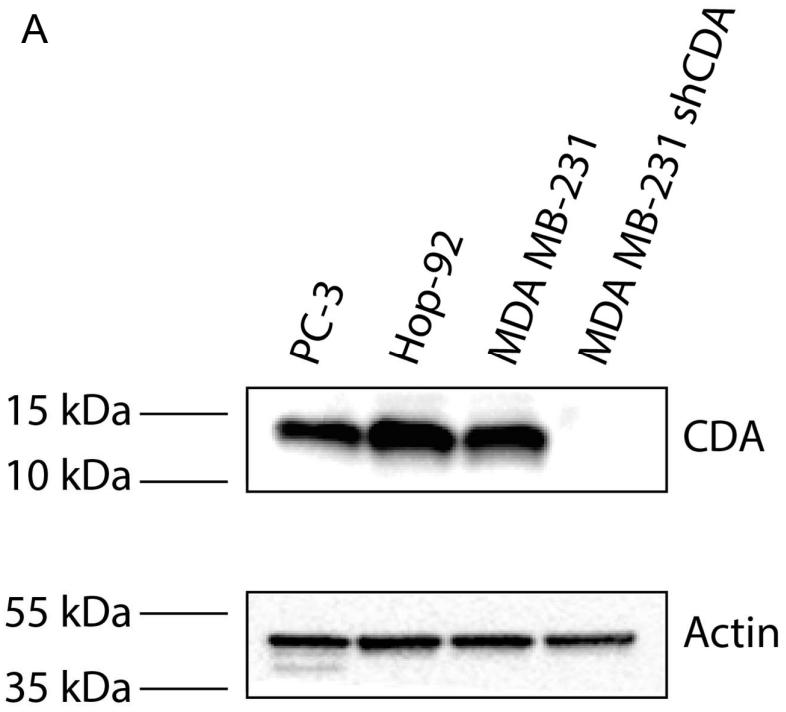
Relating to Chapter 4, "Setting up a workflow for nucleotide analysis"

Table 4 List of nucleotides retention times, their MS transitions and their LLOQs.

Nucleotide	Retention Time (min)	Best transition	*m/z* Parent Ion	*m/z* Target Ion	CE (V)	CA (V)	LLOQ (pmol)
ADP	7.4	ADP H^+	428	136	20	5	0.0316
AMP	6.4	AMP H^+	348	136	20	5	0.0316
ATP	8.1	ATP H^+	508	136	45	5	0.0316
CDP	8.2	CDP H^+	404	112	20	5	0.0316
CMP	7.8	CMP H^+	324	112	20	5	0.0316
CTP	8.1	CTP H^+	484	112	20	5	0.0316
dADP	7.6	dADP H^+	412	136	20	5	0.0316
dATP	8.1	dATP H^+	492	136	20	5	0.0316
dCDP	8.1	dCDP H^+	388	112	20	5	0.0316
dCMP	7.7	dCMP H^+	308	112	20	5	0.0316
dCTP	8.7	dCTP H^+	468	112	45	5	0.0316
dGDP	8.2	dGDP H^+	428	152	20	5	0.0316
dGMP	7.6	dGMP H^+	348	152	20	5	0.0316

dGTP	8.7	dGTP H^+	508	152	45	5	0.0316
dTDP	6.4	dTDP Na^+	425	81	30	5	0.0316
dTMP	4.8	dTMP H^+	323	81	30	5	0.0316
dTTP	7.4	dTTP Na^+	505	81	30	5	0.0316
dUDP	6.7	dUDP Na^+	411	81	30	5	0.316
dUMP	5.6	dUMP Na^+	331	81	20	5	0.0316
dUTP	7.5	dUTP H^+	469	81	30	5	0.0316
GDP	8.3	GDP H^+	444	152	20	5	0.0316
GMP	7.7	GMP H^+	364	152	20	5	0.0316
GTP	8.6	GTP H^+	524	152	20	5	1
UDP	6.9	UDP H^+	405	97	25	5	0.0316
UMP	7	UMP H^+	325	97	30	5	0.0316
UTP	6.8	UTP H^+	485	97	30	5	0.0316

Relating to Chapter 5 "Cell-based tools for investigating nucleotide metabolism"



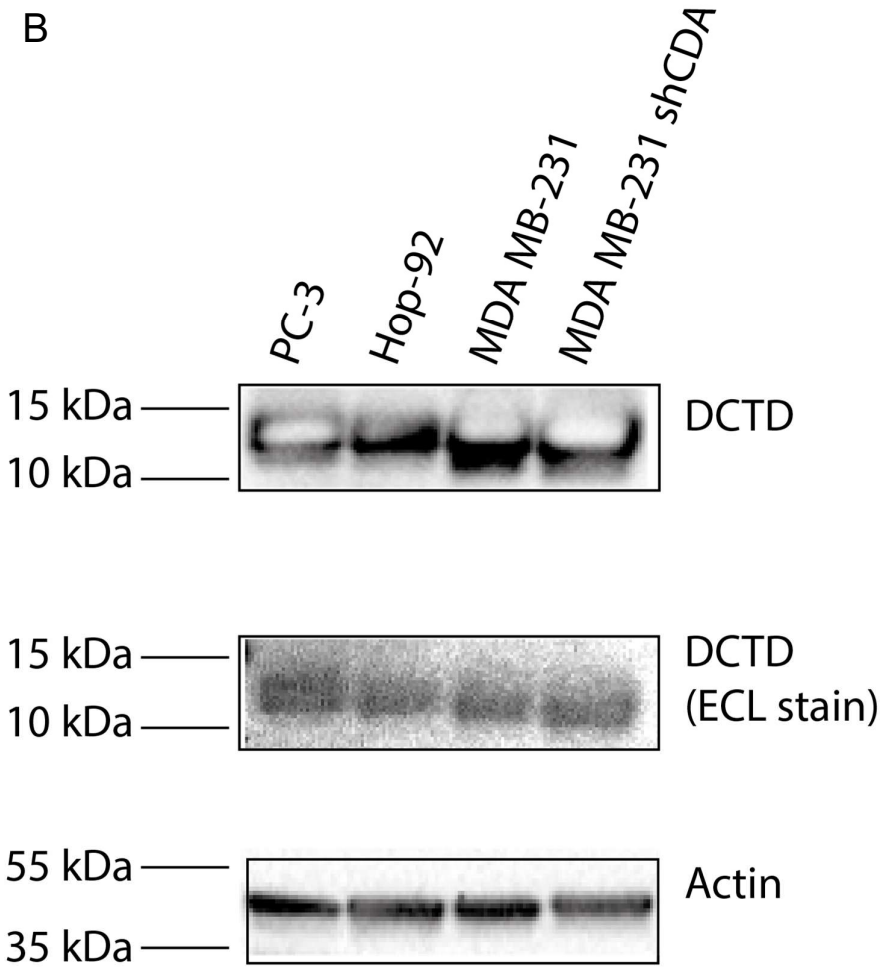


Figure 0.1 PC-3, MDA MB-231 and HOP-92 have detectable levels of CDA and DCTD protein, second replicate. *Immunoblot of CDA and DCTD for all cell lines, with β-actin blot showing loading distribution.* ***A,*** *CDA protein levels* ***B,*** *DCTD protein levels, also showing from DCTD ECL stain.*

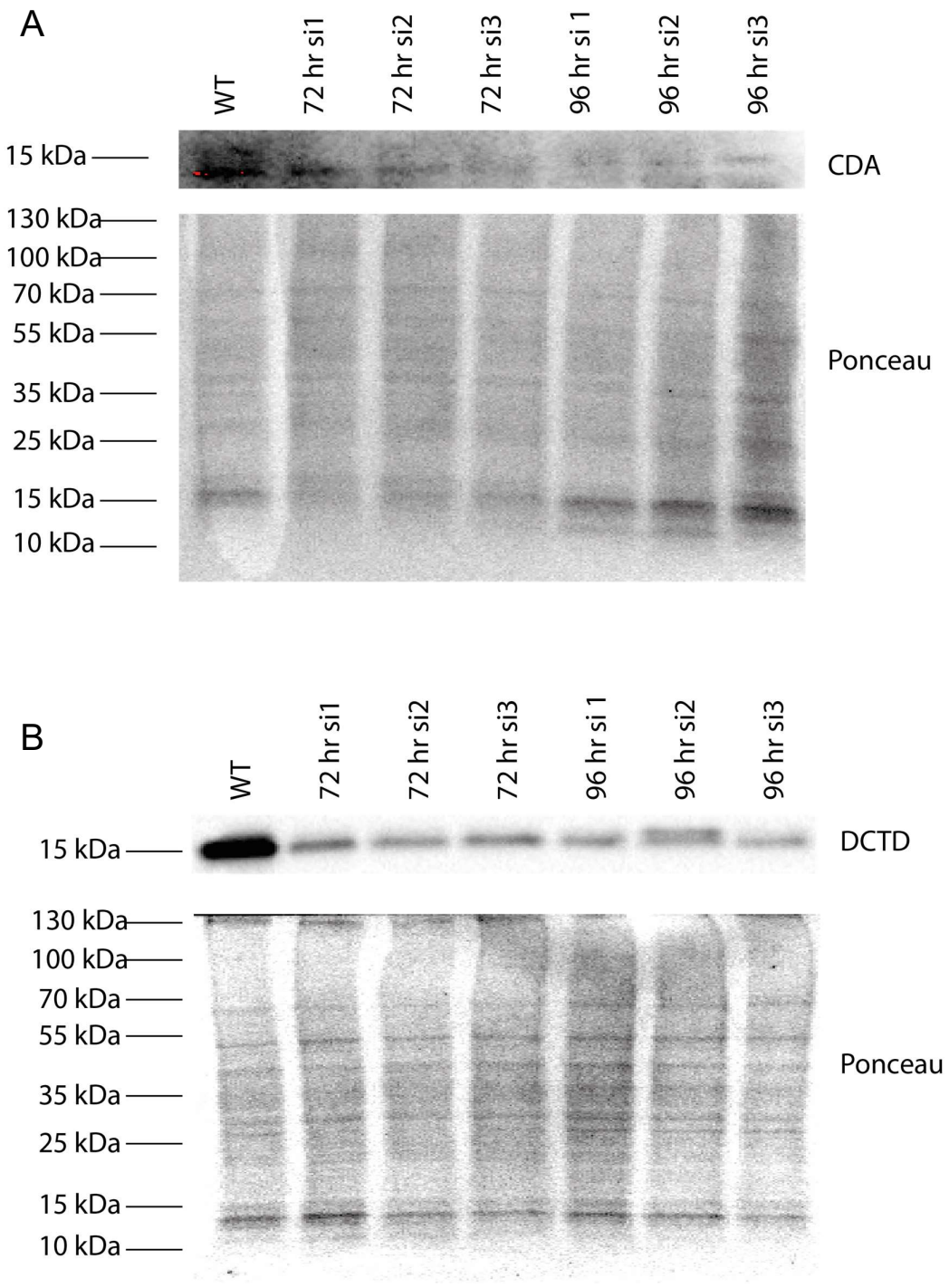


Figure 0.2 CDA and DCTD KD is stable over time in the MDA MB-231 cell line, second replicate. *Immunoblot of CDA and DCTD for MDA MB-231, with Ponceau stain showing loading distribution.* ***A****, CDA KD compared to scrambled siRNA from 72 hrs to 96 hrs post exposure to siRNAs****; B****, DCTD KD compared to scrambled siRNA from 72 hrs to 96 hrs post exposure to siRNAs****.*

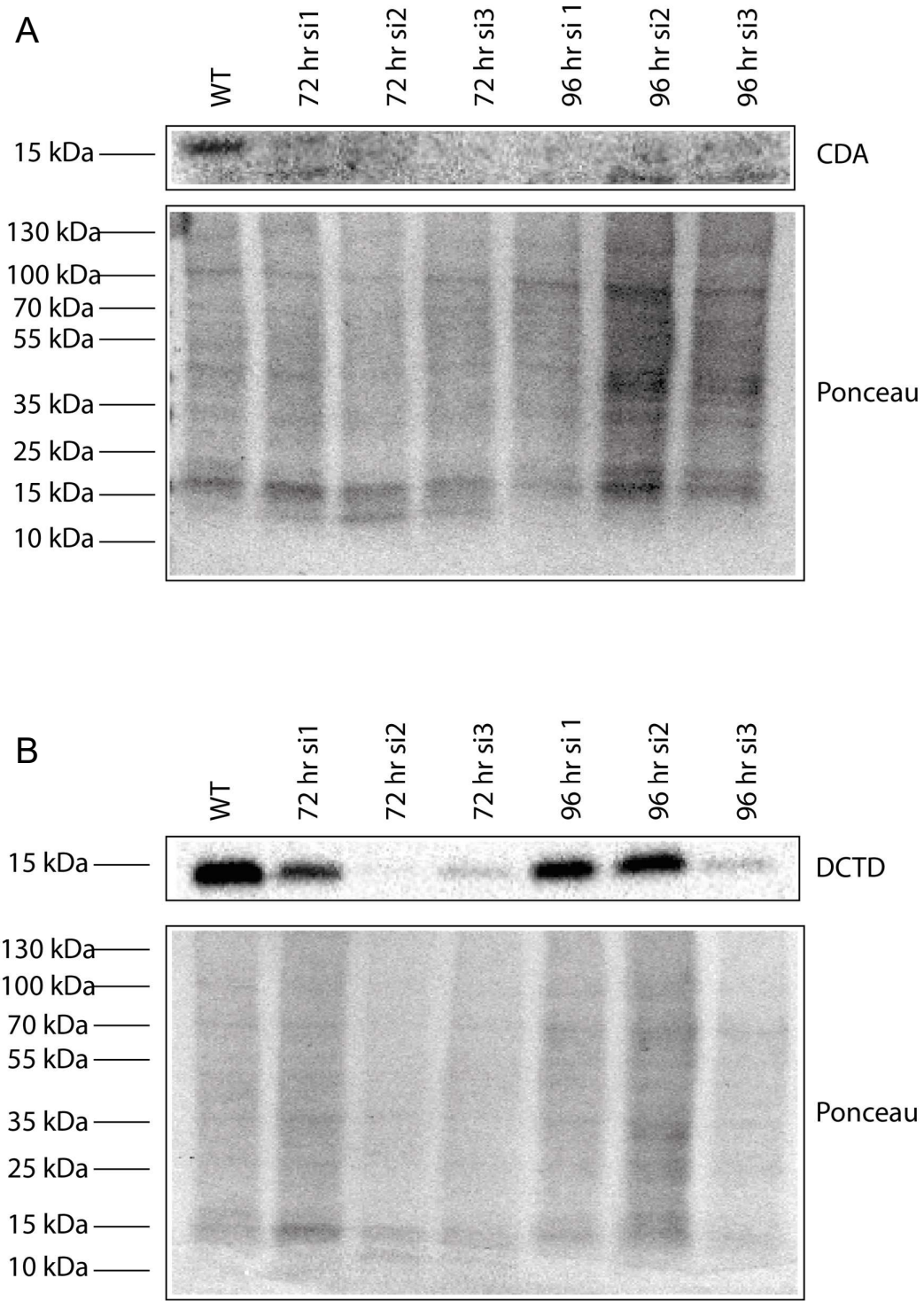


Figure 0.3 CDA and DCTD KD is stable over time in the PC-3 cell line, second replicate. *Immunoblot of CDA and DCTD for MDA MB-231, with Ponceau stain showing loading distribution.* ***A****, CDA KD compared to scrambled siRNA from 72 hrs to 96 hrs post exposure to siRNAs****; B****, DCTD KD compared to scrambled siRNA from 72 hrs to 96 hrs post exposure to siRNAs****.*