



# **An investigation into the role of Ap<sub>4</sub>A and Ap<sub>4</sub>Ns in the inhibition of the initiation phase of DNA replication**

By Freya Ferguson

PhD in Biomedical and Life Sciences

Lancaster University  
Faculty of Health and Medicine  
Biomedical Life Sciences

March 2023



## **Abstract**

### **An investigation into the role of Ap<sub>4</sub>A and Ap<sub>4</sub>Ns in the inhibition of the initiation phase of DNA replication**

By Freya Ferguson

Precise DNA replication is essential for the maintenance of a healthy genome. Although DNA replication has a very high fidelity, mutations can be introduced during this process, and the likelihood of mutation increases when a cell is under stress. Therefore, several signalling mechanisms exist, such as the ATM and ATR-dependent signalling in the DNA damage response, which stall DNA replication and promote the repair of damage. Diadenosine tetraphosphate (Ap<sub>4</sub>A) levels increase in response to cell stress, where it inhibits the initiation phase of DNA replication at micromolar levels. However, the mechanistic basis for its inhibitory activity has not yet been determined. In addition, there are other dinucleoside tetraphosphates (Ap<sub>4</sub>Ns) that have extracellular signalling roles, yet their synthesis and intracellular function have not been established. Here we present evidence suggesting that Ap<sub>4</sub>Ns can be synthesised by UBE1 in a biologically relevant manner. We have developed a technique for the extraction and detection of Ap<sub>4</sub>A, Ap<sub>4</sub>C, Ap<sub>4</sub>G and Ap<sub>4</sub>U by triple quadrupole mass spectrometry in extracts from a mammalian KBM7 cell line with a disrupted *Nudt2* gene (NuKO). Furthermore, we have applied this technique to investigate the effect of DNA damaging agents and DNA stressing agents on Ap<sub>4</sub>N levels in NuKO cells. Finally, the inhibitory effect of each Ap<sub>4</sub>N in a cell-free DNA replication system containing late G1 nuclei and S-phase cytosolic extract demonstrates that Ap<sub>4</sub>Ns have the capacity to inhibit the initiation of DNA replication. Further analysis suggests that Ap<sub>4</sub>Ns reduce initiation by reducing the chromatin localisation of the key replication proteins Mcm2 and PCNA. Together, these data suggest that both Ap<sub>4</sub>A and Ap<sub>4</sub>Ns may regulate the DNA replication process and that they could represent an additional regulatory mechanism to prevent untimely entry into S-phase during genotoxic stress.

## Table of Contents

Abstract.....	1
Table of Contents.....	2
List of Tables .....	12
List of Figures .....	13
Abbreviations.....	16
Acknowledgements .....	22
Author's declaration .....	23
Chapter 1: General Introduction .....	24
1.1 Introduction to Ap <sub>4</sub> A.....	25
1.1.1 Ap <sub>4</sub> A can be synthesised through different mechanisms.....	25
1.1.2 Hydrolysis of Ap <sub>4</sub> A is required for maintaining the appropriate Ap <sub>4</sub> A level .	30
1.1.3 Ap <sub>4</sub> A is produced in response to different types of stress .....	31
1.2 Ap <sub>4</sub> A has been implicated in several different signalling mechanisms .....	32
1.2.1 Ap <sub>4</sub> A and other Ap <sub>4</sub> Ns form protective caps on RNA transcripts.....	32
1.2.2 Ap <sub>4</sub> A modulates histidine triad nucleotide-binding protein 1 (HINT1) signalling .....	35
1.2.3 Ap <sub>4</sub> A competes with cyclic GMP–AMP (cGAMP) to interact with STING.....	37
1.2.4 Ap <sub>4</sub> A modulates the production of xanthosine 5'-monophosphate (XMP) from inosine 5'-monophosphate (IMP) in GTP biosynthesis.....	39
1.2.5 Biofilm formation.....	42
1.2.6 DNA replication.....	43
1.2.7 Ap <sub>4</sub> A often, but not always, exerts its effect through binding ATP binding sites .....	44
1.3 Other Ap <sub>4</sub> Ns also have known signalling roles .....	45

1.4 Regulation of the cell-cycle.....	46
1.5 Regulation of DNA replication .....	48
1.5.1 Formation of the pre-replication complex .....	48
1.5.2 Formation of the pre-initiation (pre-IC) complex.....	50
1.5.3 Formation of the replisome.....	53
1.6 The DNA damage response: stalled replication forks and double strand breaks (DSBs) .....	54
1.6.1 The ATR-mediated response to stalled replication forks .....	55
1.6.2 The ATM-mediated response to double strand breaks.....	57
1.6.3 Homologous recombination is initiated by CtIP, a phosphorylation target of ATM .....	60
1.7 Conclusion.....	64
Chapter 2: Materials and Methods.....	66
2.1 Synthesis and validation of Ap <sub>4</sub> Ns and Ap <sub>4</sub> dNs .....	67
2.1.1 Chemical synthesis of Ap <sub>4</sub> Ns.....	67
2.1.2 Biological synthesis of Ap <sub>4</sub> Ns .....	68
2.1.3 Validation of synthesised Ap <sub>4</sub> N identity by High Resolution Mass Spectrometry (HRMS) .....	68
2.2 Determining synthesised Ap <sub>4</sub> N concentration .....	68
2.2.1 Hydrolysis of Ap <sub>4</sub> A and Ap <sub>4</sub> N synthesis products.....	68
2.2.2 Determining the extinction coefficient for each Ap <sub>4</sub> N .....	69
2.3 Cell culture .....	69
2.3.1 Adherent cells: 3T3 and HeLa cells .....	69
2.3.2 Suspension cells – KBM7 cells.....	69
2.4 Extraction of metabolites from KBM7 and NuKO cells.....	70
2.5 HPLC analysis of mammalian cell extract .....	70



2.6 Clean up KBM7 mammalian cell extract.....	71
2.6.1 Graphite column clean-up.....	71
2.6.2 Strong cation exchange clean-up.....	71
2.7 Triple quadrupole mass spectrometry analysis of mammalian cell extract .....	72
2.7.1 Determining multiple reaction monitoring (MRM) transitions for Ap <sub>4</sub> N standards .....	72
2.7.2 Analysis of mammalian cell extract by LC-MS .....	73
2.7.3 LC-MS measurement of the effect of DNA damaging agents on DNA replication .....	73
2.8 Determining mammalian cell proliferation rate and cell cycle kinetics .....	73
2.8.1 Determining proliferation rate .....	73
2.9 Flow cytometry .....	74
2.9.1 Measurement of cell cycle kinetics .....	74
2.9.2 Measurement of apoptosis.....	75
2.10 Analysis of the $\gamma$ H2AX response to cell stress using Western blotting (WB) .....	75
2.10.1 Inducing $\gamma$ H2AX stress in KBM7 and NuKO cells.....	75
2.10.2 Analysis by Western blotting.....	76
2.11 Cell synchronisation.....	77
2.11.1 Synchronisation of 3T3 cells in mid- and late-G1 .....	77
2.11.2 Synchronisation of S-phase HeLa cells.....	77
2.11.3 Harvesting synchronised nuclei and extract.....	77
2.12 Cell-free replication assays .....	78
2.12.1 Cell-free replication .....	78
2.12.2 Preparation of WB to determine nuclear replication protein level .....	79
Chapter 3: Synthesis and purification of Ap <sub>4</sub> A and Ap <sub>4</sub> N dinucleotide tetraphosphates .....	80

3.1 Introduction to the synthesis of Ap <sub>4</sub> A and Ap <sub>4</sub> Ns.....	81
3.1.1 Ap <sub>4</sub> A structure and synthesis.....	81
3.1.2 Ap <sub>4</sub> A signalling mechanisms .....	82
3.1.3 Chapter Aims.....	83
3.2 Preparation of standards for HPLC .....	83
3.3 Optimisation of Ap <sub>4</sub> A synthesis .....	85
3.3.1 Effect of increased ATP in the reaction mixture.....	85
3.3.2 Effect of HEPES and EDC on increasing Ap <sub>4</sub> A yield.....	87
3.3.3 Effect of temperature on Ap <sub>4</sub> A yield .....	90
3.4 Synthesis of other dinucleotide tetraphosphates .....	92
3.4.1 Changes to Ap <sub>4</sub> A structure.....	92
3.4.2 Synthesis and purification of Ap <sub>4</sub> Ns .....	94
3.5 Validation of chemically synthesised Ap <sub>4</sub> Ns.....	97
3.5.1 Validation through analysis of Ap <sub>4</sub> A hydrolase products.....	97
3.5.2 Validation of Ap <sub>4</sub> Ns by HRMS .....	104
3.6 Determining concentration of synthesised Ap <sub>4</sub> Ns .....	108
3.7 Biological synthesis of Ap <sub>4</sub> A and Ap <sub>4</sub> Ns.....	109
3.8 Chapter Discussion.....	114
3.8.1 Ap <sub>4</sub> A and Ap <sub>4</sub> Ns can be analysed and purified from NTPs and AMP by HPLC .....	114
3.8.2 EDC as a coupling agent in Ap <sub>4</sub> A synthesis.....	115
3.8.3 Analysis of Ap <sub>4</sub> N hydrolysis products as a method of validating Ap <sub>4</sub> N identity .....	116
3.8.4 HRMS as a method of validating Ap <sub>4</sub> N identity.....	116
3.8.5 Ub and Ub-like activating enzymes as a source of Ap <sub>4</sub> Ns .....	117
3.9 Conclusion.....	118

Chapter 4: Identification of intracellular Ap <sub>4</sub> Ns in a mammalian cell line .....	119
4.1 Introduction .....	120
4.1.1 Ap <sub>4</sub> Ns exist intracellularly in different domains of life and are synthesised in response to cell stress.....	120
4.1.2 Ap <sub>4</sub> Ns in mammalian cells .....	120
4.1.3 Ub-activating enzyme and the synthesis of Ap <sub>4</sub> Ns.....	121
4.1.4 Chapter Aims.....	121
4.2 Analysis of metabolites by HPLC.....	122
4.2.1 FastAP Alkaline phosphatase treatment has no effect on Ap <sub>4</sub> A .....	122
4.2.2 Extraction of metabolites from mammalian cells and analysis by HPLC .....	125
4.3 Validation of Ap <sub>4</sub> A detection and analysis by LC-MS.....	130
4.3.1 Determining MRM transitions for Ap <sub>4</sub> N standards .....	130
4.4 KBM7 cell extracts .....	135
4.4.1 Investigating KBM7 cell extracts for the presence of Ap <sub>4</sub> Ns .....	135
4.4.2 Clean-up of the KBM7 extract using preparative chromatography approaches.....	138
4.5 NuKO ( <i>NUDT2</i> knock out) cell extracts.....	143
4.5.1 Identification of Ap <sub>4</sub> Ns in NuKO cells .....	143
4.5.2 Identification of Ap <sub>4</sub> dNs in NuKO cells .....	149
4.6 Chapter discussion.....	153
4.6.1 Ap <sub>4</sub> Ns can be harvested from a mammalian cell line by chloroform extraction.....	153
4.6.2 Identification of Ap <sub>4</sub> N candidates in NuKO extracts by HPLC .....	154
4.6.3 Identification of a candidate Ap <sub>4</sub> A peak in KBM7 cells by LC-MS .....	154
4.6.4 Ap <sub>4</sub> Ns and some Ap <sub>4</sub> dNs are detectable in NuKO cells.....	156
4.7 Conclusion.....	156

Chapter 5: Investigating the stress response in cells with intrinsically high Ap <sub>4</sub> N levels compared with their parental cell line .....	157
5.1 Chapter introduction .....	158
5.1.1 Ap <sub>4</sub> A and the cell cycle.....	158
5.1.2 Ap <sub>4</sub> A has been implicated in the cellular response to stress .....	159
5.1.3 Ap <sub>4</sub> A has been linked to apoptosis.....	159
5.1.4 Chapter Aims.....	160
5.2 Effect of increased Ap <sub>4</sub> N levels on cell duplication and cell cycle dynamics.....	161
5.2.1 Comparison of the proliferation rates of a high vs low Ap <sub>4</sub> N-containing cell line .....	161
5.2.2 Differences in cell cycle kinetics .....	162
5.3 Comparison of the KBM7 and NuKO response to stress .....	163
5.3.1 The $\gamma$ H2AX response to stress.....	163
5.3.2 The apoptotic response to stress .....	165
5.4 Investigating the Ap <sub>4</sub> N response to stress in NuKO cells .....	166
5.4.1 Determining the variation in area under peak readings taken by the LC-MS .....	166
5.4.2 The Ap <sub>4</sub> N response to hydroxyurea in NuKO cell extracts .....	168
5.5 The effect of DNA damaging agents on Ap <sub>4</sub> N levels.....	174
5.6 Chapter Discussion.....	183
5.6.1 Cells with intrinsically higher Ap <sub>4</sub> N levels have a slower doubling time.....	183
5.6.2 NuKO cells show similar amounts of time in each phase of the cell cycle but a prolonged transition between G1 and S-phase.....	184
5.6.3 High Ap <sub>4</sub> N-containing NuKO cells show a reduced response to replication stress.....	184
5.6.4 NuKO cells are more susceptible to apoptosis after HU-induced cell stress .....	185

5.6.5 Stressing agents may increase intracellular Ap <sub>4</sub> N levels.....	186
5.6.6 ADP levels are reduced slightly after HU and GEM treatment.....	186
5.6.7 ATP levels are increased after treatment with HU and GEM .....	187
5.6.8 ATP and ADP levels are increased after treatment with MMC and DOX....	188
5.6.9 The standardisation of Ap <sub>4</sub> N levels to internal controls ATP and ADP gives inconsistent results.....	189
5.7 Chapter conclusion .....	190
Chapter 6: The role of Ap <sub>4</sub> A and Ap <sub>4</sub> Ns in the inhibition of DNA replication initiation	192
6.1 Introduction to the role of Ap <sub>4</sub> A in DNA replication .....	193
6.1.1 DNA replication can be reconstituted <i>in vitro</i> .....	193
6.1.2 Introduction to the cell-free replication system .....	193
6.1.3 Ap <sub>4</sub> A inhibits the initiation of DNA replication.....	194
6.1.4 Aims .....	194
6.2 Preparation of nuclei and extracts for cell-free replication assays.....	195
6.2.1 Synchronisation of late G1 nuclei, and G1- and S-phase extracts.....	195
6.2.2 Testing nuclei batches and extracts in cell-free replication assays.....	196
6.3 Validation of the chemical synthesis of Ap <sub>4</sub> A and nuclei batches.....	201
6.3.1 Validation of the function of the chemically synthesised Ap <sub>4</sub> A .....	201
6.3.2 Validation of the nuclei batches through assessing the effect of Ap <sub>4</sub> A .....	203
6.4 Investigating the effect of Ap <sub>4</sub> A modification on DNA replication inhibition ....	205
6.4.1 Investigating the effect of the phosphate chain in Ap <sub>4</sub> A binding.....	205
6.4.2 Investigating the effect of fluorescent Ap <sub>4</sub> A on DNA replication initiation	207
6.5 Investigating the role of Ap <sub>4</sub> Ns in inhibiting the initiation of DNA replication ..	209
6.5.1 Initial investigation into the activity of Ap <sub>4</sub> Ns in the inhibition of the initiation of DNA replication .....	209

6.6 Detailed investigation of the effect of Ap <sub>4</sub> N in the initiation of DNA replication	211
6.6.1 Ap <sub>4</sub> A inhibits the initiation of DNA replication and the recruitment of replication proteins to chromatin	212
6.6.2 The effect of Ap <sub>4</sub> C on DNA replication and the localisation of replication proteins to chromatin remains unclear	214
6.6.3 Addition of Ap <sub>4</sub> G to S-phase extract shows a trend in reducing percentage initiation and reducing accumulation of replication proteins on chromatin	216
6.6.4 Ap <sub>4</sub> U generally shows a trend in reducing percentage initiation and reducing localisation of replication proteins to chromatin	218
6.6.5 Ap <sub>4</sub> dT shows a strong inhibition of DNA replication initiation which is not explained by changes in nuclear replication protein level	220
6.6.6 Summary of the effects of Ap <sub>4</sub> Ns on DNA replication	224
6.7 Chapter discussion	226
6.7.1 Late G1 nuclei can be synchronised and initiated to replicate <i>in vitro</i>	226
6.7.2 Ap <sub>4</sub> A inhibits the initiation of DNA replication	227
6.7.3 Modification of the phosphate chain does not affect Ap <sub>4</sub> N activity, but modification of the nucleoside components does	227
6.7.4 Ap <sub>4</sub> Ns inhibit the initiation phase of DNA replication, although the potency varies in different batches of nuclei	229
6.7.5 Inhibition of DNA replication initiation by Ap <sub>4</sub> N is associated with a trend suggesting reduced localisation of Mcm2 and PCNA to chromatin	230
6.7.6 Ap <sub>4</sub> dT strongly inhibits the initiation of DNA replication, but this is not reflected in its effect on replication protein localisation to chromatin	232
6.8 Conclusions	232
Chapter 7: General Discussion	234
7.1 Ap <sub>4</sub> Ns can be synthesised via a biologically relevant mechanism	235

7.1.1 A mechanism linking Ap <sub>4</sub> A synthesis to the stress response has previously been described .....	235
7.1.2. UBE1/Ub can synthesise Ap <sub>4</sub> A as well as other Ap <sub>4</sub> Ns .....	235
7.1.3 Further development of the method is required to enable verification of UBE1/Ub-synthesised Ap <sub>4</sub> Ns .....	236
7.2 Ap <sub>4</sub> Ns can be detected in the NuKO mammalian cell line .....	237
7.2.1 Ap <sub>4</sub> A, Ap <sub>4</sub> C Ap <sub>4</sub> G and Ap <sub>4</sub> U have previously been identified in bacteria and lower eukaryotes, but not in mammals.....	237
7.2.2 Ap <sub>4</sub> Ns can be extracted from mammalian cells and detected by triple quadrupole mass spectrometry.....	237
7.2.3 Further development of the method for identifying Ap <sub>4</sub> Ns in mammalian cells .....	238
7.3 DNA damage may increase intracellular Ap <sub>4</sub> N concentration in NuKO cells, but improved controls are required for quantitation.....	239
7.3.1 The effect of stress on Ap <sub>4</sub> N concentration .....	239
7.3.2 Unstandardised data suggests that intracellular Ap <sub>4</sub> N levels increase in response to a variety of damaging agents.....	240
7.3.3 Further method development to investigate the effect of DNA damaging agents on intracellular Ap <sub>4</sub> N abundance.....	241
7.4 Ap <sub>4</sub> Ns have differential inhibitory activities during the initiation of DNA replication.....	243
7.4.1 The role of Ap <sub>4</sub> A and Ap <sub>4</sub> Ns in DNA replication.....	243
7.4.2 Ap <sub>4</sub> Ns inhibit the initiation of DNA replication to different extents .....	244
7.4.3 Limitations and further method developments for investigating the role of Ap <sub>4</sub> Ns in DNA replication initiation.....	245
7.5 Ap <sub>4</sub> Ns reduce nuclear Mcm2 and PCNA levels in late G1 nuclei added to S-phase extract, with a greater effect on PCNA levels.....	245

7.5.1 The nuclear localisation of replicative helicase and processivity factor are affected by Ap <sub>4</sub> Ns .....	246
7.5.2 Limitations and future perspectives of Western blotting for determining the mechanism behind Ap <sub>4</sub> N-induced inhibition of DNA replication.....	249
7.6 Conclusion.....	249
Appendices and Supplementary data.....	251
References .....	283



## List of Tables

Table 1.1: A summary of the different Ap <sub>4</sub> N synthesis mechanisms.....	27
Table 1.2: A summary of the Ap <sub>4</sub> A binding partners and related signalling mechanisms.....	45
Table 2.1: Reactants for the synthesis of Ap <sub>4</sub> N and Ap <sub>4</sub> dN.....	67
Table 2.2: Eluents for triple quadrupole mass spectrometry analysis.....	72
Table 2.3: Method for LCMS-8040 analysis.....	72
Table 2.4: Primary antibody species and dilutions.....	76
Table 3.1: Retention times for each Ap <sub>4</sub> N.....	94
Table 3.2: An estimate of the purity of the synthesised Ap <sub>4</sub> Ns.....	97
Table 3.3: A summary of the expected product retention times, and the corresponding peaks in the hydrolysis products.....	103
Table 3.4: Determining the extinction coefficient for each Ap <sub>4</sub> N.....	109
Table 4.1: Summary of the three most abundant transitions for each Ap <sub>4</sub> N.....	131
Table 5.1: Summary of the mean change in abundance of the three transitions for each Ap <sub>4</sub> N.....	182
Table 6.1: Summary data for the change in percentage initiation for Ap <sub>4</sub> A, Ap <sub>4</sub> C, Ap <sub>4</sub> U, Ap <sub>4</sub> G and Ap <sub>4</sub> dT.....	224
Table 6.2: Summary data for the effect of Ap <sub>4</sub> Ns on the Mcm2 and PCNA localisation to chromatin.....	225

## List of Figures

Figure 1.1: The chemical structure of Ap <sub>4</sub> A.....	25
Figure 1.2: Aminoacyl-tRNA mediated synthesis of Ap <sub>4</sub> A.....	26
Figure 1.3 Synthesis of Ap <sub>4</sub> A by Ub- and Ubl- activating enzymes.....	29
Figure 1.4: Function of RppH and ApaH in decapping methylated or non-methylated Ap <sub>4</sub> N transcripts.....	34
Figure 1.5: Structure of the non-hydrolysable Ap <sub>4</sub> A analogue JB419.....	35
Figure 1.6: The Ap <sub>4</sub> A-induced polymerisation of HINT1.....	36
Figure 1.7: A summary of the cGAS-STING immune response.....	38
Figure 1.8: A schematic showing the <i>de novo</i> synthesis of GTP and ATP.....	40
Figure 1.9: Progression through the eukaryotic cell cycle.....	47
Figure 1.10: CDK-induced activation of the DNA helicase.....	52
Figure 1.11: An overview of the roles of ATM and ATR in the DNA damage response.....	55
Figure 1.12: The ATM response to DNA double strand breaks.....	58
Figure 1.13: Strand invasion during homologous recombination.....	61
Figure 1.14: Strand invasion during homologous recombination at the molecular level.....	63
Figure 3.1: Establishing retention times for substrates and products of the Ap <sub>4</sub> A synthesis reactions.....	84
Figure 3.2: Schematic for the chemical synthesis of Ap <sub>4</sub> A.....	85
Figure 3.3: Effect of increased ATP concentration on Ap <sub>4</sub> A synthesis.....	86
Figure 3.4: Carbodiimide coupling reaction for Ap <sub>4</sub> A synthesis.....	87
Figure 3.5: Effect of increased volume of HEPES and EDC on Ap <sub>4</sub> A synthesis.....	89
Figure 3.6: Effect of temperature on Ap <sub>4</sub> A yield.....	91
Figure 3.7: Structural differences between different Ap <sub>4</sub> Ns.....	93
Figure 3.8: Purification of Ap <sub>4</sub> Ns and Ap <sub>4</sub> dNs.....	96
Figure 3.9: Hydrolysis of Ap <sub>4</sub> Ns to NMP and NTP.....	99
Figure 3.10: Hydrolysis of Ap <sub>4</sub> Ns to NMP and NTP.....	101
Figure 3.11: Verification of Ap <sub>4</sub> Ns by mass spectrometry.....	105
Figure 3.12: Verification of Ap <sub>4</sub> dNs by mass spectrometry.....	107

Figure 3.13: <i>In vitro</i> synthesis of Ap <sub>4</sub> Ns by the ubiquitin-activating enzyme UBE1.....	111
Figure 3.14: Effect of increased incubation time on Ap <sub>4</sub> N synthesis by UBE1.....	113
Figure 4.1: The effect of phosphatase (and Ap <sub>4</sub> A hydrolase) on ATP and Ap <sub>4</sub> A..	124
Figure 4.2: Schematics for the extraction of metabolites from cells.....	126
Figure 4.3: Analysis of NuKO cell extracts by HPLC.....	128
Figure 4.4: The triple quadrupole system as shown by Kitteringham and colleagues.....	130
Figure 4.5: Determination and validation of MRM transitions for each Ap <sub>4</sub> N.....	132
Figure 4.6: Assessing the abundance of transitions common to multiple molecules using Ap <sub>4</sub> N standards.....	133
Figure 4.7: Investigating whether Ap <sub>4</sub> Ns and Ap <sub>4</sub> dNs are present in KBM7 cells.....	136
Figure 4.8: Determining the lower Ap <sub>4</sub> A concentration for detection in KBM7 extracts.....	137
Figure 4.9: The effect of the Pierce cleanup on detection of Ap <sub>4</sub> A at lower concentrations.....	140
Figure 4.10: Identification of Ap <sub>4</sub> A in a KBM7 cell extract.....	142
Figure 4.11: Identification of Ap <sub>4</sub> Ns in NuKO cells.....	144
Figure 4.12: Identification of Ap <sub>4</sub> A in NuKO cell extract.....	146
Figure 4.13: Identification of Ap <sub>4</sub> G in NuKO cell extract.....	147
Figure 4.14: Identification of Ap <sub>4</sub> C in NuKO cell extract.....	148
Figure 4.15: Identification of Ap <sub>4</sub> U in NuKO cell extract.....	149
Figure 4.16: Identification of several Ap <sub>4</sub> dNs in NuKO mammalian cell extract..	150
Figure 4.17: Analysis of fractionated NuKO extract.....	152
Figure 5.1: Differences in duplication time of KBM7 and NuKO cells.....	161
Figure 5.2: Comparing the cell cycle dynamics of KBM7 and NuKO cells.....	162
Figure 5.3: Investigating the $\gamma$ H2AX response to replication stress in KBM7 and NuKO cells.....	164
Figure 5.4: Investigating the apoptotic response to replication stress in KBM7 and NuKO cells.....	165

Figure 5.5: Variability in LC-MS analysis outputs and an approximate comparison of Ap <sub>4</sub> N abundance.....	167
Figure 5.6: The effect of HU treatment on Ap <sub>4</sub> N level in NuKO cells.....	169
Figure 5.7 Effect of increased length of HU-treatment on Ap <sub>4</sub> N concentration..	171
Figure 5.8: The effect of recovery time on Ap <sub>4</sub> N levels after HU-treatment.....	173
Figure 5.9: Effect of 18-hour treatment with damaging agents on Ap <sub>4</sub> N levels..	176
Figure 5.10: Validation of ADP and ATP MRM transitions using commercially available standards.....	178
Figure 5.11: Effect of damaging agents on internal ADP and ATP standards.....	180
Figure 6.1: Preparation of nuclei and extracts for cell-free replication assays....	196
Figure 6.2: Cell-free replication assays to test nuclei batches.....	198
Figure 6.3: Validation of the HeLa cell synchronisation process.....	200
Figure 6.4: Validation of the inhibitory activity of synthesised Ap <sub>4</sub> A relative to an Ap <sub>4</sub> A standard.....	202
Figure 6.5: The response of different nuclei batches to Ap <sub>4</sub> A.....	204
Figure 6.6: The effect of AppNppA on the initiation of DNA replication.....	206
Figure 6.7: The effect of Ap <sub>4</sub> A-FRET on the initiation of DNA replication.....	208
Figure 6.8: Initiation of replication after addition of each Ap <sub>4</sub> N to the first batch of nuclei.....	210
Figure 6.9: The role of Ap <sub>4</sub> A in inhibiting DNA replication.....	213
Figure 6.10: The role of Ap <sub>4</sub> C in the inhibition of DNA replication. ....	215
Figure 6.11: The role of Ap <sub>4</sub> G in inhibiting DNA replication.....	217
Figure 6.12: The role of Ap <sub>4</sub> U in inhibiting DNA replication.....	219
Figure 6.13: The role of Ap <sub>4</sub> dT in inhibiting DNA replication.....	221
Figure 6.14: The effect of Ap <sub>4</sub> dT on the inhibition of DNA replication initiation in cells that do not show a response to Ap <sub>4</sub> A.....	223
Figure 7.1: Essential components of an Np <sub>n</sub> N molecule capable of inhibiting the initiation of DNA replication.....	244
Figure 7.2: Stages of DNA replication that may be affected by Ap <sub>4</sub> N.....	248

## Abbreviations

A = absorbance

aaRS = aminoacyl tRNA synthetase

ADP = adenosine diphosphate

ADSS = adenylosuccinate synthase

AMP = adenosine monophosphate

ApaH = Ap<sub>4</sub>A hydrolase

APC/C<sup>CDH1</sup> = anaphase promoting complex

Ap<sub>4</sub>A = diadenosine tetraphosphate

Ap<sub>4</sub>C = adenosine cytidine tetraphosphate

Ap<sub>4</sub>dA = adenosine deoxyadenosine tetraphosphate

Ap<sub>4</sub>G = adenosine guanosine tetraphosphate

Ap<sub>4</sub>U = adenosine uridine tetraphosphate

Ap<sub>4</sub>N = dinucleotide tetraphosphate, where one nucleotide is always adenine

Arg = arginine

ATM = ataxia-telangiectasia mutated

ATR = ATM- and Rad3-Related

ATRIP = ATR-interacting protein

ATP = adenosine triphosphate

BRCT = BRCA1 C-terminal

BSA = bovine serum albumin

c = concentration

CBS = Cystathionine-β-synthase

CDK = cyclin dependent kinase

CO<sub>2</sub> = carbon dioxide

CPK = creatine phosphokinase

CTIP = (carboxy-terminal binding protein) interacting protein

CTP = cytidine triphosphate

cGAMP = cyclic guanosine monophosphate–adenosine monophosphate

cGAS = cyclic GMP-AMP synthase

CHK1 = checkpoint kinase 1

CHK2 = checkpoint kinase 2  
CMG = Cdc45-MCM-GINS  
CoA = coenzyme A  
Cryo-EM = cryo-electron microscopy  
dAMP = deoxyadenosine monophosphate  
DAPI = 4',6-diamidino-2-phenylindole  
dATP = deoxyadenosine triphosphate  
DDK = Cdc7/Dbf4-dependent kinase  
D-MEM = Dulbecco's Modified Eagle's Medium  
DNA = deoxyribonucleic acid  
dNTP = deoxynucleotide triphosphate  
DBPS = Dulbecco's phosphate buffered saline  
DMAP = 6-dimethylaminopurine  
DOX = doxorubicin  
DSB = double strand break  
DSBR = double-strand break repair  
EDC = 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide  
EDTA = Ethylenediaminetetraacetic acid  
EdU = 5-ethynyl-2'-deoxyuridine  
EMI1 = early mitotic inhibitor 1  
ESI = electrospray ionisation  
 $\epsilon$  = extinction coefficient  
FAD = flavin adenine dinucleotide  
FBS = Fetal Bovine Serum  
FHA = forkhead-associated  
FRET = Förster resonance energy transfer  
GEM = gemcitabine  
GlcNAc = N-acetylglucosamine  
GlyRS = glycyl-tRNA synthetase  
Gp<sub>4</sub>G = diguanosine tetraphosphate  
GST = glutathione S-transferase  
GTP = guanosine triphosphate

G1 = first gap phase  
G2 = second gap phase  
HBRCT = helix-BRCA1 C-terminal  
HEPES = 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid  
HILIC = hydrophilic interaction liquid chromatography  
HINT1 = histidine triad nucleotide-binding protein 1  
His = histidine  
HIT = histidine triad  
HPLC = high performance liquid chromatography  
HR = homologous recombination  
HU = hydroxyurea  
H2A K15ub = histone H2A ubiquitinated on lys-15  
H2AX = H2A histone family member X  
H<sub>2</sub>O<sub>2</sub> = hydrogen peroxide  
H4K20me2 = histone H4 dimethylated lysine-20  
IC50 = half-maximal inhibitory concentration  
IgE = immunoglobulin E  
IMDM = Iscove's Modified Dulbecco's Medium  
IMP = inosine 5'-monophosphate  
IMPDH = Inosine-5'-monophosphate dehydrogenase  
IP-RP = ion-paired reverse-phase  
IκBα = nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, alpha  
IRF3 = interferon regulatory factor 3  
Kb = kilobases  
l = path length  
LC-MS = liquid chromatography mass spectrometer  
LCMS-IT-TOF = liquid chromatography coupled to ion trap time-of-flight  
Lys = lysine  
LysRS = lysyl tRNA synthetase  
M = mitotic phase  
MCM = minichromosome maintenance complex

Mcm2–7 = minichromosome maintenance proteins 2 to 7  
 MDC1 = Mediator of DNA damage checkpoint protein 1  
 MDMX = murine double minute X  
 MDM2 = murine double minute 2  
 MEF = Mouse embryo fibroblast  
 MgCl<sub>2</sub> = Magnesium chloride  
 MITF = microphthalmia transcription factor  
 MMC = mitomycin C  
 MO = Mcm2–7-ORC  
 MOC = Mcm2–7-ORC-Cdc6  
 MRM = multiple reaction monitoring  
 MRN = Mre11-Rad50-Nbs1  
 mRNA = Messenger RNA  
 NAD<sup>+</sup>/NADH = nicotinamide adenine dinucleotide  
 NCIN = non-canonical initiating nucleotide  
 NHEJ = non-homologous end joining  
 NMP = nucleotide monophosphate  
 Np<sub>4</sub>N = dinucleotide tetraphosphate  
 Np<sub>n</sub>N = dinucleotide polyphosphate, where n=the number of phosphate groups  
 NSC = neuronal stem cells  
 NSD = amino terminal serine/threonine-rich domain  
 Nt = nucleotides  
 NTD-A = N-terminal domain A  
 NTP = nucleoside triphosphate  
 Nudt2 = Nudix hydrolase 2  
 NuKO = *Nudt2* knockout KBM7 cell line  
 OB = oligonucleotide/oligosaccharide-binding  
 OCCM = ORC-Cdc6-Cdt1-MCM  
 OCM = ORC-Cdc6-MCM  
 OCMM = ORC-Cdc6-MCM-MCM  
 PBS = phosphate buffered saline  
 PCNA = proliferating cell nuclear antigen



PEG = polyethylene glycol  
PG = phosphorylation-generated  
Phe = phenylalanine  
PI = propidium iodide  
PIKK = phosphatidylinositol-3 kinase-related kinase  
PKCI-1 = protein kinase C-interacting protein 1  
PMSF = Phenylmethylsulphonyl Fluoride  
Pol $\alpha$  = DNA polymerase alpha  
Pol  $\delta$  = DNA polymerase delta  
Pol  $\epsilon$  = DNA polymerase epsilon  
PPi = pyrophosphatase  
P/S/G = penicillin/streptomycin/glutamine  
PTIP = Pax transactivation domain-interacting protein  
PP2A = protein phosphatase 2A  
Pre-IC = pre-initiation complex  
Pre-LC = pre-loading complex  
Pre-RC = pre-replication complex  
PVDF = polyvinylidene fluoride  
QFWB = Quantitative fluorescent Western blotting  
QQQ = triple quadrupole  
Rb = Retinoblastoma tumour suppressor protein  
RFC = replication factor C  
RNA = ribonucleic acid  
RPA = replication protein A  
RppH = RNA pyrophosphohydrolase  
SDSA = synthesis-dependent strand annealing  
sRNA = small RNA  
STING = stimulator of interferon genes  
S = synthesis  
SDS-PAGE = Sodium Dodecyl Sulfate–Polyacrylamide Gel Electrophoresis  
Ser = serine  
ssDNA = single stranded DNA

TB = trypan blue  
TBK1 = TANK-binding kinase 1  
TEAA = triethylammonium acetate  
TFA = trifluoroacetic acid  
Thr = threonine  
TIRR = Tudor Interacting Repair Regulator  
TOPBP1 = Treslin and DNA topoisomerase II $\beta$ -binding protein 1  
tRNA = transfer RNA  
Tyr = Tyrosine  
Ub = Ubiquitin  
UBA1 = Ub-like modified-activating enzyme  
UBE1 = human ubiquitin activating enzyme  
UDP = uridine diphosphate  
UDR = ubiquitination-dependent recruitment  
UHPLC = Ultra-high performance liquid chromatography  
UTP = uridine triphosphate  
UV = ultraviolet  
VEGFR = vascular endothelial growth factor receptor 2  
WB = Western blotting  
WHD = winged helix domain  
XMP = xanthosine 5'-monophosphate

## Acknowledgements

Firstly, I would like to thank Dr. Nikki Copeland for all his incredible guidance and support throughout this experience. Thank you for your patience, for answering my many questions and for being a fantastic supervisor from beginning to end. I would also like to thank Dr. Mick Urbaniak for the invaluable support he has provided throughout, and particularly for his guidance on the HPLC and LC-MS aspects of my project. Many thanks also to Prof. Sandy McLennan and Dr. Nigel Jones for their support, training and the wealth of knowledge they have contributed to the project. Thank you also to Dr. David Rochester for all his support with the HRMS and LC-MS side of my project.

I would like to thank Dr. James Tollitt, Natasha Falkingham, Olivia Iwanowysch, Louise Caprani, Megan Fraser and Dr. Betül Sopaci, who have been my lab companions across my PhD years and without whom my experience would not have been the same. Special thanks to Dr. James Tollitt, who has been in the lab from the beginning of my PhD experience, and from whom I have learnt so much. In addition to his support in the lab, I would like to thank James for all the humour, random facts and musicality that he brought to the lab.

Next, I would like to thank my parents, Wendy and Jim, and my brother Rowan for their limitless kindness, love, support and encouragement throughout this PhD. Thank you for the relaxing walks in the countryside, the frequent phone calls and the home cooked meals. Thank you also to Alan, Eileen and family for their constant support from beginning to end.

Finally, a huge thanks to my fiancé Mike, who has provided unconditional love and support throughout my PhD. We began our PhDs together and have shared the ups and downs of the roller coaster that it has been from start to finish. This journey would not have been the same without you. Thank you.

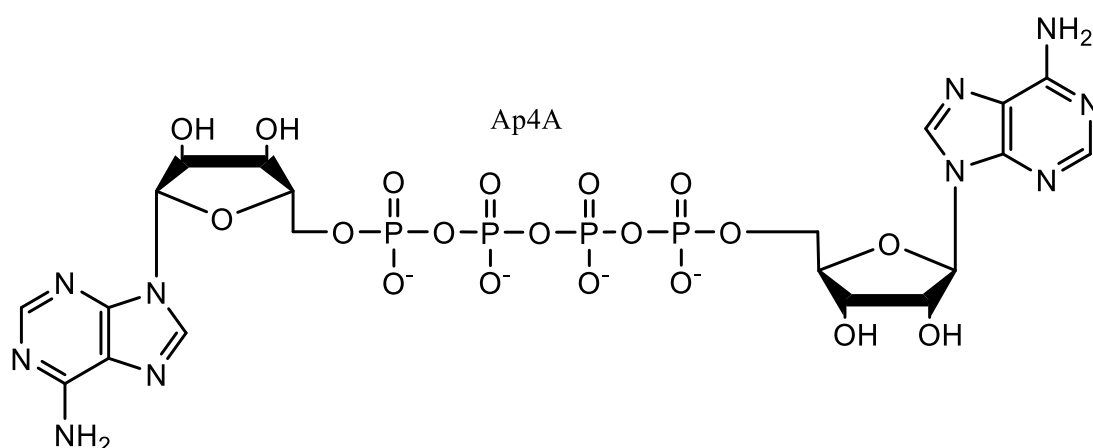
## **Author's declaration**

This thesis is my own work and has not been submitted in substantially the same form for the award of a higher degree elsewhere. Any sections of the thesis which have been published elsewhere have been clearly identified.

# **Chapter 1: General Introduction**

## 1.1 Introduction to Ap<sub>4</sub>A

Diadenosine tetraphosphate (Ap<sub>4</sub>A) is an alarmone that is ubiquitous across many domains of life which has been implicated in several signalling mechanisms. It is composed of two adenosine residues linked by a chain of four phosphates (Fig. 1.1). Ap<sub>4</sub>A was first identified in the 1960s and has frequently been linked to DNA replication, amongst a number of other signalling mechanisms (Zamecnik et al., 1966). Recently Ap<sub>4</sub>A was shown to inhibit the initiation stage of deoxyribonucleic acid (DNA) replication (Marriott et al., 2015). Because Ap<sub>4</sub>A levels are increased in response to a variety of different stressors, it is possible that Ap<sub>4</sub>A regulates DNA replication in a damage-associated manner. In this section the synthesis and degradation of Ap<sub>4</sub>A will be discussed. Subsequently, the intracellular signalling mechanisms that have been identified will be discussed in both eukaryotic and prokaryotic systems. Further detail about Ap<sub>4</sub>A and its intracellular signalling mechanisms can be found in a recent review (Ferguson et al., 2020).

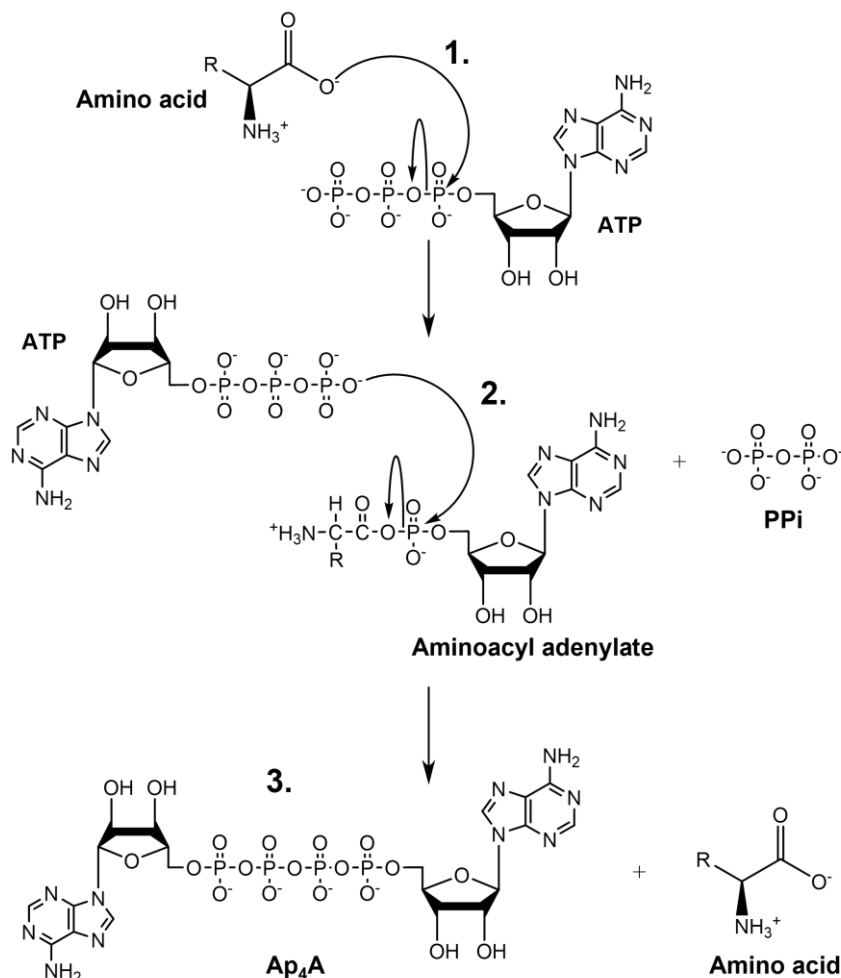


**Figure 1.1: The chemical structure of Ap<sub>4</sub>A.** Ap<sub>4</sub>A is composed of two adenosine nucleotides linked by a chain of four phosphate groups. Ap<sub>4</sub>A, diadenosine tetraphosphate.

### 1.1.1 Ap<sub>4</sub>A can be synthesised through different mechanisms

The level of Ap<sub>4</sub>A in cells is controlled by the balance between the rate of its synthesis and degradation, and it is synthesised through several biological mechanisms. Ap<sub>4</sub>A was initially identified as a product of adenosine triphosphate (ATP) reacting with the lysyl-small RNA (sRNA) synthetase in the presence of lysine and Mg<sup>2+</sup> (Zamecnik et al., 1966). It was later demonstrated that Ap<sub>4</sub>A could be produced from many different aminoacyl-tRNA synthetases (aaRSs) in both prokaryotes and eukaryotes (Goerlich et

al., 1982; Plateau et al., 1981). In this process an aminoacyl adenylate is first produced from  $\text{Mg}^{2+}$ -ATP, followed by reaction with a second ATP molecule to produce  $\text{Ap}_4\text{A}$  (Fig. 1.2) (Goerlich et al., 1982).



**Figure 1.2: Aminoacyl-tRNA mediated synthesis of  $\text{Ap}_4\text{A}$ .** 1. An amino acid reacts with ATP to form an aminoacyl adenylate. 2. A second ATP molecule adds to the aminoacyl adenylate as a substrate, releasing inorganic pyrophosphate ( $\text{PPi}$ ). 3. This results in the production of  $\text{Ap}_4\text{A}$ .

$\text{Zn}^{2+}$  is a more potent activator in reactions producing  $\text{Ap}_4\text{A}$  from phenylalanyl-tRNA synthetase and lysyl-tRNA synthetase but has no effect in reactions involving the other tRNA synthetases (Goerlich et al., 1982; Plateau et al., 1981). In some cases, such as the seryl-tRNA synthetase,  $\text{Mn}^{2+}$  ions can also stimulate  $\text{Ap}_4\text{A}$  synthesis (Belrhali et al., 1995). Unlike the other aminoacyl tRNA synthetases, glycine concentration is not

important for the synthesis of Ap<sub>4</sub>A by glycyl tRNA-synthetase, which can produce Ap<sub>4</sub>A through a direct ATP condensation mechanism (Guo et al., 2009).

Since the phosphate group, rather than the nucleotide moiety itself, is involved in the reaction, this suggests that other dinucleoside tetraphosphates, where one nucleotide is always adenosine (Ap<sub>4</sub>Ns), could be produced by aminoacyl tRNA synthetases if a different nucleoside triphosphate (NTP) was used as the substrate. Indeed, in the presence of other NTPs, other Ap<sub>4</sub>Ns are synthesised via this mechanism (Randerath et al., 1966). Several synthesis mechanisms implicated in Ap<sub>4</sub>A or Ap<sub>4</sub>N synthesis are described in Table 1.1.

**Table 1.1: A summary of the different Ap<sub>4</sub>N synthesis mechanisms.** AMP, adenosine monophosphate; Ap<sub>4</sub>A, diadenosine tetraphosphate; Ap<sub>4</sub>N, dinucleotide tetraphosphate, where one nucleotide is adenosine; ATP, adenosine triphosphate; CoA, coenzyme A; E, enzyme; NEDD8, neural precursor cell expressed developmentally down-regulated protein 8; NTP, nucleoside triphosphate; PPi, inorganic pyrophosphate; SUMO, small ubiquitin-like modifier; Ub, ubiquitin; Ubl, ubiquitin-like.

Protein	Method of Ap <sub>4</sub> A/Ap <sub>4</sub> N synthesis	References
<b>DNA and RNA ligase</b> T4 phage	$E + ATP \rightarrow E-AMP + PPi$ $E-AMP + NTP \rightarrow Ap_4N$	(Atencia et al., 1999; Madrid et al., 1998)
<b>Firefly luciferase</b> <i>Photinus pyralis</i>	$Luciferin + ATP + E \rightarrow E-luciferin-AMP + PPi$ $E-luciferin-AMP \rightarrow E-dehydroluciferin-AMP$ $E-luciferin-AMP + ATP \rightarrow Ap_4A + luciferin + E$	(Guranowski et al., 1990; Sillero and Sillero, 2000)
<b>Aminoacyl tRNA synthetases</b> LysRS, PheRS, HisRS, SerRS, IleRS, LeuRS, AspRS, TryRS, ValRS, GlyRS Prokaryotes and eukaryotes	$Amino\ acid + ATP + E \rightarrow E-aminoacyl-AMP + PPi$ $E-aminoacyl-AMP + ATP \rightarrow Ap_4A + Amino\ acid + E$	(Goerlich et al., 1982; Guo et al., 2009; Plateau et al., 1981; Zamecnik et al., 1966)
<b>Acyl-CoA synthetase</b> <i>Pseudomonas fragi</i>	$E + ATP \rightarrow E-AMP + PPi$ $E-AMP + NTP \rightarrow E + Ap_4N$ OR $E-RCO-AMP + NTP \rightarrow E + RCOOH + Ap_4N$ (in the presence of fatty acids)	(Fontes et al., 1998)
<b>Ubiquitin or Ubiquitin-like activating enzyme</b> UBA1, NEDD8, SUMO Human	$E + ATP + Ub/Ubl \rightarrow E + Ub/Ubl-AMP + PPi$ $ATP + Ub/Ubl-AMP \rightarrow Ub/Ubl + Ap_4A$	(Götz et al., 2019)

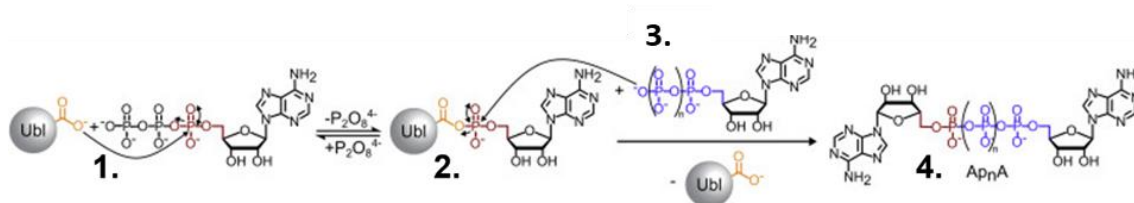


Other enzymes that are capable of Ap<sub>4</sub>A and Ap<sub>4</sub>N synthesis include several DNA and ribonucleic acid (RNA) ligases. ATP reacts with T4 DNA ligase to produce an enzyme-adenosine monophosphate (E-AMP) intermediate, which subsequently reacts with NTPs or deoxynucleoside triphosphates (dNTPs) including ATP, guanosine triphosphate (GTP) and deoxyadenosine triphosphate (dATP) to produce Ap<sub>4</sub>A, adenosine guanosine tetraphosphate (Ap<sub>4</sub>G) and adenosine deoxyadenosine tetraphosphate (Ap<sub>4</sub>dA), respectively (Madrid et al., 1998). In addition to T4 DNA ligase, in the absence of DNA, human DNA ligase III $\beta$  is also capable of synthesising Ap<sub>4</sub>A (McLennan, 2000). Similar to the T4 DNA ligase, T4 DNA polymerase can synthesise Ap<sub>4</sub>A, with ATP initially acting as a donor to form E-AMP, followed by ATP or NTP acting as an acceptor to produce Ap<sub>4</sub>A or Ap<sub>4</sub>N (Atencia et al., 1999). While the NTPs tested were capable of acting as donors for an enzyme-nucleotide monophosphate (E-NMP) complex, the relative rate of this occurring for either cytidine triphosphate (CTP) or GTP was over 100 times lower than for ATP (Atencia et al., 1999).

Another ligase capable of Ap<sub>4</sub>N synthesis is the acyl-Coenzyme A (acyl-CoA) synthetase in *Pseudomonas fragi* (Fontes et al., 1998; Sillero and Sillero, 2000). Although the primary nucleotide polyphosphates produced by this enzyme have only one nucleotide (eg. p<sub>4</sub>A), Ap<sub>4</sub>Ns can also be produced by acyl-CoA (Fontes et al., 1998; Sillero and Sillero, 2000). In this process, ATP initially acts as a donor to form an intermediate complex, followed by reaction with an NTP, which acts as an acceptor to form Ap<sub>4</sub>N (Fontes et al., 1998). Acetyl-Coenzyme A (Acetyl-CoA) in *Saccharomyces cerevisiae* functions in a similar way but has only been demonstrated to produce nucleotide polyphosphates with a single nucleotide (Guranowski et al., 1994b). Unlike in the other synthesis mechanisms which form an Enzyme-X-AMP intermediate, for both Acetyl-CoA and Acyl-CoA, synthesis can be catalysed without an acyl/acetyl intermediate (E-AMP intermediate) (Fontes et al., 1998; Guranowski et al., 1994b).

Firefly luciferase is an oxidoreductase that shares some characteristics with the ligases and is also capable of synthesising Ap<sub>4</sub>A and Ap<sub>4</sub>G in a luciferin and divalent cation (Mg<sup>2+</sup>, Mn<sup>2+</sup>, Co<sup>2+</sup> or Zn<sup>2+</sup>)-dependant manner (Guranowski et al., 1990; Sillero and Sillero, 2000). Initially it was assumed that Ap<sub>4</sub>N was produced from the

E-luciferin-AMP intermediate (Guranowski et al., 1990). While this may be possible to a small extent, it was later demonstrated that E-dehydroluciferin-AMP was the major intermediate in the synthesis of Ap<sub>4</sub>N by firefly luciferase (Sillero and Sillero, 2000). Recently the ubiquitin (Ub) activating enzymes have also been implicated in the synthesis of Ap<sub>4</sub>A. In this proposed mechanism, Ub is adenylated by the Ub-activating enzyme in the presence of ATP. The phosphate group of a second ATP molecule then attacks the adenylated complex to form Ap<sub>4</sub>A (Fig. 1.3) (Götz et al., 2019).



**Figure 1.3 Synthesis of Ap<sub>4</sub>A by Ub- and Ubl- activating enzymes.** In the hypothetical model proposed by Götz and colleagues in 2019: A Ub- or Ubl activating enzyme initiates a reaction between ATP and Ub or a Ub-like protein (1) resulting in the formation of the adenylated Ub/Ub-like intermediate (2). The  $\gamma$ -phosphate of a second ATP molecule subsequently reacts with the adenylated intermediate (3), producing Ap<sub>4</sub>A (4). Ap<sub>4</sub>A, diadenosine tetraphosphate; Ub, ubiquitin; Ubl, Ubiquitin-like. Figure reprinted from Cell Chemical Biology, Volume 26, Götz, K. H., Mex, M., Stuber, K., Offensperger, F., Scheffner, M. & Marx, A. Formation of the Alarmones Diadenosine Triphosphate and Tetraphosphate by Ubiquitin- and Ubiquitin-like-Activating Enzymes, Pages No. 1535–1543, Copyright 2019, with permission from Elsevier. Numbering was added for the purpose of this thesis.

Since it is the  $\gamma$ -phosphate of the second ATP, and not the adenosine ring, that reacts with the adenylated Ub complex, it is likely that other Ap<sub>4</sub>Ns may also be synthesised in this way. However, this has not yet been studied and further investigation into this possibility is covered in Chapter 3. Identification of the Ub activating enzymes as having the potential to synthesise Ap<sub>4</sub>A has demonstrated a mechanism for Ap<sub>4</sub>A synthesis which links its synthesis to stress conditions, as the activity of the Ub activating enzyme is increased under conditions of stress (Shang et al., 1997). This link between Ap<sub>4</sub>A synthesis and stress is strongly supported by the numerous mechanisms showing increases in intracellular Ap<sub>4</sub>A level because of cell stress. Types of stress that cause intracellular Ap<sub>4</sub>A concentration to be elevated are discussed in section 1.1.3.

### 1.1.2 Hydrolysis of Ap<sub>4</sub>A is required for maintaining the appropriate Ap<sub>4</sub>A level

To prevent a build-up of Ap<sub>4</sub>A in the cell at times when the cell is not under stress, there must also be a mechanism for the breakdown of Ap<sub>4</sub>A. In cells, Ap<sub>4</sub>A can either be hydrolysed symmetrically, forming adenosine diphosphate (ADP), or asymmetrically, forming ATP and AMP. Ap<sub>4</sub>A hydrolase (ApaH) and Nudix hydrolase 2 (Nudt2) are the major players in Ap<sub>4</sub>A hydrolysis in bacteria and mammalian cells, respectively.

In bacterial cells, ApaH breaks down Ap<sub>4</sub>A symmetrically, resulting in the production of ADP (Guranowski et al., 1983). Increased Ap<sub>4</sub>A hydrolase concentration in *Escherichia coli* is associated with a drop in Ap<sub>4</sub>N concentration, whereas the removal of Ap<sub>4</sub>A hydrolase in an ApaH<sup>-</sup> mutant resulted in an increase in Ap<sub>4</sub>A levels by at least 16-fold (Farr et al., 1989; Plateau et al., 1987a). Recently the YqeK (COG1713) protein family, which is present in a group of gram-positive bacteria lacking ApaH, has also been identified as an alternative symmetrical Ap<sub>4</sub>A hydrolase in *Bacillus subtilis* (Minazzato et al., 2020). Furthermore, Rv2613, an Ap<sub>4</sub>A phosphorylase in *Mycobacterium tuberculosis* H37Rv has structural similarity to proteins in the histidine triad (HIT) superfamily and converts Ap<sub>4</sub>A into ATP and ADP. Furthermore, although the YgdP hydrolase in *Salmonella enterica* has Ap<sub>5</sub>A as its main substrate, this can also asymmetrically hydrolyse Ap<sub>4</sub>A, demonstrating evidence of a bacterial nudix hydrolase with Ap<sub>4</sub>A-hydrolysis activity (Ismail et al., 2003; McLennan, 2006). Therefore, there are several enzymes with Ap<sub>4</sub>A-hydrolysis activity in gram-positive and gram-negative bacteria (Mori et al., 2011; Mori et al., 2010).

In *Schizosaccharomyces pombe* the *aph1* gene encodes an asymmetrical Ap<sub>4</sub>A hydrolase which, like Rv2613, shares sequence similarity with the HIT family and is capable of hydrolysing Ap<sub>4</sub>A as well as other dinucleoside polyphosphates (Huang et al., 1995; Robinson et al., 1993). The major human Ap<sub>4</sub>A hydrolase is encoded by the *Nudt2* gene and belongs to the Nudix family (Thorne et al., 1995). Ap<sub>4</sub>A interacts with this 17-kDa asymmetric human Ap<sub>4</sub>A hydrolase via a ring stacking arrangement, and disruption of this hydrolase results in an accumulation of Ap<sub>4</sub>A, increasing its intracellular concentration 175-fold (Marriott et al., 2016; Swarbrick et al., 2005). This

demonstrates that the nudix hydrolase is highly important for maintaining low Ap<sub>4</sub>A levels in mammalian cells.

### **1.1.3 Ap<sub>4</sub>A is produced in response to different types of stress**

Over the years there have been many instances where Ap<sub>4</sub>A and Ap<sub>4</sub>N levels have been linked to cell stress across a variety of different organisms. One type of stress known to elevate intracellular Ap<sub>4</sub>A levels in both prokaryotes and eukaryotes is heat shock (Brevet et al., 1989a; Coste et al., 1987; Lee et al., 1983). In *Salmonella typhimurium* LT2, Ap<sub>4</sub>A and Ap<sub>4</sub>G are both synthesised, among other dinucleotide polyphosphates, alongside heat shock proteins when temperature increases from 28 °C to 50 °C (Lee et al., 1983). A study in *E. coli* demonstrated that even in aaRS-overproducing cells, increasing the temperature from 37 °C to 48 °C still resulted in an increase in intracellular Ap<sub>4</sub>N concentration (Brevet et al., 1989a). Similar to the effect of heat shock in these bacteria, a temperature shift from 30 °C to 46 °C in *S. cerevisiae* cells resulted in an increase in all Ap<sub>4</sub>N levels as well as the level of other Np<sub>4</sub>Ns (Coste et al., 1987).

Other cases of Ap<sub>4</sub>N level increasing in response to cell stress include in the cellular response to oxidative stress. This was demonstrated in *S. typhimurium*, which had elevated dinucleotide polyphosphate levels, including Ap<sub>4</sub>A and Ap<sub>4</sub>G, after exposure to 10% ethanol (Lee et al., 1983). Ap<sub>4</sub>N and other dinucleotide tetraphosphate (Np<sub>4</sub>N) concentrations were increased in *E. coli* and *S. cerevisiae* after cadmium exposure (Coste et al., 1987). Ap<sub>4</sub>A concentration in *E. coli* is also increased by exposure to aminoglycoside antibiotics such as kanamycin (Ji et al., 2019). In this case, an increase in Ap<sub>4</sub>A levels has been linked to the efficacy of the aminoglycosides, suggesting that Ap<sub>4</sub>A may have a role in the response to stress.

In the acellular slime mould *Physarum polycephalum*, Ap<sub>4</sub>A and Ap<sub>4</sub>G levels were not found to change during progression through the cell cycle, but increased 3–7-fold after a 1-hour treatment with 0.1 mM dinitrophenol (Garrison et al., 1986). Another inducer of oxidative stress is mitomycin C (MMC), a DNA cross-linking agent used in some cancer treatments (Dapa et al., 2017; Turkez et al., 2012). Ap<sub>4</sub>N levels increase in

Chinese hamster AA8 cells, mouse embryo fibroblasts and HeLa cells in response to treatment with MMC (Marriott et al., 2015). Together these data indicate an increase in Ap<sub>4</sub>N level across multiple domains of life after exposure to agents capable of inducing oxidative and genotoxic stress.

## **1.2 Ap<sub>4</sub>A has been implicated in several different signalling mechanisms**

Until recently, the increased concentration of Ap<sub>4</sub>A in response to stress was not associated with cellular responses. The lack of evidence for its involvement in any specific signalling mechanisms led to questions surrounding whether Ap<sub>4</sub>A is indeed a signalling alarmone, or whether it is simply a metabolite produced as a result of DNA damage (Despotović et al., 2017). However, Ap<sub>4</sub>A has now been linked to several signalling mechanisms, including altering gene transcription, modulation of the cyclic guanosine monophosphate (GMP)-AMP synthase-stimulator of interferon genes (cGAS-STING) pathway, GTP biosynthesis and DNA replication (Giammarinaro et al., 2022; Guerra et al., 2020; Luciano and Belasco, 2020; Luciano et al., 2019; Marriott et al., 2015; Yu et al., 2019). These signalling mechanisms will now be discussed in more detail.

### **1.2.1 Ap<sub>4</sub>A and other Ap<sub>4</sub>Ns form protective caps on RNA transcripts**

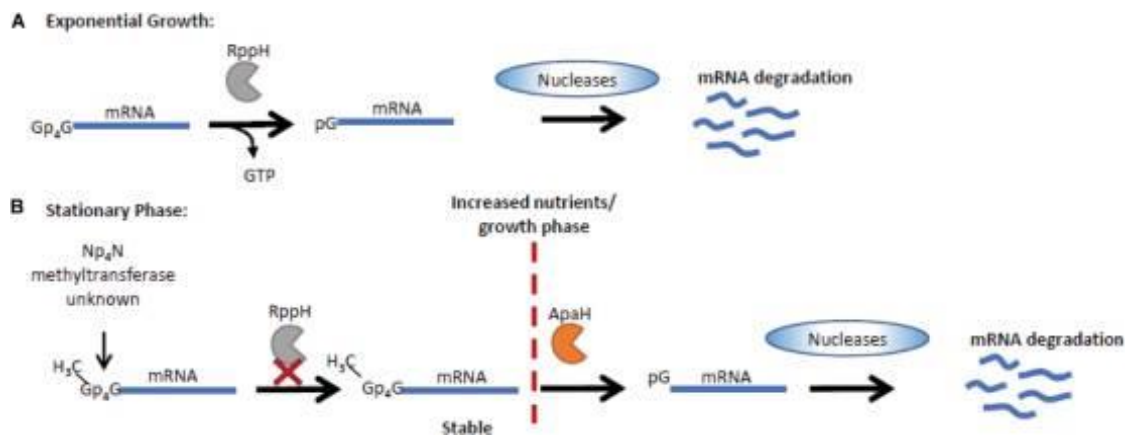
Canonical 5'-caps are co-transcriptionally produced by RNA polymerase II and protect pre-mRNA from degradation during transcription (Galloway and Cowling, 2019; Ramanathan et al., 2016). These caps have many roles in both the nucleus and the cytoplasm, including mRNA processing, nuclear export, translation initiation and mRNA pseudo-circularisation (Ramanathan et al., 2016). They also enable the immune system to discriminate between 'self' and 'non-self' RNAs, and prevent mRNAs from being degraded by 5'–3' exonuclease (Galloway and Cowling, 2019). Over recent years, a range of cellular metabolites and co-enzymes have been found that can form non-canonical 5'-RNA caps. These include the Ap<sub>4</sub>Ns and Np<sub>n</sub>Ns, as well as ADP-ribose, nicotinamide adenine dinucleotide (NAD<sup>+</sup>/NADH), flavin adenine dinucleotide (FAD), uridine diphosphate (UDP)-glucose, UDP-N-acetylglucosamine (UDP-GlcNAc) and dephospho-CoA (Doamekpor et al., 2022; Mattay, 2022; Wiedermannová et al., 2021).

Recently, two groups independently discovered the formation of dinucleotide polyphosphate caps on RNA transcripts and published their data at similar times (Hudeček et al., 2020; Luciano and Belasco, 2020; Luciano et al., 2019).

The Belasco group demonstrated that disulphide stress, known to increase Ap<sub>4</sub>N concentration, induced Np<sub>4</sub> capping of several messenger RNAs (mRNAs) and small RNAs (sRNAs) (Bochner et al., 1984; Luciano et al., 2019). The Np<sub>4</sub> caps are formed through the incorporation of Np<sub>4</sub>N by RNA polymerase during transcription initiation (Luciano and Belasco, 2020). Similarly, the Cahová group demonstrated that T7 RNA polymerase can use dinucleotide polyphosphates (Np<sub>n</sub>Ns) as substrates to produce both methylated and non-methylated Np<sub>4</sub>N caps (Hudeček et al., 2020). This occurs during transcription, where Np<sub>n</sub>Ns are accepted by T7 and *E.coli* RNA polymerases as non-canonical initiating nucleotides (NCINs) (Hudeček et al., 2020). This is similar to the mechanism by which NAD<sup>+</sup> and NADH are incorporated onto transcripts by prokaryotic and eukaryotic RNA polymerases, as these also serve as NCINs during transcription initiation (Bird et al., 2016). Interestingly, Ap<sub>4</sub>Ns are incorporated at a particularly high efficiency when a purine is in the -1 position of the DNA coding strand, and most Ap<sub>4</sub>Ns are incorporated more readily than ATP (Luciano and Belasco, 2020).

The Np<sub>4</sub>-capped *E. coli* transcripts are de-capped by both ApaH and RNA pyrophosphohydrolase (RppH), producing a di- or mono-phosphorylated product, respectively (Luciano et al., 2019). The Cahová group instead suggested that both enzymes produced a mono-phosphorylated product when capped with an Ap<sub>n</sub>N, where n=3 or 4 phosphates (Hudeček et al., 2020). The RppH hydrolysis activity is consistent with previous findings which showed the capability of RppH to hydrolyse the 5'-cap to form a 5'-monophosphate RNA (Song et al., 2013). Of these two proteins ApaH was the major contributor to decapping, while it was suggested that RppH generally removed the remaining phosphate from the dephosphorylated ApaH product (Luciano et al., 2019). The role of both ApaH and RppH is supported by the Cahová group, who demonstrated that although both enzymes are capable of decapping Np<sub>4</sub>N-capped RNA transcripts, only ApaH is capable of hydrolysing methylated caps (Fig. 1.4)

(Hudeček et al., 2020). Generally, methylated  $\text{Np}_n\text{N}$  caps are more abundant in the stationary phase than the exponential phase, suggesting that methylation may be a way for cells to retain RNA when nutrients are low (Fig. 1.4) (Hudeček et al., 2020).



**Figure 1.4: Function of RppH and ApaH in decapping methylated or non-methylated  $\text{Ap}_4\text{N}$  transcripts.** **A:** During exponential growth, RppH can remove  $\text{Np}_4\text{N}$  caps from mRNA, targeting it for degradation. **B:** During the stationary phase, where caps are methylated, RppH is incapable of removing these caps. However, ApaH retains its ability to cleave caps even when methylated and can therefore target mRNA for degradation regardless of methylation status. ApaH,  $\text{Ap}_4\text{A}$  hydrolase;  $\text{Ap}_4\text{N}$ , dinucleotide tetraphosphate containing at least one adenosine moiety; mRNA, messenger RNA;  $\text{Np}_4\text{N}$ , dinucleotide tetraphosphate; RppH, RNA pyrophosphohydrolase.

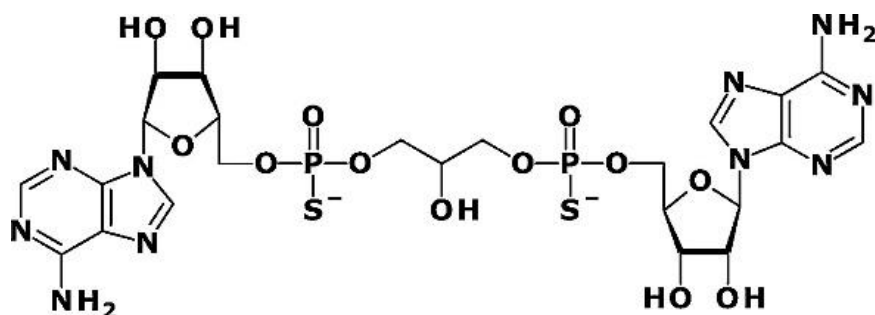
Figure reproduced without changes from *Frontiers in Molecular Biosciences*, Volume 7, Ferguson, F., McLennan, A. G., Urbaniak, M. D., Jones, N. J. & Copeland, N. A., Re-evaluation of Diadenosine Tetraphosphate ( $\text{Ap}_4\text{A}$ ) From a Stress Metabolite to Bona Fide Secondary Messenger, Copyright 2020, <https://doi.org/10.3389/fmolb.2020.606807>; available under the CC BY 4.0 licence (<https://creativecommons.org/licenses/by/4.0/legalcode>) and disclaimer of warranties within.

This incorporation of  $\text{Np}_4\text{Ns}$  into RNA caps suggests that these molecules may induce their effect through altering gene transcription, rather than by binding a specific protein. The methylated caps described by Hudeček and colleagues in 2020 could only be hydrolysed by ApaH and not RppH, suggesting that cells can control the extent of decapping by adjusting the number of transcripts with methylated caps. Furthermore, cadmium or diamide-induced stress greatly reduced the decapping activity of ApaH (Luciano et al., 2019). This suggests a further method of control at times of cell stress

because, in addition to increased Ap<sub>4</sub>A availability, there is also a reduced ability for the removal of any protective RNA caps.

### 1.2.2 Ap<sub>4</sub>A modulates histidine triad nucleotide-binding protein 1 (HINT1) signalling

For many years, Ap<sub>4</sub>A has been implicated in regulation of the HINT1- microphthalmia transcription factor (MITF) signalling pathway. MITF binds to HINT1, previously known as protein kinase C-interacting protein 1 (PKCI-1), resulting in suppression of MITF activity by HINT1 (Razin et al., 1999). In 2004, Lee and colleagues demonstrated that lysyl tRNA synthetase (LysRS), which is known to produce Ap<sub>4</sub>A, also associates with MITF and forms a complex with MITF and HINT1. They also demonstrated that upon mast cell activation, Ap<sub>4</sub>A binds to HINT, resulting in the dissociation of MITF from HINT and subsequently triggering expression of MITF-target genes (Lee et al., 2004). This association between Ap<sub>4</sub>A and HINT1 was initially supported by a report of the crystallographic structure of HINT1 bound to JB419, a non-hydrolysable analogue of Ap<sub>4</sub>A, with the structure shown in Figure 1.5 (Dolot et al., 2016). More recently this analogue was shown to bind to HINT1 and HINT2 with a similar dissociation constant (K<sub>d</sub>) (Dolot et al., 2021).



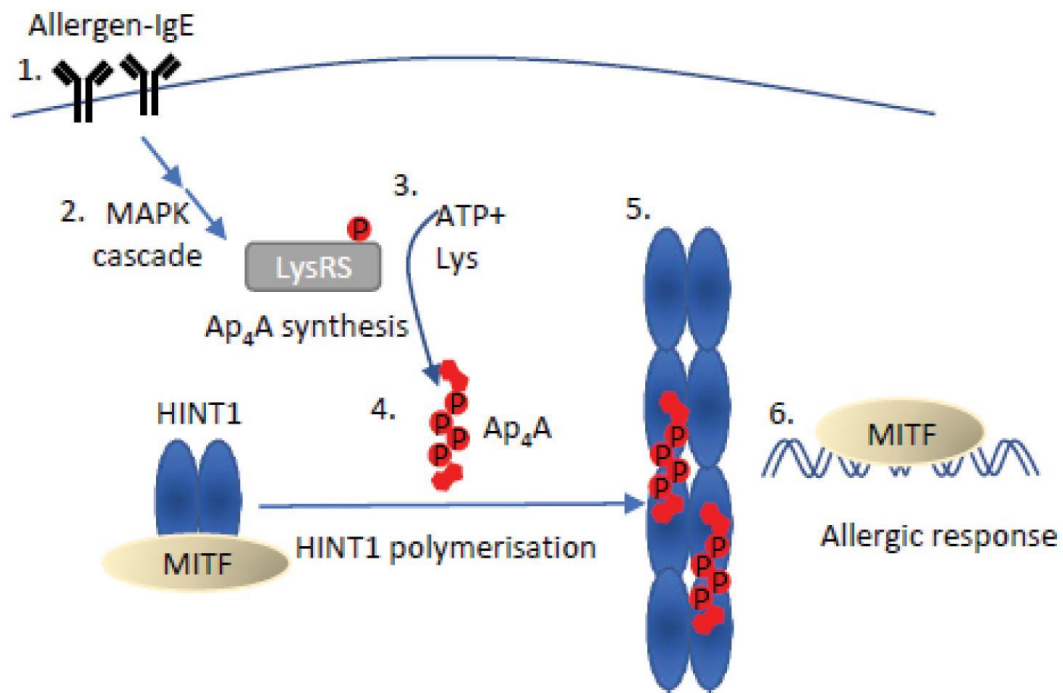
**Figure 1.5: Structure of the non-hydrolysable Ap<sub>4</sub>A analogue JB419.** JB419 is an Ap<sub>4</sub>A analogue, in which the central phosphate chain has been altered. Ap<sub>4</sub>A, diadenosine tetraphosphate.

Figure reprinted from the International Journal of Biological Macromolecules, Volume 87; Dolot, R., Kaczmarek, R., Sęda, A., Krakowiak, A., Baraniak, J. and Nawrot, B., Crystallographic studies of the complex of human HINT1 protein with a non-hydrolyzable analog of Ap<sub>4</sub>A, Pages No. 62–69, Copyright 2016, with permission from Elsevier.

Crystal structures of Ap<sub>4</sub>A bound to HINT1 have now demonstrated that one end of the Ap<sub>4</sub>A molecule binds to an adenosine binding pocket in a HINT1 dimer, while the other



end of the molecule binds to an adenosine binding pocket on a second HINT1 dimer to form a tetramer (Yu et al., 2019). The other subunit of each dimer also has the potential to bind  $\text{Ap}_4\text{A}$ , and linking it to another dimer, and so on, causing HINT1 polymerisation (Fig. 1.6) (Yu et al., 2019).



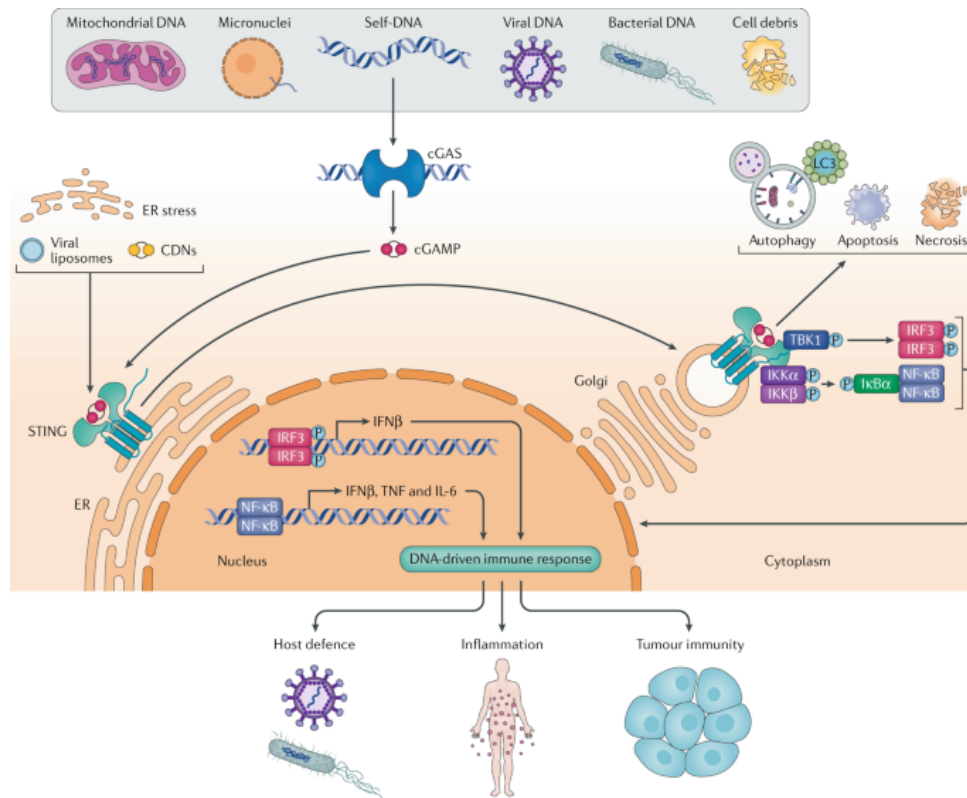
**Figure 1.6: The  $\text{Ap}_4\text{A}$ -induced polymerisation of HINT1.** An allergen (1) initiates the MAPK cascade (2) through binding to IgE. This initiates phosphorylation of LysRS (3), and the subsequent production of  $\text{Ap}_4\text{A}$  (4).  $\text{Ap}_4\text{A}$  displaces MITF from its shared binding site on HINT1. Both ends of the  $\text{Ap}_4\text{A}$  molecule can bind HINT dimers, causing formation of a HINT tetramer. A second  $\text{Ap}_4\text{A}$  molecule can then bind the other molecule in the HINT1 dimer, linking this to another dimer, and so on to form a chain (5). The displaced MITF is then available to bind DNA and induce transcription if its target genes, eliciting the allergic response (6).  $\text{Ap}_4\text{A}$ , diadenosine tetraphosphate; HINT1, Histidine Triad Nucleotide-Binding Protein 1; IgE, immunoglobulin E; MITF, microphthalmia transcription factor.

Figure reproduced without changes from *Frontiers in Molecular Biosciences*, Volume 7, Ferguson, F., McLennan, A. G., Urbaniak, M. D., Jones, N. J. & Copeland, N. A., Re-evaluation of Diadenosine Tetraphosphate ( $\text{Ap}_4\text{A}$ ) From a Stress Metabolite to Bona Fide Secondary Messenger, Copyright 2020, <https://doi.org/10.3389/fmolb.2020.606807>; available under the CC BY 4.0 licence (<https://creativecommons.org/licenses/by/4.0/legalcode>) and disclaimer of warranties within.

Furthermore, Ap<sub>4</sub>A and MITF binding sites overlap, suggesting that Ap<sub>4</sub>A may function by displacing MITF from HINT1 (Yu et al., 2019). However, despite these findings, no binding between Ap<sub>4</sub>A and either HINT1 or HINT2 was detectable by isothermal titration calorimetry (Strom et al., 2020).

### **1.2.3 Ap<sub>4</sub>A competes with cyclic GMP–AMP (cGAMP) to interact with STING**

The cGAS-STING pathway couples the sensing of DNAs and DNA:RNA hybrids, and the stimulation of type 1 interferons and other immune response genes in response to foreign DNA or cell stress (Decout et al., 2021; Guerra et al., 2020; Motwani et al., 2019). In 2013 cGAS, a member of the nucleotidyltransferase family, was first identified as a cytosolic sensor for DNA that produces the second messenger cGAMP (Sun et al., 2013). The cGAMP produced by cGAS has a 2'–5' and 3'–5' phosphodiester linkage, which enables it to bind to STING with a higher affinity than other cGAMP molecules (Ablasser et al., 2013; Zhang et al., 2013). Furthermore, the 2'–5' phosphodiester linkage was found to be essential for the activation of STING (Ablasser et al., 2013). Additionally, cGAMP activates interferon regulatory factor 3 (IRF3) and induces interferon- $\beta$  in a STING-dependent manner (Sun et al., 2013; Wu et al., 2013). A summary of the cGAS-STING immune response is pictured in Figure 1.7 (Motwani et al., 2019).



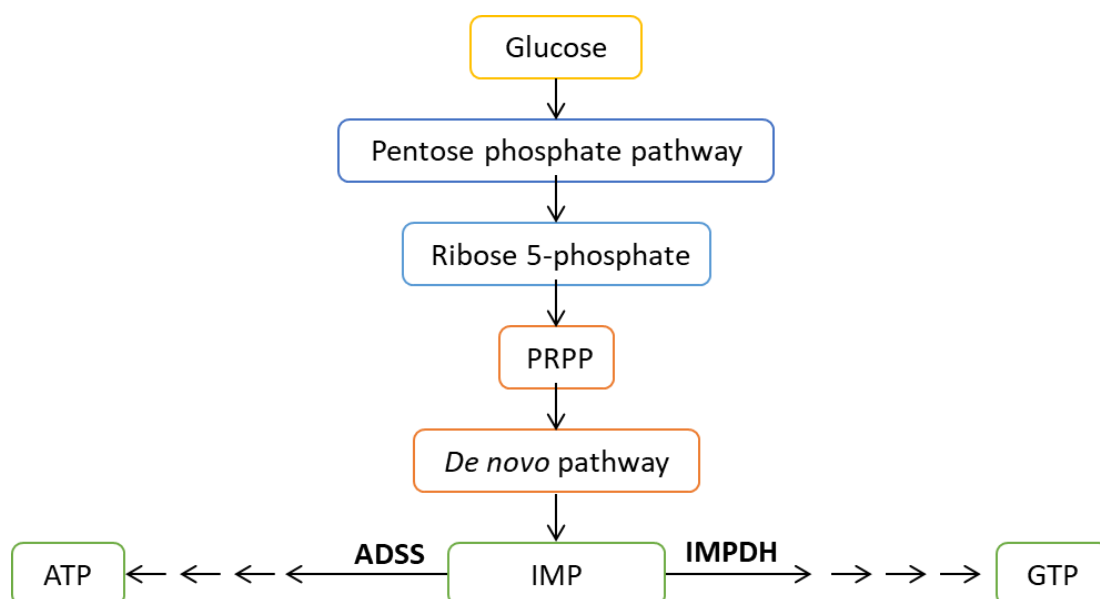
**Figure 1.7: A summary of the cGAS-STING immune response.** Upon detection of different types of DNA, cGAS produces cGAMP. Subsequently, cGAMP binds to STING in the endoplasmic reticulum, activating it. This enables relocation of STING to the golgi apparatus, where it recruits TBK1 and IKK. These molecules phosphorylate IκBα and IRF3. This results in NF-κB and IRF3 relocating to the nucleus, where they induce the transcription of target genes and an immune response. cGAMP, cyclic guanosine monophosphate-adenosine monophosphate; IKK, IκB kinase; IκBα, nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, alpha; STING, Stimulator of interferon genes; TBK1, TANK-binding kinase 1. Figure reproduced with permission from [Springer Nature](#). Motwani, M., Pesiridis, S. & Fitzgerald, K. A. DNA sensing by the cGAS–STING pathway in health and disease. *Nature Reviews Genetics*, Volume 20, Page No. 657–674, 2019, Springer Nature.

Ap<sub>4</sub>A has now been implicated in the regulation of the cGAS-STING pathway. Use of a glutathione S-transferase (GST)-STING<sub>139–378</sub> and immobilized biotinylated Ap<sub>4</sub>A demonstrated that Ap<sub>4</sub>A can bind to STING and reduce STING-dependant signalling (Guerra et al., 2020). Amber force field predicted that cGAMP and Ap<sub>4</sub>A adopt a similar 3D structure when bound to STING, and cGAMP was able to displace Ap<sub>4</sub>A from STING, consistent with Ap<sub>4</sub>A and cGAMP sharing the same binding site on STING that may modulate its activity (Guerra et al., 2020).

In addition to DNA, DNA:RNA hybrids are detected by the cGAS-STING pathway, where, upon their recognition, cGAS produces cGAMP (Mankan et al., 2014). cGAS and LysRS both bind RNA:DNA hybrids with similar affinity and knockdown of LysRS increases intracellular cGAMP levels, suggesting that LysRS delays recognition of the hybrids by cGAS. This is particularly interesting as since the 1960s LysRS has been associated with Ap<sub>4</sub>A synthesis (Zamecnik et al., 1966). Together these data suggest that LysRS and Ap<sub>4</sub>A regulate the cGAS-STING pathway by interfering with cGAS binding, both in the initial detection of RNA:DNA hybrids and also by delaying cGAS binding to STING. The involvement of LysRS in the localisation of Ap<sub>4</sub>A to STING draws a parallel with the HINT1-MITF signalling pathway where LysRS is involved in binding to MITF, suggesting a common mechanism whereby the Ap<sub>4</sub>A-producing LysRS is well positioned to cause an effect through Ap<sub>4</sub>A production.

#### **1.2.4 Ap<sub>4</sub>A modulates the production of xanthosine 5'-monophosphate (XMP) from inosine 5'-monophosphate (IMP) in GTP biosynthesis**

During *de novo* GTP biosynthesis, 20 reactions are performed requiring a total of 9 ATPs (Kofuji and Sasaki, 2020). Once IMP has been produced, the pathway can diverge into either ATP or GTP biosynthesis, depending on whether IMP is converted into S-AMP by adenylosuccinate synthase (ADSS) or into XMP by inosine-5'-monophosphate dehydrogenase (IMPDH) (Kofuji and Sasaki, 2020). A simple schematic for the *de novo* GTP and ATP synthesis pathways is shown in Figure 1.8.



**Figure 1.8: A schematic showing the *de novo* synthesis of GTP and ATP.** GTP biosynthesis initially proceeds via the pentose phosphate pathway in which there is a stepwise conversion of glucose into ribose 5-phosphate. The production of PRPP from ribose 5-phosphate marks the start of the *de novo* pathway. During this pathway PRPP goes through several steps requiring ATP, finally producing IMP. Subsequently either ADSS or IMPDH can react with IMP, triggering the pathways that convert IMP into either ATP or GTP, respectively. ATP, adenosine triphosphate; ADSS, adenylosuccinate synthase; GTP, guanosine triphosphate; IMPDH, inosine-5'-monophosphate dehydrogenase; PRPP, phosphoribosyl pyrophosphate.

IMPDH monomers consist of the catalytic domain and a  $(\beta/\alpha)_8$  barrel (Hedstrom, 2009). Most also possess a subdomain composed of two cystathionine-beta-synthase (CBS) domains, also known as a Bateman domain (Hedstrom, 2009). A study using a biotinylated Ap<sub>4</sub>A molecular hook identified Ap<sub>4</sub>A as a potential binding partner for IMPDH (Guo et al., 2011). This was supported by a second paper which identified *E. coli* IMPDH as a binding partner for Ap<sub>4</sub>A (Despotović et al., 2017). It was later demonstrated that adenosine and guanosine dinucleotide polyphosphates regulate the activity of IMPDH *in vitro* through binding to the Bateman domain and competing with adenine or guanine mononucleotides (Fernández-Justel et al., 2019). Furthermore, the binding of Ap<sub>4</sub>A to the Bateman domain is supported by another paper, which demonstrated that Ap<sub>4</sub>A bound to *B. subtilis* IMPDH in the cleft between the two CBS domains of each monomer (Giammarinaro et al., 2022).

There are two classes of bacterial IMPDHs, named class I and class II, which can be distinguished according to their kinetics and quaternary structure (Alexandre et al., 2015). By this classification, class I IMPDHs are octameric, cooperative enzymes which are activated by MgATP; while class II IMPDHs have Michaelis-Menten kinetics and are tetramers which shift to octamers in the presence of MgATP or NAD (Alexandre et al., 2015). *B. subtilis* IMPDH is a class II enzyme, and interestingly the crystal structure of Ap<sub>4</sub>A-bound *B. subtilis* IMPDH suggested that two IMPDH tetramers form an octamer, stabilised by Arg-141 and 144 in the CBS domains, which interact with the Ap<sub>4</sub>A of the joining tetramer (Giammarinaro et al., 2022). As octamer formation correlates with reduced IMPDH activity, and Ap<sub>4</sub>A-induced octamerization alters the conformation of the Cys-308-containing active site loop, the authors suggested that Ap<sub>4</sub>A inhibits IMPDH activity by altering the conformation of the active site (Giammarinaro et al., 2022). Contrary to this, in *Pseudomonas fluorescens* with null mutations in ApaH there was an increase in c-di-GMP, which was associated with increased GTP levels, suggesting that in this case Ap<sub>4</sub>A is promoting rather than inhibiting GTP biosynthesis (Monds et al., 2010).

Ap<sub>4</sub>A shares its binding site in the Bateman domain with ATP, which is present in cells at a much higher concentration. However, in *B. subtilis*, Ap<sub>4</sub>A and other Ap<sub>n</sub>As have a much higher affinity for IMPDH (Ap<sub>4</sub>A  $K_d = 7.4 \pm 2.1 \mu\text{M}$ ) compared with adenosine mononucleotides AMP ( $K_d = 18 \text{ mM}$ ), ADP (no interaction) and ATP ( $K_d = 5 \text{ mM}$ ) (Giammarinaro et al., 2022). This differs from in *E. coli*, where the affinity of Ap<sub>4</sub>A for IMPDH is only five-fold higher than that of ATP for IMPDH (Despotović et al., 2017). When considered in terms of the physiological concentration of ATP compared with Ap<sub>4</sub>A, this suggested that the binding of Ap<sub>4</sub>A to IMPDH is physiologically irrelevant in *E. coli* (Despotović et al., 2017). Interestingly, Ap<sub>4</sub>A did not affect *E. coli* IMPDH activity, whereas increased Ap<sub>4</sub>A concentration correlated with reduced *B. subtilis* IMPDH activity, suggesting that it is functioning differently in *E. coli* and *B. subtilis* (Despotović et al., 2017; Giammarinaro et al., 2022). Furthermore in *B. subtilis*, although ATP and Ap<sub>4</sub>A can both bind and restrict the flexibility of the CBS domains, in the presence of substrates Ap<sub>4</sub>A promotes octamer formation with a 250-fold lower EC<sub>50</sub> compared

with ATP, suggesting why Ap<sub>4</sub>A has a more inhibitory effect than ATP in these cells (Giammarinaro et al., 2022).

Therefore, Ap<sub>4</sub>A appears to have a function in regulating the GTP biosynthesis pathway through its interaction with IMPDH in *B. subtilis* cells. It is interesting that Ap<sub>4</sub>A appears to be important in the formation of the less active class I IMPDH octamers from the tetrameric apo-form, as this is reminiscent of the role of Ap<sub>4</sub>A in HINT1 signalling discussed in section 1.2.2, where it forms HINT tetramers and oligomers by linking dimers (Yu et al., 2019). However, unlike in HINT signalling where the adenine moieties themselves bind to and link the different dimers, in the case of IMPDH it appears that the octamers are formed by interactions between the CBS domain of one tetramer and the phosphates of the Ap<sub>4</sub>A in the second tetramer (Giammarinaro et al., 2022; Yu et al., 2019).

### 1.2.5 Biofilm formation

Another example of Ap<sub>4</sub>A influencing gene expression comes with biofilm formation. In *Streptococcus mutans*, deletion of the Ap<sub>4</sub>A hydrolase gene, *YqeK*, resulted in reduced biofilm formation and reduced water-insoluble exopolysaccharide production (Zheng et al., 2022). In the *YqeK* mutant, expression of biofilm formation-related genes *gtfB*, *gtfC* and *gbpC* were inhibited and this translated to a reduction in the abundance and activity of the GTF proteins (Zheng et al., 2022). Conversely, an *apaH* mutant of *P. fluorescens* possessing increased Ap<sub>4</sub>A levels showed a 2-fold increase in biofilm formation compared with the wild-type, suggesting that Ap<sub>4</sub>A may have a role in promoting biofilm formation in *P. fluorescens* (Monds et al., 2010). This is interesting as it provides another example of Ap<sub>4</sub>A having the opposite effect in *P. fluorescens* compared with another species, as was the case in GTP biosynthesis. Furthermore, in a *pst* mutant of *P. fluorescens* which has severely reduced biofilm formation, mutation of *apaH* resulted in biofilm formation being partially restored (Monds et al., 2010). Together, these data suggests that Ap<sub>4</sub>A may regulate biofilm formation through altered gene expression, but that the way in which Ap<sub>4</sub>A has its effect may differ between organisms.

### 1.2.6 DNA replication

Since its discovery, Ap<sub>4</sub>A has been highly implicated in the DNA replication process, although how Ap<sub>4</sub>A functions in this process is not yet fully understood. In 1978, Ap<sub>4</sub>A was reported to stimulate DNA synthesis in baby hamster kidney cells arrested in the first gap phase (G1) (Grummt, 1978b). Supporting this, the Ap<sub>4</sub>A pool in both baby hamster kidney fibroblasts and mouse 3T3 cells increased 1000-fold throughout G1 phase, reaching maximum levels in S-phase (Weinmann-Dorsch et al., 1984a). Similarly, an 8- to 30-fold increase in Ap<sub>4</sub>A level in *P. polycephalum* is seen at entry into the synthesis (S) phase (Weinmann-Dorsch et al., 1984b). This role for Ap<sub>4</sub>A in DNA replication is supported by evidence that Ap<sub>4</sub>A interacts with the 57 kDa subunit of DNA polymerase alpha (pol  $\alpha$ ) (Grummt et al., 1979). Pol  $\alpha$  is an essential component of the replisome, without which DNA cannot be synthesised (Yeeles et al., 2015). This therefore supports the hypothesis that Ap<sub>4</sub>A is involved in the DNA replication process.

However, other reports have suggested that Ap<sub>4</sub>A may not initiate DNA replication. An investigation into the role of Ap<sub>4</sub>A in sea urchin embryos reported an abrupt decrease in Ap<sub>4</sub>A levels before each S phase (Morioka and Shimada, 1985). More recently evidence has been reported which suggests that Ap<sub>4</sub>A instead inhibits the initiation of DNA replication (Marriott et al., 2015). In this paper, a cell-free DNA replication system was used to demonstrate that Ap<sub>4</sub>A inhibited the initiation of replication in G1-arrested nuclei, but had no inhibitory effect on the elongation of cells that were already replicating at the time of the Ap<sub>4</sub>A addition (Marriott et al., 2015). Unlike DNA pol  $\alpha$ , the presence of Ap<sub>4</sub>A is not essential for eukaryotic DNA replication to occur (Yeeles et al., 2015). This is demonstrated by the fact that no exogenous Ap<sub>4</sub>A, nor any the enzymes required for Ap<sub>4</sub>A synthesis, were added to the *in vitro* system used by Yeeles and colleagues (2015) to determine the essential DNA replication components. As Ap<sub>4</sub>A levels increase under stress, Ap<sub>4</sub>A may be a damage-associated regulator of DNA replication (Ferguson et al., 2020; Marriott et al., 2015). This is supported by evidence of a role for Ap<sub>4</sub>A in apoptosis during the early stages of G1/S arrest (Vartanian et al., 2003).



### 1.2.7 Ap<sub>4</sub>A often, but not always, exerts its effect through binding ATP binding sites

In the signalling mechanisms discussed above, Ap<sub>4</sub>A frequently exerts its effect by competing for nucleoside binding sites that can also be occupied by other molecules. In the cGAS-STING pathway, Ap<sub>4</sub>A competes with cGAMP for binding to STING, while in the HINT signalling pathway it competes with MITF for the adenosine binding pockets on HINT (Guerra et al., 2020; Yu et al., 2019). Similarly, in the bacterial RNA capping mechanism, Ap<sub>4</sub>Ns outcompete ATP to be incorporated as protective caps (Luciano and Belasco, 2020). ATP and Ap<sub>4</sub>A are also both capable of binding CBS domains, such as those forming the subdomain of most IMPDH enzymes; yet another example of Ap<sub>4</sub>A sharing its binding site with other nucleotides (Giammarinaro et al., 2022; Tseng et al., 2011). A summary of the Ap<sub>4</sub>A binding partners and corresponding signalling mechanisms is described in Table 1.2.

In *E. coli*, other binding partners for Ap<sub>4</sub>A have been identified, including DnaK, GroEL, E89 (ClpB) and NrdR (Despotović et al., 2017; Fuge and Farr, 1993; Johnstone and Farr, 1991). In the case of the molecular chaperone GroEL, it was demonstrated that Ap<sub>4</sub>A binds to a site that is distinct from the known ADP and ATP binding sites (Tanner et al., 2006). NrdR is responsible for the expression of ribonucleotide reductase genes in *E. coli*, and in addition to Ap<sub>4</sub>A it is also capable of binding ATP (McKethan and Spiro, 2013; Torrents, 2014). Interestingly, a recent study investigating the proteins interacting with Ap<sub>3</sub>A and Ap<sub>4</sub>A in HEK293T cells demonstrated that 46% of the identified Ap<sub>n</sub>A binding proteins were also known to bind nucleotides, and more specifically 32% of them were known to bind ATP (Krüger et al., 2021). Similarly, in *E. coli*, 55% were nucleotide binding proteins, and 40% were known ATP binding proteins (Krüger et al., 2021).

A binding partner has not yet been linked to a mechanism for the role Ap<sub>4</sub>A plays in the inhibition of DNA replication. It is possible that Ap<sub>4</sub>A affects DNA replication through its interaction with Pol  $\alpha$ . Alternatively, the high frequency of cases where Ap<sub>4</sub>A exerts its effect by binding to known nucleotide binding sites suggests that it is also possible that Ap<sub>4</sub>A may compete with the binding of ATP or another nucleotide at some point in the replication process.

**Table 1.2: A summary of the Ap<sub>4</sub>A binding partners and related signalling mechanisms.** CBS, cystathionine  $\beta$ -synthase; cGAS, cyclic GMP-AMP synthase; HINT1, Histidine triad nucleotide-binding protein 1; IMPDH, Inosine-5'-monophosphate dehydrogenase; STING, Stimulator of interferon genes.

<b>Signalling Pathway</b>	<b>Organism</b>	<b>Binding partner</b>	<b>Mode of inhibition</b>	<b>Effect</b>	<b>Reference</b>
<b>mRNA capping</b>	<i>E. coli</i>	RNA pol	Competes with ATP for incorporation into transcripts	Protects RNA transcripts from degradation	(Luciano and Belasco, 2020; Luciano et al., 2019)
<b>STING signalling</b>		STING	Competes with cGAS	Displaces or slows down interaction between cGAS and STING	(Guerra et al., 2020)
<b>HINT signalling</b>		HINT1 dimers	Binds to adenosine binding pockets	Polymerises HINT1, results in dissociation of MITF and induction of target genes	(Yu et al., 2019)
<b>GTP biosynthesis</b>		IMPDH	Binds to CBS domains which can also bind ATP	Promotes formation of less active IMPDH octamers	(Giammarina et al., 2022)
<b>DNA replication</b>	Mouse, Human	Unknown	Unknown	Inhibits the initiation of DNA replication	(Marriott et al., 2015)

### 1.3 Other Ap<sub>4</sub>Ns also have known signalling roles

Ap<sub>4</sub>A is the most extensively studied of the Ap<sub>4</sub>Ns, probably because it is present at a higher concentration in cells than the other Ap<sub>4</sub>Ns (Coste et al., 1987). However, while some of the signalling mechanisms discussed in the previous section rely on the homobifunctional structure of Ap<sub>4</sub>A to perform their function, such as HINT1 oligomerisation (Yu et al., 2019), this is not the case with others. Capping of mRNA transcripts is not limited to Ap<sub>4</sub>A, as caps were also formed from adenosine cytidine tetraphosphate (Ap<sub>4</sub>C), Ap<sub>4</sub>G and adenosine uridine tetraphosphate (Ap<sub>4</sub>U) (Luciano and Belasco, 2020). In this case, the incorporation efficiency of the different Ap<sub>4</sub>Ns varied depending on the nucleotides involved in the promoter sequence (Luciano and Belasco, 2020).

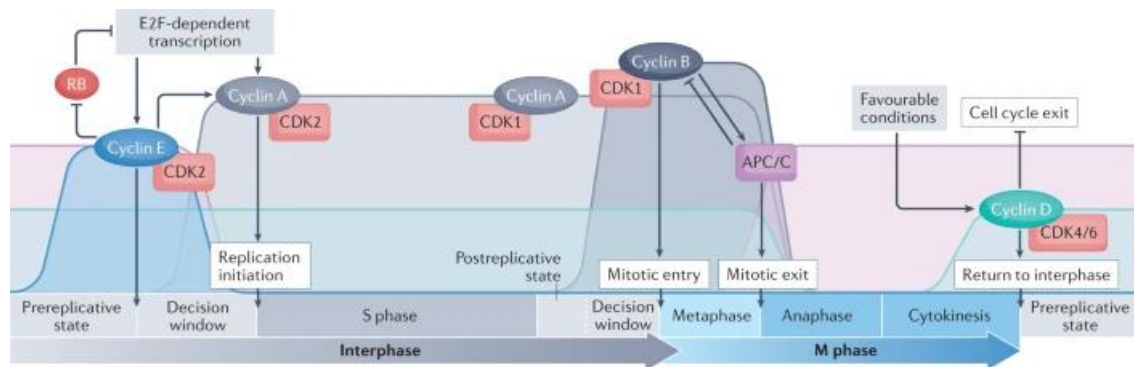
There is also evidence that other Ap<sub>4</sub>Ns are involved in other signalling mechanisms. For example, Ap<sub>4</sub>U has a documented extracellular role in the cardiovascular system, where it is synthesised by vascular endothelial growth factor receptor 2 (VEGFR2) (Jankowski et al., 2013a; Zhou et al., 2019b). Ap<sub>4</sub>U and Ap<sub>4</sub>A are also inhibitors of uridine kinase in Ehrlich ascites tumour cells, while adenosine deoxythymidine tetraphosphate (Ap<sub>4</sub>dT) inhibited thymidine kinase in acute myeloid leukaemia blast cells (Bone et al., 1986b; Cheng et al., 1986). These examples suggest evidence for Ap<sub>4</sub>N activity in both intracellular and extracellular signalling. Furthermore, many studies referring to Ap<sub>4</sub>A used an ATP luciferase-based assay for quantification of Ap<sub>4</sub>A levels. This measures ATP production after hydrolysis and therefore would not distinguish Ap<sub>4</sub>A from other Ap<sub>4</sub>Ns, suggesting that in some cases the contribution of other Ap<sub>4</sub>Ns could have previously been overlooked.

#### **1.4 Regulation of the cell-cycle**

In this project, the role of Ap<sub>4</sub>A and Ap<sub>4</sub>Ns in the regulation of DNA replication will be evaluated. As Ap<sub>4</sub>A has been associated with inhibition of the initiation phase of DNA replication, this activity could be mediated at several transitions prior to replisome assembly. This section will describe the temporal assembly of protein complexes that license and facilitate initiation of DNA replication to provide context to key regulatory events that may be affected by Ap<sub>4</sub>A. This section will then be followed by an overview of the responses to genotoxic damage and DNA repair.

Despite the high fidelity of the DNA replication process, mutation can still occur, and whole-genome studies have shown that the average mutation rate in humans is  $0.96 \times 10^{-8}$ – $1.20 \times 10^{-8}$  mutations per base pair per generation. Therefore, precise regulation of the cell cycle is essential to maintain genome stability. Without this, the cell-cycle can become deregulated, leading to sustained proliferative signalling and genome instability, both of which are known hallmarks of cancer (Hanahan and Weinberg, 2011). Therefore, to prevent mutations from being introduced into the genome, the cell duplication process is tightly controlled, with cell cycle progression relying on changing concentrations of various cyclins and cyclin-dependent kinases (CDKs) (Matthews et al., 2022). There are four stages to the eukaryotic cell cycle:

G1-phase, S-phase, the second gap phase (G2) and the mitotic phase (M) (Fig. 1.9) (Duronio and Xiong, 2013). Briefly, progression from G<sub>1</sub> phase is controlled by cyclin D-CDK4/6 and cyclin E-CDK2, but the specific timings and the extent to which these two complexes overlap is still being investigated (Fig. 1.9) (Matthews et al., 2022).



**Figure 1.9: Progression through the eukaryotic cell cycle.** The cell cycle is split into four phases – G<sub>1</sub>, S, G<sub>2</sub> and M. Different levels of the cyclin/CDK complexes are required for progression through each stage of the cell cycle. The prereplicative state occurs in G<sub>1</sub> phase. During this phase, cyclin E and CDK2 accumulate, creating a decision window in which the cell can transition into S-phase. Cyclin A then also begins to accumulate and associate with CDK2, and replication initiates. After S-phase, cyclin A/B-CDK1 complexes form and promote the progression of the cell cycle through a second decision window and into mitosis. Under favourable conditions, Cyclin D-CDK4/6 begins to accumulate at the end of mitosis, allowing the cell to enter G<sub>1</sub> once again.

Reproduced with permission from [Springer Nature](#). Matthews, H. K., Bertoli, C. & De Bruin, R. A. M., Cell cycle control in cancer. *Nature Reviews Molecular Cell Biology*, Volume 23, Pages No. 74–88, 2022, Springer Nature.

In early G<sub>1</sub>, cyclin D-CDK4/6 exclusively monophosphorylates the retinoblastoma tumour suppressor protein (Rb), forming a functionally active Rb isoform (Narasimha et al., 2014). In this model, the monophosphorylated Rb is hyperphosphorylated by cyclin E-CDK2 in late G<sub>1</sub> promoting dissociation of Rb from E2F, allowing the transcription of E2F target genes (Narasimha et al., 2014). However, recent findings suggest that CDK4/6 activity is essential for sustaining Rb phosphorylation prior to the onset of S-phase (Chung et al., 2019). CDK4/7 inhibition shows that in the absence of CDK4/6, replication stress and G<sub>1</sub> arrest is induced, replisome components are downregulated, and origin licensing is reduced (Crozier et al., 2022). Furthermore,

release from this arrest frequently results in p53-dependent exit from the cell cycle (Crozier et al., 2022). The initiation of S-phase depends on the accumulation of cyclin A-CDK2 and the inactivation of the anaphase promoting complex (APC/C<sup>CDH1</sup>) by cyclin E-CDK2 and early mitotic inhibitor 1 (EMI1), which causes the cell to fully commit to proliferation and begin DNA replication (Cappell et al., 2016; Cappell et al., 2018). After completion of S-phase, CDK1 is then activated by associating with cyclin A and cyclin B, facilitating bypass of the G2 checkpoint and subsequently entry into mitosis (Matthews et al., 2022). During mitosis, the nuclear envelope breaks down, the duplicated DNA condenses and the resulting sister chromatids are separated, leading to the generation of two identical daughter cells (Hirano, 2015).

### **1.5 Regulation of DNA replication**

Bidirectional synthesis from 30,000–50,000 replication origins allow the replication of the human genome to occur in approximately 6–8 hours (Masai et al., 2010; Méchali, 2010). Origin licencing begins with the loading of the pre-replication complex (Pre-RC) onto DNA (Pozo and Cook, 2016). The CDK and Cdc7/Dbf4-dependent kinases (DDKs) then trigger the formation of replication protein complexes, contributing to a timely activation of the replisome (Sheu and Stillman, 2006; Tanaka et al., 2007). The minimum number of proteins essential for DNA replication to occur have been determined, and consist of Cdt, the minichromosome maintenance complex (MCM) proteins 2-7 (Mcm2–7), the origin replication complex (ORC), Cdc6, DDK, Sld3/7, Cdc45, S-CDK, Dpb11, Sld2, GINS, Mcm10, Pol  $\alpha$ , DNA polymerase epsilon (Pol  $\epsilon$ ), Ctf4, replication protein A (RPA) and topoisomerase 2 (Topo II) (Yeeles et al., 2015). However, to achieve maximal replication rates Csm3/Tof1, Mrc1, proliferating cell nuclear antigen (PCNA), replication factor C (RFC) and DNA polymerase delta (Pol  $\delta$ ) are also needed (Yeeles et al., 2017).

#### **1.5.1 Formation of the pre-replication complex**

The formation of the pre-RC is a dynamic process which results in the Mcm2–7 helicase being loaded onto DNA origins (Tsakraklides and Bell, 2010). The licensing of DNA replication origins is a multistep process that begins with the origin recognition complex binding to putative origins. The origin recognition complex is capable of

recognising origins and binding DNA in an ATP-dependent process (Bell and Stillman, 1992). A cryo-electron microscopy structure for DNA-bound ORC shows ORC wrapping around the DNA through interactions with both the DNA backbone and bases (Li et al., 2018). In budding yeast, the Lys-362 and Arg-367 residues of Orc1 are essential for ORC to bind DNA (Kawakami et al., 2015). DNA-bound ORC then forms a complex with Cdc6, which displaces the Orc2 winged helix domain (WHD) from the gap in the ring-shaped structure formed by the AAA+ domains of Orc1–5 (Feng et al., 2021). As Cdc6 also contains an AAA+ domain, this results in the formation of a closed ring of AAA+ domains (Feng et al., 2021). Binding of Cdc6 to the ORC-DNA alters the positioning of a helix-turn-helix motif in Orc1 by approximately 5 Å, forming part of the ATP binding site on Cdc6 and activating its ATPase activity (Feng et al., 2021). This ATPase activity helps to modulate formation of the MCM-ORC-Cdc6 complex because ATPase is suppressed by origin DNA but not non-origin DNA and this activity is also required for dissociation of Cdc6 after formation of the pre-RC (Chang et al., 2015; Speck and Stillman, 2007).

During G1 phase, Cdt1 and Mcm2–7 accumulate in the nucleus and associate with one another (Tanaka and Diffley, 2002). After Cdc6 binds to ORC-DNA, the Cdt1-Mcm2–7 complex binds to the ORC-Cdc6-DNA complex via the WHDs of Mcm3 and Mcm7, which act as a double anchor to form a semi-attached ORC-Cdc6-Cdt1-MCM (OCCM) complex (Yuan et al., 2020). This initial binding of Mcm2–7 is made possible by the structural changes induced by Cdc6 binding, opening Mcm3 and Mcm7 binding sites (Feng et al., 2021). Subsequently, the Mcm2–7 complex binds to ORC-Cdc6, forming the pre-insertion OCCM, resulting in DNA bending by 60°, leaving it positioned between ORC and Mcm2–7, adjacent to the DNA entry gate (Yuan et al., 2020). For the helicase to load onto DNA, the Mcm2/Mcm5 entry gate needs to be opened (Samel et al., 2014). This association between Mcm2–7 and chromatin is dependent on Mcm10 and its interaction with Mcm7 (Homesley et al., 2000). The origin DNA is inserted so that it passes through Mcm2–7 ring as well as the ORC-Cdc6 ring, forming the OCCM complex prior to assembly of the double hexamer (Samel et al., 2014; Yuan et al., 2017). For dimerisation of the Mcm2–7 complex to occur, Cdt1 must first be released

from the OCCM complex (Evrin et al., 2014). Cdt1 is released from the OCCM in a Cdc6 and Orc1 ATPase-dependent manner, resulting in the ORC-Cdc6-MCM (OCM) complex which is competent for Mcm2–7 dimerisation (Evrin et al., 2014; Fernández-Cid et al., 2013). The OCM then associates with a second Mcm2–7 hexamer, to form the ORC-Cdc6-MCM-MCM (OCMM) intermediate (Sun et al., 2014).

Recently, a further intermediate was found, Mcm2–7-ORC (MO), which lacked Cdc6 and Cdt1 and in which the Mcm3/5 gate is closed (Miller et al., 2019). Release of Cdc6 from the OCM complex requires its ATPase activity (Chang et al., 2015). The MO complex is fully formed when a second ORC associates with the N-terminal side of the loaded MCM complex and is followed by recruitment of a second Cdc6 and MCM-Cdt1 to form an Mcm2–7-ORC-Cdc6 (MOC)-MC complex (Miller et al., 2019). Recruitment of second Cdt1 and ORC molecules is consistent with previous evidence which demonstrated that multiple molecules are required for the formation of the pre-RC (Coster and Diffley, 2017; Takara and Bell, 2011). Furthermore, a cryo-electron microscopy (cryo-EM) image showing MCM sandwiched between two ORC complexes demonstrates that in some cases the second ORC binds before complete dissociation of the first (Miller et al., 2019). Finally, dissociation of the other proteins is repeated, leaving two Mcm2–7 hexamers linked at their N-terminal rings, forming the inactive Mcm2–7 double hexamer (Remus et al., 2009).

### **1.5.2 Formation of the pre-initiation (pre-IC) complex**

After formation of the pre-RC, Cdc45, Sld2, Sld3, Dpb11, GINS and Pol  $\epsilon$  accumulate at the origin of replication and the Mcm2–7 double hexamer splits to become two single hexamers, forming the pre-IC (Miyazawa-Onami et al., 2017). This process depends on, and is regulated by, the phosphorylation of several molecules by different protein kinases, including DDK and CDK.

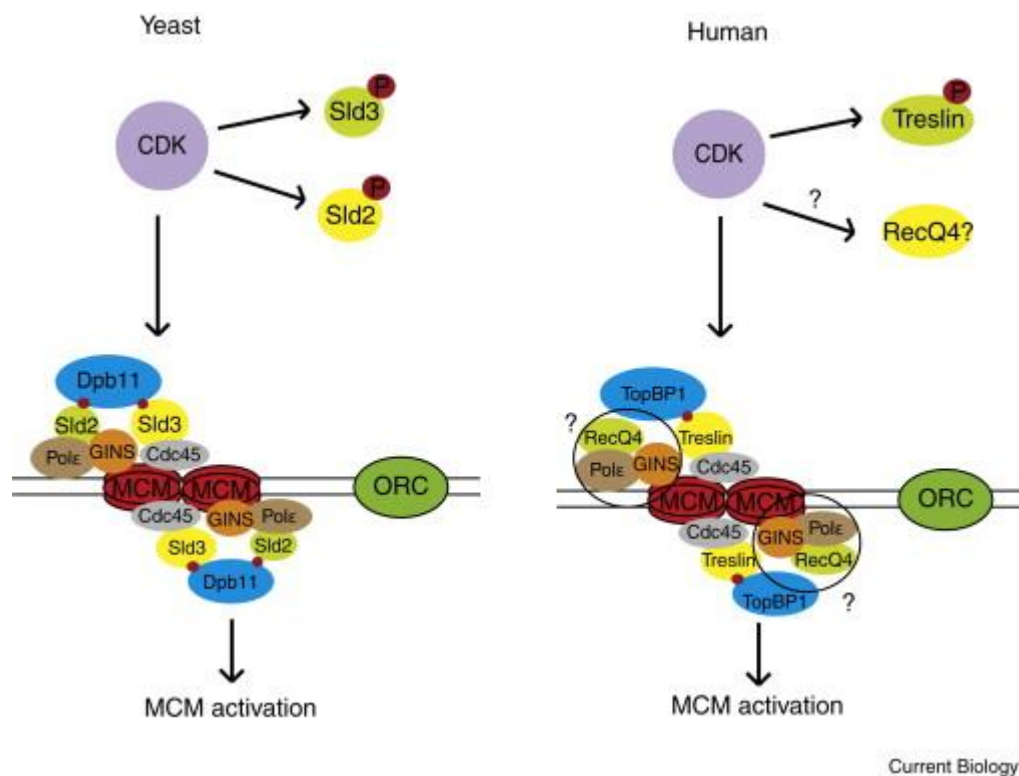
Firstly, after the pre-RC has formed it must be phosphorylated before the DDK can bind (Francis et al., 2009). Then, up to two molecules of DDK can bind independently to the Mcm2–7 complex through their Dbf4 subunit (Cheng et al., 2022; Saleh et al.,

2022). Dbf4 uses its helix-BRCA1 C-terminal (HBRCT) domain to anchor to Mcm2 and associates with the N-terminal domain A (NTD-A) subdomains of Mcm4 and Mcm6 of one hexamer and Mcm2 of the opposite hexamer, demonstrating that both hexamers are required for DDK binding (Cheng et al., 2022; Saleh et al., 2022). Upon docking, DDK phosphorylates Mcm4, removing the inhibitory effect induced by the amino terminal serine/threonine-rich domain (NSD) of Mcm4 to promote progression through S-phase (Sheu and Stillman, 2010; Sheu and Stillman, 2006). In addition to its intrinsic and phosphorylation-generated (PG) sites on Mcm4, DDK also targets intrinsic and PG sites on Mcm6 (Randell et al., 2010). This process is regulated by the checkpoint kinase Rad53, which inhibits DDK-MCM binding independently of Rad53 kinase activity (Abd Wahab and Remus, 2020). Once DDK has bound and phosphorylated the Mcm2–7 complex, Sld3 can bind to DDK-phosphorylated sites on Mcm4 and Mcm6, where it has been suggested to promote the recruitment of Cdc45 (Deegan et al., 2016). However, this is in contradiction with a previous study, which suggested that Sld3 and Cdc45 form a complex then bind to replication origins simultaneously in a mutually-dependent manner (Kamimura et al., 2001)..

The next steps in the formation of the pre-IC are regulated by S-phase CDK. Current data suggest that phosphorylation of Sld2 by S-CDK is necessary for formation of the Sld2-Dpb11 complex during S-phase (Masumoto et al., 2002). CDK-induced phosphorylation of canonical motifs regulate phosphorylation of residue Threonine84 (Thr84), which, when phosphorylated, is responsible for Sld2-Dpb11 complex formation (Tak et al., 2006). When not bound to replication origins, CDK-phosphorylated Sld2 and Dpb11 form a complex with GINS and Pol  $\epsilon$ , forming the pre-loading complex (Pre-LC) (Muramatsu et al., 2010). GINS consists of Sld5, Psf1, Psf2 and Psf3 and its association with Dpb11 and Cdc45 is essential in order for binding of Dpb11, Cdc45 or GINS to origin DNA (Takayama et al., 2003). Meanwhile, Sld3 is phosphorylated by CDK at Thr-600 and Ser-622, creating a binding site which enables the association of Sld3 with the N-terminal pair of BRCT domains on Dpb11, forming the pre-IC (Tanaka et al., 2007; Zegerman and Diffley, 2007). CDK is also important in humans, where it phosphorylates Treslin (the human homologue of Sld3) on S1000, which enables the association between Treslin and DNA topoisomerase II $\beta$ -binding



protein 1 (TOPBP1, the human homologue of Dpb11) (Kumagai et al., 2011; Mueller et al., 2011). TOPBP1 interacts directly with Cdc45 and facilitates its loading onto DNA during the G1/S transition (Schmidt et al., 2008). Unlike Sld2, the human homologue RecQ4 is not regulated by CDK (Mueller et al., 2011). A summary of the role of CDK in activating the replicative helicase is shown in Fig. 1.10. A recent report suggested that formation of the pre-IC is sufficient for the splitting of the double hexamer, contradicting former evidence which suggested that Mcm10 was needed for this to occur (Miyazawa-Onami et al., 2017; Quan et al., 2015).



**Figure 1.10: CDK-induced activation of the DNA helicase.** In yeast, CDK phosphorylates Sld2, and Sld3 allowing association of these proteins with Dpb11 and subsequent formation of the pre-LC and pre-IC. In humans, CDK also phosphorylates Treslin, the human homologue of Sld3, which then subsequently associates with TOPBP1, the human homologue of Dpb11. However, unlike in yeast, the human homologue of Sld2 (RecQ4) functions independently of CDK. Reprinted from Current Biology, Volume 21, Mueller, A. C., Keaton, M. A. and Dutta, A., DNA Replication: Mammalian Treslin–TopBP1 Interaction Mirrors Yeast Sld3–Dpb11, Pages No. R638–R640, Copyright 2011, with permission from Elsevier.

### 1.5.3 Formation of the replisome

Formation of the replisome is essential for DNA to be replicated and has been studied extensively. After formation of the pre-IC, Mcm10 binds Mcm2 between its Oligonucleotide/Oligosaccharide-Binding (OB)-fold and A subdomain, where it promotes elongation of replication and stabilises the association of the Cdc45-MCM-GINS (CMG) complex (Lööke et al., 2017). Furthermore, it has been suggested that upon interaction with CMG, Mcm10 sits at the junction of the replication fork and has a role in the control of fork regression (Mayle et al., 2019).

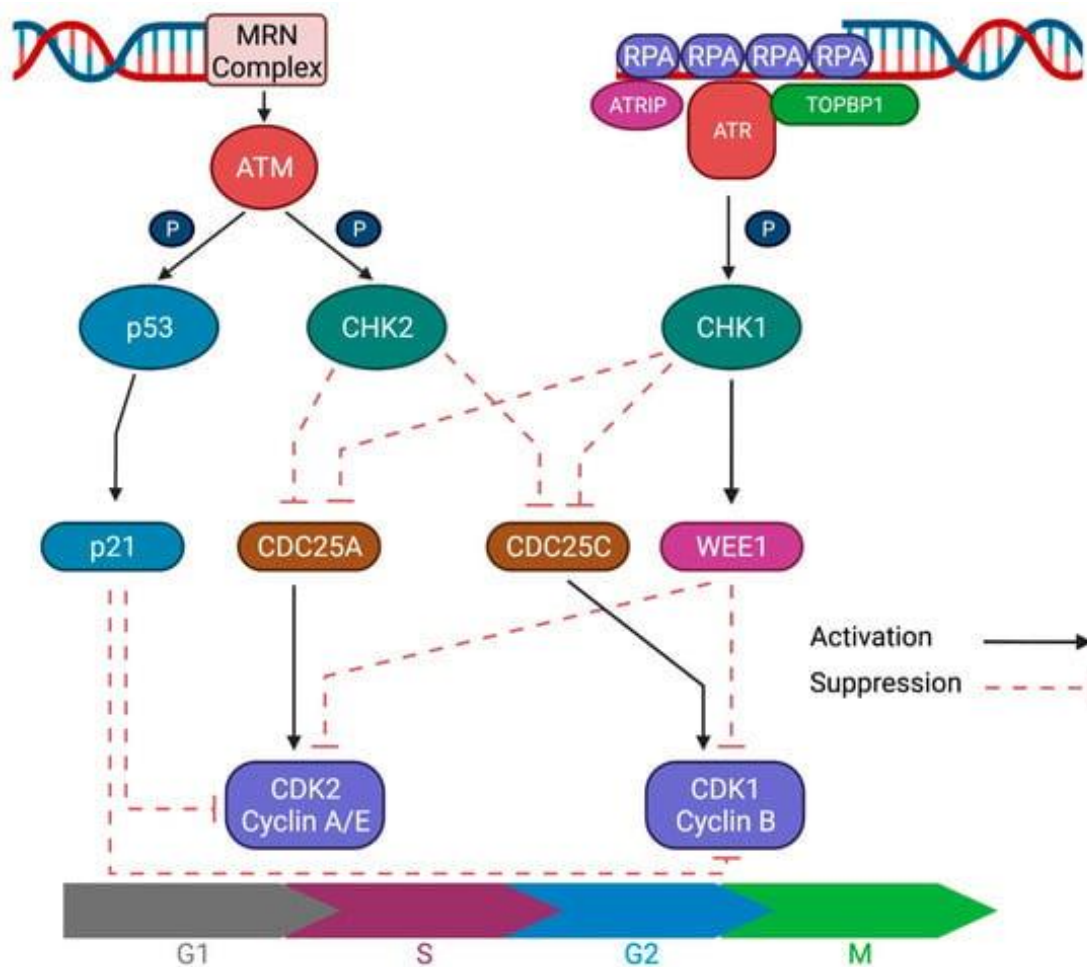
Mcm10 also plays an important part in the recruitment of RPA and the DNA polymerases, including the stabilisation and recruitment of Pol  $\alpha$  (Kanke et al., 2012; Perez-Arnaiz et al., 2017; Ricke and Bielinsky, 2004). The Pol  $\alpha$ /primase complex binds to an RNA primer and extends it with deoxyribonucleotides to produce an RNA/DNA hybrid primer (Perera et al., 2013). Subsequently Pol  $\epsilon$  participates in leading strand replication, while Pol  $\delta$  is thought to be the primary polymerase involved in lagging strand replication (Daigaku et al., 2015; Nick McElhinny et al., 2008; Pursell et al., 2007). Recent evidence suggests that Pol  $\delta$  is also involved in the initiation and termination of leading-strand synthesis (Zhou et al., 2019a). Pol  $\delta$  and Pol  $\epsilon$  are both stimulated by PCNA, with Pol  $\delta$  having a higher affinity for PCNA (Chilkova et al., 2007).

A limited component-containing leading-strand replisome composed of the CMG helicase, Pol  $\epsilon$ , RFC, PCNA and RPA is capable of replicating DNA at an average rate of  $4.4 \pm 0.6$  nucleotides (nt)/s, equivalent to 0.264 kilobases (Kb)/min (Georgescu et al., 2014). While this is much lower than the *in vivo* replication rate, it demonstrates that replication is possible with a very limited number of components. Additionally, although RPA and PCNA stimulate the reaction, they are not essential and the replication rate was halved in the absence of PCNA and RFC (Georgescu et al., 2014). A maximal replication rate of  $1.92 \text{ kb min}^{-1}$ , and bulk rate of  $1.44 \text{ kb min}^{-1}$  is made possible with the addition of further replisome components (Yeeles et al., 2017). These rates are comparable to those measured *in vivo* (Conti et al., 2007; Sekedat et al., 2010); however, *in vivo* rates do have the additional challenge of chromatin disassembly and reassembly. Studies into the replisome components required for

maximal replication rate have identified Csm3-Tof1 and Mrc1 as key replisome components for maximal replication rate in *S. cerevisiae* (Yeeles et al., 2017).

### **1.6 The DNA damage response: stalled replication forks and double strand breaks (DSBs)**

High fidelity DNA replication is essential for cell survival. To ensure this faithful duplication, many DNA damage response pathways help to repair any damage that may occur. Intrinsic to the DNA damage response are ataxia-telangiectasia mutated (ATM) and ATM- and Rad3-Related (ATR), two members of the phosphatidylinositol-3 kinase-related kinase (PIKK) family which respond to different types of damage (Blackford and Jackson, 2017; Maréchal and Zou, 2013). The ATR kinase is essential for cell survival (de Klein et al., 2000). It responds to a variety of different types of damage but is particularly important in cell proliferation and DNA replication (Saldivar et al., 2017; Sirbu and Cortez, 2013). On the other hand, ATM's key role is in the DNA damage response to double strand breaks (DSBs) and oxidative stress (Maréchal and Zou, 2013; Paull, 2015). Both kinases phosphorylate downstream proteins, resulting in signalling cascades which induce cell cycle arrest and DNA repair (Choi and Lee, 2022). A summary of these processes is shown in Figure 1.11. In this section, the roles of ATM and ATR in the response to DNA damage will be discussed.



**Figure 1.11: An overview of the roles of ATM and ATR in the DNA damage response.**

ATM is recruited to double strand breaks and activated through its interaction with the MRN complex. Upon activation ATM phosphorylates several substrates, including p53 and CHK2. ATR is targeted to single stranded DNA at replication forks by ATRIP. Upon activation, ATR activates CHK1, which induces a signalling cascade that overlaps with that in the ATM response. Both cascades result in the inhibition of Cyclin-CDKs that are required for progression through the cell cycle. ATM, ataxia-telangiectasia mutated; ATR, ATM- and Rad3-Related (ATR); ATRIP, ATR-interacting protein; CHK1, checkpoint kinase 1; CHK2, checkpoint kinase 2; DNA, deoxyribonucleic acid.

Figure reproduced without changes from the International Journal of Molecular Sciences, Volume 23, Choi, W. and Lee, S., Therapeutic targeting of DNA damage response in cancer, Copyright 2022, <https://doi.org/10.3390/ijms23031701>; available under the CC BY 4.0 licence (<https://creativecommons.org/licenses/by/4.0/legalcode>) and disclaimer of warranties within.

### 1.6.1 The ATR-mediated response to stalled replication forks

ATR is a key kinase which, in addition to its response to DSBs, acts in response to a variety of different types of damage that result in the stalling of replication forks

(Saldivar et al., 2017). More specifically, it is activated in response to damaging agents such as MMC, a chemotherapeutic alkylating agent which introduces crosslinks into DNA (Mladenov et al., 2007; Roh et al., 2008). Interestingly, there is an increase in Ap<sub>4</sub>N levels in response to MMC treatment, and Ap<sub>4</sub>Ns inhibit the initiation of DNA replication (Marriott et al., 2015). It is therefore possible that the mechanism by which Ap<sub>4</sub>A inhibits DNA replication may also have a role in supporting the ATR damage response.

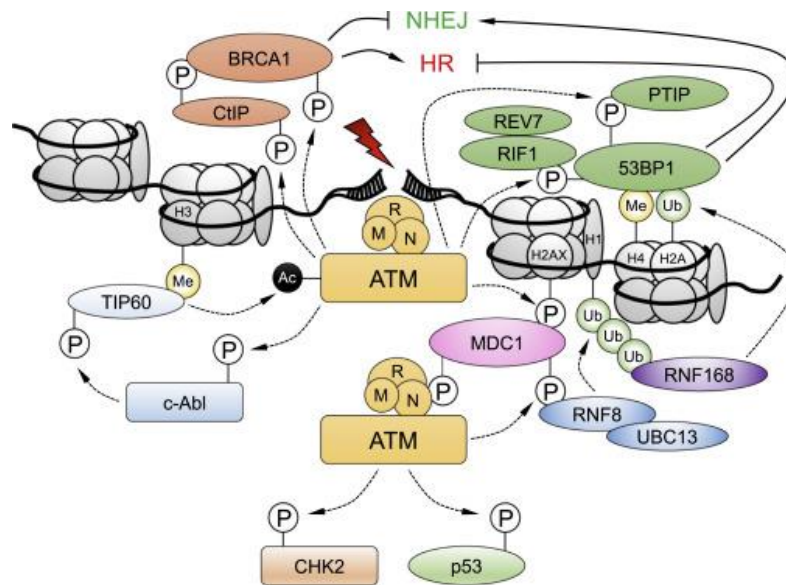
The ATR signalling response begins with RPA, which coats single stranded DNA (ssDNA) and recruits the ATR-interacting protein (ATRIP) (Zou and Elledge, 2003). This ensures that ATR localises to RPA-coated DNA through its interaction with the carboxy-terminal of ATRIP, while the amino-terminal of ATRIP associates with RPA (Cortez et al., 2001; Itakura et al., 2004). It has been suggested that RPA also induces Rad17-Rfc2-5 complex binding and promotes Rad17-Rfc2-5-induced recruitment of the Rad9-Rad1-Hus1 (9-1-1) complex (Zou et al., 2003). However, another study has suggested that TopBP1 and Pol  $\alpha$  are required for 9-1-1 complex loading in *Xenopus laevis* (Yan and Michael, 2009). TopBP1 is an activator of the ATR-ATRIP complex, and it must interact with ATRIP to bind and induce the kinase activity of ATR (Kumagai et al., 2006; Mordes et al., 2008). Interaction between BRCT 1 and 2, and phospho-serine (phospho-ser) 387 form an interaction between TopBP1 and Rad9, resulting in TopBP1 being localised to the stalled replication fork (Delacroix et al., 2007).

A key substrate of ATR is checkpoint kinase 1 (Chk1), which is phosphorylated after localisation of TopBP1 to the replication fork via its interaction with Rad9 (Delacroix et al., 2007). Chk1 phosphorylation is removed by protein phosphatase 2A (PP2A) and this mechanism works to maintain low activity of Chk1 in unstressed cells (Leung-Pineda et al., 2006). Chk1 phosphorylates Cdc25A on four residues - 123, 178, 278, and 292 – resulting in the degradation of Cdc25A, a process which is further accelerated in ionising radiation-induced damage by the kinase activity of checkpoint kinase 2 (Chk2) and ATM (Sørensen et al., 2003). This provides an example of the overlapping activity of the ATR and ATM kinases in response to DNA damage. Cdc25A is a protein phosphatase which regulates apoptosis and the increases the activity of CDKs involved

in cell cycle progression (Shen and Huang, 2012). Therefore, proteolysis of Cdc25A promotes the stalling of replication forks to allow time for DNA damage to be repaired.

### **1.6.2 The ATM-mediated response to double strand breaks**

In the ATM-mediated response, ATM is first localised to DNA DSBs by the Mre11-Rad50-Nbs1 (MRN) complex, which then induces ATM's kinase activity (Lee and Paull, 2005). Previously, ATM autophosphorylation of Ser-1981 in human cells was shown to dissociate the inactive ATM dimer into active monomers (Bakkenist and Kastan, 2003). However, this requirement for ATM autophosphorylation was not seen in yeast cells, where an S1918A mutant was able to stimulate phosphorylation of p53 and Chk2 to a similar extent as the wild-type ATM (Lee and Paull, 2005). Chk2 is a checkpoint kinase which is phosphorylated by ATM directly in response to ionizing radiation, but not in response to ultraviolet (UV) radiation or hydroxyurea (HU) (Matsuoka et al., 2000). Upon activation in response to DNA damage, Chk2 phosphorylates murine double minute X (MDMX) at both Ser-342 and Ser-367 residues, with a strong preference for the latter (Chen et al., 2005). Phosphorylation of Ser-367 on MDMX leads to its degradation by murine double minute 2 (MDM2), which in turn results in the accumulation of p53 (Chen et al., 2005). Figure 1.12 presents some of the key processes that ATM is involved in (Blackford and Jackson, 2017).



**Figure 1.12: The ATM response to DNA double strand breaks.** Upon activation by the MRN complex in response to DSBs, ATM induces a signalling cascade. This results in phosphorylation of several downstream factors, including CHK2, p53, H2AX, MDC1, 53BP1, BRCA and CtIP. CHK2, checkpoint kinase 2; CtIP, carboxy-terminal binding protein interacting protein; DSB, double strand break; H2AX, H2A histone family member X; MDC1, Mediator of DNA damage checkpoint protein 1. Reprinted from Molecular Cell, Volume 66, Blackford, A. N. & Jackson, S. P. ATM, ATR, and DNA-PK: The Trinity at the Heart of the DNA Damage Response., Pages No. 801–817, Copyright 2017, with permission from Elsevier.

ATM is also the major PIKK responsible for H2AX phosphorylation at Ser-139 in the early stages of the damage response, forming  $\gamma$ H2AX (Burma et al., 2001). Formation of  $\gamma$ H2AX enables the binding of MDC1 via the phosphoepitope on  $\gamma$ H2AX (Stucki et al., 2005). From this point, MDC1 has a number of downstream binding partners (Coster and Goldberg, 2010). Although ATM can interact with damage sites in the absence of MDC1, for it to remain at sites of DNA damage it must autophosphorylate at Ser-1981 and associate with MDC1 (So et al., 2009). This binding between H2AX, ATM and MDC1 initiates a positive feedback loop, enabling further phosphorylation of H2AX by ATM (Lou et al., 2006).

In addition to enabling a positive feedback loop for damage response signalling, MDC1 also initiates other pathways by targeting a number of downstream binding partners (Coster and Goldberg, 2010). Interaction between ATM-phosphorylated MDC1 and the

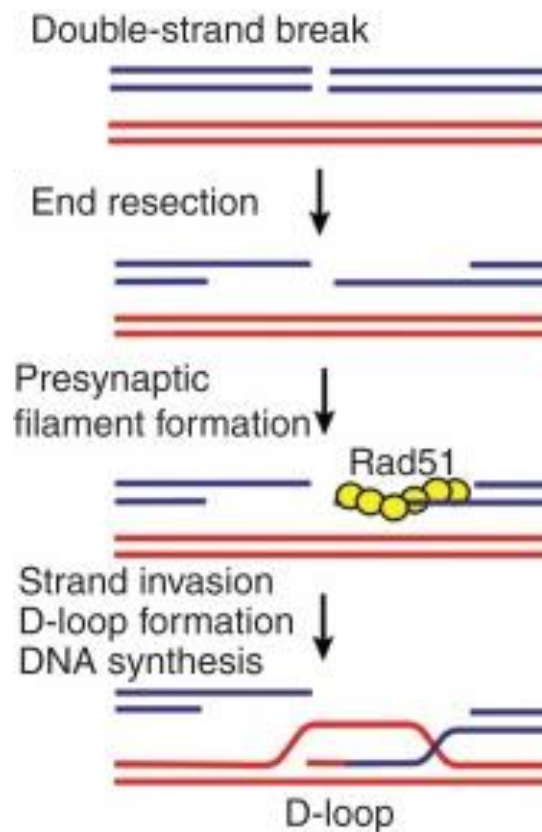
forkhead-associated (FHA) domain of RNF8 is responsible for localisation of RNF8 to sites of DNA damage (Huen et al., 2007; Kolas et al., 2007; Mailand et al., 2007). In a UBC13-interaction dependent manner, RNF8 subsequently enables the localisation and retention of 53BP1 and BRCA1 to sites of DNA damage (Huen et al., 2007; Kolas et al., 2007; Mailand et al., 2007; Plans et al., 2006). RNF8 also promotes ubiquitylation at the damage site, including ubiquitylation of histone H2A and H2AX (Mailand et al., 2007). Then RNF168, a second Ub ligase, associates with ubiquitylated H2A and aids the retention of 53BP1 and BRCA1 at DSBs by increasing lys63-linked Ub conjugate concentration (Doil et al., 2009; Stewart et al., 2009). Histone interaction is also seen between 53BP1, and histone H4 dimethylated lysine-20 (H4K20me2) and histone H2A ubiquitinated on lys-15 (H2A K15ub) via its Tudor-UDR domain and ubiquitination-dependent recruitment (UDR) motif (Fradet-Turcotte et al., 2013). Alternatively, in the absence of ATM the Tudor-domain of 53BP1 interacts with Tudor Interacting Repair Regulator (TIRR), a nudix hydrolase which reduces the localisation of 53BP1 to DSBs by preventing the interaction between 53BP1 and H4K20me2 (Drané et al., 2017). In the presence of ATM, the interaction between 53BP1 and TIRR is reduced after treatment with ionising radiation, suggesting that ATM is responsible for the dissociation of 53BP1 from TIRR after DNA damage (Zhang et al., 2017).

In G1-phase, ATM-phosphorylated 53BP1 recruits RIF and this interaction is sustained, which inhibits DNA resection and prevents homologous recombination (HR), pushing the repair pathway towards non-homologous end joining (NHEJ) (Isono et al., 2017). Although this phosphorylation of 53BP1 and recruitment of RIF also occurs in S- and G2-phase, it occurs only transiently and is followed by DNA resection and HR (Isono et al., 2017). This occurs because in G1, 53BP1 inhibits accumulation of BRCA1 at DSBs (Escribano-Díaz et al., 2013). However, in the presence of BRCA1, 53BP1 is dephosphorylated by the PP4C phosphatase, causing RIF to be released and promoting HR by allowing DNA resection to occur (Isono et al., 2017). Association of 53BP1 with nucleosomes and recruitment of RIF1 and Pax transactivation domain-interacting protein (PTIP) to DSBs is prevented by its acetylation, also resulting in the promotion of HR (Guo et al., 2018).



### **1.6.3 Homologous recombination is initiated by CtIP, a phosphorylation target of ATM**

HR is one of the two major pathways involved in the repair of DSBs. HR primarily contributes to DNA repair during DNA replication; whereas NHEJ, the other major pathway, contributes to DNA damage repair throughout the cell cycle (Karanam et al., 2012; Rothkamm et al., 2003). In this section the focus will be HR, as this primarily takes place in response to damage during replication. Briefly, HR proceeds in response to DSBs and begins with the resection of DNA around the DSB (Fig. 1.13) (Daley et al., 2014). Because of this, both ATM (which forms part of the initial response to DSBs) and ATR (which responds to RPA-coated ssDNA) can be activated in this process (Sun et al., 2020). During HR, Rad51 interacts with the ssDNA to form the presynaptic filament, which invades a homologous sequence and facilitates DNA synthesis and repair (Fig. 1.13) (Daley et al., 2014).

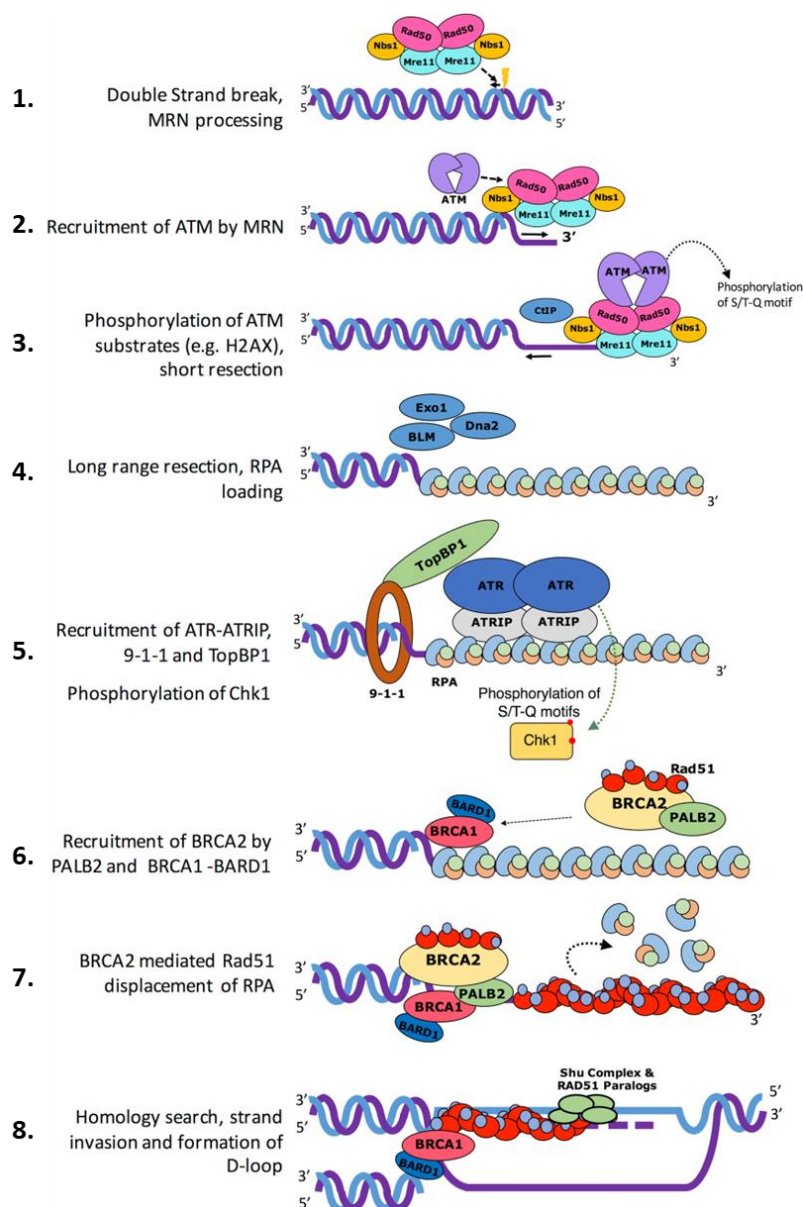


**Figure 1.13: Strand invasion during homologous recombination.** HR is one of the pathways responsible for the repair of DSBs. During the presynaptic phase of HR, the DNA around the DSB is resected. Rad51 replaces the RPA-coated ssDNA, forming the presynaptic filament. This filament invades duplex DNA at sites of homology, causing formation of the D-loop. DSB, double strand break; HR, homologous recombination; ssDNA, single stranded DNA.

Reprinted from Cold Spring Harbor Perspectives in Biology without changes, Volume 6, Daley, J. M., Gaines, W. A., Kwon, Y. & Sung, P., Regulation of DNA pairing in homologous recombination., a017954, Copyright 2014, with permission from Cold Spring Harbor Laboratory Press

CtIP is a human protein sharing sequence homology with yeast Sae2 which is recruited to DSBs in S- and G2-phase, where it binds directly to the MRN complex and is important for DNA resection at DSBs (Sartori et al., 2007). In *S. cerevisiae*, CDK-induced phosphorylation of Sae2 at its Ser-267 residue is important in the control of DSB end resection (Huertas et al., 2008). Like in yeast, phosphorylation of the Thr-847 residue in CtIP is also important for DNA resection (Huertas and Jackson, 2009). CDK-induced phosphorylation causes CtIP to associate with Nbs1 of the MRN complex, which enables phosphorylation of CtIP by ATM (Wang et al., 2003) (Fig. 1.14). CtIP also

associates with BRCA1, which enables interaction between BRCA1 and the MRN complex (Chen et al., 2008). Furthermore, in order for HR to occur, phosphorylation of the serine 327 residue on CtIP is required, further supporting evidence for the importance of phosphorylation in HR (Yun and Hiom, 2009).



**Figure 1.14: Strand invasion during homologous recombination at the molecular level.** 1. The MRN complex detects a double strand break and 2. recruits ATM, resulting in the initiation of 3. the ATM signalling cascade and the DNA damage response. CtIP is phosphorylated by ATM and promotes Rad50 dependent DNA end resection by MRE11. 4. BLM stimulates EXO1 and DNA2-mediated resection of DNA, and the resulting ssDNA is coated by RPA. 5. ATR is recruited to the RPA-coated ssDNA at the site of DNA damage by ATRIP and activated to induce its DNA damage response signalling cascade through the phosphorylation of CHK1. 6. BRCA1 targets Rad51 to the ssDNA and aids the 7. substitution of RPA for Rad51. 8. Finally, the Rad51 presynaptic filament invades duplex DNA at areas of homology. Figure adapted from Sun et al., 2020.

Figure reproduced with minor changes from Cellular and Molecular Life Sciences, Volume 77, Sun, Y., McCorvie, T. J., Yates, L. A. & Zhang, X., Structural basis of homologous recombination, Copyright 2020, doi: 10.1007/s00018-019-03365-1, available under the CC BY 4.0 licence (<http://creativecommons.org/licenses/by/4.0/>) and disclaimer of warranties within.

Thr-847-phosphorylated CtIP promotes the endonuclease activity of Mre11, and 5'-terminated double stranded DNA near DSBs is cleaved by MRN-CtIP in a RAD50 ATP-hydrolysis-dependent manner (Anand et al., 2016). Exo1, WRN, DNA2 and BLM are also involved in pathways resulting in the resection of DNA at DSBs (Tomimatsu et al., 2012). The BLM helicase stimulates DNA resection both by EXO1 and by DNA2, leading to binding of RPA to the resulting ssDNA which promotes resection (Nimonkar et al., 2008; Soniat et al., 2019). Interestingly, RPA is phosphorylated during resection, causing it to inhibit the BLM helicase and therefore inhibit DNA resection (Soniat et al., 2019). DNA resection by Exo1 and BLM enable Rad51 to promote homologous DNA pairing (Nimonkar et al., 2008). The presence of free RPA inhibits the association of Rad51 with DNA (Ma et al., 2016). However, Rad51 is targeted to the RPA-coated ssDNA by BRCA2, which works with DSS1 to enable the replacement of RPA on DNA with Rad51 and stimulates exchange of the DNA strands (Fig. 1.14) (Jensen et al., 2010; Liu et al., 2010; Thorslund et al., 2010; Zhao et al., 2015). The phase of HR up to the strand exchange is known as presynapsis, and is followed by synapsis (Li and Heyer, 2008). During synapsis, the presynaptic filament forms a D-loop as it locates sequence homology and invades the duplex DNA (Li and Heyer, 2008). Finally, postsynapsis proceeds via one of three possible pathways – the synthesis-dependent strand annealing (SDSA) pathway, the double-strand break repair (DSBR) pathway or the double Holliday junction dissolution pathway (Daley et al., 2014).

## **1.7 Conclusion**

In summary, high-fidelity DNA replication is required for the successful replication of DNA, however despite its high fidelity, mutations can be introduced during the copying process. The progression of cells through the cell cycle and through the DNA replication process is therefore highly regulated. The concentration of the alarmone Ap<sub>4</sub>A increases under various types of cell stress, including oxidative stress and stress induced by chemotherapeutic alkylating agents (Marriott et al., 2015). Ap<sub>4</sub>A has been linked to several signalling processes including DNA replication (Marriott et al., 2015). As Ap<sub>4</sub>A inhibits the initiation of DNA replication and increases in response to cell stress, Ap<sub>4</sub>A may have a role in signalling in the damage-associated replication pathways. However, how Ap<sub>4</sub>A modulates this response has not yet been determined.

In some signalling mechanisms a homobifunctional Ap<sub>4</sub>A structure is required for signalling (Yu et al., 2019); however, this is not always the case (Luciano and Belasco, 2020; Luciano et al., 2019). Therefore, the molecular determinants that enable the inhibition of DNA replication remain to be determined. This project aimed to:

- a) Optimise a method for the synthetic production of Ap<sub>4</sub>A and Ap<sub>4</sub>Ns and investigate whether Ap<sub>4</sub>Ns could be produced by a biologically relevant mechanism in human cells.
- b) Develop a method for the extraction and identification of intracellular Ap<sub>4</sub>Ns from a human cell line.
- c) Investigate the stress response in human cells with intrinsically high Ap<sub>4</sub>N levels.
- d) Determine whether, like Ap<sub>4</sub>A, Ap<sub>4</sub>Ns can also inhibit the initiation stage of DNA replication, and how these molecules influence the localisation of DNA replication proteins to chromatin.

## **Chapter 2: Materials and Methods**

## 2.1 Synthesis and validation of Ap<sub>4</sub>Ns and Ap<sub>4</sub>dNs

### 2.1.1 Chemical synthesis of Ap<sub>4</sub>Ns

Magnesium chloride (MgCl<sub>2</sub>; 1M, 3 µL), (d)NTP (100 mM, 10 µL) and (d)AMP (1M, 3 µL) were mixed and lyophilised overnight in a Christ freeze-dryer (-80°C, 0.0010 mbar; Alpha 2–4 LD plus). The appropriate AMP and NTP molecules were added depending on the target product, as shown in Table 2.1.

**Table 2.1: Reactants for the synthesis of Ap<sub>4</sub>N and Ap<sub>4</sub>dN**

Product	Reactants		Supplier
<b>Ap<sub>4</sub>A</b>	AMP	ATP	Sigma Aldrich
<b>Ap<sub>4</sub>C</b>	AMP	CTP	Sigma Aldrich
<b>Ap<sub>4</sub>G</b>	AMP	GTP	Sigma Aldrich
<b>Ap<sub>4</sub>U</b>	AMP	UTP	Sigma Aldrich
<b>Ap<sub>4</sub>dA</b>	AMP	dATP	Sigma Aldrich
<b>Ap<sub>4</sub>dC</b>	AMP	dCTP	Sigma Aldrich
<b>Ap<sub>4</sub>dG</b>	AMP	dGTP	Sigma Aldrich
<b>Ap<sub>4</sub>dT</b>	AMP	dTTP	Sigma Aldrich
<b>dAp<sub>4</sub>dA</b>	dAMP	dATP	Sigma Aldrich
<b>AppNppA</b>	AMP	AMPPNP	Roche Diagnostics, REF: 10102547001

The lyophilised product was resuspended in 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES buffer; 3.4 M, pH 6.8, 10 µL) and 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC; 5 M, 5 µL), then incubated at 25–30 °C overnight. After overnight incubation, the reaction was quenched in ammonium bicarbonate (21 mM, pH 8.6, 100 µL). The Ap<sub>4</sub>N product was analysed by high performance liquid chromatography (HPLC; 1260 Infinity II, Agilent Technologies) using a ResourceQ column (1 mL; GE17-1177-01) and an ammonium bicarbonate:water gradient increasing from 50 mM to either 1 M or 1.5 M. For initial determination of retention time, the synthesised Ap<sub>4</sub>A was compared with a commercially sourced Ap<sub>4</sub>A standard (Sigma Aldrich).



### **2.1.2 Biological synthesis of Ap<sub>4</sub>Ns**

Ub (60 µM; Boston Biochem, Cat# U-100H, Ubiquitin, human), human ubiquitin activating enzyme (UBE1; 2 mM; Boston Biochem, Cat# E-206, GST-E1, human), ATP (1 mM, Sigma Aldrich) and the appropriate NTP (for Ap<sub>4</sub>C = CTP, Ap<sub>4</sub>G = GTP and Ap<sub>4</sub>U = uridine triphosphate [UTP]; all 1 mM) were combined and incubated in buffer (Tris, 25 mM, pH 7.6; NaCl, 50 mM; MgCl<sub>2</sub>, 5 mM; dithiothreitol [DTT], 1.25 mM; total reaction volume, 20 µL) at 37 °C for 0, 24, 50 or 72 hours. Aliquots (4 µL) were taken from the reaction mixture at these timepoints and frozen at -80 °C until ready for analysis. All reactions were performed in DNA LoBind tubes to reduce risk of polyethylene glycol (PEG) contamination. After completion of all timepoints, the samples were lyophilised, diluted in water, and analysed by HPLC (ResourceQ column; ammonium bicarbonate gradient, 50 mM → 1 M).

### **2.1.3 Validation of synthesised Ap<sub>4</sub>N identity by High Resolution Mass Spectrometry (HRMS)**

Synthesised Ap<sub>4</sub>N samples were diluted to between 6 and 16 µM, then sent for analysis by HRMS. Analysis was performed on the liquid chromatography coupled to ion trap time-of-flight mass spectrometer (LCMS-IT-TOF) (Shimadzu) fronted by a Shimadzu NexeraX2 Ultra-high performance liquid chromatography (UHPLC) system, in electrospray ionisation (ESI) mode.

## **2.2 Determining synthesised Ap<sub>4</sub>N concentration**

### **2.2.1 Hydrolysis of Ap<sub>4</sub>A and Ap<sub>4</sub>N synthesis products**

Synthesised Ap<sub>4</sub>Ns were diluted in buffer (Tris HCl, 20 mM; MgCl<sub>2</sub>, 2 mM; pH 7.5 using sodium hydroxide (NaOH); filtered) to approximately 100 µM. For each Ap<sub>4</sub>N, 9.5 µL of the diluted nucleotide was added to two separate tubes. The control tube was supplemented with a further 0.5 µL of buffer, while 0.5 µL recombinant human NUDT2 Ap<sub>4</sub>A hydrolase (final concentration, 5 µg/mL; as used by Marriot et al., 2015) was added to the other. All samples were incubated at room temperature until hydrolysis had reached completion. To confirm complete hydrolysis had occurred, the hydrolysed samples were analysed by HPLC (ResourceQ, ammonium bicarbonate gradient [50 mM

→ 1 M]). The concentration of both the hydrolysed and non-hydrolysed samples were then determined using a NanoDrop 2000c spectrophotometer at a wavelength of 260 nm and 280 nm (Thermo Scientific).

### **2.2.2 Determining the extinction coefficient for each Ap<sub>4</sub>N**

An approximate extinction coefficient ( $\epsilon$ ) was first calculated by summing the previously recorded extinction coefficients for the two nucleotides in each molecule (Cavaluzzi and Borer, 2004). The percentage change in NanoDrop A260 reading after hydrolysis was then used to modify the extinction coefficient to account for the effect of base stacking on absorbance readings. These modified extinction coefficients were then used for the calculation of Ap<sub>4</sub>N concentration (mmol) using a rearrangement of the equation  $A = \epsilon lc$  ( $A$ , absorbance;  $\epsilon$ , extinction coefficient;  $l$ , path length [in this case,  $l=1$ ];  $c$ , concentration).

## **2.3 Cell culture**

### **2.3.1 Adherent cells: 3T3 and HeLa cells**

Murine 3T3 fibroblasts and HeLa cells were each maintained in Dulbecco's Modified Eagle's Medium plus GlutaMax™-I (D-MEM; 21885-025, Gibco) supplemented with 10% v/v Fetal Bovine Serum (FBS; LabTech or Biosera), and penicillin/streptomycin/glutamine (P/S/G; 10378-016, Gibco) at 37 °C, 5% CO<sub>2</sub> in a HeraCell 150i CO<sub>2</sub> incubator (Thermo Scientific). For passage, the cells were washed in pre-warmed Dulbecco's phosphate buffered saline (DPBS; Gibco, 14190-144), then incubated with 10X trypsin-ethylenediaminetetraacetic acid (EDTA) (1 mL; 15400-054, Gibco) diluted in DPBS (1:10), at 37 °C until fully detached. The trypsin was inactivated by addition of the pre-warmed supplemented medium, and the cells were split according to their density. Pre-warmed medium was added to make up to a total of 30 mL per 15 cm cell culture dish.

### **2.3.2 Suspension cells – KBM7 cells**

Human chronic myeloid leukaemia cells in blast crisis (KBM7 cells) and their *Nudt2* knockout counterparts (NuKO cells) (Marriott et al., 2016) were cultured at 37 °C, 5%

CO<sub>2</sub>, in Iscove's Modified Dulbecco's Medium (IMDM; 21980-032, Gibco) supplemented with 10% v/v FBS (LabTech or Biosera) and P/S/G (Gibco). For passage at high cell densities the appropriate amount of cell-containing medium was transferred directly to a new cell culture plate. At lower densities, the cells were centrifuged at room temperature in a Harrier 18/80 refrigerated centrifuge (500 x *g*, 5 minutes), and the medium was removed. The cell pellet was resuspended in fresh medium and split between the appropriate number of plates. In both cases, fresh medium was added to the new plates to reach the appropriate volume, equivalent to a total of 30 mL for a 15 cm cell culture dish.

#### **2.4 Extraction of metabolites from KBM7 and NuKO cells**

Prior to extraction, the KBM7 or NuKO cells were counted. The required number of cells were centrifuged (500 x *g*, 5 minutes; Fisher Scientific, accuSpin Micro 17), and the medium was removed. To reduce the likelihood of plastics contaminating the sample, the cell pellet was washed in 1X DPBS (1 mL, Gibco), and transferred to a DNA-LoBind eppendorf tube® (Ref# 0030 108 418). The sample was centrifuged (500 x *g*, 5 minutes), the DPBS was removed, and the cell pellet was resuspended in 50:50 ice-cold chloroform (319988; Sigma Aldrich) and water (1 mL total) on ice. While on ice, the sample was mixed to ensure maximum extraction of cell metabolites. The sample was then centrifuged (10,000 x *g*, 10 minutes) to remove debris, and the aqueous fraction was transferred to a fresh DNA Lo-bind tube. This process was repeated twice more but centrifuging at 20,000 x *g* for 10 minutes each time. The sample was frozen at -80 °C, then lyophilised overnight.

#### **2.5 HPLC analysis of mammalian cell extract**

Prior to analysis HPLC eluents A and B were prepared. Eluent A was composed of ammonium bicarbonate (1, 1.5 or 2 M (method adapted to maintain gradient regardless of upper concentration used); pH 9.6 using ammonium hydroxide; 09830, Sigma Aldrich), while Eluent B was MilliQ water. Eluent A was filtered to 0.22 µM. Both eluents were sonicated, then used to equilibrate the column prior to analysis. Metabolites were extracted from the mammalian cell lines as described in section 2.4,

then treated with recombinant NUDT2 Ap<sub>4</sub>A hydrolase (Marriott et al., 2015) and/or FastAP Thermosensitive Ap<sub>4</sub>A phosphatase, or frozen directly at -80 °C (Thermo Scientific, #EF0651). The mammalian cell extract was lyophilised, then resuspended in MilliQ (110 µL) and 100 µL of sample was injected into the HPLC, to maximise the amount of product detectable. The sample was run through a ResourceQ column on a steady gradient of increasing ammonium bicarbonate (50 mM to 1M):water (MilliQ), with a total run time of 22.5 minutes. LC OpenLAB software was used for visualising HPLC traces.

## **2.6 Clean up KBM7 mammalian cell extract**

### **2.6.1 Graphite column clean-up**

The lyophilised extract sample was solubilised in ammonium bicarbonate (10 mM) prior to clean-up. First, the storage buffer was removed from the spin column by centrifugation (2000 x *g*, 1 minute; accuSpin Micro 17, Fischer Scientific). The graphite spin columns (Pierce) were washed in acetonitrile (80% v/v; Sigma Aldrich, 34998) + trifluoroacetic acid (TFA; 0.1% v/v), then centrifuged (2000 x *g*, 1 minute). The column was then washed in water, then centrifuged (2000 x *g*, 1 minute) before the sample was loaded by adding it to the column and incubating for 10 minutes, with periodic vortexing to maximise binding. Unbound extract was then removed from the column by centrifugation (1,000 x *g*, 3 minutes). The column-bound sample was washed with water, centrifuged (2,000 x *g*, 1 minute), washed with acetonitrile (25% v/v), centrifuged (2,000 x *g*, 1 minute), then washed with triethyl ammonium acetate (TEAA) (50 mM; pH 6, using high grade acetic acid; prod # 1862986, Thermo Scientific) and centrifuged once more (2000 x *g*, 1 minute). Finally, the Ap<sub>4</sub>Ns were eluted by adding a mixture of acetonitrile (50% v/v), and TEAA (50 mM, pH 6) and centrifuging (2000 x *g*, 1 minute). The eluted sample was lyophilised overnight.

### **2.6.2 Strong cation exchange clean-up**

The column was prepared by incubating in ammonium bicarbonate (10 mM, 10 minutes). Next, the column was centrifuged (1000 x *g*, 3 minutes) and buffer flow through discarded. The incubation and centrifugation steps were then repeated once

more. The sample was then prepared in ammonium bicarbonate (10 mM) and added to the column. The column was centrifuged (1000 x *g*, 3 minutes), and the flow-through was collected in a clean DNA LoBind tube. The flow through was then lyophilised overnight.

## 2.7 Triple quadrupole mass spectrometry analysis of mammalian cell extract

### 2.7.1 Determining multiple reaction monitoring (MRM) transitions for Ap<sub>4</sub>N standards

Triple quadrupole (QQQ) mass spectrometry analysis was performed on the liquid chromatography mass spectrometer (LC-MS)-8040 (Shimadzu) in ESI mode. Each Ap<sub>4</sub>N sample was diluted in water to concentrations of ~20–70 µM in 100 µL total volume. The eluents were set up as described in Table 2.2.

**Table 2.2: Eluents for triple quadrupole mass spectrometry analysis** (Schulz et al., 2014). HPLC = high performance liquid chromatography; N,N-DMHA = N,N-Dimethylhexylamine.

Eluent	Composition	pH
<b>Eluent A</b>	N,N-DMHA (0.1% v/v; Sigma Aldrich, 308102-5G) Ammonium formate (1 mM; Sigma Aldrich)	pH 9.0, adjusted with formic acid (Sigma Aldrich)
<b>Eluent B</b>	100% v/v HPLC-grade Acetonitrile (Sigma Aldrich)	-

Each standard was then injected and run according to the method described in Table 2.3.

**Table 2.3: Method for LCMS-8040 analysis**

Time (minutes)	Eluent A (%)	Eluent B (%)
<b>0</b>	100	0
<b>2</b>	84	16
<b>4</b>	79	21
<b>14</b>	2	98
<b>18</b>	2	98
<b>20</b>	100	0
<b>22</b>	100	0
<b>25</b>	End of method	

The specific masses for each Ap<sub>4</sub>N were input into the LC-MS Wizard, which was then used to optimise detection of the most abundant MRM transitions for each Ap<sub>4</sub>N. The top three transitions for each Ap<sub>4</sub>N were selected and used to create the method for analysis of cell extracts. To validate the method for identification of each Ap<sub>4</sub>N, each Ap<sub>4</sub>N standard was run with the machine set to detect all selected transitions to confirm that the appropriate signals were picked up for each sample.

### **2.7.2 Analysis of mammalian cell extract by LC-MS**

Lyophilised extract samples were diluted in water (50 µL) and were run through an AdvanceBio peptide mapping column (2.1 x 150 mm; 653750-902; Agilent) at a rate of 0.3 mL/minute and a temperature of 35 °C. The system was set up to detect the three most abundant MRM transitions, as determined using the method in section 2.7.1. Column run time and eluent ratios were set up as for the detection of MRM transitions, shown in Table 2.3 above.

### **2.7.3 LC-MS measurement of the effect of DNA damaging agents on DNA replication**

NuKO cells were treated with either MMC (100 nM; Fisher BioReagents, BP2531-10), doxorubicin (DOX; 100 nM; Fisher BioReagents 10512955), HU (0.1 mM and 1 mM; Sigma Aldrich H8627), gemcitabine (GEM; 100 nM, HCl Sigma Aldrich G6423) or left untreated for 18 hours. The cells were counted, and the metabolites were independently extracted from an equal number of treated and untreated cells using the protocol described in section 2.4. The lyophilised sample was resuspended in water, then injected and analysed on the LC-MS-8040 triple quadrupole mass spectrometer, using the MRM transitions determined and parameters described in sections 2.7.1 and 2.7.2.

## **2.8 Determining mammalian cell proliferation rate and cell cycle kinetics**

### **2.8.1 Determining proliferation rate**

KBM7 and NuKO cells were each seeded at a density of  $2 \times 10^6$  cells per 10 cm plate ( $1.66 \times 10^7$  cells/mL). Each cell line was split every other day to ensure that available nutrients were not a limiting factor. The cell number for each cell line was recorded

daily for a total of seven days. Counted cell number was then multiplied up by the amount split, to reflect the total number of cells produced across the 7-day period.

## **2.9 Flow cytometry**

### **2.9.1 Measurement of cell cycle kinetics**

KBM7 and NuKO cells were incubated in medium containing 5-ethynyl-2'-deoxyuridine (EdU, 1:1000; Click-iT™ EdU Alexafluor™ 555 Imaging kit; C10338, invitrogen) for 1 hour, then centrifuged (500 x *g*, 5 minutes). The cell pellet was washed in DPBS (Gibco), then the resulting suspension was centrifuged (500 x *g*, 5 minutes). The cells were then resuspended in PBS containing bovine serum albumin (BSA; 1% w/v; BP9701-100, Fisher BioReagents™) and centrifuged once more (500 x *g*, 5 minutes). To fix the cells, the pellet was resuspended in 4% w/v formaldehyde (15 minutes, room temperature; J19943-K2, Thermo Scientific), then centrifuged (500 x *g*, 5 minutes). To wash the cells, the supernatant was removed, and the pellet was incubated in PBS containing BSA (1% w/v, 5 minutes), then centrifuged (500 x *g*, 5 minutes). This wash step was repeated twice more. The supernatant was removed before the cell pellet was permeabilised in 0.5% v/v triton in PBS for 15 minutes. After permeabilisation, the cells were washed three times in 3% w/v BSA in 1X PBS (5-minute washes), with a centrifugation step (500 x *g*, 5 minutes) to remove the supernatant after each step. The cell pellet was resuspended in EdU cocktail (Click-iT™ EdU Alexafluor™ 555 Imaging kit details; C10338, invitrogen), incubated in the dark (15 minutes, room temperature) and centrifuged (500 x *g*, 5 minutes). The cell pellet was retained and washed three times in 0.1% v/v triton in PBS (5-minute washes), with a centrifugation step and removal of the supernatant between each wash. Finally, the cell pellet was incubated in Hoescht 33342 (Click-iT™ EdU Alexafluor™ 555 Imaging kit details; C10338, Invitrogen) diluted 1:2000 in PBS containing 0.1% v/v triton (500 µL total volume). The sample was kept on ice and in the dark until analysis on the CytoFLEX flow cytometer (Beckman Coulter) using the PB-450 and PE channels.

### **2.9.2 Measurement of apoptosis**

KBM7 and NuKO cells were treated with HU for 2 hours, then allowed to recover for 24 hours. After recovery, the cells were centrifuged (500 x *g*, 5 minutes), then the resulting pellet was washed in cold PBS (1 mL). The suspension was centrifuged (500 x *g*, 5 minutes) and the pellet was resuspended in fresh medium (1 mL). The Yo-Pro 1 stock (Thermofisher Y3603) was diluted 1:10 in DMSO, then added at a dilution of 1:1000 to each tube. Propidium Iodide (PI, 1 µg/mL; (Sigma Aldrich P4170) was also added to each tube, and all samples were incubated on ice for 20–30 minutes. Compensation was performed using permeabilised cells, cells with PI staining only, and cells with Yo-Pro1 staining only for each cell line. Apoptosis was then measured with a CytoFLEX flow cytometer (Beckman Coulter), using the PE and FITC channels.

### **2.10 Analysis of the $\gamma$ H2AX response to cell stress using Western blotting (WB)**

#### **2.10.1 Inducing $\gamma$ H2AX stress in KBM7 and NuKO cells**

KBM7 and NuKO cells were each incubated in hydroxyurea (HU; 1 mM, 37 °C) for 2 hours. The cells were then centrifuged (500 x *g*, 5 minutes), and washed in pre-warmed DPBS. The cells were centrifuged once more (500 x *g*, 5 minutes), and the pellet was resuspended in fresh, pre-warmed medium. The cells were then allowed to recover by incubating in the fresh medium at 37 °C for a further 0, 1, 2, 4, or 24 hours. At the appropriate time point, the cells were centrifuged (500 x *g*, 5 minutes), and the pellet was washed in pre-warmed DPBS. The DPBS was removed by centrifugation (500 x *g*, 5 minutes), and the resulting cell pellet was resuspended in 4X Sodium Dodecyl Sulfate (SDS) loading buffer (200 mM Tris-Cl (pH 6.8), 200 mM DTT, 0.8% w/v SDS, 0.4% w/v bromophenol blue, 40% w/v glycerol) containing Phenylmethylsulphonyl Fluoride (PMSF, Sigma Aldrich; 1 mM). Finally, each sample was incubated at 95 °C for 10 minutes in a dry bath (Fisher Scientific, FB15103), and stored at -20 °C until analysis.

Samples prepared for Western blotting were thawed, then heated at 95 °C for 10 minutes. For molecular weight comparison, samples were run alongside a molecular



weight marker (Pierce™ Prestained Protein molecular weight marker, Thermo Scientific™, Cat# 26612; or PageRuler plus Prestained Protein ladder, Cat# 26619, Thermo Scientific). Samples were loaded and run through the stacking gel (5% w/v; 100 V, Bio-Rad, PowerPac™ Basic) then through the Sodium Dodecyl Sulfate–Polyacrylamide Gel Electrophoresis (SDS-PAGE) resolving gel (10% w/v, 200 V). The current was stopped prior to the dye front reaching the bottom of the gel. The samples were then transferred onto polyvinylidene fluoride (PVDF) membrane (Amersham™) using a semi-dry transfer machine (90 minutes, 63 mA/gel) connected to a PowerEase 500 (Invitrogen life technologies) power supply. The membranes were cut and blocked in blocking buffer (BSA, 1% w/v; 1 x TBS; Tween20, 0.1% v/v) for 1 hour. After blocking, the membranes were incubated in primary antibody diluted in blocking buffer at 4 °C overnight. The dilutions for each primary antibody are shown in Table 2.4

**Table 2.4: Primary antibody species and dilutions**

Antibody	Species	Dilution	Supplier
Mcm2	Mouse	1:500	BD biosciences
PCNA	Mouse	1:250	Santa Cruz Biotechnology
γH2AX	Mouse	1:2500	abcam
Histone H3	Rabbit	1:10,000	abcam
Actin	Mouse	1:10,000	Sigma-Aldrich

### 2.10.2 Analysis by Western blotting

After overnight antibody binding, excess antibody was removed by washing four times in blocking buffer (5-minute washes). The membrane was incubated in the appropriate secondary antibody (1:5000; rabbit, Sigma, (A0545) Anti-Rabbit IgG (whole molecule) -Peroxidase antibody produced in goat; or mouse, Invitrogen (A28177) Goat anti-Mouse IgG (H+L) Superclonal™ Recombinant secondary antibody, HRP) depending on species (see Table 2.4) at room temperature for 1 hour. The membranes were subsequently washed four times in wash buffer (1X TBS [diluted from 10x TBS, pH 7.6]; Tween20, 0.1% v/v; 5-minute washes) to remove excess secondary antibody. Supersignal® West Pico enhanced chemiluminescence (ECL) Western blotting substrate solutions A and B (Thermo Scientific) were mixed 1:1 and

added to the membrane immediately prior to imaging on a ChemiDoc™ MP Imaging System (Bio-Rad). For analysis, band intensity was calculated using the ImageLab software (Bio-Rad), and bands were standardised to actin, or histone when the sample was composed of nuclei only.

## **2.11 Cell synchronisation**

### **2.11.1 Synchronisation of 3T3 cells in mid- and late-G1**

3T3 cells were maintained at a density of ~40%, then the medium was changed, and the cells were incubated for 48 hours (37 °C, 5% CO<sub>2</sub>). Subsequently, the medium was removed and replaced with fresh medium. The confluent cells were incubated for a further 48 hours (37 °C, 5% CO<sub>2</sub>). The cells were then incubated with 10X trypsin (2 mL) diluted in pre-warmed DPBS (8 mL, Gibco) for 2 minutes, or until fully detached. After dissociation, the cells were split 1:4 and incubated for either 15 hours (for mid-G1 extract) or 17.5 hours (for late G1-phase nuclei) before harvesting.

### **2.11.2 Synchronisation of S-phase HeLa cells**

HeLa cells were split to ~30% confluency then incubated in supplemented D-MEM medium containing thymidine (2.5 mM) for 24 hours, to inhibit DNA replication. The thymidine-containing medium was removed, then cells were washed in pre-warmed DPBS (Gibco), and the removed medium was replaced with fresh D-MEM medium without thymidine. The cells were incubated for 8 hours to allow any non-stalled cells to reach early S-phase, then the medium was removed and replaced with medium containing thymidine (2.5 mM). The cells were incubated in the thymidine containing medium for 16 hours, to stall any remaining cells in early S-phase. Finally, the medium was replaced with fresh medium (no thymidine), and incubated for 1 hour, to allow all cells to progress fully into S-phase. After 1 hour, the cells were harvested as described in section 2.11.3.

### **2.11.3 Harvesting synchronised nuclei and extract**

Harvesting of both nuclei and cytosolic extracts was performed at 4 °C. The medium was removed, and all plates were washed in ice-cold hypotonic buffer (20 mM HEPES,

5 mM potassium acetate; 0.5 mM  $\text{MgCl}_2$ , 1 mM DTT). The cells were incubated in ice-cold hypotonic buffer for 5 minutes, then the buffer was discarded. Excess hypotonic buffer was discarded by tilting the plates at a  $45^\circ$  angle and leaving for 3–5 minutes to allow the collection and removal of excess buffer from the bottom of the plate. The cells were then harvested by scraping, and the nuclei and soluble extract fractions were extracted from the cells by Dounce homogenisation (Wheaton Dounce homogeniser, tight pestle) followed by centrifugation ( $6000 \times g$ , 5 minutes for nuclei;  $17,000 \times g$ , 10 minutes for extract). For nuclei, the pellet was retained and resuspended in an equal volume of hypotonic buffer. The resulting suspension was snap-frozen in liquid nitrogen as 10–12  $\mu\text{L}$  pellets. For the extract, the suspension was directly frozen in liquid nitrogen, as 30–40  $\mu\text{L}$  pellets. Both nuclei and extracts were stored in liquid nitrogen.

## **2.12 Cell-free replication assays**

### **2.12.1 Cell-free replication**

Pre-mix (40 mM HEPES, pH 7.8 using sodium hydroxide; 7 mM  $\text{MgCl}_2$ ; 1 mM DTT; 40 mM phosphocreatine; 3 mM ATP; and GTP, CTP, UTP, dATP, dGTP, dCTP, all 0.1 mM) was added to both G1- and S-phase extract at a ratio of 1:10.  $\text{MgCl}_2$  (0.1 M), creatine phosphokinase (CPK, 10 mg/mL; Millipore-MERCK) and biotin-dUTP (Roche) were then added 1:50 into the resulting extracts. The extracts were divided into 10  $\mu\text{L}$  aliquots for cell-free immunofluorescence assays only, or 15  $\mu\text{L}$  aliquots if done in parallel with the WB experiments. The synthesised  $\text{Ap}_4\text{Ns}$ , synthesised AppNppA or  $\text{Ap}_4\text{A-FRET}$  (provided by Professor Andreas Marx) were diluted in water or HEPES (20 mM, pH 7.8 with sodium hydroxide), then added into the appropriate tubes to a final concentration of 100  $\mu\text{M}$ . The extracts were mixed, then late G1 nuclei were added to each tube (1  $\mu\text{L}$  for cell-free only, 3  $\mu\text{L}$  for cell-free done in parallel with WB). The nuclei were resuspended in the extracts by flicking gently, then all samples were incubated in a dry bath at  $37^\circ\text{C}$  for 30 minutes. The cell extracts were resuspended in 4% w/v formaldehyde (120  $\mu\text{L}$ ) and incubated at room temperature for 15 minutes. Poly-lysine coated glass coverslips were inserted into the bottom of snaptwist #PE 6.5 mL scintillation vials and covered with 30% w/v sucrose (0.8 mL). The fixed cell extract

was then added gently, as a layer to the top of the sucrose, and centrifuged at 1000 RPM for 10 minutes. Next, the coverslips were incubated in antibody buffer (PBS, 0.1% v/v Triton X-100, 0.02% w/v SDS, 1% w/v BSA) for 5 minutes. This was followed by two more quick washes in antibody buffer. Each coverslip was then incubated in streptavidin Alexa Fluor 555 diluted 1:1000 in antibody buffer for 30 minutes in a humidity chamber at 37 °C, 5% CO<sub>2</sub>. Excess streptavidin Alexa Fluor was subsequently removed by washing three times in antibody buffer then once in PBS. Each coverslip was mounted in mounting medium containing 4',6-diamidino-2-phenylindole (DAPI; Vectashield®, H-1200; Vector Laboratories) and stored in the dark at 4 °C. Fluorescence imaging was performed on a Zeiss Axio Scope A1 microscope.

### **2.12.2 Preparation of WB to determine nuclear replication protein level**

For WB, the cell-free reaction was performed as described in section 2.12.1, up to and including the 30-minute incubation in the dry bath. At this point, instead of the fixation step, the nuclei were incubated in 0.5% v/v Triton X-100 and 1 mM PMSF on ice for 5 minutes, then centrifuged (17,000 x *g*, 5 minutes), to separate the nuclei from the extract fractions. Each sample was then resuspended in 4X SDS loading buffer containing DTT (1:6 dilution), then heated to 95 °C for 10 minutes. Samples were stored at -20 °C until use, then run on SDS-PAGE gels as described in section 2.10.2.

# **Chapter 3: Synthesis and purification of $\text{Ap}_4\text{A}$ and $\text{Ap}_4\text{N}$ dinucleotide tetraphosphates**

### 3.1 Introduction to the synthesis of Ap<sub>4</sub>A and Ap<sub>4</sub>Ns

#### 3.1.1 Ap<sub>4</sub>A structure and synthesis

Ap<sub>4</sub>A is a signalling molecule composed of two adenosine moieties linked by a chain of four phosphate groups. Ap<sub>4</sub>A has been found in many different domains of life and can be synthesised through a variety of mechanisms (Ferguson et al., 2020). Ap<sub>4</sub>A can be synthesised either by the direct condensation of two ATP molecules by human glycyl-tRNA synthetase (GlyRS) (Guo et al., 2009), or through the reaction of ATP with an intermediate such as T4 RNA ligase-AMP (Atencia et al., 1999), AcylCoA-LH<sub>2</sub>-AMP in *P. fragi* (Fontes et al., 1998), or an aminoacyladenylate in both prokaryotes and eukaryotes (Brevet et al., 1989b; Goerlich et al., 1982). Some of these reaction mechanisms are also capable of synthesising other Ap<sub>4</sub>Ns in addition to Ap<sub>4</sub>A (Atencia et al., 1999; Brevet et al., 1989b; Fontes et al., 1998). An additional mechanism for Ap<sub>4</sub>A synthesis has been found recently, where Ap<sub>4</sub>A is produced as a by-product in the activation of Ub, by Ub- and Ub-like activating enzymes (Götz et al., 2019). This is particularly interesting as the activity of Ub-activating enzyme is increased under stress conditions (Shang et al., 1997).

In addition to its biological synthesis, Ap<sub>4</sub>A can also be produced chemically via a number of different mechanisms involving either P(III) or P(V) chemistry (Appy et al., 2019a). In the P(V) mechanism, a 5'-nucleotide is activated, followed by reaction of the resulting intermediate with a second 5'-nucleotide or inorganic pyrophosphate to produce dinucleotide polyphosphates of varying lengths (Appy et al., 2019a). Methods utilising P(V) chemistry for the synthesis of dinucleotide tetraphosphates began as early as the 1960s, where Ap<sub>3</sub>A, Ap<sub>4</sub>A and Ap<sub>5</sub>A were identified as products of a dismutation reaction between adenosine 5'-phosphoromorpholidate and tributylammonium pyrophosphate in an anhydrous pyridine (Reiss and Moffatt, 1965). Another P(V) chemistry approach enables reaction in water, producing diadenosine polyphosphates from the reaction of AMP, ADP or ATP with a carbodiimide (Ng and Orgel, 1987).

Alternatively, mechanisms using P(III) chemistry can be applied to the synthesis of dinucleotide tetraphosphates. P(III) reagents result in phosphite or phosphoramidite

intermediates which then oxidise to form P(V) molecules (Appy et al., 2019a). A mechanochemistry approach has also been used for the synthesis of dinucleotide tetraphosphates, where a ball-mill was used for the coupling of 5'-adenosine phosphomorpholidate to phosphorylated ribose derivatives (Ravalico et al., 2011). More recently, a one-pot approach has been demonstrated for the synthesis of  $\text{Np}_4\text{N}$  from NMPs which makes use of readily available reagents and does not require protective groups (Depaix et al., 2017). A solvent-assisted mechanochemistry technique was later applied for the synthesis of dinucleotide tetraphosphates from nucleotide 5'-monophosphates and could produce the nucleotide tetraphosphate within short reaction times (Appy et al., 2019b).

### **3.1.2 $\text{Ap}_4\text{A}$ signalling mechanisms**

In some of its signalling mechanisms,  $\text{Ap}_4\text{A}$  appears to function by competitive inhibition. An example of this is in the STING pathway, where  $\text{Ap}_4\text{A}$  binds STING via the same binding pocket as cGAMP, delaying the cGAMP-STING interaction and therefore the interferon response (Guerra et al., 2020). Likewise,  $\text{Ap}_4\text{A}$  can bind an ATP-binding domain on HINT1, displacing MITF and causing HINT1 to polymerise (Yu et al., 2019). Interestingly, in the HINT1 signalling pathway both ends of the  $\text{Ap}_4\text{A}$  molecule are essential for polymerisation, with each end binding to an ATP-binding pocket on a separate HINT1 dimer (Yu et al., 2019).

$\text{Ap}_4\text{A}$  alters cell signalling at a transcriptional level in bacteria, where  $\text{Ap}_4\text{Ns}$  can cap RNA, protecting the mRNA from degradation and thereby influencing cell signalling (Hudeček et al., 2020; Luciano and Belasco, 2020; Luciano et al., 2019). This signalling pathway is not specific to  $\text{Ap}_4\text{A}$ , as other  $\text{Ap}_4\text{Ns}$  are also capable of forming  $\text{Np}_4$  caps on the RNA, demonstrating that only one end of the  $\text{Ap}_4\text{A}$  structure is needed for this kind of signalling to occur.

Another signalling mechanism that  $\text{Ap}_4\text{A}$  has been implicated in is the inhibition of DNA replication (Marriott et al., 2015). However, the mechanism behind how it achieves this inhibition is not known. As a step towards elucidating this pathway, a chemical biology approach was taken to determine whether both ends of the molecule were

important for signalling (as in the HINT1 signalling pathway) or whether only one end of the molecule was essential for signalling (as demonstrated by Ap<sub>4</sub>N-capping of transcripts in *E. coli*). The first step in elucidating this pathway was to begin with the synthesis of Ap<sub>4</sub>A.

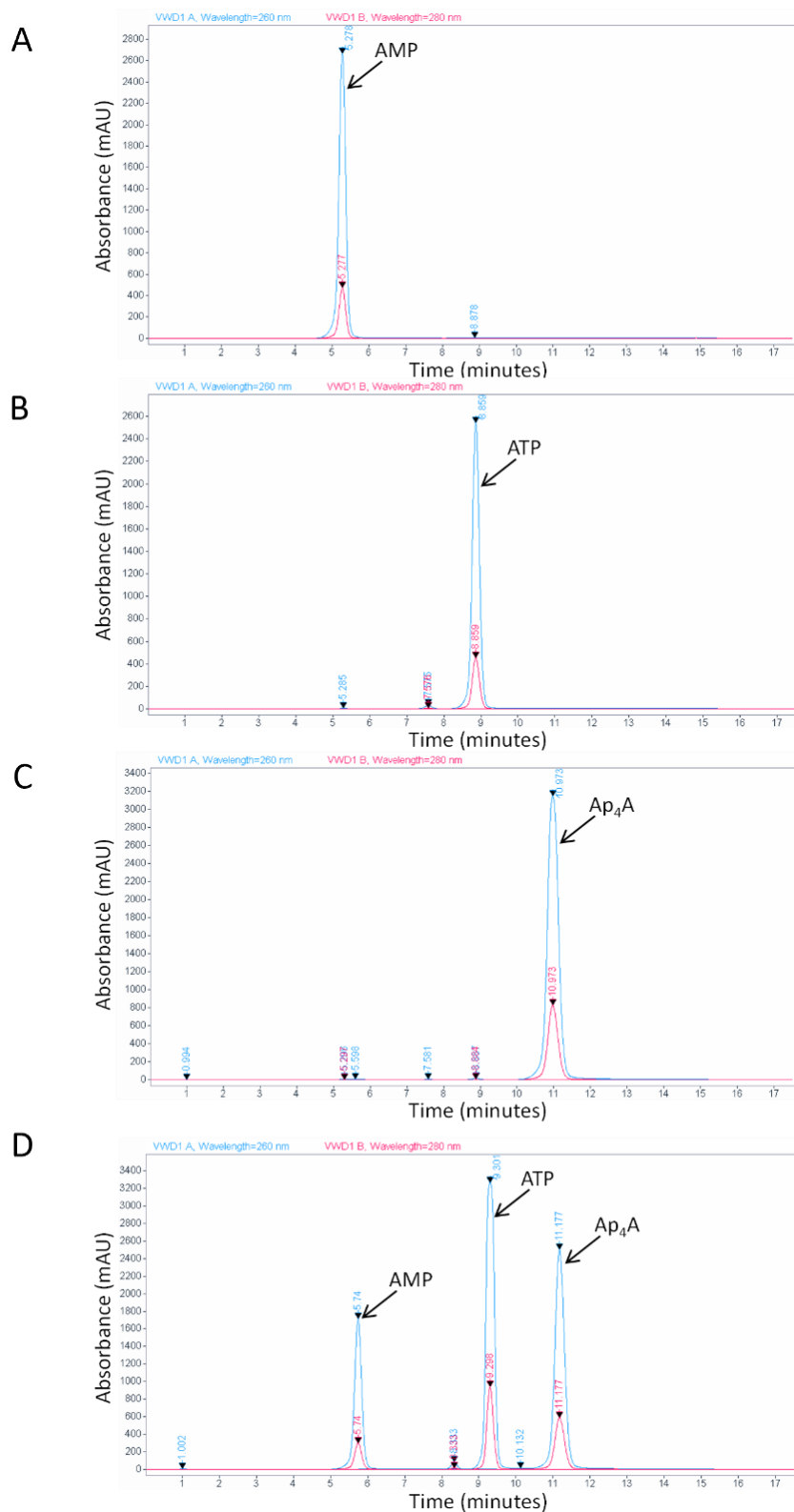
### 3.1.3 Chapter Aims

To determine the functionality of Ap<sub>4</sub>A in DNA replication, a chemical biology approach was taken to investigate the potential inhibitory and binding mechanisms that could mediate its activity. In other model systems, Ap<sub>4</sub>A can bind through either adenosine moiety with the other being solvent exposed. In this scenario, the presence of an alternative nucleoside such as C, U or G may not affect function. However, if binding is mediated by a homobifunctional Ap<sub>4</sub>A such as in the mechanism used by Ap<sub>4</sub>A to mediate HINT oligomerisation (Yu et al., 2019), a heterobifunctional Ap<sub>4</sub>N may not be active. As a first step, the chemical synthesis of Ap<sub>4</sub>A was performed for use in future experiments and to facilitate synthesis of other Ap<sub>4</sub>N molecules. The next aim was to synthesise and verify the identity a variety of different Ap<sub>4</sub>Ns and Ap<sub>4</sub>dNs, where N=A, C, G or U/T, so we could determine what effects these changes in structure would have on the function of the Ap<sub>4</sub>Ns in the context of DNA replication. Finally, we aimed to determine if the synthesis of Ap<sub>4</sub>N molecules was possible in biologically relevant systems using the Ub-activating enzyme-mediated Ap<sub>4</sub>A synthesis mechanism demonstrated by Götz and colleagues in 2019.

### 3.2 Preparation of standards for HPLC

Commercial standards for AMP, ATP and Ap<sub>4</sub>A were analysed by HPLC with a ResourceQ column to determine the retention times for each molecule (Fig. 3.1). As the ResourceQ column separates molecules according to their charge, the large changes associated with the number of phosphate groups in AMP, ATP and Ap<sub>4</sub>A resulted in a good separation of these molecules.



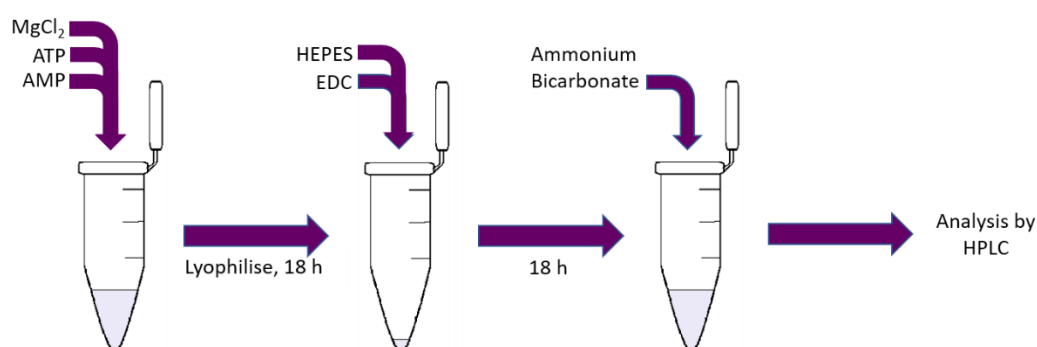


**Figure 3.1: Establishing retention times for substrates and products of the Ap<sub>4</sub>A synthesis reactions.** Standards for AMP (A), ATP (B) and Ap<sub>4</sub>A (C) or all three molecules (D) were bound to a 1 mL ResourceQ column and eluted with an increasing linear gradient of ammonium bicarbonate pH 9.6 from 0.05–1 M to determine their retention times (mins): AMP=5.3; ATP=8.9; Ap<sub>4</sub>A=11.0.

AMP eluted after 5.3 minutes (Fig. 3.1 A), ATP after 8.9 minutes (Fig. 3.1 B), and Ap<sub>4</sub>A after approximately 11 minutes (Fig. 3.1 C). When the standards were injected as a mixed sample, their retention times shifted slightly to the right - AMP eluted at 5.7 minutes, ATP at 9.3 minutes and Ap<sub>4</sub>A at 11.2 minutes (Fig. 3.1 D). The clear separation of each molecule allows the accurate prediction of the elution time for the Ap<sub>4</sub>A product in subsequent reactions.

### 3.3 Optimisation of Ap<sub>4</sub>A synthesis

Ap<sub>4</sub>A was synthesised chemically from ATP and AMP in the presence of MgCl<sub>2</sub>, HEPES and EDC, then analysed by HPLC. A schematic for this chemical synthesis method is shown in Figure 3.2. To optimise the Ap<sub>4</sub>A synthesis method, several variables were altered independently, to determine which of these influenced yield. Therefore, the effects of variables such as ATP concentration, buffer concentration, activating agent concentration and reaction temperature needed to be measured.

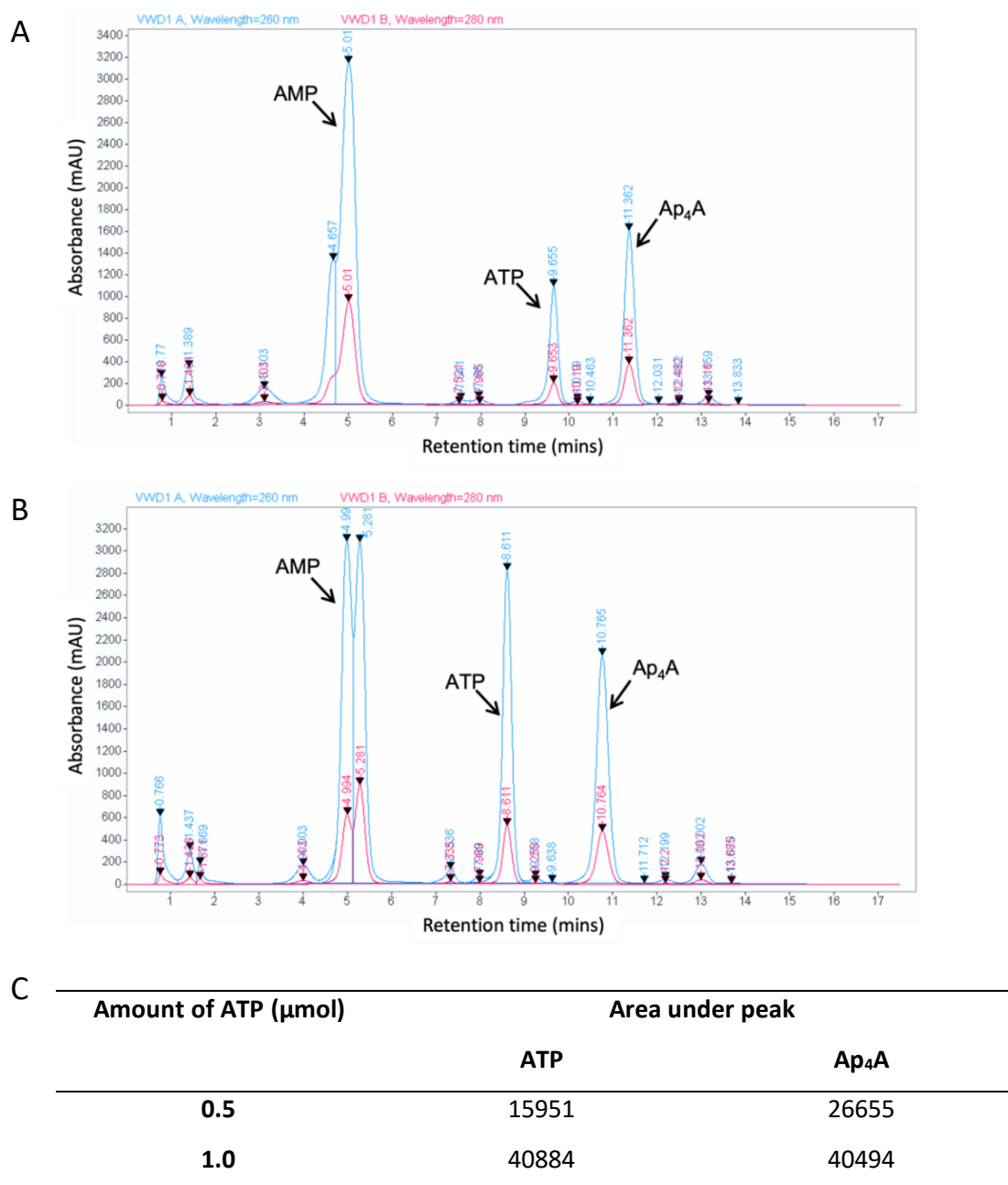


**Figure 3.2: Schematic for the chemical synthesis of Ap<sub>4</sub>A.** ATP, AMP and MgCl<sub>2</sub> were combined and lyophilised overnight to remove water. HEPES and EDC were then added to initiate the reaction. The reaction mixture was incubated overnight for at least 18 hours, then quenched in ammonium bicarbonate. A bolus was then injected into the HPLC and run through a ResourceQ column on an ammonium bicarbonate:water gradient.

#### 3.3.1 Effect of increased ATP in the reaction mixture

In the initial reaction, the molar ratio of the key reactants AMP, ATP and EDC is 6:1:50. ATP is a necessary component in the Ap<sub>4</sub>A reaction mixture, as it reacts with AMP to produce a diadenosine molecule with the appropriate number of phosphate groups.

Therefore the effect of increasing the proportion of ATP on the reaction kinetics was investigated by altering the molar ratio of AMP:ATP:EDC to 3:1:25 (Fig. 3.3).

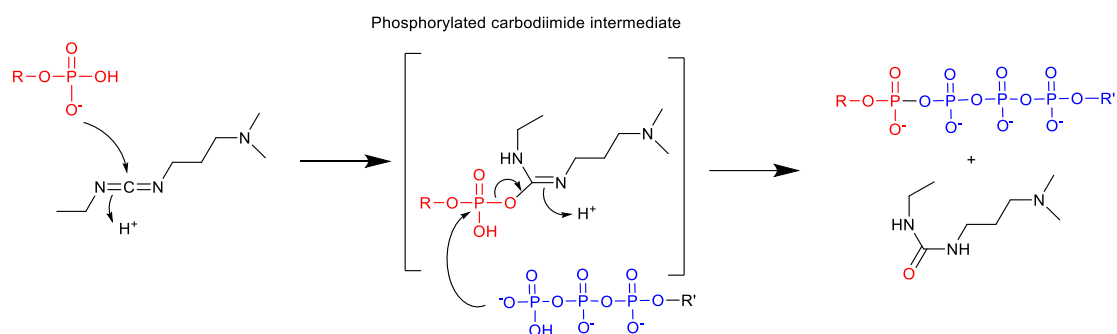


**Figure 3.3: Effect of increased ATP concentration on Ap<sub>4</sub>A synthesis.** A, B: Ap<sub>4</sub>A was synthesised chemically following the method shown in Figure 3.2, with either 0.5  $\mu\text{mol}$  or 1.0  $\mu\text{mol}$  Ap<sub>4</sub>A. A+B) HPLC traces for Ap<sub>4</sub>A reaction mix with 0.5  $\mu\text{mol}$  (A) and 1.0  $\mu\text{mol}$  (B) ATP. C: Summary of the area under the peaks corresponding to ATP and Ap<sub>4</sub>A.

Increasing the amount of ATP in the reaction mixture increased the amount of Ap<sub>4</sub>A produced in the reaction (Fig. 3.3). The area under the peak corresponding to Ap<sub>4</sub>A is a direct measure of the amount produced. This increases from 26,655 (Fig. 3.3 A and C) to 40,494 (Fig. 3.3 B and C) when the concentration of ATP is doubled. Unfortunately increasing the ATP concentration in the reaction mix resulted in an increase in the amount of unused ATP. To prevent waste, this ATP fraction could be collected and purified on the HPLC. Due to the higher amount of product obtained, 1.0  $\mu$ mol ATP/NTP was used in future Ap<sub>4</sub>N synthesis reactions. The AMP peak showed a double peak, which could be caused by several factors. As this was not seen with the molecules eluting later in the run, this is likely explained by an incompletely filled sample loop during injection. As the AMP concentration was already in excess, with a molar ratio to ATP of 6:1, altering it was unlikely to influence the reaction efficiency as this was already saturated in the reaction mixture. Therefore, it was unnecessary to experiment further with the AMP concentration.

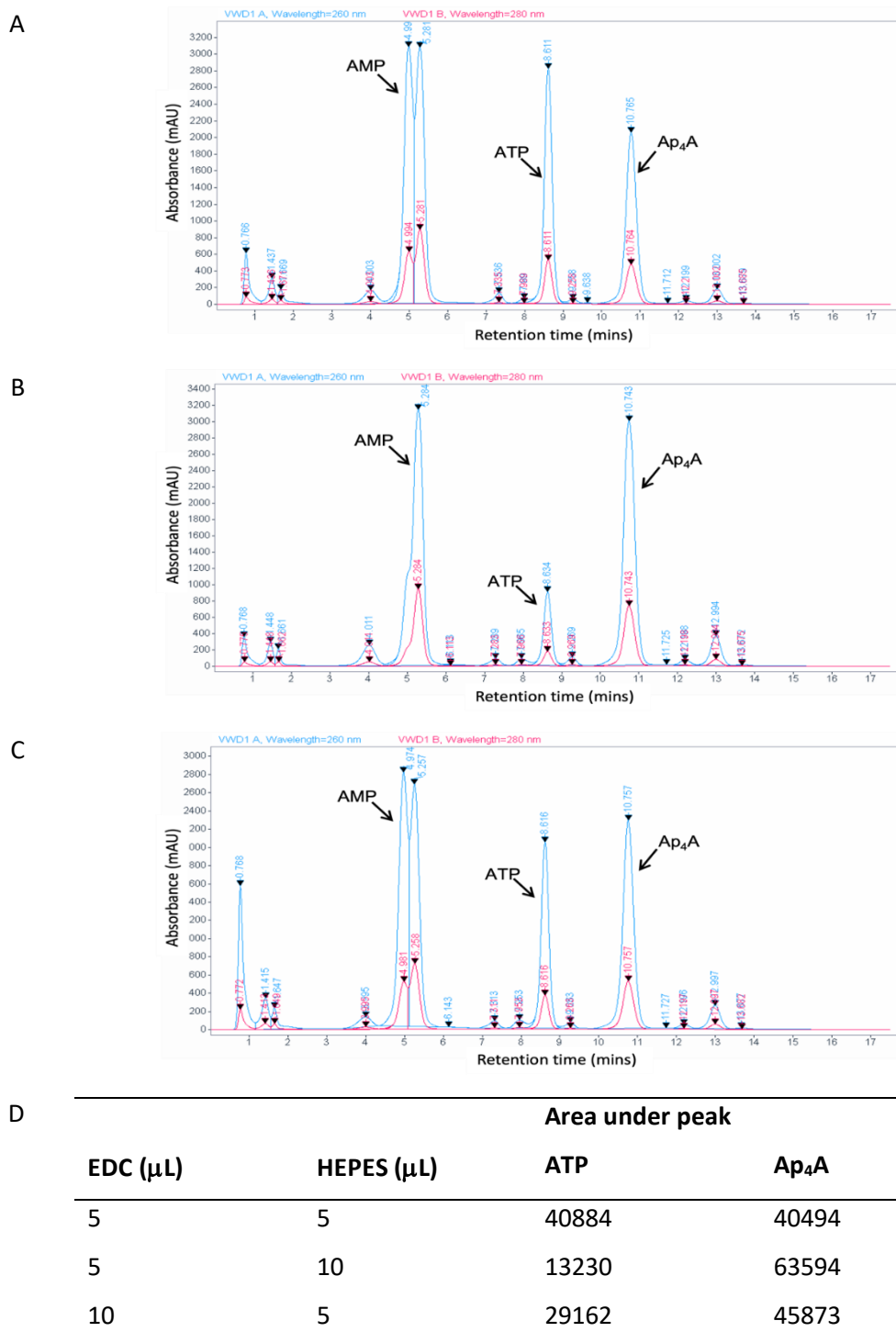
### 3.3.2 Effect of HEPES and EDC on increasing Ap<sub>4</sub>A yield

EDC is an activating agent that is essential in initiating the reaction between AMP and ATP. In this process, the AMP phosphate (or the  $\gamma$ -phosphate on ATP) reacts with the EDC to form a phosphorylated carbodiimide intermediate (Fig. 3.4). The intermediate then reacts with the  $\gamma$ -phosphate on NTP (or the AMP phosphate) to produce Ap<sub>4</sub>N and a urea by-product.



**Figure 3.4: Carbodiimide coupling reaction for Ap<sub>4</sub>A synthesis.** The mechanism for the formation of the phosphorylated carbodiimide intermediate, followed by reaction with NTP to form the Ap<sub>4</sub>N product and a urea by-product. R = Adenine nucleoside; R' = A, C, G or U nucleosides.

As EDC availability is necessary for the reaction to occur, an increased amount of EDC in the reaction mix was tested alongside the higher ATP amount (molar ratio of AMP:ATP:EDC = 3:1:50), to determine whether EDC concentration in the reaction mixture could be a limiting factor (Fig. 3.5). Furthermore, the reaction is performed in a very low volume to increase the chances of successful collisions. However, as the EDC and buffer are both required to be at very high molarity, some problems with solubility were observed. To assess whether increasing the amount of buffer could influence Ap<sub>4</sub>A yield, different volumes of HEPES were added to the reaction mix to alter reagent concentrations (molar ratio of AMP:ATP:EDC = 3:1:25) and the Ap<sub>4</sub>A produced was analysed by HPLC (Fig. 3.5).

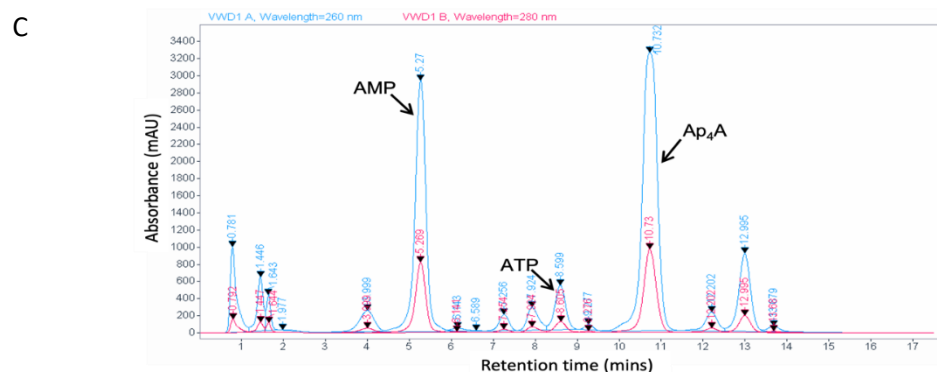
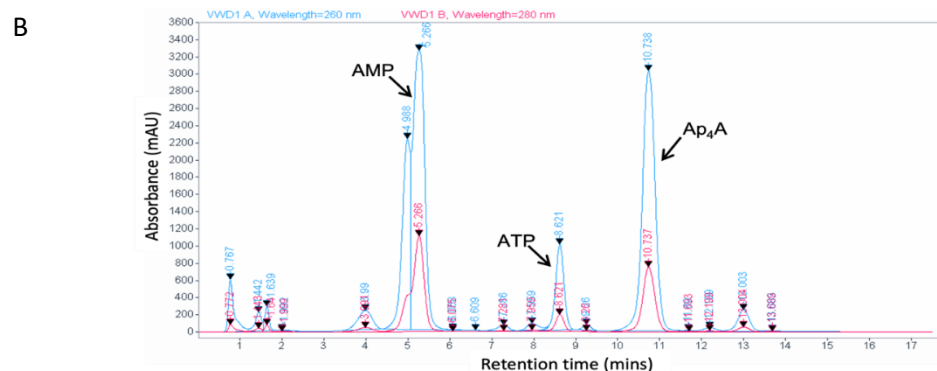
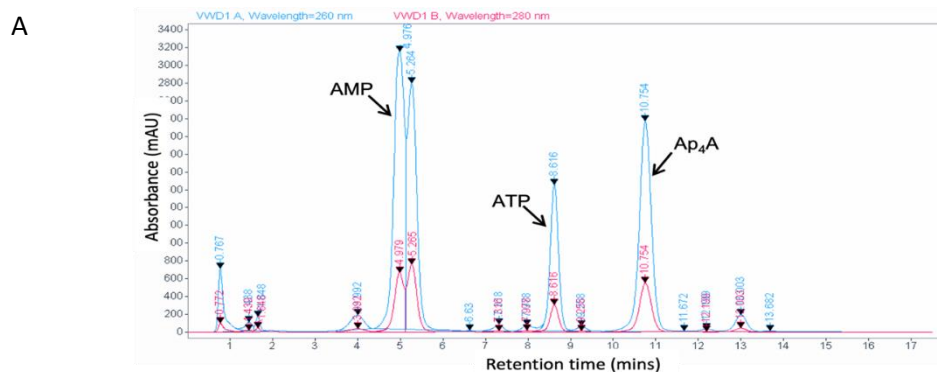


**Figure 3.5: Effect of increased volume of HEPES and EDC on Ap<sub>4</sub>A synthesis.** Ap<sub>4</sub>A was synthesised chemically following the method shown in Figure 3.2, with differing volumes of EDC or HEPES buffer. **A:** HPLC trace for Ap<sub>4</sub>A reaction mix with equal volumes of EDC and HEPES. **B, C:** HPLC traces for Ap<sub>4</sub>A reaction mix with increased volume of HEPES buffer (B) and increased volume of EDC reagent (C). **D:** Summary of the area under each peak corresponding to ATP and Ap<sub>4</sub>A.

Analysis of each reaction condition by HPLC identified that increased volume of HEPES buffer led to an increased the yield of Ap<sub>4</sub>A from 40,494 units to 63,594 units and decreased the unused ATP from 40,884 units to 13,230 units, increasing the yield of the reaction relative to ATP and reducing ATP waste (Fig. 3.5 B and D). In comparison, despite reducing the unused ATP in the reaction by 11,722 units, increasing the EDC volume only resulted in a slight increase in Ap<sub>4</sub>A yield, from 40,494 to 45,873 units (Fig. 3.5 C and D). Therefore, the increased volume of HEPES was used in all future reactions, while the volume of EDC was kept at its original volume.

### **3.3.3 Effect of temperature on Ap<sub>4</sub>A yield**

An increase in temperature generally results in an increase in rate of reaction. Therefore, to identify whether increasing the reaction temperature would increase Ap<sub>4</sub>A product yield, the Ap<sub>4</sub>A synthesis reaction was performed at different temperatures. The Ap<sub>4</sub>A reaction mix was incubated at either room temperature, 25 °C or 37 °C to determine the most efficient reaction temperature (Fig. 3.6).



D

	Area under peak	
Temperature (°C)	ATP	Ap <sub>4</sub> A
RT	22808	47262
25	14170	64724
37	9391	80842

**Figure 3.6: Effect of temperature on Ap<sub>4</sub>A yield.** The Ap<sub>4</sub>A synthesis reaction was set off at **A**: room temperature, **B**: 25 °C or **C**: 37 °C. The reaction mixtures were analysed by HPLC to determine the relative abundance of ATP and Ap<sub>4</sub>A at each temperature. **D**: Summary of the area under each peak corresponding to ATP and Ap<sub>4</sub>A.

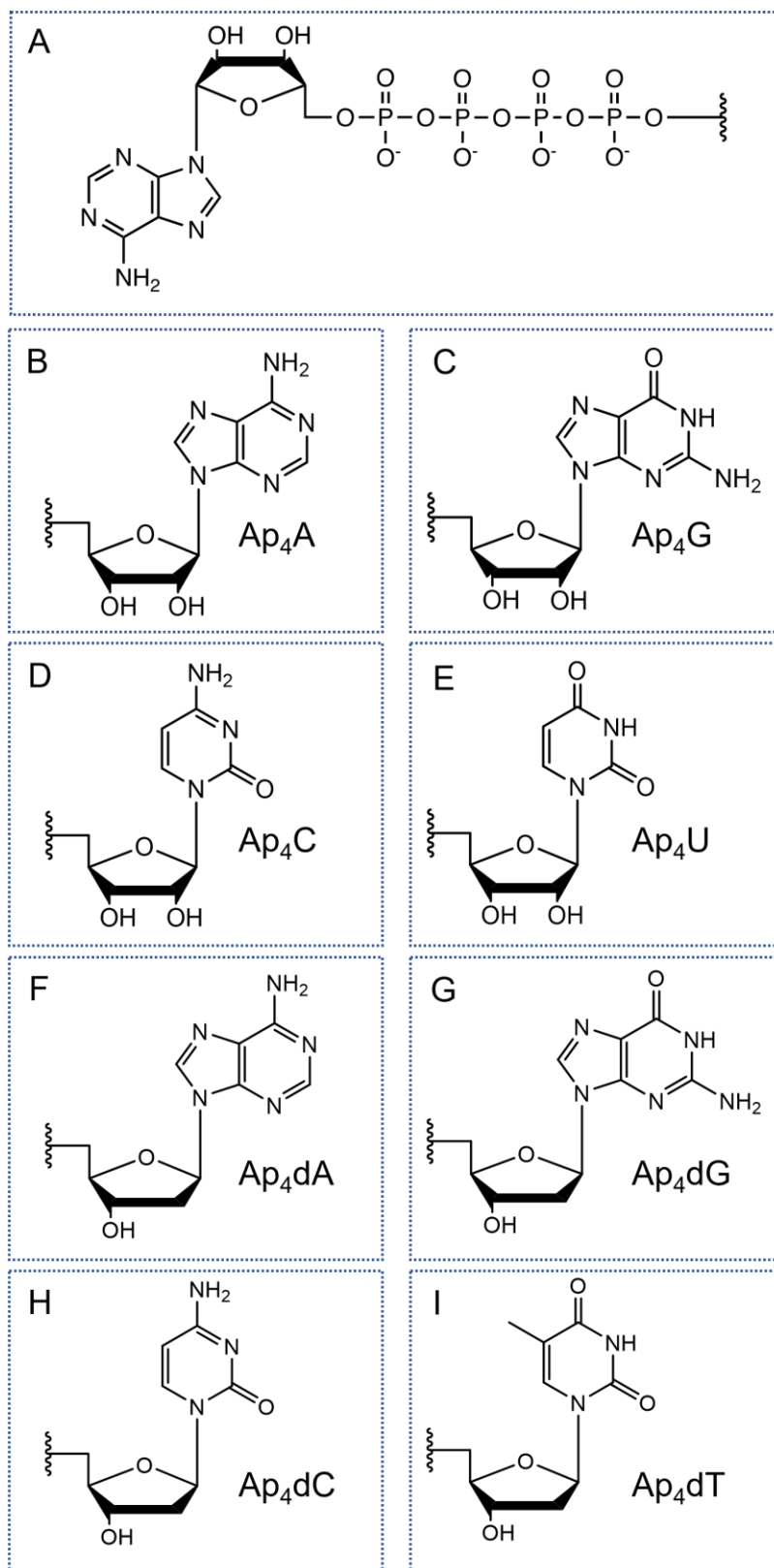


Ap<sub>4</sub>A was successfully synthesised at all temperatures (Fig. 3.6 A, B and C). Increasing the reaction temperature from room temperature to 25 °C increased Ap<sub>4</sub>A yield by 17,462 units and reduced the residual ATP by 37.8% (Figure 3.6 B and D). Increasing the reaction temperature to 37 °C increased the Ap<sub>4</sub>A yield by a further 16,118 units compared with the reaction at 25 °C and by 33,580 units compared with the reaction at room temperature (Fig. 3.6 C and D). However, the 37 °C reaction temperature also resulted in an increased number of side reactions, likely including unwanted products such as Ap<sub>6</sub>A (retention time ~13 mins) (Fig. 3.6 C). This was not a big issue, as the Ap<sub>4</sub>A peak was still clearly distinguishable from the side reactions, so Ap<sub>4</sub>A could still be purified. From this set of data, we therefore determined that future reactions should be performed at a temperature of at least 25 °C.

### **3.4 Synthesis of other dinucleotide tetraphosphates**

#### **3.4.1 Changes to Ap<sub>4</sub>A structure**

Ap<sub>4</sub>A is involved in several different signalling mechanisms (see section 3.1.2). Ap<sub>4</sub>A functions through either mono or bifunctional interactions in specific processes. To determine if one of the adenosine groups can be substituted by either purine or pyrimidine and retain activity, all ribosyl bases were used to produce Ap<sub>4</sub>A, Ap<sub>4</sub>U, Ap<sub>4</sub>G and Ap<sub>4</sub>C. Similarly, all deoxy-ribosyl bases were used to produce Ap<sub>4</sub>dA, Ap<sub>4</sub>dT, Ap<sub>4</sub>dG and Ap<sub>4</sub>dC. The production of each Ap<sub>4</sub>N and Ap<sub>4</sub>dN will enable characterisation of each molecule and their potential role in regulation of the DNA replication process. Each molecule was produced by reacting AMP with the each NTP/dNTP using the reaction conditions optimised above. A summary of the structural changes made is shown in Figure 3.7.



**Figure 3.7: Structural differences between different Ap<sub>4</sub>Ns.** The first part of the structure is conserved across all molecules (A) with the end part of the molecule changing depending on the nucleotide used for synthesis: Ap<sub>4</sub>A (B), Ap<sub>4</sub>G (C), Ap<sub>4</sub>C (D), Ap<sub>4</sub>U (E), Ap<sub>4</sub>dA (F), Ap<sub>4</sub>dG (G), Ap<sub>4</sub>dC (H), Ap<sub>4</sub>dT (I).

Previous work suggested that diguanosine tetraphosphate (Gp<sub>4</sub>G) was inactive and did not affect initiation of DNA replication (Marriott et al., 2015). Consequently, we hypothesised that at least one adenosine moiety was required for its inhibitory activity based on recent analyses showing adenosine mediating the key interactions in other pathways regulated by Ap<sub>4</sub>A (Ferguson et al., 2020; Yu et al., 2019). As adenine and guanine are both purines, Ap<sub>4</sub>G (Fig. 3.7 C) has the most similar structure to Ap<sub>4</sub>A (Fig. 3.7 B), in addition to Ap<sub>4</sub>dA and Ap<sub>4</sub>dG (Fig. 3.7 F and G). The remaining molecules are purine-pyrimidine structures linked by a tetraphosphate group (Fig. 3.7 D, E, H and I). The phosphate chain length was not investigated here as Ap<sub>3</sub>A and Ap<sub>5</sub>A do not inhibit the initiation of DNA replication (Marriott et al. 2015) suggesting that the distance between the nucleosides is important for their inhibitory activity in DNA replication assays.

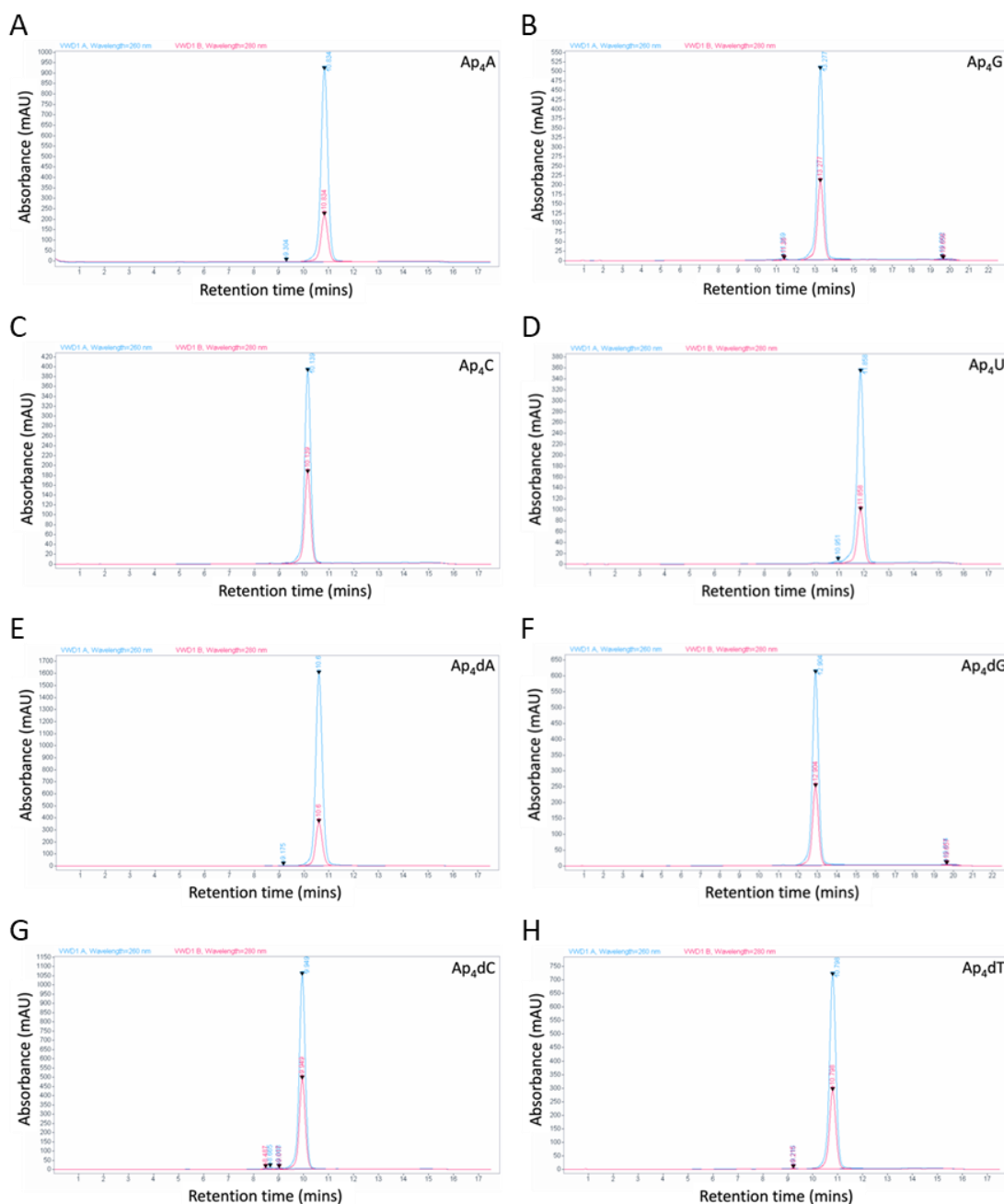
### 3.4.2 Synthesis and purification of Ap<sub>4</sub>Ns

The synthesis of each Ap<sub>4</sub>N was performed using the same chemistry as Ap<sub>4</sub>A, with each NTP or dNTP being substituted for ATP. Despite the similarity in structure, the small structural changes to each Ap<sub>4</sub>N caused them to elute from the ResourceQ column at different retention times (Table 3.1).

**Table 3.1: Retention times for each Ap<sub>4</sub>N.** Ap<sub>4</sub>Ns were chemically synthesised then analysed by HPLC. Standards for each NTP were analysed by HPLC to determine possible changes in retention time for each Ap<sub>4</sub>N.

	Elution time	
	NTP	Ap <sub>4</sub> N
Ap <sub>4</sub> A	8.9	10.7
Ap <sub>4</sub> G	10.4	13.1
Ap <sub>4</sub> C	7.8	10.0
Ap <sub>4</sub> U	9.1	11.7
Ap <sub>4</sub> T	8.3	10.7
Ap <sub>4</sub> dA	8.3	10.5
Ap <sub>4</sub> dG	10.1	12.8
Ap <sub>4</sub> dC	7.7	9.9

There were no standards available for the individual Ap<sub>4</sub>Ns and due to the scale at which we were synthesising the Ap<sub>4</sub>Ns it was not possible to verify their identity by NMR spectroscopy. Therefore, after synthesising and purifying each molecule, we planned to verify its identity first by analysis of its hydrolysis products, and then by HRMS. For each reaction, the retention time for each reagent was characterised and the retention time of the reaction products was monitored. A list of the retention times for each NTP and the Ap<sub>4</sub>N can be found in Table 3.1. After synthesis, each Ap<sub>4</sub>N was purified from the reaction mixture by injecting the full sample mixture and collecting the appropriate fraction from the HPLC. After lyophilising to remove the salt and concentrate the sample, all Ap<sub>4</sub>N samples were re-run to check their purity (Fig. 3.8).



**Figure 3.8: Purification of Ap<sub>4</sub>Ns and Ap<sub>4</sub>dNs.** Each Ap<sub>4</sub>N was synthesised and the appropriate fraction for each was separated out and collected by HPLC. Fractions were lyophilised and a sample was re-injected to determine the purity.

As expected, all reactions produced a new peak that eluted later than the AMP and NTP standard, consistent with formation of the desired Ap<sub>4</sub>N. This peak was then successfully purified, and the retention time for each Ap<sub>4</sub>N was unique: Ap<sub>4</sub>G eluted at 13.2 minutes (Figure 3.8 B), Ap<sub>4</sub>C at 10.1 minutes (Fig. 3.8 C) and Ap<sub>4</sub>U at 11.9 minutes (Fig. 3.8 D). As the Ap<sub>4</sub>dNs contained one less oxygen atom than their Ap<sub>4</sub>N

counterparts, they eluted from the column slightly earlier. Ap<sub>4</sub>dA eluted at 10.6 minutes (Fig. 3.8 E), Ap<sub>4</sub>dG eluted at 12.9 minutes (Fig. 3.8 F), Ap<sub>4</sub>dC at 9.9 minutes (Fig. 3.8 G) and Ap<sub>4</sub>dT at 10.8 minutes (Fig. 3.8 H). The purity of each Ap<sub>4</sub>N was evaluated by determining the area under the Ap<sub>4</sub>N peak at an absorbance of 260 nm as a proportion of the total area under all detected peaks. A summary of the purity estimated for each molecule can be found in Table 3.2.

**Table 3.2: An estimate of the purity of the synthesised Ap<sub>4</sub>Ns.** This estimate was determined based on Ap<sub>4</sub>N peak absorbance as a proportion of the total absorbance at 260 nm.

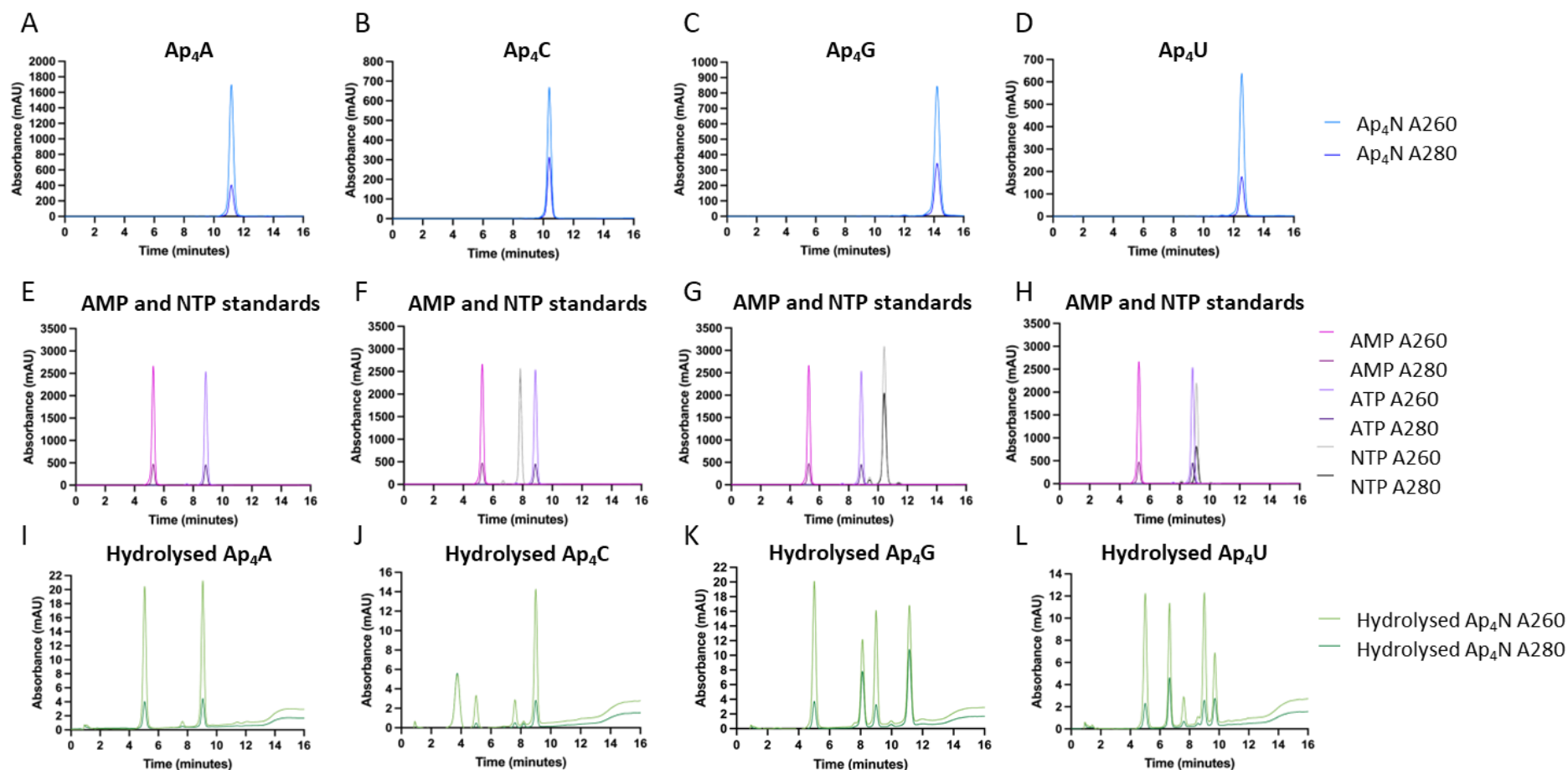
	Area under Ap <sub>4</sub> N peak	Total area under all peaks	Estimate of purity (%)
Ap <sub>4</sub> A	18664.582	18695.8588	99.8
Ap <sub>4</sub> C	6840.7842	6840.7842	100
Ap <sub>4</sub> G	11930.9414	12193.2151	97.8
Ap <sub>4</sub> U	7024.3799	7082.6893	99.2
Ap <sub>4</sub> dA	31412.5723	31458.7189	99.9
Ap <sub>4</sub> dC	18278.9023	18551.2489	98.5
Ap <sub>4</sub> dG	13901.7607	14064.9597	98.8
Ap <sub>4</sub> dT	13593.6367	13634.6085	99.7

### 3.5 Validation of chemically synthesised Ap<sub>4</sub>Ns

#### 3.5.1 Validation through analysis of Ap<sub>4</sub>A hydrolase products

To verify that the synthesised molecules were the intended Ap<sub>4</sub>Ns, each Ap<sub>4</sub>N was treated with recombinant human NUDT2 Ap<sub>4</sub>A hydrolase. Ap<sub>4</sub>A hydrolases have been found across multiple domains of life and cleave Ap<sub>4</sub>A asymmetrically into AMP and ATP (Swarbrick et al., 2005). An NMR study investigating human Ap<sub>4</sub>A hydrolase showed that an ATP binding cleft is present which contains histidine (His)-42, tyrosine (Tyr)-87 and phenylalanine (Phe)-133 aromatic side chains (Swarbrick et al., 2005). Two different conformations for ATP in this cleft have been identified, either where the adenine is stacked between Tyr-87 and Phe-133 aromatic rings, or where the adenosine lies across the top of the cleft (Swarbrick et al., 2005).

Ap<sub>4</sub>A hydrolases have been categorised into two subgroups - either the bacteria and plant subgroup or the animal and archaea subgroup. Studies with Lupin Ap<sub>4</sub>A hydrolase showed that hydrophilic attack occurs at the fourth phosphate away from the nucleoside which best fits the binding site (Guranowski et al., 1994a). Lupin Ap<sub>4</sub>A hydrolase has a preference for Ap<sub>4</sub>A but is also capable of breaking down Ap<sub>4</sub>C, Ap<sub>4</sub>G and Ap<sub>4</sub>U, each of which result in the production of AMP, ATP, NMP and NTP (where N=C, G or U, depending on the starting molecule) (Jakubowski and Guranowski, 1983). Similarly in the animal subgroup, the brine shrimp bis(5'-nucleosidyl) tetrphosphate pyrophosphohydrolase is capable of hydrolysing Np<sub>4</sub>Ns such as Ap<sub>4</sub>A, Gp<sub>4</sub>G and Ap<sub>4</sub>G (Prescott et al., 1989). This suggests that an adenosine moiety is not essential for hydrolysis to occur. Interestingly, when Ap<sub>4</sub>G is the substrate, AMP and GTP are produced 4.5-fold more than GMP and ATP. These data suggest that the human Ap<sub>4</sub>A hydrolase may be able to hydrolyse all Ap<sub>4</sub>Ns to AMP, NMP, ATP and NTP. It also suggests that, depending on the nucleosides in each molecule, there may be a preference in the ratio of the NMP and NTP products. Therefore, each of the Ap<sub>4</sub>Ns was hydrolysed to completion with human NUDT2 Ap<sub>4</sub>A hydrolase and the resulting products were analysed by HPLC to determine whether the expected AMP, NMP, ATP and NTP products were produced (Fig. 3.9).

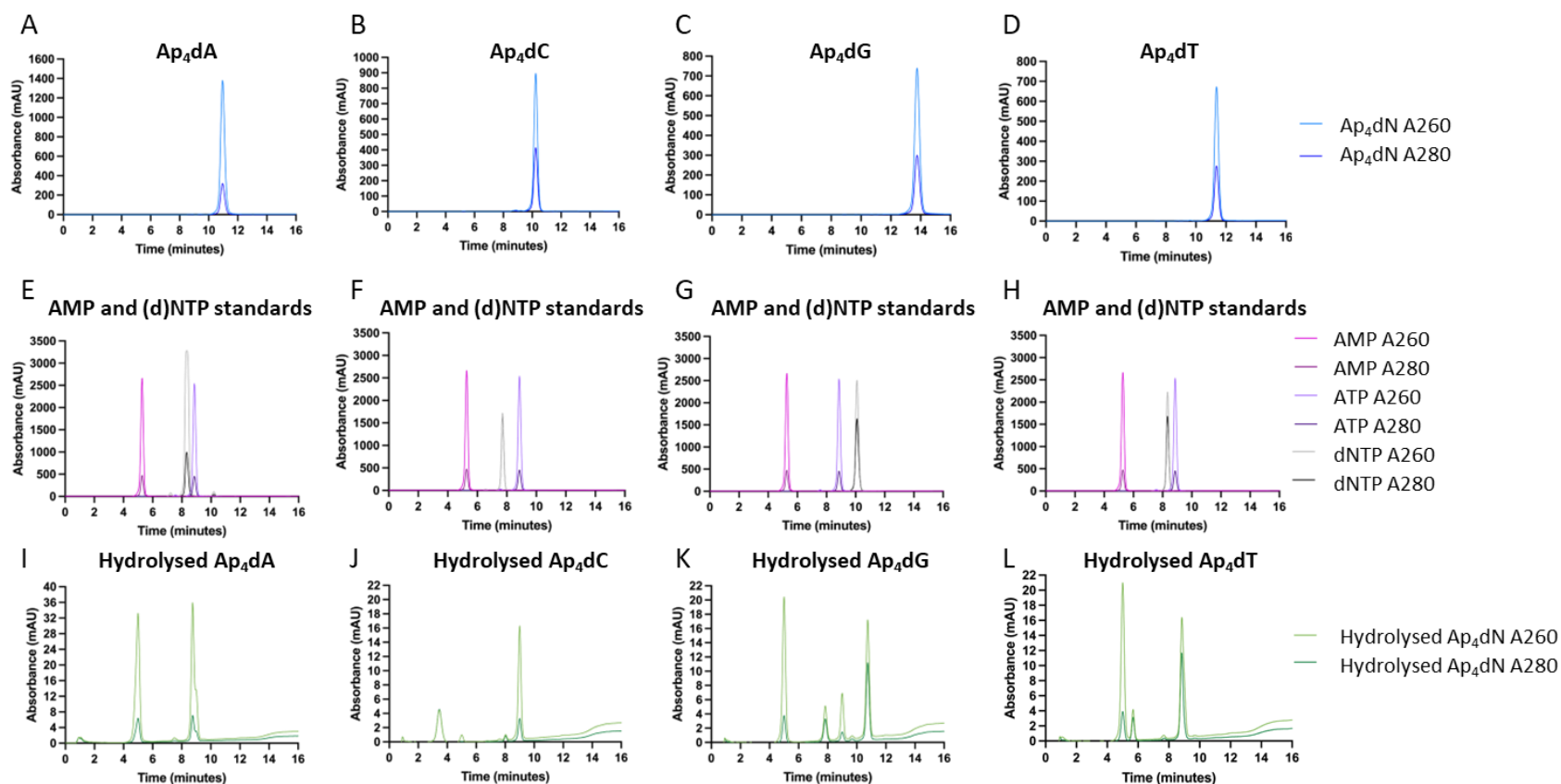


**Figure 3.9: Hydrolysis of Ap<sub>4</sub>Ns to NMP and NTP.** Each of the Ap<sub>4</sub>Ns was hydrolysed to completion, then each sample was analysed by HPLC. The retention times of the hydrolysed sample peaks were compared with the expected retention times for the corresponding Ap<sub>4</sub>N, NTP and AMP standards. **A–D:** Expected retention times for Ap<sub>4</sub>A (A), Ap<sub>4</sub>C (B), Ap<sub>4</sub>G (C) and Ap<sub>4</sub>U (D). **E–H:** Retention times of the standards corresponding to the expected AMP, ATP and NTP hydrolysis products for Ap<sub>4</sub>A (E), Ap<sub>4</sub>C (F), Ap<sub>4</sub>G (G) and Ap<sub>4</sub>U (H), overlaid on top of one another. Standards for the other NMPs are not available. **I–L:** Retention times for Ap<sub>4</sub>N products after complete hydrolysis of Ap<sub>4</sub>A (I), Ap<sub>4</sub>C (J), Ap<sub>4</sub>G (K) and Ap<sub>4</sub>U (L).



The HPLC trace for the hydrolysed Ap<sub>4</sub>A sample did not have a peak at 10.8 minutes (Fig. 3.9 I), demonstrating that it had been fully hydrolysed (Fig. 3.9 A). Instead, the hydrolysed Ap<sub>4</sub>A sample showed peaks at 5.0 and 8.7 minutes, corresponding to the 5.3- and 8.9-minute elution times previously determined for AMP and ATP (Fig. 3.9 E). This is consistent with the asymmetrical hydrolysis expected by the Ap<sub>4</sub>A hydrolase. Similarly, the hydrolysed Ap<sub>4</sub>C sample had lost the 10.1-minute peak expected if Ap<sub>4</sub>C were still present (Fig 3.9 B), but peaks corresponding to the elution times for ATP and AMP had formed (Fig. 3.9 F, J). A small peak formed at 7.4 minutes, but as this peak is also visible in the other hydrolysed Ap<sub>4</sub>N samples, it is unlikely to correspond to CTP (Fig. 3.9 J). A very small peak is visible at a retention time of just below 8 minutes and could therefore correspond to CTP, however its intensity was too low to be picked up automatically by the HPLC machine (Fig. 3.9 J). The hydrolysed Ap<sub>4</sub>G sample resulted in four peaks of similar intensity (Fig. 4.9 K), and no peak at 13.3 minutes, where Ap<sub>4</sub>G would be expected to elute (Fig. 3.9 C). The peaks at 5.0, 10.6 and 8.7 minutes correspond to the expected elution times for AMP, GTP and ATP, respectively (Fig. 3.9 G, K). The synthesised Ap<sub>4</sub>U was also successfully hydrolysed, as there was no peak at Ap<sub>4</sub>U's expected elution time of 11.9 minutes (Fig. 3.9 D). The peaks at 5.0, 9.3 and 8.7 minutes correspond to AMP, UTP, and ATP, respectively (Fig. 3.9 H, L). The additional peak in each trace (Fig. 3.9 I–L) likely corresponds to the respective NMP for each molecule; however, we did not have the standards available to confirm this.

The Ap<sub>4</sub>dNs were also treated with Ap<sub>4</sub>A hydrolase to confirm whether these molecules would break down as expected and therefore help validate their identity (Fig. 3.10).



**Figure 3.10: Hydrolysis of Ap<sub>4</sub>Ns to NMP and NTP.** Each of the Ap<sub>4</sub>dNs was hydrolysed to completion, then each sample was analysed by HPLC. The retention times of the hydrolysed sample peaks were compared with the expected retention times for the corresponding Ap<sub>4</sub>dN, (d)NTP and AMP standards. **A–D:** Expected retention times for Ap<sub>4</sub>dA (A), Ap<sub>4</sub>dC (B), Ap<sub>4</sub>dG (C) and Ap<sub>4</sub>dT (D). **E–H:** Retention times of the standards corresponding to the expected AMP, ATP and NTP hydrolysis products for Ap<sub>4</sub>A (E), Ap<sub>4</sub>C (F), Ap<sub>4</sub>G (G) and Ap<sub>4</sub>U (H), overlaid on top of one another. **I–L:** Retention times for Ap<sub>4</sub>N products after complete hydrolysis of Ap<sub>4</sub>dA (I), Ap<sub>4</sub>dC (J), Ap<sub>4</sub>dG (K) and Ap<sub>4</sub>dT (L).

As with all the Ap<sub>4</sub>Ns, each of the Ap<sub>4</sub>dNs was fully hydrolysed by the Ap<sub>4</sub>A hydrolase (Fig. 3.10 I–L). As with the hydrolysed Ap<sub>4</sub>A, two new peaks had formed after hydrolysis of Ap<sub>4</sub>dA. These peaks had retention times of 5.0 and 8.4 minutes (Fig. 3.10 I). While the 5.0-minute peak corresponds to AMP, the 8.4-minute peak could correspond to dATP or ATP, which have expected retention times of 8.3 and 8.9 minutes, respectively (Fig. 3.10 E). As this peak has a shoulder it is likely that it includes both ATP and dATP, but these molecules elute too closely together to be properly differentiated by this approach as it lacks the resolution to separate and resolve dATP from ATP.

The hydrolysed Ap<sub>4</sub>dC sample (Fig. 3.10 J) shows a similar pattern in terms of peak intensity to the hydrolysed Ap<sub>4</sub>C sample (Fig. 3.9 J). Small peaks are present at 5.0 (AMP) and 7.7 minutes (dCTP), showing that hydrolysis into AMP and dCTP has occurred (Fig. 3.10 F, J). However, the two larger peaks are at 3.7 minutes and 8.7 minutes, with the 8.7-minute peak corresponding to ATP (Fig. 3.10 J). This suggests that Ap<sub>4</sub>dC is primarily hydrolysed into ATP and dCMP. However, we did not have the necessary standard to confirm the dCMP retention time.

On the contrary, hydrolysis of Ap<sub>4</sub>dG seems to result primarily in production of AMP and dGTP, as the larger peaks occur at 5.0 and 10.2 minutes (Fig. 3.10 K). Again, another two peaks are present at 8.7 and 7.5 minutes, showing that ATP (8.7 minutes) is also produced (Fig. 3.10 G, K).

Finally, as dTTP and ATP have similar retention times, the peaks for each molecule cannot be differentiated, resulting in a single peak with a shoulder at 8.5 minutes (Fig. 3.10 L). The two peaks eluting at approximately 5 minutes likely correspond to AMP and dTMP, however of these two, only the AMP retention time has been confirmed (Fig. 3.10 H, L). Based on the ratio between the absorbances at 280:260, it is likely that the larger peak (at 5.0 mins) corresponds to AMP, as the ratio between 280:260nm absorbance is 1:5.6, and in both the ATP and AMP standards the 280:260 absorbance ratio is also 1:5.6. Together this suggests that the majority of the Ap<sub>4</sub>dT is hydrolysed into AMP and dTTP, with only a small amount being hydrolysed into dTMP and ATP. A summary of the Ap<sub>4</sub>N and Ap<sub>4</sub>dN hydrolysis products and their retention

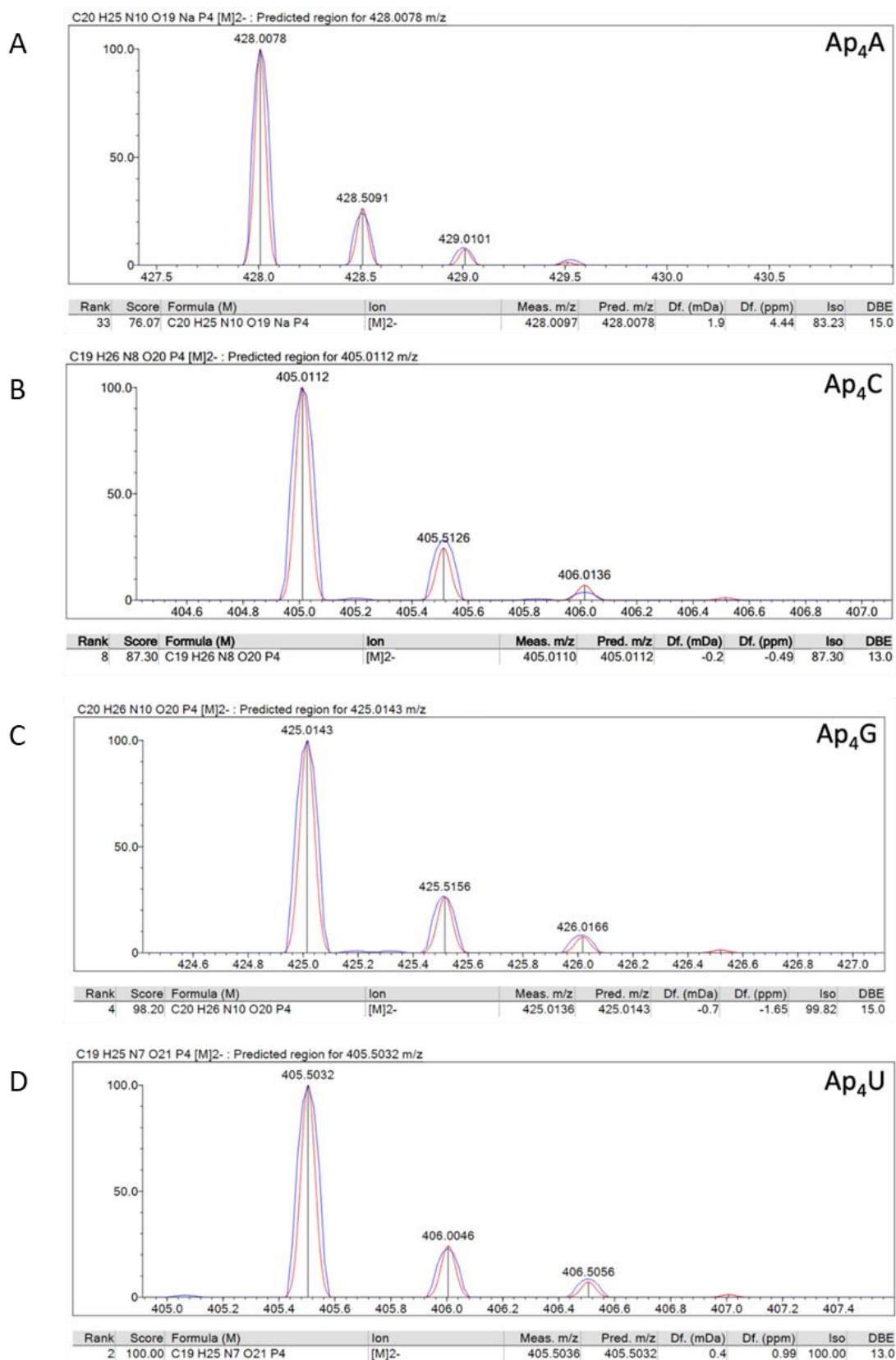
times, alongside the expected retention times for AMP, ATP and each NTP, is shown in Table 3.3. No standards were available for the NMP retention times; however, based on the expected asymmetrical hydrolase products, it is likely that the final peak in each trace corresponds to the relevant NMP.

**Table 3.3: A summary of the expected product retention times, and the corresponding peaks in the hydrolysis products.**

	<b>Expected products</b>	<b>Expected retention time (minutes)</b>	<b>Peak in hydrolysed sample?</b>	<b>Product peak retention time (minutes)</b>
<b>Ap<sub>4</sub>A</b>	AMP	5.3	✓	5.0
	ATP	8.9	✓	8.7
<b>Ap<sub>4</sub>C</b>	AMP	5.3	✓	5.0
	CTP	7.8	?	?
	ATP	8.9	✓	8.7
	CMP	-	?	4.0?
<b>Ap<sub>4</sub>G</b>	AMP	5.3	✓	5.0
	GTP	10.4	✓	10.6
	ATP	8.9	✓	8.7
	GMP	-		7.8?
<b>Ap<sub>4</sub>U</b>	AMP	5.3	✓	5.0
	UTP	9.1	✓	9.3
	ATP	8.9	✓	8.7
	UMP	-	?	6.4?
<b>Ap<sub>4</sub>dA</b>	AMP	5.3	✓	5.0
	dATP	8.3	✓	8.4
	ATP	8.9	✓	8.4
	dAMP	-	?	5.0?
<b>Ap<sub>4</sub>dC</b>	AMP	5.3	✓	5.0
	dCTP	7.7	✓	7.8
	ATP	8.9	✓	8.7
	dCMP	-	?	3.7?
<b>Ap<sub>4</sub>dG</b>	AMP	5.3	✓	5.0
	dGTP	10.1	✓	10.2
	ATP	8.9	✓	8.7
	dGMP	-	?	7.5
<b>Ap<sub>4</sub>dT</b>	AMP	5.3	✓	5.0
	dTTP	8.3	✓	8.5
	ATP	8.9	✓	8.5
	dTMP	-	?	5.6?

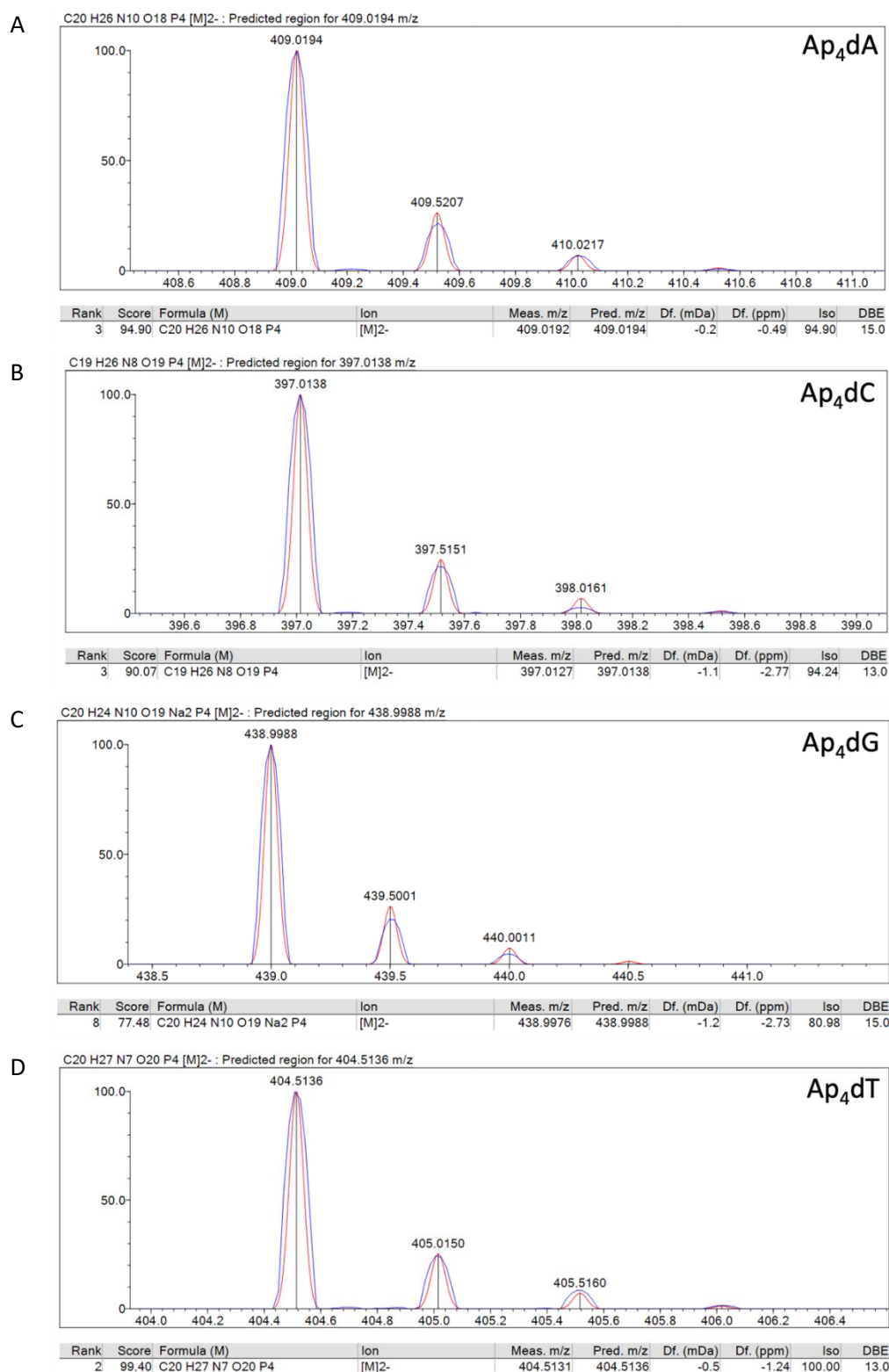
### **3.5.2 Validation of Ap<sub>4</sub>Ns by HRMS**

Although the hydrolysis products highly suggested that the synthesised molecules were the expected Ap<sub>4</sub>Ns, analysis of the expected mass for each synthesised Ap<sub>4</sub>N was performed by HRMS. All Ap<sub>4</sub>Ns and Ap<sub>4</sub>dNs tested were verified as having the correct molecular mass and isotopic pattern for each formula, verified to <5ppm (Fig. 3.11 and Fig. 3.12). The full report for each Ap<sub>4</sub>N can be found within the supplementary data (Supplementary Fig. 3.1).



**Figure 3.11: Verification of Ap<sub>4</sub>Ns by mass spectrometry.** Samples from the synthesised Ap<sub>4</sub>Ns were diluted in water and injected for analysis by HRMS. **A:** Ap<sub>4</sub>A, **B:** Ap<sub>4</sub>C, **C:** Ap<sub>4</sub>G, **D:** Ap<sub>4</sub>U. The predicted pattern for each molecule is shown in red, whereas the sample run is shown in blue.

At least two HRMS assignments were made for each Ap<sub>4</sub>N/Ap<sub>4</sub>dN, with the [M]<sup>2-</sup> ion consistently being one of the best ions for detection of all Ap<sub>4</sub>Ns. In this form, the Ap<sub>4</sub>A sample showed the appropriate isotopic pattern to 4.44 ppm when sodium was also bound to the molecule (Fig. 3.11 A, Supplementary Fig. 3.11 A). A further two assignments were made for this sample that were less than 5 ppm, in the [M]<sup>-</sup> (3.71 ppm) and [M-H]<sup>-</sup> (-2.92 ppm) forms of the ion. The Ap<sub>4</sub>C sample was assigned in the [M]<sup>2-</sup> form to -0.49 ppm and follows the expected isotopic pattern (Fig. 3.11 B, Supplementary Fig. 3.1 C). Several other assignments to within 5 ppm were also made for the Ap<sub>4</sub>C sample. Ap<sub>4</sub>G in the [M]<sup>2-</sup> form was also assigned to -1.65 ppm and had the appropriate pattern for its isotopes (Fig. 3.11 C, Supplementary Fig. 3.1 E). Two other assignments for the Ap<sub>4</sub>G sample were also made that were below 5 ppm. Similarly, the Ap<sub>4</sub>U sample was assigned in its [M]<sup>2-</sup> form to 0.99 ppm, with four other assignments to below 5 ppm being made in other ionic forms (Fig. 3.11 D, Supplementary Fig. 3.1 G). The full reports for two different assignments of each of the nucleotides synthesised are shown in the supplementary data (Supplementary Fig. 3.1).



**Figure 3.12: Verification of Ap<sub>4</sub>dNs by mass spectrometry.** Samples from the synthesised Ap<sub>4</sub>dNs were diluted in water and injected for analysis by HRMS. **A:** Ap<sub>4</sub>dA, **B:** Ap<sub>4</sub>dC, **C:** Ap<sub>4</sub>dG, **D:** Ap<sub>4</sub>dT. The predicted pattern for each molecule is shown in red, whereas the sample run is shown in blue.



As with the Ap<sub>4</sub>Ns, the [M]<sup>2-</sup> ion proved best for validating the Ap<sub>4</sub>dN samples. The Ap<sub>4</sub>dA sample followed the correct pattern with its isotopes and was assigned to within 0.49 ppm in the [M]<sup>2-</sup> ion form (Fig. 3.12 A). Likewise, Ap<sub>4</sub>C was found to within 2.77 ppm in the [M]<sup>2-</sup> form and with peaks of the expected proportions for the different isotopic masses (Fig. 3.12 B). Ap<sub>4</sub>dG was successfully assigned in the [M]<sup>2-</sup> ionic form to within 2.73 ppm and followed the appropriate isotopic pattern when sodium ions were also bound to the molecule (Fig. 3.12 C). Finally, the Ap<sub>4</sub>dT sample was also verified to within 1.24 ppm for the [M]<sup>2-</sup> ion, also showing the expected proportions of its various isotopes (Fig. 3.12 D). The full reports for two different assignments for each Ap<sub>4</sub>dN can be found in Supplementary Figure 3.1. Therefore, all Ap<sub>4</sub>Ns and Ap<sub>4</sub>dNs successfully verified by HRMS.

### **3.6 Determining concentration of synthesised Ap<sub>4</sub>Ns**

As these molecules had now been verified to be the expected molecules, next it was necessary to determine their concentration to prepare them for use in future experiments. As there is no published extinction coefficient for the Ap<sub>4</sub>Ns synthesised here, it first needed to be calculated. An approximate extinction coefficient was calculated based on the value for the sum of the extinction coefficients for each of the two nucleotides making up the molecule. However, these needed to be modified further to account for the effect of base stacking, which reduces the absorbance of the Ap<sub>4</sub>N relative to the absorbance of the bases. Therefore, the absorbance of each molecule at 260 nm was measured before and after hydrolysis. The change in absorbance was used to modify the previously calculated extinction coefficient. The modified extinction coefficient ( $\epsilon$ ) for each Ap<sub>4</sub>N was used for all further calculations of Ap<sub>4</sub>N concentration and is shown in Table 3.4.

**Table 3.4: Determining the extinction coefficient for each Ap<sub>4</sub>N.** The 260 nm absorbance was measured for each Ap<sub>4</sub>N and compared with the absorbance of the Ap<sub>4</sub>N sample after hydrolysis. The difference in absorbance was used to adjust the extinction coefficient calculated based on the extinction coefficients for the individual NMPs, to account for the effect of base stacking in the unhydrolyzed molecule. Abs., absorbance; diff., difference;  $\epsilon$ , extinction coefficient.

	Abs. diff.	1st NMP $\epsilon$	2nd NMP $\epsilon$	Previously calculated $\epsilon$	Modified $\epsilon$
<b>Ap<sub>4</sub>A</b>	0.1825	15,020	15,020	30,040	27117
<b>Ap<sub>4</sub>G</b>	0.1555	15,020	12,080	27,100	25492
<b>Ap<sub>4</sub>dA</b>	0.4130	15,020	15,060	30,080	27116
<b>Ap<sub>4</sub>dG</b>	0.1025	15,020	12,180	27,200	26173
<b>Ap<sub>4</sub>C</b>	-0.0025	15,020	7,070	22,090	22472
<b>Ap<sub>4</sub>dC</b>	0.0420	15,020	7,100	22,120	21910
<b>Ap<sub>4</sub>U</b>	-0.3085	15,020	9,660	24,680	26759
<b>Ap<sub>4</sub>dT</b>	-0.0990	15,020	8,560	23,580	24984

The extinction coefficients calculated based on the NMPs were altered to different degrees depending on the properties of the purine and pyrimidine rings in the Ap<sub>4</sub>N molecules. Ap<sub>4</sub>A, Ap<sub>4</sub>G, Ap<sub>4</sub>dA and Ap<sub>4</sub>dG were affected the most as these molecules have larger purine rings that enhanced the base stacking effect. This approach has enabled a more accurate method for calculation of the Ap<sub>4</sub>N concentration using the extinction coefficient determined here, and this was used to determine the concentration of each batch of Ap<sub>4</sub>N produced for further study. Each Ap<sub>4</sub>N had now been chemically synthesised, verified, and had its concentration measured ready for use in future experiments.

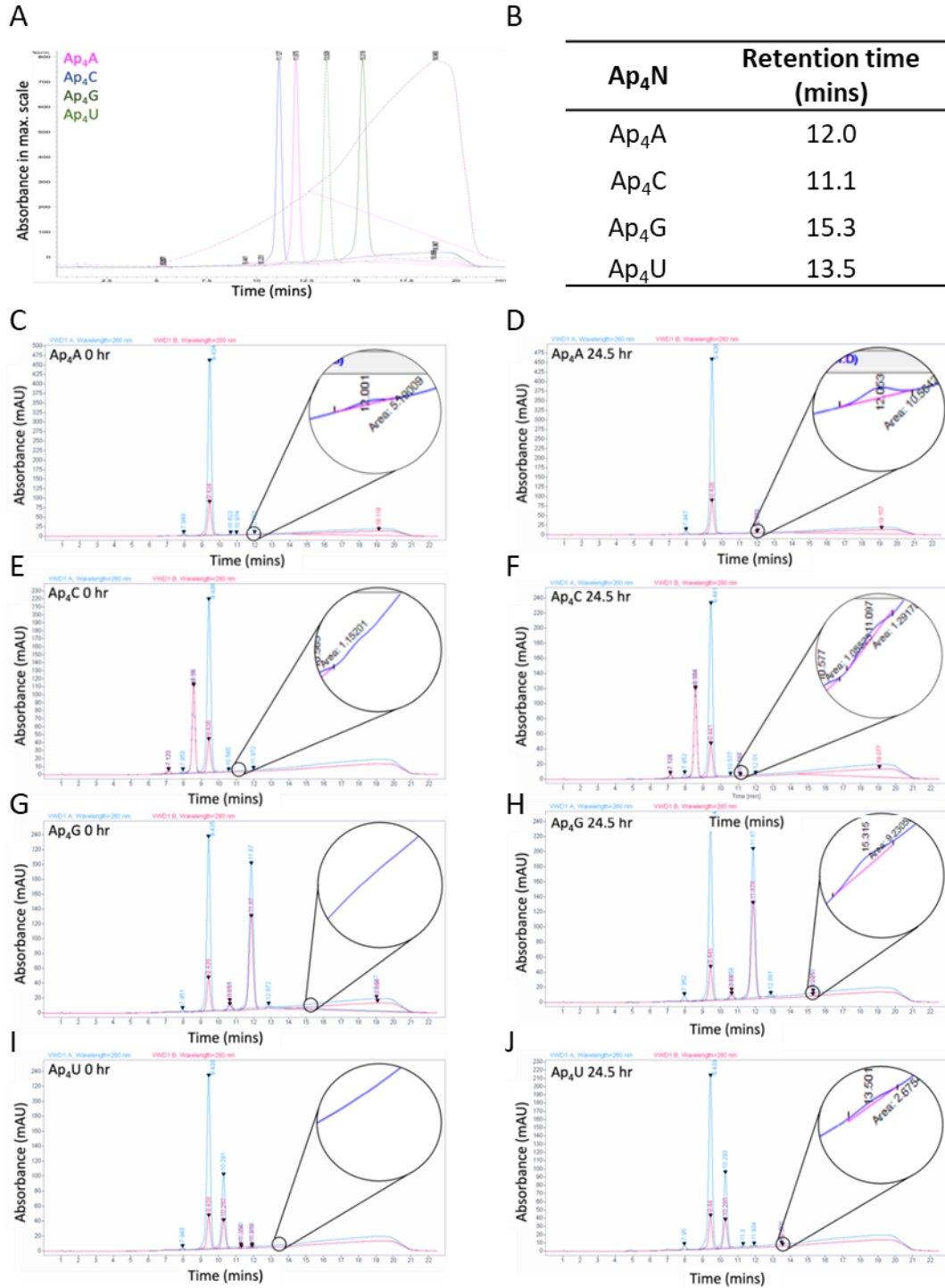
### 3.7 Biological synthesis of Ap<sub>4</sub>A and Ap<sub>4</sub>Ns

Many biological pathways can produce Ap<sub>4</sub>A. These include synthesis by the aminoacyl tRNA synthetases, T4 RNA and DNA ligase, and Acyl CoA (Atencia et al., 1999; Brevet et al., 1989b; Fontes et al., 1998; Goerlich et al., 1982; Guo et al., 2009; Guranowski et al.,

1994b; Madrid et al., 1998; Zamecnik et al., 1966). Recently, it has been shown that during the activation of Ub, Ap<sub>4</sub>A is produced as a by-product (Götz et al., 2019). In the model suggested for this mechanism, the Ub activating enzyme reacts with  $\alpha$ -phosphate on one ATP molecule to form an adenylated intermediate. The  $\gamma$ -phosphate of a second ATP group then reacts with the intermediate to form Ap<sub>4</sub>A (Fig. 1.3).

In some of the signalling mechanisms that Ap<sub>4</sub>A is involved in, some of the Ap<sub>4</sub>Ns are also involved, for example in the RNA capping mechanism in *E. coli* (Luciano et al., 2019). This demonstrates that Ap<sub>4</sub>Ns exist and have an intracellular function in bacterial cells. Some Ap<sub>4</sub>Ns have an extracellular function in mammalian cells. An example of this is Ap<sub>4</sub>U, which has an extracellular role in the cardiovascular system and can be synthesised by VEGFR2 and released from endothelial cells (Jankowski et al., 2013a; Jankowski et al., 2009). As it has now been demonstrated that Ap<sub>4</sub>A can be synthesised intracellularly by the Ub activating enzymes (Götz et al., 2019), we wanted to determine whether the other Ap<sub>4</sub>Ns could also be synthesised in this way. We hypothesised that Ap<sub>4</sub>Ns may be produced if a different NTP were available in the system, because it is the  $\gamma$ -phosphate, and not the adenosine, of the second ATP which is involved in the reaction. Therefore, the  $\gamma$ -phosphate of a different NTP molecule may be able to react with the adenylated intermediate.

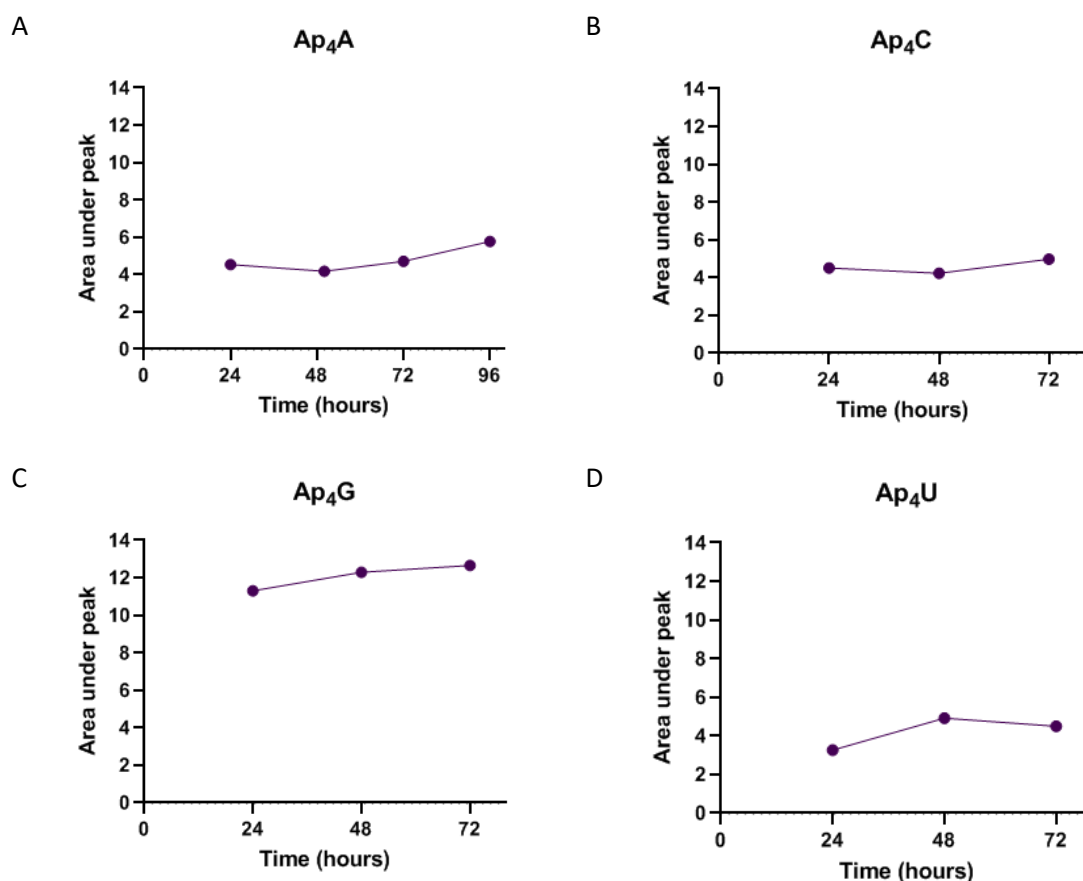
This could indicate whether these other Ap<sub>4</sub>Ns might exist in mammalian cells and therefore whether they might play a role in intracellular signalling mechanisms in these cells. Therefore the protocol used by Götz and colleagues in 2019 was adapted to determine whether the other Ap<sub>4</sub>Ns could also be synthesised as a result of Ub activating enzyme activity, by adding the other NTPs into the reaction mix. Samples were analysed by HPLC at 0-hour and 24.5-hour timepoints to determine whether a peak formed at the appropriate retention time (Fig. 3.13). As a result of pressure build up in the column, the run method had to be adapted to reduce this pressure to prevent damage to the column. Therefore the Ap<sub>4</sub>N standards were re-run to determine the new elution times. The new elution times are listed in Figure 3.13 B.



**Figure 3.13: *In vitro* synthesis of Ap<sub>4</sub>Ns by the ubiquitin-activating enzyme UBE1.** **A, B:** HPLC retention times for Ap<sub>4</sub>A, Ap<sub>4</sub>C, Ap<sub>4</sub>G and Ap<sub>4</sub>U were determined. **C–J:** Ubiquitin, UBE1 and ATP (C, D) or ATP plus either CTP (E, F), GTP (G, H) or UTP (I, J) were incubated together. The samples for each reaction were taken at 0 or 24.5 hours then analysed by HPLC. Ap<sub>4</sub>N peaks were selected manually as the concentrations were too low to be detected automatically.

Immediately after initiating the reaction a small peak eluted at 12.0 minutes, the expected retention time for Ap<sub>4</sub>A (Fig. 3.13 C). As the samples were at low concentrations and at the detection limit of the HPLC, it is likely that this peak could be background as it is unlikely that the Ap<sub>4</sub>A synthesis reaction would have had time to occur before the sample was taken. However, the size of this peak doubled after 24 hours, suggesting that Ap<sub>4</sub>A had been synthesised during this time (Fig. 3.13 D). No peak was identified at 0 hours in the Ap<sub>4</sub>C synthesis, as although a small bump is present in the trace, it was not large enough to be detected by the HPLC even when picking the peaks manually (Fig. 3.13 E). Nevertheless, a small peak is detectable at the appropriate retention time in the 24.5-hour sample, with an area of 1.29 units (Fig. 3.13 F). The Ap<sub>4</sub>G reaction mix showed no peak at the 0-hour timepoint (Fig. 3.13 G). However, after 24.5 hours a peak with an area of 9.2 units appeared at 15.3 minutes (Fig. 3.13 H). This corresponds to the retention time of the Ap<sub>4</sub>G standard produced using the HRMS-validated chemical synthesis method (Fig. 3.13 A, B). Finally, the Ap<sub>4</sub>U synthesis reaction mix showed no peak at the appropriate retention time immediately after combining all the reactants (Fig. 3.13 I). However, as with the other Ap<sub>4</sub>Ns, a small peak formed at the expected retention time for Ap<sub>4</sub>U after 24.5 hours (Fig. 3.13 J).

Peaks were formed at the expected retention times for all Ap<sub>4</sub>N UBE1-mediated synthesis reactions; however, the yield from these reactions was very small. To determine whether increased reaction time would induce an increase in Ap<sub>4</sub>N level, a time-course experiment was set up. The reactants for each Ap<sub>4</sub>N synthesis were combined, and samples were taken approximately every 24 hours (Fig. 3.14).



**Figure 3.14: Effect of increased incubation time on Ap<sub>4</sub>N synthesis by UBE1.** Ubiquitin, UBE1 and ATP (A) or ATP and either CTP (B), GTP (C) or UTP (D) were combined and incubated at 37 °C. Samples were taken every 24 hours and frozen at -80 °C. After all the timepoints had been collected, the samples were analysed by HPLC and the area under the peak was measured at the appropriate elution times for each Ap<sub>4</sub>N.

The area under the peak corresponding to Ap<sub>4</sub>A ranged from 4.2 to 5.8 units after between 24 and 96 hours at 37 °C; however, this was not a linear upwards trend (Fig. 3.14 A). In fact, between 24- and 50-hours after the start of the incubation the area under the peak decreased slightly. This suggests that these small changes may just be due to the variability in the signal to noise ratio as the products are present at low concentrations, or variability induced by the necessity for manual peak selection at these concentrations. Similarly, the area measured under the Ap<sub>4</sub>C peak was within one unit for all timepoints, with the area under the peak also dipping at 48-hours (Fig. 3.14 B). The area under the Ap<sub>4</sub>G peak increased over time, but still only

increased by a maximum of 1.3 units from the 24-hour to the 72-hour timepoints (Fig. 3.14 C). Like the other Ap<sub>4</sub>Ns, the area of the Ap<sub>4</sub>U peak did not change much, with a difference of 1.7 units between the highest and lowest peak areas (Fig. 3.14 D). The area under the Ap<sub>4</sub>U peak was largest after 48-hours of incubation, but then decreased slightly at the 72-hour timepoint.

### **3.8 Chapter Discussion**

#### **3.8.1 Ap<sub>4</sub>A and Ap<sub>4</sub>Ns can be analysed and purified from NTPs and AMP by HPLC**

In this chapter, we have demonstrated that Ap<sub>4</sub>A can be distinguished from AMP and ATP through separation by HPLC using a linear ammonium bicarbonate salt gradient. This separation method can also be applied to the other Ap<sub>4</sub>Ns and their separation from their respective NTPs and AMP.

Development of HPLC began in the 1960's as a method for the separation of complex mixtures of nucleotides (Horvath et al., 1967). As with other chromatography, the HPLC system contains a mobile and stationary phase. In HPLC, a high pressure is used to push the sample in the mobile phase through to the column, where it interacts with the stationary phase. The extent to which the sample molecules interact with the stationary phase will determine the retention time of each molecule. As different molecules have different chemical and physical properties, the extent to which they will interact with a given column will differ, allowing separation of different molecules from a complex mixture. Since its introduction, the sensitivity and variety of molecules detectable by HPLC has been increased. Different types of HPLC are now available for different applications, for example normal-phase-, reversed-phase-, ion-exchange- and bio-affinity- HPLC (Bansal, 2010). Parameters such as pore-size, internal diameter, particle size of beads inside the column, and pump pressure can also be adapted for analysis of different types of molecules (Bansal, 2010). Furthermore, several different types of detector are now available depending on the function required, including but not limited to UV-Visible, fluorescence, electrochemical and light-scattering detectors (Swartz, 2010).

There are several method options for separation of nucleotides by HPLC, including ion-paired reverse-phase (IP-RP), ion-exchange, and hydrophilic interaction (HILIC) chromatography. Ion-paired reverse phase systems are a popular choice and have been used on multiple occasions for separation of molecules such as AMP, ADP and ATP (Manfredi et al., 2002; Zur Nedden et al., 2009). Alternatively, ion-exchange columns have been used for separation of nucleotides. A DNAPac PA200 ion-exchange column with an increasing salt gradient was used recently for the separation of NMPs, NDPs and NTPs, where N=A, C, G or U (Strezsak et al., 2022). Hydrophilic interaction (HILIC) chromatography is another option for analysis of nucleotides, and the FructoShell-N HILIC column was successfully able to separate most of the nucleotides tested and had a short run time; however, other HILIC columns displayed problems with peak asymmetry (Fabino Carr et al., 2019).

Since the dinucleotide tetraphosphates differ from AMP and the NTPs in terms of the number of phosphate groups they possess, they also differ considerably in their charge. The Ap<sub>4</sub>Ns contain four phosphates and therefore carry a stronger negative charge relative to corresponding NTPs and AMP, which contain three or just one phosphate group(s), respectively. Therefore, an ion-exchange HPLC technique was used to separate out these molecules. A ResourceQ column was chosen for separation and followed a similar method to one previously used for the separation of nucleotides on a MonoQ column (Marriott et al., 2015). Furthermore, this method utilised an ammonium bicarbonate gradient, which was well-suited to purification as the highly volatile nature of its decomposition products meant that they could be fully removed after purification by lyophilisation.

### **3.8.2 EDC as a coupling agent in Ap<sub>4</sub>A synthesis**

After demonstrating that AMP, ATP and Ap<sub>4</sub>A could be successfully separated by anion-exchange HPLC, we optimised a synthesis protocol for Ap<sub>4</sub>A and other Ap<sub>4</sub>Ns. A number of possible reaction mechanisms have been developed for the synthesis of Ap<sub>4</sub>A, each with their own benefits and limitations (Appy et al., 2019a). For our synthesis of Ap<sub>4</sub>A we decided on a carbodiimide coupling reaction which results in the formation of a phosphodiester bond between the phosphate groups of the two



starting molecules: AMP and ATP. The reaction mechanism for this chemical Ap<sub>4</sub>A synthesis is shown in Figure 3.4. Ap<sub>4</sub>N synthesis was performed using EDC as the catalyst due to its simplicity and the fact that it could easily be applied to the production of the other Ap<sub>4</sub>Ns by substituting the ATP starting reactant with the other NTPs. Furthermore, the reagents needed were limited to ATP, AMP, MgCl<sub>2</sub> and EDC, all of which are commercially available.

### **3.8.3 Analysis of Ap<sub>4</sub>N hydrolysis products as a method of validating Ap<sub>4</sub>N identity**

Hydrolysis of the Ap<sub>4</sub>Ns was used as a technique for preliminarily validating the identity of the chemically synthesised Ap<sub>4</sub>Ns. The asymmetrical Nudix Ap<sub>4</sub>A hydrolase is present across different organisms and cleaves Ap<sub>4</sub>A into AMP and ATP (Ferguson et al., 2020; McLennan, 2006). It is unlikely that another molecule would have the same retention time and asymmetrical hydrolysis pattern as expected for each of the Ap<sub>4</sub>Ns, making it a useful validation method. In addition to the elution times, the absorbance of the hydrolysis products could be used as a secondary measure to determine whether they are the expected products. Of the nucleoside triphosphates, CTP absorbance peaks at 270 nm, whereas the maximal absorbance for the other nucleotides occurs closer to 260 nm (Zur Nedden et al., 2009). This supports the data in Figure 3.8 and 3.9, as the peaks produced that are expected to correspond to CMP, CTP, dCMP or dCTP after hydrolysis all have a similar absorbance at 260 nm as at 280 nm. This is not seen for the other nucleotides, which show a much larger peak at 260 nm compared with the 280 nm absorbance.

### **3.8.4 HRMS as a method of validating Ap<sub>4</sub>N identity**

The chemical synthesis of the correct Ap<sub>4</sub>Ns was successfully validated by LCMS-IT-TOF High Resolution Mass Spectrometry. LCMS-IT-TOF has been used in different scenarios for the identification and analysis of various molecules (Can, 2018; Hyakkoku et al., 2010; Kim et al., 2015; Ouyang et al., 2015). The high sensitivity and high resolution of the LCMS-IT-TOF made it ideal for the accurate identification of the synthesised Ap<sub>4</sub>Ns, since we were able to calculate the expected molecular mass for each sample.

### 3.8.5 Ub and Ub-like activating enzymes as a source of Ap<sub>4</sub>Ns

Finally, we have demonstrated that the ubiquitin activating enzyme UBE1 is capable of synthesising molecules with elution times corresponding to the retention times for HRMS-verified Ap<sub>4</sub>Ns (Fig. 3.11). This experiment was adapted from the experiment used by Götz and colleagues in 2019 which demonstrated that the ubiquitin-activating enzyme UBA1 and the ubiquitin-like activating enzymes NEDD8 and SUMO are capable of synthesising Ap<sub>4</sub>A as a by-product during ubiquitin activation. The model for the mechanism described by Götz and colleagues is depicted in Figure 1.3. This synthesis mechanism demonstrated that there is the potential for Ap<sub>4</sub>Ns to be produced intracellularly at times when cells are under stress. This correlates with the data that show an increase in Ap<sub>4</sub>N levels in response to stress (Ji et al., 2019; Luciano et al., 2019), as there is also an increase in Ub activation under stressful conditions (Shang et al., 1997).

A drawback of this system for Ap<sub>4</sub>N synthesis is that the Ap<sub>4</sub>N yield was very low, and the peaks at the appropriate retention time are only just distinguishable by HPLC (Fig. 3.12). This suggests that this synthesis mechanism may not be enough to increase the Ap<sub>4</sub>N level to concentrations high enough to see an effect. However, this experiment was performed *in vitro*, so other factors could also influence the yield. For example, perhaps the limited constituents of this reaction meant that the Ub can only be used once in this *in vitro* reaction mechanism, which would limit Ap<sub>4</sub>N synthesis. Certainly, increasing the time course for the experiment did not show much increase in the production of Ap<sub>4</sub>Ns detectable by HPLC (Fig. 3.14). Despite this, this *in vitro* method supports the conclusion made by Götz and colleagues that Ap<sub>4</sub>A can be synthesised by the ubiquitin-activating enzymes. Here the synthesis of other potential Ap<sub>4</sub>N molecules has also been demonstrated with similar efficiency to the synthesis of Ap<sub>4</sub>A, suggesting that UBE1 can also generate other dinucleoside tetraphosphates. This demonstrates a proof-of-principle that Ap<sub>4</sub>Ns can be produced intracellularly, and therefore suggests that these nucleotides may be present in mammalian cells.

### **3.9 Conclusion**

In all, we have demonstrated that Ap<sub>4</sub>A and Ap<sub>4</sub>Ns can be synthesised chemically through a carbodiimide coupling reaction. These molecules were successfully purified from their reaction mixture by HPLC and validated, first by analysis of their hydrolysis products and then by HRMS. Finally, these validated Ap<sub>4</sub>Ns could be used as standards to identify the peaks forming because of UBE1-mediated synthesis, demonstrating the capacity for UBE1 to synthesise Ap<sub>4</sub>Ns from ATP and the appropriate NTPs.

# **Chapter 4: Identification of intracellular $\text{Ap}_4\text{Ns}$ in a mammalian cell line**

## **4.1 Introduction**

### **4.1.1 Ap<sub>4</sub>Ns exist intracellularly in different domains of life and are synthesised in response to cell stress**

Dinucleoside tetraphosphates are synthesised in eukaryotes and prokaryotes in response to stress and it is now becoming evident that their activity can be mediated by multiple mechanisms. Ap<sub>4</sub>Ns have been identified in *E. coli*, where they are synthesised by aminoacyl tRNA synthetases in both heat-shocked and exponentially growing cells (Brevet et al., 1989b). Similarly, accumulation of diadenosine polyphosphates including Ap<sub>4</sub>A and Ap<sub>4</sub>G has been demonstrated in *S. typhimurium* as a result of oxidative stress (Bochner et al., 1984). The increase in the levels of these Ap<sub>4</sub>Ns under stressful conditions has been linked to mechanisms such as formation of mRNA caps for protecting RNA transcripts in *E. coli* (Luciano et al., 2019). As demonstrated in bacteria, Ap<sub>4</sub>Ns have also been identified in the yeast *S. cerevisiae*, where the cellular pool increased from 0.8  $\mu$ M to 4  $\mu$ M or higher after exposure to heat-shock or cadmium-treatment (Baltzinger et al., 1986). The presence of intracellular Ap<sub>4</sub>Ns is not limited to bacteria and yeast cells, as Ap<sub>4</sub>Ns have also been identified in *Drosophila* cells. Of the Ap<sub>4</sub>Ns, Ap<sub>4</sub>C and Ap<sub>4</sub>U signal in *Drosophila* cells was approximately 5% of the total luminescence signal, with the majority of the signal corresponding to Ap<sub>4</sub>A and Ap<sub>4</sub>G (Brevet et al., 1985b).

### **4.1.2 Ap<sub>4</sub>Ns in mammalian cells**

Ap<sub>4</sub>A was identified in mammalian cells as early as the 1970s, when it was measured in human, chick, mouse and hamster cells lines (Rapaport and Zamecnik, 1976). It has since been measured in other mammalian cells and increases in response to MMC-induced stress in 3T3, HeLa and AA8 cell lines (Marriott et al., 2015). Ap<sub>4</sub>A has also been found in HEK293T cells, a human embryonic kidney cell line expressing the SV40 large T antigen, where it increases significantly in cells stressed through treatment with MMC or hydrogen peroxide (Krüger et al., 2021). Interestingly, very little has been reported about the presence of Ap<sub>4</sub>Ns other than Ap<sub>4</sub>A intracellularly in mammalian cells. Ap<sub>4</sub>U has a documented extracellular role in the cardiovascular system, where it is released by endothelial cells after VEGFR2-mediated synthesis

(Jankowski et al., 2013b; Matsumoto et al., 2011). Ap<sub>4</sub>dT, Ap<sub>5</sub>dT and Ap<sub>6</sub>dT were all able to inhibit cytosolic thymidine kinase in patients with acute myelocytic leukaemia, with Ap<sub>5</sub>dT showing the most potent inhibition of thymidine kinase (Bone et al., 1986a). This demonstrates that there is the potential for dinucleotide tetraphosphates to have a regulatory role in mammalian cells.

#### **4.1.3 Ub-activating enzyme and the synthesis of Ap<sub>4</sub>Ns**

As discussed previously, several biological mechanisms are capable of synthesising Ap<sub>4</sub>A intracellularly in different contexts (see section 1.1.1 and 3.1.1). The synthesis of Ap<sub>4</sub>A by Ub- and Ub-like activating enzymes is particularly interesting, as it presents a relevant biological mechanism through which Ap<sub>4</sub>A could be synthesised in stressed mammalian cells (Götz et al., 2019). This biological synthesis method is supported by profiling data which highlighted UBA1 as an Ap<sub>4</sub>A-interacting protein (Krüger et al., 2021). This data is further supported by experiments described in chapter three, which utilised an adapted version of the Götz method to demonstrate that the Ubiquitin activating enzyme-mediated synthesis has the potential to produce other Ap<sub>4</sub>Ns in addition to Ap<sub>4</sub>A (Figure 3.13). The identification of a biologically relevant mechanism that is capable of synthesising Ap<sub>4</sub>Ns supports the hypothesis that these Ap<sub>4</sub>Ns can be produced and therefore these molecules are likely to be present in mammalian cells. This leads onto the aims of this chapter, which centre around the identification of Ap<sub>4</sub>Ns in mammalian cells.

#### **4.1.4 Chapter Aims**

In some cases, both ends of the Ap<sub>4</sub>A molecule are required for functionality and it can therefore be described as homobifunctional. An example of this is apparent in HINT signalling, where both ends of the Ap<sub>4</sub>A molecule are required for polymerisation of HINT dimers (Yu et al., 2019). However, in cases such as the protective capping of RNA in *E. coli*, different Ap<sub>4</sub>Ns are capable of forming Np<sub>4</sub> caps, demonstrating that only one adenosine is required for activity (Luciano and Belasco, 2020; Luciano et al., 2019). This occurs through incorporation of Ap<sub>4</sub>N during transcription initiation and is at least partially affected by the untranscribed region around the promoter (Luciano and Belasco, 2020).

Ap<sub>4</sub>Ns have a similar structure to Ap<sub>4</sub>A, with one of the adenosine groups being replaced by another of the nucleoside bases. These molecules can therefore be used to characterise the functional requirements of a signalling system because the change prevents the molecule from acting homobifunctionally, and so analysis of the functionality of these molecules could give insight into how Ap<sub>4</sub>A functions in signalling mechanisms such as the initiation of DNA replication. Furthermore, if a homobifunctional mechanism is not required, it is possible that these other Ap<sub>4</sub>Ns could also have a role in the initiation of DNA replication. Previous studies have generally focused on Ap<sub>4</sub>A as an alarmone because this molecule is present at a higher concentration than the other Ap<sub>4</sub>Ns. However, bioluminescence-based Ap<sub>4</sub>A assays used to measure Ap<sub>4</sub>A level, or altering Ap<sub>4</sub>A hydrolase activity to change Ap<sub>4</sub>A concentration, will also have a similar effect on other Ap<sub>4</sub>Ns. This suggests that in these experiments the influence of the other Ap<sub>4</sub>Ns, if present intracellularly, may have been overlooked.

In Chapter 3, biological synthesis of Ap<sub>4</sub>A and Ap<sub>4</sub>Ns was demonstrated using UBE1 ubiquitin activating enzyme. To take this further, the aim was to develop a technique to investigate whether these other Ap<sub>4</sub>Ns could also be detected intracellularly in mammalian cells. The first aim of this chapter was therefore to develop a method for harvesting the Ap<sub>4</sub>Ns from a population of mammalian cells. Once isolated, we aimed to develop and optimise a technique for the identification of Ap<sub>4</sub>A and other Ap<sub>4</sub>Ns distinctly from each other in these mammalian cell extracts.

## **4.2 Analysis of metabolites by HPLC**

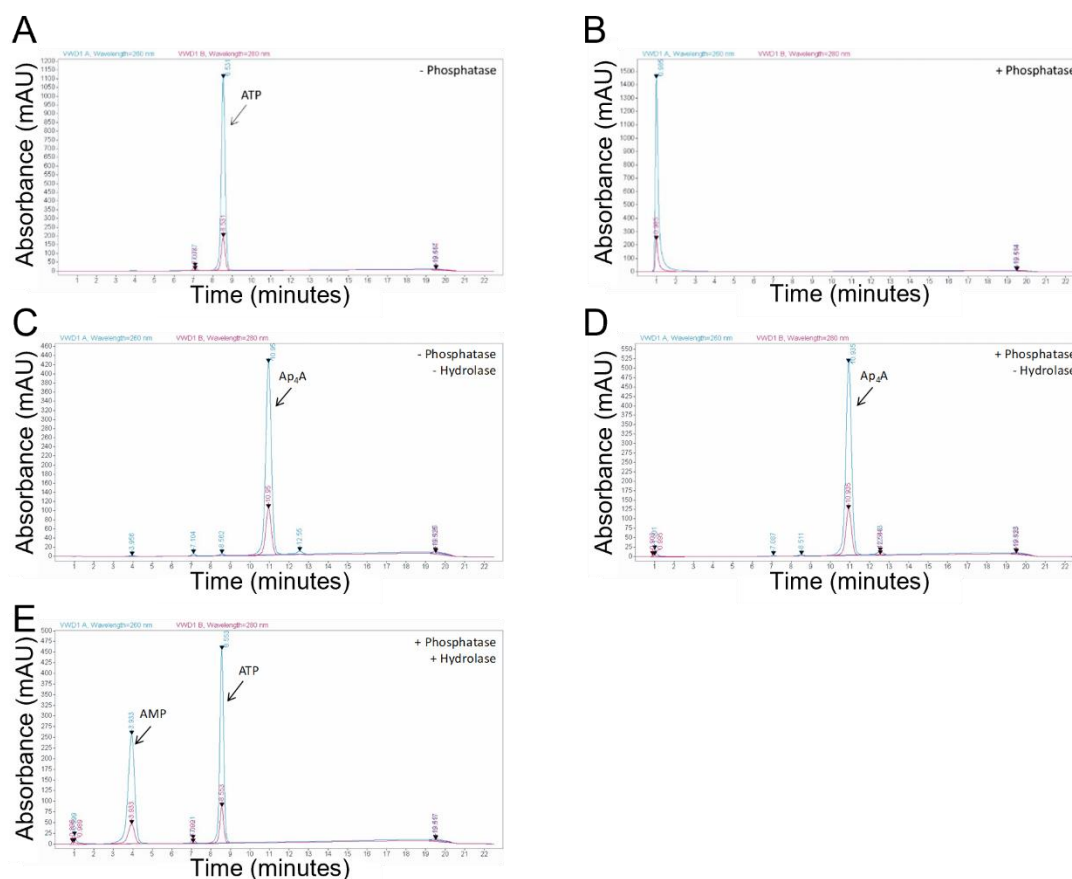
### **4.2.1 FastAP Alkaline phosphatase treatment has no effect on Ap<sub>4</sub>A**

As the extract collected from the mammalian cells would contain a large mixture of metabolites, we wanted to reduce the number of peaks that would elute to increase the likelihood of identifying Ap<sub>4</sub>A or the other Ap<sub>4</sub>Ns by HPLC. The purification of dinucleoside tetraphosphates was performed using anion exchange. As dinucleoside tetraphosphates are typically produced in the submicromolar range, competition from intracellular nucleoside triphosphates which are present in the millimolar range

(0.1–1 mM) may prevent efficient binding to the column. To reduce complexity, the use of FastAP alkaline phosphatase was used to remove the phosphate groups from the NTP and dNTPs and improve isolation of the Ap<sub>4</sub>Ns.

First, the activity of the FastAP alkaline phosphatase, which catalyses the removal of phosphate groups from nucleotides, needed to be confirmed. To confirm its functionality, the alkaline phosphatase was added to ATP, and the resulting sample was analysed by HPLC (Fig. 4.1). Then, to establish that this approach would not reduce the Ap<sub>4</sub>A levels, *in vitro* reactions were performed using alkaline phosphatase and/or Ap<sub>4</sub>A hydrolase. This approach is used for the classical determination of Ap<sub>4</sub>A concentration through extensive hydrolysis of the intracellular ATP pools with phosphatase treatments prior to Ap<sub>4</sub>A hydrolysis for quantitation of residual ATP levels (Baker and Jacobson, 1984; Marriott et al., 2015). To confirm that the phosphatase would remove exposed phosphates but not affect the Ap<sub>4</sub>Ns themselves, Ap<sub>4</sub>A was treated with phosphatase. The resulting sample was then run through the HPLC to determine whether the phosphatase affected Ap<sub>4</sub>A (Fig. 4.1).





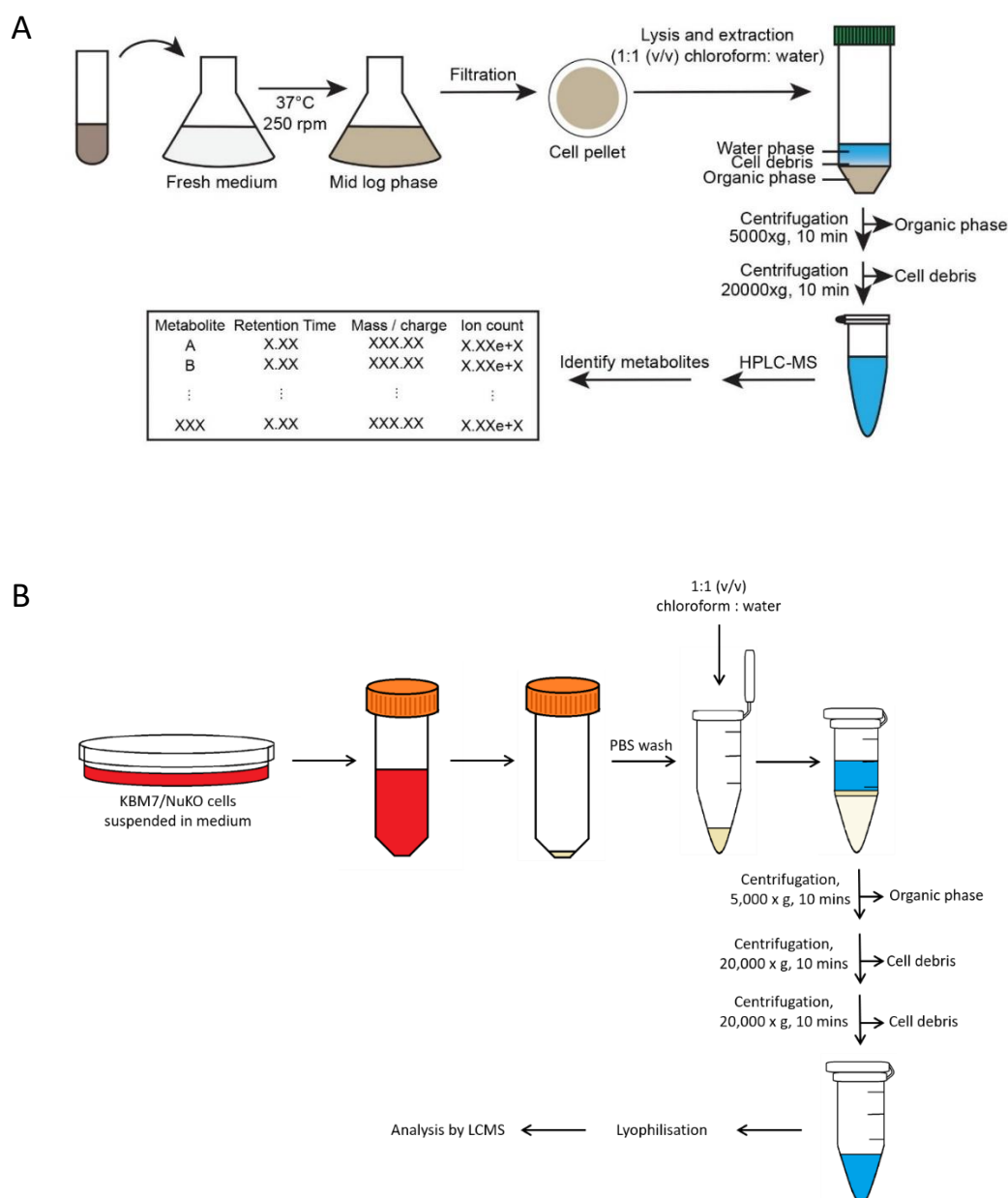
**Figure 4.1: The effect of phosphatase (and Ap<sub>4</sub>A hydrolase) on ATP and Ap<sub>4</sub>A.** A, B: Analysis of the stability of ATP by measuring retention time before (A) and after (B) phosphatase treatment. C–E: Analysis of the stability of Ap<sub>4</sub>A by measuring retention time before treatment (C), after phosphatase treatment (D) or after phosphatase and hydrolase incubation (E). Reaction products were analysed by HPLC and detected at UV wavelengths of 260 nm (blue) and 280 nm (red). Where the sample was treated with both phosphatase and hydrolase, the phosphatase was added first, then inactivated prior to addition of the hydrolase.

After treatment with alkaline phosphatase the retention time for the ATP sample was reduced from approximately 8.5 minutes to 1 minute (Fig. 4.1 A, B). This is consistent with the loss of all phosphates from the molecule, as these negatively charged phosphate groups were responsible for the increased retention time on the column. HPLC analysis revealed that untreated Ap<sub>4</sub>A had a retention time of 10.95 minutes (Fig. 4.1 C). This retention time was not affected by phosphatase treatment, demonstrating that the phosphatase had no effect on Ap<sub>4</sub>A levels (Fig. 4.1 D). As a further method of discriminating the Ap<sub>4</sub>Ns from other molecules absorbing in the 260–280nm range, we used an Ap<sub>4</sub>A hydrolase. In Chapter 3 we demonstrated that

Ap<sub>4</sub>A hydrolase is capable of cleaving Ap<sub>4</sub>Ns asymmetrically to produce NMP and NTP (Fig. 3.9 and 3.10) (Jakubowski and Guranowski, 1983; Swarbrick et al., 2005). To confirm that Ap<sub>4</sub>A could be fully hydrolysed after phosphatase treatment, the phosphatase was inactivated, then Ap<sub>4</sub>A hydrolase added to reactions. This resulted in complete hydrolysis of Ap<sub>4</sub>A into ATP and AMP (Fig. 4.1 E).

#### **4.2.2 Extraction of metabolites from mammalian cells and analysis by HPLC**

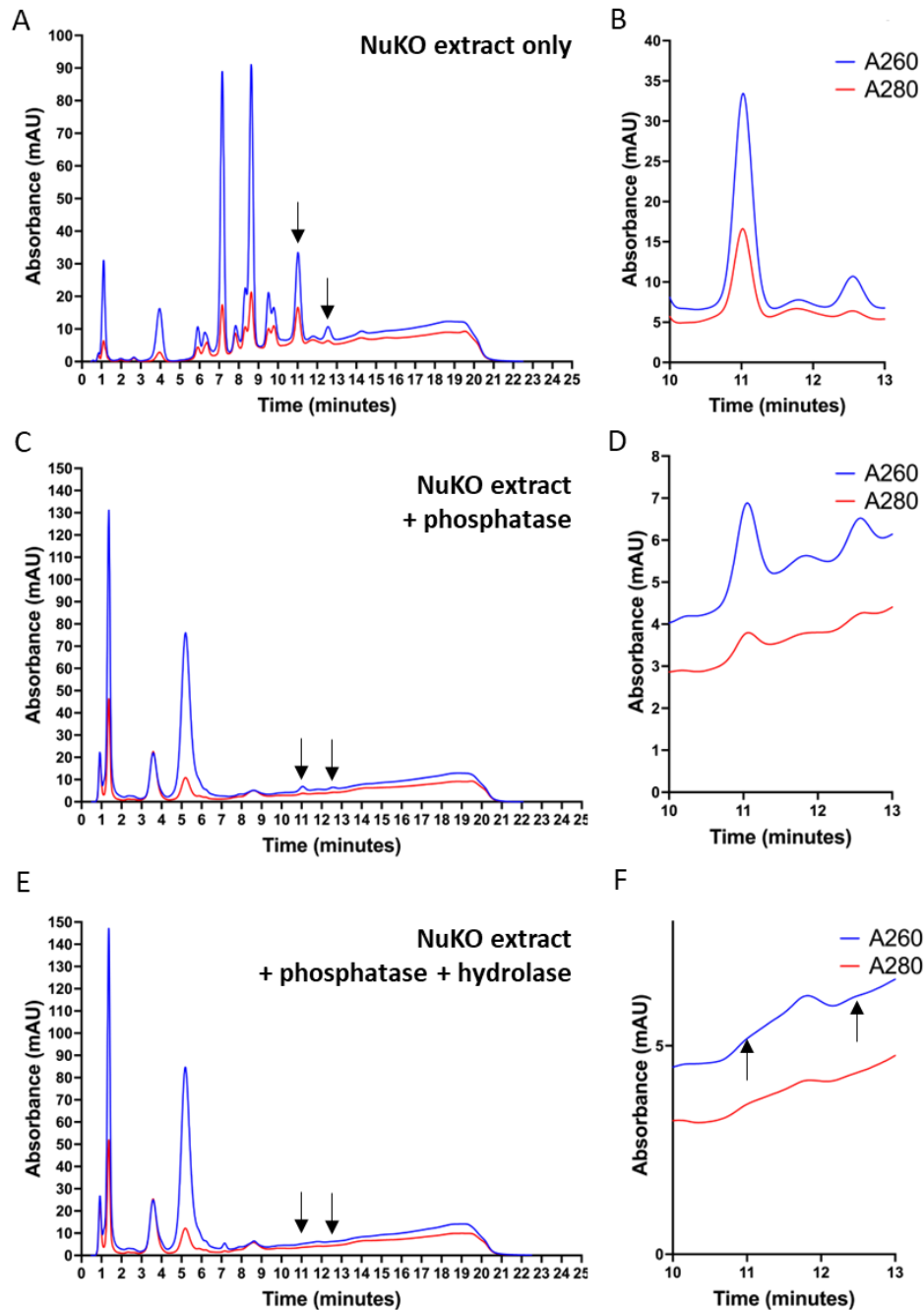
Having demonstrated that Ap<sub>4</sub>A is resistant to FastAP alkaline phosphatase activity, the approach of removing cellular NTPs from cell extracts could be used in the preparation of extracts for analysis by HPLC. The next step of the process was to extract cellular metabolites from mammalian cells for analysis. The approach would use two cell lines that have quantitatively different Ap<sub>4</sub>A levels, as determined by coupled enzyme assays (Marriott et al., 2016). The NUDT2 knockout (NuKO) cell line derived from KBM7 cells was used for this investigation as NuKO cells lack an Ap<sub>4</sub>A hydrolase, which promotes accumulation of Ap<sub>4</sub>A and presumably Ap<sub>4</sub>Ns, if synthesised. The NuKO cell line has been reported to have 175-fold higher levels of Ap<sub>4</sub>A compared with its parental KBM7 cell line (Marriott et al., 2016). Due to the nature of the experiments used, this value would also include increased levels of the other Ap<sub>4</sub>Ns, if present. The increased level of Ap<sub>4</sub>Ns in this cell line may facilitate detection using UV absorbance and HPLC to investigate and identify the presence of other Ap<sub>4</sub>Ns in mammalian cells. To extract the cellular metabolites, a chloroform extraction method was adapted from the method used by Fung and colleagues to extract metabolites including Ap<sub>4</sub>A from *B. subtilis* cells (Fung et al., 2020) (Fig. 4.2).



**Figure 4.2: Schematics for the extraction of metabolites from cells.** **A:** Schematic shown by Fung and colleagues to show the method for the extraction of metabolites from *B. subtilis* cells (Fung et al., 2020). **B:** The method for the extraction of Ap<sub>4</sub>Ns from mammalian cells was adapted from the method described by Fung and colleagues in 2020.

Figure 4.2 A reproduced without changes from Frontiers in Microbiology, Volume 11, Fung, D. K., Yang, J., Stevenson, D. M., Amador-Noguez, D. & Wang, J. D., Small Alarmone Synthetase SasA Expression Leads to Concomitant Accumulation of pGpp, ppApp, and AppppA in *Bacillus subtilis*, Copyright 2020, doi: <https://doi.org/10.3389/fmicb.2020.02083>, available under the CC BY 4.0 licence (<https://creativecommons.org/licenses/by/4.0/>) and disclaimer of warranties within.

Suspension cells were cultured, then the pellet resuspended in an ice-cold solution of chloroform and water (Fig. 4.2 B). The aqueous layer was then collected by centrifugation, lyophilised, and resuspended in water. The resulting extract was analysed by HPLC and peaks eluting with similar retention times to the synthesised Ap<sub>4</sub>Ns were detected (Fig. 4.3 A, B). Separated samples prepared simultaneously were treated with either phosphatase (Fig. 4.3 C, D) or phosphatase followed by Ap<sub>4</sub>A hydrolase (Fig. 4.3 E, F), to determine whether the identified peaks responded in a similar manner to what was expected for each Ap<sub>4</sub>N.



**Figure 4.3: Analysis of NuKO cell extracts by HPLC.** Metabolites were extracted from NuKO cells and analysed by HPLC. **A,B:** The sample was injected without further treatment. **C,D:** The extracts were treated with alkaline phosphatase prior to analysis by HPLC. **E,F:** The extracts were treated with alkaline phosphatase and  $\text{Ap}_4\text{A}$  hydrolase before HPLC analysis. Where the sample was treated with both phosphatase and hydrolase, the phosphatase was added first, then inactivated prior to addition of the hydrolase. The right-hand panels show the HPLC traces focussing on peaks eluting at similar times to the expected retention times of the  $\text{Ap}_4\text{Ns}$ .

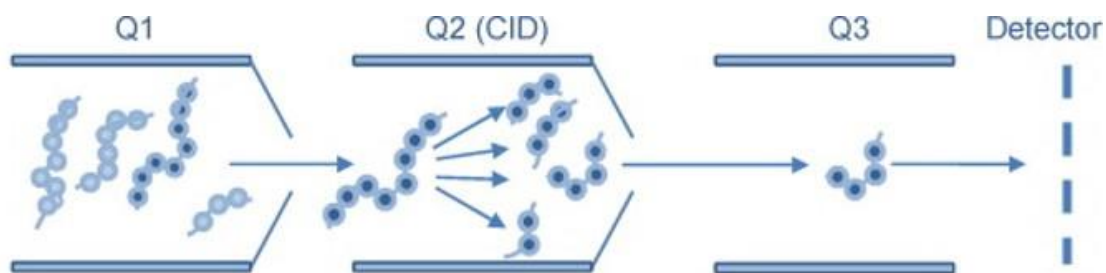
In the untreated NuKO extract, there were many peaks representing molecules that absorbed at 260 nm and 280 nm (Fig. 4.3 A). The elution times of two peaks at ~11 and ~12.5 minutes show similar elution positions to HRMS-verified Ap<sub>4</sub>A and Ap<sub>4</sub>U (Fig. 4.3 B).

To reduce the number of peaks, another sample was run that had been treated with a phosphatase to remove any exposed phosphate groups on cellular NTPs that may obscure low abundance Ap<sub>4</sub>N peaks (Fig. 4.3 C). When treated with alkaline phosphatase, most of the peaks at higher retention times disappeared, consistent with the degradation of nucleotide di- and tri-phosphate degradation to earlier eluting nucleotides (Fig. 4.3 C). In addition, the area under both candidate Ap<sub>4</sub>A and Ap<sub>4</sub>U peaks was reduced, suggesting that a proportion of the peaks at 11 and 12.5 minutes were not Ap<sub>4</sub>A or Ap<sub>4</sub>U, as they were lost during phosphatase treatment. However, although smaller, peaks remain at both 11- and 12.5-minutes retention times (Fig. 4.2 C, D).

To determine whether these remaining peaks could be due to the presence of Ap<sub>4</sub>A and Ap<sub>4</sub>U in the extract mixture, a third NuKO extract sample was treated with phosphatase. This phosphatase was inactivated, and the remaining sample was treated with Ap<sub>4</sub>A hydrolase, which would break down any Ap<sub>4</sub>N to ATP, NTP, AMP and NMP if present. Addition of Ap<sub>4</sub>A hydrolase caused both the Ap<sub>4</sub>A and Ap<sub>4</sub>U candidate peaks to disappear (Fig. 4.3 E and F) whereas the rest of the peaks in the sample seemed largely unaffected. Together, these data are consistent with the presence of Ap<sub>4</sub>A and Ap<sub>4</sub>U in the NuKO cell extracts. The approach had its limitations as the peaks corresponding to the retention times of Ap<sub>4</sub>C and Ap<sub>4</sub>G could not be distinguished in the HPLC trace. However, this could be due to their concentrations being below the sensitivity limit of the HPLC machine and so the presence of these Ap<sub>4</sub>Ns in the cell extract should not be dismissed. These data suggest that Ap<sub>4</sub>Ns are present in cellular extracts and suggest that a more sensitive and quantitative approach is required for the intracellular detection of dinucleoside tetraphosphates.

### 4.3 Validation of Ap<sub>4</sub>A detection and analysis by LC-MS

Analysis by HPLC suggested that at least some of the Ap<sub>4</sub>Ns may be detectable in mammalian cells. We therefore decided to develop a more sensitive technique for identification of the Ap<sub>4</sub>Ns from chloroform extracted cellular extracts using Triple Quadrupole Mass Spectrometry (Fig. 4.4).



**Figure 4.4: The triple quadrupole system as shown by Kitteringham and colleagues (Kitteringham et al., 2009).** The precursor ion is selected in the 1<sup>st</sup> quadrupole, then fragmented in the 2<sup>nd</sup> quadrupole. Fragments corresponding to the molecular weight of the standard fragments are selected in the 3<sup>rd</sup> quadrupole. These fragments go on to be detected for identification of the target molecule.

Reprinted from the Journal of Chromatography B, Volume 877, Kitteringham, N. R., Jenkins, R. E., Lane, C. S., Elliott, V. L. & Park, B. K., Multiple reaction monitoring for quantitative biomarker analysis in proteomics and metabolomics, Pages No. 1229–1239, Copyright 2009, with permission from Elsevier.

This technique is highly specific to a given structure as it distinguishes molecules by a combination of several characteristics - molecular weight, fragmentation pattern and chromatographic elution time.

#### 4.3.1 Determining MRM transitions for Ap<sub>4</sub>N standards

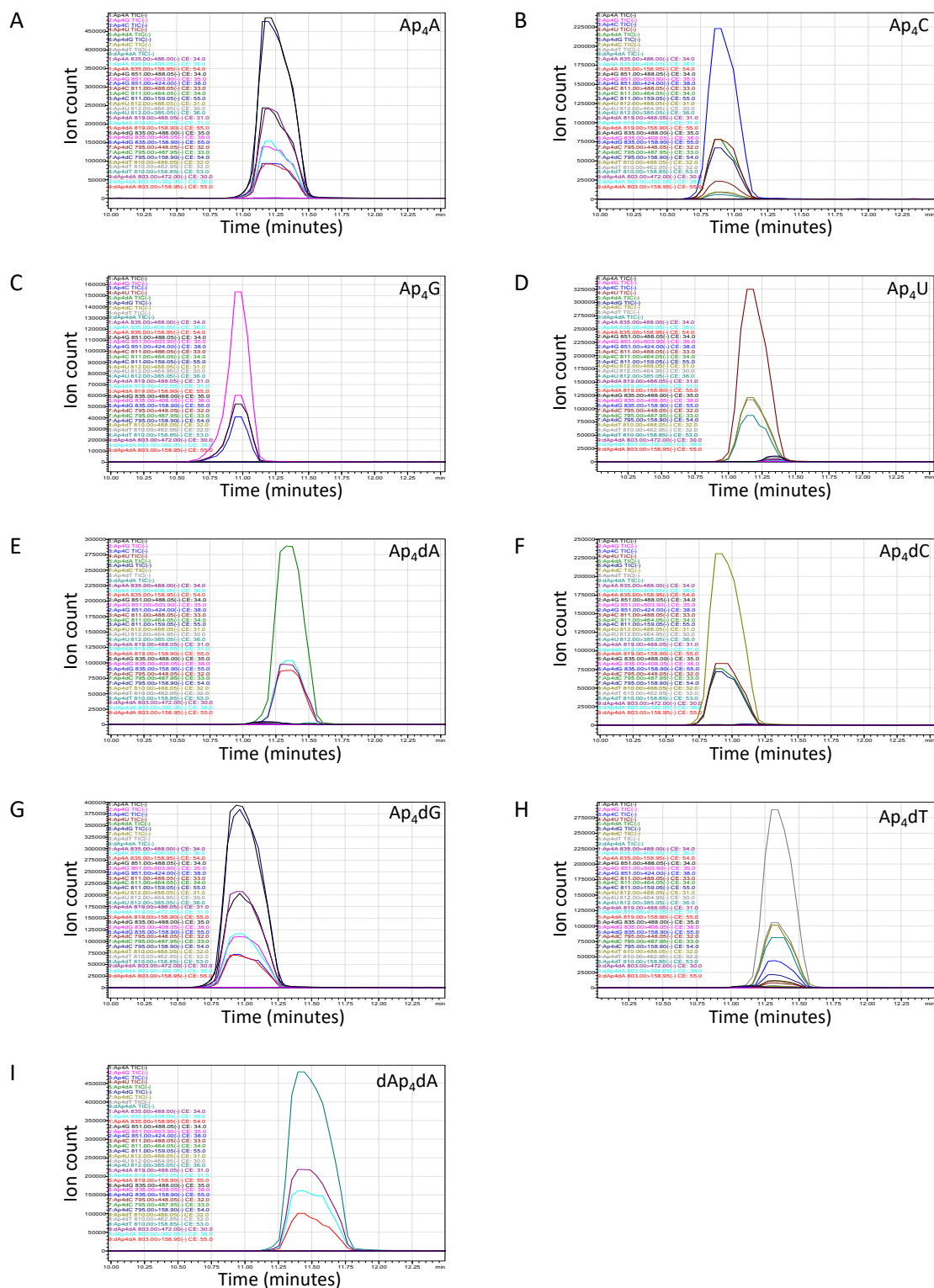
As a starting point, the chemically synthesised Ap<sub>4</sub>Ns that had been validated by HRMS were used as standards to establish the correct parameters for detection and validation. The LC-MS was set to determine the most frequently formed ion fragments for each molecule tested. The three most abundant transitions for each molecule are listed in Table 4.1.

**Table 4.1: Summary of the three most abundant transitions for each Ap<sub>4</sub>N**

	Transitions		
	1	2	3
<b>Ap<sub>4</sub>A</b>	835.00>488.00	835.00>408.05	835.00>158.95
<b>Ap<sub>4</sub>C</b>	811.00>488.05	811.00>464.05	811.00>159.05
<b>Ap<sub>4</sub>G</b>	851.00>488.05	851.00>503.90	851.00>424.00
<b>Ap<sub>4</sub>U</b>	812.00>488.05	812.00>464.05	812.00>385.05
<b>Ap<sub>4</sub>dA</b>	819.00>488.05	819.00>472.04	819.00>158.90
<b>Ap<sub>4</sub>dC</b>	795.00>448.05	795.00>487.95	795.00>158.90
<b>Ap<sub>4</sub>dG</b>	835.00>488.00	835.00>408.05	835.00>158.90
<b>Ap<sub>4</sub>dT</b>	810.00>488.05	810.00>464.95	810.00>385.05
<b>dAp<sub>4</sub>dA</b>	803.00>472.00	803.00>392.05	803.00>158.95

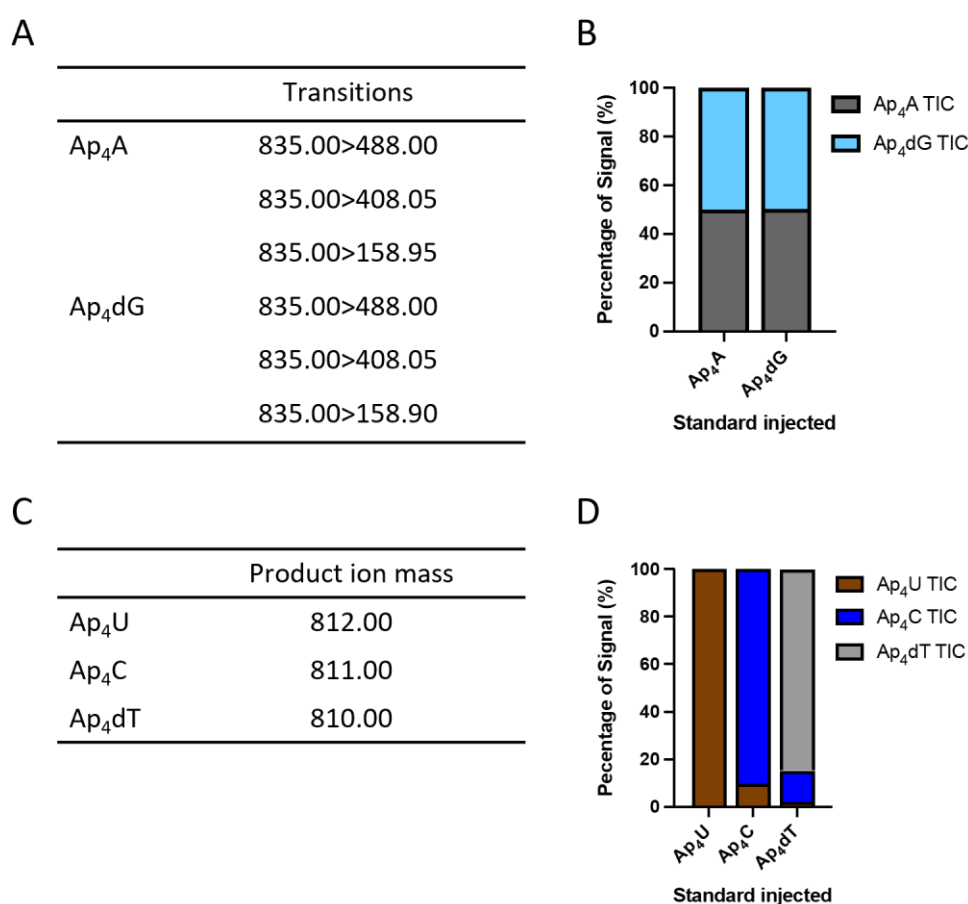
Each Ap<sub>4</sub>N sample was individually analysed, and the automated wizard set to identify ions that corresponded to the transitions for each of the Ap<sub>4</sub>Ns. The resultant chromatogram traces for each of the transitions, in addition to the total ion count (TIC), are shown in Figure 4.5.





**Figure 4.5: Determination and validation of MRM transitions for each Ap<sub>4</sub>N.** The most abundant MRM transitions were identified in HRMS validated standards. Standards were re-run with the machine set to detect the chosen MRM transitions for all Ap<sub>4</sub>Ns to confirm that they could be identified. Ap<sub>4</sub>Ns are shown in the order: Ap<sub>4</sub>A (A), Ap<sub>4</sub>C (B), Ap<sub>4</sub>G (C), Ap<sub>4</sub>U (D), Ap<sub>4</sub>dA (E), Ap<sub>4</sub>dC (F), Ap<sub>4</sub>dG (G), Ap<sub>4</sub>dT (H) and dAp<sub>4</sub>dA (I).

As the high salt concentration used for HPLC was not compatible with the LC-MS, a C18 column was used with an ammonium formate:acetonitrile gradient. When run through this column, the selected transitions successfully identified all Ap<sub>4</sub>Ns. However, due to the similarity in their structures and the type of column used, all Ap<sub>4</sub>Ns had very similar retention times of between 10.75 and 11.75 minutes (Fig. 4.5). Of the Ap<sub>4</sub>Ns, Ap<sub>4</sub>G, Ap<sub>4</sub>dA, Ap<sub>4</sub>dC and dAp<sub>4</sub>dA were identified only in the appropriate samples (Fig. 4.5 C, E, F and I). However due to the structural similarities of some of the molecules there was an overlap in the molecules identified by some of the transitions. The most apparent case of this was with Ap<sub>4</sub>A and Ap<sub>4</sub>dG, which each showed six transitions rather than the expected three transitions (Fig. 4.5 A, G; Fig. 4.6).



**Figure 4.6: Assessing the abundance of transitions common to multiple molecules using Ap<sub>4</sub>N standards.** **A:** The similarities in the most abundant transitions for Ap<sub>4</sub>A and Ap<sub>4</sub>dG. **B:** The proportion and identity of the transitions triggered by each of the Ap<sub>4</sub>A and Ap<sub>4</sub>dG standards. **C:** A comparison of the product ion masses for Ap<sub>4</sub>U, Ap<sub>4</sub>C and Ap<sub>4</sub>dT. **D:** The proportion and identity of transitions triggered by the Ap<sub>4</sub>U, Ap<sub>4</sub>C and Ap<sub>4</sub>dT standards.

Ap<sub>4</sub>A and Ap<sub>4</sub>dG have the same mass and share the same ion fragment peaks (Fig. 4.6 A), with the adenosine end of the molecule generally being present in the most abundant fragments. Unfortunately, no fragments were found in high enough abundance that were specific only to Ap<sub>4</sub>dG or to Ap<sub>4</sub>A. As the molecular and fragment ion peaks were almost identical, the method was unable to distinguish between them. Approximately 50% of the signal resulting from the Ap<sub>4</sub>A standard was detected as Ap<sub>4</sub>dG, and *vice versa* (Fig. 4.6 B). This meant that any result found in the cell extract could potentially represent either molecule. The similarity in mass and chemical structure implies that if either of these molecules were detected in the extracts, we would need to determine a method of identifying whether one or both molecules were present.

Additionally, the Ap<sub>4</sub>C and Ap<sub>4</sub>dT standards also triggered a signal for Ap<sub>4</sub>U or Ap<sub>4</sub>C (Fig. 4.5 B, H; Fig 4.6 D). Ap<sub>4</sub>U, Ap<sub>4</sub>C and Ap<sub>4</sub>dT have distinct molecular masses; however, they are each only separated by 1 g/mol (Fig. 4.6 C). Ap<sub>4</sub>U has the highest molecular weight, and so none of the other molecular signals were triggered. Approximately 10% of the total peak area in the Ap<sub>4</sub>C standard also triggered detection of Ap<sub>4</sub>U, while approximately 13% and 2% of the Ap<sub>4</sub>dT standard suggested the presence of Ap<sub>4</sub>C and Ap<sub>4</sub>U, respectively (Fig. 4.6 D). As these molecules are similar in molecular weight, the triggers are likely due to the different masses of a molecule caused by the natural presence of different isotopes, known as the isotopic envelope of a compound. The molecular mass is based on the most abundant isotope for each element in the molecule; however, other isotopes exist which have different molecular weights. For example, the most abundant Carbon isotope is <sup>12</sup>C. However, <sup>13</sup>C is also present naturally and is composed of an additional neutron, causing any molecule containing <sup>13</sup>C to have +1 g/mol added to its molecular mass. In nature, <sup>13</sup>C exists at an abundance of approximately 1.1%. As there are either 19 or 20 carbon atoms in a single Ap<sub>4</sub>N molecule, this means that any given Ap<sub>4</sub>N molecule has approximately a 20% probability of containing a single <sup>13</sup>C isotope, which would increase the mass of that molecule by 1 g/mol. This isotopic envelope can be seen for each Ap<sub>4</sub>N molecule in the HRMS spectra in Figure 3.11 and 3.12. The probability of two <sup>13</sup>C isotopes occurring in the same Ap<sub>4</sub>N molecule is much less likely, explaining why the likelihood

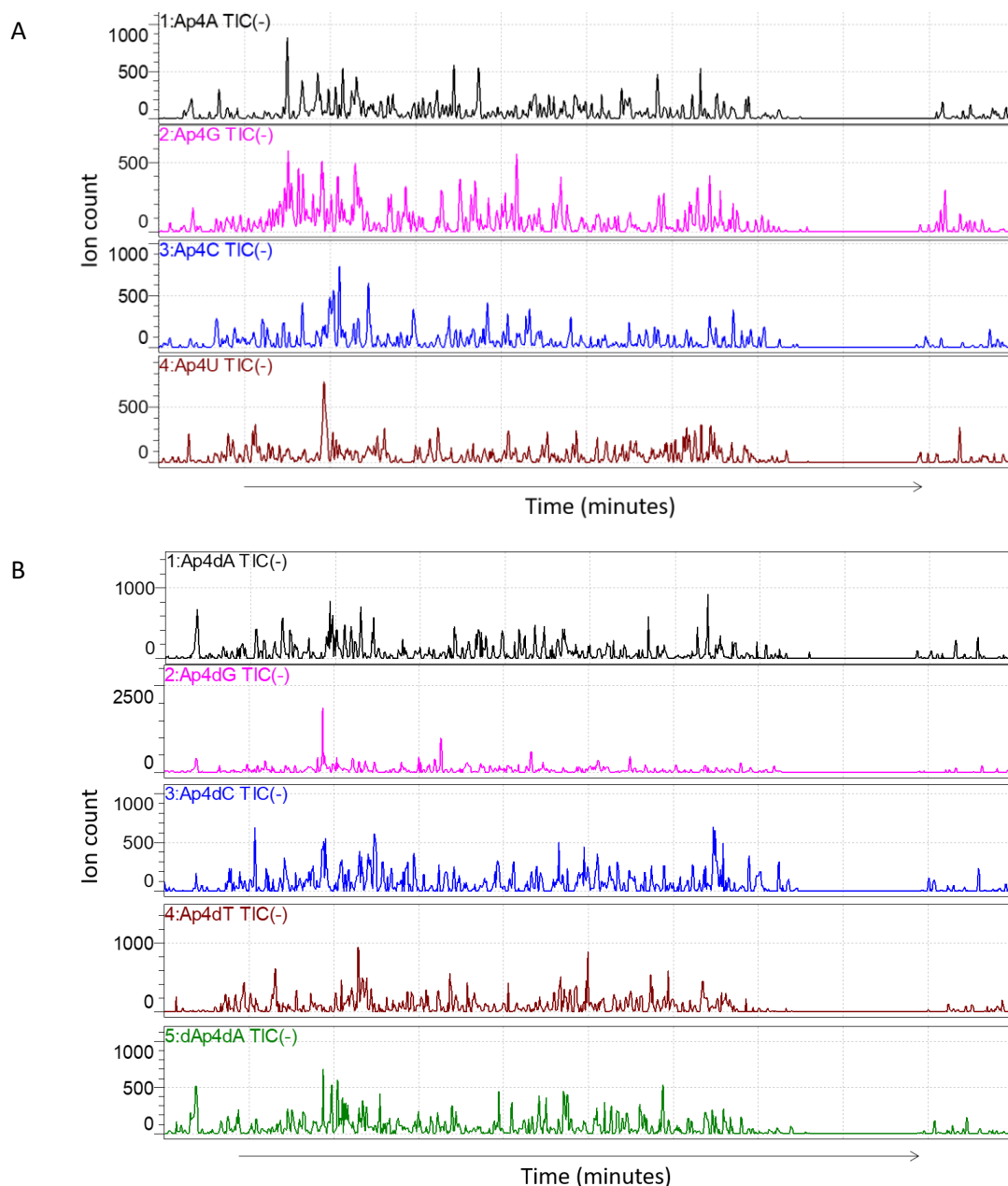
of an  $\text{Ap}_4\text{U}$  peak (molecular mass = 812.00) showing up in an  $\text{Ap}_4\text{dT}$  standard (molecular mass = 810.00) is less likely (Fig. 4.6 D).

$\text{Ap}_4\text{U}$ ,  $\text{Ap}_4\text{C}$  and  $\text{Ap}_4\text{dT}$  have similar retention times and so cannot be distinguished chromatographically. The potential for an incorrect signal to occur is therefore a limitation of this system for the detection of individual  $\text{Ap}_4\text{Ns}$ . If these molecules are detected in the cell extracts it is important to acknowledge this limitation; however, these standards suggest that most of the signal for  $\text{Ap}_4\text{C}$  and  $\text{Ap}_4\text{dT}$  is accurate. If all these molecules are present it may be possible to calculate how much of each is likely to be due to these different isotopes based on the proportion of the inappropriate signal in Figure 4.6 D.

#### **4.4 KBM7 cell extracts**

##### **4.4.1 Investigating KBM7 cell extracts for the presence of $\text{Ap}_4\text{Ns}$**

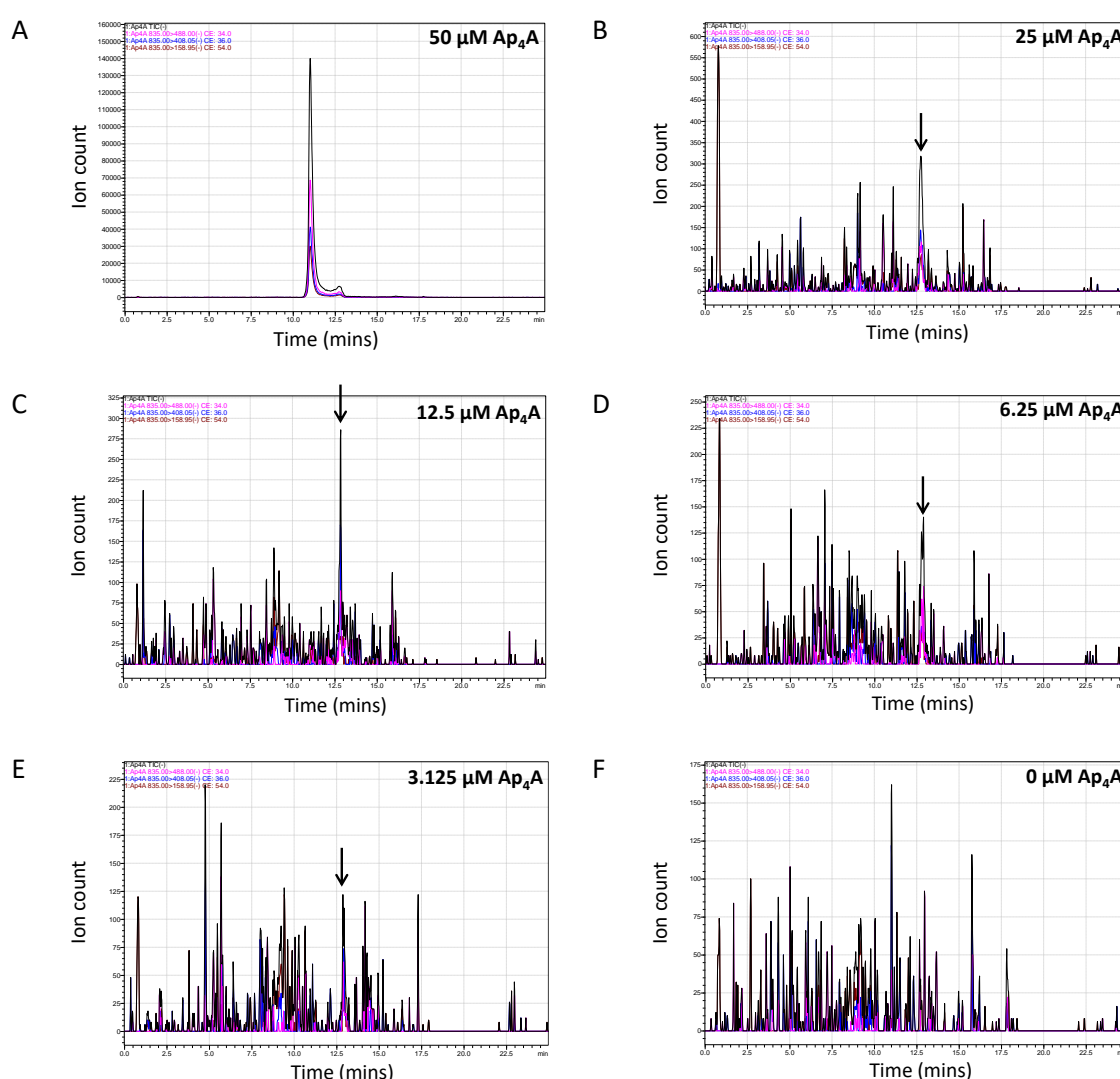
As the mass spectrometer is more sensitive than the HPLC, we started by attempting to identify intracellular  $\text{Ap}_4\text{Ns}$  in the parental KBM7 cells. The metabolites were harvested from the KBM7s following the same process as used prior to the HPLC analysis. Subsequently, the sample was analysed by LC-MS. The same sample was injected twice, first so the spectrometer could search for the  $\text{Ap}_4\text{Ns}$ , and then so it could search for the  $\text{Ap}_4\text{dNs}$  (Fig. 4.7 A and B, respectively).



**Figure 4.7: Investigating whether Ap<sub>4</sub>Ns and Ap<sub>4</sub>dNs are present in KBM7 cells.** Metabolites were harvested from KBM7 cells and analysed by LC-MS, and separated in the chromatography phase on an Advance Bio peptide mapping C18 column. The sample was injected in two consecutive runs to minimise the number of molecules the machine was looking for at once, firstly searching for the Ap<sub>4</sub>Ns (A) and secondly for the Ap<sub>4</sub>dNs (B). The total ion count for each molecule is shown here for simplicity.

None of the Ap<sub>4</sub>Ns or Ap<sub>4</sub>dNs, including Ap<sub>4</sub>A, were identifiable within the KBM7 cell extracts, as only baseline signals were observed for each molecule (Fig. 4.7). However, Ap<sub>4</sub>A is present in KBM7 cells at a concentration of  $0.21 \pm 0.2$  pmol/ $10^6$  cells (Marriott et

al., 2016). Assuming the even distribution of Ap<sub>4</sub>A in a typical mammalian cell with a volume of 2000  $\mu\text{m}^3$ , the authors roughly equated an Ap<sub>4</sub>A concentration of 1 pmol/ $10^6$  cells to an intracellular Ap<sub>4</sub>A concentration of 0.5  $\mu\text{M}$  (Marriott et al., 2015). This therefore equated to an intracellular Ap<sub>4</sub>A concentration of approximately 0.1  $\mu\text{M}$  in the KBM7 cell line. As this concentration is so low, it is possible that the LC-MS is not sensitive enough to detect it. To determine the concentration of Ap<sub>4</sub>A required for the spectrometer to identify Ap<sub>4</sub>A within the cell extract, the extract was harvested and spiked with Ap<sub>4</sub>A at a range of concentrations (Fig. 4.8).



**Figure 4.8: Determining the lower Ap<sub>4</sub>A concentration for detection in KBM7 extracts.** Metabolites were extracted from KBM7 cells using the same method as for sample preparation prior to HPLC analysis. After lyophilising the extract, the sample was spiked with either 50  $\mu\text{M}$  (A), 25  $\mu\text{M}$  (B), 12.5  $\mu\text{M}$  (C), 6.25  $\mu\text{M}$  (D), 3.125  $\mu\text{M}$  (E) or 0  $\mu\text{M}$  (F) Ap<sub>4</sub>A, then 10  $\mu\text{L}$  of each sample was injected for analysis by LC-MS.

Spiking the sample with 50  $\mu\text{M}$   $\text{Ap}_4\text{A}$  resulted in a large  $\text{Ap}_4\text{A}$  peak at approximately 11 minutes, with the peak intensity for the TIC reaching approximately 140,000 units (Fig. 4.8 A). This peak was slightly asymmetrical, with a small tailing peak showing at  $\sim 12.5$  minutes. When the sample was spiked with 12.5  $\mu\text{M}$  or 25  $\mu\text{M}$   $\text{Ap}_4\text{A}$ , the peak intensity dropped to just 300 units, but remained composed of all three of the expected  $\text{Ap}_4\text{A}$  MRM transitions (Fig. 4.8 B and C). However, this peak was not easily distinguishable from the background noise and eluted at  $\sim 12.5$  minutes, later than the large  $\text{Ap}_4\text{A}$  peak eluting at higher concentrations. The samples spiked with 6.25  $\mu\text{M}$  and 3.125  $\mu\text{M}$   $\text{Ap}_4\text{A}$  appeared initially to show only background noise (Fig. 4.8, D and E). Upon closer inspection, small peaks can be identified at the same time as those in the 12.5  $\mu\text{M}$  and 25  $\mu\text{M}$  samples; however, their total ion counts are approximately 125 units, smaller than many of the random peaks causing the background noise. As seen previously in Figure 4.7,  $\text{Ap}_4\text{A}$  was not identified in the un-spiked KBM7 extract (Fig. 4.8 F). The large difference in intensity between the 50  $\mu\text{M}$ - and 25  $\mu\text{M}$ -spiked  $\text{Ap}_4\text{A}$  samples suggests that the LC-MS is struggling to detect  $\text{Ap}_4\text{A}$  at concentrations of 25  $\mu\text{M}$  and below. This could be due to the sensitivity limit of the machine, or due to and ion suppression effect induced by other molecules in the cell extract.

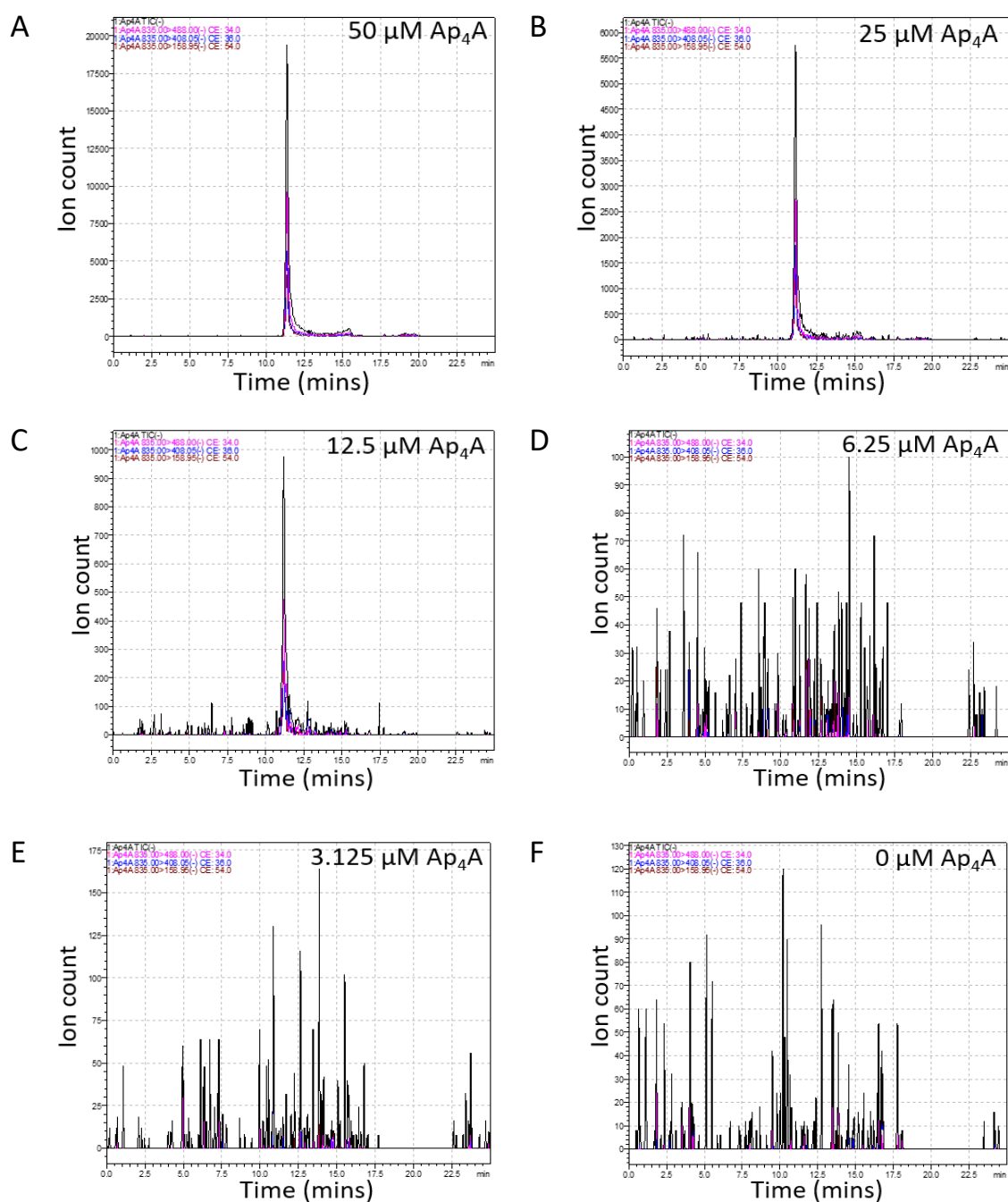
#### **4.4.2 Clean-up of the KBM7 extract using preparative chromatography approaches**

Ion suppression is an effect seen in mass spectrometry which can result in a reduced amount of the desired ion reaching the detector (Annesley, 2003). Compounds in the sample that have a low volatility droplet formation or evaporation efficiency result in reduced amounts of the charged ion reaching the detector (Annesley, 2003; King et al., 2000). Therefore, removing any unnecessary ions from the sample is important in reducing the ion suppression effect.

To attempt to lower the concentration at which  $\text{Ap}_4\text{A}$  which was detectable by LC-MS, we introduced a clean-up step to the extract preparation to reduce the number of unnecessary ions in the mixture. In this process, the KBM7 extract was harvested as previously described. Initially, both ion exchange and graphite columns were investigated as methods for the clean-up of the cell extract. Ion exchange columns can be either anion or cation-exchange and separate molecules based on charge. As the

Ap<sub>4</sub>Ns each contain four phosphate groups, they carry a strong negative charge. These molecules would therefore pass through a cation exchange column and any positively charged ions in the mixture would be retained on the column, removing them from the sample. Graphite columns bind hydrophilic peptides and retain these molecules until eluted. However, they have also been used for the purification of sugar nucleotides from cell extracts (Turnock and Ferguson, 2007). As each type of column will remove certain ions more efficiently, we aimed to determine whether either of these would sufficiently reduce ion suppression enough for the Ap<sub>4</sub>Ns to be detectable in the KBM7 cell line. Both Harvard microspin strong cation exchange spin columns and Pierce (carbon) graphite columns were tested, but neither resulted in the identification of Ap<sub>4</sub>Ns in the KBM7 sample (Supplementary Fig. 4.1). As the Harvard columns were cation exchange columns, these only allowed us to collect the flow through as they would not retain the negatively charged Ap<sub>4</sub>Ns. This meant that only positively charged ions could be removed from the mixture, rather than specifically purifying out molecules with similar qualities to the Ap<sub>4</sub>Ns. We therefore decided to continue optimisation of this method with the Pierce spin columns, as these were graphite columns to which the Ap<sub>4</sub>Ns could bind independently of charge. To determine the concentration of Ap<sub>4</sub>A that could be identified after a Pierce-column clean-up, Ap<sub>4</sub>A was spiked into KBM7 extract at concentrations between 0 and 50 µM prior to the clean-up step. The sample was lyophilised once more, and the concentrated samples were analysed by LC-MS (Fig. 4.9).

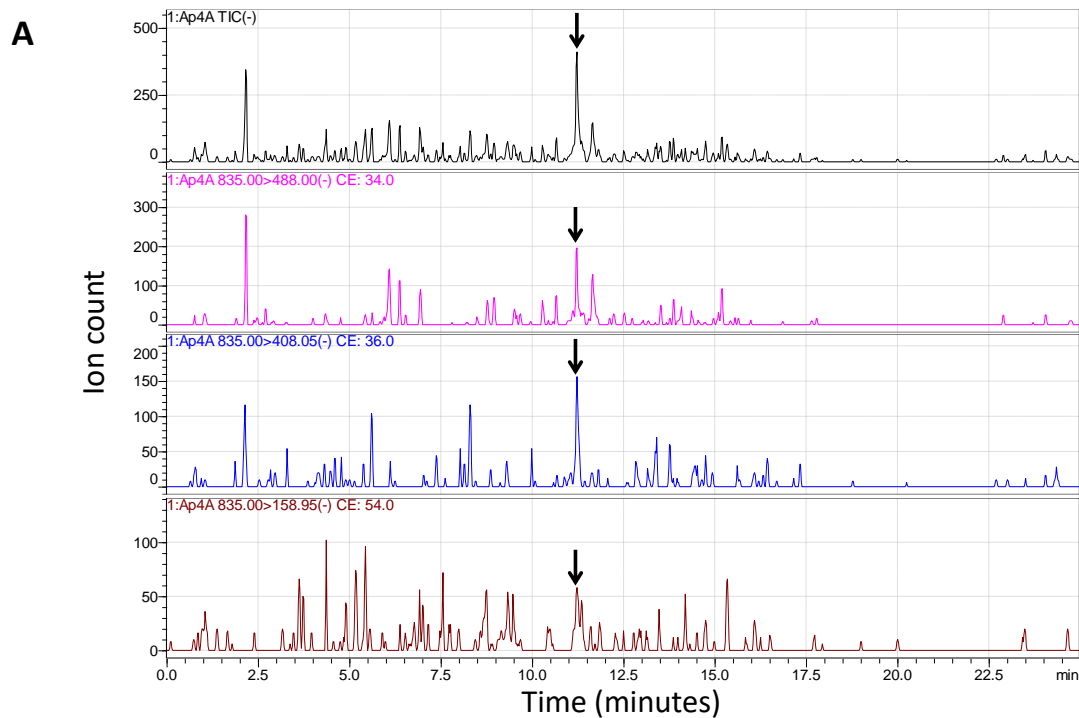




**Figure 4.9: The effect of the Pierce cleanup on detection of Ap<sub>4</sub>A at lower concentrations.** Metabolites were extracted from KBM7 cells using the same method for sample preparation as for HPLC analysis. After lyophilising the extract, the sample was prepared with additional Ap<sub>4</sub>A at different concentrations: 50  $\mu$ M Ap<sub>4</sub>A (A), 25  $\mu$ M Ap<sub>4</sub>A (B), 12.5  $\mu$ M Ap<sub>4</sub>A (C), 6.25  $\mu$ M Ap<sub>4</sub>A (D), 3.125  $\mu$ M Ap<sub>4</sub>A (E) and 0  $\mu$ M Ap<sub>4</sub>A (F). The samples were then cleaned up using a Pierce column and lyophilised once again before resuspending in water and injecting into the LC-MS for analysis.

Introducing the clean-up step reduced the ion count for the sample when 50  $\mu$ M Ap<sub>4</sub>A was spiked in, suggesting that Ap<sub>4</sub>A may have been lost in the clean-up step

(Fig. 4.9 A). Despite this, the clean-up method allowed Ap<sub>4</sub>A to be easily distinguishable from the background noise when spiked-in at concentrations of 12.5 µM or higher (Fig. 4.9 A to C). Despite helping to improve the sensitivity to Ap<sub>4</sub>A in the extract mixture, the level of Ap<sub>4</sub>A was still too low to be detected in the KBM7s when no Ap<sub>4</sub>A was spiked in (Fig. 4.9 F), or at spiked-in concentrations of 3.125 µM or 6.25 µM (Fig. 4.9 D and E). This may be linked to the detection limit of the LC-MS system as this clean-up process should have removed most of the unnecessary ions that could be inducing an ion suppression effect. To get the intracellular Ap<sub>4</sub>N concentration over the lower detectable limit for the LC-MS, a larger cell number ( $36 \times 10^6$  cells) was used for the extraction of the metabolites alongside an increased injection volume (20 µL) (Fig. 4.10).



**B**

	Retention time (minutes)		
	835>488	835>408	835>158
Ap <sub>4</sub> A Standard	11.2	11.2	11.2
KBM7 extract	11.2	11.2	-

**C**

	Ratio relative to 835>488 transition		
	835>488	835>408	835>158
Ap <sub>4</sub> A Standard	1.0	0.6	0.4
KBM7 extract	1.0	1.2	-

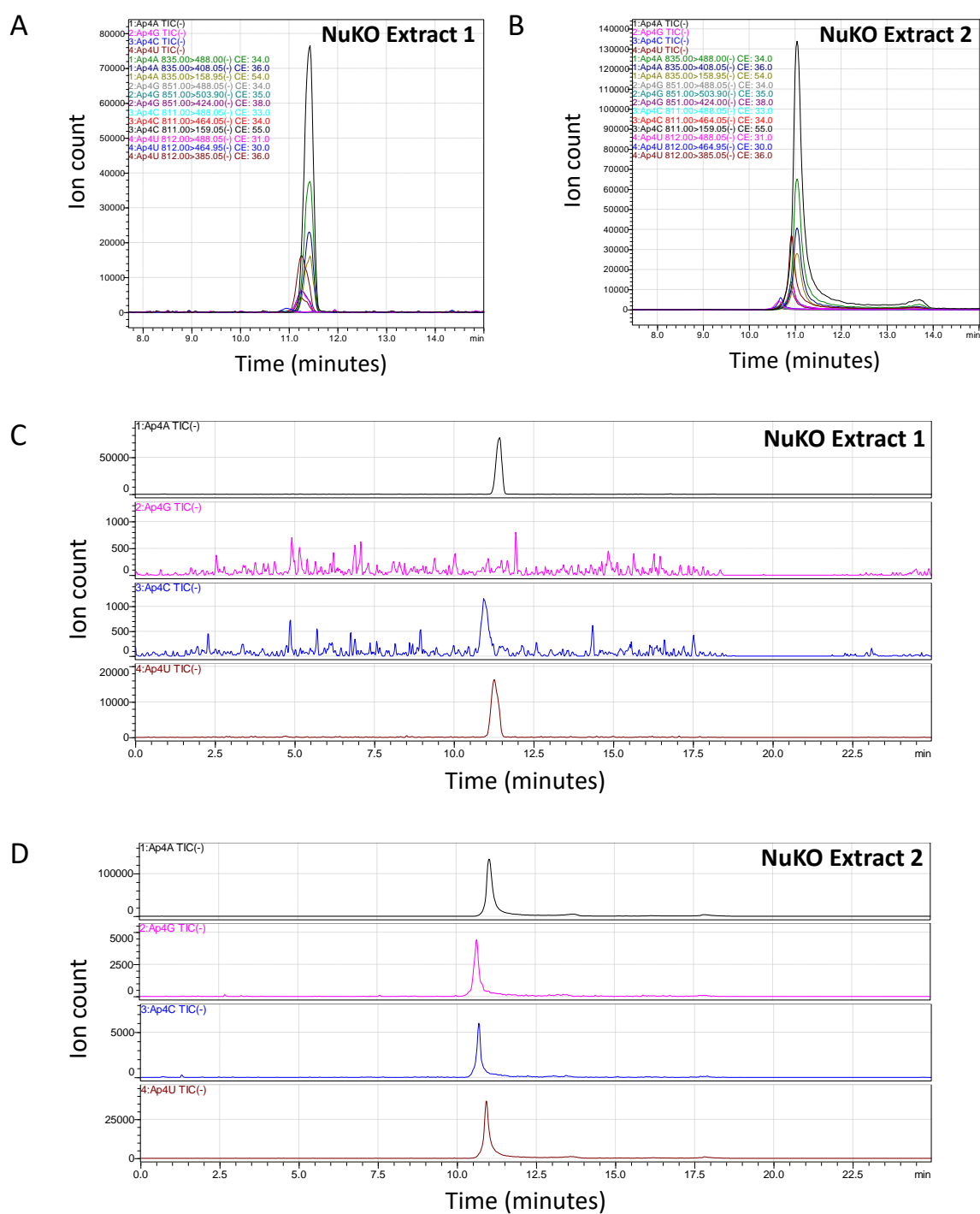
**Figure 4.10: Identification of Ap<sub>4</sub>A in a KBM7 cell extract.** Metabolites were extracted from  $36 \times 10^6$  KBM7 cells, a Pierce column cleanup was performed and the samples were analysed by LC-MS. **A:** The peaks for the total ion count (TIC) and all three transitions corresponding to Ap<sub>4</sub>A. The peak eluting at the expected retention time for each transition is shown by the black arrow. **B:** Comparison of the expected retention time for each Ap<sub>4</sub>A transition compared with the peak in the KBM7 extract. **C:** The relative ratio of the abundance of each transition, where the 835>488 transition is set equal to 1.

A small peak for two of the Ap<sub>4</sub>A MRM transitions appeared at the appropriate elution time (Fig. 4.10 B), denoted by the black arrows in Figure 4.10 A. While a small peak was also present for the third transition, the intensity of this peak is below background levels. The MRM transitions that were a high enough concentration to be measurable in the extract were not present in ratios comparable to the Ap<sub>4</sub>A standard (Fig. 4.10 C). This is likely due to the low concentration of Ap<sub>4</sub>A in the sample and it being at the lower sensitivity limit of the LC-MS. These data are therefore suggestive, but cannot confirm, that Ap<sub>4</sub>A is detectable in the KBM7 cells. If the other Ap<sub>4</sub>Ns are present in KBM7 cells, their intracellular concentration is not high enough to be detected using this method (Supplementary Figure S4.2).

#### **4.5 NuKO (*NUDT2* knock out) cell extracts**

##### **4.5.1 Identification of Ap<sub>4</sub>Ns in NuKO cells**

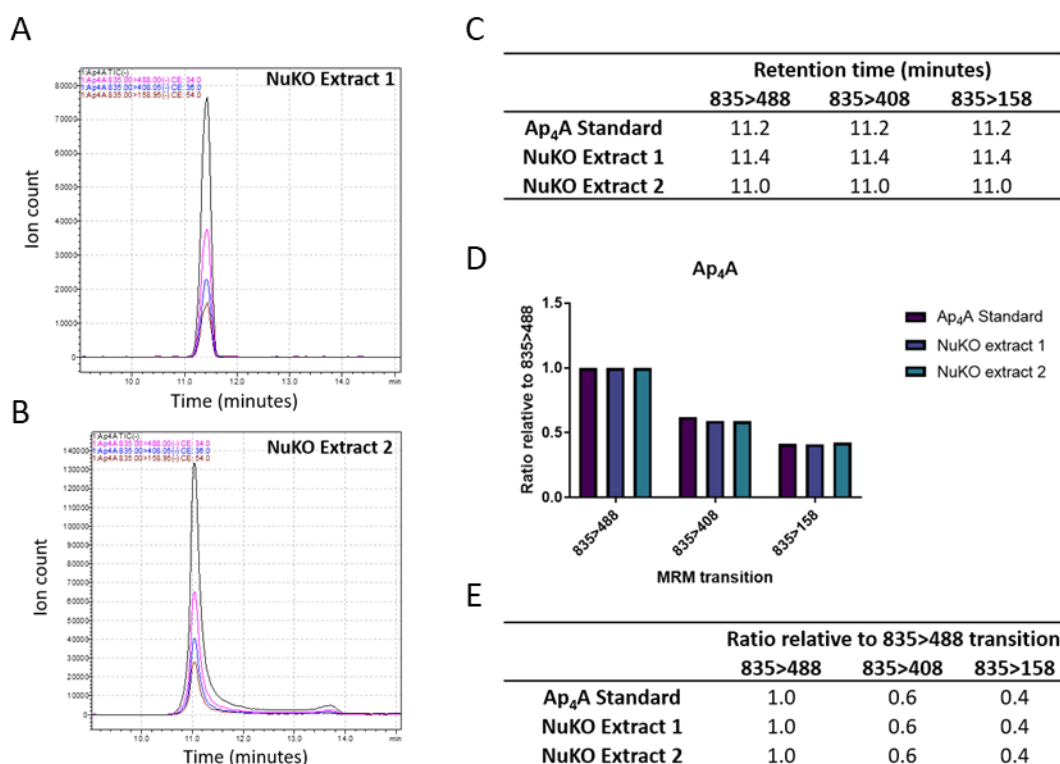
To determine if Ap<sub>4</sub>Ns are present in mammalian cells, metabolites were harvested from the Nudix hydrolase 2 knockout cells. The *Nudt2* gene has been knocked out in these cells, resulting in an increase in Ap<sub>4</sub>A, as the nudix hydrolase is the principal mechanism responsible for breaking down Ap<sub>4</sub>Ns. This results in a 175-fold increase in intracellular Ap<sub>4</sub>A (and potentially Ap<sub>4</sub>N) concentration compared with the parental KBM7 cell line (Marriott et al., 2016). We therefore deduced that if the Ap<sub>4</sub>Ns are present in KBM7 cells, there may be an increase in Ap<sub>4</sub>Ns in the NuKO cells due to the lack of Ap<sub>4</sub>A hydrolase activity. NuKO cellular extracts were prepared in the same way as described previously (Section 4.2.2). Since the Ap<sub>4</sub>N concentration should be much higher in these cells, we performed the extraction without the addition of a clean-up step, to avoid any loss of the dinucleotide tetraphosphates. Two samples of the NuKO extract sample were harvested separately and analysed for the presence of Ap<sub>4</sub>Ns (Fig. 4.11).



**Figure 4.11: Identification of Ap<sub>4</sub>Ns in NuKO cells.** Metabolites were extracted from NuKO cells, lyophilised and resuspended in water, then analysed directly on the LC-MS. **A, B:** Peaks for total ion count (TIC) and individual MRM transitions were identified for each of the Ap<sub>4</sub>Ns in two sets of NuKO cell extracts. **C, D:** Total ion counts (TICs) for the individual Ap<sub>4</sub>Ns in each NuKO extract 1 (C) and NuKO extract 2 (D) are shown in their own scale.

Initial analysis of the Ap<sub>4</sub>Ns in NuKO cells revealed peaks corresponding to Ap<sub>4</sub>A, Ap<sub>4</sub>C, and Ap<sub>4</sub>U in the mammalian knockout cell line (Fig. 4.11 A), with subsequent analysis also identifying an Ap<sub>4</sub>G peak in the extract (Fig. 4.11 B). Ap<sub>4</sub>A constituted the largest peak in both samples, with the peak area for the TIC reaching nearly  $1 \times 10^6$  units in Extract 1 (Fig. 4.11 C) and just over  $2 \times 10^6$  units in Extract 2 (Fig. 4.11 D). The Ap<sub>4</sub>U peak was the next largest, with TIC peak areas of  $2.5 \times 10^5$  and  $5.4 \times 10^5$  units (Fig. 4.11 C, D). The Ap<sub>4</sub>C peak reached  $\sim 1.6 \times 10^4$  and  $6.6 \times 10^4$  units in each extract (Fig. 4.11 C, D), while the Ap<sub>4</sub>G reached a TIC peak area of approximately  $6.7 \times 10^4$  units in NuKO extract 2, despite being absent in Extract 1 (Fig. 4.11 C and D). Accurate quantitation of each molecule is not possible using this approach, as the efficiency with which each molecule fragments will differ according to its structure. However, the peak intensity may give an indication of the relative abundance of each molecule in relation to the others and is discussed in Chapter 5. The difference in the peak intensities for all the Ap<sub>4</sub>Ns is likely due to a problem with the efficiency of the extraction process during the first harvesting experiment, as the peak volumes for all Ap<sub>4</sub>Ns are lower in NuKO extract 1 compared with NuKO extract 2.

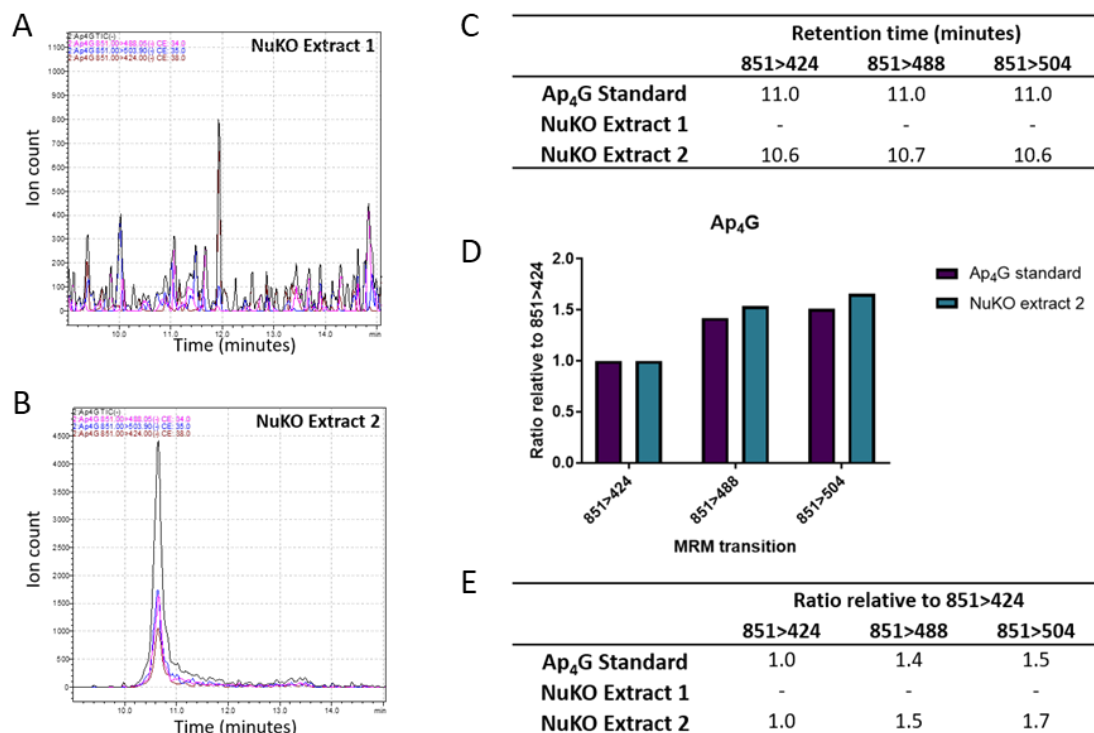
This approach has for the first time revealed that each of the ribosyl nucleoside tetraphosphates are produced in mammalian cells. However, for accurate quantification of these nucleotides the method would need to be developed further to include a standard curve for each molecule, and an internal standard to ensure that the measurements were consistent. To verify the identity of each of the Ap<sub>4</sub>Ns identified in the NuKO extract, we investigated firstly whether the candidate Ap<sub>4</sub>N peaks eluted at the appropriate retention time chromatographically (Fig. 4.12–4.16). Another method for verifying identity is to confirm whether all three of the pre-determined MRM transitions are present for each molecule in the sample. The ratio of the three MRM transitions could be used as a further measure, by comparing the ratios in the standard and the ratios of the transitions in the extract.



**Figure 4.12: Identification of Ap<sub>4</sub>A in NuKO cell extract.** NuKO cell extract was harvested on two separate occasions and injected for analysis by LC-MS **A, B**: The traces for all Ap<sub>4</sub>A transitions, including TIC, in NuKO extract 1 (**A**) and NuKO extract 2 (**B**). **C**: The retention time and **D, E**: the ratios of each transition in relation to one another were recorded and compared with the Ap<sub>4</sub>A standard.

All three Ap<sub>4</sub>A transitions were present in both samples of the NuKO extract (Fig. 4.12 A, B). The retention times for NuKO extract samples varied slightly, but each was within 0.2-minutes of the Ap<sub>4</sub>A standard (Fig. 4.12 C). A final measure used for the identification of molecules by triple quad mass spectrometry is to compare the ratios of each of the MRM transitions. The ratios of each transition in relation to one another were the same for the Ap<sub>4</sub>A standard and both NuKO extract samples (Fig. 4.12 D, E). Taken together, these data confirm the presence of Ap<sub>4</sub>A in the NuKO cell extracts.

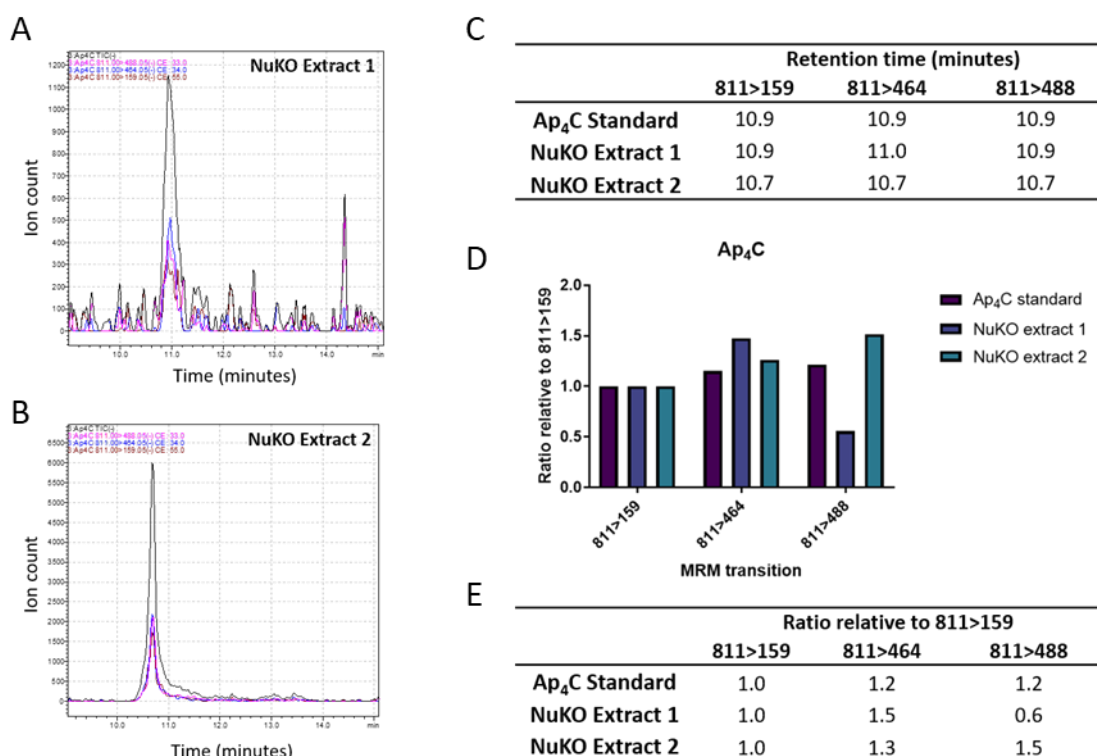
Similarly, we wanted to confirm the identity of the peaks suggesting the presence of the other Ap<sub>4</sub>Ns in the NuKO extract. Verification of the identity of these Ap<sub>4</sub>Ns is demonstrated in Figure 4.13 to 4.15.



**Figure 4.13: Identification of Ap<sub>4</sub>G in NuKO cell extract.** NuKO cell extract was harvested on two separate occasions and injected for analysis by LC-MS **A**, **B**: The traces for all Ap<sub>4</sub>G transitions, including TIC, in NuKO extract 1 (**A**) and NuKO extract 2 (**B**). **C**: The retention time and **D**, **E**: the ratios of each transition in relation to one another were recorded and compared with the Ap<sub>4</sub>G standard.

Ap<sub>4</sub>G was not present at high enough levels to be detected in NuKO extract 1 (Fig. 4.13 A); however, all three transitions were present in NuKO extract 2 (Fig. 4.13 B). In NuKO extract 2, the expected Ap<sub>4</sub>G transitions eluted at 10.6 and 10.7 minutes, compared with the 11.0-minute retention time in the Ap<sub>4</sub>G standard (Fig. 4.13 B). This shift to a slightly earlier retention time is consistent with the slightly earlier retention time seen for Ap<sub>4</sub>A in NuKO extract 2 (Fig. 4.12 C) and is likely due to small changes in salt concentration or the pH of the running buffer. The ratios between transitions in NuKO extract 2 were comparable to the ratios between transitions in the Ap<sub>4</sub>G standard (Fig. 4.13 E). Taken together, these data indicate that Ap<sub>4</sub>G is also present intracellularly in NuKO cells.

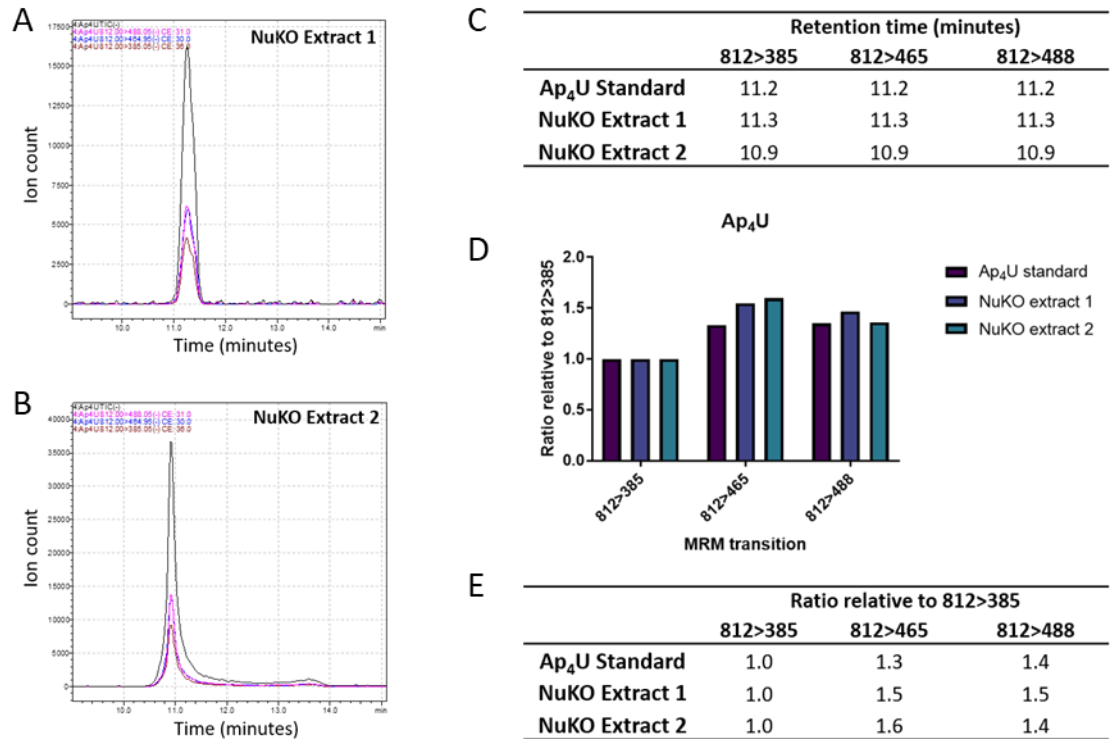




**Figure 4.14: Identification of Ap<sub>4</sub>C in NuKO cell extract.** NuKO cell extract was harvested on two separate occasions and injected for analysis by LC-MS **A, B**: The traces for all Ap<sub>4</sub>C transitions, including TIC, in NuKO extract 1 (**A**) and NuKO extract 2 (**B**). **C**: The retention time and **D, E**: the ratios of each transition in relation to one another were recorded and compared with the Ap<sub>4</sub>C standard.

Three MRM transitions were identified corresponding to Ap<sub>4</sub>C in both sets of NuKO extract (Fig. 4.14 A, B). Furthermore, the retention time for the Ap<sub>4</sub>C peak in each extract are within 0.2-minutes of the Ap<sub>4</sub>C standard (Fig. 4.14 C). Like the Ap<sub>4</sub>A and Ap<sub>4</sub>G samples, the elution time for Ap<sub>4</sub>C in NuKO extract 2 is slightly earlier than in the standard (Fig. 4.14 C). As NuKO extract 2 was analysed for all Ap<sub>4</sub>Ns simultaneously, this supports the hypothesis that a small change in buffer composition may be responsible for this change in the retention time for all Ap<sub>4</sub>Ns. Alternatively, it is possible that column performance may be degrading over time, allowing the Ap<sub>4</sub>Ns to elute more readily. Finally, the ratios of the MRM transitions in NuKO extract 2 are comparable to the Ap<sub>4</sub>C standard (Fig. 4.14 D, E). However, the ratio of the transitions in NuKO extract 1 differ from the standard and NuKO extract 2 ratios (Fig. 4.14 D, E). As the peak in NuKO extract 1 is much smaller than NuKO extract 2, it is likely that

machine sensitivity limit could have affected the detection of the MRM transitions in this sample.



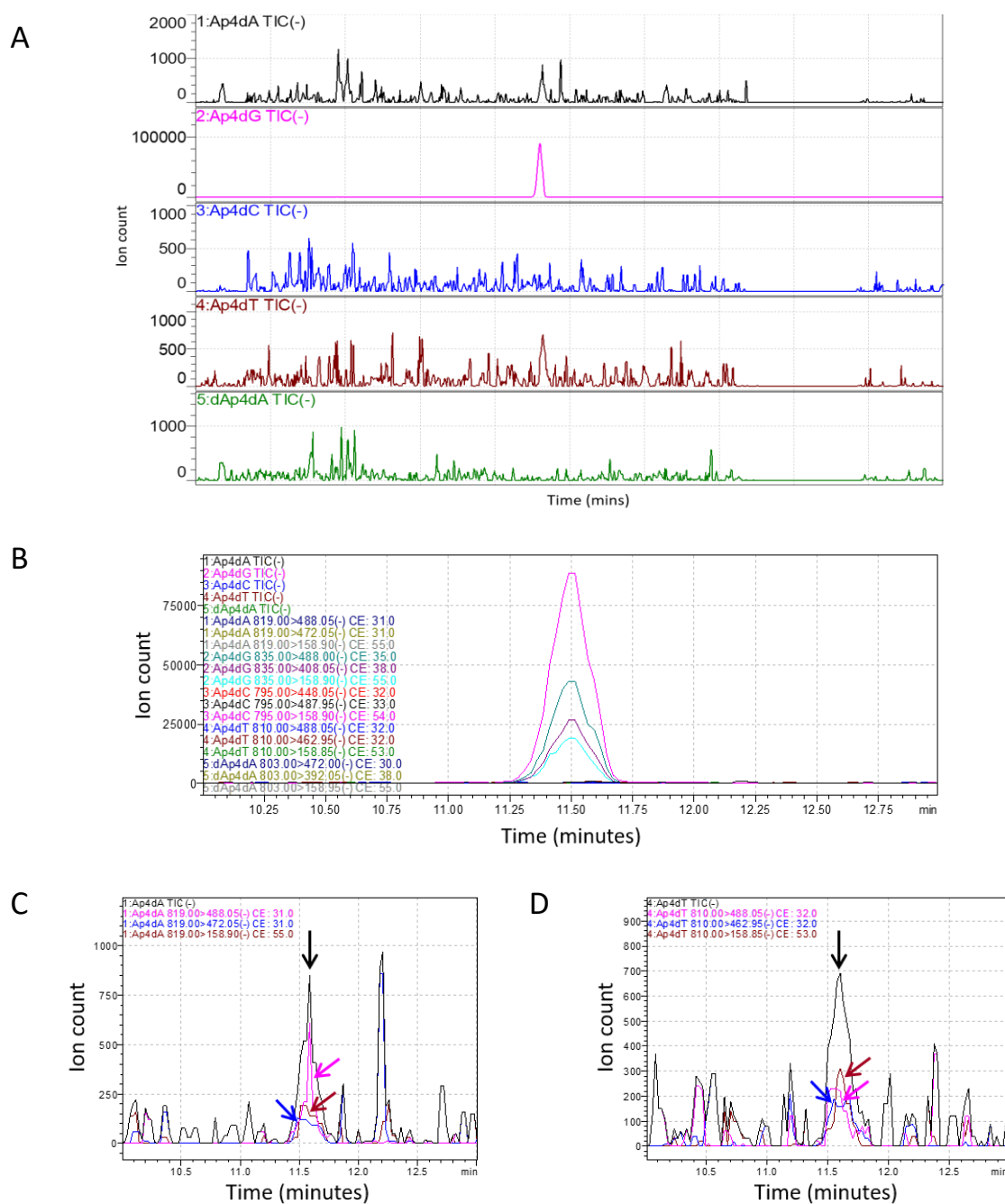
**Figure 4.15: Identification of Ap<sub>4</sub>U in NuKO cell extract.** NuKO cell extract was harvested on two separate occasions and injected for analysis by LC-MS **A, B**: The traces for all Ap<sub>4</sub>U transitions, including TIC, in NuKO extract 1 (**A**) and NuKO extract 2 (**B**). **C**: The retention time and **D, E**: the ratios of each transition in relation to one another were recorded and compared with the Ap<sub>4</sub>U standard.

Ap<sub>4</sub>U peaks composed of all three MRM transitions are present in both NuKO extracts (Fig. 4.15 **A, B**). Like the other Ap<sub>4</sub>Ns, the retention times for these peaks are comparable to the Ap<sub>4</sub>U standard, eluting up to a maximum of 0.3-minutes earlier in NuKO extract 2 Fig. 4.15 **C**). The ratio of each MRM transition compared with the others varies slightly but shows a similar pattern in general (Fig. 4.15 **D, E**). Taken together these data demonstrate that Ap<sub>4</sub>A, Ap<sub>4</sub>G, Ap<sub>4</sub>C and Ap<sub>4</sub>U are all present intracellularly in NuKO cells.

#### 4.5.2 Identification of Ap<sub>4</sub>dNs in NuKO cells

Having determined that all four of the Ap<sub>4</sub>Ns could be detected in NuKO cells, we wanted to identify whether any of the Ap<sub>4</sub>dNs could also be present. NuKO extract 1

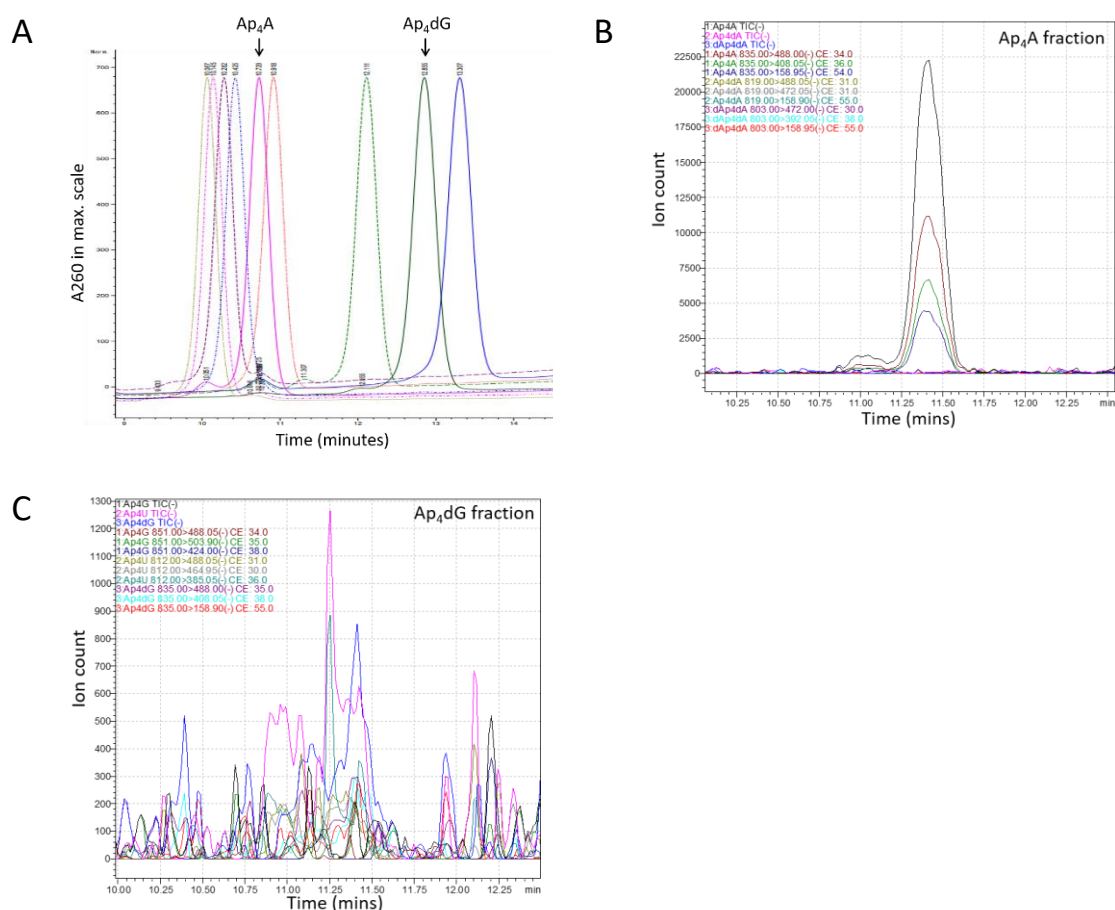
was injected for a second time, with the LC-MS set up to detect the transitions for each of the Ap<sub>4</sub>dNs (Fig. 4.16).



**Figure 4.16: Identification of several Ap<sub>4</sub>dNs in NuKO mammalian cell extract.** Metabolites were extracted from NuKO cells, lyophilised and resuspended in water, then analysed directly on the LC-MS. **A:** Peaks for total ion count (TIC) were identified for several of the Ap<sub>4</sub>dNs in a NuKO cell extract. **B–D:** The total ion counts (TICs) and transitions for Ap<sub>4</sub>dG (B), Ap<sub>4</sub>dA (C), and Ap<sub>4</sub>dT (D) found in the NuKO extract are shown in their own scale.

Small peaks composed of all three transitions were present for Ap<sub>4</sub>dA (Fig. 4.16 A, C) and Ap<sub>4</sub>dT (Fig. 4.16A, D). However, the area under the peaks for the individual transitions was too low to be quantified (Fig. 4.16 C and D). A large peak corresponding to Ap<sub>4</sub>dG was present (Fig. 4.16 A, B). However, as determined in section 4.3.1, this large Ap<sub>4</sub>dG peak could be a false positive as any Ap<sub>4</sub>A in the sample would trigger the same peaks to form. Historically, most of the literature has focussed on Ap<sub>4</sub>A as this is the most abundant dinucleotide tetraphosphate, so it is likely that it is Ap<sub>4</sub>A, rather than Ap<sub>4</sub>dG, that is causing this peak.

To confirm this, before analysis by LC-MS, the NuKO extract was first separated by HPLC using a ResourceQ column, into fractions that would contain either Ap<sub>4</sub>A or Ap<sub>4</sub>dG, if present. As this column separates according to charge, Ap<sub>4</sub>dG elutes later than Ap<sub>4</sub>A because of its more negatively charged structure (Fig. 4.17 A). The NuKO extract was separated by HPLC into an Ap<sub>4</sub>A-containing and Ap<sub>4</sub>dG-containing fraction, then each sample was analysed by LC-MS (Fig. 4.17, B and C).



**Figure 4.17: Analysis of fractionated NuKO extract.** **A:** Standards were run for all Ap<sub>4</sub>Ns and their retention time determined. Metabolite containing extract was harvested from NuKO cells using a chloroform extraction method and lyophilised overnight. **B,C:** The sample was resuspended in water and run through a ResourceQ column. Fractions were collected between 9.5 and 11.5 mins (Ap<sub>4</sub>A fraction) and 11.5–14 (Ap<sub>4</sub>dG fraction) minutes, lyophilised, then each fraction sample was solubilised in water and injected into the LC-MS.

The Ap<sub>4</sub>A and Ap<sub>4</sub>dG HPLC elution times were distinct from one another, with the Ap<sub>4</sub>A standard eluting at 10.7 minutes and Ap<sub>4</sub>dG standard eluting at approximately 12.9 minutes (Fig. 4.17 A). When separated, a large peak was present in the Ap<sub>4</sub>A-containing fraction of the extract (Fig. 4.17 B) and not in the Ap<sub>4</sub>dG fraction (Fig. 4.17 C). This fractionation process resulted in loss of Ap<sub>4</sub>Ns from the extract, so could not be used in general practice for separation of the Ap<sub>4</sub>Ns before analysis. However, separation of Ap<sub>4</sub>A and Ap<sub>4</sub>dG in the NuKO extract successfully demonstrated that it was Ap<sub>4</sub>A, and not Ap<sub>4</sub>dG that was the major contributor to the peaks identified in the NuKO cells.

## 4.6 Chapter discussion

### 4.6.1 Ap<sub>4</sub>Ns can be harvested from a mammalian cell line by chloroform extraction

In this chapter we have demonstrated that adenosine-containing dinucleotide tetraphosphate metabolites can be extracted from a human chronic myeloid leukaemia cell line following a 1:1 chloroform:water extraction method. A variety of different methods for metabolite extraction have been reported across different cell types, usually involving some combination of methanol, acetonitrile, chloroform, or water. Use of 60% methanol in the quenching of mammalian cells has been linked to the leakage of intracellular metabolites (Dietmair et al., 2010). However, a later report contended this, showing no substantial leakage for the metabolites tested, with the exception of malate, when quenched with 60% methanol followed by two 100% methanol extractions and one water extraction (Sellick et al., 2011). However, these methods described for the extraction of metabolites from mammalian cells generally discuss molecules such as various nucleotides or amino acids which are present at higher levels in the cell, and not the extraction of less abundant alarmones such as Ap<sub>4</sub>A.

Similarly, in *E. coli*, an acidic 80:20 acetonitrile:water system, or a 40:40:20 acetonitrile:methanol:water system were found to be the best options for the extraction of triphosphate molecules (Rabinowitz and Kimball, 2007). However, in *S. cerevisiae* an acidic acetonitrile-methanol system was not as successful as other methods such as boiling ethanol or a methanol/chloroform mixture in the recovery of metabolites (including several amino acids, AMP, ADP and ATP) (Canelas et al., 2009). A 40:40:20 mixture of acetonitrile:methanol:water was also used for the extraction of nucleotides such as GMP, GDP and GTP in bacteria, however this method was not suitable for the extraction of the (p)ppGpp alarmone to detectable levels (Liu et al., 2015). A modified 1:1 chloroform:water extraction procedure with improved recovery was later used by the authors, which resulted in the identification of (p)ppGpp and Ap<sub>4</sub>A in *B. subtilis* cells (Fig. 4.2 A) (Fung et al., 2020). As this method had a demonstrated capacity to extract Ap<sub>4</sub>A from bacterial cells, we chose to adapt this method for use in the mammalian suspension cell line (Fig. 4.2 B).

This method was well-suited to adaptation for the extraction of metabolites from KBM7/NuKO cells as these also grow in suspension so there is no need for detachment as would be the case with an adherent cell line. Indeed, the data show that each of the Ap<sub>4</sub>Ns were successfully extracted from the NuKO cell line using this method.

#### **4.6.2 Identification of Ap<sub>4</sub>N candidates in NuKO extracts by HPLC**

Candidate Ap<sub>4</sub>A and Ap<sub>4</sub>U peaks were identified in the NuKO extract by HPLC analysis, which provided the foundation for investigating the cell extracts more thoroughly by LC-MS. Though not as sensitive as the LC-MS technique employed by Fung and colleagues, the HPLC analysis provided a useful first step to determine whether any peaks were identifiable that could potentially represent the presence of the Ap<sub>4</sub>Ns within the mammalian cell extract. This was chosen because a suitable HPLC-method for identification of the Ap<sub>4</sub>Ns had already been set up in Chapter 3, and the expected retention times for the Ap<sub>4</sub>Ns at 260 nm and 280 nm had already been determined. This meant that, although not as sensitive, the HPLC setup was a relatively simple and low-cost method for the preliminary analysis.

Ap<sub>4</sub>A phosphatases have been used as part of a method to determine Ap<sub>4</sub>A concentration in previous experiments. In these experiments, the background ATP needs to be removed prior to hydrolysis of the Ap<sub>4</sub>A, so that the ATP formed during hydrolysis by Ap<sub>4</sub>A hydrolase is directly proportional to the Ap<sub>4</sub>A content in the cell. Use of a phosphatase was therefore an obvious choice for this purpose as it removes background levels of molecules with exposed phosphate groups such as ATP but has no effect on Ap<sub>4</sub>A (Marriott et al., 2015; Murphy et al., 2000).

#### **4.6.3 Identification of a candidate Ap<sub>4</sub>A peak in KBM7 cells by LC-MS**

A candidate Ap<sub>4</sub>A peak was successfully identified in KBM7 cell extract; however, levels of other Ap<sub>4</sub>Ns in this cell line were too low to detect. Identification of Ap<sub>4</sub>A was initially hindered by ion suppression due to the mixture of ions in the extracts isolated from the KBM7 cells. This meant that initially no Ap<sub>4</sub>A could be detected in the cells

despite previous literature quantifying Ap<sub>4</sub>A in the same cell line (Marriott et al., 2015). This ion suppression effect could have been exacerbated by the PBS wash step, as washing the cells has previously been associated with reduced metabolite signal intensities (Ser et al., 2015). However, this step was necessary to ensure that only intracellular Ap<sub>4</sub>Ns were present in the cell extract. Optimisation of the method to include a graphite column clean-up step prior to analysis resulted in a reduction in ion suppression and led to the successful identification of Ap<sub>4</sub>A in the KBM7 cell extract.

Ap<sub>4</sub>A was identified in the cell extract both by chromatographic elution time and by the detection of its expected MRM transitions in a triple quadrupole system. Prior to analysis of a complex mixture, MRM transitions are determined using a standard of a given molecule. When a complex mixture is injected, the first quadrupole functions as a filter, only allowing molecules through that have the same mass as the expected molecule. These molecules are then fragmented in the second quadrupole. Only fragments that have the same molecular weight as the fragments produced in the standard enter the third quadrupole, where they go on to be detected (Fig. 4.4). The technique using ion fragmentation for the quantitative analysis of small molecules was first used 45 years ago and these techniques are now well-established for the identification of target molecules from complex mixtures (Baty and Robinson, 1977). As we have chemically synthesised HRMS-verified standards for each of the Ap<sub>4</sub>Ns, this technique was ideal for the identification of these molecules in the complex mixture of metabolites extracted from the cells. As this method can identify low molecular weight molecules in a complex mixture, it prevented the need for time-consuming and complex purification procedures, as would be necessary for other methods such as HRMS.

Optimisation of the method and addition of a clean-up protocol successfully reduced ion suppression enough for a potential Ap<sub>4</sub>A peak to be identifiable in these cells (Fig. 4.10). However, despite the introduction of a cleanup step, unfortunately other Ap<sub>4</sub>N levels in this cell line were too low to be recognised by the LC-MS. As this was likely due to the sensitivity of the system, it was not a problem we could easily overcome. Nevertheless, the use the *Nudt2* knockout version of this cell line, which



has intrinsically higher Ap<sub>4</sub>N levels, facilitated identification of these molecules in mammalian cells using the available system.

#### **4.6.4 Ap<sub>4</sub>Ns and some Ap<sub>4</sub>dNs are detectable in NuKO cells**

Ap<sub>4</sub>A, Ap<sub>4</sub>U, Ap<sub>4</sub>C and Ap<sub>4</sub>G are all present in the mammalian NuKO cell line. Up to this point, only Ap<sub>4</sub>A has been specified in mammalian cells, even though the methods used for quantitation would often encompass other Ap<sub>4</sub>Ns in addition to Ap<sub>4</sub>A. As the MRM transitions are specific to the changes in the mass of the ions of each nucleotide, this approach enabled us to identify all four Ap<sub>4</sub>Ns as well as low amounts likely corresponding to Ap<sub>4</sub>dT and Ap<sub>4</sub>dA in the NuKO cells. As these molecules are present intracellularly, it opens the possibility that these molecules, like Ap<sub>4</sub>A, may have some purpose in mammalian cells.

#### **4.7 Conclusion**

In summary, we have adapted the technique used by Fung and colleagues in 2020 and applied it to the extraction of metabolites from a mammalian cell line. Further to this we have developed a technique for the accurate identification of the individual Ap<sub>4</sub>Ns in a mammalian NuKO cell extract, using retention time, three MRM transitions and their relative ratios as parameters for identification of each molecule. The sensitivity of the LC-MS limited the ability to detect Ap<sub>4</sub>Ns in the KBM7 cell line, despite the introduction of a clean-up step to limit ion suppression. However, the development of this technique in combination with the *Nudt2* knockout cell line (NuKO) has enabled us to identify Ap<sub>4</sub>A, Ap<sub>4</sub>G Ap<sub>4</sub>C and Ap<sub>4</sub>U, as well as peaks corresponding to low levels of Ap<sub>4</sub>dA and Ap<sub>4</sub>dT, intracellularly in this cell line. In the previous chapter we presented evidence to suggest that each of the Ap<sub>4</sub>Ns could be synthesised as a by-product of ubiquitin activation. This, combined with the evidence now shown demonstrating the existence of these molecules intracellularly opens the question of whether, like Ap<sub>4</sub>A, these Ap<sub>4</sub>Ns might also have a role in signalling in the intracellular environment.

## **Chapter 5: Investigating the stress response in cells with intrinsically high $\text{Ap}_4\text{N}$ levels compared with their parental cell line**

## 5.1 Chapter introduction

### 5.1.1 Ap<sub>4</sub>A and the cell cycle

A well-regulated cell cycle is integral to successful cell replication. The cell cycle begins with G1-phase, in which the cell prepares for DNA synthesis. From here, the cell must pass through the p53-dependent G1/S-checkpoint before it can move into S-phase, where the DNA is replicated. After DNA replication, the cell proceeds into the G2-phase, where it prepares for mitosis. If it passes the quality control checks at the G2/M-checkpoint, the cell will reach M-phase, where mitosis occurs and the cell duplicates (Vermeulen et al., 2003). Changes in cyclin level throughout the cell cycle are responsible for the activation of the different cyclin dependent kinases (CDKs) which are required in the different cell cycle phases. There have been conflicting reports surrounding the role of Ap<sub>4</sub>A and fluctuations in its concentration throughout the cell cycle. In *E. coli*, Ap<sub>4</sub>A has been shown to induce cell division, with high Ap<sub>4</sub>A levels promoting early cell division and low Ap<sub>4</sub>A levels resulting in a delay to cell division (Nishimura et al., 1997). In the acellular slime mould *P. polycephalum*, one group concluded that there was a rapid increase in Ap<sub>4</sub>A concentration when entering S-phase (Weinmann-Dorsch et al., 1984b). However, another group found no change in Ap<sub>4</sub>A or Ap<sub>4</sub>G level in the same organism (Garrison et al., 1986). Furthermore, contrary to the increase in Ap<sub>4</sub>A levels reported in *P. polycephalum* on S-phase entry, a study in *Strongylocentrotus nudus* (sea urchin) embryos instead demonstrated a decrease in Ap<sub>4</sub>A concentration prior to S-phase (Morioka and Shimada, 1985).

Several early reports suggested that Ap<sub>4</sub>A was important in stimulating DNA synthesis. This was demonstrated in baby hamster kidney cells, where addition of Ap<sub>4</sub>A to permeabilised cells caused the stimulation of discontinuous DNA synthesis in G1-arrested cells in a dose-dependent manner (Grummt, 1978b). Microinjection of Ap<sub>4</sub>A into *X. laevis* oocytes also stimulated DNA synthesis (Zourgui et al., 1984). Supporting these results is the identification of a 57 kDa subunit of the DNA polymerase  $\alpha$  protein complex, which functions as a highly specific Ap<sub>4</sub>A-binding partner (Grummt et al., 1979). However, more recently it has been demonstrated that Ap<sub>4</sub>A instead inhibits the initiation of DNA replication (Marriott et al., 2015). In this

paper, the authors demonstrated that Ap<sub>4</sub>A has this inhibitory effect on the initiation but not the elongation phase of DNA replication (Marriott et al., 2015).

### **5.1.2 Ap<sub>4</sub>A has been implicated in the cellular response to stress**

Different physiological stresses have been shown to result in the accumulation of Ap<sub>4</sub>Ns in a variety of species. One such type of stress is a change in temperature. Ap<sub>4</sub>A levels were increased in simian virus 40-transformed mouse 3T3 cells treated at an increased temperature of 45 °C (Baker and Jacobson, 1986). Heat-shock of *Drosophila* cells from 19 °C to 37 °C resulted in a 2.2-fold increase in Ap<sub>4</sub>A, but did not seem to affect Ap<sub>4</sub>G levels (Brevet et al., 1985a).

Oxidative damage also induces an accumulation of Ap<sub>4</sub>A and Ap<sub>4</sub>G, among other nucleotide polyphosphates, in *S. typhimurium* (Bochner et al., 1984). Antibiotics such as gentamicin (an aminoglycoside), ampicillin (a β-lactam) and norfloxacin (a fluoroquinolone) have all been shown to contribute to the production of hydrogen peroxide H<sub>2</sub>O<sub>2</sub> in *E. coli* (Dwyer et al., 2014). Recently, treatment with the aminoglycoside kanamycin was demonstrated to increase Ap<sub>4</sub>A levels in *E. coli* in a hydroxyl radical-dependent manner and improved the activity of the antibiotic (Ji et al., 2019). In HEK293T cells, treatment with H<sub>2</sub>O<sub>2</sub> to induce oxidative stress resulted in a significant increase in both Ap<sub>3</sub>A and Ap<sub>4</sub>A (Krüger et al., 2021). MMC is a chemotherapeutic cross-linking agent that induces an accumulation of Ap<sub>4</sub>A in several cell lines. AA8 cells, HeLa cells and MEF cells showed an increase in Ap<sub>4</sub>A level of between 7- and 9-fold after treatment with MMC (Marriott et al., 2015). This is supported by recent evidence which has demonstrated that MMC treatment results in increased Ap<sub>3</sub>A and Ap<sub>4</sub>A levels in HEK293T cells (Krüger et al., 2021).

### **5.1.3 Ap<sub>4</sub>A has been linked to apoptosis**

In addition to Ap<sub>4</sub>A levels increasing in response to cell stress, increased Ap<sub>4</sub>A levels have also been linked to apoptosis. Analysis of DNA fragmentation after Ap<sub>4</sub>A treatment showed that Ap<sub>4</sub>A, but not Ap<sub>3</sub>A, induced apoptosis in mouse VMRO and human HL60, U937 and Jurkat cells, but not in human HL299 cells, human NT2 cells, or mouse Swiss 3T3 cells (Vartanian et al., 1999). In this study, hydrolysis-resistant

analogues were unable to induce apoptosis in any of the cell lines, leading the authors to conclude that the induction of apoptosis required the hydrolysis of Ap<sub>4</sub>A (Vartanian et al., 1999). Alternatively, if the tetraphosphate moiety is important for Ap<sub>4</sub>A to induce apoptosis, the small change in the structure of the analogues could mean that they lack the necessary structure for inducing apoptosis. Consistent with a role for Ap<sub>4</sub>A in apoptosis, a more recent study found an up-regulation of several pro-apoptotic genes in NuKO cells compared with their parental cell line (Marriott et al., 2016). Furthermore, an isosteric Ap<sub>4</sub>A analogue was capable of inducing apoptosis in Fhit-positive but not Fhit-negative HEK293T cells, and has therefore been suggested as a potential anti-cancer therapy in Fhit-positive cancer cells (Krakowiak et al., 2011). However, in this case the analogue is somewhat different in structure from Ap<sub>4</sub>A and is much more stable in cell medium and cell extract, so may not be representative of the role of Ap<sub>4</sub>A in apoptosis. In contrast, in another study Ap<sub>4</sub>A was suggested to inhibit apoptosis through the prevention of cytochrome c translocation and activation of caspase-3 in a rat model for Parkinson's disease (Wang et al., 2003).

#### **5.1.4 Chapter Aims**

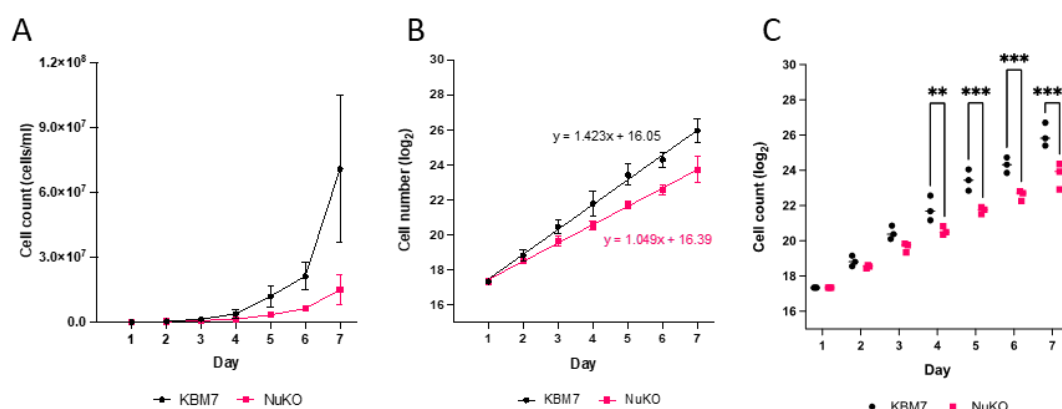
In the previous chapter, we demonstrated that Ap<sub>4</sub>A, Ap<sub>4</sub>C, Ap<sub>4</sub>G and Ap<sub>4</sub>U are all synthesised and present in a *Nudt2* knockout cell line derived from parental KBM7 cells, opening the possibility that they may have a role in cellular function. Previous studies have linked cell stress to an increase in Ap<sub>4</sub>A concentration in mammalian cells, or protective Ap<sub>4</sub>N-induced caps in bacteria (Luciano and Belasco, 2020; Luciano et al., 2019; Marriott et al., 2015). In the previous chapter we used a high-Ap<sub>4</sub>N-containing 'NuKO' cell-line to demonstrate that Ap<sub>4</sub>N molecules are present in mammalian cells. The NuKO cell line has intrinsically higher Ap<sub>4</sub>N levels than its parental KBM7 cell line and therefore provides a method of analysing how increased Ap<sub>4</sub>N levels might affect cellular function. Here, by comparing factors such as proliferation and cell cycle dynamics in the KBM7 and NuKO cell lines, we first aimed to investigate the ways in which an increased intracellular Ap<sub>4</sub>N concentration might affect cellular function in an unstressed cell population. We next wanted to assess whether having an already increased Ap<sub>4</sub>N concentration affected the cells' ability to respond to stress. Finally, we

aimed to determine whether applying several different types of stress influenced the intracellular concentration of each of the Ap<sub>4</sub>Ns.

## 5.2 Effect of increased Ap<sub>4</sub>N levels on cell duplication and cell cycle dynamics

### 5.2.1 Comparison of the proliferation rates of a high vs low Ap<sub>4</sub>N-containing cell line

To determine the effect of an increased intracellular Ap<sub>4</sub>N concentration on the normal function of unstressed cells, we used the KBM7 and NuKO cell lines discussed in the previous chapter. As alterations to the DNA replication and the cell cycle would affect proliferation rate, we first compared the duplication time for the KBM7 cell line with the NuKO, high Ap<sub>4</sub>N-containing, cell line. The number of cells for each cell line were counted every day over 7 days, then plotted on a log<sub>2</sub> scale to determine duplication time (Fig. 5.1).



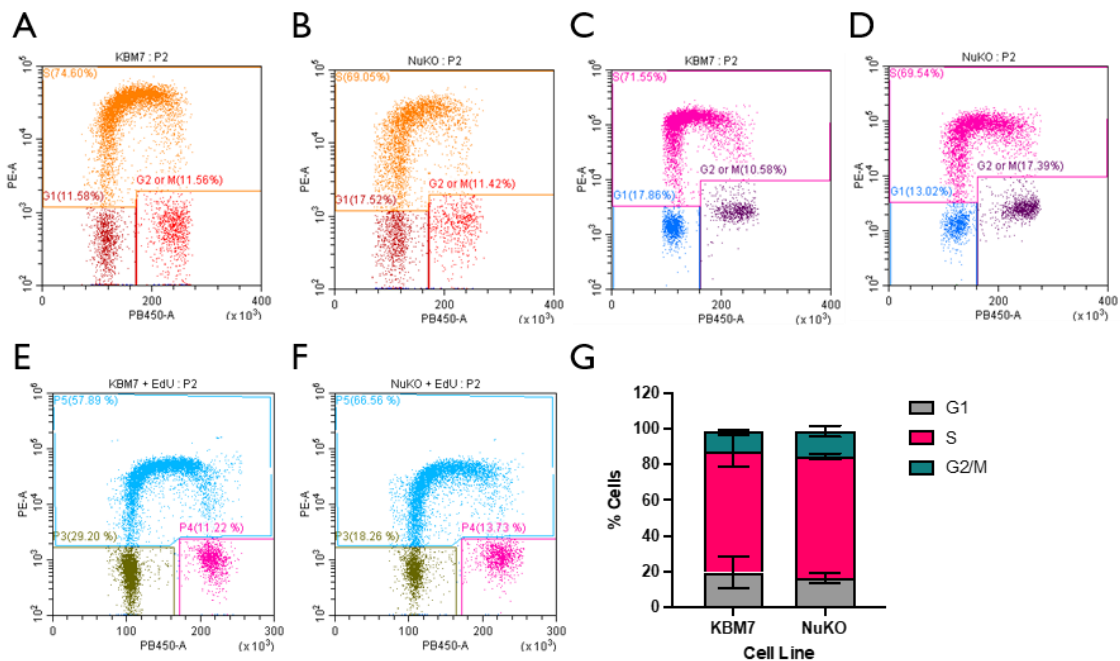
**Figure 5.1: Differences in duplication time of KBM7 and NuKO cells.** 2 million cells of each cell line were seeded, then counted daily. **A:** the raw cell number was plotted and **B, C:** the log<sub>2</sub> of the cell number was plotted to calculate duplication time and for statistical analysis using a two-way ANOVA with Šidák's multiple comparisons test (Day 4,  $p=0.0086$ ; day 5,  $p=0.0002$ ; day 6,  $p=0.0002$ ; day 7,  $p<0.0001$ ).

Over time the cell counts revealed a difference in rate of cell duplication (Fig. 5.1 A). These data were plotted as a log<sub>2</sub> function, and the gradient was used to determine duplication time (Fig. 5.1 B). These data indicate that KBM7 cells replicate at a faster rate than their *NUDT2* knockout counterparts, with a mean duplication time of approximately 16.9 hours compared with 22.9 hours in the NuKO cells. A two-way ANOVA using Šidák's multiple comparisons test demonstrated that the difference

between the cell counts for the KBM7 and NuKO cells was statistically significant from day 4 onwards (Fig. 5.1 C).

### 5.2.2 Differences in cell cycle kinetics

For cells to successfully divide they pass through four stages of the cell cycle: G1, S, G2 and M phases, and pass through several cell cycle checkpoints. As the KBM7 cells replicate faster than the high-Ap<sub>4</sub>A containing NuKO cells, the cell cycle profile of each cell line was investigated to determine whether NuKO cells were arrested at a specific point in the cell cycle. To further investigate this, unstressed cells were labelled with EdU and Hoescht and analysed by flow cytometry to determine the proportion of cells in each stage of the cell cycle (Fig. 5.2).



**Figure 5.2: Comparing the cell cycle dynamics of KBM7 and NuKO cells.** Flow Cytometry data showing the proportion of cells in G1, S and G2/M phase in KBM7 cells compared with NuKO cells. **A, B:** Repeat 1; **C, D:** repeat 2; **E, F:** repeat 3. **G:** A summary of n=3 repeats showing the approximate percentage of cells in each cell cycle phase. Gating for repeats 1 and 2 is shown in Supplementary Figure 5.1. Please note that the gating for repeat 3 is unavailable due to corruption of the file.

The flow cytometry data show no clear differences in G1, S and G2 distribution between the two cell lines (Fig. 5.2). When taken together with the proliferation rates in these cells, these data suggest that the KBM7s spend approximately 2.5 hours in G1-phase, 12.3 hours in S-phase and 1.9 hours in G2/M-phase, while the NuKO cells spend approximately 3.5 hours, 15.9 hours, and 3.3 hours in G1, S and G2/M-phase, respectively. Interestingly, the G1 and S populations appear to be more distinct in the KBM7 cells compared with the NuKO cell line for all repeats (Fig. 5.2 A, C and E versus B, D and F). This could be due to a delay at some point in the G1/S transition causing the cells to take longer to progress fully into S-phase.

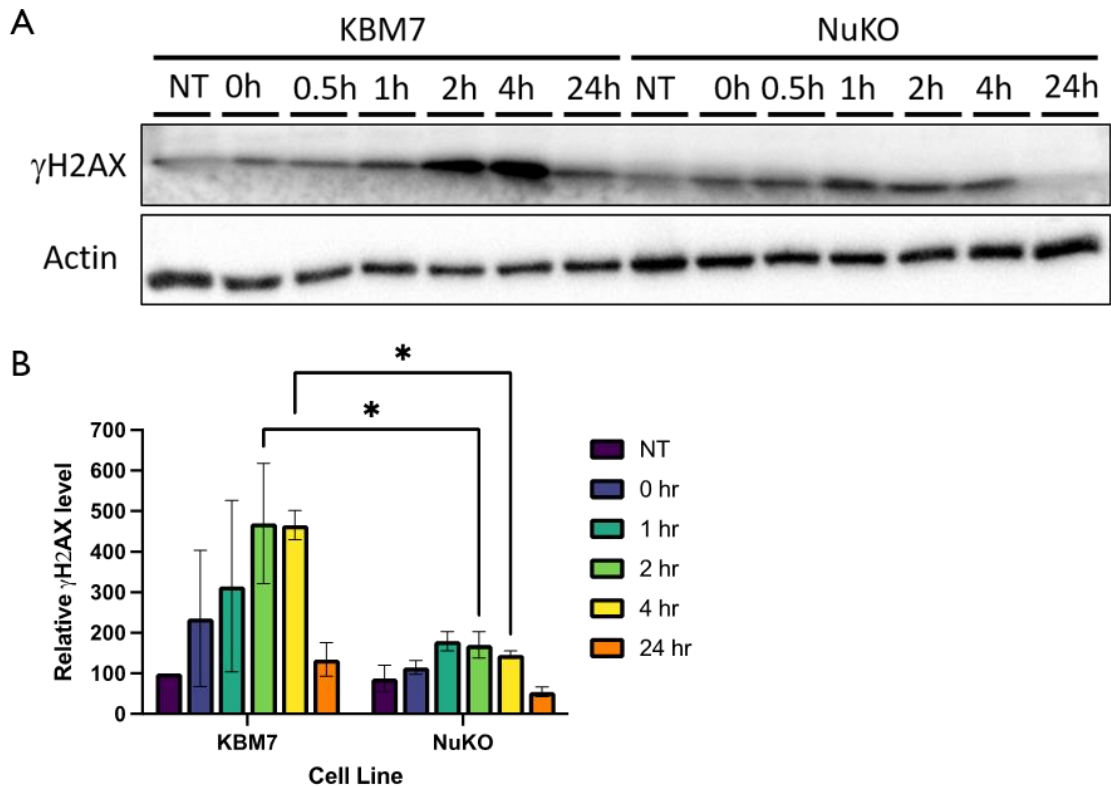
### **5.3 Comparison of the KBM7 and NuKO response to stress**

#### **5.3.1 The $\gamma$ H2AX response to stress**

Throughout multiple domains of life, Ap<sub>4</sub>A levels have been shown to increase in response to stress. This led to the question of whether a cell line with intrinsically high Ap<sub>4</sub>A levels would elicit the same response to cell stress. HU slows down DNA replication forks by inhibiting ribonucleotide reductase, causing a reduction in the levels of deoxyribonucleoside triphosphates (Singh and Xu, 2016). If the problem is not resolved then stalled replication forks can collapse, causing DNA damage (Singh and Xu, 2016).

In response to DNA damage and replication stress, the DNA damage response is activated.  $\gamma$ H2AX functions as an early marker of the DNA damage response. In the presence of double strand breaks, H2AX is phosphorylated at its serine-139 residue by ATM to form  $\gamma$ H2AX (Burma et al., 2001). Replication stress induced by hydroxyurea also induces the phosphorylation and foci-formation of H2AX in an ATR-dependent manner (Ward and Chen, 2001). Therefore,  $\gamma$ H2AX is useful as a marker of the DNA damage response. To determine whether the increased intracellular concentration of Ap<sub>4</sub>A in the NuKO cells impacts their ability to elicit a DNA damage response to stress, we treated both the KBM7 and NuKO cell lines with HU for 2 hours then allowed them to recover over a 24-hour period. Samples were taken at intervals over this time-course, then separated by Western blotting and probed for  $\gamma$ H2AX (Fig. 5.3).



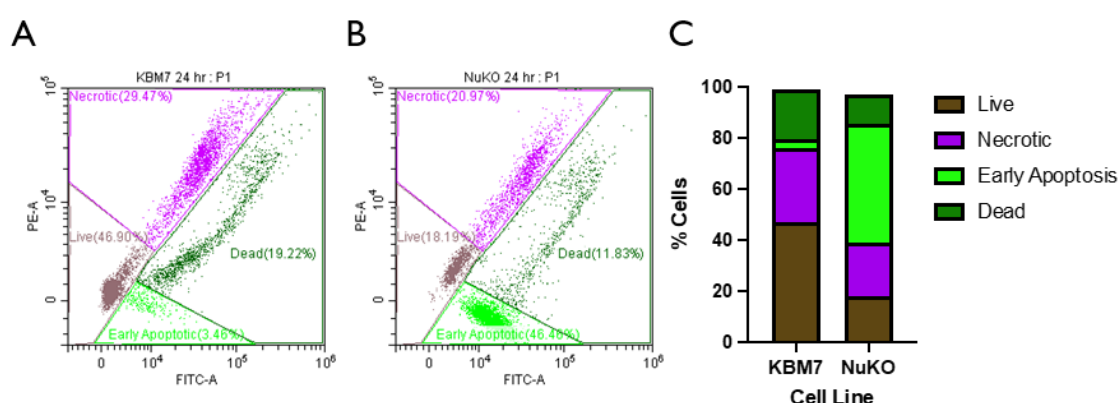


**Figure 5.3: Investigating the  $\gamma$ H2AX response to replication stress in KBM7 and NuKO cells.** KBM7 and NuKO cells were each treated with 1 mM HU for two hours, then allowed to recover for between 0 and 24 hours and analysed for differences in  $\gamma$ H2AX level. **A:** Representative Western blot showing changes in  $\gamma$ H2AX level. **B:** Summary data for n=2 Western blot repeats with the data scaled so that the untreated KBM7 level is equal to 100 and changes are shown relative to this. Statistical analysis was performed using a two-way ANOVA with Šídák's multiple comparisons post-hoc test.

Western blotting of the  $\gamma$ H2AX level after treatment with HU showed that KBM7 cells had maximal  $\gamma$ H2AX levels 2–4 hours after recovery started (Fig. 5.3). NuKO cells showed maximal  $\gamma$ H2AX levels after 1–2 hours; however, the mean  $\gamma$ H2AX level at the peak response time in the NuKO cells was 2.6 times lower than peak  $\gamma$ H2AX response in the KBM7 cells. The response of the NuKO cell line was also less varied between the two repeats. The  $\gamma$ H2AX level in both cell lines returned to within baseline levels after 24 hours.

### 5.3.2 The apoptotic response to stress

Increased Ap<sub>4</sub>A levels have been linked to increased apoptosis in response to cell stress in several cell lines (Vartanian et al., 1999). Furthermore, when the KBM7 and NuKO cell lines were previously analysed for changes to gene expression, a number of pro-apoptotic genes were up-regulated in the latter (Marriott et al., 2016). Induction of apoptosis was therefore measured in the KBM7 and NuKO cell lines to determine whether up-regulation of pro-apoptotic genes translated to increased apoptotic response to cell stress. As in the previous section, each cell line was treated with hydroxyurea for 2 hours then allowed to recover for 24 hours in fresh medium, then labelled with YO-PRO®-1 and PI. Apoptotic cells are permeant to YO-PRO®-1, a green, fluorescent carbocyanine nucleic acid stain which does not stain live cells (Allen et al., 2007; ThermoFisher Scientific, accessed 06 February 2023). On the other hand, PI does not stain live cells or early apoptotic cells (Allen and Davies, 2007). Necrotic cells are stained by both PI and YO-PRO®-1; however, fluorescence intensity is reduced compared with dead cells as the fluorescence is degraded by DNA (Allen and Davies, 2007). After 24 hours, apoptosis was measured by flow cytometry (Figure 5.4).



**Figure 5.4: Investigating the apoptotic response to replication stress in KBM7 and NuKO cells.** NuKO cells were each treated with 1 mM HU for two hours, then allowed to recover for 24 hours. Apoptotic response was measured using YO-PRO1 and compared between the two cell lines to distinguish between live, dead, necrotic and early apoptotic cells. **A:** KBM7 cells; **B:** NuKO cells. **C:** A summary of the proportion of cells that were live, necrotic, early apoptotic, or dead.

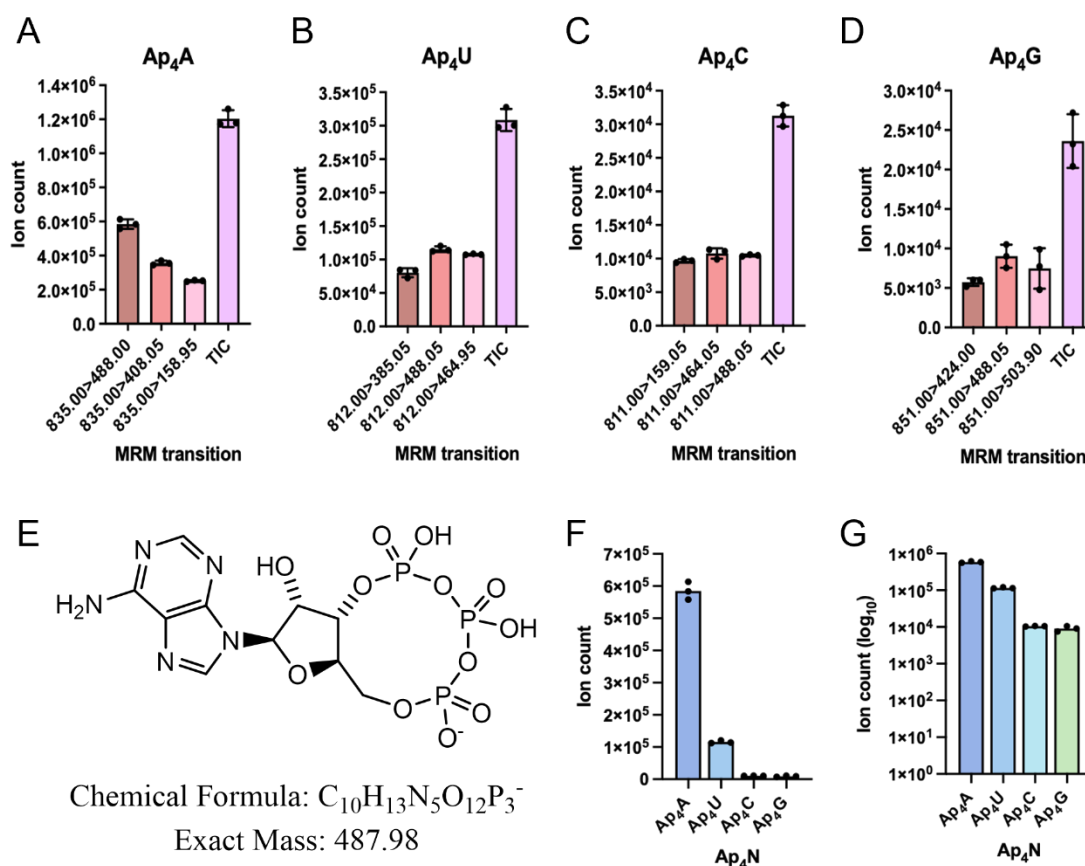
The NuKO cells had a 13-fold increase in the percentage of cells in early apoptosis (46.46%) (Fig. 5.4 B and C) compared with the KBM7 cells (3.46%) (Fig. 5.4 A and C). Likewise, a much greater proportion of the KBM7 cells remained alive (46.90%) (Fig. 5.4 A and C) compared with the number of living cells in the NuKO sample (18.19%) (Fig. 5.4 B and C). However, the percentage of dead cells was also higher in the KBM7s, with 19.22% of cells being dead (Fig. 5.4 A and C) compared with 11.83% in the NuKO cells (Fig. 5.4 B and C). The percentage of necrotic cells was slightly higher in the KBM7 cell line, with 29.47% of cells being necrotic (Fig. 5.4 A and C) compared with 20.97% in the NuKO cell line (Fig. 5.4 B and C). In summary, the primary differences between the two cell lines were that the NuKO cells showed an increased number of cells in early apoptosis and a reduced number of live cells compared with the parental cell line (Fig. 5.4 C).

## **5.4 Investigating the Ap<sub>4</sub>N response to stress in NuKO cells**

### **5.4.1 Determining the variation in area under peak readings taken by the LC-MS**

In the previous chapter we identified each of the Ap<sub>4</sub>Ns in the NuKO cell extract. Unfortunately, despite attempting to resolve Ap<sub>4</sub>Ns in the KBM7 extract using clean-up techniques, it was not possible to identify the Ap<sub>4</sub>Ns in this extract. Therefore, for the remaining analysis in this chapter, the NuKO cell line was used. However, development of the method for detection of the KBM7s remains an important next step.

To enable us to accurately measure any changes that occurred in Ap<sub>4</sub>N level after exposure to stress, we first needed to determine how variable the measurements were for Ap<sub>4</sub>N levels in separate samples. To do this, three separate extracts were prepared simultaneously from the NuKO cells and the area under each Ap<sub>4</sub>N peak was measured by LC-MS (Fig. 5.5).



**Figure 5.5: Variability in LC-MS analysis outputs and an approximate comparison of Ap<sub>4</sub>N abundance.** Extract samples were harvested simultaneously from NuKO cells, then analysed by LC-MS. Variation in the area under the peaks corresponding to **A**: Ap<sub>4</sub>A, **B**: Ap<sub>4</sub>U, **C**: Ap<sub>4</sub>C and **D**: Ap<sub>4</sub>G was recorded. **E**: The predicted structure for the ion fragment corresponding to the molecular mass of 488.05, found in all four Ap<sub>4</sub>Ns. **F**, **G**: A linear (F) and log<sub>10</sub> (G) comparison of the abundance of the 488.05 fragment as an approximate measure of relative Ap<sub>4</sub>N concentration.

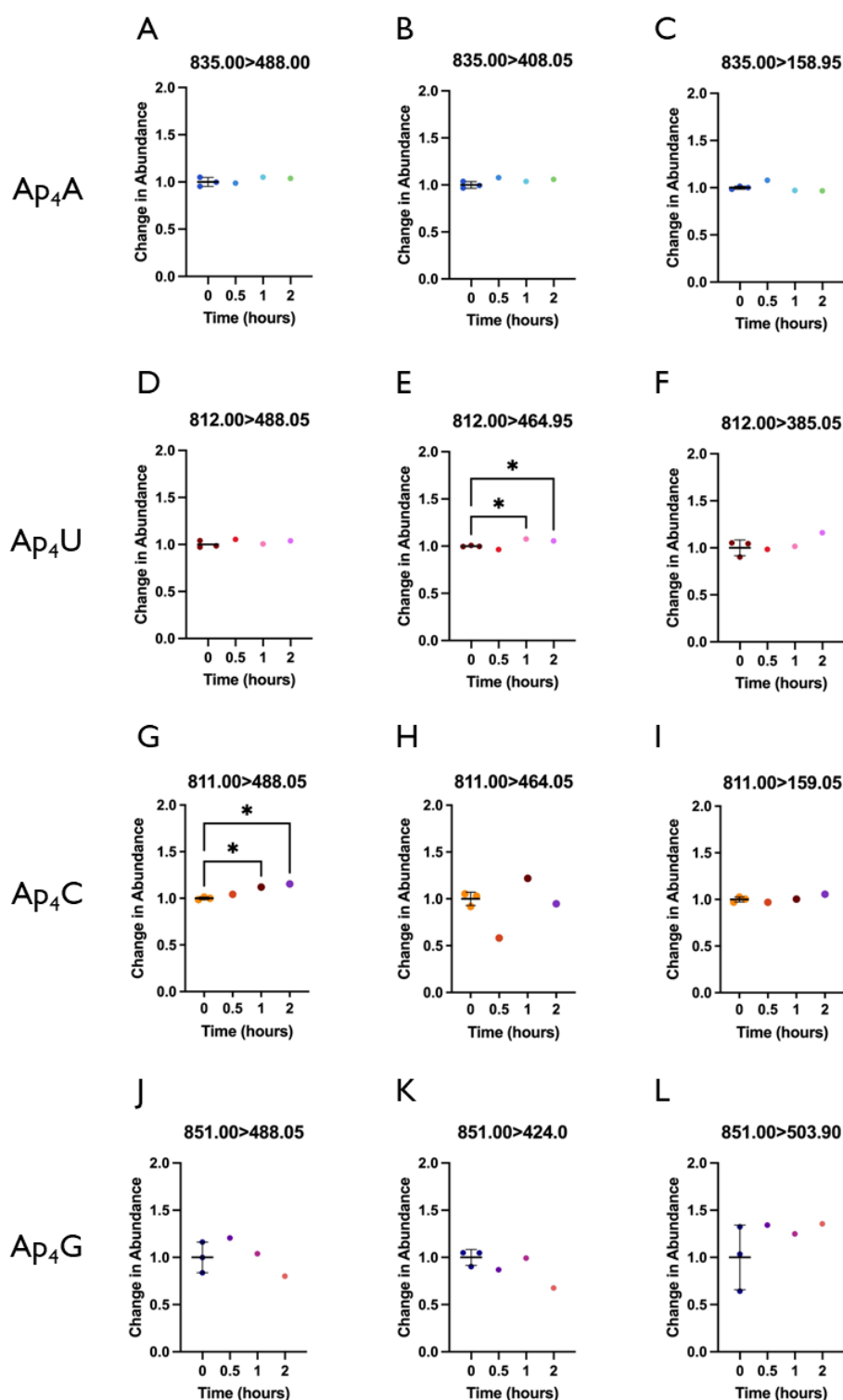
The area under the peak for each of the Ap<sub>4</sub>A transitions varied from the mean by up to 4.9%, 3.7% and 1.6% for each of the transitions, and by up to 4.7% of the mean TIC for Ap<sub>4</sub>A (Fig. 5.5 A). The area under the peak varied from the mean by up to 9.8%, 0.7% and 4.1% for each of the Ap<sub>4</sub>U transitions and by 6.2% for the Ap<sub>4</sub>U TIC (Fig. 5.5 B). The variations for the Ap<sub>4</sub>C transitions were comparable to those for Ap<sub>4</sub>U, with the 811>159.05, 811>464.05 and 811>488.05 transitions varying from the mean by up to 2.9%, 8.2% and 1.3%, respectively and the TIC for Ap<sub>4</sub>C varying by up to 5.1% (Fig. 5.5 C). The area under the peak for the Ap<sub>4</sub>G transitions was less consistent, varying by up to 9.7%, 16.3% and 35.8% (Fig. 5.5 D). Similarly, the TIC for Ap<sub>4</sub>G varied by up to 15.2% compared with the mean (Fig. 5.5 D). It is likely that the larger variation

relative to the mean seen in the Ap<sub>4</sub>G transitions is due to the lower abundance of Ap<sub>4</sub>G fragments in the sample, meaning small changes represent a bigger proportion of the total volume. To summarise, of the four different Ap<sub>4</sub>Ns, the Ap<sub>4</sub>A transitions had the highest ion count, and this correlated with a lower variation in signal (Fig. 5.5 A). Conversely, the Ap<sub>4</sub>G transitions had the lowest ion count, and this correlated with a much higher variation in signal (Fig. 5.5 D).

Additionally, of the top three transitions for each Ap<sub>4</sub>N, all included a transition from the original molecular mass of the molecule to a fragment with a mass of 488.05, which corresponds to the chemical formula of C<sub>10</sub>H<sub>13</sub>N<sub>5</sub>O<sub>12</sub>P<sub>3</sub><sup>-</sup>. A possible structure for this molecule is shown in Figure 5.5 E. As this molecular mass was present in the fragmentation of all Ap<sub>4</sub>Ns, it was chosen as an estimate of the relative abundance of each Ap<sub>4</sub>N in the unstressed NuKO cells. Comparison of the 488.05 fragment abundance for each molecule demonstrated that this Ap<sub>4</sub>A transition was approximately 5 times more abundant than the corresponding Ap<sub>4</sub>U transition, approximately 55 times more abundant than the Ap<sub>4</sub>C transition and approximately 65 times more abundant than the Ap<sub>4</sub>G transition (Fig. 5.5 F and G). It is important to note that the fragmentation efficiency differs between different molecules, and as this fragment contains the adenosine base Ap<sub>4</sub>A is probably more likely to produce this fragment than the other Ap<sub>4</sub>Ns, as the molecule could fragment in two places to form the same fragment while only one fragmentation site is present in the other Ap<sub>4</sub>N molecules.

#### **5.4.2 The Ap<sub>4</sub>N response to hydroxyurea in NuKO cell extracts**

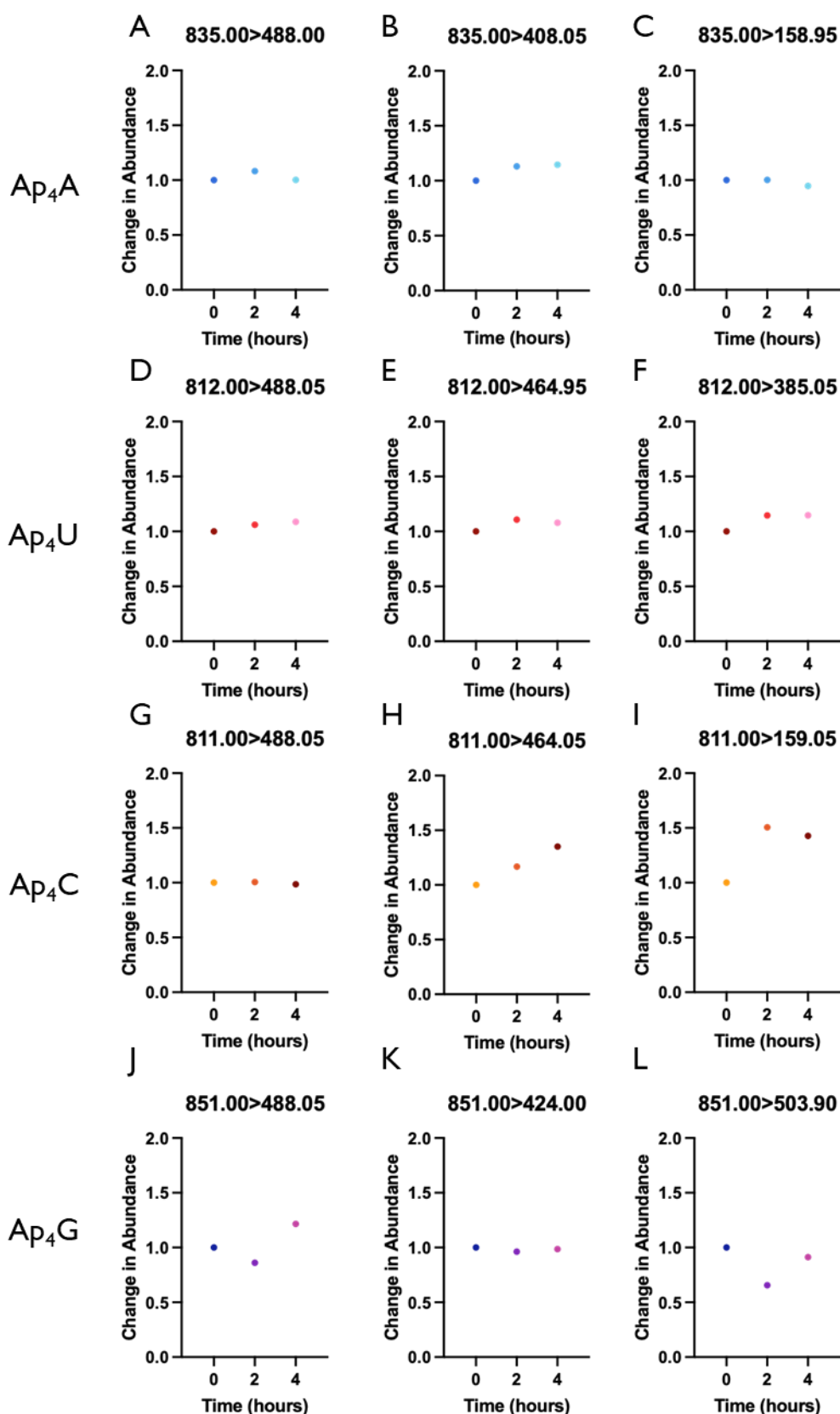
In Figure 5.3 we demonstrated that the maximal  $\gamma$ H2AX response to HU in NuKO cells occurs after approximately 1–2 hours of recovery. As  $\gamma$ H2AX is indicative of the DNA damage response, the next aim was to determine whether activation of the DNA damage response led to a change in intracellular Ap<sub>4</sub>N levels in NuKO cells. NuKO cells were treated with HU for either 0, 0.5, 1 or 2 hours to determine whether increased treatment time influenced Ap<sub>4</sub>N abundance (Fig. 5.6).



**Figure 5.6: The effect of HU treatment on Ap<sub>4</sub>N level in NuKO cells.** NuKO cells were treated with 1 mM HU for 0.5, 1 or 2 hours. Metabolite-containing extracts were harvested and subjected to LC-MS analysis. Changes in the levels of all three MRM transitions for Ap<sub>4</sub>A (A–C), Ap<sub>4</sub>U (D–F), Ap<sub>4</sub>C (G–I) and Ap<sub>4</sub>G (J–L) were analysed. For the untreated cells, the individual data points, as well as the mean values and standard deviation are shown. Statistical analysis was performed using a one-way ANOVA with Dunnett’s multiple comparisons post-hoc test to compare transition abundance of the treated versus untreated cells.

There was no significant change in the abundance of any of the Ap<sub>4</sub>A transitions after treatment with HU for 0.5, 1 or 2 hours compared with 0 hours (Fig. 5.6 A–C), which suggests that the Ap<sub>4</sub>A level in HU-treated NuKO cells is not immediately affected. Similarly, in general the abundance of each of the Ap<sub>4</sub>U transitions did not change significantly with HU-treatment (Fig. 5.6 D–F). There was one exception to this with the 812.00>464.95 transition, where the change in abundance of this transition was considered statistically significant after the 1 and 2 hours of HU treatment (adjusted P value = 0.0190 and 0.0359, respectively) compared with the untreated samples (Fig. 5.6 E). As this was not consistent with any of the other samples, this is likely due to the narrow standard deviation seen in the repeats for the untreated sample. Similarly, two of the three Ap<sub>4</sub>C transitions do not show a statistically significant change in transition abundance (Fig. 5.6 H and I). However, a statistically significant change in the abundance of the 811.00>488.05 transition is observed after 1- and 2-hour HU treatments (adjusted P value = 0.0322 and 0.0202, respectively) compared with the untreated sample (Fig. 5.6 G). As with the Ap<sub>4</sub>U transition, this is likely to be due to the narrow standard deviation for the untreated sample and may not remain significant with further repeats. Finally, although the individual points vary more for the Ap<sub>4</sub>G transitions, none of the transition abundances changed significantly (Fig. 5.6 J–L). This is probably due to the wider standard deviation of the untreated control sample, which likely occurs because the overall lower abundance of the Ap<sub>4</sub>G transitions increases the variability in detection.

In general, short-term HU treatment did not significantly affect Ap<sub>4</sub>N levels. This could be due to several factors. For example, the intrinsically high level of Ap<sub>4</sub>Ns in the NuKO cell line could be masking any smaller changes in Ap<sub>4</sub>N signal, or there may be reduced capacity for Ap<sub>4</sub>N synthesis during HU-induced stress. Alternatively, it may be that more time is needed for Ap<sub>4</sub>Ns to accumulate to detectable levels. To determine whether an extended treatment time could affect Ap<sub>4</sub>N levels, the HU-treatment time was extended to include four hours (Fig. 5.7).

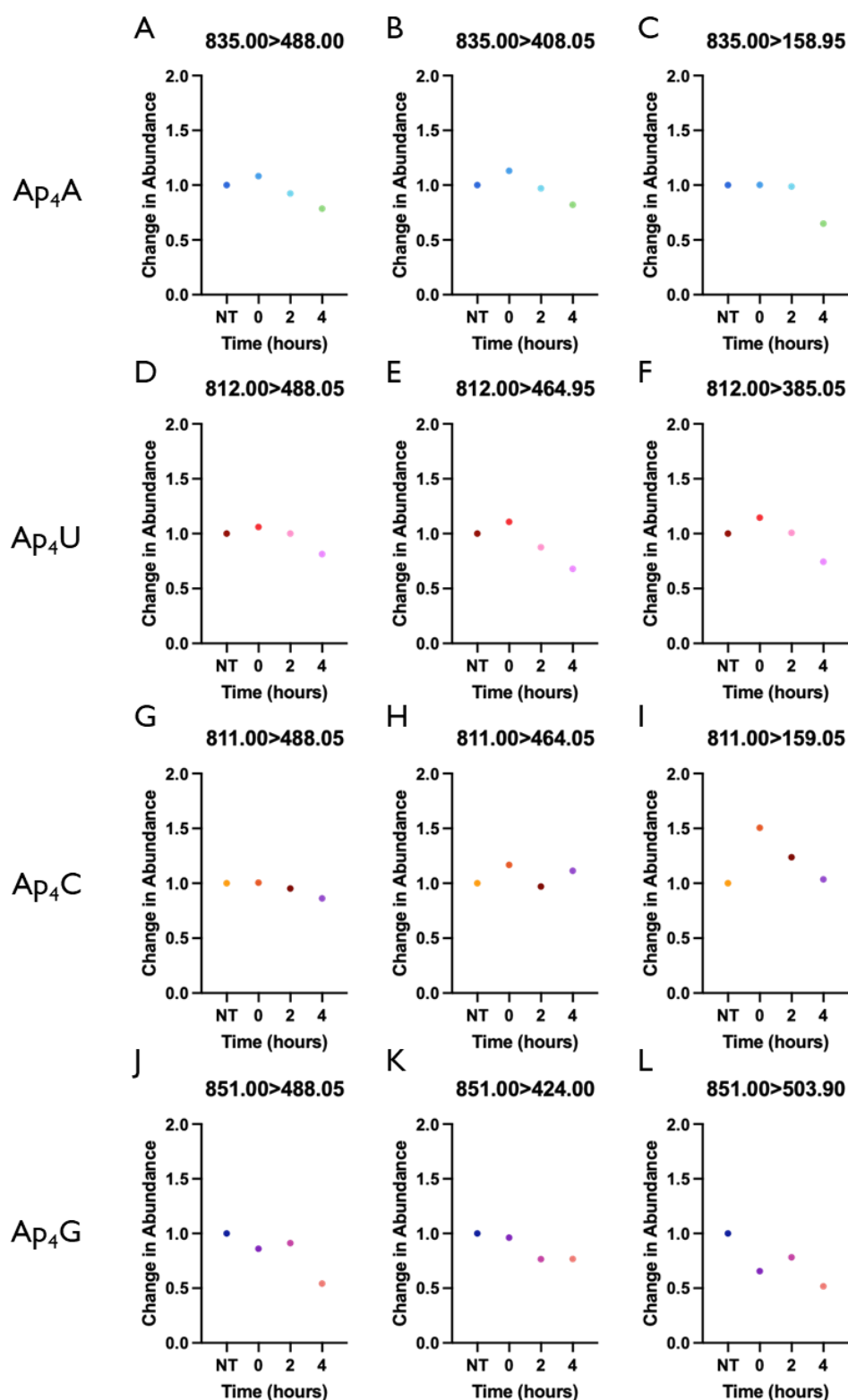


**Figure 5.7 Effect of increased length of HU-treatment on Ap<sub>4</sub>N concentration.** NuKO cells were treated with HU for increasing lengths of time then their extract was immediately harvested and analysed by LC-MS. Changes in the levels of all three MRM transitions for Ap<sub>4</sub>A (A–C), Ap<sub>4</sub>U (D–F), Ap<sub>4</sub>C (G–I) and Ap<sub>4</sub>G (J–L) were analysed.



An extended HU-treatment time of 4 hours did not generally result in a clear change to intracellular Ap<sub>4</sub>N levels; however, only one data point was performed for each transition (Fig. 5.7). The extended timepoint data points showed the least variation for Ap<sub>4</sub>A and Ap<sub>4</sub>U, for which all transitions are present at a higher abundance compared with Ap<sub>4</sub>C and Ap<sub>4</sub>G (Fig. 5.7 A–F vs. G–L). A lack of consistency in pattern between the three transitions, as seen for Ap<sub>4</sub>C or Ap<sub>4</sub>G (Fig. 5.7 G–L), is an indicator of reduced reliability, and is likely caused by the reduced accuracy of transition detection at these lower abundances.

In summary, treatment with HU for between 30 minutes and 4 hours did not result in an immediate detectable increase in Ap<sub>4</sub>N levels (Fig. 5.6 and 5.7). A potential reason for this could be that Ap<sub>4</sub>Ns may not begin to accumulate to detectable levels until after the stressing agent has been removed. To determine whether Ap<sub>4</sub>Ns would accumulate if allowed to recover from HU stress, the NuKO cells were stressed with HU for 2 hours, then allowed to recover for between 0 and 4 hours (Fig. 5.8).



**Figure 5.8: The effect of recovery time on Ap<sub>4</sub>N levels after HU-treatment.** NuKO cells were treated for 2 hours then allowed to recover for 0, 2 or 4 hours. The extract was then harvested and analysed by LC-MS, and Ap<sub>4</sub>N levels were compared with those in the extract from untreated cells. Changes in the levels of all three MRM transitions for Ap<sub>4</sub>A (A–C), Ap<sub>4</sub>U (D–F), Ap<sub>4</sub>C (G–I) and Ap<sub>4</sub>G (J–L) were analysed. Please note, this data set was produced alongside the data set in Figure 5.7. The NT and 0-hour timepoint here correspond to the 0-hour and 2-hour treatments with HU in Figure 5.7.

Generally, the abundance of each transition shows a trend of decreasing as recovery time increases (Fig. 5.8). This is consistent across transitions for Ap<sub>4</sub>A and Ap<sub>4</sub>U, increasing the reliability of these results (Fig. 5.8 A–F). The transitions for Ap<sub>4</sub>C and Ap<sub>4</sub>G are less consistent (Fig. 5.8 G vs. H. vs. I and J vs. K vs. L), meaning these data points are less reliable. The mixed results for the Ap<sub>4</sub>G sample are unsurprising as the Ap<sub>4</sub>G samples showed the most variation even an untreated cell population (Fig. 5.5). However, even in these cases the same general trend of reduced transition abundance with increasing recovery time can be seen. This reduction may be due to a reduced capacity for the cell to synthesise Ap<sub>4</sub>N, either due to the stress itself or due to reduced NTP availability. Alternatively, it is possible that the response mechanism may result in the hydrolysis of Ap<sub>4</sub>N, which would reduce Ap<sub>4</sub>N abundance. In summary, Ap<sub>4</sub>N levels were not immediately affected after 0–4 hours of HU-induced stress, regardless of stress duration. Similarly, Ap<sub>4</sub>N levels were either unaffected or reduced when the cells were allowed time to recover after HU-induced stress.

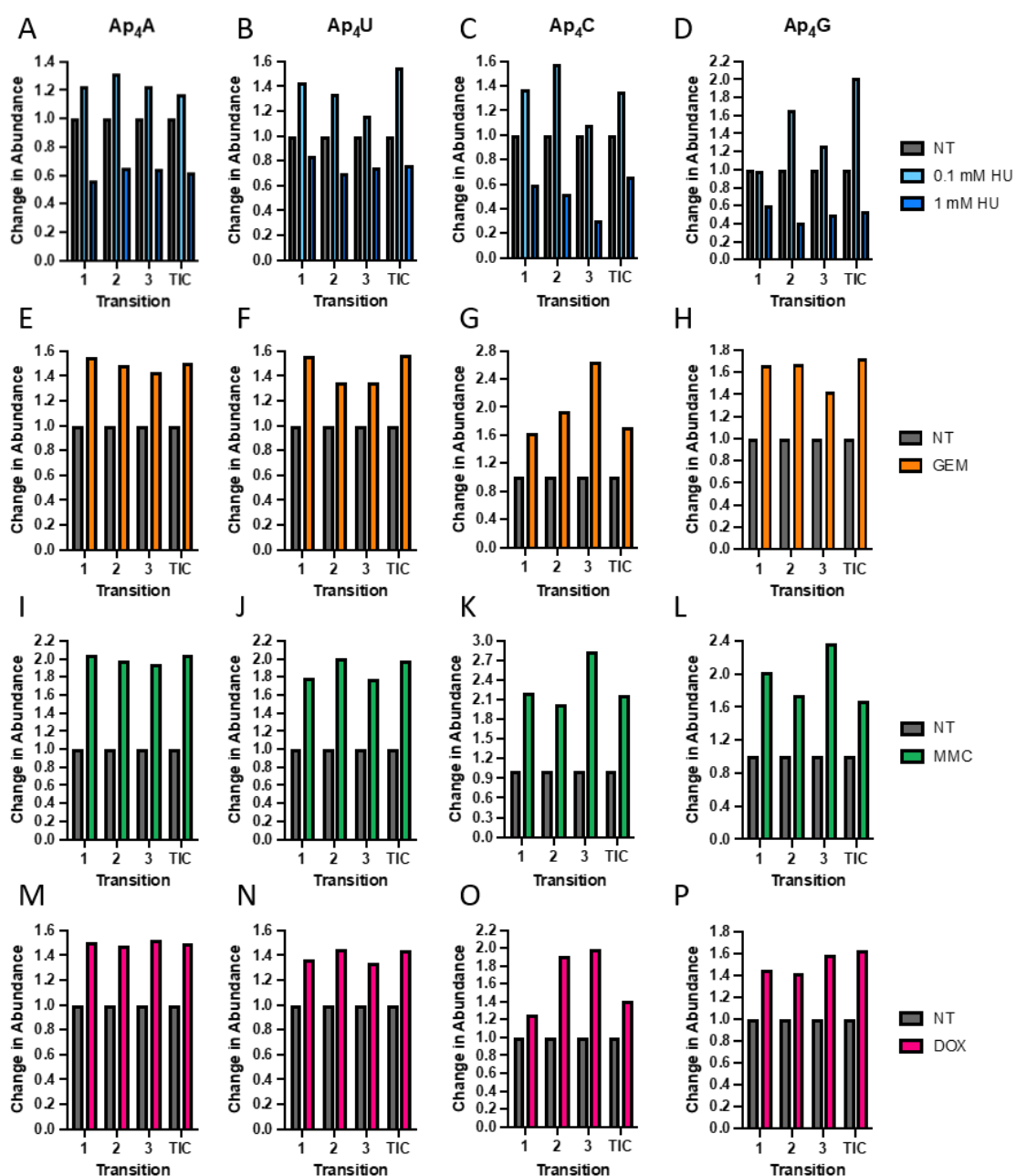
### **5.5 The effect of DNA damaging agents on Ap<sub>4</sub>N levels.**

To determine whether the Ap<sub>4</sub>N response differed depending on the type of stressing agent and its mechanism of action, the NuKO cells were subjected to several different kinds of cell stress. The additional types of cell-stressing agents chosen for analysis were doxorubicin, gemcitabine and mitomycin C.

Doxorubicin is an anthracycline that induces stress in cells through its intercalation into DNA and inhibition of topoisomerase II, or alternatively through the generation of free radicals (Taymaz-Nikerel et al., 2018). Gemcitabine (2',2'-difluoro-2'-deoxycytidine) is a deoxycytidine analogue used in the treatment of several cancers. Gemcitabine was chosen as a stressing agent because, like HU, it also stalls DNA replication by inhibiting DNA synthesis (Huang et al., 1991); however, gemcitabine causes inhibition of DNA synthesis through a different mechanism. Gemcitabine is an analogue of deoxycytidine, and in its triphosphate form it competes with dCTP for incorporation into DNA and prevents elongation through 'masked chain termination' (de Sousa Cavalcante and Monteiro, 2014; Mini et al., 2006). Finally, MMC is an antineoplastic drug that is active against a range of different cancer types (Bradner, 2001). In cells,

the quinone group of MMC can be reduced by a variety of reductases including NADPH-Cytochrome P-450, xanthine oxidase and xanthine dehydrogenase (Gustafson and Pritsos, 1992; Pan et al., 1984). Reduction of the MMC produces an unstable molecule that initiates a cascade resulting in the covalent inter- and intra-strand cross-linking of DNA (Tomasz, 1995). This cross-linking distorts DNA structure and induces DNA replication stress. Nucleotide excision repair and DSB repair through recombination have been implicated in this repair process in different organisms (De Silva et al., 2000; McHugh et al., 2001). Previous analysis has demonstrated that MMC treatment resulted in an increased intracellular Ap<sub>4</sub>A level (Marriott et al., 2015). The method used by Marriott and colleagues (2015) for the quantification of intracellular Ap<sub>4</sub>A level measured the luminescence of ATP produced by hydrolysing the sample. As ATP can be produced by the hydrolysis of all the Ap<sub>4</sub>Ns using Ap<sub>4</sub>A hydrolase (Fig. 3.9 and 3.10) (Swarbrick et al., 2005), this technique cannot distinguish between Ap<sub>4</sub>A and the other Ap<sub>4</sub>Ns. We therefore wanted to investigate whether this increase could be identified in NuKO cells using the LC-MS method, which would enable differentiation between different Ap<sub>4</sub>Ns.

As treatment times using these agents are typically longer and at lower concentration than those for HU, the treatment time and drug concentration of these stressors was adjusted accordingly (Fig. 5.9). An increased HU-treatment time was also performed alongside these experiments for comparison (Fig. 5.9 A–D), and to determine whether increased incubation time would allow Ap<sub>4</sub>Ns to accumulate to higher levels. Furthermore, since this treatment is much longer than the HU treatment discussed previously, the method was adapted such that the cells were counted after the treatment time and prior to metabolite extraction, to prevent any differences in cell number which could result from replication during the incubation time.



**Figure 5.9: Effect of 18-hour treatment with damaging agents on Ap<sub>4</sub>N levels.** Cells were treated with either Hydroxyurea (HU; A–D), Gemcitabine (GEM; E–H), Mitomycin C (MMC; I–L) or Doxorubicin (DOX; M–P) for 18 hours before extracts were harvested and analysed by LC-MS. Values were scaled so that each untreated sample was set equal to one. Transitions 1, 2 and 3 are as follows: Ap<sub>4</sub>A: 1=835.00>488.00, 2=835.00>408.05, 3=835.00>158.95; Ap<sub>4</sub>U: 1=812.00>385.05, 2=812.00>464.95, 3=812.00>488.05; Ap<sub>4</sub>C: 1=811.00>159.05, 2=811.00>464.05, 3=811.00>488.05; Ap<sub>4</sub>G: 1=851.00>424.00, 2=851.00>488.05, 3=851.00>503.90. The corresponding raw ion count values can be found in Supplementary Figure 5.7.

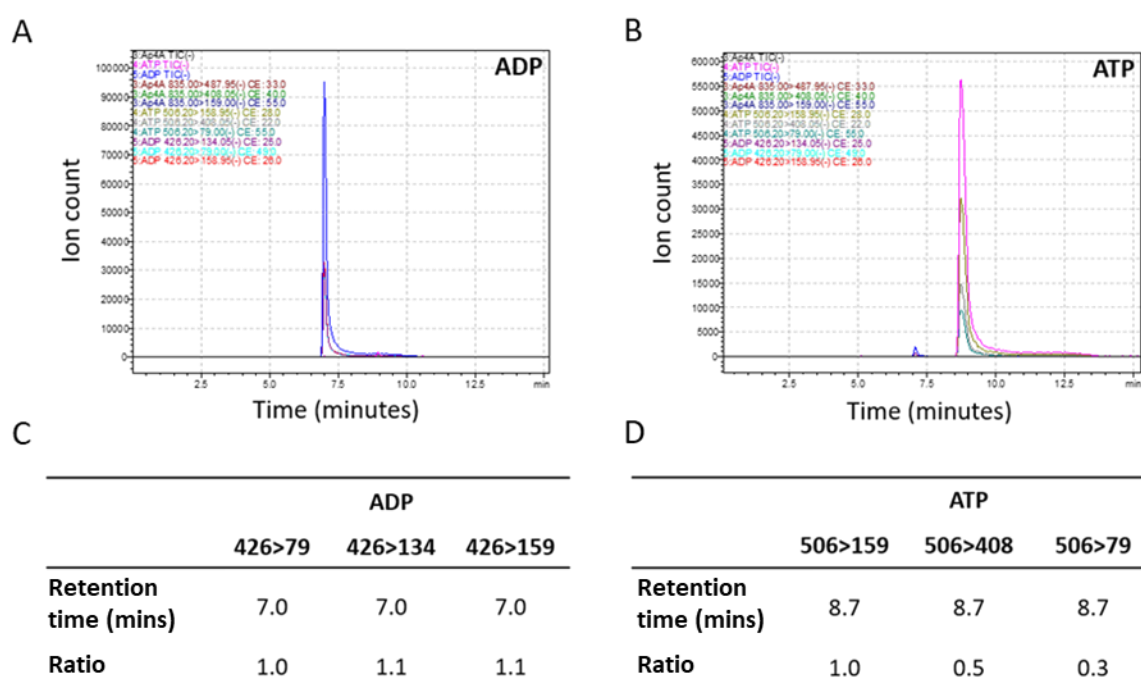
As HU is not usually used for the longer timepoints, two different concentrations were used. After treatment with 1 mM HU, the data suggested a reduction in all Ap<sub>4</sub>N levels

for all transitions (Fig. 5.9 A–D). This was likely caused by the extensive cell death in this sample and so is unlikely to reflect the effect of less severe cell damage on Ap<sub>4</sub>N levels. In general, the data suggested that the intracellular Ap<sub>4</sub>N levels are increased in the cells treated with 0.1 mM HU, with the transition abundance increasing by 23–32% for Ap<sub>4</sub>A (Fig. 5.9 A), 16–43% for Ap<sub>4</sub>U (Fig. 5.9 B) and by 8–58% for Ap<sub>4</sub>C (Fig. 5.9 C). The Ap<sub>4</sub>G transitions also generally followed this trend, except for one transition which suggested a 1% decrease in Ap<sub>4</sub>G levels, despite the other transitions increasing by 26% and 67% (Fig. 5.9 D). The abundance of all Ap<sub>4</sub>N transitions consistently increased after treatment in the 100 nM gemcitabine sample (Fig. 5.9 E–H). The percentage increase in transitions was generally larger for gemcitabine compared with HU, with transition abundance increasing by 43–55% (Ap<sub>4</sub>A, Fig. 5.9 E), 35–56% (Ap<sub>4</sub>U, Fig. 5.9 F), 63–164% (Ap<sub>4</sub>C, Fig. 5.9 G), and 43–67% (Ap<sub>4</sub>G, Fig. 5.9 H). All transitions showed a consistently large increase in abundance after MMC treatment, regardless of which Ap<sub>4</sub>N they corresponded to. Transition abundance increased by 95–104% (Ap<sub>4</sub>A, Fig. 5.9 I), 77–100% (Ap<sub>4</sub>U, Fig. 5.9 J), 102–184% (Ap<sub>4</sub>C, Fig. 5.9 K) and 75–137% (Ap<sub>4</sub>G, Fig. 5.9 L). Increases to each transition after doxorubicin treatment were generally smaller than those for MMC and GEM across all Ap<sub>4</sub>Ns and transitions. Except for transition 1 for Ap<sub>4</sub>C (Fig. 5.9 O), the increase also remained consistent across transitions for each molecule, with increases of 48–52% for Ap<sub>4</sub>A (Fig. 5.9 M), 34–45% for Ap<sub>4</sub>U (Fig. 5.9 N), 25–99% for Ap<sub>4</sub>C (Fig. 5.9 O) and 42–58% for Ap<sub>4</sub>G (Fig. 5.9 P).

Across all four damaging agents, the transitions are more varied for Ap<sub>4</sub>G and Ap<sub>4</sub>C compared with Ap<sub>4</sub>A and Ap<sub>4</sub>U. The increased range for the Ap<sub>4</sub>C and Ap<sub>4</sub>G transitions throughout this data set is reflective of a reduced detection accuracy at lower concentrations, and it is likely that this variation has occurred because of the lower abundance of these transitions compared with those for Ap<sub>4</sub>A and Ap<sub>4</sub>U in NuKO cells (Fig. 5.5 F).

In all, this data set suggested that an 18-hour treatment with any of the four agents tested may result in increased Ap<sub>4</sub>N levels in NuKO cells. This method utilised a cell count prior to extraction to ensure that a similar number of cells were analysed. However, due to the nature of a cell count, there is likely to be variability in the results.

Furthermore, a cell count prior to metabolite extraction does not account for any loss of Ap<sub>4</sub>N that could occur during the extraction process. Therefore, we next aimed to determine a method of controlling for this. To do this, an internal control using ATP or ADP concentration was developed. As a first step, the most abundant MRM transitions were successfully determined and verified using ADP and ATP standards (Fig. 5.10).



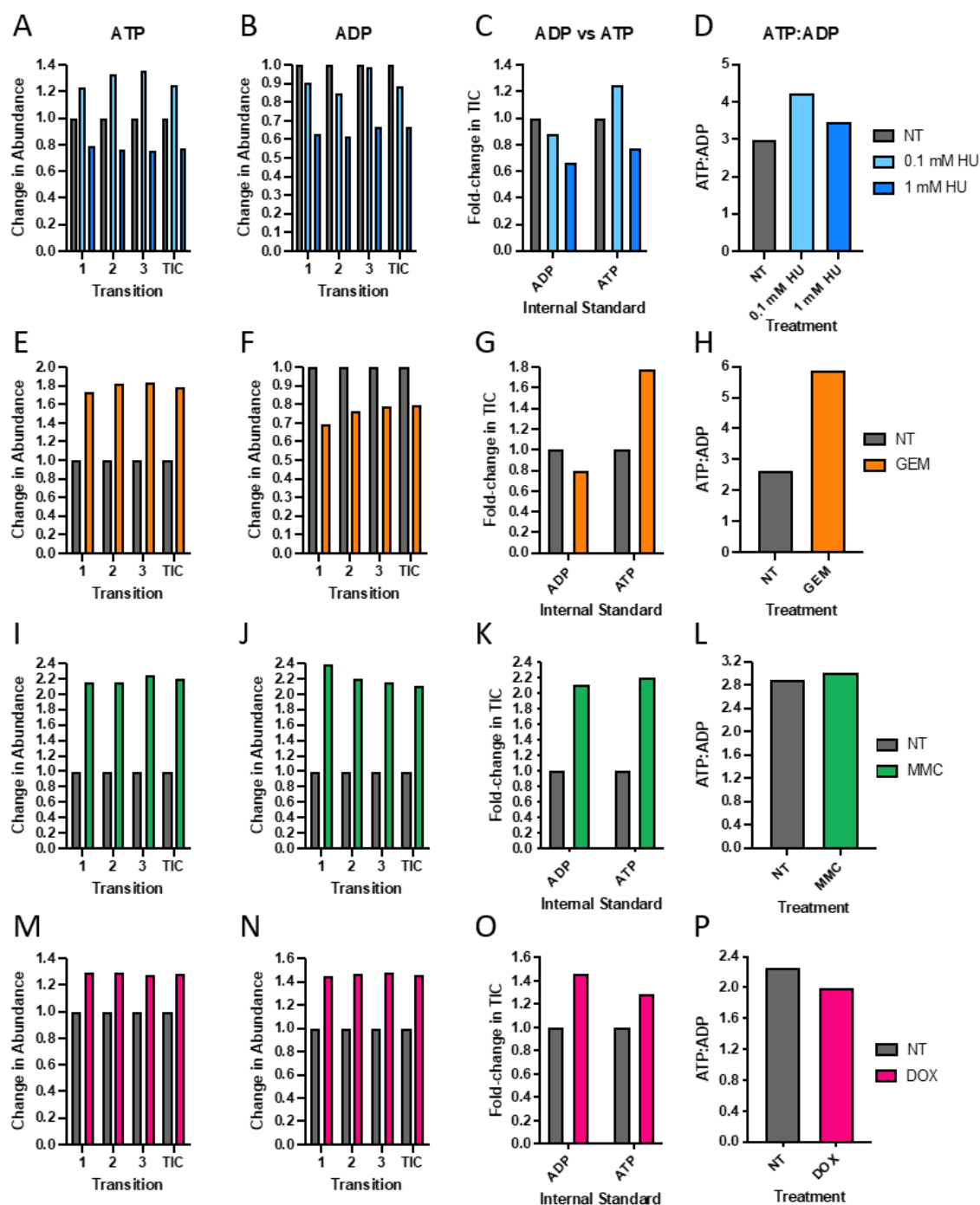
**Figure 5.10: Validation of ADP and ATP MRM transitions using commercially available standards.** Commercially available ADP (A) and ATP (B) standards were diluted and injected for analysis by triple quadrupole LC-MS. The MRM transitions, their relative ratios and their retention times were then determined for ADP (C) and ATP (D).

ADP and ATP had shorter retention times compared with the Ap<sub>4</sub>Ns, eluting at 7.0 and 8.7 minutes, respectively (Fig. 5.10). This meant that the level of ADP and ATP in each sample could be measured earlier in the run and would not interfere with the machine's capacity to measure Ap<sub>4</sub>N levels. ATP and ADP also had different profiles for the relative abundance of each MRM transition (Fig. 5.10 C and D). ADP had a similar amount of each transition (Fig. 5.10 C), while the 506>159 transition was more abundant in the ATP standard compared with the other two transitions (Fig. 5.10 D). As the relative ratio of each of the transitions depends on the fragmentation efficiency of a particular molecule and some fragments form more readily than others, these can

be used as an additional measure to confirm molecule identity in a sample versus a standard.

To make a useful internal standard that could be used for standardisation of the  $\text{Ap}_4\text{N}$  changes after DNA stress, the level of the internal standard should not change in cells exposed to these stressors. To evaluate the potential efficacy of using ATP or ADP as a standard, the abundance of the three transitions was measured for each, before and after treatment with hydroxyurea, gemcitabine, mitomycin C and doxorubicin (Fig. 5.11). The effect of stressing agents on the level of these two molecules was also compared, to determine whether one would be a better control than the other (Fig. 5.11 C, G, K, and O). Finally, the relative ratio of ATP:ADP was evaluated before and after treatment with the stressing agents. If the abundance of ATP and ADP is not changing, then their ratios should stay the same before and after treatment (Fig. 5.11 D, H, L and P).





**Figure 5.11: Effect of damaging agents on internal ADP and ATP standards.** Cells were treated with either Hydroxyurea (HU; A–D), Gemcitabine (GEM; E–H), Mitomycin C (MMC; I–L) or Doxorubicin (DOX; M–P) for 18 hours before extracts were harvested and analysed by LC-MS. Values were scaled so that each untreated sample was set equal to one. For the ATP:ADP ratio, the ADP abundance of the total ion count for each condition was set equal to one, and the ATP value is shown after scaling relative to this. Transitions 1, 2 and 3 are as follows: ATP: 1=506.20>79.00, 2=506.20>158.95, 3=506.20>408.05; ADP: 1=426.20>79.00, 2=426.20>134.05, 3=426.20>158.95. The corresponding raw ion count values can be found in Supplementary Figure 5.8.

After treatment with 1 mM HU, the levels of ATP and ADP were reduced (Fig. 5.11 A). As with the unstandardised data set, this reduction was likely caused by the high amount of cell death observed in these cells when counting, and is unlikely to be reflective of the change that would be observed with less lethal DNA damage. Therefore, the 0.1 mM treatments were the focus. ATP was more affected by HU treatment than ADP (Fig. 5.11 A–C). The ATP transitions all increased by 23–35% after treatment with 0.1 mM HU, whereas the ADP transitions all showed a reduction of between 1% and 15%. This resulted in an increase in the ratio of ATP to ADP, from 3:1 to 4.2:1 (Fig. 5.11 D), which further suggested that the balance of these molecules had been affected by HU treatment. An even greater increase in ATP transition levels was seen after GEM treatment (73–84%, Fig. 5.11 E), while ADP levels decreased by between 21% and 31% (Fig. 5.11 F). ATP transition abundance was most strongly affected by gemcitabine treatment (Fig. 5.11 G). This, combined with the reduction in ADP transition abundance, resulted in a large change in the relative ratio of ATP to ADP, from 2.6:1 to 5.9:1 (Fig. 5.11 H). The balance of ATP and ADP in the cell is strongly affected by both hydroxyurea and gemcitabine treatment, suggesting that either one or both molecules would not make a good internal standard. A possible explanation for this effect may relate to the mechanism of action of these drugs and will be addressed in the discussion section.

Mitomycin C and doxorubicin both resulted in an increase in all transitions corresponding to both ATP and ADP (Fig. 5.11 I–K and M–O). ATP transition abundance increased by 115–124% (Fig. 5.11 I) and 28–30% (Fig. 5.11 M), in MMC and doxorubicin-treated cells, respectively. Similarly, ADP transition abundance increased by 115–139% in MMC-treated cells (Fig. 5.11 J) and 45–48% in doxorubicin-treated cells (Fig. 5.11 N). For both treatments, ATP and ADP total ion counts increased similarly (Fig. 5.11 K and O). Therefore, the ratios of ATP to ADP remain similar (2.9:1 vs 3.0:1 for MMC [Fig. 5.11 L] and 2.3:1 vs 2.0:1 for doxorubicin [Fig. 5.11 P]). As the ATP:ADP ratios have remained consistent, it is difficult to determine whether the observed increase is due to disruption of both ATP and ADP levels, or whether ATP and ADP levels were increased in this sample for some other reason.

As the ATP and ADP responses are different in response to different damaging agents, they are unlikely to be ideal internal controls. For the HU and GEM treatments abundance of the ADP transitions did not change substantially (5.11 C and G); therefore, ADP has the potential to act as an internal standard when analysing the effect of these agents on Ap<sub>4</sub>N levels. However, more repeats are needed to determine whether the small reduction in ADP seen after treatment with both agents was due to variability of the detection, or whether it is a true reflection of a small reduction in ADP abundance caused by HU and GEM treatment. As ADP and ATP transition abundances increase greatly after MMC and DOX treatment, neither molecule is likely to make a good internal standard. However, the ratio of these molecules remains very similar. Therefore, this observed change may be caused by variation in cell number, or cell extraction efficiency between the two samples. Further repeats would help to identify whether there was a consistently large increase in both ATP and ADP abundance when treated with MMC or DOX. For reference, a summary of the observed effects of HU, GEM, MMC or DOX treatment on Ap<sub>4</sub>N abundance both before and after standardisation to either ADP or ATP is shown in Table 5.1.

**Table 5.1: Summary of the mean change in abundance of the three transitions for each Ap<sub>4</sub>N.** An increase or decrease in mean transition abundance is shown by an ↑ or ↓, respectively, with the mean percentage change across transitions represented in brackets. As ATP is likely a poor standard for HU and GEM, these values are shaded in red. The value of ADP as a standard for HU and GEM treatments, and either ADP or ATP for MMC and DOX treatments is yet to be confirmed. This is represented by yellow shading.

		HU	GEM	MMC	DOX
<b>Unstandardised</b>	<b>Ap<sub>4</sub>A</b>	↑ (26)	↑ (49)	↑ (99)	↑ (50)
	<b>Ap<sub>4</sub>U</b>	↑ (31)	↑ (42)	↑ (85)	↑ (38)
	<b>Ap<sub>4</sub>C</b>	↑ (34)	↑ (107)	↑ (135)	↑ (72)
	<b>Ap<sub>4</sub>G</b>	↑ (31)	↑ (59)	↑ (105)	↑ (48)
<b>Standardised to ATP</b>	<b>Ap<sub>4</sub>A</b>	↑ (1)	↓ (16)	↓ (10)	↑ (17)
	<b>Ap<sub>4</sub>U</b>	↑ (5)	↓ (18)	↓ (16)	↑ (7)
	<b>Ap<sub>4</sub>C</b>	↑ (7)	↑ (17)	↑ (7)	↑ (33)
	<b>Ap<sub>4</sub>G</b>	↑ (5)	↓ (11)	↓ (7)	↑ (15)
<b>Standardised to ADP</b>	<b>Ap<sub>4</sub>A</b>	↑ (42)	↑ (87)	↓ (6)	↑ (3)
	<b>Ap<sub>4</sub>U</b>	↑ (48)	↑ (78)	↓ (12)	↓ (5)
	<b>Ap<sub>4</sub>C</b>	↑ (52)	↑ (160)	↑ (11)	↑ (18)
	<b>Ap<sub>4</sub>G</b>	↑ (48)	↑ (99)	↓ (3)	↑ (2)

When standardised to ADP or left unstandardised, the mean abundance across the three transitions for each of Ap<sub>4</sub>A, Ap<sub>4</sub>U, Ap<sub>4</sub>C and Ap<sub>4</sub>G was increased after treatment with HU and gemcitabine. When standardising to ATP the results were mixed, with a small increase in Ap<sub>4</sub>N transition abundance after HU treatment, and a decrease in Ap<sub>4</sub>N (except for Ap<sub>4</sub>C) abundance after gemcitabine treatment. However, ATP was considered a poor method of standardisation for these treatments because of the large change in ATP abundance after HU or gemcitabine treatment. Mitomycin C appeared to result in an increase in Ap<sub>4</sub>N abundance in the absence of standardisation; however, standardisation to either ADP or ATP hinted at small reductions in Ap<sub>4</sub>A, Ap<sub>4</sub>U and Ap<sub>4</sub>G abundance. Except for Ap<sub>4</sub>U abundance when standardised to ADP, treatment with doxorubicin resulted in an increase in the mean transition abundance for each of the Ap<sub>4</sub>Ns, regardless of standardisation method. Interestingly, mean Ap<sub>4</sub>C transition abundance consistently increased after treatment with any of the stressing agents, regardless of the method used for standardisation.

## 5.6 Chapter Discussion

### 5.6.1 Cells with intrinsically higher Ap<sub>4</sub>N levels have a slower doubling time

In this chapter we have demonstrated that the KBM7 *NUDT2*<sup>-/-</sup> NuKO cell line, has a slower duplication time (22.9 hours) compared with its parental KBM7 cell line (16.9 hours) and doubling rates show a statistical significance from day 4 onwards (two-way ANOVA with Šidák's multiple comparisons test). This change in proliferation rate is similar to previous analysis in a breast cancer cell line, which demonstrated that T47D cells transfected with *NUDT2*-specific siRNA had a reduced proliferation rate compared with cells transfected with control siRNA (Oka et al., 2011). Additionally, previous analysis identified a number of differentially expressed genes relating to proliferation in NuKO cells compared with the parental KBM7 cell line (Marriott et al., 2016). Examples of some of the changes detected include down-regulation of *GFRA1* and *BRINP3* expression and up-regulation in the expression of genes such as *FOXD3* and *NKD2* (Marriott et al., 2016). GDNF-treated prostate cancer cell lines with increased proliferation correlate with high *GFRA1* expression, while the proliferation rates of cell lines with low *GFRA1* expression are not significantly affected (Huber et al., 2015). This suggests that a reduction in *GFRA1* protein level may contribute to

reduced replication rates in the NuKO cells. In gonadotrope cell pituitary adenomas BRINP3 initiates proliferation, suggesting that down-regulation of this protein may reduce cell proliferation (Shorts-Cary et al., 2007). Together evidence from the literature supports the trend showing a decrease in proliferation rate when the *NUDT2* gene is knocked out.

#### **5.6.2 NuKO cells show similar amounts of time in each phase of the cell cycle but a prolonged transition between G1 and S-phase**

Despite the reduced proliferation rate of the NuKO cells, we did not identify any clear differences in their cell cycle kinetics. However, the extended duplication time of the NuKO cells means that the NuKOs spend approximately 1 hour longer in G1, 3.6 hours longer in S-phase and 1.4 hours longer in G2/M compared with the KBM7 cells. Furthermore, the NuKO cell G1- and S-phase populations were not as distinct as the corresponding KBM7 cell populations, suggesting that perhaps there is a delay to these cells at the G1/S-transition (Fig. 5.2 A and B). As Ap<sub>4</sub>A levels are intrinsically higher in these cells, this data supports the hypothesis that Ap<sub>4</sub>A may have a role at the G1/S-transition in cells. This is further supported by data in the literature which has demonstrated that Ap<sub>4</sub>A can inhibit the initiation stage of DNA replication (Marriott et al., 2015). If Ap<sub>4</sub>A is important in regulating the initiation of DNA replication, the intrinsically high levels in the NuKO cells could mean either that the excess of Ap<sub>4</sub>A is inhibiting the initiation into S-phase, or that the process is not being regulated as it normally would be, as the Ap<sub>4</sub>A levels are not being regulated themselves in these cells.

#### **5.6.3 High Ap<sub>4</sub>N-containing NuKO cells show a reduced response to replication stress**

In this chapter, NuKO cells were found to have a reduced ability to respond to cell stress compared with the parental KBM7 cell line.  $\gamma$ H2AX has a well-documented role as a sensitive early marker of DNA damage and is produced when the Ser-139 residue of histone variant H2AX is phosphorylated in response to such damage (Mah et al., 2010).

The data demonstrate that treatment of KBM7 and NuKO cells with hydroxyurea induces the phosphorylation of H2AX into  $\gamma$ H2AX in both cell lines, but this increase is significantly reduced in the NuKO cell line compared with the KBM7 cell line after 2-and 4-hours of recovery. This suggests that the NuKO cells cannot elicit as much of a DNA damage response as the KBM7 cells. The NuKO response peaks after 1-hour of recovery, whereas the KBM7 response peaks after 2 hours of recovery time, demonstrating that it is the strength of the response and not the time taken for the response to occur that causes this difference between the KBM7 and NuKO cells. Supporting this data is evidence linking an increase in  $\text{Ap}_4\text{N}$  concentration to cell stress (Baker and Jacobson, 1986; Bochner et al., 1984; Ji et al., 2019; Krüger et al., 2021; Marriott et al., 2015). As  $\text{Ap}_4\text{Ns}$  are already intrinsically higher in the NuKO cell line, it follows that changes to the  $\text{Ap}_4\text{N}$  concentration on top of this due to stress may no longer have the expected effect.

H2AX is phosphorylated by both ATM (Burma et al., 2001) and ATR (Ward and Chen, 2001) in response to stressing agents such as hydroxyurea, and is part of the ATM and ATR signalling cascades. Therefore, a reduction in  $\gamma$ H2AX level could be reflective of reduced ATM or ATR signalling. As  $\gamma$ H2AX accumulation is reduced in the cells with altered  $\text{Ap}_4\text{N}$  levels, it is possible that  $\text{Ap}_4\text{Ns}$  are involved upstream of H2AX phosphorylation in either one or both of these damage response mechanisms. The role of  $\text{Ap}_4\text{N}$  in these responses has not been investigated in detail here and is an area that could be explored in future with further analysis.

#### **5.6.4 NuKO cells are more susceptible to apoptosis after HU-induced cell stress**

Our data also suggest that cells with intrinsically high  $\text{Ap}_4\text{N}$  levels are more susceptible to apoptosis after a 24-hour recovery time from HU-induced stress. After the 24-hour recovery time, over 10x as many cells were in apoptosis in the NuKO sample compared with the KBM7 sample, suggesting that  $\text{Ap}_4\text{Ns}$  may have a role in apoptosis. This data has been collected from a single replicate and therefore more replicates would be needed to confirm this. However, the data here are supported by other reports that have linked increased  $\text{Ap}_4\text{A}$  concentration to an increase apoptosis (Krakowiak et al., 2011; Marriott et al., 2016; Vartanian et al., 1999). In particular, a number of

up-regulated pro-apoptotic genes have been identified in the NuKO cell line compared with the parental KBM7 cell line, which strongly supports this data (Marriott et al., 2016).

#### **5.6.5 Stressing agents may increase intracellular Ap<sub>4</sub>N levels**

In this chapter we have used the triple quad mass-spectrometry method developed in Chapter 4 to analyse changes in intracellular concentration of Ap<sub>4</sub>A, Ap<sub>4</sub>C, Ap<sub>4</sub>G and Ap<sub>4</sub>U after induction of stress in NuKO cells. Initial studies suggested little change in Ap<sub>4</sub>N abundance after short-term treatment with HU. Further studies then suggested that an 18-hour treatment with hydroxyurea, gemcitabine, doxorubicin and mitomycin C may increase intracellular Ap<sub>4</sub>A, Ap<sub>4</sub>U Ap<sub>4</sub>C and Ap<sub>4</sub>G levels. This data is supported by evidence that Ap<sub>4</sub>N levels are elevated in response to cell stress (Coste et al., 1987; Lee et al., 1983; Marriott et al., 2015); however, it lacked an internal standardisation method to account for any changes to Ap<sub>4</sub>N abundance brought about during the cell counting, extraction, or lyophilisation processes.

#### **5.6.6 ADP levels are reduced slightly after HU and GEM treatment**

Assessment of ADP and ATP as potential internal standards for analysis of Ap<sub>4</sub>N abundance in response to gemcitabine or hydroxyurea treatment suggested that the abundance of both molecules may be affected by these stressors. A limitation of these standardisation methods is that they require a constant intracellular ADP or ATP concentration. Interestingly, ADP abundance decreased after treatment with HU or gemcitabine (Fig. 5.11). This is surprising, as hydroxyurea has been shown to inhibit CDP, ADP UDP and GDP reductase activity (Yeh and Tessman, 1978), and the dFdCDP product of gemcitabine also inhibits ribonucleotide reductase (Heinemann et al., 1990). As the ribonucleotide reductases are responsible for the reduction of NDP into dNDP, this would suggest that hydroxyurea and gemcitabine may result in an increased intracellular ADP concentration, as ADP is no longer being reduced by the reductases (Torrents, 2014). Supporting this, another study found that treatment of neuronal stem cells (NSCs) with hydroxyurea resulted in a slight increase in ADP levels (Daniele et al., 2016). However, evidence from cerebellar granule cells suggests that intracellular ADP levels are reduced during early apoptosis alongside an increase in

ATP levels (Atlante et al., 2005). This reflects the decrease in ADP and concomitant increase in ATP demonstrated after treatment with HU and GEM (Fig. 5.11 C and G).

An alternative possibility is that the reduction in ADP levels observed after HU and gemcitabine treatment may be due to variation in detection by the machine. To determine whether this is the case, future studies could aim to identify the expected variation in detection of ADP at this concentration in NuKO cells. ADP abundance was consistently lower across all transitions, which supports the data suggesting that ADP level was lower in this sample. However, further repeats could also be used to determine whether ADP levels are consistently lower after treatment with HU or GEM, or whether this was caused by an issue with the extraction process.

#### **5.6.7 ATP levels are increased after treatment with HU and GEM**

While reductions in ADP level are relatively small, ATP showed a particularly strong increase in abundance after both HU and GEM treatment (Fig. 5.11 C and G). As ADP levels are not affected similarly, it is likely that these large increases in ATP levels are due to the stresses rather than error in the extraction process. Ribonucleotide reductase also usually converts NTP to dNTP, so inhibition of this enzyme by HU and GEM is likely to contribute to the observed increase in ATP levels. Additionally, due to the nature of the stressing agents, some cells will have died during 18 hours of treatment. As these cells are suspension cells there was no way to remove dead cells from the sample; however, only live cells were counted when measuring cell number for analysis. Therefore, any dead cells would remain in the sample and be present in addition to the equal number of counted live cells. These cells may still retain residual ATP or ADP levels, which may disproportionately affect intracellular ATP concentration in the damaging agent-treated cells, as cytosolic ATP concentration is increased during apoptotic cell death (Zamaraeva et al., 2005). However, some evidence suggests that ATP levels can be reduced by HU treatment and does not support the effect seen here (Daniele et al., 2016; Raththagala et al., 2010). There are reports of ATP being released from erythrocytes after hydroxyurea treatment, suggesting that hydroxyurea might reduce intracellular ATP concentration (Raththagala et al., 2010). This reduction is



supported by a study in NSCs, which demonstrated that 8 mM HU significantly decreased ATP concentrations (Daniele et al., 2016).

Together, the data suggest that of the two molecules used for standardisation, the more reliable method for determining Ap<sub>4</sub>N concentration after HU or gemcitabine treatment is to standardise to ADP. However, ADP levels were also affected slightly, and there is existing evidence suggesting that ADP levels may be reduced during early apoptosis (Atlante et al., 2005). Therefore, it is yet to be determined whether ADP is a suitable molecule to standardise to. There may be other factors that could have affected ADP concentration, and further studies are needed to determine whether the small reductions in ADP are due to variation in detection or whether it is something that is happening on a physiological level due to the DNA stressors.

#### **5.6.8 ATP and ADP levels are increased after treatment with MMC and DOX**

Both MMC and Dox treatment induce an increase in both ATP and ADP levels (Fig. 5.11 I–P). Unlike with HU and GEM treatment, ATP and ADP levels are increased proportionally to each other, meaning that their relative ratios are similar before and after treatment (Fig. 5.11 L and P). MMC and DOX do not inhibit ribonucleotide reductase, so this mechanism does not provide an explanation for the increases observed here. In fact, previous studies using DOX have shown a reduction in ATP levels after DOX treatment. One study demonstrated that treatment with 1 µM doxorubicin reduces ATP levels after 14 or more hours of treatment (Jeyaseelan et al., 1997). Another study in cardiomyocytes instead found that ADP and ATP levels both decrease immediately after doxorubicin treatment, but that the level of these molecules begins to recover after 2 hours of treatment (Pointon et al., 2010). In this 2010 study by Pointon and colleagues, 12 hours after starting doxorubicin treatment both ADP and ATP concentrations were no longer significantly different from without treatment. Similarly, a previous study investigating the effect of chemotherapeutic agents on ATP level demonstrated that intracellular ATP levels decrease with increasing MMC concentrations of between 1 and 30 µM compared with 0 µM MMC (Martins et al., 2009). As these concentrations are all higher than the 100 nM used in

the experiments in this chapter, this could explain why a decrease is not seen. However, it does not explain the increased ATP levels observed in Figure 5.11.

As ADP and ATP levels were affected similarly, and the data does not follow the pattern seen in the literature, a factor secondary to the damage mechanism itself could have resulted in the observed increase. For example, live cells are more likely have been inaccurately counted in the treated samples, as these contained more dead cells. As live cells were discriminated based on membrane structure alone, this process may overestimate the number of 'live' cells. To ensure more accurate counting, future work could incorporate the use of a live/dead cell stain such as trypan blue (TB). This would eliminate the human error aspect of differentiating between live and dead cells. However, even with this technique, viability and cell population density determined using TB still vary by approximately 5% and 20%, respectively (Piccinini et al., 2017). Additionally, as any residual ATP in dead or dying cells cannot be removed because of the suspension-cell nature of the NuKO cells, this may result in an increase in observed levels of both ATP and ADP. Therefore this may be responsible for the increased ATP and ADP levels after MMC and DOX treatment. Further repeats of the data set could also help to establish whether this increase remains consistent and could therefore help to determine whether ADP or ATP have the potential to be used as a standard for MMC or DOX treatment.

#### **5.6.9 The standardisation of Ap<sub>4</sub>N levels to internal controls ATP and ADP gives inconsistent results**

As ADP and ATP levels appear to be affected by the different types of stress, without further data it cannot be confirmed whether they could be used as internal standards for these studies. Particularly in the case of ADP, further studies could be used to determine whether ADP levels are consistently changed after stress. In summary Table 5.1, standardisation of the data shows mixed results in terms of the effect of stress on Ap<sub>4</sub>N levels.

Interestingly, this standardisation of the data to either ATP or ADP suggests that MMC does not impact Ap<sub>4</sub>N levels. This is particularly interesting because luminescence

experiments have shown that in KBM7 cells, the parental cell line, Ap<sub>4</sub>N levels increase in response to MMC treatment (Marriott et al., 2015). An explanation for the difference in these findings from the published data could be that ATP and ADP are inappropriate standards to use. Alternatively, it could be that the *Nudt2* knockout in these cells increases the intracellular Ap<sub>4</sub>N level enough that any changes brought about by stresses may be obscured. This possibility emphasises the need for the development of a method which enables the detection of the Ap<sub>4</sub>Ns at levels present in the KBM7 cell line. Therefore, it would be useful for future studies to focus on developing an improved pre-quantitation clean-up technique that would sufficiently prevent an ion suppression effect, or for these experiments to be attempted on a machine with higher sensitivity.

As the current data set suggests that ADP or ATP are unlikely to be suitable internal standards for analysis, it would also be of use to identify a suitable standard. This remains a challenge due to the need for the molecule to be extracted from cells and retained at a similar efficiency as the Ap<sub>4</sub>Ns. One option would be to determine whether AMP could work as a useful internal standard, especially as a previous study found that AMP levels remained unchanged in cells in the early stages of apoptosis (Atlante et al., 2005). However, there is a limitation to the use of AMP in that AMP is a hydrolysis product of Ap<sub>4</sub>A and the other Ap<sub>4</sub>Ns (Fig. 3.9 and 3.10). This means that if the inhibitory effect of Ap<sub>4</sub>A is brought about through hydrolysis, this would result in changes to the intracellular AMP level.

## 5.7 Chapter conclusion

This chapter has demonstrated that cells with intrinsically high Ap<sub>4</sub>N levels possess slower duplication times, but similar cell cycle kinetics. Further to this we have shown that HU treatment induces a reduced  $\gamma$ H2AX response and increased apoptosis in the high-Ap<sub>4</sub>N cell line. We have used the technique developed in Chapter 4 to analyse changes in Ap<sub>4</sub>A, Ap<sub>4</sub>U Ap<sub>4</sub>C and Ap<sub>4</sub>G concentration after subjecting the high-Ap<sub>4</sub>N-containing NuKO cells to stress. Additionally, further study into the effect of each of the stress types on ADP and ATP concentrations is necessary to determine whether either of these could be used as internal standards. Unfortunately, as demonstrated by

the MMC treatment, because the cells already contained such high Ap<sub>4</sub>N levels, it is likely that this caused any small changes in concentration because of stress to be obscured. We have demonstrated proof of principle for the identification of individual Ap<sub>4</sub>Ns in the NuKO cell line. This method could therefore be very useful for identification of changes in Ap<sub>4</sub>N concentration in the KBM7 cell line, but use of a more sensitive machine or improved clean-up technique would be required.

# **Chapter 6: The role of Ap<sub>4</sub>A and Ap<sub>4</sub>Ns in the inhibition of DNA replication initiation**

## **6.1 Introduction to the role of Ap<sub>4</sub>A in DNA replication**

### **6.1.1 DNA replication can be reconstituted *in vitro***

The reconstitution of DNA replication *in vitro* has provided a reductionist approach to analyse the highly orchestrated process of DNA replication (Yeeles et al., 2015). This process used purified proteins to identify Cdt/Mcm2–7, ORC, Cdc6, DDK, Sld3/7, Cdc45, S-CDK, Dpb11, Sld2, GINS, Mcm10, Pol  $\alpha$ , Pol  $\epsilon$ , Ctf4, RPA and Topoisomerase II as the essential proteins and factors that are required for DNA replication to occur in budding yeast (Yeeles et al., 2015). Further reconstitution of the DNA replication system identified additional factors in budding yeast, Mrc1 and Csm3/Tof, that are necessary for replisome progression to match *in vivo* rates (Yeeles et al., 2017). Although this process is very useful for determining the mechanics of DNA replication in yeast, there is no recombinant system for the analysis of replication dynamics in mammalian cells. However, a cell-free system has been developed which enables the induction of DNA synthesis in late G1-phase nuclei (Coverley et al., 2002; Krude et al., 1997)

### **6.1.2 Introduction to the cell-free replication system**

In 1970, Rao and Johnson demonstrated that the fusion of G1- and S-synchronised HeLa cells resulted in the induction of DNA synthesis in G1 nuclei. In this system, the higher the proportion of S-phase nuclei in the system, the faster the induction of DNA synthesis (Rao and Johnson, 1970). In 1997 this system was further developed by Krude and colleagues, who demonstrated that S-phase extracts could be used instead of S-phase nuclei to initiate semiconservative DNA replication in G1-phase, but not G2-phase, nuclei. Furthermore, DNA replication is still initiated with the same efficiency in a system when the S-phase extract has been replaced with Cdk2 and the S-phase cyclins, cyclin E and cyclin A (Krude et al., 1997). The stimulation of DNA replication by the cyclin/Cdk complexes was inhibited by the kinase inhibitor 6-dimethylaminopurine (DMAP), and also at higher concentrations of cyclin/Cdk (Krude et al., 1997). Experiments incubating mid-G1 and late-G1 nuclei in cyclin E/Cdk2 and cyclin A/Cdk2 extracts provided initial evidence that, together with Cdc6, cyclin E promotes pre-replication complex assembly (Coverley et al., 2002). Subsequent

addition of Cyclin A/Cdk2 promotes DNA synthesis and inhibits the formation of new pre-RCs (Coverley et al., 2002). This cell-free replication system has already been applied for the identification of Ciz1, a factor suggested to promote DNA replication after formation of the replication complex (Coverley et al., 2005).

### **6.1.3 Ap<sub>4</sub>A inhibits the initiation of DNA replication**

For many years, Ap<sub>4</sub>A has been implicated in the DNA replication process (Grummt, 1978b). However, early studies suggested that Ap<sub>4</sub>A stimulated DNA replication (Grummt, 1978b; Weinmann-Dorsch et al., 1984a; Weinmann-Dorsch et al., 1984b), and it was only recently demonstrated that Ap<sub>4</sub>A instead has a role in the inhibition of DNA replication (Marriott et al., 2015). This inhibition of DNA replication was demonstrated using a cell-free system, and showed that Ap<sub>4</sub>A inhibited the initiation, but not elongation stage of DNA replication (Marriott et al., 2015). In this chapter we will utilise this system to investigate whether the other Ap<sub>4</sub>Ns might play a similar role in inhibiting the initiation phase of DNA replication.

### **6.1.4 Aims**

In the previous chapters we have demonstrated that Ap<sub>4</sub>Ns can be synthesised via a biologically relevant mechanism and that they are present in the NuKO mammalian cell line (Chapters 3–5). Furthermore, we have successfully synthesised the Ap<sub>4</sub>Ns chemically and verified their identity by HRMS (Chapter 3). Previously, Ap<sub>4</sub>A has been shown to inhibit the initiation of DNA replication, but the mechanism behind how Ap<sub>4</sub>A has this effect remains unclear (Marriott et al., 2015). As Ap<sub>4</sub>Ns are also present in mammalian cells (Chapter 4), it is possible that these molecules may also influence DNA replication. As these molecules are similar in structure to Ap<sub>4</sub>A, whether they have a role in the inhibition of the initiation of DNA replication could shed light on how Ap<sub>4</sub>A performs this function. If these molecules can also induce inhibition of DNA replication it would demonstrate that the homobifunctional structure of Ap<sub>4</sub>A, required in some mechanisms, is not required for this process.

In this chapter, elucidation of the mechanism for the role of Ap<sub>4</sub>A in the inhibition of the initiation of DNA replication will be performed using a chemical biology approach.

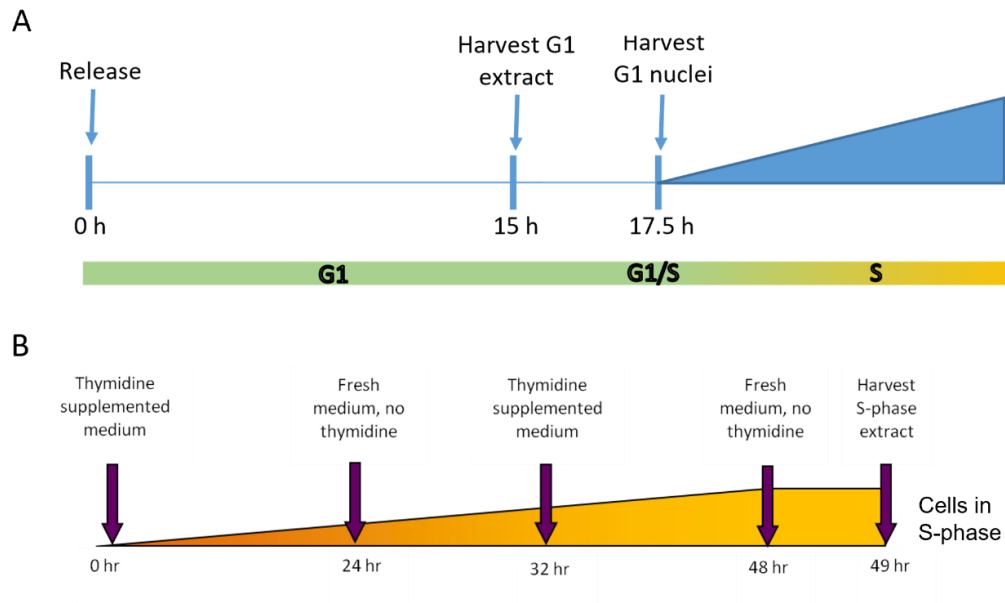
*In vitro* cell-free DNA replication assays will be used to determine if a homobifunctional structure is essential and whether changes to Ap<sub>4</sub>A structure by substitution with alternative nucleosides affects its activity. Secondly, a more detailed investigation of the inhibitory activity of Ap<sub>4</sub>A and other Ap<sub>4</sub>Ns during the initiation process will be investigated through monitoring changes in the nuclear localisation of Mcm2 and PCNA, two essential DNA replication components which act at different stages during the initiation process.

## **6.2 Preparation of nuclei and extracts for cell-free replication assays**

### **6.2.1 Synchronisation of late G1 nuclei, and G1- and S-phase extracts**

Prior to performing the cell-free replication assays, three components are required: synchronised late G1 nuclei, mid-G1 phase cytosolic extracts and S-phase cytosolic extracts, which are used as positive and negative controls respectively (Coverley et al., 2002). They are prepared using mouse fibroblast NIH3T3 cells, which are synchronised by contact inhibition and serum starvation at defined phases of the cell cycle (Fig. 6.1), and S-phase extracts from HeLa cells synchronised by double thymidine synchronisation (Fig. 6.1).





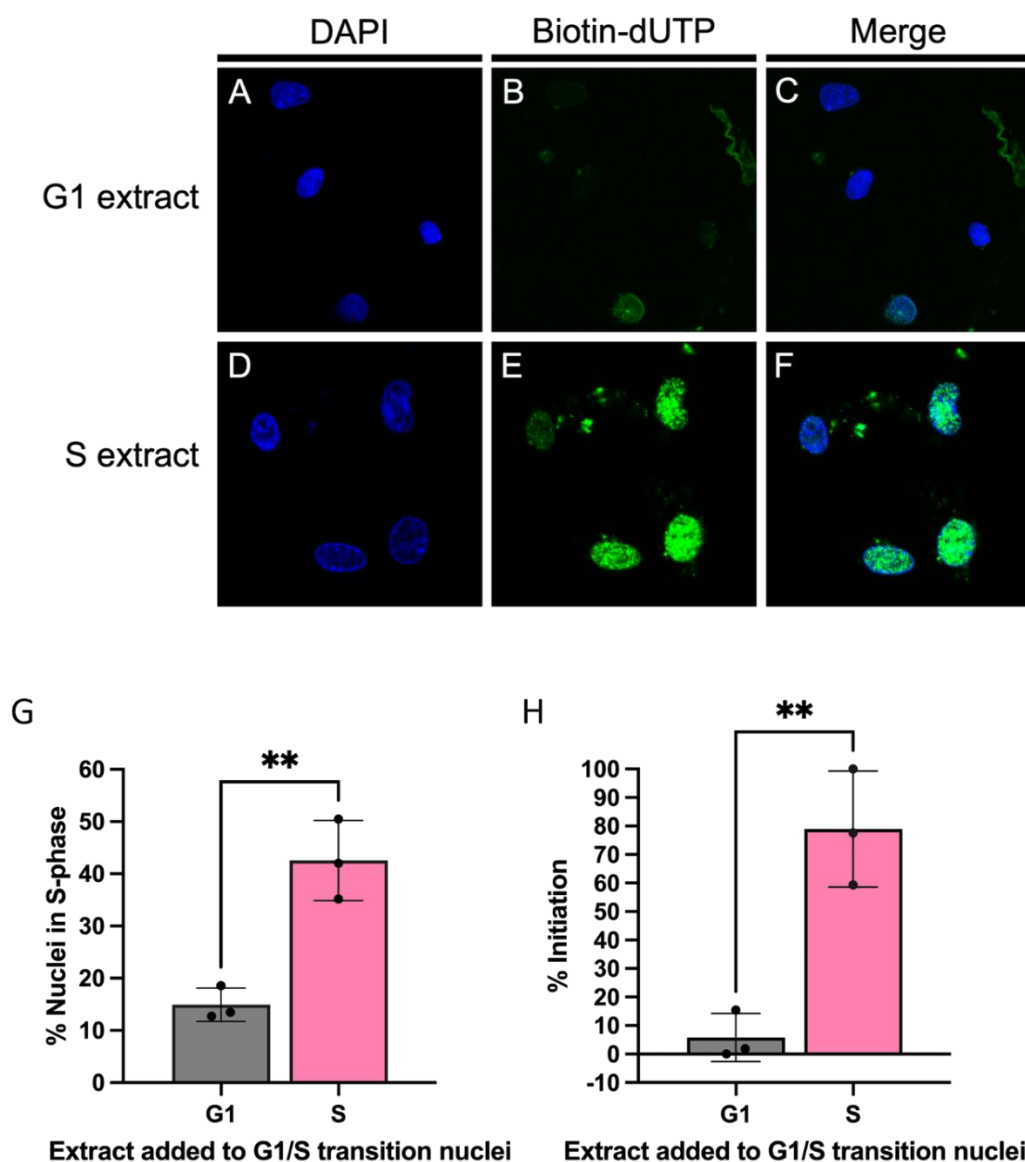
**Figure 6.1: Preparation of nuclei and extracts for cell-free replication assays. A:** Timeline for the release and harvesting of G1-phase extract (15 h) and late G1 nuclei (17.5 h). Cells were released from quiescence then harvested by scraping. Nuclei and extracts were harvested by Dounce homogenisation and frozen as pellets in liquid nitrogen. **B:** Timeline for synchronisation of HeLa cells for the preparation of cytosolic S-phase extracts following alternating thymidine-containing and thymidine-free incubations in supplemented medium.

This approach is used to generate mid-G1 cytosolic extracts isolated 15 hours after release from quiescence and late-G1 nuclei that are licensed to initiate DNA replication but require CDK activity to initiate DNA replication. This can be provided by the addition of either an S-phase HeLa cytosolic extract (Krude et al., 1997) or via recombinant cyclin A-CDK2 (Copeland et al., 2010; Coverley et al., 2002). Cytosolic extracts and nuclei are incubated with  $\text{MgCl}_2$ , a creatine phosphate/creatine phosphokinase (CPK) energy regenerating system, and all NTPs (except for dTTP and biotinylated-dUTP). Cells that are replicating are labelled with biotin-dUTP and imaged using Alexa Fluor-streptavidin. Replicating nuclei will incorporate the biotinylated-dUTP and are scored using fluorescence microscopy.

### 6.2.2 Testing nuclei batches and extracts in cell-free replication assays

Each batch of nuclei that are produced are tested extensively to ensure that there is a >2-fold increase in initiation between nuclei incubated in mid G1 extracts and S-phase extracts. There are several controls for each reaction: incubation in a G1 extract

provides the baseline for each batch of nuclei to determine the percentage of cells that are already in S-phase due to asynchrony after release from quiescence. The S-phase extracts are a composite of nuclei already in S-phase and the licensed nuclei that replicate after incubation with cyclin E-CDK2 and cyclin A-CDK2 present in the S-phase extracts. Consequently, the number of replicating nuclei is either given as a percentage of dUTP incorporating nuclei (Fig. 6.2 G) or as a percentage initiation, which uses a standardisation approach to normalise for the percentage of nuclei that initiate in S-phase cytosolic extracts (Coverley 2005; Marriott et al., 2015) (Fig. 6.2 H).

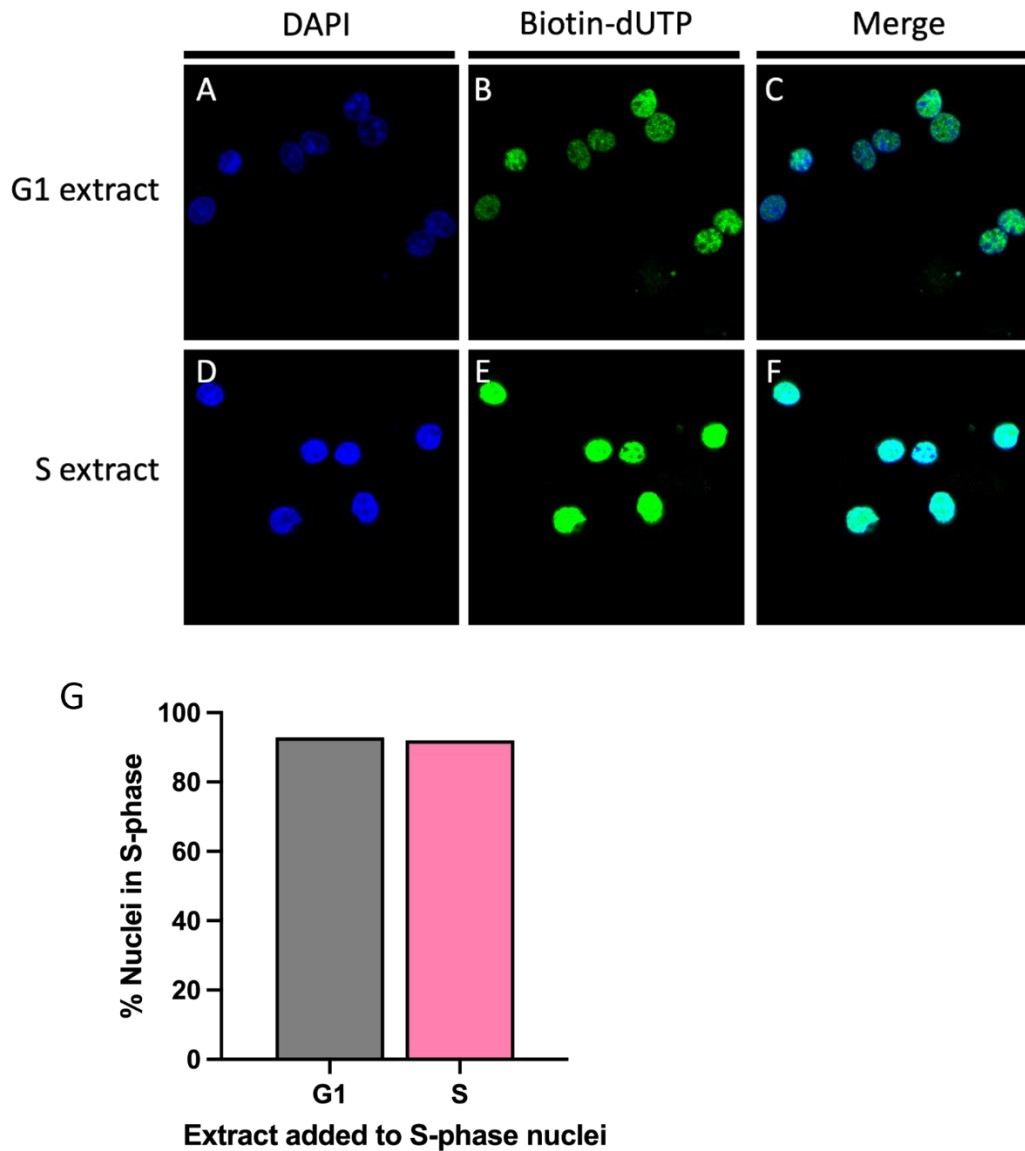


**Figure 6.2: Cell-free replication assays to test nuclei batches.** G1 nuclei were added to either G1- or S-phase extracts and labelled with a streptavidin conjugated Alexa Fluor (n=3). **A–F:** Fluorescence microscopy image comparing G1 nuclei added to either a G1-phase (A–C) or S-phase (D–F) extract. **G, H:** Replicating nuclei were scored as a percentage of the total nuclei counted (where n≥100 nuclei). Scored values for percentage of replicating nuclei in G1- and S-phase extracts (G) or standardised values to show the percentage initiation in the nuclei batch (H) were recorded. Statistical differences were measured using an unpaired, two-tailed t-test, assuming equal standard deviations. \*\*P≤0.01.

Images of the nuclei show a low level of incorporation of biotinylated-dUTP when G1 nuclei are added to G1 extract (Fig. 6.2 B and C). However, in the G1 nuclei added to

the S-phase extract there is a much higher level of biotinylated-dUTP incorporation (Fig. 6.2 E and F), which accumulates in a punctate pattern across the nucleus that is characteristic of DNA replication (Takebayashi et al., 2017) (Fig. 6.2 E and F). The percentage of nuclei in S-phase increased significantly ( $P=0.0045$ ), from a mean of 14.9% in G1 extract to 42.6% in S-phase extract. When the largest S-phase and smallest G1-phase percentage of nuclei in S-phase were scaled to 100% and 0% initiation, respectively, the mean percentage initiation increased by 73.2%, from 5.8% to 79.0% when G1 nuclei were added to an S-phase instead of G1-phase extract ( $P=0.0045$ ). This larger difference between the controls gives a larger window for analysis of the effect of the addition of any variables in the initiation of DNA replication. It is important to note that while the percentage initiation shows the number of nuclei that have initiated replication, it does not necessarily reflect the efficiency of DNA replication. DNA replication when nuclei are added to S-phase extracts is also more efficient, and this is demonstrated by the increased intensity of the signal in these nuclei (Fig. 6.2 B and E).

To confirm that the synchronisation process for the HeLa extract had successfully synchronised the HeLa cells into S-phase, nuclei were harvested from the cells at the same time as the extract. These nuclei were added to G1- and S-phase extracts in cell-free replication assays to compare the number of replicating nuclei (Fig. 6.3).



**Figure 6.3: Validation of the HeLa cell synchronisation process.** HeLa S-phase nuclei (n=1) were added to either G1- or S-phase extracts and labelled with a streptavidin-conjugated Alexa Fluor. **A-F:** Replicating nuclei in each extract were imaged and scored. **G:** Summary of raw values for the proportion of replicating nuclei when added to each extract.

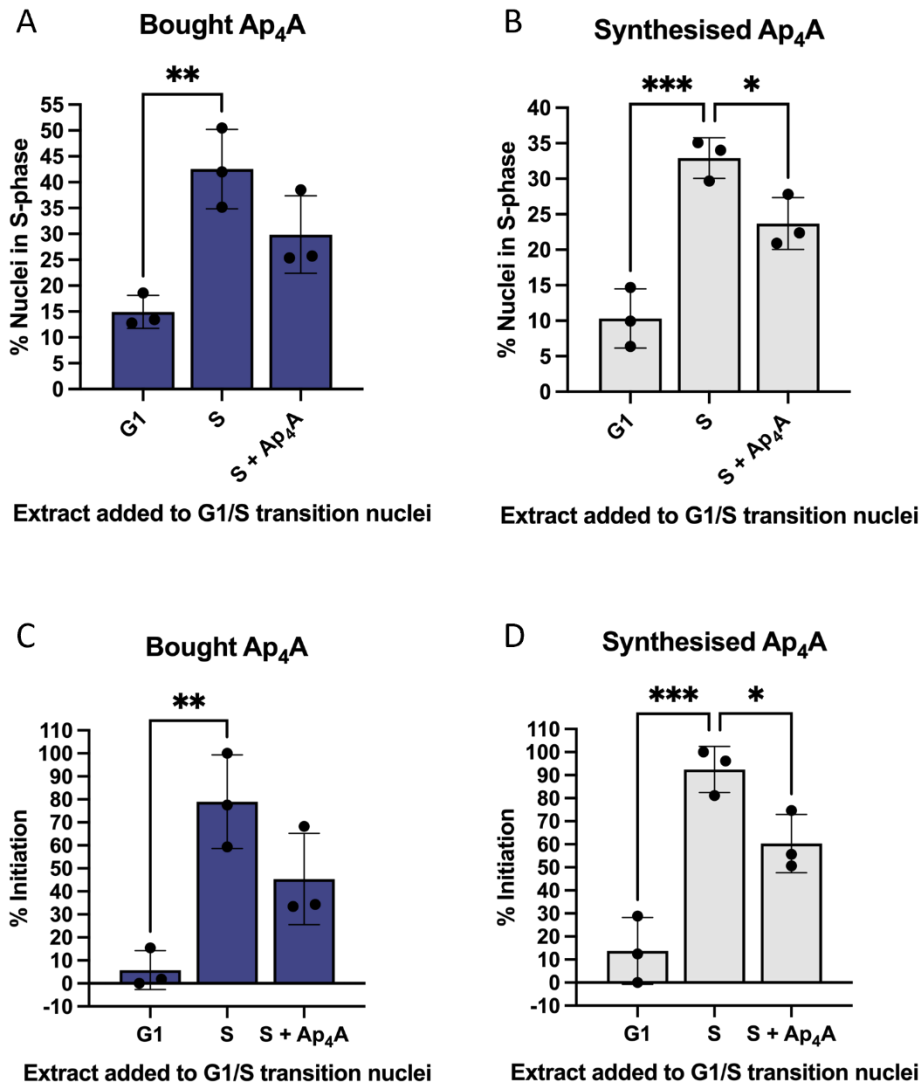
The punctate pattern of biotin-dUTP incorporation was present in the S-phase nuclei added to either G1-phase extract (Fig. 6.3 B and C) or S-phase extract (Fig. 6.3 E and F). In S-phase nuclei added to G1-phase extract, 92.9% of nuclei incorporated biotin-dUTP (Fig. 6.3 G). When added to an S-phase extract, although the intensity of the signal from the replicating nuclei increased (Fig. 6.3 B vs E), the proportion of replicating nuclei did not change (Fig. 6.3 G). These data are consistent with the synchronisation

of cells in S-phase after double thymidine treatment, as the S-phase nuclei have already initiated DNA replication and are in the elongation phase. As the replication machinery has already assembled, there is no longer a requirement for cyclin E-CDK2 and cyclin A-CDK2 to bypass the CDK threshold for the G1/S transition and replisome formation. Having established and characterised the synchronised cytosolic and nuclear fractions for cell-free initiation of DNA replication assays, the potential for dinucleoside tetraphosphates (Ap<sub>4</sub>N) in the regulation of this process could be performed.

### **6.3 Validation of the chemical synthesis of Ap<sub>4</sub>A and nuclei batches**

#### **6.3.1 Validation of the function of the chemically synthesised Ap<sub>4</sub>A**

In Chapter 3, production of the Ap<sub>4</sub>Ns was successfully validated by HRMS. To confirm that these molecules also retain their function, the effect of the synthesised Ap<sub>4</sub>A molecule on initiation of DNA replication was compared with commercial Ap<sub>4</sub>A (Sigma Aldrich) in cell-free DNA replication assays. The cell-free reactions used late G1 nuclei (17.5 hr) added to either G1 extract (15 hr), S-phase extract, or S-phase extract + 100 µM Ap<sub>4</sub>A and the proportion of replicating nuclei under each condition was scored (Fig. 6.4).



**Figure 6.4: Validation of the inhibitory activity of synthesised Ap<sub>4</sub>A relative to an Ap<sub>4</sub>A standard.** Cell-free replication assays were performed in the presence and absence of either an inhouse synthesised Ap<sub>4</sub>A or an Ap<sub>4</sub>A standard. **A, B:** Raw data showing the percentage of nuclei in S-phase under each condition. **C, D:** Scoring for percentage initiation for the Ap<sub>4</sub>A standard and the inhouse synthesised Ap<sub>4</sub>A relative to the S-phase only control. The standard Ap<sub>4</sub>A and synthesised Ap<sub>4</sub>A were tested in different nuclei batches. A one-way ANOVA was performed followed by Dunnett's multiple comparisons post hoc test to compare G1 and S+Ap<sub>4</sub>A with S-phase extract only within each batch of nuclei. \*P≤0.05; \*\*P≤0.01; \*\*\*P≤0.001.

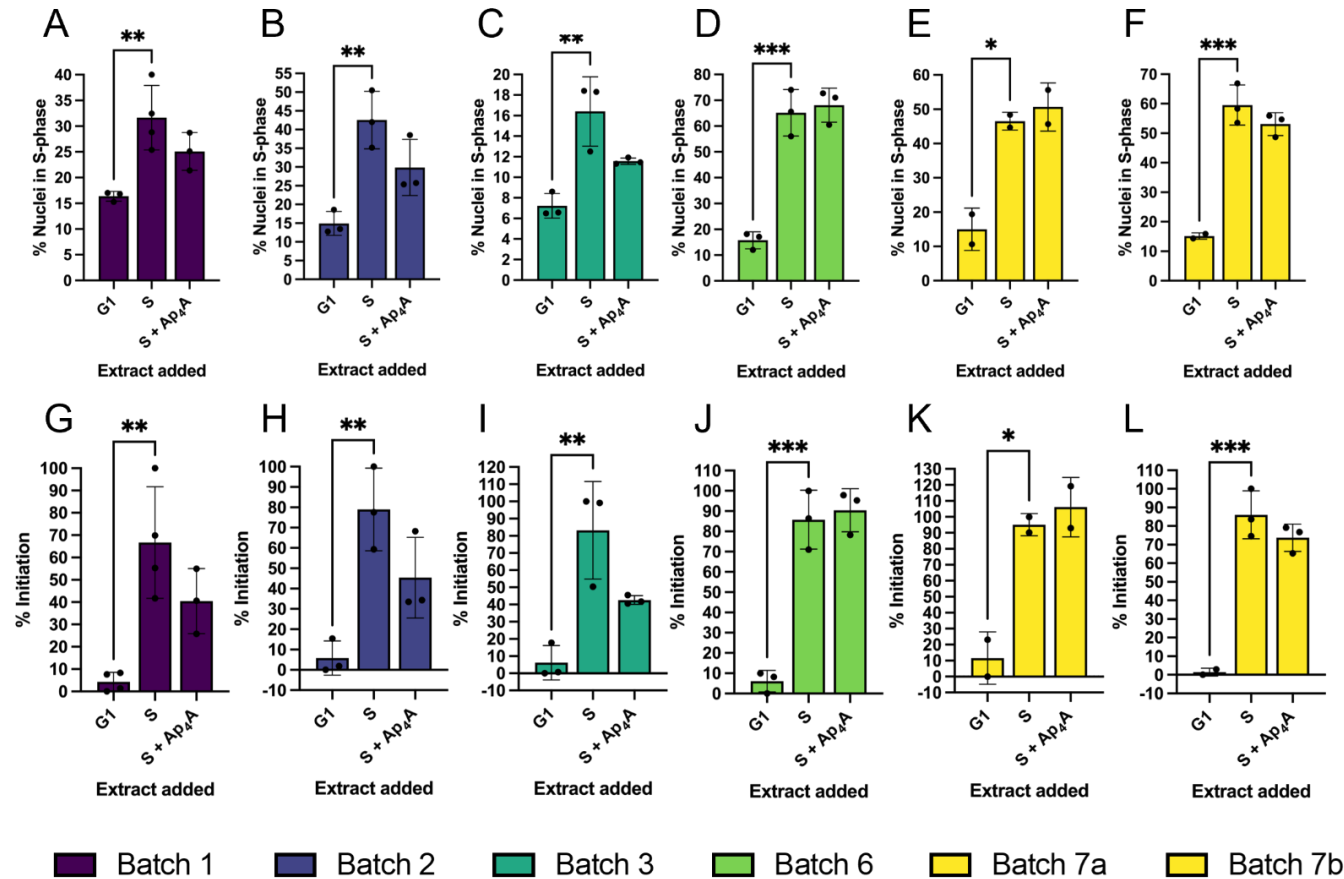
When either bought Ap<sub>4</sub>A or synthesised Ap<sub>4</sub>A were added to the reaction, there was a reduction in the number of nuclei in S-phase compared with when late G1 nuclei were added to S-phase extract only (Fig. 6.4 A and B). As the bought Ap<sub>4</sub>A and synthesised Ap<sub>4</sub>A were tested in different batches of nuclei, which have a different capacity for initiation into S-phase, the data were standardised so that the maximum S-phase

extract only value was set to 100% and the minimum G1-phase extract value was set to 0% (Fig. 6.4 C, D). This showed that when compared with the S-phase extract only sample, the mean percentage initiation was reduced by 33.6% and 32.1%, in the presence of the commercially sourced and synthesised Ap<sub>4</sub>A, respectively. This demonstrates a very similar response between the two Ap<sub>4</sub>A molecules and demonstrated that the Ap<sub>4</sub>A synthesis method produced a functional molecule.

### **6.3.2 Validation of the nuclei batches through assessing the effect of Ap<sub>4</sub>A**

As there is previous evidence that Ap<sub>4</sub>A inhibits the initiation phase of DNA replication, the individual batches of nuclei were verified based on their response to Ap<sub>4</sub>A in S-phase extract. The effect of Ap<sub>4</sub>A on the number of replicating nuclei and the percentage initiation was determined for each batch of nuclei (Fig. 6.5).





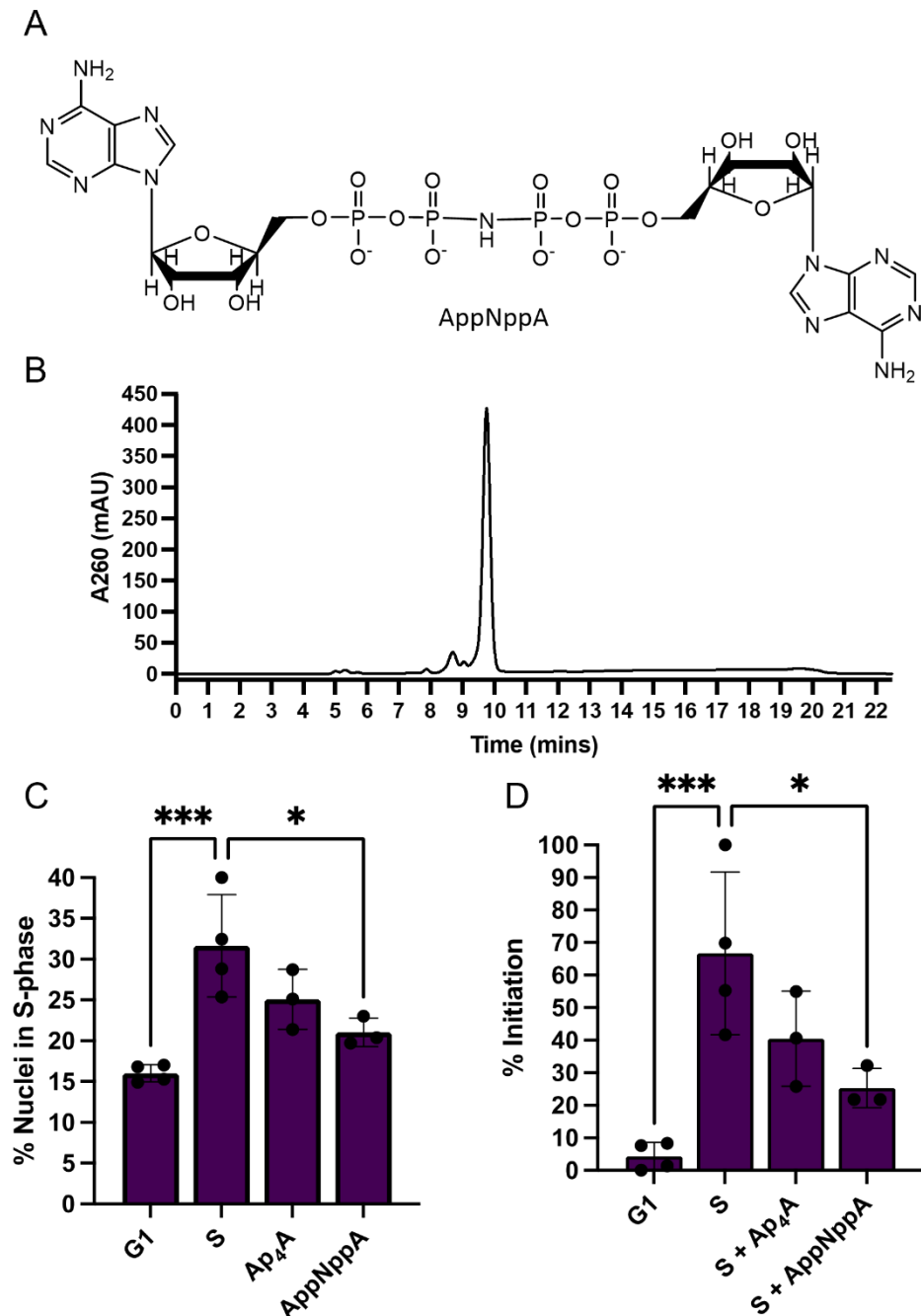
**Figure 6.5: The response of different nuclei batches to Ap<sub>4</sub>A.** The effect of Ap<sub>4</sub>A on the initiation of DNA replication was investigated using cell-free replication assays. **A-F:** the raw data set showing the percentage of nuclei in S-phase and **G-L:** the data set standardised so that the maximum S-phase only extract value is equal to 100% and the minimum G1-phase extract value is equal to 0%. A one-way ANOVA was performed followed by Dunnett's post-hoc test to compare the positive control (S) with G1 and S+Ap<sub>4</sub>A extracts within each batch of nuclei. NB: This data set is made up of a combination of bought Ap<sub>4</sub>A and synthesised Ap<sub>4</sub>A. \*P ≤ 0.05; \*\*P ≤ 0.01; \*\*\*P ≤ 0.001

Batches 1, 2 and 3 showed a consistent reduction in the mean percentage of nuclei in S-phase (-6.6% [batch 1], -12.7% [batch 2] and -4.8% [batch 3]) and percentage initiation (-26.2% [batch 1], -33.6% [batch 2], and -40.5% [batch 3]) when Ap<sub>4</sub>A was added to the S-phase extract compared with S-phase extract alone (Fig. 6.5). In some batches of nuclei, there was no response upon the addition of Ap<sub>4</sub>A (Fig. 6.5, batches 6 and 7). Therefore, these batches were omitted from future experiments. Interestingly, the nuclei batches that do not respond to Ap<sub>4</sub>A have the highest difference in percentage initiation when added to S-phase extract compared with G1-phase extract (Fig. 6.5 J–L versus G–I). The synchronisation approach used here has a degree of asynchrony and relies on harvesting nuclei within a 30-minute window to yield the highest quality nuclei. The temporal loading of factors during mid- to late-G1 phase includes the loading and activation of the helicases on to chromatin prior to assembly of the replisome. The precise mechanism involved in the inhibition of the initiation phase could include activation of the helicase or replisome assembly. The observation that nuclei that are potentially at a later stage of replication licensing and initiation in batches 6 and 7 suggests that Ap<sub>4</sub>A may not be as effective at inhibiting the DNA replication complexes involved at this later stage.

## **6.4 Investigating the effect of Ap<sub>4</sub>A modification on DNA replication inhibition**

### **6.4.1 Investigating the effect of the phosphate chain in Ap<sub>4</sub>A binding**

As a first step to investigating the role of Ap<sub>4</sub>A structure in its binding and activity, in DNA replication, we synthesised the AppNppA analogue. This analogue has a modified phosphate chain which has a nitrogen substituted for the central oxygen atom in the phosphate chain (Fig. 6.6 A). This modification maintains a similar length to the four-phosphate chain in Ap<sub>4</sub>A but makes the AppNppA analogue more stable compared with Ap<sub>4</sub>A (Viatchenko-Karpinski et al., 2016). After its synthesis, AppNppA was purified by HPLC and added to late G1 nuclei in an S-phase extract in cell-free replication assays (Fig. 6.6).



Extract added to G1/S transition nuclei    Extract added to G1/S transition nuclei

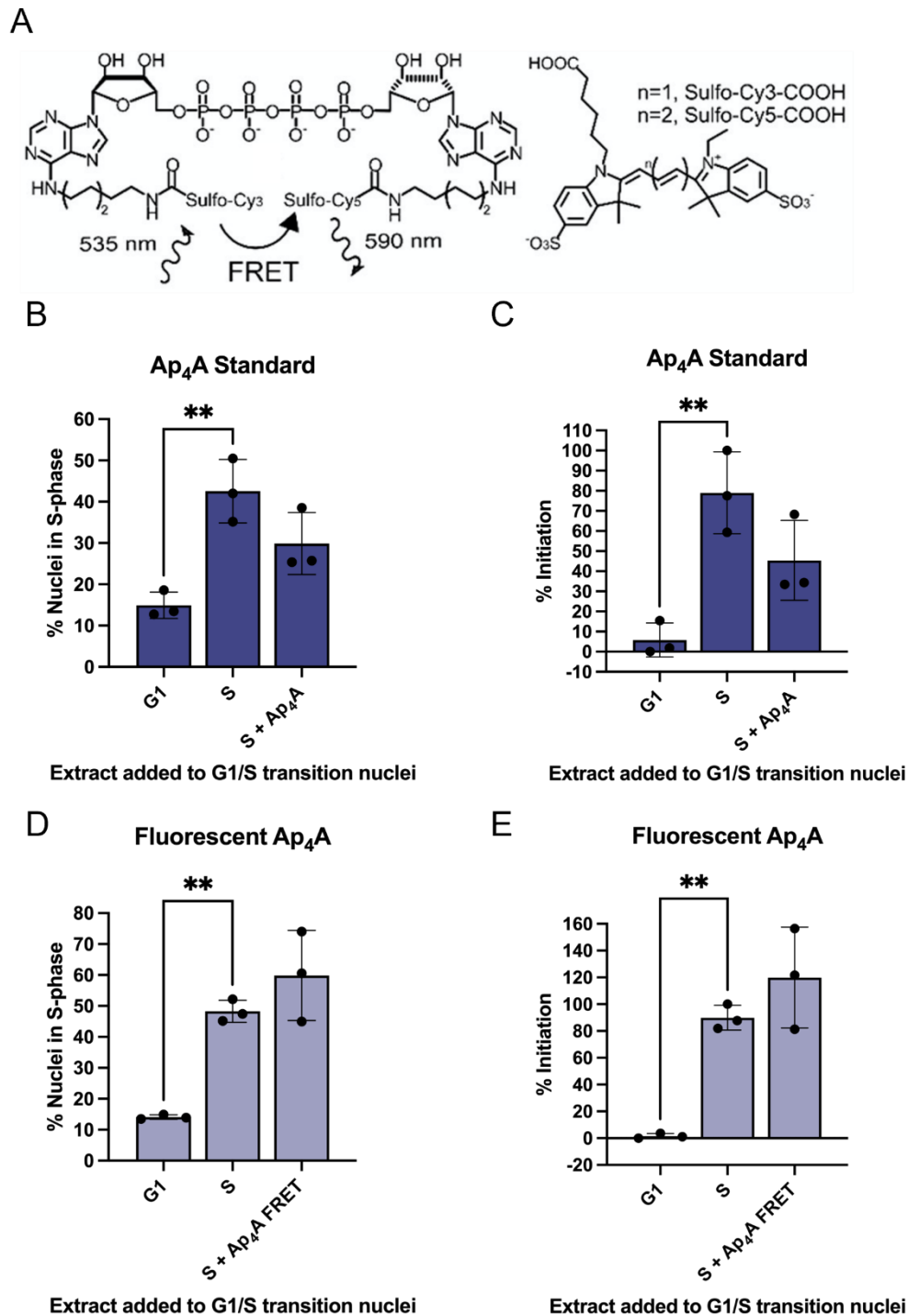
**Figure 6.6: The effect of AppNppA on the initiation of DNA replication.** Cell-free replication assays were performed, and Ap<sub>4</sub>A or AppNppA were incubated with late-G1 nuclei in S-phase extract. **A:** The structure of AppNppA. **B:** Purified AppNppA after chemical synthesis, measured at a wavelength of 260 nm. **C:** Raw data showing the percentage nuclei in S-phase after addition of Ap<sub>4</sub>A and AppNppA compared with S-phase extract only. **D:** Data from (C) scaled so that the largest percentage initiation in the S-phase only extract is equal to 100%, and the minimum level of background replicating nuclei in G1-phase extract is equal to 0%. A one-way ANOVA was performed followed by Dunnett's multiple comparisons test to compare G1, S+Ap<sub>4</sub>A, and S+AppNppA with S-phase extract only within the nuclei batch. \*P≤0.05; \*\*\*P≤0.001.

AppNppA retains the same phosphate chain length as Ap<sub>4</sub>A, but a nitrogen replaces the central oxygen atom in the phosphate chain (Fig. 6.6 A). After synthesis, AppNppA was successfully purified by HPLC, and its concentration was determined (Fig. 6.6 B). Addition of AppNppA to late G1 nuclei in S-phase extract significantly reduced the percentage of nuclei in S-phase (adjusted P value = 0.0140) and the percentage of DNA replication initiation (adjusted P value = 0.0155) compared with late G1 nuclei added to S-phase extract alone (Fig. 6.6 C and D). Although the Ap<sub>4</sub>A scoring was not significant (adjusted p value = 0.1282 and 0.1260, respectively), a similar trend was followed, suggesting the AppNppA inhibits the initiation of DNA replication to a similar extent, if not more effectively, than Ap<sub>4</sub>A (Fig. 6.6 C and D).

#### **6.4.2 Investigating the effect of fluorescent Ap<sub>4</sub>A on DNA replication initiation**

Ap<sub>4</sub>A-Förster resonance energy transfer (FRET) has previously been used to investigate the activity of inhibitors of Ap<sub>4</sub>A phosphorylase Rv2613c from *M. tuberculosis* (Götz et al., 2017). In the Ap<sub>4</sub>A-FRET molecules, a fluorophore is added to the amino group on each adenosine moiety (Götz et al., 2017). These fluorophores can undergo FRET when they are part of the same molecule, because the fluorophores are in close proximity, but not after Ap<sub>4</sub>A hydrolysis (Götz et al., 2017). Ap<sub>4</sub>A-FRET can therefore provide a measure of Ap<sub>4</sub>A hydrolysis.

As Ap<sub>4</sub>A-FRET relies on the addition of bulky groups to the Ap<sub>4</sub>A molecule (Fig. 6.7 A), it may interfere with Ap<sub>4</sub>A signalling in contexts where the adenosine moieties are important for binding. Therefore, initially this molecule was used to determine whether Ap<sub>4</sub>A-FRET retained the functionality of Ap<sub>4</sub>A in the inhibition of DNA replication initiation (Fig. 6.7). This was done with the view that if Ap<sub>4</sub>A-FRET retained the inhibitory activity of Ap<sub>4</sub>A, it might provide a method for analysis of the Ap<sub>4</sub>A mechanism and whether hydrolysis is required for its activity.



**Figure 6.7: The effect of Ap<sub>4</sub>A-FRET on the initiation of DNA replication.** **A:** The structure of Ap<sub>4</sub>A-FRET, adapted from Götz et al., 2017. **B–E:** Ap<sub>4</sub>A (B and C) or Ap<sub>4</sub>A-FRET (D and E) was added to late G1 nuclei and S-phase extract in a cell-free replication assay. The percentage of nuclei in S-phase (B and D) and the percentage initiation (C and E) were recorded. A one-way ANOVA was performed followed by Dunnett's multiple comparisons test to compare G1 and S+Ap<sub>4</sub>A extracts with S-phase extract only within the nuclei batch. \*\*P≤0.01. FRET, Förster resonance energy transfer.

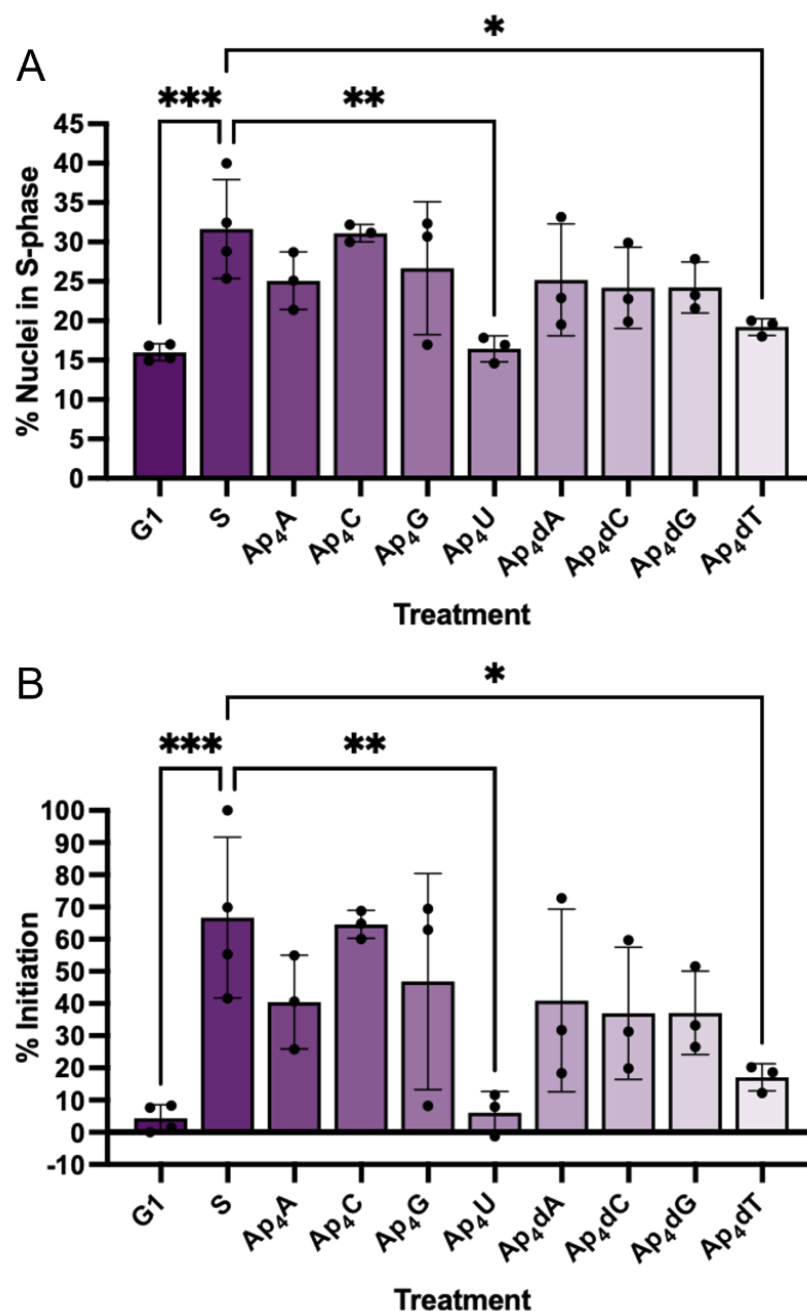
When added to late G1 nuclei in S-phase extract, Ap<sub>4</sub>A reduced the mean percentage of nuclei in S-phase (Fig. 6.7 B) and the mean percentage initiation (Fig. 6.7 C). The Ap<sub>4</sub>A-FRET molecules did not follow this trend (Fig. 6.7 D and E). Results from the cell-free experiments with Ap<sub>4</sub>A-FRET suggested that, if anything, the percentage initiation was increased compared with the S-phase only extract. This data suggests that the modifications to the Ap<sub>4</sub>A molecule necessary to produce Ap<sub>4</sub>A-FRET prevent the inhibitory function of Ap<sub>4</sub>A. This in turn suggests that the structure of the adenosine moieties may be important for Ap<sub>4</sub>A function and that only certain positions may be acceptable for modification without affecting Ap<sub>4</sub>A activity. Therefore, the next step was to determine whether a more subtle change to the adenosine structure would also prevent Ap<sub>4</sub>A activity.

## **6.5 Investigating the role of Ap<sub>4</sub>Ns in inhibiting the initiation of DNA replication**

Initiation of DNA replication is inhibited by Ap<sub>4</sub>A. However, currently there is no evidence surrounding whether the other Ap<sub>4</sub>N molecules have a regulatory role in this process. Several of the Ap<sub>4</sub>N molecules have a very similar structure to Ap<sub>4</sub>A, especially Ap<sub>4</sub>G, which also contains the purine ring and only differs from Ap<sub>4</sub>A by the addition of a carboxyl group at position C6, and the position of the amino group at C2 rather than C6. Interestingly, Gp<sub>4</sub>G does not inhibit DNA replication (Marriott et al., 2015), demonstrating that at least one adenosine nucleotide is required for Ap<sub>4</sub>A function. Therefore, investigating whether the Ap<sub>4</sub>Ns retained the inhibitory function of Ap<sub>4</sub>A could provide insight into whether the specific structure with two adenosine groups is essential for inhibition of DNA replication initiation to occur.

### **6.5.1 Initial investigation into the activity of Ap<sub>4</sub>Ns in the inhibition of the initiation of DNA replication**

After synthesis and HRMS verification of Ap<sub>4</sub>A, Ap<sub>4</sub>C, Ap<sub>4</sub>G, Ap<sub>4</sub>U, Ap<sub>4</sub>dA, Ap<sub>4</sub>dC, Ap<sub>4</sub>dG and Ap<sub>4</sub>dT in Chapter 3, their activity during the initiation of DNA replication was assessed using cell-free replication assays. A summary of the results from the first batch (batch 1) of nuclei is shown in Figure 6.8.



**Figure 6.8: Initiation of replication after addition of each Ap<sub>4</sub>N to the first batch of nuclei.** Each Ap<sub>4</sub>N was tested in a cell-free DNA replication assay and the number of replicating nuclei were scored. **A:** The raw number of replicating nuclei under each condition. **B:** The percentage initiation under each condition calculated after scaling the highest S-phase extract only value to 100% and the lowest G1-phase extract only value to 0%. NB: the process for quantifying Ap<sub>4</sub>N concentration was subsequently updated, resulting in slight alterations to the concentrations used in this experiment. Updated concentrations are: Ap<sub>4</sub>A (100  $\mu$ M), Ap<sub>4</sub>C (98.4  $\mu$ M), Ap<sub>4</sub>G (67.1  $\mu$ M), Ap<sub>4</sub>U (92.4  $\mu$ M), Ap<sub>4</sub>dA (110.9  $\mu$ M), Ap<sub>4</sub>dC (100.8  $\mu$ M), Ap<sub>4</sub>dG (104.0  $\mu$ M) and Ap<sub>4</sub>dT (94.2  $\mu$ M). The difference in the percentage of nuclei in S phase and percentage initiation was analysed using a one-way ANOVA with Dunnett's multiple comparisons post hoc test. \* $P \leq 0.05$ ; \*\* $P \leq 0.01$ ; \*\*\* $P \leq 0.001$ .

In this first set of nuclei, the mean percentage of nuclei in S-phase and mean percentage initiation was reduced for all Ap<sub>4</sub>Ns except for Ap<sub>4</sub>C (Fig. 6.8). This effect was significant upon addition of either Ap<sub>4</sub>U (adjusted P value = 0.0024) or Ap<sub>4</sub>dT (adjusted P value = 0.0145) (Fig. 6.8 B). Interestingly, apart from the difference between Ap<sub>4</sub>C and Ap<sub>4</sub>dC, the Ap<sub>4</sub>dNs appeared to reduce DNA replication initiation to a similar extent as their corresponding Ap<sub>4</sub>Ns (Fig. 6.8 B). This suggests that the oxygen removed from the ribose of each molecule does not have a significant effect on substrate binding affinity.

This data enabled the streamlining of future experiments to include investigation of the ribosyl dinucleoside tetraphosphates. As the Ap<sub>4</sub>dNs appeared to behave similarly to the Ap<sub>4</sub>Ns, they were not studied further. The exception to this was Ap<sub>4</sub>dT, which we decided to study further based on its significant inhibition of DNA replication.

### **6.6 Detailed investigation of the effect of Ap<sub>4</sub>N in the initiation of DNA replication**

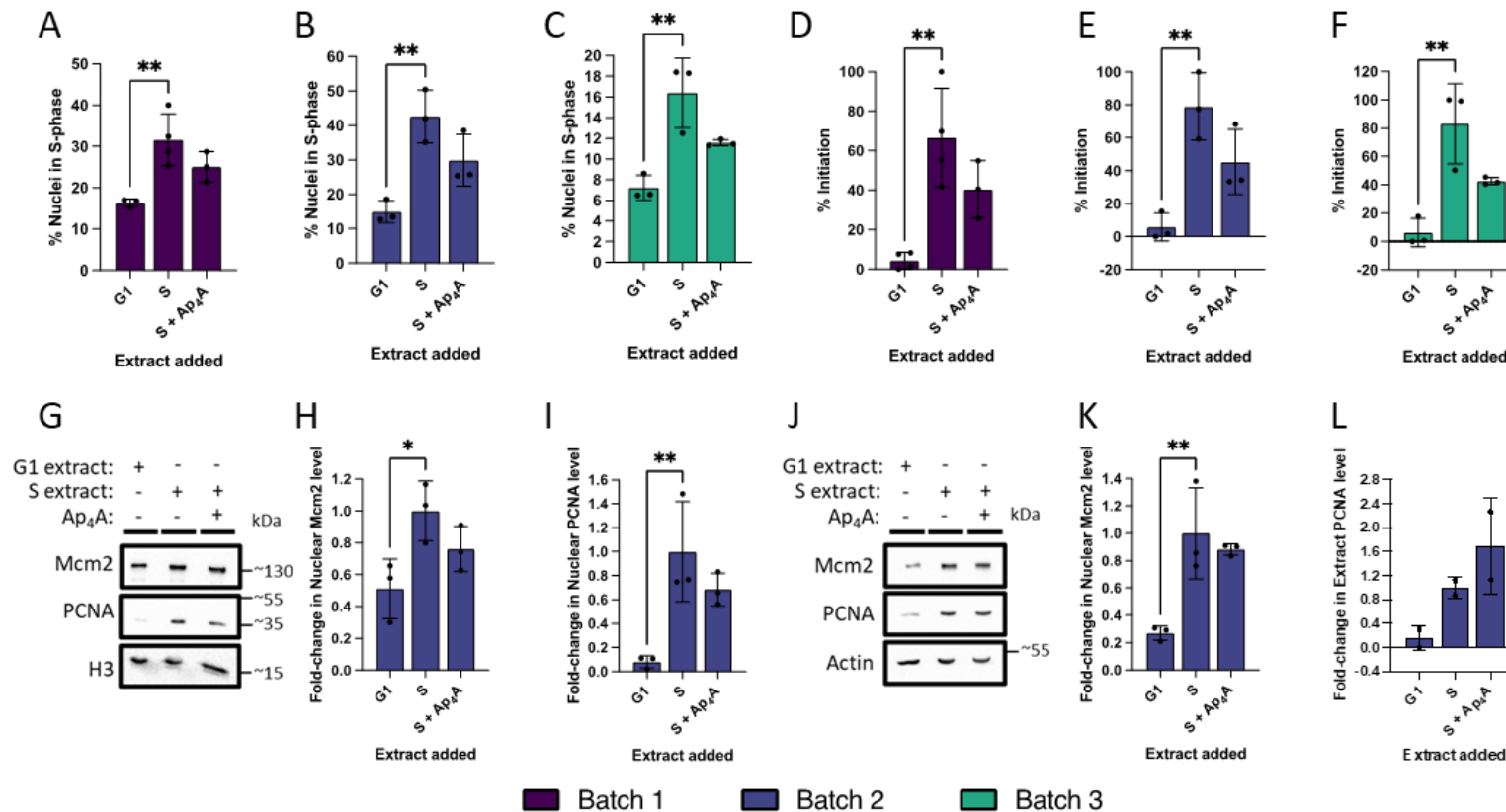
As most of the Ap<sub>4</sub>Ns were able to inhibit DNA replication in the cell-free assays, it suggested that the process for inhibiting DNA replication does not rely on a homobifunctional structure. To further elucidate this process, repeats were performed for all Ap<sub>4</sub>Ns in other validated batches of nuclei. Furthermore, to explore this process, the cell-free system was used to investigate the nuclear levels of proteins involved in the DNA replication process to see if their localisation to chromatin was affected by the Ap<sub>4</sub>Ns.

The MCM complex is key to the initiation of DNA replication as it is loaded onto chromatin along with Cdc6 and ORC, forming the pre-replication complex. DDK later binds chromatin bound Mcm2, Mcm4 and Mcm6, contributing to the initiation of DNA replication (Cheng et al., 2022; Saleh et al., 2022; Tsuji et al., 2006). Proliferating cell nuclear antigen (PCNA) is not recruited until later in the DNA replication process and is important in the loading of DNA polymerases needed for DNA synthesis. Nuclear Mcm2 and PCNA levels were measured by Western blotting for each nuclei batch in triplicate and the intensity of the bands was quantified.



### **6.6.1 Ap<sub>4</sub>A inhibits the initiation of DNA replication and the recruitment of replication proteins to chromatin**

To investigate the role of Ap<sub>4</sub>A in DNA replication further, Western blotting was performed for the batches of nuclei that had been validated by fluorescence scoring in Figure 6.5. The Western blots were quantified and the average intensity for the late-G1 nuclei added to S-phase extract was scaled to one and the average intensities for the Ap<sub>4</sub>A-treated samples are shown as a proportion of this (Fig. 6.9).

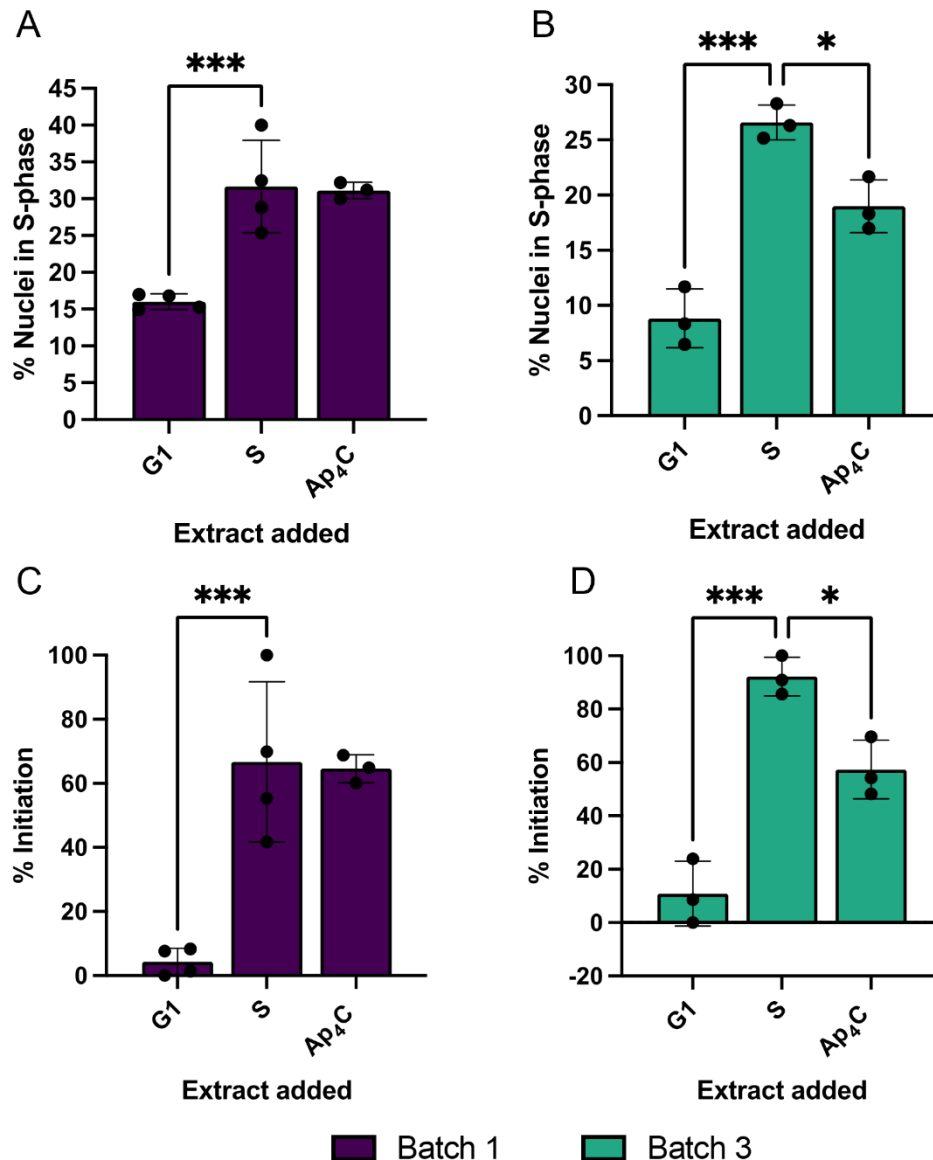


**Figure 6.9: The role of Ap<sub>4</sub>A in inhibiting DNA replication.** Cell-free replication assays were performed in the presence of Ap<sub>4</sub>A. **A–C:** The raw number of nuclei in S-phase. **D–F:** the standardised values to give percentage initiation. **G–L:** Representative Western blotting data for changes in Mcm2 and PCNA levels localising to the chromatin (**G–I**) and soluble (**J–L**) fractions. Statistical analysis for all panels was performed using a one-way ANOVA followed by Dunnett's multiple comparisons test to compare G1 and S+Ap<sub>4</sub>A extracts with S-phase extract only within each batch of nuclei. \*P≤0.05; \*\*P≤0.01. For all panels n=3, except for panel L where one repeat was removed due to poor transfer (n=2). Please note that no ladder is available for the Mcm2 and PCNA data set shown in J–L.

The percentage of nuclei in S-phase when added to an S-phase extract varies between nuclei batches, with the mean percentage of nuclei in S-phase after addition of the S-phase extract ranging from 16.4–42.6% (Fig. 6.9 A–C). Therefore, the results were scaled so that the effect of Ap<sub>4</sub>A was compared with the S-phase extract only sample in each batch, and the percentage change could be compared. Across the three nuclei batches, the percentage initiation is reduced by 26.2–40.5% in the Ap<sub>4</sub>A-treated samples compared with the S-phase extract alone (Fig. 6.9 D–F). Although the data were not significant, there was a trend suggesting a reduction in the mean level of chromatin-localised Mcm2 (24%) and PCNA level (34%) when Ap<sub>4</sub>A was incubated with the S-phase extract compared with S-phase extract only (Fig. 6.9 G–I). This reduction was not seen in the soluble extract fraction for either Mcm2 (+2%) or PCNA (+70%) (Fig. 6.9 J–L), which demonstrates that the reductions seen in the chromatin fractions are not due to a loss of protein stability. The increase seen in Figure 6.9 L may be due to the reduction in chromatin binding of PCNA, shown in Figure 6.9 I.

#### **6.6.2 The effect of Ap<sub>4</sub>C on DNA replication and the localisation of replication proteins to chromatin remains unclear**

To determine whether the reduction in Mcm2 and PCNA levels was specific to Ap<sub>4</sub>A, we also investigated the effect of the other Ap<sub>4</sub>Ns on nuclear protein localisation. First, the batches of nuclei that were responsive to Ap<sub>4</sub>A were incubated with Ap<sub>4</sub>C in cell-free replication assays (Fig. 6.10). Changes in the percentage of nuclei in S-phase and percentage initiation were recorded (Fig. 6.10).



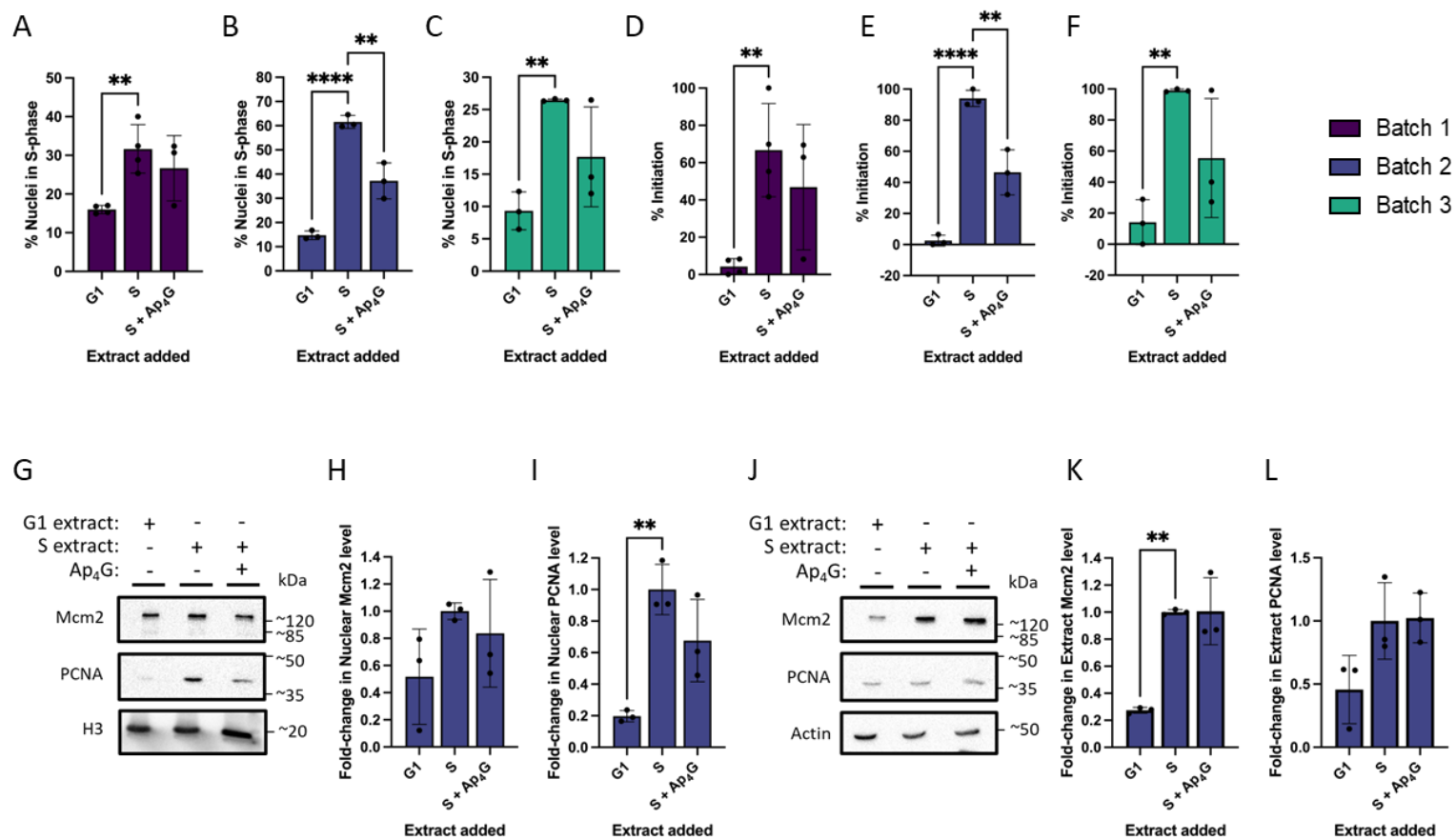
**Figure 6.10: The role of Ap<sub>4</sub>C in the inhibition of DNA replication.** Cell-free replication assays were performed in the presence of Ap<sub>4</sub>C. **A,B:** The raw number of nuclei in S-phase using two batches of nuclei and **C,D:** the standardised values to give the percentage initiation in the presence and absence of Ap<sub>4</sub>C. A one-way ANOVA was performed followed by Dunnett's multiple comparisons post-hoc test to compare G1 and S+Ap<sub>4</sub>C with S-phase extract only within each batch of nuclei. \* $P \leq 0.05$ ; \*\*\* $P \leq 0.001$ .

There were mixed results for the role of Ap<sub>4</sub>C in the initiation of DNA replication. In the first batch, the mean percentage of initiating nuclei was reduced by just 2.1% in the nuclei treated with Ap<sub>4</sub>C compared with the nuclei added to S-phase extract only (Fig. 6.10 C). This suggested that Ap<sub>4</sub>C may not influence the initiation of DNA replication. However, results from the batch 3 nuclei showed a statistically significant

reduction in the percentage of nuclei in S-phase and percentage initiation (adjusted P value = 0.0110 for both), with a mean reduction of 34.8% when Ap<sub>4</sub>C was incubated with the S-phase extract compared with S-phase extract only (Fig. 6.10 B and D). Therefore, results from the Ap<sub>4</sub>C data set suggested that there were inconsistencies related to the nuclei batch.

### **6.6.3 Addition of Ap<sub>4</sub>G to S-phase extract shows a trend in reducing percentage initiation and reducing accumulation of replication proteins on chromatin**

Of the dinucleotide tetraphosphates tested, Ap<sub>4</sub>G is the closest to Ap<sub>4</sub>A in structure, with an oxygen replacing the amino group on carbon 6 of the adenine ring, and the amino group instead being present on C2. Therefore, we investigated whether the small structural changes to Ap<sub>4</sub>G affected its ability to inhibit DNA replication or reduce localisation of DNA replication proteins to chromatin (Fig. 6.11).



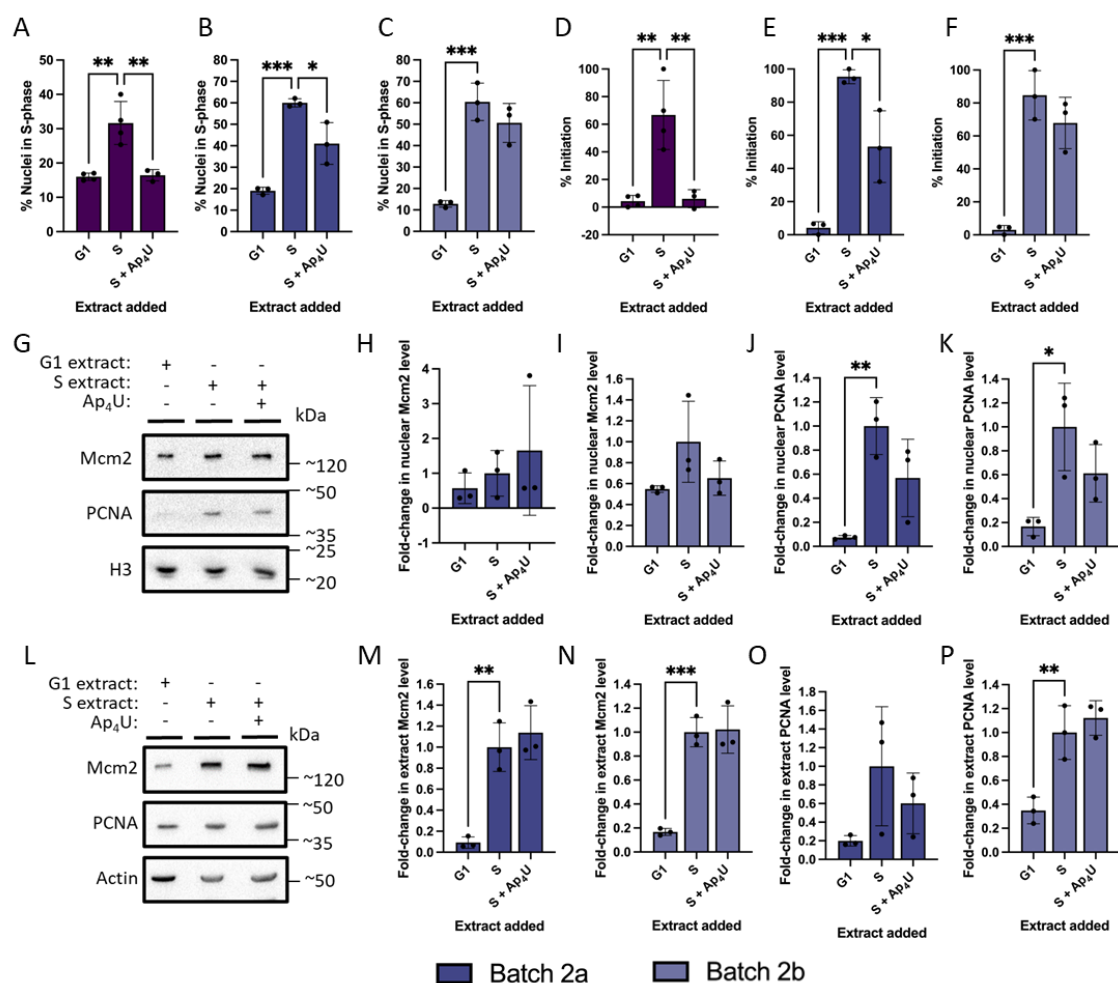
**Figure 6.11: The role of Ap<sub>4</sub>G in inhibiting DNA replication.** Cell-free replication assays were performed in the presence of Ap<sub>4</sub>G. **A–C:** The raw number of percentage nuclei in S-phase. **D–F:** The standardised values to give percentage initiation. **G–L:** Representative Western blotting data for changes in Mcm2 and PCNA levels localising to chromatin (G–I) and soluble (J–L) fractions. Statistical analysis for all panels was performed using a one-way ANOVA followed by Dunnett's multiple comparisons test to compare G1 and S+Ap<sub>4</sub>G extracts with S-phase extract only within each batch of nuclei. \*P≤0.05; \*\*P≤0.01; \*\*\*\*P≤0.0001.

For all nuclei batches, the increase in the percentage of S-phase nuclei was statistically significant when the late G1 nuclei were added to an S-phase extract compared with a G1-phase extract (adjusted P value = 0.0089, <0.0001 and 0.0082 for batches 1, 2 and 3, respectively) (Fig. 6.11 A–C). However, as demonstrated for Ap<sub>4</sub>A and Ap<sub>4</sub>C there was variation between the extent to which each batch of nuclei initiated into S-phase when added to S-phase extract compared with G1-phase extract (Fig. 6.11 A–C), so the data sets were standardised to show percentage initiation to enable a comparison between the different batches (Fig. 6.11 D–F). Statistical analysis of the standardised data showed that there was a significant reduction in the percentage initiation of batch 2 nuclei when Ap<sub>4</sub>G was added to the S-phase extract compared with S-phase extract alone (Fig. 6.11 E; adjusted P value = 0.0012). In this batch, the reduction in the mean percentage initiation was 47.6%. In batch 1 and 3, although there wasn't a significant reduction in mean percentage initiation, percentage initiation was still reduced by 19.9% and 43.5%, respectively, in the Ap<sub>4</sub>G-treated nuclei compared with the S-phase extract only nuclei.

There was also an overall reduction in the localisation of Mcm2 and PCNA to chromatin in the samples containing Ap<sub>4</sub>G compared with those in S-phase extract alone (Fig. 6.11 G–I). As with Ap<sub>4</sub>A, this reduction was greater for PCNA (32%) than for Mcm2 (16%) (Fig. 6.11 H and I). Neither Mcm2 nor PCNA levels were affected in the soluble fraction, as the relative abundance of each of these changed by +1% and +2%, respectively (Fig. 6.11 K and L). This demonstrates that the pool of Mcm2 and PCNA is available, suggesting that any change in nuclear protein abundance is likely due to an inhibition of the re-localisation of these proteins.

#### **6.6.4 Ap<sub>4</sub>U generally shows a trend in reducing percentage initiation and reducing localisation of replication proteins to chromatin**

In the initial data set, Ap<sub>4</sub>U showed the strongest inhibition of the initiation of DNA replication (Fig. 6.8). Therefore, further analysis was performed to determine whether this strong capacity for inhibition was consistent across different nuclei batches, and to investigate whether Ap<sub>4</sub>U affected nuclear localisation of the replication proteins (Fig. 6.12).



**Figure 6.12: The role of Ap<sub>4</sub>U in inhibiting DNA replication.** Cell-free replication assays were performed in the presence and absence of Ap<sub>4</sub>U. **A–C:** The raw number of nuclei in S-phase and **D–F:** the standardised values to give percentage initiation in the presence and absence of Ap<sub>4</sub>U. **G–P:** Representative Western blotting data for changes in Mcm2 and PCNA levels on chromatin (**G–K**) and soluble (**L–P**) fractions for two sets of data using the same nuclei batch with different Ap<sub>4</sub>U syntheses. Statistical analysis was performed using a one-way ANOVA followed by Dunnett's multiple comparisons test to compare G1 and S+Ap<sub>4</sub>U with S-phase extract only within each set of results. \*P≤0.05; \*\*P≤0.01; \*\*\*P≤0.001.

In the raw data set, the percentage of nuclei in S-phase increased significantly for batch 1 and 2 nuclei when added to an S-phase extract instead of G1-phase extract (Fig. 6.12 A–C). Batch 3 was excluded from this data set as there was no significant increase in the percentage of nuclei in S-phase when added to the S-phase extract compared with the G1-phase extract. There was a large amount of variation in the effect of Ap<sub>4</sub>U across the different nuclei batches (Fig. 6.12 A–C). There was a significant reduction in the percentage initiation in batch 1 (adjusted P value = 0.0019)

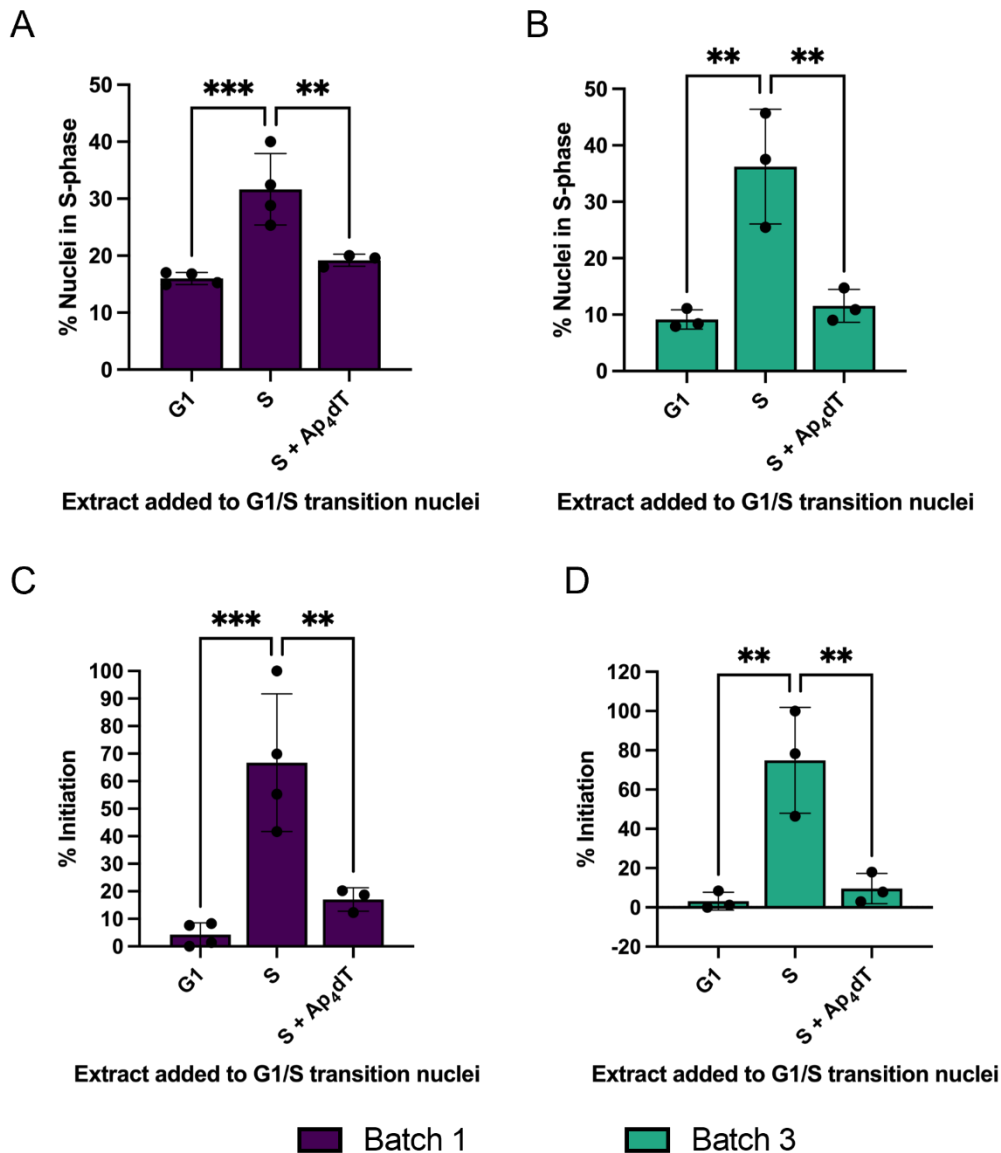


and batch 2a (adjusted P value = 0.0125), in the Ap<sub>4</sub>U-treated compared with the nuclei added to S-extract alone, with a reduction in mean percentage initiation of 60.6% and 42.2%, respectively. In batch 2b the mean percentage initiation was reduced by 16.8%, so although this wasn't statistically significant it followed a similar trend to the other data sets.

Generally, this reduced percentage initiation correlated with reduced Mcm2 and PCNA levels in the chromatin fraction (Fig. 6.12 H–K), except for a single data point in Figure 6.12 H, which showed an increase in Mcm2 levels. Because of this, the mean Mcm2 level in the chromatin fraction increased by 66% in batch 2a but decreased by 35% in batch 2b. Mean PCNA levels decreased more consistently, by 39% and 43% in batch 2a and 2b, respectively. Therefore, the greater reduction in PCNA levels compared with Mcm2 levels was consistent with the data sets for Ap<sub>4</sub>A and Ap<sub>4</sub>G in Figures 6.9 and 6.11.

#### **6.6.5 Ap<sub>4</sub>dT shows a strong inhibition of DNA replication initiation which is not explained by changes in nuclear replication protein level**

In the initial data set, Ap<sub>4</sub>dT was one of the most effective molecules at inhibiting the initiation of DNA replication, alongside Ap<sub>4</sub>U. Interestingly, these molecules have a very similar structure, with the only difference being the addition of a methyl group at position 5 on the pyrimidine ring of Ap<sub>4</sub>dT, and loss of a hydroxyl group from the ribose to form the deoxyribose structure. While Ap<sub>4</sub>U still reduced DNA replication initiation in most other batches of nuclei, this effect was not as strong as in batch 1, where the percentage initiation was comparable to the background level of nuclei replicating when added to G1-phase extract (Fig. 6.12 A–F). The effect of Ap<sub>4</sub>dT was therefore determined in a second batch of Ap<sub>4</sub>A-verified nuclei to determine whether Ap<sub>4</sub>dT was as effective in other nuclei batches as in batch 1 (Fig. 6.13)

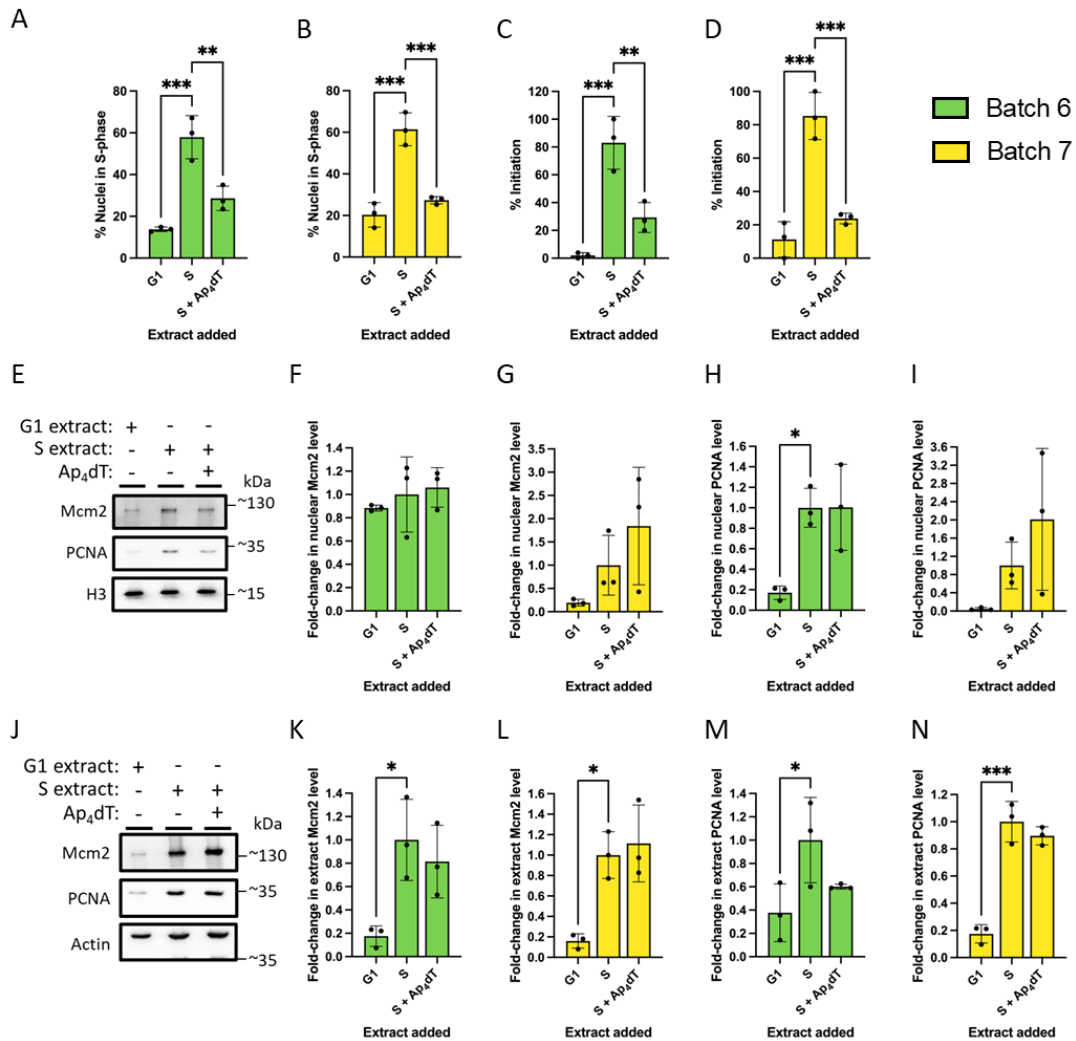


**Figure 6.13: The role of Ap<sub>4</sub>dT in inhibiting DNA replication.** Cell-free replication assays were performed in the presence of Ap<sub>4</sub>dT. **A, B:** The raw number of nuclei in S-phase using two batches of nuclei and **C, D:** the standardised values to give the percentage initiation in the presence and absence of Ap<sub>4</sub>dT. A one-way ANOVA was performed followed by Dunnett's multiple comparisons post-hoc test to compare G1 and S+Ap<sub>4</sub>dT with S-phase extract only within each batch of nuclei. \*\*P≤0.01; \*\*\*P≤0.001.

The percentage of nuclei in S-phase was significantly increased in both nuclei batches (adjusted P value = 0.0009 [batch 1] and 0.0031 [batch 3]) when added to S-phase extract compared with the G1-phase extract (Fig. 6.13 A and B). Furthermore, the percentage of nuclei in S-phase and the percentage initiation of DNA replication were

significantly inhibited in batch 1 and batch 3 nuclei (adjusted P value = 0.0059 [batch 1] and 0.0050 [batch 3]) (Fig. 6.13). This data demonstrates a consistently strong inhibitory effect of Ap<sub>4</sub>dT at a concentration of 100  $\mu$ M.

As Ap<sub>4</sub>dT consistently demonstrated a strong, statistically significant, inhibitory effect on DNA replication, we hypothesised that it may be functioning via a different mechanism to the other Ap<sub>4</sub>Ns. We therefore investigated whether Ap<sub>4</sub>dT retained its inhibitory effect even in nuclei that had previously been excluded due to the lack of effect induced by Ap<sub>4</sub>A (Fig. 6.14).



**Figure 6.14: The effect of Ap<sub>4</sub>dT on the inhibition of DNA replication initiation in cells that do not show a response to Ap<sub>4</sub>A.** **A, B:** The raw percentage of nuclei in S-phase and **C, D:** the standardised values to give percentage initiation in the presence and absence of Ap<sub>4</sub>dT. **E–N:** Representative Western blotting data for changes in Mcm2 and PCNA levels on chromatin (**E–I**) and soluble (**J–N**) fractions for two batches of nuclei. Statistical analysis was performed using a one-way ANOVA followed by Dunnett's multiple comparisons test to compare G1 and S+Ap<sub>4</sub>dT with S-phase extract only within each set of results. \*P≤0.05; \*\*P≤0.01; \*\*\*P≤0.001.

Even in the batches of nuclei where Ap<sub>4</sub>A did not inhibit the initiation of DNA replication, Ap<sub>4</sub>dT consistently induces a statistically significant inhibition of the initiation of DNA replication (adjusted p value=0.0036 and 0.0007 in batch 6 and batch 7, respectively). This statistically significant change in the percentage of nuclei in S-phase and the percentage initiation occurs across all nuclei batches (Fig. 6.13 and 6.14). However, unlike with the other Ap<sub>4</sub>Ns, this strong inhibitory effect is not

reflected by a change in the nuclear localisation of Mcm2 and PCNA (Fig. 6.14 E–I). In fact, the results from batch 7 suggest an increase in replication protein levels in the chromatin fraction in the presence of Ap<sub>4</sub>dT (Fig. 6.14 G and I). However, this result is likely due to the wide standard deviation of this data set. In the batch 6 nuclei, the available pool of Mcm2 and PCNA appeared to be reduced in the presence of Ap<sub>4</sub>dT compared with S-phase extract only (Fig. 6.14 K and M). However, this was not consistent with the other batch of nuclei which showed a similar level of Mcm2 and PCNA availability in the soluble fraction (Fig. 6.14 L and N).

### 6.6.6 Summary of the effects of Ap<sub>4</sub>Ns on DNA replication

After standardisation, it is possible to compare the inhibitory effect of the different Ap<sub>4</sub>Ns based on the change in percentage initiation of the Ap<sub>4</sub>N-treated nuclei in S-phase extract compared with S-phase extract only. A summary of the effect of each of the Ap<sub>4</sub>Ns on the inhibition of DNA replication initiation is shown in Table 6.1.

**Table 6.1: Summary data for the change in percentage initiation for Ap<sub>4</sub>A, Ap<sub>4</sub>C, Ap<sub>4</sub>U, Ap<sub>4</sub>G and Ap<sub>4</sub>dT.** The mean value for the percentage initiation in the presence and absence of the Ap<sub>4</sub>N was calculated for each experiment. The mean change in percentage initiation is represented in the table below where ‘=’ represents a change of 0–10%; ‘↓’ or ‘↑’ = 10–30%; ‘↓↓’ or ‘↑↑’ = 30–50% and ‘↓↓↓’ or ‘↑↑↑’ = >50%, where the direction of the arrow indicates whether percentage initiation decreases (↓) or increases (↑) when the Ap<sub>4</sub>N is included in the reaction mix. Where a ‘/’ separates two arrows, this indicates the results from two different sets of experiments with a particular Ap<sub>4</sub>N within a given nuclei batch. \*P≤0.05; \*\*P≤0.01.

		Nuclei batch		
		Batch 1	Batch 2	Batch 3
Change in % initiation	Ap <sub>4</sub> A	↓ (ns)	↓ ↓ (ns)	↓↓ (ns)
	Ap <sub>4</sub> C	= (ns)		↓↓ (*)
	Ap <sub>4</sub> G	↓ (ns)	↓↓ (**)	↓↓ (ns)
	Ap <sub>4</sub> U	↓↓↓ (**)	↓↓(*)/↓(ns)	
	Ap <sub>4</sub> dT	↓↓ (**)		↓↓↓ (**)

Overall, the inhibition of the initiation of DNA replication does not seem to be limited to just Ap<sub>4</sub>A. However, the extent to which each Ap<sub>4</sub>N inhibits the percentage initiation into S-phase varies with each Ap<sub>4</sub>N as well as between nuclei batches (Table 6.1). Although the results from batch 1 suggested that Ap<sub>4</sub>U was the most potent inhibitor, the other batches did not show as much of a reduction in percentage initiation after Ap<sub>4</sub>U treatment. On the other hand, Ap<sub>4</sub>dT showed a strong inhibition of the initiation of DNA replication in both batches of nuclei tested (Table 6.1).

Standardisation of the Western blotting data also enabled comparison between the effect of each of the Ap<sub>4</sub>Ns on the localisation of DNA replication proteins Mcm2 and PCNA to chromatin. In this process, the intensity of the Western blotting bands was standardised with respect to histone H3, then scaled so that the standardised band intensities for the S-phase extract only control was set to one. A summary of the effect of the different Ap<sub>4</sub>Ns on the two replication proteins is shown in Table 6.2.

**Table 6.2: Summary data for the effect of Ap<sub>4</sub>Ns on the Mcm2 and PCNA localisation to chromatin.** The mean change in Mcm2 and PCNA localisation to chromatin is represented in the table below where ‘=’ represents a change of 0–10%; ‘↓’ or ‘↑’ = 10–30%; ‘↓↓’ or ‘↑↑’ = 30–50% and ‘↓↓↓’ or ‘↑↑↑’ = >50%, where the direction of the arrow indicates whether percentage initiation decreases (↓) or increases (↑) when the Ap<sub>4</sub>N is included in the reaction mix. Where a ‘/’ separates two arrows, this indicates the results from two different sets of experiments with a particular Ap<sub>4</sub>N within a given nuclei batch.

Change in localisation to chromatin		
<b>Mcm2</b>	<b>Ap<sub>4</sub>A</b>	↓
	<b>Ap<sub>4</sub>G</b>	↓
	<b>Ap<sub>4</sub>U</b>	↑↑↑ / ↓↓
<b>PCNA</b>	<b>Ap<sub>4</sub>A</b>	↓↓
	<b>Ap<sub>4</sub>G</b>	↓↓
	<b>Ap<sub>4</sub>U</b>	↓↓↓/↓↓

Although the data were not statistically significant, there was a general trend suggesting that nuclear Mcm2 and PCNA levels were reduced by addition of Ap<sub>4</sub>N in

cell-free DNA replication reactions (Table 6.2). Additionally, when treated with a dinucleoside tetraphosphate there was generally a larger reduction in nuclear PCNA levels than nuclear Mcm2 levels within a nuclei batch (Table 6.2; Mcm2 vs. PCNA). The reduction in PCNA loading and a reduction in biotin-dUTP incorporation are consistent with Ap<sub>4</sub>Ns reducing recruitment of PCNA due to a reduction in active replication forks. This suggests that these molecules block replication complex assembly to reduce initiation of DNA replication. There is one exception to the general reduction in the localisation of nuclear proteins, where there is discrepancy in the effect of Ap<sub>4</sub>U on nuclear Mcm2 levels across the same nuclei batch (Table 6.2; Mcm2, Ap<sub>4</sub>U). One of these data sets shows an increase in the mean change in Mcm2 localisation. However, this increase is likely caused by an outlying result in one of the three replicates which greatly increases the average percentage change (Table 6.2; Fig. 6.12 H).

## **6.7 Chapter discussion**

### **6.7.1 Late G1 nuclei can be synchronised and initiated to replicate *in vitro***

This chapter has demonstrated that mouse 3T3 fibroblasts and HeLa cells can be synchronised at mid G1-phase, late G1-phase and S-phase (Fig. 6.2 and 6.3). The procedure of initiating DNA replication in G1 nuclei through the addition of S-extract is well-established, as it has been around since 1997 (Krude et al., 1997). There was some variation in the background level of nuclei already in S-phase and the extent to which different batches of nuclei could initiate replication when added to an S-phase extract. This is likely due to the specific timing of harvest of different nuclei batches after synchronisation, as the window of time for the harvesting of late G1 nuclei is relatively short. If the nuclei are harvested too early, they will be earlier in G1-phase and therefore it is possible that not all nuclei will reach S-phase within the 30-minute period during which they are incubated in S-phase extract. Likewise, if they are harvested too late, there will be a higher background level of nuclei already in S-phase prior to the incubation in S-phase extract. In addition to this, biological variation between cells that have been cultured at different times is likely to occur. It is for this reason that G1-phase and S-phase controls are used for each set of experiments.

### **6.7.2 Ap<sub>4</sub>A inhibits the initiation of DNA replication**

Although not statistically significant, there was a consistent reduction in initiation in several nuclei batches, suggesting that Ap<sub>4</sub>A inhibits the initiation of DNA replication. This trend is supported by a trend in the Western blotting data which suggested a reduction of nuclear Mcm2 and PCNA levels after incubation with Ap<sub>4</sub>A. It is also supported by the Marriot and colleagues 2015 paper, which demonstrated that Ap<sub>4</sub>A inhibited the initiation, but not elongation stage of DNA replication. Since we had demonstrated that Ap<sub>4</sub>A could inhibit DNA replication initiation in cell-free replication assays, this enabled us to investigate the Ap<sub>4</sub>A-binding properties, as the effect of Ap<sub>4</sub>A on DNA replication could be compared with other molecules with similar structures, including modifications to the phosphate chain and nucleotide sections of the molecule.

### **6.7.3 Modification of the phosphate chain does not affect Ap<sub>4</sub>N activity, but modification of the nucleoside components does**

Comparison between Ap<sub>4</sub>A and AppNppA demonstrated that AppNppA inhibits the initiation of DNA replication to at least the same extent as Ap<sub>4</sub>A, if not more so (Fig. 6.6). AppNppA has previously been used in studies as a more stable analogue of Ap<sub>4</sub>A (Viatchenko-Karpinski et al., 2016). Importantly, substitution with nitrogen in the phosphate chain only involves the replacement of one oxygen atom with a nitrogen atom, meaning that chain length is not substantially affected. By comparing the effect of Ap<sub>3</sub>A, Ap<sub>4</sub>A and Ap<sub>5</sub>A on the inhibition of DNA replication, a previous study has shown that a specific phosphate chain length is essential for inhibition of DNA replication to occur (Marriott et al., 2015). This is consistent with studies showing the effect of Ap<sub>4</sub>A in other pathways, such as HINT1 polymerisation which is induced by Ap<sub>4</sub>A but not Ap<sub>3</sub>A or Ap<sub>5</sub>A (Yu et al., 2019). This substitution therefore demonstrates that a change to the phosphate chain does not affect activity if the chain length is not affected, suggesting that the phosphate chain is less likely to be involved in Ap<sub>4</sub>A-substrate binding.

To confirm whether AppNppA has a consistently stronger inhibitory effect on the initiation of DNA replication, further repeats would need to be performed.



Additionally, experiments could be performed with alternative Ap<sub>4</sub>A analogues to determine whether these also retained the inhibitory activity of Ap<sub>4</sub>A. Examples of other stable Ap<sub>4</sub>A analogues that have been used in experiments previously include AppCH<sub>2</sub>ppA (Viatchenko-Karpinski et al., 2016), and the non-hydrolysable JB419 analogue used to investigate the interaction between Ap<sub>4</sub>A and HINT1 (Dolot et al., 2016; Dolot et al., 2021).

In contrast, the Ap<sub>4</sub>A-FRET molecule did not inhibit initiation of DNA replication (Fig. 6.7). As the Ap<sub>4</sub>A-FRET molecule has a bulky group joined to each of the nucleoside moieties (Götz et al., 2017), the lack of activity is likely due to these groups preventing Ap<sub>4</sub>A-substrate interactions from occurring. Previously, the Ap<sub>4</sub>A-FRET molecule has been used to investigate inhibitors of the Ap<sub>4</sub>A phosphorylase Rv2613c from *M. tuberculosis* (Götz et al., 2017). When the two fluorophores on the Ap<sub>4</sub>A-FRET molecule are in close proximity, in other words when they are present on the same molecule, they undergo FRET, which changes their fluorescence properties compared with when the Ap<sub>4</sub>A molecules has been hydrolysed. Initially, the theory behind testing the Ap<sub>4</sub>A FRET in cell-free replication assays was that fluorescence imaging may enable us to identify whether the Ap<sub>4</sub>A-induced inhibition of DNA replication initiation required Ap<sub>4</sub>A hydrolysis. However, unfortunately this could not be investigated because this molecule did not inhibit DNA replication initiation in the cell-free replication assays. Despite this, the initial investigation provided key data supporting the hypothesis that the nucleoside moieties are important for Ap<sub>4</sub>A binding and function, as the bulky groups in the FRET molecule prevented the Ap<sub>4</sub>A from functioning normally. It is possible that these side chains may have prevented the usual binding of Ap<sub>4</sub>A to its substrate, which would then prevent normal Ap<sub>4</sub>A function. Future studies aiming to investigate the binding mechanism and function of Ap<sub>4</sub>A should therefore carefully consider the choice to use any linkers or side chains on these moieties.

Together these data suggest that the nucleotide moieties are important for binding, whereas the structure of the phosphate chain is less important, providing chain length is not altered. This could provide a good foundation for the improvement of technical

approaches to the study of Ap<sub>4</sub>A in DNA replication, as bulkier groups could be added to the central nitrogen atom to determine whether the molecule retains its activity. If this activity is retained, this structure could be further developed to include the addition of an alkyne group to the central nitrogen molecule. Such an alkyne group could then be utilised for the addition of other groups, such as biotin or fluorophores, onto the Ap<sub>4</sub>A molecule. If these retained the inhibitory function, they could be used for further investigation into Ap<sub>4</sub>A binding partners and localisation.

#### **6.7.4 Ap<sub>4</sub>Ns inhibit the initiation phase of DNA replication, although the potency varies in different batches of nuclei**

We have demonstrated that at least one batch of all the Ap<sub>4</sub>Ns tested showed a reduction in the percentage of nuclei initiating into S-phase compared with the percentage initiation in nuclei added to S-phase extract alone (Table 6.1). Although not all batches of nuclei gave statistically significant reductions in percentage initiation, they generally followed a trend of reducing percentage initiation compared with nuclei added to S-phase extract only. This suggests that the specific Ap<sub>4</sub>A structure is not essential for its inhibitory activity.

Ap<sub>4</sub>G has a structure that is very similar to Ap<sub>4</sub>A; however, there are subtle changes in structure. These changes are limited to the repositioning of the amine group from C6 to C2, and the addition of an oxygen double bonded to C6. Our results demonstrate that these changes did not impact the ability of Ap<sub>4</sub>G to inhibit the initiation of DNA replication. Despite its similarity in structure, previous studies have generally focussed on the role of Ap<sub>4</sub>A in DNA replication. However, in the recent study investigating the role of Ap<sub>4</sub>A in DNA replication, Gp<sub>4</sub>G was also compared (Marriott et al., 2015). In this study, Gp<sub>4</sub>G did not show a significant effect on the inhibition of DNA replication in a cell-free system (Marriott et al., 2015). This suggests that at least one adenosine moiety may be essential for function.

Ap<sub>4</sub>C and Ap<sub>4</sub>U are pyrimidines and are therefore more structurally distinct from Ap<sub>4</sub>A. As these molecules were also both able to inhibit the initiation of DNA replication (Fig. 6.10 and 6.12), it further suggests that the Ap<sub>4</sub>A binding site which contributes to

the inhibition of DNA replication does not require a homobifunctional structure. As all the nucleotide tetraphosphates were able to inhibit the initiation of DNA replication, it provides evidence that these molecules may have a functional role in the DNA replication system. This hypothesis is supported by our results in Chapter 4, where we demonstrated that all the Ap<sub>4</sub>Ns exist in mammalian cells, opening the possibility that they could have an intracellular signalling role. Furthermore, Ap<sub>4</sub>U is already implicated in other signalling mechanisms such as extracellular signalling in the cardiovascular system (Jankowski et al., 2013a; Zhou et al., 2019b). This demonstrates that, at least extracellularly, it is present at concentrations high enough to induce signalling.

Interestingly, Ap<sub>4</sub>dT appeared to consistently have a significant effect on inhibition of DNA replication initiation in the cell-free replication assays (Fig. 6.13 and 6.14). Thymidine has a very similar structure to uridine, with the only changes being the addition of a methyl group to C5 of the pyrimidine ring, and loss of a hydroxyl group from the ribose. It was therefore surprising that this small, non-polar methyl group made such a large difference in the inhibition of DNA replication. A possible explanation for the consistently strong inhibition of DNA replication by Ap<sub>4</sub>dT is that it may function through some other mechanism. This suggestion is supported by the lack of effect of Ap<sub>4</sub>dT on localisation of the replication proteins to chromatin compared with the reduced localisation of these proteins after incubation with the other Ap<sub>4</sub>Ns. However, the intracellular concentration of Ap<sub>4</sub>dT is likely very low, as it was only just detectable in NuKO cells (Fig. 4.14), which show a 175-fold increase in total Ap<sub>4</sub>N concentration compared with the parental KBM7 cell line under non-stressed conditions. Therefore, as a next step it would be useful to investigate the half-maximal inhibitory concentration (IC<sub>50</sub>) value for Ap<sub>4</sub>dT, which would help with determining whether the inhibitory effect of Ap<sub>4</sub>dT could be biologically relevant.

#### **6.7.5 Inhibition of DNA replication initiation by Ap<sub>4</sub>N is associated with a trend suggesting reduced localisation of Mcm2 and PCNA to chromatin**

Addition of all Ap<sub>4</sub>Ns tested, except Ap<sub>4</sub>dT, showed a trend in reducing in the mean level of Mcm2 and PCNA on chromatin compared with the addition of nuclei to

S-phase extract only (Table 6.2). Additionally, all Ap<sub>4</sub>Ns induced a greater reduction in mean nuclear PCNA level than in mean Mcm2 level (Table 6.2). However, these data were not significant, so more repeats would be needed to reduce the likelihood that this is due to chance. Alternatively, modification of the Western blotting technique to improve sensitivity could be considered for the measurement of Mcm2 and PCNA levels. Generally, Western blotting is a well-established method, but is considered semi-quantitative because of the variability in measurements due to protein transfer, immunostaining and ECL detection (Zellner et al., 2008). Quantitative fluorescent Western blotting (QFWB) is a modified Western blotting technique that can be considered quantitative (Eaton et al., 2014). This is because it enables the protein of interest and total protein expression profile to be imaged simultaneously, and there is a strong correlation between quantitation and protein amount across an extended dynamic range (Zellner et al., 2008). Therefore, modifying the Western blotting technique to use fluorescent antibodies could improve the accuracy of quantitation, and reduce the variability between repeats.

The precise time at which Ap<sub>4</sub>A (and potentially Ap<sub>4</sub>Ns) inhibit the DNA replication process is unknown. However, it has previously been shown that Ap<sub>4</sub>A inhibits the initiation but not elongation phase of DNA replication (Marriott et al., 2015). As Mcm2 and PCNA are recruited at different stages of the DNA replication process, these data were able to help in the elucidation of the point at which Ap<sub>4</sub>N has an effect. Mcm2 is one of the first components to localise to DNA, and is loaded onto DNA during the formation of the pre-RC (Yuan et al., 2020). Conversely, PCNA isn't recruited until after the replisome begins to form, and has a role in stimulating the DNA polymerases (Chilkova et al., 2007). Therefore, these suggest that Ap<sub>4</sub>N has a small effect in the lead up to Mcm2 localisation, but exerts its primary effect after this, between Mcm2 and PCNA localisation to chromatin. This is supported by the data from Marriott and colleagues in 2015 demonstrating Ap<sub>4</sub>A's involvement in DNA replication initiation, and it further narrows down the window in which Ap<sub>4</sub>N is having an effect. Furthermore, we have demonstrated that it is specifically the localisation of these proteins to chromatin that is affected, rather than their availability, as Mcm2 and PCNA levels are comparable in the soluble fraction samples regardless of Ap<sub>4</sub>N presence or absence. To

further elucidate this process, quantification of other proteins involved in Pre-IC and early replisome formation could be performed. This could include measurement of the nuclear localisation of a pre-IC component, such as GINS, and nuclear localisation of an earlier initiator of replisome formation, such as Mcm10 or DNA pol  $\alpha$ .

#### **6.7.6 Ap<sub>4</sub>dT strongly inhibits the initiation of DNA replication, but this is not reflected in its effect on replication protein localisation to chromatin**

Ap<sub>4</sub>dT induced a particularly strong effect on the inhibition of DNA replication initiation, and this occurred in all nuclei batches tested regardless of whether Ap<sub>4</sub>A had an effect (Fig. 6.13 and 6.14). Ap<sub>4</sub>dT is the only Ap<sub>4</sub>N tested for which the reduction in percentage initiation was statistically significant for all nuclei batches tested. Interestingly, unlike the other Ap<sub>4</sub>Ns, Ap<sub>4</sub>dT did not appear to inhibit DNA replication through reducing the nuclear localisation of Mcm2 and PCNA (Fig. 6.14). This suggests that Ap<sub>4</sub>dT may be inhibiting the initiation of DNA replication through some other mechanism which is distinct from that of Ap<sub>4</sub>A. As Ap<sub>4</sub>dT is present at very low intracellular concentrations, it remains to be determined whether it could have an effect *in vivo*. However, it highlights an alternative mechanism that could be studied further and potentially be utilised for the inhibition of DNA replication.

### **6.8 Conclusions**

In summary, mouse 3T3 fibroblasts can be synchronised at early and late G1 phase using a contact inhibition and serum starvation approach. As HeLa cells are not contact inhibited, synchronisation of these cells into S-phase can be achieved through incubation with thymidine. These synchronised nuclei and extracts can then be used to initiate replication in late G1 nuclei. Addition of Ap<sub>4</sub>A to late G1 nuclei in S-phase extract resulted in a trend showing a reduction in percentage initiation compared with S-phase extract only. This trend was supported by results from a previous study (Marriott et al., 2015). Modification of the phosphate chain did not prevent the inhibitory activity of AppNppA, but additional groups on the adenosine moieties prevented Ap<sub>4</sub>A-FRET from inhibiting DNA replication initiation, suggesting that these moieties are important for Ap<sub>4</sub>A binding. However, most Ap<sub>4</sub>Ns showed a significant reduction in percentage initiation, suggesting that minor modifications to the

nucleoside structure did not prevent inhibition from occurring. Ap<sub>4</sub>dT had a particularly strong effect on the inhibition of DNA replication. However, this did not correspond to a reduction in Mcm2 or PCNA nuclear localisation, suggesting that perhaps Ap<sub>4</sub>dT can inhibit the initiation of DNA replication via a mechanism that is distinct from the Ap<sub>4</sub>A-induced inhibition. Ap<sub>4</sub>A, Ap<sub>4</sub>G and Ap<sub>4</sub>U all showed a trend of reducing the localisation of Mcm2 and PCNA to chromatin, with an effect that was generally stronger for nuclear localisation of PCNA compared with Mcm2. This suggests that these molecules may function via the same mechanism as Ap<sub>4</sub>A, and evidence from Chapter 4 demonstrating that these Ap<sub>4</sub>Ns exist in mammalian cells suggests the potential for these Ap<sub>4</sub>Ns to also have a role in the inhibition of DNA replication initiation. Furthermore, it narrows the window in which Ap<sub>4</sub>A has its primary effect on DNA replication initiation, to somewhere after Mcm2 localisation but before PCNA localisation to chromatin.

## **Chapter 7: General Discussion**

## **7.1 Ap<sub>4</sub>Ns can be synthesised via a biologically relevant mechanism**

### **7.1.1 A mechanism linking Ap<sub>4</sub>A synthesis to the stress response has previously been described**

Ap<sub>4</sub>A was first identified in 1966, and since then it has been studied extensively across a variety of signalling mechanisms (Zamecnik et al., 1966). Ubiquitous across multiple domains of life, Ap<sub>4</sub>A is synthesised in response to stress, resulting in increased intracellular Ap<sub>4</sub>A levels (Brevet et al., 1989b; Coste et al., 1987; Lee et al., 1983; Marriott et al., 2015). In *S. cerevisiae* and several bacterial strains, other Ap<sub>4</sub>Ns have been studied alongside Ap<sub>4</sub>A, and their intracellular levels also increase in response to several types of cell stress (Coste et al., 1987; Garrison et al., 1986; Lee et al., 1983). For many years we have known about a number of Ap<sub>4</sub>A- and Ap<sub>4</sub>N-producing synthesis mechanisms, including those that require various synthetases, luciferase or ligases (Fontes et al., 1998; Goerlich et al., 1982; Guranowski et al., 1990; Plateau et al., 1981; Sillero and Sillero, 2000; Zamecnik et al., 1966). However, in recent years an additional mechanism involving ubiquitin and the ubiquitin-like activating enzymes has been identified (Götz et al., 2019). As ubiquitin activating enzyme activity is also stimulated in response to cell stress (Shang et al., 1997), this mechanism for Ap<sub>4</sub>A synthesis links Ap<sub>4</sub>A production with the stress response. In the context of the stress response during DNA replication, increased activity of ubiquitin activating enzymes may contribute to the observed increase in the level of intracellular Ap<sub>4</sub>A, a molecule which is known to inhibit DNA replication (Marriott et al., 2015). However, whether the ubiquitin activating enzyme can stimulate the production of other Ap<sub>4</sub>Ns, and whether these other Ap<sub>4</sub>Ns contribute to DNA replication, has not previously been investigated.

### **7.1.2. UBE1/Ub can synthesise Ap<sub>4</sub>A as well as other Ap<sub>4</sub>Ns**

This work has demonstrated that *in vitro* reactions, where ATP and NTP are combined in the presence of ubiquitin and ubiquitin activating enzyme, result in small HPLC peaks which correspond to the expected retention time for the relevant Ap<sub>4</sub>Ns. Although this is highly indicative of ubiquitin and ubiquitin-activating enzyme-dependent Ap<sub>4</sub>N synthesis, HRMS will need to be performed for confirmation. Our



initial attempt at verifying peak identity was unsuccessful, so further development of this method would be necessary to verify peak identity by HRMS. The lack of success with the current method is likely due to sample contamination with PEGs (despite the use of DNA LoBind tubes to reduce plastic exposure), combined with the low concentration of sample retrieved from the HPLC. PEG contamination may have masked the relevant Ap<sub>4</sub>N peaks during HRMS analysis. Therefore, further development of this method could aim to increase product concentration or reduce potential contamination prior to HRMS.

### **7.1.3 Further development of the method is required to enable verification of UBE1/Ub-synthesised Ap<sub>4</sub>Ns**

One potential avenue for method development would be to increase the amount of Ap<sub>4</sub>N product formed during the reaction. Previously it has been shown that there is a positive correlation between increased UBA1 concentration and increased Ap<sub>4</sub>A synthesis (Götz et al., 2019). Therefore, increasing the concentration of the ubiquitin activating enzyme may result in increased Ap<sub>4</sub>N formation. Sequential addition of ubiquitin activating enzyme to the reaction mixture could be another potential way to increase product formation. Time-course experiments in Chapter 3 demonstrated that an increased reaction time did not have much of an effect on reaction product abundance (Fig. 3.14). It is therefore possible that there is a limited use of the ubiquitin activating enzyme in the *in vitro* system before it needs to be reset to become active again, and some other reaction component not already included in the mixture may be necessary for this to occur. Further studies could identify whether this is the case by adding a second amount of ubiquitin activating enzyme to the reaction mixture to determine whether Ap<sub>4</sub>N production increases.

Alternatively, a method for the clean-up of the HPLC extract could be determined to reduce sample contamination. PEG contamination is a known challenge in mass spectrometry, and titanium dioxide microcolumns developed for the enrichment of phosphorylated peptides from peptide mixtures have been applied for the removal of PEGs from phosphorylated peptide samples (Larsen et al., 2005; Zhao and O'Connor, 2007). Further studies could identify whether similar techniques could be applied for

the removal of PEGs from Ap<sub>4</sub>N-containing mixtures. These future developments centre around maximising product formation and limiting PEG contamination. If developed, they could be used in tandem to limit the possibility of Ap<sub>4</sub>N products being masked by contaminants.

## **7.2 Ap<sub>4</sub>Ns can be detected in the NuKO mammalian cell line**

### **7.2.1 Ap<sub>4</sub>A, Ap<sub>4</sub>C Ap<sub>4</sub>G and Ap<sub>4</sub>U have previously been identified in bacteria and lower eukaryotes, but not in mammals**

From as early as the 1960s, *E. coli* lysyl sRNA synthetase was found to synthesise Ap<sub>4</sub>Ns (Randerath et al., 1966). Since then, Ap<sub>4</sub>Ns have been identified intracellularly in the *E. coli* strain JM101Tr at a total intracellular concentration of 3.4 µM (Brevet et al., 1989b), and in the *E. coli* K-12 strain AB1157 at concentrations of 2.42 µM (Ap<sub>4</sub>A), 0.61 µM (Ap<sub>4</sub>C), 0.95 µM (Ap<sub>4</sub>G), and 1.17 µM (Ap<sub>4</sub>U) (Plateau et al., 1987b). Similarly, Ap<sub>4</sub>A, Ap<sub>4</sub>C, Ap<sub>4</sub>G and Ap<sub>4</sub>U have been identified in exponentially growing *S. cerevisiae* cells, where they are present at concentrations of 0.55 µM, 0.20 µM, 0.28 µM and 0.46 µM, respectively (Coste et al., 1987). Ap<sub>4</sub>A and Ap<sub>4</sub>G have also been found in *Drosophila* cells at concentrations of 0.25 and 0.31 µM, respectively (Brevet et al., 1985b).

Ap<sub>4</sub>A has also been documented in several mammalian cell lines, including mouse embryo fibroblasts (MEFs), HeLa cells and Chinese hamster AA8 cells (Marriott et al., 2015). However, although Ap<sub>4</sub>U has a documented extracellular role in the cardiovascular system (Jankowski et al., 2013a; Matsumoto et al., 2011), there were previously no reports identifying Ap<sub>4</sub>U or the other Ap<sub>4</sub>Ns within mammalian cells.

### **7.2.2 Ap<sub>4</sub>Ns can be extracted from mammalian cells and detected by triple quadrupole mass spectrometry**

In this thesis a method has been adapted from one used for the extraction of metabolites from bacterial cells (Fung et al., 2020) for the extraction of Ap<sub>4</sub>Ns from a mammalian cell line. Further to this, we have developed an LC-MS based technique for the specific identification of each of the Ap<sub>4</sub>Ns in extract harvested from mammalian

cells with a disrupted *Nudt2* gene (Fig. 4.11). This method is capable of distinguishing between different Ap<sub>4</sub>Ns, and was used to identify that Ap<sub>4</sub>A, Ap<sub>4</sub>C, Ap<sub>4</sub>G and Ap<sub>4</sub>U are present in NuKO cell extracts. The data resulting from this analysis suggested that each of the Ap<sub>4</sub>Ns are present at different concentrations in cells, with Ap<sub>4</sub>A and Ap<sub>4</sub>U generating the largest fragment ion counts. However, as this difference could be brought about through the difference in the fragmentation efficiency of the different Ap<sub>4</sub>N molecules, further development of the technique would be necessary to confirm relative Ap<sub>4</sub>N abundance. Method development could therefore include the production of a calibration curve using increasing concentrations of the Ap<sub>4</sub>N standard for each molecule. This could then be applied to determine the approximate concentration of each Ap<sub>4</sub>N in the NuKO cell line.

### **7.2.3 Further development of the method for identifying Ap<sub>4</sub>Ns in mammalian cells**

In the two repeats initially identifying Ap<sub>4</sub>Ns in NuKO cells, the ion count for each of the Ap<sub>4</sub>Ns varied between repeats (Fig. 4.11 C vs D). However, the extract with the higher ion count was consistently higher for all Ap<sub>4</sub>Ns, suggesting that this difference may have been due to the efficiency of the extraction process. The method could therefore be developed further by introducing a system for the measurement of the efficiency of the extraction process. Method design to account for this could include the introduction of a non-natural, unique Np<sub>4</sub>N analogue of known concentration into the medium prior to extraction. The LC-MS ion-count profile for the spiked analogue could then be compared with that expected from the concentration injected, which could be used to determine how much of the molecule was lost during the extraction. For this to work, the spiked novel analogue would need to have similar properties to the Ap<sub>4</sub>Ns, as the process would rely on the assumption that all molecules would be extracted with the same efficiency.

Alternatively, isotopic labelling of the Ap<sub>4</sub>N molecule could be performed to aid quantitative mass spectrometry analysis (Morano et al., 2008). An Ap<sub>4</sub>N molecule that was composed of different isotopes could be used in the same way as the spiked novel analogue, as the LC-MS would be able to differentiate the different mass of this molecule from the naturally occurring one. This would have the advantage of having

the same structure as the Ap<sub>4</sub>N molecule under investigation, while having the ability to distinguish between the two molecules through their mass.

### **7.3 DNA damage may increase intracellular Ap<sub>4</sub>N concentration in NuKO cells, but improved controls are required for quantitation**

#### **7.3.1 The effect of stress on Ap<sub>4</sub>N concentration**

A number of DNA damaging agents and stress conditions have been linked to increased intracellular concentrations of Ap<sub>4</sub>N, including heat shock (Brevet et al., 1989b; Coste et al., 1987; Lee et al., 1983) and oxidative stress (Coste et al., 1987; Garrison et al., 1986; Ji et al., 2019; Lee et al., 1983; Marriott et al., 2015). While most of these studies are in bacteria and lower eukaryotes, the study by Marriot and colleagues in 2015 demonstrated an increase in the intracellular Ap<sub>4</sub>A level in several mammalian cell lines in response to oxidative stress induced by MMC. This study calculated Ap<sub>4</sub>A concentration using a luminometric assay to determine ATP production (Marriott et al., 2015). However, since we have demonstrated that other Ap<sub>4</sub>Ns are present in mammalian cells (Fig. 4.11), and that Ap<sub>4</sub>Ns have a hydrolysis product that co-elutes with ATP (Fig. 3.9), this demonstrates that other Ap<sub>4</sub>Ns may have contributed to the concentration previously reported as Ap<sub>4</sub>A. This opens the possibility that there could be a function for the other Ap<sub>4</sub>Ns that may have been overlooked previously. We therefore aimed to develop a method to determine whether MMC also results in an increase in the intracellular concentration of the other Ap<sub>4</sub>Ns, as this may indicate whether they are likely to have a functional role *in vivo*.

Further to this, we aimed to investigate the Ap<sub>4</sub>A mechanism of action by expanding the types of damage that the cells were exposed to. Agents previously used to induce oxidative stress include ethanol (Lee et al., 1983), cadmium (Coste et al., 1987), aminoglycoside antibiotic (Ji et al., 2019), dinitrophenol (Garrison et al., 1986) and MMC (Marriott et al., 2015). However, there are several oxidative stresses with different mechanisms of action that have not yet been shown to increase intracellular Ap<sub>4</sub>A level. To further investigate the mechanisms through which Ap<sub>4</sub>A (and potentially Ap<sub>4</sub>N) may be involved in in response to DNA stress, we therefore also decided to

investigate several other damaging agents that are known to induce DNA replication stress via different mechanisms.

### **7.3.2 Unstandardised data suggests that intracellular Ap<sub>4</sub>N levels increase in response to a variety of damaging agents**

The unstandardised data from Chapter 5 suggests that there is an increase in each of the Ap<sub>4</sub>Ns after the induction of stress, rather than purely an increase in Ap<sub>4</sub>A concentration (Fig. 5.9). However, as this measurement was not standardised, other factors could potentially have affected the result. As cells were counted prior to harvesting extract, approximately the same number of live cells should have been present in the treated and untreated samples. However, as the NuKO cells are suspension cells, any dead cells could not be removed, and so remained present in the pellet. Any residual Ap<sub>4</sub>N abundance in the dead or dying cells could therefore have contributed to the concentration of Ap<sub>4</sub>N measured. As more cells are likely to be dead in the samples treated with the damaging agent, this has the potential to disproportionately affect the results of the damaging agent-treated samples.

To account for this, ADP and ATP were tested as potential internal standards, as they have a high intracellular abundance and have similar structures to Ap<sub>4</sub>A. However, our data demonstrated that these molecules were not ideal standards as their concentrations were also affected by the stress-inducing agents (Fig. 5.11). This effect was most prominent for the HU and GEM samples, where the ADP:ATP ratio was particularly high. This result can be explained, as both of these drugs are known to inhibit ribonucleotide reductase, which converts NTPs to dNTPs and is therefore a factor that could affect ATP concentration (Hofer et al., 2012; Singh and Xu, 2016). Furthermore, Ap<sub>4</sub>A itself binds NrdR, a modulator of ribonucleotide reductase, and prevents it from binding to the *NrdA* promoter region (Despotović et al., 2017). This presents a further mechanism through which disrupted Ap<sub>4</sub>A levels may also affect ATP concentration and is another reason why ATP may not be a suitable internal standard in this kind of assay. Furthermore, ATP is a known precursor in the synthesis of Ap<sub>4</sub>Ns, meaning its concentration may also be affected through the synthesis of the Ap<sub>4</sub>Ns themselves. Additionally, the ADP:ATP ratio has previously been used as a

marker of cell viability and apoptosis, suggesting that it is unlikely to be an ideal standard for the comparison of stressed and unstressed cells (Bradbury et al., 2000).

### **7.3.3 Further method development to investigate the effect of DNA damaging agents on intracellular Ap<sub>4</sub>N abundance**

To truly determine the effect of these stress-inducing agents on mammalian cells, it would be necessary to develop an appropriate standardisation method. Ideally, to maintain consistency between repeats and to determine whether the unstandardised data is accurate, an internal standard would be identified which is not affected by the damaging agents. However, determining what this standard could be remains a challenge, as proteins are denatured during the chloroform:water step of the extraction process. Therefore, as the standard would need to be retained during this process, it cannot be any kind of protein. Although the dNTP and NTP molecules have a similar structure to Ap<sub>4</sub>A, these molecules are also involved in the DNA replication process and are therefore likely to be affected by agents that induce DNA replication stress, meaning that they are unlikely to make good internal standards.

Additionally, it is also possible that the true changes to the intracellular Ap<sub>4</sub>N levels are being obscured by the extremely high level of Ap<sub>4</sub>Ns in the NuKO cell line. To prevent this from happening, the method for investigating Ap<sub>4</sub>N levels in the KBM7 cell line could be further developed. Using the methods described here, it was not possible to identify Ap<sub>4</sub>Ns in the KBM7 parental cell line, either because their concentration was too low or because the ion suppression effect was too great. Re-evaluation of these results using a more sensitive mass spectrometer may resolve this. As a result of this lack of sensitivity, any differences in Ap<sub>4</sub>N level in the parental cell line were not detected. Unfortunately, as the baseline was not determined, any increase in Ap<sub>4</sub>N resulting from DNA damaging agents could not be identified. Therefore, future studies could aim to develop this method further to enable detection of the Ap<sub>4</sub>Ns in KBM7 cells. Previous studies have already demonstrated that Ap<sub>4</sub>A levels are increased in KBM7 cells in response to MMC treatment (Marriott et al., 2015). Therefore, as we know MMC affects intracellular Ap<sub>4</sub>A concentration, developing an analysis procedure to evaluate Ap<sub>4</sub>N concentration in response to MMC may be a good place to start.

Avenues that could be followed include maximising the sample concentration or introducing clean-up steps to reduce the ion suppression effect. As part of this process, we have demonstrated that a graphite column was able to lower the detectable concentration of Ap<sub>4</sub>A. However, use of this column may have resulted in some loss of Ap<sub>4</sub>A from the sample and was insufficient for the detection of Ap<sub>4</sub>A at concentrations in KBM7 cells. Further optimisation of the wash steps used in combination with the graphite column could increase retention of Ap<sub>4</sub>A and may enable it to reach a detectable concentration in KBM7 cell extract analysis.

Alternatively, method optimisation could involve the use of a more sensitive mass spectrometer for analysis of Ap<sub>4</sub>N concentration. A triple quadrupole mass spectrometer was chosen for this study as it allows the identification of specific masses from a complex mixture of molecules. This was an advantage in this case, as it enabled the detection of the individual Ap<sub>4</sub>Ns from the mixture of metabolites in the cell extract harvested from treated cells. Although high resolution mass spectrometers have a much higher sensitivity, they do not allow for the identification of molecules from a complex mixture, and so would not be appropriate for this type of study unless the Ap<sub>4</sub>Ns were to be fully purified from the cell extracts first. However, other triple quadrupole mass spectrometers, including the LCMS-8050 and LCMS-8060, are available which have a higher sensitivity and faster scan speeds compared with the LCMS-8040 used in this study (SHIMADZU, 2022). The limitation here is the high expense associated with these machines.

In summary, although the presence of the Ap<sub>4</sub>Ns was not detectable in the KBM7 cells, their presence in the NuKO cell extract provides strong evidence for their existence. Although the reduction in ion suppression brought about by the clean-up methods discussed in Chapter 4 were insufficient, increased injection concentration and further development of the protocol using the graphite column may provide potential areas for improvement going forwards. Alternatively, it may be necessary for this technique to be applied to a more sensitive LC-MS system for them to be detected. Identifying which types of genotoxic stress influence Ap<sub>4</sub>A concentration could provide key evidence as to how Ap<sub>4</sub>A exerts its effect in stress responses.

## **7.4 Ap<sub>4</sub>Ns have differential inhibitory activities during the initiation of DNA replication**

### **7.4.1 The role of Ap<sub>4</sub>A and Ap<sub>4</sub>Ns in DNA replication**

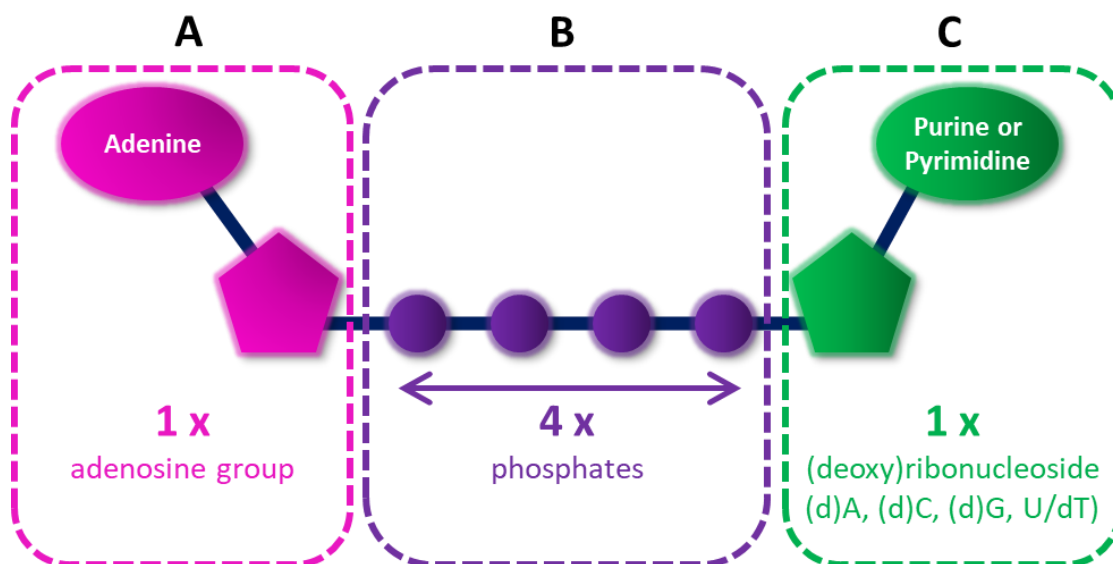
Ap<sub>4</sub>A has been implicated in the DNA replication process for many years. Initially thought to stimulate DNA replication (Grummt, 1978a), it has only recently been demonstrated to instead inhibit the initiation phase of DNA replication (Marriott et al., 2015). Nevertheless, how Ap<sub>4</sub>A inhibits the initiation of DNA replication is not fully understood. Ap<sub>4</sub>A is now known to be involved in several different signalling mechanisms, and functions differently in different contexts. In the polymerisation of HINT1, the homobifunctional structure of Ap<sub>4</sub>A is necessary as both ends of the molecule bind to adenine binding pockets on HINT1, linking two separate HINT dimers (Yu et al., 2019). Contrary to this, in the mRNA capping mechanism in *E. coli*, several Ap<sub>4</sub>Ns function as precursors to protective caps and different Ap<sub>4</sub>Ns have differences in capping efficiency depending on the promoter sequence (Luciano and Belasco, 2020; Luciano et al., 2019). In the DNA replication process, it is not known whether a homobifunctional or heterobifunctional structure is required for its activity.

In 2015, Marriott and colleagues used a cell-free replication system to demonstrate that Ap<sub>4</sub>A inhibited the initiation of DNA replication, and that this inhibition was strictly dependent on phosphate chain length, with neither diadenosine triphosphate nor diadenosine pentaphosphate able to inhibit the initiation phase of DNA replication. (Marriott et al., 2015). This study also investigated if substitution of adenosine for guanosine and using diguanosine tetraphosphate could similarly inhibit DNA replication. Gp<sub>4</sub>G was not effective at reducing initiation of DNA replication (Marriott et al., 2015). Whether or not other Ap<sub>4</sub>N molecules are capable of inhibiting DNA replication could shed some light on the mechanism by which DNA replication is inhibited.



#### 7.4.2 Ap<sub>4</sub>Ns inhibit the initiation of DNA replication to different extents

In Chapter 6, the cell-free replication assays used by Marriott and colleagues in 2015 were employed to demonstrate that Ap<sub>4</sub>C, Ap<sub>4</sub>G, Ap<sub>4</sub>U, Ap<sub>4</sub>dA, Ap<sub>4</sub>dC, Ap<sub>4</sub>dG, and Ap<sub>4</sub>dT are also capable of inhibiting DNA replication, and that they do this to different extents (Fig. 6.8, 6.10). This provides evidence for a heterobifunctional role of the Ap<sub>4</sub>Ns in the inhibition of DNA replication and is a step towards elucidation of the mechanism by which Ap<sub>4</sub>A modulates this process. Furthermore, alongside data from previous papers this helps to define the essential structural components that the Ap<sub>4</sub>N molecule must have to retain its function (Fig. 7.1)



**Figure 7.1: Essential components of an Np<sub>n</sub>N molecule capable of inhibiting the initiation of DNA replication.** To inhibit the initiation of DNA replication, three specific components of the Np<sub>n</sub>N molecule are essential. **A:** At least one adenosine group is required for functionality. **B:** A specific chain length of four phosphates is essential; however, small changes to phosphate chain structure (eg. in AppNppA) are tolerated. **C:** A second nucleoside group is required, in which both the ribose and deoxyribose form of the sugar are tolerated.

Data from Marriott and colleagues in 2015 demonstrated that Ap<sub>4</sub>A, but not Gp<sub>4</sub>G, resulted in the inhibition of DNA replication initiation. In addition to this, we have demonstrated that Ap<sub>4</sub>G is also capable of inhibiting DNA replication initiation (Fig. 6.11). Together, this suggests that at least one adenosine moiety is essential for the inhibition of DNA replication initiation (Fig. 7.1 A). Additionally, analysis with

Ap<sub>4</sub>A-FRET demonstrated that large modifications to the adenosine moieties are not tolerated for Ap<sub>4</sub>A function (Fig. 6.7). Furthermore, although a specific phosphate chain length of four phosphates is essential for Ap<sub>4</sub>A's inhibitory activity (Fig. 7.1 B) (Marriott et al., 2015), a small change to the structure of the phosphate chain is tolerated. This is demonstrated by the activity of AppNppA in inhibiting the initiation of DNA replication (Fig. 6.6). Finally, different Ap<sub>4</sub>Ns and Ap<sub>4</sub>dNs are capable of inhibiting DNA replication initiation (Fig. 6.8), which demonstrates that the second nucleoside in the molecule is not limited to adenosine and shows that both ribonucleoside and deoxyribonucleoside structures are tolerated.

#### **7.4.3 Limitations and further method developments for investigating the role of Ap<sub>4</sub>Ns in DNA replication initiation**

A limitation of these experiments is the variability between different batches of nuclei, with some batches of nuclei initiating better than others when combined with an S-phase extract. Due to the nature of the synchronisation method, only a limited number of nuclei can be produced simultaneously. This meant that not all Ap<sub>4</sub>Ns could be tested using one batch of nuclei. However, to minimise the variability brought about by different batches, multiple repeats were performed with different batches of nuclei for each Ap<sub>4</sub>N. Another limitation of this method is the possibility for human bias to occur. Therefore, after scoring the effect of Ap<sub>4</sub>N on DNA replication for the first set of nuclei, future batches were scored blind to prevent this. Future work could aim to replicate this data with further batches of synchronised nuclei. Further support for the effect of Ap<sub>4</sub>Ns on the DNA replication process could be addressed through their effect on the DNA replication machinery.

#### **7.5 Ap<sub>4</sub>Ns reduce nuclear Mcm2 and PCNA levels in late G1 nuclei added to S-phase extract, with a greater effect on PCNA levels**

The precise point at which Ap<sub>4</sub>A (and Ap<sub>4</sub>Ns) have their effect on DNA replication is currently unknown, as there are different stages to the DNA replication process which Ap<sub>4</sub>N could be affecting. In 1979, Grummt and colleagues demonstrated that Ap<sub>4</sub>A binds to the 57 kDa subunit of DNA pol  $\alpha$  (Grummt et al., 1979). This subunit was later

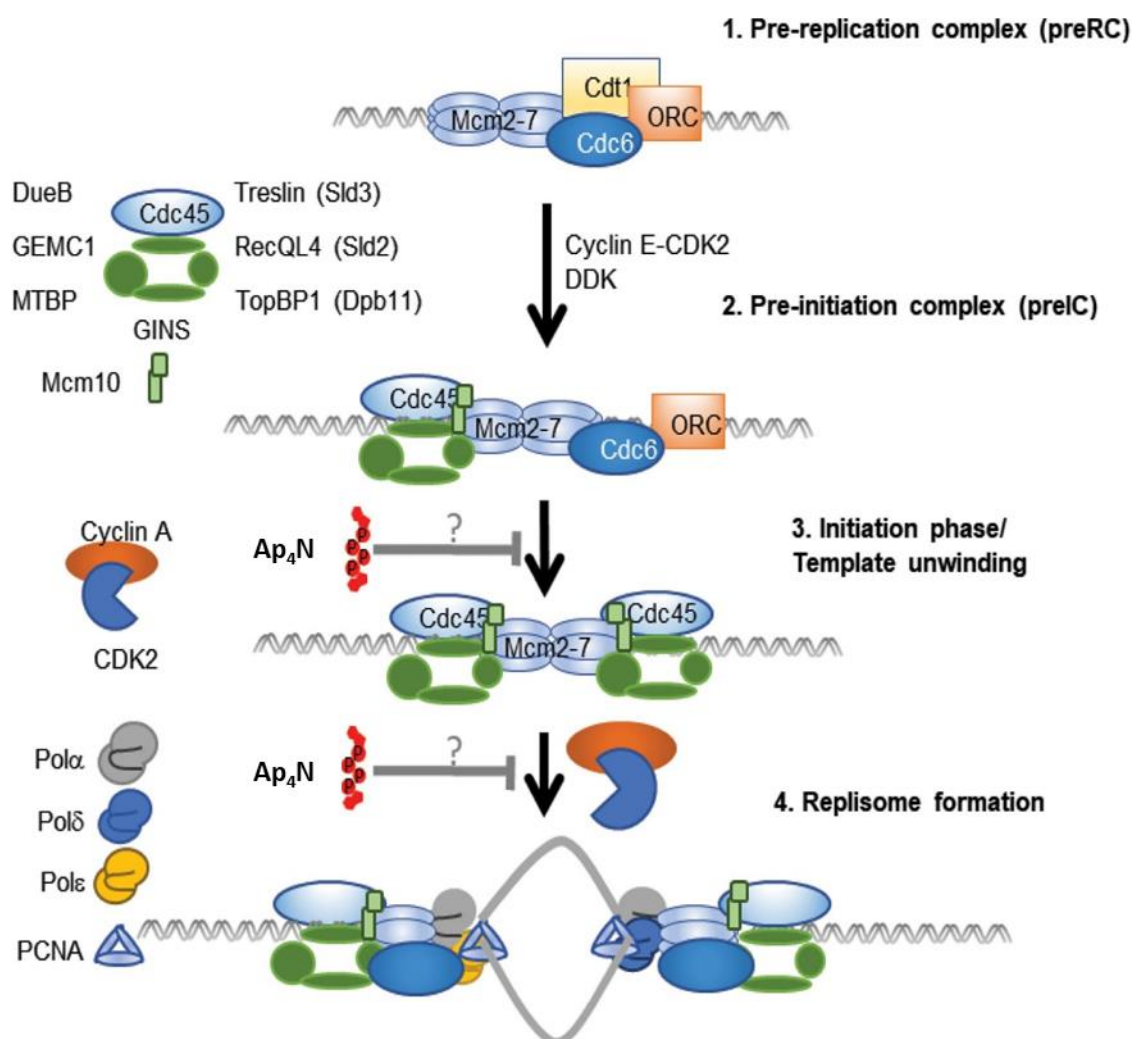
was later resolved into two polypeptides of 45 kDa and 22 kDa, with Ap<sub>4</sub>A only acting as a ligand for the former (Baxi et al., 1994). Interestingly Ap<sub>4</sub>G, but not Gp<sub>4</sub>G, was shown to compete with Ap<sub>4</sub>A for binding to this 45 kDa polypeptide (Baxi et al., 1994). This is therefore in line with our data suggesting that Ap<sub>4</sub>G can inhibit DNA replication initiation, while supporting previous evidence suggesting that Gp<sub>4</sub>G does not affect DNA replication initiation (Marriott et al., 2015). Therefore, it is possible that Ap<sub>4</sub>A and the Ap<sub>4</sub>Ns function through binding to the 45 kDa polypeptide in the DNA pol  $\alpha$  protein complex. If this were solely the case, we would not expect the recruitment of DNA replication proteins to be affected until during the formation of the replisome, after the involvement of DNA pol  $\alpha$  in the DNA replication process.

Alternatively, many of the processes involved in the formation of the replication machinery are known to rely on ATP and its hydrolysis (Bell and Stillman, 1992; Chang et al., 2015; Feng et al., 2021; Fernández-Cid et al., 2013; Speck and Stillman, 2007). As each end of the Ap<sub>4</sub>A molecule has structural homology with ATP, it is possible that Ap<sub>4</sub>A could interfere with processes requiring ATP hydrolysis by competing for the ATP binding site. Although Ap<sub>4</sub>A is present at a much lower concentration than ATP, if Ap<sub>4</sub>A were to bind with a higher affinity it is possible that it may still compete with ATP. An example of this occurs in the interaction between IMPDH and Ap<sub>4</sub>A in *B. subtilis*, where the K<sub>d</sub> for Ap<sub>4</sub>A of IMPDH is in the micromolar range compared with the millimolar range for AMP and ATP (Giammarinaro et al., 2022). Similarly, Ap<sub>4</sub>Ns compete with ATP for incorporation onto RNA transcripts as protective caps in *E. coli* (Luciano and Belasco, 2020; Luciano et al., 2019). Interestingly, a recent study investigating Ap<sub>4</sub>A binding proteins demonstrated that 46% of the proteins identified were known nucleotide binders (Krüger et al., 2021). This provides further evidence supporting a role for competition with the relevant nucleotides in the inhibition of DNA replication.

#### **7.5.1 The nuclear localisation of replicative helicase and processivity factor are affected by Ap<sub>4</sub>Ns**

To determine which stage of the initiation of DNA replication is inhibited by the Ap<sub>4</sub>Ns, late G1-phase nuclei were added to an S-phase extract and Mcm2 and PCNA levels were analysed in the presence and absence of each Ap<sub>4</sub>N. In the initial stages of the

replication process, Mcm2–7 is loaded onto DNA origins to form the pre-RC (Tsakraklides and Bell, 2010). On the other hand, PCNA is not incorporated in the replication process until the replisome assembles, where it stimulates Pol  $\delta$  and Pol  $\epsilon$  during the formation of the replisome (Fig. 7.2) (Chilkova et al., 2007).



**Figure 7.2: Stages of DNA replication that may be affected by Ap<sub>4</sub>N.** DNA replication begins with the formation of the pre-RC and localisation of Mcm2–7 to chromatin, followed by formation of the pre-IC. Subsequently, the replisome is formed in a process which involves PCNA. Ap<sub>4</sub>N inhibits the initiation of DNA replication and results in reduced localisation of replication proteins to chromatin. While some Ap<sub>4</sub>Ns reduced localisation of Mcm2 to chromatin, a greater reduction was seen in the localisation of PCNA to chromatin. This suggests that Ap<sub>4</sub>N primarily acts after Mcm2–7 localisation to chromatin, but before the involvement of PCNA in the replisome formation stage of DNA replication. Figure adapted from Ferguson et al., 2020.

Figure reproduced from Frontiers in Molecular Biosciences with minor changes, Volume 7, Ferguson, F., McLennan, A. G., Urbaniak, M. D., Jones, N. J. & Copeland, N. A., Re-evaluation of Diadenosine Tetraphosphate (Ap<sub>4</sub>A) From a Stress Metabolite to Bona Fide Secondary Messenger, Copyright 2020, <https://doi.org/10.3389/fmolb.2020.606807>; available under the CC BY 4.0 licence (<https://creativecommons.org/licenses/by/4.0/legalcode>) and disclaimer of warranties within.

Unlike Mcm2, although PCNA is required for maximal replication rates, it is not an essential component of the minimal replication complex (Yeeles et al., 2015; Yeeles et al., 2017). Similarly, although Ap<sub>4</sub>A has a role in DNA replication, it does not belong to this minimum group of components necessary for replication to occur (Marriott et al., 2016; Yeeles et al., 2015). When Ap<sub>4</sub>N was added to the cell-free replication mix, there was a trend suggesting that the nuclear abundance of both Mcm2 and PCNA was reduced, with PCNA levels being reduced to a greater extent than Mcm2 (Fig. 6.9, 6.11 and 6.12). This suggests that although Ap<sub>4</sub>A may have some effect prior to Mcm2 localisation, its primary effect seems to occur after Mcm2 has localised to the replication complex (Fig. 7.2).

#### **7.5.2 Limitations and future perspectives of Western blotting for determining the mechanism behind Ap<sub>4</sub>N-induced inhibition of DNA replication**

Although there was a clear trend towards a reduction in nuclear Mcm2 and PCNA levels, the data were not statistically significant. This was likely due to the nature and lack of accuracy of the Western blotting process. To improve these results, further steps could include repetition of these experiments so that a higher n number could be obtained. Alternatively, experiments could be performed using fluorescent antibodies, which have an extended dynamic range compared with traditional ECL techniques and can be considered quantitative (Eaton et al., 2014; Zellner et al., 2008). Other than increasing the significance of the data, further steps could be taken to narrow down the window in which Ap<sub>4</sub>N is inhibiting the DNA replication process. This could be done by probing for other DNA replication proteins to determine whether these are affected by Ap<sub>4</sub>N. Suitable replication proteins to probe for could include members of the pre-IC, which is formed between the pre-RC and replisome formation stages of DNA replication. An effect or lack of effect of Ap<sub>4</sub>N on these protein levels would narrow down the range of time over which Ap<sub>4</sub>N may be acting.

#### **7.6 Conclusion**

In summary, we have demonstrated evidence supporting a biological mechanism for the synthesis of all four different Ap<sub>4</sub>Ns by UBE1 in a biologically relevant mechanism that is activated under conditions of cell stress. The capability of this mechanism to

synthesise Ap<sub>4</sub>Ns provides evidence of a role for the Ap<sub>4</sub>Ns during times of cell stress. Further to this we have developed a technique for the extraction and identification of Ap<sub>4</sub>Ns in mammalian cells by triple quadrupole mass spectrometry. This technique enabled us to demonstrate for the first time that all four Ap<sub>4</sub>Ns and some Ap<sub>4</sub>dNs are present at varying concentrations in the NuKO cell line. With further development, this technique could be used for the accurate quantification of each Ap<sub>4</sub>N, and to investigate the abundance of each Ap<sub>4</sub>N in the parental KBM7 cell line. Additionally, we have developed a technique for measuring the change in Ap<sub>4</sub>N abundance in response to different types of stresses in mammalian NuKO cells. With further development and the identification of an appropriate standard, this could then be applied to investigate the effect of these stresses on Ap<sub>4</sub>N abundance in the KBM7 cell line, and to determine whether the Ap<sub>4</sub>N response is the same after treatment with these different types of DNA damage. Finally, we have shown that each of the Ap<sub>4</sub>Ns inhibit the initiation of DNA replication to different extents in a DNA cell-free replication system, and that this process involves a reduction in the localisation of replication proteins to chromatin, with a particularly strong effect on PCNA localisation. Together, these data provide evidence for the presence of Ap<sub>4</sub>Ns in mammalian cells, demonstrate a potential role for each of the Ap<sub>4</sub>Ns in inhibiting DNA replication, and provide evidence that Ap<sub>4</sub>Ns exert their primary effect after formation of the pre-RC but prior to replisome formation. With additional investigation, this could further narrow down the precise point at which the Ap<sub>4</sub>Ns inhibit DNA replication when cells are under stress.

# **Appendices and Supplementary data**



A

Data File: C:\LabSolutions\Data\2 University\Urbaniak, M (BLS)\Ferguson, Freya\191011 Ferguson HRMS Assignment\210428\Ap4A-210428-00...

Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Use Adduct
H	1	20	30	Na	1	0	3	Fe	2	0	0	I	3	0	0	Cl
2H	1	0	0	Si	4	0	1	Co	2	0	0	Re	2	0	0	H
B	3	0	0	P	3	0	5	Ni	2	0	0	Ir	3	0	0	Br
C	4	19	20	S	2	0	2	Cu	2	0	0	Pt	2	0	0	I
N	3	6	12	Cl	1	0	0	Br	1	0	0					CHO2
O	2	16	22	V	2	0	0	Ru	2	0	0					
F	1	0	0	Cr	2	0	0	Ag	1	0	0					

Error Margin (ppm): 20

DBE Range: -2.0 - 1000.0

Electron Ions: both

HC Ratio: unlimited

Apply N Rule: no

Use MSn Info: yes

Max Isotopes: all

Isotope RI (%): 1.00

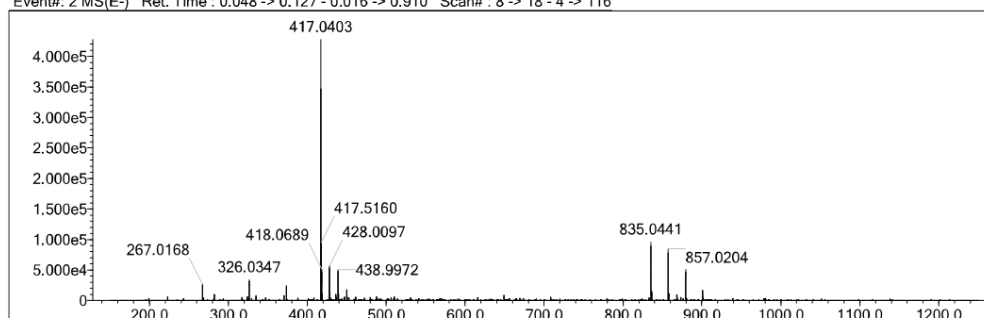
Isotope Res: 6000

MSn Iso RI (%): 75.00

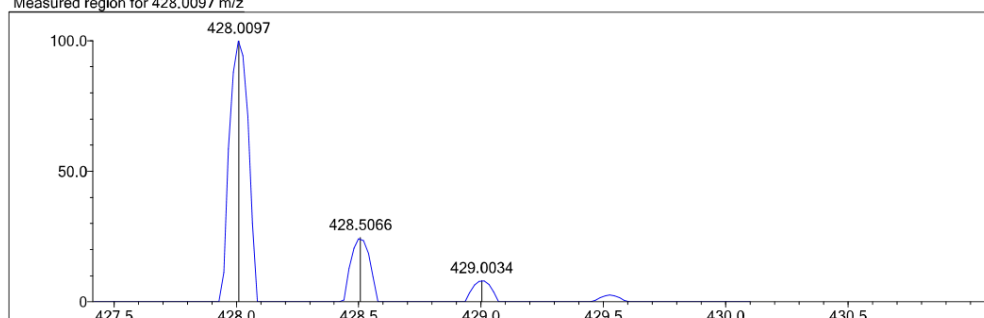
MSn Logic Mode: AND

Max Results: 500

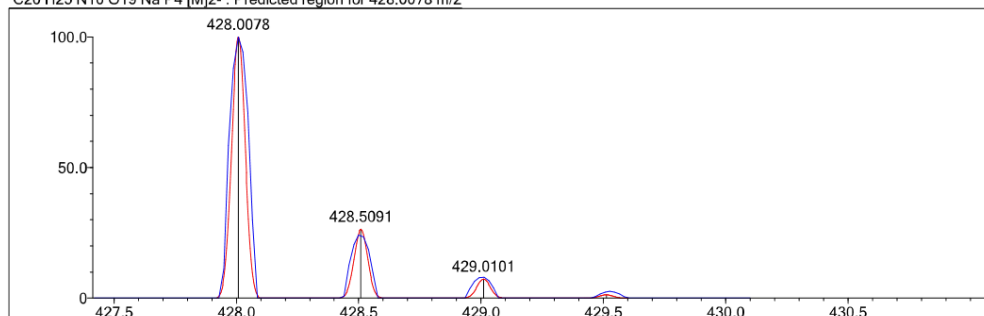
Event#: 2 MS(E-) Ret. Time : 0.048 -&gt; 0.127 - 0.016 -&gt; 0.910 Scan#: 8 -&gt; 18 - 4 -&gt; 116



Measured region for 428.0097 m/z



C20 H25 N10 O19 Na P4 [M]2- : Predicted region for 428.0078 m/z



Rank	Score	Formula (M)	Ion	Meas. m/z	Pred. m/z	Df. (mDa)	Df. (ppm)	Iso	DBE
33	76.07	C20 H25 N10 O19 Na P4	[M]2-	428.0097	428.0078	1.9	4.44	83.23	15.0

## B

Formula Predictor Report - Ap4A-210428-004.lcd

Page 1 of 1

Data File: C:\LabSolutions\Data\2 University\Urbaniak, M (BLS)\Ferguson, Freya\191011 Ferguson HRMS Assignment\210428\Ap4A-210428-00...

Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Use Adduct
H	1	20	30	Na	1	0	0	Fe	2	0	0	I	3	0	0	Cl
2H	1	0	0	Si	4	0	1	Co	2	0	0	Re	2	0	0	H
B	3	0	0	P	3	0	5	Ni	2	0	0	Ir	3	0	0	Br
C	4	19	20	S	2	0	2	Cu	2	0	0	Pt	2	0	0	I
N	3	6	12	Cl	1	0	0	Br	1	0	0					CHO2
O	2	16	22	V	2	0	0	Ru	2	0	0					
F	1	0	0	Cr	2	0	0	Ag	1	0	0					

Error Margin (ppm): 20

DBE Range: -2.0 - 1000.0

Electron Ions: both

HC Ratio: unlimited

Apply N Rule: no

Use MSn Info: yes

Max Isotopes: all

Isotope RI (%): 1.00

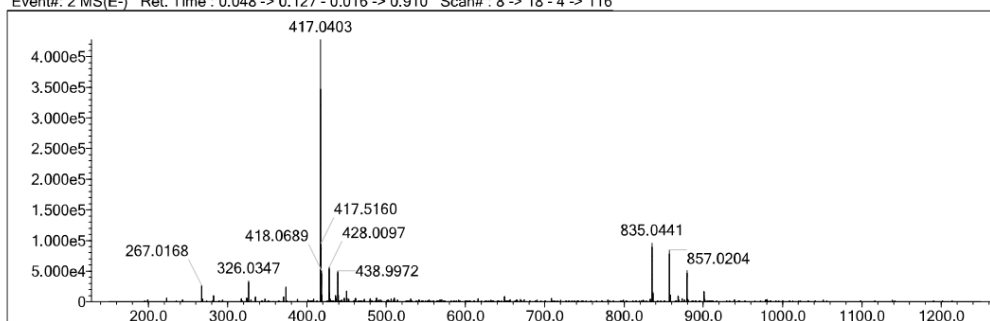
Isotope Res: 6000

MSn Iso RI (%): 75.00

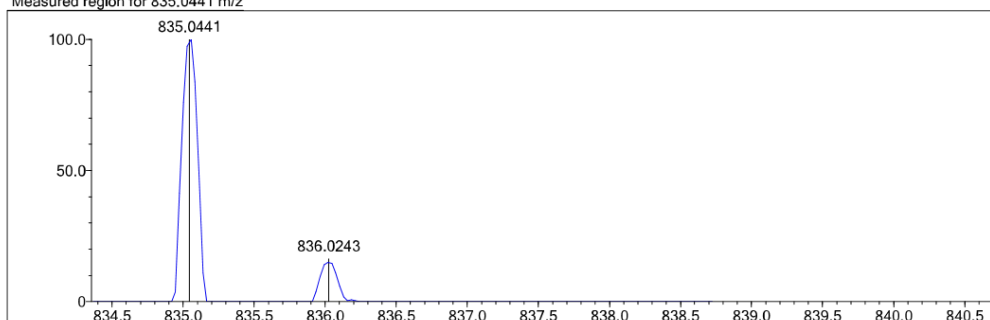
MSn Logic Mode: AND

Max Results: 500

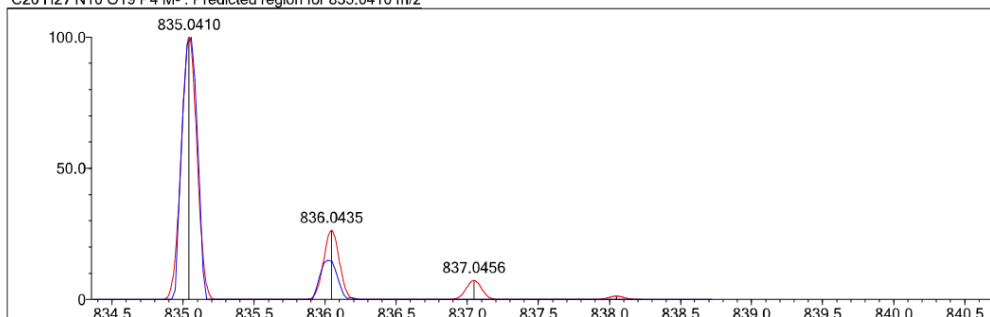
Event#: 2 MS(E-) Ret. Time : 0.048 -&gt; 0.127 - 0.016 -&gt; 0.910 Scan#: 8 -&gt; 18 - 4 -&gt; 116



Measured region for 835.0441 m/z



C20 H27 N10 O19 P4 M- : Predicted region for 835.0410 m/z



Rank	Score	Formula (M)	Ion	Meas. m/z	Pred. m/z	Df. (mDa)	Df. (ppm)	Iso	DBE
43	0.00	C20 H27 N10 O19 P4	M-	835.0441	835.0410	3.1	3.71	0.00	14.5

C

Data File: C:\LabSolutions\Data\2 University\Urbaniak, M (BLS)\Ferguson, Freya\191011 Ferguson HRMS Assignment\210428\Ap4C-210428-00...

Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Use Adduct
H	1	20	30	Na	1	0	3	Fe	2	0	0	I	3	0	0	Cl
2H	1	0	0	Si	4	0	1	Co	2	0	0	Re	2	0	0	H
B	3	0	0	P	3	4	4	Ni	2	0	0	Ir	3	0	0	Br
C	4	19	20	S	2	0	0	Cu	2	0	0	Pt	2	0	0	I
N	3	6	12	Cl	1	0	0	Br	1	0	0					CHO2
O	2	16	22	V	2	0	0	Ru	2	0	0					
F	1	0	0	Cr	2	0	0	Ag	1	0	0					

Error Margin (ppm): 20

DBE Range: -2.0 - 1000.0

Electron Ions: both

HC Ratio: unlimited

Apply N Rule: no

Use MSn Info: yes

Max Isotopes: all

Isotope RI (%): 1.00

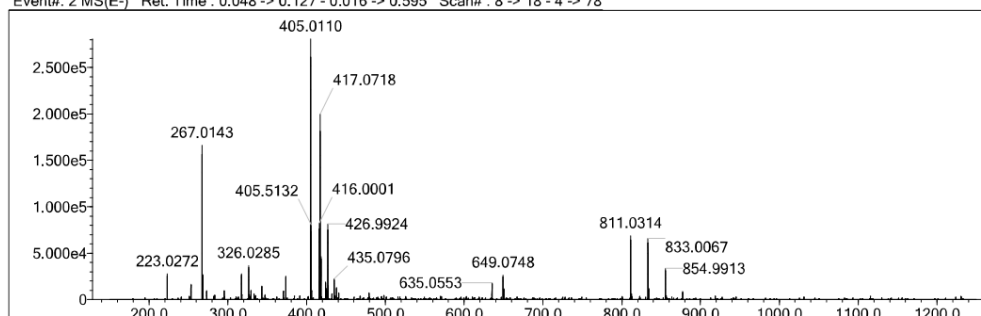
Isotope Res: 6000

MSn Iso RI (%): 75.00

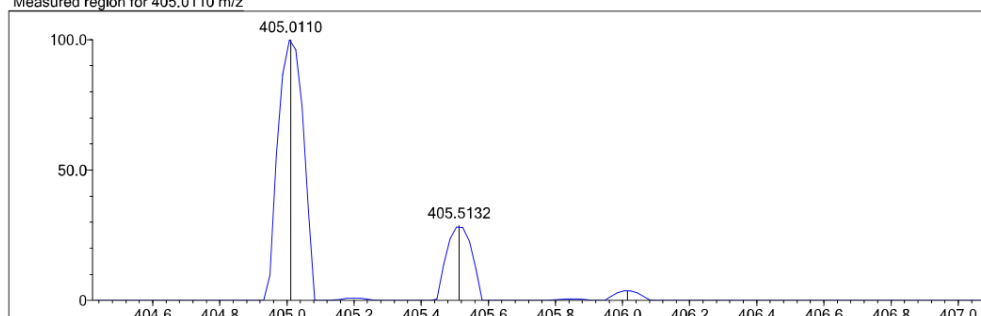
MSn Logic Mode: AND

Max Results: 500

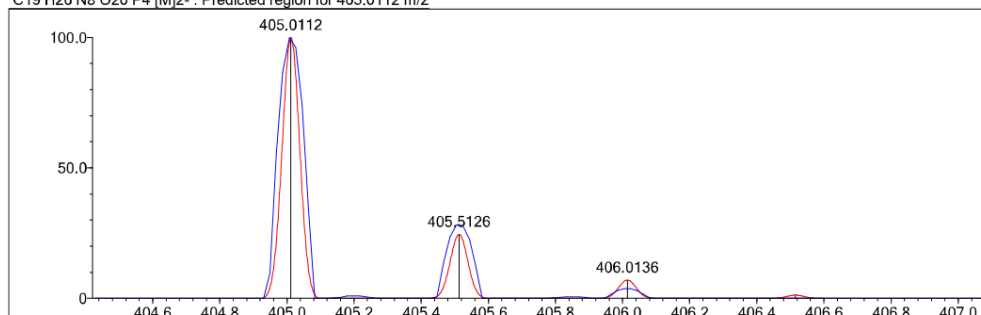
Event#: 2 MS(E-) Ret. Time : 0.048 -&gt; 0.127 - 0.016 -&gt; 0.595 Scan#: 8 -&gt; 18 - 4 -&gt; 78



Measured region for 405.0110 m/z



C19 H26 N8 O20 P4 [M]2- : Predicted region for 405.0112 m/z



Rank	Score	Formula (M)	Ion	Meas. m/z	Pred. m/z	Df. (mDa)	Df. (ppm)	Iso	DBE
8	87.30	C19 H26 N8 O20 P4	[M]2-	405.0110	405.0112	-0.2	-0.49	87.30	13.0

D

Data File: C:\LabSolutions\Data\2 University\Urbaniak, M (BLS)\Ferguson, Freya\191011 Ferguson HRMS Assignment\210428\Ap4C-210428-00...

Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Use Adduct
H	1	20	30	Na	1	0	4	Fe	2	0	0	I	3	0	0	H
2H	1	0	0	Si	4	0	1	Co	2	0	0	Re	2	0	0	Na
B	3	0	0	P	3	4	4	Ni	2	0	0	Ir	3	0	0	K
C	4	19	20	S	2	0	0	Cu	2	0	0	Pt	2	0	0	NH4
N	3	6	12	Cl	1	0	0	Br	1	0	0					
O	2	16	22	V	2	0	0	Ru	2	0	0					
F	1	0	0	Cr	2	0	0	Ag	1	0	0					

Error Margin (ppm): 20

DBE Range: -2.0 - 1000.0

Electron Ions: both

HC Ratio: unlimited

Apply N Rule: no

Use MSn Info: yes

Max Isotopes: all

Isotope RI (%): 1.00

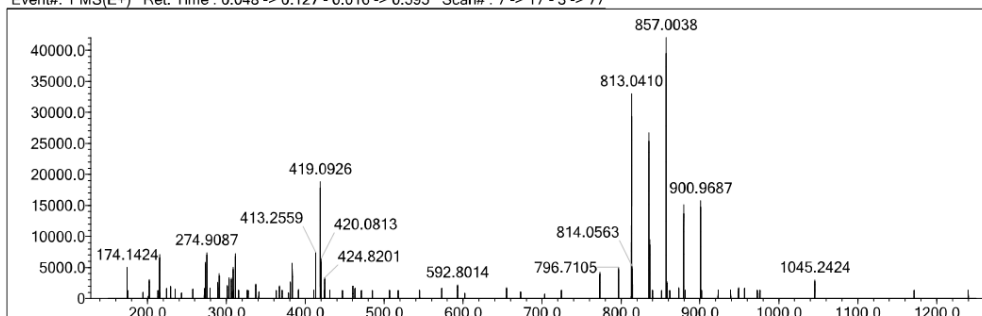
Isotope Res: 6000

MSn Iso RI (%): 75.00

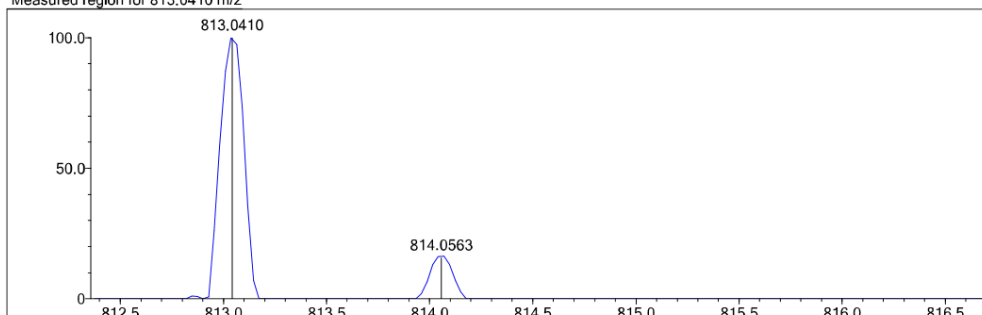
MSn Logic Mode: AND

Max Results: 500

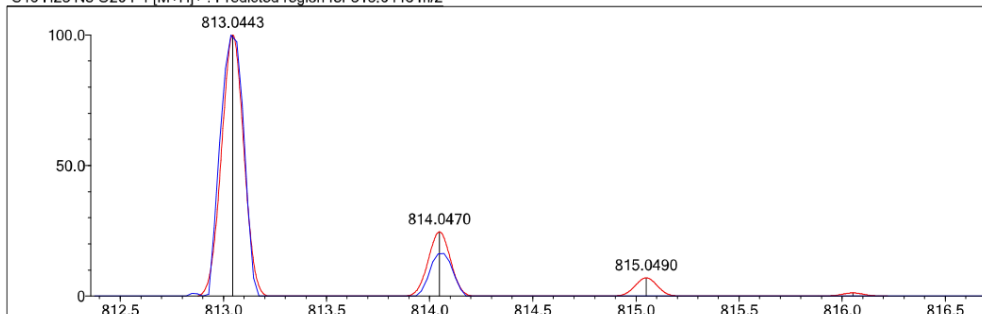
Event#: 1 MS(E+) Ret. Time: 0.048 -&gt; 0.127 - 0.016 -&gt; 0.595 Scan#: 7 -&gt; 17 - 3 -&gt; 77



Measured region for 813.0410 m/z



C19 H28 N8 O20 P4 [M+H]+ : Predicted region for 813.0443 m/z



Rank	Score	Formula (M)	Ion	Meas. m/z	Pred. m/z	Df. (mDa)	Df. (ppm)	Iso	DBE
8	0.00	C19 H28 N8 O20 P4	[M+H]+	813.0410	813.0443	-3.3	-4.06	0.00	12.0

Data File: C:\LabSolutions\Data\2 University\Urbaniak, M (BLS)\Ferguson, Freya\191011 Ferguson HRMS Assignment\210428\Ap4G-210428-00...

Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Use Adduct
H	1	20	30	Na	1	0	3	Fe	2	0	0	I	3	0	0	Cl
2H	1	0	0	Si	4	0	1	Co	2	0	0	Re	2	0	0	H
B	3	0	0	P	3	4	4	Ni	2	0	0	Ir	3	0	0	Br
C	4	19	20	S	2	0	0	Cu	2	0	0	Pt	2	0	0	I
N	3	6	12	Cl	1	0	0	Br	1	0	0					CHO2
O	2	16	22	V	2	0	0	Ru	2	0	0					
F	1	0	0	Cr	2	0	0	Ag	1	0	0					

Error Margin (ppm): 20

DBE Range: -2.0 - 1000.0

Electron Ions: both

HC Ratio: unlimited

Apply N Rule: no

Use MSn Info: yes

Max Isotopes: all

Isotope RI (%): 1.00

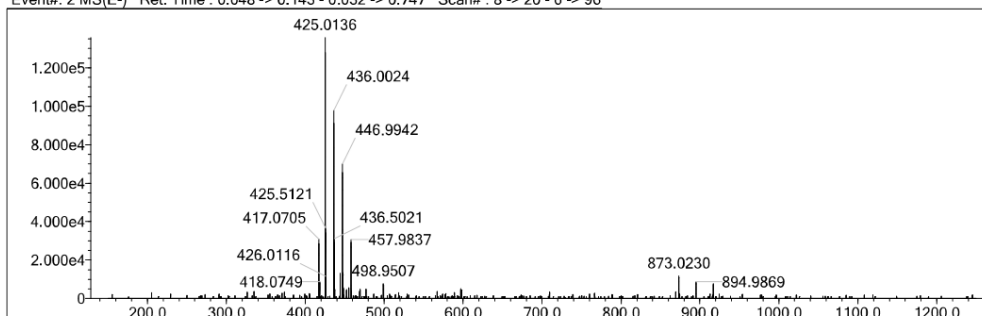
Isotope Res: 6000

MSn Iso RI (%): 75.00

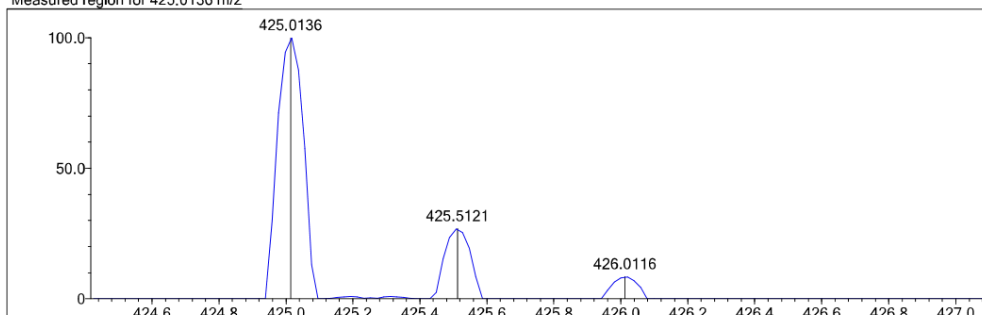
MSn Logic Mode: AND

Max Results: 500

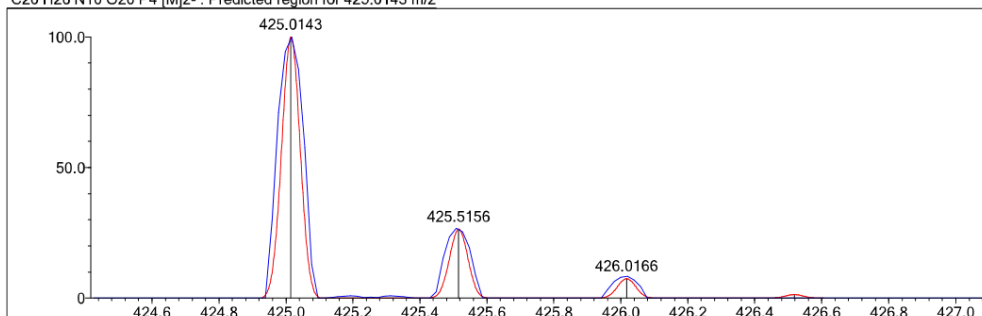
Event#: 2 MS(E-) Ret. Time : 0.048 -&gt; 0.143 - 0.032 -&gt; 0.747 Scan#: 8 -&gt; 20 - 6 -&gt; 96



Measured region for 425.0136 m/z



C20 H26 N10 O20 P4 [M]2- : Predicted region for 425.0143 m/z



Rank	Score	Formula (M)	Ion	Meas. m/z	Pred. m/z	Df. (mDa)	Df. (ppm)	Iso	DBE
4	98.20	C20 H26 N10 O20 P4	[M]2-	425.0136	425.0143	-0.7	-1.65	99.82	15.0

Data File: C:\LabSolutions\Data\2 University\Urbaniak, M (BLS)\Ferguson, Freya\191011 Ferguson HRMS Assignment\210428\Ap4G-210428-00...

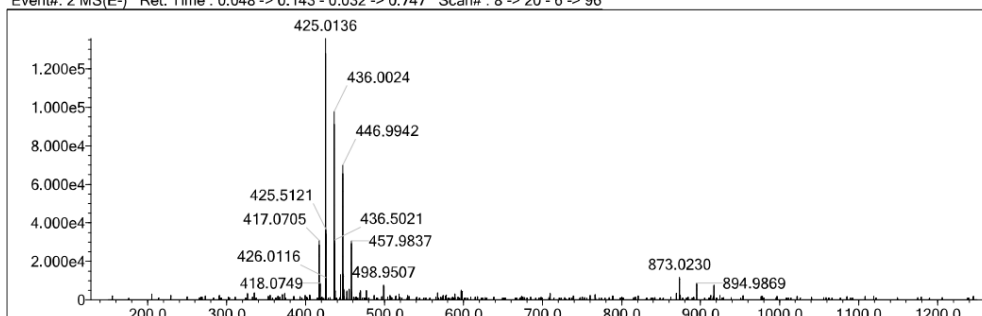
Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Use Adduct
H	1	20	30	Na	1	0	3	Fe	2	0	0	I	3	0	0	Cl
2H	1	0	0	Si	4	0	1	Co	2	0	0	Re	2	0	0	H
B	3	0	0	P	3	4	4	Ni	2	0	0	Ir	3	0	0	Br
C	4	19	20	S	2	0	0	Cu	2	0	0	Pt	2	0	0	I
N	3	6	12	Cl	1	0	0	Br	1	0	0					CHO2
O	2	16	22	V	2	0	0	Ru	2	0	0					
F	1	0	0	Cr	2	0	0	Ag	1	0	0					

Error Margin (ppm): 20  
 HC Ratio: unlimited  
 Max Isotopes: all  
 MSn Iso RI (%): 75.00

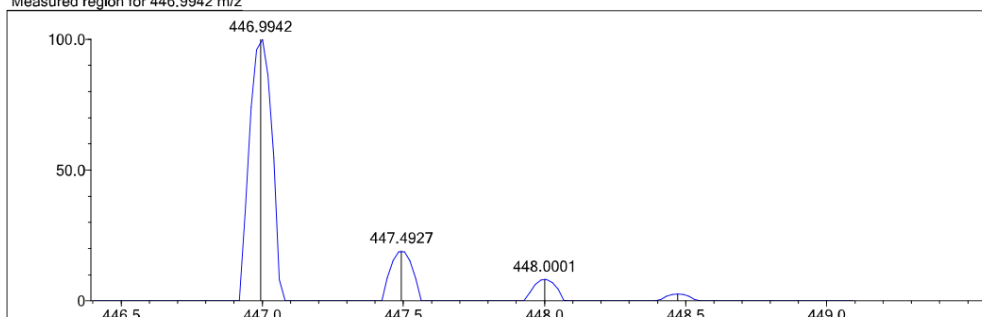
DBE Range: -2.0 - 1000.0  
 Apply N Rule: no  
 Isotope RI (%): 1.00  
 MSn Logic Mode: AND

Electron Ions: both  
 Use MSn Info: yes  
 Isotope Res: 6000  
 Max Results: 500

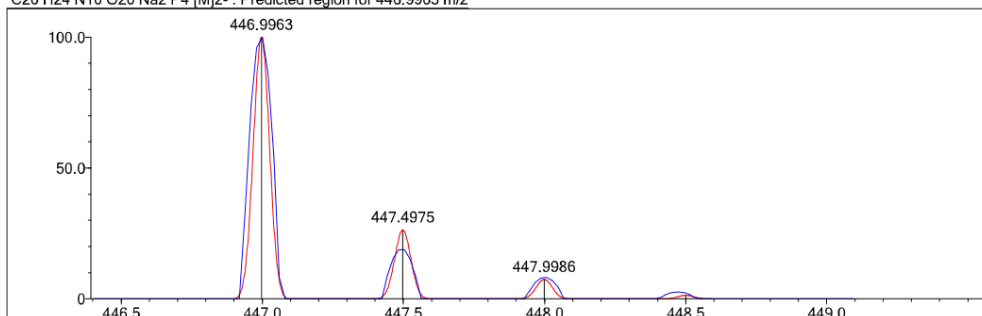
Event#: 2 MS(E-) Ret. Time : 0.048 -&gt; 0.143 - 0.032 -&gt; 0.747 Scan#: 8 -&gt; 20 - 6 -&gt; 96



Measured region for 446.9942 m/z



C20 H24 N10 O20 Na2 P4 [M]2- : Predicted region for 446.9963 m/z



Rank	Score	Formula (M)	Ion	Meas. m/z	Pred. m/z	Df. (mDa)	Df. (ppm)	Iso	DBE
3	53.02	C20 H24 N10 O20 Na2 P4	[M]2-	446.9942	446.9963	-2.1	-4.70	58.43	15.0

Data File: C:\LabSolutions\Data\2 University\Urbaniak, M (BLS)\Ferguson, Freya\191011 Ferguson HRMS Assignment\210428\Ap4V-210428-00...

Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Use Adduct
H	1	20	30	Na	1	0	4	Fe	2	0	0	I	3	0	0	Cl
2H	1	0	0	Si	4	0	1	Co	2	0	0	Re	2	0	0	H
B	3	0	0	P	3	4	4	Ni	2	0	0	Ir	3	0	0	Br
C	4	19	20	S	2	0	0	Cu	2	0	0	Pt	2	0	0	I
N	3	6	12	Cl	1	0	0	Br	1	0	0					CHO2
O	2	16	22	V	2	0	0	Ru	2	0	0					
F	1	0	0	Cr	2	0	0	Ag	1	0	0					

Error Margin (ppm): 20

DBE Range: -2.0 - 1000.0

Electron Ions: both

HC Ratio: unlimited

Apply N Rule: no

Use MSn Info: yes

Max Isotopes: all

Isotope RI (%): 1.00

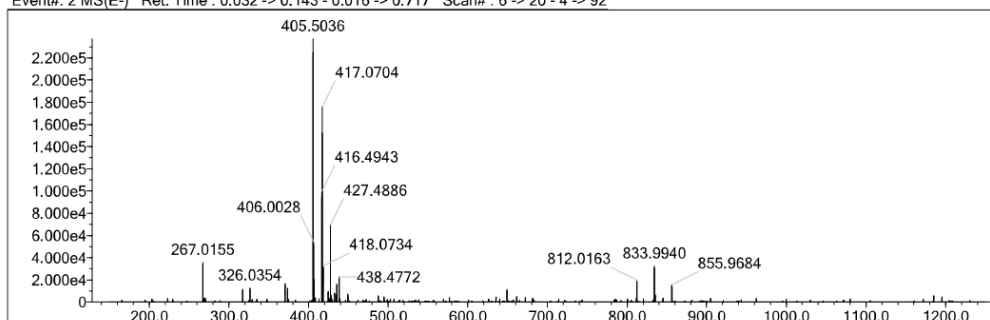
Isotope Res: 6000

MSn Iso RI (%): 75.00

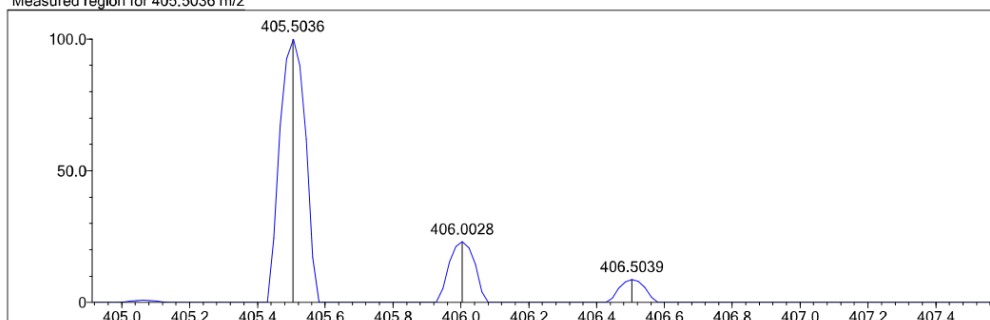
MSn Logic Mode: AND

Max Results: 500

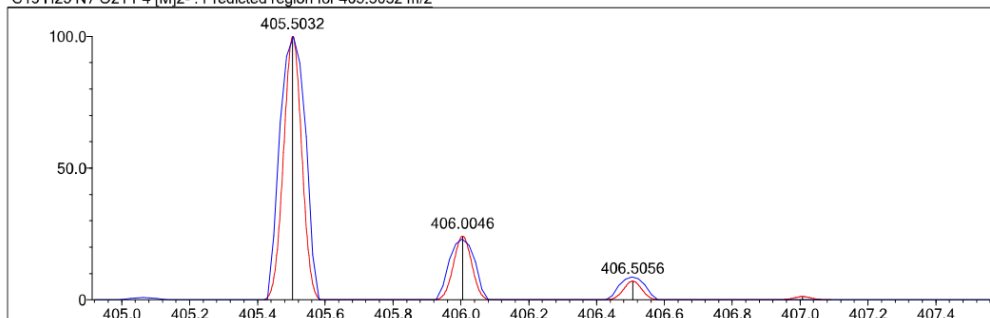
Event#: 2 MS(E-) Ret. Time : 0.032 -&gt; 0.143 - 0.016 -&gt; 0.717 Scan#: 6 -&gt; 20 - 4 -&gt; 92



Measured region for 405.5036 m/z



C19 H25 N7 O21 P4 [M]2- : Predicted region for 405.5032 m/z



Rank	Score	Formula (M)	Ion	Meas. m/z	Pred. m/z	Df. (mDa)	Df. (ppm)	Iso	DBE
2	100.00	C19 H25 N7 O21 P4	[M]2-	405.5036	405.5032	0.4	0.99	100.00	13.0

Data File: C:\LabSolutions\Data\2 University\Urbaniak, M (BLS)\Ferguson, Freya\191011 Ferguson HRMS Assignment\210428\Ap4V-210428-00...

Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Use Adduct
H	1	20	30	Na	1	0	4	Fe	2	0	0	I	3	0	0	Cl
2H	1	0	0	Si	4	0	1	Co	2	0	0	Re	2	0	0	H
B	3	0	0	P	3	4	4	Ni	2	0	0	Ir	3	0	0	Br
C	4	19	20	S	2	0	0	Cu	2	0	0	Pt	2	0	0	I
N	3	6	12	Cl	1	0	0	Br	1	0	0					CHO2
O	2	16	22	V	2	0	0	Ru	2	0	0					
F	1	0	0	Cr	2	0	0	Ag	1	0	0					

Error Margin (ppm): 20

DBE Range: -2.0 - 1000.0

Electron Ions: both

HC Ratio: unlimited

Apply N Rule: no

Use MSn Info: yes

Max Isotopes: all

Isotope RI (%): 1.00

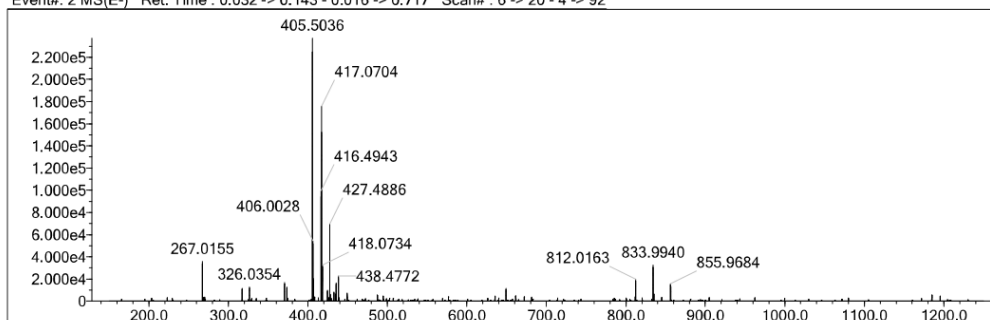
Isotope Res: 6000

MSn Iso RI (%): 75.00

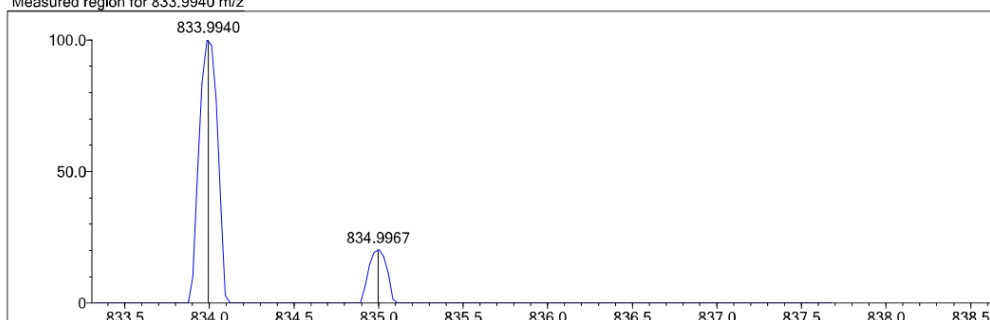
MSn Logic Mode: AND

Max Results: 500

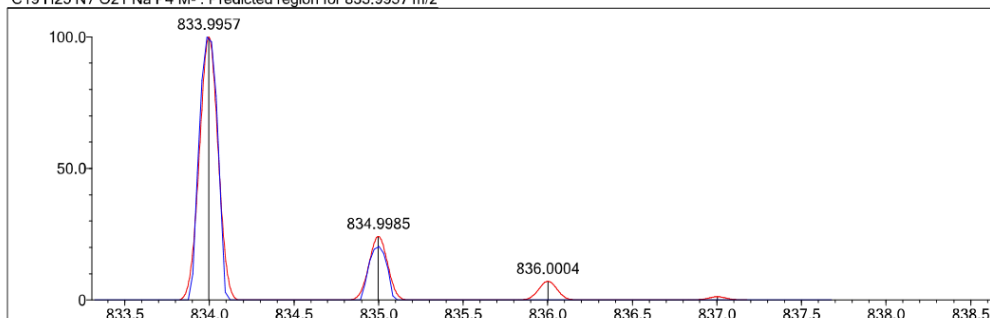
Event#: 2 MS(E-) Ret. Time : 0.032 -&gt; 0.143 - 0.016 -&gt; 0.717 Scan#: 6 -&gt; 20 - 4 -&gt; 92



Measured region for 833.9940 m/z



C19 H25 N7 O21 Na P4 M- : Predicted region for 833.9957 m/z



Rank	Score	Formula (M)	Ion	Meas. m/z	Pred. m/z	Df. (mDa)	Df. (ppm)	Iso	DBE
4	0.00	C19 H25 N7 O21 Na P4	M-	833.9940	833.9957	-1.7	-2.04	0.00	12.5



Data File: C:\LabSolutions\Data\2 University\Urbaniak, M (BLS)\Ferguson, Freya\191011 Ferguson HRMS Assignment\210428\Ap4dA-210428-0...

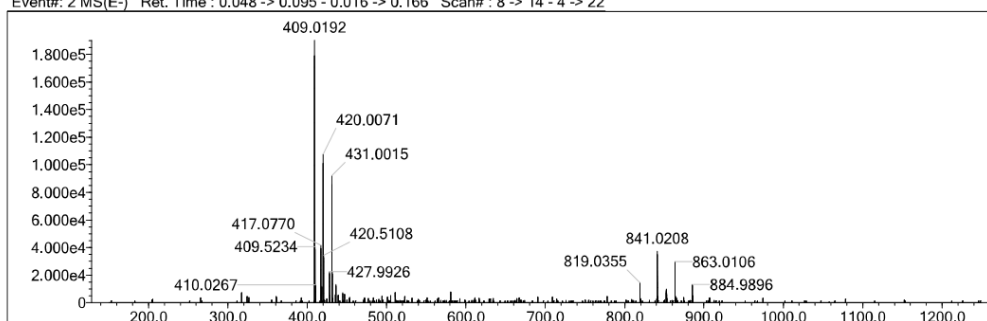
Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Use Adduct
H	1	20	30	Na	1	0	4	Fe	2	0	0	I	3	0	0	Cl
2H	1	0	0	Si	4	0	1	Co	2	0	0	Re	2	0	0	H
B	3	0	0	P	3	4	4	Ni	2	0	0	Ir	3	0	0	Br
C	4	19	20	S	2	0	0	Cu	2	0	0	Pt	2	0	0	I
N	3	6	12	Cl	1	0	0	Br	1	0	0					CHO2
O	2	16	22	V	2	0	0	Ru	2	0	0					
F	1	0	0	Cr	2	0	0	Ag	1	0	0					

Error Margin (ppm): 20  
 HC Ratio: unlimited  
 Max Isotopes: all  
 MSn Iso RI (%): 75.00

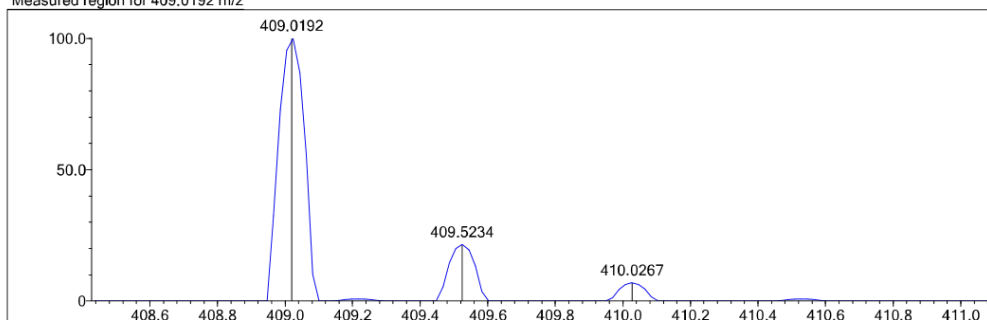
DBE Range: -2.0 - 1000.0  
 Apply N Rule: no  
 Isotope RI (%): 1.00  
 MSn Logic Mode: AND

Electron Ions: both  
 Use MSn Info: yes  
 Isotope Res: 6000  
 Max Results: 500

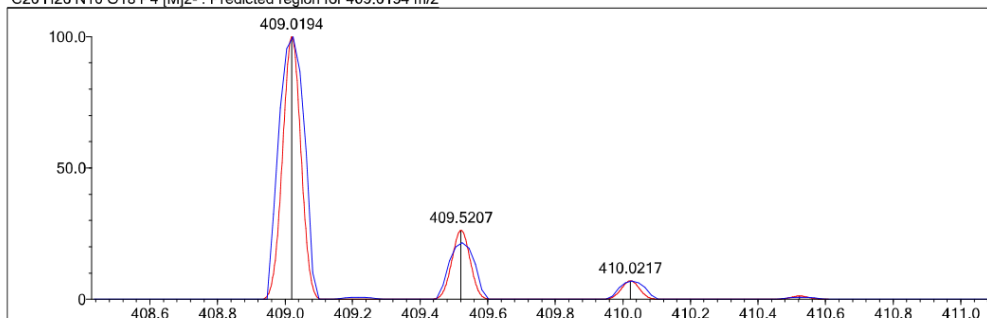
Event#: 2 MS(E-) Ret. Time : 0.048 -&gt; 0.095 - 0.016 -&gt; 0.166 Scan#: 8 -&gt; 14 - 4 -&gt; 22



Measured region for 409.0192 m/z



C20 H26 N10 O18 P4 [M]2- : Predicted region for 409.0194 m/z



Rank	Score	Formula (M)	Ion	Meas. m/z	Pred. m/z	Df. (mDa)	Df. (ppm)	Iso	DBE
3	94.90	C20 H26 N10 O18 P4	[M]2-	409.0192	409.0194	-0.2	-0.49	94.90	15.0

Data File: C:\LabSolutions\Data\2 University\Urbaniak, M (BLS)\Ferguson, Freya\191011 Ferguson HRMS Assignment\210428\Ap4dA-210428-0...

Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Use Adduct
H	1	20	30	Na	1	0	4	Fe	2	0	0	I	3	0	0	Cl
2H	1	0	0	Si	4	0	1	Co	2	0	0	Re	2	0	0	H
B	3	0	0	P	3	4	4	Ni	2	0	0	Ir	3	0	0	Br
C	4	19	20	S	2	0	0	Cu	2	0	0	Pt	2	0	0	I
N	3	6	12	Cl	1	0	0	Br	1	0	0					CHO2
O	2	16	22	V	2	0	0	Ru	2	0	0					
F	1	0	0	Cr	2	0	0	Ag	1	0	0					

Error Margin (ppm): 20

DBE Range: -2.0 - 1000.0

Electron Ions: both

HC Ratio: unlimited

Apply N Rule: no

Use MSn Info: yes

Max Isotopes: all

Isotope RI (%): 1.00

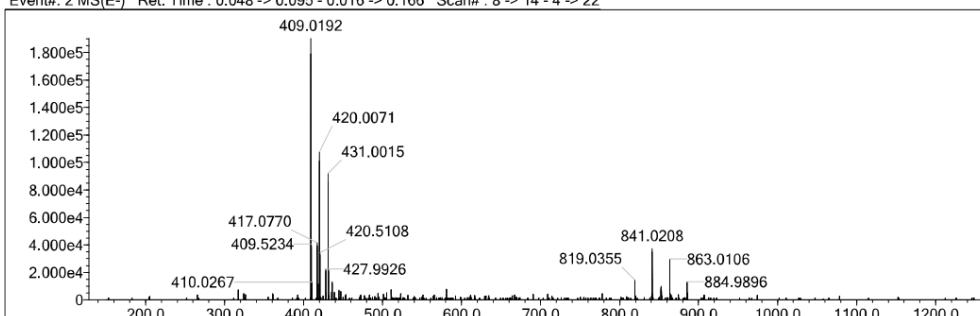
Isotope Res: 6000

MSn Iso RI (%): 75.00

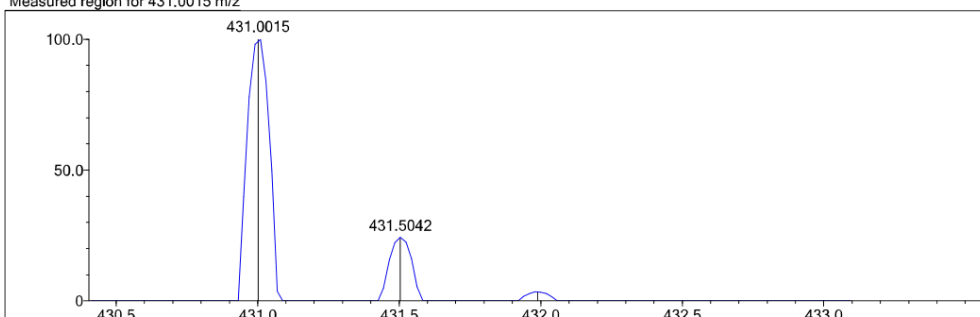
MSn Logic Mode: AND

Max Results: 500

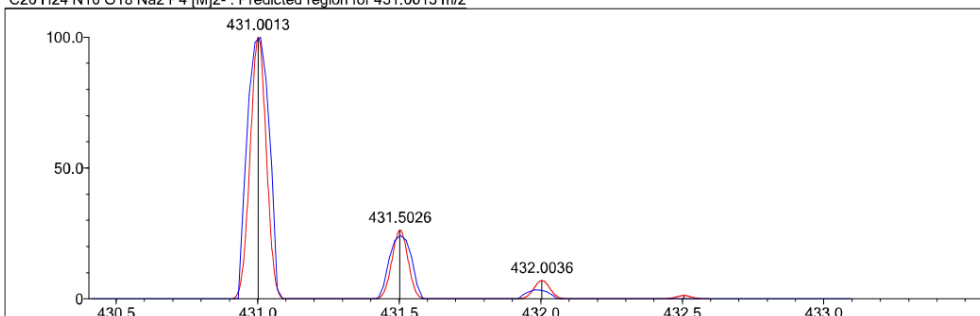
Event#: 2 MS(E-) Ret. Time : 0.048 -&gt; 0.095 - 0.016 -&gt; 0.166 Scan#: 8 -&gt; 14 - 4 -&gt; 22



Measured region for 431.0015 m/z



C20 H24 N10 O18 Na2 P4 [M]2- : Predicted region for 431.0013 m/z



Rank	Score	Formula (M)	Ion	Meas. m/z	Pred. m/z	Df. (mDa)	Df. (ppm)	Iso	DBE
8	69.24	C20 H24 N10 O18 Na2 P4	[M]2-	431.0015	431.0013	0.2	0.46	69.24	15.0

Data File: C:\LabSolutions\Data\2 University\Urbaniak, M (BLS)\Ferguson, Freya\191011 Ferguson HRMS Assignment\210428\Ap4dC-210428-0...

Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Use Adduct
H	1	20	30	Na	1	0	4	Fe	2	0	0	I	3	0	0	Cl
2H	1	0	0	Si	4	0	1	Co	2	0	0	Re	2	0	0	H
B	3	0	0	P	3	4	4	Ni	2	0	0	Ir	3	0	0	Br
C	4	19	20	S	2	0	0	Cu	2	0	0	Pt	2	0	0	I
N	3	6	12	Cl	1	0	0	Br	1	0	0					CHO2
O	2	16	22	V	2	0	0	Ru	2	0	0					
F	1	0	0	Cr	2	0	0	Ag	1	0	0					

Error Margin (ppm): 20

HC Ratio: unlimited

Max Isotopes: all

MSn Iso RI (%): 75.00

DBE Range: -2.0 - 1000.0

Apply N Rule: no

Isotope RI (%): 1.00

MSn Logic Mode: AND

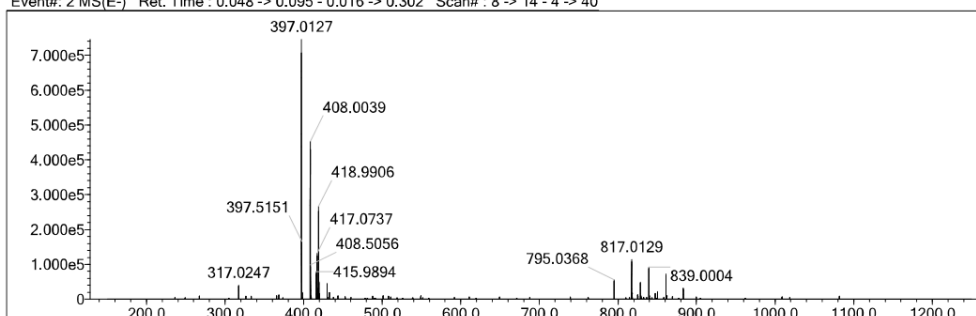
Electron Ions: both

Use MSn Info: yes

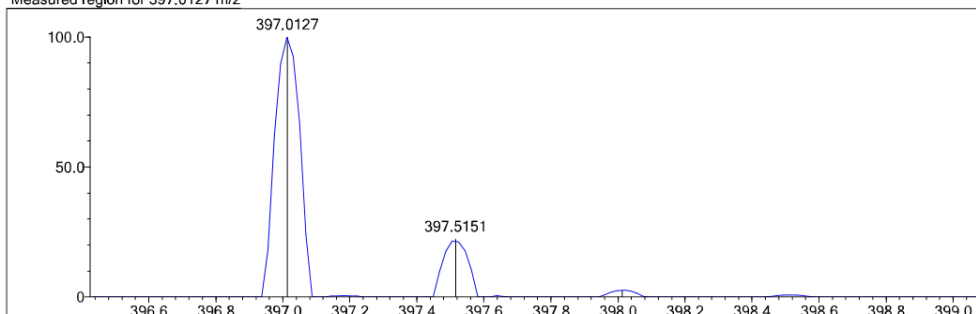
Isotope Res: 6000

Max Results: 500

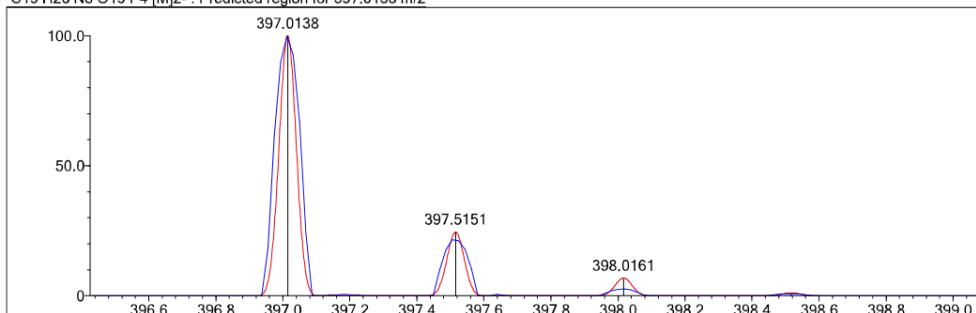
Event#: 2 MS(E-) Ret. Time : 0.048 -&gt; 0.095 - 0.016 -&gt; 0.302 Scan#: 8 -&gt; 14 - 4 -&gt; 40



Measured region for 397.0127 m/z



C19 H26 N8 O19 P4 [M]2- : Predicted region for 397.0138 m/z



Rank	Score	Formula (M)	Ion	Meas. m/z	Pred. m/z	Df. (mDa)	Df. (ppm)	Iso	DBE
3	90.07	C19 H26 N8 O19 P4	[M]2-	397.0127	397.0138	-1.1	-2.77	94.24	13.0

Data File: C:\LabSolutions\Data\2 University\Urbaniak, M (BLS)\Ferguson, Freya\191011 Ferguson HRMS Assignment\210428\Ap4dC-210428-0...

Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Use Adduct
H	1	20	30	Na	1	0	4	Fe	2	0	0	I	3	0	0	H
2H	1	0	0	Si	4	0	1	Co	2	0	0	Re	2	0	0	Na
B	3	0	0	P	3	4	4	Ni	2	0	0	Ir	3	0	0	K
C	4	19	20	S	2	0	0	Cu	2	0	0	Pt	2	0	0	NH4
N	3	8	8	Cl	1	0	0	Br	1	0	0					
O	2	16	22	V	2	0	0	Ru	2	0	0					
F	1	0	0	Cr	2	0	0	Ag	1	0	0					

Error Margin (ppm): 20

DBE Range: -2.0 - 1000.0

Electron Ions: both

HC Ratio: unlimited

Apply N Rule: no

Use MSn Info: yes

Max Isotopes: all

Isotope RI (%): 1.00

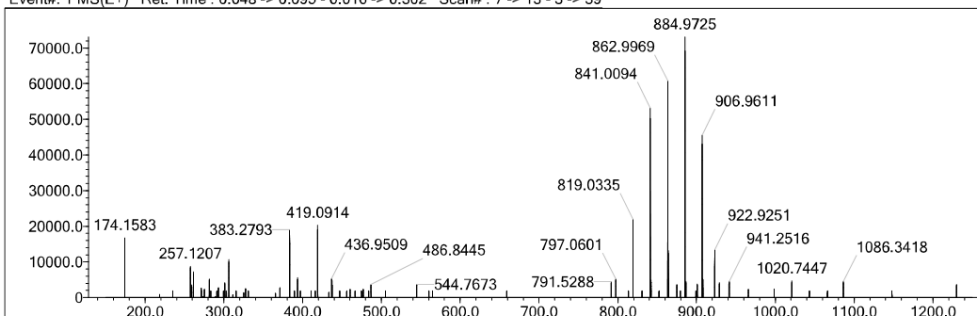
Isotope Res: 6000

MSn Iso RI (%): 75.00

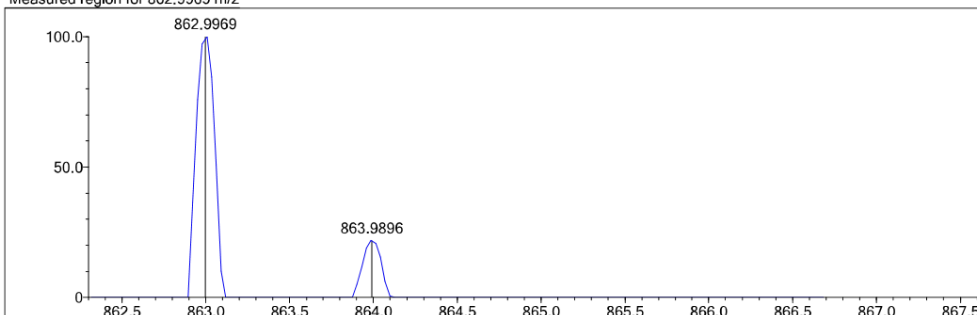
MSn Logic Mode: AND

Max Results: 500

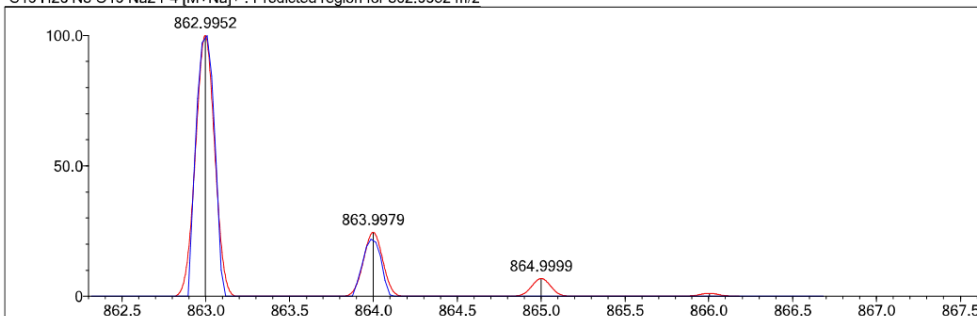
Event#: 1 MS(E+) Ret. Time : 0.048 -&gt; 0.095 - 0.016 -&gt; 0.302 Scan#: 7 -&gt; 13 - 3 -&gt; 39



Measured region for 862.9969 m/z



C19 H26 N8 O19 Na2 P4 [M+Na]+ : Predicted region for 862.9952 m/z



Rank	Score	Formula (M)	Ion	Meas. m/z	Pred. m/z	Df. (mDa)	Df. (ppm)	Iso	DBE
3	0.00	C19 H26 N8 O19 Na2 P4	[M+Na]+	862.9969	862.9952	1.7	1.97	0.00	12.0

Data File: C:\LabSolutions\Data\2 University\Urbaniak, M (BLS)\Ferguson, Freya\191011 Ferguson HRMS Assignment\210428\Ap4dG-210428-0...

Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Use Adduct
H	1	20	30	Na	1	0	4	Fe	2	0	0	I	3	0	0	Cl
2H	1	0	0	Si	4	0	1	Co	2	0	0	Re	2	0	0	H
B	3	0	0	P	3	4	4	Ni	2	0	0	Ir	3	0	0	Br
C	4	19	20	S	2	0	0	Cu	2	0	0	Pt	2	0	0	I
N	3	6	12	Cl	1	0	0	Br	1	0	0					CHO2
O	2	16	22	V	2	0	0	Ru	2	0	0					
F	1	0	0	Cr	2	0	0	Ag	1	0	0					

Error Margin (ppm): 20

DBE Range: -2.0 - 1000.0

Electron Ions: both

HC Ratio: unlimited

Apply N Rule: no

Use MSn Info: yes

Max Isotopes: all

Isotope RI (%): 1.00

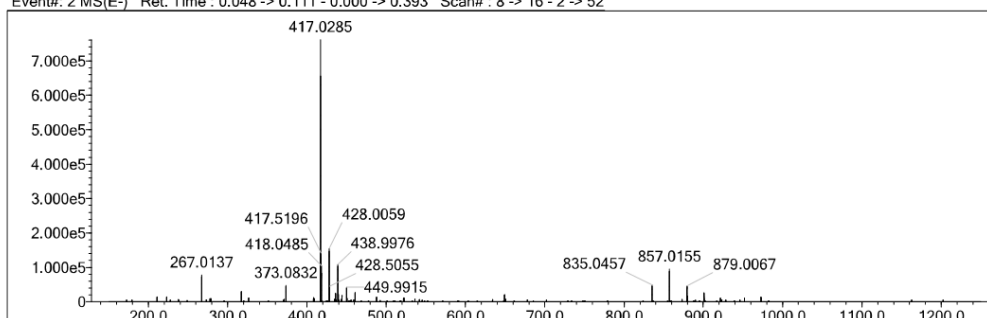
Isotope Res: 6000

MSn Iso RI (%): 75.00

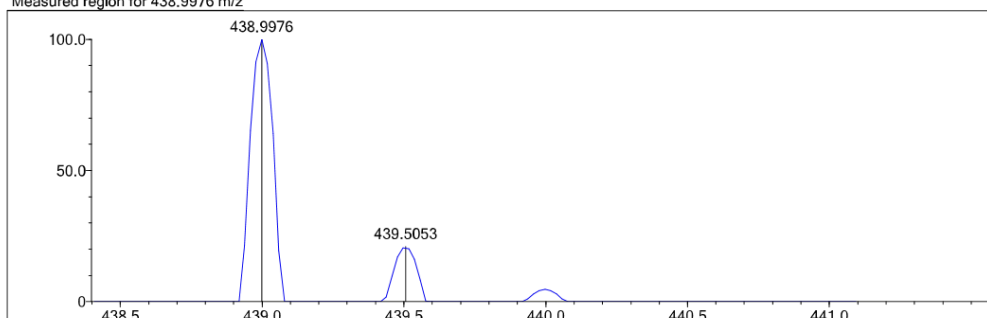
MSn Logic Mode: AND

Max Results: 500

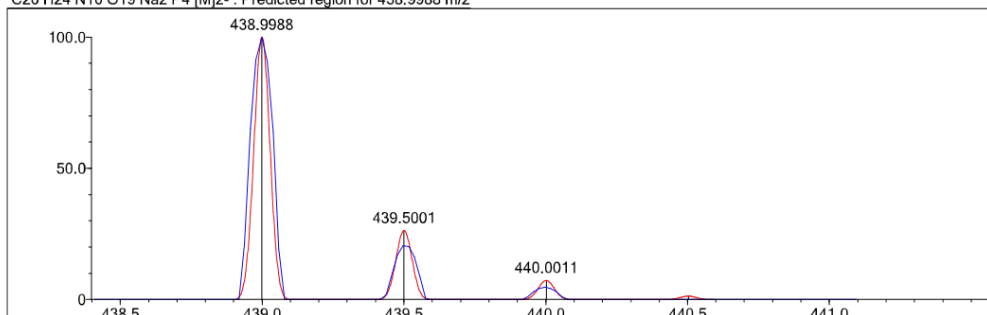
Event#: 2 MS(E-) Ret. Time : 0.048 -&gt; 0.111 - 0.000 -&gt; 0.393 Scan#: 8 -&gt; 16 - 2 -&gt; 52



Measured region for 438.9976 m/z



C20 H24 N10 O19 Na2 P4 [M]2- : Predicted region for 438.9988 m/z



Rank	Score	Formula (M)	Ion	Meas. m/z	Pred. m/z	Df. (mDa)	Df. (ppm)	Iso	DBE
8	77.48	C20 H24 N10 O19 Na2 P4	[M]2-	438.9976	438.9988	-1.2	-2.73	80.98	15.0

Data File: C:\LabSolutions\Data\2 University\Urbaniak, M (BLS)\Ferguson, Freya\191011 Ferguson HRMS Assignment\210428\Ap4dG-210428-0...

Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Use Adduct
H	1	20	30	Na	1	0	4	Fe	2	0	0	I	3	0	0	Cl
2H	1	0	0	Si	4	0	1	Co	2	0	0	Re	2	0	0	H
B	3	0	0	P	3	4	4	Ni	2	0	0	Ir	3	0	0	Br
C	4	19	20	S	2	0	0	Cu	2	0	0	Pt	2	0	0	I
N	3	6	12	Cl	1	0	0	Br	1	0	0					CHO2
O	2	16	22	V	2	0	0	Ru	2	0	0					
F	1	0	0	Cr	2	0	0	Ag	1	0	0					

Error Margin (ppm): 20

DBE Range: -2.0 - 1000.0

Electron Ions: both

HC Ratio: unlimited

Apply N Rule: no

Use MSn Info: yes

Max Isotopes: all

Isotope RI (%): 1.00

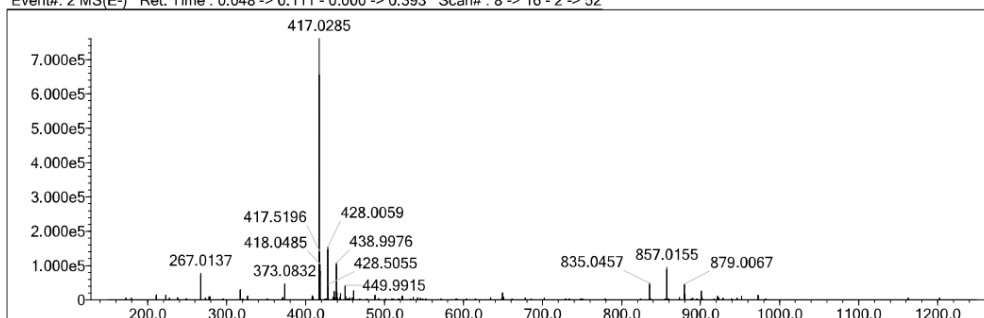
Isotope Res: 6000

MSn Iso RI (%): 75.00

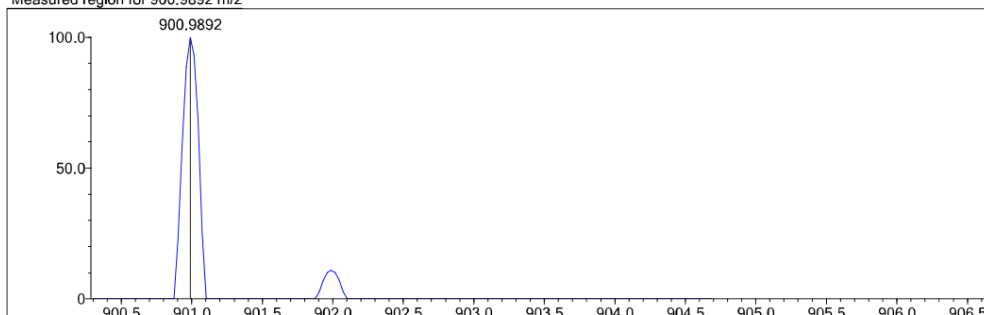
MSn Logic Mode: AND

Max Results: 500

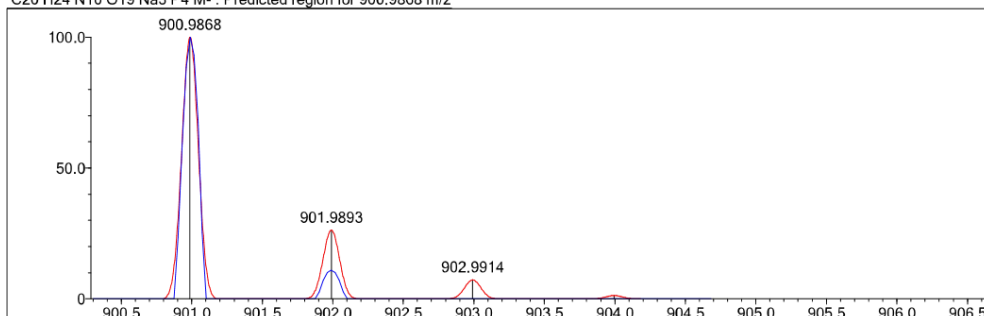
Event#: 2 MS(E-) Ret. Time : 0.048 -&gt; 0.111 - 0.000 -&gt; 0.393 Scan#: 8 -&gt; 16 - 2 -&gt; 52



Measured region for 900.9892 m/z



C20 H24 N10 O19 Na3 P4 M- : Predicted region for 900.9868 m/z



Rank	Score	Formula (M)	Ion	Meas. m/z	Pred. m/z	Df. (mDa)	Df. (ppm)	Iso	DBE
14	0.00	C20 H24 N10 O19 Na3 P4	M-	900.9892	900.9868	2.4	2.66	0.00	14.5

Data File: C:\LabSolutions\Data\2 University\Urbaniak, M (BLS)\Ferguson, Freya\191011 Ferguson HRMS Assignment\210428\Ap4dT-210428-0...

Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Use Adduct
H	1	20	30	Na	1	0	4	Fe	2	0	0	I	3	0	0	Cl
2H	1	0	0	Si	4	0	1	Co	2	0	0	Re	2	0	0	H
B	3	0	0	P	3	4	4	Ni	2	0	0	Ir	3	0	0	Br
C	4	19	20	S	2	0	0	Cu	2	0	0	Pt	2	0	0	I
N	3	7	7	Cl	1	0	0	Br	1	0	0					CHO2
O	2	16	22	V	2	0	0	Ru	2	0	0					
F	1	0	0	Cr	2	0	0	Ag	1	0	0					

Error Margin (ppm): 20

HC Ratio: unlimited

Max Isotopes: all

MSn Iso RI (%): 75.00

DBE Range: -2.0 - 1000.0

Apply N Rule: no

Isotope RI (%): 1.00

MSn Logic Mode: AND

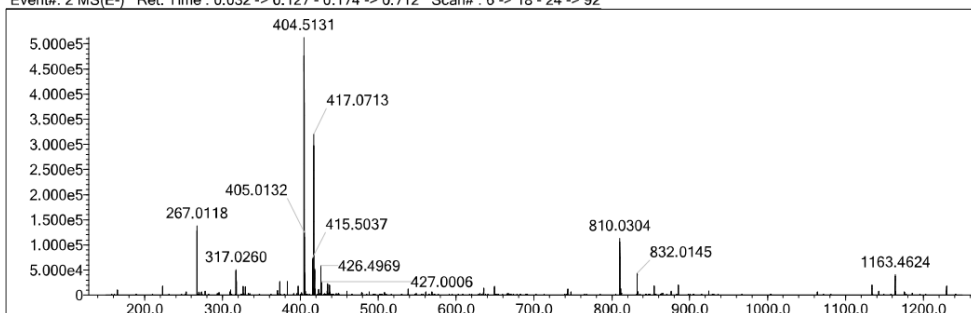
Electron Ions: both

Use MSn Info: yes

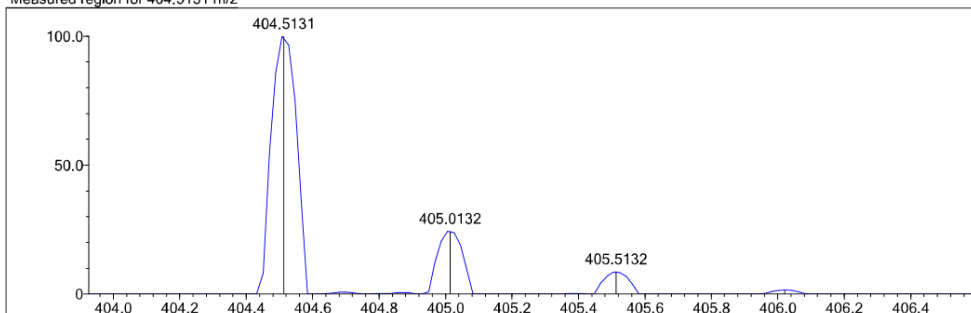
Isotope Res: 6000

Max Results: 500

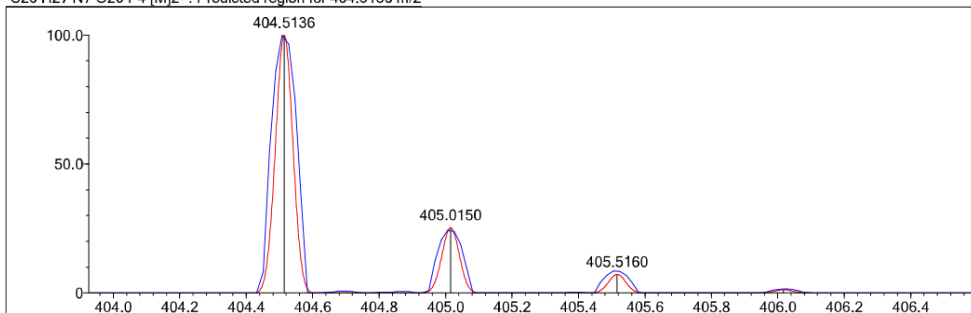
Event#: 2 MS(E-) Ret. Time : 0.032 -&gt; 0.127 - 0.174 -&gt; 0.712 Scan#: 6 -&gt; 18 - 24 -&gt; 92



Measured region for 404.5131 m/z



C20 H27 N7 O20 P4 [M]2- : Predicted region for 404.5136 m/z



Rank	Score	Formula (M)	Ion	Meas. m/z	Pred. m/z	Df. (mDa)	Df. (ppm)	Iso	DBE
2	99.40	C20 H27 N7 O20 P4	[M]2-	404.5131	404.5136	-0.5	-1.24	100.00	13.0

Data File: C:\LabSolutions\Data\2 University\Urbaniak, M (BLS)\Ferguson, Freya\191011 Ferguson HRMS Assignment\210428\Ap4dT-210428-0...

Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Elmt	Val.	Min	Max	Use Adduct
H	1	20	30	Na	1	0	4	Fe	2	0	0	I	3	0	0	H
2H	1	0	0	Si	4	0	1	Co	2	0	0	Re	2	0	0	Na
B	3	0	0	P	3	4	4	Ni	2	0	0	Ir	3	0	0	K
C	4	19	20	S	2	0	0	Cu	2	0	0	Pt	2	0	0	NH4
N	3	7	7	Cl	1	0	0	Br	1	0	0					
O	2	16	22	V	2	0	0	Ru	2	0	0					
F	1	0	0	Cr	2	0	0	Ag	1	0	0					

Error Margin (ppm): 20

HC Ratio: unlimited

Max Isotopes: all

MSn Iso RI (%): 75.00

DBE Range: -2.0 - 1000.0

Apply N Rule: no

Isotope RI (%): 1.00

MSn Logic Mode: AND

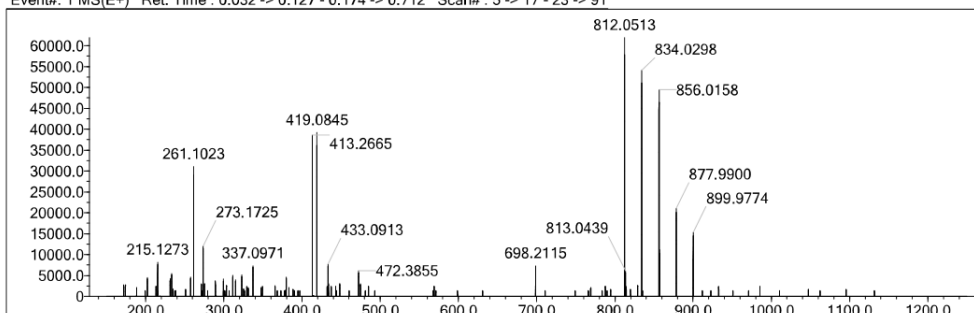
Electron Ions: both

Use MSn Info: yes

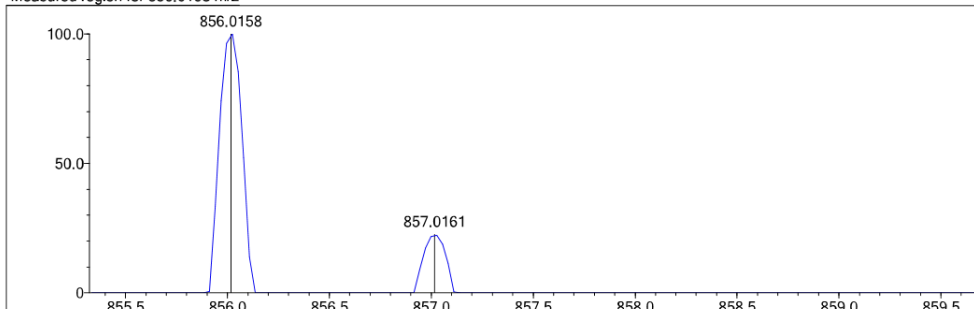
Isotope Res: 6000

Max Results: 500

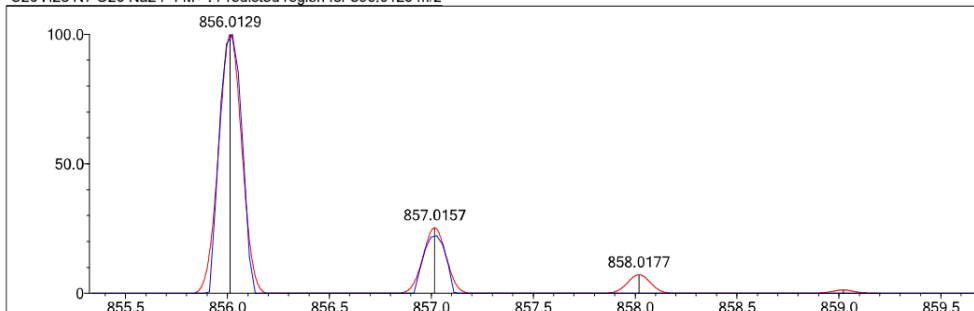
Event#: 1 MS(E+) Ret. Time : 0.032 -&gt; 0.127 - 0.174 -&gt; 0.712 Scan#: 5 -&gt; 17 - 23 -&gt; 91



Measured region for 856.0158 m/z



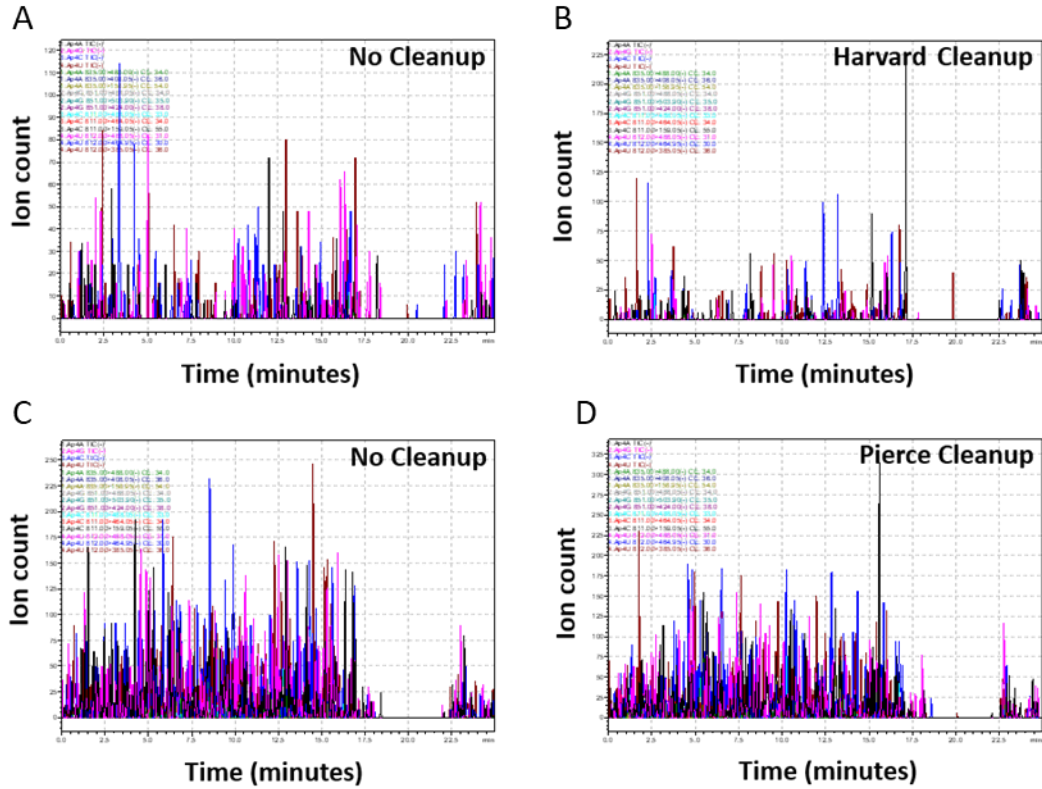
C20 H28 N7 O20 Na2 P4 M+ : Predicted region for 856.0129 m/z



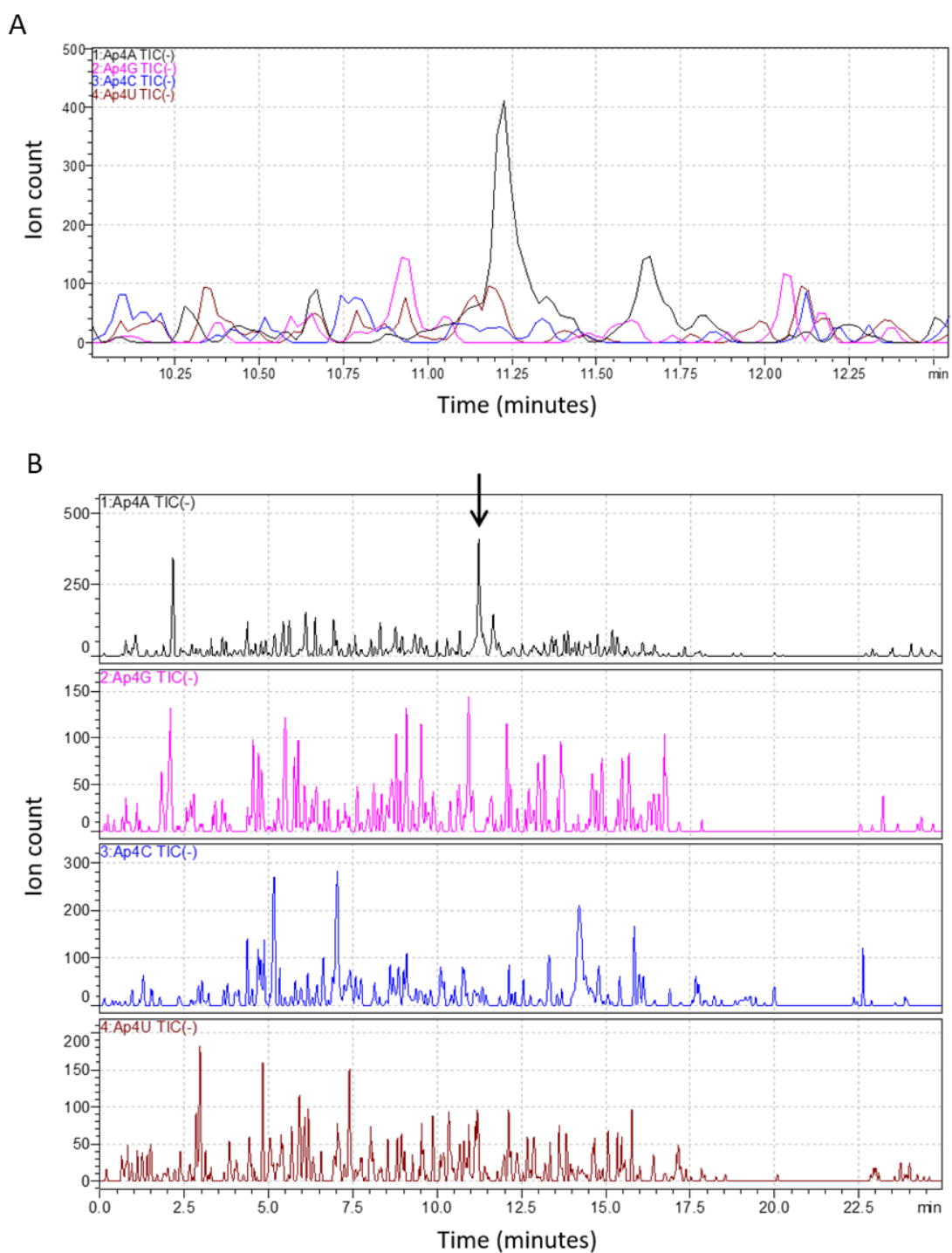
Rank	Score	Formula (M)	Ion	Meas. m/z	Pred. m/z	Df. (mDa)	Df. (ppm)	Iso	DBE
3	0.00	C20 H28 N7 O20 Na2 P4	M+	856.0158	856.0129	2.9	3.39	0.00	11.5

**Supplementary Figure 3.1: HRMS reports for chemically synthesised Ap<sub>4</sub>Ns, where N= A, C, G, U, dA, dC, dG or dT.** Ap<sub>4</sub>Ns were synthesised chemically and purified by HPLC. The ammonium bicarbonate buffer was removed by lyophilisation and the samples were analysed by HRMS.

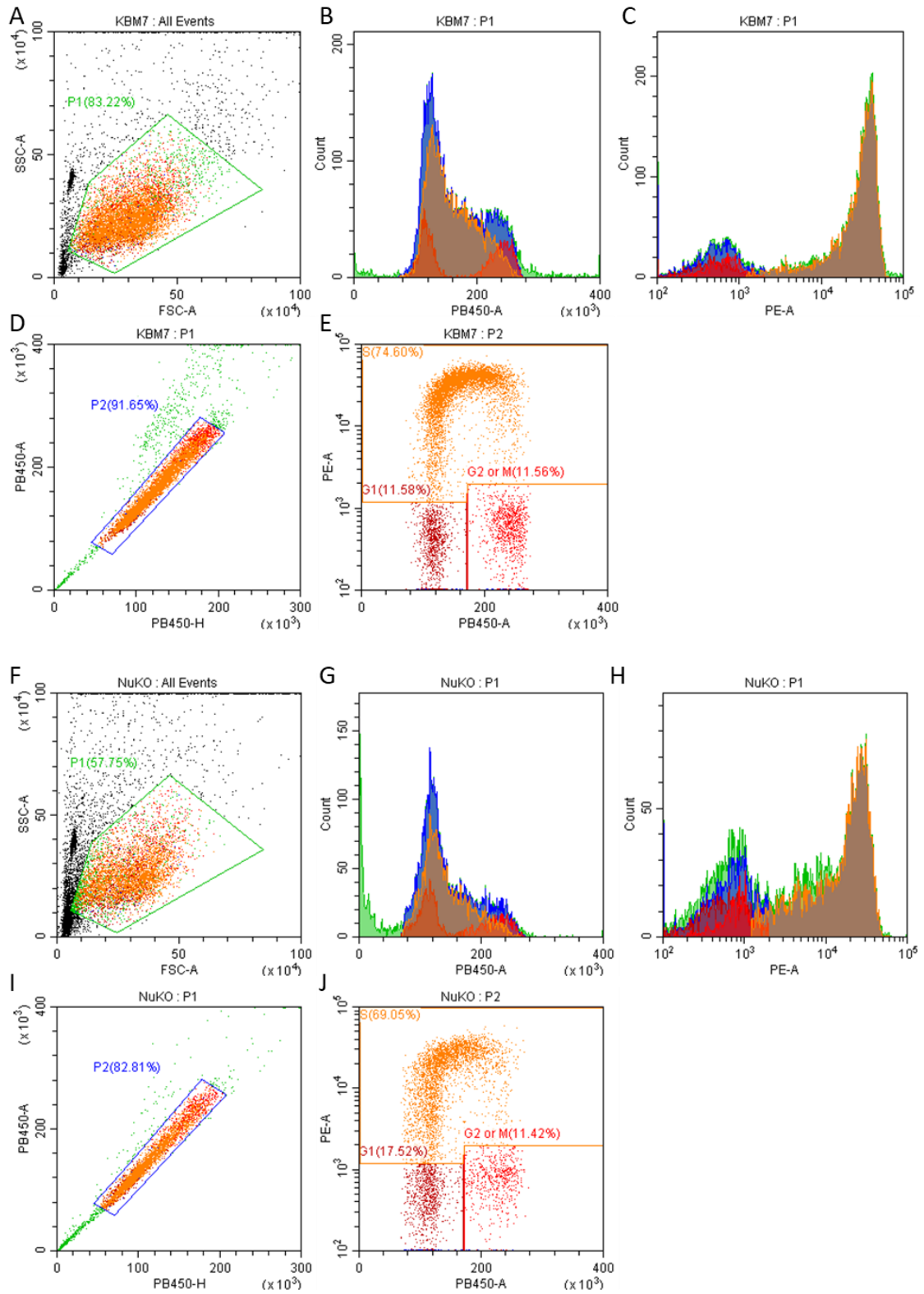


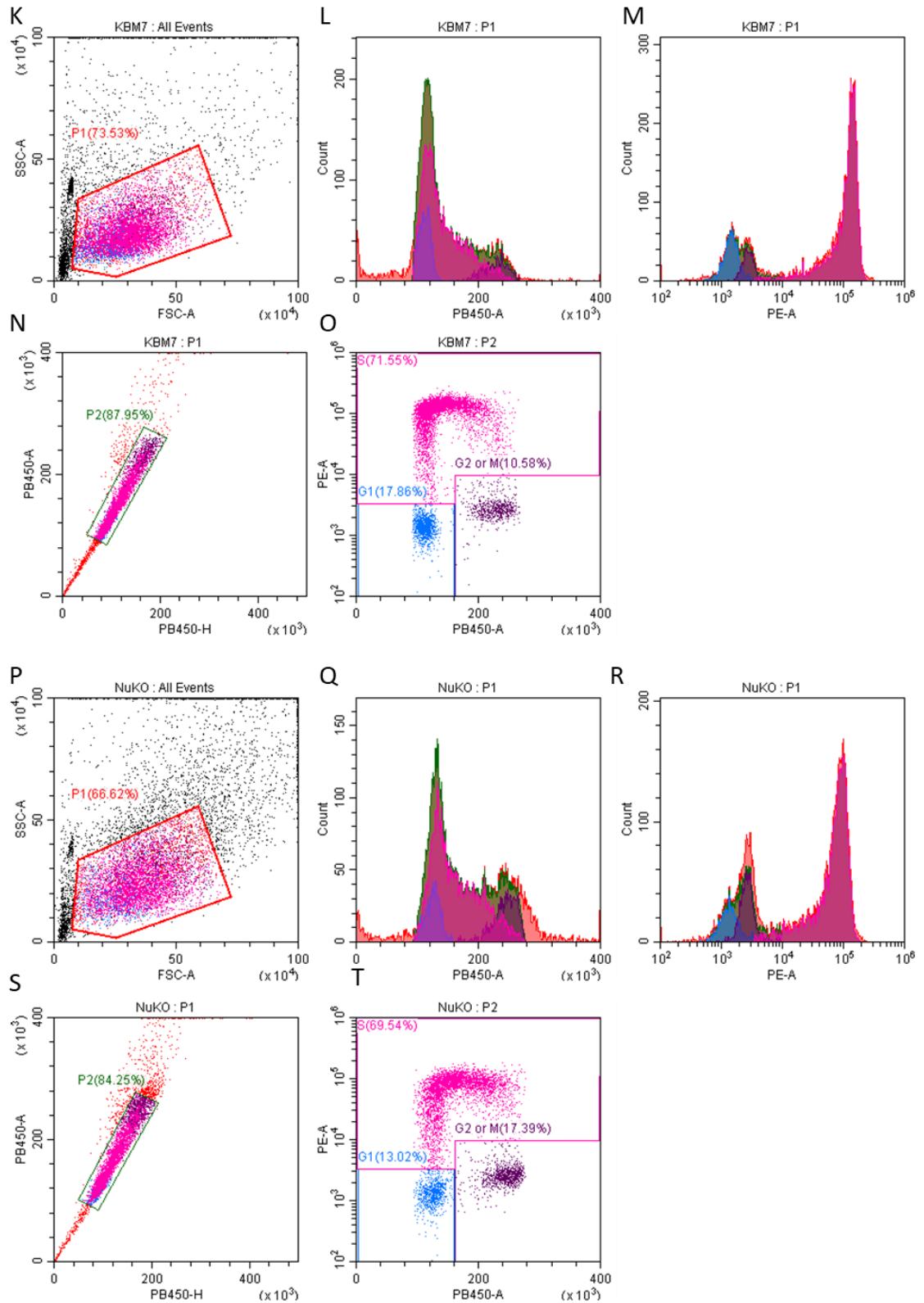


**Supplementary Figure 4.1: Effect of Harvard or Pierce Cleanup on identification of Ap<sub>4</sub>A in KBM7 cell extract.** Extract was harvested from KBM7 cells and subjected to either no cleanup (A, C), Harvard column cleanup (B), or Pierce column cleanup (D).

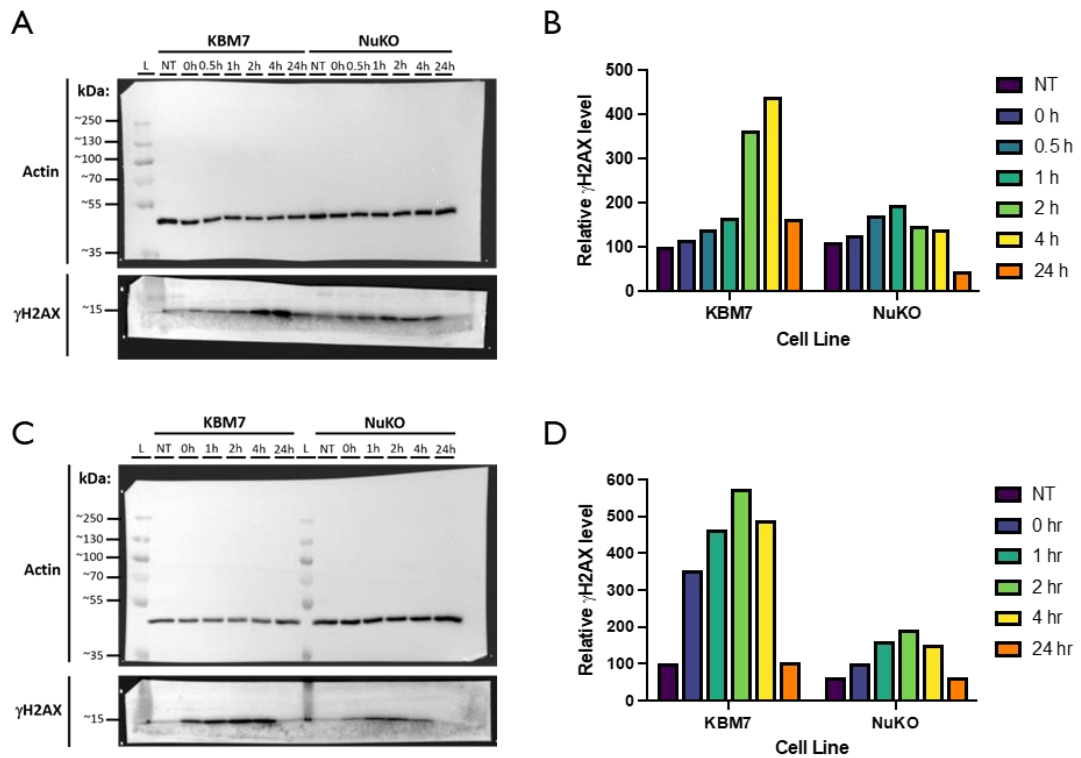


**Supplementary Figure 4.2: Investigating for the presence of Ap<sub>4</sub>Ns in KBM7 cells.** KBM7 cell extract was harvested then analysed by triple quadrupole mass spectrometry. The resulting peaks are show overlaid with one another (A) and in their own scale (B).

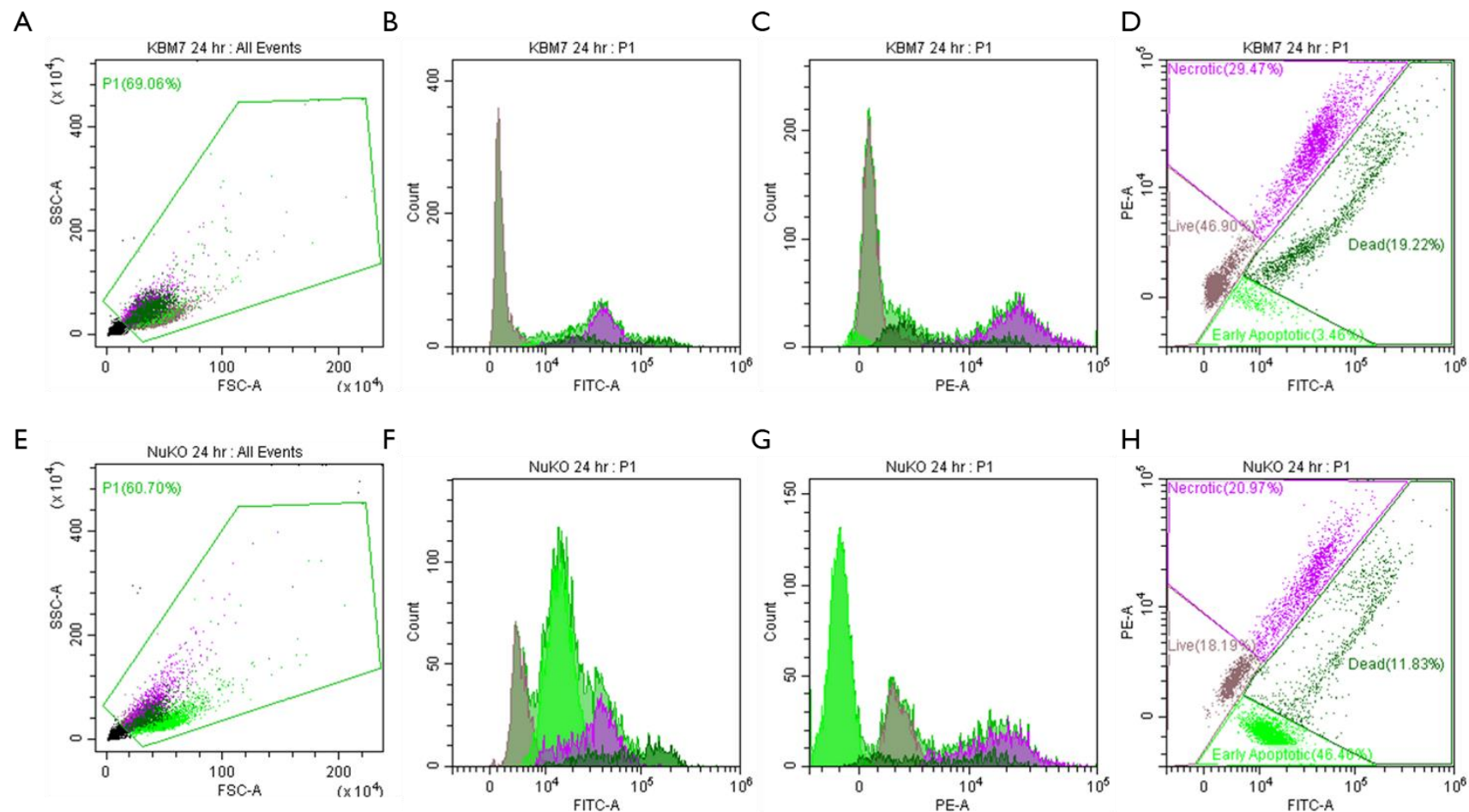




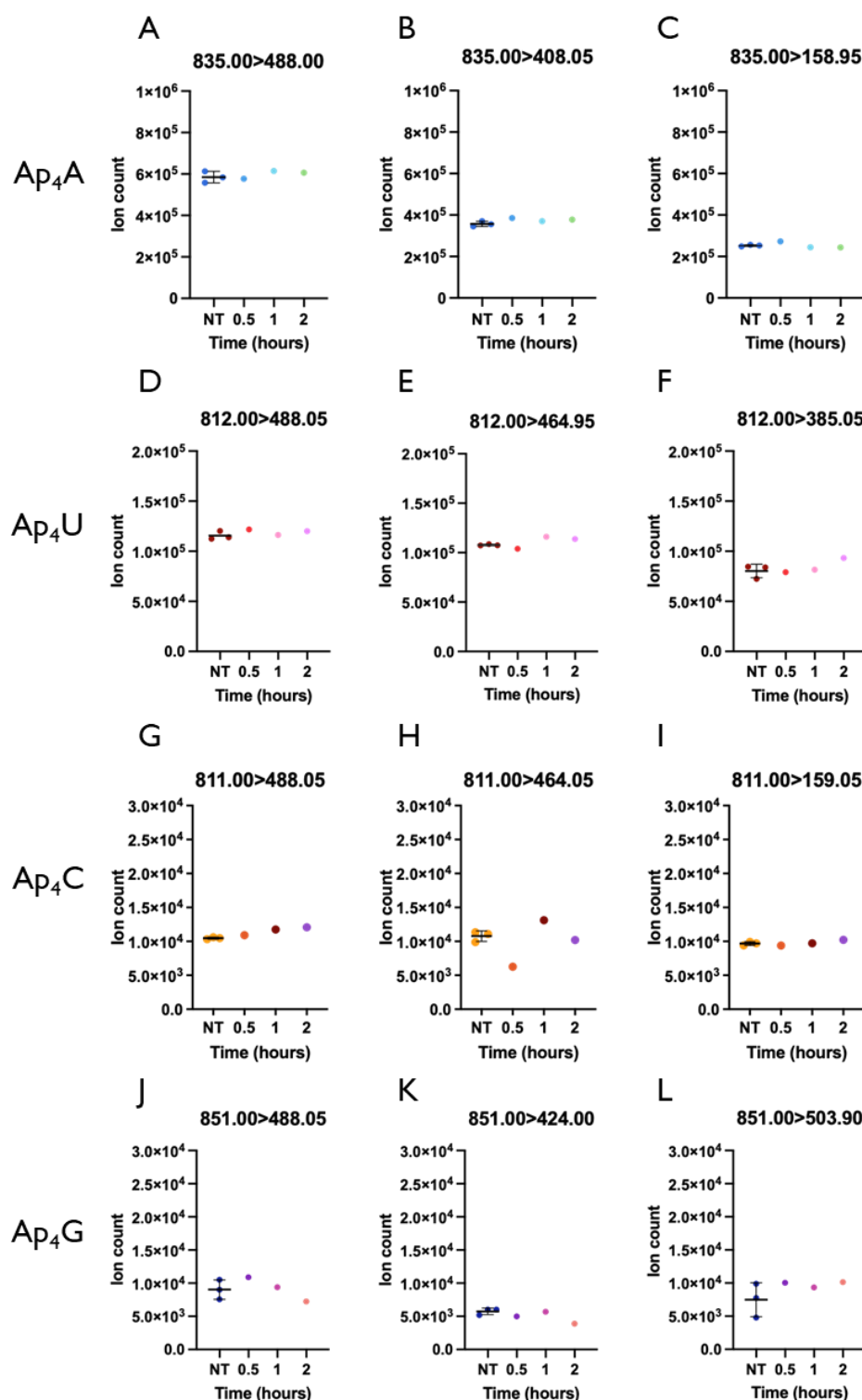
**Supplementary Figure 5.1: Gating for the analysis of the cell cycle kinetics in KBM7 and NuKO cell lines.** KBM7 and NuKO cells were labelled with EdU and Hoescht to determine the proportion of cells in each stage of the cell cycle. Two biological repeats were performed (A = Repeat 1, B = Repeat 2). EdU labelling = PE, Hoescht labelling = PB450.



**Supplementary Figure 5.2:  $\gamma$ H2AX response to HU-induced stress in KBM7 and NuKO cells.** KBM7 and NuKO cells were stressed with HU for 2 hours then allowed to recover for 0, 1, 2, 4 or 24 hours.  $\gamma$ H2AX levels were measured relative to actin then scaled so that the  $\gamma$ H2AX level in untreated KBM7 cells was equal to 100. A) Repeat 1 B) Repeat 2.

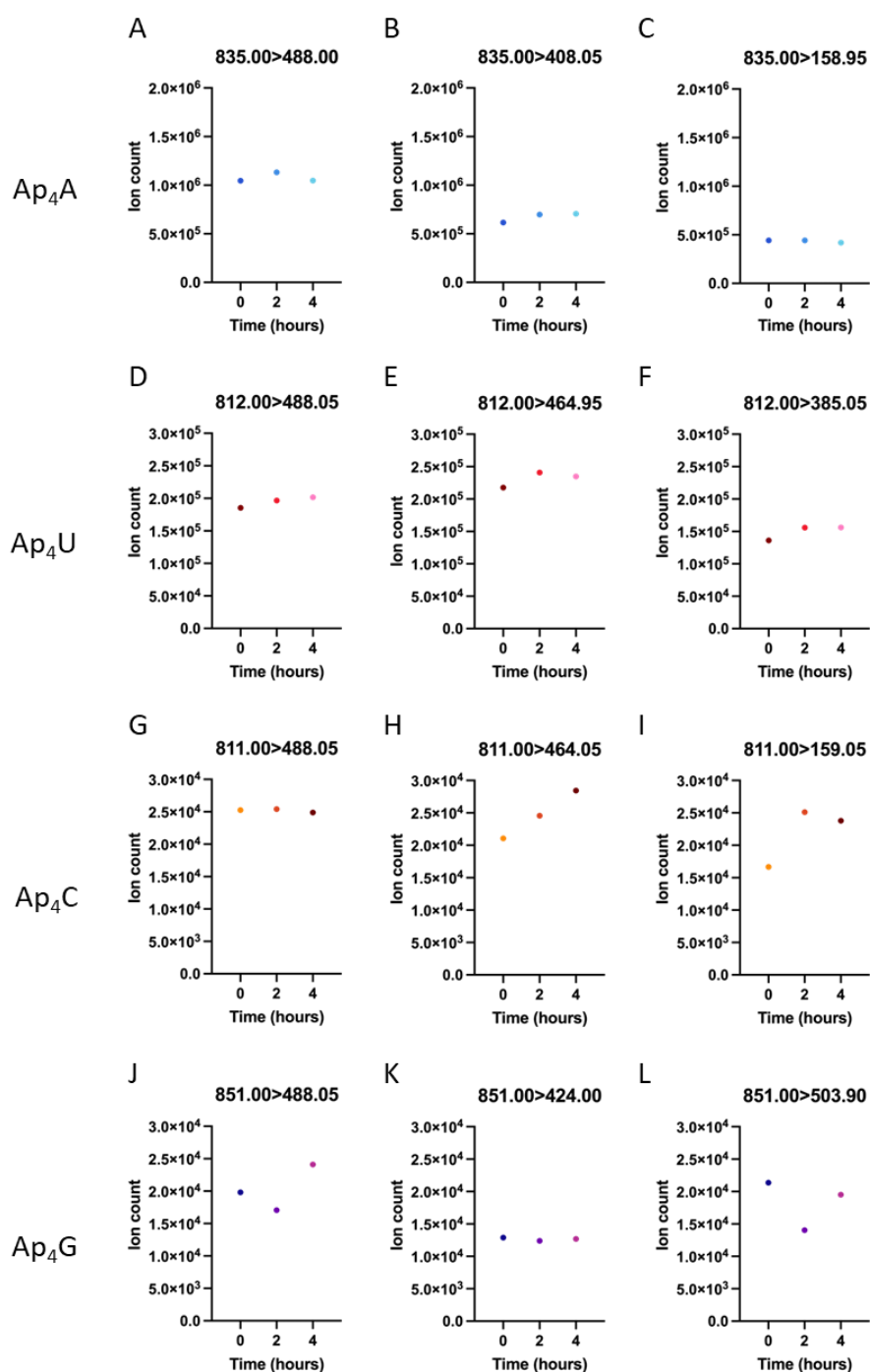


**Supplementary Figure 5.3: Gating for analysis of the apoptotic response in HU-treated KBM7 and NuKO cells.** HU-treated KBM7 or NuKO cells were allowed to recover for 24 hours then labelled with YO-PRO1 and PI and analysed by flow cytometry. Gating shown for A) KBM7 and B) NuKO cells. YO-PRO1 = FITC and PI = PE.



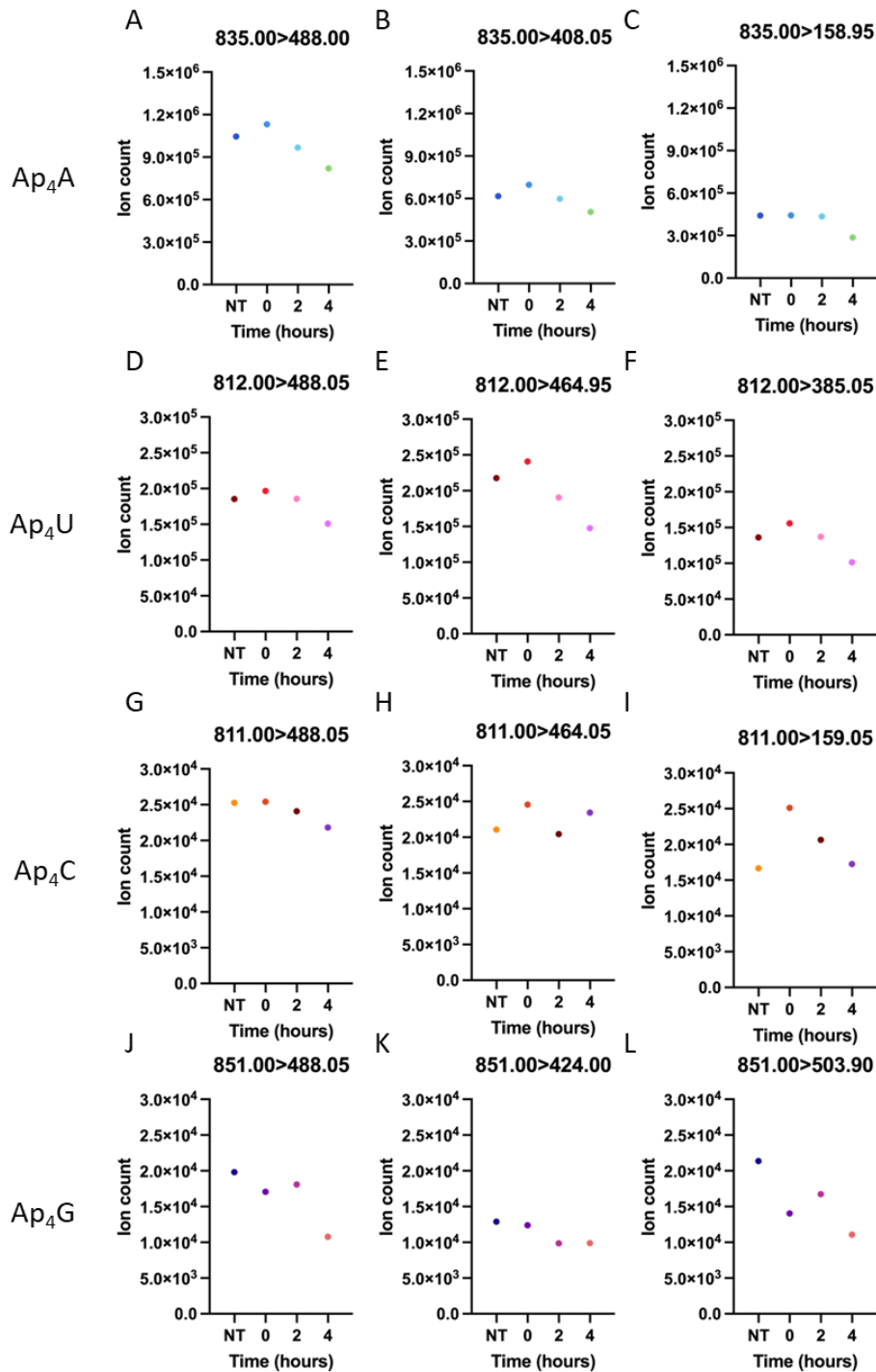
**Supplementary Figure 5.4: Raw ion count values showing the effect of HU treatment on Ap<sub>4</sub>N transition level in NuKO cells.** NuKO cells were treated with 1 mM HU for 0.5, 1 or 2 hours. Metabolite-containing extracts were harvested and subjected to LC-MS analysis. The ion counts for all three MRM transitions for Ap<sub>4</sub>A (A–C), Ap<sub>4</sub>U (D–F), Ap<sub>4</sub>C (G–I) and Ap<sub>4</sub>G (J–L) were analysed. For the untreated cells, individual data points as well as the mean values and standard deviation are shown. This data set corresponds to Figure 5.6.



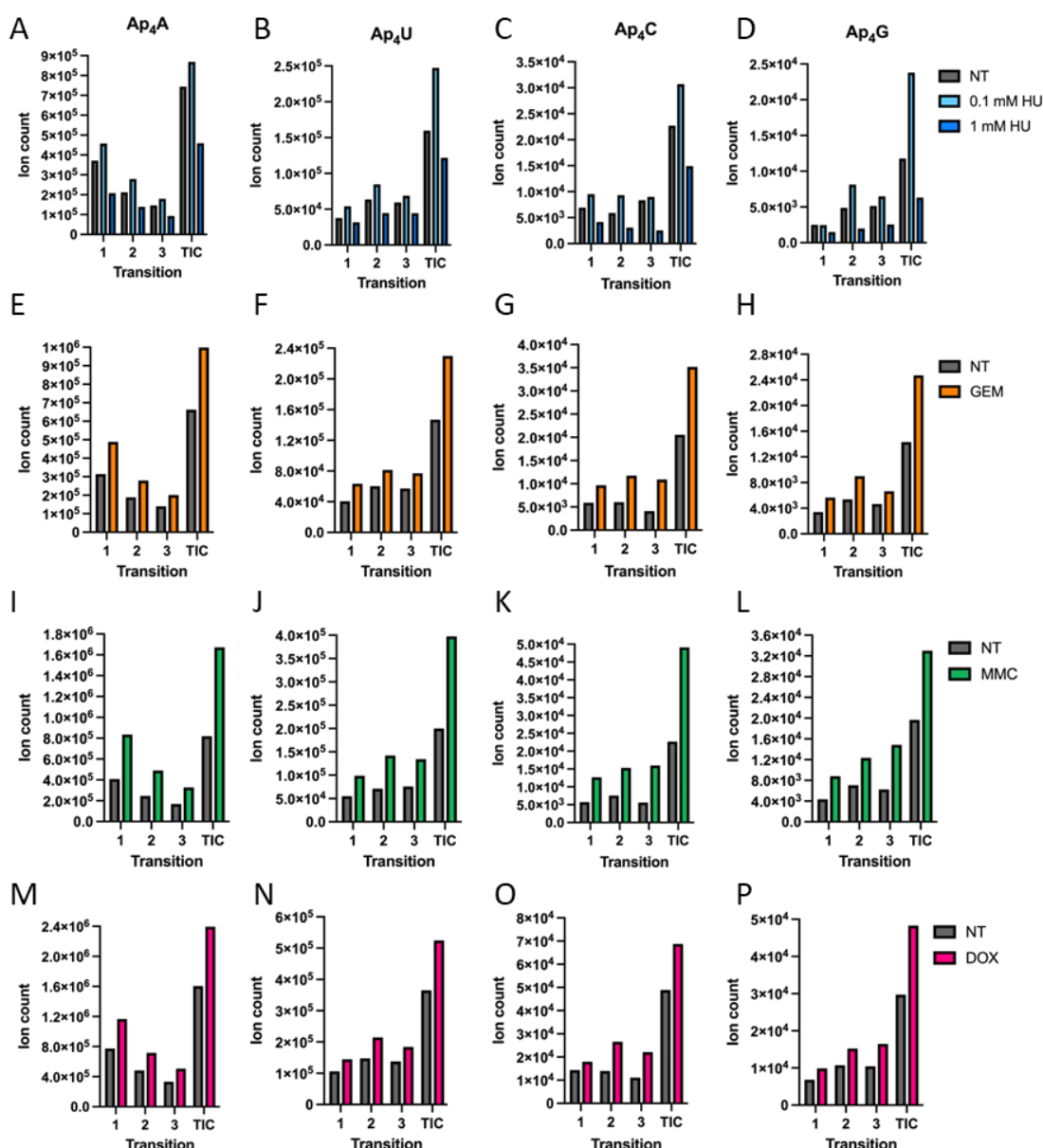


**Supplementary Figure 5.5: Raw ion count values showing the effect of increased length of HU-treatment on Ap<sub>4</sub>N concentration.** NuKO cells were treated with HU for increasing lengths of time then their extract was immediately harvested and analysed by LC-MS. Ion count abundance for all three MRM transitions for Ap<sub>4</sub>A (A–C), Ap<sub>4</sub>U (D–F), Ap<sub>4</sub>C (G–I) and Ap<sub>4</sub>U (J–L) were recorded. This data set corresponds to Figure 5.7.

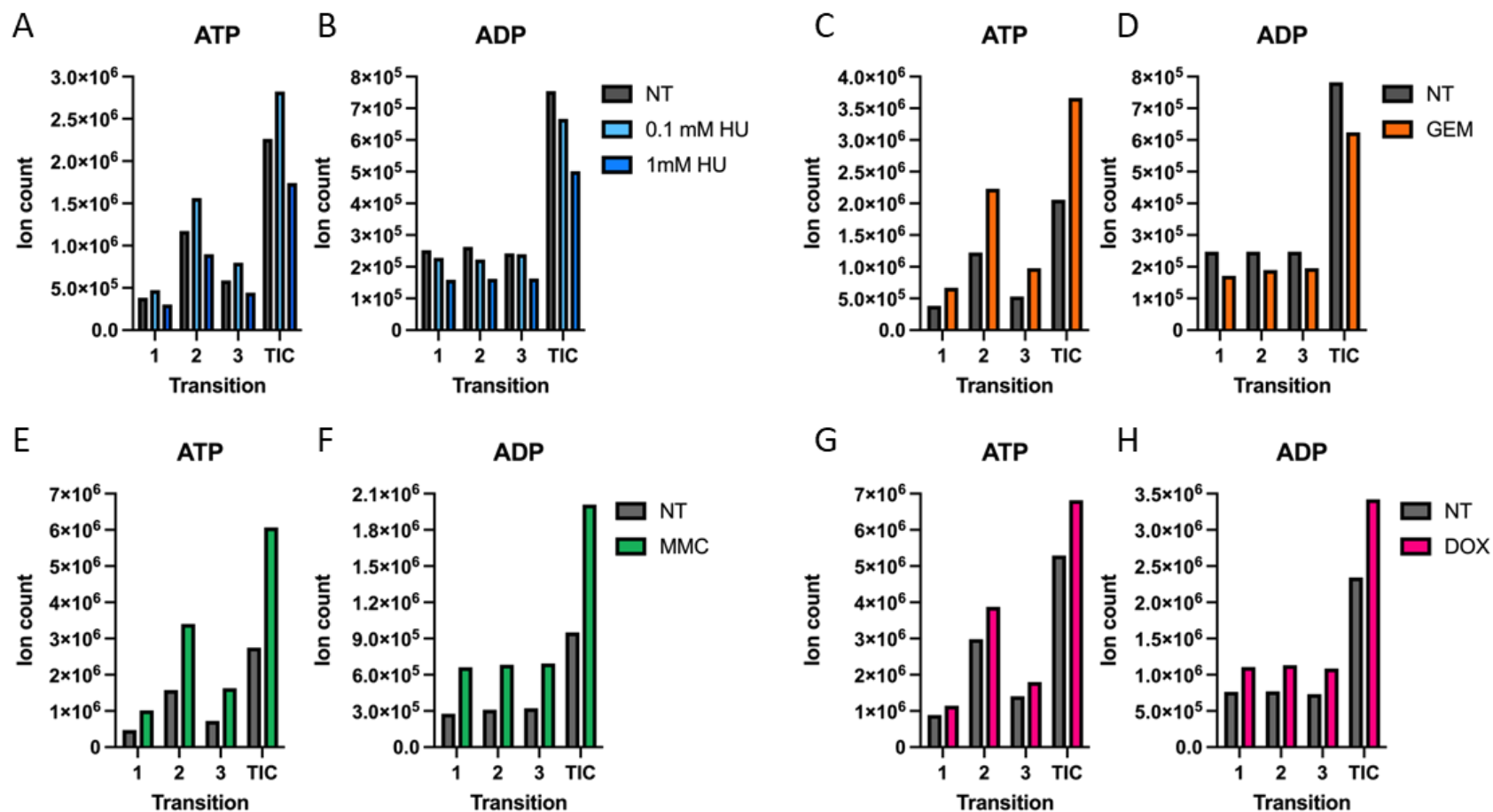




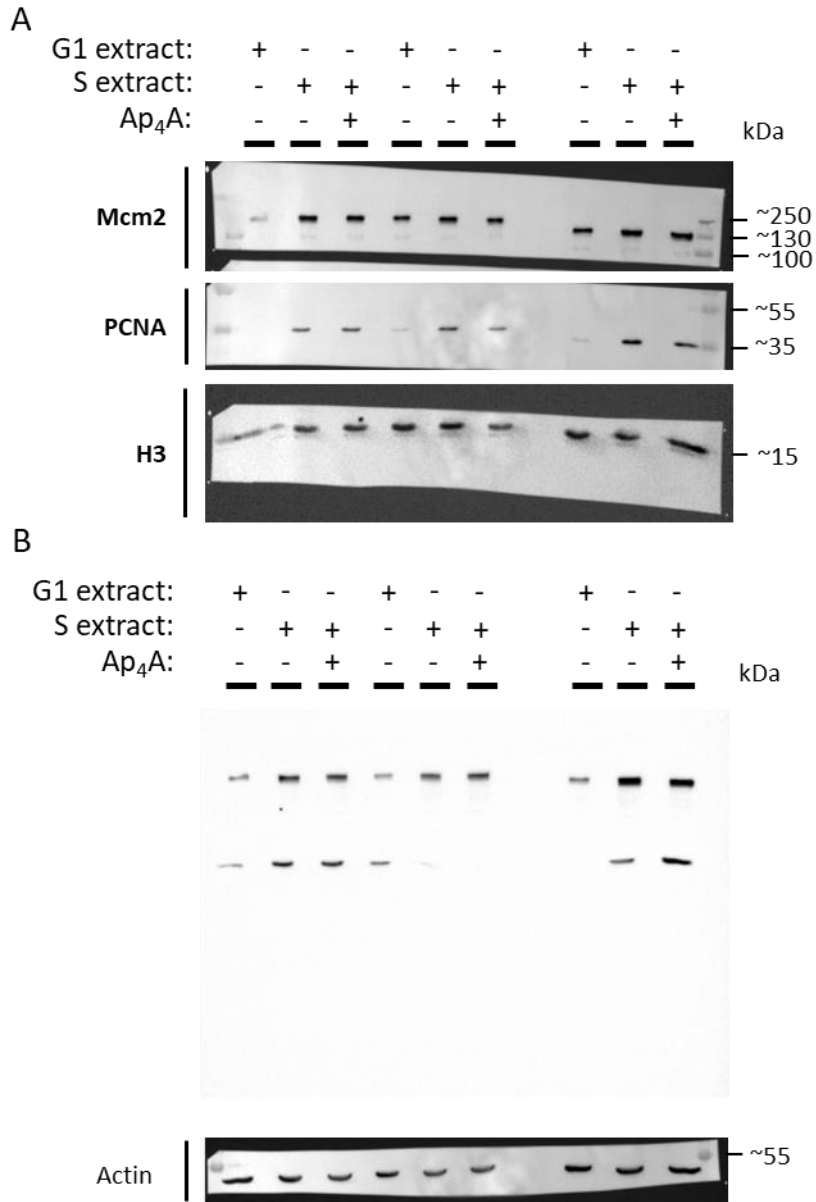
**Supplementary Figure 5.6: Raw ion count values showing the effect of recovery time on Ap<sub>4</sub>N levels after HU-treatment.** NuKO cells were treated for 2 hours then allowed to recover for 0, 2 or 4 hours. The extract was then harvested and analysed by LC-MS and Ap<sub>4</sub>N levels were compared with those in the extract from untreated cells. Changes in the levels of all three MRM transitions for Ap<sub>4</sub>A (A–C), Ap<sub>4</sub>U (D–F), Ap<sub>4</sub>C (G–I) and Ap<sub>4</sub>G (J–L) were analysed. Please note, this data set was produced alongside the data set in Figure 5.7. The NT and 0-hour time-point here correspond to the 0-hour and 2-hour treatments with HU in Figure 5.7. This data set corresponds to Figure 5.8.



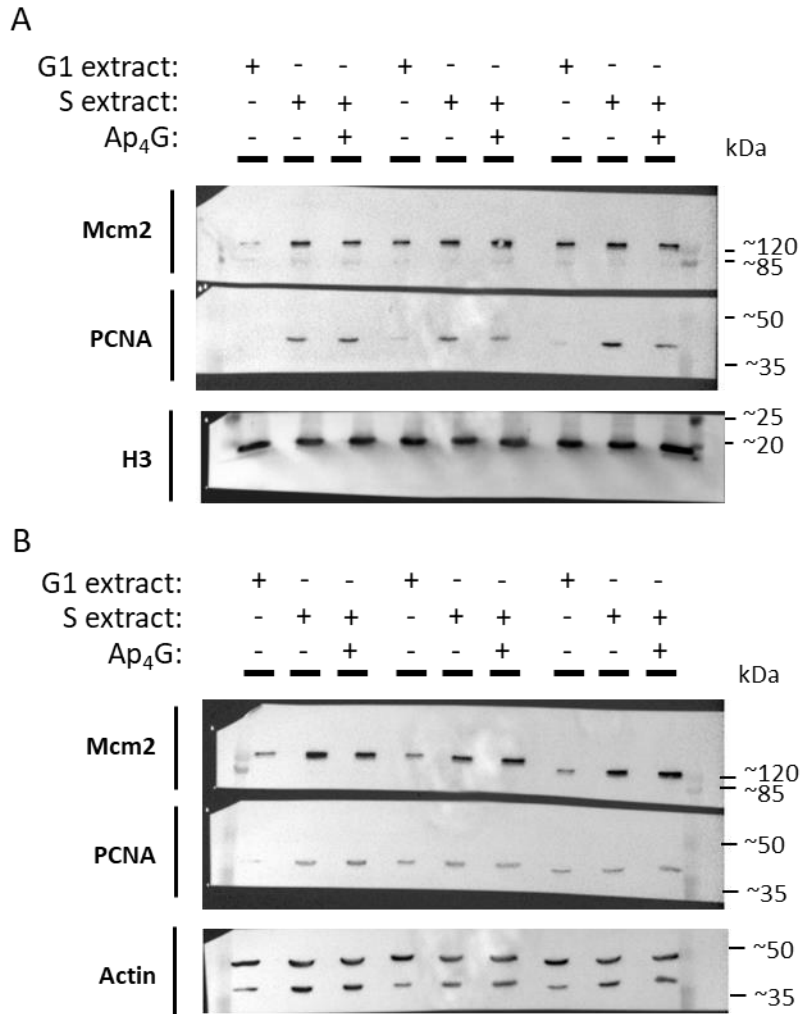
**Supplementary Figure 5.7: Raw ion count values showing the effect of 18-hour treatment with damaging agents on  $Ap_4N$  levels.** Cells were treated with either Hydroxyurea (HU; A–D), Gemcitabine (GEM; E–H), Mitomycin C (MMC; I–L) or Doxorubicin (DOX; M–P) for 18 hours before extracts were harvested and analysed by LC-MS. Transitions 1, 2 and 3 are as follows:  $Ap_4A$ : 1=835.00>488.00, 2=835.00>408.05, 3=835.00>158.95;  $Ap_4U$ : 1=812.00>385.05, 2=812.00>464.95, 3=812.00>488.05;  $Ap_4C$ : 1=811.00>159.05, 2=811.00>464.05, 3=811.00>488.05;  $Ap_4G$ : 1=851.00>424.00, 2=851.00>488.05, 3=851.00>503.90. This data set corresponds to Figure 5.9.



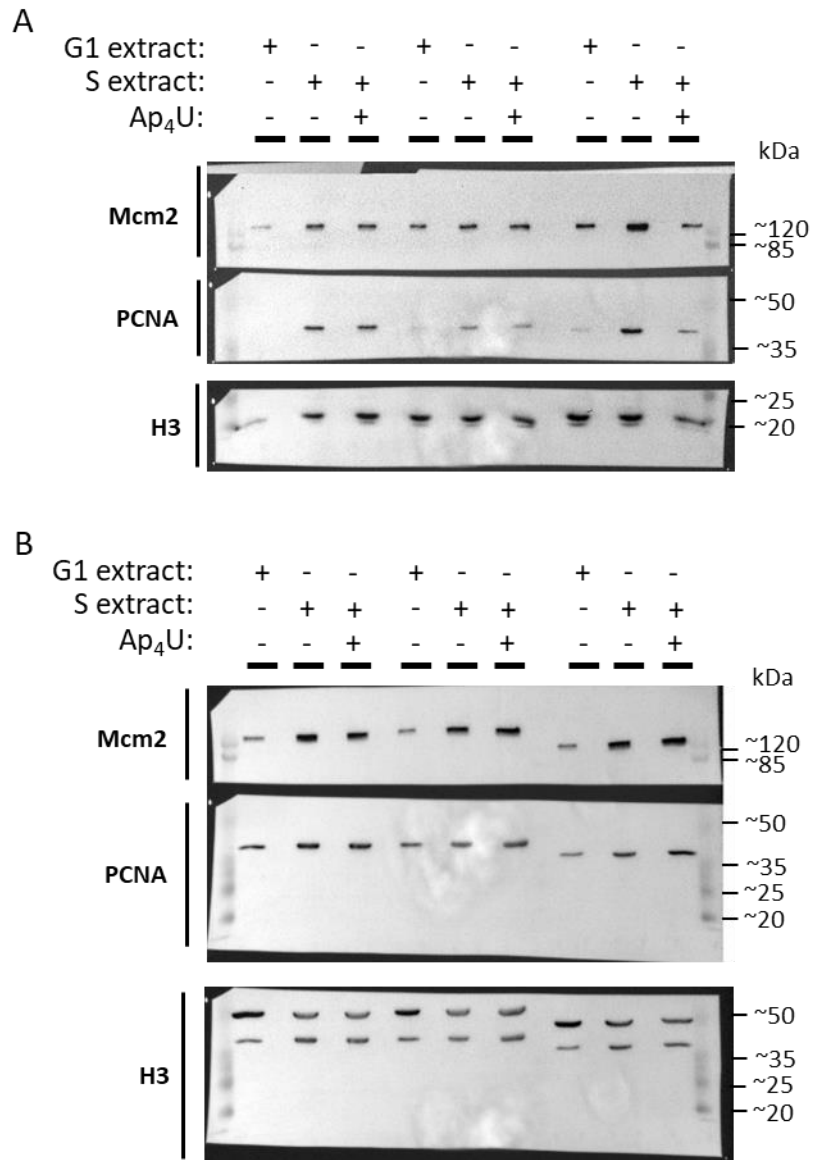
**Supplementary Figure 5.8: Effect of damaging agents on internal ADP and ATP standards.** Cells were treated with either Hydroxyurea (HU; A, B), Gemcitabine (GEM; C, D), Mitomycin C (MMC; E, F) or Doxorubicin (DOX; G, H) for 18 hours before extracts were harvested and analysed by LC-MS. Transitions 1, 2 and 3 are as follows: ATP: 1=506.20>79.00, 2=506.20>158.95, 3=506.20>408.05; ADP: 1=426.20>79.00, 2=426.20>134.05, 3=426.20>158.95. This data set corresponds to Figure 5.11.



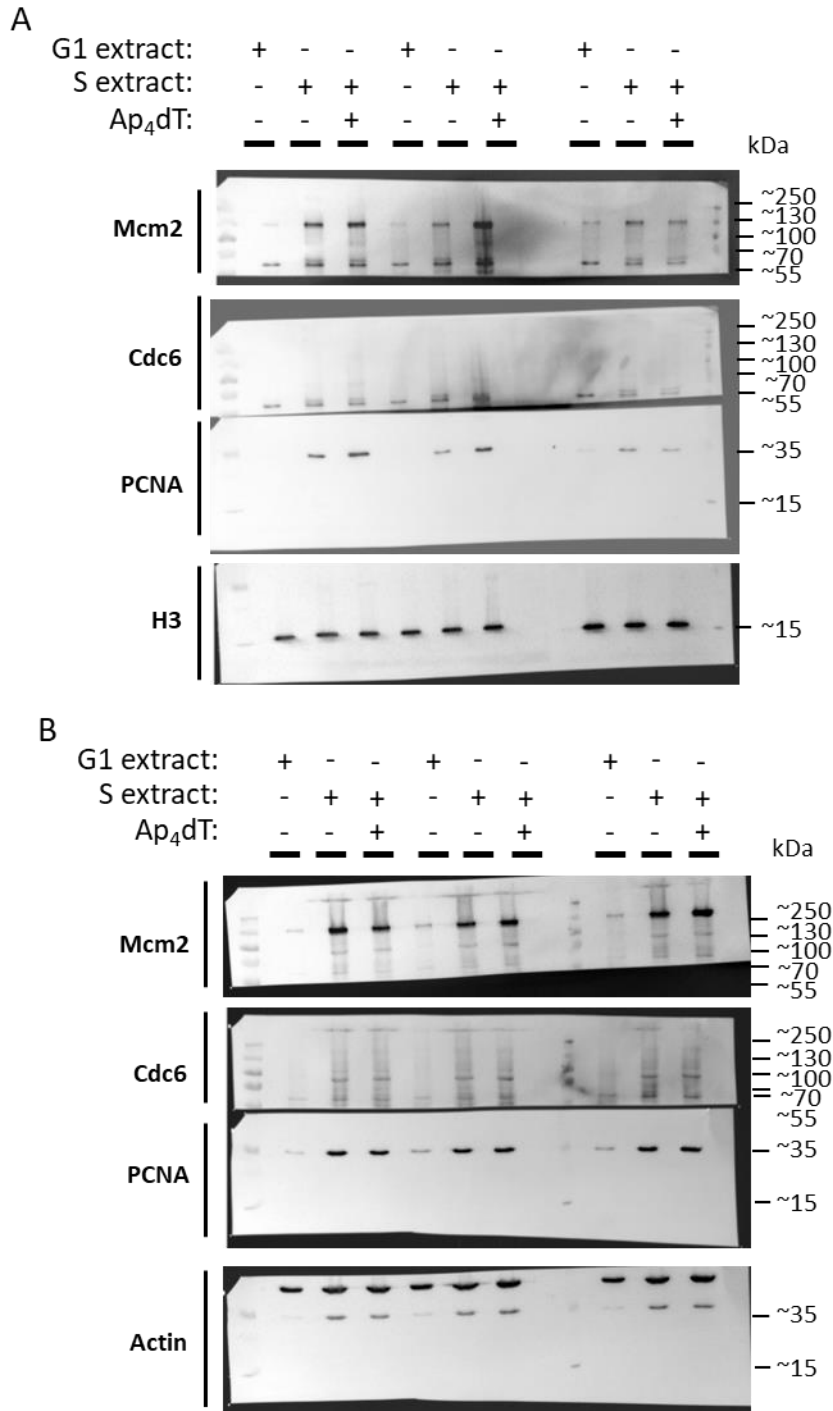
**Supplementary Figure 6.1: Full Western blotting data showing the effect of Ap<sub>4</sub>A on localisation of DNA replication proteins. A:** A comparison of level of Mcm2 and PCNA localising to chromatin after treatment with G1, S or S+Ap<sub>4</sub>A extracts, using H3 for standardisation. **B:** The level of Mcm2 and PCNA available in the soluble extract fraction after incubation in G1, S or S+Ap<sub>4</sub>A extracts. Please note that no ladder is available for the Mcm2 and PCNA data. For the PCNA data, the middle set of repeats was excluded due to poor transfer. This data set corresponds to Figure 6.9.



**Supplementary Figure 6.2: Full Western blotting data showing the effect of Ap<sub>4</sub>G on localisation of DNA replication proteins. A:** A comparison of level of Mcm2 and PCNA localising to chromatin after treatment with G1, S or S+Ap<sub>4</sub>G extracts. **B:** The level of Mcm2 and PCNA available in the soluble extract fraction after incubation in G1, S or S+Ap<sub>4</sub>G extracts. This data set corresponds to Figure 6.11.



**Supplementary Figure 6.3: Full Western blotting data showing the effect of Ap<sub>4</sub>U on localisation of DNA replication proteins. A:** A comparison of level of Mcm2 and PCNA localising to chromatin after treatment with G1, S or S+Ap<sub>4</sub>U extracts. **B:** The level of Mcm2 and PCNA available in the soluble extract fraction after incubation in G1, S or S+Ap<sub>4</sub>U extracts. This data set corresponds to Figure 6.12.



**Supplementary Figure 6.4: Full Western blotting data showing the effect of Ap<sub>4</sub>dT on localisation of DNA replication proteins. A:** A comparison of level of Mcm2 and PCNA localising to chromatin after treatment with G1, S or S+Ap<sub>4</sub>dT extracts. **B:** The level of Mcm2 and PCNA available in the soluble extract fraction after incubation in G1, S or S+Ap<sub>4</sub>dT extracts. This data set corresponds to Figure 6.14. The membrane was probed for Cdc6 prior to Mcm2, so the Cdc6 blot is also shown to demonstrate which bands represent Mcm2. Similarly, the membrane was probed for PCNA prior to actin. This data set corresponds to Figure 6.14.

## References

- Abd Wahab, S. & Remus, D. (2020) Antagonistic control of DDK binding to licensed replication origins by Mcm2 and Rad53. *eLife*, 9, e58571. 10.7554/eLife.58571.
- Ablasser, A., Goldeck, M., Cavlar, T., Deimling, T., Witte, G., Röhl, I., Hopfner, K. P., Ludwig, J. & Hornung, V. (2013) cGAS produces a 2'-5'-linked cyclic dinucleotide second messenger that activates STING. *Nature*, 498(7454), 380-4. 10.1038/nature12306.
- Alexandre, T., Rayna, B. & Munier-Lehmann, H. (2015) Two Classes of Bacterial IMPDHs according to Their Quaternary Structures and Catalytic Properties. *PLOS ONE*, 10(2), e0116578. 10.1371/journal.pone.0116578.
- Allen, P. & Davies, D. (2007) Apoptosis detection by flow cytometry. *Flow cytometry: Principles and applications*, 147-163.
- Anand, R., Ranjha, L., Cannavo, E. & Cejka, P. (2016) Phosphorylated CtIP Functions as a Co-factor of the MRE11-RAD50-NBS1 Endonuclease in DNA End Resection. *Molecular Cell*, 64(5), 940-950. <https://doi.org/10.1016/j.molcel.2016.10.017>.
- Annesley, T. M. (2003) Ion Suppression in Mass Spectrometry. *Clinical Chemistry*, 49(7), 1041-1044. 10.1373/49.7.1041.
- Appy, L., Chardet, C., Peyrottes, S. & Roy, B. (2019a) Synthetic Strategies for Dinucleotides Synthesis. *Molecules*, 24(23), 4334.
- Appy, L., Depaix, A., Bantreil, X., Lamaty, F., Peyrottes, S. & Roy, B. (2019b) Straightforward Ball-Milling Access to Dinucleoside 5',5'-Polyphosphates via Phosphorimidazolid Intermediates. *Chemistry*, 25(10), 2477-2481. 10.1002/chem.201805924.
- Atencia, E. A., Madrid, O., Günther Sillero, M. A. & Sillero, A. (1999) T4 RNA ligase catalyzes the synthesis of dinucleoside polyphosphates. *Eur J Biochem*, 261(3), 802-11. 10.1046/j.1432-1327.1999.00338.x.



- Atlante, A., Giannattasio, S., Bobba, A., Gagliardi, S., Petragallo, V., Calissano, P., Marra, E. & Passarella, S. (2005) An increase in the ATP levels occurs in cerebellar granule cells en route to apoptosis in which ATP derives from both oxidative phosphorylation and anaerobic glycolysis. *Biochim Biophys Acta*, 1708(1), 50-62. 10.1016/j.bbabbio.2005.01.009.
- Baker, J. C. & Jacobson, M. K. (1984) Determination of diadenosine 5',5'''',-P<sub>1</sub>,P<sub>4</sub>-tetrphosphate levels in cultured mammalian cells. *Anal Biochem*, 141(2), 451-60. 10.1016/0003-2697(84)90070-8.
- Baker, J. C. & Jacobson, M. K. (1986) Alteration of adenyl dinucleotide metabolism by environmental stress. *Proceedings of the National Academy of Sciences of the United States of America*, 83(8), 2350-2352. 10.1073/pnas.83.8.2350.
- Bakkenist, C. J. & Kastan, M. B. (2003) DNA damage activates ATM through intermolecular autophosphorylation and dimer dissociation. *Nature*, 421(6922), 499-506. 10.1038/nature01368.
- Baltzinger, M., Ebel, J. P. & Remy, P. (1986) Accumulation of dinucleoside polyphosphates in *Saccharomyces cerevisiae* under stress conditions. High levels are associated with cell death. *Biochimie*, 68(10-11), 1231-6. 10.1016/s0300-9084(86)80069-4.
- Bansal, V. K. (2010) HIGH PERFORMANCE LIQUID CHROMATOGRAPHY: A SHORT REVIEW. *Journal of Global Pharma Technology*, 2.
- Baty, J. & Robinson, P. (1977) Single and multiple ion recording techniques for the analysis of diphenylhydantoin and its major metabolite in plasma. *Biomedical mass spectrometry*, 4(1), 36-41.
- Baxi, M. D., McLennan, A. G. & Vishwanatha, J. K. (1994) Characterization of the HeLa cell DNA polymerase alpha-associated Ap4A binding protein by photoaffinity labeling. *Biochemistry*, 33(48), 14601-7. 10.1021/bi00252a028.

Bell, S. P. & Stillman, B. (1992) ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. *Nature*, 357(6374), 128-134. 10.1038/357128a0.

Belrhali, H., Yaremchuk, A., Tukalo, M., Berthet-Colominas, C., Rasmussen, B., Bösecke, P., Diat, O. & Cusack, S. (1995) The structural basis for seryl-adenylate and Ap4A synthesis by seryl-tRNA synthetase. *Structure*, 3(4), 341-352. [https://doi.org/10.1016/S0969-2126\(01\)00166-6](https://doi.org/10.1016/S0969-2126(01)00166-6).

Bird, J. G., Zhang, Y., Tian, Y., Panova, N., Barvík, I., Greene, L., Liu, M., Buckley, B., Krásný, L., Lee, J. K., Kaplan, C. D., Ebright, R. H. & Nickels, B. E. (2016) The mechanism of RNA 5' capping with NAD<sup>+</sup>, NADH and desphospho-CoA. *Nature*, 535(7612), 444-447. 10.1038/nature18622.

Blackford, A. N. & Jackson, S. P. (2017) ATM, ATR, and DNA-PK: The Trinity at the Heart of the DNA Damage Response. *Molecular Cell*, 66(6), 801-817. <https://doi.org/10.1016/j.molcel.2017.05.015>.

Bochner, B. R., Lee, P. C., Wilson, S. W., Cutler, C. W. & Ames, B. N. (1984) AppppA and related adenylylated nucleotides are synthesized as a consequence of oxidation stress. *Cell*, 37(1), 225-232.

Bone, R., Cheng, Y.-C. & Wolfenden, R. (1986a) Inhibition of thymidine kinase by P1-(adenosine-5')-P5-(thymidine-5')-pentaphosphate. *Journal of Biological Chemistry*, 261(13), 5731-5735.

Bone, R., Cheng, Y. C. & Wolfenden, R. (1986b) Inhibition of thymidine kinase by P1-(adenosine-5')-P5-(thymidine-5')-pentaphosphate. *J Biol Chem*, 261(13), 5731-5.

Bradbury, D. A., Simmons, T. D., Slater, K. J. & Crouch, S. P. (2000) Measurement of the ADP:ATP ratio in human leukaemic cell lines can be used as an indicator of cell viability, necrosis and apoptosis. *J Immunol Methods*, 240(1-2), 79-92. 10.1016/S0022-1759(00)00178-2.

Bradner, W. T. (2001) Mitomycin C: a clinical update. *Cancer Treatment Reviews*, 27(1), 35-50. <https://doi.org/10.1053/ctrv.2000.0202>.

Brevet, A., Chen, J., Lévêque, F., Plateau, P. & Blanquet, S. (1989a) In vivo synthesis of adenylylated bis (5'-nucleosidyl) tetraphosphates (Ap4N) by Escherichia coli aminoacyl-tRNA synthetases. *Proceedings of the National Academy of Sciences*, 86(21), 8275-8279.

Brevet, A., Chen, J., Lévêque, F., Plateau, P. & Blanquet, S. (1989b) In vivo synthesis of adenylylated bis(5'-nucleosidyl) tetraphosphates (Ap4N) by Escherichia coli aminoacyl-tRNA synthetases. *Proc Natl Acad Sci U S A*, 86(21), 8275-9. 10.1073/pnas.86.21.8275.

Brevet, A., Plateau, P., Best-Belpomme, M. & Blanquet, S. (1985a) Variation of Ap4A and other dinucleoside polyphosphates in stressed Drosophila cells. *J Biol Chem*, 260(29), 15566-70.

Brevet, A., Plateau, P., Best-Belpomme, M. & Blanquet, S. (1985b) Variation of Ap4A and other dinucleoside polyphosphates in stressed Drosophila cells. *Journal of Biological Chemistry*, 260(29), 15566-15570.

Burma, S., Chen, B. P., Murphy, M., Kurimasa, A. & Chen, D. J. (2001) ATM phosphorylates histone H2AX in response to DNA double-strand breaks. *J Biol Chem*, 276(45), 42462-7. 10.1074/jbc.C100466200.

Can, N. (2018) Development of Validated and Stability-Indicating LC-DAD and LC-MS/MS Methods for Determination of Avanafil in Pharmaceutical Preparations and Identification of a Novel Degradation Product by LCMS-IT-TOF. *Molecules*, 23(7). 10.3390/molecules23071771.

Canelas, A. B., Ten Pierick, A., Ras, C., Seifar, R. M., Van Dam, J. C., Van Gulik, W. M. & Heijnen, J. J. (2009) Quantitative Evaluation of Intracellular Metabolite Extraction Techniques for Yeast Metabolomics. *Analytical Chemistry*, 81(17), 7379-7389. 10.1021/ac900999t.

- Cappell, S. D., Chung, M., Jaimovich, A., Spencer, S. L. & Meyer, T. (2016) Irreversible APC(Cdh1) Inactivation Underlies the Point of No Return for Cell-Cycle Entry. *Cell*, 166(1), 167-80. 10.1016/j.cell.2016.05.077.
- Cappell, S. D., Mark, K. G., Garbett, D., Pack, L. R., Rape, M. & Meyer, T. (2018) EMI1 switches from being a substrate to an inhibitor of APC/CCDH1 to start the cell cycle. *Nature*, 558(7709), 313-317. 10.1038/s41586-018-0199-7.
- Cavaluzzi, M. J. & Borer, P. N. (2004) Revised UV extinction coefficients for nucleoside-5'-monophosphates and unpaired DNA and RNA. *Nucleic Acids Res*, 32(1), e13. 10.1093/nar/gnh015.
- Chang, F., Riera, A., Evrin, C., Sun, J., Li, H., Speck, C. & Weinreich, M. (2015) Cdc6 ATPase activity disengages Cdc6 from the pre-replicative complex to promote DNA replication. *eLife*, 4, e05795. 10.7554/eLife.05795.
- Chen, L., Gilkes, D. M., Pan, Y., Lane, W. S. & Chen, J. (2005) ATM and Chk2-dependent phosphorylation of MDMX contribute to p53 activation after DNA damage. *Embo j*, 24(19), 3411-22. 10.1038/sj.emboj.7600812.
- Chen, L., Nievera, C. J., Lee, A. Y. & Wu, X. (2008) Cell cycle-dependent complex formation of BRCA1.CtIP.MRN is important for DNA double-strand break repair. *J Biol Chem*, 283(12), 7713-20. 10.1074/jbc.M710245200.
- Cheng, J., Li, N., Huo, Y., Dang, S., Tye, B.-K., Gao, N. & Zhai, Y. (2022) Structural Insight into the MCM double hexamer activation by Dbf4-Cdc7 kinase. *Nature communications*, 13(1), 1-13.
- Cheng, N., Payne, R. C., Kemp, W. E., Jr. & Traut, T. W. (1986) Homogeneous uridine kinase from Ehrlich ascites tumor: substrate specificity and inhibition by bisubstrate analogs. *Mol Pharmacol*, 30(2), 159-63.
- Chilkova, O., Stenlund, P., Isoz, I., Stith, C. M., Grabowski, P., Lundström, E. B., Burgers, P. M. & Johansson, E. (2007) The eukaryotic leading and lagging strand DNA polymerases are loaded onto primer-ends via separate mechanisms but have

comparable processivity in the presence of PCNA. *Nucleic Acids Res*, 35(19), 6588-97. 10.1093/nar/gkm741.

Choi, W. & Lee, E. S. (2022) Therapeutic Targeting of DNA Damage Response in Cancer. *International Journal of Molecular Sciences*, 23(3), 1701.

Chung, M., Liu, C., Yang, H. W., Köberlin, M. S., Cappell, S. D. & Meyer, T. (2019) Transient Hysteresis in CDK4/6 Activity Underlies Passage of the Restriction Point in G1. *Mol Cell*, 76(4), 562-573.e4. 10.1016/j.molcel.2019.08.020.

Conti, C., Saccà, B., Herrick, J., Lalou, C., Pommier, Y. & Bensimon, A. (2007) Replication fork velocities at adjacent replication origins are coordinately modified during DNA replication in human cells. *Mol Biol Cell*, 18(8), 3059-67. 10.1091/mbc.e06-08-0689.

Copeland, N. A., Sercombe, H. E., Ainscough, J. F. & Coverley, D. (2010) Ciz1 cooperates with cyclin-A-CDK2 to activate mammalian DNA replication in vitro. *J Cell Sci*, 123(Pt 7), 1108-15. 10.1242/jcs.059345.

Cortez, D., Guntuku, S., Qin, J. & Elledge, S. J. (2001) ATR and ATRIP: partners in checkpoint signaling. *Science*, 294(5547), 1713-6. 10.1126/science.1065521.

Coste, H., Brevet, A., Plateau, P. & Blanquet, S. (1987) Non-adenylylated bis (5'-nucleosidyl) tetraphosphates occur in *Saccharomyces cerevisiae* and in *Escherichia coli* and accumulate upon temperature shift or exposure to cadmium. *Journal of Biological Chemistry*, 262(25), 12096-12103.

Coster, G. & Diffley, J. F. X. (2017) Bidirectional eukaryotic DNA replication is established by quasi-symmetrical helicase loading. *Science*, 357(6348), 314-318. 10.1126/science.aan0063.

Coster, G. & Goldberg, M. (2010) The cellular response to DNA damage: a focus on MDC1 and its interacting proteins. *Nucleus*, 1(2), 166-78. 10.4161/nucl.1.2.11176.

Coverley, D., Laman, H. & Laskey, R. A. (2002) Distinct roles for cyclins E and A during DNA replication complex assembly and activation. *Nature Cell Biology*, 4(7), 523-528. 10.1038/ncb813.

Coverley, D., Marr, J. & Ainscough, J. (2005) Ciz1 promotes mammalian DNA replication. *J Cell Sci*, 118(Pt 1), 101-12. 10.1242/jcs.01599.

Crozier, L., Foy, R., Mouery, B. L., Whitaker, R. H., Corno, A., Spanos, C., Ly, T., Gowen Cook, J. & Saurin, A. T. (2022) CDK4/6 inhibitors induce replication stress to cause long-term cell cycle withdrawal. *Embo j*, 41(6), e108599. 10.15252/embj.2021108599.

Daigaku, Y., Keszthelyi, A., Müller, C. A., Miyabe, I., Brooks, T., Retkute, R., Hubank, M., Nieduszynski, C. A. & Carr, A. M. (2015) A global profile of replicative polymerase usage. *Nat Struct Mol Biol*, 22(3), 192-198. 10.1038/nsmb.2962.

Daley, J. M., Gaines, W. A., Kwon, Y. & Sung, P. (2014) Regulation of DNA pairing in homologous recombination. *Cold Spring Harb Perspect Biol*, 6(11), a017954. 10.1101/cshperspect.a017954.

Daniele, S., Da Pozzo, E., Iofrida, C. & Martini, C. (2016) Human Neural Stem Cell Aging Is Counteracted by  $\alpha$ -Glycerylphosphorylethanolamine. *ACS Chemical Neuroscience*, 7(7), 952-963. 10.1021/acscchemneuro.6b00078.

Dapa, T., Fleurier, S., Bredeche, M.-F. & Matic, I. (2017) The SOS and RpoS regulons contribute to bacterial cell robustness to genotoxic stress by synergistically regulating DNA polymerase Pol II. *Genetics*, 206(3), 1349-1360.

De Klein, A., Muijtjens, M., Van Os, R., Verhoeven, Y., Smit, B., Carr, A. M., Lehmann, A. R. & Hoeijmakers, J. H. (2000) Targeted disruption of the cell-cycle checkpoint gene ATR leads to early embryonic lethality in mice. *Curr Biol*, 10(8), 479-82. 10.1016/s0960-9822(00)00447-4.

De Silva, I. U., Mchugh, P. J., Clingen, P. H. & Hartley, J. A. (2000) Defining the roles of nucleotide excision repair and recombination in the repair of DNA interstrand cross-links in mammalian cells. *Mol Cell Biol*, 20(21), 7980-90. 10.1128/mcb.20.21.7980-7990.2000.

De Sousa Cavalcante, L. & Monteiro, G. (2014) Gemcitabine: Metabolism and molecular mechanisms of action, sensitivity and chemoresistance in pancreatic cancer.

<https://doi.org/10.1016/j.ejphar.2014.07.041>.

Decout, A., Katz, J. D., Venkatraman, S. & Ablasser, A. (2021) The cGAS–STING pathway as a therapeutic target in inflammatory diseases. *Nature Reviews Immunology*, 21(9), 548-569. 10.1038/s41577-021-00524-z.

Deegan, T. D., Yeeles, J. T. & Diffley, J. F. (2016) Phosphopeptide binding by Sld3 links Dbf4-dependent kinase to MCM replicative helicase activation. *Embo j*, 35(9), 961-73. 10.15252/emboj.201593552.

Delacroix, S., Wagner, J. M., Kobayashi, M., Yamamoto, K. & Karnitz, L. M. (2007) The Rad9-Hus1-Rad1 (9-1-1) clamp activates checkpoint signaling via TopBP1. *Genes Dev*, 21(12), 1472-7. 10.1101/gad.1547007.

Depaix, A., Peyrottes, S. & Roy, B. (2017) Water-Medium Synthesis of Nucleoside 5'-Polyphosphates. *Curr Protoc Nucleic Acid Chem*, 69, 13.16.1-13.16.11. 10.1002/cpnc.30.

Despotović, D., Brandis, A., Savidor, A., Levin, Y., Fumagalli, L. & Tawfik, D. S. (2017) Diadenosine tetraphosphate (Ap4A) - an E. coli alarmone or a damage metabolite? *Febs j*, 284(14), 2194-2215. 10.1111/febs.14113.

Dietmair, S., Timmins, N. E., Gray, P. P., Nielsen, L. K. & Krömer, J. O. (2010) Towards quantitative metabolomics of mammalian cells: development of a metabolite extraction protocol. *Analytical biochemistry*, 404(2), 155-164.

Doamekpor, S. K., Sharma, S., Kiledjian, M. & Tong, L. (2022) Recent insights into noncanonical 5' capping and decapping of RNA. *Journal of Biological Chemistry*, 298(8), 102171. <https://doi.org/10.1016/j.jbc.2022.102171>.

Doil, C., Mailand, N., Bekker-Jensen, S., Menard, P., Larsen, D. H., Pepperkok, R., Ellenberg, J., Panier, S., Durocher, D., Bartek, J., Lukas, J. & Lukas, C. (2009) RNF168 binds and amplifies ubiquitin conjugates on damaged chromosomes to allow accumulation of repair proteins. *Cell*, 136(3), 435-46. 10.1016/j.cell.2008.12.041.

Dolot, R., Kaczmarek, R., Sęda, A., Krakowiak, A., Baraniak, J. & Nawrot, B. (2016) Crystallographic studies of the complex of human HINT1 protein with a non-hydrolyzable analog of Ap4A. *International Journal of Biological Macromolecules*, 87, 62-69. <https://doi.org/10.1016/j.ijbiomac.2016.02.047>.

Dolot, R., Krakowiak, A., Kaczmarek, R., Włodarczyk, A., Pichlak, M. & Nawrot, B. (2021) Biochemical, crystallographic and biophysical characterization of histidine triad nucleotide-binding protein 2 with different ligands including a non-hydrolyzable analog of Ap4A. *Biochimica et Biophysica Acta (BBA)-General Subjects*, 1865(11), 129968.

Drané, P., Brault, M. E., Cui, G., Meghani, K., Chaubey, S., Detappe, A., Parnandi, N., He, Y., Zheng, X. F., Botuyan, M. V., Kalousi, A., Yewdell, W. T., Münch, C., Harper, J. W., Chaudhuri, J., Soutoglou, E., Mer, G. & Chowdhury, D. (2017) TIRR regulates 53BP1 by masking its histone methyl-lysine binding function. *Nature*, 543(7644), 211-216. 10.1038/nature21358.

Duronio, R. J. & Xiong, Y. (2013) Signaling pathways that control cell proliferation. *Cold Spring Harb Perspect Biol*, 5(3), a008904. 10.1101/cshperspect.a008904.

Dwyer, D. J., Belenky, P. A., Yang, J. H., Macdonald, I. C., Martell, J. D., Takahashi, N., Chan, C. T. Y., Lobritz, M. A., Braff, D., Schwarz, E. G., Ye, J. D., Pati, M., Vercruysse, M., Ralifo, P. S., Allison, K. R., Khalil, A. S., Ting, A. Y., Walker, G. C. & Collins, J. J. (2014) Antibiotics induce redox-related physiological alterations as part of their lethality. *Proceedings of the National Academy of Sciences*, 111(20), E2100-E2109. doi:10.1073/pnas.1401876111.

Eaton, S. L., Hurtado, M. L., Oldknow, K. J., Graham, L. C., Marchant, T. W., Gillingwater, T. H., Farquharson, C. & Wishart, T. M. (2014) A guide to modern quantitative fluorescent western blotting with troubleshooting strategies. *J Vis Exp*(93), e52099. 10.3791/52099.

Escribano-Díaz, C., Orthwein, A., Fradet-Turcotte, A., Xing, M., Young, Jordan t. F., Tkáč, J., Cook, Michael a., Rosebrock, Adam p., Munro, M., Canny, Marella d., Xu, D. & Durocher, D. (2013) A Cell Cycle-Dependent Regulatory Circuit Composed of 53BP1-



RIF1 and BRCA1-CtIP Controls DNA Repair Pathway Choice. *Molecular Cell*, 49(5), 872-883. <https://doi.org/10.1016/j.molcel.2013.01.001>.

Evrin, C., Fernández-Cid, A., Riera, A., Zech, J., Clarke, P., Herrera, M. C., Tognetti, S., Lurz, R. & Speck, C. (2014) The ORC/Cdc6/MCM2-7 complex facilitates MCM2-7 dimerization during prereplicative complex formation. *Nucleic Acids Res*, 42(4), 2257-69. 10.1093/nar/gkt1148.

Fabino Carr, A., Patel, D. C., López, D. A., Armstrong, D. W. & Ryzhov, V. (2019) Comparison of reversed-phase, anion-exchange, and hydrophilic interaction HPLC for the analysis of nucleotides involved in biological enzymatic pathways. *Journal of Liquid Chromatography & Related Technologies*, 42, 184 - 193.

Farr, S. B., Arnosti, D. N., Chamberlin, M. J. & Ames, B. N. (1989) An apaH mutation causes AppppA to accumulate and affects motility and catabolite repression in Escherichia coli. *Proceedings of the National Academy of Sciences*, 86(13), 5010-5014. doi:10.1073/pnas.86.13.5010.

Feng, X., Noguchi, Y., Barbon, M., Stillman, B., Speck, C. & Li, H. (2021) The structure of ORC–Cdc6 on an origin DNA reveals the mechanism of ORC activation by the replication initiator Cdc6. *Nature Communications*, 12(1), 3883. 10.1038/s41467-021-24199-1.

Ferguson, F., Mclennan, A. G., Urbaniak, M. D., Jones, N. J. & Copeland, N. A. (2020) Re-evaluation of Diadenosine Tetraphosphate (Ap4A) From a Stress Metabolite to Bona Fide Secondary Messenger. *Frontiers in Molecular Biosciences*, 7. 10.3389/fmolb.2020.606807.

Fernández-Cid, A., Riera, A., Tognetti, S., Herrera, M. C., Samel, S., Evrin, C., Winkler, C., Gardenal, E., Uhle, S. & Speck, C. (2013) An ORC/Cdc6/MCM2-7 complex is formed in a multistep reaction to serve as a platform for MCM double-hexamer assembly. *Mol Cell*, 50(4), 577-88. 10.1016/j.molcel.2013.03.026.

Fernández-Justel, D., Peláez, R., Revuelta, J. L. & Buey, R. M. (2019) The Bateman domain of IMP dehydrogenase is a binding target for dinucleoside polyphosphates. *J Biol Chem*, 294(40), 14768-14775. 10.1074/jbc.AC119.010055.

Fontes, R., Sillero, M. A. & Sillero, A. (1998) Acyl coenzyme A synthetase from *Pseudomonas fragi* catalyzes the synthesis of adenosine 5'-polyphosphates and dinucleoside polyphosphates. *J Bacteriol*, 180(12), 3152-8. 10.1128/jb.180.12.3152-3158.1998.

Fradet-Turcotte, A., Canny, M. D., Escribano-Díaz, C., Orthwein, A., Leung, C. C., Huang, H., Landry, M. C., Kitevski-Leblanc, J., Noordermeer, S. M., Sicheri, F. & Durocher, D. (2013) 53BP1 is a reader of the DNA-damage-induced H2A Lys 15 ubiquitin mark. *Nature*, 499(7456), 50-4. 10.1038/nature12318.

Francis, L. I., Randell, J. C., Takara, T. J., Uchima, L. & Bell, S. P. (2009) Incorporation into the prereplicative complex activates the Mcm2-7 helicase for Cdc7-Dbf4 phosphorylation. *Genes Dev*, 23(5), 643-54. 10.1101/gad.1759609.

Fuge, E. K. & Farr, S. B. (1993) AppppA-binding protein E89 is the *Escherichia coli* heat shock protein ClpB. *J Bacteriol*, 175(8), 2321-6. 10.1128/jb.175.8.2321-2326.1993.

Fung, D. K., Yang, J., Stevenson, D. M., Amador-Noguez, D. & Wang, J. D. (2020) Small Alarmone Synthetase SasA Expression Leads to Concomitant Accumulation of pGpp, ppApp, and AppppA in *Bacillus subtilis*. *Frontiers in Microbiology*, 11. 10.3389/fmicb.2020.02083.

Galloway, A. & Cowling, V. H. (2019) mRNA cap regulation in mammalian cell function and fate. *Biochim Biophys Acta Gene Regul Mech*, 1862(3), 270-279. 10.1016/j.bbagrm.2018.09.011.

Garrison, P. N., Mathis, S. A. & Barnes, L. D. (1986) In vivo levels of diadenosine tetraphosphate and adenosine tetraphospho-guanosine in *Physarum polycephalum* during the cell cycle and oxidative stress. *Molecular and Cellular Biology*, 6(4), 1179-1186. doi:10.1128/mcb.6.4.1179-1186.1986.

Georgescu, R. E., Langston, L., Yao, N. Y., Yurieva, O., Zhang, D., Finkelstein, J., Agarwal, T. & O'donnell, M. E. (2014) Mechanism of asymmetric polymerase assembly at the eukaryotic replication fork. *Nature Structural & Molecular Biology*, 21(8), 664-670. 10.1038/nsmb.2851.

Giammarinaro, P. I., Young, M. K. M., Steinchen, W., Mais, C.-N., Hochberg, G., Yang, J., Stevenson, D. M., Amador-Noguez, D., Paulus, A., Wang, J. D. & Bange, G. (2022) Diadenosine tetraphosphate regulates biosynthesis of GTP in *Bacillus subtilis*. *Nature Microbiology*, 7(9), 1442-1452. 10.1038/s41564-022-01193-x.

Goerlich, O., Foeckler, R. & Holler, E. (1982) Mechanism of synthesis of adenosine(5')tetraphospho(5')adenosine (AppppA) by aminoacyl-tRNA synthetases. *Eur J Biochem*, 126(1), 135-42. 10.1111/j.1432-1033.1982.tb06757.x.

Götz, K. H., Hacker, S. M., Mayer, D., Dürig, J.-N., Stenger, S. & Marx, A. (2017) Inhibitors of the Diadenosine Tetraphosphate Phosphorylase Rv2613c of *Mycobacterium tuberculosis*. *ACS Chemical Biology*, 12(10), 2682-2689. 10.1021/acschembio.7b00653.

Götz, K. H., Mex, M., Stuber, K., Offensperger, F., Scheffner, M. & Marx, A. (2019) Formation of the Alarmones Diadenosine Triphosphate and Tetraphosphate by Ubiquitin- and Ubiquitin-like-Activating Enzymes. *Cell Chem Biol*, 26(11), 1535-1543.e5. 10.1016/j.chembiol.2019.08.004.

Grummt, F. (1978a) Diadenosine 5',5'''-P<sub>1</sub>P<sub>4</sub>-tetraphosphate triggers initiation of in vitro DNA replication in baby hamster kidney cells. *Proceedings of the National Academy of Sciences*, 75(1), 371-375. doi:10.1073/pnas.75.1.371.

Grummt, F. (1978b) Diadenosine 5', 5'''-P<sub>1</sub>, P<sub>4</sub>-tetraphosphate triggers initiation of in vitro DNA replication in baby hamster kidney cells. *Proceedings of the National Academy of Sciences*, 75(1), 371-375.

Grummt, F., Walzl, G., Jantzen, H. M., Hamprecht, K., Huebscher, U. & Kuenzle, C. C. (1979) Diadenosine 5',5'''-P<sub>1</sub>P<sub>4</sub>-tetraphosphate, a ligand of the 57-kilodalton subunit

of DNA polymerase alpha. *Proceedings of the National Academy of Sciences*, 76(12), 6081-6085. doi:10.1073/pnas.76.12.6081.

Guerra, J., Valadao, A. L., Vlachakis, D., Polak, K., Vila, I. K., Taffoni, C., Prabakaran, T., Marriott, A. S., Kaczmarek, R., Houel, A., Auzemery, B., Déjardin, S., Boudinot, P., Nawrot, B., Jones, N. J., Paludan, S. R., Kossida, S., Langevin, C. & Laguette, N. (2020) Lysyl-tRNA synthetase produces diadenosine tetraphosphate to curb STING-dependent inflammation. *Sci Adv*, 6(21), eaax3333. 10.1126/sciadv.aax3333.

Guo, R. T., Chong, Y. E., Guo, M. & Yang, X. L. (2009) Crystal structures and biochemical analyses suggest a unique mechanism and role for human glycyl-tRNA synthetase in Ap4A homeostasis. *J Biol Chem*, 284(42), 28968-76. 10.1074/jbc.M109.030692.

Guo, W., Azhar, M. A., Xu, Y., Wright, M., Kamal, A. & Miller, A. D. (2011) Isolation and identification of diadenosine 5',5'''-P<sub>1</sub>,P<sub>4</sub>-tetraphosphate binding proteins using magnetic bio-panning. *Bioorganic & Medicinal Chemistry Letters*, 21(23), 7175-7179. <https://doi.org/10.1016/j.bmcl.2011.09.070>.

Guo, X., Bai, Y., Zhao, M., Zhou, M., Shen, Q., Yun, C. H., Zhang, H., Zhu, W. G. & Wang, J. (2018) Acetylation of 53BP1 dictates the DNA double strand break repair pathway. *Nucleic Acids Res*, 46(2), 689-703. 10.1093/nar/gkx1208.

Guranowski, A., Brown, P., Ashton, P. A. & Blackburn, G. M. (1994a) Regiospecificity of the hydrolysis of diadenosine polyphosphates catalyzed by three specific pyrophosphohydrolases. *Biochemistry*, 33(1), 235-240.

Guranowski, A., Günther Sillero, M. A. & Sillero, A. (1994b) Adenosine 5'-tetraphosphate and adenosine 5'-pentaphosphate are synthesized by yeast acetyl coenzyme A synthetase. *J Bacteriol*, 176(10), 2986-90. 10.1128/jb.176.10.2986-2990.1994.

Guranowski, A., Jakubowski, H. & Holler, E. (1983) Catabolism of diadenosine 5',5'''-P<sub>1</sub>,P<sub>4</sub>-tetraphosphate in procaryotes. Purification and properties of diadenosine 5',5'''-P<sub>1</sub>,P<sub>4</sub>-tetraphosphate (symmetrical) pyrophosphohydrolase from *Escherichia coli* K12. *J Biol Chem*, 258(24), 14784-9.

Guranowski, A., Sillero, M. A. & Sillero, A. (1990) Firefly luciferase synthesizes P<sub>1</sub>,P<sub>4</sub>-bis(5'-adenosyl)tetraphosphate (Ap<sub>4</sub>A) and other dinucleoside polyphosphates. *FEBS Lett*, 271(1-2), 215-8. 10.1016/0014-5793(90)80409-c.

Gustafson, D. L. & Pritsos, C. A. (1992) Bioactivation of Mitomycin C by Xanthine Dehydrogenase From EMT6 Mouse Mammary Carcinoma Tumors. *JNCI: Journal of the National Cancer Institute*, 84(15), 1180-1185. 10.1093/jnci/84.15.1180.

Hanahan, D. & Weinberg, R. A. (2011) Hallmarks of cancer: the next generation. *cell*, 144(5), 646-674.

Hedstrom, L. (2009) IMP dehydrogenase: structure, mechanism, and inhibition. *Chem Rev*, 109(7), 2903-28. 10.1021/cr900021w.

Heinemann, V., Xu, Y. Z., Chubb, S., Sen, A., Hertel, L. W., Grindey, G. B. & Plunkett, W. (1990) Inhibition of ribonucleotide reduction in CCRF-CEM cells by 2',2'-difluorodeoxycytidine. *Mol Pharmacol*, 38(4), 567-72.

Hirano, T. (2015) Chromosome Dynamics during Mitosis. *Cold Spring Harb Perspect Biol*, 7(6). 10.1101/cshperspect.a015792.

Hofer, A., Crona, M., Logan, D. T. & Sjöberg, B. M. (2012) DNA building blocks: keeping control of manufacture. *Crit Rev Biochem Mol Biol*, 47(1), 50-63. 10.3109/10409238.2011.630372.

Homesley, L., Lei, M., Kawasaki, Y., Sawyer, S., Christensen, T. & Tye, B. K. (2000) Mcm10 and the MCM2-7 complex interact to initiate DNA synthesis and to release replication factors from origins. *Genes Dev*, 14(8), 913-26.

Horvath, C. G., Preiss, B. A. & Lipsky, S. R. (1967) Fast liquid chromatography: an investigation of operating parameters and the separation of nucleotides on pellicular ion exchangers. *Anal Chem*, 39(12), 1422-8. 10.1021/ac60256a003.

Huang, P., Chubb, S., Hertel, L. W., Grindey, G. B. & Plunkett, W. (1991) Action of 2',2'-difluorodeoxycytidine on DNA synthesis. *Cancer Res*, 51(22), 6110-7.

Huang, Y., Garrison, P. N. & Barnes, L. D. (1995) Cloning of the *Schizosaccharomyces pombe* gene encoding diadenosine 5',5'''-P<sub>1</sub>P<sub>4</sub>-tetraphosphate (Ap4A) asymmetrical hydrolase: sequence similarity with the histidine triad (HIT) protein family. *Biochem J*, 312 ( Pt 3)(Pt 3), 925-32. 10.1042/bj3120925.

Huber, R. M., Lucas, J. M., Gomez-Sarosi, L. A., Coleman, I., Zhao, S., Coleman, R. & Nelson, P. S. (2015) DNA damage induces GDNF secretion in the tumor microenvironment with paracrine effects promoting prostate cancer treatment resistance. *Oncotarget*, 6(4), 2134-2147. 10.18632/oncotarget.3040.

Hudeček, O., Benoni, R., Reyes-Gutierrez, P. E., Culka, M., Šanderová, H., Hubálek, M., Rulíšek, L., Cvačka, J., Krásný, L. & Cahová, H. (2020) Dinucleoside polyphosphates act as 5'-RNA caps in bacteria. *Nature Communications*, 11(1), 1052. 10.1038/s41467-020-14896-8.

Huen, M. S., Grant, R., Manke, I., Minn, K., Yu, X., Yaffe, M. B. & Chen, J. (2007) RNF8 transduces the DNA-damage signal via histone ubiquitylation and checkpoint protein assembly. *Cell*, 131(5), 901-14. 10.1016/j.cell.2007.09.041.

Huertas, P., Cortés-Ledesma, F., Sartori, A. A., Aguilera, A. & Jackson, S. P. (2008) CDK targets Sae2 to control DNA-end resection and homologous recombination. *Nature*, 455(7213), 689-92. 10.1038/nature07215.

Huertas, P. & Jackson, S. P. (2009) Human CtIP mediates cell cycle control of DNA end resection and double strand break repair. *J Biol Chem*, 284(14), 9558-65. 10.1074/jbc.M808906200.

Hyakkoku, K., Hamanaka, J., Tsuruma, K., Shimazawa, M. & Hara, H. (2010) Proteomic approach with LCMS-IT-TOF identified an increase of Rab33B after transient focal cerebral ischemia in mice. *Experimental & Translational Stroke Medicine*, 2(1), 20. 10.1186/2040-7378-2-20.

Ismail, T. M., Hart, C. A. & McLennan, A. G. (2003) Regulation of Dinucleoside Polyphosphate Pools by the YgdP and ApaH Hydrolases Is Essential for the Ability of *Salmonella enterica* serovar Typhimurium to Invade Cultured Mammalian Cells\*.

*Journal of Biological Chemistry*, 278(35), 32602-32607.  
<https://doi.org/10.1074/jbc.M305994200>.

Isono, M., Niimi, A., Oike, T., Hagiwara, Y., Sato, H., Sekine, R., Yoshida, Y., Isobe, S. Y., Obuse, C., Nishi, R., Petricci, E., Nakada, S., Nakano, T. & Shibata, A. (2017) BRCA1 Directs the Repair Pathway to Homologous Recombination by Promoting 53BP1 Dephosphorylation. *Cell Rep*, 18(2), 520-532. 10.1016/j.celrep.2016.12.042.

Itakura, E., Takai, K. K., Umeda, K., Kimura, M., Ohsumi, M., Tamai, K. & Matsuura, A. (2004) Amino-terminal domain of ATRIP contributes to intranuclear relocation of the ATR-ATRIP complex following DNA damage. *FEBS Lett*, 577(1-2), 289-93. 10.1016/j.febslet.2004.10.026.

Jakubowski, H. & Guranowski, A. (1983) Enzymes hydrolyzing ApppA and/or AppppA in higher plants. Purification and some properties of diadenosine triphosphatase, diadenosine tetraphosphatase, and phosphodiesterase from yellow lupin (*Lupinus luteus*) seeds. *J Biol Chem*, 258(16), 9982-9.

Jankowski, V., Schulz, A., Kretschmer, A., Mischak, H., Boehringer, F., Van Der Giet, M., Janke, D., Schuchardt, M., Herwig, R., Zidek, W. & Jankowski, J. (2013a) The enzymatic activity of the VEGFR2 receptor for the biosynthesis of dinucleoside polyphosphates. *J Mol Med (Berl)*, 91(9), 1095-107. 10.1007/s00109-013-1036-y.

Jankowski, V., Schulz, A., Kretschmer, A., Mischak, H., Boehringer, F., Van Der Giet, M., Janke, D., Schuchardt, M., Herwig, R., Zidek, W. & Jankowski, J. (2013b) The enzymatic activity of the VEGFR2 receptor for the biosynthesis of dinucleoside polyphosphates. *Journal of Molecular Medicine*, 91(9), 1095-1107. 10.1007/s00109-013-1036-y.

Jankowski, V., Van Der Giet, M., Mischak, H., Morgan, M., Zidek, W. & Jankowski, J. (2009) Dinucleoside polyphosphates: strong endogenous agonists of the purinergic system. *Br J Pharmacol*, 157(7), 1142-53. 10.1111/j.1476-5381.2009.00337.x.

Jensen, R. B., Carreira, A. & Kowalczykowski, S. C. (2010) Purified human BRCA2 stimulates RAD51-mediated recombination. *Nature*, 467(7316), 678-83. 10.1038/nature09399.

- Jeyaseelan, R., Poizat, C., Wu, H.-Y. & Kedes, L. (1997) Molecular Mechanisms of Doxorubicin-induced Cardiomyopathy: SELECTIVE SUPPRESSION OF REISKE IRON-SULFUR PROTEIN, ADP/ATP TRANSLOCASE, AND PHOSPHOFRUCTOKINASE GENES IS ASSOCIATED WITH ATP DEPLETION IN RAT CARDIOMYOCYTES\*. *Journal of Biological Chemistry*, 272(9), 5828-5832. <https://doi.org/10.1074/jbc.272.9.5828>.
- Ji, X., Zou, J., Peng, H., Stolle, A.-S., Xie, R., Zhang, H., Peng, B., Mekalanos, J. J. & Zheng, J. (2019) Alarmone Ap4A is elevated by aminoglycoside antibiotics and enhances their bactericidal activity. *Proceedings of the National Academy of Sciences*, 116(19), 9578-9585. doi:10.1073/pnas.1822026116.
- Johnstone, D. B. & Farr, S. B. (1991) AppppA binds to several proteins in Escherichia coli, including the heat shock and oxidative stress proteins DnaK, GroEL, E89, C45 and C40. *Embo j*, 10(12), 3897-904. 10.1002/j.1460-2075.1991.tb04959.x.
- Kamimura, Y., Tak, Y. S., Sugino, A. & Araki, H. (2001) Sld3, which interacts with Cdc45 (Sld4), functions for chromosomal DNA replication in Saccharomyces cerevisiae. *Embo j*, 20(8), 2097-107. 10.1093/emboj/20.8.2097.
- Kanke, M., Kodama, Y., Takahashi, T. S., Nakagawa, T. & Masukata, H. (2012) Mcm10 plays an essential role in origin DNA unwinding after loading of the CMG components. *Embo j*, 31(9), 2182-94. 10.1038/emboj.2012.68.
- Karanam, K., Kafri, R., Loewer, A. & Lahav, G. (2012) Quantitative live cell imaging reveals a gradual shift between DNA repair mechanisms and a maximal use of HR in mid S phase. *Mol Cell*, 47(2), 320-9. 10.1016/j.molcel.2012.05.052.
- Kawakami, H., Ohashi, E., Kanamoto, S., Tsurimoto, T. & Katayama, T. (2015) Specific binding of eukaryotic ORC to DNA replication origins depends on highly conserved basic residues. *Scientific Reports*, 5(1), 14929. 10.1038/srep14929.
- Kim, I. S., Nguyen, G. H., Kim, S.-J. & Jang, A. (2015) Qualitative analysis of the most toxic and abundant microcystin variants (LR, RR, and YR) by using LCMS–IT–TOF. *Journal of Industrial and Engineering Chemistry*, 29, 375-381.



King, R., Bonfiglio, R., Fernandez-Metzler, C., Miller-Stein, C. & Olah, T. (2000) Mechanistic investigation of ionization suppression in electrospray ionization. *Journal of the American Society for Mass Spectrometry*, 11(11), 942-950. [https://doi.org/10.1016/S1044-0305\(00\)00163-X](https://doi.org/10.1016/S1044-0305(00)00163-X).

Kitteringham, N. R., Jenkins, R. E., Lane, C. S., Elliott, V. L. & Park, B. K. (2009) Multiple reaction monitoring for quantitative biomarker analysis in proteomics and metabolomics. *Journal of Chromatography B*, 877(13), 1229-1239.

Kofuji, S. & Sasaki, A. T. (2020) GTP metabolic reprogramming by IMPDH2: unlocking cancer cells' fuelling mechanism. *J Biochem*, 168(4), 319-328. 10.1093/jb/mvaa085.

Kolas, N. K., Chapman, J. R., Nakada, S., Ylanko, J., Chahwan, R., Sweeney, F. D., Panier, S., Mendez, M., Wildenhain, J., Thomson, T. M., Pelletier, L., Jackson, S. P. & Durocher, D. (2007) Orchestration of the DNA-damage response by the RNF8 ubiquitin ligase. *Science*, 318(5856), 1637-40. 10.1126/science.1150034.

Krakowiak, A., Pęcherzewska, R., Kaczmarek, R., Tomaszewska, A., Nawrot, B. & Stec, W. J. (2011) Evaluation of influence of Ap4A analogues on Fhit-positive HEK293T cells; cytotoxicity and ability to induce apoptosis. *Bioorganic & Medicinal Chemistry*, 19(16), 5053-5060. <https://doi.org/10.1016/j.bmc.2011.06.028>.

Krude, T., Jackman, M., Pines, J. & Laskey, R. A. (1997) Cyclin/Cdk-dependent initiation of DNA replication in a human cell-free system. *Cell*, 88(1), 109-19. 10.1016/s0092-8674(00)81863-2.

Krüger, L., Albrecht, C. J., Schammann, H. K., Stumpf, F. M., Niedermeier, M. L., Yuan, Y., Stuber, K., Wimmer, J., Stengel, F. & Scheffner, M. (2021) Chemical proteomic profiling reveals protein interactors of the alarmones diadenosine triphosphate and tetraphosphate. *Nature communications*, 12(1), 1-13.

Kumagai, A., Lee, J., Yoo, H. Y. & Dunphy, W. G. (2006) TopBP1 Activates the ATR-ATRIP Complex. *Cell*, 124(5), 943-955. <https://doi.org/10.1016/j.cell.2005.12.041>.

- Kumagai, A., Shevchenko, A., Shevchenko, A. & Dunphy, W. G. (2011) Direct regulation of Treslin by cyclin-dependent kinase is essential for the onset of DNA replication. *J Cell Biol*, 193(6), 995-1007. 10.1083/jcb.201102003.
- Larsen, M. R., Thingholm, T. E., Jensen, O. N., Roepstorff, P. & Jørgensen, T. J. (2005) Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Mol Cell Proteomics*, 4(7), 873-86. 10.1074/mcp.T500007-MCP200.
- Lee, J.-H. & Paull, T. T. (2005) ATM Activation by DNA Double-Strand Breaks Through the Mre11-Rad50-Nbs1 Complex. *Science*, 308(5721), 551-554. doi:10.1126/science.1108297.
- Lee, P. C., Bochner, B. R. & Ames, B. N. (1983) AppppA, heat-shock stress, and cell oxidation. *Proceedings of the National Academy of Sciences*, 80(24), 7496-7500.
- Lee, Y. N., Nechushtan, H., Figov, N. & Razin, E. (2004) The function of lysyl-tRNA synthetase and Ap4A as signaling regulators of MITF activity in FcepsilonRI-activated mast cells. *Immunity*, 20(2), 145-51. 10.1016/s1074-7613(04)00020-2.
- Leung-Pineda, V., Ryan, C. E. & Piwnicka-Worms, H. (2006) Phosphorylation of Chk1 by ATR is antagonized by a Chk1-regulated protein phosphatase 2A circuit. *Mol Cell Biol*, 26(20), 7529-38. 10.1128/mcb.00447-06.
- Li, N., Lam, W. H., Zhai, Y., Cheng, J., Cheng, E., Zhao, Y., Gao, N. & Tye, B.-K. (2018) Structure of the origin recognition complex bound to DNA replication origin. *Nature*, 559(7713), 217-222. 10.1038/s41586-018-0293-x.
- Li, X. & Heyer, W.-D. (2008) Homologous recombination in DNA repair and DNA damage tolerance. *Cell Research*, 18(1), 99-113. 10.1038/cr.2008.1.
- Liu, J., Doty, T., Gibson, B. & Heyer, W. D. (2010) Human BRCA2 protein promotes RAD51 filament formation on RPA-covered single-stranded DNA. *Nat Struct Mol Biol*, 17(10), 1260-2. 10.1038/nsmb.1904.

- Liu, K., Myers, A. R., Pisithkul, T., Claas, K. R., Satyshur, K. A., Amador-Noguez, D., Keck, J. L. & Wang, J. D. (2015) Molecular mechanism and evolution of guanylate kinase regulation by (p)ppGpp. *Molecular cell*, 57(4), 735-749. 10.1016/j.molcel.2014.12.037.
- Lööke, M., Maloney, M. F. & Bell, S. P. (2017) Mcm10 regulates DNA replication elongation by stimulating the CMG replicative helicase. *Genes Dev*, 31(3), 291-305. 10.1101/gad.291336.116.
- Lou, Z., Minter-Dykhouse, K., Franco, S., Gostissa, M., Rivera, M. A., Celeste, A., Manis, J. P., Van Deursen, J., Nussenzweig, A., Paull, T. T., Alt, F. W. & Chen, J. (2006) MDC1 maintains genomic stability by participating in the amplification of ATM-dependent DNA damage signals. *Mol Cell*, 21(2), 187-200. 10.1016/j.molcel.2005.11.025.
- Luciano, D. J. & Belasco, J. G. (2020) Np(4)A alarmones function in bacteria as precursors to RNA caps. *Proc Natl Acad Sci U S A*, 117(7), 3560-3567. 10.1073/pnas.1914229117.
- Luciano, D. J., Levenson-Palmer, R. & Belasco, J. G. (2019) Stresses that Raise Np(4)A Levels Induce Protective Nucleoside Tetraphosphate Capping of Bacterial RNA. *Mol Cell*, 75(5), 957-966.e8. 10.1016/j.molcel.2019.05.031.
- Ma, C. J., Gibb, B., Kwon, Y., Sung, P. & Greene, E. C. (2016) Protein dynamics of human RPA and RAD51 on ssDNA during assembly and disassembly of the RAD51 filament. *Nucleic Acids Research*, 45(2), 749-761. 10.1093/nar/gkw1125.
- Madrid, O., Martín, D., Atencia, E. A., Sillero, A. & Günther Sillero, M. A. (1998) T4 DNA ligase synthesizes dinucleoside polyphosphates. *FEBS letters*, 433(3), 283-286.
- Mah, L. J., El-Osta, A. & Karagiannis, T. C. (2010)  $\gamma$ H2AX: a sensitive molecular marker of DNA damage and repair. *Leukemia*, 24(4), 679-686. 10.1038/leu.2010.6.
- Mailand, N., Bekker-Jensen, S., Faustrup, H., Melander, F., Bartek, J., Lukas, C. & Lukas, J. (2007) RNF8 ubiquitylates histones at DNA double-strand breaks and promotes assembly of repair proteins. *Cell*, 131(5), 887-900. 10.1016/j.cell.2007.09.040.

- Manfredi, G., Yang, L., Gajewski, C. D. & Mattiazzi, M. (2002) Measurements of ATP in mammalian cells. *Methods*, 26(4), 317-26. 10.1016/s1046-2023(02)00037-3.
- Mankan, A. K., Schmidt, T., Chauhan, D., Goldeck, M., Höning, K., Gaidt, M., Kubarenko, A. V., Andreeva, L., Hopfner, K. P. & Hornung, V. (2014) Cytosolic RNA:DNA hybrids activate the cGAS-STING axis. *Embo j*, 33(24), 2937-46. 10.15252/emboj.201488726.
- Maréchal, A. & Zou, L. (2013) DNA damage sensing by the ATM and ATR kinases. *Cold Spring Harb Perspect Biol*, 5(9). 10.1101/cshperspect.a012716.
- Marriott, A. S., Copeland, N. A., Cunningham, R., Wilkinson, M. C., McLennan, A. G. & Jones, N. J. (2015) Diadenosine 5', 5'''-P(1),P(4)-tetrphosphate (Ap4A) is synthesized in response to DNA damage and inhibits the initiation of DNA replication. *DNA Repair (Amst)*, 33, 90-100. 10.1016/j.dnarep.2015.06.008.
- Marriott, A. S., Vasieva, O., Fang, Y., Copeland, N. A., McLennan, A. G. & Jones, N. J. (2016) NUDT2 disruption elevates diadenosine tetrphosphate (Ap4A) and down-regulates immune response and cancer promotion genes. *PLoS One*, 11(5), e0154674.
- Martins, I., Tesniere, A., Kepp, O., Michaud, M., Schlemmer, F., Senovilla, L., S  ror, C., M  tivier, D., Perfettini, J. L., Zitvogel, L. & Kroemer, G. (2009) Chemotherapy induces ATP release from tumor cells. *Cell Cycle*, 8(22), 3723-8. 10.4161/cc.8.22.10026.
- Masai, H., Matsumoto, S., You, Z., Yoshizawa-Sugata, N. & Oda, M. (2010) Eukaryotic chromosome DNA replication: where, when, and how? *Annu Rev Biochem*, 79, 89-130. 10.1146/annurev.biochem.052308.103205.
- Masumoto, H., Muramatsu, S., Kamimura, Y. & Araki, H. (2002) S-Cdk-dependent phosphorylation of Sld2 essential for chromosomal DNA replication in budding yeast. *Nature*, 415(6872), 651-655. 10.1038/nature713.
- Matsumoto, T., Tostes, R. C. & Webb, R. C. (2011) The role of uridine adenosine tetrphosphate in the vascular system. *Adv Pharmacol Sci*, 2011, 435132. 10.1155/2011/435132.

- Matsuoka, S., Rotman, G., Ogawa, A., Shiloh, Y., Tamai, K. & Elledge, S. J. (2000) Ataxia telangiectasia-mutated phosphorylates Chk2 in vivo and in vitro. *Proc Natl Acad Sci U S A*, 97(19), 10389-94. 10.1073/pnas.190030497.
- Mattay, J. (2022) Noncanonical metabolite RNA caps: Classification, quantification, (de)capping, and function. *Wiley Interdiscip Rev RNA*, e1730. 10.1002/wrna.1730.
- Matthews, H. K., Bertoli, C. & De Bruin, R. a. M. (2022) Cell cycle control in cancer. *Nature Reviews Molecular Cell Biology*, 23(1), 74-88. 10.1038/s41580-021-00404-3.
- Mayle, R., Langston, L., Molloy, K. R., Zhang, D., Chait, B. T. & O'donnell, M. E. (2019) Mcm10 has potent strand-annealing activity and limits translocase-mediated fork regression. *Proceedings of the National Academy of Sciences*, 116(3), 798-803. doi:10.1073/pnas.1819107116.
- Mchugh, P. J., Spanswick, V. J. & Hartley, J. A. (2001) Repair of DNA interstrand crosslinks: molecular mechanisms and clinical relevance. *The Lancet Oncology*, 2(8), 483-490. [https://doi.org/10.1016/S1470-2045\(01\)00454-5](https://doi.org/10.1016/S1470-2045(01)00454-5).
- Mckethan, B. L. & Spiro, S. (2013) Cooperative and allosterically controlled nucleotide binding regulates the DNA binding activity of NrdR. *Mol Microbiol*, 90(2), 278-89. 10.1111/mmi.12364.
- Mclennan, A. G. (2000) Dinucleoside polyphosphates—friend or foe? *Pharmacology & Therapeutics*, 87(2), 73-89. [https://doi.org/10.1016/S0163-7258\(00\)00041-3](https://doi.org/10.1016/S0163-7258(00)00041-3).
- Mclennan, A. G. (2006) The Nudix hydrolase superfamily. *Cell Mol Life Sci*, 63(2), 123-43. 10.1007/s00018-005-5386-7.
- Méchali, M. (2010) Eukaryotic DNA replication origins: many choices for appropriate answers. *Nat Rev Mol Cell Biol*, 11(10), 728-38. 10.1038/nrm2976.
- Miller, T. C. R., Locke, J., Greiwe, J. F., Diffley, J. F. X. & Costa, A. (2019) Mechanism of head-to-head MCM double-hexamer formation revealed by cryo-EM. *Nature*, 575(7784), 704-710. 10.1038/s41586-019-1768-0.

- Minazzato, G., Gasparri, M., Amici, A., Cianci, M., Mazzola, F., Orsomando, G., Sorci, L. & Raffaelli, N. (2020) Functional Characterization of COG1713 (YqeK) as a Novel Diadenosine Tetraphosphate Hydrolase Family. *J Bacteriol*, 202(10). 10.1128/jb.00053-20.
- Mini, E., Nobili, S., Caciagli, B., Landini, I. & Mazzei, T. (2006) Cellular pharmacology of gemcitabine. *Annals of Oncology*, 17, v7-v12. <https://doi.org/10.1093/annonc/mdj941>.
- Miyazawa-Onami, M., Araki, H. & Tanaka, S. (2017) Pre-initiation complex assembly functions as a molecular switch that splits the Mcm2-7 double hexamer. *EMBO reports*, 18(10), 1752-1761. <https://doi.org/10.15252/embr.201744206>.
- Mladenov, E., Tsaneva, I. & Anachkova, B. (2007) Activation of the S phase DNA damage checkpoint by mitomycin C. *J Cell Physiol*, 211(2), 468-76. 10.1002/jcp.20957.
- Monds, R. D., Newell, P. D., Wagner, J. C., Schwartzman, J. A., Lu, W., Rabinowitz, J. D. & O'toole, G. A. (2010) Di-adenosine tetraphosphate (Ap<sub>4</sub>A) metabolism impacts biofilm formation by *Pseudomonas fluorescens* via modulation of c-di-GMP-dependent pathways. *J Bacteriol*, 192(12), 3011-23. 10.1128/jb.01571-09.
- Morano, C., Zhang, X. & Fricker, L. D. (2008) Multiple isotopic labels for quantitative mass spectrometry. *Anal Chem*, 80(23), 9298-309. 10.1021/ac801654h.
- Mordes, D. A., Glick, G. G., Zhao, R. & Cortez, D. (2008) TopBP1 activates ATR through ATRIP and a PIKK regulatory domain. *Genes Dev*, 22(11), 1478-89. 10.1101/gad.1666208.
- Mori, S., Shibayama, K., Wachino, J.-I. & Arakawa, Y. (2011) Structural Insights into the Novel Diadenosine 5',5'''-P<sub>1</sub>,P<sub>4</sub>-Tetraphosphate Phosphorylase from *Mycobacterium tuberculosis* H37Rv. *Journal of Molecular Biology*, 410(1), 93-104. <https://doi.org/10.1016/j.jmb.2011.04.059>.
- Mori, S., Shibayama, K., Wachino, J. & Arakawa, Y. (2010) Purification and molecular characterization of a novel diadenosine 5',5'''-P(1),P(4)-tetraphosphate phosphorylase

from *Mycobacterium tuberculosis* H37Rv. *Protein Expr Purif*, 69(1), 99-105. 10.1016/j.pep.2009.09.010.

Morioka, M. & Shimada, H. (1985) Change in the cellular level of AP4A is correlated with the initiation of DNA replication in sea urchin embryos. *Developmental Biology*, 112(1), 261-263. [https://doi.org/10.1016/0012-1606\(85\)90142-3](https://doi.org/10.1016/0012-1606(85)90142-3).

Motwani, M., Pesiridis, S. & Fitzgerald, K. A. (2019) DNA sensing by the cGAS–STING pathway in health and disease. *Nature Reviews Genetics*, 20(11), 657-674. 10.1038/s41576-019-0151-1.

Mueller, Adam c., Keaton, Mignon a. & Dutta, A. (2011) DNA Replication: Mammalian Treslin–TopBP1 Interaction Mirrors Yeast Sld3–Dpb11. *Current Biology*, 21(16), R638-R640. <https://doi.org/10.1016/j.cub.2011.07.004>.

Muramatsu, S., Hirai, K., Tak, Y. S., Kamimura, Y. & Araki, H. (2010) CDK-dependent complex formation between replication proteins Dpb11, Sld2, Pol (epsilon), and GINS in budding yeast. *Genes Dev*, 24(6), 602-12. 10.1101/gad.1883410.

Murphy, G. A., Halliday, D. & McLennan, A. G. (2000) The Fhit tumor suppressor protein regulates the intracellular concentration of diadenosine triphosphate but not diadenosine tetraphosphate. *Cancer research*, 60(9), 2342-2344.

Narasimha, A. M., Kaulich, M., Shapiro, G. S., Choi, Y. J., Sicinski, P. & Dowdy, S. F. (2014) Cyclin D activates the Rb tumor suppressor by mono-phosphorylation. *Elife*, 3. 10.7554/eLife.02872.

Ng, K. E. & Orgel, L. E. (1987) The action of a water-soluble carbodiimide on adenosine-5'-polyphosphates. *Nucleic acids research*, 15(8), 3573-3580. 10.1093/nar/15.8.3573.

Nick McElhinny, S. A., Gordenin, D. A., Stith, C. M., Burgers, P. M. & Kunkel, T. A. (2008) Division of labor at the eukaryotic replication fork. *Mol Cell*, 30(2), 137-44. 10.1016/j.molcel.2008.02.022.

Nimonkar, A. V., Özsoy, A. Z., Genschel, J., Modrich, P. & Kowalczykowski, S. C. (2008) Human exonuclease 1 and BLM helicase interact to resect DNA and initiate DNA repair.

*Proceedings of the National Academy of Sciences*, 105(44), 16906-16911.  
doi:10.1073/pnas.0809380105.

Nishimura, A., Moriya, S., Ukai, H., Nagai, K., Wachi, M. & Yamada, Y. (1997) Diadenosine 5',5'''-P<sub>1</sub>,P<sub>4</sub>-tetrphosphate (Ap<sub>4</sub>A) controls the timing of cell division in *Escherichia coli*. *Genes Cells*, 2(6), 401-13. 10.1046/j.1365-2443.1997.1300328.x.

Oka, K., Suzuki, T., Onodera, Y., Miki, Y., Takagi, K., Nagasaki, S., Akahira, J. I., Ishida, T., Watanabe, M. & Hirakawa, H. (2011) Nudix-type motif 2 in human breast carcinoma: a potent prognostic factor associated with cell proliferation. *International journal of cancer*, 128(8), 1770-1782.

Ouyang, Y.-F., Li, H.-B., Tang, H.-B., Jin, Y. & Li, G.-Y. (2015) A reliable and sensitive LCMS-IT-TOF method coupled with accelerated solvent extraction for the identification and quantitation of six typical heterocyclic aromatic amines in cooked meat products. *Analytical Methods*, 7(21), 9274-9280. 10.1039/C5AY01236H.

Pan, S.-S., Andrews, P. A., Glover, C. J. & Bachur, N. (1984) Reductive activation of mitomycin C and mitomycin C metabolites catalyzed by NADPH-cytochrome P-450 reductase and xanthine oxidase. *Journal of Biological Chemistry*, 259(2), 959-966.

Paull, T. T. (2015) Mechanisms of ATM Activation. *Annu Rev Biochem*, 84, 711-38. 10.1146/annurev-biochem-060614-034335.

Perera, R. L., Torella, R., Klinge, S., Kilkenny, M. L., Maman, J. D. & Pellegrini, L. (2013) Mechanism for priming DNA synthesis by yeast DNA polymerase  $\alpha$ . *Elife*, 2, e00482. 10.7554/eLife.00482.

Perez-Arnaiz, P., Bruck, I., Colbert, M. K. & Kaplan, D. L. (2017) An intact Mcm10 coiled-coil interaction surface is important for origin melting, helicase assembly and the recruitment of Pol- $\alpha$  to Mcm2-7. *Nucleic Acids Res*, 45(12), 7261-7275. 10.1093/nar/gkx438.



Piccinini, F., Tesei, A., Arienti, C. & Bevilacqua, A. (2017) Cell Counting and Viability Assessment of 2D and 3D Cell Cultures: Expected Reliability of the Trypan Blue Assay. *Biol Proced Online*, 19, 8. 10.1186/s12575-017-0056-3.

Plans, V., Scheper, J., Soler, M., Loukili, N., Okano, Y. & Thomson, T. M. (2006) The RING finger protein RNF8 recruits UBC13 for lysine 63-based self polyubiquitylation. *J Cell Biochem*, 97(3), 572-82. 10.1002/jcb.20587.

Plateau, P., Fromant, M. & Blanquet, S. (1987a) Heat shock and hydrogen peroxide responses of *Escherichia coli* are not changed by dinucleoside tetraphosphate hydrolase overproduction. *Journal of Bacteriology*, 169(8), 3817-3820. doi:10.1128/jb.169.8.3817-3820.1987.

Plateau, P., Fromant, M., Kepes, F. & Blanquet, S. (1987b) Intracellular 5',5'-dinucleoside polyphosphate levels remain constant during the *Escherichia coli* cell cycle. *J Bacteriol*, 169(1), 419-22. 10.1128/jb.169.1.419-422.1987.

Plateau, P., Mayaux, J. F. & Blanquet, S. (1981) Zinc(II)-dependent synthesis of diadenosine 5', 5''' -P(1) ,P(4) -tetraphosphate by *Escherichia coli* and yeast phenylalanyl transfer ribonucleic acid synthetases. *Biochemistry*, 20(16), 4654-62. 10.1021/bi00519a021.

Pointon, A. V., Walker, T. M., Phillips, K. M., Luo, J., Riley, J., Zhang, S.-D., Parry, J. D., Lyon, J. J., Marczylo, E. L. & Gant, T. W. (2010) Doxorubicin In Vivo Rapidly Alters Expression and Translation of Myocardial Electron Transport Chain Genes, Leads to ATP Loss and Caspase 3 Activation. *PLOS ONE*, 5(9), e12733. 10.1371/journal.pone.0012733.

Pozo, P. N. & Cook, J. G. (2016) Regulation and Function of Cdt1; A Key Factor in Cell Proliferation and Genome Stability. *Genes (Basel)*, 8(1). 10.3390/genes8010002.

Prescott, M., Milne, A. D. & McLennan, A. G. (1989) Characterization of the bis(5'-nucleosidyl) tetraphosphate pyrophosphohydrolase from encysted embryos of the brine shrimp *Artemia*. *Biochemical Journal*, 259(3), 831-838. 10.1042/bj2590831.

Pursell, Z. F., Isoz, I., Lundström, E. B., Johansson, E. & Kunkel, T. A. (2007) Yeast DNA polymerase epsilon participates in leading-strand DNA replication. *Science*, 317(5834), 127-30. 10.1126/science.1144067.

Quan, Y., Xia, Y., Liu, L., Cui, J., Li, Z., Cao, Q., Chen, X. S., Campbell, J. L. & Lou, H. (2015) Cell-Cycle-Regulated Interaction between Mcm10 and Double Hexameric Mcm2-7 Is Required for Helicase Splitting and Activation during S Phase. *Cell Rep*, 13(11), 2576-2586. 10.1016/j.celrep.2015.11.018.

Rabinowitz, J. D. & Kimball, E. (2007) Acidic Acetonitrile for Cellular Metabolome Extraction from *Escherichia coli*. *Analytical Chemistry*, 79(16), 6167-6173. 10.1021/ac070470c.

Ramanathan, A., Robb, G. B. & Chan, S. H. (2016) mRNA capping: biological functions and applications. *Nucleic Acids Res*, 44(16), 7511-26. 10.1093/nar/gkw551.

Randell, J. C., Fan, A., Chan, C., Francis, L. I., Heller, R. C., Galani, K. & Bell, S. P. (2010) Mec1 is one of multiple kinases that prime the Mcm2-7 helicase for phosphorylation by Cdc7. *Mol Cell*, 40(3), 353-63. 10.1016/j.molcel.2010.10.017.

Randerath, K., Janeway, C. M., Stephenson, M. L. & Zamecnik, P. C. (1966) Isolation and characterization of dinucleoside tetra- and tri-phosphates formed in the presence of lysyl-sRNA synthetase. *Biochem Biophys Res Commun*, 24(1), 98-105. 10.1016/0006-291x(66)90416-5.

Rao, P. N. & Johnson, R. T. (1970) Mammalian Cell Fusion : Studies on the Regulation of DNA Synthesis and Mitosis. *Nature*, 225(5228), 159-164. 10.1038/225159a0.

Rapaport, E. & Zamecnik, P. C. (1976) Presence of diadenosine 5', 5'''-P<sub>1</sub>, P<sub>4</sub>-tetrphosphate (Ap<sub>4</sub>A) in mamalian cells in levels varying widely with proliferative activity of the tissue: a possible positive" pleiotypic activator". *Proceedings of the National Academy of Sciences*, 73(11), 3984-3988.

Raththagala, M., Karunarathne, W., Kryziniak, M., Mccracken, J. & Spence, D. M. (2010) Hydroxyurea stimulates the release of ATP from rabbit erythrocytes through an

increase in calcium and nitric oxide production. *Eur J Pharmacol*, 645(1-3), 32-8. 10.1016/j.ejphar.2010.07.012.

Ravalico, F., Messina, I., Berberian, M. V., James, S. L., Migaud, M. E. & Vyle, J. S. (2011) Rapid synthesis of nucleotide pyrophosphate linkages in a ball mill. *Org Biomol Chem*, 9(19), 6496-7. 10.1039/c1ob06041d.

Razin, E., Zhang, Z. C., Nechushtan, H., Frenkel, S., Lee, Y.-N., Arudchandran, R. & Rivera, J. (1999) Suppression of Microphthalmia Transcriptional Activity by Its Association with Protein Kinase C-interacting Protein 1 in Mast Cells\*. *Journal of Biological Chemistry*, 274(48), 34272-34276. <https://doi.org/10.1074/jbc.274.48.34272>.

Reiss, J. R. & Moffatt, J. G. (1965) Dismutation Reactions of Nucleoside Polyphosphates. III. The Synthesis of  $\alpha,\iota$ -Dinucleoside 5'-Polyphosphates<sup>1</sup>. *The Journal of Organic Chemistry*, 30(10), 3381-3387. 10.1021/jo01021a029.

Remus, D., Beuron, F., Tolun, G., Griffith, J. D., Morris, E. P. & Diffley, J. F. (2009) Concerted loading of Mcm2-7 double hexamers around DNA during DNA replication origin licensing. *Cell*, 139(4), 719-30. 10.1016/j.cell.2009.10.015.

Ricke, R. M. & Bielinsky, A. K. (2004) Mcm10 regulates the stability and chromatin association of DNA polymerase- $\alpha$ . *Mol Cell*, 16(2), 173-85. 10.1016/j.molcel.2004.09.017.

Robinson, A. K., De La Peña, C. E. & Barnes, L. D. (1993) Isolation and characterization of diadenosine tetraphosphate (Ap<sub>4</sub>A) hydrolase from *Schizosaccharomyces pombe*. *Biochim Biophys Acta*, 1161(2-3), 139-48. 10.1016/0167-4838(93)90207-8.

Roh, D. S., Cook, A. L., Rhee, S. S., Joshi, A., Kowalski, R., Dhaliwal, D. K. & Funderburgh, J. L. (2008) DNA Cross-linking, Double-Strand Breaks, and Apoptosis in Corneal Endothelial Cells after a Single Exposure to Mitomycin C. *Investigative Ophthalmology & Visual Science*, 49(11), 4837-4843. 10.1167/iovs.08-2115.

Rothkamm, K., Krüger, I., Thompson, L. H. & Löbrich, M. (2003) Pathways of DNA double-strand break repair during the mammalian cell cycle. *Mol Cell Biol*, 23(16), 5706-15. 10.1128/mcb.23.16.5706-5715.2003.

Saldivar, J. C., Cortez, D. & Cimprich, K. A. (2017) The essential kinase ATR: ensuring faithful duplication of a challenging genome. *Nat Rev Mol Cell Biol*, 18(10), 622-636. 10.1038/nrm.2017.67.

Saleh, A., Noguchi, Y., Aramayo, R., Ivanova, M. E., Stevens, K. M., Montoya, A., Sunidhi, S., Carranza, N. L., Skwark, M. J. & Speck, C. (2022) The structural basis of Cdc7-Dbf4 kinase dependent targeting and phosphorylation of the MCM2-7 double hexamer. *Nature Communications*, 13(1), 2915. 10.1038/s41467-022-30576-1.

Samel, S. A., Fernández-Cid, A., Sun, J., Riera, A., Tognetti, S., Herrera, M. C., Li, H. & Speck, C. (2014) A unique DNA entry gate serves for regulated loading of the eukaryotic replicative helicase MCM2-7 onto DNA. *Genes Dev*, 28(15), 1653-66. 10.1101/gad.242404.114.

Sartori, A. A., Lukas, C., Coates, J., Mistrik, M., Fu, S., Bartek, J., Baer, R., Lukas, J. & Jackson, S. P. (2007) Human CtIP promotes DNA end resection. *Nature*, 450(7169), 509-514. 10.1038/nature06337.

Schmidt, U., Wollmann, Y., Franke, C., Grosse, F., Saluz, H. P. & Hänel, F. (2008) Characterization of the interaction between the human DNA topoisomerase II $\beta$ -binding protein 1 (TopBP1) and the cell division cycle 45 (Cdc45) protein. *Biochemical Journal*, 409(1), 169-177. 10.1042/BJ20070872.

Schulz, A., Jankowski, V., Zidek, W. & Jankowski, J. (2014) Highly sensitive, selective and rapid LC-MS method for simultaneous quantification of diadenosine polyphosphates in human plasma. *J Chromatogr B Analyt Technol Biomed Life Sci*, 961, 91-6. 10.1016/j.jchromb.2014.05.018.

Sekedat, M. D., Fenyő, D., Rogers, R. S., Tackett, A. J., Aitchison, J. D. & Chait, B. T. (2010) GINS motion reveals replication fork progression is remarkably uniform throughout the yeast genome. *Mol Syst Biol*, 6, 353. 10.1038/msb.2010.8.

Sellick, C. A., Hansen, R., Stephens, G. M., Goodacre, R. & Dickson, A. J. (2011) Metabolite extraction from suspension-cultured mammalian cells for global metabolite profiling. *Nature Protocols*, 6, 1241+.

Ser, Z., Liu, X., Tang, N. N. & Locasale, J. W. (2015) Extraction parameters for metabolomics from cultured cells. *Analytical biochemistry*, 475, 22-28. 10.1016/j.ab.2015.01.003.

Shang, F., Gong, X. & Taylor, A. (1997) Activity of ubiquitin-dependent pathway in response to oxidative stress. Ubiquitin-activating enzyme is transiently up-regulated. *J Biol Chem*, 272(37), 23086-93. 10.1074/jbc.272.37.23086.

Shen, T. & Huang, S. (2012) The role of Cdc25A in the regulation of cell proliferation and apoptosis. *Anticancer Agents Med Chem*, 12(6), 631-9. 10.2174/187152012800617678.

Sheu, Y.-J. & Stillman, B. (2010) The Dbf4–Cdc7 kinase promotes S phase by alleviating an inhibitory activity in Mcm4. *Nature*, 463(7277), 113-117. 10.1038/nature08647.

Sheu, Y. J. & Stillman, B. (2006) Cdc7-Dbf4 phosphorylates MCM proteins via a docking site-mediated mechanism to promote S phase progression. *Mol Cell*, 24(1), 101-13. 10.1016/j.molcel.2006.07.033.

Shimadzu (2022) *Triple Quadrupole LC-MS/MS Comparison*. Available at: <https://www.shimadzu.com/an/products/liquid-chromatograph-mass-spectrometry/triple-quadrupole-lc-msms/comparison/index.html> [Accessed 8th January 2023 2023].

Shorts-Cary, L., Xu, M., Ertel, J., Kleinschmidt-Demasters, B. K., Lillehei, K., Matsuoka, I., Nielsen-Preiss, S. & Wierman, M. E. (2007) Bone morphogenetic protein and retinoic acid-inducible neural specific protein-3 is expressed in gonadotrope cell pituitary adenomas and induces proliferation, migration, and invasion. *Endocrinology*, 148(3), 967-75. 10.1210/en.2006-0905.

- Sillero, A. & Sillero, M. a. a. G. (2000) Synthesis of dinucleoside polyphosphates catalyzed by firefly luciferase and several ligases. *Pharmacology & therapeutics*, 87(2-3), 91-102.
- Singh, A. & Xu, Y.-J. (2016) The Cell Killing Mechanisms of Hydroxyurea. *Genes*, 7(11), 99. 10.3390/genes7110099.
- Sirbu, B. M. & Cortez, D. (2013) DNA damage response: three levels of DNA repair regulation. *Cold Spring Harb Perspect Biol*, 5(8), a012724. 10.1101/cshperspect.a012724.
- So, S., Davis, A. J. & Chen, D. J. (2009) Autophosphorylation at serine 1981 stabilizes ATM at DNA damage sites. *J Cell Biol*, 187(7), 977-90. 10.1083/jcb.200906064.
- Song, M. G., Bail, S. & Kiledjian, M. (2013) Multiple Nudix family proteins possess mRNA decapping activity. *Rna*, 19(3), 390-9. 10.1261/rna.037309.112.
- Soniat, M. M., Myler, L. R., Kuo, H. C., Paull, T. T. & Finkelstein, I. J. (2019) RPA Phosphorylation Inhibits DNA Resection. *Mol Cell*, 75(1), 145-153.e5. 10.1016/j.molcel.2019.05.005.
- Sørensen, C. S., Syljuåsen, R. G., Falck, J., Schroeder, T., Rönnstrand, L., Khanna, K. K., Zhou, B. B., Bartek, J. & Lukas, J. (2003) Chk1 regulates the S phase checkpoint by coupling the physiological turnover and ionizing radiation-induced accelerated proteolysis of Cdc25A. *Cancer Cell*, 3(3), 247-58. 10.1016/s1535-6108(03)00048-5.
- Speck, C. & Stillman, B. (2007) Cdc6 ATPase activity regulates ORC x Cdc6 stability and the selection of specific DNA sequences as origins of DNA replication. *J Biol Chem*, 282(16), 11705-14. 10.1074/jbc.M700399200.
- Stewart, G. S., Panier, S., Townsend, K., Al-Hakim, A. K., Kolas, N. K., Miller, E. S., Nakada, S., Ylanko, J., Olivarius, S., Mendez, M., Oldreive, C., Wildenhain, J., Tagliaferro, A., Pelletier, L., Taubenheim, N., Durandy, A., Byrd, P. J., Stankovic, T., Taylor, A. M. & Durocher, D. (2009) The RIDDLE syndrome protein mediates a

ubiquitin-dependent signaling cascade at sites of DNA damage. *Cell*, 136(3), 420-34. 10.1016/j.cell.2008.12.042.

Strezsak, S. R., Beuning, P. J. & Skizim, N. J. (2022) Versatile separation of nucleotides from bacterial cell lysates using strong anion exchange chromatography. *J Chromatogr B Analyt Technol Biomed Life Sci*, 1188, 123044. 10.1016/j.jchromb.2021.123044.

Strom, A., Tong, C. L. & Wagner, C. R. (2020) Histidine triad nucleotide-binding proteins HINT1 and HINT2 share similar substrate specificities and little affinity for the signaling dinucleotide Ap4A. *FEBS Lett*, 594(10), 1497-1505. 10.1002/1873-3468.13745.

Stucki, M., Clapperton, J. A., Mohammad, D., Yaffe, M. B., Smerdon, S. J. & Jackson, S. P. (2005) MDC1 directly binds phosphorylated histone H2AX to regulate cellular responses to DNA double-strand breaks. *Cell*, 123(7), 1213-26. 10.1016/j.cell.2005.09.038.

Sun, J., Fernandez-Cid, A., Riera, A., Tognetti, S., Yuan, Z., Stillman, B., Speck, C. & Li, H. (2014) Structural and mechanistic insights into Mcm2-7 double-hexamer assembly and function. *Genes Dev*, 28(20), 2291-303. 10.1101/gad.242313.114.

Sun, L., Wu, J., Du, F., Chen, X. & Chen, Z. J. (2013) Cyclic GMP-AMP synthase is a cytosolic DNA sensor that activates the type I interferon pathway. *Science*, 339(6121), 786-91. 10.1126/science.1232458.

Sun, Y., Mccorvie, T. J., Yates, L. A. & Zhang, X. (2020) Structural basis of homologous recombination. *Cellular and Molecular Life Sciences*, 77(1), 3-18. 10.1007/s00018-019-03365-1.

Swarbrick, J. D., Buyya, S., Gunawardana, D., Gayler, K. R., McLennan, A. G. & Gooley, P. R. (2005) Structure and Substrate-binding Mechanism of Human Ap4A Hydrolase\*. *Journal of Biological Chemistry*, 280(9), 8471-8481. <https://doi.org/10.1074/jbc.M412318200>.

Swartz, M. E. (2010) HPLC DETECTORS: A BRIEF REVIEW. *Journal of Liquid Chromatography & Related Technologies*, 33, 1130 - 1150.

- Tak, Y. S., Tanaka, Y., Endo, S., Kamimura, Y. & Araki, H. (2006) A CDK-catalysed regulatory phosphorylation for formation of the DNA replication complex Sld2-Dpb11. *Embo j*, 25(9), 1987-96. 10.1038/sj.emboj.7601075.
- Takara, T. J. & Bell, S. P. (2011) Multiple Cdt1 molecules act at each origin to load replication-competent Mcm2-7 helicases. *Embo j*, 30(24), 4885-96. 10.1038/emboj.2011.394.
- Takayama, Y., Kamimura, Y., Okawa, M., Muramatsu, S., Sugino, A. & Araki, H. (2003) GINS, a novel multiprotein complex required for chromosomal DNA replication in budding yeast. *Genes Dev*, 17(9), 1153-65. 10.1101/gad.1065903.
- Takebayashi, S. I., Ogata, M. & Okumura, K. (2017) Anatomy of Mammalian Replication Domains. *Genes (Basel)*, 8(4). 10.3390/genes8040110.
- Tanaka, S. & Diffley, J. F. X. (2002) Interdependent nuclear accumulation of budding yeast Cdt1 and Mcm2-7 during G1 phase. *Nature Cell Biology*, 4(3), 198-207. 10.1038/ncb757.
- Tanaka, S., Umemori, T., Hirai, K., Muramatsu, S., Kamimura, Y. & Araki, H. (2007) CDK-dependent phosphorylation of Sld2 and Sld3 initiates DNA replication in budding yeast. *Nature*, 445(7125), 328-332. 10.1038/nature05465.
- Tanner, J. A., Wright, M., Christie, E. M., Preuss, M. K. & Miller, A. D. (2006) Investigation into the interactions between diadenosine 5',5'''-P1,P4-tetraphosphate and two proteins: molecular chaperone GroEL and cAMP receptor protein. *Biochemistry*, 45(9), 3095-106. 10.1021/bi052529k.
- Taymaz-Nikerel, H., Karabekmez, M. E., Eraslan, S. & Kırdar, B. (2018) Doxorubicin induces an extensive transcriptional and metabolic rewiring in yeast cells. *Scientific Reports*, 8(1), 13672. 10.1038/s41598-018-31939-9.
- Thorne, N. M., Hankin, S., Wilkinson, M. C., Nuñez, C., Barraclough, R. & McLennan, A. G. (1995) Human diadenosine 5',5'''-P1,P4-tetraphosphate pyrophosphohydrolase is a



member of the MutT family of nucleotide pyrophosphatases. *Biochem J*, 311 ( Pt 3)(Pt 3), 717-21. 10.1042/bj3110717.

Thorslund, T., Mcilwraith, M. J., Compton, S. A., Lekomtsev, S., Petronczki, M., Griffith, J. D. & West, S. C. (2010) The breast cancer tumor suppressor BRCA2 promotes the specific targeting of RAD51 to single-stranded DNA. *Nat Struct Mol Biol*, 17(10), 1263-5. 10.1038/nsmb.1905.

Tomasz, M. (1995) Mitomycin C: small, fast and deadly (but very selective). *Chem Biol*, 2(9), 575-9. 10.1016/1074-5521(95)90120-5.

Tomimatsu, N., Mukherjee, B., Deland, K., Kurimasa, A., Bolderson, E., Khanna, K. K. & Burma, S. (2012) Exo1 plays a major role in DNA end resection in humans and influences double-strand break repair and damage signaling decisions. *DNA Repair (Amst)*, 11(4), 441-8. 10.1016/j.dnarep.2012.01.006.

Torrents, E. (2014) Ribonucleotide reductases: essential enzymes for bacterial life. *Frontiers in cellular and infection microbiology*, 4, 52-52. 10.3389/fcimb.2014.00052.

Tsakraklides, V. & Bell, S. P. (2010) Dynamics of pre-replicative complex assembly. *J Biol Chem*, 285(13), 9437-9443. 10.1074/jbc.M109.072504.

Tseng, P. Y., Yu, W. P., Liu, H. Y., Zhang, X. D., Zou, X. & Chen, T. Y. (2011) Binding of ATP to the CBS domains in the C-terminal region of CLC-1. *J Gen Physiol*, 137(4), 357-68. 10.1085/jgp.201010495.

Tsuji, T., Ficarro, S. B. & Jiang, W. (2006) Essential role of phosphorylation of MCM2 by Cdc7/Dbf4 in the initiation of DNA replication in mammalian cells. *Molecular biology of the cell*, 17(10), 4459-4472.

Turkez, H., Aydin, E. & Aslan, A. (2012) Xanthoria elegans (Link) (lichen) extract counteracts DNA damage and oxidative stress of mitomycin C in human lymphocytes. *Cytotechnology*, 64(6), 679-86. 10.1007/s10616-012-9447-0.

Turnock, D. C. & Ferguson, M. A. (2007) Sugar nucleotide pools of Trypanosoma brucei, Trypanosoma cruzi, and Leishmania major. *Eukaryotic cell*, 6(8), 1450-1463.

Vartanian, A., Alexandrov, I., Prudowski, I., McLennan, A. & Kisselev, L. (1999) Ap4A induces apoptosis in human cultured cells. *FEBS Lett*, 456(1), 175-80. 10.1016/S0014-5793(99)00956-4.

Vartanian, A. A., Suzuki, H. & Poletaev, A. I. (2003) The involvement of diadenosine 5',5'''-P<sub>1</sub>P<sub>4</sub>-tetraphosphate in cell cycle arrest and regulation of apoptosis. *Biochemical Pharmacology*, 65(2), 227-235. [https://doi.org/10.1016/S0006-2952\(02\)01481-8](https://doi.org/10.1016/S0006-2952(02)01481-8).

Vermeulen, K., Van Bockstaele, D. R. & Berneman, Z. N. (2003) The cell cycle: a review of regulation, deregulation and therapeutic targets in cancer. *Cell proliferation*, 36(3), 131-149. 10.1046/j.1365-2184.2003.00266.x.

Viatchenko-Karpinski, V., Novosolova, N., Ishchenko, Y., Azhar, M. A., Wright, M., Tsintsadze, V., Kamal, A., Burnashev, N., Miller, A. D., Voitenko, N., Giniatullin, R. & Lozovaya, N. (2016) Stable, synthetic analogs of diadenosine tetraphosphate inhibit rat and human P2X<sub>3</sub> receptors and inflammatory pain. *Mol Pain*, 12. 10.1177/1744806916637704.

Wang, Y., Chang, C.-F., Morales, M., Chiang, Y.-H., Harvey, B. K., Su, T.-P., Tsao, L.-I., Chen, S. & Thiernemann, C. (2003) Diadenosine tetraphosphate protects against injuries induced by ischemia and 6-hydroxydopamine in rat brain. *Journal of Neuroscience*, 23(21), 7958-7965.

Ward, I. M. & Chen, J. (2001) Histone H2AX is phosphorylated in an ATR-dependent manner in response to replicational stress. *J Biol Chem*, 276(51), 47759-62. 10.1074/jbc.C100569200.

Weinmann-Dorsch, C., Hedl, A., Grummt, I., Albert, W., Ferdinand, F. J., Friis, R. R., Pierron, G., Moll, W. & Grummt, F. (1984a) Drastic rise of intracellular adenosine(5')tetraphospho(5')adenosine correlates with onset of DNA synthesis in eukaryotic cells. *Eur J Biochem*, 138(1), 179-85. 10.1111/j.1432-1033.1984.tb07897.x.

Weinmann-Dorsch, C., Pierron, G., Wick, R., Sauer, H. & Grummt, F. (1984b) High diadenosine tetraphosphate (Ap4A) level at initiation of S phase in the naturally

synchronous mitotic cycle of *Physarum polycephalum*. *Experimental cell research*, 155(1), 171-177.

Wiedermannová, J., Julius, C. & Yuzenkova, Y. (2021) The expanding field of non-canonical RNA capping: new enzymes and mechanisms. *R Soc Open Sci*, 8(5), 201979. 10.1098/rsos.201979.

Wu, J., Sun, L., Chen, X., Du, F., Shi, H., Chen, C. & Chen, Z. J. (2013) Cyclic GMP-AMP is an endogenous second messenger in innate immune signaling by cytosolic DNA. *Science*, 339(6121), 826-30. 10.1126/science.1229963.

Yan, S. & Michael, W. M. (2009) TopBP1 and DNA polymerase- $\alpha$  directly recruit the 9-1-1 complex to stalled DNA replication forks. *J Cell Biol*, 184(6), 793-804. 10.1083/jcb.200810185.

Yeeles, J. T., Deegan, T. D., Janska, A., Early, A. & Diffley, J. F. (2015) Regulated eukaryotic DNA replication origin firing with purified proteins. *Nature*, 519(7544), 431-5. 10.1038/nature14285.

Yeeles, J. T. P., Janska, A., Early, A. & Diffley, J. F. X. (2017) How the Eukaryotic Replisome Achieves Rapid and Efficient DNA Replication. *Mol Cell*, 65(1), 105-116. 10.1016/j.molcel.2016.11.017.

Yeh, Y. C. & Tessman, I. (1978) Differential effect of hydroxyurea on a ribonucleotide reductase system. *Journal of Biological Chemistry*, 253(5), 1323-1324. [https://doi.org/10.1016/S0021-9258\(17\)34866-4](https://doi.org/10.1016/S0021-9258(17)34866-4).

Yu, J., Liu, Z., Liang, Y., Luo, F., Zhang, J., Tian, C., Motzik, A., Zheng, M., Kang, J., Zhong, G., Liu, C., Fang, P., Guo, M., Razin, E. & Wang, J. (2019) Second messenger Ap4A polymerizes target protein HINT1 to transduce signals in Fc $\epsilon$ RI-activated mast cells. *Nature Communications*, 10(1), 4664. 10.1038/s41467-019-12710-8.

Yuan, Z., Riera, A., Bai, L., Sun, J., Nandi, S., Spanos, C., Chen, Z. A., Barbon, M., Rappsilber, J., Stillman, B., Speck, C. & Li, H. (2017) Structural basis of Mcm2-7

replicative helicase loading by ORC-Cdc6 and Cdt1. *Nat Struct Mol Biol*, 24(3), 316-324. 10.1038/nsmb.3372.

Yuan, Z., Schneider, S., Dodd, T., Riera, A., Bai, L., Yan, C., Barbon, M., Magdalou, I., Ivanov, I. & Stillman, B. (2020) Structural mechanism of helicase loading onto replication origin DNA by ORC-Cdc6. *Proceedings of the National Academy of Sciences*, 117(30), 17747-17756.

Yun, M. H. & Hiom, K. (2009) CtIP-BRCA1 modulates the choice of DNA double-strand-break repair pathway throughout the cell cycle. *Nature*, 459(7245), 460-463. 10.1038/nature07955.

Zamaraeva, M. V., Sabirov, R. Z., Maeno, E., Ando-Akatsuka, Y., Bessonova, S. V. & Okada, Y. (2005) Cells die with increased cytosolic ATP during apoptosis: a bioluminescence study with intracellular luciferase. *Cell Death & Differentiation*, 12(11), 1390-1397. 10.1038/sj.cdd.4401661.

Zamecnik, P. G., Stephenson, M. L., Janeway, C. M. & Randerath, K. (1966) Enzymatic synthesis of diadenosine tetraphosphate and diadenosine triphosphate with a purified lysyl-sRNA synthetase. *Biochemical and biophysical research communications*, 24(1), 91-97.

Zegerman, P. & Diffley, J. F. X. (2007) Phosphorylation of Sld2 and Sld3 by cyclin-dependent kinases promotes DNA replication in budding yeast. *Nature*, 445(7125), 281-285. 10.1038/nature05432.

Zellner, M., Babeluk, R., Diestinger, M., Pirchegger, P., Skeledzic, S. & Oehler, R. (2008) Fluorescence-based Western blotting for quantitation of protein biomarkers in clinical samples. *Electrophoresis*, 29(17), 3621-7. 10.1002/elps.200700935.

Zhang, A., Peng, B., Huang, P., Chen, J. & Gong, Z. (2017) The p53-binding protein 1-Tudor-interacting repair regulator complex participates in the DNA damage response. *J Biol Chem*, 292(16), 6461-6467. 10.1074/jbc.M117.777474.

Zhang, X., Shi, H., Wu, J., Zhang, X., Sun, L., Chen, C. & Chen, Z. J. (2013) Cyclic GMP-AMP containing mixed phosphodiester linkages is an endogenous high-affinity ligand for STING. *Mol Cell*, 51(2), 226-35. 10.1016/j.molcel.2013.05.022.

Zhao, C. & O'connor, P. B. (2007) Removal of polyethylene glycols from protein samples using titanium dioxide. *Anal Biochem*, 365(2), 283-5. 10.1016/j.ab.2007.03.024.

Zhao, W., Vaithiyalingam, S., San filippo, J., Maranon, David g., Jimenez-Sainz, J., Fontenay, Gerald v., Kwon, Y., Leung, Stanley g., Lu, L., Jensen, Ryan b., Chazin, Walter j., Wiese, C. & Sung, P. (2015) Promotion of BRCA2-Dependent Homologous Recombination by DSS1 via RPA Targeting and DNA Mimicry. *Molecular Cell*, 59(2), 176-187. <https://doi.org/10.1016/j.molcel.2015.05.032>.

Zheng, T., Jing, M., Gong, T., Yan, J., Zeng, J. & Li, Y. (2022) Deletion of the yqeK gene leads to the accumulation of Ap4A and reduced biofilm formation in Streptococcus mutans. *Molecular Oral Microbiology*, 37(1), 9-21.

Zhou, Z.-X., Lujan, S. A., Burkholder, A. B., Garbacz, M. A. & Kunkel, T. A. (2019a) Roles for DNA polymerase  $\delta$  in initiating and terminating leading strand DNA replication. *Nature Communications*, 10(1), 3992. 10.1038/s41467-019-11995-z.

Zhou, Z., Matsumoto, T., Jankowski, V., Pernow, J., Mustafa, S. J., Duncker, D. J. & Merkus, D. (2019b) Uridine adenosine tetraphosphate and purinergic signaling in cardiovascular system: An update. *Pharmacol Res*, 141, 32-45. 10.1016/j.phrs.2018.12.009.

Zou, L. & Elledge, S. J. (2003) Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes. *Science*, 300(5625), 1542-8. 10.1126/science.1083430.

Zou, L., Liu, D. & Elledge, S. J. (2003) Replication protein A-mediated recruitment and activation of Rad17 complexes. *Proc Natl Acad Sci U S A*, 100(24), 13827-32. 10.1073/pnas.2336100100.

Zourgui, L., Tharaud, D., Solari, A., Litvak, S. & Tarrago-Litvak, L. (1984) Stimulation of DNA synthesis by microinjection of diadenosine 5',5'''-P<sub>1</sub>, P<sub>4</sub>-tetrphosphate (Ap<sub>4</sub>A) into *Xenopus laevis* oocytes. *Developmental Biology*, 103(2), 409-413.  
[https://doi.org/10.1016/0012-1606\(84\)90328-2](https://doi.org/10.1016/0012-1606(84)90328-2).

Zur Nedden, S., Eason, R., Doney, A. S. & Frenguelli, B. G. (2009) An ion-pair reversed-phase HPLC method for determination of fresh tissue adenine nucleotides avoiding freeze-thaw degradation of ATP. *Anal Biochem*, 388(1), 108-14.  
10.1016/j.ab.2009.02.017.

Websites:

ThermoFisher Scientific, accessible via  
<https://www.thermofisher.com/order/catalog/product/Y3603> (last accessed on 6 February 2022)