

PCAF, SIRT1 and the regulation of substrate acetylation



Loane Serrano

Brasenose College
University of Oxford

This thesis is submitted to the Department of Chemistry, University of
Oxford, in partial fulfilment of the requirements for the degree of Doctor
of Philosophy
Trinity Term 2021



Brasenose
College
UNIVERSITY of OXFORD



Newcastle
University

Declaration

I declare that this thesis is entirely my own work except for listed items:

- His-HAT-BRD (EB1 and EB2) batches of protein were made and purified by Eidarus Salah (Kawamura group) for use by the author.
- RapifleX MALDI-TOF-MS and MALDI-TOF/TOF-MS data was acquired by Dr Maria Dueñas in collaboration with Matthias Trost's group at Newcastle University. Appendix Figure A.1 was made by Dr Maria Dueñas, used with permission.
- A MALDI-TOF-MS data recovery script was written by Dr Tom McAllister and used by the author.
- Protein mass spectra were acquired by Dr Anthony Tumber and Dr Joanna Bonnici (Kawamura group, Appendix Figure E.1-Appendix Figure E.5).
- Figure 5.1 was adapted from Serrano *et. al.* with permission from the authors.¹
- p1-His-HAT and p3-His-HAT-BRD plasmids were donated by the Structural Genomics Consortium
- PCAF bromodomain inhibitors were donated by Dr. Moses Moustakim (Brennan group).
- Other materials used in this work were either made by the author, purchased or donated, with the origin stated in the text and/or in Materials and Methods.

Where reference has been made to the work of others in the thesis, appropriate credit has been given.

¹ L. Serrano, G. Roper and A. Kawamura, in *The Discovery and Utility of Chemical Probes in Target Discovery*, eds. P. Brennan and S. Rodriguez Vazquez, The Royal Society of Chemistry, 2021, pp. 100–123.

Abstract

Levels of lysine acetylation (KAc) in cells are regulated by epigenetic enzymes with opposing roles; histone acetyl transferases (HATs) and histone deacetylases (HDACs). The HAT p300 CBP associating factor (PCAF) and the sirtuin HDAC SIRT1 have protein substrates in common, including histone 3 (H3) at lysine residues 9 and 14 (H3K9 and H3K14) and non-histone protein hypoxia inducible factor 1 α (HIF1 α). Studying PCAF and SIRT1 in parallel helps understand how levels of acetylation are maintained by the two classes of proteins.

In this work, a matrix assisted laser desorption ionisation (MALDI) time of flight (TOF) mass spectrometry (MS) assay was developed and used to characterise the activity of PCAF and SIRT1 proteins for their reported histone substrates (H3₁₋₂₁ and H3₁₋₂₁K9Ac respectively). MALDI-TOF-MS was used to effectively identify peptide and cofactor selectivity for the enzymes. Commercially available nonspecific HAT inhibitors were tested with MALDI-TOF-MS, identifying anacardic acid as the most potent inhibitor available against PCAF HAT and HAT-BRD catalytic acetylation activity.

Epigenetic enzymes operate in a context dependent manner in cells. Histone and non-histone proteins in cells often have multiple different post translational modifications (PTMs) at the same time (*e.g.* acetylation and methylation). These combinatorial PTMs are an important regulatory system affecting the activity of surrounding epigenetic enzymes, which is not often taken into account in peptide-based assays. For this reason, a rapifleX MALDI-TOF-MS high throughput screen (HTS) was developed to screen commercial libraries of histone peptides with combinatorial PTMs against PCAF and SIRT1 enzymes. This allows for mapping of PTM interactions with the enzymes, leading to changes in enzyme activity, referred to as PTM crosstalk. Instances of PTM crosstalk were identified, with a notable example being the activation of K18 acetylation on H3₁₋₂₁ by PCAF in the presence of a mono-methylation (Me1) PTM at K14 of the H3₁₋₂₁ substrate peptide. MALDI-TOF/TOF-MS was used to confirm site selectivity. Efforts to expand the method to screening of non-histone peptides using HIF1 α as a protein template were not successful, however, other proteins may be more suited to screening through peptide segments.

There is a lack of selective and potent inhibitors targeting PCAF HAT. As PTM modified histone peptides were shown to have inhibitory and activating effects on the PCAF enzyme, cyclic peptides were thought to be a good avenue for developing chemical probes for PCAF HAT. mRNA display was used to screen a library of cyclic peptides against PCAF HAT and HAT-BRD, however, no binding peptides were identified. It was determined that the screen failed due to nonspecific binding of PCAF to the hybrid mRNA/DNA tag used in the selection method.

In conclusion, an effective screening method has been developed to study the effects of PTM of histone peptides on epigenetic enzymes. The assay developed has potential for wider applications in screening epigenetic enzymes against a range of peptide libraries and epigenetic enzymes. In addition, the screening results provide insight into the biological activity of PCAF, and may provide a starting point for the development of PCAF selective chemical probes.

Acknowledgements

I would like to thank my supervisors Prof Akane Kawamura and Prof Paul Brennan. I could not have asked for more encouraging and helpful people to work with. Thank you for making yourselves available when needed and helping me develop my project. A particular thank you to Akane for always being very supportive and understanding. I have been lucky to get to know two great cities and survive a pandemic by your side.

Thank you to Dr Tom McAllister for his enduring patience, sharing his expertise and his advice. His contributions have been essential to the success of my project (not joint first author though, sorry).

I would like to thank everyone who has contributed to my project by donating materials or their time and advice. Eidarus Salah, Dr Oliver Coleman, Dr Joanna Bonnici, Dr John-Paul Bukowski, Dr Bhaskar Bhushan, Dr. Matthew Peake, Dr. Chiara Maniaci and Dr. Maria Dueñas in particular have all been very generous with their time.

Thank you to my parents Sylvie Lederlé and Didier Serrano for their invaluable support and encouragement. I could not have gotten here without you. Thank you to my late grandfather Jacques Lederlé, who was always encouraging and proud of me.

Thank you to my sister Maëlle Serrano for listening to me complain for 26 years (so far, and only because it took me two years to get going). Thank you to Iris for allowing me to take enough time off from storytelling to finish my thesis.

I would also like to thank all of the great friends that I have made over the last five years in both Oxford and Newcastle. Thank you Holly for agreeing to try all of my experimental cooking, and being fun, helpful, patient, and an all-around great person to suffer through

a DPhil with (I promise not to try and roast you again). Thank you Benoît for a regular, healthy dose of dark humour. Thank you Grace for exploring Newcastle with me and being a great pandemic buddy. Thank you Bhaskar and Karen for sharing my taste in epic books. Thank you to my friends at Brasenose for making me feel at home at Oxford. Thank you to present and past members of the Kawamura group for being a great bunch of people to work and go to the pub with. Thank you to the welcome trust lab members for not always talking about cricket. Thank you to the Newcastle PhD office lab members for making the last six months of my studies more light-hearted. And lastly, thank you to Margaery the MALDI for not dying on me.

COVID impact statement

The COVID-19 pandemic has had a large impact on working conditions for experimental work carried out for this thesis. My primary supervisor Prof. Akane Kawamura moved from Oxford University to The University of Newcastle in October 2019. I moved to Newcastle in January 2020 but due to significant delays in setting up of the laboratory facilities in Newcastle, no lab work could be carried out before end of February 2020. During this time, I co-authored a book chapter on the use of cyclic peptides as chemical probes.²

Due to the COVID-19 pandemic, Newcastle University shut down all onsite work on the 18th March 2020, and I was not able to return to experimental work until beginning of July 2020. I used this time to work on the introduction to my thesis and the mRNA display work (Chapter 5) which had been completed prior to my move to Newcastle University. Further time for experiments was lost due to periods of self-isolation following suspected COVID cases and visiting my family in France. To compensate for the lost time for experimental work caused by the lab move and the Coronavirus crisis, I was granted a three-term extension. Upon return to the lab, social distancing measures in place led to restricted numbers of people allowed to access work areas, restricted working hours, and restricted access to communal facilities. These restrictions remained in place during all of my remaining experimental work (July 2020 – March 2021).

The majority of the work presented in this thesis was carried out under social distancing restrictions, including development of the rapiflex MALDI-TOF-MS high throughput screen in collaboration with Dr. Maria Dueñas in Matthias Trost's lab at Newcastle University (I was not able to access the instruments due to restrictions), development of data analysis methods to process screen results in collaboration with Dr. Tom McAllister (Kawamura group), and all work relating to SIRT1. This corresponds to part of the work presented in Chapter 2, most of the work presented in Chapter 3 and all of the work presented in Chapter 4.

I would like to present the work that I have undertaken, whilst acknowledging that working and living conditions outside my control have impacted the extent of work which I have been able to carry out.

² L. Serrano, G. Roper and A. Kawamura, in *The Discovery and Utility of Chemical Probes in Target Discovery*, eds. P. Brennan and S. Rodriguez Vazquez, The Royal Society of Chemistry, 2021, pp. 100–123.

Table of Contents

Declaration.....	i
Abstract.....	ii
Acknowledgements	iii
COVID impact statement.....	v
Table of Contents.....	vi
List of figures	viii
List of tables.....	xi
Abbreviations	xiii
 Chapter 1 Introduction	 1
1.1 Epigenetics.....	1
1.2 PCAF.....	9
1.3 SIRT1	20
1.4 Complementary roles of PCAF and SIRT1.....	23
1.5 Chemical probes for HATs and HDACs	26
1.6 Project goals	38
 Chapter 2 Characterisation of the HAT and HDAC proteins PCAF and SIRT1 using a MALDI-TOF-MS assay	 41
2.1 Introduction.....	41
2.2 Characterisation of PCAF	49
2.3 Characterisation of SIRT1	75
2.5 Discussion	79
2.6 Future work	91
2.7 Conclusion	92
 Chapter 3 Development of a rapifleX MALDI-TOF-MS high throughput screen for identification of crosstalk between post translational modifications mediated by epigenetic enzymes.....	 94
3.1 Introduction.....	94
3.2 Technical information for the rapifleX MALDI-TOF-MS HTS	98
3.3 RapifleX MALDI-TOF-MS screen results for PCAF HAT and HAT-BRD	118
3.4 Results for rapifleX MALDI-TOF-MS HTS against SIRT1.....	151
3.5 Results for rapifleX MALDI-TOF/TOF-MS of HTS samples.....	154
3.6 Discussion	158

3.8	Future work	189
3.9	Conclusion	196
Chapter 4 Characterisation of PCAF and SIRT1 substrates		198
4.1	Introduction	198
4.2	Peptide synthesis	202
4.3	Characterisation of PCAF substrate selectivity	205
4.4	Characterisation of SIRT1 substrate selectivity	214
4.5	Discussion.....	218
4.6	Future work.....	226
4.7	Conclusion	229
Chapter 5 mRNA display screening of PCAF HAT/HAT-BRD for development of cyclic peptide chemical probes		231
5.1	Introduction	231
5.3	Preparation of materials for cyclic peptide mRNA-display.....	242
5.4	Cyclic peptide mRNA display screening of PCAF HAT/HAT-BRD.....	245
5.6	Discussion.....	258
5.7	Conclusions	267
Chapter 6 Conclusions		270
6.1	Summary	270
6.2	Conclusions and future work	275
Chapter 7 Materials and Methods.....		279
7.1	Materials used.....	279
7.2	General methods.....	279
7.3	Ligation independent cloning (LIC)	285
7.4	Protein expression and purification	287
7.5	Peptide synthesis	297
7.6	Peptides sequences.....	301
7.7	Procedures for MALDI-TOF-MS assays.....	303
7.8	mRNA display procedures	310
7.9	Bio Layer Interferometry (BLI)	317
Appendix		319
References		347

List of figures

Figure 1.1: Histone epigenetic post translational modifications (PTMs).....	2
Figure 1.2: Crosstalk between epigenetic PTMs on histone tails.	4
Figure 1.3: Lysine acetylation epigenetic PTM.....	5
Figure 1.4: Histone acetyl transferase (HAT) phylogenetic tree.....	8
Figure 1.5: Histone deacetylase (HDAC) phylogenetic tree.....	8
Figure 1.6: Composition of the PCAF protein and crystal structures of the HAT domains of PCAF and GCN5.	11
Figure 1.7: Mechanism of lysine substrate acetylation by PCAF HAT.....	12
Figure 1.8: Mechanism of NAD ⁺ mediated deacetylation of substrate KAc catalysed by sirtuins.	22
Figure 1.9: Bi-substrate and natural product chemical probes for HATs.	29
Figure 1.10: Small molecule chemical probes for HATs.....	32
Figure 1.11: Small molecule chemical probes for GCN5 HAT.	34
Figure 1.12: Reported chemical probes for PCAF BRD.	36
Figure 1.13: Selected chemical probes targeting SIRT1.....	37
Figure 2.1: MALDI-TOF-MS method.....	45
Figure 2.2: MALDI-TOF/TOF-MS method.....	47
Figure 2.3: Structure of the PCAF protein and PCAF HAT and HAT-BRD plasmids.	50
Figure 2.4: Protein gels showing purified PCAF protein for HAT and HAT-BRD construct variants.....	55
Figure 2.5: Development of a MALDI-TOF-MS activity assay.	57
Figure 2.6: Enzyme activity of different PCAF constructs compared.....	59
Figure 2.7: H3 ₁₋₂₁ acetylation by PCAF HAT.....	62
Figure 2.8: MS/MS spectra of H3 ₁₋₂₁ control peptide, and H3 ₁₋₂₁ acetylated by PCAF HAT (1Ac product).....	63
Figure 2.9: MS/MS spectra of H3 ₁₋₂₁ peptides acetylated using PCAF HAT (2Ac and 3Ac products).	64
Figure 2.10: Structures of acyl-CoAs tested with PCAF HAT and HAT-BRD for enzymatic acylation activity.	67
Figure 2.11: MALDI-TOF-MS spectra showing PCAF HAT activity with acetyl CoA, propionyl CoA, butyryl CoA and succinyl CoA.....	68
Figure 2.12: Acyl-CoA screen with PCAF HAT and HAT-BRD.....	69
Figure 2.13: Screening of bromodomain inhibitors with PCAF HAT-BRD.....	70
Figure 2.14: HAT inhibitors tested against PCAF HAT and HAT-BRD.....	72
Figure 2.15: Inhibition profiles of active HAT inhibitors tested against PCAF HAT and HAT-BRD using MALDI-TOF-MS.....	74
Figure 2.16: SIRT1 protein map, and SDS-PAGE gel of SIRT1 protein purification.	75
Figure 2.17: SIRT1 deacetylation of acylated H3 peptides.	77
Figure 3.1: Composition of Altabioscience histone peptide libraries SETI-IV.....	101
Figure 3.2: Time dependent assays for PCAF HAT, PCAF HAT-BRD and SIRT1 with control substrate peptides to determine end point for rapifleX MALDI-TOF-MS HTS assay....	103
Figure 3.3: Workflow developed for a rapifleX MALDI-TOF-MS high throughput screen.	107
Figure 3.4: Examples illustrating differences in parameters used for peak picking.	109

Figure 3.5: Example workflow for the data extraction and collation process involved in the analysis of results of the rapifleX MALDI-TOF-MS HTS.	114
Figure 3.6: Example spectra illustrating problems peptide quality in the histone peptide libraries.	117
Figure 3.7: Summary figures for the SETI-IV libraries screened against PCAF HAT and HAT-BRD in rapifleX MALDI-TOF-MS HTS.	126
Figure 3.8 Overview of the SETIV library screened against SIRT1 in a rapifleX MALDI-TOF-MS HTS.	153
Figure 3.9: MS/MS spectra of SETI library peptides incubated with PCAF HAT.	155
Figure 3.10: Summary of observed PCAF enzymatic activity against histone peptides with PTMs.....	182
Figure 4.1: Peptide synthesis scheme for automated synthesis on a Chorus synthesiser.	204
Figure 4.2: MALDI-TOF-MS traces for H3 ₁₋₂₁ peptides with K14 PTMs acetylated by PCAF HAT.....	207
Figure 4.3: MALDI-TOF/TOF-MS spectra of control H3 ₁₋₂₁ peptides with Ac or Me1 PTMs at K14.	208
Figure 4.4: Identification of 1Ac and 2Ac acetylation sites by PCAF HAT on H3 ₁₋₂₁ K14Me1 using MALDI-TOF/TOF-MS.	209
Figure 4.5: Identification of 1Ac acetylation sites by PCAF HAT on H3 ₁₋₂₁ K14Me2 and H3 ₁₋₂₁ K14Me3 using MALDI-TOF/TOF-MS.	210
Figure 4.6: Time resolved MALDI-TOF-MS assay of PCAF with a panel of H3 ₁₋₂₁ peptides with K14 PTMs.	212
Figure 4.7: Exploring PCAF HAT enzymatic activity for HIF1 α peptides.	213
Figure 4.8: MALDI-TOF-MS spectra showing SIRT1 deacetylation of HIF1 α peptides.	215
Figure 4.9: Time resolved MALDI-TOF-MS assay of SIRT1 with a H3 ₁₋₂₁ and HIF1 α substrate panel.	217
Figure 5.1 : mRNA display selection process and components.	238
Figure 5.2: HAT and HAT-BRD constructs of PCAF used in mRNA display.....	243
Figure 5.3: Bead loading test for PCAF HAT-bio and His-HAT-BRD-bio.	244
Figure 5.4: Optimisation of mRNA display conditions.	248
Figure 5.5: Results of qPCR monitoring of mRNA selections.....	250
Figure 5.6: Peptide synthesis and testing workflow used for cyclic peptides.	253
Figure 5.7: Peptide cyclisation scheme.....	254
Figure 5.8: Results of BLI binding and MALDI-TOF-MS activity of cyclic peptides against PCAF HAT and HAT-BRD.....	257
Figure 6.1: Visual summary of effects of PTM crosstalk on the acetylation catalytic activity of PCAF.	272

Appendix Figure A.1: Spectra acquired for a dilution series of H31-21 peptides in buffer on the rapifleX MALDI-TOF-MS.....	319
Appendix Figure B.1: HAT inhibitors showing no inhibitory of PCAF HAT/HAT-BRD activity.....	320
Appendix Figure C.1: MALDI-TOF-MS spectra for linear peptides used in PCAF characterisation.	322

Appendix Figure C.2: MALDI-TOF-MS spectra for cyclic peptides tested against PCAF for binding and inhibitory activity.	323
Appendix Figure E.1: Mass spectrum obtained for PCAF HAT protein.	344
Appendix Figure E.2: Mass spectrum obtained for PCAF His-HAT protein.....	344
Appendix Figure E.3: Mass spectrum obtained for PCAF HAT-bio protein.....	344
Appendix Figure E.4: Mass spectrum obtained for PCAF HAT-BRD protein.....	345
Appendix Figure E.5: Mass spectrum obtained for SIRT1 protein.....	345
Appendix Figure F.1: Screenshots of parameters used for auto-acquisition of MALDI-TOF-MS spectra on the Bruker microflex.....	346

List of tables

Table 1.1: Selected examples of reported biological activity of PCAF relating to cancer.	15
Table 1.2: Roles of SIRT1 in biological functions and diseases.	21
Table 3.1: Comparison of external and internal control peptides results for comparisons between libraries.	122
Table 3.2: SETIV library peptides matched with comparable peptides from other libraries.	123
Table 3.3: Library H3 ₁₋₂₁ peptides with combinatorial Ac PTMs were used to locate sites of enzymatic acetylation by PCAF HAT/HAT-BRD.	131
Table 3.4: Library H4 ₁₋₂₁ peptides with combinatorial Ac PTMs were used to locate sites of enzymatic acetylation by PCAF HAT/HAT-BRD.	134
Table 3.5: Library H2a ₁₋₂₁ peptides with combinatorial Ac PTMs were used to locate sites of enzymatic acetylation by PCAF HAT/HAT-BRD.	135
Table 3.6: Results for screening of library H2b ₁₋₂₁ peptides with combinatorial Ac PTMs with PCAF HAT/HAT-BRD.	137
Table 3.7: Effects of H3R2Me2a, H3T3Phos and H3K4Me1/2/3 on PCAF activity were examined.	139
Table 3.8: Effects of H3S10Phos and H3T11Phos PTMs on PCAF activity for H3 ₁₋₂₁ peptides.	141
Table 3.9: Effects of H3T3Phos, H4S1Phos and H2aS1Phos on PCAF activity were examined.	143
Table 3.10: Effects of H3S28Phos, H3T32Phos and H2aS32Phos on PCAF activity were examined.	146
Table 3.11: Effects of lysine mono-, di- or tri-methylation on PCAF activity were examined.	150
Table 3.12: Biotools scoring of Ac PTM site identification in MALDI-TOF/TOF-MS spectra obtained for library peptides incubated with PCAF HAT.	157
Table 3.13: List of peptides from the rapifleX MALDI-TOF-MS HTS to be tested in future work.	191
Table 4.1: Peptides chosen for substrate characterisation with PCAF and SIRT1.	203
Table 4.2: Summary of results obtained for HIF1 α peptides with PCAF and SIRT1 in the MALDI-TOF-MS assay compared with literature findings.	224
Table 4.3: Peptides to be synthesised for further characterisation of PCAF activity.	227
Table 5.1: Conditions used for mRNA display Selections 1-5 against PCAF HAT and HAT-BRD.	247
Table 5.2: Results of NGS sequencing.	252
Table 5.3: Peptide sequences synthesised for testing.	255
Table 6.1: Summary of main results obtained in the rapifleX MALDI-TOF-MS HTS against PCAF HAT and HAT-BRD (Chapter 3).	273
Table 6.2: Summary of results on PCAF HAT/HAT-BRD and SIRT1 enzyme selectivity in this study.	274
Table 7.1: Media used in microbiology and protein production work.	280
Table 7.2: Quantity of agarose needed to make up 50 mL of agarose gel at specified percentage.	283
Table 7.3: Primers used in LIC cloning for PCAF HAT and HAT-BRD constructs.	287

Table 7.4: Buffers for protein purification of PCAF proteins	290
Table 7.5: Buffers for protein purification of SIRT1 protein	290
Table 7.6: Peptides used in MALDI-TOF-MS assays for characterising PCAF HAT/HAT-BRD and SIRT1 enzyme activity.....	302
Table 7.7: Cyclic peptides synthesised following mRNA display against PCAF HAT and HAT-BRD.....	303
Table 7.8: Assay conditions for testing of inhibitors with PCAF HAT and HAT-BRD.....	306
Table 7.9: Specifications and conditions used for each of the MALDI-TOF-MS screens performed with AltaBioscience histone peptide libraries	309
Table 7.10: Primers and building blocks used in the mRNA display procedures.....	315

Appendix Table D.1: Results of rapifleX MALDI-TOF-MS HTS of the SETI library with PCAF HAT, PCAF HAT-BRD and a NEC.....	327
Appendix Table D.2: Results of rapifleX MALDI-TOF-MS HTS of the SETII library with PCAF HAT, PCAF HAT-BRD and a NEC.....	331
Appendix Table D.3: Results of rapifleX MALDI-TOF-MS HTS of the SETIII library with PCAF HAT, PCAF HAT-BRD and a NEC.	335
Appendix Table D.4: Results of rapifleX MALDI-TOF-MS HTS of the SETIV library with PCAF HAT, PCAF HAT-BRD and a NEC.	339
Appendix Table D.5: Results of rapifleX MALDI-TOF-MS HTS of the SETIV library with SIRT1 and a NEC.	343

Abbreviations

Abbreviation	Full name
1Ac	single acetylation
2Ac	double acetylation
Ac	acetyl
AcCoA	acetyl coenzyme A / acetyl CoA
AnA	anacardic acid
ATAC	ada-Two-A-containing complex
bio	C-terminal avitag of a plasmid
BirA	biotin ligase
BLI	bio-layer interferometry
BRD	bromodomain
BSA	bovine serum albumin
Bu	butyryl
CAA	canonical amino acid
CBP	CREB binding protein
CHCA	4-hydroxy- α -cyanocinnamic acid
CIAcY	chloroacetyl-tyrosine
DEL	DNA encoded library
DIC	N,N'-Diisopropylcarbodiimide
ELISA	enzyme-linked immunosorbent assay
Emb	embelin
Ep	potential energy
ESI	electrospray ionisation
fMet	formyl-methionine
FRET	Förster resonance energy transfer
GCN5	general control nonderepressible 5
GNAT	GCN5-related N-acetyltransferase
h	human
H3 ₁₋₂₁	histone 3 residues 1 to 21
HAT	histone acetyl transferase
HDA1	histone deacetylase 1
HDAC	histone deacetylase
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HIF1	hypoxia inducible factor-1
HIF1 α	hypoxia inducible factor 1 α
His	poly-histidine
HTS	high throughput screen
KAT2A	lysine acetyl transferase 2A (GCN5)
KAT2B	lysine acetyl transferase 2B (PCAF)
KAT3A	lysine acetyl transferase 3A (p300)
LCMS	liquid chromatography mass spectrometry
LIC	ligation independent cloning
LID	laser induced decay

Abbreviation	Full name
m	mouse
MALDI	matrix assisted laser desorption ionisation
Me1	mono-methyl
Me2	di-methyl
Me3	tri-methyl
MS	mass spectrometry
MYST	Morf, Ybf2/Sas3, Sas2 and Tip60
N	number of replicates
NAD+	nicotinamide adenine dinucleotide
nCAA	non-canonical amino acid
NEC	no enzyme control
NGS	next generation sequencing
NMR	nuclear magnetic resonance
OBOC	one bead one compound
PCAF	p300 CBP associating factor
Phos	phosphate
PPI	protein-protein interaction
Pr	propionyl
PSD	post source decay
RINK	4-(2',4'-Dimethoxyphenyl-Fmoc-aminomethyl)-phenoxy
SAGA	spt-Ada-GCN5 acetyltransferase
SD	standard deviation
SEC	size exclusion chromatography
SICLOPPS	split-intein circular ligation of peptides and proteins
SIR2	silent information regulator 2
SIRT	sirtuin class of proteins
SPPS	solid phase peptide synthesis
SPR	surface plasmon resonance
TAR	trans-activation response
TAT1	HIV type 1 trans- activator protein
TB	terrific broth
TEV	tobacco etch virus
TFA	trifluoroacetic acid
TOF	time of flight
TSS	transcription start sites
-Xac	X number of deacetylations
y	yeast
YEATS	Yaf9, ENL, AF9, Taf14, and Sas5
yRNA	yeast RNA
ySir2	silent information regulator 2
Y-tRNA	tyrosine tRNA
ΔAA	amino acid free

Chapter 1 Introduction

1.1 Epigenetics

1.1.1 The concept of epigenetics

Epigenetics is defined as the mechanisms which modulate chromatin, allowing for inheritance of phenotypic traits (patterns of gene expression) without changes to the DNA template. Phenotype is therefore dependent both on the DNA and epigenetic layout of cells.¹ Chromatin modulation goes through epigenetic post translational modification (PTM) of proteins, and chemical modifications of DNA and RNA.¹⁻⁴ These modifications are regulated by a network of epigenetic control enzymes. As technology used to map epigenetic activity has evolved, our understanding of epigenetic regulation has increased. It has become clear that epigenetic cellular controls can become dysregulated, leading to disease states. Aberrant epigenetic marks have been linked to cancer, neurological disorders immunological diseases and much more.⁵⁻⁸ Alterations to the epigenome can be due to either environmental or genetic factors.⁹⁻¹³ Having a good comparative understanding of epigenetics in normal and disease states is therefore essential in artificially modulating disease states.

1.1.1.1 Role of epigenetic PTMs

Epigenetic protein PTMs were first identified on histone proteins in nucleosomes, the basic units of chromatin.¹⁴ An octamer of histone proteins forms the nucleosome core, containing two of H2a, H2b, H3 and H4, respectively. H1 associates to this core, forming a full nucleosome around which two 80 base pair superhelical DNA loops are coiled.¹⁵ The *N*-terminal tails of the histone proteins remain accessible outside the nucleosome structure, and are subject to epigenetic protein modification. PTMs found on histone

tails are varied and are located at numerous sites, with new PTMs regularly identified.¹⁶ A selection of known PTMs on histone tails is shown in Figure 1.1, including acetylation, phosphorylation, methylation and ubiquitination.¹⁷ PTMs on nucleosome tails have a role in regulating nucleosome structure and dynamics,¹⁸ and regulating access to DNA through PTM dictated switches between transcriptionally active or silent chromatin states.¹⁹

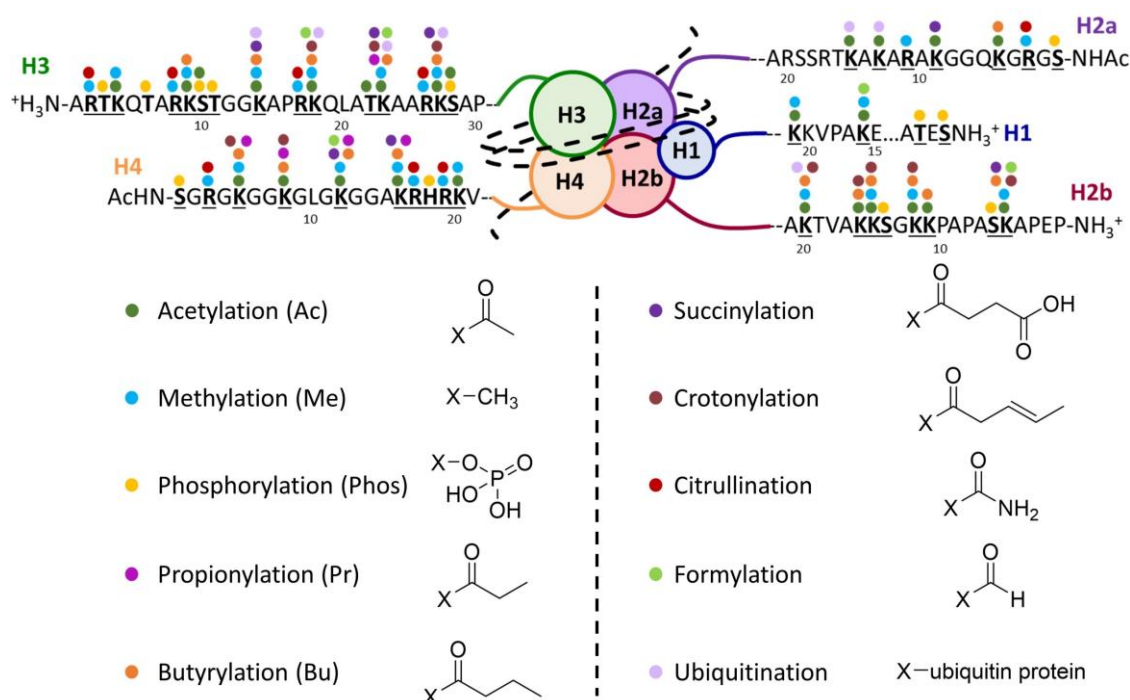


Figure 1.1: Histone epigenetic post translational modifications (PTMs).

Selected PTM types found on N-terminal histone tails of H1, H2a, H2b, H3 and H4. Structure of the PTMs are shown with X representing the amino acids onto which the PTMs are added. Figure adapted from Huang et. al.¹⁷

Although histone epigenetic PTMs are the best characterised, it has since been shown that epigenetic enzymes can modify non-histone proteins, contributing to signalling networks in all cellular compartments. This illustrates the ubiquitous role of epigenetics in the cell signalling and control mechanisms.^{20–22} New examples of non-histone proteins with epigenetic PTMs are regularly identified through proteomic studies,^{1,23} with a similar range in PTM types as found on histones.^{2,17}

Epigenetic PTMs are covalently added to protein residues by 'writer' enzymes, act as binding sites for 'reader' proteins and are removed by 'eraser' enzymes.^{1,24} These different classes of epigenetic proteins work in tandem to control the levels of epigenetic PTMs in the cell.

1.1.1.2 Crosstalk in epigenetics

Identifying and characterising individual PTMs and epigenetic proteins is important to help understand the mechanisms by which epigenetic PTMs are controlled. However, epigenetics is a vast signalling network which operates in the context of other cellular systems. There are many examples of crosstalk between signalling systems such as between epigenetics and either metabolism,^{25–27} cancer mechanisms,^{25,28,29} non-coding RNAs,³⁰ or transcription mechanisms.³¹ There is also crosstalk between different areas of epigenetics such as DNA chemical modifications co-ordinating with histone PTMs.^{32–34} These epigenetic networks serve as a mechanism for forming and reinforcing positive or negative feedback loops in cell regulation.

In addition to crosstalk between signalling systems and within epigenetics, crosstalk can be identified between individual PTMs. Histone tail PTMs are found in combinations dependent on cellular context and must therefore be considered as a dynamic interactive system as well as individual marks. There are numerous examples of positive or negative feedback between epigenetic PTMs either on a single histone tail (cis effects) or between histone tails (trans effects).^{19,35–39} Some examples of crosstalk between PTMs on histone tails are shown in Figure 1.2, including a previously unknown example of crosstalk identified in this work between histone 3 lysine 14 mono-methylation (H3K14Me1) and histone 3 lysine 18 acetylation (H3K18Ac).^{17,35,40–52}

Crosstalk between PTMs is mediated by epigenetic effector proteins (readers, writers and erasers). Combinatorial histone PTMs affect epigenetic proteins through sequence and PTM specific binding interactions, and it is through these epigenetic proteins that the PTMs affect each other.^{19,35,36} An additional level of control comes through association of different epigenetic proteins, where a single protein can contain multiple epigenetic domains and/or can be part of a complex with multiple epigenetic proteins, mixing readers, writers and erasers. The interactions between these different epigenetic effector proteins creates additional feedback, signalling and control mechanisms which play a large role in regulating the site selectivity and activity of many epigenetic proteins and enzymes.⁵³ Because of their ubiquitous roles in cellular systems, epigenetic proteins have become targets for drug discovery.^{54,55}

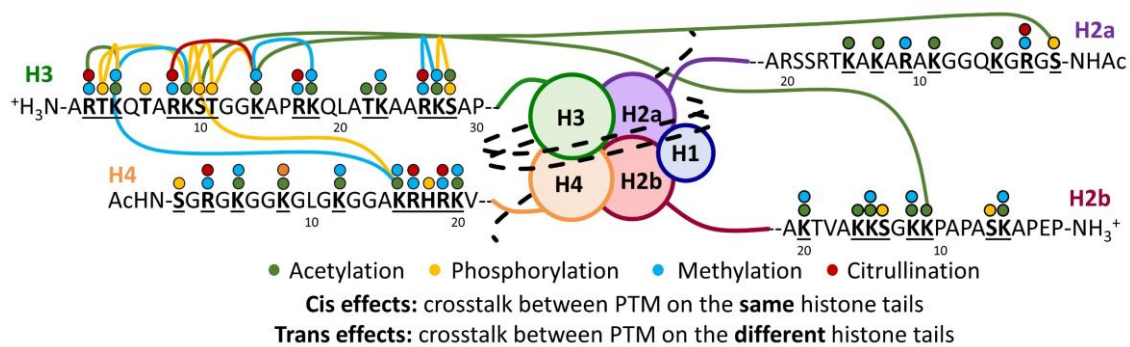


Figure 1.2: Crosstalk between epigenetic PTMs on histone tails.

PTMs are indicated on N-terminal tails of H2a, H2b, H3 and H4, colour coded by PTM type. Crosstalk between epigenetic marks is indicated by a line corresponding to the colour of the PTM having a positive or negative effect on its target site.^{17,35,40–52}

1.1.2 Lysine acetylation of proteins

Lysine acetylation was the first PTM discovered on histones in 1963.¹⁴ Protein acetylation is an important post translational modification which is involved in many cellular processes and disease states.⁵⁶ Acetylation can occur during translation (at the protein N-terminus) or post-translationally. N-terminal acetylation is found in >80% of human proteins and is an important marker which can regulate protein localisation,

enzymatic activity, interactions, stability and degradation.⁵⁷ Post-translational lysine acetylation occurs by the transfer of the acetyl (Ac) group of acetyl coenzyme A (acetyl CoA or AcCoA) onto a primary amine at the ϵ -position of a lysine residue side chain (Figure 1.3). Lysine PTM acetylation is found in the cytoplasm, nucleus and mitochondria.^{58,59} Acetylation is an important cell signalling tool, which is both influenced by and can affect metabolism (particularly regulation of AcCoA levels),⁶⁰ gene transcription and the development of disease states.^{22,61–63} Many methods for mapping acetylation sites have been developed including proteomics, genomics, mass spectrometry, and isotope labelling.^{56,64}

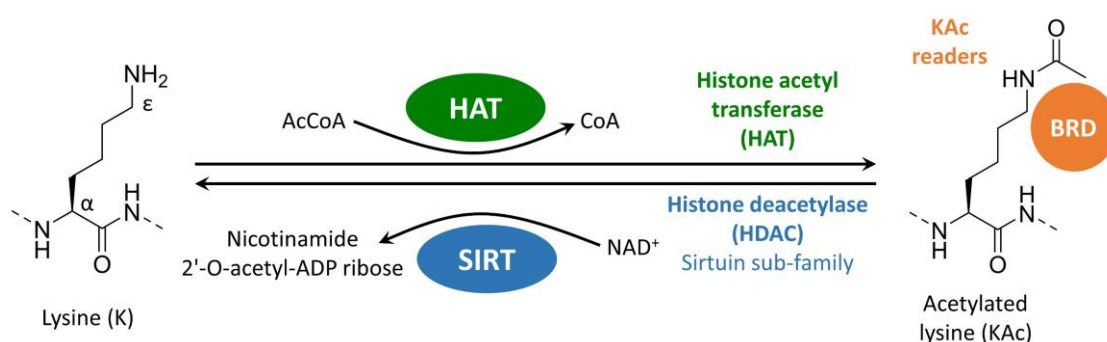


Figure 1.3: Lysine acetylation epigenetic PTM.

Biological levels of lysine acetylation (KAc) are regulated by histone acetyl transferase (HAT) writer proteins, histone deacetylase (HDAC) eraser proteins of which sirtuins (SIRT) are a subset, and KAc reader proteins including bromodomain (BRD) proteins. HATs use acetyl coenzyme A (AcCoA) to generate KAc with coenzyme A (CoA) as a side product. Sirtuins use nicotinamide adenine dinucleotide (NAD^+) as a coenzyme to generate unlabelled K with nicotinamide and 2'-O-acetyl-ADP ribose as side products.

Lysine acetylation can occur both non-enzymatically and enzymatically. The levels of non-enzymatic acetylation are dependent on pH and AcCoA levels and are low in most organelles other than mitochondria.⁶⁵ In most cases, acetylation levels are regulated by enzymes (Figure 1.3). Writers of lysine acetylation are referred to as lysine/histone acetyl transferases (KATs/HATs) and erasers of lysine acetylation are referred to as lysine/histone deacetylases (KDACs/HDACs).⁵⁶ There are also lysine acetylation reader

domains including bromodomains (BRDs) which are the most studied lysine acetylation readers,⁵⁶ but also YEATs (Yaf9, ENL, AF9, Taf14, and Sas5) domains,⁶⁶ plant homeodomain (PHD) domains,⁶⁷ and one identified double pleckstrin-homology (PH) domain.⁶⁸

HATs and HDACs are complementary enzyme classes with opposite effects on lysine acetylation, which is reflected in their largely opposite biological roles. For example, histone acetylation is generally correlated to an increase in transcriptional activity, HAT activity being correlated with active transcription sites and HDAC activity correlated with inactive transcription sites.⁶⁹ HATs and HDACs regulate transcriptional activity through acetylation/deacetylation of histones,⁷⁰ but also through direct acetylation/deacetylation of transcription factors,^{56,71} and through regulation of chromatin structure.⁷² These biological effects can be partly attributed to non-specific cumulative effects of histone acetylation, with many HATs and HDACs showing a wide range of site selectivity.⁷³ Biological effects linked to specific acetylation/deacetylation events are also reported (*e.g.* at H4K16Ac, H3K9Ac/18Ac/27Ac/56Ac).^{69,73,74} Substrate selectivity for HATs and HDACs cannot be tied solely to consensus sequences.⁷⁵ HATs and HDACs have been shown to form multiprotein complexes with various epigenetic regulators,^{53,76} and to interact with other epigenetic effector enzymes which regulate alternative PTMs.⁷⁶ Combinatorial PTM crosstalk involving acetylation PTMs, HATs and HDACs has also been widely reported (Figure 1.2).^{37,42,62,77–83} Both multiprotein complex formation and PTM crosstalk can affect substrate selectivity of HATs and HDACs.

Because of their wide range of biological roles, significant efforts have been made in characterising HAT and HDAC proteins, and in developing probes and inhibitors targeting HATs and HDACs for research and medical purposes.^{56,84–86}

1.1.3 Classification of HAT and HDAC epigenetic families

1.1.3.1 Classification of HAT enzymes

HAT proteins are classed into three sub-families based on structural similarities, catalytic mechanism and sequence conservation: GCN5 *N*-acetyltransferase (GNAT), MYST (named after the four founding HATs Morf, Ybf2/Sas3, Sas2 and Tip60) and an ‘orphan’ family containing HATs with diverse catalytic domain architecture (Figure 1.4).^{87–89} There is a shared conserved motif between HAT families of Arg/Gln-X-X-Gly-X-Gly/Ala which corresponds to an AcCoA binding site.⁹⁰ Earlier nomenclature divided HAT proteins based on their localisation in the nucleus (type A) or the cytoplasm (type B). This is outdated, with many HATs being found in both the nucleus and cytoplasm,⁵³ but protein names based on this nomenclature remain in use.

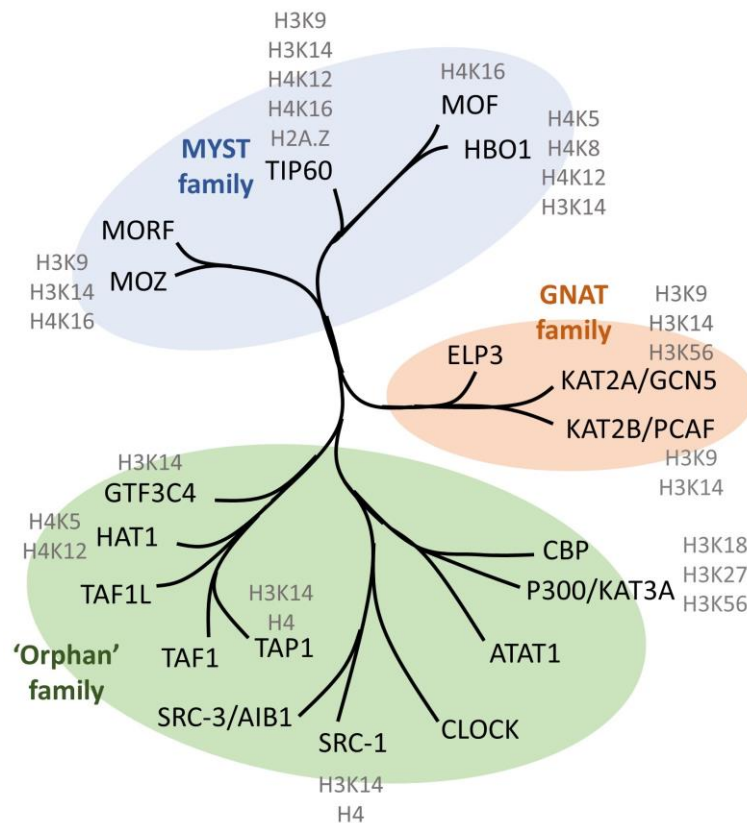


Figure 1.4: Histone acetyltransferase (HAT) phylogenetic tree.

Three families of HATs are represented: GNAT, MYST and Orphan proteins. Selected HAT proteins from each family and their histone substrate selectivity (when known) are shown.⁹¹ Figure reproduced from Di. Cerbo et. al.⁸⁷

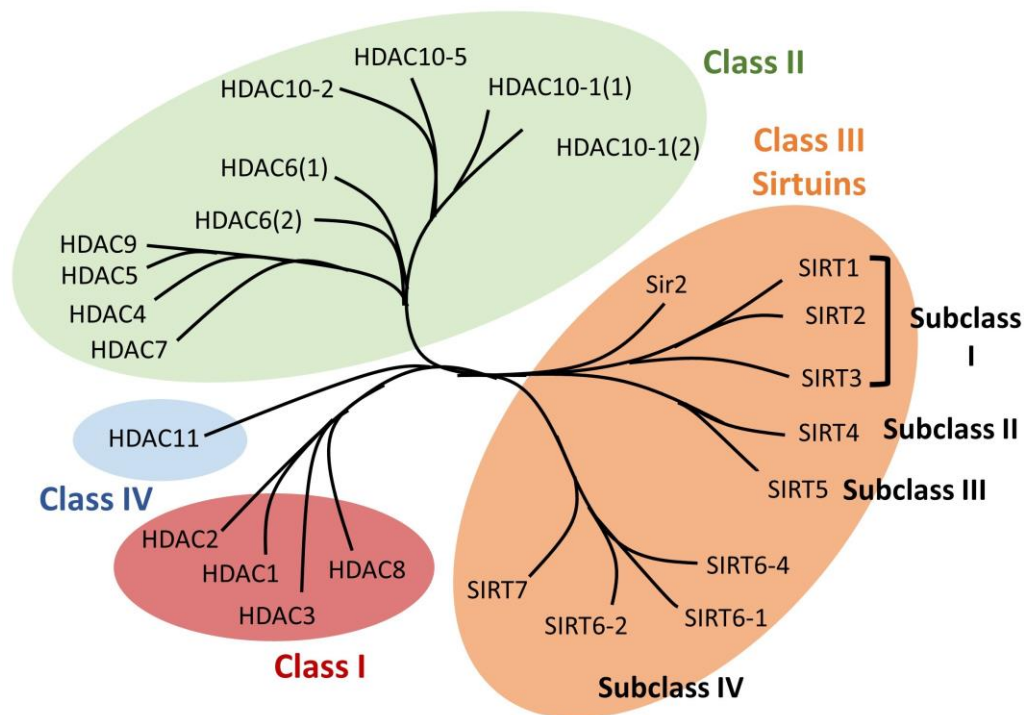


Figure 1.5: Histone deacetylase (HDAC) phylogenetic tree.

The four classes of HDACs are represented. Class III of HDACs are sirtuins, separated into subclasses I-IV. Figure adapted from Arrowsmith et. al.^{86,92,93}

1.1.3.2 Classification of HDAC enzymes

HDACs are separated into four different classes based on sequence alignment and structure similarities to previously characterised yeast proteins including silent information regulator 2 (ySir2, Figure 1.5). Class III HDACs are referred to as sirtuins (SIRT) and show sequence similarity to the ySir2 protein. Whilst class I-III HDACs belong to the arginase/deacetylase protein superfamily, sirtuins belong to the deoxyhypusine synthase like nicotinamide adenine dinucleotide (NAD⁺) or flavin adenine dinucleotide binding domain superfamily. Sirtuins are further divided into four subclasses based on sequence similarity and activity. All sirtuins share an NAD⁺ dependent catalytic mechanism, and a conserved core domain with common sequence motifs. However, not all sirtuins show deacetylation activity.^{76,86}

1.2 **PCAF**

1.2.1 **Structure of PCAF HAT and mechanism of acetylation**

p300/CBP associating factor (PCAF), also known as KAT2B, belongs to the KAT2/GNAT family of HATs (Figure 1.4). It was discovered in 1996 by sequence homology to the HAT domain of human and yeast general control nonderepressible 5 (hGCN5/yGCN5). Both proteins have a HAT domain and a BRD, with human PCAF (hPCAF) and hGCN5 showing a 75% sequence identity overall, with up to 85% identity in the region corresponding to the HAT and BRD domains.^{94–97} hPCAF also contains an *N*-terminal domain, which is only found in human, mouse and drosophila GCN5 (Figure 1.6a).^{96,98} PCAF and GCN5 are often treated as a single protein type due to a large overlap in structure, site selectivity, enzymatic activity and biological activity, but differences have been reported in the activities of PCAF and GCN5 (Section 1.2.3.6).^{97,99,100} PCAF was originally identified

through its binding to the HAT p300 (also known as KAT3A), where it was shown to compete for binding of adenoviral oncoprotein E1A with p300, therefore inhibiting cell division.⁹⁴ PCAF was reported to have acetylation activity both on nucleosome particles and H3/H4 core histones with a preference for H3.⁹⁴

PCAF contains an *N*-terminal region, a HAT domain (residues 493–653, PDB 1CM0)⁹⁵ and a *C*-terminal bromodomain (BRD, residues 725–819, PDB 1N72).¹⁰¹ Within the *N*-terminal region, there is an E3 ubiquitin ligase domain (residues 67–378 with catalytic activity in region 121–241, PDB 7BY1)¹⁰² a binding region for p300 and CREB binding protein (CBP), and a nuclear localisation signal (between residues 350–445).^{94,103} The domain architecture is shown in Figure 1.6a.

The crystal structure of PCAF HAT with bound AcCoA was published in 1999 (PDB 1CM0).⁹⁵ PCAF has a globular structure made up of α -helices and β -sheets, with a pronounced cleft in the middle of the structure in which AcCoA and the substrate bind (Figure 1.6b). Although no structure of PCAF with bound substrate has been published, the location of substrate binding is inferred by comparison with the crystal structure of tetrahymena GCN5 (tGCN5) with bound substrate and coenzyme (PDB 1PU9, Figure 1.6c).¹⁰⁴ AcCoA interacts with PCAF through >30 hydrogen bonds and Van der Waals interactions. These interactions are highly conserved between PCAF and GCN5 structures. Two less well conserved substrate interacting loops flank the cleft in the protein, which are thought to allow for structural rearrangement upon substrate binding, leading to a widening of the cleft (Figure 1.6b, indicated in orange).⁹⁵ These flexible loops may contribute to the lack of consensus sequence for substrate sites of PCAF HAT.

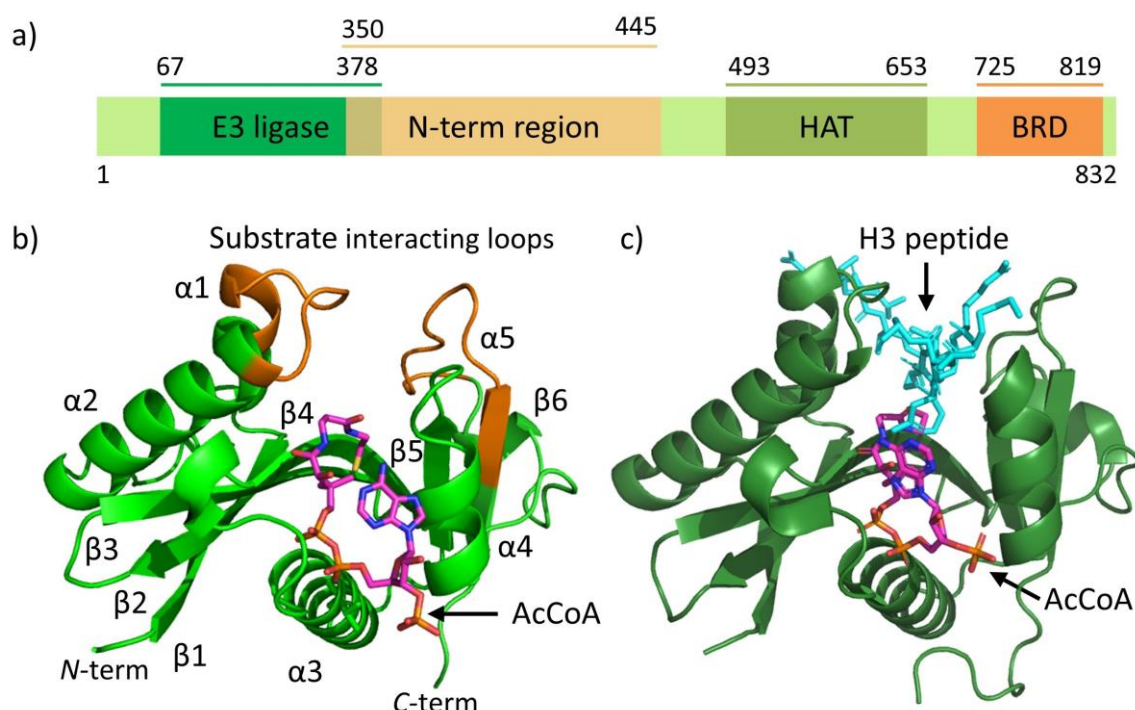


Figure 1.6: Composition of the PCAF protein and crystal structures of the HAT domains of PCAF and GCN5.

a) Map of the different domains of PCAF: E3 ligase (ubiquitin ligase domain), N-term region (contains a section which binds to p300 and a nuclear localisation signal), HAT (histone acetyl transferase domain), and BRD (bromodomain). **b)** Crystal Structure of PCAF with AcCoA. The topology of PCAF follows $\beta 1, \alpha 1, \alpha 2, \beta 2, \beta 3, \beta 4, \beta 5, \alpha 3, \beta 5, \alpha 4, \alpha 5, \beta 6$ structure elements, giving a globular fold with a cleft in the middle where AcCoA and the substrate bind. The substrate interacts with two flexible loops (orange) flanking the central cleft. Figure adapted from Clements et. al. PDB 1CM0.⁹⁵ **c)** Crystal structure of GCN5 with AcCoA and H3₅₋₂₄ peptide adapted from Clements et. al. PDB 1PU9.¹⁰⁵

PCAF HAT acetylation is reported to have an ordered sequential kinetic mechanism in which an intermediate ternary complex is formed between the protein, substrate and co-enzyme (Figure 1.7).^{95,106–108} These findings are supported by parallel studies with GCN5 crystal structures and kinetics reports.^{104,109,110} AcCoA first binds to PCAF HAT, followed by substrate binding. E570 is located in a hydrophobic patch, which contributes to raise its pKa to 7.5, allowing it to act as a general base towards the lysing ϵ -amine group on the substrate, making it more nucleophilic. The substrate amine then attacks the AcCoA carbonyl group forming a tetrahedral intermediate, which is the rate limiting step in the reaction. This intermediate collapses to release first the KAc product followed by CoA.^{95,106} C574 which is placed to form a hydrogen bond to AcCoA in the crystal

structure of PCAF/AcCoA is thought to stabilise the tetrahedral intermediate through a hydrogen bond to its backbone amide.⁹⁵ This theory is supported by GCN5 models.¹¹⁰ AcCoA was found to bind with a K_d of $0.64 \pm 0.12 \mu\text{M}$ to PCAF, and a 21-mer histone peptide comprising of residues 1-21 (H3₁₋₂₁) with an added C-terminal cysteine residue was found to bind with an apparent K_d of $116 \pm 17 \mu\text{M}$ to the PCAF/AcCoA complex.¹⁰⁶

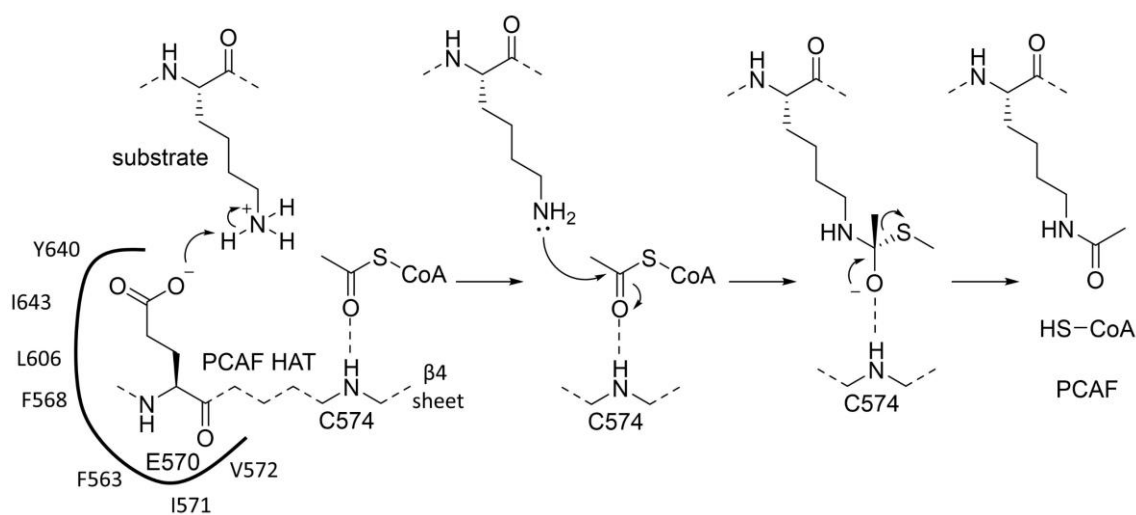


Figure 1.7: Mechanism of lysine substrate acetylation by PCAF HAT.

E570, made more basic by its position in a hydrophobic patch, deprotonates the substrate lysine residue. AcCoA then undergoes nucleophilic attack by the substrate's lysine amine group, forming a ternary complex. The KAc product and CoA are then released from the PCAF active site in an ordered mechanism. The backbone amide of C574 has a hydrogen bond interaction with AcCoA in the PCAF-AcCoA crystal structure and is theorised to stabilise the tetrahedral reaction intermediate. Figure adapted from Clements et. al.⁹⁵

HATs are known to participate in multiprotein complexes in a cellular context.⁵³ Whilst for some HAT families, complex formation is necessary for catalytic activity this is not the case for the GNAT family or for homologous HATs p300 and CREBP binding protein (CBP). PCAF, GCN5, p300 and CBP all have independent catalytic activity.¹⁰⁷ However, PCAF has been reported as part of multiple protein complexes including the Ada-Two-A-containing complex (ATAC),^{97,111} and the Spt-Ada-GCN5 acetyltransferase (SAGA) complex.^{97,112} In these cases, complex formation is thought to affect substrate selectivity and biological activity.^{107,113}

1.2.2 PCAF substrate selectivity

PCAF has been reported to catalyse the acetylation of histones H3 and H4, whilst GCN5 has been reported to catalyse the acetylation of H2a, H2b, H3 and H4 with a strong preference for H3 and H4.^{114–117} Preferred PCAF acetylation sites are reported as H3 lysine 14 (H3K14) > H3K9 > H4K8.^{50,74,90,94,97,118,119} Developing a clear understanding of the substrate scope exhibited by HAT enzymes can be complex as it can vary depending on factors such as AcCoA concentration,¹²⁰ protein complex formation,¹²¹ interaction between protein domains,¹²² and PTMs on either the enzyme¹²³ or the substrate.⁴⁷ *In vitro* studies are models for enzyme activity, which cannot fully replicate the complexity of a cellular setting. This must be taken into account when analysing results. There is also a significant amount of overlap between substrate selectivity of HAT enzymes, which may complicate analysis in cellular and animal models.^{119,124} In mouse models, deletion of both PCAF and GCN5 was required to affect levels of H3K9Ac, showing their redundant roles. However, deletion of both PCAF and GCN5 did not affect H3K14Ac levels, indicating that other HATs also affect this PTM *in vivo* (e.g. HATs E1A, MOZ, MORF). HDAC activity which counteracts HAT activity also affects analysis of HAT activity in cells and animal models.^{74,119,124}

PCAF has been reported to catalyse the acetylation of a variety of non-histone proteins including but not limited to: hypoxia inducible factor 1 α (HIF1 α , Chapter 4),¹²⁵ tumour protein p53,¹²⁶ phosphatase PTEN,¹²⁷ the RPA1 complex,¹²⁸ kinase cdk2,¹²⁹ protein EB1,¹³⁰ protein cyclinA,¹³¹ receptor ER α ,¹³² enzyme ALDH1A1,¹³³ enzyme YARS,¹³⁴ kinase MKL1,¹³⁵ kinase PKM2,¹³⁶ and transcription factors TFIIE and TFIIIF,¹²⁸ USF,¹³⁷ KLF13,¹³⁸

KLF8,¹³⁹ and HOXA10.¹⁴⁰ Non-histone substrates for PCAF and other HATs/HDACs have been reviewed in Narita *et. al.*²⁰

PCAF also has autoacetylation activity, which can affect its stability and activity in cellular and animal models. PCAF autoacetylation is not reported *in vitro*, and only occurs with the full-length protein containing the *N*-terminal region, which was shown to be essential for autoacetylation activity.¹²³

1.2.3 Biological role of PCAF

1.2.3.1 PCAF in cancer

PCAF has been extensively studied in relation to cancer and has a dual role either as a tumour promoter or a tumour repressor, depending on biological context. PCAF has been linked to brain, liver, gastric, lung, central nervous system, prostate, colon, and bladder cancer, taking part in numerous distinct signalling cascades, as shown by selected examples listed in Table 1.1. HATs are increasingly being targeted in development of cancer therapies.¹⁴¹

Role	Biological function	Ref
Tumour suppressor activity of PCAF in Cancer	Knockdown of PCAF in glioma cells was linked to reduced tumour proliferation and increased apoptosis through regulation of transcription mediated by the Hh–Gli signaling pathway.	142
	PCAF catalyses the acetylation of GLI1 leading to modulation of the GLI1/BCL2/BAX axis for increased apoptosis in hepatocellular carcinoma.	143
	PCAF acetylation of GLI1 is linked to decreased metastasis of hepatocellular carcinoma.	144
	PCAF catalyses the acetylation of H4 and is correlated with inactivation of AKT signalling leading to increased apoptosis in hepatocellular carcinoma.	145
	PCAF catalyses the acetylation of Ku70 leading to an increase in apoptosis by sequestering Bax, affecting mitochondrial activity.	146
	PCAF protein and mRNA is downregulated in gastric tumour cells. PCAF acts as a tumour suppressor through a PCAF-AE1-p16-CDK4 signalling cascade.	147
Tumour promoter activity of PCAF in Cancer	PCAF is overexpressed in cisplatin-resistant cells and gives increased resistance to apoptosis through increased expression of E2F1.	148
	PCAF acetylation of STAT3 was correlated with tumour proliferation in association with the C5a-C5aR axis of inflammation.	149
	PCAF has acetylation activity against PTEN, reducing its ability to block cell cycle arrest to inhibit tumour formation.	127
	PCAF and GCN5 acetylate the oncoprotein c-MYC, leading to decreased c-MYC degradation.	150
	Combined activity of PCAF acetylation of the methyl transferase EZH2 with EZH2 phosphorylation leads to increased EZH2 stability, increased gene suppression, increased cell migration and invasion of lung cancer, and poor patient prognosis.	151
	PCAF catalyses the acetylation of AKT1 leading to increased proliferation of human glioblastoma cells.	152
	PCAF is overexpressed in central nervous system tumours and Wilm's tumours of paediatric patients compared with benign tumours.	153
	Ubiquitination and degradation of ACLY is blocked by acetylation with PCAF. This promotes lipid synthesis, and tumour proliferation.	154
	PCAF is upregulated in prostate cancer cells, affecting androgen receptor signalling and leading to increased cell growth.	155
	PCAF catalyses the acetylation of β -catenin, increasing its stability. Knockdown of PCAF decreased β -catenin acetylation in colon cancer cells, inhibiting cell migration, and repressing tumour growth in mouse xenograft models.	156
	PCAF is overexpressed in chemotherapy resistant urothelial cancer cells. PCAF knockdown reduced Y-B1 expression and rendered cells sensitive to chemotherapy, decreasing cell growth and invasiveness.	157

Table 1.1: Selected examples of reported biological activity of PCAF relating to cancer.

PCAF has activity both as a tumour promoter and a tumour suppressor, depending on biological context.

1.2.3.3 PCAF in infectious diseases

Both the HAT and the BRD of PCAF are important in the regulation of HIV type 1 trans-activator protein (TAT1), necessary for replication of HIV. The TAT1 protein is acetylated by PCAF HAT at K28, promoting binding of TAT1 to the CDK9/P-TEFb kinase, and is acetylated by p300 at K50, promoting dissociation of Tat from trans-activation response (TAR) RNA.¹⁵⁸ PCAF BRD binds to K50Ac (acetylated by p300), competing with TAR RNA for binding.¹⁵⁹ Small molecule inhibitors of PCAF BRD have been designed which inhibit HIV replication by promoting TAR RNA binding to the TAT1 protein instead of PCAF BRD.¹⁶⁰ This shows that PCAF HAT and PCAF BRD can have different activities towards the same target, and it is therefore important to consider the role of both domains. In addition to HIV specific activity, PCAF and GCN5 are reported to negatively regulate interferon- β production, compromising innate immune signalling systems.¹⁶¹ PCAF and GCN5 were also reported to catalyse the acetylation of the SARS-CoV-2 nucleocapsid protein.¹⁶²

1.2.3.4 PCAF in the brain

Chromatin modifying enzymes are necessary for regulating hippocampal-dependent memories, and PCAF has been shown to particularly influence the formation of fear-extinction associated memories.¹⁶³ PCAF knockout mice showed altered stress responses and memory capacity throughout their lifetime, with different stages of memory formation affected at 2 or 6 months.¹⁶⁴ In addition it was reported that during memory formation stages, H2b and H4 were increasingly acetylated, which correlated with an increase in expression of PCAF and p300/CBP.¹⁶⁵ These studies show that PCAF has a role in brain development and memory formation, and it has also been shown to

affect neurodegenerative diseases. PCAF has been studied in the context of Alzheimer's disease, with PCAF knockout mice showing decreased sensitivity to β -amyloid toxicity.¹⁶⁶ Httex1p, a domain of the huntingtin protein responsible for Huntington's disease, binds to and inhibits acetylation activity of PCAF and p300/CBP, leading to hypoacetylation of H3 and H4 in cellular models.¹⁶⁷ HATs including PCAF are now being explored as possible targets for therapeutic modulation of neurodegenerative diseases.^{168,169}

1.2.3.5 PCAF in cardiovascular disease

In the context of heart disease, PCAF has been shown to catalyse the acetylation of connexin 43, which is affected in Duchenne cardiomyopathy.¹⁷⁰ PCAF is highly involved in the control of inflammation and vascularisation. It has been shown to catalyse the acetylation of HIF1 α which is involved in tumour angiogenesis.¹²⁵ PCAF HAT activity enhances the activity of nuclear factor- κ B (NF- κ B) a transcription factor which is key for inflammatory response of cells.¹⁷¹ PCAF is a regulator of arteriogenesis, with PCAF deficiency reducing recovery from ischemia in mice studies, and decreasing inflammatory response of leukocytes and vascular cells.¹⁷² PCAF is also involved in the development of intimal hyperplasia.¹⁷²

PCAF is also implicated in the regulation of type 2 diabetes through acetylation of PGC-1 α at K328 and K450. This suppresses transcriptional activity of PGC-1 α by increasing proteasomal degradation. As PGC-1 α is involved in the regulation of energy metabolism, gluconeogenesis and increased glucose output, degradation of PGC-1 α leads to improved glucose homeostasis and increased insulin sensitivity.¹⁷³ In addition to regulating hepatic gluconeogenesis through modulation of PGC-1 α , PCAF has been

linked to multiple signalling systems affecting liver inflammation, hepatic insulin resistance and hepatic metabolic syndrome.¹⁷⁴

1.2.3.6 Interconnected activities of PCAF and GCN5

PCAF and GCN5 are homologous enzymes with a large overlap in structure, biochemical selectivity, and some overlap in biological activity (Sections 1.2.1,1.2.2). However, in many cases, the biological role of these proteins differs, making it necessary to be able to target one enzyme selectively over the other in cell/animal models and in therapeutic modulation.

In mice, PCAF and GCN5 were shown to be differentially expressed in different tissues with GCN5 more prevalent in brain, thymus, spleen, testis, and embryonic tissue and PCAF more prevalent in heart, liver, kidney, and skeletal muscle.⁹⁶ This suggests that PCAF and GCN5 have context dependent roles. In PCAF knockdown mice, GCN5 is overexpressed in tissues where PCAF is typically prevalent, indicating that GCN5 can compensate for PCAF loss. GCN5 knockout mice die in the first 9 days of gestation, with PCAF only strongly expressed from day 12.5 in unmodified embryos. Double knockout mice (PCAF and GCN5) died earlier than GCN5 knockout mice. This indicates that GCN5 and PCAF may both have a role in embryogenesis but that GCN5 plays critical roles which cannot be compensated for by PCAF.¹⁰⁰ These reports suggest that PCAF and GCN5 can have both overlapping and separate roles in human biology, in development and in general cell function.

PCAF and GCN5 are reported to have the same substrate selectivity profiles towards histone proteins, acetylating H3K14 and H3K9 preferentially.⁷⁴ Although PCAF is

identified as acetylating 68 non-histone proteins in Narita *et. al.* GCN5 is identified as acetylating only 14 non-histone proteins, with no overlap between non-histone PCAF and GCN5 substrates.²⁰ A later review by Downey identified 14 non-histone proteins affected by both PCAF and GCN5, with many more proteins identified solely as PCAF substrates.¹⁷⁵ This small overlap in activity may be partially due to lack of systematic cross-checking between PCAF and GCN5 activities. However, it is clear that PCAF and GCN5 often have diverging biological roles, and therefore also have diverging substrate selectivity, as illustrated in the following examples.

hPCAF and hGCN5 can both take part in SAGA and ATAC protein complexes (involved in transcriptional regulation), in a mutually exclusive manner based on cellular context.^{112,176,177} GCN5 and PCAF were shown to have opposing effects on telomere maintenance through knockdown studies. GCN5 interacts directly with USP22 at telomeres, opposing telomere recombination, whilst PCAF is not directly located at telomeres but is involved in a signalling cascade which leads to increased telomere recombination.¹⁷⁸ In breast cancer cell lines, combined RNAi downregulation of PCAF and GCN5 led to increased expression of MDR1, which increases drug resistance in breast cancer. Downregulation of PCAF alone did not affect MDR1 expression, indicating that GCN5 downregulation is necessary to achieve this effect.¹⁷⁹ PCAF catalyses the acetylation of ACLY, a protein important in metabolism regulation, promoting tumour growth in lung cancer, whilst GCN5 has no acetylation activity against ACLY.¹⁵⁴ The nucleoprotein of influenza virus A is acetylated by PCAF at K31 and GCN5 at K90. PCAF and GCN5 expression have opposite effects on the influenza virus, with RNAi silencing

of PCAF leading to increased viral polymerase activity and RNAi silencing of GCN5 leading to decreased viral polymerase activity.¹⁸⁰

1.3 SIRT1

Human SIRT1-7 isoforms of the SIRT family of NAD⁺ dependent HDACs vary in their sub-cellular location, target specificity and biological functions. SIRT1 is mainly found in the nucleus but can relocate to the cytoplasm.¹⁸¹ SIRT1 has a very broad range of substrate specificity for histone and non-histone proteins.^{20,182} The main histone substrates of SIRT1 are H3K9Ac, H3K14Ac, H4K16Ac,^{86,182} with lower levels of deacetylation activity reported at H4K8 and H4K12, H3K56.^{182–185} Non-histone substrates include p53,¹⁸⁶ HIF1 α ,¹²⁵ p300¹⁸⁵ and PCAF¹⁸⁷ with over 107 substrates identified as of November 2019.²⁰ Like PCAF and GCN5, SIRT1 activity can be modulated by biological context. SIRT1 activity is location dependent (cytoplasm or nucleus),^{181,188,189} and can be modified by complex formation.^{190–192} In part due to its large range of substrates, SIRT1 has been implicated in many biological functions and diseases including metabolic functions, circadian rhythms, the immune system, neurodegeneration and cancer (both tumour suppressor and tumorigenic activities identified, context dependent).^{84,193} A few of the reported biological functions of SIRT1 are listed in Table 1.2.

Disease state	Reported biological role	Ref
Metabolism Diabetes	SIRT1 orchestrates changes in gene expression affecting nutrient handling, lipid storage and lipid oxidation. SIRT1 also affects insulin resistance.	194
Metabolism	SIRT1 deacetylation of LXR α reduces expression of LXR α targets controlling cholesterol levels and lipid metabolism, leading to an increased risk of atherosclerosis.	195
Circadian rhythms	SIRT1 is involved in circadian regulation of NAD ⁺ levels in a feedback loop involving CLOCK HAT.	196
Cardiovascular disease	SIRT1 single nucleotide polymorphism variants correlate with low rate of cardiovascular disease.	197
Diabetes	Increased mortality for subjects with SIRT1 haplotype 1, exacerbated by low vitamin B3 intake (VITB3 is the precursor for NAD ⁺).	198
Neurodegeneration	SIRT1 catalyses the deacetylation of phosphorylated Tau protein which is involved in dementia, Parkinson's, and Alzheimer's diseases.	199
Neurodegeneration	H4K16 deacetylation through SIRT1 activity is correlated with decreased ROCK1 expression affecting calorie restriction influence on Alzheimer's type amyloid neuropathology.	200
Neuronal development	SIRT1 translocates to the nucleus modulating neuronal differentiation.	188
Immunity Inflammation	SIRT1 catalyses the deacetylation of NF- κ B, leading to increased transcription of pro-inflammatory genes in macrophages.	201
Chronic pulmonary disease	Monocyte cells treated with cigarette smoke show decreased SIRT1 levels associated with increased acetylation of NF- κ B, increasing inflammation.	202
Immunity	SIRT1 catalytic activity negatively affects COX-2 expression leading to amelioration of macrophage functions.	203
Prostate cancer	SIRT1 activity is linked to an increase in tumour survival.	204
Prostate cancer	SIRT1 is overexpressed in prostate cancer cells and tissues. SIRT1 catalyses the deacetylation of FOXO1 contributing to cancer propagation.	205
Hepatocellular carcinoma	SIRT1 activity is linked to telomere maintenance and tumour cell growth.	206
Development and cancer	SIRT1 null mouse embryos have a low survival rate. Increased SIRT1 activity was correlated with low tumour formation.	207
Colon cancer	SIRT1 overexpression is correlated with lower rates of tumour formation. SIRT1 inhibition combined with chemotherapy drugs leads to higher rates of apoptosis.	208

Table 1.2: Roles of SIRT1 in biological functions and diseases.

The mechanism of SIRT1 deacetylation goes through a sequential binding of NAD^+ and substrate. Nucleophilic attack by the KAc acetamide oxygen of the substrate leads to formation of a C1'-O-alkylamidate intermediate, releasing nicotinamide. An active site histidine facilitates nucleophilic attack by the 2'-hydroxy group of NAD^+ ribose, forming a 1',2'-cyclic intermediate. An active site H_2O reacts with the intermediate leading to release of deacetylated lysine and 2'-O-acetyl-ADP ribose. 2'-O-acetyl-ADP ribose and 3'-O-acetyl-ADP ribose are rapidly interconverted under physiological aqueous conditions (Figure 1.8).^{86,209–211} The nicotinamide side product is often used as a tool in assays for quantification of enzyme activity, either through enzyme coupled assays or radiolabelling.^{212,213}

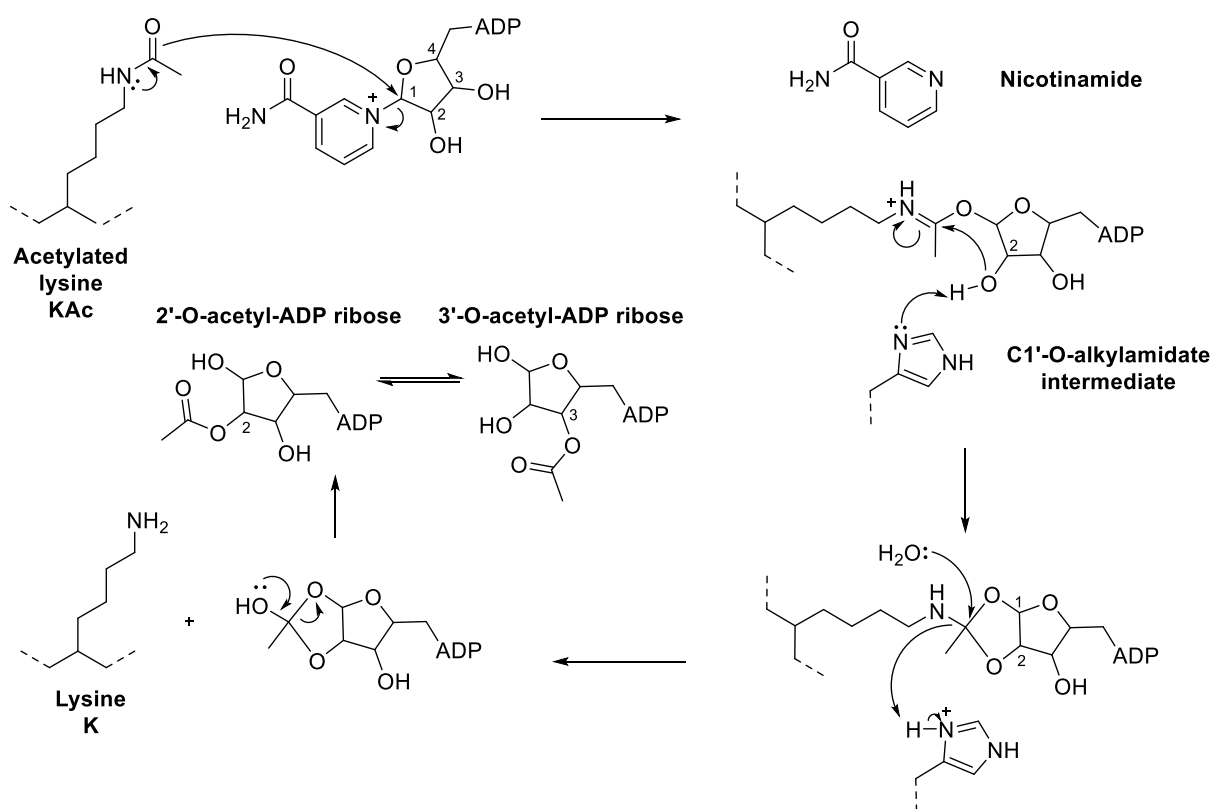


Figure 1.8: Mechanism of NAD^+ mediated deacetylation of substrate KAc catalysed by sirtuins.

1.4 Complementary roles of PCAF and SIRT1

Development of cancers can in some cases be traced back to an imbalance in acetylation levels. A large general decrease in acetylation leads to decreased expression of tumour suppressor genes, and a large increase in acetylation leads to increased expression of proto-oncogenes. These changes in KAc levels can be due to an imbalance in HAT and HDAC activity.⁸⁷ Thus, understanding of the roles of both HATs and HDACs regulating the KAc PTMs are needed to address dysregulation in KAc levels.

1.4.1 PCAF and SIRT1 activity for non-histone proteins

Because neither PCAF or SIRT1 have a clear consensus sequence for substrate sites, and both show a large substrate scope, the regulatory mechanisms that govern the activity of these proteins is unclear. It has been suggested that sub-cellular location is a factor in defining activity, as both proteins are found in the nucleus and cytoplasm.⁸⁴ PCAF and SIRT1 have significant overlap in non-histone protein substrates: p53 and HIF1 α , discussed below; Akt and PDK1 are both acetylated by PCAF and deacetylated by SIRT1, with deacetylation linked to increased tumorigenesis and cardiac hypertrophy;²¹⁴ ER α involved in metabolic signalling is acetylated by PCAF and deacetylated by SIRT1 and HDAC8, with increased acetylation correlated with increased DNA binding by ER α ;¹³² Lin28, involved in stem cell differentiation and developmental timing is acetylated by PCAF and deacetylated by SIRT1.²¹⁵ A wider list of PCAF and SIRT1 overlap in substrate selectivity for non-histone proteins can be found in Narita *et. al.*²⁰ The overlap in activity between these two proteins is not limited to shared substrates, as SIRT1 has been shown to directly catalyse the deacetylation of PCAF.¹⁸⁷

The mechanisms by which PCAF and SIRT1 regulate signalling systems linked to tumour formation have been extensively studied, especially in relation to the interconnected tumour suppressor protein p53 and the hypoxia inducible factor 1 α (HIF1 α), which is essential to cellular hypoxic response. For p53, PCAF and SIRT1 target different sites, whereas for HIF1 α they target the same site, suggesting a closer relationship between the activities of the two proteins in this regulatory mechanism.

Acetylation activates p53, increasing DNA binding of p53. PCAF catalyses the acetylation of K320 and p300 catalyses the acetylation of K373/K382 of p53. In response to DNA damaging chemicals, p53 acetylation by PCAF and p300 increases, as part of a feedback loop.¹²⁶ In addition, it was shown that p53 phosphorylation in response to DNA damage affected the activity of PCAF towards p53, in an example of PTM crosstalk affecting biological activity.^{80,83} PCAF acetylation of p53 promotes target gene activation in breast cancer cell lines. In these cell lines, PCAF itself is upregulated by p53 in a positive feedback loop.²¹⁶ SIRT1 was shown to catalyse the deacetylation of p53 at K382, with increased acetylation correlated with increase in cellular apoptosis and radiosensitivity.¹⁸⁶ Small molecule inhibitors of SIRT1 showed inhibition of SIRT1 deacetylation activity for p53.^{217,218}

Hypoxia affects the acetylation activity of PCAF towards both p53 and HIF1 α . In hypoxic conditions, PCAF acetylation of p53 at K320 decreases, affecting the ability of p53 to bind to gene promoters or cause cell apoptosis. By contrast, hypoxic conditions lead to acetylation of HIF1 α by PCAF, affecting the stability of HIF1 α and its downstream effects on cell cycle arrest and apoptosis.²¹⁹ SIRT1 was reported to reduce levels of HIF1 α acetylation under hypoxic conditions affecting transcriptional activity of HIF1 α .²²⁰ It was

subsequently shown that PCAF and SIRT1 both affect K674 of HIF1 α . SIRT1 activity against HIF1 α was shown to be downregulated in hypoxia due to low NAD⁺ levels, and upregulated through redox mediated blocking of glycolysis. SIRT1 decreased HIF1 α mediated tumour formation and angiogenesis in hypoxic mouse tissues.¹²⁵ The signalling systems of p53 and HIF1 α have been closely linked, with the effects of p53 and HIF1 α on apoptosis extensively studied.²²¹ The activities of PCAF and SIRT1 towards p53 and HIF1 α are one of the many links established between the two proteins.

1.4.2 Histone protein acetylation sites affected by PCAF and SIRT1

PCAF and SIRT1 have overlapping substrate selectivity towards histones (H3K9/H3K9Ac and H3K14/H3K14Ac).^{90,11986,182} It has been demonstrated that histone acetylation plays an important role in regulating gene expression.^{18,222} The specific role of H3K9 and H3K14 acetylation have been studied in various contexts.

H3K9 and H3K14 are both important sites of acetylation in many biological settings. In breast cancer cell lines, resistance to chemotherapy has been linked to an increase in H3K9 acetylation in the region of the MDR1 gene, and an increased level of H3K14, H3K4, H4K8, and H4K12 acetylation globally.¹⁷⁹ H3K9Ac is often found along with H3K14Me3 at active gene promoters, with H3K9Ac necessary to initiate the elongation step of polymerase II.²²³ Chromosome condensation is influenced by PTMs at H3K9, with a switch from acetylation to methylation necessary for full chromosome condensation.²²⁴ Successful regulation of the H3K9Ac to Me3 switch is also necessary to prevent formation of liver cancer.²²⁵ H3K14Ac was shown to modulate tissue specific gene expression by affecting chromatin remodelling.^{226,227} H3K14Ac levels at gene promoters was reported to affect stress response in cells.²²⁸ H3K14Ac is recognised by three

bromodomains of the protein polybromo-1, with this binding interaction necessary for its tumour suppressor role.²²⁹

In addition to having separate biological roles, combination of H3K9Ac and H3K14Ac can influence cell signalling. H3K9Ac and H3K14Ac are often found together genome wide, particularly at bivalent promoters and active enhancers. However, H3K14Ac was also found alone at inactive gene promoters, suggesting different roles for H3K14Ac alone or in combination with H3K9Ac.²³⁰ H3K9Ac and H3K14Ac were found in combination at the promoter of the LGALS9 gene associated with cervical cancer.²³¹ Either alone or in combination, H3K9Ac and H3K14Ac are generally associated with active chromatin.²³²

The levels of H3K14Ac and H3K9Ac are regulated by HATs and HDACs, including but not limited to PCAF and SIRT1. Regulation of these marks can also be linked to crosstalk between regulatory systems and protein complexes.^{34,47,233}

1.5 Chemical probes for HATs and HDACs

1.5.1 Chemical probes: definition, utility and applications

Chemical probes are compounds (small molecules, cyclic peptides *etc.*) which allow for selective modulation of protein activity, facilitating the study of mechanistic and phenotypic roles of the target protein in the context of *in vitro*, cellular and *in vivo* study models. Chemical probes are often developed in parallel to drug candidates as they have different profiles: whilst drugs candidates are developed on the basis of activity, bioavailability and physiochemical properties, chemical probes are developed on the basis of potent binding, proven selectivity and understanding of mode of action, with

activity not necessarily required. Chemical probes are developed for research purposes with the characterisation data and compound available for general use.^{234–238}

There are two major ways to monitor the biological activity of proteins: use of chemical probes or use of genetic manipulations (*e.g.* protein deletion, or mutation). Both approaches can have off-target effects in cells and animal experiments. Protein deletion can be very disruptive towards cellular function, and selective protein mutation to render proteins inactive may affect protein structure, disrupting protein-protein interactions (PPIs). Genetic methods require a lot of biological manipulation to achieve the correct cell line for study. By contrast, chemical probes provide a more flexible solution, with timing of the administration and concentration easily modified. Multiple chemical probes can also be combined to look at crosstalk between proteins and signalling pathways. Chemical probes can however have off-target effects, so using a combination of genetic and chemical methods provides the best overview of target activity.^{234,235,239}

1.5.2 Chemical probes for PCAF, GCN5 and p300/CBP HATs

1.5.2.1 Overview of selected HAT chemical probes

Small molecule chemical probes have been developed for the different families of proteins involved in regulating acetylation including HATs, HDACs and BRDs.^{92,119,240} Selected representative examples of the compounds tested against PCAF, GCN5 and/or p300/CBP HATs are described below.

The chemical probes available for HATs largely suffer from a lack of selectivity and potency, with only a few examples released in recent years showing a marked

improvement. Existing HAT chemical probes have been developed using three different approaches: bi-substrate inhibitors have been developed by covalently linking mimics of HAT substrates and the AcCoA cofactor; natural products have shown anti-HAT activity and analogues of these have been developed, with improved physiochemical properties; lastly, virtual and physical high throughput screening (HTS) has been used, with a variety of small molecule compound libraries screened.

- Bi-substrate chemical probes

The first HAT inhibitors were bi-substrate inhibitors (Figure 1.9). Lys-CoA inhibits p300 (IC₅₀ 500 nM) selectively over PCAF. H3-CoA-20 inhibits PCAF (IC₅₀ 300 nM) which is selective for H3 but does not inhibit p300 which is selective for H4. Bi-substrate inhibitors with H3 sequences shorter than 20 residues did not inhibit PCAF, showing that interactions outside the active site are important for H3 to PCAF binding and selectivity.²⁴¹

- Natural product chemical probes and their derivatives

Amongst the natural products tested for HAT activity, anacardic acid, garcinol, curcumin, embelin and their derivatives were found to have HAT inhibitory activity (Figure 1.9).²⁴¹⁻²⁴⁴

Anacardic acid (AnA) is an inhibitor of PCAF, p300, Tip60 and KAT8 with reported IC₅₀ values ranging between 5-1000 µM depending on the enzyme, the histone materials used and the assay setup.^{119,246} Analogues of anacardic acid were developed either with improved selectivity for PCAF or Tip60.^{247,248} *In vitro*, AnA derivative CTPB is inactive against PCAF and increases the activity of p300. In cells it increases transcriptional

activity.²⁴² AnA derivative SPV-106 is cell permeable, unlike AnA and many of its derivatives. *In vitro* it is a p300/CBP inhibitor, is inactive against GCN5, and is a PCAF activator; in cells, it arrests the cell cycle.²⁴⁹ Further study of this class of compounds showed the length of the alkyl chain affects the selectivity and activity of the compounds.²⁵⁰

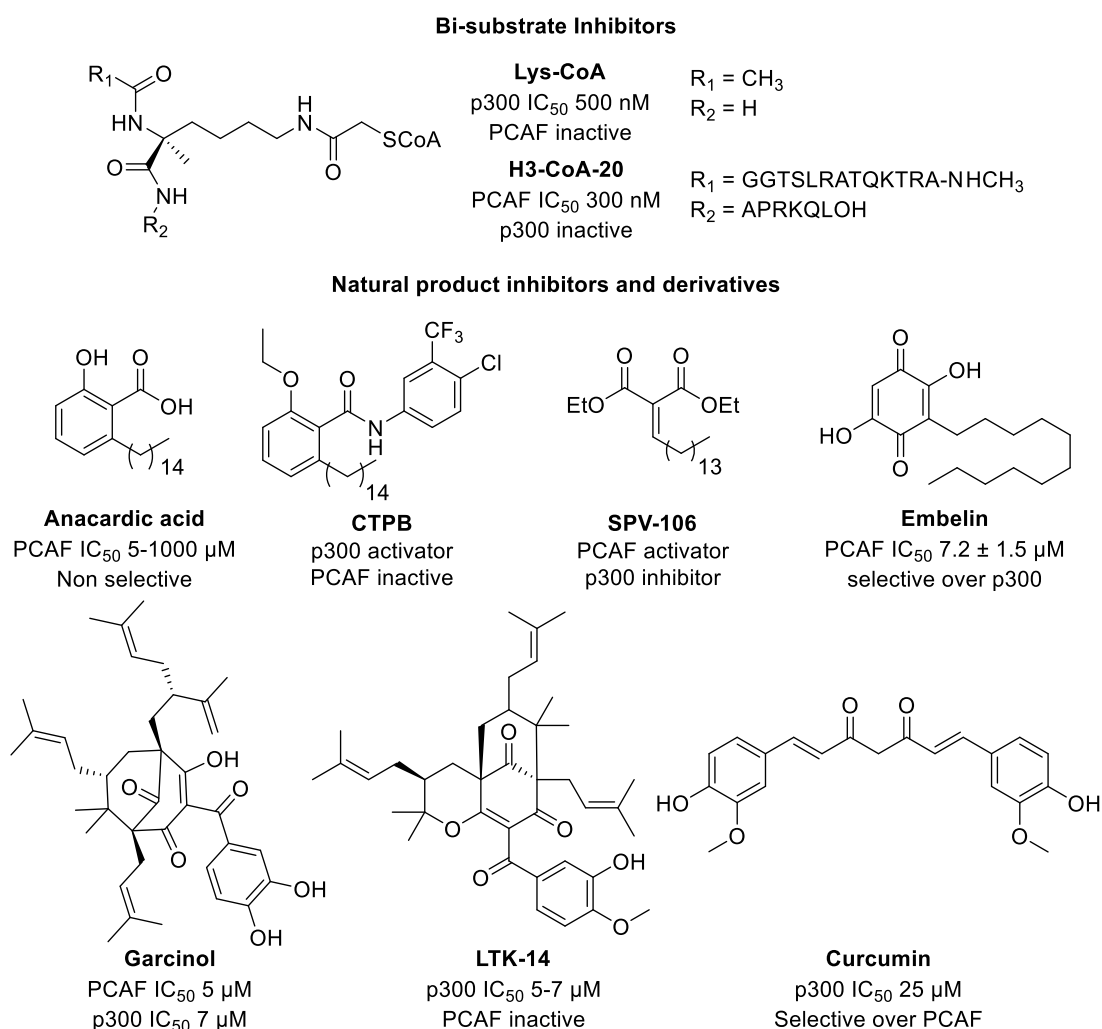


Figure 1.9: Bi-substrate and natural product chemical probes for HATs.

Embelin (Emb) is a non-competitive PCAF HAT inhibitor (IC₅₀ of 7.2 ± 1.5 μM) with selectivity over p300 and Tip60. It is active in cells, inhibiting H3K9 acetylation.²⁴⁵ The lipophilic tail is necessary for activity, though it renders it poorly water soluble.²⁵¹

Garcinol is active against PCAF (IC₅₀ 5 μ M) and p300 (IC₅₀ 7 μ M) *in vitro* and was shown to be competitive against histone protein substrates. In cells garcinol inhibited acetylation of H4 and H2a in the presence of HDAC inhibitors, and it induced apoptosis in cancer cell lines.²⁴³ Garcinol analogue LTK-14 was shown to be selective for p300 (IC₅₀ of 5-7 μ M) over PCAF, and less cytotoxic than garcinol.²⁵²

Curcumin inhibits p300 (IC₅₀ of 25 μ M), and is selective over PCAF, HDAC1 and methyl transferase 69a.²⁴⁴ However it has been shown to have inhibitory activity against an alternative panel of KATs, KDACs, methyltransferases and microRNAs.²⁵³ Some curcumin analogues have been developed and tested but were shown to also lack selectivity.¹¹⁹

The biggest issue with these natural product chemical probes and their derivatives is off target activity, which has been well recorded particularly for curcumin and embelin.^{253,254} They also suffer from poor cell permeability and low binding affinities.

- Small molecule chemical probes

High throughput screening has yielded HAT inhibitors with a variety of scaffolds (Figure 1.10). p300 is the most common target for HAT inhibitors as its structure differs to that of other HAT enzymes, increasing the chances of identifying selective compounds.⁸⁷⁻⁸⁹

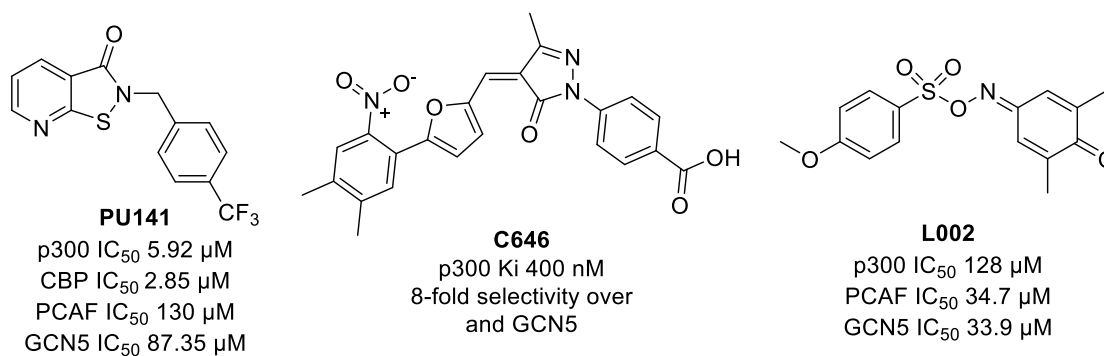
Isothiazolones were developed as covalent inhibitors for HATs as they showed activity against PCAF *in vitro* and cytotoxicity in cells. However, they were shown to react non-specifically with cysteine residues in other proteins, and the compounds initially identified were completely inactivated by incubation with BSA due to this off-target activity.²⁵⁵ Screening of isothiazolone containing compounds yielded compound PU141 selective for p300/CBP (IC₅₀ values of 5.92 μ M and 2.85 μ M respectively) over

PCAF/GCN5 (IC_{50} values of 130 μ M and 87.35 μ M respectively). PU141 was tested in a cysteine protease profiling assay, showing selective activity for HAT proteins over other cysteine containing proteins.²⁵⁶ PU141 treatment led to histone hypoacetylation in both cancer cell lines and in healthy mice.²⁵⁷

C646 was developed based on a virtual screening against p300 and was shown to inhibit p300 HAT activity against H4 peptides (K_i 400 nM). It showed a selectivity of >8 fold against the HATs AANAT, PCAF, GCN5, Sas, Moz and Rtt109. In cells, C646 treatment led to hypoacetylation and cell cycle arrest.²⁵⁸ It has been one of the most widely used chemical probes for obtaining cellular data due to its superior potency and selectivity compared to other available probes until 2017, but shows some reactivity towards off-target cysteine containing proteins and has intrinsic fluorescent activity which affects readings in fluorescent assays.^{259 119}

L002 is active against p300 (IC_{50} 128 μ M), PCAF (IC_{50} 34.7 μ M) and GCN5 (IC_{50} 33.9 μ M), but unreactive against other KAT families, HDACs and methyltransferases. It showed high toxicity against leukaemia, lymphoma and triple negative breast cancer (TNBC) but low cytotoxicity for other solid tumour cell lines. In mice, L002 led to a decrease in tumour growth and acetylation in a xenograft model.²⁶⁰

Small molecule inhibitors derived from HTS



Inhibitors selective for p300/CBP

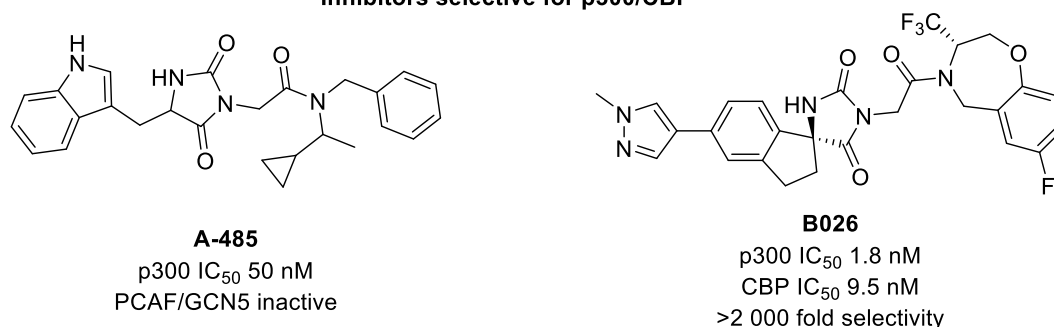


Figure 1.10: Small molecule chemical probes for HATs.

Chemical probes for HATs derived from high throughput screening and chemical probes with high selectivity and potency against p300/CBP HATs.

The most successful HAT inhibitors, A-485 and B026 were published in 2017 and 2020 respectively, both targeting p300 (Figure 1.10).^{261,262} A-485 is an inhibitor competing with AcCoA and is selective for p300 (IC₅₀ 50 nM), as shown by testing against a large panel of HATs, HDACs, BRDs, BETs and non-epigenetic proteins. It showed activity against dopamine and serotonin transporters but does not cross the blood-brain barrier. A-485 showed good activity against haematological tumours but not solid tumours and inhibited tumour growth in mice.²⁶¹ B026 shows improved binding towards p300 (IC₅₀ 1.8 nM) and CBP (IC₅₀ 9.5 nM) with 2 000-10 000 fold selectivity over other HAT proteins. B026 shows good bioavailability and physiochemical properties in mice, is active against a panel of cancer cell lines, and has superior tumour inhibition in mice over A-485.²⁶²

Compounds targeting GCN5, closest homologue to PCAF, have also been developed (Figure 1.11). Butyrolactone 3 is a selective inhibitor of GCN5 (IC_{50} 100 μ M) over CBP (IC_{50} 500 μ M), shows a reversible mode of action, but has not been tested in cells.²⁶³ Phenotypic screening in yeast identified CPTH2 and CPTH6 as binders for the HAT domain of GCN5.²⁶⁴ CPTH6 is selective for PCAF/GCN5 over p300/CBP and HDAC1, and was shown to induce hypoacetylation of H3 and apoptosis in leukaemia cell lines.²⁶⁵ MC1823 is presumed to act on GCN5 and shows decrease in H3 and α -tubulin acetylation in cells, but binding and selectivity for HAT enzymes was not established.²⁶⁶ More active chemical probes targeting GCN5 have been developed in recent years with DC-G16-11 and DC-HG24-01 identified through HTS. DC-G16-11 inhibition of GCN5 (IC_{50} 6.8 μ M, K_D 4.2 μ M) was shown by surface plasmon resonance (SPR) and nuclear magnetic resonance (NMR) *in vitro* with dose dependent cell growth inhibition and apoptosis, and reduction of H3K14 acetylation in cells.²⁶⁷ DC-HG24-01 also inhibited GCN5 (IC_{50} 3.1 ± 0.2 μ M, K_D 11.6 ± 1.3 μ M) as shown by AlphaScreen™ and SPR *in vitro* and had similar cellular activity to DC-G16-11.²⁶⁸ Selectivity over other HAT enzymes was not shown. Compound C14 was developed in virtual screening against *P. falciparum* GCN5 to be selective over hGCN5. It showed high affinity (IC_{50} 225 nM), but selectivity against other HATs was not demonstrated.²⁶⁹

HAT inhibitors targeting other HAT families (other than p300) are not mentioned here; for a wider chemical probe overview, available HAT inhibitors have been reviewed multiple times in recent years.^{85,119} Inhibitors have mainly been developed for the HATs Tip60 and MOF, and inhibitors include natural product derivatives which are selective for Tip60 and MOF over PCAF and p300.²⁴⁸

Inhibitors designed to target GCN5

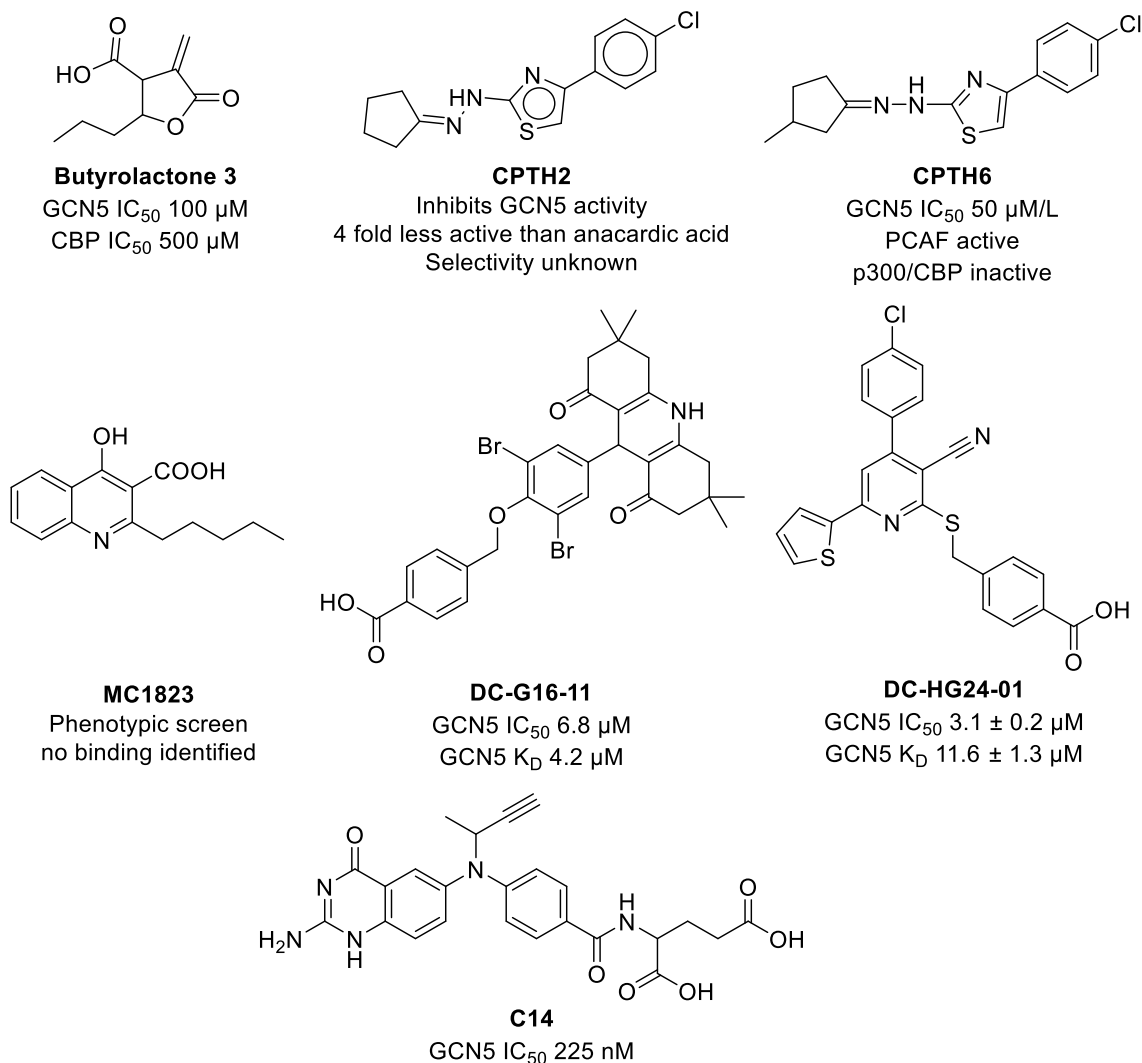


Figure 1.11: Small molecule chemical probes for GCN5 HAT.

1.5.2.2 Limitations of reported HAT chemical probes

The published HAT chemical probes either lack selectivity, or have not been widely tested for selectivity against a large panel of HATs and alternative epigenetic enzymes, except for A-485 and B026. Many of the natural products identified as chemical probes are implicated in various cellular processes in addition to modulating acetylation activity.²⁵³ Covalent inhibitors of HATs, particularly the isothiazolone derivatives, but also C646, suffer from non-specific activity towards cysteine containing proteins.^{119,255,259} Derivatives of natural products have improved specificity and activity profiles, and often have improved physiochemical properties (Figure 1.9, Figure 1.10,

Figure 1.11). As the compounds presented are the best existing chemical probes, they are often used in biological studies against PCAF, which can lead to misleading results due to lack of selectivity.

Significant steps in improvement of chemical probes available for HATs have been made with the publication of A-486 and B025, with inhibitory activity in nM region, and with a strong selectivity for p300/CBP over PCAF/GCN5 as well as other HAT and non-HAT enzymes.^{261,262} Potent inhibitors for GCN5 have also been published, but selectivity has not been demonstrated (Figure 1.11). Compounds targeting GCN5 are particularly interesting in the context of this work as GCN5 is a close homologue to PCAF. Significantly, there are no potent selective inhibitors published specifically targeting the HAT domain of PCAF. The conserved fold of the GNAT family of proteins to which PCAF and GCN5 belong and the high sequence homology between PCAF and GCN5 may make identification of selective chemical probes for PCAF difficult.²⁷⁰

1.5.3 Chemical probes for PCAF bromodomain

Understanding how the HAT and BRD domains act separately and co-operatively is essential to understanding the overall biological role of PCAF. Although HAT chemical probes for PCAF are sorely lacking, efforts to develop potent and selective PCAF BRD chemical probes have been successful.

Two small molecule chemical probes which selectively target the bromodomain of PCAF and GCN5 have been developed: *L*-Moses and GSK4027 (Figure 1.12).^{271,272} *L*-Moses is selective for PCAF (K_D 126 nM) and GCN5 (K_D 550 nM) over a panel of 45 bromodomains. GSK4027 is selective for PCAF/GCN5 (K_i 1.4 nM) over the BET family (≥ 18000 fold) and other BRDs (≥ 70 fold). Both probes showed good cell permeability, and were active in

NanoBRET target engagement assays with PCAF and H3.3. Both were reported alongside inactive enantiomers used as controls (*D*-Moses and GSK4028). However, neither were shown to affect PCAF or GCN5 HAT activity. GSK4027 was used to develop a proteolysis targeting chimera (PROTAC) chemical probe, to better explore the biological activity of PCAF. The modified GSK4027 binds to the BRD of PCAF, marking PCAF for proteasomal degradation. This approach was used to show that PCAF deficient macrophages produce fewer cytokines, an effect which is not mediated by the BRD of PCAF.²⁷³

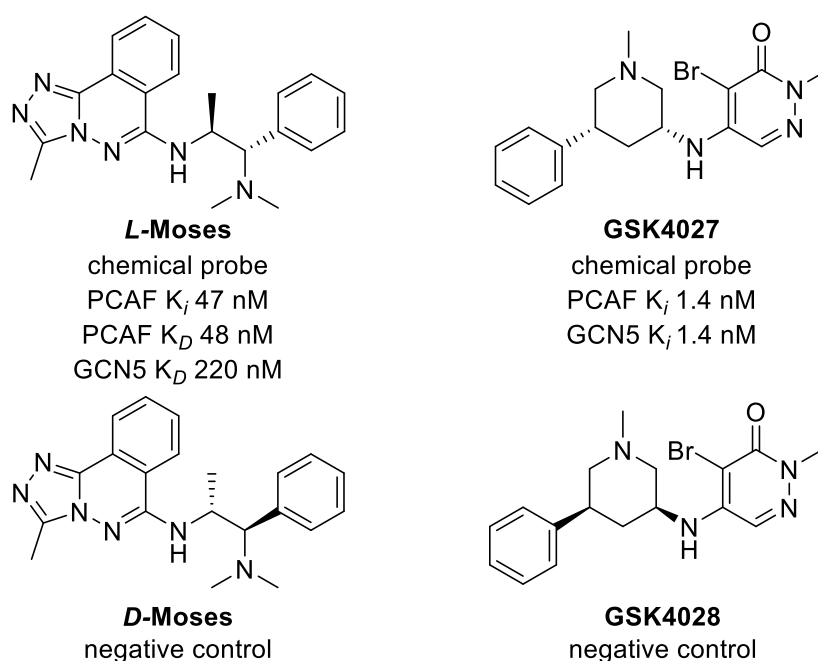


Figure 1.12: Reported chemical probes for PCAF BRD.

L-Moses and GSK4027 are both selective and potent chemical probes for PCAF BRD. *D*-Moses and GSK4028 are inactive controls.^{271,272}

1.5.4 Small molecule chemical probes targeting SIRT1

Both activating and inhibiting chemical probes have been developed for SIRT1 (Figure 1.13).^{56,84,86,274} The validity of the activation profiles of chemical probes developed has been questioned, due to off target effects and assay artifacts.^{274,275} Inhibitors of SIRT1 deacetylation activity include substrate and NAD⁺ mimics, and small molecules originating either from natural product or drug discovery platform screening.²⁷⁴ Sirtinol

was the first of an early class of inhibitors, discovered through phenotypic screening. It shows no selectivity between human and yeast sirtuins.²⁷⁶ Sirtinol derivative JGB1741 was developed with improved selectivity for SIRT1, and showed cellular activity.²⁷⁷ Affinity screening with a library of DNA encoded small molecules identified a family of thieno-pyrimidine inhibitors, which showed tight but non-specific binding to SIRT1-3.²⁷⁸ Cell based screening was used to identify tenovin-6, which showed non-specific inhibition of SIRT1-3, but had efficacy against tumour growth in mice xenograft studies.²¹⁸ EX-527/Selisistat was identified through high throughput screening, and showed good potency and selectivity for SIRT1.²⁷⁹ This compound has been taken forward into clinical trials for Huntington's disease.²⁸⁰

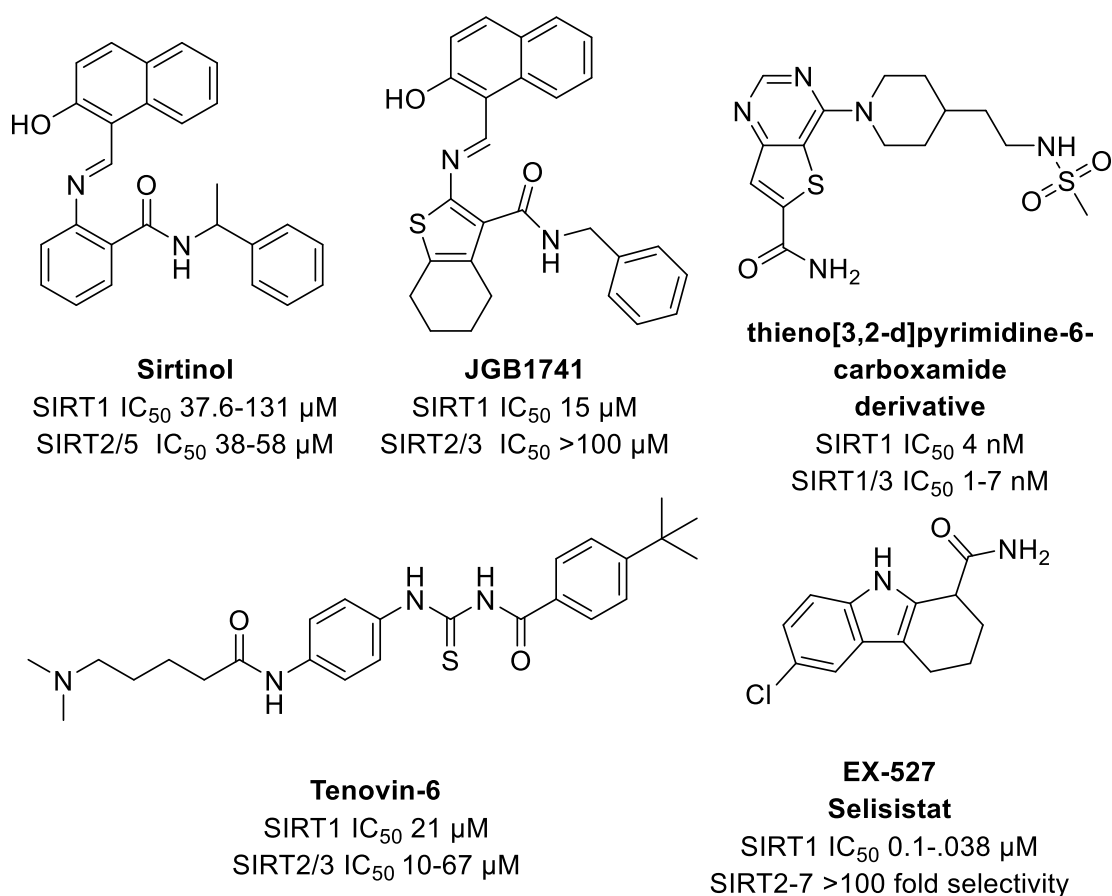


Figure 1.13: Selected chemical probes targeting SIRT1.

Examples of chemical probes which show variable levels of selectivity for SIRT1 over other SIRT enzymes.

1.5.5 Peptides as chemical probes for epigenetic enzymes.

Chemical probes are traditionally small molecules evolved from natural products or identified from libraries of chemical compounds. Peptides have emerged as a new class of chemical probes and have some advantages over the more classical small molecule chemical probes.²⁸¹ Due to their larger size and chemical composition, peptides are well suited to mimic the PPIs formed by epigenetic enzymes with their substrates.²⁸² Cyclic peptides have advantages over linear peptides in stability (resistance to peptidases), potency (fewer degrees of freedom) and as structural mimics (can be locked in protein-like secondary α -helical or β -sheet structures).²⁸¹

Peptides have previously been used to target both PCAF HAT and SIRT1. A linear peptide based on the section of the E1A adenoviral oncoprotein which binds to PCAF was shown to inhibit PCAF activity by competing with AcCoA/CoA for binding to PCAF (K_D of 6.5 μ M).²⁸³ mRNA display was used to target SIRT2, identifying cyclic peptide binders containing KAc mimic warheads (*e.g.* conversion of lysine ϵ -N acetyl to trifluoroacetate). The lead peptide showed activity against both SIRT1 (IC_{50} of 47 nM) and SIRT2 (IC_{50} of 3.2 nM), showing some selectivity for SIRT2.²⁸⁴

1.6 Project goals

The goal of this project is to gain a better understanding of PCAF and SIRT1 substrate selectivity. This can help broaden our understanding of their biological activity, as well as advance development of selective chemical probes.

Current knowledge on the substrate selectivity of PCAF and SIRT1 stems either from very broad proteomic work which has identified many non-histone substrates for PCAF and

SIRT1 without giving details on binding modes, or from studies on a relatively small number of selected substrates which shows binding interactions specific to the group of substrates studied. This means that although many substrates have been identified, binding interactions have only been thoroughly studied for a small group of substrates (usually H3 or H4 peptides), and it is still unclear how PCAF and SIRT1 achieve substrate selectivity against a large range of substrates.

As the effects of PTM crosstalk on PCAF and SIRT1 substrate selectivity have not been very widely explored, this was chosen as a focus for the project. Understanding the effect of combinatorial substrate PTMs on PCAF and SIRT1 activity can help elucidate binding modes and regulatory systems affecting the proteins. It has been shown that PTM crosstalk is an important factor in regulation of epigenetics. However, due to the dynamic and context dependent nature of the histone PTM landscape, there are difficulties in identifying relevant PTM crosstalk effects. In this work, the problem was approached by developing a high throughput screen with libraries of histone peptides with PTMs used as a simplified model (Chapter 3). Efforts were made to widen the utility of the HTS, by exploring the merits and limits of peptides as models for PPI interactions between PCAF, SIRT1 and their substrates (Chapter 4).

Selective chemical probes for PCAF HAT are lacking, which has further hampered understanding of substrate selectivity. For biological studies of PCAF, either genetic methods or non-selective probes must currently be used. To facilitate the biological study of PCAF, the development of selective and potent chemical probes is necessary. This problem was addressed here by using mRNA display of cyclic peptides against PCAF (Chapter 5).

To achieve these goals, a MALDI-TOF-MS based assay was developed to systematically characterise the protein substrates (Chapter 2). This assay was then developed into a high throughput screen which was used against histone peptide libraries with PTMs, identifying substrate selectivity and PTM crosstalk for PCAF not previously described (Chapter 3). Histone and non-histone substrate selectivity for PCAF and SIRT1 were further explored with selected targets, to validate results of the HTS, and to explore a wider applicability of the HTS (Chapter 4). Cyclic peptide mRNA display was carried out in order to develop cyclic peptide chemical probes for PCAF HAT, to allow for better characterisation of the PCAF HAT domain activity profile (Chapter 5).

Chapter 2 Characterisation of the HAT and HDAC proteins PCAF and SIRT1 using a MALDI-TOF-MS assay

2.1 Introduction

2.1.1 Biological relevance of the HAT PCAF and the HDAC SIRT1

Epigenetic acetylation networks on histone and non-histone proteins are dynamic in nature. However, the systems for regulating the global levels of acetylation and the levels of acetylation at specific biological sites are not fully understood. PCAF and SIRT1 are opposing HAT and HDAC enzymes which share biologically relevant substrates on histone proteins (H3K9 and H3K14) and non-histone proteins (*e.g.* HIF1 α and p53).^{20,119,182} In addition, PCAF and SIRT1 have a range of biological functions in cell signalling and disease states, with some overlap between their respective roles (Section 1.4). Exploring whether PCAF and SIRT1 are governed by the same signalling networks can therefore help understand how biological activity is regulated by protein specific selectivity and activity.

Because of their important biological roles, the structures and enzyme mechanisms of PCAF HAT and SIRT1 have been well characterised. Crystal structures of PCAF and the homologous GCN5 enzyme have been obtained with a number of substrates to help understand how binding modes affect substrate selectivity (PDB IDs: 1CM0, 1Q2C, 1Q2D, 1PU9, 1PUA, 1QSN, 7BY1 and 1JM4),²⁴⁻²⁹ and the enzyme mechanism of PCAF has been elucidated (Section 1.2.1).^{106,285} The enzyme mechanism of SIRT1 has also been characterised,²¹⁰ and the crystal structure of SIRT1 has been solved (PDB ID: 4I5I).²⁸⁶ More detailed information on the background of these enzymes can be found in Chapter

1. As no conserved sequence has been identified for PCAF and SIRT1 substrates, and new substrates for the proteins are regularly identified, the full range of substrates for the enzymes and the factors regulating substrate selectivity remains unclear.

Although the biological role of PCAF has been widely studied, the contribution of the BRD domain of PCAF to its biological functions has not been fully elucidated. PCAF is reported to catalyse the acetylation of HIV TAT1 protein at K28 through its HAT domain and to bind to K50Ac through its BRD domain.^{158,159} In this case the HAT and BRD domains play distinctive roles in regulating TAT1. The BRD domain of PCAF has also been reported to bind to acetylated histone peptides using saturation transfer difference NMR, showing a preference for H3K14Ac ($K_D 128 \pm 12 \mu M$) and H4K20Ac ($K_D 247 \pm 53 \mu M$) over H3K9Ac ($K_D 1051 \pm 90 \mu M$) with binding to HIV TAT1 K50Ac as a reference ($K_D 84 \pm 12 \mu M$).²⁸⁷ However, chemical probes developed to target PCAF BRD²⁷¹ do not affect the acetylation activity of PCAF towards histones when used in cells (Brennan group, unpublished results). Beyond this, the role of the BRD on the biological activity of PCAF has not been well characterised. Its contribution to the activity of PCAF HAT is likely to be substrate dependent.

2.1.2 Choice of MALDI-TOF-MS as a method for measuring HAT and HDAC activity

Acetylation and deacetylation can be monitored using different techniques. These are either direct techniques measuring the amount of acetylated product, or indirect techniques measuring the amount of side product produced (CoA in the case of HAT enzymes, and nicotinamide for SIRT1, Figure 1.3).

For HAT enzymes, direct assays developed include radiolabelling, microfluidic mobility, antibody-based assays, AlphaScreen™, Förster resonance energy transfer (FRET), NMR and mass spectrometry (MS) assays. Indirect assays include fluorescent thiol quantification and enzyme-coupling assays.^{288–291} However all of these techniques have inherent limitations. Fluorometric thiol-scavenger based assays suffer from compound interference and need large product concentrations for signal detection.^{290,291} Enzyme coupled assays are continuous assays, good for time-dependent studies, but suffer from compound interference and a need for relatively large concentrations of product.^{289,292} Detection of radiolabelled product peptide with filter binding or scintillation assays using radiolabelled AcCoA are very sensitive techniques, but the cost and dangers of radiation are prohibitive.^{289,293} AlphaScreen™, FRET, enzyme-linked immunosorbent assays (ELISA) and western blots are sensitive techniques using little reagent, but the need for selective antibodies limits their versatility.^{288,294} Fluidic mobility assays are not common due the highly specialised equipment used.²⁹⁵ NMR is sensitive but product analysis can be complicated, limiting screening of different substrates.⁴⁵

Assays assessing sirtuin HDAC activity have also been widely developed,²⁹⁶ using radiolabelling,²¹³ fluorometric,^{297–299} FRET,³⁰⁰ immunohistochemistry,^{301,302} NMR,³⁰³ and MS based methods.³⁰⁴ These have the same advantages and disadvantages as described above for equivalent methods with HAT assays.

Using mass spectrometry to detect mass changes corresponding to gain or loss of an acetyl group (± 42 Da) is a powerful technique for measuring enzyme activity of HAT and HDAC enzymes. It is a label free (optional) technique which does not rely on radiation or antibodies for readout. It is sensitive (femtomolar quantities detected), uses little

reagent, and can distinguish between products bearing single or multiple acetylations within a single assay (unlike any of the other techniques used). MS/MS can be used to further characterise products, allowing for identification of the amino acids affected by the enzymes. Depending on the instrument used, continuous assays can be developed (*e.g.* Rapidfire™), and different sizes/types of HAT substrates can be used (different methods/instruments might be used if peptide, protein or nucleosomes are the desired substrates).^{305,306} Disadvantages of MS are the need for access to specialised and expensive MS instruments, with different instruments needed for different applications.

Matrix assisted laser desorption ionization (MALDI) time of flight (TOF) mass spectrometry (MS) or MALDI-TOF-MS is a technique that is well suited for measuring mass changes in peptides (700-3500 Da) and proteins (4-20 kDa), as the soft ionization technique results in minimal degradation of peptide and proteins (unlike hard ionisation techniques such as electron impact or chemical ionisation). In addition, the ionisation technique results in formation of singly charged peptide species which is convenient for analysis.^{307–311} MALDI-TOF-MS sensitivity is dependent on the matrix, compounds and instruments used but has been shown to go down to sub-picomolar sensitivity, and small volumes ($\leq 1 \mu\text{L}$) of sample are required.³¹²

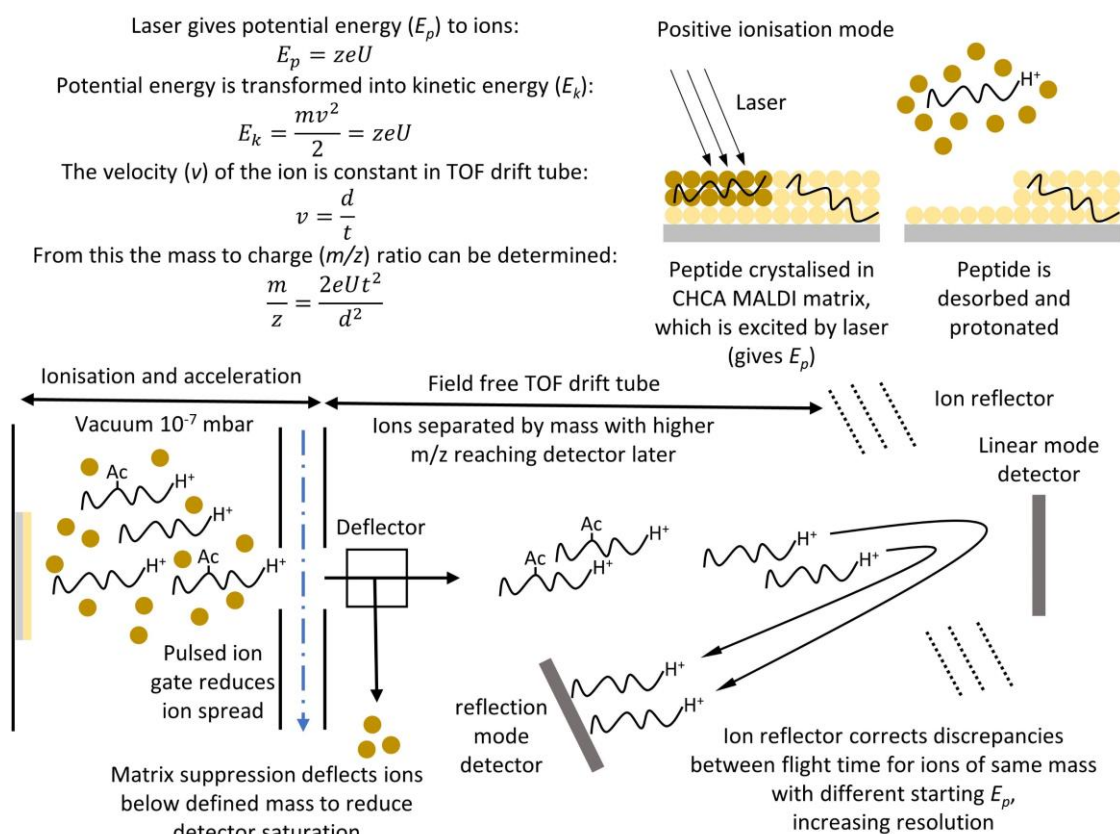


Figure 2.1: MALDI-TOF-MS method.

In MALDI-TOF-MS the mass of peptides is determined based on the mass/charge (m/z) ratio based on the acceleration voltage (U) used, the time (t) and distance (d) travelled in the time of flight (TOF) drift tube (e is the elementary charge). Peptides are first crystallised in 4-hydroxy- α -cyanocinnamic acid (CHCA) MALDI matrix recommended for peptide use, and a laser is used to desorb and ionise the peptide molecules, with a proton transferred from the matrix to the peptides in positive ionisation mode. The peptide enters the TOF drift tube and the time to reach the detector is measured with peptides separated based on their m/z ratio. Mass resolution is increased by using a pulsed laser beam and ion gate to minimise spatial and energy spread of peptides at entry of the drift tube, by using a deflector for removal matrix molecules to decrease detector saturation, and by using an ion reflector which alters the path length of particles based on their starting potential energy, correcting for discrepancies in potential energy distribution.

In MALDI-TOF-MS, the sample is mixed with a matrix solution for crystallisation onto a MALDI plate. The matrix is excited by the laser emitted by the MALDI source, transferring protons to the sample molecules (in positive mode) and allowing for desorption of the sample. The laser used in Bruker microflex and rapiflex instruments is a Bruker Smartbeam™ with a 355 nm wavelength, $\geq 100 \mu\text{K/pulse}$ and a variable repetition rate of 1-10kHz. A matrix type is chosen for optimal desolvation, desorption and ionisation properties, selected depending on the amount of energy needed to excite

the sample without causing fragmentation. 4-hydroxy- α -cyanocinnamic acid (CHCA) is the recommended matrix for acquisition of peptide MALDI-TOF-MS spectra. Mass analysis is carried out with time of flight (TOF), in which the compounds are separated based on their mass to charge (m/z) ratio. Compounds with a larger m/z ratio travel slower to the detector. Sample resolution is improved by pulsed ion extraction, a pulsed ion gate, MALDI matrix deflection to decrease saturation of the detector and the use of an ion reflection system. The ion reflection corrects differences in potential energy (E_p) transferred to peptides with identical masses from the laser by allowing particles with higher E_p to penetrate deeper into the energy field resulting in a longer flight path. Ion reflection is not suitable for high or low mass targets as high mass targets are more likely to fragment in this mode (referred to as post source decay or PSD), and low mass targets will not be correctly reflected (Figure 2.1). Using the methods described, MALDI-TOF-MS of peptides allows sufficient resolution (≤ 5 ppm) to identify isotopic peaks.^{312–314}

Bruker rapifleX instruments are equipped for MS/MS with a TOF/TOF setup, in which peptides are fragmented, and the fragments are analysed for sequence identification (Figure 2.2). In MALDI-TOF/TOF-MS, sample ions are first separated according to their mass (m/z ratio) and undergo PSD fragmentation by laser induced decay (LID) in the first TOF region. Fragments and their precursors travel at the same velocity (same initial E_p), and can therefore be selected for using a timed ion gate which is calibrated for time/mass correlation, and which selectively removes ions which do not correspond to the wanted parent mass.³¹⁵ An ion source is used to re-accelerate fragments for a second TOF separation followed by ion reflection and detection of the fragments.

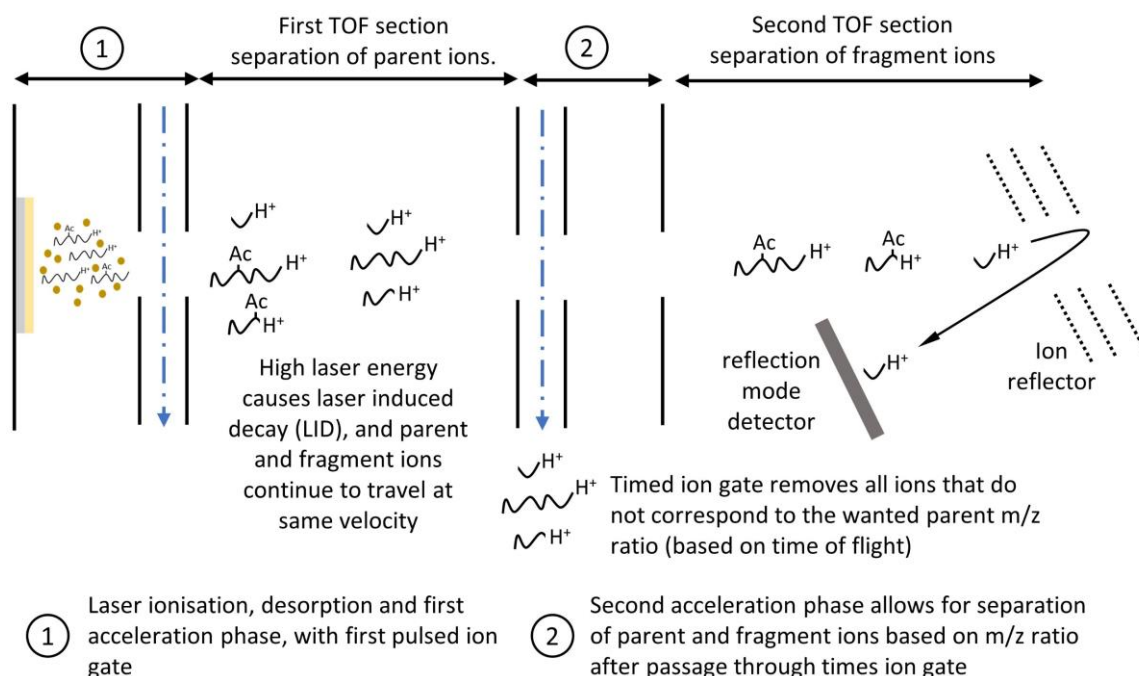


Figure 2.2: MALDI-TOF/TOF-MS method.

In the MALDI-TOF/TOF-MS method the ions are first separated in a time of flight (TOF) drift tube. A high energy initial laser source is used to give laser induced decay (LID) during the first separation. Parent and fragment ions retain the same initial velocity and therefore arrive at the same time at a timed ion gate. The m/z ratio of the wanted parent ion is determined and only fragments of the correct m/z ratio are allowed to pass through a timed ion gate. A second acceleration phase allows for separation of parent and fragment peaks based on their respective m/z ratios in the second TOF drift tube. An ion reflector is used for increased resolution and the ions are recorded by an ion detector.

Electrospray ionisation (ESI) is another soft ionisation technique which uses liquid rather than crystallised samples, avoiding issues with heterogeneity in sample crystallisation which can lead to errors in analysis of MALDI samples. However, ESI gives multiply charged species which complicates analysis of samples, and is more sensitive to changes in buffer (e.g. salt, DMSO and detergent use) and contaminants leading to sample suppression and source contamination which makes the technique less robust for assay analysis.³¹⁶ ESI MS has been widely developed by coupling with liquid chromatography (LC) systems for LC-MS analysis. The throughput of these methods has been improved by development of Rapidfire™ instruments which can sample assay mixtures and desalt, purify and concentrate samples using a solid-phase extraction cartridge before injection into the ESI-MS.^{317–319} Methods for LC-MS with MALDI have also been developed but are

much less used than LC-MS with ESI.³²⁰ However, MALDI-TOF-MS remains a method of choice for low volumes used (Rapidfire™ samples typically need >10 µL), speed of acquisition (<1s per sample for rapifleX MALDI-TOF-MS whereas Rapidfire™ LC-MS typically takes >7s per sample) making development of high throughput screens (HTS) for MALDI-TOF-MS more accessible, particularly when looking at epigenetic enzymes.^{321–323}

2.1.3 Aims

In this chapter, a MALDI-TOF-MS assay will be developed and used to carry out initial characterisation of PCAF and SIRT1 enzyme activity, which will provide context and groundwork for further work on exploring substrate selectivity (Chapter 3, Chapter 4) and inhibitor development (Chapter 5). Using two constructs of PCAF (a single HAT domain construct and a dual HAT-BRD construct) the role of the BRD of PCAF will also be studied in this chapter.

Specific aims are:

- Cloning, expression and purification of two different constructs of PCAF protein
- Expression and purification of SIRT1 protein
- Setting up of a MALDI-TOF-MS assay to monitor mass changes in H3₁₋₂₁ peptides
- Application of MALDI-TOF-MS to characterisation PCAF and SIRT1 enzyme activity
- Characterisation of PCAF substrate site selectivity and co-factor selectivity
- Characterisation of SIRT1 substrate selectivity
- Determination of the effectiveness of known HAT inhibitors towards PCAF HAT and HAT-BRD

2.2 Characterisation of PCAF

2.2.1 PCAF HAT and HAT-BRD protein production

2.2.1.1 PCAF HAT and HAT-BRD constructs

The full length PCAF protein consists of an E3 ligase domain, an *N*-terminal region which binds to p300 and contains a nuclear location signal, a HAT domain and a BRD (Figure 2.3a). The catalytic HAT domain is commonly used in biochemical studies and its crystal structure has been resolved (PDB 1CM0).⁹⁵ The crystal structure of the E3 ligase of mGCN5 (a PCAF homologue) has also been resolved (PDB 7BY1).¹⁰² The bromodomain structure of hPCAF with HIV-1 Tat-K50Ac peptide has been determined by NMR (PDB 1JM4).¹⁵⁹ However, the full length structure of PCAF has not been published, meaning that the way in which the *N*-terminal domain, HAT domain and bromodomain physically interact with each other is still unclear.

To better understand the roles of both the HAT domain and the BRD of PCAF, two protein constructs were used in this work: the single HAT construct, and the dual HAT-BRD construct (Figure 2.3b). Variations of these constructs were produced containing combinations of tags allowing for different protein purification methods, assays and screens. A poly-histidine tag (six histidines, His-tag) was used for Ni²⁺ binding and an avitag was used for protein biotinylation and subsequent streptavidin binding. The His-tag can be removed from the proteins by cleavage with tobacco etch virus (TEV) protease at a ENLYFQ\S sequence downstream of the His-tag.³²⁴ Proteins with an avitag sequence (GLNDIFEAQKIEWHE) were biotinylated enzymatically at the avitag lysine residue during protein production in *E. Coli* cells transformed with a biotin ligase (BirA) plasmid.³²⁵

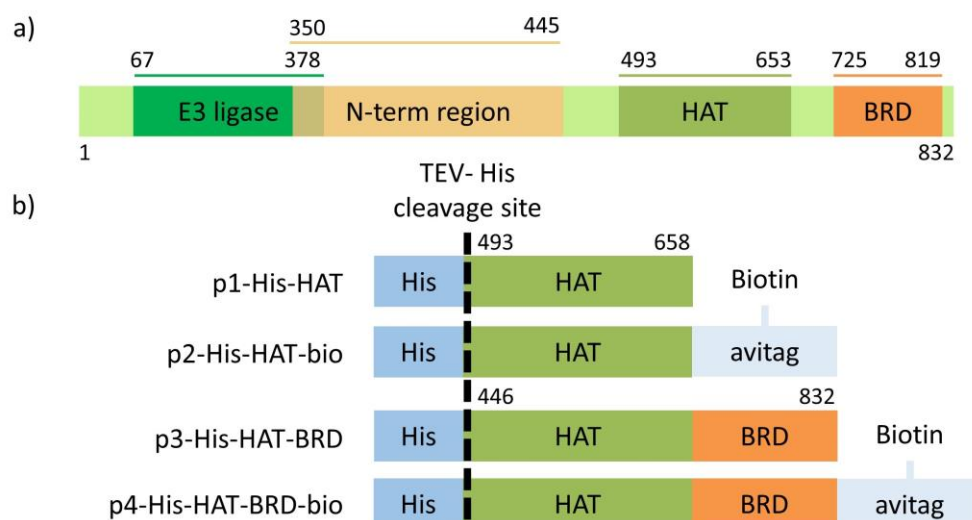


Figure 2.3: Structure of the PCAF protein and PCAF HAT and HAT-BRD plasmids.

a) Map of the different domains of PCAF: E3 ligase (ubiquitin ligase domain), N-term region (contains a section which binds to p300 and a nuclear localisation signal), HAT (histone acetyl transferase domain), and BRD (bromodomain). **b)** Two plasmids were generated for the single HAT domain and two for the dual HAT-BRD construct of PCAF, bearing either a single N-terminal polyhistidine tag (His-tag, His) or an N-terminal His-tag combined with C-terminal avitag (bio). Plasmids with an N-terminal His-tag for HAT and HAT-BRD were obtained from the structural genomics consortium. LIC cloning was used to produce constructs for HAT and HAT-BRD with an additional C-terminal avitag for biotinylation during protein expression in *E. Coli*. Plasmids are numbered p1-p4.

Plasmids containing PCAF HAT (p1-His-HAT, PCAF residues 493-658) or PCAF HAT-BRD (p3-His-HAT-BRD, PCAF residues 446-832) were donated by the structural genomics consortium. These constructs contained an N-terminal His-tag with no C-terminal tag, and were used as DNA templates for molecular cloning of PCAF HAT and PCAF HAT-BRD into pNIC-Bio vector using ligation independent cloning (LIC), to produce plasmids with an N-terminal His-tag and a C-terminal avitag (bio) (Materials and Methods 7.3).³²⁶ Briefly, primers were designed to generate PCR fragments containing the sequence of interest. The PCR fragments produced and pNIC-Bio3 vector (previously treated to remove generic insert) were incubated separately with T4 DNA polymerase and annealed due to complementary cohesive ends when combined. The new vectors were transformed into *E. Coli* for amplification, and were isolated and sequenced. This

generated two new plasmids with PCAF constructs: p2-His-HAT-bio, and p4-His-HAT-BRD-bio (Figure 2.3b).

2.2.1.2 PCAF HAT constructs - expression and purification

Several variations of the PCAF HAT domain protein were expressed and purified. The p1-His-HAT construct was used to produce unlabelled HAT enzyme for activity assays (MALDI-TOF-MS activity assays, Chapter 2, Chapter 3, Chapter 4) and His-HAT enzyme for binding assays requiring Ni^{2+} immobilisation (BLI). p2-His-HAT-bio containing a C-terminal avitag was used to produce HAT-bio enzyme for experiments requiring streptavidin-biotin immobilisation (mRNA display and BLI, Chapter 5). Expression and purification methods were based on a protocol obtained from Dr. Michael Fairhead (structural genomics consortium) and were optimized for each protein and for available equipment.

- Unlabelled HAT protein expression and purification

p1-His-HAT was expressed in Rosetta™ 2(DE3)pLysS Singles™ competent *E. Coli* cells at 37°C in autoinduction media. After reaching an OD_{600} of 0.4-0.7, cells were grown at 18°C for 18h. Cells were spun down, and a cell disruptor was used to lyse the recovered pellets. All purification steps were carried out at 4°C to avoid protein degradation. A first purification step with Ni^{2+} metal affinity chromatography was carried out by elution with an imidazole gradient (from 5 to 500 mM imidazole). To allow for an efficient His-tag cleavage by TEV protease, the collected fractions were combined, and buffer exchanged into a low imidazole buffer. Buffer exchange with desalting columns was performed but resulted in protein precipitation. Dialysis was therefore used as a more gradual buffer

exchange method. His-tag cleavage with TEV protease and dialysis was carried out simultaneously overnight with the TEV protease added to the protein solution in a snakeskin dialysis pouch placed in 5 L of low imidazole buffer. A second Ni²⁺ metal affinity chromatography column was performed to remove any remaining His-tagged protein, giving clean cleaved protein. Unlabelled HAT protein purity was assessed using sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-Page, >95% purity achieved, Figure 2.4) and protein concentration was measured using Nanodrop™ (Materials and Methods 7.4.4).

- His-HAT protein expression and purification

p1-His-HAT was expressed in *E. Coli* cells at 37°C in terrific broth (TB) media. After reaching an OD₆₀₀ of 0.4-0.7, cultures were cooled to 18°C for an hour, induced using Isopropyl β-d-1-thiogalactopyranoside (IPTG, 0.5 mM), and grown for a further 18h at 18°C. TEV cleavage was not carried out for this protein. After the initial Ni²⁺ gravity column purification step, insufficient purity was achieved, and recovered protein was further purified by size exclusion chromatography (SEC, Figure 2.4, Materials and Methods 7.4.5).

- HAT-bio protein expression and purification

The p2-His-HAT-bio plasmid was co-expressed with a plasmid coding for BiRA, allowing for biotinylation during protein expression. After reaching an OD₆₀₀ of 0.4-0.7 with growth at 37°C in TB media, cultures were cooled to 18°C for an hour, induced using IPTG (0.5 mM), and grown for a further 18h at 18°C in the presence of biotin (0.2 mM). The protein was purified firstly by Ni²⁺ gravity column, followed by buffer exchange and

TEV cleavage in dialysis tubing overnight, a second Ni²⁺ gravity column purification, and a final SEC purification. (Figure 2.4, Materials and Methods 7.4.6).

2.2.1.3 PCAF HAT-BRD constructs - expression and purification

Several variations of the PCAF HAT-BRD dual domain protein were expressed and purified. The p3-His-HAT-BRD construct was used to produce unlabelled HAT-BRD and His-HAT-BRD enzyme for activity assays and for binding assays requiring Ni²⁺ immobilisation respectively (MALDI-TOF-MS activity assays and BLI, all chapters). His-HAT-BRD recombinant protein provided by Eidarus Salah was used for enzyme characterisation. p4-His-HAT-BRD-bio was used to produce His-HAT-BRD-bio recombinant enzyme for experiments requiring streptavidin-biotin immobilisation (mRNA display and BLI, Chapter 5).

- Unlabelled HAT-BRD protein expression and purification

p3-His-HAT-BRD was transformed into Rosetta™ 2(DE3)pLysS Singles™ competent *E. Coli* cells. After reaching an OD₆₀₀ of 0.4-0.7 with growth at 37°C in TB media, cultures were cooled to 18°C for an hour, induced using IPTG (0.5 mM), and grown for a further 18h at 18°C. The protein was purified using a first Ni²⁺ gravity column, followed by desalting and TEV cleavage in dialysis tubing overnight, a second Ni²⁺ gravity column purification, and a final SEC purification. (Figure 2.4, Materials and Methods 7.4.7).

Although unlabelled HAT-BRD was found to be catalytically active (activity tested by MALDI-TOF-MS activity assay as described below), protein yield was low after purification and final concentration steps (0.18 mg/L isolated after SEC). His-HAT-BRD

protein expression and purification was further optimised by Eidarus Salah for large scale (12 L) protein expression for use in protein characterisation and screening assays.

- His-HAT-BRD-bio protein expression and purification

p4-His-HAT-BRD-bio was co-transformed with the BirA plasmid into Rosetta™ 2(DE3)pLysS Singles™ competent *E. Coli* cells. After reaching an OD₆₀₀ of 0.4-0.7 with growth at 37°C in TB media, cultures were cooled to 18°C for an hour, induced using IPTG (0.5 mM), and grown for a further 18h at 18°C in the presence of biotin (0.2 mM). The protein was purified using a first Ni²⁺ gravity column purification step. TEV cleavage was attempted but gradual loss of protein throughout purification steps led to very low protein yields. Further attempts at protein purification were therefore carried out without cleaving the His-tag. After the first Ni²⁺ gravity column, a SEC purification step was carried out. (Figure 2.4, Materials and Methods 7.4.8).

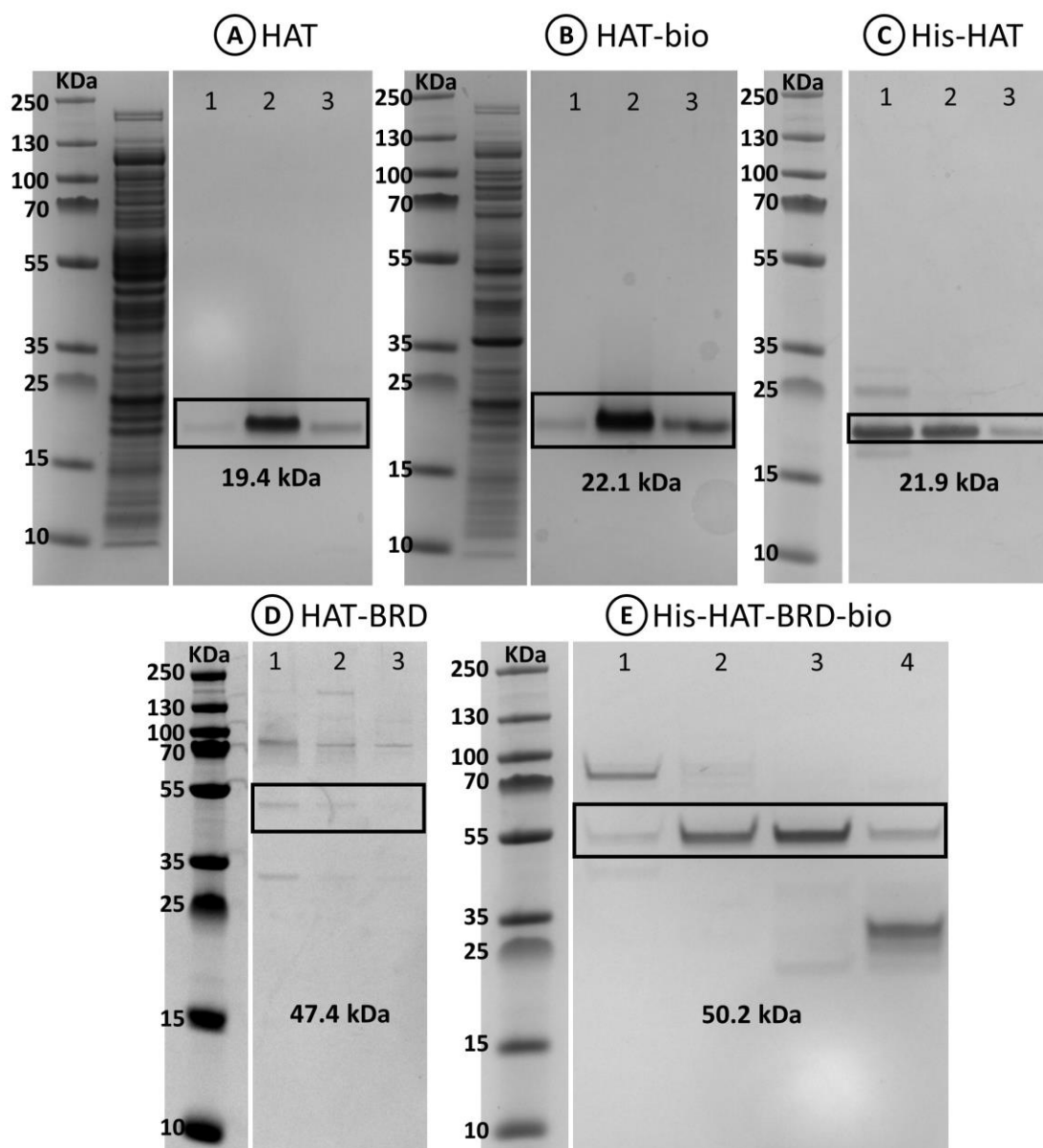


Figure 2.4: Protein gels showing purified PCAF protein for HAT and HAT-BRD construct variants.

Cropped SDS-PAGE gels for HAT, HAT-Bio, His-HAT, HAT-BRD and HAT-BRD-bio PCAF proteins produced by the author. A ladder (kDa) is shown on the left of each gel and the protein bands are indicated by a black box, and the molecular weight of the protein is given (kDa). **A** Protein gel following Ni^{2+} column. Lane 1 – 20 mM imidazole wash, lane 2 – 40 mM imidazole wash, lane 3 – 50 mM imidazole wash. **B** Protein gel following SEC purification. Lanes 1-3 – SEC fractions. **C** Protein gel following SEC purification. Lanes 1-3 – SEC fractions. **D** Protein gel following SEC purification. Lanes 1-3 – SEC fractions. **E** Protein gel following SEC purification. Lanes 1-4 – SEC fractions.

2.2.3 Development of a MALDI-TOF-MS assay to measure HAT activity

Histone peptides have widely been used to characterise epigenetic enzymes because of their ability to accurately replicate the behaviour of histone protein *N*-terminal tails (due to the high degree of flexibility exhibited by both the histone tails and the histone peptides).³²⁷ Studies characterising PCAF, GCN5 and SIRT1 using histone peptides as substrates have been previously reported (Chapter 1). Because PCAF and SIRT1 share H3K9 and H3K14 as substrate sites, peptides corresponding to histone 3 residues 1 to 21 (H3₁₋₂₁) were mainly used for enzyme characterisation.

Peptide properties (sequence, charge, secondary structure, intra- and inter-molecular interactions) determine how suitable a peptide is for MALDI-TOF-MS analysis. Peptides differ in ease of crystallisation, ionisation by the MALDI matrix, desorption, and fragmentation in the TOF mass analyser. To assess the effect of acetylation on the suitability of a H3₁₋₂₁ peptide for MALDI-TOF-MS, acetylated versions of H3₁₋₂₁ corresponding to known products of PCAF acetylation on H3 were purchased and tested alongside the unmodified H3₁₋₂₁ peptide. H3₁₋₂₁K9Ac, H3₁₋₂₁K14Ac and H3₁₋₂₁K9AcK14Ac were each mixed 1:1 with H3₁₋₂₁ at a concentration of 10 μ M and spotted onto MALDI plates. Three replicates were spotted for each mixture and three MALDI-TOF-MS spectra were acquired for each spot using a Bruker microflex instrument, giving nine replicates for each combination of peptides. The relative percentage of measured intensities of the unmodified and acetylated peptides was calculated. Each pair of peptides showed a ~50% intensity split, showing that acetylation of H3₁₋₂₁ peptides at K9, K14 or K9 and K14 does not significantly affect the relation between concentration and measured signal intensity for these peptides (Figure 2.5a). Comparison of relative intensities of

unacetylated and acetylated H3₁₋₂₁ peptides can therefore be used as a measure of PCAF enzymatic acetylation activity.

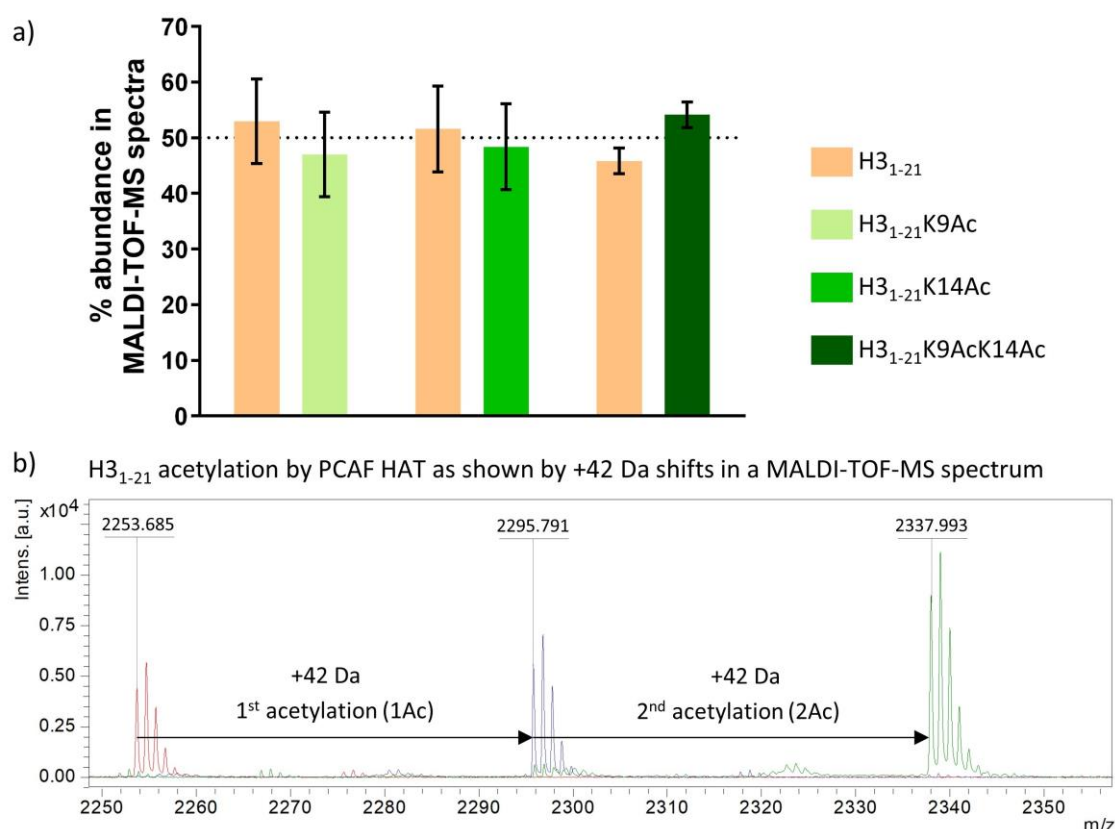


Figure 2.5: Development of a MALDI-TOF-MS activity assay.

a) H3₁₋₂₁K9Ac, H3₁₋₂₁K14Ac and H3₁₋₂₁K9AcK14Ac were each mixed 1:1 with H3₁₋₂₁ and spotted onto a MALDI plate. Quantification of relative intensities showed that acetylation did not significantly affect MALDI-TOF-MS acquisition, as shown by a 50:50 distribution of measured signal intensity for the unmodified and acetylated peptides. Results are shown as mean of %age abundance $\pm$ SD measured in 3 experimental replicates with 3 technical replicates for each spot (9 total replicates). **b)** Example spectrum of unmodified H3₁₋₂₁ peptide (red) overlayed with 1Ac (blue, +42 Da shift) and 2Ac (green, +84 Da shift) products of H3₁₋₂₁ peptide acetylation by PCAF HAT.

A MALDI-TOF-MS assay was then developed using a Bruker microflex instrument, to assess enzyme activity for the PCAF enzymes produced. The selected enzyme was incubated with AcCoA and a H3₁₋₂₁ at either 22°C or 37°C (initial experiments were carried out at room temperature/22°C, with a later switch to 37°C to better mirror physiological conditions). The buffer chosen was 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES, 50 mM) at pH 7.5 with NaCl (50 mM). At pH 7.5

the rate of non-enzymatic acetylation of histone peptides by AcCoA is reported to be negligible.²⁹⁰ Assay aliquots were taken at defined time points and quenched with 2% formic acid. Quenched assay solution (1 μ L) was mixed 1:1 with CHCA MALDI matrix and spotted onto a MALDI plate. Crystallisation was achieved by heating the plate in a hybridiser at 37°C.

MALDI-TOF-MS traces for each sample were obtained with the Flexcontrol software on a Bruker microflex MALDI-TOF-MS instrument in positive mode, with ion reflection for increased resolution. A range of laser power was defined following manual data acquisition on representative samples. The data for all samples was then acquired with an automatic method (parameters given in Appendix Figure F.1). Acquired data was analysed in the Flexanalysis software using a peak picking script (provided by Bruker). Peak intensity was used for analysis rather than peak area as it has been shown to give more reproducible results.³²⁸ Full experimental details for MALDI-TOF-MS assays can be found in materials and methods 7.7, with specific experimental conditions given for each set of results presented.

Incubation of PCAF HAT/HAT-BRD proteins with H3₁₋₂₁ showed two +42 Da shifts in the MALDI-TOF-MS traces, corresponding to two separate acetylation events (Figure 2.5b). This is consistent with reports showing PCAF acetylation of H3 protein and peptides at H3K14 and H3K9. PCAF HAT, His-HAT, HAT-Bio, HAT-BRD and His-HAT-BRD-Bio proteins produced by the author and His-HAT-BRD protein produced by Eidarus Salah were all shown to have enzymatic activity using the MALDI-TOF-MS assay. An example MALDI-TOF-MS time dependent assay comparing the activities of PCAF HAT, HAT-bio, HAT-BRD and His-HAT-BRD-bio is shown in Figure 2.6a, in which the biotinylated

proteins have lower specific activity towards H3₁₋₂₁ peptide than corresponding non-biotinylated proteins (Figure 2.6b). Assay conditions were chosen to produce a single +42 Da shift, to compare the specific activity of the different constructs towards the H3₁₋₂₁ unmodified peptide (Figure 2.6b). Activity of different batches of PCAF proteins is compared, showing some variability between batches of the same construct.

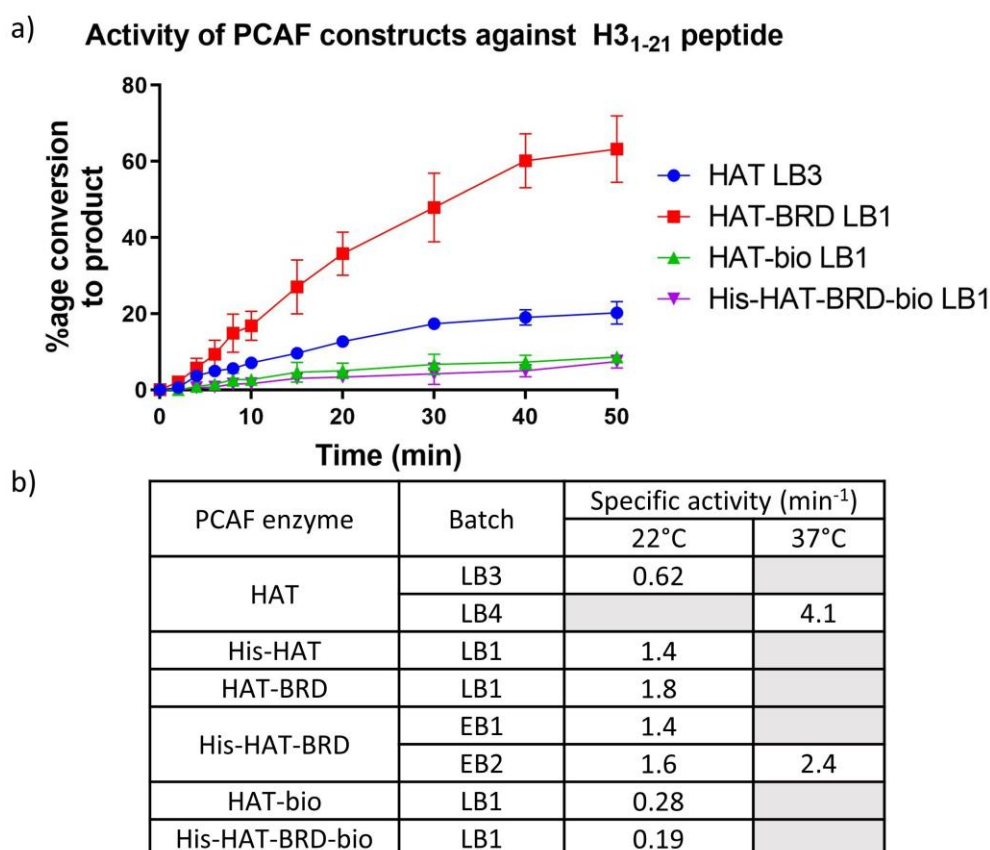


Figure 2.6: Enzyme activity of different PCAF constructs compared.

a) MALDI-TOF-MS time dependent activity assay comparing the activity of PCAF constructs HAT LB3, HAT-BRD LB1, HAT-bio LB1 and His-HAT-BRD-bio LB1 against H3₁₋₂₁ substrate peptide with a single acetylation event observed for all constructs. Assay conditions were enzyme (0.1 μM), H3₁₋₂₁ peptide (10 μM), AcCoA (50 μM) incubated in buffer (HEPES 50 mM, NaCl 50 mM) at 22°C. Results are shown as a mean ± SD of conversion to 1Ac product with three experimental replicates. **b)** Specific activity (min⁻¹) calculated for each PCAF protein at either 22°C or 37°C. Protein batches are identified as made by the author (Loane batch # or LB#) or by Eidarus Salah (Eidarus batch # or EB#)

All assays presented for enzyme characterisation were carried out at 37°C and used PCAF HAT B4 and PCAF His-HAT-BRD EB2 referred to as PCAF HAT and PCAF HAT-BRD for simplification. For data which was acquired with other protein batches this is stated.

At 37°C, PCAF HAT B4 was shown to have a higher specific activity towards H3₁₋₂₁ than His-HAT-BRD EB2 (Figure 2.6b).

2.2.4 PCAF characterisation using MALDI-TOF-MS

2.2.4.1 Site selectivity of PCAF HAT and HAT-BRD explored using MALDI-TOF/TOF-MS

MALDI-TOF-MS activity assays of PCAF with H3₁₋₂₁ as a substrate showed single acetylation (1Ac) and double acetylation (2Ac) products, consistent with reports of PCAF acetylation of H3 at K14 and K9 (Figure 2.5a). To confirm the acetylation site of PCAF on the H3₁₋₂₁ peptide, rapifleX MALDI-TOF/TOF-MS was used. H3₁₋₂₁ peptide (20 µM) was incubated with PCAF HAT or HAT-BRD (1 µM) at 37°C for 5 h to ensure complete conversion to products. Aliquots were taken at regular intervals and quenched with a 2% formic acid solution. The long reaction times resulted in formation of a third acetylation product (3Ac) as shown by a +126 Da mass shift compared to H3₁₋₂₁ peptide at 300 min (Figure 2.7). A large standard deviation (SD) between three replicates was observed for 3Ac product formation (Figure 2.7b), however MALDI-TOF/TOF-MS analysis of this peak was conclusive. Representative aliquots were identified as containing either parent peptide, single acetylation product (1Ac), double acetylation product (2Ac), or triple acetylation product (3Ac) using MALDI-TOF-MS. MS/MS spectra of the selected samples were acquired by MALDI-TOF/TOF-MS and the spectra were analysed using Bruker Biotoools software. The software identifies peptides based on recognition of fragments corresponding to theoretical masses for each possible peptide fragment generated from the input sequence. Identified fragments are classed as 'a', 'b' or 'y' fragments, based on the peptide bond being cleaved during fragmentation

(between the amino acid α -carbon and carbonyl for 'a' fragments, or between the amide carbonyl and nitrogen for *N*-terminal 'b' and *C*-terminal 'y' fragments). H3₁₋₂₁ unmodified peptide was used as a control peptide to ensure successful peptide identification through fragment mapping using MALDI-TOF/TOF-MS.

For the H3₁₋₂₁ control peptide, the peptide sequence was confirmed with high coverage of peptide fragments (>95%). The H3₁₋₂₁ peptide *C*-terminal amide, which results from cleavage of peptides synthesised on RINK amide resin, was correctly identified (Figure 2.8). This shows that the technique has sufficient sensitivity to identify PTMs on H3₁₋₂₁ peptides. For H3₁₋₂₁ peptides incubated with PCAF enzyme, three acetylation events were identified using MALDI-TOF-MS, with the third acetylation occurring between the 4h and 5h time points. The fragment masses identified for the 1Ac product were consistent with a H3₁₋₂₁K14Ac peptide sequence, identifying K14 as the preferred site of H3₁₋₂₁ acetylation by PCAF HAT/HAT-BRD (Figure 2.8). The fragment masses identified for the 2Ac product were consistent with a H3₁₋₂₁K9Ac14Ac peptide sequence, identifying K9 as the second site of H3₁₋₂₁ acetylation by PCAF HAT/HAT-BRD (Figure 2.9). The fragment masses identified for the 3Ac product were consistent with a H3₁₋₂₁K9Ac14AcK18Ac peptide sequence, identifying K18 as the third site of H3₁₋₂₁ acetylation by PCAF HAT/HAT-BRD (Figure 2.9). Equivalent results were obtained for PCAF HAT and HAT-BRD (results for HAT-BRD not shown).

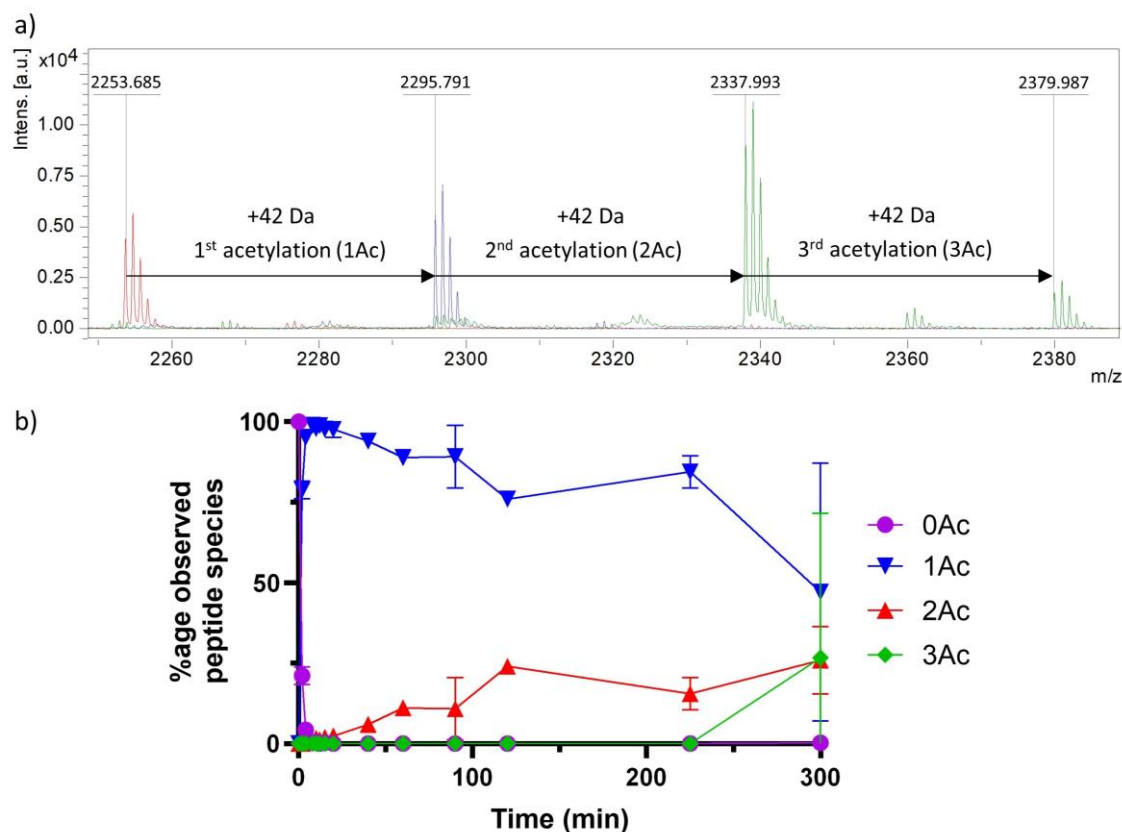
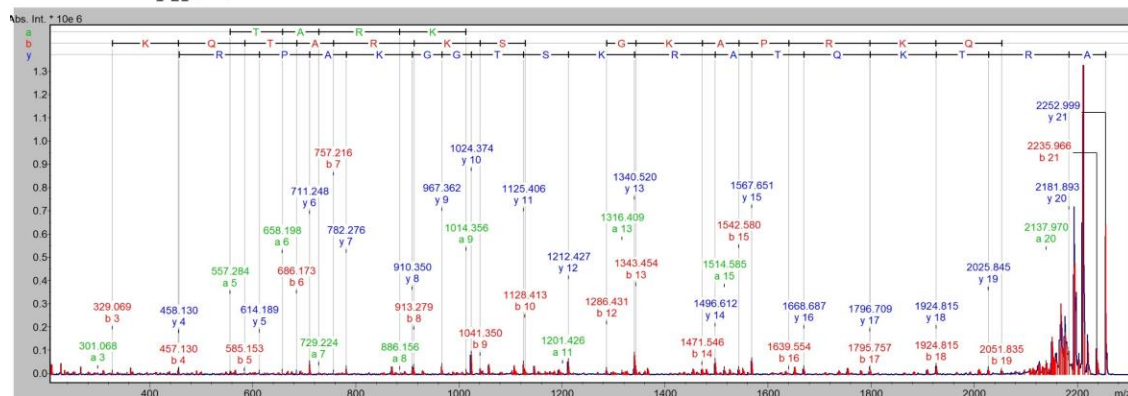


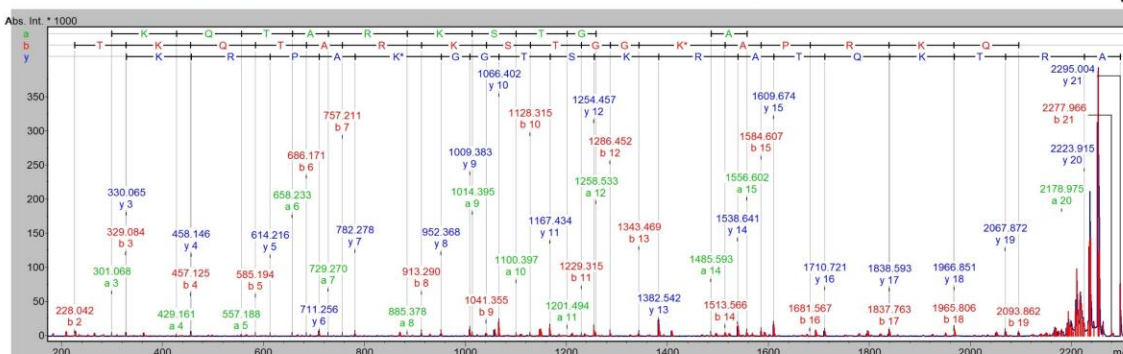
Figure 2.7: H3₁₋₂₁ acetylation by PCAF HAT.

Assay conditions are enzyme (1 μ M), AcCoA (100 μ M), H3₁₋₂₁ (20 μ M) incubated in buffer (HEPES 50 mM, NaCl 50 mM) at 37°C. **a)** Under specific conditions (high enzyme concentration, long reaction times), PCAF can acetylate H3₁₋₂₁ at three separate sites as shown by three +42 Da shifts. A MALDI-TOF-MS example spectrum for H3₁₋₂₁ incubated with PCAF HAT is shown, with NEC (red) overlaid with two H3₁₋₂₁ product spectra (blue and green) at different time points. **b)** MALDI-TOF-MS time dependent assay with PCAF HAT and H3₁₋₂₁ run in order to identify samples suitable for MALDI-TOF/TOF-MS analysis. A mean $\pm$ SD of the relative percentage of four different species of H3₁₋₂₁ peptide is plotted with 3 experimental replicates, corresponding to the parent peptide (0Ac) and three separate acetylation events (1Ac, 2Ac and 3Ac). Although SD for the 300 min time point was high, MALDI-TOF/TOF-MS showed that triple acetylation of H3₁₋₂₁ was indeed taking place.

Control: H3₁₋₂₁ peptide – unmodified



H3 ₁₋₂₁ Calc. mass: 2253.000 MH ⁺ (mono): 2253.338						H3 ₁₋₂₁ K14Ac Calc. mass: 2295.000 MH ⁺ (mono): 2295.349					
ion #	a	b	sequence	y	ion #	ion #	a	b	sequence	y	ion #
1	44.049	72.044	A	2253.337	21	1	44.049	72.044	A	2295.348	21
2	200.151	228.145	R	2182.3	20	2	200.151	228.145	R	2224.311	20
3	301.198	329.193	T	2026.199	19	3	301.198	329.193	T	2068.21	19
4	429.293	457.288	K	1925.152	18	4	429.293	457.288	K	1967.162	18
5	557.352	585.347	Q	1797.057	17	5	557.352	585.347	Q	1839.067	17
6	658.399	686.394	T	1668.998	16	6	658.399	686.394	T	1711.009	16
7	729.437	757.432	A	1567.95	15	7	729.437	757.432	A	1609.961	15
8	885.538	913.533	R	1496.913	14	8	885.538	913.533	R	1538.924	14
9	1013.633	1041.628	K	1340.812	13	9	1013.633	1041.628	K	1382.823	13
10	1100.665	1128.66	S	1212.717	12	10	1100.665	1128.66	S	1254.728	12
11	1201.712	1229.707	T	1125.685	11	11	1201.712	1229.707	T	1167.696	11
12	1258.734	1286.729	G	1024.637	10	12	1258.734	1286.729	G	1066.648	10
13	1315.755	1343.75	G	967.616	9	13	1315.755	1343.75	G	1009.627	9
14	1443.85	1471.845	K	910.594	8	14	1485.861	1513.856	KAc (+42)	952.605	8
15	1514.887	1542.882	A	782.5	7	15	1556.898	1584.893	A	782.5	7
16	1611.94	1639.935	P	711.462	6	16	1653.951	1681.946	P	711.462	6
17	1768.041	1796.036	R	614.41	5	17	1810.052	1838.047	R	614.41	5
18	1896.136	1924.131	K	458.309	4	18	1938.147	1966.142	K	458.309	4
19	2024.195	2052.19	Q	330.214	3	19	2066.205	2094.2	Q	330.214	3
20	2137.279	2165.274	L	202.155	2	20	2179.289	2207.284	L	202.155	2
21	2207.332	2235.327	A-amide	89.071	1	21	2249.343	2277.337	A-amide	89.071	1

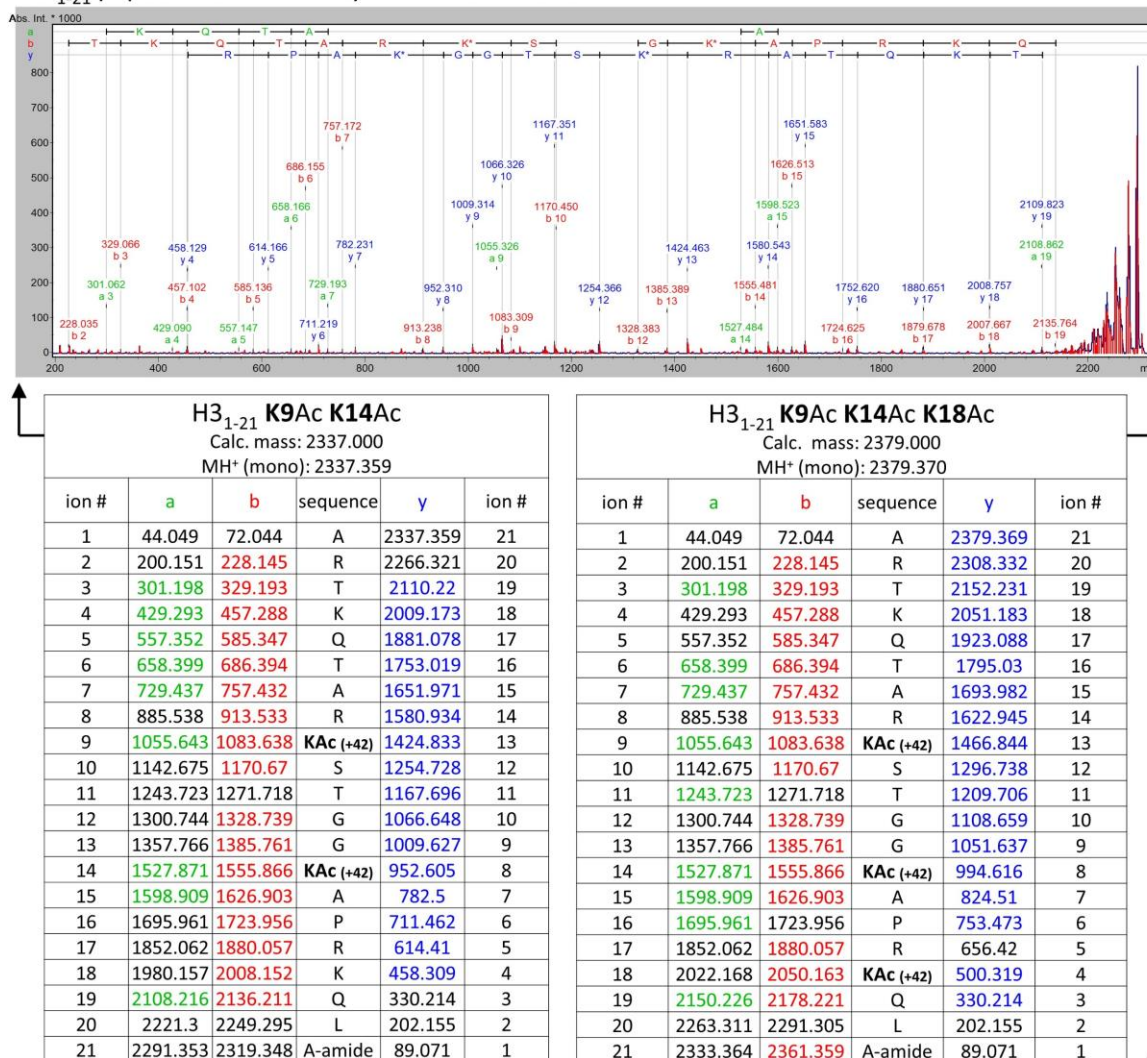


H3₁₋₂₁ peptide – First acetylation with PCAF HAT

Figure 2.8: MS/MS spectra of H3₁₋₂₁ control peptide, and H3₁₋₂₁ acetylated by PCAF HAT (1Ac product).

MS/MS spectra acquired using MALDI-TOF/TOF-MS and analysed using Biotoools are shown. Top: MS/MS spectrum of the unmodified H3₁₋₂₁ control peptide, correctly identifying the peptide sequence and C-terminal amide modification. Bottom: MS/MS spectrum of the single acetylation product of H3₁₋₂₁ peptide treated with PCAF HAT, identifying K14 as the acetylation site.

H3₁₋₂₁ peptide – Second acetylation with PCAF HAT



H3₁₋₂₁ peptide – Third acetylation with PCAF HAT

Figure 2.9: MS/MS spectra of H3₁₋₂₁ peptides acetylated using PCAF HAT (2Ac and 3Ac products).

MS/MS spectra acquired using MALDI-TOF/TOF-MS and analysed using Biotoools are shown. Top: MS/MS spectrum of the double acetylation product of H3₁₋₂₁ peptide treated with PCAF HAT, identifying K14 and K9 as acetylation sites. Bottom: MS/MS spectrum of the double acetylation product of H3₁₋₂₁ peptide treated with PCAF HAT, identifying K14, K9 and K18 as acetylation sites.

The MS/MS methodology used did not inform on dependence between H3₁₋₂₁K9, H3₁₋₂₁K14 and H3₁₋₂₁K18 acetylation. The Biotoools software identifies fragments based on the masses calculated from the peptide sequence provided, with optional PTM modification types and sites included. The software identifies the major species present at a given mass, based on coverage of a/b/y fragments identified. This does not rule out the presence of minor products with the same mass (*e.g.* H3₁₋₂₁K14Ac is identified as major 1Ac product, but H3₁₋₂₁K9Ac is not ruled out as a minor product found in the same MALDI-TOF-MS peak at 2296 Da). To test whether K9 and K14 acetylation are co-dependent, activity assays were run using H3₁₋₂₁K9Ac or H3₁₋₂₁K14Ac peptides as substrates for PCAF enzymatic acetylation. In both cases 1Ac products were identified using MALDI-TOF-MS, indicating that a pre-existing Ac PTM at either site does not inhibit PCAF enzymatic acetylation of the other site.

2.2.4.2 Evaluation of Co-factor selectivity of PCAF HAT and HAT-BRD using MALDI-TOF-MS

Next, the catalytic activities of PCAF for different acyl-CoAs were explored. A panel of acyl-CoA cofactors (Figure 2.10), including acyl-CoAs tested in reported works,^{329,330} was screened for activity of PCAF HAT or HAT-BRD with H3₁₋₂₁ peptide as substrate using MALDI-TOF-MS (Figure 2.11, Figure 2.12). A triplicate end point assay was carried out with PCAF enzyme (2 µM), H3₁₋₂₁ peptide (10 µM) and acyl-CoA (50 µM), quenched after 1h at 37°C with 2% formic acid. Enzyme activity was measured as a relative percentage of parent to product peak intensity in the MALDI-TOF-MS spectra for each of the acyl-CoA analogues tested, with average percentages and associated standard

deviations calculated. For all acyl-CoAs tested, PCAF HAT and HAT-BRD showed equal levels of activity towards H3₁₋₂₁ (Figure 2.12a). AcCoA showed ~50% single and ~50% double acetyl product (Figure 2.12a). Propionyl-CoA (PrCoA) showed full conversion to a single propionyl product, with only traces of a double propionyl product observed (Figure 2.11, Figure 2.12a). Traces of butyryl product were observed upon treatment with butyryl-CoA (Figure 2.11). In contrast isobutyryl-CoA and E-but-enoyl-CoA showed no traces of product, indicating that branching and unsaturations on the acyl chain may affect PCAF activity. None of the long-chain acyl-CoA variants showed any product, except succinyl-CoA. Succinylation of H3₁₋₂₁ was shown to be non-enzymatic, with no enzyme controls (NEC) showing the same levels of succinylation as samples incubated with PCAF HAT and HAT-BRD (Figure 2.11). Non-enzymatic succinylation of proteins has previously been identified.³³¹ The results obtained with the panel of acyl-CoAs were in line with previously reported results which showed 100-fold slower acetylation rates for butyrylation compared with propionylation of H3 peptides by PCAF HAT.³²⁹

Full length PCAF was reported to have propionylation and butyrylation activity,³³⁰ however this was observed in a radiometric activity assay that has a higher sensitivity than the MALDI-TOF-MS assay, and did not show relative rates of propionylation and butyrylation.³³⁰ In the results obtained, the bromodomain of PCAF did not appear to affect the acyl-CoA specificity of PCAF, with PCAF HAT-BRD showing equal levels of acylation activity as PCAF HAT with all acyl-CoA variants tested. A time course was run with PCAF HAT and HAT-BRD with H3₁₋₂₁ peptides as substrate and either AcCoA or PrCoA (Figure 2.12b). This showed slower acylation of H3₁₋₂₁ by PrCoA compared to AcCoA, in line with reported results for PCAF HAT.³²⁹ PCAF HAT-BRD activity showed only

a slightly slower conversion to product than PCAF HAT for both AcCoA and PrCoA, a further indication that the BRD did not significantly affect the acyl-CoA selectivity of PCAF.

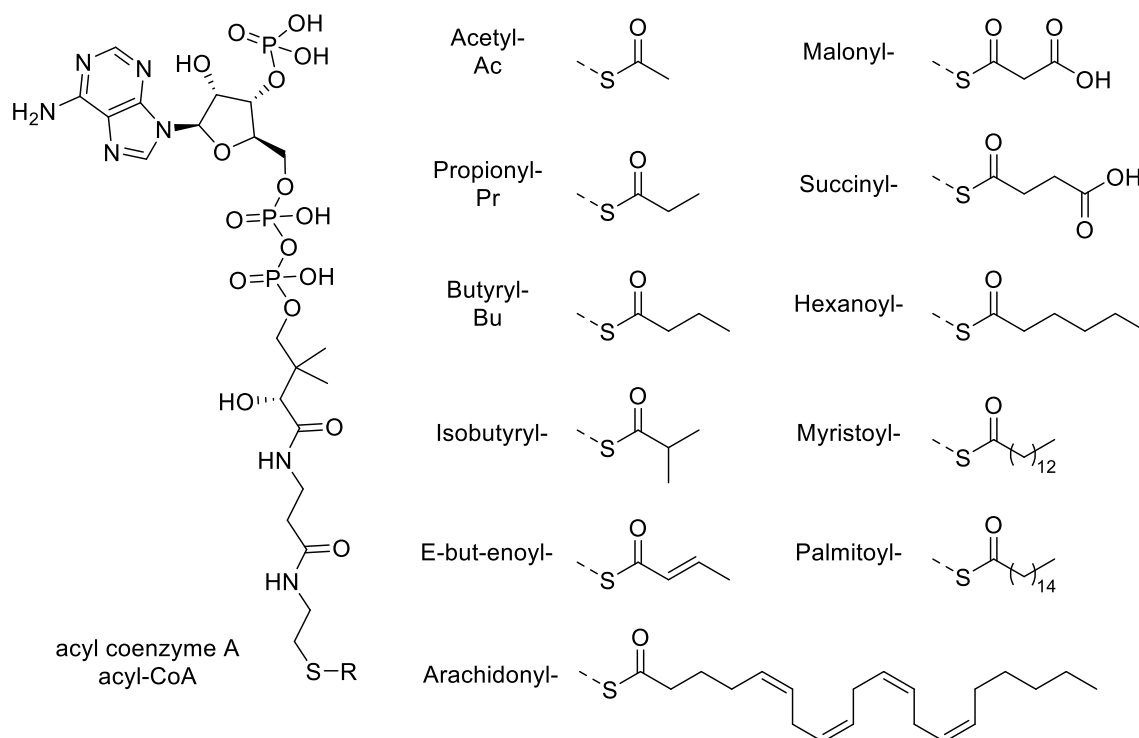


Figure 2.10: Structures of acyl-CoAs tested with PCAF HAT and HAT-BRD for enzymatic acylation activity.

Finally, to understand the site selectivity of PCAF propionylation, assays were run with purchased H3₁₋₂₁K9Ac, H3₁₋₂₁K14Ac and H3₁₋₂₁K9AcK14Ac peptides in the presence of PCAF HAT or HAT-BRD and PrCoA. H3₁₋₂₁ and H3₁₋₂₁K9Ac both showed a single propionylation product, whereas H3₁₋₂₁K14Ac and H3₁₋₂₁K9AcK14Ac gave no propionylation product. This indicated that propionylation by PCAF occurred selectively at H3₁₋₂₁K14. It also suggested that H3K9Ac did not interfere with PCAF propionylation activity. MS/MS of the propionylated products using MALDI-TOF/TOF-MS was attempted but showed poor fragmentation, giving inconclusive results. The results obtained were consistent with reports showing that PCAF HAT propionylated H3 at K14.³²⁹

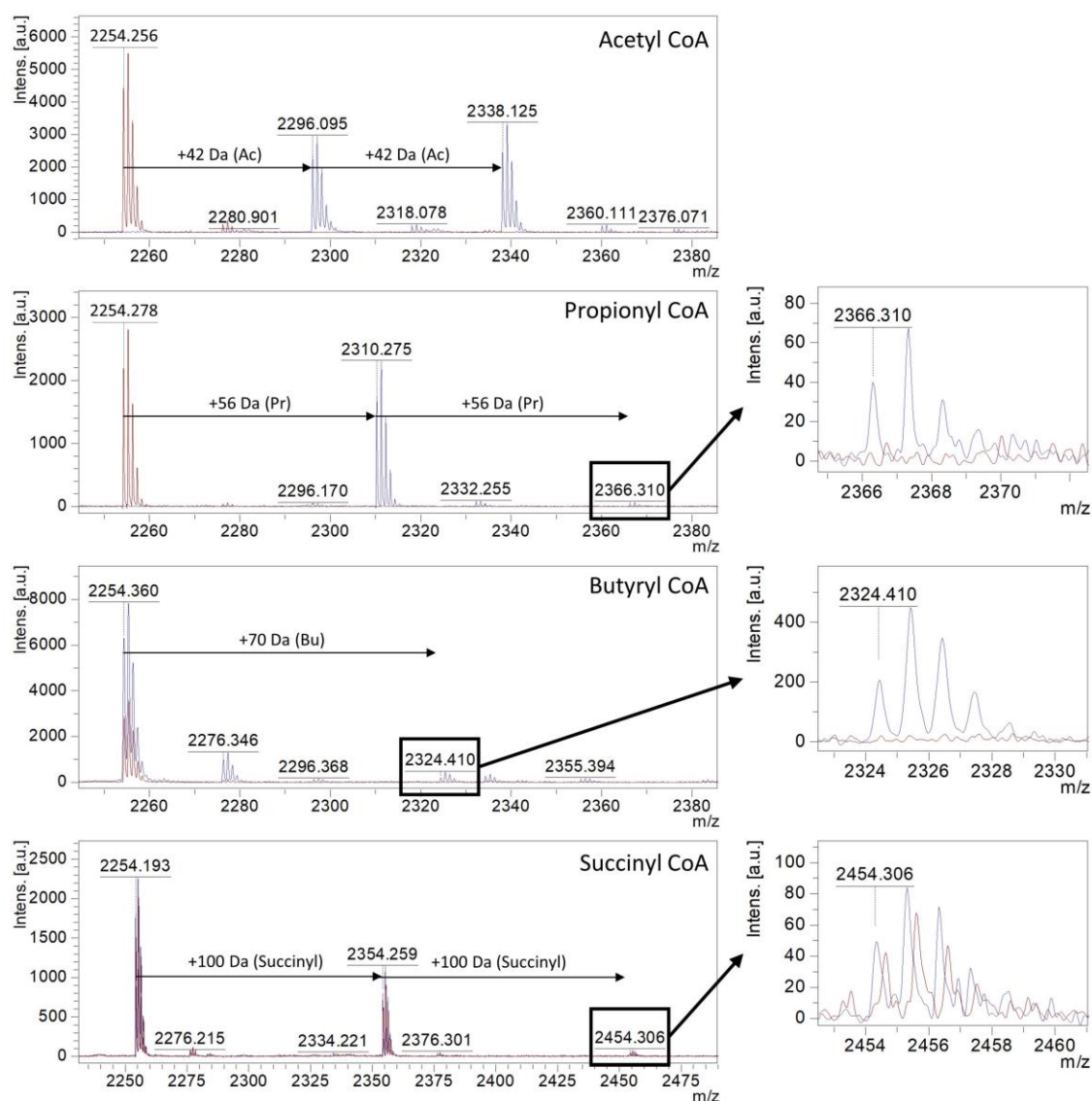


Figure 2.11: MALDI-TOF-MS spectra showing PCAF HAT activity with acetyl CoA, propionyl CoA, butyryl CoA and succinyl CoA.

An end point assay was carried out with PCAF HAT (1 μ M) HAT, H3₁₋₂₁ peptide (10 μ M) and acyl-CoA (50 μ M) incubated at 37°C for 40 min. Spectrum of NEC (red) is overlaid with active enzyme spectrum (blue) for each acyl-CoA shown. Two enzymatic modification peaks are observed for acetyl-CoA and propionyl-CoA, and one enzymatic modification is observed for butyryl-CoA. Two succinylation events occur non-enzymatically with succinylation of H3₁₋₂₁ observed in the NEC. Peaks of low intensity (trace amounts of product observed) are shown enlarged on the right-hand side. Equivalent results were obtained with PCAF HAT-BRD (results not shown).

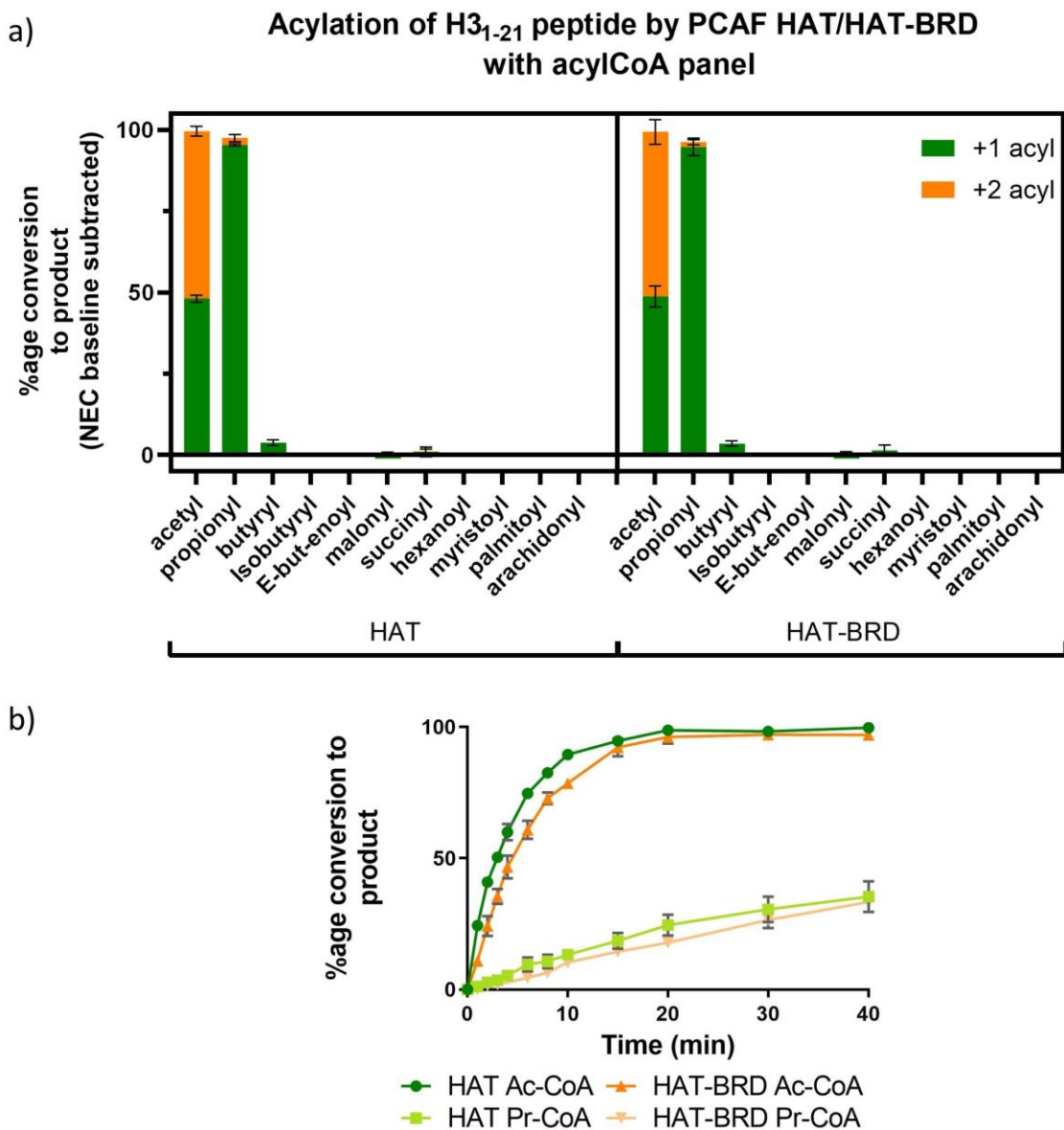


Figure 2.12: Acyl-CoA screen with PCAF HAT and HAT-BRD.

a) PCAF HAT and HAT-BRD activity with a panel of acyl-CoA variants was measured using MALDI-TOF-MS. An end point assay was carried out with enzyme (1 μ M), H3₁₋₂₁ peptide (10 μ M) and acyl-CoA (50 μ M) incubated at 37°C for 40 min. Results are shown as mean conversion to product $\pm$ SD over three experimental replicates. The screen showed equal activity levels for PCAF HAT and HAT-BRD for all acyl-CoA variants. Two acylation events were observed for acetyl-CoA. One acylation event was observed for propionyl-CoA. Trace levels of acylation were observed with butyryl-CoA. Succinylation was shown to be non-enzymatic, with no enzyme control (NEC), HAT and HAT-BRD showing equal conversion to product peptide. No other acyl-CoA tested showed any activity with PCAF HAT or HAT-BRD. **b)** PCAF HAT and HAT-BRD acylation of H3₁₋₂₁ with Acetyl-CoA (Ac-CoA) or propionyl-CoA (Pr-CoA) were measured over time using MALDI-TOF-MS. Conditions for the assays were enzyme (0.5 μ M), H3₁₋₂₁ peptide (10 μ M) and acyl-CoA (50 μ M), with time points taken over 40 min at 37°C. Results are plotted as mean of percentage conversion to 1Ac or 1Pr products $\pm$ SD for three experimental replicates, with <2% conversion to 2Ac observed (not plotted). Slower conversion to product peptide is observed for PrCoA than AcCoA with slightly slower conversion for PCAF HAT-BRD compared with PCAF HAT.

2.2.4.3 Evaluation of PCAF bromodomain chemical probes using MALDI-TOF-MS

The role of the bromodomain of PCAF has not been fully elucidated. To assess whether BRD inhibitors affected the activity of PCAF HAT, chemical probes which selectively target the BRD domain of PCAF were used in the MALDI-TOF-MS assay. L-Moses (PCAF BRD K_D 126 nM),²⁷¹ GSK4027 (PCAF/GCN5 BRD K_i 1.4 nM),²⁷² and their inactive enantiomers D-Moses and GSK4028 were obtained from Dr. Moses Moustakim (Brennan group, Structural Genomics Consortium). PCAF HAT-BRD acetylation activity against H3₁₋₂₁ was evaluated with a DMSO control or 100 μ M of either L-Moses, D-Moses, GSK4027 or GSK4028. None of the compounds significantly affected PCAF HAT-BRD activity, with compounds and DMSO control showing equal levels of conversion to a singly acetylated product (Figure 2.13). This indicated that the enzyme activity of PCAF HAT-BRD was not affected by blocking of the BRD binding site.

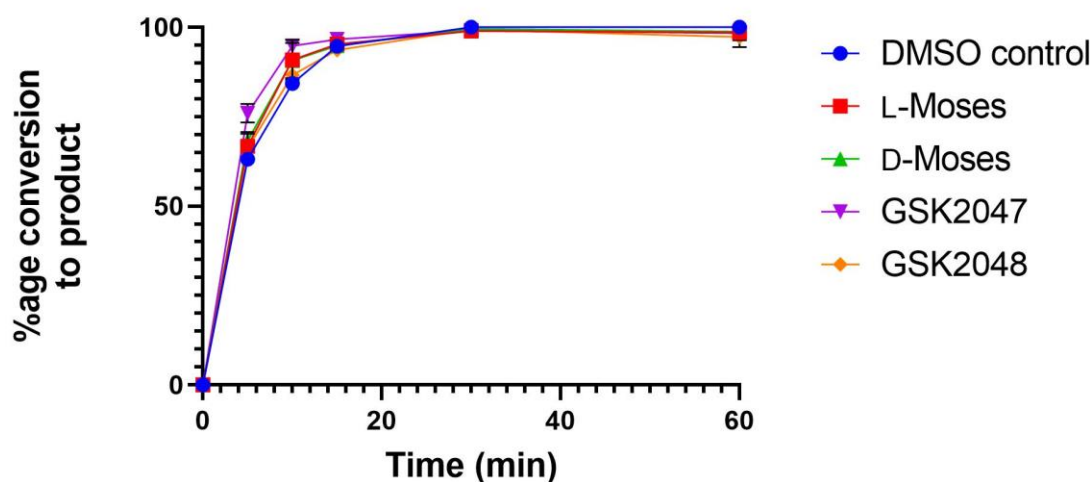


Figure 2.13: Screening of bromodomain inhibitors with PCAF HAT-BRD.

Selective PCAF bromodomain inhibitors L-Moses and GSK2047 along with their inactive analogues D-Moses and GSK2048 were tested in a time-dependent MALDI-TOF-MS activity assay with PCAF HAT-BRD and H3₁₋₂₁ peptide as substrate. Assay conditions are 1 μ M enzyme, 10 μ M H3₁₋₂₁ peptide, 50 μ M AcCoA, 100 μ M compound, 1% DMSO buffer used throughout. Percentage conversion to 1Ac product is plotted as a mean $\pm$ SD of two experimental replicates (<1% conversion to 2Ac product was observed for all data points, data not shown).

2.2.4.4 Evaluation of reported HAT chemical probes using MALDI-TOF-MS

Chemical probes targeting the HAT enzymes PCAF, GCN5, p300 and CBP have been developed using a variety of screening methods and assays (Section 1.5.2). Variation in protocols and materials used leads to inconsistency in reported results, making it difficult to select a chemical probe for further study. For these reasons, commercially available chemical probes were selected for testing based on their reported activities for PCAF, GCN5, p300 or CBP and consistently tested in the MALDI-TOF-MS assay to measure their inhibitory activity against PCAF HAT and HAT-BRD (Figure 2.14). The compounds were dissolved in DMSO to make up 10 mM stock solution and were tested against PCAF HAT and HAT-BRD at concentrations ranging from 0.78 μ M to either 100 μ M or 200 μ M.

Compounds A-485, CTPB and curcumin, which were reported to be selective for p300 over PCAF, did not inhibit acetylation activity of PCAF HAT/HAT-BRD for H3₁₋₂₁ peptides in the MALDI-TOF-MS assay, which is consistent with reported results.^{242,244,261} CPTH2 and butyrolactone 3 were both reported as GCN5 inhibitors, but showed no inhibition of PCAF activity in the MALDI-TOF-MS assay.^{263,264} SPV-106 reportedly had no effect on GCN5 activity whilst increasing PCAF HAT activity.²⁴⁹ SPV-106 showed no activation or inhibition of PCAF enzymatic acetylation in the MALDI-TOF-MS assay (Appendix Figure B.1).

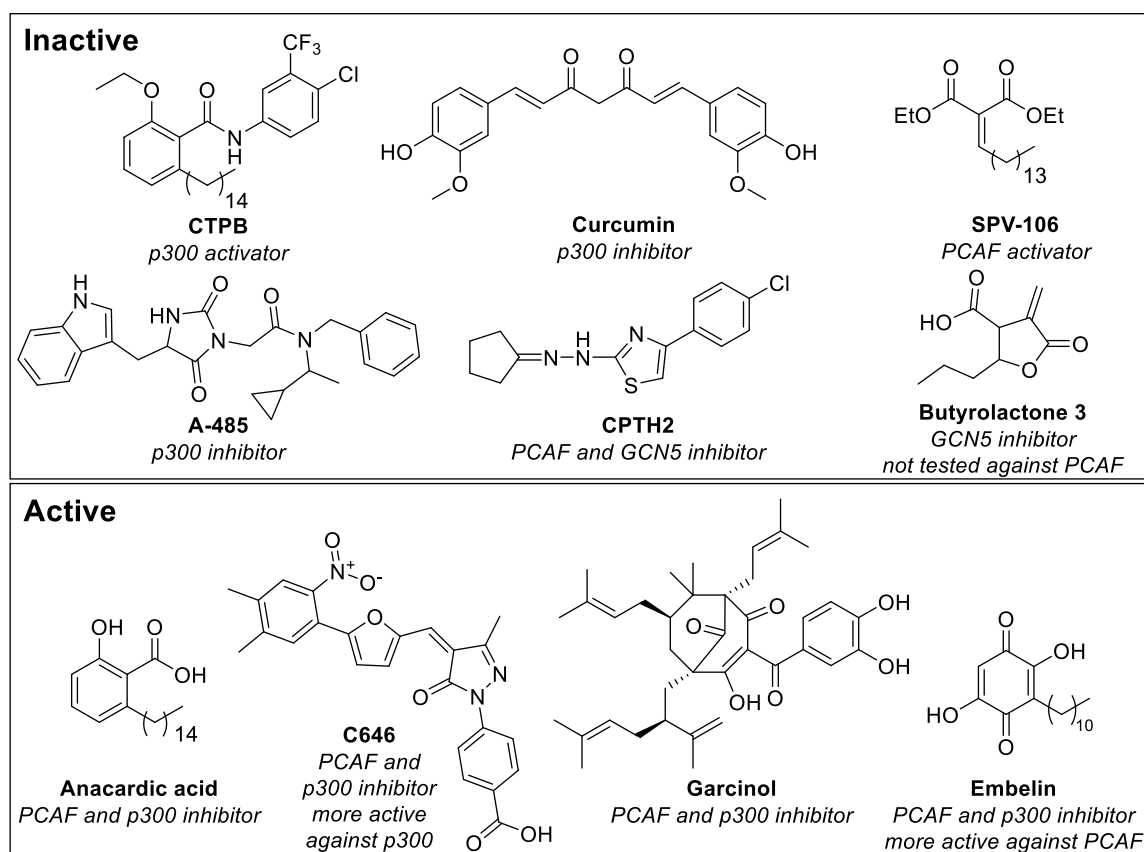


Figure 2.14: HAT inhibitors tested against PCAF HAT and HAT-BRD.

Inhibitors are separated into inactive (top) or active (bottom) categories, based on whether or not they affected PCAF HAT/HAT-BRD activity in the MALDI-TOF-MS assay with H3₁₋₂₁ peptide as substrate. The structure, name (bold), and reported activity (italics) is given for each inhibitor.

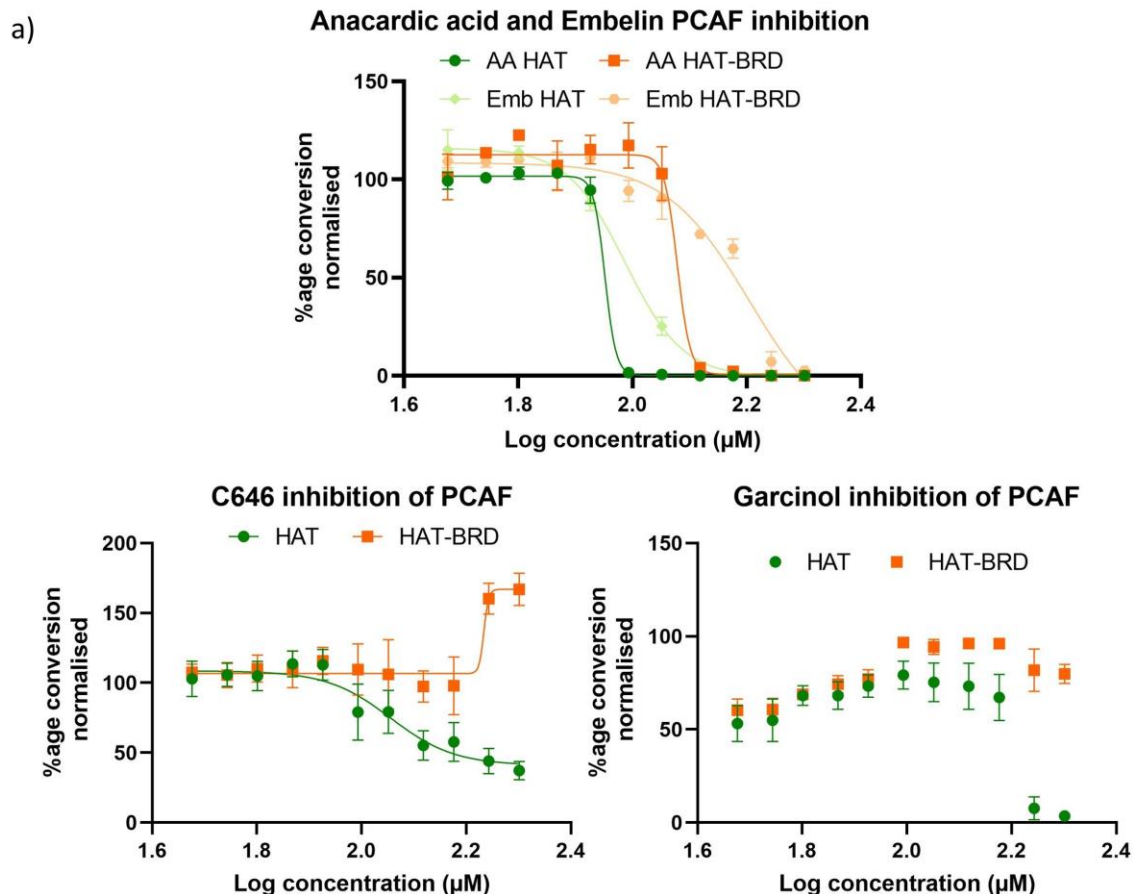
Anacardic acid (AnA),^{119,246} C646,^{243,258} garcinol^{119,245,259} and embelin (Emb)²⁵¹ are reported to show inhibition of PCAF HAT activity. Anacardic acid and embelin were tested in the MALDI-TOF-MS assay and showed straightforward inhibition profiles against PCAF HAT and HAT-BRD (Figure 2.15a). Both compounds had greater inhibition of PCAF HAT activity (AnA IC₅₀ 89 μ M, Emb IC₅₀ 113 μ M) than PCAF HAT-BRD activity (AnA IC₅₀ 121 μ M, Emb IC₅₀ 162 μ M, Figure 2.15b). C646 inhibited PCAF HAT (IC₅₀ 160 μ M). However, full inhibition of PCAF HAT activity was not achieved at a concentration of 200 μ M C646, with ~37% activity conserved compared to the no enzyme control. For HAT-BRD, an increase in enzyme activity was observed for C646 compound concentrations of 175-200 μ M (Figure 2.15a). Inhibition with garcinol showed a different activity profile compared to the other chemical probes. For

47-100 μM garcinol, PCAF HAT/HAT-BRD activity increased with increasing inhibitor concentration. For 100-200 μM garcinol, PCAF HAT activity decreased with increasing inhibitor concentration, giving a measurable IC_{50} of 160 μM . For garcinol with PCAF HAT-BRD, only a small decrease in enzyme activity was seen at 175-200 μM inhibitor concentrations. Garcinol therefore had greater inhibitory activity against PCAF HAT compared with HAT-BRD (Figure 2.15a).

Reported IC_{50} values for garcinol and embelin are $\leq 10 \mu\text{M}$. For anacardic acid, the reported IC_{50} values ranged from 1-1000 μM , depending on the assays and materials used. For C646, which was reported as a p300 selective inhibitor, a 10% decrease in PCAF activity was reported with C646 (10 μM inhibitor concentration). The IC_{50} values measured in the MALDI-TOF-MS inhibitor screen were therefore significantly higher than reported IC_{50} values for garcinol and embelin. For anacardic acid, MALDI-TOF-MS measured inhibition was towards the lower range of reported IC_{50} values. Observed results for C646 inhibition of PCAF HAT activity were consistent with reported results, as full enzyme inhibition was not observed even at concentrations of 200 μM C646 in the MALDI-TOF-MS screen (Figure 2.15).

A disparity between reported and measured IC_{50} values is expected due to the variation in IC_{50} measurements depending on substrate and enzyme forms used, buffer, relative concentrations of substrate to enzyme to co-factor, and assay type.³³² Comparing inhibitors in a single assay allows us to observe differences between PCAF HAT and HAT-BRD activity and to give a tentative ranking for inhibitors. For PCAF HAT, the inhibitor activity ranking was anacardic acid > embelin > C646 > garcinol in terms of inhibitor potency whereas for PCAF HAT-BRD, it was found to be anacardic acid >

embelin > garcinol in terms of inhibitor potency with C646 showing an unexpected activating effect at 175-200 μM concentrations (Figure 2.15).



b)

Inhibitor	MALDI assay measured IC_{50} (μM)		Literature IC_{50} (μM)
	PCAF HAT	PCAF HAT-BRD	PCAF
Anacardic acid	89 ± 10	121 ± 7	5-1000
Embelin	113 ± 11	162 ± 10	7.2
C646	114 ± 14	N/A	<10% conversion at 10 μM compound
Garcinol	160 ± 9	N/A	5

Figure 2.15: Inhibition profiles of active HAT inhibitors tested against PCAF HAT and HAT-BRD using MALDI-TOF-MS.

a) Plotted data for determination of IC_{50} values of anacardic acid, embelin, C646 and garcinol. Results are shown as mean percentage conversion to product $\pm$ SD of three experimental replicates. A four variable fit is shown where applicable in order to calculate IC_{50} values.

b) Observed IC_{50} values of compounds tested are reported as mean IC_{50} values over 9 replicates $\pm$ standard deviation. Reported IC_{50} values (or available inhibition data) are given for anacardic acid,^{242,247,248,333} embelin,²⁵¹ C646²⁵⁸ and garcinol^{243f} for comparison.

2.3 Characterisation of SIRT1

2.3.1 SIRT1 protein production and purification

A construct for SIRT1 was obtained from John Denu via Addgene (Plasmid #13735, Figure 2.16a).³³⁴ The construct contains human SIRT1 (residues 82-747) with an N-terminal His-tag. The catalytic domain of SIRT1 is contained in this construct (residues 241-516).²⁸⁶ The protein was expressed and purified following the protocol used in Hallows *et. al*, with an additional SEC purification step to increase protein purity. The protocol was optimized based on instruments and materials available.³³⁴

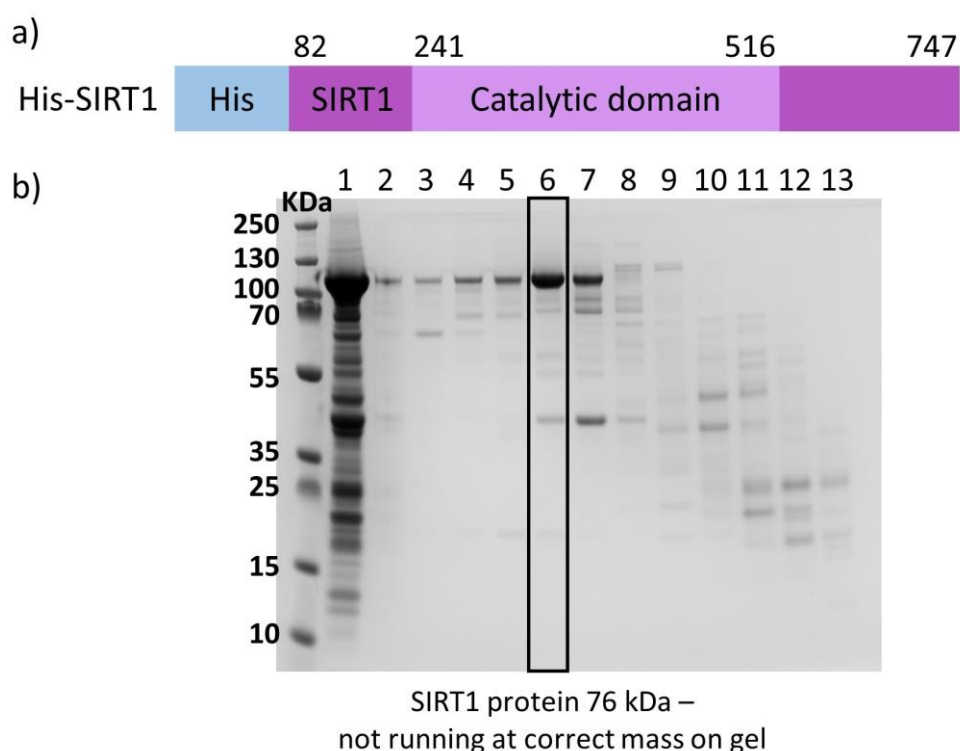


Figure 2.16: SIRT1 protein map, and SDS-PAGE gel of SIRT1 protein purification.

a) SIRT1.1 protein construct obtained from John Denu via Addgene.³³⁴ **b)** SDS-Page gel following SEC protein purification. Ladder is shown (kDa) on left, and the lane corresponding to the purified protein is shown by a black box. Lane 1 – sample injected onto SEC column 2-13 – collected SEC fractions.

His-SIRT1 was expressed into Rosetta™ 2(DE3)pLysS Singles™ competent *E. Coli* cells.

After growth at 37°C in 2YT media, cells were induced at an OD₆₀₀ of 0.4-0.7 with IPTG (0.5 mM) and grown at 25°C for 8h. Cells were spun down, and a sonicating probe was

used to lyse the resuspended pellets. All purification steps were carried out at 4°C or on ice to avoid protein degradation. A first purification step with Ni²⁺ metal affinity chromatography was carried out using a gravity His-trap column and protein was eluted using a 0-250 mM imidazole gradient. Recovered protein aliquots were pooled, concentrated, and further purified by SEC (Figure 2.16b). Protein purity was assessed using SDS-Page (>95% purity achieved) and selected protein aliquots were concentrated and quantified using Nanodrop™ (Figure 2.16c, Materials and Methods 7.4.9).

2.3.2 SIRT1 characterisation using MALDI-TOF-MS

The MALDI-TOF-MS assay developed for measuring PCAF acetylation activity (Section 2.2.2) was adapted to measure SIRT1 deacetylation activity. AcCoA was replaced by NAD⁺ as cofactor, with H3₁₋₂₁K9Ac and H3₁₋₂₁K14Ac and H3₁₋₂₁K9AcK14Ac used as peptide substrates, as SIRT1 is known to catalyse the deacetylation of H3K9 and H3K14.⁸⁴

SIRT1 was incubated with NAD⁺ and catalysed the acetylation of H3₁₋₂₁ peptides at 37°C. Buffers containing either Tris (50 mM) or HEPES (50 mM) at pH 7.5 with NaCl (250 mM) were tested, giving similar results. HEPES buffer was used to keep conditions between PCAF and SIRT1 assays as similar as possible. Time-course assays were run in the same way as PCAF activity assays with 2% formic acid solution in water used for quenching and spotting onto MALDI plates with CHCA MALDI matrix (1:1 with sample) for MALDI-TOF-MS analysis. Spectra were acquired and processed using the same parameters as for PCAF activity assays: spectra acquisition on Bruker microflex MALDI-TOF-MS instrument in positive ion reflection mode, automatic acquisition and processing with Flexanalysis software (Section 2.2.2, Material and Methods 7.7).

Results showed that the H3₁₋₂₁K9Ac and H3₁₋₂₁K14Ac substrate peptides were converted to a deacetylated H3₁₋₂₁ peptide product, consistent with a -42 Da mass shift. Diacetylated H3₁₋₂₁K9AcK14Ac substrate peptide showed a mix of mono (-42 Da) and full deacetylation (-84 Da, Figure 2.17a). These results were consistent with reported SIRT1 activity.⁸⁴

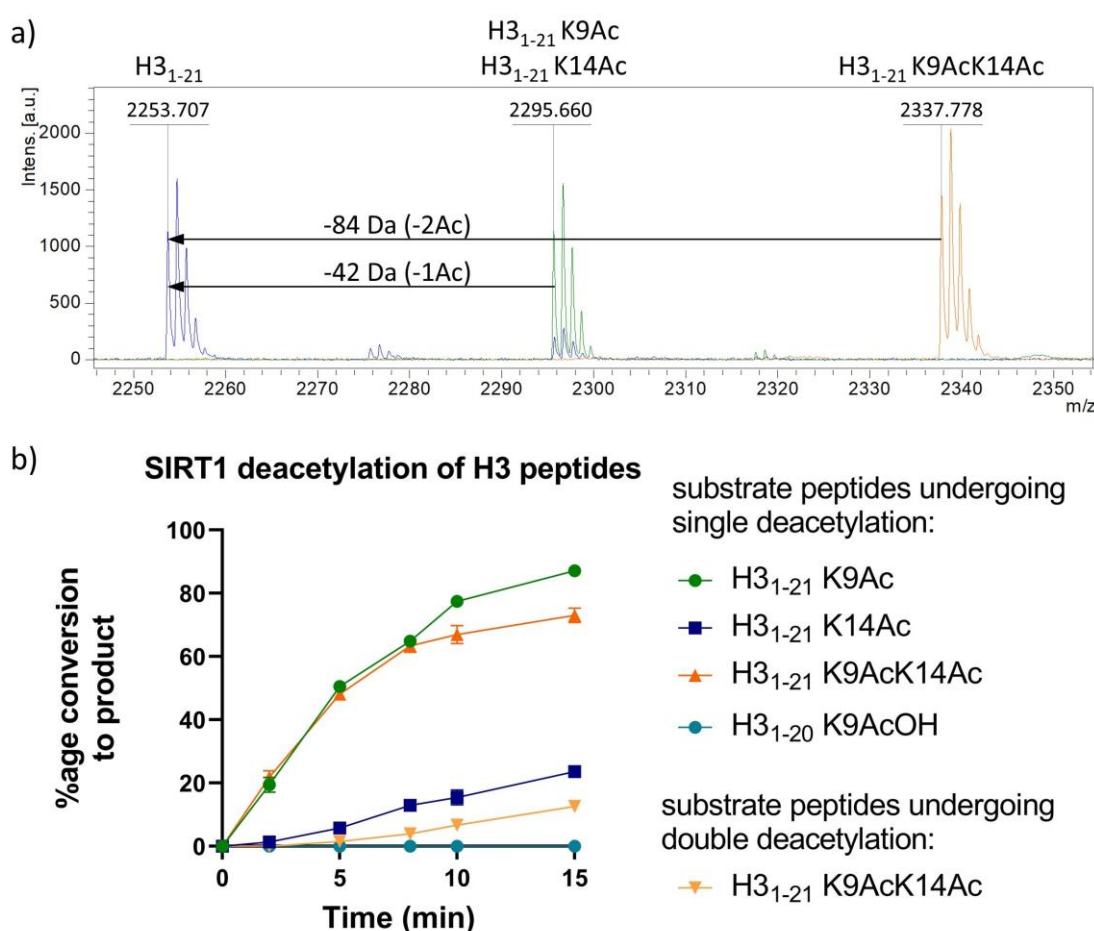


Figure 2.17: SIRT1 deacetylation of acylated H3 peptides.

a) MALDI-TOF-MS spectra for H3₁₋₂₁K9AcK14Ac (blue) superimposed with SIRT1 single deacetylation products H3₁₋₂₁K9Ac/H3₁₋₂₁K14Ac (green) and double deacetylation H3₁₋₂₁ (red). **b)** Time-course using a MALDI-TOF-MS activity assay for SIRT1 against selected acylated H3 peptides, showing either the single deacetylation product or the double deacetylation product. Assay conditions were SIRT1 (0.75 μ M), NAD⁺ (50 μ M), peptide (10 μ M) incubated at 37°C, run in triplicate. Results are given as an mean $\pm$ SD of percentage conversion to product for either single or double deacetylation products for three experimental replicates.

Comparative activity assays were run with SIRT1 (0.75 μ M), NAD⁺ (50 μ M) and selected peptide (10 μ M) over 15 min at 37°C at fixed time points (Figure 2.17). This showed that

H3₁₋₂₁K9Ac deacetylation (~85% after 15 min) was more efficient than H3₁₋₂₁K14Ac deacetylation (~25% after 15 min). H3₁₋₂₁K9AcK14Ac showed ~75% single deacetylation and ~15% double deacetylation likely corresponding to rapid deacetylation of K9Ac and slow deacetylation of K14Ac. These results indicated that H3K9Ac was a better substrate for SIRT1 than H3K14Ac.

SIRT1 has been reported to deacylate peptides with acyl PTMs other than Ac, including hexanoyl, decanoyl and myristoyl PTMs.³³⁵ Other acetyl mimics such as thioacetyl-lysine were also reported to inhibit SIRT1 activity.³³⁶ hST2, a homologue of SIRT1, was reported to deacylate propionyl, α -hydroxyacetyl and mono-, di- and tri-fluoroacetyl lysine peptides at significantly slower rates than acetyl lysine peptides.³³⁷ To study the diversity of acyl PTM substrates for SIRT1, H3₁₋₂₀K9AcOH (synthesised by Dr Roman Belle, Kawamura group) was tested against SIRT1 using the MALDI-TOF-MS activity assay. No deacetylation of the peptide by SIRT1 was observed with the conditions used (Figure 2.17b). This mirrored the reported activity for hST2 and showed selectivity of the SIRT1 protein produced for acetylation modifications.

2.5 Discussion

2.5.1 Merits of the MALDI-TOF-MS activity assay

Many of the frequently used assays for measuring HAT activity present disadvantages in terms of safety (radioactive materials), sensitivity/scale (fluorometric thiol-scavenging assays), versatility (antibody-based assays) or complicated assay set-up (FRET, fluidic mobility). A MALDI-TOF-MS activity assay was deemed to solve these issues (Sections 2.1.2, 2.2.3)

The MALDI-TOF-MS activity assay used in this chapter is versatile: it can be applied to HAT and HDAC enzymes, along with other types of writer or eraser enzymes.³³⁸ It was shown that acetylation of the target peptides did not affect the peptide concentration to signal intensity relationship in MALDI-TOF-MS acquisition (Figure 2.5a), allowing for quantification of enzyme activity using the relative intensities of parent and product peptides for both HATs and HDACs. Low peptide concentrations were used (10 μ M), and the method allows for small-scale reactions with 0.5-1 μ L reaction mixture needed for analysis at each time point. The MALDI-TOF-MS instruments used have very high accuracy with sub-Dalton resolution achieved for peptides, with isotopic peaks clearly resolved.

2.5.2 Characterisation of PCAF HAT and HAT-BRD proteins using MALDI-TOF-MS

PCAF HAT has previously been well characterised (crystal structure,⁹⁵ acetylation mechanism and kinetics,²⁸⁵ substrate selectivity,^{114–117} cofactor selectivity,^{329,330} inhibitor development^{85,266,271,272}). The aim of the work carried out in this chapter was to consolidate different aspects of PCAF HAT characterisation, as the information is

distributed between multiple studies and was collected using a variety of techniques, leading to inconsistencies. In addition, the HAT-BRD construct of PCAF has not been reported or previously characterised. Certain studies have used the full length PCAF protein (particularly in cell work and *in vivo*, where the protein is shown to be expressed, but no protein purification is necessary)³³⁹, but in such cases, it is not always possible to tell if observed PCAF activity is a result of interactions between HAT/BRD or HAT/N-terminal domain or HAT/multiprotein complexes. Using a single technique to study various aspects of PCAF activity for both the HAT and the HAT-BRD constructs can help produce a clearer, global picture of PCAF activity and identify possible roles for the BRD.

2.5.2.1 Exploration of site selectivity on H3₁₋₂₁ peptides for PCAF HAT and HAT-BRD

PCAF HAT is known to preferentially catalyse the acetylation of H3K14 and H4K8 with lower levels of activity at H3K9. The MALDI-TOF-MS assay developed showed single, double and triple acetylation of H3₁₋₂₁ peptides by PCAF HAT and HAT-BRD. Triple acetylation of H3₁₋₂₁ peptides was only observed with very long reaction times (5h, vs <1h for identification of 1Ac and 2Ac products, Figure 2.7). MALDI-TOF/TOF-MS was used as a MS/MS technique to identify the acetylation sites of PCAF on H3₁₋₂₁ peptides. H3K14 was identified as the first acetylation site, followed by acetylation of H3K9 and H3K18 (Section 2.2.4.1, Figure 2.8 and Figure 2.9). As both PCAF HAT and HAT-BRD constructs showed the same selectivity, this indicated that the BRD domain of PCAF did not affect the site selectivity of PCAF HAT on unmodified H3₁₋₂₁ peptides. PCAF HAT-BRD showed lower specific activity for H3₁₋₂₁ compared to PCAF HAT, which might be indicative of slightly different binding modes for the two constructs (Figure 2.6).

Bromodomain binding chemical probes which bind selectively to the BRD of PCAF were obtained and tested in the MALDI-TOF-MS activity assay. PCAF HAT-BRD showed the same levels of single and double acetylation with or without bromodomain chemical probes confirming that binding between the BRD and substrate does not affect PCAF HAT activity or site selectivity on H3₁₋₂₁ peptides (Section 2.2.4.3, Figure 2.13). Peptide acetylation is a simplified model for biological activity of HATs, not taking into account the context and complexity of cellular interactions. However, L-Moses reportedly did not affect PCAF HAT activity in cells (Brennan group, unpublished results), which is consistent with results obtained in the MALDI-TOF-MS assay.

PCAF and its homologue GCN5 have similar activity and selectivity for histones (Chapter 1). For GCN5, an MS study locating acetylation sites on H3 tails concluded that H3K9 acetylation was dependent on H3K14 acetylation. This was based on identification by MS of H3K14Ac and H3K9AcK14Ac peptides but not H3K9Ac peptides, after digestion of H3 protein treated with GCN5 HAT.³⁴⁰ None of the results obtained using the MALDI-TOF-MS assay could be used to confirm or disprove whether H3K9 acetylation by PCAF was dependent on a pre-existing H3K14Ac PTM. PCAF acetylation at H3₁₋₂₁K14 was clearly favoured over H3₁₋₂₁K9 acetylation; the H3₁₋₂₁ peptide was fully converted to a 1Ac product within 10 minutes under the assay conditions (Figure 2.7), and the major species identified by MS/MS in the 1Ac peak was H3₁₋₂₁K14Ac. However, this does not mean that there was no H3₁₋₂₁K9Ac peptide species in the 1Ac peak, as the MS/MS software only indicates which is the major species at a given mass, based on fragment ion identification. As the BRD chemical probes did not inhibit 1Ac or 2Ac activity of PCAF HAT-BRD, there is no clear link between H3K9/H3K14 acetylation and BRD involvement

in directing enzymatic activity. To identify whether acetylation at these two positions is independent, alanine and/or arginine substitution peptides would have to be tested.

The acetylation of H3₁₋₂₁ peptides preferentially at H3K14 over H3K9 (and H3K18 which is acetylated to a much smaller extent) shows that the PCAF HAT and HAT-BRD enzymes produced are site selective (preferentially acetylating these positions without affecting H3K4 under the assay conditions), and that substrate sites are acetylated at different rates. The mechanisms leading to site selectivity and different rates of acetylation for PCAF are poorly understood. Because there is no consensus sequence for PCAF site selectivity, selectivity and activity are thought to be governed by a combination of loose sequence preferences, binding of histone residues outside of the active site of PCAF HAT, and possibly bromodomain binding preferences.^{50,105,287} A better understanding of PCAF HAT and HAT-BRD selectivity can be achieved by a larger screen of PCAF substrates (Chapter 3).

2.5.2.2 Exploration of Acyl-CoA co-factor selectivity for PCAF HAT and HAT-BRD

Acetylation is a common and well characterised PTM found on lysine residues on all histone tails. Alternative acyl modifications such as propionylation, butyrylation, crotonylation, malonylation, succinylation and formylation have also been identified (Figure 1.1).^{17,63} The biological roles of these PTMs is of interest, as H3K14Pr and H3K14Br have both been associated with active transcription sites, with H3K14Pr shown to stimulate transcriptional activity *in vitro*.³³⁰ Propionylation (Pr) and butyrylation (Bu) activity has been shown for PCAF HAT on histone peptides, with butyrylation being 100 fold slower than propionylation.³²⁹ Acyl-CoA selectivity of full length PCAF and GCN5 *in vitro* and in cellular models is less clear. A radiometric assay has been used to show H3

propionylation and butyrylation activity against calf thymus histones catalysed by PCAF and GCN5 *in vitro*. However, *in vitro* assays against histone octamers showed only propionylation activity for GCN5. In cells, a combination of PCAF and GCN5 knockdown led to decreased H3K14Pr levels without affecting H3K14Br levels, but the extent of contribution for each protein was unclear.³³⁰

In cells, levels of acylation PTMs are linked to metabolic activity as well as epigenetic enzyme activity, with metabolism affecting cellular levels of available cofactor for the enzymes. The metabolic regulation of acetyltransferase activity by endogenous acyl-CoAs has been explored, reporting decreasing binding affinities of PCAF towards acetyl > propionyl > palmitoyl > butyryl > malonyl > succinyl > crotonoyl acyl-CoA forms.³⁴¹

An expanded panel of acyl-CoA cofactors was tested against PCAF HAT and HAT-BRD with MALDI-TOF-MS (section 2.2.4.2, Figure 2.12). In this assay the two PCAF constructs showed equal activity levels to each other, indicating that the BRD does not affect cofactor selectivity. Butyryl-CoA gave only trace levels of Bu product under the assay conditions (Figure 2.11). E-but-enoyl CoA and isobutyryl-CoA, not previously tested against PCAF HAT, gave no product. CoA derivatives with longer acyl chains also gave no product, except for succinyl-CoA which was shown to non-enzymatically succinylate the H3₁₋₂₁ peptide. It was reported that palmitoyl-CoA bound more strongly to PCAF HAT than butyryl-CoA.³⁴¹ However, whilst trace levels of butyrylated H3₁₋₂₁ peptide were observed, no product was seen with palmitoyl-CoA. This indicates that obstruction of the enzyme active site is likely a bigger factor in the inhibition of enzyme activity than

the binding affinities of the acyl-CoA variants, with longer acyl chains as well as branching and unsaturations of the acyl chain negatively affecting enzyme activity.

2.5.2.3 Testing of HAT inhibitors against PCAF HAT and HAT-BRD

Biological studies to determine the role of HATs using currently available chemical probes have been reported. However, there are inconsistencies in results between *in vitro*, cellular, and *in vivo* data due to the involvement of HATs in protein complexes, and due to overlap in activity between different HAT and HDAC enzymes. Interactions of HATs with alternative cell signalling systems can also alter their activity, leading to very different results between cell lines where the cellular context is different (*e.g.* different tissue origins).²⁵⁰ In addition, inconsistency in results of studies with available chemical probes can be due to the lack of selectivity. For example, C646, the best available p300 HAT inhibitor (K_i 0.32 μ M) until 2017, was shown to increase rather than decrease lysine acetylation in murine macrophages due to a previously unknown inhibition of HDACs 1, 3, 6 and 8 by C646 (K_i 7-25 μ M). C646 was reported to reduce pro-inflammatory gene expression, however due to lack of selectivity, the exact role of p300 in this context is still in question.³⁴²

Few selective and potent chemical probes with inhibition of HAT activity have been identified. Two of the more successful HAT inhibitors, A-485 and B026, target p300 and CBP HAT enzymes selectively over PCAF HAT.^{261,262} An overview of the chemical probes developed to target PCAF, GCN5, p300 and/or CBP was presented in Chapter 1 (Section 1.5.2). These reported chemical probes have been tested for binding or inhibition of HAT activity in different studies using varied assays, and have not been consistently screened for activity against the same form of PCAF (truncated or full length

protein constructs) or the same form of target (peptides, full histones, nucleosomes, *in vitro*, cell studies, animal models, *etc.*). This means that reported binding and inhibition data is not always comparable between chemical probes (*e.g.* the measured IC₅₀ values of anacardic acid against PCAF vary from 5 μ M to 2 mM using radiometric, immunosorbent and colorimetric assays).¹¹⁹

Certain inhibitors have been tested against GCN5 but not PCAF, assuming identical inhibitory activity for both these enzymes. Whilst PCAF and GCN5 show very close sequence and structure homology, they sometimes diverge in their biological roles.¹⁰⁰ Chemical probes and inhibitors of HAT activity which are selective for one enzyme over the other would be a useful biological tool to better understand their differences. Testing reported HAT inhibitors developed to target p300, PCAF or GCN5 with a single screening method against PCAF HAT and HAT-BRD gives an opportunity to rank them in terms of inhibitor potency, as well as allowing for better understanding the role of the BRD of PCAF. Commercially available HAT inhibitors were therefore purchased and tested using the MALDI-TOF-MS activity assay developed (Section 2.2.4.4, Figure 2.14, Figure 2.15).

Most of the inhibitors tested against PCAF HAT/HAT-BRD showed expected results. CTPB, curcumin and A-485 which are reported to be selective for p300 over PCAF, showed no inhibition of PCAF activity. Results for curcumin against PCAF activity were quite variable due to poor solubility of the compound (Appendix Figure B.1). Anacardic acid, C646, garcinol and embelin, which are reported inhibitors of PCAF activity, all showed inhibition of PCAF HAT activity in the MALDI-TOF-MS assay (Figure 2.15). However, some of the inhibitor profiles observed differed from reported results.

- Inactive inhibitors which diverge from reported results

The activating effect of SPV-106 on H3 acetylation was observed in cell extracts and was attributed to PCAF based on the loss of H3 acetylation when SPV-106 and anacardic acid were used in combination (however, anacardic acid also targets other HATs such as Tip60).²⁴⁹ The activating effect of SPV-106 on PCAF activity has since been reported with H4 peptide substrates.³⁴³ SPV-106 was also shown to increase H3K9 acetylation in mice with links to wound repair, and increase H3K9Ac and H3K14Ac levels in cells from diabetic patients.³⁴⁴ SPV-106 showed no activation or inhibition of PCAF acetylation activity in the MALDI-TOF-MS assay (Appendix Figure B.1). The lack of inhibition observed may be due to the use of H3₁₋₂₁ peptides, whilst reported results use H4 peptides or cellular data. SPV-106 binding to PCAF was not clearly established in cellular data or in *in vitro* data, so the reported effects of SPV-106 might not be the result of binding to the PCAF HAT or BRD domains, accounting for the lack of observed activity with the PCAF constructs used here.

Butyrolactone 3 and CPTH2 are both reported GCN5 inhibitors. Butyrolactone 3 was reported to inhibit GCN5 activity against purified core histone proteins with little inhibition observed for CBP.²⁶³ It was further shown that butyrolactone 3 affected GCN5 interactions with non-histone protein in cells.³⁴⁵ In *Arabidopsis thaliana* butyrolactone 3 decreased H3K14Ac levels, but the HAT responsible was not identified.³⁴⁶ No reports of butyrolactone 3 being tested against PCAF were found. CPTH2 was shown to inhibit H3 acetylation by GCN5, but was not tested against PCAF or p300/CBP. It was compared with anacardic acid for efficacy showing less efficient inhibition.²⁶⁴ Its derivative CPTH6 was shown to be selective for GCN5 and PCAF over p300/CBP with similar levels of

inhibition of GCN5 acetylation activity as butyrolactone 3.²⁶⁵ CPTH2 was subsequently shown to affect H3K9Ac levels in *Polyandrocarpa misakiensis*, which has an isoform of GCN5.⁷⁷ Neither butyrolactone 3 or CPTH2 showed any inhibition of PCAF HAT or HAT-BRD acetylation activity at concentrations of 100-200 μ M (Appendix Figure B.1). This indicated that these inhibitors are either selective for GCN5 over PCAF, or that they are not potent enough to show inhibitory activity at these concentrations in the MALDI-TOF-MS screen. This second explanation is likely for CPTH2 as CPTH6 showed no selectivity for GCN5 over PCAF and has been shown to be active only at concentrations over 800 μ M. For butyrolactone 3, an IC_{50} of 100 μ M was measured using a radioactive assay with core histone proteins.²⁶³ The use of histone peptides instead of histone proteins may account for the lack of activity observed, as the binding mode of the protein may be affected by the change in substrate. Testing GCN5 activity in the same assay format as PCAF would be needed to conclusively show inhibitor selectivity.

- Overview of Active inhibitors

Anacardic acid is often used as a reference compound in the development of new inhibitors for PCAF. A number of different assays have therefore been used to measure its inhibitory activity against PCAF, giving IC_{50} measurements between 5-1000 μ M.^{242,247,248,333} IC_{50} measurements can depend on a variety of factors including the enzymes and substrates used, buffer composition, and cofactor concentrations (important in bi-substrate enzymes).³³² The IC_{50} values observed in the MALDI-TOF-MS screen were 88 μ M for PCAF HAT and 120 μ M for PCAF HAT-BRD, which is towards the lower range of reported IC_{50} values. The difference in IC_{50} measurements between the HAT and HAT-BRD might indicate a difference in binding between the two constructs.

However this could also be due to variation in measured IC_{50} values related to differences in product conversion levels for the assay.³⁴⁷ The steep curve observed in the IC_{50} assay can be indicative of non-specific inhibition, tight binding or irreversible enzyme inhibition.³³² Anacardic acid was previously identified as a non-specific inhibitor of p300, however its binding mode to PCAF has not been elucidated.²⁴² Anacardic acid suffers from lack of selectivity and lack of membrane permeability, with efforts made towards developing derivatives to solve these problems.^{85,119,248}

Embelin was shown to inhibit the activity of full-length PCAF with purified core histones with an IC_{50} of 7.2 μ M, using a radioactive filter binding assay.²⁵¹ The higher IC_{50} values observed in the MALDI-TOF-MS assay (113 μ M for HAT and 162 μ M for HAT-BRD) may be due to the different materials or assay format used. Embelin was reported to selectively inhibit PCAF over the HATs p300 and Tip60.²⁵¹ However, Embelin is a natural product, and has been shown to modulate the activity of numerous proteins involved in apoptosis signalling cascades and other cancer related biological functions, and is therefore non-specific for PCAF inhibition.²⁵⁴

C646 was discovered through virtual screening as an inhibitor of p300. It was reported to be selective for p300, with a concentration of 10 μ M C646 giving less than 10% decrease in acetylation activity for PCAF and GCN5 against a H4 peptide in a fluorometric assay.²⁵⁸ The constructs used for PCAF and GCN5 screening were not clearly stated, making comparisons to these reported results less significant. In the MALDI-TOF-MS assay, C646 was shown to inhibit PCAF HAT (IC_{50} 114 μ M). A concentration of 200 μ M C646 inhibitor was not sufficient to fully inhibit PCAF HAT activity. This can be due to partial enzyme inhibition by the compound or can reflect low solubility of the compound

at high concentrations.³³² For PCAF HAT-BRD, an increase in activity was observed at concentrations 175-200 μ M C646. This may be due to stabilisation of the protein by C646 at higher concentrations. This could be tested using a thermal shift assay, to see if PCAF HAT-BRD is more stable in the presence of C646.^{348,349}

Garcinol was shown to inhibit activity of both p300 and PCAF against purified core histones in a radioactivity assay with measured IC_{50} values of 7 μ M and 5 μ M respectively.²⁴³ Garcinol was later shown to decrease PCAF expression and H3K9 levels in renal cells.³⁵⁰ Garcinol was shown to act as a competitive inhibitor of H3 binding to the enzymes.²⁴³ In the MALDI-TOF-MS activity assay with 47-100 μ M garcinol, PCAF activity increased with increasing inhibitor concentration. However, for concentrations between 150-200 μ M garcinol, enzyme activity decreased with increasing inhibitor concentration, giving an IC_{50} of 160 μ M against PCAF HAT. For PCAF HAT-BRD, garcinol did not show sufficient inhibition with concentrations up to 200 μ M for an IC_{50} measurement to be made with the MALDI-TOF-MS assay. In the reported radioactivity assay, this dual effect based on garcinol concentration was not observed, with p300 and PCAF enzyme activity decreasing with garcinol concentrations between 1-100 μ M. Pre-incubation of garcinol and PCAF HAT/HAT-BRD may have had some role on the differences in activity observed with lower inhibitor concentrations in the MALDI-TOF-MS assay, as PCAF activity was reportedly affected by preincubation with either AcCoA, CoA or 0.01–0.1% BSA.^{290,351} Under the right conditions (*e.g.* low concentrations of chemical probe), the enzyme might be stabilised by pre-incubation, leading to an increase in activity. This effect is highly dependent on conditions and could be lost at high compound concentrations.

- Conclusions

Although there were some disparities between reported and observed compound activities, these can be largely explained by differences in materials used, assay type and assay sensitivity which all affect reported IC₅₀ values. The goal of this screen was to compare known and suspected (if only previously tested against GCN5) commercially available PCAF inhibitors to gain insight into the difference in activity between PCAF HAT and HAT-BRD, and to rank the inhibitors in terms of potency. This helps characterise reported inhibitors of PCAF HAT activity and validates the MALDI-TOF-MS assay as a method for identifying inhibitors of PCAF activity. These results provide an assay structure and positive inhibition controls for validating any novel inhibitors of PCAF acetylation activity developed.

Anacardic acid was shown to be the most potent PCAF inhibitor of both PCAF HAT and HAT-BRD, followed by embelin, C646 and garcinol. Anacardic acid, embelin and garcinol were shown to be less potent against HAT-BRD, with C646 showing activation of HAT-BRD activity at high concentrations. The difference in activities against PCAF HAT-BRD might indicate a structural difference in the HAT domain which affects inhibitor binding. The differences in inhibitor profiles might reflect different binding sites for the inhibitors which might be differently affected by the BRD. Other commercially available inhibitors tested showed no activity of PCAF HAT or HAT-BRD at concentrations of up to 200 µM (Figure 2.14, Appendix Figure B.1). Concentrations above 200 µM were not tested as the inhibitors were dissolved in DMSO and high DMSO concentrations affect crystallisation of MALDI matrix. In addition, these would not be potent inhibitors and therefore, less relevant.

Anacardic acid was chosen as a control for PCAF HAT and HAT-BRD inhibition, and was subsequently used when testing cyclic peptides for inhibition of PCAF HAT and HAT-BRD using the MALDI-TOF-MS assay (Chapter 5).

2.5.3 Characterisation of SIRT1 protein using MALDI-TOF-MS

SIRT1 was selected for study as an opposing enzyme to PCAF, due to their overlapping biological roles and target selectivity.⁸⁴ The MALDI-TOF-MS activity assay was shown to be equally effective in measuring deacetylation of H3₁₋₂₁K9Ac, H3₁₋₂₁K14Ac and H3₁₋₂₁K9AcK14Ac peptides by SIRT1 as measuring PCAF acetylation activity on H3₁₋₂₁ (Section 2.3.2, Figure 2.17). The assay was used to compare SIRT1 activity against H3₁₋₂₁K9Ac, H3₁₋₂₁K14Ac and H3₁₋₂₁K9AcK14Ac, showing a preference for deacetylation at H3K9Ac over H3K14Ac. The substrate selectivity for acetylated peptides was explored with H3₁₋₂₀K9AcOH showing no deacylation by SIRT1. This showed substrate selectivity of the SIRT1 enzyme under the assay conditions used, in line with reported selectivity, making it a good construct for further substrate screening (Chapter 3).

2.6 Future work

To give a more complete picture of PCAF activity, it would be useful to have GCN5 recombinant protein available for parallel screening. Studying the distinct roles of these enzymes would be helpful to better understand regulation of acetylation, as well as for development of selective chemical probes and inhibitors. Currently, the lack of selectivity of chemical probes used to target PCAF, GCN5, p300 and CBP in cellular and *in vivo* studies leads to results in which the roles of different HAT enzymes cannot be clearly distinguished.

Alternative activity and binding assays would be useful for validation of obtained data for both PCAF and SIRT1. Alphascreen™ is a good assay to develop for binding studies due to its sensitivity and small scale. Fluorometric assays (enzyme coupled or quantification of side products nicotinamide or CoA) are a good alternative to MS for activity assays as they can be continuous assays. However, issues with scale and sensitivity were encountered with preliminary efforts to set up an alternate activity assay (results not shown).

2.7 Conclusion

PCAF and SIRT1 epigenetic enzymes are interesting targets for parallel study as they have opposite biological effects at target sites. A MALDI-TOF-MS activity assay was developed to characterise these enzymes. MALDI-TOF-MS is well suited to characterising PCAF and SIRT1 activity of H3 peptides as it is a soft ionisation method allowing MS of peptides without fragmentation and with sufficient resolution to distinguish a ± 42 Da mass shift with resolved isotopes, whilst also having the capacity for MS/MS when fragmentation is needed for characterisation purposes.

Previous substrate and cofactor characterisation of PCAF and SIRT1 has been undertaken through many different reported works using different assays, methods and materials. Some of these reported results have been validated and expanded on here, using the MALDI-TOF-MS assay developed, in order to broaden the enzyme characterisation of PCAF HAT and SIRT1. The previously untested HAT-BRD construct of PCAF has been screened with PCAF HAT to help understand the biological role of the bromodomain of PCAF. The results obtained have shown that MALDI-TOF-MS and MALDI-TOF/TOF-MS are accurate and versatile techniques which can be widely used to

study enzyme activity and selectivity of epigenetic writer and eraser enzymes such as HATs and HDACs. The preliminary characterisation work carried out in this chapter constitutes groundwork for a larger substrate scope screen (Chapter 3, Chapter 4) and for inhibitor development (Chapter 5).

There is currently a lack of selective and potent chemical probes targeting HAT enzymes. Of the existing commercially available small molecule chemical probes targeting PCAF, anacardic acid was identified as the most potent, inhibiting PCAF HAT and HAT-BRD with IC_{50} values of 88 μ M and 121 μ M respectively. However, this inhibitor is also known to target p300/CBP and GCN5 HATs and does not penetrate cell membranes, showing that there is a need for further development of chemical probes targeting PCAF HAT. In this chapter, a larger focus has been placed on PCAF characterisation over SIRT1 characterisation, in view of developing cyclic peptide chemical probes to targeting PCAF HAT.

Chapter 3 Development of a rapifleX MALDI-TOF-MS high throughput screen for identification of crosstalk between post translational modifications mediated by epigenetic enzymes

3.1 Introduction

3.1.1 Activity profiles of PCAF and SIRT1 epigenetic enzymes

PCAF is known to catalyse the acetylation of several sites on histone (H3K14, H3K9 and H4K8) and non-histone proteins (68 identified as of November 2018)²⁰ including p53, and HIF1 α (Section 1.2.2).^{20,119} This indicates that there is flexibility in the HAT active site and the substrate binding pocket. It is likely that PCAF has a much larger scope of histone substrates than is currently understood, as has been shown for homologous HAT enzyme GCN5 which can catalyse the acetylation of six different lysines on histone H3 alone.³⁴⁰ Similarly, SIRT1 is known to catalyse the deacetylation several sites on histone (H3K9 and H3K14) and non-histone proteins (107 identified as of November 2018),²⁰ including p53, and HIF1 α (Section 1.3).^{20,182} Identification of substrate sites for enzymatic activity is necessary to properly understand the biological role of epigenetic proteins. This can be achieved through systematic screening and detailed biochemical characterisation of interactions between histone tails and epigenetic enzymes.

Many epigenetic proteins contain a combination of writer, reader and eraser domains, and the selective contributions of each domain to the overall activity of the protein is not always understood. The full length PCAF protein consists of a HAT domain, a BRD and an E3 ligase domain.^{95,352} Whilst the BRD of homologous protein GCN5 has been

shown to alter its acetylation activity,¹²² the relationship between the BRD and HAT domains of PCAF is not well understood (Section 2.1.1). Epigenetic proteins can also associate to form complexes with multiple roles (*e.g.* SAGA complex with GCN5/PCAF). In the case of GCN5, association with the SAGA protein complex alters the HAT enzyme selectivity of GCN5 for histones compared with isolated GCN5.¹¹³ This shows that it is necessary to consider the effect of experimental context (*e.g.* protein boundaries, type of substrate used, lack of interaction with cellular environment) on observed results, and to consider designing assays which give a more holistic view of enzyme activity.

3.1.2 Crosstalk between combinatorial epigenetic PTMs

Histone tails can have multiple epigenetic PTMs at one time, referred to as combinatorial PTMs.^{19,36} The known PTMs for histone sequences have been surveyed, but new PTM types and sites are regularly identified (Section 1.1.1.1, Figure 1.1).¹⁷ These PTMs are often dynamic in a cellular context, with many epigenetic writers, readers and erasers having single or multiple substrate PTM sites on histone tails. This means that at any one time, the histone tails where PCAF and SIRT1 are active will have PTMs adjacent to the enzyme substrate site, which has been shown to affect the activity and selectivity of these enzymes.^{47,353} This is referred to as crosstalk between epigenetic marks, mediated by epigenetic enzymes, and constitutes a facet of cell signalling mechanisms (Section 1.1.1.2, Figure 1.2).^{20,37,46,354}

Crosstalk between epigenetic PTMs is not well understood as most *in vitro* studies on epigenetics concentrate on a specific PTM type and site without taking into account the state of the rest of the histone tail. This is because accounting for surrounding PTMs is difficult, as the state of the histone tail is context dependent, and there are too many

possible combinations of epigenetic PTMs to replicate when routinely studying a specific PTM *in vitro*. Cellular and animal studies are naturally affected by combinatorial epigenetic PTM crosstalk because cells are a dynamic environment. In addition, because combinatorial PTMs are context dependent they can have different outcomes in different environments (*e.g.* different tissues) which complicates our understanding of crosstalk interactions.^{37,354}

As sequence specificity is not a major factor in governing PCAF and SIRT1 substrate selectivity, other factors are likely to govern their activity levels. Crosstalk between combinatorial PTMs and PPIs (*e.g.* between the PCAF HAT and BRD domain, or between PCAF HAT and the SAGA complex) are two factors likely to play a major role in regulation of PCAF and SIRT1 selectivity.^{10,18-20} In such cases, studying combinations of PTMs in parallel with combinations of epigenetic proteins with different boundaries (*e.g.* comparison of PCAF HAT and HAT-BRD constructs) is necessary to understand the crosstalk between the PTMs and the proteins as well as the crosstalk within proteins or protein complexes.³⁵⁴ Systematic screening of PCAF HAT, PCAF HAT-BRD and SIRT1 against a panel of histone peptides with PTM modifications can therefore help better understand the selectivity of the enzymes. As PCAF and SIRT1 have opposing biological roles, identifying overlap or differences in regulation of these enzymes by PTMs can be informative as to the role of PTM regulation on maintaining acetylation levels in cells.

A few experimental or computational methods have been developed to systematically screen the effect of PTM crosstalk on epigenetic proteins. Kornacki *et. al.* screened a positional scanning peptide array of 96 peptides using SAMDI-MS, and identified H3R8 and H3K14 crosstalk in PCAF HAT activity.⁴⁷ This is an effective method, but

immobilisation of peptides on a membrane may affect enzyme activity. This method was not developed for HTS. Garske *et. al.* synthesised a one bead one compound (OBOC) library of H4₁₋₂₁ peptides covering all known PTMs randomly combined (800 library members). The binding of biotinylated hJMJD2A was assessed through activity of a streptavidin-conjugated alkaline phosphatase which resulted in a blue colouring of beads where hJMJD2A was bound.³⁵⁷ This method relied on manually identified coloured beads for selection of hit peptides, followed by peptide cleavage and MALDI-TOF/TOF-MS for sequence identification, which results in a time consuming process. The method was developed to identify reader domains rather than to monitor enzyme activity. Aparicio Pelaz *et. al.* screened libraries of tagged and immobilised semisynthetic nucleosomes with combinatorial PTMs with a fluorescence assay in which PTM identification was carried out through next generation sequencing (NGS) of the tags.³⁵⁸ The activity of the nucleosomes used might however be affected by the synthesis method, and the fluorescence assay was used to identify reader domain binding rather than monitoring enzyme activity. Chandrasekharan *et. al.* monitored effects of histone PTMs in total or nuclear cell extracts using chromatin-double immunoprecipitation and ChIP, which relies on the availability of selective antibodies.³⁵⁹ Guan *et. al.* used stable isotope labelling of amino acids in cells culture (SILAC) in tandem with LC-MS to determine levels of PTMs, which has inherent drawbacks of expensive isotope labelling.³⁶⁰ Computational methods for identifying PTMs have also been developed by Lu *et. al.* and Hang *et. al.* which use known data to develop algorithms for identifying PTM crosstalk.^{361,362}

3.1.3 Aims

The aim of this chapter is to develop and optimise a methodology for high throughput screening (HTS) of histone peptide libraries with combinatorial PTMs against epigenetic enzymes PCAF and SIRT1 using MALDI-TOF-MS.

Specific aims:

- Development of a methodology for high-throughput screening of histone peptide libraries with PTMs using MALDI-TOF-MS
- Optimisation of screening conditions
- Development and optimisation of data processing methods for analysis of results
- Testing of high- and low-throughput MS/MS data acquisition methods using MALDI-TOF/TOF-MS and analysis of MS/MS data
- Analysis of results of the MALDI-TOF-MS HTS performed against PCAF HAT, PCAF HAT-BRD and SIRT1

3.2 Technical information for the rapifleX MALDI-TOF-MS HTS

3.2.1 Choice of a RapifleX MALDI-TOF-MS screen

In Chapter 2, a MALDI-TOF-MS activity assay was developed to monitor activity of PCAF and SIRT1 against H3₁₋₂₁ peptides (Section 2.2.2, Section 2.3.2). This activity assay can be applied more widely to screen for alternative peptide substrates of PCAF and SIRT1. MALDI-TOF-MS is well suited for MS acquisition of peptides with a mass accuracy of <5 ppm for peptide mixtures which can monitor a +42 Da (acetylation) shift accurately with resolution of isotopic peaks. MALDI-TOF-MS is an advantageous screening method because multiple mass shifts can be observed simultaneously, independent of peptide

sequence or site of acetylation, unlike methods relying on antibody binding for monitoring of enzyme activity. The assays are run in solution which removes any issues with immobilisation affecting enzyme activity which can arise with SAMDI-MS.⁴⁷ In addition, the screen developed can be run with libraries of untagged peptides, removing any assay interference from tags. The low volumes required for MALDI-TOF-MS analysis ($\leq 1 \mu\text{L}$ used for spotting), make it a good method for development of a HTS without excessive use of reagents. Although MALDI-TOF-MS of peptides is not as sensitive as radiometric methods, it is more sensitive than fluorometric methods often used to measure HAT activity ($10 \mu\text{M}$ peptide used for data acquisition in Chapter 2).^{289–292}

3.2.2 Histone peptide libraries

In order to explore the effects of combinatorial PTM crosstalk on PCAF and SIRT1 substrate selectivity, four histone peptide libraries with PTMs were purchased from Altabioscience (SETI-IV). The libraries cover 20 mer fragments of Histones H3, H4, H2a and H2b with a combination of phosphorylation (Phos or P), acetylation (Ac or A), methylation (mono-methylation Me1 or 1, di-methylation Me2 or 2, tri-methylation Me3 or 3, asymmetric di-methylation Me2a or M) PTMs at lysine (K), arginine (R), serine (S) and threonine (T) residues. Although the term post translational modification (PTM) is not fully correct for modifications on peptides, as they are chemically synthesised, it is used in this chapter for simplicity, as these peptides are models for histones with PTMs. The library members are conjugated with biotin either at their *N*-terminus with an aminohexanoic linker (biotin-Ahx) or at their *C*-terminus with an aminohexanoic acid-lysine linker (Ahx-Lys-biotin) to allow for immobilisation or recognition with streptavidin linked tools (Figure 3.1).

The combinations of PTMs in the libraries are chemically synthesised combinations which are not necessarily known to be biologically relevant. However, these libraries help study how combinations of PTMs on histones might affect the binding and activity of proteins. The placement and chemical nature of the PTMs can help inform on binding modes, selectivity and structural constraints at the peptide–protein interface.

These commercial libraries are made using an automated peptide synthesis system which allows for parallel synthesis of library peptide in a 96-well plate format. In these systems, there is little optimisation for difficult peptide couplings with non-canonical amino acids (nCAA), as all the couplings are carried out in parallel with the same reaction conditions (coupling time, temperature *etc.*). After peptide synthesis, the peptides are washed, cleaved and no further purification is carried out before the peptides are lyophilised for storage. The quality of the peptides is therefore variable, depending on ease of synthesis for each peptide sequence. Side products are likely to arise due to amino acids not coupling properly, issues with side chain reactivity, bead aggregation and incomplete steps for washing, deprotection or coupling. In addition, the peptides are made on a small scale (200 nmol) and are not quantified. Peptide concentration is therefore dependent on the overall yield of the synthesis and on the quantity of side products produced (affects peptide purity). Variable peptide concentration will affect observed activity of the enzyme and the quality of the MS and MS/MS spectra obtained. The quality of the library peptides in the high-throughput screen was therefore taken into account when analysing the results obtained.

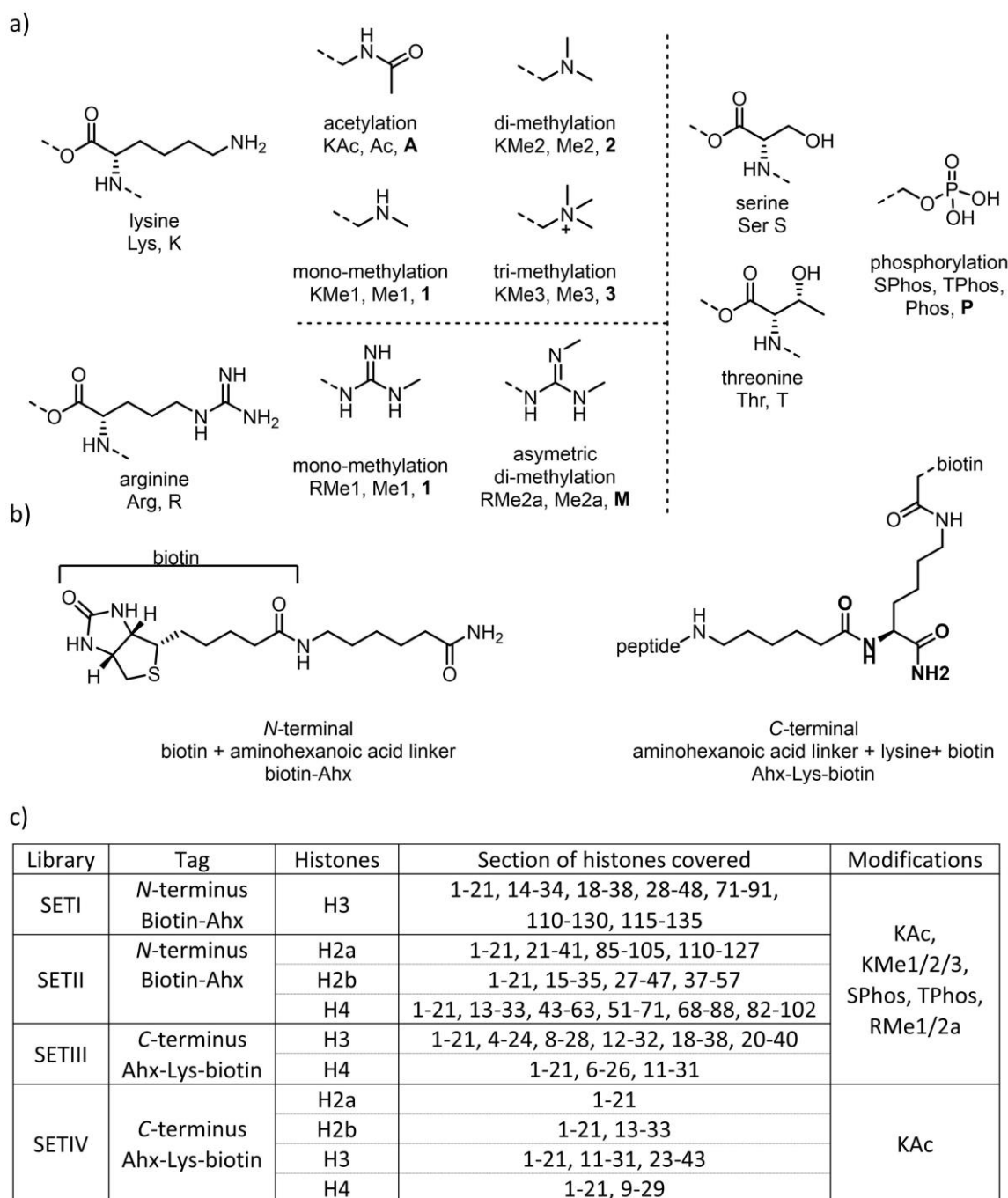


Figure 3.1: Composition of Altabioscience histone peptide libraries SETI-IV.

a) Post translational modifications (PTMs) found in the libraries, with corresponding abbreviations. **b)** Structure of the N- or C-terminal biotin-linker tags used for the library peptides. **c)** Table showing the contents of each library, listing the name, linker type, histone type, sections of histones covered in the library and types of PTMs used.

The libraries were obtained as lyophilised powders and made up with a ratio of purified water and acetonitrile depending on solubility of the peptides. Aliquots of the master plate were made at an assumed peptide concentration of 100 μ M for use in enzyme assays (Materials and Methods 7.7.6). MALDI-TOF-MS spectra for all wells of the library

plates were acquired. The quality of each library was assessed by matching peaks in MALDI-TOF-MS spectra acquired to the masses of peptides listed for the corresponding library wells. Using this method, it was found that SETIV had the lowest library quality with 83% peptides identified whilst for SETI, SETII and SET III 85%, 89% and 95% peptides were identified respectively.

3.2.3 Pre-screen optimisation of assay conditions

The library screening assays were designed as end-point MALDI-TOF-MS assays, with a similar procedure to time dependent assays (Section 2.2.2). Two constructs of PCAF (HAT and HAT-BRD) and a single construct of SIRT1 were screened. In Chapter 2, H3₁₋₂₁ was established as a substrate for PCAF HAT/HAT-BRD and H3₁₋₂₁ K9Ac for SIRT1. Assay conditions were therefore initially optimised with these peptides, and they were used as positive controls in the library screen.

The concentration of peptide to be used in the assay was determined by testing a dilution series (10-0.75 μ M, x 0.75 dilutions, three replicates) of H3₁₋₂₁, H3₁₋₂₁ K9Ac, H3₁₋₂₁ K14Ac and H3₁₋₂₁ K9AcK14Ac peptides with rapifleX MALDI-TOF-MS. Under a concentration of 2 μ M, spectra acquired showed no peaks in the mass range corresponding to the peptides. Over 2 μ M concentrations, clear spectra were acquired with increasing concentrations giving increasing peak intensities for the same acquisition parameters (Appendix Figure A.1). MS/MS using rapifleX MALDI-TOF/TOF-MS on these samples showed that quality data was acquired only for samples with ≥ 10 μ M peptide. The assay was therefore designed to achieve a final concentration of 10 μ M peptide. Because the library peptides are not purified or quantified, unlike the test peptides which were purified using HPLC and quantified using

NMR, it is unlikely that the concentrations of library peptides are accurate (Section 3.2.2). This led to variable quality of MS and MS/MS spectra obtained for the library peptides in the screen.

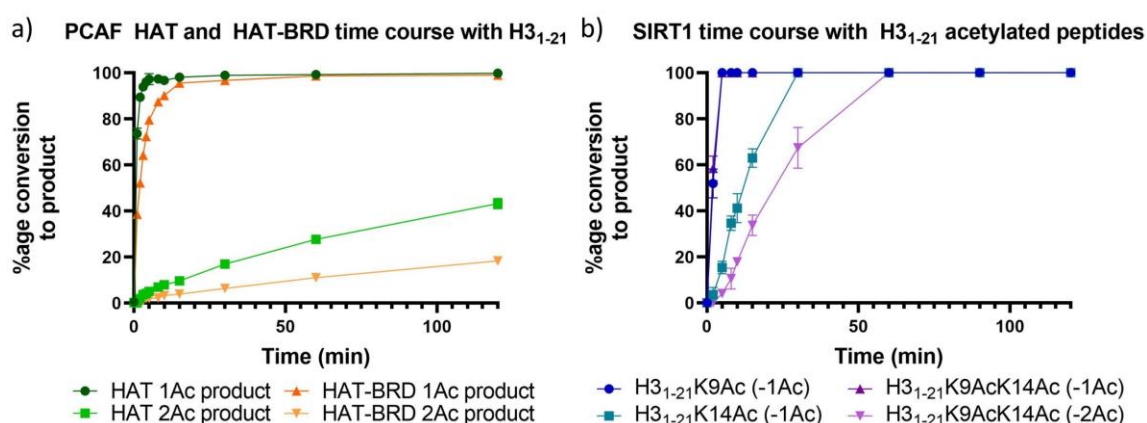


Figure 3.2: Time dependent assays for PCAF HAT, PCAF HAT-BRD and SIRT1 with control substrate peptides to determine end point for rapifleX MALDI-TOF-MS HTS assay.

a) Time course assay run with PCAF HAT/HAT-BRD (2 μ M), AcCoA (100 μ M) and control peptide H3₁₋₂₁ at 37°C with time points over 2 h. Results are shown as mean $\pm$ SD of conversion to product peptide for three experimental replicates. **b)** Time course assay run with SIRT1 (2 μ M), NAD⁺ (100 μ M) and control peptides H3₁₋₂₁K9Ac, H3₁₋₂₁K14Ac and H3₁₋₂₁K9AcK14Ac at 37°C with time points over 2 h. Results are shown as mean $\pm$ SD of conversion to product peptide for three experimental replicates.

Time courses were run to determine the optimal end point for the assay. Coenzyme concentration was set to allow for a full 5 turnovers of the peptide by the enzyme, as PCAF and SIRT1 have both been shown to have multiple substrate sites on the H3₁₋₂₁ peptides. A time course was run with 2 μ M PCAF HAT or HAT-BRD, 100 μ M AcCoA and 20 μ M H3₁₋₂₁ peptide over two hours at 37°C (Figure 3.2a). A rapid conversion to the single acetylation (1Ac) peptide product H3₁₋₂₁K14Ac (100% within 5 min for both HAT and HAT-BRD), the preferred acetylation site for PCAF (Section 2.2.4.1), was initially observed. At 60 minutes partial conversion to the double acetylation (2Ac) product H3₁₋₂₁K9AcK14Ac was observed (30% with HAT, 11% with HAT-BRD). One hour was determined as an optimal end point to distinguish between products acetylated at different rates, whilst also limiting errors due to evaporation. An equivalent time course

was run for SIRT1 with 2 μ M SIRT1, 100 μ M NAD⁺, and 20 μ M H3₁₋₂₁K9Ac, H3₁₋₂₁K14Ac or H3₁₋₂₁K9AcK14Ac (Figure 3.2b). Initial rapid single deacetylation (-1Ac) by SIRT1 was observed for H3₁₋₂₁K9Ac and H3₁₋₂₁K9AcK14Ac peptides (100%, 5 min), followed by deacetylation of H3₁₋₂₁K14Ac (100%, 30 min), and a second deacetylation event for H3₁₋₂₁K9AcK14Ac (100%, 60 min). A one hour end point was therefore also used for screening of SIRT1.

3.2.4 Screen plan and assay methodology

Screening of the library assays was run in a 96-well plate format, to match the format of the library plates, with wells A1 and H12 left available as controls. The full assay procedure can be found in materials and methods 7.7.6. Briefly, the coenzyme and library peptides are transferred to the assay plate, followed by enzyme mixture, and the plate is incubated at 37°C for 1 h. The plate is quenched with one equivalent of 2% formic acid, and the assay mixture is spotted 1:1 with CHCA matrix onto MALDI plates for MALDI-TOF-MS analysis. Analysis is carried out through peak picking and comparison of relative intensities of parent and product peptides. This method was previously established as a low-throughput method by Dr. Joanna Bonnici (Kawamura group) with manual spotting and data analysis. Development of the method into a high-throughput method with automated spotting and data recovery is described herein.

The screening method was first performed on a single library (SETIII) on a 10 μ L scale with PCAF HAT and HAT-BRD using a low-throughput method on a Bruker MALDI microflex machine. This yielded a wide spectrum of activity data across different library peptides, but manual spotting (~3 hrs for 7 plates), low-throughput of data acquisition on the microflex instrument (~30 s per sample) and largely manual data collation and

analysis (> 7 days) made the process very time consuming. Thus, a workflow to improve throughput was needed to screen other peptide libraries, starting with purchased SETII, III and IV Altabioscience libraries. To achieve this, in collaboration with Matthias Trost's lab (Newcastle University), a liquid handling robot (Mosquito LV, SPT Labtech) and a Bruker MALDI rapifleX PharmaPulse™ MALDI TOF/TOF were implemented into the workflow (Figure 3.3). The Mosquito liquid handler has a pipetting range of 25 nL – 1.2 µL and can dispense to 384 wells in ~3 min. The rapifleX has a mass accuracy of ≤ 5 ppm, an effective flight path of >300 cm, and a data acquisition speed of 10 kHz on reflector MALDI-TOF-MS mode (time per spot <1 s). These instruments allow for rapid spotting and data acquisition in MALDI-TOF-MS assays. Spotting on the Mosquito liquid handler and data acquisition using rapifleX MALDI-TOF-MS was carried out by Dr. Maria Dueñas (Trost group, Newcastle University).

The enzyme assays for the high-throughput method were run in a 96 well plate format to preserve continuity with previous results. Three assay replicates (15 µL scale) were run for each enzyme and for a no enzyme control (NEC), which was used as a peptide baseline control (t=0 spectra) and as a quality check for each library well. Quenched assay aliquots (10 µL) were transferred to a 384 well plate suitable for the Mosquito liquid handling robot. Each assay aliquot was spotted (125 nL) in quadruplicate to account for variations in liquid dispensing and spot crystallisation, covering a 1536 spot MALDI plate. A rapifleX MALDI-TOF-MS instrument was used to automatically acquire MS spectra for each of the spots.

The post-acquisition data analysis method needed to be partially automated for the high-throughput screen as large numbers of spectra (approximately 12,000 total) had to

be processed for data analysis. The first part of data analysis is peak picking for each of the spectra acquired using a peak picking script in the Flexanalysis software (Bruker) which makes a peak list for each of the spectra (Section 3.2.5.1, Figure 3.5a). For the second stage of analysis, the intensity for peaks corresponding to parent and product peptides must be recorded and the relative percentage of modified peptide to parent peptide calculated. This was performed by a custom python script written by Dr. Tom McAllister (Section 3.2.5.2, Figure 3.5b-e). Once the data was extracted, Excel was used to tabulate the data for the different enzymes and NEC to allow for comparative analysis between enzymes as well as between library members. The data was then manually analysed, with the NEC and calculated standard deviation (SD) being used to assess data reproducibility.

A second stage of HTS development involved the use of MS/MS to identify the site of enzymatic activity on the library peptides to better understand activity and specificity of the enzymes screened. As the rapliflex MALDI has MS/MS capacity and potential to run high-throughput MS/MS, efforts were made to add this to the HTS workflow (sections 3.5 and 3.6.4).

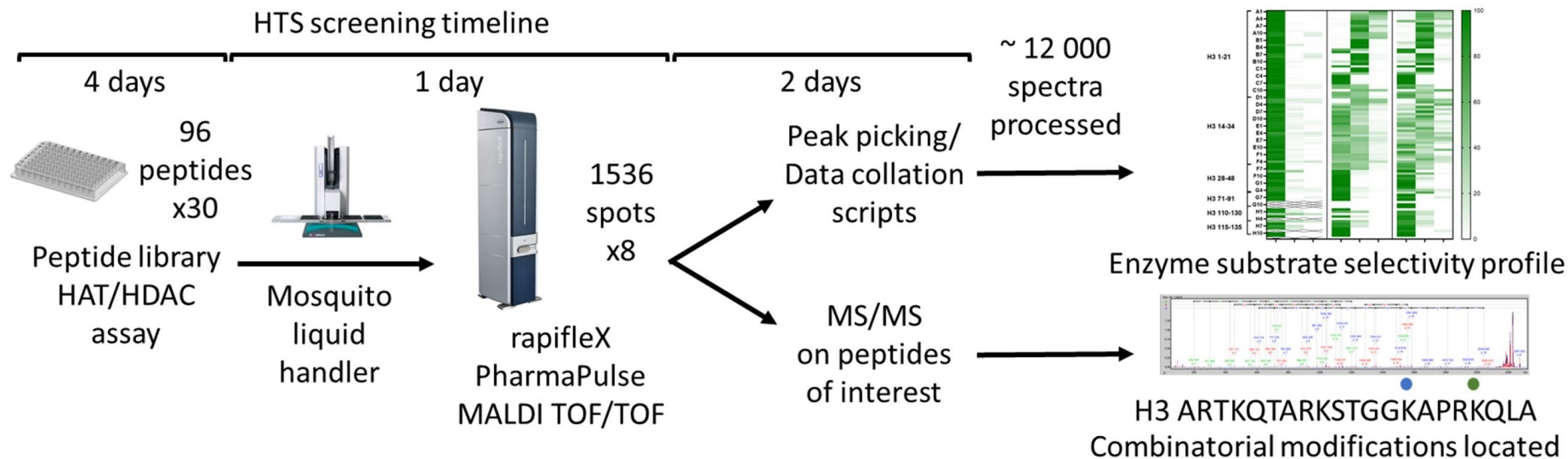


Figure 3.3: Workflow developed for a rapifleX MALDI-TOF-MS high throughput screen.

The assay is run in a 96 well plate format, the assay mixtures are then spotted on 1536 MALDI plates using a Mosquito LV liquid handler. The plates are run with a rapifleX PharmaPulse MALDI-TOF-MS mass spectrometer. Data is extracted using 2 adapted scripts and formatted into tables for analysis. MS/MS can be carried out on peptides of interest using rapifleX MALDI-TOF/TOF-MS for further analysis

3.2.5 Analysis methods for the rapifleX MALDI-TOF-MS screen

3.2.5.1 Peak picking script

Peak picking for the MALDI-TOF-MS spectra was performed on the Flexanalysis software (Bruker) with a customisable script provided by Bruker. The script was modified to suit the mass range needed and parameters (peak picking algorithm, spectra smoothing and baseline subtraction, area vs. intensity, signal to noise ratio) were optimised. For low-throughput analysis of the SETIII library a mass range of 2000 to 3500 Da was used, and for high-throughput analysis of the SETI, SETII and SETIV libraries a mass range of 1700 to 4000 Da was used. The output of this script generated a file containing data for each peak including peak mass, intensity and area (Figure 3.5a).

To properly quantify the parent and modified peptides in relation to each other, the monoisotopic masses of each peptide must be compared. SNAP and centroid peak picking methods were tested. For SNAP peak picking, each isotopic peak is assigned to its corresponding monoisotopic peak with only the mass and intensity of the monoisotopic peak appearing in the output peak list file. In comparison, centroid processing picks each of the isotopic peaks separately. This leads to a much longer peak list with many overlapping peaks, leading to issues with correctly identifying the monoisotopic peptide peak (Figure 3.4a). SNAP was therefore chosen as the peak processing method.

Peak intensity was chosen over peak area for relative quantification of parent and product peptides as it has been shown to be more reliable for comparative analysis.³²⁸

Using peak intensity in time course assays was found to give reproducible data.

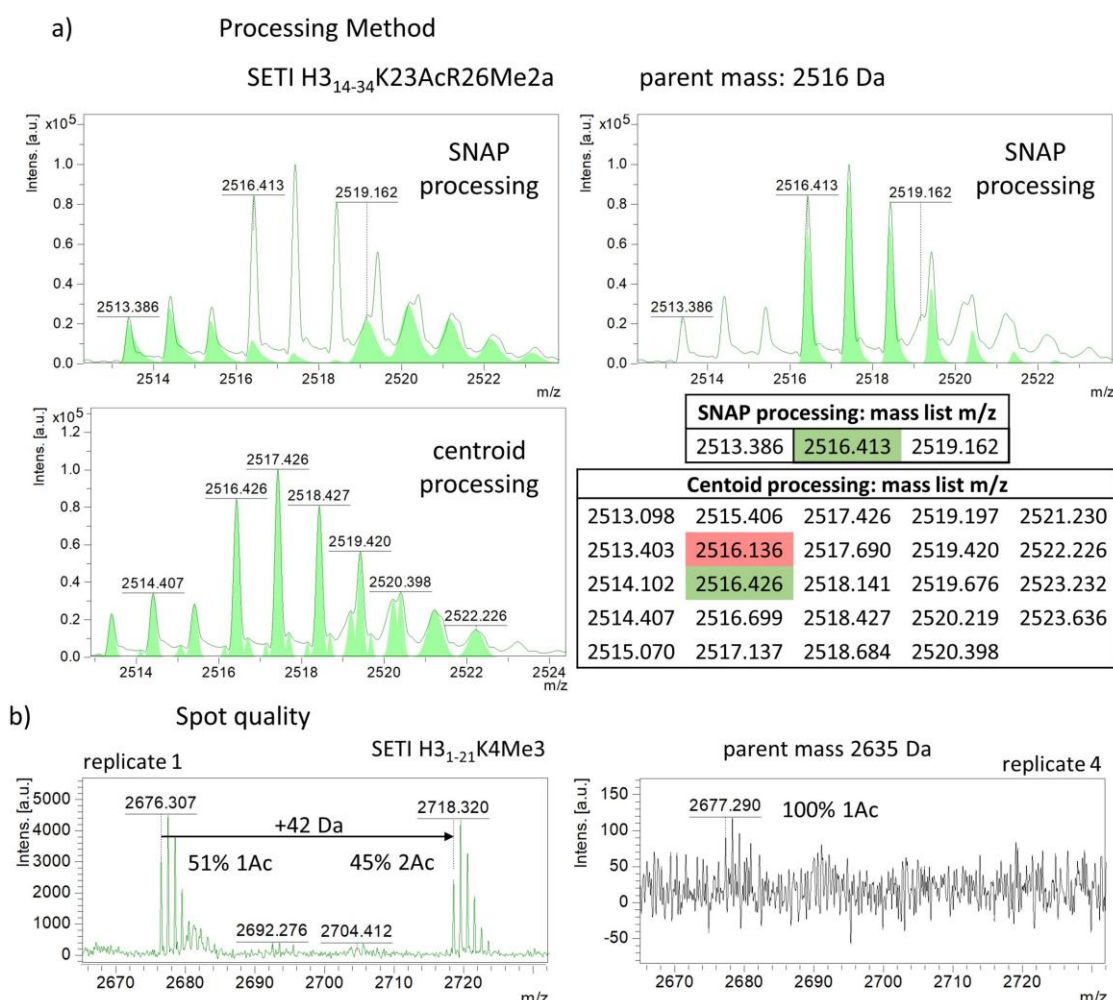


Figure 3.4: Examples illustrating differences in parameters used for peak picking.

a) Examples showing peaks in a section of the rapifleX MALDI-TOF-MS spectrum for SETI H3₁₋₃₄K23AcR26Me2a NEC, processed either with SNAP or centroid peak picking. The monoisotopic mass for the peptide is highlighted in green in the corresponding tables, whilst a peak mistakenly identified as the peptide monoisotopic mass with centroid processing is highlighted in red. **b)** Example showing peaks in sections of rapifleX MALDI-TOF-MS spectra for two replicates of SETI H3₁₋₂₁K4Me3 with PCAF HAT, illustrating how spot quality can affect MALDI-TOF-MS spectrum quality, peak picking and calculation of enzyme activity.

Issues with quantification can arise due to peak picking in spots with low-quality spectra.

Spotting with Mosquito can lead to variable spot quality giving spectra of differing quality for the same sample, due to mixing or spotting errors and variability in spot crystallisation. Peak picking for low quality spectra with low intensity peaks will only identify the most abundant peaks, with other peaks lost in the spectrum noise, leading to mis-quantification of enzyme activity (e.g. SETI H3₁₋₂₁K4Me3, Figure 3.4b). Mis-quantification affects the calculated standard deviation between replicates for each

library member. Setting a higher signal to noise threshold for peak picking can mitigate the issue, but due to variable quality of peptides and their corresponding MS spectra, the threshold must be kept relatively low to avoid false negative results. A signal to noise threshold of 2 was used for analysis.

3.2.5.2 Data recovery script

The peak picking script generates a peak list output file with the masses and intensities listed for each peak in each spectrum (Figure 3.5a). The parent and product peaks for each library member being screened must then be identified, the corresponding data recorded, and their relative intensities used to measure enzyme activity. A python script was written by Dr. Tom McAllister designed in collaboration with the author to automatically carry out data extraction and calculation of average relative percentages and SD for each peptide.

A mass list was first generated in Excel which lists the expected masses for each of the MALDI spectra in a plate (Figure 3.5b). The masses of interest were for PCAF parent peptide mass, single acetylation mass (1Ac, parent mass + 42 Da) and double acetylation mass (2Ac, parent mass + 84 Da). For SIRT1, up to 5 deacetylation events were tracked for each of the library members as the SETIV library has peptides with 0-5 acetyl groups (parent mass and parent mass -42 Da / -84 Da / -126 Da / -168 Da / -210 Da, -1Ac – -5Ac).

The script uses the location (X and Y coordinates) listed in the mass list to locate the correct MALDI-TOF-MS spectrum and its corresponding peak list in the data files. To account for variability in the exact mass recorded by MALDI-TOF-MS a 1 Da range was defined. The script identifies the closest peak to the mass looked for within a 1 Da range,

recording the exact mass and its corresponding intensity. SNAP peak picking only lists monoisotopic peaks, limiting errors in peak identification. With centroid processing for H3₁₄₋₃₄K23acR26me2a the closest peak to the listed parent mass of 2516 Da was 2516.134 Da (highlighted in red), corresponding to an isotopic peak of an impurity, rather than the monoisotopic peak of the parent peptide with a mass of 2516.426 Da (highlighted in green, Figure 3.4b). The script uses mass and intensity information to calculate the relative percentage of parent and modified peptides for the spectrum (Figure 3.4b, Figure 3.5c).

Each library member/enzyme combination had multiple replicates (N). For SETIII run with low-throughput, there were three assay replicates for PCAF HAT and two replicates for both PCAF HAT-BRD and NEC. For the other libraries run in a high-throughput manner, there were three assay replicates of each enzyme and NEC, and each assay aliquot was spotted with a further four technical replicates, bringing the total replicate number to twelve (or 8 for NEC due to limited number of 1536 plates available). Some library members had fewer than the maximum number of replicates, corresponding to flatline spectra or spectra where no peaks of interest were identified. This can be due to poor purity of the synthesised peptide, poor desorption/ionisation of the peptide under MALDI conditions or due to technical variability (pipetting errors in assay, spotting errors in liquid handling, variable crystallisation of spot).

The replicates of each library member + enzyme combinations were grouped under a single ID number in the Excel mass list, which allowed the script to identify replicates. Replicates in which none of the peptide masses could be identified were treated as failed spectra which were not counted. This removed flat spectra as well as spectra with only

impurity peaks. An average of the percentage of each peptide state (parent peptide and modified peptides) with corresponding standard deviation (SD) and number of replicates (N) was calculated for spectra with the same ID (Figure 3.5d). SD and N were important indicators of assay reproducibility in terms of enzyme activity (SD) and of peptide quality (SD and N). A large SD can be due to variability in enzyme activity between biological replicates or to low peptide quality (overlapping background peaks or peptide peaks close to spectrum noise). A low N points to issues with peptide purity or technical issues but does not inform on enzyme activity, as the parent peptide should still be identified if the enzyme is not active. Peptides with a $SD \geq 10\%$ were manually checked to identify possible sources of errors. Because different numbers of replicates were used in low-throughput and high-throughput screens, no number of successful replicates was set to consider data as reliable.

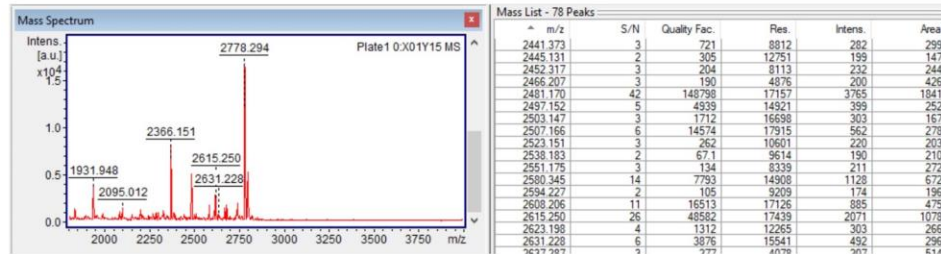
The script was run on results of each MALDI plate separately as it used plate co-ordinates with ID information based on each plate map, and could not compile data acquired from several different MALDI plates. NEC control data was compiled separately with average percentage, SD and N calculated using Excel because the NEC were spotted onto several different plates. The average percentage, SD and N for SETIII (run in low-throughput assay) were also calculated using Excel as a 96 spot MALDI plate format was used and the script is designed to identify replicates only if the data acquisition is carried out in a single 1536 spot MALDI plate format.

An example workflow for SETII with PCAF HAT, HAT-BRD and NEC shows the data analysis steps: output of the first script run in Flexanalysis (Figure 3.5a); format of an Excel mass list (Figure 3.5b); output of the second script showing recorded mass and

intensities with corresponding calculated relative percentages for each peptide state (Figure 3.5c); output of the second script with an average of the calculated relative percentages (%) with corresponding SD and N for each peptide ID (Figure 3.5d); data collation to match a peptide to each ID and combine results of each enzyme for each peptide in a single spreadsheet with results format $\% \pm \text{SD (N)}$ (Figure 3.5e). The workflow also shows how SD and N can be used for quality check. For peptide B1 with HAT the parent peptide result was 50 ± 19 (12), where the number of replicates was the maximum of 12 but the SD was high at 19%. For peptide B1 with HAT-BRD, the parent peptide gave 60 ± 8 (12), with a low SD number of 8%. This pointed towards an issue with the reproducibility of enzyme activity rather than an issue with peptide quality as the same peptide gave a more reproducible result with HAT-BRD than with HAT. Peptide B2 showed low SD and high N for all columns giving good confidence in these results. For peptide B3 most replicates gave no results (low reported N) with one HAT-BRD spectrum giving 100% 2Ac and one NEC spectrum giving 100% parent peptide. This indicated a problem with peptide purity or fitness of this peptide for MALDI-TOF-MS analysis (Figure 3.5e).

The data analysis script was designed to identify expected products listed in the mass list. Any side reactions of the enzyme which do not involve the expected mass shift cannot be identified with this data processing method.

a) **First script** (peak picking in Flexanalysis software and peak list generation)



b) **Excel mass list** (masses and peptide identification information)

peptide identification			enzyme ID		replicate ID	mass list		
1536 MALDI	384 well	library	enzyme	replicate	ID	parent	1Ac	2Ac
X25Y05	C13	B1	HAT	2	HATB1	2444	2486	2528
X26Y05	C13	B1	HAT	2	HATB1	2444	2486	2528
X27Y05	C14	B2	HAT	2	HATB2	2458	2500	2542
X28Y05	C14	B2	HAT	2	HATB2	2458	2500	2542
X29Y05	C15	B3	HAT	2	HATB3	2473	2515	2557

c)

Second script (returns exact mass and intensity for each identified peak and calculates relative percentage)

script output for each spectrum														
1536 MALDI	ID	parent	mass found	int	%age	1Ac	mass found	int	%age	2Ac	mass found	int	%age	any found
X25Y05	HATB1	2444	2444.24	2419.14	55.52	2486	2485.25	1938.41	44.48	2528	0	0	0	1
X26Y05	HATB1	2444	0	0	0	2486	2485.24	347.95	100.00	2528	0	0	0	1
X27Y05	HATB2	2458	2457.24	695.16	64.43	2500	2499.25	383.77	35.57	2542	0	0	0	1
X28Y05	HATB2	2458	2457.24	5942.86	66.67	2500	2499.24	2816.18	31.59	2542	2541.22	154.98	1.74	1
X29Y05	HATB3	2473	0	0	0	2515	0	0	0	2557	0	0	0	0

d)

Second script (Uses ID number to calculate average percentage %, standard deviation SD and number of replicates N)

1536 MALDI	ID	parent	%	SD	N	1Ac	%	SD	N	2Ac	%	SD	N
X25Y05	HATB1	2444	50	19	12	2486	50	19	12	2528	0	0	12
X26Y05	HATB1	2444	50	19	12	2486	50	19	12	2528	0	0	12
X27Y05	HATB2	2458	70	4	12	2500	30	4	12	2542	1	1	12
X28Y05	HATB2	2458	70	4	12	2500	30	4	12	2542	1	1	12
X29Y05	HATB3	2473	0	0	0	2515	0	0	0	2557	0	0	0

e)

Data collation and peptide identification (match peptide to ID number and compare results between enzymes)

ID	HAT			HAT-BRD			NEC			Peptide ID		
	parent	1Ac	2Ac	parent	1Ac	2Ac	parent	1Ac	2Ac			
B1	50 ± 19 (12)	50 ± 19 (12)	0 ± 0 (12)	60 ± 8 (12)	40 ± 8 (12)	0 ± 0 (12)	97 ± 5 (9)	3 ± 5 (9)	0 ± 0 (9)	H4	1 21	K12me
B2	70 ± 4 (12)	30 ± 4 (12)	1 ± 1 (12)	72 ± 3 (12)	27 ± 3 (12)	1 ± 1 (12)	99 ± 1 (9)	1 ± 1 (9)	0 ± 0 (9)	H4	1 21	K12me2
B3	0 ± 0 (0)	0 ± 0 (0)	0 ± 0 (0)	0 ± 0 (1)	0 ± 0 (1)	100 ± 0 (1)	100 ± 0 (1)	0 ± 0 (1)	0 ± 0 (1)	H4	1 21	K12me3

Figure 3.5: Example workflow for the data extraction and collation process involved in the analysis of results of the rapifleX MALDI-TOF-MS HTS.

a) Output of the first peak picking script run in Flexanalysis **b)** Format of an Excel mass list **c)** Output of the second script showing recorded masses and intensities with corresponding calculated relative percentage for each peptide state **d)** Output of the second script with calculated average percentage of each peptide state (%), standard deviation (SD) and number of replicates (N) for each peptide ID **e)** Data collation to match a peptide to each ID and list results of each enzyme for each peptide in a single spreadsheet with a results format of % ± SD (N).

3.2.6 Quality issues observed with the peptides in the purchased histone peptide libraries

Due to the nature of parallel peptide synthesis of the purchased histone peptide library, the quality of the synthesised library peptides was variable. Some of the issues encountered in the screen are illustrated in example spectra **A-F** (Figure 3.6). When observed, the parent and single acetylation (1Ac) product peptides were indicated on the spectra.

Acetylation was clearly observed when a relatively high purity peptide was incubated with PCAF HAT (*e.g.* SETI H3₁₋₂₁K4Me3K9Ac, Figure 3.6A). The spectrum was very clean with parent and 1Ac product peptides clearly identifiable. This was considered to be reliable data, with the datasets for this peptide resulting in $\leq 2\%$ SD and 8N. Where the peptides were of low purity, poor dataset were observed. For SETI H3₁₋₂₁K4Me3K9Me3 (Figure 3.6B) the parent substrate peak was identified as a minor peak amongst several impurity peaks in both the NEC and enzyme incubated spectra. Incubation with PCAF HAT resulted in formation of 1Ac product, also a minor peak in the spectrum. The parent and product peaks both had intensities < 1000 a.u. compared with intensities of 1000-5000 a.u. for impurities in the spectrum. Such spectra are likely due to a combination of poor synthesis of the peptide and poor flying or ionisation with MALDI of the peptide compared with impurities. The calculated enzyme activity for such spectra was less reliable than for clean spectra. In other cases, impurities in the peptides were also substrates of the enzymes and resulted in multiple product peaks. For example, incubation of SETI H3₁₄₋₃₄K27AcS28Phos with PCAF HAT (Figure 3.6C) resulted in several sets of peaks; the parent and 1Ac product peptide peaks were clearly

identified in the first set, and a second set of peaks were observed with impurity and impurity +42 Da peaks, showing that the impurity is also a substrate for PCAF HAT. This can affect levels of enzyme activity for the library peptide. Some spectra showed impurities which did not affect peak identification or calculation of enzyme activity. For example, incubation of SETI H3₁₋₂₁K14Me3 with PCAF HAT (Figure 3.6D) gave clearly identifiable parent and 1Ac product peaks. However, impurities lower the peptide concentration which can affect enzyme activity and ability to perform MS/MS on the peptides. In other cases, impurities overlapped with the peptide peaks, affecting peak picking and calculation of enzyme activity (this issue was minimised by using SNAP processing, Section 3.2.5.1, Figure 3.4a). This is seen in the acetylation of SETI H3₁₄₋₃₄K23AcR26Me2a by PCAF HAT (Figure 3.6E). All previous examples are from the SETI library and are representative of spectra in the SETI, SETII and SETIII libraries. The SETIV library is of poorer quality than the other libraries, showing fewer clean peptides. In addition to issues previously stated, analysis of this library was further complicated by contamination in the spectra by a polymer showing a 44 Da mass difference between peaks (the nature and origin of the contamination is further discussed in section 3.4). An example of this is seen in the spectrum for SETIV H3₁₁₋₃₁K14Ac peptide acetylation by PCAF HAT (Figure 3.6F). In this case, parent and 1Ac peptide peaks did not overlap with any polymer peaks, and enzyme activity could be correctly measured. However, lower quality of data was obtained in the screen of PCAF and SIRT1 with SETIV compared with screening of other libraries, due to peak overlap between polymer peaks and peptide peaks.

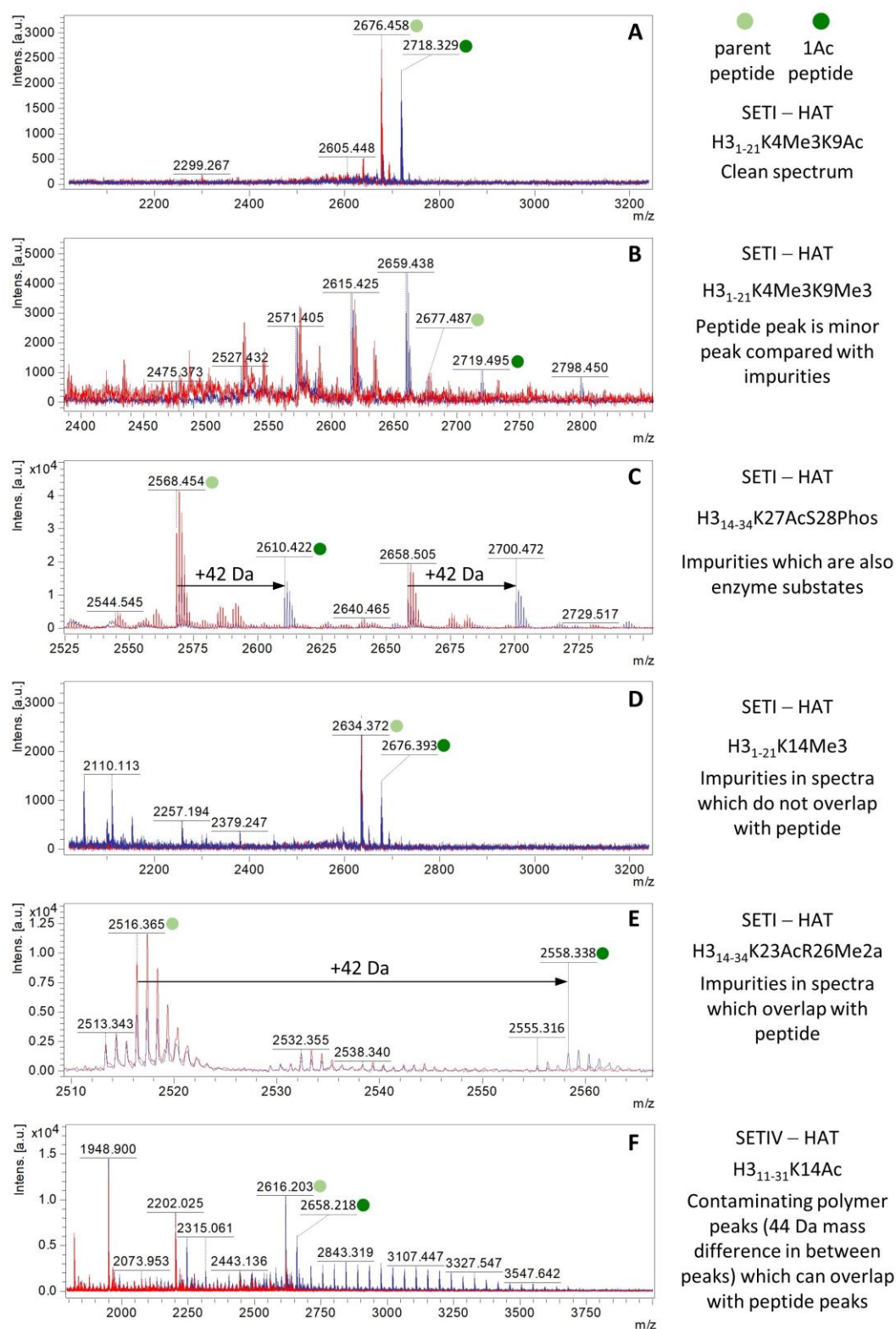


Figure 3.6: Example spectra illustrating problems peptide quality in the histone peptide libraries.

Spectra **A-F** were selected from the rapifleX MALDI MS screen of PTM histone peptide libraries against PCAF HAT (blue) overlaid with a NEC spectrum (red). The peptide masses of interest are indicated on the spectra with colour-coded dots. The peptide identity and the quality issue being illustrated is listed on the right of each spectrum.

The issues exemplified in these spectra were minimised as much as possible with adapted processing methods and quality checks with the SD and number of replicates (peaks in the NEC which overlap with the product peaks were also taken into account). However, it was often impossible to identify the source of poor data quality without checking the original spectra. In addition, impurities which did not directly affect identification and quantification of peptide peaks were not taken into account but may have affected enzyme activity (*e.g.* if impurities were better substrates) and quality of results.

3.3 RapifleX MALDI-TOF-MS screen results for PCAF HAT and HAT-BRD

3.3.1 Overview of PCAF HAT and HAT-BRD screen results

Histone peptide libraries SETI, SETII, SETIII and SETIV were screened against PCAF HAT, PCAF HAT-BRD and a NEC used for quality control. The assays were run, quantified, and analysed as set out in section 3.2. Peaks of +42 Da unit shift(s), relative to the parent, were considered as acetylated product peaks. Results are shown as relative proportion (%) of parent, single acetylation (1Ac) and double acetylation (2Ac) product peptides with associated standard deviation (SD) and number of replicates (N) under the format % $\pm$ SD (N). To gain a global picture of the screen results, the control peptides used in the libraries were examined (Table 3.1, Table 3.2), and summary figures were produced showing activity trends for each library (Figure 3.7). Full screen results can be found in Appendix Table D.1, Appendix Table D.2, Appendix Table D.3 and Appendix Table D.4.

3.3.1.1 Use of control peptides in data analysis of HTS results

When developing a HTS, increased throughput may be achieved to the detriment of data robustness. In the MALDI-TOF-MS HTS, mitigation strategies were incorporated to give confidence in the validity of the data. Control peptides were used to compare the relative activity of the enzymes between libraries or between substrates. Untagged purified H3₁₋₂₁ peptide, previously used to optimise assay conditions (section 3.2.3), was used as an 'external' control in each library, to ensure that the enzyme and reagents (*e.g.* AcCoA coenzyme, buffers) used were functioning correctly for a known substrate, and to be able to compare peptides conversion between libraries. In addition, the libraries used contained reference peptides with no PTMs to be used as 'internal' controls for comparisons with peptides of matching amino acid sequence but varying PTMs (*e.g.* biotinylated H3₁₋₂₁ can be used as an internal control for biotinylated H3₁₋₂₁K4Ac, both of which were included in the SETI library), allowing for comparison of peptide conversion between different peptides within a library.

Incubation of the external control with PCAF enzyme proceeded as expected based on optimisation assays, with 1Ac and 2Ac products of H3₁₋₂₁ observed for HAT and HAT-BRD in all four libraries screened against PCAF HAT/HAT-BRD. The peptide used was quantified by NMR and is of high purity (>95%) and should therefore be more reliable as a control between libraries than library peptides which are not purified or quantified. However, despite the consistent assay conditions used across SETI, SETII and SETIV library screens, variability in the level of 2Ac product was observed ranging from 32% to 69% conversion (note that the percentage of 1Ac product observed was dependent on

the amount of 2Ac product as the 1Ac product is substrate for a second acetylation). Screening of SETIII was carried out with different conditions in a low throughput assay, but the external control showed 2Ac conversion in the same range as the other libraries (Table 3.1).

Internal controls were part of each library, synthesised with the same biotin tags as the rest of the peptides in the library, and were therefore better references for comparisons to library peptides with PTMs than the external controls. Internal controls were adapted to the library being screened (*e.g.* SETII did not have a tagged H3₁₋₂₁ internal control because the library covered H2a, H2b and H4 but not H3), but there was overlap between the libraries and some internal controls could therefore be compared between libraries (Table 3.1). SETI, SETIII and SETIV had tagged H3₁₋₂₁ control peptides. Whilst SETI and SETIII showed similar levels of 2Ac conversion for tagged H3₁₋₂₁ (53% and 56% respectively with PCAF HAT), the same control for SETIV was not functional (flat line spectra were acquired for the peptide, indicating an issue with peptide synthesis). In the same way, the tagged H4₁₋₂₁ control peptide found in the SETII and SETIII libraries showed consistent levels of 1Ac peptide conversion (28% and 30% respectively with PCAF HAT), whilst the equivalent SETIV control was not functional. The SETII library contained tagged H2a₁₋₂₁ and H2b₁₋₂₁ internal control peptides which were used for comparisons within the SETII library, but the equivalent SETIV control peptides were not functional for comparisons between libraries. Most of the internal controls in the SETIV library were not functional with either flat-line spectra acquired (*e.g.* H3₁₋₂₁, H4₁₋₂₁ and H2b₁₋₂₁, Table 3.1) or with overlapping background peaks in the NEC (*e.g.* H2a₁₋₂₁, Table 3.1) which contributed to the lower quality of data obtained for the SETIV library.

As internal controls for SETIV could not be used due to poor library quality, a like for like comparison of peptides was used to check the relative activity between libraries. H3₁₋₂₁K14Ac (SET I and SET IV), H4₁₋₂₁K16Ac (SET II and SET IV), H2a₁₋₂₁K13Ac (SET II and SET IV) and H2b₁₋₂₁K5Ac (SET II and SET IV) were used for comparisons (Table 3.2). Whilst the HAT-BRD results were comparable between all the peptides in different libraries, the conversion levels were significantly higher for peptides of the SETIV library incubated with HAT. The reasons for this are unclear: SETIV and SETIII shared a C-terminal Ahx-Lys-biotin tag and each of the libraries had overlap of peptide sequences with SETIV. This unexplained increased activity for PCAF HAT against SETIV was taken into account when analysing the data obtained.

Due to variability in peptide conversion observed for the external controls of the four peptide libraries, comparison of absolute activities (quantitative comparisons) between libraries may not be reliable. However, a qualitative analysis was made, with internal controls used for 'rank order' of peptides and analysis of trends, and comparisons made between trends in different libraries.

H3			1 – ARTKQTARK STGGKAPRKQ LA – 21					
H4			1 – SGRGKGGKG LGKGGAKRHR KV – 21					
H2a			1 – SGRGKQGGK ARAKAKTRSS RA– 21					
H2b			1 – PEPAKSAPA PKKGSKKAVT KA – 21					
peptide length: 1-21								
# – no tag control / N – biotin-Ahx N-terminus tag / C – Ahx-Lys-biotin C-terminus tag								
Lib	Tag	H	HAT		HAT-BRD		NEC	
			1Ac	2Ac	1Ac	2Ac	1Ac	2Ac
I	#	H3	30 ± 6 (9)	69 ± 6 (9)	54 ± 4 (7)	46 ± 4 (7)	0 ± 0 (3)	0 ± 0 (3)
II	#	H3	66 ± 3 (12)	32 ± 3 (12)	83 ± 7 (12)	16 ± 7 (12)	1 ± 0 (9)	0 ± 0 (9)
III	#	H3	50 ± 15 (3)	50 ± 15 (3)	76 ± 3 (2)	24 ± 3 (2)	0 ± 0 (2)	0 ± 0 (2)
IV	#	H3	51 ± 1 (12)	48 ± 1 (12)	66 ± 3 (12)	34 ± 3 (12)	1 ± 0 (8)	1 ± 1 (8)
I	N	H3	46 ± 6 (7)	53 ± 6 (7)	70 ± 10 (12)	29 ± 10 (12)	0 ± 0 (7)	0 ± 0 (7)
III	C	H3	43 ± 15 (3)	56 ± 16 (3)	77 ± 6 (2)	22 ± 5 (2)	0 ± 0 (2)	0 ± 0 (2)
IV	C	H3	0 ± 0 (7)	0 ± 0 (7)	0 ± 0 (11)	0 ± 0 (11)	0 ± 0 (7)	0 ± 0 (7)
II	N	H4	28 ± 2 (12)	1 ± 0 (12)	27 ± 6 (12)	0 ± 0 (12)	3 ± 2 (9)	0 ± 0 (9)
III	C	H4	30 ± 6 (3)	0 ± 0 (3)	31 ± 1 (2)	0 ± 0 (2)	1 ± 0 (2)	0 ± 0 (2)
IV	C	H4	0 ± 0 (1)	0 ± 0 (1)	0 ± 0 (0)	0 ± 0 (0)	0 ± 0 (0)	0 ± 0 (0)
II	N	H2a	16 ± 3 (12)	0 ± 0 (12)	13 ± 2 (12)	0 ± 1 (12)	1 ± 0 (9)	0 ± 0 (9)
IV	C	H2a	36 ± 2 (9)	32 ± 3 (9)	84 ± 23 (12)	0 ± 0 (12)	72 ± 28 (8)	10 ± 19 (8)
II	N	H2b	8 ± 1 (12)	0 ± 0 (12)	5 ± 2 (12)	0 ± 0 (12)	3 ± 1 (9)	0 ± 0 (9)
IV	C	H2b	0 ± 0 (0)	0 ± 0 (0)	0 ± 0 (0)	0 ± 0 (0)	0 ± 0 (0)	0 ± 0 (0)

Table 3.1: Comparison of external and internal control peptides results for comparisons between libraries.

RapifleX MALDI-TOF-MS screen results for PCAF HAT, PCAF HAT-BRD and a no enzyme control (NEC) against SETI-IV histone peptide libraries with post translational modifications (PTMs). The results are shown as an average of the observed relative percentages (%) of peak intensities for parent peptide, singly acetylated peptide (1Ac) and doubly acetylated peptide (2Ac) with associated standard deviation (SD) and number of replicates (N) under the format % ± SD (N). The results for the parent peptide are not shown for clarity. A gradual colour scale is used to represent observed percentage of peptide ranging from lowest value % (red) to highest value % (green). The peptides shown are external untagged H3₁₋₂₁ peptide, and internal control tagged peptides H3₁₋₂₁, H4₁₋₂₁, H2a₁₋₂₁ and H2b₁₋₂₁ peptides for each of the libraries. Issues with the spectra of the internal control peptides for SETIV meant these could not be used as controls.

H3							1 – ARTKQTARK STGGKAPRKQ LA – 21			
H4							1 – SGRGKGGKG LGKGGAKRHR KV – 21			
H2a							1 – SGRGKQGGK ARAKAKTRSS RA– 21			
H2b							1 – PEPAKSAPA PKKGSKKAVT KA – 21			
peptide length: 1-21										
A – acetyl / N – biotin-Ahx N-terminus tag / C – Ahx-Lys-biotin C-terminus tag										
Lib	K 14	K 16	K 13	K 5	Tag	H	HAT		HAT-BRD	
							1Ac	2Ac	1Ac	2Ac
I	A				N	H3	42 ± 4 (12)	1 ± 0 (12)	33 ± 3 (12)	0 ± 0 (12)
IV	A				C	H3	82 ± 15 (10)	14 ± 16 (10)	25 ± 1 (12)	0 ± 0 (12)
II		A			N	H4	31 ± 5 (12)	0 ± 1 (12)	30 ± 6 (12)	0 ± 0 (12)
IV		A			C	H4	96 ± 3 (12)	4 ± 2 (12)	27 ± 1 (12)	0 ± 1 (12)
II			A		N	H2a	10 ± 2 (12)	0 ± 0 (12)	9 ± 1 (12)	0 ± 0 (12)
IV			A		C	H2a	43 ± 6 (12)	0 ± 0 (12)	10 ± 1 (12)	0 ± 0 (12)
II				A	N	H2b	6 ± 1 (12)	0 ± 0 (12)	5 ± 1 (12)	0 ± 0 (12)
IV				A	C	H2b	12 ± 8 (12)	0 ± 0 (12)	4 ± 0 (12)	0 ± 0 (12)

Table 3.2: SETIV library peptides matched with comparable peptides from other libraries.

RapifleX MALDI-TOF-MS screen results for PCAF HAT and HAT-BRD against SETI-IV histone peptide libraries with post translational modifications (PTMs). The results are shown as an average of the observed relative percentages (%) of peak intensities for parent peptide, singly acetylated peptide (1Ac) and doubly acetylated peptide (2Ac) with associated standard deviation (SD) and number of replicates (N) under the format % ± SD (N). The results for the parent peptide are not shown for clarity. A gradual colour scale is used to represent observed percentage of peptide ranging from lowest value % (red) to highest value % (green). The peptides shown are peptides from the SETIV peptide libraries matched with equivalent peptides from the SETI and SETII libraries for a comparison of conversion levels. The conversion levels for HAT are consistently higher in the SETIV library, whilst the conversion levels for HAT-BRD are equivalent.

3.3.1.3 Global trends observed in PCAF activity for the SETI-IV histone peptide libraries

For each library, the results were compiled into a heat map diagram showing the relative proportion of parent peptide to enzymatically acetylated peptides for each of the enzyme assays with a colour scale of 0-100% going from white to green. The percentage of parent, single acetylation (1Ac) and double acetylation (2Ac) product peptide was calculated as an average of all replicates for the peptide-enzyme combination concerned (SD and N values are not shown). Enzymatic acetylation activity by the two PCAF constructs against the range of peptides was quantified in terms of conversion to 1Ac and 2Ac products (Figure 3.7a-d). Full results corresponding to the summary figures can be found in Appendix D.

The NEC served as a control for the peptide state prior to enzyme incubation. Thus a peptide was considered potential substrate of the enzyme, only if the +42 (1Ac, 2Ac) product ratios increased relative to the parent, compared to NEC. This method should correct for any +42 mass change due to overlapping impurities, non-enzymatic acetylation (not previously observed with experimental conditions used), or formation of peptide-ACN adduct with ACN from the peptide solvent or the MALDI matrix solvent. In addition, the NEC for each of the libraries gave some idea of overall library quality. However, as only peaks which overlapped with the expected masses for the peptides were detected with the script processing method used, this representation does not give a full overview of peptide quality. A method to quantify the proportion of side products and impurities in comparison to product peptide would give a better overview of library

quality. However, the method used allowed for identification of impurity peaks which overlapped with product peptides introducing errors into the calculations of enzyme activity. By comparing peptide conversions, SD and N for enzyme assay results and the NEC, some peptides were discounted from analysis of results on the basis of impurity overlap (*e.g.* if there was a peak corresponding to the 1Ac or 2Ac peptide product in the NEC with a relative intensity of $\geq 5\%$).

SETI comprised of 20 amino acid sections of H3 peptides ranging from 1-135 (Figure 3.1c). The NEC showed an increase in overlapping impurities for peptides corresponding to the C-terminus of H3. Spectra with no detectable peptide in the NEC are indicated by grey boxes in Figure 3.7. H3₁₋₂₁ showed the highest levels of 1Ac and 2Ac activity followed by H3₁₄₋₃₄. Sections of H3 from 28-135 showed low conversion rates compared with H3₁₋₂₁ and H3₁₄₋₃₄ (Figure 3.7a). The known acetylation sites on H3 are K14 and K9, with suspected sites of acetylation at K18 and K23 and K27 based on MS studies of acetylation activity for the homologous HAT enzyme GCN5.³⁴⁰ The results obtained showed that PCAF acetylation is low for H3 positions after residue 34 which is in agreement with known and suspected sites of acetylation for PCAF on H3.

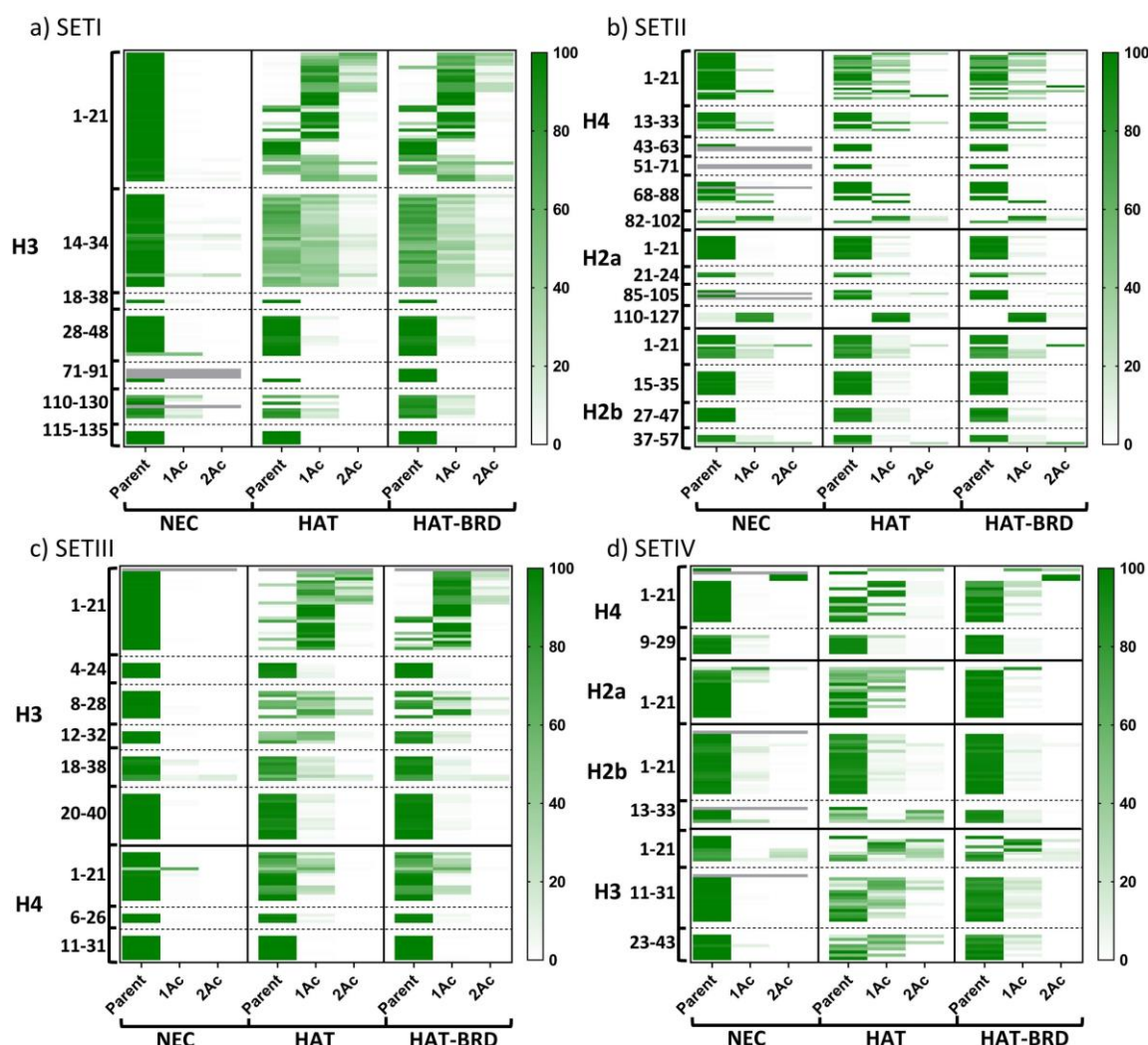


Figure 3.7: Summary figures for the SETI-IV libraries screened against PCAF HAT and HAT-BRD in rapifleX MALDI-TOF-MS HTS.

For each library, results are shown for screening of the peptides against PCAF HAT and HAT-BRD with the no enzyme control (NEC) on the left. Enzyme activity for each section (NEC, HAT or HAT-BRD) was measured in terms of observed relative percentages of parent, single acetylation (1Ac) and double acetylation (2Ac) peptides. The relative percentage is shown as a colour scale from white (0% observed peptide) to green (100% observed peptide). Full results tables, showing corresponding peptide PTMs, standard deviation and number of replicates for each peptide are available in Appendix D. **a)** Results for the SETI histone peptide library (Appendix Table D.1). **b)** Results for the SETII histone peptide library (Appendix Table D.2). **c)** Results for the SETIII histone peptide library (Appendix Table D.3). **d)** Results for the SETIV histone peptide library (Appendix Table D.4).

SETII library comprised of H4, H2a and H2b covering residues 1-102, 1-57 and 1-102 respectively. Taking into account NEC overlapping impurity peaks, this library showed 1Ac activity of PCAF for H4₁₋₂₁, with much lower activity for peptides corresponding to the C-terminus of H4. Acetylation activity of PCAF for H2a₁₋₂₁ and H2b₁₋₂₁ was observed,

with lower conversion than for H4₁₋₂₁. However, high proportions of overlapping impurity peaks in the later section of H2a and H2b peptides made analysis for these more difficult (Figure 3.7b).

SETIII comprised of H3 and H4 peptides covering residues 1-40 and 1-31 respectively. This library was clean compared to the other libraries with very few overlapping background peaks and no failed control spectra. This can be partly attributed to the peptides sequences that were covered in this library, as the *N*-terminal sections of H3 and H4 synthesise well on peptide synthesisers, are soluble in assay conditions and are generally amenable to MALDI-TOF-MS. Another contributing factor is that this library was screened with a low-throughput process in which spectra were acquired manually if automatic MALDI-TOF-MS methods failed to produce a spectrum. This library showed high PCAF 1Ac and 2Ac activity for H3₁₋₂₁. Lower acetylation activity was observed for H3₄₋₂₄ and H3₈₋₂₈, with very modest activity observed for H3₂₀₋₄₀. H4₁₋₂₁ showed 1Ac activity, with other sections of H4 showing under 5% 1Ac (Figure 3.7c). These results were consistent with identified preferred acetylation sites of PCAF H3K14, H3K9 and H4K8.

SETIV comprised of H4, H2a, H2b and H3 peptides covering residues 1-29, 1-21, 1-33 and 1-43 respectively. This was an acetylation only library with zero to five combinatorial acetylation marks on the peptides. This library showed the poorest library quality with peptide peak intensities much lower than impurity intensities in many cases, for the NEC. For PCAF acetylation activity, the overlap with background peaks was moderate. This library showed similar results to previous libraries: H3₁₋₂₁ showed the highest levels

of 1Ac and 2Ac enzyme activity; H4₁₋₂₁, H3₁₁₋₃₁ and H3₂₃₋₄₃ showed moderate 1Ac conversion; H2a and H2b peptides showed the lowest levels of conversion. In libraries SETI-III, PCAF HAT and HAT-BRD constructs showed very similar levels of activity (exceptions discussed below). For SETIV, levels of conversion were significantly and consistently higher for HAT compared with the HAT-BRD, affecting all the different histones equally (Figure 3.7d).

The overview of these four histone peptide libraries showed that the PCAF enzymes were most active on the *N*-terminal portions of the histones. The highest levels of activity were observed for H3₁₋₂₁ with 1Ac and 2Ac products observed. PCAF showed moderate activity for portions of H3 up to residue 40 and against H4₁₋₂₁ with only 1Ac product observed. PCAF activity against H2a₁₋₂₁ and H2b₁₋₂₁ was measurable but with low levels of conversion. H2a, H2b and H4 peptides covering sections after the 21st residue showed little to no acetylation activity (Figure 3.7a-d).

3.3.2 Site selectivity of PCAF HAT and HAT-BRD for H3, H4, H2a and H2b

To gain a better understanding of the screening results, it is important to determine the site of lysine acetylation by PCAF for each peptide. To date, PCAF has been reported to only catalyse the acetylation of H3K14, H3K9 and H4K8 on histones.¹¹⁹ Within the same HAT subfamily, acetylation of H2a and H2b by the GCN5 containing SAGA complex has been observed, with H2bK11 and H2bK16 identified as acetylation sites.^{114–117} No acetylation site has been recorded for H2a with GCN5, and no studies have reported PCAF HAT acetylation of H2a and H2b.

The next step in analysis of screen results therefore focussed on confirming the selectivity of PCAF for H3 and H4 peptides and identifying new acetylation site preferences on H2a and H2b. In the absence of high-throughput MS/MS to screen all peptides from the library, systematic acetylation profiling of acetylated histone peptides was used to deduce the likely sites of enzymatic modification. Screen results are given as relative percentages of parent (not shown) to product (1Ac and 2Ac) peptides. Results are shown in tables with shading ranging from red (lowest % value) to green (highest % value) with results in the format % $\pm$ SD (N). Tables here show selected results from the screen to illustrate analysis. Full results tables can be found in Appendix D.

3.3.2.1 Site selectivity of PCAF HAT and HAT-BRD for H3 peptides

The internal H3₁₋₂₁ control peptides gave 1Ac and 2Ac products when incubated with PCAF. SETI and SETIII H3₁₋₂₁ control peptide gave $\approx$ 45% 1Ac and $\approx$ 55% 2Ac product with PCAF HAT, with all parent peptide depleted. As the second acetylation reaction uses the 1Ac product as a substrate, these results meant that full turnover to 1Ac and $\approx$ 55% turnover to 2Ac was occurring. For PCAF HAT-BRD with SETI and SETIII, H3₁₋₂₁ control peptide showed $\approx$ 75% 1Ac and $\approx$ 25% 2Ac, giving full 1Ac turnover and $\approx$ 25% 2Ac turnover (Table 3.3). This was in line with previous time-dependent MALDI-TOF-MS and MALDI-TOF/TOF-MS assay results for H3₁₋₂₁ peptide incubated with PCAF HAT or HAT-BRD, where the first acetylation corresponded to H3₁₋₂₁K14Ac and the second corresponded to H3₁₋₂₁K9Ac (Chapter 2).

- Single acetylation

H3₁₋₂₁K9Ac peptides gave $\approx 100\%$ 1Ac product and no 2Ac product (HAT SETI/III $<5\%$ 2Ac, HAT SETIV $7 \pm 9\%$ 2Ac acetylation within error range; HAT-BRD SETI/III/IV $<5\%$ 2Ac), suggesting that the K14 site was fully enzymatically acetylated under the assay conditions. SETI H3₁₋₂₁K14Ac gave 42% 1Ac with HAT and 33% 1Ac with HAT-BRD, indicating that $<50\%$ of the K9 position was enzymatically acetylated under the assay conditions. H3₁₋₂₁K4Ac was only present in SETIV, giving conversions of $26 \pm 15\%$ 1Ac and $73 \pm 16\%$ 2Ac with HAT with $79 \pm 3\%$ 1Ac and $18 \pm 1\%$ 2Ac with HAT-BRD. The 2Ac conversion for HAT against H3₁₋₂₁K4Ac was higher than control levels ($\approx 55\%$ 2Ac), but the high SD of 16% may account for this discrepancy. In addition, it was previously shown that SETIV conversion for HAT was higher than for the other libraries. For HAT-BRD the SD are low and the levels of conversion for H3₁₋₂₁K4Ac 1Ac and 2Ac were in range with the H3₁₋₂₁ internal controls of SETI and SETIII. It is therefore unlikely that H3₁₋₂₁K4Ac affected PCAF activity. Given that 1Ac and 2Ac products were observed with H3₁₋₂₁K4Ac, K4 could also be ruled as a site of PCAF acetylation (Table 3.3).

H3				1 – ARTKQTARK STGGKAPRKQ LA – 21					
peptide length: 1-21									
A – acetyl									
SE T	K 4	K 9	K 14	HAT		HAT-BRD		NEC	
				1Ac	2Ac	1Ac	2Ac	1Ac	2Ac
I				46 ± 6 (7)	53 ± 6 (7)	70 ± 10 (12)	29 ± 10 (12)	0 ± 0 (7)	0 ± 0 (7)
III				43 ± 15 (3)	56 ± 16 (3)	77 ± 6 (2)	22 ± 5 (2)	0 ± 0 (2)	0 ± 0 (2)
I		A		96 ± 1 (12)	3 ± 0 (12)	95 ± 1 (12)	2 ± 1 (12)	0 ± 1 (8)	0 ± 0 (8)
III		A		98 ± 2 (3)	1 ± 1 (3)	99 ± 0 (2)	0 ± 0 (2)	0 ± 0 (2)	0 ± 0 (2)
IV		A		92 ± 9 (12)	7 ± 9 (12)	97 ± 3 (12)	0 ± 0 (12)	0 ± 0 (8)	0 ± 0 (8)
I			A	42 ± 4 (12)	1 ± 0 (12)	33 ± 3 (12)	0 ± 0 (12)	0 ± 0 (6)	0 ± 0 (6)
IV			A	82 ± 15 (10)	14 ± 16 (10)	25 ± 1 (12)	0 ± 0 (12)	0 ± 0 (7)	0 ± 0 (7)
IV	A			26 ± 15 (12)	73 ± 16 (12)	79 ± 3 (12)	18 ± 1 (12)	0 ± 0 (9)	0 ± 0 (9)
IV		A	A	4 ± 6 (12)	33 ± 22 (12)	0 ± 0 (12)	9 ± 6 (12)	0 ± 0 (8)	28 ± 28 (8)
IV	A	A		67 ± 24 (12)	33 ± 24 (12)	91 ± 4 (12)	8 ± 4 (12)	1 ± 3 (9)	18 ± 21 (9)
IV	A		A	41 ± 16 (12)	37 ± 21 (12)	16 ± 3 (12)	13 ± 8 (12)	0 ± 0 (9)	25 ± 28 (9)
IV	A	A	A	14 ± 15 (12)	10 ± 13 (12)	11 ± 5 (12)	10 ± 5 (12)	9 ± 11 (9)	9 ± 11 (9)

Table 3.3: Library H3₁₋₂₁ peptides with combinatorial Ac PTMs were used to locate sites of enzymatic acetylation by PCAF HAT/HAT-BRD.

RapifleX MALDI-TOF-MS screen results for PCAF HAT, HAT-BRD and a no enzyme control (NEC) against SETI-IV histone peptide libraries with post translational modifications (PTMs). The results are shown as an average of the observed relative percentages (%) of peak intensities for parent peptide, singly acetylated peptide (1Ac) and doubly acetylated peptide (2Ac) with associated standard deviation (SD) and number of replicates (N) under the format % ± SD (N). The results for parent peptide are not shown for clarity. A gradual colour scale is used to represent observed percentage of peptide ranging from lowest value (red) to highest value (green). The peptides shown are H3₁₋₂₁ peptides with combinations of lysine (K) acetylation (A) PTMs.

- Multiple acetylations

The SETIV library contained combinations of acetylations which showed effects of PCAF substrate selectivity on peptides with combinatorial KAc PTMs.

The acetylation levels for HAT were higher for the SETIV library than for equivalent peptides in SETI-III libraries (Section 3.3.1.1, Table 3.2). SETIV also had issues with high levels of impurity overlap in the NEC and polymer contamination (section 3.2.6). Taking the NEC into account, combinations of K9 and K14 acetylations in peptides resulted in total loss of enzyme activity. H3₁₋₂₁K9AcK14Ac showed <5% enzymatic acetylation for HAT and HAT-BRD (2Ac product was dismissed as impurity overlap since the NEC showed 28% 2Ac). H3₁₋₂₁K4AcK9AcK14Ac gave ≈10% 1Ac and 2Ac with both HAT and HAT-BRD, which was attributed to impurity overlap as the NEC showed 9% 1Ac and 2Ac. K4Ac was unlikely to contribute to observed inhibition of acetylation by H3₁₋₂₁K4AcK9AcK14Ac as H3₁₋₂₁K4Ac alone showed no inhibition of PCAF activity (Table 3.3).

Collectively, these results confirmed that PCAF HAT and HAT-BRD selectively catalysed the acetylation of H3₁₋₂₁ peptides at K14 and K9 with K14 being fully acetylated and K9 showing <50% acetylation under screen conditions. K14 showed full acetylation for both HAT and HAT-BRD. Acetylation of K9 appeared to vary between 40-50% with HAT and 20-30% with HAT-BRD, indicating that the single domain may be more active than the dual domain at H3K9. This may be due to interactions directly with the bromodomain. Alternatively, addition of the bromodomain might change the structure and available surfaces for peptide binding on the HAT domain. In addition, these results indicated that there was no crosstalk between H3K9Ac and H3K14Ac, with neither of these modifications inhibiting acetylation at the other site.

Acetylation activity was observed for other section of H3 including H3₁₄₋₃₄ showing ≈48% 1Ac and ≈14% 2Ac with both HAT and HAT-BRD, H3₈₋₂₈ showing 30-90% 1Ac and 0-27%

2Ac (HAT and HAT-BRD), H₃₈₋₂₈ showing 9-56% 1Ac and <5% 2Ac (HAT and HAT-BRD), H₃₁₈₋₃₈ showing <20% 1Ac (HAT and HAT-BRD), and H₃₂₀₋₄₀ showing ≤10% 1Ac (HAT and HAT-BRD). This indicated that other sites on the H3 peptide could be acetylated by PCAF. The screen results were not clear enough to identify any other PCAF acetylation sites on H3 and MS/MS studies would be necessary to further investigate acetylation selectivity of PCAF against these sections of H3 (Appendix Table D.1, Appendix Table D.3).

3.3.2.2 Site selectivity of PCAF HAT and HAT-BRD for H4

H₄₁₋₂₁ internal control peptides gave only 1Ac product, showing ≈30% 1Ac conversion for both SETII and SETIII H₄₁₋₂₁ peptides with PCAF HAT and HAT-BRD. The SETIV library showed conversions of up to 96% with HAT and 27% with HAT-BRD, consistent with previous results showing higher peptide conversion with HAT in SETIV compared to other libraries. Peptides acetylated at H₄₁₋₂₁K8 gave ≤6% conversion with both HAT and HAT-BRD in SETII-IV. K5Ac, K12Ac and K16Ac PTMs did not significantly affect rates of acetylation either alone or in combinations unless they were combined with a K8Ac PTM (Table 3.4).

These results confirmed that H4K8 was selectively acetylated by PCAF HAT and HAT-BRD, and showed that K5, K12 and K16 Ac PTMs did not affect PCAF activity. PCAF HAT and HAT-BRD showed similar levels of enzymatic acetylation for H4K8, implying that presence of the bromodomain did not affect binding, selectivity or activity of PCAF HAT for H4K8 on H₄₁₋₂₁ peptides.

Other sections of H4 showed significantly less acetylation with H4₆₋₂₆, H4₉₋₂₉, H4₁₃₋₃₃ giving <10% 1Ac and H4₁₁₋₃₁, H4₄₃₋₆₃, H4₅₁₋₇₁ and H4₆₈₋₈₈ giving <5% 1Ac (Appendix Table D.2, Appendix Table D.3, Appendix Table D.4).

H4					1 – SGRG <u>K</u> GG <u>K</u> G LG <u>K</u> GG <u>A</u> KRHR KV – 21			
peptide length: 1-21								
A – Ac								
SET	K 5	K 8	K 12	K 16	HAT		HAT-BRD	
					1Ac	2Ac	1Ac	2Ac
II					28 ± 2 (12)	1 ± 0 (12)	27 ± 6 (12)	0 ± 0 (12)
III					30 ± 6 (3)	0 ± 0 (3)	31 ± 1 (2)	0 ± 0 (2)
II	A		A		27 ± 3 (12)	0 ± 0 (12)	20 ± 4 (8)	0 ± 0 (8)
II		A			4 ± 1 (12)	0 ± 0 (12)	4 ± 1 (12)	0 ± 0 (12)
II				A	31 ± 5 (12)	0 ± 1 (12)	30 ± 6 (12)	0 ± 0 (12)
III	A				31 ± 15 (3)	0 ± 0 (3)	29 ± 0 (2)	0 ± 0 (2)
III	A	A	A		6 ± 1 (3)	0 ± 0 (3)	5 ± 0 (2)	0 ± 0 (2)
III		A			2 ± 1 (3)	0 ± 0 (3)	1 ± 1 (2)	0 ± 0 (2)
IV			A		91 ± 3 (12)	4 ± 2 (12)	25 ± 0 (12)	0 ± 0 (12)
IV				A	96 ± 3 (12)	4 ± 2 (12)	27 ± 1 (12)	0 ± 1 (12)
IV	A	A			5 ± 8 (12)	5 ± 10 (12)	3 ± 0 (12)	0 ± 0 (12)
IV	A		A		93 ± 6 (12)	5 ± 4 (12)	22 ± 2 (12)	0 ± 0 (12)
IV	A			A	96 ± 1 (12)	3 ± 1 (12)	25 ± 1 (12)	0 ± 0 (12)

Table 3.4: Library H4₁₋₂₁ peptides with combinatorial Ac PTMs were used to locate sites of enzymatic acetylation by PCAF HAT/HAT-BRD.

RapifleX MALDI-TOF-MS screen results for PCAF HAT and HAT-BRD against SETI-IV histone peptide libraries with post translational modifications (PTMs). The results are shown as an average of the observed relative percentages (%) of peak intensities for parent peptide, singly acetylated peptide (1Ac) and doubly acetylated peptide (2Ac) with associated standard deviation (SD) and number of replicates (N) under the format % ± SD (N). The results for parent peptide are not shown for clarity. A gradual colour scale is used to represent observed percentage of peptide ranging from lowest % value (red) to highest % value (green). The peptides shown are H4₁₋₂₁ peptides with combinations of lysine (K) acetylation (A) PTMs.

3.3.2.3 Site selectivity of PCAF HAT and HAT-BRD for H2a

H2a₁₋₂₁ internal control peptide gave only 1Ac product for SETII H2a₁₋₂₁, showing 16% 1Ac with HAT and 13% 1Ac with HAT-BRD. SETIV H2a₁₋₂₁ showed up to 69% 1Ac with HAT and up to 6% 1Ac with HAT-BRD, in holding with the increased PCAF HAT activity observed against the SETIV library compared with other libraries (Table 3.5).

H2a					1 – SGRGKQGGK ARAKAKTRSS RA– 21			
peptide length: 1-21								
A – Ac								
SET	K 5	K 9	K 13	K 15	HAT		HAT-BRD	
					1Ac	2Ac	1Ac	2Ac
II					16 ± 3 (12)	0 ± 0 (12)	13 ± 2 (12)	0 ± 1 (12)
II	A				8 ± 1 (12)	0 ± 0 (12)	6 ± 1 (12)	0 ± 0 (12)
II	A	A			0 ± 0 (12)	0 ± 0 (12)	0 ± 1 (12)	0 ± 0 (12)
II		A			2 ± 0 (12)	0 ± 0 (12)	1 ± 0 (12)	0 ± 0 (12)
II		A	A		1 ± 0 (12)	0 ± 0 (12)	1 ± 0 (12)	0 ± 0 (12)
II			A		10 ± 2 (12)	0 ± 0 (12)	9 ± 1 (12)	0 ± 0 (12)
II			A	A	8 ± 3 (12)	0 ± 0 (12)	7 ± 1 (12)	1 ± 1 (12)
II		A	A	A	0 ± 0 (12)	0 ± 1 (12)	0 ± 0 (12)	0 ± 0 (12)
IV	A		A		69 ± 11 (12)	0 ± 1 (12)	6 ± 1 (12)	0 ± 0 (12)
IV	A			A	42 ± 12 (12)	0 ± 0 (12)	6 ± 1 (12)	0 ± 0 (12)
IV			A	A	47 ± 9 (12)	0 ± 0 (12)	5 ± 1 (12)	0 ± 0 (12)
IV	A		A	A	33 ± 11 (11)	0 ± 0 (11)	4 ± 1 (12)	0 ± 0 (12)
IV	A	A			1 ± 2 (12)	0 ± 0 (12)	0 ± 0 (12)	0 ± 0 (12)
IV		A	A		0 ± 0 (12)	0 ± 0 (12)	0 ± 0 (12)	0 ± 0 (12)
IV		A		A	4 ± 3 (12)	0 ± 0 (12)	0 ± 0 (12)	0 ± 0 (12)
IV	A	A	A		0 ± 1 (12)	0 ± 0 (12)	0 ± 0 (12)	0 ± 0 (12)
IV	A	A	A	A	1 ± 2 (12)	0 ± 0 (12)	0 ± 0 (12)	0 ± 0 (12)

Table 3.5: Library H2a₁₋₂₁ peptides with combinatorial Ac PTMs were used to locate sites of enzymatic acetylation by PCAF HAT/HAT-BRD.

RapifleX MALDI-TOF-MS screen results for PCAF HAT and HAT-BRD against SETI-IV histone peptide libraries with post translational modifications (PTMs). The results are shown as an average of the observed relative percentages (%) of peak intensities for parent peptide, singly acetylated peptide (1Ac) and doubly acetylated peptide (2Ac) with associated standard deviation (SD) and number of replicates (N) under the format % ± SD (N). The results for parent peptide are not shown for clarity. A gradual colour scale is used to represent observed percentage of peptide ranging from lowest % value (red) to highest % value (green). The peptides shown are H2a₁₋₂₁ peptides with combinations of lysine (K) acetylation (A) PTMs.

H2a₁₋₂₁ acetylation at K9 inhibited PCAF activity, with all peptides bearing this PTM showing ≤4% 1Ac for SETII/IV with HAT and HAT-BRD. Because SETIV gave much higher levels of PCAF HAT activity compared to SETII, the inhibition seen with H2a₁₋₂₁K9Ac was much more pronounced: SETIV peptides that did not have a K9Ac PTM showed 1Ac levels between 33-69%. Combinations of Ac PTMs at positions K5, K13 and K15 showed some inhibitory activity in SETIV. Peptides containing K15Ac gave lower 1Ac conversion (33-42% 1Ac with HAT) than H2a₁₋₂₁K5K13 peptide (69% 1Ac with HAT). However, there

was no appropriate control for comparisons, and the SD for these peptides was of $\approx 10\%$, so these results may not be meaningful. When K5, K13 and K15 Ac PTMs were combined with K9Ac, 1Ac levels dropped to $\leq 4\%$ (Table 3.5). These results strongly indicated that H2aK9 was the substrate site for enzymatic acetylation by PCAF HAT and HAT-BRD.

3.3.2.4 Site selectivity of PCAF HAT and HAT-BRD for H2b

H2b₁₋₂₁ control peptide gave only 1Ac in SETII showing 8% 1Ac with HAT and 5% 1Ac with HAT-BRD. SETIV H2b₁₋₂₁ showed up to 12% 1Ac with HAT and up to 4% 1Ac with HAT-BRD. As SETIV did not show a large increase in H2b activity compared to SETII for the HAT construct, it is likely that PCAF acetylation activity is lower for H2b than it is for H2a. Ac PTMs at K5, K11, K12, K15 and K16 on H2b₁₋₂₁ were present in a variety of combinations. No site of Ac PTM showed a complete inhibition of activity (Table 3.6). With low activity overall, any differences in activity were not likely to be significant. To better understand the selectivity of PCAF acetylation for H2b, it would be necessary to run an assay with H2b peptides over a longer period or with higher enzyme concentration to observe significant changes in activity between combinations of peptide modifications.

H2b						1 – PEPAKSAPA PKKGSKKAVT KA – 21			
peptide length: 1-21									
A – Ac									
SET	K 5	K 11	K 12	K 15	K 16	HAT		HAT-BRD	
						1Ac	2Ac	1Ac	2Ac
II						8 ± 1 (12)	0 ± 0 (12)	5 ± 2 (12)	0 ± 0 (12)
IV	A					12 ± 8 (12)	0 ± 0 (12)	4 ± 0 (12)	0 ± 0 (12)
II	A					6 ± 1 (12)	0 ± 0 (12)	5 ± 1 (12)	0 ± 0 (12)
IV		A				10 ± 10 (11)	0 ± 0 (11)	4 ± 0 (12)	0 ± 0 (12)
II			A			11 ± 3 (12)	0 ± 0 (12)	5 ± 1 (12)	0 ± 0 (12)
IV	A			A		4 ± 3 (12)	2 ± 3 (12)	4 ± 0 (12)	0 ± 0 (12)
IV	A				A	4 ± 4 (12)	0 ± 1 (12)	4 ± 0 (8)	0 ± 0 (8)
IV		A	A			6 ± 3 (12)	1 ± 2 (12)	4 ± 0 (12)	0 ± 0 (12)
IV		A		A		4 ± 2 (12)	4 ± 6 (12)	3 ± 0 (12)	0 ± 0 (12)
IV		A			A	5 ± 3 (12)	3 ± 5 (12)	3 ± 0 (12)	0 ± 0 (12)
IV		A	A	A		4 ± 1 (12)	1 ± 2 (12)	2 ± 1 (12)	0 ± 0 (12)
IV		A	A		A	5 ± 3 (12)	2 ± 4 (12)	3 ± 0 (12)	0 ± 0 (12)
IV			A	A	A	3 ± 3 (12)	1 ± 2 (12)	1 ± 1 (12)	0 ± 0 (12)
IV	A	A	A			7 ± 6 (12)	5 ± 7 (12)	3 ± 0 (12)	0 ± 0 (12)
IV	A	A	A	A	A	3 ± 3 (12)	0 ± 0 (12)	0 ± 0 (12)	0 ± 0 (12)

Table 3.6: Results for screening of library H2b₁₋₂₁ peptides with combinatorial Ac PTMs with PCAF HAT/HAT-BRD.

RapifleX MALDI-TOF-MS screen results for PCAF HAT and HAT-BRD against SETI-IV histone peptide libraries with post translational modifications (PTMs). The results are shown as an average of the observed relative percentages (%) of peak intensities for parent peptide, singly acetylated peptide (1Ac) and doubly acetylated peptide (2Ac) with associated standard deviation (SD) and number of replicates (N) under the format % ± SD (N). The results for the parent peptide are not shown for clarity. A gradual colour scale is used to represent observed percentage of peptide ranging from lowest % value (red) to highest % value (green). The peptides shown are H2b₁₋₂₁ peptides with combinations of lysine (K) acetylation (A) PTMs.

3.3.4 Effects of combinatorial PTMs on the activity of PCAF HAT and HAT-BRD

Crosstalk between histone PTMs has been shown to profoundly affect the activity of epigenetic proteins.²⁰ The following analysis is centred on exploring the effects of histone PTMs on PCAF HAT/HAT-BRD activity and selectivity.

3.3.4.1 Combinatorial H3R2 and H3K4 PTMs

H3K4 and H3R2 methylation are both well documented PTMs.¹⁷ They have both been linked to H3K9 and H3K14 sites in distinct crosstalk mechanisms.^{46,363}

SETI and SETIII libraries contained H3₁₋₂₁ peptides with combinations of K4Me1/2/3 and R2Me2a. SETI/III H3₁₋₂₁ control peptides showed $\approx 45\%$ 1Ac and $\approx 55\%$ 2Ac with HAT, with $\approx 70\%$ 1Ac and $\approx 25\%$ 2Ac with HAT-BRD (Table 3.7). H3K4Me1/2/3 did not have a strong inhibitory effect on 1Ac or 2Ac PCAF activity. For SETI with PCAF HAT, moderate inhibition of 2Ac was observed by K4 methylation PTMs (36-43% 2Ac with a large SD of 28% for H3₁₋₂₁K4Me3 compared with $53 \pm 6\%$ 2Ac for H3₁₋₂₁ control). This effect was not reproduced in SETIII H3₁₋₂₁ peptides with PCAF HAT where acetylation levels of peptides with K4 PTMs were in line with the control except for H3₁₋₂₁K4Me3 ($31 \pm 26\%$ 2Ac) due to variation between spectra (high SD). For peptides incubated with PCAF HAT-BRD, no inhibition linked to K4 methylation was seen for either SETI or SETIII peptides, with 2Ac conversion staying very close to control levels. Combination of K4Me3 with R2Me2a PTMs on H3₁₋₂₁ peptides did not have any further inhibitory effect on PCAF HAT or HAT-BRD acetylation activity. For comparison, combinations of K4Me3 with T3Phos or K9 PTMs on H3₁₋₂₁ peptides led to strong inhibition of PCAF HAT and HAT-BRD 2Ac activity (Table 3.7).

H3						1 – ARTKQTARK STGGKAPRKQ LA – 21			
						M – Me2a / A – Ac / P – Phos / 1 – Me1 / 2 – Me2 / 3 – Me3			
Lib	R 2	T 3	K 4	K 9	Peptide length	HAT		HAT-BRD	
						1Ac	2Ac	1Ac	2Ac
I					1-21	46 ± 6 (7)	53 ± 6 (7)	70 ± 10 (12)	29 ± 10 (12)
I			1		1-21	62 ± 2 (10)	36 ± 3 (10)	72 ± 4 (8)	26 ± 4 (8)
I			2		1-21	60 ± 4 (11)	39 ± 3 (11)	72 ± 2 (8)	26 ± 3 (8)
I			3		1-21	57 ± 28 (7)	43 ± 28 (7)	73 ± 5 (8)	26 ± 4 (8)
III					1-21	43 ± 15 (3)	56 ± 16 (3)	77 ± 6 (2)	22 ± 5 (2)
III			1		1-21	50 ± 11 (3)	50 ± 12 (3)	74 ± 0 (2)	25 ± 0 (2)
III			2		1-21	41 ± 10 (3)	59 ± 10 (3)	72 ± 2 (2)	27 ± 2 (2)
III			3		1-21	33 ± 25 (3)	31 ± 26 (3)	74 ± 6 (2)	22 ± 2 (2)
I	M		3		1-21	64 ± 2 (12)	34 ± 2 (12)	76 ± 2 (8)	22 ± 2 (8)
I	M	P	3		1-21	89 ± 6 (11)	11 ± 6 (11)	94 ± 3 (8)	5 ± 2 (8)
I		P	3		1-21	89 ± 6 (11)	11 ± 6 (11)	94 ± 5 (8)	6 ± 5 (8)
III	M		3		1-21	56 ± 12 (3)	44 ± 12 (3)	83 ± 1 (2)	17 ± 1 (2)
III	M	P	3		1-21	58 ± 41 (3)	9 ± 8 (3)	95 ± 1 (2)	2 ± 2 (2)
III		P	3		1-21	83 ± 5 (3)	16 ± 5 (3)	94 ± 1 (2)	6 ± 1 (2)
III			3	A	1-21	95 ± 5 (3)	1 ± 1 (3)	100 ± 0 (2)	0 ± 0 (2)
III			3	3	1-21	100 ± 0 (3)	0 ± 0 (3)	100 ± 0 (2)	0 ± 0 (2)

Table 3.7: Effects of H3R2Me2a, H3T3Phos and H3K4Me1/2/3 on PCAF activity were examined.

RapifleX MALDI-TOF-MS screen results for PCAF HAT and HAT-BRD against SET1-IV histone peptide libraries with post translational modifications (PTMs). The results are shown as an average of the observed relative percentages (%) of peak intensities for parent peptide, singly acetylated peptide (1Ac) and doubly acetylated peptide (2Ac) with associated standard deviation (SD) and number of replicates (N) under the format % ± SD (N). The results for the parent peptide are not shown for clarity. A gradual colour scale is used to represent observed percentage of peptide ranging from lowest % value (red) to highest % value (green). The peptides shown are H3₁₋₂₁ peptides with combinations of PTMs: arginine (R) asymmetric di-methylation (M); lysine (K) acetylation (A), mono-methylation (1) and di-methylation (2); threonine (T) phosphorylation (P).

3.3.4.2 Combinatorial H3S10 and H3T11 PTMs

H3S10Phos and H3T11Phos are well documented PTMs which have been shown to impact acetylation activity of GCN5.^{17,45} Liokatis et. al. showed that H3S10Phos and H3T11Phos inhibited acetylation at H3K9 by GCN5. H3S10Phos and H3T11Phos PTMs had different reported effects on H3K14 enzymatic acetylation by GCN5, with H3T11Phos inhibiting acetylation and not H3S10Phos.

SETI and SETIII contained H3₁₋₂₁ peptides with H3S10Phos and H3T11Phos PTMs in combination with acetylation (Ac) or tri-methylation (Me3) modifications at K9 or K14. SETI/III H3₁₋₂₁ internal control peptides showed ≈45% 1Ac and ≈55% 2Ac with HAT, with ≈70% 1Ac and ≈25% 2Ac with HAT-BRD (Table 3.8).

- PCAF activity at H3K9 of H3₁₋₂₁ peptides

Both S10Phos and T11Phos PTMs on H3₁₋₂₁ peptides completely inhibited 2Ac activity (<5% 2Ac conversion compared to >20% 2Ac for control peptides) by PCAF HAT and HAT-BRD in SETI/III. When combined with K9 modifications, a decrease 1Ac activity was observed along with complete inhibition of 2Ac activity. When combined with K14Ac or K14Me3 both 1Ac and 2Ac activities were completely lost (Table 3.8). This was in line with previous observations and literature results: the 2Ac activity of PCAF corresponds to H3K9 enzymatic acetylation whilst the 1Ac activity corresponds to H3K14 enzymatic acetylation. These results confirmed that H3S10Phos and H3T11Phos PTMs inhibited H3K9 acetylation by PCAF HAT and HAT-BRD.

- PCAF activity at H3K14 of H3₁₋₂₁ peptides

Effects of H3S10Phos and H3T11Phos PTMs on H3K14 PCAF acetylation of H3₁₋₂₁ peptides could be seen by changes in the levels of conversion to 1Ac (Table 3.8). H3S10Phos PTMs on H3₁₋₂₁ did not affect 1Ac activity unless in combination with H3T11Phos: H3₁₋₂₁S10Phos, H3₁₋₂₁K9AcS10Phos and H3₁₋₂₁K9Me3S10Phos all gave ≥99% 1Ac conversion with PCAF HAT and HAT-BRD. This also showed that H3K9 PTM did not

affect H3K14 enzymatic acetylation by PCAF in agreement with previously observed results (Section 3.3.2.1, Table 3.3).

H3					1 – ARTKQTARK STGG KAPRKQ LA – 21				
					A – Ac / P – Phos / 3 – Me3				
SET	K 9	S 10	T 11	K 14	Peptide length	HAT		HAT-BRD	
						1Ac	2Ac	1Ac	2Ac
I					1-21	46 ± 6 (7)	53 ± 6 (7)	70 ± 10 (12)	29 ± 10 (12)
III					1-21	43 ± 15 (3)	56 ± 16 (3)	77 ± 6 (2)	22 ± 5 (2)
I			P		1-21	96 ± 5 (5)	2 ± 2 (5)	81 ± 2 (12)	1 ± 1 (12)
III			P		1-21	94 ± 2 (3)	4 ± 1 (3)	77 ± 2 (2)	0 ± 0 (2)
I	A		P		1-21	39 ± 3 (12)	0 ± 1 (12)	13 ± 12 (8)	0 ± 0 (8)
I	3		P		1-21	78 ± 8 (12)	0 ± 0 (12)	65 ± 2 (8)	0 ± 0 (8)
III	A		P		1-21	75 ± 9 (3)	1 ± 1 (3)	22 ± 1 (2)	0 ± 0 (2)
III	3		P		1-21	90 ± 8 (3)	0 ± 1 (3)	55 ± 3 (2)	0 ± 0 (2)
I			P	A	1-21	3 ± 2 (10)	0 ± 0 (10)	3 ± 0 (8)	0 ± 0 (8)
I			P	3	1-21	5 ± 2 (12)	0 ± 0 (12)	0 ± 2 (11)	0 ± 0 (11)
I		P			1-21	99 ± 1 (7)	1 ± 1 (7)	100 ± 1 (8)	0 ± 0 (8)
III		P			1-21	99 ± 1 (3)	1 ± 1 (3)	99 ± 1 (2)	0 ± 0 (2)
I	A	P			1-21	100 ± 1 (10)	0 ± 0 (10)	99 ± 2 (12)	0 ± 0 (12)
I	3	P			1-21	99 ± 2 (12)	0 ± 0 (12)	99 ± 2 (9)	0 ± 0 (9)
III	A	P			1-21	99 ± 1 (3)	1 ± 1 (3)	99 ± 1 (2)	0 ± 0 (2)
III	3	P			1-21	99 ± 1 (3)	0 ± 0 (3)	99 ± 1 (2)	0 ± 0 (2)
I		P	P		1-21	46 ± 4 (12)	0 ± 0 (12)	28 ± 11 (10)	0 ± 0 (10)
III		P	P		1-21	72 ± 18 (3)	0 ± 0 (3)	27 ± 3 (2)	0 ± 0 (2)
I	A	P	P		1-21	12 ± 4 (12)	0 ± 0 (12)	10 ± 1 (8)	0 ± 0 (8)
I	3	P	P		1-21	19 ± 16 (7)	0 ± 0 (7)	12 ± 18 (6)	0 ± 0 (6)
III	A	P	P		1-21	27 ± 9 (3)	0 ± 0 (3)	7 ± 1 (2)	0 ± 0 (2)
III	3	P	P		1-21	54 ± 18 (3)	0 ± 0 (3)	21 ± 2 (2)	0 ± 0 (2)
I		P	P	A	1-21	0 ± 0 (11)	0 ± 0 (11)	0 ± 0 (9)	0 ± 0 (9)
I		P	P	3	1-21	0 ± 0 (12)	0 ± 0 (12)	0 ± 0 (8)	0 ± 0 (8)

Table 3.8: Effects of H3S10Phos and H3T11Phos PTMs on PCAF activity for H3₁₋₂₁ peptides.

RapifleX MALDI-TOF-MS screen results for PCAF HAT and HAT-BRD against SET1-IV histone peptide libraries with post translational modifications (PTMs). The results are shown as an average of the observed relative percentages (%) of peak intensities for parent peptide, singly acetylated peptide (1Ac) and doubly acetylated peptide (2Ac) with associated standard deviation (SD) and number of replicates (N) under the format % ± SD (N). The results for the parent peptide are not shown for clarity. A gradual colour scale is used to represent observed percentage of peptide ranging from lowest % value (red) to highest % value (green). The peptides shown are H3₁₋₂₁ peptides with combinations of PTMs: lysine (K) acetylation (A) and tri-methylation (3); serine (S) and threonine (T) phosphorylation (P).

H3T11Phos PTM on its own on H3₁₋₂₁ peptides showed no inhibition of PCAF HAT activity ($\geq 94\%$ 1Ac conversion) but showed inhibition of PCAF HAT-BRD activity (H3₁₋₂₁T11Phos showed 77-81% 1Ac conversion with HAT-BRD, where H3₁₋₂₁S10Phos showed $\geq 99\%$ 1Ac with HAT-BRD). Combination of T11Phos with K9 PTMs showed inhibition of PCAF HAT and HAT-BRD activity with greater inhibition observed with HAT-BRD (39-90% 1Ac with HAT, 13-65% 1Ac with HAT-BRD). In addition, greater inhibition of PCAF activity was seen for H3₁₋₂₁T11PhosK9Ac (39-75% 1Ac with HAT, 13-22% 1Ac with HAT-BRD) than for H3₁₋₂₁T11PhosK9Me3 (78-90% 1Ac with HAT, 55-65% 1Ac with HAT-BRD). Combining H3T11Phos with H3S10Phos PTMs on H3₁₋₂₁ peptides also inhibited PCAF HAT and HAT-BRD activity (46-72% 1Ac with HAT, 27-28% 1Ac with HAT-BRD). The greatest levels of inhibition of PCAF activity were seen when H3T11Phos PTM was combined with both S10Phos and K9Me3 or K9Ac PTMs on H3₁₋₂₁ peptides (12-54% 1Ac with HAT, 7-21% 1Ac with HAT-BRD).

3.3.4.3 Combinatorial H3T3Phos, H4S1Phos, H2aS1Phos and H3S28Phos PTMs

There are many phosphorylation sites on histone tails at serine and threonine residues, and many of these have been implicated in crosstalk with acetylation and methylation PTMs.^{17,78} Several phosphorylation PTMs other than H3S10Phos and H3T11Phos were found to affect PCAF activity.

H3		1 – ARTK QTARK STGGKAPRKQ LA – 21											
H4		1 – SGRGK GGKG LGKGGAKRHR KV – 21											
H2a		1 – SGRGK QGGK ARAKAKTRSS RA– 21											
Peptide length: 1-21													
M – Me2a / A – Ac / P – Phos / 1 – Me1 / 3 – Me3 / * – N-term acetyl													
Lib	H3			H4			H2a		H	HAT		HAT-BRD	
	R	T	K	S	R	K	S	K		1Ac	2Ac	1Ac	2Ac
	2	3	4	1	3	5	1	5					
I									H3	46 ± 6 (7)	53 ± 6 (7)	70 ± 10 (12)	29 ± 10 (12)
III									H3	43 ± 15 (3)	56 ± 16 (3)	77 ± 6 (2)	22 ± 5 (2)
I	M		3						H3	64 ± 2 (12)	34 ± 2 (12)	76 ± 2 (8)	22 ± 2 (8)
III	M		3						H3	56 ± 12 (3)	44 ± 12 (3)	83 ± 1 (2)	17 ± 1 (2)
I	M	P	3						H3	89 ± 6 (11)	11 ± 6 (11)	94 ± 3 (8)	5 ± 2 (8)
III	M	P	3						H3	58 ± 41 (3)	9 ± 8 (3)	95 ± 1 (2)	2 ± 2 (2)
I			P	3					H3	89 ± 6 (11)	11 ± 6 (11)	94 ± 5 (8)	6 ± 5 (8)
III			P	3					H3	83 ± 5 (3)	16 ± 5 (3)	94 ± 1 (2)	6 ± 1 (2)
II									H4	28 ± 2 (12)	1 ± 0 (12)	27 ± 6 (12)	0 ± 0 (12)
III									H4	30 ± 6 (3)	0 ± 0 (3)	31 ± 1 (2)	0 ± 0 (2)
II					1				H4	34 ± 3 (12)	1 ± 1 (12)	33 ± 5 (12)	0 ± 0 (12)
III					1				H4	30 ± 6 (3)	0 ± 0 (3)	30 ± 1 (2)	0 ± 0 (2)
II				P					H4	7 ± 5 (12)	0 ± 0 (12)	8 ± 8 (12)	0 ± 0 (12)
III				P					H4	11 ± 3 (3)	0 ± 0 (3)	11 ± 0 (2)	0 ± 0 (2)
II				P	M	A			H4	8 ± 8 (12)	0 ± 0 (12)	1 ± 3 (11)	0 ± 0 (11)
II									H2a	16 ± 3 (12)	0 ± 0 (12)	13 ± 2 (12)	0 ± 1 (12)
II							P		H2a	0 ± 0 (12)	0 ± 0 (12)	1 ± 2 (11)	1 ± 1 (11)
II							P	A	H2a	0 ± 0 (12)	0 ± 0 (12)	0 ± 0 (12)	0 ± 0 (12)
II								A	H2a	8 ± 1 (12)	0 ± 0 (12)	6 ± 1 (12)	0 ± 0 (12)

Table 3.9: Effects of H3T3Phos, H4S1Phos and H2aS1Phos on PCAF activity were examined.

RapifleX MALDI-TOF-MS screen results for PCAF HAT and HAT-BRD against SETI-IV histone peptide libraries with post translational modifications (PTMs). The results are shown as an average of the observed relative percentages (%) of peak intensities for parent peptide, singly acetylated peptide (1Ac) and doubly acetylated peptide (2Ac) with associated standard deviation (SD) and number of replicates (N) under the format % ± SD (N). The results for parent peptide are not shown for clarity. A gradual colour scale is used to represent observed percentage of peptide ranging from lowest % value (red) to highest % value (green). The peptides shown are H3₁₋₂₁, H4₁₋₂₁, and H2a₁₋₂₁ peptides with combinations of PTMs: arginine (R) mono-methylation (1) and asymmetric di-methylation (M); lysine (K) acetylation (A), mono-methylation (1) and tri-methylation (3); serine (S) and threonine (T) phosphorylation (P).

H3T3Phos showed inhibition of H3K9 acetylation by PCAF on H3₁₋₂₁ peptides. This modification was not present in the libraries on its own but in combination with H3R2Me2a and H3K4Me3 PTMs. H3₁₋₂₁R2Me2aK4Me3 had very little effect on PCAF activity showing 34-44% 2Ac with HAT and 17-22% 2Ac with HAT-BRD compared with

observed conversions of $\approx 55\%$ 1Ac with HAT and $\approx 25\%$ with HAT-BRD for H3₁₋₂₁ control peptides. When the H3T3Phos PTM was introduced onto H3₁₋₂₁ peptides, acetylation levels dropped to 9-16% 2Ac with HAT and 2-6% 2Ac with HAT-BRD. This indicated a strong inhibition of H3K9 enzymatic acetylation activity by PCAF. However, H3T3Phos PTM did not appear to affect H3K14 enzymatic acetylation by PCAF on H3₁₋₂₁ peptides (the turnover to 1Ac remained $\geq 95\%$ with both HAT and HAT-BRD, with 1Ac and 2Ac products taken into account). H4S1Phos PTM had a similar inhibitory effect on H4K8 enzymatic acetylation by PCAF on H4₁₋₂₁ peptide with acetylation levels of 7-11% 1Ac with HAT and 1-11% 1Ac with HAT-BRD compared to control levels for H4₁₋₂₁ peptide of $\approx 32\%$ 1Ac with both HAT and HAT-BRD. H2aS1Phos PTM inhibited H2aK9 enzymatic acetylation by PCAF on H2a₁₋₂₁ peptides showing $\leq 1\%$ 1Ac with HAT and HAT-BRD compared to $\approx 14\%$ 1Ac with HAT and HAT-BRD for H2a₁₋₂₁ control peptide (Table 3.9).

The effects of phosphorylation of H3S28 and H3T32 were also explored on H3 peptides ranging from residues 14 to 40. The range of peptides used complicated analysis due to lack of appropriate controls, and poorer understanding of PCAF acetylation sites on these peptides. The H3₁₄₋₃₄ control peptide showed $\approx 48\%$ 1Ac and $\approx 14\%$ 2Ac with both PCAF HAT and HAT-BRD. One acetylation was likely to be situated at H3K14Ac and the other at H3K18Ac of the H3₁₄₋₃₄ peptides, based on the fact that the H3₂₀₋₄₀ peptides showed $\leq 3\%$ 1Ac for both HAT and HAT-BRD. The 1Ac product was a substrate for the second acetylation, however, the sum of the observed percentages for 1Ac and 2Ac products was $< 65\%$ showing that neither acetylation site was fully acetylated. H3S8Phos PTMs on H3 peptides appeared to inhibit PCAF HAT-BRD enzyme activity to a stronger

degree than that of PCAF HAT. H3₁₄₋₃₄S28Phos showed <4% 2Ac with both HAT and HAT-BRD but showed 37-45% 1Ac with HAT and 18-26% 1Ac with HAT-BRD. This effect could not be properly cross-verified as other peptides with the H3S28Phos modification on H3 gave poor overall conversion (SETIII H3₂₀₋₄₀ peptides showed $\leq$ 3% 1Ac with both HAT and HAT-BRD). Effects of H3T32Phos were similarly difficult to quantify due to lack of controls and poor conversions. H3₁₈₋₃₈S28PhosT32Phos showed no acetylation activity by PCAF but was the only peptide in the libraries covering that section of H3, so no comparisons could be made. H3₂₈₋₄₈T32Phos showed <2% 1Ac compared with 8-10% 1Ac for the control with both HAT and HAT-BRD, which was too small a difference to have confidence in the inhibitory effect without more data. Similarly, H3S32Phos on H2b appeared to inhibit PCAF HAT-BRD acetylation (H2b₂₇₋₄₇S28Phos and H2b₂₇₋₄₇S28PhosS36Phos gave $\leq$ 2% 1Ac compared with 12% 1Ac for the H2b₂₇₋₄₇ control with PCAF HAT-BRD), but the conversion of the control and the range for inhibition was too small to have confidence in this effect (Table 3.10).

To properly study the effects of these modifications on PCAF HAT activity, more control peptides and higher conversions would be needed. Higher conversion might be achieved with longer incubation times or higher enzyme concentrations. Identifying the site of acetylation for the peptide segments concerned would also be necessary to understand effects of the PTMs on PCAF HAT and HAT-BRD activity.

H3						14 – KAPRKQ LATKAARKSA PATGGVKKPH R – 40					
H2b						27 – KKR KRSRKESYSV YVYKVLKQ– 47					
M – Me2a / A – Ac / P – Phos / 3 – Me3 / * – N-term Ac											
Lib	H3			H2b		H	Peptide length	HAT		HAT-BRD	
	K 27	S 28	T 32	S 32	S 36			1Ac	2Ac	1Ac	2Ac
I		P	P			H3	18-38	0 ± 0 (10)	0 ± 0 (10)	0 ± 0 (7)	0 ± 0 (7)
I						H3	14-34	49 ± 2 (11)	15 ± 3 (11)	47 ± 3 (12)	13 ± 3 (12)
I	A	P				H3	14-34	37 ± 2 (12)	3 ± 2 (12)	18 ± 8 (7)	0 ± 0 (7)
I	3	P				H3	14-34	37 ± 4 (9)	3 ± 2 (9)	26 ± 3 (8)	1 ± 1 (8)
I		P				H3	14-34	45 ± 17 (12)	4 ± 3 (12)	26 ± 5 (12)	2 ± 2 (12)
I						H3	28-48	8 ± 3 (9)	0 ± 0 (9)	11 ± 2 (8)	0 ± 0 (8)
I			P			H3	28-48	1 ± 1 (9)	0 ± 0 (9)	2 ± 1 (8)	0 ± 0 (8)
III						H3	20-40*	2 ± 1 (3)	0 ± 0 (3)	2 ± 0 (2)	0 ± 0 (2)
III	A	P				H3	20-40*	2 ± 1 (3)	0 ± 0 (3)	0 ± 0 (2)	0 ± 0 (2)
III	3	P				H3	20-40*	1 ± 1 (3)	0 ± 0 (3)	0 ± 0 (2)	0 ± 0 (2)
III		P				H3	20-40*	3 ± 1 (3)	0 ± 0 (3)	0 ± 0 (2)	0 ± 0 (2)
II						H2b	27-47	10 ± 1 (11)	1 ± 1 (11)	12 ± 5 (12)	1 ± 2 (12)
II				P		H2b	27-47	10 ± 4 (12)	0 ± 0 (12)	2 ± 5 (7)	0 ± 0 (7)
II				P	P	H2b	27-47	6 ± 8 (11)	0 ± 0 (11)	1 ± 3 (12)	0 ± 0 (12)
II					P	H2b	27-47	9 ± 3 (11)	0 ± 0 (11)	7 ± 4 (12)	0 ± 0 (12)

Table 3.10: Effects of H3S28Phos, H3T32Phos and H2aS32Phos on PCAF activity were examined.

RapifleX MALDI-TOF-MS screen results for PCAF HAT and HAT-BRD against SETI-IV histone peptide libraries with post translational modifications (PTMs). The results are shown as an average of the observed relative percentages (%) of peak intensities for parent peptide, singly acetylated peptide (1Ac) and doubly acetylated peptide (2Ac) with associated standard deviation (SD) and number of replicates (N) under the format % ± SD (N). The results for parent peptide are not shown for clarity. A gradual colour scale is used to represent observed percentage of peptide ranging from lowest % value (red) to highest % value (green). The peptides shown are H3₁₈₋₃₈, H3₁₄₋₃₄, H3₂₀₋₄₀, H3₂₈₋₄₈, and H2a₂₇₋₄₇ peptides with combinations of PTMs: lysine (K) acetylation (A) and tri-methylation (3); serine (S) and threonine (T) phosphorylation (P).

3.3.4.5 Combinatorial mono-, di- and tri-methylation PTMs

Mono-, di- and tri-methylation are common epigenetic modifications and have therefore been included in the histone peptide libraries at numerous biologically relevant sites on the peptides.¹⁷ Because methylation and acetylation both occur on lysine residues, they are mutually exclusive and often have opposite biological effects.^{224,225} It is therefore important to map crosstalk between methylation and acetylation PTMs (Table 3.11).

- H3K14 methylation

H3₁₋₂₁ control peptide for SETI gave 46% 1Ac and 53% 2Ac with HAT and gave 70% 1Ac and 29% 2Ac with HAT-BRD. As previously discussed (Section 3.3.2.1, Table 3.3), H3₁₋₂₁K14Ac peptide showed complete loss of 2Ac giving 42% 1Ac with HAT and 33% 1Ac with HAT-BRD corresponding to PCAF enzymatic acetylation of H3₁₋₂₁K9Ac. Switching acetylation for mono-methylation at K14 had a strong activating effect on PCAF enzymatic acetylation with SETI H3₁₋₂₁K14Me1 peptide giving 56% 1Ac and 41% 2Ac with HAT and giving 54% 1Ac and 44% 2Ac with HAT-BRD. This activating effect was not observed for di- and tri-methylation. SETI H3₁₋₂₁K14Me2 and H3₁₋₂₁K14Me3 peptides showed $\approx$ 42% 1Ac with HAT and $\approx$ 35% 1Ac with HAT-BRD, in range with SETI H3₁₋₂₁K14Ac peptide conversion (Table 3.11).

SETIII had an equivalent set of Ac and methylation PTMs on H3₈₋₂₈ peptides, but with no internal control peptide for comparative purposes. H3₈₋₂₈K14Ac gave $37 \pm 11\%$ 1Ac with HAT and 19% 1Ac with HAT-BRD. Taking the error for PCAF HAT against H3₈₋₂₈K14Ac into

account, SETIII H3₈₋₂₈K14Ac showed slightly lower K9 conversion levels than SETI H3₁₋₂₁K14Ac (42% 1Ac with HAT and 33% 1Ac with HAT-BRD), likely due to loss of binding interactions with amino acids preceding K9 in the peptide sequence. H3₈₋₂₈K14Me1 gave 67% 1Ac and 30% 2Ac with HAT and gave 73% 1Ac and 20% 2Ac with HAT-BRD, mirroring the activating effect of the K14Me1 PTM on PCAF observed with the H3₁₋₂₁K14Me1 peptide. H3₈₋₂₈K14Me2 and H3₈₋₂₈K14Me3 showed no activating effect on PCAF HAT acetylation activity, showing conversions of $\approx$ 39% 1Ac with HAT/HAT-BRD and $\approx$ 21% 2Ac with HAT/HAT-BRD, in line with H3₈₋₂₈K14Ac acetylation levels (Table 3.11).

Observing the activating effect of the K14Me1 PTM in two different libraries (SETI and SETIII), on two different peptide lengths (H3₁₋₂₁ and H3₈₋₂₈) and with two different constructs (HAT and HAT-BRD) gave confidence in the validity of the observations, despite the low replicate number for SETIII (2-3 replicates vs. 9-10 replicates for SETI, because of low-throughput method used), and the large SD observed with HAT-BRD against H3₁₋₂₁K14Me1 in SETI (18-19% SD) and with HAT against H3₈₋₂₈K14Me1 in SETIII (10-12% SD).

- H4K12 and H2bK23 methylation

SETII H4₁₋₂₁ internal control peptide gave $\approx$ 28% 1Ac with both HAT and HAT-BRD and no 2Ac activity. H4₁₋₂₁K12Me1 gave $50 \pm 19\%$ 1Ac with HAT and $40 \pm 8\%$ 1Ac with HAT-BRD which might indicate an increase in PCAF activity, but the high SD values are suspect. H4₁₋₂₁K12Me2 gave $\approx$ 28% 1Ac with both HAT and HAT-BRD in line with acetylation levels of the control. These results could indicate activation of PCAF acetylation activity by the H4K12Me1 PTM. In SETIII however, H4₁₋₂₁control, H4₁₋₂₁K12Me1, H4₁₋₂₁K12Me2 and

H4₁₋₂₁K12Me₃ all showed 22-31% 1Ac for both HAT and HAT-BRD, with SD ≤8%. This lack of validation between the two libraries indicated that H4₁₋₂₁K12Me₁ activation of PCAF was likely to be an artifact of the screen (Table 3.11).

H2b₁₅₋₃₅ also showed possible activation of PCAF acetylation activity by a H2bK14Me₁ PTM with H2b₁₅₋₃₅control giving 6% 1Ac and H2b₁₅₋₃₅K23Me₁ giving 9% 1Ac with HAT and HAT-BRD. However, as conversions of <10% 1Ac were observed for these peptides, these results must be analysed with caution, as the difference in activity is well within observed SD ranges for the screen.

- Site specific mono-methylation effect

Increased activity with mono-methylation vs. acetylation and di- or tri-methylation at H3₁₋₂₁K14Me₁ and H3₈₋₂₈K14Me₁ was a site-specific effect. H3₁₋₂₁K9Me₁ and H3₁₄₋₃₄K27Me₁ showed no difference in activity compared to Ac, Me₂ or Me₃ PTM peptide variants. Multiple other methylation sites showed no change in activity with Me_{1/2/3}, or showed inhibition of activity at similar rates for Me_{1/2/3} (Appendix D).

H3					1 – ARTKQTARKK STGGKAPRKQ LATKAARKS – 28					
H4					1 – SGRGKGGKG LGKGGAKRHR KV – 21					
H2b					15 – KKA VT KAQKKGDKKR KRSRKE – 35					
A – Ac / 1 – Me1 / 2 – Me2 / 3 – Me3 / * – N-term acetyl										
Lib	H3		H4	H2 b	H	Peptide length	HAT		HAT-BRD	
	K 9	K 14	K 12	K 23			1Ac	2Ac	1Ac	2Ac
I					H3	1-21	46 ± 6 (7)	53 ± 6 (7)	70 ± 10 (12)	29 ± 10 (12)
I		A			H3	1-21	42 ± 4 (12)	1 ± 0 (12)	33 ± 3 (12)	0 ± 0 (12)
I		1			H3	1-21	56 ± 6 (9)	41 ± 7 (9)	54 ± 18 (10)	44 ± 19 (10)
I		2			H3	1-21	43 ± 4 (12)	2 ± 1 (12)	38 ± 2 (8)	1 ± 1 (8)
I		3			H3	1-21	42 ± 4 (8)	1 ± 1 (8)	32 ± 13 (8)	0 ± 0 (8)
III		A			H3	8-28*	37 ± 11 (3)	3 ± 3 (3)	19 ± 0 (2)	1 ± 0 (2)
III		1			H3	8-28*	67 ± 10 (3)	30 ± 12 (3)	73 ± 0 (2)	20 ± 2 (2)
III		2			H3	8-28*	39 ± 8 (3)	4 ± 2 (3)	24 ± 1 (2)	2 ± 0 (2)
III		3			H3	8-28*	40 ± 9 (3)	5 ± 3 (3)	19 ± 7 (2)	0 ± 0 (2)
I	A				H3	1-21	96 ± 1 (12)	3 ± 0 (12)	95 ± 1 (12)	2 ± 1 (12)
I	1				H3	1-21	93 ± 1 (12)	5 ± 1 (12)	95 ± 2 (12)	3 ± 1 (12)
I	2				H3	1-21	97 ± 2 (12)	3 ± 2 (12)	97 ± 2 (11)	2 ± 2 (11)
I	3				H3	1-21	95 ± 2 (12)	5 ± 2 (12)	95 ± 1 (7)	4 ± 1 (7)
II					H4	1-21	28 ± 2 (12)	1 ± 0 (12)	27 ± 6 (12)	0 ± 0 (12)
II			1		H4	1-21	50 ± 19 (12)	0 ± 0 (12)	40 ± 8 (12)	0 ± 0 (12)
II			2		H4	1-21	30 ± 4 (12)	1 ± 1 (12)	27 ± 3 (12)	1 ± 1 (12)
III					H4	1-21	30 ± 6 (3)	0 ± 0 (3)	31 ± 1 (2)	0 ± 0 (2)
III			1		H4	1-21	30 ± 7 (3)	0 ± 0 (3)	28 ± 1 (2)	0 ± 0 (2)
III			2		H4	1-21	28 ± 6 (3)	0 ± 0 (3)	28 ± 0 (2)	0 ± 0 (2)
III			3		H4	1-21	22 ± 8 (3)	0 ± 0 (3)	25 ± 4 (2)	0 ± 0 (2)
II					H2b	15-35	6 ± 2 (12)	0 ± 0 (12)	6 ± 1 (12)	0 ± 0 (12)
II				1	H2b	15-35	9 ± 3 (12)	0 ± 0 (12)	9 ± 4 (12)	0 ± 1 (12)
II				2	H2b	15-35	3 ± 3 (12)	0 ± 0 (12)	2 ± 3 (12)	0 ± 0 (12)
II				3	H2b	15-35	3 ± 4 (12)	0 ± 0 (12)	1 ± 3 (12)	0 ± 0 (12)

Table 3.11: Effects of lysine mono-, di- or tri-methylation on PCAF activity were examined.

RapifleX MALDI-TOF-MS screen results for PCAF HAT and HAT-BRD against SETI-IV histone peptide libraries with post translational modifications (PTMs). The results are shown as an average of the observed relative percentages (%) of peak intensities for parent peptide, singly acetylated peptide (1Ac) and doubly acetylated peptide (2Ac) with associated standard deviation (SD) and number of replicates (N) under the format % ± SD (N). The results for parent peptide are not shown for clarity. A gradual colour scale is used to represent observed percentage of peptide ranging from lowest % value (red) to highest % value (green). The peptides shown are peptides of selected portions of H3, H4 and H2b with combinations of PTMs: lysine (K) acetylation (A), mono-methylation (1), di-methylation (2) and tri-methylation (3).

3.4 Results for rapifleX MALDI-TOF-MS HTS against SIRT1

HDAC enzymes can also be used in the high-throughput MALDI-TOF-MS screen developed, by monitoring for -42 Da mass changes for the library peptides. SIRT1 was screened against the SETIV histone peptide library, with screen conditions for PCAF adapted for SIRT1 (Section 3.2.3). The SETIV library consists of combinatorial Ac PTMs on H3, H4, H2a and H2b.

The SETIV library showed a high number of sub-optimally synthesised peptides. For seven out of eight control peptides, the peptide masses were not found in the spectra, and showed many impurity peaks (Figure 3.8a box 1). For certain library members incubated with SIRT1, peaks were identified at masses which corresponded to deacetylation of a greater number of acetyl groups than were present on the peptide (Figure 3.8a boxes 1, 2, 3 and 4). These peaks were not identified in the NEC for the library indicating that they did not correspond to side products from peptide synthesis. Examination of the spectra showed contamination of samples by peaks separated by 44 Da (Figure 3.8b). A regular separation of peaks is indicative of a polymer contaminant, and a 44 Da separation is reported to correspond to either polyethylene glycol (PEG), triton X-100 (triton) or polyethylene glycol sorbitan monolaurate (tween).^{364–366} Tween was not used in the buffers for these assays and is an unlikely source of contamination. PEG is a surfactant found in laboratory plasticware, glue, gloves and hand sanitiser. Triton is a detergent which is an ingredient in cleaning products used in the laboratory. These are both likely sources of contamination with PEG being the more common contaminant.^{364–366} Impurity peaks were also observed in the NEC samples (Figure 3.8a)

and manual examination of spectra showed some polymer contamination in the NEC samples. Visual inspection of MALDI-TOF-MS spectra for the SETIV library plate and assays run for the SETIV library with PCAF HAT/HAT-BRD and SIRT1 showed that only samples spotted with the mosquito liquid handler and run with rapifleX MALDI-TOF-MS contained polymer contamination. In addition, no pattern could be discerned for which assay wells were contaminated, and in some instances not all technical spotting replicates showed polymer contamination. It is therefore likely that contamination was introduced through plastic 384 well plates, or consumables used for the mosquito liquid handler, or through a brand of plate seals used for which parts of the adhesive remained on the plate surface after defrosting of plates stored at -20°C.

For library members where the peptide mass was in the same range as the polymer mass, overlap between polymer peaks and peaks corresponding to peptide deacetylation was observed (Figure 3.8b). Polymer contamination therefore led to a higher rate of false positives in the screen results for SETIV compared with SETI, SETII and SETIII libraries where no polymer contamination was observed. Because polymer contamination was not equally distributed between assay replicates, polymer contamination contributed to high SD values for concerned samples, and was therefore accounted for in data analysis without need for further manipulation of spectra or data.

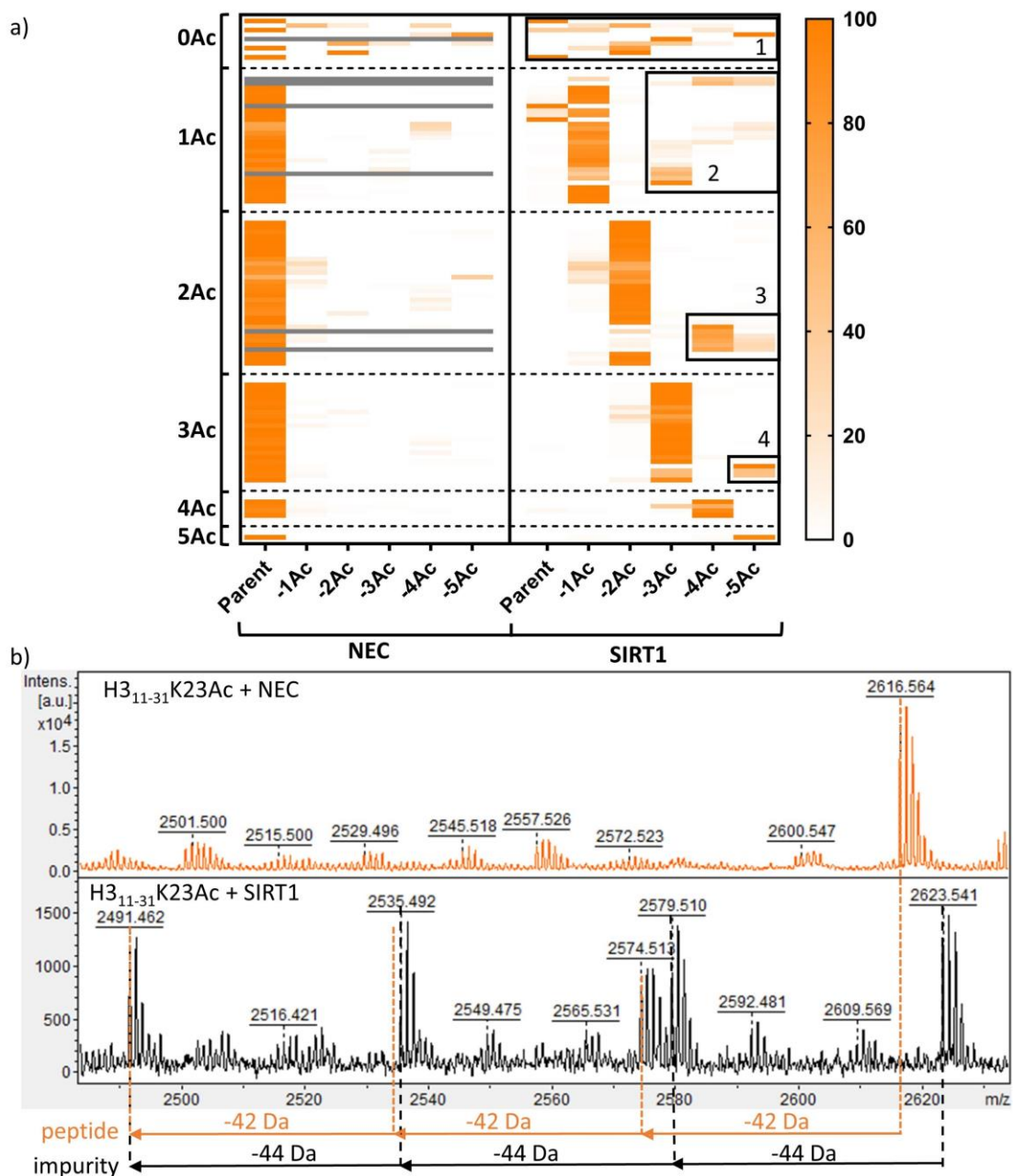


Figure 3.8 Overview of the SETIV library screened against SIRT1 in a rapifleX MALDI-TOF-MS HTS.

a) Results are shown for screening of the peptides against SIRT1 with the no enzyme control (NEC) on the left. Enzyme activity for each section (NEC, SIRT1) is being measured in terms of observed relative percentage of parent peptide and 5 potential deacetylation product peptides (-1 to -5Ac). The relative percentage is shown as a colour scale from white (0% observed peptide) to orange (100% observed peptide). Full results for this screen can be found in Appendix Table D.5. **b)** Selected NEC and product spectra for H3₁₁₋₃₁K23Ac peptide are shown, to illustrate the overlap between deacetylation peaks looked for and a polymer contaminant in the assay which led to a high rate of false-positives in the library results.

Although polymer contamination complicated data analysis, the SETIV library screen with SIRT1 showed a clear overall trend. Complete deacetylation of library peptides was

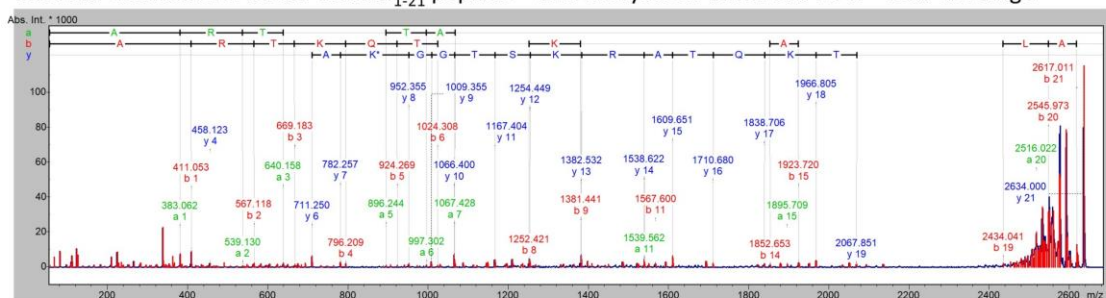
observed regardless of the number of Ac PTMs or the peptide sequence, with all sections of H3, H4, H2a and H2b being equally suitable substrates for SIRT1 under the assay conditions (Figure 3.8a). Full results for this screen can be found in Appendix Table D.5.

3.5 Results for rapifleX MALDI-TOF/TOF-MS of HTS samples

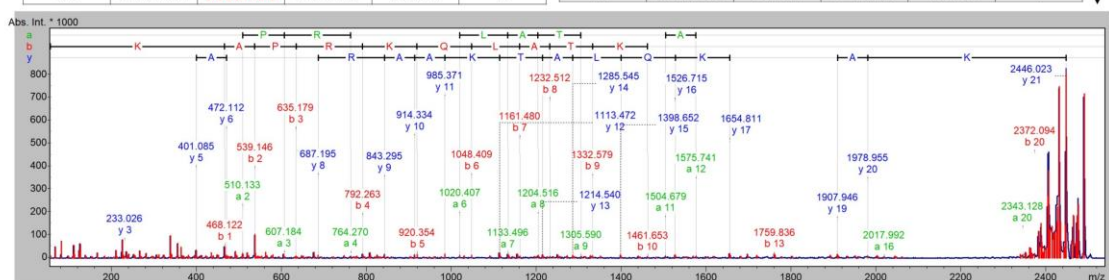
RapifleX is capable of MS/MS data acquisition (Figure 2.2), both in manual and automatic mode. MALDI-TOF/TOF-MS using manual mode for clean purified H3₁₋₂₁ peptides used in assays with PCAF HAT/HAT-BRD gave quality MS/MS data (Section 2.2.4.1). As the rapifleX has an automatic acquisition mode for MALDI-TOF/TOF-MS, the instrument has the capacity to run MS/MS as a high throughput process. This would be interesting to develop in the context of the HTS carried out for PCAF to identify the site selectivity of the PCAF HAT and HAT-BRD enzymes for library peptides.

Initial tests were run, using automatic MALDI-TOF/TOF-MS acquisition on a 384 spot MALDI plate, on which was spotted two assay replicates of a NEC with the SETI library, with two spotting replicates for each assay well. The automatic mode was set to acquire MS/MS results for the ten most intense peaks of each spectrum. For clean peptides this setting is sufficient to acquire MS/MS for each of the parent peptide and possible product peptides of PCAF acetylation. Certain low-quality peptides with numerous impurities had more than 10 impurity peaks with higher intensities than the peptide peaks, in which case MS/MS for the peptide peaks were not acquired using the automatic method. A total of 3111 MS/MS spectra were acquired for the 384 spot MALDI plate using these parameters. Of the spectra acquired, 432 MS/MS spectra were acquired for peaks with masses corresponding to library peptides.

Internal Control: SETI biotin-Ahx-H3₁₋₂₁ peptide – first acetylation with PCAF HAT – 95% coverage



Biotin-Ahx-H3 ₁₋₂₁ K14Ac						Biotin-Ahx-H3 ₁₄₋₃₄ K??Ac					
Calc. mass: 2634.000						Calc. mass: 2528.000					
MH ⁺ (mono): 2633.502						MH ⁺ (mono): 2528.425					
ion #	a	b	sequence	y	ion #	ion #	a	b	sequence	y	ion #
1	382.203	410.198	Biotin-AHX-A	2633.502	21	1	439.261	467.256	Biotin-AHX-K	2445.411	21
2	538.304	566.299	R	2224.311	20	2	510.298	538.293	A	1979.162	20
3	639.352	667.347	T	2068.21	19	3	607.351	635.346	P	1908.125	19
4	767.447	795.442	K	1967.162	18	4	763.452	791.447	R	1811.072	18
5	895.506	923.501	Q	1839.067	17	5	891.547	919.542	K	1654.971	17
6	996.553	1024.548	T	1711.009	16	6	1019.606	1047.601	Q	1526.876	16
7	1067.59	1095.585	A	1609.961	15	7	1132.69	1160.685	L	1398.818	15
8	1223.692	1251.686	R	1538.924	14	8	1203.727	1231.722	A	1285.734	14
9	1351.787	1379.781	K	1382.823	13	9	1304.775	1332.769	T	1214.696	13
10	1438.819	1466.813	S	1254.728	12	10	1432.87	1460.864	K	1113.649	12
11	1539.866	1567.861	T	1167.696	11	11	1503.907	1531.902	A	985.554	11
12	1596.888	1624.883	G	1066.648	10	12	1574.944	1602.939	A	914.517	10
13	1653.909	1681.904	G	1009.627	9	13	1731.045	1759.04	R	843.48	9
14	1824.015	1852.01	KAc (+42)	952.605	8	14	1859.14	1887.135	K	687.378	8
15	1895.052	1923.047	A	782.5	7	15	1946.172	1974.167	S	559.283	7
16	1992.105	2020.099	P	711.462	6	16	2017.209	2045.204	A	472.251	6
17	2148.206	2176.201	R	614.41	5	17	2114.262	2142.257	P	401.214	5
18	2276.301	2304.296	K	458.309	4	18	2185.299	2213.294	A	304.162	4
19	2404.359	2432.354	Q	330.214	3	19	2286.347	2314.341	T	233.124	3
20	2517.443	2545.438	L	202.155	2	20	2343.368	2371.363	G	132.077	2
21	2587.496	2615.491	A-amide	89.071	1	21	2399.405	2427.4	G-amide	75.055	1



SETI biotin-Ahx-H3₁₄₋₃₄ peptide – first acetylation with PCAF HAT – 85% coverage, acetylation site not identified.

Figure 3.9: MS/MS spectra of SETI library peptides incubated with PCAF HAT.

MS/MS spectra acquired using MALDI-TOF/TOF-MS and analysed using Biotoools are shown. Top: MS/MS spectrum of the single acetylation product of SETI biotin-ahx-H3₁₋₂₁-NH₂ peptide treated with PCAF HAT, identifying K14 as acetylation site. Bottom: MS/MS spectrum of the single acetylation product of SETI biotin-ahx-H3₁₄₋₃₄-NH₂ peptide treated with PCAF HAT, in which the acetylation site could not be definitively identified.

Biotoools software (Bruker) was previously used to analyse MALDI-TOF/TOF-MS spectra.

This software allows for manual analysis of spectra but is not adaptable to high throughput MS/MS analysis. Development of alternative analysis methods would be

necessary to achieve high-throughput MS/MS screening, as manual analysis of ~432 spectra per 384 MALDI spots is not realistic.

MALDI-TOF/TOF-MS spectra were manually acquired for selected peptides in the SETI-IV libraries screened against PCAF HAT/HAT-BRD, where site selectivity of PCAF needs to be determined for better understanding of enzyme activity. The quality of the peptides in the libraries was variable as they were not purified or quantified. Level of peptide conversion to a peptide product of interest was dependent on peptide sequence and PTMs. Both of these factors affected whether peak intensity was sufficient for quality MS/MS spectra to be acquired. The spectra acquired for the screened peptides were of variable quality (insufficient peak intensity gives insufficient fragmentation for definitive identification of PTM sites) and site selectivity of PCAF for most peptides could not be properly determined.

For the SETI tagged internal control peptide biotin-Ahx-H3₁₋₂₁-NH₂ incubated with PCAF HAT, acetylation by PCAF at K14 was correctly identified as the 1Ac site. Sequence coverage of fragments was 95% and scoring of alternative sequences showed K14 acetylation as being four times more likely than acetylation at another position. This showed that the acetylation site of PCAF on library peptides can be correctly identified with the Biotoools software, with the *N*-terminal biotin-Ahx tag and *C*-terminal amide successfully taken into account (Figure 3.9, Table 3.12). For SETI biotin-Ahx-H3₁₄₋₃₄-NH₂ incubated with PCAF HAT, a lower sequence coverage of up to 85% was attained and scoring of the different sequence alternatives did not give sensible results. Sequences with no acetylations, K14Ac, K18Ac, K23Ac, K14Ac + K18Ac and K14Ac + K23Ac

acetylation combinations were all scored as equally likely acetylation sites, even though the mass of the parent peptide corresponded to a 1Ac product. For SETII NH₂-H2a₁₋₂₁-Ahx-Lys-biotin incubated with PCAF HAT, whilst 95% sequence coverage was attained, biotools scoring for the different sequence alternatives did not give sensible results, with no acetylations being the highest scoring option. MALDI-TOF/TOF-MS spectra of other library peptides were analysed using biotools, but the data was of insufficient quality to make any observations.

SETI biotin-ahx-H3 ₁₋₂₁ -NH ₂		SETI biotin-ahx-H3 ₁₄₋₃₄ -NH ₂		SETII NH ₂ -H2a ₁₋₂₁ -ahx-lys-biotin	
Ac PTM positions	Biotoools score	Ac PTM positions	Biotoools score	Ac PTM positions	Biotoools score
No PTM	102	No PTM	1179	No PTM	49187
K4Ac	41	K14Ac	1072	K5Ac	32802
K4AcK9Ac	91	K14AcK18Ac	1055	K5AcK9Ac	4129
K4AcK9AcK14Ac	280	K14AcK18AcK23Ac	537	K5AcK9AcK13Ac	288
K4AcK9AcK14AcK18Ac	25	K14AcK18AcK23AcK27Ac	26	K5AcK9AcK13AcK15Ac	96
K4AcK9AcK18Ac	92	K14AcK18AcK27Ac	57	K5AcK9AcK15Ac	96
K4AcK14Ac	4118	K14AcK23Ac	1071	K5AcK13Ac	289
K4AcK14AcK18Ac	30	K14AcK23AcK27Ac	42	K5AcK13AcK15Ac	96
K4AcK18Ac	1046	K14AcK27Ac	114	K5AcK15Ac	97
K9Ac	354	K18Ac	1067	K9Ac	4378
K9AcK14Ac	342	K18AcK23Ac	553	K9AcK13Ac	535
K9AcK14AcK18Ac	93	K18AcK23AcK27Ac	37	K9AcK13AcK15Ac	344
K9AcK18Ac	150	K18AcK27Ac	76	K9AcK15Ac	343
K14Ac	16480	K23Ac	1179	K13Ac	4380
K18AcK18Ac	91	K23AcK27Ac	157	K13AcK15Ac	4185
K18Ac	4193	K27Ac	217	K15Ac	16476

Table 3.12: Biotoools scoring of Ac PTM site identification in MALDI-TOF/TOF-MS spectra obtained for library peptides incubated with PCAF HAT.

Library peptides shown are SETI tagged internal control peptide biotin-ahx-H3₁₋₂₁-NH₂, SETI biotin-ahx-H3₁₄₋₃₄-NH₂ and SETII NH₂-H2a₁₋₂₁-ahx-lys-biotin. N- and C-terminal modifications are set as invariable modifications for the peptide sequences and are therefore not shown in the results.

3.6 Discussion

3.6.1 Strengths and weaknesses of the rapifleX MALDI-TOF-MS HTS

The rapifleX MALDI-TOF-MS HTS developed is a powerful method allowing the screen of ≈ 400 peptides against 3 different enzymes in the space of ≈ 1 week. The ≈ 12000 spectra can be analysed in a highly automated way using two optimised scripts. In this way a large amount of data can be obtained in a short period of time. This makes it a powerful screening method. However as with all high throughput screens, data quality must be assessed, and it is important to understand the limitations and issues that come with the materials and instruments used. Any limitations with the technique must be taken into account when analysing the data and significant data must be verified with further assays.

3.6.1.1 Histone peptide libraries

The histone peptide libraries produced by Altabioscience offer a large array of PTM combinations on H3, H4, H2a and H2b, allowing for a wide screen of histone sections and modifications in 21 amino acid peptide segments. These libraries are synthesised using automated peptide synthesis with 94 peptides synthesised in parallel in a 96 well plate format on 0.2 μmol scale, and supplied as lyophilised peptide.

Peptide synthesis in this format gives rapid access to peptide libraries but can have very varied peptide quality. The amino acid content of the peptides defines whether the peptide will synthesise well or poorly: peptides with a high proportion of amino acids which do not couple easily will have a lower yield and more side products. Couplings with nCAA are more likely to be difficult than CAA coupling, depending on the nature of

the nCAA. For example, in SETII it was observed that peptides containing the nCAA KMe3 consistently showed larger SDs than peptides with KMe1 or KMe2. The library peptides synthesised are not purified, so any impurities produced can interfere with activity assays and with data analysis.

Because enzyme activity is measured using relative intensity between parent and product peptide peaks, meaningful data can still be acquired with impure peptides. However, the impurities can affect the quality of the data. If impurities overlap with the library peptide masses, intensity measurements will be affected, giving false positives. This can be mitigated by acquiring NEC spectra for comparison. If impurities are also peptide substrates (truncated peptides, certain branched peptides) this will affect the activity of the enzyme for the substrate of interest. If the impurities are more prevalent in the spectrum than the peptide of interest either because they fly better with MALDI-TOF-MS or due to poor product synthesis, the peaks of interest will be closer to the spectrum noise and relative intensity measurement will be less accurate. Finally, presence of impurities means the final concentration of product peptide is lower than the maximum of 0.2 μmol . Peptide concentration affects MS spectrum quality and enzyme activity leading to larger errors in activity measurements.

These issues were mitigated as much as possible. The processing method (SNAP vs. centroid) and signal to noise threshold of the peak picking script were adjusted. Enzyme activity was measured as the average of relative percentage of parent and product peptide peaks with associated standard deviation, number of replicates and NEC for quality control. The first measure helps minimise peak identification errors and the

second allows for distinction between high- and low-quality results. Low-quality results must be used with caution in analysis of results.

The quality of the libraries was quite variable. Peptides in the 1-21 amino acid range of each histone were generally of high quality. Later portions of the histones (especially from residue 40 to C-terminus) generally showed poorer quality peptides. The lack of quality control peptides in SETIV was also a significant issue: most control peptides in this library failed to produce spectra. Because all analysis of the libraries is relative between peptides with different modifications, not having proper control peptides complicates analysis of the data.

The libraries present a range of chemical combinations of PTMs on histone peptides. The sites chosen for modifications were chosen to be biologically relevant. However, the combinations present were not necessarily biologically relevant. Results should be validated with untagged peptides (the library tags could lead to screen artifacts) and assessed for biological relevance.

3.6.1.2 Rapliflex MALDI-TOF-MS HTS methodology

Efforts were made during assay development to minimise errors brought about by technical issues with the screening process.

The histone library peptide screen was developed first as a low-throughput screen of a single library and then adapted as a high-throughput screen of multiple libraries and enzymes. The low-throughput screen had fewer replicates: three replicates for PCAF HAT, two replicates for PCAF HAT-BRD (due to insufficient protein stocks at time of

assay), and two NEC replicates. No technical replicates for dispensing were acquired, as manual pipetting with higher volumes (1 μ L assay mixture) ensured complete sample transfer to the MALDI plate. Low-quality spots were re-spotted and low-quality spectra were re-acquired manually, giving an overall high quality of data. High-throughput screens require elimination of manual steps. Increasing the number of replicates is used instead to give confidence in the data acquired. In the high-throughput screen, three biological replicates were carried out for each enzyme type (HAT, HAT-BRD, SIRT1 and NEC). Errors with peptide spotting, carried out using a Mosquito liquid handler, were minimised by quadruplicate spotting of each sample. Rate of spotting errors was variable between libraries with SETI showing 35% failed spots, SETII showing 13% failed spots, and SETIV showing 9% failed spots. However, this failure rate does not distinguish between lack of spectra due to library issues or spotting issues. The liquid handling instrument was calibrated before each library spotting to ensure minimal error rates. SETI was the first library to be tested with high-throughput screening, which may account for the larger percentage of failed spots as the screening method was being set up.

There are several sources of experimental errors in the assay process and in the instruments used. Pipetting errors can occur whilst running the assay (the assays were run in 96 well PCR plates with a mix of manual and automatic multichannel pipettes). Automatic MALDI acquisition was carried out with rapifleX MALDI-TOF-MS, in which the laser power was automatically adjusted, and the end point was determined based on peak resolution of top peaks. This process may prioritise resolution of high intensity impurity peaks over low intensity peptide peaks for low-quality peptides. Regular

calibration of the rapifleX and microflex instruments were carried out. However, there may be some disparity between different acquisition dates and different instruments. The masses defined for each library member are round numbers, so a mass range in peak identification of 1 Da allowed for slight variation between replicates, instruments and acquisition dates.

3.6.1.3 Data compilation methodology

Errors in the analysis process were minimised as much as possible by optimising the parameters of the peak picking and data retrieval scripts. The data from the four libraries (low- and high-throughput screens) were processed using the same parameters and scripts. The data retrieval script generated an Excel file with mass and intensity data for each of the peaks of interest. The script then calculated the relative percentage of parent to product peaks for each peptide based on intensity of the peaks. An average of the percentage for each peptide state between all of the replicates for a peptide-enzyme combination was calculated with associated standard deviation (SD) and number of replicates (N). This was either carried out directly in the data retrieval script if the data acquisition was in the correct 1536 spot MALDI plate format, or using Excel. All of the data for average percentages, SD and N was compiled in tables for each library with comparative columns for each enzyme and NEC corresponding to each library member. The library members were then identified by matching the well numbers to peptide sequences. (Figure 3.5, Appendix D)

Errors linked to impurities, automated spotting, automated spectrum acquisition, peak picking and peak identification will be reflected in the calculated SD. Variation between

biological replicates is also reflected in the SD. In high-throughput methods it can be hard to identify the cause for large variation in measured activity. Some trends can however be identified. Large SD were recorded for all KMe3 containing peptides in SETII, pointing to an issue with couplings of this amino acid. Large SD for one PCAF construct compared with small SD for the other construct for a same peptide likely indicates variation in the enzymatic assay rather than issues with the peptide. If the validity of certain results was in doubt, the MS spectra for the biological and technical replicates could be manually examined. Library quality, automated spotting and experimental errors were the largest sources of errors in the screen with automatic data acquisition by the rapifleX MALDI and optimised peak picking and data retrieval being very reliable.

In this study, the data for PCAF HAT, HAT-BRD and NEC control was presented as an end-point dataset of relative peak intensities of substrate and products. In enzyme assays, the NEC values are often subtracted from the enzyme assay values to account for background levels in measurements. In this case, the observed conversion was analysed rather than a calculated product turnover, so minimal data manipulation was preferable. Quality checks for the library were based on the SD measurements between replicates for a peptide/enzyme combination. The difference in SD measurements between the enzyme and the NEC would not be accounted for in baseline subtraction calculations. In addition, there was variability in the impurities identified between the enzyme containing samples and the NEC, with impurities (likely corresponding to incomplete library peptide sequences) being acetylated by PCAF (Section 3.2.6, Figure 3.6). The data was therefore analysed without further manipulation, to minimise errors

linked to data processing, and to retain all relevant information on errors linked to the screening method.

3.6.1.4 Accounting for errors which arise in the rapifleX MALDI-TOF-MS HTS

The limitations of MALDI as a technique must be taken into account when analysing the results. In MALDI, peptides travel in the TOF drift tube based on their mass to charge ratio (m/z). Peptides which are more easily ionised and desorbed will appear more abundant in the MALDI-TOF-MS spectrum. Amino acid PTMs can affect desorption and ionisation of the peptide, and the propensity of peptides to fragment (*e.g.* H3₁₋₂₁K4Me1, H3₁₋₂₁K4Me2 and H3₁₋₂₁K4Me3 may not be equally fit for MALDI-TOF-MS data acquisition). In addition, because the absolute intensities of peaks is spectrum dependent, comparison of peak intensities between spectra is meaningless. For this reason, the relative percentage of parent to product peptide peak intensities was compared between spectra rather than the intensities themselves. It was observed for H3₁₋₂₁ that acetylation at K9, K14 or K9 and K14 did not affect the MALDI-TOF-MS measured abundance of the H3₁₋₂₁ peptides (Section 2.2.2). Based on this result, relative quantification of enzymatic acetylation of peptides with PTMs within a single sample should be an accurate measure of activity, as this should not affect peptide desorption and ionisation.

There are multiple sources of errors in the screening process, linked to the materials used, the physical process of the assay technique and processing, instrument limitations and data processing methods. These sources of error were identified and mitigated where possible. For low-throughput screening of the SETIII library with limited number

of assay replicates, manual checks at each step in the screening process ensured high data quality. For high-throughput screening of SETI, SETII and SETIV libraries multiple assay and technical replicates were instead acquired, and strategies for identifying low quality data and limiting the impact of low quality data on analysis of results were put in place (Sections 3.6.1.1-3.6.1.3). This included using SD and N values to identify low quality data, and analysing data based on comparisons with appropriate controls and by rank order of peptides.

3.6.1.5 Accounting for issues arising in rapifleX MALDI-TOF/TOF-MS

The rapifleX MALDI-TOF-MS has MS/MS capacity. The quality of MS/MS spectra was dependent on the intensity of the peptide peak at the mass chosen for fragmentation, and therefore dependent on peptide concentration. The screen assays were run at 20 μ M peptide concentration to allow for MS/MS. This assumed that the library peptide has been synthesised with 100% yield, and that the peptide in the assay was present in a single state (parent, 1Ac or 2Ac). The library peptides were not purified or quantified. Many peptide spectra showed significant amounts of side products, indicating low yield. In addition, not all library peptides were fully converted to product; the peptides observed in the MS spectra were distributed between parent state, 1Ac state and 2Ac state for libraries screened with PCAF. The final concentration for a specific state of each library peptide in each assay sample was therefore significantly less than the theoretical maximum concentration. MS/MS of library peptides with rapifleX MALDI-TOF/TOF-MS was attempted but the majority of samples tested for MS/MS showed too little fragmentation for clear analysis.

To successfully integrate automatic MS/MS into the screen, higher material concentrations would have to be used, or the peptides would have to be purified and spotted at higher concentrations. The process for automated analysis of MS/MS spectra using alternative software for analysis also needs to be fully developed and tested.

3.6.2 RapifleX MALDI-TOF/TOF-MS HTS results for screening of PCAF HAT and HAT-BRD

The rapifleX screen of histone peptide libraries with PTMs against PCAF HAT and HAT-BRD has improved the understanding of PCAF site selectivity for histone peptides, and shown some interesting examples of crosstalk between epigenetic PTMs and between epigenetic protein domains (Figure 3.10). The significance of the results observed is discussed below.

3.6.2.1 Differential activity between histones H3, H4, H2a and H2b

The histone peptide libraries screened contained peptides corresponding to sections of histones H3, H4, H2a and H2b. The screen allowed for an overview of peptide conversions which can be analysed with peptides sorted by histone type and section for each of the libraries screened with HAT and HAT-BRD (Figure 3.7a-d).

The results confirmed H3 and H4 as the principal substrate sites for PCAF acetylation activity. H3₁₋₂₁ peptides showed the highest levels of conversion with single (1Ac) and double (2Ac) acetylation observed. The *N*-terminus of H3 showed the highest levels of conversion, with little acetylation observed in peptides corresponding to H3 residues after K18. H4₁₋₂₁ showed the second highest levels of enzymatic acetylation with only

1Ac observed. Peptides corresponding to H4 residues from G9 and onwards showed low levels of acetylation. Acetylation of H2a₁₋₂₁ and H2b₁₋₂₁ was also observed, with conversion levels lower than those observed for H3₁₋₂₁ and H4₁₋₂₁.

In the screen, high concentrations of PCAF were used (2 μ M HAT or HAT-BRD) over a 1h period of incubation at 37°C. The H2a and H2b acetylation sites are less favourable than the H3 and H4 sites as indicated by the lower peptide conversions for H2a₁₋₂₁ and H2b₁₋₂₁. There is some evidence that GCN5 is able to acetylate H2a and H2b. GCN5 in the SAGA complex was shown to activate acetylation of H2a.Z (an isoform of H2a) in human cells, at an unknown site.^{114,115} GCN5 was identified as acetylating H2bK11 and H2bK16 within the yeast SAGA complex.^{116,117} No equivalent studies have shown PCAF HAT acetylation of H2a or H2b. Low levels of acetylation for H2a₁₋₂₁ and H2b₁₋₂₁ in the rapifleX screen (compared to H3₁₋₂₁ and H4₁₋₂₁ levels of acetylation) indicates that these sites are unlikely to be primary substrates for PCAF HAT in a cellular context. As the screen is a simplified model, it may be that use of full-length histone proteins or nucleosomes as substrates, or participation of PCAF in a protein complex affects the ability of PCAF HAT to acetylate H2a and H2b. Further tests are therefore necessary to determine whether the H2a and H2b acetylation events observed are biologically relevant. Aside from a possible biological significance, information on PCAF site selectivity can help understand sequence selectivity and peptide interactions that govern the activity of PCAF HAT, which may justify further exploration of H2a and H2b as PCAF substrates.

3.6.2.2 Identification of acetylation sites on H3, H4, H2a and H2b

Comparing conversion levels for peptides with combinatorial PTMs can help identify the sites of PCAF acetylation. This was the case for H3, H4 and H2a, but the acetylation site for H2b remained unclear (Figure 3.10a).

For H3₁₋₂₁control peptide two acetylation events were observed, with full conversion to 1Ac product and partial conversion to 2Ac product. PTMs at H3₁₋₂₁K9 led to loss of the 2Ac product whilst the 1Ac product was maintained. PTMs at H3₁₋₂₁K14 (Me1 excluded) led to loss of the 2Ac product and reduction of the 1Ac product to match the level of H3₁₋₂₁control 2Ac product. PTMs at H3₁₋₂₁K4 did not appear to affect the 1Ac or 2Ac product levels (with errors taken into account). These results indicate full enzymatic acetylation by PCAF of H3₁₋₂₁ at K14 and partial acetylation of at K9 (Section 3.3.2.1, Table 3.3). This was supported by MS/MS results obtained on unmodified H3₁₋₂₁ peptides in time course studies. These showed full conversion to a 1Ac product composed in majority of H3₁₋₂₁K14Ac followed by conversion to a 2Ac product composed in majority of H3₁₋₂₁K9AcK14Ac (Section 2.2.4.1). PTMs at H3K9 and H3K14 of H3₁₋₂₁ peptides appear to be independent of each other with PTMs at either of these positions not affecting PCAF acetylation of the other position. This points to PCAF acetylating H3₁₋₂₁K9 and H3₁₋₂₁K14 independently with a greater efficiency for H3₁₋₂₁K14 acetylation as indicated by full 1Ac conversion versus partial 2Ac conversion.

For H4₁₋₂₁control peptide a single acetylation event was observed, with partial conversion to a 1Ac product. All combinations of PTM which included modifications at

H4₁₋₂₁K8 led to loss of acetylation (Section 3.3.2.2, Table 3.4). These results indicated that enzymatic acetylation by PCAF of H4₁₋₂₁ occurs solely at H4K8.

For H2a₁₋₂₁control peptide, a single acetylation event was observed, with partial conversion to a 1Ac product. All combinations of PTM on H2a₁₋₂₁ peptides which included modifications at H2aK9 led to loss of acetylation (Section 3.3.2.3, Table 3.5). For the SETII library, conversion levels were low with $\leq 16\%$ 1Ac observed for H2a₁₋₂₁control. PCAF HAT showed increased conversion levels for SETIV compared with the other libraries, giving up to 69% 1Ac conversion for H2a₁₋₂₁ peptides, but showing large SD of up to 12%. Due to increased enzyme activity, the inhibitory effect of H2a₁₋₂₁K9 modifications was much more marked for the SETIV peptides compared to SETII peptides, with $< 5\%$ 1Ac observed for peptides with K9 modifications compared with $> 30\%$ 1Ac for peptides without K9 modifications. These results indicate that enzymatic acetylation by PCAF of H2a₁₋₂₁ occurs solely at H2a₁₋₂₁K9. Although there is a disparity in conversion levels for PCAF HAT between the libraries, HAT and HAT-BRD for SETII and SETIV all showed the same trend for H2a₁₋₂₁K9 PTMs, giving confidence in the results.

For H2b₁₋₂₁control peptide a single acetylation event was observed, with partial conversion to a 1Ac product. The levels of acetylation for H2b₁₋₂₁ peptides was very moderate with $\leq 12\%$ 1Ac product seen even for PCAF HAT with SETIV peptides, which generally show higher conversion than for other libraries. PTMs at K5, K11, K12, K15 and K16 were present in the libraries. Individually none of these sites showed marked inhibition of acetylation. Combinations of modification showed inhibition of acetylation

but with no discernible pattern. Literature findings would suggest acetylation sites to be K11 and K16,^{116,117} however, the site of PCAF acetylation on the H2b₁₋₂₁ peptide could not be determined with the results obtained in the rapifleX MALDI-TOF-MS screen. Repeating the assay with longer incubation or higher enzyme concentrations might yield higher peptide conversions, giving more information on inhibition. Although the site may not have high levels of acetylation in a cellular context, knowing the acetylation sites on H2b can help understand PCAF selectivity.

The acetylation sites of PCAF on H3 and H4 are well documented. Results of the screen were in agreement with reported selectivity, with H3K14, H3K9 and H4K8 determined as acetylation sites. The site of PCAF acetylation of H2a has not been reported to date, likely due to the lower conversion levels seen with H2a compared with H3 and H4. Identification of H2a₁₋₂₁K9 as an acetylation site of PCAF HAT and HAT-BRD therefore need to be thoroughly validated and studies are needed to confirm whether this is a biologically relevant site of acetylation by PCAF. These results show the usefulness of the rapifleX MALDI-TOF-MS screen in confirming and identifying new sites of enzyme activity on histone peptides.

3.6.2.3 Phosphorylation combinatorial PTMs: site specific effects on PCAF activity

The phosphorylation and acetylation biological signalling systems are parallel systems each governed by sets of writer and eraser enzymes. Where acetylation is modulated by HATs and HDACs, phosphorylation is modulated by kinases and phosphatases. What is of interest here are the close links and crosstalk between these two signalling systems.⁷⁸

Crosstalk between phosphorylation and acetylation PTMs are evident in activity of PCAF towards non-histone proteins. Phosphorylation of the tumour suppressor protein p53 increases p53 acetylation by both p300 and PCAF proteins, which in turn enhances sequence specific DNA binding by p53.^{80,83} DNA dependent protein kinase (DNA-PK) phosphorylates PCAF in response to DNA damage, promoting PCAF acetylation of the replication protein A (RPA) complex, promoting DNA repair.³⁶⁷ Effects of Ac-Phos crosstalk are also seen on histones tails, with linked modifications either on a single histone tail (cis effects), or on separate histone tails (trans effects).¹⁰⁵ On H3, S10Phos and T11Phos have been shown to affect GCN5 mediated acetylation of K9 and K14.^{40–43,45,50,51} H2aS1 phosphorylation was inhibited by PCAF acetylation of H3.⁴⁴ H3S10Phos was shown to promote H4K16 acetylation by the HAT MOF.³⁶⁸ Certain combinations of phosphorylation and acetylation PTMs are present in the histone peptide libraries which were screened during the rapifleX screen with PCAF HAT and HAT-BRD, giving an insight into PCAF regulation by phosphorylation PTMs (Figure 3.10b).

- H3S10Phos and H3T11Phos

H3S10Phos and H3T11Phos are biologically relevant PTMs. H3S10Phos has been linked to mitosis, chromosomal condensation and de-condensation, and transcriptional activation.⁵⁰ H3T11Phos has been linked to identification of androgen receptor dependent genes, mitosis, gene activation, transcription levels and acetylation levels.⁴⁵ Phosphorylation of H3S10 and H3T11 has been reported to increase levels of H3K9 and H3K14 acetylation by GCN5.^{40,41,51} H3S10 was shown to have a smaller activating effect on PCAF acetylation activity compared to GCN5, indicating a possible divergence in

selectivity between these enzymes.⁴¹ However, contradictory evidence was published, showing no activation of GCN5 acetylation activity linked to H3S10 or H3T11.^{42,43} Using time resolved NMR experiments, Liokatis et. al showed that in a mixture containing ¹⁵N labelled H3₁₋₃₃S10Phos and H3₁₋₃₃T11Phos, H3₁₋₃₃S10Phos was first acetylated at K14 at the same rate as unmodified H3₁₋₃₃ peptide, followed by H3₁₋₃₃T11Phos acetylation at K14 at a slower rate. No K9 acetylation was observed for H3₁₋₃₃S10Phos or H3₁₋₃₃T11Phos.⁴⁵ This indicated that for GCN5, phosphorylation at H3S10 inhibited H3K9 acetylation but had no effect on K14 acetylation, whilst phosphorylation at H3T11 inhibited H3K9 acetylation and to a lesser extent diminished the rate of K14 acetylation.

In the screen of H3₁₋₂₁ peptides against PCAF HAT and HAT-BRD, control peptides showed a 1Ac product corresponding to acetylation at H3K14 and a 2Ac product corresponding to acetylation at H3K9 (Section 3.3.4.2, Table 3.8). H3₁₋₂₁S10Phos and H3₁₋₂₁T11Phos peptides both showed complete loss of the 2Ac product compared to controls. This indicated that for PCAF against H3₁₋₂₁ peptides, H3S10Phos and H3T11Phos both inhibited H3K9 acetylation. H3₁₋₂₁ peptides with H3S10Phos PTMs alone or in combination with K9 PTMs showed ≥99% 1Ac, indicating that this PTM did not affect H3K14 acetylation by PCAF. This was in agreement with time-resolved NMR study results with GCN5 and H3₁₋₃₃.

H3₁₋₂₁ peptides bearing H3T11Phos PTM alone did not show significant inhibition of H3K14 acetylation by PCAF HAT (1Ac ≥ 94%). Because the screen performed was an end point assay, it could not distinguish between rates of acetylation for peptides with different PTMs where the peptides had been fully acetylated by the end of the assay.

This may account for discrepancy in observations between the rapifleX screen of PCAF activity and the time resolved NMR screen of GCN5 for T11Phos inhibition of H3K14 acetylation. Although acetylation for H3₁₋₂₁T11Phos by PCAF HAT may be slower than that of H3₁₋₂₁S10Phos, both were observed to be fully acetylated at K14 (1Ac $\geq$ 94%) by the end of the assay. Variation between rapifleX and NMR results could also be due to the difference in peptides used (H3₁₋₂₁ for rapifleX H3₁₋₃₃ for NMR) as it has been shown that the *N*-terminal region of H3 peptides is important for peptide binding to tGCN5, and by extension PCAF.^{41,50,105} For PCAF HAT-BRD acetylation of H3₁₋₂₁ peptides, H3T11Phos PTM on its own showed moderate inhibition of acetylation (1Ac $\leq$ 81% for H3₁₋₂₁T11Phos, whilst H3₁₋₂₁S10Phos gave 1Ac $\geq$ 99%).

Increased inhibition for both PCAF HAT and HAT-BRD was observed when T11Phos was combined with S10Phos and K9 PTMs on H3₁₋₂₁ peptides, with inhibition of HAT-BRD being more pronounced for all PTM combinations. Although H3S10Phos is often found in conjunction with H3K9Ac or H3K9Me3 in cellular models,^{369,370} it is not clear whether these modifications are found in conjunction with H3T11Phos. Combinations of PTM including H3K9Ac consistently showed greater inhibition of PCAF activity for H3₁₋₂₁ peptides than those with H3K9Me3. Crystal structures of γ GCN5 with H3₅₋₂₃ and H3₅₋₂₃S10Phos with K14 bound in the active site showed that S10Phos caused a large shift in H3₅₋₂₃ binding to γ GCN5, and that both S10 and K9 made numerous interactions with γ GCN5 which were altered upon phosphorylation of S10.¹⁰⁵ It is therefore likely that a change from K9Ac to K9Me3 affects an important interaction site between H3K9 and PCAF, leading to a difference in enzyme activity.

Although the cooperative inhibition of PCAF activity towards H3₁₋₂₁ peptides by H3T11Phos, H3S10Phos and H3K9Ac/Me3 may not be biologically relevant if these PTM are not found clustered together in cellular models, it shows that alterations to this section of the peptide have a profound effect on binding interactions between PCAF and H3 peptides.

- H3S28Phos

Phosphorylation occurs at specific time points of the mitotic cycle on H3 at S10, S28, T3 and T11. H3S10 and H3S28 share a sequence motif (ARKS), are both phosphorylated by the aurora-B and MSK kinases, and have very similar roles in mitosis and in transcription regulation.³⁷¹ Clear links are established between the biology of H3S10Phos with H3K9Me3, H3K9Ac and H3K14Ac as well as between H3S28Phos and H3K27Me3 and H3K27Ac in the mitotic cycle of human and drosophilae cells.^{52,354,371-373} Although K9Ac/S10Phos/K14Ac and K27Ac/S28Phos modifications on H3 are biologically linked, and are recognised as pairs by 14-3-3 reader proteins, the causality leading to the modifications being present simultaneously are not entirely clear.³⁷³ The effects of H3S10 phosphorylation on PCAF activity have been discussed above. It was shown that H3S28Phos led to a switch from H3K27 methylation to acetylation in human cells, but the HAT responsible for acetylation was not identified.⁵² Further understanding of the links between H3S28Phos and acetylation are therefore necessary.

The results of the rapifleX screen with PCAF showed 1Ac and 2Ac products for SET1 H3₁₄₋₃₄, however the sites of acetylation could not be identified due to low fragmentation in MS/MS. H3₁₄₋₃₄S28Phos led to a decrease in H3₁₄₋₃₄ acetylation

compared with a H3₁₄₋₃₄ control. Pairing S28Phos with K27Me3 or K27Ac PTMs on H3₁₄₋₃₄ peptides did not appear to modify the activity of PCAF for the peptides. This data could not be cross-checked between libraries, as the S28Phos PTM was present in SETIII, but the H3₂₀₋₄₀ peptide used showed only 2% 1Ac product for the control. These results indicated that the two acetylation events observed with SETI H3₁₄₋₃₄ were likely occurring at the K14 and K18 sites, as K23 and K27 acetylation would likely also be observed with the SETIII H3₂₀₋₄₀ peptide. This also implied that the acetylation site affected by the S28Phos PTM was not located at K27. This would in turn imply that PCAF is not involved in the S28Phos activation of K27 reported in cellular data.⁵² H3₁₄₋₃₄S28Phos showed more pronounced inhibition with PCAF HAT-BRD (18-26% 1Ac compared with 47% 1Ac for control) than with PCAF HAT (37-45% 1Ac compared with 49% 1Ac for control) reflecting results observed for PCAF acetylation of H3₁₋₂₁T11Phos. As the site of the two acetylation events on H3₁₄₋₃₄ are unknown, these observations cannot help understand sequence selectivity and peptide binding modes of PCAF HAT versus HAT-BRD. However, it does point to a trend where phosphorylation modifications affect the dual domain to a higher extent than the single domain.

- H4S1Phos, H2aS1Phos and H3T3Phos

H4S1 and H2aS1 phosphorylation has been observed in certain phases of mitosis, with the occurrence of H4S1Phos and H2aS1Phos correlating with H3S10Phos and H3S28Phos marks.³⁷⁴ H3T3Phos is also observed in mitosis but has a different spatio-temporal distribution than H3S10Phos.³⁷⁵ No links have been made between H4S1Phos/H2aS1Phos/H3T3Phos and PCAF acetylation.

In the rapifleX MALDI screen, H4S1Phos, H2aS1Phos and H3T3Phos all showed inhibitory effects against PCAF acetylation activity. H3T3Phos gave inhibition of H3K9 acetylation of H3₁₋₂₁ peptides with PCAF HAT and HAT-BRD, with a drop of activity of ~45% 2Ac for HAT and ~20% 2Ac with HAT-BRD. The effect of H3R2Me2a and H3K4Me3 modifications which are paired with H3T3Phos was much slighter particularly with HAT-BRD, indicating that only H3T3Phos was largely responsible for inhibition. H3T3Phos did not appear to affect H3K14 acetylation with turnover to 1Ac product of >95% (with conversion to 2Ac product taken into account). H4S1Phos inhibited H3K8 acetylation of H4₁₋₂₁ peptides with PCAF HAT and HAT-BRD. Effects of H4R3Me, H4R3Me2a and H4K5Ac were negligible by comparison. SETII H2a₁₋₂₁ acetylation by PCAF HAT and HAT-BRD gave a low turnover (~15% 1Ac). However, for H2a₁₋₂₁H2aS1Phos, ≤1% 1Ac was observed with HAT and HAT-BRD giving confidence in the inhibitory effect of H2aS1Phos. H2aK5Ac which was paired with H2aS1Phos, may also have some inhibitory effect, without complete loss of acetylation activity. However, the H2aS1Phos PTM alone is sufficient for complete loss of PCAF acetylation activity. Loss of acetylation activity is assumed to be at the K9 lysine on H2a₁₋₂₁, which was deduced to be the logical site of PCAF acetylation based on the effects of combinatorial Ac PTM on PCAF activity.

The first 5 residues of H4 and H2a are the same (S₁GRGK) with the site for lysine acetylation by PCAF also having a similar surrounding sequence (GGK₈G for H4 and GGK₉A for H2a), which likely explains that they share a biologically relevant S1Phos modification and have similar effects on PCAF acetylation activity. The *N*-terminal sequence of H3 is different from that of H4 and H2a. The amino acids surrounding the lysine residue targeted for catalytic activity of PCAF for H3K14 and H2aK8 are the same

(GGK₁₄A/GGK₈A) unlike that of H3K9 (ARK₉S). However, the spacing between H3T3/H3K9 (5 residues) is closer to that of H4S1/H4K8 (6 residues) and H2aS1/H2aK9 (7 residues). This might explain why phosphorylation of H3₁₋₂₁T3Phos inhibits PCAF activity at H3₁₋₂₁K9 rather than H3₁₋₂₁K14. There is little sequence similarity between the acetylation sites of PCAF on histone and non-histone substrates in general, which makes it unlikely that sequence motifs affect PCAF selectivity.^{50,105}

It is unclear whether these observations have any biological relevance. The phosphorylation marks involved all have a role in mitosis. Levels of H3K9Ac and H4K8Ac have been linked to H3T3Phos levels in the control of meiotic division in porcine oocyte maturation.³⁷⁶ H3K9Ac and H4K8Ac were also linked to increased tumour proliferation for oesophageal cancer, but with no link to phosphorylation.³⁷⁷ No links to mitosis were found for H2aK9Ac. PCAF has been shown to catalyse the acetylation of H3 and H4 at mitotic checkpoint protein (MAD1/2) promoter regions during selected mitotic phases, however no specific acetylation sites were identified.³⁷⁸ Phosphorylation and acetylation are shown to both be necessary and co-dependent in many biological processes;⁷⁸ however the nature of the interactions seen here has not yet been studied. If the PTMs discussed do not occur at the same time/location where PCAF is present, then the observations may not have any biological relevance.

3.6.2.4 Methylation combinatorial PTMs: site specific effects on PCAF activity

Methylation is a widely spread epigenetic mark, which has many known links with acetylation. Many histone lysine sites are involved in signalling with a switch from methylation to acetylation giving opposite effects.^{224,225} Most of the methylation marks

in the histone peptide libraries did not affect PCAF acetylation activity and are therefore not mentioned here but can be found in Appendix D.

- Reported biological roles of H3K14 methylation

H3K14 methylation is found in very low abundance compared with H3K14 acetylation and was therefore overlooked as a H3 PTM until it was identified as part of a bacterial mechanism for increasing efficiency of infection.^{49,379}

H3K14 methylation in mammalian cells was first identified by Rolando *et al.* in 2013. *Legionella pneumophila* produces a methyltransferase (RomA) which was shown to specifically target H3K14. Using a H3K14Me2 antibody, H3K14 methylation was confirmed *in vitro* on peptides and nucleosomes, and in human cell lines. MS/MS was used to identify trimethylation (+42 Da) on nucleosomes and peptides treated with RomA. The H3K14Me2 antibody used bound to both H3₁₋₂₁K14Me2 and H3₁₋₂₁K14Me3 but not H3₁₋₂₁K14Me1. This meant that there was no way to ascertain if all H3K14 sites were fully trimethylated. It was shown that increase in H3K14 methylation levels correlated with decreased H3K14Ac levels. H3K14 methylation occurred genome-wide with hot spots of H3K14 methylation mapping onto locations where H3K14Ac is usually found. This switch from acetylation to methylation led to reduced transcriptional activity, particularly for genes which allowed *L. pneumophila* to evade the immune system and replicate.⁴⁹

Endogenous H3K14 methylation has since been identified in eukaryotes. Various studies identified H3K14 mono-, di- and trimethylation at very low levels whilst profiling all

epigenetic marks in mouse embryonic stem cells,³⁸⁰ HeLa cells,³⁸¹ yeast,³⁸² and epithelial cells.³⁸³ A study aimed specifically at determining H3K14Me3 location and quantification in human cell lines confirmed the presence of endogenous H3K14Me3.³⁷⁹ In HEK293 cells H3K14Me3 was detected with an abundance of ~0.4% compared with an average of 20-30% H3K14Ac commonly found in human cell lines.^{379,383} H3K14Me3 was identified in HEK293, hESC and mouse B16 cells. In HEK293 and B16 cells, H3K14Me3 was identified at transcription start sites (TSS), indicating a role in active transcription for the modification. This study also identified KDM4A/C as demethylases active against H3K14Me3 *in vitro* and in cells. The antibody used in this study showed clear specificity for H3₁₋₂₁K14Me3 over H3₁₋₂₁K14Me1/2.³⁷⁹

A study investigated the role of methylation in regulation of cell proliferation for colon cancer. H3K14Me1/2/3 antibodies were used, detecting H3K14Me1/2/3 at the TSS for the tumour suppressor candidate 3 (TUSC3). H3K14Me1/2/3 were all substrates for binding of ubiquitin-like PHD and RING finger domain containing protein 1 (UHRF1) which protected H3K14 from acetylation by KAT7 (MYST2). These effects combined to downregulate TUSC3 leading to an increase in cell proliferation of colon cancer cells.³⁸⁴ These observations showed that H3K14Me1/2/3 are all biologically relevant endogenous PTMs.

- H3K14 methylation in the rapifleX MALDI-TOF-MS HTS

In the rapifleX screen, H3₁₋₂₁K14Me1 was identified to have a different effect on PCAF activity compared with H3₁₋₂₁K14Ac, H3₁₋₂₁K14Me2 and H3₁₋₂₁K14Me3 peptides. SETI library H3₁₋₂₁ internal control peptide showed 1Ac and 2Ac products corresponding to

PCAF acetylation first at K14 then K9. H3₁₋₂₁ peptides with Ac, Me2 or Me3 modifications at K14 showed only 1Ac product, with the enzymatic acetylation most likely occurring at H3K9, as the H3K14 position is no longer available. In contrast, H3₁₋₂₁K14Me1 peptide showed an increase in acetylation with $\approx 55\%$ 1Ac and $\approx 42\%$ 2Ac observed with both PCAF HAT and HAT-BRD (HAT-BRD showed large SD). This effect was also observed in the SETIII library with H3₈₋₂₈ peptides showing only 1Ac product for H3₈₋₂₈K14Ac/Me2/Me3 with PCAF HAT/HAT-BRD, but H3₈₋₂₈K14Me1 showing $\approx 70\%$ 1Ac and $\approx 25\%$ 2Ac with PCAF HAT (large SD) and HAT-BRD. Observing this activation pattern in two different libraries on different peptide lengths for both constructs of PCAF gave confidence in the observations (Section 3.3.4.4, Table 3.11, Figure 3.10c).

It is unclear why such a large difference in activity was seen between peptides with different methylation states in the HTS, but this is likely due to changes in binding modes of the peptides with PCAF. In the studies on H4K14 methylation, Taniue *et al.* used three distinct and selective antibodies, showing the presence of all three methylation states at the TSS for TUSC3.³⁸⁴ No difference in biological activity was noted between the three methylation states. The other works mentioned either focussed on H3K14Me3³⁷⁹ or H3K14Me2/3 (due to a non-selective antibody)⁴⁹, without discussing whether different methylation states might affect biological activity. It was shown that KDM4A/C demethylated H3K14Me3 both on peptides and in cells, with peptide MALDI-TOF-MS studies showing demethylation of H3₁₋₂₁K14Me3 to give H3₁₋₂₁K14Me2 but not H3₁₋₂₁K14Me1.³⁷⁹ This would indicate that distinct methylation states are found in cells and are therefore likely to have different biological roles. Further characterisation of the

interactions between H3₁₋₂₁K14Me1 and PCAF are needed to ascertain the cellular location of H3₁₋₂₁K14Me1 PTMs, which might help better inform on biological relevance.

The effect of H3K14Me1 crosstalk was shown to be a site-specific effect. The histone peptide libraries have mono-, di- and tri-methylation at a number of different sites on all four histone types including but not limited to H3K4/K9/23, H4K12/20 and H2aK23. In the SETII library an increase in 1Ac product was observed with H4₁₋₂₁K12me1 with both HAT and HAT-BRD. However, this was not reproduced in the SETIII library with the same peptides and needs further investigation. Other methylated peptides throughout the libraries either showed inhibition of PCAF activity (*e.g.* H3₁₋₂₁K9Me1/2/3 block the K9 acetylation site of PCAF leading to decreased acetylation activity compared with the H3₁₋₂₁ internal control peptide), or had no visible effect on PCAF activity.

- H3K4Me3

H3K4Me3 has been linked to increased acetylation activity of GCN5 in the SAGA complex in yeast and mammalian cells. It was reported that this activity was linked to H3K4Me3 binding by the Sgf29 tudor domain, also part of the SAGA complex. Outside of the SAGA complex, GCN5 activity was not altered by H3K4Me3.³⁸⁵ In the rapifleX results, H3K4Me3 did not affect PCAF acetylation activity. However, differences in activity of the PCAF and GCN5 truncated recombinant enzymes compared with their full-length protein counterparts within a complex might differ. This means that a negative result in the HTS does not prove lack of interaction with the protein in a biological setting. However, it indicates that like GCN5, PCAF activity might be indirectly affected by H3K4Me3 rather than PCAF having any direct binding interactions with this PTM.

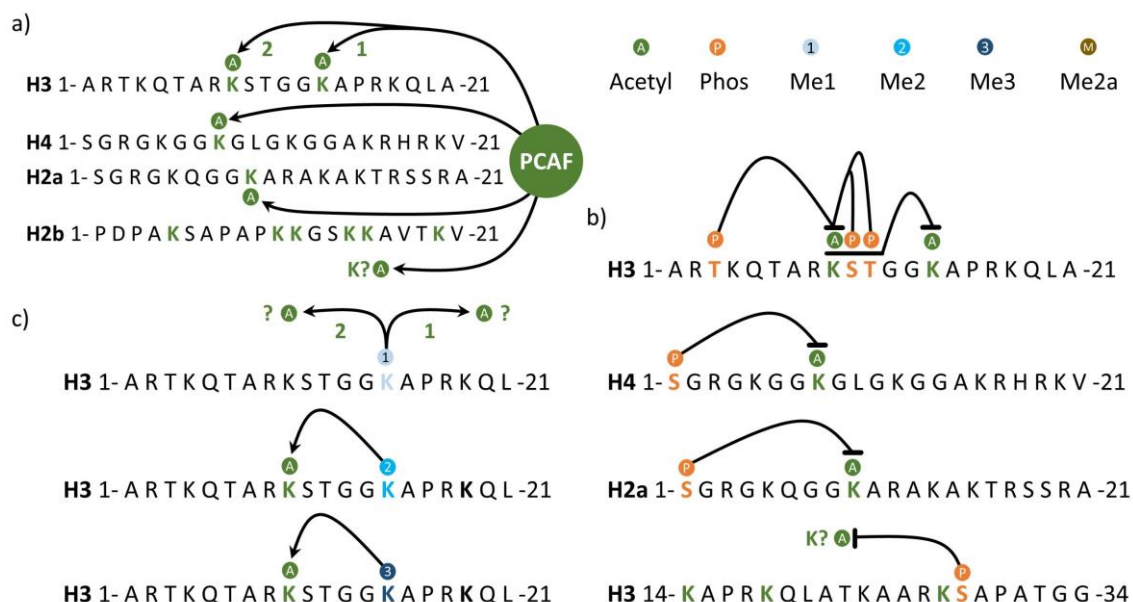


Figure 3.10: Summary of observed PCAF enzymatic activity against histone peptides with PTMs.

a) The site selectivity of PCAF for H3₁₋₂₁, H4₁₋₂₁ and H2a₁₋₂₁ peptides was deduced. H3₁₋₂₁ was acetylated first at K14 followed by K9, H4₁₋₂₁ was acetylated at K8 and H2a₁₋₂₁ was acetylated at K9. No site selectivity for PCAF acetylation of H2b₁₋₂₁ could be determined. **b)** Phosphorylation PTMs were shown to have inhibitory activity against PCAF acetylation. For H3₁₋₂₁ peptides, T3Phos, S10Phos and T11Phos inhibited PCAF acetylation of K9. In addition, T11Phos had inhibitory activity against PCAF acetylation of K14 which was increased when associated with S10Phos and K9Ac (or Me3, not shown). S28Phos inhibited PCAF acetylation of H3₁₄₋₃₄ peptides, but the site of acetylation was not identified. For H4₁₋₂₁ peptides, S1Phos PTM had inhibitory activity against PCAF acetylation at H4K8. For H2a₁₋₂₁ peptides, S1Phos PTM had inhibitory activity against PCAF acetylation at H2aK9. **c)** Mono-methylation of H3K14 activates PCAF acetylation towards H3₁₋₂₁K14Me1 peptides. Di- and tri-methylation do not show the same effect.

3.6.2.5 Differential activity between PCAF HAT and HAT-BRD

HAT and HAT-BRD PCAF constructs were screened against each of the libraries, allowing for comparisons between activity profiles of the two constructs. The single HAT construct is commonly studied; however, the HAT-BRD construct has not been previously reported. Any differences in activity involving the HAT-BRD construct may better reflect the activity of full-length PCAF. It was reported that acetylation activity of GCN5 in the ADA subcomplex was altered by a Y413A (homologous to Y809 in PCAF) mutation which disturbs the KAc binding ability of the GCN5 bromodomain without

affecting structure.¹²² Effects of the bromodomain on PCAF HAT activity have not been reported, but may mirror GCN5 bromodomain activity.

The SETIV library consistently showed increased conversion levels for PCAF HAT compared to conversion levels seen in other libraries. (Section 3.3.1, Table 3.2). Differences in PCAF HAT and HAT-BRD activities against the SETIV library were therefore discounted for this section of analysis.

PCAF HAT and HAT-BRD both showed full 1Ac conversion of H3₁₋₂₁ control peptide. In contrast, PCAF HAT consistently gave higher 2Ac product levels than HAT-BRD in SETI and SETIII (Table 3.3, Table 3.7, Table 3.9, Table 3.11). PCAF HAT and HAT-BRD consistently showed similar levels of 1Ac product of H4₁₋₂₁ peptides (SETII and SETIII, Section 3.3.2.2, Table 3.4) and H2a₁₋₂₁ peptides (SETII, Section 3.3.2.3, Table 3.5). This indicates that PCAF HAT and HAT-BRD have the same levels of activity for H3₁₋₂₁K14, H4₁₋₂₁K8 and H2a₁₋₂₁K9. However, for H3₁₋₂₁K9, PCAF HAT appears to be more active than PCAF HAT-BRD. The difference in structure between the HAT and HAT-BRD constructs may affect certain peptide-protein interactions, leading to a difference in affinities for H3₁₋₂₁K9 compared with H3₁₋₂₁K14. It was shown that the bromodomain of PCAF had a much lower binding affinity for H3₂₋₁₆K9Ac (K_d of 1051 ± 90 µM) than for H3₇₋₂₁K14Ac (K_d of 128 ± 12 µM).²⁸⁷ This might affect the difference in activities observed with the HAT-BRD construct compared with the single HAT construct.

H3₁₋₂₁T11Phos and H3₁₄₋₃₄S28Phos modifications were observed to inhibit PCAF HAT-BRD acetylation activity to a greater extent than PCAF HAT (Sections 3.3.4.2-3.3.4.3, Table 3.8, Table 3.10). It is also possible that H2bS32Phos had a similar effect, but the

low conversion observed for H2b₂₇₋₄₇ peptides (1Ac ≤12%) makes this data less reliable (Table 3.10). This difference in inhibition profiles with HAT and HAT-BRD may be due to a change in the binding mode of the peptides. The addition of the BRD to the structure can obstruct certain binding sites, change the conformation of the HAT domain, or bind the peptides itself. Because H3T11Phos and H3S28Phos are on different lengths of H3 (H3₁₋₂₁ and H3₁₄₋₃₄) and because the acetylation sites on the H3₁₄₋₃₄ peptide are not verified (but are theorised to be K14 and K18), it is difficult to extract any information on binding modes. There is no sequence alignment between these two peptides. A crystal structure of tGCN5 with CoA and H3₅₋₂₃S10Phos showed a large divergence in peptide binding interactions and structure compared to a control structure with unphosphorylated H3₅₋₂₃ peptide. In the H3₅₋₂₃S10Phos, the *N*-terminal and *C*-terminal areas of the peptide became ordered making new interactions to the tGCN5 protein flexible loops. In addition, the interactions to H3S10 and H3T11 amino acids were altered.¹⁰⁵ No structure of tGCN5 with a H3T11Phos modified peptide is available. However, it is likely that alterations in binding occur with H3T11Phos/H3S28Phos which might affect the PCAF HAT-BRD structure (no crystal structure available) differently to the PCAF HAT structure.

These results showed that there were slight differences in PCAF HAT and HAT-BRD activity. These differences in activity were dependent of the substrate site (*e.g.* H3₁₋₂₁ vs. H4₁₋₂₁ peptides) and on interactions with PTM on substrate peptides (*e.g.* H3T11Phos/H3S28Phos). However, the role of PCAF BRD remains unclear. As H3T11Phos and H3S28Phos appear to particularly affect the HAT-BRD construct, identifying the nature of interactions between the PTM and PCAF might help understand how the BRD

is involved. In a first instance, small molecule PCAF BRD chemical probes L-Moses and GSK2047 might be used to help distinguish whether binding to the bromodomain active site is somehow involved. However, the use of these BRD inhibitors in time-dependent assays of H3₁₋₂₁ with PCAF HAT-BRD did not affect the acetylation activity of the enzyme (Figure 2.13), indicating that binding of the BRD for unmodified H3₁₋₂₁ peptides does not factor into differences of activity. Although further investigation is necessary, the results obtained in the rapifleX MALDI-TOF-MS HTS constitute a step in further understanding crosstalk between the different domains of PCAF.

3.6.3 RapifleX MALDI-TOF-MS HTS results for screening of SIRT1

Sirt1 is known to deacetylate both histone and non-histone proteins.²⁰ On histone proteins, SIRT1 has been shown to deacetylate H1K26, H3K9, H3K14, H3K56 and H4K16.¹⁸² It was chosen for study in the rapifleX screen because of the shared H3K9 and H3K14 and non-histone substrates with PCAF. Any overlap between PTM crosstalk affecting PCAF and SIRT1 would therefore be interesting to investigate further. Because SIRT1 is an HDAC and cannot have any activity in the absence of Ac PTMs, only the SETIV library which is a library of combinatorial acetylation PTMs was screened.

Manual examination of the spectra showed polymer contamination (likely PEG) in many of the samples spotted with the mosquito liquid handler and analysed using rapifleX MALDI-TOF-MS. As polymer peaks sometimes overlapped with peptide peaks, this led to some false positive results in the SETIV library screen with SIRT1. Although the SETIV library screen with PCAF HAT/HAT-BRD also showed polymer contamination, the smaller range of masses screened for (parent mass +42/+84 Da) compared with the SIRT1 screen

(parent mass -42 Da to -210 Da) meant that there were lower chances of overlap between peaks of interest and polymer peaks in the PCAF HAT/HAT-BRD screens, and lower chances of false positive results. However, overall results with the SETIV library screening are likely less reliable than for the other libraries, increasing the need for rigorous validation of screen results.

Taking into account data quality, a clear trend was nevertheless observed in the SETIV screen against SIRT1, showing non-specific deacetylation of all Ac PTMs on library peptides, regardless of sequence or of PTM position. In time-course assays with H3₁₋₂₁ peptides, different rates of deacetylation were observed between H3₁₋₂₁K9Ac and H3₁₋₂₁K14Ac (Figure 2.7). However, assay conditions were chosen which ensured full deacetylation of both of the native substrates in an end point assay, assuming that non-native substrates would be deacetylated at a much slower rate. If there was any difference in deacetylation rates between the library members, it could not be observed with the assay conditions used.

These results could indicate that SIRT1 has much less sequence specificity than PCAF, and is not affected by any of the combinatorial Ac PTMs. Whether this is true for SIRT1 generally, or if this construct of SIRT1 lacked portions of the protein which give increased sequence selectivity would have to be tested.

3.6.4 RapifleX MALDI-TOF/TOF-MS results

Initial attempts were made to integrate MS/MS into the MALDI-TOF-MS HTS workflow. The rapifleX instrument has the capacity to automatically acquire MS/MS spectra for a specified number of peaks of each spectrum (*e.g.* top 10 most intense peaks). This

functionality was tested and MS/MS spectra were successfully acquired, with 432 of the 3111 MS/MS spectra acquired for a 384 well plate being matched to masses of library peptides.

MALDI-TOF/TOF-MS spectra were manually acquired and analysed for certain peptides. Analysis of an internal control peptide incubated with PCAF (SETI biotin-ahx-H3₁₋₂₁-NH₂) showed that identification of PTM sites was possible for library peptides screened, with the *N*-terminal and *C*-terminal modifications successfully identified by the biotools software used. However, MALDI-TOF/TOF-MS spectra of other peptides incubated with PCAF did not give data of sufficient quality for the PTM site to be identified. SETI biotin-ahx-H3₁₄₋₃₄-NH₂ and SETII H-H2a₁₋₂₁-ahx-lys-biotin are used as examples to show the low data quality obtained for most library peptides tested manually. In the software, a sequence is input with invariable modifications (*N*- and *C*-terminal tags), and optional modifications (*e.g.* Ac PTM at all lysines), and the software gives sequence coverage and scores for each optional sequence. The value of the scores given to each sequence is determined by the scoring algorithm (not disclosed by Bruker). Relative comparison of the scores is being used for analysis, rather than the numerical value given. Although sequence coverage of fragments identified for the H3₁₄₋₃₄ and H2a₁₋₂₁ peptides are relatively high (up to 85% and 95% respectively), similarly high sequence coverage is seen for multiple optional sequences, and scoring for the different alternative PTM positions was not sensible, unlike scores obtained for SETI biotin-ahx-H3₁₋₂₁-NH₂ alternative sequences. The no PTM sequence option had a high score for both peptides tested even though the parent mass for fragmentation corresponded to a 1Ac product of PCAF acetylation. These issues with the biotools sequence coverage and scoring are

likely due to insufficient intensity of fragments leading to noise being identified as fragments.

Results with internal control peptide SETI biotin-ahx-H3₁₋₂₁-NH₂ peptide showed that it is technically possible to use MS/MS on the library peptide to identify enzyme site selectivity. Two issues have hindered the integration of this functionality into the current HTS workflow: length of analysis and poor quality of data obtained. The poor quality of data is mostly likely due to insufficient concentration of product peptides being analysed. This is due in part to low library quality. A simple way to counter this would be to increase the concentration of library peptides in samples. This would lead to a higher consumption of materials for the screen. However, as the necessary assay volume is 10 µL/well, this might be advisable, if better quality data can be obtained. The second hurdle in developing high throughput MS/MS is the time needed for data analysis. Currently an analysis time of 10-30 min per MS/MS spectrum is needed. However, use of software designed for proteomics research (*e.g.* using MASCOT servers) is likely possible and is currently being explored as an option for high-throughput MS/MS analysis.

3.8 Future work

3.8.1 Validation of results obtained in the rapifleX MALDI-TOF-MS screen against PCAF HAT and HAT-BRD

Examples of site selectivity and PTM crosstalk not previously reported were observed for PCAF HAT and HAT-BRD. Because of the inherent limitations of HTS (prioritisation of throughput over quality means a higher possibility of false positive results), peptides identified in the screen as affecting the activity of PCAF HAT/HAT-BRD should be validated using untagged, clean and quantified peptides in MALDI-TOF-MS assays and using MS/MS to identify acetylation sites (Table 3.13).

PCAF was found to acetylate both H2a₁₋₂₁ and H2b₁₋₂₁, with the site of acetylation on H2a₁₋₂₁ identified as H2a₁₋₂₁K9 and no site of acetylation identified for H2b₁₋₂₁ due to low peptide conversion. As these are not previously recorded substrates for PCAF, further validation using MALDI-TOF-MS with untagged peptides and using MS/MS to determine site selectivity is necessary. Validation of the site selectivity of PCAF on nucleosomes and in cells is needed to determine whether activity of PCAF for H2a and H3b is biologically relevant.

A number of peptide PTMs were shown to affect the enzyme activity of PCAF. The activating effect of the K14Me1 PTM on PCAF acetylation activity for the H3₁₋₂₁K14Me1 peptide observed in the HTS needs to be validated using untagged peptides corresponding to H3₁₋₂₁K14Ac, H3₁₋₂₁K14Me1, H3₁₋₂₁K14Me2 and H3₁₋₂₁K14Me3 to confirm that the effect observed is reproducible and specific to K14Me1. In addition, the site of the two enzymatic acetylation events observed must be identified using MS/MS,

to better understand the effect of the K14Me1 PTM on PCAF selectivity and activity (Chapter 4). In the HTS, activity of PCAF against H4₁₋₂₁K12Me1 was not consistent between SETII and SETIII libraries. Because comparison between libraries can be misleading (external controls between libraries showed variable levels of conversion) and as H3₁₋₂₁K14Me1 has been shown to affect PCAF activity, independently checking the effect of H4₁₋₂₁K12Me1 on PCAF activity might be useful to confirm whether mono-methylation PTM activation of PCAF activity is specific to H3₁₋₂₁K14Me1.

Phosphorylation PTMs have also been shown to affect the enzymatic activity of PCAF. The phosphorylation PTMs of H3₁₋₂₁T3Phos, H4₁₋₂₁S1Phos and H2a₁₋₂₁S1Phos were shown to inhibit PCAF acetylation activity towards the peptides compared with unlabelled H3₁₋₂₁. For H3₁₋₂₁T3Phos, it was deduced that PCAF acetylation at K9 was inhibited whilst K14 acetylation was not affected. The inhibitory effects of these peptides and the site of PCAF acetylation for H3₁₋₂₁T3Phos should be validated. As PCAF activity was sensitive to crosstalk with phosphorylation marks, further examples of Ac/Phos crosstalk might be observed by screening an alternative library which exclusively contains combinatorial phosphorylation marks (SETV).

#	Protein	Residues	PTM	Sequence	Purpose for testing
1	H3	1-21	K14Me1	ARTKQTARKSTGGKAPRKQLA	Identify two acetylation sites of PCAF HAT and HAT-BRD using MS/MS (Chapter 4)
2	H3	1-21	K14Me2		
3	H3	1-21	K14Me3		
4	H2a	1-21		SGRGKQGKGKARAKAKTRSSRA	Identify acetylation sites using MS/MS
5	H2b	1-21		PEPAKSAPAPKKGSKKAVTKA	
6	H3	1-21	T3Phos	ARTKQTARKSTGGKAPRKQLA	Confirm inhibitory activity against H3K9 acetylation by PCAF HAT/HAT-BRD
7	H4	1-21	S1Phos	SGRGKGGKGLGKGGAHRHRKV	Confirm inhibitory activity against H4K8 acetylation by PCAF HAT/HAT-BRD
8	H2a	1-21	S1Phos	SGRGKQGKGKARAKAKTRSSRA	Confirm inhibitory activity against H2aK9 acetylation by PCAF HAT/HAT-BRD
9	H3	1-21	T11Phos	ARTKQTARKSTGGKAPRKQLA	Confirm inhibition of H3K9 acetylation Investigate how T11Phos affects the activity of PCAF HAT and HAT-BRD differentially at K14
10	H3	1-21	S10Phos	ARTKQTARKSTGGKAPRKQLA	
11	H3	1-21	S10Phos T11Phos	ARTKQTARKSTGGKAPRKQLA	
12	H3	14-34		KAPRKQLATKAARKSAPATGG	Identify acetylation sites using MS/MS
13	H3	14-34	S28Phos	KAPRKQLATKAARKSAPATGG	Confirm inhibition activity and investigate how S28Phos affects the activity of PCAF HAT and HAT-BRD differentially at K14
14	H3	14-34	T32Phos	KAPRKQLATKAARKSAPATGG	
15	H3	14-34	S28Phos T32Phos	KAPRKQLATKAARKSAPATGG	
16	H4	1-21	K12Me1	SGRGKGGKGLGKGGAHRHRKV	Check whether H4 K12Me1 activation was correctly identified as a false result

Table 3.13: List of peptides from the rapifleX MALDI-TOF-MS HTS to be tested in future work. Peptides 1-3 were synthesised and tested in Chapter 4. For all peptides listed, the purpose of further testing is given.

Differences in activity between PCAF HAT and HAT-BRD were observed towards H3₁₋₂₁ peptides with combinations of S10Phos and T11Phos PTMs, as well as with H3₁₄₋₃₄S28Phos (the role of T32Phos PTMs was not clear). As the role of the BRD of PCAF is not well defined, these peptides may present an avenue to better understand

differences in activity between the HAT and HAT-BRD constructs. Kinetic work to characterise the activity of the two PCAF enzyme construct for S10Phos, T11Phos and S28Phos containing peptides would be necessary. In addition, the site of acetylation of PCAF on H3₁₄₋₃₄ peptides is unclear and MS/MS is necessary to identify acetylation sites, and therefore better understand the inhibitory effect of the phosphorylation PTM of H3₁₄₋₃₄S28Phos. Use of BRD binding chemical probes might also help elucidate how the bromodomain is involved in crosstalk with phosphorylation PTMs.

3.8.2 Acquisition of an improved dataset for the SETIV library screens

The SETIV library was of lower quality compared with the SETI, SETII and SETIII libraries. Internal controls in this library were not reliable (most gave flatline spectra, or showed high levels overlap between peptide and impurity peaks). In addition, the general quality of the peptides was low (*e.g.* many of the peptide peaks were of lower intensity than impurity peaks), and for certain peptides, incorrect information on peptide sequence was initially provided and had to be verified using MS/MS. As this is a commercially sourced library, library quality issues cannot be solved without purchase of a new batch of the library.

It was shown that polymer contamination occurred somewhere in the process of running the assay and acquiring the data, rather than coming from the library itself. This, in addition to low quality of the library, led to low data quality of acquired data. Identification and elimination of the source of contamination (likely PEG from plasticware and consumables), would allow for acquisition of a higher quality dataset for this library. In addition, it was observed that the peptides incubated with PCAF HAT

had unusually high conversion rates, and a biological replicate would be useful to show whether this was due to an error in the assay. These additional replicates were not carried out previously, due to restricted access to the mosquito and rapifleX instruments (no direct access granted due to coronavirus restrictions, results acquired by Dr. Maria Dueñas of the Trost group), and due to expenses (library, plates, mosquito consumables).

In addition, although the data was of poorer quality than for other libraries, trends were observed in the SETIV library screen with PCAF HAT and with SIRT1. The SETIV library provided a confirmation of the site selectivity of PCAF HAT towards H2a₁₋₂₁ peptides which was previously observed in the SETII library. Validation of the site selectivity observed with untagged purified H2a₁₋₂₁ peptide is necessary and would provide more reliable data than a repeat of the screen. Any other trends observed with PCAF and the SETIV library were also validated by comparison with an alternative library. For SETIV screened against SIRT1, SIRT1 showed no selectivity towards specific sequences or PTM positions. Alternative assay conditions would likely give more informative results as to the selectivity of SIRT1.

For SIRT1 screening of SETIV, a shorter assay (or lower enzyme concentrations) might distinguish between rates of deacetylation for certain peptides. Using alternative constructs for SIRT1 might also show that this construct lacks portions of the protein which confer selectivity. Screening SIRT1 against other libraries which contain non-Ac PTMs might also show examples of cross-talk between peptide PTMs and the enzyme,

conferring selectivity to SIRT1. For PCAF, no crosstalk between Ac modifications was observed, so SIRT1 may be equally unaffected in this regard.

3.8.3 Further development of the MALDI-TOF-MS HTS with addition of MS/MS

Integration of MS/MS functionality into the MALDI-TOF-MS HTS would be an invaluable addition to the methodology developed. In the screen of the SET libraries against PCAF and SIRT1 enzymes, lack of identification of the acetylation sites (*e.g.* for H3₁₄₋₃₄S28Phos screened against PCAF) have made it more difficult to understand the selectivity of the enzymes. In addition, whilst SIRT1 did not appear to show selectivity towards the peptides tested, some of the library peptides were synthesised with an *N*-terminal acetylation and use of MS/MS would be needed to verify whether SIRT1 is involved in deacetylation of *N*-terminal Ac groups.

One of the main advantages of MS over other techniques is its ability to identify multiple enzymatic modifications as separate events on any enzyme substrate studied (unlike antibodies which are specific to selected substrates), and that further characterisation of enzyme site selectivity is possible with MS/MS. The current set up has already allowed for identification of examples of PTM crosstalk affecting the activity of epigenetic enzymes (*e.g.* H3₁₋₂₁K14Me1 activates PCAF acetylation activity). However, these results could not be immediately characterised by using MALDI-TOF/TOF-MS to identify the acetylation sites of PCAF on the H3₁₋₂₁K14Me1 peptide. Rapid acquisition and analysis of MS/MS data for all peptides screened would give invaluable site selectivity information

about enzymes screened, making the assay developed a very useful tool in the understanding of the effects of crosstalk on enzyme site selectivity.

3.8.4 Expansion of the MALDI-TOF-MS HTS developed

The rapifleX MALDI-TOF-MS HTS has been developed with commercially available histone peptide libraries screened against PCAF and SIRT1 enzymes. This was chosen as a screening method because the MALDI-TOF-MS assay developed to characterise first PCAF then SIRT1 was thought to be robust enough to identify mass changes (± 42 Da for PCAF and SIRT1) on a varied panel of histone peptides. Screening different classes of enzymes against the histone peptide libraries using this method would exemplify the possible range of the method developed. In addition, because the HTS is a model for biological activity of enzymes, identifying whether the results obtained are biologically relevant would illustrate the utility of this method as a first step in identifying novel substrate sites for enzymes and novel instances of PTM crosstalk. Current work on characterising epigenetic enzymes focusses on identifying sites of activity for the enzymes. However, systematic screening of the effects of PTM crosstalk on the activity of the enzymes characterised is needed. Epigenetic PTMs are a dynamic and interdependent regulatory system in cells, which exists as a network rather than as isolated PTMs. The methodology developed can help better understand how PTM modifications and epigenetic enzymes are involved in a larger epigenetic regulatory system.

3.9 Conclusion

Enzyme activity in a cellular context is regulated by a complex signalling network, which includes but is not limited to post translational modifications (PTMs) of histones and regulatory proteins, and interactions between and within protein complexes. As an epigenetic enzyme with multiple domains with different biological roles, PCAF activity is likely affected by both external PTM signalling and internal interaction between its domains. The aim of this work is to better understand the extent of the role of these two aspects on PCAF regulation.

The role of PTM modifications on the activity of PCAF HAT was explored by testing the activity of PCAF against a set of histone peptide libraries with combinatorial PTMs to identify inhibitory and activating effects of selected PTMs or combinations of PTMs. The effect of interplay between the PCAF HAT domain and bromodomain was explored by comparing PCAF activity of HAT and HAT-BRD constructs.

The levels of epigenetic PTMs in cells are regulated by the opposing activities of writer and eraser epigenetic enzymes for each modification site and type. PCAF is a HAT which is reported to acetylate H3K14 and H3K9. SIRT1 is an HDAC which is reported to deacetylate H3K9 and H3K14. They also share a number of non-histone protein substrates. Understanding how both PCAF and SIRT1 are regulated, and the similarities and differences between their regulatory networks would shed light on how levels of acetylation at H3K9/H3K14 is determined.

A rapifleX MALDI-TOF-MS HTS for histone peptide libraries with PTMs was developed. This was used to screen 376 peptides (4 libraries of 94 peptides) against two constructs

of PCAF (HAT and HAT-BRD) and a NEC for quality check. A further 94 peptides were screened against SIRT1 with a corresponding NEC. Counting all biological and technical replicates, approximately 12000 spectra were acquired with Bruker microflex (low-throughput) or rapiflex (high-throughput) MALDI-TOF-MS. These spectra were processed using two optimised scripts: first a peak picking script on Flexanalysis software (Bruker); second a custom script designed for this screening method, which retrieves the data needed to quantify enzyme activity against each peptide. The data from the screen was used to explore the activity profiles of PCAF and SIRT1.

The screen was used to confirm site selectivity of PCAF for H3 and H4, and identify activity of PCAF at H2aK9 and two non-identified locations on H2b. Effects of H3S10Phos and H3T11Phos PTM crosstalk reportedly observed with the homologous GCN5 protein was also seen for PCAF. A H3K14Me1 PTM was shown to activate PCAF HAT and HAT-BRD acetylation activity. In addition, H3S28Phos was shown to inhibit PCAF activity at an undefined location. Different levels of activity were observed between H3T11Phos/H3S28Phos and either PCAF HAT or HAT-BRD, hinting at a biological role for PCAF BRD. The SIRT1 construct used was shown to lack selectivity, deacetylating H3, H4, H2a and H2b at all lysine sites equally. Some of these effects were further studied in Chapter 4, to validate the results of the screen and further characterise PCAF and SIRT1 enzymes.

These results show the success of this HTS screening method in improving the understanding of crosstalk between PTMs and their interactions with epigenetic protein domains.

Chapter 4 Characterisation of PCAF and SIRT1 substrates

4.1 Introduction

4.1.1 Literature substrates for PCAF and SIRT1

HAT and HDAC enzymes, each with different substrate selectivity, dynamically regulate the levels of acetylation PTMs in cells, contributing to many biological processes.^{37,56,63,86,222,386} PCAF (HAT) and SIRT1 (HDAC) have largely opposite roles in regulating acetylation levels and in cell biology, with a large overlap in substrate proteins. This includes H3K9 and H3K14 on histone proteins, as well as non-histone proteins HIF1 α and p53.^{20,119,182} Developing a better understanding of the site selectivity on substrates and of the signalling systems that regulate PCAF and SIRT1 helps understand how these enzymes contribute to regulation of acetylation levels.

Peptides are a commonly used simplified model for studying proteins, as they successfully mimic PPI.^{387,388} Peptide sequences corresponding to portions of substrate proteins can be synthesised relatively easily, rapidly and economically on automated peptide synthesisers, allowing for sequence variation on a larger scale than when working with proteins.²⁸¹ Peptide models have previously been used to study PCAF/GCN5 and SIRT1 enzymes, including histone,³⁸⁹ HIF1 α ,¹²⁵ HIV,¹⁵⁸ and p53 peptides.^{105,126} These models can be useful to understand binding modes of epigenetic enzymes for specific substrates.^{105,159,287} However, epigenetic enzymes (*e.g.* GCN5) act in a context dependent manner, with different activity and selectivity trends identified

with peptides *in vitro* and in cells.^{122,385} Results from simplified peptide and protein model assays must therefore be validated to establish biological relevance.

The heterodimeric Hypoxia inducible factor-1 (HIF-1) transcription factor is composed of HIF1 α and HIF1 β . Whilst HIF1 β is ubiquitously expressed, HIF1 α levels are oxygen-regulated in its stability. In normoxia, HIF1 α is hydroxylated at two proline residues, which leads to its ubiquitination by the von-Hippel-Lindau (VHL) protein and subsequent degradation of HIF1 α . Because prolyl hydroxylases use oxygen as a substrate, they have reduced activity at low oxygen levels (hypoxia). This leads to increased cellular concentrations of HIF1 α and nuclear translocation of HIF1 α to form the HIF1 complex between HIF1 α and HIF1 β . Complex formation then leads to expression of genes needed for cell adaptation to hypoxic conditions. HIF1 α is often overexpressed in cancer and is therefore a common target for anti-cancer drug development.^{390–393} Acetylation of HIF1 α and its binding partners provides a hypoxia mediated regulatory system which acts co-operatively with oxygen dependent regulation, providing an additional layer of feedback.^{219,394} PCAF and SIRT1 were found to regulate levels of HIF1 α acetylation, having opposite roles on HIF1 α K674 acetylation.¹²⁵ HIF1 α was shown to be activated and stabilised by acetylation, and deacetylation of HIF1 α was correlated with increased protein degradation.^{125,220,395,396} This makes HIF1 α an interesting target to test using MALDI-TOF-MS, in a first attempt to extend the MALDI-TOF-MS HTS developed to non-histone proteins.

4.1.2 Use of MALDI-TOF-MS to validate histone and non-histone substrates of PCAF and SIRT1

To expand on histone substrate selectivity reported in literature, a rapifleX MALDI-TOF-MS HTS was used to screen a panel of histone peptides with combinatorial post translational modifications (PTMs) against PCAF and SIRT1 (Chapter 3). In this screen PCAF substrate sites and examples of PTM crosstalk affecting PCAF were tentatively identified. This included a previously unreported PTM crosstalk between H3₁₋₂₁K14Me1 and PCAF HAT. A substantial increase in acetylation activity was observed with PCAF HAT/HAT-BRD for H3₁₋₂₁K14Me1 compared with H3₁₋₂₁K14Me2/3 and H3₁₋₂₁K14Ac peptides. High throughput screens are useful because of their scope, but suffer from lower data quality than low throughput methods and can deliver false positive results. In addition, the peptides screened were tagged peptides, with variable quality of synthesis leading to many impurities, both of which can affect observed results. Further validation is therefore needed, using untagged and quality-controlled peptides. In addition, better understanding of enzyme activity for the peptides can be achieved by identifying acetylation sites using MS/MS and characterising kinetic parameters.

The scope of the MALDI-TOF-MS HTS, previously limited to histone peptide libraries, could be expanded to screening libraries of peptide sequences based on non-histone substrates for PCAF and SIRT1. HIF1 α was chosen as a model to try and expand the scope of MALDI-TOF-MS testing of PCAF and SIRT1 selectivity. PCAF and SIRT1 have both been shown to target K674 on HIF1 α .¹²⁵ HIF1 α acetylation at K532 was shown to negatively affect its activity in mice, but has not been confirmed as an acetylation site in humans.³⁹⁴

It was further reported that SIRT1 and PCAF did not affect K532 acetylation levels in human cell lines.¹²⁵ SIRT1 and PCAF were linked to p300 activity on HIF1 α , with PCAF facilitating binding of p300 to HIF1 α and SIRT1 abolishing p300 activity.¹²⁵ It was shown that p300 catalyses the acetylation of K709 on HIF1 α , a site which is not acetylated by PCAF.³⁹⁵ Based on these reports, three sections of HIF1 α were targeted for screening: residues 522-542 (K532), 664-684 (K674) and 700-721 (K709). Unmodified and acetylated peptides corresponding to these segments were synthesised for testing against PCAF HAT/HAT-BRD and SIRT1. If successful, this could be the basis for development rapifleX MALDI-TOF-MS screening of HIF1 α peptide libraries, exploring potential crosstalk between PTMs placed on HIF1 α peptides.

4.1.3 Aims

The aim of this chapter is to further characterise PCAF and SIRT1 substrate selectivity using MALDI-TOF-MS. Peptides of interest from the rapifleX MALDI-TOF-MS HTS performed in Chapter 3 will be tested for activity against PCAF to validate the results from the HTS. Selected HIF1 α peptides will be tested for activity against PCAF and SIRT1, to explore the use of MALDI-TOF-MS screening for non-histone peptides.

Specific aims are:

- Synthesis of selected histone and non-histone peptides for characterisation.
- Use of MALDI-TOF-MS time-dependent assays to characterise activity of PCAF and SIRT1 for substrates
- Use of MS/MS to identify the site of acetylation of PCAF on substrate peptides

4.2 Peptide synthesis

4.2.1 Peptides chosen for synthesis

In the rapifleX MALDI-TOF-MS HTS, PCAF HAT/HAT-BRD was shown to catalyse the acetylation of a control H3₁₋₂₁ peptide to give single (1Ac) and double acetylation (2Ac) products. The acetylation sites of PCAF on H3₁₋₂₁ peptides were proven by MALDI-TOF/TOF-MS to correspond to acetylation at K14 followed by K9 (Chapter 2). Similarly, in the MALDI-TOF-MS HTS, PCAF HAT/HAT-BRD acetylated H3₁₋₂₁K14Me1 giving 1Ac and 2Ac products. As K14 is mono-methylated in the H3₁₋₂₁K14Me1 peptide, it is possible but unlikely that an additional PTM is being added to K14. With H3₁₋₂₁K14Me2/3 peptide substrates, only a 1Ac product was observed. Although the results were considered reliable by crosschecking between libraries screened, further validation and characterisation is needed. H3₁₋₂₁K14Me1, H3₁₋₂₁K14Me2 and H3₁₋₂₁K14Me3 peptide sequences were therefore chosen for synthesis (Table 4.1).

HIF1 α peptides were chosen for study against PCAF HAT/HAT-BRD and SIRT1 based on reported literature mapping HIF1 α acetylation (Section 4.1.2). Three acetylation sites have been identified on HIF1 α which have been studied in the context of PCAF and SIRT1 activity: K523, K674 and K709. HIF1 α K523/K523Ac were reportedly not substrates (or at least, not the main substrates) for PCAF or SIRT1 activity.¹²⁵ HIF1 α K674/K674Ac were reported as substrates for PCAF and SIRT1 respectively.¹²⁵ HIF1 α K709 is reported as a substrate for p300 but not for PCAF. SIRT1 deacetylation activity at this site has not been established.³⁹⁵ 21mer peptides containing each of these HIF1 α lysine substrate sites

were synthesised with a non-acetylated and acetylated version produced to test against PCAF and SIRT1 respectively (Table 4.1).

Protein	Residues	PTM	Sequence	Background
H3	1-21	K14Me1	ARTKQTARKSTGG KMe APRKQLA	H3 ₁₋₂₁ K14Me1 Identified in the MALDI-TOF-MS screen against PCAF as PTM crosstalk
H3	1-21	K14Me2	ARTKQTARKSTGG KMe2 APRKQLA	
H3	1-21	K14Me3	ARTKQTARKSTGG KMe3 APRKQLA	
HIF1 α	522-542		YVDSDMVNEFKLELVEKLFAE	Possible substrate for PCAF
HIF1 α	522-542	K532Ac	YVDSDMVNEFK Ac LELVEKLFAE	Possible substrate for SIRT1
HIF1 α	664-684		SRTASPNRAGKGVIEQTEKSH	PCAF targets K674 ¹²⁵
HIF1 α	664-684	K674Ac	SRTASPNRAG KAc GVIEQTEKSH	SIRT1 targets K674Ac ¹²⁵
HIF1 α	700-721		VPEEELNPKILALQNAQRK	p300 targets at K709 ³⁹⁵ Not a substrate for PCAF ³⁹⁵
HIF1 α	700-721	K709Ac	VPEEELNPK Ac ILALQNAQRK	Not tested against SIRT1

Table 4.1: Peptides chosen for substrate characterisation with PCAF and SIRT1.

The peptide sequence and previous work (background) leading to these peptides being chosen substrates are shown. PTM: post translational modification.

4.2.2 Peptide synthesis

Peptide synthesis was carried out using Fmoc-based chemistry on a 0.05 mmol scale with a Chorus automated peptide synthesiser (Gyros). N,N'-Diisopropylcarbodiimide (DIC) / Oxyma was used for coupling chemistry, DIC being the coupling agent and Oxyma the activating base. This method minimises racemisation at the α -carbon of the coupling amino acid. The Fmoc-protected amino acid is activated by the carbodiimide, which then undergoes an acyl exchange with Oxyma. This forms an activated oxime ester, which favours an intramolecular reaction to form a peptide bond over intermolecular side reactions, decreasing racemisation (Figure 4.1). Cleavage was carried out on the Chorus peptide synthesiser with a trifluoroacetic acid (TFA) / scavengers mixture. The peptides were purified on an Agilent semi-prep HPLC using a 5 μ M NX-C18 110 Å column to >95% purity and the mass of the purified peptides was confirmed by MALDI-TOF-MS

(Materials and methods sections 7.5, Table 7.6, Appendix Ci). Stock solutions were made of purified peptides in D₂O and NMR quantification was used to determine accurate peptide concentrations.³⁹⁷ Peptide synthesis and purification went smoothly for all peptides except for H3₁₋₂₁KMe2. The problem was initially linked to degraded Fmoc-lys(Me2)-OH amino acid. However, issues with the synthesis persisted with newly purchased amino acid. This peptide was therefore sourced from GL Biochem.

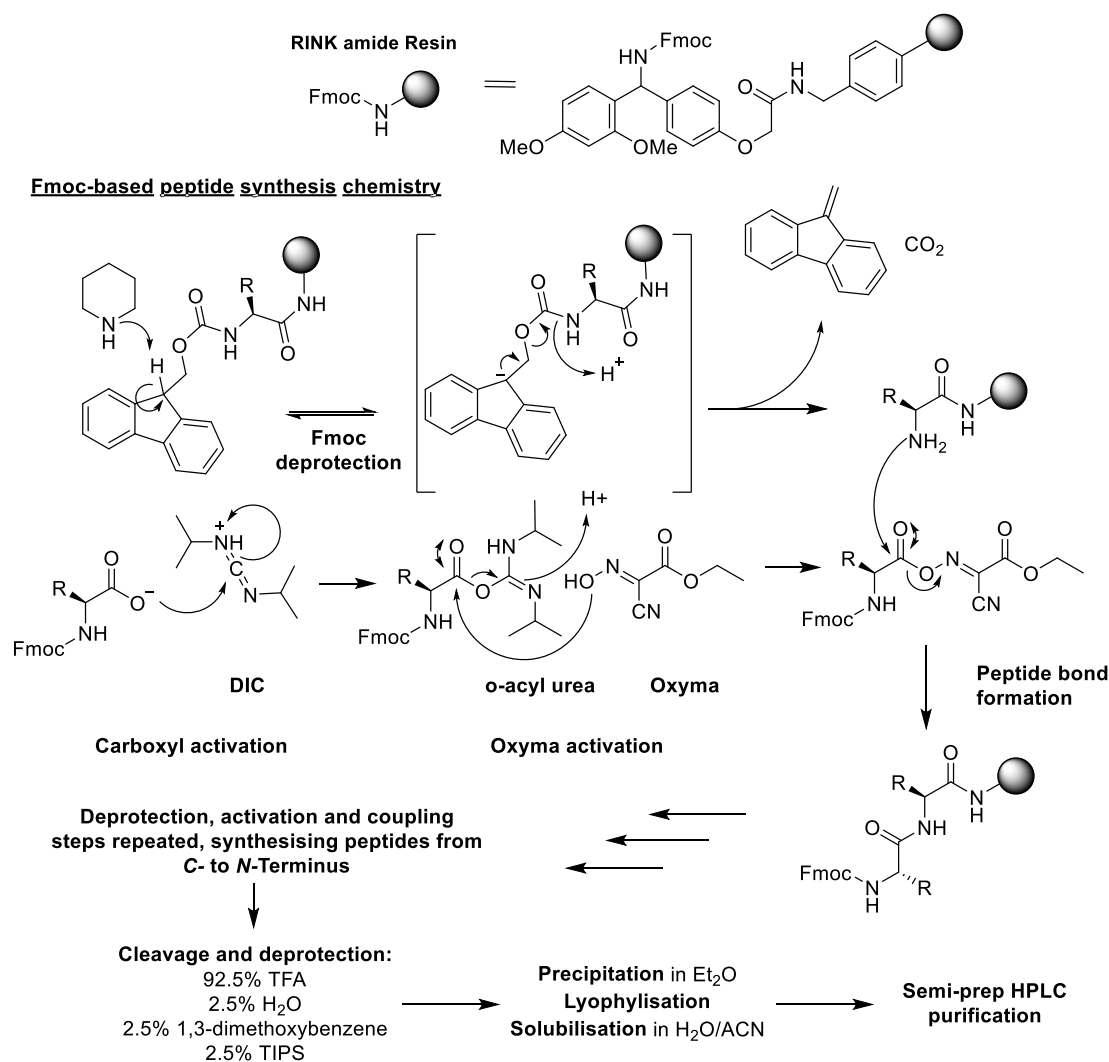


Figure 4.1: Peptide synthesis scheme for automated synthesis on a Chorus synthesiser.

Fmoc peptide chemistry was used with *N,N'*-Diisopropylcarbodiimide (DIC) as coupling agent and Oxyma as activating base on a 4-(2',4'-Dimethoxyphenyl-Fmoc-aminomethyl)-phenoxy (RINK) amide resin. TFA deprotection and cleavage was carried post-synthesis followed by purification using semi-prep HPLC.

4.3 Characterisation of PCAF substrate selectivity

4.3.1 Site selectivity of PCAF HAT and HAT-BRD for K14 methylated H3₁₋₂₁ peptides explored with MALDI-TOF/TOF-MS

In Chapter 2, the site selectivity of PCAF for the unmodified H3₁₋₂₁ peptide was identified using MALDI-TOF/TOF-MS. This showed that H3₁₋₂₁ was acetylated by PCAF HAT/HAT-BRD at K14 followed by acetylation at K9, consistent with literature reports.¹¹⁹ For H3₁₋₂₁K14Me1, two acetylation events were observed in the rapifleX HTS, whereas for H3₁₋₂₁K14Me2/3 only one acetylation event was observed. Because the K14 position on H3₁₋₂₁K14Me1 is methylated, it is likely that the two acetylation sites for this substrate are different to that of H3₁₋₂₁ acetylation sites (K14 and K9). MS/MS was therefore used to map the acetylation sites for the mono-, di- and trimethylated H3₁₋₂₁ peptides, to shed light on the effect of methylation on PCAF site selectivity.

The Bruker rapifleX instrument has MALDI-TOF/TOF-MS capabilities (Section 2.2.4.1). MS/MS samples of PTM modified peptides were prepared in the same way as H3₁₋₂₁ unmodified peptide samples (Materials and Methods 7.7.7) Briefly, H3₁₋₂₁K14Ac, H3₁₋₂₁K14Me1, H3₁₋₂₁K14Me2 and H3₁₋₂₁K14Me3 peptides (20 µM) were incubated with PCAF HAT or HAT-BRD (1 µM) at 37°C for 5 h. Representative aliquots were identified for the K14 acetyl, mono- di- and tri- methyl peptides incubated with PCAF HAT/HAT-BRD (Figure 4.2, peaks used for MS/MS identified by black boxes) containing either parent peptide, single acetylation product (1Ac) or double acetylation product (2Ac). Control peptides were tested using MALDI-TOF/TOF-MS to ensure successful peptide identification through fragment mapping. This included H3₁₋₂₁ unmodified

peptide (Figure 2.8), H3₁₋₂₁K14Ac incubated with PCAF HAT which showed a single enzymatic acetylation by PCAF at K9 as expected, and H3₁₋₂₁K14Me1 unmodified peptide for which the mono-methylation at K14 was correctly identified. In addition, the C-terminal amide which results from cleavage of peptides from RINK amide resin was consistently identified in the control peptides, showing that the method is sufficiently sensitive to identify peptide PTMs which give a small mass difference (Figure 4.3). Coverage of fragment identification for each peptide tested was sufficient to give confidence in the results obtained, with 95–100% sequence coverage for all spectra shown (Figure 4.3, Figure 4.4, Figure 4.5).

In the MALDI-TOF/TOF-MS results for H3₁₋₂₁K14Me1 acetylated by PCAF HAT and HAT-BRD, the first enzymatic acetylation peak corresponded to acetylation at K18, giving H3₁₋₂₁K14Me1K18Ac as product. The second acetylation peak corresponded to acetylation of K9, giving H3₁₋₂₁K9AcK14Me1K18Ac as product (Figure 4.4). These results pointed towards a more efficient acetylation of H3K18 and a less efficient acetylation of H3K9 by PCAF HAT and HAT-BRD in the presence of an H3K14Me1 PTM. In the MALDI-TOF/TOF-MS results for both H3₁₋₂₁K14Me2 and H3₁₋₂₁K14Me3 acetylated by PCAF HAT and HAT-BRD, a single acetylation peak was observed. The acetylation site was identified as H3K9 by MS/MS, giving H3₁₋₂₁K9AcK14Me2 and H3₁₋₂₁K9AcK14Me3 products respectively (Figure 4.5). These results indicated that the H3K14Me1 PTM had a selective activating effect on H3K18 acetylation of H3₁₋₂₁ peptides by PCAF HAT and HAT-BRD, which was not the case with H3K14Me2/3 PTMs. Results with PCAF HAT and HAT-BRD were identical, and only results with peptides acetylated by PCAF HAT are shown here.

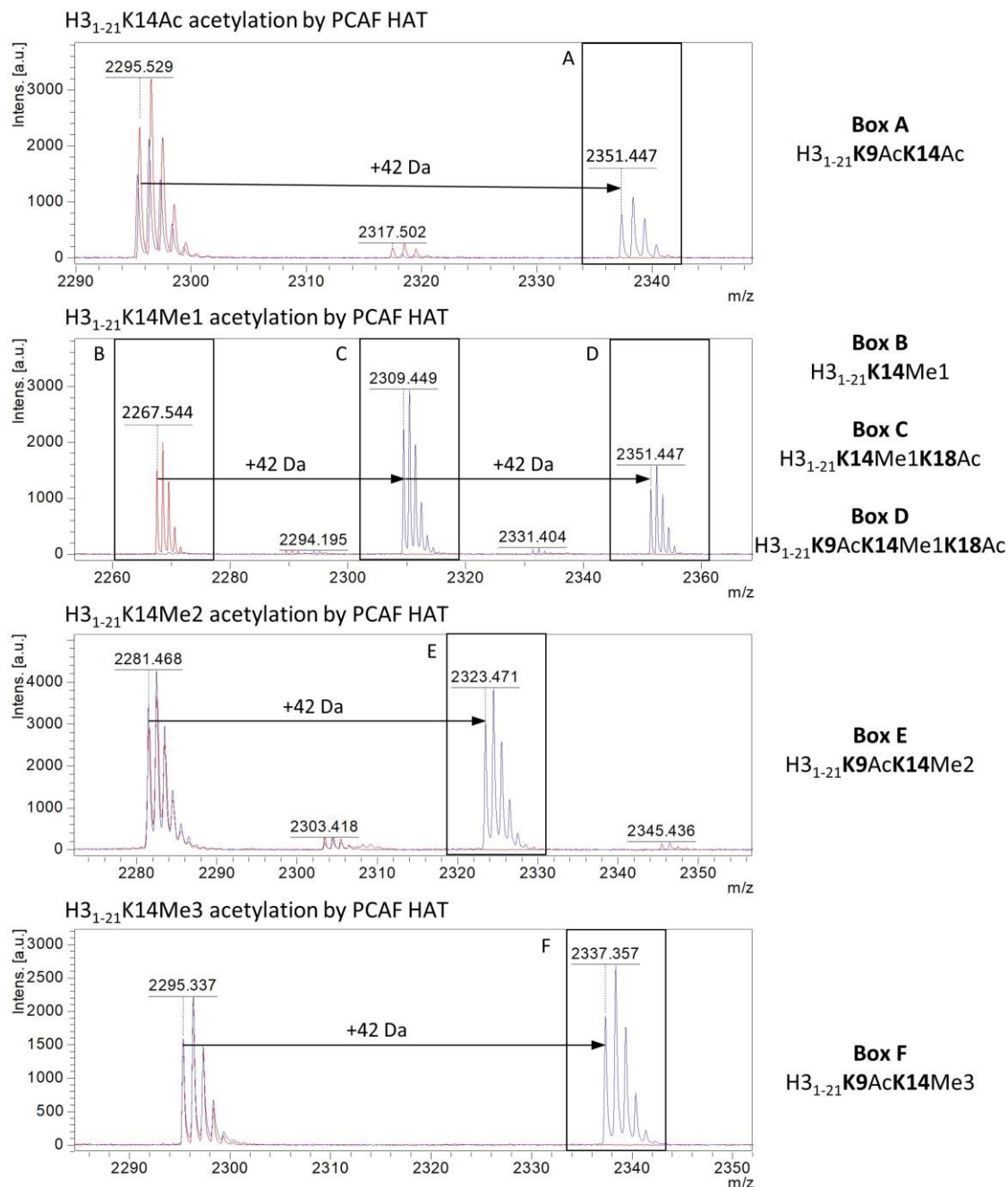
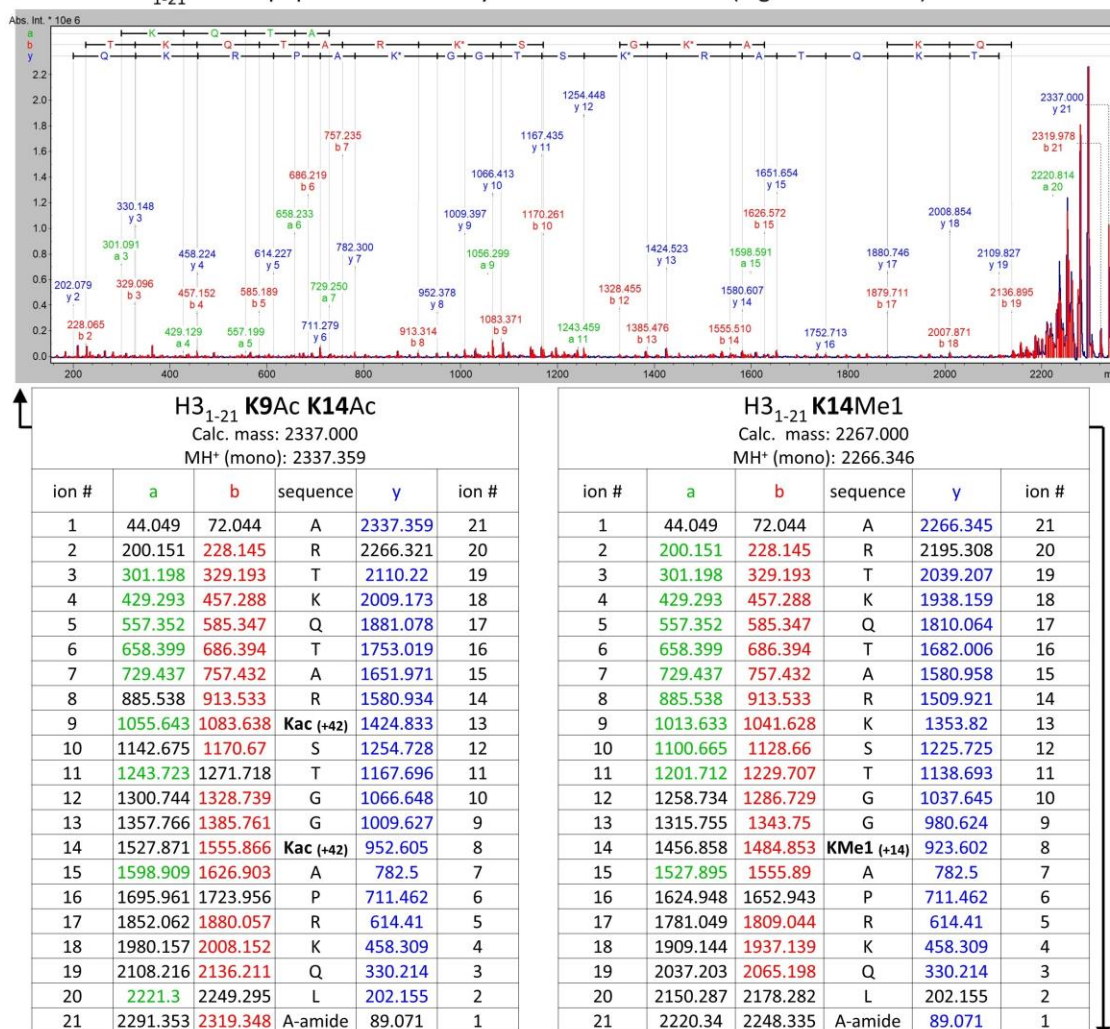


Figure 4.2: MALDI-TOF-MS traces for H3₁₋₂₁ peptides with K14 PTMs acetylated by PCAF HAT. H3₁₋₂₁K14Ac, H3₁₋₂₁K14Me1, H3₁₋₂₁K14Me2 and H3₁₋₂₁K14Me3 were incubated with PCAF HAT to obtain enzymatically acetylated product peptides for use in MALDI-TOF/TOF-MS analysis. Assay conditions were enzyme (1 μ M), AcCoA (100 μ M) and peptide (20 μ M) incubated at 37°C for 5h. Spectra are shown as NEC (red) overlaid with spectra of product peptides (blue). The peaks for which MALDI-TOF/TOF-MS analysis is shown below (Figure 4.3, Figure 4.4, Figure 4.5) are indicated with a numbered black box, and the identity of the peptide is given. Equivalent results were obtained with PCAF HAT-BRD (not shown).

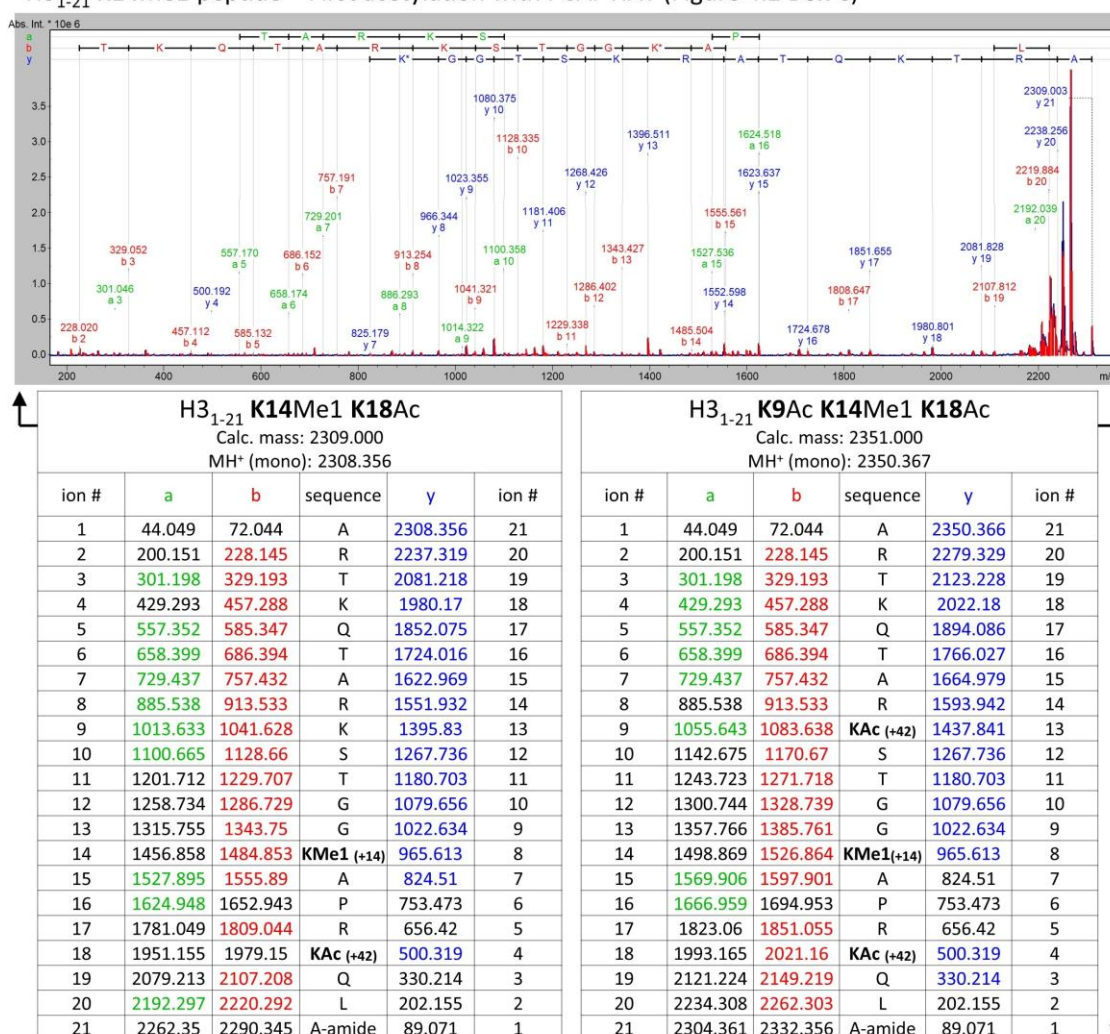
Control: H3₁₋₂₁ K14Ac peptide – First acetylation with PCAF HAT (Figure 4.2 Box A)



Control: H3₁₋₂₁ K14Me1 peptide – unmodified (Figure 4.2 Box B)

Figure 4.3: MALDI-TOF/TOF-MS spectra of control H3₁₋₂₁ peptides with Ac or Me1 PTMs at K14. MS/MS spectra acquired using MALDI-TOF/TOF-MS and analysed using Biotoools are shown. Top: MS/MS of the single acetylation product of H3₁₋₂₁K14Ac peptide treated with PCAF HAT, correctly identifying K9 as the acetylation site. Bottom: MS/MS of the unmodified H3₁₋₂₁K14Me1 peptide, correctly identifying the K14Me1 PTM. See section 4.3.2.1 for assay conditions.

H3₁₋₂₁ K14Me1 peptide – First acetylation with PCAF HAT (Figure 4.2 Box C)

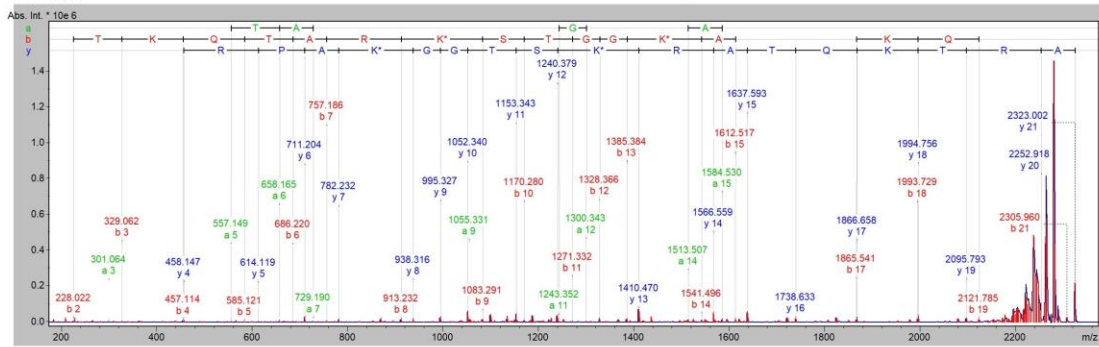


H3₁₋₂₁ K14Me1 peptide – Second acetylation with PCAF HAT (Figure 4.2 Box D)

Figure 4.4: Identification of 1Ac and 2Ac acetylation sites by PCAF HAT on H3₁₋₂₁K14Me1 using MALDI-TOF/TOF-MS.

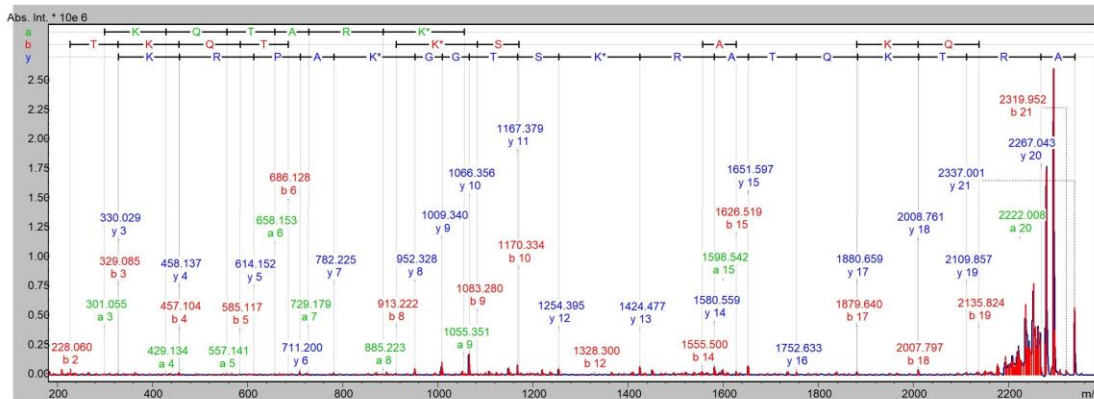
MS/MS spectra acquired using MALDI-TOF/TOF-MS and analysed using Biotoools are shown. Top: MS/MS of the single acetylation product of H3₁₋₂₁K14Me1 peptide treated with PCAF HAT, identifying K18 as the acetylation site. Bottom: MS/MS of the double acetylation product of H3₁₋₂₁K14Me1 peptide treated with PCAF HAT, identifying K9 and K18 as acetylation sites. See section 4.3.2.1 for assay conditions.

H3₁₋₂₁ K14Me2 peptide – First acetylation with PCAF HAT (Figure 4.2 Box E)



H3 ₁₋₂₁ K9Ac K14Me2					
Calc. mass: 2323.000					
MH ⁺ (mono): 2323.380					
Ion #	a	b	sequence	y	Ion #
1	44.049	72.044	A	2323.379	21
2	200.151	228.145	R	2252.342	20
3	301.198	329.193	T	2096.241	19
4	429.293	457.288	K	1995.193	18
5	557.352	585.347	Q	1867.098	17
6	658.399	686.394	T	1739.04	16
7	729.437	757.432	A	1637.992	15
8	885.538	913.533	R	1566.955	14
9	1055.643	1083.638	KAc (+42)	1410.854	13
10	1142.675	1170.67	S	1240.748	12
11	1243.723	1271.718	T	1153.716	11
12	1300.744	1328.739	G	1052.669	10
13	1357.766	1385.761	G	995.647	9
14	1513.892	1541.887	KMe2(+28)	938.626	8
15	1584.929	1612.924	A	782.5	7
16	1681.982	1709.977	P	711.462	6
17	1838.083	1866.078	R	614.41	5
18	1966.178	1994.173	K	458.309	4
19	2094.237	2122.232	Q	330.214	3
20	2207.321	2235.316	L	202.155	2
21	2277.374	2305.369	A-amide	89.071	1

H3 ₁₋₂₁ K9Ac K14Me3					
Calc. mass: 2337.000					
MH ⁺ (mono): 2337.396					
ion #	a	b	sequence	y	ion #
1	44.049	72.044	A	2337.395	21
2	200.151	228.145	R	2266.358	20
3	301.198	329.193	T	2110.257	19
4	429.293	457.288	K	2009.209	18
5	557.352	585.347	Q	1881.114	17
6	658.399	686.394	T	1753.056	16
7	729.437	757.432	A	1652.008	15
8	885.538	913.533	R	1580.971	14
9	1055.643	1083.638	KAc (+42)	1424.87	13
10	1142.675	1170.67	S	1254.764	12
11	1243.723	1271.718	T	1167.732	11
12	1300.744	1328.739	G	1066.684	10
13	1357.766	1385.761	G	1009.663	9
14	1527.908	1555.903	KMe3(+42)	952.641	8
15	1598.945	1626.94	A	782.5	7
16	1695.998	1723.993	P	711.462	6
17	1852.099	1880.094	R	614.41	5
18	1980.194	2008.189	K	458.309	4
19	2108.252	2136.247	Q	330.214	3
20	2221.336	2249.331	L	202.155	2
21	2291.389	2319.384	A-amide	89.071	1



H3₁₋₂₁ K14Me3 peptide – First acetylation with PCAF HAT (Figure 4.2 Box F)

Figure 4.5: Identification of 1Ac acetylation sites by PCAF HAT on H3₁₋₂₁K14Me2 and H3₁₋₂₁K14Me3 using MALDI-TOF/TOF-MS.

MS/MS spectra acquired using MALDI-TOF/TOF-MS and analysed using Biotoools are shown. Top: MS/MS of the single acetylation product of H3₁₋₂₁K14Me2 peptide treated with PCAF HAT, identifying K9 as the acetylation site. Bottom: MS/MS of the single acetylation product of H3₁₋₂₁K14Me3 peptide treated with PCAF HAT, identifying K9 as the acetylation site. See section 4.3.2.1 for assay conditions.

4.3.2 Activity of PCAF HAT and HAT-BRD explored with MALDI-TOF-MS

4.3.2.1 Activity of PCAF HAT and HAT-BRD for H3₁₋₂₁ peptides with K14 PTMs

Next, to compare the rates of acetylation between PCAF HAT and HAT-BRD, and how these were affected by Ac and Me1/2/3 PTMs on H3₁₋₂₁ peptides, time-dependent assays were run (Figure 4.6). PCAF HAT and HAT-BRD were incubated at 37°C with H3₁₋₂₁, H3₁₋₂₁K14Me1, H3₁₋₂₁K14Me2, H3₁₋₂₁K14Me3, and H3₁₋₂₁K14Ac and the acetylation reactions were monitored over 3 h using MALDI-TOF-MS.

The specific activity for each peptide was calculated in the assay linear range. Results showed that PCAF HAT had a much higher specific activity for unmodified H3₁₋₂₁ at K14 than for any other peptide or site. Acetylation of H3₁₋₂₁K14Me1 at K18 was the second most favoured site (with 80 fold lower specific activity than unmodified H3₁₋₂₁). Specific activity could not be calculated for HAT-BRD with H3₁₋₂₁ due to issues with assay wells in the linear range. However, at 180 min higher conversion for H3₁₋₂₁ at K14 was observed compared to H3₁₋₂₁K14Me1 at K18, showing that HAT-BRD is also more active towards unmodified H3₁₋₂₁ at K14. Specific activity of PCAF HAT and HAT-BRD at the K9 site of H3₁₋₂₁, H3₁₋₂₁K14Me1, H3₁₋₂₁K14Me2, H3₁₋₂₁K14Me3 and H3₁₋₂₁K14Ac was significantly lower than conversion of H3₁₋₂₁K14Me1 at K18. For all peptides, acetylation at K9 proceeded at a similar rate, with PCAF HAT showing higher specific activity levels compared to HAT-BRD for all peptides sites. These results indicated that K14 on H3₁₋₂₁ followed by K18 on H3₁₋₂₁KMe1 are more favourable sites for PCAF HAT/HAT-BRD acetylation than K9 on all peptides tested.

a)

Substrate peptide	Ac #	Site of Ac	Specific activity $\times 10^{-2}$ (min^{-1})		Conversion at 180 min (%)	
			HAT	HAT-BRD	HAT	HAT-BRD
H3 ₁₋₂₁	1st	K14	4800 $\pm$ 68	N/A*	100	100
	2nd	K9	4.0 $\pm$ 1.2	N/A*	25	16
H3 ₁₋₂₁ K14Me	1st	K18	59 $\pm$ 2.5	28 $\pm$ 2.0	100	90
	2nd	K9	3.6 $\pm$ 0.13	1.0 $\pm$ 0.062	34	17
H3 ₁₋₂₁ K14Me2	1st	K9	4.6 $\pm$ 0.10	1.7 $\pm$ 0.078	34	17
H3 ₁₋₂₁ K14Me3	1st	K9	5.2 $\pm$ 0.10	1.9 $\pm$ 0.073	40	22
H3 ₁₋₂₁ K14Ac	1st	K9	5.7 $\pm$ 1.2	0.88 $\pm$ 0.058	35	13

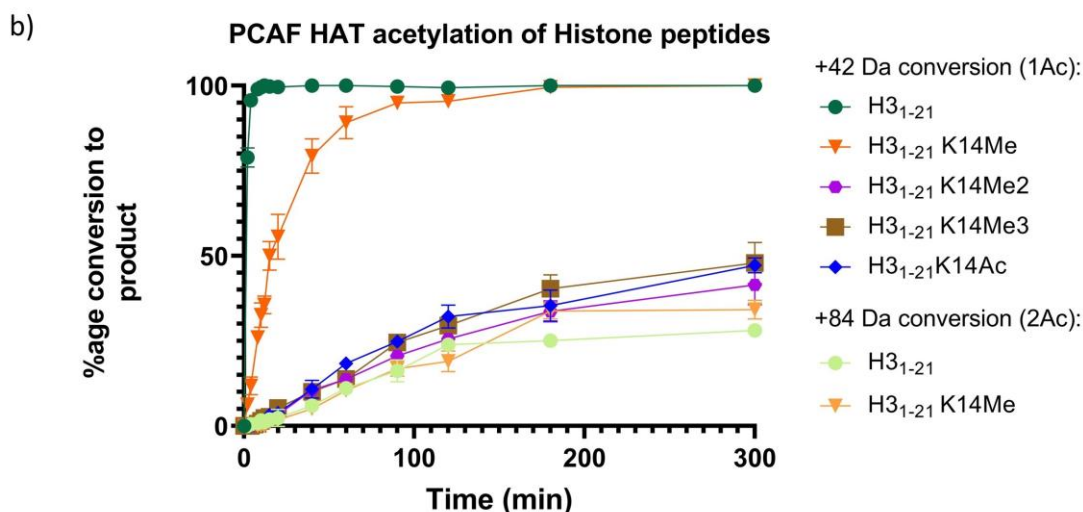


Figure 4.6: Time resolved MALDI-TOF-MS assay of PCAF with a panel of H3₁₋₂₁ peptides with K14 PTMs.

Assay conditions were PCAF enzyme (1 μM), AcCoA (100 μM), peptide (20 μM), buffer (50 mM HEPES, 50 mM NaCl) incubation at 37°C for 3h. **a)** The calculated specific activity (min^{-1}) with SD over three experimental replicates for peptide conversion with PCAF HAT and HAT-BRD is shown alongside conversion of peptide at 180 min. * = insufficient data points in linear range to calculate specific activity. Where two acetylation events were observed for a peptide, specific activity and conversion is given for the 1st and 2nd acetylation events. **b)** Plotted time course assay showing percentage conversion to product peptide $\pm$ SD for three experimental replicates vs. time (min). For H3₁₋₂₁ and H3₁₋₂₁K14Me1, enzyme conversion for two acetylation events was measured (+42 Da shift for 1st and +84 Da shift for 2nd acetylation). For the other peptides, only one acetylation event was observed.

4.3.2.2 Exploration of HIF1 α peptides as substrates for PCAF HAT and HAT-BRD using MALDI-TOF-MS

PCAF HAT and HAT-BRD were tested for activity with the synthesised HIF1 α peptides:

HIF1 α ₅₂₂₋₅₄₂, HIF1 α ₆₆₄₋₆₈₄ and HIF1 α ₇₀₀₋₇₂₁. The MALDI-TOF-MS assay was carried out

using an increased concentration of enzyme (4 μM) and incubation time (40 min)

compared to MALDI-TOF-MS assays described previously, to allow for conversion of

peptides less well suited to acetylation by PCAF than H3₁₋₂₁ peptides. Under these conditions, PCAF HAT catalyses 100% conversion of H3₁₋₂₁ peptides to a single acetylation product with a further 84% conversion to double acetylation product. However, no enzymatic acetylation by PCAF HAT was observed with any of the HIF1 α peptides tested under these conditions (Figure 4.7). PCAF HAT-BRD also did not show any acetylation activity against the HIF1 α peptides tested (results not shown).

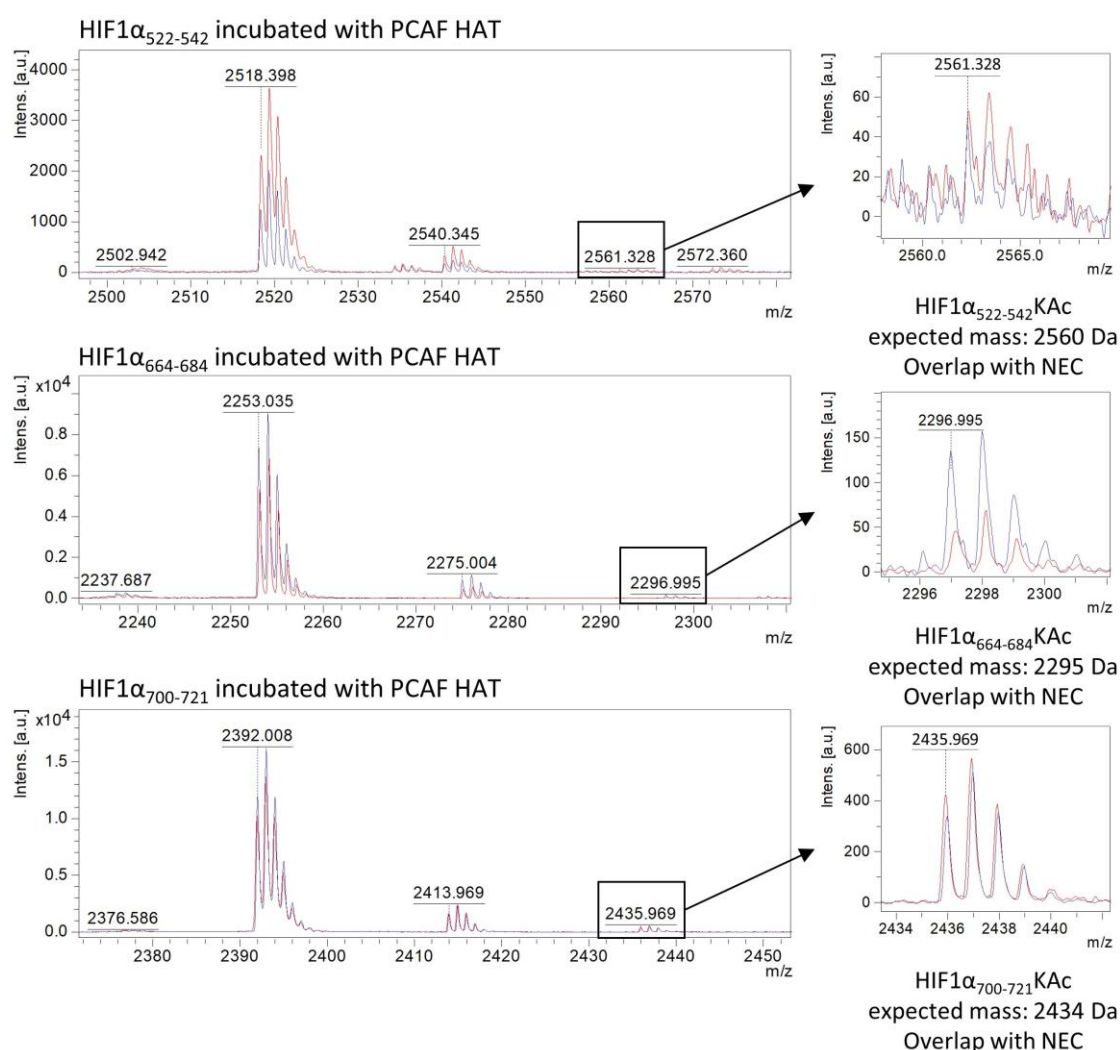


Figure 4.7: Exploring PCAF HAT enzymatic activity for HIF1 α peptides.

MALDI-TOF-MS spectra for HIF1 α ₅₂₂₋₅₄₂, HIF1 α ₆₆₄₋₆₈₄ and HIF1 α ₇₀₀₋₇₂₁ incubated with PCAF HAT (blue) overlaid with a NEC (red). The three HIF1 α peptides tested did not show conversion to acetylated products, as peaks at +42 Da for each peptide were impurities found in the NEC. Assay conditions were PCAF HAT (4 μ M), AcCoA (50 μ M), peptide (5 μ M), buffer (50 mM HEPES, 50 mM NaCl), incubated at 37°C for 40 min. Equivalent results were obtained for PCAF HAT-BRD (not shown).

4.4 Characterisation of SIRT1 substrate selectivity

4.4.1 Exploration of HIF1 α peptides as substrates for SIRT1 using MALDI-TOF-MS

SIRT1 was tested for deacetylation activity of HIF1 α ₅₂₂₋₅₄₂K523Ac, HIF1 α ₆₆₄₋₆₈₄K674Ac and HIF1 α ₇₀₀₋₇₂₁K709Ac using MALDI-TOF-MS, with deacetylation of all three peptides observed (–42 Da shift for each peptide, Figure 4.8). To compare the efficiency of SIRT1 deacetylation of the HIF1 α peptides with that of H3₁₋₂₁ peptides, comparative time courses were run monitoring substrate to product conversion by SIRT1 (0.1 μ M) over 12 minutes with incubation at 37°C (Figure 4.9b). H3₁₋₂₁K9Ac, H3₁₋₂₁K14Ac and H3₁₋₂₁K9AcK14Ac were tested alongside HIF1 α ₅₂₂₋₅₄₂K523Ac, HIF1 α ₆₆₄₋₆₈₄K674Ac and HIF1 α ₇₀₀₋₇₂₁K709Ac.

Specific activity (min^{-1}) of SIRT1 was measured for each peptide substrate (Figure 4.9). SIRT1 showed the highest specific activity for H3₁₋₂₁K9Ac, which is consistent with reports that H3K9Ac is the preferred site for SIRT1 activity on H3.^{84,184} SIRT1 showed similar specific activity for H3₁₋₂₁K14Ac and H3₁₋₂₁K9AcK14Ac peptides (Figure 4.9a,b). In this assay, only a single deacetylation event was observed for all peptides including H3₁₋₂₁K9AcK14Ac. No distinction in the MALDI-TOF-MS spectra could be made between deacetylation of the K9Ac and K14Ac sites of H3₁₋₂₁K9AcK14Ac, and the measured deacetylation product was therefore likely to be a mix of the two deacetylated species (*i.e.* either H3₁₋₂₁K9Ac or H3₁₋₂₁K14Ac). The lower rate of H3₁₋₂₁K9AcK14Ac deacetylation compared to H3₁₋₂₁K9Ac deacetylation can be due to an alteration in enzyme binding modes linked to K14Ac. If a large portion of the enzyme is bound to the K14Ac site with

a slow reaction rate, less enzyme is available for deacetylation at the K9Ac position. Alternatively, K14Ac might affect the binding mode of SIRT1, making K9Ac deacetylation less favourable.

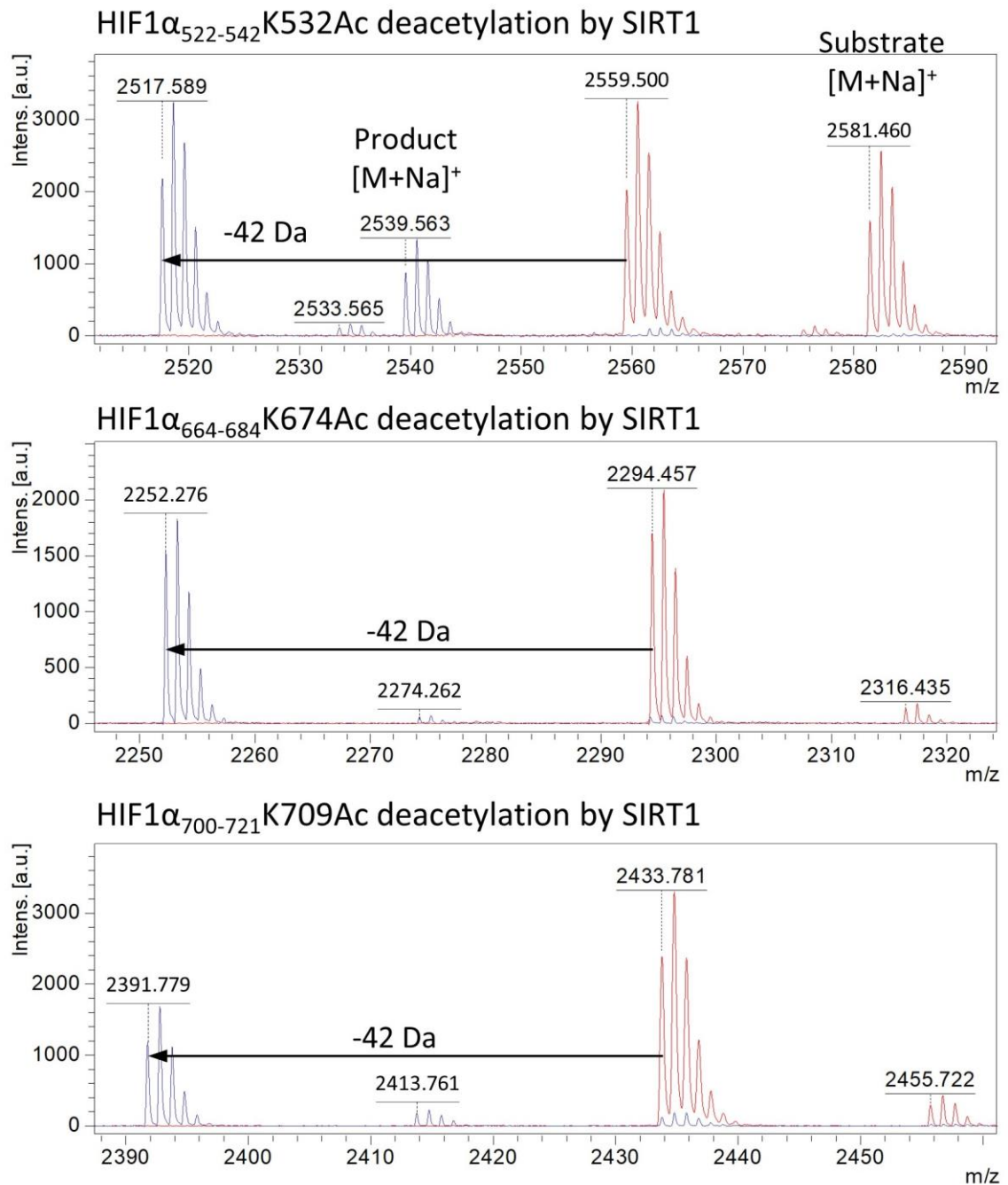


Figure 4.8: MALDI-TOF-MS spectra showing SIRT1 deacetylation of HIF1 α peptides. MALDI-TOF-MS spectra showing deacetylation by SIRT1 of HIF1 α ₅₂₂₋₅₄₂K523Ac, HIF1 α ₆₆₄₋₆₈₄K674Ac and HIF1 α ₇₀₀₋₇₂₁K709Ac substrate peptides (-42 Da mass shift). A NEC spectrum (red) is overlaid with product spectra (blue). Assay conditions were SIRT1 (0.75 μ M), NAD⁺ (50 μ M), peptide (10 μ M) incubated at 37°C for 15 min.

In the previously described rapifleX MALDI-TOF-MS HTS (Chapter 3, Figure 3.8), all acetylation sites on H1, H2, H3 and H4 acetylated peptides were equally deacetylated, including for peptides with multiple acetylations (1-5 Ac sites). The rapifleX MALDI-TOF-MS HTS was an end point assay run at a higher enzyme concentration (2 μ M vs. 0.1 μ M in this assay) over a longer period of time (1 h, vs. 12 min in this assay). Under such conditions, it was not possible to distinguish rates of deacetylation between peptides, and no site selectivity could be determined. Here, using a shorter assay with lower enzyme concentrations, SIRT1 is observed to have some site selectivity, as shown by the different deacetylation rates of H3₁₋₂₁ acetylated peptides.

SIRT1 catalysed up to ~30% deacetylation of HIF1 α ₅₂₂₋₅₄₂K523Ac compared with <6% deacetylation of HIF1 α ₆₆₄₋₆₈₄K674Ac and HIF1 α ₇₀₀₋₇₂₁K709Ac (Figure 4.9a,b). The higher specific activity seen with HIF1 α ₅₂₂₋₅₄₂K523Ac may be in part due to increased error in measurement due to the high proportion of [M+Na]⁺ ion adduct observed in the MALDI-TOF-MS (Figure 4.9c). Because both acetylated and deacetylated versions of HIF1 α ₅₂₂₋₅₄₂K523Ac showed high levels of [M+Na]⁺ ions adduct, not seen with any of the other peptides, this was taken into account when measuring enzyme activity, by averaging the measured conversion for both species (Equation 1).

$$\% \text{ product} = \frac{\text{substrate int} + \text{substrate } [M + Na]^+ \text{ int}}{\text{substrate int} + \text{substrate } [M + Na]^+ \text{ int} + \text{product int} + \text{product } [M + Na]^+ \text{ int}} \times 100$$

Equation 1: Calculation of product formation in the MALDI-TOF-MS assay taking into account [M+Na]⁺ peptide ion adducts.

Int: intensity of peak at defined mass. This was only used for HIF1 α ₅₂₂₋₅₄₂K523Ac, as other peptides showed a negligible amount of [M+Na]⁺ ion formation in the MALDI-TOF-MS spectra (<5%).

a)

Substrate peptide	Product	Specific activity (min^{-1})
H3 ₁₋₂₁ K9Ac	- 1Ac	11.03 ± 0.82
H3 ₁₋₂₁ K14Ac	- 1Ac	2.72 ± 0.18
H3 ₁₋₂₁ K9AcK14Ac	- 1Ac	2.93 ± 0.21
HIF1 α ₅₂₂₋₅₄₂ K532Ac	- 1Ac	2.22 ± 0.13
HIF1 α ₆₆₄₋₆₈₄ K674Ac	- 1Ac	0.24 ± 0.06
HIF1 α ₇₀₀₋₇₂₀ K709Ac	- 1Ac	0.39 ± 0.04

b) **SIRT1 substrate scope**

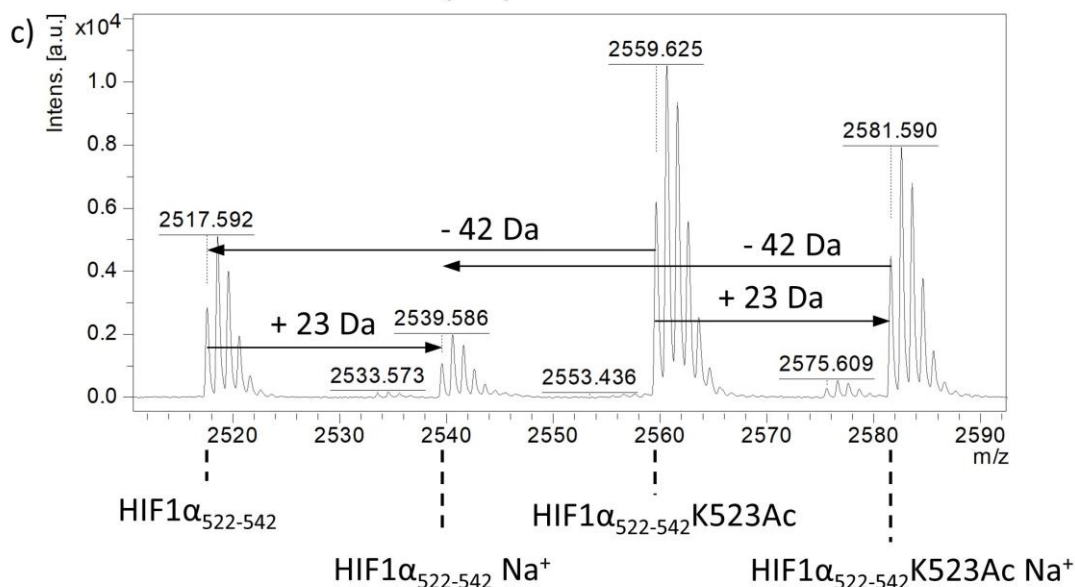
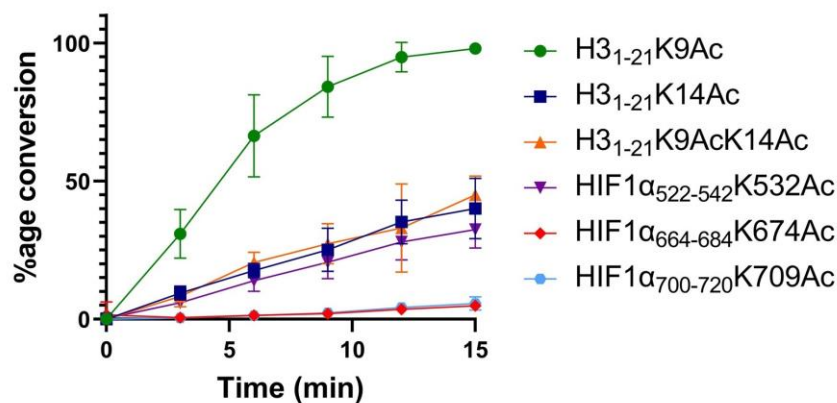


Figure 4.9: Time resolved MALDI-TOF-MS assay of SIRT1 with a H3₁₋₂₁ and HIF1 α substrate panel.

Assay conditions were SIRT1 (0.75 μM), NAD⁺ (50 μM), peptide (10 μM) incubated at 37°C for 15 min, carried out with three experimental replicates. **a)** Calculated specific activity $\pm$ SD for three experimental replicates (min^{-1}) for each peptide, calculated in the linear range of the time course assay. **b)** Plotted time course assay showing average %age conversion $\pm$ SD to singly deacetylated product peptide vs. time (min) for each peptide. **c)** Example of a MALDI-TOF-MS spectrum for HIF1 α ₅₂₂₋₅₄₂K523Ac deacetylation with SIRT1, showing a deacetylation (-42 Da) event for both the $[M+H]^+$ substrate and the $[M+Na]^+$ sodium adduct (+23 Da) of the substrate.

4.5 Discussion

4.5.1 Substrate selectivity of PCAF for H3₁₋₂₁ peptides with PTM

4.5.1.1 Effects of K14 methylation on K18 acetylation for H3₁₋₂₁ peptides

Using MALDI-TOF/TOF-MS, the sites of acetylation on the methylated H3₁₋₂₁ peptides were identified (Figure 4.3, Figure 4.4, Figure 4.5). For H3₁₋₂₁K14Me1, the first acetylation site was K18 followed by K9. For H3₁₋₂₁K14Ac (control), H3₁₋₂₁K14Me2 and H3₁₋₂₁K14Me3 only K9 is acetylated. K18 acetylation of H3₁₋₂₁ by PCAF can be observed, but only with prolonged reaction times (5h), whereas H3₁₋₂₁K14 acetylation was rapid (2 min for full conversion) under the same conditions (Chapter 2). These results indicate that K14Me1 has an activating effect on the acetylation activity of PCAF towards K18 of H3₁₋₂₁ peptides. In addition, PCAF had significantly higher specific activity towards K18 on H3₁₋₂₁K14Me1 than towards K9 on H3₁₋₂₁K14Me1/2/3/Ac.

H3K14 methylation (Me1, Me2 and Me3) has been reported in eukaryotic cells in varied settings, making it a biologically relevant epigenetic PTM, as discussed in Chapter 3 (Section 3.6.2.4). H3K18Ac is an epigenetic PTM which is associated with cancer progression in hepatocellular, thyroid, lung, pancreatic, colon prostate and breast cancer.³⁹⁸ In yeast, increase in γ GCN5 and H3K18Ac were linked and shown to have a large effect on transcription levels in stress adaptation situations.³⁹⁹ In lung cancer, GCN5 and H3K18Ac levels were also linked with simultaneous decrease of H3K18Ac and GCN5 promoting cancer development.⁴⁰⁰ H3K18Ac was identified as an acetylation site for GCN5 on H3 protein, with a quantitative MS approach determining the preference of GCN5 for H3 lysine sites to be K14>K9≈K23>K18≈K27<K36.³⁴⁰ A study of inflammation

in the development of renal injury linked PCAF and H3K18Ac levels, with high levels of both linked to increased expression of inflammatory molecules.³³⁹ These studies show that PCAF/GCN5 acetylation of H3K18 is a biologically relevant PTM, though less common than H3K14 or H3K9 PCAF acetylation.

For GCN5 taking part in the transcriptional activator alteration/deficiency in activation (ADA) subcomplex, it was reported that H3K14 acetylation was a prerequisite for H3K18 acetylation. Acetylation of H3K18 was shown to decrease when the GCN5 bromodomain was mutated to impair KAc binding. A parallel experiment in which H3K14R was incubated with unmodified GCN5 showed a decrease in H3K18 acetylation. These results indicated that acetylation of H3K18 by GCN5 in the ADA subcomplex was dependent on bromodomain recognition of a H3K14Ac PTM.¹²² In the rapifleX screen results, increased PCAF activity linked to H3K14Me1 was observed for both the HAT and HAT-BRD constructs, suggesting that bromodomain binding did not intervene in this effect (Section 3.3.4.4, Table 3.11). The activating effect observed must therefore be due to the HAT domain interacting with H3K14Me1.

It was reported that histone H3S10 phosphorylation can have a large effect on the binding mode of H3₅₋₂₃ histone peptides to γ GCN5.¹⁰⁵ In a similar manner, K14Me1 might also alter the binding mode of PCAF HAT to H3₁₋₂₁ peptides, making K18 a more favourable acetylation site when K14 is mono-methylated compared with the control peptides, H3₁₋₂₁ and H3₁₋₂₁K14Ac and di-/tri-methylated peptides H3₁₋₂₁K14Me2 and H3₁₋₂₁K14Me3. A specific binding configuration between H3₁₋₂₁K14Me1 and PCAF HAT/HAT-BRD might therefore exist. The HAT domain of hGCN5 or γ GCN5 (PCAF

homologue) is usually used for production of crystal structures, due to the difficulty in obtaining crystal structures of PCAF HAT in complex with substrate peptides. One avenue of further study would therefore be to characterise the enzyme activity of GCN5 for H3₁₋₂₁K14Me1/2/3, and if similar activity to that of PCAF HAT was observed, obtain crystal structures to help elucidate the binding modes involved.

Because the MALDI-TOF-MS assays with histone peptides are a simplified model, assays using histone proteins and nucleosomes with K14Me1/2/3 modifications would be the next step in validating the activating effect of the K14Me1 PTM on PCAF activity at H3K18. Going further, a link between H3K14Me1 and H3K18Ac PTM might exist in a biological setting. As both H3K14Me1 and H3K18Ac PTM have been linked to an increased cell proliferation in cancers,^{384,398} this might be worth further investigation. However, the low abundance of H3K14 methylation (as compared to H3K14 acetylation) might make identification of such a link difficult. It has been shown that certain circumstances such as *Legionella pneumophila* infection can lead to a large increase in H3K14 methylation.^{49,379} In addition, the distribution of H3K14Me3 was found to vary depending on cell type, indicating levels of H3K14 methylation might vary between tissue types.³⁷⁹ Specific circumstances could therefore lead to an increase in endogenous cellular H3K14 methylation levels with a relevant biological role. The identification of HDACs (KDM4A/C) active against H3K14Me3 and of H3K14Me1/2/3 reader proteins (UGRF1), coupled with the dynamic interaction reported between H3K14Me1/2/3 levels and H3K14 acetylation by KAT7 show that though levels of H3K14 methylation are low, this PTM has an active role in epigenetic cross-talk which must be further investigated.^{379,384}

4.5.1.2 Effects of K14 methylation on K9 acetylation of H3₁₋₂₁ peptides

Whether H3K9 and H3K14 acetylation by PCAF/GCN5 are co-dependent is not clearly understood. For GCN5, it was theorised that H3K9 acetylation was dependent on GCN5 interaction with H3K14Ac (discussed in Section 2.5.2.1).³⁴⁰ PCAF showed similar levels of specific activity towards K9 acetylation of H3₁₋₂₁, H3₁₋₂₁K14Ac, H3₁₋₂₁K14Me1, H3₁₋₂₁K14Me2 and H3₁₋₂₁K14Me3, whilst PCAF had much higher specific activity for K14 of H3₁₋₂₁ and K18 of H3₁₋₂₁K14Me1(Figure 4.6a). If K9 acetylation by PCAF were dependent on having a previous K14Ac modification, PCAF should have significantly different specific activity towards K9 of the H3₁₋₂₁K14 methylated peptides compared to K9 of H3₁₋₂₁ (2Ac) and H3₁₋₂₁K14Ac. Testing the activity of H3₁₋₂₁K14A or H3₁₋₂₁K14R mutant peptides with PCAF would give a more definitive answer as to the interdependence of H3K9 and K3K14 acetylation.

4.5.2 **Implications of measured SIRT1 selectivity for H3₁₋₂₁ peptides for the rapifleX MALDI-TOF-MS HTS**

In the rapifleX MALDI-TOF-MS HTS of histone peptide libraries with combinatorial PTMs, no selectivity was observed for the SIRT1 enzyme which deacetylated up to 5 acetylation PTMs on H3, H4, H2a and H2b peptides. The rapifleX MALDI-TOF-MS HTS was based on an end point assay, in which it is impossible to distinguish selectivity of deacetylation between peptides which have been fully deacetylated. In this chapter, conditions were adapted to monitor deacetylation rates for H3₁₋₂₁ and HIF1 α acetylated peptides. SIRT1 showed a higher rate of deacetylation towards H3K9Ac compared with H3K14Ac and H3K9AcK14Ac, with very low deacetylation rate observed for HIF1 α ₆₆₄₋₆₈₄K674Ac and

HIF1 α ₇₀₀₋₇₂₁K709Ac. This shows that under the right conditions in *in vitro* assays, SIRT1 has substrate selectivity. The conditions for the HTS were chosen based on results from a first screen with PCAF HAT and HAT-BRD, in which PCAF showed substrate selectivity between histone types (H3, H4, H2a and H2b) and between different portions of these histones. This suggests that SIRT1 has less inherent selectivity than PCAF, with respect to amino acid sequences. However, SIRT1 had much higher specific activity towards H3₁₋₂₁ peptides than PCAF, which might instead indicate that the conditions of the HTS were not well chosen.

4.5.3 Activity of PCAF and SIRT1 for HIF1 α peptides and implications for development of a MALDI-TOF-MS screen

4.5.3.1 Choice of HIF1 α peptides for testing as models of non-histone peptide substrates of PCAF and SIRT1

Although there are some examples of peptides acting as models for endogenous non-histone protein substrates of HATs and HDACs (HIF1 α ,¹²⁵ HIV,¹⁵⁸ p53,^{105,126} for PCAF), using peptides as models for proteins may not always successfully replicate native protein-substrate binding. Histone tails are naturally disordered protein regions, as shown by a crystal structure of the nucleosome core.^{327,401} Because of this, linear peptides, which generally have low levels of secondary structure, are a good model for enzymatic activity of PCAF and SIRT1 on histone tails.

Protein disorder has been identified as an important factor in the site selectivity of phosphorylation marks,⁴⁰² and has also been shown to affect regulation of acetylation PTMs, with protein disorder being a factor in CBP acetylation of p53, NF- κ B, STAT

proteins and HIF1 α .⁴⁰³ In addition, it was reported that PTM crosstalk is often found in sections of proteins which are disordered and have a high density of PTMs.³⁵³ Conversely, PTMs have been reported to play a role in promoting protein folding.^{353,402,404} There is some evidence to suggest that PCAF and SIRT1 target disordered regions of non-histone proteins. It is estimated that 30-40% of PTMs on proteins which interact with SIRT enzymes are located in disordered regions of proteins.³⁵³ p53 and NF- κ B are two examples of proteins with intrinsically disordered regions which are substrates for both SIRT1 and PCAF. In addition PCAF was reported to catalyse acetylation of p53 peptides, which have little secondary structure.^{105,403} Linear peptides might therefore be suited to screening non-histone protein regions of interest for substrate selectivity of PCAF and SIRT1. HIF1 α was chosen to test this theory, for the reasons listed below.

HIF1 α has a mix of ordered domains and sections for which no structure has been identified.⁴⁰⁵ Its C-terminal CAD domain is reported to be disordered (as shown by circular dichroism) when not bound to CBP.⁴⁰⁶ p300/CBP is reported to acetylate HIF1 α at K709 whilst PCAF and SIRT1 are reported to affect K674, and both of these residues are located outside defined domains of HIF1 α , with no defined structures associated with these regions. The third reported acetylation site on HIF1 α , K532, is located at the N-terminus of the NAD domain (res 531-575).⁴⁰⁵ The N-terminus of this domain might be a flexible linker region, or could be part of a secondary structure element of the NAD domain. It is therefore possible that the K532, K674 and K709 acetylation sites of HIF1 α are all in disordered regions of the protein, and linear peptides could successfully replicate PPI interactions between PCAF, SIRT1 and HIF1 α .

4.5.3.2 Results of the MALDI-TOF-MS testing of HIF1 α peptides as PCAF and SIRT1 substrates

The results obtained with HIF1 α peptides in the MALDI-TOF-MS assay did not correspond to reported activity of PCAF and SIRT1 proteins on full length HIF1 α protein (Table 4.2), unlike results obtained with H3₁₋₂₁ peptides which corresponded to reported activity of the full-length proteins (Chapter 2).

Peptide Tested	HIF1 α residue of interest	Acetylated by PCAF		Deacetylated by SIRT1	
		Reported	MALDI assay	Reported	MALDI assay
HIF1 α ₅₂₂₋₅₄₂	K523	No	No		
HIF1 α ₆₆₄₋₆₈₄	K674	Yes	No		
HIF1 α ₇₀₀₋₇₂₁	K709	No	No		
HIF1 α ₅₂₂₋₅₄₂ K523Ac	K523Ac			No	Yes
HIF1 α ₆₆₄₋₆₈₄ K674Ac	K674Ac			Yes	Yes
HIF1 α ₇₀₀₋₇₂₁ K709Ac	K709Ac			Unknown	Yes

Table 4.2: Summary of results obtained for HIF1 α peptides with PCAF and SIRT1 in the MALDI-TOF-MS assay compared with literature findings.

HIF1 α K674 is the only site which is reported to be acetylated by PCAF using full length HIF1 α protein. If the MALDI-TOF-MS assay was to successfully replicate biological results, the HIF1 α ₆₆₄₋₆₈₄ should be acetylated and the HIF1 α ₅₂₂₋₅₄₂ and HIF1 α ₇₀₀₋₇₂₁ peptides should not be acetylated by PCAF HAT/HAT-BRD. PCAF HAT and HAT-BRD were tested for acetylation activity of HIF1 α ₅₂₂₋₅₄₂, HIF1 α ₆₆₄₋₆₈₄ and HIF1 α ₇₀₀₋₇₂₁ peptides using MALDI-TOF-MS, without any acetylation being detected.

HIF1 α K674Ac is the only acetylation site on full-length HIF1 α protein that has been reported to be deacetylated by SIRT1 in cellular experiments. Using HIF1 α mutants with pulldown assays targeting KAc PTMs in cells transfected with PCAF and/or SIRT1, both unmodified HIF α protein and a HIF1 α K523R mutant showed high levels of acetylation

when a SIRT1 inhibitor nicotinamide (NAM) was used. Mutation of HIF1 α to HIF1 α K674R gave very low levels of protein acetylation.¹²⁵ These results points towards SIRT1 being inactive at HIF1 α K523KAc, whilst deacetylating HIF1 α K674KAc. Whilst it was established that p300 catalyses the acetylation of HIF1 α K709 and p300 interacts with SIRT1, SIRT1 deacetylation of HIF1 α K709Ac has not been tested.³⁹⁵ The expected results for the MALDI-TOF-MS assay of SIRT1 with synthesised HIF1 α peptides were therefore for SIRT1 to be inactive against the HIF1 α ₅₂₂₋₅₄₂K523Ac peptide, active against the HIF1 α ₆₆₄₋₆₈₄K674Ac peptide, and no expected results for HIF1 α ₇₀₀₋₇₂₁K709Ac. However, deacetylation by SIRT1 was observed for all three peptides.

SIRT 1 deacetylation of HIF1 α ₆₆₄₋₆₈₄K674Ac and HIF1 α ₇₀₀₋₇₂₁K709Ac was slower than for HIF1 α ₅₂₂₋₅₄₂K523Ac. Specific activity of SIRT1 for the latter was comparable to that of H3₁₋₂₁K14Ac deacetylation. Increased deacetylation at HIF1 α ₅₂₂₋₅₄₂K523Ac, which is not the expected deacetylation site for SIRT1, was partly attributed to errors in measuring of enzyme activity using MALDI-TOF-MS, linked to Na⁺ ion formation for the HIF1 α ₅₂₂₋₅₄₂K523Ac peptide. However, as PTMs on proteins are reported to lead to an increase in secondary protein structure formation, and as K523 is located within a reported domain of HIF1 α (NAD), it is possible that HIF1 α ₅₂₂₋₅₄₂K523Ac has some secondary structure, which could lead to better binding to SIRT1 and increased deacetylation activity at this site. Following this logic, unmodified HIF1 α ₅₂₂₋₅₄₂, HIF1 α ₆₆₄₋₆₈₄ and HIF1 α ₇₀₀₋₇₂₁ peptides might lack secondary structure, accounting for the lack of acetylation activity observed with PCAF. This would entail that PCAF and SIRT1 binding to the HIF1 α protein at K674 is mediated by secondary structure elements, or that binding at these locations is facilitated by interactions with portions of HIF1 α , not

covered in the 21mer peptides. It has previously been reported that PCAF has increased specificity towards the 19 residue peptide H3₅₋₂₃ compared with the 11 residue peptide H3₉₋₁₉.²⁹⁰ This indicates that the boundaries used for the peptides tested can be important, and might affect ability of the enzymes to bind to and modify HIF1 α peptides. To test these theories, peptides with alternative boundaries could be tested, and circular dichroism could be used to assess secondary structure elements in the HIF1 α peptides used.

4.6 Future work

None of the results obtained for H3₁₋₂₁ peptide acetylation by PCAF clearly show whether H3₁₋₂₁K9 acetylation is independent of H3₁₋₂₁K14 acetylation. To test whether K14 and K9 acetylation are truly independent, H3₁₋₂₁K14R and H3₁₋₂₁K9R mutants should be synthesised and tested. Lysine is substituted by arginine because it is the closest analogue to lysine and would cause minimal changes to peptide structure and charge. H3₁₋₂₁K14Me1/2/3 all showed similar rates of acetylation by PCAF HAT and HAT-BRD at K9, which might indicate that H3K9 acetylation is indeed independent of K14Ac. However, it may be that binding to H3K14 (independent of PTMs) is necessary for H3K9 acetylation. A H3₁₋₂₁ Δ K14A peptide should have a very different binding profile to H3₁₋₂₁K14 and could be tested to complement results for H3₁₋₂₁ Δ K14R.

#	Protein	Residues	PTM	Sequence	Purpose for testing
1	H3	1-21	ΔK14R	ARTKQTARKSTGG <u>R</u> APRKQLA	Identify whether K14Ac is a pre-requisite for K9 acetylation
2	H3	1-21	ΔK14A	ARTKQTARKSTGG <u>A</u> APRKQLA	Identify whether binding to H3K14 site is needed for K9 acetylation
3	H3	1-21	ΔK9R	ARTKQTAR <u>R</u> STGGKAPRKQLA	Identify whether K9Ac affects K14 acetylation

Table 4.3: Peptides to be synthesised for further characterisation of PCAF activity.

4.6.1 P53 peptides as alternative to HIF1α for non-histone peptide substrates of PCAF and SIRT1

HIF1α peptides used in this study were not successful in modelling the interactions between PCAF and the HIF1α protein. This may be due to unknown secondary structure elements intervening in the PPI, or binding of the enzymes to sections of HIF1α not covered in the short peptide sequences. As PCAF activity against HIF1α was determined in a cellular context,^{125,394,395} it may also be that PCAF selectivity for HIF1α is mediated by protein complex formation or other factors which are not replicated in *in vitro* assays. If secondary structure is important in binding, cyclic peptides might be more effective models than linear peptides and could be synthesised for testing. Deacetylation activity by SIRT1 was observed with acetylated HIF1α peptides. However, the peptide site on HIF1α which showed the highest levels of deacetylation was not consistent with biological results, unlike results obtained with MALDI-TOF-MS screening of PCAF and SIRT1 with H3₁₋₂₁ peptides (Figure 4.9). HIF1α peptides are therefore unlikely to be a good model for PPI interactions between PCAF/SIRT1 and the HIF1α protein.

To test whether the rapifleX MALDI-TOF-MS HTS might be applied to non-histone peptide libraries, p53 peptides might be a successful alternative to HIF1α. p53 peptides have previously been reported as active in *in vitro* assays with PCAF/GCN5 and have

been crystallised with the proteins.^{50,105,126} p53 is also subject to multiple PTMs in cells, which are found clustered in its C-terminal domains (tetramerization domain and C-terminal regulatory domain) with multiple acetylation, phosphorylation, methylation and ubiquitination PTM sites between residues 305 and 395.⁴⁰⁷ If p53 peptides (unmodified p53₃₁₀₋₃₃₀ and acetylated p53₃₇₂₋₃₉₂ K382Ac) were shown to be substrates of PCAF and SIRT1 in independent MALDI-TOF-MS assays, a library of peptides covering p53 residues 300-395 could be synthesised using a high throughput parallel peptide synthesis method and screened with PCAF and SIRT1 using rapifleX MALDI-TOF-MS HTS, exploring the effects of crosstalk between PTMs found on p53.

4.6.1.1 Enzyme kinetics

Time courses have been used to determine specific activity of PCAF HAT/HAT-BRD and SIRT1 for H3₁₋₂₁ and HIF1 α substrate peptides, allowing for comparison of enzyme activity at different sites on each peptide. Kinetic studies (K_M and K_{cat} determination) of the enzymes and peptides would provide more comparative data, allowing for a better understanding of enzyme activity for the different substrates, and the nature of PTM crosstalk identified (K18 acetylation of H3₁₋₂₁K14Me1 by PCAF). Initial work has been carried out to use MALDI-TOF-MS to determine K_M for the peptides with PCAF. Issues were encountered with data reproducibility, due to the high concentrations of peptide required (K_M of PCAF for H3₁₋₂₁ is reported as $532 \pm 81 \mu M$, determined using a pyruvate dehydrogenase coupled assay)¹⁰⁶. Use of an alternative assay (*e.g.* an enzyme coupled assay) may be necessary to obtain reliable kinetic data for PCAF and SIRT1 enzymes.

4.6.1.2 Biological relevance of PCAF and SIRT1 substrate selectivity

In this chapter, the H3₁₋₂₁K14Me1 substrate identified in the MALDI-TOF-MS HTS as having PTM crosstalk effects on the activity of PCAF HAT/HAT-BRD has been validated and characterised. This is the first step in demonstrating the usefulness of the HTS developed in identifying substrates for epigenetic enzymes and for mapping PTM crosstalk which affects them. Further work to investigate whether H3K14Me1 activation of PCAF acetylation at H3K18 is also observed in more complicated biological models (*e.g.* with full histone proteins and in cells) is the next step in linking the results observed in the screen to biological activity of the PCAF enzyme. This could also help understand the biological roles of H3K14Me1 and H3K18Ac PTMs. The methodology developed can be used in this manner to better understand effects of PTM crosstalk on epigenetic enzymes, a regulatory mechanism of epigenetics which is not well characterised.

4.7 Conclusion

The rapifleX MALDI-TOF-MS HTS screen developed has led to the discovery of interesting PTM crosstalk interactions on histone peptides affecting PCAF enzymatic acetylation activity. Mapping activity of PCAF/SIRT1 against peptides is a model for investigating PPI interactions between these proteins and their cognate substrates. The first step in validating results of the HTS is to independently reproduce the screen results using re-synthesised, unlabelled high-quality peptides in time resolved MALDI-TOF-MS assays which allow for measurement of specific activity of the enzymes for different substrates. This method was used to validate HTS results observed between PCAF and H3₁₋₂₁K14Me1. Mono-methylation of H3₁₋₂₁ peptide at K14 activates PCAF acetylation of

the peptide at H3K18, as shown by MALDI-TOF/TOF-MS. Further validation of this effect on histone proteins and nucleosomes would be the next step in exploring the biological relevance of the H3K14Me1 PTM and its effect on acetylation.

Histone peptides are widely used as a model for histone tail PPIs, and were therefore ideal as library constituents to test in the development of the rapifleX MALDI-TOF-MS HTS. As PCAF and SIRT1 both have many non-histone cognate substrates which are known to bear varied PTMs, expanding this method to screening non-histone peptide libraries would increase the scope of the HTS method developed. HIF1 α peptides were chosen as substrates to test against PCAF and SIRT1 on MALDI-TOF-MS assays due to the overlap in activity between these two enzymes for the HIF1 α protein. PCAF showed no acetylation activity towards HIF1 α peptides, whilst SIRT1 showed activity against HIF1 α peptides bearing acetylations at either K523, K674 or K709. The results observed with HIF1 α peptides did not match reported biological activity, indicating that the HIF1 α peptides used were not suitable for screening with PCAF and SIRT1. More targets should be screened to determine whether the rapifleX MALDI-TOF-MS HTS can be used to model PCAF and SIRT1 binding to non-histone substrates, with p53 peptides being promising candidates for further tests.

Chapter 5 mRNA display screening of PCAF HAT/HAT-BRD for development of cyclic peptide chemical probes

5.1 Introduction

5.1.1 Background for use of cyclic peptides as chemical probes and inhibitors

Cyclic peptides are an emerging class of molecules, attracting significant interest as useful modalities for developing selective and potent chemical probes against protein targets of interest.²⁸¹ They differ to classical small molecules in a number of ways: automated synthesis using solid phase peptide synthesis (SPPS) makes their synthesis relatively rapid and simple; they present a large surface area and can adopt 3-dimensional structures resembling proteins; using non-canonical amino acids (nCAA), a great diversity of structure and function can be achieved.^{408,409} These factors make it possible to synthesise large libraries of peptides with varied structures and chemical properties to determine binding affinity, structure activity relationship (SAR) and inhibitory activity towards selected targets.⁴¹⁰ The size of peptides presents an advantage over small molecules as there is a larger interaction surface area which can improve affinity and selectivity.⁴⁰⁹ This is particularly the case when targeting protein-protein interactions (PPIs) which cover a surface area of 1500-3000 Å² where small molecule to protein interactions cover an area of 300-1000 Å².⁴¹¹ Thus, cyclic peptides are an ideally suited modality for developing chemical probes against targets of interest.

There are aspects of cyclic peptides which are currently disadvantageous for their development as chemical probes. Low cell permeability, plasma stability impaired by

peptidases, low bioavailability and difficulties with computational prediction (*e.g.* chemical properties, binding modes, 3D structure) means design of cell-active probes is challenging. These issues are well documented, and significant work is being carried out to understand and overcome them.^{412–415} Some methods used to improve cell permeability and stability of peptides include cyclisation, *N*-methylation, intramolecular hydrogen bonds, use of a wide range of non-canonical amino acids (nCAA) and peptide stapling. These options can be used to widen the diversity of peptide libraries as well as to improve physiochemical properties.^{416–420}

Natural product cyclic peptide drugs such as oxytocin, vancomycin, cyclosporin and leuprolide were the first cyclic peptide chemical probes, used in biology and medicine.^{419,420} Synthetic cyclic peptides were then used, with peptides being stapled to adopt 3D structures mimicking portions of proteins involved in PPIs.^{421,422} This methodology is limited by the need for previous understanding of the PPIs involved. To widen the scope of cyclic peptides to study specific target biology, methods for identifying *de novo* cyclic peptides without previous knowledge of the target protein structure have been developed. Peptide screening platforms, based on the concept of HTS using small molecules, can now be used to produce a library of randomly generated peptides which are screened against any target of interest.^{281,423}

Access to large libraries of structurally and chemically diverse cyclic peptides which can be rapidly screened against any target, along with the advantages that peptides have over small molecules in targeting PPIs, make cyclic peptides a good option for

development of highly potent and selective chemical probes against protein targets of interest.

5.1.2 Screening methods for development of cyclic peptides chemical probes

An effective peptide screening method needs to be able to produce a large library of cyclic peptides and have an efficient process for identifying recovered hits.⁴²⁴ Producing peptides for screening can be carried out chemically using SPPS and split and mix methods,^{425,426} or biologically using the ribosome translation apparatus. Identifying peptides is most commonly achieved through sequencing an attached DNA tag or template. Sanger sequencing and next generation sequencing (NGS) make this a very quick, cost effective and reliable method of identifying peptide hits from screens.⁴²⁷ Using DNA tags, nCAA can be identified if the codons for canonical amino acids (CAA) are repurposed for selected nCAA. Five main peptide screening methods have been described which allow for screening of cyclic peptides. Two are chemically-based methods: DNA encoded libraries (DEL) and one bead one compound (OBOC). Three are biologically-based methods: Phage display, Split-intein circular ligation of peptides and proteins (SICLOPPS) and mRNA display.²⁸¹

Chemically synthesised DNA encoded peptide libraries have not been developed to the same extent as other methods described due to the difficulty of combining chemical peptide synthesis with DNA compatible solvents. Two methods are available for DEL in which a cyclic peptide is covalently linked to a DNA tag, with a DNA segment added to encode each amino acid in the peptide.^{428,429} OBOC screening is the only method described which does not rely on DNA encoding for hit identification. In this method,

peptides are chemically synthesised on bead using split and mix methods with SPPS and, after affinity capture, the hits are identified using LCMS or Edman degradation techniques. These peptide identification methods become difficult when dealing with nCAA and cyclic peptides.^{426,430–433}

Phage display was the first biological peptide screening method developed, for which George P. Smith and Sir Gregory P. Winter received one half of the Nobel prize in Chemistry in 2018.⁴³⁴ In this method cyclic peptides are displayed on phage particles with the encoding DNA encapsulated in the phage. Hits are identified using affinity capture, and phage amplification for enrichment or sequencing is achieved by infecting *E. Coli* cultures with the phage.⁴³⁵ SICLOPPS is a cellular method in which the peptides are encoded in plasmids used to transform cells (*E. Coli* or other).^{436,437} Each cell produces multiple copies of a unique cyclic peptide and an in-built cellular assay allows for selection of hits based on activity of the cyclic peptide.⁴³⁸ In mRNA display, the cyclic peptide is covalently bound to a hybrid DNA/RNA tag which can be amplified using PCR for enrichment or sequencing.^{439,440} The library is produced *in vitro* using purified or synthetic translation, transcription and reverse transcription enzymes, and the hits are identified by affinity capture.

mRNA display is the method chosen as it presents advantages over other screening methods. It allows access to a larger library pool than phage display or SICLOPPS as these methods are limited by the transformation efficiency of cells for library size. It is carried out fully *in vitro* allowing for a wide use of nCAA and cyclisation strategies which are not possible in cellular contexts (phage display and SICLOPPS). DEL and OBOC also allow for

use of nCAA, but cyclisation strategies are limited in DEL libraries and hit identification is difficult for OBOC as it does not use DNA tags. Finally, mRNA display uses small reaction volumes (<20 μ L), and one round of selection is carried out over 2/3 days making it a quicker screening method than the alternatives described.²⁸¹

5.1.3 Cyclic peptide mRNA display screen – background and methodology

mRNA display, initially used to identify antibody epitopes,⁴³⁹ has since been developed to display large libraries of linear or cyclic peptides (>10¹² library size) for screening against a variety of protein targets.⁴⁴⁰ Research in the Kawamura group focuses on cyclic peptide libraries, as these tend to have better physiochemical properties and binding affinities compared to linear peptides.

The mRNA library is designed using DNA primers with a variable region of determined length made using degenerate codons. A codon structure NNK is used where N = A, T, C or G and K = C or G. This encodes for all 20 CAA and one of three stop codons. Preceding the variable region, the DNA contains a start codon used to insert a chloroacetylated nCAA (*e.g.* chloroacetyl-tyrosine, ClAcY). After the variable region there is a cysteine codon, a spacer (*e.g.* three Gly-Ser repeats, (GS)₃) and a site for ligation of a puromycin linker (Figure 5.1c). Puromycin is covalently bound to a synthetic oligonucleotide containing several PEG spacer building blocks and a short DNA sequence which is ligated to the library mRNA with a ligase enzyme (Materials and Methods 7.8, Table 7.10). The role of the puromycin linker is to covalently attach the translated peptide to its corresponding DNA/mRNA tag, allowing for identification of peptide hits by DNA sequencing. Translation of the open reading frame is carried out by the ribosome until

the TAG stop codon is reached. As the reaction mix does not contain release factor 1, the ribosome is stalled and cannot release the peptide or RNA template. At this point, puromycin is able to enter the ribosome, mimicking a tyrosine-tRNA (Y-tRNA) substrate (Figure 5.1b). The ribosome catalyses the formation of an amide bond between the peptide and puromycin, covalently linking the peptide to its tag.⁴⁴⁰ The (GS)₃ spacer is used to create a spatial separation between the peptide sequence and the puromycin/RNA/DNA tag, to minimise interference of the tag with target binding. The ClAcY and cysteine amino acids which flank the variable region of the peptide are used for cyclisation. The library is incubated at pH 8-9 which allows the cysteine thiol to attack the chloro-acetyl group of ClAcY in an S_N2 reaction making a stable thioether bond and cyclising the peptide (Figure 5.1a).

mRNA display library production is carried out fully *in vitro* and each step of the process can therefore be manipulated to include non-standard biological components. The method used in the Kawamura group uses the flexizyme technology developed by H. Suga for insertion of nCAA.^{441–443} Flexizymes are a family of synthetic acylation ribozymes which ligate a nCAA to a custom tRNA. Their use allows for insertion of nCAA into library peptide sequences for diversification of the library and peptide cyclisation. Selected amino acids are excluded from the medium during the translation step of mRNA display. This leaves vacant codons which can be repurposed by addition of a custom acyl-tRNA with the correct anticodon and carrying an alternative nCAA, previously prepared using flexizyme technology. The Suga group have developed a range of flexizymes which are able to accommodate a large range of acyl-donor substrates

ranging from nCAA with artificial side chains, N-Acyl or N-Alkyl modifications as well as D-amino acids.⁴⁴²

In this chapter, flexizyme technology for the synthesis of CIAcY-tRNA was used in mRNA display library preparation in order to obtain cyclic peptides (Section 5.3.3). CIAcY was chosen as nCAA for cyclisation due to the efficiency of amino acid insertion of the CIAcY-tRNA and because tyrosine absorbs UV light which is useful for peptide quantification and HPLC purification.

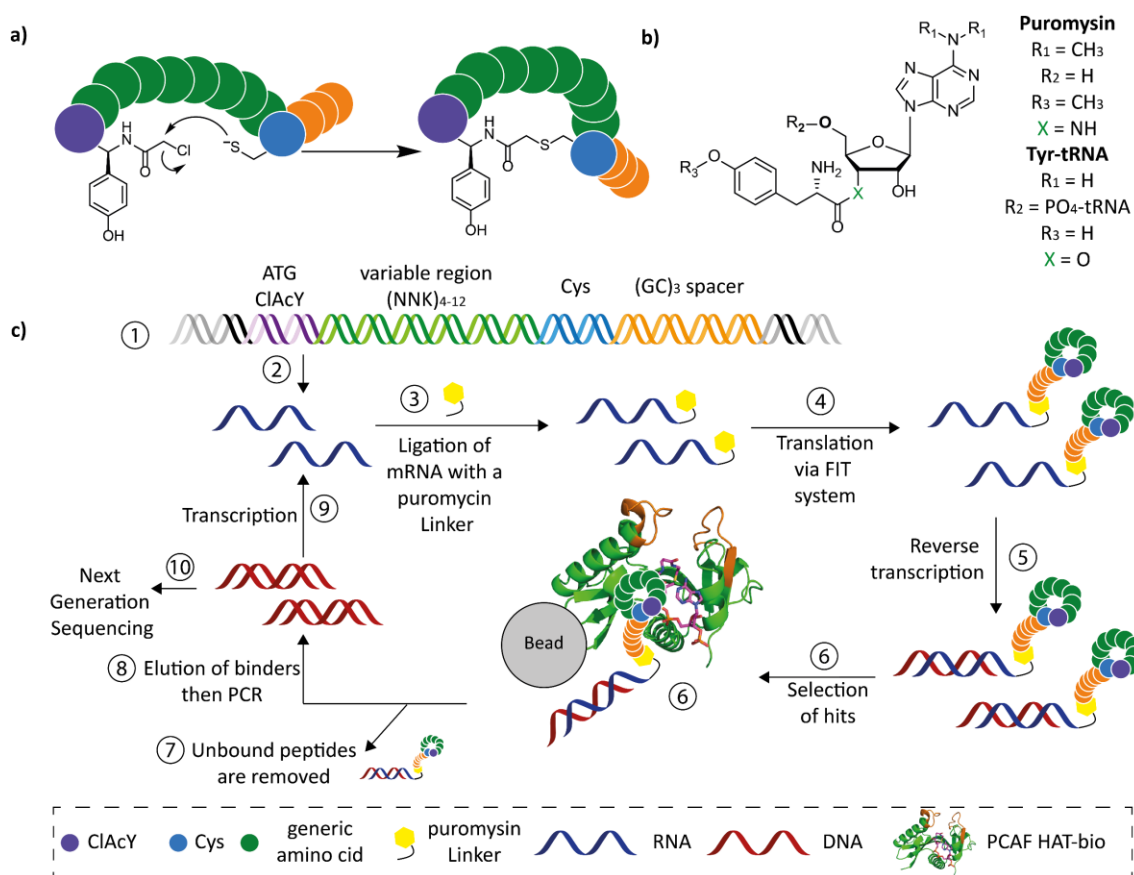


Figure 5.1 : mRNA display selection process and components.

a) Cyclisation reaction between the ClAcY and cysteine which flank the variable region of the library peptides. The cysteine thiol attacks the chloroacetyl group in an S_N2 reaction forming a stable thioether bond. **b)** Puromycin stalls the ribosome by mimicking Tyr-tRNA with a non-hydrolysable amide bond at position X. The selection tag or the tRNA are linked to the structures shown through position R_2 . **c)** Steps of the mRNA display method used (based on RaPID technology) Step 1: the DNA template contains an ATG start codon used for chloroacetyl-tyrosine (ClAcY) insertion, a variable region with a selected number of NNK degenerate codons, a Cys codon, a (GS)₃ spacer, and a site for puromycin linker ligation. Step 2: DNA is transcribed to make the mRNA starting library. Step 3: Ligation of the DNA-PEG-puromycin linker to the mRNA library. Step 4: customised in vitro translation step with methionine excluded to be replaced by ClAcY using a premade ClAcY-tRNA. After synthesis, the peptide is ligated to puromycin by the ribosome, covalently linking the peptide to its encoding mRNA sequence as a tag. Spontaneous cyclisation occurs at pH 8-9. Step 5: Reverse transcription forms the mRNA/DNA helix tag. Step 6: The library is incubated first with streptavidin-biotin-avitag coated beads to remove non-specific binders then with tagged PCAF protein which is pulled down with streptavidin coated beads. Step 7: the beads are washed to remove unbound and weakly binding peptides. Step 8: Heat elution is used to recover the DNA tag of bound peptides for qPCR quantification and PCR amplification. Step 9: The DNA is transcribed to mRNA to be used in a new round of selection in order to achieve enrichment of binding peptides. Step 10: After sufficient enrichment has been achieved, the DNA is analysed using next generation sequencing (NGS) to identify peptide binding hits. Figure adapted from Serrano et al. with permission.²⁸¹

The screening using mRNA display is a 10-step process adapted from Suga *et. al.* methodology (Figure 5.1c).^{443,444} After the DNA library is designed it is translated and a custom starting mRNA library is made up containing mRNA with variable regions of selected lengths. In the following step, the puromycin linker is ligated to the mRNA library members. Peptide translation is then carried out using a custom *in vitro* translation (IVT) system with methionine excluded and ClAcY-tRNA added for ClAcY insertion into methionine codons. The nascent peptides are bound to their tags through a covalent amide bond formed to puromycin, and the peptides cyclise spontaneously. A reverse transcription step is carried out to complete the mRNA/DNA hybrid tag. The fully formed cyclic peptide library is then incubated with unmodified streptavidin coated beads to remove any peptide sequences which bind non-specifically to either streptavidin or the beads (this step is referred to as a negative control). The remaining library is then incubated with target protein immobilised on streptavidin beads for an affinity capture step. The supernatant with the unbound library members is removed from the beads which are washed repeatedly to remove any weak binders. Heat elution is then used to recover the DNA tags of binding peptides for PCR amplification. The amplified DNA is quantified by qPCR, purified, checked for size on an agarose gel, and is fed back into another round of selection by transcription to mRNA. This process is repeated over multiple rounds in order to achieve enrichment of peptide binders. After sufficient enrichment has been measured using qPCR, the DNA output can be sequenced to identify peptide binders. Peptides identified to be binders during the selection process are synthesised free of tags and on sufficient scale for validation using binding and activity assays.

This method has previously been used in the Kawamura and Suga groups to successfully target various epigenetic enzymes including prolyl hydroxylase 2 (PHD2), TET1 and KDM4A/B/C.^{445–447} These epigenetic enzymes were chosen as targets for mRNA screens in order to develop inhibitors for the proteins (TET1, KDMs) or to develop biological tools to facilitate their study (PHD2). This demonstrates the utility of cyclic peptide screening for proteins which were previously seen as hard targets for probe development. In the examples given, cyclic peptide chemical probes were successfully generated by screening libraries of peptides with randomly generated variable regions. These libraries can also be further tailored to the protein of interest. To target the HDAC SIRT2 Yamagata *et al.* used a peptide library containing a constant ϵ -trifluoroacetylamine lysine residue as an acetyl-lysine mimic flanked by variable amino acids, producing a cyclic peptide inhibitor superior to previously described small molecule and linear peptide inhibitors.⁴⁴⁸ The increased accessibility NGS can also be used for in depth data analysis, allowing for detailed SAR analysis and informed inhibitor design.⁴⁴⁹

The activity of epigenetic enzymes is context dependent. In a biological setting, PCAF activity is affected by cellular location (cytoplasm or nucleus), signalling cascades, interactions with regulatory proteins or protein complexes, co-localisation with histone or non-histone substrate proteins and combinatorial PTM signalling on histones (Chapters 1,2,3). Whilst available generic small molecule HAT inhibitors (Chapters 1,2), target the active site of HATs inhibiting catalytic activity, they are not known to discriminate between substrates. Development of peptide chemical probes with alternative binding/inhibition modes may provide substrate specific inhibition. In addition, mRNA display can help identify potent binding sequences, that may

correspond to endogenous PCAF substrates or protein binding partners, helping to understand the substrate selectivity of PCAF.

As discussed in sections 1.5.2 and 2.5.2.3, there are no potent or selective small molecule inhibitors for the HAT domain of PCAF or its close homologue GCN5. PCAF has important biological roles which would be better understood with the availability of a selective and potent chemical probe/inhibitor. Considering reported successes achieved with targeting epigenetic enzymes, mRNA display technology was used in an attempt to provide a cyclic peptide chemical probe for PCAF HAT.

5.1.4 Aims

The aim of this chapter is to use mRNA display screening against PCAF HAT and HAT-BRD to develop cyclic peptide chemical probes for these proteins.

Specific aims are:

- Production of all materials needed for mRNA display screening
- Use of mRNA display to screen PCAF HAT and HAT-BRD
- Synthesis of peptides identified using NGS on mRNA display screen outputs
- Use of peptides synthesised for binding and activity tests against PCAF HAT and HAT-BRD

5.3 Preparation of materials for cyclic peptide mRNA-display

To carry out a cyclic peptide mRNA display selection, necessary starting materials must be assembled: purified recombinant protein with appropriate tags for immobilisation on bead, charged CIACy-tRNA for CIACy insertion into the cyclic peptides, and a starting mRNA library.

5.3.1 PCAF HAT and HAT-BRD protein construct: design and protein production

Single HAT and dual domain HAT-BRD containing PCAF constructs were designed based on protein domain boundaries that yield soluble protein. Biotin affinity tags needed for immobilisation were added (Section 2.2.1.1). Parallel screening of HAT and HAT-BRD PCAF constructs was carried out to identify peptides which bind to the HAT domain (these peptides might appear in both screens), the BRD or the HAT/BRD interface (these peptides would only be present in a screen against HAT-BRD). To date, no crystal structure of the dual HAT-BRD construct has been published for PCAF or for its close homologue GCN5, making the location of the interaction surfaces between the HAT and BRD unknown, thus a dual domain targeting cyclic peptides would also be valuable in understanding the inter-domain crosstalk / interactions. PCAF HAT and HAT-BRD recombinant proteins with appropriate tags (avitag for mRNA-display selection and/or His tag for purification) were produced for use in mRNA screening and subsequent testing (Figure 5.2, Material and methods 7.4.6, 7.4.8). PCAF HAT-bio and His-HAT-BRD-bio proteins used were shown to be active using MALDI-TOF-MS (Figure 2.6).

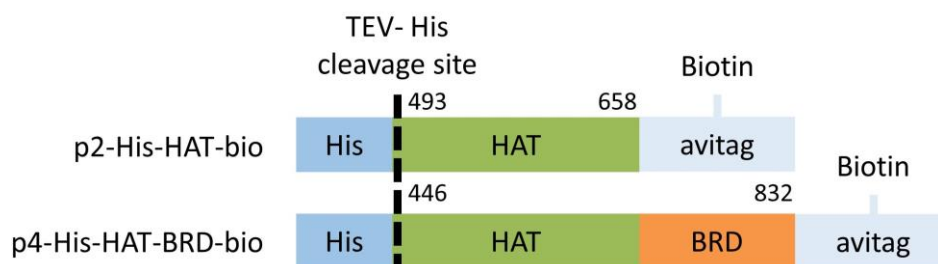


Figure 5.2: HAT and HAT-BRD constructs of PCAF used in mRNA display.

Two constructs of PCAF are shown: His-HAT-bio containing only the HAT domain and His-HAT-BRD-bio containing the HAT and bromodomains of PCAF. The His tags are used for purification and subsequently removed for p2-His-HAT-bio. The Avitag provides a biotinylation site for biotinylation during protein production.

5.3.2 Optimisation of biotinylated PCAF protein loading onto beads

The optimal protein to bead ratio was determined in loading tests with PCAF HAT-bio/His-HAT-BRD-bio and streptavidin-coated magnetic beads. The proteins were incubated with a determined amount of streptavidin coated beads and the bound and unbound protein fractions were quantified by SDS-page protein gels (Figure 5.3). In the gels, the bound and unbound fractions are examined to identify in which sample the protein has been completely bound to the beads and is no longer seen in the unbound fraction. Optimal bead loading is determined based on the protein to bead ratio in the selected sample. A loading of 1.43 pmol/ μ L bead for PCAF HAT-bio and 2 pmol/ μ L bead for PCAF His-HAT-BRD-bio was subsequently used during the selection process (Materials and methods 7.8.2).

To ensure the immobilised protein retained catalytic activity and the enzyme active site remained accessible to cyclic peptides, a MALDI-TOF-MS activity assay was carried out with immobilised PCAF HAT-bio and H3₁₋₂₁ peptide. The protein was immobilised on streptavidin beads, which were washed to remove any unbound protein. The beads were then incubated with H3₁₋₂₁ peptide with mixing at 37°C, and aliquots of

supernatant were taken at specific time points. MALDI-TOF-MS was used to assess activity by measuring conversion of H3₁₋₂₁ to H3₁₋₂₁Ac, as described in section 2.2.2 (Materials and methods 7.7). A +42 Da shift was observed in the MALDI-TOF-MS spectra, indicating that immobilised HAT-bio retained activity and suggesting that the active sites are not obstructed by immobilisation. This demonstrated that selection of peptides binding to the active site in the mRNA display screen was possible. This experiment was not carried out with His-HAT-BRD-bio due to limited available stocks of this protein.

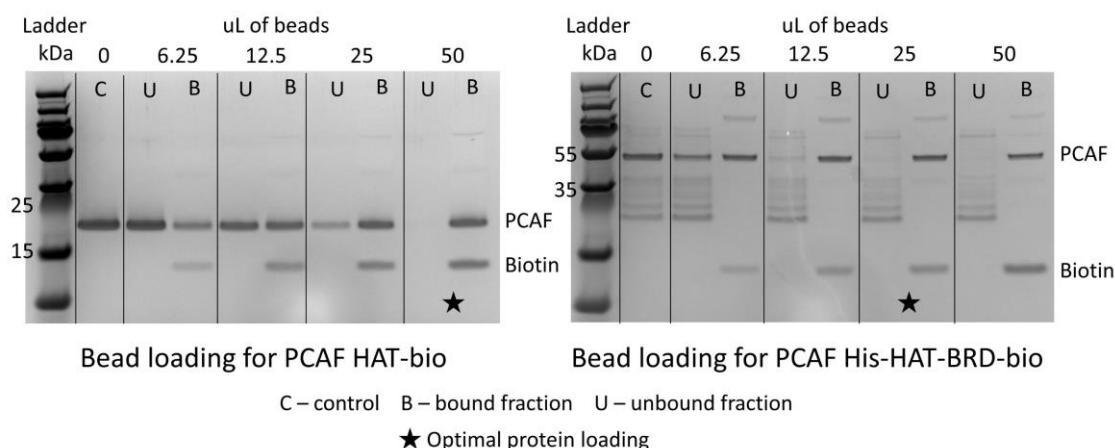


Figure 5.3: Bead loading test for PCAF HAT-bio and His-HAT-BRD-bio.

Biotinylated protein constructs are immobilised on streptavidin coated beads. The amount of protein was kept constant in each case whilst the volume of beads was increased from 0 µL to 50 µL of beads (left to right on gels). After incubation, the supernatant was removed from the beads (unbound fraction) and the bound protein was removed from the beads by heat elution in buffer (bound fraction). These fractions were run on an agarose gel, and bead loading was calculated from the first lane where the protein was completely bound to the beads (denoted with a star). A loading of 1.43 pmol/µL bead for PCAF HAT-bio and 2 pmol/µL bead for PCAF His-HAT-BRD-bio were determined as optimal.

5.3.3 tRNA and flexizyme synthesis

To cyclise the peptides a ClAcY residue was reprogrammed in the place of the initial formyl-methionine (fMet) at the start codon ATG. An fMet-tRNA and an eFx flexizyme were synthesised as described by Goto *et al.*⁴⁴² DNA primers were used to synthesise the DNA template which was translated *in vitro* to generate RNA. After synthesis, the

RNA product was purified by SEC. The synthesised flexizyme was then used to catalyse the ligation of ClAcY onto the fMet-tRNA. The ClAcY-tRNA formed was precipitated out of solution, dried and stored at -80° C for use in mRNA library preparation. See material and methods section 7.8.3 for full protocol.

5.3.4 Preparation of a starting cDNA library for use in the mRNA display screen

DNA sequences designed with a variable region of 4-12 degenerate codons (NNK 4-12) were produced by PCR using purchased oligonucleotide templates. These were transcribed to mRNA and quantified either using Nanodrop™ or a Qubit fluorimeter. The starting mRNA library for a selection was prepared using a tailored ratio of mRNA sequences containing different variable region lengths. Starting libraries with varying ratios of mRNA of different lengths were prepared for the mRNA display selections performed, aimed to produce peptides with desired properties (Table 5.1, Materials and Methods 7.8.4).

5.4 Cyclic peptide mRNA display screening of PCAF HAT/HAT-BRD

5.4.1 Results of the Cyclic peptide mRNA display selections against PCAF HAT/HAT-BRD

Cyclic peptide mRNA display screening against PCAF HAT-bio and His-HAT-BRD-bio was carried out using standard generic selection conditions used in the Kawamura group (Table 5.1, Selection 1). However, due low recovery of binding peptides, alternative conditions were tested to optimise the selection. Overall, five different selections were

carried out, with conditions modified to address issues encountered during the screening process (Table 5.1, Selections 1-5, Materials and Methods 7.8).

The first selection (Selection 1) was run using protocols previously optimised in the Kawamura group (library size of NNK8, on bead screening, buffer conditions as used in activity assays). This screen was unsuccessful due to loss of library (111 bp) and predominance of a primer dimer PCR product 'lower band' (< 100 bp), as seen in DNA gels performed at the end of selection rounds to monitor DNA recovery (Figure 5.4a). To address this problem, PCR conditions were optimised; changes in DMSO, primer and $MgCl_2$ concentrations were explored showing that a decrease in $MgCl_2$ concentration decreased primer dimer formation (Figure 5.4b). In addition, shorter primers were designed which resulted in cleaner PCR products (Figure 5.4c, primer sequences given in Table 7.10).

A second selection (Selection 2) was run using optimized PCR conditions and a pull-down variation to the selection protocol in which the library was first incubated with the protein in solution followed by addition of streptavidin coated beads to pull out the biotinylated protein with bound peptides. This selection appeared successful for His-HAT-BRD-bio which was carried through nine rounds of selection (Figure 5.5a). The second selection for the HAT-bio domain was unsuccessful with loss of library DNA by the end of round 4.

Selection	Protein	Number of NNK repeats	Number of rounds	Buffer	Conditions
Selection 1	HAT	8	4	50 mM HEPES 50 mM NaCl 0.05 % Tween-20 pH 7.6	On bead screening Long Primers 2.5 mM MgCl ₂ in PCR streptavidin control
	HAT-BRD				
Selection 2	HAT	8	4	50 mM HEPES 50 mM NaCl 0.05 % Tween-20 pH 7.6	Pull down screening Short primers 1.75 mM MgCl ₂ in PCR streptavidin control
	HAT-BRD		9		
Selection 3	HAT	9-12 Ratio: 1:1:1:1	7	50 mM HEPES 50 mM NaCl 0.05 % Tween-20 pH 7.6	Pull down screening Short primers 1.75 mM MgCl ₂ in PCR streptavidin control
	HAT-BRD				
Selection 4	HAT	8-12 Ratio: 1-2-4-8- 16	1	50 mM HEPES 50 mM NaCl 0.05 % Tween-20 pH 7.6	Pull down screening Short Primers 1.75 mM MgCl ₂ in PCR Δ AA streptavidin-biotin-avitag control 100 μ M AcCoA
	HAT-BRD				
Selection 5	HAT	8-12 Ratio: 1-2-4-8- 16	3	Control: 50 mM HEPES 0.01% Tween-20 0.01% BSA 150 mM NaCl pH 7.6 Selections: 50 mM HEPES 0.01% Tween-20 0.01% BSA 150 mM NaCl 1 mg/mL YRNA pH 7.6	Pull down screening Short Primers 1.75 mM MgCl ₂ in PCR streptavidin-biotin-avitag control 100 μ M AcCoA Increased NaCl, BSA and YRNA for blocking of non-specific binding
	HAT-BRD		4		

Table 5.1: Conditions used for mRNA display Selections 1-5 against PCAF HAT and HAT-BRD.

A summary of conditions used for each of the mRNA display selections run against PCAF HAT-bio and PCAF His-HAT-BRD-bio is given. YRNA: yeast RNA, Δ AA: no amino acids in selection.

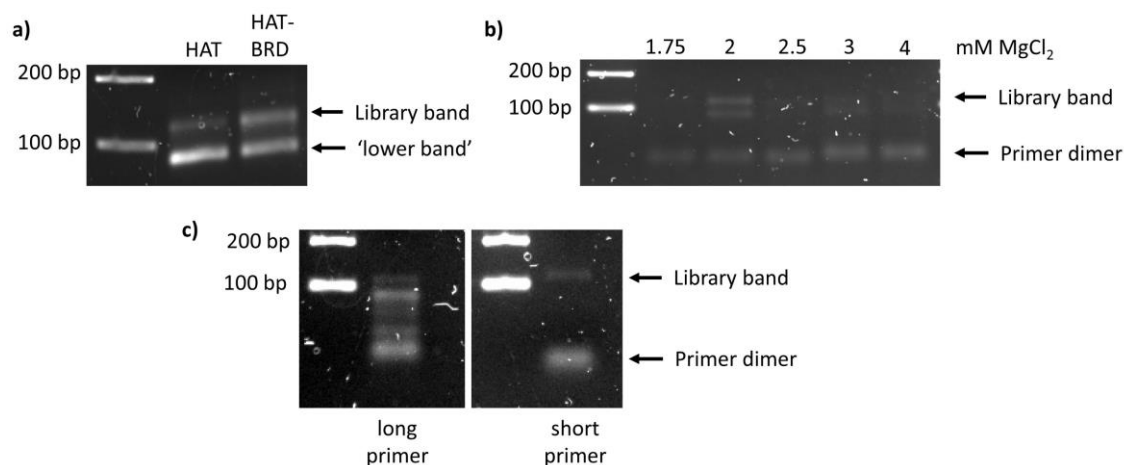


Figure 5.4: Optimisation of mRNA display conditions.

a) Agarose gel of PCR product of DNA recovered from Selection 1 round 4 (1R4) showing the upper library band (111 bp) and the unwanted 'lower band' (< 100 bp) for mRNA selections against PCAF HAT-bio and His-HAT-BRD-bio **b)** Agarose gel showing effect of different concentrations of $MgCl_2$ in PCR reactions on DNA products amplified (library band 111 bp) **c)** Agarose gel comparing the use of long and short primers in PCR amplification of DNA recovered in selections.

A third selection (Selection 3) was carried out for both HAT-bio and His-HAT-BRD-bio with the same conditions used in Selection 2, but using a starting library with a longer variable region (NNK 9-12). Larger peptides have a larger surface binding area and may yield tighter binding peptides. A higher proportion of tight binders in the library may lead to higher DNA recovery making the PCR step more successful. The output of the Selections 2 and 3 for His-HAT-BRD-bio and Selection 3 for HAT-bio were then assessed using NGS. DNA from selection rounds were chosen for NGS analysis based on enrichment of the library being observed in the qPCR results at the end of each round (Figure 5.5a). Peptides identified as potential binders in the NGS results were synthesised on 0.5 μM scale for testing as described in sections 5.4.2-5.4.4.

A fourth selection (Selection 4) was run to assess the effects of non-specific DNA/RNA binding of the DNA/RNA hybrid tag to PCAF. A control selection (with all components included) was run alongside an amino acid free (ΔAA) selection. The levels of DNA

recovery at the end of the single round of selection, as determined by qPCR quantification, showed similar levels of recovery in the control and Δ AA selections (Figure 5.5b). This indicated that issues with previous selections were in part due to non-specific DNA/RNA binding to PCAF, which must be minimised in order to obtain reliable peptide binders.

Optimisation of the selection buffer was carried out to minimise non-specific DNA/RNA binding. Parameters tested included buffer type, salt concentration, and use of additives such as bovine serum albumin (BSA), yeast RNA (yRNA), salmon sperm DNA, herring sperm DNA and heparin, a naturally occurring glycosaminoglycan. Higher NaCl concentrations decreased DNA recovery for PCAF HAT-bio. BSA, herring sperm DNA, salmon sperm DNA and heparin did not significantly affect DNA recovery. Use of 1 mg/mL yRNA lowered DNA recovery for both HAT-bio and His-HAT-BRD-bio (Figure 5.5c). Based on these results, a fifth and final selection (Selection 5) was carried out using 50 mM HEPES, 0.01% Tween-20, 0.01% BSA, 150 mM NaCl, 1 mg/mL yRNA as buffer. This selection was carried out over four rounds and the final output for His-HAT-BRD-bio assessed using NGS. For HAT-bio, the library band was lost in the third round of selection.

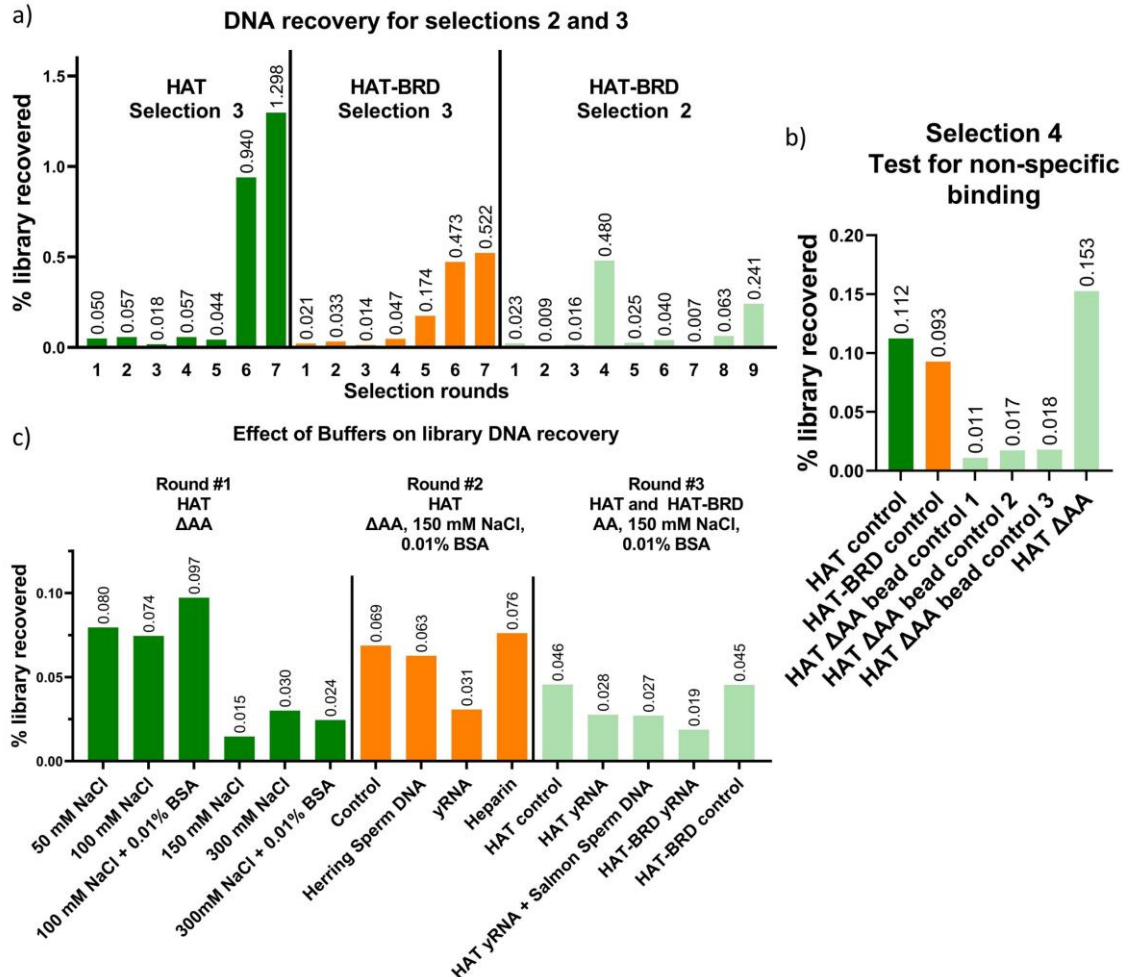


Figure 5.5: Results of qPCR monitoring of mRNA selections.

a) DNA enrichment determined by qPCR for Selection 2 (His-HAT-BRD-bio) and Selection 3 (HAT-bio and His-HAT-BRD-bio), showing increased percentage of library DNA recovery indicating enrichment of binding sequences in later rounds of each selection. **b)** Library DNA recovery measured by qPCR for Selection 4. The library DNA recovery for immobilised PCAF HAT-bio is greater in the absence of library peptides (HAT ΔAA) than with library peptides present (HAT control and HAT-BRD control), showing that PCAF HAT is binding to the DNA/RNA tag used in the selection. The recovery for the bead controls 1-3 (HAT ΔAA bead control 1-3) is minimal in absence of peptide showing that the streptavidin-biotin-avitag portion of the protein does not bind the DNA/RNA tag. **c)** qPCR quantification of library DNA recovery using different buffers over three rounds of testing. Round #1 shows recovery for different NaCl concentrations in a ΔAA (no amino acids in buffer) selection with or without 0.01% BSA using PCAF HAT-bio. 150 mM NaCl is chosen as optimal. BSA has little effect on protein to DNA binding. Round #2 shows effects of DNA blockers (all used at 1 mg/mL) in a ΔAA selection using 150 mM NaCl and 0.1% BSA in the buffer. Yeast RNA blocker shows a decrease in DNA recovery compared with the control, with Herring Sperm DNA and Heparin showing little effect on recovery. Round #3 shows a final round of tests using HAT-bio and His-HAT-BRD-bio controls with 150 mM NaCl, 0.01% BSA and with amino acids in the buffer. Yeast RNA decreases DNA recovery for both protein constructs. Addition of salmon sperm DNA does not further reduce DNA recovery.

5.4.2 NGS sequencing results for samples from mRNA display selections against PCAF HAT/HAT-BRD

Outputs from the Selections 2, 3 and 5 were sequenced using NGS (Materials and Methods 7.8.6). The results of this sequencing showed that the selection process was not successful.

In the sequencing results for the PCAF selections carried out, there were many indicators of an unsuccessful selection (Table 5.2). There was very little enrichment of peptide sequences, with the top 10 sequences accounting for an abnormally low percentage of total reads (< 1% where a successful selection will typically have 1-10% enrichment for the top peptide binding sequences). The NGS results showed a large proportion of sequences with frameshift mutations so that the intended end sequence of CGSGSGS* (* = stop) instead codes for either RQRQRQLGRGAE or AAAAAAARTGGG (the additional amino acids in these sequences originate from bases of the RNA downstream of the stop codon that are usually not translated, with the ribosome stalling at the end of the RNA strand rather than at a stop codon). Methionine codon containing sequences were often enriched, which should not occur as methionine was removed from the selection. In the absence of methionine, isoleucine is most commonly mis-incorporated at methionine codons, which can lead to synthesis of sensible peptide sequences with isoleucine being substituted for methionine.⁴⁵⁰ A large proportion of stop codons were also seen in the sequencing results. Although the amber stop codon appears in the NNK codon sequences, translation halts at the appearance of a stop codon. This produces short linear peptide sequences with low binding potential which should be phased out of the library during bead washes to remove non-binders and weak binders during the

enrichment process. This was not the case in the sequencing results for the PCAF screens. The top sequences for the Selections 2 and 3 against HAT-bio and His-HAT-BRD-bio had frameshift mutations for 7 and 3 of the top 10 sequences respectively. The top 10 sequences for Selection 5 against His-HAT-BRD-bio showed no frameshift mutations but had an increased number of methionine codons compared with the other sequencing results (Table 5.2).

	HAT S3R7	HAT-BRD S3R7	HAT-BRD S5R4
1	YCRREWTV RQRQRLGRGAE	YANAIKQGHSVC	YLTNELK*IC
2	YIRVDRGSRLR AAAAAAARTGGG	YPNDRGGQEVKC	YSLDPK M KAYKC
3	YNQVDSVDFV RQRQRLGRGAE	YASYKQVIVL AAAAAAARTGGG	YSLTCEV MM LYTC
4	YRVDRVAL RQRQRLGRGAE	YCHNNPDKGTPC	YLQNSGTRIQRDC
5	YYASGQIRENFC	YTTEHDAKQGHTC	YDWSLAHE M TPKC
6	Y*T* MDSAAAAAARTGGG	YLTCPKVASC	YPQKTRQRYL M C
7	YCRRKWTV RQRQRLGRGAE	YKQQVIVIV RQRQRLGRGAE	YKS MHMF PSSTLSC
8	YCRREWT MRQRQRLGRGAE	YTRKQGQ M SAHTC	YSLSTDSDT*HC
9	YNGQVHASTLE M QC	YTLQEVSA AAAAAARTGGG	YSIQHLLIPYTPQC
10	YYRREWTV RQRQRLGRGAE	YMAVKQCHSG*	YKKTSSLTVTRHSC

Table 5.2: Results of NGS sequencing.

The top 10 most abundant peptide sequences in the next generation sequencing (NGS) results for HAT-bio selection 3 round 7 (S3R7), His-HAT-BRD-bio selection 3 round 7 (S3R7) and His-HAT-BRD-bio selection 5 round 4 (S5R4) are shown. The initial tyrosine and terminal cysteine residues needed for cyclisation are included in the sequences. The Notation uses one letter code for amino acids with methionine codons indicated in red **M**, stop codons denoted as *, and mutations leading to frameshift mutations denoted in brown.

5.4.3 Synthesis of peptides identified in the mRNA display screen against PCAF HAT/HAT-BRD

Peptides from the sequencing results of Selection 2 (HAT-BRD S2R9, selection 2 round 9) and Selection 3 (HAT and HAT-BRD S3R7, selection 3 round 7) were chosen for synthesis to test for binding and activity against PCAF. These peptides were chosen on the basis of their enrichment in the sequencing results and on their belonging to peptide families (families of peptides determined based on sequence similarity). Twelve peptides were synthesised using SPPS. These were non-labelled peptides, removing any

interference of binding or activity linked to the GSGSGS linker, puromycin or DNA/RNA tag used during the selection process.

Peptide synthesis was carried out using Fmoc-based chemistry on a 0.05 mmol scale as described in Section 4.2.2 (Figure 4.1) using a Liberty Blue automated peptide synthesiser with microwave heating. A modified workflow was used for cyclisation of peptides before purification (Figure 5.6).

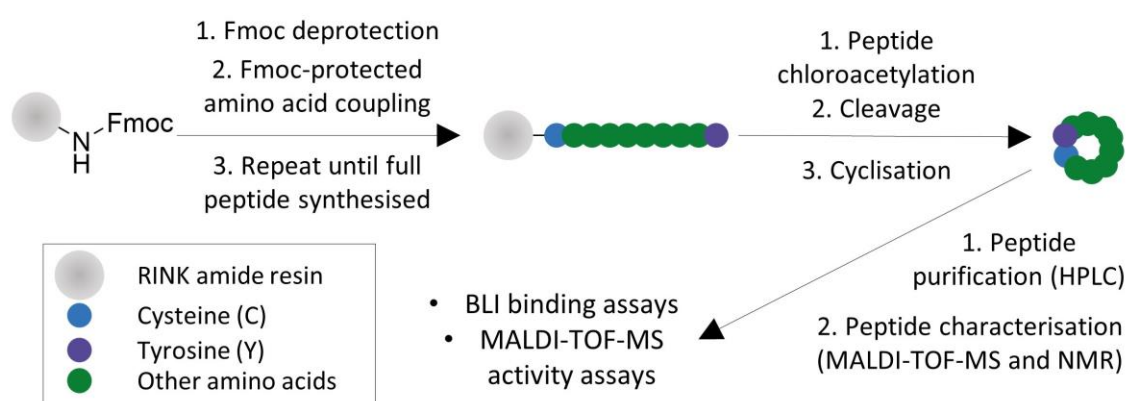


Figure 5.6: Peptide synthesis and testing workflow used for cyclic peptides.

Peptide synthesis was carried out with a Liberty Blue microwave synthesiser with RINK amide resin using Fmoc based chemistry, followed by chloro-acetylation, cleavage from beads and cyclisation as shown in Figure 5.7. After cleavage peptides were purified and characterised before use in assays.

After automated synthesis, the peptides were chloro-acetylated at the *N*-terminus whilst still on resin. A TFA cleaving mixture was then used to deprotect and cleave the peptide from the resin. The peptides were dissolved in a mixture of acetonitrile and ammonium acetate buffer adjusted to pH 8 using potassium hydroxide (2 M). Spontaneous cyclisation between the *N*-terminal chloroacetyl-tyrosine and C-terminal cysteine occurred during an incubation period of 37 °C for 1 h. (Figure 5.7). Following synthesis and cyclisation, the peptides were purified using semi-prep HPLC (reverse phase C18 column) and the mass of the peptides was confirmed by MALDI-TOF-MS

(Table 7.7, Appendix Cii). See material and methods section 7.5 for SPPS and HPLC purification protocols.

Of the twelve peptides synthesised, six were successfully purified for testing (LS3-LS6, LS8, LS11). Certain peptides (LS1, LS2, LS7, LS9, LS10, LS12) were insoluble after synthesis and cleavage/cyclisation and could not be carried forwards for purification and testing (Table 5.3).

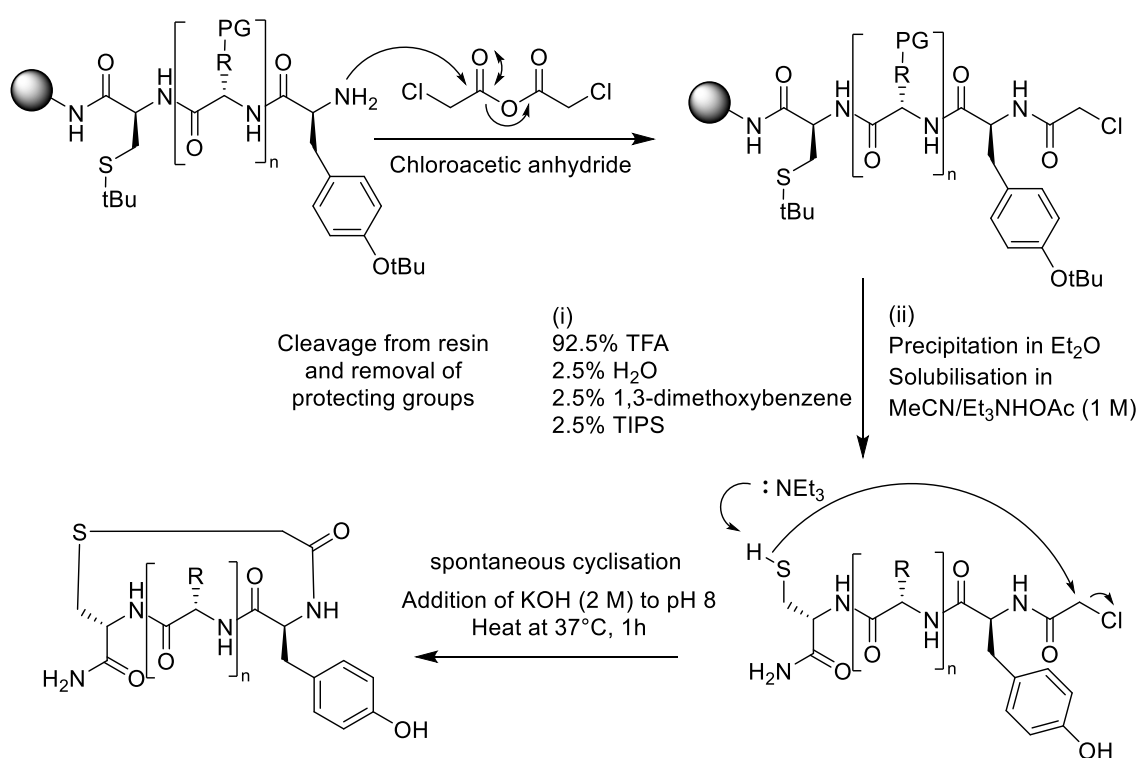


Figure 5.7: Peptide cyclisation scheme.

Peptide cyclisation was carried out manually on bead, after automated synthesis. A mix of TFA with scavengers was used for cleavage and Et₂O was used to precipitate the cleaved peptides. The peptides were then cyclised in acetonitrile/ammonium acetate buffer adjusted to pH 8 using potassium hydroxide (2 M). Spontaneous S_N2 cyclisation occurs during incubation at 37 °C for 1 h.

ID	Sequence	Length	Selection	purity	BLI binding	Inhibition of activity
LS1	YYASGQIRENFC-NH ₂	10	HAT S3R7	N/A		
LS2	YNGQVHASTLELQC-NH ₂	12	HAT S3R7	N/A		
LS3	YNGTPGHVWC-NH ₂	8	HAT-BRD S2R9	>95%	Non-binding	No inhibition
LS4	YIDPQEVSALC-NH ₂	9	HAT-BRD S3R7	>95%	Non-binding	No inhibition
LS5	YTSEAQQWEC-NH ₂	8	HAT-BRD S2R9	>95%	Non-binding	No inhibition
LS6	YPNDRGGQEVKC-NH ₂	10	HAT-BRD S3R7	>95%	Non-binding	No inhibition
LS7	YASYKQVIVLC-NH ₂	9	HAT-BRD S3R7	N/A		
LS8	YLFSDRHRNRDC-NH ₂	10	HAT S3R7	>95%	Non-specific binding	No inhibition
LS9	YANAIKQGHSVC-NH ₂	10	HAT-BRD S2R9	N/A		
LS10	YIRVDRGSRLC-NH ₂	9	HAT S3R7	N/A		
LS11	YSKSTYGTPGNC-NH ₂	10	HAT-BRD S3R7	>95%	Non-binding	No inhibition
LS12	YCRREWTESPC-NH ₂	9	HAT S3R7	N/A		

Table 5.3: Peptide sequences synthesised for testing.

The peptides were synthesised on a Liberty Blue microwave peptide synthesiser. The peptide sequence is shown, followed by the variable region length, and the selection (S) and round (R) from which the peptide was identified. Sequences are shown with an initial tyrosine and terminal cysteine necessary for cyclisation. Insoluble peptides are shown as N/A purity. Soluble peptides were purified using a JASCO HPLC semi-prep system. These peptides were tested against PCAF HAT and HAT-BRD for binding using biolayer interferometry (BLI) and for inhibition of enzyme activity using MALDI-TOF-MS. Results of these assays are shown where applicable, with the peptides showing no binding or non-specific binding, and showing no inhibition of enzyme activity at the concentrations tested.

5.4.4 BLI binding and MALDI-TOF-MS activity assays for cyclic peptides with PCAF HAT and HAT-BRD

Purified peptides were tested for binding to PCAF HAT and HAT-BRD using BLI. BLI measures binding based on the difference of optical interference of white light reflected from a reference layer and a biolayer where the target is immobilised. The biolayer with bound target is placed in a solution containing a selected compound, where binding causes a measurable change in spectral shift compared to the baseline, which is proportional to the amount of compound bound to the target. The biolayer is removed from the binder solution and placed in clean buffer solution allowing dissociation of binders to be measured. This technique allows for real-time measurement of binding

and dissociation constants and needs only one tagged binding partner for immobilisation.

HIS-tagged PCAF HAT-BRD protein was used for immobilisation on Ni²⁺ BLI sensors, and biotinylated PCAF HAT-bio protein for streptavidin BLI sensors. The peptides were tested at different concentrations from 156.25 nM to 10 µM. None of the peptides showed binding to the HAT or HAT-BRD domains of PCAF (an example negative BLI trace is shown in Figure 5.8a), except for LS8 which exhibited a non-specific binding profile with unloaded biosensors (Figure 5.8b). Full screening method can be found in materials and methods 7.9.

Purified peptides were also tested for inhibitory activity of PCAF HAT and HAT-BRD using the MALDI-TOF-MS activity assay developed with H3₁₋₂₁ peptide as substrate. The masses of the cyclic peptides tested were <2000 Da and therefore do not interfere with acquisition of MALDI-TOF-MS traces for the H3₁₋₂₁ peptide with a parent mass of 2254 Da. Anacardic acid was used as a positive inhibitory control and the peptides were tested at three concentrations (1 µM, 10 µM and 100 µM). Whilst anacardic acid showed full inhibition of PCAF HAT and HAT-BRD acetylation at a 100 µM concentration, the peptides did not show inhibitory activity for either enzyme (Figure 5.8c-d).

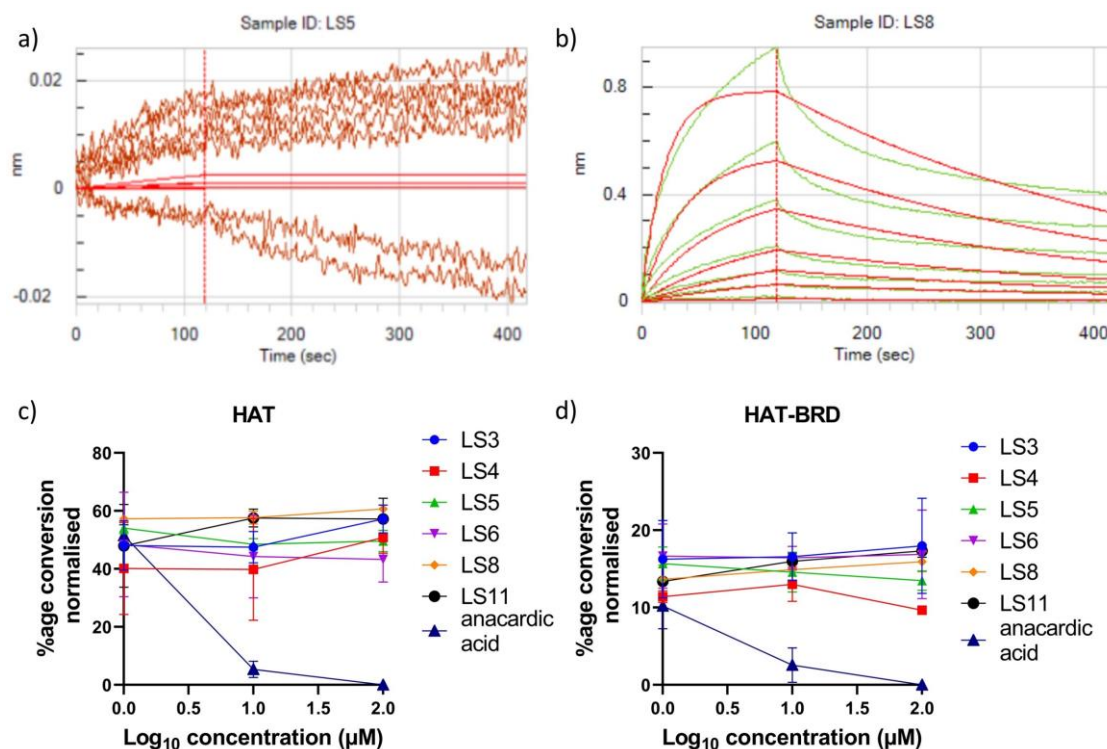


Figure 5.8: Results of BLI binding and MALDI-TOF-MS activity of cyclic peptides against PCAF HAT and HAT-BRD.

Peptides identified using mRNA display against PCAF HAT and HAT-BRD as being potential binders were tested for binding activity using biolayer interferometry (BLI) and inhibitory activity using MALDI-TOF-MS. For BLI peptides were tested between 10 μM to 156.25 nM concentrations (three experimental replicates) with a two-fold dilution series and results were fitted using a global 1:1 fit model. Assay conditions for MALDI-TOF-MS assays were PCAF enzyme (1 μM), AcCoA (50 μM), H3_{1-21} substrate peptide (10 μM), cyclic peptide or anacardic acid (1, 10 and 100 μM) in 1% DMSO buffer (HEPES 50 mM and NaCl 50 mM). Results are shown as average $\pm$ SD of 3 experimental replicates normalised to conversion of H3_{1-21} peptide in 1% DMSO buffer **a)** Example BLI trace for a non-binding peptide. The results for LS5 against PCAF His-HAT using streptavidin BLI sensors is shown. **b)** Example BLI trace for non-specific binding of a peptide to the BLI sensors (no protein added). The results for LS8 against PCAF HAT-bio using streptavidin BLI sensors is shown. **c)** MALDI-TOF-MS activity assay against PCAF HAT showing the conversion of the H3_{1-21} peptide substrate to $\text{H3}_{1-21}\text{Ac}$, using anacardic acid (AA) as positive inhibitor control. Peptides LS3, LS4, LS5, LS6, LS8 and LS11 were tested and showed no inhibitory activity of PCAF HAT. **d)** MALDI-TOF-MS activity assay against PCAF His-HAT-BRD showing the conversion of the H3_{1-21} peptide substrate to $\text{H3}_{1-21}\text{Ac}$, using anacardic acid (AA) as positive inhibitor control. Peptides LS3, LS4, LS5, LS6, LS8 and LS11 were tested and showed no inhibitory activity of PCAF His-HAT-BRD.

5.6 Discussion

mRNA display screening is a powerful method which has been key in identifying cyclic peptide binders and inhibitors against epigenetic protein targets. However, not all selections yield positive binders. Issues which arise in the use of this method are often protein-specific and can generally be attributed to protein size, protein immobilisation, protein specific properties, protein structure or non-specific binding. These issues were addressed during the selections performed against PCAF HAT-bio and His-HAT-BRD-bio.

5.6.1 Lower band problems (PCR optimisation, short primers, HANDS PCR, DNA/RNA purification)

A significant problem encountered during the selections was loss or degradation of library (DNA and RNA) between rounds of selection. This occurs when the quantity of DNA recovered is minimal (as measured by qPCR) and many rounds of PCR become necessary to produce enough material for use in subsequent rounds of selection. Loss of library was observed by agarose gel of the recovered DNA following PCR amplification: the band corresponding the molecular weight of the DNA (111 bp for NNK8, 123 bp for NNK12) becomes weak or disappears (Figure 5.4a). Loss of library was accompanied by an increase in primer dimer formation indicated by one or several bands at a lower molecular weight (< 100 bp). This issue was observed throughout the selection process against both PCAF constructs and was an indication of weakly binding peptides. To address this problem the PCR conditions used for DNA amplification were optimised.⁴⁵¹ Concentrations of DMSO, MgCl₂ and primer were varied, and the PCR parameters were adapted accordingly, with a decreased MgCl₂ concentration

subsequently used (Figure 5.4b). A common cause for primer dimer formation is primer design: a new set of shorter primers were made which showed reduced primer dimer formation, with cleaner PCR product (Figure 5.4c). Although primer dimer formation was reduced with these primers, mutations in the constant regions will not be corrected between rounds, which can affect selection outcomes. Methods of removing the unwanted PCR constructs were attempted, with HANDS PCR performed in certain rounds to stop primer dimer formation and favour library formation. In this method, the forward and reverse primer are identical, so any primer dimers formed interact intramolecularly, which stops them being further amplified, favouring amplification of the library.⁴⁵² DNA clean-up by gel excision of the library band and a PCR purification kit did not recover sufficient DNA for the selection to proceed. Purification of the mRNA library was then explored using SEC, separating the library RNA from unwanted constructs with some success. Successful methods of maintaining the library were used in combination, but as they did not address the underlying issues which lead to low enrichment and low DNA recovery, the selections remained ineffective.

5.6.2 Protein immobilisation (on bead vs. solution, on bead activity assay)

Issues can arise from protein immobilisation in biological assays. Tags are added to the C- or N-termini of the protein, changing the protein sequence (in the case of His and avitag used, this is minimal due to the small size of the tag). Issues can arise from alterations to protein structure or blockage of key binding sites, either due to binding conformation or protein density on the bead surface.⁴⁵³ To address these problems, a MALDI-TOF-MS activity assay was run with immobilised protein, showing that PCAF HAT remained active on the beads (Section 5.3.2). This indicated that the active site

remained intact and accessible to AcCoA and H3₁₋₂₁ peptide after streptavidin immobilisation. However allosteric binding sites might be blocked. The mRNA display protocol was changed from incubation of the library with pre-immobilised protein to incubation with protein in solution followed by a pull-down with streptavidin beads. This method may minimise issues with binding site occlusion: on binding to the bead, peptides already bound to the protein may be retained.

5.6.3 Use of AcCoA in the selection buffer (protein folding and activity issues)

It is unclear if PCAF HAT and HAT-BRD are fully folded in the absence of acetyl CoA. No apo crystal structures of PCAF or GCN5 have been published, and screening performed by the Structural Genomics Consortium to displace AcCoA for fragment screening was unsuccessful (unpublished data). If the protein is partially unfolded on bead, this could give access to areas of the protein that are not usually exposed such as hydrophobic patches. A selection against a partially unfolded protein could therefore yield inhibitory peptides which bind to the AcCoA free protein and lock it in an inactive form. Additionally, as PCAF shows flexibility in substrate specificity, acetylation of library peptides in the presence of AcCoA is possible. This could affect the outcome of the screen by modifying the affinity of library peptides through enzymatic acetylation. Peptide modifications cannot be detected as the screen readout is based on DNA tag sequencing and does not inform on possible post translational modifications of the peptides. The screen was therefore first performed in the absence of AcCoA. However, as flexibility in protein structure may discourage peptide binding, AcCoA was included in the buffer in later rounds in an effort to decrease stringency of the selections.

5.6.4 Minimising non-specific binding

The major problem affecting the selections was identified as being non-specific binding of the DNA/RNA tag to PCAF HAT (Figure 5.5b). As a nuclear protein, PCAF is associated with protein complexes involved in translational regulation and has been shown to interact with chromatin.^{97,111,454} It is therefore likely that it has DNA interaction sites. This problem has been encountered with other epigenetic proteins used in mRNA display screens in the Kawamura lab, and has been dealt with by tailoring the buffer conditions using DNA/RNA binding blockers.

Screening conditions which reduce non-specific DNA/RNA interactions were identified by a buffer screen. PBS was explored as an alternative buffer to HEPES, without visible benefit. A gradient of NaCl concentrations was tested with higher NaCl concentration decreasing DNA recovery in Δ AA selection tests (Figure 5.5c, Round #1). However, increasing NaCl concentration above 150 mM resulted in decreased enzyme activity as measured by MALDI-TOF-MS in a parallel activity screen (results not shown). Decreased enzyme activity is likely to indicate decreased binding of the H3₁₋₂₁ peptide to PCAF HAT or a change in conformation. If the enzyme substrate is a weaker binder in these conditions, it may disfavour binding of the mRNA library peptides. NaCl concentration was therefore increased from 50 mM to 150 mM, but no higher.

To prevent protein aggregation and non-specific peptide to protein binding, 0.01% BSA was added to the buffer, though it does not affect non-specific DNA/RNA binding (as shown by the negligible effect of BSA on DNA recovery in Δ AA selection tests, Figure 5.5c, Round #1). Peptides binding to the protein tag (avitag-biotin) or to streptavidin

were removed by preincubating the selection library with beads coated with streptavidin-biotin-avitag prior to incubation with PCAF.

Generic RNA/DNA such as Salmon sperm DNA, herring sperm DNA or yeast RNA (YRNA) and heparin can be included in the buffer to disrupt non-specific binding interactions. Yeast RNA is used as a blocking agent in northern blots and is a co-precipitant for nucleic acids. Salmon sperm DNA is used to block non-specific binding of beads in ChIP assays. Herring sperm DNA is used as a blocking agent for southern blots. Heparin is used in heparin column purification to separate DNA and proteins. The effect of these additives on library DNA recovery were tested separately and in combination. Use of Yeast RNA (1 mg/mL) showed a reduction in DNA recovery in Δ AA conditions for HAT and in normal conditions (with AA included) for both HAT and HAT-BRD. The other additives did not affect DNA recovery (Figure 5.5c, Round #2, Round #3).

The aim is to minimise binding of PCAF HAT to DNA/RNA therefore encouraging binding of protein to the peptide portion of the library. The effects of blocking agents are particularly crucial in the early rounds of selection: if non-specific binders are recovered in the first round of selection they will be disproportionately enriched in later rounds, with genuine protein binders unable to compete and therefore phased out of the library. However, buffer conditions which disfavour peptide binding will also lead to loss of weaker binders which may have been enriched during the selection. It is necessary to find conditions which minimise non-specific binding without stopping genuine peptide binders from being recovered. With the conditions used in the final mRNA selection, DNA corresponding to the library size was recovered over 5 rounds and enrichment was

observed for PCAF His-HAT-BRD-bio. However, for PCAF HAT-bio, loss of library was observed from the first round of selection and previously described methods for preserving the library DNA were unsuccessful (Section 5.6.1).

5.6.5 Protein specific effects

This difference in DNA recovery between the different PCAF constructs can be partly attributed to the relative sizes of the constructs: the larger surface area of His-HAT-BRD-bio gives more room for peptide binding. The presence of the BRD could also affect peptide binding by changing the conformation of the HAT domain in a way that favours binding of peptides over DNA/RNA (a portion of the HAT domain will become inaccessible due to intramolecular interactions with the bromodomain, and the immobilisation conformation is likely to be altered). The BRD itself could also be a more favourable binding target for peptides leading to a higher library recovery.

The structure of PCAF is very well defined around the AcCoA binding site with many protein to co-factor interactions identified in the crystal structure of PCAF with AcCoA bound. The peptide binding site is however less well defined, with two flexible loops flanking a large cleft in which the histone peptide sits (histone peptide binding site determined by comparison with GCN5 bound to H3 peptide, Figure 1.6).⁹⁵ The flexibility of the loops likely contributes to the large substrate scope exhibited by PCAF HAT. This flexibility in the active site may contribute to the lack of strongly binding peptides in the mRNA selections.

The active site flexibility and large substrate scope of PCAF HAT indicates the protein may make a number of weak interactions rather than strongly binding to any particular

residues. For instance, PTM modifications at H3R8 affect acetylation at H3K14 by PCAF HAT through an interaction between H3R8 and Y640 of PCAF HAT.⁴⁷ This exemplifies the importance of allosteric sites in peptide binding and how understanding of crosstalk between PTM modifications can give important information on protein interactions (Chapter 2 and Chapter 3). The cumulative effect of weak binding interaction between PCAF and its substrates is likely difficult to replicate, which may explain the failures of the mRNA selections.

5.6.6 Next generation sequencing results

The NGS sequencing data acquired for Selections 2 and 3 showed that there were significant problems with the selection process. The sequences showed a high incidence of internal stop codons, methionine codons and mutations leading to frameshift mutations (Table 5.2). Sequencing multiple rounds of Selection 2 showed that problematic sequences started appearing by round 2 and that these sequences were enriched throughout the rounds, unlike sequences resembling proper library members. This is likely due to non-specific binding of the protein to the DNA/RNA tag of library members. In a successful selection, mRNA sequences with these issues will be regularly phased out of the selection as they will produce short, linear peptides with little binding potential. There was no overlap in sequences enriched in Selections 2 and 3 for His-HAT-BRD-bio, which can also be an indicator of non-specific binding. These selections had different starting library compositions however, which might account for the lack of overlap in sequences.

In these selections, little genetic material was recovered at the end of each round. Tight peptide binders were not being enriched and non-specific binders are unlikely to be tightly bound to the protein, meaning most of these will be discarded during the washing steps of the selection. Because of this, many rounds of PCR (30+ compared to 20-30 rounds in successful selections) were necessary to produce enough DNA to take into the next round of selection. Excessive PCR on small amounts of DNA is likely to result in an increased mutation rate in the DNA produced, which might explain why the top 10 most abundant sequences in the NGS results present so few sensible peptide sequences. For PCAF HAT-bio, 9 out of the top 10 most abundant sequences present mutations compared with 5 for His-HAT-BRD-bio (Table 5.2). This is likely because the amount of DNA recovered for HAT-bio was consistently lower than for His-HAT-BRD-bio for reasons previously discussed (size, effects of the BRD, see section 5.6.5).

Sequencing results for Selection 5 in which conditions had been optimized to minimize non-specific binding were not significantly improved compared to previous sequencing and did not correspond to NGS results for a successful selection process. The sequences enriched in Selection 5 show a higher incidence of methionine codons but fewer frameshift mutations (Table 5.2). This may indicate a change in binding preferences with optimised buffer conditions. Although isoleucine can be inserted into methionine codons,⁴⁵⁰ such a high incidence of methionine codons is not generally seen in successful selections. The top 10 sequences shown are unlikely to correspond to genuine peptide binders. The results were deemed unreliable, and no peptides were synthesised for testing.

5.6.7 Peptide synthesis and testing

Peptides were synthesised based on the NGS results of Selections 2 and 3 (Section 5.4.2). These were chosen based on enrichment, peptide families and their sequence. Peptide families were determined based on peptides having 5 or more residues matching (frameshift mutations excluded). For the families with the greatest number of peptides members, the most abundant sequences were chosen for synthesis. Sequences chosen included sequences with no unfavourable mutations and the start of frameshift mutation sequences. Sequences with clear mutations were chosen due to their abundance and in order to compare properties of sequences with and without mutations (Table 5.3).

Many of the peptide sequences were insoluble after synthesis and could not be taken forward for cyclisation, purification or testing. Of the 5 peptides chosen from the HAT S3R7 sequencing results, 4 were insoluble (LS1, LS2, LS10 and LS12). All three peptides chosen from frameshift mutations were insoluble (LS7, LS10, LS12). The only insoluble peptide which does not fit in these two categories was the most enriched peptide from HAT-BRD S3R7. These sequences do not present as insoluble sequences as they do not contain a very high number of non-polar residues. As these sequences are likely enriched based on non-specific DNA/RNA binding to the protein it is unlikely that the corresponding peptides favour cyclisation, which might explain the high proportion of insoluble peptides (Table 5.3).

BLI was chosen to test for binding as it is an accessible technique which requires only one of the binding partners to be tagged for immobilisation and can detect binding of

compounds in allosteric sites as well as competitive inhibitors binding in the active site. Of the 6 peptides tested for binding 5 showed no binding response against both PCAF His-HAT, HAT-bio and His-HAT-BRD (Table 5.3, Figure 5.8a-b). For LS8, the binding profile observed was the same with immobilised PCAF HAT-bio and with bare sensors, indicating non-specific binding to the sensors (Figure 5.8b). This binding profile could be due to difference in immobilisation in BLI compared to bead binding in the selections, which is why both His-HAT and HAT-bio constructs were tested. Other, more sensitive techniques could be used to measure binding such as ITC and Alphascreen™. However, ITC uses a large amount of material and the Alphascreen™ method relies on displacement of the native substrate by a binder which would not detect allosteric binding. The peptides were also tested for inhibitory activity against PCAF HAT and PCAF His-HAT-BRD and did not show any inhibitory activity at 100 µM of peptide (Figure 5.8c-d), which indicates if there were binding it would not be in the active site of the enzyme. Because of this, and as further mRNA selections showed that there were significant problems with non-specific binding, no further binding assays were attempted with these peptides.

5.7 Conclusions

Chemical probe development for PCAF HAT has been largely unsuccessful, with no selective and potent inhibitors reported. PCAF has many cognate substrates including histone and non-histone proteins which bind through PPI over a large surface area.^{20,119} There is no clear consensus sequence for PCAF substrate sites, with activity and binding affected by residues up to ten amino acids up- or downstream of the acetylation site on

the substrate. In addition, activity and binding of PCAF to substrates is affected by combinatorial PTMs (Chapter 3,4), and by other context dependent factors such as interaction with other regulatory proteins. Because the binding interactions to substrate sites are not well understood, achieving selectivity between substrates is an issue. In addition, GCN5 is highly homologous to PCAF with $\geq 80\%$ identity, and their active sites are very conserved.⁹⁵ Selectivity for PCAF over GCN5 is therefore also a challenge.

Cyclic peptide mRNA display screens against PCAF were attempted to meet the need for PCAF HAT chemical probes. Cyclic peptides may be able to overcome some of the challenges outlined in identifying selective chemical probes for PCAF as they present a relatively large surface binding area compared to small molecules and can replicate some of the protein-protein binding modes typical of interactions between epigenetic proteins and their cognate substrates. However, these screens can have mixed success depending on the targets, and the screens attempted against PCAF HAT and HAT-BRD were not successful despite attempts to adapt screening conditions for this protein.

There are several factors which might make PCAF a particularly difficult target for these screens. The active site of the protein is flanked by two relatively flexible loops which might not encourage binding.⁹⁵ In addition, PCAF HAT has a wide selectivity for substrates (Chapters 1-4), which means it has many different binding modes for different sequences and 3-D structures, which are difficult to replicate with smaller binding partners. Finally, PCAF appears to interact quite strongly with the mRNA/DNA tag used in the mRNA selection process, which might be expected as it is a nuclear enzyme involved in transcription regulating multiprotein complexes and interacts with

chromatin. Many methods were attempted to overcome the challenges encountered during the screening process, but none of the methods proved effective, and no binding peptides were produced.

Future modifications to the mRNA display selection process to identify hits against PCAF might be to design a custom library with a lysine or acetyl-lysine mimic as a warhead. However, as the major problem encountered with this selection appears to be due to the use of a mRNA/DNA tag, mRNA display may not be suitable for the identification of cyclic peptide binders for PCAF. Other methods such as OBOC, phage display or SICLOPPS may fare better as they do not use DNA tags directly attached to the cyclic peptides being screened.

Chapter 6 Conclusions

6.1 Summary

PCAF and SIRT1 are complementary epigenetic HAT and HDAC proteins, with substantial substrate overlap for both histone and non-histone substrates. The biological and medical relevance of both PCAF and SIRT1 has been extensively reported (Chapter 1). Because of their overlap in substrate selectivity and biological activity, this pair of proteins were chosen for parallel study. Selective and potent chemical probes are available for SIRT1 but not for PCAF, therefore, a greater emphasis was placed on characterising PCAF, in view of developing chemical probes targeting the HAT domain of PCAF.

PCAF and SIRT1 were expressed and purified (Chapter 2), and their substrate selectivity for known histone substrates was confirmed via a MALDI-TOF-MS and MALDI-TOF/TOF-MS assay developed for monitoring acetylation and deacetylation activity via a mass change of ± 42 Da. PCAF was shown to catalyse the acetylation of H3₁₋₂₁ at K14 followed by K9 (and to a much lesser extent K18), whilst SIRT1 was shown to catalyse the deacetylation of H3₁₋₂₁K9Ac, H3₁₋₂₁K14Ac and H3₁₋₂₁K9AcK14Ac.

The activity of PCAF was further probed (Chapter 2, results summarised in Table 6.2) with different acyl-CoA cofactors, confirming reports that the rate of acylation of H3₁₋₂₁ peptides by PCAF is higher with AcCoA, but PrCoA and BuCoA are also accepted as acyl donors. Equivalent results were obtained with PCAF HAT and PCAF HAT-BRD, showing that the BRD of PCAF does not affect cofactor selectivity. Commercially available HAT inhibitors were tested for inhibition of PCAF HAT and HAT-BRD acetylation activity in the

MALDI-TOF-MS assay developed. The assay was used to identify and to rank currently available PCAF HAT inhibitors in terms of potency against PCAF HAT and HAT-BRD acetylation activity of a H3₁₋₂₁ substrate. The inhibitors tested were less potent against PCAF HAT-BRD compared with PCAF HAT. In addition, garcinol did not affect PCAF HAT and HAT-BRD equally, showing inhibition of PCAF HAT and activation of PCAF HAT-BRD (not verified in an alternative assay). This indicated that binding of inhibitors to PCAF might be affected by the different structure of PCAF HAT-BRD compared with the single HAT domain.

The scope of histone substrate selectivity of PCAF and SIRT1 was further explored with a rapifleX MALDI-TOF-MS HTS screen developed, in which commercial histone peptide libraries with PTMs (acetylation, phosphorylation and single or multiple methylations) were screened against the enzymes, and acetylation/deacetylation of the peptides was measured (Chapter 3). In this screen, SIRT1 showed no deacetylation selectivity between histone peptides with combinatorial acetylation PTMs, fully deacetylating all positions. PCAF HAT and HAT-BRD showed selective acetylation of histone peptides with 0-2 enzymatic acetylation events observed on peptides corresponding to sections of H2a, H2b, H3 and H4. PTMs on the histone peptides affected the acetylation activity and selectivity of PCAF HAT/HAT-BRD.

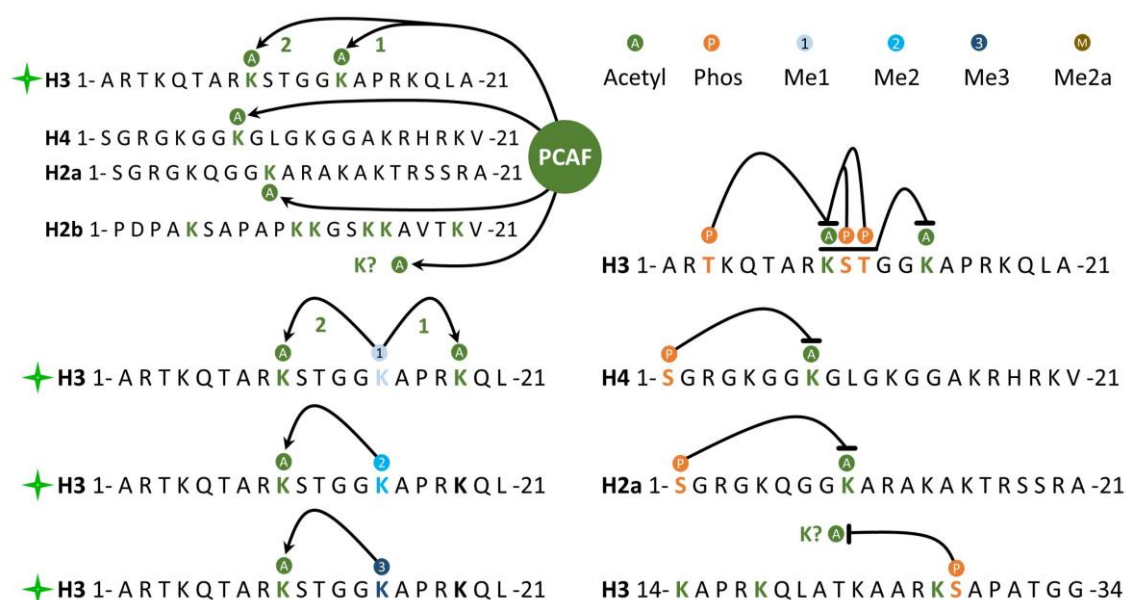


Figure 6.1: Visual summary of effects of PTM crosstalk on the acetylation catalytic activity of PCAF.

Visual representation of results obtained in the rapifleX MALDI-TOF-MS screen listed in Table 6.1. Results with a green star have been validated in MALDI-TOF-MS and MALDI-TOF/TOF-MS assays using quality-controlled peptides. Other results require further validation.

The main results of the MALDI-TOF-MS HTS against PCAF are summarised in Table 6.1 and Figure 6.1. PCAF HAT and HAT-BRD mainly showed the same activity profiles, showing that the BRD does not generally affect the activity and selectivity of the HAT domain of PCAF. Exceptions are noted in the case of H3₁₋₂₁T11Phos and H3₁₄₋₃₄S28Phos in which the enzyme activity of PCAF HAT-BRD was decreased by the PTM to a greater extent than that of PCAF HAT. The previously reported acetylation sites of PCAF on H3₁₋₂₁ and H4₁₋₂₁ were confirmed, and a previously unreported acetylation site of PCAF on H2a₁₋₂₁ was identified. Several PTMs were shown to inhibit PCAF acetylation activity, when compared with control peptides. Conversely, the K14Me1 PTM of H3₁₋₂₁K14Me1 was observed to activate the PCAF enzymes, leading to a much higher rate of H3₁₋₂₁K14Me1 peptide acetylation than was observed for H3₁₋₂₁K14Me2/3 peptides. Results for H3₁₋₂₁K14Me1/2/3 were validated using quality-control peptides with MALDI-TOF-MS. MALDI-TOF/TOF-MS was used to further explore the effect of K14Me1

on PCAF activity of H3₁₋₂₁ peptides, showing acetylation of H3₁₋₂₁K14Me1 by PCAF at K18 followed by K9. In equivalent assays, H3₁₋₂₁K14Me2/3 peptides were only acetylated at K9. These results indicate that the H3₁₋₂₁K14Me1 peptide has a different binding mode to PCAF than H3₁₋₂₁K14Me2/3 peptides, which changes the activity of the enzyme (Chapter 4).

Peptide substrates	Main results
H2a ₁₋₂₁	H2aK9 was identified as the acetylation site by PCAF (Needs checking by MS/MS)
H2b ₁₋₂₁	H2b acetylation site/s could not be identified
H3 ₁₋₂₁ S10 and T11 Phos	• Phosphorylation of S10 and T11 inhibited PCAF acetylation of H3K9 • Alone, S10Phos did not affect PCAF acetylation of H3K14 • T11Phos alone and in combination with S10Phos and K9Ac/Me3 inhibited acetylation of H3K14 by PCAF. This effect was more pronounced with HAT-BRD than HAT.
H3 ₁₋₂₁ T3Phos	H3T3Phos inhibited H3K9 acetylation by PCAF (The H3T3Phos PTM was found in combination with H3R2Me2a and H3K4Me3, but these did not affect PCAF activity)
H4 ₁₋₂₁ S1Phos	Alone, S1Phos inhibited H4K8 acetylation by PCAF
H2a ₁₋₂₁ S1Phos	H2aS1Phos inhibited acetylation of H2a by PCAF (Acetylation position on H2a thought to be H2aK9)
H3 ₁₄₋₃₄ S28Phos	H328Phos inhibited H3 peptide acetylation by PCAF HAT-BRD to a greater extent than acetylation by PCAF HAT The location of the two acetylation sites could not be determined
H3 ₁₋₂₁ K14Me1 and H3 ₈₋₂₈ K14Me1	H3K14Me1 peptide was acetylated at H3K18 followed by H3K9 by PCAF HAT/HAT-BRD H3K14Me2/3 peptides were only acetylated at H3K9 Only H3K14Me1 activated PCAF acetylation of H3K18 pointing to a selective binding mode

Table 6.1: Summary of main results obtained in the rapifleX MALDI-TOF-MS HTS against PCAF HAT and HAT-BRD (Chapter 3).

The rapifleX MALDI-TOF-MS HTS was developed using commercial libraries of histone peptides, as a MALDI-TOF-MS assay had successfully been developed to monitor PCAF and SIRT1 activity for H3₁₋₂₁ peptides. To widen the scope of the assay, peptides derived from HIF1 α , a non-histone protein substrate of both PCAF and SIRT1, were synthesised for testing in the MALDI-TOF-MS assay (Chapter 4). No acetylation of the HIF1 α peptides

was observed with PCAF HAT or HAT-BRD. Deacetylation of the HIF1 α peptides by SIRT1 was observed. However, the selectivity of SIRT1 for the HIF1 α peptides did not correspond to reported biological selectivity. Characterisation and substrate validation for PCAF and SIRT1 carried out in this study are summarised in Table 6.2.

Protein	PCAF HAT/HAT-BRD	SIRT1
Control histone substrate	H3 ₁₋₂₁	H3 ₁₋₂₁ K9Ac, H3 ₁₋₂₁ K14Ac
Site of activity on control substrate	K14>K9>K18	K9>K14
Compounds with inhibitory activity observed in MALDI-TOF-MS	For PCAF HAT: Anacardic acid > Embelin > C646 > Garcinol For PCAF HAT-BRD: Anacardic acid > Embelin	N/A
Inhibitor chosen as control	Anacardic acid	N/A
Cofactor selectivity	AcCoA > PrCoA > BuCoA	N/A
rapifleX MALDI-TOF-MS HTS results	Certain PTMs affect the activity and selectivity of PCAF HAT for histone peptide substrates (Table 6.1)	No selectivity observed between histone peptides (full deacetylation of all substrates)
HIF1 α peptide selectivity	HIF1 α ₅₂₂₋₅₄₂ HIF1 α ₆₆₄₋₆₈₄ HIF1 α ₇₀₀₋₇₂₁ No acetylation activity observed	HIF1 α ₅₂₂₋₅₄₂ K523Ac (532) HIF1 α ₆₆₄₋₆₈₄ K674Ac (664) HIF1 α ₇₀₀₋₇₂₁ K709Ac (709) Deacetylation observed 532 > 664 = 709

Table 6.2: Summary of results on PCAF HAT/HAT-BRD and SIRT1 enzyme selectivity in this study.

Results of enzyme characterisation and substrate scope carried out using MALDI-TOF-MS and MALDI-TOF/TOF-MS methods are shown for PCAF and SIRT1.

As there are no chemical probes currently available which selectively and potently target PCAF HAT, efforts were made to develop PCAF HAT selective cyclic peptide chemical probes (Chapter 5). mRNA display of cyclic peptides was chosen as a method for chemical probe development. Cyclic peptides have advantageous chemical properties (size, flexibility, 3D structure) for targeting PPI involved in epigenetics, and mRNA display of cyclic peptides has previously been used to successfully target KDMs, PHDs and

Sirtuins.¹⁻⁴ The mRNA selection against PCAF HAT/HAT-BRD was not successful. Non-specific binding of PCAF HAT/HAT-BRD to the mRNA/DNA hybrid tag used in the selection was determined as the main problem affecting the screens. Although steps were taken to reduce non-specific binding, these were not sufficient, and no binding peptides were identified.

6.2 Conclusions and future work

Epigenetic PTMs, and the proteins that regulate them, act in a very context dependent manner. PCAF and SIRT1 are both known to take part in large protein complexes in a biological context, and both have >50 histone and non-histone protein substrates, which illustrates the importance of PPI in regulating protein function.²⁰ In addition, the nature and location of epigenetic marks found on histones differs depending on chromosomal location, which allows for a regulation of epigenetic proteins based on location in the nucleus.

The MALDI-TOF-MS assay developed is an effective way of characterising enzyme activity for a single histone peptide substrate, as was shown in Chapter 2 with the screen of HAT inhibitors against PCAF HAT and HAT-BRD. However, this assay, like most assays developed to monitor enzyme activity of epigenetic enzymes, does not take into account context dependent PTM crosstalk. For this reason, the regulatory role of PTM crosstalk on epigenetic enzymes is not well understood. The rapifleX MALDI-TOF-MS HTS developed is an effective method for identifying interactions between histone peptide PTMs and epigenetic enzymes, as illustrated by the screen against PCAF HAT/HAT-BRD.

The rapifleX MALDI-TOF-MS HTS was developed using commercially available histone peptide libraries with PTMs, screened against two constructs PCAF (HAT and HAT-BRD) and one construct of the HDAC SIRT1. The HTS identified PTM modifications with activating or inhibitory effects towards the acetylation activity of PCAF HAT. In addition to providing valuable insights into the selectivity of the PCAF enzyme, these results are a proof of concept showing that this screen could be more widely applied to mapping the effect of histone PTM crosstalk on diverse epigenetic enzymes, using a variety of peptide libraries. Other methods have been developed to map PTM crosstalk, however, the HTS described here is one of the few methods which screens for enzyme activity rather than binding and has some advantages over previously developed screens (no immobilisation, does not rely on availability of antibodies or on radiolabelling, can be tag-free, potential for rapid identification of substrate sites).

Although efforts to expand the scope of the screen to non-histone peptides against PCAF and SIRT1 using HIF1 α as a template were not the most promising, this is a possible avenue for future work, exploring whether different substrates might be more amenable to the screening method (*e.g.* p53). In addition, some differences were identified between the enzymatic activities of the HAT and HAT-BRD constructs of PCAF towards histone peptides in the HTS. This method might therefore prove a useful tool in deconvoluting the respective roles of different domains in proteins which contain multiple epigenetic modules.

Not all results obtained with the HTS against PCAF HAT and HAT-BRD have been fully verified, and these should be further studied to validate the information obtained.

However, the activating effect of the K14Me1 PTM of H3₁₋₂₁ on PCAF acetylation activity has been thoroughly investigated with MALDI-TOF-MS and MALDI-TOF/TOF-MS of purified peptides. This showed a large difference in specific activity of PCAF for substrates bearing PTM variations and showed an increase of K18 acetylation on the H3₁₋₂₁K14Me1 peptide by PCAF HAT and HAT-BRD compared with H3₁₋₂₁K14Me2/3 peptides. As histone peptides are a model for histone proteins, further validation of these observations on full length histone proteins, nucleosomes, and in cells would be necessary to determine biological relevance. H3K14 methylation (mono-, di- and tri-) is an endogenous eukaryotic PTM, though found at relatively low levels and not well characterised. These findings might therefore be biologically relevant and merit further investigation. PCAF has a large range of substrates which do not have a consensus sequence, and the systems regulating the selectivity and activity of PCAF are not well understood. Mapping PTM crosstalk interactions with PCAF might help pinpoint how it is regulated.

A limiting factor in understanding epigenetic systems is the lack of selective and potent chemical probes targeting epigenetic proteins of interest. To address this issue, mRNA display of cyclic peptides was attempted, targeting the HAT domain of PCAF. Although efforts in this direction were unsuccessful, cyclic peptide chemical probe development may still be an effective avenue for targeting PCAF HAT. The HTS against PCAF HAT showed that histone peptides with PTMs can have a large activating or inhibitory effect on PCAF HAT activity, at concentrations much lower than those used for PCAF inhibition by small molecules (*e.g.* anacardic acid). Problems encountered with the mRNA screen were linked to non-specific binding between PCAF and the tag used in the screen.

Alternative methods for screening cyclic peptides which do not require an mRNA/DNA tag might therefore be effective in identifying a PCAF specific chemical probes.

Because cellular and animal studies are dynamic and complicated models compared to peptide studies looking at specific PTMs, and due to a lack of appropriate tools, protein characterisation *in vitro* often does not correspond to results obtained in cellular and animal studies. Although the HTS designed remains a simplified model, it is a first step in increasing the complexity of *in vitro* models used to characterise proteins. The ability to identify crosstalk between PTM and proteins and between protein domains in a model system can help increase our general understanding of epigenetic PTMs as an interconnected regulatory system rather than as isolated interactions. However, an understanding of PTM crosstalk and protein activity in cells is dependent on the availability of robust chemical biology tools, such as selective chemical probes. Efforts must therefore be made in both of these fields to be able to identify dysregulation of epigenetic systems and open avenues for epigenetic modulation of disease states.

Chapter 7 Materials and Methods

7.1 Materials used

All primers were purchased from Sigma-Merck (99% purity, desalted and lyophilised) and dissolved in water up to 100µM.

General chemicals and solvents were purchased from Sigma Aldrich, Merck, ThermoFisher Scientific, Fluorochem, Cambridge Reagents and Alpha Aesar and used without further purification.

TEV protease was provided by Dr Ilaria Pettinati.

L-Moses, D-Moses, GSK4027 and GSK4028 were provided by Dr. Moses Moustakim and Prof. Paul Brennan.

Large scale expression and purification of PCAF His-HAT-BRD protein was carried out by Eidarus Salah.

7.2 General methods

H₂O used was Type 1 H₂O (T1H₂O, resistivity of >18 MΩ/cm, conductivity of <0.056 µS/cm and <50 ppb of Total Organic Carbons), taken from an Elga Milli-Q™ system or an Avidity science system with a Duo 0.2 µM filter. Glassware was cleaned using a Smeg Camlab dishwasher and autoclaved using a Systex DE-65 autoclave (121°C, 15 min). All cell work was carried out using a Class II biological safety cabinet (Labard). Centrifuges used were ThermoFisher Scientific Heraeus Pico 17 centrifuge, Heraeus Pico 21 centrifuge, Heraeus Megafuge 16R centrifuge, Beckman Coulter Avanti J-E®

centrifuge, Avanti JXN-26 standing centrifuge depending on volume, location and speed needed.

7.2.1 Media preparation

Media was prepared by weighing media powder or granules (see Table 7.1 for mass for each media type), mixing with T1H₂O and autoclaved before use using a Systex DE-65 autoclave.

Media Type	Brand	Mass (g) per 1L H ₂ O
Lysogeny broth (LB)	ThermoFisher Scientific	25
Terrific Broth (TB)	ThermoFisher Scientific	47.6 + 4 mL Glycerol
Autoinduction TB (AIM TB)	Fortmedium	55.85
2YT	ThermoFisher Scientific	31
LB Agar	Applichem Panreac ITW companies	35

Table 7.1: Media used in microbiology and protein production work.

7.2.2 Buffer preparation

Buffers were prepared by weighing components out and dissolving in T1H₂O. pH of buffers was adjusted using a Mettler Toledo MP220 pH Meter with NaOH (2 M) or HCl (2 M) solutions added dropwise with mixing. Where needed, buffers were filtered with an 0.2 µM filter (for SEC purification and HPLC).

7.2.3 Antibiotic preparation and concentrations

Antibiotics were prepared at 1000 times the necessary concentration and stored at -20°C until use. Concentrations used in cell culture were the following: Kanamycin 50 µg/mL in T1H₂O, Ampicillin 100 µg/mL in T1H₂O, Spectinomycin 100 µg/mL in T1H₂O, chloramphenicol 34 µg/mL in EtOH.

7.2.4 Bacterial Cell Types

Plasmid Replication cell types: XL-10 Gold, TOP10

Protein Production cell types: Rosetta2 (DE3) pLysS, Rosetta pBirA

Notes from the manufacturer (New England Biolabs):

Rosetta2 cells are BL21 derivatives designed to enhance the expression of target eukaryotic proteins that contain codons rarely used in *E. coli*. These strains supply tRNAs for 7 codons (AGA, AGG, AUA, CUA, GGA, CCC, and CGG) on a suitable chloramphenicol-resistant plasmid. The tRNA genes are under the control of their native promoters.

DE3 indicates that the host is a lysogen of λ DE3, carrying a chromosomal copy of the T7 RNA polymerase gene under control of the lacUV5 promoter. DE3 strains are suitable for production by induction with IPTG of target proteins with corresponding gene cloned in pET vectors which contain a T7 promoter.

BL21(DE3)pLysS also contains a pLysS plasmid, carrying the gene encoding T7 lysozyme. Before IPTG induction, the background expression level of target genes with a T7 promoter is lowered by T7 lysozyme. Following IPTG induction, T7 lysozyme does not affect the levels of expression of target genes.

Rosetta pBirA cells are Rosetta2 cells transformed with a pBir1 plasmid for expression of biotin ligase by IPTG induction, and subsequent biotinylation of expressed recombinant proteins which contain an avitag sequence.

7.2.5 Transformations of plasmids into *E. Coli* cells

Competent *E. Coli* cells (10 μ L) were incubated with plasmid DNA (1 μ L, 50 to 100 ng/ μ L) on ice (30 min) then subjected to heat shock (30 s at 42°C). The cells were cooled on ice (2 min) before addition LB media (1 mL) and were incubated for a recovery period (1 h, 37°C, 200 rpm). LB Agar plates with appropriate antibiotic were prepared. Cells (50 and 150 μ L) was spread onto two agar plates. The remaining media was centrifuged until a pellet was visible, the supernatant was removed, and the cells were resuspended in LB media (50 μ L) and spread onto a third plate. The plates were incubated overnight (37°C) then visually inspected for colonies. The plates were then stored at 4°C until needed.

7.2.6 Starter Cultures

Media of choice (20 mL) supplemented with the appropriate antibiotics was inoculated with a pipette tip scraping from a colony on an agar plate culture. Cells were shaken overnight (37°C, 200 rpm).

7.2.7 Miniprep plasmid purification

Mini preps were performed with a NucleoSpin Plasmid (Mini) according to the manufacturer's (Macherey-Nagel) instructions.

7.2.8 Agarose gel preparation

%age of gel	Agarose used (g)
0.8	0.4
1	0.5
2	1
3	1.5

Table 7.2: Quantity of agarose needed to make up 50 mL of agarose gel at specified percentage.

Agarose was weighed out and dissolved in 1 × Tris/Borate/EDTA (TBE) buffer (50 mL), microwaved until clear and cooled. SYBR Safe (5 µL, Thermofisher Scientific) was added to the mix which was poured out and left to set (30 min). Samples were mixed with 6 × DNA Gel Loading Dye (Fermentas). The gels were run in 1 × TBE buffer at 150 V for 30 min. Gels were imaged using a Bio Rad ChemiDoc™ MP Imaging system. The gels were visualised with Image Lab 5.0 using SYBR Safe mode.

7.2.9 12% Acrylamide gel preparation

Urea (2.1 g), 5 × TBE (1200 µL), 40% acrylamide (1800 µL), T1H₂O (1000 µL), 10% APS (60 µL), TEMED (4 µL) were measured out and mixed by vortexing. The gel was left to set (30 min). The gel was loaded with low range ssRNA ladder (1 µL, New England Biolabs), and prepared samples (15 µL total volume, 7.5 µL sample, 7.5 µL 5 × RNA Loading Dye, Thermofisher Scientific). The gel was stained in SYBR gold (3 µL in 30 mL H₂O) for 10-15 min. Gels were imaged using a Bio Rad ChemiDoc™ MP Imaging system and viewed with Image Lab 5.0 using SYBR Gold mode.

7.2.10 Running of Protein gels

All gels shown are NuPage™ 12%/4-12% Bis-Tris Gel, Invitrogen. Gels were run with PageRuler™ Plus Prestained Protein Ladder, 10 to 250 kDa (3 µL). Protein sample to be

loaded (10 μ L) was prepared with sample (6.5 μ L), NuPage^R LDS Sample Buffer (2.5 μ L, 4 $\times$ concentration stock, Invitrogen), NuPage^R Sample reducing agent (1 μ L, 10 $\times$ concentration stock, Invitrogen). Gels were run with 1 $\times$ MES SDS running buffer at 200 V for 40 min. The gels were then soaked in InstantBlueTM (30 min, Expedon), imaged using a Bio Rad ChemiDocTM MP Imaging system visualised by colorimetric imaging on ImageLab 5.0 (Bio-Rad).

7.2.11 Ni²⁺ IMAC column preparation

Binding buffer: 10 mM HEPES, 500 mM NaCl, 5% glycerol, 0.5 mM tris(2-carboxyethyl)phosphine (TCEP), 20 mM imidazole, pH 7.6

Elution Buffer: 10 mM HEPES, 500 mM NaCl, 5% glycerol, 0.5 mM TCEP, 500 mM imidazole, pH 7.6

IMAC (immobilised metal affinity chromatography) Sepharose TM 6 Fast Flow resin (1 mL) was poured into a manual column. The storage solution was drained, and the resin was washed with T1H₂O (5 mL). NiSO₄ (0.2 M, 10 mL) was added to the resin and left to incubate at room temperature for 30 min. The resin was then drained and washed with T1H₂O (20 mL $\times$ 2). For purification the column was equilibrated with the binding buffer. Once the sample has been eluted, the column was washed with elution buffer (20 mL $\times$ 2) followed by T1H₂O (20 mL $\times$ 2). The column was then stored in 20% EtOH in H₂O (3 mL) at 4°C for reuse.

To reuse the column, the storage solution was drained, washed with T1H₂O (20 mL $\times$ 2) then equilibrated with binding buffer (20 mL $\times$ 2).

7.2.12 Polymerase Chain Reaction (PCR)

PCR was carried out using Techne TC-312 or Bio-rad C1000 Touch Thermal Cyclers. PCR reactions were performed in standard thin wall 100 μ L PCR tubes or 100 μ L x 2-8 tube PCR strips. Reaction concentrations and parameters were tailored to the experiments. Polymerases used were BioTaq (Bioline) for general DNA amplification in the production of RNAs or Q5 (NEB) for mRNA display related PCR.

7.3 Ligation independent cloning (LIC)

Construct inserts were cloned from p1-His-HAT and p3-His-HAT-BRD PCAF constructs, donated by the SGC (confirmed by Sanger sequencing), using forward and reverse primers detailed in Table 7.3, with a procedure by Dr. Oliver Coleman followed.

Insert Preparation for one 25 μ L PCR reaction: DNA template (2.5 μ L, $\sim$ 5ng/ μ L, < 1000 ng), forward primer (1.25 μ L of 10 μ M), reverse primer (1.25 μ L of 10 μ M), 5X Q5 Buffer (5 μ L), dNTPs (0.5 μ L of 10 mM), Q5 polymerase (0.25 μ L, NEB), T1H₂O (14.25 μ L, for 25 μ L final volume). DNA purification was carried out by running on 2% agarose gel (80 V for 30 minutes) and excising a band of correct weight, which was purified using a gel extraction kit (Qiagen).

Generation of cohesive ends: PCR product DNA ($\sim$ 5 ng/ μ L, 7.15 μ L), 10x T4 polymerase buffer (1 μ L), dCTP (25 mM, 1 μ L), DTT (100 mM, 0.5 μ L), BSA (NEB, 10 mg/ml, 0.1 μ L), T4 DNA polymerase (Novagen or NEB, 0.25 μ L). Mixture was incubated (30 min, 22°C) followed by a second incubation (20 min, 75°C) to inactivate the enzyme (preheat the thermocycler lid to 90°C).

Vector Preparation for 100 µl reaction volume: For pNICBio3 restriction enzyme Bsa1 is used. T1H₂O (volume adjustable according to DNA concentration), NEB Cutsmart buffer (10 µL), plasmid (5 µg, volume determined according to DNA concentration), Bsa1 restriction enzyme (3 µL, 30 units). The reaction is incubated (2h, 50°C).

Purification of the vector DNA was carried out using a PCR purification kit (Qiagen), eluted in T1H₂O (50 µL). Samples were prepared with cut purified vector (1 µL), T1H₂O (4 µL), DNA loading dye (1 µL) to verify the band by running an 0.8% agarose gel (80 V, 1 h, imaged with UV).

Generation of cohesive ends: In a PCR tube, mix: T1H₂O (11 µL), Digested plasmid (25 µL), 10 × T4 DNA polymerase buffer (5 µl, Novagen or NEB buffer 2), dGTP (25 mM, 5 µL), BSA (NEB, 20 mg/ml, 0.25 µL), DTT (filtered, 100 mM, 2.5 µL), T4 DNA polymerase (1.25 µL, LIC qualified, Novagen or NEB). Reaction is incubated (30 min, 22°C), followed by a second incubation (20 min, 75°C) to inactivate the enzyme (preheat the thermocycler lid to 90°C).

Annealing and Transformation: T4 treated vector DNA (1 µL) and T4 treated insert DNA (2 µL) were mixed (20 min, 20°C) prior to general transformation procedure. Cells were plated onto sucrose (5% w/v) LB agar plates containing the appropriate antibiotics. Corresponding plasmids were obtained by miniprep (10 mL), and were sequenced by Sanger sequencing (SourceBioscience).

	Bp	Name	Sequence
Forward Primer HAT	32	FP_c031_PCAF_LIC	TACTTCCAATCCATGGTAATTGAATTTACGCT
Reverse Primer HAT	37	RP_c031_PCAF_LIC	TATCCACCTTTACTGCTTTCTGTGTACGGGATCCGTG
Forward Primer HAT- BRD	33	FP_c024_PCAF_LIC	TACTTCCAATCCATGGGGGATATTCCGATGGAA
Reverse Primer HAT- BRD	38	RP_c024_PCAF_LIC	TATCCACCTTTACTGCTCTTGTCAATTAATCCAGCTTC

Table 7.3: Primers used in LIC cloning for PCAF HAT and HAT-BRD constructs.

7.4 Protein expression and purification

7.4.1 General Protocol for His-Tagged Proteins

Protein purification was carried out in the cold room (4°C) or on ice, unless stated otherwise. Buffers were prepared with filtered (0.2 µm) T1H₂O prior to use, and stored at 4°C (Table 7.4 for PCAF purification buffers and Table 7.5 for SIRT1 purification buffers). Isopropyl-D-thiogalactopyranoside (IPTG) was used to induce protein expression at 0.5 mM final concentration. Cells were grown by incubation at 37°C or 18°C, carried out in a New Bruker Scientific Innova® 44 or a MaxQ8000 ThermoFisher Scientific shaker incubator. Protein was extracted either via cell disruption by sonication on a Banderlin Sonoplus instrument or a Sonics® Vibracell™ model VCX500 or by cell disruption by pressure was carried out using a Constant Cell Disruption system.

Manual HisTrap purification was carried out using gravity IMAC columns loaded with Ni²⁺ (GE Healthcare) were used for His tag purification. Resin (IMAC Sepharose 6 Fast Flow, GE Healthcare) was prepared and charged with Nickel according to the manufacturer's instructions. The lysate (5-10 mL) was applied to 2 mL of resin in a disposable column (15 ml) and washed with lysis buffer (4 × 15 mL). The target protein

was eluted in IMAC elution buffer (15 mL with increased imidazole concentration with 2 final washes at 500 mM imidazole) and concentrated to 2 mL with a 10 kDa MWCO centrifugal concentrator (Amicon Ultra, Millipore).

Cleavage of His-tags was carried out using TEV protease (donated by Dr Ilaria Pettinati) was added to the concentrated protein following a Ni^{2+} IMAC in a 10:1 ratio of Protein:TEV (mg, determined by NanoDrop™) and incubated at 4°C for 12 hours, with dialysis carried out in SnakeSkin™ dialysis tubing with a 3500 molecular weight cut off (MWCO), 3.7 mL/cm, 34 mm width (ThermoFisher Scientific) in 4 L of low imidazole buffer. Cleavage was assessed by SDS-PAGE gel.

SEC was carried out on AKTA™ Start and AKTA™ Pure systems either at room temperature or at 4°C using Unicorn 6.4.1 software. Buffers for SEC were prepared in T1H₂O, filtered with 0.2 µM filter and degassed. Columns used were HiLoad™ 16/600 Superdex™ columns with either 75 pg or 200 pg depending on protein size (GE Healthcare). The column was washed with T1H₂O (1 column volume = CV) and SEC buffer (1 CV). The protein sample was applied to the SEC column in a 5 mL loop and run in SEC buffer. Fractions (4 mL) were collected automatically and then analysed by SDS-PAGE gel. Staining with InstantBlue™ was performed for 15 minutes and fractions containing the target band (> 90%) were pooled.

Protein containing factions were pooled and centrifugally concentrated using filters with a 10 kDa MWCO (Utracell® Amicon Ultra). The final protein concentration was estimated by an average of six NanoDrop™/PHERAstar® FSX LVis readings. Concentrated protein

was aliquoted on ice into PCR strips that were snap frozen in liquid N₂ for long-term storage at -80°C.

7.4.2 Plasmid encoded protein sequences

Expressed protein amino acid sequences are shown with His-tag and avitag sequences underlined.

PCAF p1-His-HAT:

MHHHHHHSSGVDLGTENLYFQSMVIEFHVVGNLSNQPNKKILMWLVGLQNVFSHQLPRMPK
EYITRLVFDPKHKTLALIKDGRVIGGICFRMFPSQGFTIVFCAVTSNEQVKGYGTHLMNHLKEYHIK
HDILNFLTIADEYAIGYFKKQGFSGKEIKIPKTKYVGYIKDYEGATLMGCELNPRIPYTE

PCAF p2-His-HAT-Bio:

MHHHHHHSSGVDLGTENLYFQSMVIEFHVVGNLSNQPNKKILMWLVGLQNVFSHQLPRMPK
EYITRLVFDPKHKTLALIKDGRVIGGICFRMFPSQGFTIVFCAVTSNEQVKGYGTHLMNHLKEYHIK
HDILNFLTIADEYAIGYFKKQGFSGKEIKIPKTKYVGYIKDYEGATLMGCELNPRIPYTESSKGGYGLND
IFEAQKIEWHE

PCAF p3-His-HAT-BRD:

MHHHHHHSSGVDLGTENLYFQSMGDIPMELINEVMSTITDPAAMLGPETNFLSAHSARDEAARL
EERRGVIEFHVVGNLSNQPNKKILMWLVGLQNVFSHQLPRMPKEYITRLVFDPKHKTLALIKDGR
VIGGICFRMFPSQGFTIVFCAVTSNEQVKGYGTHLMNHLKEYHIKHDILNFLTIADEYAIGYFKKQ
GFSKEIKIPKTKYVGYIKDYEGATLMGCELNPRIPYTEFSVIIKKQKEIKKLIERKQAAQIRKVYPGLSCF
KDGVRQIPIESIPGIRETGWKPSGKEKSKEPRDPDQLYSTLKSILQQVKSHQSAWPFMEPVKRTEAP
GYEYVIRFPMDLKTMSERLKNRYVVSKKLFMADLQRVFTNCKEYNPPESEYYKCANILEKFFFSKIKE
AGLIDK

PCAF p4-His-HAT-BRD-Bio:

MHHHHHHSSGVDLGTENLYFQSMGDIPMELINEVMSTITDPAAMLGPETNFLSAHSARDEAARL
EERRGVIEFHVVGNLSNQPNKKILMWLVGLQNVFSHQLPRMPKEYITRLVFDPKHKTLALIKDGR
VIGGICFRMFPSQGFTIVFCAVTSNEQVKGYGTHLMNHLKEYHIKHDILNFLTIADEYAIGYFKKQ
GFSKEIKIPKTKYVGYIKDYEGATLMGCELNPRIPYTEFSVIIKKQKEIKKLIERKQAAQIRKVYPGLSCF
KDGVRQIPIESIPGIRETGWKPSGKEKSKEPRDPDQLYSTLKSILQQVKSHQSAWPFMEPVKRTEAP
GYEYVIRFPMDLKTMSERLKNRYVVSKKLFMADLQRVFTNCKEYNPPESEYYKCANILEKFFFSKIKE
AGLIDKSSKGGYGLNDIFEAQKIEWHE

SIRT1 p5-SIRT1.1:

MRGSHHHHHHGSMADAAAAAGGEQEAQATAAAGEGDNPGPLQGSPREPPLADNLYDEDDDD
EGEEEEAAAAAIGYRDNLLFGDEIITNGFHSCESDEEDRASHASSDWTTPRPRIGPYTFVQQHLM
GTDPRITLKDLLPETIPPELDDMTLWQIVINILSEPPKRKRKDINTIEDAVKLLQECKKIIVLTGAGV
SVSCGIPDFRSRDGIYARLAVDFPDLPDPQAMFDIEYFRKDPRPFFKFAKEIYPGQFQPSLCHKFIAL
SDKEGKLLRNYTQNIDTLEQVAGIQRIIQCHGSFATASCLICKYKVDCEAVRGDIFNQVVP RCP RCP
ADEPLAIMKPEIVFFGENLPEQFHRAMKYDKDEVLLVIGSSLKVRPVALIPSSIPHEVPQILINREPL
PHLHFDVELLGDGCDVIINELCHRLGGEYAKLCCNPVKLSEITEKPPRTQKELAYLSELPPTPLHVSEDS
SSPERTSPDSSVIVTLLDQAAKSNDLDVSESKGCMEEKPQEVQTSRNVESIAEQMENPDLKNV
GSSTGEKNERTSVAGTVRKCPNRVAKQISRRLDGNQYLFLPPNRYIFHGAEVYSDEDDVLSSS
SCGSNSDSGTCQSPSLEEPMEDESEIEEFYNGLEDEPDVPERAGGAGFGTDGDDQEAINAISVKQ
EVTDMNYPSPNKSCRPAAKLN

7.4.3 Buffers used for protein purification

Lysis buffer	10 mM HEPES, 500 mM NaCl, 5% glycerol, 0.5 mM TCEP, 20 mM imidazole, pH 7.6
IMAC elution buffers	10 mM HEPES, 500 mM NaCl, 5% glycerol, 0.5 mM TCEP, 20/40/50/70/100/150/300/500 mM imidazole, pH 7.6
Low imidazole buffer	20 mM HEPES, 500 mM NaCl, 5% Glycerol
Wash buffer	10 mM HEPES, 500 mM NaCl, 5% glycerol, 0.5 mM TCEP, 20 mM imidazole, pH 7.6
SEC buffer	10 mM HEPES, 500 mM NaCl, 5% glycerol, 0.5 mM TCEP, pH 7.6

Table 7.4: Buffers for protein purification of PCAF proteins

Lysis buffer	50 mM Tris·HCl, 250 mM NaCl, 20 mM imidazole, 0.5 mM TCEP, pH 8
IMAC elution buffers	50 mM Tris·HCl, 250 mM NaCl, 0.5 mM TCEP, 20/50/100/150/300/500 mM imidazole, pH8
SEC buffer	50 mM Tris·HCl, 250 mM NaCl, 5% glycerol, pH8

Table 7.5: Buffers for protein purification of SIRT1 protein

7.4.4 PCAF HAT (p1-His-HAT) expression and purification

PCAF p1-His-HAT was transformed into Rosetta™ 2(DE3)pLysS Singles™ competent *E. Coli* cells (kanamycin and chloramphenicol), and starter cultures were prepared in TB (kanamycin and chloramphenicol). AIM-TB (Formedium™, 2 × 1 L) was prepared (kanamycin) and inoculated with starter culture (1 mL/L). Cells were grown for 4 h at 37°C/200 rpm followed by 48 h at 18°C/200 rpm. Cells were spun down (18600 *g*, 20 min, 4°C) and cell pellets were stored at –80°C.

Cell pellets were defrosted and resuspended in 50 mL lysis buffer containing DNaseI. The solution was lysed using a cell disruptor and cell debris was removed by centrifugation (39800 *g*, 20 min, 4°C). The supernatant was then filtered through a 0.45 µM filter followed by a 0.2 µM filter, then loaded onto a gravity IMAC column loaded with Ni²⁺ (CV = 2 mL). The column was washed with IMAC buffers of increasing imidazole concentrations. All fractions were analysed by SDS-Page and His-HAT containing fractions pooled. TEV (2 µL of 90 mg/mL) was added to the mixture for His-tag cleavage and the solution was buffer exchanged by dialysis overnight with low imidazole buffer. The dialysed protein solution was loaded onto a gravity IMAC column loaded with Ni²⁺ (CV = 2 mL), and incubated with the resin for 30 min. The flow through was collected and the resin was washed with wash buffer. Fractions were analysed by SDS-Page, HAT containing fractions were pooled and concentrated. Fractions were buffer exchanged into SEC buffer and concentrated to 174 µM (NanoDrop™). Protein aliquots were snap frozen in liquid nitrogen for long term storage at –80°C. Protein mass confirmed by mass spectrometry (Appendix Figure E.1), and enzymatic activity was assessed using a MALDI-TOF-MS assay.

PCAF HAT protein, Mass = 19437.76 Da, Mass observed = 19437 Da, $\epsilon = 20400 \text{ M}^{-1}.\text{cm}^{-1}$

7.4.5 PCAF His-HAT (p1-His-HAT) expression and purification

PCAF p1-His-HAT was transformed into Rosetta™ 2(DE3)pLysS Singles™ competent *E. Coli* cells (kanamycin and chloramphenicol), and starter cultures were prepared in TB (kanamycin and chloramphenicol). TB media (2 × 1 L) was prepared (kanamycin) and inoculated with starter culture (1 mL/L). Cells were grown to an OD₆₀₀ 0.4-0.6 for ~4 h

at 37°C/200 rpm followed by cooling to 18°C for 1h, induction with ITPG (0.5 mM) and growth for 48 h at 18°C/200 rpm. Cells were spun down (18600 *g*, 20 min, 4°C) and cell pellets were stored at –80°C.

Cell pellets were defrosted and resuspended in 50 mL lysis buffer containing DNaseI. The solution was lysed using a cell disruptor and cell debris was removed by centrifugation (39800 *g*, 20 min, 4°C). The supernatant was then filtered through a 0.45 µM filter followed by a 0.2 µM filter, then loaded onto a gravity IMAC column loaded with Ni²⁺ (CV = 2 mL). The column was washed with IMAC buffers of increasing imidazole concentrations. All fractions were analysed by SDS-Page and His-HAT containing fractions pooled. Fractions were analysed by SDS-Page, HAT containing fractions were pooled and concentrated. The concentrated fractions were further purified by SEC on a 75 µg HiLoad™ column at 1 mL/min with SEC buffer. His-HAT containing fractions were analysed by SDS-Page pooled and concentrated to 158.89 µM (NanoDrop™). Protein aliquots were snap frozen in liquid nitrogen for long term storage at –80°C. Protein mass confirmed by mass spectrometry (Appendix Figure E.2), and enzymatic activity was assessed using a MALDI-TOF-MS assay.

PCAF His-HAT protein, Mass = 21903.41 Da, Mass observed = 21903.32 Da, ϵ = 27390 M⁻¹.cm⁻¹

7.4.6 PCAF HAT-bio (p2-His-HAT-Bio) expression and purification

p2-His-HAT-Bio was transformed into pBirA-containing Rosetta2 cells, and starter cultures were prepared in TB (kanamycin and chloramphenicol). TB (2 × 1 L) was

prepared (kanamycin and spectinomycin) and inoculated with starter culture (1 mL/L). Cells were grown at 37°C/200 rpm until an OD₆₀₀ of ~0.6 was reached (~5 h). The cultures were transferred to an incubator at 18°C and after 1 h, the expression of PCAF and BirA (biotin ligase) was induced with IPTG (0.5 mM) at OD₆₀₀ 0.765 and the media was supplemented with biotin (0.2 mM). Cells were grown at 18°C for 18 h/200 rpm, spun down (18600 *g*, 20 min, 4°C) and stored at –80°C.

Cell pellets were defrosted and resuspended in 50 mL lysis buffer containing DNaseI. The solution was lysed using a cell disruptor and cell debris was removed by centrifugation (39800 *g*, 20 min, 4°C). The supernatant was then filtered through a 0.2 µM filter and loaded onto a gravity IMAC column loaded with Ni²⁺ (CV = 2 mL). The column was washed with IMAC buffers of increasing imidazole concentrations. All fractions were analysed by SDS-Page and His-HAT-Bio containing fractions were pooled. TEV (10 µL of 90 mg/mL) was added to the mixture for HIS-tag cleavage and the solution was buffer exchanged by dialysis overnight with low imidazole buffer. The dialysed solution was loaded onto a gravity IMAC column loaded with Ni²⁺ (CV=2mL) and incubated with the resin for 30 min. The flow through was collected and the resin was washed with wash buffer. Fractions were analysed by SDS-Page, HAT-bio containing fractions were pooled and concentrated. The concentrated fractions were further purified by SEC on a 75 µg column at 1 mL/min with SEC buffer. HAT-bio containing fractions were analysed by SDS-Page pooled and concentrated to 96 µM (NanoDrop™). Protein aliquots were snap frozen in liquid nitrogen for long term storage at –80°C. Protein mass confirmed by mass spectrometry (Appendix Figure E.3), and enzymatic activity was assessed using a MALDI-TOF-MS assay.

PCAF HAT-bio protein, Mass expected = 22054 Da, Mass observed = 22055.11 Da, ϵ = 27390 M⁻¹.cm⁻¹

7.4.7 PCAF HAT-BRD (p3-His-HAT-BRD) expression and purification

p3-His-HAT-BRD was transformed into Rosetta™ 2(DE3)pLysS Singles™ competent *E. Coli* cells (kanamycin and chloramphenicol), and starter cultures were prepared in TB (kanamycin and chloramphenicol). TB (2 × 1 L) was prepared (kanamycin) and inoculated with starter culture (1 mL/L). Cells were grown at 37°C/200 rpm until an OD₆₀₀ of ~0.6 was reached (5h). The cultures were transferred to an incubator at 18°C and after 1 h, the cells were induced with IPTG (0.5 mM). Cells were grown at 18°C for 20 h then spun down (18600 *g*, 20 min, 4°C) and stored at -80°C.

Cell pellets were defrosted and resuspended in lysis buffer (15 mL) containing DNaseI and complete mini EDTA-free protease inhibitor (Roche). The solution was lysed using a cell disruptor and cell debris was removed by centrifugation (39800 *g*, 20 min, 4°C). The supernatant was then filtered through a 0.2 µM filter and loaded onto a gravity IMAC column loaded with Ni²⁺ (CV=2mL). The column was washed with imidazole buffers of increasing imidazole concentrations. All fractions were analysed by SDS-Page and His-HAT-Bio containing fractions were pooled. TEV (10 µL of 90 mg/mL) was added to the mixture for HIS-tag cleavage and the solution was buffer exchanged by dialysis overnight with low imidazole buffer. The dialysed solution was loaded onto a gravity IMAC column loaded with Ni²⁺ (CV=2mL) and incubated with the resin for 30 min. The flow through was collected and the resin was washed with wash buffer. Fractions were analysed by SDS-Page, HAT-BRD containing fractions were pooled and concentrated. The

concentrated fractions were further purified by SEC on a 75 μ g column at 1 mL/min with SEC buffer. HAT-BRD containing fractions were analysed by SDS-Page, pooled and concentrated to 40.08 μ M (NanoDrop™). Protein aliquots were snap frozen in liquid nitrogen for long term storage at -80°C . Protein mass was confirmed by mass spectrometry (Appendix Figure E.4), and enzymatic activity was assessed using a MALDI-TOF-MS assay.

PCAF HAT-BRD protein, Mass = 47402.99 Da, Mass observed = 47398.00, ϵ = 44810 $\text{M}^{-1}.\text{cm}^{-1}$

7.4.8 PCAF His-HAT-BRD-Bio (p4-His-HAT-BRD-Bio) expression and purification

p4-His-HAT-BRD-Bio was transformed into pBirA-containing Rosetta *E.coli* cells, and starter cultures were prepared in TB (kanamycin and chloramphenicol). TB (1 L) was prepared (kanamycin and spectinomycin) and inoculated with starter culture (1 mL/L). Cells were grown at $37^{\circ}\text{C}/200$ rpm until an OD_{600} of ~ 0.6 was reached (5h). The cultures were transferred to an incubator at 18°C and after 1 h, the expression of PCAF and BirA (biotin ligase) was induced with IPTG (0.5 mM) and the media was supplemented with biotin (0.2 mM). Cells were grown at 18°C for 20 h then spun down (18600 g , 20 min, 4°C) and stored at -80°C .

Cell pellets were defrosted and resuspended in lysis buffer (15 mL) containing DNaseI and complete mini EDTA-free protease inhibitor (Roche). The solution was lysed using a cell disruptor and cell debris was removed by centrifugation (39800 g , 20 min, 4°C). The

supernatant was then filtered through a 0.2 μM filter, then loaded onto a gravity IMAC column loaded with Ni^{2+} (CV=2mL). The column was washed with imidazole buffers of increasing concentrations. All fractions were analysed by SDS-Page and high purity His-HAT-BRD-bio fractions were concentrated, tested for activity using MALDI-TOF-MS and concentrated to 16 μM (NanoDrop™). Protein aliquots were snap frozen in liquid nitrogen for long term storage at -80°C . Low purity His-HAT-BRD-bio fractions were concentrated down to 3 mL and further purified by gel filtration on a 75 μg column at 1 mL/min using gel filtration buffer. His-HAT-BRD-bio containing fractions were analysed by SDS-Page, pooled and concentrated to 6 μM (NanoDrop™). Protein aliquots were snap frozen in liquid nitrogen for long term storage at -80°C . Enzymatic activity was assessed using a MALDI-TOF-MS assay.

PCAF HAT-BRD-bio protein, Mass = 50156.73 Da, $\epsilon = 53290 \text{ M}^{-1}.\text{cm}^{-1}$

7.4.9 SIRT1.1

p5-SIRT1.1 was transformed into Rosetta™ 2(DE3)pLysS Singles™ competent *E. Coli* cells (ampicillin and chloramphenicol), and starter cultures were prepared in 2YT (ampicillin and chloramphenicol). 2YT ($2 \times 1 \text{ L}$) was prepared (ampicillin) and inoculated with starter culture (1 mL/L). Cells were grown at $37^{\circ}\text{C}/200 \text{ rpm}$ until an OD_{600} of ~ 0.6 was reached (5h). The cultures were transferred to an incubator at 25°C and after 1 h, the cells were induced with IPTG (0.5 mM). Cells were grown at 25°C for 12 h then spun down (18600 *g*, 20 min, 4°C) and stored at -80°C .

Cell pellets were defrosted and resuspended in lysis buffer (15 mL) containing DNaseI and complete mini EDTA-free protease inhibitor (Roche). The solution was lysed by sonication, 60% amplitude, 30 s of (0.5 s on, 0.5 s off) followed by 30 s rest, 6 cycles on ice, and cell debris was removed by centrifugation (39800 *g*, 20 min, 4°C). The supernatant was then filtered through a 0.2 µM filter and loaded onto a gravity IMAC column loaded with Ni²⁺ (CV = 2 mL). The column was washed with imidazole buffers of increasing imidazole concentrations. All fractions were analysed by SDS-Page and His-SIRT1 containing fractions were pooled. The His-SIRT1 containing fractions were further purified by SEC on a 200 µg column at 1 mL/min with SEC buffer. His-SIRT1 containing fractions were analysed by SDS-Page, pooled and concentrated to 71 µM (Nanodrop™). Protein aliquots were snap frozen in liquid nitrogen for long term storage at –80°C. Protein mass was confirmed by mass spectrometry (Appendix Figure E.5), and enzymatic activity was assessed using a MALDI-TOF-MS assay.

His-SIRT1 protein, Mass = 76301 Da, Observed mass = 76311.89 Da, $\epsilon = 40340 \text{ M}^{-1}.\text{cm}^{-1}$

7.5 Peptide synthesis

7.5.1 Peptide synthesis machines used

A Liberty Blue™ automated peptide synthesiser with HT12 resin loader (CEM) with Liberty Blue™ application software (CEM) or a PurePep® Chorus automated peptide synthesiser (Gyros protein technologies) were used for peptide synthesis, depending on availability at time of synthesis. Both synthesisers use fluorenylmethyloxycarbonyl (Fmoc) solid phase chemistry but differ in their reaction vessel heating methods with

Liberty Blue™ using microwave heating and PurePep® using induction heating. Fmoc-protected amino acids were purchased from CEM for use in peptide synthesis.

Peptides synthesised using PurePep® are listed in Table 7.6 and peptides synthesised using Liberty Blue™ are listed in Table 7.7. All synthesised and purchased peptides were prepared as C-terminal amides and stored as lyophilised powder and/or as stock solutions in T1H₂O at –20°C.

7.5.2 Peptide synthesis procedure

Synthesis:

Peptides were synthesised on 0.05 mmol scale peptide synthesis according to the manufacturer's protocols. Rink amide 4-methylbenzhydrylamine (MBHA) polymer bound (AGTC Bioproducts) was used as resin, with resin loading specified for each batch. Amino acids with an N-terminal Fmoc group and protected side chains where needed were weighed out and dissolved in dimethylformamide (DMF, 0.2 M for Liberty Blue™ or 0.25 M for PurePep™, peptide grade 99.5% purity DMF). Amide couplings were carried out with diisopropylcarbodiimide (DIC, 0.5 for Liberty Blue™ or 0.25 M for PurePep™) as activating agent and Oxyma (1 M for Liberty Blue™ or 0.25 M for PurePep™) as activating base, made up in peptide grade DMF. Single couplings (90°C, 2 min) were performed for all amino acids except cysteine and histidine with low temperature single couplings (50°C, 2 min) and arginine with a double coupling (75°C, 2 min × 2). Fmoc deprotection was carried out with 20% piperidine (5 mL, 3 min × 2) followed by washing of the resin with DMF (10 mL × 5) before each coupling. For

synthesis of cyclic peptides, chloroacetic anhydride was coupled manually onto the peptide C-terminus before cleavage from resin. The resin was resuspended in DMF (~5 mL) and incubated with chloroacetic anhydride (10 eq., 2h, 20°C) with agitation. The resin was washed with DMF (10 mL × 2) and DCM (10 mL × 3) and dried under N₂.

Cleavage and deprotection:

Cleavage from the resin and removal of amino acid side chain protecting groups was performed for linear and cyclic peptides by incubating the resin-bound protected peptides with a cleavage solution (5 mL, 3h, rt) composed of trifluoroacetic acid (TFA, 92.5%), T1H₂O (2.5%), triisopropylsilane (TIPS, 2.5%), and dimethoxybenzene (2.5%). The solution was recovered, and the beads were washed with cleavage solution (5 mL) for increased peptide recovery. The cleavage was performed either manually for peptides synthesised with Liberty Blue™ or automatically for peptides synthesised with PurePep™. The cleaved peptides were precipitated in ether (45 mL, precooled at 4°C) and collected by centrifugation (4000 rpm, 4°C, 10 min). The recovered pellet was washed with ether (45 mL × 2) with centrifugation. The pellet was dried under N₂, resuspended in T1H₂O/acetonitrile (ACN) (the ratio of T1H₂O to ACN was adapted to each peptide depending on solubility), snap frozen in liquid N₂ and lyophilised a Benchtop Pro Onitronics (SP Scientific) freeze drier. Lyophilised linear peptides was re-dissolved in T1H₂O/ACN (2–8 mL) to achieve full solubilisation of the peptide.

Optional cyclisation:

For cyclic peptides, crude peptide (10-20 mg) was dissolved in T1H₂O/ACN (~1 mL). NaOH (aq. 0.1 M) was added until pH 8-9 was achieved (10 µL increments, tested using pH paper). The solution was incubated for cyclisation (1h, 60°C). MALDI-TOF-MS was used to confirm peptide cyclisation. The reaction was quenched with ACN/T1H₂O/TFA (ratio previously determined) to give 0.1% TFA in the final solution (acidity tested by pH paper).

Purification:

Linear and cyclic peptides were purified by reverse-phase high-performance liquid chromatography (Rf-C18 HPLC) on an 1260 Infinity II (Agilent) with a 5 µM NX-C18, 110 Å, 250 x 21.2 mm AX1 column (Gemini, Phenomenex Flow). The initial mixture of solvent A to solvent B was tailored to the solubility of each peptide in water (solvent A: 0.1% TFA in T1H₂O, solvent B: 0.1% TFA in ACN). Purification methods were tailored to each peptide to give optimal separation of product peptide from impurities, over 60 minutes with progressive increase of solvent B. Fractions were collected based on elution time and peak slope, and the collected fractions were analysed by MALDI-TOF-MS. Fractions containing pure peptide were pooled, snap frozen in liquid N₂, and lyophilised.

Characterisation:

Peptide mass was confirmed using MALDI-TOF-MS (Table 7.6, Table 7.7). Peptide concentration was determined using ¹H NMR spectroscopy (Bruker Avance III HD 700

MHz) using a 3-(tri-methylsilyl)-propanoic acid (TSP) internal standard reference with a ^1H chemical shift of 0 ppm. Pure peptide was weighed and dissolved in $\text{T1H}_2\text{O}$ (5 or 10 mM presumed, depending on peptide solubility). Peptide (5 μL) and TSP (10 mg/mL, 5 μL) were added to 550 μL D_2O for ^1H NMR analysis. Spectra were generated in TopSpin 3.2 with manual phase and baseline correction applied. The spectra were analysed with manual integration in MestReNova v12.0.1 and the absolute concentration of the peptide solution was calculated using Equation 2.

$$C_x = \frac{I_x}{I_{cal}} \times \frac{N_{cal}}{N_x} \times C_{cal}$$

Equation 2: Calculation of accurate peptide concentration using ^1H NMR with an internal standard.

C = concentration, I = integral area, N = number of nuclei, cal = calibrant (TSP), x = peptide

7.6 Peptides sequences

7.6.1 Linear peptides

H3₁₋₂₁, H3₁₋₂₁K9Ac, H3₁₋₂₁K14Ac, H3₁₋₂₁K9AcK14Ac, H3₁₋₂₁K14Me₂ were purchased from GL Biochem. H3₁₋₂₁K14Me₁, H3₁₋₂₁K14Me₃, HIF1 α ₅₂₂₋₅₄₂, HIF1 α ₅₂₂₋₅₄₂K523Ac, HIF1 α ₆₆₄₋₆₈₄, HIF1 α ₆₆₄₋₆₈₄K674Ac, HIF1 α ₇₀₀₋₇₂₁, HIF1 α ₇₀₀₋₇₂₁K709Ac were synthesised using a PurePep™ peptide synthesiser, following procedure in section 7.5.2. All peptides were synthesised as C-terminal amides and were dissolved in $\text{T1H}_2\text{O}$ and quantified by ^1H NMR, and the peptide mass was checked using MALDI-TOF-MS. MALDI-TOF-MS spectra of synthesised peptides are shown in Appendix Ci.

Peptide name	Peptide sequence	Calculated [M+H] ⁺ mass (Da)	Observed [M+H] ⁺ mass (Da)
*H31-21	ARTK4QTARK9STGGK14APRK18QLA	2253.33	2253.325
*H31-21K9Ac	ARTKQTARKAcSTGGKAPRKQLA	2295.34	2295.201
*H31-21K14Ac	ARTKQTARKSTGGKAcAPRKQLA	2295.34	2295.538
*H31-21K9AcK14Ac	ARTKQTARKAcSTGGKAcAPRKQLA	2337.35	2337.618
‡H31-21K14Me1	ARTKQTARKSTGGKMeAPRKQLA	2267.34	2267.526
*H31-21K14Me2	ARTKQTARKSTGGKMe2APRKQLA	2281.36	2281.539
‡H31-21K14Me3	ARTKQTARKSTGGKMe3APRKQLA	2296.38	2296.766
‡HIF1α522-542	YVDSDMVNEFKLELVEKLFAE	2517.22	2517.421
‡HIF1α522-542K523Ac	YVDSDMVNEFKAcLELVEKLFAE	2559.26	2559.488
‡HIF1α664-684	SRTASPNRAGKGVIEQTEKSH	2252.18	2252.384
‡HIF1α664-684K674Ac	SRTASPNRAGKAcGVIEQTEKSH	2294.19	2294.454
‡HIF1α700-721	VPEEELNPKILALQNAQRK	2391.33	2391.651
‡HIF1α700-721K709Ac	VPEEELNPKAcILALQNAQRK	2433.34	2433.671

Table 7.6: Peptides used in MALDI-TOF-MS assays for characterising PCAF HAT/HAT-BRD and SIRT1 enzyme activity.

*peptide purchased from GL biochem ‡peptide synthesised as described in section 7.5.2. Observed mass [M+H]⁺ was determined by MALDI-TOF-MS.

7.6.2 Cyclic peptides

Cyclic peptides were synthesised as C-terminal amides using a Liberty Blue™ peptide synthesiser, following procedures in section 7.5.2. Peptides listed are cyclised between the C-terminal cysteine side chain and a chloro-acetyl group on the N-terminal tyrosine backbone amine forming a thioether bond. All purified peptides were dissolved in T1H₂O, quantified by ¹H NMR, and the peptide mass was checked using MALDI-TOF-MS. MALDI-TOF-MS spectra of synthesised peptides are shown in Appendix Cii.

Peptide name	Peptide sequence	Calculated [M+H] ⁺ mass (Da)	Observed [M+H] ⁺ mass (Da)
LS1	YYASGQIRENFC	1489.69	N/A
LS2	YNGQVHASTLELQC	1601.73	N/A
LS3	YNGTPGHVWC	1172.41	1172.436
LS4	YIDPQEVSAIC	1276.58	[M+Na] ⁺ 1298.461
LS5	YTSEAQQWEC	1283.49	1283.411
LS6	YPNDRGGQEVKC	1404.63	1404.576
LS7	YASYKQVIVLC	1326.67	N/A
LS8	YLFSDRHRNRDC	1621.74	1621.707
LS9	YANAIKQGHSVC	1329.63	N/A
LS10	YIRVDRGSRLC	1377.70	N/A
LS11	YSKSTYGTPGNC	1316.53	1316.524
LS12	YCRREWTESPC	1469.59	N/A

Table 7.7: Cyclic peptides synthesised following mRNA display against PCAF HAT and HAT-BRD. The peptides were synthesised as described in section 7.5.2. Peptides with no observed mass were not purified due to solubility issues.

7.7 Procedures for MALDI-TOF-MS assays

7.7.1 MALDI-TOF-MS acquisition procedure

MALDI matrix was prepared by mixing an excess α -cyano-4-hydroxycinnamic acid (CHCA) in a solvent mixture of ACN (50%), T1H₂O (49.9%) and TFA (0.1%), with sonication (10 min, rt), ensuring saturation of the solvent with CHCA. MALDI-TOF-MS spectra were obtained by spotting sample (0.5-1 μ L) with an equivalent volume of MALDI matrix. Plates were dried at 37°C prior to insertion into the MALDI-TOF-MS instrument. For assays and during peptide purification, a spot of the crude reaction mixture was used for analysis (peptide concentration of 5-10 μ M generally used). Purified peptides were diluted in T1H₂O (final concentration 5-50 μ M) and spotted. Plates used were ground steel 96 spot plates, unless otherwise specified. Spectra were acquired using Flexcontrol 3.4 on a Bruker Daltonics MALDI-TOF-MS microflex Flexseries 1.4 (flight tube voltage 2500 V, reflectron voltage 1500 V, positive ion mode), with manual adjustment of parameters including laser voltage, number of shots per spot adjusted manually. For a

small number of spots (≤ 10) spectra were acquired manually. For large number of spots and full plates, spectra were acquired automatically, with a set method in which the mass range and minimum and maximum laser power were adapted for each experiment (Appendix F). Spectra were processed and analysed using Flexanalysis 3.4 software (Bruker Daltonics). Baseline subtraction and smoothing were applied to the spectra, and automatic peak picking in SNAP processing mode was used. Spectra images were exported from Flexanalysis directly for use. For analysis of enzyme activity, spectra details were exported into excel for processing or a processing script written by Dr. Tom McAllister was used to extract required information.

7.7.2 Specifications for PCAF HAT/HAT-BRD and SIRT1 enzyme assays

Specifications for PCAF HAT/HAT-BRD assays:

- Assay buffer: 50 mM HEPES, 50 mM NaCl pH 7.5
- Coenzyme: AcCoA
- Control peptide: H3₁₋₂₁

Specifications for SIRT1 assays:

- Assay buffer: 50 mM HEPES, 50 mM NaCl pH 7.5
- Coenzyme: NAD⁺
- Control peptide: H3₁₋₂₁K9Ac

Temperature controlled assays were carried out using an Eppendorf thermomixer C for single plate assays, and either a ThermoFisher Scientific Heratherm incubator or a Techne hybridiser HB-10 for multi-plate assays.

7.7.3 Enzyme assay protocol for time dependent assays with PCAF/SIRT1 proteins.

Solutions containing either enzyme or a mixture of peptide and co-enzyme were made up at 2 × the required concentration. The enzyme, cofactor and peptide concentrations were adapted to the assay being performed. The solutions were incubated separately (37°C, 5 min), and an equal volume of each solution was transferred to a reaction tube to initiate the reaction. Aliquots of reaction mixture were taken at chosen time-points and transferred with mixing to a 96-well plate pre-prepared with formic acid (2% in T1H₂O, to give 1% after mixing with reaction solution), to quench the reaction. The samples were spotted onto MALDI plates (Section 7.7.1) and enzyme activity was monitored via MALDI-TOF-MS.

Samples were prepared and mixed in PCR tube strips (8 × 100 µL) with heating in a shaking incubator (Eppendorf thermomixer C). Quenching was carried out in 96 well PCR plates. Assays were set up for use of 8 channel pipettes. Reagent stocks were in T1H₂O or in SEC buffer for proteins and were prepared for the assay in assay buffer.

7.7.4 Inhibitor Assay protocol

PCAF inhibitors were purchased from Enzo Life Sciences (C646), Cambridge Bioscience Ltd (anacardic acid, A-485, CPTH2, garcinol), Insight Biotechnology (butyrolactone-3, curcumin), Merck Life Science (embelin, SPV-106). The chemicals were used without further purification and were dissolved in DMSO to a concentration of 10 mM.

Enzyme	PCAF HAT	PCAF HAT-BRD
Enzyme (μM)	0.6	0.8
H3 ₁₋₂₁ (μM)	10	
AcCoA (μM)	100	
Endpoint (min)	7	
Final assay Buffer	50 mM HEPES, 1% DMSO, 50 mM NaCl, pH 7.5	
DMSO free buffer	50 mM HEPES, 50 mM NaCl, pH 7.5	
Buffer for inhibitor dilutions	50 mM HEPES, 4% DMSO, 50 mM NaCl, pH 7.5	

Table 7.8: Assay conditions for testing of inhibitors with PCAF HAT and HAT-BRD.

Time courses were run according to section 7.7.3 with a buffer containing DMSO (1%), to determine appropriate reaction conditions for PCAF HAT and PCAF HAT-BRD inhibitor testing (Table 7.8). The compounds bought for testing were dissolved in DMSO to a concentration of 10 mM. The compounds were further diluted to 1 mM in DMSO free buffer. A 1:4 dilution series was made for each of the compounds to give a concentration $\times 4$ the final wanted assay concentration. DMSO containing buffer was used in the dilution series to keep the DMSO concentration constant between samples.

The compound/DMSO (2 μL) control samples were transferred to a 96-well PCR plates, to a premixed solution of H3₁₋₂₁ peptide and AcCoA (4 μL , DMSO free buffer) was added. The plate was mixed and centrifuged and the plate and enzyme solution were incubated (37°C, 2 min). Enzyme solution (2 μL , DMSO free buffer) was added to each well, the plate was incubated (37°C, 7 min) then the assay was quenched with formic acid (8 μL , 2% in T1H₂O). Each assay point was analysed by MALDI-TOF-MS.

MALDI-TOF-MS results were extracted using a custom script (Dr. Tom McAllister) to calculate percentage conversion of H3₁₋₂₁ peptide at each inhibitor concentration (in triplicate), which was normalised using the DMSO control. Results were plotted using GraphPad Prism 9.2.0, and a curve was fitted using a non-linear four-parameter fit. Where applicable, IC₅₀s were calculated automatically using Equation 3.

$$Y = \frac{\text{low } Y + (\text{top } Y - \text{low } Y)}{1 + 10^{(\text{LogIC}_{50} - X) \times \text{hillslope}}}$$

Equation 3: Calculation of IC_{50} s using Y = percentage conversion, low Y = plateau at minimum conversion, top Y = plateau at maximum conversion, X = compound concentration, hillslope = slope of curve

7.7.5 Inhibitor assay protocol for testing of cyclic peptides

Cyclic peptides produced based on mRNA selection results against PCAF HAT and HAT-BRD (Table 7.7) were tested for inhibition of PCAF HAT and HAT-BRD acetylation activity of H3₁₋₂₁ substrate peptide. Assays were run as described in section 7.7.4, with some modifications. Purified cyclic peptides were dissolved in DMSO (10 mM stock), then diluted in assay buffer (final DMSO concentration 1% in assay). Peptides and an anacardic acid control inhibitor were tested in triplicate at three concentrations (1 μ M, 10 μ M and 100 μ M final concentration in assay) with a DMSO (1%) control. Results were extracted as described in section 7.7.4, and plotted in GraphPad Prism™ with calculated standard deviation, but as no inhibition was observed for the cyclic peptides no curve fitting or IC_{50} calculations were used.

7.7.6 Histone peptide library screening using MALDI-TOF-MS

Histone libraries with PTM 21-mer biotin labelled peptides were purchased from AltaBioscience. These are presented in a 96-well deep well plate format as lyophilised peptides. The libraries were prepared by resolubilising the lyophilised peptide in T1H₂O/ACN, with T1H₂O added first with agitation and sonication of the 96-well plate, and ACN added gradually to ensure full solubilisation of peptide. If necessary TFA (1% in T1H₂O, 5 μ L increments) was added to help with solubility. The master plate (200 nmol

peptide) with resolubilised peptides was aliquoted out to make daughter plates (5 × 40 nmol). Four plates were and lyophilised a Benchtop Pro Onitronics (SP Scientific) freeze drier for storage at –80°C, and the fifth was used to produce library stock plates (96-well PCR plates) for use in the MALDI-TOF-MS screening of histone peptide libraries. Stock plates were stored at –20°C.

Conditions for screening were chosen by pre-screening PCAF and SIRT1 enzymes with control peptides (H3₁₋₂₁ and H3₁₋₂₁K9Ac respectively) as described in section 7.7.3. Specifications and conditions for low- and high-throughput screens with PCAF HAT/HAT-BRD and for high-throughput screens with SIRT1 are shown in Table 7.9. Screening was carried out in 96-well PCR plates, with position A1 and H12 free for control peptides. For PCAF and SIRT1 screens position A1 was used for H3₁₋₂₁ and for SIRT1 screens position H12 was used for H3₁₋₂₁K9Ac. The assay setup is described below using specifications and concentrations for the rapifleX MALDI-TOF-MS HTS screens.

Co-enzyme mixture (2 µL, 375 µM stock, 100 µM in assay) was dispensed into each plate using a multi-dispensing 8-channel pipette. Stock library peptide mixture (1.5 µL, 100 µM stock, 20 µM in assay) was added to each well using an 8-channel pipette. The assay plates were shaken and centrifuged for mixing. Enzyme mixture (4 µL, 3.75 µM stock, 2 µM in assay) was added to each well using an 8-channel pipette. The assay plates were shaken, centrifuged and incubated (37°C, 1h). The assay was quenched with formic acid (2% in T1H₂O, 7.5µL), and 10µL of each sample was transferred to a 384-well plate suitable for Mosquito liquid handling (SPT Labtech).

For low-throughput assays, samples were spotted onto MALDI plates and spectra acquired as described in section 7.7.1 by the author. For high-throughput assays, spotting onto MALDI plates and spectra acquisition were performed by Dr. Maria Dueñas (in a collaboration with Prof. Matthias Trost, Newcastle University). Spotting was performed using a mosquito® liquid handler onto a 1536 spot AnchorChip™ MALDI plate and spectra acquisition was performed on a Bruker rapifleX Pharmapulse™ MALDI-TOF-MS instrument.

Enzyme	PCAF HAT/HAT-BRD	PCAF HAT/HAT-BRD	SIRT1
Assay throughput	Low-throughput	High-throughput	High-throughput
MALDI-TOF-MS instrument used	microflex	rapifleX	rapifleX
Libraries screened	SETIII	SETI, SETII, SETIV	SETIV
Enzyme (μM)	2	2	2
Peptide (μM)	10	20	20
Coenzyme (μM)	50	100	100
Buffer used	50 mM HEPES, 50 mM NaCl, 1 mM TCEP, pH 7.5	50 mM HEPES, 50 mM NaCl pH 7.5	50 mM HEPES, 50 mM NaCl pH 7.5
Assay running	37°C, 1h	37°C, 1h	37°C, 1h
Number of replicates	3 assay replicates for HAT, 2 assay replicates for HAT-BRD	3 assay replicates 4 technical replicates Total: 12	3 assay replicates 4 technical replicates Total: 12
Number of NEC replicates	2 assay replicates	3 assay replicates 4 technical replicates Total: 12	3 assay replicates 4 technical replicates Total: 12

Table 7.9: Specifications and conditions used for each of the MALDI-TOF-MS screens performed with AltaBioscience histone peptide libraries

7.7.7 MALDI-TOF/TOF-MS experiments

Time course assay conditions:

- Enzyme PCAF HAT/HAT-BRD: 1 μM
- Peptides (H3₁₋₂₁, H3₁₋₂₁K14Ac, H3₁₋₂₁K9K14Me1, H3₁₋₂₁K14Me2, H3₁₋₂₁K14Me3): 20 μM
- Coenzyme AcCoA: 100 μM
- Buffer: 50 mM HEPES, 50 mM NaCl, pH 7.6

- Quench: 15% formic acid in T1H₂O, final concentration of formic acid 1%
- Incubation at 37°C for 5h

Time course assays were run as described in section 7.7.3 with conditions listed above.

Time course samples were analysed using microflex MALDI-TOF-MS and Flexanalysis to identify appropriate samples for MALDI-TOF/TOF-MS analysis. Selected samples were spotted manually onto a 384-well ground steel MALDI plate (section 7.7.1) and run with Bruker autoflex™ or rapiflex Pharmapulse™ MALDI-TOF/TOF-MS. MS/MS spectra were analysed using Bruker Biotools software with spectra and fragment identification data exported directly from the software.

7.8 mRNA display procedures

7.8.1 Phenol-Chloroform Purification of DNA/RNA

DNA/RNA products were transferred (*e.g.* 100 µL) into a standard Eppendorf tube (1.5 mL). An equivalent volume (*e.g.* 100 µL) of phenol/chloroform/isoamyl alcohol (25:24:1, Sigma Aldrich) was added to the sample (working in a fume hood) and it was mixed by vigorous shaking, then centrifuged (≥15,000 g, 5 min). The aqueous top layer was recovered by pipette and added to a new pre-labelled Eppendorf tube (1.5 mL). As protein could precipitate between the two layers, care was taken to ensure the pipette tip did not touch the interface. The same volume (*e.g.* 100 µL) of chloroform/isoamyl alcohol (24:1, Sigma Aldrich) was added and mixed by vigorous shaking, then centrifuged (≥15,000 g, 5 min). The top layer was transferred into a new tube. The volume of the extract was determined using a pipette. NaCl (3 M) was added at approximately 1/9 of the volume of the sample (*e.g.* 11 µL), giving the correct NaCl concentration for

DNA/RNA precipitation in ethanol. 2x the resultant volume (*e.g.* 220 μL) of 100% ethanol was added to the DNA/RNA solution and mixed well by vortex to precipitate the DNA/RNA, then centrifuged ($\geq 15,000$ g, 15 min) and the supernatant was removed. Half the initial volume (*e.g.* 50 μL) of 70% ethanol was added to wash the pellet by vortexing, then centrifuged ($\geq 15,000$ g, 3 min). The supernatant was carefully removed to leave the pellet intact, which was then air dried (5-10 min). The DNA/RNA pellet was dissolved in $\text{T1H}_2\text{O}$ (*e.g.* 20 μL) for downstream applications. DNA and RNA concentrations were measured by NanoDrop™ 1000 3.8.0, or on a Qubit 4 fluorometer (Invitrogen), used according to manufacture instructions.

7.8.2 Determination of protein loading on streptavidin coated beads

Loading buffer: 50 mM HEPES, 50 mM NaCl, 0.05% Tween-20, pH 7.6

Streptavidin coated magnetic beads (Life-technologies Dynabeads® M-280 Streptavidin, 100 μL) were prepared by removal of the storage solution and washed with loading buffer (100 $\mu\text{L} \times 3$). Increasing quantities of beads were placed into separate Eppendorf tubes (0 μL , 6.25 μL , 12.5 μL , 25 μL , 50 μL). The supernatant was removed from each tube and the wanted biotinylated protein (HAT-bio/His-HAT-BRD-bio, 20 μL of 2.5 μM protein) was added to the beads. The slurry was incubated with shaking (4°C, 30 min). The supernatant was removed from the beads and kept for analysis. A solution of 1:1 loading dye and reducing agent (20 μL , 1 $\times$ NuPage® LDS Sample Buffer, 1 $\times$ NuPage® Sample reducing agent) was added to the beads. The supernatant and beads were heated (95°C, 5 min) and the samples were run on a NuPage™ 12% Bis-Tris precast protein gel (200 V, 45 min, 1 $\times$ MES buffer). The gel was stained in Instant Blue™ (2h,

with rocking, Expedon) then de-stained in T1H₂O (2 h, with rocking) and imaged by colorimetric imaging on ImageLab 5.0 (Bio-Rad).

7.8.3 Flexizyme and formyl-methionine tRNA synthesis, purification and charging

Protocols published by Goto. *et al.* were followed for synthesis and purification of materials for the mRNA display selections.^{442,455} Flexizyme and formyl-methionine tRNA for insertion of the ClAc-^{D/L}Tyr-tRNA^{INI} were synthesised by PCR amplification of selected primers (Flexizyme primers: eFx-R1, eFx-R2, FxF and T7-F; tRNA primers: fMet-F, T7-F, fMet-R1, fMet-R2, fMet-R3; Extension 1: 95°C 60 s, 50°C 60 s, 72°C 60 s, 5 repeats; Extension 2: 95°C 60 s, 95°C 40 s, 50°C 40 s, 72°C 40 s, 14 repeats; Extension 3: 95°C 60 s, 95°C 40 s, 50°C 40 s, 72°C 40 s, 5 repeats).⁴⁵⁵ The DNA was checked on a 3% agarose gel and purified by a phenol chloroform extraction (section 7.8.1). The DNA was transcribed to RNA (1 mL scale; overnight incubation at 37°C; conditions for flexizyme were 5 × T7 Buffer, 10 mM DTT, 30 mM MgCl₂, 5 mM NTPs, 30 ng/μL DNA, 0.03 M KOH, 2 U/μL T7 RNA Polymerase; conditions for tRNA were 5 × T7 Buffer, 10 mM DTT, 22.5 mM MgCl₂, 5 mM GMP, 3.75 mM NTPs, 30 ng/μL DNA, 0.03 M KOH, 2 U/μL T7 RNA Polymerase). The template DNA was digested (100 mM MnCl₂, 4 U/μL DNase I, 30 min, 37°C) and the flexizyme and tRNA RNA were purified by gel filtration on a HiLoad™ 16/600 Superdex™ 75 pg column (r.t., 1 mL/min, gel filtration buffer: 1 × PBS, 0.3 M NaCl, pH 7.5). The RNA containing samples were analysed on a 12% acrylamide gel, combined and precipitated with EtOH (2 volume equivalents), then washed with 70% EtOH (1 volume equivalent).

The dried pellets were re-solubilised in T1H₂O and quantified using a Qubit 4 fluorometer.

After synthesis and purification, the tRNA was charged with ClAc^{L/D}Y by the flexizyme (500 pmol scale). A mixture of RNA in buffer (50 mM HEPES-KOH pH 7.5, 25 µM tRNA, 25 µM flexizyme, volume made up to 20 µL with T1H₂O) was heated (95°C, 5 min) then cooled (r.t., 5 min) before addition of MgCl₂ (600 mM MgCl₂). The reaction was incubated (r.t., 5 min) then cooled on ice (5 min). The ClAc^{L/D}Y amino acid (5 mM ClAc^{L/D}Y) was added, and the reaction carried out on ice (2 h, final volume 28 µL). The reaction was then quenched (80 mM NaOAc, 70% EtOH, 10 volume equivalents) and centrifuged (>15,000 g, 15 mins) and washed (3 volume equivalents, 3 × 70% EtOH with 0.1 M NaOAc, pH 5.2 followed by 1 × 70% EtOH) with centrifugation between washes (>15,000 g, 5 mins). The resulting pellets of charged tRNA ClAc^{D/L}Tyr-tRNA^{INI} were air dried and stored at -80°C until use.

The synthesised charged tRNA ClAc^{D/L}Tyr-tRNA^{INI} were tested by performing an *in vitro* translation transcription (IVTT) with a PURExpress kit (NEB) on a 5 µL scale as per the manufacturer's instructions, with a DNA test template in low binding tubes (1.5 µL SolA Δ AA, 1.5 µL SolB, 0.3 mM AAs Δ Met, 0.8 U/µL RNase Inhibitor, 100 µM ClAc^{D/L}Tyr-tRNA^{INI}, 2.5 µM DNA, 90 min at 37°C). The solution was purified by C₁₈ ZipTip BondElut OMIX (Agilent) and analysed by MALDI-TOF-MS as described in section 7.7.1.

7.8.4 mRNA cyclic peptide selection procedure

The starting mRNA library mix (customised for each selection, Table 5.1) was ligated with the Pu-linker (1 × T4 RNA ligase buffer, 10 µM mRNA, 15 µM Pu-linker, 1 U/µL RNase Inhibitor, 2.5 U/µL T4 RNA ligase, final volume 7.5 µL, 1 h at room temperature followed by 10 min at 70°C). *In vitro* transcription was then carried out using a NEB Pure Express® Δ (aa, tRNA) Kit (1.5 µL Sol A Δ AAs, 1.5 µL Sol B, 0.3 mM AAs Δ Met, 0.8 U/µL RNase Inhibitor, 100 µM ClAc-^DTyr-tRNA^{INI} prepared previously, 2.5 µM RNA carried through from previous reaction, final volume 5.85 µL, 1 h at 37°C). The reverse transcription premix was prepared (10 µM P2, 3 × M-MLV Buffer, 1.5 mM dNTPs, final volume 17.5 µL) then the reverse transcription step was carried out (*in vitro* transcription input, 1.68 µL of premix, 5.3 U/µL M-MLV RTase, final volume 10 µL, incubation for 1 h at 42°C). The resulting library was added to selection buffer (20 µL final volume, buffers customised for each selection prepared at 2 × wanted concentration, Table 5.1) and incubated with prewashed beads (volume of beads determined by protein loading test section 7.8.2, Life-technologies Dynabeads® M-280 Streptavidin) to remove any streptavidin binding peptide sequences (4°C, 3 × 30 min). After the third negative selection step, an aliquot (0.25 µL) was taken from the supernatant and diluted to 1000 µL in T1H₂O (Sample C). The library (supernatant) was then incubated with PCAF HAT-bio/His-HAT-BRD-bio (200 nM, 4°C, 30 min). After incubation the protein was pulled down with streptavidin beads (volume determined by protein loading test section 7.8.2). DNA was recovered from the beads by heating (95°C, 5 min) and DNA was quantified using qPCR with Thermofisher PowerUp™ SYBR™ Green Master Mix (section 7.8.5) and amplified by PCR with SP1 and SP2 primers (95°C 3 min, [95°C 40 s, 61°C 40 s,

72°C 40 s] × 30+). The number of PCR cycles were determined by qPCR results. Library size was checked on a 3% agarose gel (150 V, 30 min) by comparison with a DNA ladder. The library DNA was purified by a phenol chloroform extraction (section 7.8.1) and quantified using NanoDrop™. Depending on the level of DNA recovery in the selection (as measured by qPCR) the recovered library DNA was sequenced using NGS or transcribed into mRNA and used as input for another round of selection (1 × T7 RNA polymerase buffer, 10 mM DTT, 1 mM MgCl₂, 10 mM NTPs, 3 µg DNA, 1 mM KOH, 1 U/µL RNase Inhibitor, 2.5 U/µL T7 RNA polymerase, final volume 20 µL, incubations overnight at 37°C). The DNA was digested (0.01 U/µL DNase, 10 mM MnCl₂, 37°C, 30 min), the mRNA was recovered by a phenol chloroform extraction (section 7.8.1) and quantified using a Qubit 4 fluorometer.

Primer	Sequence (5' to 3')
eFx-R1	ACCTAACGCTAATCCCCTTTCGGGGCCGCGGAAATCTTTCGATCC
eFx-R2	ACCTAACGCTAATCCCCT
FxF	GTAATACGACTCACTATAGGATCGAAAGATTTCCGC
T7-F	GGCGTAATACGACTCACTATAG
fMet-F	GTAATACGACTCACTATAGGCGGGGTGGAGCAGCCTGGTAGCTCGTCGG
fMet-R1	GAACCGACGATCTTCGGGTTATGAGCCCGACGATCTTCGGG
fMet-R2	TGGTTGCGGGGGCCGGATTTGAACCGACGATCTTCGGG
fMet-R3	TGGTTGCGGGGGCCGGATTT
P1	TAATACGACTCACTATAGGGTTAACTTTAAGAAGGAGATATACATATG
P2	TTCCGCCCCCGTCCTAGCTGCCGCTGCCGCTGCCGCA
SP1	TAATACGACTCACTATAGGGTTAACTTTAAGAAGGAG
SP2	TTCCGCCCCCGTCCTAGCTGC
NGS_F	AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTTCCGATCT TAGGGTTAACTTTAAGAAGGAGATATAC
NGS_R	CAAGCAGAAGACGGCATACGAGATXXXXXXGTGACTGGAGTTCAGACGTGTGCTCT TCCGATCTGTCTAGCTGCCGCTG
DNA-PEG- Puromycin	CTCCCGCCCCCGTC-(SPC) ₅ -CC-puromycin

Table 7.10: Primers and building blocks used in the mRNA display procedures.

For NGS_R, XXXXXX represents the index, which was either: GCGGAC, TTTCAC, GGCCAC or CGAAAC. SPC is a hexaethylene glycol chain used as a spacer. Primers and building blocks were based on work described by Hayashi et. al.⁴⁵⁶

7.8.5 Quantitative PCR (qPCR)

qPCR (triplicate samples) was performed to quantitate the amount of DNA present in the recovered library DNA from the selection. A qPCR master mix was made containing all reagents except for the unknown sample ($1.05 \times$ SYBR GREEN, $0.26 \mu\text{M}$ of P1 and $0.26 \mu\text{M}$ of P2, in $\text{T1H}_2\text{O}$). qPCR master mix ($19 \mu\text{L}$) and recovered library or a known standard ($1 \mu\text{L}$) were added to each sample well in the qPCR plate. Standards used for quantitation purposes were generated by reverse transcription of a known quantity of purified library RNA, assuming a 1:1 transcription ratio. qPCR was carried out using a MJ Mini™ thermal Cycler (BioRad) with a MiniOpticon™ real time PCR System (BioRad) or on a C1000 Thermal Cycler (BioRad) using the program: 95°C (3 min) incubation followed by 40 cycles of [95°C (40 s), 61°C (40 s), 72°C (40 s), plate read].

Bio-Rad CFX Manager 3.1 was used to analyse the data. DNA quantitation was calculated with respect to the standards, using them to calculate a mean starting number of DNA molecules in each qPCR sample, taking into account dilution factors and starting volume. The percentage DNA recovery for the positive mRNA selection against PCAF protein was then calculated by taking the quantity of DNA in the positive selection sample and dividing it by the quantity of DNA in Sample C and converting the result into a percentage.

7.8.6 Next Generation Sequencing and Data Analysis

DNA recovered from the mRNA display selection was prepared for Illumina-compatible sequencing by PCR with specifically designed primers NGS_F and NGS_R ($1 \times$ Q5 buffer,

0.2 mM dNTPs, 0.5 mM NGS_F, 0.5 mM NGS_R, 1 × High GC enhancer, 0.2 U/μL Q5 High-Fidelity DNA pol, 1 μL DNA template, final reaction volume 100 μL, PCR programme 95°C 3 min, [95°C 40 s, 61°C 40 s, 72°C 40 s] × 30).⁴⁵⁷ The PCR products were purified using a PCR purification kit (Qiagen) and a Quant-iT™ PicoGreen™ assay (ThermoFisher Scientific) according to manufacturer instructions. Further steps for NGS (Paired-end sequencing was performed using a NextSeq 500 by Illumina with 80 reads in each direction) were carried out by Dr. Tom McAllister along with post sequencing data extraction. Data analysis was then carried out on excel by the author.

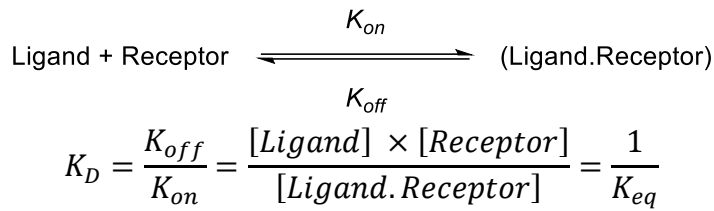
7.9 Bio Layer Interferometry (BLI)

BLI Buffers:

- Buffer 1: PBS, 0.05% Tween, 1 mM TCEP, 0.1% BSA
- Buffer 2: Buffer 1 + x 0.1% DMSO

BLI experiments were carried out on a Forte Bio OctetRed-384 System, with BLI plates (384-well tilted bottom, Forte Bio), BLI sensors (Ni-NTA, biotin, Forte Bio). The selected enzyme (His-HAT, HAT-bio, HAT-BRD-bio) with a tag corresponding to the sensors used (His for Ni²⁺ sensors and avitag-biotin for streptavidin sensors) was diluted in buffer 1 (~1 μM, volume ≤ 560 μL, on ice). The compound of interest was first diluted into buffer 1 from a concentrated stock solution (10 mM) down to the highest concentration to be tested (10 μM, volume ≤ 160 μL). DMSO (0.1%) was kept constant for the peptide dilution series using buffer 2. Compound dissociation steps were performed in wells containing only Buffer 2. Dilution series were prepared for the cyclic peptides with 2-fold dilution with 7 concentrations (10 μM to 156.25 nM), and a DMSO control (buffer 2) in the 8th position, to serve as a reference well for subtraction of Ni-NTA sensor drift in

analysis (Data Analysis v9). Ni-NTA sensors were stripped between cyclic peptides being tested by immersion in glycine (0.5 M, pH 1) and regenerated with 10 mM NiCl₂. Streptavidin sensors could not be regenerated and were single use. The software used for data acquisition was Forte Bio DataAcquisition v9 where the BLI step-procedure for one cyclic peptide in a typical experiment included: Baseline (buffer 1, 15 s); Loading (protein, threshold 1.5 nm / 180 s); Baseline (buffer 1, 15 s); Baseline (buffer 2, 15 s); Association (cyclic peptide dilution series, 120 s); Dissociation (buffer 2, 300-600 s); 3x[Wash (glycine, 3 s), (buffer 1, 3 s)]; Baseline (buffer 1, 15 s); Regeneration step only for Ni-NTA sensors (NiCl₂, 60 s); Baseline (Buffer 1, 15 s). Data were analysed using DataAnalysis v9, with the no-peptide well (8th position) used as a reference for subtraction and alignment to baseline. Data was fitted using the global fit 1:1 model (Equation 4).

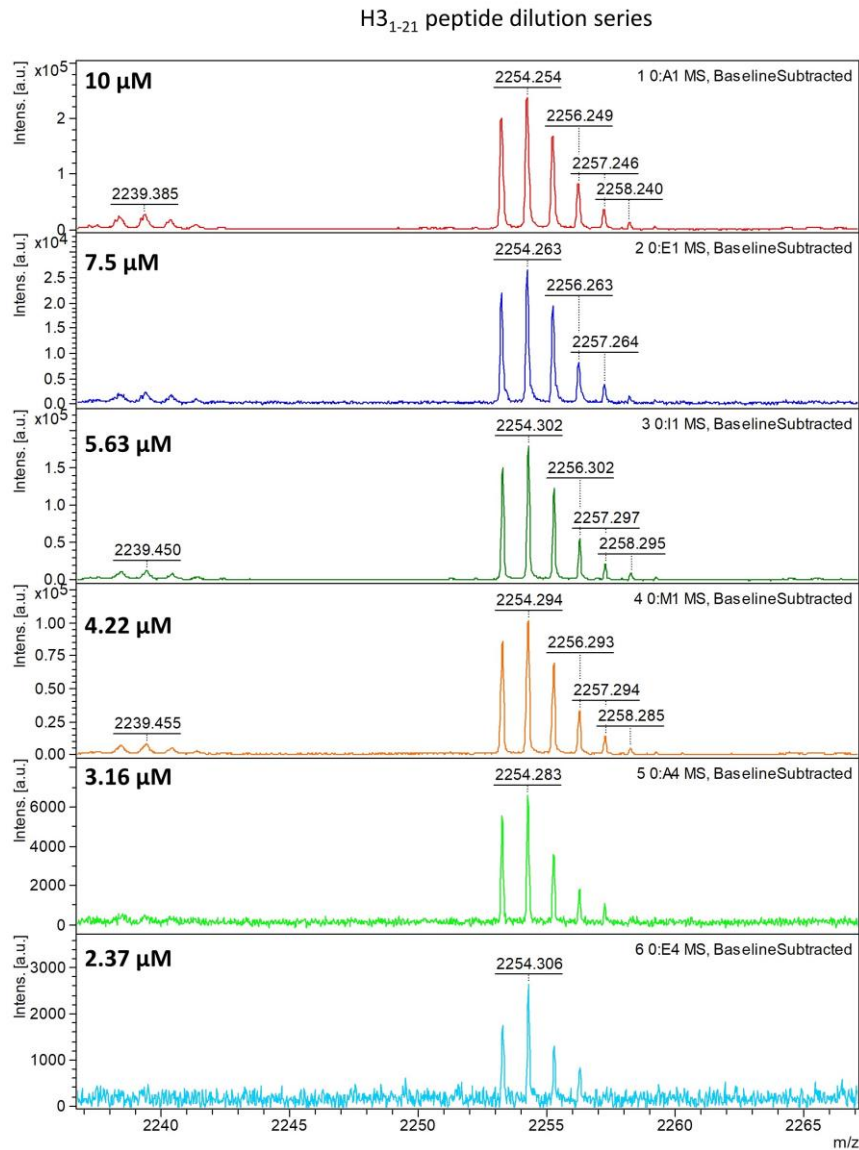


Equation 4: Global fit 1:1 model used for fitting BLI data.

K_D (M) is the equilibrium constant with K_D giving peptide concentration for 50% occupancy of target, k_{on} (M⁻¹s⁻¹) is the association rate constant, k_{off} (M⁻¹s⁻¹) is the dissociation rate constant and K_{eq} (M⁻¹) is the equilibrium constant.

Appendix

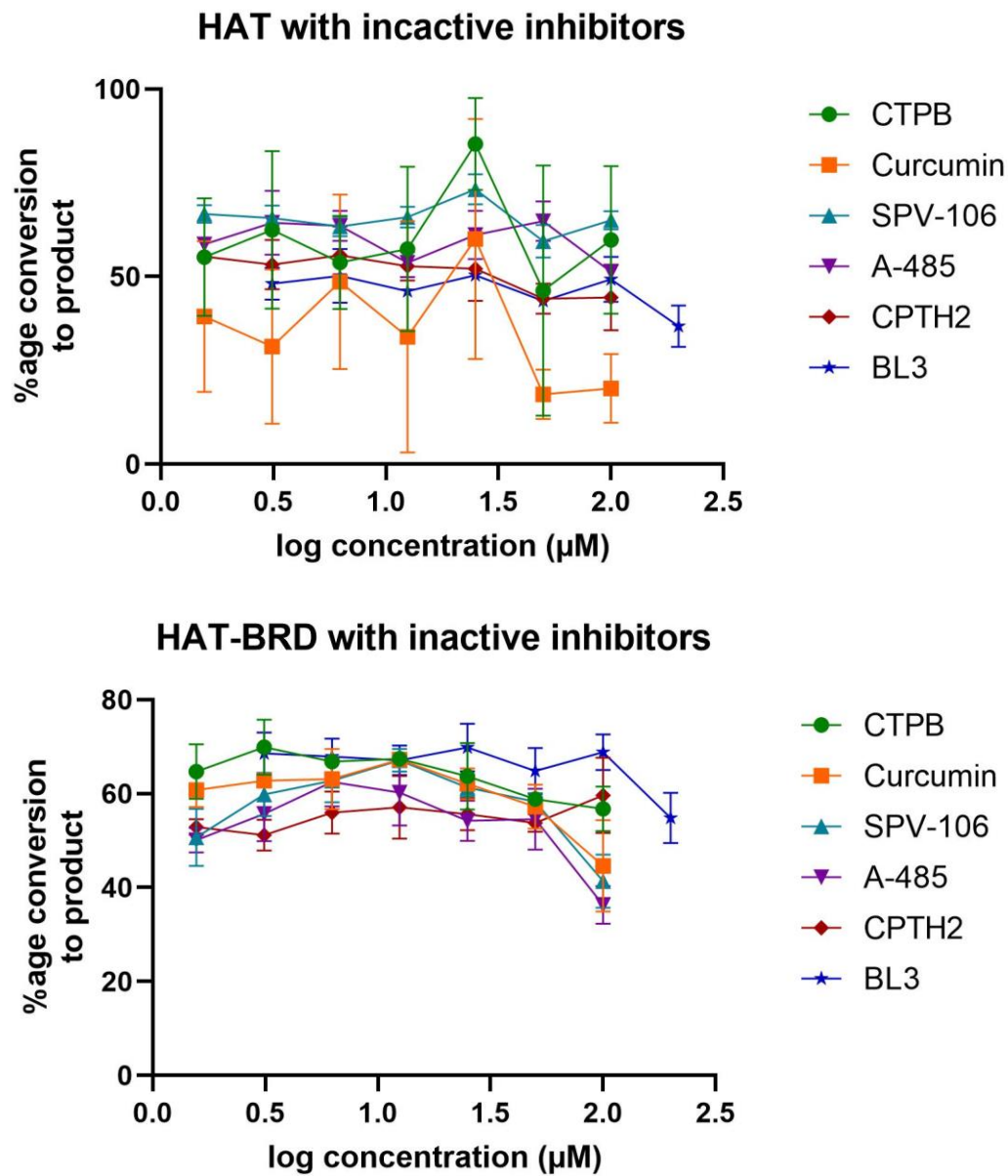
A. RapifleX MALDI-TOF-MS detection limits



Appendix Figure A.1: Spectra acquired for a dilution series of H31-21 peptides in buffer on the rapifleX MALDI-TOF-MS.

Data acquired by Dr. Maria Dueñas. Figure made by Dr. Maria Dueñas used with permission.

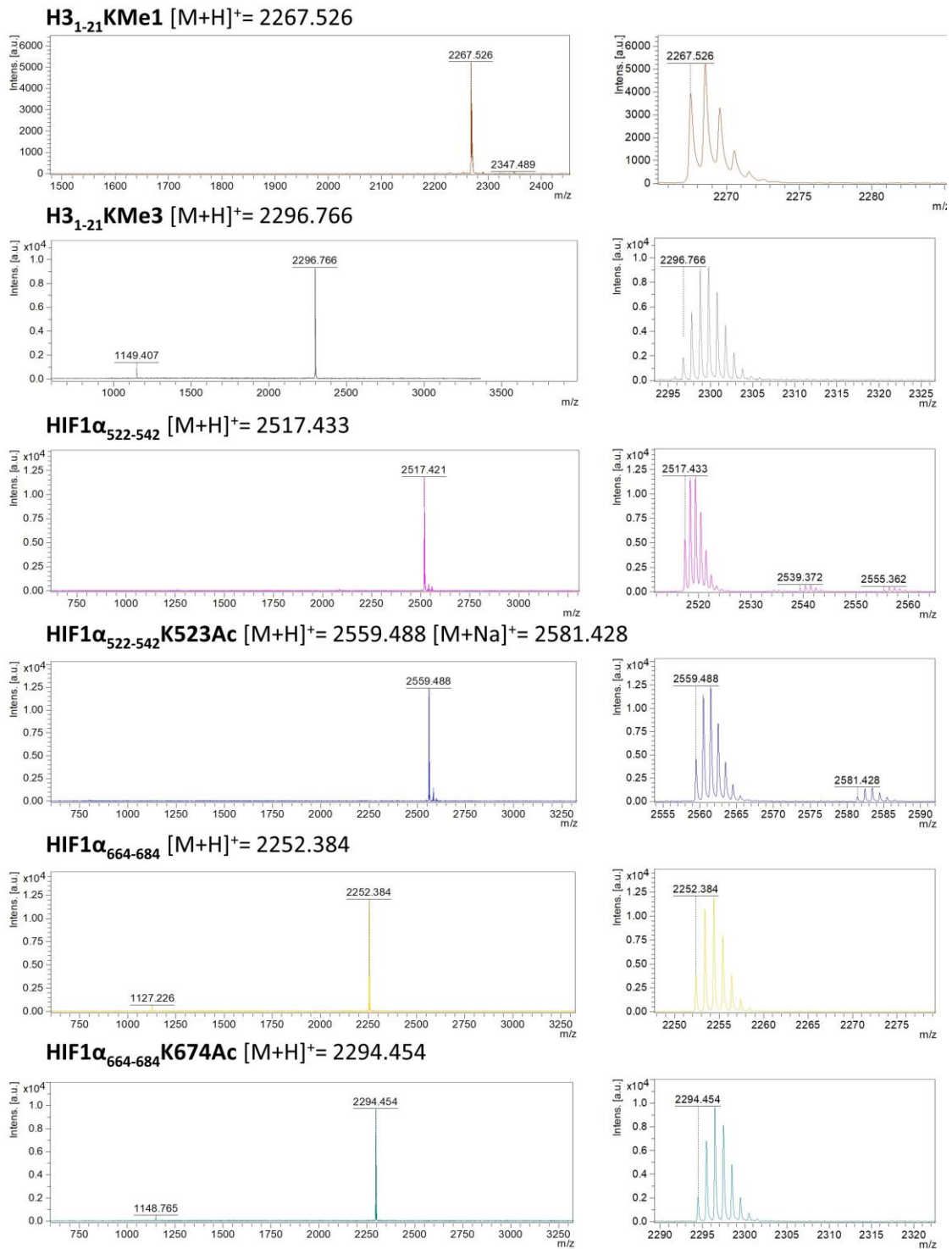
B. Inactive PCAF HAT and HAT-BRD inhibitors

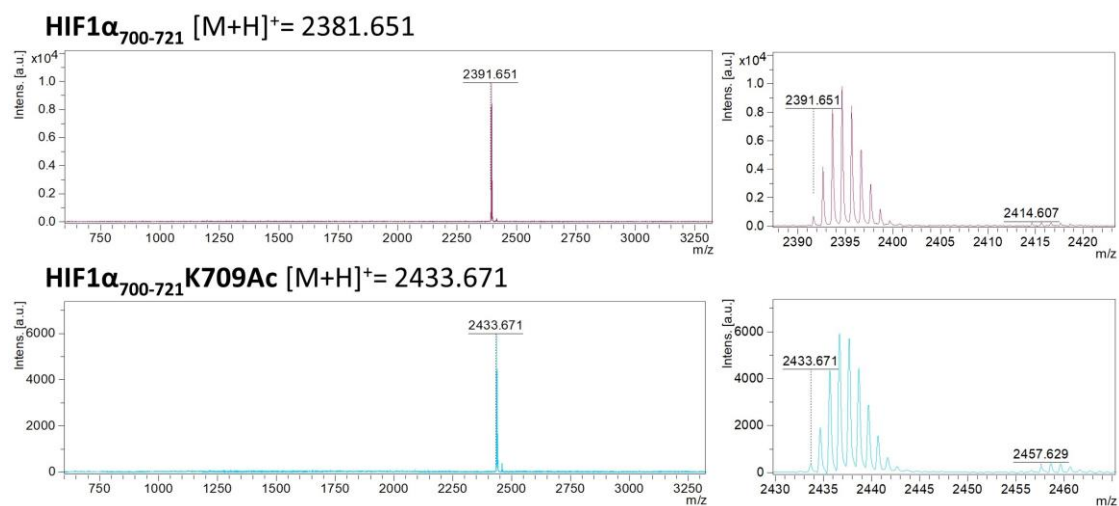


Appendix Figure B.1: HAT inhibitors showing no inhibitory of PCAF HAT/HAT-BRD activity.

C. MALDI-TOF-MS mass verification of synthesised peptides

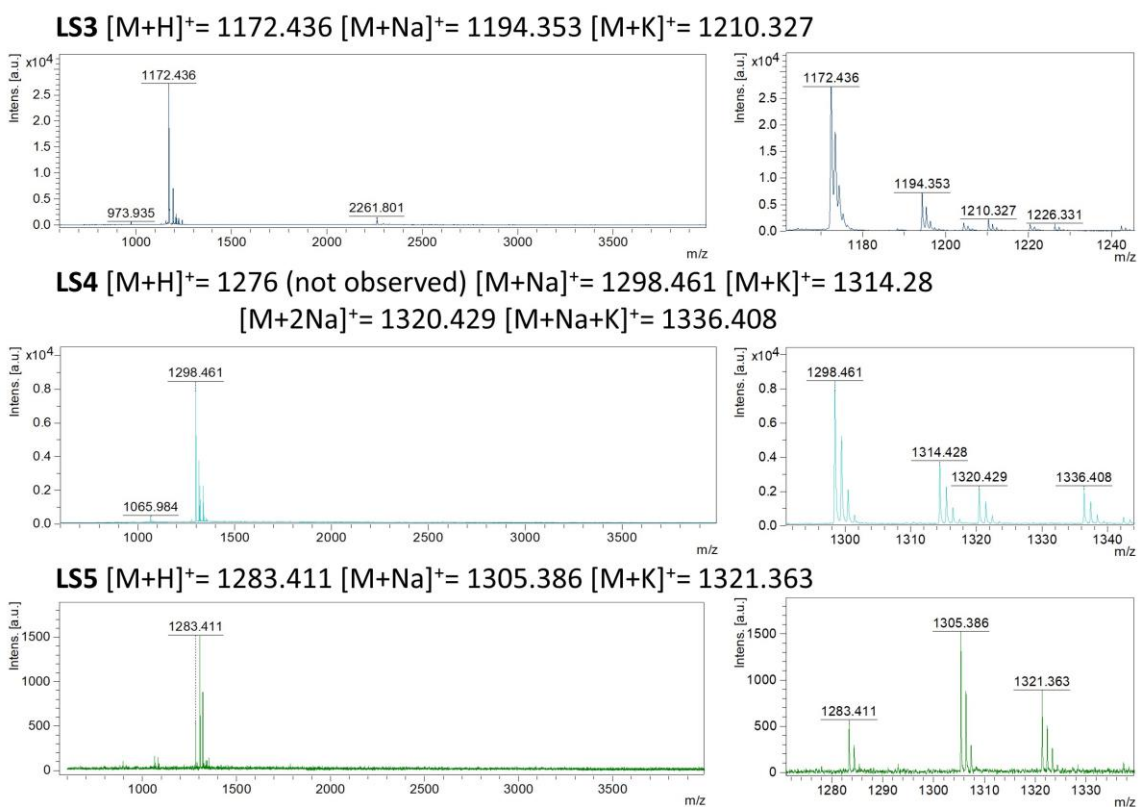
Ci. Linear peptides

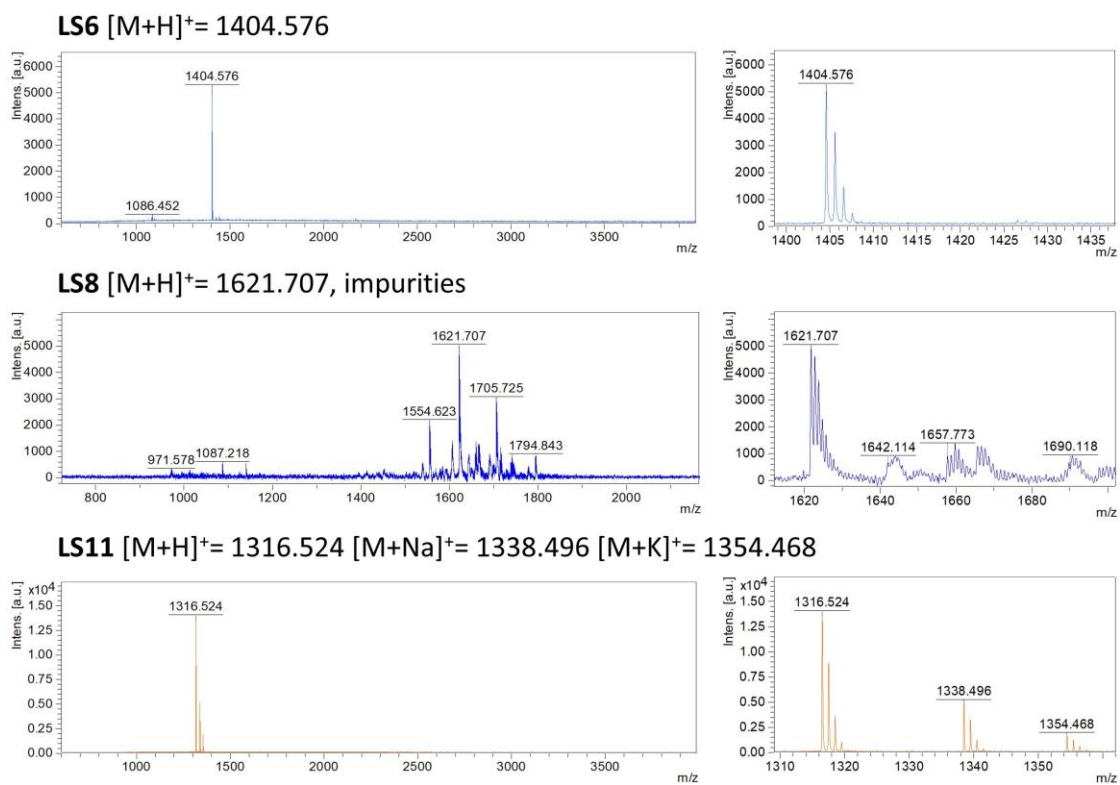




Appendix Figure C.1: MALDI-TOF-MS spectra for linear peptides used in PCAF characterisation.

Cii. Cyclic peptides





Appendix Figure C.2: MALDI-TOF-MS spectra for cyclic peptides tested against PCAF for binding and inhibitory activity.

D. RapifleX MALDI-TOF-MS HTS full result tables

Di. SETI peptide library screened with PCAF HAT and HAT-BRD and NEC

SETI library					PCAF HAT									PCAF HAT-BRD									NEC								
					parent			1Ac			2Ac			parent			1Ac			2Ac			parent			1Ac			2Ac		
#	well	hist	length	peptide PTM	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N
1	A1	H3	1-21	non biotinylated	0	0	9	30	6	9	69	6	9	0	0	7	54	4	7	46	4	7	100	0	3	0	0	3	0	0	3
2	A2	H3	1-21		1	1	7	46	6	7	53	6	7	1	1	12	70	10	12	29	10	12	100	0	7	0	0	7	0	0	7
3	A3	H3	1-21	R2Me	1	1	12	60	4	12	39	4	12	2	1	12	62	19	12	36	20	12	99	0	8	1	0	8	0	0	8
4	A4	H3	1-21	R2Me2a	3	2	12	63	4	12	34	5	12	2	1	12	70	6	12	28	6	12	100	1	8	0	1	8	0	0	8
5	A5	H3	1-21	R2Me2aT3Phos	1	1	12	85	5	12	15	5	12	50	49	8	45	44	8	5	5	8	100	0	8	0	0	8	0	0	8
6	A6	H3	1-21	R2Me2aT3PhosK4Me3	0	0	11	89	6	11	11	6	11	1	1	8	94	3	8	5	2	8	100	0	7	0	0	7	0	0	7
7	A7	H3	1-21	R2Me2aK4Me3	2	0	12	64	2	12	34	2	12	1	1	8	76	2	8	22	2	8	99	1	8	1	1	8	0	0	8
8	A8	H3	1-21	T3Phos	0	1	8	86	10	8	14	9	8	0	1	9	91	5	9	9	5	9	100	0	6	0	0	6	0	0	6
9	A9	H3	1-21	T3PhosK4Me3	0	1	11	89	6	11	11	6	11	0	1	8	94	5	8	6	5	8	100	1	7	0	1	7	0	0	7
10	A10	H3	1-21	K4Me	1	1	10	62	2	10	36	3	10	2	1	8	72	4	8	26	4	8	100	0	5	0	0	5	0	0	5
11	A11	H3	1-21	K4Me2	1	1	11	60	4	11	39	3	11	1	1	8	72	2	8	26	3	8	100	1	7	0	0	7	0	0	7
12	A12	H3	1-21	K4Me3	0	1	7	57	28	7	43	28	7	1	2	8	73	5	8	26	4	8	99	1	7	1	1	7	0	0	7
13	B1	H3	1-21	K4Me3K9Ac	1	2	12	96	3	12	3	2	12	1	1	8	98	2	8	1	1	8	100	0	8	0	0	8	0	0	8
14	B2	H3	1-21	K4Me3K9Me3	0	0	4	100	0	4	0	0	4	0	0	9	100	0	9	0	0	9	100	0	4	0	0	4	0	0	4
15	B3	H3	1-21	K9Ac	2	1	12	96	1	12	3	0	12	3	1	12	95	1	12	2	1	12	100	1	8	0	1	8	0	0	8
16	B4	H3	1-21	K9AcS10Phos	0	1	10	100	1	10	0	0	10	1	2	12	99	2	12	0	0	12	100	0	8	0	0	8	0	0	8
17	B5	H3	1-21	K9AcT11Phos	61	4	12	39	3	12	0	1	12	87	12	8	13	12	8	0	0	8	100	0	8	0	0	8	0	0	8
18	B6	H3	1-21	K9AcS10PhosT11Phos	88	4	12	12	4	12	0	0	12	90	1	8	10	1	8	0	0	8	99	1	6	1	1	6	0	0	6
19	B7	H3	1-21	K9Me	2	1	12	93	1	12	5	1	12	2	1	12	95	2	12	3	1	12	100	0	7	0	0	7	0	0	7
20	B8	H3	1-21	K9Me2	0	1	12	97	2	12	3	2	12	1	1	11	97	2	11	2	2	11	100	0	8	0	0	8	0	0	8
21	B9	H3	1-21	K9Me3	1	1	12	95	2	12	5	2	12	1	1	7	95	1	7	4	1	7	99	1	7	1	1	7	0	0	7
22	B10	H3	1-21	K9Me3T11Phos	22	8	12	78	8	12	0	0	12	35	2	8	65	2	8	0	0	8	100	0	7	0	0	7	0	0	7
23	B11	H3	1-21	K9Me3S10Phos	1	2	12	99	2	12	0	0	12	1	2	9	99	2	9	0	0	9	100	0	5	0	0	5	0	0	5

SETI library					PCAF HAT									PCAF HAT-BRD									NEC								
					parent			1Ac			2Ac			parent			1Ac			2Ac			parent			1Ac			2Ac		
#	well	hist	length	peptide PTM	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N
24	B12	H3	1-21	K9Me3S10PhosT11Phos	81	16	7	19	16	7	0	0	7	88	18	6	12	18	6	0	0	6	99	1	5	1	1	5	0	0	5
25	C1	H3	1-21	S10Phos	0	0	7	99	1	7	1	1	7	0	1	8	100	1	8	0	0	8	100	0	8	0	0	8	0	0	8
26	C2	H3	1-21	T11Phos	3	3	5	96	5	5	2	2	5	18	1	12	81	2	12	1	1	12	100	0	6	0	0	6	0	0	6
27	C3	H3	1-21	S10PhosT11Phos	54	4	12	46	4	12	0	0	12	72	11	10	28	11	10	0	0	10	100	0	7	0	0	7	0	0	7
28	C4	H3	1-21	S10PhosT11PhosK14Ac	100	0	11	0	0	11	0	0	11	100	0	9	0	0	9	0	0	9	100	0	4	0	0	4	0	0	4
29	C5	H3	1-21	S10PhosT11PhosK14Me3	100	0	12	0	0	12	0	0	12	100	0	8	0	0	8	0	0	8	100	1	6	0	1	6	0	0	6
30	C6	H3	1-21	T11PhosK14Ac	97	2	10	3	2	10	0	0	10	97	0	8	3	0	8	0	0	8	100	0	6	0	0	6	0	0	6
31	C7	H3	1-21	T11PhosK14Me3	95	2	12	5	2	12	0	0	12	100	2	11	0	2	11	0	0	11	100	0	8	0	0	8	0	0	8
32	C8	H3	1-21	K14Ac	57	4	12	42	4	12	1	0	12	67	3	12	33	3	12	0	0	12	100	0	6	0	0	6	0	0	6
33	C9	H3	1-21	K14AcR17Me2a	53	6	11	44	7	11	3	2	11	75	11	7	24	11	7	1	2	7	96	2	6	1	1	6	3	2	6
34	C10	H3	1-21	K14Me	3	2	9	56	6	9	41	7	9	2	2	10	54	18	10	44	19	10	100	0	7	0	0	7	0	0	7
35	C11	H3	1-21	K14Me2	55	4	12	43	4	12	2	1	12	61	2	8	38	2	8	1	1	8	100	0	7	0	0	7	0	0	7
36	C12	H3	1-21	K14Me3	57	5	8	42	4	8	1	1	8	68	13	8	32	13	8	0	0	8	98	1	5	2	1	5	0	0	5
37	D1	H3	1-21	K14Me3R17Me2a	45	28	7	47	29	7	7	9	7	59	5	6	39	7	6	2	5	6	93	9	7	3	4	7	4	5	7
38	D2	H3	1-21	R17Me	1	1	4	57	3	4	42	2	4	2	1	10	73	9	10	26	9	10	100	0	8	0	0	8	0	0	8
39	D3	H3	1-21	R17Me2a	1	2	11	55	6	11	43	7	11	3	2	11	68	4	11	29	4	11	93	3	8	3	2	8	4	2	8
40	D4	H3	14-34		36	5	11	49	2	11	15	3	11	41	6	12	47	3	12	13	3	12	100	0	7	0	0	7	0	0	7
41	D5	H3	14-34	R17Me2aK18Ac	63	2	12	33	2	12	4	1	12	75	1	8	24	1	8	2	0	8	99	1	8	1	1	8	0	0	8
42	D6	H3	14-34	K18Ac	47	2	12	44	1	12	9	1	12	59	3	8	37	3	8	4	1	8	100	0	7	0	0	7	0	0	7
43	D7	H3	14-34	K18AcK23Me3	69	4	10	28	3	10	2	1	10	81	1	8	18	1	8	1	0	8	98	1	8	2	1	8	0	0	8
44	D8	H3	14-34	K18AcK23Me	64	10	12	33	9	12	3	2	12	71	3	12	28	2	12	1	0	12	100	0	6	0	0	6	0	0	6
45	D9	H3	14-34	K18AcK23Ac	77	11	11	22	10	11	1	1	11	79	7	7	21	7	7	0	0	7	100	0	8	0	0	8	0	0	8
46	D10	H3	14-34	K23Ac	70	8	12	29	7	12	2	1	12	72	2	9	27	2	9	1	0	9	100	0	7	0	0	7	0	0	7
47	D11	H3	14-34	K23AcR26Me2a	67	3	11	27	3	11	5	1	11	78	2	8	18	2	8	4	2	8	96	3	7	1	1	7	3	2	7
48	D12	H3	14-34	K23AcR26Me2aK27Ac	69	27	8	26	29	8	4	3	8	79	8	11	18	8	11	3	3	11	91	1	4	4	1	4	5	0	4
49	E1	H3	14-34	K23Me	62	17	8	34	14	8	4	3	8	66	2	8	32	2	8	2	0	8	99	0	8	1	0	8	0	0	8

SETI library					PCAF HAT									PCAF HAT-BRD									NEC								
					parent			1Ac			2Ac			parent			1Ac			2Ac			parent			1Ac			2Ac		
#	well	hist	length	peptide PTM	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N
50	E2	H3	14-34	K23Me2	58	2	4	37	1	4	5	0	4	71	5	12	28	5	12	2	1	12	99	0	8	1	0	8	0	0	8
51	E3	H3	14-34	K23Me3	56	3	12	38	3	12	5	1	12	67	8	11	30	7	11	3	2	11	99	1	8	1	1	8	0	0	8
52	E4	H3	14-34	K23Me3R26Me2a	45	9	12	40	6	12	16	3	12	60	3	12	26	3	12	14	2	12	75	4	7	10	2	7	15	2	7
53	E5	H3	14-34	K23Me3R26Me2aK27Ac	68	3	12	24	2	12	8	1	12	73	2	8	19	2	8	8	0	8	85	3	5	6	2	5	9	1	5
54	E6	H3	14-34	R26Me	41	4	11	46	3	11	14	3	11	56	6	8	38	4	8	6	2	8	99	1	7	0	0	7	0	0	7
55	E7	H3	14-34	R26Me2a	37	5	12	47	2	12	16	3	12	49	7	8	41	4	8	10	4	8	93	6	7	3	3	7	4	3	7
56	E8	H3	14-34	R26Me2aK27Ac	50	7	12	40	6	12	10	2	12	47	18	10	44	19	10	9	7	10	92	3	8	3	3	8	5	3	8
57	E9	H3	14-34	R26Me2aK27AcS28Phos	57	5	12	39	5	12	5	3	12	79	0	4	20	0	4	1	0	4	100	0	4	0	0	4	0	0	4
58	E10	H3	14-34	R26Me2aS28Phos	58	8	12	37	6	12	4	3	12	67	3	8	33	3	8	0	1	8	100	0	4	0	0	4	0	0	4
59	E11	H3	14-34	K27Ac	57	13	12	37	12	12	5	3	12	62	6	8	35	5	8	3	1	8	99	0	4	1	0	4	0	0	4
60	E12	H3	14-34	K27AcS28Phos	60	4	12	37	2	12	3	2	12	82	8	7	18	8	7	0	0	7	100	0	5	0	0	5	0	0	5
61	F1	H3	14-34	K27Me	45	11	8	47	6	8	8	4	8	60	4	8	36	2	8	4	2	8	99	0	8	1	0	8	0	0	8
62	F2	H3	14-34	K27Me2	50	5	4	44	4	4	6	1	4	60	7	12	36	6	12	3	2	12	100	1	8	0	0	8	0	0	8
63	F3	H3	14-34	K27Me3	45	3	12	47	2	12	8	1	12	63	7	11	34	5	11	4	2	11	98	2	7	2	2	7	0	0	7
64	F4	H3	14-34	R26Me2aK27Me3	31	19	12	42	18	12	27	14	12	55	16	11	28	12	11	17	8	11	59	29	7	15	13	7	25	18	7
65	F5	H3	14-34	K27Me3S28Phos	60	4	9	37	4	9	3	2	9	72	3	8	26	3	8	1	1	8	100	0	5	0	0	5	0	0	5
66	F6	H3	14-34	R26Me2aK27Me3S28Phos	56	5	10	40	5	10	5	3	10	74	2	8	26	2	8	0	1	8	99	1	5	1	1	5	0	0	5
67	F7	H3	14-34	S28Phos	51	16	12	45	17	12	4	3	12	72	7	12	26	5	12	2	2	12	100	0	4	0	0	4	0	0	4
68	F8	H3	18-38	S28PhosT32Phos	100	0	10	0	0	10	0	0	10	100	0	7	0	0	7	0	0	7	96	4	4	4	4	4	0	0	4
69	F9	H3	28-48		92	3	9	8	3	9	0	0	9	89	2	8	11	2	8	0	0	8	99	1	7	1	1	7	0	0	7
70	F10	H3	28-48	T32Phos	99	1	9	1	1	9	0	0	9	98	1	8	2	1	8	0	0	8	100	0	4	0	0	4	0	0	4
71	F11	H3	28-48	K36Me	99	1	12	1	1	12	0	0	12	100	1	7	0	1	7	0	0	7	99	0	4	1	0	4	0	0	4
72	F12	H3	28-48	K36Me2	99	1	12	1	1	12	0	0	12	99	1	8	1	1	8	0	0	8	99	1	4	1	1	4	0	0	4
73	G1	H3	28-48	K36Me3	99	1	9	1	1	9	0	0	9	99	1	12	1	1	12	0	0	12	99	1	8	1	1	8	0	0	8
74	G2	H3	28-48	T32PhosK36Me3	100	0	4	0	0	4	0	0	4	100	0	11	0	0	11	0	0	11	100	1	8	0	1	8	0	0	8
75	G3	H3	28-48	K37Me	90	2	12	10	2	12	0	0	12	91	3	12	9	3	12	0	0	12	99	1	6	1	1	6	0	0	6

SETI library					PCAF HAT									PCAF HAT-BRD									NEC								
					parent			1Ac			2Ac			parent			1Ac			2Ac			parent			1Ac			2Ac		
#	well	hist	length	peptide PTM	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N
76	G4	H3	28-48	K37Me2	90	2	11	8	1	11	2	1	11	88	4	10	10	4	10	2	1	10	98	1	7	1	1	7	1	1	7
77	G5	H3	28-48	K37Me3	95	2	12	5	2	12	0	0	12	93	0	8	7	0	8	0	0	8	99	1	8	1	1	8	0	0	8
78	G6	H3	28-48	T32PhosK37Me3	99	1	11	1	1	11	0	0	11	98	1	8	2	1	8	0	0	8	100	0	7	0	0	7	0	0	7
79	G7	H3	28-48	K36Me3K37Me3	99	2	9	1	2	9	0	0	9	100	0	8	0	0	8	0	0	8	99	1	7	1	1	7	0	0	7
80	G8	H3	28-48	T32PhosK36Me3K37Me3	100	0	8	0	0	8	0	0	8	100	0	7	0	0	7	0	0	7	50	50	6	50	50	6	0	0	6
81	G9	H3	71-91		0	0	0	0	0	0	0	0	0	100	0	1	0	0	1	0	0	1			0			0			0
82	G10	H3	71-91	K79Me	0	0	0	0	0	0	0	0	0	100	0	8	0	0	8	0	0	8			0			0			0
83	G11	H3	71-91	K79Me2	0	0	0	0	0	0	0	0	0	100	0	2	0	0	2	0	0	2			0			0			0
84	G12	H3	71-91	K79Me3	100	0	4	0	0	4	0	0	4	100	0	1	0	0	1	0	0	1	99	1	4	1	1	4	0	0	4
85	H1	H3	110-130		64	25	3	36	25	3	0	0	3	75	4	7	25	4	7	0	0	7	67	8	4	33	8	4	0	0	4
86	H2	H3	110-130	K115Ac	0	0	0	0	0	0	0	0	0	86	6	8	14	6	8	0	0	8	88	1	2	11	0	2	1	1	2
87	H3	H3	110-130	K115AcT118Phos	87	13	6	13	13	6	0	0	6	87	5	8	13	5	8	0	0	8	100	0	4	0	0	4	0	0	4
88	H4	H3	110-130	T118Phos	0	0	0	0	0	0	0	0	0	91	7	8	9	7	8	0	0	8			0			0			0
89	H5	H3	110-130	K122Ac	76	17	9	24	17	9	0	0	9	82	3	8	18	3	8	0	0	8	88	9	3	12	9	3	0	0	3
90	H6	H3	110-130	T118PhosK122Ac	86	14	6	14	14	6	0	0	6	81	5	8	19	5	8	0	0	8	92	11	3	8	11	3	0	0	3
91	H7	H3	110-130	K115AcT118PhosK122Ac	82	8	7	18	8	7	0	0	7	94	6	8	6	6	8	0	0	8	79	21	4	21	21	4	0	0	4
92	H8	H3	115-135		100	1	8	0	1	8	0	0	8	100	1	9	0	1	9	0	0	9	99	1	3	1	1	3	0	0	3
93	H9	H3	115-135	R128Me	100	0	7	0	0	7	0	0	7	99	2	4	1	2	4	0	0	4	100	0	1	0	0	1	0	0	1
94	H10	H3	115-135	R128Me2a	99	2	8	1	2	8	0	0	8	99	2	8	1	2	8	0	0	8	100	0	4	0	0	4	0	0	4
95	H11	H3	115-135	K122AcR128Me2a	100	0	8	0	0	8	0	0	8	97	2	4	3	2	4	0	0	4	100	0	2	0	0	2	0	0	2
96	H12				0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			0			0			0	

Appendix Table D.1: Results of rapifleX MALDI-TOF-MS HTS of the SETI library with PCAF HAT, PCAF HAT-BRD and a NEC.

Peptide number (#), well in library and assay (well), histone type (hist), peptide length, and peptide PTMs are listed for each peptide. The results of the screen are shown as an average of the observed relative percentages (%) of peak intensities for parent peptide, singly acetylated peptide (1Ac) and doubly acetylated peptide (2Ac) with associated standard deviation (SD) and number of replicates (N) for either PCAF HAT, PCAF HAT-BRD or a NEC. Colour coding (3 tone scale, generated in Excel) is used for % (0-100%, red to green), SD (0-100%, white to blue), and N (0-12 replicates, white to orange).

Dii.SETII peptide library screened with PCAF HAT and HAT-BRD and NEC

SETII library					HAT									HAT-BRD									NEC								
					parent			1Ac			2Ac			parent			1Ac			2Ac			parent			1Ac			2Ac		
#	well	hist	length	peptide PTM	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N
1	A1	H3	1-21	non-biotinylated control	2	1	12	66	3	12	32	3	12	1	1	12	83	7	12	16	7	12	99	1	9	1	0	9	0	0	9
2	A2	H4	1-21		71	3	12	28	2	12	1	0	12	73	6	12	27	6	12	0	0	12	97	2	9	3	2	9	0	0	9
3	A3	H4	1-21	S1Phos	93	5	12	7	5	12	0	0	12	92	8	12	8	8	12	0	0	12	100	1	9	0	1	9	0	0	9
4	A4	H4	1-21	R3Me	66	3	12	34	3	12	1	1	12	67	5	12	33	5	12	0	0	12	99	1	9	1	1	9	0	0	9
5	A5	H4	1-21	R3Me2a	64	4	12	35	4	12	1	1	12	64	5	12	35	4	12	1	1	12	98	1	9	2	1	9	0	0	9
6	A6	H4	1-21	R3Me2aK5Ac	56	6	12	44	6	12	0	1	12	64	5	12	36	4	12	0	0	12	99	1	9	1	1	9	0	0	9
7	A7	H4	1-21	S1PhosR3Me2aK5Ac	92	8	12	8	8	12	0	0	12	99	3	11	1	3	11	0	0	11	100	0	9	0	0	9	0	0	9
8	A8	H4	1-21	K5Ac	41	6	12	52	5	12	8	2	12	45	8	12	47	6	12	8	3	12	70	7	9	30	7	9	0	0	9
9	A9	H4	1-21	K5AcK12Ac	73	3	12	27	3	12	0	0	12	80	4	8	20	4	8	0	0	8	99	0	9	1	0	9	0	0	9
10	A10	H4	1-21	K5AcK8Ac	98	1	12	2	1	12	0	0	12	99	1	12	1	1	12	0	0	12	100	0	9	0	0	9	0	0	9
11	A11	H4	1-21	K8Ac	96	1	12	4	1	12	0	0	12	96	1	12	4	1	12	0	0	12	99	1	9	1	1	9	0	0	9
12	A12	H4	1-21	K8AcK12Ac	98	1	12	2	1	12	0	0	12	100	1	12	0	1	12	0	0	12	100	0	9	0	0	9	0	0	9
13	B1	H4	1-21	K12Me	50	19	12	50	19	12	0	0	12	60	8	12	40	8	12	0	0	12	97	5	9	3	5	9	0	0	9
14	B2	H4	1-21	K12Me2	70	4	12	30	4	12	1	1	12	72	3	12	27	3	12	1	1	12	99	1	9	1	1	9	0	0	9
15	B3	H4	1-21	K12Me3	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	100	0	1	100	0	1	0	0	1	0	0	1
16	B4	H4	1-21	K8AcK12Me	98	0	12	2	0	12	0	0	12	99	1	12	1	1	12	0	0	12	100	0	9	0	0	9	0	0	9
17	B5	H4	1-21	K8AcK12Me3	0	0	2	100	0	2	0	0	2	0	0	2	50	50	2	50	50	2	13	33	8	88	33	8	0	0	8
18	B6	H4	1-21	K12MeK16Ac	67	3	12	30	3	12	2	1	12	68	4	12	30	4	12	2	1	12	87	5	9	11	4	9	2	1	9
19	B7	H4	1-21	K12Me3K16Ac	0	0	1	0	0	1	100	0	1	0	0	0	0	0	0	0	0	0	100	0	1	0	0	1	0	0	1
20	B8	H4	1-21	K16Ac	69	5	12	31	5	12	0	1	12	70	6	12	30	6	12	0	0	12	98	1	9	2	1	9	0	0	9
21	B9	H4	13-33		96	1	12	4	1	12	0	0	12	97	0	12	3	0	12	0	0	12	97	1	9	3	1	9	0	0	9
22	B10	H4	13-33	K20Ac	96	1	12	4	1	12	0	0	12	96	0	12	4	0	12	0	0	12	97	2	9	3	2	9	0	0	9
23	B11	H4	13-33	K16AcK20Ac	97	1	12	3	1	12	0	0	12	97	1	12	3	0	12	0	0	12	97	1	9	3	1	9	0	0	9
24	B12	H4	13-33	K16AcK20Me	96	1	12	4	1	12	0	0	12	97	1	12	3	1	12	0	0	12	97	2	9	3	2	9	0	0	9

SETII library					HAT									HAT-BRD									NEC								
					parent			1Ac			2Ac			parent			1Ac			2Ac			parent			1Ac			2Ac		
#	well	hist	length	peptide PTM	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N
25	C1	H4	13-33	K16Ack20Me3	0	0	4	75	43	4	25	43	4	38	48	8	58	46	8	5	12	8	68	40	4	32	40	4	0	0	4
26	C2	H4	13-33	K20Me	97	2	12	2	2	12	0	0	12	96	2	12	3	2	12	0	0	12	97	1	9	3	1	9	0	0	9
27	C3	H4	13-33	K20Me2	97	2	12	3	2	12	0	0	12	97	1	12	3	1	12	0	0	12	97	1	9	3	1	9	0	0	9
28	C4	H4	13-33	K20Me3	33	47	3	67	47	3	0	0	3	50	50	6	50	50	6	0	0	6	30	43	6	70	43	6	0	0	6
29	C5	H4	43-63		99	3	12	1	3	12	0	0	12	92	4	12	8	4	12	0	1	12	100	0	6	0	0	6	0	0	6
30	C6	H4	43-63	S47Phos	100	0	4	0	0	4	0	0	4	99	2	9	1	2	9	0	0	9			0			0			0
31	C7	H4	43-63	K59Ac	100	0	2	0	0	2	0	0	2	100	0	4	0	0	4	0	0	4			0			0			0
32	C8	H4	51-71		96	9	11	4	9	11	0	0	11	99	4	12	1	4	12	0	0	12			0			0			0
33	C9	H4	51-71	K59Ac	100	0	1	0	0	1	0	0	1	100	0	5	0	0	5	0	0	5			0			0			0
34	C10	H4	68-88		98	1	12	2	1	12	0	0	12	99	1	12	1	1	12	0	0	12	97	6	9	3	6	9	0	0	9
35	C11	H4	68-88	K77Ac	99	1	12	1	1	12	0	0	12	99	1	8	1	1	8	0	0	8	98	2	9	2	2	9	0	0	9
36	C12	H4	68-88	K77Ack79Ac	99	1	12	1	1	12	0	0	12	99	1	11	1	1	11	0	0	11			0			0			0
37	D1	H4	68-88	K79Ac	99	1	12	1	1	12	0	0	12	99	1	12	1	1	12	0	0	12	100	0	6	0	0	6	0	0	6
38	D2	H4	68-88	K77Ack79Me	99	1	12	1	1	12	0	0	12	99	0	12	1	0	12	0	0	12	100	1	9	0	1	9	0	0	9
39	D3	H4	68-88	K77Ack79Me3	0	0	9	100	0	9	0	0	9	0	0	0	0	0	0	0	0	0	50	50	2	50	50	2	0	0	2
40	D4	H4	68-88	K79Me	99	1	11	1	1	11	0	0	11	99	1	12	1	1	12	0	0	12	99	1	9	1	1	9	0	0	9
41	D5	H4	68-88	K79Me2	98	1	12	2	1	12	0	0	12	98	1	12	2	1	12	0	0	12	99	1	9	1	1	9	0	0	9
42	D6	H4	68-88	K79Me3	0	0	8	100	0	8	0	0	8	0	0	4	100	0	4	0	0	4	25	43	4	75	43	4	0	0	4
43	D7	H4	82-102		0	0	12	88	9	12	12	9	12	0	0	12	94	11	12	6	11	12	11	31	9	83	31	9	6	8	9
44	D8	H4	82-102	R92Me	0	0	10	82	13	10	18	13	10	0	0	12	83	15	12	17	15	12	12	32	8	79	31	8	9	12	8
45	D9	H4	82-102	R92Me2a	73	25	11	22	15	11	6	14	11	77	27	10	22	28	10	1	3	10	77	32	7	18	20	7	5	13	7
46	D10	H2a	1-21		84	3	12	16	3	12	0	0	12	87	2	12	13	2	12	0	1	12	99	1	9	1	0	9	0	0	9
47	D11	H2a	1-21	S1Phos	100	0	12	0	0	12	0	0	12	98	2	11	1	2	11	1	1	11	100	0	9	0	0	9	0	0	9
48	D12	H2a	1-21	S1PhosK5Ac	100	0	12	0	0	12	0	0	12	100	0	12	0	0	12	0	0	12	100	0	9	0	0	9	0	0	9
49	E1	H2a	1-21	K5Ac	92	1	12	8	1	12	0	0	12	94	1	12	6	1	12	0	0	12	99	1	9	1	1	9	0	0	9
50	E2	H2a	1-21	K5Ack9Ac	100	0	12	0	0	12	0	0	12	100	1	12	0	1	12	0	0	12	99	1	9	1	0	9	0	0	9

SETII library					HAT									HAT-BRD									NEC								
					parent			1Ac			2Ac			parent			1Ac			2Ac			parent			1Ac			2Ac		
#	well	hist	length	peptide PTM	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N
51	E3	H2a	1-21	K9Ac	98	0	12	2	0	12	0	0	12	99	0	12	1	0	12	0	0	12	98	2	9	2	2	9	0	0	9
52	E4	H2a	1-21	K9AcK13Ac	99	1	12	1	0	12	0	0	12	99	0	12	1	0	12	0	0	12	99	0	9	1	0	9	0	0	9
53	E5	H2a	1-21	K13Ac	90	2	12	10	2	12	0	0	12	91	1	12	9	1	12	0	0	12	99	1	9	1	1	9	0	0	9
54	E6	H2a	1-21	K13AcK15Ac	92	3	12	8	3	12	0	0	12	91	2	12	7	1	12	1	1	12	99	1	9	1	1	9	0	0	9
55	E7	H2a	1-21	K9AcK13AcK15	100	1	12	0	0	12	0	1	12	100	0	12	0	0	12	0	0	12	100	0	9	0	0	9	0	0	9
56	E8	H2a	21-41		69	9	12	29	9	12	2	0	12	69	5	12	30	5	12	2	0	12	83	4	9	16	3	9	1	1	9
57	E9	H2a	21-41	K36Ac	93	3	12	7	3	12	0	0	12	94	1	12	6	1	12	0	0	12	92	3	9	8	3	9	0	0	9
58	E10	H2a	85-105		97	1	12	3	1	12	0	0	12	97	0	12	3	0	12	0	0	12	98	1	9	2	1	9	0	0	9
59	E11	H2a	85-105	K95Ac	65	18	12	11	5	12	24	16	12	93	7	12	0	1	12	7	6	12			0			0			0
60	E12	H2a	85-105	K99Ac	91	8	12	6	4	12	3	5	12	98	2	12	2	2	12	0	0	12	100	0	2	0	0	2	0	0	2
61	F1	H2a	85-105	K95AcK99Ac	100	0	2	0	0	2	0	0	2	100	0	9	0	0	9	0	0	9			0			0			0
62	F2	H2a	110-127		0	0	11	90	15	11	10	15	11	0	0	12	93	6	12	7	6	12	11	31	9	82	28	9	7	3	9
63	F3	H2a	110-127	K118Ac	0	0	12	92	3	12	8	3	12	0	0	12	93	3	12	7	3	12	11	30	9	82	28	9	7	5	9
64	F4	H2a	110-127	K118AcT119Phos	0	0	11	97	4	11	3	4	11	0	0	12	98	3	12	2	3	12	13	29	9	82	28	9	5	4	9
65	F5	H2a	110-127	T119Phos	1	5	12	99	5	12	0	0	12	0	0	12	100	2	12	0	2	12	10	29	9	87	28	9	3	5	9
66	F6	H2b	1-21		92	1	12	8	1	12	0	0	12	95	2	12	5	2	12	0	0	12	97	1	9	3	1	9	0	0	9
67	F7	H2b	1-21	K5Ac	94	1	12	6	1	12	0	0	12	95	1	12	5	1	12	0	0	12	96	1	9	4	1	9	0	0	9
68	F8	H2b	1-21	K5Me	92	3	12	7	3	12	1	0	12	95	1	12	5	0	12	1	0	12	96	1	9	3	1	9	1	0	9
69	F9	H2b	1-21	K5Me2	92	1	12	7	1	12	1	0	12	95	2	12	5	2	12	0	0	12	96	2	9	4	2	9	0	0	9
70	F10	H2b	1-21	K5Me3	69	40	4	6	3	4	25	43	4	0	0	11	3	11	11	97	11	11	11	31	9	28	32	9	61	37	9
71	F11	H2b	1-21	K12Ac	89	3	12	11	3	12	0	0	12	95	1	12	5	1	12	0	0	12	96	1	9	4	1	9	0	0	9
72	F12	H2b	1-21	S14Phos	77	9	12	23	9	12	0	0	12	79	5	12	21	5	12	0	0	12	77	11	9	23	11	9	0	0	9
73	G1	H2b	1-21	K12AcS14Phos	80	7	11	20	7	11	0	0	11	74	3	12	26	3	12	0	0	12	81	8	9	19	8	9	0	0	9
74	G2	H2b	1-21	S14PhosK15Ac	92	9	12	8	9	12	0	0	12	79	7	12	21	7	12	0	0	12	83	14	9	17	14	9	0	0	9
75	G3	H2b	1-21	K12AcS14PhosK15Ac	84	3	12	16	3	12	0	0	12	74	6	12	26	6	12	0	0	12	79	9	9	21	9	9	0	0	9
76	G4	H2b	15-35		94	2	12	6	2	12	0	0	12	94	1	12	6	1	12	0	0	12	96	2	9	4	1	9	0	1	9

SETII library					HAT									HAT-BRD									NEC								
					parent			1Ac			2Ac			parent			1Ac			2Ac			parent			1Ac			2Ac		
#	well	hist	length	peptide PTM	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N
77	G5	H2b	15-35	K20Ac	93	4	12	7	4	12	0	0	12	97	2	12	3	1	12	0	0	12	99	1	9	1	1	9	0	0	9
78	G6	H2b	15-35	K20AcK23Me	96	1	12	4	1	12	0	0	12	97	1	12	3	1	12	0	0	12	97	1	9	3	1	9	0	0	9
79	G7	H2b	15-35	K20AcK23Me3	100	1	12	0	1	12	0	0	12	99	1	12	1	1	12	0	0	12	100	0	9	0	0	9	0	0	9
80	G8	H2b	15-35	K23Me	91	3	12	9	3	12	0	0	12	91	4	12	9	4	12	0	1	12	94	2	9	6	2	9	0	0	9
81	G9	H2b	15-35	K23Me2	97	3	12	3	3	12	0	0	12	98	3	12	2	3	12	0	0	12	98	2	9	2	2	9	0	0	9
82	G10	H2b	15-35	K23Me3	97	4	12	3	4	12	0	0	12	99	3	12	1	3	12	0	0	12	96	3	9	4	3	9	0	0	9
83	G11	H2b	15-35	K23MeK24Ac	97	3	12	3	3	12	0	0	12	99	2	12	1	2	12	0	0	12	98	1	9	2	1	9	0	0	9
84	G12	H2b	15-35	K23Me3K24Ac	99	2	12	1	2	12	0	0	12	100	0	12	0	0	12	0	0	12	100	0	9	0	0	9	0	0	9
85	H1	H2b	15-35	K24Ac	95	2	12	5	2	12	0	0	12	96	2	12	4	2	12	0	0	12	98	2	9	2	1	9	0	0	9
86	H2	H2b	27-47		89	2	11	10	1	11	1	1	11	87	5	12	12	5	12	1	2	12	97	1	9	3	1	9	0	0	9
87	H3	H2b	27-47	S32Phos	90	4	12	10	4	12	0	0	12	98	5	7	2	5	7	0	0	7	100	0	8	0	0	8	0	0	8
88	H4	H2b	27-47	S32PhosS36Phos	94	8	11	6	8	11	0	0	11	99	3	12	1	3	12	0	0	12	98	6	8	0	0	8	2	6	8
89	H5	H2b	27-47	S36Phos	91	3	11	9	3	11	0	0	11	93	4	12	7	4	12	0	0	12	100	0	9	0	0	9	0	0	9
90	H6	H2b	27-47	S36PhosK43Ac	91	8	11	9	8	11	0	0	11	83	7	10	12	6	10	5	3	10	100	0	5	0	0	5	0	0	5
91	H7	H2b	27-47	K43Ac	88	3	12	10	2	12	2	1	12	83	11	12	12	9	12	5	6	12	99	2	7	1	2	7	0	0	7
92	H8	H2b	37-57		93	4	12	7	4	12	0	0	12	92	1	12	8	1	12	0	0	12	91	4	9	9	4	9	0	0	9
93	H9	H2b	37-57	K43Me	95	4	12	5	4	12	0	0	12	92	1	12	8	1	12	0	0	12	90	4	9	10	4	9	0	0	9
94	H10	H2b	37-57	K43Me2	94	2	4	6	2	4	0	0	4	94	2	12	6	2	12	0	0	12	92	3	9	8	3	9	0	0	9
95	H11	H2b	37-57	K43Me3	72	44	11	0	1	11	27	45	11	12	32	8	27	33	8	61	38	8	25	43	4	40	43	4	35	41	4
96	H12				0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			0			0				0

Appendix Table D.2: Results of rapifleX MALDI-TOF-MS HTS of the SETII library with PCAF HAT, PCAF HAT-BRD and a NEC.

Peptide number (#), well in library and assay (well), histone type (hist), peptide length, and peptide PTMs are listed for each peptide. The results of the screen are shown as an average of the observed relative percentages (%) of peak intensities for parent peptide, singly acetylated peptide (1Ac) and doubly acetylated peptide (2Ac) with associated standard deviation (SD) and number of replicates (N) for either PCAF HAT, PCAF HAT-BRD or a NEC. Colour coding (3 tone scale, generated in Excel) is used for % (0-100%, red to green), SD (0-100%, white to blue), and N (0-12 replicates, white to orange).

Diii. SETIII peptide library screened with PCAF HAT and HAT-BRD and NEC

SETIII library					HAT									HAT-BRD									NEC								
					parent			1Ac			2Ac			parent			1Ac			2Ac			parent			1Ac			2Ac		
#	well	Hist	length	peptide PTM	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N
1	A1				0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	A2	H3	1-21		0	1	3	43	15	3	56	16	3	1	1	2	77	6	2	22	5	2	100	0	2	0	0	2	0	0	2
3	A3	H3	1-21	R2Me	0	0	3	58	19	3	42	19	3	0	0	2	75	4	2	25	4	2	100	0	2	0	0	2	0	0	2
4	A4	H3	1-21	R2Me2a	0	0	3	9	6	3	91	6	3	0	0	2	84	1	2	16	1	2	100	0	2	0	0	2	0	0	2
5	A5	H3	1-21	R2Me2aT3Phos	0	0	3	82	6	3	18	6	3	0	0	2	97	3	2	3	3	2	99	1	2	1	1	2	0	0	2
6	A6	H3	1-21	R2Me2aT3PhosK4Me3	33	47	3	58	41	3	9	8	3	3	3	2	95	1	2	2	2	2	100	0	2	0	0	2	0	0	2
7	A7	H3	1-21	R2Me2aK4Me3	0	0	3	56	12	3	44	12	3	0	0	2	83	1	2	17	1	2	99	0	2	1	0	2	0	0	2
8	A8	H3	1-21	T3Phos	0	0	3	81	8	3	19	8	3	0	0	2	93	1	2	7	1	2	100	0	2	0	0	2	0	0	2
9	A9	H3	1-21	T3PhosK4Me3	0	1	3	83	5	3	16	5	3	0	0	2	94	1	2	6	1	2	100	0	2	0	0	2	0	0	2
10	A10	H3	1-21	K4Me	0	0	3	50	11	3	50	12	3	1	1	2	74	0	2	25	0	2	100	0	2	0	0	2	0	0	2
11	A11	H3	1-21	K4Me2	0	0	3	41	10	3	59	10	3	1	0	2	72	2	2	27	2	2	100	0	2	0	0	2	0	0	2
12	A12	H3	1-21	K4Me3	36	45	3	33	25	3	31	26	3	3	3	2	74	6	2	22	2	2	100	0	2	0	0	2	0	0	2
13	B1	H3	1-21	K4Me3K9Ac	3	5	3	95	5	3	1	1	3	0	0	2	100	0	2	0	0	2	100	0	2	0	0	2	0	0	2
14	B2	H3	1-21	K4Me3K9Me3	0	0	3	100	0	3	0	0	3	0	0	2	100	0	2	0	0	2	100	0	1	0	0	1	0	0	1
15	B3	H3	1-21	K9Ac	1	1	3	98	2	3	1	1	3	1	0	2	99	0	2	0	0	2	100	0	2	0	0	2	0	0	2
16	B4	H3	1-21	K9AcS10Phos	0	1	3	99	1	3	1	1	3	1	1	2	99	1	2	0	0	2	100	0	2	0	0	2	0	0	2
17	B5	H3	1-21	K9AcT11Phos	25	9	3	75	9	3	1	1	3	78	1	2	22	1	2	0	0	2	100	0	2	0	0	2	0	0	2
18	B6	H3	1-21	K9AcS10PhosT11Phos	73	9	3	27	9	3	0	0	3	93	1	2	7	1	2	0	0	2	100	0	2	0	0	2	0	0	2
19	B7	H3	1-21	K9Me	1	0	3	95	1	3	4	1	3	0	0	2	98	0	2	2	0	2	100	0	2	0	0	2	0	0	2
20	B8	H3	1-21	K9Me2	1	1	3	97	2	3	2	1	3	0	0	2	100	0	2	0	0	2	100	0	2	0	0	2	0	0	2
21	B9	H3	1-21	K9Me3	2	3	3	97	4	3	2	1	3	0	0	2	99	1	2	1	1	2	100	0	2	0	0	2	0	0	2
22	B10	H3	1-21	K9Me3T11Phos	10	7	3	90	8	3	0	1	3	45	3	2	55	3	2	0	0	2	100	0	2	0	0	2	0	0	2
23	B11	H3	1-21	K9Me3S10Phos	1	1	3	99	1	3	0	0	3	1	1	2	99	1	2	0	0	2	100	0	2	0	0	2	0	0	2
24	B12	H3	1-21	K9Me3S10PhosT11Phos	46	18	3	54	18	3	0	0	3	79	2	2	21	2	2	0	0	2	99	1	2	1	1	2	0	0	2

SETIII library					HAT									HAT-BRD									NEC								
					parent			1Ac			2Ac			parent			1Ac			2Ac			parent			1Ac			2Ac		
#	well	Hist	length	peptide PTM	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N
25	C1	H3	1-21	S10Phos	0	0	3	99	1	3	1	1	3	1	1	2	99	1	2	0	0	2	100	0	2	0	0	2	0	0	2
26	C2	H3	1-21	T11Phos	2	2	3	94	2	3	4	1	3	23	2	2	77	2	2	0	0	2	100	0	2	0	0	2	0	0	2
27	C3	H3	1-21	S10PhosT11Phos	28	18	3	72	18	3	0	0	3	73	3	2	27	3	2	0	0	2	100	0	2	0	0	2	0	0	2
28	C4	H3	4-24	S10PhosT11PhosK14Ac	98	2	3	2	2	3	0	0	3	100	0	2	0	0	2	0	0	2	100	0	2	0	0	2	0	0	2
29	C5	H3	4-24	S10PhosT11PhosK14Me3	96	1	3	4	1	3	0	0	3	99	1	2	1	1	2	0	0	2	100	0	2	0	0	2	0	0	2
30	C6	H3	4-24	T11PhosK14Ac	95	2	3	5	2	3	0	0	3	99	1	2	1	1	2	0	0	2	100	0	2	0	0	2	0	0	2
31	C7	H3	4-24	T11PhosK14Me3	92	3	3	8	3	3	0	0	3	98	0	2	2	0	2	0	0	2	100	0	2	0	0	2	0	0	2
32	C8	H3	4-24	K9AcK14Ac	92	3	3	8	3	3	0	0	3	97	0	2	3	0	2	0	0	2	100	0	2	0	0	2	0	0	2
33	C9	H3	8-28	K14Ac	60	13	3	37	11	3	3	3	3	80	0	2	19	0	2	1	0	2	100	0	2	0	0	2	0	0	2
34	C10	H3	8-28	K14AcR17Me2a	73	9	3	26	8	3	1	1	3	90	0	2	10	0	2	0	0	2	100	0	2	0	0	2	0	0	2
35	C11	H3	8-28	K14Me	3	2	3	67	10	3	30	12	3	7	2	2	73	0	2	20	2	2	100	0	2	0	0	2	0	0	2
36	C12	H3	8-28	K14Me2	57	11	3	39	8	3	4	2	3	73	0	2	24	1	2	2	0	2	100	0	2	0	0	2	0	0	2
37	D1	H3	8-28	K14Me3	56	12	3	40	9	3	5	3	3	81	7	2	19	7	2	0	0	2	100	0	2	0	0	2	0	0	2
38	D2	H3	8-28	K14Me3R17Me2a	65	12	3	32	10	3	3	2	3	88	2	2	12	1	2	1	1	2	98	0	2	2	0	2	0	0	2
39	D3	H3	8-28	R17Me	1	0	3	75	11	3	25	12	3	1	1	2	89	0	2	11	1	2	99	0	2	1	0	2	0	0	2
40	D4	H3	8-28	R17Me2a	0	0	3	73	6	3	27	6	3	1	0	2	90	2	2	9	2	2	99	1	2	1	1	2	0	0	2
41	D5	H3	8-28	K14AcK18Ac	68	7	3	31	6	3	1	1	3	89	2	2	11	2	2	0	0	2	100	0	2	0	0	2	0	0	2
42	D6	H3	12-32	R17Me2aK18Ac	40	16	3	56	13	3	4	3	3	89	0	2	11	0	2	0	0	2	99	0	2	1	0	2	0	0	2
43	D7	H3	12-32	K18Ac	43	8	3	52	6	3	5	2	3	82	2	2	18	2	2	1	1	2	99	0	2	1	0	2	0	0	2
44	D8	H3	12-32	K18AcK23Me3	45	12	3	53	11	3	2	1	3	90	1	2	10	1	2	0	0	2	100	0	2	0	0	2	0	0	2
45	D9	H3	12-32	K18AcK23Ac	78	8	3	22	8	3	0	0	3	91	1	2	9	1	2	0	0	2	100	0	2	0	0	2	0	0	2
46	D10	H3	18-38	K23Ac	86	3	3	14	3	3	0	0	3	97	1	2	3	1	2	0	0	2	100	0	2	0	0	2	0	0	2
47	D11	H3	18-38	K23AcR26Me2a	78	6	3	19	5	3	3	1	3	89	1	2	10	0	2	1	1	2	89	5	2	10	5	2	2	0	2
48	D12	H3	18-38	K23AcR26Me2aK27Ac	89	0	3	10	0	3	0	0	3	88	1	2	12	1	2	0	0	2	93	0	2	7	1	2	0	0	2
49	E1	H3	18-38	K23Me	82	7	3	18	6	3	0	0	3	96	0	2	4	0	2	0	0	2	100	0	2	0	0	2	0	0	2
50	E2	H3	18-38	K23Me2	80	6	3	20	5	3	0	0	3	95	0	2	5	0	2	0	0	2	100	0	2	0	0	2	0	0	2

SETIII library					HAT									HAT-BRD									NEC								
					parent			1Ac			2Ac			parent			1Ac			2Ac			parent			1Ac			2Ac		
#	well	Hist	length	peptide PTM	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N
51	E3	H3	18-38	K23Me3	84	8	3	16	8	3	0	0	3	95	0	2	5	0	2	0	0	2	99	0	2	1	0	2	0	0	2
52	E4	H3	18-38	K23Me3R26Me2a	54	5	3	31	5	3	15	2	3	76	1	2	11	1	2	12	1	2	81	1	2	6	0	2	14	1	2
53	E5	H3	18-38	K23Me3R26Me2aK27Ac	76	3	3	12	2	3	11	1	3	75	4	2	11	4	2	14	1	2	81	4	2	7	1	2	12	2	2
54	E6	H3	20-40		91	3	3	9	3	3	0	0	3	95	1	2	5	1	2	0	0	2	100	0	2	0	0	2	0	0	2
55	E7	H3	20-40	R26Me	91	3	3	9	3	3	0	0	3	94	0	2	6	0	2	0	0	2	100	0	2	0	0	2	0	0	2
56	E8	H3	20-40	R26Me2a	84	4	3	14	4	3	1	1	3	93	1	2	6	1	2	1	0	2	98	1	2	2	1	2	0	0	2
57	E9	H3	20-40	R26Me2aK27Ac	97	2	3	3	2	3	0	0	3	98	2	2	2	2	2	0	0	2	98	0	2	2	0	2	0	0	2
58	E10	H3	20-40	R26Me2aK27AcS28Phos	98	1	3	2	1	3	0	0	3	98	0	2	2	0	2	0	0	2	100	0	2	0	0	2	0	0	2
59	E11	H3	20-40	R26Me2aS28Phos	97	1	3	3	1	3	0	0	3	98	0	2	2	0	2	0	0	2	100	0	2	0	0	2	0	0	2
60	E12	H3	20-40	K27Ac	96	1	3	4	1	3	0	0	3	98	0	2	2	0	2	0	0	2	100	0	2	0	0	2	0	0	2
61	F1	H3	20-40	K27AcS28Phos	98	1	3	2	1	3	0	0	3	100	0	2	0	0	2	0	0	2	100	0	2	0	0	2	0	0	2
62	F2	H3	20-40	K27Me	90	3	3	10	3	3	0	0	3	97	1	2	3	1	2	0	0	2	100	0	2	0	0	2	0	0	2
63	F3	H3	20-40	K27Me2	92	5	3	8	5	3	0	0	3	96	1	2	4	1	2	0	0	2	100	0	2	0	0	2	0	0	2
64	F4	H3	20-40	K27Me3	93	2	3	7	2	3	0	0	3	96	1	2	4	1	2	0	0	2	100	0	2	0	0	2	0	0	2
65	F5	H3	20-40	R26Me2aK27Me3	88	3	3	11	2	3	1	1	3	93	1	2	6	1	2	1	0	2	98	2	2	1	1	2	1	1	2
66	F6	H3	20-40	K27Me3S28Phos	99	1	3	1	1	3	0	0	3	100	0	2	0	0	2	0	0	2	100	0	2	0	0	2	0	0	2
67	F7	H3	20-40	R26Me2aK27Me3S28Phos	98	2	3	2	2	3	0	0	3	99	0	2	1	0	2	0	0	2	100	0	2	0	0	2	0	0	2
68	F8	H3	20-40	S28Phos	97	1	3	3	1	3	0	0	3	100	0	2	0	0	2	0	0	2	100	0	2	0	0	2	0	0	2
69	F9	H4	1-21		70	6	3	30	6	3	0	0	3	69	1	2	31	1	2	0	0	2	99	0	2	1	0	2	0	0	2
70	F10	H4	1-21	S1Phos	89	3	3	11	3	3	0	0	3	89	0	2	11	0	2	0	0	2	100	0	2	0	0	2	0	0	2
71	F11	H4	1-21	R3Me	70	6	3	30	6	3	0	0	3	70	2	2	30	1	2	0	0	2	99	0	2	1	0	2	0	0	2
72	F12	H4	1-21	R3Me2a	68	8	3	31	8	3	1	0	3	69	0	2	30	0	2	0	0	2	99	0	2	1	0	2	0	0	2
73	G1	H4	1-21	R3Me2aK5Ac	59	13	3	41	13	3	0	1	3	62	4	2	38	4	2	0	0	2	99	0	2	1	0	2	0	0	2
74	G2	H4	1-21	S1PhosR3Me2aK5Ac	38	3	3	61	2	3	1	1	3	36	0	2	62	2	2	2	2	2	33	4	2	67	4	2	0	0	2
75	G3	H4	1-21	K5Ac	68	15	3	31	15	3	0	0	3	71	0	2	29	0	2	0	0	2	99	0	2	1	0	2	0	0	2
76	G4	H4	1-21	K5AcK8AcK12Ac	94	1	3	6	1	3	0	0	3	95	0	2	5	0	2	0	0	2	94	1	2	6	1	2	0	0	2

SETIII library					HAT									HAT-BRD									NEC								
					parent			1Ac			2Ac			parent			1Ac			2Ac			parent			1Ac			2Ac		
#	well	Hist	length	peptide PTM	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N
77	G5	H4	1-21	K5AcK8Ac	99	1	3	1	1	3	0	0	3	99	0	2	1	0	2	0	0	2	99	0	2	1	0	2	0	0	2
78	G6	H4	1-21	K8Ac	98	1	3	2	1	3	0	0	3	99	1	2	1	1	2	0	0	2	99	0	2	1	0	2	0	0	2
79	G7	H4	1-21	K8AcK12Ac	99	1	3	1	1	3	0	0	3	99	1	2	1	1	2	0	0	2	99	0	2	1	0	2	0	0	2
80	G8	H4	1-21	K12Me	70	7	3	30	7	3	0	0	3	72	1	2	28	1	2	0	0	2	99	0	2	1	0	2	0	0	2
81	G9	H4	1-21	K12Me2	71	6	3	28	6	3	0	0	3	72	0	2	28	0	2	0	0	2	99	0	2	1	0	2	0	0	2
82	G10	H4	1-21	K12Me3	78	8	3	22	8	3	0	0	3	75	4	2	25	4	2	0	0	2	99	0	2	1	0	2	0	0	2
83	G11	H4	1-21	K8AcK12Me	99	1	3	1	1	3	0	0	3	98	0	2	2	0	2	0	0	2	99	0	2	1	0	2	0	0	2
84	G12	H4	1-21	K8AcK12Me3	98	0	3	2	0	3	0	0	3	98	0	2	2	0	2	0	0	2	99	0	2	1	0	2	0	0	2
85	H1	H4	6-26	K12MeK16Ac	92	4	3	8	4	3	0	0	3	95	0	2	5	0	2	0	0	2	100	0	2	0	0	2	0	0	2
86	H2	H4	6-26	K12Me3K16Ac	99	1	3	1	1	3	0	0	3	98	0	2	2	0	2	0	0	2	100	0	2	0	0	2	0	0	2
87	H3	H4	6-26	K16Ac	93	4	3	7	4	3	0	0	3	97	1	2	3	1	2	0	0	2	99	0	2	1	0	2	0	0	2
88	H4	H4	11-31		99	1	3	1	1	3	0	0	3	99	0	2	1	0	2	0	0	2	100	0	2	0	0	2	0	0	2
89	H5	H4	11-31	K20Ac	99	1	3	1	1	3	0	0	3	100	0	2	0	0	2	0	0	2	100	0	2	0	0	2	0	0	2
90	H6	H4	11-31	K16AcK20Ac	100	0	3	0	0	3	0	0	3	100	0	2	0	0	2	0	0	2	100	0	2	0	0	2	0	0	2
91	H7	H4	11-31	K16AcK20Me	100	0	3	0	0	3	0	0	3	100	0	2	0	0	2	0	0	2	100	0	2	0	0	2	0	0	2
92	H8	H4	11-31	K16AcK20Me3	100	0	3	0	0	3	0	0	3	100	0	2	0	0	2	0	0	2	100	0	2	0	0	2	0	0	2
93	H9	H4	11-31	K20Me	99	0	3	1	0	3	0	0	3	100	0	2	0	0	2	0	0	2	100	0	2	0	0	2	0	0	2
94	H10	H4	11-31	K20Me2	99	1	3	1	1	3	0	0	3	100	0	2	0	0	2	0	0	2	100	0	2	0	0	2	0	0	2
95	H11	H4	11-31	K20Me3	100	0	3	0	0	3	0	0	3	99	0	2	1	0	2	0	0	2	100	0	2	0	0	2	0	0	2
96	H12	H3	1-21	non biotinylated control	0	0	3	50	15	3	50	15	3	0	0	2	76	3	2	24	3	2	100	0	2	0	0	2	0	0	2

Appendix Table D.3: Results of rapifleX MALDI-TOF-MS HTS of the SETIII library with PCAF HAT, PCAF HAT-BRD and a NEC.

Peptide number (#), well in library and assay (well), histone type (hist), peptide length, and peptide PTMs are listed for each peptide. The results of the screen are shown as an average of the observed relative percentages (%) of peak intensities for parent peptide, singly acetylated peptide (1Ac) and doubly acetylated peptide (2Ac) with associated standard deviation (SD) and number of replicates (N) for either PCAF HAT, PCAF HAT-BRD or a NEC. Colour coding (3 tone scale, generated in Excel) is used for % (0-100%, red to green), SD (0-100%, white to blue), and N (0-3 replicates, white to orange).

Div. SETIV peptide library screened with PCAF HAT and HAT-BRD and NEC

SETIV library					HAT									HAT-BRD									NEC								
					parent			1Ac			2Ac			parent			1Ac			2Ac			parent			1Ac			2Ac		
#	well	hist	length	peptide PTM	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N
1	A1	H3	1-21		1	0	12	51	1	12	48	1	12	1	1	12	66	3	12	34	3	12	98	1	8	1	0	8	1	1	8
2	A2	H4	1-21		100	0	1	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0			0			0			0
3	A3	H4	1-21	K5Ac	0	0	0	0	0	0	0	0	0	0	0	12	0	0	12	100	0	12	0	0	5	0	0	5	100	0	5
4	A4	H4	1-21	K8Ac	0	0	0	0	0	0	0	0	0	0	0	12	0	0	12	100	0	12	0	0	1	0	0	1	100	0	1
5	A5	H4	1-21	K12Ac	4	3	12	91	3	12	4	2	12	74	0	12	25	0	12	0	0	12	99	1	9	1	1	9	0	0	9
6	A6	H4	1-21	K16Ac	1	1	12	96	3	12	4	2	12	72	1	12	27	1	12	0	1	12	100	0	8	0	0	8	0	0	8
7	A7	H4	1-21	K5AcK8Ac	90	14	12	5	8	12	5	10	12	97	1	12	3	0	12	0	0	12	99	1	8	1	1	8	0	0	8
8	A8	H4	1-21	K5AcK12Ac	2	3	12	93	6	12	5	4	12	78	2	12	22	2	12	0	0	12	99	1	9	1	1	9	0	0	9
9	A9	H4	1-21	K5AcK16Ac	1	1	12	96	1	12	3	1	12	75	1	12	25	1	12	0	0	12	98	1	8	2	1	8	0	0	8
10	A10	H4	1-21	K8AcK12Ac	96	2	12	4	2	12	0	0	12	98	0	12	2	0	12	0	0	12	99	2	9	1	2	9	0	0	9
11	A11	H4	1-21	K8AcK16Ac	96	6	12	2	4	12	1	5	12	99	0	12	1	0	12	0	0	12	99	1	9	1	1	9	0	0	9
12	A12	H4	1-21	K12AcK16Ac	40	6	11	60	6	11	0	0	11	85	1	12	15	1	12	0	0	12	99	1	8	1	1	8	0	0	8
13	B1	H4	1-21	K5AcK8AcK12Ac	94	3	12	3	2	12	3	4	12	98	0	12	2	0	12	0	0	12	98	2	9	1	1	9	1	3	9
14	B2	H4	1-21	K8AcK12AcK16Ac	94	7	12	3	4	12	3	4	12	99	1	12	1	1	12	0	0	12	99	1	9	1	1	9	0	0	9
15	B3	H4	1-21	K5AcK12AcK16Ac	31	6	12	66	7	12	3	4	12	86	0	12	14	0	12	0	0	12	100	0	9	0	0	9	0	0	9
16	B4	H4	1-21	K5AcK8AcK16Ac	89	8	12	5	5	12	6	6	12	99	0	12	1	0	12	0	0	12	97	5	8	1	1	8	2	6	8
17	B5	H4	1-21	K5AcK8AcK12AcK16Ac	96	3	12	3	1	12	1	2	12	99	1	12	1	1	12	0	0	12	99	1	9	1	1	9	0	0	9
18	B6	H4	9-29		72	42	9	28	42	9	0	0	9	100	0	3	0	0	3	0	0	3	77	23	2	23	23	2	0	0	2
19	B7	H4	9-29	K20Ac	95	7	12	4	6	12	1	2	12	95	0	12	5	0	12	0	0	12	94	7	8	6	7	8	0	0	8
20	B8	H4	9-29	K16Ac	95	7	12	4	7	12	1	2	12	97	2	12	3	2	12	0	0	12	99	2	8	1	2	8	0	0	8
21	B9	H4	9-29	K12AcK16Ac	94	7	12	5	5	12	1	3	12	96	1	12	4	1	12	0	0	12	94	7	9	6	7	9	0	0	9
22	B10	H4	9-29	K16AcK20Ac	95	6	12	4	5	12	1	2	12	95	1	12	5	1	12	0	0	12	98	3	8	2	3	8	0	0	8
23	B11	H4	9-29	K12AcK16AcK20Ac	93	5	12	5	4	12	1	2	12	97	2	12	3	2	12	0	0	12	97	4	8	3	4	8	0	0	8
24	B12	H2a	1-21		33	3	9	36	2	9	32	3	9	16	23	12	84	23	12	0	0	12	18	23	8	72	28	8	10	19	8

SETIV library					HAT									HAT-BRD									NEC								
					parent			1Ac			2Ac			parent			1Ac			2Ac			parent			1Ac			2Ac		
#	well	hist	length	peptide PTM	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N
25	C1	H2a	1-21	K5Ac	50	2	12	50	2	12	0	0	12	100	0	12	0	0	12	0	0	12	79	25	7	21	25	7	0	0	7
26	C2	H2a	1-21	K9Ac	65	14	12	35	14	12	0	0	12	100	1	12	0	1	12	0	0	12	88	12	8	12	12	8	0	0	8
27	C3	H2a	1-21	K13Ac	57	6	12	43	6	12	0	0	12	90	1	12	10	1	12	0	0	12	94	10	9	6	10	9	0	0	9
28	C4	H2a	1-21	K15Ac	48	2	12	52	2	12	0	0	12	100	0	12	0	0	12	0	0	12	82	24	8	18	24	8	0	0	8
29	C5	H2a	1-21	K5AcK9Ac	99	2	12	1	2	12	0	0	12	100	0	12	0	0	12	0	0	12	100	0	9	0	0	9	0	0	9
30	C6	H2a	1-21	K5AcK13Ac	31	11	12	69	11	12	0	1	12	94	1	12	6	1	12	0	0	12	100	0	9	0	0	9	0	0	9
31	C7	H2a	1-21	K5AcK15Ac	58	12	12	42	12	12	0	0	12	94	1	12	6	1	12	0	0	12	100	0	9	0	0	9	0	0	9
32	C8	H2a	1-21	K9AcK13Ac	100	0	12	0	0	12	0	0	12	100	0	12	0	0	12	0	0	12	100	0	9	0	0	9	0	0	9
33	C9	H2a	1-21	K9AcK15Ac	96	3	12	4	3	12	0	0	12	100	0	12	0	0	12	0	0	12	100	0	9	0	0	9	0	0	9
34	C10	H2a	1-21	K13AcK15Ac	53	9	12	47	9	12	0	0	12	95	1	12	5	1	12	0	0	12	100	0	8	0	0	8	0	0	8
35	C11	H2a	1-21	K5AcK9AcK13Ac	100	1	12	0	1	12	0	0	12	100	0	12	0	0	12	0	0	12	100	0	9	0	0	9	0	0	9
36	C12	H2a	1-21	K5AcK13AcK15Ac	67	11	11	33	11	11	0	0	11	96	1	12	4	1	12	0	0	12	100	0	9	0	0	9	0	0	9
37	D1	H2a	1-21	K9AcK13AcK15Ac	99	1	12	1	1	12	0	0	12	100	0	12	0	0	12	0	0	12	100	0	9	0	0	9	0	0	9
38	D2	H2a	1-21	K5AcK9AcK15Ac	98	4	12	2	4	12	0	0	12	100	0	12	0	0	12	0	0	12	99	1	8	1	1	8	0	0	8
39	D3	H2a	1-21	K5AcK9AcK13AcK15Ac	99	2	12	1	2	12	0	0	12	100	0	12	0	0	12	0	0	12	100	1	8	0	1	8	0	0	8
40	D4	H2b	1-21		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			0			0			0
41	D5	H2b	1-21	K5Ac	88	8	12	12	8	12	0	0	12	96	0	12	4	0	12	0	0	12	97	1	9	3	1	9	0	0	9
42	D6	H2b	1-21	K11Ac	90	10	11	10	10	11	0	0	11	96	0	12	4	0	12	0	0	12	98	3	9	2	3	9	0	0	9
43	D7	H2b	1-21	K12Ac	71	15	10	29	15	10	0	0	10	95	2	12	5	2	12	0	0	12	98	4	9	2	4	9	0	0	9
44	D8	H2b	1-21	K15Ac	83	6	12	5	3	12	12	6	12	90	0	8	4	0	8	6	0	8	91	4	9	3	3	9	7	5	9
45	D9	H2b	1-21	K5AcK11AcK16Ac	94	3	12	4	2	12	2	2	12	97	1	12	3	1	12	0	0	12	86	31	9	14	31	9	0	0	9
46	D10	H2b	1-21	K5AcK12Ac	71	5	12	27	5	12	2	3	12	86	0	12	14	0	12	0	0	12	85	6	9	15	5	9	1	1	9
47	D11	H2b	1-21	K5AcK15Ac	94	3	12	4	3	12	2	3	12	96	0	12	4	0	12	0	0	12	95	3	9	4	2	9	1	2	9
48	D12	H2b	1-21	K5AcK16Ac	95	4	12	4	4	12	0	1	12	96	0	8	4	0	8	0	0	8	98	2	9	2	1	9	0	1	9
49	E1	H2b	1-21	K11AcK12Ac	93	4	12	6	3	12	1	2	12	96	0	12	4	0	12	0	0	12	98	1	9	2	1	9	0	0	9
50	E2	H2b	1-21	K11AcK15Ac	93	6	12	4	2	12	4	6	12	97	0	12	3	0	12	0	0	12	97	2	9	2	2	9	1	2	9

SETIV library					HAT									HAT-BRD									NEC								
					parent			1Ac			2Ac			parent			1Ac			2Ac			parent			1Ac			2Ac		
#	well	hist	length	peptide PTM	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N
51	E3	H2b	1-21	K11AcK16Ac	92	6	12	5	3	12	3	5	12	97	0	12	3	0	12	0	0	12	98	2	8	2	2	8	0	0	8
52	E4	H2b	1-21	K5AcK11AcK12Ac	87	12	12	7	6	12	5	7	12	97	0	12	3	0	12	0	0	12	96	2	9	3	1	9	1	1	9
53	E5	H2b	1-21	K5AcK15AcK16Ac	84	4	12	14	3	12	2	2	12	93	2	12	7	2	12	0	0	12	91	4	8	9	5	8	0	1	8
54	E6	H2b	1-21	K11AcK12AcK15AcK16Ac	95	4	12	4	3	12	1	2	12	98	1	12	2	1	12	0	0	12	87	31	9	13	31	9	0	0	9
55	E7	H2b	1-21	K11AcK15AcK16Ac	88	11	12	11	11	12	1	1	12	92	1	12	8	1	12	0	0	12	92	2	9	8	2	9	0	0	9
56	E8	H2b	1-21	K12AcK15AcK16Ac	96	4	12	3	3	12	1	2	12	99	1	12	1	1	12	0	0	12	99	2	5	1	2	5	0	0	5
57	E9	H2b	1-21	K11AcK12AcK15Ac	96	2	12	4	1	12	1	2	12	98	1	12	2	1	12	0	0	12	97	2	9	3	2	9	0	1	9
58	E10	H2b	1-21	K11AcK12AcK16Ac	93	4	12	5	3	12	2	4	12	97	0	12	3	0	12	0	0	12	95	3	8	4	2	8	1	2	8
59	E11	H2b	1-21	K5AcK11AcK12AcK15AcK16Ac	97	3	12	3	3	12	0	0	12	100	0	12	0	0	12	0	0	12	97	3	8	3	3	8	0	0	8
60	E12	H2b	13-33		100	0	1	0	0	1	0	0	1	0	0	0	0	0	0	0	0			0			0			0	
61	F1	H2b	13-33	K20Ac	30	41	8	0	0	8	70	41	8	90	5	12	9	2	12	1	3	12	99	3	8	1	3	8	0	0	8
62	F2	H2b	13-33	K23Ac	63	34	11	5	4	11	33	37	11	93	1	12	7	1	12	0	0	12	100	0	8	0	0	8	0	0	8
63	F3	H2b	13-33	K24Ac	47	31	9	0	0	9	53	31	9	92	2	12	8	2	12	0	0	12	98	3	8	2	3	8	0	0	8
64	F4	H2b	13-33	K20AcK23AcK24Ac	32	15	12	38	8	12	30	12	12	86	7	12	14	7	12	0	0	12	67	30	8	31	29	8	3	8	8
65	F5	H3	1-21		100	0	7	0	0	7	0	0	7	100	0	11	0	0	11	0	0	11	100	0	7	0	0	7	0	0	7
66	F6	H3	1-21	K4Ac	1	2	12	26	15	12	73	16	12	3	3	12	79	3	12	18	1	12	100	0	9	0	0	9	0	0	9
67	F7	H3	1-21	K9Ac	1	2	12	92	9	12	7	9	12	3	3	12	97	3	12	0	0	12	100	0	8	0	0	8	0	0	8
68	F8	H3	1-21	K14Ac	4	8	10	82	15	10	14	16	10	75	1	12	25	1	12	0	0	12	100	0	7	0	0	7	0	0	7
69	F9	H3	1-21	K4AcK9Ac	0	0	12	67	24	12	33	24	12	1	2	12	91	4	12	8	4	12	81	20	9	1	3	9	18	21	9
70	F10	H3	1-21	K4AcK14Ac	22	11	12	41	16	12	37	21	12	72	6	12	16	3	12	13	8	12	75	28	9	0	0	9	25	28	9
71	F11	H3	1-21	K9AcK14Ac	63	17	12	4	6	12	33	22	12	91	6	12	0	0	12	9	6	12	72	28	8	0	0	8	28	28	8
72	F12	H3	1-21	K4AcK9AcK14Ac	76	28	12	14	15	12	10	13	12	79	9	12	11	5	12	10	5	12	81	21	9	9	11	9	9	11	9
73	G1	H3	11-31		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			0			0			0	
74	G2	H3	11-31	K14Ac	51	12	12	46	12	12	4	3	12	88	1	12	12	1	12	0	0	12	100	0	9	0	0	9	0	0	9
75	G3	H3	11-31	K18Ac	12	3	12	66	3	12	22	4	12	78	1	12	22	1	12	0	0	12	99	1	9	1	1	9	0	0	9
76	G4	H3	11-31	K23Ac	25	4	12	68	3	12	7	1	12	85	1	12	15	1	12	0	0	12	100	0	9	0	0	9	0	0	9

SETIV library					HAT									HAT-BRD									NEC								
					parent			1Ac			2Ac			parent			1Ac			2Ac			parent			1Ac			2Ac		
#	well	hist	length	peptide PTM	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N
77	G5	H3	11-31	K27Ac	7	5	12	61	8	12	32	8	12	79	1	12	21	1	12	0	0	12	100	1	8	0	1	8	0	0	8
78	G6	H3	11-31	K14AcK18Ac	69	12	12	27	9	12	4	8	12	91	1	12	9	1	12	0	0	12	100	0	9	0	0	9	0	0	9
79	G7	H3	11-31	K14AcK23Ac	88	2	12	12	2	12	0	1	12	96	0	12	4	0	12	0	0	12	99	1	9	1	1	9	0	0	9
80	G8	H3	11-31	K14AcK27Ac	74	3	12	26	3	12	0	0	12	91	0	12	9	0	12	0	0	12	98	3	8	2	3	8	0	0	8
81	G9	H3	11-31	K18AcK23Ac	55	6	12	43	6	12	1	2	12	89	1	12	11	1	12	0	0	12	100	0	9	0	0	9	0	0	9
82	G10	H3	11-31	K18AcK27Ac	40	5	12	53	4	12	7	3	12	84	1	12	16	1	12	0	0	12	99	2	9	1	2	9	0	0	9
83	G11	H3	11-31	K23AcK27Ac	64	5	12	35	5	12	0	1	12	89	0	12	11	0	12	0	0	12	100	0	8	0	0	8	0	0	8
84	G12	H3	11-31	K14AcK23AcK27Ac	98	3	12	2	3	12	0	0	12	100	0	12	0	0	12	0	0	12	100	0	9	0	0	9	0	0	9
85	H1	H3	11-31	K18AcK23AcK27Ac	74	7	12	25	7	12	0	1	12	91	2	12	9	2	12	0	0	12	100	0	9	0	0	9	0	0	9
86	H2	H3	11-31	K14AcK18AcK27Ac	80	4	12	18	2	12	2	3	12	93	1	12	7	1	12	0	0	12	100	0	8	0	0	8	0	0	8
87	H3	H3	11-31	K14AcK18AcK23AcK27Ac	98	3	12	1	2	12	1	2	12	100	0	12	0	0	12	0	0	12	99	2	9	1	2	9	0	0	9
88	H4	H3	23-43		14	10	12	65	4	12	22	7	12	74	2	12	26	1	12	0	1	12	100	0	8	0	0	8	0	0	8
89	H5	H3	23-43	K36Ac	57	7	12	43	7	12	0	0	12	91	1	12	9	1	12	0	0	12	100	0	9	0	0	9	0	0	9
90	H6	H3	23-43	K37Ac	4	3	12	64	5	12	31	8	12	83	1	8	17	1	8	0	0	8	100	1	8	0	1	8	0	0	8
91	H7	H3	23-43	K36AcK37Ac	67	4	12	33	4	12	0	0	12	94	0	12	6	0	12	0	0	12	89	31	9	11	31	9	0	0	9
92	H8	H3	23-43	K27Ac	53	8	12	47	8	12	0	0	12	90	0	12	10	0	12	0	0	12	100	1	8	0	1	8	0	0	8
93	H9	H3	23-43	K27AcK36Ac	97	3	12	3	3	12	0	0	12	100	1	12	0	1	12	0	0	12	100	0	9	0	0	9	0	0	9
94	H10	H3	23-43	K27AcK37Ac	54	7	12	46	7	12	0	0	12	92	0	12	8	0	12	0	0	12	100	1	8	0	1	8	0	0	8
95	H11	H3	23-43	K27AcK36AcK37Ac	97	0	12	3	0	12	0	0	12	100	0	12	0	0	12	0	0	12	100	0	9	0	0	9	0	0	9
96	H12				0			0			0			0			0			0			0			0			0		

Appendix Table D.4: Results of rapifleX MALDI-TOF-MS HTS of the SETIV library with PCAF HAT, PCAF HAT-BRD and a NEC.

Peptide number (#), well in library and assay (well), histone type (hist), peptide length, and peptide PTMs are listed for each peptide. The results of the screen are shown as an average of the observed relative percentages (%) of peak intensities for parent peptide, singly acetylated peptide (1Ac) and doubly acetylated peptide (2Ac) with associated standard deviation (SD) and number of replicates (N) for either PCAF HAT, PCAF HAT-BRD or a NEC. Colour coding (3 tone scale, generated in Excel) is used for % (0-100%, red to green), SD (0-100%, white to blue), and N (0-12 replicates, white to orange).

Dv. SETIV peptide library screened with SIRT1 and NEC

SETIV library			SIRT1																		NEC																	
			parent			-1Ac			-2Ac			-3Ac			-4Ac			-5Ac			parent			-1Ac			-2Ac			-3Ac			-4Ac			-5Ac		
well	peptide	PTM (Ac locations)	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N			
A1	H3 ₁₋₂₁	non biotinylated control	96	2	12	3	2	12	1	1	12	0	0	12	0	0	12	0	0	12	98	2	4	1	2	4	1	1	4	0	0	4	0	0	4	0	0	4
A2	H4 ₁₋₂₁		0	0	5	23	20	5	65	29	5	6	11	5	7	13	5	0	0	5	0	0	3	52	41	3	15	21	3	0	0	3	33	47	3	0	0	3
B6	H4 ₉₋₂₉		40	34	11	38	35	11	0	0	11	2	7	11	20	26	11	0	0	11	100	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0	0	1
B12	H2a ₁₋₂₁		0	0	4	0	0	4	0	0	4	0	0	4	0	0	4	100	0	4	0	0	2	0	0	2	0	0	2	0	0	2	19	19	2	81	19	2
D4	H2b ₁₋₂₁		0	0	3	0	0	3	0	0	3	100	0	3	0	0	3	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E12	H2b ₁₃₋₃₃		0	0	9	2	5	9	46	25	9	43	16	9	9	17	9	0	0	9	0	0	3	0	0	3	67	23	3	17	23	3	0	0	3	16	23	3
F5	H3 ₁₋₂₁		0	0	2	23	23	2	77	23	2	0	0	2	0	0	2	0	0	2	100	0	4	0	0	4	0	0	4	0	0	4	0	0	4	0	0	4
G1	H3 ₁₁₋₃₁		0	0	12	0	0	12	100	0	12	0	0	12	0	0	12	0	0	12	0	0	1	0	0	1	100	0	1	0	0	1	0	0	1	0	0	1
H4	H3 ₂₃₋₄₃		100	0	12	0	0	12	0	0	12	0	0	12	0	0	12	0	0	12	100	0	4	0	0	4	0	0	4	0	0	4	0	0	4	0	0	4
A3	H4 ₁₋₂₁	K5	0	0	7	29	45	7	0	0	7	0	0	7	43	49	7	29	45	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A4	H4 ₁₋₂₁	K8	0	0	9	0	0	9	0	0	9	11	31	9	55	37	9	34	34	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A5	H4 ₁₋₂₁	K12	1	1	12	97	4	12	1	1	12	0	0	12	1	2	12	1	2	12	99	1	4	1	1	4	0	0	4	0	0	4	0	0	4	0	0	4
A6	H4 ₁₋₂₁	K16	1	1	8	98	2	8	0	1	8	0	0	8	0	1	8	1	2	8	98	2	4	0	0	4	0	0	4	0	0	4	1	2	4	1	2	4
B7	H4 ₉₋₂₉	K20	4	4	12	95	7	12	2	4	12	0	0	12	0	0	12	0	0	12	100	0	4	0	0	4	0	0	4	0	0	4	0	0	4	0	0	4
B8	H4 ₉₋₂₉	K16	6	5	12	91	7	12	4	8	12	0	0	12	0	0	12	0	0	12	100	0	4	0	0	4	0	0	4	0	0	4	0	0	4	0	0	4
C1	H2a ₁₋₂₁	K5	100	0	10	0	0	10	0	0	10	0	0	10	0	0	10	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C2	H2a ₁₋₂₁	K9	20	21	12	80	21	12	0	0	12	0	0	12	0	0	12	0	1	12	100	0	4	0	0	4	0	0	4	0	0	4	0	0	4	0	0	4
C3	H2a ₁₋₂₁	K13	18	26	12	82	26	12	0	0	12	0	1	12	0	0	12	0	0	12	100	0	4	0	0	4	0	0	4	0	0	4	0	0	4	0	0	4
C4	H2a ₁₋₂₁	K15	100	0	12	0	0	12	0	0	12	0	0	12	0	0	12	0	0	12	100	0	4	0	0	4	0	0	4	0	0	4	0	0	4	0	0	4
D5	H2b ₁₋₂₁	K5	1	1	12	84	8	12	0	0	12	0	0	12	2	5	12	12	5	12	70	6	4	0	0	4	0	0	4	0	0	4	30	6	4	0	0	4
D6	H2b ₁₋₂₁	K11	0	1	12	74	17	12	0	0	12	0	0	12	8	10	12	18	9	12	72	2	4	0	0	4	0	0	4	0	0	4	28	2	4	0	0	4
D7	H2b ₁₋₂₁	K12	1	1	12	91	7	12	0	0	12	0	1	12	1	2	12	7	6	12	86	4	4	0	0	4	0	0	4	0	0	4	13	2	4	1	2	4
D8	H2b ₁₋₂₁	K15	0	1	10	88	13	10	0	0	10	0	0	10	2	6	10	9	9	10	94	4	4	0	0	4	1	1	4	0	0	4	6	5	4	0	0	4
F1	H2b ₁₃₋₃₃	K20	0	0	12	73	33	12	0	0	12	10	18	12	17	18	12	0	0	12	100	0	3	0	0	3	0	0	3	0	0	3	0	0	3	0	0	3

SETIV library			SIRT1																		NEC																	
			parent		-1Ac			-2Ac			-3Ac			-4Ac			-5Ac			parent		-1Ac			-2Ac			-3Ac			-4Ac			-5Ac				
well	peptide	PTM (Ac locations)	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N			
F2	H2b ₁₃₋₃₃	K23	1	2	12	85	29	12	1	1	12	14	30	12	0	0	12	0	0	12	100	0	1	0	0	1	0	0	1	0	0	1	0	0	1			
F3	H2b ₁₃₋₃₃	K24	1	2	12	88	12	12	1	4	12	8	8	12	2	6	12	0	0	12	92	15	4	0	0	4	0	0	4	8	15	4	0	0	4	0	0	4
F6	H3 ₁₋₂₁	K4	0	0	8	89	19	8	3	7	8	8	14	8	0	0	8	0	0	8	100	0	4	0	0	4	0	0	4	0	0	4	0	0	4	0	0	4
F7	H3 ₁₋₂₁	K9	0	1	12	94	9	12	4	6	12	2	5	12	0	0	12	0	0	12	91	9	4	6	10	4	0	0	4	3	5	4	0	0	4	0	0	4
F8	H3 ₁₋₂₁	K14	0	0	12	86	20	12	0	1	12	14	20	12	0	0	12	0	0	12	100	0	4	0	0	4	0	0	4	0	0	4	0	0	4	0	0	4
G2	H3 ₁₁₋₃₁	K14	2	6	12	52	13	12	0	0	12	46	14	12	0	0	12	0	0	12	89	19	4	0	0	4	0	0	4	11	19	4	0	0	4	0	0	4
G3	H3 ₁₁₋₃₁	K18	2	6	8	40	24	8	0	0	8	58	24	8	0	0	8	0	0	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
G4	H3 ₁₁₋₃₁	K23	1	2	12	55	25	12	0	0	12	44	26	12	0	0	12	0	0	12	97	5	4	0	0	4	0	0	4	3	5	4	0	0	4	0	0	4
G5	H3 ₁₁₋₃₁	K27	0	0	10	0	0	10	5	15	10	95	15	10	0	0	10	0	0	10	100	0	4	0	0	4	0	0	4	0	0	4	0	0	4	0	0	4
H5	H3 ₂₃₋₄₃	K36	2	1	12	98	1	12	0	0	12	0	0	12	0	0	12	0	0	12	99	1	4	1	1	4	0	0	4	0	0	4	0	0	4	0	0	4
H6	H3 ₂₃₋₄₃	K37	2	1	12	98	1	12	0	0	12	0	0	12	0	0	12	0	0	12	98	3	4	0	0	4	0	0	4	2	3	4	0	0	4	0	0	4
H8	H3 ₂₃₋₄₃	K27	1	1	12	99	1	12	0	0	12	0	0	12	0	0	12	0	0	12	98	2	4	1	1	4	1	1	4	0	0	4	0	0	4	0	0	4
H12	H4 ₁₋₂₁	K9 (control)	2	0	12	97	1	12	1	1	12	0	0	12	0	0	12	0	0	12	99	1	4	0	0	4	0	0	4	0	0	4	0	0	4	0	0	4
A7	H4 ₁₋₂₁	K5K8	0	0	10	0	1	10	99	2	10	0	0	10	0	0	10	1	2	10	100	0	4	0	0	4	0	0	4	0	0	4	0	0	4	0	0	4
A8	H4 ₁₋₂₁	K5K12	0	0	12	0	0	12	98	4	12	0	1	12	0	0	12	2	4	12	100	0	4	0	0	4	0	0	4	0	0	4	0	0	4	0	0	4
A9	H4 ₁₋₂₁	K5K16	0	0	12	1	1	12	99	1	12	0	0	12	0	0	12	0	1	12	95	5	4	0	0	4	0	0	4	0	0	4	1	2	4	3	3	4
A10	H4 ₁₋₂₁	K8K12	0	0	8	1	1	8	99	1	8	0	1	8	0	0	8	0	1	8	99	2	4	0	0	4	0	0	4	0	0	4	0	0	4	1	2	4
A11	H4 ₁₋₂₁	K8K16	0	0	12	1	1	12	96	6	12	0	0	12	0	0	12	2	6	12	100	0	4	0	0	4	0	0	4	0	0	4	0	0	4	0	0	4
A12	H4 ₁₋₂₁	K12K16	0	1	12	1	1	12	99	2	12	0	1	12	0	0	12	0	0	12	100	0	4	0	0	4	0	0	4	0	0	4	0	0	4	0	0	4
B9	H4 ₉₋₂₉	K12K16	0	0	8	3	3	8	95	4	8	2	3	8	0	0	8	0	0	8	100	0	4	0	0	4	0	0	4	0	0	4	0	0	4	0	0	4
B10	H4 ₉₋₂₉	K16K20	0	1	12	6	6	12	94	6	12	0	0	12	0	0	12	0	0	12	100	0	4	0	0	4	0	0	4	0	0	4	0	0	4	0	0	4
C5	H2a ₁₋₂₁	K5K9	0	0	8	8	6	8	91	6	8	0	0	8	1	1	8	0	0	8	89	8	4	11	8	4	0	0	4	0	0	4	0	0	4	0	0	4
C6	H2a ₁₋₂₁	K5K13	0	1	12	33	39	12	67	39	12	0	0	12	0	0	12	0	0	12	73	13	4	27	13	4	0	0	4	0	0	4	0	0	4	0	0	4
C7	H2a ₁₋₂₁	K5K15	0	0	12	38	37	12	62	37	12	0	0	12	0	0	12	0	0	12	86	8	4	14	8	4	0	0	4	0	0	4	0	0	4	0	0	4
C8	H2a ₁₋₂₁	K9K13	0	0	12	19	23	12	81	23	12	0	1	12	0	0	12	0	0	12	83	11	4	17	11	4	0	0	4	0	0	4	0	0	4	0	0	4
C9	H2a ₁₋₂₁	K9K15	0	0	8	15	10	8	85	10	8	0	0	8	0	0	8	0	0	8	61	6	4	0	0	4	0	0	4	0	0	4	0	0	4	39	6	4

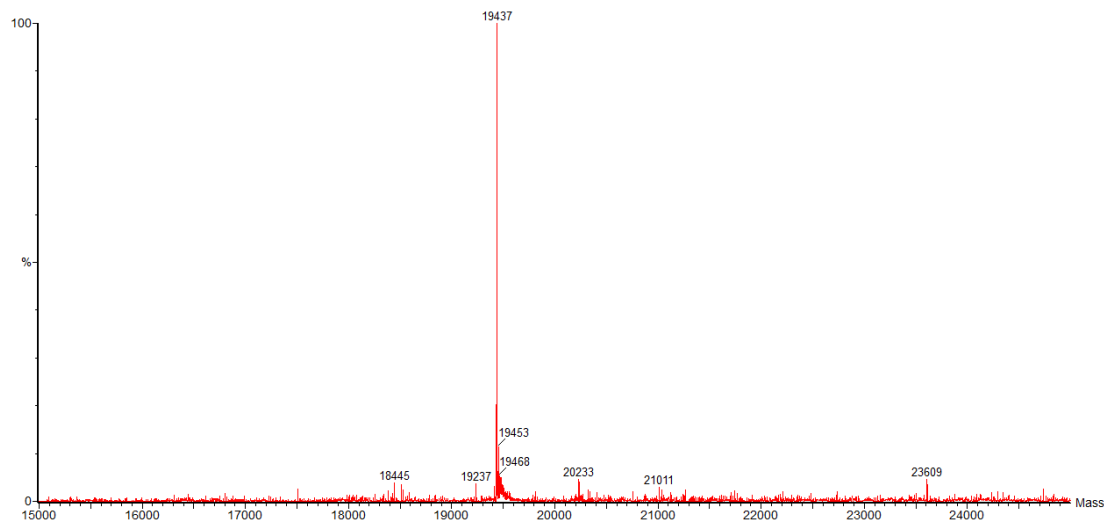
SETIV library			SIRT1																		NEC																	
			parent		-1Ac			-2Ac			-3Ac			-4Ac			-5Ac			parent		-1Ac			-2Ac			-3Ac			-4Ac			-5Ac				
well	peptide	PTM (Ac locations)	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N			
C10	H2a ₁₋₂₁	K13K15	0	1	12	23	25	12	76	25	12	0	0	12	0	0	12	0	0	12	88	17	3	12	17	3	0	0	3	0	0	3	0	0	3	0	0	3
D10	H2b ₁₋₂₁	K5K12	0	0	12	2	2	12	97	3	12	0	0	12	0	0	12	1	3	12	93	4	4	4	1	4	0	0	4	1	1	4	2	4	4	0	0	4
D11	H2b ₁₋₂₁	K5K15	0	0	12	1	1	12	99	2	12	0	1	12	0	0	12	1	2	12	95	8	4	0	0	4	0	0	4	0	0	4	5	8	4	0	0	4
D12	H2b ₁₋₂₁	K5K16	0	0	12	2	2	12	98	2	12	0	1	12	0	0	12	0	0	12	98	4	4	0	0	4	0	0	4	0	0	4	2	3	4	1	1	4
E1	H2b ₁₋₂₁	K11K12	0	0	12	3	1	12	97	1	12	0	0	12	0	1	12	0	0	12	88	1	4	0	0	4	0	0	4	0	0	4	12	1	4	0	0	4
E2	H2b ₁₋₂₁	K11K15	0	0	12	3	1	12	96	2	12	0	1	12	0	1	12	1	2	12	95	5	4	0	1	4	0	0	4	0	0	4	4	4	4	0	0	4
E3	H2b ₁₋₂₁	K11K16	0	0	12	1	1	12	97	5	12	0	1	12	0	0	12	1	5	12	92	4	4	0	0	4	0	0	4	0	0	4	8	4	4	0	0	4
F9	H3 ₁₋₂₁	K4K9	0	0	11	0	0	11	93	10	11	3	8	11	3	8	11	0	0	11	88	14	4	0	0	4	12	14	4	0	0	4	0	0	4	0	0	4
F10	H3 ₁₋₂₁	K4K14	0	0	8	0	0	8	98	3	8	1	2	8	1	2	8	0	0	8	100	0	4	0	0	4	0	0	4	0	0	4	0	0	4	0	0	4
F11	H3 ₁₋₂₁	K9K14	0	0	12	0	0	12	93	9	12	0	0	12	5	7	12	2	5	12	100	0	4	0	0	4	0	0	4	0	0	4	0	0	4	0	0	4
G6	H3 ₁₁₋₃₁	K14K18	0	0	12	0	0	12	0	0	12	4	13	12	92	17	12	4	12	12	80	13	4	16	10	4	0	0	4	0	0	4	3	5	4	0	0	4
G7	H3 ₁₁₋₃₁	K14K23	0	0	8	0	0	8	29	32	8	0	0	8	71	32	8	0	0	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
G8	H3 ₁₁₋₃₁	K14K27	0	0	10	0	0	10	0	0	10	5	16	10	75	25	10	20	25	10	95	5	4	4	5	4	0	0	4	1	1	4	0	0	4	0	0	4
G9	H3 ₁₁₋₃₁	K18K23	0	0	11	0	0	11	7	12	11	0	0	11	68	29	11	25	24	11	92	9	4	8	9	4	0	0	4	0	0	4	0	0	4	0	0	4
G10	H3 ₁₁₋₃₁	K18K27	0	0	12	0	0	12	9	13	12	0	0	12	61	23	12	30	22	12	100	0	4	0	0	4	0	0	4	0	0	4	0	0	4	0	0	4
G11	H3 ₁₁₋₃₁	K23K27	0	0	7	0	0	7	3	7	7	0	0	7	71	25	7	26	23	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
H7	H3 ₂₃₋₄₃	K36K37	0	0	12	6	1	12	94	1	12	1	1	12	0	0	12	0	0	12	100	0	4	0	0	4	0	0	4	0	0	4	0	0	4	0	0	4
H9	H3 ₂₃₋₄₃	K27K36	0	0	12	2	0	12	98	1	12	1	1	12	0	0	12	0	0	12	99	1	4	1	1	4	0	0	4	0	0	4	0	0	4	0	0	4
H10	H3 ₂₃₋₄₃	K27K37	0	0	12	7	1	12	93	2	12	0	1	12	0	0	12	0	0	12	100	0	4	0	0	4	0	0	4	0	0	4	0	0	4	0	0	4
B1	H4 ₁₋₂₁	K5K8K12	0	1	12	0	0	12	1	1	12	98	3	12	0	1	12	1	2	12	99	2	4	0	0	4	0	0	4	0	0	4	0	0	4	1	2	4
B2	H4 ₁₋₂₁	K8K12K16	0	0	12	0	0	12	1	1	12	98	2	12	1	1	12	0	0	12	100	0	4	0	0	4	0	0	4	0	0	4	0	0	4	0	0	4
B3	H4 ₁₋₂₁	K5K12K16	0	1	12	0	0	12	1	1	12	98	2	12	1	1	12	0	0	12	100	0	4	0	0	4	0	0	4	0	0	4	0	0	4	0	0	4
B4	H4 ₁₋₂₁	K5K8K16	0	1	12	0	0	12	1	1	12	98	2	12	0	1	12	0	0	12	100	0	4	0	0	4	0	0	4	0	0	4	0	0	4	0	0	4
B11	H4 ₉₋₂₉	K12K16K20	0	0	12	0	0	12	3	4	12	97	4	12	0	0	12	0	0	12	97	5	4	3	5	4	0	0	4	0	0	4	0	0	4	0	0	4
C11	H2a ₁₋₂₁	K5K9K13	0	0	12	0	0	12	18	21	12	82	21	12	0	0	12	0	0	12	100	0	4	0	0	4	0	0	4	0	0	4	0	0	4	0	0	4
C12	H2a ₁₋₂₁	K5K13K15	0	0	12	0	0	12	7	4	12	93	4	12	0	0	12	0	1	12	94	10	4	0	0	4	6	10	4	0	0	4	0	0	4	0	0	4

SETIV library			SIRT1																		NEC																	
			parent			-1Ac			-2Ac			-3Ac			-4Ac			-5Ac			parent			-1Ac			-2Ac			-3Ac			-4Ac			-5Ac		
well	peptide	PTM (Ac locations)	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N			
D1	H2a ₁₋₂₁	K9K13K15	0	0	12	0	0	12	27	34	12	73	34	12	0	0	12	0	0	12	96	3	3	4	3	3	0	0	3	0	0	3	0	0	3	0	0	3
D2	H2a ₁₋₂₁	K5K9K15	0	0	12	0	0	12	12	9	12	88	9	12	0	0	12	0	0	12	100	0	4	0	0	4	0	0	4	0	0	4	0	0	4	0	0	4
D9	H2b ₁₋₂₁	K5K11K16	0	1	11	0	0	11	2	1	11	98	2	11	0	0	11	0	0	11	92	2	4	6	1	4	2	2	4	0	0	4	0	0	4	0	0	4
E4	H2b ₁₋₂₁	K5K11K12	0	0	11	0	0	11	1	2	11	99	2	11	0	0	11	0	0	11	100	0	4	0	0	4	0	0	4	0	0	4	0	0	4	0	0	4
E5	H2b ₁₋₂₁	K5K15K16	0	0	12	0	0	12	2	1	12	98	1	12	0	0	12	0	0	12	99	1	4	0	0	4	0	0	4	0	1	4	0	0	4	0	1	4
E7	H2b ₁₋₂₁	K11K15K16	0	0	12	0	0	12	1	1	12	99	1	12	0	0	12	0	0	12	96	4	4	0	1	4	0	0	4	0	0	4	2	3	4	1	1	4
E8	H2b ₁₋₂₁	K12K15K16	0	0	12	0	0	12	1	2	12	98	3	12	0	0	12	1	2	12	93	6	4	0	0	4	0	0	4	0	0	4	6	6	4	0	1	4
E9	H2b ₁₋₂₁	K11K12K15	1	1	12	0	0	12	1	1	12	98	2	12	0	0	12	0	0	12	100	0	4	0	0	4	0	0	4	0	0	4	0	0	4	0	0	4
E10	H2b ₁₋₂₁	K11K12K16	0	0	12	0	0	12	1	2	12	99	2	12	0	0	12	0	0	12	94	3	4	0	0	4	0	0	4	0	0	4	6	3	4	0	0	4
F4	H2b ₁₃₋₃₃	K20K23K24	0	0	12	0	0	12	1	2	12	90	15	12	4	12	12	5	12	12	100	0	4	0	0	4	0	0	4	0	0	4	0	0	4	0	0	4
F12	H3 ₁₋₂₁	K4K9K14	0	0	12	0	0	12	0	0	12	99	3	12	0	0	12	1	3	12	96	5	4	0	0	4	0	0	4	0	0	4	2	3	4	2	4	4
G12	H3 ₁₁₋₃₁	K14K23K27	0	0	2	0	0	2	0	0	2	0	0	2	0	0	2	100	0	2	98	3	4	2	3	4	0	0	4	0	0	4	0	0	4	0	0	4
H1	H3 ₁₁₋₃₁	K18K23K27	0	0	9	0	0	9	0	0	9	52	22	9	0	0	9	48	22	9	98	4	4	2	4	4	0	0	4	0	0	4	0	0	4	0	0	4
H2	H3 ₁₁₋₃₁	K14K18K27	0	0	12	0	0	12	0	0	12	53	14	12	0	0	12	47	14	12	96	4	4	4	4	4	0	0	4	0	0	4	0	0	4	0	0	4
H11	H3 ₂₃₋₄₃	K27K36K37	0	0	12	0	0	12	8	2	12	92	2	12	1	1	12	0	0	12	100	0	4	0	0	4	0	0	4	0	0	4	0	0	4	0	0	4
B5	H4 ₁₋₂₁	K5K8K12K16	0	0	11	0	1	11	0	0	11	1	1	11	98	2	11	0	0	11	99	1	4	1	1	4	0	0	4	0	0	4	0	0	4	0	0	4
D3	H2a ₁₋₂₁	K5K9K13K15	0	0	12	0	0	12	0	0	12	39	38	12	60	38	12	0	1	12	95	3	4	2	2	4	1	2	4	2	2	4	0	0	4	0	0	4
E6	H2b ₁₋₂₁	K11K12K15K16	4	12	12	1	2	12	0	0	12	3	2	12	93	11	12	0	0	12	92	3	4	6	1	4	1	1	4	0	0	4	1	2	4	0	0	4
H3	H3 ₁₁₋₃₁	K14K18K23K27	0	0	9	0	0	9	0	0	9	0	0	9	100	0	9	0	0	9	97	5	4	3	5	4	0	0	4	0	0	4	0	0	4	0	0	4
E11	H2b ₁₋₂₁	K5K11K12K15K16	0	1	12	3	5	12	0	0	12	0	0	12	2	2	12	95	5	12	100	0	4	0	0	4	0	0	4	0	0	4	0	0	4	0	0	4

Appendix Table D.5: Results of rapifleX MALDI-TOF-MS HTS of the SETIV library with SIRT1 and a NEC.

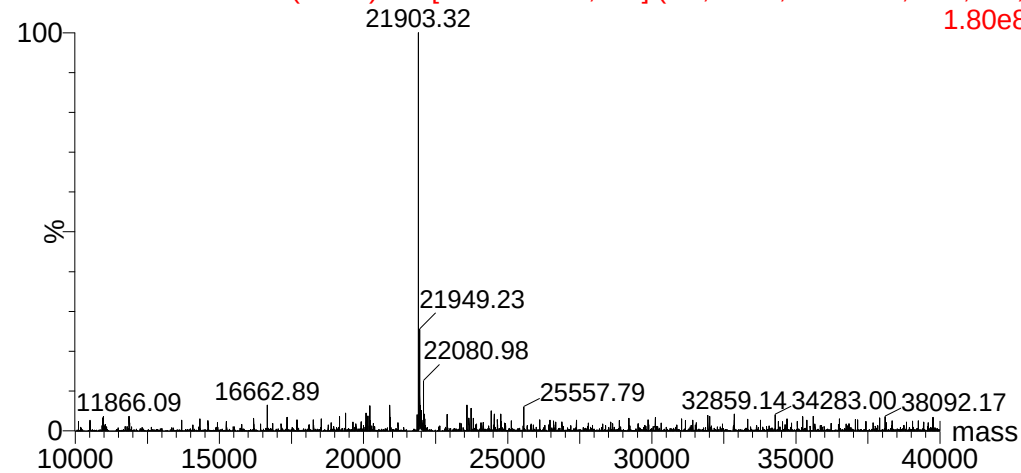
Well ID in library and assay (well), histone type and length (peptide), and site of peptide KAc PTMs are listed for each peptide. Peptides are ordered according to the number of KAc PTMs they carry, rather than by well number, matching the graphical representation of **FIG XXX**. The results of the screen are shown as an average of the observed relative percentages (%) of peak intensities for parent peptide, and peptides with 1 to 5 deacetylation events (-1Ac to -5Ac) with associated standard deviation (SD) and number of replicates (N) for either PCAF HAT, PCAF HAT-BRD or a NEC. Colour coding (3 tone scale, generated in Excel) is used for % (0-100%, red to green), SD (0-100%, white to blue), and N (0-12 replicates, white to orange).

E. Mass spectra for purified proteins

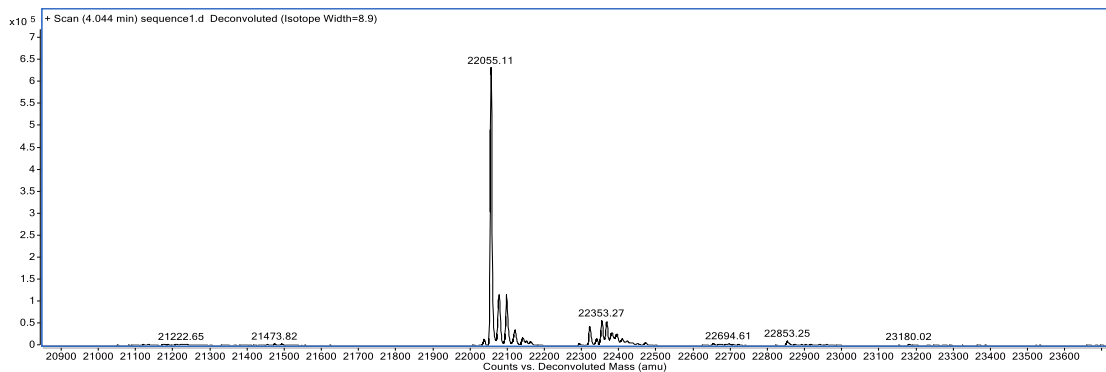


Appendix Figure E.1: Mass spectrum obtained for PCAF HAT protein.
Spectrum acquired by Dr. Anthony Tumber (Schofield group).

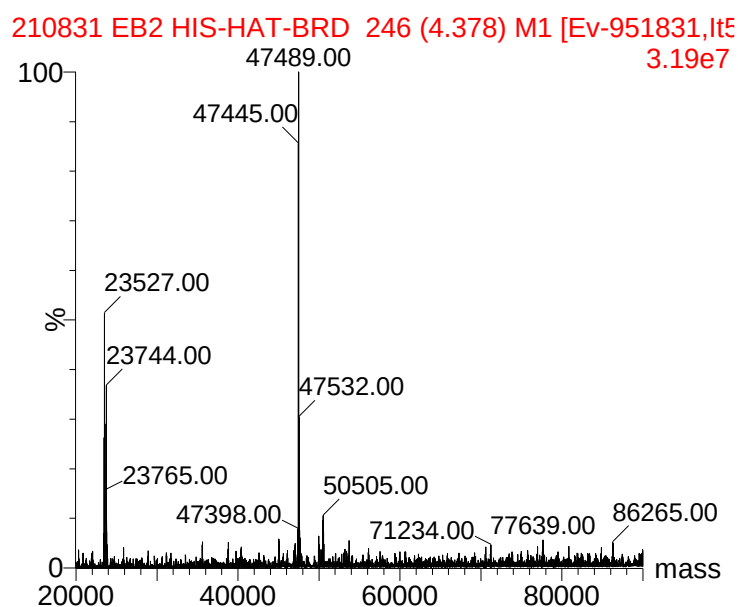
210831 HIS-HAT 246 (4.378) M1 [Ev-1255841,lt64] (Gs,0.750,566:2000,1.00,L33,F 1.80e8



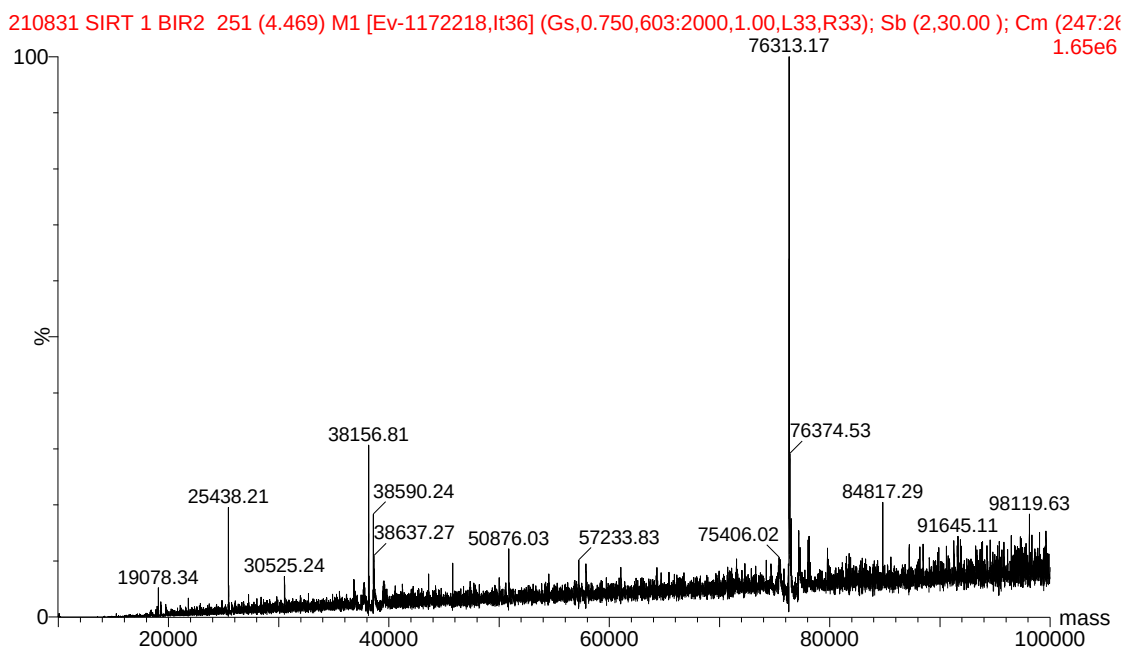
Appendix Figure E.2: Mass spectrum obtained for PCAF His-HAT protein.
Spectrum acquired by Dr. Joanna Bonnici (Kawamura group).



Appendix Figure E.3: Mass spectrum obtained for PCAF HAT-bio protein.
Spectrum acquired by Dr. Anthony Tumber (Schofield group).

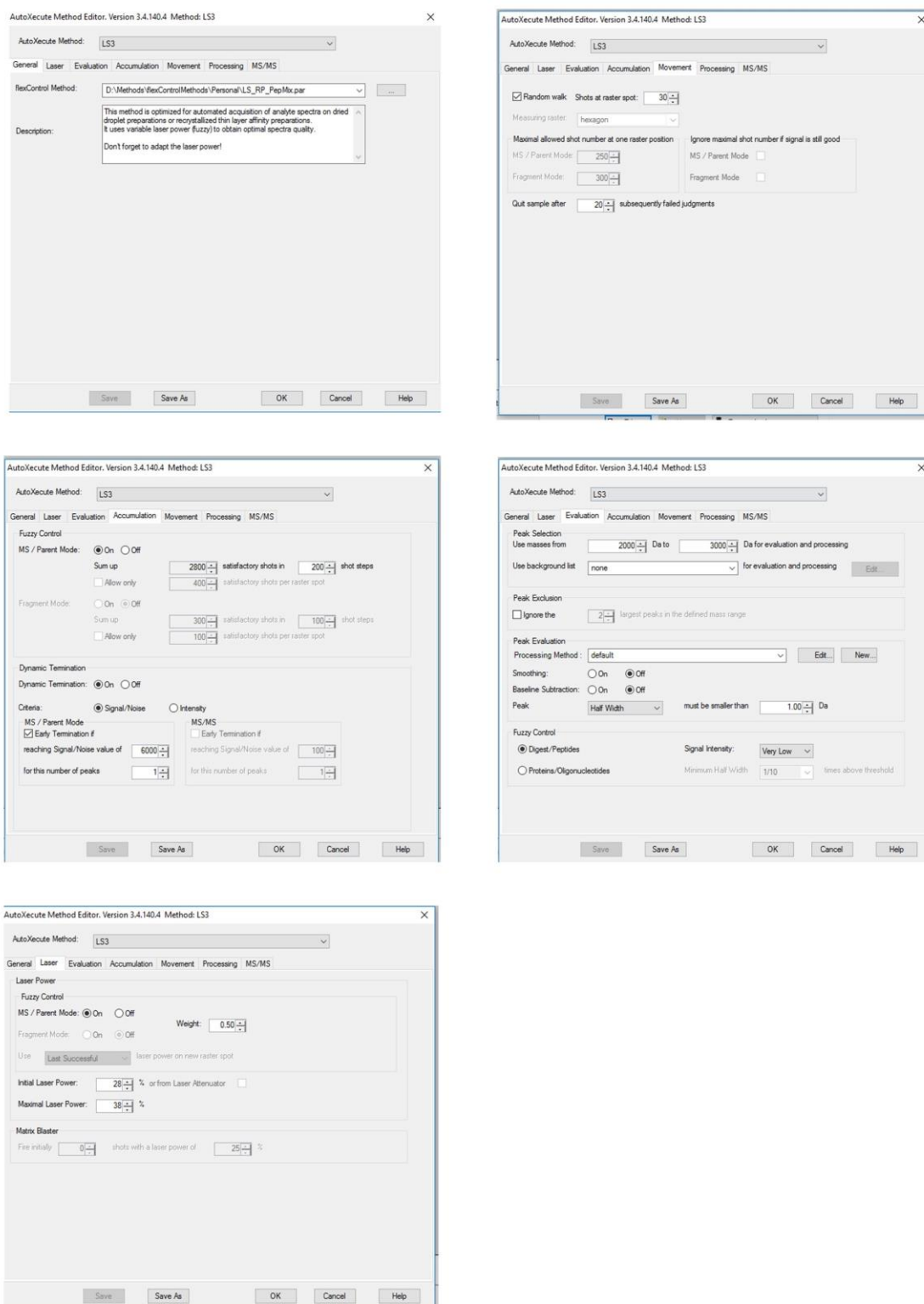


Appendix Figure E.4: Mass spectrum obtained for PCAF HAT-BRD protein.
Spectrum acquired by Dr. Joanna Bonnici (Kawamura group).



Appendix Figure E.5: Mass spectrum obtained for SIRT1 protein.
Spectrum acquired by Dr. Joanna Bonnici (Kawamura group).

F. Parameters for auto-acquisition of MALDI-TOF-MS spectra on microflex



Appendix Figure F.1: Screenshots of parameters used for auto-acquisition of MALDI-TOF-MS spectra on the Bruker microflex.

References

- 1 C. D. Allis and T. Jenuwein, The molecular hallmarks of epigenetic control, *Nat. Rev. Genet.*, 2016, **17**, 487–500.
- 2 L. Y. Zhao, J. Song, Y. Liu, C. X. Song and C. Yi, Mapping the epigenetic modifications of DNA and RNA, *Protein Cell*, 2020, **11**, 792–808.
- 3 G. P. Delcuve, M. Rastegar and J. R. Davie, Epigenetic control, *J. Cell. Physiol.*, 2009, **219**, 243–250.
- 4 Z. H. Harvey, Y. Chen and D. F. Jarosz, Protein-Based Inheritance: Epigenetics beyond the Chromosome, *Mol. Cell*, 2018, **69**, 195–202.
- 5 A. Portela and M. Esteller, Epigenetic modifications and human disease, *Nat. Biotechnol.*, 2010, **28**, 1057–1068.
- 6 S. B. Baylin and P. A. Jones, A decade of exploring the cancer epigenome-biological and translational implications, *Nat. Rev. Cancer*, 2011, **11**, 726–734.
- 7 I. Houston, C. J. Peter, A. Mitchell, J. Straubhaar, E. Rogaev and S. Akbarian, Epigenetics in the human brain, *Neuropsychopharmacology*, 2013, **38**, 183–197.
- 8 H. D. Kondilis-Mangum and P. A. Wade, Epigenetics and the adaptive immune response, *Mol. Aspects Med.*, 2013, **34**, 813–825.
- 9 B. M. Turner, Epigenetic responses to environmental change and their evolutionary implications, *Philos. Trans. R. Soc. B Biol. Sci.*, 2009, **364**, 3403–3418.
- 10 W. G. Kaelin and S. L. McKnight, Influence of metabolism on epigenetics and disease, *Cell*, 2013, **153**, 56–69.
- 11 V. N. Budhavarapu, M. Chavez and J. K. Tyler, How is epigenetic information maintained through DNA replication?, *Epigenetics and Chromatin*, 2013, **6**, 32.
- 12 K. Skvortsova, N. Iovino and O. Bogdanović, Functions and mechanisms of epigenetic inheritance in animals, *Nat. Rev. Mol. Cell Biol.*, 2018, **19**, 774–790.
- 13 E. Heard and R. A. Martienssen, Transgenerational epigenetic inheritance: Myths and mechanisms, *Cell*, 2014, **157**, 95–109.
- 14 D. M. PHILLIPS, The presence of acetyl groups of histones., *Biochem. J.*, 1963, **87**, 258–263.
- 15 R. A. Laskey and W. C. Earnshaw, Nucleosome assembly, *Nature*, 1980, **286**, 763–767.
- 16 F. H. Andrews, B. D. Strahl and T. G. Kutateladze, Insights into newly discovered marks and readers of epigenetic information, *Nat. Chem. Biol.*, 2016, **12**, 662–668.
- 17 H. Huang, B. R. Sabari, B. A. Garcia, C. David Allis and Y. Zhao, SnapShot: Histone Modifications, *Cell*, 2014, **159**, 458.e1.
- 18 G. E. Zentner and S. Henikoff, Regulation of nucleosome dynamics by histone modifications, *Nat. Struct. Mol. Biol.*, 2013, **20**, 259–266.
- 19 T. Jenuwein and C. D. Allis, Translating the histone code, *Science (80-.)*, 2001, **293**, 1074–1080.
- 20 T. Narita, B. T. Weinert and C. Choudhary, Functions and mechanisms of non-histone protein acetylation, *Nat. Rev. Mol. Cell Biol.*, 2019, **20**, 156–174.
- 21 D. E. Sterner and S. L. Berger, Acetylation of Histones and Transcription-Related Factors, *Microbiol. Mol. Biol. Rev.*, 2000, **64**, 435–459.
- 22 G. Duan and D. Walther, The Roles of Post-translational Modifications in the Context of Protein Interaction Networks, *PLoS Comput. Biol.*, 2015, **11**, 1004049.
- 23 V. Marx, Epigenetics: Reading the second genomic code, *Nature*, 2012, **491**, 143–147.
- 24 C. H. Arrowsmith, C. Bountra, P. V. Fish, K. Lee and M. Schapira, Epigenetic protein families: a new frontier for drug discovery, *Nat. Rev. Drug Discov.*, 2012, **11**, 384–400.
- 25 C. C. Wong, Y. Qian and J. Yu, Interplay between epigenetics and metabolism in

- oncogenesis: Mechanisms and therapeutic approaches, *Oncogene*, 2017, **36**, 3359–3374.
- 26 S. Hino, K. Nagaoka and M. Nakao, Metabolism-epigenome crosstalk in physiology and diseases, *J. Hum. Genet.*, 2013, **58**, 410–415.
- 27 A. Nieborak and R. Schneider, Metabolic intermediates – Cellular messengers talking to chromatin modifiers, *Mol. Metab.*, 2018, **14**, 39–52.
- 28 A. Prahallad and R. Bernards, Opportunities and challenges provided by crosstalk between signalling pathways in cancer, *Oncogene*, 2016, **35**, 1073–1079.
- 29 F. Liu, L. Wang, F. Perna and S. D. Nimer, Beyond transcription factors: How oncogenic signalling reshapes the epigenetic landscape, *Nat. Rev. Cancer*, 2016, **16**, 359–372.
- 30 S. E. Castel and R. A. Martienssen, RNA interference in the nucleus: Roles for small RNAs in transcription, epigenetics and beyond, *Nat. Rev. Genet.*, 2013, **14**, 100–112.
- 31 J. Zhang, Y. Z. Zhang, J. Jiang and C. G. Duan, The Crosstalk Between Epigenetic Mechanisms and Alternative RNA Processing Regulation, *Front. Genet.*, 2020, **11**, 998.
- 32 J. Du, L. M. Johnson, S. E. Jacobsen and D. J. Patel, DNA methylation pathways and their crosstalk with histone methylation, *Nat. Rev. Mol. Cell Biol.*, 2015, **16**, 519–532.
- 33 H. Cedar and Y. Bergman, Linking DNA methylation and histone modification: Patterns and paradigms, *Nat. Rev. Genet.*, 2009, **10**, 295–304.
- 34 C. Z. Wang and B. Zhu, You are never alone: crosstalk among epigenetic players, *Sci. Bull.*, 2015, **60**, 899–904.
- 35 J. S. Lee, E. Smith and A. Shilatifard, The Language of Histone Crosstalk, *Cell*, 2010, **142**, 682–685.
- 36 B. D. Strahl and C. D. Allis, The language of covalent histone modifications, *Nature*, 2000, **403**, 41–45.
- 37 X. J. Yang and E. Seto, Lysine Acetylation: Codified Crosstalk with Other Posttranslational Modifications, *Mol. Cell*, 2008, **31**, 449–461.
- 38 Y. Atlasi and H. G. Stunnenberg, The interplay of epigenetic marks during stem cell differentiation and development, *Nat. Rev. Genet.*, 2017, **18**, 643–658.
- 39 E. M. Jobe, A. L. McQuate and X. Zhao, Crosstalk among epigenetic pathways regulates neurogenesis, *Front. Neurosci.*, 2012, **6**, 59.
- 40 P. Cheung, K. G. Tanner, W. L. Cheung, P. Sassone-Corsi, J. M. Denu and C. D. Allis, Synergistic coupling of histone H3 phosphorylation and acetylation in response to epidermal growth factor stimulation, *Mol. Cell*, 2000, **5**, 905–915.
- 41 W. S. Lo, R. C. Trievel, J. R. Rojas, L. Duggan, J. Y. Hsu, C. D. Allis, R. Marmorstein and S. L. Berger, Phosphorylation of serine 10 in histone H3 is functionally linked in vitro and in vivo to Gcn5-mediated acetylation at lysine 14, *Mol. Cell*, 2000, **5**, 917–926.
- 42 C. J. Fry, M. A. Shogren-Knaak and C. L. Peterson, in *Cold Spring Harbor Symposia on Quantitative Biology*, Cold Spring Harb Symp Quant Biol, 2004, vol. 69, pp. 219–226.
- 43 M. A. Shogren-Knaak, C. J. Fry and C. L. Peterson, A native peptide ligation strategy for deciphering nucleosomal histone modifications, *J. Biol. Chem.*, 2003, **278**, 15744–15748.
- 44 Y. Zhang, K. Griffin, N. Mondal and J. D. Parvin, Phosphorylation of histone H2A inhibits transcription chromatin templates, *J. Biol. Chem.*, 2004, **279**, 21866–21872.
- 45 S. Liokatis, A. Stützer, S. J. Elsässer, F. X. Theillet, R. Klingberg, B. Van Rossum, D. Schwarzer, C. D. Allis, W. Fischle and P. Selenko, Phosphorylation of histone H3 Ser10 establishes a hierarchy for subsequent intramolecular modification events, *Nat. Struct. Mol. Biol.*, 2012, **19**, 819–823.
- 46 O. Binda, On your histone mark, SET, methylate!, *Epigenetics*, 2013, **8**, 457–463.
- 47 J. R. Kornacki, A. D. Stuparu and M. Mrksich, Acetyltransferase p300/CBP associated factor (PCAF) regulates crosstalk-dependent acetylation of histone H3 by distal site recognition, *ACS Chem. Biol.*, 2015, **10**, 157–164.
- 48 S. H. Ahn, R. L. Diaz, M. Grunstein and C. D. Allis, Histone H2B Deacetylation at Lysine 11

- Is Required for Yeast Apoptosis Induced by Phosphorylation of H2B at Serine 10, *Mol. Cell*, 2006, **24**, 211–220.
- 49 M. Rolando, S. Sanulli, C. Rusniok, L. Gomez-Valero, C. Bertholet, T. Sahr, R. Margueron and C. Buchrieser, Legionella pneumophila effector RomA uniquely modifies host chromatin to repress gene expression and promote intracellular bacterial replication, *Cell Host Microbe*, 2013, **13**, 395–405.
- 50 A. N. Poux and R. Marmorstein, Molecular Basis for Gcn5/PCAF Histone Acetyltransferase Selectivity for Histone and Nonhistone Substrates, *Biochemistry*, 2003, **42**, 14366–14374.
- 51 M. Shimada, H. Niida, D. H. Zineldeen, H. Tagami, M. Tanaka, H. Saito and M. Nakanishi, Chk1 Is a Histone H3 Threonine 11 Kinase that Regulates DNA Damage-Induced Transcriptional Repression, *Cell*, 2008, **132**, 221–232.
- 52 P. N. I. Lau and P. Cheung, Histone code pathway involving H3 S28 phosphorylation and K27 acetylation activates transcription and antagonizes polycomb silencing, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 2801–2806.
- 53 K. K. Lee and J. L. Workman, Histone acetyltransferase complexes: One size doesn't fit all, *Nat. Rev. Mol. Cell Biol.*, 2007, **8**, 284–295.
- 54 F. Ye, J. Huang, H. Wang, C. Luo and K. Zhao, Targeting epigenetic machinery: Emerging novel allosteric inhibitors, *Pharmacol. Ther.*, 2019, **204**, 107406.
- 55 P. F. Cartron, M. Cheray and L. Bretaudeau, Epigenetic protein complexes: The adequate candidates for the use of a new generation of epidrugs in personalized and precision medicine in cancer, *Epigenomics*, 2020, **12**, 171–177.
- 56 I. Ali, R. J. Conrad, E. Verdin and M. Ott, Lysine Acetylation Goes Global: From Epigenetics to Metabolism and Therapeutics, *Chem. Rev.*, 2018, **118**, 1216–1252.
- 57 H. Aksnes, A. Drazic, M. Marie and T. Arnesen, First Things First: Vital Protein Marks by N-Terminal Acetyltransferases, *Trends Biochem. Sci.*, 2016, **41**, 746–760.
- 58 A. J. Lindahl, A. J. Lawton, J. Baeza, J. A. Dowell and J. M. Denu, in *Methods in Molecular Biology*, Humana Press Inc., 2019, vol. 1983, pp. 79–106.
- 59 J. Baeza, J. Baeza, A. J. Lawton, A. J. Lawton, J. Fan, J. Fan, M. J. Smallegan, M. J. Smallegan, I. Lienert, T. Gandhi, O. M. Bernhardt, L. Reiter, J. M. Denu and J. M. Denu, Revealing Dynamic Protein Acetylation across Subcellular Compartments, *J. Proteome Res.*, 2020, **19**, 2404–2418.
- 60 P. Mews, G. Donahue, A. M. Drake, V. Luczak, T. Abel and S. L. Berger, Acetyl-CoA synthetase regulates histone acetylation and hippocampal memory, *Nature*, 2017, **546**, 381–386.
- 61 K. Struhl, Histone acetylation and transcriptional regulatory mechanisms, *Genes Dev.*, 1998, **12**, 599–606.
- 62 B. M. Turner, Histone acetylation and an epigenetic code, *BioEssays*, 2000, **22**, 836–845.
- 63 C. Choudhary, B. T. Weinert, Y. Nishida, E. Verdin and M. Mann, The growing landscape of lysine acetylation links metabolism and cell signalling, *Nat. Rev. Mol. Cell Biol.*, 2014, **15**, 536–550.
- 64 S. C. Kim, R. Sprung, Y. Chen, Y. Xu, H. Ball, J. Pei, T. Cheng, Y. Kho, H. Xiao, L. Xiao, N. V. Grishin, M. White, X. J. Yang and Y. Zhao, Substrate and Functional Diversity of Lysine Acetylation Revealed by a Proteomics Survey, *Mol. Cell*, 2006, **23**, 607–618.
- 65 F. Gnad, F. Forner, D. F. Zielinska, E. Birney, J. Gunawardena and M. Mann, Evolutionary constraints of phosphorylation in eukaryotes, prokaryotes, and mitochondria, *Mol. Cell. Proteomics*, 2010, **9**, 2642–2653.
- 66 Y. Li, H. Wen, Y. Xi, K. Tanaka, H. Wang, D. Peng, Y. Ren, Q. Jin, S. Y. R. Dent, W. Li, H. Li and X. Shi, AF9 YEATS domain links histone acetylation to DOT1L-mediated H3K79 methylation, *Cell*, 2014, **159**, 558–571.
- 67 L. Zeng, Q. Zhang, S. Li, A. N. Plotnikov, M. J. Walsh and M. M. Zhou, Mechanism and

- regulation of acetylated histone binding by the tandem PHD finger of DPF3b, *Nature*, 2010, **466**, 258–262.
- 68 D. Su, Q. Hu, Q. Li, J. R. Thompson, G. Cui, A. Fazly, B. A. Davies, M. V. Botuyan, Z. Zhang and G. Mer, Structural basis for recognition of H3K56-acetylated histone H3-H4 by the chaperone Rtt106, *Nature*, 2012, **483**, 104–109.
- 69 B. Li, M. Carey and J. L. Workman, The Role of Chromatin during Transcription, *Cell*, 2007, **128**, 707–719.
- 70 S. K. Kurdistani, S. Tavazoie and M. Grunstein, Mapping global histone acetylation patterns to gene expression, *Cell*, 2004, **117**, 721–733.
- 71 C. E. Brown, T. Lechner, L. A. Howe and J. L. Workman, The many HATs of transcription coactivators, *Trends Biochem. Sci.*, 2000, **25**, 15–19.
- 72 M. Grunstein, Histone acetylation in chromatin structure and transcription, *Nature*, 1997, **389**, 349–352.
- 73 M. F. Dion, S. J. Altschuler, L. F. Wu and O. J. Rando, Genomic characterization reveals a simple histone H4 acetylation code, *PNAS*, 2005, **102**, 5501–5506.
- 74 Q. Jin, L. R. Yu, L. Wang, Z. Zhang, L. H. Kasper, J. E. Lee, C. Wang, P. K. Brindle, S. Y. R. Dent and K. Ge, Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation, *EMBO J.*, 2011, **30**, 249–262.
- 75 A. Kimura, K. Matsubara and M. Horikoshi, A Decade of Histone Acetylation: Marking Eukaryotic Chromosomes with Specific Codes, *J. Biochem.*, 2005, **138**, 647–662.
- 76 X. J. Yang and E. Seto, HATs and HDACs: From structure, function and regulation to novel strategies for therapy and prevention, *Oncogene*, 2007, **26**, 5310–5318.
- 77 M. Shibuya, M. Hatano and K. Kawamura, Interactive histone acetylation and methylation in regulating transdifferentiation-related genes during tunicate budding and regeneration, *Dev. Dyn.*, 2015, **244**, 10–20.
- 78 T. Banerjee and D. Chakravarti, A Peek into the Complex Realm of Histone Phosphorylation, *Mol. Cell. Biol.*, 2011, **31**, 4858–4873.
- 79 C. Caron, C. Boyault and S. Khochbin, Regulatory cross-talk between lysine acetylation and ubiquitination: Role in control of protein stability, *BioEssays*, 2005, **27**, 408–415.
- 80 S. Margiola, K. Gerecht and M. M. Müller, Semisynthetic ‘designer’ p53 sheds light on a phosphorylation–acetylation relay, *Chem. Sci.*, 2021, **12**, 8563–8570.
- 81 J. Habibian and B. S. Ferguson, The crosstalk between acetylation and phosphorylation: Emerging new roles for HDAC inhibitors in the heart, *Int. J. Mol. Sci.*, 2019, **20**, 102.
- 82 B. D. Bryson and F. M. White, Quantitative Profiling of Lysine Acetylation Reveals Dynamic Crosstalk between Receptor Tyrosine Kinases and Lysine Acetylation, *PLoS One*, 2015, **10**, e0126242.
- 83 K. Sakaguchi, J. E. Herrera, S. Saito, T. Miki, M. Bustin, A. Vassilev, C. W. Anderson and E. Appella, DNA damage activates p53 through a phosphorylation-acetylation cascade, *Genes Dev.*, 1998, **12**, 2831–2841.
- 84 W. Stünkel and R. M. Campbell, Sirtuin 1 (SIRT1): The misunderstood HDAC, *J. Biomol. Screen.*, 2011, **16**, 1153–1169.
- 85 M. Huang, J. Huang, Y. Zheng and Q. Sun, Histone acetyltransferase inhibitors: An overview in synthesis, structure-activity relationship and molecular mechanism, *Eur. J. Med. Chem.*, 2019, **178**, 259–286.
- 86 E. Seto and M. Yoshida, Erasers of histone acetylation: The histone deacetylase enzymes, *Cold Spring Harb. Perspect. Biol.*, 2014, **6**, a018713.
- 87 V. Di Cerbo, R. Schneider, V. Di Cerbo and R. Schneider, Cancers with wrong HATs: The impact of acetylation, *Brief. Funct. Genomics*, 2013, **12**, 231–243.
- 88 A. E. Ringel and C. Wolberger, Structural basis for acyl-group discrimination by human Gcn5L2, *Acta Crystallogr. Sect. D Struct. Biol.*, 2016, **72**, 841–848.

- 89 R. Marmorstein and R. C. Trievel, Histone modifying enzymes: Structures, mechanisms, and specificities, *Biochim. Biophys. Acta - Gene Regul. Mech.*, 2009, **1789**, 58–68.
- 90 S. Y. Roth, J. M. Denu and C. D. Allis, Histone Acetyltransferases, *Annu. Rev. Biochem.*, 2001, **70**, 81–120.
- 91 L. Liu, X. T. Zhen, E. Denton, B. D. Marsden and M. Schapira, ChromoHub: A data hub for navigators of chromatin-mediated signalling, *Bioinformatics*, 2012, **28**, 2205–2206.
- 92 C. H. Arrowsmith, C. Bountra, P. V. Fish, K. Lee and M. Schapira, Epigenetic protein families: A new frontier for drug discovery, *Nat. Rev. Drug Discov.*, 2012, **11**, 384–400.
- 93 R. A. Frye, Phylogenetic classification of prokaryotic and eukaryotic Sir2-like proteins, *Biochem. Biophys. Res. Commun.*, 2000, **273**, 793–798.
- 94 X. J. Yang, V. V. Ogryzko, J. I. Nishikawa, B. H. Howard and Y. Nakatani, A p300/CPB-associated factor that competes with the adenoviral oncoprotein E1A, *Nature*, 1996, **382**, 319–324.
- 95 A. Clements, J. R. Rojas, R. C. Trievel, L. Wang, S. L. Berger and R. Marmorstein, Crystal structure of the histone acetyltransferase domain of the human PCAF transcriptional regulator bound to coenzyme A, *EMBO J.*, 1999, **18**, 3521–3532.
- 96 W. Xu, D. G. Edmondson and S. Y. Roth, Mammalian GCN5 and P/CAF Acetyltransferases Have Homologous Amino-Terminal Domains Important for Recognition of Nucleosomal Substrates, *Mol. Cell. Biol.*, 1998, **18**, 5659–5669.
- 97 Z. Nagy and L. Tora, Distinct GCN5/PCAF-containing complexes function as co-activators and are involved in transcription factor and global histone acetylation, *Oncogene*, 2007, **26**, 5341–5357.
- 98 E. R. Smith, J. M. Belote, R. L. Schiltz, X. J. Yang, P. A. Moore, S. L. Berger, Y. Nakatani and C. D. Allis, Cloning of drosophila GCN5: Conserved features among metazoan GCN5 family members, *Nucleic Acids Res.*, 1998, **26**, 2948–2954.
- 99 E. A. Koutsogiannouli, N. Wagner, C. Hader, M. Pinkerneil, M. J. Hoffmann, W. A. Schulz, E. A. Koutsogiannouli, N. Wagner, C. Hader, M. Pinkerneil, M. J. Hoffmann and W. A. Schulz, Differential Effects of Histone Acetyltransferase GCN5 or PCAF Knockdown on Urothelial Carcinoma Cells, *Int. J. Mol. Sci.*, 2017, **18**, 1449.
- 100 T. Yamauchi, J. Yamauchi, T. Kuwata, T. Tamura, T. Yamashita, N. Bae, H. Westphal, K. Ozato and Y. Nakatani, Distinct but overlapping roles of histone acetylase PCAF and of the closely related PCAF-B/GCN5 in mouse embryogenesis, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 11303–11306.
- 101 C. Dhalluin, J. E. Carlson, L. Zeng, C. He, A. K. Aggarwal and M.-M. Zhou, Structure and ligand of a histone acetyltransferase bromodomain, *Nature*, 1999, **399**, 491–496.
- 102 S. T. Fukai, R. Hibi, T. Naganuma, M. Sakai, S. Saijo, N. Shimizu, M. Matsumoto and T. Shimizu, Crystal structure of GCN5 PCAF N-terminal domain reveals atypical ubiquitin ligase structure, *J. Biol. Chem.*, 2020, **295**, 14630–14639.
- 103 L. K. Linares, R. Kiernan, R. Triboulet, C. Chable-Bessia, D. Latreille, O. Cuvier, M. Lacroix, L. Le Cam, O. Coux and M. Benkirane, Intrinsic ubiquitination activity of PCAF controls the stability of the oncoprotein Hdm2, *Nat. Cell Biol.*, 2007, **9**, 331–338.
- 104 J. R. Rojas, R. C. Trievel, J. Zhou, Y. Mo, X. Li, S. L. Berger, C. D. Allis and R. Marmorstein, Structure of Tetrahymena GCN5 bound to coenzyme A and a histone H3 peptide, *Nature*, 1999, **401**, 93–98.
- 105 A. Clements, A. N. Poux, W. S. Lo, L. Pillus, S. L. Berger and R. Marmorstein, Structural basis for histone and phosphohistone binding by the GCN5 histone acetyltransferase, *Mol. Cell*, 2003, **12**, 461–473.
- 106 K. G. Tanner, M. R. Langer and J. M. Denu, Kinetic Mechanism of Human Histone Acetyltransferase P/CAF †, *Biochemistry*, 2000, **39**, 11961–11969.
- 107 C. E. Berndsen and J. M. Denu, Catalysis and substrate selection by histone/protein lysine

- acetyltransferases, *Curr. Opin. Struct. Biol.*, 2008, **18**, 682–689.
- 108 O. D. Lau, A. D. Courtney, A. Vassilev, L. A. Marzilli, R. J. Cotter, Y. Nakatani and P. A. Cole, p300/CBP-associated factor histone acetyltransferase processing of a peptide substrate: Kinetic analysis of the catalytic mechanism, *J. Biol. Chem.*, 2000, **275**, 21953–21959.
 - 109 K. G. Tanner, R. C. Trievel, M.-H. Kuo, R. M. Howard, S. L. Berger, C. David Allis, R. Marmorstein and J. M. Denu, Catalytic Mechanism and Function of Invariant Glutamic Acid 173 from the Histone Acetyltransferase GCN5 Transcriptional Coactivator*, *J. Biol. Chem.*, 1999, **274**, 18157–18160.
 - 110 J. Jiang, J. Lu, D. Lu, Z. Liang, L. Li, S. Ouyang, X. Kong, H. Jiang, B. Shen and C. Luo, Investigation of the Acetylation Mechanism by GCN5 Histone Acetyltransferase, *PLoS One*, 2012, **7**, e36660.
 - 111 Y.-L. Wang, F. Faiola, M. Xu, S. Pan and E. Martinez, Human ATAC Is a GCN5/PCAF-containing Acetylase Complex with a Novel NC2-like Histone Fold Module That Interacts with the TATA-binding Protein, *J. Biol. Chem.*, 2008, **283**, 33808–33815.
 - 112 S. P. Baker and P. A. Grant, The SAGA continues: Expanding the cellular role of a transcriptional co-activator complex, *Oncogene*, 2007, **26**, 5329–5340.
 - 113 J. M. Espinola-Lopez and S. Tan, The Ada2/Ada3/Gcn5/Sgf29 histone acetyltransferase module, *Biochim. Biophys. Acta - Gene Regul. Mech.*, 2021, **1864**, 194629.
 - 114 M. Semer, B. Bidon, A. Larnicol, G. Caliskan, P. Catez, J. M. Egly, F. Coin and N. Le May, DNA repair complex licenses acetylation of H2A.Z.1 by KAT2A during transcription, *Nat. Chem. Biol.*, 2019, **15**, 992–1000.
 - 115 N. Au-Yeung and C. M. Horvath, Histone H2A.Z Suppression of Interferon-Stimulated Transcription and Antiviral Immunity Is Modulated by GCN5 and BRD2, *iScience*, 2018, **6**, 68–82.
 - 116 S. M. T. Hoke, J. Genereaux, G. Liang and C. J. Brandl, A Conserved Central Region of Yeast Ada2 Regulates the Histone Acetyltransferase Activity of Gcn5 and Interacts with Phospholipids, *J. Mol. Biol.*, 2008, **384**, 743–755.
 - 117 N. Suka, Y. Suka, A. A. Carmen, J. Wu and M. Grunstein, Highly specific antibodies determine histone acetylation site usage in yeast heterochromatin and euchromatin, *Mol. Cell*, 2001, **8**, 473–479.
 - 118 R. L. Schiltz, C. A. Mizzen, A. Vassilev, R. G. Cook, C. D. Allis and Y. Nakatani, Overlapping but distinct patterns of histone acetylation by the human coactivators p300 and PCAF within nucleosomal substrates, *J. Biol. Chem.*, 1999, **274**, 1189–1192.
 - 119 R. P. Simon, D. Robaa, Z. Alhalabi, W. Sippl and M. Jung, KATching-Up on Small Molecule Modulators of Lysine Acetyltransferases, *J. Med. Chem.*, 2016, **59**, 1249–1270.
 - 120 R. A. Henry, Y. M. Kuo, V. Bhattacharjee, T. J. Yen and A. J. Andrews, Changing the selectivity of p300 by acetyl-coa modulation of histone acetylation, *ACS Chem. Biol.*, 2015, **10**, 146–156.
 - 121 P. A. Grant, A. Eberharter, S. John, R. G. Cook, B. M. Turner and J. L. Workman, Expanded lysine acetylation specificity of Gcn5 in native complexes, *J. Biol. Chem.*, 1999, **274**, 5895–5900.
 - 122 A. M. Cieniewicz, L. Moreland, A. E. Ringel, S. G. Mackintosh, A. Raman, T. M. Gilbert, C. Wolberger, A. J. Tackett and S. D. Taverna, The bromodomain of gcn5 regulates site specificity of lysine acetylation on histone H3, *Mol. Cell. Proteomics*, 2014, **13**, 2896–2910.
 - 123 H. Santos-Rosa, E. Valls, T. Kouzarides, M. Martinez-Balbas, M. Martõ Ànez-Balba and M. Martinez-Balbas, Mechanism of PCAF auto-acetylation, *Nucleic Acids Res.*, 2003, **31**, 4285–4292.
 - 124 R. Marmorstein, Structure and function of histone acetyltransferases, *Cell. Mol. Life Sci.*, 2001, **58**, 693–703.

- 125 J. H. Lim, Y. M. Lee, Y. S. Chun, J. Chen, J. E. Kim and J. W. Park, Sirtuin 1 modulates cellular responses to hypoxia by deacetylating hypoxia-inducible factor 1alpha., *Mol. Cell*, 2010, **38**, 864–878.
- 126 L. Liu, D. M. Scolnick, R. C. Trievel, H. B. Zhang, R. Marmorstein, T. D. Halazonetis and S. L. Berger, p53 Sites Acetylated In Vitro by PCAF and p300 Are Acetylated In Vivo in Response to DNA Damage, *Mol. Cell. Biol.*, 1999, **19**, 1202–1209.
- 127 K. Okumura, M. Mendoza, R. M. Bachoo, R. A. DePinho, W. K. Cavenee and F. B. Furnari, PCAF modulates PTEN activity, *J. Biol. Chem.*, 2006, **281**, 26562–26568.
- 128 A. Imhof, X. J. Yang, V. V. Ogryzko, Y. Nakatani, A. P. Wolffe and H. Ge, Acetylation of general transcription factors by histone acetyltransferases, *Curr. Biol.*, 1997, **7**, 689–692.
- 129 F. Mateo, M. Vidal-Laliena, N. Canela, A. Zecchin, M. Martínez-Balbás, N. Agell, M. Giacca, M. J. Pujol and O. Bachs, The transcriptional co-activator PCAF regulates cdk2 activity., *Nucleic Acids Res.*, 2009, **37**, 7072–84.
- 130 P. Xia, Z. Wang, X. Liu, B. Wu, J. Wang, T. Ward, L. Zhang, X. Ding, G. Gibbons, Y. Shi and X. Yao, EB1 acetylation by P300/CBP-associated factor (PCAF) ensures accurate kinetochore-microtubule interactions in mitosis, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 16564–16569.
- 131 F. Mateo, M. Vidal-Laliena, N. Canela, L. Busino, M. A. Martinez-Balbas, M. Pagano, N. Agell and O. Bachs, Degradation of cyclin A is regulated by acetylation, *Oncogene*, 2009, **28**, 2654–2666.
- 132 B. J. Wilson, A. M. Tremblay, G. Deblois, G. Sylvain-Drolet and V. Giguère, An acetylation switch modulates the transcriptional activity of estrogen-related receptor α , *Mol. Endocrinol.*, 2010, **24**, 1349–1358.
- 133 D. Zhao, Y. Mo, M. T. Li, S. W. Zou, Z. L. Cheng, Y. P. Sun, Y. Xiong, K. L. Guan and Q. Y. Lei, NOTCH-induced aldehyde dehydrogenase 1A1 deacetylation promotes breast cancer stem cells, *J. Clin. Invest.*, 2014, **124**, 5453–5465.
- 134 X. Cao, C. Li, S. Xiao, Y. Tang, J. Huang, S. Zhao, X. Li, J. Li, R. Zhang and W. Yu, Acetylation promotes TyrRS nuclear translocation to prevent oxidative damage, *Proc. Natl. Acad. Sci. U. S. A.*, 2017, **114**, 687–692.
- 135 L. Yu, Z. Li, M. Fang and Y. Xu, Acetylation of MKL1 by PCAF regulates pro-inflammatory transcription, *Biochim. Biophys. Acta - Gene Regul. Mech.*, 2017, **1860**, 839–847.
- 136 K. Shanmugasundaram, B. K. Nayak, W. E. Friedrichs, D. Kaushik, R. Rodriguez and K. Block, NOX4 functions as a mitochondrial energetic sensor coupling cancer metabolic reprogramming to drug resistance, *Nat. Commun.*, 2017, **8**, 997.
- 137 R. H. F. Wong, I. Chang, C. S. S. Hudak, S. Hyun, H. Y. Kwan and H. S. Sul, A Role of DNA-PK for the Metabolic Gene Regulation in Response to Insulin, *Cell*, 2009, **136**, 1056–1072.
- 138 C. Z. Song, K. Keller, Y. Chen and G. Stamatoyannopoulos, Functional interplay between CBP and PCAF in acetylation and regulation of transcription factor KLF13 activity, *J. Mol. Biol.*, 2003, **329**, 207–215.
- 139 A. M. Urvalek, X. Wang, H. Lu and J. Zhao, KLF8 recruits the p300 and pPCAF co-activators to its amino terminal activation domain to activate transcription, *Cell Cycle*, 2010, **9**, 601–611.
- 140 L. H. Zhu, L. H. Sun, Y. L. Hu, Y. Jiang, H. Y. Liu, X. Y. Shen, X. Y. Jin, X. Zhen, H. X. Sun and G. J. Yan, PCAF impairs endometrial receptivity and embryo implantation by down-regulating β 3-integrin expression via HOXA10 acetylation, *J. Clin. Endocrinol. Metab.*, 2013, **98**, 4417–4428.
- 141 A. Farria, W. Li and S. Y. R. Dent, KATs in cancer: Functions and therapies, *Oncogene*, 2015, **34**, 4901–4913.
- 142 M. Malatesta, C. Steinhauer, F. Mohammad, D. P. Pandey, M. Squatrito and K. Helin, Histone acetyltransferase PCAF is required for hedgehog-Gli-dependent transcription

- and cancer cell proliferation, *Cancer Res.*, 2013, **73**, 6323–6333.
- 143 X. Gai, K. Tu, C. Li, Z. Lu, L. R. Roberts and X. Zheng, Histone acetyltransferase PCAF accelerates apoptosis by repressing a GIL1/BCL2/BAX axis in hepatocellular carcinoma, *Cell Death Dis.*, 2015, **6**, e1712–e1712.
- 144 Q. Li, Z. Liu, M. Xu, Y. Xue, B. Yao, C. Dou, Y. Jia, Y. Wang, K. Tu, X. Zheng and Y. Yao, PCAF inhibits hepatocellular carcinoma metastasis by inhibition of epithelial-mesenchymal transition by targeting Gli-1, *Cancer Lett.*, 2016, **375**, 190–198.
- 145 X. Zheng, X. Gai, F. Ding, Z. Lu, K. Tu, Y. Yao and Q. Liu, Histone acetyltransferase PCAF Up-regulated cell apoptosis in hepatocellular carcinoma via acetylating histone H4 and inactivating AKT signaling, *Mol. Cancer*, 2013, **12**, 96.
- 146 H. Y. Cohen, S. Lavu, K. J. Bitterman, B. Hekking, T. A. Imahiyerobo, C. Miller, R. Frye, H. Ploegh, B. M. Kessler and D. A. Sinclair, Acetylation of the C terminus of Ku70 by CBP and PCAF controls Bax-mediated apoptosis, *Mol. Cell*, 2004, **13**, 627–638.
- 147 H. J. Fei, L. D. Zu, J. Wu, X. S. Jiang, J. L. Wang, Y. E. Chin and G. H. Fu, PCAF acts as a gastric cancer suppressor through a novel PCAF-p16-CDK4 axis, *Am. J. Cancer Res.*, 2016, **6**, 2772–2786.
- 148 G. Hirano, H. Izumi, A. Kidani, Y. Yasuniwa, B. Han, H. Kusaba, K. Akashi, M. Kuwano and K. Kohno, Enhanced expression of PCAF endows apoptosis resistance in cisplatin-resistant cells, *Mol. Cancer Res.*, 2010, **8**, 864–872.
- 149 K. Cai, Y. Wan, Z. Wang, Y. Wang, X. Zhao and X. Bao, C5a promotes the proliferation of human nasopharyngeal carcinoma cells through PCAF-mediated STAT3 acetylation, *Oncol. Rep.*, 2014, **32**, 2260–2266.
- 150 J. H. Patel, Y. Du, P. G. Ard, C. Phillips, B. Carella, C.-J. Chen, C. Rakowski, C. Chatterjee, P. M. Lieberman, W. S. Lane, G. A. Blobel and S. B. McMahon, The c-MYC Oncoprotein Is a Substrate of the Acetyltransferases hGCN5/PCAF and TIP60, *Mol. Cell. Biol.*, 2004, **24**, 10826–10834.
- 151 J. Wan, J. Zhan, S. Li, J. Ma, W. Xu, C. Liu, X. Xue, Y. Xie, W. Fang, Y. E. Chin and H. Zhang, PCAF-primed EZH2 acetylation regulates its stability and promotes lung adenocarcinoma progression, *Nucleic Acids Res.*, 2015, **43**, 3591–3604.
- 152 S. Zhang, G. Sun, Z. Wang, Y. Wan, J. Guo and L. Shi, PCAF-mediated Akt1 acetylation enhances the proliferation of human glioblastoma cells, *Tumor Biol.*, 2014, **36**, 1455–1462.
- 153 P. Eguía-Aguilar, M. Solís-Paredes, P. Reyes-Cid, M. Perezpeña-Diazconti, F. C. P. De León, S. Sadowinski-Pine and F. Arenas-Huertero, Expression of histone acetylases p300 and PCAF in pediatric astrocytomas, *Child's Nerv. Syst.*, 2013, **29**, 1089–1096.
- 154 R. Lin, R. Tao, X. Gao, T. Li, X. Zhou, K. L. Guan, Y. Xiong and Q. Y. Lei, Acetylation stabilizes ATP-citrate lyase to promote lipid biosynthesis and tumor growth, *Mol. Cell*, 2013, **51**, 506–518.
- 155 A. Y. Gong, A. N. Eischeid, J. Xiao, J. Zhao, D. Chen, Z. Y. Wang, C. Y. F. Young and X. M. Chen, MiR-17-5p targets the p300/CBP-associated factor and modulates androgen receptor transcriptional activity in cultured prostate cancer cells, *BMC Cancer*, 2012, **12**, 492.
- 156 X. Ge, Q. Jin, F. Zhang, T. Yan and Q. Zhai, PCAF Acetylates β -catenin and improves its stability, *Mol. Biol. Cell*, 2009, **20**, 419–427.
- 157 M. Shiota, A. Yokomizo, Y. Tada, T. Uchiumi, J. Inokuchi, K. Tatsugami, K. Kuroiwa, K. Yamamoto, N. Seki and S. Naito, P300/CBP-associated factor regulates Y-box binding protein-1 expression and promotes cancer cell growth, cancer invasion and drug resistance, *Cancer Sci.*, 2010, **101**, 1797–1806.
- 158 R. E. Kiernan, HIV-1 Tat transcriptional activity is regulated by acetylation, *EMBO J.*, 1999, **18**, 6106–6118.

- 159 S. Mujtaba, Y. He, L. Zeng, A. Farooq, J. E. Carlson, M. Ott, E. Verdin and M. M. Zhou, Structural basis of lysine-acetylated HIV-1 Tat recognition by PCAF bromodomain, *Mol. Cell*, 2002, **9**, 575–586.
- 160 Q. Wang, R. Wang, B. Zhang, S. Zhang, Y. Zheng and Z. Wang, Small organic molecules targeting PCAF bromodomain as potent inhibitors of HIV-1 replication, *Medchemcomm*, 2013, **4**, 737–740.
- 161 Q. Jin, L. Zhuang, B. Lai, C. Wang, W. Li, B. Dolan, Y. Lu, Z. Wang, K. Zhao, W. Peng, S. Y. Dent and K. Ge, Gcn5 and PCAF negatively regulate interferon- β production through HAT-independent inhibition of TBK 1, *EMBO Rep.*, 2014, **15**, 1192–1201.
- 162 D. Hatakeyama, T. Masuda, R. Miki, S. Ohtsuki and T. Kuzuhara, In-vitro acetylation of SARS-CoV and SARS-CoV-2 nucleocapsid proteins by human PCAF and GCN5, *Biochem. Biophys. Res. Commun.*, 2021, **557**, 273–279.
- 163 W. Wei, C. M. Coelho, X. Li, R. Marek, S. Yan, S. Anderson, D. Meyers, C. Mukherjee, G. Sbardella, S. Castellano, C. Milite, D. Rotili, A. Mai, P. A. Cole, P. Sah, M. S. Kobor and T. W. Bredy, P300/CBP-associated factor selectively regulates the extinction of conditioned fear, *J. Neurosci.*, 2012, **32**, 11930–11941.
- 164 T. Maurice, F. Duclot, J. Meunier, G. Naert, L. Givalois, J. Meffre, A. Célérier, C. Jacquet, V. Copois, N. Mechti, K. Ozato and C. Gongora, Altered memory capacities and response to stress in p300/CBP-associated factor (PCAF) histone acetylase knockout mice, *Neuropsychopharmacology*, 2008, **33**, 1584–1602.
- 165 O. Bousiges, A. Pereira De Vasconcelos, R. Neidl, B. Cosquer, K. Herbeaux, I. Panteleeva, J.-P. Loeffler, J.-C. Cassel and A.-L. Boutillier, Spatial Memory Consolidation is Associated with Induction of Several Lysine-Acetyltransferase (Histone Acetyltransferase) Expression Levels and H2B/H4 Acetylation-Dependent Transcriptional Events in the Rat Hippocampus, *Neuropsychopharmacology*, 2010, **35**, 2521–2537.
- 166 F. Duclot, J. Meffre, C. Jacquet, C. Gongora and T. Maurice, Mice knock out for the histone acetyltransferase p300/CREB binding protein-associated factor develop a resistance to amyloid toxicity, *Neuroscience*, 2010, **167**, 850–863.
- 167 J. S. Steffan, L. Bodai, J. Pallos, M. Poelman, A. McCampbell, B. L. Apostol, A. Kazantsev, E. Schmidt, Y. Z. Zhu, M. Greenwald, R. Kurokawa, D. E. Housman, G. R. Jackson, J. L. Marsh and L. M. Thompson, Histone deacetylase inhibitors arrest polyglutamine-dependent neurodegeneration in Drosophila, *Nature*, 2001, **413**, 739–743.
- 168 S. K. Piroozniat and F. Elefant, Targeting specific HATs for neurodegenerative disease treatment: Translating basic biology to therapeutic possibilities, *Front. Cell. Neurosci.*, 2013, **7**, 30.
- 169 A. Schneider, S. Chatterjee, O. Bousiges, B. R. Selvi, A. Swaminathan, R. Cassel, F. Blanc, T. K. Kundu and A. L. Boutillier, Acetyltransferases (HATs) as Targets for Neurological Therapeutics, *Neurotherapeutics*, 2013, **10**, 568–588.
- 170 C. Colussi, J. Rosati, S. Straino, F. Spallotta, R. Berni, D. Stilli, S. Rossi, E. Musso, E. Macchi, A. Mai, G. Sbardella, S. Castellano, C. Chimenti, A. Frustaci, A. Nebbioso, L. Altucci, M. C. Capogrossi and C. Gaetano, N ϵ -lysine acetylation determines dissociation from GAP junctions and lateralization of connexin 43 in normal and dystrophic heart, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 2795–2800.
- 171 K.-A. Sheppard, D. W. Rose, Z. K. Haque, R. Kurokawa, E. McInerney, S. Westin, D. Thanos, M. G. Rosenfeld, C. K. Glass and T. Collins, Transcriptional Activation by NF- κ B Requires Multiple Coactivators, *Mol. Cell. Biol.*, 1999, **19**, 6367–6378.
- 172 R. C. M. de Jong, M. M. Ewing, M. R. de Vries, J. C. Karper, A. J. N. M. Bastiaansen, H. A. B. Peters, F. Baghana, P. J. van den Elsen, C. Gongora, J. W. Jukema and P. H. A. Quax, The epigenetic factor PCAF regulates vascular inflammation and is essential for intimal hyperplasia development, *PLoS One*, 2017, **12**, e0185820.

- 173 C. Sun, M. Wang, X. Liu, L. Luo, K. Li, S. Zhang, Y. Wang, Y. Yang, F. Ding and X. Gu, PCAF Improves Glucose Homeostasis by Suppressing the Gluconeogenic Activity of PGC-1 α , *Cell Rep.*, 2014, **9**, 2250–2262.
- 174 T. Wang, W. Yao, Y. Shao, R. Zheng and F. Huang, PCAF fine-tunes hepatic metabolic syndrome, inflammatory disease, and cancer, *J. Cell. Mol. Med.*, 2018, **22**, 5787–5800.
- 175 M. Downey, Non-histone protein acetylation by the evolutionarily conserved GCN5 and PCAF acetyltransferases, *Biochim. Biophys. Acta - Gene Regul. Mech.*, 2021, **1864**, 194608.
- 176 D. Helmlinger and L. Tora, Sharing the SAGA, *Trends Biochem. Sci.*, 2017, **42**, 850–861.
- 177 G. Spedale, H. T. M. Timmers and W. W. M. P. Pijnappel, ATAC-king the complexity of SAGA during evolution, *Genes Dev.*, 2012, **26**, 527–541.
- 178 M. Jeitany, D. Bakhos-Douaihy, D. C. Silvestre, J. R. Pineda, N. Ugolin, A. Moussa, L. R. Gauthier, D. Busso, M. P. Junier, H. Chneiweiss, S. Chevillard, C. Desmaze and F. D. Boussin, Opposite effects of GCN5 and PCAF knockdowns on the alternative mechanism of telomere maintenance, *Oncotarget*, 2017, **8**, 26269–26280.
- 179 M. Toth, I. M. Boros and E. Balint, Elevated level of lysine 9-acetylated histone H3 at the MDR1 promoter in multidrug-resistant cells, *Cancer Sci.*, 2012, **103**, 659–669.
- 180 D. Hatakeyama, M. Shoji, S. Yamayoshi, R. Yoh, N. Ohmi, S. Takenaka, A. Saitoh, Y. Arakaki, A. Masuda, T. Komatsu, R. Nagano, M. Nakano, T. Noda, Y. Kawaoka and T. Kuzuhara, Influenza A virus nucleoprotein is acetylated by histone acetyltransferases PCAF and GCN5, *J. Biol. Chem.*, 2018, **293**, 7126–7138.
- 181 M. Tanno, J. Sakamoto, T. Miura, K. Shimamoto and Y. Horio, Nucleocytoplasmic shuttling of the NAD⁺-dependent histone deacetylase SIRT1, *J. Biol. Chem.*, 2007, **282**, 6823–6832.
- 182 H. Jing and H. Lin, Sirtuins in Epigenetic Regulation, *Chem. Rev.*, 2015, **115**, 2350–2375.
- 183 P. Oberdoerffer, S. Michan, M. McVay, R. Mostoslavsky, J. Vann, S. K. Park, A. Hartlerode, J. Stegmuller, A. Hafner, P. Loerch, S. M. Wright, K. D. Mills, A. Bonni, B. A. Yankner, R. Scully, T. A. Prolla, F. W. Alt and D. A. Sinclair, SIRT1 Redistribution on Chromatin Promotes Genomic Stability but Alters Gene Expression during Aging, *Cell*, 2008, **135**, 907–918.
- 184 A. Vaquero, M. Scher, D. Lee, H. Erdjument-Bromage, P. Tempst and D. Reinberg, Human SirT1 interacts with histone H1 and promotes formation of facultative heterochromatin, *Mol. Cell*, 2004, **16**, 93–105.
- 185 C. Das, M. S. Lucia, K. C. Hansen and J. K. Tyler, CBP/p300-mediated acetylation of histone H3 on lysine 56, *Nature*, 2009, **459**, 113–117.
- 186 H. Vaziri, S. K. Dessain, E. N. Eaton, S. I. Imai, R. A. Frye, T. K. Pandita, L. Guarente and R. A. Weinberg, hSIR2/SIRT1 functions as an NAD-dependent p53 deacetylase, *Cell*, 2001, **107**, 149–159.
- 187 N. Pediconi, F. Guerrieri, S. Vossio, T. Bruno, L. Belloni, V. Schinzari, C. Scisciani, M. Fanciulli and M. Levrero, hSirT1-Dependent Regulation of the PCAF-E2F1-p73 Apoptotic Pathway in Response to DNA Damage, *Mol. Cell. Biol.*, 2009, **29**, 1989–1998.
- 188 S. Hisahara, S. Chiba, H. Matsumoto, M. Tanno, H. Yagi, S. Shimohama, M. Sato and Y. Horio, Histone deacetylase SIRT1 modulates neuronal differentiation by its nuclear translocation, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 15599–15604.
- 189 T. Sugino, M. Maruyama, M. Tanno, A. Kuno, K. Houkin and Y. Horio, Protein deacetylase SIRT1 in the cytoplasm promotes nerve growth factor-induced neurite outgrowth in PC12 cells, *FEBS Lett.*, 2010, **584**, 2821–2826.
- 190 S. Leng, W. Huang, Y. Chen, Y. Yang, D. Feng, W. Liu, T. Gao, Y. Ren, M. Huo, J. Zhang, Y. Yang and Y. Wang, SIRT1 coordinates with the CRL4B complex to regulate pancreatic cancer stem cells to promote tumorigenesis, *Cell Death Differ.*, 2021, 1–15.
- 191 L. Tiberi, J. Bonnefont, J. vandenAmele, S. D. LeBon, A. Herpoel, A. Bilheu, B. W. Baron

- and P. Vanderhaeghen, A BCL6/BCOR/SIRT1 Complex Triggers Neurogenesis and Suppresses Medulloblastoma by Repressing Sonic Hedgehog Signaling, *Cancer Cell*, 2014, **26**, 797–812.
- 192 S. Oka, R. Alcendor, P. Zhai, J. Y. Park, D. Shao, J. Cho, T. Yamamoto, B. Tian and J. Sadoshima, PPAR α -sirt1 complex mediates cardiac hypertrophy and failure through suppression of the err transcriptional pathway, *Cell Metab.*, 2011, **14**, 598–611.
- 193 A. Vassilopoulos, K. S. Fritz, D. R. Petersen and D. Gius, The human sirtuin family: Evolutionary divergences and functions, *Hum. Genomics*, 2011, **5**, 485–496.
- 194 M. J. Holness, P. W. Caton and M. C. Sugden, Acute and long-term nutrient-led modifications of gene expression: Potential role of SIRT1 as a central co-ordinator of short and longer-term programming of tissue function, *Nutrition*, 2010, **26**, 491–501.
- 195 X. Li, S. Zhang, G. Blander, J. G. Tse, M. Krieger and L. Guarente, SIRT1 Deacetylates and Positively Regulates the Nuclear Receptor LXR, *Mol. Cell*, 2007, **28**, 91–106.
- 196 M. M. Bellet and P. Sassone-Corsi, Mammalian circadian clock and metabolism - The epigenetic link, *J. Cell Sci.*, 2010, **123**, 3837–3848.
- 197 M. Kuningas, M. Putters, R. G. J. Westendorp, P. E. Slagboom and D. van Heemst, SIRT1 Gene, Age-Related Diseases, and Mortality: The Leiden 85-Plus Study, *Journals Gerontol. Ser. A Biol. Sci. Med. Sci.*, 2007, **62**, 960–965.
- 198 M. C. Zillikens, J. B. J. van Meurs, E. J. G. Sijbrands, F. Rivadeneira, A. Dehghan, J. P. T. M. van Leeuwen, A. Hofman, C. M. van Duijn, J. C. M. Witteman, A. G. Uitterlinden and H. A. P. Pols, SIRT1 genetic variation and mortality in type 2 diabetes: interaction with smoking and dietary niacin, *Free Radic. Biol. Med.*, 2009, **46**, 836–841.
- 199 S. W. Min, S. H. Cho, Y. Zhou, S. Schroeder, V. Haroutunian, W. W. Seeley, E. J. Huang, Y. Shen, E. Masliah, C. Mukherjee, D. Meyers, P. A. Cole, M. Ott and L. Gan, Acetylation of tau inhibits its degradation and contributes to tauopathy, *Neuron*, 2010, **67**, 953–966.
- 200 W. Qin, T. Yang, L. Ho, Z. Zhao, J. Wang, L. Chen, W. Zhao, M. Thiyagarajan, D. MacGrogan, J. T. Rodgers, P. Puigserver, J. Sadoshima, H. Deng, S. Pedrini, S. Gandy, A. A. Sauve and G. M. Pasinetti, Neuronal SIRT1 activation as a novel mechanism underlying the prevention of alzheimer disease amyloid neuropathology by calorie restriction, *J. Biol. Chem.*, 2006, **281**, 21745–21754.
- 201 T. T. Schug, Q. Xu, H. Gao, A. Peres-da-Silva, D. W. Draper, M. B. Fessler, A. Purushotham and X. Li, Myeloid Deletion of SIRT1 Induces Inflammatory Signaling in Response to Environmental Stress, *Mol. Cell. Biol.*, 2010, **30**, 4712–4721.
- 202 S. Rajendrasozhan, S. R. Yang, V. L. Kinnula and I. Rahman, SIRT1, an antiinflammatory and antiaging protein, is decreased in lungs of patients with chronic obstructive pulmonary disease, *Am. J. Respir. Crit. Care Med.*, 2008, **177**, 861–870.
- 203 R. Zhang, H. Z. Chen, J. J. Liu, Y. Y. Jia, Z. Q. Zhang, R. F. Yang, Y. Zhang, J. Xu, Y. S. Wei, D. P. Liu and C. C. Liang, SIRT1 suppresses activator protein-1 transcriptional activity and cyclooxygenase-2 expression in macrophages, *J. Biol. Chem.*, 2010, **285**, 7097–7110.
- 204 B. Wang, M. K. Hasan, E. Alvarado, H. Yuan, H. Wu and W. Y. Chen, NAMPT overexpression in prostate cancer and its contribution to tumor cell survival and stress response, *Oncogene*, 2011, **30**, 907–921.
- 205 B. Jung-Hynes, M. Nihal, W. Zhong and N. Ahmad, Role of sirtuin histone deacetylase SIRT1 in prostate cancer: A target for prostate cancer management via its inhibition?, *J. Biol. Chem.*, 2009, **284**, 3823–3832.
- 206 J. Chen, B. Zhang, N. Wong, A. W. I. Lo, K. F. To, A. W. H. Chan, M. H. L. Ng, C. Y. S. Ho, S. H. Cheng, P. B. S. Lai, J. Yu, H. K. Ng, M. T. Ling, A. L. Huang, X. F. Cai and B. C. B. Ko, Sirtuin 1 is upregulated in a subset of hepatocellular carcinomas where it is essential for telomere maintenance and tumor cell growth, *Cancer Res.*, 2011, **71**, 4138–4149.
- 207 R. H. Wang, K. Sengupta, C. Li, H. S. Kim, L. Cao, C. Xiao, S. Kim, X. Xu, Y. Zheng, B. Chilton,

- R. Jia, Z. M. Zheng, E. Appella, X. W. Wang, T. Ried and C. X. Deng, Impaired DNA Damage Response, Genome Instability, and Tumorigenesis in SIRT1 Mutant Mice, *Cancer Cell*, 2008, **14**, 312–323.
- 208 N. Kabra, Z. Li, L. Chen, B. Li, X. Zhang, C. Wang, T. Yeatman, D. Coppola and J. Chen, SirT1 is an inhibitor of proliferation and tumor formation in colon cancer, *J. Biol. Chem.*, 2009, **284**, 18210–18217.
- 209 K. G. Tanner, J. Landry, R. Sternglanz and J. M. Denu, Silent information regulator 2 family of NAD-dependent histone/protein deacetylases generates a unique product, 1-O-acetyl-ADP-ribose, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 14178–14182.
- 210 M. T. Borra, M. R. Langer, J. T. Slama and J. M. Denu, Substrate specificity and kinetic mechanism of the Sir2 family of NAD⁺-dependent histone/protein deacetylases, *Biochemistry*, 2004, **43**, 9877–9887.
- 211 B. M. Hirsch and W. Zheng, Sirtuin mechanism and inhibition: Explored with N ϵ -acetyl-lysine analogs, *Mol. Biosyst.*, 2011, **7**, 16–28.
- 212 B. C. Smith, W. C. Hallows and J. M. Denu, A continuous microplate assay for sirtuins and nicotinamide-producing enzymes, *Anal. Biochem.*, 2009, **394**, 101–109.
- 213 T. McDonagh, J. Hixon, P. S. DiStefano, R. Curtis and A. D. Napper, Microplate filtration assay for nicotinamide release from NAD using a boronic acid resin, *Methods*, 2005, **36**, 346–350.
- 214 N. R. Sundaresan, V. B. Pillai, D. Wolfgeher, S. Samant, P. Vasudevan, V. Parekh, H. Raghuraman, J. M. Cunningham, M. Gupta and M. P. Gupta, The deacetylase SIRT1 promotes membrane localization and activation of Akt and PDK1 during tumorigenesis and cardiac hypertrophy, *Sci. Signal.*, 2011, **4**, ra46–ra46.
- 215 L. xia Wang, J. Wang, T. ting Qu, Y. Zhang and Y. fei Shen, Reversible acetylation of Lin28 mediated by PCAF and SIRT1, *Biochim. Biophys. Acta - Mol. Cell Res.*, 2014, **1843**, 1188–1195.
- 216 G. S. Watts, M. M. Oshiro, D. J. Junk, R. J. Wozniak, S. J. Watterson, F. E. Domann and B. W. Futscher, The Acetyltransferase p300/CBP-Associated Factor Is a p53 Target Gene in Breast Tumor Cells, *Neoplasia*, 2004, **6**, 187–194.
- 217 J. M. Solomon, R. Pasupuleti, L. Xu, T. McDonagh, R. Curtis, P. S. DiStefano and L. J. Huber, Inhibition of SIRT1 Catalytic Activity Increases p53 Acetylation but Does Not Alter Cell Survival following DNA Damage, *Mol. Cell. Biol.*, 2006, **26**, 28–38.
- 218 S. Lain, J. J. Hollick, J. Campbell, O. D. Staples, M. Higgins, M. Aoubala, A. McCarthy, V. Appleyard, K. E. Murray, L. Baker, A. Thompson, J. Mathers, S. J. Holland, M. J. R. Stark, G. Pass, J. Woods, D. P. Lane and N. J. Westwood, Discovery, In Vivo Activity, and Mechanism of Action of a Small-Molecule p53 Activator, *Cancer Cell*, 2008, **13**, 454–463.
- 219 G. Xenaki, T. Ontikatz, R. Rajendran, I. J. Stratford, C. Dive, M. Krstic-Demonacos and C. Demonacos, PCAF is an HIF-1 α cofactor that regulates p53 transcriptional activity in hypoxia, *Oncogene*, 2008, **27**, 5785–5796.
- 220 A. Laemmle, A. Lechleiter, V. Roh, C. Schwarz and S. Portmann, Inhibition of SIRT1 Impairs the Accumulation and Transcriptional Activity of HIF-1 α Protein under Hypoxic Conditions, *PLoS One*, 2012, **7**, 33433.
- 221 A. Sermeus and C. Michiels, Reciprocal influence of the p53 and the hypoxic pathways, *Cell Death Dis.*, 2011, **2**, e164.
- 222 E. Verdin and M. Ott, 50 years of protein acetylation: From gene regulation to epigenetics, metabolism and beyond, *Nat. Rev. Mol. Cell Biol.*, 2015, **16**, 258–264.
- 223 L. A. Gates, J. Shi, A. D. Rohira, Q. Feng, B. Zhu, M. T. Bedford, C. A. Sagum, S. Y. Jung, J. Qin, M. J. Tsai, S. Y. Tsai, W. Li, C. E. Foulds and B. W. O'Malley, Acetylation on histone H3 lysine 9 mediates a switch from transcription initiation to elongation, *J. Biol. Chem.*, 2017, **292**, 14456–14472.

- 224 J. A. Park, A. J. Kim, Y. Kang, Y. J. Jung, H. K. Kim and K. C. Kim, Deacetylation and methylation at histone H3 lysine 9 (H3K9) coordinate chromosome condensation during cell cycle progression, *Mol. Cells*, 2011, **31**, 343–349.
- 225 H. Ji, Y. Zhou, X. Zhuang, Y. Zhu, Z. Wu, Y. Lu, S. Li, Y. Zeng, Q. R. Lu, Y. Huo, Y. Shi and H. Bu, HDAC3 deficiency promotes liver cancer through a defect in H3K9ac/H3K9me3 transition, *Cancer Res.*, 2019, **79**, 3676–3688.
- 226 M. Nevil and R. J. Duronio, The awesome power of histone genetics, *Mol. Cell*, 2021, **81**, 1593–1595.
- 227 I. Regadas, O. Dahlberg, R. Vaid, O. Ho, S. Belikov, G. Dixit, S. Deindl, J. Wen and M. Mannervik, A unique histone 3 lysine 14 chromatin signature underlies tissue-specific gene regulation, *Mol. Cell*, 2021, **81**, 1766–1780.
- 228 A. Johnsson, M. Durand-Dubief, Y. Xue-Franzén, M. Rönnerblad, K. Ekwall and A. Wright, HAT–HDAC interplay modulates global histone H3K14 acetylation in gene-coding regions during stress, *EMBO Rep.*, 2009, **10**, 1009–1014.
- 229 L. Liao, N. L. Alicea-Velázquez, L. Langbein, X. Niu, W. Cai, E. Cho, M. Zhang, C. B. Greer, Q. Yan, M. S. Cosgrove and H. Yang, High affinity binding of H3K14ac through collaboration of bromodomains 2, 4 and 5 is critical for the molecular and tumor suppressor functions of <scp>PBRM</scp> 1, *Mol. Oncol.*, 2019, **13**, 811–828.
- 230 K. Karmodiya, A. R. Krebs, M. Oulad-Abdelghani, H. Kimura and L. Tora, H3K9 and H3K14 acetylation co-occur at many gene regulatory elements, while H3K14ac marks a subset of inactive inducible promoters in mouse embryonic stem cells, *BMC Genomics*, 2012, **13**, 424.
- 231 E. Armenta-Castro, T. Reyes-Vallejo, D. Máximo-Sánchez, I. Herrera-Camacho, G. López-López, S. Reyes-Carmona, I. Conde-Rodríguez, I. Ramírez-Díaz, A. Aguilar-Lemarroy, L. F. Jave-Suárez, L. Milflores-Flores, G. Santos-Lopez, J. Reyes-Leyva and V. Vallejo-Ruiz, Histone H3K9 and H3K14 acetylation at the promoter of the *LGALS9* gene is associated with mRNA levels in cervical cancer cells, *FEBS Open Bio*, 2020, **10**, 2305–2315.
- 232 P. Ferrari and M. Strubin, Uncoupling histone turnover from transcription-associated histone H3 modifications, *Nucleic Acids Res.*, 2015, **43**, 3972–3985.
- 233 B. N. Albaugh and J. M. Denu, Catalysis by protein acetyltransferase Gcn5, *Biochim. Biophys. Acta - Gene Regul. Mech.*, 2021, **1864**, 194627.
- 234 B. K. Wagner, in *The Discovery and Utility of Chemical Probes in Target Discovery*, eds. P. Brennan and S. Vazquez Rodriguez, RSC, 2021, pp. 1–13.
- 235 C. H. Arrowsmith, J. E. Audia, C. Austin, J. Baell, J. Bennett, J. Blagg, C. Bountra, P. E. Brennan, P. J. Brown, M. E. Bunnage, C. Buser-Doepner, R. M. Campbell, A. J. Carter, P. Cohen, R. A. Copeland, B. Cravatt, J. L. Dahlin, D. Dhanak, A. M. Edwards, S. V. Frye, N. Gray, C. E. Grimshaw, D. Hepworth, T. Howe, K. V. M. Huber, J. Jin, S. Knapp, J. D. Kotz, R. G. Kruger, D. Lowe, M. M. Mader, B. Marsden, A. Mueller-Fahrnow, S. Müller, R. C. O'Hagan, J. P. Overington, D. R. Owen, S. H. Rosenberg, B. Roth, R. Ross, M. Schapira, S. L. Schreiber, B. Shoichet, M. Sundström, G. Superti-Furga, J. Taunton, L. Toledo-Sherman, C. Walpole, M. A. Walters, T. M. Willson, P. Workman, R. N. Young and W. J. Zuercher, The promise and peril of chemical probes, *Nat. Chem. Biol.*, 2015, **11**, 536–541.
- 236 S. V. Frye, The art of the chemical probe, *Nat. Chem. Biol.*, 2010, **6**, 159–161.
- 237 P. Cohen, Guidelines for the effective use of chemical inhibitors of protein function to understand their roles in cell regulation, *Biochem. J.*, 2010, **425**, 53–54.
- 238 A. M. Edwards, C. Bountra, D. J. Kerr and T. M. Willson, Open access chemical and clinical probes to support drug discovery, *Nat. Chem. Biol.*, 2009, **5**, 436–440.
- 239 Z. A. Knight and K. M. Shokat, Chemical Genetics: Where Genetics and Pharmacology Meet, *Cell*, 2007, **128**, 425–430.
- 240 M. Moustakim, P. G. K. Clark, D. A. Hay, D. J. Dixon and P. E. Brennan, Chemical probes

- and inhibitors of bromodomains outside the BET family, *Medchemcomm*, 2016, **7**, 2246–2264.
- 241 O. D. Lau, T. K. Kundu, R. E. Soccio, S. Ait-Si-Ali, E. M. Khalil, A. Vassilev, A. P. Wolffe, Y. Nakatani, R. G. Roeder and P. A. Cole, HATs off: Selective synthetic inhibitors of the histone acetyltransferases p300 and PCAF, *Mol. Cell*, 2000, **5**, 589–595.
- 242 K. Balasubramanyam, V. Swaminathan, A. Ranganathan and T. K. Kundu, Small molecule modulators of histone acetyltransferase p300, *J. Biol. Chem.*, 2003, **278**, 19134–19140.
- 243 K. Balasubramanyam, M. Altaf, R. A. Varier, V. Swaminathan, A. Ravindran, P. P. Sadhale and T. K. Kundu, Polyisoprenylated Benzophenone, Garcinol, a Natural Histone Acetyltransferase Inhibitor, Represses Chromatin Transcription and Alters Global Gene Expression, *J. Biol. Chem.*, 2004, **279**, 33716–33726.
- 244 K. Balasubramanyam, R. A. Varier, M. Altaf, V. Swaminathan, N. B. Siddappa, U. Ranga and T. K. Kundu, Curcumin, a novel p300/CREB-binding protein-specific inhibitor of acetyltransferase, represses the acetylation of histone/nonhistone proteins and histone acetyltransferase-dependent chromatin transcription, *J. Biol. Chem.*, 2004, **279**, 51163–51171.
- 245 Structural Genomic Consortium Website, <https://www.thesgc.org/>.
- 246 H. Wapenaar, P. E. Van Der Wouden, M. R. Groves, D. Rotili, A. Mai and F. J. Dekker, Enzyme kinetics and inhibition of histone acetyltransferase KAT8, *Eur. J. Med. Chem.*, 2015, **105**, 289–296.
- 247 M. Ghizzoni, A. Boltjes, C. De Graaf, H. J. Haisma and F. J. Dekker, Improved inhibition of the histone acetyltransferase PCAF by an anacardic acid derivative, *Bioorganic Med. Chem.*, 2010, **18**, 5826–5834.
- 248 M. Ghizzoni, J. Wu, T. Gao, H. J. Haisma, F. J. Dekker and Y. George Zheng, 6-alkylsalicylates are selective Tip60 inhibitors and target the acetyl-CoA binding site, *Eur. J. Med. Chem.*, 2012, **47**, 337–344.
- 249 G. Sbardella, S. Castellano, C. Vicidomini, D. Rotili, A. Nebbioso, M. Miceli, L. Altucci and A. Mai, Identification of long chain alkylidenemalonates as novel small molecule modulators of histone acetyltransferases, *Bioorganic Med. Chem. Lett.*, 2008, **18**, 2788–2792.
- 250 S. Castellano, C. Milite, A. Feoli, M. Viviano, A. Mai, E. Novellino, A. Tosco and G. Sbardella, Identification of Structural Features of 2-Alkylidene-1,3-Dicarbonyl Derivatives that Induce Inhibition and/or Activation of Histone Acetyltransferases KAT3B/p300 and KAT2B/PCAF, *ChemMedChem*, 2015, **10**, 144–157.
- 251 R. Modak, J. Basha, N. Bharathy, K. Maity, P. Mizar, A. V. Bhat, M. Vasudevan, V. K. Rao, W. K. Kok, N. Natesh, R. Taneja and T. K. Kundu, Probing p300/CBP associated factor (PCAF)-dependent pathways with a small molecule inhibitor, *ACS Chem. Biol.*, 2013, **8**, 1311–1323.
- 252 K. Mantelingu, B. A. A. Reddy, V. Swaminathan, A. H. Kishore, N. B. Siddappa, G. V. P. Kumar, G. Nagashankar, N. Natesh, S. Roy, P. P. Sadhale, U. Ranga, C. Narayana and T. K. Kundu, Specific Inhibition of p300-HAT Alters Global Gene Expression and Represses HIV Replication, *Chem. Biol.*, 2007, **14**, 645–657.
- 253 S. Fu and R. Kurzrock, Development of curcumin as an epigenetic agent, *Cancer*, 2010, **116**, 4670–4676.
- 254 J. H. Ko, S. G. Lee, W. M. Yang, J. Y. Um, G. Sethi, S. Mishra, M. K. Shanmugam and K. S. Ahn, The application of embelin for cancer prevention and therapy, *Molecules*, 2018, **23**, 621.
- 255 L. Stimson, M. G. Rowlands, Y. M. Newbatt, N. F. Smith, F. I. Raynaud, P. Rogers, V. Bavetsias, S. Gorsuch, M. Jarman, A. Bannister, T. Kouzarides, E. McDonald, P. Workman and G. W. Aherne, Isothiazolones as inhibitors of PCAF and p300 histone

- acetyltransferase activity, *Mol. Cancer Ther.*, 2005, **4**, 1521–1532.
- 256 S. D. Furdas, I. Hoffmann, D. Robaa, B. Herquel, W. Malinka, P. Wiątek, A. Akhtar, W. Sippl and M. Jung, Pyrido- and benzisothiazolones as inhibitors of histone acetyltransferases (HATs), *Medchemcomm*, 2014, **5**, 1856–1862.
- 257 J. M. Gajer, S. D. Furdas, A. Gründer, M. Gothwal, U. Heinicke, K. Keller, F. Colland, S. Fulda, H. L. Pahl, I. Fichtner, W. Sippl and M. Jung, Histone acetyltransferase inhibitors block neuroblastoma cell growth in vivo, *Oncogenesis*, 2015, **4**, e137.
- 258 E. M. Bowers, G. Yan, C. Mukherjee, A. Orry, L. Wang, M. A. Holbert, N. T. Crump, C. A. Hazzalin, G. Liszczak, H. Yuan, C. Larocca, S. A. Saldanha, R. Abagyan, Y. Sun, D. J. Meyers, R. Marmorstein, L. C. Mahadevan, R. M. Alani and P. A. Cole, Virtual ligand screening of the p300/CBP histone acetyltransferase: Identification of a selective small molecule inhibitor, *Chem. Biol.*, 2010, **17**, 471–482.
- 259 J. H. Shrimp, A. W. Sorum, J. M. Garlick, L. Guasch, M. C. Nicklaus and J. L. Meier, Characterizing the Covalent Targets of a Small Molecule Inhibitor of the Lysine Acetyltransferase P300, *ACS Med. Chem. Lett.*, 2016, **7**, 151–155.
- 260 H. Yang, C. E. Pinello, J. Luo, D. Li, Y. Wang, L. Y. Zhao, S. C. Jahn, S. A. Saldanha, J. Planck, K. R. Geary, H. Ma, B. K. Law, W. R. Roush, P. Hodder and D. Liao, Small-molecule inhibitors of acetyltransferase p300 identified by high-throughput screening are potent anticancer agents, *Mol. Cancer Ther.*, 2013, **12**, 610–620.
- 261 L. M. Lasko, C. G. Jakob, R. P. Edalji, W. Qiu, D. Montgomery, E. L. Digiammarino, T. M. Hansen, R. M. Risi, R. Frey, V. Manaves, B. Shaw, M. Algire, P. Hessler, L. T. Lam, T. Uziel, E. Faivre, D. Ferguson, F. G. Buchanan, R. L. Martin, M. Torrent, G. G. Chiang, K. Karukurichi, J. W. Langston, B. T. Weinert, C. Choudhary, P. De Vries, J. H. Van Drie, D. McElligott, E. Kesicki, R. Marmorstein, C. Sun, P. A. Cole, S. H. Rosenberg, M. R. Michaelides, A. Lai and K. D. Bromberg, Discovery of a selective catalytic p300/CBP inhibitor that targets lineage-specific tumours, *Nature*, 2017, **550**, 128–132.
- 262 Y. Yang, R. Zhang, Z. Li, L. Mei, S. Wan, H. Ding, Z. Chen, J. Xing, H. Feng, J. Han, H. Jiang, M. Zheng, C. Luo and B. Zhou, Discovery of Highly Potent, Selective, and Orally Efficacious p300/CBP Histone Acetyltransferases Inhibitors, *J. Med. Chem.*, 2020, **63**, 1337–1360.
- 263 M. Biel, A. Kretsovali, E. Karatzali, J. Papamatheakis and A. Giannis, Design, Synthesis, and Biological Evaluation of a Small-Molecule Inhibitor of the Histone Acetyltransferase Gcn5, *Angew. Chemie Int. Ed.*, 2004, **43**, 3974–3976.
- 264 F. Chimenti, B. Bizzarri, E. Maccioni, D. Secci, A. Bolasco, P. Chimenti, R. Fioravanti, A. Granese, S. Carradori, F. Tosi, P. Ballario, S. Vernarecci and P. Filetici, A novel histone acetyltransferase inhibitor modulating Gcn5 network: cyclopentylidene-[4-(4'-chlorophenyl)thiazol-2-yl]hydrazonone, *J. Med. Chem.*, 2009, **52**, 530–536.
- 265 D. Trisciuglio, Y. Ragazzoni, A. Pelosi, M. Desideri, S. Carradori, C. Gabellini, G. Maresca, R. Nescatelli, D. Secci, A. Bolasco, B. Bizzarri, C. Cavaliere, I. D'Agnano, P. Filetici, L. Ricci-Vitiani, M. G. Rizzo and D. Del Bufalo, CPTH6, a thiazole derivative, induces histone hypoacetylation and apoptosis in human leukemia cells, *Clin. Cancer Res.*, 2012, **18**, 475–486.
- 266 A. Mai, D. Rotili, D. Tarantino, P. Ornaghi, F. Tosi, C. Vicidomini, G. Sbardella, A. Nebbioso, M. Miceli, L. Altucci and P. Filetici, Small-molecule inhibitors of histone acetyltransferase activity: Identification and biological properties, *J. Med. Chem.*, 2006, **49**, 6897–6907.
- 267 H. Xiong, J. Han, J. Wang, W. Lu, C. Wang, Y. Chen, Fulin Lian, N. Zhang, Y. C. Liu, C. Zhang, H. Ding, H. Jiang, W. Lu, C. Luo and B. Zhou, Discovery of 1,8-acridinedione derivatives as novel GCN5 inhibitors via high throughput screening, *Eur. J. Med. Chem.*, 2018, **151**, 740–751.
- 268 H. Tao, J. Wang, W. Lu, R. Zhang, Y. Xie, Y. C. Liu, R. Liu, L. Yue, K. Chen, H. Jiang, Y. Zhang, X. Xu and C. Luo, Discovery of trisubstituted nicotinonitrile derivatives as novel human

- GCN5 inhibitors through AlphaScreen-based high throughput screening, *RSC Adv.*, 2019, **9**, 4917–4924.
- 269 A. Kumar, K. Bhowmick, K. S. Vikramdeo, N. Mondal, N. Subbarao and S. K. Dhar, Designing novel inhibitors against histone acetyltransferase (HAT: GCN5) of *Plasmodium falciparum*, *Eur. J. Med. Chem.*, 2017, **138**, 26–37.
- 270 M. W. Vetting, L. P. Luiz, M. Yu, S. S. Hegde, S. Magnet, S. L. Roderick and J. S. Blanchard, Structure and functions of the GNAT superfamily of acetyltransferases, *Arch. Biochem. Biophys.*, 2005, **433**, 212–226.
- 271 M. Moustakim, P. G. K Clark, L. Trulli, A. L. Fuentes de Arriba, M. T. Ehebauer, A. Chaikuad, E. J. Murphy, J. acqui Mendez-Johnson, D. Daniels, C.-F. D. Hou, Y.-H. Lin, J. R. Walker, R. Hui, H. Yang, L. Dorrell, C. M. Rogers, O. P. ctovia Monteiro, O. leg Fedorov, K. V M Huber, S. Knapp, J. ag Heer, D. J. arren Dixon, P. E. Brennan, M. Moustakim, D. L. uentes de Arriba, D. JDixon, D. GKClark, L. Trulli, D. T. Ehebauer, D. E. J Murphy, D. Chaikuad, P. rof DrS Knapp, J. Mendez-Johnson, D. D. -F Hou, Y. -H Lin, D. R. Alker, D. R. Hui, D. Y. Ang and D. Heer, Discovery of a PCAF Bromodomain Chemical Probe, *Angew. Chemie - Int. Ed.*, 2017, **56**, 827–831.
- 272 P. G. Humphreys, P. Bamborough, C. W. Chung, P. D. Craggs, L. Gordon, P. Grandi, T. G. Hayhow, J. Hussain, K. L. Jones, M. Lindon, A. M. Michon, J. F. Renaux, C. J. Suckling, D. F. Tough and R. K. Prinjha, Discovery of a potent, cell penetrant, and selective p300/CBP-associated factor (PCAF)/general control nonderepressible 5 (GCN5) bromodomain chemical probe, *J. Med. Chem.*, 2017, **60**, 695–709.
- 273 Z. I. Bassi, M. C. Fillmore, A. H. Miah, T. D. Chapman, C. Maller, E. J. Roberts, L. C. Davis, D. E. Lewis, N. W. Galwey, || Kirsty, E. Waddington, V. Parravicini, A. L. Macmillan-Jones, C. Gongora, P. G. Humphreys, I. Churcher, R. K. Prinjha and D. F. Tough, Modulating PCAF/GCN5 Immune Cell Function through a PROTAC Approach, *ACS Chem. Biol.*, 2018, **13**, 2862–2867.
- 274 M. Rahnasto-Rilla, J. Tyni and M. Lahtela-Kakkonen, in *Topics in Medicinal Chemistry*, Springer, 2020, vol. 33, pp. 55–92.
- 275 M. Pacholec, J. E. Bleasdale, B. Chrnyk, D. Cunningham, D. Flynn, R. S. Garofalo, D. Griffith, M. Griffor, P. Loulakis, B. Pabst, X. Qiu, B. Stockman, V. Thanabal, A. Varghese, J. Ward, J. Withka and K. Ahn, SRT1720, SRT2183, SRT1460, and resveratrol are not direct activators of SIRT1, *J. Biol. Chem.*, 2010, **285**, 8340–8351.
- 276 C. M. Grozinger, E. D. Chao, H. E. Blackwell, D. Moazed and S. L. Schreiber, Identification of a Class of Small Molecule Inhibitors of the Sirtuin Family of NAD-dependent Deacetylases by Phenotypic Screening, *J. Biol. Chem.*, 2001, **276**, 38837–38843.
- 277 A. M. Kalle, A. Mallika, J. Badiger, Alinakhi, P. Talukdar and Sachchidanand, Inhibition of SIRT1 by a small molecule induces apoptosis in breast cancer cells, *Biochem. Biophys. Res. Commun.*, 2010, **401**, 13–19.
- 278 J. S. Disch, G. Evindar, C. H. Chiu, C. A. Blum, H. Dai, L. Jin, E. Schuman, K. E. Lind, S. L. Belyanskaya, J. Deng, F. Coppo, L. Aquilani, T. L. Graybill, J. W. Cuzzo, S. Lavu, C. Mao, G. P. Vlasuk and R. B. Perni, Discovery of thieno[3,2-d]pyrimidine-6-carboxamides as potent inhibitors of SIRT1, SIRT2, and SIRT3, *J. Med. Chem.*, 2013, **56**, 3666–3679.
- 279 A. D. Napper, J. Hixon, T. McDonagh, K. Keavey, J. F. Pons, J. Barker, W. T. Yau, P. Amouzegh, A. Flegg, E. Hamelin, R. J. Thomas, M. Kates, S. Jones, M. A. Navia, J. O. Saunders, P. S. DiStefano and R. Curtis, Discovery of indoles as potent and selective inhibitors of the deacetylase SIRT1, *J. Med. Chem.*, 2005, **48**, 8045–8054.
- 280 S. D. Süssmuth, S. Haider, G. B. Landwehrmeyer, R. Farmer, C. Frost, G. Tripepi, C. A. Andersen, M. Di Bacco, C. Lamanna, E. Diodato, L. Massai, D. Diamanti, E. Mori, L. Magnoni, J. Dreyhaupt, K. Schiefele, D. Craufurd, C. Saft, M. Rudzinska, D. Ryglewicz, M. Orth, S. Brzozy, A. Baran, G. Pollio, R. Andre, S. J. Tabrizi, B. Darpo and G. Westerberg, An

- exploratory double-blind, randomized clinical trial with selisistat, a SirT1 inhibitor, in patients with Huntington's disease, *Br. J. Clin. Pharmacol.*, 2015, **79**, 465–476.
- 281 L. Serrano, G. Roper and A. Kawamura, in *The Discovery and Utility of Chemical Probes in Target Discovery*, eds. P. Brennan and S. Rodriguez Vazquez, The Royal Society of Chemistry, 2021, pp. 100–123.
- 282 T. A. F. Cardote and A. Ciulli, Cyclic and Macrocyclic Peptides as Chemical Tools To Recognise Protein Surfaces and Probe Protein-Protein Interactions, *ChemMedChem*, 2016, **11**, 787–794.
- 283 S. Shi, K. Liu, Y. Chen, S. Zhang, J. Lin, C. Gong, Q. Jin, X. J. Yang, R. Chen, Z. Ji and A. Han, Competitive Inhibition of Lysine Acetyltransferase 2B by a Small Motif of the Adenoviral Oncoprotein E1A, *J. Biol. Chem.*, 2016, **291**, 14363–14372.
- 284 J. Morimoto, Y. Hayashi and H. Suga, Discovery of macrocyclic peptides armed with a mechanism-based warhead: Isoform-selective inhibition of human deacetylase SIRT2, *Angew. Chemie - Int. Ed.*, 2012, **51**, 3423–3427.
- 285 K. G. Tanner, M. R. Langer, Y. Kim and J. M. Denu, Kinetic mechanism of the histone acetyltransferase GCN5 from yeast, *J. Biol. Chem.*, 2000, **275**, 22048–22055.
- 286 X. Zhao, D. Allison, B. Condon, F. Zhang, T. Gheyi, A. Zhang, S. Ashok, M. Russell, I. Macewan, Y. Qian, J. A. Jamison, J. G. Luz and J. Gately Luz, The 2.5 Å crystal structure of the SIRT1 catalytic domain bound to nicotinamide adenine dinucleotide (NAD⁺) and an indole (EX527 analogue) reveals a novel mechanism of histone deacetylase inhibition, *J. Med. Chem.*, 2013, **56**, 963–969.
- 287 L. Zeng, Q. Zhang, G. Gerona-Navarro, N. Moshkina and M. M. Zhou, Structural Basis of Site-Specific Histone Recognition by the Bromodomains of Human Coactivators PCAF and CBP/p300, *Structure*, 2008, **16**, 643–652.
- 288 S. Gadhia, J. H. Shrimp and J. L. Meier, *Histone Acetyltransferase Assays in Drug and Chemical Probe Discovery*, 2018.
- 289 C. E. Berndsen and J. M. Denu, Assays for mechanistic investigations of protein/histone acetyltransferases, *Methods*, 2005, **36**, 321–331.
- 290 R. C. Trievel, F. Y. Li and R. Marmorstein, Application of a fluorescent histone acetyltransferase assay to probe the substrate specificity of the human p300/CBP-associated factor, *Anal. Biochem.*, 2000, **287**, 319–328.
- 291 T. Gao, C. Yang and Y. G. Zheng, Comparative studies of thiol-sensitive fluorogenic probes for HAT assays, *Anal. Bioanal. Chem.*, 2013, **405**, 1361–1371.
- 292 Y. Kim, K. G. Tanner and J. M. Denu, A continuous, nonradioactive assay for histone acetyltransferases, *Anal. Biochem.*, 2000, **280**, 308–314.
- 293 S. Ait-Si-Ali, S. Ramirez, P. Robin, D. Trouche and A. Harel-Bellan, A rapid and sensitive assay for histone acetyl-transferase activity, *Nucleic Acids Res.*, 1998, **26**, 3869–3870.
- 294 H. Falk, T. Connor, H. Yang, K. J. Loft, J. L. Alcindor, G. Nikolakopoulos, R. N. Surjadi, J. D. Bentley, M. K. Hattarki, O. Dolezal, J. M. Murphy, B. J. Monahan, T. S. Peat, T. Thomas, J. B. Baell, J. P. Parisot and I. P. Street, An efficient high-throughput screening method for MYST family acetyltransferases, a new class of epigenetic drug targets, *J. Biomol. Screen.*, 2011, **16**, 1196–1205.
- 295 A. W. Sorum, J. H. Shrimp, A. M. Roberts, D. C. Montgomery, N. K. Tiwari, M. Lal-Nag, A. Simeonov, A. Jadhav and J. L. Meier, Microfluidic Mobility Shift Profiling of Lysine Acetyltransferases Enables Screening and Mechanistic Analysis of Cellular Acetylation Inhibitors, *ACS Chem. Biol.*, 2016, **11**, 734–741.
- 296 M. T. Borra and J. M. Denu, Quantitative Assays for Characterization of the Sir2 Family of NAD⁺-Dependent Deacetylases, *Methods Enzymol.*, 2003, **376**, 171–187.
- 297 D. Wegener, C. Hildmann, D. Riester and A. Schwienhorst, Improved fluorogenic histone deacetylase assay for high-throughput-screening applications, *Anal. Biochem.*, 2003, **321**,

- 202–208.
- 298 B. Heltweg, F. Dequiedt, E. Verdin and M. Jung, Nonisotopic substrate for assaying both human zinc and NAD⁺-dependent histone deacetylases, *Anal. Biochem.*, 2003, **319**, 42–48.
 - 299 Y. Liu, R. Gerber, J. Wu, T. Tsuruda and J. D. McCarter, High-throughput assays for sirtuin enzymes: A microfluidic mobility shift assay and a bioluminescence assay, *Anal. Biochem.*, 2008, **378**, 53–59.
 - 300 P. A. Marcotte, P. R. Richardson, J. Guo, L. W. Barrett, N. Xu, A. Gunasekera and K. B. Glaser, Fluorescence assay of SIRT protein deacetylases using an acetylated peptide substrate and a secondary trypsin reaction, *Anal. Biochem.*, 2004, **332**, 90–99.
 - 301 T. Mitiaë and J. S. McKay, Immunohistochemical Analysis of Acetylation, Proliferation, Mitosis, and Apoptosis in Tumor Xenografts Following Administration of a Histone Deacetylase Inhibitor—A Pilot Study, *Toxicol. Pathol.*, 2005, **33**, 792–799.
 - 302 D. Rauh, F. Fischer, M. Gertz, M. Lakshminarasimhan, T. Bergbrede, F. Aladini, C. Kambach, C. F. W. Becker, J. Zerweck, M. Schutkowski and C. Steegborn, An acetylome peptide microarray reveals specificities and deacetylation substrates for all human sirtuin isoforms, *Nat. Commun.*, 2013, **4**, 2327.
 - 303 M. Sankaranarayananpillai, W. P. Tong, D. S. Maxwell, A. Pal, J. Pang, W. G. Bornmann, J. G. Gelovani and S. M. Ronen, Detection of histone deacetylase inhibition by noninvasive magnetic resonance spectroscopy, *Mol. Cancer Ther.*, 2006, **5**, 1325–1334.
 - 304 K. Tanabe, J. Liu, D. Kato, H. Kurumizaka, K. Yamatsugu, M. Kanai and S. A. Kawashima, LC-MS/MS-based quantitative study of the acyl group- and site-selectivity of human sirtuins to acylated nucleosomes, *Sci. Rep.*, 2018, **8**, 1–11.
 - 305 Y. M. Kuo, R. A. Henry and A. J. Andrews, A quantitative multiplexed mass spectrometry assay for studying the kinetic of residue-specific histone acetylation, *Methods*, 2014, **70**, 127–133.
 - 306 P. T. Rye, L. E. Frick, C. C. Ozbal and W. A. Lamarr, Advances in Label-Free Screening Approaches for Studying Histone Acetyltransferases, *J. Biomol. Screen.*, 2011, **16**, 1186–1195.
 - 307 M. Karas, D. Bachmann and F. Hillenkamp, Influence of the Wavelength in High-Irradiance Ultraviolet Laser Desorption Mass Spectrometry of Organic Molecules, *Anal. Chem.*, 1985, **57**, 2935–2939.
 - 308 M. Karas, M. Glückmann and J. Schäfer, Ionization in matrix-assisted laser desorption/ionization: singly charged molecular ions are the lucky survivors, *J. Mass Spectrom.*, 2000, **35**, 1–12.
 - 309 K. Tanaka, H. Waki, Y. Ido, S. Akita, Y. Yoshida, T. Yoshida and T. Matsuo, Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry, *Rapid Commun. Mass Spectrom.*, 1988, **2**, 151–153.
 - 310 R. C. Beavis, B. T. Chait and K. G. Standing, Matrix-assisted laser-desorption mass spectrometry using 355 nm radiation, *Rapid Commun. Mass Spectrom.*, 1989, **3**, 436–439.
 - 311 R. Zenobi and R. Knochenmuss, Ion formation in maldi mass spectrometry, *Mass Spectrom. Rev.*, 1998, **17**, 337–366.
 - 312 M. L. Vestal, Modern MALDI time-of-flight mass spectrometry, *J. Mass Spectrom.*, 2009, **44**, 303–317.
 - 313 S. Hosseini and S. O. Martinez-Chapa, *Fundamentals of MALDI-TOF-MS Analysis*, Springer Verlag, 2017.
 - 314 B. A. Mamyrin, V. I. Karataev, D. V. Shmikk and V. A. Zagulin, The mass-reflectron, a new nonmagnetic time-of-flight mass spectrometer with high resolution, *J. Exp. Theor. Phys.*, 1973, **37**, 35.

- 315 N. Bradbury and R. Nielsen, Absolute Values of the Electron Mobility in Hydrogen, *Phys. Rev.*, 1936, **49**, 388–393.
- 316 A. El-Aneed, A. Cohen and J. Banoub, Mass Spectrometry, Review of the Basics: Electrospray, MALDI, and Commonly Used Mass Analyzers, *Appl. Spectrosc. Rev.*, 2009, **44**, 210–230.
- 317 W. S. Sawyer, N. Srikumar, J. Carver, P. Y. Chu, A. Shen, A. Xu, A. J. Williams, C. Spiess, C. Wu, Y. Liu and J. C. Tran, High-throughput antibody screening from complex matrices using intact protein electrospray mass spectrometry, *PNAS*, 2020, **117**, 9851–9856.
- 318 T. Bretschneider, C. Ozbal, M. Holstein, M. Winter, F. H. Buettner, S. Thamm, D. Bischoff and A. H. Luippold, RapidFire BLAZE-Mode Is Boosting ESI-MS Toward High-Throughput-Screening, *SLAS Technol.*, 2019, **24**, 386–393.
- 319 I. D. G. Campuzano, T. San Miguel, T. Rowe, D. Onea, V. J. Cee, T. Arvedson and J. D. McCarter, High-Throughput Mass Spectrometric Analysis of Covalent Protein-Inhibitor Adducts for the Discovery of Irreversible Inhibitors, *J. Biomol. Screen.*, 2016, **21**, 136–144.
- 320 K. Y. C. Fung, S. Askovic, F. Basile and M. W. Duncan, A simple and inexpensive approach to interfacing high-performance liquid chromatography and matrix-assisted laser desorption/ionization-time of flight-mass spectrometry, *Proteomics*, 2004, **4**, 3121–3127.
- 321 C. Haslam, J. Hellicar, A. Dunn, A. Fuetterer, N. Hardy, P. Marshall, R. Paape, M. Pemberton, A. Resemann and M. Leveridge, The Evolution of MALDI-TOF Mass Spectrometry toward Ultra-High-Throughput Screening: 1536-Well Format and beyond, *J. Biomol. Screen.*, 2016, **21**, 176–186.
- 322 K. Beeman, J. Baumgärtner, M. Laubenheimer, K. Hergesell, M. Hoffmann, U. Pehl, F. Fischer and J. C. Pieck, Integration of an In Situ MALDI-Based High-Throughput Screening Process: A Case Study with Receptor Tyrosine Kinase c-MET, *SLAS Discov.*, 2017, **22**, 1203–1210.
- 323 K. Guitot, S. Scarabelli, T. Drujon, G. Bolbach, M. Amoura, F. Burlina, A. Jeltsch, S. Sagan and D. Guianvarc’h, Label-free measurement of histone lysine methyltransferases activity by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, *Anal. Biochem.*, 2014, **456**, 25–31.
- 324 L. Wei, X. Cai, Z. Qi, L. Rong, B. Cheng and J. Fan, In vivo and in vitro characterization of TEV protease mutants, *Protein Expr. Purif.*, 2012, **83**, 157–163.
- 325 M. Fairhead and M. Howarth, Site-specific biotinylation of purified proteins using BirA, *Methods Mol. Biol.*, 2015, **1266**, 171–184.
- 326 C. Aslanidis and P. J. de Jong, Ligation-independent cloning of PCR products (LIC-PCR), *Nucleic Acids Res.*, 1990, **18**, 6069–6074.
- 327 K. Luger, A. W. Mäder, R. K. Richmond, D. F. Sargent and T. J. Richmond, Crystal structure of the nucleosome core particle at 2.8 Å resolution, *Nature*, 1997, **389**, 251–260.
- 328 B. Burat, J. Gonzalez, F. L. Sauvage, H. Aouad, H. Arnion, E. Pinault, P. Marquet and M. Essig, Sum of peak intensities outperforms peak area integration in iTRAQ protein expression measurement by LC-MS/MS using a TripleTOF 5600+ platform, *Biosci. Rep.*, 2019, **39**, 1–16.
- 329 H. Leemhuis, L. C. Packman, K. P. Nightingale and F. Hollfelder, The human histone acetyltransferase P/CAF is a promiscuous histone propionyltransferase, *ChemBioChem*, 2008, **9**, 499–503.
- 330 A. F. Kebede, A. Nieborak, L. Z. Shahidian, S. Le Gras, F. Richter, D. A. Gómez, M. P. Baltissen, G. Meszaros, H. De Fatima Magliarelli, A. Taudt, R. Margueron, M. Colomé-Tatché, R. Ricci, S. Daujat, M. Vermeulen, G. Mittler and R. Schneider, Histone propionylation is a mark of active chromatin, *Nat. Struct. Mol. Biol.*, 2017, **24**, 1048–1056.
- 331 G. R. Wagner, D. P. Bhatt, T. M. O’Connell, J. W. Thompson, L. G. Dubois, D. S. Backos, H.

- Yang, G. A. Mitchell, O. R. Ilkayeva, R. D. Stevens, P. A. Grimsrud and M. D. Hirschey, A Class of Reactive Acyl-CoA Species Reveals the Non-enzymatic Origins of Protein Acylation, *Cell Metab.*, 2017, **25**, 823-837.e8.
- 332 R. A. Copeland, *Evaluation of Enzyme Inhibitors in Drug Discovery*, Wiley, 2013.
- 333 J. Wu, N. Xie, Z. Wu, Y. Zhang and Y. G. Zheng, Bisubstrate Inhibitors of the MYST HATs Esa1 and Tip60, *Bioorganic Med. Chem.*, 2009, **17**, 1381–1386.
- 334 W. C. Hallows, S. Lee and J. M. Denu, Sirtuins deacetylate and activate mammalian acetyl-CoA synthetases, *PNAS*, 2006, **103**, 10230–10235.
- 335 J. L. Feldman, K. E. Dittenhafer-Reed, N. Kudo, J. N. Thelen, A. Ito, M. Yoshida and J. M. Denu, Kinetic and Structural Basis for Acyl-Group Selectivity and NAD⁺ Dependence in Sirtuin-Catalyzed Deacetylation, *Biochemistry*, 2015, **54**, 22.
- 336 D. G. Fatkins, A. D. Monnot and W. Zheng, N ϵ -Thioacetyl-lysine: A multi-facet functional probe for enzymatic protein lysine N ϵ -deacetylation, *Bioorganic Med. Chem. Lett.*, 2006, **16**, 3651–3656.
- 337 B. C. Smith and J. M. Denu, Sir2 deacetylases exhibit nucleophilic participation of acetyl-lysine in NAD⁺ cleavage, *J. Am. Chem. Soc.*, 2007, **129**, 5802–5803.
- 338 L. J. Walport, R. J. Hopkinson, R. Chowdhury, R. Schiller, W. Ge, A. Kawamura and C. J. Schofield, Arginine demethylation is catalysed by a subset of JmJc histone lysine demethylases, *Nat. Commun.*, 2016, **7**, 11974.
- 339 J. Huang, D. Wan, J. Li, H. Chen, K. Huang and L. Zheng, Histone acetyltransferase PCAF regulates inflammatory molecules in the development of renal injury, *Epigenetics*, 2015, **10**, 62–72.
- 340 Y.-M. Kuo and A. J. Andrews, Quantitating the Specificity and Selectivity of Gcn5-Mediated Acetylation of Histone H3, *PLoS One*, 2013, **8**, 54896.
- 341 D. C. Montgomery, A. W. Sorum, L. Guasch, M. C. Nicklaus and J. L. Meier, Metabolic Regulation of Histone Acetyltransferases by Endogenous Acyl-CoA Cofactors, *Chem. Biol.*, 2015, **22**, 1030–1039.
- 342 T. Van Den Bosch, A. Boichenko, N. G. J. Leus, M. E. Ourailidou, H. Wapenaar, D. Rotili, A. Mai, A. Imhof, R. Bischoff, H. J. Haisma and F. J. Dekker, The histone acetyltransferase p300 inhibitor C646 reduces pro-inflammatory gene expression and inhibits histone deacetylases, *Biochem. Pharmacol.*, 2016, **102**, 130–140.
- 343 C. Milite, S. Castellano, R. Benedetti, A. Tosco, C. Ciliberti, C. Vicidomini, L. Bouilly, G. Franci, L. Altucci, A. Mai and G. Sbardella, Modulation of the activity of histone acetyltransferases by long chain alkylidenemalonates (LoCAMs), *Bioorganic Med. Chem.*, 2011, **19**, 3690–3701.
- 344 F. Spallotta, C. Cencioni, S. Straino, G. Sbardella, S. Castellano, M. C. Capogrossi, F. Martelli and C. Gaetano, Enhancement of lysine acetylation accelerates wound repair, *Commun. Integr. Biol.*, 2013, **6**, e25466.
- 345 H. Xu, Q. Zhou, Q. Yi, B. Tan, J. Tian, X. Chen, Y. Wang, X. Yu and J. Zhu, Islet-1 synergizes with Gcn5 to promote MSC differentiation into cardiomyocytes, *Sci. Rep.*, 2020, **10**, 1–12.
- 346 F. Aquea, T. Timmermann and A. Herrera-Vásquez, Chemical inhibition of the histone acetyltransferase activity in *Arabidopsis thaliana*, *Biochem. Biophys. Res. Commun.*, 2017, **483**, 664–668.
- 347 M. G. Acker and D. S. Auld, Considerations for the design and reporting of enzyme assays in high-throughput screening applications, *Perspect. Sci.*, 2014, **1**, 56–73.
- 348 K. Huynh and C. L. Partch, Analysis of protein stability and ligand interactions by thermal shift assay, *Curr. Protoc. protein Sci.*, 2015, **79**, 28.9.1-28.9.14.
- 349 F. H. Niesen, H. Berglund and M. Vedadi, The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability, *Nat. Protoc.*, 2007, **2**, 2212–

- 2221.
- 350 S. Chung, S. Kim, M. Son, M. Kim, E. Koh, S. Shin, C. Park and H.-S. Kim, Inhibition of p300/CBP-Associated Factor Attenuates Renal Tubulointerstitial Fibrosis through Modulation of NF- κ B and Nrf2, *Int. J. Mol. Sci.*, 2019, **20**, 1554.
- 351 J. E. Herrera, M. Bergel, X.-J. Yang, Y. Nakatani and M. Bustin, The Histone Acetyltransferase Activity of Human GCN5 and PCAF Is Stabilized by Coenzymes, *J. Biol. Chem.*, 1997, **272**, 27253–27258.
- 352 J. E. Morgan and S. F. Greer, Pulling a Ligase out of a (HAT): pCAF Mediates Ubiquitination of the Class II Transactivator, *Int. J. Cell Biol.*, 2017, 8093813.
- 353 A. Papantonis, V. N. Uversky, O. Sorokina, S. K. Banerjee, S. Aggarwal, N. Chandra Talukdar and A. Kumar Yadav, Post-translational Modification Crosstalk and Hotspots in Sirtuin Interactors Implicated in Cardiovascular Diseases, *Front. Genet.*, 2020, **11**, 1–15.
- 354 S. Nagarajan and S. A. Johnsen, in *Chromatin Signaling and Diseases*, Elsevier Inc., 2016, pp. 217–239.
- 355 J. A. Latham and S. Y. R. Dent, Cross-regulation of histone modifications, *Nat. Struct. Mol. Biol.*, 2007, **14**, 1017–1024.
- 356 K. E. Gardner, C. D. Allis and B. D. Strahl, Operating on chromatin, a colorful language where context matters, *J. Mol. Biol.*, 2011, **409**, 36–46.
- 357 A. L. Garske, G. Craciun and J. M. Denu, A Combinatorial H4 Tail Library for Exploring the Histone Code, *Biochemistry*, 2008, **47**, 8094–8102.
- 358 D. Aparicio Pelaz, Z. Yerkesh, S. Kirchg   ner, H. Mahler, V. Kharchenko, D. Azhibek, M. Jaremko, H. D. Mootz,  . Jaremko, D. Schwarzer and W. Fischle, Examining histone modification crosstalk using immobilized libraries established from ligation-ready nucleosomes, *Chem. Sci.*, 2020, **11**, 9218–9225.
- 359 M. B. Chandrasekharan, F. Huang and Z. W. Sun, Decoding the trans-histone crosstalk: Methods to analyze H2B ubiquitination, H3 methylation and their regulatory factors, *Methods*, 2011, **54**, 304–314.
- 360 X. Guan, N. Rastogi, M. R. Parthun and M. A. Freitas, Discovery of histone modification crosstalk networks by stable isotope labeling of amino acids in cell culture mass spectrometry (SILAC MS), *Mol. Cell. Proteomics*, 2013, **12**, 2048–2059.
- 361 Z. Lu, Z. Cheng, Y. Zhao and S. L. Volchenboum, Bioinformatic analysis and post-translational modification crosstalk prediction of lysine acetylation, *PLoS One*, 2011, **6**, e28228.
- 362 Y. Huang, B. Xu, X. Zhou, Y. Li, M. Lu, R. Jiang and T. Li, Systematic characterization and prediction of post-translational modification cross-talk, *Mol. Cell. Proteomics*, 2015, **14**, 761–770.
- 363 Y. Qiu, L. Liu, C. Zhao, C. Han, F. Li, J. Zhang, Y. Wang, G. Li, Y. Mei, M. Wu, J. Wu and Y. Shi, Combinatorial readout of unmodified H3R2 and acetylated H3K14 by the tandem PHD finger of MOZ reveals a regulatory mechanism for HOXA9 transcription, *Genes Dev.*, 2012, **26**, 1376–1391.
- 364 B. O. Keller, J. Sui, A. B. Young and R. M. Whittall, Interferences and contaminants encountered in modern mass spectrometry, *Anal. Chim. Acta*, 2008, **627**, 71–81.
- 365 H. Tong, D. Bell, K. Tabei and M. M. Siegel, Automated data massaging, interpretation, and e-mailing modules for high throughput open access mass spectrometry, *J. Am. Soc. Mass Spectrom.*, 1999, **10**, 1174–1187.
- 366 M. J. Rardin, Rapid Assessment of Contaminants and Interferences in Mass Spectrometry Data Using Skyline, *J. Am. Soc. Mass Spectrom.*, 2018, **29**, 1327–1330.
- 367 M. Zhao, R. Geng, X. Guo, R. Yuan, X. Zhou, Y. Zhong, Y. Huo, M. Zhou, Q. Shen, Y. Li, W. Zhu and J. Wang, PCAF/GCN5-Mediated Acetylation of RPA1 Promotes Nucleotide Excision Repair, *Cell Rep.*, 2017, **20**, 1997–2009.

- 368 A. Zippo, R. Serafini, M. Rocchigiani, S. Pennacchini, A. Krepelova and S. Oliviero, Histone Crosstalk between H3S10ph and H4K16ac Generates a Histone Code that Mediates Transcription Elongation, *Cell*, 2009, **138**, 1122–1136.
- 369 D. H. Khan, S. Healy, S. He, D. Lichtensztejn, L. Klewes, K. L. Sharma, V. Lau, S. Mai, G. P. Delcuve and J. R. Davie, Mitogen-induced distinct epialleles are phosphorylated at either H3S10 or H3S28, depending on H3K27 acetylation, *Mol. Biol. Cell*, 2017, **28**, 817–824.
- 370 Y. S. Jeong, S. Cho, J. S. Park, Y. Ko and Y. K. Kang, Phosphorylation of serine-10 of histone H3 shields modified lysine-9 selectively during mitosis, *Genes to Cells*, 2010, **15**, 181–192.
- 371 A. Sawicka and C. Seiser, Histone H3 phosphorylation - A versatile chromatin modification for different occasions, *Biochimie*, 2012, **94**, 2193–2201.
- 372 W. A. Kellner, E. Ramos, K. Van Bortle, N. Takenaka and V. G. Corces, Genome-wide phosphoacetylation of histone H3 at Drosophila enhancers and promoters, *Genome Res.*, 2012, **22**, 1081–1088.
- 373 S. Winter, W. Fischle and C. Seiser, Cell Cycle Modulation of 14-3-3 interaction with phosphorylated histone H3 by combinatorial modification patterns, *Cell Cycle*, 2008, **7**, 1336–1342.
- 374 C. M. Barber, F. B. Turner, Y. Wang, K. Hagstrom, S. D. Taverna, S. Mollah, B. Ueberheide, B. J. Meyer, D. F. Hunt, P. Cheung and C. D. Allis, The enhancement of histone H4 and H2A serine 1 phosphorylation during mitosis and S-phase is evolutionarily conserved, *Chromosoma*, 2004, **112**, 360–371.
- 375 H. Polioudaki, Y. Markaki, N. Kourmouli, G. Dialynas, P. A. Theodoropoulos, P. B. Singh and S. D. Georgatos, Mitotic phosphorylation of histone H3 at threonine 3, *FEBS Lett.*, 2004, **560**, 39–44.
- 376 Z. Cao, T. Xu, X. Tong, D. Zhang, C. Liu, Y. Wang, D. Gao, L. Luo, L. Zhang, Y. Li and Y. Zhang, HASPIN kinase mediates histone deacetylation to regulate oocyte meiotic maturation in pigs, *Reproduction*, 2019, **157**, 501–510.
- 377 J. Cao, W. Lv, L. Wang, J. Xu, P. Yuan, S. Huang, Z. He and J. Hu, Ricolinostat (ACY-1215) suppresses proliferation and promotes apoptosis in esophageal squamous cell carcinoma via miR-30d/PI3K/AKT/mTOR and ERK pathways, *Cell Death Dis.*, 2018, **9**, 1–11.
- 378 H. Li, C. Cuenin, R. Murr, Z.-Q. Wang and Z. Herceg, HAT cofactor Trrap regulates the mitotic checkpoint by modulation of Mad1 and Mad2 expression, *EMBO J.*, 2004, **23**, 4824–4834.
- 379 B. Zhao, W. Xu, B. Rong, G. Chen, X. Ye, R. Dai, W. Li, J. Chen, J. Cai, L. Song, Z. Q. Luo, R. Zeng, Y. Shi, J. D. J. Han and F. Lan, H3K14me3 genomic distributions and its regulation by KDM4 family demethylases, *Cell Res.*, 2018, **28**, 1118–1120.
- 380 H. R. Jung, D. Pasini, K. Helin and O. N. Jensen, Quantitative mass spectrometry of histones H3.2 and H3.3 in Suz12-deficient mouse embryonic stem cells reveals distinct, dynamic post-translational modifications at Lys-27 and Lys-36, *Mol. Cell. Proteomics*, 2010, **9**, 838–850.
- 381 D. Bonenfant, H. Towbin, M. Coulot, P. Schindler, D. R. Mueller and J. van Oostrum, Analysis of dynamic changes in post-translational modifications of human histones during cell cycle by mass spectrometry, *Mol. Cell. Proteomics*, 2007, **6**, 1917–1932.
- 382 B. A. Garcia, S. B. Hake, R. L. Diaz, M. Kauer, S. A. Morris, J. Recht, J. Shabanowitz, N. Mishra, B. D. Strahl, C. D. Allis and D. F. Hunt, Organismal differences in post-translational modifications in histones H3 and H4, *J. Biol. Chem.*, 2007, **282**, 7641–7655.
- 383 S. Sidoli, C. Lu, M. Coradin, X. Wang, K. R. Karch, C. Ruminowicz and B. A. Garcia, Metabolic labeling in middle-down proteomics allows for investigation of the dynamics of the histone code, *Epigenetics Chromatin*, 2017, **10**, 34.
- 384 K. Taniue, T. Hayashi, Y. Kamoshida, A. Kurimoto, Y. Takeda, L. Negishi, K. Iwasaki, Y. Kawamura, N. Goshima and T. Akiyama, UHRF1-KAT7-mediated regulation of TUSC3

- expression via histone methylation/acetylation is critical for the proliferation of colon cancer cells, *Oncogene*, 2020, **39**, 1018–1030.
- 385 A. E. Ringel, A. M. Cieniewicz, S. D. Taverna and C. Wolberger, Nucleosome competition reveals processive acetylation by the SAGA HAT module, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, E5461–E5470.
- 386 T. Kouzarides, Chromatin Modifications and Their Function, *Cell*, 2007, **128**, 693–705.
- 387 H. Benyamini and A. Friedler, Using peptides to study protein-protein interactions, *Future Med. Chem.*, 2010, **2**, 989–1003.
- 388 J. D. Lear, Z. R. Wasserman and W. F. Degrado, Synthetic amphiphilic peptide models for protein ion channels, *Science (80-.)*, 1988, **240**, 1177–1181.
- 389 M. H. Kuo, J. E. Brownell, R. E. Sobel, T. A. Ranalli, R. G. Cook, D. G. Edmondson, S. Y. Roth and C. D. Allis, Transcription-linked acetylation by Gcn5p of histones H3 and H4 at specific lysines, *Nature*, 1996, **383**, 269–272.
- 390 R. Chowdhury, J. I. Candela-Lena, M. C. Chan, D. J. Greenald, K. K. Yeoh, Y. M. Tian, M. A. McDonough, A. Tumber, N. R. Rose, A. Conejo-Garcia, M. Demetriades, S. Mathavan, A. Kawamura, M. K. Lee, F. Van Eeden, C. W. Pugh, P. J. Ratcliffe and C. J. Schofield, Selective small molecule probes for the hypoxia inducible factor (HIF) Prolyl Hydroxylases, *ACS Chem. Biol.*, 2013, **8**, 1488–1496.
- 391 P. Jaakkola, D. R. Mole, Y. M. Tian, M. I. Wilson, J. Gielbert, S. J. Gaskell, A. Von Kriegsheim, H. F. Hebestreit, M. Mukherji, C. J. Schofield, P. H. Maxwell, C. W. Pugh and P. J. Ratcliffe, Targeting of HIF- α to the von Hippel-Lindau ubiquitylation complex by O₂-regulated prolyl hydroxylation, *Science (80-.)*, 2001, **292**, 468–472.
- 392 A. L. Harris, Hypoxia - A key regulatory factor in tumour growth, *Nat. Rev. Cancer*, 2002, **2**, 38–47.
- 393 G. L. Semenza, Targeting HIF-1 for cancer therapy, *Nat. Rev. Cancer*, 2003, **3**, 721–732.
- 394 S. H. Kim, J. A. Park, J. H. Kim, J. W. Lee, J. H. Seo, B. K. Jung, K. H. Chun, J. W. Jeong, M. K. Bae and K. W. Kim, Characterization of ARD1 variants in mammalian cells, *Biochem. Biophys. Res. Commun.*, 2006, **340**, 422–427.
- 395 H. Geng, Q. Liu, C. Xue, L. L. David, T. M. Beer, G. V Thomas, M.-S. Dai and D. Z. Qian, HIF1 Protein Stability Is Increased by Acetylation at Lysine 709, *J. Biol. Chem.*, 2012, **287**, 35496–35505.
- 396 S. F. Leiser and M. Kaeberlein, A Role for SIRT1 in the Hypoxic Response, *Mol. Cell*, 2010, **38**, 779–780.
- 397 S. K. Bharti and R. Roy, Quantitative ¹H NMR spectroscopy, *Trends Anal. Chem.*, 2012, **35**, 5–26.
- 398 M. Hałasa, A. Wawruszak, A. Przybyszewska, A. Jaruga, M. Guz, J. Kałafut, A. Stepulak and M. Cybulski, H3K18Ac as a Marker of Cancer Progression and Potential Target of Anti-Cancer Therapy, *Cells*, 2019, **8**, 485.
- 399 Y. Xue-Franzén, J. Henriksson, T. R. Bürglin and A. P. H. Wright, Distinct roles of the Gcn5 histone acetyltransferase revealed during transient stress-induced reprogramming of the genome, *BMC Genomics*, 2013, **14**, 479.
- 400 Z. Cheng, X. Li, S. Hou, Y. Wu, Y. Sun and B. Liu, K-Ras-ERK1/2 accelerates lung cancer cell development via mediating H3^{K18ac} through the MDM2-GCN5-SIRT7 axis, *Pharm. Biol.*, 2019, **57**, 701–709.
- 401 R. Marmorstein, Protein modules that manipulate histone tails for chromatin regulation, *Nat. Rev. Mol. Cell Biol.*, 2001, **2**, 422–432.
- 402 L. M. Iakoucheva, P. Radivojac, C. J. Brown, T. R. O'Connor, J. G. Sikes, Z. Obradovic and A. K. Dunker, The importance of intrinsic disorder for protein phosphorylation, *Nucleic Acids Res.*, 2004, **32**, 1037–1049.
- 403 H. J. Dyson and P. E. Wright, Role of intrinsic protein disorder in the function and

- interactions of the transcriptional coactivators CREB-binding Protein (CBP) and p300, *J. Biol. Chem.*, 2016, **291**, 6714–6722.
- 404 H. Xie, S. Vucetic, L. M. Iakoucheva, C. J. Oldfield, A. K. Dunker, Z. Obradovic and V. N. Uversky, Functional anthology of intrinsic disorder. 3. Ligands, post-translational modifications, and diseases associated with intrinsically disordered proteins, *J. Proteome Res.*, 2007, **6**, 1917–1932.
- 405 S. E. Wilkins, M. I. Abboud, R. L. Hancock and C. J. Schofield, Targeting Protein-Protein Interactions in the HIF System, *ChemMedChem*, 2016, **11**, 773–786.
- 406 S. A. Dames, M. Martinez-Yamout, R. N. De Guzman, H. Jane Dyson and P. E. Wright, Structural basis for Hif-1 α /CBP recognition in the cellular hypoxic response, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 5271–5276.
- 407 C. Dai and W. Gu, P53 post-translational modification: Deregulated in tumorigenesis, *Trends Mol. Med.*, 2010, **16**, 528–536.
- 408 R. B. Merrifield, Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide, *J. Am. Chem. Soc.*, 1963, **85**, 2149–2154.
- 409 T. A. F. Cardote and A. Ciulli, Cyclic and Macrocyclic Peptides as Chemical Tools To Recognise Protein Surfaces and Probe Protein-Protein Interactions, *ChemMedChem*, 2016, **11**, 787–794.
- 410 R. Obexer, L. J. Walport and H. Suga, Exploring sequence space: harnessing chemical and biological diversity towards new peptide leads, *Curr. Opin. Chem. Biol.*, 2017, **38**, 52–61.
- 411 J. A. Wells and C. L. McClendon, Reaching for high-hanging fruit in drug discovery at protein-protein interfaces., *Nature*, 2007, **450**, 1001–9.
- 412 H. Bruzzoni-Giovanelli, V. Alezra, N. Wolff, C. Z. Dong, P. Tuffery and A. Rebollo, Interfering peptides targeting protein–protein interactions: the next generation of drugs?, *Drug Discov. Today*, 2018, **23**, 272–285.
- 413 M. Trellet, A. S. J. Melquiond and A. M. J. J. Bonvin, in *Computational Peptidology*, Humana Press, New York, 2015, pp. 221–239.
- 414 M. Ciemny, M. Kurcinski, K. Kamel, A. Kolinski, N. Alam, O. Schueler-Furman and S. Kmiecik, Protein–peptide docking: opportunities and challenges, *Drug Discov. Today*, 2018, **23**, 1530–1537.
- 415 L. Nevola and E. Giralt, Modulating protein-protein interactions: the potential of peptides, *Chem. Commun.*, 2015, **51**, 3302–15.
- 416 J. Ottl, L. Leder, J. V. Schaefer and C. E. Dumelin, Encoded library technologies as integrated lead finding platforms for drug discovery, *Molecules*, 2019, **24**, 1629.
- 417 A. Alex, D. S. Millan, M. Perez, F. Wakenhut and G. A. Whitlock, Intramolecular hydrogen bonding to improve membrane permeability and absorption in beyond rule of five chemical space, *Medchemcomm*, 2011, **2**, 669–674.
- 418 M. Fouché, M. Schäfer, J. Berghausen, S. Desrayaud, M. Blatter, P. Piéchon, I. Dix, A. Martingarcia and H. J. Roth, Design and Development of a Cyclic Decapeptide Scaffold with Suitable Properties for Bioavailability and Oral Exposure, *ChemMedChem*, 2016, **11**, 1048–1059.
- 419 A. Henninot, J. C. Collins and J. M. Nuss, The Current State of Peptide Drug Discovery: Back to the Future?, *J. Med. Chem.*, 2018, **61**, 1382–1414.
- 420 D. Faulds, K. L. Goa and P. Benfield, Cyclosporin, a review of its pharmacodynamic and pharmacokinetic properties and therapeutic use in immunoregulatory Disorders, *Drugs*, 1993, **45**, 953–1040.
- 421 P. M. Cromm, J. Spiegel and T. N. Grossmann, Hydrocarbon Stapled Peptides as Modulators of Biological Function, *ACS Chem. Biol.*, 2015, **10**, 1362–1375.
- 422 F. Bernal, A. F. Tyler, S. J. Korsmeyer, L. D. Walensky and G. L. Verdine, Reactivation of the p53 tumor suppressor pathway by a stapled p53 peptide, *J. Am. Chem. Soc.*, 2007,

- 129, 2456–2457.
- 423 C. Heinis and G. Winter, Encoded libraries of chemically modified peptides, *Curr. Opin. Chem. Biol.*, 2015, **26**, 89–98.
- 424 A. Galán, L. Comor, A. Horvatić, J. Kuleš, N. Guillemin, V. Mrljak and M. Bhide, Library-based display technologies: Where do we stand?, *Mol. Biosyst.*, 2016, **12**, 2342–2358.
- 425 K. Lam, S. Salmon, E. Hersh, V. Hruby, W. Kazmierski and R. Knapp, A new type of synthetic peptide library for identifying ligand-binding activity, *Nature*, 1991, **354**, 82–84.
- 426 R. A. Houghten, C. Pinilla, S. E. Blondelle, J. R. Appel, C. T. Dooley and J. H. Cuervo, Generation and use of synthetic peptide combinatorial libraries for basic research and drug discovery, *Nature*, 1991, **354**, 84–86.
- 427 J. Shendure and H. Ji, Next-generation DNA sequencing, *Nat. Biotechnol.*, 2008, **26**, 1135–1145.
- 428 Z. Zhu, A. Shaginian, L. C. Grady, T. O’keeffe, X. E. Shi, C. P. Davie, G. L. Simpson, J. A. Messer, G. Evindar, R. N. Bream, P. P. Thansandote, N. R. Prentice, A. M. Mason and S. Pal, Design and Application of a DNA-Encoded Macrocyclic Peptide Library, *ACS Chem. Biol.*, 2018, **13**, 53–59.
- 429 B. N. Tse, T. M. Snyder, Y. Shen and D. R. Liu, Translation of DNA into a Library of 13 000 Synthetic Small-Molecule Macrocycles Suitable for in Vitro Selection, *J. Am. Chem. Soc.*, 2008, **130**, 15611–15626.
- 430 N. Yao, W. Xiao, X. Wang, J. Marik, S. H. Park and K. S. Lam, Discovery of Targeting Ligands for Breast Cancer Cells Using the One-Bead One-Compound Combinatorial Method, *J. Med. Chem.*, 2009, **52**, 126–133.
- 431 R. Liu, J. Marik and K. S. Lam, A novel peptide-based encoding system for ‘one-bead one-compound’ peptidomimetic and small molecule combinatorial libraries, *J. Am. Chem. Soc.*, 2002, **124**, 7678–7680.
- 432 R. Liu and K. S. Lam, Automatic Edman microsequencing of peptides containing multiple unnatural amino acids, *Anal. Biochem.*, 2001, **295**, 9–16.
- 433 T. Kodadek and K. Bachhawat-Sikder, Optimized protocols for the isolation of specific protein-binding peptides or peptoids from combinatorial libraries displayed on beads, *Mol. Biosyst.*, 2006, **2**, 25–35.
- 434 G. P. Smith, Filamentous Fusion Phage: Novel Expression Vectors That Display Cloned Antigens on the Virion Surface, *Science (80-.)*, 1985, **228**, 1315–1317.
- 435 G. P. Smith and V. A. Petrenko, Phage Display, *Chem. Rev.*, 1997, **97**, 391–410.
- 436 C. P. Scott, E. Abel-Santos, M. Wall, D. C. Wahnnon and S. J. Benkovic, Production of cyclic peptides and proteins in vivo, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 13638–13643.
- 437 A. Tavassoli, SICLOPPS cyclic peptide libraries in drug discovery, *Curr. Opin. Chem. Biol.*, 2017, **38**, 30–35.
- 438 A. Tavassoli and S. J. Benkovic, Genetically Selected Cyclic-Peptide Inhibitors of AICAR Transformylase Homodimerization, *Angew. Chemie Int. Ed.*, 2005, **44**, 2760–2763.
- 439 R. W. Roberts and J. W. Szostak, RNA-peptide fusions for the in vitro selection of peptides and proteins, *Proc. Natl. Acad. Sci. U. S. A.*, 1997, **94**, 12297–12302.
- 440 K. Josephson, A. Ricardo and J. W. Szostak, mRNA display: from basic principles to macrocycle drug discovery, *Drug Discov. Today*, 2014, **19**, 388–399.
- 441 H. Saito, D. Kourouklis and H. Suga, An in vitro evolved precursor tRNA with aminoacylation activity, *EMBO J.*, 2001, **20**, 1797–1806.
- 442 Y. Goto, T. Katoh and H. Suga, Flexizymes for genetic code reprogramming, *Nat. Protoc.*, 2011, **6**, 779–790.
- 443 T. Passioura and H. Suga, A RaPID way to discover nonstandard macrocyclic peptide modulators of drug targets, *Chem. Commun*, 2017, **53**, 1931–1940.
- 444 C. J. Hipolito and H. Suga, Ribosomal production and in vitro selection of natural product-

- like peptidomimetics: The FIT and RaPID systems, *Curr. Opin. Chem. Biol.*, 2012, **16**, 196–203.
- 445 T. E. McAllister, T. L. Yeh, M. I. Abboud, I. K. H. Leung, E. S. Hookway, O. N. F. King, B. Bhushan, S. T. Williams, R. J. Hopkinson, M. Münzel, N. D. Loik, R. Chowdhury, U. Oppermann, T. D. W. Claridge, Y. Goto, H. Suga, C. J. Schofield and A. Kawamura, Non-competitive cyclic peptides for targeting enzyme-substrate complexes, *Chem. Sci.*, 2018, **9**, 4569–4578.
- 446 K. Nishio, R. Belle, T. Katoh, A. Kawamura, T. Sengoku, K. Hanada, N. Ohsawa, M. Shirouzu, S. Yokoyama and H. Suga, Thioether Macrocyclic Peptides Selected against TET1 Compact Catalytic Domain Inhibit TET1 Catalytic Activity, *ChemBioChem*, 2018, **19**, 979–985.
- 447 A. Kawamura, M. Münzel, T. Kojima, C. Yapp, B. Bhushan, Y. Goto, A. Tumber, T. Katoh, O. N. F. King, T. Passioura, L. J. Walport, S. B. Hatch, S. Madden, S. Müller, P. E. Brennan, R. Chowdhury, R. J. Hopkinson, H. Suga and C. J. Schofield, Highly selective inhibition of histone demethylases by de novo macrocyclic peptides, *Nat. Commun.*, 2017, **8**, 1–10.
- 448 K. Yamagata, Y. Goto, H. Nishimasu, J. Morimoto, R. Ishitani, N. Dohmae, N. Takeda, R. Nagai, I. Komuro, H. Suga and O. Nureki, Structural basis for potent inhibition of SIRT2 deacetylase by a macrocyclic peptide inducing dynamic structural change, *Structure*, 2014, **22**, 345–352.
- 449 J. M. Rogers, T. Passioura and H. Suga, Nonproteinogenic deep mutational scanning of linear and cyclic peptides, *PNAS*, 2018, **115**, 10959–10964.
- 450 C. J. Hipolito, Y. Tanaka, T. Katoh, O. Nureki and H. Suga, A macrocyclic peptide that serves as a cocrystallization ligand and inhibits the function of a MATE family transporter, *Molecules*, 2013, **18**, 10514–10530.
- 451 E. M. Farrell and G. Alexandre, Bovine serum albumin further enhances the effects of organic solvents on increased yield of polymerase chain reaction of GC-rich templates, *BMC Res. Notes*, 2012, **5**, 257.
- 452 J. Brownie, S. Shawcross, J. Theaker, D. Whitcombe, R. Ferrie, C. Newton and S. Little, The elimination of primer-dimer accumulation in PCR, *Nucleic Acids Res.*, 1997, **25**, 3235–3241.
- 453 G. Faccio, From protein features to sensing surfaces, *Sensors (Switzerland)*, 2018, **18**, 1204.
- 454 J. C. G. Blanco, S. Minucci, J. Lu, J. C. G. Blanco, S. Minucci, J. Lu, X. Yang, K. K. Walker, H. Chen, R. M. Evans, Y. Nakatani and K. Ozato, The histone acetylase PCAF is a nuclear receptor coactivator The histone acetylase PCAF is a nuclear receptor coactivator, *Genes Dev.*, 1998, **12**, 1638–1651.
- 455 Y. Goto, T. Katoh and H. Suga, *Preparation of materials for flexizyme reactions and genetic code reprogramming*, Nature Publishing Group, 2011.
- 456 Y. Hayashi, J. Morimoto and H. Suga, In vitro selection of anti-Akt2 thioether-macrocyclic peptides leading to isoform-selective inhibitors, *ACS Chem. Biol.*, 2012, **7**, 607–613.
- 457 T. E. McAllister, T.-L. L. Yeh, M. I. Abboud, I. K. H. H. Leung, E. S. Hookway, O. N. F. F. King, B. Bhushan, S. T. Williams, R. J. Hopkinson, M. Münzel, N. D. Loik, R. Chowdhury, U. Oppermann, T. D. W. W. Claridge, Y. Goto, H. Suga, C. J. Schofield, A. Kawamura, M. Munzel, N. D. Loik, R. Chowdhury, U. Opperman, T. D. W. W. Claridge, Y. Goto, H. Suga, C. J. Schofield and A. Kawamura, Non-competitive cyclic peptides for targeting enzyme-substrate complexes, *Chem. Sci.*, 2018, **9**, 4569–4578.