Investigation of Nucleosome Dynamics by Genetic Code Expansion

Dissertation
for the award of the degree
„Doctor rerum naturalium“
of the Georg-August-Universität Göttingen

within the doctoral program *Biomolecules: Structure-Function-Dynamics*
of the Georg-August University School of Science (GAUSS)

submitted by
Liljan E. Hahn
from Oldenburg

Göttingen, 2015

Thesis Committee

Prof. Dr. Heinz Neumann (Supervisor),
Angewandte Synthetische Biologie, Universität Göttingen

Prof. Dr. Ulf Diederichsen,
Institut for Organische und Biomolekulare Chemie, Universität Göttingen

Prof. Dr. Ralf Ficner,
Institut für Mikrobiologie und Genetik, Universität Göttingen

Members of the examination board

Referee:

Prof. Dr. Heinz Neumann,

Angewandte Synthetische Biologie, Universität Göttingen

2nd Referee:

Prof. Dr. Ulf Diederichsen,

Institut for Organische und Biomolekulare Chemie, Universität Göttingen

Prof. Dr. Ralf Ficner,

Institut für Mikrobiologie und Genetik, Universität Göttingen

Further members of the examination board

Prof. Dr. Kai Tittmann, Molekulare Enzymologie, Universität Göttingen

Prof. Dr. Matthias Dobbelstein, Molekulare Onkologie, Universität Göttingen

Prof. Dr. Jörg Stülke, Institut für Mikobiologie und Genetik, Universität Göttingen

Date of oral examination: 10.03.2015

Affidavit

Herewith I declare that I prepared this thesis "Investigation of nucleosome dynamics by genetic code expansion" independently with no other sources and aids than quoted.

Göttingen, 29.01.2015

Liljan Hahn

„Jedes Naturgesetz, das sich dem Beobachter offenbart,
lässt auf ein höheres noch unerkanntes schließen.“

Alexander von Humbold

0 Summary

The DNA of eukarytic cells is compacted into a complex macromolecular structure called chromatin. Histones are the small structural proteins that form central octamers around which the DNA is wrapped to form the basic repeating units of chromatin, the nucleosomes. The histone proteins are highly altered by a plethora of posttranslational modifications (PTMs) which are involved in altering the chromatin structure affecting various cellular mechanisms. Acetylation of lysine residues neutralizes the positive charge thereby disturbing contacts with DNA or other histones. Recently, other acyl modifications, including lysine propionylation, crotonylation or butyrylation have been discovered. In this thesis the site-specific incorporation of those novel acyl modifications was addressed using genetic code expansion. This technique is used to incorporate unnatural amino acids (UAAs) in response to an amber stop codon. The codon is recognized by a pair of evolved aminoacyl-tRNA synthetase/tRNA$_{CUA}$ that are orthogonal components for the recognition of specific UAAs. Mutant synthetases were evolved to recognize acyl lysine derivatives as a substrate and were selected from a MbPylS library.

The histone H4 tail was reported to be most extensively modified by the different acyl modifications. Since H4 was hitherto only unsuccessfully addressed by genetic code expansion, a new strategy which overcame these limitations is presented here. Furthermore the yield of site-specifically modified H4 was efficient enough to allow for the reconstitution of octamers and nucleosomal arrays carrying the modified H4. Thereby, we provide a basis to study effects caused by these acyl modifications.

We could identify deacylation activity for the *E. coli* HDAC CobB, indicating different reaction velocities depending upon the type of acyl modification.

In further explorations using genetic code expansion technologies, we set out to explore the highly dynamic nucleosome structure by monitoring time resolved structural changes by fluorescence resonance energy transfer (FRET). However, this approach requires the installation of two distinct fluorophores, which remains challenging. In this thesis, the site-specific installation of FRET pairs on histones is addressed based on a combinatorial approach combining the labeling of an UAA carrying a functional group suitable for derivatization by click chemistry with the thiol labeling of an individually placed cysteine.

Table of Contents

0 Summary ... I

Table of Contents ... III

Abbreviations ... VII

1 Introduction ... 1

1.1 The genetic code ... 1

1.2 An expanded genetic code ... 1

1.3 DNA compaction by histone proteins ... 8

1.4 Post-translational modifications ... 10

1.4.1 Novel acyl PTMs ... 12

1.4.2 Acetylation regulating enzymes ... 14

1.5 Investigations of protein dynamics using genetic code expansion ... 16

1.6 Aims ... 20

2 Results ... 21

2.1 An avenue towards new lysine modifications ... 21

2.1.1 Incorporation of acyl modifications on histones ... 26

2.2 First attempts towards protein labeling using GCE ... 42

2.2.1 Introduction of two distinct UAAs for protein labeling ... 53

2.2.2 Direct introduction of a fluorescent UAA ... 55

3 Summary and Conclusions ... 59

3.1 Novel lysine acyl modifications and their investigations ... 59

3.2 Attempts towards protein labeling with fluorophores ... 63

4 Materials and Methods ... 69

4.1 Materials ... 69

4.1.1 Machines ... 69

4.1.2 Other materials ... 70

4.1.3 Chemicals ... 71

4.1.4 Unnatural amino acids ... 72
4.1.5 Fluorophores ... 72
4.1.6 Kit systems ... 73
4.1.7 Chromatography columns and resins ... 73
4.1.8 Buffers and solutions ... 73
4.1.9 Composition of SDS gels (for six gels) ... 75
4.1.10 Media and culture plates ... 75
4.1.11 Antibiotics and work concentrations ... 75
4.1.12 Enzymes, buffer and protein/DNA size standards ... 75
4.1.13 Cell lines ... 76
4.1.14 Plasmides ... 77
4.1.15 Antibodies ... 78
4.1.16 DNA oligonucleotids ... 78
4.2 Molecular biology methods ... 80
4.2.1 Purification of plasmid DNA “Miniprep” ... 80
4.2.2 Large scale purification of plasmid DNA “Giga Prep” ... 80
4.2.3 Agarose gel electrophoresis ... 80
4.2.4 DNA-elution from agarose gels ... 80
4.2.5 Polymerase chain reaction ... 81
4.2.6 Mutagenesis PCR “QuickChange” ... 81
4.2.7 Inverse PCR ... 83
4.2.8 Restriction digestion of DNA ... 84
4.2.9 Ligation reactions ... 84
4.2.10 Ethanol precipitation of DNA ... 84
4.2.11 Determination of nucleic acid concentration ... 85
4.2.12 Sequencing ... 85
4.2.13 Preparation of DNA templates for chromatin assembly ... 86
4.2.14 Preparation of the pCDF-PylT-AraP-His$_6$-H3(Δ93-98)-TEV-H4 plasmid ... 87

4.3 Cell biology methods 87

4.3.1 Competent cells for chemical transformations 87

4.3.2 Competent cells for electroporation 87

4.3.3 Transformation of competent bacteria 88

4.3.4 Electroporation 88

4.4 Protein biochemical methods 88

4.4.1 SDS gel electrophoresis 88

4.4.2 Native PAGE analysis 89

4.4.3 Western blot 89

4.4.4 Stripping of membranes for reprobing 90

4.4.5 Estimation of protein concentrations using Bradford reagent 90

4.4.6 TEV cleavage 90

4.4.7 Dialysis 91

4.4.8 Chemical modification of lysine residues on WT histone H4 91

4.5 Expression and Purification of proteins 91

4.5.1 Expression and purification of recombinant WT histone proteins by ion exchange chromatography 91

4.5.2 Purification of his$_6$-tagged proteins via Ni^{2+}-batch 93

4.5.3 Production and purification of recombinant myoglobin 93

4.5.4 Expression and purification of recombinant acyl-modified histone H4 proteins 94

4.5.5 Expression and purification of recombinant histones containing an UAA 94

4.5.6 Expression and purification of GFP 95

4.5.7 Production of TEV protease 95

4.5.8 Expression and purification of CobB 96

4.5.9 Expression and purification of Hst2 96

4.5.10 Expression and purification of GST-MBP 97

4.6 Recombinant chromatin 98

4.6.1 Histone octamer refolding ... 98

4.6.2 Reconstitution of mononucleosomes via salt dialysis ... 99

4.6.3 Reconstitution of nucleosome arrays via salt dialysis ... 99

4.6.4 Analysis of nucleosome arrays ... 99

4.6.5 Nucleosome array self-association assay ... 100

4.6.6 Assay lysine deacylation with CobB ... 100

4.7 Pyrrolysine mutational library ... 102

4.7.1 Library creation ... 102

4.7.2 Library selection ... 103

4.8 Analysis of proteins by mass-spectrometry ... 105

4.8.1 Analysis of total proteins by ESI-MS ... 105

4.8.2 Protein peptide analysis by MS/MS ... 105

4.9 Protein labeling ... 108

4.9.1 Cu(I) catalyzed click reaction ... 108

4.9.2 Copper free click reaction ... 108

4.9.3 Labeling cysteine residues with maleimide dyes ... 108

4.9.4 Methanol-chloroform precipitation to remove free dye ... 109

5 References ... 111

6 List of Figures ... 127

7 List of Tables ... 129

8 Appendix ... 131

8.1 Purification of acylated octamers ... 131

9 Curriculum Vitae ... 133

Abbreviations

AA	amino acid
Ac	acetyl
AcK	N(ε)acetyl-L-lysine
ACN	acetonitrile
Amp	ampicillin
APS	ammonium peroxydisulfate
AzF	p-Azido-L-phenylalanine
BocK	N(ε)*tert*-Butyloxycarbonyl-L-lysine
bp	base pairs
BSA	bovine serumalbumin
BuK	N(ε)butyryl-L-lysine
CIP	calf alkaline phosphatase
Cm	chloramphenicol
CrK	N(ε)crotonyl-L-lysine
DCM	dichloromethane
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotidetriphosphats
DOC	Na-deoxycholate
DTT	dithiothreitol
E. coli	*Escherichia coli*
ECL	enhanced chemiluminescence
EDTA	ethylenediaminetetraacetic acid
ESI	electrospray ionization
Et	ethyl
FRET	fluorescence resonance energy transfer
GFP	green fluorescent protein
GST	Glutathion-S-Transferase
H4mod	Histone H4 containing an acyl group at position K16
HEPES	2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid
HRP	horseradish peroxidase
IPTG	isopropyl β-D-1-thiogalactopyranoside
Kan/K	kanamycin
kDa	kilodalton

LB medium	Luria broth medium
mal	maleimide
Mb	methanosarcina bakeri
MBP	Maltose-binding-protein
Me	methyl
MNase	micrococcoal nucelase
MS	mass spectrometry
MW	molecular weight
MWCO	molecular weight cutoff
NBD-K	(7-nitrobenz-2-oxa-1,3-diazole-4-yl)-lysine
NMR	nuclear magnetic resonance
NTA	nitrilotriacetic acid
OD_{600}	optical density at 600 nm wavelenght
ON	overnight
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PIC	protease inhibitor cocktail
PMSF	phenylmethanesulfonylfluoride
ppm	parts per million
PrgK	N(ε)propargyloxycarbonyl-L-lysine
PrK	N(ε)propionyl-L-lysine
PVDF	polyvenylidene fluoride
PylS	pyrrolysine aminoacyl-tRNA synthetase
PylT	pyrrolysine tRNA
rcf	relative centrifugal force
RNA	ribonucleic acid
rpm	revolutions per minute
RT	room temperature
SDS	sodium dodecyl sulfate
Spec/Sm	spectinomycin
ßME	ß-Mercaptoethanol
TBE-buffer	tris-borat-EDTA-buffer
TBS	tris-buffered saline

TCA	trichloroacetic acide
TCEP	tris(2-carboxyethyl)phosphine
TEMED	tetramethylethylenediamine
Tet/T	tetracycline
TEV	*Tobacco etch virus*
TFA	trifluoroaceticacide
THF	tetrahydrofurane
TLC	thin layer chromatography
Tris	tris(hydroxymethyl)aminomethane
UAA	unnatural amino acid
UV	ultraviolet
WB	Western blot
WT	wild type

Amino acids are indicated either by the three letter or one letter code.

1 Introduction

1.1 The genetic code

The genetic code is universally conserved among all organisms. It is based upon DNA sequences, which encode the genetic information essential for life. The DNA sequences consists of alternating nucleotide units which encompass one of the four nucleobases, adenine (A), thymine (T), cytosine (C) and guanine (G) in combination with desoxyribose and a phosphate group. Three consecutive nucleobases form so-called codons, which read as instructions for the proper translations of proteins. The four bases can be combined to result in 64 different codons. 61 of them are assigned to decode for 20 canonical amino acids. Each codon reads for an amino acid during protein synthesis, except three combinations that signal termination of translation. These stop codons are named amber (UAG), ochre (UAA) and opal (UGA).

Proteins play critical roles in all biological systems, as they possess a remarkable range of functionalities, though assembled by just 20 amino acids. The unique amino acid sequence defines the individual function of a protein and even minor changes within the sequence can alter the functionality or perturb protein folding.

1.2 An expanded genetic code

The functionalities provided by this set of 20 native amino acids, is wide but still limited. Cells have developed different strategies to overcome these limitations. One possibility, used by different enzymes, is the binding of cofactors. Cofactors include metal ions or small organic molecules. Like Mn which is used by superoxide dismutase or Cu and Zn which are required for function of cytochrome-c-oxidase.[1] Among the small organic cofactors are NAD, FAD, ADP, or CoA which function as carrier for electrons, acyl-, or phosphate groups.[2] Another amplification of functionalities is offered by the plethora of posttranslational modifications already reported today.

However, the genetic code is not limited to the 20 canonical amino acids but a 21st amino acid, selenocysteine and a 22nd, pyrrolysine were discovered.[3,4] Selenocysteine (SeC) incorporation was found in eubacteria, archaebacteria and eukaryotes. The incorporation is in response to an in-frame opal codon (UGA) during translation requiring a special mRNA sequence.[1] Thus, this mechanism is not based on an orthogonal tRNA, but instead a tRNA charged with serine, followed by enzymatic conversion into SeC.[3,5]

In contrast, pyrrolysine is encoded by an amber codon (UAG) via a unique orthogonal aminoacyl-tRNA synthetase (AARS)/tRNA pair that suppresses the stop codon. Both suppressor tRNA and AARS do not confer recognition of endogenous amino acids but the AARS charges only its cognate tRNA with pyrrolysine, delivers it to the ribosome and is there incorporated in the nascent amino acid chain in response to the amber codon.

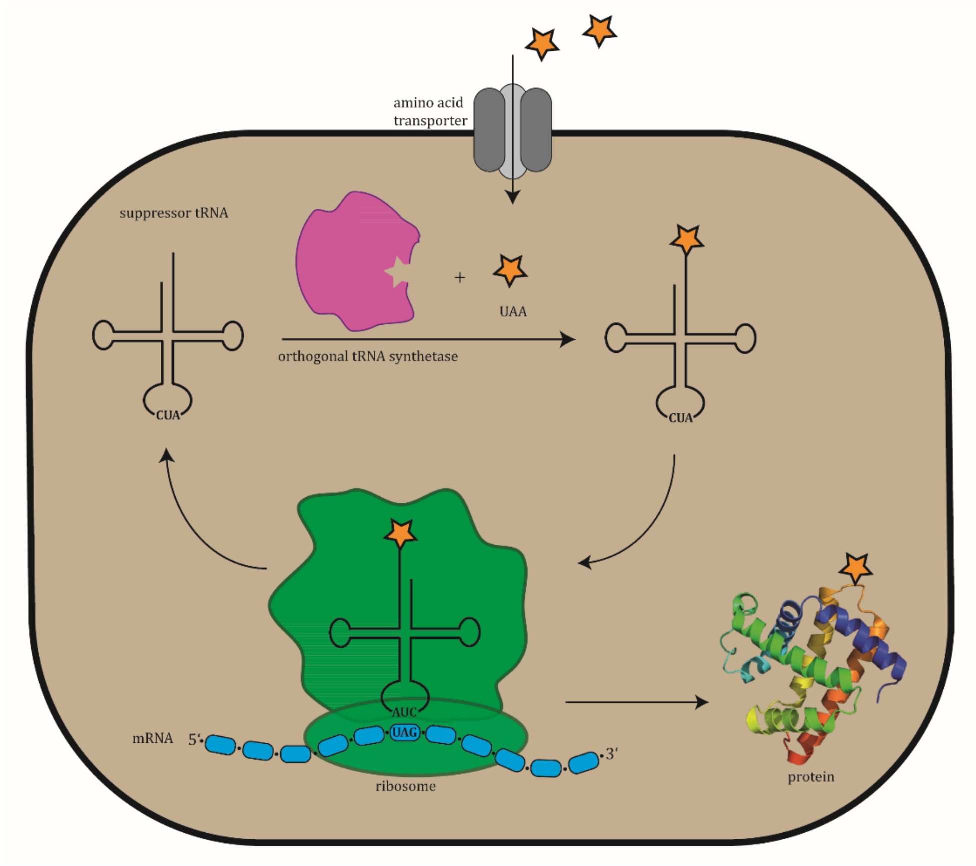


Figure 1 ***Incorporation of an unnatural amino acid (UAA).*** An orthogonal amber suppressor tRNA synthetase charges an orthogonal suppressor tRNA with the UAA. This UAA is then incorporated at the ribosome in the nascent amino acid chain in response to the amber codon. Thereby creating proteins with a unique functionality at the position of interest.

Since nature already started to modify the genetic code, this is of great interest for scientists as a site-directed modification of the genetic code would facilitate new perspectives towards the investigation of cellular processes, function of biomolecules or for the development of new drugs. Based on methods, provided by nature with the introduction of pyrrolysine, this method was adapted and modified for the site-specific incorporation of amino acids carrying functionalities not provided by the canonical ensemble. Those amino acids are termed unnatural amino acids (UAAs) and amber suppressor pairs were discovered for their site specific incorporation at the genetic level. The general scheme, how an UAA is incorporated by amber suppression is depicted in Figure 1.

The orthogonal pairs must fulfill certain requirements.[6,7] The suppressor tRNA must be aminoacylated by the AARS with only the desired UAA, thus not charging the tRNA with any of the canonical amino acids. Concomitantly the UAA must not be a substrate for any of the endogenous synthetases. And finally the UAA has to be efficiently transported into the cells.

This was mainly achieved by adapting tRNA/AARS pairs from other organsims. For use in *Escherischia coli,* pairs were imported from e.g. *Saccharomyces cerevisiae* and evolved for specific use.[7,8]

A first attempt for incorporation of an UAA was carried out by Further in 1998 using *p*-fluoro-phenylalanine (*p*-F-Phe), a substrate suitable for ^{19}F-NMR studies. Therefore a nonessential heterologous $\text{tRNA}^{\text{Phe}}_{\text{CUA}}$/PheRS pair from *S. cerevisiae* was used in *E. coli*. The synthetase accepts *p*-F-Phe as substrate without being further evolved. However, this system was not truly orthogonal due to the recognition of both substrates. In order to efficiently produce protein with *p*-F-Phe the substrate had to be used in high excess to outcompete the native Phe substrate.[9]

The first truly orthogonal pair was discovered by Wang et al. by importing a tRNA/AARS pair from *Methanococcus jannaschiii* in *E. coli* for the incorporation of *O*-methyl-L-tyrosine.[10] They evolved the tRNA^{Tyr} towards responding to the UAG stop codon and the TyrRS to accepting only *O*-methyl-L-tyrosine as new substrate to produce dihydrofolate reductase (DHFR) protein containing the new substrate at one specific position.

Once researchers developed functional orthogonality, the true power of the system became apparent. The synthetase specificity could be altered towards recognizing any unique UAA of interest to finally yield in proteins containing novel functionalities. This was achieved by creating libraries which contain randomized mutations within the recognition site of the synthetase. These libraries of mutants could then be selected for mutants with the ability to selectively incorporate the UAA of interest (depicted in Figure 2). Alternating rounds of positive and negative selection were performed in this regard.[8] For the positive selection, the mutant library was transformed into *E. coli* cells that contain a chloramphenicol acetyl transferase (CAT) gene, perturbed by amber codons.[7,11] The cells were grown in the presence of the UAA. Synthetases, capable of charging the suppressor tRNA, with either the UAA or a native amino acid, suppress the submissive amber codon and allow full-length expression from the CAT gene, thus surviving of the cells in the presence of chloramphenicol.[6,12] In the second round a negative selection was performed. Synthetases selected in the first positive round were co-transformed in *E. coli* cells together with a plasmid harboring a toxic barnase gene, perturbed by three amber codons.[13] The full-length barnase protein is expressed in all clones that aminoacylate endogenous amino acids, causing cell death as the protein contains a toxic ribonuclease activity.[14–16] This round was performed in the absence of the UAA, causing death of undesired mutants and only desired mutants survived. The herewith obtained mutants were used in an iterative positive selection round. Leading towards the evolution of an evolved synthetase, specific to a desired UAA as its substrate.

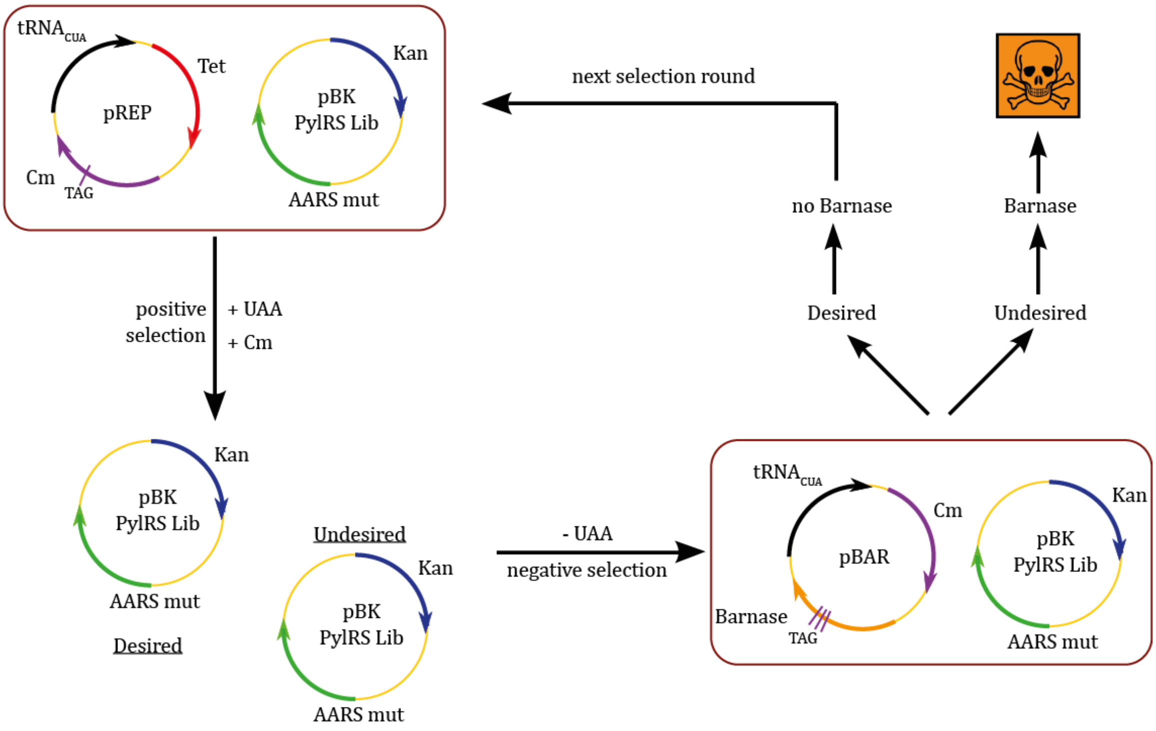

$tRNA_{CUA}$
Tet
pREP
Cm
TAG
Kan
pBK
PylRS Lib
AARS mut
next selection round
no Barnase
Barnase
Desired
Undesired
positive selection
+ UAA
+ Cm
Kan
pBK
PylRS Lib
AARS mut
Desired
Undesired
Kan
pBK
PylRS Lib
AARS mut
- UAA
negative selection
$tRNA_{CUA}$
Cm
pBAR
Barnase
TAG
Kan
pBK
PylRS Lib
AARS mut

Figure 2 ***Selection of a PylS mutant library***. A library of PylS synthetases with some randomized positions, is applied to alternating rounds of positive and negative selection to identify one mutant that specifically recognizes one unique UAA. In the first positive selection round the library is transformed into *E. coli* cells together with a reporter plasmid, carrying a CAT gene perturbed by an amber codon. The selection is done in the presence of the UAA. Synthetases that can charge the UAA or any of the endogenous AA suppress the amber codon and survive chloramphenicol (Cm) pressure. Obtained mutants are performed in a negative selection and therefore combined with another reporter plasmid, carrying a barnase gene perturbed by three amber codons. Barnase is toxic to the cells and as the selections is processed in the absence of the UAA, only mutants charging the tRNA with any of the endogenous AA express the toxic barnase protein and die. Finally, the positive selection is repeated.

With the selection (as described in Figure 2) of a mutant PylS library, another pair, the pyrrolysyl tRNA synthetase of *Methanosarcina barkeri* and its cognate tRNA (MbPylRS/tRNA) was evolved by Neumann et al. for substrate specificity and site specific incorporation of N(ε)-acetyllysine in myoglobin.[17]

With this technique, numerous UAAs, with different functionalities, could already be incorporated not only in *E. coli*, but also in *S. cerevisiae,* Xenopus oocytes, mammalian cells and includes even animals, namely *Caenorhabditis elegans* and *Drosophila melanogaster*.[18–25]

Novel functionalities that could be exploited to address biological question are different crosslinkers to map protein interactions, site-specifically introduced post-translational modifications, some fluorescent amino acids and samples suitable for IR- or NMR spectroscopy.

A very suitable crosslinker is the UV inducible *p*-benzoyl-phenylalanine (BPA), which was evolved by Chin et al.[26] This crosslinker was successfully used by Wilkins et al. to capture how a cascade of histone modifications induces chromatin condensation in living yeast cells.[27]

Various PTMs could likewise be encoded thereby accessing possibilities to investigate PTMs site-specifically. Among them, the already mentioned acetylation, which was used by Neumann et al. to investigate the effect of K56 acetylation on histone H3 on nucleosome stability. Thereby revealing a DNA breathing effect on the nucleosome core particle induced by the acetylation.[28] Further genetically encoded are lysine methylations. Whereby the methylation encountered to be not incorporated straight forward as the chemical difference between lysine and N(ε)-methyl-lysine are too small to result in specific synthetases. This was solved for the monomethylation by encoding the boc protected N(ε)-Boc-N(ε)methyl-lysine, followed by chemical cleavage of the Boc protection group, performed on the purified protein.[29] For accessing dimethylated lysine the protection group chemistry had to be applied broader. Again the lysine of interest was incorporated as protected derivative, this time BocK. All other lysine residues of the purified protein were covered by another protection group, followed by deprotection of boc on the individual lysine residue and chemical dimethylation, before removal of remaining protection groups.[30] Other genetically encoded PTMs are tyrosine phosphorylation mimics[31,32], phosphorylated serine[33] and sulfonated tyrosine[34].

Another modification is the by oxidative damage caused nitrolation of tyrosine, which is a vascular disease marker. The site-specific incorporation of 3-nitro tyrosine in manganese superoxide dismutase at one position reported to be modified disease related, could be used to deduce the dramatically reduced activity caused by this modification.[35]

Further photocaged versions of Tyr, Cys, Ser and Lys were established.[36,37] The presence of a photocaged derivative of one amino acid blocks one particular function, which can be restored by de-caging the amino acid using a pulse of light, thereby activating the protein inside the living cell to elucidate resulting changes. The photocaged serine was used to determine the effect of phosphorylation in the yeast transcription factor Pho4, affecting its export from the nucleus.[38] Gautier et al. used a photocaged lysine, placed within a nuclear localization sequence, in combination with a fused GFP to follow protein localization.[39]

Another expanding area of functionalities applies by installing probes with different spectroscopic properties or functionalities that can be modified by bioorthogonal labeling methods, like click reactions. UAAs with spectroscopic applications include a broad range of methods. ^{19}F-labeled UAAs are a powerful tool to monitor protein conformational changes by NMR.[40,41] Li et al. could show that this method is also applicable for in cell NMR studies in *E. coli*.[42] More recently also stable spin labels for use in EPR studies were encountered.[43] A technique used to monitor even faster conformational changes is IR spectroscopy. Suitable samples, incorporated site-specifically are derivatives of Phe bearing nitriles or azides, which have very unique vibrational signatures.[44,45] The *p*-azido-phenylalanine (AzF) was used by Ye et al. to investigate conformational changes in rhodopsin.[46] Another powerful tool to investigate protein dynamics is by using fluorescence spectroscopy which is discussed in chapter 1.5.

However, all the different functional UAAs are mainly enabled by just four AARS/tRNA pairs.[47] The *M. jannaschii* TyrRS/tRNA pairs which is orthogonal only in *E. coli* [48] and the PylS/tRNA system from *M. bakeri,* which is orthogonal in bacteria, yeast and mammals [17,19,21,24,49]. The remaining pairs consisting of the *E. coli* TyrRS/tRNA and *E. coli* LeuRS/tRNA that were evolved to be orthogonal in yeast and mammalian cells, complete the list.[20,22,37]

The amber stop codon is the most commonly used stop codon for these techniques, as it is the least abundant in *E. coli*. Other blank codons like ochre and opal together with four-base codons were also accessed for genetic code expansion.[50–52]

Genetic code expansion is a valuable tool, but it requires blank codons for utilization of new functionalities, which are in terms of reprogrammed stop codons limited by a maximum of two distinct functionalities due to the lack of additional stop codons for reassignment. Additionally orthogonal AARS/tRNA pairs need be established for their recognition. Two distinct UAAs were incorporated by Anderson et al. into myoglobin by combining orthogonal tRNAs for amber suppression with a four-base suppressor tRNA.[53] Wan et al. evolved the PylT towards encoding of an ochre stop codon and combined this system with the evolved *M. jannaschii* TyrRS/tRNA$_{\mathrm{CUA}}$ and possessed therewith installation of two distinct UAAs with functionalities suitable for labeling with fluorescent dyes onto one protein.[50]

Another step, to further expand the genetic code, is the approach to generate more blank codons. This goal was addressed by Chin and coworkers, targeting the ribosome itself. They created a nonessential orthogonal ribosome which operates parallel to the natural ribosome, as it differs in the Shine-Dalgarno sequence from the natural ribosome.[18] These ribosomes could be evolved towards new functions. Mutations in the A-site created ribosomes highly efficient for amber suppression.[54] One step further, the ribosome could be evolved to decode quadruplet codons and one amber codon.[55] The quadruplet codons provide theoretically 256 new blank codons, which would allow the step from proteins with enhanced functionalities towards polypeptides completely consisting of UAAs.

1.3 DNA compaction by histone proteins

The DNA of eukaryotes must be densely packed to fit the entire genome into the small volume of the nucleus. Therefore the DNA is compacted in the cells into chromatin.[1] The chromatin structure is composed of small repeating subunits, the nucleosome.

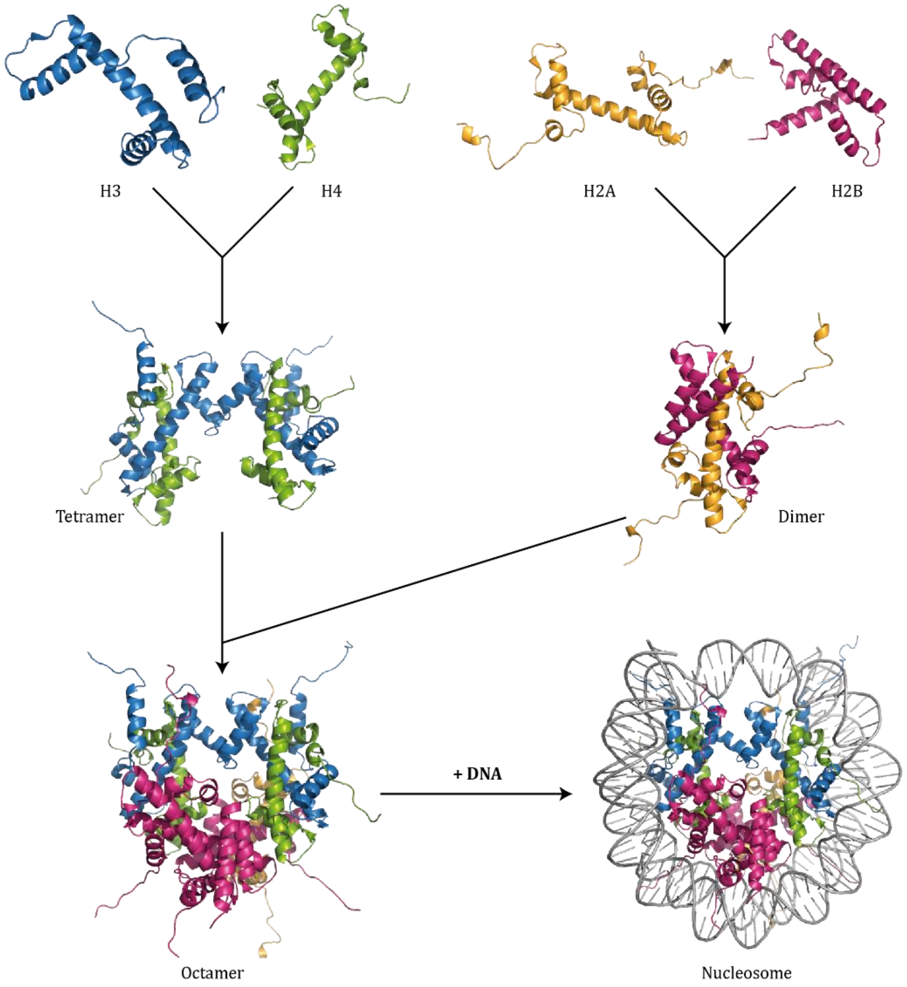


Figure 3 ***Nucleosome assembly.*** The core histones H3 and H4 form a central tetramer. Histones H2A and H2B form two heterodimers respectively, which together with the tetramer built the octamer. By addition of DNA the nucleosome is assembled. The picture was created based on the crystal structure PDB-Id: 1AOI.[56]

The nucleosome itself is assembled by DNA closely associated around an octamer of histone proteins. Histones are small basic proteins, containing a high percentage of lysine and arginine residues.[1] The core histones are H2A, H2B, H3 and H4 and are completed by H1, a linker histone. Two copies of each histone associate into an octamer. Each octamer is composed of a central H3 and H4 tetramer flanked by two H2A, H2B heterodimers.[57,58] The octamer tightly binds 147 base pairs (bp) of DNA that wrap around the core structure (shown in Figure 3). Linker histone H1 then binds the DNA at its entry-exit point. The histone N-terminal tail domains protrude out of the core nucleosome unit, allowing histone access to the extranucleosomal spaces. The DNA between nucleosomes, linker DNA, connects nucleosomes, forming arrays and the "beads on a string" appearance. This 10 nm fiber is considered the primary structure of chromatin.[59] Nucleosomes undergo short-range interactions, forming higher ordered "secondary structures" (also referred to as 30-nm fibers), long-term interactions contribute to further compaction into "tertiary structures"[60] and their final arrangement represents the maximally folded chromatin fiber[61]. This is illustrated in Figure 4. However, chromatin structure is highly heterogeneous, with regions that are densely packed (heterochromatin) and regions that are relatively uncondensed and transcriptionally active (euchromatin).[62] DNA accessibility is a critical parameter for processes such as transcription, replication, recombination and DNA repair.[63] It is controlled by various factors, with PTMs, both on the histone N-terminal tails and within the core region playing a key regulatory role.[64]

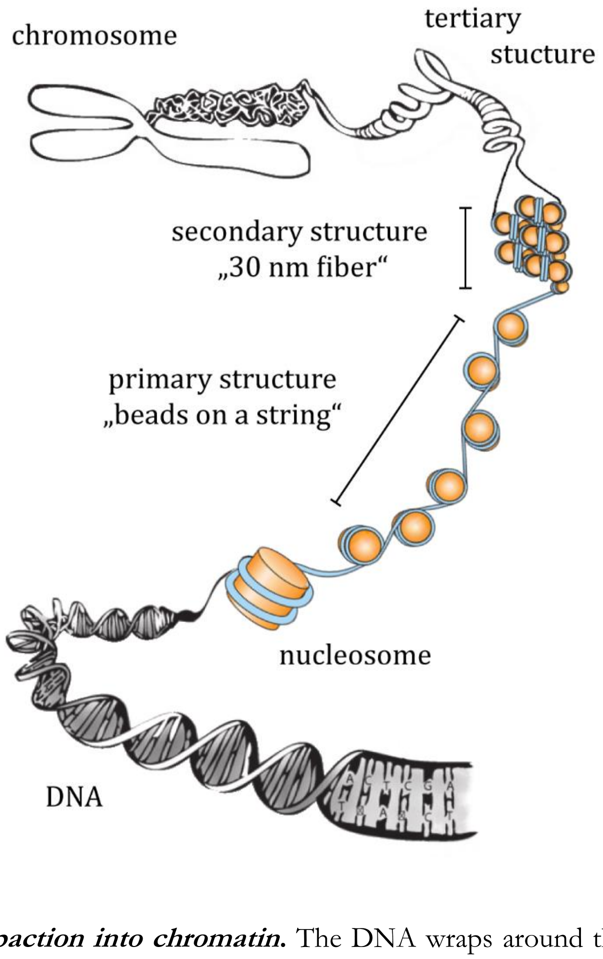


Figure 4 ***DNA compaction into chromatin.*** The DNA wraps around the histone octamer which consists of four core histones, to form the basic repeating unit the nucleosome. The nucleosomes further compact into primary, secondary and ternary structures to finally form chromatin. The picture was modified from Allis et al. and Lodish et al.[65,66]

1.4 Post-translational modifications

Chromatin structure is highly regulated by histone posttranslational modifications (PTM). PTMs are epigenetic marks that are generally defined as influencing changes in gene expression without affecting the DNA sequence. These posttranslational modifications are of extreme interest as they are involved in dynamic changes in chromatin structure affecting cellular processes including transcription, replication and DNA repair. The majority of PTMs are located at the histone tails where they alter chromatin structure by recruiting effector proteins or directly interfering histone DNA contacts[63], as depicted in Figure 5. PTMs include, but are not limited to, acetylation, mono-, di- or trimethylation, phosphorylation, ubiquitination, and sumoylation. While these are the most commonly studied PTMs, the list of newly discovered PTMs is constantly growing. All the different modifications are carefully regulated within the cell by subsets of different enzymes, like histone acetyl transferases (HATs) and histone deacetylases (HDACs) regulating the acetylation of proteins. The different modifications are involved in many cellular functions, thereby are not only individual modifications crucial, but crosstalk between different modifications, or the formation of whole patterns essentially influencing their regulatory role, thus creating a whole new level of complexity.

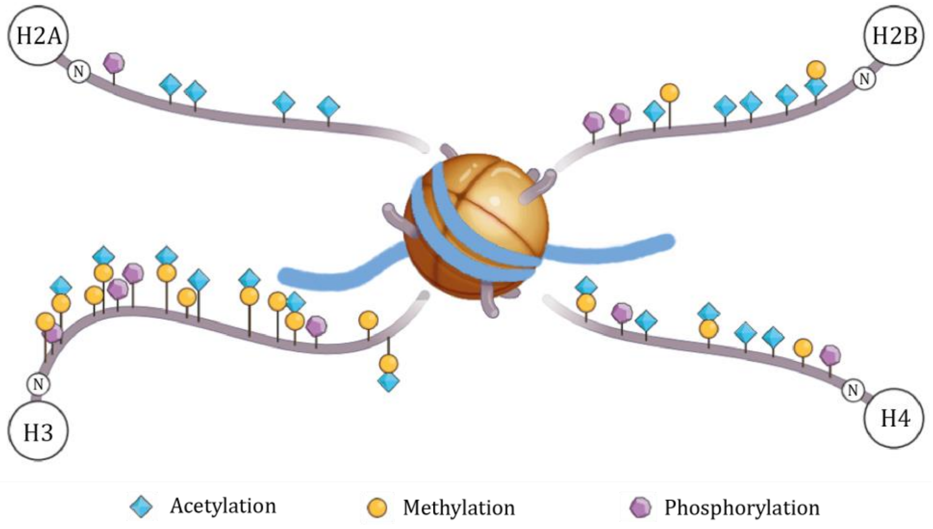


Figure 5 ***PTMs on histone tails.*** The core histone tails are subject to many different PTMs, with acetylation, methylation and phosphorylation being shown here. The picture was modified from Helin et al.[67]

Acetylation of lysine side chains neutralizes the positive charge of lysine residues thus disturbing contacts with other histones or DNA, thereby influencing chromatin structure.[68]

Acetyl modifications are generally correlated to transcriptionally active regions. For example, acetylation located in the nucleosome core at H4 K91 destabilizes the octamer by disturbing interactions between H3/H4 tetramer and H2A/H2B dimers.[69] It has also been reported that acetylation at K122 on histone H3 is an important modification for transcription as it modulates histone-DNA binding.[70] The histone tails of H3 and H4 are specifically acetylated on newly synthesized histones and deactylated later in the cell cycle.[69,71] The dual acetylation of H4 K5 and K12 are the most prominently studied pattern occurring on newly synthesized histones with no functional roles for viability or chromatin assembly. But when these modifications are combined with a non acetylatable mutation at K91, a hypersensitivity towards replication stress and DNA damaging results or combined with mutation of K56 the viability is greatly decreased.[72] Acetylation of H4 K16 is linked to transcriptionally active chromatin, hampers the formation of higher order (30 nm like) fibers and is a key substrate for Sir2.[73,74]

Methylation of residues does not alter the charge of a protein. Methylation of residues is rather complex. Regarding histones it occurs on lysine residues as mono-, di-, or trimethylation, and at arginine residues as monomethylation or symmetrically or asymmetrically dimethylation.[75] The function of methylation is highly dependent on the methylation site, signaling either activation or repression.[75] Furthermore the functionality depends on the methylation state, as reported for X inactivation.[76] Methylation of H3 K9 is linked both to transcription activation and DNA methylation.[77] H4 K20 methylation is reported to be involved in cell cycle regulation and DNA damage response.[76,78,79]

Another theory implies methylation marks as binding sites for effector proteins that mediate downstream effects.[68,80]

Phosphorylation in histones occurs on serines, threonins and tyrosine residues. The phosphorylation adds a negative charge to the modified residue. It is a versatile modification that is involved in many cellular processes including transcription, DNA repair, apoptosis and chromosome condensation.[81] Phosphorylation events are often coupled to acetylations thereby activating transcription.[82–85]

Phosphorylation on the same position can have diverging effects depending on context such as phosphorylation at H3 S10 or H3 S28 are associated with both, chromatin condensation in the cell cycle and transcriptional activation.[82,86,87]

Ubiquitination is a much larger modification compared to the other marks, as it is a 76 amino acids long polypeptide that can be dynamically attached to lysines. A protein can be mono or polyubiquitinated, with polyubiquitination generally targeting proteins for degradation.[88] Monoubiquitination of histones seems to be highly relevant though detailed site and in turn consequences remain elusive.[71] However, ubiquitination of H2A at K119 could be linked to gene silencing[89] whereas ubiquitination at H2B K123 is involved in transcription initiation and elongation[71,90].

In addition, many of the different modifications act together, thereby regulating complex functionalities. This crosstalk between different modifications or the formation of "modification patterns" is referred to as "histone code".[68,80,91–93]

1.4.1 Novel acyl PTMs

In the last decade, many new sites and types of modifications were discovered by taking advantage of mass spectrometric approaches for their identification.[94,95] Findings concluded that acetylation occurs not only on Lys residues but also on Ser and Thr.[96] Furthermore, a variety of novel histone Lys modifications were reported. They include the acyl modifications formylation, propionylation, butyrylation, crotonylation as well as succinylation and malonylation.[97–101] The structures of the novel acyl modifications are depicted in Figure 6.

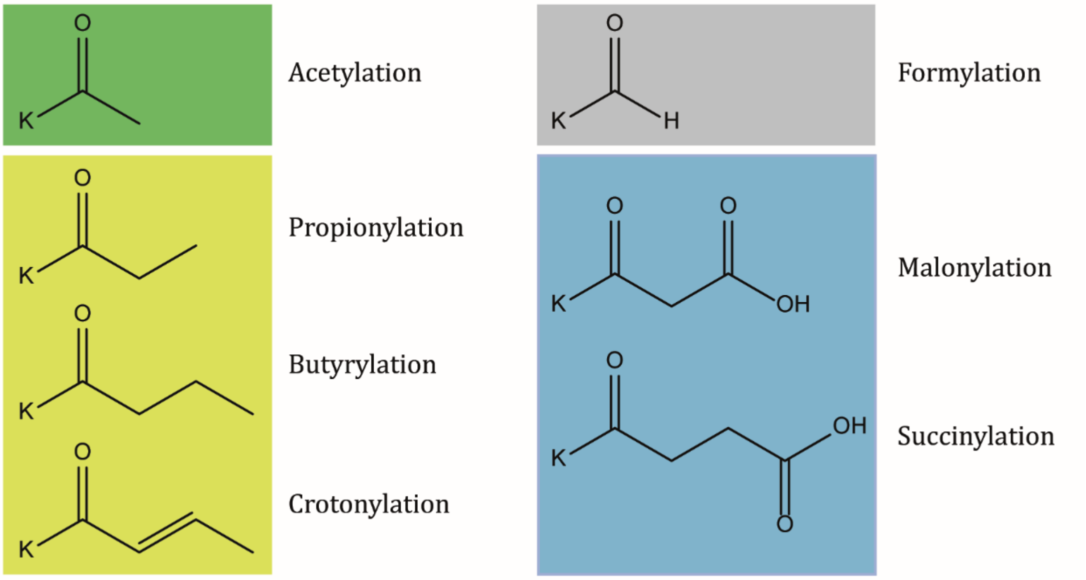


Figure 6 ***Novel histone lysine acyl modifications***. The well-established acetylation AcK is depicted in green. In yellow propionylation (PrK), butyrylation (BuK) and crotonylation (CrK) which are addressed in this thesis are shown. Formylation (FoK) is underlayed in grey and Malonylation (MaK) and succinylation (SuK) are depicted in blue.

Acetylation as a PTM was first reported by Allfrey et al. and its role in many cellular processes, including transcription, replication, repair, or chromosome condensation, has been extensively studied.[102] The distinct roles of the novel acyl modifications still has to be elucidated, some already reported functions are discussed below.

The formylation was reported as secondary modification arising from oxidative stress[97] occurring on residues known to be acetylated or methylated, thereby interfering with normal histone modifications leading to dysregulation and disease of the cell.[103,104]

Lysine succinylation and malonylation were identified by combining MS/MS approaches with western blot analysis using pan antibodies and isotopic labeling[101,105] and reported to be PTMs relevant on histones[106].

Crotonylation of lysines was found by Tan et al. by LC-MS approaches and validated by western blot using a modification specific pan-antibody, thereby reporting many crotonylation sites on histones H2B, H3 and H4.[100] The crotonylation is localized in the genome marking either active promoters or potential enhancers. Enrichment of crotonylation was found in haploid male germ cells immediately following meiosis in genes that escape sex chromosome inactivation designating its functional role in protecting genes from repression.[107]

Propionylation and butyrylation of histones were discovered by the group of Zhao using mass spectrometry and validating the modifications by peptide mapping, western blot with modification specific pan antibodies and coelution of peptides by HPLC.[98,99]

Acetylation neutralizes the positive charge of the lysine site-chain, due to their structural similarity it is likely, that the other acyl modifications similarly weaken histone-DNA contacts.[64] However their specific role or if they assign redundant functions as the acetylation remains to be determined.

1.4.2 Acetylation regulating enzymes

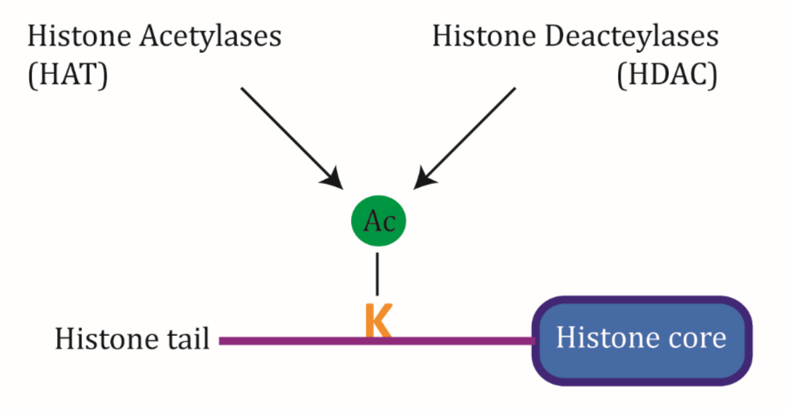


Figure 7 ***Acetyl modifying enzymes***. The acetylation is installed on the amino group of a lysine residue by HATs and removed by HDACs.

The acetylation of organisms is regulated by histone acetyl transferases (HATs) and histone deacetylases (HDACs) as shown in Figure 7. HAT enzymes are separated in three different classes: Gcn5-related N-acetyltransferases (GNATs), p300/CREB binding-protein and MYST proteins.[108] The deacetylating enzymes are divided in two families. The histone deacetylases (HDACs) and the NAD^+ dependent Sir2 (silent information regulator)-like proteins, called sirtuins.[109] The function of sirtuins is dependent on NAD^+, where nicotinamide is cleaved and the acetyl group is subsequently transferred onto ADP-ribose, forming *O*-acetyl-ADP-ribose (*O*AADPr). A schematic mechanism of a sirtuin catalyzed deacetylation reaction is depicted in Figure 8. The sirtuin family includes human Sirt1-7, and homologues of other organisms, among them *E. coli* CobB and yeast Sir2 and Hst2.[110–112]

Sirtuins differ in deacetylation efficiency, with Sirt4-7 being reported to perform only poorly in deacetylation reactions. Together with the discovery of new acyl modifications, this raises questions, whether these sirtuins target other acyl modifications than the acetylation itself.[113]

Specific deacylases, elusive towards one of the novel modifications were not discovered so far. However, some HDAC seem to have robust deacylation activity. Indications that propionyl and butyryl modifications can be removed by known deacetylases was found by Zhang et al. when culturing yeast cells with HDAC inhibitors. Comparison of modification levels by western blot showed significantly enhanced signals for propionyl- and butyryl in the presence of inhibitors.[98] Sirt5 (a human HDAC), which has only low deacetylase activity, addresses efficiently succinyl and malonyl groups.[101,114] The *E. coli* HDAC CobB, was found to have desuccinylation activity, that is comparable to that of deacetylation.[115] Other sirtuins were likewise reported to have depropionylation and debutyrylation activity including Sirt1, Sirt2 and Sirt3.[116,117]

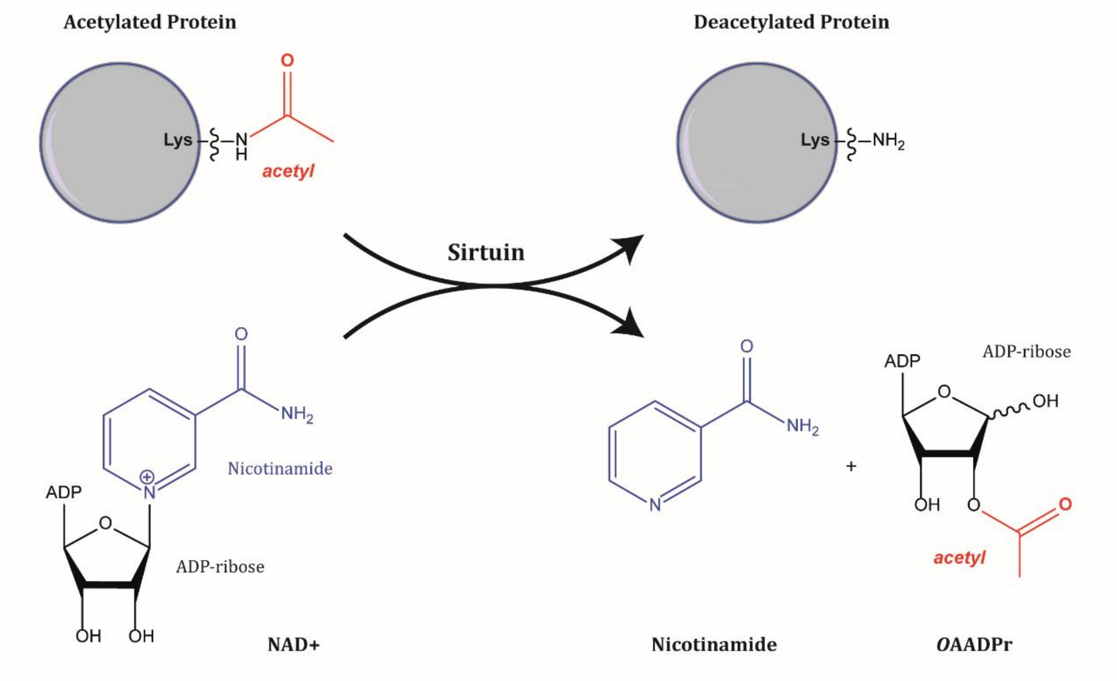

Acetylated Protein
Lys
N
H
O
acetyl
Deacetylated Protein
Lys
NH2
Sirtuin
O
NH2
Nicotinamide
ADP
N
ADP-ribose
OH OH
NAD+
O
NH2
N
+
Nicotinamide
ADP
ADP-ribose
OH
OH
O
O
acetyl
OAADPr

Figure 8 ***Sirtuin catalyzed deacetylation reaction.*** The acetyl group of a substrate is transferred onto ADP-ribose forming *O*-acetyl-ADP-ribose (*O*AADPr) through hydrolysis of NAD^+ thereby generating nicotinamide.

Acetylation of lysines is achieved in cells by histone acetyl transferases (HATs) using acetyl-CoA as substrate. Most of the acylations found are already reported as CoA intermediates present in cells pinpointing towards being as well substrates for histone acetyl transferases too. The known acetyl-transferases p300 and CREB-binding protein were reported to catalyze propionylation and butyrylation in vitro, using radioactive labeled propionyl-CoA and butyryl-CoA as substrates.[99] Human PCAF was found to propionylate or butyrylate H3 peptides.[118] Furthermore the activity of the propionyl-CoA synthetase is regulated by propionylation at K592, with the propionylation rendering the synthetase inactive.[119] The propionylation is introduced by bacterial Gcn5 and can in turn be removed by bacterial or human Sir2.

1.5 Investigations of protein dynamics using genetic code expansion

Investigations of protein structure and function is still a major challenge, as it requires site-specific labeling of proteins to monitor dynamic structural changes. Nonetheless, the introduction of homogenous modifications on proteins is difficult to achieve under physiological conditions. One approach to installing a fluorescent label onto a protein is done by creating fusions with a fluorescent protein like the green fluorescent protein.[120] However, this is limited to the N- or C-terminus of a target protein and its relatively large size can cause perturbations within the structure. Another strategy is to directly synthesize smaller peptides using solid phase synthesis.[121] This method has the advantage of being able to include a wide variety of functionalities, but the peptide-size is limited due to low purification yield. The subsequent verification of the structure and functionality of the peptide is also problematic. Another strategy combines the addition of a synthetic peptide, carrying a specific functional group, to a recombinant expressed protein fragment via native-chemical ligation. This allows the production of larger proteins with functional modifications (semisynthetic protein), but still limits the specific functionality to being close to either N- or C-termini.[122] Modifications can also be introduced via side-chain chemistry of the native functional groups. Examples of this include reacting cysteins with maleimide dyes or using reactive NHS esters for modification of lysines. These strategies can prove difficult because proteins rarely contain only one of the amino acid functionality of interest for labeling.[123–126]

Each of these limitations can be overcome when using amber suppressor pairs to genetically encode desired modifications onto any chosen position of a protein of interest.

The best solution to introduce fluorescent labels would be provided by the direct incorporation of small fluorescent UAAs. However the number of directly encoded fluorescent amino acids is very limited so far[47,127] including only coumarinyl glycine (HceG, structure depicted in Figure 9), dansyl alanine and naphthyl alanine.[13,128,129] A synthetic biology toolbox, to achieve labeling of proteins on a position of interest, is provided by incorporating UAAs bearing functionalities suitable for bioorthogonal derivatization reactions. Reactions are termed bioorthogonal when they are chemoselective under physiological conditions with reactants that are stable at conditions given within the cell. They form stable covalent bonds and are nontoxic.[130,131]

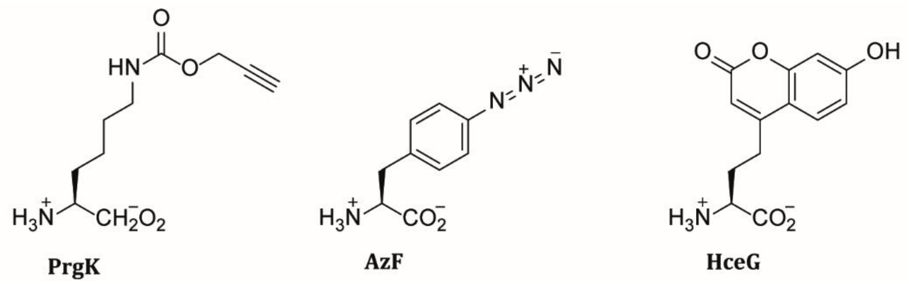


Figure 9 ***Structures of UAAs suitable for bioorthogonal labeling reactions.*** PrgK: N(ε)-propargyloxycarbonyl-L-lysine, AzF: *p*-Azidophenylalanine, HceG, L-(7-hydroxycoumarin-4-yl)-ethylglycerine.

Bioorthogonal reactions have emerged as a premier biological application, with the click reaction being the most prominent. A click reaction occurs between an azide and an alkyne functionality in a [3+2]-cycloaddition, forming a triazole linker (Figure 10). The reactions are accessible for proteins as both, azide and alkyne functional UAAs can be introduced by genetic code expansion. Those include a N(ε)-propargyloxycarbonyl-L-lysine (PrgK) that harbors an alkyne moiety and *p*-Azidophenylalanine (AzF), which has an azide side-chain group. Each of these UAAs are efficient substrates for click reaction and were also used within this work (structures are depicted in Figure 9).[50,132,133]

The typical click reaction is catalyzed by Cu(I), which is cyctotoxic thereby limiting the reaction to *in vitro* applications.[134] However, further developments allowed avoidance of copper as the catalyst by using cyclic alkynes, where the ring-strain promotes the formation of the triazole linker (Figure 10). The copper free bioorthogonal reaction of proteins is of great interest as it has better compatibility with fragile protein structures, and provides applications not only *in vitro* but also *in vivo*.[135]

Additional UAAs, which react efficiently with tetrazines in an inverse electron-demand Diels-Alder cycloaddition, expand the list of functionalities used by genetic code expansion. These are based on strained alkenes and alkynes, including norbonene, bicyclononynes or *trans*-cyclooctynes.[135–139]

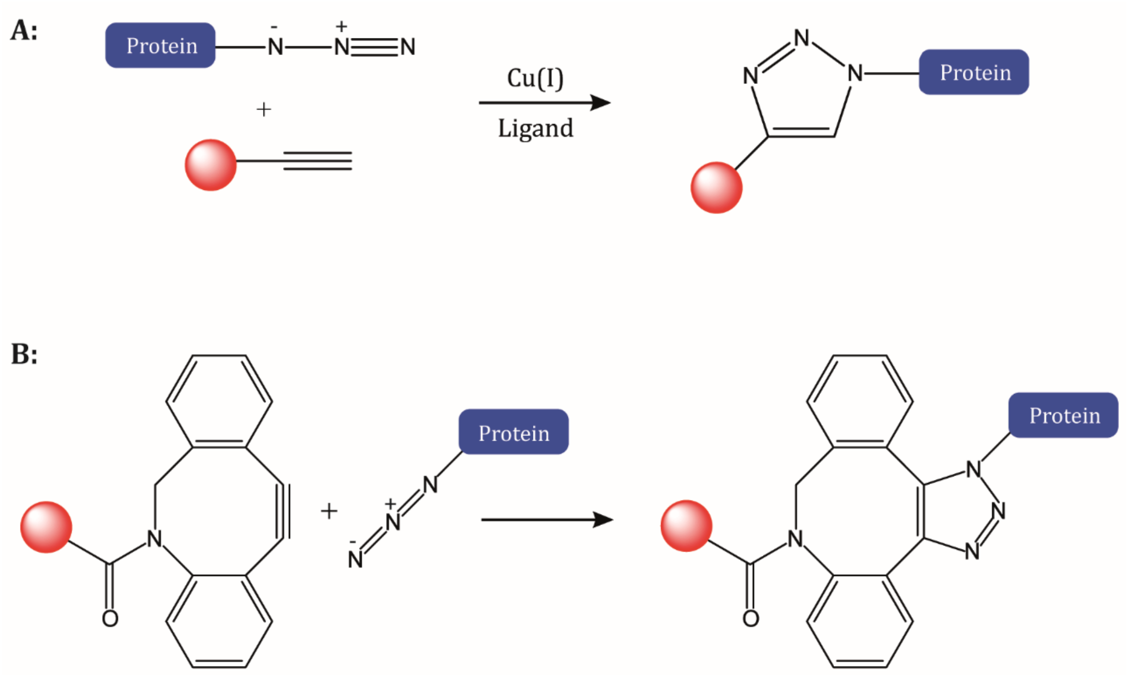


Figure 10 ***Scheme of different click reactions used for labeling of biomolecules.*** A: Cu(I) catalyzed click reaction. B: Strain promoted click reaction, using a dibenzoylcyclooctyne-(DBCO)-dye.

These bioorthogonal reactions can be used to introduce fluorescent labels site-specifically on proteins which allows for their photometrical investigation. The monitoring of time resolved structural changes by Förster/Fluorescence resonance energy transfer (FRET) has proven valuable in this regard. However it requires not only one fluorescent label but two distinct fluorophores with overlapping spectral properties. The donor fluorophore, which is excited using a distinct wavelength, can transfer the energy onto an acceptor fluorophore which then emits photons that are detected. The transfer occurs when donor and acceptor fluorophores are in close proximity. It is based on long-range dipole-dipole interactions and the transfer itself is radiationless. As the efficiency of the energy transfer is highly distance dependent, it allows for time resolved monitoring of distances between the two fluorophores.[140,141]

The technique of FRET is a useful tool to investigate the highly dynamic nucleosome structure. So far nucleosome dynamics was addressed by FRET techniques by installing the fluorophores as fluorescent protein tags, by DNA labeling or by labeling of individually introduced cysteine residues with maleimide or a combination thereof.[142–144]

Herewith using bulk or single-pair FRET, the very dynamic organization of individual nucleosomes that modulates DNA accessibility were monitored. Including DNA breathing, transient unwrapping of nucleosomal DNA, the release of H2A/H2B dimers or sliding of the octamer relative to DNA position.[142,143,145] DNA breathing and nucleosome stability can be modulated by histone or DNA sequence and by PTMs on histone proteins.[140,144,146]

Furthermore, the complex level of chromatin compaction has been addressed using FLIM (fluorescence lifetime imaging microscopy)-FRET, created by fusions of different fluorescent proteins (GFP and mCherry) to histone H2B, and allowed for monitoring of chromatin compaction during mitosis in HeLa cells, thereby stating the maximum compaction during anaphase.[147,148]

In summary, FRET is a powerful tool to investigate protein dynamics and is frequently used to unravel dynamic changes within biomolecules, including studies addressing the nucleosome. However, no findings addressing nucleosome or chromatin dynamics were done so far by using labeling strategies based upon genetic code expansion. This provides a great variety of novel functionalities that allow for site-specific labeling of proteins, thereby expanding current limitations. However, detailed investigations to optimize labeling efficiencies based upon the conditions best suitable for individual proteins, like histones, would promote the applicability of this technique for use in FRET studies.

1.6 Aims

A huge number of UAAs has been successfully introduced onto proteins using genetic code expansion, including different PTMs, fluorophores or reactive handles for chemical derivatization. Although some PTMs could already be installed using this technique, the number of newly discovered modifications are likewise expanding. In this thesis the recently reported novel acyl PTMs should be the subjects to be installed site-specifically onto a protein of interest. This should be done by screening of a library created of *Mb*PylS mutants for their ability to recognize any of the acyl UAAs PrK, CrK or BuK, thereby enlarging accessible functionalities for genetic code expansion.

As for many of the newly reported PTMs their remodeling enzymes are unknown, to create proteins carrying those modifications would greatly help in unraveling their function. Histone proteins, the building blocks of chromatin, are the most prominent targets for PTMs. The novel acyl modifications should therefore address them as main subjects. Furthermore not only the histones H2A, H2B and H3, which were already established targets for GCE, but also H4, which was so far difficult and only unsuccessfully addressed by UAA incorporation, should be exploited as possible substrate.

Chromatin and with it the nucleosomes are very dynamically organized to provide both stability and accessibility of the DNA, with the histone tail domains playing a vital role. The stability and therewith dynamic properties are extensively regulated by various mechanisms. A good technique to investigate dynamic properties is provided by FRET investigations. However FRET requires for the installation of two fluorophores. This was previously done to investigate on chromatin or nucleosomes by combining fluorescent protein tags, DNA labeling and labeling of thiols as provided by cysteines. In this thesis, functionalities should be introduced on histones that can react in click reactions and best suitable conditions for their derivatization reactions should be identified. The combination of labeling distinct functionalities in a click reaction together with labeling of an individually placed cysteine should result in a FRET pair that provides the potential to be placed at any positions of interest throughout the histone proteins.

Recent studies enabling the incorporation of two distinct UAAs by GCE could provide an avenue to create site-defined FRET pairs (by combining one UAA and cysteine labeling) with one of the novel acyl PTM thereby creating a possibility for examining their distinct effects on nucleosome stability.

2 Results

2.1 An avenue towards new lysine modifications

During the last decade, numerous new post-translational modifications have been discovered. Among them were several lysine acyl modifications, including propionyl-, butyryl-, crotonyl-, succinyl- and malonyl-groups.[94,99–101,106] To investigate the novel acyl modifications concerning their physiological impact on proteins, it is of great importance to produce proteins that contain the modification selectively on the position of interest. One method to obtain site-specific modified proteins is genetic code expansion, a technique that incorporates an UAA of interest in response to an amber stop codon directly in the living cells by using the endogenous translation machinery, utilizing an evolved pyrrolysine synthetases (PylS) system. Therefore, a library of *M. bakeri* PylS mutants with randomized positions in the active site was created in order to evolve PylS mutants with high selectivity for the lysine acyl modifications.[174] The positions were chosen based on the crystal structure of *M. mazei* PylS.[149] The active site residues are conserved between *M. mazei* and *M. bakeri* and the chosen positions are Ala267, Tyr271, Leu274, Cys313 and Met315 (described in 4.7.1.), highlighted in Figure 11.

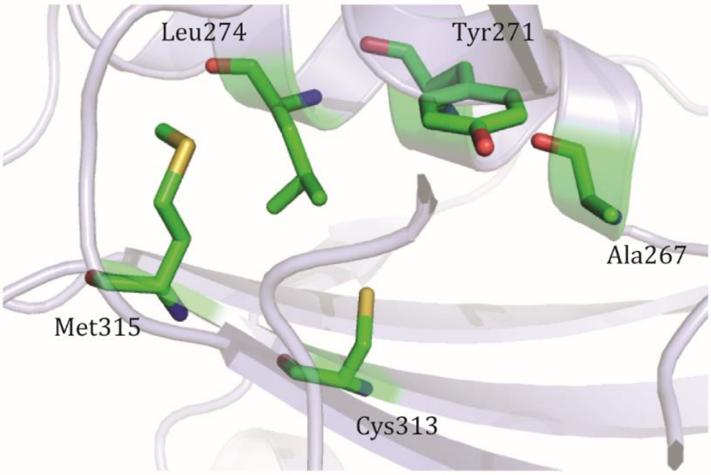


Figure 11 ***Structure of the active site of M. mazei.*** The active site residues are conserved between *M. mazei* and *M. bakeri* and for the library randomized positions are highlighted. The picture was created using PyMol, based on the crystal structure of *M. mazei,* PDB ID 2Q7H.

The library contained nearly thirty million mutants and covered a diversity of about 64%. This library was screened for mutants incorporating propionyl- or crotonyllysine by three rounds of positive and negative selection, using either chloramphenicol acetyltransferase or barnase, perturbed by amber codons, as reporter constructs (outlined in 4.7.2). The mutants were individually investigated by cotransforming again with chloramphenicol acetyltransferase reporter in *E. coli* DH10B cells and growing them in the presence and absence of PrK or CrK. After 3 h, the cells were plated on agarose plates containing increasing concentrations of chloramphenicol, again in the presence and absence of PrK or CrK. The most promising mutants were selected and the procedure repeated.

The results for the PrK selection is depicted in Figure 12. Without Cm all colonies grew independent of the presence of PrK. When Cm was present, only colonies on the positive PrK plates survived on plates containing 100 mM Cm. At a concentration of 200 mM Cm no mutants survived. The same was performed for CrK and the results are shown in Figure 13.

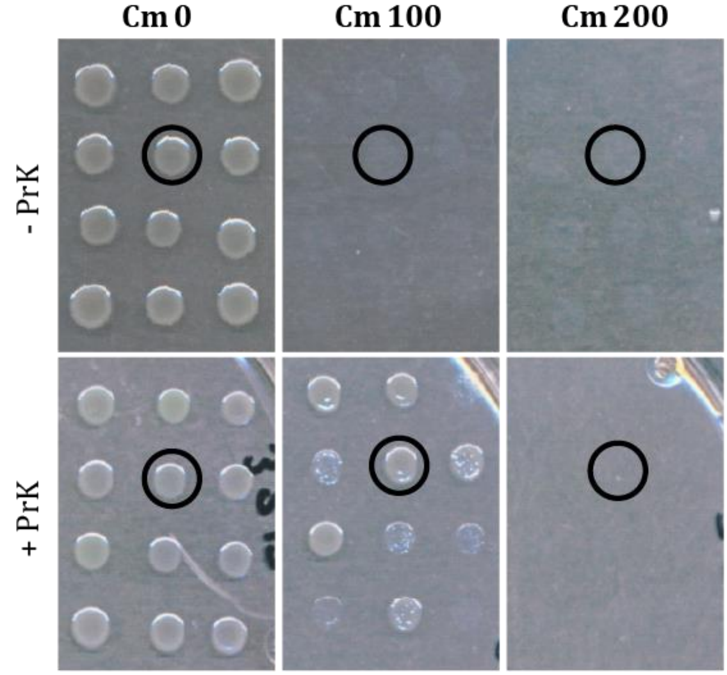


Figure 12 ***Selection of the PylS library on PrK*** After a selection including one round of positive, one negative and another positive, individual colonies were picked and used in a chloramphenicol-assay with increasing Cm concentrations. The colonies that showed no growth on plates with Cm and without PrK and that grew on plates with PrK and Cm were used for protein test expressions. The black circle marks the mutant that performed best in protein expression and is later referred to as PrKRS.

Again colonies could survive on negative CrK plates only when no Cm present. In contrast to the PrK selection, some colonies showed growth, not only on plates with 100 mM Cm, but also on plates with 200 mM Cm. The respective positive colonies from both selections were used for protein test expressions (black circles in Figure 12 indicate individual mutants that performed best during the following experiments). Verification that the selected mutant incorporates PrK or CrK and not one of the native AA, was done by a test expression using a his_6-tagged myoglobin. This expression was performed in the presence and absence of the synthetic lysine acyl modification. Proteins were isolated from cultures normalized to cell count and then separated by SDS-PAGE and analyzed by western blot using α-His antibody. Mutants displaying the strongest UAA dependent protein expression were used for further experiments (depicted in Figure 14).

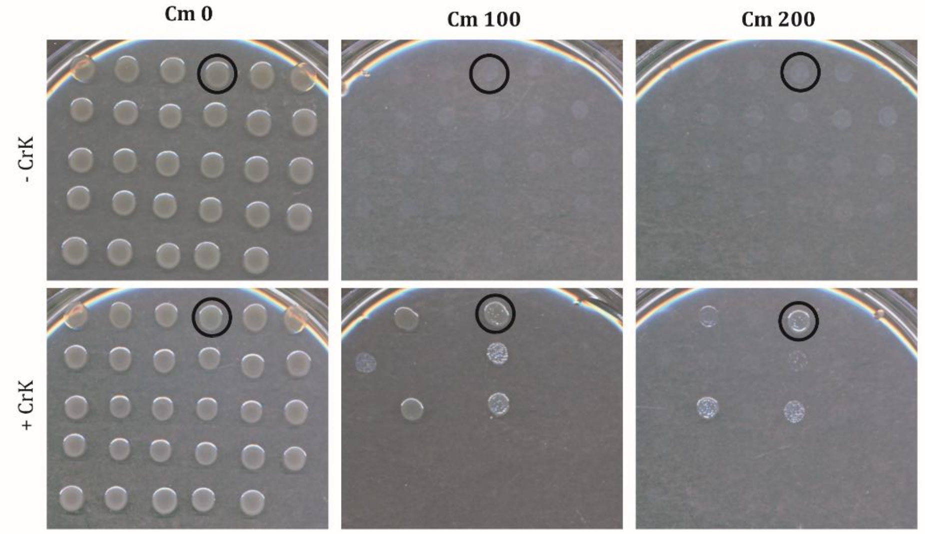


Figure 13 ***Selection of the PylS library on mutants for CrK incorporation***. After a selection-cycle including one round of positive, one negative and another positive selection. Individual colonies were picked and used in a chloramphenicol-assay with increasing Cm concentrations. The colonies that showed no growth on plates with Cm and without CrK and that grew on plates including CrK and Cm were used for protein test expressions. The black circle marks the mutant that performed best in later protein expression experiments.

The altered active sites for the evolved synthetases, as compared to the WT PylS are depicted in Table 2-1. The evolved PrKRS13 (referred to as PrKRS throughout the remainder of this work) was found to have Tyr271 replaced by Phe and Cyc313 by Thr. The evolved CrKRS6 (referred to as CrKRS throughout the remainder of this work) had two changes in the AA sequence compared to the WT. Cyc313 was replaced by Val and Met315 was replaced by Tyr.

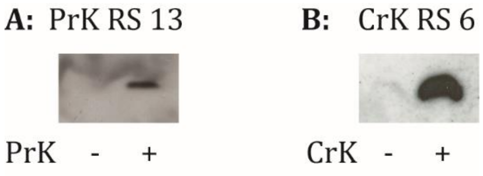


Figure 14 ***Western blot of his_6-tagged myoglobin4TAG test expression with the selected synthetases for incorporation of PrK and CrK***. The medium was either supplemented with 2 mM PrK or CrK or no UAA, and the expression run for 3 h. The samples were run on a 15% SDS-PAGE and transferred onto nitrocellulose membrane. Protein was detected by blotting against his_6-tag and visualized via chemiluminescence. A: Mutant with best incorporation efficiency for PrK and B: Synthetase with the best incorporation efficiency for CrK.

Table 2-1 ***Changes in sequence of the selected synthetases compared to WT PylS***. Changes compared to the WT PylS are indicated in red.

Wt	Ala267	Tyr271	Leu274	Cyc313	Met315
PrK RS 13	Ala	Phe	Leu	Thr	Met
CrK RS 6	Ala	Tyr	Leu	Val	Tyr

The acyl modifications have a great similarity, differing only by the number of carbon atoms in the length of the acyl chain. The WT PylS was previously used for the incorporation of different UAAs, for example the substrate analog BocK.[21,55,150] Additionally, the performance of the evolved sythetases was compared to the WT PylS and the previously by Neumann et al. evolved AcKRS3.[28] Therefore, sperm whale myoglobin containing an amber codon at position 4 was expressed in *E. coli* BL21 (DE3) cells, comparing PrK and CrK. For analysis of the incorporation efficiencies, the AcKRS3 incorporating AcK and WT PylS, possessing the very efficient substrate analog BocK, were expressed for comparison. All amino acids were used at a final concentration of 2 mM. Cells were normalized to OD and then harvested, the resulting proteins were then analyzed by SDS-PAGE and western blot using α-His antibody. For PrK incorporation the evolved PrKRS performed best although both WT PylS, AcKRS and CrKRS were able to incorporate PrK, however less efficiently (Figure 15 A). Both WT PylS and CrKRS could incorporate CrK with almost the same efficiencies, whereas PrKRS performed less efficient and AcKRS failed to successfully install the amino acid (B). Subsequently, the different synthetases were tested for their ability to incorporate BuK (C). Here, the WT PylS performed most efficiently, the CrKRS performed also well, however the AcKRS and PrKRS failed to work. Furthermore the evolved AARS were tested upon their ability to perform on AcK (D). WT PylS, CrKRS and PrKRS showed no signal above background and approved the evolved AcKRS as only functional AARS for AcK incorporation. Thus, the best efficiency for this AARS is achieved when using the amino acid in a concentration of 10 mM (indicated by *). Accordingly, in the following experiments PrKRS was used for the incorporation of PrK. The WT PylS was utilized for BuK and CrK incorporation and AcKRS was used to perform with AcK hence in a concentration of 10 mM.

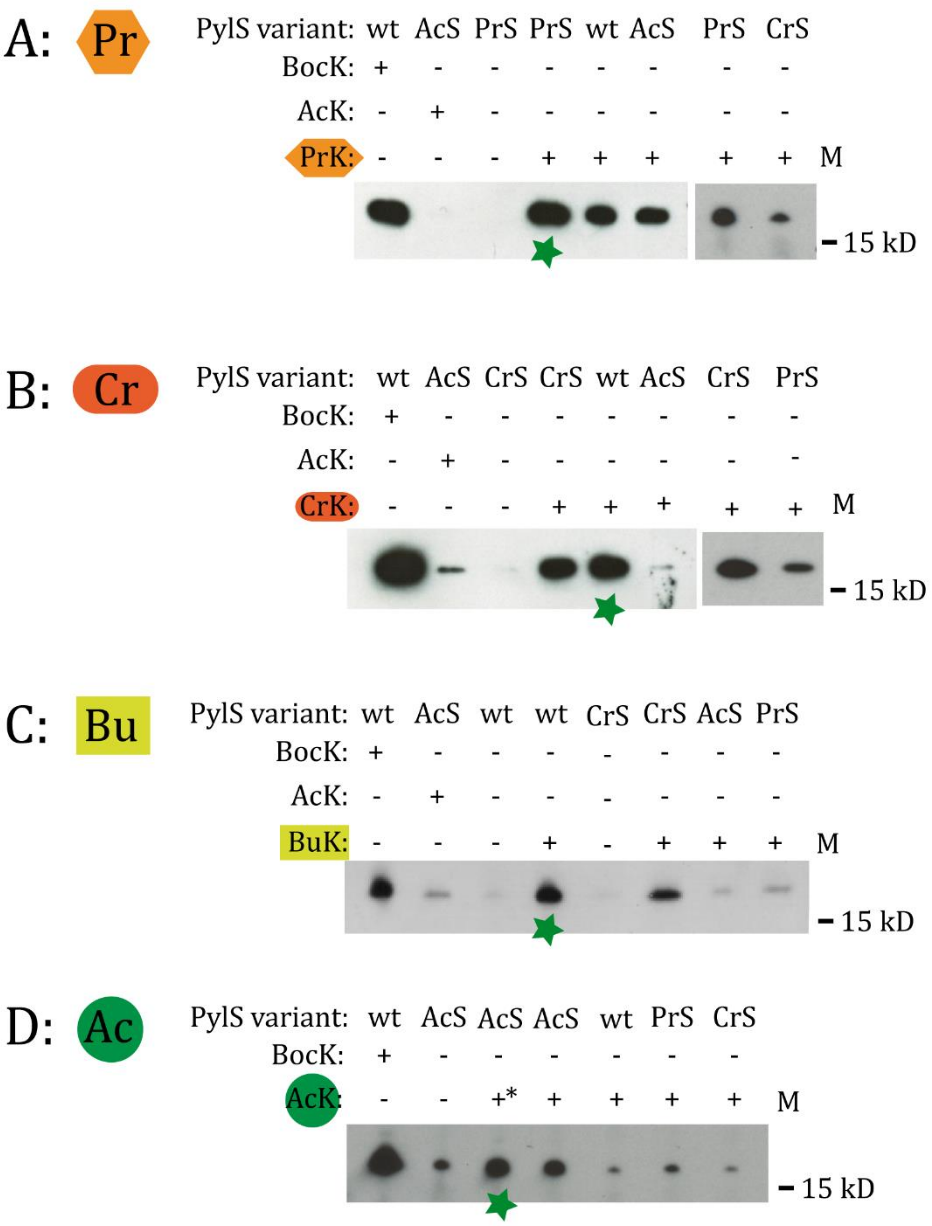

A: Pr
PylS variant: wt AcS PrS PrS wt AcS PrS CrS
BocK: + - - - - - - -
AcK: - + - - - - - -
PrK: - - - + + + + + M
15 kD
B: Cr
PylS variant: wt AcS CrS CrS wt AcS CrS PrS
BocK: + - - - - - - -
AcK: - + - - - - - -
CrK: - - - + + + + + M
15 kD
C: Bu
PylS variant: wt AcS wt wt CrS CrS AcS PrS
BocK: + - - - - - - -
AcK: - + - - - - - -
BuK: - - - + - + + + M
15 kD
D: Ac
PylS variant: wt AcS AcS AcS wt PrS CrS
BocK: + - - - - - -
AcK: - - +* + + + + M
15 kD

Figure 15 ***Specificity of the synthetases was tested by expression of his$_6$-tagged myoglobin4TAG with different combinations of synthetases and amino acids.***[174] The efficiency of the synthetases was monitored by western blot analysis against his-antibody. Concentration of the amino acids was 2 mM. A normalized OD 2 was harvested. The green asterix indicates the respective most efficient synthetase, which was used for further expressions. A: Selection for PrK incorporation. B: Selection for CrK incorporation. C: Selection for BuK incorporation. D: Test for AcK incorporation (*) indicates a concentration of 10 mM. AcS:Acetyllysine synthetase; PrS:Propionyllysine synthetase; CrS:Crotonyllysine synthetase.

Homogeneous populations of AcK, PrK, CrK and BuK were validated by analyzing purified Myo4mod protein (4.5.3) via ESI-TOF measurements (described in 4.8.1, measurements performed by Dr. H. Frauendorf).[174] These results are highlighted in Figure 16, verifying the specific incorporation of all four acyl modifications on myoglobin.

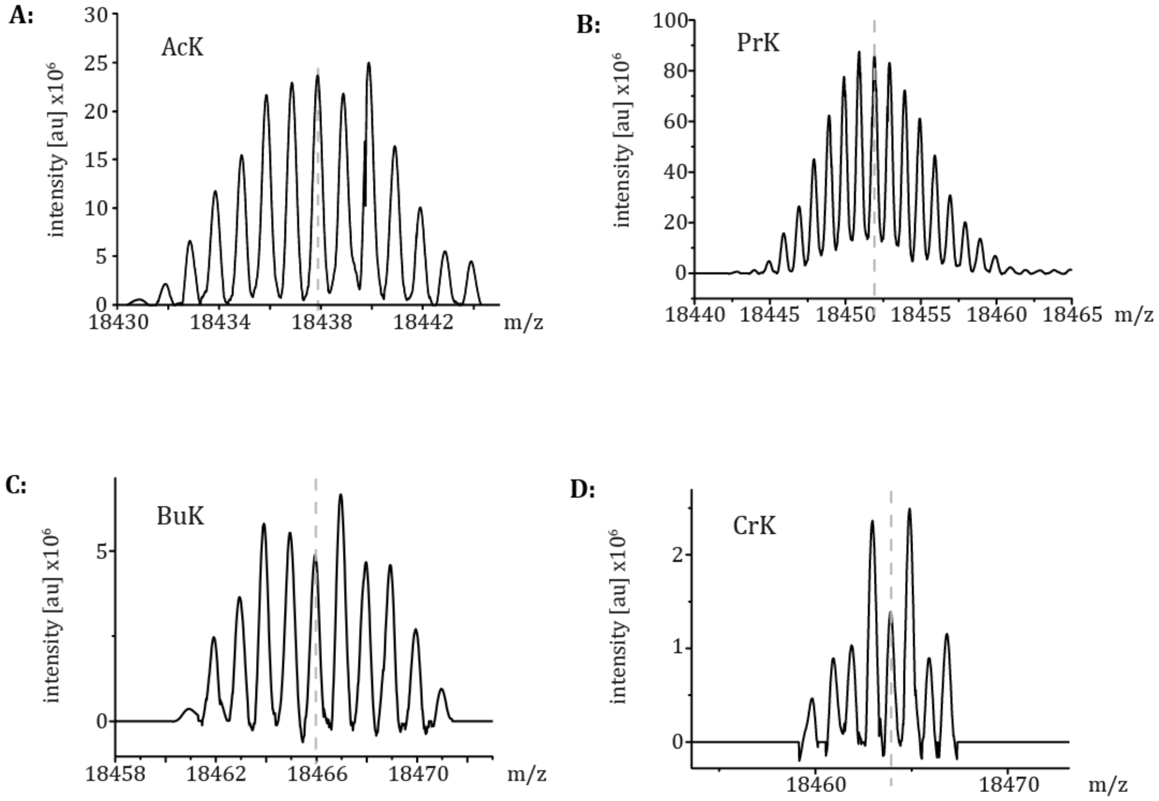


Figure 16 ***Analysis of acyl modified myoglobin by ESI-TOF.*** The incorporation of the acyl modifications respectively, was verified by analyzing purified Myo4mod by ESI-TOF. A: Myo4Ac Expected mass: 18438 Da, found mass: 18438 Da. B: Myo4Pr expected mass: 18452 Da, found mass 18452 Da. C: Myo4Bu expected mass 18466 Da, found mass 18466 Da. D: Myo4Cr expected mass 18464 Da, found mass 18464 Da.

With the evolved synthetases for the incorporation of the novel acyl UAAs, we could next address the histone proteins as target.

2.1.1 Incorporation of acyl modifications on histones

2.1.1.1 Acyl modified histone H3

Histones are proteins which are highly regulated by PTMs.[71,151] Therefore, with the synthetases at hand, we decided to produce histones, which carry the desired acylation at specific positions. Histone H3 is an established model for the incorporation of UAAs, and therefore the protein of choice to verify UAA incorporation on a histone.[28,29] Acylations are reported to occur on histone H3 tails, including acetylation at K9, K14, K18; crotonylation at positions K4, K9 and K18, butyrylation at K14 and K27 and propionylation at K23 (summarized in Figure 17).

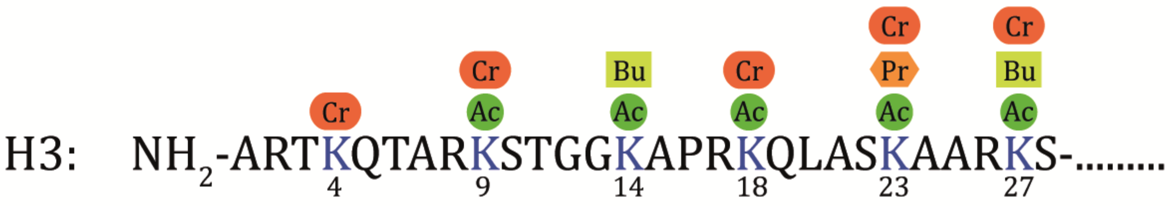


Figure 17 ***Acyl modifications on the histone H3 tail.*** Summarized from publications by Tan et al.[100] and Albaugh et al.[152]

First, PrK and CrK incorporation was performed on histone H3 at positions K9 and K14 respectively and compared to AcK at the same positions (4.5.5). This results in the production of histone H3 with different acylation at the respectively indicated position with good yields (data not shown). In contrast to similar expressions performed with this system, NAM was first not utilized to prevent the acyl group from being cleaved by CobB, as previously reported by Neumann et al. for the acetylation.[17] The proteins were purified as described in 4.5.5 and total protein was analyzed by ESI-MS (4.8.1). The results showed correct masses for the incorporation of acetylation (performed in the presence of NAM, as previously described), but propionylation and crotonylation were partially, to almost completely cleaved when produced without the CobB inhibitor (data not shown). Thus, NAM was further used in all expressions of acyl modified proteins to inhibit the acyl group cleavage by CobB.

2.1.1.2 Acyl modified histone H4

The production of modified histones H2A, H2B and H3 has been widely applied whereas the production of histone H4 with UAAs, following the same procedure, has been difficult and has yet not yield detectable amounts of protein. Since the novel acyl modifications are highly prominent on the histone H4 tail (depicted in Figure 18), we decided to optimize genetic code expansion suppression techniques for H4 expression and, in turn, acyl modifications on histone H4.[174]

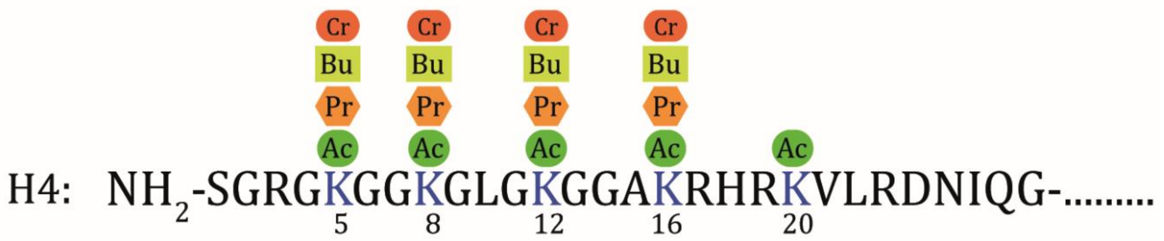


Figure 18 ***Acyl modifications on histone H4 tail.*** Acylations, summarized from publications by Tan et al. and Albaugh et al.[100,152]

One hypothesis as to why the production of H4 using genetic code expansion has been difficult is that H4 might be less insoluble than the other core histones and therefore being degraded when expressed in only low amounts. In an effort to make H4 more insoluble an expression vector was constructed that fused the histone H4 gene to a histone H3 gene, because H3 is a well characterized insoluble protein.

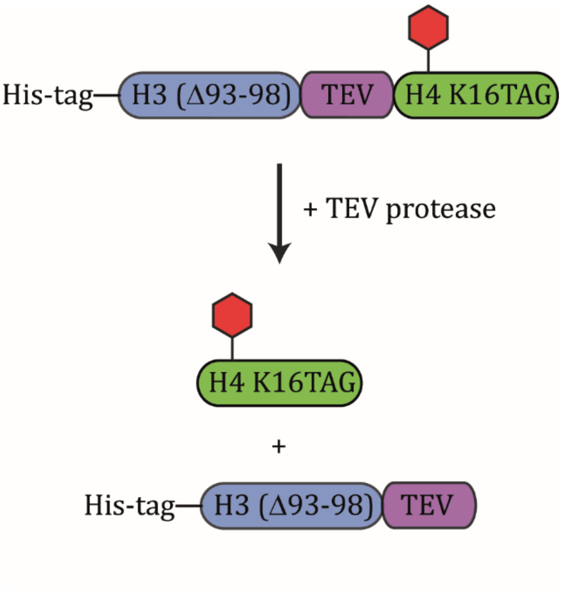


Figure 19 ***Overview construction of histone H4 including an UAA.*** The construct used for expression of H4 histone is a his$_6$-tagged fusion protein that combines a shortened H3 (Δ93-98) and the H4 protein containing an amber codon at position K16 via a TEV cleavage side. After expression and his$_6$-purification the desired H4 K16mod is obtained by TEV cleavage of the construct. The truncated H3 does not refold into octamers and therefore does not require further separation from H4 product.

The construct was created by Dr. Bryan Wilkins on a pCDF PylT backbone as described in 4.2.14. It contains an N-terminal his$_6$-tag and a TEV cleavage site inserted between the two ORFs. Furthermore, the H3 gene has deletions of the codons encoding for amino acids 93-98. This was done to prevent the H3 TEV cleaved product from assembling into octamers during reconstitution reactions (scheme depicted in Figure 19). An amber codon was placed at position K16 on H4 for incorporation of the acyl modifications.

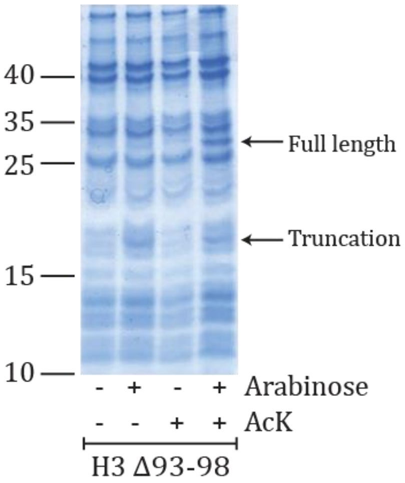


Figure 20 ***Test expression of the H3(Δ93-98)-TEV-H4 K16TAG construct.*** Expressions were run for 3 h with 10 mM AcK, 20 mM NAM and induced with 0.2% arabinose. Whole cell lysates were analyzed by 15% SDS-PAGE. The full length protein was only expressed after induction and when AcK was supplemented to the medium (experiment performed by Bryan Wilkins).[174]

Test expressions in *E. coli* BL21 (DE3) cells showed many impurities and needed further improvement to give valuable amounts of protein (data not shown). For that purpose, an amber depletion cell line was used. This cell line required a replacement of the T7 promoter by an ara*BAD* promoter (creation of the plasmid described in 4.2.14). Test expressions of this construct in *E. coli* C321.ΔA.exp[153] cells together with the AcKRS synthetase showed, that the full length protein is formed only when AcK is present in the media and protein expression is induced (Figure 20). To prevent the acetyl group from being cleaved by CobB, NAM is used as an inhibitor in the expressions. Different amino acid depletions were tested but the Δ93-98 proofed best and was further used. (Data not shown).

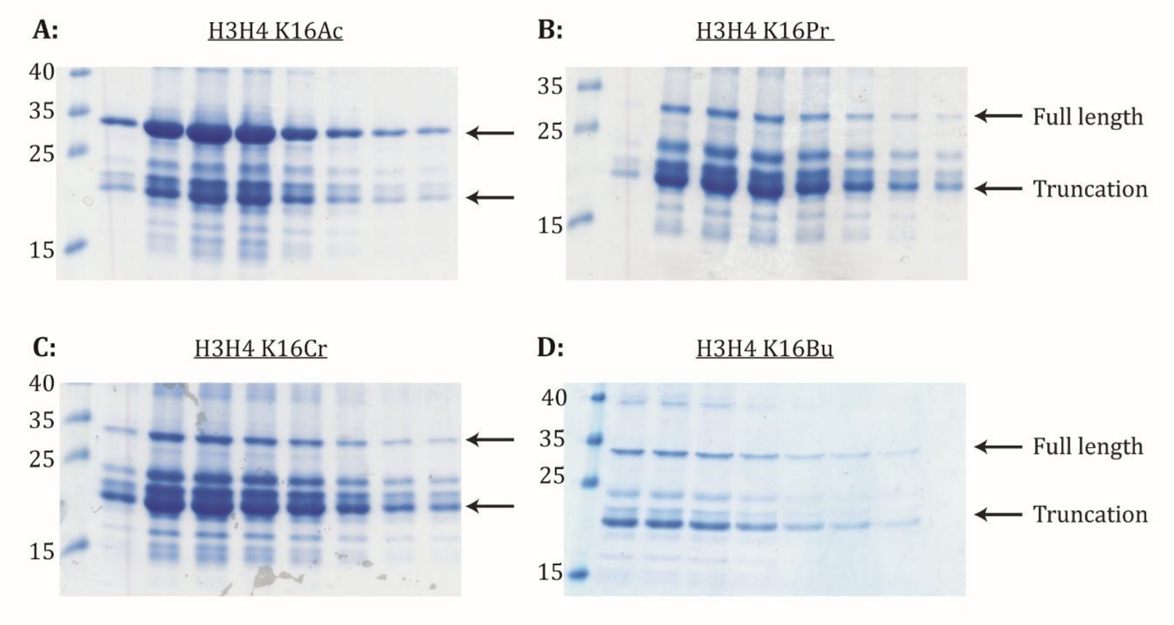


Figure 21 ***Purified His$_6$-H3(Δ93-98)TEV-H4 K16mod.*** The upper band is the full length protein containing either one of the modifications, whereas below is the truncation when translation terminated at the stop codon. A: H3H4 K16Ac purified from 2 L); B: H3H4 K16Pr (purified from 2 L); C: H3H4 K16Cr (purified from 2 L); D: H3H4 K16Bu (purified from 1 L).

The described construct was used together with the evolved synthetases for the direct incorporation on histone H4 containing either one of the acyl modifications acetyl, propionyl, crotonyl or butyryl at position K16. The proteins were expressed and purified as described in 4.5.4. SDS-PAGE of eluted fractions from the purified H3H4 fusion proteins are depicted in Figure 21. The truncation, resulting from protein expression termination at the amber stop codon, is the lower band (~19 kDa) whereas the full length protein is present in the upper band (~29 kDa). Fractions containing full-length protein were pooled, dialyzed against βME and then TEV cleaved. The TEV cleavage is depicted in Figure 22, showing the ability to produce individual histone H4 proteins including the different acyl modifications. Typical yields of H4mod were around 1 mg protein from 1 L expression culture, except levels for propionylation were found to be lower.

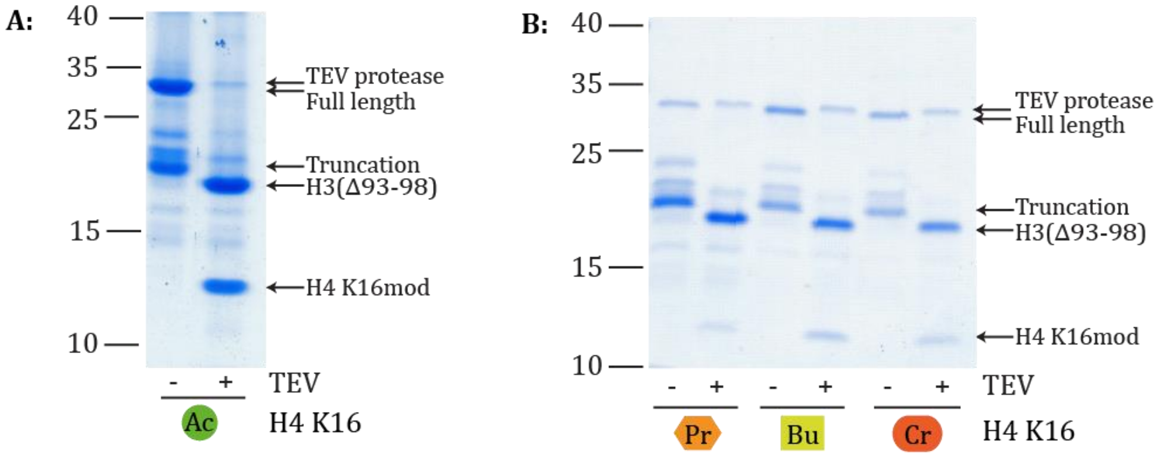
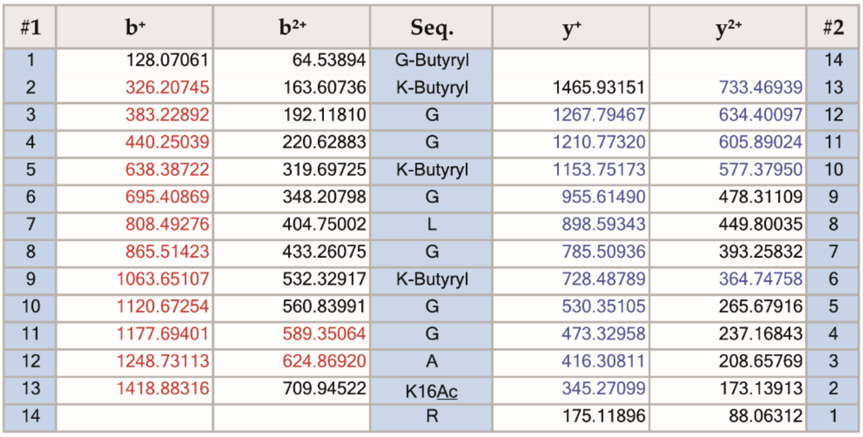

A:
40
35
25
15
10
TEV protease
Full length
Truncation
H3(Δ93-98)
H4 K16mod
- +
TEV
Ac
H4 K16
B:
40
35
25
15
10
TEV protease
Full length
Truncation
H3(Δ93-98)
H4 K16mod
- + - + - +
TEV
Pr
Bu
Cr
H4 K16

Figure 22 ***TEV cleavage of the purified H3(Δ93-98)-TEV-H4mod*.** 15% SDS-PAGE of the H3H4 constructs before and after TEV cleavage are visualized. A: TEV cleavage of H3H4 K16Ac; in B: TEV cleavage of H3H4 with Pr, Bu and Cr are depicted.

The K16 acetylated histone H4 was analyzed by mass spectrometry. However, analysis of tryptic fragments of histone proteins is difficult to achieve, as the high abundance of lysine and arginine residues generates very short, irreproducible fragments that are difficult to analyze.[154] One method to obtain longer tryptic fragments was published by Garcia et al. They used chemical derivatization of lysine residues, thereby preventing trypsin cleavage at the lysine residues and resulting in fragments only from C-terminal Arg cleavage.[155] We applied this method to analyze the purified histone H4 with acetylation at K16, after chemical butyrylation of remaining lysine residues (4.8.2.1). Following trypsin digestion, resulting peptide ends were again butyrylated and ESI-MSMS2 analyzed by Dr. O. Valerius (4.8.2.2). The obtained chromatogram is depicted in Figure 23, showing successful site-specific acetylation at position K16. The details for the resulting fragment ion series are listed in Table 2-2.

With the here described method we were able to establish the installation of novel acyl lysine UAAs using genetic code expansion on histone H4.

Table 2-2 ***H4 K16Ac peptide data from ESI-MSMS2.*** All lysine residues and all N-terminal peptide ends were chemically butyrylated.



#1	b^+	b^{2+}	Seq.	y^+	y^{2+}	#2
1	128.07061	64.53894	G-Butyryl			14
2	326.20745	163.60736	K-Butyryl	1465.93151	733.46939	13
3	383.22892	192.11810	G	1267.79467	634.40097	12
4	440.25039	220.62883	G	1210.77320	605.89024	11
5	638.38722	319.69725	K-Butyryl	1153.75173	577.37950	10
6	695.40869	348.20798	G	955.61490	478.31109	9
7	808.49276	404.75002	L	898.59343	449.80035	8
8	865.51423	433.26075	G	785.50936	393.25832	7
9	1063.65107	532.32917	K-Butyryl	728.48789	364.74758	6
10	1120.67254	560.83991	G	530.35105	265.67916	5
11	1177.69401	589.35064	G	473.32958	237.16843	4
12	1248.73113	624.86920	A	416.30811	208.65769	3
13	1418.88316	709.94522	K16Ac	345.27099	173.13913	2
14			R	175.11896	88.06312	1

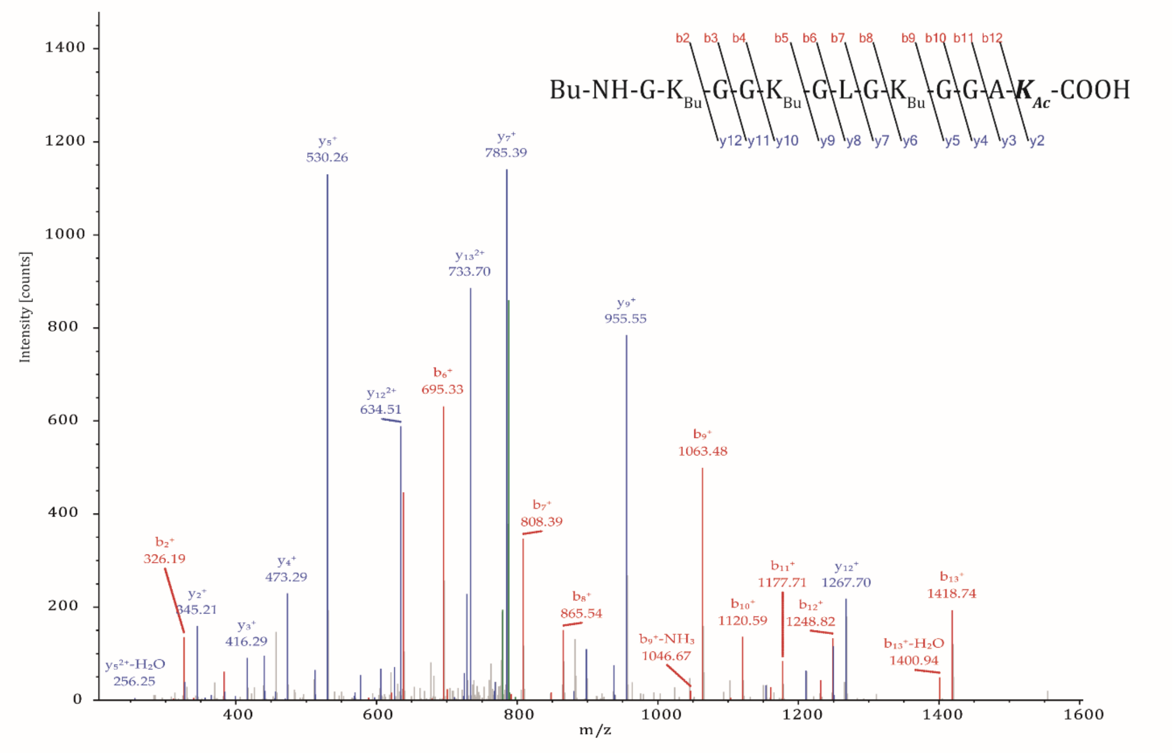


Figure 23 ***Analysis of tryptical fragments from H4 K16ac by mass spectrometry.*** All lysine residues (except for K16) were chemically butyrylated using butyryl-anhydride. Before analysis by mass-spectrometry, the butyrylated histone H4 was trypsin digested, resulting ends again butyrylated and fragments analyzed. The obtained spectrum is depicted plotting the charge in [m/z] against the intensity in [counts]. Red labelled peaks represent the fragment-ions from the b-series, whereas blue peaks indicate fragment-ions from the y-series.

2.1.1.3 Antibody recognition and specificity towards acyl modifications

Antibodies are available not only as general pan antibodies to recognize acetylation but also towards acetylation at certain positions, like H4 K16Ac. Furthermore antibodies are not limited to acetyl, but more recently also available for other acyl modifications, for example crotonyl groups, hence recognizing the modification without specificity towards one particular position. As the discrimination between the different acyl modifications is difficult regarding their structural similarity, questions arise as to if the antibodies designed to recognize only one specific acyl modification, can cleanly differentiate between the acyl modifications.

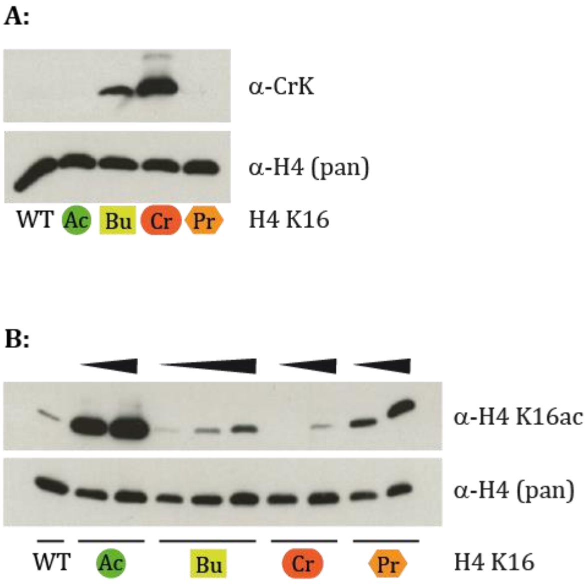


Figure 24 ***Antibody specificity test.*** Purified histone H4mod were separated by SDS-PAGE, wet blotted on PVDF membrane and the membrane incubate with either α-CrK antibody or α-H4K16ac antibody and additionally α-H4 as control and bands were visualized by chemiluminescence. The crotonyl antibody recognizes besides the crotonoylation the butyrylation but less strong. The α-H4K16ac antibody detects the acetylation, the propionylation weaker and the butyrylation poorly but not the crotonylation.

This was addressed by assaying the acyl modified histone H4 K16 proteins by western blot and then detecting the modifications with antibodies.[174] The antibodies tested included α-H4, α-H4K16ac and α-CrK. (details listed in Table 4-6). Histones H4 K16mod were separated on 15% SDS-PAGE, transferred onto PVDF membrane by wet blotting and incubated with either α-H4 (pan)-, α-H4K16ac- or α-CrK antibody and detected by chemiluminesence. The resulting western blots are depicted in Figure 24 (for the final figure presented here, WB and detection were performed by Svenja Heitmüller). Part A represents the comparison of a α-

CrK antibody for the different acyl modifications and a α-H4 control. The α-CrK antibody recognizes the crotonyl modification nicely, the butyrylation weaker, and not the acetyl or propionyl groups. Part B details the same comparison of modifications, thus using a α-H4K16ac antibody. This antibody most efficiently recognizes the acetylation. The propionylation is recognized less efficiently and the butyrylation quite poorly. The crotonylation is not detected with this antibody.

To exclude false positives due to some intracellular processes, like enzymatic interconversions, of the modification, producing a mixture of butyryl and crotonyl, the antibodies were tested on chemically modified H4 proteins. Therefore WT histone H4 was chemically altered with propionyl, crotonyl or butyryl groups, using the respective anhydrides for acylation of lysines throughout the whole H4 protein, resulting in extensively altered H4mod (depicted in Figure 25).

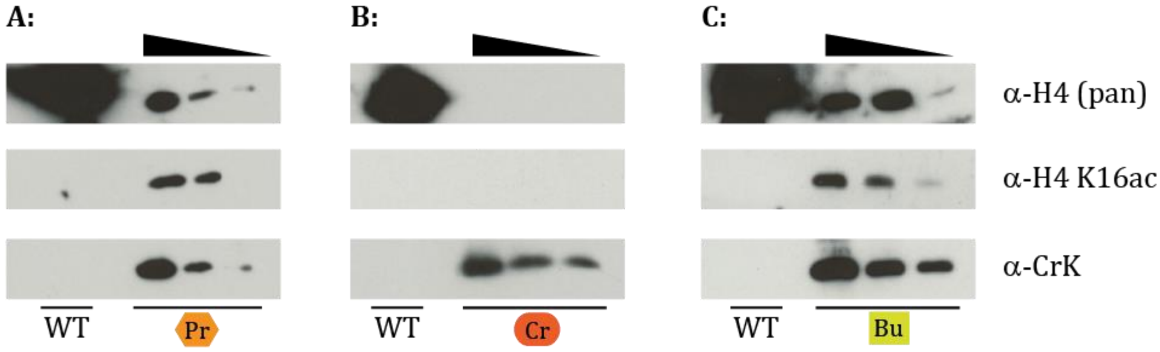


Figure 25 ***Antibody specificity test using chemically altered H4mod.*** Purified histone H4 WT was poly-acylated using either (A) propionyl-, (B) crotonyl-, or (C) butyryl anhydride. The obtained acylated histone protein was separated by 15% SDS-PAGE and transferred onto PVDF membrane. After incubation with α-H4 (pan), α-H4K16ac and α-CrK antibodies, signals were visualized by chemiluminescence. The α-CrK antibody recognizes all three chemically highly altered acyl modified histone H4 proteins. The α-H4K16 antibody recognizes acetyl and propionyl modified protein and the α-H4 (pan) antibody recognizes the highly modified proteins much weaker as compared to the not modified protein.

The method is outlined in 4.4.8. In the final step, dried samples were dissolved and separated by SDS-PAGE and transferred onto PVDF membrane by wet blot. Membranes were incubated with α-CrK, α-H4K16ac and α-H4 and visualized via chemiluminescence. The α-H4K16ac antibody again recognizes propionyl and butyryl groups and the α-CrK recognizes crotonylation, butyrylation and to some degree propionylation, when extensively altered.

This indeed indicates a certain substrate tolerance of the antibodies and in further experiments we therefore used the α-H4K16ac antibody to recognize acetyl, propionyl and butyryl and α-CrK for crotonylation.

2.1.1.4 Acylated octamers

The H3 truncation that results from the production of histone H4 via H3 fusion protein after TEV cleavage was not removed by an additional purification step. Whether an additional purification step was required was investigated by collecting the precipitate that formed during reconstitution of octamers by dialysis into high-salt buffer. The precipitate and fully purified octamers, after size-exclusion chromatography, with and without acetylation at position K16, were run on 15% SDS-PAGE and either stained by Instant Blue or transferred onto PVDF membrane and incubated with α-H4K16ac antibody and visualized by chemiluminescence. The results are depicted in Figure 26 and show that the truncation is fully precipitated and removed by ultrafiltration and does not reconstitute into octamers.[174]

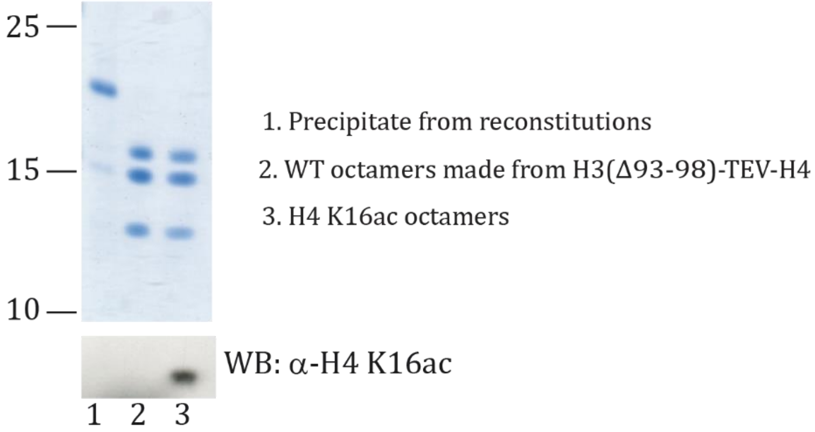

25 —
15 —
10 —
1. Precipitate from reconstitutions
2. WT octamers made from H3(Δ93-98)-TEV-H4
3. H4 K16ac octamers
WB: α-H4 K16ac
1 2 3

Figure 26 ***Effect of the precipitated H3 truncation on the formation of histone octamers***. The precipitate forming during dialysis when reconstitution of octamers including H4, produced using the H3H4 fusion protein. Compared are the collected precipitate from the refolding with the purified reconstituted octamers with and without acetylation at K16. The samples were run on 15% SDS-PAGE and either stained by instant blue or transferred onto PVDF membrane, incubated with α-H4 K16ac antibody and visualized by chemiluminescence. The precipitated H3 truncation is fully removed by precipitation and the followed size-exclusion chromatography step (picture provided by Bryan Wilkins).

Thus, we were able to produce octamers (as described in 4.6.1) that were cleanly acylated at position K16. After salt dialysis, the reconstituted octamers were separated from additionally formed dimers by size-exclusion chromatography. An example of a chromatogram from a typical reconstitution and the purification of octamers including H4 K16bu is depicted in Figure 27. Fractions A6 and A7 showed the proper ratios of each histone protein and were collected and combined. Similarly, octamers containing the H4 K16 acylations: acetyl, propionyl, butyryl and crotonyl were obtained in the same manner. (The results for the other acylated octamers are shown in chapter 8.1.)

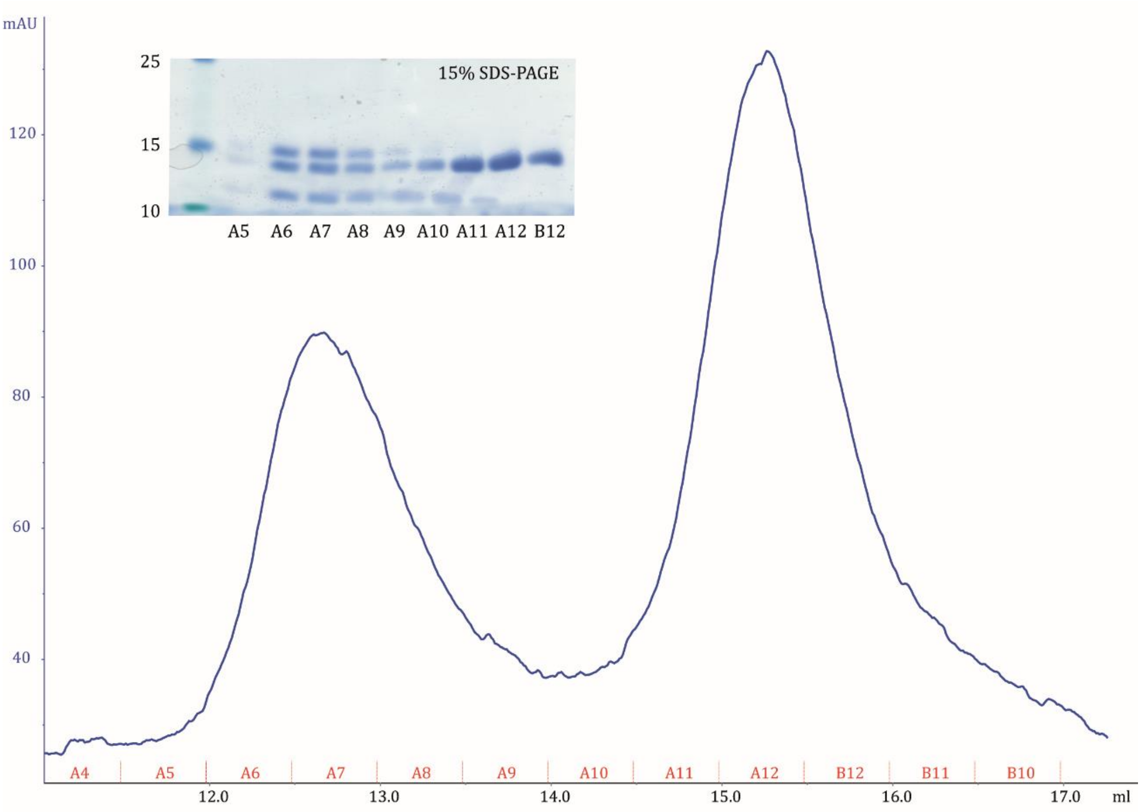


Figure 27 ***Purification of refolded octamers including H4 K16Bu*.** The by dialysis against high-salt buffer refolded octamers, including H4 K16bu were purified by size-exclusion chromatography using a S200 column. The resulting chromatogram depicts the volume in ml on the x-Axis and the absorption at 280 nm at the y-axis. Collected fractions are shown in red. The peak at around 12 ml represents the desired octamers whereas at volumes of around 15 ml dimers are eluted from the column. The small picture shows the obtained fractions separated by 15% SDS-PAGE and visualized by Instant Blue staining. Fractions A6 and A7 were pooled and used in further experiments.

The acylated octamers, including a WT octamer were separated by 15% SDS-PAGE, transferred onto PVDF membranes and the specificity of antibodies towards the different acylation was investigated (compare Figure 28), as described above when tested on histone H4 alone and on chemically modified histone H4. The results clarify that the α-H4K16ac antibody recognizes, not only the acetylation, but also the butyrylation and, to lesser extent propionylation. Again, the crotonylation is not recognized. The α-CrK antibody gives a strong signal for crotonyl but recognizes butyrylation only poorly when used in low concentrations (1:20 000 compared to 1:4 000 in chapter 2.1.1.3).

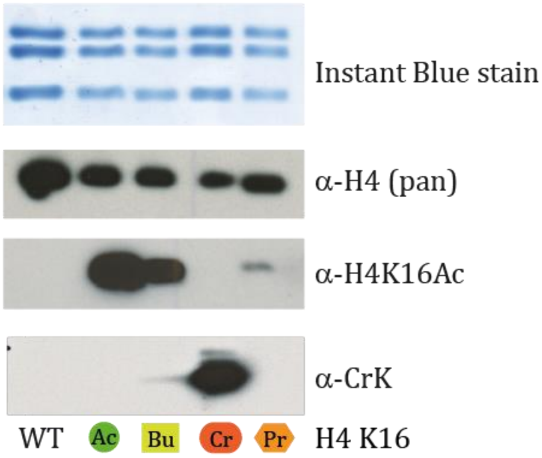


Figure 28 ***Recognition of acyl modified octamers by different antibodies.*** Octamers including H4 K16mod were separated by 15% SDS-PAGE and analyzed with antibodies specific for α-H4, α-H4K16ac and α-CrK. The Instant Blue stain shows the typical pattern for all octamers independent of the acyl modification. The α-H4 blots shows signals for the H4 proteins prominent for WT and all acylated H4s. The α-H4K16ac antibody recognizes acetylation, butyrylation and propionylation, though the latter one weaker. The antibody specific to α-CrK recognizes crotonylation and butyrylation but only poorly.

This sections shows, that the produced modified H4 was successfully used for octamer refolding and required no further purification step. The modifications were verified by WB with the different antibodies and the previously stated substrate tolerance of the different antibodies was confirmed.

2.1.1.5 Impact of acyl modifications on nucleosome array compaction

Nucleosomal arrays can be induced to self-associate in a salt-dependent manner, demonstrating the role of histone tails for chromatin higher ordered folding events.[156] The histone tails mediate DNA binding, with contributions of all four core histone N-terminal tails in this event. The largest contributions come from the H3 and H4 tails.[157,158] Lysine acetylation of histone tails is generally associated with active transcription, as acetylations neutralize the lysine positive charges and abrogate tail function by altering interactions with the DNA.[159] Along with that, lysine acetylation inhibits self-association of nucleosome arrays as they perturb interactions of the positively charged residues in the tail with acidic patches from other histones.[160,161] This is commonly occurring between the H4 tail and the acidic patch of H2A,[56,59] thereby causing the need for higher ionic strength to make precipitation occur. With the acetylation proven to strongly influence the tertiary structure, the effects of other acyl modifications still needs to be elucidated.

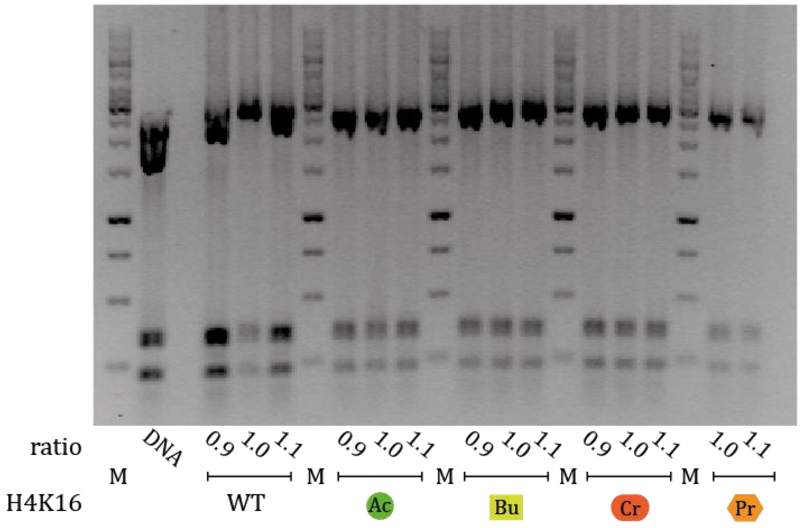


Figure 29 ***Nucleosome arrays assembled from H4 K16mod octamers***. For the nucleosome assembly including the differently acyl modified H4 proteins, different ratios of DNA:Octamer were used. The different octamers were mixed with DNA and assembled by slow exchange of high-salt with low-salt buffer via dialysis using a gradient. The resulting nucleosomes were run on a 0.5% agarose gel and DNA was visualzed by UV absorption after GelRed staining. Scavenger DNA runs on the lower part of the gel whereas nucleosome arrays run higher than the 12x200 bp DNA band.

Nucleosome arrays were prepared as described in 4.6.2 using 12x200 bp DNA, which was compounded from the plasmid pUC18_12x200x601 (kindly provided by Daniela Rhodes) as outlined in 4.2.13. The DNA already contained 147 bp scavenger DNA. The nucleosomes were obtained using the octamers including H4 K16mod by titrating octamer to DNA and assembled by slowly exchanging high-salt with low-salt buffer via dialysis. Figure 29 depicts the nucleosome arrays, which were run on 0.5% agarose gel and visualized via the DNA bands by GelRed staining. Arrays were compared to free DNA, indicating larger gel shifts (slower mobility) for the readily assembled arrays compared to the free 12x200 bp DNA. The arrays resulting from the titration ratio of 1:1 DNA:octamer were further validated by digesting them with MNase. MNase digests the arrays in individual nucleosomes creating a ladder of evenly spaced nucleosomal lengths. Before the beginning of the MNase digest, only the DNA bands for the nucleosome array and the scavenger DNA is visible. While proceeding, the digestion bands of the array decrease and a “ladder” of signals, referring to different cutting stages of the arrays, occurs, finally, resulting in mononucleosomes (Figure 30).

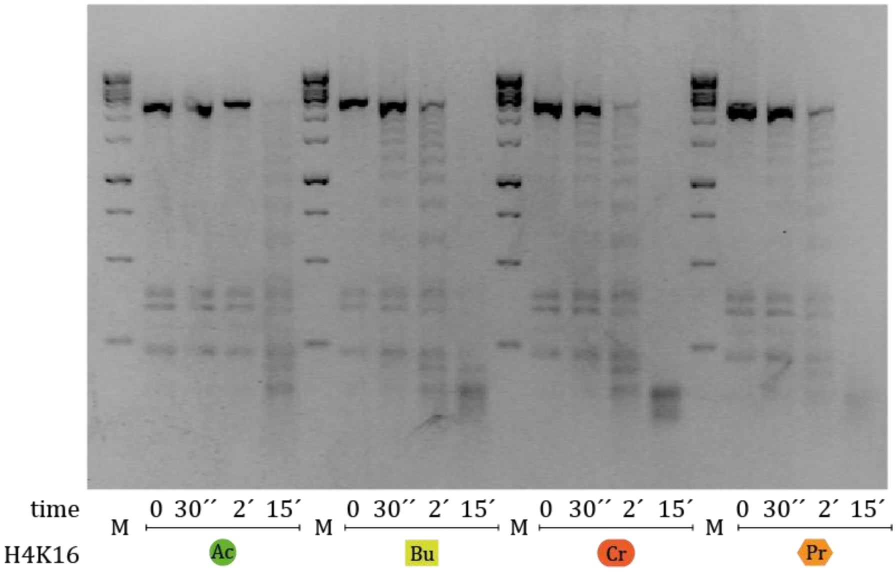


Figure 30 ***MNase digest of H4mod nucleosome arrays.*** The different acyl modified nucleosome arrays formed using a DNA:octamer ratio of 1:1 were analyzed by MNase digest. The arrays were incubated with MNase and samples collected after 30 sec, 2 min and 15 min and compared to a sample without addition of MNase. The collected samples were quenched by addition of buffer from the gel purification kit and heated for 10 min to 95 °C. The samples were run on a 2% agarose gel and visualized by GelRed. After 15 min, the MNase had cut the nucleosome arrays into mononucleosomes.

The compaction assay is based on salt-induced self-association and was done by mixing nucleosome arrays with increasing Mg^{2+} concentrations. Precipitations were removed from the reactions by ultracentrifugation before measuring the amount of array left in solution via absorption at 260 nm (experimental details described in 4.6.5). The results from this assay compare the influence of acyl modifications to the effect of acetylation is indicated in Figure 31. For the WT nucleosome arrays, 50% precipitation is occurring at 1.5 mM Mg^{2+}. For the acetylation at K16 this value is shifted to 4.5 mM Mg^{2+}. Nucleosome arrays that were either propionyl-, butyryl- or crotonylated at K16 show a value of 50% precipitation around 3 mM Mg^{2+} and are between the strong effect caused by acetylation and the WT arrays. Nevertheless, the observed slope is much steeper for WT and K16 acetylated arrays compared to the other acyl modified arrays.

In summary, here we show that the acylated octamers can be used to reconstitute nucleosomal arrays. The arrays were verified by MNas digest and we started out to elucidate the specific functions of the novel acyl modifications on formation of higher order structures, indicating an effect similar to the acetylation though less strongly.

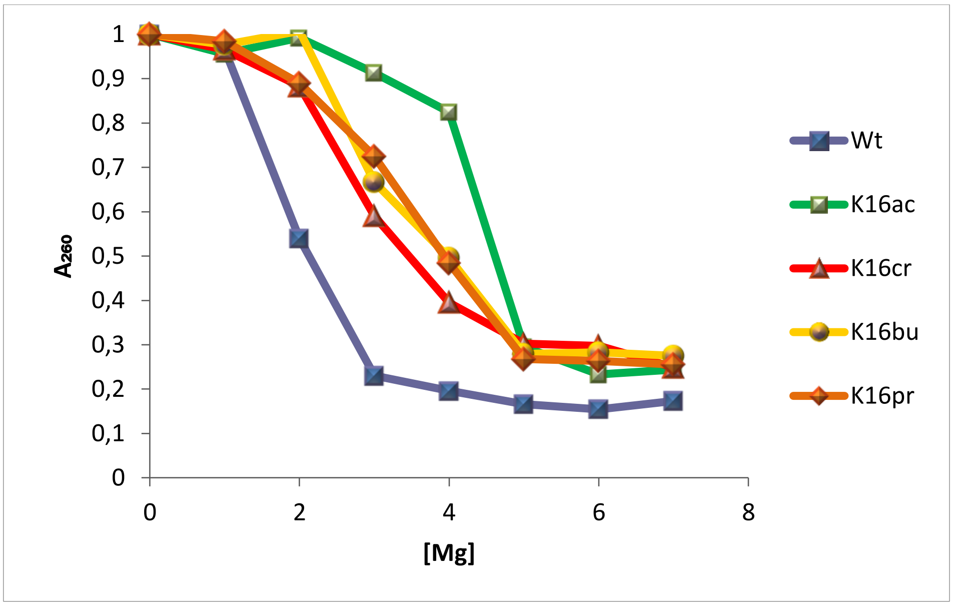


Figure 31 ***Nucleosome self-association assay*.** Different nucleosome arrays containing either one of the acyl modifications were precipitated by addition of increasing concentrations of Mg^{2+}. After addition of Mg^{2+}, samples were incubated for 20 min at RT and precipitated material was removed from reaction by centrifugation. Remaining soluble arrays were measured at A260 nm via Nanodrop.

2.1.1.6 Removal of novel acyl modifications by HDACs

The acetylation of proteins is regulated by acetyl transferases (HATs) and histone deacetylases (HDACs). The bacterial HDAC is CobB.[112] CobB is reported to remove acetyl groups and can be inhibited by NAM.[162,163] In chapter 2.1.1.1, we could already show, that CobB lacks a certain specificity to discriminate against longer acyl chains and hence addresses other modifications when inhibition of CobB was deficient.

We wanted to investigate the catalytic differences of deacylation efficiency of CobB on full length histone H4, containing the acyl modification at position K16. This project was done as collaboration with Svenja Heitmüller. She prepared the plasmids needed for CobB expression and established the expression and purification of CobB, as described in 4.5.8. The eluted fractions were run on 12% SDS-PAGE and visualized by Instant Blue staining (depicted in Figure 32). The resulting purified protein fractions were pooled and stored for usage.

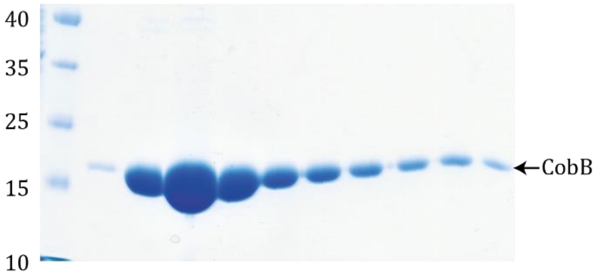


Figure 32 ***Purified CobB***. The purified CobB fractions were separated on 12% SDS-PAGE and stained with Instant Blue. Fractions containing pure protein were pooled and stored for usage.

At first, the activity of CobB towards the different acyl modification on full-length H4 proteins was addressed. Therefore acyl modified protein was mixed with CobB using a 30x molar excess of the histone protein. Samples were collected before addition of CobB and after incubation with CobB for 20 min. Further activity of CobB was quenched by addition of 50 mM NAM and boiling of the samples for 5 min at 95 °C. The samples were run on 15% SDS-PAGE, transferred onto PVDF membrane by wet-blot (described in 4.4.3) and membranes were then incubated with α-H4 K16ac and α-H4 (pan) antibody respectively for the individual acyl modified samples. Resulting signals were visualized by chemiluminescence and are depicted in Figure 33. The results show, that the modification at H4 K16 is fully removed by CobB after 20 min, independent of the different acylation type.

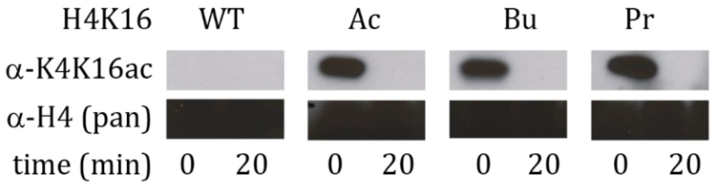
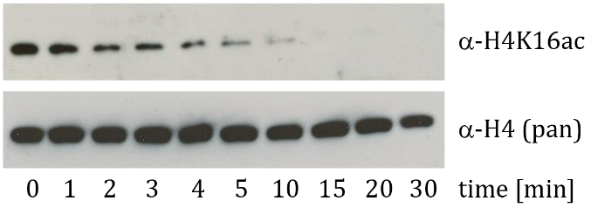


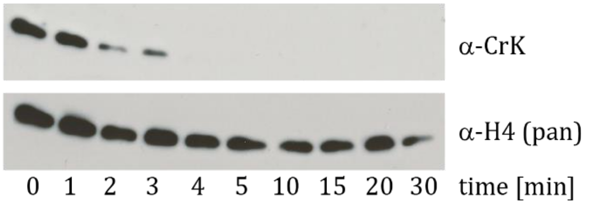
Figure 33 ***Removal of acyl modifications at K16 on histone H4 by CobB.*** The purified H4 K16mod proteins were incubated with CobB for 20 min and afterwards quenched by addition of 50 mM NAM and heating to 95 °C for 5 min. Samples before addition and after CobB incubation were run on 15% SDS-PAGE, transferred onto PVDF membranes and respectively for each modification incubated with α-H4 K16ac and α-H4 (pan) antibody. Signals were visualized by chemiluminescence, showing that the signal for the acyl modification had fully disappeared after 20 min incubation with CobB.

Nonetheless, this experiment still lacks cohesive evidence as to whether there is a different velocity of the reaction, dependent on the length of the acyl chain, as indicated by Denu et al.[117] For that purpose, Svenja and I established and performed a deacylation assay, as described in 4.6.6. Histone H4 K16mod was mixed with CobB using a 30x molar excess of the histone protein, to ensure dependency of turnover rates only on the deacetylase. Samples were taken over a time-span of 30 min. Resulting samples were processed to PVDF membranes by wet-blot and analyzed using either α-H4K16ac or α-CrK antibody. Membranes were stripped as indicated in 4.4.4 and signals reinvestigated with α-H4 antibody. The results are depicted in Figure 34. The acylations are fully removed within less than 10 min, and indicate different times depending on the acyl chain length. Hereby, we suggest a correlation of an decrease of the time required for deacylation with an increasing acyl chain length.

A: ***H4 K16Ac***



B: ***H4 K16Bu***



C: ***H4 K16Cr***

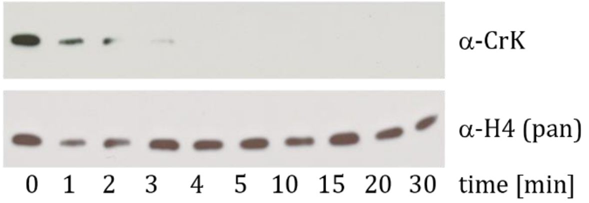


Figure 34 ***Removal of acyl group by CobB***. H4 K16mod were incubated with CobB. A 30x molar access of the histone proteins was used and samples were taken covering a time-span of 30 min. Samples were quenched by addition of preheated NAM and boiled at 95 °C for 5 min. Afterwards samples were separated by 15% SDS-PAGE, transferred onto PVDF membrane by wet-blot and membranes incubated with either α-H4K16ac or α-CrK following signal visualization by chemiluminescence were the membranes stripped and incubated with α-H4 antibody and signals were again detected by chemiluminescence. A: Removal of acetylation, B: removal of butyrylation and C: removal of crotonylation by CobB. The results show, that the acylation is completely removed by CobB. Hence, indicating different velocities of the reaction depending on the type of acylation.

2.2 First attempts towards protein labeling using GCE

Investigations of protein dynamics can be done using FRET. This requires the incorporation of two distinct fluorophores. Achieving site specific incorporation of two fluorophores onto one protein is difficult since it requires functionalities which are only rarely present within the protein of choice. One well established method is the labeling of lysine or cysteine residues with fluorescent dyes. As histone proteins rarely contain cysteines, installation of one fluorescent label on a cysteine is possible. Combinative cysteine labeling with a unique functionality introduced to the protein by genetic code expansion is one possibility to create a FRET pair. This was addressed by targeting histone protein H3. Histone H3 contains one cysteine at position C110, which lies directly within the center of the H3´-H3 four helix bundle. Labeling of this position was reported to result in complete loss of octamer formation ability.[164]

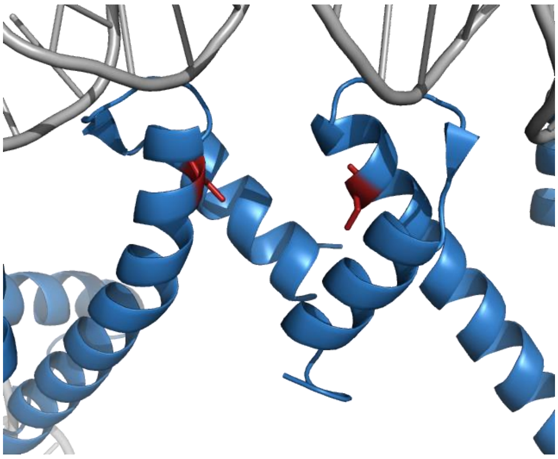


Figure 35 ***Cysteine position on histone H3.*** Histone H3 contains one cysteine at position 110 which lays within the H3´-H3 four helix bundle (highlighted in red). The picture was created using Pymol based on the structure PDB: 1AOI.

Therefore the cysteine was replaced by alanine and the cysteine reintroduced at other distinct positions to test whether these were suitable for labeling without perturbing octamer formation. The different histone H3 mutants contained the cysteine at one of the following positions R52, Q76, F78, T80, L109, Q125 and R128. The different H3 variants were expressed and purified as described in 4.5.5. The cysteines were labeled with Alexa 546-C2-maleimide (procedure outlined in 4.9.3), refolded into octamers and the octamers subsequently purified. Correct formation of octamers was found for all mutants tested, except for the mutant L109C. (Overview in Table 2-3).

Table 2-3 ***H3 mutants with replaced cysteine, subsequently labeled and tested upon ability to form octamers.***

Mutant	R52C	Q76C	F78C	T80C	L109C	Q125C	R128C
Octamer refolding	+	+	+	+	-	+	+

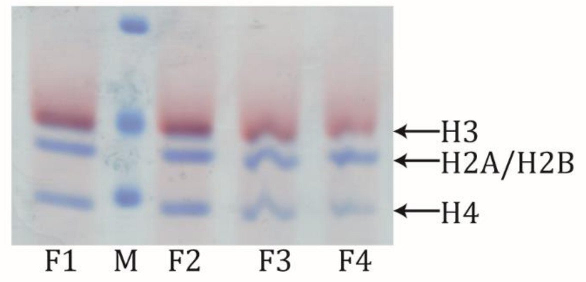


Figure 36 ***Octamer H3 Q76C-Alexa546.*** The octamer was assembled using H3 Q76C-Alexa546, purified by size-exclusion chromatography and resulting fractions separated by 15% SDS-PAGE. The Alexa546 dye was visualized by using Typhoon Imager (532 nm excitation wavelength, emission filter 580BP30, PTM 400 V) and protein bands were afterwards stained by Instant Blue. The Figure depicts an overlay of protein and fluorescence signals.

Those mutants were further combined with amber mutations at different positions on the histone tail, thereby replacing either position T6, A24 or V35. The amber mutations were used for the incorporation of an UAA, suitable for labeling via click reactions. For this purpose AzF was chosen.[44,133]

However, before addressing double labeling on the histone protein itself, the conditions for labeling of the AzF needed to be further investigated. In this regard the GFP protein was chosen for optimization tests. GFP was expressed as both WT and a mutant containing AzF at position 150. Labeling tests using DBCO545 as fluorescent dye revealed no unspecific labeling of GFP without AzF present, but for GFP containing AzF, a fluorescence signal specific for DBCO545 could be detected (shown in Figure 37).

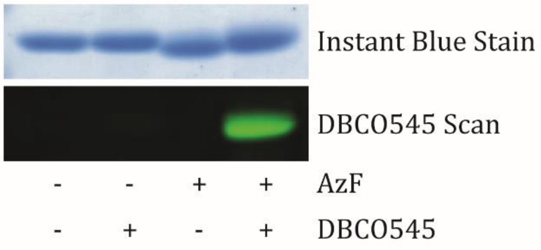


Figure 37 ***Labeling of GFP with DBCO-dye.*** GFP WT and GFP-AzF were reacted with equimolar ratios of DBCO545 dye, run on 12% SDS-PAGE, and the fluorescence signal of the dye was analyzed by Typhoon Imager (settings: excitation 532 nm, emission filter 580BP30, PTM 300 V). Afterwards, proteins were stained by Instant Blue.

Next, the ratio of protein to dye was varied, testing the ratios of 0, 1, 1.5, 2 and 5 of molar equivalents of dye for labeling of the GFP-AzF. The reaction was performed at RT for 30 min and the samples subsequently separated by 12% SDS-PAGE, without previous boiling of the sample. The fluorescence signal of the DBCO545 dye was detected by Typhoon Imager and afterwards protein bands were visualized by Instant Blue staining. The result is depicted in Figure 38, shown by an overlay of protein and fluorescence signals. For the GFP WT no unspecific labeling, even with higher concentrations of dye was visible. Protein containing the azido functionality revealed fluorescence signals with increasing intensity at higher amounts of dye.

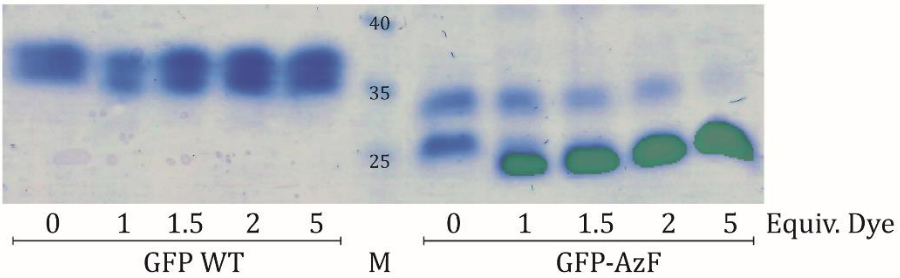


Figure 38 ***GFP labeling tested with different dye concentrations.*** GFP-AzF was labeled with different molar ratios of DBCO545 dye and compared to GFP WT. As ratios 1 eq, 1.5 eq, 2 eq and 5 eq of dye to protein were tested. Further a control without dye was included. The labeling was performed in PBS buffer at RT and 30 min reaction time. Samples of each of the different labeling reactions were separated by 12% SDS-PAGE, fluorescence signals of the dye measured by Typhoon Imager (Ex. 532 nm, Em. filter 580BP30, PTM 400 V) and afterwards Instant Blue stained. The figure presents an overlay of fluorescence and protein signals.

For application of labeling on the histone H3 protein, the H3 variants, including the repositioned cysteine together with an amber mutation for incorporation of AzF, needed to be produced. The positions chosen for AzF incorporation were on the N-terminal tail T6, A24 and V35. All the different histone variants were expressed in the presence of 1 mM AzF and purified as indicated in 4.5.5. One example representing the purified H3 Q76C V35AzF is shown in Figure 39. After purification the His_6-tag was cleaved by TEV protease and the H3 stored for further usage.

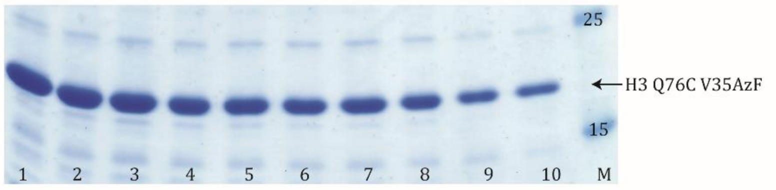


Figure 39 ***Purification of H3 Q76C V35AzF.*** The H3 construct was purified from 1 L using Ni^{2+}-batch purification. The fractions obtained by elution of proteins from the beads were separated by 15% SDS-PAGE and visualized by Instant Blue staining. Fractions containing the pure protein were pooled, subsequently dialyzed in ßME, and either TEV cleaved or directly used in further experiments.

The histone H3 variant with the mutations C110A, Q76C and V35AzF was used for investigations of the labeling conditions for derivatization of the azido functionality with a dibenzoyl-cyclooctyne dye (DBCO545), a strained alkyne for copper free click reactions. The reaction was followed by taking samples at different time points. Accordingly, molar ratios of protein to dye and the reaction temperature were tested (Figure 40). Molar ratios of protein to dye concentrations of 1, 5, 10 and 20 mol were covered. A control sample at each time point was included, using equimolar ratios of protein to dye, but with quenching of the azido function by addition of excess TCEP prior to the labeling reaction. The temperatures investigated were 25°C, 37°C, 16°C and 4°C. All samples were run on 15% SDS-PAGE and subsequently analyzed for the fluorescence signal by Typhoon Imager and protein bands by Instant Blue staining. The control samples revealed no attachment of the dye due to the absence of a fluorescence signal, but showed a clear protein band independent of the reaction temperature. In the comparison of different molar equivalents of dye, a decrease of the fluorescence signal together with a decrease of protein signal was observed at increasing dye concentrations for all temperatures investigated. Accordingly, longer reaction times yielded higher fluorescence signals at equimolar protein to dye ratios. Furthermore this effect was strongest at RT.

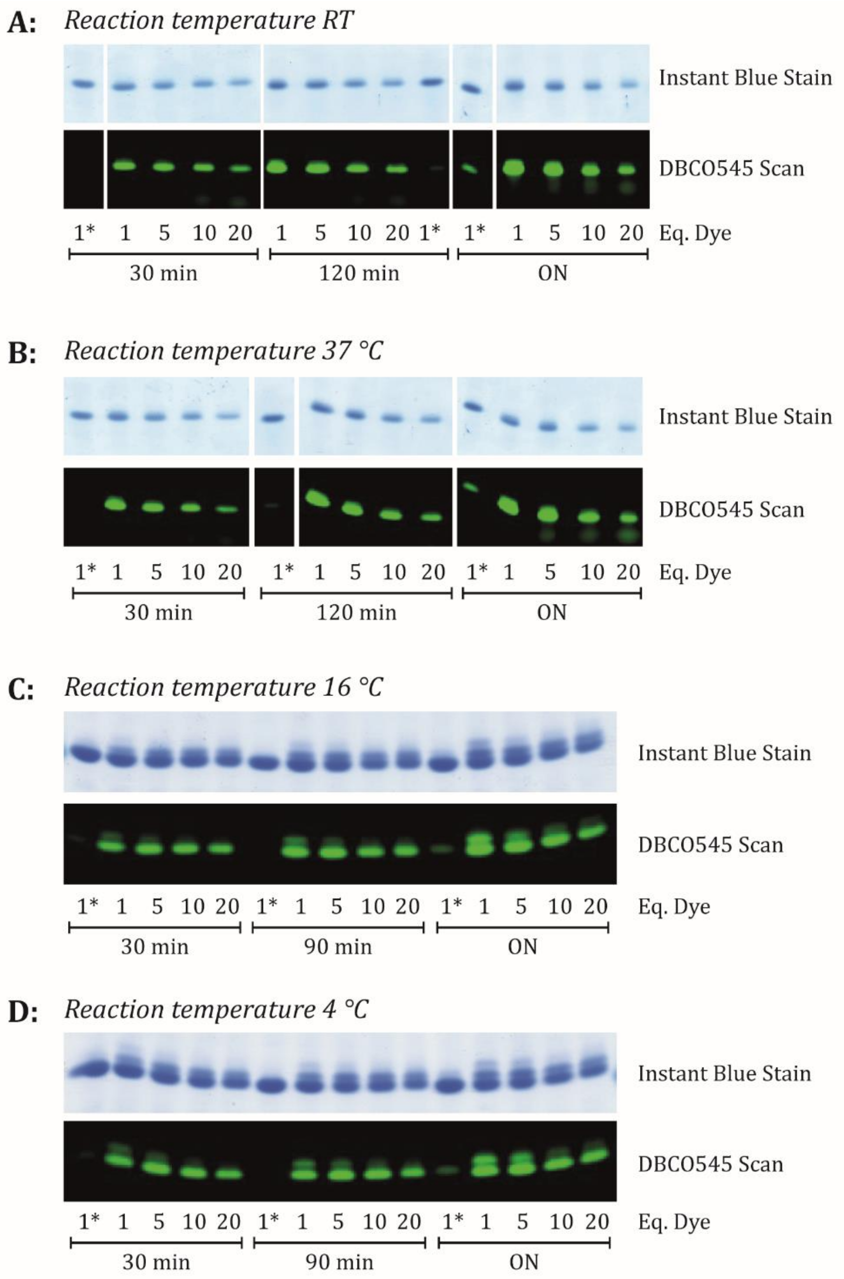

A: Reaction temperature RT
Instant Blue Stain
DBCO545 Scan
1* 1 5 10 20 1 5 10 20 1* 1* 1 5 10 20 Eq. Dye
30 min
120 min
ON
B: Reaction temperature 37 °C
Instant Blue Stain
DBCO545 Scan
1* 1 5 10 20 1* 1 5 10 20 1* 1 5 10 20 Eq. Dye
30 min
120 min
ON
C: Reaction temperature 16 °C
Instant Blue Stain
DBCO545 Scan
1* 1 5 10 20 1* 1 5 10 20 1* 1 5 10 20 Eq. Dye
30 min
90 min
ON
D: Reaction temperature 4 °C
Instant Blue Stain
DBCO545 Scan
1* 1 5 10 20 1* 1 5 10 20 1* 1 5 10 20 Eq. Dye
30 min
90 min
ON

Figure 40 ***Investigation of reaction conditions for DBCO labeling.*** Purified H3 Q76C V35AzF was reacted with DBCO545 dye using different molar ratios of protein to dye. Furthermore different reaction temperatures were investigated and the progress was monitored taking samples at various timepoints. The (*) indicates a control sample that was quenched by addition of excess TCEP before addition of the dye for labeling. The samples were run on 15% SDS-PAGE, the fluorescence signal was detected using Typhoon Imager and afterwards protein signals were visualized by Instant Blue staining. The settings used for DBCO545 analysis were: DBCO545: Ex. 532 nm, Em. 580BP30, PTM350 V and Alexa647-mal: Ex. 633 nm, Em. 670BP30, PTM 350 V. The different temperatures investigated were in A: RT, B: 37°C, C: 16°C and D: 4°C.

In summary, the best conditions for DBCO labeling were concentrations of near equimolar protein to dye and a temperature of 25°C (RT). These reactions required an incubation time of greater than 120 min.

The combination of azide labeling with DBCO and Cys labeling with a maleimide dye proved challenging. The simultaneous labeling with both dyes is not possible, since the Cys labeling requires TCEP as reducing agent in the buffer, which in turn quenches the azido function. Thus both labeling reactions had to be performed in successive steps, starting with the DBCO labeling.

Several procedures were tested to perform the buffer exchange of the histone H3 DBCO labeled protein into PBS containing TCEP. The exchange of buffers via Nap10 column was not possible as only the unlabeled protein passed through the column but the labeled protein stuck to the material. Other attempts, using the his$_6$-tagged protein and either performing the labeling reaction first with subsequent binding to Ni^{2+}-beads or vice versa was likewise unsuccessful. The protein was impossible to recover from the beads through elution or TEV cleavage from the beads. Successful buffer exchange was achieved by performing short dialysis steps subsequent to the DBCO labeling.

Consequently, the combination of both labels could be achieved starting with the DBCO labeling. Then the protein was dialyzed into PBS-TCEP (dialysis steps: 1 h, 2x 30 min) followed by cysteine labeling overnight at 4°C.

With this established procedure, different dye concentrations were tested thus not exceeding a two molar ratio of each dye respectively (depicted in Figure 41).

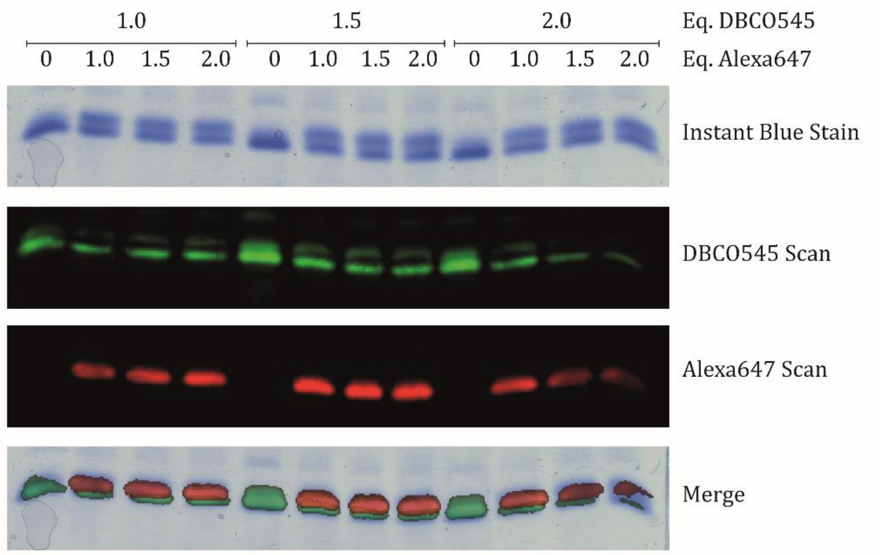


Figure 41 ***Test of different dye concentrations for double labeling.*** Histone H3 Q76C V35AzF was first labeled at the azido function with different molar ratios of DBCO545 dye, rebuffered via dialysis in PBS-TCEP buffer and afterwards reacted with different molar ratios of Alexa647-mal. The samples were separated by 15% SDS-PAGE and the different dyes analyzed by Typhoon Imager. DBCO545: Ex. 532 nm, Em. 580BP30, PTM 350 V. Alexa647-mal: Ex. 633 nm, Em. 670BP30, PTM 350 V.

At first, the DBCO dye was reacted with the azido function in molar ratios of 1, 1.5 and 2. The samples were rebuffered via dialysis and subsequently labeled with Alexa647-mal using again molar ratios of 1, 1.5 and 2 of the second dye. Samples were separated by 15% SDS-PAGE, and then the fluorescence signals analyzed by Typhoon Imager scan with settings appropriate for the different properties of the dyes. The protein was visualized by Instant Blue staining. Increasing the DBCO concentration to a 1.5 molar excess resulted in brighter fluorescence signals whereas 2 molar excess showed no further improvement. In combination with the Alexa647 dye concentrations of 2 molar excess DBCO and increasing molar ratios of Alexa647 accompanied decreasing fluorescence signals. Using equal concentrations of protein and DBCO combined with increasing Alexa647 concentrations revealed increasing brightness in the Alexa647 fluorescence signal, hence the DBCO signal was weaker as compared to 1.5 molar ratios of DBCO. With this DBCO ratio, all Alexa647 concentrations seemed to be almost equal. Therefore both dyes were further used in 1.5 molar excesses.

This double labeled histone H3 was tested for the ability to efficiently reconstitute into octamers. Therefore WT H2A, H2B and H4 were combined with labeled H3 (DBCO545 and Alexa647 both 1.5 molar ratios from Figure 41). The labeled H3 was mixed 1: 10 with H3 WT to avoid the incorporation of two double labeled H3 into one octamer (shown in Figure 42).

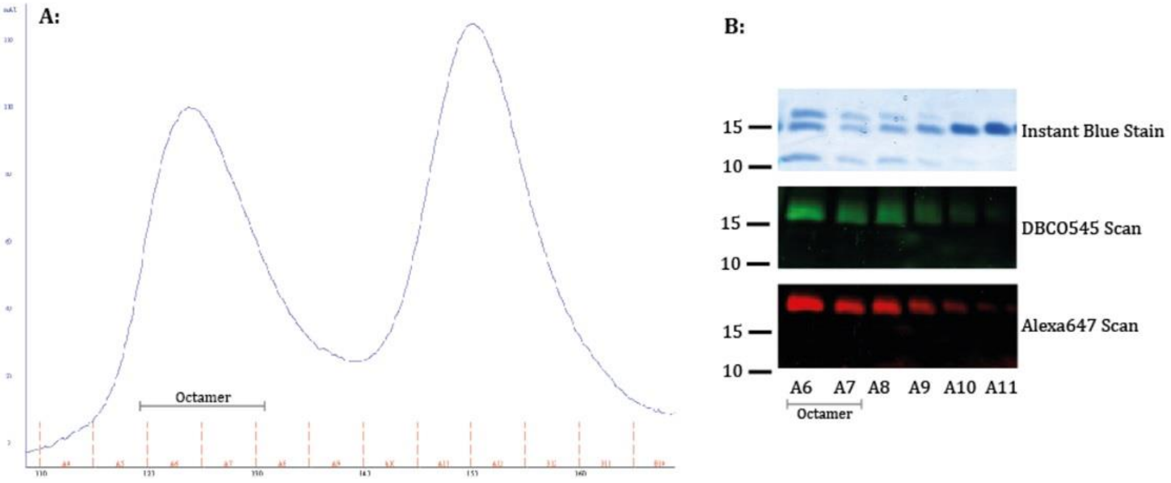


Figure 42 ***Octamer refolding including double labeled H3.*** The histone octamer was refolded using WT H2A, H2B and H4. For histone H3 the double labeled H3 Q76CV35AzF was mixed 1:10 with WT H3. The octamer was purified by size-exclusion chromatography and the resulting chromatogram is depicted in A. The different fractions from A were analyzed by 15% SDS-PAGE (shown in B) with detection if the octamer contained the different fluorophores by Typhoon Imager scan with settings appropriate to the dye respectively. DBCO545: Ex. 532 nm, Em. 580BP30, PTM400 V. Alexa647-mal: Ex. 633 nm, Em. 670BP30, PTM 400 V. Protein bands were visualized by Instant Blue staining.

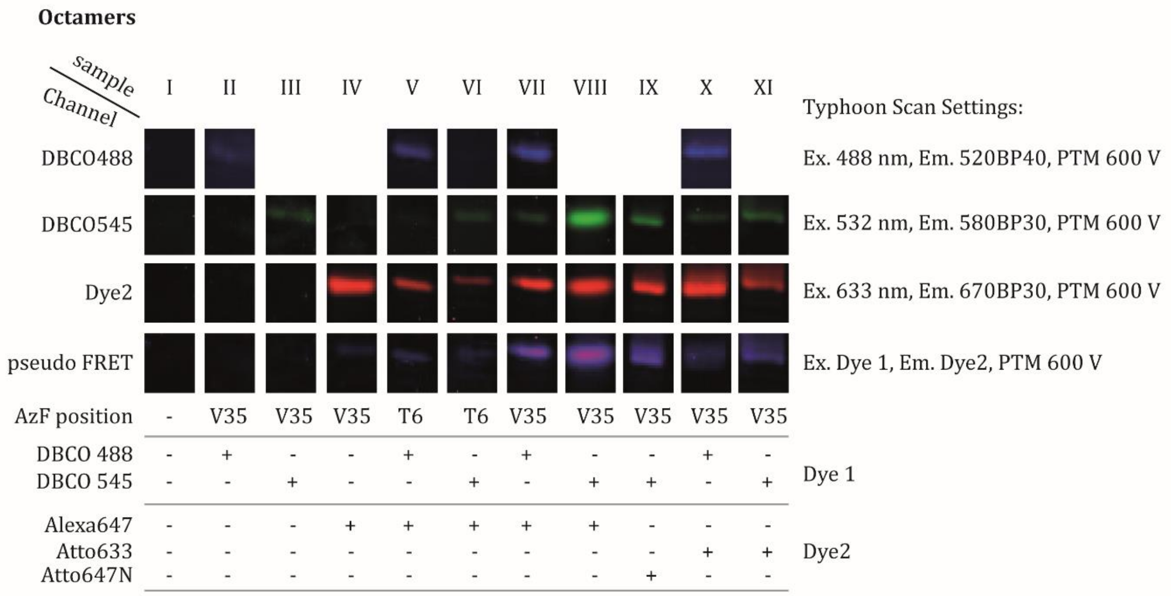
The resulting octamers were refolded and purified as described in 4.6.1 (depicted in Figure 42 A) and fractions containing protein analyzed by 15% SDS-PAGE (Figure 42 B). Fluorescently labeled H3 octamers were analyzed by Typhoon Imager using settings appropriate for either DBCO545 or Alexa647. Besides correct formation of the octamer, the H3 band showed fluorescence signals of both dyes.

Consequently, different dye combinations were tested on H3 variants with a replaced cysteine, combined with the AzF at either position T6 or V35. Potential FRET pairs were obtained by testing different dye combinations with the labeled H3 during octamer reconstitution. The different variations of dyes and positions of the AzF are outlined in Table 2-4.

Table 2-4 ***Variations of AzF position and dye combinations.***

Sample	Protein	Cys position	AzF position	Cys Label	AzF Label2
I	H3 C110A	Q76C	-	-	-
II	H3 C110A	Q76C	V35AzF	-	DBCO488
III	H3 C110A	Q76C	V35AzF	-	DBCO545
IV	H3 C110A	Q76C	V35AzF	Alexa647	-
V	H3 C110A	Q76C	T6AzF	Alexa647	DBCO488
VI	H3 C110A	Q76C	T6AzF	Alexa647	DBCO545
VII	H3 C110A	Q76C	V35AzF	Alexa647	DBCO488
VIII	H3 C110A	Q76C	V35AzF	Alexa647	DBCO545
IX	H3 C110A	Q76C	V35AzF	Atto647N	DBCO545
X	H3 C110A	Q76C	V35AzF	Atto633	DBCO488
XI	H3 C110A	Q76C	V35AzF	Atto633	DBCO545

Octamers containing the indicated labels were refolded, again mixing the labeled H3 1:10 with WT H3 in combination with WT H2A, H2B and H4. The octamers were purified, separated by 15% SDS-PAGE and analyzed for the different fluorescence signals. The fraction of eluted protein that contained the respectively highest octamer concentration were compared (depicted in Figure 43). The potential FRET donor was installed by labeling the azido functionality with DBCO. Therefore either DBCO488 or DBCO545 was used. Due to differing spectral properties, the settings for the fluorescence scan were adjusted accordingly and signals shown in separate channels (DBCO488: Ex. 488 nm, Em. filter 520BP40; DBCO545: Ex. 532 nm, Em. filter 580BP30). FRET acceptor "red" dyes were used and scanned in a channel called dye2 (Ex. 633 nm, Em. Filter 670BP30), with the appropriate setting. The red dyes tested include Alexa647, Atto633 and Atto647N. Furthermore the samples were analyzed for an in-gel FRET signal indicated as a pseudo-FRET channel. In the pseudo-FRET channel the donor dye (dye1) was excited with a suitable excitation wavelength and the emission filter adjusted as indicated for the acceptor (dye2). The octamer containing no labels showed no fluorescence signals with all settings. For the octamers containing only one label (either DBCO488, DBCO545 or Alexa647), a fluorescence signal was only detected within the respective channel and no background was observed when scanned by using the remaining settings. Comparing DBCO488 installed at either T6 (V) or V35 (VII) and combined with Alexa647, the pseudo-FRET signal for V35 (VII) is significantly higher. The same applies by comparing DBCO545 labeled T6 (VI) and V35 (VIII) with Alexa647. Hence for V35 DBCO545 and Alexa647 all individual fluorescence signals are brighter. DBCO545 at V35 combined with Atto647N (IX) showed both signals for the individual channels of the dyes and a pseudo-FRET signal. Comparing either DBCO488 or DBCO545 at V35 with Atto633 labeling revealed only a very weak pseudo-FRET signal for DBCO545 (XI) but almost no signal in case of DBCO488 (X). In general the observed pseudo-FRET signals were stronger in combinations where Alexa647 was used compared to the other red dyes Atto633 and Atto647N.


Octamers
Sample
Channel
I
II
III
IV
V
VI
VII
VIII
IX
X
XI
Typhoon Scan Settings:
DBCO488
DBCO545
Dye2
pseudo FRET
Ex. 488 nm, Em. 520BP40, PTM 600 V
Ex. 532 nm, Em. 580BP30, PTM 600 V
Ex. 633 nm, Em. 670BP30, PTM 600 V
Ex. Dye 1, Em. Dye2, PTM 600 V
AzF position
DBCO 488
DBCO 545
Dye 1
Alexa647
Atto633
Atto647N
Dye2

Sample	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	
AzF position	-	V35	V35	V35	T6	T6	V35	V35	V35	V35	V35	
DBCO 488	-	+	-	-	+	-	+	-	-	+	-	Dye 1
DBCO 545	-	-	+	-	-	+	-	+	+	-	+	
Alexa647	-	-	-	+	+	+	+	+	-	-	-	Dye2
Atto633	-	-	-	-	-	-	-	-	-	+	+	
Atto647N	-	-	-	-	-	-	-	-	+	-	-	

Figure 43 ***Overview of labeled Octamers.*** Octamers were refolded with modified H3 containing a Cys at position Q76 and AzF at either position T6 or V35 which was labeled with different dyes or combinations thereof. The purified octamer fractions were separated by 15% SDS-PAGE. The fluorescence signals were detected by Typhoon Imager using settings suitable for DBCO488 (Ex. 488 nm, Em. Filter 520BP40 and PTM 600 V), DBCO545 (Ex. 532 nm, Em. Filter 580BP30 and PTM 600 V), or Dye2 (including either Alexa647, Atto633 or Atto647N; Ex. 633 nm, Em. Filter 670BP30, PTM 600 V). Detection for pseudo-FRET was done by exiting the DBCO dye respectively (Ex. 488 or 532 nm) and detecting for Dye2 (Em. Filter 670BP30).

A subset of the labeled octamers was selected for nucleosome assembly (as described in 4.6.2). Resulting nucleosome bands are compared in Figure 44.

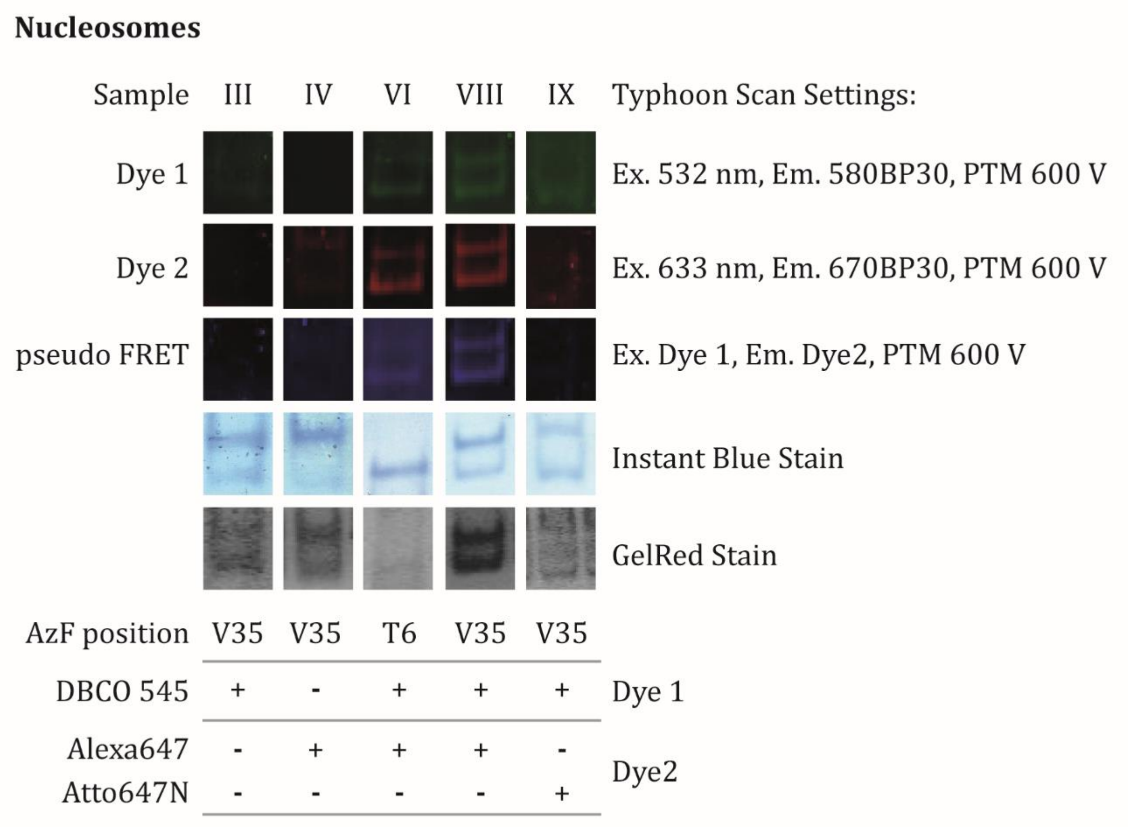

Nucleosomes
Sample
III
IV
VI
VIII
IX
Typhoon Scan Settings:
Dye 1
Ex. 532 nm, Em. 580BP30, PTM 600 V
Dye 2
Ex. 633 nm, Em. 670BP30, PTM 600 V
pseudo FRET
Ex. Dye 1, Em. Dye2, PTM 600 V
Instant Blue Stain
GelRed Stain
AzF position
V35 V35 T6 V35 V35
DBCO 545
+ - + + +
Dye 1
Alexa647
- + + + -
Atto647N
- - - - +
Dye2

Figure 44 ***Overview of labeled Nucleosomes.*** Nucleosomes were assembled by salt dialysis using the octamers containing histone H3 with different mutations and dyes and 187 bp DNA containing the '601' positioning sequence[165]. The nucleosomes were run on 6% native PAGE and fluorescence analyzed by Typhoon Imager scans with settings suitable for the respective fluorescent dye and pseudo FRET. Dye1: Ex. 532 nm, Em. Filter 580BP30, PTM 600 V. Dye2: Ex. 633 nm, Em. Filter 670BP30, PTM 600 V. Pseudo-FRET: Ex. 532 nm, Em. Filter 670BP30, PTM 600 V. Afterwards protein bands were visualized by Instant Blue staining and DNA by GelRed staining.

All samples tested resulted in signals for both protein and DNA staining, verifying the formation of nucleosomes. Both mono labeled controls were applied on the H3 variant Q76C and V35AzF (III DBCO545 and IV Alexa647). Only very weak signals in the respective fluorescence scan are visible. Fluorescence signals in all three channels were observed when comparing potential FRET pairs formed by DBCO545 and Alexa647 with the DBCO label placed at either T6AzF or V35AzF. Hence, the signals detected with the V35 DBCO labeled variant were brighter. The combination of DBCO545 with Atto647N on H3 V35AzF showed only a weak signal for the DBCO545 fluorescence but almost no signal for Atto647N and accordingly no pseudo-FRET.

For detailed analysis of the labeling efficiency, possible methods to remove the free dye from the reaction were addressed. As the labeled histones bind tightly to Ni^{2+}-beads or sepharose material, alternative methods were investigated. These were based on protein precipitation. The labeled H3 Q76C V35AzF with either DBCO545, Alexa647 or both were precipitated by MeOH/$HCCl_3$[166], TCA[167], TCA/DOC[168] and acetone[169]. Precipitations of the protein by TCA or acetone failed to yield significant results. Samples precipitated with MeOH/$HCCl_3$ and TCA/DOC are depicted in Figure 45. For labeled H3 precipitation with TCA/DOC the yield of recovered protein was poor. In contrast precipitation with MeOH/$HCCl_3$ showed improved protein recovery.

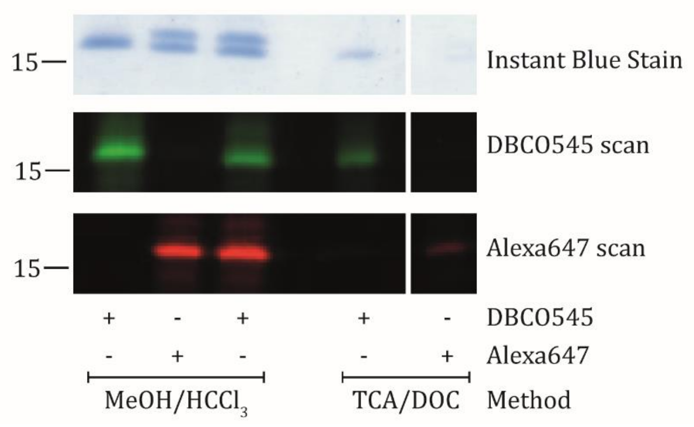


Figure 45 ***Separation of free dye by protein precipitations.*** H3 Q76C V35AzF containing the indicated dyes were precipitated either by MeOH/$HCCl_3$ or by TCA/DOC. Subsequently to the precipitation procedure samples were run on 15% SDS-PAGE, fluorescence signals analyzed by Typhoon Imager (DBCO545: Ex. 532 nm, Em. filter 580BP30, PTM 400 V; Alexa647: Ex. 633 nm, Em. Filter 670BP30, PTM 400 V) and protein bands visualized by Instant Blue staining.

The precipitation of labeled H3 allows for quantitative analysis of the different labeling steps maybe revealing optimization potential, which remains to be elusive.

2.2.1 Introduction of two distinct UAAs for protein labeling

The introduction of two distinct UAAs into one protein requires two tRNA/AARS pairs. Great advances were achieved in the Chin lab with the evolution of orthogonal ribosomes that allows for the introduction of two different UAAs.[54,170] However the efficiency is still limited, as the use of two tRNA/AARS pairs, along with a plasmid containing the gene of interest are required. The use of many different plasmids along with different selective antibiotics causes high metabolic stress in the host cells. The optimization of this system was addressed by Dr. C. Lammers whereby he balanced the levels of the different compounds by fine-tuning the expression level of PylS with different promoters. In addition, a modular plasmid was created, containing both tRNA/AARS pairs thereby reducing the number of plasmids required by one. This increased the host cell viability as shown via growth assays with this system.[171] With the different changes, further optimization of the system towards the introduction of two distinct UAAs into one protein was achieved.

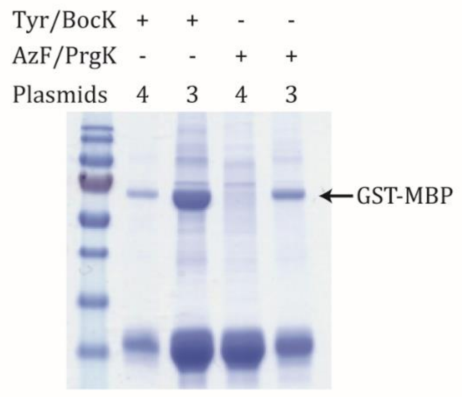


Figure 46 ***GST-MBP expressed and purified using different plasmids and amino acid combinations.*** The GST-MBP containing the different amino acid combinations were expressed either using the initial 4 plasmid combination or using the optimized plasmid that requires a total of only 3 plasmids. Obtained proteins were purified, separated by 12% SDS-PAGE and visualized via Instant Blue staining. The protein yields were significantly higher using the 3 plasmid system than by using the initial 4 plasmid system.

We used the modular plasmid for introducing two different UAAs in the GST-tagged maltose-binding-protein (GST-MBP). Position Y17 of MBP was replaced by a quadruplet codon and N234 of the same protein by an amber codon. To monitor the functionality of this system with UAAs, the protein was expressed together with two different amino acid combinations to label the protein with two different UAAs. First, *E. coli* cells were transformed with either the initial 4 plasmid combination or by using the modular plasmid with only a total number of 3 plasmids. These were cultured and used for small volume protein expressions. The protein was purified, separated by 12% SDS-PAGE and Instant Blue stained (depicted in Figure 46). Using the modular plasmid, the protein yield was significantly higher as compared to the 4 plasmid system.

GST-MBP was then produced for later labeling reactions, using the three plasmid system as described by Lammers et al.[171] introducing either Tyr and BocK as controls or PrgK and AzF as functional amino acids for chemical modification. The proteins were purified as described in 4.5.10, separated by 12% SDS-PAGE and stained using Instant Blue. The results are depicted in Figure 47.

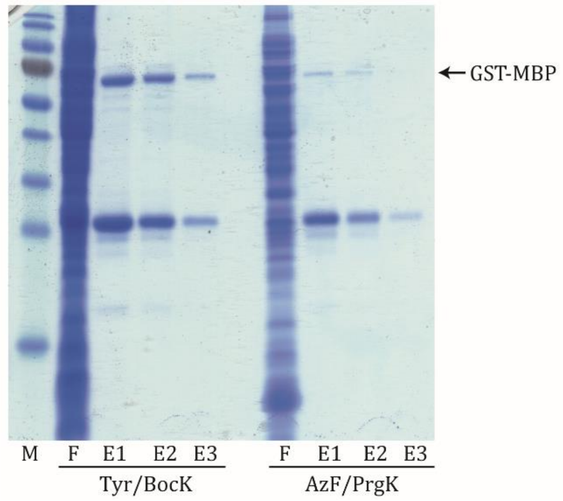


Figure 47 ***Purified GST-MBP with different UAA combinations.*** For control experiments the protein was expressed and purified with the amino acids Tyr and BocK. For labeling the amino acid combination of AzF and PrgK was chosen. Samples from flow through (F) and different elution fractions (E) were separated by 12% SDS-PAGE and visualized by Instant Blue staining.

The purified GST-MBP proteins were then labeled using Alexa-647-azide for labeling of the alkyne moiety of PrgK performing a Cu(I) catalyzed click reaction (details outlined in 4.9.1), whereas DBCO-488 was attached to the azide functionality of the AzF by strain promoted click reaction (method described in 4.9.2). The dyes were also added to the protein carrying only nonfunctional UAA as comparison. The labeled samples were separated by 12% SDS-PAGE and analyzed using Typhoon Fluorescence Imager with excitation wavelength and emission filter adjusted to the properties of the dye respectively. The resulting picture is shown in Figure 48. The nonfunctional UAAs do not react with the dye. The functional UAAs reacted only with the respectively appropriate dye but showed no cross reactivity showing respectively functional labeling for both UAAs.

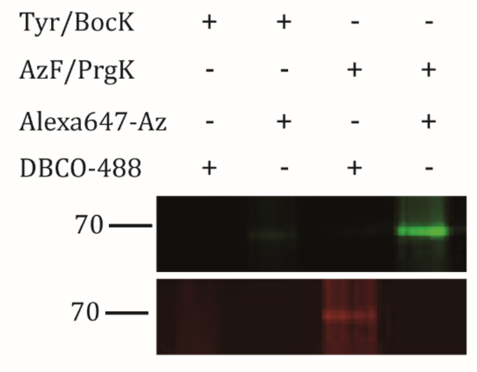
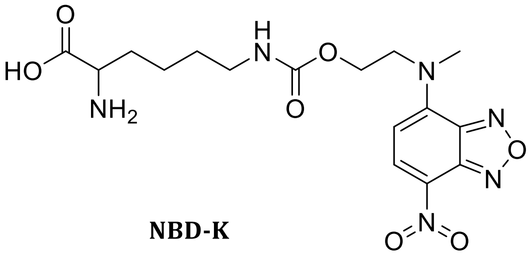


Figure 48 ***Labeled GST-MBP.*** GST-MBP protein containing either Tyr and BocK or AzF and PrgK were separately labeled with Alexa647 and DBCO488. Afterwards the proteins were separated by 12% SDS-PAGE and the dyes visualized using Typhoon fluorescence imager with excitation wavelength and emission filter adjusted to the properties of the respective dye. No signal for the control experiments are detectable, whereas the functional amino acids show a clear fluorescence signal in the respective channel.

2.2.2 Direct introduction of a fluorescent UAA

So far, only very few fluorescent UAAs could be introduced on proteins directly. We wanted to investigate whether (7-nitrobenz-2-oxa-1,3-diazole-4-yl)-lysine (NBD-K) was a possible candidate. The fluorescent NBD is widely applied in studying membrane processes when combined with different lipids, as it is only weakly fluorescent in water and it is environmental sensitive when transferred to a hydrophobic medium.[172] The structure of NBD-K is depicted in Figure 49.



NBD-K

Figure 49 ***Structure of NBD-K.***

The UAA was synthesized and kindly provided to us from the group of Prof. S. Hell (MPI, Göttingen). To render an UAA suitable for genetic code expansion, the UAA has to be efficiently taken up by the cells [6]. To analyze this criterion, *E. coli* cells were grown for 3 h in the presence of 2 mM NBD-K and subsequently visualized using fluorescence microscopy, as shown in Figure 50. The *E. coli* bacteria showed a fluorescence signal upon excitation with one specific wavelength.

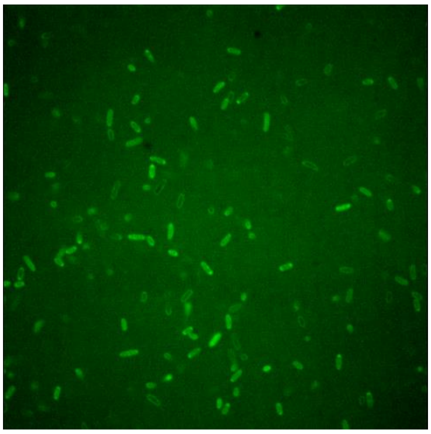


Figure 50 ***NBD-K uptake in E. coli cells.*** *E. coli* cells were grown in the presence of 2 mM NBD-K and cells were visualized using fluorescence microscopy.

Next, the library was used to select for a mutant that possibly recognizes NBD-K. This was done as described in 4.7.2. The resulting mutants from the selection are depicted in Figure 51. Without Cm all mutants survived independent of the presence of NBD-K. For Cm concentration of 100 mM and 200 mM without NBD-K, all mutants (except for one) died. Whereas in the presence of NBD-K various mutants survived both Cm concentrations. At a Cm concentration of 400 mM no more mutants survived (data not shown).

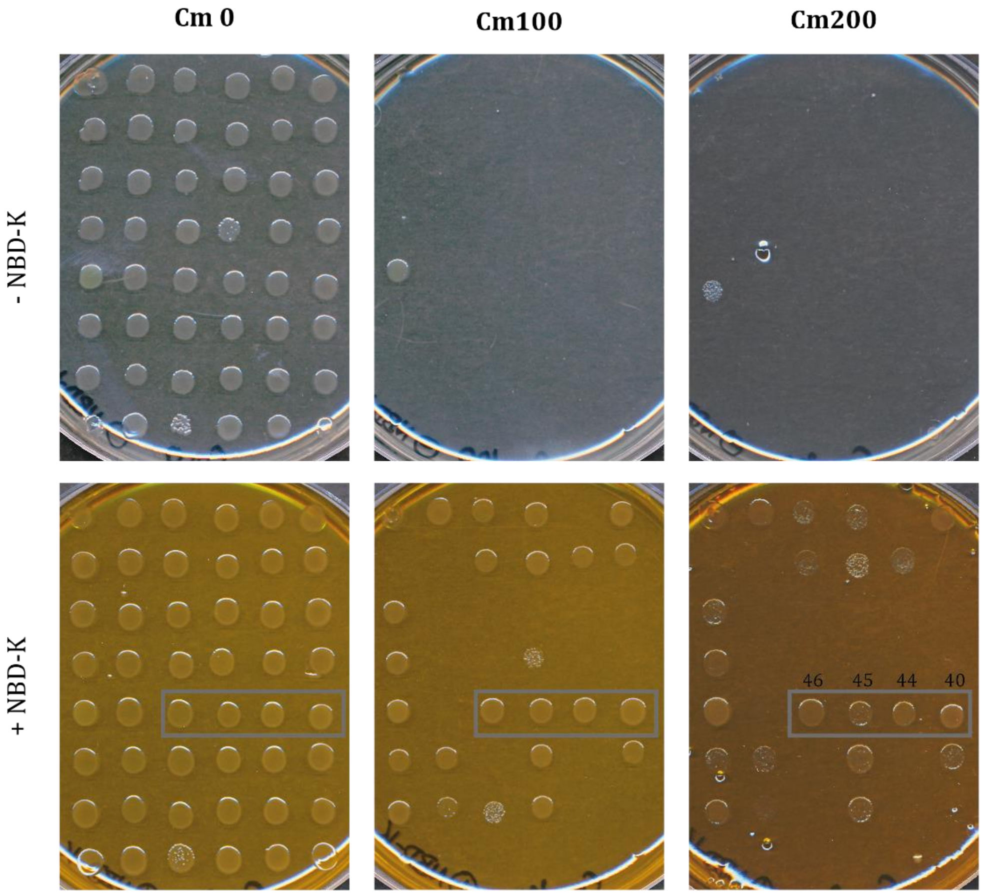


Figure 51 ***Library selection for a mutant which recognizes NBD-K.*** Mutants selected with positive and negative reporters. Mutants from the second negative round were again transformed together with a positive reporter plasmid and tested for their Cm dependent growth. The mutants with number 40, 44, 45 and 46 are highlighted and were subsequently used for protein test expressions.

Test expressions of myoglobin in the presence and absence of NBD-K is shown in Figure 52. The PylS mutants NBD-RS45 and NBD-RS 46 showed amino acid dependent formation of myoglobin, whereas NBD-RS40 and NBD-RS44 showed no specific recognition of NBD-K.

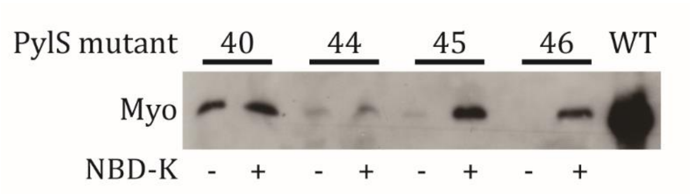


Figure 52 ***Test expression of myoglobin in the presence and absence of NBD-K.*** The mutants 45 and 46 showed NBD-K dependent formation of myoglobin. Whereas mutants 40 showed unspecific AA incorporation and 44 failed to produce full-length protein.

The sequences found for the mutants tested in myoglobin expression are outlined in Table 2-5. The mutant NBD-RS 45, which proved best for protein expression, contained the following variations compared to the WT PylS. Leu274 was changed to Val, Cys313 was changed to Val, Met315 was replaced by Cys. Additionally, a nonprogrammed mutation changed Thr248 to Ser.

Table 2-5 ***Sequence changes in mutants selected to recognize NBD-K.*** The highlighted mutant NBDK-RS 45 was used in further experiments for protein expressions.

WT	Ala267	Tyr271	Leu274	Cys313	Met315	Other
NBDK 40	His	Tyr	Cys	Lys	Val	
NBDK 44	Ala	Val	Arg	Leu	Ala	
NBDK 45	Ala	Tyr	Val	Val	Cys	Thr248Ser
NBDK 46	Ala	Tyr	Ala	Ala	Leu	

Myoglobin was expressed using a 50 mL culture supplemented with 1 mM NBD-K and appropriate antibiotics. The cells were cracked instead of using sonification by supplementing the lysis buffer 1:10 with BugBuster® (Novagene). Then protein purification was carried out following the standard protocol described in 4.5.2, except that because of the small volume the purification was performed in an Eppendorf tube. After addition of the buffer volume the Ni^{2+}-beads were sedimented via centrifugation and the remaining supernatant removed, before subsequent application of washing/elution steps. The result is depicted in Figure 53, showing the purified myoglobin in the eluted fraction and indicated successful incorporation of the NBD-K. However, as a next step the detailed verification by mass remains to be performed.

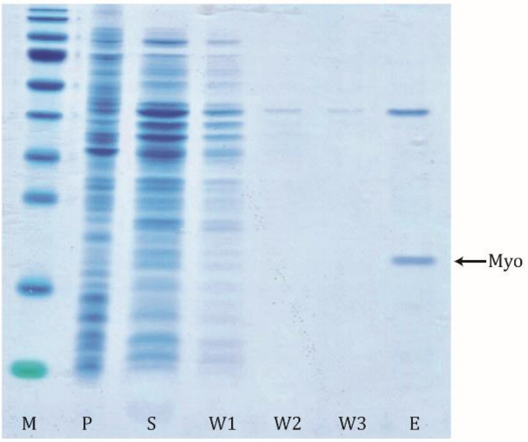


Figure 53 ***Purification of myoglobin containing NBD-K at position 4.*** The different fractions resulting during purification were run on 15% SDS-PAGE and stained by instant-blue. M: Marker, P: Pellet, S: Supernatant, W: Wash, E: Elution.

3 Summary and Conclusions

3.1 Novel lysine acyl modifications and their investigations

At the onset of this doctoral work, the acyl modifications propionyl, butyryl and crotonyl were only just recently discovered.[99,100] Since then, acyl modification identifications have expanded to include by formyl-, succinyl-, and malonyl groups during the last years.[97,101,105] We decided to investigate their site-specific incorporation on histone proteins by genetic code expansion to enable investigation of their physiological role.

We created a PylS mutant library (Figure 11) and selected successfully for PylS mutants, which accepted either PrK or CrK as possible substrates. After performing test expressions with the selected mutants, we could determine mutants that performed the most efficiently (Figure 14).

Due to the structural similarity of the acyl amino acids and the substrate tolerance of the WT PylS[173–175], we tested the cross selectivity of the evolved acylRS mutants and PylS WT for the different acyl modifications. We further tested those PylS variants on their ability to incorporate BuK. We found that the evolved PrKRS performed best with PrK. For CrK, WT PylS and the evolved CrKRS performed with approximately similar efficiency. The WT PylS furthermore accepted PrK and BuK as substrate but did not tolerate AcK. The BuK was also accepted by CrKRS as a substrate though the WT seemed to perform slightly better. As for the AcK, neither one of the newly evolved synthetases accepted AcK as a substrate. The incorporation of AcK was more efficient when using a concentration of 10 mM[28] as compared to 2 mM (Figure 15 A,B,C) when the yields were significantly lower.

Our results matched those by Gattner et al. who used the WT *M. mazei* PylS for the incorporation of propionyl-, butyryl and crotonyl-lysine.[175] Our selected mutants for PrK and CrK incorporation contained only two changes in the amino acid sequence as compared to the WT PylS. For PrKRS the changes were Tyr271Phe which removes the polar hydroxylgroup and Cys313Thr so that at this position probably the methyl group interacts with the aliphatic site-chain. For CrKRS Cys313 is replaced by the aliphatic Val and Met315 is replaced by Tyr. A MbPylS synthetases was also evolved for CrK recognition by Kim et al. containing the three mutations Leu274Ala, Cys313Phe and Tyr349Phe.[176] Noticeably the polar Cys313 was in all variations replaced. A structural investigation on how the WT *M.mazei* PylS accommodates the different acyl modification within the active site shows variations of the ε-N positions dependent on the length of the aliphatic chain of the acyl group and due to their flexibility results in a “curling up” within the pocket.[173,177]

We could confirm successful incorporation of all four acyl modifications (AcK, PrK, CrK and BuK) in Myoglobin by ESI-TOF mass spectrometry (Figure 16). The introduction of these acyl modifications was next applied on the histone H3 at positions K9 and K14. Investigations with mass spectrometry revealed cleavage of the acyl modifications when produced in the absence of CobB inhibitor NAM. Pinpointing towards functional recognition of the other acyl modifications by CobB. This is further discussed below.

All the different acyl modifications have been reported to exist on histone H3, though only individual acyl modifications for each position respectively (compare Figure 17). This complicates studies addressing the physiological impact of the acyl modifications. In contrary, most of the positions reported to be acetylated on the H4 tail, contained matches for all other acyl modifications as well (Figure 18).[100,152] With the exception of K20, where none of the novel acyl modifications was reported hitherto. This rendered H4 highly valuable for studying the effects caused by varying acyl modifications. However, H4 has been a difficult target for genetic code expansion. We solved those problems by constructing an expression vector that fused a H3 gene, with depleted positions 93-98 via a TEV-cleavage site to the H4 gene. We furthermore included an amber codon at the position K16, where the acyl lysine should be located.

Amber suppression competes with the release factor1 (RF1) which natively binds to the stop codon and initiates transcription termination.[6] To optimize amber suppression, *E. coli* strains have been constructed, with replacing the amber stop codons (UAG) by ochre stop codons (UAA) which are recognized by RF2 and allowed for the deletion of RF1.[153,178] In a comparison of DH10B cells with the RF1 deletion cell line C321.ΔA.exp[153], we showed a 4-fold improvement of protein yield.[174] Taken utility of this *E. coli* strain, we could successfully produce H4 with K16 acylations (Figure 21, Figure 23). The TEV cleaved H4 K16mod (Figure 22) was used for octamer refolding without further removal of the truncated H3. The H3 truncation precipitated during dialysis and could be removed via centrifugation through Spin-X® filters and did not incorporate into the octamers (depicted Figure 26). Octamers, containing the acyl modified H4 were successfully reconstituted and purified (Figure 27 and Figure 28). Pointing out that the main achievement with this methods is the relatively high yields of modified H4, which allow for octamer reconstitution with the modified H4 protein and even nucleosome assembly.

As the synthetases possess a certain substrate tolerance, we investigated if the antibodies, specifically discriminate between the different acyl modifications. We tested this for α-H4K16ac and α-CrK antibodies on site-specifically acylated H4 (Figure 24), revealing the tolerance of various acyl modifications. Though the α-H4K16ac antibody did not recognize crotonylation and the α-CrK antibody detected besides the crotonylation also butyrylation but concentration dependent. With a concentration of the α-CrK of 1:4 000 the butyrylation is nicely recognized whereas when diluting the antibody to concentrations of 1:20 000 it almost selectively recognized crotonylation. This was also affirmed on chemically acylated H4 (Figure 25), with the extensively acylated lysines residues weakened the interactions with the pan α-H4 antibody and further on the histone octamers harboring the site-specific modifications (Figure 28).

The histone tails are important modulators of interactions with the DNA during formation of higher order structures. Especially the histone H4 tail is important in this regard.[179] More detailed, position K16 was discovered as playing a vital role, with acetylation at this position impeding formation of higher order structures and changing interactions with chromatin-associated proteins.[74,179–181] Acetylation of the lysine residue neutralizes its positive charge thereby perturbing interactions of the H4 tail with the acidic patch from H2A.[57,181]

We wanted to compare the effects of differently acetylated H4 K16 on chromatin structure. We produced 12-mer nucleosome arrays containing the differently acylated H4 (Figure 29) and validated by MNase digest (Figure 30).

The effect of the acyl modifications were tested by performing a nucleosome self-association assay, which was induced by increasing Mg^{2+} concentrations.[182] The array containing WT H4 precipitated at Mg^{2+} concentrations of 1.5 mM which matches previously reported values of 1.0-2.0 mM Mg^{2+} for saturated 12-mer array to form the maximally folded fiber[182]. For arrays containing H4 K16ac, the required Mg^{2+} concentrations was significantly higher with ~4.5 mM for array precipitation. This is also confirming already published values, thus the site-specific acetylation was introduced via a native-chemical ligation strategy.[74]

The Mg^{2+} concentrations required for precipitation of array containing the H4 K16 propionyl, crotonyl and butyryl were with 3 mM between the WT and the strong effect observed for the acetylated arrays. This indicated a similar effect on higher order structure formation as the acetylation though less strong. However, the slope of the curves observed for WT and K16ac is much steeper as for the other acyl modifications, which could point towards an effect caused by subsaturation of the arrays instead of interruptions through the acyl modifications. A subsaturation of the arrays compared to saturated ones increase the Mg^{2+} concentration required for array precipitation.[160] To make a clear statement, on what effects the other acyl modifications cause as compared to the acetylation on K16, the experiments have to be repeated, including triplicate measurements of each sample and the saturation of the arrays need to be determined carefully prior to self-association assay to ensure functional comparison of the differently acylated arrays.

The *E. coli* HDAC CobB was indicated to have an effect not only as deacetylase but also as more general deacylase, as found by incomplete acylation when expressions were carried out lacking the CobB inhibitor NAM.[17,175,176] We wanted to verify this effect using an assay where we incubated differently acylated H4 with purified CobB and analyzed remaining acylation signals by comparing modification detection with pan α-H4 and α-H4K16ac antibodies. There was a complete loss of the acylation after 20 min (Figure 33). We evaluated this effect in more detail, by taking samples from the deacylation reactions following a time-span of 30 min. Thereby we found that all acylations were fully removed within 10 min and variations of the catalytic activity of CobB was dependent on acylation type, showing a decreasing reaction time with increasing acyl chain length (Figure 34).

These results are contradictory to previous findings by Denu et al. who compared activities of sirtuins regarding different acylated peptides. They showed decreasing activity with increasing chain length.[117] Bheda et al. compared deacylation of acetyl and propionyl groups of Sir2 *Thermotoga maritima* on acylated p53 peptides and found depropionylase activity, though with slightly lower activity as for the deacetylation.[183] The deacetylases described all belong to the sirtuin family, which commonly perform deacetylation NAD^+ dependent as described in Figure 8. The catalytic core between the different sirtuins is conserved, thus impeding a conformational recognition rather than a sequenced based one. However their variable N- and C-terminus might possibly confer to substrate specificity and protein-specific functions.[163,184]

It is difficult to compare our results with the results by other studies, as those were based on acylated peptides (by Denu et al. H3 peptides with K14mod[117]) whereas we used full-length H4 with the modification on K16. Furthermore the different findings on acyl chain length dependent deacylation velocities were based on different sirtuin species which contribute to divergent results. This is underpinned by Garrity et al. who investigated depropionylase activity on the *salmonella enterica* propionyl-CoA-synthetase by different sirtuins including human Sirt2, Sirt3, bacterial CobB, Sir2 of *Thermotoga maritima* ($Sir2_{Tm}$) and murine Sirt1.[119] Not all sirtuins revealed depropionylase activity and those that did showed significant variations. $Sir2_{Tm}$ performed four times faster compared to CobB and the human Sirt2 was 2.5 fold faster than $Sir2_{Tm}$. But Sirt3 was only 1.5 fold faster than CobB and only 0.3 times as fast as Sirt2 and murine Sirt1 was not active in depropionylation.[119] Sirt5 that was discovered as more active demalonylase and desuccinylase than deacetylase.[101,114]

It is therefore very likely that some other deacetylases maybe harbor a higher specificity towards other acylations than the specific acetylation. Our approach to produce site-specifically acylated full-length H4 provides a valuable contribution towards probing deacylation efficiencies on full-length proteins instead of only peptides and furthermore would also allow for experiments addressing effects of the acyl modifications on octamers, mononucleosomes or nucleosomal arrays. This would contribute to unraveling the specific functionalities provided by the wide variety of acyl groups.

3.2 Attempts towards protein labeling with fluorophores

The dynamical properties of the nucleosome is the target of ongoing research. Thereby often using FRET as one particular technique that provides the possibility to follow structural changes by distance-dependent fluorescence transfer. For addressing nucleosome dynamics, the FRET pairs were often achieved by fluorescent protein tags coupled with labeling of DNA or cysteine rsidues.[142–144,147] Hereby provides only the Cys labeling positional freedom. Other attempts including the combination of Cys labeling with a functional UAA have been previously addressed by Brustad et al. and Milles et al. using *p*-acetylphenylalanine, a ketone that was labeled via hydrazide or alkoxyamino functinonal dyes on T4 ligase or Nucleoporing153FG.[185,186] In another study, Nguyen et al. combined cysteine labeling with the labeling of the genetically encoded 1,2-aminothiol UAA with 2-cyanobenzothiazole (CBT) on T4 ligase and additionally calmodulin.[187] In addition, Milles et al. investigated different azide UAAs in their performance on click reactions and combined with maleimide labeling on RanBP3.[133]

We wanted to apply a similar strategy, based on the combination of thiol and genetically encoded functional UAA on histones. Histone H3 was used as it has previously been established for UAA incorporation and thiol labeling[28,142], for the installation of a FRET pair and subsequently reconstitution of octamers and nucleosomes. Summarized in Figure 54.

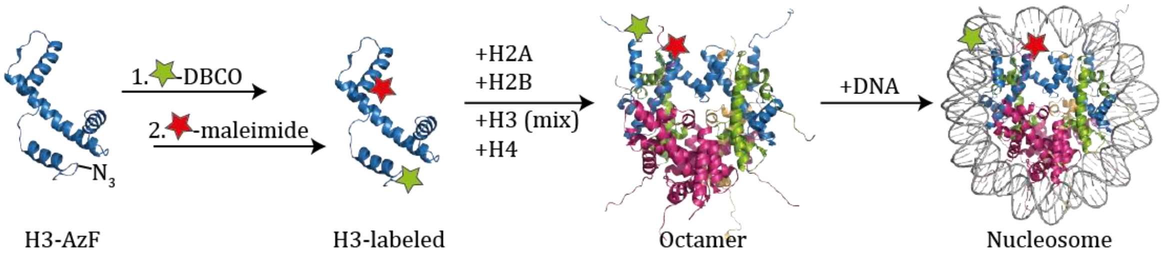


Figure 54 ***Summary creation of nucleosomes containing a FRET pair.***

The donor dye was installed on an azido UAA (AzF) and the dye attached via click chemistry as performed with a strained DBCO dye instead of the Cu(I) catalyzed variant. The acceptor should be provided by thiol labeling of the cysteine.

Histone H3 contains one native cysteine at position C110 but labeling at this position was reported to prevent octamer formation.[164] So we first investigated possible positions for labeling of the thiol functionality with a maleimide dye that still allowed for efficient octamer formation. We showed that, of the different positions tested for labeling, only position C109, which still lies within the H3-H3´ four helix bundle, resulted in loss of octamer formation. All the other positions tested revealed efficient octamer formation (summarized in Figure 55).

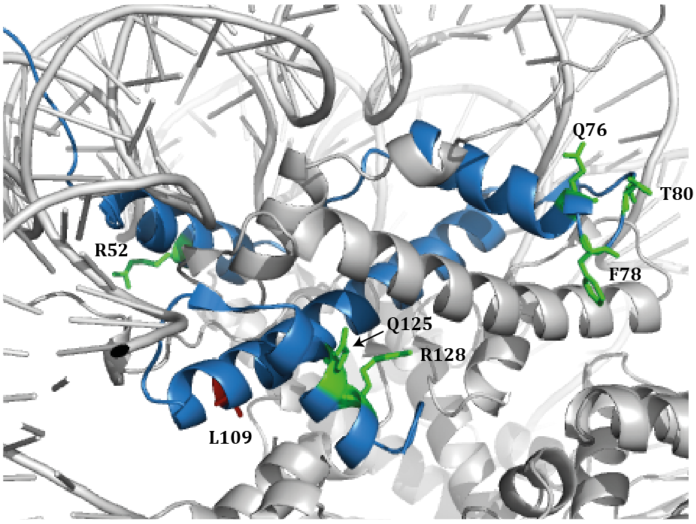


Figure 55 ***Overview of positions replaced by Cys and tested for labeling.*** Histone H3 is presented in blue, all other histones are shown in grey for simplification. The positions highlighted in green revealed efficient octamer formation whereas position L109C (indicated in red) prevented octamer formation. The picture was created using Pymol with PDB: 1AOI.

Cys residues are commonly reduced prior to labeling.[125] This is in general achieved by using either DTT or TCEP. DTT requires an additional purification step to be removed, as it otherwise reacts with the thiol functionality. In contrast TCEP is considered to be nonreactive towards thiol functionalities[123] and was therefore not removed prior to maleimide labeling.

However, TCEP is a phosphine, which reacts with azides in a Staudinger ligation. Staudinger ligations between genetically encoded azides that were reacted with phosphines also contribute to the toolbox of bioorthogonal reactions.[134,188,189] But as the azide group in our approach should be labeled via a click reaction, which has the advantage to be approximately 10 times faster than the Staudinger ligation[190], the quenching of the azide functionality by TCEP was undesired. This excluded simultaneous labeling with both dyes and furthermore defined the order of the successive labeling steps, as to perform the DBCO click reaction prior to the cysteine labeling.

The high specificity of the DBCO reaction was verified as a proof of principle experiment by labeling GFP protein as either WT or with AzF at position 150, as shown in Figure 37. Histone H3 was produced with AzF positioned at either T6, A24 or V35 (as shown in Figure 39).

This allowed for investigation of the conditions for labeling with DBCO on the target protein. Equal concentrations of protein to dye worked best and reactions performed at RT showed the most efficient progress by comparison of different reaction times. A high excess of dye caused precipitation of H3 displayed by decreased signals for both fluorescence and protein concentration (Figure 40). This effect was observed for all reaction temperatures tested. A strategy which allowed for buffer exchange between the two successive labeling steps was explored and conditions for double labeling examined. Again protein to dye ratios with only slight excess of dye (about 1.5 molar excess) for both labeling steps performed best.

The doubly labeled histone H3 efficiently formed octamers when used in refolding reactions. The octamer contains two copies of each histone[57], but two FRET pairs within one octamer might complicate interpretation of FRET values.[142] To prevent that from happening, the labeled H3 species was mixed with WT for octamer refolding reactions. Even with this low amount of labeled protein within the refolding reaction, we obtained octamers which showed good fluorescence signals as measured “in-gel” (Figure 42).

As the different commercially available dyes differ greatly in their spectroscopic properties and environmental affection[191,192], we tested two different donor dyes, DBCO488 and DBCO545 in combination with “red” acceptor dyes (Alexa647, Atto647N and Atto633) in refolded octamers. DBCO488 is not optimal as a FRET donor in combination with red dyes, as seen by less efficient pseudo-FRET, especially the combination with Atto633 revealed only weak pseudo-FRET. This is explained by the smaller spectroscopic overlap of DBCO488 compared to DBCO545 with the properties of red acceptors. Therefore only DBCO545 samples were used as donor for nucleosome assembly. The successful formation of nucleosomes could be shown though in the lower concentrations of the nucleosomes on the gel, the advantage of Alexa647 being brighter compared to Atto647N became obvious.

The positions that were replaced by amber codons for incorporation of AzF were T6, A24 and V35. Position V35 is located closely to the DNA entry-exit point[57], and a position frequently chosen by other studies that created FRET pairs by combining DNA labeling with

V35C replacement and maleimide labeling.[142,145] Position T6 lies in the unstructured, highly flexible regions of the N-terminal tail, thus complicating predictions if FRET might occur. This matches our observations, as for “in-gel” measurements of octamer and nucleosomes no pseudo-FRET signal occurred.

Nonetheless, a clear statement on the quality of the applied methodology is currently difficult, as this requires determination the labeling efficiency. Effects like reduction of the azides to the non-reactive amine[193,194] and TCEP being not as inactive as previously considered[195,196], might contribute to a reduced labeling efficiency. To evaluate those effects we compared different strategies to remove the free dye from the reaction, which would allow for the quantification of individual reaction efficiencies. Thereby showing that precipitation of the labeled H3 using MeOH/$HCCl_3$ leads to removal of free dye with almost quantitative protein recovery.

In summary, we could present a method that allowed for the installation of a FRET pair site-specifically on histone H3 with good position-freedom. However detailed investigation of the quantitative labeling efficiencies remains elusive. Single-pair FRET (spFRET) measurements allow to exclude inefficiently labeled species and can select for specific sub-populations to investigate them individually.[142] Since our approach was meant to present a method to create samples for spFRET measurements this would tolerate some non-stochiometrically labeled samples.

For histone proteins, that only rarely contain cysteine residues, this method provides a suitable labeling pathway. However disadvantages like the requirement of reducing agents thereby requiring successive labeling steps hamper the efficiency of this system. Those limitation can be overcome when introducing two distinct UAAs which would provide a greater variability in functionalities and a more accurate choice of labels.

Therefore, besides the amber codon, the ochre codon or four-base codons were explored for UAA incorporation.[50,53,197] An orthogonal ribosome as developed in the Chin group, allowed amber suppression with high efficiency and decodes four-base codons.[54,55,198] Nonetheless the efficiency for incorporation of two distinct UAAs is poor. Additionally this creates high metabolic stress for the cells as two tRNA/AARS pair plasmids of different selective antibiotics are involved. The optimization of this system was addressed by Christoph Lammers, who optimized the expression levels of the individual components.[171] He further constructed a modular genetic tool, a plasmid that combined MbPylS/t and MjYRS/t, thereby requiring only three plasmids instead of four. This system reduced the metabolic burden of the cells and resulted in more versatile cells.[171] Taking advantage of this system we could introduce not only Tyr and BocK, but also successfully produce GST-MBP carrying two distinct functional UAAs for chemical derivatization (AzF and PrgK). Comparing protein yields, as obtained from expression with the components on individual plasmids (four in total) and with the optimized version (three plasmids in total, and PylS controlled by P_{lac} promoter) showed significantly enhanced yields with the three plasmid system as compared to the initial four plasmids (Figure 46) for both amino acid combinations. The two functional UAAs AzF and PrgK were reacted separately with fluorescent dyes performing either a Cu(I) catalyzed click reaction or a strain promoted click reaction using an DBCO-dye. Both functionalities were selectively labeled as shown by specific fluorescence signals when excited with appropriate wavelengths (Figure 48). However, those two reactions can only be performed separately, as combining both dyes simultaneously with the protein would lead to reactions between the two dyes, hampering functional labeling. With these two UAAs, double labeling on one protein requires successive labeling steps, including removal of free dye from the reaction before performing the second labeling step. Which would again require optimization of the individual steps.

The click reaction between azide and a strained alkyne can not only be performed by incorporation of AzF but an inverted variant, based on the incorporation of strained cyclooctynes[199] with azide dyes are possible. However, the advantage of the labeling method described in this thesis, is its accessibility as the compounds, AzF and several DBCO-dyes, are already commercially available as holds true for the maleimide dyes. Other compounds, that react efficiently with tetrazines on timescales faster than that of the strained azide-alkyne click reaction[190], might be considered as alternatives to expand the feasibility of this system.

An even better approach would be the direct encoding of a fluorescent UAA which eliminates additional modification steps, thereby providing protein quantitatively carrying the fluorescence label, only dependent upon incorporation efficiency. So far, however, only very few fluorescent UAAs are accessible to be directly encoded.[13,128,129] We aimed to expand the number by the environmental sensitive fluorophore NBD-K.[200]

Using fluorescence microscopy we could show efficient uptake of the UAA by *E. coli*. Directed evolution to obtain a NBD-K specific synthetase revealed several possible candidates. Some were selected and tested in protein expressions. The two candidates, which selectively incorporated NBD-K in response to the amber codon showed that Leu274 was replaced by a smaller, aliphatic amino acid (Ala or Val) which also replaced the polar Cys313. Met315 was replaced by a more compact amino acid being either Cys or Leu. As expected, all those changes offer more space to accommodate the bulkier NBD within the active-site. The nonprogrammed mutation Thr248Ser can be considered as having no functional contribution, as this position is not located within the binding pocket. The synthetase NBDKRS45 was further used in small scale expressions for myoglobin, with the UAA incorporated at position 4. In summary these experiments are very promising, as we found an additional fluorescent UAA for direct genetic encoding. Thus, more detailed studies and verification by mass spectrometry have to be performed.

4 Materials and Methods

4.1 Materials

4.1.1 Machines

AbiPrism 3100 DNA Sequencer	Applied Biosystems, Darmstadt
Accumax Pipet Help	Accumax, India
Agarose gel electrophoresis chamber	PeqLab, Erlangen
	GE Healthcare, München
Äkta Prime	GE Healthcare, Freiburg
Äkta Purifier	GE Healthcare, Freiburg
Autoclave HST 4-5-8	Zirbus, Bad Grund
Bunsen burner Fuego basis	WLD-Tec, Göttingen
Centrifuge Allegra 2IR	Beckman Coulter, Krefeld
Centrifuge Avanti J-20 XPIJA-20	Beckman Coulter, Krefeld
Centrifuge HERAEUS Pico 17	Thermo Scientific, Langenselbold
Centrifuge 5415R	Eppendorf, Hamburg
Concentrators MWCO 3 k, 10 k	Sartorius, Göttingen
Concentrator 5301	Eppendorf, Hamburg
Electroblotter semi-dry	PeqLab, Erlangen
Electroblotter wet	BioRad, München
Electroporator Easyject Prima	EquiBio, Willstätt
FLUOstar Omega	BMG Labtech, Offenburg
Gel Doc gel documentation machine	BioRad, München
Gel shaker Rotamax 120	Heidolph, Kelheim
Hamilton Syringe 50 µL	Hamilton, USA
Incubator Mytron	Schütt, Göttingen
Labcycler	SensoQuest, Göttingen

Magnetic stirrer MR Hei-Standard	Heidolph, Kelheim
Microfluidizer 110 S	Microfluidics, USA
Nano Drop ND-1000 Spectrophotometer	Thermo Scientific, Langenselbold
Optimax X-ray film processor	Protec, Oberstenfeld
Orbital shaker Rotamax 120	Heidolph, Schwabach
pH meter PT-15	Satorius, Göttingen
Photometer	Biometra, Göttingen
Photometer	Eppendorf, Hamburg
Pumpdrive 5001	Heidolph, Kehlheim
Power Supply EV231	Consort, Belgium
Rotor *JA-20*, *JLA-8.1000*, *JA-30*	Beckman Coulter, Krefeld
Sonifier 250	Branson, USA
Special accuracy balance	Sartorius, Göttingen
Superloops (50 mL, 150 mL)	Amersham Pharmacia Biotech, Freiburg
Thermoshaker comfort	Eppendorf, Hamburg
Unitron Incubators	Infors HT, Einsbach
Vertical Double Gel Systems	PeqLab, Erlangen
Vertical gel Electrophoresis chamber-	GP_Kunststofftechnik, Kassel
H10 Mini Vortex generator VV3	VWR, Hannover

4.1.2 Other materials

Coli Rollers Plating Beads	Novagen, USA
Disposal bags	Sarstedt, Nümbrecht
Dialysis tubing cellulose membrane 10 mm	Sigma-Aldrich, München
Electroporation Cuvettes, 2mm gap	PeqLab, Erlangen
Falcon tubes (15 mL, 50 mL)	Sarstedt, Nümbrecht
Gas Permeable Adhesive Seals	Thermo Scientific, Langenselbold

Hyperfilm ECL	GE Healthcare, UK
Immobilon-P Transfer Membrane (PVDF)	Merck Millipore, Darmstadt
Instant Blue	Biozol, Eching
Low Binding E Tips 10 µL	Sorenson, BioScience, USA
Low Binding 1.7 mL cups	Corning Costar, USA
Micro tubes 1.5 mL, 2.0 mL	Sarstedt, Nümbrecht
Ni^{2+}-NTA column	Qiagen, Hilden
Nitrocellulose Transfer Membrane, Protran	Whatman, Dassel
PCR Softtubes (0.2 mL)	Biozym, Austria
Petri dish 92 x 16 mm with cams	Sarstedt, Nümbrecht
Petri dishes 145/20 with vent	Greiner Bio One, Solingen
Pipets Research plus (10, 100, 1000 µL)	Eppendorf, Hamburg
Pipet tips	Sarstedt, Nümbrecht
Plate 96 DW Square U	VWR, Hannover
Spectra/Por Dialysis Membrane 50 mm	Spectrum Laboratories, Netherlands
Spectra/Por Closures 23 mm, 55 mm	Spectrum Laboratories, Netherlands
Spin-X Centrifuge Tube Filter	Costar, USA
UV Cuvettes	Eppendorf, Hamburg
Whatman 3MM Chr paper	Whatman plc, UK
96-well Microplates	BD Falcon, USA

4.1.3 Chemicals

GelRed	Biotium, USA
Imidazole	Roth, Karlsruhe
Urea	Merck Millipore, Darmstadt
Guanidinium hydrochloride	Sigma (Fluka)
Nicotinamide	Sigma, Steinheim

All other chemicals were purchased from either AppliChem (Darmstadt), Sigma-Aldrich (Steinheim), Merck (Darmstadt) or Roth (Karlsruhe) in the quality *pro analysis*.

4.1.4 Unnatural amino acids

Table 4-1 ***Unnatural amino acids, concentrations and plasmids used for incorporation***.

Unnatural amino acid	Plasmid used for incorporation	Work concentration	Company
Acetyllysine (AcK)	pBK AcKRS3[28]	10 mM	Bachem, Switzerland
Butyryllysine (BuK)	pBK PylS	2 mM	ChemImpex (USA)
Crotonyllysine (CrK)	pBK CrKRS 6/PylS	2 mM	ChemImpex (USA)
Propionyllysine (PrK)	pBK PrKRS 13	2 mM	ChemImpex (USA)
N(epsilon)Boc-lysine (BocK)	pBK PylS	1 mM	Bachem, Switzerland
Azidophenylalanine (AzF)	pCNPheRS[201]	2 mM	Bachem, Switzerland
Propagyllysine (PrgK)	pBK PylS	2 mM	Synthezied after Nguyen et al.[132]

4.1.5 Fluorophores

Table 4-2 ***Fluorescent dyes.***

Dye	Abs Max [nm]	Em. Max [nm]	Conjugate	Company
DBCO-545	546	565	Strained alkyne	Jena Biosciences
DBCO-488	501	525	Strained alkyne	Jena Biosciences
Alexa-647-C2-mal	651	671	maleimide	Thermo Scientific, Langenselbold
Alexa-647-Az	651	671	azide	Thermo Scientific, Langenselbold
Atto-647N	644	669	maleimide	Atto-Tec, Siegen
Atto633	629	657	maleimide	Atto-Tec, Siegen

4.1.6 Kit systems

PEQ Gold Plasmid Mini Prep Kit	PeqLab, Erlangen
QIAquick gel extraction kit	QIAGEN, Hilden
DNA Gel Extraction Kit	Merck Millipore, Darmstadt
QIAquick PCR purification kit	QIAGEN, Hilden
Plasmid Giga Kit	QIAGEN, Hilden
Pierce ECL Plus Western Blotting Substrate	Thermo Scientific, USA
ECL Prime WB Detection Reagent	GE Healthcare, UK
ECL Select WB Detection Reagent	GE Healthcare, UK

4.1.7 Chromatography columns and resins

HisTrap FF, 5 mL	GE Healthcare, Freiburg
HiTrap Q XL, 5 mL	GE Healthcare, Freiburg
HiTrap SP FF, 5 mL	GE Healthcare, Freiburg
HisPur™Ni-NTA Resin	Thermo Scientific, München
Superdex 200 10/300 GL	GE Healthcare, Freiburg
Superdex 75 26/70	GE Healthcare, Freiburg
Superdex 200 26/70	GE Healthcare, Freiburg

4.1.8 Buffers and solutions

10x DNA Loading Dye for agarose gels	20 % Glycerol 0.25 % Bromphenole blue 0.25 % Xylene cyanol
4x Loading buffer for SDS gel	50 mM Tris-HCl pH 6.8 100 mM DTT 2 % SDS solution 0.1 % Bromphenol blue 10 % Glycerol

Ni wash	10 mM Tris-HCl pH 7.5 20 mM Imidazole 200 mM NaCl
1x PBS	137 mM NaCl 2.7 mM KCl 10 mM Na_2HPO_4 1.76 mM KH_2PO_4 pH 7.4
1000x Protease Inhibitor Cocktail	75 mM Pefablock 150 μM Leupeptin hemisulfate salt 37.5 mM *O*-Phenanthroline hydrochloride-monohydrate 500 μM Pepstatin A
Ponceau S	0.5 % (w/v) Ponceau S 5 % Trichloroacteic acid
10x Running buffer	250 mM Tris base 1920 mM Glycine 35 mM SDS
1x Transfer buffer (western blot)	25 mM Tris base 192 mM Glycine 3.5 mM SDS 20 % Methanol
TB	89 mM Tris base 89 mM Boric acide
TBE	89 mM Tris base 89 mM Boric acide 2.5 mM ETDA
TBS	50 mM Tris-HCl pH 7.5 150 mM NaCl

4.1.9 Composition of SDS gels (for six gels)

Table 4-3 ***Composition of SDS gels.***

	Stacking gel 4 %	**Separation gel 12.5 %**	**Separation gel 15 %**	**Separation gel 18.5 %**
H_2O	12 mL	21.8 mL	16 mL	9.8 mL
Tris-HCl	5.0 mL (0.625 M, pH 6.8)	17.6 mL (1.5 M, pH 8.8)	17.6 mL (1.5 M, pH 8.8)	17.6 mL (1.5 M, pH 8.8)
10 % SDS	200 µL	750 µL	750 µL	750 µL
Arylamide (37.5:1)	2.6 mL	29.2 mL	35.0 mL	41.2 mL
10 % APS	100 µL	750 µL	750 µL	750 µL
TEMED	30 µL	60 µL	60 µL	60 µL

4.1.10 Media and culture plates

1 L LB medium contains: 10 g Tryptone/Peptone, 5 g Yeast extract and 5 g NaCl. For preparation of agar plates the medium was supplemented with 1.5 % agar and autoclaved. The selective antibiotics were added before pouring the plates.

4.1.11 Antibiotics and work concentrations

Table 4-4 ***Antibiotics and work concentrations.***

Antibiotic	Work Concentration	Company
Ampicillin	50 µg/mL	AppliChem, Darmstadt
Chloramphenicol	50 µg/mL	AppliChem, Darmstadt
Kanamycin	50 µg/mL	AppliChem, Darmstadt
Spectinomycin	75 µg/mL	Sigma, Steinheim
Tetracycline	25 µg/mL	AppliChem, Darmstadt

4.1.12 Enzymes, buffer and protein/DNA size standards

Gene Ruler^{TM} DNA Ladder Mix	Thermo Scientific, USA
Gene Ruler 100 bp Plus DNA Ladder	Thermo Scientific, USA
Quickload 1 kb DNA Ladder	New England Biolabs, USA

Unstained Protein Molecular Weight Marker	Thermo Scientific, USA
PageRuler Prestained Protein Ladder	Thermo Scientific, USA
Lysozyme from chicken egg	Serva, Heidelberg
T4 DNA Ligase	Thermo Scientific, USA
T4 DNA Ligase buffer	Thermo Scientific, USA
Polymerase Expand High Fidelity	Roche, Grenzach-Wyhlen
Expand High Fidelity Buffer (10x)	Roche, grenzach-Wyhlen
Pfu Turbo DNA Polymerase	Agilent, Böblingen
CIP	New England Biolabs, Frankfurt
DpnI, PstI, NdeI, SalI, MNase	Thermo Scientific, USA
BsaI, DdeI, HaeII, EcoRI, BfuCI	New England Biolabs, Frankfurt
AvaI Fast digest	Thermo Scientific, USA
NEB buffer 4	New England Biolabs, Frankfurt
NEB 10x buffer	New England Biolabs, Frankfurt

4.1.13 Cell lines

Table 4-5 ***Cell lines.***

Strain	Genotype	Reference
DH10B	F$^-$ endA1 recA1 galE15 galK16 nupG rpsL ΔlacX74 Φ80lacZΔM15 araD139Δ(ara,leu) 7697 mcrA Δ(mrr-hsdRMS-mcrBC) λ$^-$	Invitrogen (18297010)
BL21 (DE3)	F$^-$ *ompT* *gal dcm lon* hsdS$_B$(r$_B^-$ m$_B^-$) λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5])	Invitrogen (C600003)
Rosetta 2 (DE3) (Novagen)	F$^-$ *ompT hsdS*$_B$(r$_B^-$ m$_B^-$) *gal dcm* pRARE2 (Cam^R)	Merck Millipore (71402)
C321.ΔA.exp	MG1655 Δ(ybhB-bioAB)::zeoR ΔprfA; all 321 UAG codons changed to UAA. mutS+	Addgene (49018)
dam$^-$/dcm$^-$ competent *E. coli*	ara-14 leuB6 fhuA31 lacY1 tsx78 glnV44 galK2 galT22 mcrA dcm-6 hisG4 rfbD1 R(zgb210::Tn10) *Tet*^{S} endA1 rspL136 *(Str*^{R}*)* dam13::Tn9 *(Cam*^{R}*)* xylA-5 mtl-1 thi-1 mcrB1 hsdR2	NEB (C2925H)

4.1.14 Plasmides

pBAD sfGFP

pBAD sfGFP150TAG

pBar PylT

pBK pCNPheRS

pBK PylS

pBK PylS Lib

pBK-AcKRS3[28]

pCDF Duet1 CobB (Svenja Heitmüller)

pCDF Duet1 Hst2 (Svenja Heitmüller)

pCDF H3 opt. PylT A24TAG

pCDF H3 opt. PylT T6TAG

pCDF H3 opt. PylT V35TAG

pCDF H3 opt. PylT R52TAG

pCDF-PylT-AraC-His$_6$-H3(Δ93-98)-TEV-H4 (Bryan Wilkins)

pCDF-PylT-AraC-His$_6$-H3(Δ93-98)-TEV-H4K16TAG (Bryan Wilkins)

pDULE pCNPheRS

pET H2A

pET H2B

pET H3

pET H4

pET24 TEV (kindly provided by Achim Dieckmanns)

p*Myo*4TAGPylT-*his$_6$*[17]

pREP PylT

pUC18_12x200x601[165](kindly provided by Daniela Rhodes)

4.1.15 Antibodies

Table 4-6 *Antibodies*.

Target	Host	Supplier/Catalog number	Dilution
Primary Antibodies			
6 x HIS epitope	Mouse	GE (27-4710-01)	1:6 000 in 3% BSA-PBS
Histone H4	Rabbit	Abcam, UK (ab7311)	1:5 000 in 3% BSA-TBS
Histone H4 acetyl Lys 16	Rabbit	Active Motif (39167)	1:15 000 in 3% Milk-TBS
Crotonyllysine	Mouse	PTM-Biolabs (PTM501)	1:5 000 in 5% BSA-TBS*
Secondary Antibodies			
Mouse HRP conjugated	Goat	Sigma (A4911)	1:10 000 in 5% Milk-PBS *1:10 000 in 5% BSA-TBS
Rabbit HRP conjugated	Goat	Sigma (A0545)	1:10 000 in 3% Milk-TBS

4.1.16 DNA oligonucleotids

The oligonucleotids used in this work were purchased from Sigma-Aldrich in desalting purity. They were diluted following the manufacturer´s instructions. A solution of 10 mM was prepared for use.

Name	Sequence 5´-3´	Tm (°C)
PylSAYLf (42)	GCGCAGGTCTCAATGCTGNNKCCGACCCTGNNKAACTATNNKCGTAAACTGGATCGTATTCTGCCGGGTC	89.9
PylSSAYLr (45)	GCGCAGAGTAGGTCTCAGCATCGGACGCAGGCACAGGTTTTTATC	86.8
PylSCMf (43)	GCGCAGGTCTCAAACTTTNNKCAANNKGGCAGCGGCTGCACCCGTGAAAACCTGGAAG	92.0
PylSCMr (46)	GCGCAGAGTAGGTCTCAAGTTAACCATGGTGAATTCTTCCAGGTGTTCTTTG	84.0
T7 Promoter	TAATACGACTCACTATAGG	46.2
pBKseqf (50)	CTCGGGTTGTCAGCCTGTC	66.0
pBKseqr (51)	CAGAGATCATGTAGGCCTG	58.3
H3_T6TAG f (141)	GGGCCGTACCAAACAGTAGGCGCGTAAAAGCACC	82.2
H3_T6TAG r (142)	CGCCGGTGCTTTTACGCGCCTACTGTTTGGTACG	83.4
H3_A24TAG f (143)	CAGCTGGCCACCAAATAGGCCCGTAAAAGCGCG	84.8
H3_A24TAG r (144)	CGGCGCGCTTTTACGGGCCTATTTGGTGGCCAGC	87.8
H3_V35TAG f	CCGGCGACCGGCGGTTAGAAAAAACCGCATCGTTATCG	87.4

H3_V35TAG r	CGATGCGGTTTTTTCTAACCGCCGGTCGCCGGCGC	90.4
H3opt R52TAGneuf	CGTGAAATTTAGCGTTATCAGAAAAGCACCGAACT	75.4
H3opt R52TAGneur	CTGATAACGCTAAATTTCACGCAGCGCCACGGTGC	83.6
H3opt F78TAGf	TGCGCAGGATTAGAAAACCGATCTGCGTTTTCA	80.4
H3opt F78TAGr	GATCGGTTTTCTAATCCTGCGCAATTTCACGCACC	81.3

N = A+C+G+T, K = G+T

4.2 Molecular biology methods

4.2.1 Purification of plasmid DNA “Miniprep”

Bacteria were grown on agar plates containing a selective antibiotic. One single colony was picked with a sterile pipette tip and inoculated in 4 mL LB media with the same antibiotic. Following an overnight incubation at 37 °C 2 mL of this culture were harvested by centrifugation (13.300 rpm, 10 min) at RT. Plasmid DNA isolation and purification was performed according to the manufacturer´s instructions using PEQ Gold Plasmid Mini Prep Kit.

4.2.2 Large scale purification of plasmid DNA “Giga Prep”

E. coli cells were transformed with the plasmid of interest and grown overnight in LB with selective antibiotics. The next day 5 L of LB medium containing the selective antibiotics were inoculated and grown overnight. The cells were harvested and the plasmid isolated according to manufacturer´s instructions using a Plasmid Giga Kit (QUIAGEN).

4.2.3 Agarose gel electrophoresis

Agarose gels were prepared by melting 1% (w/v) agarose in 0.5 x TBE buffer supplemented with “GelRed” in a concentration of 1:20 000. The DNA samples were diluted 5:1 with loading dye. Gels were run at constant 120 V at room temperature with 0.5 x TBE as running buffer. DNA bands were visualized by UV light with wavelengths of 254 nm for analytical and 365 nm for preparative purposes.

4.2.4 DNA-elution from agarose gels

After preparative restriction digestion of the DNA, the fragments were separated by agarose gel electrophoresis. The bands of correct size were cut out of the gel under UV light (365 nm) and elution was performed using a QIAquick Gel Extraction Kit following the manufacturer´s protocol. The purified DNA was eluted from the column using 50 µL EB buffer and stored at –20 °C.

4.2.5 Polymerase chain reaction

Genes and fragments were amplified by polymerase chain reaction (PCR).[202]

The reaction cycle includes five steps:

1. Initial denaturing	to ensure that all templates are denatured before starting the cycles
2. Denaturing	the reaction mixture is heated to 95 °C yielding in single stranded DNA
3. Annealing	the temperature is decreased to 5 °C lower that the melting point of the oligonucleotids allowing them to hybridize
4. Elongation	the temperature is changed to the optimum for the synthetase to produce the complementary DNA strand
5. Final elongation	this final step ensures that all remaining single strands are fully extended

Steps 2 to 4 were performed in 20 to 30 cycles using the new generated DNA strands as targets for the following cycle. Following step 5 the reaction mixture was cooled to 4 °C until further usage. The PCR program was optimized for each template and primer mixtures.

Table 4-7 ***General set up for a PCR reaction.***

	Volume
Template DNA	2 µL
Forward primer	2 µL
Reverse primer	2 µL
dNTPs	1 µL
10x buffer	5 µL
Taq polymerase	1 µL
ddH$_2$O	To 50 µL

4.2.6 Mutagenesis PCR “QuickChange”

Site directed mutagenesis PCR incorporates a desired bp change in the sequence. Two complementary primers were designed by flanking the desired mutation at 25 bp. Primers were chosen approximately 10 bp before the mutation and 15 afterwards for each sense and antisense primer. The enzymatic set up is described in Table 4-8.

Table 4-8 ***Enzymatic set up for quick change reactions*.**

	Volume
Template DNA	1.0 µL
Forward primer	1.0 µL
Reverse primer	1.0 µL
dNTPs	1.5 µL
Pfu buffer	5.0 µL
ddH_2O	39.5 µL
Turbo Pfu	1.0 µL

The temperature program for quick change reactions is listed in Table 4-9.

Table 4-9 ***Temperature profile for quick change program.***

Step	**Temperature**	**Time**	
1. Initial denaturing	95 °C	3 min	
2. Denaturing	95 °C	30 sec	15 cycles
3. Annealing	60 °C	30 sec	
4. Elongation	72 °C	8 min	
5. Final elongation	72 °C	10 min	

Following the PCR, DpnI (see 4.2.8) was added to the mixture, incubated and plasmids transformed into *E. coli* DH10B cells using heat shock method (described in 4.3.3). The cells were plated onto agar plates containing selective antibiotics and the DNA was purified as described in 4.2.1. Sequencing was done with T7 Promoter (4.2.12) clones containing the desired mutation were identified by DNA sequencing and finally used for protein expression.

4.2.7 Inverse PCR

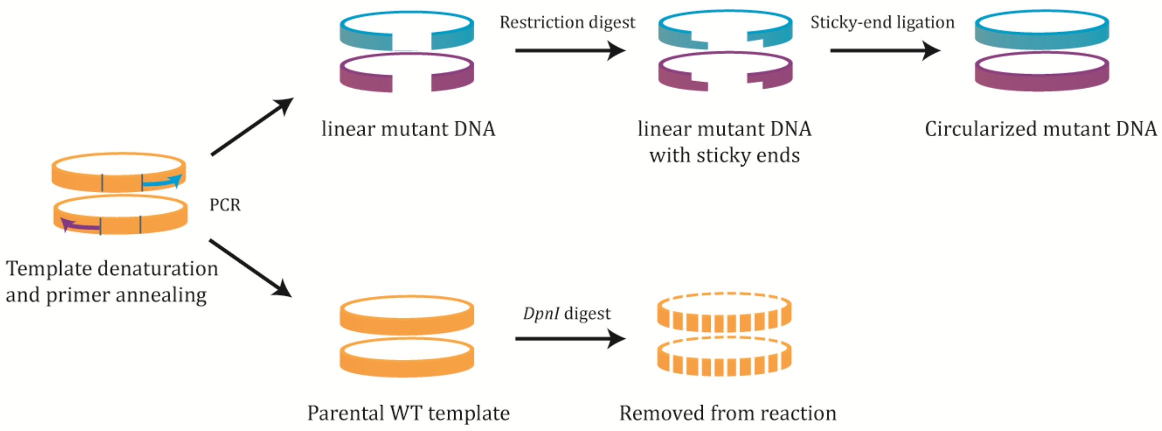

Restriction digest
Sticky-end ligation
linear mutant DNA
linear mutant DNA with sticky ends
Circularized mutant DNA
PCR
Template denaturation and primer annealing
DpnI digest
Parental WT template
Removed from reaction

Figure 56 ***Schematic procedure of the inverse PCR.*** The template is amplified, yielding in blunt end mutants. Template DNA is removed by DpnI digest and sticky ends are created upon linear mutant by BsaI digest. Finally sticky ends are ligated. Picture was modified from Williams et al.[203].

Inverse PCR[204] is a technique similar to common PCR but amplifies up to 10 kbp. The primers are designed to flank a region of a known sequence orientated reversely to the usual orientation. The recognition sequence of the primers is removed while creating a linear mutant DNA with blunt ends. The parental template is removed by DpnI digestion. Sticky ends are obtained by BsaI digestion. Sticky end ligation is performed yielding in circularized DNA.

Table 4-10 ***Set up for inverse PCR.***

	Volume
Template DNA	20 µL
Primer 1	8 µL
Primer 2	8 µL
dNTPs	20 µL
Expand HF Buffer (10x)	100 µL
ddH_2O	824 µL
Expand HF polymerase	20 µL

The reaction mixture was split into 50 µL portions followed by PCR.

Each tube was supplemented with 5 µL NEB (10x) buffer and 1.25 µL DpnI. After digestion for 1.5 h at 37 °C, 1 µL of BsaI was added and digestion continued (3 h, 50 °C). Ligation was performed overnight at 16 °C using T4 DNA ligase.

4.2.8 Restriction digestion of DNA

Plasmids in mutagenesis PCR were cut with DpnI which recognizes methylated DNA and removes methylated template DNA from the reaction. 100 µL of the PCR reaction were supplemented with 10 µL NEB (10x) buffer and 2.5 µL of DpnI. The reaction was performed for 1.5 h at 37 °C. After digestion the reaction product was either used in another digest and ligation reaction or directly transformed into *E. coli* DH10B cells and cultivated on agarose plates containing selective antibiotics.

Typ II restriction endonucleases cut DNA within their recognition sequence at defined positions producing either overhanging ends (sticky) or blunt ends. These ends can hybridize with other complementary strands and therefore allow insertion of a specific DNA fragment into a vector. As a general set up a total volume of 50 µL was used, including 5 µL buffer, 5 µL enzyme and 30-40 µL plasmid. The total volume was obtained by adding ddH_2O. Digestion was performed at 37 °C for 2 h. For double digests 20 µL of the first digestion were mixed with 2.5 µL buffer and 2.5 µL of a second enzyme and processed as described above. Afterwards 5`-phosphate groups were removed by adding *CIP* (calf alkaline phosphatase) and incubated for 0.5 h. The enzyme was finally deactivated by heating to 65 °C for 15 min.

4.2.9 Ligation reactions

For ligation of DNA fragments 2 µL vector DNA were mixed with insert DNA at a ratio of 1:3 in 1 µL ligase buffer supplemented with 0.5 µL T4 DNA ligase and incubated at 16 °C overnight. The next day the ligation reaction was directly transformed into *E. coli* DH10B cells and resulting colonies were sequenced to find the desired clone.

4.2.10 Ethanol precipitation of DNA

Large-scale ligation reactions were ethanol precipitated before transformation via electroporation. The ligation product (500 µL) was supplemented 10:1 with 3 M sodium acetate solution (50 µL, pH 5.2). Twice the total volume of the reaction mixture of abs. ethanol (1100 µL) was added and incubated on ice for 30 min. The DNA was pelleted via centrifugation (4 °C, 15 min, 13 200 rpm) and washed twice with 1 mL of 70% ethanol. The supernatant was discarded and the pellet was allowed to dry for 5 min, resuspended in 20 µL

ddH_2O and incubated at 50-60 °C for 10 minutes. The resulting DNA was transformed into *E.coli* DH10B cells.

4.2.11 Determination of nucleic acid concentration

The concentration of aqueous nucleic acid solutions was detected by measuring the absorption at 260 nm (A_{260}) and referred to pure water. The quotient of A_{260} to A_{280} showed the purity of the DNA. A quotient of 1.8 to 2.0 indicated reasonably pure DNA. A lower quotient revealed contamination with proteins.

4.2.12 Sequencing

For in-house facility sequencing, sequencing was performed according to Sanger´s dideoxy chain termination reaction. During amplification, dideoxynucleotide triphosphate (ddNTPs) are incorporated at random positions. These contain fluorophors of different emission wavelengths and terminate the chain elongation due to a lacking 3´-OH group. Based on the statistics, fragments with different lengths are contained in the resulting mixture. These fragments can be separated and analyzed. *Seq-MixBigDye Terminatorv* 1.1 (Appl. Biosystems) was used allowing the reaction to be performed in one tube. The enzymatic set up for sequencing reactions is listed in Table 4-11.

Table 4-11 ***Enzymatic set up for sequencing reaction*.**

Sequencing set up	
200-400 ng	Template DNA
8 pmol	Primer
1 µL	Seq-Mix
1 µL	Seq-buffer
Add to 10 µL	ddH_2O

Table 4-12 ***Temperature profile for sequencing PCR, cycles being repeated 25 times.***

Step	Temperature	Time	
Denaturing	96 °C	10 sec	25 cycles
Anneling	55 °C	15 sec	
Elongation	60 °C	4 min	

Following the PCR, the mixture was purified by adding 1 µL 125 mM EDTA, 1 µL 3 M sodium acetate and 50 µL ethanol (96%). The mixture was gently mixed, incubated for 5 min and centrifuged (20 °C, 5 min, 16100 x g). The pellet was washed with 70 µL ethanol (70%) and centrifuged as above. After drying the pellet (2 min, *speedvac*) it was redissolved in 15 µL formamide (99.5%) and analyzed using a capillary sequencing instrument (4.1.1).

Alternatively samples were sent to SEQLAB, Göttingen or GATC Biotech.

4.2.13 Preparation of DNA templates for chromatin assembly

DNA for nucleosome array reconstitution was prepared from pUC18_12x200x601[165] plasmid, which was kindly provided by Daniela Rhodes. The DNA was large scale amplified in *E. coli* C2925H cells and isolated (as described 4.2.2). It was digested overnight with DdeI, DpnI, HaeII and EcoRI (300 U/mg DNA respectively) in an appropriate buffer (CutSmart, NEB) at 37 °C with shaking. The vector backbone was separated from desired array DNA fragments by PEG 6000 precipitation. Digested DNA was mixed with NaCl to a final concentration of 0.5 M. Then, 3% PEG was slowly added, incubated for 15 min on ice and centrifuged for 15 min at 15 °C (full speed). The supernatant was removed and applied to the next round of precipitation, covering 3-10% PEG plus an additional step with 15% PEG. The obtained DNA pellets were washed with 70% EtOH, air dried, resuspended in water and, if necessary, additionally precipitated with *i*Propanol. The resulting fractions were analyzed by agarose gel electrophoresis using 1% agarose gel. Fractions containing the 2 kb fragment were pooled and used for assembly of nucleosome arrays.

4.2.14 Preparation of the pCDF-PylT-AraP-His$_6$-H3(Δ93-98)-TEV-H4 plasmid

The Bacterial expression vector, pCDF-PylT-AraP-His$_6$-H3(Δ93-98)-TEV-H4, was created by Bryan Wilkins as described in[174]. First, the *Xenopus laevis* histone H3 sequence (codon optimized for Bacteria) was PCR amplified without the stop codon. The PCR product was digested with BamH I and EcoR I and cloned into the same sites of pCDF PylT[28]. A DNA fragment that coded for a TEV protease recognition sequence was created using PCR and fused to *Xenopus laevis* histone H4 (bacterial codon optimized) and additionally flanked by EcoR I and Xho I sites. These were used to insert the fragment into the previous construct, yielding pCDF PylT His$_6$-H3-TEV-H4. The Nco I restriction site in the ORF of H4 was removed via QuickChange mutagenesis and then the T7 promoter was replaced by subcloning of the AraC ORF and *ara*BAD promoter from the plasmid pBAD/His A (Invitrogen) with Pfo I and Nco I. Finally, the H3(Δ93-98) deletion, and H4 K16 mutations were introduced by QuikChange mutagenesis.

4.3 Cell biology methods

4.3.1 Competent cells for chemical transformations

To obtain chemical competent cells, a 20 mL overnight culture (as described in chapter 4.2.1) was prepared and used to inoculate 250 mL prewarmed LB media (1:100). The main culture was incubated at 37 °C and at an OD_{600} of 0.6-0.8 cells were harvested by centrifugation (4 °C, 15 min, 4800 rpm). The pellets were washed with 10 mL cold calcium chloride solution (10 mM PIPES/KOH (pH 7.0), 60 mM $CaCl_2$ and 15% glycerol) and centrifuged as described above. During the next washing round cells were incubated for 1 h on ice before centrifugation. The supernatant was discarded and cells concentrated in 2 mL. Cell aliquots of 200 μL were frozen in liquid nitrogen or directly used.

4.3.2 Competent cells for electroporation

Cells for electroporation were prepared as described in chapter 4.2.1. Before centrifugation (4800 rpm, 4 °C, 7 min), they were cooled on ice for 15 min. To remove remaining salts, washing with 10 mL 15% glycerol was performed twice, keeping the cells cool during the whole process. After that cells were dissolved in remaining solvent, pooled and directly used.[205]

4.3.3 Transformation of competent bacteria

Transformation of plasmids in competent bacteria was performed using the heat shock method. For *in vivo* amplification *E. coli* DH10B and for protein expression *E. coli* BL21 (DE3), *E. coli* Rosetta 2 cells or *E. coli* C321.ΔA.exp cells were used (as indicated in section 4.5, protocols for protein expressions). 50 µL of competent cells were mixed with 3-5 µL plasmid and incubated on ice for 10 min. The cells were subjected to a heat shock for 2 min at 42 °C. Cells were returned on ice, 1 mL of LB was added, recovered (37 °C, 1 h, 450 rpm) and plated on agar plates containing appropriate antibiotics or used to inoculate an overnight culture.

4.3.4 Electroporation

For electroporation 600-800 ng of ethanol precipitated DNA was mixed with 100 µL competent bacteria, transferred to electroporation cuvettes, shocked (2.5 kV) and supplemented with 1 mL LB medium. Cells were recovered (1.5 h, 37 °C, 450 rpm) and subsequently used to inoculate 250 mL LB medium containing appropriate antibiotics.[205] Additionally, a dilution series was plated to test the diversity of cells.

4.4 Protein biochemical methods

4.4.1 SDS gel electrophoresis

The size and purity control of the proteins was performed using SDS gel electrophoresis. SDS is an anionic detergent that denatures proteins and negatively charges them. SDS-PAGE gels, as described by Laemmli[206], were used. The gels consist of a stacking gel with lower acrylamide concentration were the proteins are concentrated and a separating gel with acrylamide concentrations of 12% or 15%. Protein samples were mixed with 1x LB buffer for whole cell extracts or 4x LB buffer for protein solutions and heated to 92 °C for 5-10 min. Whole cell extracts were centrifuged and 10 µL loaded onto the gel. The prepared gel was placed into the

SDS gel electrophoresis chamber, running buffer added and the electrophoresis performed at 200 V for 60-120 min depending on the acrylamide concentration.

Separated proteins on the SDS gel were visualized by staining with Instant Blue. Therefore, the gel was placed into a staining bath for 30 min and afterwards rinsed with water.

4.4.2 Native PAGE analysis

Native PAGE gels were used to analyze mono-nucleosomes. The composition of two 6 % native PAGE gels is listed in Table 4-13.

Table 4-13 ***Instructions for 2 native PAGE gels***.

Ingredient	Volume
10x TB	0.5 mL
Rotiphoresis (29:1)	3 mL
ddH_2O	16.5 mL
APS	200 µL
TEMED	16 µL
Total	20 mL

6 µL sample was mixed with 3 µL glycerol or orange dye and the gels were run in 0.25x TB buffer at 80-90 V, for 150 min.

4.4.3 Western blot

Immunublotting of proteins was done after separation according to their molecular weight by SDS gel electrophoresis. The proteins were transferred onto nitrocellulose or PVDF membranes using either semi-dry or wet blotting. The gel was sandwiched between in transfer buffer soaked Whatman paper and transferred by an electric currency (semidry: 250 mA, 30-45 min, RT or wet blot: 100 V, 30 min, 4 °C) onto the membrane while maintaining the placement they occupied within the gel. Successful protein transfer was verified by staining with Ponceau S. Unspecific binding was blocked by incubating the membrane with BSA or milk for 10-30 min at RT or at 4 °C overnight. Primary and secondary antibody binding was performed as listed in Table 4-6. Primary antibodies were allowed binding for 1 h at RT or overnight at 4 °C. Before secondary antibody incubation (1 h at RT) the membranes were washed with TBS (PBS) + 0.1 % Tween 20 for 10 min and afterwards the washing step was repeated twice with Tween and once without. All secondary antibodies were HRP conjugated and detected via chemiluminescence using ECL substrate (4.1.6).

4.4.4 Stripping of membranes for reprobing

For reprobing of membranes with a different first antibody the western blot membranes were stripped following the “mild stripping protocol” (abcam). The membrane was incubated in stripping buffer (1 L: 15 g glycine, 1 g SDS, 10 mL Tween 20 and ddH_2O, pH 2.2) for 10 min twice, then washed twice with PBS for 10 min, and then again twice washed with TBS + 0.1 % Tween 20. Stripping success was checked by incubation of the membrane with a chemiluminescent detection reagent, before performing the blocking reaction, followed by different first antibody incubations.

4.4.5 Estimation of protein concentrations using Bradford reagent

Bradford reagent exists in the two different colors red and blue. Red absorbs at 470 nm whereas the blue form absorbs at 595 nm. Upon protein binding the red form converts to blue.[207] The sample was prepared by mixing 10 μL protein solution with 500 μL Bradford and the absorbance at 595 nm was measured against a blank. Furthermore, a calibration curve of absorbance versus concentration was prepared and therewith the concentration of the protein sample calculated.

4.4.6 TEV cleavage

N-terminal His_6-tags were cleaved using TEV-protease (4.5.7). Before cleavage the protein was dialyzed two times against 5 mM βME (100-fold volume). Tris-HCl (pH 7.5) was added up to 50 mM and supplemented 1:100 with 1 mg/mL TEV protease. The reaction was performed for 3 h at RT. Successful cleavage was monitored by SDS-PAGE followed by removal of salts by dialysis, as described below, and finally the protein was lyophilized.

4.4.7 Dialysis

Dialysis of proteins was performed using dialysis tubing with a MWCO of about half the size of the protein, retaining the proteins within the tubing whereas smaller molecules/ions can traverse the membrane. For exchange of buffer the dialysis was done against the chosen solvent in 100x volume excess three times.

4.4.8 Chemical modification of lysine residues on WT histone H4

Dried WT histone H4 were essentially acylated as described by Gracia et al.[155]. Therefore the protein was dissolved to ~1 mg/mL in 100 mM ammonium bicarbonate. 25 µg of histone was aliquoted twice for modification with either propionyl, butyryl or crotonyl respectively. Ammonium hydroxide (2 µL) was added and 100 µL of propionylation, butyrylation or crotonylation reagent (75 µL propionic/butyric/crotonic anhydride plus 25 µL MeOH, vortexed and spun down shortly), vortexed and centrifuged shortly before verification that pH is >8. If pH was too low ammonium hydroxide was added until pH >8 was achieved. The reaction was incubated for 20 min at 51 °C and afterwards concentrated in a *SpeedVac* to ~5 µL. For completion of the reaction this procedure was performed twice. After the second round of modification the samples were dried to completion.

4.5 Expression and Purification of proteins

4.5.1 Expression and purification of recombinant WT histone proteins by ion exchange chromatography

Histone proteins H2A, H2B, H3 and H4 overexpression was performed using *E. coli* Rosetta 2 cells with Cm resistance according to published protocols.[56,208]

The overexpressed histone proteins were purified by the preparation of inclusion bodies. For that the cell pellet was thawed, washed with 30 mL PBS per 2 L pellet, centrifuged (RT, 20 min, 4800 rpm) and resuspended in another 30 mL PBS supplemented with 1 mM DTT. Cell lysis was performed using the Fluidizer and afterwards the IBs were pelleted (18000 rpm, 4 °C, 20 min). IBs were resuspended in PBS containing 1 % Triton X and 1 mM DTT and centrifuged as indicated above, washed again with PBS and eventually frozen overnight. The sample was soaked in 1 mL DMSO for 30 min at RT with shaking. Slowly 20 mL of guanidinium solution were added to extract histone proteins from the pellet, incubated for 1 h at 37 °C with shaking and centrifuged as above. The supernatant was separated and dialyzed against urea dialysis buffer overnight. The precipitations were removed by centrifugation and the histones purified via ion exchange chromatography using an ÄKTA prime.

Therefore, a positively charged Q column and a negatively charged SP column were set up in a tandem in which the anion trap Q column was passed first and the SP cation trap column, where the desired protein binds, was passed subsequently. The columns were equilibrated with selected salt concentrations for each histone. Then the samples were loaded on the columns, the Q column was removed and proteins eluted by increasing the salt concentration following the published protocol by Tsukiyama[208] using a flow rate of 2 mL/min. Proteins were detected by 15 % SDS gel electrophoresis, pure fractions were pooled and dialyzed against 5 mM βME, before lyophilizing and storage at -20 °C until use.

Table 4-14 ***Buffers for purification of histones by ion exchange chromatography*****.**

Urea dialysis buffer	**Urea buffer A**	**Urea buffer B**
7 M Urea	7 M Urea	7 M Urea
10 mM Tris-HCl pH 8.0	10 mM Tris-HCl pH 8.0	10 mM Tris-HCl pH 8.0
1 mM EDTA pH 8.0	1 mM EDTA pH 8.0	1 mM EDTA pH 8.0
0.1 M NaCl	1 mM DTT	1 M NaCl
5 mM βME	0.2 mM PMSF	1 mM DTT
0.2 mM PMSF		0.2 mM PMSF

Urea was incubated for 1 h with anion/cation exchange resin and filtered before adding further reagents.

Table 4-15 ***Salt gradients for elution of histones from SP column as published by Tsukiyama*****[208].**

H2A, H2B		H3		H4	
Vol (mL)	%B	Vol (mL)	%B	Vol (mL)	%B
0	10	0	20	0	20
10	10	10	20	10	20
135	35	20	25	135	45
140	40	145	50	155	60
140.1	100	155	60	155.1	100
190	100	155.1	100	205	100
		205	100		

4.5.2 Purification of his$_6$-tagged proteins via Ni^{2+}-batch

Proteins containing a his$_6$-tag were purified via Ni^{2+}-NTA beads. First, the cells were disrupted by fluidizer or sonication. Cell debris was removed via centrifugation and the soluble protein allowed binding to the Ni^{2+}-NTA beads for 1 h at 4 °C with gentle agitation. The slurry was transferred onto a Ni^{2+}-NTA column and washed with buffer, as indicated for the individual protein. Proteins were then eluted by supplementing the indicated buffer with 200 mM imidazole and 1 mL fractions were collected. Finally fractions were analyzed by SDS-PAGE.

4.5.3 Production and purification of recombinant myoglobin

For expression of acylated sperm whale myoglobin *E. coli* DH10B cells were transformed with p*Myo*4TAGPylT-*his$_6$*[17] and a plasmid containing the synthetase for incorporation of the respective acyl modification as indicated in Table 4-16. The expression was performed using LB broth supplemented with Kan and Tet. For the expression LB medium supplemented with antibiotics and 20 mM nicotinamide were inoculated from an overnight culture. The acyl amino acid was subsequently added (Table 4-16) and cells grown at 37 °C with shaking. At OD_{600}~0.8 protein expression was induced by addition of 0.2% arabinose and the cells were harvested 4 h after induction. The pellets were resuspended in buffer A (pH 8.0: 200 mM NaCl, 20 mM Imidazole, 10 mM Tris-HCl, 20 mM NAM) supplemented with 1x protease inhibitor cocktail and 0.2 mg/mL lysozyme, incubated 20 min on ice and cracked via Sonication. The lysate was clarified via centrifugation (15 min, 18000 rpm, Beckmann JA-20, 4 °C), Ni^{2+}-NTA beads added and incubated with agitation for 1 h at 4 °C. The beads were transferred onto a Ni^{2+}-NTA column, washed with buffer A and eluted by buffer A supplemented with 200 mM imidazole. Finally, proteins were analyzed via 15% SDS-PAGE.

Table 4-16 ***Plasmids for incorporation of acyl modification and concentrations of acyl amino acids.***

Plasmid	Acyl amino acid being incorporated	Concentration used for protein expression
pBK-PylS WT	BuK and CrK	2 mM
pBK-AcKRS3[28]	AcK	10 mM
pBK-PrKRS	PrK	2 mM

4.5.4 Expression and purification of recombinant acyl-modified histone H4 proteins

Acylated H4 K16 fusion proteins were produced in *E. coli* C321.ΔA.exp cells, transformed with pCDF-PylT-AraP-His$_6$-H3(Δ93-98)-TEV-H4 K16TAG and additionally with one of the plasmids indicated in Table 4-16. Expressions were carried out in standard LB broth supplemented with Kan, Spec and 20 mM NAM. Cells were inoculated from an overnight culture and allowed to grow at 37°C with shaking. After approximately 2 h the acyl amino acid (Table 4-16) was added and cells further grown until an OD$_{600}$ ~ 0.8 before induction with 0.2% arabinose. Protein expression was carried out at 37 °C with shaking for 16 h.

Cells were collected and lysed in 30 ml PBS supplemented with 20 mM NAM, 1 mM DTT, 1 mM PMSF, 1 x PIC protease cocktail, and 0.2 mg/ml lysozyme. The cells were incubated at 37 °C for 20 min, with shaking, and then lysed via high-pressure homogenization, on ice. The lysates were clarified by centrifugation (15 min, 18000 rpm, Beckmann JA-20) and then washed in PBS supplemented with 20 mM NAM, 1 mM DTT and 1% TritonX-100. The inclusion bodies were collected as described above and then washed in PBS with 20 mM NAM, 1 mM DTT. The insoluble fraction was pelleted again and then macerated in 1 mL DMSO and incubated at room temperature for 30 min. Proteins were extracted from inclusion bodies by adding 30 ml of guanidinium buffer (6 M guanidinium chloride, 20 mM Tris (pH 8.0), 2 mM DTT) and then incubated at 37 °C for 1 hr, with shaking, centrifuged as described above and purified by his$_6$-tag as stated in 4.5.2. The washing buffer consisted of: 8 M urea, 100 mM NaH_2PO_4, and 1 mM DTT (pH 6.2) and the elution buffer of: 7 M urea, 20 mM sodium acetate, 200 mM NaCl, and 1 mM DTT (pH 4.5). The eluates were dialyzed against 5 mM βME and TEV cleaved (4.4.6). When the reaction was completed, salts were removed by dialyzing back into 5 mM βME. The concentration of H4 proteins was determined by a titration against an H4 standard via SDS-PAGE electrophoresis, and the samples then lyophilized.

4.5.5 Expression and purification of recombinant histones containing an UAA

For Histone H3 overexpressions *E. coli* BL21 (DE3) cells were transformed with a pCDF PylT plasmid carrying the ORF for the histone with amber codons at the desired positions and additionally one pBK plasmid decoding the synthetase suitable for incorporation of the

UAA of interest (UAA-plasmid combinations and UAA concentrations are described in Table 4-1). Expression and purification was performed as described in 4.5.4 with some changes. Selective antibiotics were Kan and Spec. NAM was used only for acyl amino acids and for purification of histones containing AzF no DTT was added to the buffers.

4.5.6 Expression and purification of GFP

E. coli DH10B cells were transformed with either pBAD sfGFP or pBAD sfGFP 150TAG and a plasmid for incorporation of an UAA of interest (Table 4-1). 0.5 L standard LB broth was inoculated from the overnight culture and supplemented with Amp or Amp plus selective antibiotics for the synthetase and the UAA. The cells were grown to OD ~0.8 and protein expression induced with 0.2 % arabinose. After 3 h cells were harvested, cracked (20 mM Tris-HCl (pH 7.2) 250 mM NaCl, 10 mM imidazole supplemented with 1x PIC and 0.2 mM PMSF) and purified by his$_6$-tag (4.5.2). The indicated buffer was used for washing and for elution was supplemented with 200 mM imidazole. Collected fractions were analyzed by 12 % SDS-PAGE.

4.5.7 Production of TEV protease

E. coli Rosetta 2 (DE3) cells were transformed with pET24 TEV and grown on an Amp plate. A single colony was picked and used for inoculation of an overnight culture. 2 L LB broth medium with Amp were inoculate from the overnight culture, grown to OD ~1.5 and expression induced by addition of 0.5 mM IPTG, protein expression was run for 3 h and then the cells were collected. The purification was carried out in a buffer containing 50 mM NaH_2PO_4/Na_2HPO_4 (pH 8.0) and 100 mM NaCl. For lysis via fluidizer, the buffer was supplemented with 1 mM DTT, 1x PIC and 0.2 mM PMSF. Afterwards the lysate was clarified by centrifugation (18k rpm, 30 min, 4 °C, JA-20) and the supernatant containing the soluble protein solution purified as described in 4.5.2. Using the described buffer supplemented with 1 mM DTT. For protein elution the buffer was supplemented with 300 mM imidazole. Fractions were analyzed by SDS-PAGE. Fractions containing protein were pooled, the concentration determined by Bradford and a maximum of 35 mg protein per run was subjected to a further purification by size-exclusion chromatography using a S75 column and the above indicated buffer but containing 1 mM DTT. Pure protein fractions were pooled, concentrated and stored at –80 °C in 50% glycerol.

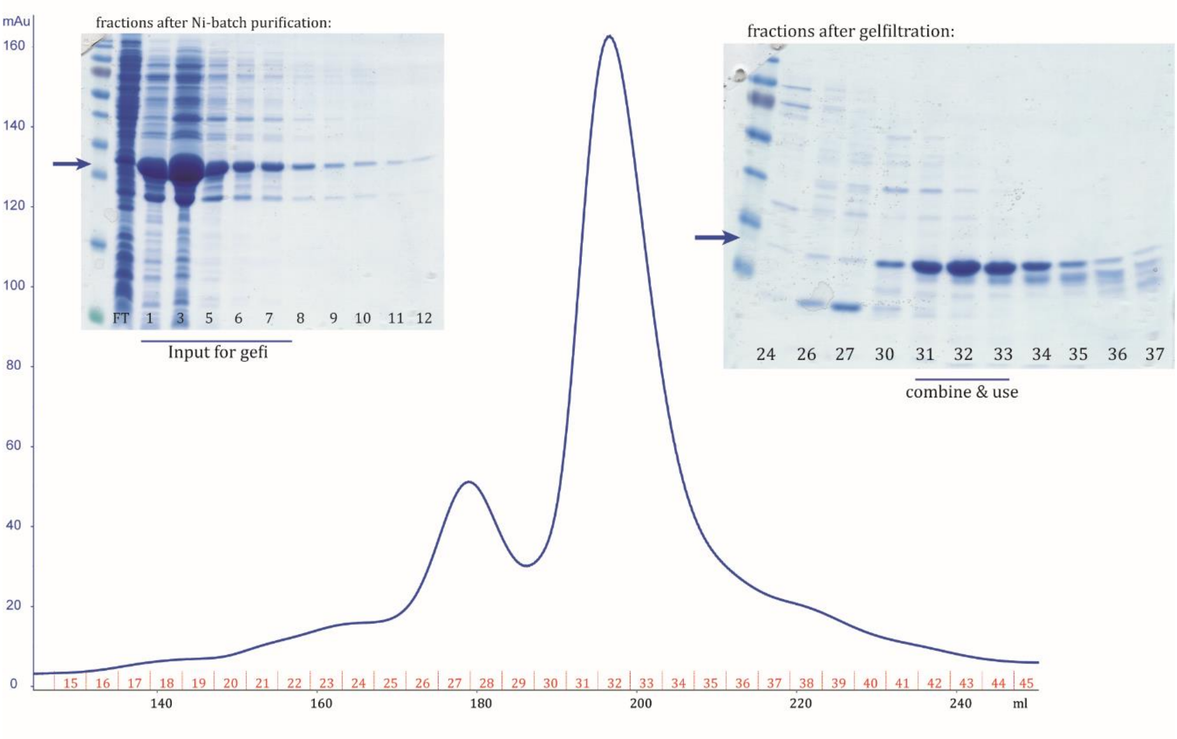


Figure 57 ***Purification of TEV protease.*** Left small picture fractions from his$_6$-purification. Fractions that were used for further purification are indicated. The size-exclusion chromatography was performed on a S75 column and the resulting chromatogram is depicted. Fractions were analyzed by 15% SDS-PAGE (small picture right) and stained by Instant Blue.

4.5.8 Expression and purification of CobB

The pCDF Duet-1 CobB plasmid was transformed in *E. coli* BL21 (DE3) cells. One single colony was used to inoculate an overnight culture containing Spec. This culture was used to inoculate 1 L LB broth containing Spec, grown to OD ~1.0 and protein expression induced by 0.5 mM IPTG. The expression was run for ~4 h before cells were collected. The protein was purified by Ni^{2+}-batch (4.5.1) using Ni-Wash buffer (4.1.8). For analysis of the eluted fractions 12% SDS-PAGE was used.

4.5.9 Expression and purification of Hst2

E. coli BL21 (DE3) cells were transformed with the plasmid pCDF-Duet Hst2 and grown on an agarose plate containing Spec. A single colony was picked and used for inoculation of an overnight culture. For the expression 2 L of standard LB broth containing Spec and 1 mM $ZnCl_2$ were inoculated from the overnight culture and grown at 37 °C until OD ~1. Then expression was induced by 0.5 mM IPTG and the temperature decreased to 16 °C. After production of Hst2 overnight, cells were collected, cracked via fluidizer in 75 mM HEPES, 200 mM NaCl, 10 mM DTT buffer supplemented with 0.2 mM PMSF and 1x PIC. The following steps were all performed at 4 °C. The lysate was clarified by centrifugation (18 000 rpm, 4 °C, 20 min, JA-20) and the soluble protein fraction separated and purified as described in 4.5.2. The protein was eluted in HEPES buffer (as indicated) containing 200 mM imidazole. The fractions were analyzed by 15 % SDS-PAGE, fractions containing Hst2 in good amounts were pooled and further purified by gel filtration chromatography using a S75 column in 20 mM HEPES, 100 mM NaCl and 10 mM DTT.

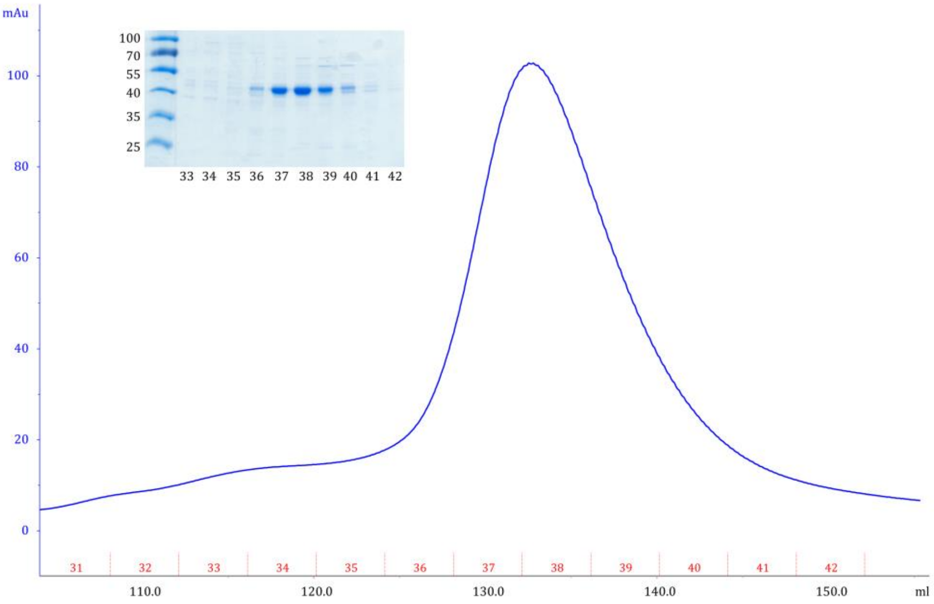


Figure 58 ***Hst2 purification by size-exclusion chromatography using a S75 column.*** X-axis volume in mL, y-axis absorption at 280 nm. Fractions are depicted in red. The small picture represents the analysis of the fractions by 12% SDS-PAGE and subsequent Instant Blue staining.

4.5.10 Expression and purification of GST-MBP

For GST-MBP purifications, overnight cultures of *E. coli* DH10B cells were transformed with the appropriate plasmid combinations[171] and used to inoculate 50 mL LB medium to OD_{600} of 0.1. The UAA combinations (BocK 1 mM or AzF 5 mM and PrgK 2 mM) were added and protein expression was induced after 3 h of incubation by the addition of 1 mM IPTG for P_{lac} constructs. After 4 hours, cells were pelleted and proteins purified by batch using GSH-beads (GE Healthcare) following published procedures.[54] Proteins were eluted from beads by PBS buffer containing 30 mM Glutathion (AppliChem).

4.6 Recombinant chromatin

4.6.1 Histone octamer refolding

Histone octamers were reconstituted from recombinant WT histones from *Xenopus laevis*. The lyophilized recombinant proteins were thawed and diluted in unfolding buffer to approximately 2 mg/mL. Unfolding was allowed for 1 h and the concentrations were determined by measuring the absorption at OD_{276} against unfolding buffer. By reconstitutions with histones that contain any UAA, the appropriate histone replaced the respective WT histone.

Table 4-17 ***Molecular weight of the different histone proteins and their absorption at 276 nm***.

Histone	MW [Da]	ε (cm^{-1} M^{-1}), 276 nm
H2A	13960	4050
H2B	13774	6070
H3	15273	4040
H4	11236	5040

The four histone proteins were mixed together in equimolar ratios and adjusted to a final concentration of 1 mg/mL. Refolding was proceeded by dialysis against three times refolding buffer at 4 °C. The reaction mixture was concentrated to 0.5-1 mL (4800 rpm, 4 °C), filtered using a spin-X column and purified by using a preequilibrated gel filtration S200 column.

Table 4-18 ***Buffer composition for Unfolding- and Refolding buffer.***

Unfolding buffer	**Refolding buffer**
7 M Guanidinium hydrochloride	2 M NaCl
20 mM Tris-HCl pH 7.5	10 mM Tris-HCl pH 7.5
10 mM DTT	1 mM Na-EDTA
	5 mM βME

The obtained octamers were purified by size-exclusion chromatography using an ÄKTA purifier. The samples were kept at 4 °C during the whole refolding and purification process. A Superdex column (4.1.7) was equilibrated with refolding buffer, the concentrated sample (about 0.7 mL) was injected onto the column and separation was achieved at a constant flow rate of 0.5 mL/min. Fractions of 0.5 mL were collected. Success of refolding was analyzed by 15% SDS-PAGE.

4.6.2 Reconstitution of mononucleosomes via salt dialysis

Mononucleosomes were assembled using 187 bp DNA which contained the 147 bp '601' positioning sequence[165] (generous gift by the group of Wolfgang Fischle, MPI Göttingen). Histone octamers were titrated against DNA in molar ratios of 0.9, 1.0, 1.1, 1.2 and 1.5: 1.0, histone to DNA. 5 M NaCl was mixed with the DNA and RB high buffer (Table 4-19) and octamer was added. The samples were transferred into Slide-A-Lyzer Mini Dialysis Devices (7K MWCO) and floated in 400 mL RB high. Then the buffer was exchanged to RB low (Table 4-19) by continuous dialysis with the pump speed adjusted to approximately 1.3 mL/min over 24 h (against at least 1.6 L RB low).[56] The whole procedure was carried out at 4 °C. Mononucleosomes were subsequently analyzed by native PAGE (4.4.2).

Table 4-19 ***Buffers for nucleosome assembly.***

RB high	**RB low**
2 M NaCl	10 mM NaCl
10 mM Tris-HCl (pH 7.5)	10 mM Tris-HCl (pH 7.5)
1 mM EDTA	1 mM EDTA
1 mM DTT	1 mM DTT

4.6.3 Reconstitution of nucleosome arrays via salt dialysis

Assembly of nucleosomal arrays was done using the Widom 601-200-12 DNA template[165] by titration of octamers against template in molar ratios of 0.9, 1.0, 1.1, 1.2 and 1.5:1.0 histone to template. Scavenger DNA, 147 bp, was included in the assembly reaction with equal concentrations as the template. 5 M NaCl was mixed with the DNA and RB hight buffer (Table 4-19) and finally octamer was added. The dialysis was carried out in Slide-A-Lyzer Mini Dialysis Devices (7K MWCO) and the buffer was exchanged to RB low (Table 4-19) by continuous dialysis with a pump speed of approximately 1.3 mL/min over 24 h at 4 °C. As described by Luger et al.[56]

4.6.4 Analysis of nucleosome arrays

Reconstituted nucleosome arrays were run on a 0.5% agarose gel in 0.2x TB buffer for analysis of quality. 100 µg of DNA template was loaded and 50-100 µg of reconstituted nucleosomes in refolding buffer mixed with 10% glycerol. The gel was run at 100 V for 1 h.

4.6.4.1 MNase digest

For the MNase digest a reaction volume of 500 µL was prepared containing 1000 ng nucleosome array in freshly prepared MNase digest buffer (50 mM Tris-HCl pH 8.0 containg 1 mM $CaCl_2$). The reaction mixture was equilibrated to RT and 100 µL sample at time point “0” was taken, mixed with an equal volume of binding buffer from a gel purification kit and stored on ice. The reaction mixture was equilibrated to 37 °C and 0.04 U/mL MNase was added. Samples were taken after 30 sec, 2 min and 15 min. After each time point the reaction was quenched by addition of an equal volume of binding buffer and stored on ice. The samples were purified according to the manufacturer´s instructions from gel extraction kit, eluted in 20 µL elution buffer and digest evaluated by a 1.5% agarose gel.

4.6.4.2 AvaI digest

Digest of nucleosome arrays to mono-nucleosomes was done in a total of 20 µL reaction volume. FD green buffer (Thermo Scientific) was mixed with water, then 500 ng array was added and finally 1 U AvaI fast digest enzyme. The reaction was incubated at 37 °C for 1.5 h, samples loaded on a 2% agarose gel (including GelRed) and run 1 h at 120 V.

4.6.5 Nucleosome array self-association assay

Before performing nucleosome self-precipitation assays, the concentration of the arrays was determined by measuring A_{260} with (A_{260} = 0.45 =1 mg/mL). The arrays were diluted to reach a starting value of ~0.4 in 10 mM Tris-HCl (pH 7.5) and 10 mM NaCl. The samples were mixed with an equal volume of 2x $MgCl_2$ solution in the indicated buffer covering Mg^{2+} concentrations between 0-8 mM. The samples were incubated at 25 °C for 20 min and precipitations removed by centrifugation (15 min, full speed). The supernatant of each sample was measured for A260 and plottet against the corresponding Mg^{2+} concentration and normalized to a sample containing no Mg^{2+}.

4.6.6 Assay lysine deacylation with CobB

The H4 K16acyl proteins were dissolved to 30 ng/µL in reaction buffer (500 mM Tris-HCl (pH7.5), 140 mM NaCl, 20 mM $Mg(OAc)_2$, 80 µM NAD^+). The deacylation reaction was started by addition of 2 ng/µL CobB and performed at RT. Samples were taken after 0, 1, 2, 3, 4, 5, 10, 15, 20 and 30 min, 7.5 µL per time point respectively. The sample was quenched by addition of 2.5 µL preheated 4x SB buffer supplemented with 48 mM NAM and kept at 95°C for 5 min then stored on ice. Samples of the different time points were separated by 15% SDS-PAGE, transferred onto PVDF membranes (4.4.3) and acylation signals detected by either α-H4K16ac (acetyl, propionyl, butyrylation) or α-CrK (crotonylation). Membranes were stripped (4.4.4) and subsequently analyzed for H4 signals. Antibody details in 4.1.15.

4.7 Pyrrolysine mutational library

4.7.1 Library creation

The primers used for the first round were PylSCMf and PylSCMr and for the second round PylSAYLf and PylSAYLr. Sequences are provided in chapter 4.1.16. The enzymatic conditions are described in 4.2.7 and the temperature program for the inverse PCR is outlined in Table 4-20.

Table 4-20 ***Temperature program for inverse PCR reaction.***

Step	Temperature	Time	
1. Initial denaturing	95 °C	1 min	
2. Denaturing	95 °C	20 sec	5 cycles
3. Annealing	65 °C	30 sec	
4. Elongation	68 °C	4 min	
5. Denaturing	95 °C	20 sec	25 cycles
6. Annealing	60 °C	30 sec	
7. Elongation	68 °C	4 min	
8. Final elongation	68 °C	5 min	

PCR products were purified using a QIAquick PCR purification kit and the DNA was eluted in 100 µL EB buffer. Parental template DNA digestions were performed using 1.25 µL DpnI at 37 °C for 1 h followed by a digest with 3 µL Bsa1 at 50 °C for 3 h. Purification was performed as described above and sticky ends ligated according to the conditions in 4.2.7.

Table 4-21 ***Enzymatic set up for ligation reaction.***

	Volume
Digest	1000 µL
ddH_2O	700 µL
Ligase buffer	200 µL
T4 DNA Ligase	100 µL

Ligation reactions were performed overnight at 16 °C. The resulting circularized DNA was ethanol precipitated as described in 4.2.10. DNA was pooled and transformed into freshly prepared electrocompetent *E. coli* DH10B cells (4.3.2). Twelve randomly chosen clones were sequenced to show the diversity of the library.

4.7.2 Library selection

Freshly prepared *E. coli* DH10B pREP-PylT cells were transformed with the mutational library. About 600 ng of pBK-PylS library DNA was transformed per 100 μL cells and grown overnight (37 °C, 210 rpm) in 100 mL LB media supplemented with a final concentration of 25 μg/mL Kan and 12.5 μg/mL Tet (LB-KT). A dilution series for the transformation revealed an efficiency of greater than 10^8 transformants. The overnight culture was diluted 1:50 into fresh LB-KT supplemented with 1 mM UAA and incubated (3-4 h, 37 °C, 210 rpm). This culture was plated twice on LB-KT agar plates containing 1 mM UAA and 50 μg/mL Cm. The plates were incubated (48 h, 37 °C) and the resulting cell colonies were scraped for DNA isolation. The library plasmid DNA was separated from the pREP-PylT plasmid DNA by agarose gel electrophoresis, cut out of the gel and extracted for purification (see 4.2.3 and 4.2.4).

About 600 ng of the purified DNA was transformed twice via electroporation into freshly prepared *E. coli* DH10B pBar-PylT cells (100 μL) for the negative selection round. The cells were recovered in LB media (2 h, 37 °C) with gentle shaking, concentrated (RT, 13300 rpm, 30 sec) and then plated (200 μL) onto two LB agar plates containing 25 μg/mL Kan, 25 μg/mL Cm and 0.2 % arabinose. After incubation for 24 h at 37 °C the colonies were harvested by stripping the plates as described above. The plasmids were isolated, separated by agarose gel electrophoresis and the PylS Lib DNA was extracted from the gel.

For the third round of the selection the pBK-PylS Lib plasmid was again transformed into *E. coli* DH10B pREP-PylT cells, grown and plated as describe for the first round. But unlike the previous round the cells were not collected from the plates, instead individual colonies were picked and inoculated in 1 mL of LB-KT. The cultures were grown overnight and 200 μL were diluted 1:10 into LB-KT and then split into two identical 1 mL cultures. One of these was supplemented with 1 mM UAA the other one was not. Both cultures were incubated for 5 h at 37 °C with 210 rpm. Meanwhile LB-KT agar plates with or without 1 mM UAA were prepared additionally supplemented with increasing concentrations of Cm. One drop of each culture was placed onto the appropriate agar plate with or without UAA and incubated overnight at 37 °C. The clones that exhibited a strong UAA dependence towards Cm resistance were isolated.

The reporter plasmid was separated from the library by digestion with SalI enzyme that digests only the pREP-PylT plasmid but not within pBK-PylS Lib. The isolated plasmids were replated to ensure that the observed Cm resistance did not result from mutations within the genome or the reporter plasmid. For that purpose *E. coli* DH10B pREP-PylT cells were transformed with the selected pBK-PylS plasmids and the growth test on Cm repeated.

4.8 Analysis of proteins by mass-spectrometry

4.8.1 Analysis of total proteins by ESI-MS

4.8.1.1 Sample preparation

For total mass analysis of myoglobin carrying the acyl modifications, the purified proteins were desalted and concentrated via ZipTip C_{18} (Millipore) according to manufacturer´s protocol and eluted in 50% MeOH, 3% FA.

4.8.1.2 High-resolution mass spectrometry

High-resolution mass spectrometry was performed by Holm-Frauendorf (Institute for Organic Chemistry, University of Göttingen), using a 7 T-FTICR-MS instrument (APEX IV, Bruker Daltonics, Billerica, USA) equipped with an APOLLO electrospray ion source and a syringe pump (74900 series, Cole-Parmer, Vernon Hills, USA) with a flow rate of 2 µL min^{-1} for sample injection. For desolvation the dry gas temperature was set to 250 °C and the capillary voltage to 4.2 kV with a capillary exit voltage of 100 V. The ions were accumulated in the hexapole for 0.5 s, transferred subsequently into the ICR cell and detected within a mass range of 250-2000 m/z.

4.8.2 Protein peptide analysis by MS/MS

4.8.2.1 In gel butyrylation of H4 K16ac and tryptic digest for analysis via nanoLC-nanoESI-MS/MS2

The procedure was performed according to Gattner et al.[175]. Purified H4 K16ac protein was run on a SDS-PAGE, stained, bands excised from gel and cut into small pieces. Then gel pieces were prepared for capping by washing with 100 µL ddH_2O and incubating 15 min at RT. Subsequently 50 mM NH_4HCO_3/acetonitrile (ACN) (each 100 µL) was added and incubated again for 15 min. All liquids were removed and gel pieces covered with 100 µL ACN. The ACN was removed and slices rehydrated with NH_4HCO_3/ACN (100 µL each) for 5 min, further 100 µL of ACN were added and incubated for 15 min. All liquids were removed and slices shrinked by 100 µL ACN. After 5 min all acetonitrile was removed and gel slices dried at RT for 10 min. Then the capping of unmodified lysines was done to create longer peptides for analysis by mass as trypsine only recognizes arginine residues afterwards. Therefore the dried gel pieces were incubated with 20 µL butyryl-anhydride and 40 µL 2 M $NH_3(H_2O)$ for 30 min at 51 °C (twice). Afterwards, samples were washed by addition of NH_4HCO_3/ACN (100 µL each), for 1 h at RT and again for 15 min before adding 100 µL ddH_2O and removal of all solvents. Again 100 µL ddH_2O were added and incubated for 10 min, then 100 µL of ACN for 10 min and dried. For trypsin digest 50 µL of 25 mM NH_4HCO_3 containing 5 ng/µL trypsine were added and digestion performed overnight at 37 °C. The supernatant was removed for separate storage and the capping procedure repeated to cap the originated peptide-ends. Therefore, the gel pieces were incubated twice with 20 µL butyryl-anhydride and 40 µL 2M $NH_3(H_2O)$ at 51 °C, 30 min. After removal of solvents, washing with 100 µL ddH_2O twice, the peptides were extracted from the gel. First, gel pieces were dissolved in 100 µL 25 mM NH_4HCO_3 and sonicated for 15 min, 100 µL of ACN were added and another 15 min sonicated. The supernatant was removed and another round of sonication was performed with the gel pieces (15 min with 100 µL 5% formic acid then 15 min with 100 µL ACN). In the end all peptide solutions were combined, concentrated and used for nanoLC-nanoESI-MS/MS2 analysis.

4.8.2.2 Mass spectrometry analysis with nanoLC-nanoESI-MS/MS2

The ESI-MS/MS2 analysis was done by Oliver Valerius, Institute for Microbiology, University of Göttingen.

4.8.2.2.1 Nano LC -*RSLCnano Ultimate 3000* system (Thermo Scientific)

Peptides of 1-5 µl sample solution were trapped and washed with 0.05 trifluoroacetic acid 2.6% acetonitrile on an *Acclaim® PepMap 100* column (75 µm x 2 cm, C18, 3 µm, 100 Å, P/N164535 Thermo Scientific) at a flow rate of 25 µl/min for 5 min. Analytical peptide separation by reverse phase chromatography was performed on an *Acclaim® PepMap RSLC* column (75 µm x 25 cm, C18, 3 µm, 100 Å, P/N164534 Thermo Scientific) running a 40 minutes gradient from 100% solvent A (0.1% formic acid) to 65% solvent B (80% acetonitrile, 0.1% formic acid) and further to 95% solvent B within 1 min at flow rates of 300 nl/min (Fisher Chemicals).

4.8.2.2.2 Nano ESI mass spectrometry - *OrbitrapVelos Pro* (Thermo Scientific)

Chromatographically eluting peptides were on-line ionized by nano-electrospray (nESI) using the *Nanospray Flex Ion Source* (Thermo Scientific) at 2.4 kV and continuously transferred into the mass spectrometer. Full scans within m/z of 300-1850 were recorded with the Orbitrap-FT analyzer at a resolution of 30.000 with parallel data-dependent top 10 MS2-fragmentation in the *LTQ Veleo Pro* linear ion trap. LCMS method programming and data acquisition was performed with the software *Xcalibur 2.2* (Thermo Fisher). MS/MS2 data processing for protein analysis and PTM identification was done with the *Proteome Discoverer 1.4* (*PD*, Thermo Scientific) and the *Discoverer DeamonDaemon 1.4* (Thermo Scientific) software using the SequestHT search engine (Thermo Scientific) and organism-specific taxon-defined protein databases extended by the most common contaminants. Acetylation, butyrylation, crotonylation and propionylation at lysine residues were considered as dynamical modifications for the respective samples. False discovery rates of considered peptide sequence matches were in maximum 0.01.

4.9 Protein labeling

4.9.1 Cu(I) catalyzed click reaction

The alkyne moiety of PrgK was modified in a Cu(I) catalyzed click reaction as described by Nguyen et al.[132]. The protein was dissolved in 50 mM sodium phosphate buffer (pH 8.3) to a concentration of 50 mM. Stock solutions of 100 mM $CuSO_4$ in water, 100 mM ascorbate in water and 50 mM *o*-phenanatroline in DMSO were prepared. 1 eq. ascorbic acide solution was mixed with 1 eq. $CuSO_4$ solution and 2 eq. of *o*-phenanatroline were added quickly, incubated at RT for 5 min to show a precipitation. The protein containing the PrgK was reacted with 25 eq. moles of Alexa Fluor 647 azide, triethylammonium salt (Invitrogen) and 1 mM $CuSO_4$, 1 mM ascorbic acid and 2 mM *o*-phenanatroline for 15 min at RT. Afterwards, the reaction was continued at 4 °C overnight. The samples were loaded on a SDS gel and fluorescence visualized using a fluorescence scanner and eventually stained with Instant Blue.

4.9.2 Copper free click reaction

The azido functionality of AzF was modified in a copper free click reaction with 50 eq. moles of DBCO488 (Jena Bioscience) in 50 mM sodium phosphate buffer pH 8.3 for 1 h at RT. Samples were mixed with 10% glycerol and analyzed by 15% SDS-PAGE and fluorescence bands were detected by Typhoon 9400 Variable Mode Imager (Amersham) and subsequently Instant Blue staining.

4.9.3 Labeling cysteine residues with maleimide dyes

The thiol function of a cysteine residue was reacted with a maleimide dye. The histone protein was first rebuffered using a Nap10 column. The column was equilibrated with PBS containing 1 mM TCEP. Then about 500 µg of protein solution was applied to the column and eluted in five fractions of 500 µL. Fractions containing protein (usually fractions 2+3) were pooled and the amount was estimated using Bradford Reagent (as described in 4.4.5). The dye was dissolved to 10 mg/mL in DMSO and the protein was mixed with dye in equimolar ratios. The reaction was allowed to perform overnight at 4 °C protected from light.

4.9.4 Methanol-chloroform precipitation to remove free dye

The precipitation procedure was performed at RT in an eppendorf cup.[166] The sample in aquous solution (150 µL) was mixed with 4 volumes of MeOH and 1 volume of chloroform and vortexed resulting in only one phase. Then 3 volumes of ddH_2O were added and mixed thoroughly following centrifugation (1 min, full speed). The upper organic phase was removed without disturbing the interface that contains the proteins and collected separately. Another 3 volumes of MeOH were added and vortexed again followed by centrifugation (1-2 min, full speed). The supernatant was removed carefully and the pellet allowed to air-dry before taken up in adequate buffer.

5 References

(1) Nelson, D. L., and Cox, M. M. (2009) Lehninger Biochemie. Springer Berlin Heidelberg, Berlin, Heidelberg.

(2) Horton, H. R., Moran, L. A., Scrimgeour, K. G., Perry, M. D., and Rawn, J. D. (2008) Biochemie. *J. Am. Chem. Soc.* 4th Editio. Pearson Studium, München.

(3) Böck, A., Forchhammer, K., Heider, J., Leinfelder, W., Sawers, G., Veprek, B., and Zinoni, F. (1991) Selenocysteine: the 21st amino acid. *Mol. Microbiol. 5*, 515–20.

(4) Srinivasan, G., James, C. M., and Krzycki, J. A. (2002) Pyrrolysine encoded by UAG in Archaea: charging of a UAG-decoding specialized tRNA. *Science 296*, 1459–62.

(5) Forchhammer, K., and Böck, A. (1991) Selenocysteine synthase from Escherichia coli. Analysis of the reaction sequence. *J. Biol. Chem. 266*, 6324–8.

(6) Wang, L., and Schultz, P. G. (2002) Expanding the genetic code. *Chem. Commun.* 1–11.

(7) Pastrnak, M., Magliery, T. J., and Schultz, P. G. (2000) A New Orthogonal Suppressor tRNA/Aminoacyl-tRNA Synthetase Pair for Evolving an Organism with an Expanded Genetic Code. *Helv. Chim. Acta 83*, 2277–2286.

(8) Liu, D. R., and Schultz, P. G. (1999) Progress toward the evolution of an organism with an expanded genetic code. *Proc. Natl. Acad. Sci. U. S. A. 96*, 4780–5.

(9) Furter, R. (1998) Expansion of the genetic code: site-directed p-fluoro-phenylalanine incorporation in Escherichia coli. *Protein Sci. 7*, 419–26.

(10) Wang, L., Brock, A., Herberich, B., and Schultz, P. G. (2001) Expanding the genetic code of Escherichia coli. *Science 292*, 498–500.

(11) Xie, J., and Schultz, P. G. (2005) Adding amino acids to the genetic repertoire. *Curr. Opin. Chem. Biol. 9*, 548–54.

(12) Liu, C. C., and Schultz, P. G. (2010) Adding new chemistries to the genetic code. *Annu. Rev. Biochem. 79*, 413–44.

(13) Wang, J., Xie, J., and Schultz, P. G. (2006) A genetically encoded fluorescent amino acid. *J. Am. Chem. Soc. 128*, 8738–9.

(14) Wang, L., and Schultz, P. G. (2005) Die Erweiterung des genetischen Codes. *Angew. Chemie 117*, 34–68.

(15) Schreiber, G., Frisch, C., and Fersht, A. R. (1997) The role of Glu73 of barnase in catalysis and the binding of barstar. *J. Mol. Biol. 270*, 111–22.

(16) Hartley, R. W. (1988) Barnase and barstar. Expression of its cloned inhibitor permits expression of a cloned ribonuclease. *J. Mol. Biol. 202*, 913–5.

(17) Neumann, H., Peak-Chew, S. Y., and Chin, J. W. (2008) Genetically encoding N(epsilon)-acetyllysine in recombinant proteins. *Nat. Chem. Biol. 4*, 232–4.

(18) Chin, J. W. (2014) Expanding and reprogramming the genetic code of cells and animals. *Annu. Rev. Biochem. 83*, 379–408.

(19) Hancock, S. M., Uprety, R., Deiters, A., and Chin, J. W. (2010) Expanding the genetic code of yeast for incorporation of diverse unnatural amino acids via a pyrrolysyl-tRNA synthetase/tRNA pair. *J. Am. Chem. Soc. 132*, 14819–24.

(20) Chin, J. W., Cropp, T. A., Anderson, J. C., Mukherji, M., Zhang, Z., and Schultz, P. G. (2003) An expanded eukaryotic genetic code. *Science 301*, 964–7.

(21) Mukai, T., Kobayashi, T., Hino, N., Yanagisawa, T., Sakamoto, K., and Yokoyama, S. (2008) Adding l-lysine derivatives to the genetic code of mammalian cells with engineered pyrrolysyl-tRNA synthetases. *Biochem. Biophys. Res. Commun. 371*, 818–22.

(22) Parrish, A. R., She, X., Xiang, Z., Coin, I., Shen, Z., Briggs, S. P., Dillin, A., and Wang, L. (2012) Expanding the genetic code of Caenorhabditis elegans using bacterial aminoacyl-tRNA synthetase/tRNA pairs. *ACS Chem. Biol. 7*, 1292–302.

(23) Ye, S., Riou, M., Carvalho, S., and Paoletti, P. (2013) Expanding the genetic code in Xenopus laevis oocytes. *Chembiochem 14*, 230–5.

(24) Greiss, S., and Chin, J. W. (2011) Expanding the genetic code of an animal. *J. Am. Chem. Soc. 133*, 14196–9.

(25) Bianco, A., Townsley, F. M., Greiss, S., Lang, K., and Chin, J. W. (2012) Expanding the genetic code of Drosophila melanogaster. *Nat. Chem. Biol. 8*, 748–50.

(26) Chin, J. W., Martin, A. B., King, D. S., Wang, L., and Schultz, P. G. (2002) Addition of a photocrosslinking amino acid to the genetic code of Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A. 99*, 11020–4.

(27) Wilkins, B. J., Rall, N. A., Ostwal, Y., Kruitwagen, T., Hiragami-Hamada, K., Winkler, M., Barral, Y., Fischle, W., and Neumann, H. (2014) A cascade of histone modifications induces chromatin condensation in mitosis. *Science 343*, 77–80.

(28) Neumann, H., Hancock, S. M., Buning, R., Routh, A., Chapman, L., Somers, J., Owen-Hughes, T., van Noort, J., Rhodes, D., and Chin, J. W. (2009) A method for genetically installing site-specific acetylation in recombinant histones defines the effects of H3 K56 acetylation. *Mol. Cell 36*, 153–63.

(29) Nguyen, D. P., Garcia Alai, M. M., Kapadnis, P. B., Neumann, H., and Chin, J. W. (2009) Genetically encoding N(epsilon)-methyl-L-lysine in recombinant histones. *J. Am. Chem. Soc. 131*, 14194–5.

(30) Nguyen, D. P., Garcia Alai, M. M., Virdee, S., and Chin, J. W. (2010) Genetically directing ε -N, N-dimethyl-L-lysine in recombinant histones. *Chem. Biol. 17*, 1072–6.

(31) Serwa, R., Wilkening, I., Del Signore, G., Mühlberg, M., Claussnitzer, I., Weise, C., Gerrits, M., and Hackenberger, C. P. R. (2009) Chemoselective Staudinger-phosphite reaction of azides for the phosphorylation of proteins. *Angew. Chem. Int. Ed. Engl. 48*, 8234–9.

(32) Xie, J., Supekova, L., and Schultz, P. G. (2007) A genetically encoded metabolically stable analogue of phosphotyrosine in Escherichia coli. *ACS Chem. Biol. 2*, 474–8.

(33) Park, H.-S., Hohn, M. J., Umehara, T., Guo, L.-T., Osborne, E. M., Benner, J., Noren, C. J., Rinehart, J., and Söll, D. (2011) Expanding the genetic code of Escherichia coli with phosphoserine. *Science 333*, 1151–4.

(34) Liu, C. C., and Schultz, P. G. (2006) Recombinant expression of selectively sulfated proteins in Escherichia coli. *Nat. Biotechnol. 24*, 1436–40.

(35) Neumann, H., Hazen, J. L., Weinstein, J., Mehl, R. A., and Chin, J. W. (2008) Genetically encoding protein oxidative damage. *J. Am. Chem. Soc. 130*, 4028–33.

(36) Deiters, A., Groff, D., Ryu, Y., Xie, J., and Schultz, P. G. (2006) A genetically encoded photocaged tyrosine. *Angew. Chem. Int. Ed. Engl. 45*, 2728–31.

(37) Wu, N., Deiters, A., Cropp, T. A., King, D., and Schultz, P. G. (2004) A genetically encoded photocaged amino acid. *J. Am. Chem. Soc. 126*, 14306–7.

(38) Lemke, E. A., Summerer, D., Geierstanger, B. H., Brittain, S. M., and Schultz, P. G. (2007) Control of protein phosphorylation with a genetically encoded photocaged amino acid. *Nat. Chem. Biol. 3*, 769–72.

(39) Gautier, A., Nguyen, D. P., Lusic, H., An, W., Deiters, A., and Chin, J. W. (2010) Genetically encoded photocontrol of protein localization in mammalian cells. *J. Am. Chem. Soc. 132*, 4086–8.

(40) Jackson, J. C., Hammill, J. T., and Mehl, R. A. (2007) Site-specific incorporation of a (19)F-amino acid into proteins as an NMR probe for characterizing protein structure and reactivity. *J. Am. Chem. Soc. 129*, 1160–6.

(41) Hammill, J. T., Miyake-Stoner, S., Hazen, J. L., Jackson, J. C., and Mehl, R. a. (2007) Preparation of site-specifically labeled fluorinated proteins for 19F-NMR structural characterization. *Nat. Protoc. 2*, 2601–7.

(42) Li, C., Wang, G.-F., Wang, Y., Creager-Allen, R., Lutz, E. A., Scronce, H., Slade, K. M., Ruf, R. A. S., Mehl, R. A., and Pielak, G. J. (2010) Protein (19)F NMR in Escherichia coli. *J. Am. Chem. Soc. 132*, 321–7.

(43) Schmidt, M. J., Borbas, J., Drescher, M., and Summerer, D. (2014) A genetically encoded spin label for electron paramagnetic resonance distance measurements. *J. Am. Chem. Soc. 136*, 1238–41.

(44) Chin, J. W., Santoro, S. W., Martin, A. B., King, D. S., Wang, L., and Schultz, P. G. (2002) Addition of p-azido-L-phenylalanine to the genetic code of Escherichia coli. *J. Am. Chem. Soc. 124*, 9026–7.

(45) Schultz, K. C., Supekova, L., Ryu, Y., Xie, J., Perera, R., and Schultz, P. G. (2006) A genetically encoded infrared probe. *J. Am. Chem. Soc. 128*, 13984–5.

(46) Ye, S., Zaitseva, E., Caltabiano, G., Schertler, G. F. X., Sakmar, T. P., Deupi, X., and Vogel, R. (2010) Tracking G-protein-coupled receptor activation using genetically encoded infrared probes. *Nature 464*, 1386–1389.

(47) Davis, L., and Chin, J. W. (2012) Designer proteins: applications of genetic code expansion in cell biology. *Nat. Rev. Mol. Cell Biol. 13*, 168–82.

(48) Xie, J., and Schultz, P. G. (2005) An expanding genetic code. *Methods 36*, 227–38.

(49) Schwarzer, D. (2009) Hacking the genetic code of mammalian cells. *Chembiochem 10*, 1602–4.

(50) Wan, W., Huang, Y., Wang, Z., Russell, W. K., Pai, P.-J., Russell, D. H., and Liu, W. R. (2010) A facile system for genetic incorporation of two different noncanonical amino acids into one protein in Escherichia coli. *Angew. Chem. Int. Ed. Engl. 49*, 3211–4.

(51) Odoi, K. a, Huang, Y., Rezenom, Y. H., and Liu, W. R. (2013) Nonsense and sense suppression abilities of original and derivative Methanosarcina mazei pyrrolysyl-tRNA synthetase-tRNA(Pyl) pairs in the Escherichia coli BL21(DE3) cell strain. *PLoS One 8*, e57035.

(52) Niu, W., Schultz, P. G., and Guo, J. (2013) An expanded genetic code in mammalian cells with a functional quadruplet codon. *ACS Chem. Biol. 8*, 1640–5.

(53) Anderson, J. C., Wu, N., Santoro, S. W., Lakshman, V., King, D. S., and Schultz, P. G. (2004) An expanded genetic code with a functional quadruplet codon. *Proc. Natl. Acad. Sci. U. S. A. 101*, 7566–71.

(54) Wang, K., Neumann, H., Peak-Chew, S. Y., and Chin, J. W. (2007) Evolved orthogonal ribosomes enhance the efficiency of synthetic genetic code expansion. *Nat. Biotechnol. 25*, 770–7.

(55) Neumann, H., Wang, K., Davis, L., Garcia-Alai, M., and Chin, J. W. (2010) Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome. *Nature 464*, 441–4.

(56) Luger, K., Rechsteiner, T. J., and Richmond, T. J. (1999) Preparation of nucleosome core particle from recombinant histones. *Methods Enzymol. 304*, 3–19.

(57) Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F., and Richmond, T. J. (1997) Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature 389*, 251–60.

(58) Wolffe, a. P., and Hayes, J. J. (1999) Chromatin disruption and modification. *Nucleic Acids Res. 27*, 711–720.

(59) Luger, K., Dechassa, M. L., and Tremethick, D. J. (2012) New insights into nucleosome and chromatin structure: an ordered state or a disordered affair? *Nat. Rev. Mol. Cell Biol. 13*, 436–47.

(60) Woodcock, C. L., and Dimitrov, S. (2001) Higher-order structure of chromatin and chromosomes. *Curr. Opin. Genet. Dev. 11*, 130–135.

(61) Hansen, J. C. (2002) Conformational dynamics of the chromatin fiber in solution: determinants, mechanisms, and functions. *Annu. Rev. Biophys. Biomol. Struct. 31*, 361–92.

(62) Campos, E. I., and Reinberg, D. (2009) Histones: annotating chromatin. *Annu. Rev. Genet. 43*, 559–99.

(63) Molina-Serrano, D., and Kirmizis, A. (2013) Beyond the histone tail: acetylation at the nucleosome dyad commands transcription. *Nucleus 4*, 343–8.

(64) Zentner, G. E., and Henikoff, S. (2013) Regulation of nucleosome dynamics by histone modifications. *Nat. Struct. Mol. Biol. 20*, 259–66.

(65) Allis, C. D., and Muir, T. W. (2011) Spreading chromatin into chemical biology. *Chembiochem 12*, 264–79.

(66) Lodish, H., Berk, A., Matsudaira, P., Kaiser, C. A., Krieger, M., Scott, M. P., Zipursky, L., and Darnell, J. (2004) Molecular Cell Biology 5th editio. W. H. Freeman.

(67) Helin, K., and Dhanak, D. (2013) Chromatin proteins and modifications as drug targets. *Nature 502*, 480–8.

(68) Gardner, K. E., Allis, C. D., and Strahl, B. D. (2011) Operating on chromatin, a colorful language where context matters. *J. Mol. Biol. 409*, 36–46.

(69) Ye, J., Ai, X., Eugeni, E. E., Zhang, L., Carpenter, L. R., Jelinek, M. A., Freitas, M. A., and Parthun, M. R. (2005) Histone H4 lysine 91 acetylation a core domain modification associated with chromatin assembly. *Mol. Cell 18*, 123–30.

(70) Tropberger, P., Pott, S., Keller, C., Kamieniarz-Gdula, K., Caron, M., Richter, F., Li, G., Mittler, G., Liu, E. T., Bühler, M., Margueron, R., and Schneider, R. (2013) Regulation of transcription through acetylation of H3K122 on the lateral surface of the histone octamer. *Cell 152*, 859–72.

(71) Bannister, A. J., and Kouzarides, T. (2011) Regulation of chromatin by histone modifications. *Cell Res. 21*, 381–95.

(72) Ge, Z., Nair, D., Guan, X., Rastogi, N., Freitas, M. a, and Parthun, M. R. (2013) Sites of acetylation on newly synthesized histone H4 are required for chromatin assembly and DNA damage response signaling. *Mol. Cell. Biol. 33*, 3286–98.

(73) Oppikofer, M., Kueng, S., Martino, F., Soeroes, S., Hancock, S. M., Chin, J. W., Fischle, W., and Gasser, S. M. (2011) A dual role of H4K16 acetylation in the establishment of yeast silent chromatin. *EMBO J. 30*, 2610–21.

(74) Shogren-Knaak, M., Ishii, H., Sun, J.-M., Pazin, M. J., Davie, J. R., and Peterson, C. L. (2006) Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science 311*, 844–7.

(75) Zhang, Y., and Reinberg, D. (2001) Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. *Genes Dev. 15*, 2343–60.

(76) Martin, C., and Zhang, Y. (2005) The diverse functions of histone lysine methylation. *Nat. Rev. Mol. Cell Biol. 6*, 838–49.

(77) Sims, R. J., Nishioka, K., and Reinberg, D. (2003) Histone lysine methylation: a signature for chromatin function. *Trends Genet. 19*, 629–39.

(78) Greeson, N. T., Sengupta, R., Arida, A. R., Jenuwein, T., and Sanders, S. L. (2008) Di-methyl H4 lysine 20 targets the checkpoint protein Crb2 to sites of DNA damage. *J. Biol. Chem. 283*, 33168–74.

(79) Sanders, S. L., Portoso, M., Mata, J., Bähler, J., Allshire, R. C., and Kouzarides, T. (2004) Methylation of histone H4 lysine 20 controls recruitment of Crb2 to sites of DNA damage. *Cell 119*, 603–14.

(80) Strahl, B. D., and Allis, C. D. (2000) The language of covalent histone modifications. *Nature 403*, 41–5.

(81) Grant, P. A. (2001) A tale of histone modifications. *Genome Biol. 2*, REVIEWS0003.

(82) Rossetto, D., Avvakumov, N., and Côté, J. (2012) Histone phosphorylation: a chromatin modification involved in diverse nuclear events. *Epigenetics 7*, 1098–108.

(83) Cheung, P., Tanner, K. G., Cheung, W. L., Sassone-Corsi, P., Denu, J. M., and Allis, C. D. (2000) Synergistic Coupling of Histone H3 Phosphorylation and Acetylation in Response to Epidermal Growth Factor Stimulation. *Mol. Cell 5*, 905–915.

(84) Clayton, A. L., Rose, S., Barratt, M. J., and Mahadevan, L. C. (2000) Phosphoacetylation of histone H3 on c-fos- and c-jun-associated nucleosomes upon gene activation. *EMBO J. 19*, 3714–26.

(85) Lo, W. S., Duggan, L., Emre, N. C., Belotserkovskya, R., Lane, W. S., Shiekhattar, R., and Berger, S. L. (2001) Snf1--a histone kinase that works in concert with the histone acetyltransferase Gcn5 to regulate transcription. *Science 293*, 1142–6.

(86) Zhang, B., Dong, Q., Su, H., Birchler, J. a, and Han, F. (2014) Histone phosphorylation: its role during cell cycle and centromere identity in plants. *Cytogenet. Genome Res. 143*, 144–9.

(87) Cheung, P., Allis, C. D., and Sassone-Corsi, P. (2000) Signaling to chromatin through histone modifications. *Cell 103*, 263–71.

(88) Hershko, A. (2005) The ubiquitin system for protein degradation and some of its roles in the control of the cell division cycle. *Cell Death Differ. 12*, 1191–7.

(89) Wang, H., Wang, L., Erdjument-Bromage, H., Vidal, M., Tempst, P., Jones, R. S., and Zhang, Y. (2004) Role of histone H2A ubiquitination in Polycomb silencing. *Nature 431*, 873–8.

(90) Lee, J.-S., Shukla, A., Schneider, J., Swanson, S. K., Washburn, M. P., Florens, L., Bhaumik, S. R., and Shilatifard, A. (2007) Histone crosstalk between H2B monoubiquitination and H3 methylation mediated by COMPASS. *Cell 131*, 1084–96.

(91) Sims, R. J., and Reinberg, D. (2008) Is there a code embedded in proteins that is based on post-translational modifications? *Nat. Rev. Mol. Cell Biol. 9*, 815–20.

(92) Turner, B. M. (2000) Histone acetylation and an epigenetic code. *BioEssays 22*, 836–845.

(93) Jenuwein, T., and Allis, C. D. (2001) Translating the histone code. *Science 293*, 1074–80.

(94) Arnaudo, A. M., and Garcia, B. A. (2013) Proteomic characterization of novel histone post-translational modifications. *Epigenetics Chromatin 6*, 24.

(95) Chen, Y., Chen, W., Cobb, M. H., and Zhao, Y. (2009) PTMap--a sequence alignment software for unrestricted, accurate, and full-spectrum identification of post-translational modification sites. *Proc. Natl. Acad. Sci. U. S. A. 106*, 761–6.

(96) Tweedie-Cullen, R. Y., Brunner, A. M., Grossmann, J., Mohanna, S., Sichau, D., Nanni, P., Panse, C., and Mansuy, I. M. (2012) Identification of combinatorial patterns of post-translational modifications on individual histones in the mouse brain. *PLoS One 7*, e36980.

(97) Jiang, T., Zhou, X., Taghizadeh, K., Dong, M., and Dedon, P. C. (2007) N-formylation of lysine in histone proteins as a secondary modification arising from oxidative DNA damage. *Proc. Natl. Acad. Sci. U. S. A. 104*, 60–5.

(98) Zhang, K., Chen, Y., Zhang, Z., and Zhao, Y. (2009) Identification and verification of lysine propionylation and butyrylation in yeast core histones using PTMap software. *J. Proteome Res. 8*, 900–6.

(99) Chen, Y., Sprung, R., Tang, Y., Ball, H., Sangras, B., Kim, S. C., Falck, J. R., Peng, J., Gu, W., and Zhao, Y. (2007) Lysine propionylation and butyrylation are novel post-translational modifications in histones. *Mol. Cell. Proteomics 6*, 812–9.

(100) Tan, M., Luo, H., Lee, S., Jin, F., Yang, J. S., Montellier, E., Buchou, T., Cheng, Z., Rousseaux, S., Rajagopal, N., Lu, Z., Ye, Z., Zhu, Q., Wysocka, J., Ye, Y., Khochbin, S., Ren, B., and Zhao, Y. (2011) Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell 146*, 1016–28.

(101) Peng, C., Lu, Z., Xie, Z., Cheng, Z., Chen, Y., Tan, M., Luo, H., Zhang, Y., He, W., Yang, K., Zwaans, B. M. M., Tishkoff, D., Ho, L., Lombard, D., He, T.-C., Dai, J., Verdin, E., Ye, Y., and Zhao, Y. (2011) The first identification of lysine malonylation substrates and its regulatory enzyme. *Mol. Cell. Proteomics 10*, M111.012658.

(102) Allfrey, V. G., Faulkner, R., and Mirsky, A. E. (1964) Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. *Proc. Natl. Acad. Sci. U. S. A. 51*, 786–94.

(103) Wisniewski, J. R., Zougman, A., and Mann, M. (2008) Nepsilon-formylation of lysine is a widespread post-translational modification of nuclear proteins occurring at residues involved in regulation of chromatin function. *Nucleic Acids Res. 36*, 570–7.

(104) Edrissi, B., Taghizadeh, K., and Dedon, P. C. (2013) Quantitative analysis of histone modifications: formaldehyde is a source of pathological n(6)-formyllysine that is refractory to histone deacetylases. *PLoS Genet. 9*, e1003328.

(105) Zhang, Z., Tan, M., Xie, Z., Dai, L., Chen, Y., and Zhao, Y. (2011) Identification of lysine succinylation as a new post-translational modification. *Nat. Chem. Biol. 7*, 58–63.

(106) Xie, Z., Dai, J., Dai, L., Tan, M., Cheng, Z., Wu, Y., Boeke, J. D., and Zhao, Y. (2012) Lysine succinylation and lysine malonylation in histones. *Mol. Cell. Proteomics 11*, 100–7.

(107) Montellier, E., Rousseaux, S., Zhao, Y., and Khochbin, S. (2012) Histone crotonylation specifically marks the haploid male germ cell gene expression program: post-meiotic male-specific gene expression. *Bioessays 34*, 187–93.

(108) Lee, S. (2013) Post-translational modification of proteins in toxicological research: focus on lysine acylation. *Toxicol. Res. 29*, 81–6.

(109) Grozinger, C. M., and Schreiber, S. L. (2002) Deacetylase Enzymes: Biological Functions and the Use of Small-Molecule Inhibitors. *Chem. Biol. 9*, 3–16.

(110) Thiagalingam, S., Cheng, K.-H., Lee, H. J., Mineva, N., Thiagalingam, A., and Ponte, J. F. (2003) Histone Deacetylases: Unique Players in Shaping the Epigenetic Histone Code. *Ann. N. Y. Acad. Sci. 983*, 84–100.

(111) Zhao, K., Chai, X., Clements, A., and Marmorstein, R. (2003) Structure and autoregulation of the yeast Hst2 homolog of Sir2. *Nat. Struct. Biol. 10*, 864–71.

(112) Zhao, K., Chai, X., and Marmorstein, R. (2004) Structure and substrate binding properties of cobB, a Sir2 homolog protein deacetylase from Escherichia coli. *J. Mol. Biol. 337*, 731–41.

(113) Feldman, J. L., Dittenhafer-Reed, K. E., and Denu, J. M. (2012) Sirtuin catalysis and regulation. *J. Biol. Chem. 287*, 42419–27.

(114) Du, J., Zhou, Y., Su, X., Yu, J. J., Khan, S., Jiang, H., Kim, J., Woo, J., Kim, J. H., Choi, B. H., He, B., Chen, W., Zhang, S., Cerione, R. A., Auwerx, J., Hao, Q., and Lin, H. (2011) Sirt5 is a NAD-dependent protein lysine demalonylase and desuccinylase. *Science 334*, 806–9.

(115) Colak, G., Xie, Z., Zhu, A. Y., Dai, L., Lu, Z., Zhang, Y., Wan, X., Chen, Y., Cha, Y. H., Lin, H., Zhao, Y., and Tan, M. (2013) Identification of lysine succinylation substrates and the succinylation regulatory enzyme CobB in Escherichia coli. *Mol. Cell. Proteomics 12*, 3509–20.

(116) Cheng, Z., Tang, Y., Chen, Y., Kim, S., Liu, H., Li, S. S. C., Gu, W., and Zhao, Y. (2009) Molecular characterization of propionyllysines in non-histone proteins. *Mol. Cell. Proteomics 8*, 45–52.

(117) Smith, B. C., and Denu, J. M. (2007) Acetyl-lysine analog peptides as mechanistic probes of protein deacetylases. *J. Biol. Chem. 282*, 37256–65.

(118) Leemhuis, H., Packman, L. C., Nightingale, K. P., and Hollfelder, F. (2008) The human histone acetyltransferase P/CAF is a promiscuous histone propionyltransferase. *Chembiochem 9*, 499–503.

(119) Garrity, J., Gardner, J. G., Hawse, W., Wolberger, C., and Escalante-Semerena, J. C. (2007) N-lysine propionylation controls the activity of propionyl-CoA synthetase. *J. Biol. Chem. 282*, 30239–45.

(120) Tsien, R. Y. (1998) The green fluorescent protein. *Annu. Rev. Biochem. 67*, 509–44.

(121) Merrifield, R. B. (1963) Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *J. Am. Chem. Soc. 85*, 2149–2154.

(122) Muralidharan, V., and Muir, T. W. (2006) Protein ligation: an enabling technology for the biophysical analysis of proteins. *Nat. Methods 3*, 429–38.

(123) (2010) Thiol-Reactive Probes, in *The Molecular Probes Handbook* 11th Editi. Invitrogen.

(124) Chalker, J. M., Bernardes, G. J. L., Lin, Y. a, and Davis, B. G. (2009) Chemical modification of proteins at cysteine: opportunities in chemistry and biology. *Chem. Asian J. 4*, 630–40.

(125) Kim, Y., Ho, S. O., Gassman, N. R., Korlann, Y., Landorf, E. V, Collart, F. R., and Weiss, S. (2008) Efficient site-specific labeling of proteins via cysteines. *Bioconjug. Chem. 19*, 786–91.

(126) Hayworth, D. Amine-reactive Crosslinker Chemistry. *Thermo Sci.*

(127) Neumann, H. (2012) Rewiring translation - Genetic code expansion and its applications. *FEBS Lett. 586*, 2057–64.

(128) Summerer, D., Chen, S., Wu, N., Deiters, A., Chin, J. W., and Schultz, P. G. (2006) A genetically encoded fluorescent amino acid. *Proc. Natl. Acad. Sci. U. S. A. 103*, 9785–9.

(129) Wang, L., Brock, A., and Schultz, P. G. (2002) Adding l -3-(2-Naphthyl)alanine to the Genetic Code of E. coli. *J. Am. Chem. Soc. 124*, 1836–1837.

(130) Lang, K., and Chin, J. W. (2014) Cellular incorporation of unnatural amino acids and bioorthogonal labeling of proteins. *Chem. Rev. 114*, 4764–806.

(131) Jewett, J. C., and Bertozzi, C. R. (2010) Cu-free click cycloaddition reactions in chemical biology. *Chem. Soc. Rev. 39*, 1272–9.

(132) Nguyen, D. P., Lusic, H., Neumann, H., Kapadnis, P. B., Deiters, A., and Chin, J. W. (2009) Genetic encoding and labeling of aliphatic azides and alkynes in recombinant proteins via a pyrrolysyl-tRNA Synthetase/tRNA(CUA) pair and click chemistry. *J. Am. Chem. Soc. 131*, 8720–1.

(133) Milles, S., Tyagi, S., Banterle, N., Koehler, C., VanDelinder, V., Plass, T., Neal, A. P., and Lemke, E. A. (2012) Click strategies for single-molecule protein fluorescence. *J. Am. Chem. Soc. 134*, 5187–95.

(134) Böhrsch, V., Serwa, R., Majkut, P., Krause, E., and Hackenberger, C. P. R. (2010) Site-specific functionalisation of proteins by a Staudinger-type reaction using unsymmetrical phosphites. *Chem. Commun. (Camb). 46*, 3176–8.

(135) Kaya, E., Vrabel, M., Deiml, C., Prill, S., Fluxa, V. S., and Carell, T. (2012) A genetically encoded norbornene amino acid for the mild and selective modification of proteins in a copper-free click reaction. *Angew. Chem. Int. Ed. Engl. 51*, 4466–9.

(136) Lang, K., Davis, L., Wallace, S., Mahesh, M., Cox, D. J., Blackman, M. L., Fox, J. M., and Chin, J. W. (2012) Genetic Encoding of bicyclononynes and trans-cyclooctenes for site-specific protein labeling in vitro and in live mammalian cells via rapid fluorogenic Diels-Alder reactions. *J. Am. Chem. Soc. 134*, 10317–20.

(137) Lang, K., Davis, L., Torres-Kolbus, J., Chou, C., Deiters, A., and Chin, J. W. (2012) Genetically encoded norbornene directs site-specific cellular protein labelling via a rapid bioorthogonal reaction. *Nat. Chem. 4*, 298–304.

(138) Devaraj, N. K., Weissleder, R., and Hilderbrand, S. A. (2008) Tetrazine-based cycloadditions: application to pretargeted live cell imaging. *Bioconjug. Chem. 19*, 2297–9.

(139) Blackman, M. L., Royzen, M., and Fox, J. M. (2008) Tetrazine ligation: fast bioconjugation based on inverse-electron-demand Diels-Alder reactivity. *J. Am. Chem. Soc. 130*, 13518–9.

(140) Tóth, K., Böhm, V., Sellmann, C., Danner, M., Hanne, J., Berg, M., Barz, I., Gansen, A., and Langowski, J. (2013) Histone- and DNA sequence-dependent stability of nucleosomes studied by single-pair FRET. *Cytometry. A 83*, 839–46.

(141) Lakowicz, J. R. (2006) Principles of Fluorescence Spectroscopy Third Edit. Springer Scince+Buisness Media.

(142) Buning, R., and van Noort, J. (2010) Single-pair FRET experiments on nucleosome conformational dynamics. *Biochimie 92*, 1729–40.

(143) Li, G., Levitus, M., Bustamante, C., and Widom, J. (2005) Rapid spontaneous accessibility of nucleosomal DNA. *Nat. Struct. Mol. Biol. 12*, 46–53.

(144) Gansen, A., Tóth, K., Schwarz, N., and Langowski, J. (2009) Structural variability of nucleosomes detected by single-pair Förster resonance energy transfer: histone acetylation, sequence variation, and salt effects. *J. Phys. Chem. B 113*, 2604–13.

(145) Li, G., and Widom, J. (2004) Nucleosomes facilitate their own invasion. *Nat. Struct. Mol. Biol. 11*, 763–9.

(146) Tóth, K., Brun, N., and Langowski, J. (2006) Chromatin compaction at the mononucleosome level. *Biochemistry 45*, 1591–8.

(147) Llères, D., James, J., Swift, S., Norman, D. G., and Lamond, A. I. (2009) Quantitative analysis of chromatin compaction in living cells using FLIM-FRET. *J. Cell Biol. 187*, 481–96.

(148) Mora-Bermúdez, F., Gerlich, D., and Ellenberg, J. (2007) Maximal chromosome compaction occurs by axial shortening in anaphase and depends on Aurora kinase. *Nat. Cell Biol. 9*, 822–31.

(149) Kavran, J. M., Gundllapalli, S., O'Donoghue, P., Englert, M., Söll, D., and Steitz, T. A. (2007) Structure of pyrrolysyl-tRNA synthetase, an archaeal enzyme for genetic code innovation. *Proc. Natl. Acad. Sci. U. S. A. 104*, 11268–73.

(150) Yanagisawa, T., Ishii, R., Fukunaga, R., Kobayashi, T., Sakamoto, K., and Yokoyama, S. (2008) Multistep engineering of pyrrolysyl-tRNA synthetase to genetically encode N(epsilon)-(o-azidobenzyloxycarbonyl) lysine for site-specific protein modification. *Chem. Biol. 15*, 1187–97.

(151) Pick, H., Kilic, S., and Fierz, B. (2014) Engineering chromatin states: chemical and synthetic biology approaches to investigate histone modification function. *Biochim. Biophys. Acta 1839*, 644–56.

(152) Albaugh, B. N., Arnold, K. M., and Denu, J. M. (2011) KAT(ching) metabolism by the tail: insight into the links between lysine acetyltransferases and metabolism. *Chembiochem 12*, 290–8.

(153) Lajoie, M. J., Rovner, A. J., Goodman, D. B., Aerni, H.-R., Haimovich, A. D., Kuznetsov, G., Mercer, J. a, Wang, H. H., Carr, P. a, Mosberg, J. a, Rohland, N., Schultz, P. G., Jacobson, J. M., Rinehart, J., Church, G. M., and Isaacs, F. J. (2013) Genomically recoded organisms expand biological functions. *Science 342*, 357–60.

(154) Plazas-Mayorca, M. D., Zee, B. M., Young, N. L., Fingerman, I. M., LeRoy, G., Briggs, S. D., and Garcia, B. A. (2009) One-pot shotgun quantitative mass spectrometry characterization of histones. *J. Proteome Res. 8*, 5367–74.

(155) Garcia, B. A., Mollah, S., Ueberheide, B. M., Busby, S. A., Muratore, T. L., Shabanowitz, J., and Hunt, D. F. (2007) Chemical derivatization of histones for facilitated analysis by mass spectrometry. *Nat. Protoc. 2*, 933–8.

(156) Schwarz, P. M., Felthauser, a, Fletcher, T. M., and Hansen, J. C. (1996) Reversible oligonucleosome self-association: dependence on divalent cations and core histone tail domains. *Biochemistry 35*, 4009–15.

(157) Gordon, F., Luger, K., and Hansen, J. C. (2005) The core histone N-terminal tail domains function independently and additively during salt-dependent oligomerization of nucleosomal arrays. *J. Biol. Chem. 280*, 33701–6.

(158) Tse, C., and Hansen, J. C. (1997) Hybrid trypsinized nucleosomal arrays: identification of multiple functional roles of the H2A/H2B and H3/H4 N-termini in chromatin fiber compaction. *Biochemistry 36*, 11381–8.

(159) Hebbes, T. R., Thorne, A. W., and Crane-Robinson, C. (1988) A direct link between core histone acetylation and transcriptionally active chromatin. *EMBO J. 7*, 1395–402.

(160) Wang, X., and Hayes, J. J. (2008) Acetylation mimics within individual core histone tail domains indicate distinct roles in regulating the stability of higher-order chromatin structure. *Mol. Cell. Biol. 28*, 227–36.

(161) Tse, C., Sera, T., Wolffe, A. P., and Hansen, J. C. (1998) Disruption of higher-order folding by core histone acetylation dramatically enhances transcription of nucleosomal arrays by RNA polymerase III. *Mol. Cell. Biol. 18*, 4629–38.

(162) Bitterman, K. J., Anderson, R. M., Cohen, H. Y., Latorre-Esteves, M., and Sinclair, D. a. (2002) Inhibition of silencing and accelerated aging by nicotinamide, a putative negative regulator of yeast sir2 and human SIRT1. *J. Biol. Chem. 277*, 45099–107.

(163) Yuan, H., and Marmorstein, R. (2012) Structural basis for sirtuin activity and inhibition. *J. Biol. Chem. 287*, 42428–35.

(164) Lindsey, G. G., Thompson, P., Pretorius, L., and von Holt, C. (1985) Fluorescent labelling of histone H3: effect on histone-histone interaction and core particle assembly. *FEBS Lett. 192*, 230–4.

(165) Lowary, P. T., and Widom, J. (1998) New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol. 276*, 19–42.

(166) Wessel, D., and Flügge, U. I. (1984) A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Anal. Biochem. 138*, 141–3.

(167) Sanchez, L. (2001) TCA precipitation protocol.

(168) Bensadoun, A., and Weinstein, D. (1976) Assay of proteins in the presence of interfering materials. *Anal. Biochem. 70*, 241–250.

(169) Wessel, D., and Flügge, U. I. (1984) A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Anal. Biochem. 138*, 141–3.

(170) Rackham, O., and Chin, J. W. (2005) A network of orthogonal ribosome x mRNA pairs. *Nat. Chem. Biol. 1*, 159–66.

(171) Lammers, C., Hahn, L. E., and Neumann, H. (2014) Optimized plasmid systems for the incorporation of multiple different unnatural amino acids by evolved orthogonal ribosomes. *Chembiochem 15*, 1800–4.

(172) Haldar, S., and Chattopadhyay, A. (2013) Fluorescent Methods to Study Biological Membranes (Mély, Y., and Duportail, G., Eds.), pp 37–50. Springer Berlin Heidelberg, Berlin, Heidelberg.

(173) Flügel, V., Vrabel, M., and Schneider, S. (2014) Structural basis for the site-specific incorporation of lysine derivatives into proteins. *PLoS One 9*, e96198.

(174) Wilkins, B. J., Hahn, L. E., Heitmüller, S., Frauendorf, H., Valerius, O., Braus, G. H., and Neumann, H. (2015) Genetically Encoding Lysine Modifications on Histone H4. *ACS Chem. Biol.*

(175) Gattner, M. J., Vrabel, M., and Carell, T. (2013) Synthesis of ε-N-propionyl-, ε-N-butyryl-, and ε-N-crotonyl-lysine containing histone H3 using the pyrrolysine system. *Chem. Commun. (Camb). 49*, 379–81.

(176) Kim, C. H., Kang, M., Kim, H. J., Chatterjee, A., and Schultz, P. G. (2012) Site-specific incorporation of ε-N-crotonyllysine into histones. *Angew. Chem. Int. Ed. Engl. 51*, 7246–9.

(177) Guo, L., Wang, Y., Nakamura, A., Eiler, D., Kavran, J. M., Wong, M., Kiessling, L. L., Steitz, T. A., O'Donoghue, P., and Söll, D. (2014) Polyspecific pyrrolysyl-tRNA synthetases from directed evolution. *Proc. Natl. Acad. Sci. U. S. A. 111*.

(178) Isaacs, F. J., Carr, P. A., Wang, H. H., Lajoie, M. J., Sterling, B., Kraal, L., Tolonen, A. C., Gianoulis, T. A., Goodman, D. B., Reppas, N. B., Emig, C. J., Bang, D., Hwang, S. J., Jewett, M. C., Jacobson, J. M., and Church, G. M. (2011) Precise manipulation of chromosomes in vivo enables genome-wide codon replacement. *Science 333*, 348–53.

(179) Dorigo, B., Schalch, T., Bystricky, K., and Richmond, T. J. (2003) Chromatin Fiber Folding: Requirement for the Histone H4 N-terminal Tail. *J. Mol. Biol. 327*, 85–96.

(180) Dion, M. F., Altschuler, S. J., Wu, L. F., and Rando, O. J. (2005) Genomic characterization reveals a simple histone H4 acetylation code. *Proc. Natl. Acad. Sci. U. S. A. 102*, 5501–6.

(181) Shogren-Knaak, M., and Peterson, C. L. (2014) Switching on Chromatin: Mechanistic Role of Histone H4-K16 Acetylation. *Cell Cycle 5*, 1361–1365.

(182) Schwarz, P. M., and Hansen, J. C. (1994) Formation and stability of higher order chromatin structures. Contributions of the histone octamer. *J. Biol. Chem. 269*, 16284–9.

(183) Bheda, P., Wang, J. T., Escalante-Semerena, J. C., and Wolberger, C. (2011) Structure of Sir2Tm bound to a propionylated peptide. *Protein Sci. 20*, 131–9.

(184) Pan, M., Yuan, H., Brent, M., Ding, E. C., and Marmorstein, R. (2012) SIRT1 contains N- and C-terminal regions that potentiate deacetylase activity. *J. Biol. Chem. 287*, 2468–76.

(185) Brustad, E. M., Lemke, E. A., Schultz, P. G., and Deniz, A. A. (2008) A general and efficient method for the site-specific dual-labeling of proteins for single molecule fluorescence resonance energy transfer. *J. Am. Chem. Soc. 130*, 17664–5.

(186) Milles, S., and Lemke, E. a. (2011) Single molecule study of the intrinsically disordered FG-repeat nucleoporin 153. *Biophys. J. 101*, 1710–9.

(187) Nguyen, D. P., Elliott, T., Holt, M., Muir, T. W., and Chin, J. W. (2011) Genetically encoded 1,2-aminothiols facilitate rapid and site-specific protein labeling via a bio-orthogonal cyanobenzothiazole condensation. *J. Am. Chem. Soc. 133*, 11418–21.

(188) Tsao, M.-L., Tian, F., and Schultz, P. G. (2005) Selective Staudinger modification of proteins containing p-azidophenylalanine. *Chembiochem 6*, 2147–9.

(189) Kiick, K. L., Saxon, E., Tirrell, D. A., and Bertozzi, C. R. (2002) Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation. *Proc. Natl. Acad. Sci. U. S. A. 99*, 19–24.

(190) Lang, K., and Chin, J. W. (2014) Bioorthogonal reactions for labeling proteins. *ACS Chem. Biol. 9*, 16–20.

(191) Buschmann, V., Weston, K. D., and Sauer, M. (2003) Spectroscopic study and evaluation of red-absorbing fluorescent dyes. *Bioconjug. Chem. 14*, 195–204.

(192) Berlier, J. E., Rothe, a., Buller, G., Bradford, J., Gray, D. R., Filanoski, B. J., Telford, W. G., Yue, S., Liu, J., Cheung, C.-Y., Chang, W., Hirsch, J. D., Beechem Rosaria P. Haugland, J. M., and Haugland, R. P. (2003) Quantitative Comparison of Long-wavelength Alexa Fluor Dyes to Cy Dyes: Fluorescence of the Dyes and Their Bioconjugates. *J. Histochem. Cytochem. 51*, 1699–1712.

(193) Staros, J. V, Bayley, H., Standring, D. N., and Knowles, J. R. (1978) Reduction of aryl azides by thiols: Implications for the use of photoaffinity reagents. *Biochem. Biophys. Res. Commun. 80*, 568–572.

(194) Sasmal, P. K., Carregal-Romero, S., Han, A. a, Streu, C. N., Lin, Z., Namikawa, K., Elliott, S. L., Köster, R. W., Parak, W. J., and Meggers, E. (2012) Catalytic azide reduction in biological environments. *Chembiochem 13*, 1116–20.

(195) Tyagarajan, K., Pretzer, E., and Wiktorowicz, J. E. (2003) Thiol-reactive dyes for fluorescence labeling of proteomic samples. *Electrophoresis 24*, 2348–58.

(196) Getz, E. B., Xiao, M., Chakrabarty, T., Cooke, R., and Selvin, P. R. (1999) A comparison between the sulfhydryl reductants tris(2-carboxyethyl)phosphine and dithiothreitol for use in protein biochemistry. *Anal. Biochem. 273*, 73–80.

(197) Chatterjee, A., Sun, S. B., Furman, J. L., Xiao, H., and Schultz, P. G. (2013) A versatile platform for single- and multiple-unnatural amino acid mutagenesis in Escherichia coli. *Biochemistry 52*, 1828–37.

(198) Wang, K., Sachdeva, A., Cox, D. J., Wilf, N. W., Lang, K., Wallace, S., Mehl, R. a, and Chin, J. W. (2014) Optimized orthogonal translation of unnatural amino acids enables spontaneous protein double-labelling and FRET. *Nat. Chem. 6*, 393–403.

(199) Plass, T., Milles, S., Koehler, C., Schultz, C., and Lemke, E. A. (2011) Genetically encoded copper-free click chemistry. *Angew. Chem. Int. Ed. Engl. 50*, 3878–81.

(200) Fery-Forgues, S., Fayet, J.-P., and Lopez, A. (1993) Drastic changes in the fluorescence properties of NBD probes with the polarity of the medium: involvement of a TICT state? *J. Photochem. Photobiol. A Chem. 70*, 229–243.

(201) Young, D. D., Young, T. S., Jahnz, M., Ahmad, I., Spraggon, G., and Schultz, P. G. (2011) An evolved aminoacyl-tRNA synthetase with atypical polysubstrate specificity. *Biochemistry 50*, 1894–900.

(202) Mullis, K. B., and Faloona, F. A. (1987) Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods Enzymol. 155*, 335–50.

(203) Williams, M., Louw, A. I., and Birkholtz, L. (2007) Deletion mutagenesis of large areas in Plasmodium falciparum genes: a comparative study. *Malar. J. 6*, 64.

(204) Ochman, H., Gerber, A. S., and Hartl, D. L. (1988) Genetic applications of an inverse polymerase chain reaction. *Genetics 120*, 621–3.

(205) Dower, W. J., Miller, J. F. J., and Ragsdale, C. W. C. (1988) High efficiency transformation of E. coli by high voltage electroporation. *Nucleic Acids Res. 16*, 6127–6145.

(206) Laemmli, U. K. (1970) Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature 227*, 680–685.

(207) Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem. 72*, 248–54.

(208) Tsukiyama. Expression , Purification and Reconstitution of Recombinant Yeast Histones.

6 List of Figures

Figure 1 *Incorporation of an unnatural amino acid (UAA)*......2

Figure 2 *Selection of a PylS mutant library*......4

Figure 3 *Nucleosome assembly.*......8

Figure 4 *DNA compaction into chromatin.*......9

Figure 5 *PTMs on histone tails.*......10

Figure 6 *Novel histone lysine acyl modifications.*......12

Figure 7 *Acetyl modifying enzymes.*......14

Figure 8 *Sirtuin catalyzed deacetylation reaction*......15

Figure 9 *Structures of UAAs suitable for bioorthogonal labeling reactions.*......17

Figure 10 *Scheme of different click reactions used for labeling of biomolecules*......18

Figure 11 *Structure of the active site of M. mazei*......21

Figure 12 *Selection of the PylS library on PrK*......22

Figure 13 *Selection of the PylS library on mutants for CrK incorporation.*......23

Figure 14 *Western blot of his$_6$-tagged myoglobin4TAG test expression with the selected synthetases for incorporation of PrK and CrK*......23

Figure 15 *Specificity of the synthetases was tested by expression of his$_6$-tagged myoglobin4TAG with different combinations of synthetases and amino acids.*......25

Figure 16 *Analysis of acyl modified myoglobin by ESI-TOF.*......26

Figure 17 *Acyl modifications on the histone H3 tail.*......27

Figure 18 *Acyl modifications on histone H4 tail.*......27

Figure 19 *Overview construction of histone H4 including an UAA.*......28

Figure 20 *Test expression of the H3(Δ93-98)-TEV-H4 K16TAG construct.*......28

Figure 21 *Purified His$_6$-H3(Δ93-98)TEV-H4 K16mod.*......29

Figure 22 *TEV cleavage of the purified H3(Δ93-98)-TEV-H4mod*......30

Figure 23 *Analysis of tryptical fragments from H4 K16ac by mass spectrometry*......31

Figure 24 *Antibody specificity test.*......32

Figure 25 *Antibody specificity test using chemically altered H4mod.*......33

Figure 26 *Effect of the precipitated H3 truncation on the formation of histone octamers*......34

Figure 27 *Purification of refolded octamers including H4 K16Bu.*......35

Figure 28 *Recognition of acyl modified octamers by different antibodies*......36

Figure 29 *Nucleosome arrays assembled from H4 K16mod octamers*......37

Figure 30 *MNase digest of H4mod nucleosome arrays.*......38

Figure 31 *Nucleosome self-association assay.* 39
Figure 32 *Purified CobB.* 40
Figure 33 R*emoval of acyl modifications at K16 on histone H4 by CobB.* 40
Figure 34 R*emoval of acyl group by CobB.* 41
Figure 35 *Cysteine position on histone H3.* 42
Figure 36 *Octamer H3 Q76C-Alexa546.* 43
Figure 37 *Labeling of GFP with DBCO-dye.* 43
Figure 38 *GFP labeling tested with different dye concentrations.* 44
Figure 39 *Purification of H3 Q76C V35AzF.* 45
Figure 40 *Investigation of reaction conditions for DBCO labeling.* 46
Figure 41 *Test of different dye concentrations for double labeling.* 47
Figure 42 *Octamer refolding including double labeled H3.* 48
Figure 43 *Overview of labeled Octamers.* 50
Figure 44 *Overview of labeled Nucleosomes.* 51
Figure 45 *Separation of free dye by protein precipitations.* 52
Figure 46 *GST-MBP expressed and purified using different plasmids and amino acid combinations.* 53
Figure 47 *Purified GST-MBP with different UAA combinations.* 54
Figure 48 *Labeled GST-MBP.* 54
Figure 49 *Structure of NBD-K.* 55
Figure 50 *NBD-K uptake in E. coli cells.* 55
Figure 51 *Library selection for a mutant which recognizes NBD-K.* 56
Figure 52 *Test expression of myoglobin in the presence and absence of NBD-K.* 56
Figure 53 *Purification of myoglobin containing NBD-K at position 4.* 57
Figure 54 *Summary creation of nucleosomes containing a FRET pair.* 63
Figure 55 *Overview of positions replaced by Cys and tested for labeling.* 64
Figure 56 *Schematic procedure of the inverse PCR.* 83
Figure 57 *Purification of TEV protease.* 96
Figure 58 *Hst2 purification by size-exclusion chromatography using a S75 column.* 97
Figure 59 *Purification of WT octamers.* 131
Figure 60 *Purification of octamers including H4 K16ac.* 131
Figure 61 *Purification of octamer including H4 K16Cr.* 132
Figure 62 *Purification of octamers including H4 K16Pr.* 132

7 List of Tables

Table 2-1 *Changes in sequence of the selected synthetases compared to WT PylS.* 24

Table 2-2 *H4 K16Ac peptide data from ESI-MSMS* 31

Table 2-3 *H3 mutants with replaced cysteine, subsequently labeled and tested upon ability to form octamers.* 43

Table 2-4 *Variations of AzF position and dye combinations.* 49

Table 2-5 *Sequence changes in mutants selected to recognize NBD-K.* 57

Table 4-1 *Unnatural Amino Acids, concentrations and plasmids used for incorporation.* 72

Table 4-2 *Fluorescent dyes.* 72

Table 4-3 *Composition of SDS gels.* 75

Table 4-4 *Antibiotics and work concentrations.* 75

Table 4-5 *Cell lines.* 76

Table 4-6 *Antibodies.* 78

Table 4-7 *General set up for a PCR reaction* 81

Table 4-8 *Enzymatic set up for quick change reactions.* 82

Table 4-9 *Temperature profile for quick change program.* 82

Table 4-10 *Set up for inverse PCR.* 83

Table 4-11 *Enzymatic set up for sequencing reaction* 85

Table 4-12 *Temperature profile for sequencing PCR, cycles being repeated 25 times.* 86

Table 4-13 *Instructions for 2 native PAGE gels.* 89

Table 4-14 *Buffers for purification of histones by ion exchange chromatography.* 92

Table 4-15 *Salt gradients for elution of histones from SP column as published by Tsukiyama*[208]. 92

Table 4-16 *Plasmids for incorporation of acyl modification and concentrations of acyl amino acids.* 93

Table 4-17 *Molecular weight of the different histone proteins and their absorption at 276 nm.* 98

Table 4-18 *Buffer composition for Unfolding- and Refolding buffer.* 98

Table 4-19 *Buffers for nucleosome assembly.* 99

Table 4-20 *Temperature program for inverse PCR reaction.* 102

Table 4-21 *Enzymatic set up for ligation reaction.* 102

8 Appendix

8.1 Purification of acylated octamers

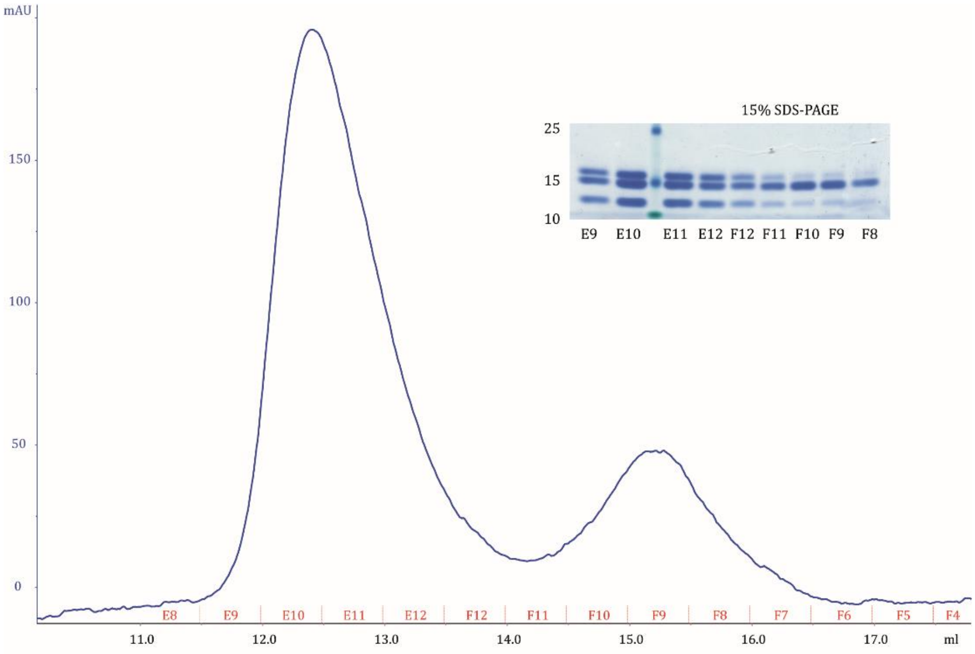


Figure 59 ***Purification of WT octamers.*** Octamers were purified by size-exclusion chromatography using a S200 column. X-axis volume in mL, Y-axis absorption at 280 nm, fraction indicated in red. Small picture presents analysis of fractions by 15% SDS-PAGE visualized by Instant Blue.

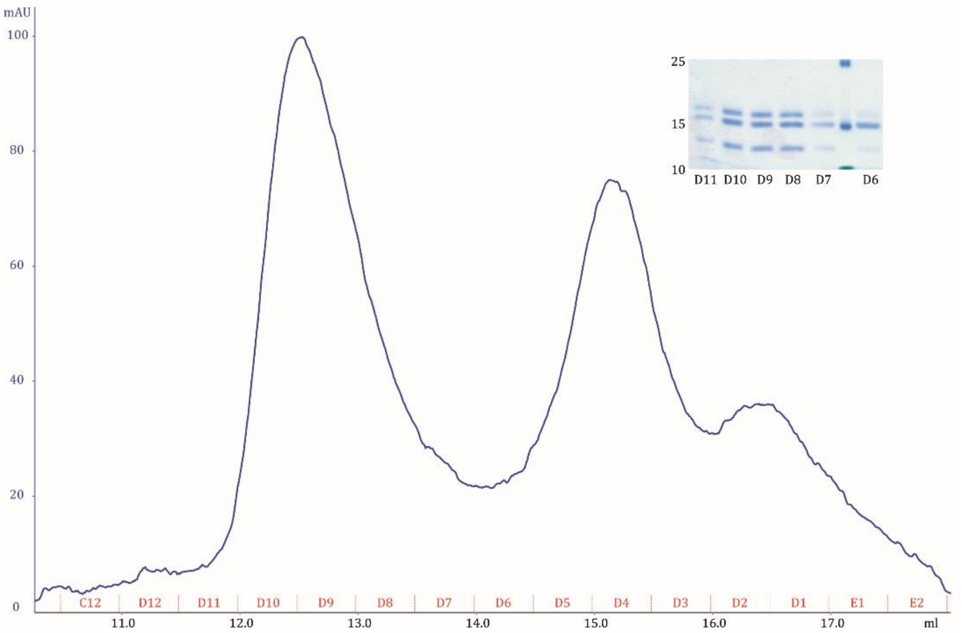


Figure 60 ***Purification of octamers including H4 K16ac.*** Octamers were purified by size-exclusion chromatography using a S200 column. X-axis volume in mL, Y-axis absorption at 280 nm, fraction indicated in red. Small picture presents analysis of fractions by 15% SDS-PAGE visualized by Instant Blue.

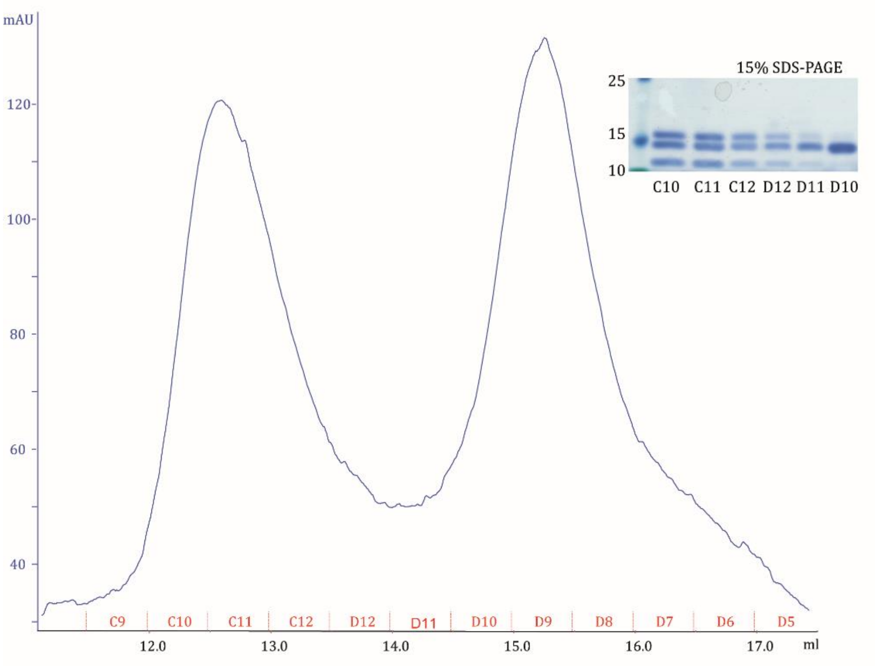


Figure 61 ***Purification of octamer including H4 K16Cr.*** Octamers were purified by size-exclusion chromatography using a S200 column. X-axis volume in mL, Y-axis absorption at 280 nm, fraction indicated in red. Small picture presents analysis of fractions by 15% SDS-PAGE visualized by Instant Blue.

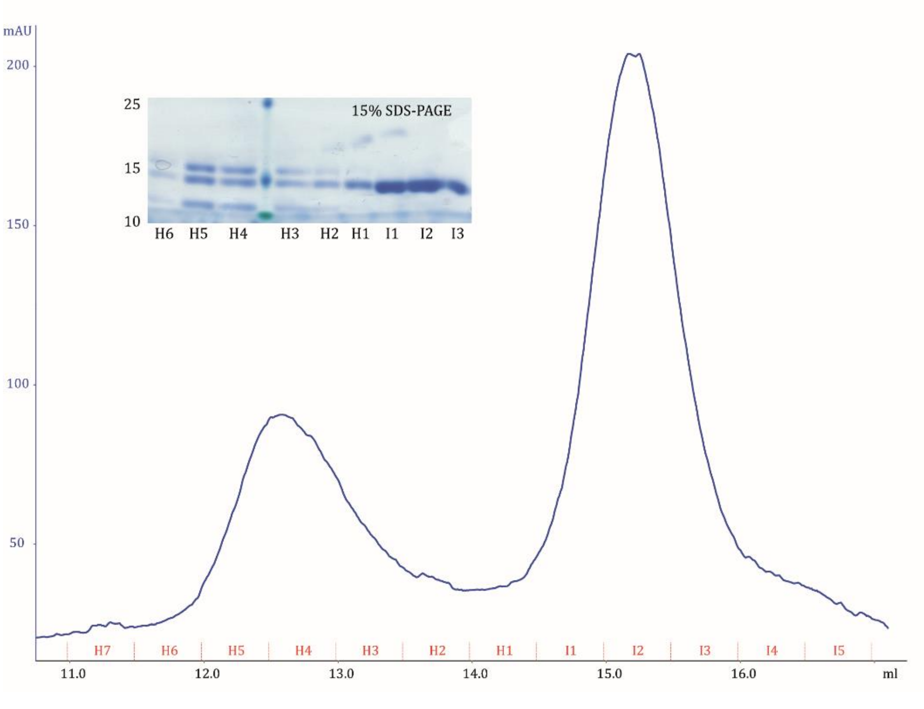


Figure 62 ***Purification of octamers including H4 K16Pr.*** Octamers were purified by size-exclusion chromatography using a S200 column. X-axis volume in mL, Y-axis absorption at 280 nm, fraction indicated in red. Small picture presents analysis of fractions by 15% SDS-PAGE visualized by Instant Blue.

9 Curriculum Vitae

Personal Data

Name: Liljan E. Hahn
Date and place of birth: 17.07.1984, Oldenburg, Germany
Nationality: German
E-Mail: lhahn1@gwdg.de

Education

02/2011-03/2015	Doctoral thesis entitled "*Investigation of Nucleosome Dynamics by Genetic Code Expansion*" Supervisor Prof. H. Neumann, University of Göttingen
12/2010	Diploma in chemistry, University of Göttingen
03-10/2010	Diploma Thesis "*Incorporation of fluorescence-active probes by genetic code expansion*" supervisor Prof. H. Neumann, University of Göttingen
04-09/2009	Erasmus exchange, Stockholm University, Supervisor Prof. G. Widmalm
10/2005-12/2010	Academic studies of chemistry at the University of Göttingen
06/2004	Allgemeine Hochschulreife (Abitur)

Publications

Wilkins BJ[1], **Hahn LE**[1], Heitmüller S, Frauendorf H, Valerius O, Braus GH, Neumann H. (2015) Genetically Encoding Lysine Modifications on Histone H4, *ACS Chem Biol.* 10, 939-44.

Lammers C, **Hahn LE**, Neumann H. (2014) Optimized plasmid systems for the incorporation of multiple different unnatural amino acids by evolved orthogonal ribosomes, *Chembiochem 15*, 1800-1804.

Rönnols J, Pendrill R, Fontana C, Hamark C, d'Ortoli TA, Engström O, Ståhle J, Zaccheus MV, Säwén E, **Hahn LE**, Iqbal S, Widmalm G. (2013) Complete (1)H and (13)C NMR chemical shift assignments of mono- to tetrasaccharides as basis for NMR chemical shift predicitions of oligosaccharides using the computer program CASPER, *Carbohydr Res.* 380, 156-66.

---

[1] Equal contribution