



Structural Basis for the Site-Specific Incorporation of Lysine Derivatives into Proteins

Veronika Flügel¹, Milan Vrabel², Sabine Schneider^{1*}

1 Department of Chemistry, TU Munich, Garching, Germany, **2** Department of Chemistry, Ludwig-Maximilians University, Munich, Germany

Abstract

Posttranslational modifications (PTMs) of proteins determine their structure-function relationships, interaction partners, as well as their fate in the cell and are crucial for many cellular key processes. For instance chromatin structure and hence gene expression is epigenetically regulated by acetylation or methylation of lysine residues in histones, a phenomenon known as the 'histone code'. Recently it was shown that these lysine residues can furthermore be malonylated, succinylated, butyrylated, propionylated and crotonylated, resulting in significant alteration of gene expression patterns. However the functional implications of these PTMs, which only differ marginally in their chemical structure, is not yet understood. Therefore generation of proteins containing these modified amino acids site specifically is an important tool. In the last decade methods for the translational incorporation of non-natural amino acids using orthogonal aminoacyl-tRNA synthetase (aaRS):tRNA^{AaCUA} pairs were developed. A number of studies show that aaRS can be evolved to use non-natural amino acids and expand the genetic code. Nevertheless the wild type pyrrolysyl-tRNA synthetase (PylRS) from *Methanosarcina mazei* readily accepts a number of lysine derivatives as substrates. This enzyme can further be engineered by mutagenesis to utilize a range of non-natural amino acids. Here we present structural data on the wild type enzyme in complex with adenylated ϵ -N-alkynyl-, ϵ -N-buteryl-, ϵ -N-crotonyl- and ϵ -N-propionyl-lysine providing insights into the plasticity of the PylRS active site. This shows that given certain key features in the non-natural amino acid to be incorporated, directed evolution of this enzyme is not necessary for substrate tolerance.

Citation: Flügel V, Vrabel M, Schneider S (2014) Structural Basis for the Site-Specific Incorporation of Lysine Derivatives into Proteins. PLoS ONE 9(4): e96198. doi:10.1371/journal.pone.0096198

Editor: Nediljko Budisa, Berlin Institute of Technology, Germany

Received: January 12, 2014; **Accepted:** April 3, 2014; **Published:** April 23, 2014

Copyright: © 2014 Flügel et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by the Excellence Cluster CIPSM, the German Research Foundation (DFG: SFB749) and the Fonds der chemischen Industrie (FCI). This work was supported by the DFG and the Technische Universität München within the funding programme Open Access Publishing. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: sabine.schneider@mytum.de

Introduction

The regulation of many cellular key processes, such as gene expression, protein activity and stability as well as molecular recognition relies on the posttranslational modification (PTM) of proteins. Lysine is the main target for PTMs particularly in the context of chromatin structure, remodeling and thus epigenetic gene regulation, known as the 'histone code' [1,2]. Histone acetyltransferases (HATs) and histone deacetylases (HDACs), which are enzymes responsible for lysine modifications, are regarded as important drug targets [3]. Recently it became clear that lysine residues in histones can not only be acetylated or methylated, but also malonylated, succinylated, butyrylated, propionylated and crotonylated [4,5,6,7,8,9]. These modifications change the net charge of the residue from positive to negative or neutral as well as alter the hydrophobicity and flexibility of the modified protein. The interaction properties, landscape and partners of the protein are therefore altered, ultimately impacting on chromatin structure and gene expression [2]. Despite the fact that some of the lysine modifications like butyrylation, crotonylation and propionylation are very similar in their chemical structure (Fig. 1), their effect on gene expression differ depending on the context. For instance, a direct link between activation of gene expression through crotonylation of Lys on histone H3 has been demonstrated

[6,10]. Additionally, various human diseases like cancer [11], disorders of the central nervous system [12], and autoimmune diseases [13] are associated with misregulation of histone PTMs. Lysine PTMs are not restricted to histones – for example, the tumor suppressor protein p53 is regulated by phosphorylation and ubiquitination at its C-terminal lysine residues [14,15]. In order to be able to understand the epigenetic regulation of cellular key processes, it is essential to elucidate the distinct temporal and special patterns of PTMs. Thus, lysine modifying enzymes, interaction partners and their regulatory proteins need to be identified. It is therefore crucial to generate the target protein containing site-specific modifications. In the last decade methods to genetically encode amino acids beyond the canonical 20 amino acids have been developed. Here orthogonal aminoacyl-tRNA synthetase (aaRS):tRNA^{AaCUA} pairs are used, such as the pyrrolysyl tRNA synthetase (PylRS) from *Methanosarcina mazei*, *Methanosarcina barkeri* and tyrosyl-tRNA synthetase (TyrRS) from *Methanococcus jannaschii*, which incorporate pyrrolysine (Pyl) and tyrosine. A number of studies show that aaRS can be further evolved to accept non-natural, functionalized or chemical modifiable amino acids. By introducing an amber stop codon in a given gene, the host's endogenous translational machinery can be employed to incorporate non-natural amino acids site-specifically, thus allowing the functionalization of the target protein

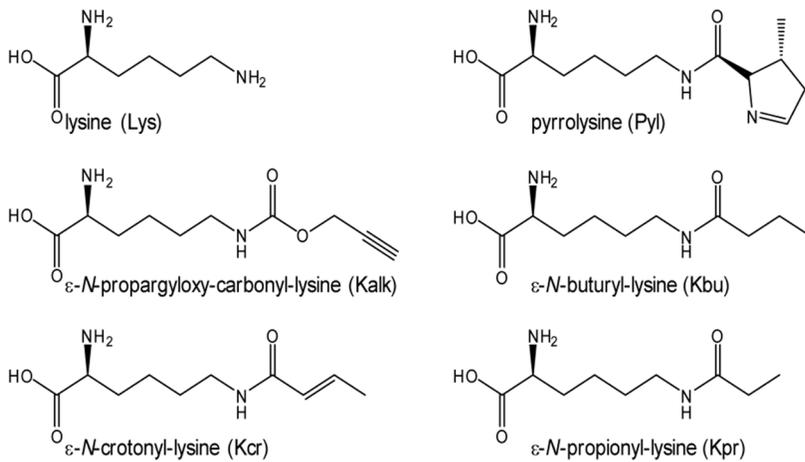


Figure 1. Lysine, pyrrolysine and analogs used in this study.
doi:10.1371/journal.pone.0096198.g001

[16,17,18,19,20,21,22,23,24]. Structures of PylRS in complex with its natural substrate [25,26,27] as well as evolved mutants containing non-natural amino acids are already available

[28,29,30]. Using an evolved PylRS (L274A C313A Y349F) from *M. barkeri* ϵ -*N*-crotonyl-lysine was introduced into histones [17]. However the wild type PylRS from *M. mazei*, which shares

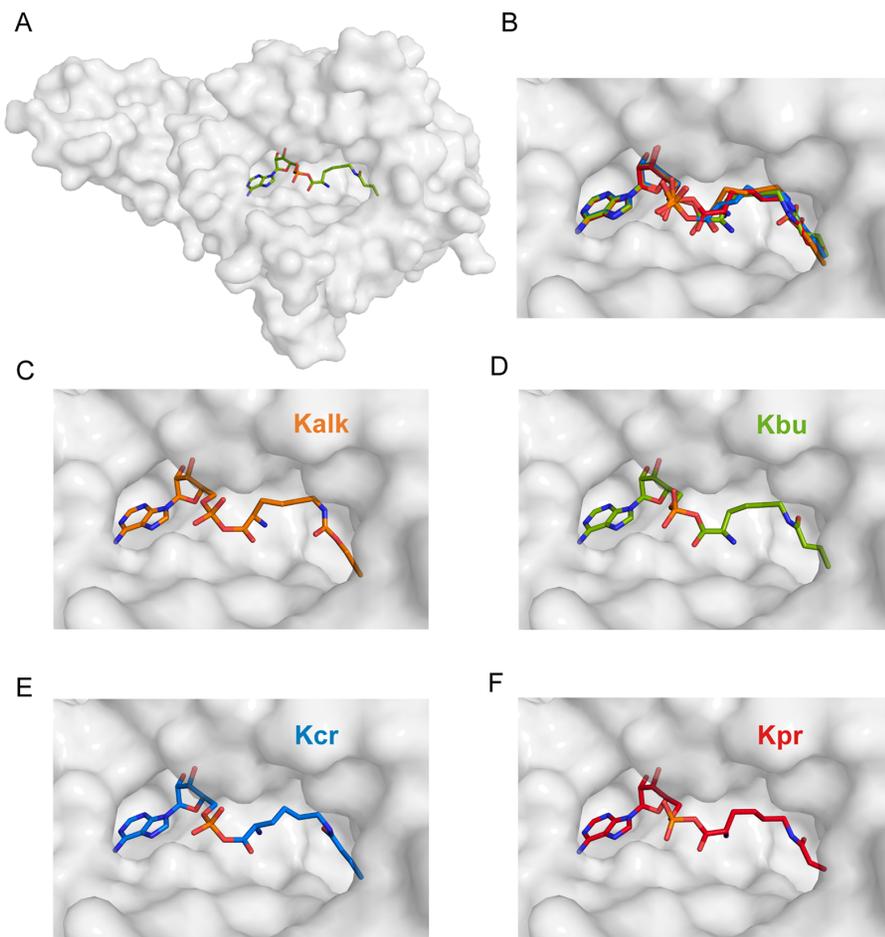


Figure 2. Cartoon representation of the overall structure of the catalytic domain of PylRS. (A) Type-II tRNA-synthetase folding topology of the tRNA synthetase domain from PylRS, shown as cartoon representation, overlaid with its semi-transparent surface. The adenylated Kbu is highlighted as green stick model in the active site. (B-F) Zoom in the active site: PylRS in complex with Kalk (orange), Kbu (green), Kcr (blue) and Kpr (red), drawn as stick models. (r. m. s. d ~ 0.35 Å). All four non-natural amino acids bind in the same position.
doi:10.1371/journal.pone.0096198.g002

71% sequence identity (85% sequence similarity) with the homologue from *M. barkeri*, readily utilizes ϵ -*N*-butyryl-(Kbu), Kcr and ϵ -*N*-propionyl-lysine (Kpr) as substrates (Fig. 1) and can therefore be used directly to incorporate these different PTMs into histones [31]. In addition, ϵ -*N*-propargyloxy-carbonyl-lysine (Kalk) is also accepted, which allows functionalization after heterologous expression using Cu(I)-catalyzed azide-alkyne cycloaddition [32,33,34,35,36,37]. Here we present the X-ray crystal structures of the PylRS from *M. mazei* in complex with adenylated Kalk, Kbu, Kcr and Kpr (Fig. 1). The data show that the amino acids are held in the PylRS active site such that variations at the ϵ -*N* position are allowed and key features can be deduced. The factors limiting the derivatives that can be accommodated by the wild type PylRS are length, flexibility and charge properties: The binding pocket is predominantly lined by hydrophobic residues and analogues which are extended compared to Pyl, but flexible can 'curl up' in order to fit in the active site. Taken together this provides insights into the plasticity of the active site of this enzyme and shows that non-natural amino acids with certain key features can be readily incorporated. Thus PylRS evolution to utilize such non-natural amino acids for the incorporation in a given target protein is not strictly necessary.

Results and Discussion

Structures of lysine derivatives bound by the catalytic domain of PylRS

The crystal structures of PylRS in complex with adenylated Kalk, Kbu, Kcr and Kpr (Fig. 1) were solved up to 2.1 Å resolution (Table 1). All four adenylated amino acids occupy the

same space in the active site of the aaRS domain (Fig. 2). While the adenosine and phosphate moieties superimpose almost perfectly, the amino group can take up alternative conformations by rotating 180° around the C-C α bond, as seen in the complex structures of Kalk, Kcr and Kpr. This was previously observed in the structure of the engineered enzyme (PylRS Y306A Y384F) binding ϵ -*N*-o-azidobenzoyloxycarbonyl lysine (PDB code 2ZIO) [30]. In the Wt-Pyl complex the α -amino group of Pyl interacts with the hydroxyl group of Tyr 384 which is located in the β 7- β 8 hairpin [27]. However, the Pyl analogs lack the ability to form a comparable H-bond and hence this loop is flexible and not defined in the electron density. This flexibility is independent of the bound substrate and it was postulated that this hairpin protects the unstable pyrrolysyladenylate intermediate [25,26]. In addition, mutation of Tyr 384 to Phe showed that this H-bond interaction between the loop and the substrate is not essential for enzyme function, and its absence may even enhance utilization of non-natural substrates [16,30]. In the vicinity of the α -phosphate some positive difference density is visible which could correspond to a partly occupied pyrophosphate, reminiscent from the adenylation reaction. Only in the Kpr-complex structure the observed density warranted modeling of the pyrophosphate at 60% occupancy. Compared to the pyrroline ring in Pyl the alkyne, butyryl, crotonyl or propionyl functional groups in Kalk, Kbu, Kcr and Kpr are smaller. Thus more space and a higher degree of flexibility is possible and not all atoms of the methylene spacer carrying the functional group are fully defined in the electron density (Fig. 3).

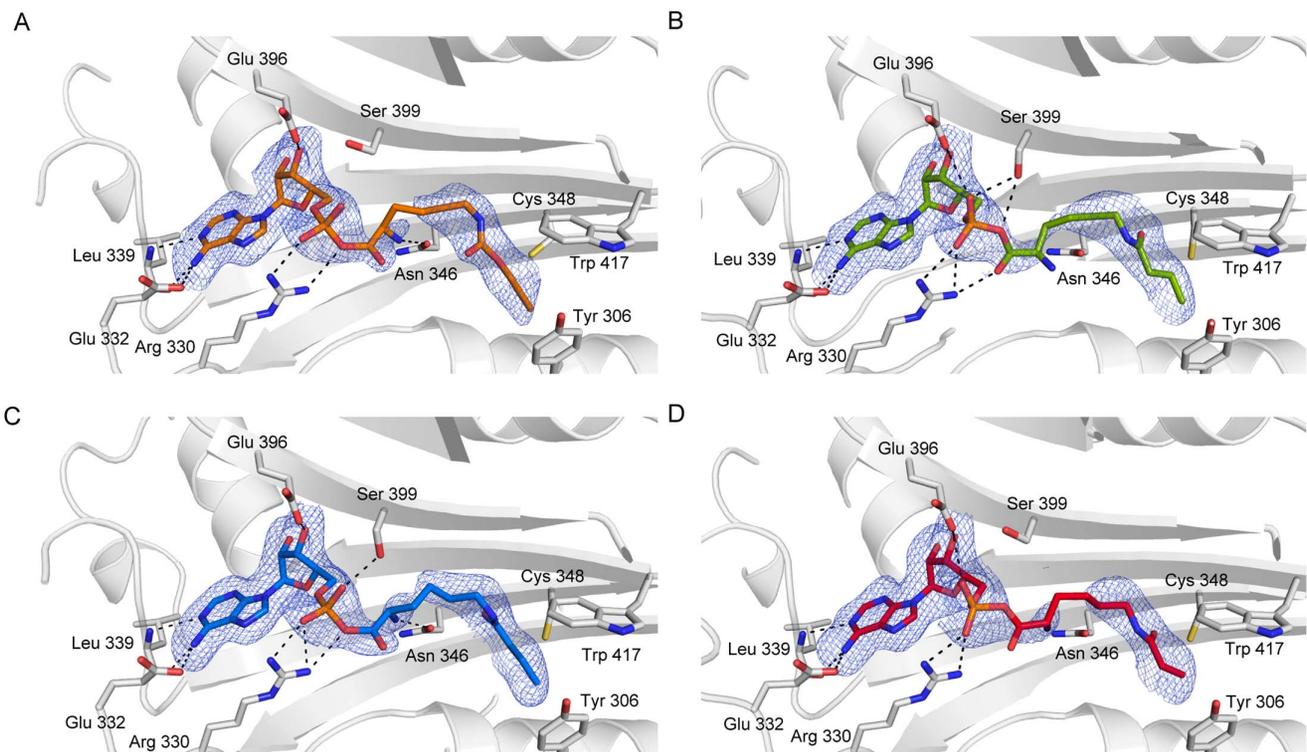


Figure 3. Simulated annealing-omit Fo-DFc electron density contoured at 2 σ of PylRS in complex with (A) Kalk (orange), (B) Kbu (green), (C) Kcr (blue) and (D) Kpr (red). The protein is shown as cartoon, overlaid with its semi-transparent surface representation. Amino acids providing key interactions are drawn as sticks, hydrogen-bonds as dashed lines.
doi:10.1371/journal.pone.0096198.g003

Table 1. Data collection, processing and structure refinement statistics.

	Kalk	Kbu	Kcr	Kpr
	(PDB ID 4CH6)	(PDB ID 4CH3)	(PDB ID 4CH4)	(PDB ID 4CH5)
Wavelength (Å)	1.0	1.0	1.0	1.0
Resolution range (Å)	38.3–2.05 (2.12–2.05)	38.4–2.28 (2.36–2.28)	42.4–2.16 (2.24–2.16)	42.3–2.2 (2.28–2.20)
Space group	P6 ₄	P6 ₄	P6 ₄	P6 ₄
Unit cell	104.9 104.9 71.0	105.3 105.3 71.4	105.0 105.0 71.8	105.4 105.4 70.8
Total reflections	271,274 (18,606)	123,272 (12,794)	247,942 (23,845)	228,986 (21,051)
Unique reflections	28,018 (2,789)	20,634 (2,045)	24,308 (2,439)	22,651 (2,185)
Multiplicity	9.7 (6.7)	6.0 (6.3)	10.2 (9.8)	10.1 (9.6)
Completeness (%)	100.0 (99.8)	99.9 (100.0)	100.0 (100.0)	99.7 (96.9)
Mean I/sigma(I)	45.7 (3.3)	24.7 (3.9)	27.2 (4.5)	23.8 (3.0)
Wilson B-factor	43.9	44.3	36.4	43
R-merge	0.027 (0.593)	0.044 (0.450)	0.0611 (0.534)	0.0622 (0.699)
R-meas	0,029	0,048	0,064	0,066
CC1/2	1 (0.877)	1 (0.889)	0.999 (0.923)	0.999 (0.837)
CC*	1 (0.967)	1 (0.97)	1 (0.98)	1 (0.955)
R-work	0.174 (0.271)	0.176 (0.231)	0.166 (0.330)	0.174 (0.225)
R-free	0.186 (0.296)	0.197 (0.245)	0.209 (0.349)	0.190 (0.257)
Number of atoms	2,296	2,252	2,310	2,276
macromolecules	2,158	2,151	2,120	2,157
ligands	21	21	21	30
water	117	80	169	89
Protein residues	264	268	259	264
RMS(bonds)	0.011	0.01	0.01	0.01
RMS(angles)	1.51	1.38	1.4	1.44
Ramachandran favored (%)	97	96	97	98
Ramachandran outliers (%)	0	0	0	0.4
Clashscore	0.23	1.17	0.71	1.62
Average B-factor	53.1	49.6	40.1	47.8
macromolecules	52.9	49.4	39.3	47.6
ligands	58.5	60.3	50.2	58.2
solvent	56.0	50.1	49.6	50.1

Numbers in parentheses correspond to the high resolution shell.
doi:10.1371/journal.pone.0096198.t001

Kalk, Kbu, Kcr and Kpr possess key features required for recognition by PylRS

Both PylRS and lysyl-tRNA-synthetase (LysRS) belong to the type II family of aaRS and thus share the same overall folding topology. A structure of the LysRS from *Bacillus stearothermophilus* (PDB code 3A74), *Bulkholderia thailandensis* (PDB code 4EX5 [38]) and *Escherichia coli* (PDB code 1E22 [39]) are available. The three LysRS share about 53% sequence identity, the catalytic domains superimpose with an r.m.s.d. of 1.2 Å and the residues lining the active site are highly conserved. In comparison, the sequence identity to the catalytic domain of PylRS from *M. mazei* are 17%, 19% and 20%, respectively, and superimpose with an r. m. s. d. of about 2.1 Å. The ATP binding site in both PylRS and LysRS are comparable. The LysRS employs an elaborate interaction network with the Lys and a tight binding pocket (Fig. 4A+D). In contrast, PylRS provides only some key interactions with Pyl in a more spacious, active site lined with hydrophobic residues (Fig. 4B+E). The alkyne group of Kalk,

the least flexible of the used derivatives, protrudes slightly deeper into the pocket than Pyl but lies in the same plane as the pyrroline ring. Thus favorable $\pi - \pi$ interactions between the amino acid side chain with Tyr 306 as well Trp 417 are possible. (Fig. 4E) PylRS can therefore accommodate a wider range of substrates, if they possess certain key prerequisites: If present, the carbamate, carbonyl or amide moiety, can either interact with Asn 346 or Cys 348, even if this interaction is not permanent, as seen in Pyl [26,27,40] and a norbornene containing Pyl analog [28]. For Pyl, an additional hydrogen bond accepting imine nitrogen in the pyrroline ring further increases the activation efficiency [41]. Our data show that analogs containing only the peptide-bond motif, can efficiently bind to the active site of the wild type PylRS. Thus Kalk, Kbu, Kcr, and Kpr are recognized as substrates and fit well in the pocket (Fig. 4F). The preferred substrate head group is

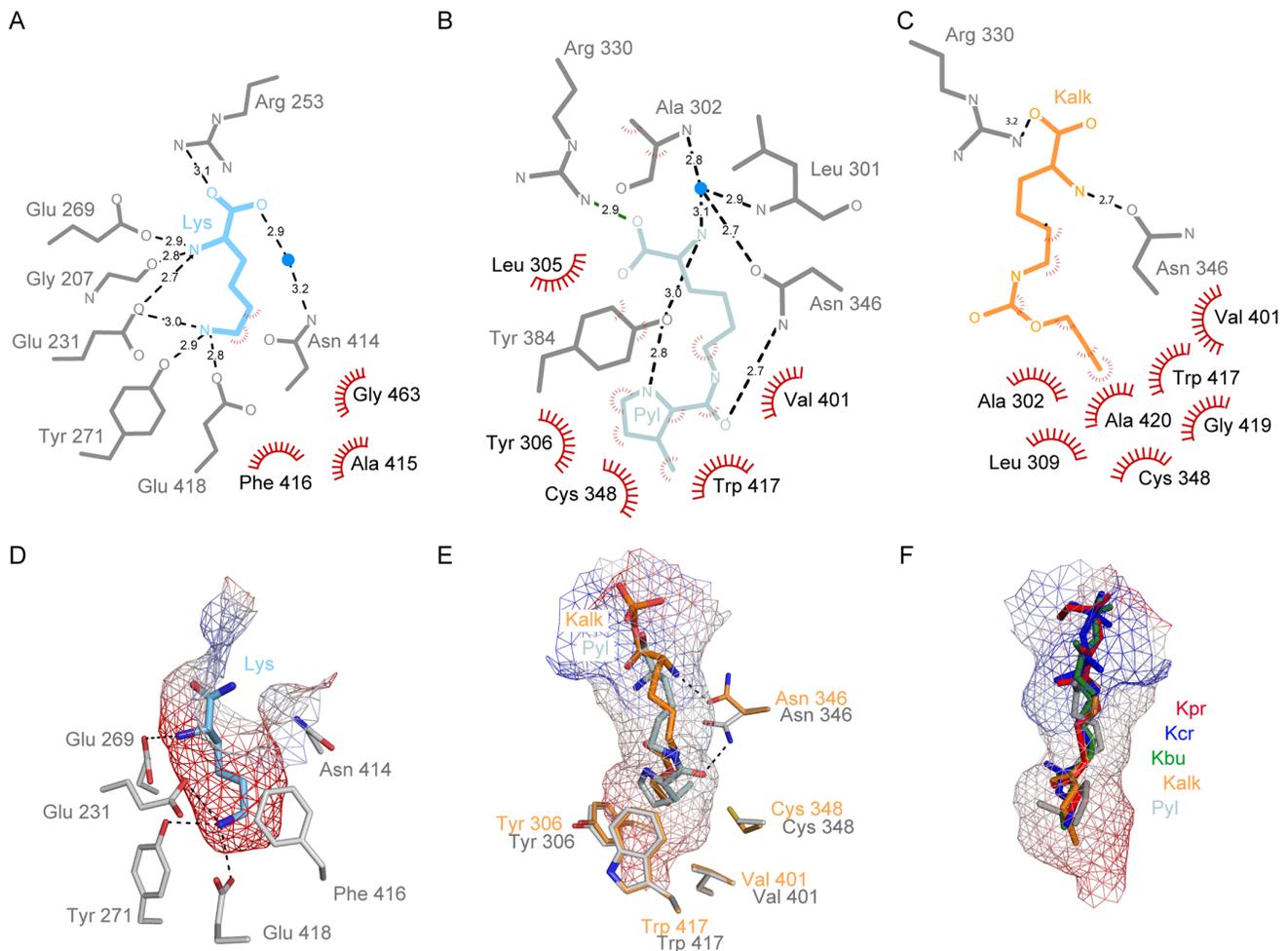


Figure 4. Comparison of the binding pockets of PylRS and LysRS. Two-dimensional plot [52] of residues interacting with Lys (A), Pyl (B) and Kalk (C). Van-der-Waals contacts are shown as red half-spheres. Surface representation of the binding pockets of LysRS (D) and PylRS (E) with Lys, Pyl and Kalk, respectively. (F) Superposition of Pyl (grey), Kalk (orange), Kbu (green), Kcr (blue) and Kpr (red). The surface of the binding pocket is shown as mesh, with the charge distribution indicated by coloring (red = negative, blue = positive). The surrounding residues are drawn as sticks. (PDB codes PylRS: 2Q7H and LysRS: 3A74). doi:10.1371/journal.pone.0096198.g004

hydrophobic and up to four atom bonds in length can be accepted in the binding pocket. Additionally, branched functional groups like a *tert*-butyloxycarbonyl group [30] or small ring systems, such as cyclopentane can be accommodated in the spacious pocket. However, re-engineering of the enzyme active site would be required to accommodate non-natural amino acids lacking these key features. The hot spots which were targeted so far by directed evolution include amino acids whose side chains are either lining the binding pocket (Tyr 306, Leu 309, Asn 346, Cys 348, Trp 417) or resides in the active site closing loop (Ile 417, Phe 384). (Corresponding residues in *M. bakeri*: Leu 274, Asn 311, Cys 313 and Tyr 349) [17,29,30,32]. For instance in order to efficiently use larger and more bulky Pyl-analogues like norbornene or ϵ -*N*-benzyloxycarbonyl-L-lysine, more room is needed. This can be made available by replacing Tyr 306 at the bottom of the binding pocket by a Gly or Ala [28,30]. Structural data on PylRS mutants show that even replacing this large amino acid appears not to perturb the overall structural architecture.

Conclusions

In the last few years numerous PTMs of lysine modifications beyond acetylation and methylation were identified. To understand their function in the epigenetic regulation of biological key processes, the lysine modifying enzymes and their recognition partners need to be identified and their interplay characterized. To address this, an important tool is the site-specific incorporation of these modifications into histones, using orthogonal PylRS:tRNA^{Aaa}CUA pairs. Here we elucidate the structural basis for the activation of the lysine derivatives Kalk, Kbu, Kcr and Kpr by the wild-type PylRS. This provides further insights into the required features of non-natural amino acids to be utilized by PylRS. Depending on the non-natural amino acid to be incorporated, engineering of PylRS is not necessarily required.

Material and Methods

Protein expression, purification and crystallization

Kalk, Kbu, Kcr and Kpr were synthesized as published by Gattner et al and Kaya et al [31,32]. The sequence encoding

the catalytic domain (residues 185–454) of the PylRS from *M. mazei* was PCR amplified and cloned into pET28a, expressed in *E. coli* Rosetta DE3 cells (Novagen) and purified as described previously [27,28]. In order to co-crystallise PylRS with the Kalk, Kbu, Kcr and Kpr amino acids, the protein was diluted (1 mg mL^{-1}) and incubated for 2 h with the respect amino acid (2 mM) and ATP (1 mM; Sigma–Aldrich) in protein storage buffer [HEPES (10 mM, pH 7.4), NaCl (300 mM), MgCl_2 (5 mM), dithiothreitol (1 mM)]. The protein was concentrated to 10 mg mL^{-1} prior to crystallisation. Crystals appeared overnight in lithium acetate (100 mM) and PEG3350 (10–14 (w/v) %). All crystals were cryoprotected with well solution supplemented with ethylene glycol (30% w/v) before flash-freezing, and then stored in liquid nitrogen until data collection.

Data collection and structure determination

Diffraction data were collected at the synchrotron beam lines PXI and PXIII (Swiss Light Source, Villigen, Switzerland). The data were processed with XDS [42] to 2.05 Å (Kalk), 2.28 Å (Kbu), 2.15 Å (Kcr) and 2.2 Å (Kpr) spacing, respectively, ensuring consistent indexing and choosing the same set of free reflections. The Crystals belong to the same space group as reported previously for the wild type PylRS [27] (P6₄, unit cell dimensions: $a = b = 105 \text{ Å}$, $c = 71 \text{ Å}$). The structure was solved by molecular replacement using the PylRS coordinates (PDB code: 4BW9) in PHASER [43,44]. For the Kalk, Kcr and Kpr complexes the coordinates of the protein atoms of the PylRS-adenylated Kbu complex were used in rigid body refinement in REFMAC [45]. In order to reduce model bias, all non-protein atoms as well as the loop region around Phe 384 were removed

from the model prior to molecular replacement/rigid body refinement, and the temperature factors were reset, followed by simulated annealing in PHENIX [46]. Clear peaks for AMP-Kalk, AMP-Kbu, AMP-Kcr and AMP-Kpr were visible in the simulated-annealing omit Fo-DFc map. Rounds of model building and refinement were carried out in COOT [47] and REFMAC. The refinement parameter file for the adenylated amino acids was generated with prodr2 [48], as implemented in COOT, and the TLSMD [49] server was used to determine TLS groups for TLS refinement [50]. Diffraction data and refinement statistics are summarized in Table 1. Structural superposition were done with SSM [51]. Two-dimensional interaction plots were carried out with LIGPLOT [52]. All structural figures were prepared with PyMol (Delano Scientific, San Carlos, CA). Atomic coordinates were submitted to the Protein Data Bank (<http://www.ebi.ac.uk/pdbe/>) with the PDB codes: 4CH6 (PylRS-Kalk), 4CH3 (PylRS-Kbu), 4CH4 (PylRS-Kcr), 4CH5 (PylRS-Kpr)

Acknowledgments

We thank the beam line scientists at the Swiss Light Source for setting up the beam lines for data collection. We also thank Michael Groll and Thomas Carell for critical reading of the manuscript and helpful discussions.

Author Contributions

Conceived and designed the experiments: VF SS. Performed the experiments: VF MV SS. Analyzed the data: VF SS. Contributed reagents/materials/analysis tools: VF MV SS. Wrote the paper: VF SS.

References

- Khorasanizadeh S (2004) The nucleosome: from genomic organization to genomic regulation. *Cell* 116: 259–272.
- Tropberger P, Schneider R (2013) Scratching the (lateral) surface of chromatin regulation by histone modifications. *Nat Struct Mol Biol* 20: 657–661.
- Cole PA (2008) Chemical probes for histone-modifying enzymes. *Nature chemical biology* 4: 590–597.
- Chen Y, Sprung R, Tang Y, Ball H, Sangras B, et al. (2007) Lysine propionylation and butyrylation are novel post-translational modifications in histones. *Mol Cell Proteomics* 6: 812–819.
- Garrity J, Gardner JG, Hawse W, Wolberger C, Escalante-Semerena JC (2007) N-lysine propionylation controls the activity of propionyl-CoA synthetase. *J Biol Chem* 282: 30239–30245.
- Tan M, Luo H, Lee S, Jin F, Yang JS, et al. (2011) Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell* 146: 1016–1028.
- Xie Z, Dai J, Dai L, Tan M, Cheng Z, et al. (2012) Lysine succinylation and lysine malonylation in histones. *Mol Cell Proteomics* 11: 100–107.
- Zhang Z, Tan M, Xie Z, Dai L, Chen Y, et al. (2011) Identification of lysine succinylation as a new post-translational modification. *Nature chemical biology* 7: 58–63.
- Arnaudo AM, Garcia BA (2013) Proteomic characterization of novel histone post-translational modifications. *Epigenetics Chromatin* 6: 24.
- Montellier E, Rousseaux S, Zhao Y, Khochin S (2012) Histone crotonylation specifically marks the haploid male germ cell gene expression program: post-meiotic male-specific gene expression. *Bioessays* 34: 187–193.
- Li GM (2013) Decoding the Histone Code: Role of H3K36me3 in Mismatch Repair and Implications for Cancer Susceptibility and Therapy. *Cancer Res* 73: 6379–6383.
- Pena-Altamira LE, Polazzi E, Monti B (2013) Histone post-translational modifications in Huntington's and Parkinson's diseases. *Curr Pharm Des* 19: 5085–5092.
- Quintero-Ronderos P, Montoya-Ortiz G (2012) Epigenetics and autoimmune diseases. *Autoimmune Dis* 2012: 593720.
- Kruse JP, Gu W (2009) Modes of p53 regulation. *Cell* 137: 609–622.
- Du J, Zhou Y, Su X, Yu JJ, Khan S, et al. (2011) Sirt5 is a NAD-dependent protein lysine demalonylase and desuccinylase. *Science* 334: 806–809.
- Kaya E, Vrabel M, Deiml C, Prill S, Fluxa VS, et al. (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: 4466–4469.
- Kim CH, Kang M, Kim HJ, Chatterjee A, Schultz PG (2012) Site-specific incorporation of epsilon-N-crotonyllysine into histones. *Angewandte Chemie* 51: 7246–7249.
- Wang Z, Rizzo CJ (2001) Synthesis of the C8-deoxyguanosine adduct of the food mutagen IQ. *Org Lett* 3: 565–568.
- Wang YS, Fang X, Wallace AL, Wu B, Liu WR (2012) A rationally designed pyrrolysyl-tRNA synthetase mutant with a broad substrate spectrum. *Journal of the American Chemical Society* 134: 2950–2953.
- Wang YS, Russell WK, Wang Z, Wan W, Dodd LE, et al. (2011) The de novo engineering of pyrrolysyl-tRNA synthetase for genetic incorporation of L-phenylalanine and its derivatives. *Molecular bioSystems* 7: 714–717.
- Zhang Y, Wang L, Schultz PG, Wilson IA (2005) Crystal structures of apo wild-type *M. jannaschii* tyrosyl-tRNA synthetase (TyrRS) and an engineered TyrRS specific for O-methyl-L-tyrosine. *Protein Sci* 14: 1340–1349.
- Santoro SW, Wang L, Herberich B, King DS, Schultz PG (2002) An efficient system for the evolution of aminoacyl-tRNA synthetase specificity. *Nat Biotechnol* 20: 1044–1048.
- Chen PR, Groff D, Guo J, Ou W, Cellitti S, et al. (2009) A facile system for encoding unnatural amino acids in mammalian cells. *Angewandte Chemie* 48: 4052–4055.
- Perona JJ, Hadd A (2012) Structural diversity and protein engineering of the aminoacyl-tRNA synthetases. *Biochemistry* 51: 8705–8729.
- Yanagisawa T, Ishii R, Fukunaga R, Kobayashi T, Sakamoto K, et al. (2008) Crystallographic studies on multiple conformational states of active-site loops in pyrrolysyl-tRNA synthetase. *J Mol Biol* 378: 634–652.
- Yanagisawa T, Sumida T, Ishii R, Yokoyama S (2013) A novel crystal form of pyrrolysyl-tRNA synthetase reveals the pre- and post-aminoacyl-tRNA synthesis conformational states of the adenylate and aminoacyl moieties and an asparagine residue in the catalytic site. *Acta Crystallogr D Biol Crystallogr* 69: 5–15.
- Kavran JM, Gundllapalli S, O'Donoghue P, Englert M, Soll D, et al. (2007) Structure of pyrrolysyl-tRNA synthetase, an archaeal enzyme for genetic code innovation. *Proc Natl Acad Sci U S A* 104: 11268–11273.
- Schneider S, Gattner MJ, Vrabel M, Flugel V, Lopez-Carrillo V, et al. (2013) Structural Insights into Incorporation of Norbornene Amino Acids for Click Modification of Proteins. *ChemBiochem: a European journal of chemical biology*.
- Takimoto JK, Dellas N, Noel JP, Wang L (2011) Stereochemical basis for engineered pyrrolysyl-tRNA synthetase and the efficient in vivo incorporation of structurally divergent non-native amino acids. *ACS Chem Biol* 6: 733–743.

30. Yanagisawa T, Ishii R, Fukunaga R, Kobayashi T, Sakamoto K, et al. (2008) Multistep engineering of pyrrolysyl-tRNA synthetase to genetically encode N(epsilon)-(o-azidobenzyloxycarbonyl) lysine for site-specific protein modification. *Chem Biol* 15: 1187–1197.
31. Gattner MJ, Vrabel M, Carell T (2013) Synthesis of epsilon-N-propionyl-, epsilon-N-butyl-, and epsilon-N-crotonyl-lysine containing histone H3 using the pyrrolysine system. *Chem Commun (Camb)* 49: 379–381.
32. Kaya E, Gutschiedl K, Vrabel M, Muller M, Thumbs P, et al. (2009) Synthesis of threefold glycosylated proteins using click chemistry and genetically encoded unnatural amino acids. *Chembiochem: a European journal of chemical biology* 10: 2858–2861.
33. Nguyen DP, Lusic H, Neumann H, Kapadnis PB, Deiters A, et al. (2009) Genetic encoding and labeling of aliphatic azides and alkynes in recombinant proteins via a pyrrolysyl-tRNA Synthetase/tRNA(CUA) pair and click chemistry. *Journal of the American Chemical Society* 131: 8720–8721.
34. Rostovtsev VV, Green LG, Fokin VV, Sharpless KB (2002) A stepwise huisgen cycloaddition process: copper(I)-catalyzed regioselective “ligation” of azides and terminal alkynes. *Angewandte Chemie* 41: 2596–2599.
35. Tornøe CW, Christensen C, Meldal M (2002) Peptidotriazoles on solid phase: [1,2,3]-triazoles by regioselective copper(I)-catalyzed 1,3-dipolar cycloadditions of terminal alkynes to azides. *J Org Chem* 67: 3057–3064.
36. Milles S, Tyagi S, Banterle N, Koehler C, VanDelinder V, et al. (2012) Click strategies for single-molecule protein fluorescence. *Journal of the American Chemical Society* 134: 5187–5195.
37. Wan W, Huang Y, Wang Z, Russell WK, Pai PJ, et al. (2010) A facile system for genetic incorporation of two different noncanonical amino acids into one protein in *Escherichia coli*. *Angewandte Chemie* 49: 3211–3214.
38. Baugh L, Gallagher LA, Patrapuvich R, Clifton MC, Gardberg AS, et al. (2013) Combining functional and structural genomics to sample the essential *Burkholderia* struotome. *PLoS One* 8: e53851.
39. Desogus G, Todone F, Brick P, Onesti S (2000) Active site of lysyl-tRNA synthetase: structural studies of the adenylation reaction. *Biochemistry* 39: 8418–8425.
40. Yanagisawa T, Ishii R, Fukunaga R, Nureki O, Yokoyama S (2006) Crystallization and preliminary X-ray crystallographic analysis of the catalytic domain of pyrrolysyl-tRNA synthetase from the methanogenic archaeon *Methanosarcina mazei*. *Acta Cryst F* 62: 1031–1033.
41. Li WT, Mahapatra A, Longstaff DG, Bechtel J, Zhao G, et al. (2009) Specificity of pyrrolysyl-tRNA synthetase for pyrrolysine and pyrrolysine analogs. *Journal of molecular biology* 385: 1156–1164.
42. Kabsch W (2010) Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr D Biol Crystallogr* 66: 133–144.
43. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, et al. (2007) Phaser crystallographic software. *J Appl Crystallogr* 40: 658–674.
44. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, et al. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 67: 235–242.
45. Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, et al. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* 67: 355–367.
46. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, et al. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66: 213–221.
47. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66: 486–501.
48. Schuttelkopf AW, van Aalten DM (2004) PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr D Biol Crystallogr* 60: 1355–1363.
49. Painter J, Merritt EA (2006) TLSMD web server for the generation of multi-group TLS models. *J Appl Crystallogr* 39: 109–111.
50. Winn MD, Murshudov GN, Papiz MZ (2003) Macromolecular TLS refinement in REFMAC at moderate resolutions. *Methods Enzymol* 374: 300–321.
51. Krissinel E, Henrick K (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr* 60: 2256–2268.
52. Wallace AC, Laskowski RA, Thornton JM (1995) LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng* 8: 127–134.