bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

SUV39 SET domains mediate crosstalk of heterochromatic histone marks

Alessandro Stirpe[1,2], Nora Guidotti[3], Sarah Northall[2], Sinan Kilic[3,4], Alexandre Hainard[5], Oscar Vadas[6,7], Beat Fierz[3], Thomas Schalch[1,2%]

[1] Department of Molecular Biology, Faculty of Science, Science III, University of Geneva, CH-1211 Geneva 4, Switzerland.
[2] Current address: Leicester Institute for Structural and Chemical Biology, Department of Molecular and Cell Biology, University of Leicester, Leicester, LE1 9HN, UK.
[3] Laboratory of Biophysical Chemistry of Macromolecules, Institute of Chemical Sciences and Engineering (ISIC), Ecole Polytechnique Fédérale de Lausanne (EPFL) , 1015 Lausanne , Switzerland.
[4] Current address: Novo Nordisk Foundation Center for Protein Research, University of Copenhagen
[5] Proteomics Core Facility, Faculty of Medicine, University Medical Center, University of Geneva, Geneva, Switzerland.
[6] School of Pharmaceutical Sciences, University of Geneva, 1211 Geneva, Switzerland.
[7] Current address: Protein Core Facility, Faculty of Medicine, University of Geneva, 1211 Geneva, Switzerland

% Corresponding author:

Email: thomas.schalch@le.ac.uk, Tel: +44 116 229 7120

Short Title: Regulation of SET domains by H3K14ub

Keywords: heterochromatin, histone methyltransferase, ubiquitin, posttranslational modifications, histones, allosteric activation, protein complex, *S.pombe*

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Abstract

The SUV39 class of methyltransferase enzymes deposits histone lysine 9 di- and tri-methylation (H3K9me2/3), the epigenetic hallmark of constitutive heterochromatin, which serves as the central recruitment platform for the heterochromatic silencing machinery. How these enzymes are regulated to mark specific genomic regions as heterochromatic is not well understood. Clr4 is the sole H3K9me2/3 methyltransferase in the fission yeast *S.pombe* and recent evidence suggests that ubiquitination of lysine 14 on H3 tail (H3K14) plays a key role in H3K9 methylation. However, the molecular mechanism of this regulation and its role in heterochromatin formation remains to be determined. Here we present a structure-function approach to understanding how the H3K14ub mark stimulates Clr4 activity *in cis*. These results show that the H3K14ub substrate binds specifically and tightly to the catalytic domain of Clr4, and thereby activates the enzyme by 250-fold. Mutations that disrupt this mechanism lead to a loss of H3K9me2/3 and abolish heterochromatin silencing similar to a *clr4* deletion. Our work reveals a sensor for H3K14 ubiquitylation in the SET domain of Clr4, which mediates the licensing of heterochromatin formation by an epigenetic cross-talk. This sensor is also active in the human SUV39H2 enzyme.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Introduction

Histone methylation plays a critical role in regulating and organizing genomes to maintain genome integrity and establish appropriate gene regulation programs [1,2]. Histone H3 lysine 9 di- and trimethylation (H3K9me2/3) are highly conserved hallmarks of heterochromatin, where they provide a platform for the recruitment of chromatin effector complexes. Dysregulation of H3K9me2/3 marks has been implicated in various cancers, in neurological diseases and viral latency and are targeted for therapeutic approaches [3].

The SUV39 clade of the SET-domain family of protein lysine methyltransferases represents the enzyme class depositing H3K9me2/3 [4]. They share a SET domain as a catalytic core, which builds on a series of curved β-sheets that are sandwiched between distinct pre- and post-SET domains, providing structural elements based on zinc-finger motifs. The SUV39 enzymatic activity facilitates the transfer of a methyl group from S-adenosyl-L-methionine (SAM) to the ε-amino group of lysine 9 on histone H3 with high specificity, thereby converting the cofactor to S-adenosyl-L-homocysteine (SAH).

The mechanisms regulating SUV39 proteins to specifically deposit H3K9me2/3 in heterochromatic regions are not well understood. The fission yeast *Schizosaccharomyces pombe* heterochromatin system, closely related to the metazoan systems, with its sole SUV39 protein Clr4 serve as a paradigm for studying the role of this enzyme class in heterochromatin establishment and maintenance. Ectopic tethering experiments with Clr4 have established that H3K9me2/3 can support stable epigenetic transmission of transcriptional states [5,6]. Clr4 has further served as a valuable tool to understand its animal orthologs, which share a very similar domain architecture consisting of an N-terminal chromodomain and a C-terminal catalytic SET domain (Fig

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

1A). The chromodomain binds the H3K9me2/3 mark and is critical for spreading the H3K9me2/3 mark across genomic domains (Melcher et al., 2000; Müller et al., 2016; Zhang et al., 2008). Recent evidence also suggests that an autoinhibitory loop in the post-SET domain regulates spreading of H3K9me2/3 by Clr4 (Iglesias et al., 2018).

Clr4 is a key part of *S. pombe* heterochromatin, which is found at pericentromeres, telomeres and the mating type locus and depends on the tightly interconnected pathways of the nuclear RNA interference machinery and an array of chromatin modifiers [7]. Small RNAs are generated from a limited set of repeat sequences in heterochromatic regions and guide the RNAi machinery in the form of the RNA-induced initiation of transcriptional gene silencing (RITS) to nascent heterochromatic transcripts [8–11]. RITS is an essential pathway for the recruitment of Clr4 and its associated Cullin4-RING ubiquitin ligase complex, the CLRC complex [12,13]. Recruitment of Clr4 underpins the deposition of H3K9me2/3 marks, providing a platform for binding of HP1 proteins, which in turn recruit further transcriptional gene silencing complexes [14–16]. Heterochromatin formation is further strongly dependent on deacetylation of H3K14 [17–20], with H3K14 acetylation being strongly associated with active gene expression in euchromatin [21].

The CLRC complex plays an essential role in the *S.pombe* heterochromatin system, but the mechanistic understanding of CLRC's action and its ubiquitination target in heterochromatin formation have remained elusive [12,22–24]. In CLRC, Clr4 associates with the cullin protein Cul4, the WD-40 β-propeller proteins Dos1 and Rik1, the RING finger protein Pip1, and Dos2 [25–29]. CLRC features the hallmarks of an intact CRL4-type E3 ligase, which uses the subunit Dos1 as a substrate adapter [26,27,30,31]. Recent evidence suggests that the preferred substrate for ubiquitylation by CLRC might be

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

lysine 14 on histone H3, yielding H3K14ub [32]. The latter work suggests that CLRC deposits H3K14ub and thereby controls the activity of Clr4, ultimately contributing to heterochromatin formation.

We combined biochemical and structural methods to decipher the molecular mechanism governing the stimulation of Clr4 by H3K14ub. These experiments identified a surface patch on the catalytic domain of Clr4, which mediates specific H3K14ub binding, and mutations that disrupt this interaction interface lose heterochromatin function. This reveals a critical regulatory mechanism that uses the SET domain of Clr4 for control of heterochromatin formation by an epigenetic crosstalk.

Results

The catalytic domain of Clr4 senses the presence of H3K14-linked ubiquitin

To understand the enzymatic reaction underlying the stimulation of Clr4 by H3K14ub, we set out to determine the domains of Clr4 that mediate this effect. Clr4's domain architecture comprises an N-terminal chromodomain, which is connected by a linker region to the catalytic lysine methyltransferase (KMT) domain (Fig. 1A). The KMT domain contains the N-terminal, pre-SET, SET and post-SET regions and its structure has been determined [33,34]. To determine the role of the chromodomain and linker region we recombinantly expressed both the full length Clr4 and the isolated KMT domain and quantified their methyltransferase activity using two different assays: (1) Incorporation of tritium-labelled methyl groups from a ^{3}H-SAM donor into histone proteins and (2) fluorescence-based measurement of SAH generation (Cisbio EPIgeneous Methyltransferase assay kit). As substrates we used unmodified H3

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

peptide (1-21) and H3 peptide ubiquitinated on K14 (H3K14ub) or K18 (H3K18ub) generated by native chemical ligation procedures. These branched peptides are identical to a native ubiquitin linkage except for a glycine to alanine change in the C-terminal residue of ubiquitin that is covalently linked to K14 (Fig. S1A).

Figure 1B shows autoradiographs of methyltransferase assays for Clr4 full length and the KMT domain, which both manifest a similar degree of strong stimulation by H3K14ub (Fig. 1C, Fig. S1D-F). Conversely, H3K18ub, a different ubiquitin mark, is a significantly poorer substrate than H3K14ub, demonstrating the specificity of Clr4 for H3K14ub. Quantification of the stimulation by the different substrates further showed that H3K14ub elicits a 40-50-fold boost in activity at a substrate concentration of 20 μM (Fig. 1C). We therefore conclude that the H3K14ub substrate triggers a dramatic and specific increase in the methyltransferase activity of Clr4, and that the KMT domain is sufficient to mediate this regulatory mechanism. We also note that the KMT domain shows consistently higher activity than the full length construct in both radiometric and fluorescence-based assays (Fig. S1G, H), which suggests that the N-terminus plays a regulatory role in the sensing of H3K14ub by the catalytic domain. However, since the primary effect of H3K14ub traces to the catalytic domain, we decided to further investigate the stimulation mechanism using the isolated KMT domain.

To fully characterize the H3K14ub-mediated stimulation we measured enzyme kinetics of Clr4KMT on H3K14ub vs. the unmodified peptide (Fig. 1D). An approximately hundred-fold difference for the Michaelis-Menten (K_M) constants was measured with 0.32 ± 0.07 μM for H3K14ub and 28.1 ± 7.9 μM for H3. Also, the maximum turnover number (k_{cat}) was about three times higher in presence of

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

H3K14ub (48.2 ± 2.0 h^{-1}) compared with unmodified H3 (16.6 ± 1.4 h^{-1}). This leads to an increase in overall enzymatic efficiency by 253.9 ± 94.8 fold. The kinetic measurements indicate that the presence of ubiquitin on lysine 14 leads to a tighter binding of the H3 substrate and that it elicits an allosteric effect in the Clr4 KMT domain, which drives a more efficient catalytic turnover.

To differentiate whether Clr4 has an independent allosteric site for ubiquitin or whether Clr4 stimulation specifically depends on K14-linked ubiquitin, we added increasing amounts of free ubiquitin to the methyltransferase reaction and measured its influence on the reaction rate. While we observed no significant increase in activity for unmodified H3, we observed a drop in the activity on H3K14ub at high concentrations of free ubiquitin (Fig.1E). These results suggest that the HMT domain of Clr4 recognizes the H3K14ub modification *in cis* and that formation of the Clr4-H3K14ub complex stimulates the catalytic activity of Clr4.

Stirpe et al Fig. 1

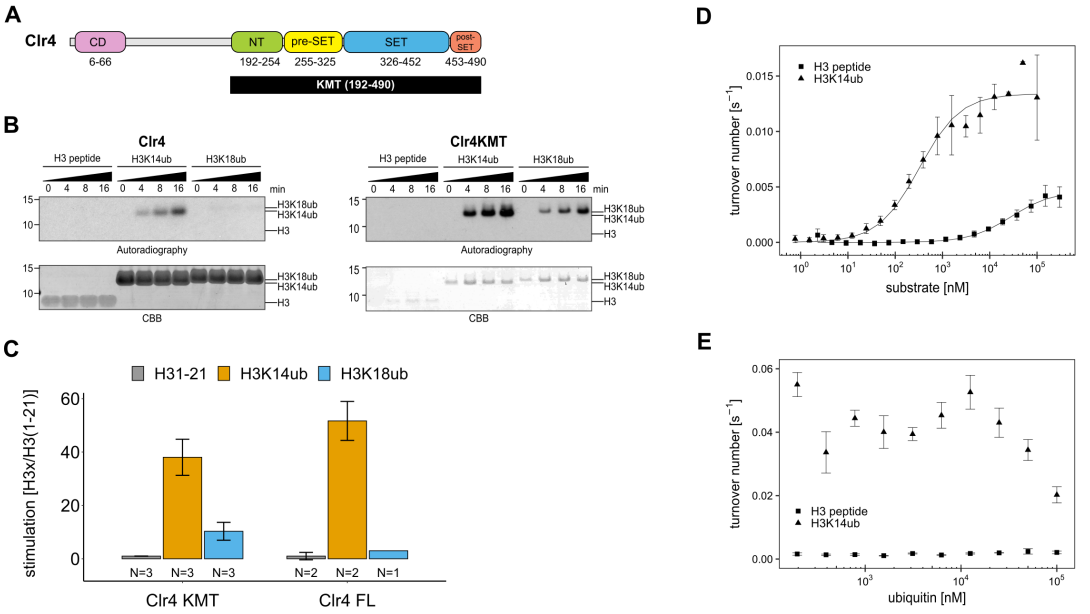

Figure 1: The catalytic domain of Clr4 senses the H3K14ub mark. (A) Domain organization of the Clr4 protein. CD: chromodomain; NT: N-terminal domain; SET: Su-var39/Enhancer of zeste/Trithorax domain; KMT: lysine methyltransferase domain. (B) Methyltransferase time

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

course on ubiquitinated versus unmodified H3 peptides shows strong stimulation by H3K14-linked ubiquitin on full-length Clr4 and the isolated KMT domain. 3H-SAM was used as methyl donor with 20 nM enzyme and analyzed by SDS-PAGE. (C) Stimulation of methyltransferase activity of Clr4 (20 nM) comparing indicated substrates with unmodified H3 peptide (substrates at 20 μM) by radiometric filter binding assay with ^{3}H-SAM as methyl donor. (D) Michaelis-Menten kinetics of SAH production of the Clr4 KMT domain on unmodified versus K14ub H3 peptides. Measured using TR-FRET competition assay (Cisbio EPIgeneous Methyltransferase Assay kit). (E) Ubiquitin competition assay demonstrates the specificity of Clr4 for ubiquitin and shows that covalent linkage in cis is required for activation. C-E) Error bars indicate standard error of the mean, N=3 unless indicated.

H3K14 ubiquitin mark affects the structural dynamics of Clr4

To understand the structural basis for the regulation of Clr4 by H3K14ub we performed hydrogen/deuterium exchange coupled to mass spectrometry (HDX-MS) analysis on the free Clr4 KMT domain and on Clr4KMT in complex with H3- and H3K14ub peptides. HDX-MS measures protein dynamics based on the rate of exchange of protein amide protons with the solvent [35]. Changes in HDX rates upon complex formation identify regions of the protein that are affected by the formation of the complex. When we compared the dynamics for Clr4KMT alone and in the presence of an excess of H3K14ub or unmodified H3 peptides, we observed a strong and unique reduction of the HDX rate for residues 243-261 of Clr4 in the presence of H3K14ub (Fig. 2A, B, Fig. S2, Tables S5-7). A further region between residues 291-305 showed amide protection by both H3 and H3K14ub substrates. These results suggest that the H3 peptide interacts with residues 291-305, while the ubiquitin moiety binds to a region involving residues 243-261. The H3K14ub binding region identified by HDX-MS maps to the NT domain just before it transitions into the pre-SET domain (Fig 2A). This

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

region corresponds to a stretch of residues that form a ridge along the "back" of Clr4, opposite to the active site pockets where cofactor and substrate peptide bind (Fig. 2C).

We also compared HDX-MS rates of isolated H3K14ub peptide with the H3K14ub-Clr4 complex and found that amino acids 25-43 on ubiquitin were protected from exchange upon interaction with Clr4 (Fig. 2D,E, Fig. S2, Table S6). These residues map to the α-helix of ubiquitin, indicating that this surface interacts with Clr4.

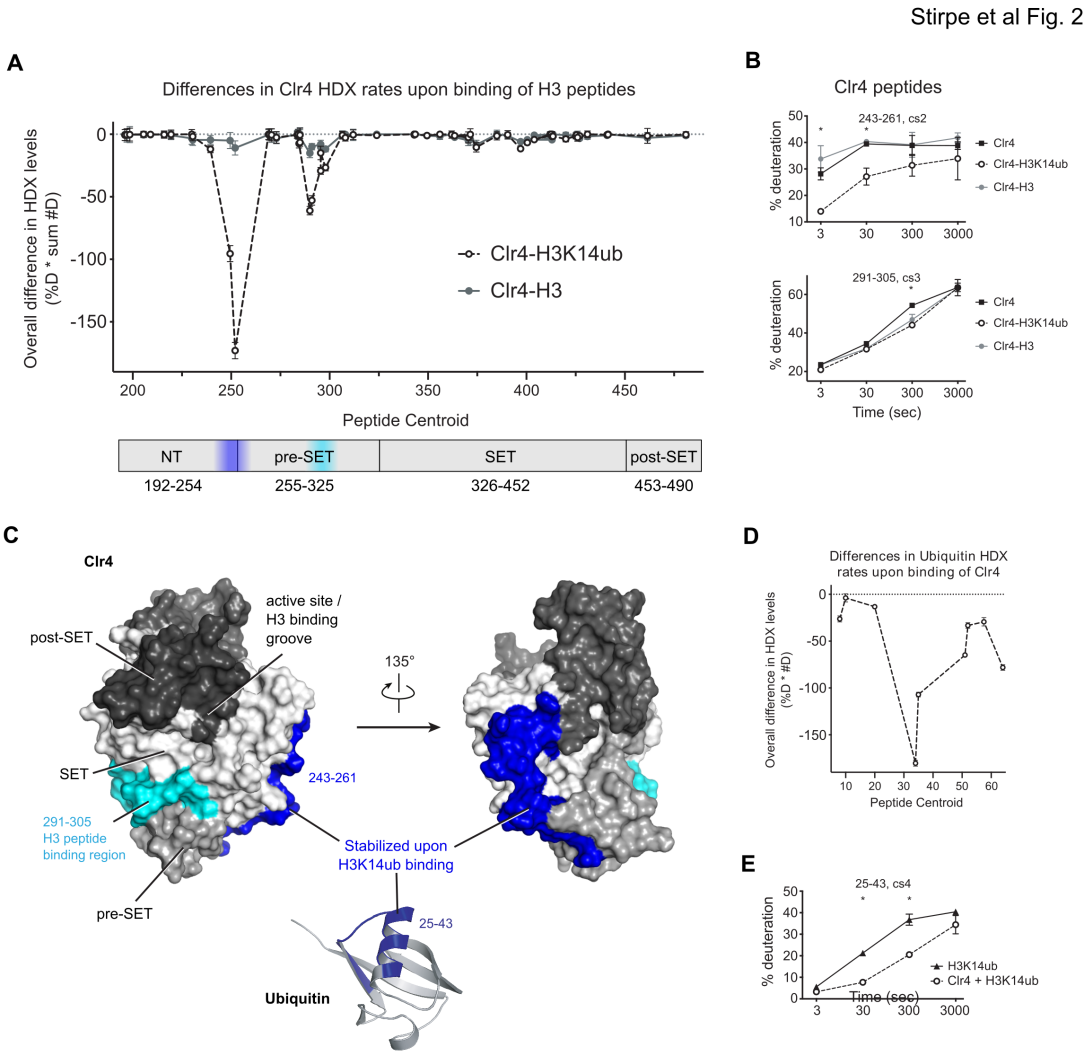

Figure 2: H3K14ub binding stabilizes residues 243-262 on Clr4. (A) Differences in Clr4 HDX rates upon binding of H3 and H3K14ub peptides are shown for each analyzed peptide. Values are plotted as the product of [deuteration percentage] * [number of deuterons] to minimize the influence of peptide length on the results. Domain diagram indicates the regions

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

showing stabilization upon interaction with H3 peptides in general (cyan) and more specifically with H3K14ub (blue). (B) Uptake plot for two peptides representative of regions showing differences in HDX rate of Clr4 upon peptide binding. (C) Surface representation of Clr4 structure (PDBID: 6BOX) and cartoon representation of ubiquitin (PDBID: 1UBQ). (D) Differences in ubiquitin HDX rates for the H3-K14ub peptide in the absence and presence of Clr4. Results are shown as in (A). (E) Uptake plot for peptide 25-43 of ubiquitin linked to H3K14 showing a marked reduction in HDX rate at 30 sec and 300 sec incubation time in deuterated buffer.

Clr4 surface mutants lose affinity for H3K14ub

The K_M for the methyltransferase activity of Clr4 on the H3K14ub substrate and the HDX-MS results suggest the formation of a tight substrate-enzyme complex. To characterize the binding energy of this interaction, we used isothermal titration calorimetry (ITC) and we found that Clr4 binds to the H3K14ub peptide with a dissociation constant (K_d) of 80 ± 6 nM in a reaction that is dominated by enthalpy (Fig. 3A). The ITC data thus confirm that H3K14ub binds with high affinity to the KMT domain of Clr4.

To create loss-of-function mutations for functional studies of the H3K14ub-Clr4 interface we mutated residues 243-261 of Clr4, the region that showed the largest H3K14ub-specific changes by HDX-MS. Substitution of residues 253-256 (sequence DPNF to GGSG, referred to as Clr4GS253) yielded a stable protein. When assayed by ITC, Clr4GS253 showed a severe loss of binding to H3K14ub (Fig. 3B). We then analyzed the Clr4 structure for potential surface-exposed residues that might be critical for the Clr4-H3K14ub interaction and we chose to further mutate three conspicuous phenylalanines located in the proximity of the GS253 mutations (F256A, F310A and F427A, referred to as Clr4F3A). ITC measurements of the Clr4F3A mutant showed that it, too, had lost its high-affinity binding to H3K14ub (Fig. 3C).

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Since the Clr4F3A mutant showed favourable biochemical behaviour we decided to determine its structure by X-ray crystallography in order to determine the effect of the mutations on the folding of Clr4. These efforts yielded a 2.46 Å structure of Clr4F3A (Table 1). The packing of the molecules in these crystals is similar to the packing observed in the structure of the autoinhibited Clr4, despite a difference in space group (PDBID:6BOX) [34]. When comparing with previous X-ray structures of Clr4 we observed no significant difference in global or local protein folding between the wild type Clr4 KMT domains and the F3A mutant (RMSD = 0.37Å) (Fig. 3D,E, Fig S3). This indicates that the F3A mutations do not induce significant structural changes.

We then tested the Clr4GS253 and Clr4F3A mutants in methyltransferase assays using ^{3}H-SAM and substrate concentrations of 20 μM. These conditions are complementary to the binding studies by ITC and focus on methyltransferase activity under conditions where substrate affinity is less critical. Based on the gel-based and filter binding assays the mutants lose activity on H3K14ub when compared to wild type Clr4KMT (Fig. 3F, G). Concomitantly, we observe ~30% loss of activity for the mutants on the unmodified H3 peptide, and the stimulatory effect of H3K14ub is therefore only reduced in the GS253 mutant, while the F3A is similar to wild type (Fig. S3C). Based on the ITC and activity assays we conclude that both mutants lose high-affinity binding for H3K14ub, and that they likely impede an allosteric mechanism that has H3K14ub-dependent and -independent components. When comparing the two mutants, the defects are more pronounced in Clr4GS253.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

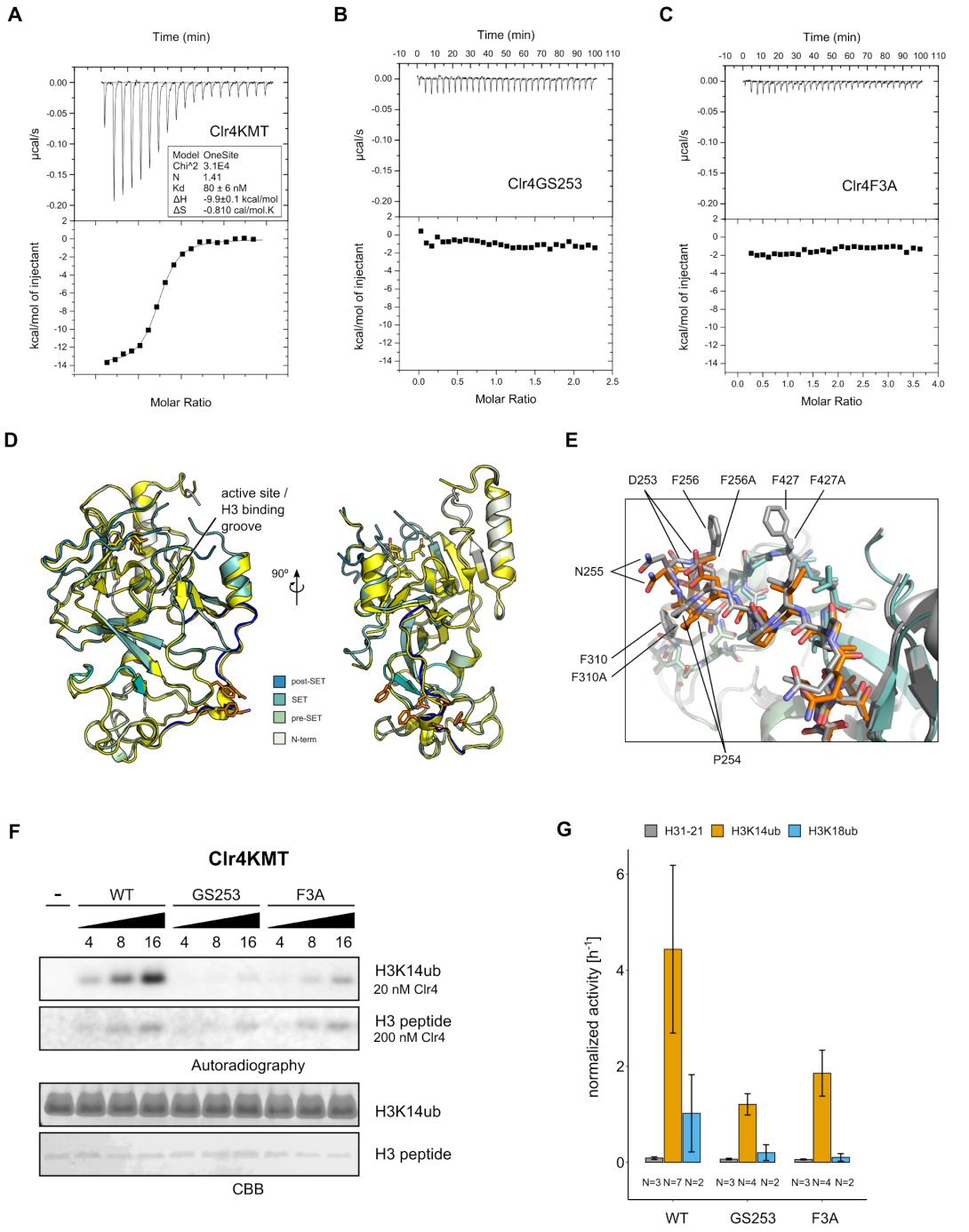

Figure 3: Mutants in the Clr4-H3K14ub interface defective for substrate binding. (A-C) Thermodynamic parameters of H3K14ub substrate binding to wild type and mutant Clr4KMT protein were measured using ITC. Heat rates (top panel) were integrated and plotted as a binding isotherm (bottom panel). Fit to a one site model is shown as a solid line where fitting was possible. (D) Overview and detail view of superimposition of Clr4F3A crystal structure (shades of teal) to wild type Clr4 (yellow) (PDBID: 6BOX) indicate absence of major differences in global folding. Residues 243-262 are colored in blue. Mutated residues (orange)

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

and SAH co-factor are shown in stick representation. (E) Superposition of mutated region in stick representation for Clr4F3A (orange) and Clr4 (gray) (PDBID:6BOX). (F) Methyltransferase kinetics of wild type and Clr4KMT mutants measured by detection of ^{3}H incorporation on SDS-PAGE gels shows loss of stimulation by H3K14ub for GS253 and F3A mutant. To observe sufficient signal differing enzyme concentrations of 20 and 200 nM were used for H3K14ub and H3 peptides, respectively. (G) Stimulation by H3K14ub and H3K18ub is shown as relative rates of ^{3}H incorporation quantified by filter binding and scintillation counting using the same experimental setup as in F. Error bars indicate standard error of mean.

Clr4 loss of function mutants affect heterochromatin

In *S.pombe* heterochromatin formation depends strongly on the H3K9me2/3 methyltransferase activity of Clr4. Clr4 knock-out and mutant strains are defective for transcriptional gene silencing and display hyperacetylated heterochromatic regions [24,36,37,34]. The Clr4GS253 and F3A mutations, therefore, provide a unique opportunity to investigate the functional importance of the H3K14ub-mediated stimulation for the *S.pombe* heterochromatin system. We introduced the Clr4GS253 and Clr4F3A mutations at the endogenous *clr4* locus and crossed them into a dual reporter (*imr1L:ura4*$^{+}$/*otr1R:ade6*$^{+}$) background for evaluating heterochromatin silencing using comparative growth assays [23]. Strains with a functional silencing machinery are able to repress transcripts from these two centromeric reporters, and wild type cells therefore show restricted growth on medium lacking uracil and turn red on medium containing low concentrations of adenine. Elevated levels of *ura4*, however, render the strains sensitive to growth on media containing 5-fluoro-uracil FOA, which is converted into the toxic product fluorodeoxyuridine when Ura4 is expressed at elevated levels. The *clr4GS253* mutant showed increased growth on medium lacking uracil, no growth on FOA plates and white colonies on low adenine, comparable to the *clr4Δ* mutant, which indicates a severe loss of transcriptional gene silencing (Fig. 4A). The F3A mutant also

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

showed a growth phenotype very similar to *clr4Δ*. However, the slightly pinkish color on low Ade plates suggests that the silencing defect is less severe than in GS253 or *clr4Δ*. The Clr4 mutants are expressed at normal levels and Co-IP experiments with Rik1 show that they remain associated with the CLRC complex (Fig. 4B,C). To get a quantitative measure of the loss of transcriptional gene silencing, we analyzed endogenous heterochromatic transcripts at centromeric *dg/dh* repeats and subtelomeric *tlh1* locus by RT-qPCR (Fig. 4D). The levels of these transcripts were greatly elevated in the *clr4GS253* mutant, similar to *clr4Δ*. In contrast, the F3A mutant showed a more nuanced silencing defect with high transcript levels for centromeric dh and telomeric transcripts, and a modest ~10-fold increase for centromeric dg. These findings are consistent with the weaker silencing defect of *clr4F3A* observed in the growth assays and with the weaker loss of enzymatic stimulation *in vitro*. To test if the loss of gene silencing is associated with loss of H3K9 methylation we performed Chromatin Immunoprecipitation (ChIP) against H3K9me1, H3K9me2, and H3K9me3 (Fig. 4E). Both H3K9me2 and H3K9me3 were completely abolished in *clr4GS253* and *clr4F3A* strains at centromeric *dh* repeats, being indistinguishable from *clr4Δ*. *clr4GS253* and *clr4Δ* also showed elevated levels of H3K9me1 when compared to wild type *clr4+*, while this was not observed for *clr4F3A*. The elevated H3K9me1 marks support the hypothesis that another unidentified methyltransferase could be depositing this mark [38]. We further found that in agreement with the increased transcript levels RNA Polymerase II occupancy at *dh* repeats was strongly increased.

In summary, we observe defects in transcriptional gene silencing and heterochromatin formation for the GS253 and F3A Clr4 mutants that are very similar to *clr4Δ*. Furthermore, the severity of the phenotype correlates with the degree of loss of enzymatic function on H3K14ub *in vitro*, further supporting that mediating the cross-

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

talk between H3K14ub and H3K9me2/3 is an essential property of the Clr4 KMT domain for heterochromatin formation and maintenance.

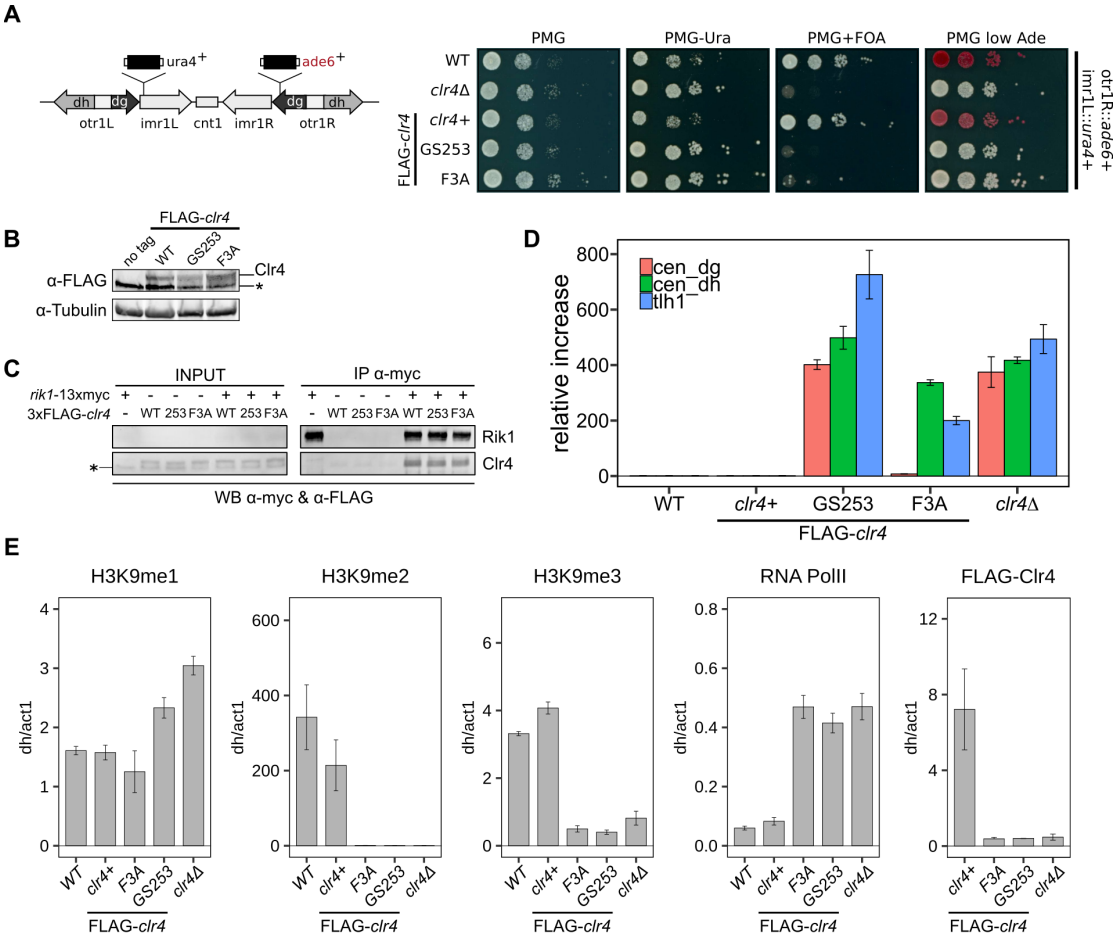

Figure 4: Clr4 mutants lose heterochromatin. (A) Serial dilution growth assays of wild type and the Clr4 mutants. Strains were assessed for growth on PMG media, PMG-ura to monitor *imr1*L::*ura4*$^{+}$ expression, and PMG+FOA to monitor silencing of *imr1*L::*ura4*$^{+}$. Expression of *otr1*R::*ade6*$^{+}$ was tested on PMG containing low Adenine. (B) Immunoblot for FLAG-Clr4 on cell lysates from *clr4* mutant strains. (C) Co-IP experiment to assess the stability of the CLRC complex in *clr4GS253* (label "253") and *clr4F3A* mutants. (D) Changes in steady-state transcript levels in *clr4* mutant strains relative to wild type cells were measured by RT-qPCR for centromeric *dg/dh* repeats and *tlh1* transcripts at telomeres. *act1* was used as an internal standard for all measurements. (E) ChIP for wild type and indicated mutant strains against FLAG-Clr4, histone marks H3K9me1, H3K9me2 and H3K9me3, and RNA Polymerase II. *act1* was used as an internal standard for all measurements. Mean and standard errors in (D) and (E) were calculated from a minimum of 3 independent biological experiments.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

H3K14ub stimulation is conserved in mammalian SUV39H2

To investigate if stimulation of H3K9 methylation by H3K14ub is a conserved process we purified the KMT domains of both human G9a and SUV39H2, as well as plant SUVH4/Kryptonite and performed methyltransferase assays to investigate the substrate preference of these enzymes by radiometric and fluorescence-based assays (Fig.5A, S5). Of our candidate enzymes, SUV39H2 consistently showed stimulation by H3K14ub, while we could not detect stimulation for G9a or SUVH4. These results suggest that the H3K14ub-mediated stimulation is conserved in a subset of SUV39 family proteins in higher eukaryotes, and that H3K14ub might be implicated in the regulation of H3K9 methylation in mammalian heterochromatin.

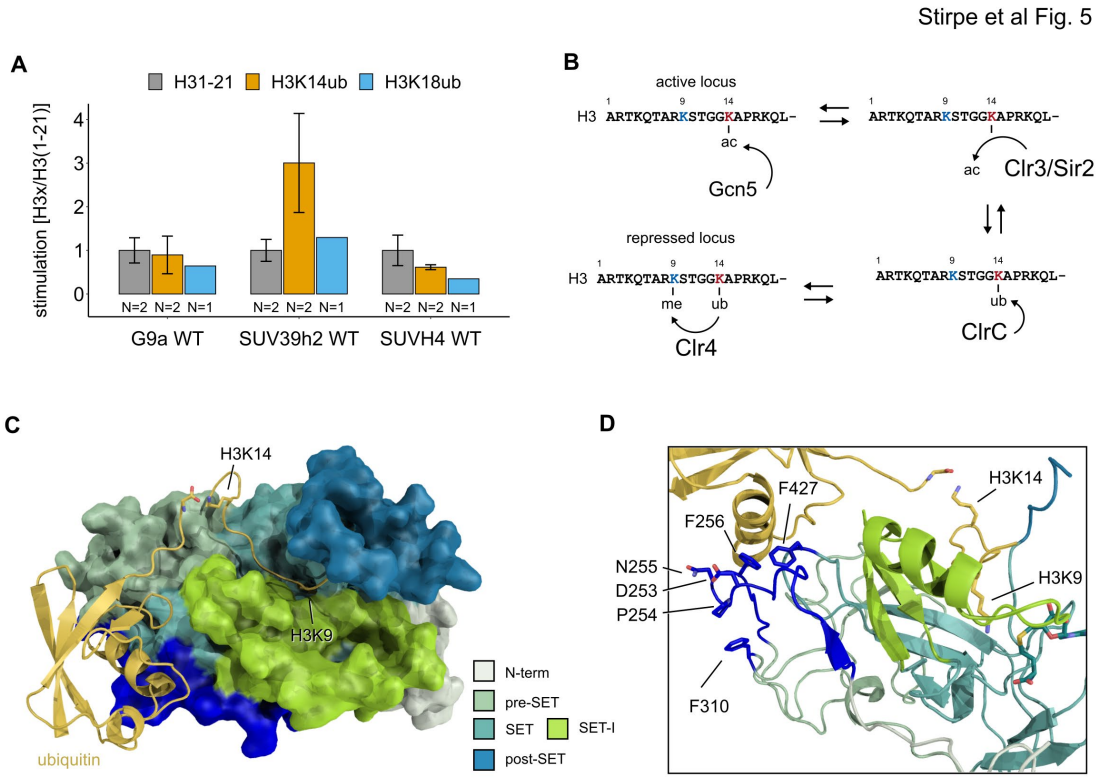

Figure 5: Conservation and mechanistic model of H3K14ub stimulation in the SUV39 methyltransferase family. (A) Methyltransferase rates of human G9a and human SUV39H2 and arabidopsis SUVH4 on indicated substrates were determined by quantifying ^{3}H-methyl incorporation using filter binding assays. Error bars correspond to standard error of

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

mean. (B) Schematic of how H3K14 ubiquitylation licenses H3K9me2/3 deposition and how H3K14ac specifically prevents this. (C, D) Overview and detail view of Clr4-H3K14ub model calculated with HADDOCK using restraints obtained by HDX-MS and mutagenesis.

Discussion

Histone methylation and acetylation are key post-translational modifications involved in the regulation of chromatin accessibility and transcription. We show here that the poorly understood histone modification H3K14ub leads to the strong stimulation of the methyltransferase activity of the fission yeast SUV39 family enzyme Clr4 through a sensing mechanism in its catalytic domain. The H3K14ub modification has, therefore, the biochemical potential to direct H3K9 methylation by Clr4 with high specificity. These results support a recently proposed regulatory role for H3K14ub in heterochromatin formation [32]. Based on these and our findings we propose that H3K14ub serves as a licensing mechanism for the deposition of H3K9me2/3 marks and heterochromatin formation (Fig. 5B). In this model, acetylation of H3K14, which is strongly associated with euchromatin, assumes a much more specific role than normally associated with histone acetylation: it serves to prevent ubiquitination of K14 similar to a protecting group in organic chemical synthesis. The H3K14ac-specific HDACs Clr3 and Sir2 therefore indirectly control H3K9me2/3 by regulating the availability of K14 for ubiquitination by the CLRC ubiquitin ligase complex. Consistent with this model, mutations of H3K14 lead to loss of H3K9 methylation [39], which is phenocopied by double deletion of Clr3 and Sir2 [17]. The genetic data further suggests that ubiquitination is essential for the establishment of heterochromatin since mutants of the CLRC complex are completely devoid of heterochromatin [24,26–29]. In the *clr4GS253* and *clr4F3A* mutants, we observe a severe loss of heterochromatin, which validates that recognition of H3K14ub by Clr4 is important for the establishment and maintenance of heterochromatin. Since the GS253 and FA3 mutants show a ~30% reduction in activity

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

on unmodified H3 peptides, we compared our results to the results obtained by mutants in the post-SET domain, which have a similar defect on methyltransferase activity on unmodified H3 [34]. These mutants (Clr4 K472R or K455R/K472, for example) show reduced, but still substantial levels of H3K9me3 at pericentromeres, while H3K9me3 is not detectable in the GS253 and F3A mutants. We therefore attribute the strong heterochromatin defects to the loss of affinity for H3K14ub in these mutants.

We have consistently observed H3K14ub-mediated stimulation in reactions with the isolated KMT domain of Clr4 and we have shown that the N-terminus of Clr4 is not essential for recognition of H3K14ub. Our observations are difficult to reconcile with the results reported in Oya et al [32], where an intact N-terminus was required for stimulation. We speculate that this discrepancy is caused by the use of different protein expression constructs. While Oya et al use GST-tagged proteins we have removed the tags for all constructs. An influence on biochemical activity by the GST-tag has been observed previously for SUV39H2 [40], and it is therefore conceivable that a GST-tag also influences the behaviour of Clr4. However, we do observe an inhibitory role for the N-terminus of Clr4 in the stimulation by H3K14ub. Oya et al. have reported a ubiquitin binding site in the CD-adjacent region, leading us to suggest that this region might act as a decoy, which prevents non-specific activation of the SET domain by random ubiquitinated peptides. Given the regulatory role of the N-terminus, there is a possibility that H3K9me2/3 binding by the CD of Clr4 will influence H3K14ub regulation. How the two mechanisms interact in heterochromatin spreading remains to be determined.

The H3K14ub binding region identified by HDX-MS is generally poorly defined in the electron density of the available crystal structures for Clr4 as well as in structures of

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

human SUV39H2. The sequence of this region is not conserved between species, which prevents the identification of a specific sequence motif. Nevertheless the HDX-MS and mutagenesis results provide a valuable set of constraints that we employed to predict the atomic structure of the Clr4-H3K14ub complex by protein-protein docking using HADDOCK [41]. The resulting model (Fig. 5C) shows that the H3K14ub peptide can comfortably fit the H3 peptide into the binding groove and simultaneously cover Clr4 residues 253-256 and F427 with its K14-linked ubiquitin moiety. The model suggests that F310 is unlikely to be involved directly in the interaction with ubiquitin as the N-terminal residues 253-256 shield F310 from interacting with ubiquitin (Fig. 5D).

Comparing this structural model with the best understood regulatory mechanisms of other SET domain proteins reveals interesting parallels and suggests a mechanism for how H3K14ub might stimulate the enzymatic activity of Clr4. EZH2, the catalytic subunit of the polycomb repressive complex PRC2, which methylates H3K27, is stimulated through conformational changes triggered by binding of its own H3K27me product to the neighboring subunit EED in the PRC2 complex [42–45]. Intriguingly, the 243-261 region that we identified by HDX-MS corresponds structurally to the "SET activation loop" (SAL) observed in PRC2 structures, which stabilizes the SET-I domain, a region inside the SET domain that provides one side of the peptide binding groove [42,43,46]. In a further example, the mixed lineage leukemia complexes with catalytic SET domain proteins MLL1 or MLL3 are activated by structural rearrangement of the regulatory subunits RBBP5 and ASH2L. RBBP5 thereby assumes a similar position as the SAL in PRC2 and stabilizes the SET-I domain in coordination with ASH2L [47]. In these experiments, apo MLL3 and the MLL3-ASH2-RBBP5 complex showed little structural differences in the positioning of SET-I, but complex formation significantly reduced the conformational dynamics of the SET-I domain. The

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

structural similarities between these regulatory mechanisms identify the SET-I domain as a potential mediator of the H3K14ub stimulation. In our model, the SET-I domain is sandwiched between the H3K14-linked ubiquitin moiety and the H3 peptide bound in the active site (Fig. 5C,D). Therefore, binding of H3K14ub to Clr4 might not only contribute to high substrate affinity but also to an optimally organized active site.

The function of H3K14ub in human cells is poorly understood, but it has been detected by mass spectrometry [48]. In combination with the conservation of the H3K14ub mediated stimulation in SUV39H2 this suggests that the cross-talk observed in fission yeast is conserved in mammals. Further studies will be required to determine the prevalence of the crosstalk between H3K9me2/3 and H3K14ub, and the role it plays in the biology of eukaryotic organisms.

Materials and Methods

Generation of *S.pombe* strains

S.pombe strains were grown and manipulated using standard techniques unless differently stated. The list of strains used in this study can be found in Table S1. Strains were crossed and analysed either by tetrad dissection or random spore analysis. For tetrad dissection, only spores from complete tetrads were further analysed. In absence of any selection markers, the genotype was analysed by colony PCR. The strain *clr4Δ::kanMX* was generated by transforming a wild type strain with a DNA fragment containing the kanMX cassettes plus 80bp of sequence found upstream and downstream of *clr4* ORF. The transformants were selected on YES plates containing 100 µg/µl of G418, and correct integration was confirmed by PCR and sequencing. The strain *3xFLAG-clr4* was generated by CRISPR mutagenesis of a wild type strain using as

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

homologous repair template a fragment of DNA containing the 3xFLAG tag plus 500 bp of homology flanking the cleavage site and a silent point mutation disrupting the PAM site. The *3xFLAG-clr4GS253* strain was generated with the same approach by transforming the *3xFLAG-clr4* strain with DNA containing the desired mutations and homologies. The sgRNA sequences used for CRISPR are indicated in Table S4. The *clr4Δ::rpl42-natMX* strain was generated by transformation of a *rpl42.sP56Q* host strain with a DNA fragment containing the rpl42-natMX cassette plus 500 bp of homology for sequences flanking *clr4* ORF. The integrants were selected by rounds of positive selection on rich media containing 100 µg/ml nourseothricin, and followed by negative selection on plates containing 100 µg/ml cycloheximide. The integrations were tested by PCR and sequencing. To generate the *3xFLAG-clr4_F3A* mutant strain, the *clr4Δ::rpl42-natMX* strain was transformed with a DNA fragment containing the desired mutations and 500 bp homology arms outside the *rpl42-natMX* cassette. The strain was selected with rounds of positive and negative selection on YES plates containing cycloheximide and nourseothricin, respectively. The strain *rik1-13xmyc-kanR* was generated by transformation of a wild type strain with a *13xmyc-kanR* cassette with homology for the 3'end of *rik1*. The strain was selected on plates containing G418. All transformants were analysed by PCR, sequenced, and the strains were outcrossed at least once.

For plate based silencing assays, reporter strains were generated by crossing the strains of interest with strains containing the marker genes *ura4+* and *ade6+* integrated at the centromere I (*imr1L::ura4+*, *otr1R::ade6+*).

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Silencing assays in *S.pombe*

Strains were grown overnight at 30 °C in YES to a density of $5x10^6$ cells/ml. Cells were pelleted by centrifugation and resuspended in water. Ten-fold serial dilutions were spotted onto PMG (PMG + Ade + Leu + His + Ura), PMG-Ura (PMG + Ade + His + Leu), and PMG + 1xFOA (PMG + Ade + Leu + 50mg/l Uracil + 1mg/ml 5-FOA), with $1x10^4$ cells in the highest density spots.

RT-qPCR

RT qPCR was performed as previously described [16]. Briefly, Strains were grown in YES to a density of 5 × 10^6 cells/mL, pelleted by centrifugation, washed with water and frozen at -70°C until further usage. RNA was purified by Trizol extraction followed by phenol/chloroform extraction and ethanol precipitation. Contaminant DNA was removed by DNaseI treatment and phenol/chloroform extraction followed by ethanol precipitation. cDNA was generated with Evo-Script Universal cDNA Master kit (Roche) and qPCR performed using SYBR Green I Master kit on LightCycler 480 instrument(Roche). qPCR primers used in this study are listed in Supplemental Table S2 and *act1*+ was used for normalization. Data were analyzed using the ΔΔCt method [49].

Chromatin immunoprecipitation (ChIP)

Chromatin immunoprecipitation was performed as previously described with some modifications[16]. Briefly, cells for ChIP were grown in 50 mL of YES medium to a density of 2 × 10^7 cells/mL. For RNA Polymerase II, H3K9me1, H3K9me2, H3K9me3, cells were fixed in 1% formaldehyde for 30 min. For 3xFLAG-Clr4, a dual-crosslinking approach was employed as previously described (Tian et al. 2012). Cells

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

were incubated at 18°C for 2 h, resuspended in 5 mL PBS and crosslinked at room temperature with 1.5 mM ethylene glycol bis-succinimidyl succinate (EGS, Thermo Scientific). After 30 min of incubation, 1% formaldehyde was added and cells were crosslinked for further 30 min. The residual formaldehyde was quenched with 125 mM Glycine for 5 min. Cells were washed and stored at -70°C until further usage.

Pellets were resuspended in ChIP Buffer (50 mM HEPES/KOH pH 7.6, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% Na-Deoxycholate, 1 mM PMSF, 1× cOmplete (Roche), and lysed by bead-beating. Chromatin was enriched by centrifugation, resuspended in lysis buffer and sonicated for 15 min (30 sec/30 sec ON/OFF) in a Bioruptor Pico. Cell debris was removed by centrifugation. For each IP, 50 µL of sheared soluble chromatin was diluted with 450 µl ChIP Buffer, mixed with 1 µg of antibody (H3K9me1 [ab8896], H3K9me2 [ab1220], RNA PolII [ab817], and FLAG-M2 [F1804]) and incubated for 2 h, followed by 45 min incubation with 30µl of Dynabeads Protein A or G. For H3K9me3, 0.75 µg of anti-H3K9me3 (Diagenode C15500003) was pre-incubated with 30µl of Dynabeads MyOne Streptavidin C1, followed by blocking with 5 µM biotin, according to manufacturer instructions. The loaded beads were then added to the diluted chromatin and incubated for 2 hours.

The bead/protein complex was washed three times with ChIP Buffer, once with high salt buffer (50 mM HEPES/KOH pH 7.6, 500 mM NaCl, 1 mM EDTA), once with LiCl buffer (5 mM Tris-Cl pH 8, 250 mM LiCl, 0.5% Triton X-100, 0.5% Na-Deoxycholate, 0.05% Tween 20), and once with TE (10 mM Tris-Cl pH 8, 1 mM EDTA). For ChIP experiments directed against FLAG tagged proteins, Na-Deoxycholate was omitted from ChIP lysis buffer. The protein-DNA complex was eluted in elution buffer (50 mM Tris-Cl pH 8, 10 mM EDTA, 1% SDS) at 65°C for 15 min, and the crosslinking reversed overnight at 65°C. The samples were then treated with proteinase K and DNA

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

was purified by phenol-chloroform. qPCR was performed using primers given in Table S2; *act1*+ was used as internal control.

Co-Immunoprecipitation from *S.pombe* cells

Cells were grown overnight at 30°C to a density of 1.0*10^7 cells/ml, washed once with PBS and the pellet frozen at -70°C until further usage. Pellets were resuspended in co-IP buffer (50 mM HEPES/KOH pH 7.5, 300 mM NaCl, 1 mM EDTA, 5 mM CHAPS, 10 mM DTT) supplemented with cOmplete EDTA free (Roche). The cells were mixed with the same volume of 0.5 mm glass beads and ruptured by bead beating in a Fastprep24 (MP Biomedical). Cell debris was removed by centrifugation and the supernatant incubated with Myc-Trap magnetic beads (ChromTek) for 1 hour at 4°C. The unbound fraction was removed, and the beads were washed 5 times with co-IP buffer. The bound proteins were eluted by boiling in PGLB (125 mM Tris-HCl, 4% SDS, 50% Glycerol, 0.2% Orange G) supplemented with 100 mM DTT. The proteins of interest were analysed by western blot.

Western blotting

Samples were run on 10% Bis-Tris gels and transferred to nitrocellulose membrane (Biorad). Proteins were detected by Western blotting using antibodies against either tag or protein of interest, followed by incubation with secondary antibodies labeled with the DyLight system, and scanning with the Odyssey Imaging System (LI-COR).

Protein expression in E.coli

All Clr4 protein constructs were cloned as N-terminal fusion to a 6xHIS-SUMO-tag into pET28b-SUMO vectors from Lima's lab using Gibson cloning. The plasmid containing 6xHIS-SUMO-SUVH4 was provided by Steven Jacobsen. Plasmids for

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

SUV39H2 (Addgene plasmid #25115) and G9a (Addgene plasmid #25503) as N-terminal fusion to a 6xHIS-tag were provided by Cheryl Arrowsmith.

All constructs were expressed in Rosetta2(DE3) strain. Cells were grown at 37°C to log-phase in 2xYT media supplemented with 2% ethanol, cooled down on ice and induced with 0.4mM IPTG at 16°C for 18 hours. Pellets were resuspended in NiNTA wash buffer (50 mM Tris-HCl pH 8.0, 300 mM NaCl, 20 mM Imidazole pH 7.5, 1 mM β-mercaptoethanol) supplemented with protease inhibitor cocktail and frozen at -80°C until further usage.

Protein purification from E.coli

For purification, frozen cells were quickly thawed at 37 °C, supplemented with protease inhibitor cocktail, benzonase, and ruptured by using either sonication or the Emulsiflex homogenizer. Cell debris was removed by centrifugation. The cleared lysate was filtered through a 0.45µm filter and loaded on a HisTrap HP column (GE Healthcare). The full length Clr4 was enriched from the cleared lysate by ammonium sulfate fractionation with 40% NH4SO4. The pellet containing Clr4 was then resuspended in NiNTA wash buffer (50mM Tris-HCl pH 8.0, 300mM NaCl, 20mM Imidazole pH 7.5, 1mM β-mercaptoethanol), filtered through a 0.45µm filter and loaded on a HisTrap HP column. After washing with NiNTA wash buffer, the proteins were eluted by either gradient or step elution in elution buffer (50mM Tris-HCl pH 8.0, 300mM NaCl, 300mM Imidazole pH 7.5, 1mM β-mercaptoethanol). All proteins containing a 6xHIS-SUMO-tag were treated with Ulp1 while being dialysed overnight in cleavage buffer (5mM Tris-HCl pH 8.0, 150mM NaCl, 2mM β-mercaptoethanol). SUV39h2 and G9a were treated with Thrombin to remove the 6xHIS-tag and dialysed in cleavage buffer. The cleaved tags and Ulp1 were removed by negative purification on a HisTrap HP

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

column. All proteins were further purified by SEC (Superdex75 or Superdex200, GE Healthcare) in gel filtration buffer (5mM Tris-HCl pH 8.0, 150mM NaCl, 5mM DTT). Peak fractions were concentrated as needed in Amicon Ultra spin concentrators. For methyltransferase assays, the concentrated proteins were mixed 1:1 with glycerol, frozen in liquid nitrogen and stored at -80°C until further usage. Clr4FL, SUV39h2, and SUVH4 were frozen without glycerol.

Materials for peptide synthesis

All solvents and reagents were purchased from commercial sources and used without further purification. All amino acid derivatives, 2-chlorotrityl chloride (2-Cl Trt) resin and 2-(7-Aza-1H-benzotriazole-1-yl)-1,1,3,3- tetramethyluronium hexafluorophosphate (HATU) were purchased from Novabiochem, Merck (Darmstadt, Germany). N,N-Dimethylformamide (DMF), N,N-diisopropylethylamine (DIEA) and piperidine were from Acros Organics (Geel, Belgium). O-(6-Chlorobenzotriazol-1-yl)-N,N,N',N'-tetramethyluronium hexafluorophosphate (HCTU) was from Carl Roth GmbH (Karsruhe, Netherlands). Hydrazine monohydrate was purchased from Alfa Aesar (Heysham, UK), acetonitrile (ACN) from Avantor Performance Materials (USA). Tentagel S RAM resin, hydroxybenzotriazole (HOBt), tris(2-carboxyethyl)phosphine (TCEP), diethylether, phenylsilane, tetrakis(triphenylphosphine)palladium(0), trifluoroacetic acid (TFA), dichloromethane (DCM), triisopropylsilane (TIS), L-glutathione reduced (GSH), sodium diethyldithiocarbamate trihydrate and methyl thioglycolate (MTG) were from Sigma Aldrich (Taufkirchen, Germany). 2,2'-Azobis[2-(2- imidazolin-2-yl)propane]dihydrochloride (VA-044) was purchased from Wako Pure Chemical Industries, Ltd. (Osaka, Japan). All other commonly used chemical reagents and buffer

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

components were from Applichem (Darmstadt, Germany) and Fisher Scientific (Reinach, Switzerland).

Instrumentation for peptide synthesis and purification

Reaction vessels for manual peptide synthesis as well as the automated Tribute peptide synthesizer were from Protein Technologies Inc. Analytical RP-HPLC analysis was performed on an Agilent 1260 series instrument using an analytical Agilent Zorbax C18 column (column dimensions: 150 mm x 4.6 mm, 5 µm particle size) at a flow rate of 1 mL/min. All RP-HPLC analyses were done with 0.1 % (v/v) TFA in H_2O (RP-HPLC solvent A) and 90 % ACN and 0.1 % (v/v) TFA in H_2O (RP-HPLC solvent B) as mobile phases. Typically, a gradient from 0-70 % solvent A to solvent B over 30 min was used for analytical RP-HPLC analyses unless otherwise stated. Purification of proteins on a semi-preparative scale were performed on an Agilent 1260 series instrument using a semi-preparative Agilent Zorbax C18 column (column dimensions: 250 mm x 9.4 mm, 5 µm particle size) at a flow rate of 4 mL/min. Lyophilization was carried out with a Telstar LyoQuest freeze dryer. Electrospray ionisation mass spectrometric (ESI-MS) analysis was conducted on a Shimadzu MS2020 single quadrupole instrument connected to a Nexera UHPLC system. Mass spectra were acquired by electrospray ionisation in positive ion mode in the mass range of 200-2000 m/z.

Preparation of Fmoc-aa-hydrazine-Cl-trityl-resin

Preparation of Fmoc-Arg(Pbf)-NHNH-Cl-Trt-resin was performed as reported previously [50]. In general, 0.50 g 2-Cl-Trt-resin (substitution: 1.63 mmol/g, 1.00 eq., 0.82 mmol) were swollen in 3.00 mL DMF for 30 min at room temperature. Subsequently, the resin was allowed to cool to 0° C and 1.00 mL of a solution

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

containing DIEA (3.00 eq., 2.45 mmol, 427 µl) and hydrazine monohydrate (2.00 eq., 1.64 mmol, 80 µl) in DMF was added dropwise. The reaction mixture was stirred 1 h at room temperature. 100 µl methanol (MeOH) was added, the resin was stirred 10 min at room temperature, transferred to a reaction vessel for manual peptide synthesis and washed thoroughly with DMF.

Due to the low stability of the hydrazine resin, the first amino acid Fmoc-Arg(Pbf)-OH was coupled manually by standard Fmoc chemistry [51]. The loading of the 2-Cl-Trt-resin with hydrazine was assumed to be 50% (1.00 eq., 0.41 mmol). Fmoc-Arg(Pbf)-OH (5.00 eq., 2.05 mmol) was activated with 3.90 mL of a 0.50 M HATU solution (4.76 eq., 1.95 mmol) in DMF followed by 2 min incubation at room temperature. Then, 714 µL DIEA (10.00 eq., 4.10 mmol) were added and the reaction mixture was incubated another 1 min at room temperature. The activated amino acid was added to the resin, incubated 30 min at room temperature and washed with DMF. To ensure high-yield, the coupling procedure was repeated. Finally, the resin was washed with DMF, DCM and MeOH and dried under vacuum. Resin substitution was determined by treating a defined amount of resin with 20% piperidine in DMF for 30 min at RT, followed by spectrophotometrical quantification of released dibenzofulven-related species (resin loading: 0.51 mmol/g).

Automated Solid Phase Peptide Synthesis (SPPS)

The sequences of the two peptides synthesized by Fmoc-SPPS are summarized in Table S3. The synthetic procedures and analytical data is presented in the following. General protocol: the peptides were synthesised by the Tribute peptide synthesiser (PTI) on the previously prepared Fmoc-Arg(Pbf)-hydrazine-Cl-Trt-resin (peptide P1) or on Tentagel S RAM resin (peptide P2) yielding peptides with C-terminal hydrazide

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

(peptide P1) or amide (peptide P2). The syntheses were performed on 0.1 mmol scale using Fmoc chemistry. The following base-resistant groups were employed to protect amino acid side chains: Arg(Pbf), Thr(tBu), Lys(Boc), Gln(Trt), Ans(Trt), Ser(tBu). Of note, Lys14 (peptide P1) and Lys18 (peptide P2) were protected with Alloc protecting group. To maximise synthesis yield, pseudoproline dipeptide building blocks were used and amino acids were double coupled where necessary.

Briefly, the N-terminal Fmoc-group was deprotected with 20% (v/v) piperidine in DMF. Activation of amino acid (5.00 eq., 0.50 mmol) was achieved by addition of HCTU (4.76 eq. 0.48 mmol) and DIEA (10.00 eq., 1.00 mmol). The coupling step was performed by adding the activated amino acid to the resin, followed by 30 min incubation at room temperature. When the full-length peptides were assembled, the peptidyl-resin was washed with DMF, DCM and MeOH and dried under vacuum. Importantly, the N-terminal Fmoc protecting group was not removed at this stage to allow further manipulation of the peptide.

Alloc deprotection

The peptidyl-resin was swollen for 30 min in DCM. Alloc deprotection of Lys14 (peptide P1) or Lys18 (peptide P2) was initiated by addition of 1.00 mL of dry DCM and $PhSiH_3$ (24 eq., 24 mmol), followed by $Pd(PPh_3)_4$ (0.25 eq., 0.025 mmol) in 3.00 mL dry DCM. The peptidyl-resin was incubated for 30 min at room temperature and washed with DCM. The deprotection reaction with $PhSiH_3$ and $Pd(PPh_3)_4$ was repeated two more times. The resin was thoroughly washed with DCM followed by washing with 0.5% (v/v) DIEA in DMF; 0.5% (w/v) sodium-diethyldithiocarbamate in DMF; 50% (v/v) DCM in DMF; 0.5% (w/v) HOBt in DMF and extensively washed with DMF.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Manual coupling of Boc-Cys(Trt)-OH

The cysteine residue was coupled manually to the e-amino group of Lys14 (peptide P1) or Lys18 (peptide P2), yielding peptides P1' and P2' respectively (Figure S1). Boc-Cys(Trt)-OH (5 eq., 0.50 mmol) was activated by addition of 0.95 mL of a 0.50 M HATU solution (4.76 eq., 0.48 mmol) in DMF, followed by 2 min incubation at room temperature. 172 µL DIEA (10 eq., 1 mmol) were added and the reaction mixture was incubated another 1 min at room temperature. The activated amino acid was added to the peptidyl-resin, incubated 30 min at room temperature and washed with DMF. To ensure high-yield, the coupling procedure was repeated. Finally, the resin was washed thoroughly with DMF.

N-terminal Fmoc deprotection and cleavage from the resin

Fmoc-deprotection was achieved by treating the peptidyl-resin with 10.00 ml 20% (v/v) piperidine in DMF for 5 min. The deprotection step was repeated.
The peptides were cleaved from the resin using either 95% (v/v) TFA, 2.5% (v/v) TIS, 2.5% (v/v) H_2O. The crude peptides were precipitated by addition of ice-cold diethyl ether, recovered by centrifugation, dissolved in 50% (v/v) acetonitrile in H_2O, flash-frozen and lyophilised.

Preparation of Ubiquitin-MESNa

Ubiquitin (G76C) was expressed in Rosetta2(DE3) cells as a C-terminal fusion to the N-terminal half of the Npu split intein, followed by a hexahistidine tag [52]. Bacterial cells were resuspended in lysis buffer (50mM Na-phosphate buffer pH 7.0, 300mM NaCl, 40mM Imidazole pH 7.0) and ruptured by sonication. Ubiquitin-NpuN-6xHIS was purified by Ni-affinity chromatography and eluted with 300mM imidazole. The

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

ubiquitin thioester was generated by intein derivatization with 100mM Sodium 2-mercaptoethanesulfonate (MESNa). The reaction was allowed to proceed for 16h at room temperature. Imidazole was removed by diafiltration and Ubiquitin-MESNa was further purified by negative Ni-affinity chromatography to remove the NpuN-6xHIS tag. Finally, ubiquitin-MESNa was diafiltrated against 0.1% TFA, lyophilised and stored at -20°C.

One pot ligation and desulfurization

Ubiquitin-MESNa (1.00 eq., 1.16 µmol) was dissolved in 115.80 µL ligation buffer (6.00 M GmdHCl, 0.20 M sodium phosphate, pH 7, degassed) to a final concentration of 10.00 mM and 0.77 µL MTG (7.50 eq., 8.70 µmol) were added to the solution. The thiol-thioester exchange reaction was allowed to proceed for 20 min at room temperature. Ligation between ubiquitin-MTG thioester and the cysteine residue coupled to Lys14 (peptide P1) or Lys18 (peptide P2) was initiated by addition of peptide P1' or peptide P2' (1.30 eq., 1.51 µmol), followed by addition of 5.80 µl of 0.50 M TCEP solution in ligation buffer. The ligation mixture was incubated at 25 °C for 16h. The progress of the reaction was monitored by RP-HPLC and ESI-MS analysis.

When the ligation was complete, radical desulfurization of the cysteine at the ligation site was performed in the same reaction tube without prior purification of the ligation product. TCEP desulfurization buffer (0.50 M TCEP, 6.00 M GdmHCl, 0.20 M phosphate, pH 7) was added to a final TCEP concentration of 0.25 M. The desulfurization reaction was initiated by addition of VA-044 and GSH to a final concentration of 30 mM and 40 mM, respectively. The reaction mixture was incubated at 42° C for 6h and the progress of the reaction was monitored by RP-HPLC and ESI-MS analysis, yielding ubiquitin adducts P1'' and P2'' (Figure S1).

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Purification of H3K14ub (P1'') and H3K18ub (P2'')

The ubiquitin adduct peptides P1'' and P2'' were purified by semi-preparative RP-HPLC on a linear gradient from 25 to 55% solvent B over 40 min. Pure fractions were pooled, lyophilised and analysed by analytical RP-HPLC and ESI-MS (Figure S1B,C).

TR-FRET Histone Methyltransferase Assay

The assays were performed using the EPIgeneous Methyltransferase kit (Cisbio). The reactions were performed by mixing the Clr4 constructs (30nM) with the indicated substrates (1nM-200µM for K_M measurements in figure 1D, 2µM in figure 1E, 20µM in figure S5) in reaction buffer (25mM Sodium Phosphate pH 8.0, 100mM NaCl, 10% v/v glycerol, 10µM SAM) and incubated at 30°C for 15min, unless differently stated. The reactions were stopped and measured using a TECAN SPARK plate reader. Detection was performed at 620 and 665 nm wavelength. The 665/620 signal ratio was used to extrapolate the concentration of SAH from a standard curve.

Tritium-based Histone Methyltransferase Assay

The HMT assays were performed in a total volume of 10 µl by mixing the substrates with the indicated constructs in reaction buffer (25mM Sodium Phosphate pH 8.0, 100mM NaCl, 10% v/v glycerol, 2.9µM (0.5 µCi) *S*-[Methyl-3H]-Adenosyl-L-methionine (17.2 mCi/mL, Perkin Elmer)) and incubated at 30 °C. Measurements were performed with 20µM of substrate, and 20nM or 200nM of each enzyme. Assay specific concentrations and incubation times are indicated in the text and captions. For gel based assays, the reactions were stopped by addition of 5 µl of PGLB + 100 mM DTT and the unreacted SAM was removed by SDS-PAGE. The gels were stained with

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

coomassie, imaged, and incubated for 30 min in EN3HANCE (Perkin Elmer) prior drying. The methylated histones were detected by autoradiography.

For filter binding assays, the reactions were stopped by addition of 5µl 75% acetic acid, and 10µl of each mixture was spotted onto phosphocellulose filter paper disks. Each filter was washed for 5 min in 2ml 100mM sodium bicarbonate pH 9.0, for a total of three times. The activity on each disk was quantified in a liquid scintillation counter after 1 hour incubation in 3ml ULTIMA GOLD F cocktail.

Hydrogen Deuterium Exchange Mass Spectrometry (HDX-MS)

HDX reactions were done in 50 µl volumes with a final protein concentration of 3 µM of enzyme. The dynamics of Clr4 was compared in three different conditions: i) soluble Clr4, ii) complex of Clr4 bound to ubiquitinated H3 peptide, iii) Clr4 in presence of an excess of H3 peptide. Dynamics of ubiquitinated H3 peptide was investigated by comparing its D2O incorporation alone or when in complex with Clr4. Protein or protein complexes were prepared in 5 µl volume and deuterium exchange was initiated by addition of 45 µl of deuterated buffer (10 mM Tris-HCl pH 7.4/ 150 mM NaCl/ 10 mM DTT). For the condition of Clr4 with an excess of H3 peptide, the deuterated buffer was supplemented with H3 peptide (from lyophilized powder). Exchange was carried out at room temperature for 4 timepoints (3s, 30s, 300s, 3000s) and terminated by the addition of 20 µl ice-cold quench buffer (3M guanidine – HCl/ 100mM NaH2PO4 pH 2.5/ 1% formic acid). All experiments were repeated in duplicates. Samples were immediately frozen in liquid nitrogen and stored at -80 °C for up to two weeks.

Protein samples were thawed and injected in a UPLC system immersed in ice. The protein was digested via two immobilized pepsin columns (Thermo #23131), and

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

peptides were collected onto a VanGuard precolumn trap (Waters). The trap was subsequently eluted and peptides separated with a C18, 300Å, 1.7 µm particle size Fortis Bio column 100 x 2.1 mm over a gradient of 8 – 30 %B over 18 min at 90 µl/min (Buffer A: 0.1% formic acid; buffer B: 100% acetonitrile/ 0.1% formic acid). Mass spectra were acquired on an Orbitrap Velos Pro (Thermo), for ions from 400 to 2200 m/z using an electrospray ionization source operated at 160 °C, 5 kV of ion spray voltage. Peptides were identified by data-dependent acquisition after MS/MS and data were analyzed by Mascot. Deuterium incorporation levels were quantified using HD examiner software (Sierra Analytics), and the quality of every peptide was checked manually. Results are presented as a percentage of maximal deuteration, with a theoretical maximal deuteration level of 86.3 %.

Isothermal titration calorimetry

ITC experiments were performed at 23 °C using a MicroCal VP-ITC calorimeter (Malvern Panalytic). All proteins and peptides used were dialyzed overnight against ITC buffer (25 mM Sodium Phosphate pH 8.0, 100 mM NaCl, 10% v/v Glycerol, 0.5 mM TCEP) prior to experiments. 10 µl of the indicated Clr4 constructs at 50 µM were injected in 180 s time intervals in the cell containing the H3K14ub peptide at 5 µM. After subtracting heat enthalpies for titrations of the respective proteins into buffer, the ITC data were analyzed using the Origin software provided by the manufacturer.

Molecular Docking

Molecular docking was performed using Clr4KMT$^{192\text{-}490}$ (PDB:1MVH), ubiquitin (PDB:1UBQ) and a peptide comprising residues 7-17 of histone H3. The input peptide was derived by extension of the H3 peptide bound to *Neurospora crassa*

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

DIM-5 (PDB:1PEG). Docking was performed with a local version of HADDOCK 2.0 using the peptide identified by HDX-MS as restraints.

Crystallisation of Clr4F3A

Clr4F3A was purified as described above. Clr4F3A at 20mg/ml was mixed with H3K14ub peptide and S-Adenosyl-homocysteine to a ratio of 1:1.2:10 (protein:peptide:SAH) and diluted to a concentration of 10mg/ml. Protein was diluted 1:1 with reservoir solution (71.4 mM MES, 28.6 mM Imidazole, 20% PEG 10K, 30mM Magnesium Acetate, 6.6 % v/v MPD, 6.6 % v/v PEG 1000, 6.6 % v/v PEG 3350) and incubated at 18°C in sitting-drop vapour diffusion setup. Crystals were directly frozen in liquid nitrogen.

Crystallographic data collection and structure determination

X-ray diffraction data were collected at the Diamond Light Source beamline I04 (Didcot, Oxford, UK) at a wavelength of 0.97950 Å. The datasets were analysed using XDS and the CCP4 suite [53,54]. Data were corrected for anisotropy using the STARANISO server (http://staraniso.globalphasing.org). Molecular replacement was performed using MOLREP and the coordinates from PDB ID 6BOX (chain A) as search model [55]. Refinement was performed using REFMAC, and Coot was used to build the Clr4F3A model [56,57]. Structural alignments and figures were generated in PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC. Data collection and refinement statistics are presented in Table 1.

Accession codes

Coordinates and structure factors have been deposited in the Protein Data Bank under accession code 6Z2A.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

References

1. Greer, E. L. & Shi, Y. Histone methylation: a dynamic mark in health, disease and inheritance. *Nat. Rev. Genet.* **13**, 343–357 (2012).

2. Janssen, A., Colmenares, S. U. & Karpen, G. H. Heterochromatin: Guardian of the Genome. *Annu. Rev. Cell Dev. Biol.* **34**, 265–288 (2018).

3. Kaniskan, H. Ü., Martini, M. L. & Jin, J. Inhibitors of Protein Methyltransferases and Demethylases. *Chem. Rev.* **118**, 989–1068 (2018).

4. Dillon, S. C., Zhang, X., Trievel, R. C. & Cheng, X. The SET-domain protein superfamily: protein lysine methyltransferases. *Genome Biol.* **6**, 227 (2005).

5. Audergon, P. N. C. B. *et al.* Restricted epigenetic inheritance of H3K9 methylation. *Science* **348**, 132–135 (2015).

6. Ragunathan, K., Jih, G. & Moazed, D. Epigenetic inheritance uncoupled from sequence-specific recruitment. *Science* **348**, 1258699 (2015).

7. Grewal, S. I. RNAi-dependent formation of heterochromatin and its diverse functions. *Curr. Opin. Genet. Dev.* **20**, 134–141 (2010).

8. Verdel, A. *et al.* RNAi-Mediated Targeting of Heterochromatin by the RITS Complex. *Science* **303**, 672–676 (2004).

9. Buhler, M., Haas, W., Gygi, S. P. & Moazed, D. RNAi-Dependent and -Independent RNA Turnover Mechanisms Contribute to Heterochromatic Gene Silencing. *Cell* **129**, 707–721 (2007).

10. Schalch, T., Job, G., Shanker, S., Partridge, J. F. & Joshua-Tor, L. The Chp1–Tas3 core is a multifunctional platform critical for gene silencing by RITS. *Nat. Struct. Mol. Biol.* **18**, 1351–1357 (2011).

11. Shimada, Y., Mohn, F. & Bühler, M. The RNA-induced transcriptional silencing complex targets chromatin exclusively via interacting with nascent transcripts.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Genes Dev. **30**, 2571–2580 (2016).

12. Zhang, K., Mosch, K., Fischle, W. & Grewal, S. I. S. Roles of the Clr4 methyltransferase complex in nucleation, spreading and maintenance of heterochromatin. *Nat. Struct. Mol. Biol.* **15**, 381–388 (2008).
13. Bayne, E. H. *et al.* Stc1: A Critical Link between RNAi and Chromatin Modification Required for Heterochromatin Integrity. *Cell* **140**, 666–677 (2010).
14. Motamedi, M. R. *et al.* HP1 Proteins Form Distinct Complexes and Mediate Heterochromatic Gene Silencing by Nonoverlapping Mechanisms. *Mol. Cell* **32**, 778–790 (2008).
15. Fischer, T. *et al.* Diverse roles of HP1 proteins in heterochromatin assembly and functions in fission yeast. *Proc. Natl. Acad. Sci.* **106**, 8998–9003 (2009).
16. Leopold, K., Stirpe, A. & Schalch, T. Transcriptional gene silencing requires dedicated interaction between HP1 protein Chp2 and chromatin remodeler Mit1. *Genes Dev.* **33**, 565–577 (2019).
17. Alper, B. J. *et al.* Sir2 is required for Clr4 to initiate centromeric heterochromatin assembly in fission yeast. *EMBO J.* **32**, 2321–2335 (2013).
18. Bjerling, P. *et al.* Functional Divergence between Histone Deacetylases in Fission Yeast by Distinct Cellular Localization and In Vivo Specificity. *Mol Cell Biol* **22**, 2170–2181 (2002).
19. Buscaino, A. *et al.* Distinct roles for Sir2 and RNAi in centromeric heterochromatin nucleation, spreading and maintenance. *EMBO J.* **32**, 1250–1264 (2013).
20. Yamada, T., Fischle, W., Sugiyama, T., Allis, C. D. & Grewal, S. I. S. The Nucleation and Maintenance of Heterochromatin by a Histone Deacetylase in Fission Yeast. *Mol. Cell* **20**, 173–185 (2005).

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

21. Wang, Y. *et al.* Histone H3 Lysine 14 Acetylation Is Required for Activation of a DNA Damage Checkpoint in Fission Yeast. *J. Biol. Chem.* **287**, 4386–4393 (2012).

22. Allshire, R. C., Nimmo, E. R., Ekwall, K., Javerzat, J. P. & Cranston, G. Mutations derepressing silent centromeric domains in fission yeast disrupt chromosome segregation. *Genes Dev* **9**, 218–233 (1995).

23. Ekwall, K., Olsson, T., Turner, B. M., Cranston, G. & Allshire, R. C. Transient Inhibition of Histone Deacetylation Alters the Structural and Functional Imprint at Fission Yeast Centromeres. *Cell* **91**, 1021–1032 (1997).

24. Nakayama Ji, J., Rice, J. C., Strahl, B. D., Allis, C. D. & Grewal, S. I. Role of Histone H3 Lysine 9 Methylation in Epigenetic Control of Heterochromatin Assembly. *Science* **15**, 15 (2001).

25. Hong, E.-J. E., Villén, J., Gerace, E. L., Gygi, S. P. & Moazed, D. A cullin E3 ubiquitin ligase complex associates with Rik1 and the Clr4 histone H3-K9 methyltransferase and is required for RNAi-mediated heterochromatin formation. *RNA Biol.* **2**, 106–111 (2005).

26. Horn, P. J., Bastie, J.-N. & Peterson, C. L. A Rik1-associated, cullin-dependent E3 ubiquitin ligase is essential for heterochromatin formation. *Genes Dev.* **19**, 1705–1714 (2005).

27. Jia, S., Kobayashi, R. & Grewal, S. I. S. Ubiquitin ligase component Cul4 associates with Clr4 histone methyltransferase to assemble heterochromatin. *Nat Cell Biol* **7**, 1007–1013 (2005).

28. Li, F. *et al.* Two Novel Proteins, Dos1 and Dos2, Interact with Rik1 to Regulate Heterochromatic RNA Interference and Histone Modification. *Curr. Biol.* **15**, 1448–1457 (2005).

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

29. Thon, G. *et al.* The Clr7 and Clr8 Directionality Factors and the Pcu4 Cullin Mediate Heterochromatin Formation in the Fission Yeast Schizosaccharomyces pombe. *Genetics* **171**, 1583–1595 (2005).

30. Buscaino, A. *et al.* Raf1 Is a DCAF for the Rik1 DDB1-Like Protein and Has Separable Roles in siRNA Generation and Chromatin Modification. *PLoS Genet* **8**, e1002499 (2012).

31. Kuscu, C. *et al.* CRL4-like Clr4 complex in Schizosaccharomyces pombe depends on an exposed surface of Dos1 for heterochromatin silencing. *Proc. Natl. Acad. Sci.* **111**, 1795–1800 (2014).

32. Oya, E. *et al.* H3K14 ubiquitylation promotes H3K9 methylation for heterochromatin assembly. *EMBO Rep.* **20**, e48111 (2019).

33. Min, J., Zhang, X., Cheng, X., Grewal, S. I. S. & Xu, R.-M. Structure of the SET domain histone lysine methyltransferase Clr4. *Nat Struct Mol Biol* **9**, 828–832 (2002).

34. Iglesias, N. *et al.* Automethylation-induced conformational switch in Clr4 (Suv39h) maintains epigenetic stability. *Nature* **560**, 504–508 (2018).

35. Kochert, B. A., Iacob, R. E., Wales, T. E., Makriyannis, A. & Engen, J. R. Hydrogen-Deuterium Exchange Mass Spectrometry to Study Protein Complexes. in *Protein Complex Assembly: Methods and Protocols* (ed. Marsh, J. A.) 153–171 (Springer, 2018). doi:10.1007/978-1-4939-7759-8_10.

36. Bannister, A. J. *et al.* Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410**, 120–124 (2001).

37. Gerace, E. L., Halic, M. & Moazed, D. The Methyltransferase Activity of Clr4Suv39h Triggers RNAi Independently of Histone H3K9 Methylation. *Mol. Cell* **39**, 360–372 (2010).

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

38. Jih, G. *et al.* Unique roles for histone H3K9me states in RNAi and heritable silencing of transcription. *Nature* **547**, 463–467 (2017).

39. Mellone, B. G. *et al.* Centromere Silencing and Function in Fission Yeast Is Governed by the Amino Terminus of Histone H3. *Curr. Biol.* **13**, 1748–1757 (2003).

40. Schuhmacher, M. K., Kudithipudi, S., Kusevic, D., Weirich, S. & Jeltsch, A. Activity and specificity of the human SUV39H2 protein lysine methyltransferase. *Biochim. Biophys. Acta BBA - Gene Regul. Mech.* **1849**, 55–63 (2015).

41. Dominguez, C., Boelens, R. & Bonvin, A. M. J. J. HADDOCK: A Protein−Protein Docking Approach Based on Biochemical or Biophysical Information. *J. Am. Chem. Soc.* **125**, 1731–1737 (2003).

42. Jiao, L. & Liu, X. Structural basis of histone H3K27 trimethylation by an active polycomb repressive complex 2. *Science* **350**, aac4383 (2015).

43. Justin, N. *et al.* Structural basis of oncogenic histone H3K27M inhibition of human polycomb repressive complex 2. *Nat. Commun.* **7**, 11316 (2016).

44. Qian, C. & Zhou, M.-M. SET domain protein lysine methyltransferases: Structure, specificity and catalysis. *Cell. Mol. Life Sci. CMLS* **63**, 2755–2763 (2006).

45. Wu, H. *et al.* Structural Biology of Human H3K9 Methyltransferases. *PLOS ONE* **5**, e8570 (2010).

46. Brooun, A. *et al.* Polycomb repressive complex 2 structure with inhibitor reveals a mechanism of activation and drug resistance. *Nat. Commun.* **7**, 1–12 (2016).

47. Li, Y. *et al.* Structural basis for activity regulation of MLL family methyltransferases. *Nature* **530**, 447–452 (2016).

48. Harrison, J. S. *et al.* Hemi-methylated DNA regulates DNA methylation

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

inheritance through allosteric activation of H3 ubiquitylation by UHRF1. *eLife* **5**, e17101 (2016).

49. Yuan, J. S., Reed, A., Chen, F. & Stewart, C. N. Statistical analysis of real-time PCR data. *BMC Bioinformatics* **7**, 85 (2006).
50. Stavropoulos, G., Gatos, D., Magafa, V. & Barlos, K. Preparation of polymer-bound trityl-hydrazines and their application in the solid phase synthesis of partially protected peptide hydrazides. *Lett. Pept. Sci.* **2**, 315–318 (1996).
51. Atherton, E. & Sheppard, R. C. *Solid phase peptide synthesis: a practical approach*. (IRL Press, 1989).
52. Kilic, S., Boichenko, I., Lechner, C. C. & Fierz, B. A bi-terminal protein ligation strategy to probe chromatin structure during DNA damage. *Chem. Sci.* **9**, 3704–3709 (2018).
53. Kabsch, W. XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).
54. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242 (2011).
55. Vagin, A. & Teplyakov, A. Molecular replacement with MOLREP. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 22–25 (2010).
56. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
57. Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 355–367 (2011).

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Author Contributions

Conceptualization, T.S and B.F.; Methodology, A.S., T.S., B.F., O.V, S.N.; Investigation, A.S., O.V., N.G., S.K., S.N.; Writing – Original Draft, A.S., T.S., B.F., O.V.; Writing – Review & Editing, A.S., T.S., B.F. O.V.; Resources, T.S., B.F., O.S., A.H.; Funding Acquisition, T.S., B.F.; Supervision, T.S., B.F.

Acknowledgements

We thank Janet Partridge, Robin Allshire and Shiv Grewal for providing strains. We thank Steven Jacobsen for plasmids. We thank John Challiss and Raj Patel for support with radio-assays. We would like to thank Diamond Light Source for beamtime (proposal mx19880), and the staff of beamlines I04 for assistance with crystal testing and data collection.

Funding

This work was supported by the Swiss National Science Foundation SNF Professorship grants [PP00P3_139137, PP00P3_163760_1, PP00P3_172904 to T.S.]; Fondation Ernst et Lucie Schmidheiny, Fonds Constantin Topali and Société Académique de Genève (T.S.). OV was supported by a SNF Ambizione fellowship [PZ00P3_148269].

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.30.177071; this version posted July 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Tables

Table 1: Crystallographic Table

	Native
Data	
Wavelength (Å)	0.97950
Resolution range (Å)	70.68 - 2.46 (2.68 - 2.46)
Space group	P 21 21 2
Unit-cell parameters (Å, °)	92.44, 110.29, 70.68, 90.00, 90.00, 90.00
Total reflections	120264 (3409)
Unique reflections	19571 (979)
Multiplicity	6.1 (3.5)
Completeness (%) spherical	72.2 (16.2)
Completeness (%) ellipsoidal	93.1 (56.9)
Mean *I/σ(I)*	6.6 (1.4)
Wilson B factor (Å^2)	
R_{merge}	0.214 (0.898)
R_{meas}	0.233 (1.046)
R_{pim}	0.093 (0.525)
CC1/2	0.993 (0.627)
CC*	0.998 (0.878)
Refinement	
Resolution range	2.46
Total number of reflections	18561
Number of reflections in test set	973
R_{work} (%)	23.8
R_{free} (%)	24.9
CC(work)	0.923
CC(free)	0.910
No. of non-hydrogen atoms	4495
Macromolecule	4411
Ligands	62
Solvent	22
No. of protein residues	549
R.m.s.d., bonds (Å)	0.014
R.m.s.d., angles (°)	1.73
Ramachandran favored (%)	96.54
Ramachandran outliers (%)	0
Ramachandran allowed (%)	3.46
Clash score	2.42
Average B factor (Å^2)	
Macromolecule	31.85
Ligands	24.61
Solvent	21.63

*Statistics for the highest-resolution shell are shown in parentheses.