Structural studies provide new insights into the role of lysine acetylation on substrate recognition by CARM1 and inform the design of potent peptidomimetic inhibitors

Zhang, Y.; Marechal, N.; Haren, M.J. van; Troffer-Charlier, N.; Cura, V.; Cavarelli, J.; Martin, N.I.

Structural Studies Provide New Insights into the Role of Lysine Acetylation on Substrate Recognition by CARM1 and Inform the Design of Potent Peptidomimetic Inhibitors

Yurui Zhang[+],[a] Nils Marechal[+],[b] Matthijs J. van Haren[+],[a] Nathalie Troffer-Charlier,[b] Vincent Cura,[b] Jean Cavarelli,*[b] and Nathaniel I. Martin*[a]

The dynamic interplay of post-translational modifications (PTMs) in chromatin provides a communication system for the regulation of gene expression. An increasing number of studies have highlighted the role that such crosstalk between PTMs plays in chromatin recognition. In this study, (bio)chemical and structural approaches were applied to specifically probe the impact of acetylation of Lys[18] in the histone H3 tail peptide on peptide recognition by the protein methyltransferase coactivator-associated arginine methyltransferase 1 (CARM1). Peptidomimetics that recapitulate the transition state of protein arginine *N*-methyltransferases, were designed based on the H3 peptide wherein the target Arg[17] was flanked by either a free or an acetylated lysine. Structural studies with these peptidomimetics and the catalytic domain of CARM1 provide new insights into the binding of the H3 peptide within the enzyme active site. While the co-crystal structures reveal that lysine acetylation results in minor conformational differences for both CARM1 and the H3 peptide, acetylation of Lys[18] does lead to additional interactions (Van der Waals and hydrogen bonding) and likely reduces the cost of desolvation upon binding, resulting in increased affinity. Informed by these findings a series of smaller peptidomimetics were also prepared and found to maintain potent and selective CARM1 inhibition. These findings provide new insights both into the mechanism of crosstalk between arginine methylation and lysine acetylation as well as towards the development of peptidomimetic CARM1 inhibitors.

Introduction

Post-translational modifications (PTMs) on the *N*-terminal tails of histones are involved in the activation or silencing of gene expression and in the signalling of readers and writers. PTMs come in a broad variety including phosphorylation, glycosylation, acetylation, and methylation or larger modifications such as ubiquitination or SUMOylation. PTMs are often reversible and interconnected, resulting in a complex code of modifications, known as crosstalk, in which one modification can result in the blocking, promoting, or recruitment of another.[1,2] Examples of crosstalk in histones include the effect of serine phosphorylation on lysine acetylation and the effect of lysine acetylation on arginine methylation in histone H3.[3,4] In addition, crosstalk can even occur between entirely different regions of chromatin as shown by the crosstalk found between lysine methylation in histone H3 and lysine acetylation in histone H4, the crosstalk between DNA methylation and histone H3 methylation, and the effect of ubiquitination on histone H2B on lysine methylation in H3 and lysine acetylation in histone H2A.[5–8] Recent years have witnessed an increasing awareness of the roles played by this complex communication system in a variety of processes in both healthy and diseased states.[9–11]

In this investigation we focussed our attention on examining the impact of lysine acetylation in histone H3 on the recognition of neighbouring arginine residues by coactivator-associated arginine methyltransferase 1 (CARM1). Previous reports on lysine acetylation/arginine methylation crosstalk have shown that the acetylation of lysine residues Lys[18] and Lys[27] in histone H3 tails promote subsequent CARM1-mediated methylation of the neighbouring arginine residues Arg[17] and Arg[26] respectively.[12,13] Specifically, the methylation of H3 Arg[17] was shown to be enhanced through acetylation of Lys[18] and to a lesser extent also through acetylation of Lys[14] or Lys[23]. In addition, the affinity of CARM1 has been reported to be greater for substrate peptides containing Lys[18]Ac and Lys[23]Ac (but not Lys[14]Ac), suggesting that acetylation of Lys[18] and Lys[23] enhances binding of the H3 substrate for CARM1, leading to increased Arg[17] methylation.[12] However, kinetic analysis of this methylation process revealed that the increased catalytic efficiency of CARM1 for the H3 substrate acetylated at Lys[18] is rather driven by an increase in turnover number (k_{cat}) with no significant change in affinity (K_M).[14] The CARM1-mediated methyl transfer reaction is facilitated by several highly con-

[a] Y. Zhang,[+] Dr. M. J. van Haren,[+] Prof. N. I. Martin
Biological Chemistry Group, Institute of Biology Leiden
Leiden University, Sylviusweg 72, 2333 BE Leiden (The Netherlands)
E-mail: n.i.martin@biology.leidenuniv.nl
[b] Dr. N. Marechal,[+] Dr. N. Troffer-Charlier, Dr. V. Cura, Prof. J. Cavarelli
Department of Integrated Structural Biology
Institut de Génétique et de Biologie Moléculaire et Cellulaire
Université de Strasbourg, CNRS UMR 7104, INSERM U 1258
Illkirch, 67404 (France)
E-mail: jean.cavarelli@igbmc.fr

[+] These authors contributed equally to this work.

Supporting information for this article is available on the WWW under https://doi.org/10.1002/cbic.202100506

served active site residues. Notably, two glutamate residues (E^{258} and E^{267}, known as the "double E-loop") serve to position the guanidine moiety in close proximity to the methyl group of the *S*-adenosyl-L-methionine (SAM) cofactor. Additionally, a specific histidine residue (H^{415}) found in the so-called THW-loop, is crucial for the deprotonation of the guanidine, which in turn allows for the methyl group transfer to occur (Figure 1A). The explanation proposed by the authors for the observed increase in k_{cat} for H3 peptide substrates containing an acetylated lysine next to the target arginine is based on the local electrostatic environment in which a neutral (acetylated) residue will lower the pK_a of the catalytic histidine (H^{415}) and aspartic acid (D^{166}) residues, thereby stabilizing the transition state and facilitating the proton transfer necessary for the methyl group transfer.[14] For the methylation of $H3Arg^{26}$ a similar observation was made, wherein mutation of the neighbouring positively charged lysine to a neutral methionine residue (K27M), enhanced the methylation of $H3Arg^{26}$ to a similar extent as acetylation on Lys^{27}.[13] To compliment these biochemical studies, we here describe structural investigations employing H3-based peptidomimetics designed to directly probe the role of lysine acetylation on substrate recognition by CARM1.

Results and Discussion

To gain additional insights into the impact of lysine acetylation on arginine methylation by CARM1, we performed structural studies using a transition-state peptidomimetic strategy recently developed by our group (Figure 1B).[15] By covalently linking the adenosine moiety of the methyl donor SAM to the arginine side chain of a substrate peptide it is possible to generate conjugates that mimic the transition state of the first methylation step performed by the family of protein arginine *N*-methyltransferases (PRMTs). These peptidomimetics facilitate structural studies with PRMTs by circumventing the need to add SAM mimics (typically SAH or sinefungin) and the formation of a ternary complex with substrate peptides.[15] In synthesizing these peptidomimetics the adenosine group is introduced via the arginine guanidine moiety using a convenient on-resin modification procedure wherein the target arginine is initially installed as an Alloc-protected ornithine residue (Scheme 1). After assembly of the peptide using solid phase peptide synthesis (SPPS), the Alloc group is selectively removed leaving the other protecting groups unaffected and the peptide bound to the resin. The free ornithine side chain amine is subsequently coupled with a Pbf-protected thiourea-linked adenosine building block leading to formation of the arginine guanidino group directly linked to the adenosine moiety.[15] Capping of the *N*-terminus with acetic anhydride followed by deprotection and cleavage from Rink amide resin yields the modified peptide with amide groups on both the *N* and *C*-terminus, mimicking those present in the natural substrate.

For this study, two pairs of peptidomimetics were prepared based on residues 1–41 and 13–31 of the histone H3 tail peptide (Scheme 1). In these peptidomimetics the Arg^{17} residue was covalently linked to an adenosine moiety via a 3-carbon linker previously shown to be the optimal length for the recognition of such peptidomimetics.[15] To directly examine the influence of lysine acetylation, both sequences were also prepared as the Lys^{18}Ac variants which were readily prepared by introduction of the corresponding acetylated lysine building block during the SPPS. The two pairs of peptidomimetics thus

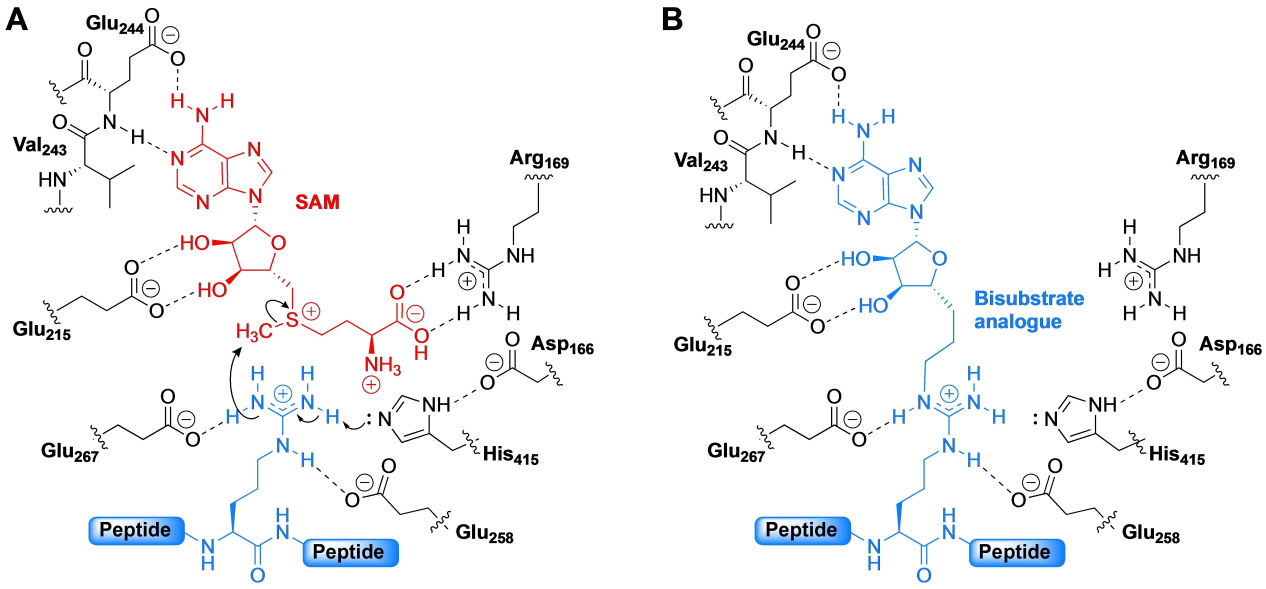
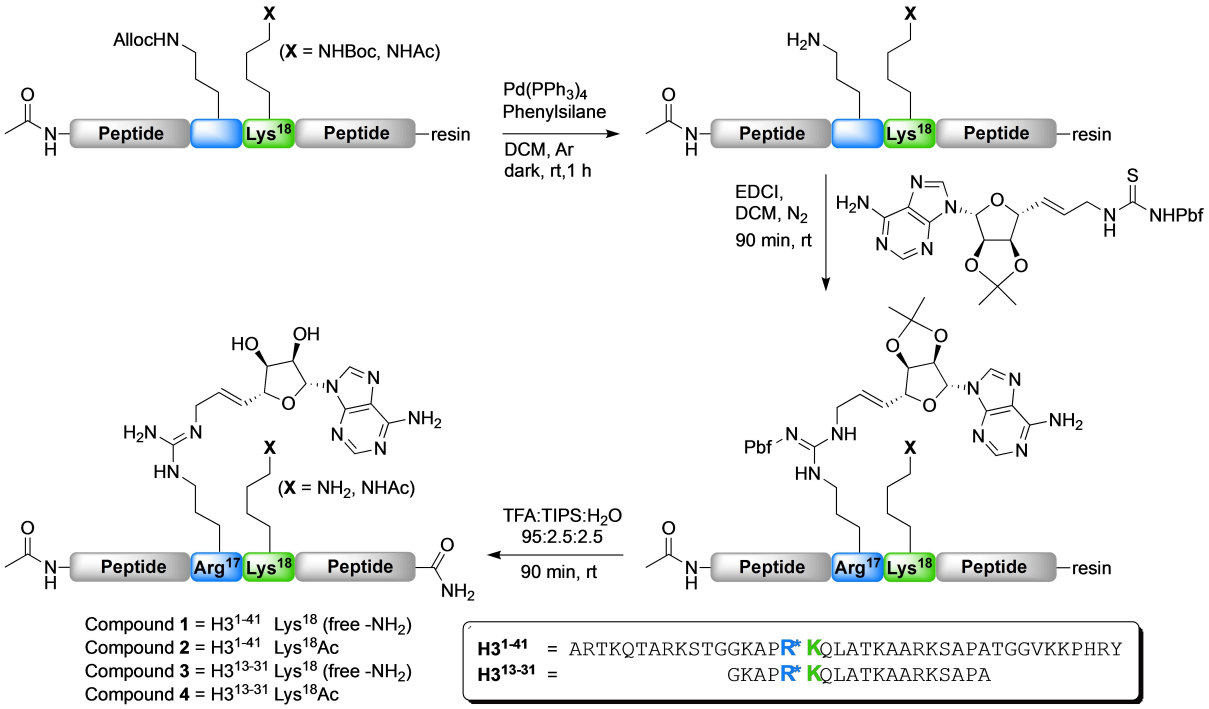


Figure 1. A) CARM1 active site with key active residues interacting with cofactor SAM and the target arginine of a peptide substrate. The double E-loop consists of glutamate residues Glu^{258} and Glu^{267}. His^{415} is involved in substrate recognition as part of the THW-loop and interacts with Asp^{166} for the deprotonation of the guanidine moiety facilitating methyl group transfer. B) Design strategy used in preparing bi-substrate analogues for structural studies and peptidic inhibitors of CARM1.



Scheme 1. General synthetic scheme for the preparation of transition state peptidomimetics with the adenosine moiety covalently linked to the side chain of the CARM1 target arginine. Also indicated is the neighbouring lysine residue in either acetylated or nonacetylated state. Details of the synthesis of the specific H3 peptidomimetics prepared as well as the preparation of the Pbf-protected adenosine thiourea building block are provided in the Supporting Information.

obtained where designed to address two aspects of H3 substrate recognition by CARM1: for both the $H3^{1–41}$ and $H3^{13–31}$ constructs the presence of free Lys^{18} or $Lys^{18}Ac$ was expected to provide insight into the role of crosstalk between substrate acetylation and methylation. In addition, the larger $H3^{1–41}$ constructs were prepared with the aim of also obtaining additional structural insights into long distance interactions known to be crucial for CARM1 substrate recognition.[16,17]

With peptidomimetics **1**–**4** in hand, co-crystallization studies were performed using an isolated catalytic domain of mmCARM1 (*Mus musculus* CARM1, residues 130–497). Peptidomimetics **1**–**4** were initially crystallized using PEG as the main crystallizing agent in line with previous structural studies with CARM1.[15,16] All structures were solved and refined (depending on crystals, resolution ranging from 2.0 to 2.7 Å at ESRF or SOLEIL synchrotron beamlines) in the space group $P2_12_12$ with one copy of the CARM1 tetramer in the asymmetric unit (see Supporting Information, appendix Table S1). While the resulting structures were solved and refined, the electron density maps displayed poor density beyond the previously established minimal binding sequence,[15,18] indicating high disorder or low occupancy for the peptidomimetics. Our previous experience in solving a number of different PRMT structures (PRMT4, PRMT2, PRMT6) has shown that in some cases PEG molecules can map the peptide binding site and in doing so inhibit, or strongly affect, peptide-binding.[19] To address this challenge we also explored the use of sodium malonate as the primary crystallization reagent instead of PEG. In total, 33 crystal structures of mmCARM1 in complex with the H3 peptidomimetics were solved and refined with PEG as the primary crystallization reagent along with an additional 12 structures obtained when using sodium malonate.[19] These studies revealed sodium malonate to be a superior crystallization reagent for obtaining high quality structures of CARM1 in complex with peptidomimetics **1**–**4** that were successfully solved and refined in the same space group. The highest resolution structures were obtained with $H3^{13–31}$ peptidomimetics **3** and **4** (2.54 Å for **3** (Lys^{18}–NH_2) and 2.2 Å for **4** ($Lys^{18}Ac$)). While the electron density maps obtained with **3** and **4** clearly revealed the conformation of 10 residues in all CARM1 complexes (amino acids 13 to 22) the same was not the case for the longer $H3^{1–41}$ peptidomimetics **1** and **2**. In the case of **1** and **2**, the peptidomimetics were found to occupy only two of the active sites of the mmCARM1 tetramer and are unable to displace all SAH molecules natively bound to the protein (the purified mmCARM1 construct naturally contains SAH molecule bound in the active site).

As noted, the $H3^{13–31}$ peptidomimetics **3** and **4** gave well-resolved structures for the first 10 amino acids. Beyond that however, residues 23 to 31 were never seen in the electron density maps, likely due to high levels of disorder. In the structures solved with both **3** and **4**, Leu^{20} of the $H3^{13–31}$ peptidomimetic is the last residue that is clearly seen to be interacting with CARM1 via Van der Waal interactions at Leu^{413}. Beyond that, the positioning of Ala^{21} and Thr^{22} indicates that residues 23–31 of the peptidomimetics are likely located in a region that has no interactions with CARM1 (Figure 2). While the longer $H3^{1–41}$-based peptidomimetics failed to give addi-

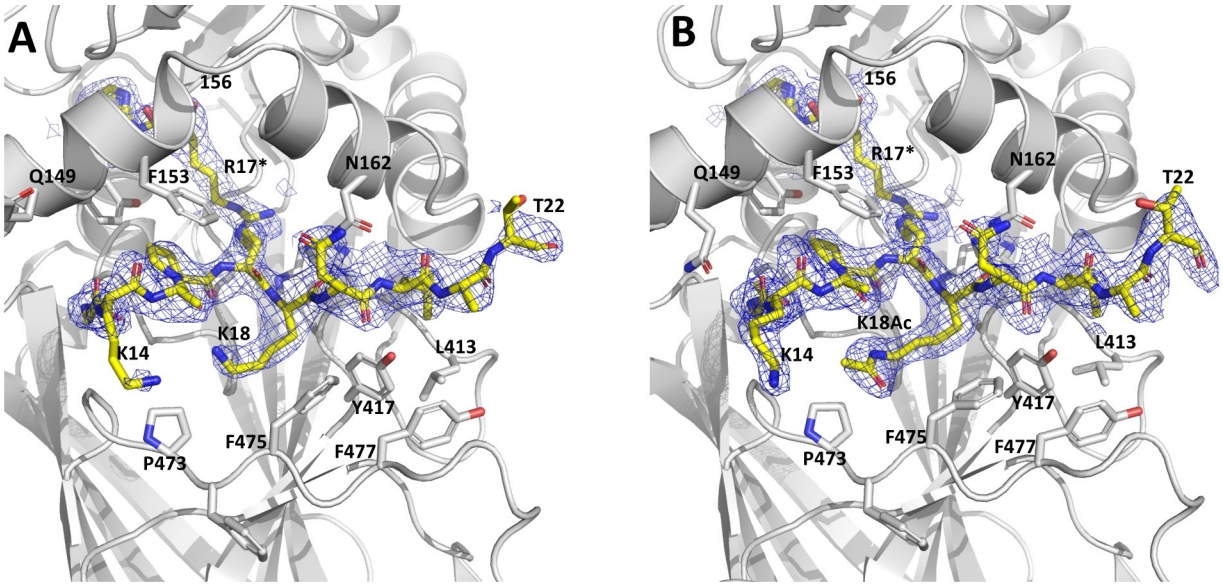


Figure 2. Electron density (2*F*obs-*F*calc) weighted maps for subunit A of mmCARM1 bound to: A) peptidomimetic **3** ($H3^{13-31}$ Lys^{18}–NH_2), PDB code 7OS4 and B) peptidomimetic **4** ($H3^{13-31}$ Lys^{18}Ac), PDB code 7OKP. CARM1 is represented as cartoon and H3 peptidomimetics are represented as stick. Maps are represented as a mesh contouring level set to 1σ.

tional structural information regarding long distance substrate interactions with CARM1, the $H3^{13-31}$ constructs did provide insights into the conformational behaviour of the substrate peptides and the impact of lysine acetylation. In keeping with previous reports, the peptide segments of transition-state mimetics **3** and **4** adopt a conformation similar to that observed in the structure of CARM1 bound to sinefungin and a linear $H3^{13-30}$ peptide (see Supporting Information Figure S1).[18]

Interestingly, little conformational change is observed for either CARM1 or the substrate peptidomimetics upon Lys^{18} acetylation (Figure 3 and Supporting Information Figure S2), with both the intra-peptide and peptide-CARM1 interactions observed with Lys^{18} peptidomimetic **3** largely maintained with Lys^{18}Ac peptidomimetic **4**. The conformation of the peptide is stabilized by an intra-peptide hydrogen bond between the Nz atom of Lys^{18} and the backbone oxygen of Ala^{15} and by additional Van der Waals interactions with Tyr^{262}, Tyr^{417} and Phe^{475} in the CARM1 active site (Figure 3). While subtle, acetylation of Lys^{18} does lead to some additional interactions: (i) a weak C–H–O hydrogen involving the O atom of the acetyl functional group and the Cβ atom of Phe^{475} and (ii) Van der Waals interactions with the proline ring of Pro^{473} and CH_3 group of Ala^{15} (Figure 3). In addition to these stabilizing interactions, acetylation of Lys^{18} may reduce the cost of desolvation of the peptides prior to binding and therefore produce an energetic gain in complex formation.

As noted above, Yue and co-workers have previously proposed that Lys^{18} acetylation stabilizes the transition state of

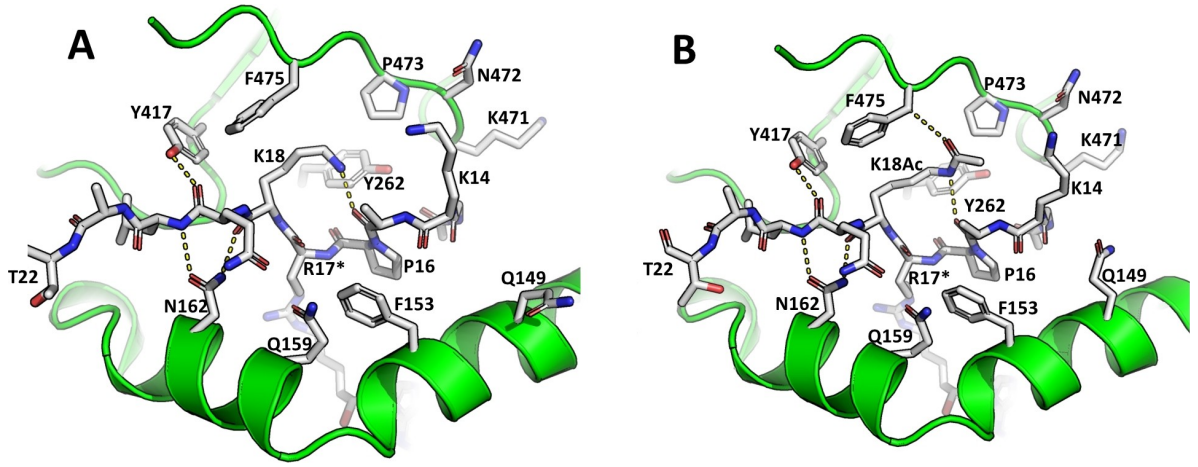


Figure 3. Recognition of peptidomimetics **3** and **4** by mmCARM1. Interactions shown for: A) compound **3** ($H3^{13-31}$ Lys^{18}–NH_2) PDB code 7OS4 and B) compound **4** ($H3^{13-31}$ Lys^{18}Ac) PDB code 7OKP. H-bonds are shown as dash lines with cartoon and stick representation of the peptidomimetics bound to mmCARM1.

the methylation transfer.[14] Our crystal structures do not, however, support this hypothesis as the side chain of Lys^{18} is found to be located more than 12 Å away from the active site centre. Rather, the structural data presented here indicate that gain in substrate affinity associated with lysine acetylation is likely due to additional interactions (van der Waals and weak hydrogen bonding) as well as a possible reduction of the desolvation penalty.

Informed by our structural findings obtained with the $H3^{13–31}$-based peptidomimetics **3** and **4**, we next prepared a series of smaller peptidomimetics and evaluated their inhibitory activity against CARM1. These peptidomimetics were centred around Arg^{17} which was again covalently linked to an adenosine group via its side chain guanidine moiety. Two peptidomimetics based on $H3^{10–25}$ (compounds **5** and **6**) were first prepared and assessed as inhibitors of CARM1 assessed (Table 1). The potent inhibition observed for both **5** and **6**, led us to also investigate shorter peptidomimetics by sequentially omitting N- and C-terminal residues to generate the corresponding deca-, octa-, hexa-, and tetra-peptide analogues **7**–**14**. Again, each of these truncated peptidomimetic were prepared with and without acetylation of the neighbouring Lys^{18} residue to probe the interplay between peptide sequence and lysine acetylation on recognition by/inhibition of CARM1. Inhibition studies subsequently revealed that all compounds retain potent inhibition with IC_{50} values in the nM range. Interestingly, the most potent inhibition measured was for the acetylated hexapeptide-based peptidomimetic **12**. This hexapeptide motif appears to be an optimum for achieving inhibition as either elongation to the octapeptide or truncation to the tetrapeptide was found to result in measurable increases in IC_{50} values. Interestingly, lysine acetylation also reduces the capacity of these peptidomimetics to engage with other PRMTs. To assess selectivity, peptidomimetics **5**–**14** were evaluated against PRMT1, which in all cases revealed a high degree of selectivity for CARM1 inhibition. These findings are in line with expectations given that the H3 peptide sequence used in this study is known to be methylated by CARM1 and not by PRMT1.[20]

As shown in Table 1, Lys^{18} acetylation led to a decrease in IC_{50} for compounds **6**, **10**, and **12** suggesting an increase in binding affinity. As noted above, in addition to stabilizing interactions with the enzyme active site, acetylation of Lys^{18} may reduce the cost of desolvation of the peptide prior to binding and therefore produce an energetic gain in complex formation. Notable is the potent inhibition obtained for hexapeptide **12** ($H3^{15–20}$ $K^{18}Ac$) which retains the main interactions with CARM1 and intra-peptide interactions revealed by our co-crystal structures. It is plausible that the larger peptidomimetics display a lowered inhibition/reduced affinity because they must pay a high desolvation penalty (particularly for Lys^{14}) in order to bind that is not compensated for by additional interactions with CARM1. We do note that in the case of decapeptide analogues **7** and **8** the finding that the acetylated species exhibits a slightly higher IC_{50} does not adhere to this explanation and remains to be understood. Our structural insights also provide an explanation for the reduced inhibition observed for the tetrapeptide analogues **13** and **14**: deletion of Ala^{15} is likely to significantly destabilize peptide binding as intra-peptide interactions between Ala^{15} and Lys^{18} (which stabilize the tight turn conformation of the peptide) are lost and in this context, acetylation of Lys^{18} is not sufficient to restore binding affinity. Also of note for peptidomimetics **5**–**14** is the finding that acetylation of Lys^{18} consistently results in an increased inhibitory selectivity towards CARM1 vs PRMT1 (Table 1). This finding points to the intriguing possibility that crosstalk between lysine acetylation and arginine methylation may also serve to reinforce PRMT specificity beyond the primary sequence of the peptide substrate.

While our studies provide new *in vitro* insights, the structural basis of crosstalk between $H3K^{18}$ acetylation and CARM1 methylation remains to be further elucidated *in vivo*. Notable in this regard is recent work by O'Malley and co-workers who combined cryo-electron microscopy and biochemical approaches in studying the ER-coactivator complex.[21] These investigations revealed that CARM1 recruitment induces p300 conformational change and promotes $H3K^{18}Ac$ and that increased histone $H3K^{18}$ acetylation in turn enhanced CARM1-mediated $H3R^{17}$ methylation.

Table 1. IC_{50} values and for compounds **5–14** against CARM1 and PRMT1.[a]

Compound		Peptidomimetic sequence	IC_{50} values [μM][a] CARM1	PRMT1
5	$H3^{10–25}$	Ac–STGGKAP**R***KQLATKAA–NH2	0.290 ± 0.015	> 2.5
6	$H3^{10–25}(K^{18}Ac)$	Ac–STGGKAP**R***K(Ac)QLATKAA–NH2	0.155 ± 0.007	> 5
7	$H3^{13–22}$	Ac–GKAP**R***KQLAT–NH2	0.121 ± 0.007	> 5
8	$H3^{13–22}(K^{18}Ac)$	Ac–GKAP**R***K(Ac)QLAT–NH2	0.155 ± 0.012	> 5
9	$H3^{14–21}$	Ac–KAP**R***KQLA–NH2	0.287 ± 0.034	> 2.5
10	$H3^{14–21}(K^{18}Ac)$	Ac–KAP**R***K(Ac)QLA–NH2	0.211 ± 0.023	> 25
11	$H3^{15–20}$	Ac–AP**R***KQL–NH2	0.143 ± 0.014	> 2.5
12	$H3^{15–20}(K^{18}Ac)$	Ac–AP**R***K(Ac)QL–NH2	**0.072 ± 0.008**	> 25
13	$H3^{16–19}$	Ac–P**R***KQ–NH2	0.346 ± 0.031	> 5
14	$H3^{16–19}(K^{18}Ac)$	Ac–P**R***K(Ac)Q–NH2	0.699 ± 0.081	> 25

[a] IC_{50} values reported in μM from duplicate data obtained from a minimum of 7 different concentrations ± standard error of the mean (s.e.m.). Corresponding K_i values are also provided in the Supporting Information, see Table S2. The **R*** indicates the Arg^{17} residue where the adenosine group is incorporated.

Conclusion

We here report the use of peptide-based transition state mimetics centred around the Arg^{17}/Lys^{18} of the histone H3 tail peptide to study crosstalk between lysine acetylation and arginine methylation and its impact on substrate recognition by CARM1. Structural studies with these peptidomimetics and the catalytic domain of CARM1 reveal that little conformational change is observed in the protein and on the peptide substrates conformations upon Lys^{18} acetylation. Rather, the increase in affinity associated with Lys^{18} acetylation is likely due to additional weak interactions with mmCARM1, intra-peptide interactions that stabilize the active conformation of the substrate peptide, and a possible reduction of the desolvation cost associated with substrate binding when Lys^{18} is acetylated. Building from these findings, shorter peptidomimetics were also synthesized and evaluated as CARM1 inhibitors. The truncation approach used led to the discovery of potent inhibitors containing only two residues flanking the central Arg-Lys pair on either side with peptidomimetics **11** and **12** exhibiting IC_{50} values of 143 and 72 nM respectively. Taken together, the findings reported in this study provide valuable new insights both into the mechanistic understanding of crosstalk and its role in CARM1 mediated methylation as well as in the design of potent CARM1-selective peptidomimetic inhibitors.

Experimental Section

General procedures

All reagents employed were of American Chemical Society grade or finer and were used without further purification unless otherwise stated. The final compounds were purified via preparative HPLC performed on a BESTA-Technik system with a Dr. Maisch Reprosil Gold 120 C18 column (25×250 mm, 10 μm) and equipped with a ECOM Flash UV detector monitoring at 214 nm. The following solvent system, at a flow rate of 12 mL/min, was used: solvent A: 0.1% TFA in water/acetonitrile 95/5; solvent B: 0.1% TFA in water/acetonitrile 5/95. Gradient elution was as follows: 95:5 (A/B) for 5 min, 95:5 to 0:100 (A/B) over 40 min, 0:100 (A/B) for 5 min, then reversion back to 95:5 (A/B) over 2 min, 95:5 (A/B) for 8 min.

Purity was confirmed to be ≥95% by LCMS performed on a Shimadzu LC-20AD system with a Shimadzu Shim-Pack GIST-AQ C18 column (3.0×150 mm, 3 μm) at 30 °C and equipped with a UV detector monitoring at 214 and 254 nm. This system was connected to a Shimadzu 8040 triple quadrupole mass spectrometer (ESI ionisation). The following solvent system, at a flow rate of 0.5 mL/min, was used: solvent A, 0.1% formic acid in water; solvent B, acetonitrile. Gradient elution was as follows: 95:5 (A/B) for 2 min, 95:5 to 0:100 (A/B) over 23 min, 0:100 (A/B) for 1 min, then reversion back to 95:5 (A/B) over 1 min, 95:5 (A/B) for 3 min.

HRMS analyses were performed on a Shimadzu Nexera X2 UHPLC system with a Waters Acquity HSS C18 column (2.1×100 mm, 1.8 μm) at 30 °C and equipped with a diode array detector. The following solvent system, at a flow rate of 0.5 mL/min, was used: solvent A, 0.1% formic acid in water; solvent B, 0.1% formic acid in acetonitrile. Gradient elution was as follows: 95:5 (A/B) for 1 min, 95:5 to 15:85 (A/B) over 6 min, 15:85 to 0:100 (A/B) over 1 min, 0:100 (A/B) for 3 min, then reversion back to 95:5 (A/B) for 3 min. This system was connected to a Shimadzu 9030 QTOF mass spectrometer (ESI ionisation) calibrated internally with Agilent's API-TOF reference mass solution kit (5.0 mM purine, 100.0 mM ammonium trifluoroacetate and 2.5 mM hexakis(1*H*,1*H*,3*H*-tetrafluoropropoxy)phosphazine) diluted to achieve a mass count of 10000.

Synthetic procedures

Compounds **1**–**14** were synthesized by using a methodology developed in our group enabling the on-resin preparation of peptides containing substituted arginine residues.[15] Specifically, Histone H3-derived peptides were synthesized by using standard Fmoc solid-phase peptide synthesis (SPPS) techniques after which the adenosine group was introduced. The peptides were synthesized on 0.1 mmol scale using Rink Amide AM resin (146 mg with a resin loading of 0.684 mmol/g). The arginine in the sequence was replaced with an alloc-protected ornithine. The lysine was introduced as Fmoc-Lys(Boc)–OH to obtain the free lysine or as Fmoc-Lys(Ac)–OH to obtain the peptides with the acetylated lysine residue. Peptide couplings were performed using standard Fmoc amino acids (4.0 eq), BOP (4.0 eq) and DiPEA (8.0 eq) in DMF (7.5 mL) at ambient temperature for 1 hour. The Fmoc deprotection was performed in two runs by using 20% piperidine in DMF (6 mL) for 5 minutes and 30 minutes, consecutively. After SPPS, the *N*-terminus was acetylated on resin using acetic anhydride (0.5 mL) and DiPEA (0.85 mL) in DMF (10 mL) for 1 hour at room temperature with nitrogen bubbling. The peptides were kept on the resin for next step.

The peptides were Alloc-deprotected on the resin using tetrakis (triphenylphosphine)-palladium(0) and phenylsilane in DCM following a literature procedure.[22] Upon the completion of Alloc-deprotection, the adenosine thiourea building block[15] (105 mg, 0.13 mmol, 1.3 eq) was coupled to the amine group of ornithine side-chain using 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDCI) (34.5 mg, 0.15 mmol, 1.5 eq) in DCM (10 mL). The mixture was stirred for 1.5 hours at room temperature, drained and the resin was washed with DCM (3×10 mL), DMF (3×10 mL) and DCM (2×10 mL). Peptides were deprotected and cleaved from the resin using cleavage cocktail (TFA/TIPS/H_2O 95:2.5:2.5). Precipitation in MTBE/Petroleum ether (1:1) yielded the crude peptide, which was purified by preparative HPLC. The purity and identity were confirmed by analytical HPLC and High-resolution Mass Spectrometry, the results of which are presented in the Supporting Information for all final compounds.

Enzymatic activity assays

The commercially available PRMT1 and CARM1 chemiluminescent assay kits (BPS Bioscience, Dan Diego, CA, USA) were used for evaluation of methyltransferase inhibition as previously described.[22] The enzymatic reactions were performed in duplicate at room temperature using 96-wells plates precoated with histone substrates. The reaction volume is 50 μL containing proprietary assay buffer, 20 μM SAM, enzyme: PRMT1 (10 ng per reaction) and CARM1 (200 ng per reaction). Against CARM1, the inhibitors were dissolved in water and tested at varying concentration ranging from 0.0128 to 200 μM. For selectivity, inhibitors were tested against PRMT1 at three fixed concentrations (2.5, 5 and 25 μM). Positive controls were performed by addition of water instead of inhibitor. Blank and substrates controls were performed in the absence of enzyme and SAM, respectively. Before the reactions were initiated by the addition of SAM, the inhibitors were incubated with the enzyme for 15 min at room temperature. After

incubation for one hour with PRMT1 or two hours with CARM1, the wells were washed and blocked and incubated with primary antibody (1:100) for 1 h. After washing and blocking, the wells were incubated with secondary HRP-labelled antibody (1:1000) for 30 minutes. After a final washing and blocking, the HRP chemiluminescent substrate mixture was added to the wells and the luminescence was measured immediately using a Tecan spark plate reader. All the measurements were performed in duplicate and the data was analysed using GraphPad Prism 9.

All the luminescence data were corrected with the blank values and the data was subsequently normalized with the highest value in the concentration range defined as 100% activity. The percentage of inhibition activity was plotted as a function of inhibitor concentration and fit using non-linear regression analysis of the sigmoidal dose -response curve generated using the normalized data and a variable slope following Equation (1):

$$Y = \frac{100}{(1 + 10^{((logIC50 - X) \times Hill\ Slope)})} \quad (1)$$

where Y = percentage activity, X = the logarithmic concentration of the inhibitors, Hill Slope = slope factor or Hill coefficient. The IC_{50} value was determined by the half maximal inhibitory concentration. The IC_{50} values measured for SAH, which served as a reference compound, are similar to those reported.[23] Full IC_{50} curves and comparative Ki values for compounds **5**–**14** and SAH are presented in the Supporting Information.

CARM1 cloning, expression, and purification

The *Mus musculus* CARM1 gene sequence corresponding to the PRMT core (residues 130 to 497, mmCARM1$^{130–497}$) were amplified by PCR from the original GST-CARM1 construct.[24] The sequences were cloned in the pDONR207™ (Invitrogen) vector using a BP reaction (Gateway® Cloning, Life Technologies). The positive clones were confirmed by sequencing (GATC). The sequences were subcloned in a pDEST20™ vector using a LR reaction. The resulting recombinant protein harbour an amino-terminal glutathione S-transferase (GST) tag followed by a Tobacco etch virus (TEV) protease cleavage site. DH10Bac competent cells containing the baculovirus genome were transformed with the pDEST20™-CARM1 plasmids and plated onto LB agar media containing 15 mg.mL^{-1} tetracycline, 7 mg.mL^{-1} gentamicin, 50 mg.mL^{-1} kanamycin, 25 mg.mL^{-1} X-Gal and 40 mg mL^{-1} IPTG. Bacmid DNA purified from recombination-positive white colonies was transfected into Sf9 cells using the Lipofectin reagent (Invitrogen). Viruses were harvested 10 days after transfection. Sf9 cells were grown at 300 K in suspension culture in Grace medium (Gibco) using Bellco spinner flasks. 1 L of sf9 cell culture (at 0.8×10^6 cells.mL^{-1}) was infected with recombinant GST-mmCARM1 virus with an infection multiplicity of 1. Cells were harvested 48 h post-infection. Cell lysis was performed by sonication in 50 mL buffer *A* [50 mM Tris-HCl pH 8.0, 250 mM NaCl, 5% glycerol, 5 mM TCEP, 0.01% NP40 and anti-proteases (Roche, Complete™, EDTA-free)] and cellular debris were sedimented by centrifugation of the lysate at 40,000 × *g* for 30 min. The supernatant was incubated overnight at 277 K with 2 mL glutathione Sepharose resin (GE Healthcare). After a short centrifugation, the supernatants were discarded, and the beads were poured in an Econo-column (Bio-Rad). After two wash steps with 10 mL buffer *A*, 2 mL buffer *A* supplemented with in-house produced TEV protease were applied to the columns and digestion was performed 4 hours at 303 K with gentle mixing. The digest was concentrated with an Amicon Ultra 10 K (Milipore), loaded on a gel-filtration column (HiLoad 16/60 Superdex S200, GE Healthcare) and eluted at 1 mL.min^{-1} with buffer *B* [20 mM Tris-HCl pH 8.0, 100 mM NaCl, 1 mM TCEP] using an ÄKTA Purifier device (GE Healthcare). Fractions containing mmCARM1$^{130–497}$ were pooled and concentrated to 7.75 mg.mL^{-1}.

Crystallization

Transition state mimics were solubilized in water before addition to the protein solution (2 mg.mL^{-1}) at the final concentration of 2 mM. The protein-peptide solution was incubated 30 minutes at room temperature before use. Vapor diffusion method utilizing hanging drop trays with a 0.5 mL reservoir was used for crystallization. Typically, 2 μL of protein-ligand solution were added to 1 μL of well solution consisting of 1–1.5 M disodium malonate, 100 mM MES pH 5.5–7 and 200 mM NaCl. Crystals grew in a few days at 293 K.

X-ray structure determination

Crystals were flash-frozen in liquid nitrogen after a brief transfer to 5 μL reservoir solution containing 25% (v/v) Glycerol as a cryoprotectant and were stored in liquid nitrogen. The diffraction data sets were collected using CBI X-ray home source (Rigaku FR–X and EIGER 4 M), SOLEIL PROXIMA1 and ESRF ID30-B beamlines, using a Pilatus 6 M, EIGER 4 M, EIGER X4M (Dectris) detector and processed with XDS[25] and HKL-2000.[26] The crystals belonged to the $P2_12_12$ space group with four monomers of CARM1 in the asymmetric unit. The structures were solved by molecular replacement using CARM1 structure as a probe.[16] Model building and refinement were carried out using Coot[27] and PHENIX.[28] TLS refinement with 6 groups per polypeptide chain was used. All other crystallographic calculations were carried out with the CCP4 package.[29] Structure figures were generated with PyMOL (http://www.pymol.org).

Supporting information

The Supporting Information document contains molecular formula strings, analytical data including HRMS values and HPLC traces, IC_{50} curves, and supplemental table and figures for structural studies.

Abbreviations

Alloc, allyloxycarbonyl; BOC, *tert*-butyloxycarbonyl; CARM1, coactivator associated arginine methyltransferase 1; LC-MS, liquid chromatography mass spectrometry; Pbf, 2,2,4,6,7-pentamethyl-dihydro-benzofuran-5-sulfonyl, PEG, polyethylene glycol; PRMT, protein arginine methyltransferase; PTM, post-translational modification; RP-HPLC, reversed phase high performance liquid chromatography; SAH, *S*-adenosyl-L-homocysteine; SAM, *S*-adenosyl-L-methionine; SPPS, solid-phase peptide synthesis; SUMO, Small Ubiquitin-like Modifier; TFA, trifluoroacetic acid.

Data availability

The atomic coordinates and experimental data have been deposited at the Protein Data Bank (CARM1-H3$^{13–31}$ K_{18}Ac =

7OKP, CARM1-H3$^{13-31}$ K_{18}-NH2 = 7OS4). The authors will release the atomic coordinates and experimental data upon article publication.

Acknowledgements

Y.Z. gratefully acknowledges the financial support provided by a scholarship from the Chinese Scholarship Council (CSC, file No. 201706210082). J.C., V.C., N.M., and N.T-C. are supported by grants from CNRS, Université de Strasbourg, INSERM, Instruct-ERIC, part of the European Strategy Forum on Research Infrastructures (ESFRI) supported by national member subscriptions as well as the French Infrastructure for Integrated Structural Biology (FRISBI) [ANR-10-INSB-005, grant ANR-10-LABX-0030-INRT, a French State fund managed by the Agence Nationale de la Recherche under the frame program Investissements d'Avenir labelled ANR-19-CE11-0010-01 JC and IGBMC and grants from Association pour la Recherche contre le Cancer (ARC) (ARC 2016, n° PJA 20161204817) and grants from "Ligue d'Alsace contre le Cancer". We thank Paolo Innocenti for performing HRMS analyses and Luc Bonnefond and Alastair McEwen for help in the structural studies. We thank members of SOLEIL Proxima1 beamlines and the European Synchrotron Radiation Facility-European Molecular Biology Laboratory joint Structural Biology groups for the use of beamline facilities and for assistance during X-ray data collection.

Conflict of Interest

The authors declare no competing financial interest.

Keywords: arginine methylation · CARM1 · crosstalk · crystallography · inhibitors · peptidomimetics

[1] T. Hunter, *Mol. Cell.* **2007**, *28*, 730–738.
[2] A. S. Venne, L. Kollipara, R. P. Zahedi, *Proteomics* **2014**, *14*, 513–524.
[3] W. S. Lo, R. C. Trievel, J. R. Rojas, L. Duggan, J. Y. Hsu, C. D. Allis, R. Marmorstein, S. L. Berger, *Mol. Cell.* **2000**, *5*, 917–926.
[4] K. P. Nightingale, S. Gendreizig, D. A. White, C. Bradbury, F. Hollfelder, B. M. Turner, *J. Biol. Chem.* **2007**, *282*, 4408–4416.
[5] L. Li, Y. Wang, *J. Biol. Chem.* **2017**, *292*, 11951–11959.
[6] O. Castillo-Aguilera, P. Depreux, L. Halby, P. B. Arimondo, L. Goossens, *Biomol. Eng.* **2017**, *7*.
[7] Y. Shi, M. J. Long, M. M. Rosenberg, S. Li, A. Kobjack, P. Lessans, R. T. Coffey, L. Hedstrom, *ACS Chem. Biol.* **2016**, *11*, 3328–3337.
[8] F. Wojcik, G. P. Dann, L. Y. Beh, G. T. Debelouchina, R. Hofmann, T. W. Muir, *Nat. Commun.* **2018**, *9*, 1394.
[9] R. S. Blanc, S. Richard, *Mol. Cell.* **2017**, *65*, 8–24.
[10] Z. Zhao, A. Shilatifard, *Genome Biol.* **2019**, *20*, 245.
[11] B. M. Lorton, D. Shechter, *Cell. Mol. Life Sci.* **2019**, *76*, 2933–2956.
[12] S. Daujat, U. M. Bauer, V. Shah, B. Turner, S. Berger, T. Kouzarides, *Curr. Biol.* **2002**, *12*, 2090–2097.
[13] Z. Zhang, B. C. Nikolai, L. A. Gates, S. Y. Jung, E. B. Siwak, B. He, A. P. Rice, B. W. O'Malley, Q. Feng, *Nucleic Acids Res.* **2017**, *45*, 9348–9360.
[14] W. W. Yue, M. Hassler, S. M. Roe, V. Thompson-Vale, L. H. Pearl, *EMBO J.* **2007**, *26*, 4402–4412.
[15] M. J. van Haren, N. Marechal, N. Troffer-Charlier, A. Cianciulli, G. Sbardella, J. Cavarelli, N. I. Martin, *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 3625–3630.
[16] N. Troffer-Charlier, V. Cura, P. Hassenboehler, D. Moras, J. Cavarelli, *EMBO J.* **2007**, *26*, 4391–4401.
[17] J. Mailliot, *Étude Structurale de l'histoneméthyltransférase "CARM1" et de Ses Complexes Biologiquement Significatifs: Des Structures 3D Vers La Conception Rationnelle de Composés à Action Pharmacologique, PhD Thesis*, University of Strasbourg. **2013**.
[18] P. A. Boriack-Sjodin, L. Jin, S. L. Jacques, A. Drew, C. Sneeringer, M. P. Scott, M. P. Moyer, S. Ribich, O. Moradei, R. A. Copeland, *ACS Chem. Biol.* **2016**, *11*, 763–771.
[19] N. Marechal, *Etude Structurale Des Protéines Arginine Méthyltransferases, PhD Thesis*, University of Strasbourg, **2018**.
[20] B. T. Schurter, S. S. Koh, D. Chen, G. J. Bunick, J. M. Harp, B. L. Hanson, A. Henschen-Edman, D. R. Mackay, M. R. Stallcup, D. W. Aswad, *Biochemistry* **2001**, *40*, 5747–5756.
[21] P. Yi, Z. Wang, Q. Feng, C. K. Chou, G. D. Pintilie, H. Shen, C. E. Foulds, G. Fan, I. Serysheva, S. J. Ludtke, M. F. Schmid, M. C. Hung, W. Chiu, B. W. O'Malley, *Mol. Cell* **2017**, *67*, 733–743 e734.
[22] Y. Zhang, M. J. van Haren, N. I. Martin, *Methods* **2020**, *175*, 24–29.
[23] M. van Haren, L. Q. van Ufford, E. E. Moret, N. I. Martin, *Org. Biomol. Chem.* **2015**, *13*, 549–560.
[24] D. Chen, H. Ma, H. Hong, S. S. Koh, S. M. Huang, B. T. Schurter, D. W. Aswad, M. R. Stallcup, *Science* **1999**, *284*, 2174–2177.
[25] W. Kabsch, *Acta Crystallogr. Sect. D* **2010**, *66*, 125–132.
[26] Z. Otwinowski, W. Minor, *Methods Enzymol.* **1997**, *276*, 307–326.
[27] P. Emsley, K. Cowtan, *Acta Crystallogr. Sect. D* **2004**, *60*, 2126–2132.
[28] P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, *Acta Crystallogr. Sect. D* **2010**, *66*, 213–221.
[29] M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin, K. S. Wilson, *Acta Crystallogr. Sect. D* **2011**, *67*, 235–242.

Manuscript received: September 23, 2021
Revised manuscript received: September 25, 2021
Accepted manuscript online: September 27, 2021
Version of record online: October 14, 2021