**Structural insights into SETD3-mediated histidine methylation on β-actin**

Qiong Guo[1,#], Shanhui Liao[1,#,*], Weronika Tomaka[2], Sebastian Kwiatkowski[2], Gao Wu[1], Jakub Drozak[2], Jinrong Min[3,4], Chao Xu[1,*]

1. Hefei National Laboratory for Physical Sciences at Microscale and School of Life Sciences, University of Science and Technology of China, Hefei, Anhui, 230027, People's Republic of China.

2. Department of Metabolic Regulation, Faculty of Biology, University of Warsaw, Warsaw, Poland.

3. Structural Genomics Consortium, University of Toronto, 101 College St., Toronto, Ontario, M5G 1L7, Canada.

4. Department of Physiology, University of Toronto, Toronto, Ontario, M5S 1A8, Canada.

#: equal contribution

Running title: The structure of SETD3 bound to a β-actin fragment

*: To whom correspondence should be addressed.

Chao Xu, xuchaor@ustc.edu.cn; Shanhui Liao, ajsod@mail.ustc.edu.cn.

**Keywords:** X-ray crystallography, post translational modifications, β-actin, SET domain, $N^3$-methylhistidine

## ABSTRACT

SETD3 is a member of SET (Su(var)3-9, Enhancer of zeste, and Trithorax) domain involved in hypoxic pulmonary hypertension by reducing VEGF expression, muscle differentiation, and carcinogenesis. In a previous paper (Jakub Drozak *et al.*, eLife 2018, accepted), we identified SETD3 as an actin-specific methyltransferase that methylates the $N^3$ of His73 on β-actin. Here we presented two complex structures of AdoHcy-bound SETD3, with one bound to an unmodified β-actin peptide and the other bound to the methylated peptide. Structural analysis indicated that the recognition and methylation of β-actin by SETD3 is highly sequence specific. Besides, both SETD3 and β-actin adopt conformational changes upon binding to each other. Our structural research, supported by the enzyme assay, allows us to propose a catalytic mechanism of SETD3-mediated histidine methylation on β-actin, which would also facilitate the design of small molecule inhibitors for SETD3.

## INTRODUCTION

Microfilaments, as the building blocks of cytoskeleton, are composed of actin proteins (dos Remedios et al., 2003; Theriot and Mitchison, 1991). Six different isoforms of actin are identified in mammals, which could be divided into three groups: alpha, beta and gamma (Gunning et al., 1983; Herman, 1993). Among the three types, β-actin is ubiquitously expressed and plays important roles in forming cytoskeleton, mediating cell motility and maintaining cell stability (Leterrier et al., 2017; Nudel et al., 1983).

Many different types of post-translational modifications (PTM) are found in actin, such as acetylation, methylation, SUMOylation, ubiquitination, etc. (Terman and Kashina, 2013). $N^3$-methylation of His73 in β-actin is of special interest since its identification in 1967 (Johnson et al., 1967), not only because it is highly conserved in almost all eukaryotic actins except the actin from *Naegleria gruberii* (Sussman et al., 1984), but also because it affects hydrolysis rate of the actin-bound ATP through mediating the inter-domain mobility around the terminal phosphate of the ATP molecule (Kabsch et al., 1990), evidenced by the higher ATP exchange rate caused by the actin-H73A mutant (Nyman et al., 2002). Aside from actin proteins, $N^3$-methylhistidine is also reported to

exist in other proteins, such as myosin (Asatoor and Armstrong, 1967) and yeast Rpl3 (Webb et al., 2010). The effects of post translational modifications on proteins could be classified as follows: some modifications lead to a conformational change of the target molecule, while others may exclude existing binding partners and/or recruit new ligands (Bonasio et al., 2010). Methylation of the His73 residue of β-actin was reported to influence polymerization of the actin molecules and rate of ATP hydrolysis (Nyman et al., 2002; Terman and Kashina, 2013). Although some actin methyltransferases had been identified decades ago (Raghavan et al., 1992; Vijayasarathy and Rao, 1987), whether these enzymes specifically catalyze the methylation of His73 of β-actin, and the molecular mechanism of $N^3$-methylation on His73, are largely unknown.

In all types of methylation, the molecular mechanism, as well as the biological functions, of histone lysine methylation has been well established because of its crucial roles in gene regulation and cell development (Martin and Zhang, 2005; Sims et al., 2003). Histone lysine methylation is reversible and dynamically controlled by histone methyltransferases and demethylases (Whetstine et al., 2006). Histone lysine N-methyltransferases (HKMTs), which install the methyl moiety on histone lysines, are also found to catalyze methylation of specific lysine residues of non-histone proteins (Hamamoto et al., 2015; Rathert et al., 2008). So far all lysine methyltransferases except DOT1L, the histone H3K79 methyltransferase, contain a SET domain (Min et al., 2003). The SET domain is a ~130 aa motif and is initially identified in Trithorax, with its acronym derived from **S**u(var)3-9, **E**nhancer of zeste, and **T**rithorax (Dillon et al., 2005). It utilizes AdoMet as the co-factor by transferring the methyl moiety of AdoMet to substrates.

There are ~50 SET domain proteins in the human genome (Dillon et al., 2005) and whether all of them act as lysine methyltransferases, are largely unknown. SETD3 is a protein abundantly expressed in muscles and its knock down impairs differentiation of muscle cells (Eom et al., 2011). Previously, SETD3 has been identified as a histone methyltransferase that methylates H3K4 and H3K36 based on bioinformatics and

peptide analyses (Eom et al., 2011; Wagner and Carpenter, 2012) and recently it is also found to be associated with lymphoma (Chen et al., 2013). After extensive screening on those putative methyltransferases, we found that SETD3 specifically recognized a His73containing fragment of β-actin, and catalyzed the N3-methylation of His73. Furthermore, it has been recently postulated that SETD3 is indeed an actin-specific $N^3$-histidine methyltransferase *in vivo* (https://doi.org/10.1101/266882).

To understand how SETD3 methylates β-actin, we crystallized and solved two structures of AdoHcy-bound SETD3, one with the peptide containing unmodified His73 and the other with the same peptide containing methylated His73 (His73me). On the basis of our two β-actin bound SETD3 complex structures, we confirmed that SETD3 is a histidine methyltransferase and it specifically methylates His73 of β-actin. By structural, mutagenesis and enzymatic analysis, we pinpointed some key active residues of SETD3 important for β-actin His73 methylation and substrate selectivity, and uncovered its methylation mechanism. This study would not only provide insights into the role of SETD3 in mediating the function of actins, but also guide the future design of small molecules towards SETD3.

## RESULTS

### SETD3 binds to and methylates β-actin

Since SETD3 has been identified as a histone lysine methyltransferase catalyzing methylation of H3K4 and H3K36, we purified the core region of SETD3 (aa 2-502) (Figure 1A), and synthesized two histone H3 peptides, containing histone H3K4 and H3K36 residues, respectively. Surprisingly, our ITC binding results revealed that neither of two peptides displayed detectable binding to SETD3 (Supplementary table S1). Furthermore, we also checked the methylation activity of SETD3 on H3K4 (1-10) and H3K36 (31-44) peptides by the mass spectrometry technique, and found that in contrast to β-actin, neither of these two peptides could be methylated by SETD3, at least under our experimental conditions (Supplementary Figure S1A-S1B).

SETD3 has been shown to play an important role in muscle cell differentiation (Eom et al., 2011). Besides, the interactions between actin and myosin are also essential for smooth muscle cell differentiation (Skalli et al., 1986). Therefore, β-actin His73 methylation by SETD3, which are critical in mediating actin polymerization, directly link the methylation activity of SETD3 with actin functions. We set out to explore if β-actin acts as the substrate of SETD3. Markedly, we found by ITC that SETD3 indeed displayed high binding affinity towards a His73-containing peptide of β-actin (aa 66-88) (Kd = 0.17 μM) (Figure 1B), and our mass spectrometry based enzymatic assay shows that this peptide could be efficiently methylated (Supplementary Figure S1C). The same β-actin peptide with His73 mutated to alanine (H73A) binds to SETD3 only slightly weaker (Kd = 0.45 μM)(Figure 1B), however, we failed to detect any methylated products from this mutant peptide (Supplementary Figure S1D), implicating that it is His73 of β-actin, which was specifically methylated by SETD3.

**Overall structure of SETD3**

To uncover the underlying recognition and methylation mechanisms of β-actin by SETD3, we tempted to crystalize SAM-bound SETD3 (2-502) with full length β-actin, but failed to get any diffractable crystals. Since SETD3 bound to β-actin peptide (66-88) and methylated His73 on it, we crystallized and determined a 1.95 Å structure of SAM-bound SETD3 (2-502) with β-actin (aa 66-88) (Table 1). There are 4 molecules in a crystallographic asymmetric unit and residues 22-501 of SETD3 and 66-84 of β-actin are visible in the complex structures (Supplementary Figure S2A). Markedly, based on the density map, we found that His73 of β-actin had been methylated in the complex structure (Supplementary Figure S3A), although we used an unmethylated peptide. The reason could be that SETD3 bound to the methyl donor AdoMet from the E. coli cells, which we used to express our protein, and the AdoMet-bound SETD3 methylated His73 during crystallization. Therefore, this complex represents the snapshot of post-methyl transfer state, which prompts us to crystallize the substrate bound enzyme next.

To avoid the methylation reaction, SETD3 was purified with adding 5-fold excess of AdoHcy in the buffer to compete off AdoMet. Then the unmethylated peptide was mixed with the purified SETD3 to form the complex. The complex was crystallized and solved at a resolution of 2.15 Å (Table 1). In this complex structure, the density map of the peptide indicates that it was unmodified, suggesting the methylation did not occur (Supplementary Figure S3B). The complex structure hereby reflects the image of the pre-methyl transfer state.

In both structures, the SETD3 core region encompassing N-SET (N-terminal to the SET domain), SET, iSET (an insertion in the SET domain) and C-SET (C-terminal to the SET domain) domains, adopts a V-shape architecture (Figure 1C). The N-terminal lobe consists of the N-SET (α1-α3), SET (β1-β12) and iSET (α4-α11) motifs (Supplementary Figure S4A). The N-SET motif consists of three α helices (α1-3), and encompasses the SET and iSET motifs (Supplementary Figure S4A). While α1 and α2 pack against α10-α11 of iSET to form a 4-helix bundle, α3 of N-SET is localized at the SET-iSET interface, contacts the anti-parallel β strands (β1-β2) of SET and α8 of iSET (Figure 1C). The SET domain adopts a canonical SET domain fold similar to that of LSMT (Trievel et al., 2002), with twelve β strands arranged into four anti-parallel β sheets ( β 1-β2, β3-β11, β4-β10-β9, β5-β7-β6) and one parallel β sheet (β8-β12) (Supplementary Figures S4A and S5). The iSET domain is a helical region (α4-α11) inserted between β5 and β6 of the SET domain (Supplementary Figures S4A and S5). The C-terminal lobe, only consisting of the C-SET domain, is almost entirely helical (α12-α19) except two anti-parallel strands (β13-β14), with α12 and the N-terminal end of α19 packed against α10 and α9 of iSET, respectively (Supplementary Figures S4A and S5). Consistently, database search using the DALI server (Holm and Rosenstrom, 2010) revealed that the overall fold of SETD3 is highly similar to those of LSMT (Z-core: 31.4, RMSD: 3.8 Å for 425 Cα atoms) and SETD6 (Z-core: 28.7, RMSD: 2.8 Å for 421 Cα atoms) despite the low sequence identities (24%-25%).

**AdoHcy binding pocket within SETD3**

In both complex structures, SETD3 binds AdoHcy in a cleft formed between SET and C-SET and buttressed by iSET at the bottom (Figure 1C). Specifically, the adenine ring of AdoHcy is sandwiched between the side chain of Glu104 and the aromatic ring of Phe327, forming stacking and π-π interactions, respectively (Figure 2A). The Ade $N^6$ and $N^7$ atoms are hydrogen bonded to the main chain carbonyl and amide groups of His279, respectively, while the Ade $C^8$ atom forms one C-H...O hydrogen bond with the hydroxyl group of Tyr313 (Figure 2A).

In addition to the adenine specific interactions, ribose O3' and O4' atoms of AdoHcy form hydrogen bonds with the main chain carbonyl of Ser325 and the side chain carboxyl of Asn278, respectively (Figure 2A). The side chain carboxyl of Asn278 makes another hydrogen bond with the AdoHcy amide, while the latter is also H-bonded to the main chain carbonyl of Phe106 (Figure 2A). The AdoHcy carboxylate forms four hydrogen bonds with the SETD3 residues, one with the main chain amide of Phe106, one with the guanidino group of Arg254, and the other two with the guanidino group of Arg75 (Figure 2A). All the AdoHcy binding residues are conserved in other SETD3 orthologs, suggesting a conserved function of SETD3 (Supplementary Figure S4A).

The AdoHcy binding mode of SETD3 is similar to those observed in some other SET domain structures, such as those of LSMT (Figure 2B) (Trievel et al., 2002) and SETD6 (Figure 2C) (Chang et al., 2011). Among those AdoHcy binding residues, Asn278, His279 and Phe327 are conserved in LSMT and SETD6, while Glu104 and Arg254 are conserved only in LSMT, but not in SETD6 (Figures 2B-2C and Supplementary Figure S4B). Arg75 is a unique AdoHcy binding residue only found in SETD3 (Supplementary Figure S4B).

**β-actin binding mode of SETD3**

In both complexes, the β-actin peptide lies in a narrow groove formed by SET, iSET and C-SET, with the unmodified or methylated His73 residue accommodated in a

pocket which is built by α11 of iSET, β6 and β12 of SET, and α12 of C-SET, and further buttressed by η5 of SET (Figures 1C and 3A). When we compared the ternary complex structure of SETD3 with AdoHcy and unmodified β-actin peptide with the previously determined AdoMet-SETD3 binary structure (PDB id: 3SMT), we found that, despite the overall similarity between these two structures (with a RMSD of 0.66 Å in Cα positions), the region containing two β-sheets of the SET domain (β4-β10-β9 and β3-β11) takes a conformational change upon peptide binding, due to its direct interactions with the N-terminal end of the peptide, with the three loops preceding β4, β9 and β11 shift 4.2 Å, 4.4 Å and 6.8 Å towards the peptide, respectively (Supplementary Figure S6). The detailed interactions between SETD3 and the N-terminal end of the peptide will be described in more detail next.

**A unique histidine recognition pocket in SETD3 confers SETD3 its histidine methyltransferase activity**

In the substrate complex, residues Leu67-Glu83 of the peptide are visible, with His73 occupying in a hydrophobic pocket (Figure 3A). The main chain of His73 forms several hydrogen bonds with SETD3, with its main chain amide, Cα and main chain carbonyl H-bonded to the main chain carbonyl of Tyr313, the side chain carboxyl of Asn256, and the guanidino of Arg316, respectively (Figure 3A). The imidazole ring of His73 is parallel to the aromatic ring of Tyr313, with its orientation determined by two hydrogen bonds, one between the $N^1$ atom of the ring and the guanidino of Arg316 mediated by a water molecule, the other between the $N^3$ atom and the main chain carbonyl of Asp275 (Figure 3A). In addition to the stacking interactions with Tyr313, His73 also makes hydrophobic contacts with Trp274 and Ile311 of SETD3. The distance between the $N^3$ atom of His73 and the sulphur of the AdoHcy is 3.7 Å, suggesting that it mimics the pre-methyl transfer state (Figure 3A).

In the product complex, residues Leu67-Lys84 and the main chain of Ile85 of the peptide are visible, with His73 methylated to $N^3$-methylated His73 (His73me) (Supplementary Figure S7). Although the peptide adopts a $3^{10}$ helix at its C-terminal

end, the overall structures of the two complexes are highly similar, with a RMSD of 0.19 Å over protein Cα atoms and a RMSD of 0.32 Å over peptide Cα atoms.

His73me is inserted into the same pocket of SETD3 as shown in the product complex, However, Superposition of the two complexes clearly indicated that while the main chain carbonyl of Tyr313 and the guanidino of Arg316 are still hydrogen bonded to the main chain of His73me, the imidazole ring of His73me rotates by ~90° comparing to that of the unmodified His73, with its $C^{\delta}H$ pointing toward the ring of Tyr313 (Figure 3B). Accordingly, Asn256 of SETD3 flips its side chain by 180° and makes one hydrogen bond with the $N^1$ atom of His73me imidazole ring, which determines the orientation of the imidazole ring of His73 and allows it to make stacking and hydrophobic interactions with Tyr313, and Trp274 and Ile311, respectively (Figure 3B).

However, the flipping of the Asn256 side chain not only disrupts the hydrogen bond between Asn256 and the main chain Cα of His73me, but also makes the side chain amide groups of Asn256 close to the His73me Cα atom with a H-H distance of 1.7 Å (Supplementary Figures S8a-S8b), which would destabilize the enzyme-product complex and lead to subsequent release of the product.

In comparison with the other SET domain containing lysine methyltransferases, we found that there are two main reasons that SETD3 is not likely to be a lysine methyltransferase. First, the histidine binding pocket of SETD3 is too shallow to accommodate the aliphatic side chain of lysine. Replacement of His73 with a lysine, not only disrupts the water mediated hydrogen bond formed between $N^1$ of His73 and Arg316, but also leads to the steric clash between the $N^{\varepsilon}$ of the lysine and the main chain of the bottom carbonyl cage residue, Asp275. Second, phenylalanine of LSMT and SETD6 that makes cation-π interaction with the lysine, is replaced by asparagine (Asn256) in SETD3, and Asn256 is critical for substrate binding and histidine methylation (Figure 3 and Supplementary Figure S4B). Therefore, the substitution establishes the imidazole ring-specific interaction with histidine by impairing the lysine

specific interaction.

**SETD3 exhibits extensive interactions with β-actin other than H73**

Besides His73, the other residues of the β-actin peptide interacts extensively with the N-terminal lobe of SETD3, with Leu67-Glu72 and Gly74-Glu83 mainly contacting the SET and iSET motifs, respectively (Figure 4 and Supplementary Figure S9A). Leu67 and Pro70 of β-actin make hydrophobic contacts with Ile284 of SETD3 (Figure 4 and Supplementary Figure S9A). Additional hydrophobic interactions also exist between Tyr69 of β-actin and Pro259, Tyr288 and Leu290 of SETD3 (Figure 4 and Supplementary Figure S9A). Two main chain hydrogen bonds are formed between Tyr69 of β-actin and Tyr288 of SETD3, and between Ile71 of β-actin and Thr286 of SETD3, which induces the shift of the two sheets (β4-β10-β9 and β3-β11) in SETD3 mentioned above (Supplementary Figure S6). Specifically, Leu283-Thr285 of SETD3 adopts β strand in the structure of AdoMet bound SETD3 without peptide, but decomposes upon binding to the N-terminal side of β-actin, which shortens the β9 strand of SETD3 in the peptide-bound structure (Supplementary Figure S6). Besides forming main chain hydrogen bonds with SETD3, Ile71 also makes hydrophobic interactions with Ile258, Ile271, Trp274, Thr286, Tyr288 and Cys295 of SETD3 (Figure 4 and Supplementary Figure S9A). The carboxylate group of Glu72 makes one hydrogen bond to the guanidino group of Arg316, which is the only C-SET domain residue that contacts the N-terminal side of theβ-actin peptide (Figure 4 and Supplementary Figure S9A).

Regarding the residues C-terminal to His73 of β-actin, the main chain amide group of Gly74 is hydrogen bonded to the main chain carbonyl group of Gln255, while the side chain of Val76 stacks with the imidazole ring of His324 (Figure 4 and Supplementary Figure S9A). In contrast to Gly74 and Val76, the Ile75 side chain is solvent exposed, thereby not involved in the interactions with SETD3 (Figure 4). Thr77 not only makes hydrophobic contact with Leu268, but also forms two hydrogen bonds through its hydroxyl group with the side chain amide groups of Asn154 and Gln255, respectively

(Figure 4 and Supplementary Figure S9A). The side chain amide group of Asn154 makes another hydrogen bond with the main chain carbonyl of Asn78. The indole ring of Trp79 is inserted into a open hydrophobic pocket composed of Ile155, Val248, Val251 and Met252, with its main chain carbonyl H-bonded to the side chain amide of Gln216, while the carboxyl group of Asp80 is hydrogen bonded to the side chain amide of Arg215 (Figure 4 and Supplementary Figure S9A). Asp81 of β-actin makes three hydrogen bonds with SETD3, one with the side chain amide of Gln216 and two with the guanidino group of Arg215 (Figure 4 and Supplementary Figure S9A). The extensive interactions between the C-terminal residues of β-actin and SETD3 explains why deletion of the C-terminal residues of β-actin reduces its binding to SETD3 significantly (Kd = 2.9 μM, Figure 1B), suggesting that the C-terminal residues far from the target His73 residue also contribute to binding. Of note, the peptide sequence is also conserved in α- and γ- actins, suggesting that the other actins may be potential SETD3 substrates as well.

**Key residues of SETD3 involved in β-actin binding and methylation**

To evaluate the roles of the SETD3 residues around the active site in β-actin methylation, we made several single mutants and tested their binding affinities to the β-actin peptide. All SETD3 mutants displayed weaker binding affinities towards β-actin peptide (Supplementary table S1). Among these residues, the R215A and R316A mutants exhibited the most significantly reduced binding affinities by 21- and 42- folds, respectively. Our complex structures showed that both Arg215 and Arg316 make several hydrogen bonds with β-actin, implicating that they play a critical role in β-actin recognition and methylation (Supplementary table S1). N256A, which disrupts the hydrogen bond to the main chain of the His73 Cα atom, also reduces the β-actin binding affinity by 12-folds (Kd = 2.1μM) (Supplementary table S1). In the substrate complex, the imidazole ring of His73 stacks with the aromatic ring of Tyr313, with the $N^3$ atom of His73 in proximity to the hydroxyl group of Tyr313, which partially neutralizes the charge of $N^3$ and stabilizes the complex (Supplementary table S1). The favorable charge-charge interaction was also observed in the complex structures of protein with

$m^7G$ base (Xu et al., 2016), verified by the fact that mutating Tyr313 to phenylalanine diminishes binding affinity by 17-fold (Kd = 3.0 μM) (Supplementary table S1).

Next we performed mutagenesis and ITC studies to corroborate the roles of β-actin residues in binding to SETD3. Since the β-actin His73 is mainly recognized by SETD3 via several main chain hydrogen bonds, it is not unexpected that H73A only slightly reduces the SETD3 binding by ~2-folds (Kd = 0.36 μM) (Supplementary Table S1). Tyr69, Ile71 and Met82 of β-actin all contact SETD3 via hydrophobic interactions, with replacement by Ala dramatically reducing the binding affinities for SETD3 by 5-90 folds (Supplementary table S1). Consisting with the fact that both Asp80 and Asp81 of β-actin make hydrogen bonds with SETD3, D80A and D81A impair the binding to SETD3 by 20- and 11- folds, respectively (Supplementary table S1). In addition, truncation of Asp81-His88 of the peptide also reduces the SETD3 binding affinity by ~17 folds (Figure 1B and Supplementary table S1). Taken together, flanking sequences of His73, rather than His73 itself on the β-actin peptide, play critical roles in binding to SETD3, suggesting that the substrate recognition by SETD3 is highly sequence-selective.

To gain further insight into how those mutants that weakened the substrate binding ability compromise catalysis ability, we used the mass spectrometry technique to compare the histidine methylation activity of SETD3 with its mutants in the presence of AdoMet. While we did detect the peak of methylated β-actin (66-88) when it was added with wild type SETD3 (2-502) and AdoMet, no such peak was detected for peptide only in the presence of AdoMet (Supplementary Figure S1C). All four SETD3 mutants (R215A, N256A, Y313F and R316A) displayed weaker histidine methylation activity towards the β-actin peptide since the peaks of products catalyzed by the SETD3 mutants were much lower than that observed for wild type SETD3 (Supplementary Figure S1E), owing to the weaker substrate binding affinities of the mutants (Supplementary table S1). Collectively, SETD3 mutants impairing the substrate binding also compromise subsequent catalysis.

To understand whether full-length β-actin is catalyzed in the same way as the β-actin fragment, we purified full-length recombinant β-actin (FL β-actin) *in vitro* and studied the enzyme activity of SETD3 (2-502) variants on β-actin protein. As expected, wild type SETD3 (2-502) displayed high activity towards FL β-actin (~1660.8 pmol/min/mg), and no activity towards the H73A mutant. The four SETD3 mutants that show reduced peptide binding affinities and peptide catalytic activity, also display much weaker activity towards the full length β-actin protein (164.4-245.3 pmol/min/mg) (Supplementary Figure S10). In addition, Arg75 and Asn278 of SETD3 are key AdoMet binding residues, with their replacement by alanine abolishing the catalytic activity (14.7-35.8 pmol/min/mg). Collectively, our data suggest that perturbation of β-actin binding affinity significantly impacts on the SETD3-mediated methylation activity.

## DISCUSSION

### Catalytic mechanism of SETD3

Actins form microfilaments in *vivo*, which is one of the building blocks of cytoskeleton. Different types of post translational modifications (PTM) were found in actins (Terman and Kashina, 2013). $N^3$-methylhistidine was identified in β-actin decades ago (Johnson et al., 1967) and was reported to mediate the polymerization and hydrolysis of actin (Kabsch et al., 1990). However, how the mark is incorporated, remains enigmatic. Our identification of SETD3 as a histidine methyltransferase catalyzing the His73 methylation of β-actin, very well answers both of the unsolved questions.

Although SETD3 contains canonical SET domain and its AdoHcy binding mode is similar to those observed in other SET domain lysine methyltransferases. The predicted SET domain fold of SETD3 could not help to uncover the molecular mechanisms of the substrate recognition or histidine methylation, since SET domain itself, as a short motif of ~110 aa, always need iSET and C-SET domains to complete its function (Chang et al., 2011; Trievel et al., 2002). Therefore, two complex structures presented

here not only for the first time uncover the molecular mechanism by which a SET domain protein acts as the $N^3$-histidine methyltransferase, as revealed by the histidine specific binding pocket in both complexes, but also provide insights into the pre- and post- methyl transfer states of the histidine methylation by the same enzyme.

Our structural analysis, assisted by mutagenesis and biochemical experiments, suggested that the recognition of β-actin and subsequent catalysis by SETD3 is different from those observed from other SET domain lysine methyltransferases in following aspects: i) histidine recognition pocket of SETD3 is shallower and key residues of SET3, which form hydrogen bonds with the His73 imidazole ring, are not conserved in other SET domain members (Figure 5A-5B and Supplementary Figure S4B). ii) In two solved SETD3 complex structures, the methylated histidine rotates its side chain by 90° to enable the release of the product (Figure 3), which is not reported in lysine methyltransferase complexes before. (iii) The recognition of β-actin by SETD3 is highly dependent on flanking sequences of His73 (Figure 4), consisting with the high binding affinity between H73A and SETD3 (Figure 1B). Although $N^3$-methylhistidine is also found in other proteins, such as mammalian myosin (Johnson et al., 1967) and *Saccharomyces cerevisiae* Rpl3 (Webb et al., 2010), the sequence preference suggests that SETD3 probably only work on His73 of β-actin, as well as on corresponding histidines in α- and γ-actins. One implication is that no SETD3 ortholog is found in baker's yeast.

Recently, the structures of CARNMT1 bound to analogs of $N^1$-methylhistidine was also reported (Cao et al., 2018), which prompts us to compare the $N^3$-histidine methylation by SETD3 with the $N^1$-histidine methylation by CARNMT1 through comparing two complexes (Figures 5A and 5C). We found the two histidine methyltransferases adopt distinct folds, which lead to different AdoHcy binding modes. Furthermore, SETD3 and CARNMT1 display different histidine methylation mechanism. While CARNMT1 utilizes Asp316 to form two hydrogen bonds with the imidazole ring of the histidine to facilitate the deprotonation of the histidine imidazole ring (Figure 5C), while in the

SETD3-substrate complex, histidine deprotonation is enabled by the Tyr313 acting as the general base and by Asp275 forming main chain hydrogen bond with the histidine imidazole ring (Figures 5A and 5D). The important role of Tyr313 in both binding and catalysis is further corroborated by the mass spectrum analysis that mutating Tyr313 to Phe markedly decreased the β-actin binding affinity, as well as the methylation activity (Supplementary Table S1 and Supplementary Figure S1E). Last but not least, unexpected 90-degree rotation of imidazole ring of His73me does not occur in CARNMT1-mediated histidine methylation.

**Implications in conformational changes of actin upon binding to SETD3**

In natively purified rabbit β-actin, His73 exists as the methylated form and it is localized close to the phosphate groups of ATP binding to actin (Kabsch et al., 1990). Although our complexes only contain a fragment of β-actin, they could still provide clues on the role of SETD3 in mediating actin functions. The residues 66-84 of β-actin consist of three β strands followed by an α-helice (Supplementary Figure S11), however, the β-actin peptide adopts an extended conformation upon binding to SETD3 (Supplementary Figure S11), suggesting that β-actin would probably endure local structural remodeling during methylation, such as beta-sheet decomposition. To complete the methylation, remodeling of β-actin should occur at the very beginning to expose the partially buried His73 and make it accessible to SETD3, which is an energy-consumption step.

Our ITC binding data indicate that SETD3 binds to β-actin peptide with extreme high binding affinity (Kd = 170 nM) and the binding would provide energy to decompose the local secondary structures of β-actin. It is not unlikely that His73 is directly recognized by SETD3 as the first step considering the β-actin binding by SETD3 largely depends on the flanking sequences of His73. Of note, Leu67-Tyr69 of β-actin is exposed to the solvent, while Tyr69 of the peptide makes extensive hydrophobic contacts with SETD3, hence the decomposition and binding of β-actin to SETD3 might firstly occur at the N-terminal side of His73, which may help to “pull out” His73 to make it exposed to SETD3. However, we could not exclude the possibility that some

factors, such as the CCT complex, may bind actin to facilitate the process (Martin-Benito et al., 2002). The structures of SETD3 with native actin may be required to provide more details on the dynamics of actin.

Since SETD3 is reported to be associated with cancer malignancy because of its higher level of expression in liver tumor (Cheng et al., 2017), uncovering the role of SETD3 in methylating actin may pave the way for understanding its role in carcinogenesis through influencing the functions of actin and cytoskeleton. Thus our results not only will facilitate the understanding the roles of actin methylation catalyzed by SETD3 in human cancer, but also would provide structural basis for designing SETD3 inhibitors in near future.

## EXPERIMENTAL PROCEDURES

### Cloning, protein expression and purification of SETD3

Gene encoding the core region of SETD3 (2-502) were synthesized by Sangon Biotech (Shanghai) and cloned into a modified pET28-MHL (Genbank accession number: EF456735)(27). Then the plasmid was transformed into E. coli BL21 (DE3) and recombinant protein was over-expressed at 16 °C for 20 hours in the presence of 0.5 mM IPTG. Recombinant SETD3 proteins were purified by fast flow Ni-NTA column (GE). N-terminal 6xHis-tag of recombinant proteins were removed by Tobacco Etch Virus (TEV) protease. Gel filtration and ion-exchange were employed for further purification. Purified SETD3 proteins, with their purities exceeding 90% as judged by SDS-PAGE, were concentrated to 15 mg/ml and stored at −80°C before further use. The SETD3 mutants were constructed by conventional PCR using the MutanBEST kit (TaKaRa) and further verified by DNA sequencing. The SETD3 mutants were expressed and purified in the same way as the wild type protein.

### Overexpression and Purification of the Recombinant β-actin Inclusion-body Protein

Plasmid pCOLD I encoding human β-actin (ACTB, GenBank: NM_001101.4) was a

kind gift of Dr. Minoru Tamura (Ehime University, Japan) and was prepared as described (Tamura, 2018). For β-actin production, Escherichia coli BL21(DE3) (Agilent, USA) cells were transformed with the DNA construct and a single colony was selected to start an over-nigh pre-culture. 500 mL of LB broth (with 100 μg/mL ampicilin) was inoculated with 50 ml of the pre-culture and incubated at 37 °C and 200 rpm until an OD600 of 0.5 was reached. The culture was placed on ice for 20 min. (cold-shock) and IPTG was added to a final concentration of 0.2 mM to induce protein expression. Cells were incubated for 20 h at 15 °C, 200 rpm, and harvested by centrifugation (6000 × g for 10 min). The cell paste was resuspended in 27.5 ml lysis buffer consisted of 20 mM Hepes, pH 7.5, 1 mM DTT, 1 mM ADP, 0.5 mM PMSF, 2 μg/ml leupeptin and 2 μg/ml antipain, 0.2 mg/ml hen egg white lysozyme (BioShop), 1000 U Viscolase (A&A Biotechnology, Poland). The cells were lysed by freezing in liquid nitrogen and after thawing and vortexing, the extracts were centrifuged at 4° C (20,000 × g for 30 min).

The pellet, containing inclusion bodies, was completely resuspended in buffer A (20 mM Hepes, pH 7.5, 2M urea, 0.5 M NaCl, 5 mM DTT, 2 mM EDTA) with the use of Potter-Elvehjem homogenizer and centrifuged at 4° C (20,000 × g for 10 min). The resulting pellet was then subjected to a further two rounds of sequential wash in buffer B (20 mM Hepes, pH 7.5, 0.5 M NaCl, 5 mM DTT, 2 mM EDTA) and buffer C (20 mM Hepes, pH 7.5, 0.5 M NaCl). The washed inclusion bodies were finally solubilized in loading buffer (20 mM Tris-HCl, pH 7.5, 6 M guanidine HCl, 0.5 M NaCl, 10 mM imidazole) and applied on a HisTrap FF column (5 ml) equilibrated with the same buffer. The column was washed with 20 ml of loading buffer and the bound protein was refolded by the use of a linear gradient of 6-0 M guanidine HCl in loading buffer (40 ml for 20 min). Next, the column was washed with 15 ml of loading buffer without guanidine HCl and the retained proteins were eluted with a stepwise gradient of imidazole (25 ml of 40 mM, 25 ml of 60 mM and 20 ml of 500 mM). The recombinant proteins were eluted with 500 mM imidazole in homogeneous form as confirmed by SDS-PAGE (Supplementary Fig. 1). The β-actin preparation was immediately desalted

on PD-10 columns equilibrated with 20 mM Tris-HCl, pH 7.5, 1 mM DTT, 6% sucrose, 2 μg/ml leupeptin and 2 μg/ml antipain. The purified β-actin was stored at -70°C.

**Isothermal titration calorimetry (ITC)**

All peptides were synthesized by GL Biochem (Shanghai) Ltd. and were dissolved in water as a stock of 5-12 mM, with the pH of stock solutions adjusted to pH 7.5. Peptides and concentrated proteins were diluted with ITC buffer (20 mM Tris, pH 7.5 and 150 mM NaCl). ITC experiments were performed by titrating 2 μl of peptide (1.2-1.5 mM) into cell containing 50 μM proteins on MicroCal™ (Malvern Panalytical, UK) at 25°C, with a spacing time of 160 s and a reference power of 10 μCal/s. Control experiments were performed by injection of peptides into buffer. Binding isotherms were plotted, analyzed and fitted in a one-site binding model by MicroCal PEAQ-ITC Analysis Software (Malvern Panalytical, UK) after subtraction of respective controls. The dissociation constants (Kds) were from a minimum of two experiments (mean ± sd).

**Crystallization, data collection and structure determination**

All crystals were grown by using sitting drop vapor diffusion method at 18°C. For crystallization of SETD3 with methylated peptide, SETD3 (12 mg/ml) was pre-incubated with synthesized β-actin peptide (66-88) (GL Biochem Ltd.) and AdoMet at a molar ratio of 1:3:4, and mixed with the crystallization buffer containing 0.1 M Sodium cacodylate trihydrate pH 6.5, 0.2M Magnesium acetate tetrahydrate, and 20% v/v Polyethylene glycol 8000. For the crystallization of SETD3 with unmodified peptide, SETD3, in the concentration of 12 mg/ml, was pre-incubated with synthesized β-actin peptide (66-88) and AdoHcy at a molar ratio of 1:3:4, and mixed with the crystallization buffer containing 0.1M HEPES sodium pH 7.5, 2% v/v Polyethylene glycol 400, and 2.0 M Ammonium sulfate. Before flash-freezing crystals in liquid nitrogen, all crystals were soaked in a cryo-protectant consisting of 90% reservoir solution plus 10% glycerol. The diffraction data were collected on BL17U1 at the Shanghai Synchrotron Facility (28) at Shanghai Synchrotron Facility (SSRF). Data sets

were collected at 0.9789 Å or 0.9789 Å, and were processed by using the HKL2000 program (29).

The initial structures of the SETD3-actin complexes were solved by molecular replacement in PHENIX (30) with previously solved SETD3 structure (PDB: 3SMT) as the search model. Then all the models were refined manually and built with Coot (31). The final structures were further refined by PHENIX (30). The statistics for data collection and refinement are summarized in Table 1.

**Mass Spectrometry**

Reversed-phase microcapillary/tandem mass spectrometry (LC/MS/MS) was performed using an Easy-nLC nanoflow HPLC (Proxeon Biosciences) with a self-packed 75 μm id x 15 cm C18 column connected to either an QE-Plus (Thermo Scientific) in the data-dependent acquisition and positive ion mode at 300 nL/min. Passing MS/MS spectra were manually inspected to be sure that all b- and y- fragment ions aligned with the assigned sequence and modification sites. A 25-ul reaction mixture containing 2 μM SETD3 or SETD3 mutants (final concentration), 20 μM peptide (final concentration) in a buffer containing 10mM Tris-HCl, pH 7.5, 20mM NaCl and 10μM AdoMet. The reaction was incubated at 37℃ for 2h before being quenched (at 70℃ for 10-15 mins). Then quenched reactions were analyzed by LC/MS/MS and Proteomics Browser Software, with the relative abundances of substrate and product reflecting the methylation activities of proteins.

**Assay of the Actin-specific Histidine N-methyltransferases Activity**

The enzyme activity was determined by measuring the incorporation of [$^3$H]methyl group from S-[methyl-3H]adenosyl-L-methionine ([$^3$H]SAM) into homogenous recombinant human (mammalian) β-actin or its mutated form in which histidine 73 was replaced by alanine residue (H73A). The standard incubation mixture (0.11 ml) contained 25 mM Tris-HCl, pH 7.2, 10 mM KCl, 1 mM DTT, 2 μM protein substrate and 1 μM [$^1$H+$^3$H] SAM (≈ 400 × 103 cpm). When appropriate, the incubation mixture

was supplemented with recombinant S-adenosyl-L-homocysteine (SAH) nucleosidase and adenine deaminase as indicated in legends to figures and tables. The reaction was started by the addition of enzyme preparation and carried out at 37°C for 15 min unless otherwise described. Protein methylation was linear for at least 15 min under all conditions studied. By analogy to assays of nonribosomal peptide synthetase activity [Richardt et al., 2003; Drozak et al., 2014], the incubation was stopped by the addition of 0.1 ml of the reaction mixture to 0.025 ml of BSA (1 mg/ml) and 0.8 ml of ice-cold 10 % (w/v) trichloroacetic acid (TCA). After 10 min on ice, the precipitate was pelleted and washed twice with ice-cold 10 % TCA. The pellet was finally dissolved in pure formic acid.

## ACCESSION NUMBERS

The coordinates and structure factors of AdoHcy bound SETD3 with unmodified β-actin peptide and methylated β-actin peptide haven been deposited into Protein Data Bank (PDB) with accession numbers 6ICV and 6ICT, respectively.

## SUPPLEMENTARY INFORMATION

Supplemental Information can be found online at XXXX

## ACKNOWLEDGEMENT

We are grateful to the staff members at beam lines BL17U1, BL18U1 and BL19U1 at Shanghai Synchrotron Radiation Facility for assistance in data collection. This work was supported by National Natural Science Foundation of China Grants 31570737 and 31770806 (to C. X.) and 31500601 (to S. L.) and the "Thousand Young Talent program" (to C. X.). The Structural Genomics Consortium is a registered charity (no. 1097737) that receives funds from AbbVie; Bayer Pharma AG; BoehringerIngelheim; Canada Foundation for Innovation; Eshelman Institute for Innovation; Genome Canada through the Ontario Genomics Institute; the Innovative Medicines Initiative (European Union/European Federation of Pharmaceutical Industries and Associations; Unrestricted Leveraging of Targets for Research Advancement and Drug Discovery

[ULTRA-DD] grant no. 115766); Janssen, Merck, and Company; Novartis Pharma AG; Ontario Ministry of Economic Development and Innovation; Pfizer; São Paulo Research Foundation (FAPESP); Takeda; and the Wellcome Trust (to J.M.).

## CONFLICT OF INTEREST

The authors declare that they have no conflict of interest related to the publication of this manuscript.

**AUTHOR CONTRIBUTIONS:** G.Q., S.L., W.T., S.K., G.W, and J.D. performed experiments and data analysis. C.X. conceived the project, J.M. and C.X supervised the project and wrote the manuscript. All authors contributed to editing the manuscript.

## REFERENCES

Asatoor, A.M., and Armstrong, M.D. (1967). 3-methylhistidine, a component of actin. Biochem Biophys Res Commun *26*, 168-174.

Bonasio, R., Tu, S., and Reinberg, D. (2010). Molecular signals of epigenetic states. Science *330*, 612-616.

Cao, R., Zhang, X., Liu, X., Li, Y., and Li, H. (2018). Molecular basis for histidine N1 position-specific methylation by CARNMT1. Cell Res *28*, 494-496.

Chang, Y., Levy, D., Horton, J.R., Peng, J., Zhang, X., Gozani, O., and Cheng, X. (2011). Structural basis of SETD6-mediated regulation of the NF-kB network via methyl-lysine signaling. Nucleic Acids Res *39*, 6380-6389.

Chen, Z., Yan, C.T., Dou, Y., Viboolsittiseri, S.S., and Wang, J.H. (2013). The role of a newly identified SET domain-containing protein, SETD3, in oncogenesis. Haematologica *98*, 739-743.

Cheng, X., Hao, Y., Shu, W., Zhao, M., Zhao, C., Wu, Y., Peng, X., Yao, P., Xiao, D., Qing, G.*, et al.* (2017). Cell cycle-dependent degradation of the methyltransferase SETD3 attenuates cell proliferation and liver tumorigenesis. J Biol Chem *292*, 9022-9033.

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

dos Remedios, C.G., Chhabra, D., Kekic, M., Dedova, I.V., Tsubakihara, M., Berry, D.A., and Nosworthy, N.J. (2003). Actin binding proteins: regulation of cytoskeletal microfilaments. Physiol Rev *83*, 433-473.

Eom, G.H., Kim, K.B., Kim, J.H., Kim, J.Y., Kim, J.R., Kee, H.J., Kim, D.W., Choe, N., Park, H.J., Son, H.J.*, et al.* (2011). Histone methyltransferase SETD3 regulates muscle differentiation. J Biol Chem *286*, 34733-34742.

Gunning, P., Ponte, P., Okayama, H., Engel, J., Blau, H., and Kedes, L. (1983). Isolation and characterization of full-length cDNA clones for human alpha-, beta-, and gamma-actin mRNAs: skeletal but not cytoplasmic actins have an amino-terminal cysteine that is subsequently removed. Mol Cell Biol *3*, 787-795.

Hamamoto, R., Saloura, V., and Nakamura, Y. (2015). Critical roles of non-histone protein lysine

methylation in human tumorigenesis. Nat Rev Cancer *15*, 110-124.
Herman, I.M. (1993). Actin isoforms. Curr Opin Cell Biol *5*, 48-55.
Holm, L., and Rosenstrom, P. (2010). Dali server: conservation mapping in 3D. Nucleic Acids Res *38*, W545-549.
Johnson, P., Harris, C.I., and Perry, S.V. (1967). 3-methylhistidine in actin and other muscle proteins. Biochem J *105*, 361-370.
Kabsch, W., Mannherz, H.G., Suck, D., Pai, E.F., and Holmes, K.C. (1990). Atomic structure of the actin:DNase I complex. Nature *347*, 37-44.
Leterrier, C., Dubey, P., and Roy, S. (2017). The nano-architecture of the axonal cytoskeleton. Nat Rev Neurosci *18*, 713-726.
Martin-Benito, J., Boskovic, J., Gomez-Puertas, P., Carrascosa, J.L., Simons, C.T., Lewis, S.A., Bartolini, F., Cowan, N.J., and Valpuesta, J.M. (2002). Structure of eukaryotic prefoldin and of its complexes with unfolded actin and the cytosolic chaperonin CCT. EMBO J *21*, 6377-6386.
Martin, C., and Zhang, Y. (2005). The diverse functions of histone lysine methylation. Nat Rev Mol Cell Biol *6*, 838-849.
Min, J., Feng, Q., Li, Z., Zhang, Y., and Xu, R.M. (2003). Structure of the catalytic domain of human DOT1L, a non-SET domain nucleosomal histone methyltransferase. Cell *112*, 711-723.
Nudel, U., Zakut, R., Shani, M., Neuman, S., Levy, Z., and Yaffe, D. (1983). The nucleotide sequence of the rat cytoplasmic beta-actin gene. Nucleic Acids Res *11*, 1759-1771.
Nyman, T., Schuler, H., Korenbaum, E., Schutt, C.E., Karlsson, R., and Lindberg, U. (2002). The role of MeH73 in actin polymerization and ATP hydrolysis. J Mol Biol *317*, 577-589.
Raghavan, M., Lindberg, U., and Schutt, C. (1992). The use of alternative substrates in the characterization of actin-methylating and carnosine-methylating enzymes. Eur J Biochem *210*, 311-318.
Rathert, P., Dhayalan, A., Murakami, M., Zhang, X., Tamas, R., Jurkowska, R., Komatsu, Y., Shinkai, Y., Cheng, X., and Jeltsch, A. (2008). Protein lysine methyltransferase G9a acts on non-histone targets. Nat Chem Biol *4*, 344-346.
Sims, R.J., 3rd, Nishioka, K., and Reinberg, D. (2003). Histone lysine methylation: a signature for chromatin function. Trends Genet *19*, 629-639.
Skalli, O., Ropraz, P., Trzeciak, A., Benzonana, G., Gillessen, D., and Gabbiani, G. (1986). A monoclonal antibody against alpha-smooth muscle actin: a new probe for smooth muscle differentiation. J Cell Biol *103*, 2787-2796.
Sussman, D.J., Sellers, J.R., Flicker, P., Lai, E.Y., Cannon, L.E., Szent-Gyorgyi, A.G., and Fulton, C. (1984). Actin of Naegleria gruberi. Absence of N tau-methylhistidine. J Biol Chem *259*, 7349-7354.
Tamura, M. (2018). Production of Human beta-Actin Using a Bacterial Expression System with a Cold Shock Vector. Curr Protoc Protein Sci *93*, e61.
Terman, J.R., and Kashina, A. (2013). Post-translational modification and regulation of actin. Curr Opin Cell Biol *25*, 30-38.
Theriot, J.A., and Mitchison, T.J. (1991). Actin microfilament dynamics in locomoting cells. Nature *352*, 126-131.
Trievel, R.C., Beach, B.M., Dirk, L.M., Houtz, R.L., and Hurley, J.H. (2002). Structure and catalytic mechanism of a SET domain protein methyltransferase. Cell *111*, 91-103.
Vijayasarathy, C., and Rao, B.S. (1987). Partial purification and characterisation of S-adenosylmethionine:protein-histidine N-methyltransferase from rabbit skeletal muscle. Biochim Biophys Acta *923*, 156-165.

Wagner, E.J., and Carpenter, P.B. (2012). Understanding the language of Lys36 methylation at histone H3. Nat Rev Mol Cell Biol *13*, 115-126.
Webb, K.J., Zurita-Lopez, C.I., Al-Hadid, Q., Laganowsky, A., Young, B.D., Lipson, R.S., Souda, P., Faull, K.F., Whitelegge, J.P., and Clarke, S.G. (2010). A novel 3-methylhistidine modification of yeast ribosomal protein Rpl3 is dependent upon the YIL110W methyltransferase. J Biol Chem *285*, 37598-37606.
Whetstine, J.R., Nottke, A., Lan, F., Huarte, M., Smolikov, S., Chen, Z., Spooner, E., Li, E., Zhang, G., Colaiacovo, M., *et al.* (2006). Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases. Cell *125*, 467-481.
Xu, C., Ishikawa, H., Izumikawa, K., Li, L., He, H., Nobe, Y., Yamauchi, Y., Shahjee, H.M., Wu, X.H., Yu, Y.T., *et al.* (2016). Structural insights into Gemin5-guided selection of pre-snRNAs for snRNP assembly. Genes Dev *30*, 2376-2390.

## TABLES

Table 1. Data collection and refinement statistics

## FIGURES LEGENDS

**Figure 1. SETD3 core region specifically recognizes a fragment of β-actin containing His73.** (A) Domain architecture of full-length human SETD3 (aa 1-594) and the sequence of the β-actin peptide (66-88), with His73 of β-actin highlighted. (B) Representative ITC binding curves for the binding of the SETD3 (aa 2-502) to β-actin peptides of different lengths. Molecular ratios, derived Kds and respective standard deviations are also indicated. (C) Overall structure of AdoHcy bound SETD3 with unmodified β-actin peptide. The SETD3 are colored in the same mode as shown in Figure 1a, with the N-SET, SET, iSET and C-SET regions of SETD3 colored in grey, blue, purple and salmon, respectively. The peptide are shown in yellow cartoon, while His73 and AdoHcy are shown in yellow and cyan sticks, respectively. His73 of actin and the AdoHcy are labeled

**Figure 2. Comparison of the AdoHcy binding mode of SETD3 with those observed in LSMT and SETD6.** (A) A detailed depiction of intermolecular interactions between key residues of SETD3 and AdoHcy. AdoHcy is shown in cyan sticks, while the AdoHcy binding residues of SETD3 are colored based on the regions they reside, as shown in Figure 1A, and are shown in sticks. (B) Detailed interactions between AdoHcy and LSMT (PDB id: 1MLV), with AdoHcy binding residues of LSMT and AdoHcy

shown in green and cyan sticks, respectively. (C) Detailed interactions between AdoMet and SETD6 (PDB id: 3QXY), with AdoMet binding residues of SETD6 and AdoMet shown in orange and cyan sticks, respectively. In Figures 2B and 2C, hydrogen bonds are indicated by black dashes.

**Figure 3. The histidine binding pocket of SETD3.** (A) The β-actin peptide is bound in a long groove on the surface of SETD3 N-lobe region, with His73 of β-actin positioned into a hydrophobic pocket. His73 and AdoHcy are shown in yellow and cyan sticks, respectively. The SETD3 residues involved in binding to His73 are colored in the same way as shown in Figure 2A. (B) Detailed interactions between His73me and SETD3 in the post-methyl transfer complex. His73me and AdoHcy are shown in yellow and cyan sticks, respectively. The SETD3 residues involved in binding to His73me are colored in the same way as shown in Figure 2A. (C) Superposition of the two complexes on the histidine/methylhistidine binding pocket. For the SETD3-His73me complex, His73me binding residues of SETD3 are shown in the same way as in Figure 2A, with His73me and AdoHcy shown in yellow and cyan sticks, respectively. For the SETD3-His73 complex, His73 binding residues and AdoHcy are shown in grey sticks, while His73 are shown in green sticks to show that it rotates 90 degree after catalysis. After catalysis, the hydrogen bond between Asn256 and Cα of His73 (grey dash) is disrupted, and one new hydrogen bond is formed between Asn256 and $N^1$ atom of His73me (red dash).

**Figure 4. Detailed interactions between SETD3 and flanking sequences of His73.** The β-actin peptide mainly contacts N-terminal lobe of SETD3, with its N-terminal and C-terminal sides contact SET and iSET domains, respectively. In the top panel, SETD3 are shown in ribbon and colored in the same way as shown in Figure 2A. β-actin peptide is shown in yellow sticks. Detailed interactions between SETD3 and Leu67-Glu72, and between SETD3 and Gly74 and Glu83, are shown in the left bottom and right bottom panels, respectively. The SETD3 residues involved in β-actin binding are shown in sticks.

**Figure 5. Comparison of the substrate recognition pocket of SETD3 with those of LSMT and CARNMT1.** (A) histidine recognition by SETD3 in the presence of AdoHcy; (B) lysine recognition by LSMT(PDB id: 1OZV) in the presence of AdoHcy; (C) carnosine recognition by CARNMT1 (PDB id: 5YF1) in the presence of SFG. Histidine, lysine and carnosine are shown in yellow sticks, while the protein residues involved in binding are shown in red sticks. AdoHcy and its mimic, SFG, are shown in cyan sticks. (D) Catalytic mechanism of actin histidine methylation by SETD3, with Tyr313 of SETD3 acting as a general base to facilitate the deprotonation of the $N^3$ atom of His73. Specifically, the imidazole ring of His73me rotated 90 degree after methylation.

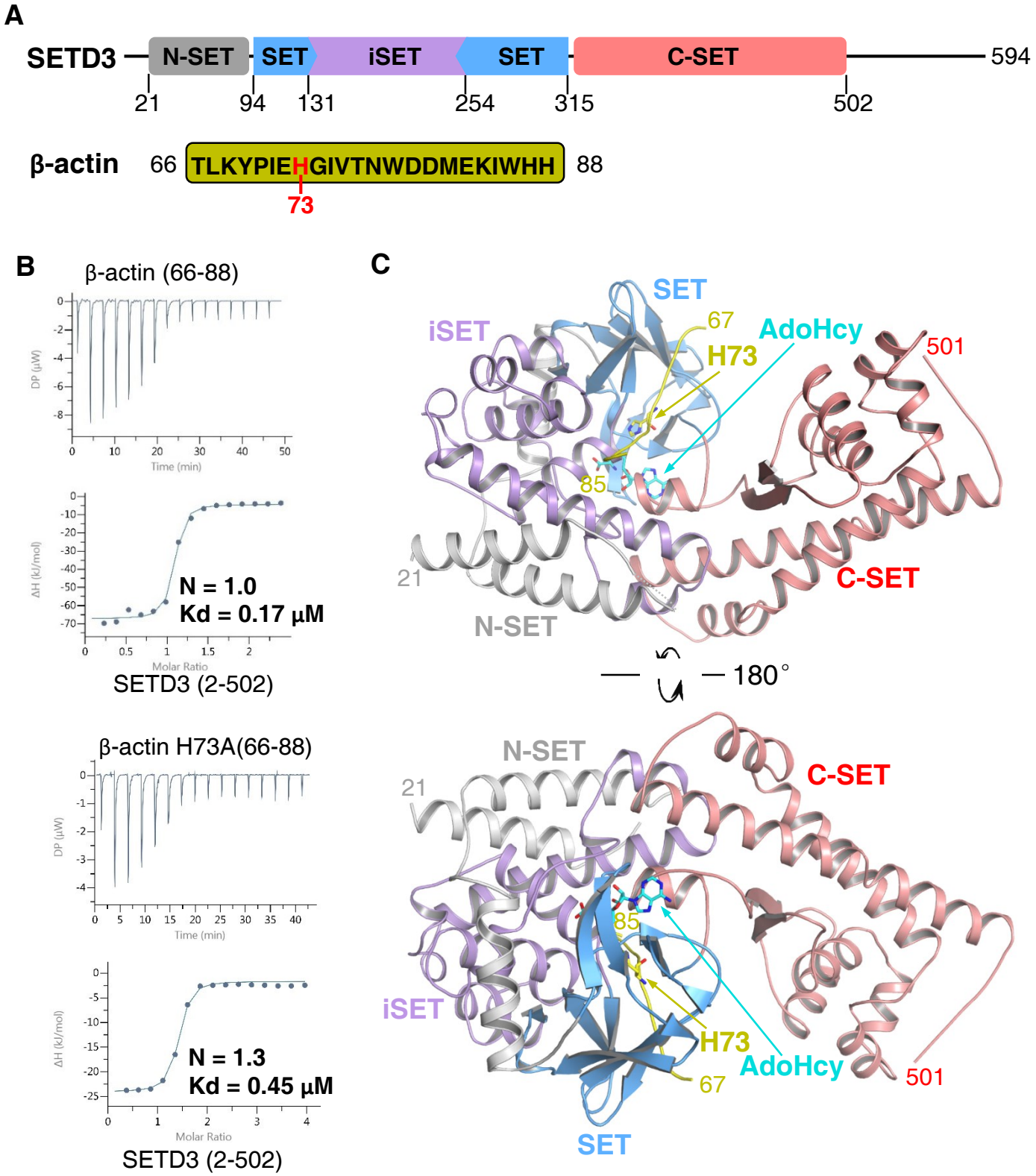

A
SETD3
N-SET
SET
iSET
SET
C-SET
594
21
94
131
254
315
502
β-actin
66
TLKYPIEHGIVTNWDDMEKIWHH
88
73
B
β-actin (66-88)
DP (μW)
Time (min)
ΔH (kJ/mol)
Molar Ratio
N = 1.0
Kd = 0.17 μM
SETD3 (2-502)
β-actin H73A(66-88)
DP (μW)
Time (min)
ΔH (kJ/mol)
Molar Ratio
N = 1.3
Kd = 0.45 μM
SETD3 (2-502)
C
SET
iSET
67
AdoHcy
H73
501
85
21
C-SET
N-SET
180°
N-SET
21
C-SET
85
iSET
H73
AdoHcy
501
67
SET

**Figure 1**

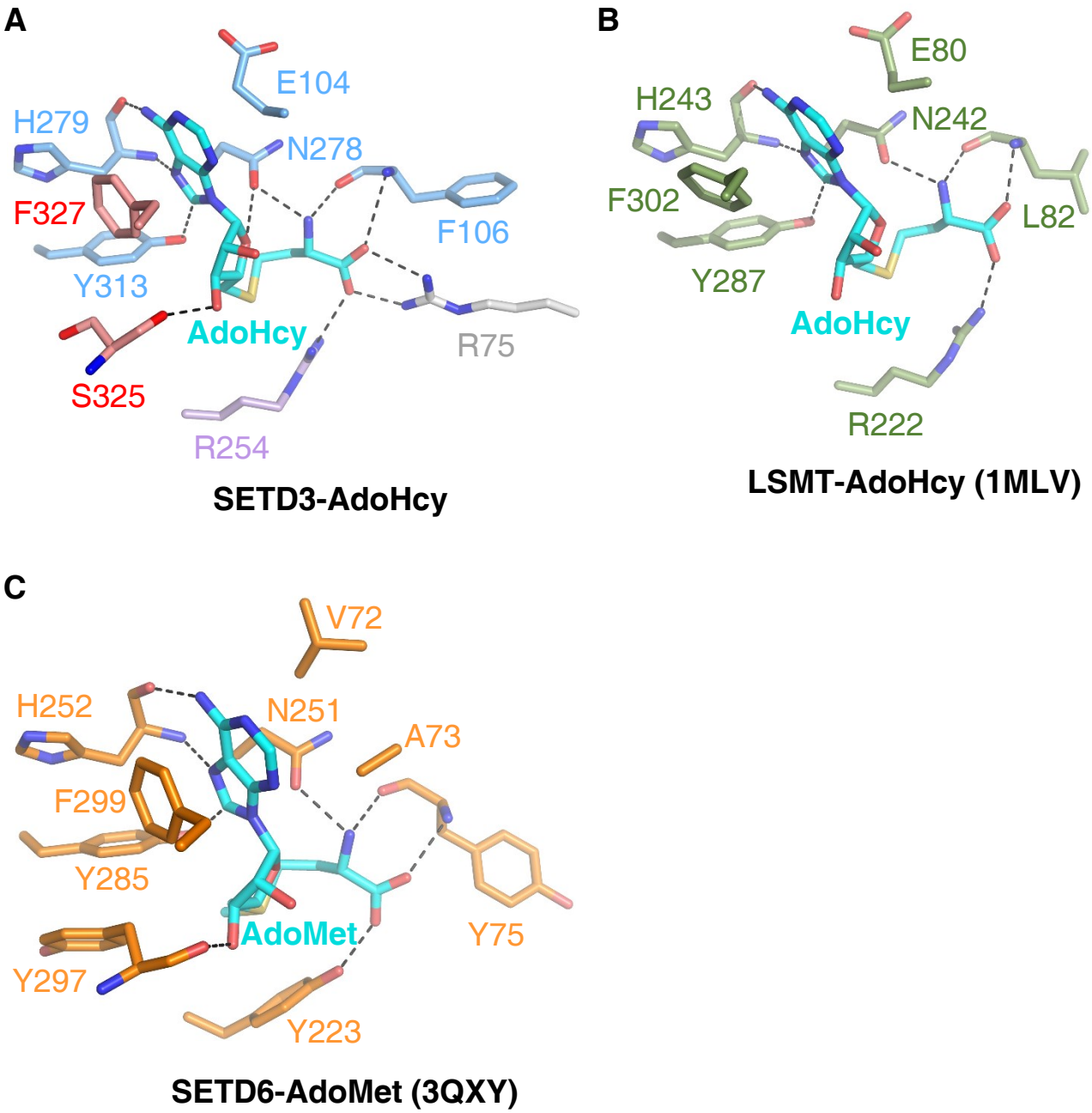

A
E104
H279
N278
F327
F106
Y313
AdoHcy
R75
S325
R254
SETD3-AdoHcy
B
E80
H243
N242
F302
L82
Y287
AdoHcy
R222
LSMT-AdoHcy (1MLV)
C
V72
H252
N251
A73
F299
Y285
AdoMet
Y75
Y297
Y223
SETD6-AdoMet (3QXY)

**Figure 2**

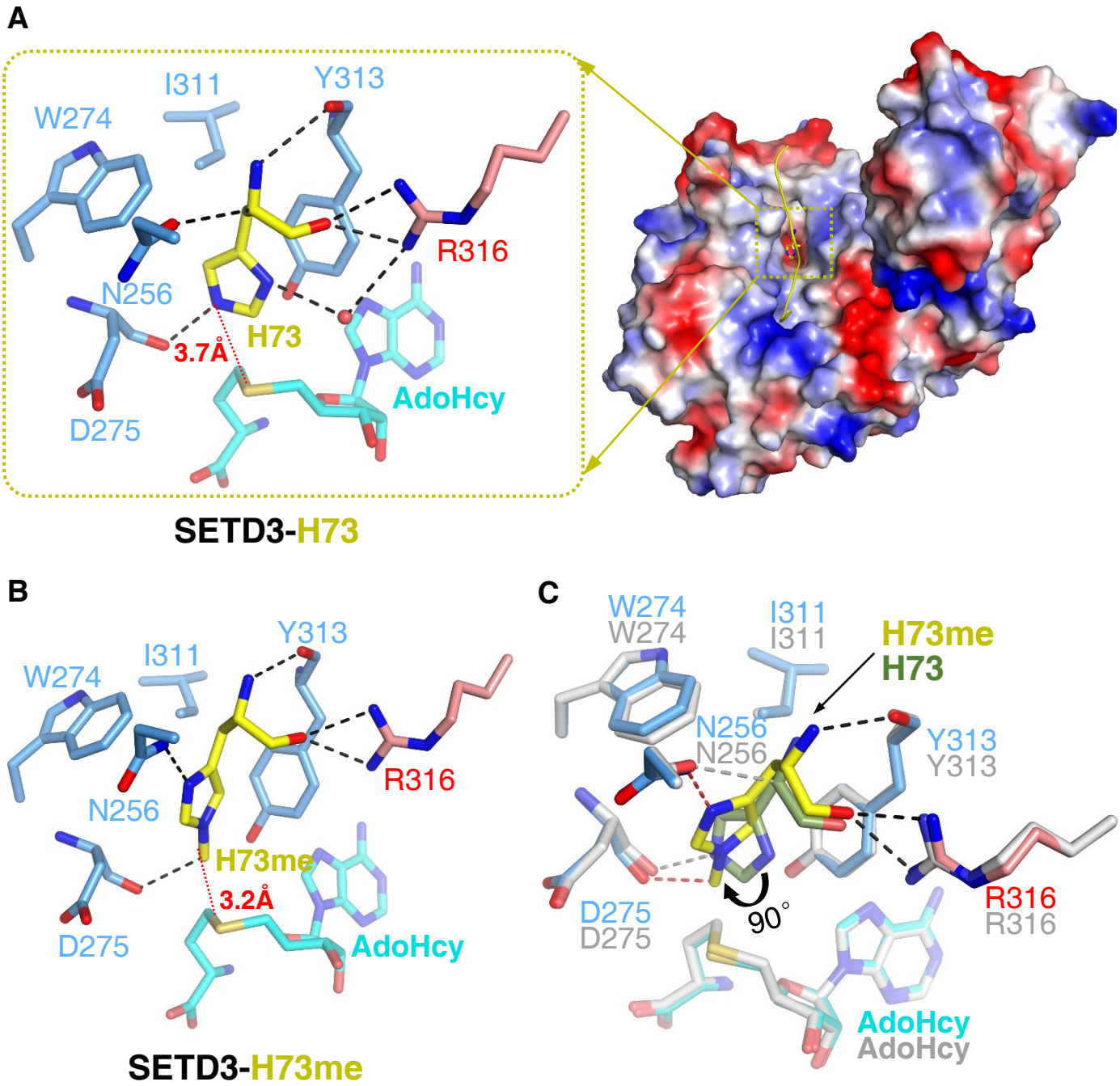

A
I311
Y313
W274
R316
N256
H73
3.7Å
AdoHcy
D275
SETD3-H73
B
Y313
I311
W274
R316
N256
H73me
3.2Å
D275
AdoHcy
SETD3-H73me
C
W274
W274
I311
I311
H73me
H73
N256
N256
Y313
Y313
D275
D275
90°
R316
R316
AdoHcy
AdoHcy

**Figure 3**

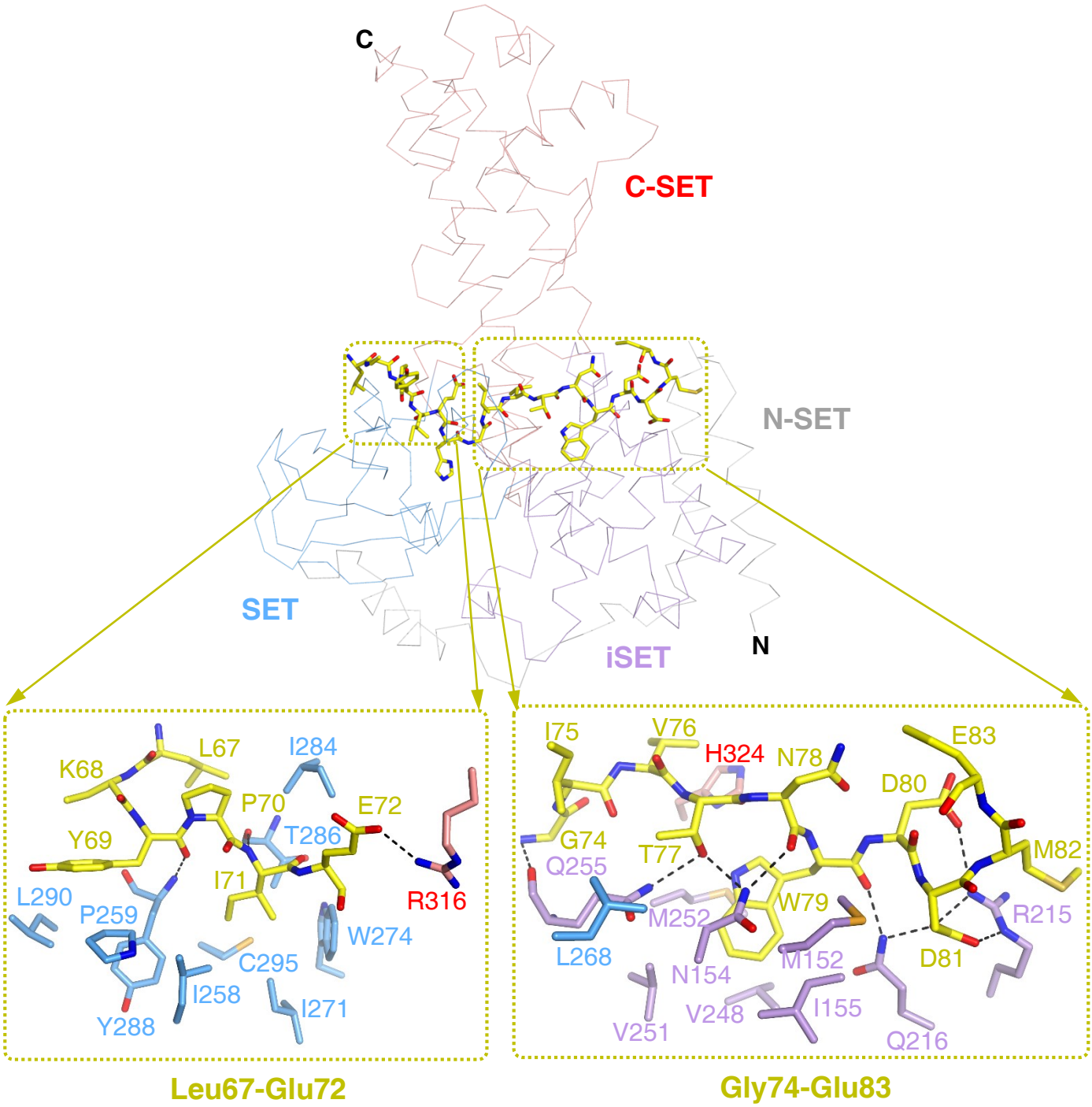

C
C-SET
N-SET
SET
iSET
N
K68
L67
I284
P70
E72
Y69
T286
I71
R316
L290
P259
W274
C295
I258
I271
Y288
Leu67-Glu72
I75
V76
H324
N78
E83
D80
G74
Q255
T77
M82
W79
M252
R215
L268
M152
D81
N154
V248
I155
V251
Q216
Gly74-Glu83

**Figure 4**

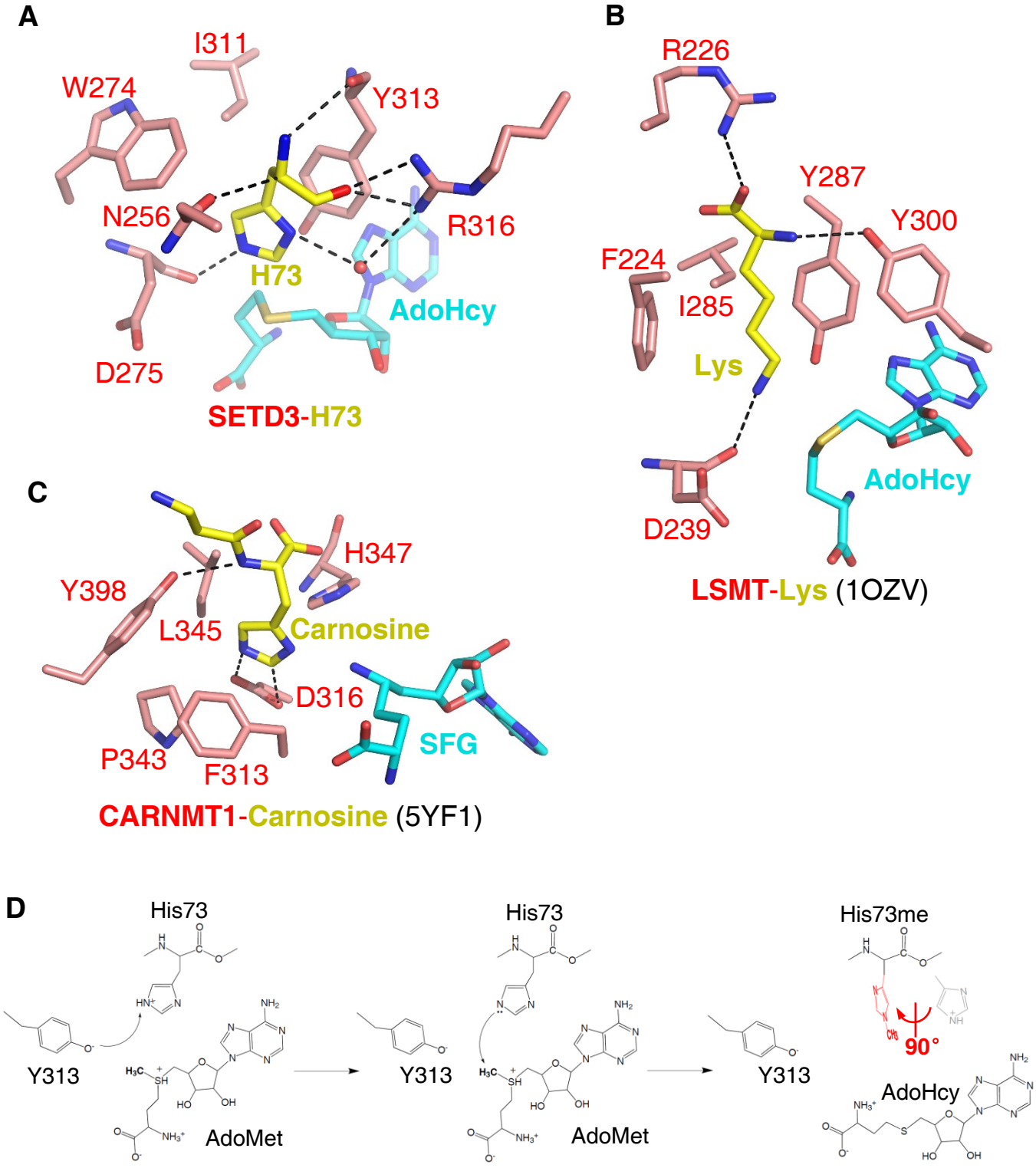

A
I311
W274
Y313
N256
R316
H73
AdoHcy
D275
SETD3-H73
B
R226
Y287
Y300
F224
I285
Lys
AdoHcy
D239
LSMT-Lys (1OZV)
C
H347
Y398
L345
Carnosine
D316
SFG
P343
F313
CARNMT1-Carnosine (5YF1)
D
His73
Y313
AdoMet
His73
Y313
AdoMet
His73me
90°
Y313
AdoHcy

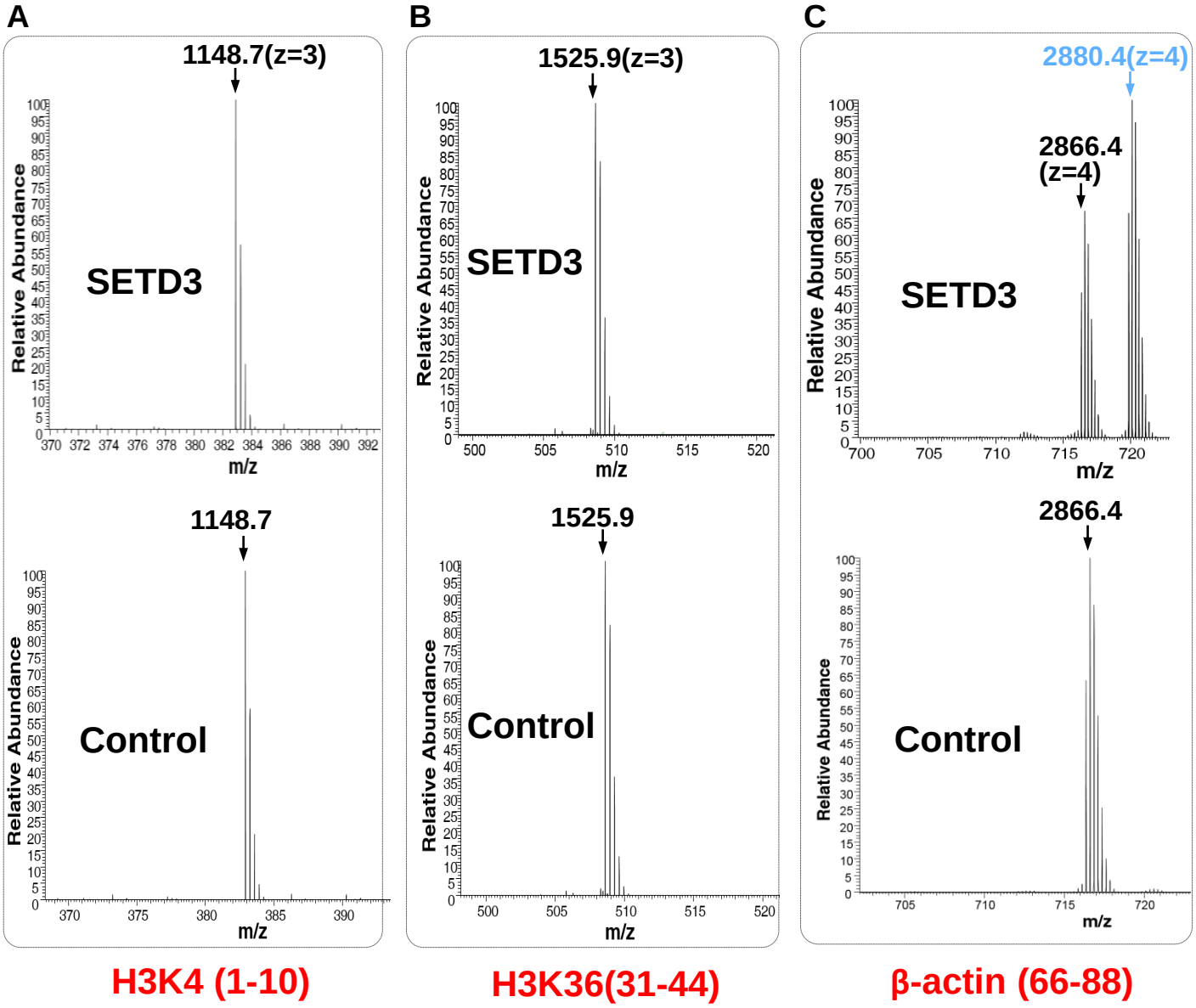
**Figure 5**

## Table 1. Data collection and refinement statistics

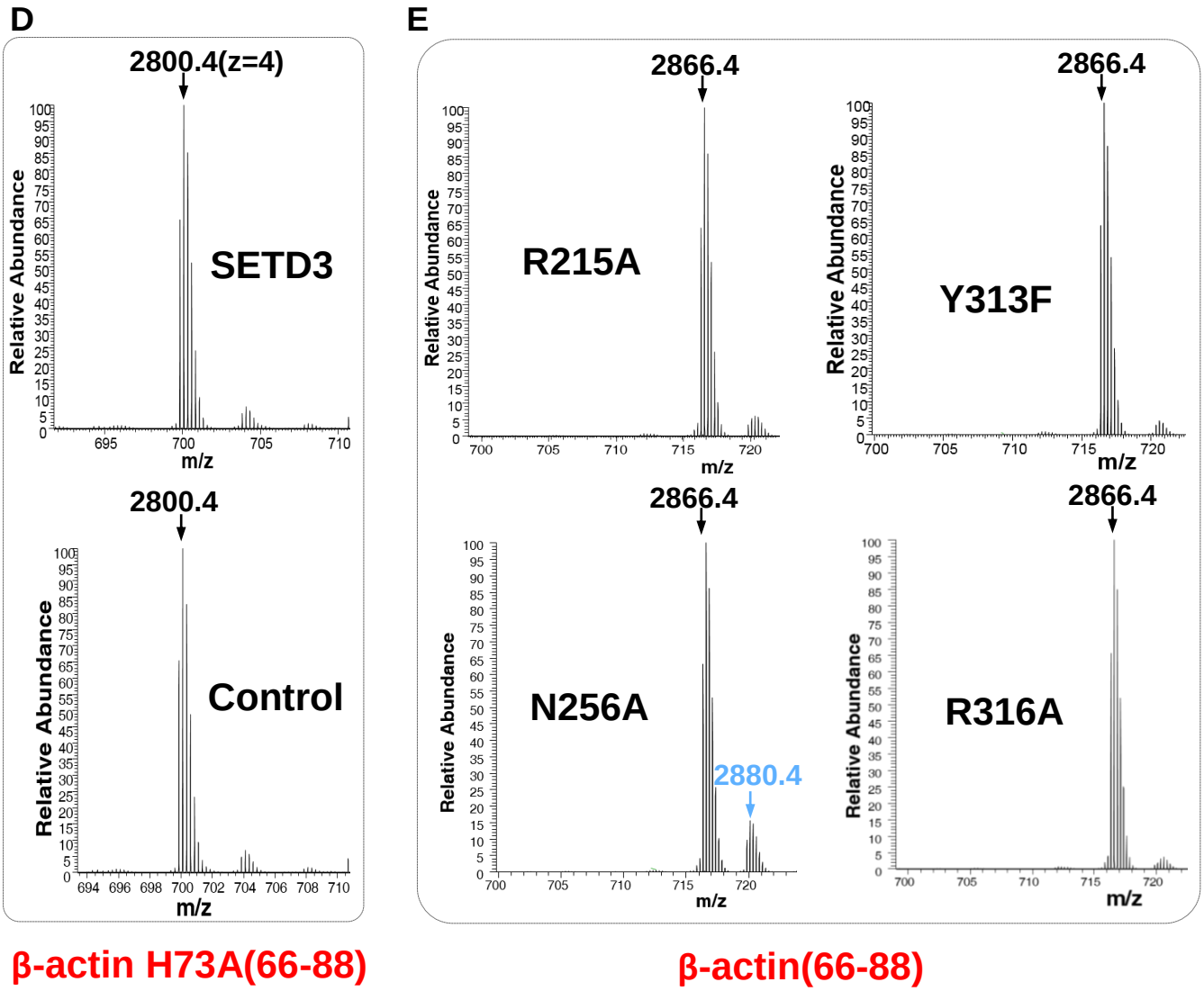
| | SETD3-β-actin | SETD3-methylated β-actin |
|---|---|---|
| **PDB id** | 6ICV | 6ICT |
| **Data collection** | | |
| Radiation wavelength (Å) | 0.9789 | 0.9792 |
| Space group | P 1 21 1 | P 1 21 1 |
| Cell dimensions | | |
| *a*, *b*, *c* (Å) | 60.19, 175.17, 66.50 | 59.98,176.69,125.89 |
| α, β, γ (°) | 90, 92.57, 90 | 90,93.37,90 |
| Resolution (Å) | 35.00-2.15(2.23-2.15) | 72.27-1.95(2.06-1.95) |
| $R_{merge}$ | 0.135(0.609) | 0.169(0.868) |
| $I / \sigma I$ | 14(3.4) | 8.1(2.6) |
| CC1/2 | 0.988(0.884) | 0.991(0.827) |
| Completeness (%) | 100(100) | 99.1(99.7) |
| Redundancy | 6.4(6.4) | 7.0(7.0) |
| **Refinement** | | |
| Resolution (Å) | 35.00-2.15 | 59.87-1.95 |
| No. of unique reflections | 74952 | 187364 |
| $R_{work}/R_{free}$ | 0.168/0.205 | 0.178/0.214 |
| Number of atoms/B-factor ($Å^2$) | 8493/30.7 | 16956/33.7 |
| Protein | 7574/30.3 | 15147/33.1 |
| Peptide | 276/32.9 | 591/38.4 |
| AdoHcy | 52/19.3 | 104/20.0 |
| Solvent | 591/36.2 | 1114/40.0 |
| RMSD bonds (Å)/angles (°) | 0.007/0.84 | 0.007/0.87 |
| Ramachandran Plot | | |
| favored/allowed/outliers (%) | 98.76/1.24/0 | 98.98/1.02/0 |

Values in parentheses are for highest-resolution shell.

# Supplemental Data

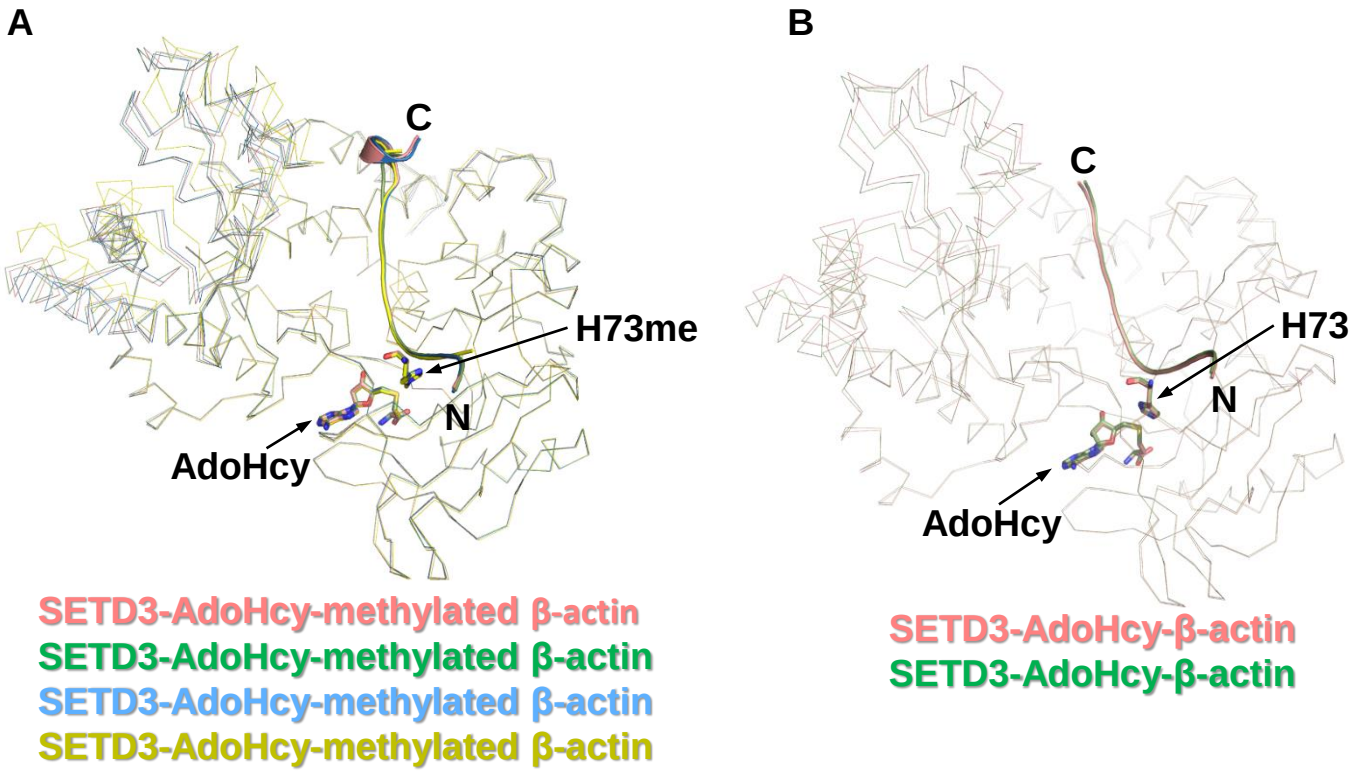



**Supplementary Figure S1 (To be continued)**



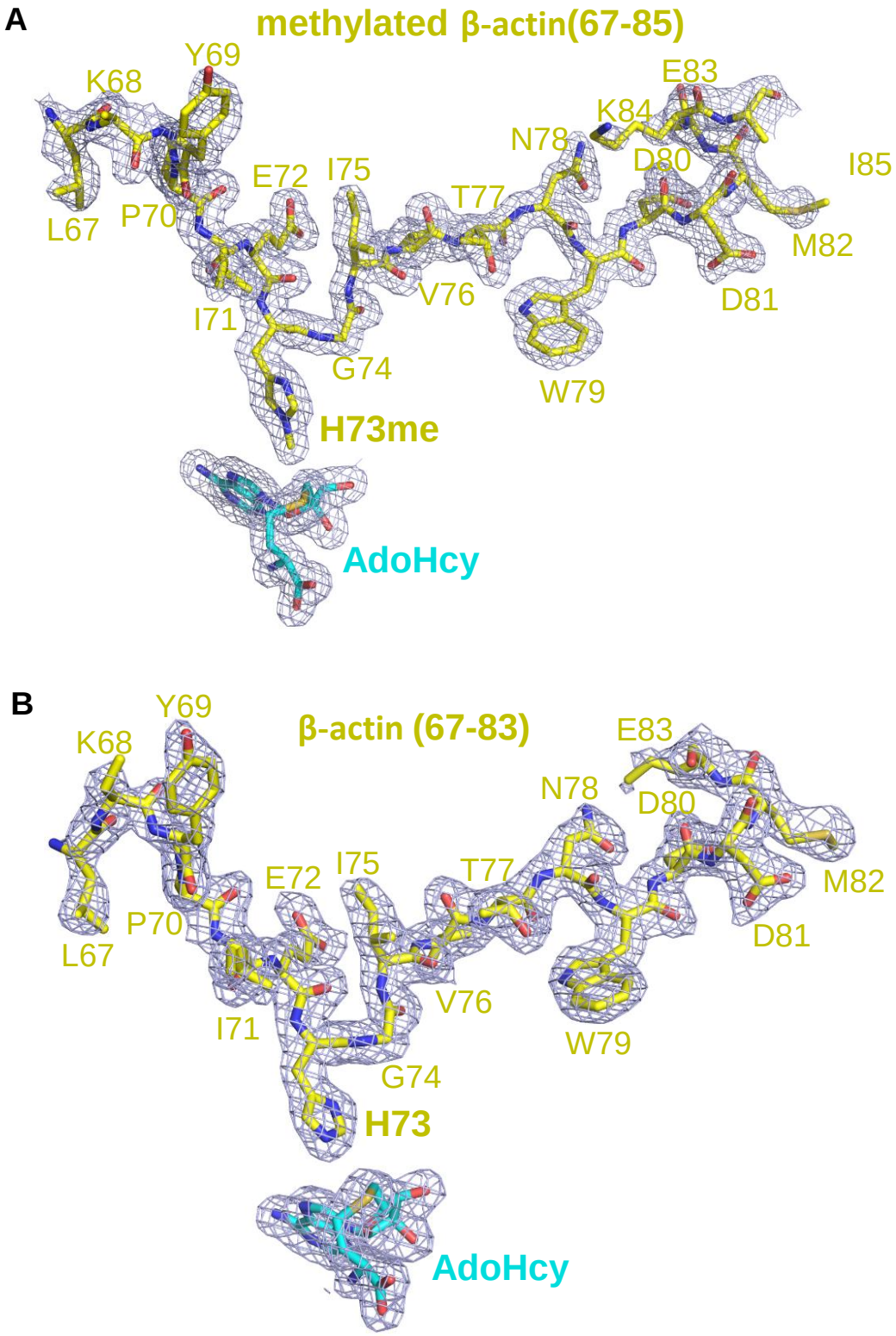

**Supplementary Figure S1.** The methylation activity of SETD3 (2-502) on the peptides, (A) H3K4 (1-10) and (B) H3K36 (31-44), (C) β-actin (66-88) and (D) β-actin H73A (66-88), as detected by mass spectrometry with peptide only as controls. (E) The activity of SETD3 mutants on β-actin (66-88) are shown for comparison with the wild type. The molecule weights of unmodified and methylated peptides are labeled in black and blue, respectively.

**Supplementary Figure S1 (To be continued)**




**Supplementary Figure S2.** (A) Superposition of four molecules of SETD3 bound to methylated β-actin in the same asymmetric unit, with the four molecules shown in red, green, blue and yellow ribbon, respectively. His73me of β-actin and AdoHcy are shown in sticks. The peptide adopts a $3^{10}$ helix at C-terminal end. (B) Superposition of two molecules of SETD3 bound to unmodified β-actin in the same asymmetric unit, with the two molecules shown in red and green ribbon, respectively. His73 of β-actin and AdoHcy are shown in sticks.

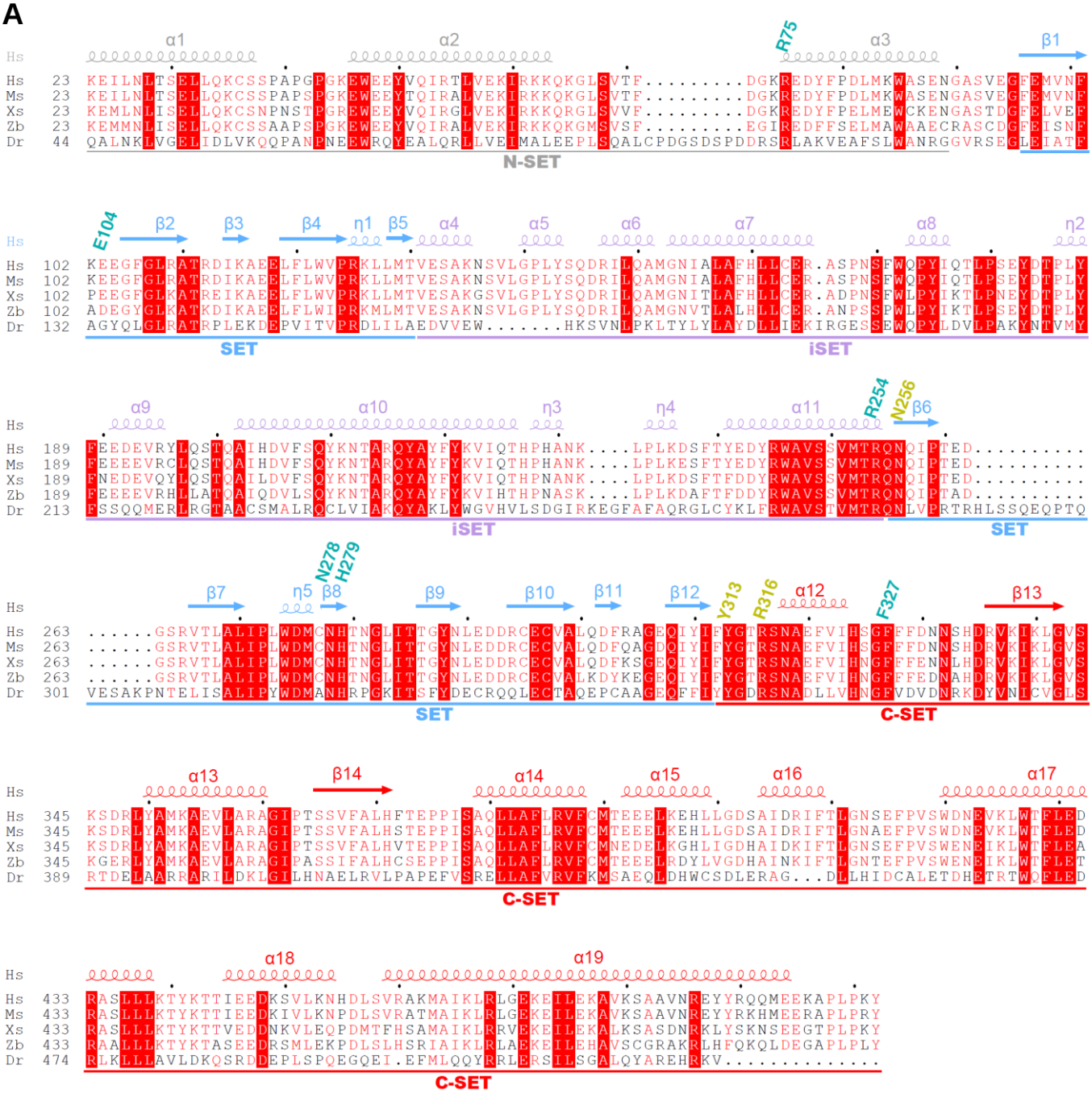
**Supplementary Figure S2**



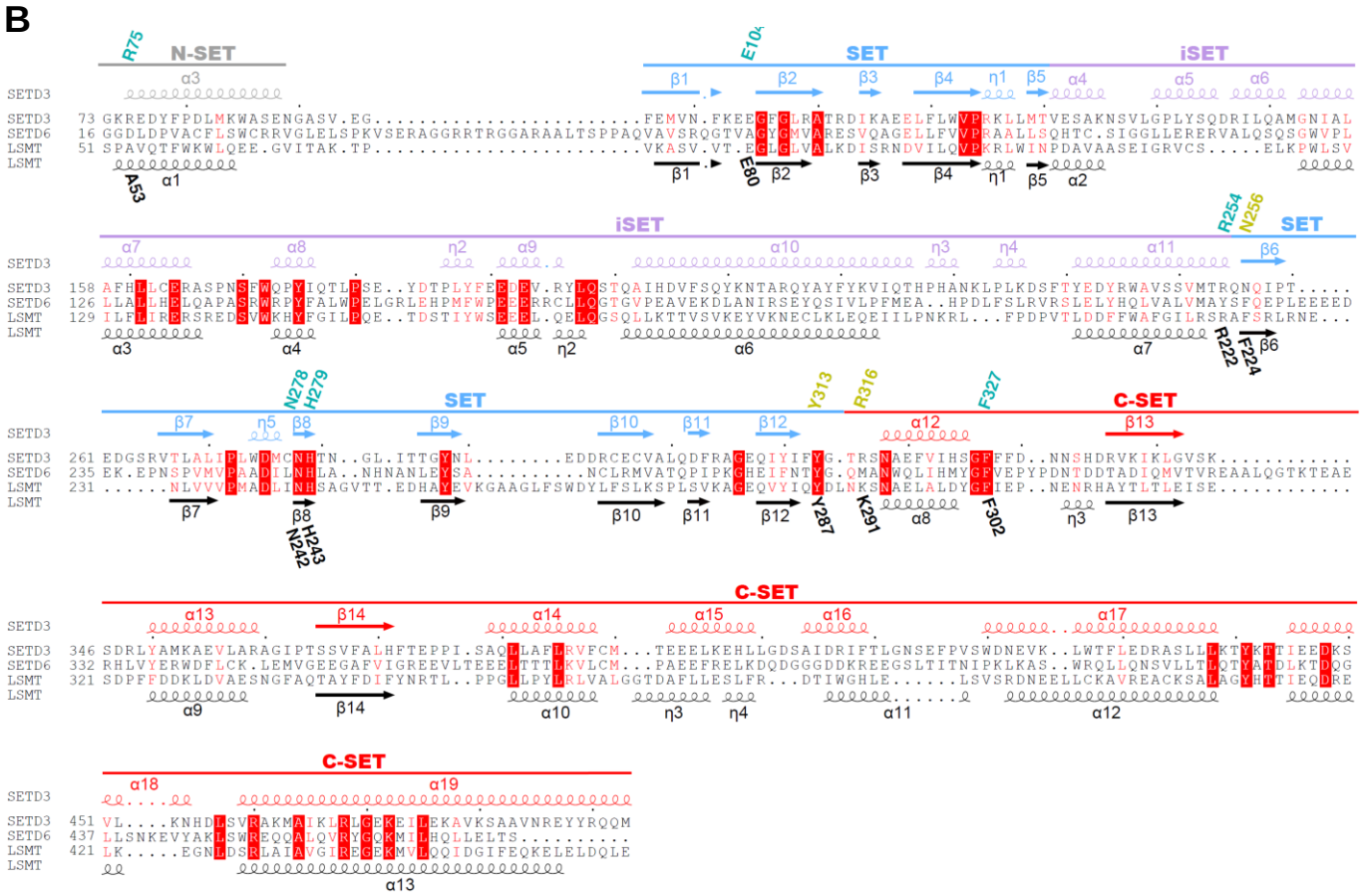

**Supplementary Figure S3.** The 2|Fo|–|Fc| σ-weighted maps of (A) methylated β-actin peptide and (B) unmodified β-actin peptide in the presence of AdoHcy, are contoured at 1.2 σ (blue cage), respectively. The peptides are shown in yellow sticks.

**Supplementary Figure S3**

**A**

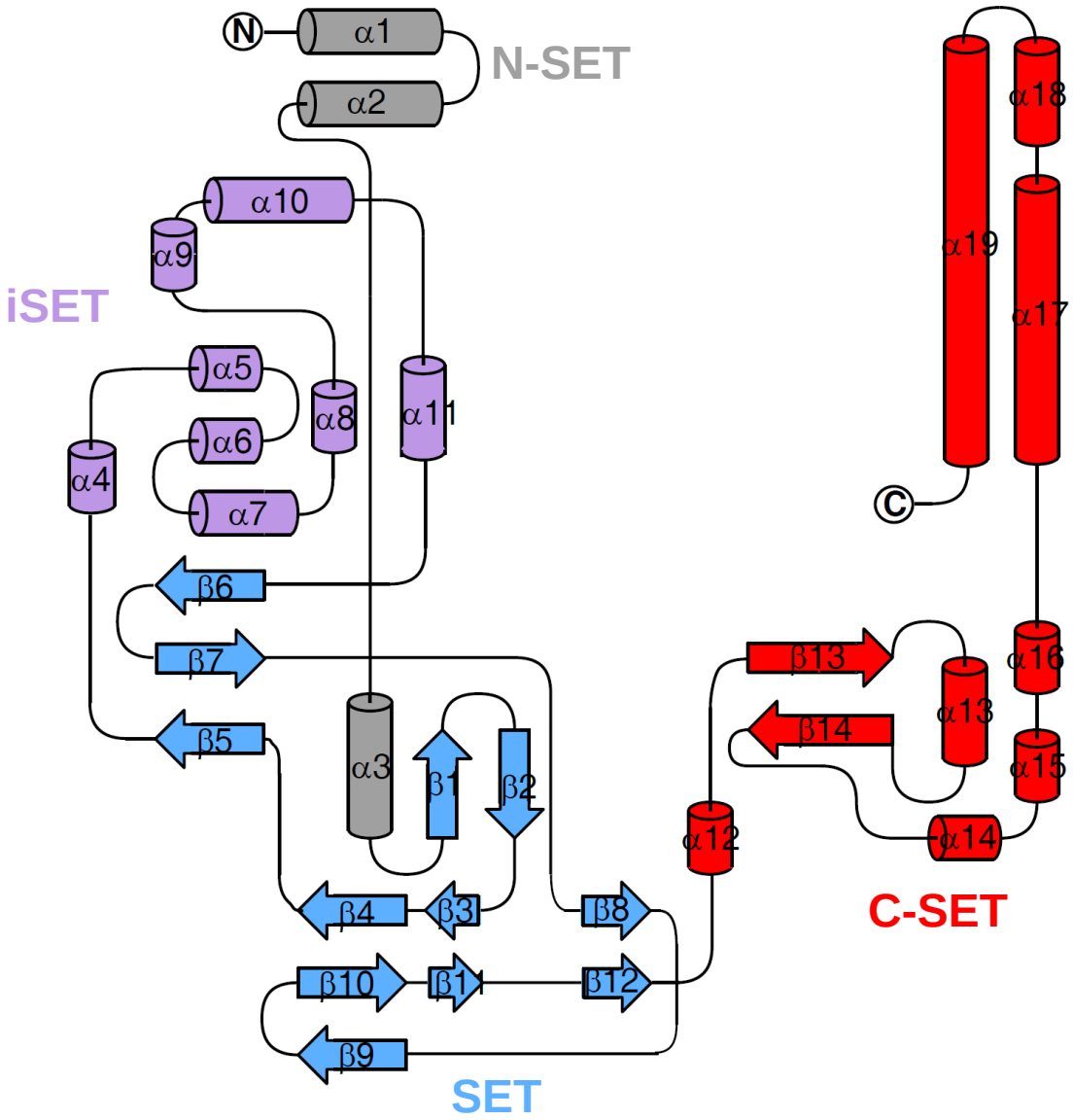



**Supplementary Figure S4A**



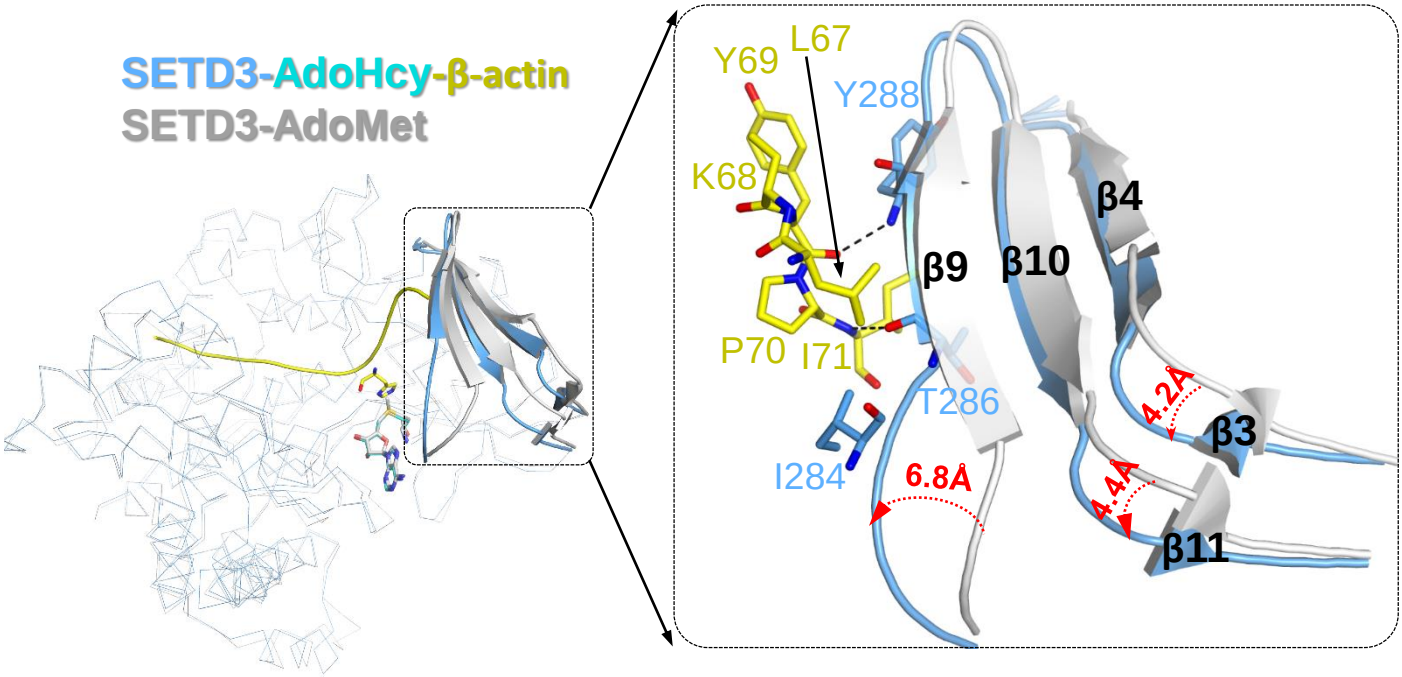

**Supplementary Figure S4.** Sequence alignment of human SETD3 with its orthologs or other SET domain proteins (SETD6 and LSMT). (A) Sequence alignment of Human SETD3(Hs, NP_115609.2), mouse SETD3(Ms, NP_082538.2), Xenopus SETD3(Xs, NP_001016577), Zebrafish SETD3 (Zb, NP_956348.1) and Drosophila (Dr, XP_017048262.1). The secondary structures of human SETD3 are colored in the same way as shown in Figure 1A. The residues involved in binding to AdoHcy and His73 are labeled in cyan and yellow, respectively. (B) Sequence aligment of human SETD3, human SETD6 (NP_001153777.1) and *Pisum sativum* LSMT (AAA69903.1). The secondary structures, AdoHcy binding residues, and His73 binding residues, are labeled in the same way as shown in Supplementary Figure S4A.

**Supplementary Figure S4B**



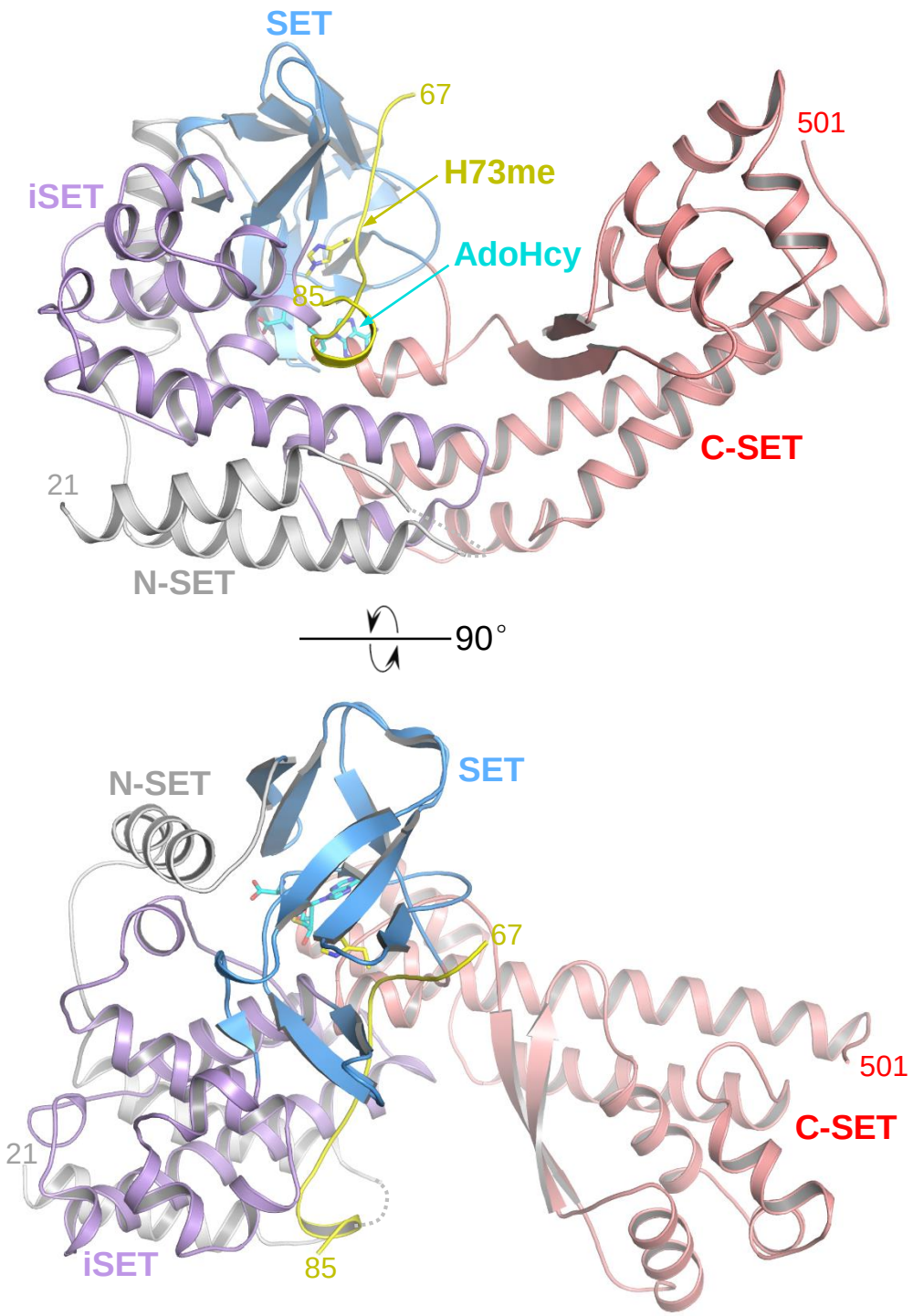

**Supplementary Figure S5.** Topology of SETD3 (2-502), with secondary structures marked and colored in the same way as shown in Supplementary Figure S4A.

**Supplementary Figure S5**



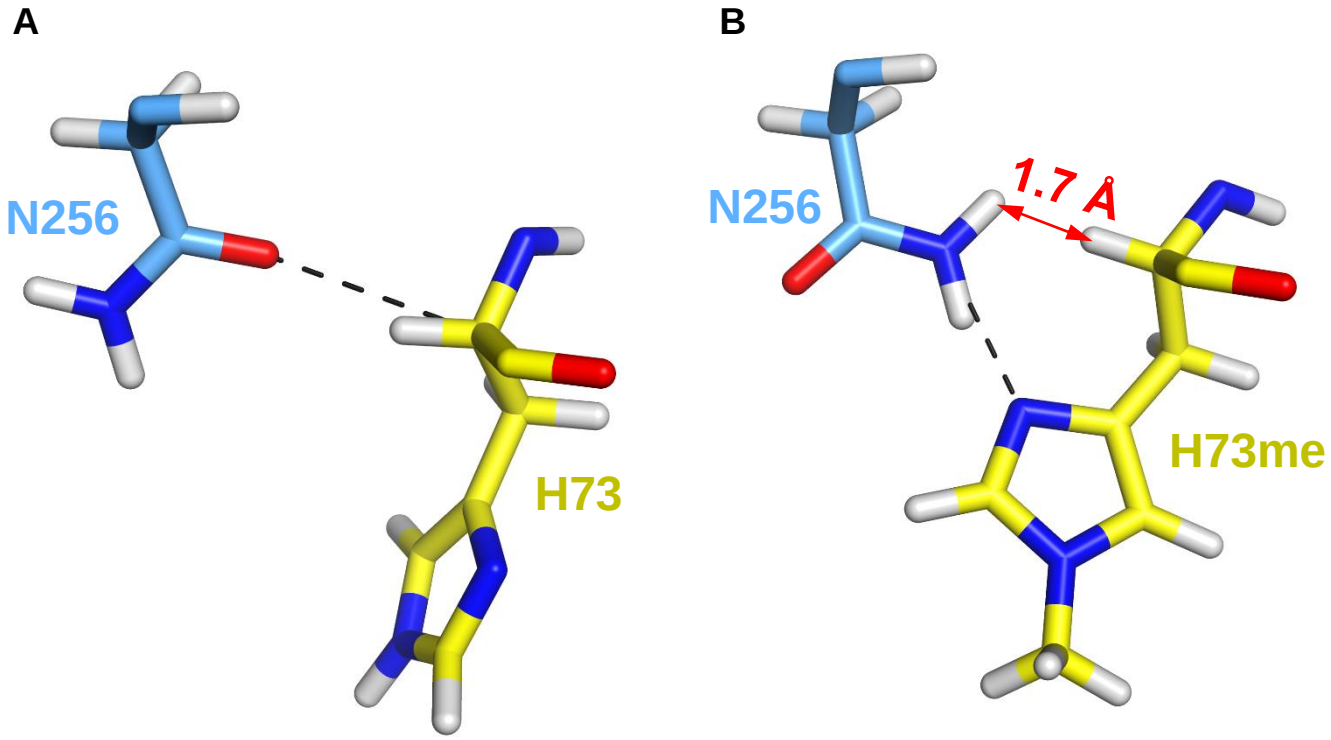

**Supplementary Figure S6.** Superposition of AdoMet bound SETD3 (PDB id: 3SMT) with AdoHcy bound SETD3-actin. In the structure of AdoMet bound SETD3, the protein is shown in grey ribbon and AdoMet is shown in grey sticks. In the structure of AdeHcy bound SETD3 in the presence of actin, protein is shown in blue ribbon and AdoHcy is shown in yellow sticks. β4-β9-β10 and β3-β11 changed conformations significantly upon binding to actin peptide, which are shown in cartoon representation. Leu67-Ile71 of β-actin and their interaction residues in SETD3, are shown in sticks. Three loops that precede β4, β9 and β11 shift 4.2Å, 4.4Å, and 6.8Å, respectively.

**Supplementary Figure S6**



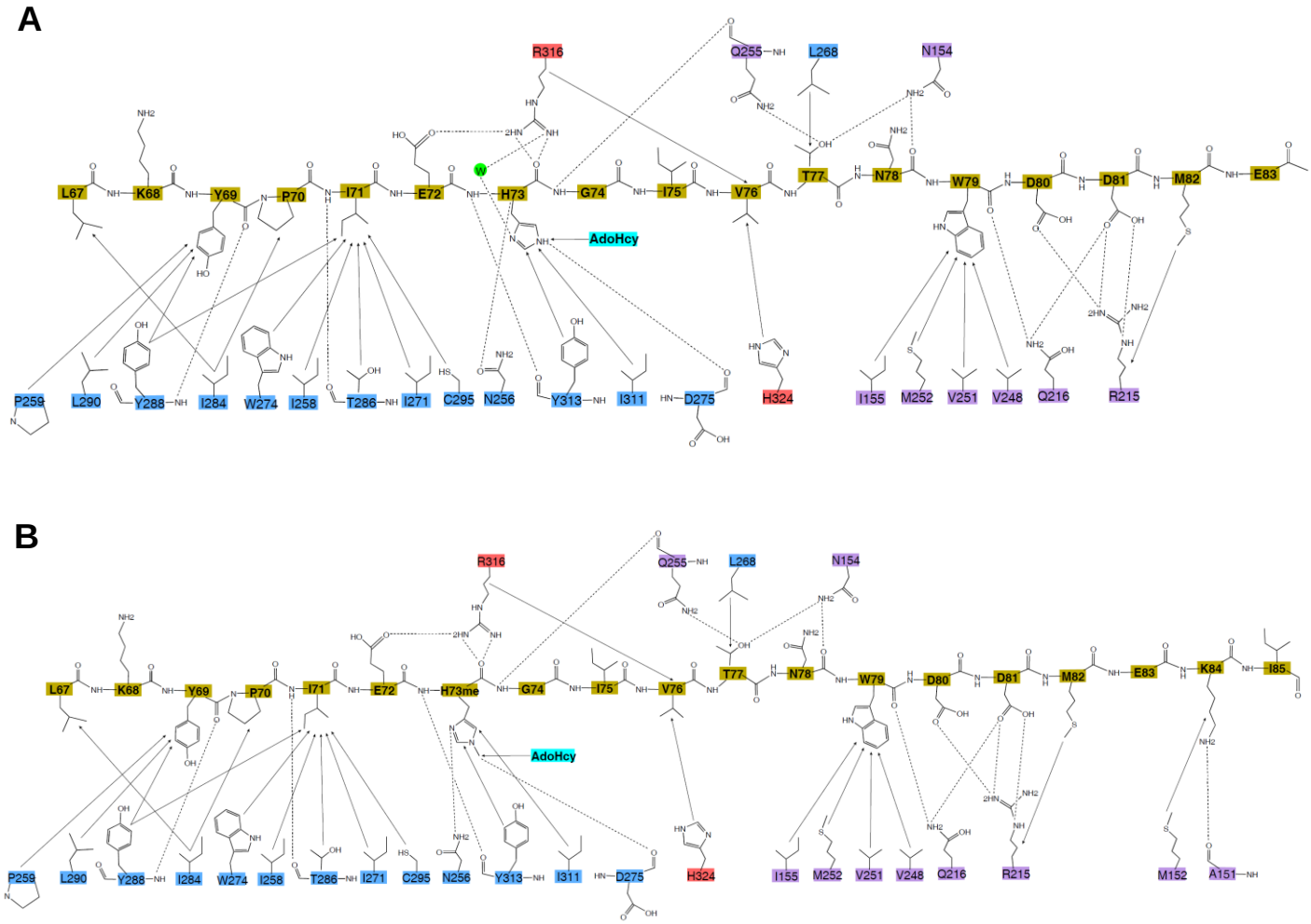

**Supplementary Figure S7.** Overall structure of SETD3 with methylated β-actin in the presence of AdoHcy. SETD3, which are colored in the same mode as shown in Figure 1C. Different domains of SETD3 are labeled. AdoHcy and His73 are also labeled.

**Supplementary Figure S7**



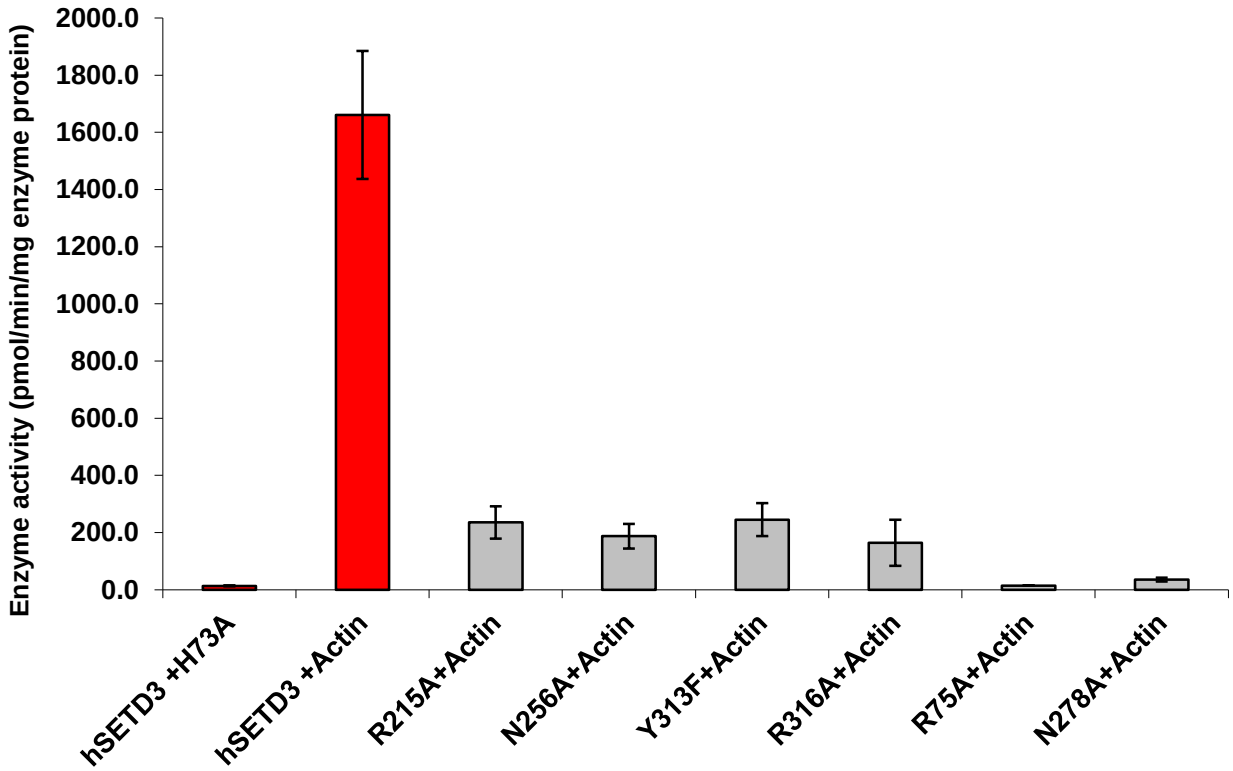

**Supplementary Figure S8.** Asn256 of SETD3 flips side chain after catalysis. (a) Asn256 of SETD3 forms one main chain hydrogen bond with Cα of β-actin His73 via its carboxyl group. (b) Asn256 flips side chain by 180 degree, making one side chain hydrogen bond with $N^1$ atom of His73me via its side chain amide. The distance between the side chain NH of Asn256 and the Hα of His73me, is only 1.7 Å, which would destabilize the complex.

**Supplementary Figure S8**



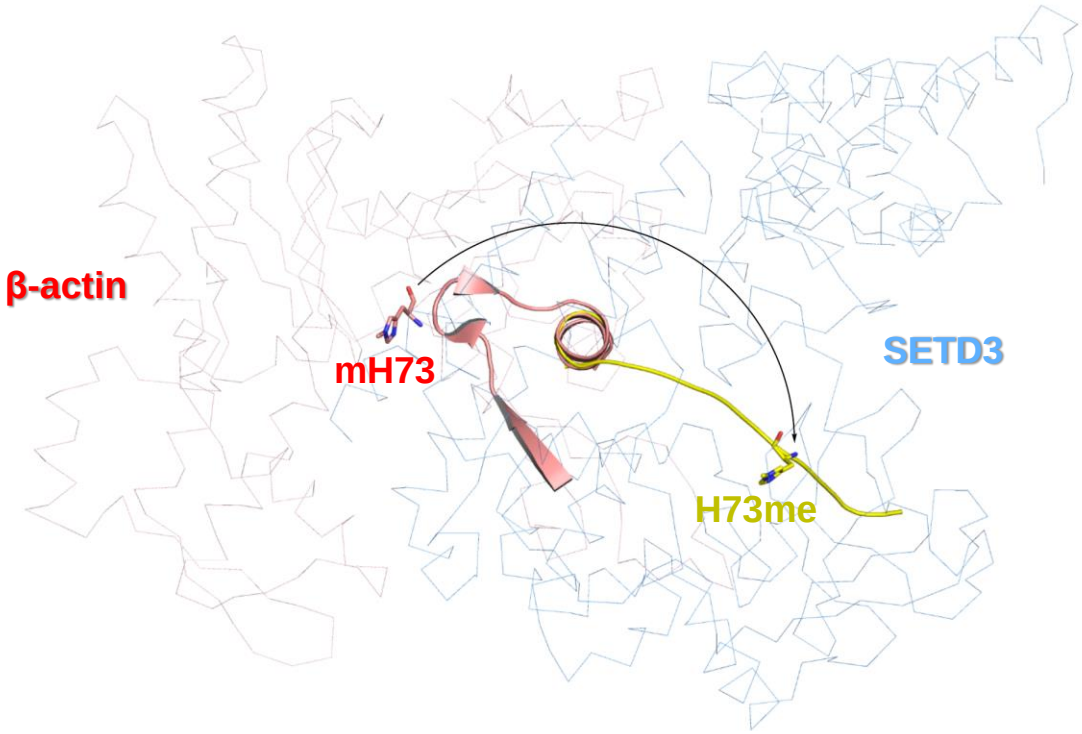

**Supplementary Figure S9.** Schematic of the detailed interactions (A) between SETD3 and unmodified β-actin (B) and between SETD3 and methylated β-actin. The SETD3 residues are colored as shown in Figure 1A, while the peptide residues are colored in yellow. The hydrophobic interactions and hydrogen bonds between protein and peptide are shown in black solid arrows and black dash arrows, respectively.

**Supplementary Figure S9**



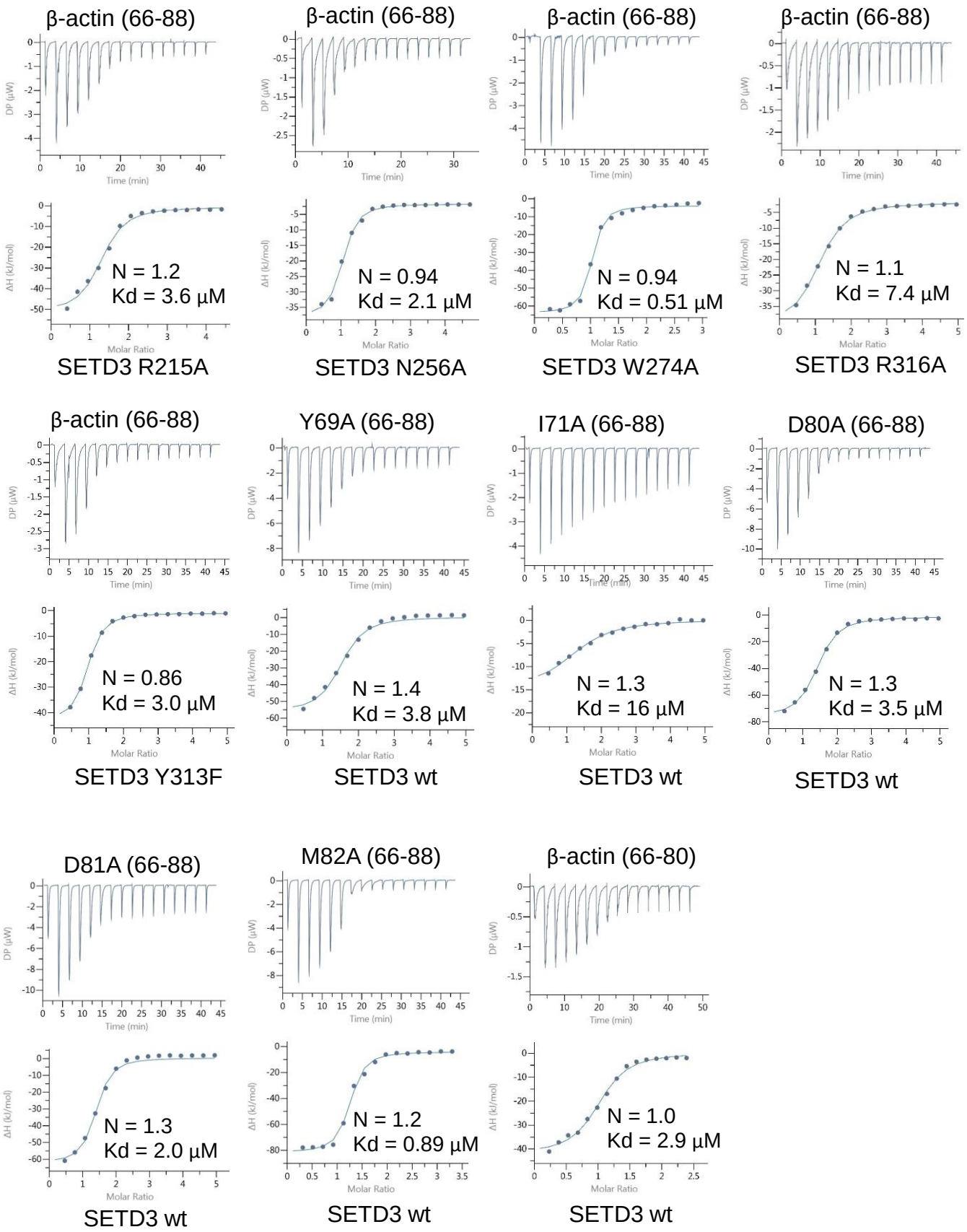

**Supplemental Figure S10.** SETD3 (2-502) displayed high enzyme activity towards FL wild type β-actin, but not the H73A mutant. The mutants exhibiting weaker peptide binding affinies also displayed weaker activity towards β-actin protein.

**Supplementary Figure S10**

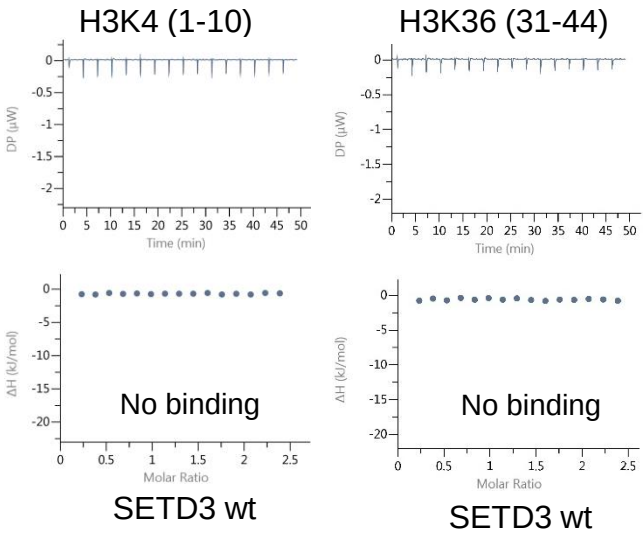



**Supplemental Figure S11.** Actin endures conformational changes upon binding to SETD3. β-actin peptide (66-88, yellow cartoon) is over laid with the same fragment in the structure of native β-actin (red cartoon) (PDB id: 1HLU). native β-actin decomposes its local secondary structures upon SETD3 binding. Native β-actin and SETD3 are shown in red and blue ribbon, respectively.

**Supplementary Figure S11**




**Supplementary Figure S12 (To be continued)**




**Supplemental Figure S12.** ITC binding curves for the binding measurements reported in Supplementary Table S1.

**Supplementary Figure S12**

**Supplementary Table S1. Binding of SETD3 proteins to β-actin peptides.**

| β-actin Peptide | SETD3 (2-502) | Stoichiometric coefficient (N) | Kd (μM) |
|---|---|---|---|
| $^{66}$TLKYPIEHGIVTNWDDMEKIWHH$^{88}$ | Wild type | 1.0 | 0.17 ± 0.04 |
| $^{66}$TLKYPIEAGIVTNWDDMEKIWHH$^{88}$ **(H73A)** | Wild type | 1.3 | 0.45 ± 0.08 |
| $^{1}$ARTKQTARKS$^{10}$ **(H3K4)** | Wild type | $^{\#}$N/A | *NB |
| $^{31}$ATGGVKKPHRYRPG$^{44}$ **(H3K36)** | Wild type | N/A | NB |
| $^{66}$TLKYPIEHGIVTNWDDMEKIWHH$^{88}$ | R215A | 1.2 | 3.6 ± 1.0 |
| $^{66}$TLKYPIEHGIVTNWDDMEKIWHH$^{88}$ | N256A | 0.94 | 2.1 ± 0.4 |
| $^{66}$TLKYPIEHGIVTNWDDMEKIWHH$^{88}$ | Y313F | 0.86 | 3.0 ± 0.1 |
| $^{66}$TLKYPIEHGIVTNWDDMEKIWHH$^{88}$ | R316A | 1.1 | 7.4 ± 0.9 |
| $^{66}$TLKAPIEHGIVTNWDDMEKIWHH$^{88}$ **(Y69A)** | Wild type | 1.4 | 3.8 ± 0.8 |
| $^{66}$TLKYPAEHGIVTNWDDMEKIWHH$^{88}$ **(I71A)** | Wild type | 1.3 | 16 ± 3 |
| $^{66}$TLKYPIEHGIVTNWADMEKIWHH$^{88}$ **(D80A)** | Wild type | 1.3 | 3.5 ± 0.4 |
| $^{66}$TLKYPIEHGIVTNWDAMEKIWHH$^{88}$ **(D81A)** | Wild type | 1.3 | 2.0 ± 0.5 |
| $^{66}$TLKYPIEHGIVTNWDDAEKIWHH$^{88}$ **(M82A)** | Wild type | 1.2 | 0.89 ± 0.08 |
| $^{66}$TLKYPIEHGIVTNWD$^{80}$ | Wild type | 1.0 | 2.9 ± 0.7 |

$^{\#}$: N/A: Not applicable; *: No detectable binding affinity by ITC.

Dissociation constants (Kds) were from a minimum of two experiments (mean ± sd). The original binding curves are shown in Supplementary Figure S12.