Rubisco in complex with Rubisco large subunit methyltransferase

Stefan Raunser[a,1], Roberta Magnani[b], Zhong Huang[c,2], Robert L. Houtz[b], Raymond C. Trievel[d], Pawel A. Penczek[c], and Thomas Walz[a,e]

[a]Department of Cell Biology and [e]Howard Hughes Medical Institute, Harvard Medical School, Boston, MA 02115; [b]Department of Horticulture, University of Kentucky, Lexington, KY 40546; [c]Department of Biochemistry and Molecular Biology, University of Texas Medical School, Houston, TX 77030; [d]Department of Biological Chemistry, University of Michigan, Ann Arbor, MI 48109

Edited by David J. DeRosier, Brandeis University, Waltham, MA, and approved December 31, 2008 (received for review October 20, 2008)

SET domain protein lysine methyltransferases (PKMT) are a structurally unique class of enzymes that catalyze the specific methylation of lysine residues in a number of different substrates. Especially histone-specific SET domain PKMTs have received widespread attention because of their roles in the regulation of epigenetic gene expression and the development of some cancers. Rubisco large subunit methyltransferase (RLSMT) is a chloroplast-localized SET domain PKMT responsible for the formation of trimethyl-lysine-14 in the large subunit of Rubisco, an essential photosynthetic enzyme. Here, we have used cryoelectron microscopy to produce an 11-Å density map of the Rubisco–RLSMT complex. The atomic model of the complex, obtained by fitting crystal structures of Rubisco and RLSMT into the density map, shows that the extensive contact regions between the 2 proteins are mainly mediated by hydrophobic residues and leucine-rich repeats. It further provides insights into potential conformational changes that may occur during substrate binding and catalysis. This study presents the first structural analysis of a SET domain PKMT in complex with its intact polypeptide substrate.

electron microscopy | LSMT | SET domain | single particle

SET [SU(VAR)3–9, E(Z), and TRX] domain protein lysine methyltransferases (PKMTs) are a structurally unique class of enzymes that catalyze the formation of site-specific methylated lysine residues in a number of different polypeptide substrates including cytochrome *c* (1, 2), histones (3), ribosomal proteins (3, 4), p53 (5), TAF10 (6), γ-tocopherol methyltransferase (7), and ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) (8). All SET domain PKMTs have a unique and conserved structural motif that contains separate binding sites for the target protein substrate and the methyl donor, *S*-adenosylmethionine (AdoMet) (9). The mechanisms by which these enzymes achieve site-specific methylation is of great interest and histone-specific SET domain protein methyltransferases (HKMTs) have received widespread attention because of their roles in the regulation of epigenetic gene expression and the development of some cancers (9, 10). The specificity of several SET domain PKMTs has been examined through structural and biochemical analyses of ternary complexes with bound polypeptide substrates, but all of these studies used short synthetic polypeptide mimetics of the intact polypeptide substrate and therefore do not account for the potential influence of areas outside the immediate residues flanking the target lysine methylation site (11–16). Rubisco large subunit methyltransferase (RLSMT) is a chloroplast-localized SET domain PKMT responsible for the formation of trimethyl-lysine-14 in the large subunit of Rubisco (8, 17), an essential photosynthetic enzyme with a hexadecameric structure and large molecular mass (≈534 kDa). Detailed information is available regarding the structure, catalytic mechanism, active site residues, and the kinetic reaction mechanism for pea RLSMT (7, 8, 17, 18) as well as Rubisco (for review see, ref. 19). In contrast to other PKMTs all of the existing biochemical characterization of pea RLSMT has used its intact polypeptide substrate, Rubisco, and these studies suggest that there are distinct interactions between the 2 proteins that create a tight and specific interaction before methylation of Lys-14 in the large subunit of Rubisco (8, 18). Here, we present a structure of the complex formed by pea RLSMT with spinach Rubisco obtained by single-particle cryoelectron microscopy (EM) that reveals extensive contact regions between the 2 proteins and provides insights into potential conformational changes that may occur during substrate binding and catalysis. Our findings may also apply to the interactions of other SET domain PKMTs with their intact polypeptide substrates, such as the interaction between histone PKMTs and nucleosome core particles.

Results and Discussion

The macromolecular complex between Rubisco and RLSMT represents an ideal model for a single-particle EM analysis of the interaction between a SET domain PKMT and its intact polypeptide substrate for a number of reasons. (*i*) High-resolution crystal structures are available for both Rubisco, mostly for the spinach enzyme (19), and RLSMT from pea (8, 17, 20), (*ii*) spinach Rubisco and pea RLSMT interact very tightly [K_D = ≈0.1 μM (18)], (*iii*) Rubisco is a sufficiently large protein (≈534 kDa) to be amenable to cryo-EM, and (*iv*) the core structure of Rubisco is a D4 symmetrical octamer of large subunits with the target lysine residue located in a random coil, solvent-accessible N-terminal region (19). Current structures describing the interactions between SET domain PKMTs and protein substrates are limited to short polypeptide substrates encompassing the target lysine methylation site (10). It is widely believed, however, that the recognition of polypeptide substrates by SET domain PKMTs extends well beyond the immediate lysine binding site and flanking cleft, to more remote regions of PKMTs (12). Two molecular docking models exist that support the idea of large contact regions between SET domain PKMTs and polypeptide substrates, one portraying the interaction of RLSMT with Rubisco (8), and the other the interaction of SET8 with a nucleosome core particle (12). Both models suggest that the N- and C-terminal flanking motifs (nSET and cSET) and the inserted α-helical domain

Author contributions: S.R., P.A.P., and T.W. designed research; S.R. and R.M. performed research; S.R., R.M., Z.H., R.L.H., R.C.T., P.A.P., and T.W. analyzed data; and S.R., R.L.H., R.C.T., P.A.P., and T.W. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

[1]To whom correspondence should be addressed at the present address: Max Planck Institute of Molecular Physiology, 44227 Dortmund, Germany. E-mail: raunser@mpi-dortmund.mpg.de.

[2]Present address: Institute of Molecular Biophysics, Florida State University, Tallahassee, FL 32306.

This article contains supporting information online at www.pnas.org/cgi/content/full/0810563106/DCSupplemental.

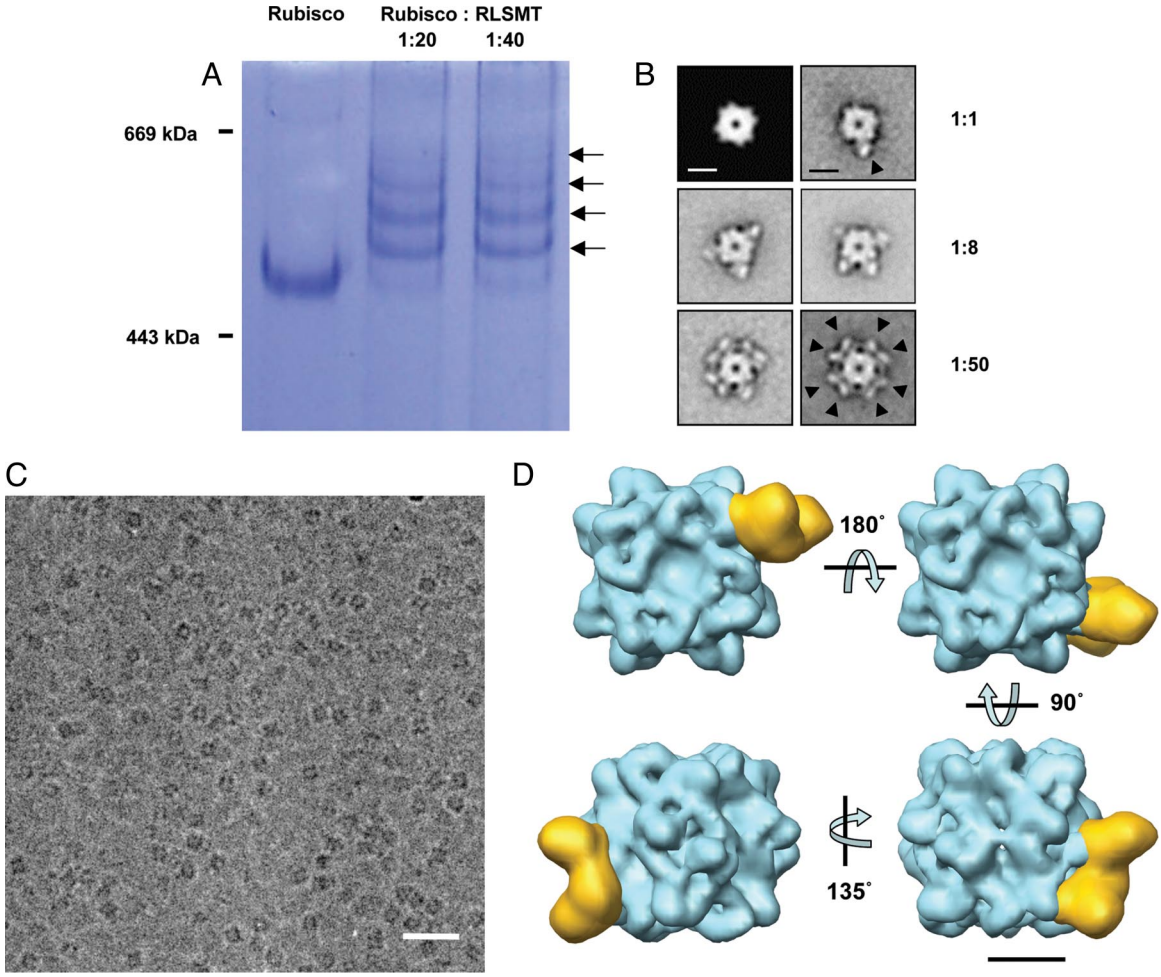

Fig. 1. Stoichiometry of the Rubisco–RLSMT complex and 3-dimensional reconstruction of the 1:1 complex. (*A*) Rubisco and RLSMT were incubated at a molar ratio of 1:20 and 1:40 with the cross-linker BMH for 30 min at room temperature. The cross-linked products were resolved on a native PAA-gel and stained with Coomassie blue. The bands corresponding to heterooligomers are marked with black arrows. (*B*) Single-particle electron microscopy of negatively stained Rubisco–RLSMT complexes. Rubisco and RLSMT were incubated at molar ratios of 1:1, 1:8, and 1:50 for 1 h at 4 °C before staining for EM. Shown are representative class averages of Rubisco–RLSMT complexes, each containing 100–300 particles. The black arrow heads in two of the class averages point to densities corresponding to RLSMT molecules bound to Rubisco. For comparison a volume was calculated from the spinach Rubisco crystal structure (PDB ID code 1RXO) (40), low-pass filtered to 25 Å and projected (*Top Left*). (Scale bar, 10 nm.) (*C*) A typical electron micrograph area of Rubisco–RLSMT complexes in vitrified ice. (Scale bar, 50 nm.) (*D*) Views of the Rubisco–RLSMT density map. The Rubisco complex is shown in blue and RLSMT in yellow. (Scale bar, 5 nm.)

bisecting the SET domain (iSET) could play important roles in substrate binding. RLSMT also contains a C-terminal lobe domain that comprises 176 residues, which is unique to RLSMT and may also be involved in substrate recognition (8).

To study its interaction with Rubisco, we used pea RLSMT, the only homolog that can be produced recombinantly with good methylation activity and that allowed the determination of its 3-dimensional structure by X-ray crystallography (8). However, pea Rubisco does not bind with any apparent or detectable affinity to methylated Rubisco, the only form of Rubisco that can be isolated from pea. We therefore used Rubisco purified from spinach for our studies, because it is not methylated in vivo (although spinach possesses an RLSMT homolog) and is an excellent substrate for pea RLSMT, which catalyzes its methylation on Lys-14 in the presence of AdoMet (18). Because the Rubisco complex consists of 8 symmetrically (D4) arranged small and large subunits, it has 8 putative binding sites for RLSMT. Cross-linking of the complex and subsequent analysis by native PAGE showed a ladder of increasing molecular mass entities beginning with native Rubisco, indicating that at least 4 RLSMTs can simultaneously bind to Rubisco (Fig. 1*A*). Because cross-linking studies are done at low protein concentrations to avoid nonspecific cross-linking artifacts, the likelihood is high that full occupancy is not reached, especially for complexes with a low binding affinity. We therefore examined Rubisco–RLSMT complexes by single-particle EM of negatively stained particles, which allowed us to use higher protein concentrations. The averages revealed that at a molar excess of RLSMT up to 8 RLSMT molecules can bind to Rubisco; thus, all binding sites on Rubisco can simultaneously be occupied by RLSMT [Fig. 1*B* and supporting information (SI) Fig. S1]. Because the RLSMT molecules bound to Rubisco are represented by clear and strong densities in the class averages, the observed interactions are likely to be specific. If RLSMT bound to Rubisco in a random, nonspecific manner, averaging of the images would have caused the densities representing the RLSMT molecules to smear out.

Full decoration of Rubisco with RLSMT required a high molar excess of RLSMT, but the resulting high background of unbound RLSMT made imaging the complex by cryo-EM impossible. Attempts to remove unbound RLSMT by gel filtration chromatography caused the complex to dissociate. We therefore used a Rubisco:RLSMT mixture at a molar ratio of 1:1

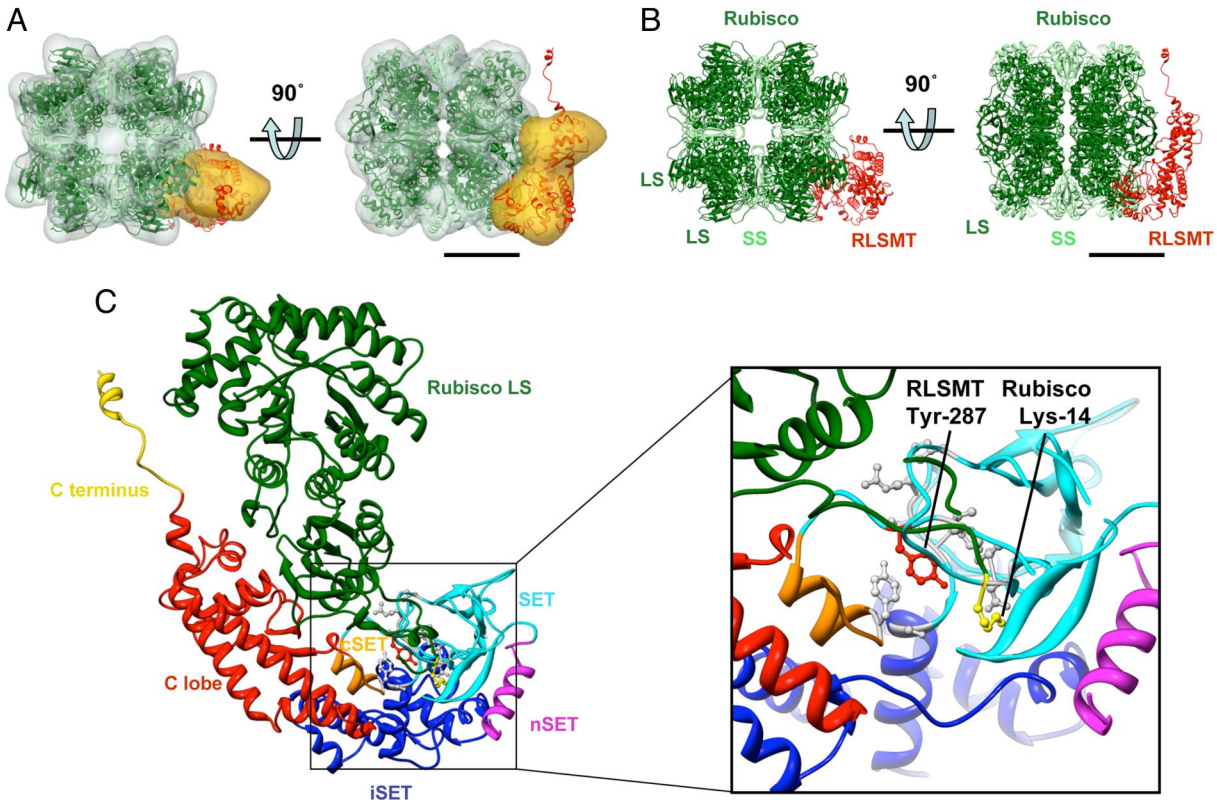

Fig. 2. Docking of the crystal structures into the EM density map. (*A* and *B*) The crystal structures of Rubisco (40) and RLSMT (8) were fit into the 3D reconstruction of the Rubisco–RLSMT complex. Rubisco large and small subunits are dark and light green, respectively. RLSMT is red. The C terminus of RLSMT protrudes from the density map, but it may be folded back toward the C lobe of the protein as modeled in Fig. S4. (Scale bar, 5 nm.) (*C*) A detailed view of the Rubisco–RLSMT interface at the catalytic site. The RLSMT nSET, iSET, cSET, and SET regions are magenta, blue, orange, and cyan, respectively. The C-terminal lobe is shown in red, and the domain-swapped C-terminal extension in gold. The Rubisco substrate Lys-14 and the invariant active site Tyr-287 in RLSMT are shown in yellow and red ball-and-stick representations, respectively. RLSMT residues involved in substrate binding are shown in gray ball-and-stick representation. A stereoview of *C* is shown in Fig. S6.

for imaging by cryo-EM (Figs. 1*C* and S2*A*). Image processing showed that ≈30% of Rubisco had one RLSMT bound at this mixing ratio (see *Materials and Methods*). Extensive image processing was performed (a detailed description is provided in *Materials and Methods*) to produce a density map from 29,000 single-particle images at 11-Å resolution (Fig. 1*D* and S2*B*), according to Fourier shell correlation (Fig. S2*D*). By fitting the crystal structures of Rubisco and RLSMT into this density map, we were able to build an atomic model of the complex (Fig. 2).

A caveat of our work is that we are studying the interaction of recombinant pea RLSMT with Rubisco isolated from a different species, spinach. Although RLSMT homologs are expressed in all plants, Rubisco's large subunit is not always methylated at Lys-14 in vivo. Rubisco is methylated, for example, in pea and tobacco, but not in corn, spinach, or wheat. To show in vitro binding of pea RLSMT with its natural substrate, unmethylated pea Rubisco, it would be necessary to isolate Rubisco from pea plants, in which RLSMT has been knocked down by RNA interference (RNAi). Transformation of pea plants is not well established, however. We therefore used tobacco plants, which are easily transformed and whose Rubisco, like in pea, is methylated in vivo. Tobacco Rubisco isolated from RLSMT knockdown plants proved to be a good substrate for pea RLSMT, which catalyzed methylation of the large subunit with kinetic parameters (K_m = 9.89 μM) similar to those when spinach Rubisco was used as substrate (K_m = 1.4 μM) (Fig. S3 *A* and *B*). Moreover, pea RLSMT formed a comparably tight complex with the unmethylated form of tobacco Rubisco (K_D = ≈0.4 μM) as with spinach Rubisco (K_D = ≈0.1 μM) (Fig. S3*C*) (18). These results thus suggest that the interaction of pea RLSMT with naturally unmethylated Rubiscos, such as spinach Rubisco, is representative of the interaction that occurs in species, in which Rubisco is naturally methylated.

Our model consisting of Rubisco with one bound RLSMT reveals that RLSMT binds primarily to the large subunit of Rubisco and not to the small subunit located at the 2 ends of the Rubisco molecule. Although the N-terminal SET domain of RLSMT is located close to the Rubisco Lys-14 site, the elongated C-lobe stretches along a pair of Rubisco large subunits.

Whereas both crystal structures fit well into the EM reconstruction, the C terminus of RLSMT protrudes from the density map (Fig. 2*A*). In the crystal structure, RLSMT formed a trimer that was held together by the C termini extending into the interlobe cleft of a neighboring molecule (8). Because RLSMT exists as a monomer in solution or in the Rubisco–RLSMT complex, it is very likely that this flexible domain folds back onto its own interlobe cleft as modeled in Fig. S4.

Our 3D density map of the Rubisco–RLSMT complex also reveals why 8 RLSMT molecules can bind to Rubisco without sterical hindrance (Figs. 1*B* and S1*C*). The N-terminal region of a Rubisco large subunit is paired with the C-terminal region of an adjacent large subunit, forming 2 catalytic sites (21). For any pair of large subunits the 2 target methylation sites at Lys-14 are thus located at opposite ends of the dimer, with the contact region for one RLSMT being formed by the C- and N-terminal domains of adjacent large subunits. The tadpole shape of RLSMT further helps to avoid steric constraints, and as a result all eight of the Lys-14 methylation sites can simultaneously be

occupied by RLSMT. These observations suggest that RLSMT has evolved to recognize pairs of large subunits.

The rigid body docking of the Rubisco and RLSMT crystal structures into the EM density map resulted in a very good overall fit, although some steric clashes can be seen between RLSMT and the N and C termini of the 2 Rubisco large subunits it interacts with (Fig. S5). In our docking, Lys-14 of Rubisco localizes proximal to the putative active site (Tyr-287) of RLSMT (Fig. 2*C* and Fig. S6). The distance between the hydroxyl group of RLSMT Tyr-287 and the nitrogen in the zeta position of Rubisco Lys-14 is ≈9 Å. Compared with the distance between these 2 residues of 3.5 Å in the crystal structure of RLSMT in ternary complex with lysine and AdoHcy (17), in our model Rubisco Lys-14 is not as well aligned with RLSMT Tyr-287. This finding and the observed steric clashes indicate that a conformational change has to occur to facilitate recognition of the large-subunit N terminus by the protein substrate-binding cleft of RLSMT. Structural and functional studies of Rubisco have revealed that the N-terminal tail of the large subunit is very mobile, which could thus easily move to accommodate binding to RLSMT (22–24). Modeling experiments predict that the most likely binding path for the N terminus of the large subunit is along the surface of the thread–loop motif in the SET domain (9). This path would enable the residues flanking Lys-14 in the large subunit to insert into the RLSMT substrate-binding site in a registry consistent with a previously published model for the interactions between the large-subunit N terminus and RLSMT (17). The residues preceding the Lys-14 methylation site may form additional interactions with RLSMT or may project away from the complex into solution.

Our model also reveals that the residues surrounding the entrance to the cofactor-binding cleft of RLSMT interact with Rubisco and partially occlude the AdoMet site (Fig. 2*C* and Fig. S6). These interactions presumably diminish the on-rate for AdoMet binding. In contrast, the protein substrate-binding site of RLSMT is relatively unobstructed, which would facilitate binding of the N terminus of the Rubisco large subunit. Therefore, a high-affinity collision complex initially formed by the 2 proteins would result in a low K_D value for Rubisco and RLSMT, whereas conformational changes that accompany productive binding of the N terminus of the Rubisco large subunit in the protein substrate-binding cleft of RLSMT would produce a higher K_m value. The same would be the case for AdoMet binding. Interestingly, K_m values for Rubisco and AdoMet were found to be 10- to 20-fold higher than their corresponding binding constants (Rubisco: K_m = 1.4 μM, K_D = ≈0.1 μM; AdoMet: K_m = 6.0 μM, K_D = 0.3 μM) (8, 20).

Calculations of the contact surface area between RLSMT and Rubisco (Fig. 3*A*) reveal that ≈4,156 Å^2 or 20% of the total surface area of RLSMT is directly involved in the interaction. According to the Kyte–Doolittle hydropathy index, of the residues involved in this interaction almost half are nonpolar with at least one third being hydrophobic (Table S1). Multiple sequence alignments of RLMSTs (Fig. S7) and Rubisco large subunits from different species (Fig. S8) and examination of the residues forming the contact surface area (Figs. S7–S9) show that residues mediating the interaction are conserved (69% identical and 83% homologous for the Rubisco large subunit and 32% identical and 53% homologous for RLSMT), suggesting that there may be common structural and/or amino acid sequence determinants in the specific interaction between RLSMT and Rubisco. The few amino acids that are variable between different Rubiscos do not seem to be specific for methylated and unmethylated forms of Rubisco. Therefore, the fact that Rubisco is not methylated in certain plant species does not appear to be due to mutations in residues that are required for its interaction with RLSMT. This conclusion is supported by the observation that pea RLSMT can methylate naturally unmethylated forms of Rubisco in vitro.

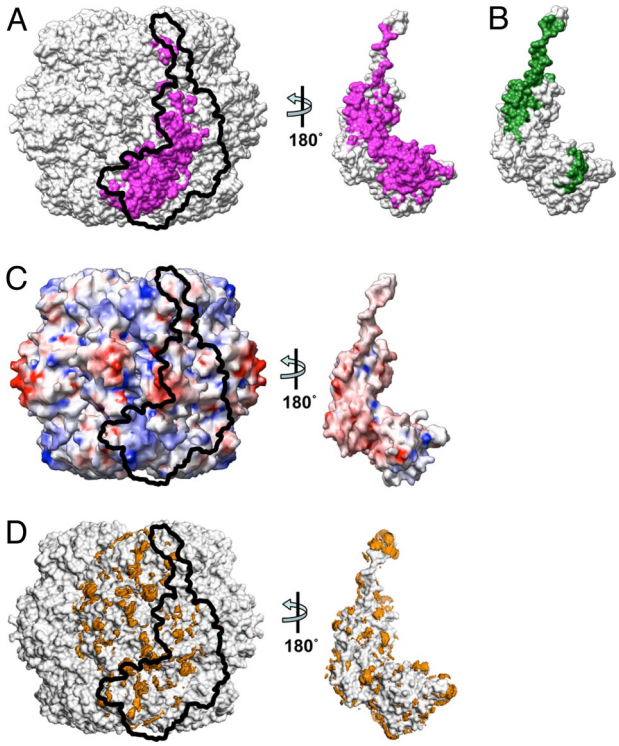

Fig. 3. The Rubisco–RLSMT interface. (*A*) The interface surface between Rubisco and RLSMT has been calculated by Intersurf (41). The contributing residues are magenta. (*B*) The 5 imperfect leucine-rich repeats are shown in green. (*C*) Calculated surface potential on Rubisco and RLSMT at neutral pH, with positive-charge density in blue and negative-charge density in red. (*D*) Hydrophobic potentials of the Rubisco large subunit and RLSMT surfaces calculated with the program GRID (39). Hydrophobic patches are shown in orange.

The kinetic reaction mechanism for RLSMT has been described as hybrid ping-pong, in which RLSMT remains bound to Rubisco through multiple methyl group additions to the target lysine residue (18). RLSMT's relatively large binding surface outside the immediate catalytic site may be an important component of this mechanism. Two previous observations support the idea of contact areas between Rubisco and RLSMT located outside the catalytic site as important in contributing to the interaction suggest that these may be hydrophobic in nature. First, alternative substrates for RLSMT composed of fusion protein constructs between the first 23 aa from the large subunit of Rubisco and human carbonic anhydrase II show an increase in the K_m for Rubisco from 1.4 μM to 1.2 mM (7). Second, affinity purification protocols for native RLSMT using immobilized Rubisco used 1 M KCl washes, which did not result in disassociation of the Rubisco–RLSMT complex (25). Similar observations have been reported for SET8, because its catalytic activity for short polypeptide substrates is substantially inhibited by 200 mM NaCl, whereas activity measured by using nucleosomes is largely unaffected (12). We measured the catalytic activity of RLSMT in the presence of increasing levels of NaCl and observed no effect up to a concentration of 800 mM, at which the activity was reduced by 25%. However, a full kinetic analysis of activity in the presence of 1 M NaCl revealed that there were increases in the K_m for Rubisco (from 1.4 μM to 9.5 μM) and AdoMet (from 6 μM to 46 μM) which, when accounted for by increased substrate levels, resulted in k_{cat} values equal to that observed in the absence of NaCl (Fig. 4). Thus, it seems unlikely that electrostatic interactions contribute significantly to the interaction between Rubisco and RLSMT. Electrostatic surface potential calculations reveal, how-

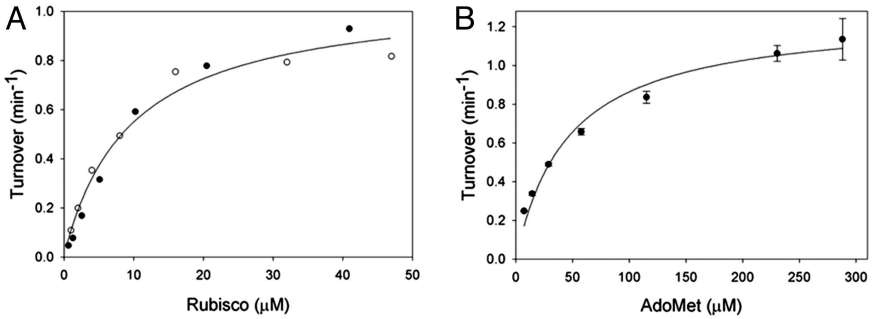

Fig. 4. Kinetic analyses of RLSMT activity in the presence of 1 M NaCl. (*A*) RLSMT turnover as a function of Rubisco concentration at 288 μM AdoMet, k_{cat} = 1.1 min^{-1}, K_m = 9.5 μM. (*B*) RLSMT turnover as a function of AdoMet concentration at 22 μM Rubisco, k_{cat} = 1.3 min^{-1}, K_m = 46.2 μM. In the absence of NaCl the RLSMT kinetic parameters were k_{cat} = 1.1 min^{-1}, K_m for Rubisco 1.4 μM, and K_m for AdoMet 6.0 μM. Data in *A* are from 2 separate experiments with different concentrations of Rubisco, and data in *B* are from 3 separate experiments using the same AdoMet concentrations. Kinetic parameters (K_m and k_{cat}) were obtained from fitted Michaelis-Menten equations using SigmaPlot 10.0.

ever, that the interface surfaces may also complement on charge, so that some contribution from electrostatic interactions cannot be excluded (Fig. 3*C*).

Because the resolution of our EM density map does not suffice to identify individual interacting residues, we focused our analysis on surface regions involved in the hydrophobic interaction between RLSMT and Rubisco. We calculated hydrophobic potentials across the surfaces of the Rubisco large subunit and RLSMT and identified hydrophobic patches at the Rubisco–RLSMT interface (Figs. 3*D* and Figs. S7 and S8). We found that most of the hydrophobic regions of the 2 proteins at the interface overlap and contain conserved residues. This finding agrees with our conclusion that hydrophobic interactions play an important role in the Rubisco–RLSMT interaction.

All RLSMTs examined to date contain 5 sequences resembling leucine-rich repeats (LRR), which are frequently associated with protein–protein interactions (26). All 5 LRR motifs contribute to the contact surface with Rubisco (Fig. 3*B* and Fig. S7). One is located in the N-terminal portion of the SET domain, and the remaining 4 LRRs are all in the C-terminal lobe region. Sequence analysis has not revealed the presence of LRRs in other SET domain PKMTs and the role of LRRs in the recognition and binding of polypeptide substrates may thus be unique to RLSMT.

Earlier molecular docking models suggested that the iSET region might be responsible for differences in polypeptide substrate specificity between different SET domain PKMTs (12). In our model of the complex, the iSET region in RLSMT is, however, located on the opposite side of where it interacts with Rubisco. It thus appears that the specificity of RLSMT for Rubisco is mainly mediated by the C-terminal lobe region and the SET domain itself, a notion that is supported by an examination of conserved residues in RLSMT (Figs. S7 and S9). This conclusion differs from a model of the SET8–nucleosome complex, in which the enzyme's iSET α-helix plays a prominent role in the interactions with the histone substrate (12). As might be expected, the N-terminal region of Rubisco containing the Lys-14 target methylation site is also highly conserved in nearly all higher plant forms of Rubisco (27) (Fig. S8).

The large contact area between RLSMT and Rubisco also provides potential insights into the temporal relationship between the synthesis of the Rubisco large subunit and its methylation at Lys-14. Because the RLSMT contact region includes both the C- and N-terminal domains of pairs of Rubisco large subunits, methylation of Lys-14 most likely takes place after synthesis of the large subunit and its assembly at least into dimers. Similarly, because RLSMT does not seem to interact with the Rubisco small subunit, it would be possible that methylation occurs before the addition of the small subunits to the octameric core formed by the large subunits. Finally, our structure of the Rubisco–RLSMT complex may be relevant to other important SET-domain PKMTs, such as those specific for histone polypeptides in nucleosome core particles, because the interactions these PKMTs form with their substrates may also involve and be regulated by extensive contacts distant from the immediate vicinity of the target lysine methylation site.

Materials and Methods

Detailed experimental procedures are provided in *SI Materials and Methods*.

Protein Production and Kinetics. Pea RLSMT was recombinantly expressed in *Escherichia coli* and purified as described in refs. 8 and 17. Rubisco was purified from spinach or tobacco leaves (28). Kinetic enzyme assays used previously determined conditions, which were optimized for linearity with time and enzyme concentration (18). Binding assays and determination of K_D for the association of pea RLSMT with PVDF-immobilized tobacco Rubisco from RLSMT knockdown plants was as described in ref. 18. Cross-linking of the complexes was done with 1,6-bismaleimidohexane (BMH), and the cross-linked products were resolved by native gel electrophoresis.

Electron Microscopy. Rubisco–RLSMT complexes mixed at different ratios were prepared for electron microscopy by negative staining with uranyl formate as described in ref. 29. For cryo-EM, Rubisco–RLSMT mixtures at a molar ratio of 1:1 were applied to holey carbon grids and frozen in liquid ethane. Images were recorded with FEI Tecnai T12 and Tecnai F20 electron microscopes on imaging plates for negatively stained specimens and on Kodak SO-163 film for vitrified specimens by using low-dose conditions. Magnifications were 67,000× and 50,000×, respectively. Electron micrographs were digitized by using a Zeiss SCAI scanner with a 7-μm step size. The 3 × 3 pixels were averaged, yielding a pixel size of 4.1 Å on the specimen scale. CTFTILT (30) was used to determine defocus values for all of the digitized micrographs.

Image Processing. Images of negatively stained specimens were processed by using the Spider software package (31). Images of vitrified complexes were processed by using SPARX (32). The mixture of asymmetric and D4-symmetric objects required a protocol specifically designed to determine the 3D structure of the Rubisco–RLSMT complex (see flow chart depicted in Fig. S10). To confirm the quality of our dataset we first calculated a structure of vitrified Rubisco by using an ab initio structure determination method (33) (Fig. S11). We then analyzed the variance in the 2D data to localize where RLSMT binds to Rubisco. 2D variance maps were aligned such that the resulting density fields reached maximum density when back-projected into a 3D volume (Fig. S12). To take advantage of the D4 symmetry of the Rubisco core of the complex, while at the same time reconstructing the entire Rubisco–RLSMT complex, we modified the 3D projection alignment procedure and used local symmetrization of the reconstructed 3D object. We first computed a 3D reconstruction by using the negative stain dataset (Fig. S13). We then calculated a 3D reconstruction with the cryo-EM dataset by using the volume obtained with the images of the negatively stained sample as the initial model (Fig. S2). To separate our dataset into images representing Rubisco alone and Rubisco–RLSMT complexes, we performed a multireference 3D projection alignment by using the reconstructions of Rubisco and the Rubisco–RLSMT complex as templates (34) (Fig. S14).

The crystal structures of RLSMT and Rubisco were docked into the final map using rigid body fitting in Situs (35). Figures were generated by using the programs PyMOL (36) and Chimera (37). Electrostatic surface potentials were calculated with the program GRASP (38), and hydrophobic potentials on the surface of Rubisco and RLSMT with the program GRID (39).

ACKNOWLEDGMENTS. We thank L. M. A. Dirk for research assistance and I. R. Vetter for help with the GRID software. This work was supported by National Institutes of Health Grant GM60635 (to P.A.P.), National Institutes of Health Grant GM073839 (to R.C.T.), and Department of Energy Grant DE-FG02-92ER20075 (to R.L.H.). T.W. is an investigator of the Howard Hughes Medical Institute. S.R. is a fellow of the German Academy of Sciences Leopoldina (BMBF-LPD 9901/8-163). The molecular EM facility at Harvard Medical School was established with a generous donation from the Giovanni Armenise Harvard Center for Structural Biology and is supported by National Institutes of Health Grant GM62580 (to S. C. Harrison).

1. Polevoda B, Martzen MR, Das B, Phizicky EM, Sherman F (2000) Cytochrome c methyltransferase, Ctm1p, of yeast. *J Biol Chem* 275(27):20508–20513.
2. Porras-Yakushi TR, Whitelegge JP, Clarke S (2007) Yeast ribosomal/cytochrome c SET domain methyltransferase subfamily: Identification of Rpl23ab methylation sites and recognition motifs. *J Biol Chem* 282(17):12368–12376.
3. Dillon SC, Zhang X, Trievel RC, Cheng X (2005) The SET-domain protein superfamily: Protein lysine methyltransferases. *Genome Biol* 6(8):227.
4. Porras-Yakushi TR, Whitelegge JP, Clarke S (2006) A novel SET domain methyltransferase in yeast: Rkm2-dependent trimethylation of ribosomal protein L12ab at lysine 10. *J Biol Chem* 281(47):35835–35845.
5. Chuikov S, *et al.* (2004) Regulation of p53 activity through lysine methylation. *Nature* 432(7015):353–360.
6. Kouskouti A, Scheer E, Staub A, Tora L, Talianidis I (2004) Gene-specific modulation of TAF10 function by SET9-mediated methylation. *Mol Cell* 14(2):175–182.
7. Magnani R, Nayak NR, Mazarei M, Dirk LM, Houtz RL (2007) Polypeptide substrate specificity of PsLSMT. A set domain protein methyltransferase. *J Biol Chem* 282(38):27857–27864.
8. Trievel RC, Beach BM, Dirk LM, Houtz RL, Hurley JH (2002) Structure and catalytic mechanism of a SET domain protein methyltransferase. *Cell* 111(1):91–103.
9. Xiao B, Wilson JR, Gamblin SJ (2003) SET domains and histone methylation. *Curr Opin Struct Biol* 13(6):699–705.
10. Couture JF, Trievel RC (2006) Histone-modifying enzymes: Encrypting an enigmatic epigenetic code. *Curr Opin Struct Biol* 16(6):753–760.
11. Xiao B, *et al.* (2003) Structure and catalytic mechanism of the human histone methyltransferase SET7/9. *Nature* 421(6923):652–656.
12. Xiao B, *et al.* (2005) Specificity and mechanism of the histone methyltransferase Pr-Set7. *Genes Dev* 19(12):1444–1454.
13. Couture JF, Collazo E, Brunzelle JS, Trievel RC (2005) Structural and functional analysis of SET8, a histone H4 Lys-20 methyltransferase. *Genes Dev* 19(12):1455–1465.
14. Couture JF, Collazo E, Hauk G, Trievel RC (2006) Structural basis for the methylation site specificity of SET7/9. *Nat Struct Mol Biol* 13(2):140–146.
15. Qian C, *et al.* (2006) Structural insights of the specificity and catalysis of a viral histone H3 lysine 27 methyltransferase. *J Mol Biol* 359(1):86–96.
16. Zhang X, *et al.* (2003) Structural basis for the product specificity of histone lysine methyltransferases. *Mol Cell* 12(1):177–185.
17. Trievel RC, Flynn EM, Houtz RL, Hurley JH (2003) Mechanism of multiple lysine methylation by the SET domain enzyme Rubisco LSMT. *Nat Struct Biol* 10(7):545–552.
18. Dirk LM, *et al.* (2007) Kinetic manifestation of processivity during multiple methylations catalyzed by SET domain protein methyltransferases. *Biochemistry* 46(12):3905–3915.
19. Andersson I, Backlund A (2008) Structure and function of Rubisco. *Plant Physiol Biochem* 46(3):275–291.
20. Couture JF, Hauk G, Thompson MJ, Blackburn GM, Trievel RC (2006) Catalytic roles for carbon-oxygen hydrogen bonding in SET domain lysine methyltransferases. *J Biol Chem* 281(28):19280–19287.
21. Andersson I (1996) Large structures at high resolution: The 1.6 A crystal structure of spinach ribulose-1,5-bisphosphate carboxylase/oxygenase complexed with 2-carboxyarabinitol bisphosphate. *J Mol Biol* 259(1):160–174.
22. Houtz RL, Mulligan RM (1991) Protection of tryptic-sensitive sites in the large subunit of ribulosebisphosphate carboxylase/oxygenase by catalysis. *Plant Physiol* 96(1):335–339.
23. Mulligan RM, Houtz RL, Tolbert NE (1988) Reaction-intermediate analogue binding by ribulose bisphosphate carboxylase/oxygenase causes specific changes in proteolytic sensitivity: The amino-terminal residue of the large subunit is acetylated proline. *Proc Natl Acad Sci USA* 85(5):1513–1517.
24. Duff AP, Andrews TJ, Curmi PM (2000) The transition between the open and closed states of rubisco is triggered by the inter-phosphate distance of the bound bisphosphate. *J Mol Biol* 298(5):903–916.
25. Wang P, Royer M, Houtz RL (1995) Affinity purification of ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit epsilon N-methyltransferase. *Protein Expression Purif* 6(4):528–536.
26. Kobe B, Deisenhofer J (1995) Proteins with leucine-rich repeats. *Curr Opin Struct Biol* 5(3):409–416.
27. Houtz RL, Magnani R, Nayak NR, Dirk LM (2008) Co- and post-translational modifications in Rubisco: Unanswered questions. *J Exp Bot* 59(7):1635–1645.
28. McCurry SD, Gee R, Tolbert NE (1982) Ribulose-1,5-bisphosphate carboxylase/oxygenase from spinach, tomato, or tobacco leaves. *Methods Enzymol* 90(Pt E):515–521.
29. Ohi M, Li Y, Cheng Y, Walz T (2004) Negative Staining and Image Classification - Powerful Tools in Modern Electron Microscopy. *Biol Proced Online* 6:23–34.
30. Mindell JA, Grigorieff N (2003) Accurate determination of local defocus and specimen tilt in electron microscopy. *J Struct Biol* 142(3):334–347.
31. Frank J, *et al.* (1996) SPIDER and WEB: Processing and visualization of images in 3D electron microscopy and related fields. *J Struct Biol* 116(1):190–199.
32. Hohn M, *et al.* (2007) SPARX, a new environment for Cryo-EM image processing. *J Struct Biol* 157(1):47–55.
33. Penczek PA, Zhu J, Frank J (1996) A common-lines based method for determining orientations for N $>$ 3 particle projections simultaneously. *Ultramicroscopy* 63(3–4):205–218.
34. Brink J, *et al.* (2004) Experimental verification of conformational variation of human fatty acid synthase as predicted by normal mode analysis. *Structure* 12(2):185–191.
35. Wriggers W, Milligan RA, McCammon JA (1999) Situs: A package for docking crystal structures into low-resolution maps from electron microscopy. *J Struct Biol* 125(2–3):185–195.
36. DeLano WL (2004) Use of PyMOL as a communications tool for molecular science. *Abstr Papers Am Chem Society* 228:U313–U314.
37. Pettersen EF, *et al.* (2004) UCSF Chimera–A visualization system for exploratory research and analysis. *J Comput Chem* 25(13):1605–1612.
38. Petrey D, Honig B (2003) GRASP2: Visualization, surface properties, and electrostatics of macromolecular structures and sequences. *Methods Enzymol* 374:492–509.
39. Goodford PJ (1985) A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J Med Chem* 28(7):849–857.
40. Taylor TC, Andersson I (1997) The structure of the complex between rubisco and its natural substrate ribulose 1,5-bisphosphate. *J Mol Biol* 265(4):432–444.
41. Ray N, Cavin X, Paul JC, Maigret B (2005) Intersurf: Dynamic interface between proteins. *J Mol Graphics Model* 23(4):347–354.