Project Title:	Mechanism and Significance of Post-Translational Modifications in the Large (LS) and Small (SS) Subunits of Ribulose-1,5 Bisphosphate Carboxylase/Oxygenase
DOE Contract:	DE-FG02-92ER20075
Principal Investigator:	Dr. Robert L. Houtz, Professor Department of Horticulture, Plant Physiology/Biochemistry/ Molecular Biology Program, University of Kentucky
Project Objectives:	1) Molecular and Biochemical Characterization of Rubisco LS and SS Methyltransferases (Rubisco LSMT/SSMT) 2) Definition of the Functional Significance of Lys-14 and Met-1 Methylation in the LS and SS respectively

This project is focused on a molecular and biochemical characterization of the protein methyltransferases responsible for methylation of the LS and SS in Rubisco, and the associated functional consequences accompanying these modifications. The identification of Rubisco Large Subunit Methyltransferases (Rubisco LSMT) as a SET domain protein methyltransferase (5), launched an intense scientific interest in this class of protein methyltransferases, especially as important determinants in the "histone code"(1, 2, 4). Our results have provided some of the most informative structural and mechanistic understandings of SET domain protein methyltransferases (5, 6). These results positioned us to provide the first unambiguous assignment of the kinetic reaction mechanism for SET-domain protein methyltransferases, and to design and engineer an alternative substrate for Rubisco LSMT, enabling substrate specificity and functional significance studies. The kinetic reaction mechanism results are in a final manuscript draft and will be submitted for publication shortly. The construction of the artificial substrate which mimics the N-terminal region of the LS of Rubisco, as well as results from preliminary studies are described below. These studies represent efforts to construct some highly specific and unique tools which we hope will provide information necessary to define functional and structural aspects of Lys-14 methylation in the LS of Rubisco. They have admittedly been time-consuming, but worth the efforts.

We recently demonstrated that the minimal substrate recognized by Rubisco LSMT is free lysine as well as monomethyllysine (6). This was corroborated both by structural analyses as well as enzymatic activity and subsequent product distribution analyses. Ternary complexes between Rubisco LSMT and free lysine compared to complexes with monomethyllysine demonstrate that the structural basis for multiple methyl group additions is a consequence of hydrogen-bond driven spatial shifts in the ε-amino group of Lys-14, which maintains the direct in-line geometry necessary for S_N2 nucleophilic attack. The structural observations are also consistent with the previous proposal that the multiplicity of methyl group additions takes place through a processive mechanism, with successive methyl group additions to an enzyme protein complex which does not disassociate prior to the formation of trimethyllysine. This mechanism has important implications, since the regulation of gene expression by SET domain histone methyltransferases is not only

dependent on site-specific lysine methylation, but also the degree of methylation (2). We examined the kinetic reaction mechanism for three different types of SET domain protein methyltransferases, each under conditions supporting mono-, di-, or trimethyllysine formation corroborated by product analyses. Additionally, the tight initial binding of Rubisco LSMT to Rubisco also allowed us to design a novel immobilized complex between Rubisco and Rubisco LSMT, which allowed for an unambiguous demonstration of the requirement for trimethyllysine formation prior to disassociation of the Rubisco LSMT:Rubisco complex, and therefore proof of the processive mechanism for methyl group transfer. Partial results from this study are shown in Figure 1. These kinetic studies also demonstrated that an important factor has been overlooked in all kinetic analyses of SET domain protein methyltransferases reported to date. This factor is the influence of the low turnover number for SET domain protein methyltransferases and how, relative to the time-frame of kinetic enzyme assays, this can generate changes in kinetic profiles shifting reciprocal plot patterns from random/ordered bi-bi to the real kinetic reaction mechanism plots of ping-pong. The results from these studies have just recently been submitted for publication.

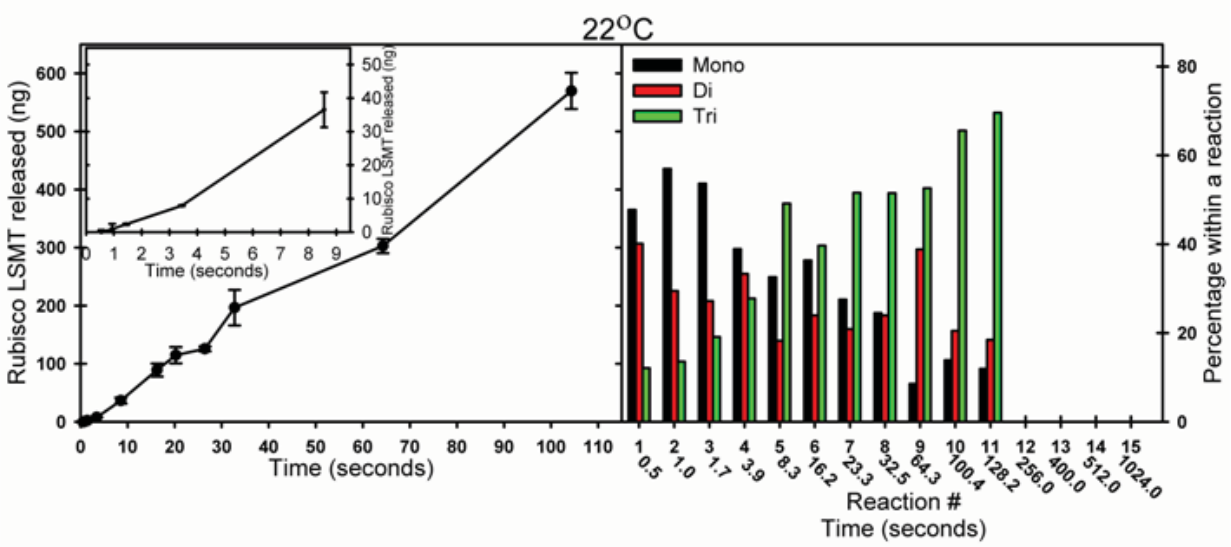

Figure 1. – *Release of Bound Rubisco LSMT from PVDF-Immobilized Rubisco as a Function of Trimethyllysine-14 Formation.* Ternary complexes between Rubisco LSMT and PVDF-immobilized Rubisco were prepared as previously described and immersed in buffered solutions containing 50 uM [^{3}H-*methyl*]AdoMet. The appearance of Rubisco LSMT in solution was quantified by enzyme assays (left panel), and the distribution of radiolabel into mono, di, and trimethyllysine (right panel) determined after complete hydrolysis of the PVDF-immobilized Rubisco with 6N HCl (110C, 22 h). Radiolabel distribution in the right panel is shown on a non-linear time scale to aid in the visual assessment of distribution between methylated lysine derivatives at short as well as longer times without using an inset figure as in the left panel.

Although the ternary complexes of Rubisco LSMT with S-Adenosylhomocysteine and lysine and monomethyllysine were informative in regard to reaction mechanism, they were not helpful in identifying the mechanism used by Rubisco LSMT for determining substrate specificity. We were unsuccessful at obtaining ternary complexes of Rubisco LSMT with bound synthetic polypeptide substrates, as has been reported for several histone methyltransferases. However, we

were able to model a polypeptide sequence corresponding to the N-terminal region of the LS of Rubisco into the apparent substrate binding cleft in Rubisco LSMT. Knowledge of the determinants of polypeptide substrate specificity are important for identifying possible alternate substrates, as well as the possibility of generating more desirable substrates amenable to site-directed mutagenesis experiments unlike Rubisco. We determined that Rubisco LSMT is capable of methylating synthetic polypeptide mimics of the N-terminal region of the LS, both free as well as conjugated to keyhole limpet hemacyanin, but with considerable less efficiency than intact holoenzyme. Nevertheless, utilizing synthetic polypeptide substrates in substrate specificity studies where a number of residues and combinations of residues will be altered is economically undesirable. Therefore we constructed an artificial substrate that could be used to map the determinants of polypeptide substrate specificity using site-directed mutagenesis, and act as a research tool to potentially identify the role of the N-terminus of the LS and methylation status as a determinant in binding and recruitment of other chloroplast proteins. Human carbonic anhydrase (hCA) is a protein which expresses well in bacterial cells and has convenient affinity-purification techniques independent of genetically engineered tags. The sequence coding for the first 25 amino acid residues from the N-terminus of the LS of Rubisco was fused to the N-terminus of a hCA cDNA. The fusion protein construct expressed at high levels in bacterial cells, and was easily purified without the use of affinity tags. The fusion protein exhibited the expected increase in molecular mass and was immunologically reactive with an antibody specific for the N-terminus of the LS. More importantly the fusion protein was a substrate for Rubisco LSMT (Figure 2).

Interestingly, kinetic analyses of the methylation of the LS-hCA fusion protein demonstrated that the K_m for the LS-CA fusion was approximately 1000-fold higher than the K_m for the Rubisco holoenzyme, while the turnover was relatively unaffected (.04 s^{-1} versus .03 s^{-1}). Thus, other major determinants for protein substrate binding reside outside the N-terminal region of the LS, a possibility which we had hypothesized earlier based on docking models (3, 6). The LS-hCA fusion protein is ideally suited for the determination of Rubisco LSMT substrate specificity since residues in the region of the Lys-14 methylation site are easily changed through primer design. We have already tested two constructs representing Lys-14/Arg-14, as a control given the location of other lysyl residues in the LS sequence as well as in hCA, and a double mutant Val-11/Ile-11;Val-17/Ala-17 representing a sequence found in RNA helicase DEAD-box proteins. The Lys-14/Arg-14 is completely inactive as a substrate demonstrating that only Lys-14 is subject to methylation. The double mutant, Val-11/Ile-11;Val-17/Ala-17 was active as a substrate albeit with only half the activity of the LS-hCA fusion construct. The kinetic characterization on this double mutant is not complete, but the observation of catalytic methylation demonstrates flexibility in the active site of Rubisco LSMT for small side-chain amino acid substitutions in the N-terminal region of Rubisco surrounding Lys-14. We are currently using the LS-hCA fusion protein to construct a substrate specificity map for Rubisco LSMT, as well as for obtaining information in regards to how these changes affect kinetic parameters and protein binding affinities.

Additionally, the information derived from these studies indirectly identifies potential alternative substrates for Rubisco LSMT, since BLAST searches on altered sequences around Lys-14 can identify other proteins which contain similar and/or identical sequences. The LS-hCA fusion construct can also be used as a tool to investigate the potential association of chloroplast proteins with the N-terminal region of the LS as a function of methylation status. We have not yet initiated these studies since we wanted information on the relative kinetic changes in the methylation of this artificial substrate first. These studies represent one of two efforts designed to provide information in regard to the functional consequences of Lys-14 methylation. There are many proteins that

contain chromodomains which contain specific binding sites for methylated lysyl residues. The demonstration of this as a primary function of lysyl methylation in the regulation of gene activity by histone specific SET-domain protein methyltransferases, suggest that this may be a reasonable possibility for the function of Lys-14 methylation in the N-terminus of the LS of Rubisco.

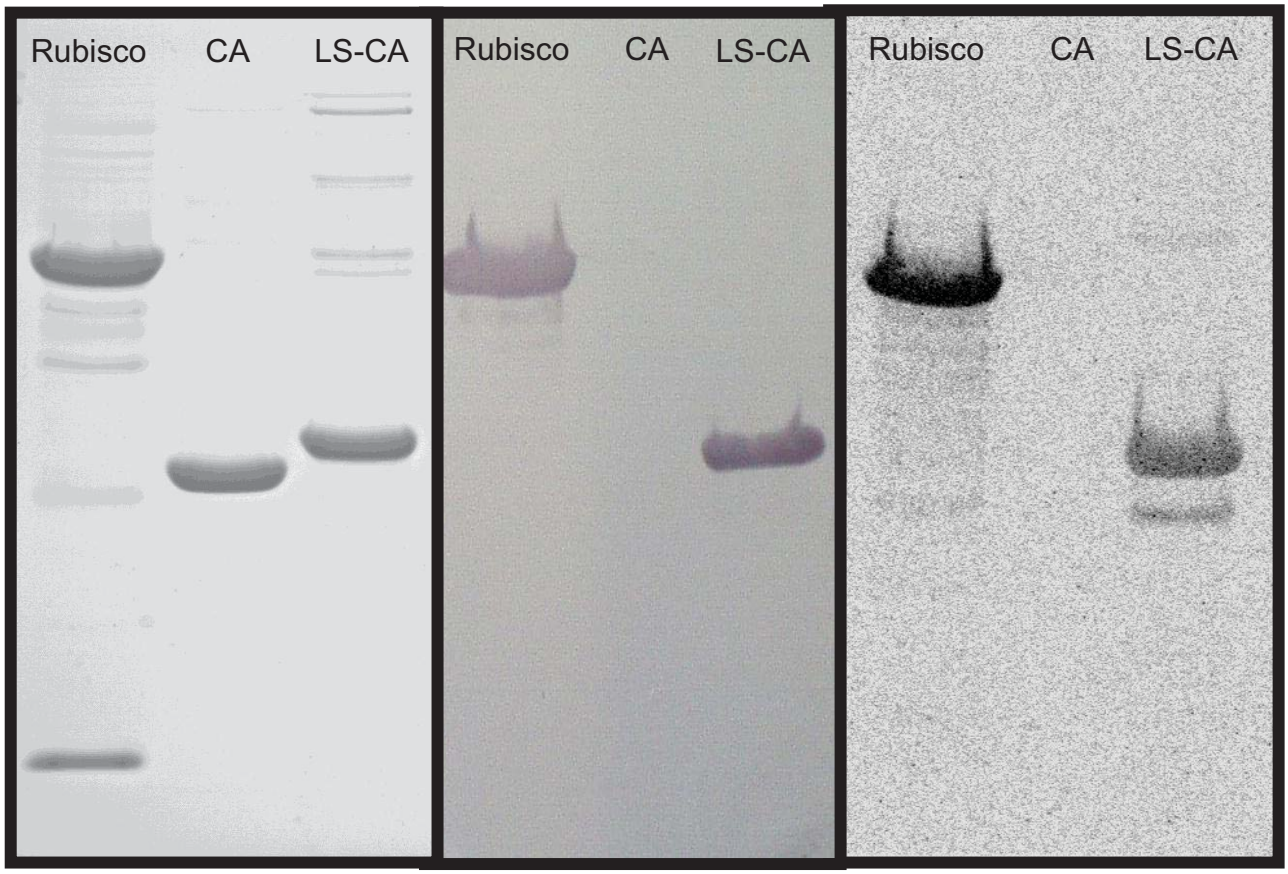

Figure 2. – *Evaluation of a fusion protein construct between the first 25 amino acid residues from the LS of Rubisco and hCA (LS-CA) as an alternative substrate for Rubisco LSMT.* Each panel contains three lanes with 10 ug each of polypeptides as identified above each lane. *Left Panel*, Coomasie-stained SDS-PAGE gel (12%); *Center Panel*, PVDF-membrane blot probed with polyclonal antibodies specific for the first 25 amino acids from the N-terminus of the LS of Rubisco; *Right Panel*, Phosphor-image of radio-label incorporation from [^{3}H-*methyl*]AdoMet catalyzed by Rubisco LSMT.

References

1. **de, l. C., X, S. Lois, S. Sanchez-Molina, and M. A. Martinez-Balbas**. 2005. Do protein motifs read the histone code? Bioessays 27:164-175.

2. **Dutnall, R. N.** 2003. Cracking the histone code: one, two, three methyls, you're out! Mol.Cell 12:3-4.

3. **Houtz, R. L. and A. R. Portis, Jr.** 2003. The life of ribulose 1,5-bisphosphate carboxylase/oxygenase--posttranslational facts and mysteries. Arch.Biochem.Biophys. 414:150-158.

4. **Jenuwein, T. and C. D. Allis**. 2001. Translating the histone code. Science 293:1074-1080.

5. **Trievel, R. C., B. M. Beach, L. M. Dirk, R. L. Houtz, and J. H. Hurley**. 2002. Structure and catalytic mechanism of a SET domain protein methyltransferase. Cell 111:91-103.

6. **Trievel, R. C., E. M. Flynn, R. L. Houtz, and J. H. Hurley**. 2003. Mechanism of multiple lysine methylation by the SET domain enzyme Rubisco LSMT. Nat.Struct.Biol. 10:545-552.

Publications during this project period acknowledging DOE support:

1.Magnani, R., Dirk, L. M., Trievel, R. C., and Houtz, R. L. (2010) Calmodulin methyltransferase is an evolutionarily conserved enzyme that trimethylates Lys-115 in calmodulin, Nature communications 1, 43.

2.Whitney, S. M., Houtz, R. L., and Alonso, H. (2011) Advancing our understanding and capacity to engineer nature's CO2-sequestering enzyme, Rubisco, Plant physiology 155, 27-35.

3.Del Rizzo, P. A., Couture, J. F., Dirk, L. M., Strunk, B. S., Roiko, M. S., Brunzelle, J. S., Houtz, R. L., and Trievel, R. C. (2010) SET7/9 catalytic mutants reveal the role of active site water molecules in lysine multiple methylation, The Journal of biological chemistry 285, 31849-31858.

4.Whitney, S. M., Kane, H. J., Houtz, R. L., and Sharwood, R. E. (2009) Rubisco oligomers composed of linked small and large subunits assemble in tobacco plastids and have higher affinities for CO2 and O2, Plant physiology 149, 1887-1895.

5.Raunser, S., Magnani, R., Huang, Z., Houtz, R. L., Trievel, R. C., Penczek, P. A., and Walz, T. (2009) Rubisco in complex with Rubisco large subunit methyltransferase, Proceedings of the National Academy of Sciences of the United States of America 106, 3160-3165.

6.Couture, J. F., Dirk, L. M., Brunzelle, J. S., Houtz, R. L., and Trievel, R. C. (2008) Structural origins for the product specificity of SET domain protein methyltransferases, Proceedings of the National Academy of Sciences of the United States of America 105, 20659-20664.

7.Dirk, L. M., Schmidt, J. J., Cai, Y., Barnes, J. C., Hanger, K. M., Nayak, N. R., Williams, M. A., Grossman, R. B., Houtz, R. L., and Rodgers, D. W. (2008) Insights into the substrate specificity of plant peptide deformylase, an essential enzyme with potential for the development of novel biotechnology applications in agriculture, The Biochemical journal 413, 417-427.

8.Houtz, R. L., Magnani, R., Nayak, N. R., and Dirk, L. M. (2008) Co- and post-translational modifications in Rubisco: unanswered questions, Journal of experimental botany 59, 1635-1645.

9.Dinkins, R. D., Majee, S. M., Nayak, N. R., Martin, D., Xu, Q., Belcastro, M. P., Houtz, R. L., Beach, C. M., and Downie, A. B. (2008) Changing transcriptional initiation sites and alternative 5'- and 3'-splice site selection of the first intron deploys Arabidopsis protein isoaspartyl methyltransferase2 variants to different subcellular compartments, The Plant journal : for cell and molecular biology 55, 1-13.

10.Magnani, R., Nayak, N. R., Mazarei, M., Dirk, L. M., and Houtz, R. L. (2007) Polypeptide substrate specificity of PsLSMT. A set domain protein methyltransferase, The Journal of biological chemistry 282, 27857-27864.

11.Dirk, L. M., Flynn, E. M., Dietzel, K., Couture, J. F., Trievel, R. C., and Houtz, R. L. (2007) Kinetic manifestation of processivity during multiple methylations catalyzed by SET domain protein methyltransferases, Biochemistry 46, 3905-3915.

12.Hou, C. X., Dirk, L. M., Pattanaik, S., Das, N. C., Maiti, I. B., Houtz, R. L., and Williams, M. A. (2007) Plant peptide deformylase: a novel selectable marker and herbicide target based on essential cotranslational chloroplast protein processing, Plant biotechnology journal 5, 275-281.

13.Lynnette M.A. Dirk, Raymond C. Trievel, and Robert L. Houtz. 2006. The Enzymes, Vol. 24 Protein methyltransferases, Steven G. Clarke and Fuyuhiko Tamanoi, eds. Chapter 8, Non-Histone Protein Lysine Methyltransferases - structure and catalytic roles.

14.Shepherd, R. W., Bass, W. T., Houtz, R. L., and Wagner, G. J. (2005) Phylloplanins of tobacco are defensive proteins deployed on aerial surfaces by short glandular trichomes, The Plant cell 17, 1851-1861.

15.Trievel, R. C., Flynn, E. M., Houtz, R. L., and Hurley, J. H. (2003) Mechanism of multiple lysine methylation by the SET domain enzyme Rubisco LSMT, Nature structural biology 10, 545-552.

16.Houtz, R. L., and Portis, A. R., Jr. (2003) The life of ribulose 1,5-bisphosphate carboxylase/oxygenase--posttranslational facts and mysteries, Archives of biochemistry and biophysics 414, 150-158.

Patent Applications submitted during this project report period acknowledging DOE support:

U.S. Provisional Application Serial No. 60/468,966 - A Modified Rubisco Large Subunit εN-methyltransferase Useful for Targeting Molecules to the Active-site Vicinity of Ribulose-1, 5-Bisphosphate - Inventor **Robert L. Houtz** - Filed May 7, 2004.