Hierarchical assembly of the MLL1 core complex within a biomolecular condensate regulates H3K4 methylation

Kevin E.W. Namitz[1], Song Tan[2], Michael S. Cosgrove[1]*

[1]SUNY Upstate Medical University, Department of Biochemistry and Molecular Biology, Syracuse, New York.
[2] Penn State University, Department of Biochemistry and Molecular Biology, University Park, Pennsylvania

*To whom correspondence should be addressed: Michael S. Cosgrove, Ph.D., Department of Biochemistry and Molecular Biology, SUNY Upstate Medical University, 750 East Adams Street, Syracuse, NY 13210, Phone: (315) 464-7751, E-mail: cosgrovm@upstate.edu

ABSTRACT

The enzymes that regulate histone H3 lysine 4 (H3K4) methylation are required for cellular differentiation and development and are often mutated in human disease. Mixed Lineage Leukemia protein-1 (MLL1) is a member of the SET1 family of histone H3 lysine 4 methyltransferases, which require interaction with a conserved sub-complex consisting of WDR5, RbBP5, Ash2L and DPY30 ($WRAD_2$) for maximal activity. It is currently unclear how assembly of SET1 family complexes is involved in the spatiotemporal control of H3K4 methylation in eukaryotic genomes. In this investigation, we systematically characterized the hydrodynamic and kinetic properties of a reconstituted human MLL1 core complex and found that its assembly is highly concentration and temperature dependent. Consistent with a hierarchical assembly pathway, we found that the holo-complex assembles through interactions between the MW and RAD_2 sub-complexes, which is correlated with enzymatic activity. Surprisingly, we found that the disassembled state is favored at physiological temperatures, and that this thermodynamic barrier can be overcome under conditions that induce high-local concentrations of subunits in phase separated compartments. Combining this data with the observation that MLL1 primary sequence contains large regions of intrinsic disorder, we propose a “swinging-domain” model in which the interaction between a tethered MW subcomplex and multiple nucleosome-RAD_2 complexes is regulated by the rapid formation or dissolution of biomolecular condensates, such as occurs in transcription factories. This model provides an elegant “switch-like” mechanism for spatiotemporal control of H3K4 methylation within eukaryotic genomes.

INTRODUCTION

Cellular identity in multicellular organisms is maintained in part by enzymes that regulate the degree of histone H3 lysine 4 (H3K4) methylation (1). Di- and trimethylation of H3K4 (H3K4me2,3) are enriched in gene bodies and promoters of active genes (2-4) respectively, and function to recruit nucleosome-remodeling complexes that regulate transcription (5-9). H3K4 monomethylation (H3K4me1) is associated with active gene enhancers (10-12), but is also associated with gene silencing (13-17). Because genome-wide alterations in the patterns of H3K4 methylation are linked to the aberrant transcriptional programs in developmental disorders and cancers (18-28), there is significant interest in understanding how different H3K4 methylation states are established and maintained.

Mixed Lineage Leukemia protein-1 (MLL1, ALL1, HRX, KMT2C) is a member of the SET1 family of H3K4 methyltransferases and is frequently altered in poor prognosis acute leukemias (29). MLL1 is a large protein with 3,969 amino acids and assembles into a super-complex with ~30 subunits (30-33). Subunits shared among all SET1 family members include WDR5, RbBP5, Ash2L, and two copies of DPY-30 ($WRAD_2$), which associate into a sub-complex that interacts with the C-terminal SuVar, Ez, Trx (SET) domain of MLL1 (34-38). *In vitro* studies have shown that the MLL1 SET domain predominantly catalyzes H3K4 monomethylation (36), whereas multiple methylation depends on interaction of MLL1 with $WRAD_2$, forming what is known as the MLL1 core complex (also known as human COMPASS, or $MWRAD_2$) (34,36,39). The requirement of full $MWRAD_2$ complex for optimal enzymatic activity suggests that H3K4 methylation may be regulated at the level of subunit assembly in the cell. Consistent with this hypothesis, genome-wide studies show that, while MLL1 localizes to thousands of genes in mammalian genomes, multiple methylation of H3K4 is mainly correlated with the subset of genes where MLL1 co-localizes with $WRAD_2$ subunits (40). In addition, disease-specific missense mutations have been shown to disrupt MLL family core complexes (41), suggesting that aberrations in complex assembly may be associated with human disease. More recently, several labs have shown that perturbation of MLL1 core complex assembly with protein-protein interaction inhibitors may have utility as a novel therapeutic approach for treating malignancies (42-44). Together, these results suggest that knowledge of the molecular mechanisms controlling MLL1 core complex assembly will be crucial for understanding of how different H3K4 methylation states are regulated in mammalian genomes. However, progress has been impeded by the lack of understanding of the biophysical and thermodynamic mechanisms that underlie MLL1 core complex assembly.

Biochemical reconstitution studies using a minimal MLL1 SET domain construct show that the stoichiometry of the MLL1 core complex consists of one copy of the MLL1, WDR5, RbBP5 and Ash2L subunits, and 2 copies of the DPY-30 subunit ($MWRAD_2$)- forming a complex with a mass of ~205 kDa (36). Direct interactions have been observed between MLL1 and WDR5 (35,37,45), WDR5 and RbBP5 (46,47), RbBP5 and Ash2L (36), and Ash2L and DPY30 (36,48,49). While these pairwise interactions suggest a linear arrangement of subunits, several lines of evidence indicate a more intricate quaternary structure. For example, while MLL1 does not interact with RbBP5 or Ash2L in pairwise experiments (36), an investigation of SET domain-associated Kabuki syndrome missense mutations suggests a direct interaction with the RbBP5/Ash2L heterodimer within the context of the holo-complex (41). The WDR5 subunit functions to stabilize this interaction by directly binding to the MLL1 WDR5 interaction (Win) motif (35,37,45) and RbBP5 (34,36). Binding experiments show that the weakest pairwise interaction occurs between the WDR5 and RbBP5 subunits (36), suggesting the complex may be hierarchically assembled. All of these interactions have been confirmed in recent Cryo-EM and X-ray crystal structures of related SET1 family complexes (50-53). Together, these results suggest that complex assembly is hierarchical in nature, with the requirement for the formation of distinct sub-complexes before assembly of the higher-order quaternary structure. The choreographic details of this assembly pathway are unknown.

In this investigation, we systematically characterized the hydrodynamic and kinetic properties of a reconstituted human MLL1 core complex under a variety of conditions. We found that MLL1 core complex assembly is highly concentration and temperature dependent. Consistent with the hypothesized hierarchical assembly pathway, we found that the holo-complex assembles through interactions between the MW and RAD_2 sub-complexes, and that $MWRAD_2$ formation is directly correlated with enzymatic activity. Surprisingly, we found that the disassembled state is favored at physiological temperatures and at concentrations typically used in steady-state enzymatic assays. In contrast, sub-physiological ionic strength dramatically increases enzymatic activity, which is associated with the formation of induced high-local concentrations of the MLL1 core complex in phase-separated droplets. Based on these results, we propose a model in which the thermodynamic barrier to complex assembly is overcome in the cell under conditions that induce high-local concentrations of subunits, such as those found in transcription factories. Together, these results are consistent with the hypothesis that regulated assembly of the MLL1 core complex underlies an important mechanism for establishing different H3K4 methylation states in mammalian genomes.

RESULTS

MLL1 core complex assembly is concentration and temperature dependent

To better understand MLL1 core complex assembly, we purified human recombinant $MWRAD_2$ as described in Methods and characterized its oligomeric behavior by size exclusion chromatography (SEC) and sedimentation velocity analytical ultracentrifugation (SV-AUC). SEC revealed that the purified complex eluted as a single symmetrical peak (Fig. 1A) and SDS-PAGE of the indicated fractions showed the presence of all five subunits with the expected stoichiometry (Fig. 1B). We note that the complex elutes later than expected based on its theoretical mass, which is likely due to the significant shape asymmetry of the particle. We then chose SV-AUC to characterize the concentration and temperature dependence of the complex in solution. SV-AUC is a first-principle technique that measures the time course of sedimentation of macromolecules in a gravitational field in a way that maintains the equilibrium of reversible associations – allowing extraction of equilibrium and kinetic properties of interactions (54,55). Sedimentation boundaries formed as the particle sediments over time were fit using a finite element analysis of Lamm equation solutions (Fig. 1C) (56) to give the diffusion-free sedimentation coefficient distribution *c(s)* (Fig. 1D). The *c(s)* plot of $MWRAD_2$ at 5 μM loading concentration at 5°C revealed a large peak accounting for almost 90% of the signal with an $s_{20,w}$ (*S*) value of 7.2 and two minor peaks at 2.9 and 4.7 *S* that each account for 4-5% of the signal (noted with arrows in Fig. 1D). The major peak at 7.2 *S* corresponds to the fully assembled MLL1 core complex, which we previously showed assembles with a stoichiometry of 1:1:1:1:2 for the $MWRAD_2$ subunits, respectively (36). In addition, the *S*-value of $MWRAD_2$ is independent of loading concentration (Fig. 2A), indicating that the complex is stable at 5°C and has a relatively long lifetime compared to the timescale of sedimentation (57). Using the derived weight-averaged frictional coefficient (f/f_0) of 1.7, the calculated molecular mass from this *S* value was 209,561 Daltons, which is within error of the expected mass (205,402) based on the amino acid sequence of the holo-complex subunits at the indicated stoichiometry.

The minor peaks observed in the *c(s)* distribution in Fig. 1D could represent trace contaminants in the sample or minor populations of dissociated sub-complexes and/or subunits. To distinguish these hypotheses, we compared *c(s)* distributions of $MWRAD_2$ at concentrations ranging from 0.25 – 5 μM at 5°C (Fig. 2A) and 30°C (Fig. 2B). If the minor peaks represent non-interacting contaminants, then the relative amount of signal between the major and minor peaks will not vary as the loading concentration is decreased. In contrast, if the complex is dissociating into sub-complexes, then the relative amount of signal in the major and minor peaks will change as the loading concentration is varied. The results were consistent with the latter possibility. For

example, while the effect at 5°C was modest, when the loading concentration of the complex was decreased from 5 μM to 0.25 μM, the amount of signal corresponding to the holo-complex decreased from ~88% to ~83% of the total signal, with a corresponding increase in both minor peak signals (Fig. 2A). The effect was more obvious at 30°C, which showed that the signal corresponding to the minor peaks increased from 35% to 75% of the total signal as the loading concentration was decreased (Fig. 2B). These results suggest that the minor peaks represent dissociated sub-complexes and/or subunits. Furthermore, because the *S*-values of the minor peaks show varying degrees of concentration dependence, they likely represent reaction boundaries of sub-complexes as opposed to individual non-interacting subunits. These data suggest that the holo-complex assembles from predominantly two sub-complexes in a temperature and concentration-dependent manner.

The disassembled state of the MLL1 core complex is favored at physiological temperature

To further explore the thermodynamics of MLL1 core complex assembly, we compared the temperature dependence of $MWRAD_2$ formation at several different loading concentrations using SV-AUC (Fig. 3). Each *c(s)* profile was integrated and the relative amount of signal corresponding to the *S* value of the holo-complex was plotted as a function of temperature and total loading concentration (Fig. 3F). At the highest loading concentration (5 μM), little variation in the amount of holo-complex was observed between 5° and 25°C (Fig. 3A, F), with a peak that accounted for 81-92% of the total signal (Table S1). In contrast, at temperatures greater than 25°C, the amount of holo-complex decreased precipitously until only ~3% of the signal could be observed at 37°C (Figs. 3A and F, Table S1). The effect of temperature on MLL1 core complex stability became increasingly more severe as the loading concentration was decreased. For example, at the lowest loading concentration (0.25 μM), only the 5°C and 10°C runs showed ~80% holo-complex (Fig. 3E, F; Table S1); whereas at higher temperatures, the signal corresponding to the holo-complex decreased from ~63% at 15°C - to ~2% of the total signal at 37°C (Fig. 3F; Table S1). At 37°C, most of the signal is instead dominated by the two sub-complex peaks with *S*-values of ~3 and 4.7 (Fig. 3G). These data are consistent with the hypothesis that the holo-MLL1 core complex assembles from interaction of two sub-complexes, the equilibrium of which is highly concentration and temperature-dependent.

Surprisingly, at all the concentrations tested, very little holo-complex with an *S* value of 7.2 was observed at physiological temperature (37°C) (Fig. 3G). This suggests that the

disassembled state of the MLL1 core complex may predominate in cells, and that other factors are required to stabilize the assembled state. In support of this hypothesis, closer examination of the *c(s)* profiles of the complex at 37°C revealed evidence that increased protein concentration promotes complex formation. For example, while similar amounts of signal are observed in the two sub-complex peaks at the 0.25 μM loading concentration (cyan line, Fig. 3G), the relative amount of signal in the two peaks changes with progressively higher concentrations. The intensity of the larger peak increased at the expense of the smaller peak and began to show evidence of concentration-dependent shifting to higher *S*-values. This hydrodynamic behavior is consistent with a reaction boundary composed of free and bound reactants that interconvert under a rapid kinetic regime that cannot be resolved within the signal-to-noise of the experiment (57). These results suggest that, unlike the long lifetime of the assembled complex observed at 5°C, the kinetics of the interaction have changed at 37°C such that the complex now has a short lifetime compared to the timescale of sedimentation.

We next analyzed the concentration series at each temperature to derive binding isotherms. We integrated each *c(s)* profile (between 0.5 and 9.5 *S*) to determine the weight-average sedimentation coefficients (s_w) (58), which were then plotted against $MWRAD_2$ concentration and fit to derive the apparent dissociation constant (K_d^{app}) for each isotherm (Fig. 4A). Given that the majority of signal in each *c(s)* profile could be assigned to three peaks, we applied the A + B ⇆ AB hetero-association model in the program SEDPHAT (59) and obtained reasonable fits (Table 1). The derived K_d^{app} values ranged from 7 nM at 5°C to ~6,200 nM at 37°C (Table 1). A van't Hoff analysis showed that complex formation is exothermic, which is offset by the negative entropy change as the complex subunits become more ordered (Fig. 4B and C). However, the van't Hoff plot reveals a non-linear relationship between K_{eq} and temperature, indicating a change in the heat capacity of the system at higher temperatures (Fig. 4B). These data suggest at least two mechanisms for complex assembly, which differ by temperature. At low temperatures (≤ 25°C), the equilibrium favors complex formation, with a relatively long lifetime that is stable on the timescale of sedimentation. Under this mechanism, the interaction is dominated by enthalpic contributions to the free energy (Fig. 4C). At high temperatures (> 25°C), the equilibrium is shifted into the rapid kinetic regime with a short complex lifetime where dissociation is more likely. While there is little difference in the Gibbs free energy between mechanisms, there is a difference in the contributions between the enthalpic and entropic terms. At higher temperatures, the entropic penalty to complex formation was increased 7-fold compared to that of the lower temperature mechanism, while the

difference in the enthalpic contribution was only increased by 3.8-fold (Fig. 4C). These results suggest that, at physiological temperature, one or more of the subunits samples alternate conformational states, some of which are not competent for complex assembly. However, given the observation that some holo-complex forms in a concentration-dependent manner, increased local concentration of subunits may be a mechanism that cells use to overcome the increased entropic cost of complex formation at 37°C.

The MLL1 core complex assembles from MW and RAD$_2$ sub-complexes

Previous experiments suggested that the holo-complex is assembled by pairwise interactions as follows: M ⇌ W ⇌ R ⇌ A ⇌ D_2 (36). Since the weakest pairwise interaction occurs between WDR5 and RbBP5 (36), we predicted that the complex assembles by first forming MW and RAD_2 sub-complexes, which then interact to form the holo-complex (Scheme 1). However, we reasoned that there are at least two additional reaction schemes that could potentially give rise to the three boundaries observed in the holo-complex *c(s)* profiles (Schemes 2 and 3). To distinguish among these schemes, we chose to use a Bayesian approach to analyze the SV-AUC data of the holo-complex collected at 25°C. The Bayesian approach is a variant of the standard maximum entropy regularization method utilized in the *c(s)* analysis in that, instead of assuming a uniform probability for the occurrence of species at every *S*-value in a distribution, it utilizes prior information to assign different probabilities in different regions of *S*-values (60). A key feature of the Bayesian implementation in SEDFIT is that, because it maintains the same degrees of freedom used in the standard *c(s)* analysis, imperfections in the expected values will result in additional features in the $c^{(p)}(s)$ plots in order to maintain the quality of the fit (60). The Bayesian analysis therefore allows us to determine which reaction scheme gives a $c^{(p)}(s)$ profile that best fits the experimental data.

$$Scheme\ 1: \quad M + W \rightleftharpoons MW, \qquad R + AD_2 \rightleftharpoons RAD_2, \qquad MW + RAD_2 \rightleftharpoons MWRAD_2$$

$$Scheme\ 2: \quad M + W \rightleftharpoons MW + R \rightleftharpoons MWR + AD_2 \rightleftharpoons MWRAD_2$$

$$Scheme\ 3: \quad W + R \rightleftharpoons WR + AD_2 \rightleftharpoons WRAD_2 + M \rightleftharpoons MWRAD_2$$

To obtain the expected *S*-values for each of the predicted sub-complexes or subunits in each reaction scheme, we mixed stoichiometric amounts of their respective subunits and characterized their concentration dependence by SV-AUC at 25°C (Fig. S1; Table S2). We then

used each of the *S*-values collected at 0.25 μM as prior expectations in the Bayesian analysis of the holo-complex. As shown in Fig. 5A, when the independently determined *S*-values for MW, RAD_2 and the $MWRAD_2$ species were used as prior expectations in the Bayesian analysis of the holo-complex at 0.25 μM (black dotted line), three peaks in the $c^{(p)}(s)$ plot were observed that were in excellent agreement with the expectations (cyan line). Indeed, good agreement was observed using the same *S*-values as prior expectations for Bayesian fits of the experimental data collected at higher holo-complex concentrations (Fig. 5A). The only deviation observed was for the position and amplitude of the holo-complex peak, which at 25°C shifts from 6.8 to 7.2 *S* in a concentration-dependent manner (Fig. 5A). In contrast, when a similar analysis was conducted instead using the expected *S*-values for the MWR and AD_2 sub-complexes predicted by Scheme 2, additional features in the $c^{(p)}(s)$ plot with an *S*-value of ~5.3 were observed at all loading concentrations that did not match the prior expectations (Fig. 5B, red arrow). Similarly, using the expected *S*-values for M and $WRAD_2$ as predicted by Scheme 3, the $c^{(p)}(s)$ plot showed little evidence of a species matching the expected value of free MLL1 at 2.3 *S*, and also showed additional features at ~3.5 *S* that did not match expectations (Fig. 5C, red arrow). To test whether the holo-complex assembles in a concerted fashion from individual subunits, we also performed a similar Bayesian analysis using the predetermined *S* values for M, W, R, AD_2, and $MWRAD_2$ as prior expectations (AD_2 is treated as a discrete species since it does not appreciably dissociate under the range of concentrations that can be detected by the absorbance optical system used in these experiments (36)). The $c^{(p)}(s)$ plot showed additional features with an *S*-value of ~5.2 that did not match expectations (Fig. 5D, red arrow). Together, these results are consistent with the hypothesis that MLL1 core complex is hierarchically assembled by association of MW and RAD_2 sub-complexes.

Enzymatic activity of the MLL1 core complex is directly related to complex assembly

To determine the impact of concentration and temperature on the enzymatic activity of the MLL1 core complex, we incubated $MWRAD_2$ (0.25 – 5 μM) with a fixed concentration of histone H3 peptide (10 μM) and saturating amounts of AdoMet (250 μM) at various temperatures. We then measured methylation using a label-free quantitative MALDI-TOF mass spectrometry assay (36). MALDI spectra were integrated and the relative amount of each peptide species was plotted as a function of time. Data were fit using a numerical integration of rate equations approach implemented in KinTek Explorer software (61), which allowed us to test the ability of different reaction schemes to fit the data.

Using the simplest irreversible consecutive reactions model (Fig. 6, Scheme 4), while acceptable fits were obtained for reaction progress curves collected at the highest concentration (5 μM) between temperatures 5 – 30°C (5°C is shown in Fig. 6A), the rest of the fits were poor (an example is shown in Fig. 6B). Since we previously showed that the complex uses a non-processive mechanism for multiple lysine methylation (36), we revised the model to incorporate binding of peptide substrate to the enzyme-AdoMet complex (E_1) and release of the H3K4me1 product after the first methylation event, followed by binding of the H3K4me1 substrate to a distinct site on the enzyme (E_2) for the dimethylation reaction. The latter step is predicated on our previous observation that the MLL1 core complex has a cryptic second active-site independent of the SET domain that is required for the H3K4 dimethylation reaction (36,62,63). Since the binding and release rates of substrates and product are currently unknown, these values were fixed to be non-rate limiting. This model allowed us to incorporate an additional term to test the impact of reversible complex disassembly, which results in negligible activity of both enzymes under these assay conditions (Fig. 6, Scheme 5) (36,37). Initial values for the ratio (k_{off}/k_{on}) for complex assembly were set to be equal to the K_d^{app} derived from each SV-AUC isotherm experiment.

The resulting simulations showed that adding a reversible complex disassembly step to the reaction scheme only modestly improved fits to the lower temperature data (Fig. 6C), but did not improve the fits of the higher temperature data (Fig. 6D). In addition, Fitspace confidence contour analysis (64) showed that the derived k_{off} value for the complex dissociation step was not constrained by the data (not shown), suggesting that the model is more complex. Closer examination of the high temperature data showed that several reactions failed to go to completion, suggesting the enzyme rapidly inactivates at higher temperatures. We therefore revised the working model to incorporate an irreversible enzyme inactivation step (k_{inact}) (Figure 6, Scheme 6). The resulting simulations resulted in good fits to both the low and high temperature datasets shown in Figs. 6E and 6F, respectively. In addition, Fitspace analysis showed that the derived pseudo-first order rate constants for monomethylation (k_{me1}), and dimethylation (k_{me2}) reactions were reasonably well-constrained by the data (Fig. 6G and H). Furthermore, the rate of enzyme inactivation (k_{inact}) was constrained by the data in the higher temperature experiments (Fig. 6H) but not in the lower temperature experiments (Fig. 6G), where enzyme inactivation is negligible. Figure 7 shows that the use of Scheme 6 produces good fits for all datasets.

Based on these results, we then used the fits to Scheme 6 to compare the impacts of temperature and concentration on the enzymatic activity of the MLL1 core complex (Fig. 8). The

obtained pseudo-first order rate constants for monomethylation (k_{me1}), dimethylation (k_{me2}) and the rate of enzyme inactivation (k_{inact}) are summarized in Tables 2-4, respectively. At most of the tested enzyme concentrations, activity increased linearly as the temperature increased from 5°C to 20°C (Fig. 8A and C). However, above 20°C, non-Arrhenius behavior was observed, as the rate of irreversible enzyme inactivation (k_{inact}) rivaled or exceeded the rates of turnover (Tables 2-4), resulting in reactions that failed to go to completion (Fig. 7). These results are consistent with the conclusions from the SV-AUC analysis, which suggested that as the complex dissociates at higher temperatures, one or more of the subunits undergoes an irreversible conformational change that is not competent for catalysis. We therefore plotted k_{me1} and k_{me2} rates (Ln(k_n)) as a function of temperature (1/T) between 5°C and 20°C to fit the data to the Arrhenius equation (Fig. 8 B and D, respectively). Linear fitting of the Arrhenius plots revealed similar values for the energy of activation (E_a) between the tested concentrations. The average E_a values were 10.9 ± 2.0 kcal K^{-1} mol^{-1} and 17.8 ± 4.7 kcal K^{-1} mol^{-1} for the monomethylation and dimethylation reactions, respectively.

The minimum enzyme concentration resulting in complete conversion into the mono- and then dimethylated forms was 1.0 μM at 15°C (Fig. 7). Slightly higher activity was observed at the same enzyme concentration at 20°C, but with evidence of significant enzyme inactivation resulting in failure to go to completion. Increased concentration extended the range of temperatures under which complete conversion could be observed. For example, at 5 μM enzyme concentration, complete conversion of the peptide into the dimethylated form was observed between 5°C and 30°C, with evidence of modest H3K4 trimethylation activity (7% - 15%) between 10°C and 25°C (Fig. 7). However, at 37°C, only ~25% of the peptide was converted into the dimethylated form before the enzyme was completely inactivated.

In general, the apparent pseudo first-order rate constants for mono- and dimethylation were correlated with the amount of holo-complex in the assay at every temperature between 5°C and 30°C, with Pearson correlation coefficients (r) ranging from 0.57-0.94 for monomethylation, and 0.46-0.86 for dimethylation. At 37°C, the correlation was less obvious due to the lack of detectable activity at the lowest concentrations. (r = 0.17 and 0.29 for mono and dimethylation, respectively). In contrast, the parameter that was most highly correlated with the amount of holo-complex present in the assay at all temperatures was the rate of irreversible enzyme inactivation (k_{inact}), with Pearson r values ranging between -0.84 and -0.99, depending on the concentration tested. These results are consistent with the conclusions from the s_w isotherm analysis, in that holo-complex formation prevents individual subunits from sampling

potential non-productive folding intermediates, some of which lead to irreversible enzyme inactivation. These results also raise questions about how cells manage to prevent loss of enzymatic activity at physiological temperatures.

Induced high local concentration within a biomolecular condensate alters the assembly and enzymatic activity of the MLL1 core complex

Both hydrodynamic and enzymatic assays suggested that higher local concentrations of subunits would promote complex formation and enzymatic activity at physiological temperatures. However, given the low concentration of MLL1 in cells (which has been estimated to be femtomoles per mg of nuclear extract (65)), it is likely other factors are required to promote complex assembly. MLL1 has been shown to localize in discrete puncta in mammalian cell nuclei (66), raising the possibility that it could be regulated by induced high local concentration in liquid-liquid phase-separated (LLPS) particles, such as those found in transcription factories (67,68). Liquid-liquid phase separation has been shown to increase local protein concentration of proteins and ligands by up to 10,000-fold (69). Common features of proteins that undergo phase separation include primary sequences with regions of low complexity, or intrinsically disordered regions, that provide the numerous transient multivalent interactions required for liquid-liquid de-mixing (70). Indeed, examination of the primary sequence of MLL1 by IUPred (71) reveals that the majority of its sequence is predicted to be intrinsically disordered (Fig. 9A). In addition, the MLL1 construct used in this investigation and each $WRAD_2$ subunit shows significant regions of predicted disorder (Fig. S2). To determine if the catalytic module of the MLL1 core complex may also be regulated by phase separation, we examined $MWRAD_2$ using differential interference contrast (DIC) microscopy at concentrations up to 75 mg/ml but observed no evidence for phase separation (not shown). However, since a previous investigation showed increased enzymatic activity of the MLL1 core complex with reduced ionic strength (72), we tested whether reduced ionic strength may also regulate the LLPS properties of the MLL1 core complex.

First, we compared MLL1 core complex activity at several different ionic strengths at 25°C using quantitative MALDI-TOF mass spectrometry. Consistent with the previous report, we found that the enzymatic activity was significantly increased in buffers with sub-physiological ionic strength (Fig. 9B, D). While there was relatively little difference in mono- or dimethylation activities between 200-100 mM NaCl, mono- and dimethylation activity was increased 15- and 12-fold, respectively, when the NaCl concentration was reduced from 100 to 25 mM (Fig. 9B, D). To better understand the reason for increased enzymatic activity at lower ionic strength, we

compared the hydrodynamic properties of the 5 μM $MWRAD_2$ complex at 100 mM and 25 mM NaCl at 25°C using SV-AUC (Fig. 9C). Strikingly, comparison of *c(s)* profiles showed hydrodynamic changes in the complex that resembled those of HP1α that was induced to undergo phase separation (73). The relatively monodisperse peak of the MLL1 core complex at physiological ionic strength (Fig. 9C, purple line) becomes more polydisperse when ionic strength is reduced (Fig 9C, blue line), with peaks at 8.1 *S* (54%), 10.0 *S* (21%) and 12.3 *S* (~10%), along with several higher molecular weight species that collectively account for ~16% of the total signal. We also noticed that the relative distribution among these species shifts to larger *S*-values in a concentration-dependent manner, which is more pronounced with even lower ionic strength (Fig. S3). These results suggest that the increased activity of the MLL1 core complex with lower ionic strength is associated with hydrodynamic alterations of the complex that could include conformational alterations, oligomerization, aggregation, and/or phase separation.

Because the standard *c(s)* analysis uses a single weight-average frictional coefficient of all particles to fit the data (74), the polydispersity of the sample at low ionic strength shown in Fig. 9C prevents accurate molecular weight estimates of each species – and thus our ability to distinguish among the different hypotheses. We therefore performed a two-dimensional size and shape distribution analysis *($c(s,f_r)$)* of the SV-AUC data, which allows estimation of the frictional coefficients and average molar masses of each species in a complex distribution (75). The *$c(s,f_r)$* distribution of $MWRAD_2$ at ~100mM NaCl showed a single peak with the typical experimental *s**-value of the complex, but encompassing a fairly broad range of frictional ratios between 1.0 and 3.0, with a weight average frictional coefficient of ~1.5 (Fig. 9E). The estimated average molecular mass using this frictional coefficient and *S*-value was ~190 kDa, which is in fairly good agreement with the theoretical mass of the monomeric complex (205 kDa). In contrast, in low ionic strength buffer, the *$c(s,f_r)$* distribution showed that the majority of the signal is divided among several peaks with larger *S*-values that ranged between 9 and 16, with evidence of several larger molecular weight species ranging between 20-70 *S* (Fig. 9F). Several of the peaks between 9 and 13 *S* had frictional ratios that range between 1.1-1.2, which gave mass estimates between 140-230 kDa. Because these species have relatively similar molar mass estimates, these *S*-values likely correspond to species with increasingly compact conformations of the monomeric MLL1 core complex. The peak at ~16 *S* gives a mass estimate of ~350 KDa, which is indicative of a reaction boundary between monomeric and dimeric complexes. These results suggest that lower ionic strength allows the complex to sample different conformational states, some of which are more compact, and some that allow

oligomerization of the MLL1 core complex. Consistent with this interpretation, these larger *S*-value species become increasingly more populated in an $MWRAD_2$ concentration-dependent manner (Fig, S3).

The *c(s,f$_r$)* analysis also showed several discrete species with *S*-values between 20-70 *S* with a broad range of frictional ratios ranging between 3-5 (Fig. 9F). Integration of these peaks gave mass estimates starting at ~3.7 MDa, which approximates an 18-mer of $MWRAD_2$, with each discrete species at higher *S*-values approximating the addition of one $MWRAD_2$ dimer. This hydrodynamic behavior is indicative of fiber-like material (76) and could reflect various sizes of insoluble aggregates, or the fiber-like polymerization that is predicted to precede the formation of phase separated droplets (Fig. 9G) (70,77). To distinguish these hypotheses, we examined enzymatic reaction mixtures at 100mM or 25 mM NaCl using DIC microscopy. Surprisingly, despite using a relatively low concentration of enzyme (5 μM), the low ionic strength reaction mixture showed evidence of spherical LLPS droplets (Fig. 10B) that were absent in the 100mM NaCl reaction mixture (Fig. 10A). No visible evidence of protein precipitation was observed. The droplets were small and mobile, but did not appear to fuse, which is a common feature of particles induced to undergo LLPS (77). However, addition of a crowding agent (dextran; 7% w/v) to the reaction mixture resulted in LLPS droplets with larger diameters and observable fusion events that could be detected by DIC microscopy (Fig. 10C and movie S1). Importantly, the droplets disappeared in the presence of 5% 1,6 hexanediol (Fig. 9D), which has been shown to disrupt LLPS droplets formed by other proteins (78). We also note that in the presence of dextran, similar LLPS droplets (Fig. S4) and hydrodynamic behavior Fig. S5) could be observed at concentrations of NaCl that more closely approximated physiological ionic strength.

Since we observed that higher concentrations of the histone H3 peptide alone showed evidence of phase separation (not shown), we next determined whether the LLPS droplets we observed contained only histone H3 or if they also contained the MLL1 core complex. To do this, we assembled the MLL1 core complex with fluorescently-labeled WDR5 (W*) or RbBP5 (R*) subunits and, after purification by SEC, we tested for their ability to phase separate using fluorescence microscopy. SEC elution profiles were similar to that of unlabeled complex (Fig S6A) and SDS-PAGE showed that each fluorescent subunit eluted in a stoichiometric complex with unlabeled subunits (Fig. S6B). In addition, control experiments with each complex showed that the fluorescent tag had minimal effect on enzymatic activity (Fig. S6C, D). When reactions were examined using fluorescent microscopy, both fluorescently labeled complexes were present in the buffer and inside the droplets (Fig. 10E, F and movies S2, S3). These results

suggest that the catalytic module of the MLL1 core complex is in an equilibrium between phases both inside and outside of the LLPS droplets.

Lastly, to determine if LLPS formation rescues enzymatic activity at physiological temperature, we compared methylation kinetics of different concentrations of the MLL1 core complex among reaction mixtures containing 200 mM or 25 mM NaCl at 37°C. As described above, at near physiological ionic strength, none of the reactions went to completion, even after 24-hour incubation, mainly due to rapid enzyme inactivation at 37°C (Fig. 11, left column). In contrast, in low ionic strength buffer, most of the tested concentrations showed at least 80% conversion to the dimethylated form of H3K4 after only 5 minutes (Fig. 11, right column). At the highest concentrations tested (5 μM) the pseudo-first order rate constants for mono- and di-methylation increased 62- and 50-fold, respectively, with no evidence of enzyme inactivation (Table 5). Lastly, unlike the reactions using higher ionic strength, at low ionic strength, the reactions better approximated true single-turnover conditions with rates that were strictly dependent on enzyme concentration and not substrate concentration (79) (Fig. S7), as would be expected upon induced high-local concentration of enzyme within a biomolecular condensate.

All together, these results are consistent with the hypothesis that induced high-local concentration within a biomolecular condensate overcomes the thermodynamic barrier for MLL1 core complex assembly at physiological temperatures.

DISCUSSION

Numerous studies have established the role of MLL1 in the regulation of the degree of H3K4 methylation in mammalian gene expression and human disease. While it has been shown that the SET domain has intrinsic H3K4 monomethylation activity, several studies have shown that multiple methylation depends on interaction of MLL1 with the $WRAD_2$ sub-complex. While the molecular details for this product specificity switch are still in question, the idea that regulated complex assembly controls the spatial and temporal deposition of different H3K4 methylation states has significant experimental support. The importance of understanding the molecular details of this mode of regulation is demonstrated by studies showing targeted inhibition of the Win motif-WDR5 protein-protein interaction within the MLL1 core complex selectively reduces proliferation of MLL1-translocation leukemias and other cancer cells (42-44,80). These results suggest that molecules mimicking the Win motif, collectively called Win motif inhibitors, may be useful alternative or complementary therapeutics for cancer.

However, progress in exploiting this potential has been impeded by the lack of understanding of the biophysical and thermodynamic mechanisms that underlie MLL1 core complex assembly. The lack of standardized *in vitro* assay conditions has resulted in different conclusions regarding the mechanisms of multiple lysine methylation by SET1 family complexes and identification of the best inhibitors. For example, we previously found that the same Win motif inhibitor gives IC_{50} values that vary by more than an order of magnitude when assayed over a relatively narrow concentration range of the MLL1 core complex (0.5-1.8 μM) (81), suggesting complex assembly is relatively labile. Missing is a complete understanding of the conditions under which the complex is assembled when assayed *in vitro*. This is crucial not only for our ability to compare the potency and specificity of different inhibitors, but also for establishing a baseline for understanding how the dynamics of MLL1 core complex assembly is regulated in cells.

In this investigation, we systematically characterized the hydrodynamic and kinetic properties of a reconstituted human MLL1 core complex under a variety of assay conditions. As expected, we found that complex assembly is highly concentration and temperature dependent. Consistent with the hypothesized hierarchical assembly pathway, we found that the holo-complex assembles through interactions between the MW and RAD_2 sub-complexes, and that this assembly correlated with enzymatic activity. However, unexpectedly, we also found that the disassembled state of the complex is favored at physiological temperatures and at the sub-micromolar enzyme concentrations typically used in steady-state enzymatic assays (in which the substrate is in vast excess compared to the concentration of enzyme). We found that the complex disassembly results in rapid and irreversible enzyme inactivation under these conditions, likely because one or more subunits samples unproductive conformational states. Consistent with this conclusion, it was previously shown that overexpression of C-terminal fragments from the human SETd1A protein in mammalian cells depletes $WRAD_2$ subunits from the endogenous SETd1A and SETd1B paralogs, resulting in their degradation (82). It is possible that in the cell, unproductive folding intermediates are limited by interaction with chaperones. Consistent with this hypothesis, HSP70 and HSP90 proteins have been found to co-purify with MLL1 super-complexes (31,33). In addition, HSP90 has been shown to be required for the stability of human MLL1 and the *Drosophila melanogaster* ortholog, *Trithorax*, which is important for homeotic gene expression (83). It remains to be determined if these or other chaperones interact with and regulate folding of the subunits of the catalytic module.

Our data suggest that the MW and RAD_2 sub-complexes interact with a K_d^{app} of ~6 μM at 37°C, raising the question of how complex forms in cells that contain relatively few molecules of

MLL1, which has been estimated to be femtomoles per mg of nuclear extract (65). $WRAD_2$ subunits appear to be present in cells in vast excess compared to that of MLL1 (65), which could help overcome the thermodynamic barrier to complex assembly. However, our previous demonstration that a stoichiometric excess of WDR5 inhibits the enzymatic activity of MLL3 (84) and MLL1 (unpublished) core complexes argues against this possibility. Our data suggest that inhibition by excess WDR5 results from saturation of available binding sites on the RAD_2 sub-complex, which would prevent its interaction with the MW sub-complex. These results suggest that cellular pools of WDR5 may need to be compartmentalized to prevent this form of inhibition of SET1 family complexes. This may explain why WDR5 over-expression is associated with several poor outcome malignancies, including bladder, breast, colon, and prostate cancers, leukemias and hepatocellular carcinomas (80,85-89).

Alternative possibilities to overcome the barrier to complex formation in cells include interaction with other unknown proteins, cofactors, nucleic acids, post translational modifications, and/or by inducing a high local concentration of $MWRAD_2$ subunits within a phase separated compartment. While there is evidence that phosphorylation and long non-coding RNAs regulate the function of MLL family complexes (90,91), it is currently unclear if these mechanisms would overcome the barrier to MLL1 core complex assembly at physiological temperatures. Our data suggests that the barrier to complex formation is overcome in cells by concentration of subunits in biomolecular condensates, such as those found in transcription factories (67). Biomolecular condensates are membraneless liquid-like organelles, or intracellular phase-separated compartments, that function to concentrate proteins and nucleic acids to regulate a variety of biological processes (77,92). This form of compartmentalization has been shown to have variable effects on the activity of enzymes, ranging from a 2-70-fold stimulation in the rate of enzyme or ribozyme-catalyzed cleavage reactions, to inhibition of catalyzed reactions, protein conformational alterations and increased thermal resistance (79,93-97). While there are a number of recent examples of chromatin and chromatin-associated proteins that undergo LLPS in mechanisms that may regulate heterochromatic gene silencing (69,73,98-100), to our knowledge, there is currently no evidence demonstrating LLPS regulation of enzymatic activity of a histone modification enzyme.

Our data suggest that concentration of the MLL1 core complex in a biomolecular condensate overcomes the barrier to complex assembly at physiological temperatures, resulting in histone methyltransferase activity that is increased by at least 30-60-fold (Table 5), depending on the enzyme concentration in the assays. However, the molecular mechanism for how compartmentalization stimulates $MWRAD_2$ activity is likely more complex. This is because the

564 hydrodynamic properties of the complex change under phase separation conditions, likely
565 involving conformational changes and oligomerization that may be prerequisites for the
566 multivalent interactions required for LLPS. It is interesting to note that in the absence of a
567 crowding agent, these hydrodynamic changes begin to occur at the lower boundary of
568 physiological ionic strength. This suggests a plausible regulatory mechanism in which small
569 changes in ionic strength, possibly through compartmentalization, could have a large impact on
570 MLL1 core complex activity. However, we also note that further lowering the ionic strength of the
571 buffer (< 50 mM) resulted in detection of up to six methylation events on the same peptide (Fig.
572 S8), suggesting reduced enzyme specificity. This result, may help explain contradictory results
573 from different labs using different assay conditions. In addition, differences in the stability of MLL
574 family complexes may underlie different conclusions about their relative activities. For example,
575 we and others have observed that the MLL3 core complex is significantly more stable than the
576 other MLL family complexes (53,84,101), which may account for observations suggesting that
577 the MLL3 core complex is more active (72,102). However, we have found that when comparing
578 enzymes under conditions where complexes are at least 80% assembled, there is little
579 difference in the overall rate of H3K4 monomethylation among SET1 family complexes (84). Our
580 results here underscore the importance of assaying enzymes under conditions where
581 complexes are fully assembled, which in several cases may preclude the use of low enzyme
582 concentrations typically used in steady-state kinetics studies.

583 Our results suggest a model in which MLL1 enzymatic activity is regulated in the cell at
584 the level of complex assembly within a phase-separated transcription factory. Several lines of
585 experimental evidence are consistent with this hypothesis. Early confocal microscopy studies
586 showed that transcription occurs in a defined number of discrete sites within the cell called
587 transcription factories (103,104), each containing a protein-rich core that encompasses RNA
588 polymerase (Pol) II, co-activators, chromatin remodelers, transcription factors, histone
589 modification enzymes, ribonucleoproteins, RNA helicases, splicing and processing factors
590 (105). Indeed, peptides derived from WDR5 and DPY30, the two most abundant MLL1 core
591 complex subunits (65), were found in purified RNA Pol II transcription factories (105). A phase
592 separation model may explain, in part, immunofluorescence experiments showing that MLL1
593 has a punctate distribution within mammalian cell nuclei (66), which is a common feature of
594 proteins that undergo LLPS (77). Furthermore, use of the PScore (106) and CatGRANULE
595 (107) LLPS prediction programs show that MLL1, as well as all human MLL family proteins,
596 have high phase separation probabilities (Table S3), as does Ash2L and Ash2L-containing sub-
597 complexes (Table S4). In addition, it was recently demonstrated that the multivalent interactions

provided by the carboxyl-terminal domain (CTD) of RNA polymerase (Pol) II are sufficient for formation of RNA Pol II LLPS clusters (78). Since several studies suggest that RNA Pol II interacts directly with MLL1 (108,109), it is possible they function together within phase-separated transcription factories. Consistent with this model, ChIP studies show that MLL1 and RNA Pol II co-localize at nucleosomes throughout the promoters and open reading frames of actively-transcribed genes (109). However, a puzzling aspect of this model is that, despite a study showing that MLL1 can be pulled-down from nuclear extracts with a recombinant GST-CTD fusion protein (109), Pol II appears to be absent in purified MLL1 super-complexes (30-33). It may be that co-localization within the same transcription factory is required for the interaction.

Combining our results on the assembly of the catalytic module with the observation that it follows a large region of predicted intrinsic disorder in the primary sequence of MLL1 (Fig. 9A), we propose a “swinging domain” model for the mechanism of action of the MLL1 core complex within cellular transcription factories (Fig. 12). A swinging domain is a common feature of enzyme complexes involved in multistep assembly pathways and are characterized by a structured mobile domain tethered to other components by conformationally flexible linker regions (110). This may explain why the low complexity region is conserved not only among MLL1 orthologs, but also in the primary sequences in all human SET1 family members, with the main differences being the length of the linker regions that precedes the SET domain (Fig. S9). This observation suggests that a swinging domain may be a conserved feature of SET1 family complexes (Fig. S10D) and linker length differences could be a unique regulatory feature that limits the range of nucleosomes that can be reached within different transcriptional compartments. This hypothesis deserves further investigation.

A swinging domain model where the SET domain-WDR5 complex swings to different nucleosomes provides a satisfying explanation for how the relatively few molecules of MLL1 in the cell could methylate multiple nucleosomes in the promoter and open reading frames of genes as they move through the transcription factory (Fig. 12). This model also provides a plausible explanation for the observation of MLL1 and RNA Pol II co-localization in ChIP experiments without the necessity of a physical interaction. Given that RAD_2 subunits are relatively abundant in cells and that the RAD_2 sub-complex interacts with nucleosomes in the absence of the MW sub-complex (manuscript in preparation), concentration of both sub-complexes within a transcription factory could provide the energy required to overcome the barrier for holo-complex formation at physiological temperatures, resulting in activation of the histone methyltransferase activity of the MLL1 core complex. This model provides an elegant “switch-like” mechanism for spatiotemporal control of H3K4 methylation through the rapid

formation or dissolution of biomolecular condensates, which would ultimately regulate the hierarchical assembly of the MLL1 core complex.

ACKNOWLEDGEMENTS

We thank John Sfakis for help with the Fluorescent labeling experiments and Connie Mitra for assistance with low-salt MALDI-TOF activity assays. We thank Patty Kane for access to the Zeiss light microscope and Brian Haarer for assistance with the microscopy. We also thank Steve Hanes, Bruce Knutson and Alaji Bah for helpful suggestions and Ashley Canning and Michael Connelly for critical reading of the manuscript. This work was supported in part by R01C140522 (to M.S.C.).

METHODS

Protein Expression and Purification

Each of the human genes for the MLL1 SET domain (a.a. 3745-3969 – Uniprot #: Q03164), WDR5 (2-334 – P61964), RbBP5 (1-538 – Q15291) and Ash2L (1-534 – Q9UBL3-3) (111) were cloned into the pST44 polycistronic vector (112). The WDR5 subunit was cloned with an N-terminal 6x-Histidine tag followed by a Tobacco Etch Virus (TEV) protease cleavage site. Plasmids were transformed into Rosetta pLysS BL21 *E. coli* cells and plated on LB agar supplemented with 50 μg/mL carbenicillin and 20 μg/mL chloramphenicol (both from Gold Biotechnology). Individual colonies were used to inoculate a seed culture of 50mL of Terrific Broth II (MP Biomedicals), again supplemented with carbenicillin and chloramphenicol and grown overnight at 30°C. 20mL of the seed culture were used to inoculate 1L of Terrific Broth II media in baffled 2800mL flasks, maintaining the antibiotic resistance. Cultures were then grown for 2-4hrs at 37°C and 200RPM shaking until the $O.D._{600}$ reached ~1. Cultures were then chilled for 1hr at 4°C followed by induction with 1mM Isopropyl β-D-1-thiogalactopyranoside (IPTG – Gold Biotechnology), after which cells were grown for an additional 20-22hrs at 16°C with constant shaking. Cells were harvested by centrifugation at 4°C and pellets were flash frozen in liquid nitrogen and stored at -80°C until they could be lysed. Frozen cells were thawed and resuspended in 50 mL of lysis buffer (50mM Tris-HCl, pH 7.5; 300mM NaCl; 30mM Imidazole; 3mM dithiothreitol (DTT) and 1μM $ZnCl_2$, supplemented with one tablet of EDTA-free protease inhibitor cocktail (Roche)), lysed with a microfluidizer, and cleared by centrifugation at 17,000RPM at 4°C for 30min. The supernatant was diluted to 250 mL in Buffer 1 (50mM Tris-HCl, pH 7.5; 300mM NaCl; 30mM Imidazole; 3mM DTT and 1μM $ZnCl_2$) and flowed over a

HisTrap 5mL nickel affinity column (GE) using an AKTA Purifier FPLC (GE) at a rate of 0.5 mL/minute. Bound complex was washed with 10 column volumes (CV) of Buffer 1 at 1 mL/min., and then eluted with a 25-CV linear gradient of Buffer 2 (Buffer 1 with 500 mM imidazole). Fractions containing the MWRA complex were pooled, supplemented with GST-6x-His-TEV protease to a final concentration of 0.1 mg/mL and dialyzed against Buffer 1 with three changes. The complex was then passed over a re-equilibrated HisTrap column and fractions from the flow-through containing the cleaved MWRA sample were collected, concentrated by ultrafiltration using a 30 kDa cutoff membrane to ~12 mg/mL, and further purified by size-exclusion chromatography (SEC) using a Superdex 200 (16/60) column (GE) pre-equilibrated with Buffer 3 (20mM Tris-HCl, pH 7.5; 300mM NaCl; 1mM TCEP and 1μM $ZnCl_2$). A two-fold Molar excess of Human DPY-30 (1-99 – Q9C005), expressed and purified as previously described (36), was added to the MWRA sample and the resultant complex was purified with multiple rounds of SEC in buffer 3. Fractions containing purified $MWRAD_2$ were concentrated to 12 mg/mL, aliquoted, flash frozen, and stored at -80°C until use. Individual subunits for Bayesian experiments were purified as previously described (36).

Sedimentation Velocity-Analytical Ultracentrifugation

Experimental Procedures. All stock protein samples were thawed on ice, diluted to the desired concentration, and spun at 15,000RPM for 15min. at 4°C using a Thermo Scientific tabletop refrigerated centrifuge to remove any debris. Protein concentrations were measured with a NanoDrop spectrophotometer using the extinction coefficient ε^{280} of 248,954 M^{-1} cm^{-1}, which was predicted from the amino acid sequence using ProtParam (113). 100 or 400 μL of diluted protein samples were then loaded into AUC cells containing 3- or 12-mm two-sector charcoal-Epon centerpieces (SpinAnalytical) assembled with quartz or sapphire windows. Matching buffer was loaded into the reference sector of each cell. AUC cells were then loaded into a Ti-60 4-hole Beckman-Coulter rotor, pre-equilibrated to the specific run temperature for at least 4hrs. Rotors were then inserted into the chamber of the centrifuge and allowed to re-equilibrate to experimental temperature for a minimum of 2hrs before initiation of the run. Sedimentation velocity analytical ultracentrifugation (SV-AUC) was performed using a Beckman-Coulter Proteomelab XL-A analytical ultracentrifuge equipped with absorbance optics. Each run was preceded by a 3000-rpm wavelength scan to detect cell leakage and to select the appropriate wavelength to ensure a starting absorbance of between 0.25 and 1.2 OD units. Wavelengths at or near the maximal absorbance for aromatics of 280 nm or peptide backbone of 230 nm were selected, depending on the protein concentration and pathlength of the centerpiece. Without

699 slowing the rotor, a method scan of 50,000-rpm was initiated, and 200 scans/cell were collected
700 with the time interval between scans set to zero. Each experiment was replicated in duplicate or
701 triplicate.
702
703 *Data Analysis.* Lamm equation modeling of all SV-AUC results was performed using the
704 continuous distribution (*c(s)*) method in SEDFIT (56). Maximum entropy (ME) regularization
705 using a confidence level of P = 0.68 was performed to identify the most parsimonious
706 distribution consistent with the data, and the fits for each experiment gave acceptable RMSD
707 values ranging between 0.003 and 0.01. Density, viscosity and partial specific volume values
708 were estimated by inputting the temperature, buffer reagents, and amino acid sequences of all
709 five complex components (assuming a DPY-30 dimer) into the SEDNTERP program (114), and
710 the values used are listed in Table S5. The resulting *c(s)* distributions were displayed and
711 further analyzed using GUSSI (115). To determine the amount of holo-complex under each
712 condition, distributions were integrated between *S*-values 6.8 and 7.6, which represents one
713 standard deviation from the mean *S*-value of the holo-complex peak over all conditions, which
714 was 7.2 +/- 0.4. For binding analyses, *c(s)* distributions were integrated from 0.5 to 9.5 *S* to
715 derive the corresponding signal-weighted average sedimentation coefficients (s_w), which were
716 plotted as a function of loading concentration at each temperature and fit with mass action law
717 models using the program SEDPHAT (116).
718 For Bayesian analyses of c(s) distributions, expected sedimentation coefficients were
719 derived from separate SV-AUC experiments of individual subunits or assembled sub-
720 complexes, which were each run at concentrations ranging from 0.25 to 5 μM at 25°C (the data
721 for 0.25 μM runs are shown in Fig. S1). These values were then used in ME regularization as
722 prior expectation restraints to give $c^{(p)}(s)$ distributions of the holo-complex at 25°C. Prior
723 expectations for sub-complexes or individual subunits were implemented as Gaussians in
724 SEDFIT for Bayesian analysis, with a peak width of sigma = 0.2 *S* and centered at the weight-
725 average *S*-value of the main peak observed in the individual experiments with an amplitude of
726 0.05 OD units. Since the prior expected *S*-values for WDR5 or RbBP5 overlapped when run in
727 individual experiments, they were used as prior expectations in $c^{(p)}(s)$ distributions to test the
728 concerted assembly mechanism with the same weight average *S*-value but with an amplitude
729 that was doubled (Fig. 5D). Each $c^{(p)}(s)$ distribution was fit with the same prior expectation for
730 $MWRAD_2$, which used the weight-average *S*-value determined at 25°C and 0.25 μM with a width
731 of sigma = 0.4 *S* and an amplitude of 0.3 OD units.

For the $c(s,f_r)$ analysis, the data were first imported into SEDFIT with reduced radial resolution (0.006cm compared to the default 0.003cm) and loading every second scan, to reduce the computational power required (60). These were fit using the $c(s, f_r)$ method in SEDFIT with resolutions of 50 for both the sedimentation coefficient and frictional ratio dimensions.

MethylItransferase Activity Assay

$MWRAD_2$ complex was assayed using a label-free quantitative MALDI-TOF mass spectrometry assay (36). Each 20 μL reaction consisted of varying concentrations of $MWRAD_2$, 250 μM S-adenosylmethionine (AdoMet) and reaction buffer (50 mM Tris, pH 9.0; 200 mM NaCl; 5% (v/v) glycerol; 1 μM $ZnCl_2$; 3 mM DTT), which were preincubated for 5 minutes at the experimental temperature in a thermocycler. Reactions were initiated by the addition of temperature-pre-equilibrated histone H3 peptide (residues 1-20, with an additional C-terminal GGK-biotin moiety) to a final concentration of 10 μM. At various timepoints, a 2 μL aliquot was removed and quenched by mixing with 2 μL of 1% trifluoroacetic acid (TFA). Quenched reactions were stored at -20°C until they could be analyzed. Upon analysis, samples were thawed and 1 μL of each was mixed with 4 μL of α-cyano-4-hydroxycinnamic acid in 0.05% TFA and 50% acetonitrile. 2 μL of this mixture for each time point was spotted onto a ground steel target plate and allowed to dry at room temperature for 3-12 hours. Spectra were acquired on a Bruker Autoflex III MALDI-TOF mass spectrometer in reflectron mode. Each spectrum was the sum of at least 1000 individual laser shots, obtained from five different positions around the spot, with 200 shots at each position. Using FlexAnalysis software (Bruker), the intensities of the unmodified (m/z 2651 Da), mono- (m/z 2665 Da), di- (m/z 2679 Da), and trimethylated (m/z 2693 Da) species were summed to obtain the total intensity. The relative amount of each species was then determined by dividing the intensity of each methylation state by the total intensity at each time point and multiplied by the starting substrate concentration (10μM) to give the micromolar concentration of each methylation state. These data were then plotted as a function of time for kinetics analyses.

Fitting of the data was performed using the numerical integration of rate equations approach implemented in KinTek Explorer software version 6.3 (61). For reaction schemes incorporating the complex dissociation step, the ratio (k_{off}/k_{on}) was constrained to be equal to estimated K_d^{app} for complex dissociation at each temperature determined from the sedimentation velocity s_w isotherm analysis, with the k_{on} fixed at the limit of diffusion. All other

765 non-variable parameters were fixed with non-rate limiting values. Confidence contour analysis
766 using a Chi^2 threshold of 0.9 was used to obtain estimates for the extent to which each variable
767 parameter was constrained by the data.
768
769 ***Labeling and assembly of fluorescent MWRAD$_2$ complexes***
770 Recombinant WDR5 or RbBP5 were expressed and purified as previously described
771 (37). Purified proteins at ~14 mg/ml were dialyzed into labeling buffer composed of 20 mM
772 HEPES, pH 7.0; 300 mM NaCl; 1 mM TCEP and 1 μM $ZnCl_2$. The neutral pH was chosen to
773 facilitate selective labeling of the free amino terminus of the protein, which has a lower pK_a than
774 the primary amines of the lysine side chains (117). The protein was mixed with AlexaFluor™
775 488 NHS Ester (Invitrogen) in a 1:6 (for WDR5) or 1:5 (for RbBP5) molar excess of label and
776 reacted for 3 hours at 4°C. The entire reaction volume for each protein was then loaded onto a
777 Superdex™ 200 10/300 GL size-exclusion column (GE) to separate the labeled protein from the
778 unreacted fluorophore. The labeled protein fractions were then combined and concentrated by
779 ultrafiltration in a 10,000 MWCO concentrator (Millipore). Once concentrated, the degree of
780 labeling was determined using the equations shown below:
781
782 $$A_{protein} = A_{280} - A_{max} * (correction\ factor)$$
783
784 $$A_{protein} \,/\, (pathlength * \varepsilon_{protein}) = [protein]$$
785
786 $$Degree\ of\ labeling = (A_{max} * (protein\ MW)) \,/\, ([protein] * \varepsilon_{dye})$$
787
788 The degree of labeling for WDR5 (W*), was found to be 1.1, or ~ 1 molecule of fluorophore for
789 each molecule of WDR5. The degree of labeling determined for RbBP5 (R*) was 1.9, or ~2
790 molecules of fluorophore per molecule of RbBP5. Each labeled protein was then mixed in
791 equivalent molar ratios with the other recombinant, unlabeled complex components and loaded
792 onto a Superdex™ 200 10/300 GL size-exclusion column and fractions containing stoichiometric
793 complex were pooled, concentrated, and stored at -80°C until use.
794
795 ***Liquid-Liquid Phase Separation (LLPS) Assays***
796 MWRAD$_2$ at a concentration of 5 μM was mixed with $H3^{1-20}$ peptide (100 – 500 μM) and
797 250 μM SAM in either physiological (~100-150 mM) or sub-physiological (~25-50 mM) NaCl

798 buffers containing (50 mM Tris, pH 9.0, 1 μM $ZnCl_2$, 3 mM DTT and 5% (w/v) glycerol) in the
799 presence or absence of 7% (w/v) Dextran Sulfate (avg. M.W. = 500,000 Da) as a crowding
800 agent. 1 μL of each sample was pipetted into the depression of 12-well precleaned frosted end
801 Bioworld microscope slide, covered by a cover slip, and observed on a Zeiss light microscope in
802 DIC mode at 40x magnification. Single images and movies were taken using a Hamamatsu
803 camera connected to the microscope. All images taken are of samples at room temperature
804 (~23°C). In addition to DIC, M(W*)RAD_2 or MW(R*)AD_2 were imaged with the FITC filter
805 activated. As a control for phase separation, reaction mixtures were compared in the presence
806 and absence of 5%1,6 hexanediol.
807

REFERENCES

1. Ruthenburg, A. J., Allis, C. D., and Wysocka, J. (2007) Methylation of lysine 4 on histone H3: intricacy of writing and reading a single epigenetic mark. *Mol Cell* **25**, 15-30
2. Liu, C. L., Kaplan, T., Kim, M., Buratowski, S., Schreiber, S. L., Friedman, N., and Rando, O. J. (2005) Single-nucleosome mapping of histone modifications in S. cerevisiae. *PLoS Biol* **3**, e328
3. Pokholok, D. K., Harbison, C. T., Levine, S., Cole, M., Hannett, N. M., Lee, T. I., Bell, G. W., Walker, K., Rolfe, P. A., Herbolsheimer, E., Zeitlinger, J., Lewitter, F., Gifford, D. K., and Young, R. A. (2005) Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell* **122**, 517-527
4. Santos-Rosa, H., Schneider, R., Bannister, A. J., Sherriff, J., Bernstein, B. E., Emre, N. C., Schreiber, S. L., Mellor, J., and Kouzarides, T. (2002) Active genes are tri-methylated at K4 of histone H3. *Nature* **419**, 407-411
5. Wysocka, J., Swigut, T., Xiao, H., Milne, T. A., Kwon, S. Y., Landry, J., Kauer, M., Tackett, A. J., Chait, B. T., Badenhorst, P., Wu, C., and Allis, C. D. (2006) A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* **442**, 86-90
6. Santos-Rosa, H., Schneider, R., Bernstein, B. E., Karabetsou, N., Morillon, A., Weise, C., Schreiber, S. L., Mellor, J., and Kouzarides, T. (2003) Methylation of histone H3 K4 mediates association of the Isw1p ATPase with chromatin. *Mol Cell* **12**, 1325-1332
7. Sims, R. J., 3rd, Chen, C. F., Santos-Rosa, H., Kouzarides, T., Patel, S. S., and Reinberg, D. (2005) Human but not yeast CHD1 binds directly and selectively to histone H3 methylated at lysine 4 via its tandem chromodomains. *J Biol Chem* **280**, 41789-41792
8. Pray-Grant, M. G., Daniel, J. A., Schieltz, D., Yates, J. R., 3rd, and Grant, P. A. (2005) Chd1 chromodomain links histone H3 methylation with SAGA- and SLIK-dependent acetylation. *Nature* **433**, 434-438
9. Kim, T., and Buratowski, S. (2009) Dimethylation of H3K4 by Set1 recruits the Set3 histone deacetylase complex to 5' transcribed regions. *Cell* **137**, 259-272
10. Heintzman, N. D., Stuart, R. K., Hon, G., Fu, Y., Ching, C. W., Hawkins, R. D., Barrera, L. O., Van Calcar, S., Qu, C., Ching, K. A., Wang, W., Weng, Z., Green, R. D., Crawford, G. E., and Ren, B. (2007) Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet* **39**, 311-318
11. Hu, D., Gao, X., Morgan, M. A., Herz, H. M., Smith, E. R., and Shilatifard, A. (2013) The MLL3/MLL4 branches of the COMPASS family function as major histone H3K4 monomethylases at enhancers. *Mol Cell Biol* **33**, 4745-4754
12. Lee, J. E., Wang, C., Xu, S., Cho, Y. W., Wang, L., Feng, X., Baldridge, A., Sartorelli, V., Zhuang, L., Peng, W., and Ge, K. (2013) H3K4 mono- and di-methyltransferase MLL4 is required for enhancer activation during cell differentiation. *eLife* **2**, e01503
13. Cheng, J., Blum, R., Bowman, C., Hu, D., Shilatifard, A., Shen, S., and Dynlacht, B. D. (2014) A role for H3K4 monomethylation in gene repression and partitioning of chromatin readers. *Mol Cell* **53**, 979-992

14. van Dijk, K., Marley, K. E., Jeong, B. R., Xu, J., Hesson, J., Cerny, R. L., Waterborg, J. H., and Cerutti, H. (2005) Monomethyl histone H3 lysine 4 as an epigenetic mark for silenced euchromatin in Chlamydomonas. *Plant Cell* **17**, 2439-2453
15. Bryk, M., Briggs, S. D., Strahl, B. D., Curcio, M. J., Allis, C. D., and Winston, F. (2002) Evidence that Set1, a factor required for methylation of histone H3, regulates rDNA silencing in S. cerevisiae by a Sir2-independent mechanism. *Curr Biol* **12**, 165-170
16. Krogan, N. J., Dover, J., Khorrami, S., Greenblatt, J. F., Schneider, J., Johnston, M., and Shilatifard, A. (2002) COMPASS, a histone H3 (Lysine 4) methyltransferase required for telomeric silencing of gene expression. *J Biol Chem* **277**, 10753-10755
17. Nislow, C., Ray, E., and Pillus, L. (1997) SET1, a yeast member of the trithorax family, functions in transcriptional silencing and diverse cellular processes. *Mol Biol Cell* **8**, 2421-2436
18. Wong, S. H., Goode, D. L., Iwasaki, M., Wei, M. C., Kuo, H. P., Zhu, L., Schneidawind, D., Duque-Afonso, J., Weng, Z., and Cleary, M. L. (2015) The H3K4-Methyl Epigenome Regulates Leukemia Stem Cell Oncogenic Potential. *Cancer Cell* **28**, 198-209
19. Dou, Y., and Hess, J. L. (2008) Mechanisms of transcriptional regulation by MLL and its disruption in acute leukemia. *Int J Hematol* **87**, 10-18
20. Dalgliesh, G. L., Furge, K., Greenman, C., Chen, L., Bignell, G., Butler, A., Davies, H., Edkins, S., Hardy, C., Latimer, C., Teague, J., Andrews, J., Barthorpe, S., Beare, D., Buck, G., Campbell, P. J., Forbes, S., Jia, M., Jones, D., Knott, H., Kok, C. Y., Lau, K. W., Leroy, C., Lin, M. L., McBride, D. J., Maddison, M., Maguire, S., McLay, K., Menzies, A., Mironenko, T., Mulderrig, L., Mudie, L., O'Meara, S., Pleasance, E., Rajasingham, A., Shepherd, R., Smith, R., Stebbings, L., Stephens, P., Tang, G., Tarpey, P. S., Turrell, K., Dykema, K. J., Khoo, S. K., Petillo, D., Wondergem, B., Anema, J., Kahnoski, R. J., Teh, B. T., Stratton, M. R., and Futreal, P. A. (2010) Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes. *Nature* **463**, 360-363
21. Morin, R. D., Mendez-Lago, M., Mungall, A. J., Goya, R., Mungall, K. L., Corbett, R. D., Johnson, N. A., Severson, T. M., Chiu, R., Field, M., Jackman, S., Krzywinski, M., Scott, D. W., Trinh, D. L., Tamura-Wells, J., Li, S., Firme, M. R., Rogic, S., Griffith, M., Chan, S., Yakovenko, O., Meyer, I. M., Zhao, E. Y., Smailus, D., Moksa, M., Chittaranjan, S., Rimsza, L., Brooks-Wilson, A., Spinelli, J. J., Ben-Neriah, S., Meissner, B., Woolcock, B., Boyle, M., McDonald, H., Tam, A., Zhao, Y., Delaney, A., Zeng, T., Tse, K., Butterfield, Y., Birol, I., Holt, R., Schein, J., Horsman, D. E., Moore, R., Jones, S. J., Connors, J. M., Hirst, M., Gascoyne, R. D., and Marra, M. A. (2011) Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. *Nature* **476**, 298-303
22. Parsons, D. W., Li, M., Zhang, X., Jones, S., Leary, R. J., Lin, J. C., Boca, S. M., Carter, H., Samayoa, J., Bettegowda, C., Gallia, G. L., Jallo, G. I., Binder, Z. A., Nikolsky, Y., Hartigan, J., Smith, D. R., Gerhard, D. S., Fults, D. W., VandenBerg, S., Berger, M. S., Marie, S. K., Shinjo, S. M., Clara, C., Phillips, P. C., Minturn, J. E., Biegel, J. A., Judkins, A. R., Resnick, A. C., Storm, P. B., Curran, T., He, Y., Rasheed, B. A., Friedman, H. S., Keir, S. T., McLendon, R., Northcott, P. A., Taylor, M. D., Burger, P. C., Riggins, G. J., Karchin, R., Parmigiani, G., Bigner, D. D., Yan, H., Papadopoulos, N., Vogelstein, B., Kinzler, K. W., and Velculescu, V. E. (2011) The genetic landscape of the childhood cancer medulloblastoma. *Science* **331**, 435-439

23. Pasqualucci, L., Trifonov, V., Fabbri, G., Ma, J., Rossi, D., Chiarenza, A., Wells, V. A., Grunn, A., Messina, M., Elliot, O., Chan, J., Bhagat, G., Chadburn, A., Gaidano, G., Mullighan, C. G., Rabadan, R., and Dalla-Favera, R. (2011) Analysis of the coding genome of diffuse large B-cell lymphoma. *Nat Genet* **43**, 830-837
24. Grasso, C. S., Wu, Y. M., Robinson, D. R., Cao, X., Dhanasekaran, S. M., Khan, A. P., Quist, M. J., Jing, X., Lonigro, R. J., Brenner, J. C., Asangani, I. A., Ateeq, B., Chun, S. Y., Siddiqui, J., Sam, L., Anstett, M., Mehra, R., Prensner, J. R., Palanisamy, N., Ryslik, G. A., Vandin, F., Raphael, B. J., Kunju, L. P., Rhodes, D. R., Pienta, K. J., Chinnaiyan, A. M., and Tomlins, S. A. (2012) The mutational landscape of lethal castration-resistant prostate cancer. *Nature* **487**, 239-243
25. Jones, W. D., Dafou, D., McEntagart, M., Woollard, W. J., Elmslie, F. V., Holder-Espinasse, M., Irving, M., Saggar, A. K., Smithson, S., Trembath, R. C., Deshpande, C., and Simpson, M. A. (2012) De novo mutations in MLL cause Wiedemann-Steiner syndrome. *American journal of human genetics* **91**, 358-364
26. Shilatifard, A. (2012) The COMPASS Family of Histone H3K4 Methylases: Mechanisms of Regulation in Development and Disease Pathogenesis. *Annu Rev Biochem* **81**, 65-95
27. Bogershausen, N., and Wollnik, B. (2013) Unmasking Kabuki syndrome. *Clinical genetics* **83**, 201-211
28. Flavahan, W. A., Gaskell, E., and Bernstein, B. E. (2017) Epigenetic plasticity and the hallmarks of cancer. *Science* **357**
29. Ziemin-van der Poel, S., McCabe, N. R., Gill, H. J., Espinosa, R., III, Patel, Y., Harden, A., Rubinelli, P., Smith, S. D., LeBeau, M. M., Rowley, J. D., and et al. (1991) Identification of a gene, MLL, that spans the breakpoint in 11q23 translocations associated with human leukemias. *Proc Natl Acad Sci U S A* **88**, 10735-10739
30. Nakamura, T., Mori, T., Tada, S., Krajewski, W., Rozovskaia, T., Wassell, R., Dubois, G., Mazo, A., Croce, C. M., and Canaani, E. (2002) ALL-1 is a histone methyltransferase that assembles a supercomplex of proteins involved in transcriptional regulation. *Mol Cell* **10**, 1119-1128
31. Dou, Y., Milne, T. A., Tackett, A. J., Smith, E. R., Fukuda, A., Wysocka, J., Allis, C. D., Chait, B. T., Hess, J. L., and Roeder, R. G. (2005) Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF. *Cell* **121**, 873-885
32. Yokoyama, A., Wang, Z., Wysocka, J., Sanyal, M., Aufiero, D. J., Kitabayashi, I., Herr, W., and Cleary, M. L. (2004) Leukemia proto-oncoprotein MLL forms a SET1-like histone methyltransferase complex with menin to regulate Hox gene expression. *Mol Cell Biol* **24**, 5639-5649
33. Wysocka, J., Myers, M. P., Laherty, C. D., Eisenman, R. N., and Herr, W. (2003) Human Sin3 deacetylase and trithorax-related Set1/Ash2 histone H3-K4 methyltransferase are tethered together selectively by the cell-proliferation factor HCF-1. *Genes Dev* **17**, 896-911
34. Dou, Y., Milne, T. A., Ruthenburg, A. J., Lee, S., Lee, J. W., Verdine, G. L., Allis, C. D., and Roeder, R. G. (2006) Regulation of MLL1 H3K4 methyltransferase activity by its core components. *Nat Struct Mol Biol* **13**, 713-719

35. Patel, A., Dharmarajan, V., and Cosgrove, M. S. (2008) Structure of WDR5 bound to Mixed Lineage Leukemia Protein-1 peptide. *J Biol Chem* **283**, 32158-32161
36. Patel, A., Dharmarajan, V., Vought, V. E., and Cosgrove, M. S. (2009) On the mechanism of multiple lysine methylation by the human mixed lineage leukemia protein-1 (MLL1) core complex. *J Biol Chem* **284**, 24242-24256
37. Patel, A., Vought, V., Dharmarajan, V., and Cosgrove, M. S. (2008) A conserved arginine containing motif crucial for the assembly and enzymatic activity of the Mixed Lineage Leukemia protein-1 core complex. *J Biol Chem* **283**, 32162-32175
38. Roguev, A., Schaft, D., Shevchenko, A., Pijnappel, W. W., Wilm, M., Aasland, R., and Stewart, A. F. (2001) The Saccharomyces cerevisiae Set1 complex includes an Ash2 homologue and methylates histone 3 lysine 4. *Embo J* **20**, 7137-7148
39. Southall, S. M., Wong, P. S., Odho, Z., Roe, S. M., and Wilson, J. R. (2009) Structural basis for the requirement of additional factors for MLL1 SET domain activity and recognition of epigenetic marks. *Mol Cell* **33**, 181-191
40. Scacheri, P. C., Davis, S., Odom, D. T., Crawford, G. E., Perkins, S., Halawi, M. J., Agarwal, S. K., Marx, S. J., Spiegel, A. M., Meltzer, P. S., and Collins, F. S. (2006) Genome-Wide Analysis of Menin Binding Provides Insights into MEN1 Tumorigenesis. *PLoS Genet* **2**, e51, 0406-0419
41. Shinsky, S. A., Hu, M., Vought, V. E., Ng, S. B., Bamshad, M. J., Shendure, J., and Cosgrove, M. S. (2014) A non-active site SET domain surface crucial for the interaction of MLL1 and the RbBP5/Ash2L heterodimer within MLL family core complexes. *J. Mol. Biol.* **426**, 2283-2299
42. Grebien, F., Vedadi, M., Getlik, M., Giambruno, R., Grover, A., Avellino, R., Skucha, A., Vittori, S., Kuznetsova, E., Smil, D., Barsyte-Lovejoy, D., Li, F., Poda, G., Schapira, M., Wu, H., Dong, A., Senisterra, G., Stukalov, A., Huber, K. V., Schonegger, A., Marcellus, R., Bilban, M., Bock, C., Brown, P. J., Zuber, J., Bennett, K. L., Al-Awar, R., Delwel, R., Nerlov, C., Arrowsmith, C. H., and Superti-Furga, G. (2015) Pharmacological targeting of the Wdr5-MLL interaction in C/EBPalpha N-terminal leukemia. *Nat Chem Biol* **11**, 571-578
43. Zhu, J., Sammons, M. A., Donahue, G., Dou, Z., Vedadi, M., Getlik, M., Barsyte-Lovejoy, D., Al-awar, R., Katona, B. W., Shilatifard, A., Huang, J., Hua, X., Arrowsmith, C. H., and Berger, S. L. (2015) Gain-of-function p53 mutants co-opt chromatin pathways to drive cancer growth. *Nature* **525**, 206-211
44. Cao, F., Townsend, E. C., Karatas, H., Xu, J., Li, L., Lee, S., Liu, L., Chen, Y., Ouillette, P., Zhu, J., Hess, J. L., Atadja, P., Lei, M., Qin, Z. S., Malek, S., Wang, S., and Dou, Y. (2014) Targeting MLL1 H3K4 methyltransferase activity in mixed-lineage leukemia. *Mol Cell* **53**, 247-261
45. Song, J. J., and Kingston, R. E. (2008) WDR5 interacts with mixed lineage leukemia (MLL) protein via the histone H3-binding pocket. *J Biol Chem* **283**, 35258-35264
46. Avdic, V., Zhang, P., Lanouette, S., Groulx, A., Tremblay, V., Brunzelle, J., and Couture, J. F. Structural and biochemical insights into MLL1 core complex assembly. *Structure* **19**, 101-108
47. Odho, Z., Southall, S. M., and Wilson, J. R. Characterization of a novel WDR5-binding site that recruits RbBP5 through a conserved motif to enhance methylation of histone H3 lysine 4 by mixed lineage leukemia protein-1. *J Biol Chem* **285**, 32967-32976

48. South, P. F., Fingerman, I. M., Mersman, D. P., Du, H. N., and Briggs, S. D. (2010) A conserved interaction between the SDI domain of Bre2 and the Dpy-30 domain of Sdc1 is required for histone methylation and gene expression. *J Biol Chem* **285**, 595-607
49. Cho, Y. W., Hong, T., Hong, S., Guo, H., Yu, H., Kim, D., Guszczynski, T., Dressler, G. R., Copeland, T. D., Kalkum, M., and Ge, K. (2007) PTIP associates with MLL3- and MLL4-containing histone H3 lysine 4 methyltransferase complex. *J Biol Chem* **282**, 20395-20406
50. Qu, Q., Takahashi, Y. H., Yang, Y., Hu, H., Zhang, Y., Brunzelle, J. S., Couture, J. F., Shilatifard, A., and Skiniotis, G. (2018) Structure and Conformational Dynamics of a COMPASS Histone H3K4 Methyltransferase Complex. *Cell* **174**, 1117-1126 e1112
51. Hsu, P. L., Li, H., Lau, H. T., Leonen, C., Dhall, A., Ong, S. E., Chatterjee, C., and Zheng, N. (2018) Crystal Structure of the COMPASS H3K4 Methyltransferase Catalytic Module. *Cell* **174**, 1106-1116 e1109
52. Takahashi, Y. H., Westfield, G. H., Oleskie, A. N., Trievel, R. C., Shilatifard, A., and Skiniotis, G. (2011) Structural analysis of the core COMPASS family of histone H3K4 methylases from yeast to human. *Proc Natl Acad Sci U S A* **108**, 20526-20531
53. Li, Y., Han, J., Zhang, Y., Cao, F., Liu, Z., Li, S., Wu, J., Hu, C., Wang, Y., Shuai, J., Chen, J., Cao, L., Li, D., Shi, P., Tian, C., Zhang, J., Dou, Y., Li, G., Chen, Y., and Lei, M. (2016) Structural basis for activity regulation of MLL family methyltransferases. *Nature* **530**, 447-452
54. Lebowitz, J., Lewis, M. S., and Schuck, P. (2002) Modern analytical ultracentrifugation in protein science: a tutorial review. *Protein science : a publication of the Protein Society* **11**, 2067-2079
55. Schuck, P. (2013) Analytical Ultracentrifugation as a Tool for Studying Protein Interactions. *Biophys Rev* **5**, 159-171
56. Schuck, P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys J* **78**, 1606-1619
57. Schuck, P. (2005) Diffusion-deconvoluted sedimentation coefficient distributions for the analysis of interacting and non-interacting protein mixtures. in *Analytical Ultracentrifugation: Techniques and Methods* (Scott, D. J., Harding, S. E., and Rowe, A. J. eds.), Royal Society of Chemistry, Cambridge, UK. pp 26-49
58. Schuck, P. (2003) On the analysis of protein self-association by sedimentation velocity analytical ultracentrifugation. *Anal Biochem* **320**, 104-124
59. Dam, J., Velikovsky, C. A., Mariuzza, R. A., Urbanke, C., and Schuck, P. (2005) Sedimentation velocity analysis of heterogeneous protein-protein interactions: Lamm equation modeling and sedimentation coefficient distributions c(s). *Biophys J* **89**, 619-634
60. Brown, P. H., Balbo, A., and Schuck, P. (2007) Using prior knowledge in the determination of macromolecular size-distributions by analytical ultracentrifugation. *Biomacromolecules* **8**, 2011-2024
61. Johnson, K. A., Simpson, Z. B., and Blom, T. (2009) Global kinetic explorer: a new computer program for dynamic simulation and fitting of kinetic data. *Anal Biochem* **387**, 20-29

62. Patel, A., Vought, V. E., Dharmarajan, V., and Cosgrove, M. S. (2011) A novel non-SET domain multi-subunit methyltransferase required for sequential nucleosomal histone H3 methylation by the mixed lineage leukemia protein-1 (MLL1) core complex. *J Biol Chem* **286**, 3359-3369
63. Patel, A., Vought, V. E., Swatkoski, S., Viggiano, S., Howard, B., Dharmarajan, V., Monteith, K. E., Kupakuwana, G., Namitz, K. E., Shinsky, S. A., Cotter, R. J., and Cosgrove, M. S. (2013) Automethylation activities within the Mixed Lineage Leukemia-1 (MLL1) core complex reveal evidence supporting a "two-active site" model for multiple histone H3 lysine 4 methylation. *J Biol Chem*
64. Johnson, K. A., Simpson, Z. B., and Blom, T. (2009) FitSpace explorer: an algorithm to evaluate multidimensional parameter space in fitting kinetic data. *Anal Biochem* **387**, 30-41
65. van Nuland, R., Smits, A. H., Pallaki, P., Jansen, P. W., Vermeulen, M., and Timmers, H. T. (2013) Quantitative dissection and stoichiometry determination of the human SET1/MLL histone methyltransferase complexes. *Mol Cell Biol* **33**, 2067-2077
66. Yano, T., Nakamura, T., Blechman, J., Sorio, C., Dang, C. V., Geiger, B., and Canaani, E. (1997) Nuclear punctate distribution of ALL-1 is conferred by distinct elements at the N terminus of the protein. *Proc Natl Acad Sci U S A* **94**, 7286-7291
67. Rieder, D., Trajanoski, Z., and McNally, J. G. (2012) Transcription factories. *Front Genet* **3**, 221
68. Papantonis, A., and Cook, P. R. (2013) Transcription factories: genome organization and gene regulation. *Chem Rev* **113**, 8683-8705
69. Gibson, B. A., Doolittle, L. K., Schneider, M. W. G., Jensen, L. E., Gamarra, N., Henry, L., Gerlich, D. W., Redding, S., and Rosen, M. K. (2019) Organization of Chromatin by Intrinsic and Regulated Phase Separation. *Cell* **179**, 470-484 e421
70. Mitrea, D. M., and Kriwacki, R. W. (2016) Phase separation in biology; functional organization of a higher order. *Cell Commun Signal* **14**, 1
71. Dosztanyi, Z., Csizmok, V., Tompa, P., and Simon, I. (2005) The pairwise energy content estimated from amino acid composition discriminates between folded and intrinsically unstructured proteins. *J Mol Biol* **347**, 827-839
72. Vedadi, M., Blazer, L., Eram, M. S., Barsyte-Lovejoy, D., Arrowsmith, C. H., and Hajian, T. (2017) Targeting human SET1/MLL family of proteins. *Protein science : a publication of the Protein Society* **26**, 662-676
73. Larson, A. G., Elnatan, D., Keenen, M. M., Trnka, M. J., Johnston, J. B., Burlingame, A. L., Agard, D. A., Redding, S., and Narlikar, G. J. (2017) Liquid droplet formation by HP1alpha suggests a role for phase separation in heterochromatin. *Nature* **547**, 236-240
74. Dam, J., and Schuck, P. (2004) Calculating sedimentation coefficient distributions by direct modeling of sedimentation velocity concentration profiles. *Methods Enzymol* **384**, 185-212
75. Brown, P. H., and Schuck, P. (2006) Macromolecular size-and-shape distributions by sedimentation velocity analytical ultracentrifugation. *Biophys J* **90**, 4651-4661
76. Chaton, C. T., and Herr, A. B. (2015) Elucidating Complicated Assembling Systems in Biology Using Size-and-Shape Analysis of Sedimentation Velocity Data. *Methods Enzymol* **562**, 187-204

77. Banani, S. F., Lee, H. O., Hyman, A. A., and Rosen, M. K. (2017) Biomolecular condensates: organizers of cellular biochemistry. *Nat Rev Mol Cell Biol* **18**, 285-298
78. Boehning, M., Dugast-Darzacq, C., Rankovic, M., Hansen, A. S., Yu, T., Marie-Nelly, H., McSwiggen, D. T., Kokic, G., Dailey, G. M., Cramer, P., Darzacq, X., and Zweckstetter, M. (2018) RNA polymerase II clustering through carboxy-terminal domain phase separation. *Nat Struct Mol Biol* **25**, 833-840
79. Strulson, C. A., Molden, R. C., Keating, C. D., and Bevilacqua, P. C. (2012) RNA catalysis through compartmentalization. *Nat Chem* **4**, 941-946
80. Neilsen, B. K., Chakraborty, B., McCall, J. L., Frodyma, D. E., Sleightholm, R. L., Fisher, K. W., and Lewis, R. E. (2018) WDR5 supports colon cancer cells by promoting methylation of H3K4 and suppressing DNA damage. *BMC cancer* **18**, 673
81. Alicea-Velazquez, N. L., Shinsky, S. A., Loh, D. M., Lee, J. H., Skalnik, D. G., and Cosgrove, M. S. (2016) Targeted Disruption of the Interaction between WD-40 Repeat Protein 5 (WDR5) and Mixed Lineage Leukemia (MLL)/SET1 Family Proteins Specifically Inhibits MLL1 and SETd1A Methyltransferase Complexes. *J Biol Chem* **291**, 22357-22372
82. Lee, J. H., Tate, C. M., You, J. S., and Skalnik, D. G. (2007) Identification and characterization of the human Set1B histone H3-Lys4 methyltransferase complex. *J Biol Chem* **282**, 13419-13428
83. Tariq, M., Nussbaumer, U., Chen, Y., Beisel, C., and Paro, R. (2009) Trithorax requires Hsp90 for maintenance of active chromatin at sites of gene expression. *Proc Natl Acad Sci U S A* **106**, 1157-1162
84. Shinsky, S. A., Monteith, K. E., Viggiano, S., and Cosgrove, M. S. (2015) Biochemical reconstitution and phylogenetic comparison of human SET1 family core complexes involved in histone methylation. *J Biol Chem* **290**, 6361-6375
85. Kim, J. Y., Banerjee, T., Vinckevicius, A., Luo, Q., Parker, J. B., Baker, M. R., Radhakrishnan, I., Wei, J. J., Barish, G. D., and Chakravarti, D. (2014) A role for WDR5 in integrating threonine 11 phosphorylation to lysine 4 methylation on histone H3 during androgen signaling and in prostate cancer. *Mol Cell* **54**, 613-625
86. Chen, X., Xie, W., Gu, P., Cai, Q., Wang, B., Xie, Y., Dong, W., He, W., Zhong, G., Lin, T., and Huang, J. (2015) Upregulated WDR5 promotes proliferation, self-renewal and chemoresistance in bladder cancer via mediating H3K4 trimethylation. *Sci Rep* **5**, 8293
87. Ge, Z., Song, E. J., Kawasawa, Y. I., Li, J., Dovat, S., and Song, C. (2016) WDR5 high expression and its effect on tumorigenesis in leukemia. *Oncotarget* **7**, 37740-37754
88. Dai, X., Guo, W., Zhan, C., Liu, X., Bai, Z., and Yang, Y. (2015) WDR5 Expression Is Prognostic of Breast Cancer Outcome. *PLoS One* **10**, e0124964
89. Cui, Z., Li, H., Liang, F., Mu, C., Mu, Y., Zhang, X., and Liu, J. (2018) Effect of high WDR5 expression on the hepatocellular carcinoma prognosis. *Oncol Lett* **15**, 7864-7870
90. Zhang, P., Chaturvedi, C. P., Tremblay, V., Cramet, M., Brunzelle, J. S., Skiniotis, G., Brand, M., Shilatifard, A., and Couture, J. F. (2015) A phosphorylation switch on RbBP5 regulates histone H3 Lys4 methylation. *Genes Dev* **29**, 123-128
91. Wang, K. C., Yang, Y. W., Liu, B., Sanyal, A., Corces-Zimmerman, R., Chen, Y., Lajoie, B. R., Protacio, A., Flynn, R. A., Gupta, R. A., Wysocka, J., Lei, M., Dekker, J., Helms, J. A., and Chang, H. Y. (2011) A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* **472**, 120-124

92. Li, P., Banjade, S., Cheng, H. C., Kim, S., Chen, B., Guo, L., Llaguno, M., Hollingsworth, J. V., King, D. S., Banani, S. F., Russo, P. S., Jiang, Q. X., Nixon, B. T., and Rosen, M. K. (2012) Phase transitions in the assembly of multivalent signalling proteins. *Nature* **483**, 336-340
93. Koga, S., Williams, D. S., Perriman, A. W., and Mann, S. (2011) Peptide-nucleotide microdroplets as a step towards a membrane-free protocell model. *Nat Chem* **3**, 720-724
94. Dhar, A., Samiotakis, A., Ebbinghaus, S., Nienhaus, L., Homouz, D., Gruebele, M., and Cheung, M. S. (2010) Structure, function, and folding of phosphoglycerate kinase are strongly perturbed by macromolecular crowding. *Proc Natl Acad Sci U S A* **107**, 17586-17591
95. Pozdnyakova, I., and Wittung-Stafshede, P. (2010) Non-linear effects of macromolecular crowding on enzymatic activity of multi-copper oxidase. *Biochim Biophys Acta* **1804**, 740-744
96. Aumiller, W. M., Jr., Davis, B. W., Hatzakis, E., and Keating, C. D. (2014) Interactions of macromolecular crowding agents and cosolutes with small-molecule substrates: effect on horseradish peroxidase activity with two different substrates. *J Phys Chem B* **118**, 10624-10632
97. Pastor, I., Vilaseca, E., Madurga, S., Garces, J. L., Cascante, M., and Mas, F. (2011) Effect of crowding by dextrans on the hydrolysis of N-Succinyl-L-phenyl-Ala-p-nitroanilide catalyzed by alpha-chymotrypsin. *J Phys Chem B* **115**, 1115-1121
98. Sanulli, S., Trnka, M. J., Dharmarajan, V., Tibble, R. W., Pascal, B. D., Burlingame, A. L., Griffin, P. R., Gross, J. D., and Narlikar, G. J. (2019) HP1 reshapes nucleosome core to promote heterochromatin phase separation. *Nature*
99. Wang, L., Gao, Y., Zheng, X., Liu, C., Dong, S., Li, R., Zhang, G., Wei, Y., Qu, H., Li, Y., Allis, C. D., Li, G., Li, H., and Li, P. (2019) Histone Modifications Regulate Chromatin Compartmentalization by Contributing to a Phase Separation Mechanism. *Mol Cell*
100. Plys, A. J., Davis, C. P., Kim, J., Rizki, G., Keenen, M. M., Marr, S. K., and Kingston, R. E. (2019) Phase separation of Polycomb-repressive complex 1 is governed by a charged disordered region of CBX2. *Genes Dev* **33**, 799-813
101. Shinsky, S. A., and Cosgrove, M. S. (2015) Unique Role of the WD-40 Repeat Protein 5 (WDR5) Subunit within the Mixed Lineage Leukemia 3 (MLL3) Histone Methyltransferase Complex. *J Biol Chem* **290**, 25819-25833
102. Zhang, P., Lee, H., Brunzelle, J. S., and Couture, J. F. (2012) The plasticity of WDR5 peptide-binding cleft enables the binding of the SET1 family of histone methyltransferases. *Nucleic Acids Res* **40**, 4237-4246
103. Jackson, D. A., Hassan, A. B., Errington, R. J., and Cook, P. R. (1993) Visualization of focal sites of transcription within human nuclei. *EMBO J* **12**, 1059-1065
104. Wansink, D. G., Schul, W., van der Kraan, I., van Steensel, B., van Driel, R., and de Jong, L. (1993) Fluorescent labeling of nascent RNA reveals transcription by RNA polymerase II in domains scattered throughout the nucleus. *The Journal of cell biology* **122**, 283-293
105. Melnik, S., Deng, B., Papantonis, A., Baboo, S., Carr, I. M., and Cook, P. R. (2011) The proteomes of transcription factories containing RNA polymerases I, II or III. *Nat Methods* **8**, 963-968

106. Vernon, R. M., Chong, P. A., Tsang, B., Kim, T. H., Bah, A., Farber, P., Lin, H., and Forman-Kay, J. D. (2018) Pi-Pi contacts are an overlooked protein feature relevant to phase separation. *eLife* **7**
107. Bolognesi, B., Lorenzo Gotor, N., Dhar, R., Cirillo, D., Baldrighi, M., Tartaglia, G. G., and Lehner, B. (2016) A Concentration-Dependent Liquid Phase Separation Can Cause Toxicity upon Increased Protein Expression. *Cell reports* **16**, 222-231
108. Guenther, M. G., Jenner, R. G., Chevalier, B., Nakamura, T., Croce, C. M., Canaani, E., and Young, R. A. (2005) Global and Hox-specific roles for the MLL1 methyltransferase. *Proc Natl Acad Sci U S A* **102**, 8603-8608
109. Milne, T. A., Dou, Y., Martin, M. E., Brock, H. W., Roeder, R. G., and Hess, J. L. (2005) MLL associates specifically with a subset of transcriptionally active target genes. *Proc Natl Acad Sci U S A* **102**, 14765-14770
110. Perham, R. N. (2000) Swinging arms and swinging domains in multifunctional enzymes: catalytic machines for multistep reactions. *Annu Rev Biochem* **69**, 961-1004
111. UniProt, C. (2019) UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res* **47**, D506-D515
112. Tan, S., Kern, R. C., and Selleck, W. (2005) The pST44 polycistronic expression system for producing protein complexes in Escherichia coli. *Protein Expr Purif* **40**, 385-395
113. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S. e., Wilkins, M. R., Appel, R. D., and Bairoch, A. (2005) Protein Identification and Analysis Tools on the ExPASy Server. in *The Proteomics Protocols Handbook*. pp 571-607
114. Laue, T. M., Shah, B., Ridgeway, T. M., and Pelletier, S. L. (1992) Computer-aided interpretation of analytical sedimentation data for proteins. in *Analytical Ultracentrifugation in Biochemistry and Polymer Science* (Harding, S. E., Rowe, A. J., and Horton, L. C. eds.), Royal Society of Chemistry, Cambridge. pp 90-125
115. Brautigam, C. A. (2015) Calculations and Publication-Quality Illustrations for Analytical Ultracentrifugation Data. *Methods Enzymol* **562**, 109-133
116. Chaturvedi, S. K., Ma, J., Zhao, H., and Schuck, P. (2017) Use of fluorescence-detected sedimentation velocity to study high-affinity protein interactions. *Nat Protoc* **12**, 1777-1791
117. Toseland, C. P. (2013) Fluorescent labeling and modification of proteins. *J Chem Biol* **6**, 85-95

Table 1: Summary of apparent dissociation constants for MLL1 core complex assembly at different temperatures*

Temperature (°C)	K_d^{app} (nM)	Confidence interval (1σ)
5	7	6 – 9
10	7	5 – 11
15	20	13 – 30
20	30	22 – 39
25	62	51 – 72
30	290	204 – 417
37	6200	4900 – 7900

* Dissociation constants and error estimates were obtained from fitting $MWRAD_2$ concentration versus signal weight average sedimentation coefficient (s_w) using the A + B ⇋ AB hetero-association model in SEDPHAT (59).

Table 2. Pseudo-first order rate constants for H3K4 monomethylation (k_{me1}) catalyzed by $MWRAD_2$ at the indicated concentration and temperature*

Temperature:	**5°C**	**10°C**	**15°C**	**20°C**	**25°C**	**30°C**	**37°C**
[$MWRAD_2$], μM	k_{me1}, min^{-1}	k_{me1}, min^{-1}	k_{me1}, min^{-1}	k_{me1}, min^{-1}	k_{me1}, min^{-1}	k_{me1}, min^{-1}	k_{me1}, min^{-1}
0.25	0.01 ± 0.01	0.03 ± 0.02	0.03 ± 0.04	0.03 ± 0.03	0.0 ± 0.05	N/A[a]	N/A
0.5	0.06 ± 0.01	0.10 ± 0.02	0.15 ± 0.03	0.18 ± 0.03	0.11 ± 0.05	0.03 ± 0.02	N/A
0.75	0.06 ± 0.01	0.16 ± 0.02	0.16 ± 0.02	0.24 ± 0.02	0.12 ± 0.04	0.07 ± 0.05	0.00 ± 0.02
1.0	0.13 ± 0.01	0.19 ± 0.04	0.25 ± 0.04	0.29 ± 0.04	0.19 ± 0.04	0.18 ± 0.04	0.07 ± N.D.[b]
5.0	0.13 ± 0.02	0.26 ± 0.04	0.31 ± 0.04	0.32 ± 0.04	0.30 ± 0.04	0.28 ± 0.04	0.13 ± 0.06

* Each is the rate constant +/- the Standard Error determined from duplicate measurements.
[a]N/A, Not applicable – no methylation observed under the indicated condition.
[b] N.D., error estimates are not defined.

Table 3. Pseudo-first order rate constants for H3K4 dimethylation (k_{me2}) catalyzed by $MWRAD_2$ at the indicated concentration and temperature*

Temperature:	**5°C**	**10°C**	**15°C**	**20°C**	**25°C**	**30°C**	**37°C**
[$MWRAD_2$], µM	k_{me2}, min^{-1}	k_{me2}, min^{-1}	k_{me2}, min^{-1}	k_{me2}, min^{-1}	k_{me2}, min^{-1}	k_{me2}, min^{-1}	k_{me2}, min^{-1}
0.25	N/A[a]	N/A	N/A	N/A	N/A	N/A	N/A
0.5	0.01 ± 0.00	0.01 ± 0.01	0.02 ± 0.01	0.03 ± 0.01	0.01 ± 0.04	N/A	N/A
0.75	0.01 ± 0.00	0.03 ± 0.01	0.04 ± 0.01	0.07 ± 0.02	0.03 ± 0.02	0.01 ± 0.06	N/A
1.0	0.02 ± 0.00	0.05 ± 0.02	0.06 ± 0.02	0.09 ± 0.03	0.06 ± 0.04	0.04 ± 0.01	N/A
5.0	0.03 ± 0.01	0.08 ± 0.01	0.09 ± 0.01	0.12 ± 0.02	0.12 ± 0.02	0.11 ± 0.02	0.09 ± 0.06

* Each is the rate constant +/- the Standard Error determined from duplicate measurements.
[a]N/A, Not applicable – no methylation observed under the indicated condition.

Table 4. Pseudo-first order rate constants for $MWRAD_2$ irreversible inactivation (k_{inact})

Temperature:	**5°C**	**10°C**	**15°C**	**20°C**	**25°C**	**30°C**	**37°C**
[$MWRAD_2$], µM	k_{inact}, min^{-1}	k_{inact}, min^{-1}	k_{inact}, min^{-1}	k_{inact}, min^{-1}	k_{inact}, min^{-1}	k_{inact}, min^{-1}	k_{inact}, min^{-1}
0.25	0.01 ± 0.02	0.02 ± 0.02	0.04 ± 0.08	0.05 ± 0.05	0.08 ± 1.07	>1700[b]	>2000[b]
0.5	0.01 ± 0.01	0.002 ± 0.01	0.01 ± 0.02	0.02 ± 0.01	0.06 ± 0.04	0.09 ± 0.07	>390[b]
0.75	0.00 ± 0.01	0.009 ± 0.01	0.02 ± 0.01	0.03 ± 0.01	0.04 ± 0.03	0.09 ± 0.08	0.13 ± 1.00
1.0	0.00 ± N.D.[a]	0.01 ± 0.02	0.01 ± 0.01	0.04 ± 0.02	0.04 ± 0.03	0.07 ± 0.03	0.24 ± 0.11
5.0	0.01 ± 0.01	0.05 ± 0.02	0.00 ± 0.00	0.09 ± 0.02	0.07 ± 0.03	0.07 ± 0.02	0.21 ± 0.13

* Each is the rate constant +/- the Standard Error determined from duplicate measurements.
[a] N.D., error estimates are not defined.
[b] k_{inact} lower bound. In Kintek Explorer software, k_{me1} was fixed to the value predicted by the Arrhenius equation at the indicated temperature and k_{inact} was floated to estimate the lower bound required for the observed loss of activity.

Table 5. Pseudo-first order rate constants for 5 μM $MWRAD_2$ at 37°C in high (200mM) and low (25mM) NaCl reaction buffer*

[$MWRAD_2$], μM	k_1 (min^{-1})		k_2 (min^{-1})		k_3 (min^{-1})		k_{inact} (min^{-1})	
	[NaCl], mM		[NaCl], mM		[NaCl], mM		[NaCl], mM	
	200	*25*	*200*	*25*	*200*	*25*	*200*	*25*
0.25	N/A[a]	0.3 (0.01)	N/A	N/A	N/A	N/A	N/A	0.00 (N.D.)
0.5	N/A	1.4 (0.10)	N/A	1.1 (0.20)	N/A	N/A	N/A	0.00 (N.D.)
0.75	0.00 (0.02)[c]	1.5 (0.10)	N/A	1.2 (0.09)	N/A	N/A	0.13 (1.00)	0.00 (N.D.)
1.0	0.07 (N.D.[b])	2.3 (0.20)	N/A	1.9 (0.10)	N/A	0.01 (0.01)	0.24 (0.11)	0.00 (N.D.)
5.0	0.13 (0.06)	8.1 (1.80)	0.09 (0.06)	4.5 (0.30)	N/A	0.03 (N.D.)	0.21 (0.13)	0.00 (N.D.)

* Each is the rate constant with Standard Error determined from duplicate measurements in parentheses.
[a]N/A, Not applicable – no methylation observed.
[b]N.D., error estimates are not defined.
[c] rates below 5×10^{-3} were rounded to 0.00.

Figure 1 – Namitz, Tan and Cosgrove

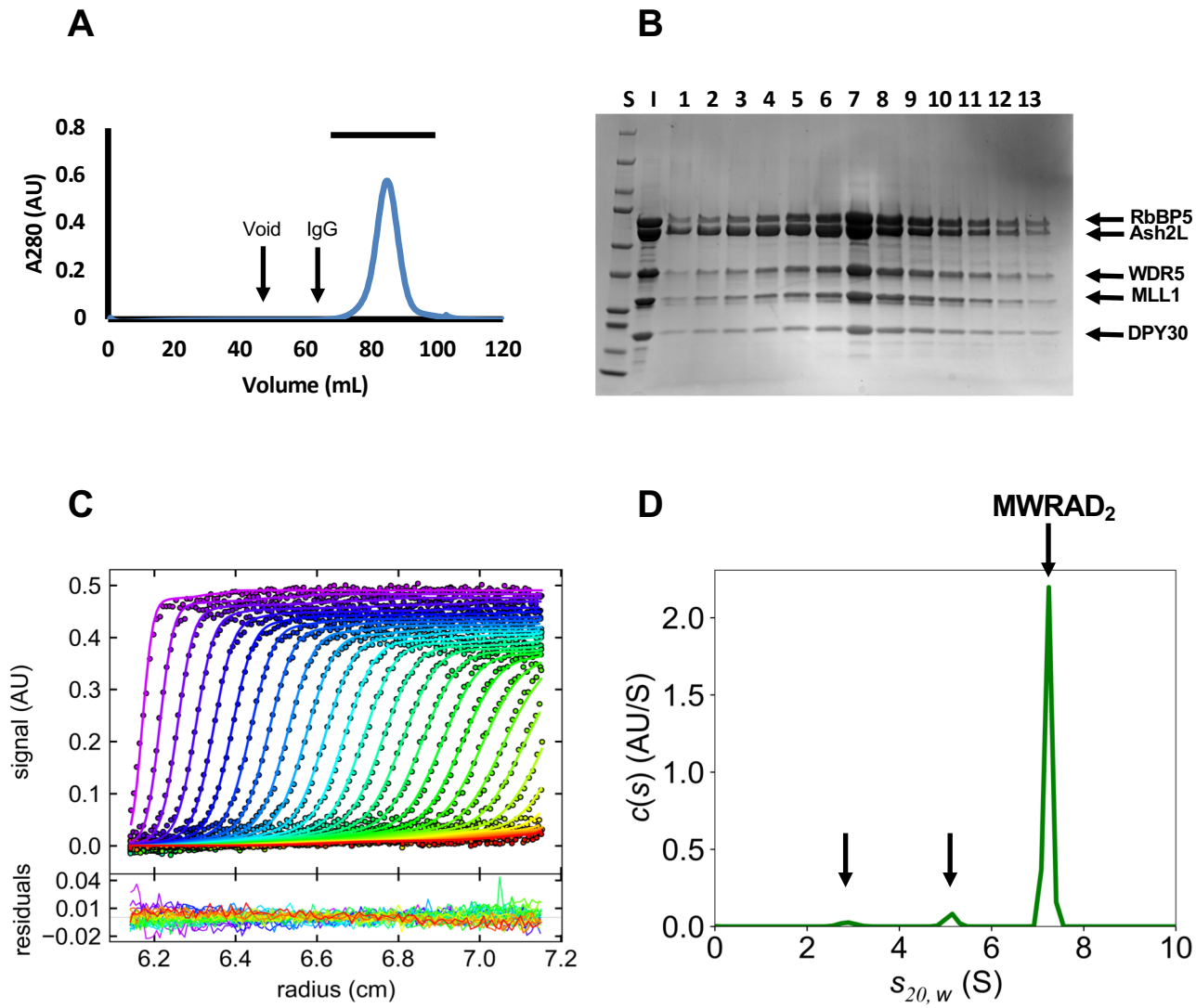


Figure 1: Purification and characterization of the MLL1 core complex. (A) Chromatogram of S200 SEC purified $MWRAD_2$. The void volume and elution volume of IgG (M_r 158 kDa) are indicated. The horizontal bar above the peak indicates fractions shown on the Coomassie stained SDS-PAGE gel in (B). (C) Upper Panel: SV-AUC run showing raw data (points) and fits using the continuous distribution (*c(s)*) method by the program SEDFIT (solid lines) (56). The lower panel shows the residuals derived from the fit. Shown is a typical run of 5 μM $MWRAD_2$ taken at 5°C. (D) Diffusion-deconvolved sedimentation coefficient distribution (*c(s)*) obtained using the fits to the raw data shown in (C). All profiles are shown with experimental *s** values corrected to standard conditions at 20°C in water ($s_{20,w}$ (*S*)). The positions of $MWRAD_2$ and the two minor peaks are indicated with arrows.

Figure 2 – Namitz, Tan and Cosgrove

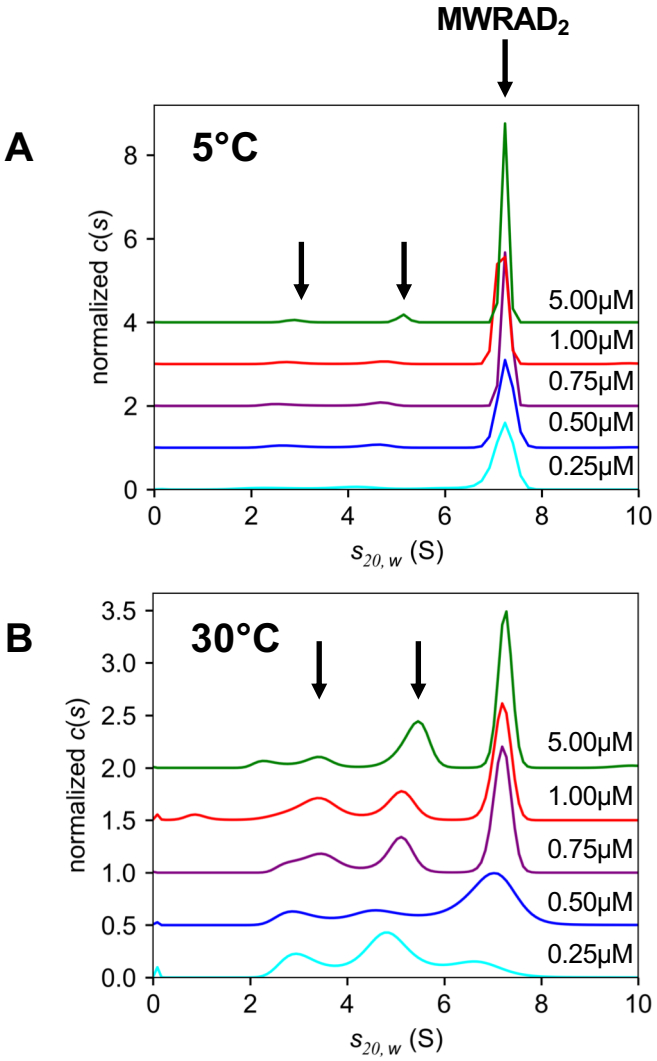


Figure 2: The holo-MLL1 core complex assembles from predominantly two sub-complexes. (A) *c(s)* distributions of $MWRAD_2$ at 5°C at five different concentrations: 0.25 μM (cyan), 0.5 μM (blue), 0.75 μM (purple), 1.0 μM (red) and 5.0 μM (green). Each profile was normalized by total integrated area under the peaks. (B) The same as in (A), but at 30°C. The unlabeled arrows in both (A) and (B) indicate the positions of the putative sub-complex peaks at the highest concentration.

Figure 3 – Namitz, Tan and Cosgrove

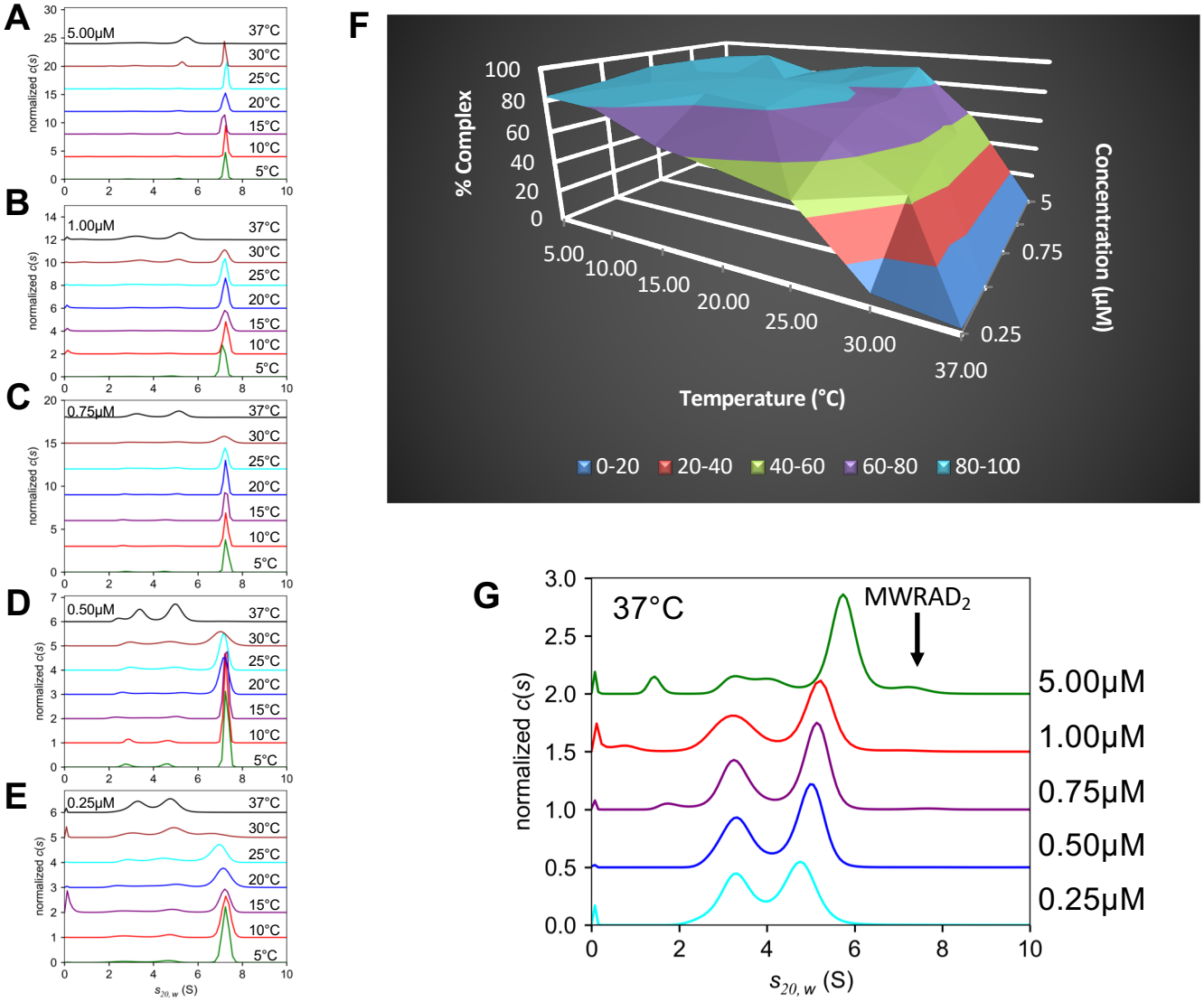
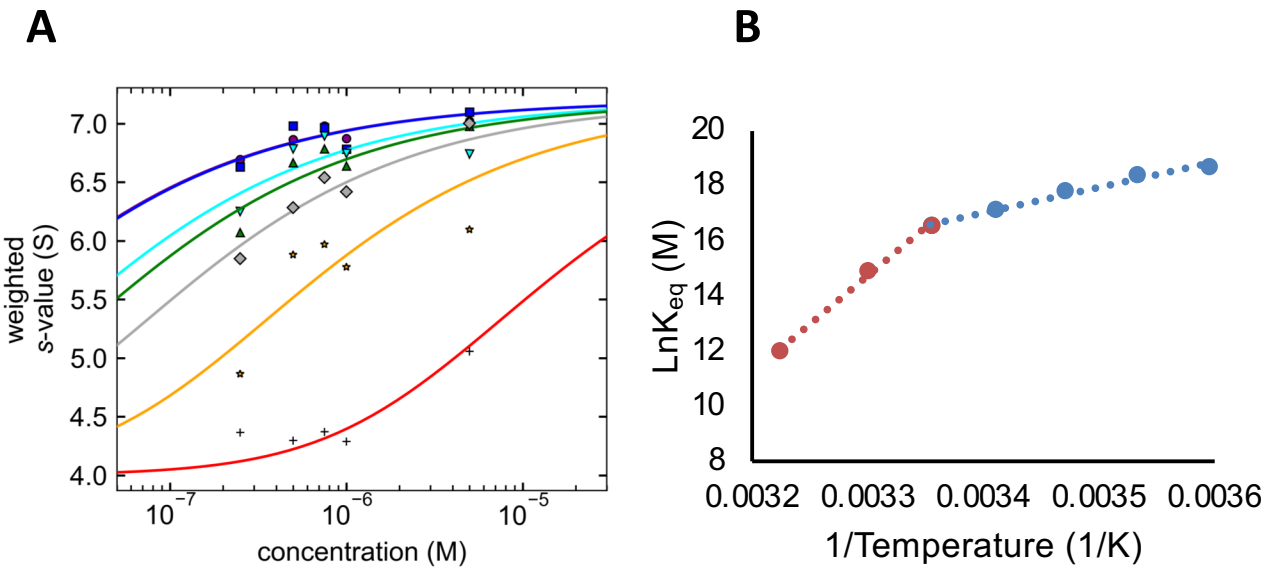


Figure 3: Temperature dependence of MLL1 core complex assembly. (A-E) Representative *c(s)* distributions of the MLL1 core complex at the indicated temperatures and loading concentrations. Each distribution was normalized for total integrated area. (F) Surface plot summarizing the percentage of signal in SV-AUC experiments corresponding to the *S* value of the MLL1 core complex as a function of temperature and concentration (see also Table S1). These values were obtained as described in Methods. (G) *c(s)* distributions from five $MWRAD_2$ concentrations at 37°C normalized by total integrated area (note: each distribution corresponds to the black line from the respective concentration panel in A – E). The position of holo-$MWRAD_2$ at 7.2 *S* is indicated with the arrow.

Figure 4 – Namitz, Tan and Cosgrove



C

Temp range	ΔG (25°C) (kcal/mol K)	ΔH (25°C) (kcal/mol K)	ΔS (25°C) (kcal/mol K)	ΔH/ΔS (25°C)
Low (5-25°C)	-9.8	-18	- 0.029	634
High (25-37°C)	-9.9	-70	-0.202	347

Figure 4: Thermodynamic characterization of MLL1 core complex assembly. (A) Signal-weighted (s_w) isotherms of MWRAD$_2$ were obtained for each temperature, plotted against loading concentrations and fit to an A + B ⇋ AB hetero-association model using SEDPHAT (114). The lines represent the fits for each isotherm, which were conducted at 5°C (blue),10°C (purple), 15°C (cyan), 20°C (green), 25°C (grey), 30°C (orange) and 37°C (red). K_d^{app} values are summarized in Table 1. (B) van't Hoff plot derived from the apparent K_{eq} values. Linear regression was used to independently fit the data for the high temperature range (red, 25-37°C) and low temperature range (blue, 5-25°C). (C) Summary of thermodynamic parameters for MLL1 core complex assembly under high and low temperature regimes derived from the van't Hoff analysis in (B).

Figure 5 – Namitz, Tan and Cosgrove

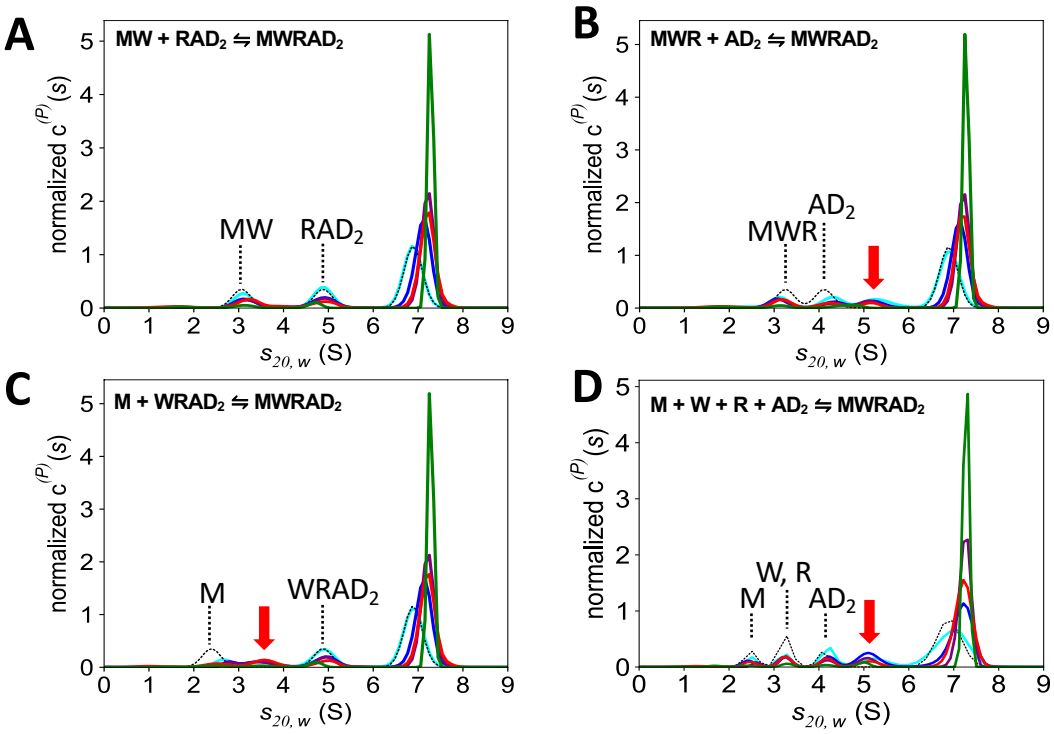


Figure 5: The holo-MLL1 core complex assembles from MW and RAD_2 sub-complexes. Bayesian analysis of $MWRAD_2$ SV-AUC data collected at 25°C. $MWRAD_2$ concentrations were 0.25 μM (cyan), 0.5 μM (blue), 0.75 μM (purple), 1.0 μM (red) and 5.0 μM (green). Maximum entropy regularizations were restrained with expected values (indicated with the dotted line) for each indicated sub-complex derived from separate experiments (Fig. S1 and Table S2) to give $c^{(P)}(s)$ distributions (colored lines), which were normalized by total integrated area. Concentrations of $MWRAD_2$ in each run were: 0.25 μM (cyan), 0.5 μM (blue), 0.75 μM (purple), 1.0 μM (red) and 5.0 μM (green). The $c^{(P)}(s)$ distributions used the following S values as prior expectations: (A) MW (3.0 S), RAD_2 (4.4 S), and $MWRAD_2$ (6.9 S) (scheme 1); (B) MWR (3.3 S), AD_2 (4.1 S), and $MWRAD_2$ (6.9 S) (scheme 2). (C) M (2.3 S), $WRAD_2$ (4.4 S), and $MWRAD_2$ (6.9 S) (scheme 3). (D) M(2.3 S), W(3.2 S), R(3.4 S), AD_2(4.1 S), and $MWRAD_2$ (6.9 S) (concerted assembly scheme).

Figure 6 – Namitz, Tan and Cosgrove

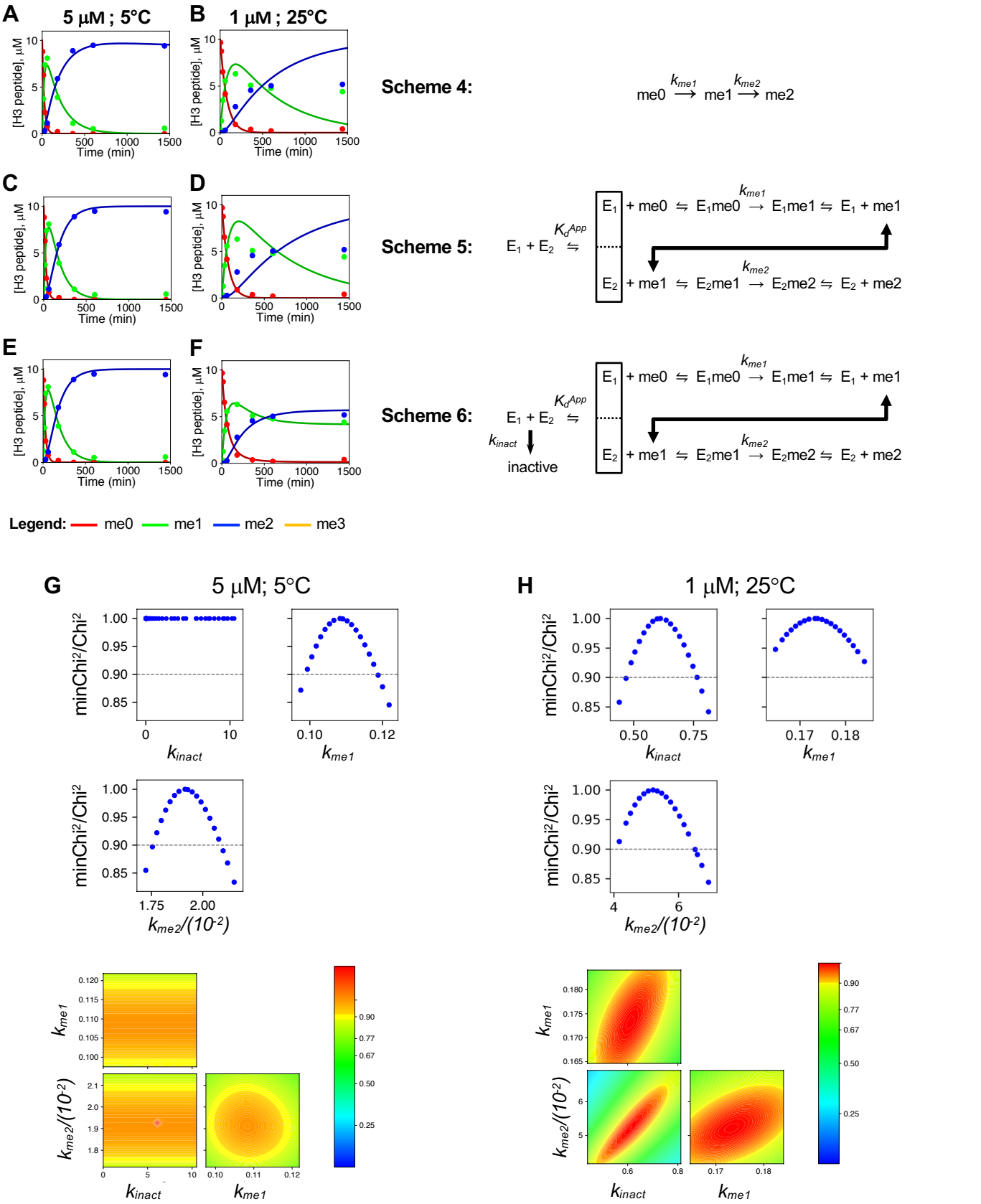


Figure 6: Comparison of minimal reaction pathways. (A, C, E) show the fits (solid lines) for the concentrations of each peptide species (me0, me1, or me2) during the reaction time course catalyzed by 5 μM $MWRAD_2$ at 5°C. Each panel shows the fits to the same data using Scheme 4 (A), Scheme 5 (C), or Scheme 6 (E). Panels (B, D, F) show fits for Schemes 4-6, respectively, for the reaction time course catalyzed by 1 μM $MWRAD_2$ at 25°C. (G) Fitspace confidence contour analysis for the reaction catalyzed by 5 μM $MWRAD_2$ at 5°C fit with Scheme 6. k_{inact} is not constrained by the data, mainly due to the absence of detectable enzyme inactivation during the reaction time course at 5°C. (H) Fit space confidence contour analysis of the fit of Scheme 6 to the reaction catalyzed by 1 μM $MWRAD_2$ at 25°C. k_{inact} is now constrained by the data.

Figure 7 – Namitz, Tan and Cosgrove

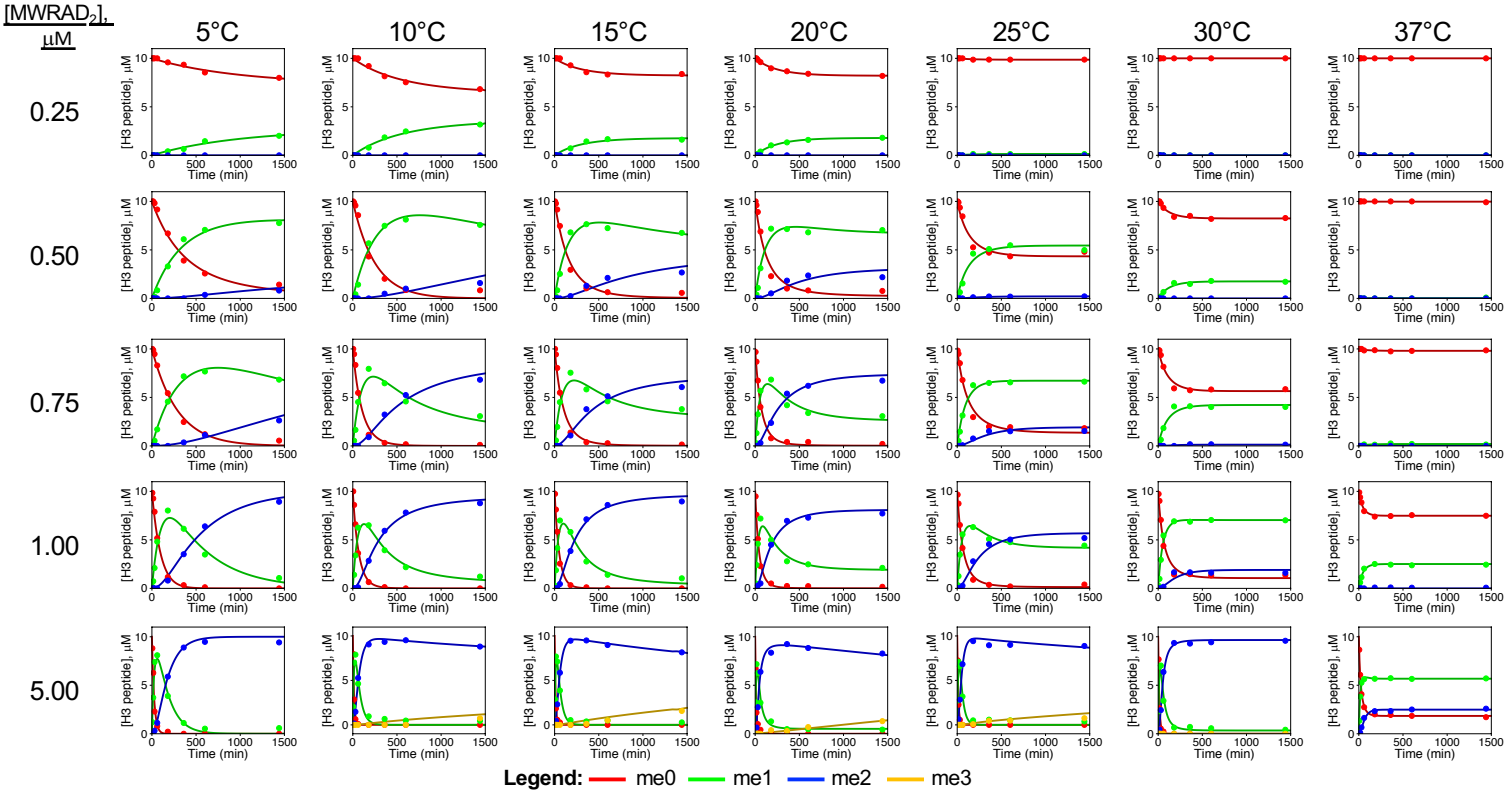


Figure 7: Temperature and concentration dependence of MLL1 core complex enzymatic activity. Time courses for reactions at the indicated $MWRAD_2$ concentrations and temperatures were plotted and fit using Scheme 6. Each time point represents the average from two independent experiments. Concentrations of each peptide species were plotted in red for H3K4me0, green for H3K4me1, blue for H3K4me2. For reactions showing small amounts of H3K4me3 (yellow), Scheme 6 was modified to incorporate an additional turnover step followed by product release.

Figure 8 – Namitz, Tan and Cosgrove

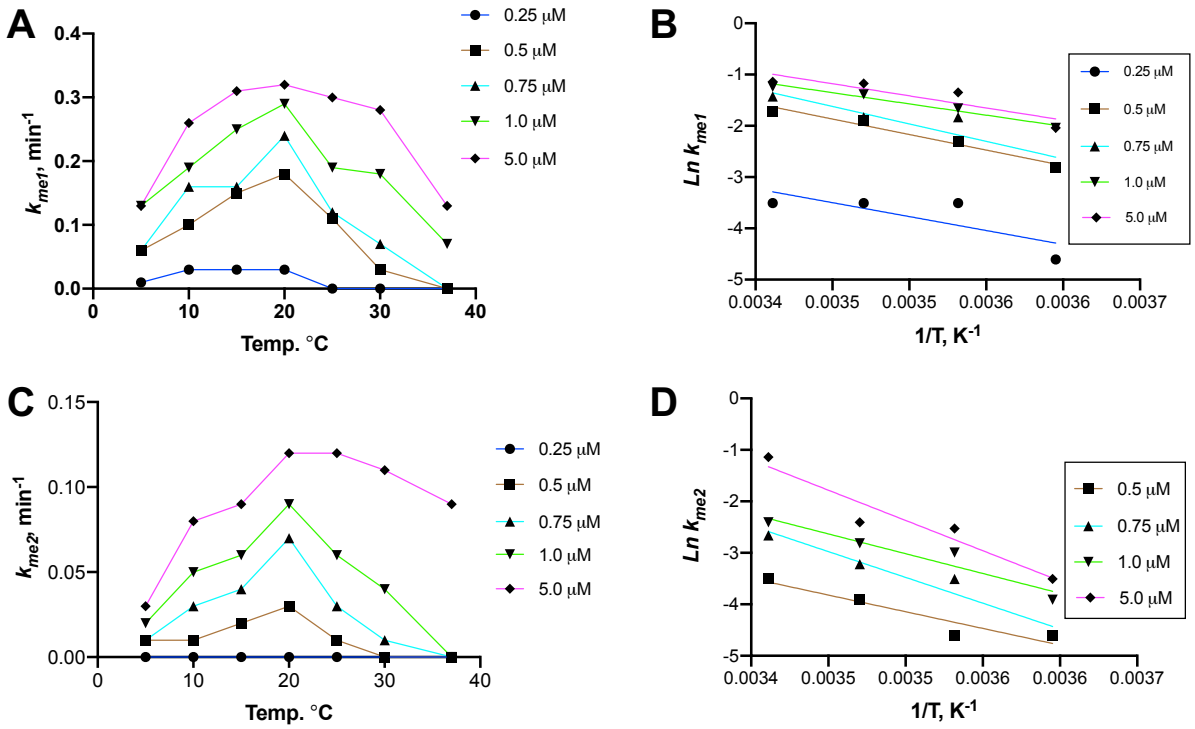


Figure 8: Effect of temperature on MWRAD$_2$ enzymatic activity. (A and C), rates of H3K4 mono- (A) and dimethylation (C) plotted as a function of temperature. Arrhenius behavior (defined as a doubling of the rate for every 10°C increase in temperature) was observed between 5°C and 20°C for most concentrations. (B and D), Arrhenius plots for H3K4 mono- (B) and dimethylation (D) for the data collected between 5°C and 20°C. The lines represent linear regression fits to the data collected at the indicated MWRAD$_2$ concentrations. E_a values were obtained from the slope of the Arrhenius fits, where $slope = -(E_a/R)$ at each enzyme concentration.

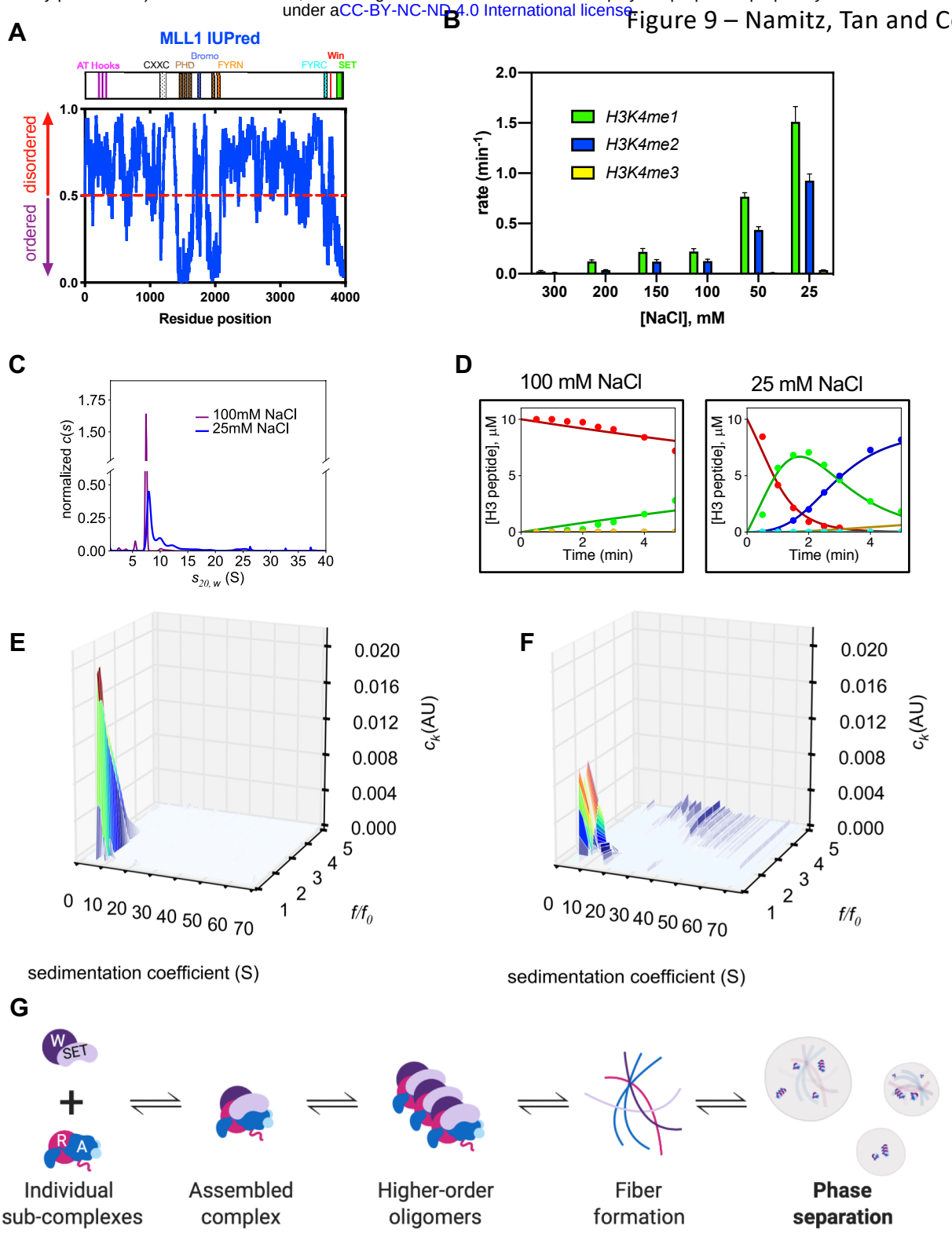


Figure 9: MLL1 core complex enzymatic activity is increased under conditions that induce phase separation. (A) IUPred disorder prediction (71) for the full-length MLL1 protein. Uniprot sub-domain boundaries are shown in the schematic above and are summarized in Table S6. (B) Comparison of 5 μM MLL1 core complex enzymatic activity at different ionic strengths at 25°C. (C) SV-AUC comparison of 5.0 μM $MWRAD_2$ *c(s)* distributions at 100 mM (purple) and 25 mM (blue) NaCl. (D) Reaction time courses of 5.0 μM $MWRAD_2$ at 100 mM NaCl (left panel) and 25 mM NaCl (right panel) at 25°C. Each time point represents the mean concentration of each peptide species, and solid lines show the fit using Scheme 6. Peptide species were H3K4me0 (red), H3K4me1 (green), H3K4me2 (blue), and H3K4me3 (yellow). (E and F) Size and shape analyses ($c(s,f_r)$) from SV-AUC runs of 5.0 μM $MWRAD_2$ in buffer with either 100 mM (E) or 25 mM (F) NaCl, each at 25°C. (G) A schematic of the build-up of higher-order oligomers and subsequent fiber formation preceding phase separation (Created with BioRender.com).

Figure 10 – Namitz, Tan and Cosgrove

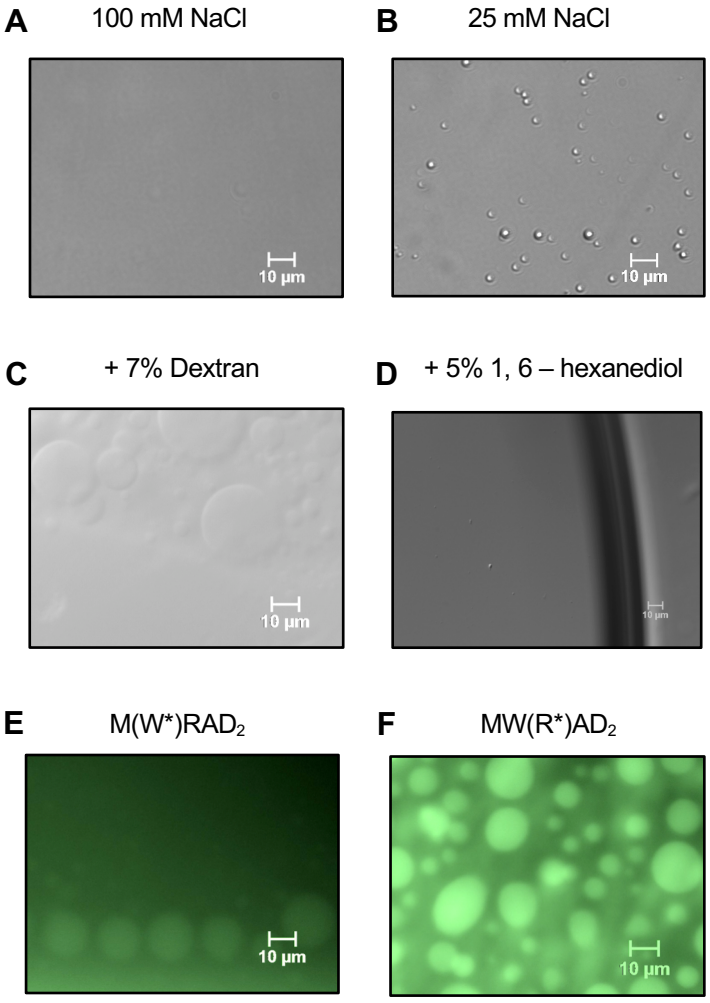


Figure 10: The MLL1 core complex phase separates at a low concentration and physiological ionic strength. (A,B) DIC microscopy images of MLL1 core complex enzymatic reactions at 100 mM (A) or 25 mM (B) NaCl. Each reaction contained 5.0 μM $MWRAD_2$, 100 μM $H3^{1-20}$ peptide and 250 μM AdoMet in reaction buffer at 25°C. (C) The same as in (B) but with 7% dextran (see also Supplementary movie S1). (D) Same as in (C) but with 5% 1,6 hexanediol. (E,F) Fluorescence microscopy images of the MLL1 core complex assembled with AlexaFluor 488-labeled WDR5 (E) or RbBP5 (F) subunits (see also supplementary movies S3 and S4). The conditions were 5.0 μM gel filtration-purified complex (see supplementary Fig.S7) in reaction buffer with 10 μM $H3^{1-20}$ peptide, 250 μM AdoMet, and 150 mM NaCl.

Figure 11 – Namitz, Tan and Cosgrove

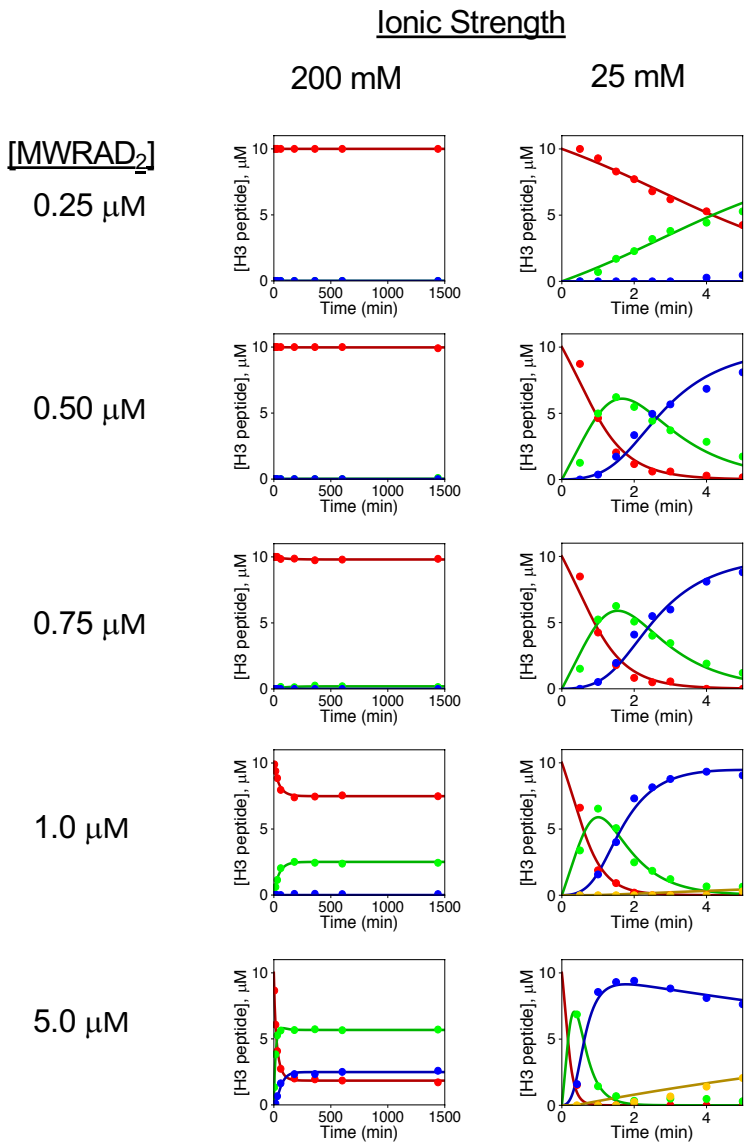


Figure 11: Enzymatic activity of the MLL1 core complex at physiological temperature under phase separation conditions. Comparison of MLL1 core complex enzymatic activity at the indicated concentrations at 37°C in high (200 mM NaCl) vs. low (25 mM NaCl) ionic strength reaction buffers. The 200mM NaCl panels (left) from Fig. 7 are shown again here for the purpose of comparison. Each time point represents the mean concentration of each peptide species and solid lines show the fit using Scheme 6 (Fig.6) or a modified form of Scheme 6 to account for trimethylation. The resulting pseudo-first order rate constants are summarized in Table 5. Peptide species were H3K4me0 (red), H3K4me1 (green), H3K4me2 (blue), and H3K4me3 (yellow). Note the time scale differences required for the high vs. low ionic strength reactions.

Figure 12 – Namitz, Tan and Cosgrove

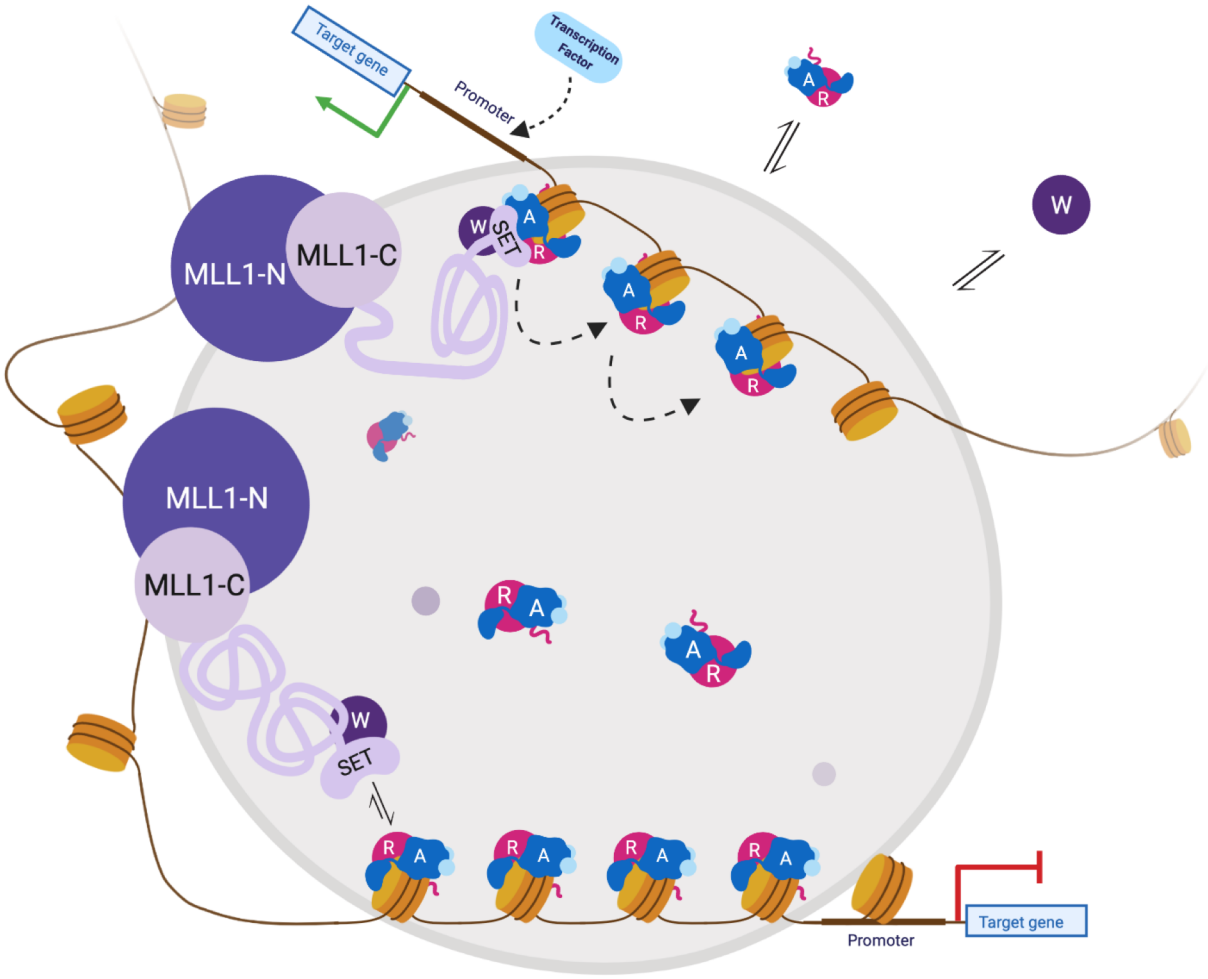


Figure 12: Swinging Domain Model for regulation of MLL1 core complex assembly and enzymatic activity within a transcription factory. MLL1 N-terminal region (MLL1-N) binds to DNA in or near a transcription factory using its DNA and chromatin-recognition domains. The C-terminal region (MLL1-C), which contains the SET domain, binds to WDR5 (W) to create the MW sub-complex, either before or after entry into the factory. The RbBP5, Ash2L, DPY-30 (RAD_2) sub-complex binds nucleosomes. MW then uses a "swinging domain" mechanism within the phase condensate to move the SET domain-WDR5 around the transcription factory, where the high local concentration forces the assembly of the full MLL1 core complex and allows for H3K4 methylation of nucleosomes within the factory that have RAD_2 already bound. This can be repeated multiple times within the factory, resulting in extensive H3K4 methylation of nucleosomes that go into the condensate. This methylation results in removal of nucleosomes and recruitment of transcription factors that, in turn, recruit RNA Polymerase II for transcription initiation. Once the chromatin leaves the factory, the reduction in local concentration results in the loss of the RAD_2 sub-complex, as well as a high kinetic barrier to reassembly of $MWRAD_2$, preventing any ectopic methylation. This figure was created with BioRender.com