The Set1 N-terminal domain and Swd2 interact with RNA polymerase II CTD to recruit COMPASS

Hyun Jin Bae1,6, Marion Dubarry2,6, Jongcheol Jeon1,3, Luis M. Soares1,5, Catherine Dargemont4, Jaehoon Kim3, Vincent Geli2 & Stephen Buratowski1

Methylation of histone H3 lysine 4 (H3K4) by Set1/COMPASS occurs co-transcriptionally, and is important for gene regulation. Set1/COMPASS associates with the RNA polymerase II C-terminal domain (CTD) to establish proper levels and distribution of H3K4 methylations. However, details of CTD association remain unclear. Here we report that the Set1 N-terminal region and the COMPASS subunit Swd2, which interact with each other, are both needed for efficient CTD binding in *Saccharomyces cerevisiae*. Moreover, a single point mutation in Swd2 that affects its interaction with Set1 also impairs COMPASS recruitment to chromatin and H3K4 methylation. A CTD interaction domain (CID) from the protein Nrd1 can partially substitute for the Set1 N-terminal region to restore CTD interactions and histone methylation. However, even when Set1/COMPASS is recruited via the Nrd1 CID, histone H2B ubiquitylation is still required for efficient H3K4 methylation, indicating that H2Bub acts after the initial recruitment of COMPASS to chromatin.
Eukaryotic gene transcription is regulated by posttranslational modifications of histone tails, which include phosphorylation, acetylation, ubiquitylation, and methylation. Specific methylation of histone lysine residues correlates with transcription activity or repression. For instance, methylation on H3K4, K36, and K79 are enriched over genes actively transcribed by RNA polymerase II (RNAPII), whereas H3K9 and H3K27 methylations are highest in transcriptionally inactive regions. Methylation of H3K4 (H3K4me) has aroused particular interest. In mammals, this mark can be deposited by multiple complexes (Setd1a, Setd1b, Mll1, Mll2, Mll3, and Mll4), all of which share a module comprising WDR5, RbBP5, ASH2L, and Dpy-30 (called WRAD) that associates with the catalytic SET domain. Each module is endowed with additional proteins that determine their recruitment and biological functions.

In budding yeast, all H3K4 methylation is catalyzed by a single Set1 complex (called COMPASS) in which Set1 acts as a scaffold for seven subunits (Bre2, Sdc1, Shg1, Spp1, Swd1, Swd2, and Swd3). Of these, Swd2 is the only subunit that is essential for viability. However, this requirement stems from its additional role as a component of the RNA 3′ end processing and termination complex, APT (Associated with Paf1)10,11. It is unclear whether Swd2 plays the same role in both APT and COMPASS. Protein interaction studies have shown that the Set1 SET domain associates with the WRAD homologs Swd1, Swd3, Bre2, and Sdc1, the N-SET domain associates with Spp1, the Set1 central region binds Shg1, and the Set1 N-terminal region contacts Swd2,12,12-15. Swd1 also shows some interactions with the Set1 N-terminal region14,16. This organization of COMPASS structure was confirmed by cryo-electron microscopy17-19.

Control of COMPASS activity involves a complex set of interactions. Remarkably, higher level H3K4 methylation by Set1 requires prior ubiquitylation on histone H2B (H2Bub)20,21. Set1 contact with Swd1/Swd3 is crucial for H2Bub-dependent H3K4 methylation16. Deletion or depletion of individual COMPASS subunits differentially impairs Set1 stability and the pattern of H3K4 methylation along active genes12,14,22-24. Particularly relevant to this study, depletion of the WD40 repeat protein Swd2 strongly destabilizes Set1 and reduces H3K4 methylation7,10,25,26. Set1 activity is positively regulated by the Set1 double RNA recognition motif (dRRM), but inhibited by a centrally located auto-inhibitory domain7,27,28. Set1 dRRM binding to nascent RNA may affect COMPASS distribution along transcription units and subsequent deposition of the H3K4me3 mark29. Surprisingly, combined deletion of the N-terminal, dRRM, and central domains leads to overexpression of a truncated Set1 protein with mistargeted H3K4 methylation30,31. Finally, mutations in H3K4 lead to Set1 degradation, indicating a feedback mechanism to control enzyme levels30.

The C-terminal domain (CTD) of Rpb1, the largest subunit of RNAPII, consists of multiple Tyr1-Ser2-Pro3-Thr4-Ser5-Pro6-Ser7 (YSPTSPS) heptad repeats32. Chromatin immunoprecipitation (ChIP) with antibodies against individual phosphorylated CTD residues shows that phosphorylation of Serine 5 (Ser5P) peaks near the promoter, whereas Serine 2 phosphorylation (Ser2P) increases later during elongation33. Set1/COMPASS co-transcriptionally associates with Ser5-phosphorylated RNAPII34 to produce a gradient of H3K4 methylation that begins at the +1 nucleosome and tails off with distance from the promoter12. This 5′ targeting is compounded over multiple transcription cycles35, leading to the canonical peak of H3K4me3 near the promoter, followed by H3K4 dimethylation (H3K4me2) and monomethylation (H3K4me1) further downstream.

Binding of COMPASS to the RNAPII CTD has been proposed to rely on the Pafl complex36,37. However, direct interactions of COMPASS with either Pafl complex or RNAPII have yet to be characterized. The recruitment of COMPASS to transcribed genes was also proposed to rely on interaction between Swd2 and H2Bub23. However, this model is not supported by in vitro reconstitution experiments showing that COMPASS lacking Swd2 still methylates H3K4 in an H2B ubiquitylation-dependent manner14, Therefore, it remains unclear how COMPASS interacts with RNAPII, and whether other proteins participate in the interaction.

Here, we provide biochemical and yeast two-hybrid evidence showing that the Set1 N-terminal region and Swd2 together mediate interaction with the Rpb1 CTD. Deletion of the first 200 amino acids of Set1 results in both loss of RNAPII CTD binding and reduction of H3K4me2 and H3K4me3 levels. These effects can be partially reversed by replacing this region of Set1 with the Nrd1 CTD-interacting domain (CID), which specifically binds Ser5P CTD38. The role of Swd2 in COMPASS recruitment is further supported by the effects of a point mutation in Swd2 that compromises its interaction with Set1, the recruitment of the complex to chromatin, and H3K4 methylation. Finally, when COMPASS is recruited by the Nrd1 “hysteresis” mechanism, the Pafl complex and H2Bub are still required for H3K4 methylation, suggesting that H2Bub acts downstream of initial COMPASS recruitment to elongation complexes.

Results

The Set1 N-terminal domain mediates RNAPII association. Our previous analysis of Set1 N-terminal truncations found that deletion of the first 200 amino acids of Set1 strongly reduced both H3K4me2 and me3. This region of Set1 interacts with the Swd2 subunit of COMPASS14. Interestingly, the mammalian Swd2 homolog Wdr82 preferentially binds Ser5P-CTD in vitro, and has thus been proposed to mediate Setd1A complex recruitment to elongating RNAPII15.

To test if the drop in H3K4me levels upon Set1 N-terminal truncation correlates with reduced association of COMPASS with Ser5P-CTD, a strain lacking the SET1 gene (set1Δ) was transformed with plasmids expressing either full-length Set1 or a truncation lacking 100 or 200 N-terminal amino acids (Set1Δ100 and SA200, Fig. 1a). Consistent with our previous results30, Set1 and SA100 supported H3K4me3 and H3K4me2, but SA200 did not (Fig. 1b). To monitor binding to Ser5P-CTD, the epitope-tagged Set1 proteins were immunoprecipitated using anti-FLAG conjugated beads (a-FLAG), followed by immunoblotting for Ser5P-CTD. As shown in Fig. 1c, N-terminal deletion of Set1 (SA200) attenuated Set1 association with Ser5P-CTD and total RNAPII. This result suggests that the N-terminal domain of Set1 (aa 1-200) that interacts with Swd2 is also important for the direct or indirect association of Set1 with RNAPII.

The N-terminal region of Set1 interacts with the Rpb1 CTD. Evidence for a direct interaction between COMPASS and the CTD of RNAPII subunit Rpb1 has been lacking. A genome-wide yeast two-hybrid (Y2H) screen was performed using full length Set1 fused to the Gal4 DNA binding domain (BD) as bait. Remarkably, a 14-CTD repeat fragment fused to the Gal4 activation domain (AD) was isolated as a robust interactor. The 14-CTD repeat construct was then tested with each of the other COMPASS subunits, but Set1 was the only subunit that showed a positive Y2H interaction (Supplementary Fig. 1a). To further map the CTD interaction, Set1 was divided into five fragments encompassing previously reported functional domains (Fig. 2a)14. Immunoblotting with antibodies against Gal4 BD and Set1 confirmed fusion protein expression (Supplementary Fig. 1b, c). Y2H analysis showed that full-length Set1, the F1 (aa 1–236), and Set1 (1–200) fusions activated the Gal4-driven HIS3 and ADE2
Fig. 1 The Set1 N-terminal domain is needed for RNApII interaction and proper H3K4 methylation. a Schematic of Set1 constructs used, with numbers below diagram indicating the amino acid residues from the wild-type proteins. A FLAG-tag at the N-terminus of all constructs is not shown. b Full length Set1 (Set1) and N-terminal truncated mutants (SΔ100, SΔ200) were transformed into set1Δ cells. Whole cell lysates were separated by SDS-PAGE and analyzed by immunoblotting using the indicated antibodies. Lysates from cells transformed with empty vector plasmid (set1Δ) served as a negative control. TATA-binding protein (TBP) and histone H3 were used as loading controls. c FLAG-tagged full-length (Set1) or Set1Δ200 (SΔ200) were expressed in set1Δ cells. Set1 proteins were immunoprecipitated with anti-FLAG beads (IP:FLAG) from whole cell lysates and analyzed by immunoblotting using antibodies for Ser5P (SSP CTD: 3E8) or total Rpb1 (Pol II: BWG16). Immunoprecipitations from cells transformed with empty vector plasmid (set1Δ) served as a negative control. Bottom panels show input samples, and upper panels show proteins bound to FLAG beads after precipitation. Source data are provided as a Source data file.

Fig. 2 The Set1 N-terminal region binds the Rpb1 CTD. a Schematic of Set1 fragments fused to GAL4 BD for yeast two hybrid assay. b GAL4 BD–Set1 fusions were co-expressed with GAL4 AD–CTD in Y2H strain PJ69-4A, which has both HIS3 and ADE2 reporters for Gal4 activation. Plasmid expressing only the GAL4 AD was used as a negative control. Cells were grown for 2 days on synthetic complete (SC) media plates lacking the indicated amino acids. Five or 20 mM of 3-aminotriazol (3AT) was added to media as indicated to increase H3K4 methylation. Add 20 mM 3AT to the media selects for higher levels or His3. n/a indicates that the genotype was not tested. -Ade, -His plates to monitor the reporter. Additionally, empty vector plasmid (SET1) or Set1Δ100 served as a negative control. However, Set1Δ100 and introduced into cells, the fusion protein (SSP CTD) was incubated with purified RNApII. Bottom panels show input samples, and upper panels show proteins bound to FLAG beads after precipitation. Source data are provided as a Source data file.

Therefore, we conclude that the Set1 N-terminal region is required for RNApII binding.

Swd2 contributes to interaction between COMPASS and RNApII CTD. Wdr82, the mammalian homolog of Swd2, can directly bind Ser5P–CTD in vitro, but also binds the N-terminal region of mammalian Setd1A. We therefore considered the possibility that the amino terminal domain of Set1 interacts with the CTD only indirectly via Swd2. However, an Swd2–CTD interaction was not seen by Y2H (Supplementary Fig. 1a), arguing against this model. Furthermore, when Swd2 was directly fused to SΔ200 and introduced into cells, the fusion protein (SSΔ200) did not restore H3K4 methylation (Supplementary Fig. 2a).

One possible explanation for these results is that CTD actually directly binds Ser5P–CTD in vitro, but also binds the N-terminal region of mammalian Setd1A. Therefore, we conclude that the Set1 N-terminal region is required for RNApII binding.
a CTD-binding conformation in the other. To test this possibility, the Y2H interaction between the Gal4 AD–CTD fusion and Gal4 BD fused to full-length Set1 or segment F1 was compared in isogenic reporter strains that were either wild-type SWD2 or swdΔ2. The lethality of swdΔ2 was suppressed, as previously described, by expression of a fragment from the termination factor Sen1 (amino acids 1890–2092, annotated as Sen1(202))25. The control SWD2 strain was also transformed with the same Sen1 construct to maintain isogenicity. Immunoblotting once again showed that Gal4 fused to Set1 F1 is expressed at higher levels than the full-length Set1 fusion, but also that swdΔ2 does not affect amounts of either protein (Supplementary Fig. 2b). In the Y2H assay, activation of the Gal4-responsive HIS3 reporter was severely reduced, but not abolished by swdΔ2 (Fig. 3). In agreement, swdΔ2 strongly reduces H3K4me3 in cells with wild-type Set1, but not in cells with the Set1 N-terminal truncation (SA200) (Supplementary Fig. 2c). These results suggest that Swd2, while not absolutely required, strongly promotes proper interaction between the Set1 and CTD fusion proteins.

Additional deletions more precisely mapped Swd2 binding to aa 124–229 of Set116. In agreement with this finding, deleting amino acids 200–210 from the Set1-F1 fragment (F1Δ200–210) weakened its yeast two-hybrid interaction with both Swd2 and CTD Gal4 AD fusions (Supplementary Fig. 3a). In the context of full length Set1, Δ200–210 only slightly affected Set1 protein levels, but H3K4me3 levels (Supplementary Fig. 3b) and ChIP of Set1 to the 5′ region of PMAL1 gene were reduced (Supplementary Fig. 3c). These results further show that Swd2 contributes to the COMPASS–RNApII interaction.

The Nrd1 CID can substitute for the Set1 N-terminal domain. If the primary role of the Set1 N-terminal region and associated Swd2 subunit is to recruit COMPASS to the elongation complex via CTD binding, it should be possible to replace this domain with another Ser5P-binding protein. The CID from Nrd1 targets the snoRNA termination machinery to 5′ ends by direct binding to Ser5P–CTD38. Accordingly, we created a fusion where the CID replaces the N-terminal 200aa of Set1 (Fig. 4a). Strikingly, the Nrd1(CID)–Δ200 fusion (NSΔ200) partially rescued both bulk H3K4me2 and me3 (Fig. 4b), as well as the ability to co-precipitate Ser5-phosphorylated RNApII (Fig. 4c). To prove that restoration of H3K4 methylation patterns by the Nrd1 CID fusion is mediated by CTD binding, we created two Set1 fusions with Nrd1 CID point mutants (D70R or I130R) known to disrupt the CTD interaction38. Co-immunoprecipitation experiments confirmed that D70RSA200 and I130RSA200 fusion proteins were expressed normally, but had lost the ability to stably bind RNApII (Fig. 4c). Neither mutant restored H3K4 methylation relative to SA200, as expected if CTD binding is essential for proper COMPASS targeting (Fig. 4b).

To determine if this CID-mediated rescue of H3K4 methylation reflects normal or aberrant distribution along genes, H3K4me2 and me3 patterns were analyzed genome-wide by ChIP and high-throughput sequencing (ChIP-Seq). Spiked-in S. pombe chromatin was used as an internal control. Both individual gene heat maps (Fig. 4d, e, left and middle panels) and averaged “meta-gene” anchor plots (Supplementary Fig. 4a, b) show that SA200 significantly reduced H3K4me3 peaks. These effects can also be seen in representative Mochiview39 genome browser tracks for individual genes (Supplementary Fig. 4c, d), as well as in heat maps quantitating the differences between SA200 and Set1 FL (Supplementary Fig. 4e, f, left panels). Relative to SA200, the NSΔ200 fusion increased both promoter–proximal H3K4me3 and downstream H3K4me2 (Fig. 4d, Supplementary Fig. 4). Thus, the NSΔ200 methylation patterns are intermediate between wild-type Set1 and SA200 cells, consistent with improved recruitment of COMPASS by the Nrd1 CID. We have previously shown that trimethylation occurs over multiple rounds of transcription35, so in mutants with reduced COMPASS occupancy or activity, promoter–proximal nucleosomes maximally attain H3K4me2, while downstream nucleosomes that would normally have H3K4me2 only reach H3K4me1, making it appear that the H3K4 methylation gradient has shifted upstream35.

Because the NSΔ200 fusion replaces the Set1 Swd2-binding region with the Nrd1 CID, its interaction with RNApII is predicted to be independent of Swd2. To test this, we compared the behaviors of wild-type Set1, SA200, and NSΔ200 in SWD2/ set1Δ versus swdΔ2/swdΔ1 cells. As previously seen25,30, full-length Set1 is degraded in cells lacking Swd2. In contrast, levels of SA200 or NSΔ200 were not reduced in swdΔ2 cells (Fig. 4f). H3K4me3, assayed either in bulk (Fig. 4f) or by ChIP (Supplementary Fig. 2c), was strongly stimulated by Swd2 in SET1 cells, but not in SA200 or NSΔ200. Deletion of SWD2 also strongly diminished interaction between RNApII and wild-type Set1. In contrast, co-immunoprecipitation of RNApII with NSΔ200, while lower than wild-type Set1, was independent of Swd2 (Fig. 4g). These observations further suggest that Swd2 and the Set1 N-terminal region cooperate in COMPASS targeting.

The Swd2 WD40 domain is important for COMPASS recruitment. To characterize the role of Swd2 WD40 domain in COMPASS–RNApII interactions, a point mutant was created at phenylalanine 250, located at the center of the WD40 domain at the tip of propeller blade 6 (Supplementary Fig. 5a, b). Protein levels of
H3K4me levels within 1500 bps from transcription start site (TSS) are shown. ChIP-seq reads from two technical repeats were averaged and plotted.

f  

spiked-in sample of were transformed into 

H3K4me2 (using indicated antibodies. Cells transformed with empty plasmid (Δswd2) served as a negative control. d, e Chip-Seq heat maps of H3K4me3 (d) and H3K4me2 (e) from strains carrying full-length Set1 (Set1), Set1Δ200 (Δswd2), or Nrd1 CID-Set1Δ200 fusion (NSΔ200). After normalizing to an internal spiked-in sample of S. pombe chromatin, SPMR values were mapped for individual RNApII transcriptional units and ordered according to H3K4me3 values in Set1 cells. H3K4me levels within 1500 bps from transcription start site (TSS) are shown. Chip-seq reads from two technical repeats were averaged and plotted. f, g Whole-cell lysates or anti-FLAG immunoprecipitates were separated by SDS-PAGE and analyzed by immunoblotting using the indicated antibodies.

Histone H3 was used as a loading control. Source data are provided as a Source data file.

the F250A mutant were close to those of wild-type Swd2. In contrast, H3K4me2 and me3 levels were much lower in the Swd2 mutant, indicating partial disruption of COMPASS activity (Fig. 5a). As we previously reported for other mutants with reduced H3K4 methylation, Set1 levels were also reduced (Fig. 5a). Immunoprecipitation of COMPASS via either Spp1 or Swd3, whose levels were unaffected in the mutant, or Swd2 itself, also showed that Swd2 F250A was defective for association with COMPASS (Supplementary Fig. 5c). To rule out that loss of interaction with Swd2 F250A was simply an indirect effect of Set1 degradation, we expressed HA-tagged WT or F250A Swd2 in the presence of FLAG-tagged Set1 and untagged WT Swd2, thereby maintaining a supply of functional and stable COMPASS methylation activity. Immunoprecipitation of Set1 confirms that the F250A mutation strongly reduces Swd2 association with Set1 (Fig. 5b).

The yeast two-hybrid interaction between Set1 and the CTD was compared in cells containing WT or F250A Swd2. Similar to the Swd2 deletion (Fig. 3), the reduced interaction in the mutant background suggests that the Swd2 WD40 domain facilitates Set1/COMPASS binding to RNApII (Fig. 5c). Immunoprecipitation with an antibody for CTD Ser5P pulled down less Spp1 and Swd2 in the F250A mutant, further supporting this conclusion (Supplementary Fig. 5d). Interestingly, ChIP-qPCR shows that Swd2 F250A is normally localized (Supplementary Fig. 5g), although this may reflect Swd2 in APT rather than free Swd2. However, Swd2 F250A diminishes enrichment of Spp1 and H3K4me3 at the PMA1 5′ region (Supplementary Fig. 5f, i), while Pol II occupancy is normal (Supplementary Fig. 5h). Supporting these in vivo results, reconstituted recombinant Set1/COMPASS complex incorporated lower levels of F250A than wild-type Swd2 (Fig. 5d) and exhibited reduced H3K4 methyltransferase activity on nucleosomes in vitro (Fig. 5c). Although these in vitro effects are less severe than in vivo, this is likely because H3K4 methylation defects in vivo are further amplified by Set1 degradation. Altogether, the F250A data implicate the Swd2 WD40 domain in COMPASS integrity and recruitment to early elongation complexes.

H2Bub is required for H3K4 methylation independently of Swd2. Ubiquitylation of histone H2B on lysine 123 (H2Bub)
facilitates higher level H3K4 methylations by COMPASS. H2Bub is targeted to transcribed regions by the Paf1 elongation factor complex (Paf1C), which recruits the H2Bub ubiquitination complex Rad6-Bre1. Deleting any of these components causes loss of H2Bub and reduced H3K4me2 and me3. It has been proposed that Swd2 mediates recognition of H2Bub near 5′ ends of genes to facilitate binding of COMPASS. If the primary role of H2Bub was mediated via Swd2, the Swd2-independent fusion of Set1 to the Nrd1 CID might also bypass the requirement for H2Bub.

As expected, immunoblotting of cell extracts showed that H3K4me2 and me3 were reduced below the level of detection in paf1Δ or rtf1Δ cells (Fig. 6a), as well as in bre1Δ cells (Supplementary Fig. 6a). Replacing the first 200 amino acids of Swd2-HA with the F250A mutant did not rescue the H3K4me2 and me3 reduction in paf1Δ or rtf1Δ cells (Fig. 6b). Therefore, the Swd2 WD40 domain is critical for COMPASS function.

**Fig. 5** The Swd2 WD40 domain is critical for COMPASS function. 

**a** Swd2-HA or Swd2-F250A–HA were expressed in an swd2Δ deletion strain. Protein levels were analyzed by immunoblotting using the indicated antibodies. Actin and Histone H3 were used as loading controls. Note that samples were run on the same blot, but intervening lanes were removed (dashed line, full blots appear in Source data). 

**b** Plasmid-encoded Swd2-HA and Swd2-F250A–HA were expressed in FLAG-Set1 containing cells that also express the endogenous untagged SWD2 gene in the chromosome. Proteins were immunoprecipitated with anti-FLAG beads (IP:FLAG) and then immunoblotted for Set1 and Swd2-HA. 

**c** The Y2H interaction between the Gal4BD–Set1 and Gal4AD–CTD fusions was tested in strain PJ69-4A expressing wild-type (WT) or mutant (F250A) Swd2. The empty Gal4 vectors served as negative controls. Activation of the Gal4-activated HIS3 reporter was tested on plates lacking histidine, with the addition of 5 mM 3AT in the last panel for additional stringency.

**d** Analysis of recombinant baculovirus-expressed Set1/COMPASS made with WT Swd2 or the Swd2–F250A mutant. SDS-PAGE/Coomassie blue staining shows total protein, while immunoblotting with anti-FLAG and anti-HA show Set1 and Swd2 levels, respectively. 

**e** H2Bub chromatin was subjected to in vitro histone methyltransferase assays with Swd2 WT and Swd2–F250A mutant-containing complexes shown in panel (d). H3K4 methylation status was monitored by immunoblotting with indicated antibodies. Histone H3 was used as an internal loading control. Source data are provided as a Source data file.
Set1 with the Nrd1 CID partially restored the ability of Set1 to co-immunoprecipitate CTD–Ser5P, even in the presence of Paf1C or Bre1 mutants (Fig. 6b, Supplementary Fig. 6b), but did not restore methylation (Fig. 6a, Supplementary Fig. 6a). The lack of methylation rescue was further confirmed by ChIP-qPCR on the YEF3 gene (Supplementary Fig. 7). These results echo our earlier demonstration that H3K4 methylation by a fusion of Set1 to the Rpb4 polymerase subunit is also still dependent upon H2Bub35. Therefore, artificial recruitment of COMPASS does not bypass the requirement for H2Bub, indicating this modification promotes H3K4 methylation independently of COMPASS association with the RnapII elongation complex. This conclusion has been supported by several structures of nucleosome-bound COMPASS that were published while this paper was in review (see Discussion below)42–44.

**Discussion**

The details of how H3K4 methylation is coordinated with active transcription are still not completely understood. Although Set1/COMPASS targeting is known to involve CTD Ser5 phosphorylation34, it has been unclear whether COMPASS directly binds the CTD, whether RnapII domains other than CTD are involved, and which subunits of COMPASS mediate the interaction. Here, we present evidence that the N-terminal region of Set1 and Swd2 cooperatively interact with RnapII–CTD to promote proper recruitment of COMPASS to transcription elongation complexes at 5′ ends of genes. This model is based on the following observations:

i. Deleting the first 200 amino acids of Set1 (SA200) strongly reduced H3K4me3 and the ability to co-precipitate RnapII.

ii. Swd2, which interacts with amino acids 124–229 of Set116,19, is needed for strong CTD binding by Set1.

iii. Deleting residues 200–210 of Set1 abrogates Y2H interaction with both the Rbp1 CTD and Swd2.

iv. An Swd2 WD40 domain point mutation impairs its interaction with both Set1 and RnapII, recruitment of COMPASS to chromatin, and H3K4 methylation activity.

v. Replacing amino acids 1–200 of Set1 with the Nrd1 CID (NSA200) partially restores H3K4me and Ser5P–CTD binding, showing that CTD binding is a major function of this region.

The Nrd1 CID fusion only partially substitutes for the Set1 N-terminal domain, which might reflect lower CTD affinity. More likely, Set1(1–200) and Swd2 have additional functions, beyond CTD interaction, in regulating COMPASS activity, as indicated by physical interaction studies15, protein crosslinking, and low-resolution cryo-EM structural analysis19. It appears that the Swd2–Set1 N-terminal module may fold over the catalytic body of COMPASS to contact other subunits and more C-terminal Set1 regions. An interesting speculative model is that CTD binding moves the Swd2–Set1 N-terminal module, relieving autoinhibition by making the Set1 active site more accessible.

Several facts led us to initially suspect that Swd2 would be the primary contact point with the CTD. Yeast Swd2 is also a component of the APT complex, which functions in the yeast Nrd1–Nab3–Sen1 (NNS) termination pathway for short non-polyadenylated transcripts. Like COMPASS, NNS termination is also directed to early stage RNAPII elongation complexes45. In addition, the mammalian Swd2 homolog Wdr82 helps recruit the Setd1A to chromatin and can directly bind to CTD Ser5P in vitro15. Echoing our results, this interaction was stimulated by the N-terminal region of mammalian Set1. Despite our expectations, we could not detect a CTD–Swd2 interaction in the Y2H assay, nor did direct fusion of Swd2 to Set1Δ200 rescue H3K4 methylation. Therefore, yeast Swd2 is apparently not sufficient for CTD interaction. Altogether, these results suggest that the N-terminal region of Set1 and Swd2 may form a composite binding site for tethering COMPASS to the RnapII CTD. It may be that both proteins make CTD interactions, perhaps RNApII itself, through Swd226. Interestingly, S. pombe has two Swd2 homologs, each specific for one of the two complexes45. In addition, the mammalian Swd2 homolog Wdr82 helps recruit the Setd1A to chromatin and can directly bind to CTD Ser5P in vitro15. Echoing our results, this interaction was stimulated by the N-terminal region of mammalian Set1. Despite our expectations, we could not detect a CTD–Swd2 interaction in the Y2H assay, nor did direct fusion of Swd2 to Set1Δ200 rescue H3K4 methylation. Therefore, yeast Swd2 is apparently not sufficient for CTD interaction. Altogether, these results suggest that the N-terminal region of Set1 and Swd2 may form a composite binding site for tethering COMPASS to the RnapII CTD. It may be that both proteins make CTD contacts, or alternatively, that one triggers a conformation that renders the other competent for CTD binding.

In contrast to Set1Δ200, Set1 derivatives further C-terminally truncated to amino acids 70035 or 76227,31 can tri-methylate H3K4. However, the resulting H3K4me3 is mislocalized, spreading from the 5′ region into the gene bodies of active genes. This inaccurate distribution of H3K4me3 still requires Spp1, as removal of the Spp1 interaction site by Set1 truncation to 780–1080 leads to loss of H3K4me318,27,31,46,47. Thus, other
modes of COMPASS recruitment to chromatin must exist in the absence of the Set1 region 1–762, likely via direct interactions of the COMPASS catalytic domain with the nucleosome42,44, or perhaps through transient interactions between Swd2 and COMPASS subunits Sppl and Swd119. In the context of full-length Set1, these alternative modes of recruitment may be inhibited by the central autoinhibitory domain of Set117.

In addition to CTD binding, co-transcriptional methylation by COMPASS requires H2Bub. PaflC associated with transcription elongation complexes directly recruits the SWR1 complex and then recruits COMPASS, and this recruitment is dependent on the presence of H2Bub36,37,45–49. Although the Nrd1 CID fusion to Set1 can restore CTD binding, it does not bypass the requirement for H2Bub. Therefore, H2Bub must also remain upon fusion of Set1 to RNApII subunit Rpb435 or strong overexpression of the hyperactive Set1 N-terminal truncation36. The H2Bub requirement for restoration of CTD binding is also remained upon fusion of Set1 to RNAPII subunit Rpb435 or strong overexpression of the hyperactive Set1 N-terminal truncations36. After normalizing for reduced Set1 protein levels in the absence of H3K4 methylation36,35, the ability of Set1 to co-precipitate RNAPII was also unaffected by loss of PaflC or Radder-Bre1 ubiquitination complex (Fig. 6b). Therefore, H2Bub must act at a step after initial recruitment of COMPASS to the RNAPII elongation complex. Indeed, several structures published while this paper was under review show how ubiquitin contacts and allosterically activates COMPASS42,44. While H2B-linked ubiquitin does not affect COMPASS affinity for nucleosomes14, its interactions with Set1 and other subunits rearrange the methyltransferase catalytic site into an active conformation.

A model thus emerges in which nucleosome binding, H2Bub sensing, and methyltransferase activity map to the C-terminal N-SET/SET region of Set1 and associated COMPASS subunits, while an N-terminal domain combines with Swd2 to create a CTD-targeting module. Future biochemical and structural studies will eventually reveal how these two modules interact, and how they are regulated by an intervening region of Set1 linked to autoinhibition and COMPASS degradation. The C-terminal N-SET/SET domains and WRAD subunits are the most conserved among the Set1/MLL family, consistent with their common H3K4 methyltransferase activities. Our experiments suggest that the non-conserved N-terminal domains are likely to target individual family members to different genomic locations through distinct protein interactions.

Methods

Strains, plasmids, and primers. Yeast strains used in this study are listed in Supplementary Table 1. Plasmids used for fusion protein analysis and yeast two hybrid assay are described in Supplementary Tables 2 and 3, respectively. DNA encoding the Cid from Nrd1 (amino acids 1–153) was amplified from yeast genomic DNA and inserted into the indicated Set1 constructs using isothermal assembly65. Primers used for cloning or ChIP assays are listed in Supplementary Table 4.

Antibodies. The following antibodies were used in this study: anti-H3K4me1 (Millipore Sigma 07-436, Burlington, MA, 1:2000), anti-H3K4me2 (Millipore Sigma 07-030, 1:1000), anti-H3K4me3 (Millipore Sigma 07-473 or 04-745, 1:2000), anti-H3K9me1 (Abcam, Cambridge, UK, 1:3000), rat monoclonal anti-H3K9me2 anti-CTD (3E10, Dick Eick, 1:1000), rat monoclonal anti-H3K9me3 CTD (3E8, Dick Eick, 1:1000), mouse monoclonal anti-H3K9me3 CTD for total Rpb1 (8WG6, Buratowski lab, 1:3000), anti-Myc (MMS-150R-500, Covance, 1:2000), anti-H3K4me1 (Buratowski lab, 1:3000), mouse monoclonal anti-CTD for total Rpb1 (8WG6, Buratowski lab, 1:2000), anti-H3K4me2, or anti-H3K4me3 and 10 μg of poly(ADP-ribose) polymerase (PARP) (from human cell lysate, 1:1000), mouse monoclonal anti-H3K4me3 (Millipore Sigma 07-473 or 04-745, 1:2000), mouse monoclonal anti-H3K4me2, or anti-H3K4me3 and 10 μg of poly(ADP-ribose) polymerase (PARP) (from human cell lysate, 1:1000). DNA samples were stained with 1 μg/ml of YOYO-1 (Invitrogen). Primary antibodies were detected using goat anti-mouse IgG conjugated to Alexa Fluor 647 (ThermoFisher Scientific, 1:2000) or goat anti-rabbit IgG conjugated to Alexa Fluor 488 (ThermoFisher Scientific, 1:2000).

Preparation of ChIP-Seq Libraries. Sequencing libraries were prepared using the following procedure, as previously described10. The concentration of immunoprecipitated DNA was measured using Qubit dsDNA HS Assay Kit (ThermoFisher Scientific). Barcoded sequencing libraries were generated from 1 ng of immunoprecipitated DNA32. Briefly, immunoprecipitated DNA was end repaired using T4 DNA polymerase, T4 PNK, and DNA polymerase I Large (Klenow) fragment (New England Biolabs, Ipswich, MA). A single adenosine was added to the 3′ end of the DNA by treating it with T4 DNA polymerase I Large (Klenow) fragment (New England Biolabs, Ipswich, MA) followed by ethanol-precipitation for further analysis.

Bioinformatics analysis. Analysis of ChIP-seq data was performed using the pipeline described in Soares et al.35. Sequence reads were demultiplexed using SABRE (https://github.com/najoshi/sabre.git) allowing for one mismatch in the barcode. For S. pombe spike-in normalization, demultiplexed reads were first aligned to the S. pombe genome (version ASM24v2.1). The unaligned reads were then subsequently aligned to the S. cerevisiae genome (version R64-1-1). Only sequence reads that could be exclusively assigned to each genome were considered for total number of reads. Normalization factors were calculated by first calculating the proportion of reads of S. pombe versus total reads in input reads, and then dividing the value by the square root (to account for the sequence tags per million reads (SPMR) normalization) of the proportion of reads of S. pombe vs. total reads in each immunoprecipitation reads. All alignment was performed using BOWTIE 1.1.1-2 excluding the first base and multi-aligned reads. Conversion of alignment files was performed using SAMTOOLS 1.2.54. Using MACS2,1.4.1 function, pileup tracks were calculated and duplicate reads were removed. Subsequently, tags were extended to 150, and values were normalized to SPMR. Final coverage outputs were converted to high density wig files and
analyzed using custom Python3.4 scripts (http://www.sciencelint.org/, https://github.com/LuisSoares/Manuscript). The GEO accession number for the data reported in this paper is GSE138281.

**Yeast two-hybrid assay (Y2H).** The SET1 open reading frame was cloned as a bait into vector pB66 (N-GAL4 BD-bait fusion). Initial interaction screening of a S. cerevisiae genomic library was performed by Hybrigenics, SA, Paris (https://www.hybrigenics-services.com/content/our-services/interaction-discovery/ultimate-y2h-2). For subsequent Y2H validation and further analysis, S. cerevisiae reporter strains P69-4A or CG1945 were used. Gal4 DBD and Gal4 AD plasmids used in Y2H are listed in Supplementary Table 3. Interactions were scored 2 or 3 days after spotting on SC media plates lacking the appropriate amino acids for reporter gene activation (Gal4-dependent UAS upstream of HIS3 or ADE2). For more stringent selection, 5 or 20 μM 3-amino triazole (3-AT), a competitive inhibitor of the HIS3-gene product, was added to media as indicated.

**Quantitative PCR (qPCR) analysis.** For individual gene analysis, quantitative PCR reactions of DNAs from ChIPs (see above) were done with a BioRad CFX384 using the following parameters: 5 min at 95°C, 40 cycles of 15 s at 95°C, 15 s at 50°C, and 40 s at 72°C, followed by 10 min at 95°C. Oligonucleotides used for PCR reactions are listed in Supplementary Table 4.

**COMPASS purification and histone methyltransferase assay.** Purification of Set1 complex, H2Bub chromatin assembly, and in vitro methyltransferase assay were performed as described previously. Briefly, cDNAs amplified from yeast genomic DNA were subcloned into pFASTBAC1 (ThermoFisher Scientific) with or without an epitope tag and baculoviruses were generated according to the manufacturer’s instruction. Set1 complexes containing Swd2 wild-type or F250A mutant were reconstituted from S9 cells which were infected with combinations of baculoviruses. Proteins/complexes were affinity purified using M2 agarose (Millipore Sigma, Burlington, MA). H2Bub chromatin assembly was performed using the recombinant ACN/NAPI system. For recombinant chromatin methyltransferase assays, reactions containing 350 ng (based on histone amount) recombinant chromatin (35 μl, assembled as above), purified Set1 complexes and 100 μM SAM (S-adenosyl methionine, New England Biolabs) were adjusted to 40 μl with HEG buffer (25 mM HEPES [pH 7.6], 0.1 mM EDTA and 10% glycerol) and incubated at 30°C for 1 h. Proteins were resolved by SDS-PAGE and subjected to immunoblotting.

**Statistics and reproducibility.** All western blots were repeated at least three times and representative images were shown in this paper. Unpaired t test was used for statistical analysis. For ChIP-Seq, mean values from two biological replicates are represented.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

**Data availability** The ChIP-seq data sets generated and analyzed in Fig. 4 and Supplementary Fig. 4 are available in the GEO repository, accession number GSE138281 (https://www.ncbi.nlm.nih.gov/geo/query/acc.fcgi?acc=GSE138281). All other relevant data supporting the key findings of this study are available within the article and its Supplementary Information files or from the corresponding authors upon reasonable request. Uncropped images for gels and blots in all figures are provided in the Source data file. A reporting summary for this Article is available as a Supplementary Information file.

Received: 1 October 2019; Accepted: 14 April 2020; Published online: 01 May 2020

**References**

1. Kouzarides, T. Chromatin modifications and their function. Cell **128**, 693–705 (2007).
2. Li, B., Carey, M. & Workman, J. L. The role of chromatin during transcription. Cell **128**, 707–719 (2007).
3. Shilatifard, A. Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. Annu. Rev. Biochem. **75**, 243–269 (2006).
4. Ernst, P. & Vakoc, C. R. WRAD: enabler of the SET1-family of H3K4 methyltransferases. Brief. Funct. Genomics **11**, 217–226 (2012).
5. Smith, E. & Shilatifard, A. The chromatin signaling pathway: diverse mechanisms of recruitment of histone-modifying enzymes and varied biological outcomes. Mol. Cell **40**, 689–701 (2010).
6. Briggs, S. D. et al. Histone H3 lysine 4 methylation is mediated by Set1 and required for cell growth and rDNA silencing in Saccharomyces cerevisiae. Genes Dev. **15**, 3386–3395 (2001).
7. Rogue, A. et al. The Saccharomyces cerevisiae Set1 complex includes an Ash2 homologue and methylates histone 3 lysine 4. EMBO J. **20**, 7137–7148 (2001).
8. Miller, T. et al. COMPASS: a complex of proteins associated with a trithorax-related SET domain protein. Proc. Natl Acad. Sci. USA **98**, 12902–12907 (2001).
9. Nagy, P. L., Griesenbeck, J. & Cleary, M. L. A trithorax group complex purified from Saccharomyces cerevisiae is required for histone methylation. Proc Natl Acad Sci USA **99**, 90–94 (2002).
10. Dichtl, B., Aasland, R. & Keller, W. Functions for S. cerevisiae Swi2p in 3’ end formation of specific mRNAs and snoRNAs and global histone 3 lysine 4 methylation. RNA **10**, 965–977 (2004).
11. Neda, E. et al. Organization and function of APT, a subcomplex of the yeast cleavage and polyadenylation factor involved in the formation of mRNA and small nuclear RNA 3'-ends. J. Biol. Chem. **278**, 33000–33010 (2003).
12. Dehe, P. M. et al. Protein interactions within the Set1 complex and their roles in the regulation of histone 3 lysine 4 methylation. J. Biol. Chem. **281**, 35404–35412 (2006).
13. Halbach, A. et al. Cotranslational assembly of the yeast SET1C histone methyltransferase complex. EMBO J. **28**, 2959–2970 (2009).
14. Kim, J. et al. The n-SET domain of Set1 regulates H2B ubiquitination-dependent H3K4 methylation. Mol. Cell. **49**, 1121–1133 (2013).
15. Lee, J. H. & Skalnik, D. G. Wdr82 is a C-terminal domain-binding protein that recruits the Setd1A Histone H3-Lys4 methyltransferase complex to transcription start sites of transcribed human genes. Mol. Cell. **28**, 609–618 (2008).
16. Jeon, J., McGinty, R. K., Muir, T. W., Kim, J. A. & Kim, J. Cross-talk among SET domains maintains histone H3 lysine 4 methylation and gene silencing in yeast. Nature **418**, 104–108 (2002).
17. Soares, L. M. & Buratowski, S. Histone crostalk: H2Bub and H3K4 Methylation. Mol. Cell. **49**, 1019–1020 (2013).
18. Lee, J. S. et al. Histone crostalk between H2B monoubiquitination and H3 methylation mediated by COMPASS. Cell. **131**, 1084–1096 (2007).
19. Vitaliano-Prunier, A. et al. Ubiquitination of the COMPASS component Swd2 links H2B ubiquitination to H3K4 trimethylation. Nat. Cell Biol. **10**, 1365–1371 (2008).
20. Neda, E. et al. The Glc7 phosphatase subunit of the cleavage and polyadenylation factor is essential for transcription termination on snoRNA genes. Mol. Cell. **29**, 577–587 (2008).
21. Soares, L. M. & Buratowski, S. Yeast Swd2 is essential because of antagonism between Set1 histone methyltransferase complex and APT (associated with PolyA termination factor). J. Biol. Chem. **287**, 15239–15241 (2012).
22. Schlichter, A. & Cairns, B. R. Histone trimethylation by Set1 is coordinated by the RRM, autoinhibitory, and catalytic domains. EMBO J. **24**, 1222–1231 (2005).
23. Tresaugues, L. et al. Structural characterization of Set1 RNA recognition motifs and their role in histone H3 lysine 4 methylation. J. Mol. Biol. **359**, 1170–1181 (2006).
24. Luciano, P. et al. Binding to RNA regulates Set1 function. Cell Discov. **3**, 17040 (2017).
25. Soares, L. M., Radman-Livaja, M., Lin, S. G., Rando, O. J. & Buratowski, S. Feedback control of Set1 protein levels is important for proper H3K4 methylation patterns. Cell Rep. **6**, 961–972 (2014).
26. Thornton, J. I. et al. Context dependency of Set1/COMPASS-mediated histone H3 Lys4 trimethylation. Genes Dev. **28**, 115–120 (2014).
27. Cordén, J. L. RNA polymerase II C-terminal domain: tethering transcription to transcript and template. Chem. Rev. **113**, 8423–8455 (2013).
28. Buratowski, S. Progression through the RNA polymerase II CTD cycle. Mol. Cell. **36**, 541–546 (2009).
29. Ng, H. H., Robert, F., Young, R. A. & Struhl, K. Targeted recruitment of histone H3 methyltransferase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. Mol. Cell. **11**, 709–719 (2003).
30. Soares, L. M. et al. Determinants of histone H3K4 methylation patterns. Mol. Cell. **68**(773–785), e776 (2017).
31. Wood, A., Schneider, J., Dover, J., Johnston, M. & Shilatifard, A. The Pacl complex is essential for histone monoubiquitination by the Rad6-Bre1
complex, which signals for histone methylation by COMPASS and Dot1p. J. Biol. Chem. 278, 34775–34784 (2003).
37. Krogan, N. J. et al. The Pafl complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation. Mol. Cell 11, 721–729 (2003).
38. Vasileva, L., Kim, M., Mutschler, H., Buratowski, S. & Meinhart, A. The Nrd1-Nab3-Sen1 termination complex interacts with the Ser5-phosphorylated RNA polymerase II C-terminal domain. Nat. Struct. Mol. Biol. 15, 795–804 (2008).
39. Homann, O. R. & Johnson, A. D. MochiView: versatile software for genome browsing and DNA motif analysis. BMC Biol. 8, 49 (2010).
40. Squazzo, S. L. et al. The Pafl complex physically and functionally associates with transcription elongation factors in vivo. EMBO J. 21, 1764–1774 (2002).
41. Van Oss, S. B., Cucinotta, C. E. & Arndt, K. M. Emerging Insights into the Roles of the Pafl Complex in Gene Regulation. Trends Biochem. Sci. 42, 788–798 (2017).
42. Hsu, P. L. et al. Structural basis of H2B ubiquitination-dependent H3K4 methylation by COMPASS. Mol. Cell 76(712–723), e714 (2019).
43. Worden, E. J., Zhang, X. & Wolberger, C. Structural basis for COMPASS recognition of an H2B-ubiquitinated nucleosome. elife https://doi.org/10.7554/eLife.53199 (2020).
44. Xue, H. et al. Structural basis of nucleosome recognition and modification by MLL methyltransferases. Nature 573, 445–449 (2019).
45. Roguev, A. et al. A comparative analysis of an orthologous proteome environment in the yeasts Saccharomyces cerevisiae and Schizosaccharomyces pombe. Mol. Cell 7, 125–132 (2004).
46. Fingerman, J. M., Wu, C. L., Wilson, B. D. & Briggs, S. D. Global loss of Set1-mediated H3 Lys4 trimethylation is associated with silencing defects in Saccharomyces cerevisiae. J. Biol. Chem. 280, 28761–28765 (2005).
47. Acquaviva, L. et al. The COMPASS subunit Spt1 links histone methylation to initiation of meiotic recombination. Science 339, 215–218 (2013).
48. Ng, H. H., Dole, S. & Struhl, K. The Rif1 component of the Pafl transcriptional elongation complex is required for ubiquitination of histone H2B. J. Biol. Chem. 278, 33625–33628 (2003).
49. Battaglia, S. et al. RNA-dependent chromatin association of transcription elongation factors and Pol II CTD kinases. elife https://doi.org/10.7554/elife.25637 (2017).
50. Kim, J. et al. RAD6-Mediated transcription-coupled H2B ubiquitination directly stimulates H3K4 methylation in human cells. Cell 137, 459–471 (2009).
51. Gibson, D. G. et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods 6, 343–345 (2009).
52. Wong, K. H., Jin, Y. & Moqtaderi, Z. Multiplex Illumina sequencing using DNA barcoding. Curr. Protoc. Mol. Biol. https://doi.org/10.1002/0471142727. mb0711s101 (2013).
53. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 10, R25 (2009).
54. Li, H. et al. The sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079 (2009).
55. Feng, J., Liu, T., Qin, B., Zhang, Y. & Liu, X. S. Identifying ChIP-seq enrichment using MACS. Nat. Protoc. 7, 1728–1740 (2012).
56. Kim, J. & Roeder, R. G. Nucleosomal H2B ubiquitination with purified factors. Methods 54, 331–338 (2011).

Acknowledgements
H.B. was supported by a postdoctoral fellowship from the Basic Science Research Program of the National Research Foundation of Korea (NRF), funded by the Ministry of Education (2015R1A6A3A03017730). This work was supported by grants from the National Research Foundation of Korea (NRF-2019R1A2C209830) to J.K., from “Ligue Nationale Contre le Cancer” (LNCC) (Equipe labellisée) to M.D. and V.G., and grants GM046498 and GM056663 from the U.S. National Institutes of Health to S.B. We thank D. Eick (CIPS, Munich) for CTD antibodies, S. Briggs (Purdue) for the parent FLAG-Sen1 construct, C. Gwizdek (Dargemont lab) for creating the Swd2 F250A mutant, and N. Adachi (SBRC, KEK, Japan) for purified RNApolII.

Author contributions
H.B. and M.D. performed all the experiments, with the exception of COMPASS reconstitution experiments performed by J.J. and J.K., and computational analyses of genomic data by L.S. and H.B. C.D. identified and initially characterized the SWD2 F250A mutant. H.B., M.D., V.G., and S.B. participated in experimental conception and design, and wrote the paper with input from all other authors.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41467-020-16082-2.

Correspondence and requests for materials should be addressed to V.G. or S.B.

Peer review information Nature Communications thanks Yali Dou and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2020