Published in final edited form as:

Nat Struct Mol Biol. 2008 August ; 15(8): 795–804. doi:10.1038/nsmb.1468.

The Nrd1–Nab3–Sen1 termination complex interacts with the Ser5-phosphorylated RNA polymerase II C-terminal domain

Lidia Vasiljeva[1], **Minkyu Kim**[1,3], **Hannes Mutschler**[2,3], **Stephen Buratowski**[1], and **Anton Meinhart**[2]

[1]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Avenue, Boston, Massachusetts 02115, USA.

[2]Department of Biomolecular Mechanisms, Max Planck Institute for Medical Research, Jahnstrasse 29, 69120 Heidelberg, Germany.

Abstract

RNA polymerase II (Pol II) in *Saccharomyces cerevisiae* can terminate transcription via several pathways. To study how a mechanism is chosen, we analyzed recruitment of Nrd1, which cooperates with Nab3 and Sen1 to terminate small nucleolar RNAs and other short RNAs. Budding yeast contains three C-terminal domain (CTD) interaction domain (CID) proteins, which bind the CTD of the Pol II largest subunit. Rtt103 and Pcf11 act in mRNA termination, and both preferentially interact with CTD phosphorylated at Ser2. The crystal structure of the Nrd1 CID shows a fold similar to that of Pcf11, but Nrd1 preferentially binds to CTD phosphorylated at Ser5, the form found proximal to promoters. This indicates why Nrd1 cross-links near 5′ ends of genes and why the Nrd1–Nab3–Sen1 termination pathway acts specifically at short Pol II–transcribed genes. Nrd1 recruitment to genes involves a combination of interactions with CTD and Nab3.

Transcription by Pol II is coordinated with other processes such as mRNA capping, splicing, polyadenylation, and RNA surveillance and export from the nucleus to ensure the efficiency and accuracy of gene expression1,2. The CTD of the Pol II largest subunit can couple transcription and mRNA processing by recruiting factors to transcribing Pol II3. The CTD contains tandem repeats of a heptad sequence (Tyr1-Ser2-Pro3-Thr4-Ser5-Pro6-Ser7) that are dynamically phosphorylated or dephosphorylated on Ser5 (Ser5P) and Ser2 (Ser2P) over the course of transcription. Various chromatin-modifying enzymes and mRNA-processing factors interact with specific CTD-phosphorylated forms that predominate at different stages of transcription.

Correspondence should be addressed to S.B. (steveb@hms.harvard.edu).

[3]These authors contributed equally to this work.

AUTHOR CONTRIBUTIONS L.V. performed CTD peptide pull-downs, Nrd1–Nab3–Pol II co-precipitations and RNA analysis; M.K. performed the ChIP experiments and constructed several yeast strains; H.M. performed the fluorescence anisotropy experiments and calculated K_d values; A.M. crystallized and solved the structure of the Nrd1 CID. L.V., A.M. and S.B. directed the research and wrote the paper.

Accession codes Protein Data Bank: Model coordinates and structure factor amplitudes for $Nrd1_{6-151}$ are deposited under accession code 3CLJ.

Ser5P is highest early in elongation, where it recruits mRNA capping enzyme4,5 and the histone H3 Lys4 (H3K4) methyltransferase Set1 (ref. 6). In contrast, the histone H3 Lys36 (H3K36) methyltransferase Set2 preferentially binds the doubly phosphorylated CTD (Ser2P/Ser5P), which is characteristic of elongating Pol II6. The Ctk1 kinase and Ser2P are important for co-transcriptional recruitment of Pcf11, an essential factor for mRNA polyadenylation and transcription termination7–9. Pcf11 contains a CID that preferentially binds Ser2P CTD, although it also binds nonphosphorylated CTD10. Structures of the Pcf11 CID show that the CTD nestles in a surface pocket, but that the phosphate group of Ser2P observable in the structure does not directly contact Pcf11 (refs. 11,12).

There are at least two termination pathways for Pol II in *S. cerevisiae*, both of which require Pcf1 1 and one other CID protein13. In the mRNA pathway, cleavage at the poly(A) site triggers degradation of the still-elongating downstream RNA transcript. This degradation by the exonuclease Rat1 (also known as Xrn2) somehow triggers transcription termination14,15. The CID protein Rtt103 is part of this termination complex14,15. The function of this Ser2P binding protein is unclear; it is not essential for viability, but it may help to recruit Rat1 to the transcribing polymerase.

The second termination pathway, used at small nucleolar RNAs (snoRNAs) and other short Pol II transcripts, requires the CID protein Nrd1 (refs. 13,16,17). Nrd1 binds RNA in a sequence-specific manner via an RNA recognition motif (RRM)18–20. Other components of the snoRNA termination complex include the RNA binding protein Nab3, the Sen1 helicase and the cap binding complex consisting of Cbp20 and Cbp80 (ref. 21). It is thought that Sen1 may disrupt the elongation complex, leading to termination. Also associated with Nrd1 are the 3′-to-5′ exonuclease complex known as the exosome and the exosome-activating complex TRAMP21. Both Nrd1 and TRAMP stimulate the exosome's ability to degrade RNAs21,22. Because the exosome trims snoRNA 3′ ends, this association couples termination and 3′ end maturation at these genes.

High-resolution studies of Pol II distribution across the yeast genome revealed that a *SEN1* mutation causes defective termination at most snoRNA genes, short mRNA genes (fewer than 600 nucleotides (nt)) and a few mRNA premature-termination (that is, attenuation) sites, as well as within some previously nonannotated intergenic areas23. Recent yeast microarray expression experiments also revealed a surprising number of cryptic transcripts that were much more abundant in exosome mutants24–27. The Nrd1–Sen1 complex has been implicated in transcription termination of these cryptic unstable transcripts (CUTs)23,28–32. Therefore, the association of the exosome with the Nrd1–Sen1 termination pathway also provides a connection between transcription and RNA surveillance. Depending on the extent of exosome degradation, the Nrd1–Sen1 termination pathway can lead to either 3′ end trimming (as at snoRNAs) or complete degradation (CUTs).

To explore how Pol II chooses between the mRNA and snoRNA termination pathways, we focus here on the recruitment mechanisms of *S. cerevisiae* Nrd1. The crystal structure of the Nrd1 CID adopts a fold similar to that of Pcf11. Unexpectedly, CTD interaction studies showed that, unlike the other yeast CID proteins Pcf11 and Rtt103, Nrd1 binds preferentially to CTD phosphorylated at Ser5. This helps explain targeting of the Nrd1–Sen1 complex to

5′ regions of genes and why this termination pathway preferentially acts at short transcription units23. The interaction between Nrd1 and Nab3 is also crucial for recruitment of the complex and suggests that a combination of CTD Ser5P binding and RNA sequence recognition by Nrd1 and Nab3 channel particular RNAs into this termination and processing pathway.

RESULTS

Structure of the Nrd1 CID

The N-terminal region of Nrd1 has sequence similarity to the CIDs of Pcf11 and Rtt103 (refs. 11,33), so we determined the crystal structure of Nrd1 residues 6–151 (Fig. 1a–c). $Nrd1_{6–151}$ folds into a right-handed superhelical arrangement similar to the CID of Pcf11 (Fig. 1d). The residues of the CTD binding pocket are highly conserved (Fig. 1c, dark green). However, relative to Pcf11, $Nrd1_{6–151}$ has an insertion located within the loop region between helices 1 and 2. Notably, a sulfate ion derived from the crystallization solution binds this loop region (Fig. 1b,c). The sulfate is coordinated by backbone amide groups of Lys21 and Ser22 and a water molecule that is bound by the backbone carbonyl group of Ile24. The only direct side chain interaction to the sulfate ion is with the γ-oxygen atom of Ser22. There is also a water-mediated contact between the sulfate and a symmetry-related Nrd1 molecule. However, the sulfate placement is unlikely to be a crystal-packing artifact, because the CID of *Schizosaccharomyces pombe* Nrd1 also has a sulfate at this position with completely different packing (A.M., unpublished data).

Another expanded region in Nrd1 is found in helix 4, which is extended by an additional fifth helical turn. Pcf11 residues involved in CTD binding are clustered in helix 4, and these are conserved in Nrd1, placing the Nrd1 extension adjacent to the predicted main CTD binding pocket (Fig. 1). The loop lying just C-terminal to helix 4 is also extended in Nrd1. This nonconserved loop is rich in asparagine and serine residues and is disordered in the crystal structure. No electron density could be assigned for residues Ser84 to Ser87. Finally, whereas the Pcf11 CID has a single helix 8, the nonconserved C-terminal region of the Nrd1 CID is split into two helices, here designated 8a and 8b.

Nrd1 preferentially binds the Ser5-phosphorylated CTD

The Nrd1 protein interacts with mouse CTD in a yeast two-hybrid assay34. Given that residues within the CTD interaction pocket of Pcf11 are conserved in Nrd1, it seemed likely that Nrd1 would show preferential binding to CTD-Ser2P. To further examine this interaction, Nrd1 was incubated with synthetic CTD peptides immobilized on beads. The differentially phosphorylated 28-mer CTD peptides consist of four heptad repeats and were either unmodified or phosphorylated at Ser2, Ser5 or both residues. The peptide beads were incubated with purified recombinant protein (rNrd1) or with tandem-affinity purified (TAP) Nrd1 complex from yeast (yNrd1). The bound material was eluted and analyzed by SDS-PAGE and immunoblotting (Fig. 2a). Unexpectedly, Nrd1 bound strongly to the CTD-Ser5P and the CTD-Ser2P/Ser5P peptides, but not to CTD-Ser2P (Fig. 2a, above). This binding pattern contrasted with that of Rtt103 (ref. 14; Fig. 2a, below) and Pcf11 (refs. 35–37), which specifically bind to CTD-Ser2P peptides.

To confirm the Nrd1 preference in a quantitative solution binding assay, we carried out fluorescence anisotropy experiments by titrating $Nrd1_{6–151}$ against a labeled Ser5P-CTD peptide consisting of two repeats. Unlabeled peptides were used for competition experiments and equilibrium dissociation constants (K_d) were calculated from the displacement of the binding curves (Fig. 2b and Table 1). Whereas a CTD-Ser2P peptide bound with weak affinity (K_d = 390 μM), Ser5P improved binding at least ten-fold (K_d = 40 μM for CTD-Ser5P). Adding Ser2P to Ser5P increased binding about two-fold (K_d = 16 μM for CTD-Ser2P/Ser5P peptides). No major difference in affinity was observed between two- and four-repeat CTD-Ser5P peptides (Table 1). Therefore, Nrd1 shows high affinity for both CTD-Ser5P and CTD-Ser5P/Ser2P *in vitro*.

The relevance of this *in vitro* binding was tested *in vivo*. A Nrd1-TAP fraction, which contains associated Pol II21, was probed with antibodies specifically recognizing different phosphorylated forms of the CTD. The purified Nrd1 complex contains Pol II that reacts with H14 antibody recognizing Ser5P (Fig. 2c). In contrast, little reactivity was seen with H5, an antibody that primarily reacts with Ser2P but also weakly with Ser5P9. Further arguing that Ser2P is not essential for Nrd1 recruitment, cross-linking of Nrd1 to the *snR33* gene was not affected by deletion of the Ser2 kinase Ctk1 (Fig. 2d). Nrd1 cross-linking was also unaffected at two other mRNA genes (data not shown). On the basis of these experiments, we conclude that, although Nrd1 can bind CTD-Ser5P or CTD-Ser5P/Ser2P *in vitro*, Ser5P is the main determinant of Nrd1 binding *in vivo*.

Although Nrd1 does not show the same phosphorylation preference as Pcf11, there is strong conservation between Nrd1 and Pcf11 of residues that bind to the CTD β-turn in the Pcf11 structure (Supplementary Fig. 1 online). The phosphate group on Ser2 does not contact the Pcf11 CID11, and Pcf11 can bind to the nonphosphorylated CTD35. Therefore, we predicted that CTD binding by the two proteins should be similar. Using the Pcf11 CTD structure11 as a guide, a CTD-Ser2P peptide was modeled in the presumed binding pocket of Nrd1 (Fig. 3, left). To test the validity of this docking model, we mutated Asp70. The corresponding aspartate in Pcf11 forms an important hydrogen bond with the CTD Tyr1 (refs. 11,12). Nrd1 Asp70 also makes a salt bridge to Arg74. The Nrd1 D70R mutant loses the ability to bind the CTD (Table 1). Similarly, mutation of Ile130, predicted to disrupt a contact with CTD residue Pro3, severely reduced binding. Therefore, it is likely that the conserved pocket of Nrd1 binds the β-turn of the CTD in much the same way as Pcf11. However, these interactions alone do not explain the specificity for different phosphorylation states by either protein.

As noted above, a sulfate ion is bound in a shallow hole close to the conserved CTD binding pocket of $Nrd1_{6–151}$ (Fig. 1). We postulated that this sulfate might identify a position normally occupied by a phosphate group, either from the CTD or a phosphorylation site within another part of Nrd1 (ref. 18). The N-terminal CTD residue in the Pcf11 cocrystal structure was Pro6, but other CTD residues were modeled on the Nrd1 structure by overlaying an extended β-strand conformation seen with other CTD-Ser5P binding proteins38,39. In this model, the phosphate group from Ser5P overlaps the observed sulfate ion (Fig. 3, right). To test whether this region contributes to CTD-Ser5P recognition, mutations were generated in Leu20, Lys21 or Ser22. In the fluorescence anisotropy assay,

affinities for CTD-Ser5P were reduced in Lys21 and Ser22 mutant CID proteins, consistent with this hypothesis (Table 1 and Supplementary Fig. 2 online). The crystal structure of the K21P mutant shows that this substitution slightly distorts the peptide backbone to make the sulfate contact less favorable (data not shown). The S22D structure shows that the aspartate side chain occupies the sulfate site (data not shown). In total, the binding experiments indicate that the conserved CID pocket of Nrd1 is essential for CTD binding, but other contacts outside the pocket are likely to contribute to specificity.

To determine whether regions of Nrd1 outside the CID (Fig. 4a) contribute to CTD interaction, additional Nrd1 deletion proteins were tested for the ability to bind CTD peptides (Table 1 and Supplementary Fig. 3 online). Nrd1 derivatives consisting of residues 6–214 or 6–224 bound with the same affinity as $Nrd1_{6–151}$. The region C-terminal to the CID interacts with Nab3 in a yeast two-hybrid screen18, so a Nab3 fragment sufficient for Nrd1 interaction (see below) was added to test for allosteric effects on binding to a CTD-Ser5P tetrarepeat peptide. We observed no change in affinity (Table 1). Furthermore, a Nrd1 deletion lacking residues 150–214 bound the CTD similarly to full-length protein (data not shown). Finally, $Nrd1_{307–560}$, containing only the RNA binding region and the C-terminal region, showed no CTD binding (data not shown). These results indicate that CTD recognition by Nrd1 is entirely contained within the CID.

Nrd1 interactions with Nab3 and Pol II mediate 5′ recruitment

Although CTD binding is important for Nrd1 recruitment, Nrd1 also interacts with the sequence-specific RNA binding protein Nab3. To examine the relative contributions of these interactions *in vivo*, two deletion alleles (Δ6–214 or Δ151–214) were shuffled into yeast from which the wild-type *NRD1* allele was removed (Fig. 4b). Notably, deletion of residues 6–214 was lethal, indicating that this region provides one or more functions essential for viability. The Nrd1Δ151–214 strain (lacking the Nab3 interaction domain) showed a slow growth phenotype at room temperature and inability to grow at 37°1C. In a separate experiment, in which Nrd1 deletion alleles were integrated into the genome (Fig. 4c), the Nrd1Δ151–214 strain again showed slow growth and temperature sensitivity. The Nrd1Δ6–150 strain (lacking only the CID) grew similarly to the wild-type parental strain, although a similar deletion has been reported to cause slow and conditional growth in a different background18. Therefore, both Nab3 interaction and the CID are important, and the lethality of the combined deletion suggests partial redundancy of these domains in Nrd1 recruitment.

The region between residues 169 and 245 was previously shown to interact with Nab3 in a yeast two-hybrid screen18. To confirm that residues 151–214 of Nrd1 are important for Nab3 association, we precipitated the Nrd1Δ6–214 and Nrd1Δ151–214 proteins and monitored the presence of Nab3 by immunoblotting (Fig. 4d). Neither of the mutants was able to bind to Nab3. Further confirming a direct interaction, purified recombinant $Nrd1_{6–224}$ stably interacted with $Nab3_{204–248}$ in pull-down assays and analytical size-exclusion chromatography (Supplementary Fig. 4a online). No interaction was seen between $Nrd1_{6–151}$ and $Nab3_{204–248}$ (data not shown). Isothermal calorimetry titrations (Supplementary Fig. 4b) show that $Nrd1_{6–224}$ and $Nab3_{204–248}$ interact with an apparent K_d of 160 nM and a stoichiometry value *n* of 1, suggesting that the two proteins bind to each

other in an equimolar ratio. The reconstituted complex of $Nrd1_{6–224}$ and $Nab3_{204–248}$ migrated in analytical size-exclusion experiments with an apparent mass expected for a heterodimer (Supplementary Fig. 4c). Finally, a complex of full-length Nrd1 and Nab3 sediments as a heterodimer during analytical centrifugation40.

Both Nrd1Δ6–150 and Nrd1Δ151–214 deletion mutants showed greatly reduced association with Pol II in extracts, indicating that both CTD interaction and Nab3 association with Nrd1 contribute to the interaction of the Nrd1–Sen1 complex with polymerase (Fig. 4e). This assertion was further supported by chromatin immunoprecipitation (ChIP) experiments showing that Nrd1Δ151– 214 cross-linking to the *snR13* gene was strongly reduced relative to wild-type (Fig. 5a). Furthermore, the recruitment of Nrd1 to the 5′ region of two mRNA genes was lost when Nrd1 lacked the Nab3 interaction domain (Nrd1Δ151–214) or the CID (Nrd1Δ6–150) (Fig. 5b,c). Therefore, at least three mechanisms contribute to recruitment of Nrd1 to genes: interaction with CTD-Ser5P via the CID, recognition of specific RNA sequences via the Nrd1 RRM and interaction with Nab3 (which also binds to specific RNA sequences via an RRM domain16,18). It is likely that different genes rely more or less strongly on one or more of these mechanisms.

Functional importance of Nrd1 interactions

The Nrd1 complex is involved in both transcription termination and 3′ end processing of snoRNAs13,21. To determine how the different Nrd1 domains contribute to these processes, expression of two snoRNA genes (Fig. 6a) was monitored by northern blotting in various *nrd1*-mutant backgrounds (Fig. 6b). When Nrd1 is inactivated using a temperature-sensitive point mutant, the *snR13* gene produces a read-through transcript that indicates a termination defect. In contrast, the *snR33* gene relies on Nrd1 for 3′ end processing and primarily produces a 3′ extended precursor RNA upon Nrd1 inactivation13,21.

In cells lacking the Nrd1 CID (Nrd1Δ6–150; Fig. 6b, lanes 3 and 8) we obsereved no termination defects at either *snR13* or *snR33*. However, we did observe accumulation of the *snR33* precursor RNA in this strain (Fig. 6b, lane 8), arguing that the Nrd1 CID is likely to be important for recruitment of the exosome for 3′ end trimming at this gene. In contrast, loss of the Nab3 interaction region (Nrd1Δ151–214) results in appearance of the snR13-TRS31 read-through transcript (Fig. 6b, lanes 4 and 5). On the *snR33* gene, deletion of the Nab3 interaction domain had no effect at the permissive temperature of 23°C but caused accumulation of both the snR33 precursor and snR33-YCR015c read-through transcript at 37°C (Fig. 6b, lanes 9 and 10). However, this result was complicated by the observation that shifting a wild-type *NRD1* strain to 37°C also increased levels of the snR33 precursor (Fig. 6b, lane 7).

To avoid temperature shift, we used a strain in which the endogenous *NRD1* promoter was replaced with the *GAL1* promoter30. This strain is grown in galactose, and a shift to glucose leads to reduced levels of Nrd1 within 2 h. This loss of Nrd1 results in accumulation of the snR13-TRS1 read-through transcript and the snR33 precursor (Fig. 6c,d, lanes 1–4). Notably, snR33 precursor levels dropped when cells remained in glucose for longer time periods (Fig. 6d, lanes 3 and 4). This may indicate that the precursor transcripts are unstable and/or that transcription rates drop as these cells die. The effects of glucose shift were

blocked when the strain also contained a copy of wild-type *NRD1* on a plasmid (Fig. 6c, lanes 17–20, and Fig. 6d, lanes 13–16). Plasmids expressing various *NRD1* mutants expressed from the *NRD1* promoter were introduced into the Nrd1-depletion strain. Confirming the above results, Nrd1 lacking the Nab3 interacting region (Nrd1Δ151–214) had partially defective snR13 termination (Fig. 6c, lanes 9–12). The larger deletion lacking both the CID and Nab3 interaction domain (Nrd1Δ6–214) could not rescue snR13 termination or snR33 processing (Fig. 6c,d, lanes 5–8).

Nrd1 lacking the RRM domain could not rescue either the snR13 transcription termination or snR33 RNA processing defects (Fig. 6c, lanes 13–16, and Fig. 6d, lanes 9–12). The defects with this mutant were less severe than with Nrd1Δ6–214, but there was some read-through even before glucose shift, suggesting that the RRM deletion may have a dominant-negative effect when coexpressed with the wild-type protein. The DRRM mutant may compete with full-length Nrd1 for CTD and Nab3 binding, but may not be fully functional.

DISCUSSION

Transcription by yeast Pol II terminates by at least two mechanisms: the Rat1-dependent ‘torpedo’ pathway and the Sen1–Nrd1 pathway13. The Rat1 pathway works at mRNA genes, whereas the Sen1–Nrd1 pathway functions at snoRNAs, CUTs and some short mRNAs13,23,28,30,41. Both pathways involve interactions between the Pol II CTD and CID proteins. Rtt103 and Pcf11 function in the mRNA pathway and preferentially bind CTD Ser2P10–12,14. Pcf11, which also binds nonphosphorylated CTD, is required for both mRNA and snoRNA termination13,37,42, suggesting a function that is common for both pathways.

Here we show that the Nrd1 CID resembles the Pcf11 CID structurally, but has a different phosphorylation preference for the CTD. *In vitro*, Nrd1 binds strongly to CTD-Ser5P and slightly better to CTD-Ser2P/Ser5P. However, several findings indicate that Ser2 phosphorylation is not crucial *in vivo*. Pol II associated with Nrd1 *in vivo* reacts with antibody H14 (recognizing Ser5P), but not H5 (primarily recognizing Ser2P). Furthermore, whereas the Nrd1 CID is required for recruitment of Nrd1 to the 5′ ends of genes, deletion of the Ser2 kinase Ctk1 has no effect. Therefore, we conclude that Ser5P is the primary determinant of CTD interaction for Nrd1 *in vivo*.

The unexpected specificity of Nrd1 for CTD-Ser5P, the promoter-proximal phosphorylation state4, explains several observations. Whereas both Rtt103 and Pcf11 cross-link at 3′ ends of Pol II– transcribed genes, Nrd1 cross-links strongly at 5′ ends and, to some extent, at 3′ ends (Fig. 5)13,43. Mutation of Sen1 causes termination defects at snoRNA genes and mRNA genes shorter than 600 nt23. Both Sen1 and Nrd1 are necessary for suppression of CUTs, the short unstable transcripts produced by cryptic promoters throughout the yeast genome23,29–32. When a Nrd1-dependent terminator sequence is moved further downstream, where Ser2P predominates and Ser5P levels are lower, it no longer functions properly44,45.

The different CTD specificities of Nrd1, Pcf11 and Rtt103 are also consistent with genetic observations suggesting that the Ser2 kinase Ctk1 acts in opposition to the Nrd1–Sen1–Nab3

complex18. A cold-sensitive allele of *NAB3* is suppressed by deletion of *CTK1*. Nab3 overexpression exacerbates cold sensitivity caused by *CTK1* deletion, whereas the *nrd1-102* allele weakly suppresses *ctk1*Δ18. Finally, increasing CTD-Ser2P levels by mutating the CTD phosphatase Fcp1 increases levels of read-through transcripts at a Nrd1-dependent terminator45. These observations suggest competition between the two termination pathways, with Ser5P early in elongation favoring the Sen1 pathway via Nrd1 and Ser2P at later times favoring the poly-adenylation/torpedo pathway via Rtt103 and Pcf11.

It is unclear what leads the CIDs of Pcf11 and Nrd1 to have different specificities. The Nrd1 CID is similar to Pcf11 in overall conformation, and a central CTD binding pocket seems to be conserved11. A superposition of Nrd1 and Pcf11 was used to model a possible Nrd1-CTD interaction (Fig. 3 and Supplementary Fig. 1). Mutagenesis studies indicate that conserved residues in the CID pocket are necessary for Nrd1 binding to the CTD (Fig. 3 and Table 1). Binding of the CTD in this conserved pocket may be independent of the CTD-phosphorylation status. Although Pcf11 binding to the CTD is enhanced by Ser2P, Pcf11 also binds unphosphorylated and doubly phosphorylated CTD10. In the Pcf11 structure, the single observed Ser2 phosphate does not contact the CID11. The Ser2P preference may be in part due to a hydrogen bond between the CTD Ser2 phosphate and the CTD Thr4 side chain that stabilizes the β-turn11, but the unphosphorylated CTD shows an intrinsic propensity to form β-turns within the Ser2-Pro3-Thr4-Ser5 motif 46. Thus, the specificity of CIDs for different CTD-phosphorylation sites may be determined by additional contacts outside the central binding pocket. This idea is supported by recent cocrystal structures of phosphorylated CTD bound to the CID of the mammalian SCAF8 protein, where CID surface residues directly contact the CTD phosphates47. A sulfate ion bound to Nrd1 may represent a Ser5P CTD interaction site (Fig. 3). Consistent with this idea, point mutations in this region reduce affinity for CTD-Ser5P peptides (Table 1).

CTD binding is only one of several mechanisms for recruiting the Nrd1–Nab3–Sen1 complex to RNAs (Supplementary Fig. 5 online).Neither the CID nor the Nab3 interaction domain is essential for viability, but both contribute to interaction with the polymerase and Nrd1 recruitment (Fig. 4 and Fig 5). There may be partial redundancy, because deletion of both domains is lethal18 (Fig. 4b). Both Nrd1 and Nab3 are sequence-specific RNA binding proteins that can be targeted to specific transcripts carrying the appropriate recognition sequences13,16,20,21,26,28,30,40,41,44. This may explain cross-linking of Nrd1 observed to regions downstream from the promoter, where Ser5P levels are likely to be lower.

Once targeted to the RNA, the Nrd1–Nab3–Sen1 complex terminates transcription by a mechanism that may involve the helicase activity of Sen1. Sen1-mediated termination is coupled to RNA 3′ processing and degradation events mediated by the TRAMP–exosome complex. We previously demonstrated a physical interaction between the exosome–TRAMP and Nrd1 complexes and showed that this interaction recruits the exosome to RNAs containing Nrd1 binding sites21. At snoRNAs, the recruitment of exosome results in 3′ end trimming21. For CUTs and certain mRNAs, this pathway results in complete degradation of the transcript (reviewed in ref. 29).

It remains to be seen how the *S. cerevisiae* Nrd1–Sen1–exosome pathway relates to gene expression in higher eukaryotes. Metazoan genomes have multiple CID proteins, several of which also carry RRMs. Furthermore, there is a mammalian Sen1-like protein called Senataxin that has been implicated in several ataxia syndromes. In mammals, most snoRNAs are processed from mRNA introns, so this pathway may be used primarily for termination and degradation of cryptic transcripts rather than for snoRNA biogenesis. Recent transcript mapping and Pol II cross-linking studies in higher eukaryotes suggest that transcription is surprisingly widespread throughout the genome and that most of these transcripts do not correspond to coding genes or stable noncoding transcripts48. Therefore, suppression of cryptic transcription by co-transcriptional targeting of termination and degradation machineries may be even more important in higher eukaryotes.

METHODS

Plasmids and yeast strains

S. cerevisiae strains used are described in Supplementary Table 1 online. Plasmids are listed in Supplementary Table 2 online.

Expression of recombinant proteins

We expressed and purified recombinant proteins from constructs pET21b-Nrd1$_{307–560}$, pET-Nrd1, pET-Nrd1Δ6–214, pET-Nrd1Δ151–214, pET-Nrd1Δ39–169 and pET41a(+)-nrd1 derivatives with point mutations as previously described21.

Recombinant Nrd1$_{6–151}$, Nrd1$_{6–224}$ and Nab3$_{204–248}$ proteins were expressed in *E. coli* BL21 (D*E3*) CodonPlus RIL cells (Stratagene) by inducing with 0.5 mM IPTG overnight at 20°C. Selenomethionine labeling of Nrd1$_{6–151(L37M\ L77M)}$ was as described in ref. 49. Cells were harvested and resuspended in suspension buffer (SB; 50 mM Tris-HCl, pH 8.0, 500 mM KCl and 10 mM β-mercaptoethanol). Cells were sonicated and debris cleared by centrifugation. Soluble lysate was run over a HisTrap FF column (GE Healthcare) equilibrated with SB. After washing with high-salt buffer (50 mM Tris-HCl, pH 8.0, 1 M NaCl and 10 mM β-mercaptoethanol), bound proteins were eluted with a 10 CV gradient (0–500 mM imidazole) of elution buffer (50 mM Tris-HCl, pH 8.0, 250 mM NaCl, 10 mM β-mercaptoethanol, plus imidazole). To improve purity, variant Nrd1$_{6–224}$ was diluted with dilution buffer (50 mM MES, pH 6.5, 1 mM EDTA and 1 mM dithioerytritol (DTE)) and loaded onto a MonoS column equilibrated with dilution buffer containing 100 mM NaCl. Nrd1$_{6–224}$ was eluted with a gradient of 15 CV (dilution buffer plus 100–600 mM NaCl). Concentrated peak fractions were applied to a Superose-6 column equilibrated with SEC buffer (25 mM HEPES-NaOH, pH 8.0, 100 mM NaCl, 2 mM EDTA and 1 mM DTE).

For purification of a recombinant glutathione S-transferase (GST) fusion of Nab3$_{204–248}$, cleared cell lysate was loaded onto glutathione-Sepharose equilibrated with SB. After extensive washing, bound proteins were eluted with SB containing 20 mM reduced glutathione. Protein fractions were dialyzed against cleavage buffer (50 mM Tris-HCl, pH 7.3, 100 mM NaCl, 2 mM $CaCl_2$ and 1 mM DTE), and Nab3$_{204–248}$ peptide was cleaved from GST using thrombin. GST and peptide were separated by size-exclusion

chromatography over Superose-12 equilibrated with size-exclusion buffer (50 mM HEPES, pH 7.5, 150 mM NaCl, 50 mM $(NH_4)_2SO_4$ and 1 mM EDTA). All purified proteins were 99% pure, judged by Coomassie-stained SDS-PAGE. For crystallization, pure proteins were concentrated to 20 mg ml^{-1}.

Tandem-affinity purification and CTD-affinity chromatography

Cell extracts and TAP purifications were prepared as previously described21. TAP pull-downs include RNase A treatment of extracts to reduce RNA-mediated associations. For CTD-affinity chromatography, recombinant Nrd1 proteins (5 μg), TAP-purified Nrd1 complex or Rtt103-hemagglutinin–tagged whole-cell extract (0.5 mg) was assayed as described50. Biotinylated CTD peptides14 were bound to streptavidin-coated magnetic beads (Dynabeads M-280; Dynal) in binding buffer (25 mM Tris-HCl, pH 8.0, 1 mM DTT, 5% (v/v) glycerol, 0.03% (v/v) Triton X-100 and 50 mM NaCl). Beads were saturated with biotinylated peptide and then washed with binding buffer. Bound proteins were eluted with 0.5M NaCl and analyzed by SDS-PAGE and immunoblotting using anti-Nrd1, anti-hemagglutinin (monoclonal 12CA5) or 6×His antibodies (monoclonal; BD Bioscience Clontech). Polyclonal rabbit antiserum against Nrd1 was from D. Brow and E. Steinmetz20. Monoclonal mouse antibody 2F12 against Nab3 was from M. Swanson51 via J. Corden.

Northern blotting and chromatin immunoprecipitation

We performed northern blotting as previously described13,21. Primers for ChIP and to generate probes for snR13 and snR33 detection were also previously described13. ChIP experiments were performed according to ref. 52.

Fluorescence anisotropy measurements

Measurements were carried out in a fluorescence spectrometer in T-configuration (Model FL322, Jobin Yvon) at 10°C. Samples were excited with vertically polarized light at 477 nm, and both vertical and horizontal emissions were recorded at 525 nm. To avoid any effects caused by N-terminal labeling of peptides, the assay was designed as Nrd1 titration experiments in which 2 μM of 5,6-carboxyfluorescein-labeled Ser5P peptides were competed with unlabeled peptide (up to 100 μM for measurements with wild-type $Nrd1_{6–151}$, $Nrd1_{6–214}$, or a complex of $Nrd1_{6–224}$ and $Nab3_{204–248}$ and 200 μM for measurements with mutated variants of $Nrd1_{6–151}$). Reactions were performed in a buffer containing 50 mM HEPES, pH 7.5, 150 mM NaCl, 1 mM EDTA and 1 mM DTE. Data were fitted to the cubic equation applying a 1:1 competitive binding mode as described53 Estimation of the K_d standard errors from reference and competition experiments were obtained by data set resampling54. We refitted 100 sample data sets including the uncertainties of experimental data, protein and peptide concentrations, for which the variance could be estimated experimentally.

Crystallization and data collection

Crystals of $Nrd1_{6–151}$ and selenomethionine-labeled $Nrd1_{6151(L37M\ L77M)}$ were grown at 20°C by the hanging drop vapor diffusion method with a reservoir solution containing 100 mM sodium citrate buffer, pH 5.5, 1.4 M $(NH_4)_2SO_4$. Single individuals grew within 1 week

to a size of 0.25 × 0.25 × 0.20 mm^3. Crystals were transferred into the reservoir solution with additional 20% (v/v) glycerol and were flash-cooled in liquid nitrogen. Synchrotron diffraction data were collected at the beamline X10SA, SLS, Villigen. Data were complete to 2.1 - resolution for native $Nrd1_{6–151}$. MAD data were collected from a selenomethionine-labeled $Nrd1_{6–151(L37M\ L77M)}$ protein crystal to 2.9-Å resolution. Data were processed with XDS55. Crystals belong to space group $P3_221$ with unit cell dimensions a = 80.20 Å , b = 80.20 Å and c = 62.97 Å , and contain one molecule per asymmetric unit.

Structure determination

Selenium sites of $Nrd1_{6–151(L37M\ L77M)}$ data were located with SOLVE56 at a resolution of 2.9 Å , and phases were improved with RESOLVE57. A preliminary model was built with O58 and refined with CNS59 Phase extension to native data with a resolution of 2.1 Å was performed following a rigid body protocol59 with phases derived from the preliminary model. In cycles of manual building and refinement with REFMAC60, the model for $Nrd1_{6–151}$ was further improved. The refined model has excellent stereochemical quality and an R-factor of 19.3% for the working set. Statistics for data quality and refinement are given in Table 2.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We thank J. Corden (Johns Hopkins University), D. Libri (Centre National de la Recherche Scientifique) and E. Steinmetz and D. Brow (Univeristy of Wisconsin, Madison) for yeast strains, plasmids and antibodies. We also thank D. Libri, J. Corden, D. Brow, R. Shoeman, Y. Groemping, J. Reinstein, B. Loll and I. Schlichting for helpful discussions, encouragement and support. We are grateful to M. Gebhardt for technical support, I. Vetter for support of the crystallographic software, W. Blankenfeldt for help during data collection, and the scientific staff for support at the beamline X10SA, Paul Scherrer Institute (Villigen, Switzerland). This research was supported by grants to S.B. from the US National Institutes of Health and to A.M. from the German Research Foundation. M.K. is supported by the Charles A. King Trust Postdoctoral Fellowship. L.V. is a recipient of a Special Fellowship from the Leukemia and Lymphoma Society.

References

1. Buratowski S. Connections between mRNA 3′ end processing and transcription termination. Curr. Opin. Cell Biol. 2005; 17:257–261. [PubMed: 15901494]
2. Li X, Manley JL. Cotranscriptional processes and their influence on genome stability. Genes Dev. 2006; 20:1838–1847. [PubMed: 16847344]
3. Buratowski S. The CTD code. Nat. Struct. Biol. 2003; 10:679–680. [PubMed: 12942140]
4. Komarnitsky P, Cho EJ, Buratowski S. Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. Genes Dev. 2000; 14:2452–2460. [PubMed: 11018013]
5. Schroeder SC, Schwer B, Shuman S, Bentley D. Dynamic association of capping enzymes with transcribing RNA polymerase II. Genes Dev. 2000; 14:2435–2440. [PubMed: 11018011]
6. Hampsey M, Reinberg D. Tails of intrigue: phosphorylation of RNA polymerase II mediates histone methylation. Cell. 2003; 113:429–432. [PubMed: 12757703]
7. Ahn SH, Kim M, Buratowski S. Phosphorylation of serine 2 within the RNA polymerase II C-terminal domain couples transcription and 3′ end processing. Mol. Cell. 2004; 13:67–76. [PubMed: 14731395]

8. Bird G, Zorio DA, Bentley DL. RNA polymerase II carboxy-terminal domain phosphorylation is required for cotranscriptional pre-mRNA splicing and 3′-end formation. Mol. Cell. Biol. 2004; 24:8963–8969. [PubMed: 15456870]
9. Cho EJ, Kobor MS, Kim M, Greenblatt J, Buratowski S. Opposing effects of Ctk1 kinase and Fcp1 phosphatase at Ser 2 of the RNA polymerase II C-terminal domain. Genes Dev. 2001; 15:3319–3329. [PubMed: 11751637]
10. Licatalosi DD, et al. Functional interaction of yeast pre-mRNA 3′ end processing factors with RNA polymerase II. Mol. Cell. 2002; 9:1101–1111. [PubMed: 12049745]
11. Meinhart A, Cramer P. Recognition of RNA polymerase II carboxy-terminal domain by 3′-RNA-processing factors. Nature. 2004; 430:223–226. [PubMed: 15241417]
12. Noble CG, et al. Key features of the interaction between Pcf11 CID and RNA polymerase II CTD. Nat. Struct. Mol. Biol. 2005; 12:144–151. [PubMed: 15665873]
13. Kim M, et al. Distinct pathways for snoRNA and mRNA termination. Mol. Cell. 2006; 24:723–734. [PubMed: 17157255]
14. Kim M, et al. The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. Nature. 2004; 432:517–522. [PubMed: 15565157]
15. West S, Gromak N, Proudfoot NJ. Human 5′ → 3′ exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. Nature. 2004; 432:522–525. [PubMed: 15565158]
16. Carroll KL, Pradhan DA, Granek JA, Clarke ND, Corden JL. Identification of cis elements directing termination of yeast nonpolyadenylated snoRNA transcripts. Mol. Cell. Biol. 2004; 24:6241–6252. [PubMed: 15226427]
17. Steinmetz EJ, Conrad NK, Brow DA, Corden JL. RNA-binding protein Nrd1 directs poly(A)-independent 3′-end formation of RNA polymerase II transcripts. Nature. 2001; 413:327–331. [PubMed: 11565036]
18. Conrad NK, et al. A yeast heterogeneous nuclear ribonucleoprotein complex associated with RNA polymerase II. Genetics. 2000; 154:557–571. [PubMed: 10655211]
19. Steinmetz EJ, Brow DA. Repression of gene expression by an exogenous sequence element acting in concert with a heterogeneous nuclear ribonucleoprotein-like protein, Nrd1, and the putative helicase Sen1. Mol. Cell. Biol. 1996; 16:6993–7003. [PubMed: 8943355]
20. Steinmetz EJ, Brow DA. Control of pre-mRNA accumulation by the essential yeast protein Nrd1 requires high-affinity transcript binding and a domain implicated in RNA polymerase II association. Proc. Natl. Acad. Sci. USA. 1998; 95:6699–6704. [PubMed: 9618475]
21. Vasiljeva L, Buratowski S. Nrd1 interacts with the nuclear exosome for 3′ processing of RNA polymerase II transcripts. Mol. Cell. 2006; 21:239–248. [PubMed: 16427013]
22. Houseley J, LaCava J, Tollervey D. RNA-quality control by the exosome. Nat. Rev. Mol. Cell Biol. 2006; 7:529–539. [PubMed: 16829983]
23. Steinmetz EJ, et al. Genome-wide distribution of yeast RNA polymerase II and its control by Sen1 helicase. Mol. Cell. 2006; 24:735–746. [PubMed: 17157256]
24. David L, et al. A high-resolution map of transcription in the yeast genome. Proc. Natl. Acad. Sci. USA. 2006; 103:5320–5325. [PubMed: 16569694]
25. Davis CA, Ares M Jr. Accumulation of unstable promoter-associated transcripts upon loss of the nuclear exosome subunit Rrp6p in *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA. 2006; 103:3262–3267. [PubMed: 16484372]
26. Houalla R, et al. Microarray detection of novel nuclear RNA substrates for the exosome. Yeast. 2006; 23:439–454. [PubMed: 16652390]
27. Samanta MP, Tongprasit W, Sethi H, Chin CS, Stolc V. Global identification of noncoding RNAs in *Saccharomyces cerevisiae* by modulating an essential RNA processing pathway. Proc. Natl. Acad. Sci. USA. 2006; 103:4192–4197. [PubMed: 16537507]
28. Arigo JT, Eyler DE, Carroll KL, Corden JL. Termination of cryptic unstable transcripts is directed by yeast RNA-binding proteins Nrd1 and Nab3. Mol. Cell. 2006; 23:841–851. [PubMed: 16973436]
29. Lykke-Andersen S, Jensen TH. CUT it out: silencing of noise in the transcriptome. Nat. Struct. Mol. Biol. 2006; 13:860–861. [PubMed: 17021618]

30. Thiebaut M, Kisseleva-Romanova E, Rougemaille M, Boulay J, Libri D. Transcription termination and nuclear degradation of cryptic unstable transcripts: a role for the Nrd1-Nab3 pathway in genome surveillance. Mol. Cell. 2006; 23:853–864. [PubMed: 16973437]
31. Wyers F, et al. Cryptic Pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. Cell. 2005; 121:725–737. [PubMed: 15935759]
32. Vanacova S, Stef R. The exosome and RNA quality control in the nucleus. EMBO Rep. 2007; 8:651–657. [PubMed: 17603538]
33. Meinhart A, Kamenski T, Hoeppner S, Baumli S, Cramer P. A structural perspective of CTD function. Genes Dev. 2005; 19:1401–1415. [PubMed: 15964991]
34. Yuryev A, et al. The C-terminal domain of the largest subunit of RNA polymerase II interacts with a novel set of serine/arginine-rich proteins. Proc. Natl. Acad. Sci. USA. 1996; 93:6975–6980. [PubMed: 8692929]
35. Barilla D, Lee BA, Proudfoot NJ. Cleavage/polyadenylation factor IA associates with the carboxyl-terminal domain of RNA polymerase II in *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA. 2001; 98:445–450. [PubMed: 11149954]
36. Kim M, Ahn SH, Krogan NJ, Greenblatt JF, Buratowski S. Transitions in RNA polymerase II elongation complexes at the 3′ ends of genes. EMBO J. 2004; 23:354–364. [PubMed: 14739930]
37. Sadowski M, Dichtl B, Hubner W, Keller W. Independent functions of yeast Pcf11p in pre-mRNA 3′ end processing and in transcription termination. EMBO J. 2003; 22:2167–2177. [PubMed: 12727883]
38. Fabrega C, Shen V, Shuman S, Lima CD. Structure of an mRNA capping enzyme bound to the phosphorylated carboxy-terminal domain of RNA polymerase II. Mol. Cell. 2003; 11:1549–1561. [PubMed: 12820968]
39. Verdecia MA, Bowman ME, Lu KP, Hunter T, Noel JP. Structural basis for phosphoserine-proline recognition by group IV WW domains. Nat. Struct. Biol. 2000; 7:639–643. [PubMed: 10932246]
40. Carroll KL, Ghirlando R, Ames JM, Corden JL. Interaction of yeast RNA-binding proteins Nrd1 and Nab3 with RNA polymerase II terminator elements. RNA. 2007; 13:361–373. [PubMed: 17237360]
41. Arigo JT, Carroll KL, Ames JM, Corden JL. Regulation of yeast NRD1 expression by premature transcription termination. Mol. Cell. 2006; 21:641–651. [PubMed: 16507362]
42. Zhang Z, Fu J, Gilmour DS. CTD-dependent dismantling of the RNA polymerase II elongation complex by the pre-mRNA 3′-end processing factor, Pcf11. Genes Dev. 2005; 19:1572–1580. [PubMed: 15998810]
43. Nedea 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 nucleolar RNA 3′-ends. J. Biol. Chem. 2003; 278:33000–33010. [PubMed: 12819204]
44. Steinmetz EJ, Ng SB, Cloute JP, Brow DA. *Cis*- and *trans*-acting determinants of transcription termination by yeast RNA polymerase II. Mol. Cell. Biol. 2006; 26:2688–2696. [PubMed: 16537912]
45. Gudipati RK, Villa T, Boulay J, Libri D. Phosphorylation of the RNA polymerase II C-terminal domain dictates transcription termination choice. Nat. Struct. Mol. Biol. 2008 Jul 27. advance online publication.
46. Kumaki Y, Matsushima N, Yoshida H, Nitta K, Hikichi K. Structure of the YSPTSPS repeat containing two SPXX motifs in the CTD of RNA polymerase II: NMR studies of cyclic model peptides reveal that the SPTS turn is more stable than SPSY in water. Biochim. Biophys. Acta. 2001; 1548:81–93. [PubMed: 11451441]
47. Becker R, Loll B, Meinhart A. Snapshots of the RNA processing factor SCAF8 bound to different phosphorylated forms of the carboxy-terminal domain of RNA-polymerase II. J. Biol. Chem. 2008 Jun 11. published online.
48. Kapranov P, Willingham AT, Gingeras TR. Genome-wide transcription and the implications for genomic organization. Nat. Rev. Genet. 2007; 8:413–423. [PubMed: 17486121]
49. Van Duyne GD, Standaert RF, Karplus PA, Schreiber SL, Clardy J. Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. J. Mol. Biol. 1993; 229:105–124. [PubMed: 7678431]

50. Pei Y, Hausmann S, Ho CK, Schwer B, Shuman S. The length, phosphorylation state, and primary structure of the RNA polymerase II carboxyl-terminal domain dictate interactions with mRNA capping enzymes. J. Biol. Chem. 2001; 276:28075–28082. [PubMed: 11387325]
51. Wilson SM, Datar KV, Paddy MR, Swedlow JR, Swanson MS. Characterization of nuclear polyadenylated RNA-binding proteins in *Saccharomyces cerevisiae*. J. Cell Biol. 1994; 127:1173–1184. [PubMed: 7962083]
52. Keogh MC, Buratowski S. Using chromatin immunoprecipitation to map cotranscriptional mRNA processing in *Saccharomyces cerevisiae*. Methods Mol. Biol. 2004; 257:1–16. [PubMed: 14769992]
53. Reinstein J, et al. Fluorescence and NMR investigations on the ligand binding properties of adenylate kinases. Biochemistry. 1990; 29:7440–7450. [PubMed: 2223775]
54. Efron, B. Society of Industrial and Applied Mathematics CBMS-NSF Monographs. Vol. Vol 38. New York, NY,: Cambridge University Press; 1982. The Jacknife, the Bootstrap, and other resampling plans.
55. Kabsch W. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. J. Appl. Crystallogr. 1993; 26:795–800.
56. Terwilliger TC, Berendzen J. Automated MAD and MIR structure solution. Acta Crystallogr D Biol. Crystallogr. 1999; 55:849–861. [PubMed: 10089316]
57. Terwilliger TC. Automated structure solution, density modification and model building. Acta Crystallogr. D Biol. Crystallogr. 2002; 58:1937–1940. [PubMed: 12393925]
58. Jones TA, Zou JY, Cowan SW, Kjeldgaard M. Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A. 1991; 47:110–119. [PubMed: 2025413]
59. Brunger AT, et al. Crystallography & NMR system: a new software suite for macro-molecular structure determination. Acta Crystallogr. D Biol. Crystallogr. 1998; 54:905–921. [PubMed: 9757107]
60. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D Biol. Crystallogr. 1997; 53:240–255. [PubMed: 15299926]
61. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 1994; 22:4673–4680. [PubMed: 7984417]

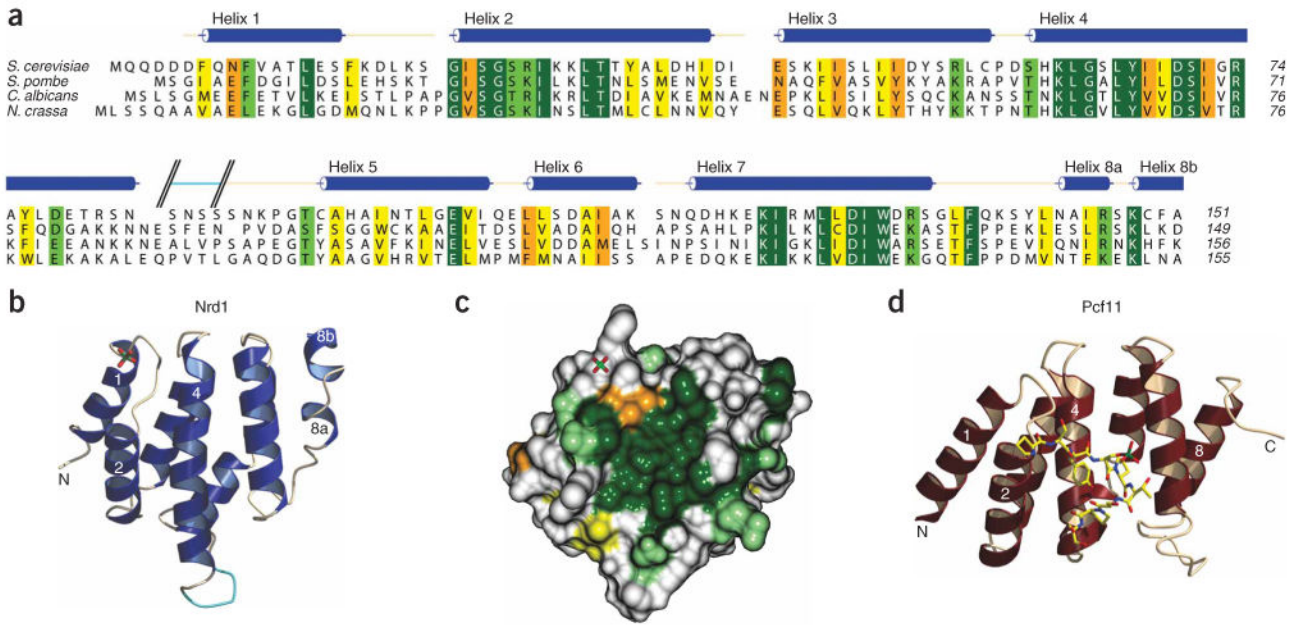

Figure 1.

Structure of the Nrd1 CID. (**a**) The CID regions of Nrd1 proteins from several yeast species were aligned with ClustalW61. Identical amino acid residues are shown in dark green, and declining sequence similarity is shown using light green, orange and yellow, in that order. Helices within the Nrd1 CID model are depicted in blue above the sequence alignment. Loop regions are shown in beige, and the loop region missing in the Nrd1 CID model in cyan. (**b**) Ribbon illustration of the Nrd1 CID polypeptide chain. The bound sulfate ion contacting the helix 1-helix 2 loop is represented as a stick model, and the missing loop between helices 4 and 5 is modeled as random coil in cyan. (**c**) Surface representation of Nrd1 CID colored according to the sequence conservation as in **a**. The bound sulfate ion is represented as a stick model. (**d**) Ribbon model of the Pcf11 CID domain (PDB 1SZA), with bound CTD-Ser2P represented as a stick model11.

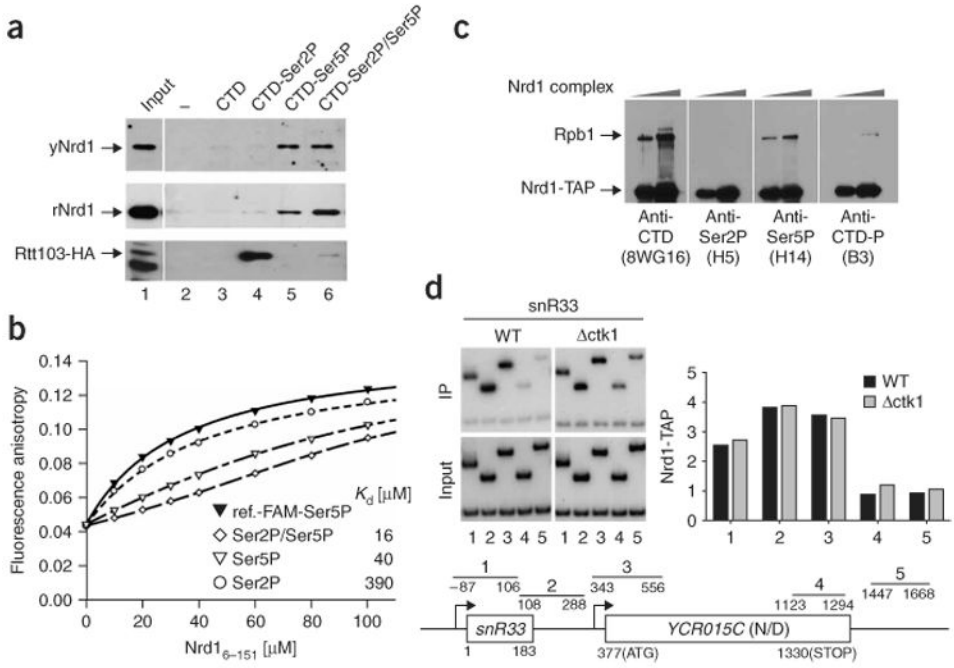

Figure 2.

Nrd1 binds preferentially to CTD-Ser5P. (**a**) Binding to four repeat CTD peptides *in vitro*. Unmodified, Ser2P, Ser5P or Ser2P/Ser5P peptides were immobilized on streptavidin-conjugated magnetic beads and incubated with 5 μg of recombinant Nrd1 (rNrd1), 10 ng of TAP-purified yeast Nrd1 complex (yNRD1) or 500 μg of whole-cell extract from an Rtt103-hemagglutinin (HA) strain (YSB815). Bound proteins were eluted, separated by SDS-PAGE and detected by immunoblotting using either anti-Nrd1 or anti-HA antibodies. Recombinant Rtt103 also specifically bound to CTD-Ser2P (not shown). (**b**) $Nrd1_{6–151}$ was titrated with fluorescently labeled CTD-Ser5P (two repeats) and binding was measured by fluorescence anisotropy (black triangles; ref.-FAM, 5,6-carboxyfluorescein labeled reference). The same experiment was then done in the presence of competing unlabeled CTD-Ser2P (circles), CTDSer5P (white triangles) or CTD-Ser2P/Ser5P (diamonds). Equilibrium dissociation constants (K_d) were calculated from the best fit to the data. (**c**) Nrd1 is associated with Ser5-phosphorylated Pol II *in vivo*. Nrd1 was purified via the TAP tag, and the phosphorylation status of the associated polymerase was monitored by immunoblotting using anti-CTD (8WG16), anti-Ser2P (H5), anti-Ser5P (H14) or an antibody that can recognize both Ser2P and Ser5P (B3)9. (**d**) Ctk1 kinase is not required for recruitment of Nrd1 to genes *in vivo*. Cross-linked chromatin was prepared from Nrd1-TAP–containing cells that were wild-type (WT) or deleted (Δctk1) for the *CTK1* gene. Following precipitation with IgG agarose, chromatin was amplified with primers across the *snR33* locus, as diagrammed below. Immunoprecipitated samples (IP) were compared against input chromatin (Input) and quantified (right). The upper band in each lane is the *snR33* product and the lower band is a nontranscribed control region. Similar results were obtained for the *PMA1* and *ADH1* genes (not shown).

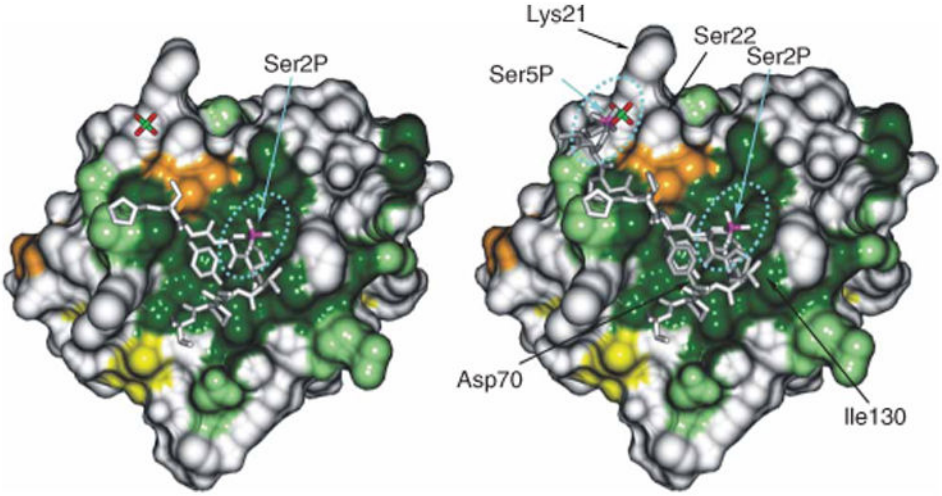

Figure 3.

Structural models for a Nrd1-CTD interaction. Left, the CTD-Ser2P peptide is modeled bound to Nrd1 by simple manual superposition of the Nrd1 CID onto the Pcf11 CTD. Right, the longer CTD model was created by superimposing the extended CTD-Ser5P bound to mRNA capping enzyme (PDB 1P16) onto the extended region of the CTD-Ser2P bound to Pcf11. The overlapping stretch of the two CTDs consists of Ser7-Tyr1-Ser2. Notably, the position of the phosphate moiety from Ser5P coincides with the bound sulfate ion observed in the Nrd1 crystal structure. Blue arrows point out phosphorylated CTD residues; black arrows show residues mutated and tested for effects on CTD binding in Table 1.

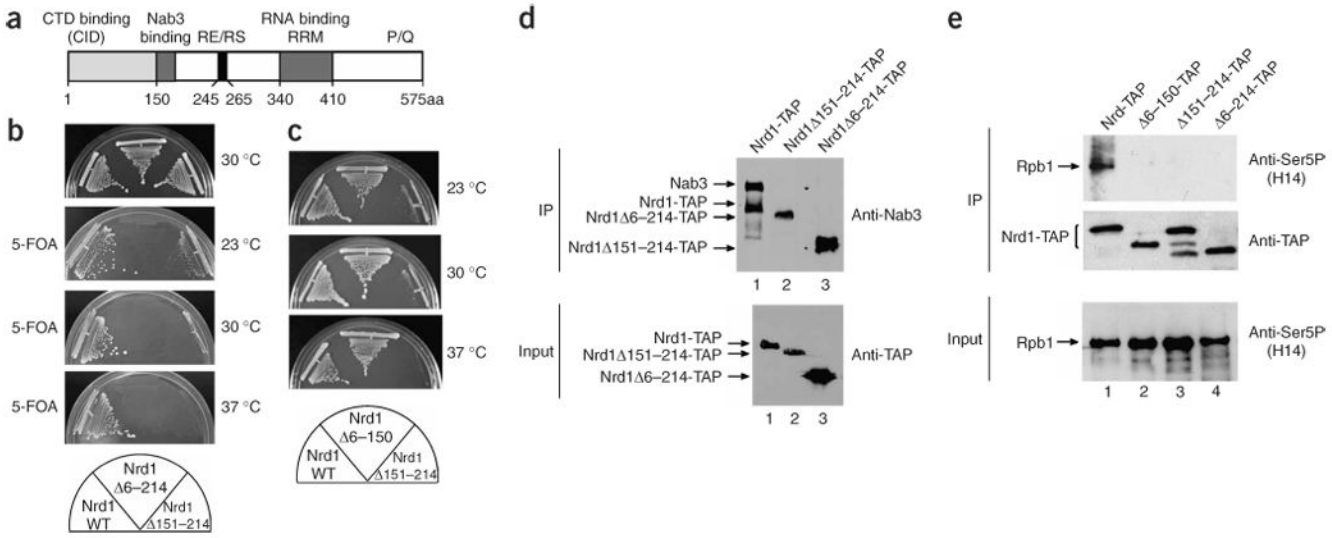

Figure 4.

The CTD and Nab3 interaction domains of Nrd1 are both important for interaction with Pol II. (**a**) Schematic diagram of Nrd1. RE/RS, arginine-, serine- and glutamate-rich region; P/Q, proline- and glutamine-rich region. (**b**) Phenotypic analysis of *nrd1Δ6–214* and *nrd1Δ151–214* deletions. The *NRD1* plasmid shuffling strain EJS101-9d was transformed with pJC580, pRS415-Nrd1Δ6–214 or pRS415-Nrd1Δ151–214, and the wild-type *NRD1/URA3* plasmid (pRS316-NRD1) was shuffled out on 5-fluoroorotic acid (5-FOA) medium at the indicated temperatures. (**c**) Phenotypic analysis of *nrd1Δ6–150* and *nrd1Δ151–214* alleles integrated into the genome. (**d**) Nrd1 residues 151–214 are responsible for interaction with Nab3. Expression of the wild-type (WT) and mutant TAP-tagged proteins Nrd1, Nrd1Δ6–214 and Nrd1Δ151–214 in extracts was monitored by immunoblotting for the Protein A module of the TAP tag (below, α-TAP). Nrd1 protein complexes were purified using IgG resin, and association with Nab3 was analyzed with anti-Nab3 antibody (above). Note that the Protein A module on Nrd1 reacts with the secondary antibody. (**e**) The Nab3 and CTD binding regions of Nrd1 contribute to its interaction with Pol II *in vivo*. Nrd1, Nrd1Δ6–150, Nrd1Δ151–214 or Nrd1Δ6–214 protein complexes were IgG-purified from whole-cell extracts and monitored for association with Pol II by immunoblotting with anti-CTD Ser5P (H14).

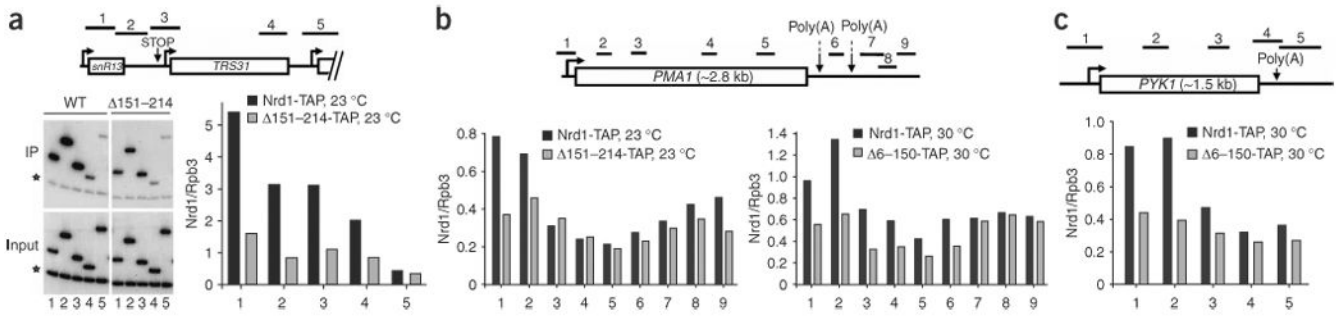

Figure 5.

Efficient Nrd1 recruitment to 5′ ends of Pol II–transcribed genes requires both the CID and Nab3 interaction domain. (**a**) Deletion of Nrd1 residues 151–214 reduces Nrd1 recruitment to the *SNR13* gene. Schematic representation of *snR13* is shown above. Coding regions are shown as boxes; arrows indicate promoters; numbered bars show positions of the PCR products used in ChIP analysis; asterisks indicate a control band amplified from a nontranscribed region of Chromosome V. ChIP (IP) results are shown in the upper panels, and PCR from non-IPed chromatin samples is shown below (Input). Quantification of the ChIP data are shown to the right. The same chromatin preparations were used to analyze Pol II subunit Rpb3. The *y*-axis of the graph shows Nrd1 levels as a ratio to Rpb3 levels. The *x*-axis refers to the numbered primer pairs. (**b,c**) ChIP analysis of full-length Nrd1, Nrd1 lacking the CID (Δ6–150), or Nrd1 lacking the Nab3 interaction region (Δ151–214) was carried out on the *PMA1* and *PYK1* genes as in **a**. Graphs show the levels of Nrd1 normalized to Rpb3 levels.

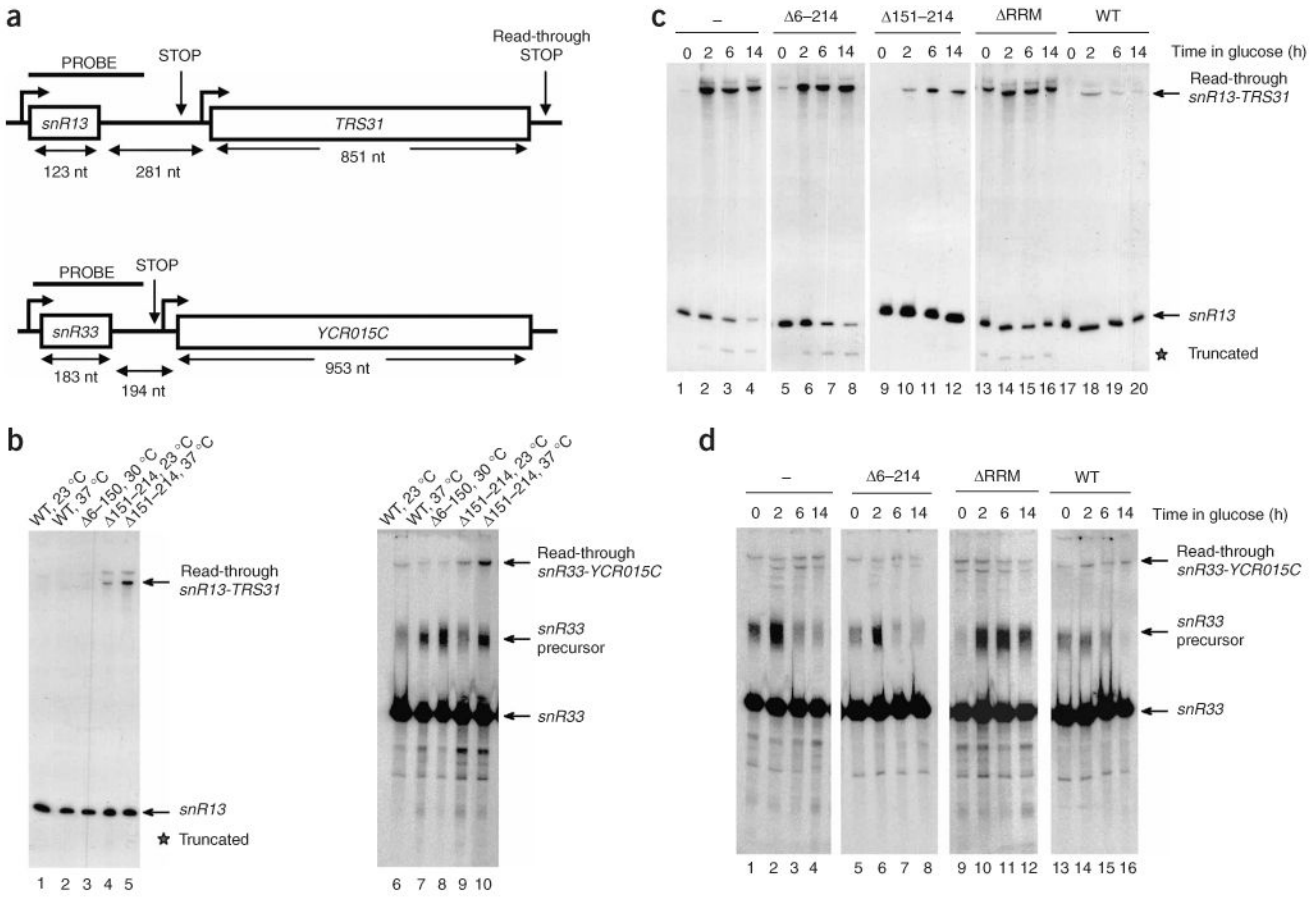

Figure 6.

Effect of Nrd1 deletions on snoRNA termination and processing. (**a**) Schematic diagram of *snR13* and *snR33*. Coding regions are shown as boxes, transcription start sites as bent arrows, and transcription termination regions are designated as STOP. The positions of the probes are shown above the gene. (**b**) Northern blot analysis of *snR13* and *snR33* RNA. Total RNA was isolated from *NRD1*, *nrd1Δ6–150* and *nrd1Δ151–214* cells and analyzed as described previously13,21. Positions of the transcription read-through transcripts, unprocessed precursor and mature snoRNAs are indicated with arrows; a previously observed truncated species is marked with an asterisk. (**c**) Northern blot analysis of *snR13* RNA from cells expressing Nrd1, Nrd1Δ151–214 and Nrd1Δ6– 214 proteins. In this experiment, wild-type Nrd1 is expressed from a galactose-inducible promoter, and depletion occurs upon shift to glucose media. The Nrd1 deletions are expressed from plasmids. The '–' panel shows cells carrying only vector. (**d**) Expression of the *snR33* gene was analyzed as in **c**.

Table 1

Equilibrium K^d values for Nrd1 and mutated variants*

	Double-repeat Ser5P [μM]	Double-repeat Ser2P [μM]	Double-repeat Ser2/Ser5P [μM]	Tetra-repeat Ser5P [μM]
$Nrd1_{6–151}$ wild-type	40 (+3/−3)	390 (+30/−30)	16 (+1/−1)	39 (+3/−3)
$Nrd1_{6–214}$ wild-type	48 (+3/−3)	n.d.	n.d.	31 (+1/−2)
$Nrd1_{6–224}$ $Nab3_{204–238}$	n.d.	n.d.	n.d.	42 (+2/−3)
$Nrd1_{6–151}$ L20D	39 (+3/−3)	560 (+70/−100)	19 (+1/−1)	n.d.
$Nrd1_{6–151}$ K21P	70 (+10/−10)	380 (+70/−70)	48 (+6/−6)	n.d.
$Nrd1_{6–151}$ K21D	95 (+10/−15)	610 (+90/−80)	70 (+10/−10)	n.d.
$Nrd1_{6–151}$ S22D	140 (+20/−15)	470 (+80/−60)	100 (+15/−10)	n.d.
$Nrd1_{6–151}$ D70R	680 (+90/−90)	n.d.	100 (+20/−10)	n.d.
$Nrd1_{6–151}$ I130R	215 (+30/−30)	550 (+60/−80)	27 (+2/−2)	n.d.

*Note that comparisons of K_d above 400 μM with this assay have large confidence intervals, and differences above this number are unlikely to be relevant *in vivo*. n.d., not determined.

Table 2

Data collection, phasing and refinement statistics for MAD (SeMet) structures

	$Nrd1_{6–151}$ (native)	**$Nrd1_{6–151\ (L37M\ L77M)}$ (MAD)**		
Data collection				
Space group	$P3_221$		$P3_221$	
Cell dimensions				
a, *b*, *c* (Å)	80.20, 80.20, 62.97		80.49, 80.49, 62.42	
α, β, γ (°)	90, 90, 120		90, 90, 120	
		Peak	*Inflection*	*Remote*
Wavelength	1.07176	0.97949	0.97904	0.92524
Resolution (Å)	20.0–2.1 (2.15–2.10)	20.00–2.90 (2.95–2.90)	20.00–2.90 (2.95–2.90)	20.00–2.90 (2.95–2.90)
R_{sym}	3.5 (23.6)	6.2 (24.6)	6.2 (24.8)	4.8 (8.8)
$I / \sigma I$	19.3 (4.7)	27.7 (10.0)	27.8 (10.2)	18.2 (11.1)
Completeness (%)	96.2 (82.1)	99.6 (99.4)	99.5 (99.0)	97.4 (99.2)
Redundancy	3	11	11	3
Refinement				
Resolution (Å)	20.0–2.1 (2.15–2.10)			
No. reflections	13,456 (752)			
R_{work} / R_{free}	19.3 / 22.2			
No. atoms				
Protein	1175			
Ligand/ion	29			
Water	56			
B-factors				
Protein	46.7			
Ligand/ion	67.6			
Water	42.8			
r.m.s. deviations				
Bond lengths (Å)	0.021			
Bond angles (°)	1.714			

*Values in parentheses are for highest-resolution shell.