Structure and Carboxyl-terminal Domain (CTD) Binding of the Set2 SRI Domain That Couples Histone H3 Lys\textsuperscript{36} Methylation to Transcription\*\textsuperscript{1}

During mRNA elongation, the SRI domain of the histone H3 methyltransferase Set2 binds to the phosphorylated carboxyl-terminal domain (CTD) of RNA polymerase II. The solution structure of the yeast Set2 SRI domain reveals a novel CTD-binding fold consisting of a left-handed three-helix bundle. NMR titration shows that the SRI domain binds an Ser\textsuperscript{2}/Ser\textsuperscript{5}-phosphorylated CTD peptide comprising two heptapeptide repeats and three flanking NH\textsubscript{2}-terminal residues, whereas a single CTD repeat is insufficient for binding. Residues that show strong chemical shift perturbations upon CTD binding cluster in two regions. Both CTD tyrosine side chains contact the SRI domain. One of the tyrosines binds in the region with the strongest chemical shift perturbations, formed by the two NH\textsubscript{2}-terminal helices. Unexpectedly, the SRI domain fold resembles the structure of an RNA polymerase-interacting domain in bacterial C factors (domain \(\alpha_5\) in \(\alpha\)-helix).

Gene transcription by RNA polymerase II (Pol II) is physically and functionally coupled to other nuclear events, most notably mRNA processing (1–7). Transcription-coupled events generally depend on the carboxyl-terminal repeat domain (CTD)\textsuperscript{4} of the largest Pol II subunit, which binds many nuclear factors during transcription elongation. The CTD forms a mobile extension from the structural core of Pol II (8) and consists of heptapeptide repeats and three flanking NH\textsubscript{2}-terminal residues, whereas a single CTD repeat is insufficient for binding. Residues that show strong chemical shift perturbations upon CTD binding cluster in two regions. Both CTD tyrosine side chains contact the SRI domain. One of the tyrosines binds in the region with the strongest chemical shift perturbations, formed by the two NH\textsubscript{2}-terminal helices. Unexpectedly, the SRI domain fold resembles the structure of an RNA polymerase-interacting domain in bacterial C factors (domain \(\alpha_5\) in \(\alpha\)-helix).

Gene transcription by RNA polymerase II (Pol II) is physically and functionally coupled to other nuclear events, most notably mRNA processing (1–7). Transcription-coupled events generally depend on the carboxyl-terminal repeat domain (CTD)\textsuperscript{4} of the largest Pol II subunit, which binds many nuclear factors during transcription elongation. The CTD forms a mobile extension from the structural core of Pol II (8) and consists of heptapeptide repeats and three flanking NH\textsubscript{2}-terminal residues, whereas a single CTD repeat is insufficient for binding. Residues that show strong chemical shift perturbations upon CTD binding cluster in two regions. Both CTD tyrosine side chains contact the SRI domain. One of the tyrosines binds in the region with the strongest chemical shift perturbations, formed by the two NH\textsubscript{2}-terminal helices. Unexpectedly, the SRI domain fold resembles the structure of an RNA polymerase-interacting domain in bacterial C factors (domain \(\alpha_5\) in \(\alpha\)-helix).

Gene transcription by RNA polymerase II (Pol II) is physically and functionally coupled to other nuclear events, most notably mRNA processing (1–7). Transcription-coupled events generally depend on the carboxyl-terminal repeat domain (CTD)\textsuperscript{4} of the largest Pol II subunit, which binds many nuclear factors during transcription elongation. The CTD forms a mobile extension from the structural core of Pol II (8) and consists of heptapeptide repeats and three flanking NH\textsubscript{2}-terminal residues, whereas a single CTD repeat is insufficient for binding. Residues that show strong chemical shift perturbations upon CTD binding cluster in two regions. Both CTD tyrosine side chains contact the SRI domain. One of the tyrosines binds in the region with the strongest chemical shift perturbations, formed by the two NH\textsubscript{2}-terminal helices. Unexpectedly, the SRI domain fold resembles the structure of an RNA polymerase-interacting domain in bacterial C factors (domain \(\alpha_5\) in \(\alpha\)-helix).

Gene transcription by RNA polymerase II (Pol II) is physically and functionally coupled to other nuclear events, most notably mRNA processing (1–7). Transcription-coupled events generally depend on the carboxyl-terminal repeat domain (CTD)\textsuperscript{4} of the largest Pol II subunit, which binds many nuclear factors during transcription elongation. The CTD forms a mobile extension from the structural core of Pol II (8) and consists of heptapeptide repeats and three flanking NH\textsubscript{2}-terminal residues, whereas a single CTD repeat is insufficient for binding. Residues that show strong chemical shift perturbations upon CTD binding cluster in two regions. Both CTD tyrosine side chains contact the SRI domain. One of the tyrosines binds in the region with the strongest chemical shift perturbations, formed by the two NH\textsubscript{2}-terminal helices. Unexpectedly, the SRI domain fold resembles the structure of an RNA polymerase-interacting domain in bacterial C factors (domain \(\alpha_5\) in \(\alpha\)-helix).

* This work was supported in part by the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

\(\alpha\)-helixes in a left-handed bundle (Fig. 1). The NH\textsubscript{2}-terminal helix \(\alpha_1\) is slightly kinked at residues Phe\textsuperscript{650} and Val\textsuperscript{651}, and the linker between helices \(\alpha_1\) and \(\alpha_2\) includes a short \(\beta\)-helical turn at residues Ser\textsuperscript{656}–
Gln<sup>652</sup>. A hydrophobic core is formed by numerous residues located at the inter-
face between the three helices, including four residues in the two regions linking
the helices (Fig. 1<sup>C</sup>). Consistently the heteronuclear <sup>1</sup>H-<sup>15</sup>N NOE measurements
demonstrate that the polypeptide backbone in all three helices and the connecting
linker regions is rigid (Fig. 1<sup>C</sup> and supplemental Fig. S1). The hydrophobic core
residues are generally conserved across species (Fig. 1<sup>C</sup>), demonstrating that our
structure is a good model for SRI domains in Set2 of other species.

The SRI Domain Defines a Novel CTD-binding Fold—Comparison with the
five known structures of CTD-binding domains reveals that the SRI domain
defines a novel CTD-binding fold. Other CTD-binding domains include FF
domains, CTD-interacting domains, WW domains, BRCT domains, and a
domain in the Cgt1 subunit of the 5<sup>′</sup>-H1 capping enzyme (reviewed in Ref. 7). Of
these, FF and CTD-interacting domains also form helical bundles (33, 34), but, in
contrast to the SRI domain, the superhelical arrangement in these two domains is
right-handed (supplemental Fig. S2). Thus the six CTD-binding domains that
have been structurally characterized use different folds for specific CTD
recognition.

The SRI Domain Binds a Two-repeat CTD Phosphopeptide—To character-
ize the CTD-binding determinants of the SRI domain, we performed NMR
titration experiments with Ser<sup>2</sup>/Ser<sup>5</sup>-phosphorylated CTD peptides (Fig. 1C). A phosphopeptide consisting of a single CTD repeat (YpSPTpSPS, pS =
phosphoserine; Fig. S3<sup>A</sup>) did not perturb chemical shifts in a two-dimensional
<sup>1</sup>H,<sup>15</sup>N HSQC spectrum, indicating that there is no significant binding (data
not shown). However, titration with a peptide that comprised two CTD
repeats and three flanking NH<sub>2</sub>-terminal residues (SPS-YpSPTpSPS-YpSPTpSPS)
resulted in many strong chemical shift perturbations (Fig. 1<sup>C</sup> and supple-
mental Fig. S3). From the titration data the dissociation constant is estimated
to be in the low micromolar range, comparable with the reported approximate
affinity of 6 μM for a CTD phosphopeptide comprising three repeats (22).

Regions in the SRI Domain That Interact with the CTD—Residues that show
strong chemical shift perturbations of their backbone NH groups cluster in
two regions on the SRI domain structure (Fig. 2A). The first region includes
residues Lys<sup>634</sup>, Phe<sup>635</sup> in α<sub>1</sub>, and Ala<sup>662</sup>, Val<sup>666</sup>, Lys<sup>667</sup>, Thr<sup>670</sup>, Thr<sup>671</sup>, and
Glu<sup>673</sup> in α<sub>2</sub>-α<sub>3</sub> linker, whereas the second region includes residues Phe<sup>653</sup>, His<sup>655</sup>,
Glu<sup>656</sup> in the α<sub>1</sub>-α<sub>2</sub> linker, and residue Ile<sup>705</sup> in α<sub>3</sub> (Figs. 1C and 2A and
supplemental Fig. S3). With the exception of Ile<sup>705</sup>, the strongest perturbations
upon peptide binding were observed in region 1 (Phe<sup>653</sup>, Ala<sup>662</sup>, Val<sup>666</sup>, Lys<sup>667</sup>,
and Glu<sup>673</sup>). In this region, the side chain NH<sub>2</sub> groups of residues Asn<sup>657</sup> and

FIGURE 1. Structure and CTD binding of the yeast Set2 SRI domain. A, ensemble of final NMR structures. The three α-helices are shown in green, and a short 3<sub>10</sub>-helix is shown in pink. B, ribbon diagram of the lowest energy structure in A. C, alignment of SRI domain sequences and NMR structure determination and CTD binding data. The secondary structure is shown above the sequence. Solvent-protected amide protons that show slow H/D exchange are indicated by filled circles; Secondary chemical shifts Δδ(Cα-Cβ) are indicated by black bars. Residues that experience large chemical shift perturbations upon addition of the CTD two-repeat phosphopeptide SPS-YpSPTpSPS-YpSPTpSPS (pS = phosphoserine) are indicated above the alignment with crosses and circled crosses for backbone and side chain amides, respectively. Yellow stars indicate residues Ala<sup>662</sup> and Val<sup>666</sup> that are implicated in binding of a CTD tyrosine side chain. Residues that are identical and conserved in fungal Set2 homologues are on red background and in red, respectively. Hydrophobic core residues are marked with a black square.

ACCELERATED PUBLICATION: Structure and CTD Binding of the Set2 SRI Domain

14 JOURNAL OF BIOLOGICAL CHEMISTRY
VOLUME 281 • NUMBER 1 • JANUARY 6, 2006
domain. We speculate that the eukaryotic TFIIEα NH2-terminal domain, which may contact promoter DNA, and the Set2 SRI domain, which binds the negatively charged phospho-CTD, both evolved from the bacterial σ7 factor.

Acknowledgments—We thank C. Buchen, L. Lariviére, and other members of the Cramer laboratory (Gene Center Munich) for help. We thank G. Stier and A. Lingel (EMBL, Heidelberg, Germany) and K. Kizer (University of North Carolina) for help.

REFERENCES

1. Dahmus, M. E. (1995) Biochim. Biophys. Acta 1261, 171–182
2. Palancade, B., and Bensaude, O. (2003) Eur. J. Biochem. 270, 3859–3870
3. Sims, R. J., III, Mandal, S. S., and Reinberg, D. (2004) Curr. Opin. Cell Biol. 16, 263–271
4. Zorio, D. A., and Bentley, D. L. (2004) Exp. Cell Res. 296, 91–97
5. Proudfoot, N. (2004) Curr. Opin. Cell Biol. 16, 272–278
6. Maniatis, T., and Reed, R. (2002) Nature 416, 499–506
7. Meinhart, A., Kamenski, T., Hoeppner, S., Baumbi, S., and Cramer, P. (2005) Genes Dev. 19, 1401–1415
8. Cramer, P., Bushnell, D. A., and Kornberg, R. D. (2001) Science 292, 1863–1876
9. Ho, C. K., Sriskanda, V., McCracken, S., Bentley, D., Schwer, B., and Shuman, S. (1988) J. Biol. Chem. 263, 1419–1425
10. Komarnitsky, P., Cho, E. J., and Buratowski, S. (2000) Genes Dev. 14, 2452–2460
11. McCracken, S., Fong, N., Rosonina, E., Yankulov, K., Brothers, G., Siderovski, D., Hessel, A., Foster, S., Shuman, S., and Bentley, D. L. (1997) Genes Dev. 11, 3306–3318
12. Cho, E. J., Takagi, T., Moore, C. R., and Buratowski, S. (1997) Genes Dev. 11, 3319–3326
13. Cho, E. J., Kobor, M. S., Kim, M., Greenblatt, J. F., and Buratowski, S. (2001) Genes Dev. 15, 3319–3329
14. Gerber, M., and Shilatifard, A. (2003) J. Biol. Chem. 278, 26303–26306
15. Hampsry, M., and Reinberg, D. (2003) Cell 113, 429–441
16. Morillon, A., Karabetsou, N., Nair, A., and Mellor, J. (2005) Mol. Cell 18, 733–734
17. Krogan, N. J., Kim, M., Tong, A., Golshani, A., Cagney, G., Canadien, V., Richards, D. P., Beattie, B. K., Emili, A., Boone, C., Shilatifard, A., Buratowski, S., and Greenblatt, J. (2003) Mol. Cell Biol. 23, 4207–4213
18. Ng, H. H., Robert, F., Young, R. A., and Struhl, K. (2003) Mol. Cell 11, 709–719
19. Strahl, B. D., Grant, P. A., Briggs, S. D., Sun, Z. W., Bone, J. R., Caldwell, J. A., Mollah, S. C., Cook, R. G., Shabahathan, J., Hunt, D. F., and Allis, C. D. (2002) Mol. Cell Biol. 22, 1298–1306
20. Xiao, T., Hall, H., Kizer, K. O., Shiibata, Y., Hall, M. C., Borchers, C. H., and Strahl, B. D. (2003) Genes Dev. 17, 654–663
21. Li, B., Howe, L., Anderson, S., Yates, J. R., III, and Workman, J. L. (2003) J. Biol. Chem. 278, 8897–8903
22. Kizer, K. O., Phatnani, H. P., Shiibata, Y., Hall, Greenblatt, J. F., and Strahl, B. D. (2005) Mol. Cell Biol. 25, 3005–3016
23. Phatnani, H. P., Jones, J. C., and Greenleaf, A. L. (2004) Biochemistry 43, 15702–15719
24. Delaglio, F., Grzesiek, S., Vuister, G., Zhu, G., Pfeifer, J., and Bax, A. (1995) J. Biomol. NMR 6, 277–293
25. Johnson, B. A., and Blevins, R. A. (1994) J. Biol. Chem. 269, 603–614
26. Sattler, M., Schleucher, J., and Griesinger, C. (1999) Prog. NMR Spectrosc. 34, 93–158
27. Cornilescu, G., Delaglio, F., and Bax, A. (1999) J. Biol. Chem. 274, 289–302
28. Farre, N. A., Mubandatu, R., Singer, A. U., Pascual, S. M., Kay, C. M., Gehr, G., Shoelson, S. E., Panson, T., Forman-Kay, J. D., and Kay, L. E. (1994) Biochemistry 33, 5984–6003
29. Linge, J. P., O’Donoghue, J., and Nilges, M. (2001) Methods Enzymol. 339, 71–90
30. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunzеле, R. W., Jiang, J. S., Kuzewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warran, G. L. (1998) Acta Crystallogr. Sect. D Biol. Crystallogr. 54, 905–921
31. Linge, J. P., Williams, M. A., Sprock, C. A., Bonvin, A. M., and Nilges, M. (2003) Proteins 50, 496–506
32. Laskowski, R. A., Ruhlmann, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) J. Biomol. NMR 8, 477–486
33. Allen, M., Friesel, A., Schon, O., and Bycroft, M. (2002) J. Mol. Biol. 323, 411–416
34. Meinhart, A., and Cramer, P. (2004) Nature 430, 223–226
35. Morris, S. A., Shihata, Y., Norita, K., Takumoto, Y., Warren, E., Tempfe, B., Grewal, I. S., and Strahl, B. D. (2005) Eur. J. Cell Biol. 84, 1446–1454
36. Adhvaryu, K. K., Morris, S. A., Strahl, B. D., and Selker, E. U. (2005) J. Biol. Chem. 280, 7039–7046
37. Fabrega, C., Shen, V., Shuman, S., and Lima, C. D. (2002) Mol. Cell Biol. 11, 1549–1561
38. Verdeca, M. A., Bowman, M. E., Lu, K. P., Hunter, T., and Noel, J. P. (2000) Nat. Struct. Biol. 7, 639–643
39. Holm, L., and Sander, C. (1995) Trends Biochem. Sci. 20, 478–480
40. Murakami, K. S., Massutu, S., and Darst, S. A. (2002) Science 296, 1280–1284
41. Gross, C. A., Chan, C., Dombroski, A., Gruber, T., Sharp, M., Tuppy, J., and Young, B. (1998) Cold Spring Harbor Symp. Quant. Biol. 63, 141–155
42. Forget, D., Langefier, M.-F., Therien, C., Trinh, V., and Coulombe, B. (2004) Mol. Cell Biol. 24, 1122–1131
43. Okhuma, Y., Sumimoto, H., Hoffmann, A., Shimasaki, S., Horiokushi, M., and Roe- der, R. (1991) Nature 354, 398–401
44. Meinhart, A., Blobel, J., and Cramer, P. (2003) J. Biol. Chem. 278, 48267–48274