

Chromodomains read the arginine code of post-translational targeting

Iris Holdermann¹, N Helge Meyer^{2,3}, Adam Round⁴, Klemens Wild¹, Michael Sattler^{2,3} & Irmgard Sinning¹

Chromodomains typically recruit protein complexes to chromatin and read the epigenetic histone code by recognizing lysine methylation in histone tails. We report the crystal structure of the chloroplast signal recognition particle (cpSRP) core from *Arabidopsis thaliana*, with the cpSRP54 tail comprising an arginine-rich motif bound to the second chromodomain of cpSRP43. A twinned aromatic cage reads out two neighboring nonmethylated arginines and adapts chromodomains to a non-nuclear function in post-translational targeting.

Chromodomain (chromatin organization modifier)-containing proteins have pivotal roles in the regulation of gene expression and chromatin remodeling¹. These chromodomains typically act as adaptors for protein-protein interactions in multiprotein complexes and were first identified in two distinct regulators of chromatin structure (heterochromatin protein 1 (HP1) and Polycomb)². They are involved in deciphering the epigenetic histone code³ by recognizing and binding to methylated lysines in histone tails⁴. Although these canonical chromodomains all localize to the nucleus, only a few are found in cytosolic proteins and have been related to different aspects of translation⁵. The signal recognition particle protein cpSRP43 localizes in the stroma of chloroplasts and is the only known example of a protein with chromodomains implicated in protein targeting⁶.

cpSRP43 comprises three chromodomains and four ankyrin repeats⁷ (Fig. 1a) and adapts the universally conserved SRP system in chloroplasts to a post-translational function⁸. Whereas cytosolic SRP is a

ribonucleoprotein complex that targets nascent polypeptide chains during translation to the translocation machinery, cpSRP is a heterodimer consisting of cpSRP43 and cpSRP54 and lacks the SRP RNA⁹. Its main cargo is nuclear-encoded light-harvesting chlorophyll *a*, *b* binding proteins (LHCPs) that are sequestered into a soluble transit complex¹⁰. The ankyrin repeats bind LHCPs with high sequence specificity⁸ and have a chaperone function¹¹. Chromodomain 2 (CD2) stably binds cpSRP54 to form cpSRP¹² and also serves in recruiting cpSRP to the thylakoid membrane by interaction with the Alb3 insertase¹³. Interaction studies between cpSRP54 and cpSRP43 previously revealed a conserved RRKR motif in the C-terminal tail of cpSRP54 as the major binding site for cpSRP43 (ref. 14) (Fig. 1a). Here we describe the structure of cpSRP43 in complex with the C-terminal tail of cpSRP54, demonstrating how chromodomains have been adapted to a function outside the nucleus.

As a first step to understand how cpSRP is formed by the interaction of cpSRP54 with cpSRP43 CD2, we determined the crystal structure of *A. thaliana* cpSRP43ΔCD3 with the RRKR motif from cpSRP54

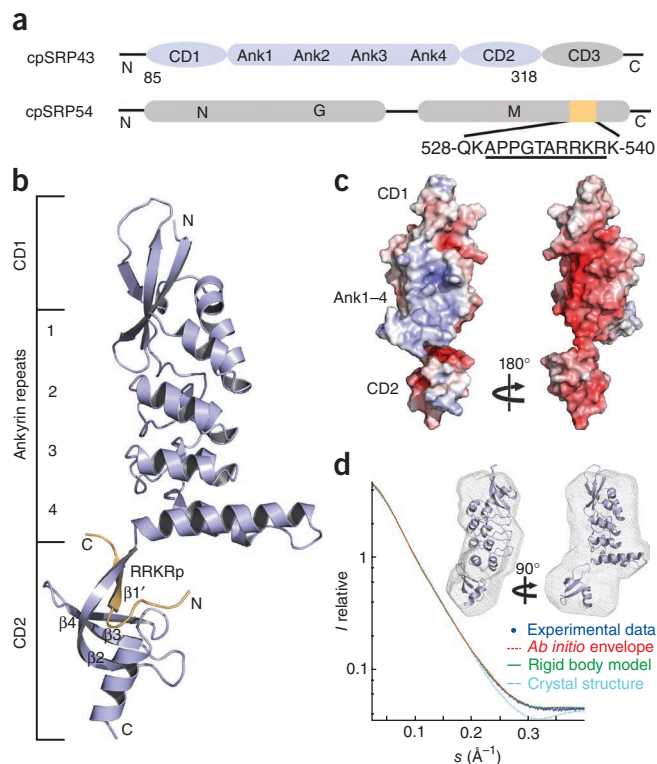


Figure 1 Structure of cpSRP43 in complex with the cpSRP54 tail. (a) Domain architecture of cpSRP43 and cpSRP54. Domains present in the crystal structure are given by residue numbers and are highlighted in blue and orange. (b) Overall structure of cpSRP43ΔCD3 (blue) in complex with RRKRp (orange). RRKRp locates to the interface of Ank4 and CD2 and binds by β-completion. (c) Electrostatic surface potential of cpSRP43ΔCD3. (d) cpSRP43ΔCD3 analyzed by SAXS. Experimental scattering (blue dots) is compared to the calculated scattering of the *ab initio* model (red), the rigid-body model (green) and the crystal structure (cyan). The rigid body model is superimposed with the envelope of the *ab initio* model (inset, shown in mesh).

¹Heidelberg University Biochemistry Center, Heidelberg, Germany. ²Institute of Structural Biology, Helmholtz Zentrum München, Neuherberg, Germany. ³Munich Center for Integrated Protein Science and Chair of Biomolecular NMR, Department Chemie, Technische Universität München, Garching, Germany. ⁴European Molecular Biology Laboratory, Grenoble Outstation, Grenoble, France. Correspondence should be addressed to I.S. (irmi.sinning@bzh.uni-heidelberg.de).

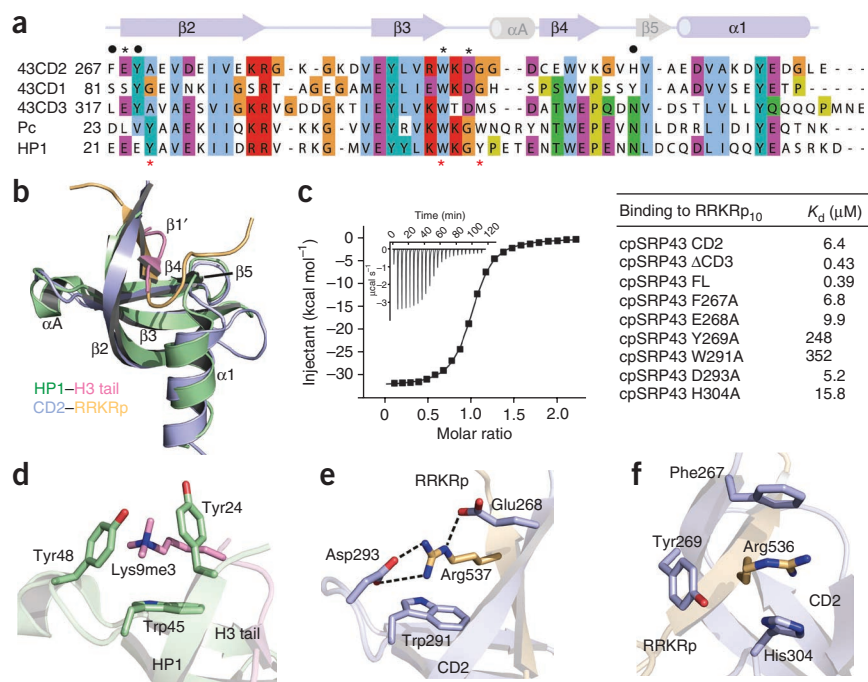
Received 22 August 2011; accepted 7 November 2011; published online 8 January 2012; doi:10.1038/nsmb.2196

Figure 2 Specific readout of the RRKR motif by a twinned cage in cpSRP43.

(a) Structure-based sequence alignment of cpSRP43 chromodomains (43CD1 to 43CD3) and the chromodomains of HP1 and Polycomb (Pc). Secondary structure elements are depicted in blue for CD2, and additional elements of canonical chromodomains in gray. Aromatic cage residues in canonical chromodomains are marked with a red asterisk. Corresponding residues in cpSRP43 chromodomains are marked with a black asterisk (cage 1), residues involved in the noncanonical cage (cage 2) with a black dot.

(b) Superposition of cpSRP43 CD2 (blue) and bound cpSRP54 tail (orange) with HP1 (green) in complex with histone tail H3K9me3 (pink) (PDB 1KNE; ref. 15). Both substrates use the same binding mode and complete the β -barrel of the chromodomain. (c) Characterization of the interaction between cpSRP43 and RRKR_{p10} by ITC. Titration of cpSRP43 (full length, FL) and RRKR_{p10} is shown as an example. Dissociation constants in the table show the contribution of cage-forming residues to binding of RRKR_p.

(d) Binding mode of H3K9me3 in the classical aromatic cage of HP1. (e) Recognition of RRKRp Arg537 by the modified aromatic cage in CD2 (cage 1). (f) Recognition of RRKRp Arg536 by the second cage of CD2 (cage 2).



(RRKR_p, residues 528–540) at 3.2 Å resolution (see **Supplementary Methods**, **Supplementary Table 1** and **Supplementary Figs. 1** and **2**). The structure comprises CD1, the four ankyrin repeats (Ank1–Ank4) and CD2 (**Fig. 1b**). CD2 shows the typical chromodomain architecture with a three-stranded antiparallel β -sheet and a perpendicularly packed C-terminal α -helix. Negatively charged residues cluster on the back side of CD2 and extend the negatively charged surface of cpSRP43, thereby adding to the previously observed RNA mimicry⁸ (**Fig. 1c**). Whereas CD1 is rigidly fused to the ankyrin repeats, CD2 does not form any tertiary interactions with the N-terminal domains of cpSRP43 and is added like a bead on a string. Therefore, the domain arrangement observed in the crystal structure may be influenced by crystal packing. Analysis of cpSRP43 Δ CD3 in solution by small-angle X-ray scattering (SAXS), however, confirms the overall domain arrangement observed in the crystal structure and further indicates flexibility concerning the position of CD2 (**Fig. 1d**). SAXS data with bound RRKR_p show that CD2 is positioned more closely to Ank4 (**Supplementary Fig. 3**). Flexibility of cpSRP43 might be beneficial for accommodating different interaction partners and especially for binding the transmembrane domains of LHCPs in the transit complex.

CD2 shares high sequence identity with canonical, nuclear chromodomains such as HP1 or Polycomb (**Fig. 2a**), and superposition with HP1 shows the high structural similarity with an overall r.m.s. deviation of 1.19 Å (for all C α atoms)¹⁵ (**Fig. 2b**). RRKR_p binds at the interface between CD2 and the fourth ankyrin repeat (1,740 Å² buried surface area) by means of so-called β -completion (**Fig. 1b** and **Supplementary Fig. 4**). Using isothermal titration calorimetry (ITC), we determined that the decamer peptide (RRKR_{p10}) binds to CD2 alone (K_d of 6.4 μ M) much more weakly than to full-length cpSRP43 (K_d of 400 nM), which reflects the contribution of the Ank4-CD2 interface to the interaction (**Fig. 2c** and **Supplementary Table 2**). Canonical chromodomains typically interact with histone tails with

K_d values in the range of 1–10 μ M (ref. 16). The tail contributes a β -strand to the chromodomain and completes the β -barrel in *trans* (strand β 1') by sandwiching itself between β 2 and β 5 (**Fig. 2b**). The cpSRP43-cpSRP54 interaction follows this paradigm and involves highly conserved residues of the cpSRP54 tail, including the RRKR motif (residues 530–539) (**Supplementary Fig. 4**). The pattern of β -completion is basically identical to the one observed in the complex of HP1 with the histone H3 tail containing trimethylated lysine 9 (HP1–H3K9me3 complex)¹⁵ (**Fig. 2b** and **Supplementary Fig. 5**). We used heteronuclear NMR spectroscopy of CD2 with RRKR_{p10} to analyze this interaction in solution (**Supplementary Fig. 6**). Both intramolecular (CD2) and intermolecular (CD2–RRKR_{p10}) NOE correlations observed in solution are fully consistent with our crystal structure. Our NMR data confirm that CD2 also adopts the typical chromodomain fold in solution and that the RRKR_p interaction follows the general principle of β -completion.

In the crystal structure, RRKR_p adopts an extended conformation, except for a type II β -turn (residues 531–534) at its N terminus, which wraps around Tyr269 within the Ank4-CD2 linker region and threads the tail into the hydrophobic binding groove of CD2. Insertion of RRKR_p as strand β 1' involves Arg537 (0 position), which is accommodated in an incomplete aromatic cage (see below), as well as the preceding four residues (Gln533–Thr534–Ala535–Arg536, positions –4 to –1), with the last three being identical to the corresponding residues in the HP1–H3K9me3 complex (**Supplementary Fig. 5**). However, β -completion by RRKR_p extends up to position +2 and therefore involves the entire RRKR motif. Specific readout of the RRKR motif relies on the two neighboring arginines, Arg536 and Arg537 (positions –1 and 0). Replacing either of them abolishes complex formation and drastically reduces LHCP targeting¹⁴. The 0 position in classical chromodomain interactions corresponds to the methylated lysine of the histone tail, which is accommodated within the aromatic cage formed by three aromatic residues that surround the

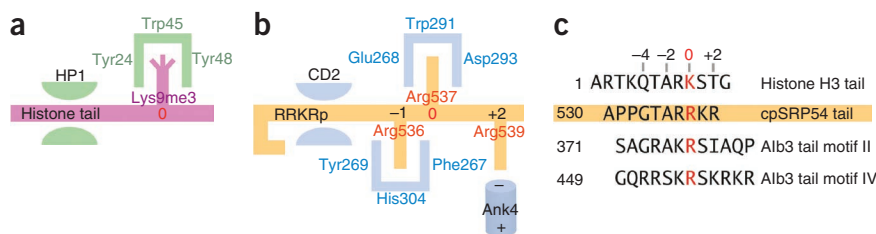


Figure 3 Substrate recognition by chromodomains. (a,b) Scheme of decoding lysine methylation by canonical chromodomains (a) and arginine recognition in the cpSRP43 twinned aromatic cage (b). (c) Sequences recognized by chromodomains with the critical position (0) shown in red.

methyl ammonium group (Fig. 2d). Cation- π stacking interactions between the positively charged lysine and the electron-rich aromatic ring systems decode the level of methylation¹⁷. Lower methylation states of histone tails are read out by domains where acidic residues line a modified aromatic cage³. Notably, the corresponding cage in CD2 (cage 1) is incomplete and is formed by only a single aromatic side chain (Trp291) and two charged residues (Glu268 and Asp293) that replace the other two aromates (Fig. 2a,e). All three residues are highly conserved (Supplementary Fig. 7), and this adaptation of the cage reflects the change in the substrate, with the methylated lysine being replaced by a nonmethylated arginine (Arg537). The positive charge of Arg537 still enables cation- π interactions with Trp291; however, the guanidinium group of Arg537 is involved in salt bridges to the acidic side chains of Glu268 and Asp293. As a unique addition to canonical chromodomain interactions, Arg536 in the -1 position is accommodated in a second aromatic cage formed by three residues: Phe267, His304 and Tyr269 (Fig. 2f and Supplementary Fig. 5). Here the guanidinium group is sandwiched between Phe267 from the Ank4-CD2 linker region and His304 from CD2 strand β 5. Furthermore, Tyr269, which usually lines the classical aromatic cage, is flipped over by 180° and completes the second cage. The change in tyrosine orientation is caused by a single-residue insertion that is conserved in all three chromodomains of cpSRP43.

In order to test the contribution of the residues forming the twinned cage to cpSRP54 tail binding, all six residues were mutated individually and analyzed by ITC (Fig. 2c and Supplementary Table 2). Mutation of Tyr269 to alanine drastically reduces the binding affinity (by a factor of about 600) as measured by ITC using full-length cpSRP43 and RRKRp₁₀. Mutation of Trp291 to alanine results in a similar reduction in affinity, whereas a D293A mutation that disrupts the salt bridge to Arg537 is less critical (reducing affinity by only a factor of 13). The guanidinium group of Arg539 (position +2) is placed at the negative end of the dipole of the C-terminal Ank4 helix, explaining its importance for the interaction¹⁴. Taken together, these results show that the C-terminal tail of cpSRP54 comprising the RRKR peptide participates in a delicate network of interactions with CD2, involving a unique twinned cage that specifically reads out two nonmethylated neighboring arginines (Fig. 3).

The chromodomains of HP1 and Polycomb are highly conserved, but they selectively interact with conserved motifs in histone tails that only vary in a few residues¹⁶. This underlines the high specificity of chromodomain interactions. The three chromodomains of cpSRP43 are also very similar; however, pull-down and ITC measurements showed that the cpSRP54 C-terminal tail only binds to CD2 (refs. 12,18). Binding studies with full-length cpSRP43 and cpSRP43 Δ CD3 resulted in identical values (Supplementary Table 2). Also, NMR titrations using the tandem chromodomains of cpSRP43 (CD2-CD3) do not show any spectral changes for residues in CD3

upon titration with the RRKRp₁₀ peptide (data not shown). These data are consistent with previous observations that CD3 is indeed dispensable for the interaction¹². Both CD1 and CD3 deviate from CD2 in residues involved in the twinned cage and are unable to position the type II turn. The unique Ank4-CD2 interface also contributes to the interaction. Taken together, our data explain the high binding affinity and specificity of the cpSRP54 tail for CD2, which is the basis for recruiting cpSRP54 to a function in post-translational targeting.

We have previously shown that CD2 and CD3 are involved in the interaction with the membrane insertase Alb3 (ref. 13). Two conserved motifs in the Alb3 C-terminal tail are required for tethering cpSRP43 to the thylakoid membrane and are reminiscent of ARK(S/T) motifs in histone tails¹³. Based on the structure of the cpSRP43-RRKRp complex, which shows how CD2 decodes two neighboring arginines (Fig. 3), and consistent with previous mutagenesis data¹³, we expect that cpSRP43 reads out the Alb3 C-terminal tail in a similar manner. This enables specific selection of the correct membrane insertase, as the Alb3 homolog Alb4 lacks both motifs and therefore does not bind to cpSRP43. We envision that the modular architecture of cpSRP43, with the two C-terminal chromodomains, allows for specific recognition of cpSRP54 and Alb3 and therefore defines the molecular framework of post-translational targeting by cpSRP in chloroplasts. What is the evolutionary origin of these non-nuclear chromodomains? The earliest occurrences of cpSRP43 in chlorophytes coincide with the appearance of chromodomain-containing proteins and chromovirus retrotransposons in plants¹⁹. Strong sequence dissimilarities already at this stage do not allow for precise conclusions about the evolutionary origin of cpSRP43 chromodomains (Supplementary Fig. 7). However, cpSRP43 is only found in higher plants, streptophytes and some chlorophytes that also encode cpSRP54 with the C-terminal RRKR motif. This indicates that cpSRP43 has coevolved with cpSRP54 and the appearance of cpSRP43 coincides with the loss of the SRP RNA as a prime regulatory element²⁰ (Supplementary Fig. 7). The transition of chromodomains from reading the epigenetic histone code to reading the arginine code of protein targeting extends the spectrum of chromodomain interactions and allows the use of both chromodomains and the SRP system in a new functional context.

Accession codes. Protein Data Bank: coordinates and structure factors have been deposited under accession code 3UI2.

Note: Supplementary information is available on the Nature Structural & Molecular Biology website.

ACKNOWLEDGMENTS

This work was supported by grants from the Deutsche Forschungsgemeinschaft (DFG) SFB638 and the Graduiertenkolleg GRK1188 (to I.S.). We thank J. Kopp and C. Siegmann from the crystallization platform of the Biochemiezentrum der Universität Heidelberg (BZH)/Cluster of Excellence:CellNetworks for support in protein crystallization, R. Lindner for support in evolutionary analysis, and S. Falk, G. Bange and D. Schünemann for stimulating discussions. We acknowledge the European Synchrotron Radiation Facility for providing synchrotron radiation and for assistance in using beamlines ID14eh2 and ID14eh3. M.S. and H.M. acknowledge support from the Cluster of Excellence, Center for Integrated Protein Science, Munich (CiPS^M). I.S. is an investigator of the Cluster of Excellence:CellNetworks.

AUTHOR CONTRIBUTIONS

I.H., K.W. and I.S. designed experiments and analyzed the data. I.H. conducted experiments. N.H.M. and M.S. provided NMR spectroscopy data and analysis.

A.R. processed and analyzed SAXS data. I.H., K.W. and I.S. wrote the manuscript. All authors commented on the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at <http://www.nature.com/nsmb/>.

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>.

1. Brehm, A., Tufteland, K.R., Aasland, R. & Becker, P.B. *Bioessays* **26**, 133–140 (2004).
2. Paro, R. & Hogness, D.S. *Proc. Natl. Acad. Sci. USA* **88**, 263–267 (1991).
3. Taverna, S.D., Li, H., Ruthenburg, A.J., Allis, C.D. & Patel, D.J. *Nat. Struct. Mol. Biol.* **14**, 1025–1040 (2007).
4. Bannister, A.J. *et al. Nature* **410**, 120–124 (2001).
5. Andersen, C.B. *et al. Nature* **443**, 663–668 (2006).
6. Richter, C.V., Bals, T. & Schunemann, D. *Eur. J. Cell Biol.* **89**, 965–973 (2010).
7. Klimyuk, V.I. *et al. Plant Cell* **11**, 87–99 (1999).
8. Stengel, K.F. *et al. Science* **321**, 253–256 (2008).
9. Grudnik, P., Bange, G. & Sinning, I. *Biol. Chem.* **390**, 775–782 (2009).
10. Schuenemann, D. *et al. Proc. Natl. Acad. Sci. USA* **95**, 10312–10316 (1998).
11. Falk, S. & Sinning, I. *J. Biol. Chem.* **285**, 21655–21661 (2010).
12. Goforth, R.L. *et al. J. Biol. Chem.* **279**, 43077–43084 (2004).
13. Falk, S., Ravaut, S., Koch, J. & Sinning, I. *J. Biol. Chem.* **285**, 5954–5962 (2010).
14. Funke, S., Knechten, T., Ollesch, J. & Schunemann, D. *J. Biol. Chem.* **280**, 8912–8917 (2005).
15. Jacobs, S.A. & Khorasanizadeh, S. *Science* **295**, 2080–2083 (2002).
16. Fischle, W. *et al. Genes Dev.* **17**, 1870–1881 (2003).
17. Hughes, R.M., Wiggins, K.R., Khorasanizadeh, S. & Waters, M.L. *Proc. Natl. Acad. Sci. USA* **104**, 11184–11188 (2007).
18. Sivaraja, V. *et al. J. Biol. Chem.* **280**, 41465–41471 (2005).
19. Gorinsek, B., Gubensek, F. & Kordis, D. *Cytogenet. Genome Res.* **110**, 543–552 (2005).
20. Rosenblad, M.A., Larsen, N., Samuelsson, T. & Zwieb, C. *RNA Biol.* **6**, 508–516 (2009).