Crystal Structure of Human Symplekin-Ssu72-CTD Phosphopeptide Complex

In eukaryotes, 3'-end processing, which includes cleavage and polyadenylation, is crucial for the stability, nuclear export and translation of mRNAs (1). In the large 3'-end processing machinery, symplekin is believed to be a scaffold that mediates interactions among several protein factors (2). For example, the yeast homolog of symplekin (Pta1) is known to interact with Ssu72 (3), a Ser-5 phosphatase of the carboxy-terminal domain (CTD) of the largest subunit of RNA polymerase II (Pol II). However, the molecular basis for how symplekin mediates these interactions was not known.

The crystal structure of the N-terminal domain of human symplekin was solved at 1.9 Å resolution with diffraction data collected at SSRL. It consists of seven pairs of anti-parallel α helices, similar to structures of Arm/HEAT repeats (Figure 1A, cyan). Proteins having such structural features usually serve as scaffolds and participate in protein-protein interactions, consistent with the function of symplekin in the 3'-end processing machinery.

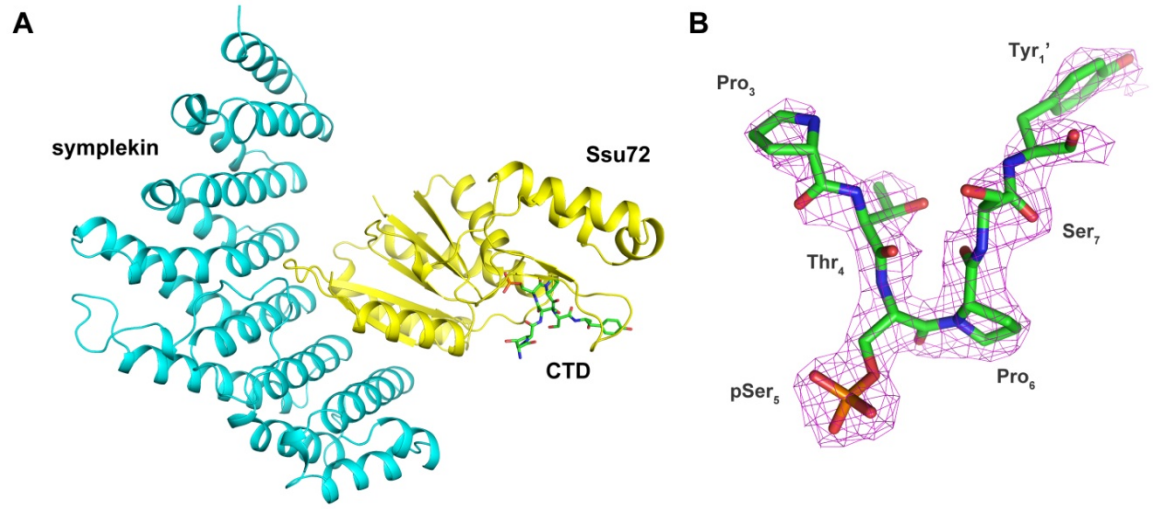

Figure 1. **A**. Schematic representation of the structure of the human symplekin N-terminal domain (cyan)-Ssu72 (yellow)-CTD phosphopeptide (green) ternary complex. **B**. Electron density for the CTD phosphopeptide at 2.4 Å resolution. The *cis* peptide bond between pSer5 and Pro6 is clearly visible.

We then established the interaction between symplekin and human Ssu72 and determined the crystal structures of their complex alone as well as with a substrate Pol II CTD phosphopeptide of Ssu72 (Figure 1A). Ssu72 makes contact with the concave side of symplekin. Interestingly, despite the long distance (25 Å) between the symplekin-Ssu72 interface and the active site of Ssu72, symplekin is able to stimulate Ssu72's phosphatase activity in our *in vitro* assays, suggesting for the first time that symplekin may also have a regulatory function rather than being simply a passive scaffold.

The most intriguing feature lies in the conformation of the substrate peptide bound to Ssu72. Ssu72 belongs to a group of proline-directed serine phosphatases, which targets the pSer-Pro structure. The peptide bond between the serine and the proline can adopt either *trans* or *cis* configuration, and all phosphatases in this category are believed to recognize only the *trans* configuration based on previous studies. However, we found that the pSer-Pro peptide bond of the substrate is in the *cis* configuration and only this configuration can be accommodated in the active site of Ssu72 (Figure 1B). This is the first example of a phosphatase that is specific for the *cis* configuration of the pSer-Pro bond in the substrate. This unique specificity implies an important role for Ssu72 in transcription regulation by selecting only the *cis* CTD configuration. This effect may also be essential for co-transcriptional 3'-end processing as CTD is pivotal in coordinating these two events.

Supported by NIH R01 GM077175.

Primary Citation

Xiang, K., Nagaike, T., Xiang, S., Kilic, T., Beh, M. M., Manley, J. L., and Tong, L. (2010) Crystal structure of the human symplekin–Ssu72–CTD phosphopeptide complex Nature, 467, 729-733.

References

1. Mandel, C. R., Bai, Y., and Tong, L. (2008) Protein factors in pre-mRNA 3'-end processing. *Cell. Mol. Life Sci.* **65**, 1099-1122.
2. Takagaki, Y., and Manley, J. L. (2000) Complex protein interactions within the human polyadenylation machinery identify a novel component. *Mol. Cell. Biol.* **20**, 1515-1525.
3. Ghazy, M. A., He, X., Singh, B. N., Hampsey, M., and Moore, C. (2009) The essential N terminus of the Pta1 scaffold protein is required for snoRNA transcription termination and Ssu72 function but is dispensible for pre-mRNA 3'-end processing. *Mol. Cell. Biol.* **29**, 2296-2307.

SSRL is primarily supported by the DOE Offices of Basic Energy Sciences and Biological and Environmental Research, with additional support from the National Institutes of Health, National Center for Research Resources, Biomedical Technology Program, and the National Institute of General Medical Sciences.