

## Crystal structure of self-association sequence and nuclear localization signal of HCF-1

Jihye Park<sup>1,\*</sup>, Fabienne Lammers<sup>2</sup>, Winship Herr<sup>2</sup> and Ji-Joon Song<sup>1</sup><sup>1</sup>Department of Biological Sciences and Graduate School of Nanoscience and Technology (World Class University), KAIST (Korea Advanced Institute of Science and Technology) Institute for the BioCentury, KAIST, Daejeon 305-701, Korea<sup>2</sup>Center for Integrative Genomics, University of Lausanne, Génopode, 1015 Lausanne, Switzerland

### 1 Introduction

HCF-1 is synthesized as a large precursor and cleaved into two subunits, HCF-1N and HCF-1C. The two subunits remain noncovalently associated via two self-association sequences (SAS) after cleavage. SAS1 is composed of N-terminal self-association sequence called SAS1N and C-terminal self-association sequence called SAS1C. SAS1C was predicted to form fibronectin type 3 (fn3) repeat structure [1]. The association of HCF-1N and HCF-1C through SAS is important for the function of HCF-1 in cell. However, the molecular mechanism of self-association was unknown. Here, we report the crystal structure of SAS1 with C-terminal nuclear localization signal (NLS).

### 2 Experiment

SAS1N and SAS1C–NLS were coexpressed in *E. coli*. The SAS1–NLS complex was purified by affinity chromatography, ion exchange chromatography, and size exclusion chromatography. SAS1–NLS crystals were grown by hanging-drop vapor diffusion method in crystallization solution containing 2.23 M ammonium sulfate and 5% (vol/vol) isopropanol. The crystal was soaked in a solution containing 3.6 M ammonium sulfate and 10% (vol/vol) glycerol for cryoprotection. X-ray diffraction data was collected at NW12A beamline in Photon Factory. Diffraction data was processed with HKL2000 and phases were determined using SOLVE/RESOLVE. Initial model was built using the program O and the model was refined with CNS and PHENIX.

### 3 Results and Discussion

The overall structure of SAS1–NLS is shown in Figure 1A. The crystal structure shows that SAS1N and SAS1C are tightly associated by forming interdigitated fn3 repeat structure. SAS1N and N-terminal part of SAS1C are folded together to form one fn3 domain (FN3-1). The c-terminal part of SAS1C forms another fn3 domain (FN3-2) (Figure 1A). This interdigitated structure of SAS1 is unexpected because SAS1N and SAS1C are apart more than 1400 amino acids. Thus, the crystal structure of SAS1–NLS reveals that HCF-1N and HCF-1C subunits are associated by forming interdigitated fn3 repeat structure via SAS1.

In the SAS1–NLS crystal structure, two SAS1–NLS molecules are in an asymmetric unit. In one molecule, C-terminal NLS and linker between SAS1C and NLS is disordered (Figure 1B, green). In the other molecule, however, part of NLS is structure and located in between two fn3 domains (Figure 1A and figure 1B, gray). Interestingly, the two fn3 domains adopt more closed conformation in a molecule having structured NLS. As a result, it seems that the two fn3 domains pinch NLS like sandwich. This sandwiched conformation suggests a regulation mechanism of nuclear localization of HCF-1 by pinching the NLS with the two fn3 domains and by releasing the NLS from the two fn3 domains.

In summary, this result provide structural insight into molecular mechanism of HCF-1 self-association and the conformation of NLS.

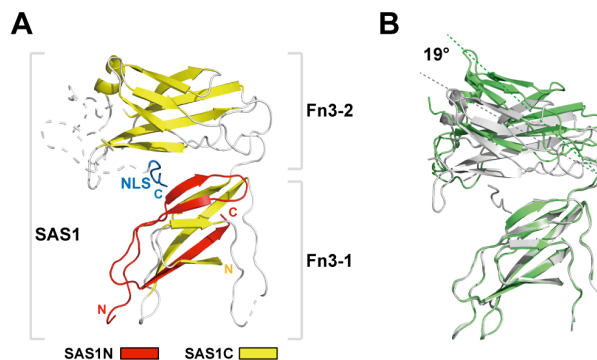


Fig. 1: Structure of HCF-1 self-association sequence and C-terminal nuclear localization signal. (A) Overall structure of SAS1–NLS. (B) Superimposition of two SAS1–NLS conformations.

### References

- [1] Wilson AC *et al.*, Mol Cell Biol. 20(18):6721-30 (2000)

\* sandeulbi@kaist.ac.kr