International Journal of Mass Spectrometry 420 (2017) 57–66

Contents lists available at ScienceDirect

International Journal of Mass Spectrometry

journal homepage: www.elsevier.com/locate/ijms

Structural characterization of the yeast CF IA complex through a combination of mass spectrometry approaches

Johann Stojko[a,b,1], Adrien Dupin[c,d,1], Stéphane Chaignepain[e,1], Lionel Beaurepaire[c,d], Amélie Vallet-Courbin[d], Alain Van Dorsselaer[a,b], Jean-Marie Schmitter[e], Lionel Minvielle-Sébastia[d], Sébastien Fribourg[c,d,**], Sarah Cianférani[a,b,*]

[a] BioOrganic Mass Spectrometry Laboratory (LSMBO), IPHC, Université de Strasbourg, 25 rue Becquerel, 67087 Strasbourg, France
[b] IPHC, CNRS, UMR7178, 67087 Strasbourg, France
[c] IECB, Université de Bordeaux, IECB, 33607 Pessac, France
[d] INSERM U1212, UMR CNRS 5320, 33076 Bordeaux, France
[e] CBMN, UMR 5248, CNRS, Université de Bordeaux, INP Bordeaux, 33607 Pessac, France

ARTICLE INFO

Article history:
Received 14 May 2016
Received in revised form 16 August 2016
Accepted 16 August 2016
Available online 17 August 2016

Keywords:
Native mass spectrometry
Crosslinking mass spectrometry
Ion mobility-mass spectrometry
Polyadenylation
CF IA

ABSTRACT

The cleavage/polyadenylation factor IA (CF IA) is a yeast multiprotein complex that consists of Rna14, Rna15, Pcf11 and Clp1 proteins, and is involved in the 3′-end maturation of mRNAs. Structural data have been reported for the individual protein partners and binary complexes; however, little is known about the molecular architecture of the entire CF IA assembly. Here, we report a thorough characterization of complete recombinant CF IA assembly and its subcomplexes using a combination of mass spectrometry (MS) approaches. We first focused on the Rna14p:Rna15p and Pcf11p:Clp1p subcomplexes in order to obtain a detailed picture of their interactions. Native MS and crosslinking MS showed that the intact CF IA assembly exists in solution as pentameric and hexameric species, composed of two copies of Rna14p, one each of Pcf11p and Clp1p, and one or two of Rna15p, respectively. Partial denaturation experiments followed by native MS along with crosslinking analysis revealed two building blocks: Rna14p:Rna15p multimer subcomplexes assemble with Pcf11p:Clp1p heterodimers to form the CF IA complex. We then used ion mobility-MS (IM-MS) to investigate the conformational changes induced upon CF IA assembly. The new information on the CF IA assembly process provided by this combination of MS approaches (native MS, crosslinking MS and IM-MS) allowed us to discuss a topological model of the CF IA assembly.

1. Introduction

Eukaryotic RNA polymerase II synthesizes mRNA precursors (pre-mRNAs) that have to be processed before being exported to the cytoplasm and translated into proteins. An essential modification occurs on every pre-mRNA at their 3′-ends (with the exception of replication-dependent histone mRNAs in metazoans). This maturation involves the addition of poly(A) tails that are important for mRNA export, translation and stability. The pre-mRNA is initially cleaved in the 3′-untranslated region (3′-UTR). A poly(A) tail is then added at the 3′-hydroxyl end of the upstream cleavage product, the 3′ fragment being degraded by Rat1p (Xrn2 in humans). Pre-mRNA processing and transcription termination rely on the recognition of *cis*-acting elements by the polyadenylation machinery [1,2].

In yeast, most of the proteins required to accomplish polyadenylation are found in two large multiprotein complexes: the cleavage/polyadenylation factor IA (CF IA) [3,4] and the cleavage and polyadenylation factor (CPF). The additional proteins required to complete the process are Hrp1p/Nab4p, involved in the poly(A) site selection [3,5], and the poly(A)-binding proteins Pab1p and Nab2p, which control the length of the poly(A) tail and the subsequent export of mature mRNA [4,6–8]. CF IA consists of Rna14p, Rna15p, Pcf11p and Clp1p subunits. Initially thought to be a heterotetramer [9], CF IA seems to be more complex than believed and there is no high resolution 3D structure available for entire CF IA. Crystal structures of *E. cuniculi* and *M. musculus* CstF-77 (orthologue of yeast Rna14p), and recently *K. lactis* Rna14p:Rna15p, have shown that CstF-77/Rna14p associates in a tight homodimer

* Corresponding author at: BioOrganic Mass Spectrometry Laboratory (LSMBO), IPHC, Université de Strasbourg, 25 rue Becquerel, 67087 Strasbourg, France.
** Corresponding author at: INSERM U1212, UMR CNRS 5320, 33076 Bordeaux, France.
E-mail addresses: sebastien.fribourg@inserm.fr (S. Fribourg), sarah.cianferani@unistra.fr (S. Cianférani).
[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.ijms.2016.08.005
1387-3806/

[10–14]. The solution structure of the Rna14p:Rna15p minimal heterodimer [14], similar to that of *K. lactis* Rna14p:Rna15p [13], has completed the overall description of the Rna14p:Rna15p complex. However, due to the presence of long unfolded linkers in the Rna14 protein, the relative position of the Rna14p and Rna15p subunits could not be determined unambiguously [13]. Structural information are also available concerning Clp1p in interaction with a small region of Pcf11p [37,38]. Conversely to Rna14p:Rna15p subcomplex, Pcf11p:Clp1p does not seem to dimerize. Additional information mainly obtained from direct protein–protein interaction assays performed in pull-down experiments with full-length recombinant proteins and/or isolated domains allowed providing a model for CF IA topology. These experiments showed that Pcf11p directly interacts with Rna14p and Rna15p, the latter through its C-terminal domain [9,49] and that Clp1p makes weak but significant contacts with Rna14p [42]. Nevertheless, these high-resolution subcomplexes structures and direct protein–protein interactions are not sufficient to draw a global topology of CF IA. Recently, based on size exclusion chromatography (SEC), analytical ultra centrifugation (AUC) and small-angle X-ray scattering analysis (SAXS), the stoichiometry of this factor was determined to be 2:2:1:1 for Rna14p:Rna15p:Pcf111p:Clp1p, respectively, and a model for its architecture was proposed [15].

As crystallographic approaches failed, we used a combination of structural mass spectrometry (MS) approaches to gain a deeper understanding of CF IA assembly and subunit stoichiometry, in order to scrutinize the proposed model from a different angle. Structural MS has emerged as a valuable tool to study large protein assemblies when classical biophysical techniques fail [16,17]. A complete MS-based toolbox is available to gain structural information on protein complexes. Native MS has proved an efficient probe of binding stoichiometry, affinity, specificity and cooperativity in solution for non-covalently linked systems [18–24]. Another key benefit of this technique is its ability to reveal the dynamics of assembly/disassembly in real-time [25,26]. Recently also, an additional separation based on ion charge and shape has been exploited in ion mobility (IM-MS) studies of protein complexes [27–29]. Finally, crosslinking MS or H/D exchange experiments can complement native MS and IM-MS, providing direct information on interacting regions [30–34].

In this study, we combined native, crosslinking and IM-MS to characterize the structure and dynamics of yeast CF IA. The topology of the CF IA complex was investigated comprehensively by native and crosslinking MS, including the CF IA subcomplexes and the entire CF IA stoichiometry. We demonstrate the coexistence of pentameric and hexameric CF IA assemblies, consisting of one Pcf11p, one Clp1p and two Rna14p subunits, with either one or two Rna15p subunits. We next used native IM-MS to characterize CF IA assembly and the associated conformational changes. Altogether, our results support the previously proposed topological model of the CF IA complex [15].

2. Materials and methods

2.1. Cloning and expression of the complex

Each full-length protein-coding sequence was obtained by polymerase chain reaction performed on yeast genomic DNA and cloned into modified versions of pET-15b, pET28b, pCDF and pLysS bacterial expression plasmids [35,36]. Due to a gain in stability in comparison to the wild-type protein, Pcf11p was deleted from the 20 glutamines stretch located between residues 234–253 (Pcf11pΔQ_{20}). The integrity of each clone was verified by nucleotide sequence analysis, with only one point mutation detected (K90R for Rna15p) with respect to database sequences. In order to explore the Clp1p-Pcf11p interaction, we used the Pcf11p construct described earlier [37,38].

For co-expression, BL21(DE3) cells (Novagen) were transformed with the relevant combination of plasmids and plated on Petri dishes with half the antibiotic concentration. Mini-cultures were grown in 10 ml of LB medium at 37 °C with antibiotics. Protein expression was induced by adding 1 mM IPTG and the cultures were kept overnight at 15 °C. The cells were harvested by centrifugation and lysed by sonication with a buffer containing 1.5× PBS, 1 mM $MgAc_2$, 0.1% NP-40, 20 mM imidazole and 10% (v/v) glycerol. After centrifugation, the supernatant was incubated with His-affinity resin (Sigma) for 30 min at 4 °C. The resin was extensively washed and the bound proteins were analysed by SDS-PAGE following the addition of loading buffer.

2.2. Sample purification

For large-scale protein purification, Terrific Broth (2 l, Conda) was inoculated with the relevant pre-culture and incubated as previously described. After harvesting the cells, the cell pellet was lysed via three passes through an EmulsiFlex C3 homogenizer (Avestin). The crude extract was centrifuged at 50,000*g* for 1 h at 4 °C. The supernatant was incubated with His-affinity resin (Sigma) and loaded onto an AK 16 column (GE Healthcare). Contaminants were removed by extensive washing with a buffer containing 25 mM Tris-HCl at pH 7.5 and 150 mM NaCl. Proteins were eluted from the column with an imidazole gradient up to 250 mM. The eluted proteins (either CF IA ΔQ_{20}, His-Rna15p or His-Rna15p:Rna14p) were then loaded onto a 5 ml HP-Heparin column (GE Healthcare) and separated using a NaCl gradient up to 1 M in 25 mM Tris-HCl at pH 7.5. The His-Pcf11pΔQ_{20}:Clp1p, His-Pcf11p[454-563]:Clp1p and Rna14p complexes were purified on a Mono Q column (GE Healthcare) and eluted with a NaCl gradient up to 500 mM in 25 mM Tris-HCl at pH 7.5. The pool of fractions was concentrated and loaded onto a HR 10/30 200 pg column (GE Healthcare) and equilibrated in 25 mM Tris-HCl at pH 7.5 and 250 mM NaCl.

2.3. Chemical crosslinking followed by high-mass MALDI-MS analysis

The (amine) crosslinkers used were DSS (disuccinimidyl suberate), BS3 (bis(sulfosuccinimidyl)suberate) and DSG (disuccinimidyl glutarate) obtained from ThermoFisher Scientific; CBDPS H8/D8 (CyanurBiotinDimercaptoPropionylSuccinimide) from Creative molecules Inc.; and glutaraldehyde from Sigma Aldrich. Apart from BS3 and glutaraldehyde, which were solubilized in water, the crosslinkers were solubilized in pure dimethylformamide. The entire CF IA ΔQ_{20} complex was crosslinked in the 1.25–10 μM range in a volume of 10 μl using crosslinkers at a final concentration of 0.2 mg/ml (except for glutaraldehyde, which was used at 0.1% (w/v). After adding the crosslinker, the samples were incubated for 1, 3, 6 or 15 h at 25 °C. The shorter incubation times and lower crosslinker/complex ratios were used to favor the stabilization of subcomplexes; longer incubations and higher crosslinker concentrations were used to achieve a complete reaction. The samples were mixed with a matrix solution (1:1 v/v) of sinapic acid (10 mg/ml) containing 50% acetonitrile (v/v) diluted in deionized water with 0.1% (v/v) TFA. Each mixture (1 μl thereof) was then deposited on the MALDI target using the dried-droplet method. As a control, 1 μl of each sample was analysed before adding crosslinker using the same deposition method. High-mass MS analyses were carried out on a MALDI TOF/TOF mass spectrometer (Ultraflex III, Bruker) used in linear mode and equipped with a HM2 high-mass detector (CovalX AG, Zürich, Switzerland), which allows the sensitive (sub-μM) detection of macromolecules up to 1500 kDa