Review Article

Purification of replication factors using insect and mammalian cell expression systems

Shuji Uno, Zhiying You, Hisao Masai *

Department of Genome Medicine, Tokyo Metropolitan Institute of Medical Science, 2-1-6 Kamikitazawa, Setagaya-ku, Tokyo 156-8506, Japan

ARTICLE INFO

ABSTRACT

Purification of factors for DNA replication in an amount sufficient for detailed biochemical characterization is essential to elucidating its mechanisms. Insect cell expression systems are commonly used for purification of the factors proven to be difficult to deal with in bacteria. We describe first the detailed protocols for purification of mammalian Mcm complexes including the Mcm2/3/4/5/6/7 heterohexamer expressed in insect cells. We then describe a convenient and economical system in which large-sized proteins and multi-factor complexes can be transiently overexpressed in human 293T cells and be rapidly purified in a large quantity. We describe various expression vectors and detailed methods for transfection and purification of various replication factors which have been difficult to obtain in a sufficient amount in other systems. Availability of efficient methods to overproduce and purify the proteins that have been challenging would facilitate the enzymatic analyses of the processes of DNA replication.

© 2012 Elsevier Inc. All rights reserved.

1. Introduction

Studies of DNA replication in various biological systems lead to identification of conserved factors required for initiation and elongation of DNA replication as well as for cellular responses to replication stress [1–3]. However, their precise functions are still largely elusive. For characterization of their biochemical properties and ultimate reconstitution of the entire processes with purified factors, it would be crucial to purify each of these factors involved in a quantity sufficient for detailed enzymatic characterization [4–6]. Unfortunately, most of these factors are generally large in sizes and are sometimes present as a component of larger multi-factor complexes. Thus, isolation of these factors in a soluble, functional form is sometimes challenging.

Escherichia coli expression systems would be normally the first choice to generate recombinant proteins, but the large size and different code preference often cause problems in the level of expression or solubility of the expressed protein [7]. Insect cell expression systems circumvent some of these problems and facilitate the purification of large sized proteins or multi-subunit complexes. A novel baculovirus expression termed MultiBac has recently developed, and has been proven to be useful for the production and purification of eukaryotic multiprotein complexes [8]. An insect cell

expression system have been successfully used to express and purify various forms of Mcm assemblies [9–13]. We have also established a transient mammalian expression system using 293T cells for overexpression and purification of replication proteins [14]. In this system, a cDNA of interest is cloned under the strong EF1α promoter with an efficient Kozak sequence positioned prior to ATG. The vector containing the SV40 origin would be transiently replicated in 293T cells expressing the viral-derived T-antigen, increasing its copy number and further contributing to the increased expression. The expression vector DNA would be transfected using a high-molecular-mass polyethylenimine (PEI) reagent which can be prepared in house for easy and economical transfection. Using this system, we were able to express and purify a varieties of replication factors which have been difficult to purify in a functional, full-length form.

In this article, we will first describe the use of the insect cell expression system for purification of the Mcm complexes. Next, a rapid and economical method for purification of large-sized proteins and complexes using a mammalian expression system will be described.

2. Description of methods

2.1. Purification of Mcm assemblies using an insect cell expression system

Mcm is composed of six subunits conserved from yeasts to human [15]. All the six subunits are required for the process of DNA replication, and it is assumed that they constitute a heterohexameric

---

complex at the replication fork. *In vitro*, different combinations of the six subunits generate various stable assemblies [16–18]. Among them, the Mcm4/6/7 complex was shown to possesss DNA helicase activity and its helicase activity was inhibited by the Mcm2 subunit and the Mcm3/5 complex *in vitro* [9,19,20]. A larger complex containing Mcm2∼7, Cdc45 and GINS (CMG helicase) has been identified and was shown to possess processive DNA helicase activity [21,22]. Isolation of intact complexes composed of multiple subunits of Mcm would be essential to unravel the precise biochemical mechanisms of operation of replication fork machinery.

2.1.1. Construction of insect cell expression vectors for Mcm subunits

Mcm subunits are relatively large, and their expression in bacteria tends to result in insoluble proteins. Therefore, we turned to the insect cell expression system. In the followings, Mcm refers to mouse Mcm. Expression of individual subunits in Sf9 cells resulted in recovery of Mcm2 and Mcm6 as soluble forms, whereas Mcm4 was insoluble. To facilitate the isolation and purification of soluble Mcm protein complexes, co-infection of individual Mcm viruses was attempted. Although co-infection of Mcm4, 6, and 7 viruses somewhat improved the solubility of the Mcm4 potentially due to complex formation with Mcm6 and 7, it was not satisfactory. We then devised a co-expression of two subunits on a dual expression vector. 6His-tagged Mcm4 and non-tagged Mcm6 were simultaneously cloned at the *Eco*RI and *Bam*HI sites, respectively, of the pAcUW31 vector (Pharmingen) carrying the baculovirus p10 and polyhedrin promoters. Similarly, 6His-tagged Mcm7 and non-tagged Mcm2 were simultaneously cloned at the *Eco*RI and *Bam*HI sites, respectively, of the pAcUW31. Mcm7-Flag was also cloned at the *Bam*HI site of pAcUW31 vector, permitting the expression of a single subunit of Mcm7. Mcm5-6His-Flag and non-tagged Mcm3 were cloned at the *Xho*I and *Eco*RI sites, respectively, of the pFast-Bac dual vector (Invitrogen) containing both the baculovirus p10 and polyhedrin promoters. These vectors were used to generate recombinant baculoviruses. We then attempted to express various subassemblies of Mcm complexes in insect cells by infection of different combinations of these recombinant baculoviruses.

2.1.2. Generation of recombinant baculoviruses

To generate recombinant baculoviruses expressing the combinations of Mcm2 and 7 or Mcm4 and 6 proteins, pAcUW31 DNA carrying Mcm2 + 7 or Mcm4 + 6 cDNA and BaculoGold Autographa californica nuclear polyhedrosis virus DNA (BaculoGold AcNPV; Pharmingen) were cotransfected into Sf9 cells and recombinant baculoviruses were isolated by plaque purification as recommended by the manufacturer. The recombinant viruses expressing the Mcm proteins were identified by infecting 2×10^6 Sf9 cells with the plaque supernatants and immunoblot analyses of the cell lysate. To generate the recombinant baculoviruses expressing Mcm3 and 5, the FastBac Dual vector DNA carrying MCM3 + 5 was transformed into an *E. coli* DH10Bac cells to obtain the recombinant Bacmid DNA, which was used for transfection into Sf9 insect cells in accordance with the manufacturers' directions. After two or three cycles of amplification of baculoviruses, the high-titer viral solutions were generated, which were used for protein production in High Five cells. Infected cells typically display the cessation of cell growth and granular appearance under microscope inspection. Sf9 and High Five insect cells were cultured at 27 °C in Sf-900 II SFM (Invitrogen) and EX-CELL 405 medium (JRH Biosciences), respectively.

2.1.3. Purification of the Mcm4/6/7 complex

Two methods have been developed for purification of Mcm4/6/7. One is to express the Mcm2, 4, 6 and 7 subunits by using the recombinant baculoviruses *Mcm2 + 7 and Mcm4 + 6* in insect cells [9] and separate Mcm4/6/7 from the Mcm2/4/6/7 complex. The other

is to express the three subunits by using the recombinant baculoviruses carrying the *Mcm4 + 6* and *Mcm7* [23–25].

2.1.3.1. Purification from coexpression of Mcm2, 4, 6 and 7 (Fig. 1A).

To purify the Mcm4/6/7 protein complexes, the four Mcm proteins (Mcm2, Mcm4, Mcm6, and Mcm7) were expressed simultaneously by infecting eight 150 mm plates each containing1.5 x 10^7 High Five insect cells (Invitrogen) with *6HisMcm4 + 6* and *Mcm2 + 6His-MCM7* recombinant baculoviruses at a multiplicity of infection of approximately 10 (1.0 ml of each virus stock per dish). Scale-up production of the recombinant Mcm proteins is possible by using spinner bottles which gave similar yield (data not shown). Cells were collected at 2 days postinfection and the recombinant proteins containing Mcm2, 4, 6 and 7 were recovered from the lysed cells by Ni–NTA (Quiagen) affinity column chromatography, as described below. The infected cells were washed once in ice-cold phosphate-buffered saline and then suspended in 8 ml of lysis buffer consisting of 10 mM Tris–HCl (pH 7.5), 130 mM NaCl, 1% Triton X-100, 10 mM NaF, 10 mM Na phosphate buffer, 10 mM $Na_4P_2O_7$, and protease inhibitor cocktail (Pharmingen) or the complete protease inhibitors cocktail (Roche). After incubation for 40 min on ice, insoluble materials were removed by centrifugation at 40,000 rpm (Beckman, 50.2 Ti rotor) for 40 min at 4 °C. To one volume of the clarified lysate, 1/10 volume of Ni–NTA-agarose that had been washed twice with 10 bed volumes of buffer A (50 mM Na-phosphate buffer [pH 6.0], 300 mM NaCl, and 10% glycerol) was added, and the mixture was incubated for 1 h at 4 °C on a rocking platform. The beads were then collected by centrifugation and washed with buffer A containing 30 mM imidazole until the absorbance at 280 nm (A_{280}) of the supernatant became less than 0.01. Next, the beads were washed once with buffer B (50 mM Na phosphate buffer [pH 8.0], 300 mM NaCl, and 10% glycerol) containing 30 mM imidazole, and the proteins bound to the beads were eluted by one bed volume of buffer B containing 200 mM imidazole (twice) and that containing 400 mM imidazole (once).

The partially purified materials were further purified by a histone H3/H4-Sepharose column chromatography [19,26]. The Mcm-containing fractions eluted from Ni–NTA–agarose were combined and then loaded onto a histone-Sepharose column equilibrated with buffer C (20 mM Tris–HCl [pH 7.5], 0.5 mM EDTA, 1 mM dithiothreitol, and 0.1 mM phenylmethylsulfonyl fluoride) containing 10% glycerol and 0.3 M NaCl. After washing the column with more than 10× bed volume of the same buffer, the bound proteins were eluted by a linear gradient of 0.3–2 M NaCl [26]. The peak of the Mcm4/6/7 complex eluted at 0.75 M NaCl, while Mcm2 containing fractions appeared after 0.85 M NaCl. It has been known that among Mcm subunits Mcm2, 4 6 and 7 bind to histones, and Mcm2 has the highest affinity notably with Histone H3 [19,26]. Therefore, the Mcm4/6/7 can be separated from the Mcm2/4/6/7 heterotetramer through the histone column.

The histone column fractions containing the Mcm4/6/7 proteins were pooled and then concentrated about 10-fold with Centricon 30 (Amicon). The concentrated sample was diluted to 0.15 M NaCl with buffer C containing 10% glycerol and 0.01% Triton X-100 and then concentrated again to approximately 1 mg/ml with the Centricon 30 apparatus. The concentrated Mcm4/6/7 complex was further fractionated by glycerol gradient centrifugation at 36,000 rpm for 16 h (Beckman, TLS55 rotor) in a 15% to 35% linear glycerol gradient in buffer C containing 0.15 M NaCl and 0.01% Triton X-100, and peak fractions (out of 15 fractions each containing 133 μl) were pooled as the final preparation.

2.1.3.2. Purification from the cells coexpressing Mcm4, 6 and 7.

The Mcm4/6/7 complex was similarly expressed in High Five cells by coinfection of *6HisMcm4 + 6* and Mcm7-Flag recombinant baculoviruses at a multiplicity of infection of approximately 10 (1.0 ml