Methods 126 (2017) 54–65

Contents lists available at ScienceDirect

Methods

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

Overexpression and purification of Dicer and accessory proteins for biochemical and structural studies

Niladri K. Sinha *, Brenda L. Bass *

Department of Biochemistry, University of Utah, Salt Lake City, UT 84112, USA

ARTICLE INFO

Article history:
Received 15 May 2017
Received in revised form 4 July 2017
Accepted 14 July 2017
Available online 16 July 2017

Keywords:
RNAi
Dicer
Loqs-PD
R2D2
Baculovirus
Protein purification

ABSTRACT

The Dicer family of ribonucleases plays a key role in small RNA-based regulatory pathways by generating short dsRNA fragments that modulate expression of endogenous genes, or protect the host from invasive nucleic acids. Beginning with its initial discovery, biochemical characterization of Dicer has provided insight about its catalytic properties. However, a comprehensive understanding of how Dicer's domains contribute to substrate-specific recognition and catalysis is lacking. One reason for this void is the lack of high-resolution structural information for a metazoan Dicer in the apo- or substrate-bound state. Both biochemical and structural studies are facilitated by large amounts of highly purified, active protein, and Dicer enzymes have historically been recalcitrant to overexpression and purification. Here we describe optimized procedures for the large-scale expression of Dicer in baculovirus-infected insect cells. We then outline a three-step protocol for the purification of large amounts (3–4 mg of Dicer per liter of insect cell culture) of highly purified and active Dicer protein, suitable for biochemical and structural studies. Our methods are general and are extended to enable overexpression, purification and biochemical characterization of accessory dsRNA binding proteins that interact with Dicer and modulate its catalytic activity.

1. Introduction

The Dicer family of class 3 Ribonuclease III (RNase III) enzymes cleaves double-stranded RNA (dsRNA) substrates into short dsRNA fragments that regulate expression of endogenous genes, or protect cells from foreign nucleic acids [1]. Dicer enzymes contain an amino-terminal DExD/H-box helicase domain, followed by a domain of unknown function (DUF 283), the Platform and PAZ domains, two signature RNase III motifs, and a dsRNA-binding motif (dsRBM) at the carboxyl terminus [2–6]. The catalytic core of Dicer, comprising the tandem RNase III domains and the dsRBM, is derived from bacterial Class 1 RNase III enzymes [7].

Some organisms, including *Homo sapiens* and *Caenorhabditis elegans*, express a single isoform of Dicer that processes both long dsRNAs and precursor microRNA (miRNA) hairpins into mature small interfering RNAs (siRNAs) and miRNAs, respectively [1,8]. On the other hand, arthropods such as *Drosophila melanogaster* express two Dicer paralogs that exhibit functional specialization and generate different types of small regulatory RNAs [9,10]. The Dicer-1 enzyme (dmDcr-1) processes pre-miRNA hairpins by cleaving a portion of the terminal stem-loop structure to generate a small dsRNA consisting of the mature miRNA base-paired to miRNA*, which is subsequently discarded [9,11]. Dicer-2 (dmDcr-2) on the other hand, cleaves long dsRNAs of either cellular or viral origin into siRNAs that mediate an endogenous or antiviral RNA interference (RNAi) response [9,12,13]. Endogenous siRNAs (endo-siRNAs) are produced by dmDcr-2 from long dsRNAs of cellular origin that arise from transposon-derived sense-antisense pairs, convergent transcription of protein-coding genes, or long stem-loop structures [13]. This function of dmDcr-2 is modulated in vivo by accessory dsRNA binding proteins (dsRBPs) such as Loquacious-PD (Loqs-PD) [13–15]. Apart from the *endo*-siRNA pathway, dmDcr-2 also cleaves long dsRNAs from exogenous sources such as viral replication intermediates, or ectopically-introduced long dsRNAs. The resulting exogenous siRNAs (*exo*-siRNA) trigger an antiviral RNAi response that silences viral transcripts [12,16–18]. The products of dmDicer-1 cleavage, the mature miRNA/miRNA* duplexes, are sorted to the *Drosophila melanogaster* Argonaute-1 (dmAgo-1) RNA-induced silencing complex (RISC) where they mediate translational repression of protein-coding transcripts. By contrast, endo- and exo-siRNAs are preferentially sorted to the Argonaute-2 (dmAgo-2) RISC, where they mediate sequence-specific gene silencing by target cleavage and degradation [8,19,20].

* Corresponding authors.
E-mail addresses: niladri.sinha@biochem.utah.edu (N.K. Sinha), bbass@biochem.utah.edu (B.L. Bass).

http://dx.doi.org/10.1016/j.ymeth.2017.07.012
1046-2023/

Initial biochemical characterization of Dicer was performed in vitro using *Drosophila* cell and embryo-extracts – these early experiments demonstrated that Dicer activity in lysates processed radiolabeled dsRNAs into siRNAs, ~21–23 nucleotides (nts) in length, in an ATP-dependent manner [3,21]. The structural features of the product siRNAs – 5′ phosphates and 3′ hydroxyls, with two nt overhangs at the 3′ end, was the first hint that they were produced by a member of the RNase III family of enzymes [2]. Dicer activities were also reported in worms [4,5,22] and plants [23].

Biochemical characterization of Dicer using purified, recombinant protein was first demonstrated by Provost, Radmark and colleagues [24] and Filipowicz and colleagues [25] using *Homo sapiens* Dicer (hsDcr). Both groups purified recombinant hsDcr using the baculovirus expression system in insect cells. Several key observations were made by these groups, foremost among these, the recapitulation of processing of dsRNA into siRNA using purified recombinant Dicer. Filipowicz and colleagues [25] further demonstrated that limited proteolytic digestion of hsDcr enhanced its cleavage efficiency. Importantly, experiments with purified recombinant hsDcr showed that processing of dsRNA did not require ATP [24,25], contrary to studies performed with *Drosophila* cell lysates or purified dmDcr-2 [3,21,26]. Consistent with this observation, subsequent kinetic analyses of hsDcr revealed that the amino-terminal helicase domain attenuates cleavage of dsRNA or pre-miRNA substrates with 2-nt 3′overhanging termini [27]. Specifically, deletion of the helicase domain substantially enhances the catalytic efficiency of hsDcr [27]. Further biochemical analyses revealed that hsDcr processes pre-miRNA substrates more efficiently compared to longer dsRNA substrates [28]. Unlike hsDcr, studies performed with purified recombinant dmDcr-2 showed a profound effect of the helicase domain on substrate-specific kinetics [29–32]. For example, dsRNA with blunt termini, promoted an optimal reaction mode of dmDcr-2, whereby the helicase domain enabled ATP-dependent, processive cleavage. In contrast, dsRNA with 3′ overhanging termini, which mimic the termini of pre-miRNAs, promoted a suboptimal, distributive cleavage reaction [29,32].

A comprehensive understanding of how Dicer's domains contribute to substrate-specific recognition and cleavage requires high-resolution structural information. However, structural information on Dicer is currently limited to Dicer enzymes from lower eukaryotes [33], low-resolution electron microscopic (EM) reconstructions of human Dicer [34–36], and crystal structures of isolated domains [37,38]. The first structural insight towards understanding the three-dimensional architecture of Dicer came from the crystal structure of *Giardia intestinalis* Dicer [33]. While *Giardia* Dicer lacks the amino-terminal helicase domain, the crystal structure of this 'mini' Dicer revealed that the unique spatial arrangement of the PAZ domain, which anchors one end of a dsRNA, and catalytic RNase III domains, enables Dicer to function as a molecular ruler and produce dsRNA fragments of a specific length [33]. Further structural insights derive from low resolution (≥20–30 Å) EM studies of hsDcr, which revealed that hsDcr has an "L-shaped" architecture with the amino terminal helicase domain occupying the base of the L-shape [34–36]. Using a combination of domain-specific antibodies, tagged-proteins and deletion constructs, these studies also revealed that the terminus-recognition modules comprising the Platform and PAZ domains occupy the cap or the "tip" of the L-shape, whereas the catalytic core comprising the tandem RNase III domains occupy the central core of the enzyme [34,35]. However, there are currently no high-resolution structures of a full-length metazoan Dicer, and no RNA-bound structures of Dicer are available for any organism.

A bottleneck to biochemical and structural characterization of metazoan Dicers has been the difficulty in overexpressing and purifying large amounts of the active enzyme [39,40]. We overcame this bottleneck by systematically optimizing overexpression protocols of dmDcr-2 in baculovirus-infected insect cells [32]. Next, we established and optimized an efficient three-step purification protocol for dmDcr-2 that enabled us to isolate multimilligram amounts of highly purified dmDcr-2 per liter of baculovirus-infected insect cells – yields that far exceeded those of prior protocols [32]. The details of our expression and purification protocols are described below. Our protocols are straightforward, utilize commercially available resources, and have been optimized based on diverse trials and attention to detail. These protocols have greatly facilitated biochemical and structure-based experiments with Dicer.

2. Methods

2.1. Expression of Dicer proteins using an optimized baculovirus expression system

The Bac-to-Bac Baculovirus Expression System (Thermo Fisher Scientific) was used to generate high-titer recombinant baculoviruses expressing various proteins involved in gene-silencing, with modifications as described elsewhere [41–43]. For simplicity, the protocol described below focuses on generating high-titer recombinant baculovirus for *Drosophila melanogaster* Dicer-2 (dmDcr-2), but the methods described here are general, and with minimal optimization may be extended to other proteins. For example, we have used the same protocols to express and purify mutant variants and truncations of dmDcr-2, hsDcr, and accessory dsRBPs (Loqs-PD and R2D2 [26]) and dmAgo-2.

1. The pFastBac™1 vector (Thermo Fisher Scientific) was modified to include an N-terminal One-STrEP-FLAG (OSF) tag and a PreScission protease (GE Healthcare Life Sciences) cleavage site in frame with a Sma I (New England BioLabs Inc.) restriction endonuclease site: (MASWSHPQFEKGGGARGGSGGGSWSHPQFEKGF*DYKDDDDK*GTRSP**LEVLFQGP**; Strep-tag® II sequences (IBA Lifesciences) underlined, FLAG-tag sequence (Sigma-Aldrich) in italics, PreScission protease cleavage site in bold). The modified pFastBac™1 vector was a kind gift from Dr. Wesley I. Sundquist, University of Utah.
2. The pFastBac™1 vector was linearized by SmaI digestion, and the coding sequence for dmDcr-2 was seamlessly inserted using Gibson Assembly® (New England BioLabs Inc.) following the manufacturer's protocol. All plasmid constructs were verified by sequencing.
3. Recombinant bacmid DNA, created by site-specific transposition of the expression cassette from the pFastBac™1 vector into a baculovirus shuttle vector (bacmid), was generated, isolated, and verified as described in the Bac-to-Bac Expression System instruction manual (Thermo Fisher Scientific).
4. The pUC/M13 forward and reverse primers were used to verify the full-length dmDcr-2 cDNA in the recombinant bacmid DNA by PCR (Fig. 1A). This protocol is described in detail in the manufacturer's instruction manual and will not be repeated here.

Note: Recombinant bacmid DNA can be stored for up to 2 weeks at 4 °C in TE (10 mM TRIS pH 8.0, 1 mM EDTA) buffer. We do not recommend storing the purified bacmid DNA at −20 °C as the DNA is sheared during freeze/thaw cycles [44].

2.2. Generation of high-titer recombinant baculoviruses

2.2.1. General considerations for growing Spodoptera frugiperda (Sf9) cells

Suspension cultures of Sf9 cells (Cat# 94-001S), growing in ESF 921 media (Cat# 96-001), were purchased from Expression