Virus Research 245 (2018) 7–16

Contents lists available at ScienceDirect

Virus Research

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

Gp41, a superfamily SF2 helicase from bacteriophage BFK20

Nora Halgasova, Radka Matuskova, Daniel Kraus, Adela Tkacova, Lenka Balusikova, Gabriela Bukovska*

Department of Genomics and Biotechnology, Institute of Molecular Biology, Slovak Academy of Sciences, Dubravska cesta 21, 845 51, Bratislava, Slovakia

ARTICLE INFO

Keywords:
Bacteriophage
Corynebacterium
DNA replication
Helicase
Protein expression
Replication protein

ABSTRACT

Gp41 is one of two helicases encoded by the genome of bacteriophage BFK20. The gp41 sequence contains conserved motifs from the SF2 family of helicases. We prepared and studied three recombinant proteins: gp41HN, a wild type-like protein with an N-terminal His-Tag; gp41HC, with an S2A mutation and a C-terminal His-Tag; and gp41dC, a mutant protein with a deleted C-terminal region and His-Tags on both N- and C-termini. We tested the enzymatic activities and DNA binding abilities of these isolated proteins. We found that both gp41HN and gp41HC had strong DNA-dependent ATPase activities, but that the ATPase activity of gp41dC was significantly lower regardless of the presence of DNA. The preferred substrates for the NTP hydrolysis reactions were ATP and dATP. gp41HC and gp41HN exhibited a low helicase activity in a fluorescence-based assay using dsDNA substrates with a 3′ overhang and with a forked end in the presence of ATP. We infer that the C-terminal region of gp41 may be involved in DNA binding, since removing this region in gp41dC reduced the protein's DNA binding ability.

1. Introduction

Helicases are molecular motor proteins that use the energy from nucleoside triphosphate hydrolysis to catalyse the separation of the complementary strands of double-stranded nucleic acids. Based on conserved motifs and comparative structural and functional analysis (Gorbalenya and Koonin, 1993; Singleton et al., 2007), helicases have been divided into six superfamilies (SF). The two largest groups, SF1 and SF2, comprise enzymes that operate as monomers or dimers, while families SF3 to SF6 contain helicases that function as ring-shaped oligomers, mostly hexamers (Singleton et al., 2007). SF1 and SF2 helicases contain a conserved helicase core consisting of two RecA-like folds with SF1 and SF2 conserved motifs, and accessory N- or C-terminal domains. Recently, Fairman-Williams et al. (2010a,b) proposed a new categorization of the SF1 and SF2 helicases based on both the helicase core and the accessory domains.

SF2 is the largest helicase superfamily. Helicases from this family are involved in many aspects of nucleic acid metabolism. SF2 DNA helicases participate in DNA replication, recombination and repair, maintaining genome stability, chromatin remodelling, and Holliday junction movement (Briggs et al., 2004; Rezazadeh, 2011; Gabbai and Marians, 2010; LeRoy et al., 2005; Saha et al., 2006). Mutations in genes coding for SF2 helicases result in serious human diseases, including premature aging and several forms of cancer (Harrigan and Bohr, 2003; Suhasini and Brosh, 2013). SF2 family helicases from different kinds of organisms (eukaryotes, prokaryotes, archaea, bacteriophages) have many common features that are related to their function and specificity.

Bacteriophages are a favourite model system to study DNA replication in prokaryotes, and they exhibit examples of every theoretically possible replication mechanism (Weigel and Seitz, 2006). The different molecular mechanisms driving phage DNA replication arise from the high diversity of phage-encoded replication proteins. The replication mechanism of a given bacteriophage depends not only on its own replication machinery, but also on its ability to recruit replication proteins from its bacterial host. Bacteriophage genomes have a modular structure, and phage genes encoding replication functions tend to be located close to each other in many phage genomes, resulting in phage replication modules. Four major types of phage replication modules have been identified: those containing (1) initiator genes, (2) DNA polymerase genes, (3) ΦP4α-type helicase-primase genes and (4) filamentous phage modules (Weigel and Seitz, 2006). Type (2) appears in phages that encode individual DNA polymerases, and their genomes contain genes encoding one or two helicases and a primase. Most frequently, one of these helicases is an SF4 replicative helicase, while the second belongs to the SF2 family; occasionally genes encoding helicases similar to the ΦP4α protein or from the SF1 family can also be found (Weigel and Seitz, 2006). The SF4 family helicases and those similar to

* Corresponding author.
E-mail address: Gabriela.Bukovska@savba.sk (G. Bukovska).

https://doi.org/10.1016/j.virusres.2017.12.005
Received 25 July 2017; Received in revised form 30 October 2017; Accepted 8 December 2017
Available online 14 December 2017
0168-1702/

Table 1
Oligonucleotides used for the preparation of helicase substrates.

Oligonucleotide	Sequence (5′→3′)	Fluorophore	Labelled end
A	TAGTACCGCCACCCTCAGAACCTTTTTTTTTTTTTTTTTTTTT	FAM	5′
B	TAGTACCGCCACCCTCAGAACC	FAM	5′
C	GGTTCTGAGGGTGGCGGTACTA	BHQ1	3′
D	TTTTTTTTTTTTTTTTTTTTTTGGTTCTGAGGGTGGCGGTACTA	BHQ1	3′
E	TTTTTTTTTTTTTTTTTAAAAAGGTTCTGAGGGTGGCGGTACTA	BHQ1	3′
Capture strand I	TAGTACCGCCACCCTCAGAACC		
Capture strand II	TAGTACCGCCACCCTCAGAACCTTTTT		

the phage P4 alpha primase-helicase serve as replicative helicases for phage DNA replication, while the SF1 and, mainly, the SF2 helicases have auxiliary functions in replication, restart, DNA repair and recombination.

SF2 helicase genes are found in many phage genomes, however, information about the structure and function for most of these proteins is limited. The most heavily studied protein is UvsW, an SF2 family helicase, encoded by the T4 phage genome. UvsW plays a unique role in phage DNA replication, recombination and repair, and, in addition to replicative helicase gp41 from the SF4 family, and the SF1 family Dda protein, is the third helicase encoded by the T4 genome. UvsW has ssDNA-dependent ATPase and branched-DNA unwinding activities (Carles-Kinch et al., 1997). The same authors found that UvsW is a functional analog of the bacterial DNA helicase RecG, and is able to dissociate RNA from R-loops. More recently, it was demonstrated that UvsW operates as a molecular switch in T4 DNA replication. T4 early origin-dependent replication uses R-loops as initiation sites. UvsW unwinds R-loops from T4 origins using its RNA-DNA helicase activity and causes a transition from origin-dependent replication to late recombination-dependent replication (Dudas and Kreuzer, 2001). UvsW unwinds a variety of DNA/DNA and RNA/DNA substrates with a preference for stalled replication forks and recombination intermediates and, additionally, contains single-stranded DNA (ssDNA) annealing activity (Nelson and Benkovic, 2007).

In recent years, two SF2 family helicases from bacteriophage T5 have been characterized. D2 is an unusual bipolar helicase with both 3′→5′and 5′→3′ unwinding activities. Unwinding of DNA substrates in the 3′→5′ direction is more robust and can be distinguished from the 5′→3′ activity by a number of features, including helicase complex stability, salt sensitivity and the length of the ssDNA overhang required to initiate helicase activity (Wong et al., 2013). D10 has sequence and functional similarities with T4 UvsW, and possesses branch migration and DNA unwinding activities. Although the DNA binding and DNA-dependent ATPase activity of D10 did not show any sequence specificity, initiation of substrate unwinding did show some sequence specificity (Wong et al., 2016). In addition to D2 and D10, a DnaB-like protein, similar to the *Escherichia coli* replicative helicase DnaB, is encoded by the T5 genome (Wong et al., 2013). Based on replication module composition, both T4 and T5 phages seem to be of type (2).

The T4 DNA replication machinery is different from both the bacterial and the archaeal/eukaryotic one, with some components distantly related to bacterial proteins (e.g., the primase and replicative helicase), while others are more related to eukaryotic ones (e.g., the polymerase, clamp and clamp loader, DNA ligase, and type II DNA topoisomerase) (Forterre, 2013). However, in both cases the phage proteins exhibit low sequence similarities to their cellular homologues, being only members of the same protein family, such as Toprim primase or family B DNA polymerase. The study of the DNA replication machinery of complex DNA viruses encoding their own replication proteins would provide important information for identifying novel proteins and mechanisms even within cellular systems, and would enlarge our picture of the world of replicons.

Bacteriophage BFK20 is a lytic phage of the L-lysine producer *Brevibacterium flavum* CCM 251. The genome of this phage has been completely sequenced and annotated (EMBL AJ278322, Bukovska et al., 2006), and potential ORFs have been identified. Clusters of functionally related putative replication, regulatory, structural and lytic genes were defined. The genes ORF29-ORF46 appear to be a replication module for BFK20. Like bacteriophages T4 and T5, the BFK20 replication module belongs to type (2), encoding a putative DNA polymerase A (gp44), a RepA-like protein with a prim-pol domain and an SF4-type helicase domain (gp43), and an SF2 family helicase (gp41). However, although the replication modules of bacteriophages T4, T5 and BFK20 are of the same type and the proteins they encode have related functions, these proteins exhibit very different sequences and different closest homologues, meaning they most likely also have different origins. It is therefore very likely that their replication mechanisms and the specific functions of individual proteins, including those of the helicases, are different. Proteins gp41 and gp43 were characterised as putative BFK20 helicases. Recently, we detected ssDNA-dependent ATPase and helicase activities for gp43, and we suggested that this protein functions as an SF4 family replicative helicase (Halgasova et al., 2015). SF4 replicative helicases unwind the template DNA for polymerases at the replication fork and are responsible for unwinding the majority of dsDNA genomes, but accessory replicative helicases, mainly from the SF2 family, aid replication repair and restart. The second helicase-like protein from BFK20 is gp41. This protein contains conserved motifs from the SF2 family and a C-terminal region of unknown function. In our previous work, we detected interactions between gp41 and the host proteins DnaZX, DnaN, Dnaδ, DnaG, and SSB, and we showed that the strongest interaction was between gp41 and DnaG (Solteszova et al., 2015). In this work, we demonstrate that gp41 has ATPase and helicase activities, and we show that the gp41 C-terminal region has a DNA binding function. This more detailed characterization of gp41 contributes to understanding the helicase's function in BFK20 genome replication, and, in addition, the function of SF2 family helicases in general.

2. Materials and methods

2.1. Bacteria, bacteriophage, growth conditions

Brevibacterium flavum CCM 251 (hse^-, Aec^r) were used for propagation of bacteriophage BFK20 according to Koptides et al. (1992). Isolation of BFK20 phage particles and phage DNA was performed according to Sambrook and Russel (2001).

Escherichia coli XL1 Blue (Stratagene), used for cloning, were grown in LB medium supplemented with 100 µg/ml kanamycin at 37 °C. *Escherichia coli* BL21 (DE3) (Novagen), used for recombinant protein expression, were grown in TB medium supplemented with 100 µg/ml kanamycin at 30 °C or 37 °C.

2.2. Cloning, expression and isolation of recombinant proteins

For expressing recombinant gp41HC, we prepared plasmid pET28-41HC. The putative gene ORF41 was amplified using PCR with BFK20 DNA as a template on a T-Gradient thermal cycler (Whatman Biometra). The forward primer 5′-TTCCATGGCTGTGAAGCCCCG