Protein Expression and Purification 95 (2014) 28–37

Contents lists available at ScienceDirect

Protein Expression and Purification

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

Strategies for purifying variants of human rhinovirus 14 2C protein

Tomáš Sára [a], Robert Konrat [b], Tim Skern [a,*]

[a] Max F. Perutz Laboratories, Medical University of Vienna, Dept. of Medical Biochemistry, Dr. Bohr-Gasse 9/3, A-1030 Vienna, Austria
[b] Max F. Perutz Laboratories, University of Vienna, Dept. of Structural and Computational Biology, Campus Vienna Biocenter 5, A-1030 Vienna, Austria

ARTICLE INFO

Article history:
Received 8 August 2013
and in revised form 14 November 2013
Available online 5 December 2013

Keywords:
Picornaviruses
Meta-structure analysis
Insoluble protein purification
Zinc finger
AAA+ superfamily
Helicase

ABSTRACT

The positive strand RNA genome of picornaviruses, including human rhinovirus (HRV), poliovirus (PV) and foot-and-mouth disease virus, is translated immediately into a polyprotein that is cleaved by virally encoded proteinases into 10–13 mature proteins. These include the four proteins required to assemble the viral particle as well as $3D^{pol}$ (the viral RNA polymerase) and 2C, an ATPase and putative helicase. 2C is a protein which is responsible, together with 2B and 3A, for anchoring the replication complexes to membranous structures in the infected cell on which RNA replication takes place. Additionally, expression of 2C and its precursor 2BC in mammalian cells leads to vesicle formation observed in infected cells. 2C is encoded by all picornaviruses; nevertheless, its exact role in viral replication remains unclear. A contributing factor is the absence of structural data for this hydrophobic protein the generation of which has been hampered by an inability to produce soluble and stable material. Here, we compare 2C from several genera and show that the 2C protein has considerable heterogeneity. Using protein structure meta-analysis, we developed models of HRV14 2C that should be useful for mutational analysis. Based on these analyses, we expressed and purified two domains of HRV14 2C using three different protocols and examined the folding by thermal denaturation or ^{1}H NMR. Both domains were concentrated sufficiently to allow crystal screens or NMR pilot experiments to be performed. This work provides a platform to explore 2C proteins from all picornaviral genera to generate candidates for structural analysis.

Introduction

The family of picornaviruses contains important pathogens that cause severe diseases in humans and animals such as hepatitis A virus (HAV)[1], foot-and-mouth disease virus (FMDV), human rhinovirus (HRV) and poliovirus (PV) [1]. Until a few years ago, only five genera of this family had been defined. Recent developments in sequencing technology have led to identification of a further 10 genera with a variety of host and tissue tropisms [2]. Even though the positive strand RNA genome of picornaviruses is translated into one single polyprotein that is subsequently processed to between 10 and 13 mature proteins, there is little knowledge linking the properties of the proteins to host and tissue tropism. It will be a major challenge for the field to identify such connections for all genera in the coming years.

The picornaviral polyprotein of the open reading frame consists of three major regions. The first, termed P1, encodes the four proteins that build the viral capsid whereas the two other regions, termed P2 and P3, contain viral enzymes and other non-structural proteins necessary for RNA replication or for influencing the host cell. The polyprotein is then stepwise processed by viral proteases within the polyprotein to generate the mature proteins [3].

The P2 polyprotein precursor contains proteins $2A^{pro}$, 2B and 2C. In enteroviruses (i.e., PV, coxsackieviruses and HRVs; these were added to the enterovirus genus 2012 [4]), the initial cleavage between the P1 and P2 regions is performed by the proteinase 2A [5]; subsequently, the proteinase 3C frees the C-terminus of 2C from the N-terminus of 3A to generate the 2ABC precursor. 3C further processes 2ABC to generate 2A, 2B, 2BC and 2C. Cleavage of 2BC is very slow, so that this protein has been termed a long-lived precursor [6]. In other genera, such as aphthoviruses and cardioviruses, the separation of 2A from 2B is achieved by a ribosomal skip mechanism, leaving 2A attached to VP1 at the end of P1. 3C again cleaves at the C-terminus to generate 2BC and 2C [7].

In all picornaviruses studied, both 2BC and 2C are a part of the viral RNA replication complex, which localises to clusters of membrane vesicles formed by host cell endoplasmic reticulum and Golgi membranes [8,9]. Expression of 2BC and 2C in various mammalian cells demonstrated their relevance for the replication

* Corresponding author. Tel.: +43 1 4277 61620; fax: +43 1 4277 9616.
E-mail address: timothy.skern@meduniwien.ac.at (T. Skern).

[1] *Abbreviations used:* MBP, maltose binding protein; EC, size exclusion chromatography; IPTG, isopropyl-β-d-thiogalactopyranoside; HSQC, heteronuclear single quantum spectroscopy; DLS, dynamic light scattering; DSF, differential scanning fluorescence; HAV, hepatitis A virus; FMDV, foot-and-mouth disease virus; HRV, human rhinovirus; PV, poliovirus; SF3, superfamily 3 of helicases.

1046-5928/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.pep.2013.11.012

of the virus and led to the formation of membrane vesicles [10,11]. However, only when 2BC and 2C are expressed together are vesicles produced that resemble those found during viral replication [12,13]. Membrane binding by PV 2C is supposedly mediated by the amphipathic helix at its N-terminus [14,15]; on its own, this domain can form an independent structural element that can tether the protein to a membrane surface [16]. The 2C protein from enteroviruses has also been proposed to bind RNA, as a cysteine-rich motif at the 2C C-terminus of enteroviruses has been shown to bind zinc, invoking the possibility of an RNA binding zinc finger [17]. This interaction with RNA has been proposed to be highly specific [18].

Global sequence alignments have classified picornaviral 2C as a member of AAA+ superfamily of ATPases [19,20]. These are often referred to as novel chaperones, performing a large variety of functions such as protein quality control, cytoskeletal movement and nucleic acid unwinding. Fittingly, several groups have described ATPase activity for PV 2C; a GTPase activity has also been described [19,21–23]. As many members of the AAA+ superfamily are nucleic acid helicases [24], a helicase function has also been proposed for the 2C; however, helicase activity has not yet been demonstrated [23,25]. AAA+ helicases form oligomeric rings as a prerequisite for ATPase activity [26]; oligomerization has been reported for both FMDV and PV 2C [27,28].

In spite of its vitality to the life cycle of picornaviruses and the importance as a potential drug target, picornaviral 2C still remains a poorly characterised protein compared to other non-structural picornaviral proteins [29] and its exact role in viral RNA replication has not been elucidated [30]. Indeed, reports on the production of soluble 2C are rare and attention has been paid almost exclusively to PV 2C. Nevertheless, only two reports [22,27] describe soluble PV 2C protein using MBP and GST tags, respectively. The FMDV protein was produced as with a deletion of the N-terminal amphipathic helix [20]. As a consequence, Springer et al. [19] resorted to *in vitro* translation of 2C in *Escherichia coli* extracts in the presence of nano-discs to determine biochemical parameters of PV 2C. However, such amounts are not suitable for structural studies.

The inability to express large amounts of 2C and the resulting paucity of biochemical and structural data stems mostly from the intrinsic insolubility and instability of the protein, making it difficult to purify and concentrate to levels required for structural analysis. Here, we attempted to solve some of these obstacles with the 2C protein from HRV14 and develop protocols that allow selected domains based partly on bioinformatic analysis to be purified to homogeneity and used at concentrations suitable for structural analysis.

Materials and methods

In silico analysis

Protein sequences were obtained from the UniProt database [31], aligned using the T-Coffee server [32] and rendered in Jalview desktop [33]. Single protein alignments were performed using the BLAST server [34]. Tertiary structure models were generated by the Swiss model server [35]. Structural visualizations were processed in PyMol [36].

Generation of meta-structure parameters

Meta-structure parameters for human rhinovirus 14 2C protein were calculated as described by Konrat [37] and as exemplified by Mayer et al. [38]. In brief, in this approach, proteins are conceived as networks in which individual amino acids represent nodes whereas edges connecting two nodes indicate spatial proximity in 3D structures. The mutual topological relationship θ of two residues is quantified by the shortest path length connecting the two residues in the network. The value of θ between two residues (A,B) characteristically depends on the amino acid types (A,B) and their primary sequence distance, l_{AB}. This characteristic topological relationship is statistically evaluated following a well established statistical procedure using a subset of the PDB structural database and stored as pairwise statistical distribution functions $r(\theta,A,B,l_{AB})$. The calculation of meta-structure parameters is as follows. The input primary sequence is used to predict for each possible amino acid pair (of residue type A and B and separated by l_{ab} in the sequence) in the protein and based on $r(q,A,B,l_{AB})$ an average topology parameter, d_{ij} which is related to the most probable shortest path length between residues i and j. The subsequent summation (for all residue pairs) provides the residue-specific compactness value C_i. For the prediction of local secondary structure elements only next neighbour distribution functions (restricted to primary sequence differences between residue pairs $\leqslant 4$) are used. Numerical details of the calculation procedure can be found in the original manuscript [37].

DNA constructions

For the preparation of 14CΔ38 (containing amino acids 39–330 of the HRV14 2C sequence) and 14C(119–249) (containing amino acids 119–249 of the HRV14 2C sequence), the respective sequences were amplified by PCR from pHRV14 plasmid [39] and inserts were ligated into the pETM11 plasmid (EMBL) carrying an N-terminal cleavable His-tag; for the construction of MBP-14CΔ38, the plasmid pETMBP_1a (EMBL) was used (Supplementary Fig. S1). All primers used carried upstream and downstream restriction sites for *Nco*I and *Xho*I restriction enzymes, respectively. All three plasmids contain the TEV protease cleavage site separated by a 9 amino acid linker from their respective N-terminal tags.

Expression

All three plasmids were transformed into *E. coli* BL21(DE3); single colonies were picked and cells grown in LB medium at 37 °C to OD_{600} 0.5–0.6. For expression of 14CΔ38 and MBP-14CΔ38, the cultures were incubated on ice for 30 min, induced with 0.1 mM IPTG and incubated overnight at 15 °C. Cultures expressing the 14C(119–249) plasmid were induced with 0.7 mM IPTG and incubated at 37 °C for 4–12 h. For the purposes of HSQC NMR measurement, the 14C(119–249) transformed cells were grown in M9 minimal medium containing ^{15}N-labelled ammonium chloride as a sole source of nitrogen; protein expression was carried out as for non-labelled material. All cultures were harvested by centrifugation at 5000*g* and pellets corresponding to 1 L of bacterial culture were stored at −20 °C.

Purification of 14CΔ38 protein

The bacterial pellet was resuspended in 10 ml 50 mM sodium phosphate, pH 6.5 and 10% glycerol; the suspension was lysed using a French press. 0.1% Triton X-100 and 2 μg/ml DNAse I were added to the crude lysate which was then incubated on ice for 30 min followed by centrifugation at 35,000*g* for 30 min at 4 °C.

For nickel affinity chromatography, 50 mM imidazole was added to the soluble fraction of the bacterial lysate and this was applied onto a HisTrap FF 1 ml column (GE Healthcare). Unbound proteins were washed off with 20 column volumes of the lysis buffer containing 80 mM imidazole and the elution was carried out in five 2 ml fractions of the lysis buffer with 300 mM imidazole added.