Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Preliminary crystallographic characterization of an RNA helicase from Kunjin virus

Eloise Mastrangelo,[a] Michela Bollati,[a] Mario Milani,[a] Nadège Brisbarre,[b] Xavier de Lamballerie,[b] Bruno Coutard,[c] Bruno Canard,[c] Alexander Khromykh[d] and Martino Bolognesi[a]*

[a]Department of Biomolecular Sciences and Biotechnology, CNR–INFM, University of Milano, Via Celoria 26, 20133 Milano, Italy, [b]Unité des Virus Emergents, Faculté de Médecine, 27 Boulevard Jean Moulin, 13005 Marseille, France, [c]Laboratoire Architecture et Fonction des Macromolécules Biologiques, UMR 6098 CNRS ESIL, Case 932, 163 Avenue de Luminy, 13288 Marseille CEDEX 9, France, and [d]RNA Virology Laboratory, School of Molecular and Microbial Sciences, The University of Queensland, Brisbane, QLD 4072, Australia

Correspondence e-mail: martino.bolognesi@unimi.it

Received 26 June 2006
Accepted 24 July 2006

Kunjin virus is a member of the *Flavivirus* genus and is an Australian variant of West Nile virus. The C-terminal domain of the Kunjin virus NS3 protein displays helicase activity. The protein is thought to separate daughter and template RNA strands, assisting the initiation of replication by unwinding RNA secondary structure in the 3′ nontranslated region. Expression, purification and preliminary crystallographic characterization of the NS3 helicase domain are reported. It is shown that Kunjin virus helicase may adopt a dimeric assembly in absence of nucleic acids, oligomerization being a means to provide the helicases with multiple nucleic acid-binding capability, facilitating translocation along the RNA strands. Kunjin virus NS3 helicase domain is an attractive model for studying the molecular mechanisms of flavivirus replication, while simultaneously providing a new basis for the rational development of anti-flaviviral compounds.

1. Introduction

Positive-strand RNA viruses include the large *Flaviviridae* family responsible for widespread diseases of man and animals. The virus genome consists of a single-stranded RNA with positive polarity; following infection of the host cells, the viral genome is transcribed to negative-strand RNA. The daughter virus genomic RNA is subsequently synthesized using the negative strand as template. To maintain proper viral replication, the nascent transcripts must be unwound from their complementary template RNA soon after transcription. However, both the positive and negative strands of template RNA are highly structured, especially at the 5′ and 3′ termini. Thus, to synthesize RNA the virus-replication machinery must be able to unwind the RNA secondary structure in the template RNAs.

The *Flaviviridae* family is composed of three different genera (*Hepacivirus*, *Pestivirus* and *Flavivirus*), the latter being the largest and including more than 70 different viruses, of which more than half have been shown to be pathogenic to humans. The flavivirus genome encodes a 370 kDa polyprotein precursor, which is inserted into the membrane of the endoplasmic reticulum and processed by cellular and viral proteases to yield three mature structural proteins (C, M and E) and seven nonstructural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5) (Fields *et al.*, 2001). The C-terminal 440 amino acids of NS3 code for a protein endowed with helicase activity (Ferron *et al.*, 2005). Although the precise biological functions of this encoded protein are unknown, it is thought to separate RNA daughter and template strands, assisting the initiation of replication by unwinding the RNA secondary structure in the 3′ nontranslated region (NTR; Takegami *et al.*, 1995). Little is known about the unwinding mechanism of flavivirus helicases, including their specificity for DNA and/or duplex RNA and requirement for a 3′ or 5′ overhang.

Kunjin virus (KUNV) is an Australian variant of West Nile virus. KUNV was first isolated from *Culex annulirostris* mosquitoes collected in north Queensland in 1960 (Scherret *et al.*, 2001). Multiple sequence alignments show that KUNV-helicase displays the seven conserved sequence motifs of the helicase superfamily 2 associated with NTP hydrolysis and nucleic acid binding (Caruthers *et al.*, 2002) (Fig. 1). In particular, the most conserved motifs, I (GAGKTRR) and II (DEAH), also known as Walker A and Walker B motifs (Walker *et*

al., 1982), are located in the amino-terminal region of KUNV-helicase, as reported for dengue virus helicase, where they interact with the NTP substrate and Mg^{2+}, respectively (Xu *et al.*, 2005).

We have expressed, purified and crystallized the C-terminal helicase domain of KUNV NS3 (434 residues; KUNV-helicase). Here, we report the characterization of the protein crystals together with data from preliminary X-ray diffraction experiments. We show that contrary to other flavivirus helicases reported to date (Wu *et al.*, 2005; Xu *et al.*, 2005), KUNV-helicase may be dimeric even in the absence of nucleic acids. Oligomerization has been proposed as a means to provide helicases with multiple nucleic acid-binding capability, facilitating the translocation of the protein along the RNA strands (Cho *et al.*, 1998; Velankar *et al.*, 1999). Insight into the crystal structure of KUNV-helicase may shed new light on the molecular mechanisms of flavivirus replication and provide a new basis for the rational development of anti-flaviviral compounds.

2. Materials and methods

2.1. Preparation of biological material

KUNV strain MRM-16 was propagated into SW13 human cells cultured as monolayers at 310 K under 5% CO_2 in Eagle's minimum essential medium with 10% calf serum and penicillin G, kanamycin

```
kunjin          KKQITVLDLHPGAGKTRRILPQIIKEAINRRLRTAVLAPTRVVAAEMAEALRGLPIRYQT 60
west            KKQITVLDLHPGAGKTRKILPQIIKEAINKRLRTAVLAPTRVVAAEMSEALRGLPIRYQT 60
dengue          KKRLTIMDLHPGAGKTKRYLPAIVREAIKRGLRTLILAPTRVVAAEMEEALRGLPIRYQT 60
yellow          KGMTTVLDFHPGAGKTRRFLPQILAECARRRLRTLVLAPTRVVLSEMKEAFHGLDVKFHT 60
tick-borne      KGQITVLDMHPGSGKTHRVLPELIRQCIDRRLRTLVLAPTRVVLKEMERALNGKRVRFHS 60
                *   *::*:***:***:: ** :: :.  : *** :******* ** .*:.*  :::::
                                  I                   Ia

kunjin          SAVAREHNGNEIVDVMCHATLTHRLMSPHRVPNYNLFVMDEAHFTDPASIAARGYISTRV 120
west            SAVHREHSGNEIVDVMCHATLTHRLMSPHRVPNYNLFIMDEAHFTDPASIAARGYIATKV 120
dengue          PAIRAEHTGREIVDLMCHATFTMRLLSPIRVPNYNLIIMDEAHFTDPASIAARGYISTRV 120
yellow          QAFSAHGSGREVIDAMCHATLTYRMLEPTRVVNWEVIIMDEAHFLDPASIAARGWAAHRA 120
tick-borne      PAVSDQQAGGAIVDVMCHATYVNRRLLPQGRQNWEVAIMDEAHWTDPHSIAARGHLYTLA 120
                 *.  .  *  ::* ***** . * : *    *::: :*****: ** ******     .
                                                  II

kunjin          ELGEAAAIFMTATPPGTSDPFPESNAPISDLQTEIPDRAWNSGYEWITEYIGKTVWFVPS 180
west            ELGEAAAIFMTATPPGTSDPFPESNAPISDMQTEIPDRAWNTGYEWITEYVGKTVWFVPS 180
dengue          EMGEAAGIFMTATPPGSRDPFPQSNAPIMDEEREIPERSWNSGHEWVTDFKGKTVWFVPS 180
yellow          RANESATILMTATPPGTSDEFPHSNGEIEDVQTDIPSEPWNTGHDWILADKRPTAWFLPS 180
tick-borne      KENKCALVLMTATPPGKSEPFPESNGAITSEERQIPDGEWRDGFDWITEYEGRTAWFVPS 180
                . .:.* ::*******. : **.**. * . : :**.  *. *.:*:      *.**:**
                           III                                    IV

kunjin          VKMGNEIALCLQRAGKKVIQLNRKSYETEYPKCKNDDWDFVVTTDISEMGANFKASRVID 240
west            VKMGNEIALCLQRAGKKVIQLNRKSYETEYPKCKNDDWDFVITTDISEMGANFKASRVID 240
dengue          IKTGNDIAACLRKNGKRVIQLSRKTFDSEYVKTRTNDWDFVVTTDISEMGANFKAERVID 240
yellow          IRAANVMAASLRKAGKSVVVLNRKTFEREYPTIKQKKPDFILATDIAEMGANLCVERVLD 240
tick-borne      IAKGGAIARTLRQKGKSVICLNSKTFEKDYSRVRDEKPDFVVTTDISEMGANLDVSRVID 240
                :  .. :*  *:: ** *: *. *::: :*   : .. **:::***:*****: ..**:*
                                                          V

kunjin          SRKSVKPTIITEGEGRVILGEPSAVTAASAAQRRGRTGRNPSQAGDEYCYGGHTNEDDSN 300
west            SRKSVKPTIIEEGDGRVILGEPSAITAASAAQRRGRIGRNPSQVGDEYCYGGHTNEDDSN 300
dengue          PRRCMKPVILTDGEERVILAGPMPVTHSSAAQRRGRIGRNPRNENDQYIYMGEPLENDED 300
yellow          CRTAFKPVLVDEGR-KVAIKGPLRISASSAAQRRGRIGRNPNRDGDSYYYSEPTSENNAH 299
tick-borne      GRTNIKP---EEVDGKVELTGTRRVTTASAAQRRGRVGRQDGR-TDEYIYSGQCDDDDSG 296
                 *  .**    :   :* :  .  :: :******** **:  .  *.* *     :::
                                      VI

kunjin          CAHWTEARIMLDNINMPNGLIAQFYQPEREKVYTMDGEYRLRGEERKNFLELLRTADLPV 360
west            FAHWTEARIMLDNINMPNGLVAQLYQPEREKVYTMDGEYRLRGEERKNFLEFLRTADLPV 360
dengue          CAHWKEAKMLLDNINTPEGIIPSIFEPEREKVDAIDGEYRLRGEARKTFVDLMRRGDLPV 360
yellow          HVCWLEASMLLDNMEVRGGMVAPLYGVEGTKTPVSPGEMRLRDDQRKVFRELVRNCDLPV 359
tick-borne      LVQWKEAQILLDNITTLRGPVATFYGPEQDKMPEVAGHFRLTEEKRKHFRHLLTHCDFTP 356
                 . * ** ::***:    * :. ::  *  *     *. **  : ** * .::   *:.


kunjin          WLAYKVAAAGVSYHDRRWCFDGPRTNTILEDNNE-VEVITKLGERKILRPRWIDARVYSD 419
west            WLAYKVAAAGISYHDRKWCFDGPRTNTILEDNNE-VEVITKLGERKILRPRWADARVYSD 419
dengue          WLAYKVAAEGINYADRRWCFDGTRNNQILEENVE-VEIWTKEGERKKLKPRWLDARIYSD 419
yellow          WLSWQVAKAGLKTNDRKWCFEGPEEHEILNDSGETVKCRAPGGAKKPLRPRWCDERVSSD 419
tick-borne      WLAWHVAANVSSVTDRSWTWEGPEANAVDEASGDLVTFRSPNGAERTLRPVWKDARMFKE 416
                **:::**    .  ** * ::*.. : : : . : *   :  * .: *:* * * *: .:


kunjin          HQALKSFKDFASGKR 434
west            HQALKSFKDFASGKR 434
dengue          PLALKEFKEFAAGRK 434
yellow          QSALSEFIKFAEGRR 434
tick-borne      GRDIKEFVAYASGRR 431
                   :..*  :* *::
```

Figure 1
Structural alignment of *Flavivirus* helicase domain of Kunjin virus (kunjin), West Nile virus (west), dengue virus (dengue), yellow fever virus (yellow) and tick-borne encephalitis virus (tick-borne). The conserved motifs among superfamily 2 helicases are indicated in pink (motif I), green (motif II) and grey (motifs Ia, III, IV, V and VI).

and streptomycin at 100 IU ml^{-1}, 100 µg ml^{-1} and 100 µg ml^{-1}, respectively. Viral RNA was isolated using the EZ1 Virus Mini Kit (Qiagen) onto Biorobot EZ1 automat (Qiagen) according to the manufacturer's recommendations.

2.2. Expression and purification

Classical one-step RT-PCR reactions were carried out with the Access RT-PCR System (Promega) and primers KUN-NS3-S1 (5′-AACATCAGGCTCGCCCATAG, located in the protease domain of NS3) and KUN-NS3-R2 (5′-TAGCTCCTCCAAAGCCATTCT, located in the NS4A gene). A 1600 bp PCR product was obtained that includes the complete helicase domain. It was re-amplified using Gateway-modified primers (KUN-He-HN-S, 5′-GGGGACAAGTTTGTACAAAAAAGCAGGCTTCGAAGGAGATAGAACCATGAAACATCACCATCACCATCAC**AAAAAACAGATCACCGTTTTGGATCTT**; KUN-He-HN-R, 5′-GGGGACCACTTTGTACAAGAAAGCTGGGTATTA**GCGCTTCCCTGAGGCGAAGTCCTT**; nucleotides in bold correspond to viral sequences) designed to encode an N-terminal His_6-tagged recombinant protein spanning amino acids 186–620 of the NS3 gene (*i.e.* the complete helicase domain up to the NS3-NS4A cleavage site). This product was subsequently cloned into pDEST 14 HN expression vector (Invitrogen). The construct was transformed into *Escherichia coli* Rosetta (DE3) pRos strain. The cells were grown at 310 K in 3 l SB medium containing 100 µg ml^{-1} ampicillin and 34 µg ml^{-1} chloramphenicol to an OD_{600} of 0.5. Upon induction with 0.5 m*M* isopropyl β-D-thiogalactopyranoside (IPTG), the incubation temperature was lowered to 290 K and cells were grown for an additional 16 h. Cells harvested by centrifugation were resuspended in 80 ml lysis buffer [300 m*M* NaCl, 50 m*M* Tris–HCl pH 8, 25 m*M* imidazole, 0.25 mg ml^{-1} lysozyme, Complete EDTA-free (Roche)] and stored at 193 K for 30 min. The cells were thawed and lysed by sonication after the addition of 10 µg ml^{-1} DNAse and 20 m*M* $MgSO_4$. The lysate was clarified by centrifugation at 23 000*g* for 1 h at 277 K. The supernatant was loaded on a metal-affinity HisTrap HP column (GE Healthcare); recombinant KUNV-helicase, eluted at 250 m*M* imidazole, was subsequently loaded onto a HiLoad 16/60 Superdex 200 column. The protein was then eluted using 100 m*M* NaCl, 1 m*M* DTT, 10 m*M* Bicine pH 8.5 and concentrated to 4 mg ml^{-1} for crystallization tests using an Amicon Ultra centrifugal filter (10 kDa cutoff).

2.3. Dynamic light scattering

The purified KUNV-helicase was centrifuged at 13 000*g* for 10 min prior to DLS analysis; all measurements were carried out at 293 K using a DynaPro instrument (Protein Solutions, Charlottesville, USA). The DLS data showed that the protein was 13.9% polydisperse and thus suitable for crystallization, as described by Ferré-D'Amaré & Burley (1994) and Zulauf & D'Arcy (1992), displaying a hydrodynamic radius of 4.0 nm (Fig. 2). The extrapolated molecular weight of the protein is approximately twice that of a single KUNV-helicase chain.

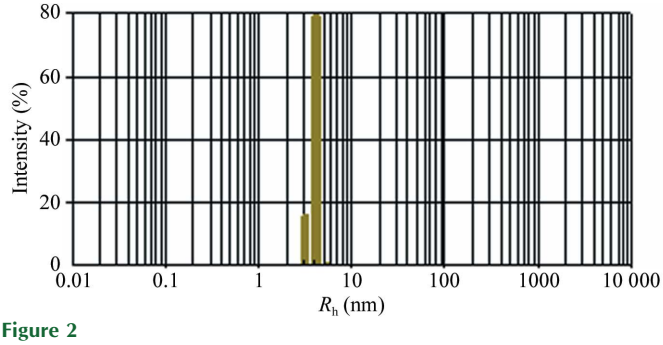

Figure 2
Representative DLS intensity particle-size distribution of KUNV-helicase (4 mg ml^{-1}) in 100 m*M* NaCl, 1 m*M* DTT, 10 m*M* Bicine pH 8.5.

Table 1
X-ray data-collection and refinement statistics for KUNV-helicase.

Values in parentheses are for the highest resolution shell.

Space group	*C*2
Unit-cell parameters (Å, °)	a = 172.1, b = 38.1, c = 148.1, β = 96.9
Resolution (Å)	40.00–3.00 (3.16–3.00)
Mosaicity (°)	0.6
No. of unique reflections	19765 (2865)
Completeness (%)	99.8 (99.8)
Redundancy	3.3
R_{sym}†	16.7 (74.6)
Average $I/\sigma(I)$	10.3 (2.8)
R factor‡ (%)	33.8
R_{free}§	42.6

† $R_{sym} = \sum_{\mathbf{h}} \sum_l |I_{\mathbf{h}l} - \langle I_{\mathbf{h}} \rangle| / \sum_{\mathbf{h}} \sum_l (I_{\mathbf{h}l})$, where I_l is the lth observation of reflection **h** and $\langle I_{\mathbf{h}} \rangle$ is the average intensity for all observations l of reflection **h**. ‡ R factor = $\sum |F_o - F_c| / \sum |F_o| \times 100$. § R_{free} is calculated from 5% randomly selected reflections for cross-validation.

2.4. Crystallization

Crystals of KUNV-helicase were grown in vapour-diffusion crystallization experiments assembled using an Oryx-6 crystallization robot (Douglas Instruments, East Garston, UK) from a 2:1 mixture of protein and reservoir solution. The 0.3 µl crystallization droplets were deposited in 96-well Corning Crystalex Round plates (Hampton Research); each well contained 100 µl of the selected precipitant solution. The screening solutions used for the experiments were those of Crystal Screens I and II and Index Screen from Hampton Research (Aliso Viejo, CA, USA) and of JBScreening Classic 2 and 4 (Jena Bioscience GmbH, Jena, Germany). All crystallization trials were performed at 293 K. Very small needle-shaped crystals were obtained after 1 d of vapour diffusion against 30%(*w*/*v*) PEG 4000 and 0.2 *M* sodium acetate, 0.1 *M* Tris–HCl pH 8.5. After optimization of the growth conditions [reservoir containing 10%(*w*/*v*) PEG 4000, 0.2 *M* sodium acetate, 0.1 *M* Tris–HCl pH 8.5, using the same experimental setup described above], KUNV-helicase crystals grew as thin elongated plates in 5–7 d to dimensions of up to approximately 200 × 30 × 10 µm.

3. Results and discussion

3.1. X-ray data collection and crystal characterization

For data collection, crystals were harvested into a cryoprotectant solution [20%(*w*/*v*) PEG 4000, 0.2 *M* sodium acetate, pH 8.5 and 20% glycerol] before being flash-cooled in liquid nitrogen. Diffraction intensities were recorded on beamline ID23-1 at the European Synchrotron Radiation Facility (ESRF-Grenoble, France). The native crystals diffracted to a maximum resolution of 3.0 Å. The diffraction data were processed with *MOSFLM* (Steller *et al.*, 1997) and intensities were merged using *SCALA* (Collaborative Computational Project, Number 4, 1994). Analysis of the diffracted intensities indicates that KUNV-helicase crystals belong to the monoclinic space group *C*2, with unit-cell parameters a = 172.1, b = 38.1, c = 148.1 Å, β = 96.9°. The calculated crystal-packing coefficient (V_M = 2.5 Å^3 Da^{-1}) indicates the likely presence of two KUNV-helicase molecules in the asymmetric unit, with a corresponding solvent

content of 49.7% (Matthews, 1968). X-ray data-collection statistics are summarized in Table 1.

3.2. Molecular-replacement search

The crystal structure of KUNV-helicase could be solved by the molecular-replacement method using the program *MOLREP* (Vagin & Teplyakov, 1997). The search model coordinates were those of the dengue virus helicase structure (PDB code 2bmf; the full protein model was used without residue mutations), locating two molecules of KUNV-helicase in the asymmetric unit in agreement with the DLS analysis results. The two independent molecules were subjected to rigid-body refinement and subsequently partially refined ($R_{\rm gen}$ = 33.8%, $R_{\rm free}$ = 42.6%, at 3 Å resolution) using *REFMAC*5 (Winn *et al.*, 2001).

References

Caruthers, J. M., Johnson, E. R. & McKay, D. B. (2002). *Proc. Natl Acad. Sci. USA*, **97**, 13080–13085.

Cho, H. S., Ha, N. C., Kang, L. W., Chung, K. M., Back, S. H., Jang, S. K. & Oh, B.-H. (1998). *J. Biol. Chem.* **273**, 15045–15052.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Ferré-D'Amaré, A. R. & Burley, S. K. (1994). *Structure*, **25**, 357–359.

Fields, B. N., Howley, P. M., Griffin, D. E., Lamb, R. A., Martin, M. A., Roizman, B., Straus, S. E. & Knipe, D. M. (2001). *Fields Virology*, 4th ed. Philadelphia, PA, USA: Lippicott Williams & Wilkins.

Ferron, F., Rancurel, C., Longhi, S., Cambillau, C., Henrissat, B. & Canard, B. (2005). *J. Gen. Virol.* **86**, 743–749.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Scherret, J. H., Poidinger, M., Mackenzie, J. S., Broom, A. K, Deubel, V., Lipkin, W. I., Briese, T., Gould, E. A. & Hall R. A. (2001). *Emerg. Infect. Dis.* **7**, 697–705.

Steller, I., Bolotovsky, R. & Rossmann, M. (1997). *J. Appl. Cryst.* **30**, 1036–1040.

Takegami, T., Sakamuro, D. & Furukawa, T. (1995). *Virus Genes*, **9**, 105–112.

Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.

Velankar, S. S., Soultanas, P., Dillingham, M. S., Subramanya, H. S. & Wigley, D. B. (1999). *Cell*, **97**, 75–84.

Walker, J. E., Saraste, M., Runswick, M. J. & Gay, N. J. (1982). *EMBO J.* **1**, 945–951.

Winn, M. D., Isupov, M. N. & Murshudov, G. N. (2001). *Acta Cryst.* D**57**, 122–133.

Wu, J., Bera, A. K., Kuhn, R. J. & Smith, J. L. (2005). *J. Virol.* **79**, 10268–10277.

Xu, T., Sampath, A., Chao, A., Wen, D., Nanao, M., Chene, P., Vasudevan, S. G. & Lescar, J. (2005). *J. Virol.* **79**, 10268–10277.

Zulauf, M. & D'Arcy, A. (1992). *J. Cryst. Growth*, **122**, 102–106.