Published in final edited form as:

Biomol NMR Assign. 2009 December ; 3(2): 219–222. doi:10.1007/s12104-009-9179-z.

¹H, ¹³C, ¹⁵N backbone and side chain NMR resonance assignments for E73 from *Sulfolobus* spindle-shaped virus ragged hills, a hyperthermophilic crenarchaeal virus from Yellowstone National Park

Casey Schlenker^{1,2}, Smita Menon^{1,2}, C. Martin Lawrence^{1,2}, and Valérie Copié^{1,2}

Valérie Copié: vcopie@chemistry.montana.edu

¹ Department of Chemistry and Biochemistry, Montana State University, 103 Chemistry and Biochemistry Building, Bozeman, MT 59717-3400, USA

² Thermal Biology Institute, Montana State University, 103 Chemistry and Biochemistry Building, Bozeman, MT 59717-3400, USA

Abstract

Crenarchaeal viruses are commonly found in hyperthermal acidic environments such as those of Yellowstone National Park. These remarkable viruses not only exhibit unusual morphologies, but also display extreme genetic diversity. However, little is known about crenarchaeal viral life cycles, virus–host interactions, and their adaptation to hyperthermophilic environments. In an effort to better understand the functions of crenarchaeal viruses and the proteins encoded by their genomes, we have undertaken detailed structural and functional studies of gene products encoded in the open reading frames of *Sulfolobus spindle-shaped* virus ragged hills. Herein, we report (¹⁵N, ¹³C, ¹H) resonance assignments of backbone and side chain atoms of a 19.1 kDa homodimeric E73 protein of SSVRH.

Keywords

Crenarchaea; Viruses; Hyperthermophiles; NMR resonance assignments; E73 protein; *Sulfolobus* spindle shaped viruses; Molecular adaptation

Biological context

Crenarchaeal viruses are commonly found in hyperthermal acidic environments such as those of Yellowstone National Park. These remarkable viruses not only exhibit unusual morphologies, but also show extreme genetic diversity, leading to the recent recognition of seven new viral families (Prangishvili et al. 2006; Lawrence et al. 2009). The best studied are the *Fuselloviridae*, which are ubiquitous to acidic hot springs (T > 70°C, 1 < pH < 4) worldwide, and include *Sulfolobus spindle-shaped* virus 1 (SSV1) and *Sulfolobus spindle-shaped* virus ragged hills (SSVRH) (Wiedenheft et al. 2004). These viruses are characterized by approximately 60 × 90 nm spindle- or lemon-shaped viral particles with tail fibers that emanate from one end. They contain double stranded circular DNA genomes that encode 34 and 37 open reading frames (ORFs), respectively (Wiedenheft et al. 2004), where the names of the ORF and corresponding protein are derived from the reading frame (A–F) and number of

encoded amino acids. Thus, ORF E73 in SSVRH encodes a 73 amino acid gene product in reading frame E.

Early analyses of the SSV1 genome revealed only two ORFs similar to proteins of known function. Consistent with its lysogenic cycle, D335 is a viral integrase (Clare and Stedman 2007), while B251 exhibits limited similarity to DnaA (Koonin 1992). Improved bioinformatic methods, however, have noted additional similarities: E51 and C80 are potential ribbon-helix-helix transcriptional regulators, C₂H₂ zinc-finger motifs are present in A45, A79 and B129, while B115 is annotated as a helix-turn-helix-type of transcriptional regulator (Prangishvili et al. 2006; Frols et al. 2007). However, 26 of 34 ORFs are not reliably identified by bioinformatics.

Greater insight into the viral gene products is clearly essential for a deeper comprehension of the SSV1 life cycle and interactions with its *Sulfolobus* host, and for a better understanding of the biology of crenarchaeal viruses in general. Towards this goal, molecular analyses of the SSV1 viral particle identify five proteins: VP1, VP2, VP3, C792 and D244 (Reiter et al. 1987; Menon et al. 2008). VP1, VP3 and C792 are membrane embedded structural proteins, while VP2 is a packaged DNA binding protein.

Importantly, structural studies are uncovering distant evolutionary relationships that are not apparent in the primary sequence of these proteins (Lawrence et al. 2009). Structural annotation now suggests that D-63 is an adaptor protein similar to Repressor of Primer, while F112 and F93 are putative transcription factors (Menon et al. 2008; Lawrence et al. 2009).

To assess the validity of the bioinformatics prediction that SSV1 E51 and its SSVRH homolog, E73, are ribbon-helix-helix type DNA binding proteins, and to gain deeper structural and functional insights, we have undertaken biochemical characterization and multidimensional solution NMR studies of SSVRH E73. Here, we report the ¹⁵N, ¹³C, and ¹H NMR resonance assignments for the backbone and side chain atoms of recombinant E73 and describe its secondary structural content.

Methods and experiments

Cloning

SSVRH ORF E73 was amplified using a nested polymerase chain reaction (PCR). SSVRH DNA was purified as previously described (Menon et al. 2008) and used as template. The PCR primers introduced a Shine–Dalgarno sequence, an N-terminal 6X His tag and attB sites to facilitate ligase free subcloning using site-specific recombination (Gateway-Invitrogen). The internal forward and reverse primers were: 5'-

ATGCATCACCATCACCATCACTTGGTTGAAAGTAAAAAGAT-3' and 5'-GTACAAGAAAGCTGGGTCCTATTTTGCGGGAAAAAACCGCT-3', while the external forward and reverse primers were 5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTTCGAAGGAGATAGAACCATGCATCACCATCACCATCAC-3' and 5'-

GGGGACCACTTTGTACAAGAAAGCTGGGTCCTA-3', respectively. The PCR product was inserted into pDONR201 via a site-specific recombination reaction. The sequence of the resulting entry clone was verified using ABI BigDye Terminator Cycle sequencing. The entry clone was then transferred into pDEST14 (Invitrogen) with a second site-specific recombination reaction, yielding the expression vector pEXP14-E73.

As a result of the cloning, the recombinant E73 protein that we have investigated contains a non-native Met-(His)₆ N-terminal extension linked to the native Met1-Val2-Glu3-sequence of

E73-SSVRH. Residue numbering in Fig. 1 and the BMRB database starts with Met1 of the native sequence.

Protein expression

For E73 protein expression, BL21(DE3)-RIL *Escherichia coli* cells (Stratagene) were transformed with pEXP14-E73. A single colony was used to inoculate 5 ml of MDG minimal non-inducing media + 100 µg/ml ampicillin, and cells were grown for 16 h. For isotopic labelling, 275 µl of the overnight culture was added to 275 ml of N-5052 media for ¹⁵N labelling of E73, or C-750501 media for ¹⁵N/¹³C protein labelling, and the resulting cell cultures were grown for 22 h (Studier 2005). Cells were subsequently harvested by centrifugation at 3,000×g (Sorvall RC2-B) for 20 min and the pellets were stored at -20°C.

Protein purification

Cell pellets were thawed and resuspended at 5 ml/gram of cell pellet in lysis buffer (10 mM Tris, pH 8.0, 5 mM imidazole, 300 mM NaCl) with freshly prepared phenylmethylsulfonyl fluoride (0.1 mM PMSF). The cell suspension was vortexed vigorously to resuspend the cells, which were then lysed using an M-110L Microfluidizer (Microfluidics, Newton, MA). The resulting cell lysate was then incubated at 65°C for 20 min to denature *E. coli* proteins, and clarified by centrifugation at 22,000×g for 20 min. The supernatant was applied to a column containing a 3–5 ml bed volume of HIS-Select™ Nickel Affinity Gel (Sigma-Aldrich). The column was washed with 10 column volumes of wash buffer (10 mM Tris, pH 8.0, 10 mM imidazole, 300 mM NaCl) and E73 was eluted in 10 mM Tris, pH 8.0, 200 mM imidazole, 50 mM NaCl. The E73 protein fraction was then applied to a calibrated Superdex S-75 (Amersham-Pharmacia) column equilibrated with 10 mM Tris (pH 8.0) and 50 mM NaCl. Protein concentrations were determined as previously described (Menon et al. 2008). Following dialysis in NMR buffer (50 mM potassium phosphate, 1 mM EDTA, 0.1 mM PMSF, 0.01% sodium azide, at pH 5.0), protein concentrations were adjusted to 1 mM dimeric E73. Samples were prepared in 95% H₂O/5% D₂O or in 100% D₂O (for acquisition of the 3D ¹³C-edited ¹H-¹H TOCSY and NOESY NMR spectra).

Gel filtration chromatography, dynamic light scattering, and ultracentrifugation experiments were conducted to determine the oligomeric state of SSVRH E73. While the E73 construct used in these studies encodes a polypeptide with a calculated mass of 9,578 Da, the elution profile on the gel filtration column was consistent with a mass of ~20,000 Da, suggesting that E73 is present as a homodimer. Similarly, analytical ultracentrifugation indicated a mass of ~19,000 Da, with no apparent equilibrium between monomer and dimer, while dynamic light scattering indicated a particle size of 3.8 nm, again consistent with a homodimer.

The homodimeric state of E73 in solution was supported by estimates of amide ¹H resonance linewidths of ~16 Hz in 2D ¹H-¹⁵N HSQC correlation spectra consistent with dimer formation (we would expect to observe amide ¹H NMR resonance linewidths of ~8 Hz if the protein was monomeric in solution). We concluded that E73 is a symmetric homodimer (as opposed to being an asymmetric or heterodimer) due to the observations of only one set of NH resonances in 2D ¹H-¹⁵N HSQC correlation NMR spectra corresponding to a single polypeptide chain.

All NMR spectra were acquired at 312 K on a four-channel Bruker DRX-600 spectrometer, with an inverse triple (¹⁵N, ¹³C, ¹H) resonance probe equipped with triple axis gradients, as done for other proteins of interest (Taubner et al. 2006). All data were processed and analyzed using NMRPipe, PIPP, and Sparky software packages. Sequential ¹H/¹³C backbone and side-chain chemical shift assignments were extracted from standard heteronuclear NMR experiments (HNCA, HNCACB, CBCA(CO)NH, C(CO)NH, HBHA(CO)NH, HC(CO)NH,

and ^{13}C -edited ^1H - ^1H TOCSY and NOESY). ^1H , ^{15}N , and ^{13}C chemical shift dimensions were indirectly referenced to DSS.

Extent of assignments and data deposition

Excluding E73's three proline residues, only the amide resonance of Met 1 remains unassigned. $\text{Ca}/\text{C}\beta$ chemical shifts were assigned for all residues except the $\text{C}\beta$ of Ser 66. Near complete assignment of ^{13}C and ^1H resonances for the side chains were made and ~97% of the $\text{Ha}/\text{H}\beta$ resonances were also assigned. Backbone and side chain (^1H , ^{15}N , ^{13}C) resonance assignments have been deposited in the BioMagResBank in Madison, (WI, USA) under the accession number 16,177.

The two-dimensional ^1H - ^{15}N -correlation HSQC spectrum of SSVRH E73 is shown in Fig. 1. The dispersion of chemical shifts in the spectrum and the linewidths of the NMR resonances are consistent with E73 SSVRH being primary α -helical in structure and homodimeric in solution, respectively. CSI analysis points (Fig. 2) to the presence of three α -helices (in contrast to the two predicted by bioinformatics) and one extended β -strand in the E73 protomer, with helices 1, 2, and 3, spanning residues 17–32, 34–52, and 61–69, respectively, while β -strand 1 is composed of residues 10–16, running antiparallel to the equivalent strand ($\beta 1'$) in the second subunit of the E73 homodimer. We established that the beta strands are antiparallel in the E73 dimer based on characteristics NOE and hydrogen bonding patterns obtained from 3D ^{15}N -edited ^1H - ^1H NOESY spectra and identification of backbone amides protected from $^1\text{H}/^2\text{H}$ exchange, respectively. These experiments are part of a repertoire of NMR experiments currently being conducted to establish the high-resolution three-dimensional structure of E73-SSVRH in solution.

Acknowledgments

This work was supported by the National Science Foundation (grant number MCB 0444056 to V.C. and MCB 0628732 to C.M.L.). The NMR experiments were recorded at Montana State University on a DRX600 Bruker solution NMR spectrometer, purchased in part with funds from the NIH shared instrumentation grant program (SIG grant # 1-S10RR13878-01), and the NSF-EPSCOR program for the State of Montana.

Abbreviations

NH_4Cl	Ammonium chloride
CSI	Chemical shift index
IPTG	Isopropylthiogalactoside
EDTA	Ethylenediamine tetraacetic acid
HSQC	Heteronuclear single quantum correlation
MDG	Non inducing media
NMR	Nuclear magnetic resonance
PMSF	Phenylmethylsulfonyl fluoride
ORF	Open reading frame

References

- Clore AJ, Stedman KM. The SSV1 viral integrase is not essential. *Virology* 2007;361:103–111.
- Frois S, Gordon PM, Panlilio MA, Schleper C, Sensen CW. Elucidating the transcription cycle of the UV-inducible hyperthermophilic archaeal virus SSV1 by DNA microarrays. *Virology* 2007;365:48–59.

- Koonin EV. Archaeobacterial virus SSV1 encodes a putative DnaA-like protein. *Nucl Acid Res* 1992;20:1143.
- Lawrence CM, Menon S, Eilers BJ, Bothner B, Khayat R, Douglas T, Young MJ. Structural and functional studies of archaeal viruses. *J Biol Chem* 2009;284(19):12599–12603. [PubMed: 19158076]
- Menon SK, Maaty WS, Corn GJ, Kwok SC, Eilers BJ, Kraft P, Gillitzer E, Young MJ, Bothner B, Lawrence CM. Cysteine usage in *Sulfolobus* spindle-shaped virus 1 and extension to hyperthermophilic viruses in general. *Viol* 2008;376:270–278.
- Prangishvili D, Vestergaard G, Haring M, Aramayo R, Basta T, Rachel R, Garrett RA. Structural and genomic properties of the hyperthermophilic archaeal virus ATV with an extracellular stage of the reproductive cycle. *J Mol Biol* 2006;359:270–278.
- Reiter WD, Palm P, Henschen A, Lottspeich F, Zillig W, Grampp B. Identification and characterization of the genes encoding three structural proteins of the *Sulfolobus* virus-like particle SSV1. *Mol Genet Genom* 1987;206:144–153.
- Studier WF. Protein production by auto-induction in high density shaking cultures. *Prot Exp Purif* 2005;41:207–234.
- Taubner L, McGuirl MA, Dooley DM, Copié V. Structural studies of Apo NosL, an accessory protein of the nitrous oxide reductase system: insights from structural homology with MerB, a mercury resistance protein. *Biochemistry* 2006;45:12240–12252. [PubMed: 17014077]
- Wiedenheft B, Stedman K, Roberto F, Willits D, Gleske AK, Zoeller L, Snyder J, Douglas T, Young M. Comparative genomic analysis of hyperthermophilic archaeal *Fuselloviridae* viruses. *J Virol* 2004;78:1954–1961. [PubMed: 14747560]

Fig. 1.

Two-dimensional ^1H - ^{15}N -correlation (HSQC) NMR spectrum of uniformly $^{15}\text{N}/^{13}\text{C}$ labeled homodimeric E73 SSVRH recorded at 600 MHz (^1H Larmor frequency) magnetic field strength, at 312 K. NMR resonances connected by a *thin dotted line* correspond to signals from side chain NH_2 of glutamine and asparagine residues. An NMR signal for the indole NH of W46 is observed at 11.2 ppm but is not shown in the figure

Fig. 2.

Chemical shift index plot derived from consensus chemical shift indexing (i.e., based on the analysis of $H\alpha$, $C\alpha$, and $C\beta$ chemical shift deviations from random coil values) delineating elements of secondary structure as a function of residue number. Alpha helices and the extended beta strand are depicted as *coils* and *arrow*, respectively, above the CSI versus residue plot