Published in final edited form as:

Biomol NMR Assign. 2015 October ; 9(2): 333–336. doi:10.1007/s12104-015-9604-4.

NMR assignments for the insertion domain of bacteriophage CUS-3 coat protein

Therese N. Tripler¹, Mark W. Maciejewski², Carolyn M. Teschke^{1,3,*}, and Andrei T. Alexandrescu^{1,*}

¹Department of Molecular and Cell Biology, University of Connecticut, 91 N. Eagleville Road, Storrs, CT 06269-3125, USA

²Department of Molecular, Microbial, and Structural Biology, University of Connecticut Health, 263 Farmington Ave., Farmington, CT 06030-3305, USA

³Department of Chemistry, University of Connecticut, 55 N. Eagleville Rd., Storrs, CT 06269-3060, USA

Abstract

CUS-3 is a P22-like tailed dsDNA bacteriophage that infects *Escherichia coli* serotype K1. The CUS-3 coat protein, which forms the icosahedral capsid, has a conserved HK97-fold but with a non-conserved accessory domain known as the insertion domain (I-domain). Sequence alignment of the coat proteins from CUS-3 and P22 shows higher sequence similarity for the I-domains (35%) than for the HK97-cores, suggesting the I-domains play important functional roles. The I-domain of the P22 coat protein, which has an NMR structure comprised of a six-stranded β barrel, has been shown to govern the assembly, stability and size of the resulting capsid particles. Here, we report the ¹H, ¹⁵N, and ¹³C assignments for the I-domain from the coat protein of bacteriophage CUS-3. The secondary structure and dynamics of the CUS-3 I-domain, predicted from the assigned NMR chemical shifts, agree with those of the P22 I-domain, suggesting the CUS-3 and P22 I-domains may have similar structures and functions in capsid assembly.

Keywords

bacteriophage; I-domain; CUS-3; viral assembly; procapsid

Biological context

CUS-3 bacteriophage is a dsDNA virus, which infects *Escherichia coli* serotype K1, and belongs to the group of P22-like phages (Casjens and Thuman-Commike 2011). Recent work has classified 151 P22-like bacteriophages into 3 sub-groups based on coat protein

*To whom correspondence should be addressed: Carolyn M. Teschke, teschke@uconn.edu, Tel:860-486-3992, fax: 860-486-4331. Andrei T. Alexandrescu, andrei@uconn.edu, Tel: 860-486-4414, fax: 860-486-4331.

Ethical standards

All experiments complied with all laws of the United States of America

Conflict of Interest

The authors declare no conflict of interest

sequence identity: CUS-3-like, P22-like, and Sf6-like (Casjens and Thuman-Commike 2011; Parent et al. 2014). In addition to a core coat protein structure based on the HK97-like fold, all three phage sub-groups have an accessory domain within the coat protein called the inserted domain or “I-domain” (Parent et al. 2012; Parent et al. 2014; Rizzo et al. 2014). The I-domains of CUS-3 and P22 have 35% sequence identity compared to 28% identity for the entire coat proteins, as determined with the Clustal W program. This level of sequence conservation suggests that the I-domains are likely to have similar structures and functional roles in the CUS-3 and P22 phages.

We recently determined the NMR structure of the P22 I-domain and showed that the protein folds into a six-stranded β -barrel ((Rizzo et al. 2014); PDB 2M5S). The β -barrel has an asymmetric charge distribution, with a positively charged face that is used to dock the I-domain to a complementary negatively charged surface on the HK97-core structure. The flexible D-loop between strands 1 and 2 of the β -barrel forms ion pair interactions between adjacent coat proteins in the capsid. The flexible S-loop abuts the A-domain of the HK97-core and may modulate the curvature of the coat protein, which in turn could govern capsid size. Additional functions for the P22 I-domain include roles as a folding nucleus, stabilization of the coat protein, and serving as an interaction partner of the portal complex (Rizzo et al. 2014; Suhanovsky and Teschke 2013).

Cryo-electron imaging of CUS-3 bacteriophage shows the capsids assemble into T=7 icosahedral shells, similar to P22 phage (Parent et al. 2014). The I-domains are displayed on the surfaces of the capsids with similar orientations for P22 and CUS-3, but there is a more marked difference for the I-domain of Sf6 phage. To better understand how the HK97 fold has adapted to support folding and assembly of a variety of viruses we have undertaken comparative studies of the I-domains from P22-like dsDNA tailed phages. Here we report the NMR assignments of the I-domain from CUS-3, paving the way to determine the NMR structure and dynamics of this accessory module.

Methods and experiments

Protein expression and purification

The gene encoding the CUS-3 I-domain, comprising amino acids 223–337 of the full-length protein, was cloned into a *pET21a* plasmid that included a C-terminal His₆-tag for purification purposes (Novagen, Madison, WI). The construct was confirmed through DNA sequencing by Genewiz (South Plainfield, New Jersey). The recombinant vector was transformed into BL21 *Escherichia coli* cells for expression. Cells were grown at 37 °C in M9 medium supplemented with 3 g/L of ¹³C-D-glucose and/or 1 g/L of ¹⁵NH₄Cl. Expression was induced for 16 hours at 30 °C by adding IPTG to a final concentration of 1 mM. Cells were harvested at 3300 x g by centrifugation in a Sorvall SH-3000 swinging bucket rotor. The cell pellet was re-suspended in 20 mM sodium phosphate buffer, pH 7.6, with the following additions: 1:100 dilution of protease inhibitor cocktail (Sigma catalog number P8849), 0.1 % Triton, 200 µg/mL lysozyme, 5.8 mM MgSO₄, 0.58 mM CaCl₂, 115 µg/mL deoxyribonuclease and ribonuclease. Cells were lysed with a Misonix sonicator set to an amplitude of 37, a total process time of 3 minutes, and a pulse time of 15 seconds with 30 sec between pulses. Cell debris was removed by centrifugation at 38,621 x g in a Thermo

Scientific F18 12x50 rotor. Membranes were removed by ultracentrifugation at 162,635 x g in a Sorvall T-865 rotor. The supernatant was loaded onto a TALON metal affinity column (Clontech Laboratories, Mountain View, CA) that was charged with 50 mM CoCl₂ for purification. Fractions containing the I-domain, based on 15 % SDS-PAGE gels, were pooled and the protein was precipitated with 60% ammonium sulfate. The precipitated protein was harvested by centrifugation at 38,621 x g in a F18 12X50 rotor. The protein pellet was resuspended in 20 mM sodium phosphate buffer, pH 6.0, and dialyzed 3 times against 2 L of the same buffer. For NMR samples, the protein was concentrated using a 3,000 molecular weight cutoff Amicon Ultra filter unit (Millipore, Billerica, MA) to a final concentration of 1.9 mM.

The CUS-3 I-domain construct contains the extraneous amino acids MAS at the N-terminus and LEHHHHHH at the C-terminus due to the cloning vector. With the exception of AS at the N-terminus and L at the C-terminus, the additional residues were not observed in 2D ¹H-¹⁵N HSQC experiments due to fast solvent exchange, and are therefore likely to be disordered. Our numbering scheme, running from residues 223–337, references only the amino acids native to the CUS-3 protein.

NMR spectroscopy

NMR experiments were collected on VARIAN 800 and 600 MHz NMR instruments at 25 °C using pulse sequences from the Varian Protein-Pack. Samples for NMR contained 1.9 mM CUS-3 protein in 20 mM sodium phosphate, buffered to pH 6.0. Backbone chemical shift assignments were made using ¹H-¹⁵N HSQC, HNCACB, HN(CO)CA, HN(CA)CO, and HNCB spectra (Cavanagh 2006). Aliphatic side-chain assignments were made using ¹H-¹⁵N HSQC, ¹⁵N NOESY-HSQC, ¹⁵N TOCSY-HSQC, HBHA(CO)NH, HCCH-TOCSY, and CCH-TOCSY experiments (Cavanagh 2006). Aromatic side-chain assignments were from 2D NOESY, DQF-COSY, and TOCSY spectra run on a sample in D₂O. Amide side-chain assignments were made using ¹⁵N-HSQC and ¹⁵N NOESY-HSQC data. Data were processed using the FELIX-NMR program (San Diego, CA) and analyzed using CcpNMR (Vranken et al. 2005).

Assignments and data deposition

Figure 1 shows the assigned ¹H-¹⁵N HSQC spectrum of the CUS-3 I-domain. Backbone assignments are more than 98% complete. Side-chain assignments were obtained for 91% of carbons, 98% of hydrogens, and 82% of nitrogens. The chemical shifts have been deposited in the Biological Magnetic Resonance Bank (<http://www.bmrb.wisc.edu/>) with the accession number 25263.

Figure 2a compares the secondary structure of the CUS-3 I-domain predicted from the amino acid sequence using the Jpred server (Cole et al. 2008) to that calculated from the assigned chemical shifts with the program TALOS+ (Shen et al. 2009). The predicted β-strands have been numbered according to the NMR structure of the P22 I-domain (Fig. 2b). Black dots show amide protons that survive exchange in D₂O for 2 h, and are likely to be involved in hydrogen-bonded secondary structure. The solvent protection data generally agree with the secondary structure predicted by the TALOS+ program (Shen et al. 2009),

except for the region A320-K325 that shows protection but no clear secondary structure preferences. The S^2 order parameters predicted by the TALOS+ program (Shen et al. 2009) from the assigned chemical shifts, suggest two dynamic regions corresponding to segments V238-N250 and D282-Q286.

Figure 2b summarizes the secondary structure and dynamics of the P22 I-domain (Rizzo et al. 2014) for comparison with the homologous domain from CUS-3. The I-domains have similar secondary structure elements and dynamics. We therefore anticipate that the CUS-3 I-domain will share the 6-stranded β -barrel fold of the P22 domain, and functionally important dynamic segments analogous to the D- and S-loops in the P22 I-domain structure (Rizzo et al. 2014). Possible differences, at the current level of analysis, are segregated to the C-terminus of CUS-3, where strand β iii of the small β -sheet outside of the β -barrel may be missing. Strand β 6 of the β -barrel appears to be shorter in CUS-3, and gives small predicted S^2 order parameters suggestive of flexibility. More detailed information on similarities and differences in how the I-domains interact with the HK97 core segments of the respective coat proteins, as well as functional analyses, will require the NMR structure determination of the CUS-3 I-domain, which is currently in progress.

Acknowledgments

We wish to thank Dr. K. Parent for supplying CUS-3 bacteriophage and Drs. A.D. Shuyker, M.R. Gryk, and J.C. Hoch for the use of the software platform NMRbox, <http://nmrbox.org/>, version 0.4a. This work was supported by NIH grant R01 GM076661 to C.M.T. and a supplement to C.M.T. and A.T.A.

References

- Casjens SR, Thuman-Commike PA. Evolution of mosaically related tailed bacteriophage genomes seen through the lens of phage P22 virion assembly. *Virology*. 2011; 411:393–415.10.1016/j.virol.2010.12.046 [PubMed: 21310457]
- Cavanagh, J.; Fairbrother Wayne, J.; Palmer; Arthur, G.; Mark, Rance; Skelton Nicholas, J. Protein NMR Spectroscopy. 2. Elsevier Inc; Amsterdam: 2006. Protein NMR Spectroscopy Principles and Practice.
- Cole C, Barber JD, Barton GJ. The Jpred 3 secondary structure prediction server. *Nucleic Acids Res*. 2008; 36:W197–201. [PubMed: 18463136]
- Parent KN, Gilcrease EB, Casjens SR, Baker TS. Structural evolution of the P22-like phages: comparison of Sf6 and P22 procapsid and virion architectures. *Virology*. 2012; 427:177–188.10.1016/j.virol.2012.01.040 [PubMed: 22386055]
- Parent KN, et al. Three-dimensional reconstructions of the bacteriophage CUS-3 virion reveal a conserved coat protein I-domain but a distinct tailspike receptor-binding domain. *Virology*. 2014; 464–465:55–66.10.1016/j.virol.2014.06.017
- Rizzo AA, et al. Multiple Functional Roles of the Accessory I-Domain of Bacteriophage P22 Coat Protein Revealed by NMR Structure and CryoEM Modeling Structure. 2014; 22:830–841.10.1016/j.str.2014.04.003
- Shen Y, Delaglio F, Cornilescu G, Bax A. TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR*. 2009; 44:213–223.10.1007/s10858-009-9333-z [PubMed: 19548092]
- Suhanovsky MM, Teschke CM. An Intramolecular Chaperone Inserted in Bacteriophage P22 Coat Protein Mediates Its Chaperonin-independent Folding. *The Journal of biological chemistry*. 2013; 288:33772–33783.10.1074/jbc.M113.515312 [PubMed: 24126914]

Fig. 1. ^1H - ^{15}N HSQC showing NMR assignments for the CUS-3 coat protein I-domain. The spectrum was recorded on a 1.9 mM sample, at 25° C, pH 6.0. Backbone signals are labeled in blue. Sidechain resonances are indicated in magenta, with dashed lines and the superscript 'SC'. Residues are numbered according to the full-length CUS-3 coat protein. Amino acids, L338 and A221, enclosed in parenthesis are cloning artifacts that are not part of the native protein.

Fig. 2. Comparison of secondary structure and dynamics in the **(a)** CUS-3 and **(b)** P22 I-domains. **(a)** CUS-3 I-domain. From top to bottom: amino acid sequence, residues 223–337; backbone amide protons protected from hydrogen exchange for 2 h at 25° C, pH 6.0 (black circles); JPred secondary structure prediction; TALOS+ secondary structure prediction with the β -strands in the β -barrel colored magenta and those in the accessory β -sheet colored orange; TALOS+ S^2 order parameter prediction, with flexible regions having S^2 values lower than 0.75 indicated with green “x” symbols. **(b)** Secondary structure and dynamics of the P22 I-domain (Rizzo et al. 2014). Segments of secondary structure and S^2 order parameters determined from the NMR structure and ^{15}N relaxation measurements, are color-coded as above.