doi:10.1016/j.jmb.2006.12.046 J. Mol. Biol. (2007) 366, 1538–1544

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

The Coiled-coil Domain Structure of the Sin Nombre Virus Nucleocapsid Protein

Sergei P. Boudko, Richard J. Kuhn and Michael G. Rossmann*

Department of Biological Sciences, Purdue University 915 W. State Street, West Lafayette, IN 47907-2054, USA

Hantaviruses can cause hemorrhagic fever with a renal syndrome and hantavirus pulmonary syndrome when transmitted to humans. The nucleocapsid protein of hantaviruses encapsidates viral genomic RNA and associates with transcription and replication complexes. Both the amino and carboxy termini of the nucleocapsid protein had been predicted to form trimers prior to the formation of the ribonucleoprotein. Crystal structures of amino-terminal fragments of the nucleocapsid protein showed the formation of intramolecular antiparallel coiled coils, but not intermolecular trimers. Thus, the amino-terminal part of the nucleocapsid protein is probably insufficient to initiate the trimerization of the full-length molecule.

Keywords: crystal structure; Sin Nombre virus; hantavirus; antiparallel coiled coil; mini-fibritin

*Corresponding author

Introduction

The term hantavirus encompasses any of the more than 20 distinct agents within the genus *Hantavirus* in the family *Bunyaviridae*. About half of the known members of the genus *Hantavirus* are pathogenic in rodents and humans. Hantaviruses of the Old World can cause hemorrhagic fever, whereas hantaviruses of the New World can cause pulmonary syndromes (HPS).[1] The major causative agent of HPS in North America is the Sin Nombre virus (SNV) carried by the deer mouse, *Peromyscus maniculatus*.

All members of the viral family *Bunyaviridae* consist of enveloped spherical particles with a helical nucleocapsid and use a genome consisting of three negative-stranded or ambisense RNAs (Figure 1). The three negative-sense RNA segments of the hantaviruses are designated L (large; about 6500 nucleotides (nt)), M (middle; about 3600 to 3700 nt), and S (small; 1700 to 2100 nt). The proteins that they encode are an RNA-dependent RNA polymerase from the L segment (L proteins); a glycoprotein precursor that is processed into two transmembrane glycoproteins (Gn and Gc) from the M segment; and a nucleocapsid (N) protein from the S segment. Hantaviral particles have a lipid envelope formed by a membrane derived from the host cell and Gn-Gc heterodimers. Within the envelope, the three genomic segments and the nucleocapsid form three separate, filamentous, 2.5 nm wide ribonucleoproteins.[2]

Like nucleoproteins of many negative-strand RNA viruses, hantaviral N protein is a multifunctional molecule involved in various interactions during the life cycle of the virus. It has essential functions in viral RNA replication, encapsidation, and also in virus assembly.[3] Trimerization of the N protein[4–6] is a crucial step in these processes[3] and was found to play a role in the discrimination between viral and non-viral RNA.[7] The N protein of hantaviruses contains from 428 (as in SNV) to 433 amino acid residues and has a molecular mass of approximately 50 kDa. A comparison of the amino acid sequences of hantavirus N proteins shows that there are three conserved domains separated by two more variable regions between residues 50 to 80 and 230 to 310. Residues 175 and 217 in the central RNA binding domain of Hantaan virus nucleocapsid[8,9] have high affinity for viral RNAs.

Using two-hybrid analyses, the hantavirus homotypic interactions were mapped to the amino-terminal[4,10] and the carboxy-terminal[10,11] regions. A three-dimensional reconstruction from electron microscopy images of recombinant N protein showed that it assembled into trimers in which individual N proteins form a curved structure.[12] Structure prediction algorithms suggested that N

Present address: S.P. Boudko, Research Department, Shriners Hospital for Children, 3101 SW Sam Jackson Park Road, Portland, OR 97239, USA.

Abbreviations used: mf, minifibritin; N protein, nucleocapsid protein; PEG, polyethylene glycol; SeMet, selenomethionine; SNV, Sin Nombre virus.

E-mail address of the corresponding author: mr@purdue.edu

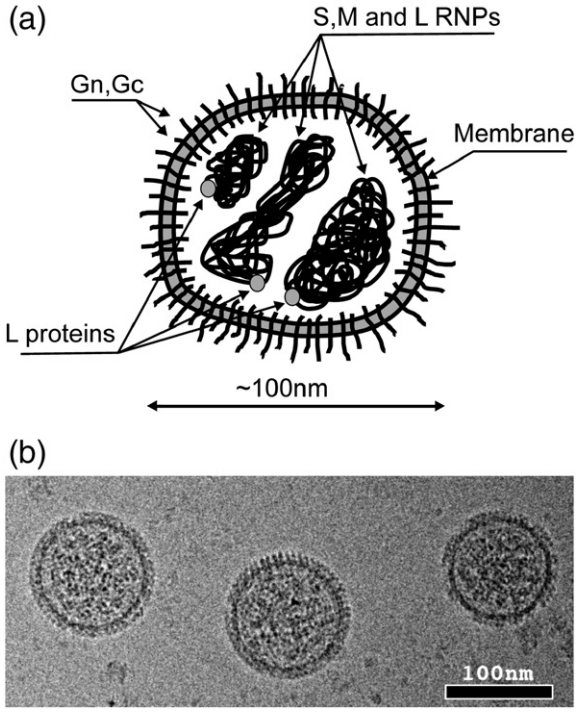

Figure 1. The *Bunyaviridae* family. (a) Schematic drawing of a viral particle.[2] The N protein binds all three segments of viral RNA and forms S, M, and L ribonucleoproteins. (b) Cryo-EM micrograph of vitrified Hantaan virus particles (courtesy of Paul R. Chipman and Dr Colleen B. Jonsson, unpublished results).

protein residues 3–75 form two coiled-coil segments separated by an intervening kink or turn sequence.[5] The properties of three chemically synthesized peptides covering residues 3–35, 43–75, and 3–75 have been examined. Peptide 3–35 assembled into trimeric coiled coils at high concentrations and low temperature, whereas peptide 43–75 trimerized efficiently at low concentration, implying that it carries a coiled-coil trigger sequence. However, the longer peptide, 3–75, assembled into dimers and/or trimers at high concentration, although at low concentration it appeared to adopt an intramolecular antiparallel coiled-coil configuration. Based on these results, a 3D structure prediction of an antiparallel coiled-coil domain spanning residues 1 to 77 of the N protein of Tula hantavirus was done very recently.[13]

Here, we report the first atomic resolution structure of a *Bunyavirus* protein. The structure shows that trimerization of the N protein is not initiated by its amino-terminal region.

Results

Design and production of recombinant proteins

A set of truncated SNV mutants, 1–75, 1–93, 1–152, 1–170, 1–300, and 1–393, and the 1–428 full-length N protein were designed for crystallization. All mutants were fused with a helper molecule (HT-mf-thromb), where HT is a His tag (MGHHHHH-HGSG), mf is a 120-residue mini-fibritin (mf) sequence,[14–16] and a linker with a thrombin cleavage site (GSSGSGLVPRGS) (Figure 2). The helper molecule, which aids trimerization, produces a high level of expression, and facilitates purification, was located at the amino termini of all mutants and could be removed after cleavage with thrombin. All fusion constructs, except 1–152, had a very high level of expression, ranging from 30 to 300 mg of purified protein from 1 l of culture media. Mutants 1–75, 1–93, and 1–170 were fully soluble, whereas 20 to 50% of the 1–300, 1–393, and 1–428 mutants were insoluble. All of these mutants were purified and subjected to thrombin cleavage. The fragments 1–300 and 1–393 precipitated during thrombin cleavage. However, the cleaved fragments 1–75, 1–93, and 1–170 were fully soluble and were separated from the helper molecule. The three purified mutants, as well as the helper molecule, were used in crystallization trials. The full-length protein was also successfully cleaved with thrombin, but it slowly aggregated and was not suitable for crystallization, much like the separately expressed protein (Said Addel-Fattah Ali Taha, personal communication). Diffraction quality crystals were obtained of the 1–75, 1–93, and HT-mf-thromb fragments. Although the crystal structure of the mf had been solved earlier to 2.0 Å resolution (PDB accession code 1OX3),[16] the resultant structure verified its trimeric character subsequent to the mf, notwithstanding the various modifications.

Crystal structures of 1–75 and 1–93

The orthorhombic crystals of the 1–75 fragment contained two molecules (chain A and chain B) per crystallographic asymmetric unit. The structure was determined by selenomethionine (SeMet) multiwavelength anomalous dispersion (MAD) phasing after determining the position of one of the two Se atoms per asymmetric unit. The monoclinic crystal structure of 1–93, which also had two molecules per asymmetric unit, was subsequently determined by molecular replacement. Both structures contained the same non-crystallographic dimer. The structure of the 1–75 monomer consists of a helix-turn-helix motif, in which two α-helices, linked by 35-DPD-37, form an antiparallel coiled coil (Figure 3). Most of the residues in the α-helix *a* and *d* positions that form much of the helix-to-helix contact region are hydrophobic (Figure 4). However, His17 and Arg50 at *d* positions have adopted conformations where the

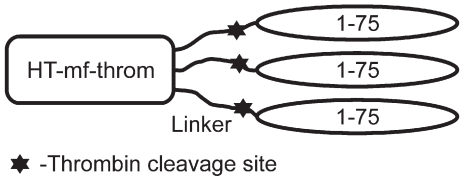

Figure 2. Fusion construct used for expression of truncated mutant 1–75 of the N protein.

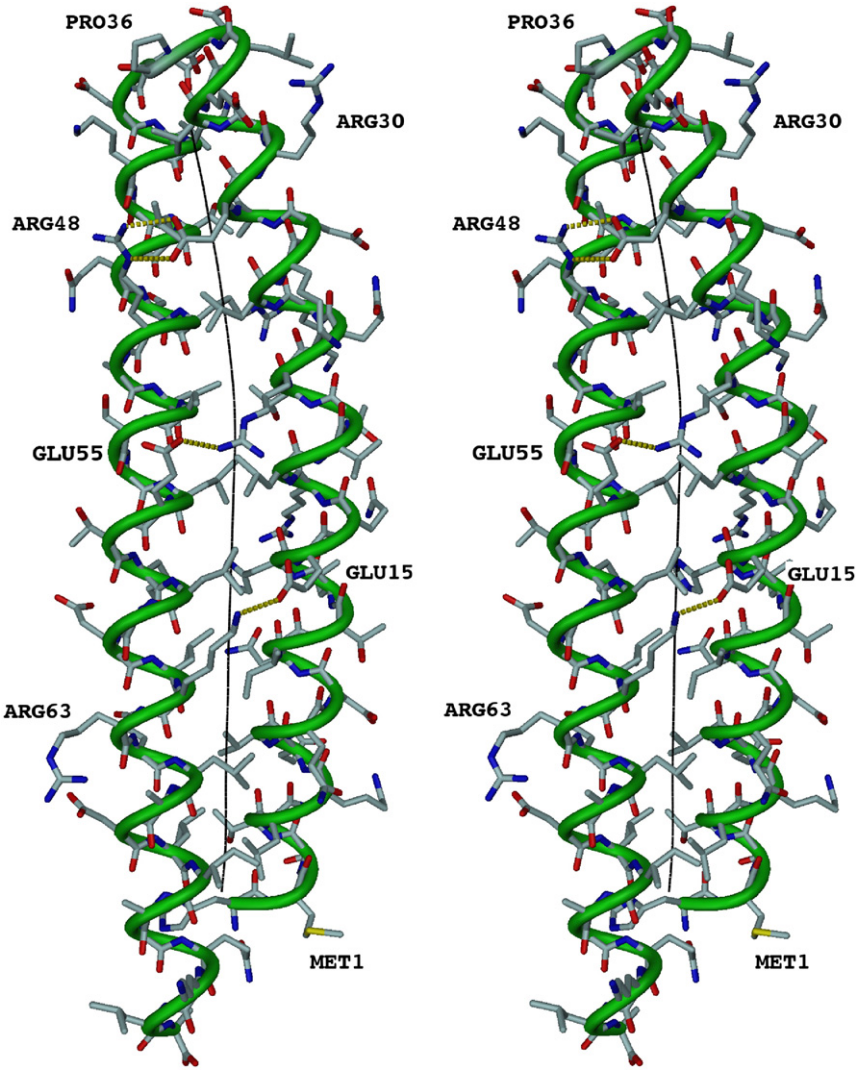

Figure 3. Stereo pair showing chain B in the crystallographic asymmetric unit of the 1–75 fragment. The α-helix backbones are shown as a green tube. A black line represents the axis of the coiled coil. Ionic bridges between two α-helices are shown as yellow broken lines. Several residues are labeled for clarity.

hydrophobic part of their side-chains is involved in hydrophobic contacts.

In addition to hydrophobic interactions, the coiled-coil structure is stabilized by ionic bridges between residues Glu15 and Lys62, Arg22 and Glu55, and Glu29 and Arg48 (the last one is not observed in chain A, fragment 1–75). All these residues are in *e* positions and all interactions are located on one side of the coiled coil (Figure 3). The geometry of the coiled coil is not ideal; the coiled-coil axis calculated as a median line between two axes of the α-helices is not a straight line (Figure 3), and the angle between two ends of the axis is about 17°. The radius of the coiled coil gradually decreases towards the turn from 5.0 Å to 4.4 Å, and the pitch decreases from 120 to 75 residues per super-helical turn.

There are two non-crystallographic 2-fold axes almost parallel to the crystal 2-fold screw axis along the crystal *b* direction. There are two types of dimers associated with these non-crystallographic 2-fold axes, creating "type I" and "type II" dimers. The monomer-to-monomer contact area in dimer I is about 7% and in dimer II about 10% of the area of an isolated monomer. Although the contact area for the type I dimer is smaller, the interactions stabilizing it include four pairs of ionic bridges: Lys24 to Asp35 and Asp37, Arg47 to Asp37 (two bonds). In the type II dimer, the interactions include ionic bridges between Glu55 and Lys41, Glu33 and Lys26 (only on one side of the non-crystallographic axis) and hydrogen bonds between Lys41 and Ser52. Dimers I and II form an infinite band of head-to-head packed molecules that lie in a plane roughly perpendicular to the crystal *b* direction.

Although the 1–93 fragment is 18 amino acid residues longer than the 1–75 fragment, the crystal structure did not show any additional residues. Well washed crystals on SDS-PAGE showed some limited proteolysis, indicating the presence of a shorter, about 75 amino acid residue long, fragment. About 80% of the material was still the whole 1–93 fragment. Thus, the additional 18 amino acid residues from 76 to 93 were disordered and prone to proteolysis. Possibly these residues would be well ordered if the rest of the molecule were present. The rms deviation between equivalent C^{α} positions when superimposing the different polypeptide chains of 1–75 and 1–93 are all less than 1.0 Å. These small differences can be attributed to residues involved in crystal packing. In the 1–93 crystals, there are type I dimers, but not type

```
    a b c d e f g

-2                  GSHMST  3
 4  L K E V Q D N           10
11  I T L H E⁻Q R           17
18  L V T T R⁺Q K           24
25  L K D A E⁻R A           31
32  V E L                   34
35                  DPD     37
38    D V N K S T           43
44  L Q S R R⁺A A           50
51  V S A L E⁻T K           57
58  L G E L K⁺R E           64
65  L A D L I A A           71
72  Q K L A                 75
```

Figure 4. Coiled-coil heptad repeat alignment of the 1–75 fragment amino acid sequence. Positions ***a*** and ***d*** are predominantly occupied by hydrophobic residues. Charges are shown for residues in *e* positions, which form interhelical ionic bridges. Numbering is as in the N protein. Residues −2 to 0 are the amino acids that remain after cleavage of the helper molecule.

II dimers, again demonstrating that the type I dimer is somewhat more stable. Thus, there are no multimeric infinite bands in the 1–93 crystal form. The type I dimers in the two crystal structures can be superimposed with an rms difference of 1.1 Å between equivalent C^{α} atoms.

The crystal structure of the HT-mf-thromb helper protein determined to 1.3 Å resolution was found to have an rms difference between equivalent C^{α} atoms of the mf component of 0.6 Å with respect to the previously reported structure[16] determined to 2.0 Å resolution. Furthermore, the protein forms a trimer as expected. Thus, this helper protein structure confirmed the correct folding of the mf component and its trimerization activity.

Discussion

The structures of the slightly different SNV N protein truncations 1–75 (crystals were grown at pH 5.4) and 1–93 (crystals were grown at 7.5) were essentially identical. The polypeptide consists of two α-helices (residues 1 to 34 and 38 to 75), which form an antiparallel coiled coil with linker residues being 35-DPD-37 (Figure 3). There is no evidence that these proteins have a trimeric association in the crystal structures. Dynamic light scattering shows that the average molecular diameter is 37.9 Å for the fusion proteins and 31.5 Å for the isolated helper protein (M_r of the trimer would be about 40 kDa). Since the helper molecule is a trimer, the molecular mass of a trimeric fusion molecule would be 40(37.9/31.5)3 or 70 kDa, as compared to the calculated value of 66 kDa for a trimer. This implies that the fusion proteins are also trimers. However, in the crystal structures, the isolated N protein fragments are found to be monomers or possibly dimers that are not associated by the anticipated coiled-coil structure. On the other hand, dynamic light scattering shows that the average molecular diameter of the isolated N protein fragments is about 57.6 Å, which suggests the formation of large aggregates, apparently inconsistent with earlier observations.[5] This inconsistency might be because of additional N-terminal residues or differences in buffer concentrations in the present results.

It is difficult to understand why the isolated proteins form aggregates, whereas the fusion proteins exist as trimers. There are two possible rationalizations of the above results (Figure 5). Possibly the fusion proteins are essentially trimeric coiled coils linked to the fibritin trimerization domain. In that case, the N protein fragments dissociate after thrombin cleavage, making individual molecules with antiparallel structures. Therefore, it is these antiparallel structures that make the large aggregates that would form after cleavage. Alternatively, the monomers within the trimeric fusion proteins are associated only in the fibritin section of the polypeptide, and the N protein section is already folded into an antiparallel coiled coil. However, in this case, it is more difficult to rationalize why the fusion protein trimers do not themselves aggregate. Possibly the amino-terminal region of the full-length N protein forms a trimeric parallel coiled coil, as suggested by the first hypothesis above, whereas the carboxy-terminal region of the molecule initiates trimerization.

Type I dimer formation found in crystal structures of the 1–75 and 1–93 fragments indicates stability of the dimer, since these two constructs were crystallized at two different pH ranges, acidic and neutral. It is possibly the dimer that was observed in the chemically synthesized peptide 3–75 by Alfadhli *et al.*[5] In the 1–75 crystal structure determined at acidic pH, which mimics the N protein environment, type I and type II dimers form an infinite band of head-to-head packed proteins. This might be a biologically relevant structure that would allow the full-length N protein to be packed into planar

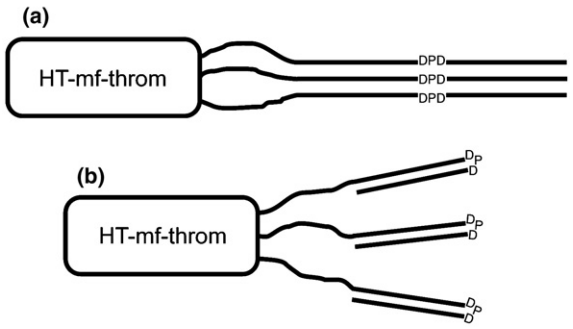

Figure 5. Two possible conformations of the 1–75 fragment as a part of the fusion molecule. (a) All three chains form a trimeric coiled coil with an insertion loop around residues 35-DPD-37. (b) Each chain independently forms an antiparallel coiled coil.

multimers and interact with viral RNA to form a filamentous ribonucleoprotein complex.

Materials and Methods

Cloning, protein expression and purification, crystallization, and particle size determination

The gene encoding HT-mf-thromb helper protein was recloned from the plasmid pHisMf[15] into pET23d(+) (Novagen) using restriction sites NcoI and BamHI. The new plasmid (pET23-HisMf) has multiple cloning sites just after the HT-mf-thromb gene.

The plasmid encoding the full-length N of Sin Nombre virus, strain CC107, was a gift from Dr Coleen B. Jonsson (Southern Research Institute, Drug Discovery and Homeland Security Division, and Department of Biochemistry and Molecular Biology, University of Alabama, Birmingham, AL). The genes encoding the various N protein constructs (Figure 1) were PCR amplified and cloned into the pET23-HisMf vector using restriction sites BamHI and SalI. The DNA inserts were verified by Sanger dideoxy DNA sequencing.

The recombinant proteins were expressed at 37 °C in the *Escherichia coli* BL21 (DE3) host strain (Novagen) after IPTG induction (final concentration 1 mM). The SeMet mutant of 1–75 was expressed in modified M9 medium in the presence of SeMet using the B834 (DE3) strain of *E. coli*. Purification of the fusion proteins by immobilized metal affinity chromatography on a HisTrap™ HP column (Amersham Biosciences) and separation of the N protein fragments after thrombin cleavage were carried out as described in the manufacturer's instructions. Before thrombin cleavage, the fusion proteins were additionally purified on an anion exchange HiTrap™ Q HP column (Amersham Biosciences). Thrombin cleavage was performed at 20 °C for 16 h with thrombin (Novagen).

The best crystals of the 1–75 fragment were obtained in solution no. 10 of Crystal Screen kit (Hampton Research), 0.2 M ammonium acetate, 0.1 M sodium acetate (pH 4.6), 30% (w/v) polyethylene glycol (PEG) 4000, whereas the crystallization conditions of the SeMet 1–75 were 0.01 M calcium chloride, 0.2 M ammonium sulfate, 0.1 M sodium acetate (pH 5.2), 20% (w/v) PEG 4000. Addition of calcium chloride significantly increased the reproducibility and quality of crystals, although no Ca or Cl ions were found in the crystal structure. The crystals grew to a size of 0.6 mm × 0.6 mm × 0.2 mm in hanging drops after about one to two weeks at 20 °C. The crystals were briefly dipped into a cryo-protectant solution containing the well solution and 1,2-propanediol at 5% (v/v) concentration and then flash frozen in a vaporized nitrogen stream at 100 K.

The best crystals of the 1–93 fragment were obtained in 0.15 M potassium bromide, 0.1 M Tris-HCl (pH 7.5), 34%

Table 1. Crystallographic data

Crystal type	Native 1–75	SeMet 1–75	Native 1–93	Native HT-mf-throm
Data collection[a]				
Wavelength (Å)	0.90020	0.97964, 0.97948, 0.95373	0.90020	1.00465
Detector	Quantum 315	MARmosaic 300	Quantum 315	MARmosaic 300
Beamline	Sector 14, BioCARS	Sector 23, GM/CA CAT	Sector 14, BioCARS	Sector 23, GM/CA CAT
Resolution range (Å)	10.00–1.15	25.00–1.33	30.00–2.00	35.00–1.32
Space group	$P2_12_12_1$	$P2_12_12_1$	$C2$	$P6_3$
Cell dimensions (Å)	$a=28.12$, $b=49.52$, $c=100.56$	$a=28.01$, $b=50.09$, $c=100.83$ $a=28.00$, $b=50.05$, $c=100.81$ $a=28.02$, $b=50.12$, $c=100.85$	$a=85.77$, $b=42.65$, $c=58.00$, $\beta=121.1°$	$a=b=53.68$, $c=106.46$
R_{merge} (%)	5.2 (23.8)	7.1 (22.3), 7.6 (20.4), 7.2 (23.6)	4.1 (21.2)	5.6 (21.6)
$I/\sigma I$	27.4 (6.6)	16.1 (8.2), 17.5 (6.1), 17.5 (6.3)	32.7 (7.5)	31.7 (5.2)
Completeness (%)	99.8 (99.4)	85.8 (52.2), 85.8 (51.5), 88.8 (62.2)	99.8 (99.5)	98.1 (91.0)
Multiplicity	5.7 (4.2)	7.4 (3.7), 7.3 (3.0), 7.4 (3.6)	4.0 (3.8)	6.6 (4.2)
Wilson plot B factor (Å^2)	9.0	11.7, 11.9, 11.5	36.1	15.7
Matthews coefficient (Å^3 Da^{-1})	2.0	2.0	2.1	3.2
Solvent content (%)	39	39	42	61
Phasing				
SOLVE Z-score/figure of merit		7.93/0.50		
Refinement				
Number of reflections used in refinement/for R_{free}	48,117/2572 (3496/159)		13,669/729 (862/37)	37,992/2001 (2591/129)
R factor/R_{free} (%)	15.4/17.7 (15.2/20.5)		24.7/32.8 (30.2/38.8)	16.0/17.0 (17.4/20.5)
Number of protein/water atoms	1184/245		1134/92	853/197
RMSD of bonds from idealized values (Å)	0.021		0.027	0.006

[a] Data in parentheses show the results in the highest resolution range.

(w/v) PEG mono-methyl ether (PEG MME) 2000. The crystals grew to a final size of 0.4 mm × 0.4 mm × 0.15 mm in hanging drops after about two to three weeks at 20 °C. The crystals were briefly dipped into a cryo-protectant solution containing the well solution and propylene glycol (1,2-propanediol) at 5% (v/v) concentration and then flash frozen in a vaporized nitrogen stream at 100 K.

The best crystals of HT-mf-thromb were obtained at solution no. 25 of a Wizard-I crystallization kit (Emerald Biosystems), 30% (v/v) PEG 400, 0.1 M TrisHCl (pH 10.5), 0.2 M magnesium chloride. The crystals were directly flash frozen in the vaporized nitrogen stream at 100 K.

The particle size of constructs was measured by a light-scattering method using photon correlation spectroscopy with the instrument Zetasizer 3000 (Malvern Instruments, Malvern, UK).

Structure solution and refinement

An SeMet crystal was used for a three-wavelength, multiwavelength anomalous dispersion data collection (Table 1) procedure.[18] The programs DENZO and SCALEPACK were used to index and integrate the diffraction data sets.[17] The program SOLVE[19] showed that there was only one Se atom per crystallographic asymmetric unit. This site was later found to correspond to Met1 of one chain, whereas Met1 of the other chain was in a disordered part of the structure. Phases were computed to 1.33 Å resolution by the program SOLVE based on the single Se atom per 156 residues. The phases were then improved by solvent flattening, and an atomic model (about 70% of all residues of both chains) was built with the program RESOLVE.[20] The remaining ordered residues were built manually using the program XtalView[21] after applying density averaging based on the non-crystallographic symmetry. The initial crystallographic refinement of the atomic model was accomplished with the program CNS[22] by using the data collected to 1.15 Å resolution from a native crystal. The program Refmac5[23] was used for the final refinement using geometrical restraints (Table 1). The ordered solvent structure contained 245 water molecules. Anisotropic temperature factors were used in the final stages of refinement. The number of observations exceeded the number of refinable parameters by about a factor of 5 in the final refinement.

The structures of 1–93 and the helper protein were solved by molecular replacement using the AmoRe program from the CCP4 suite.[23,24] Chain A of the 1–75 fragment was used as a search model to find each of the two chains in the 1–93 crystal structure. The structure refinement was performed as described above. Isotropic temperature factors were used in the refinement of the 1–93 fragment. Anisotropic temperature factors for protein atoms and isotropic temperature factors for water molecules were used in the final stages of the refinement of the helper molecule.

Figure preparation

The following programs were used in Figure preparation. Figure 3, DINO: Visualizing Structural Biology, 2002 (version 0.9.1†) and Figures 1, 2, and 5, INKSCAPE: Open source scalable vector graphics editor‡.

† http://www.dino3d.org
‡ http://www.inkscape.org

Protein Data Bank accession codes

The refined atomic models and the observed structure factors have been deposited with the RCSB Protein Data Bank , and are available under accession numbers 2IBL, 2IC6, and 2IC9.

Acknowledgements

We thank Cheryl Towell, Sharon Wilder, and Sheryl Kelly for their help in the preparation of the manuscript. We thank Dr Colleen B. Jonsson for providing the DNA plasmid encoding the N protein of SNV. We thank the staff of beam lines 14 (BioCARS) and 23 (GM/CA CAT) of the Advanced Photon Source for excellent support of our data collections. The work was supported by a National Institutes of Health grant (AI55672) to R.J.K. and M.G.R.

References

1. Schmaljohn, C. & Hjelle, B. (1997). Hantaviruses: a global disease problem. *Emerg. Infect. Dis.* **3**, 95–104.
2. Elliott, R.M., Bouloy, M., Calisher, C.H., Goldbach, R., Moyer, J.T., Nichol, S.T., *et al*. (2000). Family *Bunyaviridae*. In *Virus Taxonomy* (van Regenmortel, M.H.V., Fauquet, C.M., Bishop, D.H.L., Carsten, E.B., Estes, M.K. & Lemon, S.M. et al., eds), pp. 599–621, Academic Press, San Diego.
3. Kaukinen, P., Vaheri, A. & Plyusnin, A. (2005). Hantavirus nucleocapsid protein: a multifunctional molecule with both housekeeping and ambassadorial duties. *Arch. Virol.* **150**, 1693–1713.
4. Alfadhli, A., Love, Z., Arvidson, B., Seeds, J., Willey, J. & Barklis, E. (2001). Hantavirus nucleocapsid protein oligomerization. *J. Virol.* **75**, 2019–2023.
5. Alfadhli, A., Steel, E., Finlay, L., Bachinger, H. P. & Barklis, E. (2002). Hantavirus nucleocapsid protein coiled-coil domains. *J. Biol. Chem.* **277**, 27103–27108.
6. Kaukinen, P., Koistinen, V., Vapalahti, O., Vaheri, A. & Plyusnin, A. (2001). Interaction between molecules of hantavirus nucleocapsid protein. *J. Gen. Virol.* **82**, 1845–1853.
7. Mir, M. A. & Panganiban, A. T. (2004). Trimeric hantavirus nucleocapsid protein binds specifically to the viral RNA panhandle. *J. Virol.* **78**, 8281–8288.
8. Severson, W., Xu, X., Kuhn, M., Senutovitch, N., Thokala, M., Ferron, F. *et al.* (2005). Essential amino acids of the Hantaan virus N protein in its interaction with RNA. *J. Virol.* **79**, 10032–10039.
9. Xu, X., Severson, W., Villegas, N., Schmaljohn, C. S. & Jonsson, C. B. (2002). The RNA binding domain of the Hantaan virus N protein maps to a central, conserved region. *J. Virol.* **76**, 3301–3308.
10. Kaukinen, P., Vaheri, A. & Plyusnin, A. (2003). Mapping of the regions involved in homotypic interactions of Tula hantavirus N protein. *J. Virol.* **77**, 10910–10916.
11. Yoshimatsu, K., Lee, B. H., Araki, K., Morimatsu, M., Ogino, M., Ebihara, H. *et al.* (2003). The multimerization of hantavirus nucleocapsid protein depends on type-specific epitopes. *J. Virol.* **77**, 943–952.
12. Kaukinen, P., Kumar, V., Tulimäki, K., Engelhardt, P.,

Vaheri, A. & Plyusnin, A. (2004). Oligomerization of hantavirus N protein: C-terminal α-helices interact to form a shared hydrophobic space. *J. Virol.* **78**, 13669–13677.
13. Alminaite, A., Halttunen, V., Kumar, V., Vaheri, A., Holm, L. & Plyusnin, A. (2006). Oligomerization of hantavirus nucleocapsid protein: analysis of the N-terminal coiled-coil domain. *J. Virol.* **80**, 9073–9081.
14. Bachmann, A., Kiefhaber, T., Boudko, S., Engel, J. & Bachinger, H. P. (2005). Collagen triple-helix formation in all-*trans* chains proceeds by a nucleation/growth mechanism with a purely entropic barrier. *Proc. Natl Acad. Sci. USA*, **102**, 13897–13902.
15. Boudko, S. P. & Engel, J. (2004). Structure formation in the C terminus of type III collagen guides disulfide cross-linking. *J. Mol. Biol.* **335**, 1289–1297.
16. Boudko, S. P., Strelkov, S. V., Engel, J. & Stetefeld, J. (2004). Design and crystal structure of bacteriophage T4 mini-fibritin NCCF. *J. Mol. Biol.* **339**, 927–935.
17. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326.
18. Hendrickson, W. A. & Ogata, C. M. (1997). Phase determination from multiwavelength anomalous diffraction measurements. *Methods Enzymol.* **276**, 494–523.
19. Terwilliger, T. C. & Berendzen, J. (1999). Automated MAD and MIR structure solution. *Acta Crystallog. sect. D*, **55**, 849–861.
20. Terwilliger, T. C. (2002). Automated structure solution, density modification and model building. *Acta Crystallog. sect. D*, **58**, 1937–1940.
21. McRee, D. E. (1999). XtalView/Xfit- a versatile program for manipulating atomic coordinates and electron density. *J. Struct. Biol.* **125**, 156–165.
22. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W. *et al.* (1998). *Crystallography and NMR system*: a new software suite for macromolecular structure determination. *Acta Crystallog. sect. D*, **54**, 905–921.
23. Collaborative Computational Project Number 4 (1994). The *CCP4* suite: programs for protein crystallography. *Acta Crystallog. sect. D*, **50**, 760–763.
24. Navaza, J. (1994). AMoRe- an automated package for molecular replacement. *Acta Crystallog. sect. A*, **50**, 157–163.

Edited by A. Klug

(Received 3 September 2006; received in revised form 9 October 2006; accepted 15 December 2006)
Available online 23 December 2006