Journal of Biomolecular NMR (2005) 31: 351–356
DOI 10.1007/s10858-005-0945-7

NMR structure note

NMR solution structure of the monomeric form of the bacteriophage λ capsid stabilizing protein gpD

Hideo Iwai[a,b,c,*], Patrik Forrer[b], Andreas Plückthun[b] & Peter Güntert[a]
[a]*Tatsuo Miyazawa Memorial Program, RIKEN Genomic Sciences Center, Tsurumi, Yokohama, 230-0045, Japan;* [b]*Biochemisches Institut der Universität Zürich, CH-8057, Zürich, Switzerland;* [c]*Department of Chemistry, University of Saskatchewan, Saskatoon, SK, S7N 5C9, Canada*

Received 18 November 2004; Accepted 6 January 2005

Key words: bacteriophage λ, capsid, gpD protein, NMR structure

Abbreviations: RMSD – root-mean-square deviation; NOE – nuclear overhauser effect

Biological context

During virus assembly, viral precursor proteins form an empty capsid (procapsid) that undergoes large structural changes to become a mature capsid. In the case of *Escherichia coli* bacteriophage λ, an icosahedral virus of triangulation number $T=7$ with a non-contractile tail, the procapsid, is comprised mainly of the major capsid protein gpE. The expansion of the prohead occurs upon DNA packing. Without the subsequent binding of the gpD protein, the capsid cannot accommodate the full length DNA (Sternberg and Weisberg, 1977). Hence, gpD has been proposed to function in stabilizing the λ-head against the pressure imposed by packaged DNA. Thimble-shaped protrusions on the expanded head visible by cryo-electron microscopy were interpreted as trimers of gpD (Dokland and Murialdo, 1993). This was confirmed at higher resolution: the crystal structure of gpD is found to be a homo-trimer, apparently identical to the capsid-bound gpD trimers observed by cryo-electron microscopy of empty capsids at 15 Å resolution (Yang et al., 2000). However, the N-terminal region of gpD that is required for the binding to the capsid could not be observed in the crystals. In contrast to this trimeric capsid-bound form and to the trimeric form observed in the crystal structure, gel-filtration experiments suggested that gpD is a stable monomer in solution (Imber et al., 1980; Forrer et al., 2004). It was therefore of interest to examine this monomeric form in solution in order to determine whether there are any differences between the structures that might prevent its trimerization. As a first step, we have determined the NMR structure of the monomeric form of gpD in solution to help understand the structural basis of the trimerization during capsid assembly as well as the mechanism of capsid stabilization by gpD.

*To whom correspondence should be addressed. E-mail: h.iwai@usask.ca

Methods and results

The full length gpD protein and a truncated variant of gpD, gpDΔN1, of bacteriophage λ were expressed in *E. coli* and purified as published previously (Yang et al., 2000). Isotopically labeled proteins were prepared from cells grown in M9-based minimal medium supplemented with $^{15}NH_4Cl$ as the sole nitrogen source. The final NMR samples contained approximately 1.3 mM uniformly ^{15}N-labelled gpD dissolved in 90% H_2O/10% D_2O, or 100% D_2O, containing 20 mM sodium phosphate, pH 6.0. All NMR measurements were performed at 25° on Bruker DRX-600 or Bruker AV800 spectrometers. All pulse sequences

in H_2O incorporated the WATERGATE sequence using 3-9-19 composite pulses for the water suppression (Piotto et al., 1992). The resonance assignments of gpD have been reported previously (Iwai et al., 2004).

For structure determination, a three-dimensional ^{15}N-resolved [^{1}H, ^{1}H]-NOESY spectrum with a mixing time of 60 ms and a two-dimensional [^{1}H, ^{1}H]-NOESY spectrum in D_2O with a mixing time of 60 ms recorded on Bruker AV800 or DMX-600 spectrometers were used. A total of 2396 peaks from the two NOESY spectra and the previously reported resonance assignments of gpD (Iwai et al., 2004) were used as input for combined automated NOESY assignment and structure calculation with the program CYANA (Güntert et al., 1997; Herrmann et al., 2002; Güntert, 2003), which resulted in 1698 upper distance bounds for the final structure calculation. In addition, 78 dihedral angle constraints were used, which were derived from $^3J_{HN\alpha}$ coupling constants measured by using inverse Fourier transform fitting of in-phase multiplets (Szyperski et al., 1992). No hydrogen bond constraints were applied. No violations of NOE distance constraints larger than 0.2 Å were observed in the final 20 conformers. The atomic coordinates have been deposited in the Protein Data Bank with accession code 1VD0.

^{15}N relaxation times, T_1, T_2, and heteronuclear ^{15}N{^{1}H}-NOEs were measured and analyzed as described previously (Luginbühl et al., 1997). All experiments for the ^{15}N relaxation rates were performed at 25 °C, using a Bruker AV600 spectrometer operating at 600 MHz ^{1}H frequency. The pulse sequence used for measurements of $T_1(^{15}N)$ was adapted from Farrow et al. (1994) by replacing the H^N-selective 180° proton pulse in the ^{15}N relaxation delay by a 'hard' 180° pulse to minimize the influence of the ^{1}H radiofrequency carrier position on the suppression of the cross-correlation between dipolar relaxation and chemical shift anisotropy relaxation (Farrow et al., 1994). The T_1 relaxation delays at 600 MHz were 6, 16, 76, 117, 197, 298, 497, and 698 ms. Measurement of $T_2(^{15}N)$ was based on a CPMG-type sequence so that effectively a spin-locked transverse relaxation time was measured (Peng et al., 1991). The delay between the ^{15}N 180° pulses in the spin-lock train was 4 ms, which provided weakly spin-locked transverse ^{15}N magnetization. The T_2 relaxation delays were 8, 16, 33, 57, 82, 115, and 172 ms.

The solution structure of gpD was determined by combined use of automated NOE assignment (Herrmann et al., 2002) and torsion angle dynamics for the structure calculation (Güntert et al., 1997) (Table 1). Figure 2a provides the superposition of the final 20 conformers that represent the solution structure of gpD. The first 18 N-terminal residues are found to be disordered, whereas the rest of gpD is well-defined with a RMSD value relative to the mean coordinates of 0.56 Å for the backbone atoms of residues 19–109, which consists of 7 β-strands of residues 21–23, 35–39, 46–48, 57–61, 72–74, 78–81 and 106–109, and one short helix of residues 94–99 (Figure 2b).

From ^{15}N relaxation experiments, an average ratio of the longitudinal/transversal relaxation times of $<T_1/T_2> = 5.0 \pm 0.7$ at 600 MHz was obtained, excluding the flexible N-terminal residues 1–18 and residues 41 and 42, for which the relaxation data indicate chemical exchange processes. This value could be translated into an apparent global rotational correlation time $\tau_c = 6.3 \pm 0.6$ ns using the program DASHA with the assumption of a spherical top (Orekhov et al., 1995). This value corresponds to a molecular weight of approximately 12 kDa at 25 °C. Thus, the data obtained from the ^{15}N relaxation experiments strongly indicate that gpD exists as a

Table 1. Statistics for the NMR solution structure of the gpD protein

NOE distance constraints	
Short-range, $\|i-j\| \leq 1$	812
Medium-range, $1 < \|i-j\| < 5$	206
Long-range, $\|i-j\| \geq 5$	680
Total	1698
Maximal violation	0.09 Å
$^3J_{HN\alpha}$ scalar couplings	78
Maximal violation	0.16 Hz
Final CYANA target function value	0.07 Å^2
RMSD to mean coordinates	
Backbone N, C$^\alpha$, C′ of residues 19–109	0.56 Å
All heavy atoms of residues 19–109	0.91 Å
PROCHECK Ramachandran plot statistics	
Most favorable regions	76.8%
Additional allowed regions	23.1%
Generously allowed regions	0.1%
Disallowed regions	0.0%

Average values for the 20 final CYANA conformers.

```
          1                   20                  40
gpD        TSKETFTHY QPQGNSDPAH TATAPGGLSA KAPAMTPLML DTSSRKLVAW DGTTDGAAVG
gpDΔN1                 GSDPAH TATAPGGLSA KAPAMTPLML DTSSRKLVAW DGTTDGAAVG
SHP     VTKTITEQRAEV RIFAGNDPAH TATGSSGISS PTPALTPLML DEATGKLVVW DGQKAGSAVG

         60                  80                  100
gpD       ILAVAADQTS TTLTFYKSGT FRYEDVLWPE AASDETKKRT AFAGTAISIV
gpDΔN1    ILAVAADQTS TTLTFYKSGT FRYEDVLWPE AASDETKKRT AFAGTAISIV
SHP       ILVLPLEGTE TALTYYKSGT FATEAIHWPE SV.DEHKKAN AFAGSALSHAALP
```

Figure 1. Amino acid sequences of gpD, gpDΔN1, and SHP.

monomer in solution even at millimolar concentration. This observation is in line with previously reported gel-filtration experiments, analytical ultracentrifugation, dynamic light scattering (Imber et al., 1980; Yang et al., 2000; Forrer et al., 2004), as well as with the NMR solution structure of gpD.

In the crystal structure as well as on the capsid gpD is found as a homo-trimer with a large interaction surface of about 1250 $Å^2$ between monomers. The backbone RMSD between the mean coordinates of the X-ray structure and the NMR structure of gpD is 1.13 Å for residues 19–109, indicating an essentially identical structure for the globular domain in solution and in the crystal (Figure 2d). Differences between the two structures were observed for the ill-defined regions of the 18 N-terminal residues and the loop between residues 40–44 (Figures 2a and 3a). The first N-terminal 18 amino acids are largely disordered in the NMR structure, whereas in the crystal residues 1–14 were not observed and residues 15–17 have an extended conformation (Figure 2d). For residues 41 and 42 the ^{15}N relaxation times T_1/T_2 suggest the presence of conformational exchange, which is in line with the larger global displacement observed for this region (Figure 3b).

A truncated form of gpD lacking the flexible N-terminal 14 amino acids (gpDΔN1) had been found unable to accomplish the function of native gpD of stabilizing the capsid, which implicates the functional importance of the N-terminal region (Yang et al., 2000). However, almost all the peaks in the HSQC spectra of gpD and gpDΔN1 could be overlaid at the same chemical shift positions within 0.02 ppm for ^1H atoms and 0.2 ppm for ^{15}N atoms, except for the missing N-terminal residues and their immediate vicinity (data not shown). This indicates that the N-terminal region does not interact with the rest of the gpD structure and that the structures of the globular domains of gpD and gpDΔN1 are very likely to be identical. Negative and smaller values of the $^{15}N\{^1H\}$ heteronuclear NOEs are indicative of extensive internal motions in the N-terminal region, which is consistent with the disordered conformation in the NMR structure (Figure 3c).

Discussion and conclusions

The N-terminal tail, flexible in the solution structure, is required for binding to the capsid. Intriguingly, the homologous protein SHP from phage 21 can take over the function of gpD in the presence of Mg^{2+} (Yang et al., 2000; Wendt and Feiss, 2004), even though the sequence identity of the flexible N-terminal region is very low (Figure 1). One side of the trimer is much more conserved than the other between these two proteins, and cryo-electron microscopy suggests that it is this trimer surface that is interacting with the phage capsid. Nevertheless, the N-terminal region must be present on either gpD or SHP, even though its exact sequence does not seem to be critical for the capsid stabilization. One possible mode of interaction would be the insertion of the flexible N-terminal tail into holes or canyons in the capsid so that the gpD protein is anchored and the capsid is stabilized. This model is supported by the fact that a chimeric gpD/SHP protein, in which the N-terminal 19 amino acids of SHP are replaced with the N-terminal 14 amino acids of gpD, can stabilize λ capsid more effectively than SHP in the absence of Mg^{2+} (Wendt and Feiss, 2004). As proteins have been fused to the N-terminus of gpD and could be displayed on the λ capsid (Sternberg and Hoess, 1995; Mikawa et al., 1996), the N-terminus must find a way to the surface. The gpD protein binds to three different sites of

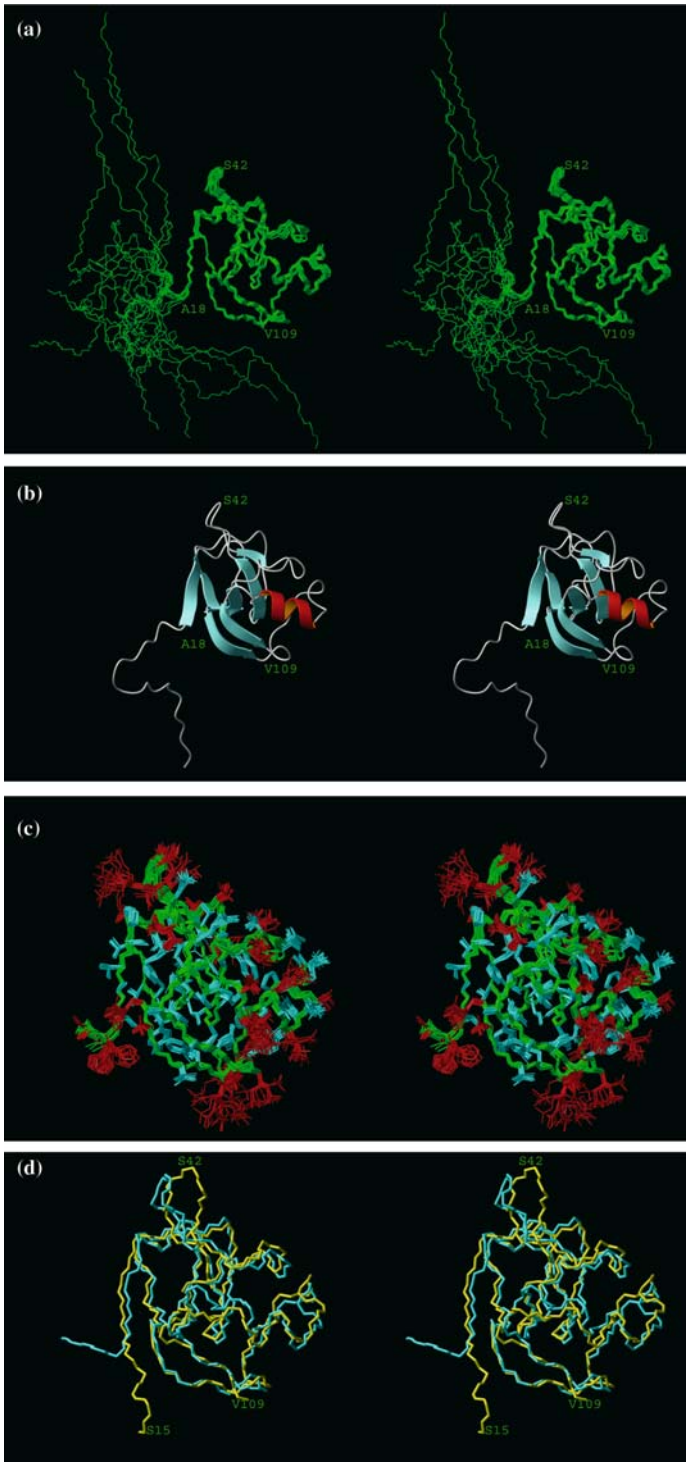

Figure 2. The NMR structure of gpD shown with stereoviews. (a) Superposition of the 20 conformers of the NMR solution structure of gpD fitted for minimal RMSD of the backbone atoms of residues 19–109. (b) Schematic representation of one of the final 20 NMR structures. (c) Same superposition of the 20 conformers of the NMR solution structure of gpD as in (a) but without residues 1–17 showing also the side-chain heavy atoms. The backbone, the 'best-defined' side-chains (i.e. side-chains for which global displacements are smaller than 0.8 Å), and the remaining side-chains are colored in green, blue, and red, respectively. (d) Superposition of the mean NMR structure (blue) and the mean X-ray structure (yellow). These figures were produced with the program MOLMOL (Koradi et al. 1996).

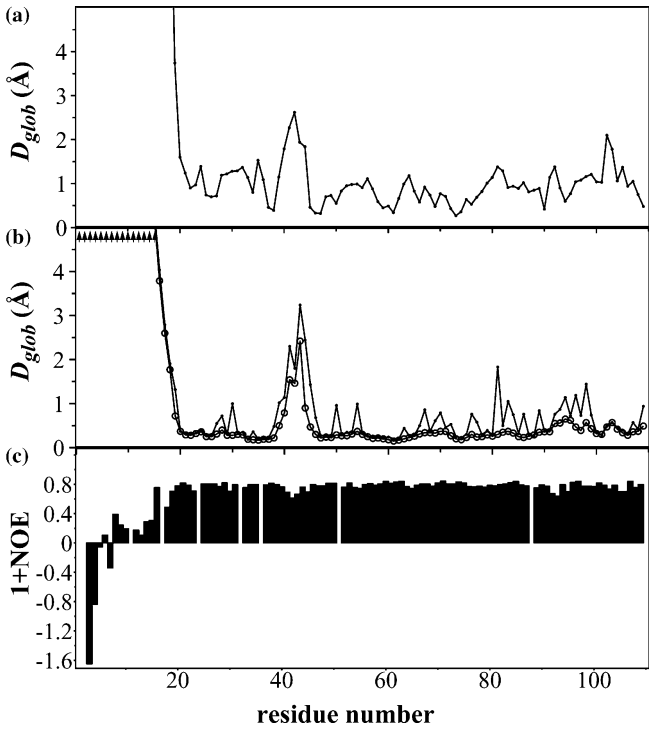

Figure 3. (a) Global displacements per residue of the backbone heavy atoms between the mean X-ray structure and the mean NMR structure after superposition of the backbone heavy atoms of residues 19–109 for minimal RMSD. (b) Mean of the pairwise global displacements per residue of the backbone heavy atoms (circles) and all heavy atoms (dots) of the 20 final NMR conformers relative to their mean coordinates calculated after superposition of the backbone heavy atoms of residues 19–109 for minimal RMSD. (c) Plot of the heteronuclear $^{15}N\{^{1}H\}$-NOEs *versus* the sequence.

threefold symmetry, one of them in the center of each icosahedral facet, the others at points of local threefold symmetry (Yang et al., 2000). The flexible N-terminal tail might thus be necessary to adapt to be able to fit to these slightly different locations on the λ capsid.

The gpD protein is a monomer in solution, and only trimerizes in the crystal and on the surface of the capsid. Our study has shown that there is no discernable change between the folded part of gpD in the monomeric and trimeric state. Therefore, we cannot invoke a conformational change with a high energetic barrier between the monomeric and trimeric state. Furthermore, there is no apparent evidence for an interaction between the flexible N-terminal region and the folded part of the protein. There is also no obvious difference between the structure of the gpDΔN1 variant lacking the flexible N-terminal region and wt gpD. Finally, the gpDΔN1 variant and wt gpD crystallize to a virtually identical trimer (Chang et al., 2004). Taken together, these results mean that the folded part of gpD and the flexible N-terminus behave as essentially independent entities.

Recently, the high resolution structure of SHP was determined (Forrer et al., 2004), which is a trimer in solution. It has a trimerization interface of virtually identical dimensions, but a much higher number of intersubunit H-bonds compared to trimeric gpD (27 *versus* 9 per trimer), which are tightly interlocked. This is probably the reason why this protein is kinetically stable against unfolding from the trimeric state. The N-terminus is ordered for a few more residues, showing that the N-terminus points away from the threefold axis to the side of the barrel. This makes it plausible how an N-terminally fused protein could become exposed on the surface of the phage and be used in λ-phage display.

As a monomer–trimer equilibrium is very sensitive to small changes in the interaction energy, at the concentration used in our assays (micro- to millimolar), gpD is essentially all monomeric, while SHP is all trimeric. Nevertheless, in their normal function, they are converted to the trimer

state upon association on the phage capsid. In order to fully understand the mechanism of λ capsid stabilization by binding of gpD or SHP, detailed information about the interaction between gpE and gpD upon the expansion of the procapsid still remains to be elucidated.

Acknowledgments

We thank the Irchel NMR Center at the University of Zürich and Dr. Oliver Zerbe for their support. This work was supported by the Tatsuo Miyazawa Memorial Program of RIKEN Genomic Sciences Center.

References

Chang, C. Plückthun, A. and Wlodawer, A. (2004) *Proteins*, **57**, 866–868.
Dokland, T. and Murialdo, H. (1993) *J. Mol. Biol.*, **233**, 682–694.
Farrow, N.A. Zhang, O. Forman-Kay, J.D. and Kay, L.E. (1994) *J. Biomol. NMR*, **4**, 727–734.
Forrer, P. Chang, C. Ott, D. Wlodawer, A. and Plückthun, A. (2004) *J. Mol. Biol.*, **344**, 179–193.
Güntert, P. (2003) *Prog. NMR Spectrosc.*, **43**, 105–125.
Güntert, P. Mumenthaler, C. and Wüthrich, K. (1997) *J. Mol. Biol.*, **273**, 283–298.
Herrmann, T. Güntert, P. and Wüthrich, K. (2002) *J. Mol. Biol.*, **319**, 209–227.
Imber, R. Tsugita, A. Wurtz, M. and Hohn, T. (1980) *J. Mol. Biol.*, **139**, 277–295.
Iwai, H. Forrer, P. Plückthun, A. and Güntert, P. (2004) *J. Biomol. NMR*, **28**, 89–90.
Koradi, R. Billeter, M. and Wüthrich, K. (1996) *J. Mol. Graph.*, **14**, 51–55.
Luginbühl, P. Pervushin, K.V. Iwai, H. and Wüthrich, K. (1997) *Biochemistry*, **36**, 7305–7312.
Mikawa, Y.G. Maruyama, I.N. and Brenner, S. (1996) *J. Mol. Biol.*, **262**, 21–30.
Orekhov, V.Y. Nolde, D.E. Golovanov, A.P. Korzhnev, D.M. and Arseniev, A.S. (1995) *Appl. Magn. Reson.*, **9**, 581 588.
Peng, J.W. Thanabal, V. and Wagner, G. (1991) *J. Magn. Reson.*, **94**, 82–100.
Piotto, M. Saudek, V. and Sklenar, V. (1992) *J. Biomol. NMR*, **2**, 661–665.
Sternberg, N. and Hoess, R.H. (1995) *Proc. Natl. Acad. Sci. USA*, **92**, 1609–1613.
Sternberg, N. and Weisberg, R. (1977) *J. Mol. Biol.*, **117**, 733–759.
Szyperski, T. Güntert, P. Otting, G. and Wüthrich, K. (1992) *J. Magn. Reson.*, **99**, 552–560.
Yang, F. Forrer, P. Dauter, Z. Conway, J.F. Cheng, N.Q. Cerritelli, M.E. Steven, A.C. Plückthun, A. and Wlodawer, A. (2000) *Nat. Struct. Biol.*, **7**, 230–237.
Wendt, J.L. and Feiss, M. (2004) *Virology*, **326**, 41–46.