PROTEINS: Structure, Function, and Bioinformatics 57:866–868 (2004)

STRUCTURE NOTE

Crystal Structure of a Truncated Version of the Phage λ Protein gpD

Changsoo Chang,[1] Andreas Plückthun,[2] and Alexander Wlodawer[1]*
[1]*Macromolecular Crystallography Laboratory, National Cancer Institute at Frederick, Frederick, Maryland*
[2]*Biochemisches Institut, Universität Zürich, Zürich, Switzerland*

Introduction. Bacteriophage λ contains a linear double-stranded DNA genome enclosed within an icosahedral capsid of triangulation number T = 7, and a flexible, non-contractile, tail.[1] It has been adapted for phage display, providing a valuable alternative to filamentous phages. As a lytic phage,[2] the phage coat is directly assembled in the cytoplasm, and does not pass through the secretion machinery and the oxidizing periplasm, thereby providing a display system for proteins not suitable for secretion. Both the head protein, gpD, and the tail protein, gpV, have been used as fusion partners in applications of phage display.[3–5] To further develop the technology of phage display, the structures of the proteins utilized in this technique are of considerable interest.

The high-resolution crystal structure of gpD, the capsid-stabilizing protein of bacteriophage λ, which sits as protrusions at trigonal sites, has been solved previously.[6] It provides a structural basis for the understanding of how gpD could play the role of a display platform. Molecules of gpD were found to be arranged in non-crystallographic trimers with substantial inter-subunit interfaces. Their C termini are well ordered and located on one side of the trimer, relatively far from its threefold axis. By contrast, the N termini are disordered up to Ser15, which is close to the threefold axis and on the same side as the C termini.

Even though these 14 N-terminal residues are disordered, they appear to have an important function, since an N-terminal truncated version of gpD (gpDΔN1) does not complement a phage mutant devoid of gpD.[6] Both gpD and gpDΔN1 are monomeric in solution, and gpD trimerizes on the phage and upon crystallization. It was, therefore, of interest to characterize gpDΔN1 more closely. While gpDΔN1 contains all visible atoms of the gpD trimer, it was conceivable that the N-terminal stretch might play an important role in trimer formation, which might occur concomitant with binding to the phage procoat. Furthermore, Pro17 and His19 form a characteristic ring in the trimer, and this might be compromised when the first 14 residues are absent. Finally, we considered that if crystals of gpDΔN1 isomorphous to those of the full-length protein could be obtained, comparison of the low-resolution diffraction of both forms might conceivably elucidate the location of the N terminus of the molecule, even if disordered, by applying an approach similar to the X-ray contrast variation used to outline the protein molecules.[7]

We thus solved the structure of gpDΔN1 at the resolution of 1.85 Å. Despite having crystallized in a larger cell with two trimers in the asymmetric unit, the structure of this truncated version of gpD is virtually identical to the structure of the unmodified protein.

Materials and Methods. The N-terminally truncated gpD (gpDΔN1, residues 15–111) was expressed and purified as described.[6] The protein was crystallized after refinement of the conditions originally reported for full-length gpD. The optimized crystallization conditions were 100 mM Tris-HCl pH 8.5, 28% PEG 4000, 200 mM Mg acetate, and 10% glycerol. Crystals shaped as thick rhombic plates were grown to a length of ~0.7 mm within 10 days. Since the crystallization mother liquor could serve directly as a cryoprotectant, crystals were directly mounted in the loops and flash-frozen in a nitrogen stream for data collection at 100 K.

The crystal of gpDΔN1 used for data collection belonged to the monoclinic space group $P2_1$ with unit cell parameters $a = 58.2$ Å, $b = 99.0$ Å, $c = 59.3$ Å, $\beta = 119.0°$ and with $V_M = 2.54$ Å^3/Da. Six molecules are present in the asymmetric unit. Data extending to 1.85 Å were collected at 100 K using the ADSC Quantum 4 CCD detector on beamline X9B, BNL, and were processed with HKL2000.[8] Data collection statistics are summarized in Table I. These crystals differ from the previously described crystals of full-length gpD.[6]

The complete trimer of gpD (PDB accession code 1c5e) was used as a search model for gpDΔN1. A fully automated script for AMoRe[9] was run with data in the resolution range of 12–3 Å. It resulted in a solution consisting of two

*Correspondence to: Macromolecular Crystallography Laboratory, National Cancer Institute at Frederick, Frederick, MD 21702. E-mail: wlodawer@ncifcrf.gov

Received 18 May 2004; Accepted 7 June 2004

Published online 12 August 2004 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.20254

TABLE I. Data Collection and Refinement Statistics

Wavelength (Å)	0.98
Space group	$P2_1$
Unit cell parameters (Å)	$a = 58.2$, $b = 99.0$, $c = 59.3$, $\beta = 119.0°$
Resolution (Å)	25–1.85
Number of reflections (unique/total)	47,862 (238,157)
Completeness (last shell)	95.2 (74.1)%
R_{merge}	2.7 (19.7)%
No. of molecules in a.u.	6
No. of protein atoms	4,146
No. of solvent molecules	685
R_{cryst}	19.1%
R_{free} (10%)	23.2%
r.m.s. deviations from ideality	
Bond lengths	0.008 Å
Angles	1.3°

trimers, with the correlation coefficient of 0.725 and the R-factor of 31.1%. The model for gpDΔN1 was refined using CNS 1.0[10] at the resolution range of 20–1.85 Å and rebuilt with O.[11] The non-crystallographic symmetry (NCS) restraints with an energy barrier of 300 kcal $mol^{-1}Å^{-2}$ between the two trimers were maintained during refinement. In addition to protein atoms, 685 water molecules have been added to the model. The final R for all reflections (20 to 1.85 Å) is 19.1 % (R_{free} 23.2%). The coordinates and structure factors have been submitted to the Protein Data Bank (accession code 1TCZ).

Results and Discussion. Although the resolution of data used in this work (1.85 Å) is considerably lower than for the full-length gpD (1.1 Å), the structure can still be considered to be high-resolution and the accuracy of the model of gpDΔN1 is acceptable, with the mean positional errors in atomic coordinates estimated by the Luzzati plot as 0.21 Å. The refinement statistics and the indicators of model quality are listed in Table I. All non-glycine and non-proline residues except for one in each molecule lie in either the most favorable or in the additionally allowed regions of the Ramachandran plot. The only residue to occupy the generously allowed region of the plot is Met34, and its non-standard torsion angles are fully supported by very clear electron density and agree with the previous observation from the very high-resolution structure of full-length gpD.

While the full-length gpD contains a single trimer in the asymmetric unit, the six molecules of gpDΔN1 present in the asymmetric unit are organized as two almost identical trimers. Individual superpositions of the first and second trimer of gpDΔN1 with the trimer of gpD resulted in r.m.s. deviations for all atoms of 0.58 and 0.59 Å, respectively. These values are similar to the r.m.s. deviations between individual molecules in the structure of gpDΔN1, indicating that truncation of the N-terminal residues has not resulted in any significant changes in the remaining part of the molecule. The similarity also extended to the solvent structure. After superposition of the gpD trimer on the individual trimers of gpDΔN1, the number of water molecules located within 1 Å of each other was 377 (out of 685 total). For comparison, superposition of the solvent surrounding each trimer of gpDΔN1 yields 220 common solvent sites.

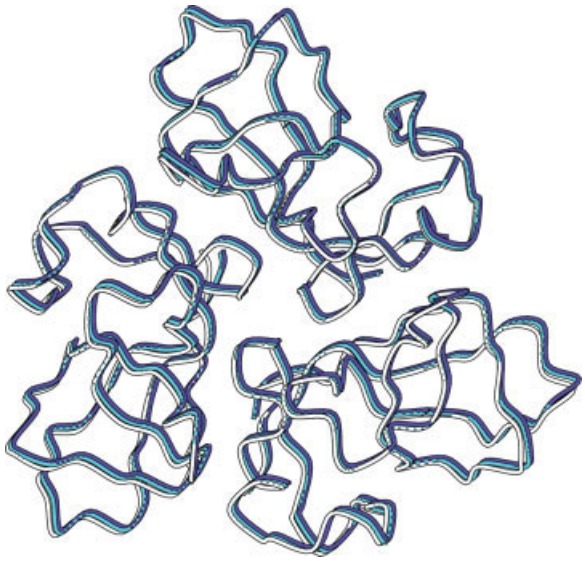

Fig. 1. Superposition of the trimers of full-length gpD and of the truncated form of this protein. Ribbon tracing of the full-length gpD is drawn in white, molecule 1 of gpDΔN1 in cyan, and molecule 2 of gpDΔN1 in blue.

All main-chain atoms, with the exception of the first visible residues on the N terminus (amino acids 15 and 16), are within 0.5 Å in the two structures. For the side-chain atoms, only a few showed larger deviations from the corresponding atoms in full-length gpD. Those side chains are located in the surface region of a trimer with high temperature factors and do not affect the assembly of molecules. Thus the tertiary structure of gpDΔN1 is virtually the same as that of gpD (Fig. 1), including the intermolecular contacts between the monomers, and even the N-terminally located Pro17 and His 19. We can, thus, conclude that even though the crystal forms for the full-length and the truncated gpD differ, with two trimers present in the asymmetric unit of gpDΔN1 and only one in gpD, the absence of the N-terminal 14 residues does not affect the trimeric structure of this protein. However, the difference between the crystal forms of the full-length and truncated protein did not allow us to use the low-resolution data for locating the area occupied by the disordered N terminus of the full-length protein.

Acknowledgments. The authors thank Dr. Patrik Forrer for purification of the gpDΔN1 protein, helpful discussions, and critical reading of the manuscript.

REFERENCES

1. Campbell AM. Bacteriophages. In: Neidhardt FC, editor. *Escherichia coli* and *Salmonella*: cellular and molecular biology. Washington DC: ASM Press; 1996. p 2325–2338.
2. Hendrix RW, Garcea RL. Capsid assembly of dsDNA viruses. Semin Virol 1994;5:15–26.
3. Santini C, Brennan D, Mennuni C, Hoess RH, Nicosia A, Cortese R, Luzzago A. Efficient display of an HCV cDNA expression

library as C-terminal fusion to the capsid protein D of bacteriophage lambda. J Mol Biol 1998;282:125–135.
4. Mikawa YG, Maruyama IN, Brenner S. Surface display of proteins on bacteriophage lambda heads. J Mol Biol 1996;262:21–30.
5. Sternberg N, Hoess RH. Display of peptides and proteins on the surface of bacteriophage lambda. Proc Natl Acad Sci USA 1995;92: 1609–1613.
6. Yang F, Forrer P, Dauter Z, Conway JF, Cheng N, Cerritelli ME, Steven AC, Plückthun A, Wlodawer A. Novel fold and capsid-binding properties of the lambda-phage display platform protein gpD. Nat Struct Biol 2000;7:230–237.
7. Bragg WL, Perutz MF. The external form of the haemoglobin molecule. I. Acta Cryst 1952;5:277–283.
8. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol 1997;276: 307–326.
9. Navaza J. An automated package for molecular replacement. Acta Crystallogr 1994;A50:157–163.
10. Brünger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography and NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr 1998;D54: 905–921.
11. Jones TA, Kjeldgaard M. Electron-density map interpretation. Methods Enzymol 1997;277:173–208.