Journal of
Nanoscience and Nanotechnology
Vol. 7, 2616-2622, 2007

Crystallization of Phi29 Spindle-Shaped Nano-Bar Anti-Receptor with Glycosidase Domain

Alexander J. DiMauro[1], Dawei Lin[3], Songchuan Guo[2], Dale B. Karr[4], John J. Tanner[5,*], and Peixuan Guo[1,2,*]

[1] *Weldon School of Biomedical Engineering and* [2] *Department of Comparative Pathobiology, Purdue University, West Lafayette, IN 47907, USA*
[3] *Bioinformatics Core, Genome Center, University of California-Davis, Davis, CA 95616, USA*
[4] *Structural Biology Core and* [5] *Departments of Chemistry and Biochemistry, University of Missouri-Columbia, Columbia, MO 65211, USA*

Bacteriophage phi29 is a small, well-characterized dsDNA virus that infects *Bacillus subtilis*. The anti-receptor of phi29 consists of oligomers of the 854-residue protein gp12 and plays an essential role in infection initiation by binding to the receptor on the host cell surface. Oligomers of gp12 exhibit a narrow spindle-shaped configuration 15 nm in length as revealed by electron microscopy and thus are potentially useful nanoscale tools, building blocks, or motor arms. To understand the mechanism of viral infection initiation and to provide a basis for engineering recombinant gp12 for nanotechnology applications, we have initiated structural and bioinformatics studies of gp12. We report here the growth of crystals of gp12 that diffract to 3.0 Å resolution. The space group is $P3_121$ or $P3_221$ with unit cell lengths of $a = 84.4$ Å and $c = 167.6$ Å. The asymmetric unit is predicted to contain one gp12 molecule and 32% solvent ($V_M = 1.8$ Å^3/Da). Domain boundary analysis revealed that gp12 may harbor three domains besides a 24 residue auto-cleave region. The N-terminal half of gp12 contains a domain with about 400 residues that held 44% sequence identity to endopolygalacturonase, a fungal glycosyl hydrolase that catalyzes hydrolysis of the polygalacturonic acid α1-4 glycosidic linkage found in plant cell walls. Interestingly, the cell wall of *Bacillus subtilis* contains a polysaccharide component made from two sugar monomers, N-acetylmuramic acid and N-acetylglucosamine, which resemble α-galacturonic acid in that they possess a six-membered pyranose ring. Hence, polygalacturonic acid of plant cell walls and peptidoglycan of bacterial cell walls may offer a similar topography in relation to the polysaccharides. These results suggest a function for gp12 as a cell-wall degrading enzyme in addition to its role in recognition of the host receptor.

Keywords: Viral Receptor Anti-Receptor Interaction, Phi29, Crystallization, Domain Boundary Analysis.

RESEARCH ARTICLE

1. INTRODUCTION

The extensive administration of chemical antibiotics has led to the increasing growth and emergence of antibiotic-resistant bacterial strains. There is an escalating trend in the United States for the use of biological molecules such as small peptides or bacteriophage for the treatment of bacterial infection. Indeed, bacteriophage therapy has been explored since the 1920's and has been used for many decades in some areas as an alternative to commonly used antibiotics. Phi29 is a virus that infects *Bacillus subtilis*. Viral infection is initiated by the specific binding of the viral anti-receptor to the host receptor on the cell surface. In order for a virus to enter or inject its genetic material into a host cell, it is necessary to bind and then penetrate the bacterial cell wall. Elucidation of the mechanism of specific interaction between the phi29 anti-receptor and host bacteriophage cell receptor will provide the basis for the understanding of phage infecting mechanisms, for the design of biological components that bind and penetrate the bacterial cell wall, and for the development of biological antibiotics acting on the bacteria targeted by the appropriate bacteriophage anti-receptor.

Phi29 is simple in structure and quite small in size.[1–3] It is a particularly well-studied dsDNA virus.[4,5] Up to 90% of the phi29 genomic DNA added can be packaged into

*Authors to whom correspondence should be addressed.

the purified procapsid *in vitro*.[6–9] The DNA-filled procapsid can then be converted to 10^9 infectious virions per milliliter *in vitro* with the addition of tail protein gp9, upper collar protein gp11, anti-receptor gp12, and morphogenetic factor gp13.[10, 11] A highly sensitive *in vitro* phage phi29 assembly and assay system has been developed.[10–12] Replacement of any one of the components with an inactive one precludes all plaque formation. This system has an eighth to ninth-order sensitivity and thus provides a convenient assay for the stoichiometric determination and functional analysis of phi29 components.

The structural component gp12 of phi29 that is displayed on the outer surface of the viral particle is the antireceptor.[13–15] The purified recombinant gp12 protein has a narrow spindle-shaped configuration, which has important implications for nanotechnology. It could be used as a basic building block, since various nano-technological applications require a variety of shapes and sizes,[16–18] and spindle-shaped materials are rare. STEM images of the purified gp12 can be seen (Fig. 1), along with the assembled mature phi29 virus which has successfully incorporated the recombinant gp12. Given both its spindle-shaped structure and its two-domain configuration, gp12 would be highly useful in single-molecule microscopic techniques using optical trapping or fluorescence microscopy, since it would be possible to distinguish the direction of movement if one end is modified for substrate binding, and the other end is modified for linking to a fluorescent microsphere. This will overcome the problem in single molecule rotation studies of nanomotors of being incapable of detecting the direction of rotation due to the resolution limit of hundreds of nanometers as limited by the wavelength of light.

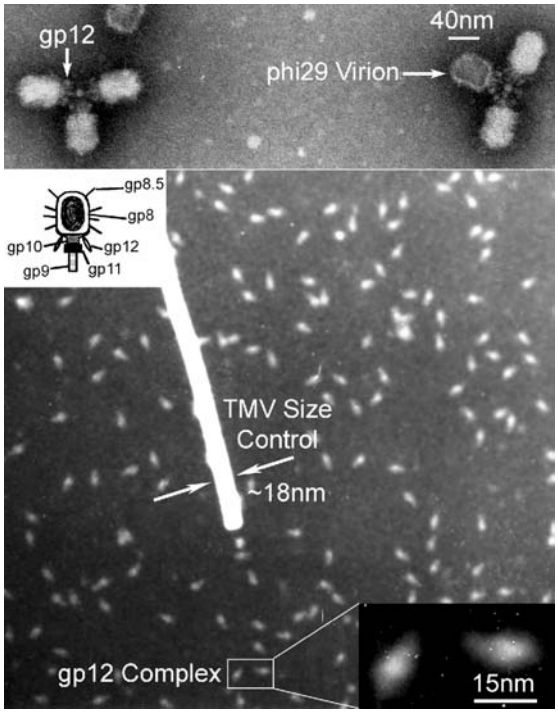


Fig. 1. Negative stain electron images of phi29 virions (A). the purified recombinant antireceptor gp12 oligomer by STEM (Scanning Transmission Electron Microscopy) (B), the schematic of the structure of phi29 virion (left inset), the tobacco mosaic virus control (the bar), and the closeup view of the antireceptor (right inset).

2. MATERIALS AND METHODS

2.1. Virion Assembly Assays

The activity of gp12 in phi29 virion assembly was measured as described previously.[19] Briefly, 10 μl of diluted purified His-procapsid (3 μg) was mixed with 1 μ1 pRNA (150 ng/μl) and dialyzed on a 0.025 mm pore size type VS filter (Millipore Corp.) against TBE for 15 min and then dialyzed against TMS for 30 min at ambient temperature. These pRNA-enriched His-procapsids were mixed with 3 μ1 reaction buffer (10 mM ATP, 6 mM spermidine, 3 mM mercaptoethanol in TMS), 5 μ1 phi29 DNA-gp3 that had been dialyzed against TMS for 40 min at ambient temperature, and 6 μl DNA-gp3, 6 μl gp16 dissolved in PEG buffer. After incubation at ambient temperature for 1 h, the final mixture was then incubated with 10 μl each of solution containing gp9, gp11-13, and gp12 to complete the assembly of infectious phage, which were assayed by standard plaque formation.

2.2. Crystallization of Recombinant gp12

An improved procedure was developed for obtaining highly pure recombinant gp12 for crystallization. The recombinant protein was overexpressed in *E. coli* stain HMS174-DE3 at 25 °C. Cells were grown in LB broth containing 100 μg/ml ampicillin until the optical density measured at 600 nm reached 0.5–0.7. The cell cultures were induced by adding isopropyl-β-D-thiogalactoside to a final concentration of 0.5 mM, and incubated an additional 3.5 h at 30 °C. The cells were harvested by centrifuging at 10,700 *g* for 15 minutes at 4 °C.

The *E. coli* cell pellet was resuspended in 0.1 M NaCl, 0.25 mM sodium EDTA, 10 mM $MgCl_2$, 50 mM Tris, pH 7.8, (5 ml buffer per g cell paste). The resuspended cells were broken by two passes through a French pressure cell adjusted to 16000 psi in a 40 k rapid fill cell and a flow rate of 3–4 ml/minute. The broken cell suspension was clarified be centrifugation at 35,000 g for 45 min at 4 °C. The resulting supernatant was next partitioned by ammonium sulfate precipitation at 4 °C. The protein fraction precipitating in the range 30–50% saturation contained gp12. This fraction was collected by centrifugation at 12,000 g for 15 min at 4 °C, and the resulting pellet was dissolved in 10 mM $MgCl_2$, 50 mM Tris, pH 7.8 and dialyzed overnight at 4 °C (2 changes). The dialyzed sample was filtered through a 0.2 μm PVDF membrane and applied to a HiTrapQ (5 ml) anion exchange column equilibrated in 10 mM $MgCl_2$, 50 mM Tris, pH 7.8. A gradient

RESEARCH ARTICLE

of 0–1 M NaCl was applied over 80 column volumes at a flow rate of 2 ml/min. Analysis of the eluted fractions by SDS-PAGE showed that gp12 eluted in the range 100–200 mM NaCl. Fractions containing gp12 were pooled and concentrated to 20 ml by ultrafiltration using a stirred cell fitted with a PES membrane (MWCO 30,000). The resulting protein solution was diluted 1:1 with 0.5 mM EDTA, 50 mM Tris, pH 7.5, and applied to CM Affi-Blue (12 ml, Bio-Rad) cation exchange column equilibrated in 75 mM NaCl, 0.5 mM EDTA, 50 mM Tris, pH 7.5. The gp12 protein eluted in the flow through and in subsequent fractions obtained by washing with 75 mM NaCl, 0.5 mM EDTA, 50 mM Tris, pH 7.5. These fractions were pooled, concentrated as described above, and further purified by HiTrap phenyl sepharose fast flow chromatography as follows. The gp12 protein sample was adjusted to 0.75 M ammonium sulfate, filtered through 0.22 μm PVDF membrane and applied to the HiTrap phenyl sepharose (5 ml) column, which had been equilibrated in 0.75 M ammonium sulfate, 0.5 mM EDTA, 50 mM Tris, pH 7.5. The gp12 protein was eluted from the phenyl sepharose using a descending gradient of 0.75–0 M ammonium sulfate. Fractions containing gp12 were pooled and dialyzed into 50 mM NaCl, 0.5 mM sodium EDTA, 50 mM Tris, pH 7.5, and concentrated to 10.5 mg/ml using ultrafiltration using a stirred cell (PES membrane, MWCO 30,000), followed by centrifugal ultrafiltration using an Amicon 4 ml Ultrafree filtertube (BIOMAX–MWCO 30,000). Protein concentration was measured by both the BCA and Bradford (Coomassie Plus, Pierce) methods. The yield of purified protein from 26 g of cell paste (12 L cultured cells) was 18 mg.

All crystallization trials were performed using the hanging drop method of vapor diffusion crystallization at 21 °C using Nextal hanging drop trays. Drops were formed by mixing 2 μl of the protein sample with 2 μl of the reservoir solution.

3. RESULTS AND DISCUSSION

3.1. Purification and Crystallization of gp12 as Well as X-ray Diffraction Data Collection

The gene coding for gp12 has been cloned into a plasmid.[19] An improved procedure was developed for producing recombinant gp12 having sufficient purity to allow growth of high quality crystals. This new procedure differs substantially from the published protocol[19] and involved four major purification steps as described in Materials and Methods: ammonium sulfate fractionation, anion exchange chromatography, cation exchange chromatography, hydrophobic interaction chromatography. This method differs from the previous protocol in the use of cation exchange chromatography and hydrophobic interaction chromatography. We found that these additional purification steps were necessary for obtaining crystallizable protein. Using the improved procedure, approximately 18 mg of highly purified protein could be obtained from 26 g of cell paste (12 L cultured cells). Virion assembly assays were performed to confirm that the purified recombinant gp12 is indeed the biologically relevant antireceptor gp12.

Initial crystallization conditions were identified using commercially available crystal screen kits. Positive results were obtained from screen reagents containing Tacsimate as the precipitating agent (Hampton Research). These initial conditions were used as the starting point for optimization trials in which the Tacsimate concentration, type of buffer, and pH were systematically varied to identify conditions that gave diffraction quality crystals. Optimization trials resulted in the growth of single crystals over reservoir solutions consisting of 34–38% Tacsimate, 0.1 M imidazole, pH 8.0. A typical crystal grown from

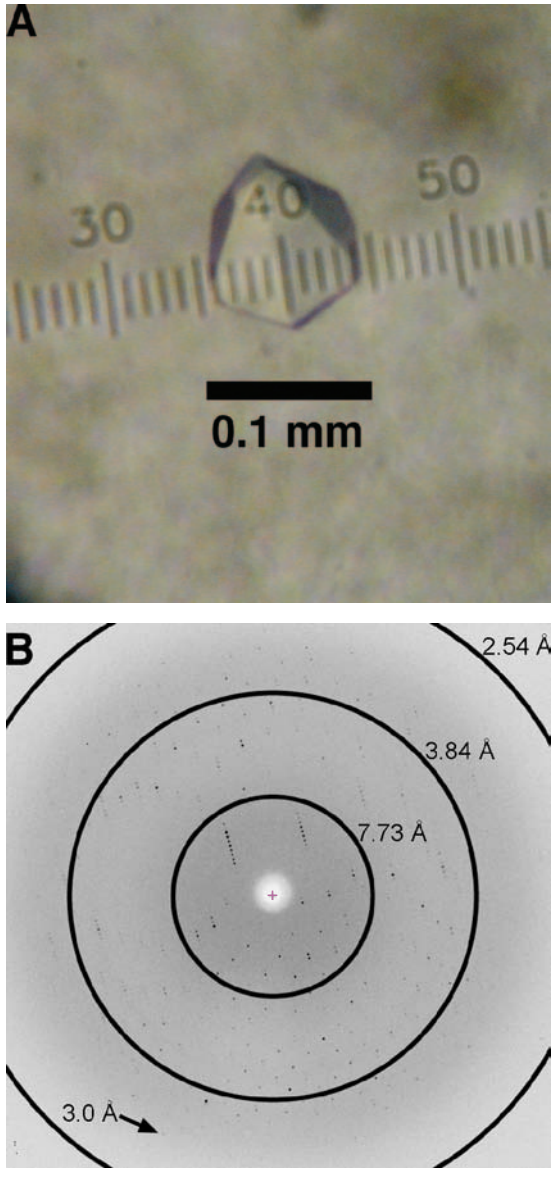


Fig. 2. (A) A crystal of recombinant gp12 grown in the presence of Tacsimate and imidazole buffer. (B) Diffraction image from a gp12 crystal.

RESEARCH ARTICLE

such conditions is shown in Figure 2(A). These crystals had hexagonal pyramid morphology and grew to a size of approximately 0.1 mm (Fig. 2(A)). The crystals were cryoprotected by soaking in 42% Tacsimate, 0.1 M imidazole, pH 8.0, and 20% PEG 200. The cryoprotected crystals were picked up with Hampton mounting loops and plunged into liquid nitrogen.

Diffraction data were collected at Advanced Light Source beamline 4.2.2 using a NOIR-1 CCD detector (Fig. 2(B)). A 3.0 Å resolution data set consisting of 180 frames was collected with an exposure time of 10 s/frame, oscillation width of 0.5°/frame and detector distance of 200 mm. Processing of the data set in *d**TREK.[20] suggested a trigonal/hexagonal lattice with unit cell lengths of $a = 84.4$ Å and $c = 167.6$ Å. Merging statistics and analysis with dtcell suggested Laue group -3 ml and probable space group $P3_121$ or $P3_221$. Using the method of Matthews[21] as implemented by Kantardjieff and Rupp,[22] the asymmetric unit is predicted to contain 1 gp12 molecule, which corresponds to 32% solvent and $V_M = 1.8$ Å^3/Da. Data collection and processing statistics for this data set are summarized in Table I. The structure of gp12 will be determined using experimental phasing methods owing to a lack of suitable search models for molecular replacement.

3.2. Domain Boundary and Functional Domain Predictions

The gp12 amino acid sequence was queried against the non-redundant sequence database (NR) of NCBI used to search BLAST.[23] The search result revealed (Fig. 3(A), marked with red) a region from residue 105 to 504 containing 400 amino acids with 44% similarity to endopolygalacturonase (evalue $= 4e^{-86}$), a protein from a fungus that infects plant cells. SIM alignment with endopolygalacturonase is shown in Figure 3(B).

Endopolygalacturonase is an enzyme that belongs to family 28 of glycosyl hydrolases from fungi[24] which catalyze the random hydrolysis of the polygalacturonic acid α1-4 glycosidic linkage. Polygalacturonic acid is a major constituent of the main chain of the homogalacturonan region of pectin in the plant cell wall. The endopolygalacturonases are thought to act by a general acid catalysis mechanism in which two amino acid residues participate. The hydrolysis may result in inversion (single-displacement reaction) or retention (double-displacement reaction) of configuration at the anomeric carbon atom of the hydrolyzed glycoside.[25] The host cell of phi29 is *Bacillus subtilis*, the cell wall of which contains peptidoglycan (polysaccharides + protein). The polysaccharide components are made from two sugar monomers, N-acetylmuramic acid (NAM) and N-acetylglucosamine (NAG). NAM and NAG resemble α-galacturonic acid in that they possess a six-member pyranose ring (Fig. 4). Hence, polygalacturonic acid of plant cell wall and peptidoglycan of bacterial cell wall may offer a similar topography in relation to the polysaccharides. It may be expected that the hydrolyzing enzymes acting on these two groups of polysaccharides may show considerable similarity in their structures. However, since the structure of an enzyme-substrate complex has not yet been obtained, the substrate binding mechanism of endopolygalacturonase is still unclear. It is predicted that gp12 residues 105–504 form a domain with endopolygalacturonase enzymatic activity, which is likely to contribute to cell surface receptor binding, polysaccharide bond degradation, and cell wall digging.

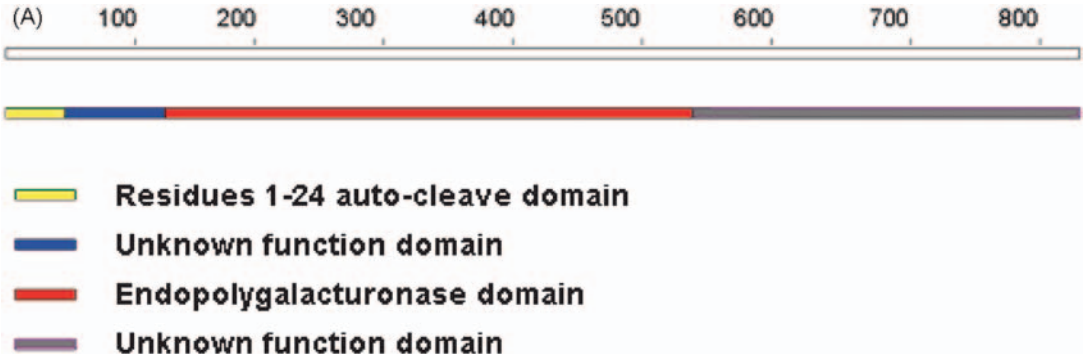
Table I. X-ray diffraction data collection and processing statistics.*

Beamline (Å)	ALS 4.2.2
Wavelength (Å)	0.9793
Laue group	-3 ml
Unit-cell dimensions (Å)	$a = 84.4$, $c = 167.6$
Diffraction resolution (Å)	45.4–3.0 (3.1–3.0)
No. of observations	77659
No. of unique reflections	14454
Redundancy	5.4 (5.5)
Completeness (%)	99 (99)
Average $I/\sigma(I)$	10.3 (4.3)
$R_{sym}(I)$	0.126 (0.350)

*Values for the outer resolution shell of data are given in parenthesis.

To construct a peptide drug for bacterial cell wall binding or digging, a minimal size particle is desired. Precise domain boundary information is crucial to the study of gp12 in a divide-and-conquer manner. Domain boundary prediction is the first step needed to initiate small peptide construction. Given that the normal size of a protein domain is around 150–300 residues, gp12, an 854-residue protein, is predicted to be a multi-domain protein. That information will allow cutting individual domains or domain combinations out of gp12 for function studies with the hope that the smaller size is more tractable than the whole protein. Since gp12 is an anti-receptor, it should contain at least two major functional domains, one binding to the procapsid of phi29, and another binding to the receptor on the surface of the host cell wall. To obtain precise domain boundary information, different types of bioinformatics predictions were employed. The integrative analysis of the computational results together with the experimental data from literature sources suggests that gp12 may have four domains (Fig. 3(A)).

The first approach was protein tertiary structure prediction using MetaSever at (www.bioinfo.pl).[26] A reasonably good model can be built from the region of residues 105–440 based on the structural templates suggested by 3D-Jury[26] (Fig. 5). The prediction to have a domain in this region is consistent with the sequence homology search and predicted 3D structure comparison, which revealed that this region of gp12 is significantly similar to endopolygalacturonase. The region within residues 500–600 that connects this domain with the region near the C-terminus end is thought to be a domain linker region because of the short length.



(B)

```
43.7% identity in 412 residues overlap; Score: 767.0; Gap frequency: 3.2%
gp12,    105 VSVKTYGAKGDGVTDDIRAFEKAIESGF-PVYVPYGTFMVSRGIKLPSNTVLTGAGKRNA
endopo,  107 VDVKDYGAVGDGVTDDTLAFKEALADGHRQVHVPAGTYLV-RGIRLPSYTILTGDGKGQT
             * ** *** *******  **  *   *   * ** **  * *** *** * *** **

gp12,    164 VIKFMDSVGRGESLMYNQNVTTGNENIFLSSFTLDGNNKRLGQGI-SGIGGSRESNLSIR
endopo,  166 ILKLHDAAPKGRRLITNQNYLLGNHHLLVERMTLDWNIERIGNAKKSSTWGNHSSCLTYA
               *  *    *  *  ***   **        *** *  * *    *   *   * *

gp12,    223 ACHNVYIRDIEAVDCTLHGIDITCGGLDYPYLGDGTTAPNPSENIWIENCEATGFGDDGI
endopo,  226 HVTYGWVFDVDAVNPGLHCFDISTPY--YNYNGDGARANLSSSFIWFDGLTGSGFGDDGI
                   *  **   **  **       * * ***  *   *  **        *******

gp12,    283 TTHHSQYINILNCYSHDPRLTA------NCNGFEIDDGSRHVVLSNNRSKGCYGGIEIKA
endopo,  284 TTHHSDHLFISNCFMHDPSGRAHAAGFSNSNGIEIDDGSRDVWLFNNATTRCFGGLEIKA
             *****    * **  ***   *       * ** ******* * * **    * ** ****

gp12,    337 HGDAPAAYNISINGHMSVEDVRSYNFRHIGHHAATDPQSVSAKNIVASNLVSIRPNNKRG
endopo,  344 HATSSAASTVKIVGHLSVDDHRSFNFRHIGHHQATDPTSQTAYNIIATRLIAQTPLYTPL
             *    **    * ** ** * ** ******** **** *  * ** *  *    *

gp12,    397 FQDNATPRVLAVSAYYGVVINGLTGYTDDP-NLLTETVVSVQFRARNCSLNGVGLTGFSN
endopo,  404 YQ-NSKPRALIVSAYRNVVIHDFTVLGDPYYDYQEQPMIGIQYKARNVTLSQVKMSGFRQ
              * *  ** * ****  ***   *   *                 *  ***  *  *   **

gp12,    456 SDNGIYVIGGSRGGDAVNISNVTLNNSGRYGVSIGSGIENVSITNISGIGDG
endopo,  463 AKTDIQVFGGDNRAEDVRLIDVRIEDSAKHGITIGPNIEHVRLTNVAIKGDG
                 * * **      *    *    *   *  **  ** *  **    ***
```

Fig. 3. (A) Predicted domains in gp12 protein of bacteriophage phi29. (B) SIM alignment result of gp12 and Endopolygalacturonase. Amino acids with * refer to an exact match. (reproduced from Progress Report to NIH in 2004).

The second approach was based on sequence analyses. DomCut[27] is a program that uses the amino acid composition differences between domains and domain linker regions to predict domain boundaries (Fig. 5(C)). The purpose of sequence feature predictions is to collect data on various sequence features to get more clues about domain boundaries. The program of disEMBL[28] was used to predict the disorder region (Fig. 5(D)), which often occurs in domain linker regions. Similarly, a search for a low complexity region in gp12 was performed (Fig. 5(E)). These regions have simple repeats of amino acids that are often found to be disordered. The predicted region overlaps with the same disorder region predicted by other programs, strengthening the domain linker or loop region prediction.

The third approach was a functional analysis of gp12 predicted domains by biological function assay based on homology search or sequence alignment to peptides or protein domains with known function. Protein sequencing of

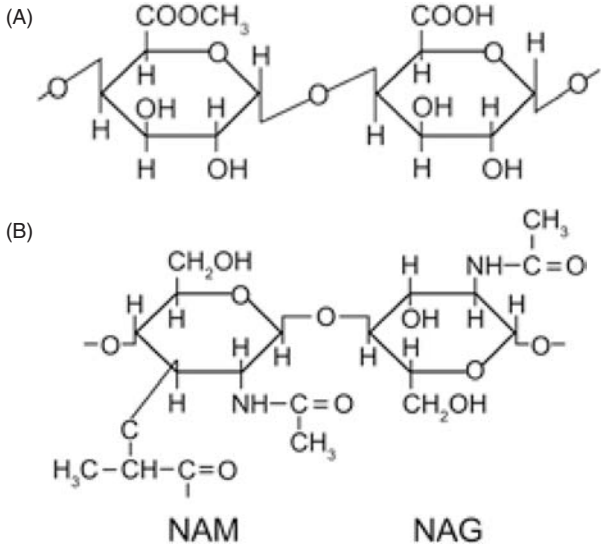


Fig. 4. Structure of polysaccharides: (A) α-1,4 polygalacturonic acid, and (B) N-acetylmuramic acid (NAM) and N-acetylglucosamine (NAG).

RESEARCH ARTICLE

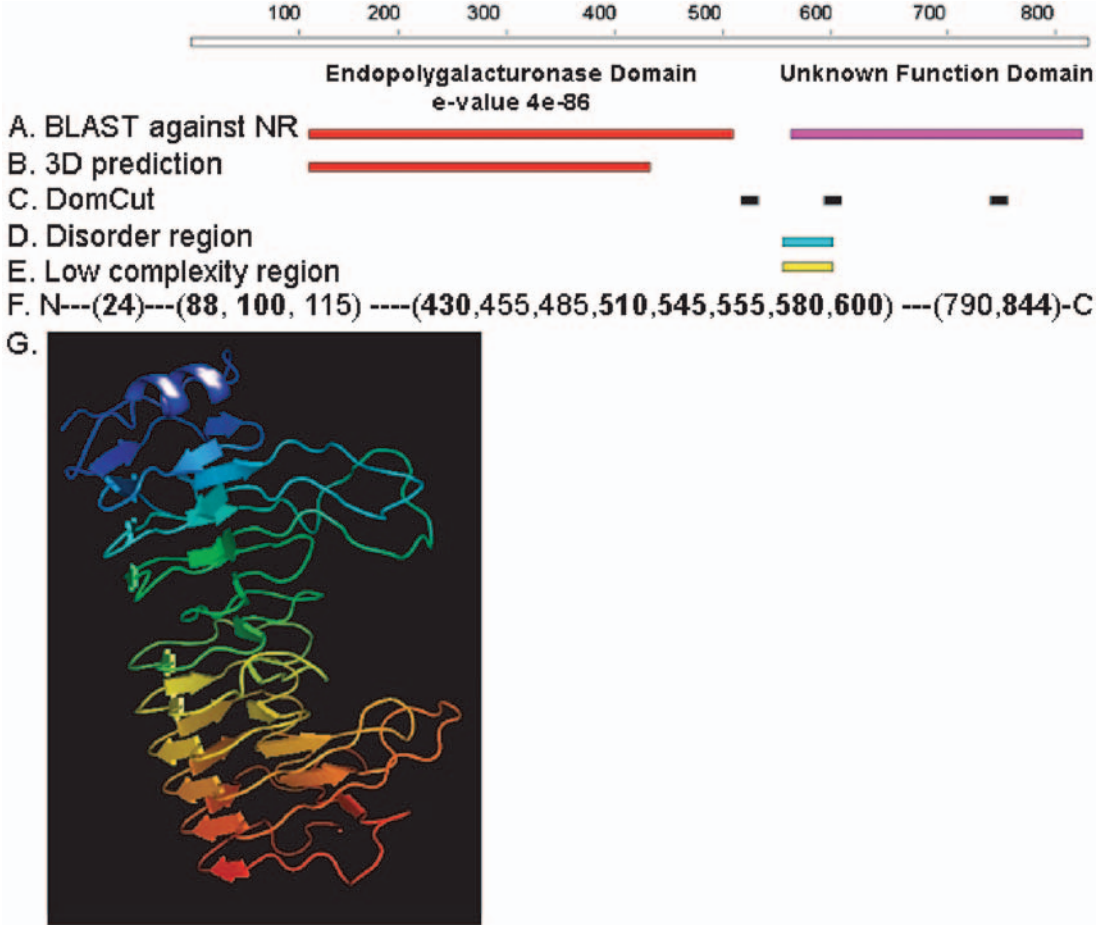


Fig. 5. (A)–(E) Domain boundary and sequence feature predictions for gp12. (F) Domain cutting sites with those considered important in bold. (G) 3D structure prediction for gp12 region 105–440. This model is built Based on PDB 1BHE. The quality scores from ProQ are LGscore: 2.7 and Max Sub: 0.18, which are considered "very good" and better than "fairly good" in ProQ classifications.

the purified recombinant gp12 revealed that the N-terminus sequence starts with N-LeuProThrAlaPheAspGluSer,[19] suggesting that the first 24 residues are cleaved after expression as compared to wild type gp12, and that loss of these residues does not affect the function of gp12.[19] So it was predicted that the first 24 residues are separated functionally from the rest of the sequence. Considering that there are about 80 residues between residue 24 and the start of the endopolygalacturonase domain at residue #105, the region of residues 25–105 may represent another domain, which could be deleted for functional studies on the endopolygalacturonase activity of gp12. Based on the predictions, other cutting sites that can be used to sub-clone individual domains or domain combinations for function studies are listed (Fig. 5(F)) with bold numbers considered important sites.

Acknowledgments: Thank Dr. Daisuke Kihara for bioinformatics analysis; Feroz Khan for comments on endopolygalacturonase prediction; John Turek for EM images of the phi29 virion trimer in Figure 1(A); Martha Simon for STEM image of the antireceptor in Figure 1(B); Jay Nix and Darren Sherrell for assistance at ALS beamline 4.2.2. This work was partially supported by NIH grant EB00370 to PG.

References and Notes

1. D. L. Anderson, H. H. Hickman, and B. E. Reilly, *J. Bact.* 91, 2081 (**1966**).
2. J. L. Carrascosa, E. Vinuela, and M. Salas, *Virology* 56, 291 (**1973**).
3. W. J. J. Meifer, J. A. Horcajadas, and M. Salas, *Microbiol. Mol. Biol Rev.* 65, 261 (**2001**).
4. P. Guo and M. Trottier, *Seminars in Virol.* 5, 27 (**1994**).
5. P. Guo, *Prog. Nucl. Acid Res. Mole. Biol.* 72, 415 (**2002**).
6. P. Guo, S. Grimes, and D. Anderson, *Proc. Natl. Acad. Sci. USA* 83, 3505 (**1986**).
7. P. Guo, C. Peterson, and D. Anderson, *J. Mol. Biol.* 197, 219 (**1987**).
8. P. Guo, B. Rajogopal, D. Anderson, S. Erickson, and C.-S. Lee, *Virology* 185, 395 (**1991**).
9. S. Grimes and D. Anderson, *J. Mol. Biol.* 209, 91 (**1989**).
10. C. S. Lee and P. Guo, *Virology* 202, 1039 (**1994**).
11. C. S. Lee and P. Guo, *J. Virol.* 69, 5018 (**1995**).
12. C. S. Lee and P. Guo, *J. Virol.* 69, 5024 (**1995**).
13. J. L. Carrascosa, A. Camacho, E. Vinuela, and M. Salas, *FEBS Lett.* 44, 317 (**1974**).
14. N. Villanueva, J. M. Lazaro, and M. Salas, *Eur. J. Biochem.* 117, 499 (**1981**).

15. M. E. Tosi, B. E. Reilly, and D. L. Anderson, *J. Virol.* 16, 1282 **(1975)**.
16. Z. Zhang and J. Buitenhuis, *Small* 3, 424 **(2007)**.
17. T. Douglas and M. Young, *Science* 312, 873 **(2006)**.
18. A. S. Blum, C. M. Soto, C. D. Wilson, T. L. Brower, S. K. Pollack, T. L. Schull, A. Chatterji, T. Lin, J. E. Johnson, C. Amsinck, P. Franzon, R. Shashidhar, and B. R. Ratna, *Small* 1, 702 **(2005)**.
19. S. Guo, D. Shu, M. Simon, and P. Guo, *Gene* 315, 145 **(2003)**.
20. J. W. Pflugrath, *Acta Crystallogr. D. Biol. Crystallogr.* 55, 1718 **(1999)**.
21. B. W. Matthews, *J. Mol. Biol.* 33, 491 **(1968)**.
22. K. A. Kantardjieff and B. Rupp, *Protein Sci.* 12, 1865 **(2003)**.
23. S. F. Altschul, T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman, *Nucleic Acids Res.* 25, 3389 **(1997)**.
24. B. Henrissat, *Biochem. J.* 280, 309 **(1991)**.
25. M. L. Sinnott, *Chem. Rev.* 90, 1171 **(1990)**.
26. K. Ginalski, A. Elofsson, D. Fischer, and L. Rychlewski, *Bioinformatics* 19, 1015 **(2003)**.
27. M. Suyama and O. Ohara, *Bioinformatics* 19, 673 **(2003)**.
28. R. Linding, L. J. Jensen, F. Diella, P. Bork, T. J. Gibson, and R. B. Russell, *Structure (Camb.)* 11, 1453 **(2003)**.

Received: 20 March 2007. Revised/Accepted: 28 March 2007.