Molecular Characterization of a Variant of *Bacillus anthracis*-Specific Phage AP50 with Improved Bacteriolytic Activity⁷†

Shanmuga Sozhamannan, 1* Michael McKinstry, 1 Shannon M. Lentz, 1 Matti Jalasvuori, 2 Farrell McAfee,¹ Angela Smith,¹ Jason Dabbs,¹ Hans-W. Ackermann,³ Jaana K. H. Bamford,² Alfred Mateczun, $¹$ and Timothy D. Read¹</sup>

Naval Medical Research Center, Silver Spring, Maryland 20910¹; Department of Biological and Environmental Science and
Nanoscience Center, P.O. Box 35, 40104-University of Jyväskylä, Jyväskylä, Finland²; and Department *Medical Biology, Faculty of Medicine, Laval University, Quebec, Quebec, Canada G1K 7P4*³

Received 19 May 2008/Accepted 29 August 2008

The genome sequence of a *Bacillus anthracis***-specific clear plaque mutant phage, AP50c, contains 31 open reading frames spanning 14,398 bp, has two mutations compared to wild-type AP50t, and has a colinear genome architecture highly similar to that of gram-positive** *Tectiviridae* **phages. Spontaneous AP50c-resistant** *B. anthracis* **mutants exhibit a mucoid colony phenotype.**

Bacillus anthracis, a category A biothreat agent, is a sporeforming gram-positive bacterium of the *Bacillus cereus* sensu lato group. It is a zoonotic soil bacterium that infects animals and occasionally humans, causing the disease anthrax. The standard diagnostic tests for suspected *B. anthracis* recommended by the Centers for Disease Control and Prevention (CDC) include testing for γ phage sensitivity (5). Recently, the inherent binding specificity and lytic action of the γ phageencoded lysin enzyme have been exploited for rapid detection and killing of *B. anthracis* (18). Although γ phage exhibits a fairly narrow host range, several *B. cereus* strains (e.g., ATCC 4342) have been shown to be sensitive to infection by this phage (1, 4, 6, 18).

In this study, we have characterized a mutant *B. anthracis*specific phage of the family *Tectiviridae*, AP50, isolated from soil using *B. anthracis* Sterne as the host in 1972 (2, 11). AP50 possesses a typical capsid architecture of the family *Tectiviridae*, i.e., isometric nucleocapsids with icosahedral symmetry and a capsid shell composed of two layers (Fig. 1a) (15). Originally it was thought to be an RNA phage but was later shown to contain double-stranded DNA and phospholipid (15). The wild-type AP50 phage, designated AP50t, was shown to have a narrow host range; only one-third of the 34 *B. anthracis* strains and none of the 52 strains belonging to six different *Bacillus* spp. were susceptible to infection by AP50 (13). Nine major structural proteins were identified on sodium dodecyl sulfatepolyacrylamide gel electrophoresis gels. The molecular mass of the phage DNA was estimated to be 9×10^6 daltons (14). Treatment with organic solvents such as chloroform (5%) and ether (25%) for 30 min inactivated the phage to a survival of about 1×10^{-4} PFU/ml (12).

Isolation of AP50c and determination of its burst size. AP50t produces turbid plaques characteristic of lysogenic phages, and spontaneous clear plaque variants designated AP50c were found in lysates after repeated propagation (Fig. 1b). The clear plaques were larger after incubation at 25°C than after incubation at 37°C and larger in phage assay agar (20) than in Luria-Bertani or blood agar plates (data not shown). The burst sizes at 25°C and 37°C reached about 200 to 300 particles/cell compared to about 3 to 5 particles/cell at 42°C.

Comparison of host ranges of phages AP50c and γ **. AP50c** infected 111 of the 115 *B. anthracis* strains $(\sim)97\%$ and none of the 100 *B. cereus* sensu lato strains (16) tested in this study, indicating a narrow host range. A side by side comparison of the infectivities of *B. anthracis* strains by phages AP50c and γ indicated that only 4 out of 115 strains were resistant to AP50c whereas 10 out of 115 strains were resistant to γ phage. Nine of the 10 strains that were resistant to infection by γ phage were susceptible to infection by AP50c, 3 of the 4 AP50cresistant strains were sensitive to γ , and one strain was resistant to both AP50c and γ phages (see Table S1 in the supplemental material).

AP50c genome sequencing and comparative analyses to other gram-positive bacterium-specific tectiviral genomes. The genome sequence of phage AP50c was determined by pyrosequencing technology using a Roche/454 Life Sciences GS20 sequencer (10). AP50c DNA isolated from purified phage particles was used to sequence to an average depth of coverage of 175 reads per base. The final assembled sequence of the AP50c genome was determined to be 14,398 bp.

The $G+C$ content of AP50c is comparable to that of its host, i.e., 38% compared to 35% for the *B. anthracis* chromosome and 32% and 33% for plasmids pXO1 and pXO2, respectively. The AP50c genome has short (28-bp) almost perfect inverted terminal repeats, 19 of which are identical, and contains 31 putative open reading frames (ORFs) (Fig. 2a). AP50c exhibits limited similarity at the DNA level to other sequenced tectiviral phages. *Bacillus thuringiensis* phages Bam35, pGil01, and pGil16c are closely related to each other (86 to 99% identity),

^{*} Corresponding author. Mailing address: Genomics Department, Biological Defense Research Directorate Annex, Naval Medical Research Center, 12300 Washington Avenue, Rockville, MD 20852. Phone: (301) 231-6713. Fax: (301) 231-6799. E-mail: Shanmuga.Sozhamannan @med.navy.mil.

[†] Supplemental material for this article may be found at http://aem .asm.org/. ∇ Published ahead of print on 12 September 2008.

FIG. 1. Transmission electron micrograph of AP50c phage particles. (a) Uranyl acetate staining at a magnification of 297,000. Scale bar, 100 nm. (b) Plaque morphology of turbid and clear plaques in a mixed lysate.

and *B. cereus* element pBclin15 is more closely related to AP50c (61%) than to the *B. thuringiensis* elements (59%) (Fig. 2b).

gram-positive bacterium-specific tectiviral phage genomes of ~53%, and extensive similarities exist (15% to 77%) throughout the genome. All the putative proteins of AP50c and the locations of their coding sequences in comparison to the Gil16,

AP50c proteins have an overall percent identity to other

13012

FIG. 2. (a) Genome map of AP50. Three gene clusters based on functional grouping and similarities to other *Tectiviridae* phages infecting gram-positive bacteria are shown. ORF boxes are color coded to indicate the degree of amino acid identity with proteins of other tectiviral phages. The color of the ORFs and percentage identities are as follows: green, 60 to 80%; red, 40 to 60%; purple, 15 to 40%; blue, 15%. ITR, inverted terminal repeat; HVR, highly variable region. Open arrowheads indicate the locations of the mutations in AP50c phages. (b) Visualization summary of whole-genome nucleotide alignments of gram-positive tectiviral phages. The ClustalW alignment file generated from Multifasta alignment was visualized in Base-By-Base (3). In this type of alignment, if two sequences have insertions or deletions relative to one another, the output looks different depending on which of the two sequences is used as the base sequence. White, perfect nucleotide homology; blue, single nucleotide polymorphism; red, deletions in the indicated phage; green, insertions in the indicated phage. The GenBank accession numbers for the sequences used in the alignment are as follows: Bam35c, NC_005258; pBth35646, NZ_AAJM00000000; Gil01, AJ536073; Gil16c, AY701338; AP50, EU408779; pBclin15, AE01878. (c) Sequence changes in AP50c and AP50t genomes. The mutation in the noncoding region just upstream of ORF 1 at nucleotide position 164 is indicated. The second mutation is in ORF 28 at position 12881 and changes amino acid residue 91 (an isoleucine in AP50c to a valine in AP50t).

TABLE 1. Comparison of predicted ORFs and their functions in the AP50c genome to those of other *Tectiviridae* genomes

AP50 ORF	No. of residues of AP50 protein (position of coding sequence)	AP50 $G + C$ content $(\%)$	Identity $(\%)$ vs:			ORF (no. of amino acids encoded) of:			PRD1 protein (no. of amino	Possible function and/or type ^{a}
			Bam35	Gil16	pBClin15	Bam35	Gil16	pBClin15	acids)	
1 $\mathfrak{2}$ 3	$61(378 - 560)$ 180 (556-1095) 74 (1169-1390)	39.89 38.22 38.29	22.6 72.8 43.2	22.2 77.2 45.9	49.2 71.7 51.5	1(58) 2(167) 3(74)	1(56) 2(166) 3(75)	1(63) 2(167) 3(71)	P12(160)	MerR-type regulator, S ssDNA binding protein, S S
4 5	233 (1456-2154) 731 (2171-4364)	31.62 36.43	30.2 47	32.7 45.4	63.5 78.7	4(245) 5(735)	4(245) 5(753)	4(233) 5(729)	P8 (259) P1(553)	Genome terminal protein, S DNA polymerase, S
6 7	67 (4368-4568) 45 (4587-4721)	40.30 28.89	67.2 44.0	67.2 36.2	70.1 53.1	6(66) 7(50) 8(46)	6(66) 7(58)	6(66) 7(49)		LexA-type repressor, S S
8	29 (4725-4811)					9(57)	8(84)			
9 10	77 (4827-5057) 57 (5073-5243)	31.25 30.00								S S S
11 12 13	114 (5244-5584) 235 (5428-6132) 83 (5981-6229)	43.86 46.67 44.98	40.0 37.5 49.4	55.2 38.6 48.2	44.2 43.3 54.2	10(145) 11(252) 12(80)	9(125) 10(248) 11(80)	8(115) 9(243) 10(81)	P6(166) P10(203)	Capsid protein/DNA packaging, S Assembly, S S
14	212 (6241-6876)	41.35	63.7	61.3	72.2	13(102) 14(212)	12(102) 13(212)	11(106) 12(216)	P9 (227)	DNA packaging ATPase, S
15 16	46 (6901-7038) 49 (7047-7193)	32.61 34.69	60.9 44.9	60.9 40.8	60.9 53.1	15(46) 16(46)	14(46) 15(45) 16(46)	13(46) 14(45)	P20(42)	TM DNA packaging, TM
17	354 (7196-8257)	39.55	64.9	62.9	63.7	17(84) 18(356)	17(89) 18(356)	15(355)	P3 (395)	Major capsid protein, S
18 19	74 (8306-8527) 56 (8536-8704)	37.84 35.59	26.9 19.1	29.5 46.4	35.1 17.5	19(76) 20(68)	19(76) 20(52) 21(35)	16(76) 17(79)		TM TM
20 21 22 23	157 (8783-9253) 58 (9256-9429) 48 (9432-9575) 91 (9591-9863)	38.85 35.63 37.50 43.22	33.1 55.2 52.1 27.7	35.6 55.2 52.1 29.8	30.6 51.7 62.5 25.5	21(143) 22(58) 23(48) 24(91)	22(143) 23(58) 24(48) 25(91)	18(140) 19(58) 20(48) 21(94)	P ₁₈ (90) P32(54) P34(68) P30(84)	DNA delivery, TM DNA delivery, TM TM Minor capsid protein, S
24 25 26 27	210 (9870-10499) 218 (10502-11155) 175 (11155-11679) 304 (11693-12604)	41.75 42.11 46.29 42.24	16.8 56.9 60.0 57.5	15.3 57.7 58.9 53.3	18.5 54.9 61.1 21.6	25(204) 26(250) 27(170) 28(304)	26(204) 27(250) 28 (170) 29(297)	22(197) 23(243) 24(173) 25(264)	P ₁₁ (207) P7(265) P31 (126) P ₅ (340)	DNA delivery, TM DNA delivery/glucosaminidase, TM Pentameric base of spike/peptidase, TM Trimeric spike protein, S
28	133 (12611-13009)	34.84				29(293)	30(293)	26(597)		S
29 30	118 (13027-13380) 49 (13385-13531)	32.62 36.05						27(95)		S S
31	252 (13536-14291)	42.99		45.7		30(265)	31 (288)	28 (227)		Endolysin, S Endolysin

^a ssDNA, single-stranded DNA; S, soluble; TM, transmembrane.

Bam35, and pBclin15 protein coding sequences are listed in Table 1. The AP50c genome is colinear with other known gram-positive *Tectiviridae* phage genomes (7, 17, 19, 21, 22) and has a cassette-like organization. The left end of the genome contains genes involved in genome replication (ORFs 2 to 5) and lysogeny control, and the middle part contains virion structural and DNA packaging genes (ORFs 11 to 27). At least five capsid proteins can be predicted based on similarities to other *Tectiviridae* phages: ORFs 11, 17, 23, 26, and 27. The right end contains the host recognition and lysis genes. Similarly to other *Tectiviridae* phages, 21 of the AP50c ORFs were predicted to encode soluble proteins, while 10 of the proteins had one or two transmembrane regions.

Several ORFs of AP50c encoded proteins that showed no significant similarity to any proteins in the databases, and they have been designated as highly variable regions. The ORFs encoding these proteins are located in two distinct regions of the genome: the first between bp 4827 and 5242 (ORFs 8 to 10) and the second between bp 12611 and 13531 (ORFs 28 to 30). Except for these ORFs, all the ORFs have orthologs in other *Tectiviridae* phages.

There are two ORFs in the AP50c genome that encode putative regulatory proteins. ORF 1 encodes a MerR-like protein, and ORF 6 encodes a LexA-like repressor protein. AP50c ORF 1 is similar to ORF 1 of the Bam35, Gil16c, and pBClin15 genomes, which exhibited similarities to truncated MerR-like regulatory proteins. A role for the LexA-like repressor (ORF 6 product) in lysogeny maintenance in Gil01 and Gil16 phages has been determined (21). Clear plaque variants of these phages possessed mutations in the ORF 6-ORF 7 region.

ORF 31 of AP50c encodes a putative endolysin. A protein alignment of ORF 31 to similar ORFs in other tectiviral phage genomes revealed several interesting features. ORF 31 of the AP50c genome is very similar to ORF 31 of Gil16c: the encoded proteins are 65% identical and 79% positive over a 186-amino-acid region. They possess the conserved endolysin domain of the amidase-3 family (*N*-acetylmuramoyl-L-alanine amidase), involved in degradation of bacterial cell walls. Also, the putative endolysins of Bam35 (encoded by ORF 30) and pBclin15 (encoded by ORF 28) are similar to each other (67% identical/77% positive over a 226-amino-acid region). They possess an Acm (1,4- β -*N*-acetylmuramidase)/glycosyl hydrolase family 25 domain. As expected, neither the AP50c nor the Gil16c ORF, encoding an amidase, is similar to the Bam35 or pBclin15 ORF, encoding a muramidase, despite the fact that overall Gil16c is genetically closer to Bam35 than it is to AP50

FIG. 3. Colony morphologies of *B. anthracis* Sterne strain 34F₂ after infection with AP50t and morphology of AP50c-resistant 34F₂ mutants. (a) Uninfected 34F2 cells diluted and plated on phage assay agar plates. (b) AP50t-infected culture, diluted and plated. The smooth and round white colonies are uninfected nonlysogens (solid arrowhead), and the flat wrinked colonies (open arrowhead) are lysogens. (c) Scanning electron micrograph of a wild-type $34F_2$ cell infected with AP50c. The arrowheads indicate AP50c particles attached to the outer surface of the bacterium (magnification, 60,000; Dennis Kunkel Microscopy, Inc.). (d) Scanning electron micrograph of 34F₂ AP50c^r mutant infected with AP50c showing the presence of extracellular material coating the outer cell surface and the absence of attached phage particles (magnification, 60,000; Dennis Kunkel Microscopy, Inc.).

(Fig. 2b). Therefore, it seems possible that there has been an interspecies exchange of endolysin domains between AP50/ Gil16c and pBclin15/Bam35, resulting in divergent endolysins in otherwise closely related *B. thuringiensis* phages (Bam35 and Gil16c). Such nonorthologous replacements have been found to be common among tailed phages (8). Interestingly, the carboxy-terminal regions of endolysins, involved in cell wall binding and host recognition (9), of Gil16 and Bam35 endolysins are somewhat similar (37% identical/58% positive over 70 amino acids at the C terminus), as they share the same host, *B. thuringiensis*. Also, the carboxyl termini of AP50, pBclin15, and Bam35/Gil16 endolysins are different, probably due to differing host specificities.

Identification of the genetic changes in turbid/clear plaque variants. DNA sequencing of the entire genomes of five turbid plaque phages revealed two changes in the genome: a single base substitution (T to C) at position 154 in a noncoding region at the left end of the AP50 genome and a nonsynonymous change at position 12881 within ORF 28 of unknown function which alters an isoleucine at position 91 in the AP50c protein to leucine in the AP50t phage protein (Fig. 2c).

AP50c exhibits very high bacterial killing efficiency. Bacterial survivors of phage AP50t infection arose at a very high frequency (50% of the initial CFU prior to infection), and the colonies were morphologically different from the uninfected cells: flat and wrinkled (Fig. 3a and b). These survivors were the result of lysogen formation. On the other hand, AP50c elicited a very high killing efficiency (5- to 6-log decrease in bacterial viability) since it propagates exclusively lytically. The survivors of AP50c infection appeared at a frequency of \sim 1 \times 10^{-6} CFU/ml. These AP50 c^r cells exhibited other phenotypes: while wild-type $34F_2$ cells settled and formed a pellet at the bottom of the tube upon overnight incubation statically at room temperature, AP50c^r cells formed a weak pellet and the colonies were mucoid. Furthermore, $AP50c^r$ cells appeared to secrete an extracellular flocculent material. Indeed, scanning electron microscopic analysis of AP50c^r cells revealed the presence of a thin coating of an extracellular material on the outer cell surface (Fig. 3c and d) and the absence of AP50c particle attachment. The AP50 c^r cells were sensitive to infection by γ phage.

Thus, AP50c, as described here, exhibits specificity and sen-

sitivity comparable to those of γ . Phages AP50c and γ probably have different target bacterial receptors since some of the naturally occurring and laboratory-isolated strains resistant to γ infection are sensitive to AP50 and vice versa (see Table S1 in the supplemental material; S. Sozhamannan, unpublished data). Thus, a combination of these two phages might be a better alternative for phage-based diagnostics, therapeutics for *B. anthracis* infection, and environmental clean-up of *B. anthracis*-contaminated areas.

Nucleotide sequence accession number. The nucleotide sequence of phage AP50c reported here has been deposited in the NCBI GenBank database under accession number EU408779.

This work was supported by funds from the Defense Threat Reduction Agency, Department of Defense of the U.S. government (BDRD authors; grant no. 8.10084_08_NM_B), and by Finnish Centre of Excellence Program (2006-2011) grant 1213467 (J.K.H.B.).

The views expressed in this article are those of the authors and do not necessarily reflect the official policy or position of the Department of the Navy, Department of Defense, or the U.S. government.

REFERENCES

- 1. **Abshire, T. G., J. E. Brown, and J. W. Ezzell.** 2005. Production and validation of the use of gamma phage for identification of *Bacillus anthracis*. J. Clin. Microbiol. **43:**4780–4788.
- 2. **Ackermann, H.-W., R. Roy, M. Martin, M. R. Murthy, and W. A. Smirnoff.** 1978. Partial characterization of a cubic *Bacillus* phage. Can. J. Microbiol. **24:**986–993.
- 3. **Brodie, R., A. J. Smith, R. L. Roper, V. Tcherepanov, and C. Upton.** 2004. Base-By-Base: single nucleotide-level analysis of whole viral genome alignments. BMC Bioinformatics **5:**96.
- 4. **Brown, E. R., and W. B. Cherry.** 1955. Specific identification of *Bacillus anthracis* by means of a variant bacteriophage. J. Infect. Dis. **96:**34–39.
- 5. **Centers for Disease Control and Prevention.** 2002. Anthrax Q & A: diagnosis. http://www.bt.cdc.gov/agent/anthrax/faq/diagnosis.asp.
- 6. **Davison, S., E. Couture-Tosi, T. Candela, M. Mock, and A. Fouet.** 2005. Identification of the *Bacillus anthracis* (γ) phage receptor. J. Bacteriol. 187: 6742–6749.
- 7. **Grahn, A. M., J. K. Bamford, M. C. O'Neill, and D. H. Bamford.** 1994. Functional organization of the bacteriophage PRD1 genome. J. Bacteriol. **176:**3062–3068.
- 8. **Hendrix, R. W., J. G. Lawrence, G. F. Hatfull, and S. Casjens.** 2000. The origins and ongoing evolution of viruses. Trends Microbiol. **8:**504–508.
- 9. **Loessner, M. J., K. Kramer, F. Ebel, and S. Scherer.** 2002. C-terminal domains of *Listeria monocytogenes* bacteriophage murein hydrolases determine specific recognition and high-affinity binding to bacterial cell wall carbohydrates. Mol. Microbiol. **44:**335–349.
- 10. **Margulies, M., M. Egholm, W. E. Altman, S. Attiya, J. S. Bader, L. A. Bemben, J. Berka, M. S. Braverman, Y. J. Chen, Z. Chen, S. B. Dewell, L. Du, J. M. Fierro, X. V. Gomes, B. C. Godwin, W. He, S. Helgesen, C. H. Ho,**

G. P. Irzyk, S. C. Jando, M. L. Alenquer, T. P. Jarvie, K. B. Jirage, J. B. Kim, J. R. Knight, J. R. Lanza, J. H. Leamon, S. M. Lefkowitz, M. Lei, J. Li, K. L. Lohman, H. Lu, V. B. Makhijani, K. E. McDade, M. P. McKenna, E. W. Myers, E. Nickerson, J. R. Nobile, R. Plant, B. P. Puc, M. T. Ronan, G. T. Roth, G. J. Sarkis, J. F. Simons, J. W. Simpson, M. Srinivasan, K. R. Tartaro, A. Tomasz, K. A. Vogt, G. A. Volkmer, S. H. Wang, Y. Wang, M. P. Weiner, P. Yu, R. F. Begley, and J. M. Rothberg. 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature **437:**376–380.

- 11. **Nagy, E.** 1974. A highly specific phage attacking *Bacillus anthracis* strain Sterne. Acta Microbiol. Acad. Sci. Hung. **21:**257–263.
- 12. **Nagy, E., and G. Ivanovics.** 1982. Anthrax-specific "AP 50-like" phages isolated from *Bacillus cereus* strains. Acta Microbiol. Acad. Sci. Hung. **29:** 89–98.
- 13. **Nagy, E., and G. Ivanovics.** 1977. Association of probable defective phage particles with lysis by bacteriophage AP50 in *Bacillus anthracis*. J. Gen. Microbiol. **102:**215–219.
- 14. **Nagy, E., O. Herczegh, and N. Ivanova.** 1982. Lipid-containing anthrax phage AP50: structural proteins and life cycle. J. Gen. Virol. **62:**323–329.
- 15. **Nagy, E., B. Pragai, and G. Ivanovics.** 1976. Characteristics of phage AP50, an RNA phage containing phospholipids. J. Gen. Virol. **32:**129–132.
- 16. **Priest, F. G., M. Barker, L. W. Baillie, E. C. Holmes, and M. C. Maiden.**

2004. Population structure and evolution of the *Bacillus cereus* group. J. Bacteriol. **186:**7959–7970.

- 17. **Ravantti, J. J., A. Gaidelyte, D. H. Bamford, and J. K. Bamford.** 2003. Comparative analysis of bacterial viruses Bam35, infecting a gram-positive host, and PRD1, infecting gram-negative hosts, demonstrates a viral lineage. Virology **313:**401–414.
- 18. **Schuch, R., D. Nelson, and V. A. Fischetti.** 2002. A bacteriolytic agent that detects and kills *Bacillus anthracis*. Nature **418:**884–889.
- 19. **Stro¨msten, N. J., S. D. Benson, R. M. Burnett, D. H. Bamford, and J. K. Bamford.** 2003. The *Bacillus thuringiensis* linear double-stranded DNA phage Bam35, which is highly similar to the *Bacillus cereus* linear plasmid pBClin15, has a prophage state. J. Bacteriol. **185:**6985–6989.
- 20. **Thorne, C. B.** 1968. Transducing bacteriophage for *Bacillus cereus*. J. Virol. **2:**657–662.
- 21. **Verheust, C., N. Fornelos, and J. Mahillon.** 2005. GIL16, a new grampositive tectiviral phage related to the *Bacillus thuringiensis* GIL01 and the *Bacillus cereus* pBClin15 elements. J. Bacteriol. **187:**1966–1973.
- 22. **Verheust, C., G. Jensen, and J. Mahillon.** 2003. pGIL01, a linear tectiviral plasmid prophage originating from *Bacillus thuringiensis* serovar israelensis. Microbiology **149:**2083–2092.