Structure of a group A streptococcal phage-encoded virulence factor reveals a catalytically active triple-stranded β-helix

Nicola L. Smith*†‡, Edward J. Taylor†§, Anna-Marie Lindsay*, Simon J. Charnock*¶, Johan P. Turkenburg§, Eleanor J. Dodson§, Gideon J. Davies§‖, and Gary W. Black*‖

*Chemical Biology Research Group, School of Applied Sciences, Northumbria University, Newcastle upon Tyne NE1 8ST, United Kingdom; and §York Structural Biology Laboratory, Department of Chemistry, University of York, York YO31 5YW, United Kingdom

Edited by Janet Thornton, European Bioinformatics Institute, Cambridge, United Kingdom, and approved October 10, 2005 (received for review June 8, 2005)

Streptococcus pyogenes **(group A Streptococcus) causes severe invasive infections including scarlet fever, pharyngitis (streptococcal sore throat), skin infections, necrotizing fasciitis (flesh-eating disease), septicemia, erysipelas, cellulitis, acute rheumatic fever, and toxic shock. The conversion from nonpathogenic to toxigenic strains of *S. pyogenes* is frequently mediated by bacteriophage infection. One of the key bacteriophage-encoded virulence factors is a putative ''hyaluronidase,'' HylP1, a phage tail-fiber protein responsible for the digestion of the *S. pyogenes* hyaluronan capsule during phage infection. Here we demonstrate that HylP1 is a hyaluronate lyase. The 3D structure, at 1.8-Å resolution, reveals an unusual triple-stranded β-helical structure and provides insight into the structural basis for phage tail assembly and the role of phage tail proteins in virulence. Unlike the triple-stranded β-helix assemblies of the bacteriophage T4 injection machinery and the tailspike endosialidase of the *Escherichia coli* K1 bacteriophage K1F, HylP1 possesses three copies of the active center on the triple-helical fiber itself without the need for an accessory catalytic domain. The triple-stranded β-helix is not simply a structural scaffold, as previously envisaged; it is harnessed to provide a 200-Å-long substrate-binding groove for the optimal reduction in hyaluronan viscosity to aid phage penetration of the capsule.**

crystal | enzyme | lyase | streptococcus | bacteriophage

Streptococcus pyogenes (group A Streptococcus) causes a wide range of acute infections in the respiratory tract and skin with invasive infections causing the life-threatening ''flesh-eating'' disease, necrotizing fasciitis, and the pyrogenic exotoxin-associated toxic shock syndrome (1). In 1927 it was demonstrated that *S. pyogenes* could be converted from a noninfective to a toxigenic state through the application of a soluble extract of scarlet fever isolates (2), which was subsequently shown to contain bacteriophage. Several of the virulence factors produced by *S. pyogenes* strains are now known to be bacteriophage-encoded. These factors include pyrogenic exotoxins A and C (3, 4), DNase (5), and ''hyaluronidase'' (6). Indeed, the involvement of phage-encoded virulence factors is now known to be a feature common to many bacterial pathogens (7, 8) with bacteriophage implicated in genome rearrangement (9) and, in the case of *Streptococcus*, interstrain transfer of pyrogenic exotoxin (*Spe*) genes (5, 10, 11). Spes, notably virulence factors SpeC and SpeA, are ''superantigens'' that directly cross-link the major histocompatibility complex class II on the antigen-presenting cells with T cell receptors, stimulating between 5% and 30% of the total T cell population (5, 12). The increased amounts of inflammatory cytokines result in the symptoms of scarlet fever and, in some cases, toxic shock (1, 12, 13). *Spe* genes are expressed at prophage induction (5), and the simultaneous induction of SpeC and prophage was shown to be mediated by a ''soluble factor'' by coculturing streptococci with pharyngeal cells (14). The precise events inducing conversion from lysogenic to a lytic cycle remain, however, unclear.

The genome sequencing of S. *pyogenes* strain SF370, revealed three full-length prophage sequences (15) belonging to the *Siphoviridae* family of bacteriophage (λ-like bacteriophage with long noncontractile tails). Of particular interest is the hyaluronidase ORF, HylP1, annotated among the tail fiber genes within the SF370.1 prophage (Fig. 1*a*). The likely role of this hyaluronidase is to allow phage penetration of the streptococcal hyaluronic acid (HA) capsule (16) (Fig. 1*b*) during transfection. To gain insight into the phage-mediated conversion to pathogenicity, in particular the role of the putative tail-fiber hyaluronidase, biochemical and structural analysis of HylP1 (from the only prophage of *S. pyogenes* SF370 that has been shown to be inducible, prophage SF370.1) was undertaken. We confirm that HylP1 is in fact a hyaluronate lyase, catalyzing a β-elimination reaction as first proposed by Baker *et al.* (17) and not a hyaluronan hydrolase. The 3D structure reveals a central triply wound β-helix on a protein >220 Å in length. All previous observations of this unusual fold have been phage-tail proteins, giving considerable credence to a similar role in phage SF370.1. It is widely believed that the triple-stranded β-helix (TSβH) motif is characteristic of structural proteins from viral tails, yet in marked contrast to the triple-stranded β-helical structures of the bacteriophage T4 injection machine (18) and the phages infecting the polysialic-acid-encapsulated human pathogen *Escherichia coli* K1 (19), the catalytic center of HylP1 does not reside on an accessory catalytic domain but on the triple-stranded β-helical core itself, whose organization may shed light on the assembly of the *Siphoviridae* tail fibers themselves.

Experimental Procedures

Cloning, Overexpression, and Site-Directed Mutagenesis. The amino sequence of HylP1 (AAK33657.1) was checked for the presence and location of signal peptide cleavage sites by using the SignalP 3.0 server, (www.cbs.dtu.dk/services/SignalP) (20). The entire coding sequence (CDS) *hylP1* (*SPy0701*) was amplified from the

Conflict of interest statement: No conflicts declared.

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: HA, hyaluronic acid; MAD, multiple anomalous difference; TSβH, triple-stranded β-helix.

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 2C3F).

†N.L.S. and E.J.T. contributed equally to this work.

‡Present address: Institute for Cell and Molecular Biosciences, University of Newcastle upon Tyne, Newcastle upon Tyne NE2 4HH, United Kingdom.

¶Present address: Megazyme International Ireland Limited, Bray, County Wicklow, Republic of Ireland.

‖To whom correspondence may be addressed. E-mail: gary.black@northumbria.ac.uk or davies@ysbl.york.ac.uk.

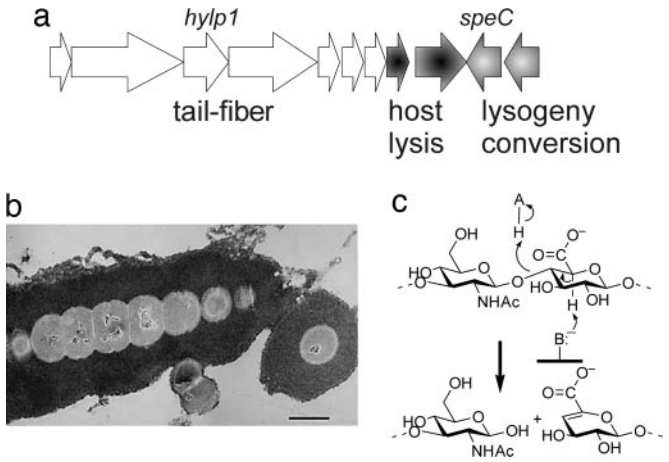

Fig. 1. Partial genetic map of *S. pyogenes* prophage SF370.1 and the hyaluronic acid capsule of streptococci. (*a*) The gene encoding HylP1 is found in the region assigned to the phage tail genes, with *speC*, encoding the phage supertoxin, located in the region associated with lysogenic conversion. (*b*) An electron micrograph of streptococcal cells surrounded by a thick layer of hyaluronan. Typically, the hyaluronan capsule is one to three times the diameter of the cell body. The anionic hyaluronan was visualized with cationic ferritin particles (≈10 nm) containing electron-dense iron cores. (Scale bar: 1 μm, an ≈27,000-fold magnification.) [Reproduced with permission from Glycoforum (www.glycoforum.gr.jp) (Copyright 2004, Glycoforum).] (*c*) The β-elimination reaction catalyzed by a hyaluronate lyase. This reaction results in the formation of a Δ4,5 unsaturated product, which may be detected spectrophotometrically at 232 nm.

genomic DNA of *S. pyogenes* SF370, American Type Culture Collection 700294 (15). This procedure required a two-stage process, because the 3′ end of CDS *hylP3* (*SPy1445*) is identical to *hylP1*. The first set of primers used were 5′-ATggCAATCAATgggTCAAg-3′ and 5′-ggATCCCTATTTTTTTAgTATgAg-3′. The amplified product was cloned into pGEM-T Easy (Promega) and designated pHylP1 + 5′. The *hylp1* ORF was amplified from pHylP1 + 5′ by using 5′-CATATgAgTgAAAATATACCgCTg3-′ and 5′-ggATCCCTATTTTTTTAgTATgAg-3′ and cloned into pET28a (Novagen) on an NdeI–BamHI fragment and designated pHylP1. The conditions used for amplification were as specified for use with Platinum *Pfx* DNA polymerase (Invitrogen). Plasmids pD137A and pY149A were generated from pHylP1 by using the QuikChange site-directed mutagenesis kit (Stratagene) and mutagenic primer pairs 5′-gTggAgCggTCAATATTgCgTTgTCgTCTACCAgAg3-′ and 5′-CTCTggTAgACgACAACgCAATATTgACCgCTCCAC3-′, and 5′-ggTgCTggTgTTgTTgTCgCgTCTgACAATgATACCAgTg-3′ and 5′-CACTggTATCATTgTCAgACgCgACAACAACACCAgCAC C-3′, respectively.

Production and Purification of HylP1 and Derivatives. N-terminally hexahistidine-tagged HylP1 was produced through induction of pHylP1-containing *E. coli* BL21 (DE3; Novagen) cultures grown at 30°C. HylP1 was purified by means of affinity chromatography using nickel-charged Sepharose chelating fast flow resin (GE Healthcare, Chalfont St. Giles, U.K.), followed by gel filtration using a HiLoad 16/60 Superdex 200 prep grade column and concentrated and exchanged into 10 mM Hepes (pH 7.4)/50 mM $CaCl_2$ using a 10-kDa cut-off concentrator unit (Vivascience, Hannover, Germany). D137A and Y149A were produced and purified as for the native enzyme, except they were produced from pD137A- and pY149A-containing *E. coli* BL21, respectively. Seleno-methionine HylP1 was produced and purified as for the native protein, except it was produced from *E. coli* B834 (DE3; Novagen), and 10 mM DTT was incorporated to all buffers before affinity chromatography. Purity of HylP1 was judged by SDS/PAGE, and static nanospray ion trap mass spectrometry (LCQ Advantage, ThermoFinnegan, San Jose, CA) of a tryptic digest was used to confirm the identity of the protein.

HylP1 Assays. Assays were performed spectrophotometrically at 37°C in triplicate at 232 nm by using a 1-cm pathlength quartz cuvette containing 2 mg·ml^{-1} substrate in 50 mM NH_4 Ac buffer (pH 6.5) unless otherwise stated. A molar absorptivity value of 5.5×10^3 liter·mol^{-1}·cm^{-1} was used for concentration calculations (21). All polymeric substrates, hyaluronan (K salt) from umbilical cord, chondroitin 4-sulfate (Na salt) from bovine trachea, and chondroitin 6-sulfate (Na salt) from shark cartilage, were treated with EDTA and dialyzed extensively before use (22). The range of hyaluronan concentrations used for determining k_{cat} and K_m values was 1.3–3 mg·ml^{-1}. Goodness-of-fit statistical analysis of the linear trendlines of the resulting Lineweaver–Burke plots produced R^2 values of $\geq$0.9935. High-performance anion-exchange chromatography was performed according to the method of Lauder *et al.* (23).

Circular Dichroism (CD). A J-810 spectropolarimeter (Jasco U.K. Ltd., Great Dunmow, Essex, U.K.) was used for CD analysis of native and mutant derivatives of HylP1. CD spectra (260–210 nm) averaged from 15 scans were collected at 20°C by using a bandwidth setting of 3 nm, a scanning speed of 100 nm·min^{-1}, and a 1-mm-pathlength cuvette. Unfolding curves of the proteins in the presence of 0–5 M guanidinium hydrochloride were generated by using the 218-nm ellipticity values (see Fig. 5, which is published as supporting information on the PNAS web site). The protein concentration used was 4.0 mg·ml^{-1} in 10 mM Hepes (pH 7.4)/50 mM $CaCl_2$, and the free energy of the folded and unfolded forms of each protein was determined by least-squares analysis of the data points from the transition region of the unfolding curves of the proteins.

Crystallization and Structure Solution. HylP1 was crystallized from 1 μl of HylP1 (at 10–20 mg·ml^{-1}) with 1–2 μl of the mother liquor (3.25 M sodium formate); 2-methyl-2,4-pentanediol was incorporated into the mother liquor to both improve crystal quality and the ease of freezing. An additional 15% 2-methyl-2,4-pentanediol was included in the stabilizing solution before flash freezing in liquid nitrogen. X-ray data were collected from single crystals at 120 K at the European Synchrotron Radiation Facility. Multiple anomalous difference (MAD) data were collected on beamline ID 14-EH4 and native data on ID 14-EH1. X-ray data were processed with MOSFLM (24), with subsequent computing using programs from the CCP4 suite (25) unless explicitly stated. Crystals belong to the rhombohedral space group H32 with the approximate cell dimensions of $a = b = 58.5$ Å, $c = 586.6$ Å. The structure was solved by using seleno-methionine multiple-wavelength data, with final phasing incorporating native data. Initial phasing used the seleno-methionine MAD data, through MLPHARE, with sites determined by using SHELXD (26). Final phasing incorporated native data, and phases were subsequently improved and extended to 1.8 Å by using DM (27) from the CCP4 suite. Five percent of the data were flagged for cross-validation, and subsequently REFMAC (28) in conjunction with ARPWARP (29) (both CCP4 suite) were used to build the sequence into the electron density automatically. QUANTA and X-FIT (Accelrys, Inc., San Diego) were used to make manual corrections to the model. Solvent molecules were added using X-SOLVATE routines of XFIT and checked manually. The final structure, refined at 1.8-Å resolution, has $R_{cryst} = 0.19$, $R_{free} = 0.23$ with deviations from stereochemical target values of 0.012 Å for bonds and 1.3° for angles (Table 1).

BIOPHYSICS

Table 1. X-ray data and structure quality statistics for the *S. pyogenes* SF370 HylP1

Statistic	HylP1 native	HylP1 MAD peak	HylP1 MAD remote	HylP1 MAD inflection
ESRF radiation source	ID 14-EH1	ID 14-EH4	ID 14-EH4	ID 14-EH4
Wavelength, Å	0.93400	0.97950	0.93930	0.97960
Resolution of data (outer shell), Å	195.52–1.80 (1.90–1.80)	43.44–2.50 (2.64–2.50)	50.64–2.50 (2.64–2.50)	196.27–2.50 (2.64–2.50)
R_{merge}*	0.097 (0.259)	0.062 (0.145)	0.064 (0.223)	0.055 (0.153)
Mean $I/\sigma I$ (outer shell)	14.0 (2.8)	18.3 (7.4)	19.4 (5.2)	19.7 (6.9)
Completeness, %	98.8 (98.8)	99.9 (99.9)	99.9 (99.9)	99.8 (99.8)
Multiplicity	5.5 (4.1)	5.2 (5.3)	5.7 (5.7)	5.0 (5.0)
No. protein atoms	2479			
No. ligand atoms	2 (Na^+ ions)			
No. solvent waters	267			
R_{cryst}†	0.19			
R_{free}	0.22			
rms deviation 1–2 bonds, Å	0.012			
Rms deviation 1–3 bonds, °	1.346			
Avg. protein B, Å^2	29			
Avg solvent B, Å^2	39			

Outer resolution bin statistics are given in parentheses. ESRF, European Synchrotron Radiation Facility.
*$R_{merge} = \Sigma_{hkl}\Sigma_i|I_{hkli} - \langle I_{hkli}\rangle|/\Sigma_{hkl}\Sigma_i\langle I_{hkli}\rangle$.
†$R_{cryst} = \Sigma||F_{obs}| - |F_{calc}||/\Sigma|F_{obs}|$.

Results and Discussion

Carbohydrate-active enzymes, notably glycoside hydrolases, carbohydrate esterases, glycosyltransferases, and polysaccharide lyases, have been grouped together into a large number of sequence-based families. There are almost 100 glycoside hydrolase families but merely 15 families of polysaccharide lyase, this discrepancy reflecting the absolute requirement for uronic-acid substrates for the latter enzyme class. Prior to this work, glycoside hydrolase CAZy (http://afmb.cnrs-mrs.fr/CAZY) family 69 (GH69) contained 26 ORFs, all from bacteriophage that infect streptococcal strains (predominantly *S. pyogenes* but also *S. equi*) that were predominantly annotated as hydrolytic hyaluronidases (EC 3.2.1.35). A feature of this enzyme family is its very high degree of sequence conservation

Phage Infection of *S. pyogenes* Involves the Action of an *Endo*-Acting Hyaluronan-Specific Lyase. *The S. pyogenes* M1 group A Streptococcus SF370 virulence factor hyaluronidase, HylP1, was expressed as a full-length 38.4-kDa N-terminally hexahistidine-tagged protein. HylP1 is active in the degradation of HA with a k_{cat} of 7.2 ± 0.9 s^{-1} and K_m = 1.47 ± 0.02 mg·ml^{-1}. Increase in absorbance at 232 nm is consistent with the action of a lyase (β-eliminase) whose action results in the production of a Δ4,5 unsaturated product (Fig. 1*c*). The enzyme is not a hydrolase, as implied by the enzyme commission number, but instead HylP1 [and its sequence homologs such as the *S. pyogenes* bacteriophage H4489A enzyme (17)] are polysaccharide lyases. Consequently, all the former GH69 enzymes have been reassigned to a new PL family, PL16.

In contrast to the unrelated hyaluronate lyases from group B streptococci, whose function is the degradation of the host-organism connective tissue to facilitate pathogen invasion (31–34), HylP1 is inactive on other connective tissue glycosaminoglycans such as chondroitin 4-sulfate and chondroitin 6-sulfate. The *in vivo* function of HylP1 is further revealed by product analysis. In stark contrast to the group B streptococcal enzymes, indeed to any previously characterized carbohydrate-active enzymes, the end products of HylP1 digestion of HA are long oligosaccharides, predominantly ΔHA_6 and ΔHA_8, with small amounts of ΔHA_4, ΔHA_{10} and longer (Fig. 2). HylP1 thus has an *endo* (random internal cut) mechanism of action with a greatly extended substrate-binding region, which requires at least four to six kinetically significant subsites on either side of the cleavage site. The function of HylP1 is therefore unlikely to be the complete degradation of connective tissue, or even HA itself. Instead, it would appear that the role of HylP1 is the introduction of widely spaced cuts in HA to facilitate a local reduction in capsule viscosity to aid phage invasion of the bacterial host.

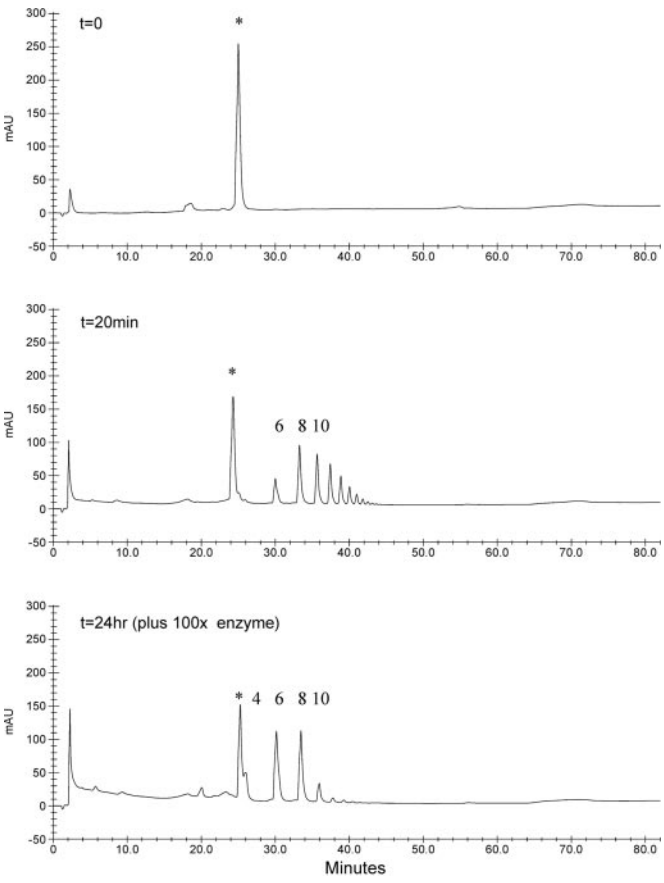

Fig. 2. Catalytic activity of *S. pyogenes* HylP1. High-performance anion exchange chromatography (see ref. 23) analysis of the products of a digestion of hyaluronic acid with HylP1 (20 min incubation with HylP1 at 2 μg·ml^{-1}). Identification of hyaluronan oligosaccharides, over a range of time points from 0 min to 24 h, was achieved by using electrospray mass spectrometry (42) with the degree of polymerization of product shown. *, artifact.

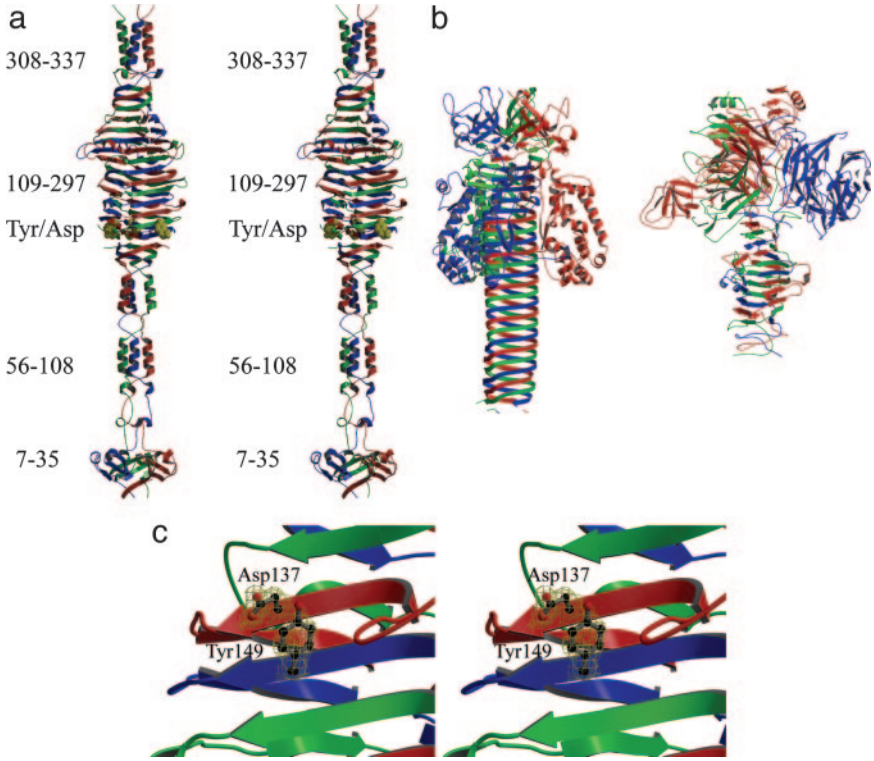

Fig. 3. The 3D structure of HylP1. (*a*) Divergent (wall-eyed) stereoview of HylP1 colored according to chain with the Asp-137/Tyr-149 couple shown in cpk representation (yellow). (*b*) The TSβHs of the bacteriophage T4 gp5 and gp12 proteins and the *E. coli* K1 bacteriophage poly(2,8-sialidase). HylP1 is unique in possessing catalytic activity within the TSβH itself, in marked contrast to the appended catalytic lysozyme and sialidase of the bacteriophage T4 and *E. coli* K1 bacteriophage proteins. (*c*) Close-up of the Asp-137/Tyr-149 pair whose mutation renders HylP1 inactive while still maintaining structural integrity. These images were drawn with MOLSCRIPT (43) and BOBSCRIPT (30).

The 3D Structure of HylP1 Reveals a Catalytically Active TSβH. The 3D structure of HylP1 was solved by using three-wavelength multiple-wavelength anomalous dispersion data from a selenomethionyl-derived protein, and harnessing native data to 1.8 Å, HylP1 crystallizes in the H32 (hexagonal setting) rhombohedral cell with unusual cell dimensions of $a = b = 58.5$ Å, $c = 586.6$ Å. The structure of HylP1 is composed of three intertwined polypeptides centered on a crystallographic threefold axis. The structure (Fig. 3*a*) can be traced from residue 6 through to the C terminus, with the sole exception of the loop from Ala-122 to Ser-128, which is disordered and which may (see below) also be implicated in catalysis. The N-terminal domain (7–35) forms a mixed globular α/β capping region that is followed by a coiled region and segmented α-helical coiled-coils to residue 108. The central "core" of HylP1 then follows, a single extended right-handed TSβH in the form of an irregular triangular tube, where each of the three faces of the tube is composed of alternating β-strands from each of the three polypeptides, with the β-strands orthogonal to the long axis of the enzyme. The faces of the strands are deeply concave, contributing to an extended polysaccharide-

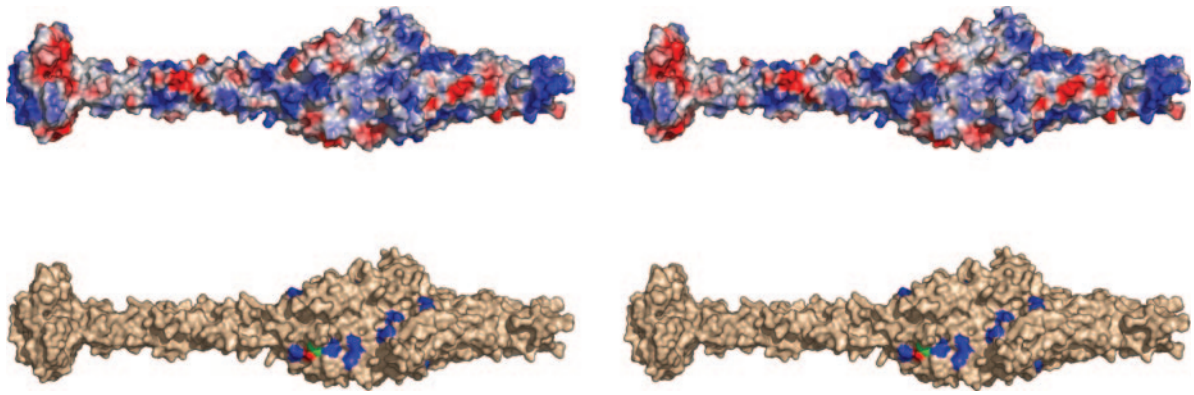

Fig. 4. The substrate-binding groove of HylP1. (*Upper*) Electrostatic surface potential of HylP1 (in divergent stereo) revealing an extended substrate binding cleft that is predominantly positively charged, complementing the negative charge of the substrate hyaluronic acid. (*Lower*) Complementarity with the alternating glucuronic acid moieties of the substrate along the putative substrate groove is provided by a ''barber's pole'' of positively charged residues (colored blue) including Arg-163, -190, -216, -277, and -279 and Lys-117, -119, -166, -176, and -226. Asp-137 and Tyr-149 are colored red and green, respectively. This figure was drawn with PYMOL (http://pymol.sourceforge.net).

binding surface (Fig. 4; discussed below). The TSβH of HylP1 is composed of residues Lys-108 through to Leu-281 and is $\approx$65 Å in length. The repeat on each side of the TSβH varies between 8 and 19 residues per turn. At the C terminus of the TSβH there is a short stretch from Ser-282 to Lys-297 where the β-sheets do not intertwine but interdigitate. The remainder of the 220-Å-long structure is composed of random coil and an α-helical "nose."

To our knowledge, the TSβH structural element has been observed in three other bacteriophage tail proteins: the T4 short tail fiber gp12 (35) and cell-puncturing device gp5 (36) and the core structure of the *E. coli* K1 bacteriophage K1F endo-2,8-sialidase (19). The latter two proteins also are involved in phage infection of the bacterial host through degradation of the cell-wall peptidoglycan and capsular polysialic acid, respectively. Threefold symmetry is thus a feature of tail-fiber proteins, manifesting itself in another such protein, tailspike gp9 of bacteriophage P22, except in this case not as a TSβH but as three single-stranded parallel β-helices (37). The width of each face of the triangular tube formed by the HylP1 TSβH varies from 49 Å at the widest point (Asn-274) to 20 Å at its narrowest. Thus, in contrast to gp5 and gp12, whose shape is regular and "smooth," the TSβH domain of HylP1 is both irregular and concave.

Beyond the simple topology, HylP1 is dissimilar to any reported protein structures (all DALI alignments have *Z*-scores of $\leq$1.2), and all primary sequence alignments of significance (*E*-values < 1.7) were also of phage-encoded hyaluronate lyases of *S. pyogenes* strains. This protein structure is therefore highly unique and highly specialized to *S. pyogenes* bacteriophage. The fact that there is no structural similarity between the TSβH domain of gp5 and gp12 with that of HylP1 and that the bacteriophage that encode these proteins, T4 and SF370.1, respectively, belong to different families suggests that the utilization of this structural element by tail-fiber proteins may be an example of convergent evolution. Indeed, in a topographical, as opposed to topological, sense, HylP1 is more similar to the single-stranded parallel β-helices of known carbohydrate-active enzymes such as several pectinases (for example, ref. 38), chondroitinase B (39), and tailspike gp9 (37). These domains are irregular, contain grooves, are catalytically active, and, in the case of gp9, also a component of a noncontractile phage tail.

Catalytic Center of HylP1. Perhaps the single most important difference between HylP1 and TSβH domains observed previously is that HylP1 catalysis requires no additional catalytic domains, such as the T4 lysozyme domain appended to gp12 and the bacteriophage K1F endosialidase (Fig. 3*b*). We have not obtained complex crystals of HylP1, but we are able to provide insight into the regions critical for catalysis. A triple-helical domain-truncated construct, from residues Asn-106–Asp-310, is catalytically active, confirming that the catalytic center resides in the TSβH domain itself. The TSβH domain provides an extended groove with which to bind long stretches of HA. Every three to four β-strands, positively charged Arg or Lys residues protrude into this grove, spaced between 9 and 14 Å apart, consistent with the spacing of the negatively charged glucuronic acid of the alternating GlcA-GlcNAc polymer that is HA (Fig. 4). Closer inspection of the structure of HylP1 revealed an Asp–Tyr pair (Fig. 3*c*), Asp-137, Tyr-149 (adjacent to a loop that is disordered in the native structure), reminiscent of similar pairings at the center of structurally unrelated hyaluronate and chondroitin lyases in which the Tyr residues act as the catalytic base for proton abstraction from the C5 of the uronic acid with the Asp involved in p*K*a maintenance (for example, ref. 40). Tyr-149 is positioned $\approx$20 and 60 Å from each end of the TSβH, consistent with the location of the active center determined through product analysis. Both D137A and Y149F mutants are completely inactive, whereas CD analysis indicated that there are no significant structural differences between the native and mutant proteins; indeed, the free energies of the folded and unfolded forms of each protein are also very similar at approximately -1 kcal·mol^{-1}. We therefore propose that Asp-137 and Tyr-149 have a catalytic function, but given the length of the binding groove, we cannot exclude a solely binding role for these residues.

The unique domain-structure of HylP1, viewed in light of the gp12 and gp9 structures, allows speculation on a possible orientation within the phage tail. The N-terminal globular domain of HylP1 resembles the globular collar domain of gp12 (41). The gp12 globular collar domain is adjacent to its receptor-binding domain, located at the C terminus of gp12. HylP1 has no such receptor-binding domain, because the globular domain of HylP1 is at the N terminus of the enzyme. This distinction is important because it does not rule out the possibility that there is another protein linked to HylP1 by means of the N-terminal globular domain, and hence the tail fiber of SF370.1 may not terminate with HylP1.

HylP1 is a previously undescribed catalytically active TSβH protein. The structure confirms its origins as part of a bacteriophage tail assembly whose function is as an "enzymatic needle" that digests the hyaluronic acid capsule during phage invasion of the *S. pyogenes* host. The role of HylP1 is not efficient hyaluronan degradation *per se* but efficient reduction of the viscosity of hyaluronan in the immediate vicinity of a single phage particle, to allow phage penetration to the cell wall. HylP1 has therefore recruited an extremely long structural protein whose extended binding surface provides exactly the unusual properties demanded for phage penetration of the capsule. The concave sides of the TSβH harness an exceedingly large number of subsites, which both prevent the enzyme from becoming "side-tracked" by complete degradation of the hyaluronic acid and help avoid product inhibition. It is particularly significant that the catalytic center resides within the TSβH. It had been assumed, at least tacitly, that the TSβH domain only provided a structural scaffold onto which catalytic domains were subsequently grafted, as illustrated by the bacteriophage T4 injection apparatus and the *E. coli* K1 phage-encoded sialidase. The HylP1 structure demonstrates that this assumption is not necessarily the case and suggests that catalytically active TSβHs will exist elsewhere. Given the increasing resistance to conventional antibiotics, we hope that insight into phage-host specificity and infection processes will have an impact on the subsequent design of novel bacteriophage-derived antibiotics.

We thank Ms. Meng Zhang for technical help with the mass spectrometry, and the European Synchrotron Radiation Facility for provision of beamlines. G.J.D. is a Royal Society University Research Fellow.

1. Cunningham, M. W. (2000) *Clin. Microbiol. Rev.* **13,** 470–511.
2. Frobisher, M. & Brown, J. (1927) *Bull. Johns Hopkins Hospital* **41,** 167–173.
3. Bohach, G. A., Fast, D. J., Nelson, R. D. & Schlievert, P. M. (1990) *Crit. Rev. Microbiol.* **17,** 251–272.
4. Weeks, C. R. & Ferretti, J. J. (1984) *Infect. Immun.* **46,** 531–536.
5. Broudy, T. B., Pancholi, V. & Fischetti, V. A. (2002) *Infect. Immun.* **70,** 2805–2811.
6. Hynes, W. L. & Ferretti, J. J. (1989) *Infect. Immun.* **57,** 533–539.
7. Canchaya, C., Fournous, G., Chibani-Chennoufi, S., Dillmann, M. L. & Brussow, H. (2003) *Curr. Opin. Microbiol.* **6,** 417–424.
8. Casjens, S. (2003) *Mol. Microbiol.* **49,** 277–300.
9. Nakagawa, I., Kurokawa, K., Yamashita, A., Nakata, M., Tomiyasu, Y., Okahashi, N., Kawabata, S., Yamazaki, K., Shiba, T., Yasunaga, T., *et al.* (2003) *Genome Res.* **13,** 1042–1055.
10. Broudy, T. B. & Fischetti, V. A. (2003) *Infect. Immun.* **71,** 3782–3786.
11. Cleary, P. P., LaPenta, D., Vessela, R., Lam, H. & Cue, D. (1998) *Infect. Immun.* **66,** 5592–5597.
12. Papageorgiou, A. C. & Acharya, K. R. (2000) *Trends Microbiol.* **8,** 369–375.
13. Marrack, P. & Kappler, J. (1990) *Science* **248,** 705–711.
14. Broudy, T. B., Pancholi, V. & Fischetti, V. A. (2001) *Infect. Immun.* **69,** 1440–1443.

15. Ferretti, J. J., McShan, W. M., Ajdic, D., Savic, D. J., Savic, G., Lyon, K., Primeaux, C., Sezate, S., Suvorov, A. N., Kenton, S., *et al.* (2001) *Proc. Natl. Acad. Sci. USA* **98,** 4658–4663.
16. Hynes, W. L., Hancock, L. & Ferretti, J. J. (1995) *Infect. Immun.* **63,** 3015–3020.
17. Baker, J. R., Dong, S. & Pritchard, D. G. (2002) *Biochem. J.* **365,** 317–322.
18. Rossmann, M. G., Mesyanzhinov, V. V., Arisaka, F. & Leiman, P. G. (2004) *Curr. Opin. Struct. Biol.* **14,** 171–180.
19. Stummeyer, K., Dickmanns, A., Muhlenhoff, M., Gerardy-Schahn, R. & Ficner, R. (2005) *Nat. Struct. Mol. Biol.* **12,** 90–96.
20. Bendtsen, J. D., Nielsen, H., von Heijne, G. & Brunak, S. (2004) *J. Mol. Biol.* **340,** 783–795.
21. Yamagata, T., Saito, H., Habuchi, O. & Suzuki, S. (1968) *J. Biol. Chem.* **243,** 1523–1535.
22. Brown, I. E., Mallen, M. H., Charnock, S. J., Davies, G. J. & Black, G. W. (2001) *Biochem. J.* **355,** 155–165.
23. Lauder, R. M., Huckerby, T. N. & Nieduszynski, I. A. (2000) *Glycobiology* **10,** 393–401.
24. Leslie, A. G. W. (1992) in *Joint CCP4 and ESF-EACMB Newsletter on Protein Crystallography*, eds. Wolf, W. M. and Wilson, K. S. (Daresbury Lab., Warrington, U.K.), Vol. 26.
25. Collaborative Computational Project, Number 4 (1994) *Acta Crystallogr. D* **50,** 760–763.
26. Schneider, T. R. & Sheldrick, G. M. (2002) *Acta Crystallogr. D* **58,** 1772–1779.
27. Cowtan, K. D. & Main, P. (1996) *Acta Crystallogr. D* **49,** 148–157.
28. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997) *Acta Crystallogr. D* **53,** 240–255.
29. Perrakis, A., Morris, R. & Lamzin, V. S. (1999) *Nat. Struct. Biol.* **6,** 458–463.
30. Esnouf, R. M. (1997) *J. Mol. Graphics. Model* **15,** 132–134.
31. Baker, J. R. & Pritchard, D. G. (2000) *Biochem. J.* **348,** 465–471.
32. Shain, H., Homer, K. A. & Beighton, D. (1996) *J. Med. Microbiol.* **44,** 372–380.
33. Pritchard, D. G., Lin, B., Willingham, T. R. & Baker, J. R. (1994) *Arch. Biochem. Biophys.* **315,** 431–437.
34. Nukui, M., Taylor, K. B., McPherson, D. T., Shigenaga, M. K. & Jedrzejas, M. J. (2003) *J. Biol. Chem.* **278,** 3079–3088.
35. van Raaij, M. J., Schoehn, G., Burda, M. R. & Miller, S. (2001) *J. Mol. Biol.* **314,** 1137–1146.
36. Kanamaru, S., Leiman, P. G., Kostyuchenko, V. A., Chipman, P. R., Mesyanzhinov, V. V., Arisaka, F. & Rossmann, M. G. (2002) *Nature* **415,** 553–557.
37. Steinbacher, S., Seckler, R., Miller, S., Steipe, B., Huber, R. & Reinemer, P. (1994) *Science* **265,** 383–386.
38. Yoder, M. D., Keen, N. T. & Jurnak, F. (1993) *Science* **260,** 1503–1507.
39. Huang, W., Matte, A., Li, Y., Kim, Y. S., Linhardt, R. J., Su, H. & Cygler, M. (1999) *J. Mol. Biol.* **294,** 1257–1269.
40. Lunin, V. V., Li, Y., Linhardt, R. J., Miyazono, H., Kyogashima, M., Kaneko, T., Bell, A. W. & Cygler, M. (2004) *J. Mol. Biol.* **337,** 367–386.
41. Thomassen, E., Gielen, G., Schutz, M., Schoehn, G., Abrahams, J. P., Miller, S. & van Raaij, M. J. (2003) *J. Mol. Biol.* **331,** 361–373.
42. Price, K. N., Tuinman, A., Baker, D. C., Chisena, C. & Cysyk, R. L. (1997) *Carbohydrate Res.* **303,** 303–311.
43. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24,** 946–950.