Biophysical Characterization and Activity of Lymphostatin, a Multifunctional Virulence Factor of Attaching and Effacing Escherichia coli

Received for publication, December 10, 2015, and in revised form, January 11, 2016. Published, JBC Papers in Press, January 19, 2016, DOI 10.1074/jbc.M115.709600

Robin L. Cassady-Cain,1 Elizabeth A. Blackburn,1 Husam Alsarra2, Emil Dedic,2 Andrew G. Bease,3 Bettina Böttcher, René Jørgensen, Martin Wear, and Mark P. Stevens

From the 1Roslin Institute, Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Easter Bush, Midlothian EH25 9RG, United Kingdom, the 2Centre for Translational and Chemical Biology (CTCB), University of Edinburgh, Michael Swann Building, King’s Buildings, Edinburgh EH9 3BF, United Kingdom, and the 3Department of Microbiology & Infection Control, Statens Serum Institut, Artillevjel 5, DK-2300 Copenhagen S, Denmark

Attaching and effacing Escherichia coli cause diarrhea and typically produce lymphostatin (LifA), an inhibitor of mitogen-activated proliferation of lymphocytes and pro-inflammatory cytokine synthesis. A near-identical factor (Efa1) has been reported to mediate adherence of E. coli to epithelial cells. An amino-terminal region of LifA shares homology with the catalytic domain of the large clostridial toxins, which are retaining glycosyltransferases with a DXD motif involved in binding of a metal ion. Understanding the mode(s) of action of lymphostatin has been constrained by difficulties obtaining a stably transformed plasmid expression clone. We constructed a tightly inducible clone of enteropathogenic E. coli O127:H6 lifA for affinity purification of lymphostatin. The purified protein inhibited mitogen-activated proliferation of bovine T lymphocytes in the femtomolar range. It is a monomer in solution and the molecular envelope was determined using both transmission electron microscopy and small-angle x-ray scattering. Domain architecture was further studied by limited proteolysis. The largest proteolytic fragment containing the putative glycosyltransferase domain was tested in isolation for activity against T cells, and was not sufficient for activity. Tryptophan fluorescence studies indicated that lymphostatin binds uridine diphosphate-N-acetylglucosamine (UDP-GlcNAc) but not UDP-glucose (UDP-Glc). Substitution of the predicted DXD glycosyltransferase motif with alanine residues abolished UDP-GlcNAc binding and lymphostatin activity, although other biophysical properties were unchanged. The data indicate that lymphostatin has UDP-sugar binding potential that is critical for activity, and is a major leap toward identifying the nature and consequences of modifications of host cell factors.

Enterohemorrhagic Escherichia coli (EHEC) and enteropathogenic E. coli (EPEC) are enteric bacterial pathogens of worldwide importance. Both pathotypes colonize intestinal mucosa via the formation of “attaching and effacing” lesions in a manner that requires a Type III protein secretion system, as well as accessory virulence factors including surface adhesins (1). Lymphostatin (also known as LifA) is a chromosomally encoded protein with a predicted molecular mass of 365 kDa produced by most EPEC and non-O157 EHEC strains (2). We have previously demonstrated that lymphostatin is required for intestinal colonization of calves by non-O157 EHEC serogroups O5, O111 (3), and O26 (4), and it also promotes colonization of the murine intestines and colonic hyperplasia by the attaching and effacing pathogen Citrobacter rodentium (5). Chlamydia species also contain a family of proteins that have homology to lymphostatin, and which have been implied to act as cytotoxins (6).

Lymphostatin was first described as the factor required for inhibition of mitogen-activated proliferation of lymphocytes by enteropathogenic E. coli O127:H6 lysates (2). This activity has been demonstrated against lymphocytes from peripheral blood and the intestines (3, 7) and is not associated with direct cytotoxicity. Peripheral blood mononuclear cells from mice, cattle, and humans are sensitive to lymphostatin (2, 3, 5). Lymphostatin has also been reported to inhibit the production of pro-inflammatory cytokines including IL-2, -4, -5, and interferon-γ (7), and it has been suggested that it may therefore interfere with the induction of innate and adaptive immune responses. In the same year as LifA was described in EPEC, a near identical factor was described in EHEC O111:H− that was associated with bacterial adherence to cultured epithelial cells. The authors named the factor EHEC factor for adherence 1 (Efa1), however, it has 97.4% amino acid identity to lymphostatin, and they are likely equivalent proteins (8). Although a direct role of
Lymphostatin Characterization and Activity

Efa1 in adherence has been reported using rabbit EPEC (9), mutations in some strains impair expression and secretion of Type III secreted proteins required for attaching and effacing lesion formation (3). Furthermore, it has recently also been reported that lymphostatin can be secreted via the type III secretion system, but its functions once injected into host cells are unknown (10).

Understanding of the mode of action of lymphostatin has been constrained by the instability of plasmid clones and difficulties in obtaining full-length purified protein (2). Furthermore, even plasmid-driven soluble expression of smaller fragments of lymphostatin has proven to be difficult (11). Bioinformatic analysis has identified homology between the amino-terminal of LifA/Efa1 and the catalytic glycosyltransferase domain of the LCTs (2, 8). These clostridial cytotoxic molecules are large proteins whose catalytic domain glycosylates Rho-family GTPases that regulate the actin network (12). They are retaining enzymes with a GT-A fold, which belong to glycosyltransferase family 44 and are characterized by having a DXD (Asp-Xaa-Asp) motif involved in metal ion binding together with the diphosphate moiety of the UDP-sugar donor (13). It is a multistep reaction dependent on the DXD motif, where sugar binding and hydrolysis is followed by transfer of the released sugar to an acceptor protein (reviewed in Ref. 12).

In addition, recent studies show that attaching and effacing E. coli express another novel protein, NleB, which is an effector glycosyltransferase injected into host cells upon infection. NleB blocks death receptor-induced apoptosis and promotes intestinal colonization (14, 15), as part of a suite of E. coli effectors that influence NF-kB signaling in mammalian cells (reviewed in Ref. 16). NleB uses uridine diphosphate N-acetylglucosamine (UDP-GlcNAc) as its donor sugar and covalently N-links a GlcNAc to a conserved arginine residue in death domain containing adaptor proteins (14, 15). This modification had not previously been described and most GlcNAc addition to proteins occurs via O-linkage on serine/threonine residues (14, 15). Apart from the glycosyltransferase domain, searches also identified a Versinia YopT-like cysteine protease (CP) motif in the sequence of lymphostatin (17). These features represent a small portion of the primary sequence of lymphostatin, and are restricted to the N-terminal third of the protein. Although one report claimed that deletion of the predicted glycosyltransferase and cysteine protease motifs attenuated C. rodentium in mice (5), close inspection reveals that stop codons were introduced that resulted in protein truncation at the deletion site rather than in-frame mutations, making the results difficult to interpret (4). Given its large size and the paucity in understanding how lymphostatin is able to carry out its activities, we sought to produce a full-length recombinant lymphostatin and characterize its structural and biophysical features, as well as its effects on activated T cells. Here we show that full-length lymphostatin can be expressed as a folded, biologically active recombinant protein that inhibits bovine T cell proliferation at very low concentrations. Furthermore, like other glycosyltransferases, it has sugar-binding potential, and both its biological activity and sugar binding is dependent on a DXD motif, which has been implicated in the activity of glycosyltransferases. Although the activity of lymphostatin was identified more than a decade ago, this is the first significant step forward in understanding the mechanisms underlying its intrinsic activities.

Experimental Procedures

Cloning—The full-length gene encoding lymphostatin (~9.6 kb) was cloned using the commercially available Expresso Rhamnose cloning and expression system (Lucigen Inc.) from the prototype enteropathogenic E. coli serotype O127:H6 strain E2348/69. The lymphostatin gene was amplified using genomic DNA from E2348/69 as lymphostatin activity was first described in this strain (2). The pRham vector incorporates a C-terminal in-frame His6 tag. Amplicons were generated using the primers EXPLiFA_FOR, 5’-GAAGAGATATACTATATGAGACTGCGAGAGAAAAATCTTTT-3’ and EXPLiFA_REV, 5’-GTGATGGTGATGATGATGGTTAAAGGTTGTCC-ATT-3’ with PHUSION proofreading polymerase (Thermo Scientific). Amplicons of the appropriate size were isolated by agarose gel electrophoresis and purified using GeneClean II (MP Biomedicals, Inc.). The amplicons were cloned into pre-linearized pRham vector by homologous recombination in E. coli 10G® (Lucigen Inc.) chemically competent cells. Subsequent transformants were screened by colony PCR using primers supplied by the manufacturer: pRham Forward, 5’-GTCTTTTTTAGACTGTCGTAGGGA-3’ and pETe Reverse, 5’-CTCAAGCCCCCTTATAGGCC-3’. The sequence of two independent clones (pRHAM-LiFA-6xH) were confirmed to be identical to the published liFA sequence of E2348/69 (gene E2348C_3234; (18)) by full-length Sanger sequencing on both strands using primers every 500 bp (GATC-Biotech).

Generation of a DTD to AAA Lymphostatin Substitution Mutant—To remove the DXD motif the amino acids DTD at position 557–559 were substituted with AAA using the QuikChange II XL site-directed mutagenesis kit (Agilent Technologies) according to the manufacturer’s directions. Briefly, using the pRHAM-LiFA-6xH plasmid as a template, the primers LiFA-DXD-1 (Forward), GGATGTATATCCTTTAAGAGACATGGTGTATTTATACACCGCGCGCCGATGATGCGCTGAC-ATACTTAAACAGGATATTATTAAAAA, and LiFA-DXD-2 (Reverse), TTCTAAAAACTCTGTTATGATTGCACGTCATCGCAGCTCGCCCGCGCGCTGTATAAAATCCACCTGCTCTTTTGAAGATACCATCC, were used to introduce AAA to the sequence, by total replication of the plasmid, followed by digestion of the parent plasmid. A NotI restriction endonuclease site (indicated in bold) was also introduced at the site of mutation. Putative mutant plasmids were screened by restriction digest with NotI, and verified by Sanger sequencing on both strands (GATC-Biotech).

Recombinant Lymphostatin Expression and Purification—Recombinant His-tagged lymphostatin (rLiFA) was overexpressed in E. coli® cells cultured in lysogeny broth at 37 °C, 250 rpm shaking to A600 nm, 0.8. Expression of the protein was induced by the addition of l-rhamnose to 0.2% (w/v) and cultured for a further 3 h at 30 °C. Cells were pelleted by centrifugation, re-suspended in 20 mM sodium phosphate (NaH2PO4), pH 7.6, 300 mM sodium chloride, 500 mM non-detergent sulfobetaine (NDSB201), 20 mM imidazole, 5% (v/v) glycerol, 1 mM dithiothreitol (DTT), 100 μM phenylmethylsulfonyl fluoride, 1
complete protease inhibitor tablet/3 g cell mass (Roche), 0.1% (v/v) Tween 20 and lysed by high pressure lysis by single passage, at 30 kpsi, through a Constant Systems TS 1.1 kW Benchtop Cell disruptor. All chromatography was performed on an ÄKTA Explorer 10 UV900 LC system (GE Healthcare) at 4 °C. The lysate was clarified by centrifugation (50,000 × g at 4 °C) and loaded onto a Ni²⁺ ion-metal affinity chromatography (IMAC) column (HisTrap FF; GE Healthcare) pre-equilibrated in 20 mM sodium phosphate, pH 7.6, 300 mM sodium chloride, 20 mM imidazole, 5% (v/v) glycerol, 1 mM DTT, 0.1% (v/v) Tween 20 at 2 ml/min. This was followed by washing with 15 column volumes (cv) of the same buffer (Buffer A) and 15 cv of Buffer A supplemented with 4% Buffer B (20 mM sodium phosphate, pH 7.6, 300 mM sodium chloride, 500 mM imidazole, 5% (v/v) glycerol, 1 mM DTT, 0.1% (v/v) Tween 20). rLifA was eluted by increasing the concentration of imidazole up to 500 mM over 5 cv (gradient over 5 cv); all at 2 ml/min. rLifA was passed over a size exclusion column pre-equilibrated in 20 mM sodium phosphate, pH 7.6, 300 mM sodium chloride, 5% (v/v) glycerol, 1 mM DTT, 0.1% (v/v) Tween 20 to separate low molecular weight contaminants (Superose-6pg XK16/60; GE Healthcare) at 1 ml/min. Fractions containing rLifA were buffer exchanged into 15 mM sodium phosphate, pH 7.6, 50 mM sodium chloride, 5% (v/v) glycerol, 1 mM DTT, 0.05% (v/v) Tween 20 at 8 ml/min and further purified to homogeneity by anion exchange (HiPrep desalt 26/10; Mono-Q 5/50 GL; GE Healthcare) at 1 ml/min. The protein was eluted by running a gradient from 15 to 500 mM NaCl over 30 cv; the protein eluted at a salt concentration of 160 mM. AAA mutant full-length lymphostatin (rLifA<sup>DTDD/AAA</sup>) was purified using the same strategy as rLifA protein, and the chromatograms were indistinguishable. The induction, size, and stability of expressed proteins were assessed by 3–8% Tris acetate SDS-PAGE and/or by Western blotting with monoclonal antibody specific for His<sub>6</sub>-tagged proteins. The identity of full-length recombinant lymphostatin was confirmed by in-gel protein digest and peptide analysis. Excised gel-bands were incubated at a porcine trypsin:lymphostatin ratio of ∼1:30, in 50 mM ammonium bicarbonate overnight at 32 °C (Promega). Peptides were identified by matrix-assisted laser desorption ionization (MALDI) mass spectroscopy on a Voyager DE-STR MALDI-TOF mass spectrometer (Applied Biosystems) using an α-cyano-4-hydroxycinnamic acid matrix. The spectral data were processed using Data Explorer software and the Mascot NCBInr database searched against the peptide mass map (Matrix Science). To investigate the domain structure of lymphostatin, purified protein was incubated with trypsin at a ratio of 375:1, at 21 °C, to give limited digestion. Aliquots were removed at 1, 2, 3, and 4 h and the reaction stopped by boiling samples adjusted with 2 mM EDTA and 2 mM PMSF in SDS-PAGE loading buffer. Digest products were separated by SDS-PAGE and individual bands were subjected to in-gel tryptic digestion and MALDI-TOF mass spectroscopy as described above. Peptide masses were compared with the sequence of full-length rLifA using GPMAW 9.2 software, mass tolerance 50 ppm (19). Fragment F1 was purified to homogeneity from other digest products by
Lymphostatin Characterization and Activity

ion-exchange chromatography (Mono-Q 5/50 GL; GE Healthcare) as described above.

Size Exclusion Chromatography-Multi-angled Light Scattering (SEC-MALS)—Size exclusion chromatography coupled to UV, static light scattering, and refractive index detection were used to determine the molecular mass of pure rLifa in solution and to estimate the detergent load (Viscotec SEC-MALS 20 and Viscotec RI Detector:VE3580; Malvern Instruments). One hundred μl of 1 mg/ml of lymphostatin was passed over a size exclusion column pre-equilibrated in 20 mM sodium phosphate, pH 7.6, 150 mM sodium chloride, 1 mM DTT, 0.1% Tween 20 (Superose6 10/300 GL, GE Healthcare) at 24 °C. Light scattering, refractive index (RI), and Tween 20 (Superose6 10/300 GL, GE Healthcare) at 24 °C. Light scattering, refractive index (RI), and detergent load (Viscotec SEC-MALS 20 and Viscotec RI Detector:VE3580; Malvern Instruments). One hundred μl of 1 mg/ml of lymphostatin was passed over a size exclusion column pre-equilibrated in 20 mM sodium phosphate, pH 7.6, 150 mM sodium chloride, 1 mM DTT, 0.1% Tween 20 (Superose6 10/300 GL, GE Healthcare) at 24 °C. Light scattering, refractive index (RI), and 0.1% Tween 20 (Superose6 10/300 GL, GE Healthcare) at 24 °C. Light scattering, refractive index (RI), and 0.1% Tween 20 detergent, 0.97 and 0.01 absorbance unit, ml/mg, respectively; RI for protein, 0.187 ml/g; and Tween 20, 0.145 ml/g (Malvern Instrument software).

Circular Dichroism (CD)—The far UV CD spectrum of full-length lymphostatin (0.11 μm), rLifa ΔTD/A (0.11 μm), and the digest fragment F1 (0.35 μm) were recorded at 10 nm/min; data pitch, 0.1 nm; response time, 2 s between 185 and 285 nm in a 0.1-cm path length quartz cuvette at 25 °C (JASCO-810 spectrometer). The proteins were exchanged into 10 mM sodium phosphate, pH 7.6, 150 mM sodium fluoride prior to analysis (HiTrap desalt column, GE) at 4 ml/min. Spectra were corrected by subtracting a buffer baseline, each an average of 5 spectra. Secondary structure was estimated using the Dichroweb CD secondary structure analysis server (20) including the methods CONTIN, SELCON3, and CDSSTR (21–24) and reference data sets SP175 and 7 (25).

Bioinformatic Analysis—The secondary structure of lymphostatin was predicted from sequence using PredictProtein (26) and PSIPRED (27). Proteins with similar structural elements and homologues were identified with PHYRE (28) and BLASTp (29).

Intrinsic Tryptophan Fluorescence—Binding of uridine diphosphate-glucose (UDP-Glc) and UDP-GlcNAc to wild-type LifA and LifAΔTD/A was determined by ligand-induced changes in intrinsic tryptophan fluorescence. Fluorescence measurements were performed on a SPEX Fluoromax 3 spectrometer (Horiba) in a 3-ml stirred cuvette by titrating the UDP-sugar into 0.2 μm lymphostatin at 20 °C in 20 mM sodium phosphate, pH 7.6, 150 mM sodium chloride, 5% glycerol, 0.1% Tween 20, 1 mM DTT. Samples were allowed to equilibrate for 5 min after the addition of each aliquot. The final volume added did not exceed 2% of the initial volume. Tryptophan was excited at 295 nm and emission spectra were recorded from 310 to 400 nm, with a 1-nm interval. Excitation and emission slits were set at 5 nm, with an integration time of 1 s. The equilibrium dissociation constant was obtained from fitting the fluorescence intensity at 340 nm to a binding model corrected for collisional quenching (Kaleidagraph, Synergy Software).

Negative Staining and Electron Microscopy—Four hundred mesh carbon-coated copper grids were glow discharged for 1 min in a Quorum Tech sputter coater with a current of 25 μA and used within 1 h. rLifa (4 μl at ~8 μg/ml concentration) was applied to a glow discharged grid and incubated for 1–2 min. Next, the grid was washed with 2 drops of water and 2 drops of 2% uranyl acetate followed by staining with a further drop of 2% uranyl acetate for 4–5 min. Finally, excess liquid was blotted from the edge of the grid with filter paper (Whatman No. 5). Dried grids were imaged with an FEI F20 field emission gun electron microscope equipped with an 8k x 8k TVIPS CMOS camera (F816). The camera format was binned by a factor of 2 giving a calibrated binned pixel size of 3.06 Å/pixel at the specimen level. Micrographs were acquired semi-automatically with EM-tools (TVIPS GmbH) under low dose conditions (20 e/Å², at 200 kV).

EM Image Processing—The defocus of each micrograph was determined with ctffind3 (30). Particles were selected semi-automatically using e2boxer (31). Selected particle images were extracted and normalized with RELION (32), with a box size of 128 pixels. The radius for background normalization was 55 pixels. Extracted particle images were classified into 100 two-dimensional classes using RELION. Particle images, which did not align with an accuracy of better than 4 degrees in the two-dimensional classification were excluded from further processing. This reduced the number of particle images in the data set from 27,431 to 25,244. The relative spatial orientations of two-dimensional class averages of the remaining particle images were determined by sinogram correlation in IMAGIC (33). After determining the relative orientations of the class averages a three-dimensional map was calculated by weighted back projection. The orientations of the class averages were further refined by angular reconstitution using projections of the three-dimensional map as anchor set followed by calculating an improved three-dimensional map by back-projection. The resulting three-dimensional map was used as reference in Relion for auotrefinement of the data set. The resolution was estimated by Fourier-Shell correlation between two independently processed data sets (34) and was 23 Å at a Fourier-Shell correlation of 0.14 (35). The absolute hand of the map is unknown.

Small-angle X-ray Scattering (SAXS)—Synchrotron radiation data were collected at the 1911-4 SAXS beamline at the MAX-II Laboratory (Lund, Sweden) as 4 x 30-s exposures of a 20–30-μl sample and scattering profiles were compared with detect radiation damage. Lymphostatin was buffer exchanged into 20 mM NaH₂PO₄, pH 7.6, 300 mM NaCl, 2 mM DTT, and 5% (v/v) glycerol (an identical buffer to the size-exclusion storage buffer but without Tween 20) prior to SAXS measurements using Amicon Ultra 0.5-ml centrifugal filters (Merck). Data were collected at 0.91-Å wavelength at 10 °C with a Dectris hybrid pixel Pilatus 1M detector. To detect concentration-dependent inter-particle effects, measurements were collected at multiple protein concentrations in the range of 0.67–1.6 mg/ml, and the 1 mg/ml scattering curve was used for all subsequent analysis. Background buffer scattering was subtracted using PRIMUS (36), part of the ATSAS package (37). Pair distance distribution function P(r) and the maximum particle dimension Dmax were computed using GNOM (38). The Porod volume was calculated using ATSAS AUTOPOROD (39) and used for molecular weight estimation. Ab initio shape envelope was developed using 10 independent DAMMIF (40) computations in P1 symmetry. DAMAVER (41) was used to align and
compare the resulting models. The most representative model was subsequently refined using DAMMIN (42). The resulting bead model was used to compute a surface envelope with Situs pdb2vol (43). The DAMMIN model was manually aligned with the EM density map using USCF Chimera (44).

Statistical Analysis of ED$_{50}$ of rLifA and rLifADTD/AAA—The effective dose 50 (ED$_{50}$) for rLifA and rLifADTD/AAA was determined using drc in R (45). One way analysis of variance was used to determine statistically significant differences between both ED$_{50}$ and dose-response curves using Minitab (46), with $p$ values ≤0.05 taken to be significant. Post hoc Tukey test was used to calculate 95% confidence intervals and confirm significance (46).

Results

Full-length Lymphostatin Can Be Produced as a Stable, Tagged Recombinant Protein—Published studies have noted instability of lymphostatin clones (2, 47) or inability to assemble the full-length gene from amplicons (8), possibly because of toxicity and gene size. To overcome this, an amplicon of the EPEC O127:H6 E2348/69 full-length lifA gene (9.6 kb) was cloned in pRham with a carboxyl-terminal histidine tag. In this Expresso Rhamnose cloning system, tight control of expression was achieved using a combination of glucose to repress transcription from the rhamnose-inducible promoter and a 3-h induction with 0.2% (v/v) L-rhamnose. Purification was optimized as described under “Experimental Procedures,” and Fig. 1, A–C, illustrates the steps in the purification strategy. No evidence of lymphostatin production was detected in uninduced cultures of *E. coli* harboring sequence-verified lifA, however, L-rhamnose induction led to the expression of rLifA, as detected by Coomassie staining and Western blotting with an anti-His tag antibody (Fig. 1, inset).

To confirm that the expressed protein is full-length lymphostatin, in-gel tryptic digestion and MALDI mass spectrometry were carried out. The resulting peptides aligned to 41% of the lymphostatin primary sequence, from residue 52 to 3174 (of 3229); representing good coverage for such a large protein. The predicted N terminus of the protein is relatively rich in basic residues. Lack of tryptic peptide coverage at the N terminus could be accounted for by extensive digestion at charged residues generating many low mass fragments that are not detectable by MALDI. The C-terminal end can be inferred to be intact because of detection of the histidine tag and by IMAC purification and Western blotting. The purified protein was analyzed by dynamic light scattering. The sample contained species with a Stokes radius of 6.5 nm, consistent with monomeric rLifA. There was no evidence of aggregation or insolubility (data not shown).

rLifA Is Biologically Active against Bovine T Cells—Lymphostatin-containing bacterial lysates have been reported to be
Lymphostatin Characterization and Activity

capable of inhibiting the proliferation of mitogen-stimulated PBMCs using cells from humans, mice (2, 7), and cattle (3, 11). Using an enriched bovine T cell population increased the signal-to-noise ratio compared with use of bulk PBMCs, giving a wider dynamic range to the assay. Using T cells obtained from four independent donors, purified rLifA inhibited ConA-stimulated proliferation, showing inhibition from the low femtomolar range with concentration-dependent titration of activity with a sigmoidal curve (Fig. 2). Measurements of the ConA-stimulated response were typically 2–5-fold higher than cells alone. The carrier buffer for rLifA was determined to have no effect on ConA stimulation of cells on its own. Furthermore, a colorimetric assay to detect release of cytosolic lactate dehydrogenase found no evidence of cell lysis when primary T cells were treated with inhibitory concentrations of rLifA, indicating that the inhibitory effect of lymphostatin on T cells is unlikely to be a consequence of direct cytotoxicity.

Lymphostatin Is a Monomer in Solution—Having established that rLifA is biologically active, we examined its structural characteristics. SEC-MALS was used to calculate the average molecular weight of rLifA in solution and hence the oligomeric assembly. Lymphostatin eluted as a well resolved single peak. Light scattering, RI, and UV were analyzed by a conjugate protein model, and the mass of lymphostatin was predicted to be 330 ± 15 kDa, slightly lower than the predicted molecular mass for monomeric protein of 365 kDa (Fig. 1D). Given the mapping of the N-terminal to at least 52 residues from the beginning of the putative start of the protein, and that the C-terminal is intact, the size differential to the calculated size is unlikely to be due to truncated protein, and is within acceptable error for the technique at the present time, taking into account the glycerol and detergent-rich buffer. Detergent micelle eluted later than protein with an estimated mass of 76 ± 5 kDa (Fig. 1D). Taken together the data indicate that lymphostatin is a monomer in solution and that it is not necessary for the protein to fully or partially be enclosed within a micelle to remain in solution.

Limited Proteolysis Identified 3 Putative Structural Domains—With no fine structural information available on lymphostatin we set out to predict putative physical domains using limited proteolysis; the premise being that loops and flexible regions of the protein are more susceptible to digestion than buried residues. Limited tryptic proteolysis of rLifA over several hours revealed a defined and consistent pattern of cleavage products. Five major fragments were identified by mass spectrometry, F1–F5 (Fig. 3A). Mass spectrometry of tryptic peptides from each species and alignment of peptides against/with the lymphostatin primary sequence suggested 3 major domains, with the additional two species representing truncated versions of two of the major digestion products, as indicated in Fig. 3B. Identification of peptides by MALDI-TOF mass spectrometry from the in-gel tryptic digestion of fragments initially generated by limited proteolysis does not precisely identify protein domain boundaries but may be considered a useful guide. Limited proteolysis fragments are likely to be somewhat longer at both the N and C terminus than the most N-terminal and C-terminal residues identified from their tryptic digestion and MS analysis. The largest intact fragment, designated F1, represents a ~1435 amino acid (aa) region at the N terminus of

![FIGURE 2. Concentration-dependent inhibition of T cell proliferation by full-length rLifA. Effect of concentration of lymphostatin against ConA-stimulated peripheral bovine T cells. Data shown are the average of 4 independent donors, done in triplicate. Error bars indicate S.E. The purity of the cells used was ~85% in each experiment, as measured by CD3 staining using flow cytometry. Calculated ED_50 was 25 ± 4.6 pg/ml (68 fm); 100 ng/ml is equivalent to ~273 pm. Concentrations on the x axis are displayed on a log scale.](image)

![FIGURE 3. Domain organization of lymphostatin. A, limited proteolysis of lymphostatin gave a consistent digest pattern. Five fragments, F1–F5, were identified by mass spectrometry. B, analysis of the primary amino acid sequence of lymphostatin predicts a glycosyltransferase (GT) and cysteine-type endopeptidase domain (CP). Black bar illustrates full-length protein (FL). Fragment F1 contains the putative GT and CP domains. Fragment F2 maps more closely to the GT domain with a C-terminal extension of ~140 aa.](image)
Lymphostatin Characterization and Activity

Lymphostatin, starting about 177 aa from the N terminus, and encompassing both the putative glycosyltransferase (GT) and cysteine protease (CP) motifs. The two smaller fragments are separated from the N-terminal fragment by about 500 aa, and are approximately ~776 aa (F3) and 271 aa (F5), respectively. In addition, the two C-terminal proximal fragments are separated by a short stretch of 43 aa. Residues 1600–2100 are not represented in the major early digest products but were very well represented in the full tryptic digest during mass spectrometry analysis, suggesting this region is intact in the full-length protein. The abundance of fragment F1 reduces as the digest time increases, whereas fragment F2, an N and C terminally truncated subfragment of F1, becomes relatively more abundant. Fragment F2 maps more closely to the GT domain with a C-terminal extension of ~140 aa. The N-terminal ~100 aa of F1 is richer in basic amino acids than the full-length lymphostatin. This property enabled us to separate F1 from other digest products by anion exchange chromatography.

The Predicted Glycosyltransferase Domain of Lymphostatin—To identify the boundaries of the glycosyltransferase domain of lymphostatin we performed BLAST analysis of the primary sequence. We identified homology with the N termini of LCTs (the first 541 amino acids of TcdA) encompassing their catalytic glycosyltransferase domains. The alignment indicates that the sequence of lymphostatin is most similar to that of the LCTs for residues that define the binding site of the glycosyl donor substrate, often described at the catalytic core of the GT domain (Fig. 4A, gray shading) (13). The catalytic core residues of the LCTs are not contiguous and have insertions that form helical bundles surrounding the core (Fig. 4B, Toxin A, orange highlights). These are thought to confer specificity for the protein that becomes glycosylated by the glycosyltransferase. The most striking differences in sequences in the LCT GT domains are within these insertions. To further examine the potential of lymphostatin to form a GT fold we generated a model using the PHyre2 server (49). The best scoring model is based on Toxin A from Clostridium difficile, which forms a GT-A-fold common to all LCTs (Fig. 4C; model of lymphostatin Ala242-Arg769). Lymphostatin has similar insertions that are predicted from the sequence to be coiled and helical in character. It is worth noting that differences to the helical bundles mean that even if an alignment is restricted to the catalytic core of the clostridial toxins (residues Thr105-Trp534 in TcdA, which align to residues Thr499-Phe860 in lymphostatin) the sequence identity is only 20%, a figure that belies their common features.

Lymphostatin displays good conservation for residues that make key non-covalent interactions with the UDP-sugar (Fig. 4A, stars above sequence; bold typeface marks identity). The LCTs and lymphostatin contain the signature DxD motif, seen in most GT domains, through which the aspartate carboxylates coordinate a divalent cation and the donor substrate (Fig. 4A, red box). Three residues at the base of the β-hairpin in the sugar donor binding site, XNX, are thought to confer sugar specificity (Figs. 4A, black box, and 5, A and C). Toxins A, B, and L all have INQ in this position; this means that the enzymes can accommodate UDP-Glc but not UDP-GlcNAc. α-Toxin and TpeL bind UDP-GlcNAc; to accommodate the acetyl group, INQ is replaced by SNA and ANQ, respectively. Lymphostatin has leucine in position 1 and glycine in position 3, LNG. Our model suggests this would make it possible for UDP-GlcNAc to bind (Fig. 5B).

Given that digest fragment F1 spans both the putative glycosyltransferase motif as well as the cysteine protease motif, we set out to test the F1 fragment in isolation in the bovine T cell proliferation assay, to explore whether it retained its inhibitory activity against lymphocytes. Using equimolar amounts of protein in the ConA-stimulated T cell proliferation assay, comparison of the full-length protein alongside purified F1 revealed that the F1 fragment was insufficient to inhibit T cell proliferation in isolation (Fig. 6). This may be because the F1 fragment N-terminal boundary is at ~177 aa and the N-terminal amino acids are important for interaction with the target protein. More likely is that the C-terminal domains are crucial for cell binding and uptake, as has been shown for the large clostridial toxins (50).

Lymphostatin Is an Ordered Protein with a High α-Helical Content in the Putative Glycosyltransferase Domain—To confirm that rLifA was folded and to estimate the content of secondary structural elements we measured the CD spectrum of full-length and the F1 fragment that encompasses the putative GT and the CP domains. The CD spectrum of full-length protein was consistent with a folded protein and indicated 37% α-helix and 17% β-sheet (Fig. 7, A and B); very similar to PSIPRED predictions (~35% α-helical and ~22% β-sheet). The N-terminal half of the protein was predicted to contain the majority of the α-helical content, whereas the C-terminal third of the protein was predicted to be rich in β-sheet (Fig. 7C). The samples showed no evidence of aggregation by dynamic light scattering (ZetasizerAPS, Malvern).

Given the lack of activity of the isolated F1 fragment in the T cell proliferation assay, and to rule out that the protein was degraded and/or not folded, CD analysis was carried out on purified F1 protein. Furthermore, given its location at the N-terminal part of lymphostatin, where the primary sequence is predicted to have a higher percentage of α-helices, it is expected that the percentage of α-helical content of F1 would be higher than that of the full-length protein. The F1 fragment was predicted by PSIPRED to consist of 52% α-helical content and analysis of CD data inferred 51% α-helical content. The CD data is broadly in agreement with the secondary structure prediction for F1 but suggested a slightly higher proportion of β-sheet, 13% as opposed to 5%.

Three-dimensional Shape of Lymphostatin—Negatively stained rLifA showed a homogeneous distribution of elongated, slightly kinked particles (Fig. 8). From images of ~25,000 of these particles, we determined a three-dimensional map of lymphostatin at 23-Å resolution. The map shows an elongated, L-shaped molecule (Fig. 8). The arms of the “L” are 130–140 Å long and between 60 and 90 Å thick. To further investigate the structure of lymphostatin, we also performed SAXS (51, 52). The linearity of the Guinier region confirms the absence of inter-particle effects (Fig. 9A). The longest particle dimension Dmax is 218 Å, which is not dissimilar to the longest dimension of the EM map (~197 Å) (Fig. 9B). The Porod volume (645 nm3) indicates a particle molecular mass of 358 kDa further suggesting that lymphostatin (365 kDa based on primary sequence)
behaves as a monomer in solution, and that the recombinantly purified full-length protein is stable. Next, we determined the ab initio shape analysis with DAMMIF (40) using 10 independent computations producing models related with an average normalized spatial discrepancy of 0.793. The representative and DAMMIN (42) refined model shows an elongated L-shape with dimensions (218 Å/H11003116 Å/H11003100 Å) slightly larger than the EM density map (197 Å/H11003110 Å/H1100389 Å). The alignment of the EM density map with the SAXS envelope highlights the similarities in dimensions as well as shape (Fig. 9C). The kink of the L-shaped molecule is situated in a similar position in the two envelopes, at roughly half the length of the particle, however, the kink is more defined in the EM map. In addition, the EM map contains a larger volume within one of the arms of the L.

Lymphostatin Binds UDP-N-acetylglucosamine—Lymphostatin contains 37 tryptophan residues, of which 7 are found within the putative GT domain; Trp320, Trp534, and Trp863 most likely are in close proximity to the active site (Fig. 4). Titration of lymphostatin with UDP-Glc gave a linear Stern-Volmer plot that would be consistent with collisional quenching of solvent-exposed tryptophan rather than any specific interaction of sugar with lymphostatin (Fig. 10). In contrast UDP-GlcNAc enhances fluorescence in a non-linear concentration-dependent manner at lower concentrations; there is

**FIGURE 4.** The putative glycosyltransferase domain of lymphostatin. A, alignment of sequence representing the putative GT domain of lymphostatin with those of Toxin A and B from *C. difficile*, Toxin L from *Clostridium sordellii*, and α-toxin from *Clostridium novyi*. Sequences were aligned using the program Clustal Omega (66). Red bars represent α-helical regions and blue arrows represent β-sheet from the crystallographic analysis of Toxin A from *C. difficile* (Protein Data Bank code 4DMW (53)). Amino acid identities between the LCTs and lymphostatin are indicated in bold. Residues that make key interactions with the UDP-Glc donor substrate binding site of the LCTs is indicated with gray shading in both the sequence alignment and structural representations. The numbering scheme refers to lymphostatin. F2 indicated the N-terminal boundary of digest fragment F2. B, representation of the crystal structure of toxin A from *C. difficile* bound to UDP-Glc (Protein Data Bank code 4DMW) (67). C, PHYRE2 (49) generated model of the GT domain of lymphostatin based on Toxin A from *C. difficile*. UDP-Glc extracted from Protein Data Bank code 4DMW was modified to UDP-GlcNac and fitted into the putative binding site using Coot (48).
Lymphostatin Characterization and Activity

Since its discovery, two predicted features have been known that are essential for the biological activity of lymphostatin: namely a glycosyltransferase domain and a YopT-like cysteine protease domain (8, 17). Although both motifs were implicated in colonization of mice and colonic hyperplasia by C. rodentium (5), the mutations caused truncation of the full-length protein is stable. Lymphostatin contains predicted signal sequence at the N terminus. 

ConA-stimulated peripheral bovine T cells. Data shown are the average of 4 independent donors, done in triplicate. Error bars indicate the S.E. The purity of the input cells used was >85% in each experiment, as measured by CD3 staining using flow cytometry. Concentrations on the x axis are displayed on a log scale.

FIGURE 5. Modeling donor substrate specificity of lymphostatin by comparison to clostridial toxins. A, alignment of the sequence of lymphostatin with those LCT residues thought to contribute to specificity for UDP-sugar. B, PHYRE2 (49) generated model of the GT domain of lymphostatin in complex with UDP-GlcNAc based on Toxin A from C. difficile, LeuG667 and Gly669 are shown in blue and Asp557 and Asp559 in red. C, toxin A from C. difficile in complex with UDP-Glc, Ile581 and Gin584 are shown in blue and Asp285 and Asp287 in red.

FIGURE 6. The putative glycosyltransferase domain of lymphostatin is not sufficient in isolation to inhibit T cell proliferation. Effect of concentration of rLifA and the purified partial tryptic digestion fragment F1 against ConA-stimulated peripheral bovine T cells. Data shown are the average of 4 independent donors, done in triplicate. Error bars indicate the S.E. The purity of the input cells used was >85% in each experiment, as measured by CD3 staining using flow cytometry. Concentrations on the x axis are displayed on a log scale.

evidence of collisional quenching at higher concentrations. This may be explained by UDP-GlcNAc binding in the GT catalytic site and changing the environment of one or more tryptophan residues. At high concentrations of UDP-GlcNAc after the active site has become saturated there is a linear relationship between fluorescence and UDP-GlcNAc concentration that could be attributed to collisional quenching of tryptophan outside the binding site. We have fitted a mixed binding model to the UDP-GlcNAc data that takes into account specific binding and collisional quenching. The affinity of lymphostatin for UDP-GlcNAc is estimated to be 120 ± 30 μM. The protein was expressed in the presence of a variety of divalent cations including Ca²⁺, Mg²⁺, and Mn²⁺ that were present in the growth media. Good reproducibility of affinity data between batches was achieved by incubating the protein sample with 50 μM MnCl₂ after purification. Manganese was selected as the divalent cation based on thermal shift experiments (differential scanning fluorimetry) that suggested MnCl₂ enhanced the thermal stability of lymphostatin and the additional stabilizing effect of UDP-GlcNAc to a greater extent than MgCl₂ (data not presented) (53). The Km for full-length TcdA from C. difficile for UDP-Glc in the presence of Mg²⁺ is 142 μM and has a Kᵢ of 45 ± 10 μM (54). The Km for the TcdA-GT domain has been reported as 36.3 ± 3.6 μM and Kᵢ of 11.4 ± 0.9 μM (53). Km for full-length TcdB from C. difficile for UDP-Glc in the presence of Mg²⁺ is 154 μM (55).

The DTD Motif within the GT Domain of Lymphostatin Is Important in Sugar Binding and Activity against Bovine Lymphocytes—Given that the DXD motif has been implicated in glycosyltransferase activity in other similar molecules (reviewed in Ref. 13), we examined the effect of a DTD to AAA substitution at position 557–559 on both sugar binding and activity of lymphostatin. rLifADTD/AAA behaved similarly to rLifA, both during expression and purification, as well as in CD and DLS lending confidence that the residue substitutions had not disrupted the original structure of the protein.

Testing rLifADTD/AAA using the tryptophan fluorescence assay and titrating either UDP-Glc or UDP-GlcNAc as for rLifA gave linear Stern-Volmer plots indicative of nonspecific quenching and lack of sugar binding (Fig. 11A). Furthermore, titration of the rLifADTD/AAA protein in the bovine T cell proliferation assay shows a drastic reduction in its ability to inhibit proliferation by almost 4 orders of magnitude (ED₅₀ rLifA = 0.014 ng/ml ± 0.0015, rLifADTD/AAA = 922 ng/ml ± 270) (Fig. 11B). These two observations imply that not only is the DTD motif important for sugar binding, but is also critical for the inhibitory activity of lymphostatin on T cells.

Discussion

Lymphostatin plays an important role in intestinal colonization by attaching and effacing E. coli and is therefore a potentially attractive target for intervention. Although functions have been assigned to this large molecule, its mode of action remains largely unknown. It is one of relatively few bacterial molecules reported to interfere with the function of adaptive immune cells. Historically, it has proven difficult to obtain a stable clone to express and purify full-length lymphostatin (2, 8, 47). Using a tightly inducible prokaryotic expression system we were able to express and purify lymphostatin with potent activity against mitogen-stimulated T lymphocytes. Mass spectrometry of tryptic peptides and SAXS indicate that the purified full-length protein is stable. Lymphostatin contains predicted non-canonical prepilin peptidase sites (8), however, it is not clear whether it is processed, at least in the laboratory-adapted E. coli K-12 strain used for expression, nor is there a predicted signal sequence at the N terminus.

Since its discovery, two predicted features have been known in lymphostatin: namely a glycosyltransferase domain and a YopT-like cysteine protease domain (8, 17). Although both motifs were implicated in colonization of mice and colonic hyperplasia by C. rodentium (5), the mutations caused trunca-
tion of lymphostatin and subsequent studies with in-frame deletions suggested the motifs were not essential for intestinal colonization of cattle by enterohemorrhagic \textit{E. coli} O26:H11002 (4). However, the effect of lymphostatin and motif mutations on mucosal lymphocytes and induction of adaptive immunity was not studied and a role for the motifs in pathogenesis therefore cannot be excluded. Using purified lymphostatin we have not observed autocatalytic cleavage via the putative cysteine protease domain, including under low pH conditions or in the presence of inositol hexakisphosphate, the co-factor known to trigger autocatalytic cleavage of LCTs (reviewed in Ref. 56). If lymphostatin is able to autocatalytically cleave using the CP motif, either it requires some as yet unidentified co-factor, or it is not triggered by pH change, and merits further closer investigation.

Although we have obtained low resolution EM and SAXS envelopes for lymphostatin, a high resolution structure has been elusive. A full-length diffracting protein crystal has not yet been obtained from pilot crystallization trials. Given the behavior of the recombinant protein in solution, high resolution cryo-EM might be a promising approach to analyze the lymphostatin structure. Alternatively, given the data from the partial tryptic digestion, an approach combining crystallization of predicted domains combined with EM could be used. Circular dichroism spectroscopy of full-length and the N-terminal putative glycosyltransferase domain has confirmed the secondary structure predictions that show the N-terminal third of the protein to be largely α-helical. In addition, the C-terminal third is very rich in β-sheet, both features shared with the LCTs. Interestingly, unlike the LCTs, lymphostatin lacks the multimodular cell wall/choline-binding repeat that form the receptor-binding domain (57).

For the first time we report that lymphostatin is able to bind UDP-GlcNAc, but not UDP-Glc in a manner dependent on a conserved DXD motif within its predicted glycosyltransferase domain and that this motif is also required for its activity against bovine T lymphocytes. Furthermore, inhibition of lymphocyte function was not associated with cell lysis as measured by release of a cytosolic enzyme or effects on the actin cytoskeleton, in contrast to LCTs (data not shown) (58, 59). The target of sugar modification by lymphostatin is unknown, and is likely to be difficult to identify. For both the LCTs and NleB, there was some initial insight into rational candidate acceptor proteins or signaling pathways prior to their identification. However, this is not currently the case for lymphostatin. Furthermore, whereas potential targets of NleB activity were identified by immunoprecipitation and yeast 2-hybrid approaches, enabling specific sugar transfer onto putative recombinant targets to be tested (14, 15), this approach is likely to be technically challenging for lymphostatin, because of the challenges of cloning the gene in the vectors required for yeast 2-hybrid screening. Interacting partners for lymphostatin are presently unknown and a similar
approach may be needed to find its target(s), given the challenge of detecting addition of labeled GlcNAc to cellular proteins against the large number of endogenous modifications expected.

Although lymphostatin has recently been demonstrated to be secreted via Type III secretion (10), the fact that it is active in an isolated recombinant form, or when produced in E. coli that lack a type III secretion system, indicates that injection into cells is not vital for activity against lymphocytes. It is plausible that it may have both Type III secretion-dependent and -independent activities. Our evidence indicates that the full-length protein is required for inhibition of lymphocyte function as the isolated F1 fragment encompassing the glycosyltransferase domain showed no activity, despite being folded and soluble.

FIGURE 9. SAXS analysis of rLifA. A, the experimental SAXS scattering curve is shown as a logarithmic intensity I (s) versus inverse scattering angle (Å⁻¹). The linear region of the Guinier plot (log I(s) versus s²) is shown as an inset. B, the distance distribution function P(r) with SAXS parameters shown in the inset. Rg (radius of gyration); the Mw (monomer) is the mass calculated from the primary sequence; Mw (SAXS) is the estimated mass calculated from the Porod volume. C, the refined DAMMIN ab initio bead model with indicated dimensions (blue) and its ab initio surface envelope (blue) aligned with the EM density map (gray), shown in two views related by a 90 degree rotation around the horizontal axis. The SAXS envelope was made transparent with Chimera (44) to visualize the alignment.

FIGURE 10. Lymphostatin binds UDP-N-acetylglucosamine. Stern-Volmer plot showing the intrinsic tryptophan fluorescence (340 nm) of lymphostatin, measured at increasing concentrations of UDP-N-acetylglucosamine (UDP-GlcNAc) and UDP-glucose (UDP-Glc) (F₀/F; where F₀ is the fluorescent intensity before the addition of UDP-sugar). Data points represent the means of data in triplicate. Kᵣ for UDP-GlcNAc was determined to be 120 ± 30 µM by fitting the raw fluorescent intensity data to a one-site binding model corrected for collisional quenching (inset).

FIGURE 11. Mutation of the DxD motif of lymphostatin abolishes both sugar binding and inhibitory activity. A, Stern-Volmer plot showing the quenching of intrinsic lymphostatin tryptophan fluorescence (340 nm) measured at increasing concentrations of UDP-GlcNAc and UDP-Glc (F₀/F; where F₀ is the fluorescent intensity before the addition of UDP-sugar). B, effect of concentration of rLifA and rLifADTD/AAA against ConA-stimulated proliferation of peripheral bovine T cells. The data shown are the average of 3 independent donors, done in triplicate. Error bars indicate S.E. The calculated ED50 was 0.014 ± 0.0015 ng/ml (rLifA), and 922 ± 270 ng/ml (rLifADTD/AAA). Concentrations on the x axis are displayed on a log scale.
Lymphostatin Characterization and Activity

with large clostridial toxins the C-terminal domain of lymphostatin may be required for cellular uptake and the GT domain is unable to act in isolation (60, 61).

Where EPEC and non-O157 EHEC almost invariably express lymphostatin, in serogroup O157 EHEC strains, lymphostatin exists as two truncated open reading frames (z4332 and z4333, encoding proteins identical to residues 1–433 and 435–710 of full-length lymphostatin) (62, 63). In addition, the pO157 plasmid encodes a putative 365-kDa homologue of lymphostatin named ToxB (L7095), which shares 28% identity and 47% amino acid similarity to LifA/Efa1 (64, 65). E. coli O157 strains have a lymphostatin-like activity (3, 15) that has been associated with the pO157 plasmid (2). Mutation of the truncated lymphostatin or toxB genes in a Shiga toxin-deficient E. coli O157 strain did not markedly affect the ability of bacterial lysates to inhibit bovine lymphocyte proliferation (11). However, the assay used relied on crude bacterial lysates and is insensitive compared with the highly purified protein tested herein, with lysates causing lymphostatin-independent inhibition of lymphocyte function at higher concentrations (4, 11).

The expression and assay systems described here are likely to be suitable to assign activities to ToxB and truncated lymphostatin. The expression and assay systems described here are likely to be compatible with T cell proliferation for full-length lymphostatin and, indeed a family of homologous cytotoxins described in suitable to assign activities to ToxB and truncated lymphostatin. The expression and assay systems described here are likely to be suitable to assign activities to ToxB and truncated lymphostatin, and indeed a family of homologous cytotoxins described in suitable to assign activities to ToxB and truncated lymphostatin, and indeed a family of homologous cytotoxins described in suitable to assign activities to ToxB and truncated lymphostatin.

Acknowledgments—We thank Tomás Plivelic at SAXS beamline I911–4 MAX-II Laboratory (Lund, Sweden) for help with SAXS data collection. Access to synchrotron beam time was made possible by I911–4 MAX-II Laboratory (Lund, Sweden) for help with SAXS data collection. Access to synchrotron beam time was made possible by the EM staining, image acquisition and analysis to produce the three-dimensional reconstruction of lymphostatin, and contributed to writing the first draft of the manuscript; R. J. carried out SAXS analysis and contributed to writing the manuscript; M. P. S. contributed to writing the manuscript. All authors contributed to analyzing the data, editing the manuscript, and approved the final form.

Author Contributions—R. L. C. C., E. A. B., M. W., R. J., and M. P. S. contributed to conception and design of study; R. L. C. C. cloned the original expression construct and initial protein studied, carried out the T cell proliferation assays for full-length lymphostatin and the F1 fragment, analyzed data, constructed figures, and carried out statistical analysis, wrote the first draft of the manuscript, and coordinated revision and submission; E. A. B. optimized protein production and purification, carried out biophysical characterization and analysis, domain bioinformatic analysis, all CD analysis, sugar binding assays, and contributed to writing the first draft of the manuscript; H. A. carried out sugar binding assays and SAXS analysis for full-length lymphostatin; E. D. carried out SAXS analysis for full-length lymphostatin and contributed to writing the first draft of the manuscript; A. B. prepared the AAA mutant of lymphostatin, helped with optimization of production and purification, carried out proliferation and sugar binding assays using the mutant protein; B. B. carried out the EM staining, image acquisition and analysis to produce the three-dimensional reconstruction of lymphostatin, and contributed to writing the first draft of the manuscript; R. J. carried out SAXS analysis and contributed to writing the manuscript; M. P. S. contributed to writing the manuscript. All authors contributed to analyzing the data, editing the manuscript, and approved the final form.

References

1. Croxon, M. A., and Finlay, B. B. (2010) Molecular mechanisms of Escherichia coli pathogenicity. Nat. Rev. Microbiol. 8, 26–38
2. Klapproth, J. M., Scaletsky, I. C., McNamara, B. P., Lai, L. C., Malstrom, C., James, S. P., and Donnenberg, M. S. (2000) A large toxin from pathogenic Escherichia coli strains that inhibits lymphocyte activation. Infect. Immun. 68, 2148–2155
3. Stevens, M. P., van Diemen, P. M., Frankel, G., Phillips, A. D., and Wallis, T. S. (2002) Efa1 influences colonization of the bovine intestine by Shiga toxin-producing Escherichia coli serotypes O5 and O111. Infect. Immun. 70, 5158–5166
4. Deacon, V., Dziva, F., van Diemen, P. M., Frankel, G., and Stevens, M. P. (2010) Efa-1/LifA mediates intestinal colonization of calves by enterohemorrhagic Escherichia coli O26:H1 in a manner independent of glycosyltransferase and cysteine protease motifs or effects on type III secretion. Microbiology 156, 2527–2536
5. Klapproth, J.-M., Sasaki, M., Sherman, M., Babin, B., Donnenberg, M. S., Fernandes, P. J., Scaletsky, I. C., Kalman, D., Nusrat, A., and Williams, I. R. (2005) Citrobacter rodentium lifa/efa1 is essential for colonizing and crypt cell hyperplasia in vivo. Infect. Immun. 73, 1441–1451
6. Belland, R. J., Scidmore, M. A., Crane, D. D., Hogan, D. M., Whitmore, W., McClarty, G., and Caldwell, H. D. (2001) Chlamydia trachomatis cytotoxicity associated with complete and partial cytotoxin genes. Proc. Natl. Acad. Sci. U.S.A. 98, 13984–13989
7. Klapproth, J. M., Donnenberg, M. S., Abraham, J. M., and James, S. P. (1996) Products of enteropathogenic Es. coli inhibit lymphokine production by gastrointestinal lymphocytes. Am. J. Pathol. 127, 8481–848
8. Nicholls, L., Grant, T. H., and Robins-Browne, R. M. (2000) Identification of a novel genetic locus that is required for in vitro adhesion of a clinical isolate of enterohaemorrhagic Escherichia coli to epithelial cells. Mol. Microbiol. 35, 275–288
9. Badea, L., Doughty, S., Nicholls, L., Sloan, J., Robins-Browne, R. M., and Hartland, E. L. (2003) Contribution of Efa1/LifA to the adherence of enteropathogenic Escherichia coli to epithelial cells. Microb. Pathog. 34, 205–215
10. Deng, W., Yu, H. B., de Hoog, C. L., Stoynov, N., Li, Y., Foster, L. J., and Finlay, B. (2012) Quantitative proteomic analysis of type III secretome of enteropathogenic Escherichia coli reveals an expanded effector repertoire for attaching/effacing bacterial pathogens. Mol. Cell. Proteomics 11, 692–709
11. Abu-Median, A.-B., van Diemen, P. M., Dziva, F., Vlisidou, I., Wallis, T. S., and Stevens, M. P. (2006) Functional analysis of lymphostatin homologues reveals an expanded effector repertoire for attaching/effacing bacterial pathogens. Mol. Cell. Proteomics 11, 692–709
12. Badea, L., Doughty, S., Nicholls, L., Sloan, J., Robins-Browne, R. M., and Hartland, E. L. (2003) Contribution of Efa1/LifA to the adherence of enteropathogenic Escherichia coli to epithelial cells. Microb. Pathog. 35, 43–49
13. Janik, T., Belyi, Y., and Aktories, K. (2015) Bacterial glycosyltransferases toxins. Cell Microbiol. 17, 1752–1765
14. Lairson, L. L., Henriisat, B., Davies, G. J., and Withers, S. G. (2008) Glycosyltransferases: structures, functions, and mechanisms. Annu. Rev. Biochem. 77, 521–555
15. Li, S., Zhang, L., Yao, Q., Li, L., Dong, N., Rong, J., Gao, W., Ding, X., Sun, L., Chen, X., Chen, S., and Shao, F. (2013) Pathogen blocks host death receptor signalling by arginine GlcNAcylation of death domains. Nature 501, 242–256
16. Pearson, J. S., Giogha, C., Ong, S. Y., Kennedy, C. L., Kelly, M., Robinson, K. S., Lung, T. W., Mansell, A., Riedmaier, P., Oates, C. V., Zaid, A., Mühlen, S., Crepin, V. F., Marches, O., Ang, C.-S., Williamson, N. A., O'Reilly, L. A., Bankovacki, A., Nachbur, U., Infusini, G., Webb, A. I., Silke, J., Strasser, A., Frankel, G., and Hartland, E. L. (2013) A type III effector antagonizes death receptor signalling during bacterial gut infection. Nature 501, 247–251
17. Stevens, M. P., and Frankel, G. M. (2014) The locus of enterocyte effacement and associated virulence factors of enterohemorrhagic Escherichia coli. Microbiol. Spectr. 2, HEC-0007–2013
18. Shao, F., Merritt, P. M., Bao, Z., Innes, R. W., and Dixon, J. E. (2002) A Yersinia effector and a Pseudomonas avirulence protein define a family of cysteine proteases functioning in bacterial pathogenesis. Cell 109, 575–588

5814 JOURNAL OF BIOLOGICAL CHEMISTRY

ASBMB

VOLUME 291 • NUMBER 11 • MARCH 11, 2016
Lymphostatin Characterization and Activity

D. J., Mayhew, G. F., Evans, P. S., Gregor, J., Kirkpatrick, H. A., Pósfai, G., Hackett, J., Klink, S., Boutin, A., Shao, Y., Miller, L., Grotbeck, E. J., Davis, N. W., Lim, A., Dimalanta, E. T., Potamousis, K. D., Apodaca, J., Antharaman, T. S., Lin, J., Yen, G., Schwartz, D. C., Welch, R. A., and Blattner, F. R. (2001) Genome sequence of enterohaemorrhagic Escherichia coli O157:H7. Nature 409, 529–533

64. Burland, V., Shao, Y., Perna, N. T., Plunkett, G., Sofia, H. J., and Blattner, F. R. (1998) The complete DNA sequence and analysis of the large virulence plasmid of Escherichia coli O157:H7. Nucleic Acids Res. 26, 4196–4204

65. Makino, K., Ishii, K., Yasunaga, T., Hattori, M., Yokoyama, K., Yutsudo, C. H., Kubota, Y., Yamaichi, Y., Iida, T., Yamamoto, K., Honda, T., Han, C. G., Ohtsubo, E., Kasamatsu, M., Hayashi, T., Kubara, S., and Shinagawa, H. (1998) Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an enterohemorrhagic Escherichia coli O157:H7 derived from Sakai outbreak. DNA Res. 5, 1–9

66. Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D., and Higgins, D. G. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. 7, 539

67. Delano, W. L. (2014) The PyMOL Molecular Graphics System, version 1.7.6.6, Schrödinger, LLC, New York