SUPPLEMENTAL DATA

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Cloning of Mub repeats— L. reuteri 1063 was obtained from the American Type Culture Collection (strain ATCC 53608) and maintained in MRS broth (per litre, Oxoid peptone 10 g, Oxoid "Lab-Lemco" 8 g, Difco yeast extract 5 g, glucose 20 g, Tween-80 1 ml, K_2HPO_4 2 g, sodium acetate·$3H_2O$ 5 g, tri-ammonium citrate 2 g, $MgSO_4 \cdot 7H_2O$ 0.58 g, $MnSO_4 \cdot 4H_2O$ 0.14 g, pH 6.5). Oligonucleotide primers for PCR amplification of DNA molecules encoding individual or multiple repeats of the *L. reuteri* 1063 Mub protein were designed to anneal to specific Mub domain border regions, as defined in (1,2), which differ from those described in the original GenBank Accession Number AF120104 (Table S1). Cultures of *L. reuteri* 1063 were grown from -80 °C glycerol stocks in 20 ml MRS broths in 1 oz. bottles at 30 °C for 24 h without shaking. Aliquots of culture (3 ml) were centrifuged and the cell pellets resuspended with 1 ml sterile ultrapure water for use as template (10 µl) in 50 µl PCRs with HotStarTaq Master Mix Kit (Qiagen). Amplified DNA fragments were purified by Qiaquick PCR purification or Qiaquick gel extraction kits (Qiagen) and inserted directly into the combined cloning/expression vector, pETBlue-1 AccepTor (Novagen), following the manufacturer's instructions. After transformation into the screening host, *E. coli* NovaBlue (Novagen), plasmid DNAs were purified by plasmid mini purification kit (Qiagen). The integrity and orientation of cloned sequences were checked by automated DNA sequencing (John Innes Genome Centre, Norwich, UK) using appropriate sequencing primers (Table S1). pETBlue-1:Mub-repeat-containing vectors were transformed into the expression host, *E. coli* Tuner (DE3)pLacI (Novagen), following the manufacturer's instructions.

Site-directed Mutagenesis—The leucine to methionine site-directed mutant L48M of Mub-R5 was generated using the QuikChange Site-Directed Mutagenesis Kit (Stratagene) with the gene-specific oligonucleotides listed in Table S1 and vector pETBlue-1:Mub-R5 as template. Mutated plasmid was transformed into *E. coli* XL1-Blue (Stratagene) for DNA isolation and sequencing.

Expression and Purification of Recombinant Proteins—Small-scale (50 ml) cultures in LB broth supplemented with 50 µg/ml carbenicillin, 34 µg/ml chloramphenicol and 1 % (w/v) glucose were induced with 1 mM IPTG for 3 h at 37 °C. For protein expression screening of single Mub repeats, three cycles of freeze-thaw extraction were carried out with cell pellets from 3 ml of culture (3) finally suspending the treated cells in 50 µl 10 mM sodium phosphate buffer (pH 7.5) to produce an extract of ~60 µl. Large-scale (1 l) LB cultures were also induced with 1 mM IPTG for 3 h prior to cell harvest and extraction. In this case, duration of the three freeze-thaw cycles was extended to 10 min freezing and 20 min thawing, finally suspending the treated cells with 3 ml 10 mM sodium phosphate buffer (pH 7.5) per gram wet weight cells. Large-scale freeze-thaw extracts were clarified by centrifugation at 23,300 $\times g$ for 30 min at 4 °C. The three recombinant single Mub repeats Mub-RI, Mub-R5 and Mub-R6 were expressed in soluble form in *E. coli* at total yields of ~25 mg/g wet weight cells. In freeze-thaw soluble extracts, the single domain proteins were ~80-90 % pure, as determined by SDS-PAGE, and represented ~50 % of the total soluble recombinant protein produced. Protein expression screening of the triple Mub-RI-III repeat was carried out using cell pellets from 50 ml of induced culture. For Mub-RI-III, with a predicted molecular mass of ~64 kDa, freeze-thaw extraction of whole cells proved less efficient, as previously reported for high molecular weight proteins (3). In this case, soluble Mub-RI-III was extracted from freeze-thawed cells using a detergent mixture (BugBuster) with a total yield of 5-10 mg/g wet weight cells. Proteins were extracted from freeze-thawed cells by gently shaking with 1.3 ml BugBuster HT (Novagen), containing Complete™ Protease Inhibitor Cocktail (Roche), for 20 min at room temperature, followed by centrifugation at 17,000 $\times g$ for 20 min at 4 °C. For large-scale extractions of Mub-RI-III, freeze-thawed cell pellets from 1 l cultures were treated as above with 5 ml BugBuster HT, containing Complete Protease Inhibitor Cocktail, per gram wet weight cells. Extracts were clarified by centrifugation at 23,300 $\times g$ for 20 min at 4 °C and dialysed in Spectra/Por MWCO 3,500 tubing (Spectrum Laboratories, Inc.) with two changes against 4 l of 10 mM sodium phosphate buffer (pH 7.5) to remove excess detergents.

Protein purification was performed by means of an ÄKTA FPLC system controlled by UNICORN v3.20 software (GE Healthcare). Both freeze-thaw and dialysed BugBuster HT protein extracts were

filtered through 0.22 µm low protein binding PVDF membranes (MILLEX-GV, Millipore) and loaded on to a Mono Q HR 10/10 ion-exchange column (GE Healthcare) pre-equilibrated with 10 mM sodium phosphate buffer (pH 7.5). Proteins were eluted with a linear gradient of 10 mM sodium phosphate buffer (pH 7.5) containing 0-0.3 M NaCl at 1 ml/min over 78 min and 1.5 ml fractions collected. Peak fractions of FPLC-purified Mub proteins were pooled and stored at 4 °C.

pETBlue-1:Mub-R5 and pETBlue-1:Mub-R5(L48M) plasmid DNAs were co-transformed with pLacI (Novagen) into *E. coli* B834(DE3) (Novagen) using double selection with carbenicillin (50 µg/ml) and chloramphenicol (34 µg/ml) in the presence of 1 % (w/v) glucose. Prior to selenomethionine labelling the proteins, the timecourse of induction with 1 mM IPTG was followed over a 42 h period in SelenoMet Medium Base and Nutrient Mix (Athena Enzyme Systems) containing 40 µg/ml L-methionine at 37 °C. Following optimisation of the induction conditions, each strain was grown in 2 × 1 l of the above medium supplemented with 40 µg/ml L-SelenoMethionine solution (Athena Enzyme Systems) instead of methionine and cells harvested after 12 h induction. Recombinant selenomethionine-labelled proteins were freeze-thaw extracted and purified as described above.

Protein Gel Electrophoresis—SDS-PAGE of reduced/denatured proteins was carried out in 4-12 % NuPAGE Bis-Tris gels (Invitrogen) in MES-SDS electrophoresis buffer for 35 min at 200 V and proteins sized against Broad Range Protein Molecular Weight Markers (Promega). Isoelectric focussing gel electrophoresis for pI determination of desalted protein samples was carried out in Novex IEF pH 3-7 gels (Invitrogen) against IEF Markers 3-10, SERVA Liquid Mix (Invitrogen), at room temperature for 1 h at 100 V, 1 h at 200 V and 30 min at 500 V. All gels were stained with Coomassie Colloidal Blue (Invitrogen).

Mass Spectrometry of Purified Proteins—For ESI-MS analysis, FPLC-purified proteins were buffer-exchanged into Ultrapure H_2O by dialysis or PD-10 gel filtration to a concentration of ~1 mg/ml. Samples (50 µl) were introduced by flow injection into a micrOTOF mass spectrometer (Bruker Daltonics Ltd.) which had been calibrated with 10 nM sodium formate. The flow rate of the running solvent, MilliQ water, was 0.2 ml/min. Data were acquired in positive ionisation mode at a capillary voltage of 4200 V and over a scan range of 250-3000 *m/z*. Background-subtracted mass spectra were deconvoluted using Bruker Daltonics DataAnalysis v3.4 deconvolution software to calculate the mean molecular mass of each species.

Proteins for MALDI-ToF-MS analysis were excised from Coomassie Colloidal Blue-stained SDS-PAGE gels and trypsinised *in situ* using a ProGest Protein Digester (Genomic Solutions) (4). Acidified digests, were automatically spotted onto pre-spotted anchor chip plastic MALDI PAC384 target plates (Bruker Daltonics Ltd.) using a MALDI-AutoPrep (MAP II) spotting robot (Bruker Daltonics Ltd.). Each sample (0.4 µl) was applied to the matrix-only spots. When the samples had dried the whole plate was dipped into 1.5 l of 10 mM ammonium phosphate, 0.1 % (w/v) trifluoroacetic acid for 3-5 s, removed and air-dried. Analysis of the sample digests was carried out on an Ultraflex MALDI-ToF/ToF mass spectrometer (Bruker Daltonics Ltd.). A 200 Hz nitrogen laser was used to desorb/ionise the matrix/analyte material and ions were detected in positive ion reflectron mode.

All spectra were acquired automatically using the Bruker fuzzy logic algorithm. Successful spectral packets for each sample were summed, calibrated, annotated and background subtracted using FlexAnalysis v3.0. The results were then automatically set up for searching against the relevant sequence database using the pre-assigned Biotools method which submits searches to the Mascot 2.1 search engine (Matrix Science Ltd.).

Aliquots of the same trypsinised samples were also analysed by LC-MS/MS in an LTQ-Orbitrap mass spectrometer (Thermo Fisher). Peptides were desalted using a pre-column (C18 pepmap100, Dionex Corp.) which was then switched in-line to a 12 cm analytical column (Picotip, 75 µm id, 15 µm tip, New Objective, Inc.) self-packed with Symmetry300 C18, 5 µm, (Waters). Samples were applied via a Surveyor HPLC system running at a flow rate of ~0.25 µl/min to an LTQ-Orbitrap™ mass spectrometer. Peptides were eluted with a gradient of 5-45 ml/100 ml acetonitrile in water/0.1 % (w/w) formic acid at a rate of 1 ml to 100 ml/min. The mass spectrometer was operated in positive ion mode at a capillary temperature of 200 ºC. The source voltage and focusing voltages were tuned for the transmission of MRFA peptide (*m/z* 524) (Sigma-Aldrich). Data dependent analysis was carried out in Oribtrap-IT parallel mode using CID fragmentation on the four most abundant ions in each

cycle. Collision energy was 35, and an isolation width of 2 was used. The Orbitrap was run with a resolution of 30,000 over the MS range from 400-1800 *m/z* and an MS target of 10^6 and 1 s maximum scan time. The MS2 was triggered by a minimal signal of 7000 with an AGC target of 3×10^4 ions and 150 ms scan time. For selection of 2^+ and 3^+ charged precursors, charge state and monoisotopic precursor selection was used. Dynamic exclusion was set to 2 counts and 60 s exclusion with an exclusion mass window of -0.5 to +1.5. MS scans were saved in profile mode while MS/MS scans were saved in centroid mode. Raw files were processed in Bioworks (Thermo Fisher) to generate dta-files which were merged into an mgf-file using a Perl script. The mgf-file was submitted for a database search using an in-house Mascot 2.2 Server (Matrix Science Ltd.). Searches were performed on the SPTrEMBL database using 5 ppm precursor tolerance, 0.5 Da fragment tolerance, 2 missed cleavages, oxidation (M) as variable and carbamidomethylation (C) as fixed modification.

Circular Dichroism Spectroscopy—FPLC-purified proteins were buffer-exchanged into ultrapure H_2O by PD-10 gel filtration to a concentration of 1 mg/ml and a portion transferred to a 0.1 mm path-length split cuvette. A total of six far UV CD spectra were accumulated and averaged per sample in a JASCO J-710 spectropolarimeter (Great Dunmow, Cambs, UK). A scan speed of 50 nm/min was used over a scan range of 260-185 nm with a band width of 1.0 nm, a response time of 2.0 s and a data pitch of 0.5 nm. Spectra of ultrapure H_2O were also collected for background subtraction. Data were manipulated with JASCO Spectra Manager 32 v1.40.00a software and CONTINLL (5) from the CDPro suite of programs was used to calculate the spectra and the proportion of each type of secondary structure (using IBasis reference set 3).

FITC-labelling of Proteins—FPLC-purified proteins were buffer-exchanged by dialysis or PD-10 gel filtration into 0.1 M sodium carbonate buffer (pH 9.3). FITC (Sigma) was added to 0.5 ml of ~0.5-10.0 mg/ml protein solution at a 10- to 12-fold molar excess and incubated with stirring at 4 °C for 16-18 h in the dark. Labelled proteins were purified from unincorporated FITC by gel filtration through a PD-10 column equilibrated with PBS (pH 7.4) and collected in 0.5 ml fractions. The molar labelling ratio (F/P) was calculated from absorbance values at 495 nm and 280 nm of ×1/600 diluted samples, using an extinction coefficient for FITC at pH 7.4 of 63,000.

Crystallisation and Crystal Structure Determination—Single crystals of selenomethionyl-labelled Mub-R5(L48M) mutant were grown at 4 °C by vapour diffusion in hanging drops containing 2 µl of protein at 2 mg/ml in water mixed with 2 µl of a solution containing 0.2 M ammonium formate and 22 % (w/v) PEG 3350. Crystals were cryoprotected by increasing the concentration of PEG 3350 in the drops to 30 % (w/v) before harvesting and storage in liquid nitrogen. One such crystal of approximate dimensions $200 \times 50 \times 20$ µm^3 was then transferred to the N_2 stream of a cryocooler at 100 K on beamline ID02 at the Diamond synchrotron. Indexing of diffraction images showed the crystals to belong to spacegroup $P2_12_12_1$ with unit cell parameters a = 44.9 Å, b = 45.7 Å and c = 191.4 Å and to contain two copies of mubR5 in the asymmetric unit. Data for SAD phasing were then collected to 2.0 Å resolution from a single crystal at wavelength, λ = 0.979 Å (the selenium K-edge) and processed using MOSFLM (6). The location of heavy atom sites and calculation of initial phase estimates was performed with SOLVE (7) using data to a resolution of 2.0 Å. These phase estimates were improved by solvent flattening using RESOLVE (8) at a nominal solvent content of 48 % (v/v). The mean figure of merit for acentric (centric) reflections to this resolution was 0.91 (0.84) after solvent flattening. Fourier maps calculated with these phases showed clear and contiguous electron density for the majority of both molecules. The data was of sufficient quality to autotrace more than 80 % of the residues of the repeat domain in both copies of the asymmetric unit using warpNtrace in ARP/wARP (9). Cycles of model building performed using the program COOT (10), simulated annealing with PHENIX (11) and maximum likelihood refinement with REFMAC (12) followed. This resulted in a structural model lacking only the C-terminal alanine residues in each copy of the protein. This model had an R_{cryst} of 28.2 % and R_{free} of 33.3 % at 2.0 Å resolution against the selenium K-edge peak dataset.

Crystals of the native Mub-R5 protein were grown in an essentially identical fashion to those of the selenomethionine-labelled mutant protein. The best quality crystals grew at 4 °C from a 1:1 mixture of native protein at 2 mg/ml in water with 0.2 M ammonium formate and 24 % (w/v) PEG 3350. Again, crystals could be cryoprotected by increasing the concentration of PEG 3350 in the drop to 30 % (w/v) before transfer to liquid nitrogen for storage. A single crystal of approximate dimensions $180 \times 40 \times 20$ µm^3 was transferred to the N_2 stream of a cryocooler at 100 K on beamline BM14 at the

ESRF, Grenoble. X-ray diffraction data to 1.8 Å resolution were collected from a single crystal at wavelength, λ = 1.033 Å and processed using HKL2000 (13). Indexing of diffraction images revealed the crystal to be almost isomorphous to those of the labeled mutant and to belong to spacegroup $P2_12_12_1$ with unit cell parameters a = 44.9 Å, b = 45.8 Å and c = 191.6 Å, again with two molecules of Mub-R5 in the asymmetric unit. The structure was solved by molecular replacement using MOLREP employing a single copy of the partially-refined structure of Se-Met Mub-R5(L48M) as search model. Refinement followed using a combination of simulated annealing refinement with PHENIX followed by cycles of manual rebuilding using COOT and restrained refinement using REFMAC to give a final structural model with R_{cryst}= 23.0 % (R_{free}= 28.2 %) at 1.8 Å resolution. ARP/wARP (9) was used to place water molecules into significant residual electron density. The final structure refined to give an R_{cryst} of 20.2 % and R_{free} 25.9 % calculated using all data to 1.8 Å resolution. When analyzed for stereochemical quality using MOLPROBITY (14) the protein has 98.9 % of residues in the most favoured regions of the Ramachandran plot with the remainder falling into additional allowed regions. No residues are found in disallowed regions. Software from the CCP4 programme suite (15) was used throughout the structure solution process.

SUPPLEMENTAL REFERENCES

1. Boekhorst J., Helmer, Q., Kleerebezem, M., and Siezen, R. J. (2006) *Microbiology* **152**, 273-280
2. Bumbaca, D., Littlejohn, J. E., Nayakanti, H., Lucas, A. H., Rigden, D. J., Galperin, M. Y., and Jedrzejas, M. J. (2007) *Proteins*, **66**, 547-558
3. Johnson, B. H., and Hecht, M. H. (1994) *Bio/Technology* **12,** 1357-1360
4. Polley, A. C., Mulholland, F., Pin, C., Williams, E. A., Bradburn, D. M., Mills, S. J., Mathers, J. C., and Johnson I. T. (2006) *Cancer Res.* **66**, 6553-6562
5. van Stokkum, I. H. M., Spoelder, H. J. W., Bloemendal, M., van Grondelle, R., and Groen, F. C. A. (1990) *Anal. Biochem.* **191,** 110-118
6. Leslie, A.G.W. (1992) Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography, No. 26
7. Terwilliger, T., and Berendzen, J. (1999) *Acta Crystallogr. D.* **55**, 849-861
8. Terwilliger, T.C. (2003) *Methods Enzymol.* **374**, 22-37
9. Perrakis, A., Morris, R., and Lamzin, V.S. (1999) *Nat. Struct. Biol.* **6**, 458-463
10. Emsley P. and Cowtan K. (2004) *Acta Crystallogr. D.* **60**, 2126-2132
11. Adams, P.D., Grosse-Kunstleve, R.W., Hung, L.-W., Ioerger, T.R., McCoy, A.J., Moriarty, N.W., Read, R.J., Sacchettini, J.C., Sauter, N.K., and Terwilliger, T.C. (2002) *Acta Crystallogr. D* **58**, 1948-1954
12. Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997) *Acta Cryst. D.* **53**, 240-255
13. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276**, 307-326
14. Davis, I.W., Leaver-Fay, A., Chen, V.B., Block, J.N., Kapral, G.J., Wang, X., Murray, L.W., Arendall, W.B., Snoeyink, J., Richardson, J.S., and Richardson, D.C. (2007) *Nucleic Acids Res.* **35**, W375-W383
15. Collaborative Computational Project (1994) *Acta Crystallogr. D.* **50**, 760-763
16. http://emboss.sourceforge.net/apps/release/5.0/emboss/apps/pepstats.html
17. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., Bairoch, (2005) *The Proteomics Protocols Handbook*, Humana Press, John M. Walker (ed),. 571-607

SUPPLEMENTAL TABLES

TABLE S1

Residue limits of the individual repeat units in the *L. reuteri* 1063 mucus binding protein

Mub1 repeat	Residues[a]	Mub2 repeat	Residues[a]
RI	557–749 (193)	R1	1325–1510 (186)
RII	750–946 (197)	R2	1511–1694 (184)
RIII	947–1134 (188)	R3	1695–1878 (184)
RIV	1135–1324 (190)	R4	1879–2062 (184)
RV	2799–2981 (183)	R5	2063–2246 (184)
RVI	2982–3187 (206)	R6	2247–2430 (184)
		R7	2431–2614 (184)
		R8	2615–2798 (184)

[a] Starting and ending amino acid residues for each repeat derived according to (1). Number of residues in the repeat is in brackets.

TABLE S2

Primers for amplification, mutagenesis and sequencing of Mub repeats

Mub repeat(s)	Oligonucleotide	Sequence[a]	Nucleotide annealing position[b]	Comments
RI	MucB1-RIF	5'-***ATG***CAAGAAGCTGCCATCAG-3'	1789-1805	
	MucB2-RIR	5'-***TTA***CGTGTCAGCGGTATAGT-3'	2351-2367	
RI-III	MucB1-RIF	5'-***ATG***CAAGAAGCTGCCATCAG-3'	1789-1805	
	MucB2-RIIIR	5'-***TTA***GGCATTCCCAACATAAAC-3'	3505-3522	
	MubRI-seq1	5'-TATGACAACGATACGCAAGCA-3'	1969-1989	Sequencing primers for Mub-RI-III
	MubRII-seq2	5'-TATGACAACAATGACCAAACG-3'	2560-2580	
	MubRIII-seq3	5'-TATGACGATGATGACGTTAAG-3'	3151-3171	
	MubRIII-seq4	5'-GTCCGATCAACCTTTTCACCA-3'	3039-3059	
	MubRII-seq5	5'-TCATATACCGCGTCAGCTGG-3'	2545-2564	
R5 & R6	MucB1-R4F	5'-***ATG***CAAAAGGTTCACGTTCAA-3'	6307-6324 & 6859-6876	Will also amplify Mub-R4
	MucB2-R4R	5'-***TTA***GGCATCAGCCGTGTAGA-3'	6842-6858 & 7394-7410	
	R5-L3F	5'-GGCGACGGTTATGAA**ATG**TTCAAGGACAACTTCC-3'	6430-6463	Used to generate Mub-R5 mutant L48M
	R5-L3R	5'-GGAAGTTGTCCTTGAA**CAT**TTCATAACCGTCGCC-3'	6430-6463	

[a] Start and stop codons are indicated in bold italicised text. Leucine to methionine mutagenic sites are indicated in bold underlined text.
[b] Based on DNA sequence coordinates from GenBank Accession Number AF120104.

TABLE S3

Mass and MW data for recombinant Mub repeats and SeMet-labelled proteins

Mub repeat(s)	Predicted mass (Da)[a]	Observed mass (Da)[b]	Apparent denatured MW (Da)[c]	% relative abundance of SeMet
R5	20494.16	20493.68 ± 0.40 (*n* = 4)	22916.55 ± 477 (*n* = 11)	NA[d]
R6	20493.17	20494.13 ± 0.35 (*n* = 5)	23342.75 ± 466 (*n* = 4)	NA
RI	21308.08	21308.52 ± 0.26 (*n* = 6)	28136.75 ± 907 (*n* = 4)	NA
RI-III	64060.81	64060.81[e]	75018.57 ± 894 (*n* = 7)	NA
SeMet-R5	20541.06 [1][f]	20542.19 ± 0.84 (*n* = 20)	NA[d]	2.19 ± 0.29 (*n* = 3)
	20587.96 [2]	20588.40 ± 0.35 (*n* = 45)		97.58 ± 0.32 (*n* = 3)
R5(L48M)	20512.20	20512.05 ± 0.56 (*n* = 15)	NA	NA
SeMet-R5(L48M)	20559.10 [1][f]	ND[g]	NA	ND[g]
	20606.00 [2]	20607.12 ± 0.90 (*n* = 14)		2.54 ± 0.17 (*n* = 2)
	20652.90 [3]	20653.37 ± 0.48 (*n* = 31)		92.10 ± 6.44 (*n* = 2)

[a] Calculated from amino acid sequence using the Pepstats program in EMBOSS v4.0.0 (16)and ProtParam tool in ExPASy (17).
[b] From ESIMS micrOTOF data ± SD.
[c] Calculated from electrophoretic mobility in reducing/denaturing SDS-PAGE using TotalLab software (Nonlinear Dynamics, Newcastle, UK) ± SD.
[d] Not applicable.
[e] From MALDI-ToF-MS and Orbitrap data.
[f] Number of SeMet residues incorporated per molecule is indicated in square brackets.
[g] Not detected.

TABLE S4

The geometry of the calcium binding site

Ligand	Distances (Å)[a]
Asp62 O	2.28 ± 0.03
Asn65 O	2.39 ± 0.03
Asp60 OD1	2.46 ± 0.07
Asp60 OD2	3.32 ± 0.31
W233 O	2.34 ± 0.03
Gln2 OE2	2.36 ± 0.10
W256 O	2.42 ± 0.05

[a] Average coordination distances taken from the two copies of structure of the native Mub-R5 repeat in the crystallographic asymmetric unit (± SD).

TABLE S5

Fluorescence signal intensities of f-Mubs and f-PpL bound to 20 µg target proteins immobilised on PVDF membranes

Target protein	Normalised mean fluorescence intensity (± SD)[a]				
	f-Mub-R5 [F/P 1.05; 9.8 µM][b]	f-Mub-R6 [F/P 0.99; 9.8 µM]	f-Mub- RI [F/P 1.41; 9.4 µM]	f-Mub-RI-III [F/P 2.37; 3.1 µM]	f-PpL [F/P 0.63; 0.5 µM]
IgA	502.74 (± 0.00)	92.00 (± 0.50)	294.29 (± 13.34)	206.89 (± 21.23)	56038.04 (± 1460.95)
IgM	463.46 (± 0.00)	108.14 (± 5.70)	168.11 (± 0.00)	108.52 (± 0.00)	57883.66 (± 918.93)
IgG	1099.06 (± 0.00)	244.34 (± 0.00)	630.37 (± 15.98)	1152.16 (± 5.43)	49796.41 (± 1193.42)
IgG-Fab	1136.24 (± 32.85)	192.87 (± 0.00)	608.50 (± 2.11)	867.06 (± 49.68)	46187.59 (± 676.57)
IgG-Fc	55.39 (± 0.00)	9.90 (± 0.00)	17.55 (± 0.00)	51.58 (± 25.91)	242.03 (± 0.00)
BSA	134.39 (± 0.00)	6.52 (± 0.00)	0.00 (± 0.00)	81.20 (± 33.51)	120.80 (± 0.00)

[a] Mean background-subtracted intensity values were normalised to an F/P ratio of 1.0 and a protein concentration of 1 µM as described in EXPERIMENTAL PROCEDURES (± SD, $n = 3$).
[b] F/P ratio and protein concentration are given in square brackets for each fluorescein-protein conjugate.

SUPPLEMENTAL FIGURE LEGENDS

FIGURE S1. **Purification of recombinant Mub domains.** SDS-PAGE of samples A, from freeze-thaw (FT) and BugBuster (BB) extracts and B, following FPLC purification on Mono Q. Samples were electrophoresed on 4-12 % NuPAGE Bis-Tris gels for 35 min at 200 V and gels stained with Coomassie Colloidal Blue. Arrowheads indicate positions of recombinant Mubs. *C*, FPLC purification of ~4 mg Mub-R5 on Mono Q.

FIGURE S2. **Far UV CD spectra of Mub domains and secondary structure determination.** FPLC-purified proteins were buffer-exchanged into Ultrapure H_2O to a concentration of 1 mg/ml and a portion transferred to a 0.1 mm path-length split cuvette. Six spectra were collected and averaged per sample over a scan range of 260-185 nm in a JASCO J-710 spectropolarimeter. Data were manipulated with JASCO Spectra Manager 32 v1.40.00a software and CONTINLL from the CDPro suite of programs was used to calculate the spectra and the proportion of each type of secondary structure. *A*, Mub domains R5, R6, RI and RI-III; *B*, SeMet- and unlabelled (UL) Mub-R5/Mub-R5(L48M); *C*, secondary structure determination.

FIGURE S3. **The environment of the calcium binding site in the B1 domain of Mub-R5.** The $(2F_o\text{-}F_c)$ electron density map is shown at 1.8 Å resolution and contoured at the 1.1 σ level. Ligand interactions in the octahedral coordination sphere of the calcium ion (grey sphere) are shown as dashed lines and the residues from which they originate are labelled. Coordination distances can be found in Supplemental Table S4.

FIGURE S4. **Alignment of the structures of Mub-R5 domains with the pilin protein, SpaB. A**, Residues 122-230 of the N2 domain of SpaB (GBS52), the minor pilin from the Gram-positive pathogen S*treptococcus agalactiae. B, C,* The B2 (R5-C) and B1 (R5-N) domains of mubR5, respectively. Secondary structural elements in the Mub domains are shown in red (α-helices) and yellow (β-strands). For SpaB, β-strands are coloured magenta. β-strands in the Mub domains which are structurally equivalent to the 3-stranded sheet in SpaB are labelled.

FIGURE S5. **Binding of Mubs and PpL to Igs immobilised on PVDF membranes.** The method is described fully in EXPERIMENTAL PROCEDURES. Probe F/P ratio and concentration for each fluorescein-conjugated protein were: f-Mub-R5, F/P 1.05, 9.8 μM; f-Mub-R6, F/P 0.99, 9.8 μM; f-Mub-RI, F/P 1.41, 9.4 μM; f-Mub-RI-III, F/P 2.37, 3.1 μM; f-PpL, F/P 0.63, 0.5 μM.