THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 286, NO. 36, pp. 31512–31521, September 9, 2011

Compensating Stereochemical Changes Allow Murein Tripeptide to Be Accommodated in a Conventional Peptide-binding Protein*[S]

Received for publication, June 1, 2011 Published, JBC Papers in Press, June 24, 2011, DOI 10.1074/jbc.M111.267179

Abbas Maqbool[‡], Vladimir M. Levdikov[§], Elena V. Blagova[§], Mireille Hervé[¶‖], Richard S. P. Horler[‡1], Anthony J. Wilkinson[§2], and Gavin H. Thomas[‡3]

From the [‡]Department of Biology (Area 10) and [§]York Structural Biology Laboratory, Department of Chemistry, University of York, York YO10 5YW, United Kingdom, the [¶]University Paris-Sud, Laboratoire des Enveloppes Bactériennes et Antibiotiques, IBBMC, UMR8619, Orsay F-91450, France, and the [‖]Centre National de la Recherche Scientifique (CNRS), Orsay F-91405, France

The oligopeptide permease (Opp) of *Escherichia coli* is an ATP-binding cassette transporter that uses the substrate-binding protein (SBP) OppA to bind peptides and deliver them to the membrane components (OppBCDF) for transport. OppA binds conventional peptides 2–5 residues in length regardless of their sequence, but does not facilitate transport of the cell wall component murein tripeptide (Mtp, L-Ala-γ-D-Glu-*meso*-Dap), which contains a D-amino acid and a γ-peptide linkage. Instead, MppA, a homologous substrate-binding protein, forms a functional transporter with OppBCDF for uptake of this unusual tripeptide. Here we have purified MppA and demonstrated biochemically that it binds Mtp with high affinity ($K_D \sim 250$ nM). The crystal structure of MppA in complex with Mtp has revealed that Mtp is bound in a relatively extended conformation with its three carboxylates projecting from one side of the molecule and its two amino groups projecting from the opposite face. Specificity for Mtp is conferred by charge-charge and dipole-charge interactions with ionic and polar residues of MppA. Comparison of the structure of MppA-Mtp with structures of conventional tripeptides bound to OppA, reveals that the peptide ligands superimpose remarkably closely given the profound differences in their structures. Strikingly, the effect of the D-stereochemistry, which projects the side chain of the D-Glu residue at position 2 in the direction of the main chain in a conventional tripeptide, is compensated by the formation of a γ-linkage to the amino group of diaminopimelic acid, mimicking the peptide bond between residues 2 and 3 of a conventional tripeptide.

Free-living bacteria have evolved routes to scavenge nutrients efficiently and often utilize peptides present in their environment (1). High-affinity bacterial transporters for di-, tri-, tetra-, and oligopeptides are well studied in bacteria and allow them to take up any peptide of a particular length (1, 2). The high-affinity peptide transporters are all members of the ATP-binding cassette (ABC) transporter family and utilize a substrate-binding protein (SBP),[4] which defines the substrate specificity of the transporter (1). The peptide-binding SBPs are larger than typical SBPs, but use a similar "Venus flytrap" mechanism of substrate capture whereby substrate recognition is accompanied by domain closure via a hinge region so that the peptide is completely buried in the binding site. On the basis of structural similarity the peptide-binding SBPs are assigned as cluster C proteins in the recent classification scheme of ABC SBPs by Berntsson *et al.* (3).

Escherichia coli has two well characterized high-affinity peptide transporters with broad specificity for substrates: the Opp system that recognizes peptides between 2 and 5 amino acid residues in length with highest affinity for tripeptides (4) and the Dpp system, which has a preference for dipeptides (5) (Fig. 1). The molecular basis of the sequence independence of peptide binding by these proteins was revealed by the solution of the crystal structures of OppA from *Salmonella typhimurium* and DppA from *E. coli* (6–9). Analysis of the binding site of *S. typhimurium* OppA revealed that the SBP forms strong interactions only to the main chain component of the bound peptide with the side chains projecting into hydrated pockets. The α-amino and α-carboxyl groups are salt bridged so as to partially restrict the length of the peptide that can be bound (6, 10).

The sequencing and annotation of the *E. coli* genome (11, 12) revealed the presence of a number of orthologues of the *oppA* and *dppA* genes on the chromosome including a number of "orphan" SBP coding sequences, which are genes encoding SBPs that are not part of an operon with the other components of an ABC transporter. Genetic data suggest that the orphan SBP, MppA, functions with the OppBCDF ABC transporter (Fig. 1) (13). The *mppA* gene was identified in 1998 as an additional locus in *E. coli* that was required for growth on the cell wall-derived murein tripeptide (L-Ala-γ-D-Glu-*meso*-Dap, where *meso*-Dap is *meso*-diaminopimelate) as a source of Dap for peptidoglycan biosynthesis (13). This MppA/OppBCDF

* This work was supported by a Higher Education Commission Pakistan studentship (to A. M.), Wellcome Trust Grant 082829/Z/07/Z (to V. L.), and European Commission SPINE2-COMPLEXES project Grants LSHG-CT-2006-031220 and BASYSBIO LSHG-CT-2006-037469 (to E. B.).

We dedicate this paper to the memory of our late colleague Berni Strongitharm.

⌘ *Author's Choice*—Final version full access.

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. S1–S6.

[1] Present address: Dept. of Microbiology, University of Illinois at Urbana-Champaign, Urbana, IL 61801.

[2] To whom correspondence may be addressed. E-mail: ajw@ysbl.york.ac.uk.

[3] To whom correspondence may be addressed. E-mail: ght2@york.ac.uk.

[4] The abbreviations used are: SBP, substrate-binding protein; GdmHCl, guanidinium HCl; Mtp, murein tripeptide; Dap, diaminopimelate; CV, column volume; ES-MS, electrospray mass spectrometry; PDB, Protein Data Bank.

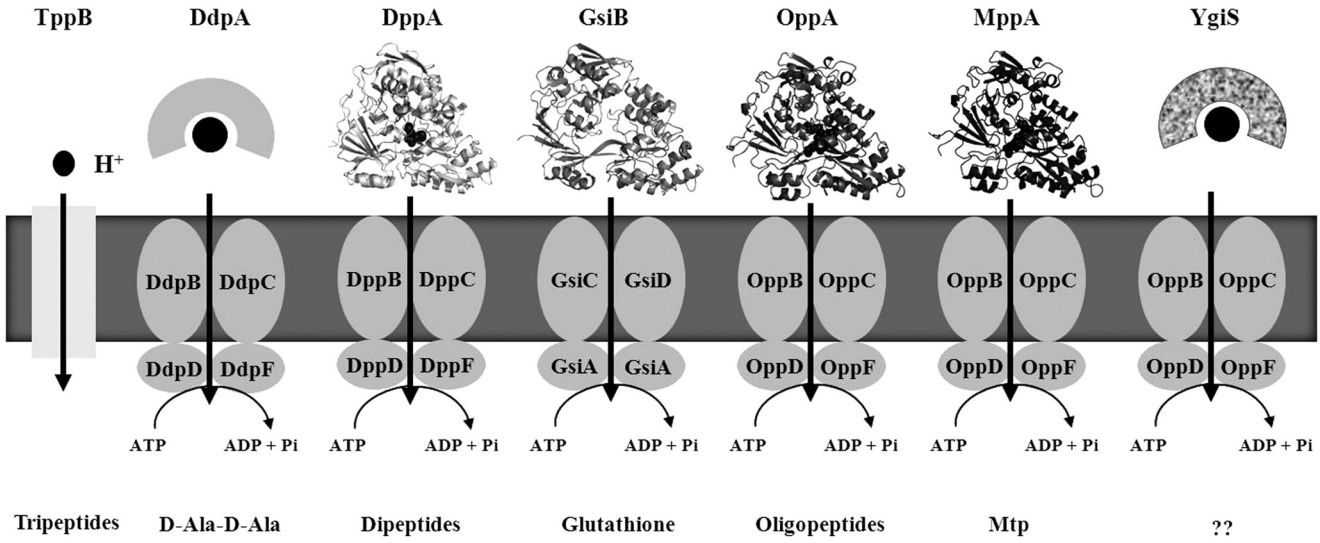

FIGURE 1. **Schematic of the known peptide transporters present in *E. coli* K-12.** This bacterium has a wide range of peptide transporters, mainly high-affinity SBP-dependent ABC systems, with a single secondary carrier TppB for tripeptides. The complete ABC transporters identified are a putative stationary phase-induced dipeptide permease (DdpABCDF), the dipeptide permease (DppABCDF), the glutathione-specific permease (GsiABCD), and the oligopeptide permease (OppABCDF) that can accept peptides from 2 to 5 residues in length. In ABC transporters, the SBP captures peptides from the periplasm and delivers them to membrane components for translocation. DdpA, DppA, GsiB, and OppA are the SBPs that work in conjunction with their cognate membrane components, whereas MppA and YgiS are orphan SBPs and are thought to use the OppBCDF permease components for solute translocation (13).

transporter presumably allows *E. coli* to recycle Mtp from either exogenous sources or Mtp that is released by the AmiD outer membrane amidase into the periplasm during cell growth (14). *E. coli* also contains a third characterized peptide ABC transporter, the Ddp system, which is thought to have evolved from the general dipeptide (Dpp) transporter system to recognize specifically D-Ala-D-Ala, another component of the peptidoglycan that can be recycled by the cell (Fig. 1) (15). Finally, a glutathione-specific transporter of the oligopeptide permease family has been characterized genetically in *E. coli*, the GsiABCD system (16) (Fig. 1). Therefore it is clear that the cluster 3 oligopeptide SBP family includes proteins that recognize general peptides, as well as others that have evolved to recognize specific peptides. In addition, peptides often have non-nutritional functions in signaling, one example being pheromone-dependent conjugation in *Enterococcus faecalis*, in which a small peptide sex pheromone is secreted by donor cells and taken up by recipient cells via a specific cluster 3 type SBP called PrgZ (17).

The structural basis of sequence independent peptide binding in transporter proteins has been well studied and much is understood. Here we are interested in how the atypical cell wall peptide Mtp is specifically accommodated within the framework of a SBP-fold. Given the importance of Mtp as a structural component of most bacterial cell walls, remarkably little is known about its structure. To gain insight into these questions we herein describe the biophysical and structural characterization the interactions of Mtp with the *E. coli* MppA protein.

EXPERIMENTAL PROCEDURES

Cloning, Expression, and Purification of MppA—The *mppA* gene was amplified from genomic DNA of *E. coli* K-12 MG1665 using the PCR primers: *mppA*F, 5′-CCAGGGACCAGCAATGGCAGAAGTTCCGAGCGGCACAGTACTGG-3′ and *mppA*R, 5′-GAGGAGAAGGCGCGTTATCAATGCTTCACAATATACATAGTCCGACTGTACG-3′ and cloned using a ligation-independent technique into the vector pET-YSBLIC3C (18) producing plasmid pAM609, which encodes a version of MppA lacking its native signal peptide and resulting in cytoplasmic expression. To construct A274E and R402A mutants of MppA, site-directed mutagenesis was carried out using the one-step site-directed and site-saturation mutagenesis protocol of Zheng *et al.* (19). The expression vector pAM609 was used as a template with the following primers: *mppAA274E*F, 5′-TATTATGAATTTAACACGCAAAAGGGCCCGACG-3′, *mppAA274E*R, 5′-GTTAAATTCATAATAATAGGTCCCGAGCTGCGG-3′, *mppAR402A*F, 5′-GATAGCGCGAACACCGGCAATTTTGATGTTATC-3′, and *mppAR402A*R, 5′- GGTGTTCGCGCTATCGATATAGGTCTTCCATTCCTG-3′. The sequences of mutants were confirmed by DNA sequencing of the plasmid.

For expression of recombinant genes, each construct was introduced into BL21(DE3) cells. The transformed cells were grown aerobically in 5 ml of Lennox broth (LB) for 8 h, and used to inoculate 50 ml of LB for overnight aerobic growth at 37 °C. This was then used to inoculate 625 ml of LB at 37 °C. Cells were allowed to grow aerobically to an A_{650} of 0.4–0.6 before production of recombinant protein was induced with 0.5 mM isopropyl 1-thio-β-D-galactopyranoside, followed by further overnight aerobic incubation at 37 °C. Cells were harvested by centrifugation at 4,430 × *g* for 15 min at 4 °C and resuspended in 50 mM potassium phosphate, pH 7.8, supplemented with 20% glycerol. The cells were lysed by sonication and cell debris was removed by centrifugation at 38,000 × *g* for 30 min at 4 °C and the supernatant containing overexpressed protein was collected. Nickel affinity chromatography was used for the purification of the hexahistidine-tagged protein using gravity-flow columns. The soluble fraction containing the protein of interest was mixed with nickel-nitrilotriacetic acid resin equilibrated

with buffer A (50 mM potassium phosphate, pH 7.8, 200 mM NaCl, 20% glycerol) supplemented with 10 mM imidazole. The mixture was incubated for 1 h at room temperature. Subsequently, the resin was washed with buffer A containing 50 mM imidazole, for 20 column volumes (CV) to remove weakly binding contaminants. Bound His-tagged protein was eluted with elution buffer containing 500 mM imidazole. To remove the His tag, purified protein was incubated with HRV3C protease in a ratio of 1:100 (protease to protein) at 4 °C overnight and then separated from the protease by a second Ni^{2+}-affinity chromatography step (5-ml His Trap column, GE Healthcare) where the protease is retained on the column through an N-terminal His tag. Four residues from the affinity tag (GPAM) remain attached to MppA following HRV3C cleavage. Protein concentrations were determined from the absorbance at 280 nm (A_{280}) using a calculated molar extinction coefficient for MppA of 99,570 cm^{-1} M^{-1}. All experiments used the cleaved MppA. OppA was expressed and purified according to procedures described previously (20).

Preparation of Ligand-free MppA—Co-purified ligand (Mtp) was removed from MppA using guanidinium HCl (GdmHCl) as described previously (21) to enable the ligand-free protein (termed MppA*) to be used for binding studies with peptide ligands. The protein was partially unfolded while bound to the nickel-nitrilotriacetic acid resin, by including additional wash steps (in buffer A with 30 mM imidazole) of 40 CV with 2 M GdmHCl, 4 CV with 1.5 M GdmHCl, 4 CV with 1 M GdmHCl, 4 CV with 0.5 M GdmHCl, and finally 8 CV with 0 M GdmHCl. The protein (MppA*) was eluted with elution buffer as described previously.

Protein Analysis by Mass Spectrometry—A Micromass LCT Premier XE mass spectrometer using a nanospray source was used for electrospray mass spectrometry (ES-MS). For analysis of MppA under denaturing conditions, purified MppA (after cleavage of the His tag) was dialyzed into 50 mM potassium phosphate buffer, pH 7.8, and concentrated to 100 μM. Concentrated protein was diluted to 1 μM in (1:1) acetonitrile/water containing 0.1% formic acid. The sample was passed into the ES-MS and data were collected over 3 min within an m/z range of 500–2,000. To investigate the mass of any endogenous ligand bound to MppA, a spectrum was obtained under native conditions. For analysis under native conditions purified MppA was dialyzed into dH_2O and then concentrated to 100 μM. Data were collected over 5 min within an m/z range of 500–10,000. The raw m/z data were deconvoluted to mass spectra using the MaxEnt1 routine in the MassLynx software provided by the manufacturer (Waters).

Crystallization and Data Collection—For crystallization, MppA was concentrated to 16 mg/ml in 20 mM MES, pH 6.5. Crystallization trials were performed using a range of commercially available screens: PACT, Hampton I, II, Index, and Clear Strategy Screens CSSI and CSSII (22), by hanging-drop vapor diffusion procedures in 96-well plates setup with a Mosquito nanoliter pipetting robot (TTP Labtech). The plates were sealed to avoid evaporation and then incubated at 20 °C and examined regularly for the appearance of crystals. MppA produced crystals in drops suspended over wells containing 50 mM zinc acetate, 20% (w/v) PEG3350. The crystals were transferred to a precipitant solution containing 25% glycerol for cryoprotection, mounted in a nylon loop, and flash cooled in liquid nitrogen. One of the crystals that gave good diffraction in-house was used for x-ray diffraction data collection on beamline ID14-4 at the European Synchrotron Radiation Facility (ESRF), Grenoble (Table 1). Data processing was carried out using the HKL2000 program. The crystal belonged to space group P6 with unit cell parameters $a = b = 165.1$ Å, $c = 38.2$ Å.

TABLE 1
MppA x-ray data collection and refinement statistics

Data collection	
Wavelength (Å)	0.9335
Resolution range (Å)	47.65–2.07
Space group	P6
Unit cell parameters (Å)	$a = b = 165.1$ Å, $c = 38.2$ Å, $\alpha = \beta = 90$, $\gamma = 120$
Number of unique reflections, overall/outer shell[a]	34,597/633
Completeness (%), overall/outer shell[a]	92/53.5
Redundancy, overall/outer shell[a]	9.7/4.3
$I/\sigma(I)$, overall/outer shell[a]	20.6/1.6
R_{merge}[b] (%), overall/outer shell[a]	10.7/55.3
Refinement and model statistics	
R_{factor}[c] (R_{free})[d]	16.7 (21.9)
Reflections (working/free)	32,851/1743
Outer shell R_{factor}[c] (R_{free})[d]	24.2 (30.4)
Outer shell[e] reflections (working/free)	1697/74
Molecules/asymmetric unit	1
Number of protein nonhydrogen atoms	4046
Number of ligand atoms	50
Number of water molecules	366
Root mean square deviation from target[f]	
Bond lengths (Å)	0.02
Bond angles (°)	1.84
Average B-factor (Å^2)	25
Ramachandran plot[g]	91.3/8.5/0.2/0.0

[a] The outer shell corresponds to 2.10–2.06 Å.
[b] $R_{merge} = \Sigma_{hkl}\Sigma_i|I_i - \langle I\rangle|/\Sigma_{hkl}\Sigma_i\langle I\rangle$, where I_i is the intensity of the ith measurement of a reflection with indexes hkl and $\langle I\rangle$ is the statistically weighted average reflection intensity.
[c] $R_{factor} = \Sigma||F_o| - |F_c||/\Sigma|F_o|$, where F_o and F_c are the observed and calculated structure factor amplitudes, respectively.
[d] R_{free} is the R_{factor} calculated with 5% of the reflections chosen at random and omitted from refinement.
[e] Outer shell for refinement corresponds to 2.073–2.126 Å.
[f] Root mean square deviation of bond lengths and bond angles from ideal geometry.
[g] Percentage of residues in most favored/additionally allowed/generously allowed/disallowed regions of the Ramachandran plot, according to PROCHECK.

Structure Determination, Model Building, and Refinement—The structure of MppA was determined to a resolution of 2.1 Å by the molecular replacement method using the coordinates of the oligopeptide-binding protein from *Yersinia pestis* (PDB code 2Z23) as a search model. Molecular replacement was carried out using the MOLREP program (23). Manual model building was carried out in the program COOT (24) and the structure was refined using REFMAC5 (25) to give final R_{factor} and R_{free} values of 16.7 and 21.9%, respectively. The data collection and refinement statistics are listed in Table 1. Structural data have been deposited in the RCSB Protein Data Bank under the accession code 3O9P.

Size Exclusion Chromatography Multi-angle Laser Light Scattering (SEC-MALLS)—For SEC-MALLS, samples of MppA (120 μl of 1 mg/ml in 20 mM MES buffer, pH 6.5) were applied to a Superdex 200 10/300 GL SEC column (GE Healthcare) at 0.5 ml/min with an HPLC system (Shimadzu), linked to a Wyatt Dawn Heleos MALLS detector and a Wyatt Optilab rEX unit, to determine the refractive index. The buffer was 30 mM MES, pH

6.5, 200 mM NaCl. Data were analyzed with the program Wyatt ASTRA version 5.3.4.14.

Preparation of Murein Tripeptide—The tripeptide L-Ala-γ-D-Glu-*meso*-Dap was generated from the corresponding precursor UDP-MurNac-tripeptide as previously described (26). First, mild acid hydrolysis (10 min at 100 °C in 0.1 M HCl) of UDP-MurNac-tripeptide-generated MurNac-tripeptide, which was purified by HPLC as a mixture of the two anomers α and β; then, MurNac-tripeptide was hydrolyzed to MurNac and tripeptide by incubation with purified *E. coli* AmiD, an *N*-acetylmuramoyl-L-alanine amidase (27). The released tripeptide was subsequently purified by HPLC on a 3-μm ODS-Hypersil column (0.46 × 25 cm) in 0.05% TFA at a flow rate of 0.5 ml min^{-1} and quantified by amino acid analysis with a Hitachi L8800 analyzer (ScienceTec) after hydrolysis in 6 M HCl for 16 h at 95 °C.

Fluorescence Spectroscopy—All protein fluorescence experiments used a FluoroMax 4 fluorescence spectrometer (Horiba Jobin-Yvon) with connecting water bath at 25 °C. Purified MppA* protein was used at a concentration of 0.05 μM in 50 mM potassium phosphate, pH 7.8, and excited at 297 nm with slit widths of 3.5 nm. Emission was monitored at 330 nm with slit widths of 3.5 nm. To examine alternative ligands for MppA, potential ligands were added at concentrations up to 1 mM. To determine the K_D for ligand binding, the protein was titrated with increasing concentrations of the ligand and the corresponding fluorescence change was monitored in time acquisition mode. The cumulative fluorescence change was plotted in SigmaPlot and the K_D was calculated from the hyperbolic fit of the binding curve. L-Ala-γ-D-Glu-L-Lys (AEK) and PFK were synthesized commercially by Alta Bioscience (Birmingham, UK).

Isothermal Titration Calorimetry—Calorimetry experiments were carried out using a VP-ITC instrument (MicroCal Inc., GE Heath Sciences). In a typical experiment, the calorimeter cell contained 1.4 ml of MppA, and the syringe contained 300 μl of peptide ligands. The concentration of ligand in the syringe was typically 10 times that in the cell, whereas the cell concentration was chosen according to *c* value of greater than 10, where c = [macromolecule]/(predicted) K_D. Experiments were carried out in 50 mM potassium phosphate buffer, pH 7.8, at 25 °C. Both cell and syringe solutions were degassed at 20 °C for 10 min before use. The titrations were performed as follows. A single preliminary injection of 5 μl of ligand solution was followed by 28 injections (10 μl), delivered at an injection speed of 10 μl s^{-1}. Injections were made at 3-min intervals with a stirring speed of 307 rpm. For titrations involving PFK a heat of dilution control experiment, where equivalent volumes of peptide were injected into buffer, was performed. This ruled out the possibility of heats of dilution masking weak binding. Raw titration data were integrated and fit to a one-site model of binding using MicroCal Origin version 7.0.

RESULTS

Expression and Purification of MppA—The *E. coli mppA* gene was expressed in *E. coli* producing a cytoplasmic N-terminal His_6-tagged protein, which was purified using Ni^{2+}-affinity chromatography. The His_6 tag was removed using HRV3C protease producing native MppA with four additional amino acid residues at the N terminus. The sample was ~99% pure as judged by Coomassie staining of SDS-PAGE gels (Fig. 2*A*, *inset*). The levels of protein expression were high and the yield obtained after Ni^{2+}-affinity chromatography was 26 mg/liter. Using denaturing ES-MS we determined the molecular weight of the protein to be 57,979 ± 6 Da (Fig. 2*A*), which is consistent with the predicted mass of the protein after removal of the His_6 tag (57,972 Da). We examined the purified protein using SEC-MALLS (supplemental Fig. S1) from which we were able to calculate a native mass for MppA of 60.6 ± 1.5 kDa. This is close to the calculated molecular mass for the MppA polypeptide at 58 kDa, suggesting that like other peptide-binding SBPs, it is a monomer. The SEC-MALLS data also demonstrated that the sample was monodisperse and suitable for crystallization.

Overall Structure of MppA—To understand how MppA functions as an Mtp-binding protein, we crystallized the recombinant protein as purified from the *E. coli* cytoplasm without attempting to influence the ligation state. Solute-binding proteins of ABC transporters invariably co-purify with their cognate ligands as a consequence of the high ligand affinity and the fact that the proteins are maintained at high concentrations throughout purification. X-ray diffraction data were collected from these crystals and the structure was solved by molecular replacement using the coordinates of a homologous oligopeptide-binding protein (PDB code 2Z23) as a search model. The refined model contains a continuous polypeptide chain from Thr-7 to the C-terminal residue His-515 (Fig. 3*A*). The first 6 residues are not represented in the electron density maps and are assumed to be disordered. MppA consists of three domains, with domain 1 encompassing residues 7–46, 165–263, and 487–515, domain 2 encompassing residues 264–486, and domain 3 encompassing residues from 47 to 164 (Fig. 3*A*). A DALI search (28) of the Protein Data Bank (PDB) revealed that, as expected, MppA has the highest structural similarity to the oligopeptide-binding proteins (OppAs) of *Y. pestis* (PDB code 2Z23; *Z*-score = 55.3, with 508 equivalent Cα atoms superimposing with a rmsΔ of 1.1 Å) and *Salmonella typhimurium* (PDB code 2OLB; *Z*-score = 55.2, with 509 equivalent Cα atoms superimposing with an rmsΔ of 1.0 Å). This structural similarity spans the full-length of the protein (supplemental Fig. S2).

The Structure of MppA Contains Bound Mtp—Examination of the interdomain region in MppA revealed electron density that could not be accounted for by protein atoms. Usually the electron density for bound peptides in peptide-binding proteins is ambiguous because a heterogeneous mixture of the peptides is present (6, 29, 30), but here we were able to unambiguously model and refine Mtp into the electron density (Fig. 3*B*), consistent with the idea that MppA binds Mtp selectively.

To confirm that the bound species was Mtp, we performed native electrospray mass spectrometry analysis on MppA, which are under conditions in which SBPs often retain endogenously bound ligands (31, 32). Here we observed a clear peak in the mass spectrum (Fig. 2*B*) that was not present in the denaturing spectrum (Fig. 2*A*). This additional peak accounts for nearly half of the total protein signal. This species has a mass of 58,361, being 389 mass units larger than that of the ligand-free protein (57,972 mass units in this sample). Given that the mass

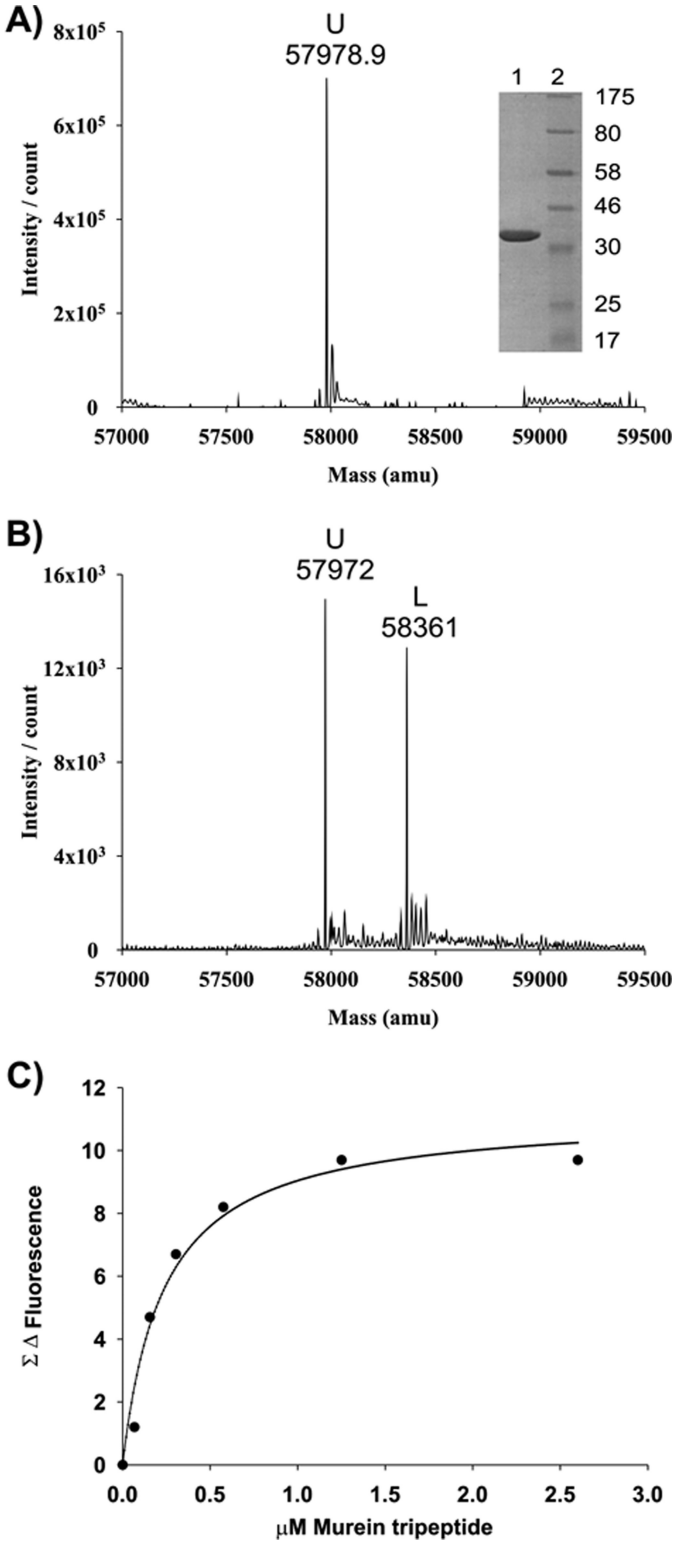

FIGURE 2. **Biophysical measurements of Mtp binding to MppA.** *A*, ES-MS spectrum of MppA under denaturing conditions. The molecular mass of the unbound protein (*U*) is indicated at 57,978.9 atomic mass units (*amu*). *Inset*, Coomassie-stained SDS-PAGE gel containing 3 μg of purified MppA (*lane 1*) and molecular weight markers (*lane 2*). *B*, ES-MS spectrum of MppA under native conditions in water. Peak *U* gives the molecular weight of ligand-free MppA and peak *L* gives the molecular weight of MppA bound to a ligand (Mtp) of 390 Da. Other peaks are adducts of phosphates. *C*, cumulative changes in tryptophan fluorescence of MppA upon titration with Mtp.

of Mtp is 389.38, this is strong evidence that the protein binds a single molecule of Mtp specifically and combined with the structural data provides conclusive evidence that MppA is an Mtp-binding protein.

The Mtp Binding Site in MppA—The crystal structure reveals MppA in a "closed" conformation with a single molecule of Mtp completely buried within the protein, as seen with other peptides in oligopeptide-binding protein structures (6, 8, 29, 30). The Mtp sits in a clear binding pocket, forming multiple interactions with the protein (Fig. 4*A*). The N-terminal amino group of the peptide forms a salt bridge with the side chain carboxylate of Asp-417 and a hydrogen bond to the main chain carbonyl of Val-415 in MppA. A prominent interaction is a bivalent salt bridge that is formed between the free α-carboxylate from the γ-linked D-Glu component of Mtp and Arg-402 in MppA. Additional charge-charge and charge-dipole interactions are made to the first carboxylate group in the *meso*-Dap component via Lys-20, Arg-411, and Ser-413 in the protein. At the C terminus of the tripeptide, the second carboxylate of the Dap is coordinated by two water molecules. Meanwhile the C-terminal amino group interacts with two water molecules and the side chain carbonyl of Gln-267 (Fig. 4*A*).

Comparison of the Binding Sites in MppA to OppA—To compare the binding site of a general oligopeptide-binding protein with that of MppA, we superposed the backbones of OppA and MppA and analyzed the binding of the two ligands, KEK for OppA and Mtp for MppA (Fig. 4*B* and supplemental Fig. S3 for a structure based sequence alignment). The proteins are 85% similar overall at the sequence level and hence certain parts of the binding site are very well conserved between these proteins, the first being the Asp-417 residue in OppA (Asp-419), which forms a salt bridge to the α-ammonium group of the peptide. This "anchors" the N terminus of the peptide and in OppA it is known that a protonated N-terminal amino group on the ligand is essential for transport, suggesting that this is also likely to be a key interaction in MppA (33). It is apparent that the overall shape of Mtp in the binding pocket of MppA is surprisingly similar to that of KEK in OppA, such that the γ-linked D-Glu adopts a similar conformation to the more typical peptide linkage seen in the KEK tripeptide (Fig. 4, *B* and *C*).

The Arg-402 residue that binds to the carboxylate group of the γ-linked D-Glu is also conserved in OppA (Arg-404). However, in OppA, it is positioned differently as a result of its interaction with Glu-276 (Fig. 4) so that it does not form a direct contact to the peptide. In OppA, Arg-404 coordinates H_2O molecules that in turn form polar contacts to the ligand (6). Water molecules, which are abundant in the OppA binding pocket, allow the protein to accommodate variability in its ligand side chains. This is because the water network can be adapted/altered according to the particular ligand present. However, in MppA, where Glu-276 is replaced by alanine (Ala-274), the Arg-402 side chain is more prominent in the binding pocket so that its guanidine group forms what are expected to be strong ion-pairing interactions with the α-carboxylate group of the D-Glu in Mtp (Fig. 4). The substitution of Ala for Glu at this position is a clear distinguishing feature of the MppA and OppA proteins (supplemental Fig. S4).

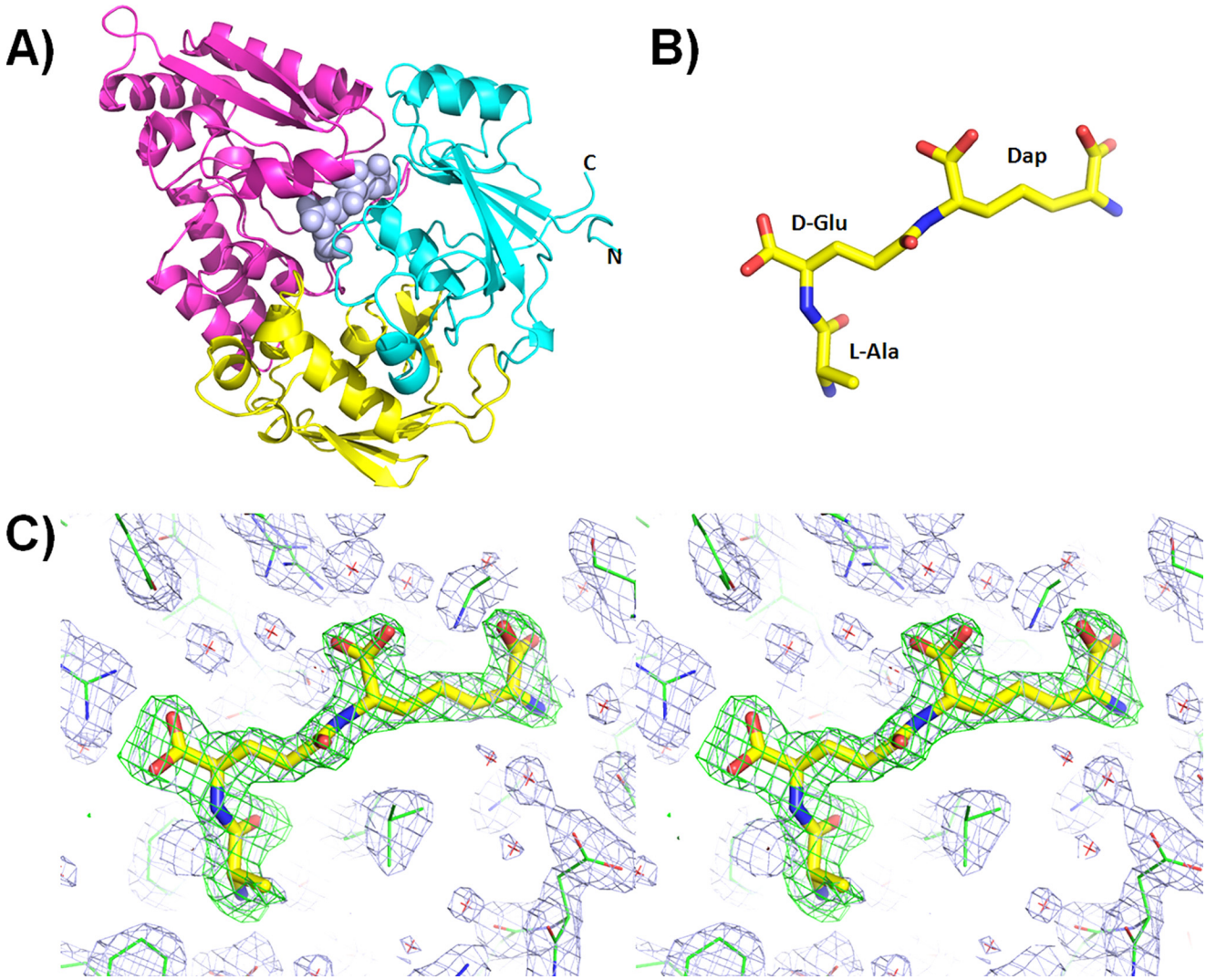

FIGURE 3. **Crystal structure of the MppA·Mtp complex.** *A*, ribbon representation of the MppA structure. The three-domain organization; domains I, II, and III are colored in *cyan*, *magenta*, and *gold*, respectively. The amino (N) and carboxyl (*C*) termini are indicated. The Mtp ligand is displayed in space-filling format. *B*, the structure of Mtp is represented as *stick* format. Carbon atoms are colored *yellow*, oxygen in *red*, and nitrogen in *blue*. *C*, stereo image of OMIT maps calculated with coefficients $F_{obs} - F_{calc}$ (*green*) contoured at the 3 σ level and with coefficients $2F_{obs} - F_{calc}$ (*light blue*) contoured at 1 σ level, and displayed in the vicinity of the Mtp ligand. Temperature factors for Mtp atoms in the final model were refined with full occupancies and are in the range of 19–33 Å^2, similar to that of surrounding protein atoms.

A second conserved residue Arg-411 (Arg-413 in OppA) also interacts strongly with the ligand. In OppA, Arg-413 either forms a salt bridge to the C-terminal carboxylate of the tripeptide KEK or a hydrogen bond to the main chain carbonyl of residue 3 in the tetrapeptide KKKA (20). In MppA where there is not a typical amino acid at the third position in the peptide, it is the first carboxylate of the *meso*-Dap that now forms a salt bridge to the corresponding Arg residue. The side chain ϵ-NH_3^+ group of Lys-20 also forms electrostatic interactions with this carboxylate in MppA. The corresponding residue is a glycine in OppA so this particular contact to the ligand is unique to MppA.

MppA Binds Mtp with High Affinity—To learn more about the interaction of MppA with Mtp, we investigated ligand binding to the protein using tryptophan fluorescence spectroscopy, a method that exploits the significant conformational change that accompanies ligand binding. For these experiments, we prepared ligand-free protein (MppA*), the absence of ligand being confirmed using native electrospray mass spectrometry (data not shown). Binding of Mtp to MppA causes an ~7% enhancement in fluorescence upon binding. Titration experiments yielded a K_D of 0.24 ± 0.02 μM (Fig. 2*C*).

We also used isothermal titration calorimetry to assess binding of Mtp to MppA, which yielded a K_D of 0.30 ± 0.06 μM, consistent with the value obtained from tryptophan fluorescence spectroscopy (Fig. 5*A*). The ITC data also confirm the 1:1 Mtp:MppA binding stoichiometry (0.99 ± 0.01) and demonstrate a positive enthalpy change upon Mtp binding (ΔH° of 2092 ± 27 kJ mol^{-1} and ΔS° of 36.9 J·mol^{-1}·K^{-1}) (Fig. 5*A*), suggesting that binding is entropy driven as observed previously for OppA (10).

Mtp Binding to MppA Is Abolished in Arg-402 and Ala-274 Mutants—To investigate the role of the Arg-402 residue in MppA whose side chain binds to the carboxylate group of the γ-linked D-Glu, we mutated this residue to Ala and investigated the effects on ligand binding. We also mutated Ala-274 to Glu

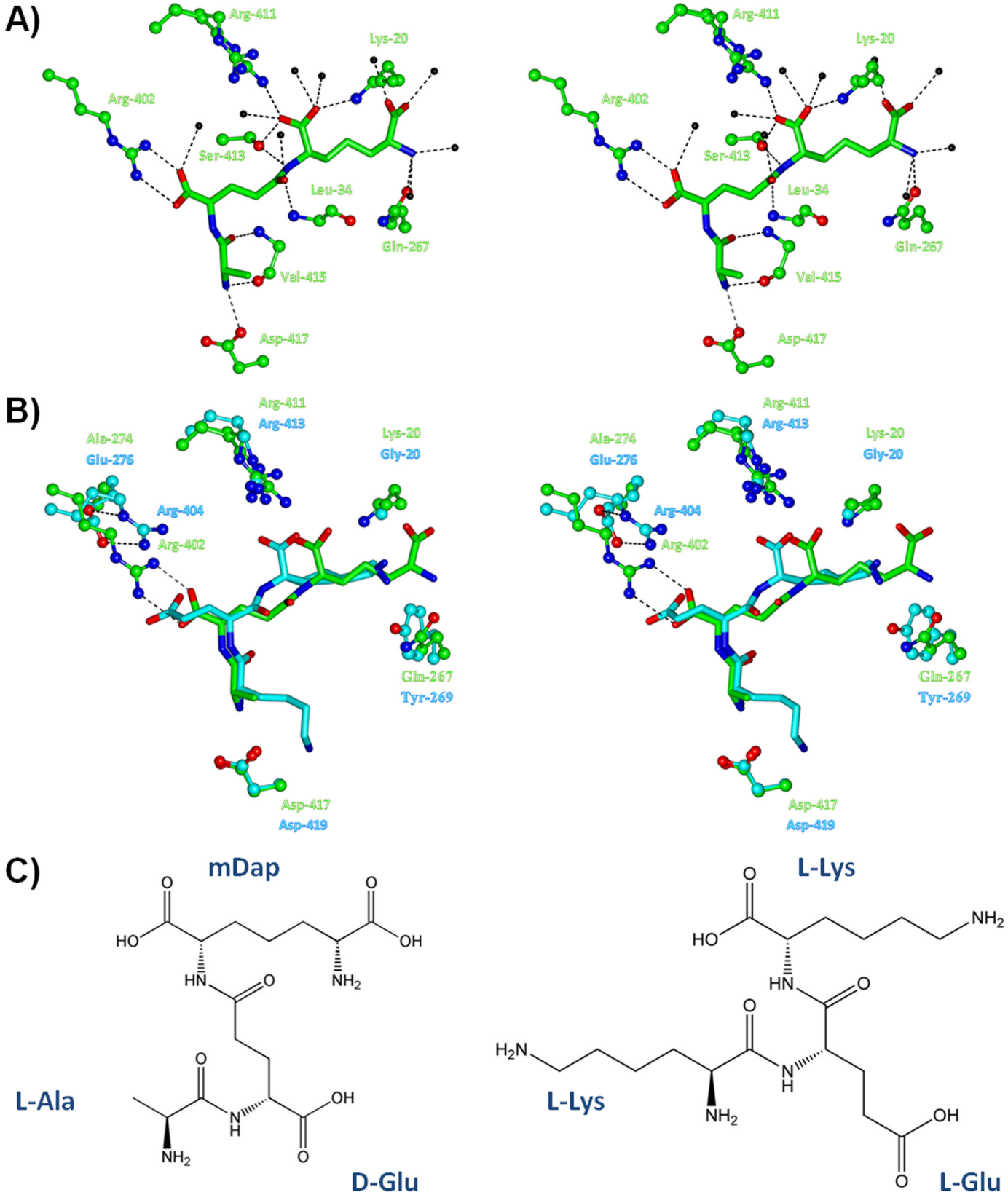

FIGURE 4. **Ligand binding in MppA and OppA.** *A*, binding site of MppA containing Mtp. The amino acids involved in coordinating the ligand are represented in *ball and stick* format and are labeled, and the ligand is represented in *cylinder format*. Hydrogen bonding interactions are indicated with *dashed lines*. Water molecules interacting with ligand are represented as *black spheres*. *B*, structural superposition of the ligand binding sites of MppA and OppA. The amino acid residues in the region of substrate-binding sites of MppA (*green*) and OppA (*cyan*) are compared and labeled by their positions in PDB files 3O9P and 1JEU, respectively. Mtp and KEK are represented in *cylinder format*, and amino acid residues of MppA and OppA are represented in *ball and stick format*. *C*, chemical structures of Mtp and the KEK tripeptide. The component amino acids are labeled.

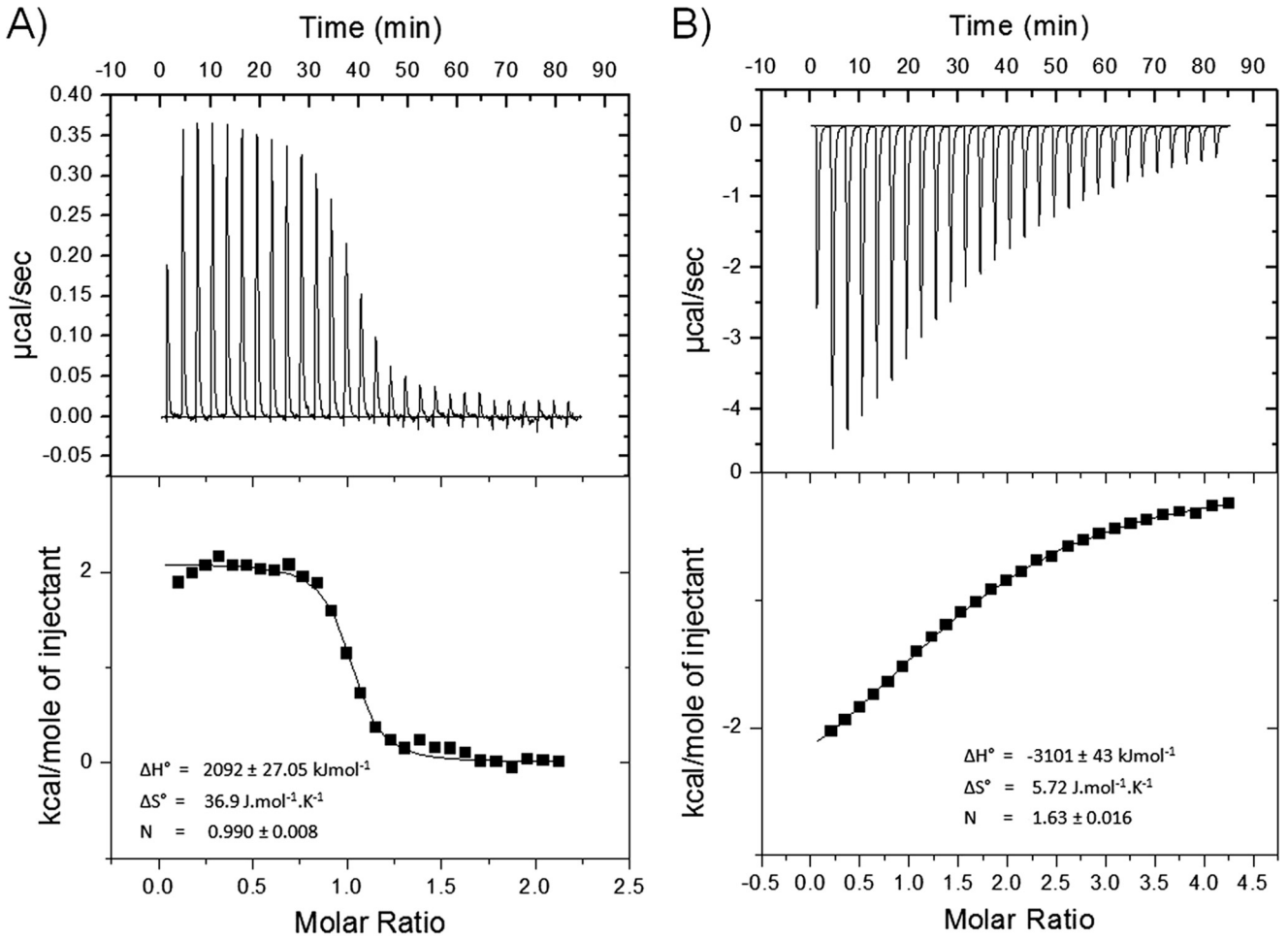

FIGURE 5. **Binding isotherms for the interaction of MppA with Mtp (*A*) and PFK (*B*).** The *top panels* show heat differences upon injection of ligands and *lower panels* show integrated heats of injection (■) and the best fit (*solid line*) to a single site binding model using Microcal Origin.

to assess whether an OppA-like residue at this position could sequester the Arg-402 side chain and prevent it from making direct contact with Mtp. Both mutations resulted in MppA proteins that were unable to bind Mtp as judged by tryptophan fluorescence spectroscopy. Also, the purified mutant proteins did not purify as complexes with Mtp as judged by ES-MS (supplemental Fig. S5), suggesting that both residues are important for MppA recognition of Mtp. As we could not detect binding of ligand to either mutant, we recorded the circular dichroism spectra of the mutant proteins. Both mutants give very similar spectra to the wild-type MppA (data not shown) indicating that failure to bind Mtp is not caused by protein misfolding or instability.

Specific Recognition of Mtp by MppA—To determine the specificity of the interaction between Mtp and MppA we assessed whether MppA was able to bind a number of closely related ligands. We first examined the binding of the synthetic peptide, L-Ala-γ-D-Glu-L-Lys (AEK), in which the Dap in Mtp is replaced by lysine. As measured by tryptophan fluorescence spectroscopy, this tripeptide bound to MppA with a K_D of 130 ± 10 μM, which is ~500-fold lower affinity than for Mtp. This points to an important role for the side chain carboxylate of Dap in binding to MppA. We next examined a range of conventional tripeptides that have been used in studies on OppA including KEK, KKK, AAA, KKKA, and AKK. For none of these peptides were we able to detect binding to MppA (data not shown).

To assess whether OppA can bind Mtp and similar peptides, we purified ligand-free OppA protein as previously described (20). The protein was able to bind a typical tripeptide, KEK with submicromolar affinity but we could not detect binding of Mtp by ITC (supplemental Fig. S5), supporting the genetic data that suggest that the Opp system with OppA as the SBP is unable to transport Mtp (13).

We also investigated a peptide L-Ala-D-Glu-*meso*-Dap, where the D-stereochemistry of residue 2 is not accompanied by a γ-peptide linkage. We were unable to detect binding of this ligand to either MppA or OppA (data not shown).

It has been demonstrated previously, using genetic methods, that a peptide transporter using MppA as the SBP is able to support growth of an *E. coli* proline auxotroph on the tripeptide L-Pro-L-Phe-L-Lys (PFK) (13). We examined direct binding of this tripeptide to MppA, but were unable to detect binding even using concentrations up to 1 mM using tryptophan fluorescence (data not shown). It is possible that PFK binding to MppA does not result in a net change in the tryptophan fluorescence of the protein, and so we measured PFK binding to MppA using isothermal titration calorimetry (Fig. 5). We were able to detect an

interaction between MppA and PFK, although the binding was weak with a K_D greater than 300 μM and the binding isotherm was markedly different from that observed for Mtp binding (ΔH° of -3100 ± 43 kJ mol^{-1} and ΔS° of 5.72 J $\cdot$ mol^{-1} $\cdot$ K^{-1}) (Fig. 5*B*), suggesting that the binding mechanism is different, perhaps indicative of nonspecific binding to the protein. It is certainly not consistent with a previous report that suggested a K_D of MppA for PFK less than 50 μM (34). Together these data suggest that high-affinity binding in the submicromolar range is seen only for Mtp itself and even small modifications of the ligand structure result in significant reductions in affinity.

DISCUSSION

Short peptides present in the environment of many bacteria are recognized with high affinity by OppABCDF oligopeptide transporters. *E. coli* contains paralogues of OppA, which are orphan SBPs that combine with the OppBCDF membrane components to form complete transporters. MppA is one such orphan, and for this SBP there is additional genetic evidence that it can function with both the Opp and Dpp systems (34). In this study we have shown conclusively, both biochemically and structurally, that MppA is a murein tripeptide-binding protein.

OppA is a versatile peptide-binding protein that binds peptides of variable length with extensive sequence diversity. Genetic data demonstrate, however, that it is unable to support the transport of Mtp (13), a conclusion consistent with the biochemical analysis in this work. Mtp is unusual as it contains an L-Ala linked to a D-Glu, which in turn has a γ-linkage to the diaminopimelate (*meso*-Dap) moiety at the C terminus (Figs. 3*B* and 4*C*). It is synthesized non-ribosomally using the MurC-F ligases (14). Intuitively, one would not expect a peptide with D-stereochemistry at position 2 to be accommodated within an OppA-like protein-fold. A change in amino acid chirality at an internal position in a peptide is a major alteration that would be expected to have dramatic structural consequences. In Mtp, the residue 2 carboxylate projects in the direction otherwise occupied by the side chain and vice versa. In the MppA·Mtp complex, extension of such a peptide through an α-linkage would lead to clashes with Arg-402. Similarly, a peptide with a γ-linkage to a Glu will differ markedly from a conventional peptide. In Mtp, the γ-linkage to Glu-2 on its own would cause clashes with the MppA in the vicinity of Arg-402. Remarkably, as the structure of Mtp in the MppA·Mtp complex shows, the combination of these two changes within one peptide represents a much smaller structural perturbation than either change alone, and the murein tripeptide is readily accommodated in the OppA protein framework. The changes are in fact compensating. The free main chain carboxylate of Glu-2 is accommodated by MppA in what would be the position 2 side chain binding pocket of OppA, whereas the side chain of this Glu takes the direction followed by the main chain in OppA. This is of minimal consequence as this side chain takes on main chain character through its amide linkage to Dap (Fig. 4*B*).

Despite the overall positioning of the ligands in the binding site being shared in MppA and OppA, there are differences in the mode of binding. Tame *et al.* (6) proposed that OppA achieved essentially sequence independent peptide binding by minimizing the formation of direct interactions between the protein and ligand side chains. This is because favorable interactions with one side chain type are likely to be unfavorable with a different side chain type leading to ligand discrimination. This binding model has been borne out in the structures of other general peptide-binding proteins (29, 30). It is striking that in MppA, the binding site contains an additional specific electrostatic interaction between Mtp and the protein, namely the interaction between Arg-402 and the carboxylate of the D-Glu. MppA, whose distribution is limited to the enterobacteriaciae, has most likely evolved via gene duplication and diversification from an ancestral *oppA* gene. One of the sequence changes that appears to be important in the evolution of MppA from an OppA-like ancestor is the replacement of Glu-279 by Ala. This substitution "releases" Arg-404 (now Arg-402 in MppA) from an ion pair with the Glu in OppA so that it can adopt a conformation in MppA that enables it to form a direct side chain interaction with the ligand and so restrict the peptides that can be recognized by this SBP. Our mutagenesis studies demonstrate that both residues are essential for binding of Mtp by MppA, which support the predictions that the contacts to the free carboxylate of the D-Glu are important for binding.

As well as MppA, there are other cluster C SBPs that are likely to recognize specific peptides as the recent characterization of the glutathione ABC transporter from *E. coli* GsiABCD illustrates (16). Although there is a structure for the SBP from this system (PDB code 1UQW) it is not ligand bound and so the basis of its specificity is not known. Interestingly, some amino acid-binding SBPs (cluster F) have also evolved to recognize specific dipeptides. The GmpC protein from *Staphylococcus aureus* binds specifically to the dipeptide glycylmethionine, yet is related in sequence to the *E. coli* methionine-binding protein, MetQ (35). Although this is an interesting case, it is unlikely that the cluster F SBPs would accommodate larger peptides like the cluster C SBPs. The flexibility of the cluster C SBPs that allows them to bind large ligands has also been demonstrated recently by the discovery of SBPs from this cluster that are oligosaccharide-binding proteins. These are mainly from Archaea but others have been recently characterized from *Thermatoga maritima*, although they may well have transferred laterally from the Archaea (36). In the structure of the TM0031 protein from *T. maritima* (37) the ligand cellopentaose (which is 5β(1$\rightarrow$4)-linked glucose molecules) sits in a similar position to the oligopeptides in OppA and other cluster C SBPs.

Our *in vitro* analysis of ligand binding to MppA clearly supports the hypothesis that Mtp is the physiological ligand for this protein, however, a number of other studies have suggested additional substrates for MppA using a range of genetic and biochemical approaches. It has been suggested that MppAOppBCDF can transport the standard tripeptide Pro-Phe-Lys (PFK) (13, 34). However, we were unable to detect binding of PFK by fluorescence spectroscopy and could detect only weak binding by ITC. It is possible that PFK binds to a site on the protein distinct from that to which Mtp is bound as the thermodynamics of PFK binding were quite different from Mtp. Whereas Mtp binding to MppA and tripeptide binding to OppA are entropy driven, PFK binding to MppA is enthalpy driven. Certainly, we could see no evidence for high-affinity

binding of PFK to MppA, which does not agree with the previous studies (13, 34). It has also been reported that MppA can bind heme and its precursor δ-aminolevulinic acid, however, whether this is of physiological relevance is unclear as heme cannot diffuse across the outer membrane of *E. coli* and its binding is competitive with PFK, which we have already demonstrated binds much more weakly to MppA that Mtp. Indeed we were unable to observe binding of δ-aminolevulinic acid to MppA via tryptophan fluorescence spectroscopy (data not shown). Also, the significant drop in binding affinity that we observed with the AEK peptide, which alters only the water-mediated hydrogen bond network in the binding site, suggests that recognition of Mtp is very specific. Together these data suggest that MppA substrates are restricted to Mtp under normal conditions in the cell, which is consistent with our simple observation that MppA copurifies from *E. coli* with Mtp bound.

Although the specificity of MppA for Mtp is clearly established, the physiological role of the MppA protein is not so well understood (13), although it most likely serves to allow *E. coli* to scavenge Mtp that is either released during turnover of its own cell wall or is available in the environment. This suggestion is supported by the fact that *E. coli* is known to secrete Mtp during growth (38). It is also clear that there must be a free pool of Mtp inside *E. coli* as cytoplasmically expressed recombinant MppA was purified in high yield as a complex with Mtp. This is consistent with the suggestion that Mtp is the ultimate breakdown product of murein, which is recycled as a building block for further synthesis (14). Although the majority of murein is recycled via AmpG in the form of the muropeptide, there is a known outer membrane-localized amidase, AmiD, that releases free tetrapeptide and tripeptide (14, 39).

It is not clear why bacteria use these unusual peptides to cross-link their peptidoglycan. Our MppA structure, the first structure of a protein complex with free Mtp, reveals that Mtp adopts a three-dimensional structure much more similar to that of a conventional α-linked peptide, than we would have expected. Perhaps Mtp has evolved to mimic the structure of an ancestral α-linked cross-linking peptide but one that now confers properties that prevent the action of antimicrobials or peptidases that target the cross-linking peptides.

Acknowledgments—We thank Andrew Leech and Berni Strongitharm from the Technology Facility, Department of Biology, University of York, for help with MS and analysis of ITC data.

REFERENCES

1. Detmers, F. J., Lanfermeijer, F. C., and Poolman, B. (2001) *Res. Microbiol.* **152,** 245–258
2. Doeven, M. K., Kok, J., and Poolman, B. (2005) *Mol. Microbiol.* **57,** 640–649
3. Berntsson, R. P., Smits, S. H., Schmitt, L., Slotboom, D. J., and Poolman, B. (2010) *FEBS Lett.* **584,** 2606–2617
4. Guyer, C. A., Morgan, D. G., and Staros, J. V. (1986) *J. Bacteriol.* **168,** 775–779
5. Smith, M. W., Tyreman, D. R., Payne, G. M., Marshall, N. J., and Payne, J. W. (1999) *Microbiology* **145,** 2891–2901
6. Tame, J. R., Murshudov, G. N., Dodson, E. J., Neil, T. K., Dodson, G. G., Higgins, C. F., and Wilkinson, A. J. (1994) *Science* **264,** 1578–1581
7. Sleigh, S. H., Tame, J. R., Dodson, E. J., and Wilkinson, A. J. (1997) *Biochemistry* **36,** 9747–9758
8. Dunten, P., and Mowbray, S. L. (1995) *Protein Sci.* **4,** 2327–2334
9. Nickitenko, A. V., Trakhanov, S., and Quiocho, F. A. (1995) *Biochemistry* **34,** 16585–16595
10. Sleigh, S. H., Seavers, P. R., Wilkinson, A. J., Ladbury, J. E., and Tame, J. R. (1999) *J. Mol. Biol.* **291,** 393–415
11. Blattner, F. R., Plunkett, G., 3rd, Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B., and Shao, Y. (1997) *Science* **277,** 1453–1462
12. Riley, M., Abe, T., Arnaud, M. B., Berlyn, M. K., Blattner, F. R., Chaudhuri, R. R., Glasner, J. D., Horiuchi, T., Keseler, I. M., Kosuge, T., Mori, H., Perna, N. T., Plunkett, G., 3rd, Rudd, K. E., Serres, M. H., Thomas, G. H., Thomson, N. R., Wishart, D., and Wanner, B. L. (2006) *Nucleic Acids Res.* **34,** 1–9
13. Park, J. T., Raychaudhuri, D., Li, H., Normark, S., and Mengin-Lecreulx, D. (1998) *J. Bacteriol.* **180,** 1215–1223
14. Park, J. T., and Uehara, T. (2008) *Microbiol. Mol. Biol. Rev.* **72,** 211–227
15. Lessard, I. A., Pratt, S. D., McCafferty, D. G., Bussiere, D. E., Hutchins, C., Wanner, B. L., Katz, L., and Walsh, C. T. (1998) *Chem. Biol.* **5,** 489–504
16. Suzuki, H., Koyanagi, T., Izuka, S., Onishi, A., and Kumagai, H. (2005) *J. Bacteriol.* **187,** 5861–5867
17. Leonard, B. A., Podbielski, A., Hedberg, P. J., and Dunny, G. M. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93,** 260–264
18. Fogg, M. J., and Wilkinson, A. J. (2008) *Biochem. Soc. Trans.* **36,** 771–775
19. Zheng, L., Baumann, U., and Reymond, J. L. (2004) *Nucleic Acids Res.* **32,** e115
20. Tame, J. R., Dodson, E. J., Murshudov, G., Higgins, C. F., and Wilkinson, A. J. (1995) *Structure* **3,** 1395–1406
21. Lanfermeijer, F. C., Picon, A., Konings, W. N., and Poolman, B. (1999) *Biochemistry* **38,** 14440–14450
22. Brzozowski, A. M., and Walton, J. (2001) *J. Appl. Crystallogr.* **34,** 97–101
23. Vagin, A., and Teplyakov, A. (2000) *Acta Crystallogr. D Biol. Crystallogr.* **56,** 1622–1624
24. Emsley, P., and Cowtan, K. (2004) *Acta Crystallogr. D Biol. Crystallogr.* **60,** 2126–2132
25. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr. D Biol. Crystallogr.* **53,** 240–255
26. Hervé, M., Boniface, A., Gobec, S., Blanot, D., and Mengin-Lecreulx, D. (2007) *J. Bacteriol.* **189,** 3987–3995
27. Pennartz, A., Généreux, C., Parquet, C., Mengin-Lecreulx, D., and Joris, B. (2009) *Antimicrob. Agents Chemother.* **53,** 2991–2997
28. Holm, L., Kääriäinen, S., Rosenström, P., and Schenkel, A. (2008) *Bioinformatics* **24,** 2780–2781
29. Berntsson, R. P., Doeven, M. K., Fusetti, F., Duurkens, R. H., Sengupta, D., Marrink, S. J., Thunnissen, A. M., Poolman, B., and Slotboom, D. J. (2009) *EMBO J.* **28,** 1332–1340
30. Levdikov, V. M., Blagova, E. V., Brannigan, J. A., Wright, L., Vagin, A. A., and Wilkinson, A. J. (2005) *J. Mol. Biol.* **345,** 879–892
31. Severi, E., Randle, G., Kivlin, P., Whitfield, K., Young, R., Moxon, R., Kelly, D., Hood, D., and Thomas, G. H. (2005) *Mol. Microbiol.* **58,** 1173–1185
32. Thomas, G. H., Southworth, T., León-Kempis, M. R., Leech, A., and Kelly, D. J. (2006) *Microbiology* **152,** 187–198
33. Payne, J. W. (1971) *Biochem. J.* **123,** 245–253
34. Létoffé, S., Delepelaire, P., and Wandersman, C. (2006) *Proc. Natl. Acad. Sci. U.S.A.* **103,** 12891–12896
35. Williams, W. A., Zhang, R. G., Zhou, M., Joachimiak, G., Gornicki, P., Missiakas, D., and Joachimiak, A. (2004) *Biochemistry* **43,** 16193–16202
36. Nanavati, D. M., Thirangoon, K., and Noll, K. M. (2006) *Appl. Environ. Microbiol.* **72,** 1336–1345
37. Cuneo, M. J., Beese, L. S., and Hellinga, H. W. (2009) *J. Biol. Chem.* **284,** 33217–33223
38. Goodell, E. W., and Schwarz, U. (1985) *J. Bacteriol.* **162,** 391–397
39. Uehara, T., and Park, J. T. (2007) *J. Bacteriol.* **189,** 5634–5641