



# Snapshots during the catalytic cycle of a histidine acid phytase reveal an induced-fit structural mechanism

Received for publication, September 4, 2020, and in revised form, October 10, 2020. Published, Papers in Press, October 14, 2020, DOI 10.1074/jbc.RA120.015925

Isabella M. Acquistapace<sup>1</sup> , Monika A. Ziętek<sup>1</sup> , Arthur W. H. Li<sup>1</sup>, Melissa Salmon<sup>1</sup>, Imke Kühn<sup>2</sup>, Mike R. Bedford<sup>3</sup>, Charles A. Brearley<sup>1</sup> , and Andrew M. Hemmings<sup>1,4,\*</sup>

From the <sup>1</sup>School of Biological Sciences and <sup>4</sup>School of Chemistry, University of East Anglia, Norwich, United Kingdom, the <sup>2</sup>AB Vista, Darmstadt, Germany, and the <sup>3</sup>AB Vista, Marlborough, Wiltshire, United Kingdom

Edited by Wolfgang Peti

Highly engineered phytases, which sequentially hydrolyze the hexakisphosphate ester of inositol known as phytic acid, are routinely added to the feeds of monogastric animals to improve phosphate bioavailability. New phytases are sought as starting points to further optimize the rate and extent of dephosphorylation of phytate in the animal digestive tract. Multiple inositol polyphosphate phosphatases (MINPPs) are clade 2 histidine phosphatases (HP2P) able to carry out the stepwise hydrolysis of phytate. MINPPs are not restricted by a strong positional specificity making them attractive targets for development as feed enzymes. Here, we describe the characterization of a MINPP from the Gram-positive bacterium *Bifidobacterium longum* (BIMINPP). BIMINPP has a typical HP2P-fold but, unusually, possesses a large  $\alpha$ -domain polypeptide insertion relative to other MINPPs. This insertion, termed the U-loop, spans the active site and contributes to substrate specificity pockets underpopulated in other HP2Ps. Mutagenesis of U-loop residues reveals its contribution to enzyme kinetics and thermostability. Moreover, four crystal structures of the protein along the catalytic cycle capture, for the first time in an HP2P, a large ligand-driven  $\alpha$ -domain motion essential to allow substrate access to the active site. This motion recruits residues both downstream of a molecular hinge and on the U-loop to participate in specificity subsites, and mutagenesis identified a mobile lysine residue as a key determinant of positional specificity of the enzyme. Taken together, these data provide important new insights to the factors determining stability, substrate recognition, and the structural mechanism of hydrolysis in this industrially important group of enzymes.

Phytic acid (*myo*-inositol hexakisphosphate; InsP<sub>6</sub>) (Fig. S1) is the major storage form of phosphorous (50–80% of total P) in grains, oil seeds, and beans used in common animal feeds (1). Phytases are the phosphomonoesterases that catalyze the sequential dephosphorylation of phytate necessary to release

this phosphorous in its utilizable form, orthophosphate. Animals rely on phytases produced by their commensal microbiota. However, in monogastrics, such as pigs and poultry, the capacity of endogenous phytases to break down dietary phytic acid is limited. Passage of undigested phytic acid to the environment in areas of intensive animal husbandry can lead to the formation of algal blooms in aquatic ecosystems from nutrient overloading. This in turn leads to eutrophication, which has been shown to reduce benthic biomass and biodiversity (2).

To increase the efficiency of conversion of phytic acid into available dietary phosphate, animal feeds are routinely supplemented with exogenous phytases. Of the major classes of phytase, those belonging to clade 2 of the histidine phosphatase superfamily (Pfam code: PF00328, His\_Phos\_2; HP2) have found most widespread use. This is due both to their high specific activity toward InsP<sub>6</sub> and to possessing pH optima in the acid regime. Notable examples are the periplasmic phytase, AppA, from *Escherichia coli*, and the secreted phytases produced by *Aspergilli* such as PhyA of *Aspergillus niger* (3, 4). This area is the subject of ongoing interest in attempts to uncover both new enzymes and engineered variants with high activity, tailored specificity and enhanced stability to environmental assault (5, 6); the latter frequently associated with the high temperatures to which these enzymes are exposed during the feed pelleting process.

HP2 phytases (HP2P) can be grouped according to the specific position of the phosphate ester group of the substrate at which hydrolysis is initiated, e.g. as 1D-3-phytases (EC 3.1.3.8) or 1D-4-phytases (EC 3.1.3.26). The EC 3.1.3.26 signifier refers to seminal characterization of the enantiomerism of products generated by the action of wheat bran phytase (reviewed in Ref. 7). 1D-6-phytases, exemplified by *E. coli* AppA, are acid phosphatases (EC 3.1.3.2) and act via an obligatory phosphohistidine intermediate (Fig. S2) (8). The crystal structures of a variety of HP2Ps have been solved (Table S1) from a variety of both bacterial (9–13) and fungal (14–16) sources. The active site cleft of these lies between two structural domains: an  $\alpha$ - and an  $\alpha/\beta$ -domain. The fold of the latter is well-conserved in HP2P, whereas the  $\alpha$ -domain is subject to enzyme-specific changes (17). At the base of this active site cleft are found two amino acid sequence motifs, the first possesses a consensus RHGXRXh sequence motif (where h represents a hydrophobic amino acid and X is any amino acid) involved in substrate binding and containing the eponymous nucleophilic histidine observed in all HP2 family members (8, 18–20). The catalytic proton donor is found in a second, short, conserved HD sequence motif

This article contains supporting information.

✂ Author's Choice—Final version open access under the terms of the Creative Commons CC-BY license.

\* For correspondence: Andrew M. Hemmings, a.hemmings@uea.ac.uk.

Present address for Isabella M. Acquistapace: Istituto Italiano Di Tecnologia, Genova, Italy.

Present address for Monika A. Ziętek: Abcam plc, Cambridge Biomedical Campus, Cambridge, United Kingdom.

Present address for Arthur W. H. Li: Centre for Genome Damage & Stability, University of Sussex, Brighton, United Kingdom.

Present address for Melissa Salmon: Earlham Institute, Norwich Research Park, Norwich, United Kingdom.

## Results

### The U-loop: a polypeptide insertion in extracellular MINPPs

Phylogenetic analysis of HP2Ps reveals the expected separation of MINPP and nonMINPP sequences (Fig. 1A). The non-MINPPs can be identified as those from fungal (eHP2P) or bacterial (bHP2P) sources. The MINPPs can be similarly divided into those from *Eukaryota* or *Bacteria*. The bacterial MINPPs are further divided in two clades, named here clades 1 and clade 2. Three groups of polypeptide inserts in MINPP sequences were detected (we refer to these as U-loops) and given the identifiers A, B, or C based on insert length. The sequences of the three loop types do not align well with each other, save for a conserved region toward the C-terminal end (Figs. 1B, Fig. S3). This region is followed by a characteristic tetrapeptide DAAM motif, which is absent in sequences that do not contain a U-loop. The type A U-loop is the longest, containing two conserved cysteine residues. It is found in clade 2 MINPPs as extracellular membrane-anchored enzymes of Gram-positive bacteria. Of these, *Bifidobacterium longum* is a major human gut bacterium present from infancy through to adult life (34). Type B is a medium length loop also found in clade 2 MINPPs. It can be viewed as resulting from a deletion of the amino acids between the two conserved cysteine residues of the type A insert (Fig. 1B). Enzymes containing this loop are mostly predicted to be membrane-anchored lipoproteins, characterized by a SEC/SPII signal peptide. Type C is a short U-loop found in both clade 1 and 2 MINPPs. The majority of enzymes with type C inserts appear to be either lipoproteins (possessing a SEC/SPII signal peptide) or have a signal peptide (SEC/SPI N-terminal signal peptide).

The extracellular type A U-loop-containing MINPP from *B. longum* subsp. *infantis* ATCC 15697 (*B*MINPP) is reported to be relatively thermostable, preserving 44% of activity after incubation at 80 °C for 15 min (35). It has a sequence identity of only 23% compared with the mesophilic, nonU-loop-containing MINPP secreted in lipid vesicles from the previously characterized Gram-negative human gut bacterium *B. thetaiotaomicron* (*B*tMINPP) (22). It also resides in a different clade. It is larger than the *Bacteroides* enzyme, bearing nearly 100 additional amino acids in multiple polypeptide insertions and, based on the known crystal structure of the latter, at least one of these was presumed to face the active site cleft. For these reasons we decided to probe the structure-function basis for the elevated thermostability of this type A U-loop enzyme.

### *B*MINPP displays typical MINPP catalytic positional specificity

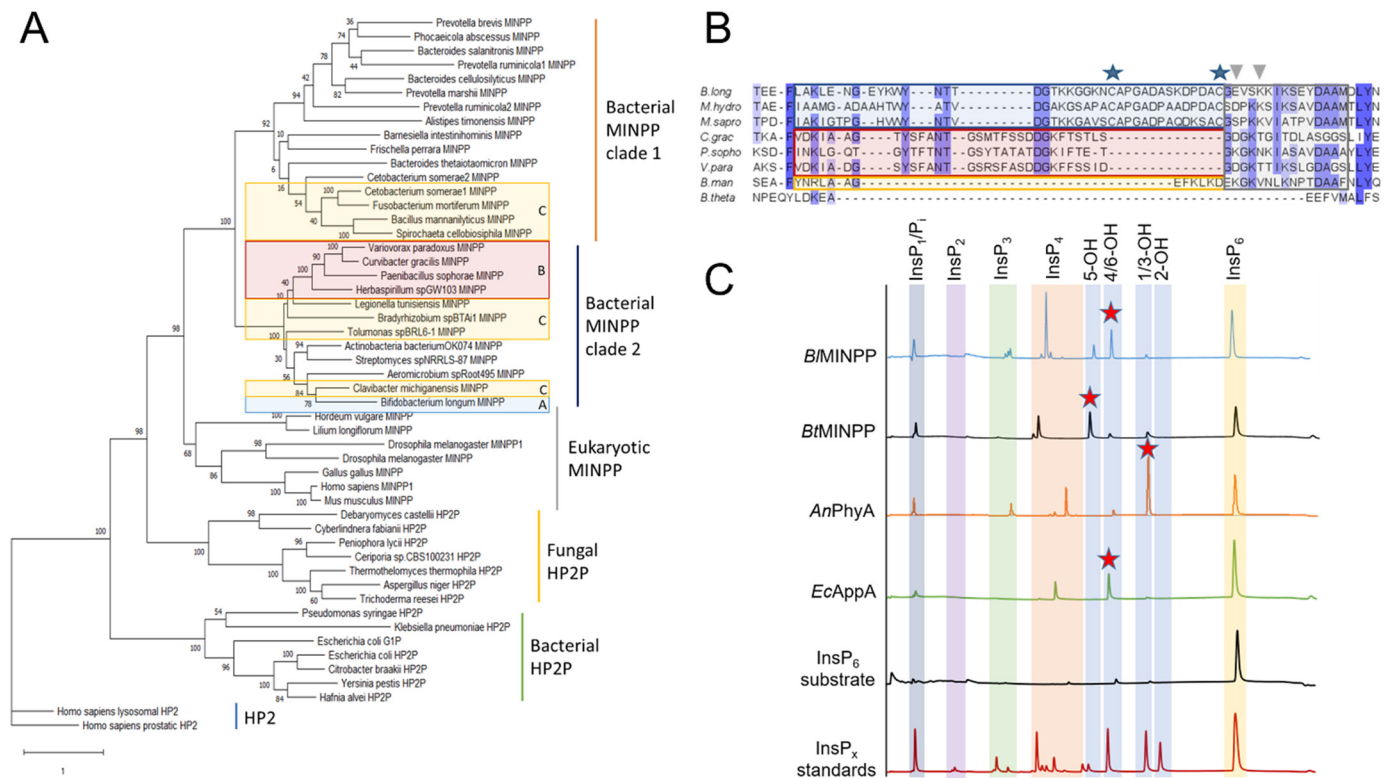
The N-terminal signal peptide and a C-terminal sortase-dependent cell wall-anchoring L(P/A)XTG domain were removed from the cloned *B*MINPP construct, and the enzyme expressed intracellularly in *E. coli* and purified. The recombinant enzyme has a pH profile of activity toward phytate, which is typical of HP2Ps, displaying a maximum activity at pH 5.5 (35). HPLC separation of hydrolysis products following digestion of  $\text{InsP}_6$  reveals the lower positional specificity toward this substrate that is characteristic of MINPPs (Fig. 1C) relative to other HP2Ps. In this way, the major  $\text{InsP}_5$  observed is 1D- and/or 1L- $\text{Ins}(1,2,3,5,6)\text{P}_5$ , hereafter [4/6-OH] $\text{InsP}_5$  (note that the

positioned such that the aspartic acid residue is the consensus proton donor in the catalytic mechanism (9, 20–23). The proton donor is required for the release of the lower inositol phosphorylated product of hydrolysis and orthophosphate with regeneration of the catalytic histidine.

The multiple inositol polyphosphate phosphatases (MINPPs) constitute a distinct evolutionary group within clade 2 of the histidine phosphatase superfamily (21, 24). Examples are found in *Bacteria* and *Eukarya* but MINPPs have not yet been identified in the domain *Archaea* (25). Like other HP2P enzymes, MINPPs carry the RHGXR<sub>x</sub>H sequence motif involved in substrate binding and catalysis. However, instead of an HD proton donor motif, the residues of an amino acid triplet, frequently with the sequence HAE, are presumed to provide an equivalent function (22). MINPPs were so-named because of their broad substrate specificity when compared with other HP2P. This allows, for example, the removal of the 3-phosphate of 2,3-bisphosphoglycerate, by so doing expanding the regulatory capacity of the Rapoport–Luebering glycolytic shunt in *Dictyostelium*, birds, and mammals (26). In the same vein, unlike HD-motif containing HP2P enzymes, MINPPs lack a strong initial positional hydrolytic specificity toward phytic acid, producing a variety of  $\text{InsP}_5$ s and lower  $\text{InsP}_x$  (22, 27, 28). As with the eukaryotic enzymes, bacterial MINPPs also lack a strong initial positional specificity, generating in the process a variety of  $\text{InsP}_5$ s and lower inositol polyphosphates. These partially dephosphorylated intermediates may have a variety of functions. For example, an extracellular MINPP released in outer membrane vesicles by the major Gram-negative human gut symbiont, *Bacteroides thetaiotaomicron* (22) has been suggested to participate in cross-kingdom cell-to-cell signaling by promoting intracellular  $\text{Ca}^{2+}$  signaling in intestinal epithelial cells. Despite this, the function of many MINPPs is still uncertain despite evidence of a role of these enzymes in a variety of cellular processes and organisms (29–32).

Extracellular enzymes generally show high thermostability (33) and so, based on the premise that bacterial extracellular MINPPs may represent a useful source of enzymes as next-generation animal feed phytases, we carried out an analysis of their amino acid sequences. In the process, we identified a subset bearing large  $\alpha$ -domain polypeptide insertions (termed U-loops). To gain a wider understanding of the implications of these insertions on the properties of these enzymes, we carried out a crystal structure and mutagenic analysis of a representative member, *B*MINPP, a moderately thermophilic, membrane-anchored extracellular enzyme. Interestingly, whereas the presence of an intra-U-loop disulfide bridge increased overall protein thermostability by 10 °C, it was found that residues of the loop unexpectedly also contributed to substrate specificity. Furthermore, four crystal structures were determined providing snapshots along the catalytic cycle of the enzyme revealing a large, ligand-driven domain motion previously unseen in the HP2P family. These results suggest evolution of polypeptide insertions may present a route for enhanced thermostability of extracellular phytases, but their presence imposes additional requirements for enhanced molecular flexibility necessary for catalysis to occur.

# Phytase Induced-Fit Structural Mechanism



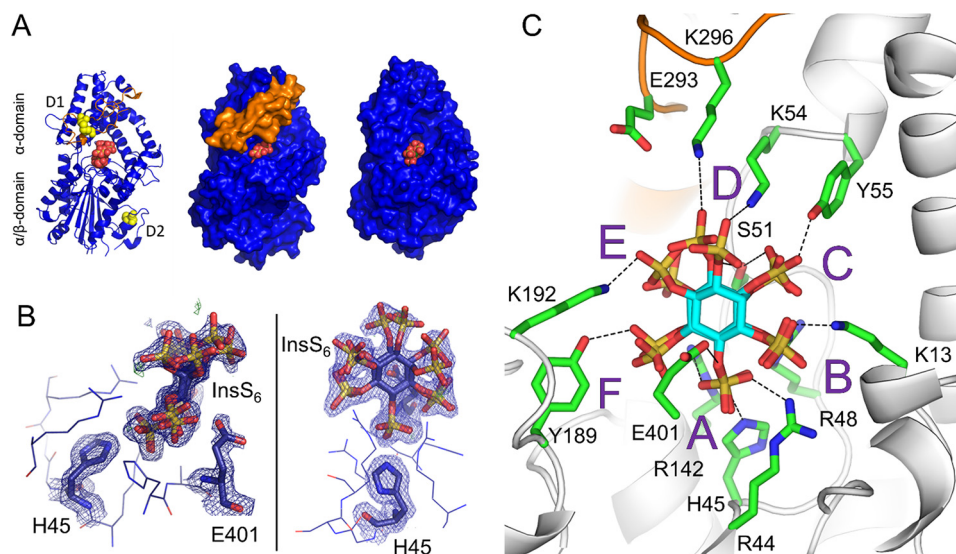
**Figure 1.** A, maximum likelihood phylogenetic tree resulting from the alignment of 51 clade 2 histidine phosphatases (Pfam family: his\_phos\_2). The sequences correspond mainly to phytases with the exception of two histidine phosphatases without phytase activity (HP2, light blue) on which the tree is rooted. Different branches were identified as representatives of family subgroups: bacterial phytases (green), fungal phytases (yellow), eukaryotic MINPPs (gray), and two clades of bacterial MINPPs: clade 1 (blue) and clade 2 (orange). Sequences were inspected for the presence of a U-loop insertion and highlighted as follows: blue boxes, sequences contain a type A U-loop; red boxes, contain a type B U-loop; yellow boxes, contain a type C U-loop. B, sequence alignment of representative U-loop-containing MINPPs. All three U-loop types contain a conserved region, highlighted in a gray shaded box, which includes the residues potentially involved in contacts with substrate. The remaining regions are less conserved and can be either long (type A U-loop, blue box, characterized by an insertion between two cysteines potentially forming a disulfide bridge), medium (type B U-loop, red box), or short (type C U-loop, yellow box). Most sequences containing a U-loop contain a “DAAM” motif downstream of the insertion. U-loops are found in bacterial MINPPs from either clade A or B. Stars highlight positions of cysteine residues conserved in type A U-loops, whereas inverted triangles point to residues contributing to phytate specificity pockets. C, HPLC chromatograms of the products of  $\text{InsP}_6$  hydrolysis by *BtMINPP*, *BtMINPP*, *A. niger* PhyA (*AnPhyA*), and *E. coli* AppA (*EcAppA*). x axis, retention time (min); y axis, intensity of UV absorbance at 290 nm. Chromatograms are labeled by enzyme. The major  $\text{InsP}_5$  product for each enzyme is indicated by a red star. Chromatograms of the undigested substrate ( $\text{InsP}_6$  substrate) and an acid hydrolysate of the substrate ( $\text{InsP}_x$  standards) are shown for reference. The elution volume ranges for the various inositol polyphosphate products are highlighted by vertical colored backgrounds (note that the notation for the  $\text{InsP}_5$  products is based on the identity of the free hydroxyl group of the intermediate).

enantiomers are not resolvable). The *meso*-compound  $\text{Ins}(1,2,3,4,6)\text{P}_5$ , hereafter [5-OH] $\text{InsP}_5$ , is also produced but at a lower level (the ratio of these products is roughly 2:1). A much smaller amount of 1D- and/or 1L- $\text{Ins}(1,2,4,5,6)\text{P}_5$ , hereafter [1/3-OH] $\text{InsP}_5$ , is also observed. This differs subtly from *BtMINPP* which produces [5-OH] $\text{InsP}_5$  as its major product (22). Hydrolytic activity also appears to stop with production of  $\text{InsP}_3$ s (35), differing again from the *Bacteroides* enzyme that produces lower inositol mono/polyphosphates. (Although  $\text{InsP}_2$ s are not observed in chromatograms separating the products of action of *BtMINPP*, it is important to note that  $\text{InsP}_1$  and  $\text{P}_i$  are not resolvable using our HPLC method and so hydrolysis beyond  $\text{InsP}_3$  intermediates may be occurring.) This behavior contrasts with that of other bHP2Ps such as *E. coli* AppA, which is accepted to possess high positional specificity, generating a majority 1D- $\text{Ins}(1,2,3,4,5)\text{P}_5$  product, hereafter [6-OH] $\text{InsP}_5$  with a small amount of 1D- $\text{Ins}(1,2,4,5)\text{P}_5$ , [1/3-OH] $\text{InsP}_5$  (Fig. 1C) and a well-characterized 6/1/3/4/5 dephosphorylation pathway, ultimately yielding  $\text{InsP}_2$  (36). In an attempt to cast light on the structural basis for the differing positional

specificities of *BtMINPP* and *BtMINPP* we embarked on an X-ray crystal structure determination of the *B. longum* enzyme.

## Phytate specificity subsites from the crystal structure of the *BtMINPP-InsP6* complex

A crystal of the purified recombinant enzyme grown in the presence of an excess  $\text{InsP}_6$ , a nonhydrolyzable phytate mimic, diffracted to 1.84 Å resolution. These data were used to solve the X-ray crystal structure of the complex by molecular replacement with the  $\text{InsP}_6$ -bound structure of *BtMINPP* employed as search model. Despite low sequence identity, the overall folds of the two MINPPs are very similar. In keeping with previously reported crystal structures of HP2P family members, the structure of *BtMINPP* consists of two domains, an  $\alpha/\beta$ -domain and an  $\alpha$ -domain (Fig. 2A). For both *BtMINPP* and *BtMINPP* the  $\alpha$ -domain resembles that seen in the crystal structures of eHP2Ps from the closely related fungi *Aspergillus niger* (14) and *Aspergillus fumigatus* (37). The catalytic center of *BtMINPP* consists of the nucleophilic histidine (His-45), three arginine residues (Arg-44, Arg-48, and Arg-142), which coordinate the scissile



**Figure 2. The crystal structure of the complex of *BtMINPP* with the substrate analog inhibitor, *InsS<sub>6</sub>*.** *A*, the U-loop obscures access of inhibitor to active center. *Left panel*: a cartoon representation of the fold of *BtMINPP* shown in *blue* except for the residues of the U-loop (257–300), which are colored *orange*. Atoms of the inhibitor bound between the lower  $\alpha/\beta$ -domain and the upper  $\alpha$ -domain are shown in *sphere format* and colored *orange* (sulfur) and *red* (oxygen). The four cysteine residues forming two disulfide bridges (labeled *D1* and *D2*) are shown as *spheres* and colored *yellow*. *Center panel*, a molecular surface representation of the *BtMINPP* structure. Orientation and coloring are as described in the *left panel*. *Right panel*, molecular surface representation of the structure of *BtMINPP* in complex with *InsS<sub>6</sub>* (PDB 4FDU). Orientation and coloring are as described in the *left panel*. *B*, orthogonal views of the electron density for selected active site residues and the bound inhibitor molecule. Inhibitor, two conformations. Residues His-45 and Glu-401 are shown in *stick format*. Other residues of the two active site signature sequence motifs RHGX<sub>R</sub> (residues 44–48) and HAE (residues 399–401), along with Arg-142 are shown in *wire format*. *Dark blue hatching* shows the difference Fourier ( $2mF_o - DF_c$ ) electron density map contoured at  $1.0\sigma$ . *C*, closeup view of bound inhibitor and definition of specificity subsite nomenclature. Enzyme shown in *cartoon format* and colored *gray*, except for the residues of the U-loop (257–300), which are colored *orange*. Bound *InsS<sub>6</sub>* (two conformations) are shown with carbon colored *cyan*, active site residues with carbon colored *green* and labeled. Polar interactions are indicated by *dashed lines*. Specificity subsites are labeled *A–F* such that the analog of the scissile phosphate occupies subsite *A* and the remaining subsites are arrayed in a counterclockwise sense when observed from the viewpoint adopted in this figure.

phosphate during catalysis, and the amino acid triplet HAE (residues 339–401) where the glutamic acid residue acts as the presumed proton donor during catalysis (22). Glutamic acid here replaces the aspartic acid seen in a corresponding HD sequence motif commonly found in nonMINPP HP2P family members (20). The inhibitor binds in the active site cleft in a manner presumed to mimic that of phytic acid. However, inspection of a molecular surface representation of the complex suggests that egress of the ligand is obstructed (Fig. 2A). This obstruction would presumably also hinder diffusion of partially dephosphorylated inositol polyphosphate hydrolysis products. This is in contrast to the situation observed with *BtMINPP* (Fig. 2A) and with all bHP2P for which corresponding crystal structures are available.

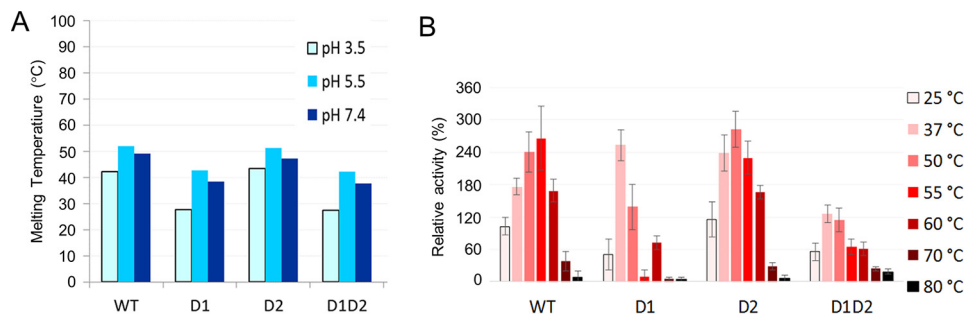
The structure of the *BtMINPP-InsS<sub>6</sub>* complex reveals static disorder in the inhibitor, which is bound in two orientations (Fig. 2B, Fig. S4). The first of these has the sulfate group at position 1D-4 of the inositol ring positioned proximal to the nucleophilic histidine. The other has the 1D-6 sulfate in this orientation. The pairs of inhibitor molecules bound to the two protein monomers in the crystallographic asymmetric unit have occupancy ratios 0.63:0.37 and 0.53:0.47 (S4:S6) suggesting little intrinsic preference between the two orientations of binding. Although noting that [4-OH]1D-*InsP<sub>5</sub>* and [6-OH]1D-*InsP<sub>5</sub>* = [4-OH]1L-*InsP<sub>5</sub>*, enantiomers cannot be separated by our HPLC approach, our observation of disorder in binding of a substrate-mimic may reflect a dual 1D-4/6-positional specificity for the enzyme. Certainly, the presence of conformational disorder (*i.e.* multiple bound conformations) in the *InsS<sub>6</sub>* ligand

is consistent with the lower positional stereospecificity observed for this enzyme relative to that seen for the more specific canonical HP2P (36, 38–40).

The structure also provides a model for enzyme-substrate interactions. Specificity pockets for the binding of the six phosphates of *InsP<sub>6</sub>* can be inferred by identifying all amino acids within 6 Å of each sulfate of the ligand in the complexed structure (the result is essentially identical irrespective of the orientation of binding). In this scheme, the residues of pocket *A* bind the scissile phosphate group and represent the catalytic center. This pocket is symmetrical in structure and charge, presenting guanidino groups from two arginine residues (Arg-48 and Arg-142) on each side of the phosphate. From a vantage point positioned behind the inositol ring and looking through it toward the *A*-subsite and arginines 48 and 142, the remaining specificity subsites are then labeled *B–F* in a counterclockwise fashion, following the order of decreasing sulfate number attached to the *myo*-inositol ring (Fig. 2C). It is likely that these are representative of the subsites involved in recognition and binding of *InsP<sub>6</sub>*. Of note is the fact that specificity pockets *C*, *D*, and *E* are more highly populated by active site residues than the equivalent sites in *BtMINPP* and other HP2P such as *EcAppA* (Table S2). U-loop residues account for this difference in the case of specificity subsite *D*. Specificity subsites *C* and *E* are also more highly populated in *BtMINPP*. However, this appears to be due more to the presence of bulkier subsite residues in *BtMINPP* than to local topology changes.

*BtMINPP* is significantly larger than *BtMINPP* and the majority of the additional residues are present as random coil. The

## Phytase Induced-Fit Structural Mechanism



**Figure 3. The role of the U-loop in determining enzyme stability.** *A*, melting temperatures of *B*/MINPP disulfide bridge knock-out mutants. Key to substitutions: *D1*, C278A/C291A; *D2*, C483A/C501; *D1D2*, C278A/C291A/C483A/C501. Results of differential scanning calorimetry at pH 3.5, 5.5, and 7.4 of *B*/MINPP WT (*WT*) and mutants *D1*, *D2*, and *D1D2* are shown. *y* axis, melting temperature (°C); *x* axis, protein samples. Reactions were carried after incubation for 30 min at 25, 37, 50, 55, 60, 70, and 80 °C. Mutants *D1* and *D1D2* shows a 10 °C lower melting temperature when compared with the WT enzyme. Due to DSC sample requirements only a single measurement was made at each pH. As such, these results should be considered tentative. *B*, recovery after heating assays of *B*/MINPP disulfide bridge knock-out mutants. The enzyme was incubated at pH 5.5 and residual phytase activity determined after cooling to room temperature. Activity expressed as a percentage (%) relative to that recovered by of the WT enzyme after heating to 25 °C.

exception to this are the 44 amino acids of the U-loop (residues 257 to 300) that lie on top of the active site, wrapped around the  $\alpha$ -domain. This insertion is stabilized in position by interaction of U-loop residues residues Asn-266 and Asp-289 with the main chain amide and carbonyl groups, respectively, of Tyr-53. Tyr-53 lies only a few residues C-terminal to the consensus RHGXRXh active site sequence motif. Only three further polar interactions are found, these involving the side chains of Trp-264, Asn-277, and Asp-287, with the remainder of the interface predominantly hydrophobic in nature. No direct interactions are seen with active site residues. U-loop residues that close within 5 Å of the substrate analog are Glu-293 and Lys-296 (Fig. 2C) and form part of specificity subsite D. The U-loop thus helps define the active site, contributing to interactions in specificity subsite D and also partially obscuring product egress from it (Fig. 2A). These interactions are missing in the shorter, nonU-loop *Bacteroides* enzyme structure and simple sequence analysis suggests they are also absent in other important proteins in the family, such as the nonU-loop containing eMINPP proteins.

Thus, through its interactions with the bound substrate analog, the U-loop contributes to substrate recognition and, when compared with nonU-Loop enzymes, changes the shape and charge distribution of the active site cavity. The structure of *B*/MINPP in complex with InsS<sub>6</sub> also allows us to speculate on the likely roles of enzymes from the shorter U-loop classes. Inserts of types B and C are of sufficient length to provide equivalents to those residues from the leading strand of the hairpin that are able to interact with the substrate. Thus, the residues in *B*/MINPP, which interact with the substrate analog in the D-specificity pocket, are predicted to be present in proteins from all U-loop classes and we would predict these loops to play analogous roles in determining substrate recognition and binding.

### The influence of the U-loop on protein stability and catalytic efficiency

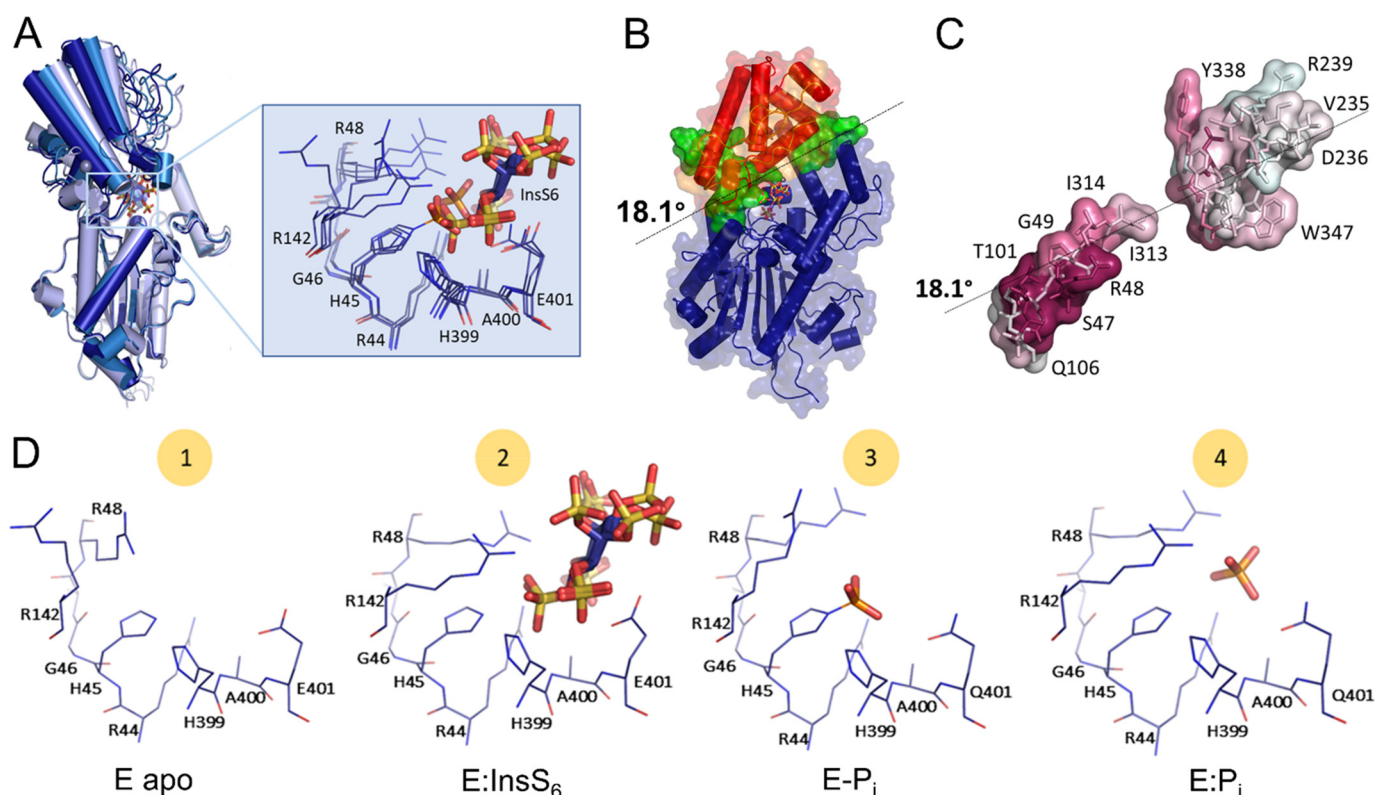
Two cysteines on the U-loop, Cys-278 and Cys-291, form a disulfide bridge. A further disulfide (Cys-483–Cys-501) is observed in the  $\alpha/\beta$ -domain, roughly 27 Å distant from the nucleophilic histidine and the active site (Fig. 2A). To test the

roles of these disulfides on structure stabilization, the corresponding cysteine residue pairs were individually replaced by alanine mutagenesis and the resulting mutants assessed for stability by measurement of  $T_m$  by differential scanning calorimetry, and by recovery of phytase activity after heating. Three disulfide deletion mutants were produced removing the Cys-278–Cys-291 disulfide (*D1* mutant), the Cys-483–Cys-501 disulfide (*D2* mutant), or both (*D1D2* mutant). Differential scanning calorimetry revealed the melting temperatures of mutants *D1* and *D1D2* to be both reduced by ~10 °C relative to the WT enzyme whereas that of mutant *D2* was unchanged (Fig. 3A). MINPPs typically exhibit phytase activity maxima around pH 3.5, 5.5, and 7.5 (22). The depression in melting temperature observed for mutants *D1* and *D1D2* was effectively consistent across these pH values. These results suggest an involvement of the disulfide bridge Cys-278–Cys-291 and of the U-loop more generally in *B*/MINPP structure stabilization. The recovery after heating of phytase activity of mutants *D1* and *D1D2* was also impaired (Fig. 3B). This enhanced sensitivity to heating following deletion of the U-loop disulfide may result from a general destabilization of the structure or simply from local melting of the U-loop structure leading to loss of stabilizing contacts of U-loop residues with substrate.

Only U-loop residues Glu-293 and Lys-296 approach within 5 Å of the bound substrate analog, the side chain of the former lying 4.6 Å from the sulfate in pocket D (Fig. 2C). We probed the roles of these residues by alanine mutagenesis (Fig. S5) (Table S3).  $K_m$  was unperturbed for both mutants.  $k_{cat}$  was reduced for K296A, whereas for E293A it was almost 70% higher than the WT enzyme, suggesting that the U-loop contributes to fine tuning of catalytic activity, possibly by involvement with formation of the ES complex and/or product release.

### Ligand-driven domain movement along the catalytic cycle

The three-dimensional structures of many HP2P are known. These have been described either in the apo- or product (*i.e.* P<sub>i</sub>)-bound form, or in complex with InsS<sub>6</sub> (for examples see Table S1). Enzymes for which pairs of apo- and inhibitor complex structures are available do not display discernible domain movements upon ligand binding, including an array of HP2 phytases (11, 12, 37, 41). This is also true for *Bt*MINPP (22).



**Figure 4. Ligand-driven domain movement along the structural catalytic cycle of *B1MINPP*.** *A*, superposition of the four *B1MINPP* structures solved in this study shown in cartoon format: *blue*, apo-enzyme and phosphohistidine intermediate; *sky-blue*, phosphate-bound form; *light blue*, *InsS<sub>6</sub>*-bound form.  $\alpha/\beta$ - and  $\alpha$ -domains are indicated. A closeup of the catalytic core residues and inhibitor is also shown. Amino acids in conserved sequence motifs are labeled. *InsS<sub>6</sub>*, *P<sub>i</sub>*, and the phosphate group of phosphohistidine are represented in stick format. Large positional shifts are detected for Arg-48, Arg-142, and Glu-401 upon ligand binding. *B*, *B1MINPP* regions involved in the domain movement (the lid). Areas highlighted: *blue*, fixed domain; *red* and *orange*, moving domain; *green*, hinge residues. The U-loop is *orange*-colored. The hinge rotational axis ( $18.1^\circ$  from the apo-form in the structure above) is represented by *dotted lines*. *C*, hinge residues colored according to conservation as determined using ConSurf (45) according to a multisequence alignment of *MINPP* representative sequences. Darker surface colors indicate higher residue conservation. Residues most conserved are the pairs Ser-47, Arg-48, and Thr-101, Leu-100. The latter residue pair is part of a conserved "GXLTX<sub>2</sub>G" sequence motif. *D*, four snapshots along the catalytic cycle of *B1MINPP* active site: 1) *E apo*: apo-protein; 2) *E:InsS<sub>6</sub>*: a model for the substrate-bound complex (*B1MINPP* in complex with *InsS<sub>6</sub>*); 3) *E-P<sub>i</sub>*: a catalytic intermediate (*B1MINPP* E401Q-phosphohistidine intermediate); and 4) *E:Pi*: the product complex (*B1MINPP* in complex with *P<sub>i</sub>*). Residues of the two active site signature sequence motifs RHGX<sub>R</sub> (residues 44–48) and HAE (residues 399–401) along with Arg-142 are shown in wire format.

Nevertheless, more localized structural changes have been detected. For example, in *E. coli* AppA, binding of *InsS<sub>6</sub>* induces a local rearrangement of residues immediately downstream of the RHGX<sub>R</sub>Xh catalytic sequence motif (9). Amino acids displaying the largest conformational change were Thr-23 and Lys-24, which "close" the active site cleft over the substrate. This movement allows the side chain of Lys-24 to rotate so as to contact the substrate analog.

To test whether the presence of the steric obstruction presented by the U-loop in *B1MINPP* might necessitate more profound conformational changes to allow substrate access to and/or product egress from the active site, structures of further representative states along the catalytic pathway were sought. To this end, crystals of the apoenzyme and of an inactive mutant in which the presumed catalytic proton donor Glu-401 was replaced by a glutamine (E401Q) were prepared, the latter in an attempt to trap the catalytic phosphohistidine intermediate. The structures of the apoenzyme and product complex were solved at 1.65 and 1.71 Å resolution, respectively, whereas a data set collected at 2.40 Å resolution from a crystal of the inactive E401Q mutant yielded a structure of the His-45 phosphohistidine derivative.

Analysis of the resulting structures indicates that, unlike previously characterized HP2P, *B1MINPP* possesses an unusual inherent flexibility (Fig. 4A). In the apo-state the enzyme exists in an open conformation with many of the specificity pockets incompletely formed. On binding the substrate analog, the enzyme moves to a closed conformation where the full array of interactions with ligand are present. This movement is not limited to the active center, as is seen in *E. coli* AppA, but rather it propagates to a large region of the  $\alpha$ -domain, which we will refer to as the molecular "lid," which rotates on binding of substrate toward the  $\alpha/\beta$ -domain. The lid comprises the majority of the  $\alpha$ -domain residues, excluding only the residues of helices Ala-222–Ile-231 and Ala-337–Lys-357, which line one side of the active site cleft and the latter of which contributes residues to specificity pocket B. DynDom (42) identified a maximum rotational movement of the lid of  $18.1^\circ$  upon ligand binding, corresponding to 82.3% closure of the moving domain (Fig. 4B).

The  $\alpha/\beta$ -domain and the remainder of the  $\alpha$ -domain undergo only limited changes and are considered fixed (RMS deviation 0.55 Å). Lid movements are also seen in the transition from the substrate analog-bound to phosphohistidine intermediate and

## Phytase Induced-Fit Structural Mechanism

then to the product-bound form. Unexpectedly, the phosphohistidine intermediate state resembles the open conformation, whereas the product-bound form of the enzyme exhibits a lid rotation of 10° (degree of closure 68.6%) back to a half-closed conformation. Taken together, these snapshots suggest a mechanism whereby the enzyme undergoes a complex structural catalytic cycle during which the lid closes on binding of substrate, then opens fully to expel the first stage product (a lower phosphorylated inositol). Presumably, a second lid motion to a half-closed state then follows to allow hydrolysis of the phosphohistidine intermediate and generation of the bound second stage product (orthophosphate). Finally, lid rotation to regain the open state allows diffusional loss of P<sub>i</sub> and reattainment of the resting state of the enzyme.

As interdomain screw axes are located in the proximity of bending residues, these amino acids can be considered to be a mechanical hinge with the interdomain screw axis as hinge axis. Two such mechanical hinges were identified by DynDom: the first involving residues clustered on two adjacent loops connecting the  $\alpha/\beta$ - and  $\alpha$ -domains, and the second involving two short sequences at the termini of  $\alpha$ -helices on the other side of the  $\alpha$ -domain, away from the bound substrate analog (Fig. 4B). Multiple sequence alignments of representative MINPPs and analysis using ConSurf (43) showed high conservation of the residues of the first of these hinges. The two strands of this hinge contain, respectively, the catalytic signature motif RHGXRXh (beginning at residue 44) and a GXLTX<sub>2</sub>G motif (beginning at residue 98) also conserved in HP2 enzymes (Figs. 4C, Fig. S6).

### Scissile phosphate interaction with Arg-48 is a candidate for initiation of lid closure

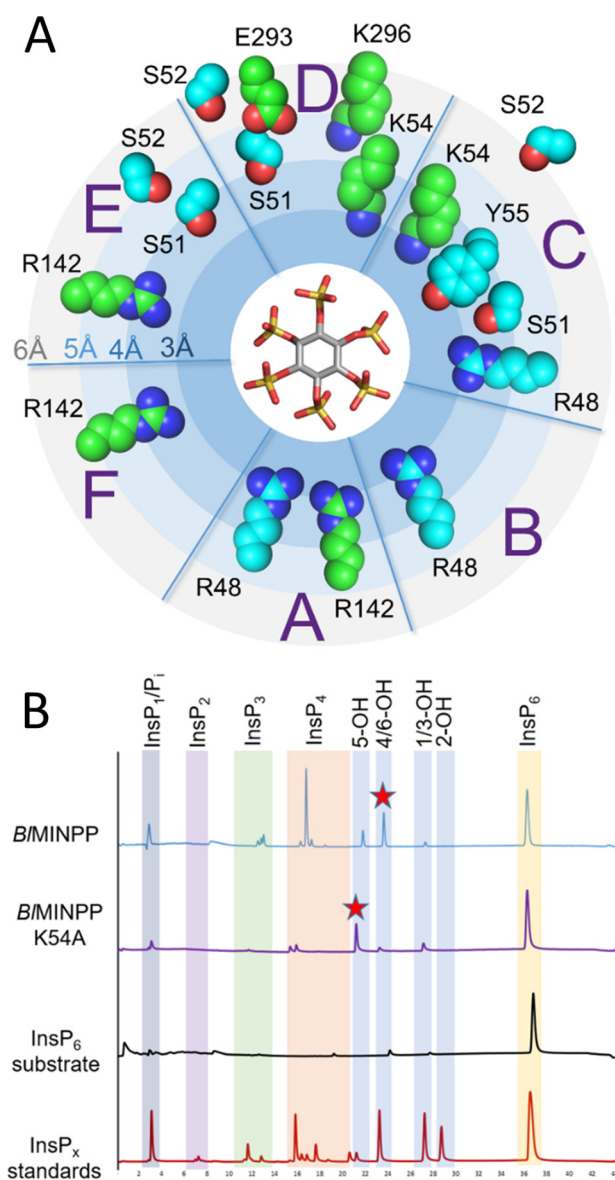
Considering its role in substrate positioning, Arg-48 is presumably one of the residues that initiates lid movement, most likely following engagement of a phosphate into the A-specificity pocket of the enzyme (Fig. 4D). Docking of a phosphate into pocket A leads to rotation about Arg-48 and propagation of a physical shift to the following polypeptide such that Ser-51 engages with phosphates in specificity pockets C and E, Lys-54 in pockets C and D, and Tyr-55 in pocket C. This region of polypeptide is hydrogen bonded to the second strand in the hinge, whose motion is transmitted to Arg-142, which swings in to engage with phosphates in pockets A and F. Through its interactions with Tyr-53, this motion can be transmitted to the lid subdomain, rotation of which leads to large displacements of residues of the U-loop such that residues Glu-293 and Lys-296 enter specificity pocket D, thus completing coordination of the substrate. This proposal makes ligand-binding the driving force for  $\alpha$ -domain movement. In *E. coli* AppA the arginine (Arg-20) corresponding to Arg-48 in *Bt*MINPP has essentially the same role except that movement is propagated only to the residues immediately following it in the sequence. In *Bt*MINPP, however, this small conformational change can be amplified through connections with the U-loop to the lid subdomain with important consequences for interactions with the substrate in specificity pockets C, D, and E.

In the phosphohistidine intermediate, Arg-48 and Arg-142 adopt open-state conformations, neither making contact with the His-45 phospho-group. In fact, this group does not make contact with hinge residues and makes direct contact only with the main chain amide of A400 (A400 forms part of the HAE proton donor motif). It also forms a weaker interaction with Arg-44. The loss of the majority of the interactions observed with the scissile phosphate in the substrate complex are consistent with the relaxation of the phosphohistidine intermediate enzyme to the fully open state. In the phosphate-bound structure, the final stage of the structural catalytic cycle, Arg-48 and Arg-142 swing back to contact the product ion. This ion also makes contact with Arg-44 and Glu-401. Save for a small drift of the ion toward Arg-48, its placement and the interactions it makes with coordinating residues are very similar to that seen for the A-pocket sulfate in the complex with InsS<sub>6</sub>. These interactions require rotation of the lid but can be formed without the necessity for the full extent of lid rotation observed in the closed state. As a consequence, the enzyme exhibits a half-closed state. Presumably, this state will more closely resemble that required for hydrolysis of the phosphohistidine intermediate than the fully open state as alluded to above.

In contrast to the situation for *Bt*MINPP, the structures of the complexes of *Bt*MINPP with InsS<sub>6</sub> (PDB 4FDU) and with P<sub>i</sub> (PDB 4FDT) are very similar (RMS deviation 0.24 Å for 333 C $\alpha$  atoms) (22). Interestingly, these structures most closely resemble the P<sub>i</sub>-bound form of *Bt*MINPP, *i.e.* they adopt a half-closed conformation. No residues in the equivalent of the lid subdomain contribute to specificity pockets in *Bt*MINPP and as a consequence further closure of the  $\alpha$ -domain is presumably unnecessary.

### Roles of mobile residues in determining *Bt*MINPP positional specificity

The presence of the U-loop in *Bt*MINPP leads to a range of motions during the catalytic cycle both subtle (for example, rotation of the RHGXRXh motif to complete coordination of the scissile phosphate in the A-subsite) and marked (most noticeably the 10 Å swing in the U-loop itself leading to residues contacting the bound substrate in specificity pocket D) (Fig. 5A). We therefore decided to investigate the roles of specific residues in these regions in determining *Bt*MINPP positional stereospecificity. U-loop residues that on lid rotation approach to within 5 Å of the substrate analog were Glu-293 and Lys-296. No discernible changes in primary hydrolytic positional specificity were observed when these residues were mutated to alanines (Fig. S7). Ser-51 and Lys-54 lie downstream of hinge residues at the domain interface and, on substrate binding move to interact with phosphates in subsites C, D, and E (Fig. 5A). An analysis of representative MINPP sequences showed that a serine or a threonine is present at the equivalent of position 51 in at least 80% of cases, with a consensus sequence RHGXRXL(S/T)SXK (residues 51 and 54 in bold) (Fig. S6). In the crystal structure of *Bt*MINPP:InsS<sub>6</sub>, the side chain of Ser-51 interacts with the substrate analog in sites C, D, and E, whereas Lys-54 contacts in sites C and D. Although the mutation S51A leads to an unchanged positional specificity, the



**Figure 5. Conformational change and its effect on the composition of specificity pockets and positional specificity of B/MINPP.** A, a representation of the specificity subsites of B/MINPP labeled A–F along with those active site residues that undergo significant movement (residues which move between 2 and 4 Å have carbon atoms colored cyan and those moving more than 4 Å on ligand binding are colored green). Residues that move less than 2 Å are 2 Å from their position in the apo-B/MINPP structure to that found in the complex with InsP<sub>6</sub>. Residues shown in sphere format and colored so are not shown. The placement of the residues indicates the minimum polar contact distance with InsP<sub>6</sub> (see Å scale on the left-hand side). In the center of the image sits a stick representation of InsP<sub>6</sub> as a proxy for phytate positioned so that the sulfate attached to C6 is located in specificity pocket A. This orientation places the axial 2-sulfate in specificity pocket E. B, HPLC chromatograms are compared for reactions that represent a depletion of InsP<sub>6</sub> by around 60% of the initial concentration. The major InsP<sub>5</sub> product in each case is indicated by a red star. Mutant K54A shows a predominant 5-hydroxy InsP<sub>5</sub> peak. Chromatograms of the undigested substrate (InsP<sub>6</sub> substrate) and an acid hydrolysate of the substrate (InsP<sub>x</sub> standards) are shown for reference. The elution volume ranges for the various inositol polyphosphates are highlighted by vertical colored backgrounds (note that the notation for the InsP<sub>5</sub> products is based on the identity of the free hydroxyl group of the intermediate).

K54A mutant shows preference for initial cleavage at the 1D-5-phosphate of phytic acid, rather than for the 1D-4/6-phosphate as displayed by the WT enzyme (Fig. 5B). The positional speci-

ficity of the K54A mutant resembles closely that observed for the MINPP from *B. thetaiotaomicronn*, a predominant 5-phytase (Fig. 1C). This is the first experimental observation of an engineered change in positional specificity by a MINPP and suggests that specificity subsites C and/or D may play a role in determining the distribution of lower inositol polyphosphates generated by hydrolysis of InsP<sub>6</sub> by MINPPs.

## Discussion

In a search for new phytases with elevated thermostability, we have identified a subfamily of predominantly extracellular MINPPs that bear characteristic sequence inserts relative to other HP2P. We name these insertions U-loops. Biochemical, biophysical, and structural characterization of one of these enzymes, the MINPP from *B. longum* subsp. *infantis* ATCC 15697, has provided new perspectives into the roles of this insertion in protein stability and ligand binding. We find that U-loop residues influence thermal stability, recovery of activity after heating, and kinetic parameters for hydrolysis of phytate. Furthermore, this structural feature suggests a basis by which recognition of phytic acid can be extended to specificity subsites underutilized in other HP2P. However, a consequence of the presence of the U-loop is the requirement for large conformational changes during the catalytic cycle. Driven by engagement of a scissile phosphate ion in the A-pocket, the U-loop closes over the active site upon ligand-binding, acting as a “lid extension” of the  $\alpha$ -domain, helping to define the enzyme active site and contributing residues to specificity pockets C, D, and E. Although in the closed conformation, the U-loop shields the bound substrate, and presumably hinders diffusional escape of the first stage hydrolysis product. Restoration of the open state conformation in the phosphohistidine intermediate allows product egress, before return to a half-closed state to allow breakdown of the intermediate and generation of the second stage product, orthophosphate.

Phytases, along with other members of the histidine phosphatase superfamily, hydrolyze phosphomonoesters in two distinct steps (21, 44). In the first, the formation of an obligatory phosphohistidine intermediate through nucleophilic attack leads to release of the first stage product, in the case of phytases, a lower phosphorylated inositol. The second step involves recruitment of a water molecule and hydrolysis of the intermediate leading to release of orthophosphate. The structures of HP2 phytases at various stages along this catalytic pathway provide useful insights into the interplay between protein structure and catalytic mechanism. Previous studies have reported details of apo-, substrate analog-, and product-bound forms of various family members but the considerable majority of these have revealed no significant enzyme conformational changes (8, 45). The exception to this is the *E. coli* phytase, AppA, where more pronounced conformational changes occur upon ligand binding, particularly effecting residues 20–25 (Fig. S8) (9). Although Arg-20 in *EcAppA* moves significantly to form a contact with the scissile phosphate, in other HP2P the corresponding arginine residue is effectively already in a “bound” conformation in the apo-structure. In addition, the main chain of Lys-24 moves by 4.7 Å, leading to a 15 Å shift of the N<sub>z</sub> atom of its side chain



## Phytase Induced-Fit Structural Mechanism

to make a contact with the ligand in the C and D specificity pockets. However, despite the large changes involving this residue, the conformational changes are localized. A more extensive conformational change occurs on ligand binding by *B*MINPP. As part of this, the main chain of Lys-54 moves 5.8 Å to adopt its ligand-bound position where it contributes to the same specificity pockets as Lys-24 in *Ec*AppA. Although Lys-24 in *Ec*AppA and Lys-54 in *B*MINPP are not aligned in terms of sequence (their  $\alpha$ -domains have very different folds) their N<sub>z</sub> atoms lie only 3.6 Å apart when their InsS<sub>6</sub>-bound structures are superimposed suggesting similar roles. That we have shown Lys-54 to play a role in determining positional specificity in *B*MINPP may suggest a similar role for Lys-24 in *Ec*AppA.

Exoenzymes find application in a wide range of biotechnological and industrial processes, and are frequent targets for enzyme discovery (46). Extreme environments are commonly exploited to discover new and robust enzymes well-suited for use in industrial applications. The extracellular MINPPs revealed by phylogenetic analysis in this study are found in a variety of often extreme environments. MINPPs bearing a type A U-loop are typically extracellular membrane-anchored enzymes of Gram-positive bacteria such as *Microbacterium hydrocarbonoxidans* and *B. longum*. *M. hydrocarbonoxidans* is an actinobacterium adapted to harsh environments. It is able to survive, for example, in oil-contaminated soil (47) or toluene filters (48). Enzymes containing the type B loop are found in a variety of environmental Gram-negative bacteria including *Cupriavidus gracilis*, a heavy-metal resistant bacterium first found in industrial biotopes (49), and *Variovorax paradoxus*, a Gram-negative,  $\beta$  proteobacterium able to utilize a wide array of recalcitrant organic pollutant and heavy metals (50).

The preferred substrate and function of MINPPs *in vivo* are still uncertain, despite there being evidence of a role for these enzymes in a variety of cellular processes and organisms (29–32). *B. longum* subsp. *infantis* ATCC 15697 is a human gut commensal known for its positive role in the early development of the infant gut. It decreases intestinal permeability, displays anti-inflammatory activity in intestinal cells, and has been associated with a lower risk of necrotizing enterocolitis in premature infants (51). The bacterium is genetically well-adapted to coexistence with its human host and it is able to digest human milk oligosaccharides due to the presence of a Human Milk Oligosaccharides gene cluster (HMO cluster I) encoding multiple oligosaccharide transporters and glycosyl hydrolases not found in other bifidobacteria (51). *B*MINPP is membrane-anchored and extracellular, and thus exposed to phosphate-containing substrates in human milk, for example, *myo*-inositol polyphosphates (52) or casein phosphopeptides (53). Considering the low and tightly regulated phosphorus levels in human milk compared with other mammals (54), it may be the case that a role for this MINPP is to hydrolyze *myo*-inositol polyphosphates, leading to an increase in the local concentration of bioavailable free phosphate whereas maintaining phosphate homeostasis in the gut (55). However, *B*MINPP only hydrolyzes phytic acid to produce *myo*-inositol trisphosphates as an end point (35). This may be related to the unusual population of

specificity subsites C, D, and E relative to other MINPPs and HP2P in general.

By far the most common application of microbial HP2P in animal nutrition has been their use as additives to animal feeds (3, 56, 57). Genetically modified phytase crops (58, 59) and transgenic animals (*e.g.* pigs (60)) are other approaches. Multiple inositol polyphosphate phosphatases have yet to find application in any of these areas although chicken MINPP has been suggested as a possible vehicle for the development of transgenic chicken (61). The structural basis for the role of the U-loop in *B*MINPP in enhancing protein thermostability and in tailoring the recognition of inositol polyphosphate substrates as described herein will be of use in efforts to alter the specificity and stability of HP2Ps by protein engineering. Indeed, given their prevalence in the environment and inherent catalytic flexibility, MINPPs have the potential to provide a rich source of new enzymes for development for animal feed enzyme applications. Given their lack of positional specificity toward phytate, MINPPs may prove attractive targets for development as a new avenue for feed enzymes, most reasonably utilized in conjunction with highly active but more specific conventional HP2P (62). Certainly, the ability to alter the positional specificity of a MINPP, as demonstrated in this work for *B*MINPP, holds promise for the development of highly efficient animal feed enzymes acting synergistically to effect the complete dephosphorylation of phytic acid.

## Experimental procedures

### Phylogenesis

The evolutionary history of clade 2 histidine phosphatases was inferred using the Maximum Likelihood method based on the JTT matrix-based model (63). Amino acid sequences were aligned with MUSCLE (64) and Jalview (65) was used for manual editing of the resulting multiple sequence alignments. Initial trees for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (4 categories; +G, parameter = 2.1429). The analysis involved 51 amino acid sequences. All positions with less than 5% site coverage were eliminated. That is, fewer than 95% alignment gaps, missing data, and ambiguous bases were allowed at any position. There were a total of 720 positions in the final data set. Evolutionary analyses were conducted in MEGA7 (66).

### Cloning and site-directed mutagenesis

The MINPP gene of *B. longum* subsp. *infantis* ATCC 15697 (*B*MINPP; GenBank™ ACJ51391.1) was supplied by Vicente Monedero (IATA-CSIC, Spain). *B*MINPP was recloned into the isopropyl thio- $\beta$ -D-galactopyranoside (IPTG)-inducible pOPINF (pOPIN Vector Suite, Protein Production, UK) and pET28a expression vectors. The sequence was cloned in a truncated form (residues 33–545 of the 623 residue full-length protein) excluding the signal peptide and C-terminal sortase-dependent cell wall-anchoring region, and fused to an N-terminal

3C-protease cleavable His<sub>6</sub> tag (residues MAHHHHHSS-GLVLFQ|GP, where | indicates the 3C-protease cleavage site; pOPINF vector) or an N-terminal thrombin cleavable His<sub>6</sub> tag (residues MGSSHHHHHSSGLVPR|GSHMAS, where | again indicates the cleavage site; pET28a vector). All experiments involving BMINPP with WT sequence utilized purified recombinant protein produced by means of the pOPINF construct, whereas experiments involving site-directed variants used protein generated using the pET28a construct and the modified QuikChange site-directed mutagenesis method reported by Liu and Naismith (67).

### Expression and purification

Transformed Rosetta(DE3) pLysS cell cultures were incubated at 37 °C and 180 rpm. On reaching an OD<sub>600</sub> of 0.6, induction was performed using 0.5 mM IPTG. Cells were chilled to 25 °C and left to grow overnight. They were then harvested by centrifugation at 4 °C, 5500 rpm for 20 min. Pellets were resuspended in 30 ml of NaH<sub>2</sub>PO<sub>4</sub> lysis buffer (50 mM NaH<sub>2</sub>PO<sub>4</sub>, pH 7.8, 300 mM NaCl, 200 mM imidazole, 0.5% (v/v) Triton) or Tris lysis buffer (50 mM Tris-HCl, pH 8, 300 mM NaCl, 10 mM imidazole), snap frozen in liquid nitrogen, and stored at -80 °C before thawing and lysis of cells by means of a French Press. The soluble fraction of the sample was separated by centrifugation at 4 °C, 15,000 rpm for 30 min and the overexpressed protein was isolated by Ni-NTA IMAC over a 20-500 mM imidazole gradient at pH 8.0.

pOPINF constructs containing a 3C-protease recognition site were dialyzed overnight in Tris buffer (20 mM Tris-HCl, pH 8.4, 150 mM NaCl, 2.5 mM CaCl<sub>2</sub>) in the presence of a 3C-His-tagged protease at a concentration 40× lower than that of the overexpressed protein. A second Ni-NTA IMAC was carried out for the separation of the 3C-His-tagged protein from the sample. pET28a constructs containing a thrombin recognition site were dialyzed overnight in Tris buffer (50 mM Tris-HCl, pH 7.4, 0.5 M NaCl, 20 mM imidazole). The buffer was replaced and the sample dialyzed again for 2 days in the presence of thrombin at a concentration of 2 units/mg of protein of interest. In both cases, cleaved BMINPP was concentrated using an Amicon Ultracentrifuge filter unit (10 kDa cutoff) and gel filtered using a HiLoad 16/600 Superdex 75 pg column (GE Healthcare) and a running buffer containing 20 mM HEPES pH 7.4, 150 mM NaCl. Protein samples with a purity of at least 99% as estimated by SDS-PAGE were collected. Estimates of the enzyme concentrations were made from absorbance measurements at 280 nm using a NanoDrop One Microvolume UV Spectrophotometer (ThermoScientific).

### Protein crystallization

Crystals of apoprotein (BMINPP), of BMINPP in complex with *myo*-inositol hexakisulfate (BMINPP:InsS<sub>6</sub>), of the phosphorylated mutant BMINPP E401Q (BMINPP-Pi) and BMINPP E401Q in complex with orthophosphate (BMINPP:Pi) were obtained by sitting drop vapor diffusion using a protein concentrated to 9 mg/ml in 20 mM HEPES, pH 7.4, 150 mM NaCl storage buffer. For the crystallization of BMINPP:InsS<sub>6</sub>, co-crystallization plates were set up with the addition of 1 mM

InsS<sub>6</sub> to the protein sample (InsS<sub>6</sub>, *myo*-inositol hexakisulfate hexapotassium salt, product number ACM28434255, was obtained from Alfa Chemistry, Protheragen Inc.). Drops consisted of 50% protein sample and 50% reservoir solution. BMINPP apo, BMINPP:Pi, and BMINPP:Pi crystals grew in drops with a reservoir solution composed of 0.1 M MES, pH 6.5, 10 mM zinc chloride, 18% (w/v) PEG 6000. Crystals of the BMINPP:InsS<sub>6</sub> complex grew in drops with a reservoir solution of 10% (w/v) PEG 1000 and 10% (w/v) PEG 8000. Crystals were harvested, cryo-protected by the addition of 30% (v/v) glycerol or PEG 400 to the mother liquor and frozen in liquid nitrogen.

### X-ray diffraction data collection and crystal structure determination

Diffraction experiments were performed at the beamlines I02 and I03 of the Diamond Light Source (Oxfordshire, UK) using a Pilatus3 6M detector and BART sample changer. Data reduction was performed with xia2 (68). Initial phases for the structure of the BMINPP:InsS<sub>6</sub> complex were determined by molecular replacement using the program Phaser (69) and the structure of the complex of BtMINPP with InsS<sub>6</sub> (PDB ID 4FDU). The phasing of the structure of the BMINPP:InsS<sub>6</sub> complex was made difficult by a pronounced conformational change and was achieved by decreasing the resolution cutoff to 4 Å and performing separate molecular replacement searches for each of the BMINPP domains. Initial phase estimates were used to generate difference maps in Coot (70). Convergence of cycles of rebuilding in Coot and refinement using phenix.refine (71) gave a refined structure for the two copies of the enzyme in the asymmetric unit. Calculation of Polder OMIT maps (38) revealed significant residual electron density in both active sites corresponding to bound InsS<sub>6</sub>. Careful inspection revealed static disorder at both sites and the inhibitor was added to the model in two orientations presenting either the 4- and 6-sulfate bound at the catalytic center. Further refinement yielded a final structural model with *R*<sub>work</sub> 20.9% and *R*<sub>free</sub> 23.8% for all data to 1.84 Å resolution. Structure solution and refinement of the apo-, orthophosphate-bound, and phosphohistidine intermediate forms followed by essentially the same methods except that the structure of BtMINPP in complex with orthophosphate (PDB code 4FDT) was used as search model for molecular replacement. All refined structures were validated using MolProbity (72) and the wwPDB Validation Service (RRID:SCR\_018135). All data collection and refinement statistics are reported in Table 1.

### Cloning, expression, and purification of BtMINPP and reference HP2 phytases

BtMINPP and the well-studied reference HP2P from *A. niger* (AnPhyA) and *E. coli* (EcAppA) were cloned and purified as part of this study.

The AnPhyA gene was codon optimized for expression in *Pichia pastoris* and synthesized by GenScript. For Gateway cloning the gene fragment was amplified with an N-terminal 3C protease site, a C-terminal His<sub>6</sub> tag, and Gateway recombination adapters by a two-step PCR using the entry vector pDONR207 and the destination vector pPICZα-DEST. The

# Phytase Induced-Fit Structural Mechanism

**Table 1**

**X-ray data collection and structure refinement statistics**

Numbers in parentheses refer to high resolution data.

Protein	Apo	InsS <sub>6</sub> complex	Phosphohistidine intermediate	P <sub>i</sub> complex
PDB ID	6RXD	6RXE	6RXF	6RXG
<b>Data collection</b>				
Beamline	Diamond I02	Diamond I02	Diamond I03	Diamond I03
Wavelength (Å)	0.9795	0.9795	0.9763	0.9762
Space group	P1	P2 <sub>1</sub>	P1	P1
Cell parameters				
<i>a</i> , <i>b</i> , <i>c</i> (Å)	54.7, 72.9, 88.2	70.8, 106.4, 76.5	55.5, 73.1, 89.0	54.6, 72.8, 87.3
$\alpha$ , $\beta$ , $\gamma$ (°)	71.6, 72.1, 77.2	90.0, 112.6, 90.0	71.3, 73.0, 79.0	71.4, 72.2, 76.8
Resolution limit (Å)	46.53-1.65 (1.69-1.65)	70.57-1.84 (1.91-1.84)	68.9-2.40 (2.44-2.40)	68.23-1.71 (1.75-1.71)
<i>R</i> merge	0.059 (0.45)	0.159 (0.625)	0.233 (0.826)	0.063 (0.525)
CC1/2	0.994 (0.645)	0.979 (0.603)	0.925 (0.328)	0.984 (0.503)
$\langle I \rangle / \sigma(I)$	8.4 (1.9)	4.6 (1.6)	2.3 (1.5)	7.5 (1.4)
Completeness (%)	96.3 (94.9)	99.65 (98.9)	94.6 (88.8)	95.5 (95.0)
Multiplicity	2.2 (2.1)	3.4 (3.3)	3.4 (3.3)	1.9 (1.9)
Overall temperature factor (Å <sup>2</sup> )	18.6	10.9	25.5	17.3
<b>Refinement</b>				
Protein monomers per asymmetric unit	2	2	2	2
Number of protein residues	509	509	509	509
Total atoms	9,328	9,374	8,117	15,951
Water molecules	1,287	1,447	286	805
<i>R</i> <sub>work</sub>	15.3%	20.9%	18.5%	14.7%
<i>R</i> <sub>free</sub>	17.7%	23.8%	23.9%	19.4%
Ramachandran analysis (%)				
Most favored	98.11%	98.1%	96.8%	97.7%
Outliers	0.28%	0.2%	0.2%	0.3%
RMS deviations				
Bonds (Å)	0.006	0.004	0.006	0.009
Angles (°)	1.008	0.70	0.80	0.927
Planes (Å)	0.004	0.004	0.005	0.006
Mean atomic <i>B</i> -value (Å <sup>2</sup> )	23.0	16.99	28.56	27.9
Macromolecule	NA <sup>a</sup>	15.16	28.31	27.3
Ligands	NA	15.85	45.06	
Solvent	26.95	26.95	34.77	38.8

<sup>a</sup>NA, not applicable.

resulting construct was linearized and transformed into *P. pastoris* KM71H (OCH1::G418R) by electroporation and spread onto 6-well LB plates containing 100 µg/ml of kanamycin and 100 µg/ml of zeocin, and incubated at 30 °C for 3-5 days. A single colony was selected and used to inoculate 5 ml of BMGY media (1% yeast extract, 2% peptone, 100 mM potassium phosphate buffer, pH 5.0, 1.34% yeast nitrogen base, 4 × 10<sup>-5</sup>% biotin, 1% glycerol, 100 µg/ml of kanamycin). After overnight incubation at 30 °C with shaking at 200 rpm, the cells were resuspended in 5 ml of BMMY media (1% yeast extract, 2% peptone, 100 mM potassium phosphate buffer, pH 5.0, 1.34% yeast nitrogen base, 4 × 10<sup>-5</sup>% biotin, 1% methanol, 100 µg/ml of kanamycin) and incubated for a further 4 days at 30 °C with shaking at 200 rpm. As the recombinant AnPhyA protein is secreted into the media, supernatants were harvested by centrifugation for 10 min at 4,000 rpm. Prior to purification, the pH of the supernatants was adjusted to pH 8.0 by addition of 10 N NaOH; and precipitate was removed by centrifugation for 10 min at 4,000 rpm. AnPhyA was subsequently purified from the supernatant by Ni-NTA metal affinity chromatography and stored at -80 °C.

The gene encoding *E. coli* AppA (*EcAppA*) was amplified from the genome of BL21 (DE3) pLysS and cloned into pOPINB. The constructs was designed for the cytoplasmic IPTG-inducible expression of an N-terminal cleavable His tag protein. The sequence was confirmed by sequencing and the protein expressed in soluble form using the *E. coli* B strain Shuffle Express T7 (73). The protein was then purified by Ni-NTA metal affinity chromatography followed by gel filtration

through a HiLoad 16/600 Superdex 75 pg (GE Healthcare) in 200 mM sodium acetate, 150 mM NaCl, pH 4.5, and stored at -80 °C. *BtMINPP* was expressed and purified according to previously established methods (22).

### Phosphate release assay

This assay allows the determination of the free phosphate released by hydrolysis of InsP<sub>6</sub> by the molybdenum blue reaction (74). The absorbance of molybdenum blue is measured at 700 nm and is proportional to the P<sub>i</sub> concentration. A typical calibration curve shows the assay to be in the linear range from 10 µM to 2.5 mM orthophosphate. 5 mM phytic acid dipotassium salt (>95% pure) was used as substrate. Reactions were performed at room temperature at pH 3.5, 5.5, or 7.4. Reactions of 50 or 100 µl volume were stopped by the addition of equal amounts of a freshly prepared solution made of 4 parts of reagent A (12 mM ammonium molybdate tetrahydrate, 5.4% saturated sulfuric acid) and 1 part of reagent B (0.4 M iron(II) sulfate heptahydrate plus a few drops of saturated sulfuric acid). Absorbance was measured at 700 nm after 30 min using a Hidex Sense plate reader. Control reactions of buffer only, substrate only, and enzyme only were set up simultaneously as well as a calibration curve of increasing concentration of orthophosphate.

### Phytase activity recovery after heating of disulfide mutants

The phytase activity of mutants D1, D2, and D1D2 were assessed and compared with the activity of *BtMINPP* WT after

30 min incubation at a range of temperatures from 25 to 80°C. After incubation, the samples were cooled to room temperature and mixed with 5 mM InsP<sub>6</sub> at pH 5.5 and reactions were allowed to proceed for 30 min at room temperature. Phosphate release was measured by means of the Phosphate Release Assay. Results are displayed as micromole of P<sub>i</sub> per minute per mg of protein.

#### Measurement of enzyme kinetic parameters

Reactions of 50 μl were set up in triplicate at fixed concentrations of enzymes (25 nM) and increasing concentration of substrate (50, 100, 200, 400, 600, 800, 1200, and 2500 μM). They were incubated for 5 min at room temperature. The buffer chosen was 200 mM sodium acetate, pH 5.5, 0.15 M NaCl. Reactions were inactivated by addition of molybdenum blue reagent in equal parts and the absorbance at 700 nm was measured after 30 min incubation of the samples with the stopping reagent. Data were processed with the “nls” function provided in R 41 (see <https://stat.ethz.ch/R-manual/R-devel/library/stats/html/nls.html>), that determines the nonlinear least-squares estimates of the parameters of a nonlinear model. In this analysis, the nonlinear model is the Michaelis-Menten equation. The goodness of fit of the model was confirmed by checking residual error values and *t* test.

#### Identification of inositol polyphosphates by HPLC

*myo*-Inositol 1,2,3,4,5,6-hexakisphosphate, dodecasodium salt (InsP<sub>6</sub>, 1 mM, *Zea mays*, Merck, 99% pure, confirmed by HPLC), was used as substrate. Enzymes were used at 25 nM in 200 mM sodium acetate, pH 5.5, 150 mM NaCl. Reactions were stopped at 5 and 10 min by boiling samples at 100°C for 10 min. Samples were diluted 5× before injection. Inositol polyphosphate standards were generated by the hydrolysis of InsP<sub>6</sub> in 1 M HCl, 120°C for 24 h. The HPLC system consisted of a first pump for sample injection (Jasco PU-2089 I Plus—Quaternary inert Pump) connected in series to two CarboPAC PA200 columns (3 × 50 mm, 3 × 250 mm) in which the InsP<sub>x</sub> species were efficiently separated (enantiomers, however, cannot be resolved) before reaching a chamber in which they were chaotically mixed with a reagent (0.1% Fe(NO<sub>3</sub>)<sub>2</sub>, 2% HClO<sub>4</sub>), which was injected by a second pump (Jasco PU-1585 Intelligent HPLC Pump). This allows UV absorbance detection at 290 nm (range 1.28 nm, Jasco UV 1575 Intelligent UV-visible detector: 16 μl cell). Samples were separated in a methane sulfonic acid gradient (0–0.6 M), flow rate 0.4 ml/min, with water as a counter eluent, reagents were injected at a flow rate of 0.2 ml/min. The total run time for each sample was 50 min: 25 min of gradient, 14 min of 0.6 M methane sulfonic acid, 11 min of water. The peak areas were calculated by integration using the software provided by Jasco (ChromNAV, version 1.19.01). The identities of inositol polyphosphates generated during hydrolysis were determined by reference to the retention times of peaks resulting from a standard sample of chemically hydrolyzed InsP<sub>6</sub> (HCl, 120°C, 24 h).

#### Calorimetry experiments

A VP-DSC (Microcal Inc.) was used for all calorimetry experiments. Initially, 20 buffer readings were taken to build

the thermal history of the instrument. The last runs were used as baseline in subsequent data analysis. The temperature gradient was set to 10–110°C with a scan rate of 200°C/h, which assured high sensitivity without excessive sharpness of peaks. A pre-scan of 5 min was added at the beginning of each read. 350 μl of sample solution at a concentration of 1 mg/ml was used for each run. Experiments were carried out at pH 3.5, 5.5, and 7.5 (pH 3.5, 0.2 M glycine-HCl; pH 5.5, 0.2 M sodium acetate; pH 7.5, 0.2 M HEPES), representing the typical pH optima observed for MINPP phytase activity (22).

#### Data availability

Coordinates and diffraction data for the crystal structures of *B. longum* MINPP in the apo form, in the phosphohistidine intermediate form and in complexes with inositol hexasulfate and phosphate have been deposited in the PDB with accession codes 6RXD, 6RXE, 6RXF and 6RXG.

**Acknowledgments**—We thank Monica Haros and Vincente Monodero (Instituto de Agroquímica y Tecnología de Alimentos, Valencia, Spain) for providing a clone of *PhyA* (BLON\_0263) from *B. longum* subsp. *infantis* ATCC 15697. We thank Diamond Light Source for beamtime under proposals MX7641, MX9475, and MX13467 and the staff of beamlines I02 and I03 for assistance with crystal testing and data collection.

**Author contributions**—I. M. A. data curation; I. M. A. validation; I. M. A., M. A. Z., A. W. L., M. S., C. A. B., and A. M. H. investigation; I. M. A. visualization; I. M. A., C. A. B., and A. M. H. methodology; I. M. A., M. A. Z., A. W. L., M. S., I. K., M. R. B., C. A. B., and A. M. H. writing-review and editing; I. K., M. R. B., C. A. B., and A. M. H. conceptualization; I. K., C. A. B., and A. M. H. supervision; I. K., M. R. B., C. A. B., and A. M. H. funding acquisition; C. A. B. and A. M. H. project administration; A. M. H. formal analysis; A. M. H. writing-original draft.

**Funding and additional information**—This work was supported by the Biotechnology and Biological Sciences Research Council and AB Vista Ltd. through IPA Award BB/M022978/1 and joint UEA-AB Vista Ltd. Ph.D. studentships (to I. M. A. and M. Z.).

**Conflict of interest**—The authors declare that they have no conflict of interest with the contents of this article.

**Abbreviations**—The abbreviations used are: InsP<sub>6</sub>, *myo*-inositol hexakisphosphate; HP2, clade 2 histidine phosphatases; HP2P, HP2 with phytase activity; InsP<sub>x</sub>, *myo*-inositol polyphosphates; InsS<sub>6</sub>, *myo*-inositol hexakisulfate; MINPP, multiple inositol polyphosphate phosphatase; RMS deviation, root mean square deviation; DSC, differential scanning calorimetry; PDB, Protein Data Bank; IPTG, isopropyl thio-β-D-galactopyranoside; Ni-NTA, nickel-nitrilotriacetic acid.

#### References

- Humer, E., Schwarz, C., and Schedle, K. (2015) Phytate in pig and poultry nutrition. *J. Anim. Physiol. Anim. Nutr. (Berl.)* **99**, 605–625 [CrossRef](#) [Medline](#)

## Phytase Induced-Fit Structural Mechanism

- Mallin, M. A., and Cahoon, L. B. (2020) The hidden impacts of phosphorus pollution to streams and rivers. *Bioscience* **70**, 315–329 [CrossRef](#)
- Romano, N., and Kumar, V. (2018) Phytase in animal feed. in *Enzymes in Human and Animal Nutrition: Principles and Perspectives*, pp. 73–88, Elsevier Inc., The Netherlands [CrossRef](#)
- Dersjant-Li, Y., Awati, A., Schulze, H., and Partridge, G. (2015) Phytase in non-ruminant animal nutrition: A critical review on phytase activities in the gastrointestinal tract and influencing factors. *J. Sci. Food Agric.* **95**, 878–896 [CrossRef](#) [Medline](#)
- Yao, M. Z., Zhang, Y. H., Lu, W. L., Hu, M. Q., Wang, W., and Liang, A. H. (2012) Phytases: Crystal structures, protein engineering and potential biotechnological applications. *J. Appl. Microbiol.* **112**, 1–14 [CrossRef](#) [Medline](#)
- Shivange, A. V., and Schwaneberg, U. (2017) Recent advances in directed phytase evolution and rational phytase engineering. in *Directed enzyme evolution: advances and applications*, pp. 145–172, Springer International Publishing, [CrossRef](#)
- Irving, G. C. J. (1980) Intermediates in the dephosphorylation of P6-inositols by phytase enzymes. in *Inositol Phosphates: their Chemistry, Biochemistry and Physiology* (Cosgrove, D. J., ed) pp. 99–117, Elsevier Scientific Publishing Company, New York (Accessed July 24, 2020)
- Ostanin, K., Harms, E. H., Stevis, P. E., Kuciel, R., Zhou, M. M., and Van Etten, R. L. (1992) Overexpression, site-directed mutagenesis, and mechanism of *Escherichia coli* acid phosphatase. *J. Biol. Chem.* **267**, 22830–22836 [Medline](#)
- Lim, D., Golovan, S., Forsberg, C. W., and Jia, Z. (2000) Crystal structures of *Escherichia coli* phytase and its complex with phytate. *Nat. Struct. Biol.* **7**, 108–113 [CrossRef](#) [Medline](#)
- Sanchez-Romero, I., Ariza, A., Wilson, K. S., Skjøt, M., Vind, J., de Maria, L., Skov, L. K., and Sanchez-Ruiz, J. M. (2013) Mechanism of protein kinetic stabilization by engineered disulfide crosslinks. *PLoS ONE* **8**, e70013 [CrossRef](#)
- Ariza, A., Moroz, O. V., Blagova, E. V., Turkenburg, J. P., Waterman, J., Roberts, S. M., Vind, J., Sjöholm, C., Lassen, S. F., De Maria, L., Glitsov, V., Skov, L. K., and Wilson, K. S. (2013) Degradation of phytate by the 6-phytase from *Hafnia alvei*: a combined structural and solution study. *PLoS ONE* **8**, e65062 [CrossRef](#) [Medline](#)
- Böhm, K., Herter, T., Müller, J. J., Borriss, R., and Heinemann, U. (2010) Crystal structure of *Klebsiella* sp. ASR1 phytase suggests substrate binding to a preformed active site that meets the requirements of a plant rhizosphere enzyme. *FEBS J.* **277**, 1284–1296 [CrossRef](#) [Medline](#)
- Dhatwalia, R., Singh, H., Reilly, T. J., and Tanner, J. J. (2015) Crystal structure and tartrate inhibition of *Legionella pneumophila* histidine acid phosphatase. *Arch. Biochem. Biophys.* **585**, 32–38 [CrossRef](#) [Medline](#)
- Kostrewa, D., Wyss, M., D'Arcy, A., and van Loon, A. P. (1999) Crystal structure of *Aspergillus niger* pH 2.5 acid phosphatase at 2.4 Å resolution. *J. Mol. Biol.* **288**, 965–974 [CrossRef](#) [Medline](#)
- Xiang, T., Liu, Q., Deacon, A. M., Koshy, M., Kriksunov, I. A., Lei, X. G., Hao, Q., and Thiel, D. J. (2004) Crystal structure of a heat-resilient phytase from *Aspergillus fumigatus*, carrying a phosphorylated histidine. *J. Mol. Biol.* **339**, 437–445 [CrossRef](#) [Medline](#)
- Ragon, M., Hoh, F., Aumelas, A., Chiche, L., Moulin, G., and Boze, H. (2009) Structure of *debaryomyces castellii* CBS 2923 phytase. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **65**, 321–326 [CrossRef](#) [Medline](#)
- Oh, B. C., Choi, W. C., Park, S., Kim, Y. O., and Oh, T. K. (2004) Biochemical properties and substrate specificities of alkaline and histidine acid phosphatases. *Appl. Microbiol. Biotechnol.* **63**, 362–372 [CrossRef](#) [Medline](#)
- Etten, R. L. V. (1982) Human prostatic acid phosphatase: a histidine phosphatase. *Ann. N. Y. Acad. Sci.* **390**, 27–51 [CrossRef](#) [Medline](#)
- Van Etten, R. L., Davidson, R., Stevis, P. E., MacArthur, H., and Moore, D. L. (1991) Covalent structure, disulfide bonding, and identification of reactive surface and active site residues of human prostatic acid phosphatase. *J. Biol. Chem.* **266**, 2313–2319 [Medline](#)
- Ostanin, K., and Van Etten, R. L. (1993) Asp304 of *Escherichia coli* acid phosphatase is involved in leaving group protonation. *J. Biol. Chem.* **268**, 20778–20784 [Medline](#)
- Rigden, D. J. (2008) The histidine phosphatase superfamily: structure and function. *Biochem. J.* **409**, 333–348 [CrossRef](#) [Medline](#)
- Stentz, R., Osborne, S., Horn, N., Li, A. W. H., Hautefort, I., Bongaerts, R., Rouyer, M., Bailey, P., Shears, S. B., Hemmings, A. M., Brearley, C. A., and Carding, S. R. (2014) A bacterial homolog of a eukaryotic inositol phosphate signaling enzyme mediates cross-kingdom dialog in the mammalian gut. *Cell Rep.* **6**, 646–656 [CrossRef](#) [Medline](#)
- Romano, P. R., Wang, J., O'Keefe, R. J., Puzas, J. E., Rosier, R. N., and Reynolds, P. R. (1998) HiPER1, a phosphatase of the endoplasmic reticulum with a role in chondrocyte maturation. *J. Cell Sci.* **111**, 803–813 [Medline](#)
- Chi, H., Tiller, G. E., Dasouki, M. J., Romano, P. R., Wang, J., O'Keefe, R. J., Puzas, J. E., Rosier, R. N., and Reynolds, P. R. (1999) Multiple inositol polyphosphate phosphatase: Evolution as a distinct group within the histidine phosphatase family and chromosomal localization of the human and mouse genes to chromosomes 10q23 and 19. *Genomics* **56**, 324–336 [CrossRef](#) [Medline](#)
- Kilaparty, S. P., Singh, A., Baltosser, W. H., and Ali, N. (2014) Computational analysis reveals a successive adaptation of multiple inositol polyphosphate phosphatase 1 in higher organisms through evolution. *Evol. Bioinform.* **10**, 239–250 [CrossRef](#)
- Cho, J., King, J. S., Qian, X., Harwood, A. J., and Shears, S. B. (2008) Dephosphorylation of 2,3-bisphosphoglycerate by MIPP expands the regulatory capacity of the Rapoport-Luebering glycolytic shunt. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 5998–6003 [CrossRef](#) [Medline](#)
- Nogimori, K., Hughes, P. J., Glennon, M. C., Hodgson, M. E., Putney, J. W., and Shears, S. B. (1991) Purification of an Inositol (1,3,4,5)-tetrakisphosphate 3-phosphatase activity from rat liver and the evaluation of its substrate specificity. *J. Biol. Chem.* **266**, 16499–16506 [Medline](#)
- Craxton, A., Caffrey, J. J., Burkhart, W., Safrany, S. T., and Shears, S. B. (1997) Molecular cloning and expression of a rat hepatic multiple inositol polyphosphate phosphatase. *Biochem. J.* **328**, 75–81 [CrossRef](#) [Medline](#)
- Yu, J., Leibiger, B., Yang, S. N., Caffery, J. J., Shears, S. B., Leibiger, I. B., Barker, C. J., and Berggren, P. O. (2003) Cytosolic multiple inositol polyphosphate phosphatase in the regulation of cytoplasmic free Ca<sup>2+</sup> concentration. *J. Biol. Chem.* **278**, 46210–46218 [CrossRef](#) [Medline](#)
- Windhorst, S., Lin, H., Blechner, C., Fanick, W., Brandt, L., Brehm, M. A., and Mayr, G. W. (2013) Tumor cells can employ extracellular Ins(1,2,3,4,5,6)P6 and multiple inositol-polyphosphate phosphatase 1 (MINPP1) dephosphorylation to improve their proliferation. *Biochem. J.* **450**, 115–125 [CrossRef](#) [Medline](#)
- Cheng, Y. L. L., and Andrew, D. J. J. (2015) Extracellular Mipp1 activity confers migratory advantage to epithelial cells during collective migration. *Cell Rep.* **13**, 2174–2188 [CrossRef](#) [Medline](#)
- Kilaparty, S. P., Agarwal, R., Singh, P., Kannan, K., and Ali, N. (2016) Endoplasmic reticulum stress-induced apoptosis accompanies enhanced expression of multiple inositol polyphosphate phosphatase 1 (Minpp1): a possible role for Minpp1 in cellular stress response. *Cell Stress Chaperones* **21**, 593–608 [CrossRef](#) [Medline](#)
- Turner, P., Mamo, G., and Karlsson, E. N. (2007) Potential and utilization of thermophiles and thermostable enzymes in biorefining. *Microb. Cell Fact.* **6**, 9 [CrossRef](#) [Medline](#)
- Arbolea, S., Watkins, C., Stanton, C., and Ross, R. P. (2016) Gut bifidobacteria populations in human health and aging. *Front. Microbiol.* **7**, [CrossRef](#)
- Tamayo-Ramos, J. A., Sanz-Penella, J. M., Yebra, M. J., Monedero, V., and Haros, M. (2012) Novel phytases from *Bifidobacterium pseudocatenulatum* ATCC 27919 and *Bifidobacterium longum* subsp. infantis ATCC 15697. *Appl. Environ. Microbiol.* **78**, 5013–5015 [CrossRef](#) [Medline](#)
- Greiner, R., Carlsson, N. G., and Alminger, M. L. (2000) Stereospecificity of myo-inositol hexakisphosphate dephosphorylation by a phytate-degrading enzyme of *Escherichia coli*. *J. Biotechnol.* **84**, 53–62 [CrossRef](#) [Medline](#)
- Liu, Q., Huang, Q., Lei, X. G., and Hao, Q. (2004) Crystallographic snapshots of *Aspergillus fumigatus* phytase, revealing its enzymatic dynamics. *Structure* **12**, 1575–1583 [CrossRef](#) [Medline](#)
- Ragon, M., Aumelas, A., Chemardin, P., Galvez, S., Moulin, G., and Boze, H. (2008) Complete hydrolysis of myo-inositol hexakisphosphate by a novel phytase from *Debaryomyces castellii* CBS 2923. *Appl. Microbiol. Biotechnol.* **78**, 47–53 [CrossRef](#) [Medline](#)

39. Barrientos, L., Scott, J. J., and Murthy, P. P. N. (1994) Specificity of hydrolysis of phytic acid by alkaline phytase from lily pollen. *Plant Physiol.* **106**, 1489–1495 [CrossRef Medline](#)
40. Greiner, R., Alminger, M. L., and Carlsson, N. G. (2001) Stereospecificity of *myo*-inositol hexakisphosphate dephosphorylation by a phytate-degrading enzyme of baker's yeast. *J. Agric. Food Chem.* **49**, 2228–2233 [CrossRef Medline](#)
41. Oakley, A. J. (2010) The structure of *Aspergillus niger* phytase PhyA in complex with a phytate mimetic. *Biochem. Biophys. Res. Commun.* **397**, 745–749 [CrossRef Medline](#)
42. Qi, G., and Hayward, S. (2009) Database of ligand-induced domain movements in enzymes. *BMC Struct. Biol.* **9**, 1–9 [CrossRef Medline](#)
43. Ashkenazy, H., Abadi, S., Martz, E., Chay, O., Mayrose, I., Pupko, T., and Ben-Tal, N. (2016) ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res.* **44**, W344–W350 [CrossRef](#)
44. Wang, Y., Liu, L., Wei, Z., Cheng, Z., Lin, Y., and Gong, W. (2006) Seeing the process of histidine phosphorylation in human bisphosphoglycerate mutase. *J. Biol. Chem.* **281**, 39642–39648 [CrossRef Medline](#)
45. Ortlund, E., LaCount, M. W., and Lebioda, L. (2003) Crystal structures of human prostatic acid phosphatase in complex with a phosphate ion and  $\alpha$ -benzylaminobenzylphosphonic acid update the mechanistic picture and offer new insights into inhibitor design. *Biochemistry* **42**, 383–389 [CrossRef Medline](#)
46. Riedel, K., and Grunau, A. (2011) *Exoenzymes*. in *Encyclopedia of Geobiology* (Reitner, J., and Thiel, V., eds) pp. 355–359, Springer, Netherlands, Dordrecht [CrossRef](#)
47. Schippers, A., Bosecker, K., Spröer, C., and Schumann, P. (2005) *Microbacterium oleivorans* sp. nov., and *Microbacterium hydrocarbonoxydans* sp. nov., novel crude-oil-degrading Gram-positive bacteria. *Int. J. Syst. Evol. Microbiol.* **55**, 655–660 [CrossRef Medline](#)
48. Prenafeta-Boldú, F. X., Guivernau, M., Gallastegui, G., Viñas, M., de Hoog, G. S., and Elías, A. (2012) Fungal/bacterial interactions during the biodegradation of TEX hydrocarbons (toluene, ethylbenzene and p-xylene) in gas biofilters operated under xerophilic conditions. *FEMS Microbiol. Ecol.* **80**, 722–734 [CrossRef Medline](#)
49. Goris, J., Vos, P. D., Coenye, T., Hoste, B., Janssens, D., Brim, H., Diels, L., Mergeay, M., Kersters, K., and Vandamme, P. (2001) Classification of metal-resistant bacteria from industrial biotopes as *Ralstonia campinensis* sp. nov., *Ralstonia metallidurans* sp. nov., and *Ralstonia basiliensis* Steinle et al. 1998. *Int. J. Syst. Evol. Microbiol.* **51**, 1773–1782 [CrossRef Medline](#)
50. Satola, B., Wübbeler, J. H., and Steinbüchel, A. (2013) Metabolic characteristics of the species *Variovorax paradoxus*. *Appl. Microbiol. Biotechnol.* **97**, 541–560 [CrossRef Medline](#)
51. Underwood, M. A., German, J. B., Lebrilla, C. B., and Mills, D. A. (2015) *Bifidobacterium longum* subspecies *infantis*: champion colonizer of the infant gut. *Pediatr. Res.* **77**, 229–235 [CrossRef Medline](#)
52. Indyk, H. E., Saldo, S. C., White, P. M., Dole, M. N., Gill, B. D., and Woolard, D. C. (2016) The free and total *myo*-inositol contents of early lactation and seasonal bovine milk. *Int. Dairy J.* **56**, 33–37 [CrossRef](#)
53. Ferranti, P., Traisci, M. V., Picariello, G., Nasi, A., Boschi, V., Siervo, M., Falconi, C., Chianese, L., and Addeo, F. (2004) Casein proteolysis in human milk: tracing the pattern of casein breakdown and the formation of potential bioactive peptides. *J. Dairy Res.* **71**, 74–87 [CrossRef Medline](#)
54. Dror, D. K., and Allen, L. H. (2018) Overview of nutrients in human milk. *Adv. Nutr.* **9**, 278–294 [CrossRef](#)
55. Ejnar Lenstrup, B. (1926) The phosphorous content of human milk and cow's milk. *J. Biol. Chem.* **70**, 193–202
56. Lei, X. G., Weaver, J. D., Mullaney, E., Ullah, A. H., and Azain, M. J. (2013) Phytase, a new life for an “old” enzyme. *Annu. Rev. Anim. Biosci.* **1**, 283–309 [CrossRef Medline](#)
57. Menezes-Blackburn, D., Gabler, S., and Greiner, R. (2015) Performance of seven commercial phytases in an *in vitro* simulation of poultry digestive tract. *J. Agric. Food Chem.* **63**, 6142–6149 [CrossRef Medline](#)
58. Ma, X. F., Tudor, S., Butler, T., Ge, Y., Xi, Y., Bouton, J., Harrison, M., and Wang, Z. Y. (2012) Transgenic expression of phytase and acid phosphatase genes in alfalfa (*Medicago sativa*) leads to improved phosphate uptake in natural soils. *Mol. Breed.* **30**, 377–391 [CrossRef Medline](#)
59. Zhao, Y., Zhu, L., Lin, C., Shen, Z., and Xu, C. (2019) Transgenic soybean expressing a thermostable phytase as substitution for feed additive phytase. *Sci. Rep.* **9**, 1–7 [CrossRef Medline](#)
60. Forsberg, C. W., Meidinger, R. G., Liu, M., Cottrill, M., Golovan, S., and Phillips, J. P. (2013) Integration, stability and expression of the *E. coli* phytase transgene in the Cassie line of Yorkshire Enviropig. *Transgenic Res.* **22**, 379–389 [CrossRef Medline](#)
61. Cho, J., Choi, K., Darden, T., Reynolds, P. R., Petite, J. N., and Shears, S. B. (2006) Avian multiple inositol polyphosphate phosphatase is an active phytase that can be engineered to help ameliorate the planet's “phosphate crisis.” *J. Biotechnol.* **126**, 248–259 [CrossRef Medline](#)
62. Wyss, M., Brugger, R., Kronenberger, A., Rémy, R., Fimbel, R., Oesterhelt, G., Lehmann, M., and Van Loon, A. P. G. M. (1999) Biochemical characterization of fungal phytases (*myo*-inositol hexakisphosphate phosphohydrolases): catalytic properties. *Appl. Environ. Microbiol.* **65**, 367–373 [CrossRef Medline](#)
63. Jones, D. T., Taylor, W. R., and Thornton, J. M. (1992) The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* **8**, 275–282 [CrossRef Medline](#)
64. Edgar, R. C. (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**, 113 [CrossRef Medline](#)
65. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M., and Barton, G. J. (2009) Jalview Version 2: a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191 [CrossRef Medline](#)
66. Kumar, S., Stecher, G., Tamura, K., and Dudley, J. (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. E* **33**, 1870–1874 [CrossRef Medline](#)
67. Liu, H., and Naismith, J. H. (2008) An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC Biotechnol.* **8**, 91 [CrossRef Medline](#)
68. Winter, G., Lobley, C. M. C., and Prince, S. M. (2013) Decision making in xia2. *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1260–1273 [CrossRef Medline](#)
69. McCoy, A. J. (2007) Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **63**, 32–41 [CrossRef Medline](#)
70. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 486–501 [CrossRef Medline](#)
71. Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H., and Adams, P. D. (2012) Toward automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **68**, 352–367 [CrossRef Medline](#)
72. Liebschner, D., Afonine, P. V., Moriarty, N. W., Poon, B. K., Sobolev, O. V., Terwilliger, T. C., and Adams, P. D. (2017) Polder maps: Improving OMIT maps by excluding bulk solvent. *Acta Crystallogr. Sect. D Struct. Biol.* **73**, 148–157 [CrossRef Medline](#)
73. Lobstein, J., Emrich, C. A., Jeans, C., Faulkner, M., Riggs, P., and Berkmen, M. (2012) SHuffle, a novel *Escherichia coli* protein expression strain capable of correctly folding disulfide bonded proteins in its cytoplasm. *Microb. Cell Fact.* **11**, 56 [CrossRef Medline](#)
74. Nagul, E. A., McKelvie, I. D., Worsfold, P., and Kolev, S. D. (2015) The molybdenum blue reaction for the determination of orthophosphate revisited: opening the black box. *Anal. Chim. Acta* **890**, 60–82 [CrossRef Medline](#)