(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
6 October 2016 (06.10.2016)

WIPO | PCT

(10) International Publication Number
WO 2016/156807 A1

(51) **International Patent Classification**:
C07K 14/245 (2006.01)

(21) **International Application Number**:
PCT/GB2016/050814

(22) **International Filing Date**:
23 March 2016 (23.03.2016)

(25) **Filing Language**: English

(26) **Publication Language**: English

(30) **Priority Data**:
1505256.6 27 March 2015 (27.03.2015) GB

(71) **Applicant**: **MEDICAL RESEARCH COUNCIL** [GB/GB]; 2nd Floor, David Philips Building, Polaris House, North Star Avenue, Swindon SN2 1FL (GB).

(72) **Inventors**: **WEBB, Martin R.**; The Francis Crick Institute, Mill Hill Laboratory, The Ridgeway, Mill Hill, London NW7 1AA (GB). **SOLSCHEID, Claudia**; Tentamus QTS Analytical Ltd, Pesticide Residue Specialists, Building 170, Abbott Drive, Kent Science Park, Sittingbourne, Kent ME9 8AZ (GB). **KUNZELMANN, Simone**; The Francis Crick Institute, Mill Hill Laboratory, The Ridgway, Mill Hill, London NW7 1AA (GB).

(74) **Agents**: **RICHARDS, William** et al.; SCRIPT IP LIMITED, Turnpike House, 18 Bridge Street, Frome Somerset BA11 1BB (GB).

(81) **Designated States** *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) **Designated States** *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

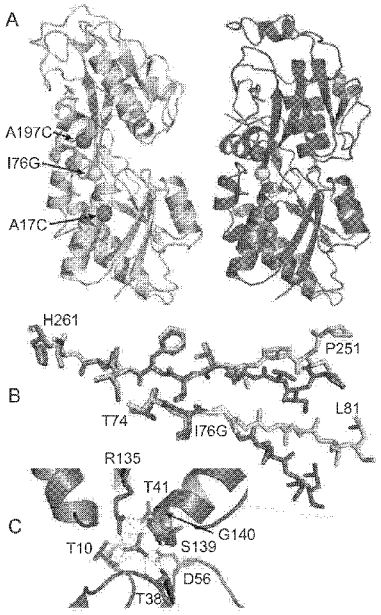
Declarations under Rule 4.17:

— *of inventorship (Rule 4.17(iv))*

[Continued on next page]

(54) **Title**: WIDE RANGE INORGANIC PHOSPHATE BIOSENSOR

FIGURE 1



(57) **Abstract**: An inorganic phosphate-binding molecule is described comprising a mutant phosphate-binding protein (PBP) which undergoes a conformational change from a first conformation to a second conformation upon binding of inorganic phosphate (Pi), said mutant PBP comprising: (i) at least one label; and (ii) at least one mutation that increases flexibility in the hinge region of the PBP as compared to a PBP that does not comprise the at least one mutation in the hinge region; wherein said mutant PBP has a lower affinity for Pi as compared to a reference PBP; and wherein the label emits a detectable signal upon changing from said first conformation to said second conformation.

WO 2016/156807 A1

Published:

— *with international search report (Art. 21(3))*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

WIDE RANGE INORGANIC PHOSPHATE BIOSENSOR

TECHNICAL FIELD

This invention relates to reagents and assays for inorganic phosphate (Pi), particularly the detection and quantification of Pi in biological solutions. More particularly, the present invention relates to a modified phosphate binding protein, and to the use of such a protein in a phosphate assay.

BACKGROUND ART

Pi is a by-product of numerous reactions in the cell, including metabolic reactions like fatty acid metabolism, energy transducing ATPases and in cell signalling, such as GTPases and protein phosphatases. It is desirable to be able to measure the concentration of Pi and the changes in such concentration in biological systems. Phosphate assays, which measure Pi concentration, are useful in a number of diagnostic methods, as well as in research into the functioning of biological systems. Fluorescent reagentless biosensors provide one method of doing this: they are single molecular species that respond to the particular analyte of interest with a change in fluorescence (*1*). This approach circumvents some of the complexities of coupled enzyme assays, for example, in which multiple species are required as additives in the assay mix. Reagentless biosensors require a minimum of recognition element, such as a binding protein and a fluorophore in the same molecule.

The periplasmic phosphate binding protein (PBP) from *Escherichia coli* is a highly specific phosphate scavenger that has been used previously as the recognition element for fluorescent biosensors for P_i (*2, 3*). Fluorophores, covalently bound to surface cysteines, were the reporters for P_i binding, responsive to Pi concentration in the medium. A diethylaminocoumarin-labeled version of the protein (MDCC-PBP) typically has a 7-fold signal change (*2*), whereas a tetraethylrhodamine-labeled biosensor (rho-PBP) results in up to an 18-fold signal increase (*3*). Both of these biosensors bind P_i at rates suggesting diffusion control and with dissociation constants in the nanomolar range. These have been particularly useful to measure P_i production or release in real time in a wide range of enzymatic systems *in vitro,* particularly in transient kinetic assays (*4-8*).

While this fast response and the tight binding of the original PBP-based biosensors enables the study of reactions that release P_i rapidly, these same properties also necessitate the sensor to be present at an excess over P_i, as essentially all P_i binds to the protein in almost all assay conditions. The tight binding also renders it sensitive to even low levels of P_i contamination, which can be present in solutions, on surfaces and particularly occur with such molecules as nucleotides (*9, 10*). The relatively high consumption of stoichiometric biosensors may also be a limiting factor for high-throughput assays and P_i contamination may require extra precautions or purifications of reagents used.

Like the other members of the periplasmic binding protein superfamily, the PBP is formed of two domains linked by a hinge region and the P_i binding site is located in the cleft between these domains (*11*). The protein undergoes a large conformational change upon ligand binding including a bending motion around the hinge (*12*). To produce a biosensor, fluorophores can be attached to cysteines introduced on the surface so that their environment changes when this P_i-induced conformation change occurs (*3, 13*). In MDCC-PBP, a single diethylaminocoumarin was attached to one domain at the top of the cleft and its interaction with the protein changes upon the conformation change. In rhodamine-PBP a tetramethylrhodamine (TMR) was attached on either side of the cleft such that in the apo form the rhodamines can stack, exhibiting very low fluorescence (*14*). The P_i-induced conformation change causes partial disruption of the stacking interaction and the fluorescence increases (*3*).

WO2007/026155 describes that by adding multiple labels to PBP, improvements in detectable changes that occur upon Pi binding can be achieved.

Mutant PBPs that retain phosphate binding have been described in the art, and these mutants can be used together with the present invention if desired. For the *E. coli* protein, for instance: Yao *et al.* (1996) *Biochemistry* 35:2079-85 discloses a mutant PBP with Asp-137 replaced by Asn, Gly or Thr, with little effect on phosphate affinity; Wang e*t* al. (1997) *Nat Struct Biol 4:519* and Wang et al. (1994) *J Biol Chem* 269:25091-4 disclose a Thr-141-Asp mutant, with the aim of changing phosphate affinity; EP-A-0715721, Brune *et al* (1994) *Biochemistry* 33:8262-71, Salins e*t al* (2004) *Sensors and Actuators B* 97:81-9 and Hirshberg *et al* (1998) *Biochemistry* 37:10381-5 disclose an Ala-197-Cys mutant of the *E. coli* PBP; Ledvina et al. (1998) *Protein Sci.* 7:2550-9 discloses an Ala-197-Trp mutant; and Wang e*t* al. (1997) *Nat Struct Biol 4:519* disclose an Asp-56-Asn mutant.

There is a need in the art for biosensors with different properties to expand the range of applications and types of assay. The present invention seeks to address this need.

ASPECTS AND EMBODIMENTS OF THE INVENTION

Mutants of PBP are described that, advantageously, have much lower affinity for Pi, whilst retaining a fluorescence change on Pi binding. The inventors attempted various different strategies to weaken binding but without losing signal on Pi binding. Two approaches were to disrupt either the binding of the phosphate itself or the associated conformational change, namely cleft closure. A third approach was to mutate the hinge between the domains. Variants of each type of approach are described in the Examples with their relative success. Overall for the binding site mutations, large changes in affinity were accompanied by considerable loss of fluorescence change on Pi binding and were not successful. By targeting the cleft closure, in all cases tested the Pi-dependent fluorescence change was almost completely lost and was not successful. The third approach, in which the flexibility of the hinge region was altered, proved highly successful. Flexibility in the hinge region was increased by forming a triple glycine

stretch. This I76G variant advantageously responded to Pi with both a good fluorescence change and a reduction in affinity as compared to a reference PBP. Advantageously, this combination of characteristics, in addition to the other advantageous characteristics described herein, makes this mutant a highly desirable biosensor for the detection of Pi.

Aspect and embodiments of the present invention are set forth in the accompanying claims.

In one aspect, there is provided an inorganic phosphate-binding molecule comprising a mutant phosphate-binding protein (PBP) which undergoes a conformational change from a first conformation to a second conformation upon binding of inorganic phosphate (Pi), said mutant PBP comprising: (i) at least one label; and (ii) at least one mutation that increases flexibility in the hinge region of the PBP as compared to a PBP that does not comprise the at least one mutation in the hinge region; wherein said mutant PBP has a lower affinity for Pi as compared to a reference PBP; wherein said mutant PBP has a lower affinity for Pi as compared to a reference PBP; and wherein the label emits a detectable signal upon changing from said first conformation to said second conformation.

Suitably, the mutation is located at the point of flexion associated with rotational movement during the conformational change of the PBP.

Suitably, the mutant hinge region of the PBP comprises or consists of the amino acid sequence motif: PXXGGGXXX.

Suitably, the mutant hinge region comprises or consists of the amino acid sequence motif: $PX^1X^2GGGX^3VX^4$, wherein X^1 is S, A, T or M, X^2 is A, V or I, X^3 is V, T or I and X^4 is P or L.

Suitably, the mutant hinge region comprises or consists of the amino acid sequence set forth in SEQ ID NOs: 7, 12, 19, 25, 28, 31, 34, 37 or 40.

Suitably, the mutant PBP comprises or consists of an amino acid sequence selected from the group consisting of: SEQ ID NOs: 3, 4, 5, 6, 10, 11, 17, 18, 24, 27, 30, 33, 36 or 39 or an amino acid sequence with at least 30% identity thereto.

Suitably, the mutant PBP is encoded by the nucleotide sequence set forth in any of SEQ ID NOs: 8, 14, 15, 21 or 22 or a polynucleotide sequence with at least 30% identity thereto.

Suitably, the PBP is or is derived from a member of the genus *Xylella*, *Xanthomonas*, *Pasteurella*, *Haemophilus*, *Escherichia*, *Shigella, Salmonella, Erwinia* or *Mesorhizobium*.

Suitably, the PBP is or is derived from *Xylella fastidiosa* 9a5c, *Xanthomonas campestris pv. campestris* str. 8004, *Pasteurella multocida subsp. multocida str.* Pm70, *Haemophilus influenzae* Rd KW20, *Escherichia coli* K-12, *Shigella flexneri serotype 5b, Salmonella schwarzengrund, Erwinia amylovora* or *Mesorhizobium loti*.

Suitably, the protein comprises at least two labels, with at least one label attached to each side of the cleft of the PBP.

Suitably, the protein comprises at least two cysteine substitutions, for attachment of first and second labels.

Suitably, the protein comprises mutations, suitably, cysteine substitutions, at amino acid positions equivalent to amino acid positions selected from the group consisting of: 17, 197, 222, 226, 229, 247, 299 and 302 of SEQ ID NO: 1 or combinations of two or more thereof.

Suitably, the protein comprises pairs of mutations, suitably, one or more pairs of cysteine substitutions, at amino acid positions equivalent to amino acid positions selected from the group consisting of: 17 and 197, or 229 and 302, or 247 and 299, or 222 and 299, or 226 and 299 of SEQ ID NO: 1 or combinations of two or more thereof.

Suitably, the first and second labels comprise fluorophores.

Suitably, the first and second labels include a rhodamine, suitably, a tetramethlyrhodamine.

Suitably, the rhodamine is 6-IATR.

Suitably, the first and second labels shift from a stacked conformation towards an unstaked conformation upon binding of Pi and the signal from the label increases.

Suitably, the mutant PBP of the inorganic phosphate-binding molecule has a lower affinity for Pi as compared to a reference PBP. Suitably, the affinity is reduced by about 50-fold, about 100-fold, about 150-fold or at least about 200-fold as compared to a reference PBP.

Suitably, the mutant PBP of the inorganic phosphate-binding molecule has a (K_d) of between about 5 μM and about 9 μM, suitably between about 6 μM and about 8 μM, suitably about 7 μM.

Suitably, the at least one label of the inorganic phosphate-binding molecule emits a detectable signal upon changing from said first conformation to said second conformation.

Suitably, the mutant PBP of the inorganic phosphate-binding molecule has a lower affinity for Pi as compared to a reference PBP as described herein and the signal from the label(s) can increase around 5-fold, 6-fold, 7-fold, 8-fold or 9-fold or more on binding with Pi, as compared to an inorganic phosphate-binding molecule comprising a reference PBP.

Suitably, the mutant PBP of the inorganic phosphate-binding molecule has a lower affinity for Pi as compared to a reference PBP as described herein and the mutant PBP has a (K_d) of between about 5 μM and about 9 μM, suitably between about 6 μM and about 8 μM, suitably about 7 μM.

Suitably, the at least one label of the inorganic phosphate-binding molecule emits a detectable signal upon changing from said first conformation to said second conformation.

Suitably, the inorganic phosphate-binding molecule has no significant signal response from the addition of Pi analogues, pyrophosphate and nucleotides, as set forth in Table 3.

Suitably, the inorganic phosphate-binding molecule does not respond to sodium arsenate.

Suitably, when measuring the activity of ATPase of PcrA, a DNA helicase, using the inorganic phosphate-binding molecule, the K_m is between about 70 nM and about 100 nM, suitably about 88 ±13 nM, and the k_{cat} is about 30 1 s^{-1}, suitably about 32 ± 1 s^{-1}.

Suitably, the affinity of the mutant PBP of the inorganic phosphate-binding molecule is not affected by pH.

In a further aspect, there is provided a mutant PBP comprising the amino acid sequence selected from the group consisting of: SEQ ID NOs: 7, 12, 19, 25, 28, 31, 34, 37 or 40.

In a further aspect, there is provided a mutant PBP comprising or consisting of the amino acid sequence selected from the group consisting of: SEQ ID NOs: 3, 4, 5, 6, 10, 11, 17, 18, 24, 27, 30, 33, 36 or 39 or an amino acid sequence with at least 30% identity thereto.

In a further aspect, there is provided an inorganic phosphate-binding molecule comprising the mutant PBP as described herein and at least one label, or comprising or consisting of the mutant PBP as described herein and at least one label.

In a further aspect, there is provided an isolated polynucleotide sequence encoding the mutant PBP as described herein or a polynucleotide sequence with at least 30% identity thereto.

In a further aspect, there is provided an isolated polynucleotide sequence comprising or consisting of polynucleotide sequence set forth in any of SEQ ID NOs: 8, 14, 15, 21 or 22 or a polynucleotide sequence with at least 30% identity thereto.

In a further aspect, there is provided a polynucleotide construct comprising the polynucleotide sequence as described herein.

In a further aspect, there is provided a host cell transformed with the polynucleotide construct of as described herein.

In a further aspect, there is provided a method of producing an inorganic phosphate-binding molecule comprising: (a) culturing said host cell as described herein in a suitable culture medium under suitable conditions to produce the mutant PBP; (b) obtaining said produced mutant PBP; and optionally (c) purifying said mutant PBP to provide a purified mutant PBP.

In a further aspect, there is provided a method for monitoring changes in the levels of inorganic phosphate concentration in a sample comprising: (a) contacting said sample with an inorganic phosphate-binding molecule as described herein; and (b) determining changes in conformation of said inorganic phosphate-binding molecule, wherein changes in conformation of said inorganic phosphate-binding molecule indicate changes in the concentration of inorganic phosphate in said sample.

Suitably, the mutant PBP is present at sub-stoichiometric concentrations, suitably between about 0.1 μM and 5 μM.

Suitably, the change is detected in the mixture over a time span of from 1 about 1 ms to about 1 hour.

Suitably, a phosphate mop is absent from the mixture.

In a further aspect, there is provided the use of an inorganic phosphate-binding molecule or a mutant PBP as described herein for measuring Pi in a sample.

In a further aspect, there is provided a method for identifying a mutation in a PBP for use in measuring Pi comprising the steps of: (i) incorporating at least one label into a PBP and incorporating at least one mutation into the hinge region of the PBP; (ii) measuring the affinity of the mutant PBP for Pi; (iii) measuring the signal emitted by the label upon binding of Pi to the

mutant PBP; and (iv) identifying a mutation in the PBP that confers: (a) a lower affinity for Pi as compared to a reference PBP; and (b) the label emits a change in signal upon binding of Pi.

In a further aspect, there is provided a method of incorporating at least one mutation into a PBP comprising the steps of: (i) identifying a mutation in a PBP as described herein; and (ii) incorporating said mutation into a PBP.

In a further aspect, there is provided a PBP obtained or obtainable by the method(s) as described herein.

In a further aspect, there is provided an inorganic phosphate-binding molecule, a mutant PBP, a polynucleotide, a construct, a host cell, a method or a use substantially as described herein and with reference to the accompanying drawings.

BRIEF DESCRIPTION OF DRAWINGS

Figure 1. Structure of *E. coli* PBP, highlighting active site residues and mutations tested. A Complete structure, indicating residues altered in rho-PBPw. Apo PBP is cyan and P_i-bound PBP dark blue (*11, 12*). The mutation site in the hinge region (I76G) is in yellow, the sites for rhodamine attachment (A17C, A197C) are in purple. B. The hinge region from the same structures, showing the movement between apo (cyan) and P_i-bound PBP (dark blue). I76 is shown in yellow. **C**. Binding site residues modified are shown as red sticks, putative bonds between residues and P_i are shown as yellow dashed lines.

Figure 2. Absorbance spectra of rho-PBPw. The spectra of 1 μM rho-PBPw with 6.25 – 400 μM P_i or phosphate mop (0.1 U ml^{-1} purine nucleoside phosphorylase, 200 μM 7-methylguanosine) (-P_i) in 10 mM Pipes pH 7.0, 100 mM NaCl. Inset: the isosbestic point was 526 nm.

Figure 3. Fluorescence spectra of rho-PBPw. Spectra of 1 μM rho-PBPw with either 200 μM P_i (+P_i) or phosphate mop (0.1U ml^{-1} Purine nucleoside phosphorylase, 200 μM 7-methylguanosine) (-P_i) were obtained in 20 mM Pipes pH 7.0, 100 mM NaCl. Excitation spectra were measured by emission at 577 nm. Emission spectra were obtained by excitation at 555 nm.

Figure 4. Titration of rho-PBPw with P_i. 2 μM rho-PBPw were titrated against P_i at 20 °C in 20 mM Pipes pH 7.0, 100 mM NaCl and 5 μM BSA. Aliquots of P_i, adjusted to pH 7.0, were added and fluorescence intensity corrected for the dilution. Data were fitted to a binding hyperbola with a K_d of 28 ± 3 μM and an 8.7-fold fluorescence change.

Figure 5. Comparison of fluorescence intensities of rho-PBPw, following stopped-flow mixing with P_i and end-point titration. For the titration (circles), aliquots of P_i were added to 1 μM rho-PBPw in 20 mM Tris.HCl pH 7.5, 150 mM NaCl and 5 μM BSA at 20 °C in a cuvette.. Data were fitted to a hyperbola to obtain a K_d of 50 ±8 μM. Constant intensities were obtained on a stopped-flow apparatus with the same final concentrations in the mixing chamber (triangles). To compare data using the different optics, the data were adjusted to the same values at 1 mM P_i.

Figure 6. Steady-state ATPase activity of PcrA. Measurements were carried out in 20mM Tris.HCl pH 7.5, 150 mM NaCl and 3 mM $MgCl_2$. Reactions contained 4 nM PcrA, 3 μM rho-PBPw, 100 μM ATP and dT_{35} at various concentrations. The concentration dependence of initial rates was fitted to the Michaelis-Menten equation to give a K_m for dT_{35} of 88 ±13 nM and a k_{cat} of 32 ±1 s^{-1}.

Figure 7 shows the structures of various rhodamines including 5-IATR and 6-IATR that are suitable for use with the invention.

Figure 8 shows the alignment *of* bacterial PBP/PstS proteins from Xylella (*Xylella fastidiosa* 9a5c), Xanthomonas (*Xanthomonas campestris pv. campestris* str. 8004), Pasteurella (*Pasteurella multocida subsp. multocida str.* Pm70), Haemophilus (*Haemophilus influenzae* Rd KW20), E. coli (*Escherichia coli* K-12), Shigella *(Shigella flexneri serotype 5b (strain 8401)),* Salmonella *(Salmonella schwarzengrund* (strain CVM19633)), Erwinia (*Erwinia amylovora* (strain ATCC 49946)) and Mesorhizobium (*Mesorhizobium loti* MAFF303099). All proteins in this alignment are Type II PBPs (Periplasmic Binding Protein Type II superfamily). The hinge region is highlighted in bold and underline.

Figure 9 shows the accession numbers and the percent identity between the *E. coli* sequence shown in Figure 8 against each of the other sequences shown in Figure 8.

ADVANTAGES

The mutants described herein have a good fluorescence change and a reduced affinity for Pi as compared to a reference PBP which makes the mutants highly desirable biosensors for the detection of Pi.

The mutants described herein can be used as biosensors in sub-stoichiometric amounts relative to Pi making them more economical, suitably between about 0.1 μM and 5 μM.

The mutants described herein can be used as biosensors in assays where higher micromolar amounts of Pi are released. This allows reactions to be measured over a longer time span, extending the assay range, for which the biosensor can be used. For example, the assays can be run for up to about 1 hour.

The mutants described herein can be used as biosensors to monitor a larger range of Pi, suitably, up to about 50 μM, about 60 μM, about 70 μM or about 80 μM or more.

The mutants described herein have lowered affinity for Pi meaning that they are less sensitive to contaminating Pi that can be present in many biological solutions and buffers *etc.* Suitably, no phosphate mop is required to remove contaminating Pi in assays when the mutants are used as biosensors. Contaminating Pi can increase the level of background fluorescence and also decrease the amount of biosensor that is free to detect Pi formation.

The affinity of the mutants described herein is little affected by pH which makes the mutants more attractive for use as biosensors.

No significant fluorescence response is obtained from the addition of a selection of Pi analogues, pyrophosphate and nucleotides, as set forth in Table 3, indicating that the mutants when used as biosensors have a high level of specificity for Pi.

DETAILED DESCRIPTION

The technical terms and expressions used within the scope of this application are generally to be given the meaning commonly applied to them in the pertinent art of molecular biology and protein engineering. All of the following term definitions apply to the complete content of this application. The word "comprising" does not exclude other elements or steps, and the indefinite article "a" or "an" does not exclude a plurality. A single step may fulfil the functions of several features recited in the claims. The terms “about”, “essentially” and “approximately” in the context of a given numerate value or range refers to a value or range that is within 20%, within 10%, or within 5%, 4%, 3%, 2% or 1% of the given value or range.

Phosphate binding protein (PBP)

The disclosure utilises a ‘phosphate binding protein’, which is the name commonly given to the primary phosphate receptor of the ABC transport system found in bacteria, also known as the periplasmic phosphate binding receptor. PBPs are also present in eukaryotes. PBPs are part of the active phosphate transfer system and reversibly bind and release Pi. They are members of the protein superfamily of extracellular solute-binding receptors and consist of two domains linked by a hinge region. The phosphate-binding site is located at the interface between the two domains. The proteins typically adopt two conformations: a phosphate-free open form and a phosphate-bound closed form, which interconvert via a hinge-bending mechanism upon phosphate binding. Native PBP is formed after cleavage of a precursor, and PBPs can be lipoproteins. The PBPs are robust to denaturation and bind to Pi specifically and tightly.

PBPs have been described for a number of bacteria and in mammals. A sequence alignment of wild type PBPs from different organisms is shown in Figure 7 of WO2007/026155. Wild type PBPs from *Xylella fastidiosa*, *Xanthomonas campestris*, *Pasteurella multocida*, *Haemophilus influenzae*, *Escherichia coli*, *Shigella flexneri, Salmonella schwarzengrund, Erwinia amylovora* and *Mesorhizobium loti* are examples of PBPs that can be mutated in accordance with the present disclosure because the sequence of the hinge region can be identified. The hinge-bending mechanism is encoded by two distinct regions of the amino acid sequence of PBP which are positioned in the 5’ and the 3’ regions of the sequence. For the purposes of the present disclosure, the 5’ hinge region is of most importance and is targeted for mutagenesis as described herein. Typically, the 3’ hinge region is not mutated although the use of a 3’ hinge region that has been mutated is not excluded.

The primary phosphate receptors of the gram-negative bacterial ABC transport system are Periplasmic Binding Proteins. Periplasmic Binding Proteins form one of the largest protein families in eubacterial and archaebacterial genomes and are considered to be derived from a common ancestor based on similarity of three-dimensional structure, mechanism of ligand binding and gene operon structure. Periplasmic Binding Proteins share common features of three-dimensional structure and patterns of ligand binding despite large length variation and low sequence identity. Periplasmic Binding Proteins consist of two globular domains of mainly α/β type. The ligand is bound in a cleft between the two domains and engulfed by both. A hinge-bending motion between the two domains is accompanied by ligand binding. Suitably, the phosphate receptors used in the present disclosure have these three features.

The genes for the ABC transport system have also been discovered in bacteria without a periplasmic space, such as gram-positive *Mycobacteria*. Primary phosphate receptors from *Mycobacteria* and other Gram-positive bacteria have a tether to anchor them to the membrane and have a similar function to the periplasmic primary phosphate receptors. The function of the similar protein(s) in mammals is unknown.

Periplasmic Binding Proteins are classified as type I or type II based in the topological arrangement of the central β-sheets in their core structure. Suitably, the PBPs of the present disclosure are Type II wherein the sheet topology of both protein domains takes the form $\beta_2\beta_1\beta_3\beta_n\beta_4$ where β_n represents the strand just after the first crossover from the N-terminal domain to the C-terminal domain, and *vice versa*.

The gene nomenclature of PBPs is typically PstS (from 'Pi-Specific Transport'), but the protein has also been referred to as PhoS, nmpA, phoR2, R2pho and phoR2a or phosphate binding protein. In *Mycobacterium tuberculosis* the protein has been referred to as 'protein antigen B' (PAB).

A particularly suitable protein is the *E. coli* PstS protein, because it has been extensively studied. The sequence of wild type *E. coli* PstS is set forth in SEQ ID NO: 2. This 346-mer is a precursor for the mature protein, which is formed by cleavage of the N-terminal 25 residues (underlined). The disclosure suitably uses the mature protein after loss of signal peptide as set forth in SEQ ID NO:1. The expressed protein can include a methionine at position -1, as shown in for example, SEQ ID NO:3.

Other suitable proteins include, but are not limited to, PBP from *Xanthomonas campestris pv. campestris* str. 8004 (the amino acid sequence of the wild type protein is set forth in SEQ ID NO: 9) and PBP from *Erwinia amylovora* (the amino acid sequence of the wild type protein is set forth in SEQ ID NO: 16), PBP from *Xylella fastidiosa* 9a5c (the amino acid sequence of the wild type protein is set forth in SEQ ID NO: 23), PBP from *Pasteurella multocida* (the amino acid sequence of the wild type protein is set forth in SEQ ID NO: 26), PBP from *Haemophilus influenza* (the amino acid sequence of the wild type protein is set forth in SEQ ID NO: 29), PBP from *Shigella flexneri* (the amino acid sequence of the wild type protein is set forth in SEQ ID

NO: 32), PBP from *Salmonella schwarzengrund* (the amino acid sequence of the wild type protein is set forth in SEQ ID NO: 36) and PBP from *Mesorhizobium loti* (the amino acid sequence of the wild type protein is set forth in SEQ ID NO: 38).

Polypeptides, including mutant PBPs, can be provided in purified or isolated form.

In one aspect, there is provided an inorganic phosphate-binding molecule comprising a mutant phosphate-binding protein (PBP) which undergoes a conformational change from a first conformation to a second conformation upon binding of inorganic phosphate (Pi), said mutant PBP comprising: (i) at least one label; and (ii) at least one mutation that increases flexibility in the hinge region of the PBP as compared to a PBP that does not comprise the at least one mutation in the hinge region; wherein said mutant PBP has a lower affinity for Pi; and wherein the label emits a detectable signal upon changing from said first conformation to said second conformation.

Mutant PBP

The disclosure provides mutant PBPs and variants thereof. Mutations in the nucleotide sequences and polypeptides described herein can include man made mutations, synthetic mutations and genetically engineered mutations. Mutations in the nucleotide sequences and polypeptides described herein can be mutations that are obtained or obtainable via a process which includes intervention by man. The mutant PBPs are 'non-naturally occurring' PBPs which means that they are not found in nature and therefore expressly exclude entities that exist in nature. Such non-naturally occurring mutants can be structurally modified, synthesised or manipulated by man.

Suitably, at least one mutation is introduced into the (5') hinge region of the PBP. The objective of the mutation is to increase the flexibility of the hinge region. The hinge region consists of two peptides, approximately parallel and linking the two domains. Cleft closure is achieved by a bending rotation of the hinge. Figure 8 discloses the alignment of PBPs from various strains of bacteria, namely *Xylella fastidiosa*, *Xanthomonas campestris*, *Pasteurella multocida*, *Haemophilus influenzae*, *Escherichia coli*, *Shigella flexneri, Salmonella schwarzengrund, Erwinia amylovora* and *Mesorhizobium loti*. The alignment of the hinge region located at amino acid positions equivalent to amino acid positions 73 to 81 of SEQ ID NO:1 is shown in bold and underlined. The hinge region can be defined by the motif: PXXXGGXXX and is located at amino acid positions equivalent to amino acid positions 73 to 81 of SEQ ID NO:1. The consensus hinge region motif sequence amongst the *Xylella fastidiosa*, *Xanthomonas campestris*, *Pasteurella multocida*, *Haemophilus influenzae*, *Escherichia coli*, *Shigella flexneri, Salmonella schwarzengrund, Erwinia amylovora* and *Mesorhizobium loti* sequences as set forth in Figure 8 is $PX^1X^2X^3GGX^4VX^5$, wherein X^1 is S, A, T or M, X^2 is A, V or I, X^3 is I or M, X^4 is V, T or I and X^5 is P or L. Suitably, X^3 is the site for mutagenesis, suitably X^3 is mutated, suitably, to a glycine to create a triple glycine stretch. In one embodiment, to increase flexibility in the hinge,

the isoleucine at the amino acid position equivalent to amino acid position 76 of SEQ ID NO:1 is mutated, suitably to a glycine resulting in a triple glycine stretch in the hinge.

Amino acid mutations are selected based on the knowledge that certain amino acids and combinations of amino acids can increase the flexibility of the hinge region. For example, replacing an isoleucine with a glycine to create a triple glycine stretch will increase flexibility.

The isoleucine at position 76 of SEQ ID NO:1, or the isoleucine at position 76 of SEQ ID NO: 9, or the isoleucine at position 76 of SEQ ID NO: 16 or the isoleucine at position 76 of SEQ ID NO: 23, or the isoleucine at position 76 of SEQ ID NO: 26, or the isoleucine at position 76 of SEQ ID NO: 29, or the serine at position 76 of SEQ ID NO: 32, or the isoleucine at position 76 of SEQ ID NO: 35, or the methionine at position 76 of SEQ ID NO: 38 is located at the point of flexion associated with the rotational movement during the conformational change. Amino acid sequences of exemplary mutants are set forth in SEQ ID NOs: 3, 4, 5, 6, 10, 11, 17, 18, 24, 27, 30, 33, 36, and 39. Exemplary mutant hinge regions are set forth in SEQ ID NOs: 7, 12, 19, 25, 31, 34, 37 and 40. Nucleotide sequences of exemplary mutants are set forth in SEQ ID NOs: 8, 14, 15, 21 and 22. The corresponding amino acid residues in other PBPs can be identified based on sequence homology *e.g.* using the alignment of Figure 7 of WO2007/026155 or Figure 8 herein.

The use of mutants in which the one or more labels are attached is particularly suitable. This will frequently require suitable amino acid residue(s) (for example, one or more, or two or more Cys residues) to be introduced at a desired position in the structure. By way of example, labels can be attached to amino acid positions selected from the group consisting of: amino acids located at positions equivalent to amino acid positions 17 and 197 of SEQ ID NO:1, or at amino acid positions located at positions equivalent to amino acid positions 229 and 302 of SEQ ID NO:1, or at amino acid positions located at positions equivalent to amino acid positions 247 and 299 of SEQ ID NO:1, or at amino acid positions located at positions equivalent to amino acid positions 222 and 299 of SEQ ID NO:1, or at amino acid positions located at positions equivalent to amino acid positions 226 and 299 of SEQ ID NO:1, or a combination of one or more thereof. Accordingly, labels can be attached at these equivalent positions in any of the amino acid sequences described herein as desired.

The corresponding amino acid residues in other PBPs can be identified based on sequence homology, for example, using the alignment of Figure 7 of WO2007/026155.

An exemplary mutant is set forth in SEQ ID NO: 4 corresponding to the mutant A17C, I76G, A197C. In this mutant, two Cys residues are incorporated at position 17 (A17C) and position 197 (A197C). Another exemplary mutant is set forth in SEQ ID NO: 5 corresponding to the mutant I76G, N226C and S299C. In this mutant, two Cys residues are incorporated at position 226 (N226C) and position 299 (S299C). Another exemplary mutant is set forth in SEQ ID NO: 6 corresponding to the mutant I76G, K229C and E302C. In this mutant, two Cys residues are incorporated at position 229 (K229C) and position 203 (E302C). Another exemplary mutant is

set forth in SEQ ID NO: 11 corresponding to the mutant S17C, I76G and S197C. In this mutant, two Cys residues are incorporated at position 17 (S17C) and position 197 (S197C). Another exemplary mutant is set forth in SEQ ID NO: 18 corresponding to the mutant N17C, I76G, A197C. In this mutant, two Cys residues are incorporated at position 17 (N17C) and position 197 (A197C). As will be understood, any of the mutations described herein for labelling – such as one or more cysteine mutations – can be incorporated into any of the amino acid sequences described herein.

In one aspect, there is disclosed a mutant PBP comprising or consisting of the amino acid sequence selected from the group consisting of: SEQ ID NOs: 3, 4, 5, 6, 10, 11, 17, 18, 24, 27, 30, 33, 36 or 39 or an amino acid sequence with at least 30% identity thereto.

An inorganic phosphate-binding molecule comprising the mutant PBP(s) and at least one label is also disclosed.

In another aspect, there is disclosed a mutant PBP comprising the amino acid sequence selected from the group consisting of: SEQ ID NOs: 7, 12, 19, 25, 28, 31, 34, 37 or 40. An inorganic phosphate-binding molecule comprising this amino acid sequence and at least one label is also disclosed.

Function of mutant PBPs

Increasing the flexibility of the hinge region results in mutant PBPs that have a number of unexpected and advantageous properties, as described herein.

The mutant PBP has a lower affinity for Pi as compared to a reference PBP. The affinity can be reduced by about 50-fold, about 100-fold, about 150-fold or at least about 200-fold.

The mutants can have a dissociation constant in the micromolar range, depending on salt concentration. The mutants can have a (K_d) of between about 5 μM and about 9 μM, suitably between about 6 μM and about 8 μM, suitably about 7 μM.

P_i affinity can be readily determined by the person skilled in the art, for example, by titration of 1 μM of the mutant to be tested with P_i in 10 mM Pipes pH 7.0, as described herein.

The mutant PBP comprises at least one label that emits a detectable signal upon changing from said first conformation to said second conformation. The signal from the label(s) can increase around 5-fold, 6-fold, 7-fold, 8-fold or 9-fold or more on binding with Pi, as compared to a reference PBP.

The mutants therefore have both a large signal change and a desirable reduction in affinity. Suitably, the affinity for Pi is reduced by about 200-fold and the signal from the label upon changing from the first conformation to the second conformation increases by around 9-fold.

It was also advantageously found that no significant fluorescence response was obtained from the addition of a selection of Pi analogues, pyrophosphate and nucleotides, as set forth in Table 3. The mutants also do not respond to sodium arsenate. Accordingly, the mutants described

herein have no significant fluorescence response to Pi analogues, pyrophosphate and nucleotides selected from the group consisting of: ATP, ADP, PPi, sodium arsenate and sodium vanadate, wherein ATP, ADP, PPi are treated with a phosphate mop to reduce contaminating Pi prior to use.

When measuring the activity of ATPase of PcrA, a DNA helicase, the K_m is between about 70 nM and about 100 nM, suitably about 88 ±13 nM, and the k_{cat} is about 30 1 s^{-1}, suitably about 32 ± 1 s^{-1}.

Suitably, the affinity of the mutants is not affected by pH.

Reference PBP

Suitably, comparisons concerning activity and/or function of the mutant PBPs are made herein with respect to a reference PBP that acts as a control. The reference PBP can be a wild-type PBP. Suitably, the reference PBP will correspond to the wild-type version of the hinge region of the PBP without the mutation(s) that are in the hinge region of the mutant version of the PBP under test. The reference PBP may include one or more mutations outside of the hinge region that are also in the mutant PBP under test. For example, the reference PBP may include one or more mutations outside of the hinge region in order to attach the same label(s) that are attached at the same position(s) to the mutant PBP. Thus, for example, when comparing the function or activity of the mutant PBP as set forth in SEQ ID NO: 4 comprising the mutations A17C, I76G and A197C, the reference PBP will comprise the mutations A17C and A197C but not the hinge mutation at I76G. This ensures that comparisons that are made between the hinge mutant PBP and the reference PBP are not influenced by differences in the labels or in the positions of the labels.

Reference Sequence

When particular amino acid residues are referred to herein using numeric addresses, the numbering is taken with reference to the mature wild type amino acid sequence after loss of signal peptide (or to the polynucleotide sequence encoding same) from the respective microorganism as set forth in, for example, SEQ ID NO: 1. This is to be used as is well understood in the art to locate the residue of interest. This is not always a strict counting exercise - attention must be paid to the context. For example, if the protein of interest is of a slightly different length, then location of the correct residue in that sequence may require the sequences to be aligned and the equivalent or corresponding residue picked. This is well within the ambit of the skilled reader.

Mutating has it normal meaning in the art and may refer to the substitution or truncation or deletion of one or more residues, motifs or domains. Mutation may be effected at the polypeptide level, for example, by synthesis of a polypeptide having the mutated sequence, or may be effected at the nucleotide level, for example, by making a polynucleotide encoding the

mutated sequence, which polynucleotide may be subsequently translated to produce the mutated polypeptide. Suitably, the mutations to be used are as set out herein.

Sequence Variation

The polypeptides described herein may comprise sequence changes relative to the wild type sequence in addition to the key mutations described in more detail herein. Specifically the polypeptides described herein may comprise sequence changes at sites which do not significantly compromise the function or operation of the polypeptides described herein. The sequence changes may be at the polypeptide or the nucleotide level.

Polypeptide function may be easily tested using the methods as set out in the examples section, in order to verify that function has not been abrogated or significantly altered. Thus, provided that the polypeptide retains its function which can be easily tested as set out herein, sequence variations may be made in the polypeptide relative to the wild type reference sequence.

Polypeptides include variants produced by introducing any type of alterations (for example, insertions, deletions, or substitutions of amino acids; changes in glycosylation states; changes that affect refolding or isomerizations, three-dimensional structures, or self-association states), which can be deliberately engineered. The variant may have alterations which produce a silent change and result in a functionally equivalent polypeptide. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity and the amphipathic nature of the residues as long as the secondary binding activity of the substance is retained. For example, negatively charged amino acids include aspartic acid and glutamic acid; positively charged amino acids include lysine and arginine; and amino acids with uncharged polar head groups having similar hydrophilicity values include leucine, isoleucine, valine, glycine, alanine, asparagine, glutamine, serine, threonine, phenylalanine, and tyrosine. Conservative substitutions may be made, for example according to the Table below. Amino acids in the same block in the second column and suitably in the same line in the third column may be substituted for each other:

ALIPHATIC	Non-polar	Gly Ala Pro Ile Leu Val
	Polar - uncharged	Cys Ser Thr Met Asn Gly
	Polar - charged	Asp Glu Lys Arg
AROMATIC		His Phe TrpTyr

In considering what mutations, substitutions or other such changes might be made relative to

the wild type sequence, retention of the function of the polypeptide is important. Typically conservative amino acid substitutions would be less likely to adversely affect the function.

When engineering mutants to increase flexibility of the hinge region it is known that certain amino acids and stretches of amino acids will confer increased flexibility.

Pi-binding in the *E. coli* PstS involves amino acids 10, 11, 38, 56, 137, 139 and 140. Mutagenesis should avoid these critical residues.

Mutagenesis should also avoid the introduction of side chains that will interfere with access to the binding cleft.

Mutagenesis should also avoid residues where an attached label will interfere with the binding cleft.

The alignment shown in Figure 7 of WO2007/026155 shows that PBPs from different organisms display both conserved and non-conserved amino acids. The Figure 7 alignment, the Figure 8 alignment herein, and others alignments created using further PBPs, can be used to identify candidate amino acid residues for mutagenesis. Residues which are less conserved between proteins are more likely to tolerate mutation.

Sequence Homology/Identity

Although sequence homology can also be considered in terms of functional similarity (that is, amino acid residues having similar chemical properties/functions), in the context of the present disclosure it is preferred to express homology in terms of sequence identity.

Sequence comparisons can be conducted by eye or, more usually, with the aid of readily available sequence comparison programs. These publicly and commercially available computer programs can calculate percent homology (such as percent identity) between two or more sequences.

Percent identity may be calculated over contiguous sequences, i.e., one sequence is aligned with the other sequence and each amino acid in one sequence is directly compared with the corresponding amino acid in the other sequence, one residue at a time. This is called an "ungapped" alignment. Typically, such ungapped alignments are performed only over a relatively short number of residues (for example less than 50 contiguous amino acids).

Although this is a very simple and consistent method, it fails to take into consideration that, for example in an otherwise identical pair of sequences, one insertion or deletion will cause the following amino acid residues to be put out of alignment, thus potentially resulting in a large reduction in percent homology (percent identity) when a global alignment (an alignment across the whole sequence) is performed. Consequently, most sequence comparison methods are designed to produce optimal alignments that take into consideration possible insertions and deletions without penalising unduly the overall homology (identity) score. This is achieved by inserting "gaps" in the sequence alignment to try to maximise local homology/identity.

These more complex methods assign "gap penalties" to each gap that occurs in the alignment so that, for the same number of identical amino acids, a sequence alignment with as few gaps as possible - reflecting higher relatedness between the two compared sequences - will achieve a higher score than one with many gaps. "Affine gap costs" are typically used that charge a relatively high cost for the existence of a gap and a smaller penalty for each subsequent residue in the gap. This is the most commonly used gap scoring system. High gap penalties will of course produce optimised alignments with fewer gaps. Most alignment programs allow the gap penalties to be modified. However, it is preferred to use the default values when using such software for sequence comparisons. For example when using the GCG Wisconsin Bestfit package (see below) the default gap penalty for amino acid sequences is -12 for a gap and -4 for each extension.

Calculation of maximum percent homology therefore firstly requires the production of an optimal alignment, taking into consideration gap penalties. A suitable computer program for carrying out such an alignment is the GCG Wisconsin Bestfit package (University of Wisconsin, U.S.A; Devereux et al., 1984, *Nucleic Acids Research* 12:387). Examples of other software than can perform sequence comparisons include, but are not limited to, the BLAST package, FASTA (Altschul et al., 1990, *J. Mol. Biol.* 215:403-410) and the GENEWOR S suite of comparison tools.

Although the final percent homology can be measured in terms of identity, the alignment process itself is typically not based on an all-or-nothing pair comparison. Instead, a scaled similarity score matrix is generally used that assigns scores to each pairwise comparison based on chemical similarity or evolutionary distance. An example of such a matrix commonly used is the BLOSUM62 matrix - the default matrix for the BLAST suite of programs. GCG Wisconsin programs generally use either the public default values or a custom symbol comparison table if supplied. It is preferred to use the public default values for the GCG package, or in the case of other software, the default matrix, such as BLOSUM62. Once the software has produced an optimal alignment, it is possible to calculate percent homology, preferably percent sequence identity. The software typically does this as part of the sequence comparison and generates a numerical result.

In the context of the present disclosure, a homologous amino acid sequence is taken to include the one or more mutations that are described herein and has an amino acid sequence which is at least 15, 20, 25, 30, 40, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 90, 91, 92, 93, or 94% identical, suitably at least 95%, 96%, 97%, 98% or 99% identical at the amino acid level. Suitably this identity is assessed over at least 50 or 100, preferably 200, 300, or even more amino acids with the relevant polypeptide sequence(s) disclosed herein, most suitably with the respective wild type sequence as set forth in SEQ ID NOs: 1, 9 or 16, 23, 26, 29, 32, 35 and 38. Suitably, this identity is assessed over the full length

sequence with the relevant polypeptide sequence(s) disclosed herein, most suitably with the respective wild type sequences as described herein.

Suitably, homology should be considered with respect to one or more of those regions of the sequence known to be essential for protein function rather than non-essential neighbouring sequences. This is especially important when considering homologous sequences from distantly related organisms.

Fragments of the polypeptide sequences that retain the activity of the full length sequence are also disclosed.

The same considerations apply to polynucleotide sequences.

Polynucleotides

Polynucleotides can be incorporated into a recombinant replicable vector. The vector may be used to replicate the polynucleotide in a compatible host cell. Thus in a further embodiment, the dislcosure provides a method of making polynucleotides by introducing a polynucleotide into a replicable vector, introducing the vector into a compatible host cell, and growing the host cell under conditions which bring about replication of the vector. The vector may be recovered from the host cell. Suitable host cells include bacteria such as *E. coli*.

Suitably, a polynucleotide in a vector is operably linked to a control sequence that is capable of providing for the expression of the coding sequence by the host cell, i.e. the vector is an expression vector. The term "operably linked" means that the components described are in a relationship permitting them to function in their intended manner. A regulatory sequence "operably linked" to a coding sequence is ligated in such a way that expression of the coding sequence is achieved under condition compatible with the control sequences.

Vectors may be transformed or transfected into a suitable host cell as described to provide for expression of a protein described herein. This process may comprise culturing a host cell transformed with an expression vector as described above under conditions to provide for expression by the vector of a coding sequence encoding the protein, and optionally recovering the expressed protein.

The vectors may be for example, plasmid or virus vectors provided with an origin of replication, optionally a promoter for the expression of the said polynucleotide and optionally a regulator of the promoter. The vectors may contain one or more selectable marker genes, for example an ampicillin resistance gene in the case of a bacterial piasmid. Vectors may be used, for example, to fransfect or transform a host cell.

Control sequences operably linked to sequences encoding the protein described herein include promoters/enhancers and other expression regulation signals. These control sequences may be selected to be compatible with the host cell for which the expression vector is designed to be used in. The term promoter is well-known in the art and encompasses polynucleotide regions

ranging in size and complexity from minimal promoters to promoters including upstream elements and enhancers.

Exemplary polynucleotide sequences are set forth in SEQ ID NOs: 8, 13, 14, 15, 20, 21 and 22.

Polynucleotides can be provided in isolated form.

Protein Expression and Purification

Proteins are typically made by recombinant means, for example as described below and in the examples. However they may also be made by synthetic means using techniques well known†o skilled persons such as solid phase synthesis. Proteins described herein may also be produced as fusion proteins, for example to aid in extraction and purification. Examples of fusion protein partners include glutathione-S-transferase (GST), 6xHis, GAL4 (DNA binding and/or transcriptional activation domains) and β-galactosidase. If may also be convenient to include a proteolytic cleavage site between the fusion protein partner and the protein sequence of interest to allow removal of fusion protein sequences. Suitably, the fusion protein selected should not hinder the function of the polypeptides described herein.

Host cells comprising polynucleotides may be used to express the polypeptides described herein. Host cells may be cultured under suitable conditions which allow expression of the polypeptides described herein. Expression of the polypeptides described herein may be constitutive such that they are continually produced, or inducible, requiring a stimulus to initiate expression. In the case of inducible expression, protein production can be initiated when required by, for example, addition of an inducer substance to the culture medium, for example dexamethasone or IPTG.

Polypeptides described herein can be extracted from host cells by a variety of techniques known in the art, including enzymatic, chemical and/or osmotic lysis and physical disruption.

Labels

The polypeptides described herein carry labels. Particularly suitable labels are those that can exhibit molecular stacking, which will thus include aromatic rings. These include the rhodamine labels. Where two labels “exhibit molecular stacking”, this typically means that their emission and/or excitation spectra are substantially identical to those of a stacked dimer.

Dye stacking is a non-covalent interaction between two chromophores having planar aromatic rings, and it occurs when the rings are separated by a distance that is short enough to allow them to interact *e.g.* to form dimers or trimers. The detectable signal of the stacked molecules is different from that of the unstacked molecules (*e.g.* stacking can cause quenching of signals, and so stacked chromophores will typically show a decreased fluorescence signal intensity relative to the individual unstacked chromophores), and this difference can be used to detect the presence or absence of stacking. Stacked chromophores can have absorption spectra with

(i) a characteristic decrease in the principal absorption peak as chromophore concentration increases and (ii) a characteristic shoulder peak ('band splitting').

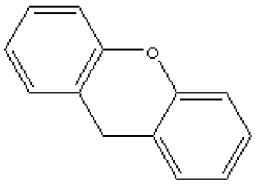
For example, rhodamine chromophores can form dimers at high concentrations in solution. The dimer (λ_{max} ~520 nm) has a different absorbance spectrum from the monomer (λ_{max} ~550 nm), and has little or no fluorescence in comparison with the monomer. This optical difference between free monomer and dimer in solution can be retained when two labels interact when attached to a protein. Two rhodamine chromophores attached to suitable positions in the protein can form dimers, whose interaction is altered when ligand binds to the protein. The difference between the Pi-free and Pi-bound conformations of PBP can be spectroscopically detected.

Labels that can undergo molecular stacking are well known in the art. Stacking can occur between identical chromophores, and can also occur between different chromophores.

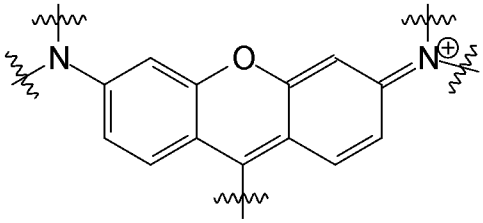
Labels used herein can give various signals. Suitable labels are luminescent labels. Luminescent labels include both fluorescent labels and phosphorescent labels. However, the use of other labels is envisaged. For example, electrochemical labels could be used wherein alteration in the environment of the labels will give rise to a change in redox state. Such a change may be detected using an electrode.

The use of fluorescent labels, which may be excited to fluoresce upon exposure to certain wavelengths of light, is preferred. The fluorescent label can be selected from the group consisting of rhodamines, cyanines, pyrenes and derivatives thereof.

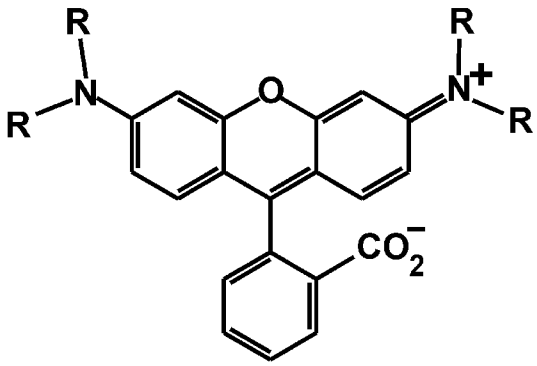
Suitable fluorescent labels are based on a xanthene nucleus, which can readily undergo stacking to form dimers:



Such labels include the rhodamine fluorophores, which include the following core structure:



In addition to the xanthene and the two amino groups, the rhodamine core generally includes a further aromatic ring with a carboxylic substitution, as shown below:



Examples of specific rhodamine fluorophores that can be used are shown in Figure 7. Suitable rhodamine labels are functionalised to give high selectivity for reaction with thiols, such as the haloacetamidotetramethylrhodamine (XATR) molecules, even more suitably iodoacetamidotetramethylrhodamine (IATR) and bromoacetamidotetramethylrhodamine (BATR) molecules. Particularly suitable labels are 5-IATR and 6-IATR, shown in Figure 7.

In certain embodiments, one or more rhodamine labels are preferred as it is capable of producing a large fluorescence response and it has a high photo stability.

Where labels can have different isomers, it is preferred to use a single isomer. Thus, for example, where a rhodamine label is capable of existing as different structural isomers (*e.g.* 5-IATR and 6-IATR), a single isomer in a single PBP is suitably used.

Where two labels are attached to a single PBP, the magnitude of the detectable change seen on Pi binding is suitably greater than the magnitude of the detectable change seen on Pi binding to a PBP with either of the two labels attached without the other being present.

The use of two stackable labels to detect a conformational change in a protein is not restricted to PBPs. For instance, the labels can be used with any periplasmic binding proteins, including those that bind leucine, isoleucine, valine, L-arabinose, glucose, galactose, D-ribose, lactose, purine, histidine, lysine, arginine, ornithine, glutamine, spermidine, putrescine, maltose, D-maltodextrin or sulphate.

Conformational change

On binding to phosphate, PBPs undergo a conformational change. The cleft containing the Pi binding site closes, causing a change in the relative distance and/or orientation of the protein's two globular domains. These alterations in structure, from an initial conformation to a final conformation, are exploited in the present disclosure.

The disclosure can exploit the conformational change by attaching labels such that their separation distance increases or decreases, or such that they rotate relative to each other. Where two labels are attached, the movement can be used to change their ability to exhibit molecular stacking, as described above. Thus the orientation of the first and second labels changes between the initial conformation and the final conformation, and suitably their separation increases which leads to an increase in signal.

Attachment of labels

The PBPs have labels attached to them. The covalent attachment of extrinsic labels to proteins is well known.

Different cysteine residues show different reactivities to labelling reagents, which can be assessed using DTNB (5,5'-dithio-bis(2-nitrobenzoic acid). For PBPs, reactivity can also be affected by the presence of bound Pi. In such cases, a phosphate mop (see below) can be used during labelling, to ensure that protein is in a Pi-free conformation.

Labels can be attached via amines or carboxyl residues on amino acid side chains, but it is preferred to use covalent linkage via thiol groups on a cysteine residue. Where more than one label is attached to a protein, these are suitably attached to separate amino acid residues.

If appropriate, a natural cysteine residue in the PBP can be used for attachment of the label. As *E. coli* PstS does not include any cysteine residues, these must be artificially introduced *e.g.* by site-directed or random mutagenesis.

Where more than one cysteine residue is to be introduced, the same criteria apply. If attached chromophores are to interact, however, the residues must be selected such that (a) they are in proximity to each other, and (b) the conformational change that occurs on Pi-binding affects one or both of the residues to cause a change in position or orientation or electronic environment of a label attached thereto. Amino acids that move apart on Pi-binding are potential sites for label attachment. The residues may be close to each other in the PBP's primary sequence, or may be far away, but the available 3D structures can be used to determine the spatial proximity of chromophores (which will also have known structures) attached to any particular pair of amino acids, both before and after Pi-binding, enabling assessment of likely molecular stacking. Typically, the α-carbons on two residues chosen for label attachment will be separated by between 0.7-2.2 nm (*e.g.* 0.8-1.3 nm) in either the Pi-bound or Pi-free protein, and by a larger distance in the other form.

Suitably, residues chosen for label attachment are surface located. Such residues are more easily accessible for labelling purposes and are less likely to disrupt the tertiary structure of the protein when labelled.

Typical PBPs have two globular domains. Where two residues are chosen these may both be in the same globular domain, or there may be one per globular domain.

For example, PstS crystal structure analysis shows that, as the cleft between the domains closes on phosphate binding, amino acids located on either side of the phosphate-binding cleft get closer in the Pi-bound structure than in the Pi-free structure. However, this movement is also transmitted to structural changes in other parts of the protein. The hinge consists of two extended pieces of the polypeptide, located centrally in the protein. On Pi-binding, the cleft closes on one side of the hinge to produce a rocking motion of the protein domains relative to each other, exposing a new 'cleft' on the opposite side of the protein.

In one embodiment, labels are attached to amino acid residues in a region of the protein remote from the binding site. Suitably, such amino acid residues are not involved in binding Pi (*i.e.* directly coordinate with Pi or indirectly via one other amino acid) or on the surface of the binding cleft. Additionally, or alternatively, labels are attached to amino acid residues on opposite sides of the binding cleft.

Using *E. coli* PstS, eight suitable amino acid residues for substitution by cysteine are, numbered from the N-terminus of the mature PstS as set forth in SEQ ID NO:1 : Ala-17, Ala-197, Glu-222, Asn-226, Lys-229, Glu-247, Ser-299, Glu-302. Where a pair of cysteine residues is introduced, five suitable pairings are: 17 & 197, 229 & 302; 247 & 299; 222 & 299; 226 & 299.

Ala-17 and Ala-197 can both mutated to cysteine residues (*e.g.* SEQ ID NO: 4).

Asn-226 and Ser-299 can both mutated to cysteine residues (*e.g.* SEQ ID NO: 5).

Lys-229 and Glu-302 can both mutated to cysteine residues (*e.g.* SEQ ID NO: 6).

Other possible attachment pairs include Glu-222 & Asp-298, Glu-62 & Lys-235, Asn-226 & Gly-230 and Lys-229 & Ser-299.

The corresponding amino acid residues in other PBPs can be identified based on sequence homology *e.g.* using the alignment of Figure 7 of WO2007/026155 and are described herein.

Fluorophores will rarely be attached to an amino acid directly, but will instead be attached via a linker. The choice of linker can also have an effect on the way the labelled PBP functions, as the size, shape and flexibility of the linker can change the ability of a linker to come into proximity with other groups. Haloacetamide linkers have been found to be useful.

Labels are suitably attached to the PBP in a manner that does not introduce a new chiral centre. Thus the label-protein adduct does not exist in diastereomeric form. This can be achieved by the use of linkers such as the haloacetamides (suitably iodoacetamides).

After attachment of the label, labelled protein will usually be purified to separate it from free label and from any mis-labelled protein. The mis-labelled protein may be unlabelled protein with which label did not react or protein where label has attached in the wrong position (either in place of or in addition to the desired label). During purification of the labelled protein, treatment with a thiol reagent may be included, such as β-mercaptoethanol, dithiothreitol or sodium 2-mercaptoethanesulfonate as this can improve the fluorescence response of the protein.

Where more than one label can be attached, it is preferred to use the protein in homogenous form. A homogenous form, e.g. pure double-labelled species, may be purified (*e.g.* by ion exchange and/or hydrophobic interaction chromatography) to obtain homogenous, double-labelled species. Single and double labelled PBPs can be distinguished by methods such as electrospray mass spectrometry.

Assay methods

The labelled PBPs can be used in assays for detecting Pi in a sample. These assays can be qualitative or quantitative. The disclosure is particularly useful for following the kinetics of

reactions. The assays can be for general biochemical use, or for diagnostic use *e.g.* for diagnosis of disease. For example, measurements of Pi may be used in diagnosis of hyper vitaminosis D, hypoparathyroidism, renal failure, rickets and Fanconi syndrome, as well as for monitoring the causes and treatment of these diseases.

The labelled PBPs may also be useful for the identification and development of drugs against phosphate-associated diseases, such as those in which phosphatase inhibitors might be useful. For example, over-expression of the receptor-like human protein tyrosine 'phosphatase a' (PTPa) results in persistent activation of pp60C-SRC with concomitant cell transformation and tumourigenesis. PTPa may function as an oncogene. Tumours such as human colon carcinoma exhibit an elevated level of pp60C-SRC kinase activity. Inhibitors of PTPa are therefore of use in the treatment of tumours. A high throughput screen assaying for Pi can be used for the identification of suitable lead compounds.

The sample may be from any source, including serum, urine, saliva, sweat, tissue culture, cell extracts, cell lines, food, beverages, pharmaceuticals and environmental (e.g. water). If concentrations of Pi in the sample are high, samples may be diluted as necessary to achieve accurate quantification of Pi levels.

These methods can be performed *in vitro* or *in vivo*, but will typically be *in vitro* assays.

Thus the disclosure provides a method for monitoring changes in the levels of inorganic phosphate concentration in a sample comprising: (a) contacting said sample with an inorganic phosphate-binding molecule as described herein; and (b) determining changes in conformation of said inorganic phosphate-binding molecule, wherein changes in conformation of said inorganic phosphate-binding molecule indicate changes in the concentration of inorganic phosphate in said sample. The disclosure also provides a method for detecting Pi in a sample, comprising the steps of: (i) mixing the sample with a inorganic phosphate-binding molecule as described herein, and (ii) detecting a change in the mixture arising from interaction between the Pi and the PBP. The change detected in step (ii) can be related to the concentration of Pi in the sample.

Identifying mutations

Methods for identifying other mutation(s) in a PBP that afford the advantageous properties of the mutant PBP of the present disclosure are also disclosed. In the method, at least one label and at least one mutation are incorporated into the hinge region of the PBP – suitably the hinge region of SEQ ID NO:1. The affinity of the mutant PBP for Pi is measured using the methods described herein. The signal emitted by the label upon binding of Pi to the mutant PBP is measured using the methods described herein. One or more mutations in the PBP that confer: (a) a lower affinity for Pi as compared to a reference PBP; and (b) the label emits a change in signal upon binding of Pi, as described herein are then selected for further study. If desired, the at least one mutation can be incorporated into a PBP. Suitably, the PBP can then be

incorporated into an inorganic phosphate-binding molecule comprising label. Suitably, the inorganic phosphate-binding molecule can then be used in accordance with the present disclosure.

The disclosure also provides an inorganic phosphate-binding molecule for use in an assay of Pi.

The invention is further described in the Examples below, which are provided to describe the invention in further detail. These examples, which set forth a preferred mode presently contemplated for carrying out the invention, are intended to illustrate and not to limit the invention.

EXAMPLES

Example 1 – Materials & Methods

Materials. Bacterial purine nucleoside phosphorylase was obtained from Sigma and dissolved to 1000 U ml^{-1}. PcrA helicase was purified as described (*15, 16*). Oligonucleotide dT_{35} and nucleotides were from Sigma. 6-iodoacetamidotetramethylrhodamine (6-IATR) was a gift from Dr J. Corrie (NIMR, London) (*17, 18*).

Expression and Purification of Phosphate Binding Protein. PBP mutants were created in pET22b harboring the gene for mature *E. coli* (A17C,A197C)PBP between Nde1 and Xho1 sites in the MCS using a Quikchange site-directed mutagenesis kit (Stratagene) according to manufacturer's instructions. A stop codon was inserted at the end of the *PstS* ORF, so that the encoded His_6-tag was not added to the polypeptide chain. Plasmids were sequenced (GATC Biotech) to confirm the presence of the desired mutation(s). Previously PBP was expressed from the full gene, induced by P_i starvation, and so included the N-terminal signal peptide that was lost in the mature protein (*3, 19*). The pET22 vector described above produced a protein identical to the mature PBP except for an additional N-terminal methionine. It has an advantage of simple induction by IPTG. The amount of purified (A17C, A197C)PBP from this new construct was comparable to the previous method and typically 300 mg from 4 l of *E. coli* culture. An equivalent construct produced similar amounts of (A197C)PBP for MDCC labeling.

Plasmid pET22b carrying the desired mutations within the *PstS* ORF encoding PBP was transformed into BL21 (DE3) and used for preparing rho-PBP variants. An overnight culture was grown in LB medium containing 100 µg ml^{-1} ampicillin at 37 °C with aeration by vigorous shaking. This culture was diluted 50-fold into 500 ml aliquots of fresh medium and grown to an OD_{600} ~0.8 before expression was induced with 500 µM IPTG. After 4 h induction, cells were harvested by 20 min centrifugation at 2500 g and 4 °C. Cells were resuspended in 20 mM Tris.HCl pH 8.0 and stored at -80 °C.

For purification, cells from 500 ml culture were thawed and sonicated 4 x 30 s at 200 W with a 5 s on/off pulse cycle. The lysate was cleared by centrifugation at 142000 g and 4 °C for 45 min. A 5-ml HiTrapQ FF column (GE Healthcare) was equilibrated in 10 mM Tris.HCl pH 8.0, 1 mM dithiothreitol (DTT). The conductivity of the supernatant was adjusted to that of the buffer before

applying it to the column. Protein was eluted in a 50 ml gradient of 0-200 mM NaCl in 10 mM Tris.HCl pH 8.0. Fractions containing PBP were pooled and concentrated in a Vivaspin 20 concentrator (MWCO 10 kDa, GE Healthcare), yielding ~130 mg of PBP per litre culture.

To determine the quarternary structure of (A17C, I76G, A197C)PBP it was applied to a Superdex 200pg 16/60 size exclusion column equilibrated in 10 mM Tris.HCl pH 8, 150 mM NaCl, 1 mM NaN_3. The protein ran as single species corresponding to the size of the monomer.

Labeling. Purified PBP was labeled with 6-IATR as described previously (*3*) in 10 mM Tris.HCl pH 8.0, 100 mM NaCl. The mixture was then slowly diluted to ~3 mM NaCl and concentrated prior to separation of free label and labeled protein. Precipitate was removed from the soluble protein fraction by centrifugation at 16000 g for 10 min at 4 °C and the supernatant filtered through a 0.2 µM polysulfone membrane (PALL Life Sciences). The protein was then applied to a 1 ml MonoQ HR 5/5 column (GE Healthcare) equilibrated in 10 mM Tris.HCl pH 8.0. The protein was eluted with a 30 ml gradient of 0-100 mM NaCl. The elution profile showed three peaks, with the major, second peak eluting at around 20 mM NaCl.

As determined by mass spectrometry and the ratio of absorbance of label (528 nm) and protein (280 nm) this fraction corresponds to the double-labeled PBP. It was concentrated and further analyzed as described below. The variant used for further study was (A17C, I76G, A197C)PBP, labeled with 6-IATR and this is termed rho-PBPw. The concentration of rho-PBP was calculated using an extinction coefficient of 108 $mM^{-1}cm^{-1}$ at 526 nm (*3*).

Absorbance and Fluorescence Measurements. Absorbance spectra were obtained on a JASCO V-550 UV/VIS spectrophotometer. Fluorescence spectra and titrations were obtained on a Cary Eclipse fluorimeter with xenon lamp. Stopped-flow experiments were performed using a HiTech SF61MX apparatus with mercury-xenon lamp and HiTech Kinetic Studio software (TgK Ltd, UK). There was a monochromator and 4 mm slits on the excitation light (548 nm) and a 570 nm cut off filter on the emission. The concentrations given are those in the mixing chamber, unless otherwise stated and data were fitted to theoretical curves using HiTech software and Grafit 7 (*20*).

Steady-state ATPase rate measurements. Rho-PBPw was used to measure the steady-state ATP hydrolysis of the DNA helicase PcrA. Measurements were carried out in 50 mM Tris.HCl pH 7.5, 150 mM NaCl, 3 mM $MgCl_2$ containing 4 nM PcrA, 100 µM ATP, 5µM BSA, 3 µM rho-PBPw and dT_{35} at concentrations ranging from 25 to 300 nM. Reactions were followed using the fluorescence detection setting on a CLARIOstar Microplate Reader (BMG Labtech).

Example 2 - Design Approach

In order to extend the useful range of the existing rhodamine-labeled phosphate biosensor (*3*) for a wider concentration range, the affinity of PBP for P_i was lowered by mutation. This could, in principle, be accomplished by several means, given that the two domains are largely unchanged internally by P_i binding: only their relative position and orientation changes as they

enclose the bound P_i. Amino acids associated with P_i binding could be changed, disrupting binding interactions. Secondly, the associated conformational change could be targeted by modifying amino acids that are in the cleft and modify its closure, but are not involved in P_i binding. Thirdly, the conformation change, and therefore P_i binding, might be affected by modifying the flexibility of the hinge between the two domains. Several mutation sites in these three areas were identified, based on crystal structures (*11, 12, 21*) (Figure 1 and Table 1) and each approach is described in turn. Each PBP variant was labeled on the two cysteines by 6-iodoacetaminotetramethylrhodamine (6-IATR): the fluorescence of this product was used to assess the response to P_i.

Example 3 - Binding site mutations.

The crystal structures of PBP, bound with P_i, show a large number of potential hydrogen bonds between amino acids and the four oxygen atoms of P_i (*11*). Previously PBP with the binding site mutation T141D was used to obtain the crystal structure of the apo protein, as this was shown to have weakened, but highly pH-dependent, binding of P_i (*12*). A study of the PBP from the cyanobacterium *Synechococcus sp.* identified several active site residues affecting phosphate affinity (*22*). Although sharing only 37% sequence identity, it was predicted to have high structural homology with the *E. coli* protein, with all but one active site residue between the species being conserved. In *Synechococcus sp.*, replacement of active-site residues lowered phosphate affinity by up to five orders of magnitude.

Based on the crystal structures and these data, single or double mutations were introduced into the *E. coli* PBP as listed in Table 1 and shown in Figure 1C. With ~1000-fold decrease compared to the original tight binding PBP sensor, S38A (K_d 93 μM) and G140A (65 μM) showed the largest change in P_i affinity. Residue R135, forming contacts with two of the phosphate's oxygens and other residues close to the active site (*11*), was also mutated. While this R135A variant showed the largest fluorescence change between apo and P_i bound PBP, binding remained tight. The T141D variant of MDCC-PBP was also tested, but the fluorescence change and affinity were highly dependent on pH, as shown previously (*5, 12*). This would make application of that variant as a biosensor difficult. In addition, some pairs of mutations were also tested in combination, but failed to give a significant fluorescence response, and so their phosphate affinity could not easily be determined. Overall for these variants with binding site mutations, large changes in affinity were accompanied by considerable loss of fluorescence change on P_i binding.

The second approach was based on targeting the cleft closure by affecting the interactions between the facing, inner surfaces of the two domains (*23-25*). This might be achieved by changing amino acids that have interactions across the closed cleft with residues on the opposite side or by changing residue size. Several residues on the domain surfaces were

mutated (Table 1). In all cases the P_i-dependent fluorescence change was almost completely lost and so the affinities of these variants for P_i were not determined.

The final approach, in which the flexibility of the hinge region was altered, proved most successful. The hinge consists of two peptides, approximately parallel and linking the two domains. Cleft closure is achieved by a bending rotation of the hinge (Figure 1B). A double glycine in the sequence provides flexibility and to increase this, a neighboring isoleucine was replaced with a further glycine resulting in a triple glycine stretch. This residue is located at the point of flexion associated with the rotational movement during the conformational change. The I76G variant of the TMR-labeled (A17C, A197C)PBP responded to P_i with a large fluorescence change and showed the desired reduction in affinity. The combination of these characteristics made this variant, from here on described as rho-PBPw, the candidate for further characterization.

Example 4 - Absorbance and fluorescence properties of rho-PBPw.

On binding of P_i to rho-PBPw, the absorbance spectrum of the rhodamine changed (Figure 2), with the peak at 515 nm decreasing and the peak at 555 nm increasing. This is typical of the two rhodamines shifting from a stacked configuration towards unstacked (*26*) and is similar to that observed in the original tight binding rho-PBP (*3*). The isosbestic point was 526 nm (Figure 2 inset).

The spectra showed a maximum fluorescence increase of around nine-fold on binding P_i (Figure 3), with excitation and emission maxima of 556 nm and 577 nm, respectively. This fluorescence change was used to measure P_i affinity, by titrating P_i into the apo protein (Figure 4). This gave a dissociation constant (K_d) of 28 μM.

Example 5 - Variation in response with solution conditions.

Both the affinity of rho-PBPw for P_i and the fluorescence changed on binding P_i varied with pH and salt concentration (Table 2). The main features are that the amplitude of the signal change decreased with pH, but did not change significantly with salt concentration. Conversely, the dissociation constant did not vary much with pH, but increased with salt concentration. The maximum signal change was observed in 20 mM Pipes pH 6.5, 150 mM NaCl.

Example 6 - Limit of P_i binding kinetics.

The tight-binding variants, rho-PBP and MDCC-PBP, bound P_i in two steps, interpreted as diffusion-controlled binding to the open conformation, followed by the cleft-closing conformation change (*3, 13*). It is likely that the weaker binding by rho-PBPw is accompanied by a combination of slower binding and/or faster dissociation to produce the reduced affinity. To test this, the time course was measured, following rapid mixing of rho-PBPw with P_i, using fluorescence stopped-flow. At all the concentrations of P_i tested, no fluorescence change

occurred during the observed time course. However, the fluorescence intensity, although constant with time, increased with P_i concentration with a similar dependence as the steady-state titrations (Figure 5). This suggests that the transient fluorescence change on binding was complete within the dead time of the stopped-flow instrument (~2 ms) and only the final, Pi-bound fluorescence was observable. This is described further in the Discussion.

Example 7 - Specificity

Wild-type PBP and the original phosphate biosensor show high selectivity for inorganic phosphate (*3, 11*). In order to establish whether lowering the affinity had an effect on the specificity, rho-PBPw was titrated with selection of P_i analogues, pyrophosphate and nucleotides (Table 3). No significant fluorescence response was obtained from addition of the main competing ligands. In contrast to the tight-binding P_i sensor, rho-PBPw did not respond to sodium arsenate over the concentration range tested.

Example 8 - Comparison of fluorescence intensity levels of TMR

The rhodamine fluorescence intensity of rho-PBPw was compared with the tight binding variant and tetramethylrhodamine in solution. The fluorescence quantum yield is difficult to measure accurately when the emission spectrum extends to high wavelength, where corrections are large. To circumvent this, a simpler measure was used by dividing the corrected fluorescence emission intensity at the maximum wavelength by the absorbance at the excitation maximum. This ratio was then used to give an approximate comparison between different species containing the same tetramethylrhodamine fluorophore (Table 4). This shows that this rhodamine is much less fluorescent when attached to the (A17C, I76G, A197C)PBP than in free solution or attached to a small molecule thiol, but does not change greatly on P_i binding.

Example 9 - Steady-state assay of P_i production.

To test rho-PBPw, the ATPase activity of PcrA, a DNA helicase, was measured in a multi-well plate format (Figure 6). PcrA couples ATP hydrolysis to translocation along single-stranded DNA, using one ATP molecule per translocation step of one base (*27*). This ATPase activity is well characterized (*28*) and so provided a good test of the new biosensor. Both the K_m (88 nM) for dT_{35} and k_{cat} (32 s^{-1}) obtained using the new biosensor are similar to previous measurements. Using the new biosensor a larger range of P_i could be monitored (up to 80 μM). In addition, the lowered affinity rendered the sensor less sensitive to contaminating Pi, so that no phosphate mop was required to remove such contaminant (*9*).

Example 10 - Discussion

The affinity of the original P_i biosensor, rho-PBP (*3*), was altered by inserting mutations in strategic positions of the PBP scaffold. Kinetic and other measurements on the fluorescent,

tight-binding PBP suggested a rapid, possibly diffusion controlled initial binding of P_i to the open conformation, followed by a rate-limiting cleft closure (*3, 13*). In principle, the affinity could be altered by changing either step of this mechanism. Three parts of the structure were targeted for mutations that might affect the binding, namely the binding site, inter-domain cleft surfaces and the hinge region between the two domains.

The P_i-binding site is highly specific, being able to discriminate against similar molecules such as sulfate. It contains twelve residues that form hydrogen bonds with the four oxygens of the P_i (*11*). While most of these residues are hydrogen donors, D56 is a hydrogen bond acceptor and plays a crucial role in substrate recognition and discrimination. On the whole for binding site mutations, a large reduction in affinity was accompanied by a large reduction in the signal change on binding P_i. This meant that mutations in the binding site did not give a suitable candidate weak-binding variant. Moreover, there were significant differences in effects with the *E. coli* protein, used here, from those described for *Synechococcus* PBP, predicted to be a close structural homologue (*22*). For example, the T10A mutation, which resulted in a large loss of affinity (5 orders of magnitude) in the *Synechococcus* PBP, had ~500-fold lower affinity in the *E. coli* protein. The difference in effect of active site mutations cannot, therefore, be exclusively due to changes in the active site residues, as these are conserved between the two species, even though overall the proteins share only moderate sequence homology. Active site mutations may have varied secondary effects on the two proteins, due to differences in their amino acid composition. This, together with the different signal element used in the two biosensors, based on the two PBPs from these sources, provides a likely explanation of the differences observed.

Mutations to the cleft surface resulted in small or no signal change upon P_i addition, so the affinity could not be readily determined. Although the predicted structure of the rhodamines in the apo structure suggests little contact with surface amino acids (*14*), this aspect is little understood and mutations in the cleft might affect indirectly the structural changes that disrupt the rhodamine interaction. In other words, mutations that weaken the interaction across the cleft may mean that cleft closure is not favored thermodynamically. Even in the absence of such weakening mutations, the equilibrium constant assigned to the cleft closure conformation change is ~40 (*13*).

The most successful strategy for lowering P_i affinity, while maintaining a fluorescence signal, was alteration of the flexibility of the hinge between two domains making up PBP. Insertion of a third glycine residue at position 76 in the hinge (Figure 1B) lowered P_i affinity up to several hundred-fold, compared to the tight binding sensor, but maintained a reasonably large signal change (Table 1). This extra mutation is on the opposite side of the protein from the rhodamines (Figure 1A), so may be unlikely to affect their fluorescence directly.

Measurements with the tight binding variant, rho-PBP, suggest that the cleft closing conformation change on rho-PBP·P_i occurs subsequent to P_i binding itself and has an equilibrium constant of 40 in favor of closure. Presumably the conformation change is still favored thermodynamically in rho-PBPw, in order to result in net cleft closure and produce the fluorescence change. So the main effect of I76G would be on the binding step itself, either slowing binding or increasing dissociation rate constants.

Stopped-flow kinetic measurements were unable to give precise kinetic information as the binding kinetics were too fast to measure. However, the data suggested that the sum of rate constants for binding and dissociation, >1800 s^{-1}, for the reaction to be almost complete within the 2-ms dead time of the stopped-flow instrument, even at low P_i concentration. This in turn suggests that a large increase in dissociation rate constant occurs from that for the tight-binding rho-PBP. Because the fluorescence change is likely to depend on the cleft-closing conformation change, here may be a limit on what can be achieved in terms of weakening binding by mutation for this rhodamine-protein adduct. An equilibrium constant favoring cleft closure must be maintained to result in a concomitant signal change.

Table 4 shows a measure of the relative fluorescence of various TMR adducts, including 6-IATR free in different solvents, as its adduct with a small molecule, MESNA, and covalently attached to the tight binding, as well as weak binding Pi sensor. By this measure, the ratio of fluorescence emission to absorbance each at its maximum wavelength, TMR in aqueous solution is half as fluorescent as in ethanol. Even in the presence of P_i, TMR has low fluorescence when in rho-PBP, in both the weak and the tight binding variant. There is only a relatively small difference in this fluorescence ratio of rho-PBP in presence and absence of P_i, which is qualitatively in line with the expected mechanism for the fluorescence increase. On binding P_i the TMR labels become at least partially unstacked, leading to an increase in absorbance of the rhodamines. This, in turn, is accompanied by an increase at the fluorescence excitation maximum. The observed fluorescence intensity change on binding P_i is a combination of this greater absorbance and increase in fluorescence per unit absorbance. However, the absorbance spectra are complex with discrete maxima at 516 and 556 nm. In addition the higher wavelength peak with P_i has a shoulder. The predicted structure, based on molecular modeling, suggests that the two rhodamines in the apo PBP are not exactly parallel, so the stacking is imperfect (*14*), while the relative size of the two rhodamine peaks suggests incomplete unstacking when P_i binds. In contrast, monomeric TMR in free solution has the 555 nm peak larger than the 515 nm one. These factors suggest that the stacking/unstacking between the two conformations of rho-PBP is more complex than seen in free solution.

The I76G variant of rho-PBP, rho-PBPw, extends the useable range of the biosensor so P_i can be readily measured in the tens of micromolar range. The fluorescence response to P_i was ionic strength and pH dependent with the maximum response was 9-fold. While the affinity of rho-

PBPw was little affected by pH, there was a decrease with ionic strength (Table 2), which is in accordance with findings that wild-type PBP affinity for P_i decreases with increasing ionic strength (Okoh, Hunter et al. 2006). These findings highlight the requirement to obtain a calibration for each experimental condition in any assay based on fluorescence. Importantly, the additional, weakening mutation did not affect specificity: none of the molecules tested resulting in a significant fluorescence response (Table 3).

Both previous, high-affinity, sensors based on the PBP scaffold have been used in a range of *in vitro* studies (*4, 7, 27-29*). The ability of rho-PBPw to be used to detect P_i in a steady-state assay in real time was demonstrated in a test system. The ATPase activity of the helicase, PcrA, showed comparable results using rho-PBPw to those in the literature (*28*).

The new biosensor has several properties differentiating it from the existing MDCC-PBP and tight-binding rho-PBP. Rhodamines are much more photostable than diethylaminocoumarin, an important property for assays relying on high intensity excitation, such as single molecule and high-throughput studies. High-throughput assays, in particular, also benefit from the extended useable range in two ways. Firstly, the lower affinity allows for the sensor to be used at sub-stoichiometric concentrations, making it more economical, whereas the tight binding biosensors must be present in significant excess over the highest concentration of P_i. Secondly, the higher P_i detection range allows reactions to be measured over a longer time span, extending the assay range, for which the sensor can be used. The tight-binding MDCC-PBP and rho-PBP remain particularly suited for rapid reactions, such as single-turnover measurements of P_i release, where sensitivity and rapid response are essential. The substoichiometric use and weaker affinity mean that rho-PBPw is likely to be little affected by typical levels P_i contamination that can be present in many biological solutions, buffers etc (*10*). In contrast, the tight-binding version binds contaminant P_i stoichiometrically under most assay conditions, so that a significant proportion of the biosensor can be P_i-bound at the start of the assay, increasing the background fluorescence and also decreasing the biosensor free to detect Pi formation.

The development of the weak-binding version of this biosensor, based on the phosphate binding protein, illustrates the variety of mutational approaches possible to achieve this. They contrast in the capability of mutations in different parts of the protein to achieve a combination of weaker affinity, while maintaining a useful signal change. The addition of a third flexible glycine in the hinge region achieved this aim, providing a different set of properties that widen the range of uses for this type of biosensor.

ABBREVIATIONS

6IATR, 6-iodoacetamidotetramethylrhodamine; DTT, dithiothreitol; MEG, 7-methylguanosine; MESNA, 2-mercaptoethanesulphonate; MDCC, *N*-[2-(1-maleimidyl)ethyl]-7-(diethylamino)coumarin-3-carboxamide; MDCC-PBP, (A197C)PBP, labeled with MDCC; PBP,

phosphate binding protein; rho-PBP, (A17C,A197C)PBP, labeled with 6-IATR; rho-PBPw, (A17C,I76G,A197C), labeled with 6-IATR; PNPase, purine nucleoside phosphorylase; TMR, tertramethylrhodamine.

REFERENCES

1. Kunzelmann, S., Solscheid, C., and Webb, M. R. (2014) Fluorescent biosensors: design and application to motility proteins, in *Fluorescent Methods Applied to Molecular Motors, Experientia Supplementum* (Toseland, C. P., and Fili, N., Eds.), Springer, Basel, *105*, 25-47.
2. Brune, M., Hunter, J. L., Corrie, J. E. T., and Webb, M. R. (1994) Direct, real-time measurement of rapid inorganic phosphate release using a novel fluorescent probe and its application to actomyosin subfragment 1 ATPase, *Biochemistry 33*, 8262-8271.
3. Okoh, M. P., Hunter, J. L., Corrie, J. E. T., and Webb, M. R. (2006) A biosensor for inorganic phosphate using a rhodamine-labeled phosphate binding protein, *Biochemistry 45*, 14764-14771.
4. White, H. D., Belknap, B., and Webb, M. R. (1997) Kinetics of nucleoside triphosphate cleavage and phosphate release steps by associated rabbit skeletal actomyosin, measured using a novel fluorescent probe for phosphate, *Biochemistry 36*, 11828-11836.
5. Webb, M. R. (2007) Development of fluorescent biosensors for probing the function of motor proteins, *Mol. Biosyst. 3*, 249 - 256.
6. Gilbert, S. P., Webb, M. R., Brune, M., and Johnson, K. A. (1995) Pathway of processive ATP hydrolysis by kinesin, *Nature 373*, 671-676.
7. He, Z.-H., Chillingworth, R. K., Brune, M., Corrie, J. E. T., Webb, M. R., and Ferenczi, M. A. (1999) The efficiency of contraction in rabbit skeletal muscle fibres, determined from the rate of release of inorganic phosphate, *J. Physiol. 517*, 839-854.
8. Hanes, J. W., and Johnson, K. A. (2008) Real-time measurement of pyrophosphate release kinetics, *Anal. Biochem. 372*, 125-127.
9. Nixon, A. E., Hunter, J. L., Bonifacio, G., Eccleston, J. F., and Webb, M. R. (1998) Purine nucleoside phosphorylase: its use in a spectroscopic assay for inorganic phosphate and to remove inorganic phosphate with the aid of phosphodeoxyribomutase, *Anal. Biochem. 265*, 299-307.
10. Webb, M. R. (2003) A fluorescent sensor to assay inorganic phosphate, in *Kinetic Analysis of Macromolecules: a Practical Approach* (Johnson, K. A., Ed.), Oxford University Press, Oxford, U.K., 131-152.
11. Luecke, H., and Quiocho, F. A. (1990) High specificity of a phosphate transport protein determined by hydrogen bonds, *Nature 347*, 402-406.

12. Ledvina, P. S., Yao, N., Choudhary, A., and Quiocho, F. A. (1996) Negative electrostatic surface potential of protein sites specific for anionic ligands, *Proc. Natl. Acad. Sci. U.S.A. 93*, 6786-6791.

13. Brune, M., Hunter, J. L., Howell, S. A., Martin, S. R., Hazlett, T. L., Corrie, J. E. T., and Webb, M. R. (1998) Mechanism of inorganic phosphate interaction with phosphate binding protein from *Escherichia coli*, *Biochemistry 37*, 10370-10380.

14. Goncalves, M. B., Dreyer, J., Lupieri, P., Barrera-Patino, P. C., Ippoliti, E., Webb, M. R., Corrie, J. E. T., and Carloni, P. (2013) Structural prediction of a rhodamine-based biosensor and comparison with biophysical data, *Phys. Chem. Chem. Phys. 15*, 2177-2183

15. Bird, L. E., Brannigan, J. A., Subramanya, H. S., and Wigley, D. B. (1998) Characterisation of *Bacillus stearothermophilus* PcrA helicase: evidence against an active rolling mechanism, *Nucleic Acids Res. 26*, 2686-2693.

16. Slatter, A. F., Thomas, C. D., and Webb, M. R. (2009) PcrA helicase tightly couples ATP hydrolysis to unwinding double-stranded DNA, modulated by the replication initiator protein, RepD., *Biochemistry 48*, 6326-6334.

17. Corrie, J. E. T., and Craik, J. S. (1994) Synthesis and characterisation of pure isomers of iodoacetamidotetramethylrhodamine, *J. Chem. Soc. Perkin Trans. I*, 2967-2973.

18. Munasinghe, V. R. N., and Corrie, J. E. T. (2006) Optimised synthesis of 6-iodoacetamidotetramethylrhodamine, *ARKIVOC (ii)*, 143-149.

19. Morita, T., Amemura, M., Makino, K., Shinagawa, H., Magota, K., Otsuji, N., and Nakata, A. (1983) Hyperproduction of phosphate-binding protein, phoS, and pre-phoS proteins in *Escherichia coli* carrying a cloned phoS gene, *Eur. J. Biochem. 130*, 427-435.

20. Leatherbarrow, R. J. (2009) *GraFit Version 7*, Erithacus Software Ltd, Horley, U.K.

21. Hirshberg, M., Henrick, K., Haire, L. L., Vasisht, N., Brune, M., Corrie, J. E. T., and Webb, M. R. (1998) The crystal structure of phosphate binding protein labeled with a coumarin fluorophore, a probe for inorganic phosphate, *Biochemistry 37*, 10381-10385.

22. Gu, H., Lalonde, S., Okumoto, S., Looger, L. L., Scharff-Poulsen, A. M., Grossman, A. R., Kossmann, J., Jakobsen, I., and Frommer, W. B. (2006) A novel analytical method for in vivo phosphate tracking, *FEBS Lett. 580*, 5885-5893.

23. Marvin, J. S., and Hellinga, H. W. (2001) Manipulation of ligand binding affinity by exploitation of conformational coupling, *Nature Struct. Biol. 8*, 795-8.

24. Millet, O., Hudson, R. P., and Kay, L. E. (2003) The energetic cost of domain reorientation in maltose-binding protein as studied by NMR and fluorescence spectroscopy, *Proc. Natl. Acad. Sci. U.S.A. 100*, 12700-5.

25. Dwyer, M. A., and Hellinga, H. W. (2004) Periplasmic binding proteins: a versatile superfamily for protein engineering, *Curr. Opin. Struct. Biol. 14*, 495-504.

26. Chambers, R. W., Kajiwara, T., and Kearns, D. R. (1974) Effect of dimer formation on the electronic absorption and emission spectra of ionic dyes, *J. Phys. Chem. 78*, 380–387.

27. Dillingham, M. S., Wigley, D. B., and Webb, M. R. (2000) Demonstration of unidirectional single-stranded DNA translocation by PcrA helicase: measurement of step size and translocation speed, *Biochemistry 39*, 205-212.

28. Toseland, C. P., Martinez-Senac, M. M., Slatter, A. F., and Webb, M. R. (2009) The ATPase cycle of PcrA helicase and its coupling to translocation on DNA, *J. Mol. Biol. 392*, 1020-1032.

29. Ansell, K. H., Jones, H. M., Whalley, D., Hearn, A., Taylor, D. L., Patin, E. C., Chapman, T. M., Osborne, S. A., Wallace, C., Birchall, K., Large, J., Bouloc, N., Smiljanic-Hurley, E., Clough, B., Moon, R. W., Green, J. L., and Holder, A. A. (2014) Biochemical and antiparasitic properties of inhibitors of the Plasmodium falciparum calcium-dependent protein kinase PfCDPK1, *Antimicrob. Agents Chemother. 58*, 6032-43.

Any publication cited or described herein provides relevant information disclosed prior to the filing date of the present application. Statements herein are not to be construed as an admission that the inventors are not entitled to antedate such disclosures. All publications mentioned in the above specification are herein incorporated by reference. Various modifications and variations of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology, protein engineering or related fields are intended to be within the scope of the claims.

SEQUENCES

The numbering of amino acids herein is based on the mature protein wild type PstS protein from (SEQ ID NO:1), so amino acid number 1 is E.

SEQ ID NO:1

Wild type PstS from *Escherichia coli* K-12 – mature protein after loss of signal peptide

EASLTGAGATFPAPVYAKWADTYQKETGNKVNYQGIGSSGGVKQIIANTVDFGASDAPLSDE**K**LAQEGLF
QFPTVIGGVVLAVNIPGLKSGELVLDGKTLGDIYLGKIKKWDDEAIAKLNPGLKLPSQNIAVVRRADGSG
TSFVFTSYLAKVNEEWKNNVGTGSTVKWPIGLGGKGNDGIAAFVQRLPGAIGYVEYAYAKQNNLAYTKLI
SADGKPVSPTEENFANAAKGADWSKTFAQDLTNQKGEDAWPITSTTFILIHKDQKKPEQGTEVLKFFDWA
YKTGAKQANDLDYASLPDSVVEQVRAAWKTNIKDSSGKPLY

SEQ ID NO:2

Wild type PstS from *Escherichia coli* K-12 – protein as "naturally" expressed, before loss of signal peptide (underlined)

MKVMRTTVATVVAATLSMSAFSVFAEASLTGAGATFPAPVYAKWADTYQKETGNKVNYQGIGSSGGVKQI
IANTVDFGASDAPLSDEKLAQEGLFQFPTVIGGVVLAVNIPGLKSGELVLDGKTLGDIYLGKIKKWDDEA
IAKLNPGLKLPSQNIAVVRRADGSGTSFVFTSYLAKVNEEWKNNVGTGSTVKWPIGLGGKGNDGIAAFVQ
RLPGAIGYVEYAYAKQNNLAYTKLISADGKPVSPTEENFANAAKGADWSKTFAQDLTNQKGEDAWPITST
TFILIHKDQKKPEQGTEVLKFFDWAYKTGAKQANDLDYASLPDSVVEQVRAAWKTNIKDSSGKPLY

SEQ ID NO:3

PstS from *Escherichia coli* K-12 comprising an I76G mutation: The expression produces a methionine at position -1 (underlined). Mutations are shown in bold and underlined.

MEASLTGAGATFPAPVYAKWADTYQKETGNKVNYQGIGSSGGVKQIIANTVDFGASDAPLSDEKLAQEGL
FQFPTV**G**GGVVLAVNIPGLKSGELVLDGKTLGDIYLGKIKKWDDEAIAKLNPGLKLPSQNIAVVRRADGS
GTSFVFTSYLAKVNEEWKNNVGTGSTVKWPIGLGGKGNDGIAAFVQRLPGAIGYVEYAYAKQNNLAYTKL
ISADGKPVSPTEENFANAAKGADWSKTFAQDLTNQKGEDAWPITSTTFILIHKDQKKPEQGTEVLKFFDW
AYKTGAKQANDLDYASLPDSVVEQVRAAWKTNIKDSSGKPLY

SEQ ID NO:4

PstS from *Escherichia coli* K-12 comprising A17C, I76G and A197C mutations: The expression produces a methionine at position -1 (underlined). Mutations are shown in bold and underlined.

MEASLTGAGATFPAPVY**C**KWADTYQKETGNKVNYQGIGSSGGVKQIIANTVDFGASDAPLSDEKLAQEGL
FQFPTV**G**GGVVLAVNIPGLKSGELVLDGKTLGDIYLGKIKKWDDEAIAKLNPGLKLPSQNIAVVRRADGS
GTSFVFTSYLAKVNEEWKNNVGTGSTVKWPIGLGGKGNDGIAAFVQRLPGAIGYVEY**C**YAKQNNLAYTKL
ISADGKPVSPTEENFANAAKGADWSKTFAQDLTNQKGEDAWPITSTTFILIHKDQKKPEQGTEVLKFFDW
AYKTGAKQANDLDYASLPDSVVEQVRAAWKTNIKDSSGKPLY

SEQ ID NO: 5

PstS from *Escherichia coli* K-12 comprising I76G, N226C and S299C mutations. The expression produces a methionine at position -1 (underlined). Mutations are shown in bold and underlined.

EASLTGAGATFPAPVYAKWADTYQKETGNKVNYQGIGSSGGVKQIIANTVDFGASDAPLSDEKLAQ
EGLFQFPTV**G**GGVVLAVNIPGLKSGELVLDGKTLGDIYLGKIKKWDDEAIAKLNPGLKLPSQNIAV
VRRADGSGTSFVFTSYLAKVNEEWKNNVGTGSTVKWPIGLGGKGNDGIAAFVQRLPGAIGYVEYAY
AKQNNLAYTKLISADGKPVSPTEENFA**C**AAKGADWSKTFAQDLTNQKGEDAWPITSTTFILIHKDQ
KKPEQGTEVLKFFDWAYKTGAKQANDLDYASLPD**C**VVEQVRAAWKTNIKDSSGKPLY

SEQ ID NO: 6

PstS from *Escherichia coli* K-12 comprising I76G, K229C and E302C mutations. The expression produces a methionine at position -1 (underlined). Mutations are shown in bold and underlined.

EASLTGAGATFPAPVYAKWADTYQKETGNKVNYQGIGSSGGVKQIIANTVDFGASDAPLSDEKLAQ
EGLFQFPTV**G**GGVVLAVNIPGLKSGELVLDGKTLGDIYLGKIKKWDDEAIAKLNPGLKLPSQNIAV
VRRADGSGTSFVFTSYLAKVNEEWKNNVGTGSTVKWPIGLGGKGNDGIAAFVQRLPGAIGYVEYAY
AKQNNLAYTKLISADGKPVSPTEENFANAA**C**GADWSKTFAQDLTNQKGEDAWPITSTTFILIHKDQ
KKPEQGTEVLKFFDWAYKTGAKQANDLDYASLPDSVV**C**QVRAAWKTNIKDSSGKPLY

SEQ ID NO: 7

Mutated I76G hinge region of PstS from *Escherichia coli* K-12. Mutation is shown in bold and underlined.

PTV**G**GGVVL

SEQ ID NO: 8

Nucleotide sequence encoding the amino acid sequence of SEQ ID NO:4. Mutations are shown in bold and underlined.

atggaagcaagcctgacaggtgcaggtgcaaccttccctgcgccggtgtat**tgc**aaatgggctgacactt
accagaaagaaaccggtaataaagttaactaccagggtatcggttcttccggtggcgtaaaacagattat
cgctaataccgttgattttggtgcctctgacgcgccgctgtctgacgaaaaactggctcaggaaggtctg
ttccagttcccgaccgtg**ggt**ggcggcgtggtgctggcggttaacattccagggctgaagtctggcgaac
tggtgctggatggtaaaaccctcggcgacatctacctgggcaaaatcaagaagtgggatgatgaagccat
cgccaaactgaatccgggtctgaaactgccttcacaaaacattgctgtagtacgccgcgcagatggctcc
gggacttccttcgtcttcaccagctacctggcgaaagtgaacgaagagtggaaaaacaacgttggtactg
gctctaccgtaaaatggccgatcggtctgggcggtaaaggtaacgacggtatcgccgcgttcgttcagcg
tctgccgggtgcaattggttatgttgaatat**tgt**tacgcgaagcagaacaacctggcgtacaccaaactg
atctccgctgatggtaaaccggttagtccgaccgaagaaaacttcgctaatgcagcaaaaggtgcagact
ggagcaaaaccttcgctcaggatctgaccaaccagaaaggcgaagatgcatggcctattacctctaccac
gttcattctgatccacaaagatcagaagaaaccagaacaaggcacagaagtgctgaaattcttcgactgg
gcgtacaaaaccggggctaaacaggcgaacgacctggattacgccagcctgccggatagtgtagttgaac
aggttcgcgctgcgtggaagaccaatattaaagacagtagcggtaagccgctgtactaa

Glycine can be encoded by ggn, wherein n = a, c, g or t.

Cysteine can be encoded by tgt or tgc.

SEQ ID NO: 9

Wild type full length amino acid sequence (before loss of predicted signal peptide - underlined) of PBP from *Xanthomonas campestris pv. campestris* str. 8004.

MRRTPLPCNGVLRDVIPIATRSCSVISSIKSRLAVGVLAAALAMGAQAADVTGAGASFIYPVMSKWSADY
NTATKKQVNYQSIGSGGGIAQIKAASVDFGSSDAPLKPEELAAAGLAQFPSVIGGVVPVINVPGIAAGAV
KLDGKTLGDIFLGKVTTWNDAAIVALNPGVKLPDSKITVVHRSDGSGTSFNFTNYLSKVNPDWKSKVGEG
TAVQWPTGIGGKGNEGVAAYVKQIKGGIGYVELSYALQNKMAYTAMKNAAGKFVQPSDETFAAAANSADW
GSSKDFYLVMTNAAGDNAWPITATNFILVQKKPKNPAGLKNTLEFFRWVYSKGDAQAKALDYVPLPDTLV
SQIEAYWAKTLPR

SEQ ID NO: 10

Amino acid sequence of I76G mature PBP mutant from *Xanthomonas campestris pv. campestris* str. 8004. Mutations are shown in bold and underline.

QAADVTGAGASFIYPVMSKWSADYNTATKKQVNYQSIGSGGGIAQIKAASVDFGSSDAPLKPEELAAAGL
AQFPSV**G**GGVVPVINVPGIAAGAVKLDGKTLGDIFLGKVTTWNDAAIVALNPGVKLPDSKITVVHRSDGS
GTSFNFTNYLSKVNPDWKSKVGEGTAVQWPTGIGGKGNEGVAAYVKQIKGGIGYVELSYALQNKMAYTAM
KNAAGKFVQPSDETFAAAANSADWGSSKDFYLVMTNAAGDNAWPITATNFILVQKKPKNPAGLKNTLEFF
RWVYSKGDAQAKALDYVPLPDTLVSQIEAYWAKTLPR

SEQ ID NO: 11

Amino acid sequence of S17C I76G S197C mature PBP mutant from *Xanthomonas campestris pv. campestris* str. 8004. Mutations are shown in bold and underline.

QAADVTGAGASFIYPVM**C**KWSADYNTATKKQVNYQSIGSGGGIAQIKAASVDFGSSDAPLKPEELAAAGL
AQFPSV**G**GGVVPVINVPGIAAGAVKLDGKTLGDIFLGKVTTWNDAAIVALNPGVKLPDSKITVVHRSDGS
GTSFNFTNYLSKVNPDWKSKVGEGTAVQWPTGIGGKGNEGVAAYVKQIKGGIGYVEL**C**YALQNKMAYTAM
KNAAGKFVQPSDETFAAAANSADWGSSKDFYLVMTNAAGDNAWPITATNFILVQKKPKNPAGLKNTLEFF
RWVYSKGDAQAKALDYVPLPDTLVSQIEAYWAKTLPR

SEQ ID NO: 12

Mutated I76G hinge region of PBP from *Xanthomonas campestris pv. campestris* str. 8004. Mutation is shown in bold and underline.

PSV**G**GGVVP

SEQ ID NO: 13

Nucleotide sequence encoding SEQ ID NO:9.

ttgcgcaggacgccattgccctgcaatggcgtcctgcgggacgtcatccctatcgctacaaggagctgtt
ccgtgatctcttccatcaaatcccgtctggctgtcggcgtgctcgctgccgcactggccatgggcgccca
ggctgccgatgtcaccggcgccggcgcatcgttcatttacccggtcatgtcgaagtggtcggccgactac
aacaccgccaccaagaagcaggtcaactatcagtcgatcggctccggtggcggcattgcgcagatcaagg
ccgccagcgtggatttcggttcttccgatgccccgctcaagcccgaagagttggctgccgccggactggc
gcagtttccctcggtgatcggcggcgtggtgccggtgatcaacgtgccgggcatcgctgccggtgcggtc
aagctggacggcaagaccctgggcgatatcttcctgggcaaggtcaccacctggaacgatgcggccatcg
ttgcgctgaacccgggcgtgaagctgcccgacagcaagatcaccgtggtgcatcgctccgacggttcggg
caccagcttcaacttcaccaactacctgtccaaggtcaatccggactggaagagcaaggtcggcgaaggc
acggcggtgcagtggccgaccggcattggcggcaagggcaacgaaggcgtggccgcctacgtgaagcaga
tcaagggcggcatcggctatgtcgagctgtcgtacgcgctgcagaacaagatggcctacaccgcgatgaa
gaatgcggccggcaagttcgtgcagccgtccgatgagaccttcgctgcggccgccaacagtgccgactgg
ggcagcagcaaggacttctacctggtgatgaccaatgccgccggcgacaacgcctggccgatcaccgcca
ccaatttcatcctggtgcagaagaagccgaagaatccggccggcctgaagaacaccctggagttcttccg
ctgggtctacagcaagggcgatgcgcaggccaaggcgctggactacgtgccgctgccggacacgctggtc
agccagatcgaagcctactgggccaagaccctgccccgctaa

SEQ ID NO: 14

Nucleotide sequence encoding SEQ ID NO:10. Mutations are shown in bold and underlined.

gccgcactggccatgggcgcccaggctgccgatgtcaccggcgccggcgcatcgttcatttacccggtca
tgtcgaagtggtcggccgactacaacaccgccaccaagaagcaggtcaactatcagtcgatcggctccgg
tggcggcattgcgcagatcaaggccgccagcgtggatttcggttcttccgatgccccgctcaagcccgaa
gagttggctgccgccggactggcgcagtttccctcggtg**ggc**ggcggcgtggtgccggtgatcaacgtgc
cgggcatcgctgccggtgcggtcaagctggacggcaagaccctgggcgatatcttcctgggcaaggtcac
cacctggaacgatgcggccatcgttgcgctgaacccgggcgtgaagctgcccgacagcaagatcaccgtg
gtgcatcgctccgacggttcgggcaccagcttcaacttcaccaactacctgtccaaggtcaatccggact
ggaagagcaaggtcggcgaaggcacggcggtgcagtggccgaccggcattggcggcaagggcaacgaagg
cgtggccgcctacgtgaagcagatcaagggcggcatcggctatgtcgagctgtcgtacgcgctgcagaac
aagatggcctacaccgcgatgaagaatgcggccggcaagttcgtgcagccgtccgatgagaccttcgctg
cggccgccaacagtgccgactggggcagcagcaaggacttctacctggtgatgaccaatgccgccggcga
caacgcctggccgatcaccgccaccaatttcatcctggtgcagaagaagccgaagaatccggccggcctg
aagaacaccctggagttcttccgctgggtctacagcaagggcgatgcgcaggccaaggcgctggactacg
tgccgctgccggacacgctggtcagccagatcgaagcctactgggccaagaccctgccccgctaa

Glycine can be encoded by ggn, wherein n = a, c, g or t.

SEQ ID NO: 15

Nucleotide sequence encoding SEQ ID NO:11. Mutations are shown in bold and underlined.

Gccgcactggccatgggcgcccaggctgccgatgtcaccggcgccggcgcatcgttcatttacccggtca
tg**tcg**aagtggtcggccgactacaacaccgccaccaagaagcaggtcaactatcagtcgatcggctccgg
tggcggcattgcgcagatcaaggccgccagcgtggatttcggttcttccgatgccccgctcaagcccgaa
gagttggctgccgccggactggcgcagtttccctcggtg**ggc**ggcggcgtggtgccggtgatcaacgtgc
cgggcatcgctgccggtgcggtcaagctggacggcaagaccctgggcgatatcttcctgggcaaggtcac
cacctggaacgatgcggccatcgttgcgctgaacccgggcgtgaagctgcccgacagcaagatcaccgtg
gtgcatcgctccgacggttcgggcaccagcttcaacttcaccaactacctgtccaaggtcaatccggact
ggaagagcaaggtcggcgaaggcacggcggtgcagtggccgaccggcattggcggcaagggcaacgaagg
cgtggccgcctacgtgaagcagatcaagggcggcatcggctatgtcgagctg**tgc**tacgcgctgcagaac
aagatggcctacaccgcgatgaagaatgcggccggcaagttcgtgcagccgtccgatgagaccttcgctg
cggccgccaacagtgccgactggggcagcagcaaggacttctacctggtgatgaccaatgccgccggcga
caacgcctggccgatcaccgccaccaatttcatcctggtgcagaagaagccgaagaatccggccggcctg

aagaacaccctggagttcttccgctgggtctacagcaagggcgatgcgcaggccaaggcgctggactacg
tgccgctgccggacacgctggtcagccagatcgaagcctactgggccaagaccctgccccgctaa

Glycine can be encoded by ggn, wherein n = a, c, g or t.

Cysteine can be encoded by tgt or tgc.

SEQ ID NO: 16

Wild type full length amino acid sequence (before loss of signal peptide - underlined) of PBP from *Erwinia amylovora* (strain ATCC 49946).

MTSMHKTLAQCVALTLSLSAVSALAATNLTGAGGTFPAPVYNKWAAEYHTATGSQVNYQGIGSSGGVKQI
IAKTADFGASDAPMKDEDLAKNGLFQFPTVIGGVVLAVNIPGIKSGELTLDGKTVGDIYLGTVKKWNDPA
ITKLNPGVKLPDANINVVRRADGSGTSFVFTSYLSKVNKDWSSKVGKGSTVNWPVGLGGKGNDGVAAFVQ
RLPGSVGYVEYAYAKQNSLAYTKLVDADGKAIAPSEKSFSDAAKGADWSTSFAQDLTFQKGDNAWPITST
TFILVHKEQANTAKGAAVLQFFDWAYKNGGKTTSALDYASLPAPVVEQIRAAWKSNVKDSSGKALY

SEQ ID NO: 17

Amino acid sequence of I76G mature PBP mutant from *Erwinia amylovora* (strain ATCC 49946). Mutations are shown in bold and underline.

ATNLTGAGGTFPAPVYNKWAAEYHTATGSQVNYQGIGSSGGVKQIIAKTADFGASDAPMKDEDLAKNGLF
QFPTV**G**GGVVLAVNIPGIKSGELTLDGKTVGDIYLGTVKKWNDPAITKLNPGVKLPDANINVVRRADGSG
TSFVFTSYLSKVNKDWSSKVGKGSTVNWPVGLGGKGNDGVAAFVQRLPGSVGYVEYAYAKQNSLAYTKLV
DADGKAIAPSEKSFSDAAKGADWSTSFAQDLTFQKGDNAWPITSTTFILVHKEQANTAKGAAVLQFFDWA
YKNGGKTTSALDYASLPAPVVEQIRAAWKSNVKDSSGKALY

SEQ ID NO: 18

Amino acid sequence of N17C I76G A197C mature PBP mutant from *Erwinia amylovora* (strain ATCC 49946). Mutations are shown in bold and underline.

ATNLTGAGGTFPAPVY**C**KWAAEYHTATGSQVNYQGIGSSGGVKQIIAKTADFGASDAPMKDEDLAKNGLF
QFPTV**G**GGVVLAVNIPGIKSGELTLDGKTVGDIYLGTVKKWNDPAITKLNPGVKLPDANINVVRRADGSG
TSFVFTSYLSKVNKDWSSKVGKGSTVNWPVGLGGKGNDGVAAFVQRLPGSVGYVEY**C**YAKQNSLAYTKLV
DADGKAIAPSEKSFSDAAKGADWSTSFAQDLTFQKGDNAWPITSTTFILVHKEQANTAKGAAVLQFFDWA
YKNGGKTTSALDYASLPAPVVEQIRAAWKSNVKDSSGKALY

SEQ ID NO: 19

Mutated I76G hinge region of PBP from *Erwinia amylovora* (strain ATCC 49946). Mutation is shown in bold and underlined.

PTV**G**GGVVL

SEQ ID NO: 20

Nucleotide sequence encoding SEQ ID NO:16.

atgacttcgatgcacaaaactctggctcaatgcgttgcattaacccttctctgagcgctgtttctgcgc
tggctgcgacgaatctgaccggcgctggcggcacctttccggctccggtttataacaagtgggcggcaga
gtaccacaccgccacgggcagccaggttaactatcagggtatcggctcttccggcggcgtgaagcaaatt
atcgcgaaaaccgccgactttggtgcgtcagatgcaccgatgaaagacgaagatttggctaaaaatggcc
tgttccagttcccgacggtgattggcggggtggtactggcggtcaatattcctggtatcaaatccgggga
gctgacgctggatggtaaaaccgtgggggatatctaccttggtactgtcaagaagtggaacgatccggct
atcaccaaactcaacccgggcgtaaagttgcctgatgccaatatcaacgtggtacgccgcgctgatggtt
ccgggacctcatttgtctttaccagctatctgtctaaagtgaacaaagactggagcagcaaagtcggcaa
aggcagcaccgttaactggccggtgggcctgggcggtaaaggtaatgacggcgttgctgcattcgtacag
cgtttgccaggatctgtcggctatgtcgaatacgcctatgctaaacagaacagcctcgcctacaccaaac
tggtcgatgccgacggcaaagcgattgccccgagcgagaaaagcttcagcgatgcggctaaaggcgcaga
ctggagcacctcgttcgctcaggatctgaccttccagaaaggcgataacgcctggccgatcacctctacc
acctttatcctggttcataaagagcaggccaacaccgcgaagggcgcagccgtgttgcagttctttgact
gggcttataaaaacggtggcaaaaccaccagcgcgctggactatgcatcgctgccagctcccgtcgtgga
gcagatccgcgccgcctggaaaagcaacgtgaaagacagttccggcaaggcgttgtactga

SEQ ID NO: 21

Nucleotide sequence encoding SEQ ID NO:17. Mutations are shown in bold and underlined.

gcgacgaatctgaccggcgctggcggcacctttccggctccggtttataacaagtgggcggcagagtacc
acaccgccacgggcagccaggttaactatcagggtatcggctcttccggcggcgtgaagcaaattatcgc
gaaaaccgccgactttggtgcgtcagatgcaccgatgaaagacgaagatttggctaaaaatggcctgttc
cagttcccgacggtg**ggt**ggcggggtggtactggcggtcaatattcctggtatcaaatccggggagctga
cgctggatggtaaaaccgtgggggatatctaccttggtactgtcaagaagtggaacgatccggctatcac
caaactcaacccgggcgtaaagttgcctgatgccaatatcaacgtggtacgccgcgctgatggttccggg
acctcatttgtctttaccagctatctgtctaaagtgaacaaagactggagcagcaaagtcggcaaaggca
gcaccgttaactggccggtgggcctgggcggtaaaggtaatgacggcgttgctgcattcgtacagcgttt
gccaggatctgtcggctatgtcgaatacgcctatgctaaacagaacagcctcgcctacaccaaactggtc
gatgccgacggcaaagcgattgccccgagcgagaaaagcttcagcgatgcggctaaaggcgcagactgga
gcacctcgttcgctcaggatctgaccttccagaaaggcgataacgcctggccgatcacctctaccacctt
tatcctggttcataaagagcaggccaacaccgcgaagggcgcagccgtgttgcagttctttgactgggct
tataaaaacggtggcaaaaccaccagcgcgctggactatgcatcgctgccagctcccgtcgtggagcaga
tccgcgccgcctggaaaagcaacgtgaaagacagttccggcaaggcgttgtactga

Glycine can be encoded by ggn, wherein n = a, c, g or t.

SEQ ID NO: 22

Nucleotide sequence encoding SEQ ID NO:18. Mutations are shown in bold and underlined.

gcgacgaatctgaccggcgctggcggcacctttccggctccggtttat**tgc**aagtgggcggcagagtacc
acaccgccacgggcagccaggttaactatcagggtatcggctcttccggcggcgtgaagcaaattatcgc
gaaaaccgccgactttggtgcgtcagatgcaccgatgaaagacgaagatttggctaaaaatggcctgttc
cagttcccgacggtg**ggt**ggcggggtggtactggcggtcaatattcctggtatcaaatccggggagctga
cgctggatggtaaaaccgtgggggatatctaccttggtactgtcaagaagtggaacgatccggctatcac
caaactcaacccgggcgtaaagttgcctgatgccaatatcaacgtggtacgccgcgctgatggttccggg
acctcatttgtctttaccagctatctgtctaaagtgaacaaagactggagcagcaaagtcggcaaaggca
gcaccgttaactggccggtgggcctgggcggtaaaggtaatgacggcgttgctgcattcgtacagcgttt
gccaggatctgtcggctatgtcgaatac**tgc**tatgctaaacagaacagcctcgcctacaccaaactggtc
gatgccgacggcaaagcgattgccccgagcgagaaaagcttcagcgatgcggctaaaggcgcagactgga
gcacctcgttcgctcaggatctgaccttccagaaaggcgataacgcctggccgatcacctctaccacctt
tatcctggttcataaagagcaggccaacaccgcgaagggcgcagccgtgttgcagttctttgactgggct
tataaaaacggtggcaaaaccaccagcgcgctggactatgcatcgctgccagctcccgtcgtggagcaga
tccgcgccgcctggaaaagcaacgtgaaagacagttccggcaaggcgttgtactga

Glycine can be encoded by ggn, wherein n = a, c, g or t.

Cysteine can be encoded by tgt or tgc.

SEQ ID NO: 23

Wild type full length amino acid sequence (before loss of predicted signal peptide - underlined) of PBP from *Xylella fastidiosa* 9a5c.

MKVYFAGFTLLCLCTAVTITGCKPSNDNQSTGVSQDGNSTTPPSAEQTKSVKISGAGASFIYPLISQWSA
DYNAATGNKINYQSIGSGGGIAQIKAATIDFGSSDKPLDSSELTQAGLGQFPSAIGGVVPVVNLDNIEPG
KLRLTGPLLADIFLGKISKWNDAAIISANPGLHLPDTKINIVHRSDGSGTTFNFSNYLSKVSAEWKQKVG
EGTSVQWPGGVGGKGNEGVASYVQQIKGSIGYVELAYALQNKMSYTALQNAAGQWVQPSAESFAAAASNA
DWSNAKDFNLVITNATGEAAWPITATNFILMRKQTKDAAQRKATLDFFKWSFENGQKQANELHYVPLPPN
LVKQIEAYWASEFK

SEQ ID NO: 24

Amino acid sequence of mature I76G PBP mutant from *Xylella fastidiosa* 9a5c. Mutation is shown in bold and underline

SVKISGAGASFIYPLISQWSADYNAATGNKINYQSIGSGGGIAQIKAATIDFGSSDKPLDSSELTQAGLG
QFPSA**G**GGVVPVVNLDNIEPGKLRLTGPLLADIFLGKISKWNDAAIISANPGLHLPDTKINIVHRSDGSG
TTFNFSNYLSKVSAEWKQKVGEGTSVQWPGGVGGKGNEGVASYVQQIKGSIGYVELAYALQNKMSYTALQ
NAAGQWVQPSAESFAAAASNADWSNAKDFNLVITNATGEAAWPITATNFILMRKQTKDAAQRKATLDFFK
WSFENGQKQANELHYVPLPPNLVKQIEAYWASEFK

SEQ ID NO: 25

Mutated hinge region of PBP from *Xylella fastidiosa* 9a5c. Mutation is shown in bold and underline.

PSA**G**GGVVP

SEQ ID NO: 26

Wild type full length amino acid sequence (before loss of predicted signal peptide - underlined) of PBP from *Pasteurella multocida subsp. multocida* str. Pm70.

MTKRAMLLAILFSTFALTTQAQTITGAGASFPYPIYAKWASMYEKQTGNKVNYQSIGSGGGQQQIIAKTI
DFGASDDPMKAELLAQHQLLQFPAIIGGTVPVVNLPEITAGQLKLSGEVLADIFLGKIKKWNDPAIAKLN
QGANLPDKAIIVVHRSDGSGTTFGWTNYLSKVSTEWKETVGQGKSVKWPTGQGGKGNEGVAAYVSKIKYS
IGYVEYAYAKQNQLAWASLQNKAGQFVQPSAESFMAAAANAQWESAVGMGVILTNEEGDTSWPVTAASFI
LLHKHAEKPEITKAVFDFFDWAFKQGKVAATELDYVPLPEEVIQKIQAQWKTEVKSSDGKTLWQ

SEQ ID NO: 27

Predicted amino acid sequence of mature I76G PBP mutant from *Pasteurella multocida subsp. multocida* str. Pm70. Mutation is shown in bold and underline.

AQTITGAGASFPYPIYAKWASMYEKQTGNKVNYQSIGSGGGQQQIIAKTIDFGASDDPMKAELLAQHQLL
QFPAI**G**GGTVPVVNLPEITAGQLKLSGEVLADIFLGKIKKWNDPAIAKLNQGANLPDKAIIVVHRSDGSG
TTFGWTNYLSKVSTEWKETVGQGKSVKWPTGQGGKGNEGVAAYVSKIKYSIGYVEYAYAKQNQLAWASLQ
NKAGQFVQPSAESFMAAAANAQWESAVGMGVILTNEEGDTSWPVTAASFILLHKHAEKPEITKAVFDFFD
WAFKQGKVAATELDYVPLPEEVIQKIQAQWKTEVKSSDGKTLWQ

SEQ ID NO: 28

Mutated hinge region of PBP from *Pasteurella multocida subsp. multocida* str. Pm70. Mutation is shown in bold and underline.

PAI**G**GGTVP

SEQ ID NO: 29

Wild type full length amino acid sequence (before loss of predicted signal peptide - underlined) of PBP from *Haemophilus influenzae* Rd KW20.

MKKKSYYVLTLGTLPFAQANSITGAGASFPYPIYAKWASLYEKETGNKVNYQSIGSGGGQQQIIAKTVDF
GASDDPMKSELLQQHQLVQFPAVIGGIVPVVNLPEIKPGKLKLSGKLLAEIFLGKIKKWNDPDLVALNPT
LPLPNKNIIVIHRSDGSGTTFGFTNYLSKISNDWKNQVGEGKSVKWLTGQGGKGNEGVASYVRQMKYSIG
YVEYAYAKQNQLAWISLQNQAGQFVQPSNESFMAAASHAKWHEKAGMGVILTNETGEKSWPITAASFILL
NKYSDNPETTKNVLAFFDWAFSRGQDAATELDYVPIPADVVSTIKSQWKTELKQ

SEQ ID NO: 30

Amino acid sequence of mature I76G PBP mutant from *Haemophilus influenzae* Rd KW20. Mutation is shown in bold and underline.

ANSITGAGASFPYPIYAKWASLYEKETGNKVNYQSIGSGGGQQQIIAKTVDFGASDDPMKSELLQQHQLV
QFPAV**G**GGIVPVVNLPEIKPGKLKLSGKLLAEIFLGKIKKWNDPDLVALNPTLPLPNKNIIVIHRSDGSG
TTFGFTNYLSKISNDWKNQVGEGKSVKWLTGQGGKGNEGVASYVRQMKYSIGYVEYAYAKQNQLAWISLQ
NQAGQFVQPSNESFMAAASHAKWHEKAGMGVILTNETGEKSWPITAASFILLNKYSDNPETTKNVLAFFD
WAFSRGQDAATELDYVPIPADVVSTIKSQWKTELKQ

SEQ ID NO: 31

Mutated hinge region of PBP from *Haemophilus influenzae* Rd KW20. Mutation is shown in bold and underline.

PAV**G**GGIVP

SEQ ID NO: 32

Wild type full length amino acid sequence (before loss of signal peptide - underlined) of PBP from *Shigella flexneri serotype 5b* (strain 8401).

MKVMRTTVATVVAATLSMSAFSVFAEASLTGAGATFPAPVYAKWADTYQKETGNKVNYQGIGSSGGVKQI
IANTVDFGASDAPLSDEKLAQEGLFQFPTVIGGVVLAVNIPGLKSGELVLDGKTLGDIYLGKIKKWDDEA
IAKLNPGLKLPSQNIAVVRRADGSGTSFVFTSYLAKVNEEWKNNVGTGSTVKWPIGLGGKGNDGIAAFVQ
RLPGAIGYVEYAYAKQNNLAYTKLISADGKPVSPTEENFANAAKGADWSKTFAQDLTNQKGEDAWPITST
TFILIHKDQKKPEQGTEVMKFFDWAYKTGAKQANDLDYASLPDSVVEQVRAAWKTNIKDSSGKPLY

SEQ ID NO: 33

Amino acid sequence of mature S76G PBP mutant from *Shigella flexneri serotype 5b (strain 8401).*Mutation is shown in bold and underline.

EASLTGAGATFPAPVYAKWADTYQKETGNKVNYQGIGSSGGVKQIIANTVDFGASDAPLSDEKLAQEGLF
QFPTV**G**GGVVLAVNIPGLKSGELVLDGKTLGDIYLGKIKKWDDEAIAKLNPGLKLPSQNIAVVRRADGSG
TSFVFTSYLAKVNEEWKNNVGTGSTVKWPIGLGGKGNDGIAAFVQRLPGAIGYVEYAYAKQNNLAYTKLI
SADGKPVSPTEENFANAAKGADWSKTFAQDLTNQKGEDAWPITSTTFILIHKDQKKPEQGTEVMKFFDWA
YKTGAKQANDLDYASLPDSVVEQVRAAWKTNIKDSSGKPLY

SEQ ID NO: 34

Mutated hinge region of PBP from *Shigella flexneri serotype 5b* (strain 8401). Mutation is shown in bold and underline.

PTV**G**GGVVL

SEQ ID NO: 35

Wild type full length amino acid sequence (before loss of signal peptide - underlined) of PBP from *Salmonella schwarzengrund* (strain CVM19633)

MKVMRTTVATVVAATLSMSAFSAFAAASLTGAGATFPAPVYAKWADTYQKETGNKVNYQGIGSSGGVKQI
TANTVDFGASDAPLSDEKLNQEGLFQFPTVIGGVVLAVNIPGLKSGELVLDGKTLGDIYLGKIKKWDDEA
ITKLNPGVKLPSQNIAVVRRADGSGTSFVFTSYLAKVNDEWKSKVGAGSTVNWPTGLGGKGNDGIAAFVQ
RLPGAIGYVEYAYAKQNNLAYTKLVSADGKPVSPTEESFSNAAKGADWSKTFAQDLTNQKGDDVWPITST
TFILVHKAQKKPEQGAEVLKFFDWAYKNGAKQANDLDYASLPDNVVEQVRTAWKTSIKDSNGKALY

SEQ ID NO: 36

Amino acid sequence of mature I76G PBP mutant from *Salmonella schwarzengrund* (strain CVM19633). Mutation is shown in bold and underline.

AASLTGAGATFPAPVYAKWADTYQKETGNKVNYQGIGSSGGVKQITANTVDFGASDAPLSDEKLNQEGLF
QFPTV**G**GGVVLAVNIPGLKSGELVLDGKTLGDIYLGKIKKWDDEAITKLNPGVKLPSQNIAVVRRADGSG
TSFVFTSYLAKVNDEWKSKVGAGSTVNWPTGLGGKGNDGIAAFVQRLPGAIGYVEYAYAKQNNLAYTKLV

SADGKPVSPTEESFSNAAKGADWSKTFAQDLTNQKGDDVWPITSTTFILVHKAQKKPEQGAEVLKFFDWA
YKNGAKQANDLDYASLPDNVVEQVRTAWKTSIKDSNGKALY

SEQ ID NO: 37

Mutated hinge region of PBP from *Salmonella schwarzengrund* (strain CVM19633). Mutation is shown in bold and underline.

PTV**G**GGVVL

SEQ ID NO: 38

Wild type full length amino acid sequence (before loss of signal peptide - underlined) of PBP from *Mesorhizobium loti* MAFF303099.

MRHFIRSAAVAIAMATASTFTLSAAIAADLSGAGSTFIYPVFAKWADTYKKDTGVGLNYQSIGSGGGIKQ
VIAKTVTFGATDKPMSDADLEKNGLVQFPMVMGGIVPIVNLTGVKPGELVLDGKTLAQIYLGAITTWDDA
AIKALNPSLTLPSTAIAVVHRSDGSGTTFNFTNYLVKLSPDWKDKVGSDTAVEWPTGVGAKGSEGVANTV
KQTDGGIGYVEYAYAKQNNLSYSKMLNAAGKVVEPSLESFGAAASNADFKGAKNFNVIITNEPGDTTWPI
AASTWVLIHKAPDDAAATGEALKFFAWAYKDGKETAKALDYVSIPDSVVDLIKASWKADIQAGGKPVYAG
E

SEQ ID NO: 39

Amino acid sequence of mature M76G PBP mutant from *Mesorhizobium loti* MAFF303099. Mutation is shown in bold and underline.

AADLSGAGSTFIYPVFAKWADTYKKDTGVGLNYQSIGSGGGIKQVIAKTVTFGATDKPMSDADLEKNGLV
QFPMV**G**GGIVPIVNLTGVKPGELVLDGKTLAQIYLGAITTWDDAAIKALNPSLTLPSTAIAVVHRSDGSG
TTFNFTNYLVKLSPDWKDKVGSDTAVEWPTGVGAKGSEGVANTVKQTDGGIGYVEYAYAKQNNLSYSKML
NAAGKVVEPSLESFGAAASNADFKGAKNFNVIITNEPGDTTWPIAASTWVLIHKAPDDAAATGEALKFFA
WAYKDGKETAKALDYVSIPDSVVDLIKASWKADIQAGGKPVYAGE

SEQ ID NO: 40

Mutated hinge region of PBP from *Mesorhizobium loti* MAFF303099. Mutation is shown in bold and underline.

PMV**G**GGIVP

Table 1. Summary of mutations to weaken binding.

Amino acids are numbered, based on the sequence of the mature PBP. Cysteine mutations for fluorophore attachment are indicated as "label sites", while "mutations" denote residues altered to lower P_i affinity. The location denotes the general area of the protein where the mutation was introduced. The fluorescence change is the ratio between the signal at saturating P_i concentrations to that containing P_i mop. For variants showing a significant fluorescence change, Tthe P_i affinity was determined by titration of 1 μM rho-PBP variant with P_i in 10 mM Pipes pH 7.0. The maximum signal change obtained and the average K_d determined typically from three independent titrations.

label sites	mutations	location	fluorescence ratio (+P_i/-P_i)	K_d [μM]
	T10A	active site	2.2	32 ± 2
	S38A	active site	1.5	68 ± 9
	R135A	active site	4.5	Tight[1]
	G140A	active site	2.8	70 ± 17
A17C, A197C	T10A, G140A	active site	< 1.1	ND
	S139A, T141A	active site	< 1.2	ND
	I76G	hinge	5.9	7.4 ± 0.4
	K200A	cleft	2.1	ND
	G140A, A225Q	cleft	< 1.1	ND
	G140A, A225Y	cleft	1.0	ND
	S39D, S164K, G140A	cleft	1.0	40 ± 4
K229C, E302C	T10A	active site	1.2	ND
	G140A	active site	2.5	220 ± 11

[1]This showed tight binding, with dissociation constant ~1 μM or lower

Table 2. Fluorescence change and affinity at different pH values and salt concentrations. Fluorescence emission spectra and titrations were carried out as described in Figures 3 and 4 . Measurements at pH 6.5 and 7 were in 20 mM Pipes, those at pH 7.5-8.5 in 50 mM Tris.HCl. The fluorescence ratio is between the signal at saturating P_i concentrations to that of the solution prior to P_i addition. The value displayed corresponds to the maximum signal change obtained during all experiments. Measurements were done on at least three experimental replicates and two separate protein preparations.

pH	NaCl [mM]	Fluorescence ratio ($+P_i/-P_i$)	K_d [µM]
6.5	20	7.6	8.3 ± 0.6
	150	9.5	34 ± 3
7.0	0	5.9	7.4 ± 0.3
	20	6.8	15 ± 1
	150	8.8	57 ± 2
7.5	20	5.2	27 ±2
	150	6.1	122 ± 6
8.0	20	2.4	46 ± 3
	150	3.3	213 ±9
8.5	20	1.9	46 ± 5
	150	2.2	224 ± 19

Table 3. Response of rho-PBPw to P_i analogues and to nucleotides.

The species tested were added to 5 μM rho-PBPw in 20 mM Pipes pH 7.0, 100 mM NaCl. Fluorescence emission spectra of the solution ±100 μM of the respective species were measured as described for Figure 4. Saturating P_i was subsequently added to each mixture to verify sensor response. The fluorescence ratio was calculated from the baseline fluorescence of the solution prior to substrate addition and the resultant fluorescence after addition. The highest signal change was observed for GDP but may be explained by contaminating P_i in the nucleotide stocks.

Species	Fluorescence ratio
ATP*	1.1
ADP*	1.1
GDP*	1.5
PP_i*	0.9
Sodium arsenate	1.1
Sodium vanadate	1.0

* These were treated with phosphate mop (1 U ml^{-1} purine nucleoside phosphorylase, 200 μM 7-methylguanosine) (*9*) for 15 min at room temperature to reduce contaminating P_i, prior to addition to rho-PBPw. The phosphorylase activity may be inhibited by the presence of purine nucleotides at high concentration, limiting the effectiveness of this treatment.

Table 4. Relative fluorescence intensities of tetramethylrhodamines

Apart from the 6-IATR sample in ethanol, all measurements were in 10 mM Pipes pH 7.0, 100 mM NaCl using solutions with a maximal absorbance of <0.05. The fluorescence spectra were then corrected for the photomultiplier profile and baseline. The fluorescence intensity at maximum emission was divided by the absorbance at maximum excitation. The relative fluorescence was normalized to that of 6IATR in ethanol.

Species	Relative fluorescence
6IATR in EtOH	1
6IATR in buffer	0.42
6IATR-MESNA	0.51
rho-PBPw, P_i	0.08
rho-PBPw, no P_i*	0.03
rho-PBP, P_i	0.13
rho-PBP, no P_i*	0.07

*Solutions were treated with phosphate mop as described in Figure 3.

CLAIMS

1. An inorganic phosphate-binding molecule comprising a mutant phosphate-binding protein (PBP) which undergoes a conformational change from a first conformation to a second conformation upon binding of inorganic phosphate (Pi), said mutant PBP comprising:

(i) at least one label; and

(ii) at least one mutation that increases flexibility in the hinge region of the PBP as compared to a PBP that does not comprise the at least one mutation in the hinge region;

wherein said mutant PBP has a lower affinity for Pi as compared to a reference PBP; and

wherein the label emits a detectable signal upon changing from said first conformation to said second conformation.

2. The inorganic phosphate-binding molecule of claim 1, wherein the mutation is located at the point of flexion associated with rotational movement during the conformational change of the PBP, suitably, wherein said mutation is located at an amino acid position equivalent to amino acid position 76 of SEQ ID NO:1.

3. The inorganic phosphate-binding molecule of claim 1 or claim 2, wherein the mutant hinge region of the PBP comprises or consists of the amino acid sequence motif: PXXGGGXXX, suitably, wherein the mutant hinge region comprises or consists of the amino acid sequence motif: $PX^1X^2GGGX^3VX^4$, wherein X^1 is S, A, T or M, X^2 is A, V or I, X^3 is V, T or I and X^4 is P or L, suitably, wherein the mutant hinge region comprises or consists of the amino acid sequence set forth in SEQ ID NOs: 7, 12, 19, 25, 28, 31, 34, 37 or 40.

4. The inorganic phosphate-binding molecule of any of the preceding claims, wherein the mutant PBP comprises or consists of an amino acid sequence selected from the group consisting of: SEQ ID NOs: 3, 4, 5, 6, 10, 11, 17, 18, 24, 27, 30, 33, 36 or 39 or an amino acid sequence with at least 30% identity thereto; or wherein the mutant PBP is encoded by the nucleotide sequence set forth in any of SEQ ID NOs: 8, 14, 15, 21 or 22 or a polynucleotide sequence with at least 30% identity thereto.

5. The inorganic phosphate-binding molecule of any of the preceding claims, wherein the PBP is or is derived from a member of the genus *Xylella*, *Xanthomonas*, *Pasteurella*, *Haemophilus*, *Escherichia*, *Shigella, Salmonella, Erwinia* or *Mesorhizobium*, suitably wherein the PBP is or is derived from *Xylella fastidiosa* 9a5c, *Xanthomonas campestris pv. campestris* str. 8004, *Pasteurella multocida subsp. multocida str.* Pm70, *Haemophilus influenzae* Rd KW20, *Escherichia coli* K-12, *Shigella flexneri serotype 5b, Salmonella schwarzengrund, Erwinia amylovora* or *Mesorhizobium loti*.

6. The inorganic phosphate-binding molecule of any of the preceding claims, wherein the protein comprises at least two labels, with at least one label attached to each side of the cleft of the PBP.

7. The inorganic phosphate-binding molecule of any of the preceding claims, wherein the protein comprises at least two cysteine substitutions, for attachment of first and second labels.

8. The inorganic phosphate-binding molecule of claim 6 or claim 7, wherein the first and second labels comprise fluorophores.

9. The inorganic phosphate-binding molecule of any of claims 6 to 8, wherein the first and second labels include a rhodamine, suitably, a tetramethlyrhodamine.

10. The inorganic phosphate-binding molecule of claim 9, wherein the rhodamine is 6-IATR.

11. The inorganic phosphate-binding molecule of any one of claims 6 to 10, wherein the first and second labels shift from a stacked conformation towards an unstaked conformation upon binding of Pi and the signal from the label increases.

12. A mutant PBP comprising the amino acid sequence selected from the group consisting of: SEQ ID NOs: 7, 12, 19, 25, 28, 31, 34, 37 or 40.

13. A mutant PBP comprising or consisting of the amino acid sequence selected from the group consisting of: SEQ ID NOs: 3, 4, 5, 6, 10, 11, 17, 18, 24, 27, 30, 33, 36 or 39 or an amino acid sequence with at least 30% identity thereto.

14. An inorganic phosphate-binding molecule comprising the mutant PBP of claim 12 and at least one label, or comprising or consisting of the mutant PBP of claim 13 and at least one label.

15. An isolated polynucleotide sequence encoding the mutant PBP according to claim 12 or claim 13 or a polynucleotide sequence with at least 30% identity thereto.

16. An isolated polynucleotide sequence comprising or consisting of polynucleotide sequence set forth in any of SEQ ID NOs: 8, 14, 15, 21 or 22 or a polynucleotide sequence with at least 30% identity thereto.

17. A polynucleotide construct comprising the polynucleotide sequence of claim 15 or claim 16.

18. A host cell transformed with the polynucleotide construct of claim 17.

19. A method of producing an inorganic phosphate-binding molecule comprising:

(a) culturing said host cell according to claim 18 in a suitable culture medium under suitable conditions to produce the mutant PBP;

(b) obtaining said produced mutant PBP; and optionally

(c) purifying said mutant PBP to provide a purified mutant PBP.

20. A method for monitoring changes in the levels of inorganic phosphate concentration in a sample comprising:

(a) contacting said sample with an inorganic phosphate-binding molecule according to any of claims 1 to 11 and 14; and

(b) determining changes in conformation of said inorganic phosphate-binding molecule,

wherein changes in conformation of said inorganic phosphate-binding molecule indicate changes in the concentration of inorganic phosphate in said sample.

21. The method according to claim 20, wherein the mutant PBP is present at sub-stoichiometric concentrations, suitably between about 0.1 µM and 5 µM.

22. The method according to claim 20 or claim 21, wherein the change is detected in the mixture over a time span of from 1 about 1 ms to about 1 hour.

23.The method according to any of claims 20 to 22, wherein a phosphate mop is absent from the mixture.

24. Use of an inorganic phosphate-binding molecule according to any of claims 1 to 11 and 14 or the mutant PBP according to claims 12 or 13 for measuring Pi in a sample.

25. A method for identifying a mutation in a PBP for use in measuring Pi comprising the steps of:

(i) incorporating at least one label into a PBP and incorporating at least one mutation into the hinge region of the PBP;

(ii) measuring the affinity of the mutant PBP for Pi;

(iii) measuring the signal emitted by the label upon binding of Pi to the mutant PBP; and

(iv) identifying a mutation in the PBP that confers: (a) a lower affinity for Pi as compared to a reference PBP; and (b) the label emits a change in signal upon binding of Pi.

26. A method of incorporating at least one mutation into a PBP comprising the steps of:

(i) identifying a mutation in a PBP according to the method of claim 25; and

(ii) incorporating said mutation into a PBP.

27. A PBP obtained or obtainable by the method of claim 26.

28. An inorganic phosphate-binding molecule, a mutant PBP, a polynucleotide, a construct, a host cell, a method or a use substantially as described herein and with reference to the accompanying drawings.

FIGURE 1

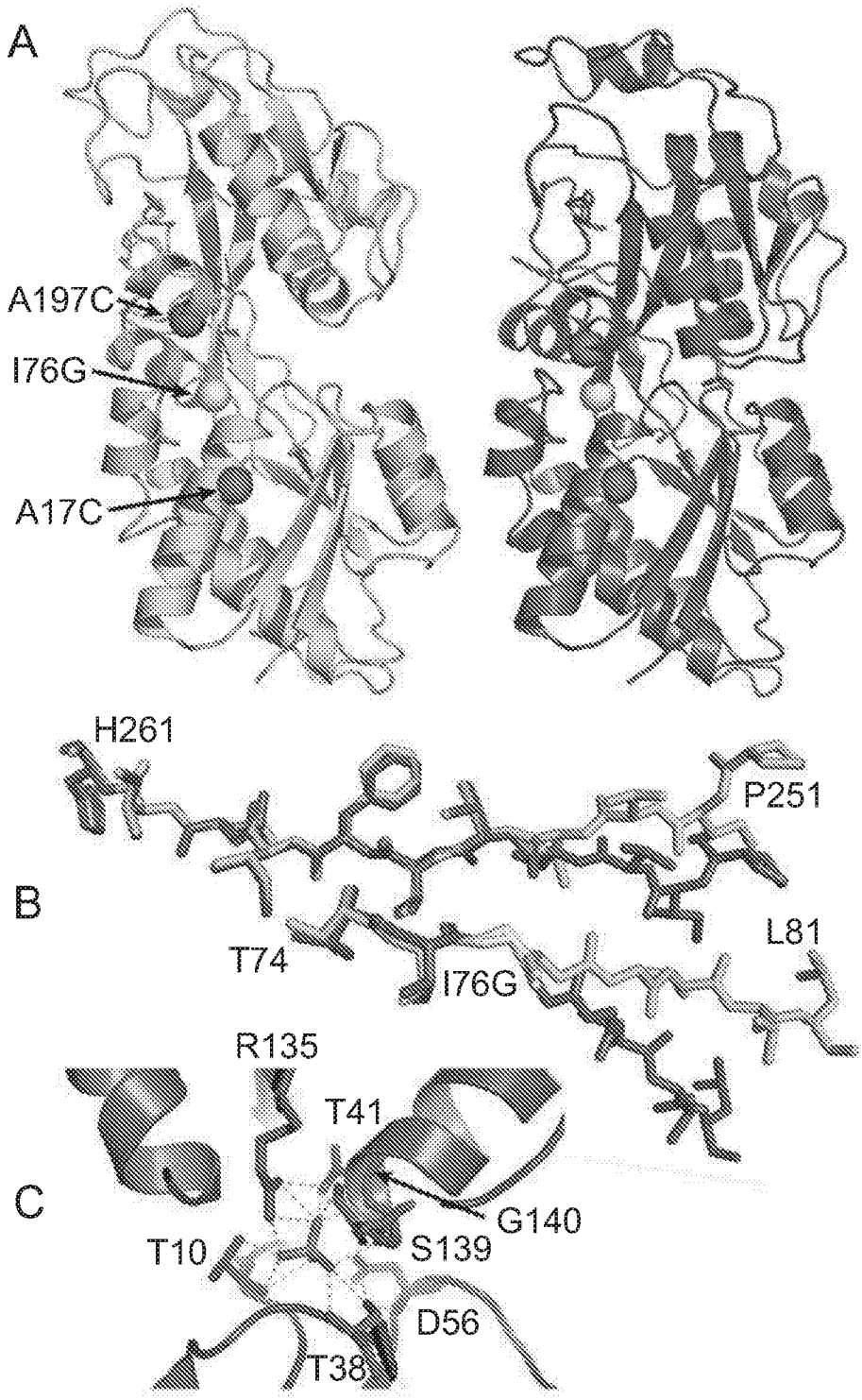

A
A197C
I76G
A17C
B
H261
P251
T74
I76G
L81
C
R135
T41
G140
T10
S139
D56
T38

FIGURE 2

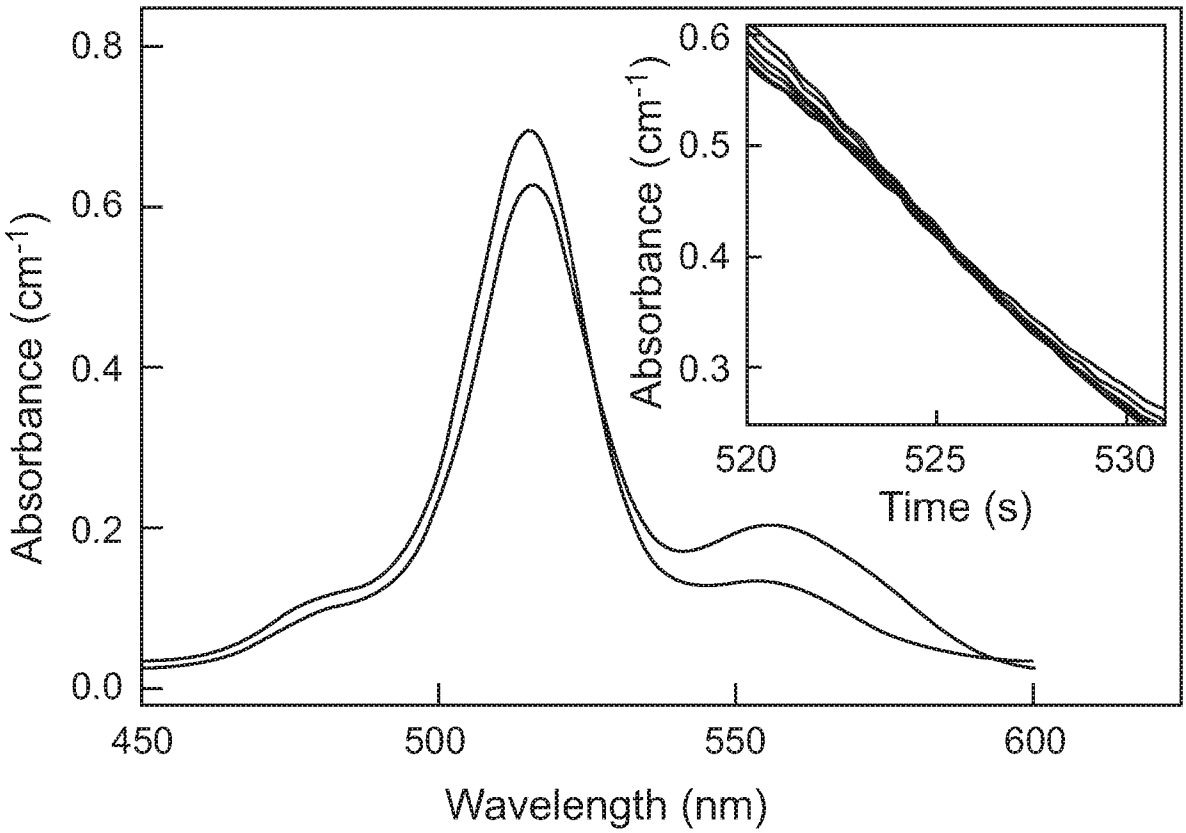

0.8
0.6
0.4
0.2
0.0
Absorbance (cm-1)
450
500
550
600
Wavelength (nm)
0.6
0.5
0.4
0.3
Absorbance (cm-1)
520
525
530
Time (s)

FIGURE 3

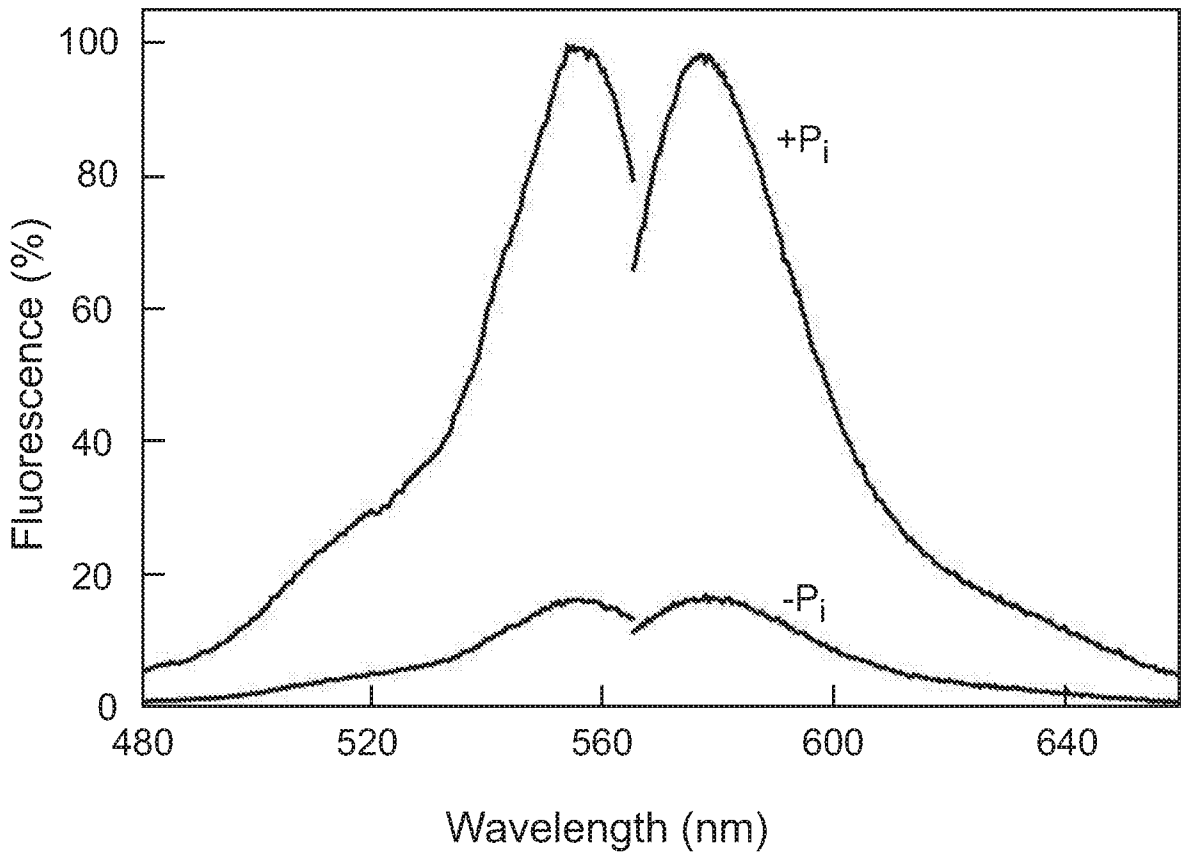

100
80
60
40
20
0
Fluorescence (%)
+Pi
-Pi
480
520
560
600
640
Wavelength (nm)

FIGURE 4

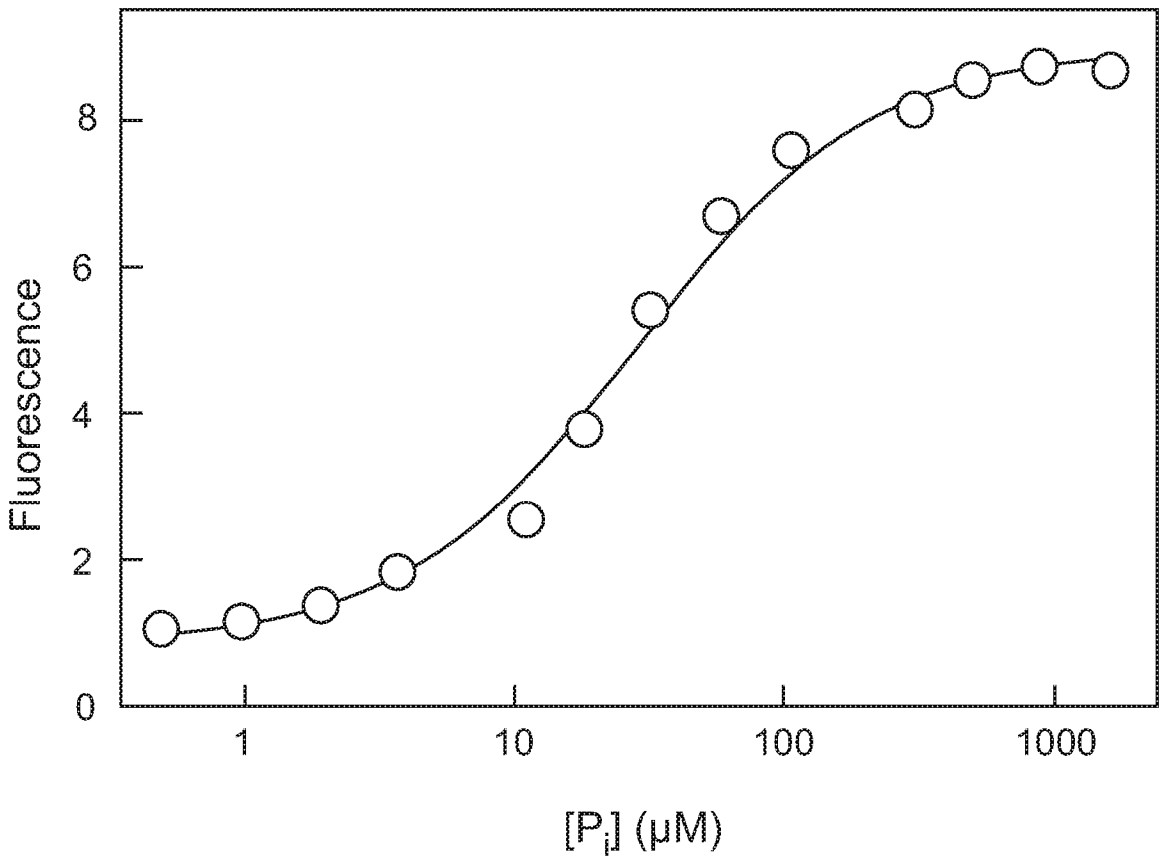


FIGURE 5

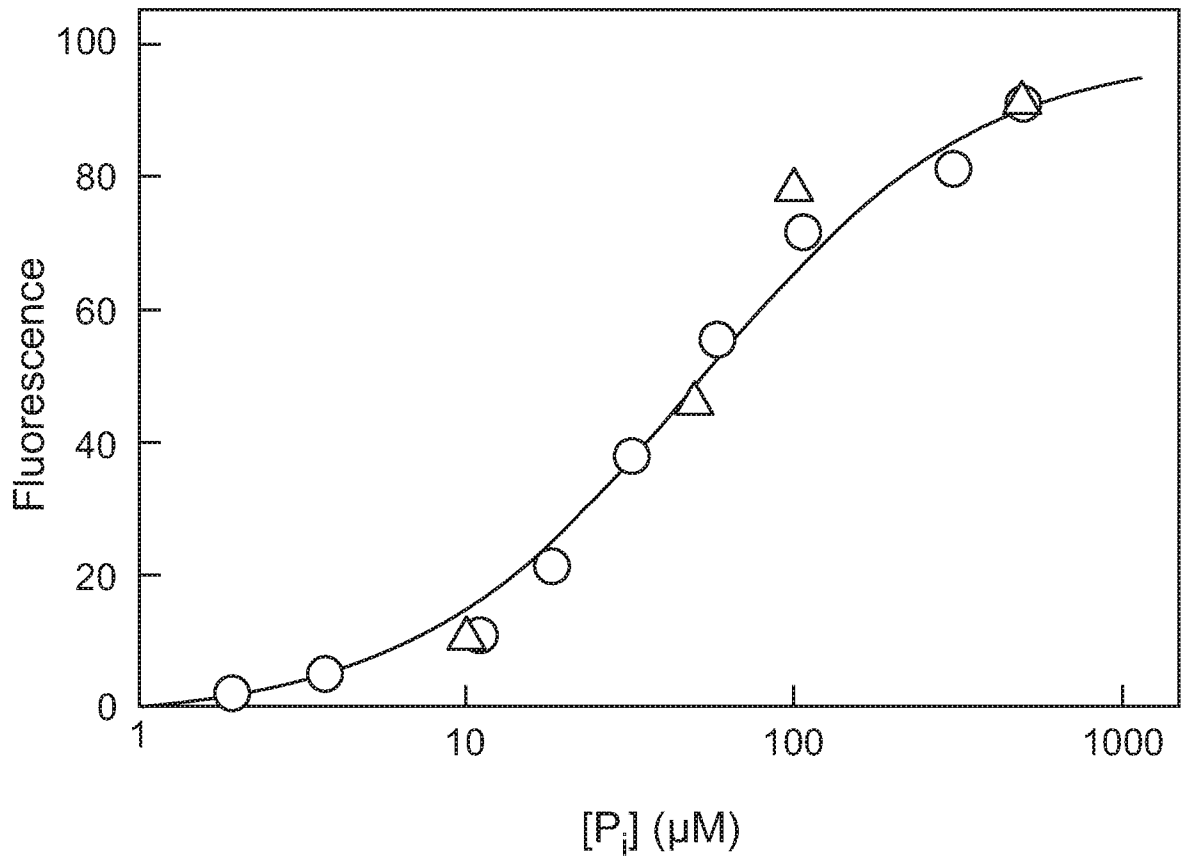


FIGURE 6

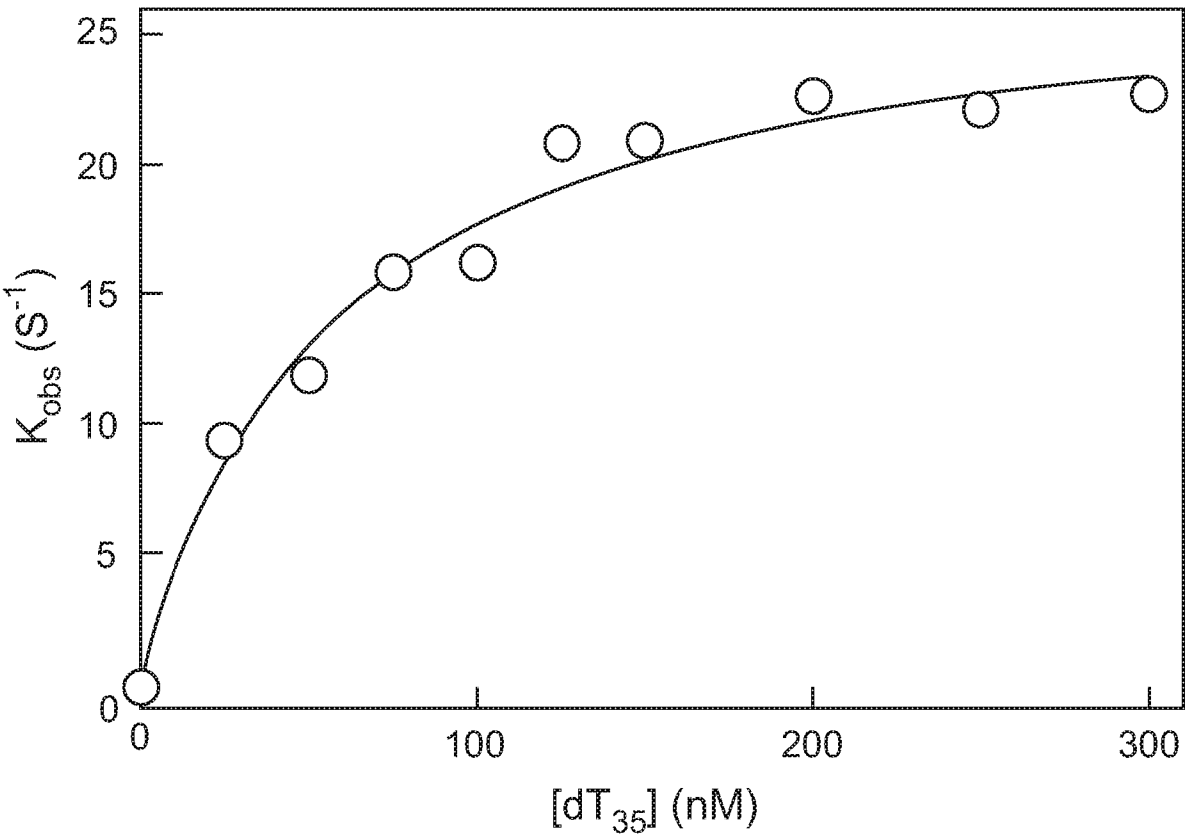
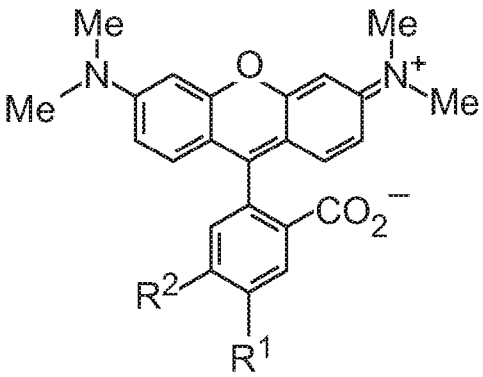


FIGURE 7

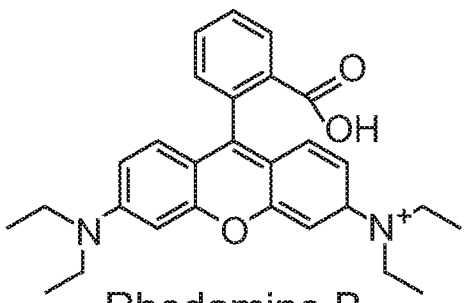


5-IATR

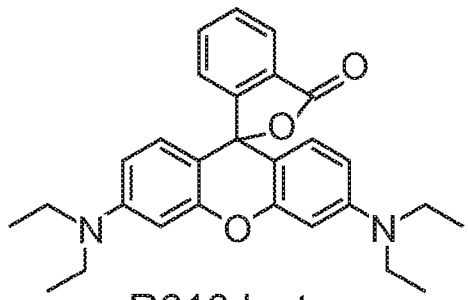
$R^1 = ICH_2CONH$, $R^2 = H$

6-IATR

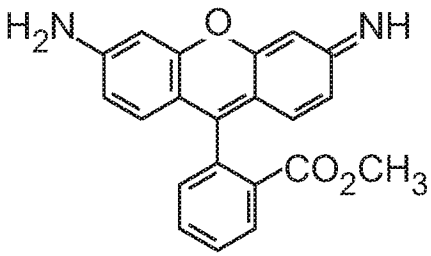
$R^1 = H$, $R^2 = ICH_2CONH$



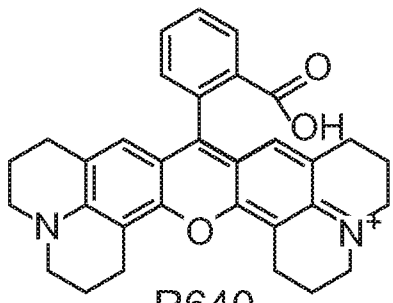
Rhodamine B



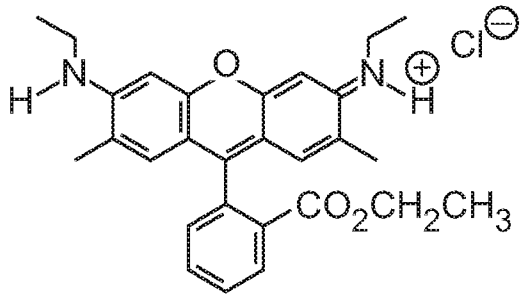
R610 lactone



Rhodamine 123



R640



Rhodamine 6G

FIGURE 8

```
Xylella          -MKVYFAGFTLLCLCTAVTITGCKPSNDNQSTGVSQDGNSTTPPSAEQTKSVKISGAGAS 59
Xanthomonas      MRRTPLPCNGVLRDVIPIATRSCSVISSIKSR---LAVGVLAAALAMGAQAADVTGAGAS 57
Pasteurella      ----------------------------MTKRA---MLLAILFSTFALTTQAQTITGAGAS 30
Haemophilus      ----------------------------MKKK---SYYVLTLGTLPFA-QANSITGAGAS 28
E.coli           ---------------------------MKVMRTTVATVVAATLSMSAFSVFAEASLTGAGAT 35
Shigella         ---------------------------MKVMRTTVATVVAATLSMSAFSVFAEASLTGAGAT 35
Salmonella       ---------------------------MKVMRTTVATVVAATLSMSAFSAFAAASLTGAGAT 35
Erwinia          ----------------------------MHKTLAQCVALTLSLSAVSALAATNLTGAGGT 32
Mesorhizobium    ---------------------MRHFIRSAAVAIAMATASTFTLSAAIAADLSGAGST 36
                                      :                    ::***.:

Xylella          FIYPLISQWSADYNAATGNKINYQSIGSGGGIAQIKAATIDFGSSDKPLDSSELTQAGLG 119
Xanthomonas      FIYPVMSKWSADYNTATKKQVNYQSIGSGGGIAQIKAASVDFGSSDAPLKPEELAAAGLA 117
Pasteurella      FPYPIYAKWASMYEKQTGNKVNYQSIGSGGGQQQIIAKTIDFGASDDPMKAELLAQHQLL 90
Haemophilus      FPYPIYAKWASLYEKETGNKVNYQSIGSGGGQQQIIAKTVDFGASDDPMKSELLQQHQLV 88
E.coli           FPAPVYAKWADTYQKETGNKVNYQGIGSSGGVKQIIANTVDFGASDAPLSDEKLAQEGLF 95
Shigella         FPAPVYAKWADTYQKETGNKVNYQGIGSSGGVKQIIANTVDFGASDAPLSDEKLAQEGLF 95
Salmonella       FPAPVYAKWADTYQKETGNKVNYQGIGSSGGVKQITANTVDFGASDAPLSDEKLNQEGLF 95
Erwinia          FPAPVYNKWAAEYHTATGSQVNYQGIGSSGGVKQIIAKTADFGASDAPMKDEDLAKNGLF 92
Mesorhizobium    FIYPVFAKWADTYKKDTGVGLNYQSIGSGGGIKQVIAKTVTFGATDKPMSDADLEKNGLV 96
                 *  *:  :*:  *.  *   :***.***.**  *: * :  **:;* *:.   *    *
```

FIGURE 8 continued

```
Xylella          QFPSAIGGVVPVVNLDNIEPGKLRLTGPLLADIFLGKISKWNDAAIISANPGLHLPDTKI 179
Xanthomonas      QFPSVIGGVVPVINVPGIAAGAVKLDGKTLGDIFLGKVTTWNDAAIVALNPGVKLPDSKI 177
Pasteurella      QFPAIIGGTVPVVNLPEITAGQLKLSGEVLADIFLGKIKKWNDPAIAKLNQGANLPDKAI 150
Haemophilus      QFPAVIGGIVPVVNLPEIKPGKLKLSGKLLAEIFLGKIKKWNDPDLVALNPTLPLPNKNI 148
E.coli           QFPTVIGGVVLAVNIPGLKSGELVLDGKTLGDIYLGKIKKWDDEAIAKLNPGLKLPSQNI 155
Shigella         QFPTVIGGVVLAVNIPGLKSGELVLDGKTLGDIYLGKIKKWDDEAIAKLNPGLKLPSQNI 155
Salmonella       QFPTVIGGVVLAVNIPGLKSGELVLDGKTLGDIYLGKIKKWDDEAITKLNPGVKLPSQNI 155
Erwinia          QFPTVIGGVVLAVNIPGIKSGELTLDGKTVGDIYLGTVKKWNDPAITKLNPGVKLPDANI 152
Mesorhizobium    QFPMVMGGIVPIVNLTGVKPGELVLDGKTLAQIYLGAITTWDDAAIKALNPSLTLPSTAI 156
                 ***  :** *  :*:  : .* : * *  :.:*:** :..*:*  :   *    **.  *

Xylella          NIVHRSDGSGTTFNFSNYLSKVSAEWKQKVGEGTSVQWPGGVGGKGNEGVASYVQQIKGS 239
Xanthomonas      TVVHRSDGSGTSFNFTNYLSKVNPDWKSKVGEGTAVQWPTGIGGKGNEGVAAYVKQIKGG 237
Pasteurella      IVVHRSDGSGTTFGWTNYLSKVSTEWKETVGQGKSVKWPTGQGGKGNEGVAAYVSKIKYS 210
Haemophilus      IVIHRSDGSGTTFGFTNYLSKISNDWKNQVGEGKSVKWLTGQGGKGNEGVASYVRQMKYS 208
E.coli           AVVRRADGSGTSFVFTSYLAKVNEEWKNNVGTGSTVKWPIGLGGKGNDGIAAFVQRLPGA 215
Shigella         AVVRRADGSGTSFVFTSYLAKVNEEWKNNVGTGSTVKWPIGLGGKGNDGIAAFVQRLPGA 215
Salmonella       AVVRRADGSGTSFVFTSYLAKVNDEWKSKVGAGSTVNWPTGLGGKGNDGIAAFVQRLPGA 215
Erwinia          NVVRRADGSGTSFVFTSYLSKVNKDWSSKVGKGSTVNWPVGLGGKGNDGVAAFVQRLPGS 212
Mesorhizobium    AVVHRSDGSGTTFNFTNYLVKLSPDWKDKVGSDTAVEWPTGVGAKGSEGVANTVKQTDGG 216
                  :::*:*****:* ::.** *:. :*.. ** ..:*:*  * *.**.:*:*  * :   .
```

FIGURE 8 continued

```
Xylella          IGYVELAYALQNKMSYTALQNAAGQWVQPSAESFAAAASNADWSNAKDFNLVITNATGEA 299
Xanthomonas      IGYVELSYALQNKMAYTAMKNAAGKFVQPSDETFAAAANSADWGSSKDFYLVMTNAAGDN 297
Pasteurella      IGYVEYAYAKQNQLAWASLQNKAGQFVQPSAESFMAAAANAQWESAVGMGVILTNEEGDT 270
Haemophilus      IGYVEYAYAKQNQLAWISLQNQAGQFVQPSNESFMAAASHAKWHEKAGMGVILTNETGEK 268
E.coli           IGYVEYAYAKQNNLAYTKLISADGKPVSPTEENFANAAKGADWS--KTFAQDLTNQKGED 273
Shigella         IGYVEYAYAKQNNLAYTKLISADGKPVSPTEENFANAAKGADWS--KTFAQDLTNQKGED 273
Salmonella       IGYVEYAYAKQNNLAYTKLVSADGKPVSPTEESFSNAAKGADWS--KTFAQDLTNQKGDD 273
Erwinia          VGYVEYAYAKQNSLAYTKLVDADGKAIAPSEKSFSDAAKGADWS--TSFAQDLTFQKGDN 270
Mesorhizobium    IGYVEYAYAKQNNLSYSKMLNAAGKVVEPSLESFGAAASNADFKGAKNFNVIITNEPGDT 276
                 :**** :** **.:::  : .  *: : *: :.*  **  *.:    :   :*   *:

Xylella          AWPITATNFILMRKQTKDAAQRKATLDFFKWSFENGQKQANELHYVPLPPNLVKQIEAYW 359
Xanthomonas      AWPITATNFILVQKKPKNPAGLKNTLEFFRWVYSKGDAQAKALDYVPLPDTLVSQIEAYW 357
Pasteurella      SWPVTAASFILLHKHAEKPEITKAVFDFFDWAFKQGKVAATELDYVPLPEEVIQKIQAQW 330
Haemophilus      SWPITAASFILLNKYSDNPETTKNVLAFFDWAFSRGQDAATELDYVPIPADVVSTIKSQW 328
E.coli           AWPITSTTFILIHKDQKKPEQGTEVLKFFDWAYKTGAKQANDLDYASLPDSVVEQVRAAW 333
Shigella         AWPITSTTFILIHKDQKKPEQGTEVMKFFDWAYKTGAKQANDLDYASLPDSVVEQVRAAW 333
Salmonella       VWPITSTTFILVHKAQKKPEQGAEVLKFFDWAYKNGAKQANDLDYASLPDNVVEQVRTAW 333
Erwinia          AWPITSTTFILVHKEQANTAKGAAVLQFFDWAYKNGGKTTSALDYASLPAPVVEQIRAAW 330
Mesorhizobium    TWPIAASTWVLIHKAPDDAAATGEALKFFAWAYKDGKETAKALDYVSIPDSVVDLIKASW 336
                  **;:::.::*:.*   ..     .: ** * :. *   :. *.*..:*  ::. :.: *
```

FIGURE 8 continued

Xylella	ASEFK----------	364
Xanthomonas	AKTLPR----------	363
Pasteurella	KTEVKSSDGKTLWQ--	344
Haemophilus	KTELKQ----------	334
E.coli	KTNIKDSSGKPLY---	346
Shigella	KTNIKDSSGKPLY---	346
Salmonella	KTSIKDSNGKALY---	346
Erwinia	KSNVKDSSGKALY---	343
Mesorhizobium	KADIQAG-GKPVYAGE	351

FIGURE 9

	Accession no.	Percent identity*
Xylella fastidiosa 9a5c	Q9PBK3	52%
Xanthomonas campestris pv. campestris str. 8004.	Q4UT67	54%
Pasteurella multocida subsp. multocida str. Pm70	Q9CNJ4	55%
Haemophilus influenzae Rd KW20	P45192	54%
Escherichia coli K-12	P0AG82	
Shigella flexneri serotype 5b (strain 8401)	Q0SYS3	98%
Salmonella schwarzengrund (strain CVM19633)	B4TN22	91%
Erwinia amylovora (strain ATCC 49946)	D4IGH7	74%
Mesorhizobium loti MAFF303099	Q98FL2	56%

*Percent identity is compared against the *E. coli* sequence.

The algorithm excludes non-aligning termini.

The accession numbers are from Uniprot.

INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2016/050814

A. CLASSIFICATION OF SUBJECT MATTER
INV. C07K14/245
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, BIOSIS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	GU H ET AL: "A novel analytical method for in vivo phosphate tracking", FEBS LETTERS, ELSEVIER, AMSTERDAM, NL, vol. 580, no. 25, 30 October 2006 (2006-10-30), pages 5885-5893, XP028061443, ISSN: 0014-5793, DOI: 10.1016/J.FEBSLET.2006.09.048 [retrieved on 2006-10-30]	27,28
Y	abstract; table 1 page 5886, left-hand column, last paragraph - page 5887, left-hand column, last paragraph	3-11, 13-26

-/--

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
20 June 2016	27/07/2016

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Strobel, Andreas

2

Form PCT/ISA/210 (second sheet) (April 2005)

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	OKOH MICHAEL P ET AL: "A Biosensor for Inorganic Phosphate Using a Rhodamine-Labeled Phosphate Binding Protein", BIOCHEMISTRY, vol. 45, no. 49, 15 November 2006 (2006-11-15), pages 14764-14771, XP002758965, DOI: 10.1021/bi060960j	27,28
A	abstract; table 1	1-26

X	KUNZELMANN SIMONE ET AL: "Fluorescent, Reagentless Biosensors for Phosphate: Altering Properties, Expanding Usefulness", BIOPHYSICAL JOURNAL, vol. 102, no. 3, 28 February 2012 (2012-02-28), XP028893400, ISSN: 0006-3495, DOI: 10.1016/J.BPJ.2011.11.2546	28
A	abstract	1-27

Y	BENSON D E ET AL: "Design of bioelectronic interfaces by exploiting hinge-bending motions in proteins", SCIENCE, AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE, US, vol. 293, no. 5535, 31 August 2001 (2001-08-31), pages 1641-1644, XP002231572, ISSN: 0036-8075, DOI: 10.1126/SCIENCE.1062461 the whole document	1,3-11, 13-28

Y	MARVIN JONATHAN S ET AL: "Manipulation of ligand binding affinity by exploitation of conformational coupling", NATURE STRUCTURAL BIOLOGY, NATURE PUBLISHING GROUP, NEW YORK, NY, US, vol. 8, no. 9, 1 September 2001 (2001-09-01), pages 795-798, XP002451002, ISSN: 1072-8368, DOI: 10.1038/NSB0901-795 abstract; figure 1 page 796, left-hand column, line 13 - line 14	1,3-11, 13-28

	-/--	

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	MARVIN J S ET AL: "THE RATIONAL DESIGN OF ALLOSTERIC INTERACTIONS IN A MONOMERIC PROTEIN AND ITS APPLICATIONS TO THE CONSTRUCTION OF BIOSENSORS (RATIONAL PROTEIN DESIGN/MALTOSE-BINDING PROTEIN/ALLOSTERY)", PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, NATIONAL ACADEMY OF SCIENCES, US, vol. 94, no. 9, 1 April 1997 (1997-04-01), pages 4366-4371, XP001108836, ISSN: 0027-8424, DOI: 10.1073/PNAS.94.9.4366 abstract; table 1	1,3-11, 13-28
A	DWYER M A ET AL: "Periplasmic binding proteins: a versatile superfamily for protein engineering", CURRENT OPINION IN STRUCTURAL BIOLOGY, ELSEVIER LTD, GB, vol. 14, no. 4, 1 August 2004 (2004-08-01) , pages 495-504, XP027347978, ISSN: 0959-440X [retrieved on 2004-08-13] abstract; figure 1 page 496, right-hand column, paragraph 4; figure 2	1-28
A	WEBB MARTIN R: "Development of fluorescent biosensors for probing the function of motor proteins", MOLECULAR BIOSYSTEMS, vol. 3, no. 4, April 2007 (2007-04), pages 249-256, XP002758966, abstract page 251, right-hand column, line 21 - page 252, middle column, line 3	1-28
X,P	SOLSCHEID CLAUDIA ET AL: "Development of a Reagentless Biosensor for Inorganic Phosphate, Applicable over a Wide Concentration Range.", BIOCHEMISTRY, vol. 54, no. 32, 18 August 2015 (2015-08-18), pages 5054-5062, XP002758967, ISSN: 1520-4995 the whole document	1-28

2