



(51) International Patent Classification:

C12N 9/00 (2006.01) *G01N 21/76* (2006.01)
C12Q 1/00 (2006.01)

(21) International Application Number:

PCT/GB2016/050818

(22) International Filing Date:

23 March 2016 (23.03.2016)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

1505266.5 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.**; National Institute for Medical Research, The Ridgeway, London NW7 1AA (GB).
VANCRAENENBROECK, Renee; National Institute for Medical Research, The Ridgeway, London NW7 1AA (GB).

(74) Agents: **SCRIPT IP LIMITED** et al.; 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))

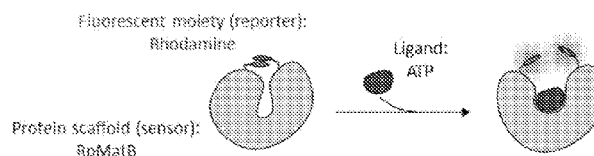
Published:

— 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))

(54) Title: ATP SENSOR

FIGURE 18



(57) Abstract: The invention relates to an ATP binding molecule comprising a polypeptide which undergoes a conformational change from a first conformation to a second conformation upon binding of ATP, said polypeptide comprising amino acid sequence corresponding to at least amino acids 5 to 497 of SEQ ID NO:1, said polypeptide having at least 21% sequence identity to SEQ ID NO:1, said polypeptide having a value of RMSD <4 Å relative to RpMat B in the ATP-bound conformation, wherein said polypeptide comprises a cysteine residue for attachment of at least one reporter moiety, and comprises at least one reporter moiety attached thereto, wherein said cysteine residue for attachment of at least one reporter moiety is positioned such that said reporter moiety undergoes a change in fluorescence upon changing from said first conformation to said second conformation. The invention also relates to use of such a sensor, and to methods for assay of ATP.



ATP SENSOR

BACKGROUND TO THE INVENTION

5 ATP is an intracellular energy source, for example involved in active transport, cell motility and biosynthesis. It is also an important extracellular signalling agent in neurotransmission (Burnstock 2012) and inflammation (Idzko, Ferrari et al. 2014).

10 ATP is generated through several pathways such as glycolysis, the Krebs cycle and oxidative phosphorylation. This makes it an important assay target and monitoring.

ATP production is widely used to measure enzyme activity in biochemical and cell-based applications.

15 Various ATP assays are known in the art.

However some of those, such as radioisotope assays, are not continuous, which is a problem.

20 Coupled-enzyme assays, including the luciferase-luciferin system (Patergnani, Baldassari et al. 2014), require several reagents, which is a drawback.

Others, like aptamer-based methods (Feng, Dai et al. 2014), require a long incubation time, which is a disadvantage.

25 Some, like organic probes (Zhou, Xu et al. 2011), show low selectivity for ATP over other nucleotides and only give a color change and no concentration-dependent read-out, which limits their usefulness.

30 Others such as (micro)electrode biosensors (Dale and Frenguelli 2012) measure ATP indirectly, which risks introducing errors.

The present invention seeks to overcome problem(s) associated with the art.

SUMMARY OF THE INVENTION

A fluorescent, reagentless biosensor for ATP is described. Such biosensors for a target molecule are a single molecular species that consists minimally of a recognition element and a reporter. In this case, the recognition element is a protein that interacts with the target, ATP, namely an ANL superfamily protein. Most suitably the ANL superfamily protein malonyl-coenzymeA synthetase from *Rhodopseudomonas palustris* (RpMatB) is used. Amino acid sequence derived from this protein is coupled covalently to reporter fluorophore(s) to give a fluorescence change on ATP binding.

In more detail, a fluorescent reagentless biosensor for ATP was developed based on malonyl-coenzyme A synthetase from *Rhodopseudomonas palustris* (RpMatB) as the protein scaffold and recognition element. In one embodiment two 5-iodoacetamidotetramethylrhodamines were covalently bound to RpMatB to provide the readout. This adduct couples ATP binding to a 3.7-fold increase in fluorescence intensity with excitation at 553 nm and emission at 575 nm. It has micromolar sensitivity for ATP and is highly selective for ATP relative to ADP. Its ability to monitor ATP production was demonstrated in a steady-state kinetic assay in which ATP is a product.

Thus in one aspect the invention provides an ATP binding molecule comprising a polypeptide which undergoes a conformational change from a first conformation to a second conformation upon binding of ATP, said polypeptide comprising amino acid sequence corresponding to at least amino acids 5 to 497 of SEQ ID NO:1, said polypeptide having at least 21% sequence identity to SEQ ID NO:1, said polypeptide having a value of RMSD <4 Å relative to RpMatB in the ATP-bound conformation, wherein said polypeptide comprises a cysteine residue for attachment of at least one reporter moiety, and comprises at least one reporter moiety attached thereto, wherein said cysteine residue for attachment of at least one reporter moiety is positioned such that said reporter moiety undergoes a change in fluorescence upon changing from said first conformation to said second conformation.

Suitably the RMSD is calculated using the PyMOL program.

Suitably the RMSD is calculated using the “align” command in the PyMOL program.

Suitably RpMatB in the ATP-bound conformation is taken as PDB number '4FUT'. Thus suitably said polypeptide has a value of RMSD $<4 \text{ \AA}$ relative to PDB number '4FUT' (i.e. RpMatB in the ATP-bound conformation).

- 5 In one embodiment, suitably sequence identity to SEQ ID NO:1 is for the amino acid residues corresponding to those shown in column II of table A. Column II of table A corresponds to 'exposed' residues.

Suitably the polypeptide comprises a first cysteine residue at a position corresponding
10 to a position selected from **R286**, A282, D283, H285, E287, S289, A290, **K385**, L383, G384, I386 and D287 of SEQ ID NO: 1,
and wherein the polypeptide comprises a second cysteine residue at a position corresponding to a position selected from
Q457, A456, V458, H460, G461, **G461**, Q457, H460, L462, G464, Q465, **K470**, L466,
15 F469 and M471 of SEQ ID NO: 1.

Suitably the polypeptide comprises a first cysteine residue at a position corresponding to a position selected from **R286** and **K385** of SEQ ID NO: 1,
and wherein the polypeptide comprises a second cysteine residue at a position
20 corresponding to a position selected from **Q457**, **G461** and **K470** of SEQ ID NO: 1.

Suitably the polypeptide comprises a first cysteine residue at a position corresponding to position R286 of SEQ ID NO: 1,
and wherein the polypeptide comprises a second cysteine residue at a position
25 corresponding to position G461 of SEQ ID NO: 1.

Suitably the polypeptide comprises a first cysteine residue at a position corresponding to position K385 of SEQ ID NO: 1,
and wherein the polypeptide comprises a second cysteine residue at a position
30 corresponding to position K470 of SEQ ID NO: 1.

Suitably the polypeptide comprises a first cysteine residue at a position corresponding to position R286 of SEQ ID NO: 1,
and wherein the polypeptide comprises a second cysteine residue at a position
35 corresponding to position Q457 of SEQ ID NO: 1.

Suitably said molecule comprises the amino acid sequence of SEQ ID NO:4, SEQ ID NO:5 or SEQ ID NO:6.

5 Suitably said molecule further comprises at least two tetramethylrhodamine moieties attached thereto. Suitably each of said at least two tetramethylrhodamine moieties is independently selected from the group consisting of 5-tetramethylrhodamine and 6-tetramethylrhodamine.

10 In one aspect, the invention relates to an ATP binding molecule as described above wherein said cysteine residue for attachment of a reporter moiety is at a position corresponding to a position selected from **E439**, I403, P433, G438, G440, **N492**, K491, V493, R495, E496 and T497 of SEQ ID NO: 1.

Suitably said cysteine residue for attachment of a reporter moiety is at a position corresponding to **E439** or **N492** of SEQ ID NO: 1.

15 Suitably said molecule comprises the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:3.

Suitably said molecule further comprises at least one diethylaminocoumarin moiety attached thereto.

20 Suitably said diethylaminocoumarin moiety is independently selected from the group consisting of (*N*-[2-(1-maleimidyl)ethyl]-7-diethylaminocoumarin-3-carboxamide and *N*-[2-(iodoacetamido)ethyl]-7-diethylaminocoumarin-3-carboxamide).

25 In one aspect, the invention relates to an ATP binding molecule as described above wherein said polypeptide comprises an amino acid other than Threonine at the position corresponding to T167 of SEQ ID NO: 1. Suitably said polypeptide comprises alanine or serine at the position corresponding to T167 of SEQ ID NO: 1, preferably alanine.

30 In one aspect, the invention relates to an ATP binding molecule as described above wherein said polypeptide comprises an amino acid other than Serine at the position corresponding to S170 of SEQ ID NO: 1. Suitably said polypeptide comprises alanine at the position corresponding to S170 of SEQ ID NO: 1.

35 In one aspect, the invention relates to an ATP binding molecule as described above wherein said polypeptide comprises an amino acid other than Threonine at the position corresponding to T303 of SEQ ID NO: 1. Suitably said polypeptide comprises alanine or serine at the position corresponding to T303 of SEQ ID NO: 1, preferably alanine.

In one aspect, the invention relates to an ATP binding molecule as described above wherein said polypeptide comprises an amino acid other than cysteine at the position corresponding to C106 of SEQ ID NO: 1.

- 5 In one aspect, the invention relates to an ATP binding molecule as described above wherein said polypeptide comprises alanine at the position corresponding to C106 of SEQ ID NO: 1. (*C106A*)

- 10 In one aspect, the invention relates to an ATP binding molecule as described above wherein said polypeptide comprises an active site mutation such that the polypeptide is catalytically inactive for ATP hydrolysis.

- 15 In one aspect, the invention relates to an ATP binding molecule as described above wherein said active site mutation comprises an amino acid other than lysine at the position corresponding to K488 of SEQ ID NO: 1.

- 20 In one aspect, the invention relates to an ATP binding molecule as described above wherein said polypeptide comprises alanine at the position corresponding to K488 of SEQ ID NO: 1. (*K488A*)

In one aspect, the invention relates to an ATP binding molecule as described above further comprising the sequence of SEQ ID NO: 8.

- 25 In one aspect, the invention relates to a nucleic acid having a nucleotide sequence encoding the polypeptide portion of an ATP binding molecule as described above.

- 30 In one aspect, the invention relates to a method for monitoring changes in ATP concentration in a sample comprising contacting said sample with an ATP binding molecule as described above and determining changes in conformation of said ATP binding molecule, wherein changes in conformation of said ADP binding molecule indicate changes in the concentration of ATP in said sample.

Suitably the conformation of said ATP binding molecule is monitored by measurement of changes in fluorescence of a fluorophore comprised by said ATP binding molecule.

Suitably the sample comprises divalent Magnesium ion (Mg^{2+}).

- 35 In one aspect, the invention relates to use of an ATP binding molecule as described above in the determination of ATP concentration in a sample.

In one aspect, the invention relates to an ATP binding molecule comprising a polypeptide which undergoes a conformational change from a first conformation to a second conformation upon binding of ATP, said polypeptide comprising amino acid
5 sequence corresponding to at least amino acids 5 to 497 of SEQ ID NO:1,
said polypeptide having at least 34% sequence similarity to SEQ ID NO:1,
wherein said polypeptide comprises a cysteine residue for attachment of at least one reporter moiety, and comprises at least one reporter moiety attached thereto,
wherein said cysteine residue for attachment of at least one reporter moiety is
10 positioned such that said reporter moiety undergoes a change in fluorescence upon changing from said first conformation to said second conformation.

In one aspect, the invention relates to an ATP binding molecule comprising a polypeptide which undergoes a conformational change from a first conformation to a second conformation upon binding of ATP, said polypeptide comprising amino acid
15 sequence corresponding to at least amino acids 5 to 497 of SEQ ID NO:1,
said polypeptide having at least 21% sequence identity to SEQ ID NO:1,
wherein said polypeptide comprises a cysteine residue for attachment of at least one reporter moiety, and comprises at least one reporter moiety attached thereto,
20 wherein said cysteine residue for attachment of at least one reporter moiety is positioned such that said reporter moiety undergoes a change in fluorescence upon changing from said first conformation to said second conformation.

DETAILED DESCRIPTION OF THE INVENTION

25 The invention provides a single solution to the problem of provision of an ATP biosensor. In one embodiment, the biosensor is singly labelled. In another embodiment, the biosensor is doubly labelled. In the singly labelled embodiment, the single dye molecule attached to the biosensor interacts with the protein, and this
30 interaction affects the signal such as fluorescence. Binding of ATP changes the conformation of the biosensor, and thus alters the interaction the dye molecule with the polypeptide. This ATP bound states exhibits a difference in reporter activity such as fluorescence and therefore serves to read out the event of ATP binding. In a double
labelled embodiment, the two dye molecules attached to the biosensor interact with one
35 another, for example via dye stacking. When ATP binds to the biosensor, the conformation changes and this conformation change affects the readout such as fluorescence of the two dye molecules. This change in reporter activity such as

fluorescence again reads out the presence of ATP. It must be emphasised that whether the biosensor is singly labelled or doubly labelled as outlined, the invention is based upon the same physical conformation change, and the dye(s) is/are attached to the biosensor in the different embodiments are responding to this same conformational change. Therefore, the principle of the invention is the same regardless of whether the
5 single labelled embodiment or the doubly labelled embodiment is considered.

In other words, the single label or dual labelling approaches both rely on a signal such as fluorescence signal due to the closing of the “lid domain” on ATP binding. The single
10 label detects the movement through interactions with the protein surface. The dual (e.g. rhodamine) label detects the relative position of the two labels, which changes on the closure.

For some applications, a doubly labelled biosensor of the invention is especially
15 suitable. Suitably the biosensor comprises a first label and a second label which can exhibit molecular stacking and wherein the molecular stacking is altered on changing from one conformation to the other. Suitably the first and second labels can exhibit molecular stacking either (a) in the first conformation but not in the second conformation, or (b) in the second conformation but not in the first conformation. In
20 one embodiment the first and second labels exhibit molecular stacking in the first conformation. In one embodiment the first and second labels exhibit molecular stacking in the second conformation.

A doubly labelled biosensor suitably uses rhodamine dye molecules. These
25 embodiments of the invention can produce the largest signal. In addition, dyes such as rhodamine can exhibit superior stability during fluorescence measurements, such as greater stability under irradiation. For example, rhodamine is more photostable than coumarins. For example, rhodamine has a longer wavelength excitation, often making it easier to use. Thus suitably the dye used in the biosensors of the invention is
30 rhodamine.

In arriving at their choice of donor molecule from which to create the sensors of the invention, the inventors faced numerous challenges. One such challenge was the availability of 3-dimensional structures from which to study the candidate
35 polypeptides. For example, two crystal structures for RpMatB were used, an apo conformation and an ATP-bound conformation. Although the published structures were crystal structures, in reality this is a flexible and dynamic protein. In particular

the protein exhibits a moving cap feature. However, there are no NMR structures (which tend to show such flexibility) available. This presented challenges which were overcome by the present invention.

- 5 Considerable search effort was made by the inventors in arriving at the choice of RpMatB and ANL superfamily proteins as suitable sensor molecules.

The choices of suitable sites for dye attachment within the sensors of the invention were intellectually challenging. The inventors overcame numerous problems in
10 arriving at the teachings presented herein.

Without wishing to be bound by theory, it is believed that sensors of the invention may exist in two conformations. It is thought that these conformations may reflect monomeric and dimeric forms of the protein. It appears that the monomeric form is
15 the most effective at ATP binding. Nevertheless, it is clear from the practical experiments provided in the application that whether or not this theory is correct, the sensors perform very well regardless of the precise description of their monomeric/dimeric forms.

20 ANL SUPERFAMILY

Suitably other proteins in the ANL superfamily may be used as ATP sensors according to the invention when appropriately labelled.

- 25 In more detail, the inventors have studied the structural similarity amongst the ANL superfamily members, and we teach that the information from the exemplary sensor based on RpMatB allows positioning of acceptor amino acids such as cysteines (i.e. cysteine substitutions), and hence reporter moieties (e.g. fluorophores), on other ANL superfamily members that bind ATP and show similar ATP-dependent conformation
30 change. Thus suitably the polypeptide of a sensor of the invention may comprise a structural homologue of RpMatB, such as an ANL superfamily protein.

The term 'ANL superfamily' is well known in the art.

- 35 In case any guidance is needed, the ANL superfamily of adenyating enzymes contains acyl- and aryl-CoA synthetases, firefly luciferase, and the adenylation domains of the modular Non-Ribosomal Peptide Synthetases (NRPSs). Members of this family catalyse

two partial reactions, the initial adenylation of a carboxylate to form an acyl-AMP intermediate, followed by a second partial reaction, most commonly, the formation of a thioester.

- 5 This is described in more detail in the art, such as in Gulick 2009 (ACS Chem Biol. 2009 October 16; 4(10): 811–827. doi:10.1021/cb900156h), which is incorporated herein by reference specifically for the information relating to the ANL superfamily.

In case any further information is helpful, reference is made to the Pfam database. The
10 Pfam database is a large collection of protein families, each represented by multiple sequence alignments and hidden Markov models (HMMs). (The Pfam protein families database: R.D. Finn, A. Bateman, J. Clements, P. Coghill, R.Y. Eberhardt, S.R. Eddy, A. Heger, K. Hetherington, L. Holm, J. Mistry, E.L.L. Sonnhammer, J. Tate, M. Punta Nucleic Acids Research (2014) Database Issue 42:D222-D230). Regarding the Pfam
15 database, suitably the release number of the database referred to is Pfam 27.0 (March 2013, 14831 families).

The Pfam description of the ANL superfamily (<http://pfam.xfam.org/ clan/ANL>) mentions:

20

“This superfamily consists of enzymes including luciferase, long chain fatty acid Co-A ligase, acetyl-CoA synthetase and various other closely-related synthetases as well as a plant auxin-responsive promoter family. The name ANL derives from from three of the subfamilies - Acyl-CoA synthetases, the NRPS adenylation domains, and the Luciferase
25 enzymes [Gulick AM; ACS Chem Biol. 2009;4:811-827.: Conformational dynamics in the Acyl-CoA synthetases, adenylation domains of non-ribosomal peptide synthetases, and firefly luciferase.]. Members of this superfamily catalyse the initial adenylation of a carboxylate to form an acyl-AMP intermediate, followed by a second partial reaction, most commonly the formation of a thioester [Gulick AM; ACS Chem Biol. 2009;4:811-
30 827.: Conformational dynamics in the Acyl-CoA synthetases, adenylation domains of non-ribosomal peptide synthetases, and firefly luciferase.].”

In structural terms, the “lid” domain, upon which the conformational change on ATP binding is based, is common to ANL superfamily proteins.

35

In addition, ANL superfamily members have two partial reactions.

Thus in a broad aspect the invention relates to ATP sensors which comprise an ANL superfamily protein. Suitably said ANL superfamily protein comprises variants of RpMatB having amino acid substitutions as described, or comprises RpMatB as exemplified.

5

Guidance regarding the sequence identity/similarity between ANL superfamily members is presented below.

Suitably sequence identity/similarity is assessed along the whole length of the amino acid sequence present in the polypeptide component of the ATP sensor, unless otherwise specified.

An example of low sequence homology with RpMatB for proteins in the PDB belonging to the ANL superfamily is for 3DHV which has

- 15 - A sequence identity of 21% with RpMatB
- A sequence similarity of 38% (calculated using the BLOSUM62 matrix with a gap penalty of 10.0 and an extend penalty of 0.5).
- An RMSD value of 2.9 Å compared to 4FUT (both proteins are in the “closed” ATP-bound conformation and the RMSD value was calculated in PyMOL using the “align” command).

20

Sequence similarity is well known in the art and takes account of conservative substitutions (see table below).

25 However, more importantly, the ANL superfamily proteins have the same structural fold as calculated via the RMSD (root mean square deviation) value. Suitably the polypeptide ATP sensors of the invention have this fold.

RMSD (Root Mean Square Deviation) is the square root of the mean of the square of the distances between the matched atoms:

30

$$\text{RMSD} = \sqrt{\{\text{SUM}(d_{ii})^2\}/N}$$

where d_{ii} is the distance between the i^{th} atom of structure 1 and the i^{th} atom of structure 2 and N is the number of atoms matched in each structure.

35 An RMSD is calculated between structures (rather than sequences).

The RMSD can be used to compare protein three-dimensional structures. The RMSD is 0 for identical structures, and its value increases as the two structures become more different. RMSD values are considered as reliable indicators of variability when applied to very similar proteins. Generally, a value of RMSD $<4 \text{ \AA}$ indicates significant structural similarities.

Different programs give different RMSD values, not because they use a different formula but because they use different cut-off parameters to determine which atoms included in the formula. Therefore, the program and command used to calculate the RMSD value must be taken into account.

For example, PyMOL (used in Example 12) calculates it after cycles of refinement in order to reject structural outliers found during the fit – at least when you use the “align” command. Another program, SuperPose does not follow this procedure. Its RMSD is usually higher than the one calculated via PyMOL (it takes more atoms into account). The skilled worker will be aware of these differences.

Suitably the RMSD is calculated using the PyMOL program.
Suitably RMSD values mentioned herein are calculated using the PyMOL program.

The PyMOL program is available for example from <https://www.pymol.org/>, or from Schrödinger, 101 SW Main Street, Suite 1300, Portland, OR 97204, USA.

Suitably the RMSD is calculated using the “align” command in the PyMOL program.
Suitably RMSD values mentioned herein are calculated using the “align” command in the PyMOL program.

RpMatB exists in multiple conformations (see below), so it is important to calculate the RMSD values for the structure pairs that have the same conformation.
Suitably RMSD values mentioned herein are for the ATP bound conformation(s).
Suitably the reference structure is RpMatB in the ATP-bound conformation.
See for example Figure 16 (Supplemental Table 3).

Suitably the polypeptide component of the ATP sensor of the invention comprises an ANL superfamily protein. Suitably said polypeptide comprises a value of RMSD $<4 \text{ \AA}$

relative to RpMatB; more suitably said polypeptide comprises a value of RMSD $<3 \text{ \AA}$ relative to RpMatB.

5 It is straightforward for the skilled worker to identify residues in such protein(s) corresponding to the labelling/substitution positions in RpMatB. In case any further information is required, guidance is provided below, particularly in the examples section.

RpMatB

10

RpMatB was chosen as the exemplary recognition element of the ATP sensors described because of several properties, including high yield expression and purification, good stability and high affinity and selectivity for ATP. RpMatB belongs to the AMP-forming acyl-coenzymeA synthetase family (PF00501 (Finn, Bateman et al. 15 2014)) and the ANL superfamily containing acyl- and aryl-coenzymeA synthetases, the adenylation domains of nonribosomal peptide synthetases and firefly luciferase (Gulick 2009). RpMatB catalyzes the conversion of malonate and coenzymeA to malonyl-coenzymeA via a ping-pong mechanism consuming ATP through a malonyl-AMP intermediate. Its products are AMP, pyrophosphate and malonyl-coenzymeA.

20 RpMatB has been crystallized in two conformations, that is an open form of the apoprotein and a closed form with MgATP bound (Crosby, Rank et al. 2012) (Figure 1). The ligand-binding pocket is between the N- and C-terminal lobes. Upon binding MgATP, the C-terminal lobe rotates $\sim 20^\circ$ to close the binding cleft. This conformational change was used to create a series of rationally designed RpMatB 25 mutants with cysteine point mutations of surface amino acid residues in order to incorporate thiol-reactive fluorophores.

We describe different design strategies to create a fluorescence signal that can be used to report the ATP concentration. One is based on the introduction of one environmentally sensitive fluorophore, namely diethylaminocoumarin. The other relies 30 on reversible stacked dimer formation between a pair of identical fluorophores, more particular two tetramethylrhodamines. Both fluorophores have been used to develop biosensors. Examples include an inorganic phosphate biosensor based on the E. coli phosphate binding protein (Brune, Hunter et al. 1994, Okoh, Hunter et al. 2006), a single stranded DNA biosensor based on the E. coli single stranded DNA binding 35 protein (Dillingham, Tibbles et al. 2008) and an ADP biosensor based on the bacterial actin homologue, ParM (Kunzelmann and Webb 2009, Kunzelmann and Webb 2010).

The exemplary ATP biosensor described herein is an adduct of RpMatB and tetramethylrhodamine, which specifically responds to ATP with a maximum 3.7-fold fluorescence increase. Its sensitivity lies in the micromolar range. Its ability to monitor ATP production was demonstrated with a steady-state kinetic assay to measure the time course of enzymatic ATP production.

In order to introduce the taught mutations into the particular ANL superfamily protein which it is designed to use, then the corresponding sites for mutation should be identified by reference to the exemplary RpMatB sequence (SEQ ID NO: 1).

Suitably the polypeptide of a sensor of the invention comprises amino acid sequence corresponding to RpMatB amino acid sequence, comprising substitutions as described.

wild-type RpMat sequence (accession number: Genbank CAE25665.1)

MNANLFAFLFDKLDLDPHKLAIETAAGDKISYAEIVARAGRVANVLVARGLQVGDRVAAQTEKSVEALVLY
LATVRAGGVYLPNTAYTLHELDYFITDAEPKIVVCDPSKRDGIAAIAAKVGATVETLGPDGRGSLTDAA
AGASEAFATIDRGADDLAAILYTS GTTGRSKGAMLSHDNLASNSLTLDVYWRFTPDDVLIHALPIYHTHG
LFVASNVTLFARGSMIFLPKFDPKILDLMARATVLMGVPTFYTRLLQSPRLTKETTGHMRLFISGSAPL
LADTHREWSAKTGHAVLERYGMTETNMNTSNPYDGDVRVPGAVGPALPGVSARVTDPETGKELPRGDIGMI
EVKGPNVFKGYWRMPEKTKSEFRDDGFFITGDLGKIDERGYVHILGRGKDLVITGGFNVPKEIESEIDA
MPGVVESAVIGVPHADFGEGVTAVVVRDKGATIDEAQVLHGLDGLAQKAKMPKKVIFVDDLPRNTMGKVQ
KNVLRETYKDIYK (**SEQ ID NO: 1**)

RpMatB polypeptides

The polypeptide components of the molecules of the invention are based on ANL superfamily polypeptide sequences such as the exemplary RpMatB sequence. In particular, amino acid addresses given in the application correspond to the numbering of the RpMatB reference sequence of SEQ ID NO:1. Where truncated or extended forms of RpMatB are used as polypeptides in molecules of the invention (e.g. where a his tag is added or where a section of the polypeptide is deleted) then the amino acid numbering should be treated as corresponding to the equivalent section of the full length RpMatB reference sequence and not as an 'absolute' or rigidly inflexible numeric address. By way of explanation, if the description mentions a substitution of K488, this means amino acid 488 of the RpMatB reference sequence of SEQ ID NO: 1. If the polypeptide used is truncated by deletion of the first 10 amino acids, the address given will still be K488 (rather than e.g. K478) – this will be easily understood by the skilled reader to refer to the amino acid of the corresponding context with reference to the full length RpMatB sequence of SEQ ID NO:1, as is conventional in the art.

Clearly there are elements of the RpMatB wild type sequence which we teach are important to mutate, such as by substitution, to achieve certain advantages. However, there are also numerous residues which may or may not be mutated depending on operator choice. Clearly there are also numerous residues which should not be mutated in case such mutation would interfere with the function of the polypeptide. Typically it would be expected that if the skilled operator had a concern whether or not a particular residue could be mutated or not, they could make the mutation and then test the resulting polypeptide to ensure that the desired property was retained in the mutated version. However, in order to provide further guidance on this point, the following comments are made:

The exemplary biosensor polypeptide RpMatB has been studied in detail and each residue has been classified as set out in Table A below. The classification is as follows, making use of information on surface exposure:

Buried: Residues in the core of the structure where mutations are likely to affect the function of the sensor.

Exposed: Residues on the surface of the protein. Mutations of these residues are likely to retain the sensor function. Possibly any mutation may be allowed

Intermediate: Change in sensor function when mutating these residues is possible, but may be difficult to predict and conservative mutations are likely to be more successful e.g. routine testing of resulting mutants is particularly preferred when mutating these residues. Most of these are partially buried.

Residues marked with an asterisk (*) designate active site residues. Active site residues may suitably be specifically mutated in order to optimise the sensor functions (such as to impair or eliminate ATP hydrolysis such as for example mutating K488). These are discussed in more detail in the text.

In more detail, the relative accessible surface area (ASA) of a residue is its ASA calculated using DSSP (Kabsch, W. and C. Sander (1983). "Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features." Biopolymers 22(12): 2577-637.) divided by its nominal maximum area as defined by (Chothia, C. (1976). "The nature of the accessible and buried surfaces in proteins." J Mol Biol 105(1): 1-12). Here, a residue is defined as exposed if its relative ASA is at least 40% of its nominal maximum area. A residue is defined as buried if its

relative ASA is less than 10% of its nominal maximum area. Residues marked with an asterisk (*) designate active site residues (within 6 Å of ATP and Mg²⁺).

Table A – RpMatB residues

5

Buried (I)		Exposed (II)		Intermediate (III)	
Numbering	Amino acid	Numbering	Amino acid	Numbering	Amino acid
4	N	1	M	2	N
5	L	8	R	3	A
6	F	11	D	13	L
7	A	12	K	16	P
9	L	14	D	19	L
10	F	15	D	23	T
20	A	17	H	30	S
21	I	18	K	32	A
22	E	24	A	35	V
29	I	25	A	36	A
31	Y	26	G	37	R
34	L	27	D	40	R
38	A	28	K	43	N
39	G	33	E	49	G
41	V	46	V	53	G
42	A	47	A	54	D
44	V	48	R	55	R
45	L	51	Q	60	T
50	L	52	V	62	K
56	V	61	E	63	S
57	A	86	A	65	E
58	A	88	T	84	N
59	Q	89	L	85	T
64	V	90	H	87	Y
66	A	97	T	91	E
67	L	100	E	93	D
68	V	102	K	94	Y
69	L	108	P	98	D
70	Y	109	S	107	D
71	L	111	R	110	K
72	A	112	D	113	G
73	T	116	A	115	A
74	V	119	A	121	V
75	R	120	K	123	A
76	A	122	G	124	T
77	G	130	P	125	V
78	G	131	D	134	G
79	V	133	R	135	S
80	Y	138	D		

81	L	141	A	137	T
82	P	142	G	139	A
83	L	144	S	140	A
92	L	145	E	143	A
95	F	146	A	147	F
96	I	148	A	149	T
99	A	151	D	150	I
101	P	153	G	152	R
103	I	155	D	154	A
104	V	164	S (*)	156	D
105	V	165	G (*)	157	L
106	C	167	T (*)	162	Y
114	I	168	G (*)	163	T (*)
117	I	169	R (*)	170	S (*)
118	A	189	D	171	K (*)
126	E	192	R	174	M
127	T	194	T	178	D
128	L	195	P	182	S
129	G	196	D	185	L
132	G	207	H (*)	186	T
136	L	229	P	188	V
158	A	230	K	190	Y
159	A	232	D	198	V
160	I	234	D	204	P
161	L	235	K	206	Y
166	T (*)	238	D	220	F
172	G	241	A	222	R
173	A	242	R	224	S
175	L	251	T	231	F
176	S	255	R	233	P
177	H	258	Q	244	T
179	N	260	P	249	V
180	L	264	K	250	P
181	A	265	E	254	T
183	N	268	G	257	L
184	S	269	H	259	S
187	L	277	S (*)	261	R
191	W	278	A (*)	263	T
193	F	279	P (*)	266	T
197	D	283	D	271	R
199	L	286	R	281	L
200	I	287	E	282	A
201	H	289	S	294	H
202	A	290	A	298	E (*)
203	L	291	K	301	G (*)
205	I	293	G	302	M (*)

208	T (*)	295	A	303	T (*)
209	H	314	D	304	E (*)
210	G	315	G	312	P
211	L	316	D	313	Y
212	F	317	R	318	V
213	V	320	G	319	P
214	A	327	P	322	V (*)
215	S	328	G	324	P
216	N	330	S	325	A
217	V	336	P	329	V
218	T	337	E	331	A
219	L	338	T	332	R
221	A	340	K	335	D
223	G	341	E	339	G
225	M	343	P	342	L
226	I	344	R	347	I
227	F	345	G	349	M
228	L	346	D	354	G
236	I	353	K	355	P
237	L	359	K	364	M
239	L	363	R	368	T
240	M	365	P	371	E
243	A	366	E	372	F
245	V	367	K	373	R
246	L	369	K	376	G
247	M	370	S	377	F
248	G	374	D	379	I (*)
252	F	375	D	381	G
253	Y	385	K	382	D (*)
256	L	388	E	386	I
262	L	389	R	387	D
267	T	397	R (*)	390	G
270	M	399	K	393	H
272	L	403	I	394	I (*)
273	F	405	G	395	L
274	I	406	G	396	G (*)
275	S (*)	407	F	398	G
276	G (*)	408	N	400	D
280	L (*)	412	K	401	L
284	T	413	E	410	Y
285	H	416	S	415	E
288	W	417	E	419	D
292	T	420	A	421	M
296	V	422	P	423	G
297	L	425	V	426	E
299	R (*)	433	P	434	H

300	Y (*)	435	A	437	F
305	T	436	D	439	E
306	N	438	G	446	V
307	M	447	R	451	A
308	N	448	D	459	L
309	T	449	K	466	L
310	S	450	G	467	A
311	N	452	T	469	F
321	A	454	D	470	K
323	G	455	E	471	M
326	L	456	A	476	I
333	V	457	Q	477	F
334	T	460	H	478	V
348	G	461	G	482	P
350	I	463	D	484	N (*)
351	E	464	G	489	V
352	V	465	Q	490	Q (*)
356	N	468	K	491	K
357	V	473	K	495	R
358	F	474	K	498	Y
360	G	479	D	501	I
361	Y	480	D	502	Y
362	W	483	R		
378	F	485	T		
380	T	486	M		
383	L	487	G		
384	G	488	K (*)		
391	Y	492	N		
392	V	493	V		
402	V	496	E		
404	T	497	T		
409	V	499	K		
411	P	500	D		
414	I	503	K		
418	I				
424	V				
427	S				
428	A				
429	V				
430	I				
431	G				
432	V				
440	G				
441	V				
442	T				
443	A				

444 V
 445 V
 453 I
 458 V
 462 L
 472 P
 475 V
 481 L
 494 L

Buried Residues

A few conservative changes may be possible in the buried region. Suitably residues in the polypeptide part of an ATP sensor molecule of the invention have at least 90% sequence identity to RpMatB (SEQ ID NO:1). Suitably any differences are conservative substitutions. Suitably residues in the polypeptide part of an ATP sensor molecule of the invention have 100% similarity to RpMatB (SEQ ID NO:1). More suitably residues shown as 'buried' in table A are not mutated. Thus suitably residues in the polypeptide part of an ATP sensor molecule of the invention which correspond to RpMatB residues shown as 'buried' in table A are not mutated relative to RpMatB (SEQ ID NO:1). In other words, suitably residues in the polypeptide part of an ATP sensor molecule of the invention shown as 'buried' in table A comprise the same residue as at the corresponding position in RpMatB (SEQ ID NO:1).

Thus in some embodiments, the polypeptide component of the ATP sensor molecule of the invention suitably comprises amino acid sequence having 100% sequence identity to those residues shown as 'buried' in table A.

Intermediate Residues

Suitably residues shown as 'intermediate' in table A may be mutated. Thus suitably residues in the polypeptide part of an ATP sensor molecule of the invention which correspond to RpMatB residues shown as 'intermediate' in table A may be mutated relative to RpMatB (SEQ ID NO:1). In other words, suitably residues in the polypeptide part of an ATP sensor molecule of the invention shown as 'intermediate' in table A may comprise a different residue (or no residue) from the corresponding position in RpMatB (SEQ ID NO:1).

Suitably a biosensor of the invention has at least 60% sequence identity to SEQ ID NO: 1. In some embodiments, the polypeptide component of the ATP sensor of the

invention suitably comprises amino acid sequence having at least 60% sequence identity to those residues shown as 'intermediate' in table A. In some embodiments, the polypeptide component of the ATP sensor of the invention suitably comprises amino acid sequence having at least 68% sequence identity to those residues shown as

5 'intermediate' in table A, suitably least 70% sequence identity, suitably least 74% sequence identity, suitably least 78% sequence identity, suitably least 82% sequence identity, suitably least 86% sequence identity, suitably least 90% sequence identity, suitably least 94% sequence identity, suitably least 98% sequence identity to those residues shown as 'intermediate' in table A.

10

However, it should be noted that these 'intermediate' residues are in fact partially buried. Thus, suitably these 'intermediate' residues are only mutated by substitution with a conservative residue relative to RpMatB. In other words, suitably the non-identical residues noted above comprise only conservative substitutions relative to the

15 corresponding residue in RpMatB.

20

Unless otherwise apparent from the text, mentions of 'sequence homology' suitably refer to sequence identity. However, occasionally it is more appropriate to refer to sequence similarity. Sequence similarity takes account of sequence identity and also takes account of conservative substitutions (i.e. non-identical residues but where the residue is similar or conserved compared to the original residue). Assessing sequence similarity is well known in the art. Examples are provided in the examples section. In case any further guidance is needed, suitably the following parameters are used in the algorithm for calculating sequence similarity: BLOSUM62 matrix, gap penalty 10.0,

25 gapextend penalty 0.5; more suitably BLOSUM62 matrix, gapopen 10.0, gapextend 0.5, endopen 10.0, endextend 0.5, pairwise alignment. Suitably the polypeptide component of the ATP sensor of the invention has at least 34% sequence similarity to SEQ ID NO: 1, suitably at least 38% sequence similarity, suitably at least 40% sequence similarity, suitably at least 50% sequence similarity, suitably at least 60% sequence similarity,

30 suitably at least 65% sequence similarity, suitably at least 70% sequence similarity, suitably at least 75% sequence similarity, suitably at least 80% sequence similarity, suitably at least 85% sequence similarity, suitably at least 90% sequence similarity, suitably at least 95% sequence similarity, suitably at least 98% sequence similarity, suitably at least 99% sequence similarity, most suitably 100% similarity to SEQ ID NO:

35 1.

Conservative substitutions may be made for example according to the table below.

Amino acids in the same block in the second column and preferably in the same line in the third column may be substituted for each other:

ALIPHATIC	Non-polar	G A P
		I L V
	Polar - uncharged	C S T M
		N Q
	Polar - charged	D E
		K R
AROMATIC		H F W Y

More suitably residues shown as 'intermediate' in table A are not mutated. Thus
 5 suitably residues in the polypeptide part of an ATP sensor of the invention which
 correspond to residues shown as 'intermediate' in table A are not mutated relative to
 RpMatB (SEQ ID NO:1). In other words, suitably residues in the polypeptide part of an
 ATP sensor of the invention shown as 'intermediate' in table A comprise the same
 residue as at the corresponding position in RpMatB (SEQ ID NO:1).

10

Exposed Residues

Suitably residues shown as 'exposed' in table A (i.e. column II of table A) may be
 mutated. Thus suitably residues in the polypeptide part of an ATP sensor of the
 invention which correspond to residues shown as 'exposed' in table A may be mutated
 15 relative to RpMatB (SEQ ID NO:1). In other words, suitably residues in the polypeptide
 part of an ATP sensor of the invention shown as 'exposed' in table A may comprise a
 different residue (or no residue) from the corresponding position in RpMatB (SEQ ID
 NO:1).

20 In some embodiments, the polypeptide component of the ATP sensor of the invention
 suitably comprises amino acid sequence having at least 21% sequence identity to those
 residues shown as 'exposed' in table A, suitably at least 30% sequence identity, suitably
 at least 40% sequence identity, suitably at least 50% sequence identity, suitably at least
 60% sequence identity, suitably at least 65% sequence identity, suitably at least 70%
 25 sequence identity, suitably at least 75% sequence identity, suitably at least 80%
 sequence identity, suitably at least 85% sequence identity, suitably at least 90%
 sequence identity, suitably at least 95% sequence identity, suitably at least 98%

sequence identity, suitably at least 99% sequence identity, most suitably 100% identity to those residues shown as 'exposed' in table A.

AMINO ACID RESIDUES

5

For dye conjugation we teach specific examples of specific amino acid residues for dye attachment for the sensor molecules of the invention. Details of these are provided herein.

10 It may also be possible to use other residues in the vicinity of those specified for dye attachment. By "in the vicinity" we mean a near neighbour of the specified amino acid. A near neighbour of the specified amino acid may mean an adjacent amino acid i.e. the amino acid before or the amino acid after the one specified.

15 Alternatively, neighbouring amino acid may be used. A neighbouring amino acid may refer to an amino acid two residues either side of the residue specified, for example if the residue specified is 398 then residues 400 and 396 would be considered neighbouring amino acids. For example, if amino acid 398 was specified, then residues 399 and 397 would be considered near neighbours. Alternatively, a further neighbour
20 of the amino acid may be specified, for example, if amino acid 398 is taught for attachment then a further neighbour might be three amino acid residues away such as residue 401 or residue 395. More distant amino acids may be used if desired. In all cases, it is advisable to check the performance of the sensor using the assays as taught herein.

25

More suitably, an amino acid which is "in the vicinity" of a specified amino acid is one which is present in a physically adjacent 3-dimensional space. For example, an amino acid specified on a part of an α helix will have a neighbouring amino acid in the vicinity such as the residue at the same position on the next turn of that α helix. For example,
30 amino acid 457 is on one side of an α helical section of the protein; thus amino acid 461 is a neighbouring amino acid in the vicinity of amino acid 457 since it is on the corresponding side of the next turn of the same α helix. Thus, when considering the 3-dimensional structure of the protein, an amino acid which might seem "distant" in terms of the number of intervening residues may actually be a neighbouring amino acid
35 in the vicinity of the specified amino acid if it is close in 3-dimensional space.

Amino acid residues close in space to those exemplified for labelling (see below and in Examples section) may also be suitable for labelling. Examples of surface accessible residues & neighbouring amino acids are now discussed to aid understanding.

5

Based on the exemplified cysteine substitution (mutation) sites, Table 2 below shows examples of amino acids which are neighbouring in sequence or in space. In the latter case, this represents amino acids in a similar location on the next turn of a helix, on the same side of a beta-sheet or on an adjacent sheet of a

10

Table 2: The neighbouring residues in sequence and space for the amino acid mutations of the diethylaminocoumarin (MDCC and IDCC) and tetramethylrhodamine (5-ATR and 6-ATR) labeled RpMatB cysteine mutant sensor molecules:

15

Amino acid mutation	Neighbouring residues in sequence and space
R286C	A282, D283, H285, E287, S289 and A290
K385C	L383, G384, I386 and D287
E439C	I403, P433, G438 and G440
Q457C	A456, V458, H460 and G461
G461C	Q457, H460, L462, G464 and Q465
K470C	L466, F469 and M471
N492C	K491, V493, R495, E496 and T497

These are further shown within the sequence of MatB and in structural representations below. Suitably the or each reporter moiety (such as fluorophore) is attached to the polypeptide via an amino acid residue

20

Exemplary Sensors of the Invention

The following variants provide a good signal (see Examples & Table 1), with a minimum of 60% increase in fluorescence between apo and ATP-bound form (relative to which of these forms had lower fluorescence):

5 **Single Reporter (fluorophore):**

MDCC-His6-RpMatB C106A/K488A/N492C

IDCC-His6-RpMatB C106A/K488A/N492C

MDCC-His6-RpMatB C106A/E439C/K488A

10 **Dual Reporter (double fluorophore):**

(6-ATR)2-His6-RpMatB C106A/R286C/G461C/K488A

(6-ATR)2-His6-RpMatB C106A/K385C/K470C/K488A

(6-ATR)2-His6-RpMatB C106A/R286C/Q457C/K488A

(5-ATR)2-His6-RpMatB C106A/R286C/Q457C/K488A

15

Exemplary Sensor Sequences based on RpMat

(Substitutions (Mutations) shown in bold)

20

Single Label (Single cysteine):

RpMatB C106A/K488A/N492C **SEQ ID NO: 2**

25 MNANL FARLFDKLD DPHKLA IETAAGDKISY AELVARAGRVANVLVARGLQVGDRVAAQTEKSVEALVLY
LATVRAGGVYLP LNTAYTLHELDYFITDAEPKIVV**AD**PSKRDGIAAIAAKVGATVETLGP DGRGSLTDAA
AGASEAFATIDRGADDLAAILY TSGTTGRSKGAMLSHDNLA SDSLTLVDYWRFTPDDVLIHALPIYH THG
LFVASNVT L FARGSMIFLPKFDPDKILDLMARATVLMGVPTFYTRLLQSPRLTKETTGHMRLFISGSAPL
LADTHREWSAKTGHAVLERYGMTETNMNTSNPYD GDRVPGAVGPALPGVSARVTD PETGKELPRGDIGMI
30 EVKGP NVFKGYWRMPEKTKSEFRDDGFFITGDLGKIDERGYVHILGRGKDLVITGGFNVYPKEIESEIDA
MPGVVESAVIGVPHADFGEGVTAVVVRDKGATIDEAQVLHGLDGQLAKFKMPKKVIFVDDLPRNTMG**AVQ**
KCVLRETYKDIYK

RpMatB C106A/E439C/K488A **SEQ ID NO: 3**

35 MNANL FARLFDKLD DPHKLA IETAAGDKISY AELVARAGRVANVLVARGLQVGDRVAAQTEKSVEALVLY
LATVRAGGVYLP LNTAYTLHELDYFITDAEPKIVV**AD**PSKRDGIAAIAAKVGATVETLGP DGRGSLTDAA
AGASEAFATIDRGADDLAAILY TSGTTGRSKGAMLSHDNLA SDSLTLVDYWRFTPDDVLIHALPIYH THG
LFVASNVT L FARGSMIFLPKFDPDKILDLMARATVLMGVPTFYTRLLQSPRLTKETTGHMRLFISGSAPL
40 LADTHREWSAKTGHAVLERYGMTETNMNTSNPYD GDRVPGAVGPALPGVSARVTD PETGKELPRGDIGMI
EVKGP NVFKGYWRMPEKTKSEFRDDGFFITGDLGKIDERGYVHILGRGKDLVITGGFNVYPKEIESEIDA
MPGVVESAVIGVPHADFG**CG**VTAVVVRDKGATIDEAQVLHGLDGQLAKFKMPKKVIFVDDLPRNTMG**AVQ**
KNVLR ETYKDIYK

45 **Dual Label (Two cysteines):**

RpMatB C106A/R286C/G461C/K488A SEQ ID NO: 4

MNANLFARLFDKLDDPHKLAIETAAGDKISYAELVARAGRVANVLVARGLQVGDRVAAQTEKSVEALVLY
 LATVRAGGVYLPNTAYTLHELDYFITDAEPKIVV**ADPS**KRDGIAAIAAKVGATVETLGPDGRGSLTDAA
 5 AGASEAFATIDRGADDLAAILYTS GTTGRSKGAMLSHDNLA NSLTLVDYWRFTPDDVLIHALPIYHTHG
 LFVASNVTLFARGSMIFLPKFDPDKILDLMARATVLMGVPTFYTRLLQSPRLTKETTGHMRLFISGSAPL
 LADTH**CEWS**AKTGHAVLERYGMTETNMNTSNPYDGD RVP GAVGPALPGVSARVTD PETGKELPRGDIGMI
 EVKGPNVFKGYWRMPEKTKSEFRDDGFFITGDLGKIDERG YVHILGRGKDLVITGGFNVPK EIESEIDA
 10 MPGVVESAVIGVPHADFGEGVTAVVVRDKGATIDEA QVLH**CLD**GQLAKFKMPKKVIFVDDLPRNTMG**AVQ**
 KNLVRETYKDIYK

RpMatB C106A/K385C/K470C/K488A SEQ ID NO: 5

MNANLFARLFDKLDDPHKLAIETAAGDKISYAELVARAGRVANVLVARGLQVGDRVAAQTEKSVEALVLY
 15 LATVRAGGVYLPNTAYTLHELDYFITDAEPKIVV**ADPS**KRDGIAAIAAKVGATVETLGPDGRGSLTDAA
 AGASEAFATIDRGADDLAAILYTS GTTGRSKGAMLSHDNLA NSLTLVDYWRFTPDDVLIHALPIYHTHG
 LFVASNVTLFARGSMIFLPKFDPDKILDLMARATVLMGVPTFYTRLLQSPRLTKETTGHMRLFISGSAPL
 LADTHREWSAKTGHAVLERYGMTETNMNTSNPYDGD RVP GAVGPALPGVSARVTD PETGKELPRGDIGMI
 EVKGPNVFKGYWRMPEKTKSEFRDDGFFITGDLG**C**IDERG YVHILGRGKDLVITGGFNVPK EIESEIDA
 20 MPGVVESAVIGVPHADFGEGVTAVVVRDKGATIDEA QVLHGLDGQLAKF**CM**PKKVIFVDDLPRNTMG**AVQ**
 KNLVRETYKDIYK

RpMatB C106A/R286C/Q457C/K488A SEQ ID NO: 6

MNANLFARLFDKLDDPHKLAIETAAGDKISYAELVARAGRVANVLVARGLQVGDRVAAQTEKSVEALVLY
 25 LATVRAGGVYLPNTAYTLHELDYFITDAEPKIVV**ADPS**KRDGIAAIAAKVGATVETLGPDGRGSLTDAA
 AGASEAFATIDRGADDLAAILYTS GTTGRSKGAMLSHDNLA NSLTLVDYWRFTPDDVLIHALPIYHTHG
 LFVASNVTLFARGSMIFLPKFDPDKILDLMARATVLMGVPTFYTRLLQSPRLTKETTGHMRLFISGSAPL
 LADTH**CEWS**AKTGHAVLERYGMTETNMNTSNPYDGD RVP GAVGPALPGVSARVTD PETGKELPRGDIGMI
 30 EVKGPNVFKGYWRMPEKTKSEFRDDGFFITGDLGKIDERG YVHILGRGKDLVITGGFNVPK EIESEIDA
 MPGVVESAVIGVPHADFGEGVTAVVVRDKGATIDEA **CVL**HGLDGQLAKFKMPKKVIFVDDLPRNTMG**AVQ**
 KNLVRETYKDIYK

As will be clear from the text, these exemplary sequences may have additions/deletions
 35 (e.g. N- or C- terminal truncations) or other substitutions as described.

SEQ ID NO: 7 (see Figure 17) shows exemplary sequence annotated to illustrate
 exemplary substitutions which may be used in the invention.

40 Suitably these exemplary sequences may further comprise an N-terminal addition
 MSYYHHHHHH DYDIPTSENLYFQGAS (**SEQ ID NO: 8**) added directly before the
 first Methionine of the sequences above. This N-terminal addition comprises a 6His
 tag useful in purification.

45 **TRUNCATIONS/INSERTIONS/DELETIONS**

In a broad aspect the invention relates to an ANL superfamily polypeptide for use as an
 ATP sensor molecule or ATP binding molecule. Such an ANL superfamily polypeptide
 may be full length (i.e. comprising all 503 amino acid residues corresponding to SEQ

ID NO:1 (whether or not substitutions relative to SEQ ID NO:1 are made in the particular amino acids present)) or truncated. Suitably truncated forms are those which lack a small number of amino acid residues from the N- or C- terminus of the polypeptide relative to wild type. Suitably a small number is 10 or fewer.

5

In more detail, ANL superfamily proteins such as RpMatB have flexible C- and N-terminal ends. We teach that small truncations might be made at either or both ends. Suitably some or all of the amino acids from these flexible sections may be deleted without adversely affecting the remaining structure and hence retaining sensor

10

function.

In more detail, N-terminal amino acids which may be deleted include those corresponding to M1, N2, A3, N4 of SEQ ID NO: 1.

15

In more detail, C-terminal amino acids which may be deleted include those corresponding to E498, K499, D500, I501, Y502, K503 of SEQ ID NO: 1.

In case it is necessary to check the sensor function after such deletion(s), the skilled worker may simply test the sensor as outlined herein.

20

Similarly, small insertions or deletions may be made in the protein without disrupting function, in particular insertions or deletions are suitably not made in the region of the reporter/dye attachment points, nor in the ATP binding section of the protein. Suitably insertions or deletions are suitably not made in secondary structure elements of the polypeptide such as alpha helices or beta sheets. Suitably insertions/deletions are not made in dual labelled polypeptides in the section of the polypeptide between the dye attachment points, so as to preserve the dye spacing.

25

TAGS

30

It is often useful to tag proteins of the invention, for example to facilitate their purification after recombinant production.

Suitably tags may be placed at the extreme C-terminus or the extreme N-terminus of the sensor molecule.

35

Most suitably tags are placed at the N-terminus of the protein.

Suitably one tag per protein molecule is used. Multiple tags per protein molecule may be used if desired, including multiple copies of the same tag or two or more different tags, for example it may be desirable to use a 6 His-tag for purification and in addition
5 to use a Myc-tag for detection.

Tags may be removed from the sensor protein, for example by proteolytic cleavage, or may be retained on the sensor protein during use.

10 A hexahistidine tag (6his) may be added to the polypeptide part of the ATP binding molecule of the invention to simplify purification; most suitably a N-terminal hexahistidine tag is used. 6 His is a particularly useful tag for purification on nickel substrates. However, any suitable tag known in the art may be used.

15 Alternatively, the sensor molecule of the invention may be tagless. Tagless purification (if needed) is well known in the art.

BACKBONE MUTATIONS

20 ANL superfamily proteins may have catalytic activity. For example, RpMatB may consume ATP when carrying out its enzymatic activity. However, RpMatB has no significant ATPase activity (forming ADP and Pi). The reaction for RpMatB is:

-ATP + malonate to malonyl-AMP + PPi

-Malonyl-AMP + CoASH to Malonyl-CoA and AMP.

25

It is possible to use sensors of the invention with an intact catalytic activity. In this case, it is important that no substrates of the enzyme are present in the assay mixture. The reason is that if substrates are present, this will drive ATP consumption and may distort the data provided by the sensor. For this reason, suitably the sensor molecules
30 of the invention are catalytically inactive. Suitably the sensor molecules of the invention do not comprise ATPase activity. Suitably the sensor molecules of the invention are catalytically inactive for ATP consumption.

The sensor molecule may be catalytically inactivated by mutation of the active site. For
35 example, suitably the sensor of the invention is mutated to render it catalytically inactive. When the sensor comprises an ANL superfamily member such as RpMatB, suitably a K488 mutation is present. In other words, suitably the amino acid

corresponding to the wild type K488 is changed to any amino acid other than K. Suitably K488V may be used. Suitably K488A may be used. Thus, suitably the sensor of the invention comprises K488X¹, wherein X¹ is not K. Suitably the sensor of the invention comprises K488V. Most suitably the sensor of the invention comprises K488A.

The conjugation of dye to the polypeptide of interest may target naturally occurring cysteine residues. To this end, it may be useful to remove unwanted cysteine residues by a process of mutation in order to eliminate background signal. When the sensor of the invention comprises an ANL superfamily member such as RpMatB, suitably cysteine 106 of the wild type protein is mutated to be other than cysteine. For example, suitably the sensor of the invention comprises a C106X² mutation, wherein X² is not C. Suitably the sensor of the invention comprises a C106A mutation. Background signal from C106 may be as high as 6%. Therefore, C106X² (where X² is not C) mutants such as C106A provide the advantage of eliminating this background signal.

A sensor which is wild type at position 106 i.e. comprising C106 may still be used – in this case the values collected should be adjusted for any background signal as necessary.

20

MANUFACTURE AND PREPARATION

The polypeptide components of the ATP sensor molecules of the invention may be produced by standard recombinant techniques, such as creating a nucleic acid encoding the amino acid sequence of the polypeptide, and then expressing the polypeptide in a host such as *E.coli*. Alternatively an *in vitro* translation may be used. Alternatively the polypeptide itself may be chemically synthesised.

The polypeptide(s) may be purified by any suitable method known in the art, such as 6His tagging the protein then purification using Ni-NTA beads.

To introduce the desired amino acid substitutions into the polypeptide of the invention, any suitable technique may be used such as site directed mutagenesis. For example, mutant PCR primers or oligonucleotides containing the desired nucleotide sequence may be annealed to a template and ligated, extended or amplified to produce a mutated nucleotide sequence encoding the desired substitution. Alternatively the desired nucleotide sequence may be synthesised chemically.

Exemplary techniques are presented in the Examples below.

REPORTER MOIETIES / DYES

5

Suitably the ATP binding molecules of the invention comprise at least one reporter moiety attached thereto. The reporter moiety may be any suitable chemical group or structure capable of reading out change(s) in the conformation of said ATP binding molecule. Most suitably the reporter moiety comprises one or more fluorophore(s) such as coumarin or rhodamine.

10

Reporter moieties used in the invention can give various signals, but preferred 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. Most suitably fluorescent labels which may be excited to fluoresce upon exposure to certain wavelengths of light are used. The fluorescent label can be selected from the group consisting of rhodamines, cyanines, pyrenes and derivatives thereof. Preferred fluorescent fluorophores are based on a xanthene nucleus, which can readily undergo stacking to form dimers. Especially suitable are rhodamine fluorophores.

15

20

Suitably the reporter moiety comprises any usable fluorescent label. Fluorescent labels with environmentally sensitive fluorescence are most suitable. When a cysteine is the site of attachment, then the moiety needs thiol-reactivity for attachment. In other embodiments, an amine-sensitive label on a non-Cys amino acid may be employed.

25

In some embodiments, reporter moieties may be those that can exhibit molecular stacking, which will thus include aromatic rings. These include the rhodamine labels. In other embodiments labels which do not stack may be used, such as coumarin labels.

30

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

35

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').

5

For example, rhodamine chromophores can form dimers at high concentrations in solution. The dimer has a different absorbance spectrum from the monomer, and has little or no fluorescence in comparison with the monomer. Two rhodamine chromophores attached to suitable positions in the protein can form dimers, whose
10 interaction is altered when ligand binds to the protein. The invention can spectroscopically detect the difference between the ATP-free and ATP-bound conformations of ATP binding molecule. Molecular stacking takes place through the physical interaction of ground states of the two moieties. Labels that can undergo molecular stacking are well known in the art. Stacking can occur between identical
15 chromophores, and can also occur between different chromophores.

Suitably the reporter moiety is a dye. Suitably the reporter moiety is or comprises a fluorophore. Suitably said fluorophore is attached at a position on the polypeptide such that conformational change of the polypeptide upon ATP binding causes a
20 corresponding change in fluorescence of said fluorophore.

We describe ATP sensor molecules with single labelling. Most suitably, when the ATP sensor is singly labelled a coumarin type dye is used as the label.

25 We describe ATP sensors having double labelling. Double labelling means two dye molecules per sensor molecule. Suitably each of the two dye molecules is attached to a separate amino acid residue on the sensor molecule. For doubly labelled ATP sensors. Suitably the dyes are rhodamine type dyes.

30 For dual labelling, any stacking rhodamine type dye is useful.

Cy dyes may be useful in the invention. However, these might require different attachment points from those taught for rhodamine attachment due to different dye molecule sizes. In other words, the dye molecules may need to be placed closer
35 together or further apart in space than the corresponding rhodamines. In some embodiments, suitably Cy dyes are not used, the reason is that Cy dyes can suffer from the drawback of tending to provide high fluorescence for much of the time. This can

make it more difficult to observe the fluorescence changes which are useful in the invention.

Acrylodan dyes may be useful in the invention.

5

Dual Labelling

More specifically, for the two fluorophore stacking quenching (i.e. dual labelling as used for rho-MatB with two rhodamines):

10

In principle any rhodamine dye that can be specifically linked to surface thiols could be used. This includes the 5- and 6-isomers of tetramethylrhodamine.

15

Rhodamine dyes are available in different isomers. For example, 5-tetramethylrhodamine (5-ATR) and/or 6-ATR may be used to label ATP sensors of the invention. It is possible to use a mixture of rhodamine isomers to label the sensor of the invention. This may result in a heterogeneous population of sensor molecules. For example, if a mixture of 5-ATR and 6-ATR dye is used to label the sensor protein, the following species will be generated –

20

- 5-ATR, 5-ATR
- 5-ATR, 6-ATR
- 6-ATR, 5-ATR
- 6-ATR, 6-ATR

25

The sensor of the invention may comprise any of these species of molecules. The sensor of the invention may comprise a mixture of more than one of these sensor molecules. The sensor of the invention may comprise a mixture of all four of these labelled sensor molecules.

30

Without wishing to be bound by theory, it may be technically possible to find a difference in the absorption/emission behaviour of sensors according to the invention having different isomers of dye attached. However, in practice the use of a mixture of sensors having alternate isomer dyes attached performs well in the assays of the invention.

35

Single Labelling

For single labelled sensors of the invention, any suitable dye may be used, most suitably a dye which provides a coplanar alignment when attached to the sensor protein. Other
 5 conformations are possible, for example a twisted conformation, but those often tend to provide a lower fluorescence. A lower fluorescence may be usable but would need to be checked on a case by case basis.

For singly labelled sensors for the invention, any coumarin type dye is expected to
 10 provide good results.

In principle, any fluorophore known to have an environmentally sensitive fluorescence may be useful in the invention.

15 In more detail, for single labelling, such as with a single environmentally sensitive dye, any coumarin may be used, including the iodoacetamide- and maleimide- linked diethylaminocoumarins (*N*-[2-(1-maleimidyl)ethyl]-7-diethylaminocoumarin-3-carboxamide and *N*-[2-(iodoacetamido)ethyl]-7-diethylaminocoumarin-3-carboxamide).

20 Other fluorophore types, known to have fluorescence intensity, depending on physical environment, such as interactions with protein surfaces, include

CPM (7-(diethylamino)-3-[4'-(1-maleimidyl)phenyl]-4-methylcoumarin)
 25 Alexafluor 350 – C5 maleimide
 Acrylodan (6-acryloyl-2-dimethylaminonaphthalene)
 Badan (6-bromoacetyl-2-dimethylaminonaphthalene)
 MANS (2-(4'-maleimidylanilino)naphthalene-6-sulfonic acid)
 IAEDANS (5-[2-[(2-Iodo-1-oxoethyl)amino]ethylamino]-1-naphthalenesulfonic acid)
 30 Alexa Fluor 488 maleimide
 Fluorescein-5-maleimide
 Cy3-maleimide (1-(6-{[2-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)ethyl]amino}-6-oxohexyl)-2-[(1E,3E)-3-(1-ethyl-3,3-dimethyl-5-sulfo-1,3-dihydro-2H-indol-2-ylidene)prop-1-enyl]-3,3-dimethyl-3H-indolium)
 35 Pyrene-maleimide

Conjugation Chemistry

Reporter moieties or labels such as fluorophores may be attached to the ATP binding molecule of the invention by any suitable means known in the art.

- 5 Suitable amino acid residues may be engineered into the polypeptide. Although the invention is exemplified using cysteines such as cysteine substitutions, and attachment of reporters (e.g. fluorophores) such as via maleimide or iodoacetamide, it should be borne in mind that alternate technologies such as peptide ligation (e.g. chemical synthesis of protein or attachment of synthetic peptides to other polypeptides), and/or
- 10 introduction of unnatural amino acids using mutated tRNAs and/or tRNA synthetases may be used. For example, amino acids having reactive azide groups may be introduced to take advantage of 'click chemistry' (or vice versa) or other conjugation techniques may be used. For example, lysine based unnatural amino acids may be introduced to achieve this goal (e.g. Nguyen et al 2009 (J Am Chem Soc. 2009 Jul
- 15 1;131(25):8720-1); e.g. Lang et al 2012 Nature Chemistry 4, 298–304 (2012)).

Suitably the fluorophores are attached by conventional conjugation techniques such as covalent attachment via a cysteine residue in the polypeptide component of the ATP binding molecule.

20

The covalent attachment of extrinsic reporter moieties 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)).

- 25 Reporter moieties can be attached via amines or carboxyl residues on amino acid side chains, but it is more suitable 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.
- 30 Where a cysteine residue has to be introduced, either by insertion or substitution, a number of factors should be considered. These are discussed in more detail herein. Exemplary sites for introduction of Cys residues and thus for label attachment are disclosed in detail herein.
- 35 If attached chromophores are to interact, the residues must be selected such that they are in proximity to each other, and that the conformational change that occurs on ATP-binding affects one or both of the residues to cause a change in position or orientation

or electronic environment of a label attached thereto. Exemplary pairs of attachment sites are set out above.

Exemplary attachment techniques are presented in the examples section.

5

ASSESSMENT OF FLUORESCENCE

A key concept of the invention is that the ATP binding molecules (sensor proteins) are configured so that they undergo a conformational change upon ATP binding. It is
10 detection of this conformational change which allows the ATP binding status of the molecules of the invention to be determined. Clearly, where the technology permits, single molecule conformational changes may be determined. However, for most embodiments of the invention, determination of the conformational status of the ATP binding molecule is suitably assessed as a population effect. In other words, assessing
15 the conformational change of an ATP binding molecule of the invention may be carried out by determining the conformational change of a population of ATP binding molecules of the invention. Considering those embodiments where the conformational change is read out or detected by monitoring changes in the behaviour of a reporter moiety coupled to the ATP binding molecule of the invention, this idea may be
20 explained as follows. For the purposes of illustration, the reporter moiety will be considered to be a fluorophore. The fluorophore is attached to the ATP binding molecule of the invention. ATP binding leads to a conformational change of the ATP binding molecule. This conformational change can lead to a change in fluorescence. This change in fluorescence may be an increase or a decrease upon ATP binding
25 depending on the particular labelling strategy used. For any given application having a fixed amount of sensor protein, the change in fluorescence will be consistently associated with the corresponding change in ATP binding. ATP binding is proportional to the concentration of ATP present in the sample being studied. Therefore, changes in ATP binding provide information about changes in the ATP concentration in the
30 sample being studied. Thus, changes in fluorescence which are catalysed by conformational changes in the ATP binding molecule of the invention brought about by ATP binding directly provide information about the concentration of ATP in the sample being studied. For the great majority of applications or embodiments of the invention, conformational changes will be detected for a population of ATP binding molecules
35 according to the invention. In practical terms, this means that a certain amount of the ATP binding molecule of the invention will be added to the sample being studied. The fluorescence of this population of ATP binding molecules of the invention will then be

monitored. Thus, changes in the fluorescence of these molecules represents an indication of the ATP binding status of a proportion or a population of those molecules present in the sample.

5 Thus, it will be clearly understood that the level of fluorescence varies directly with the amount of ATP binding of the ATP binding molecules of the invention. Thus, this will give the effect of an almost continuously variable level of fluorescence depending on the level of ATP binding (and thus the ATP concentration) within the sample. At one extreme, there will be no ATP binding to any of the ATP binding molecules of the
10 invention. At the other extreme, all of the ATP binding molecules of the invention present in the sample will each be bound to ATP such that there is saturation and complete (or near complete) binding and therefore a 100% signal. Thus, within these two extremes, the particular level of fluorescence is directly correlated to a particular concentration of ATP in the sample. Thus, in some embodiments of the invention, a
15 standard curve may be constructed by measuring the fluorescence of a constant amount of the ATP binding molecule of the invention in the presence of differing known concentrations of ATP. This standard curve may then be used in order to read out or convert measured fluorescence values to absolute concentrations of ATP present in a sample.

20

In another embodiment, the readout of the invention may be advantageously calibrated by inclusion of samples having known ATP concentrations in the analysis being undertaken. In these embodiments, the samples containing known concentrations of ATP may be regarded as “internal controls”. This permits accurate estimation of ATP
25 concentrations in experimental settings where reference to a standard curve is less appropriate, for example in complex reaction mixtures in which other components might perturb the readouts, or might not have been present during the construction of a standard curve, thereby making such comparisons potentially inappropriate.

30 It will be noted that some of the sensors provided herein show increasing fluorescence in the presence of ATP, and some show decreasing fluorescence in the presence of ATP. Either type of sensor is useful.

Advantageously, sensors showing increased fluorescence upon binding ATP are used.
35 These provide the advantage of avoiding confounding factors which might otherwise reduce fluorescence, for example photo-bleaching or other degradation of the dye.

Occasionally a “percentage change” is discussed in the context of the invention. For sensors which increase their fluorescence upon the ATP binding, the percentage change is the percentage increase in fluorescence from the unbound to the bound state. For those sensors which exhibit a fluorescence decrease upon ATP binding, the percentage change is calculated upwards i.e. taking the decreased fluorescence observed on binding and comparing that to the higher level of fluorescence observed in the absence of ATP gives a percentage increase. For example, a sensor whose fluorescence decreases from 1.0 to 0.6 upon ATP binding has a percentage fluorescence change of 67% (i.e. the difference in fluorescence intensity of 0.4 divided by the fluorescence on ATP binding of 0.6 equals 67% change).

Suitably sensors of the invention exhibit at least 50% fluorescence change upon ATP binding, more suitably 60%, more suitably 70%, more suitably 80%, more suitably 90%, more suitably 100% or even more (such as a multiple of fluorescence in the unbound state).

DEFINITIONS

The term ‘comprises’ (comprise, comprising) should be understood to have its normal meaning in the art, i.e. that the stated feature or group of features is included, but that the term does not exclude any other stated feature or group of features from also being present.

An ATP binding molecule is a molecule capable of binding ATP. Use of the term ATP binding molecule does not imply or require that ATP is present. ATP binding molecule means molecule capable of binding ATP.

Abbreviations:

Rhodopseudomonas palustris malonyl-coenzymeA synthetase (RpMatB) - Protein Data Bank (PDB) - 5-iodoacetamidotetramethylrhodamine (5-IATR) - 7-diethylamino-3-(((2-maleimidyl)ethyl)amino)carbonyl)coumarin (MDCC) - 7-diethylamino-3-(((2-iodoacetomido)ethyl)amino)carbonyl)coumarin (IDCC)) - size-exclusion chromatography coupled to multi-angle laser light scattering (SEC-MALLS) - nicotinamide adenine dinucleotide (NADH) - deoxyadenosine triphosphate (dATP) - adenosine 5'-(γ-thio)triphosphate (ATPγS) - adenosine 5'-(β,γ-imido)triphosphate (AMP-PNP)

SLU: 5'-O-[N-(DEHYDROLUCIFERYL)-SULFAMOYL] ADENOSINE

ADVANTAGES

It is an advantage that the sensor of the invention is faster than any existing ATP sensor molecule such as those based on polypeptides.

5

An existing ATP sensor molecule known as “ATeam” has a 30 fold higher affinity for ATP than for ADP. The sensors of the invention advantageously have approximately 67 fold higher affinity for ATP than for ADP. This is a significant advantage offered by sensors of the invention.

10

In more detail, in assay scenarios in which ATP is measured, ADP will almost always be present. Therefore, this difference in affinity for ATP compared to ADP provides a much better sensor than prior art sensors. The practical advantage is that sensors can tolerate much higher concentrations of ADP whilst still performing well as ATP sensors.

15

The exemplary biosensors shown have the advantage of larger signal.

Suitably the ATP binding molecule of the invention is, and may be used as, a reagentless biosensor.

20

The invention advantageously provides a reagentless biosensor for ATP.

The invention advantageously provides a fluorescent reagentless biosensor for ATP.

25

CO-FACTORS

Divalent cation such as divalent metal ion is required for ATP binding to the sensors of the invention.

30

Suitably one or more of Cd^{2+} (Cadmium), Mn^{2+} (manganese) and/or Mg^{2+} (Magnesium) should be present together with the sensor of the invention in use. Most suitably the divalent cation is Mg^{2+} (Magnesium).

35

The skilled person knows that cations are not provided in isolation, but are provided as a salt with corresponding anions in solution. A typical way of providing divalent cations such as Mg^{2+} is to add magnesium chloride (MgCl_2) to the reaction

mixture/buffer. However, any other acceptable salt of the above mentioned divalent metal ions may equally be used provided it does not compromise the action of the assay. This is easily tested as set out below in the example section whilst varying the divalent cation (i.e. the salt) which is incorporated into the assay.

5

In one embodiment the invention relates to a kit comprising an ATP sensor molecule as described above together with a source of divalent Magnesium ion (Mg^{2+}). Suitably the source of Mg^{2+} is Magnesium Chloride.

10 Most importantly, in the presence of divalent cation such as divalent metal ion e.g. Mg^{2+} , ATP will be present as MgATP. Suitably Mg^{2+} is present in excess over ATP in the assay of the invention. Suitably Mg^{2+} is present at 1mM or more in the assay of the invention. Suitably ATP is present as MgATP in the assay of the invention.

15 pH

The sensors of the invention have the advantage of being usable under a wide range of pH conditions.

20 Suitably the pH of the assay is in the range 6.0 to 9.0. More suitably the pH of the assay is in the range 7.0 to 7.5.

It is an advantage of the invention that the sensor molecules do not react to nucleotides other than ATP, nor to ATP analogues. In other words, it does not matter if these
25 chemical entities are also present in the assay mixture, the sensor has the advantage of only reacting to ATP and therefore the presence of these other molecular species does not perturb the assay of ATP concentration according to the invention.

FURTHER APPLICATIONS AND EMBODIMENTS

30

The invention provides a fluorescent, reagentless biosensor for ATP, suitably based on or derived from malonyl-coenzyme A synthetase.

Additional mutations may be made to the sensors as described above to acquire further
35 advantages.

Examples of mutations which may be made are shown in Table C (see example 13). Table C shows the characteristics bestowed on the sensor by making of the mutations

as shown. Therefore, depending on the application(s) of the biosensor intended by the skilled worker, mutations from Table C should be chosen accordingly, or not made, as desired. In Table C, 'Parent' means exemplary Rho-MatB; however, the mutations in Table C are equally applicable to any other biosensor disclosed herein without requiring the specific set of mutations of Rho-MatB - these additional mutations of Table C are merely exemplified in the Rho-MatB background to illustrate the advantages and help understand the effects of the mutations which are described individually and may be applied to biosensors of the invention individually or in combination. Suitably mutations disclosed in Table C are applied individually. Suitably the invention provides a biosensor as described above further comprising one further mutation selected from those disclosed in Table C.

In more detail, especially useful mutations are discussed more fully below.

15 T167

Mutating T167 provides the advantage of increased K_d for ATP whilst retaining excellent fluorescence ratio (F⁺/F⁻). This has the further technical benefit of widening the range of ATP concentrations that can be measured by the biosensor.

Suitably T167 is mutated to a small amino acid.

20 Suitably T167 is mutated to Alanine (A), Serine (S) or Glycine (G).

Glycine can make the chain flexible which may or may not be desired. Thus suitably T167 is mutated to Alanine (A) or Serine (S), avoiding flexibility effect(s) of Glycine.

Thus suitably said polypeptide comprises an amino acid other than T at the position corresponding to 167 of SEQ ID NO: 1. Thus suitably said polypeptide comprises Alanine (A), Serine (S) or Glycine (G) at the position corresponding to 167 of SEQ ID NO: 1. More suitably said polypeptide comprises Alanine (A) or Serine (S) at the position corresponding to 167 of SEQ ID NO: 1. More suitably said polypeptide comprises Serine (S) at the position corresponding to 167 of SEQ ID NO: 1, which has the advantage of enhanced fluorescence ratio and enhanced K_d compared to Rho-MatB. Most suitably said polypeptide comprises Alanine (A) at the position corresponding to 167 of SEQ ID NO: 1, which has the advantage of greatly enhanced K_d whilst retaining the same advantageous fluorescence ratio as Rho-MatB.

S170

35 Mutating S170 provides the advantage of increased fluorescence ratio (F⁺/F⁻) whilst retaining excellent K_d for ATP. This has the further technical benefit of a decreased

ADP affinity, thereby increasing selectivity of ATP over ADP to >200 fold. This is especially advantageous in measuring ATP when ADP is also present.

Suitably S170 is mutated to a small amino acid.

Suitably S170 is mutated to Alanine (A) or Glycine (G).

- 5 Glycine can make the chain flexible which may or may not be desired. Thus suitably S170 is mutated to Alanine (A), avoiding flexibility effect(s) of Glycine.

Thus suitably said polypeptide comprises an amino acid other than S at the position corresponding to 170 of SEQ ID NO: 1. Thus suitably said polypeptide comprises Alanine (A), or Glycine (G) at the position corresponding to 170 of SEQ ID NO: 1. More

- 10 suitably said polypeptide comprises Alanine (A) at the position corresponding to 170 of SEQ ID NO: 1. Most suitably said polypeptide comprises Alanine (A) at the position corresponding to 170 of SEQ ID NO: 1, which has the advantage of greatly increased fluorescence ratio (F+/F-) compared to Rho-MatB.

15 T303

Mutating T303 provides the advantage of increased Kd for ATP whilst retaining excellent fluorescence ratio (F+/F-). This has the further technical benefit of widening the range of ATP concentrations that can be measured by the biosensor.

Suitably T303 is mutated to a small amino acid.

- 20 Suitably T303 is mutated to Alanine (A), Serine (S) or Glycine (G).

Glycine can make the chain flexible which may or may not be desired. Thus suitably T303 is mutated to Alanine (A) or Serine (S), avoiding flexibility effect(s) of Glycine.

Thus suitably said polypeptide comprises an amino acid other than T at the position corresponding to 303 of SEQ ID NO: 1. Thus suitably said polypeptide comprises

- 25 Alanine (A), Serine (S) or Glycine (G) at the position corresponding to 303 of SEQ ID NO: 1. More suitably said polypeptide comprises Alanine (A) or Serine (S) at the position corresponding to 303 of SEQ ID NO: 1. More suitably said polypeptide comprises Serine (S) at the position corresponding to 303 of SEQ ID NO: 1, which has the advantage of enhanced Kd compared to Rho-MatB. Most suitably said polypeptide
- 30 comprises Alanine (A) at the position corresponding to 303 of SEQ ID NO: 1, which has the advantage of enhanced Kd whilst also retaining an advantageous high fluorescence ratio comparable to that of Rho-MatB.

- Suitably positions corresponding to T167 and T303 of SEQ ID NO: 1 are not both
- 35 mutated in the same sensor molecule. Without wishing to be bound by theory, it is possible that double mutation of positions 167 and 303 may weaken ATP binding in such a double mutant. Thus suitably if a position corresponding to T167 is mutated,

then the position corresponding to T303 is wild type, suitably T. Thus suitably if a position corresponding to T303 is mutated, then the position corresponding to T167 is wild type, suitably T.

- 5 Most suitably the additional mutations described are made in conjunction with the mutations of RhoMatB such as shown in SEQ ID NO: 6. Suitably the dyes and any other additions to the sensor molecule are as for RhoMatB.

INDUSTRIAL APPLICATION

10

The invention finds use as a reagent in a research setting.

The invention finds application in the assessment of food contamination, such as bacterial contamination.

15

The invention finds application in drug screening methodology.

The invention finds application in any setting in which it is desired to assay ATP.

- 20 The invention finds application in the assessment of surface contamination such as in a clinical/hospital/medical environment.

The invention finds application in the assessment of surface contamination such as in food preparation areas.

25

Further particular and preferred aspects are set out in the accompanying independent and dependent claims. Features of the dependent claims may be combined with features of the independent claims as appropriate, and in combinations other than those explicitly set out in the claims.

30

Where an apparatus feature is described as being operable to provide a function, it will be appreciated that this includes an apparatus feature which provides that function or which is adapted or configured to provide that function.

BRIEF DESCRIPTION OF THE DRAWINGS

35

Embodiments of the present invention will now be described further, with reference to the accompanying drawings, in which:

Figure 1: C-terminal domain rotation upon MgATP binding and position of mutations in RpMatB. (A) Shown is a cartoon representation of RpMatB. The N-terminal domain (amino acids 1-399) in the apo conformation (PDB 4FUQ (Crosby, Rank et al. 2012)) is coloured grey. The C-terminal domain (amino acids 400-503) in the apo conformation is shown in pink. The C-terminal domain in the MgATP-bound conformation (PDB 4FUT (Crosby, Rank et al. 2012), after superimposing the N-terminal domain of 4FUT on the N-terminal domain of 4FUQ) is shown in green. ATP (in ball and stick conformation and coloured by CPK convention) and Mg^{2+} (depicted as an orange sphere) are shown in the active site. The C-terminal domain rotation upon MgATP-binding is clearly visible. The positions of mutations are shown as spheres and labeled: C106A in the N-terminal domain (yellow), K488A in the C-terminal domain (red), R286C in the N-terminal domain and Q457C in the C-terminal domain (both blue). (B) As described above but the C-terminal domain in the malonyl-coenzymeA and AMP bound conformation of *Streptomyces coelicolor* MatB (PDB 3NYR (Hughes and Keatinge-Clay 2011), after superimposing the N-terminal domain of 3NYR on the N-terminal domain of 4FUQ) is shown in blue. Malonyl-coenzymeA and AMP (in ball and stick conformation and coloured by CPK convention) are shown in the active site.

Figure 2: Characterization of the fluorescence and absorbance spectra of Rho-MatB. (A) Fluorescence excitation and emission spectra of 1 μM Rho-MatB in 50 mM Hepes pH 7.0, 100 mM NaCl, 10 mM $MgCl_2$ and 0.3 mg ml^{-1} bovine serum albumin in the absence and the presence of 175 μM ATP. Excitation was at 553 nm for the emission spectra. Emission was measured at 575 for the excitation spectra. Data were corrected for dilution. (B) Absorbance spectra of 1 μM Rho-MatB in 50 mM Hepes pH 7.0, 100 mM NaCl, 10 mM $MgCl_2$, 0.3 mg ml^{-1} bovine serum albumin in the absence and presence of 119 μM ATP. Data were corrected for dilution.

Figure 3: Characterization of nucleotide binding affinity to Rho-MatB. (A) Titration of ATP (open circles), ADP (solid circles), dATP (open triangles), ATP γ S (solid triangles) and AMP-PNP (open squares) to 0.5 μM Rho-MatB in 50 mM Hepes pH 7.0, 100 mM NaCl, 10 mM $MgCl_2$, 0.3 mg ml^{-1} bovine serum albumin. Aliquots of ligand were added and the fluorescence intensity was measured at 571 nm (exciting at 553 nm) at 20 °C. The data were corrected for dilution and normalized to 1 for the fluorescence intensity of Rho-MatB in the absence of ligand. The dissociation constants were obtained using a quadratic binding curve (see Materials and methods). Shown here is one representative experiment. The dissociation constants listed in Table 1 are the average of at least three repeat experiments. (B) 1 μM Rho-MatB in the absence (open circles) and presence of 100 μM ADP (solid circles), 100 μM dATP (open triangles), 100 μM ATP γ S (solid triangles) or 100 μM AMP-PNP (open squares) in 50 mM Hepes pH 7.0, 100 mM NaCl, 10 mM $MgCl_2$ and 0.3 mg ml^{-1} bovine serum albumin was titrated with an ATP / ADP, dATP, ATP γ S or AMP-PNP / Rho-MatB solution with ADP, dATP, ATP γ S or AMP-PNP and Rho-MatB of the same concentration as ADP, dATP, ATP γ S or AMP-PNP and Rho-MatB in the cuvette. That means that [ADP], [dATP], [ATP γ S] or [AMP-PNP] and [Rho-MatB] were held constant. The

fluorescence intensity was measured at 571 nm (exciting at 553 nm) at 20 °C. The data were normalized to 1 for the fluorescence intensity of Rho-MatB in the absence of ATP. The dissociation constants for ATP were obtained using a quadratic binding curve (see Materials and methods). Shown here is one representative experiment. The dissociation constants listed in Table 1 are the average of two repeat experiments. (C) 2.5 μM Rho-MatB in the absence (open circles) and presence of 10 μM ADP (solid circles), 50 μM ADP (solid triangles) or 100 μM ADP (solid squares) in 50 mM Hepes pH 7.0, 100 mM NaCl, 10 mM MgCl₂ and 0.3 mg ml⁻¹ bovine serum albumin was titrated with an ATP / Rho-MatB solution with ATP and Rho-MatB of the same concentration as ADP and Rho-MatB in the cuvette. That means that the total nucleotide concentration ([ATP] + [ADP]) and [Rho-MatB] was held constant and only the proportion of ATP was changed. The fluorescence intensity was measured at 571 nm (exciting at 553 nm) at 20 °C. The data were normalized to 1 for the fluorescence intensity of Rho-MatB in the absence of ATP. The fluorescence response could be fitted to a line up to 6 μM ATP with - for the experiment shown - slopes of 0.178 ± 0.010 , 0.180 ± 0.004 , 0.153 ± 0.005 and $0.124 \pm 0.004 \mu\text{M}^{-1}$ at 0, 10, 50 and 100 μM ADP, respectively.

Figure 4: Association and dissociation kinetics of Rho-MatB and ATP. (A, B) Association kinetics were measured under pseudo-first order conditions with respect to ATP by mixing 0.25 μM Rho-MatB with different concentrations of ATP at 25 °C in 50 mM Hepes pH 7.0, 100 mM NaCl, 10 mM MgCl₂, 0.3 mg ml⁻¹ bovine serum albumin using a stopped-flow apparatus. Two time windows were used in order to accurately measure the fast phase and the slow phase. A representative set of fluorescence traces is shown in black and the labels are the micromolar ATP concentrations. Data were fitted using the Kinetic Studio software. The fitted traces are shown in red. The short time scale fluorescence signal observed after mixing Rho-MatB and ATP was well fitted using a single exponential with a linear term. The long time scale fluorescence signal fitted to a double exponential function using a fixed rate constant for the fast phase determined previously from the short time scale traces. The average rate constant determined for the slow phase is $0.88 \pm 0.13 \text{ s}^{-1}$. The observed rate constants for the fast phase ($k_{\text{obs,fast}}$) are plotted against ATP concentration (C). Linear regression gave an association rate constant of $1.83 \pm 0.02 \mu\text{M}^{-1} \text{ s}^{-1}$ and the intercept was $8.15 \pm 0.17 \text{ s}^{-1}$. (D) Association kinetics were measured under pseudo-first order conditions with a 5 or 10-fold excess of Rho-MatB over ATP (0.25 μM ATP and 1.25 μM Rho-MatB; 0.50 μM ATP and 2.50 μM Rho-MatB; 0.375 μM ATP and 3.75 μM Rho-MatB; 0.500 μM ATP and 5.00 μM Rho-MatB; 0.625 μM ATP and 6.25 μM Rho-MatB; 0.750 μM ATP and 7.50 μM Rho-MatB or 1.00 μM ATP and 10.0 μM Rho-MatB). A representative set of fluorescence traces, recorded at 25 °C in 50 mM Hepes pH 7.0, 100 mM NaCl, 10 mM MgCl₂, 0.3 mg ml⁻¹ bovine serum albumin using a stopped-flow apparatus, is shown in black. Each trace is normalized to 100% for the initial intensity but offset by 2% from each other for clarity. The concentrations mentioned in the graph are those of Rho-MatB in the mixing chamber. Data were fitted with a single exponential using the Kinetic Studio software. The fitted traces are shown in red. The observed rate constants are plotted against Rho-MatB concentration (E). Linear regression gave an association rate constant of $0.72 \pm 0.03 \mu\text{M}^{-1} \text{ s}^{-1}$ and the intercept was $7.95 \pm 0.18 \text{ s}^{-1}$. (F) To determine the dissociation rate constants, we mixed the preformed ATP (5 μM) and Rho-MatB (0.25 μM) complex with 25 μM (solid line) or 50 μM (dashed line) His₆-RpMatB C106A/K488A at 25 °C in 50 mM Hepes pH 7.0, 100 mM NaCl, 10 mM

MgCl₂, 0.3 mg ml⁻¹ bovine serum albumin in a stopped-flow apparatus. Representative time courses are shown. The kinetics of the fluorescence change are limited by the ATP dissociation, as shown by varying the concentration of His₆-RpMatB C106A/K488A. Data were fitted with a double exponential using the Kinetic Studio software. The average rate constants determined are $6.94 \pm 0.28 \text{ s}^{-1}$ (~85 %) and $1.50 \pm 0.07 \text{ s}^{-1}$ (~15 %). The percentages in parentheses are the relative intensities of the fast and slow phases for double-exponential fits, these were approximately constant as the concentration varied.

Figure 5: (A) Conformational selection model. Rho-MatB1 and Rho-MatB2 correspond to two different conformational states of Rho-MatB. Only Rho-MatB2 can bind ATP to form ATP·RhoMatB2. **(B) Induced fit model.** Rho-MatB binds ATP to form ATP·RhoMatB. ATP·RhoMatB* corresponds to ATP-bound Rho-MatB in a different conformational state.

Figure 6: The production of ATP by pyruvate kinase as monitored by Rho-MatB. **(A)** The calibration curve to define the relationship between fluorescence intensity and ATP concentration. The fluorescence intensity of a solution containing 2.5 μM Rho-MatB in 50 mM Tris HCl pH 7.5, 100 mM NaCl, 10 mM MgCl₂, 100 mM KCl, 0.3 mg ml⁻¹ bovine serum albumin, 250 μM ADP, 2.5 mM phosphoenolpyruvate and ATP was measured in a microplate reader. A linear relationship between fluorescence intensity and ATP concentration at these conditions was observed up to 5 μM ATP. The slope obtained by linear regression analysis is shown. **(B)** Time courses of fluorescence change upon ATP production by pyruvate kinase at different phosphoenolpyruvate concentrations as monitored by Rho-MatB using a microplate reader. Reaction mixtures contained 50 mM Tris HCl pH 7.5, 100 mM NaCl, 10 mM MgCl₂, 100 mM KCl, 0.3 mg ml⁻¹ bovine serum albumin, 250 μM ADP, 2.5 μM Rho-MatB and various phosphoenolpyruvate concentrations. All reactions were started by the addition of pyruvate kinase (0.025 U ml⁻¹), and the change in fluorescence was monitored for several minutes at room temperature (~20 °C). The initial rates (v_i) were determined by linear regression (from 200 to 400 s) using the slope obtained from the calibration curve and were plotted versus phosphoenolpyruvate concentration **(C)**. The parameters K_m and V_{max} were obtained from a curve fit according to the Michaelis-Menten equation and are - for the experiment shown - $75.5 \pm 3.4 \text{ μM}$ and $0.019 \pm 0.002 \text{ μM s}^{-1}$, respectively. **(D)** Time courses of ATP production by pyruvate kinase as monitored by Rho-MatB in a stopped-flow apparatus. Reaction mixtures contained 50 mM Tris HCl pH 7.5, 100 mM NaCl, 10 mM MgCl₂, 100 mM KCl, 0.3 mg ml⁻¹ bovine serum albumin, 250 μM ADP, 100 μM phosphoenolpyruvate and 2.5 μM Rho-MatB. All reactions were started by the addition of pyruvate kinase (0.025, 0.50, 0.75, 1.0, 2.0, 4.0, 5.0 or 6.0 U ml⁻¹), and the change in fluorescence was monitored for several seconds at 25 °C. ATP concentrations were calculated from the fluorescence signal using the calibration method. Briefly, the signal was calibrated by consecutively introducing 2.5 μM Rho-MatB alone in 50 mM Tris HCl pH 7.5, 100 mM NaCl, 10 mM MgCl₂, 100 mM KCl, 0.3 mg ml⁻¹ bovine serum albumin, 250 μM ADP, 100 μM

phosphoenolpyruvate; then the same solution containing 0.5, 1.0, 2.0, 3.0, 4.0 or 5.0 μM ATP. At each stage the fluorescence was measured. A representative set of traces at different pyruvate kinase concentrations is shown. Traces are offset by 0.20 μM ATP from each other at zero time for clarity. The initial rates (v_i) were determined by linear regression and were plotted versus pyruvate kinase concentration (E).

Figure 7 (Supplemental figure 1): Chromatogram of the purification of Rho-MatB via ion exchange chromatography. Detection was performed via $A_{280\text{nm}}$. The conductivity (red) is shown on the secondary vertical axis. The conductivity signal gives an indication of the applied gradient. See Materials and methods for further details.

We observed that the protein eluted in two peaks. Overall, our data suggested that the protein was in equilibrium between two “states”. For instance, when one peak was re-run using the same conditions, we could see the same elution profile again (data not shown). Also, there is no difference in the fluorescence change upon ATP-binding and the affinity for ATP for both peaks (data not shown).

Figure 8 (Supplemental figure 2): Characterization of the fluorescence spectra of Rho-MatB. (A) Fluorescence emission spectra of 1 μM Rho-MatB in 50 mM Hepes pH 7.0, 100 mM NaCl, 10 mM MgCl_2 and 0.3 mg ml^{-1} bovine serum albumin (black line) supplemented with 1 mM malonate (red line), 0.25 mM coenzymeA (blue line), 1 mM malonate and 0.25 mM coenzymeA (green line) or 1 mM malonate, 0.25 mM coenzymeA and 1 mM ATP (orange line). Data were corrected for dilution. **(B)** Fluorescence emission spectra of 1 μM Rho-MatB in 50 mM Hepes pH 7.0, 100 mM NaCl, 0.5 mM EDTA and 0.3 mg ml^{-1} bovine serum albumin (black line) supplemented with 250 μM ATP (red line) or 250 μM ATP and 10 mM MgCl_2 (blue line). Data were corrected for dilution.

The fluorescence emission spectra were measured by exciting protein solutions at 553 nm.

Figure 9 (Supplemental figure 3): Simulations of ATP-binding to Rho-MatB. (A) Simulation using the conformational selection model (Figure 5A). Simulated time traces shown are, from bottom to top, 0.25 μM ATP : 1.25 μM Rho-MatB; 0.375 μM ATP: 3.75 μM Rho-MatB; 0.625 μM ATP: 6.25 μM Rho-MatB and 1.00 μM ATP: 10.0 μM Rho-MatB. The simulated time traces are normalized to 100% for the initial signal but offset by 2% from each other for clarity. The observed fluorescence signal was generated by combined fluorescence from all species using the following formula: $a \times ([\text{Rho-MatB1}] + [\text{Rho-MatB2}]) + b \times [\text{ATP} \cdot \text{RhoMatB2}]$ with $a = 1$ and $b = 3.7$. Rho-MatB1 and Rho-MatB2 concentrations were both set at 50 % of the total [Rho-MatB]. Rate constants were $k_{+1} = 0.88 \text{ s}^{-1}$, $k_{-1} = 0.88 \text{ s}^{-1}$, $k_{+2} = 1.83 \mu\text{M}^{-1} \text{ s}^{-1}$ and $k_{-2} = 8.15 \text{ s}^{-1}$. **(B)** Simulation using the induced fit model (Figure 5B). Simulated time traces shown are, from bottom to top, 0.25 μM ATP : 1.25 μM Rho-MatB; 0.375 μM ATP: 3.75 μM Rho-MatB; 0.625 μM ATP: 6.25 μM Rho-MatB and 1.00 μM ATP: 10.0 μM Rho-MatB. The simulated time traces are normalized to 100% for the initial signal but offset by 2% from each other for clarity. The observed fluorescence signal is generated by

combined fluorescence from all species using the following formula: $a \times [\text{Rho-MatB}] + b \times [\text{ATP} \cdot \text{RhoMatB}] + c \times [\text{ATP} \cdot \text{RhoMatB}^*]$ with $a = 1$, $b = 3$ and $c = 3.7$. Rate constants were $k_{+1} = 1.83 \mu\text{M}^{-1} \text{s}^{-1}$, $k_{-1} = 8.15 \text{s}^{-1}$, $k_{+2} = 0.88 \text{s}^{-1}$ and $k_{-2} = 0 \text{s}^{-1}$.

- 5 **Figure 10 (Supplemental figure 4): Rho-MatB activity assay.** We measured the time course and extent of NADH oxidation in the presence of Rho-MatB to determine the residual activity of Rho-MatB. Reaction mixtures (200 μl) contained 50 mM Hepes buffer pH 7.5, 25 mM NaCl, 25 mM KCl, 0.3 mg ml^{-1} bovine serum albumin, 0.5 mM ATP, 0.5 mM coenzymeA, 10 mM MgCl_2 , 3 mM phosphoenolpyruvate, 0.2 mM
- 10 NADH, 0.01 U μl^{-1} pyruvate kinase, 0.05 U μl^{-1} adenylate kinase, 0.015 U μl^{-1} lactate dehydrogenase and 2 mM malonate. The reaction was started by the addition of 0.03 μM Rho-MatB and the change in $A_{340\text{nm}}$ was monitored in time. No activity was observed. As a positive control, 0.03 μM (5-ATR)₂-His₆-RpMatB C106A/R286C/Q457C was added to the reaction mixture. The specific activity for (5-
- 15 ATR)₂-His₆-RpMatB C106A/R286C/Q457C was calculated using $\epsilon_{340\text{nm},\text{NADH}}$ and was $37 \pm 4 \mu\text{mol AMP min}^{-1} \text{mg}^{-1}$.

- Figure 11 (Supplemental figure 5): The production of ATP by pyruvate kinase as monitored by a coupled-enzyme assay. (A)** Reaction mixtures (200 μl) contained 50
- 20 mM Tris HCl pH 7.5, 100 mM NaCl, 10 mM MgCl_2 , 100 mM KCl, 0.3 mg ml^{-1} bovine serum albumin, 250 μM ADP, 200 μM NADH, 5 U ml^{-1} lactate dehydrogenase and various concentrations of phosphoenolpyruvate. All reactions were started by the addition of pyruvate kinase (0.10 U ml^{-1}) and the change in $A_{340\text{nm}}$ was monitored for several minutes at room temperature ($\sim 20^\circ\text{C}$) using a microplate reader. The initial
- 25 rates (v_i) were determined by linear regression (from 200 to 700 s) using $\epsilon_{340\text{nm},\text{NADH}}$ and were plotted versus phosphoenolpyruvate concentration. **(B)** The parameters K_m and V_{max} were obtained from a curve fit according to the Michaelis-Menten equation and are - for the experiment shown - $173 \pm 19 \mu\text{M}$ and $0.063 \pm 0.002 \mu\text{M s}^{-1}$, respectively.

- 30 **Figure 12 (Supplemental figure 6): The quaternary structures of His₆-RpMatB C106A/R286C/Q457C/K488A and Rho-MatB were analyzed via SEC-MALLS.** About 15 μM Rho-MatB (blue) and His₆-RpMatB C106A/R286C/Q457C/K488A (red) were analyzed via SEC-MALLS at room temperature at a flow rate of 0.5 ml min^{-1} in the absence of ATP (30 mM Tris HCl pH 7.5, 100 mM NaCl, 3 mM NaN_3 , graph on the left) and in the presence of ATP and Mg^{2+} (30 mM Tris HCl pH 7.5, 100 mM NaCl, 5
- 35 mM MgCl_2 , 50 μM ATP, 3 mM NaN_3 , graph on the right). The differential refractive index and the molar mass distributions, determined throughout the elution of each peak, were plotted against time.

- Briefly, in the absence of ATP or Mg^{2+} , we observe a broad peak with fronting corresponding to a molar mass ranging from ~ 60 to 80 kDa for His₆-RpMatB C106A/R286C/Q457C/K488A and a molar mass ranging from ~ 60 to 90 kDa for Rho-
- 40 MatB. At higher concentrations of His₆-RpMatB C106A/R286C/Q457C/K488A, the equilibrium shifts toward the 60 kDa form (data not shown). Addition of 50 μM ATP in the presence of 5 mM Mg^{2+} shifts the equilibrium towards the 60 kDa species for the

unlabeled protein. However, no clear shift is observed for Rho-MatB, i.e. its molar mass ranges from ~60 to 80 kDa.

To conclude, we observe an equilibrium between two species. This can be monomer and dimer protein. However, it is also possible that the protein exists as two different conformations in solution. This equilibrium is not only concentration-dependent but also ligand-dependent.

Figure 13 (Table 1): Fluorescence changes (F_+/F_-) and dissociation constants (K_d) for binding of ATP and other ligands to Rho-MatB. The fluorescence change upon ligand binding and the equilibrium dissociation constants were obtained from fluorescence titrations as described in Figure 3A and Figure 3B.

Figure 14 (Supplemental table 1): Fluorescence changes (F_+/F_-) and dissociation constants (K_d) for binding of ATP and ADP to diethylaminocoumarin (MDCC and IDCC) and tetramethylrhodamine (5-ATR and 6-ATR) labeled RpMatB cysteine mutants. Data are from a survey without complete optimization of the labeling and purification for each mutant. The fluorescence changes upon ligand binding were obtained from fluorescence emission spectra at 20 °C in 50 mM Tris HCl pH 7.5, 100 mM NaCl, 10 mM MgCl₂, 0.3 mg / ml bovine serum albumin using 1 μM protein and excess ATP or ADP. The equilibrium dissociation constants were obtained from fluorescence titrations at 20 °C in 50 mM Tris HCl pH 7.5, 100 mM NaCl, 10 mM MgCl₂, 0.3 mg / ml bovine serum albumin using 0.5 μM protein and various concentrations of ATP or ADP.

Figure 15 (Supplemental table 2): Fluorescence changes (F_+/F_-) and dissociation constants (K_d) for binding of ATP to Rho-MatB in different buffer conditions. The fluorescence changes upon ATP binding were obtained from fluorescence emission spectra at 20 °C in the buffer mentioned using 1 μM protein and excess ATP in the presence of 10 mM MgCl₂ and 0.3 mg ml⁻¹ bovine serum albumin. The equilibrium dissociation constants were obtained from fluorescence titrations at 20 °C in the buffer mentioned using 0.5 μM protein and various concentrations of ATP.

Figure 16 (Supplemental Table 3): Structure comparison between apo and holo (ATP (analogue)-bound) ANL superfamily proteins.

To investigate the degree of the conformational change after ATP-binding within the ANL superfamily, we collected 5 apo-holo protein pairs (based on (Gulick 2009), including a more up-to-date table on <http://labs.hwi.buffalo.edu/gulick/table.html>). The structural similarities/differences between apo and holo structures were investigated using the root mean square deviation (RMSD) of the alpha carbon atoms of aligned residues using SuperPose (Maiti, Van Domselaar et al. 2004). The table shows that not only RpMatB but also other ANL superfamily members undergo a conformational change after ATP (analogue) binding.

We also analyzed the relative interdomain orientations through manual inspection of the placement of the C-terminal domain with respect to its N-terminal domain after superposing the N-terminal domain. For the holo molecules, we found 9 out of 11

structures for which the interdomain geometry is conserved. On the contrary, the superposition of the apo molecules showed large variations in the positioning of the C-terminal domain relative to the N-terminal domain.

(%) In total, we analyzed 11 holo molecules and 25 apo molecules. We found that the C-terminal domain was not present in 11 out of 25 analyzed apo crystal structures. It was either unresolved hence assumed to be flexible (like PDB 3IVR and PDB 4LGC) or not included in the purified protein (like PDB 3T5B and PDB 3WV4). For simplicity, only the 5 apo-holo pairs are shown in this table.

(+) Ratio of the average B-factor of the C-terminal domain using all atoms and the average B-factor of the N-terminal domain using all atoms.

(*) This crystal structure contained multiple chains that had a different conformation. Chains representing both conformations were used in the analysis.

(\$) Only partially resolved C-terminal domain.

(#) This crystal structure contained multiple chains but all had the same conformation. Only one was used in the analysis.

References to publications can be found using the links at the Protein Data Bank.

Figure 17 shows an exemplary sequence of the invention referred to as RpMatB, which is annotated to show examples of substitutions and/or additions useful in sensors of the invention (SEQ ID NO: 7).

Figure 18 shows summary of the invention.

Figure 19 shows structural comparison of the proteins in Table B. Location of amino acids closest to those used in RpMatB as cysteine points of attachment.

Figure 20 shows C-terminal domain rotation upon MgATP binding and position of mutations in RpMatB.

Figure 21 shows Crystal structure of RpMatB in the MgATP bound conformation, showing conserved motifs and position of binding site mutations.

(A) Protein structure showing the conserved core motifs and MgATP (Crosby, Rank et al. 2012). The highly conserved core motifs (see also Table S1) are shown: structural motifs are coloured green, binding site motifs are pink. (B) ATP binding site, showing the position of the three mutations.

Figure 22 shows Fluorescence excitation and emission spectra of variants of Rho-MatB.

(A) 1 μ M Rho-MatB T167A with and without 5 mM ATP; (B) 1 μ M Rho-MatB T303A with and without 3 mM ATP; (C) 1 μ M Rho-MatB S170A with and without 0.5 mM ATP. These ATP concentrations were saturating for the variant. Solutions were in 50 mM Hepes pH 7.0, 100 mM NaCl, 10 mM MgCl₂, 0.3 mg ml⁻¹ bovine serum albumin at 20 °C.

Figure 23 shows Nucleotide affinity to variants of Rho-MatB.

Titration of ATP (circles) and ADP (triangles) (A) 0.5 μ M Rho-MatB T167A, (B) 0.5 μ M Rho-MatB T303A; (C) 0.5 μ M Rho-MatB S170A. Measurements were in 50 mM Hepes pH 7.0, 100 mM NaCl, 10 mM MgCl₂, 0.3 mg ml⁻¹ bovine serum albumin at 20 °C. The dissociation constants and fluorescence ratios were obtained using a quadratic binding equation (see Methods) and are listed in Table D. The drop lines are to indicate the dissociation constants.

Figure 24 shows The effect of ADP on ATP binding.

Calibrations were determined by measuring the fluorescence as a function of ATP with different amounts of ADP present, but with the total nucleotide concentration (ADP + ATP) constant. (A) 1 μ M Rho-MatB T167A; (B) 1 μ M Rho-MatB T303A; (C) 1 μ M Rho-MatB S170A. The total nucleotide concentration is shown in micromolar. For Rho-MatB S170A, 500 μ M ADP was added at each ATP concentration. Solution conditions were as in Figure 23. The data were linear fit to demonstrate approximate linear dependence over the range measured.

Figure 25 shows Association kinetics of ATP binding to Rho-MatB variants

Fluorescence time courses were measured by rapidly mixing different concentrations of ATP with 0.25 μ M Rho-MatB with a large excess of ATP (micromolar concentration shown) in 50 mM Hepes pH 7.0, 100 mM NaCl, 10 mM MgCl₂, 0.3 mg ml⁻¹ bovine serum albumin at 20 °C. Time courses are for two different time scales were obtained to show fast and slow phases. Slow phases are shown in Figure 28 (Figure S3). Note that the dead time of the stopped-flow instrument is ~2 ms, so that the traces of the fast phase only record changes from that time. (A) Example traces for Rho-MatB T167A, T303A and S170A variants. (B) The fast phases were fit to single exponentials to give rate constants (k_{obs}), increasing linearly with ATP concentration. A simple binding model for pseudo-first order conditions gives $k_{\text{obs}} = k_+[ATP] + k_-$: the slope (second order rate constant for association, k_+) and intercept (dissociation rate constant, k_-) are shown in Table D.

Figure 26 (Figure S1) shows core motifs sequences Sequence logos were created for ANL superfamily proteins, using WebLogo 3.4 (<http://weblogo.threeplusone.com/>). Sequence conservation is indicated as the total height of each stack (measured in bits), while the relative height of bases in a stack reflects base frequencies at that position. The numbers correspond to the alignment position. The colour scheme is based on hydrophobicity: R, K, D, E, N, Q are blue; S, G, H, T, A, P are green; Y, V, M, C, L, F, I, W are black. The motifs shown are ones in which mutations were prepared: the sequence of RpMatB is also shown for each.

Figure 27 (Figure S2) shows Absorbance spectra of variants of Rho-MatB with and without ATP (A) 1 μ M Rho-MatB T167A with and without 5 mM ATP; (B) 1 μ M Rho-MatB T303A with and without 3 mM ATP; (C) 1 μ M Rho-MatB S170A with and without 0.5 mM ATP. These ATP concentrations were saturating for the variant. Solutions were in 50 mM Hepes pH 7.0, 100 mM NaCl, 10 mM MgCl₂, 0.3 mg ml⁻¹ bovine serum albumin at 20 °C.

Figure 28 (Figure S3) shows Association kinetics of variants of Rho-MatB with excess ATP Example time courses were obtained as in Figure 25 at various ATP concentrations, shown in micromolar. Rho-MatB T167A, T303A and S170A variants. While Figure 25 shows the fast phases of each time course, the equivalent slow phase are shown here. These were fit to single exponentials, whose rate constants varied little with ATP concentration. The average rate constants, measuring a conformation change as described in the main text, are in Table D.

25 EXAMPLES

Experimental procedures

Selection of protein scaffold

All available structures of proteins expressed in *E.coli* with either ATP or an ATP analogue bound were extracted from the Protein Data Bank (PDB) ([Berman, Westbrook et al. 2000](#)). To obtain the corresponding ligand-free structure for each ATP or ATP analogue -bound structure, the sequences of all available PDB structures of proteins expressed in *E.coli*. were compared with the ATP and ATP analogue -bound sequences using CD-HIT2D ([Huang, Niu et al. 2010](#)). This program identifies the sequences in one database that are similar to the other database. Sequences with 90% sequence identity were considered to be the same protein and allowed ligand-bound

and ligand-free structures to be compared using PyMOL (The PyMOL Molecular Graphics System, Version 1.3, Schrödinger, LLC).

Plasmids

5 Plasmid pRpMatB39 (plasmid pTEV5 containing the coding sequence of RpMatB with the point mutation K488A and an N-terminal His₆-tag) was provided by J.C. Escalante-Semerena (Crosby, Rank et al. 2012). The QuikChange site-directed mutagenesis protocol and the QuikChange Lightning Multi site-directed mutagenesis kit (Stratagene) were used for single-site or multi-site mutations of the pRpMatB39
10 plasmid, respectively. The primers to introduce the point mutation C106A in the pRpMatB39 plasmid are 5'-ccgaagatcggtggccgatccgccaagcg-3' and 5'-cgcttgacggatcgccaccacgatcttcgg-3'. The primers to introduce R286C and Q457C, i.e. the mutations in the most preferred RpMatB biosensor, are respectively 5'-gctgccgatacgcatggaatggtcg-3' and 5'-acgatcgacgaagcgtgctgctgcacggcctc-3'.

15

Protein expression and purification

RpMatB variants were synthesized in *E. coli* OverExpress C41(DE3) cells (Lucigen). The cell pellet of a 5 ml overnight culture, grown at 37 °C and supplemented with 100 µg ml⁻¹ ampicillin, was resuspended in 0.5 l of lysogeny broth medium, supplemented
20 with ampicillin, in 5 l shaker flasks and cultured via vigorous shaking at 30 °C to an optical density at 600 nm of 0.6 - 0.8 cm⁻¹. Then, 0.5 mM isopropyl-β-D-thiogalactoside was added to start induction at 30 °C for 16 h. After induction, the cultures were cooled down to 4 °C and centrifuged at 3500 rpm for 30 min at 4 °C (rotor JS 4.2, Beckman). The cell pellet was washed with 30 ml of ice-cold buffer (10
25 mM Tris HCl pH 7.5, 300 mM NaCl), centrifuged at 3500 rpm for 30 min at 4 °C (rotor JS 4.2, Beckman), the supernatant discarded and the pellet stored at -80 °C until use. About 3 g wet weight of *E. coli* cells were harvested from 0.5 l culture in a typical preparation.

The cell pellet was resuspended in 35 ml 30 mM Tris HCl, 300 mM NaCl, 10 mM
30 imidazole, 3 mM tris(2-carboxyethyl)phosphine, 2 mM phenylmethanesulfonyl fluoride, pH 8.0 and sonicated on ice using an ultra-sonicator (VC505, Sonics) at 200 W for 5 times 30 s with a 5 s on / 5 s off pulser. The soluble fraction was collected by centrifugation at 35000 rpm for 45 min at 4 °C (rotor 45 Ti, Beckman). The His₆-tagged protein was purified at 4 °C on an immobilized metal ion affinity chromatography (1 ml
35 HisTrap HP column, GE Healthcare) using an Äkta system (GE Healthcare). The resin was equilibrated with Buffer A (30 mM Tris HCl, 300 mM NaCl, 10 mM imidazole, 1 mM tris(2-carboxyethyl)phosphine, pH 8.0). The sample was filtered (0.45 µm

Minisart NML filter, Sartorius) and loaded onto the column at 0.5 ml min⁻¹. The column was washed with 20 ml Buffer A and additionally with 20 ml of 95 % Buffer A and 5 % Buffer B (30 mM Tris HCl, 300 mM NaCl, 250 mM imidazole, 1 mM tris(2-carboxyethyl)phosphine, pH 8.0) at a flow rate of 1 ml min⁻¹. The protein was eluted
5 with 20 ml of Buffer B at a flow rate of 1 ml min⁻¹. Protein fractions were pooled (~2-4 ml) and further purified via size exclusion chromatography at 4 °C using the HiLoad 16/60 Superdex 200 prep grade column (GE Healthcare) equilibrated with 30 mM Tris HCl, 100 mM NaCl, 0.5 mM ethylenediaminetetraacetic acid, 5 mM dithiothreitol, 1 mM NaN₃. The flow rate was 1 ml min⁻¹ and the loading volume ranged from 2 to 4 ml.
10 Fractions containing the protein were pooled and concentrated (VivaSpin 20 MWCO 10 kDa cut off, GE Healthcare) to ~10 mg ml⁻¹.

The protein concentration was determined from the absorbance at 280 nm using the extinction coefficient at 280 nm of 46300 M⁻¹ cm⁻¹ calculated from the sequence via Expasy ProtParam (Wilkins, Gasteiger et al. 1999). The protein was drop-frozen in
15 liquid nitrogen and stored at -80 °C. Typically, 45 mg of protein was obtained from 3 g wet weight of cells.

Labeling of RpMatB with fluorescent dyes

We will first describe the detailed procedure for labeling and purification of the most
20 preferred sensor variant RpMatB (His₆-RpMatB C106A/R286C/Q457C/K488A). Deviations from this procedure during labeling and purification of RpMatB variants at previous stages of the sensor development are mentioned afterwards.

Dithiothreitol was removed from ~40 mg of protein using a PD10 desalting column (GE Healthcare) pre-equilibrated with Buffer L (30 mM Tris HCl pH 7.5, 100 mM NaCl) at
25 20 °C. 50 µM protein was incubated at 20 °C with 225 µM 5-iodoacetamidotetramethylrhodamine (5-IATR, AnaSpec, CA) in Buffer L using an end-over-end mixer for 90 min. Afterwards, 2 mM sodium-2-mercaptoethanesulfonate was added and incubation continued for 15 min. After centrifugation at 3500 rpm for 15 min at 4 °C (Heraeus Biofuge), the supernatant was filtered through a 0.2 µm syringe
30 filter (Acrodisc Syringe Filter with HT Tuffryn Membrane, Pall Life Sciences) and loaded onto a PD10 desalting column, equilibrated with Buffer Q1 (30 mM Tris HCl pH 8.0, 25 mM NaCl) at ~20 °C to remove free label.

The most preferred RpMatB biosensor was further purified via ion exchange chromatography at 4 °C using a 1 ml HiTrap Q HP column (GE Healthcare),
35 equilibrated in Buffer Q1. The flow rate was 1 ml min⁻¹ during the whole purification. After sample loading, the column was washed with 90 ml Buffer Q1. The protein was eluted using a gradient from 100 % Buffer Q1 to 50 % Buffer Q1 and 50 % Buffer Q2 (30

mM Tris HCl pH 8.0, 1 M NaCl) over 25 ml followed by a gradient from 50 % Buffer Q1 and 50 % Buffer Q2 to 100 % Buffer Q2 over 10 ml. Fractions containing the protein were pooled and concentrated to $\sim 5 \text{ mg ml}^{-1}$ using a concentrator (Amicon Ultra-4 10 kDa cut off, Millipore).

- 5 While screening for a fluorescence change upon ATP binding, labeling was performed on a scale of $\sim 2 \text{ mg}$ of protein. RpMatB variants ($100 \mu\text{M}$) were incubated at 20°C with 2-fold (7-diethylamino-3-((((2-maleimidyl)ethyl)amino)carbonyl)coumarin (MDCC) or 7-diethylamino-3-((((2-iodoacetomido)ethyl)amino)carbonyl)coumarin (IDCC)) or 4-fold (5-IATR or 6-IATR (synthesized in-house (Corrie and Craik 1994))) excess of
10 fluorophore over RpMatB for 90 (tetramethylrhodamine), 35 (MDCC) or 120 (IDCC) min. During screening of all the labeled RpMatB variants, proteins were not further purified via ion exchange chromatography but tested after removal of free label using a PD10 desalting column.

The labeled protein concentrations were determined using the following extinction
15 coefficients: RpMatB: $\epsilon_{280\text{nm}}$ ($46300 \text{ M}^{-1} \text{ cm}^{-1}$), tetramethylrhodamine: $\epsilon_{280\text{nm}}$ ($31000 \text{ M}^{-1} \text{ cm}^{-1}$) and $\epsilon_{528\text{nm}}$ ($52000 \text{ M}^{-1} \text{ cm}^{-1}$) (Corrie and Craik 1994), MDCC: $\epsilon_{280\text{nm}}$ ($7470 \text{ M}^{-1} \text{ cm}^{-1}$) and $\epsilon_{430\text{nm}}$ ($46800 \text{ M}^{-1} \text{ cm}^{-1}$) and IDCC: $\epsilon_{280\text{nm}}$ ($7470 \text{ M}^{-1} \text{ cm}^{-1}$) and $\epsilon_{430\text{nm}}$ ($44800 \text{ M}^{-1} \text{ cm}^{-1}$). The protein was drop-frozen in liquid nitrogen and stored at -80°C . Labeling yields were up to 35 %.

- 20 Using mass spectrometry, unlabeled RpMatB (His₆-RpMatB C106A/R286C/Q457C/K488A) had a mass of 57335.9 Da. The theoretical molecular weight is 57324.2 Da, assuming loss of the N-terminal methionine. Presumably, other post-translational modifications, such as partial oxidation of surface-accessible methionines, occur (Guan, Yates et al. 2003). The most preferred RpMatB biosensor,
25 Rho-MatB ((5-ATR)₂-His₆-RpMatB C106A/R286C/Q457C/K488A)), had a mass of 58216.4 Da, conforming to the theoretical molecular weight of the RpMatB protein labeled with two 5-IATRs ($57335.9 + 2 \times 441.5$).

- The solution molecular weight was analyzed using size exclusion chromatography coupled to multi-angle laser light scattering (SEC-MALLS). Protein (1 mg ml^{-1}) was
30 applied in a volume of 100 μl to a Superdex 200 10/300 GL column (GE Healthcare) connected to a Jasco PU-1580 HPLC at a flow rate of 0.5 ml min^{-1} . The HPLC system was connected to a Dawn Heleos II light scattering instrument (Wyatt Technology) and Optilab T-rex differential refractometer (Wyatt Technology). The solution molecular weight was determined from the combined data from both detectors using the ASTRA
35 software version 6.1.1.17 (Wyatt Technology) with the refractive index increment set to 0.1860 ml g^{-1} .

RpMatB activity assay

RpMatB specific activity was quantified using a nicotinamide adenine dinucleotide (NADH) assay (Crosby, Rank et al. 2012). Reaction mixtures (200 µl) contained 50 mM Hepes buffer pH 7.5, 25 mM NaCl, 10 mM MgCl₂, 25 mM KCl, 0.3 mg ml⁻¹ bovine serum albumin, 0.5 mM ATP, 0.5 mM coenzymeA, 3 mM phosphoenolpyruvate, 0.2 mM NADH, 0.01 U µl⁻¹ pyruvate kinase (rabbit muscle (Sigma)), 0.05 U µl⁻¹ adenylate kinase (chicken muscle (Sigma)), 0.015 U µl⁻¹ lactate dehydrogenase (rabbit muscle (Sigma)) and 2 mM malonate. All reactions were started by the addition of RpMatB (0.03 µM), and the change in A_{340nm} was monitored for several min at 25 °C on a Jasco V-550 UV-Vis Spectrophotometer (Jasco).

Fluorescence measurements

Fluorescent measurements were obtained on a Cary Eclipse Spectrofluorometer (Agilent Technologies), using a 3-mm path-length quartz cuvette (Hellma), unless otherwise mentioned. Excitation and emission slits were set at 5 nm. Protein and nucleotide concentrations, buffer conditions and excitation and emission wavelengths used are given in the figure legends.

For titrations to measure nucleotide binding, the fluorescence was excited at 553 nm and emission was measured at 571 nm. Data were analyzed with a quadratic binding curve using Grafit software (Leatherbarrow 2009):

$$F = F_{min} + (F_{max} - F_{min}) \times \frac{(K_d + L + P - \sqrt{(K_d + L + P)^2 - 4 \times P \times L})}{2 \times P}$$

where P and L are the total concentrations of protein and ligand, respectively, K_d is the dissociation constant, and F_{min} and F_{max} are the fluorescence intensities of the free and ligand- bound protein, respectively.

Stopped-flow measurements

Stopped flow experiments were carried out using a HiTech SF-61DX2 apparatus (TgK Scientific, Bradford-on-Avon, UK) with a xenon-mercury lamp and operated by Kinetic Studio software (TgK Scientific). The excitation wavelength was 548 nm and there was an OG570 cut-off filter on the emission. The concentrations mentioned in the text and figures are those in the mixing chamber. Data were fitted to theoretical equations using the Kinetic Studio software.

KinTek Global Kinetic Explorer software (version 4.0) ([Johnson, Simpson et al. 2009](#), [Johnson, Simpson et al. 2009](#)) was used to obtain global fits and to generate simulation data for possible kinetic schemes.

5 Steady-state analysis of ATP production by pyruvate kinase

Absorbance: Steady-state activity measurements of pyruvate kinase (rabbit muscle from Sigma) were obtained on a CLARIOstar microplate reader (BMG Labtech) using a 96-well polystyrene microplate (black, clear flat bottom, Corning). Reaction mixtures (200 μ l) contained 50 mM Tris HCl pH 7.5, 100 mM NaCl, 10 mM MgCl₂, 100 mM KCl, 10 0.3 mg ml⁻¹ bovine serum albumin, 250 μ M ADP, 0.2 mM NADH, 5 U ml⁻¹ lactate dehydrogenase and various phosphoenolpyruvate concentrations. All reactions were started by the addition of pyruvate kinase (0.1 U ml⁻¹), and the change in A_{340nm} was monitored for at 20 °C. The initial rates were obtained by linear regression, using an ϵ_{340nm} for NADH of 6220 M⁻¹ cm⁻¹.

15 Fluorescence: Steady-state measurements of pyruvate kinase activity were obtained on a CLARIOstar microplate reader (BMG Labtech) using a 96-well polystyrene microplate (black, F-bottom, Greiner). Reaction mixtures (200 μ l) contained 50 mM Tris HCl pH 7.5, 100 mM NaCl, 10 mM MgCl₂, 100 mM KCl, 0.3 mg ml⁻¹ bovine serum albumin, 250 μ M ADP, 2.5 μ M Rho-MatB and various phosphoenolpyruvate 20 concentrations. The emission at 580 nm (10 nm bandwidth) after excitation at 545 nm (10 nm bandwidth) was recorded. A calibration curve was determined using various concentrations of ATP added to the solution above in the presence of 2.5 mM phosphoenolpyruvate. Reactions were started by the addition of pyruvate kinase (0.025 U ml⁻¹), and the change in fluorescence signal was monitored at 20 °C. Linear 25 regression analysis was used to determine the initial velocity.

Stopped-flow: Steady-state measurements of pyruvate kinase activity were obtained using a stopped-flow apparatus, as above. The excitation wavelength was 548 nm and there was an OG570 cut-off filter on the emission. Reaction mixtures contained 50 mM Tris HCl pH 7.5, 100 mM NaCl, 10 mM MgCl₂, 100 mM KCl, 0.3 mg ml⁻¹ bovine serum 30 albumin, 250 μ M ADP, 100 μ M phosphoenolpyruvate and 2.5 μ M Rho-MatB. A calibration curve was determined using various concentrations of ATP added to the solution above. Reactions were started by the addition of pyruvate kinase (up to 6.0 U ml⁻¹) and the change in fluorescence signal was monitored for several seconds at 25 °C. Linear regression analysis was used to determine the initial velocity.

35

Example 1: Design of the biosensor, based on RpMatB

Finding a suitable candidate protein recognition element for an ATP biosensor requires comparing ATP- or ATP-analogue- bound protein structures with their corresponding ligand-free protein structures. If that comparison revealed a significant conformational change upon ligand binding, the protein was seen as a potential candidate for biosensor development. Such conformational changes can be harnessed to transduce ligand binding to a fluorescence change to a fluorophore reporter, local to that region of the protein so that it responds to the change in structural environment. Functional parameters were considered next, for example affinity and selectivity for ATP, or known mutants that block enzymatic activity, such as ATPase activity. Following this analysis, RpMatB was chosen as the most suitable candidate for further ATP-biosensor development.

The crystal structures for the apo and MgATP-bound states of RpMatB suggested that there is a conformational change upon MgATP-binding (Crosby, Rank et al. 2012). In particular, the C-terminal domain rotates $\sim 20^\circ$ towards the N-terminal domain closing the active site cleft (Figure 1A and Figure 1B).

Cysteine mutations were introduced as sites for labeling onto a background of the C106A and K488A mutations in the wild-type protein. The K488A RpMatB variant does not catalyze the adenylation half-reaction, that is it cannot convert ATP and malonate to malonyl-AMP and pyrophosphate (Crosby, Rank et al. 2012). C106 in the wild-type protein (Figure 1A) is situated in the N-terminal domain, distant from the active site but slightly solvent accessible. Having shown that there is a low, but significant, degree of background labeling at this position (6% with MDCC), C106 was mutated to alanine.

Several variants were designed and prepared with either one or two cysteine residues introduced in order to label C106A/K488A RpMatB with one diethylaminocoumarin or two tetramethylrhodamines, respectively.

After examination of the ligand-bound and ligand-free RpMatB structures, sites for diethylaminocoumarin labeling were chosen, situated around the ligand-binding pocket on the C-terminal lobe, so that the fluorophore might experience an environmental change when MgATP binds. These labeled variants were tested for the fluorescence change upon ATP -binding in the presence of Mg^{2+} (Supplemental Table 1).

Sites for tetramethylrhodamine labeling were chosen so that stacking of the two fluorophores might be possible in the apo conformation and that dissociation of these stacked tetramethylrhodamines could occur on the conformational change with MgATP binding. In particular, positions were chosen with a suitable distance (~ 1.5 nm) and orientation between them in the apo conformation. In addition, the distance and

orientation between the chosen positions changed when MgATP binds. Results from seven different combinations are in Supplemental Table 1.

The labeled variant with the largest fluorescent increase upon MgATP-binding was (5-ATR)₂-His₆-RpMatB C106A/R286C/Q457C/K488A, hereinafter referred to as Rho-MatB. Both positions are well defined in the apo and MgATP-bound structure as shown in Figure 1A. R286 is located in the N-terminal domain and Q457 in the C-terminal domain.

Example 2: Fluorescence and absorbance properties of Rho-MatB with ATP

Figure 2A shows fluorescence excitation and emission spectra of Rho-MatB upon addition of increasing amounts of ATP. The fluorescence spectral shape did not change, but the fluorescence intensity increased 3.7-fold.

Absorbance spectra of Rho-MatB were also measured for a range of ATP concentrations (Figure 2B). In the absence of nucleotide, the maximum absorbance was at 518 nm with a smaller peak at 553 nm. This is characteristic of tetramethylrhodamine stacking, corresponding qualitatively to the absorbance spectra for other stacked rhodamines (Selwyn and Steinfelz 1972, Chambers, Kajiwarra et al. 1974). Binding ATP causes an absorbance decrease at 518 nm and increase at 553 nm, indicating that nucleotide binding reduces the stacking, as monomeric rhodamine had a larger peak at ~550 nm.

The isosbestic point was at 532 nm.

The affinity for ATP was determined by measuring the fluorescence at different concentrations of ATP in a solution of Rho-MatB and $K_{d,ATP}$ was 6.4 μ M (Figure 3A).

Example 3: Affinity of Rho-MatB for other nucleotides and potential ligands

The fluorescence of Rho-MatB responded to addition of nucleotides other than ATP. There was an increase in fluorescence intensity upon addition of ADP, deoxyadenosine triphosphate (dATP), adenosine 5'-(γ -thio)triphosphate (ATPyS) and adenosine 5'-(β , γ -imido)triphosphate (AMP-PNP). In contrast, AMP, GDP and GTP did not have a significant effect. Also, RpMatB substrates other than ATP (i.e. malonate and coenzymeA) had no influence on the fluorescence and their presence did not inhibit the MgATP-induced fluorescence increase (Supplemental Figure 2A).

The affinity for the responsive nucleotides, namely ADP, dATP, ATPyS and AMP-PNP was determined by measuring the fluorescence at different concentrations of the nucleotide in a solution of Rho-MatB (Figure 3A and Table 1). The dissociation constants for ADP, dATP, ATPyS and AMP-PNP were 428 μ M, 440 μ M, 16.2 μ M and 253 μ M, respectively. So the binding of ADP, dATP and AMP-PNP to RpMatB was

much weaker than ATP. The affinity of ATP γ S to Rho-MatB was similar to ATP. The maximum fluorescence increase upon binding these nucleotides was smaller than with ATP.

In most biological reactions ATP is likely to be produced from ADP. Consequently, ADP will be present in the assay solutions, so the influence of ADP on the fluorescence signal change and the affinity of ATP was assessed. Titration curves for ATP were obtained in the presence of a fixed concentration of ADP. The apparent K_d for ATP was similar in the absence and presence of 100 μ M ADP (Figure 3B and Table 1).

In addition, the influence on ATP affinity of ATP analogues was assessed. Titration curves for ATP were obtained in the presence of a fixed concentration of dATP, ATP γ S and AMP-PNP. Only ATP γ S had a remarkable influence on the apparent K_d for ATP by increasing it ~5-fold (Figure 3B and Table 1).

It simplifies application of a fluorescence biosensor, if it can be used over a range for which the response to concentration is approximately linear. Figure 3C shows there is a good linear fit up to 6 μ M ATP. The fluorescence response to ATP was also measured in the presence of different ADP concentrations but at constant total nucleotide concentration to mimic ADP conversion to ATP. Although the slope decreased with increasing ADP concentration, the fluorescence response to ATP remained linear in the presence of ADP over the same range (Figure 3C).

Example 4: The influence of solution conditions on the fluorescent properties of Rho-MatB

The fluorescence response to ATP was measured in different buffers, over a pH range from 6.0 to 9.0 and different ionic strengths with salt from 50 to 200 mM (Supplemental Table 2). Although the size of the fluorescence response changed with solution conditions, there were rather small effects on the K_d for ATP.

Mg²⁺ is required for the fluorescence change, suggesting that MgATP is bound (Supplemental Figure 2B), as might be expected as Mg²⁺ is a cofactor for the enzyme.

Example 5: Binding kinetics

The kinetics of ATP binding and dissociation were measured by stopped-flow fluorescence in order to assess the range of rates over which Rho-MatB is suitable for real-time measurements.

Binding kinetics were first measured under pseudo-first-order conditions by rapidly mixing different concentrations of ATP, in large excess, with Rho-MatB (Figure 4A & 4B). The time course of the subsequent fluorescence response was biphasic, fitting well to a double exponential. The observed rate constant for the fast phase increased linearly

up to 150 μM ATP (Figure 4C), giving a slope of $1.83 \mu\text{M}^{-1} \text{s}^{-1}$ and the intercept was 8.2s^{-1} . Assuming this phase represents binding, the slope is the association rate constant, the intercept is the dissociation rate constant. The dissociation constant, calculated from these values ($4.5 \mu\text{M}$) agrees well with the value from equilibrium binding data (Figure 3A and Table 1). The observed rate constants for the slow phase did not vary significantly over the ATP concentration range with an average value of 0.88s^{-1} . The biochemical basis of these rate constants will be discussed later.

In order to explain the biphasic kinetic data for ATP binding in excess ATP, the data were fitted to two possible two-step binding mechanism as shown in Figure 5. In the first model a pre-existing conformational equilibrium in the apoprotein was followed by association of MgATP to one of the conformations (conformational selection, Figure 5A). The second model involves a structural transition that follows MgATP-binding (induced fit, Figure 5B). Both mechanisms were possible.

To dissect between the two schemes, simulations of nucleotide-binding experiments using excess Rho-MatB were performed using KinTek Global Kinetic Explorer. By conducting simulations at varying initial concentrations of excess Rho-MatB for both mechanisms, single phase binding kinetics in the case of conformational selection and biphasic binding kinetics in the case of induced fit was observed (Supplemental Figure 3A and Supplemental Figure 3B).

ATP binding was measured with Rho-MatB present in large excess over the nucleotide. Rapid mixing produced a monophasic, exponential increase over the whole range of Rho-MatB concentrations used (Figure 4D). The observed rate constant increased linearly with [Rho-MatB], giving an association rate constant of $0.72 \mu\text{M}^{-1} \text{s}^{-1}$ and the intercept was 8.0s^{-1} , similar to the values obtained in the excess of ATP (Figure 4E). Hence, only the conformational selection mechanism fits our data.

Dissociation kinetics were measured directly starting from Rho-MatB.ATP complex and trapping dissociated ATP with a large excess of unlabeled protein (Figure 4F). The fluorescence decreased with time and the curves required a double exponential with rates 6.9s^{-1} and 1.5s^{-1} . These do not vary when the concentration of the unlabeled protein varies, indicating that the process is ATP-dependent.

Example 6: Enzymatic activity

In order to confirm that the most preferred RpMatB construct lacked enzymatic activity, as expected as it contains K488A, ATPase activity was measured using a coupled-enzyme assay with NADH consumption (Crosby, Rank et al. 2012). Rho-MatB showed no residual activity under the conditions tested (Supplemental Figure 4). In comparison, the similar Rho-MatB variant but without the K488 mutation ((5-ATR)₂-

His₆-RpMatB C106A/R286C/Q457C) had a specific activity of $37 \pm 4 \mu\text{mol AMP mg}^{-1} \text{ min}^{-1}$.

Example 7: Test assay: Steady-state production of ATP by pyruvate kinase

5 The ATP biosensor was tested in a steady-state assay in which ATP was produced from ADP and phosphoenolpyruvate in a reaction catalyzed by pyruvate kinase (Figure 6A and Figure 6B). This was chosen as there is a coupled-enzyme assay for this enzyme, that is well established and described below, which could be used to validate the biosensor results. The rate of ATP formation was measured at different concentrations
10 of phosphoenolpyruvate using Rho-MatB. The resulting fluorescence traces were analyzed, using a calibration curve of fluorescence between 0 and 5 μM ATP, obtained under similar experimental conditions. Rates of ATP formation were plotted as a function of phosphoenolpyruvate concentration and fitted to Michaelis-Menten kinetics (Figure 6C), giving a K_m of 103 μM and $V_{\max}$ of 0.76 $\mu\text{M s}^{-1} \text{ U}^{-1} \text{ ml}^{-1}$.

15 As a comparison, pyruvate formation was measured using a coupled-enzyme assay under the same conditions to compare the parameters directly. This coupled the pyruvate kinase reaction to that of lactate dehydrogenase, in which NADH reduces the pyruvate as it is converted to NAD⁺. This assay gave a K_m of 251 μM and a $V_{\max}$ of 0.66 $\mu\text{M s}^{-1} \text{ U}^{-1} \text{ ml}^{-1}$ (Supplemental Figure 5). The $V_{\max}$ was similar to that obtained using
20 Rho-MatB; reasons why the K_m is different for the two assays probably relates to the different extents of reaction required to perform the two assays.

In order to test the ability of Rho-MatB to assay more rapid ATP production, measurements were done with higher concentrations of pyruvate kinase. Because the time scale was below that of manual mixing, the reactions were done in a stopped flow
25 apparatus (Figure 6D). The observed rate was proportional to the amount of pyruvate kinase added (Figure 6E).

Example 8: Annotated RpMatB sequence

30 Figure 17 shows an exemplary annotated RpMatB sequence.

Numbering scheme and position of mutations:

35 The N-terminal His₆-tag is highlighted in *red/italics*. The mutations are highlighted in blue/underlined and labeled. The residues in close proximity (sequence and space) of the labeled positions are highlighted in *green/bold*. They are summarised in Table 2. Suitably truncated forms are those which lack a small number of amino acid residues from the N- or C- terminus of the polypeptide relative to wild type MatB.

Example 9

40 Design of the biosensor

After deciding that RpMatB was a suitable recognition element for ATP, we introduced either one or two cysteine residues in order to label RpMatB with diethylaminocoumarin or tetramethylrhodamine. We tested all variants for their fluorescence change upon MgATP addition and found different fluorescence responses to MgATP binding, ranging from a decrease to an increase in fluorescence intensity (Supplemental Table 1).

Hence, not all fourteen rationally designed RpMatB mutants had identical properties, probably due to a combination of several factors. However, those having the desired and most advantageous properties could be identified as described.

Without wishing to be bound by theory, reasons for the range of observations may include:

Firstly, introducing a fluorophore might have an effect on protein flexibility and therefore on the documented conformational change upon ligand binding (either because of interaction with amino acids or because of interactions between fluorophores).

Secondly, whereas comparisons between apo and ATP (analogue) -bound structures indicate that most of the ANL superfamily members undergo a conformational change after ATP (analogue) binding, we observed that the interdomain geometry in apo structures of ANL superfamily members is less conserved than in ATP (analogue) -bound structures (Supplemental Table 3). In addition, several observations provide evidence that the C-terminal domain is also more flexible than the N-terminal domain. This is based on the absence of the C-terminal domain in a lot of apo structures within the ANL superfamily and the higher average B-factor for the C-terminal domain compared to the N-terminal domain in crystal structures of proteins from the ANL superfamily (Supplemental Table 3). Both findings, namely the structurally less conserved apo conformation and a flexible C-terminal domain, would interfere with our rational design.

Characterization of Rho-MatB

Rho-MatB can be used as a biosensor for ATP under various pH and salt conditions (Supplemental Table 2) in the presence of Mg²⁺ (Supplemental Figure 2B). The maximum fluorescence increase upon ATP-binding was 3.7-fold (Figure 2A). The change in the absorbance spectrum provides support that this increase is due to the unstacking of the two tetramethylrhodamines upon ATP-binding (Figure 2B).

Rho-MatB binds ATP with a dissociation constant of 6.4 μ M (Figure 3A and Table 1), higher than RpMatB K488A (0.31 μ M) (Crosby, Rank et al. 2012). This difference

might be caused by the point mutations introduced or the tetramethylrhodamine fluorophores attached to RpMatB.

In most biological reactions ATP is likely to be produced from ADP. Consequently, ADP will be present in the assay solutions as substrate and it is important that the ATP biosensor discriminates against ADP. Rho-MatB has 67-fold greater affinity for ATP than ADP (Figure 3A and Table 1). Even at 100 μ M ADP, the response of the biosensor to ATP is only reduced slightly (Figure 3B and Figure 3C and Table 1).

RpMatB substrates, other than ATP, that is malonate and coenzymeA, have no influence on the fluorescence signal and do not interfere with an MgATP-induced fluorescence increase (Supplemental Figure 2A). Presumably they do not induce the same conformational change as MgATP binding.

The preferred Rho-MatB biosensor exhibited a low detection limit (~ 0.25 μ M).

Binding kinetics

The ATP biosensor responds relatively fast to ATP. Biphasic binding kinetics between Rho-MatB and ATP were observed in excess ATP (Figure 4B).

The fast phase is ATP-dependent – in a linear way (Figure 4C); the slow phase not. Linear concentration dependence of the observed binding rates indicates simple first-order binding. However, we do have to point out that there is a deviation from linearity for $k_{\text{obs,fast}}$ at higher ATP concentrations, suggesting two-step binding. However, no valid rate constants can be obtained for ATP concentrations of more than 200 μ M at 25 $^{\circ}$ C due to the dead time of the instrument. Therefore, no accurate kinetic parameters for a two-step binding mechanism are available.

At this point, we presumed the slow phase could be caused either by a rate-limiting conformational change prior to ATP-binding (conformational selection, Figure 5A) or by a structural transition that follows MgATP-binding (induced fit, Figure 5B) (Vogl, Pozzi et al. 2014).

To dissect between the two schemes, we carried out nucleotide-binding experiments using excess Rho-MatB. Since simulations conducted at varying initial concentrations of excess Rho-MatB predicted single phase binding kinetics in the case of conformational selection and biphasic binding kinetics in the case of induced fit (Supplemental Figure 3A and Supplemental Figure 3B). We observed single phase binding kinetics (Figure 4D) and the observed rate constant is Rho-MatB-dependent (Figure 4E). This corresponds to the conformational selection mechanism (Figure 5A).

Therefore, we believe that Rho-MatB exists in two different conformations prior to ATP-binding. ATP binds to only one conformation. In excess Rho-MatB, there is “enough” Rho-MatB in the “correct” conformation to bind ATP and we observe only

ATP-binding, hence single phase binding kinetics. In excess ATP, there is not “enough” Rho-MatB in the “correct” conformation to bind all ATP immediately and we observe ATP-binding and the Rho-MatB conformational change.

Of note, even the conformational selection scheme does not exactly fit all experimental results. Additional steps to the scheme might be required, for example a multi-step binding in which both conformational selection and induced fit are present. However, finding a detailed mechanism for Rho-MatB is not the scope of this document. We have described a minimal model to fit all of the data.

We have to add that a rate-limiting conformational change prior to ATP-binding (conformational selection) is consistent with the SEC-MALLS data that clearly show two different species in the absence of ATP (Supplemental Figure 6).

The sensor works well and a positive demonstration of its action is provided herein.

Use of the biosensor

The ATP biosensor was tested in a steady-state assay in which ATP was produced from ADP and phosphoenolpyruvate in a reaction catalyzed by pyruvate kinase (Figure 6). In addition, pyruvate formation was measured using a coupled-enzyme assay with lactate dehydrogenase to compare the parameters directly (Supplemental Figure 5). While the $V_{\max}$ was similar, the K_m was different for the two assays. According to us, this might relate to the extent of reaction required to perform the two assays. In one, you measure ATP production by pyruvate kinase directly, in the other, you measure pyruvate production through lactate dehydrogenase.

Advantages compared to other assays to monitor ATP production *in vitro*

The biosensor facilitates measuring ATP *in vitro* with relatively high time resolution. This makes it suitable for a range of real-time, kinetic assays, which can be done in close to physiological conditions, for example pH and ionic strength. A major advantage of reagentless biosensors of the invention is that only a single species is required to monitor the analyte, in this case ATP, thereby minimizing the likelihood of interference by added reagents.

It is also very specific for ATP relative to ADP. Preferred sensor Rho-MatB has 67-fold higher affinity for ATP than ADP. The highest specificity so far for a protein-based biosensor was measured for ATeam (30-fold higher affinity for ATP than ADP (Imamura, Nhat et al. 2009)).

Finally, tetramethylrhodamine has high photostability. Its fluorescence is unlikely to interfere with (or be affected by) the system being studied because of excitation around 550 nm and emission around 570 nm.

Thus, this ATP biosensor has many advantages relative to other ATP assays.

To sum up, the present work shows that Rho-MatB can be used as a sensitive probe to measure ATP. It shows a linear response in the micromolar range and is selective for ATP. It can be used to elucidate the mechanisms of ATP production.

Example 10: ATP biosensors using polypeptide sequence from members of the ANL superfamily other than RpMatB

This protein family contains acyl- and aryl-coenzymeA synthetases, the adenylation domains of nonribosomal peptide synthetases and firefly luciferases. Structural information (see Table B below) suggests a similar ligand-mediated conformational change for family members, even if the structures of ANL superfamily proteins lacking ligands are highly variable in the position of the C-terminus. This may be exploited to generate a family of ATP biosensors with possibly different properties (such as ATP sensitivity, fluorescence signal) using protein engineering techniques, based on straightforward structure and sequence principles. This approach has been successfully applied in the past to construct a biosensor family based on bacterial periplasmic binding proteins (de Lorimier, R. M., J. J. Smith, M. A. Dwyer, L. L. Looger, K. M. Sali, C. D. Paavola, S. S. Rizk, S. Sadigov, D. W. Conrad, L. Loew and H. W. Hellinga (2002). "Construction of a fluorescent biosensor family." *Protein Sci* **11**(11): 2655-2675).

This may be achieved as follows:

1. Identify a protein from the ANL superfamily either with a similar sequence or with a known structure with a similar conformation to that of RpMatB.
2. Identify sites within this protein, to which reporter fluorophores can be attached, based on homology to RpMatB, either sequence or structural homology
 - a. **Sequence-based**, using bioinformatics software:
 - i. Align the sequence of the candidate with the sequence of RpMatB. Of note, the quality of the alignment will improve when performing a multiple sequence alignment instead of a pairwise alignment.
 - ii. Compare the aligned sequences. Identify the amino acids in the candidate that are close to the labeling sites in RpMatB.
 - b. **Structure-based**, using software for protein structure analysis and visualization:
 - i. Align the structure of the candidate with the structure of RpMatB. Of note, the quality of the superposition will improve when performing a multiple structure alignment instead of a pairwise protein structure alignment.
 - ii. Compare the superposed structures. Identify the amino acids in the candidate that are close to the labeling sites, identified in RpMatB.

In this case, 6 proteins from the ANL superfamily (2 acyl- and aryl-coenzymeA synthetase members, 2 adenylation domains of nonribosomal peptide synthetase members and 2 firefly luciferase members) were chosen because crystal structures have been solved for the closed conformation. The selection was based on a review of the structures (Gulick, A. M. (2009). "Conformational dynamics in the Acyl-CoA synthetases, adenylation domains of non-ribosomal peptide synthetases, and firefly luciferase." ACS Chem Biol 4(10): 811-827, <http://labs.bwi.buffalo.edu/gulick/table.html>). To identify sites within "the candidate" that undergo local conformational change and to which reporter functions can be attached, a structure-based approach has been applied to this set of 6 proteins, as in the Figure. This gives amino acids in each case that are in similar locations to the Cysteine mutation sites of RpMatB (first line of the Table)

TABLE B: ANL superfamily members identified for structural similarity with the closed conformation of RpMatB and the proposed mutation sites, based on structural comparison as in Figure 19

Protein	PDB ID	Ligands	Proposed mutations based on	
			MDCC- RpMatB	Rho-MatB
RpMatB, an acyl- and aryl-coenzymeA synthetases (green)	4FUT	MgATP	E439	R286 and Q457
<i>Homo sapiens</i> ACSM2A, an acyl- and aryl-coenzymeA synthetases (cyan)	3C5E	MgATP	E503	E347 and (Q528, Q529 or H 530)
<i>Streptomyces sp</i> CytC1, an adenylation domain of nonribosomal peptide synthetase (dark pink)	3VNQ	ATP	L458	R290 and E476
<i>Photinus pyralis</i> luciferin 4-monooxygenase, a firefly luciferase (yellow)	4G36, chain A	SLU	E479	E325 and (D500 or Y501)
<i>Burkholderia xenovorans</i> benzoate-coenzyme A ligase, an acyl- and aryl-coenzymeA synthetases (light pink)	2V7B, chain A	Benzoate	V468	E318 and E489
<i>Bacillus cereus</i> DltA, an adenylation domain of nonribosomal peptide synthetase (purple)	3FCC	MgATP	Y438	R279 and A461
<i>Luciola cruciata</i> luciferin 4-monooxygenase, a firefly luciferase (orange)	2D1Q	MgATP	E481	E327 and D501 or Y503

References to publications can be found using the links at the Protein Data Bank.

Suitably the polypeptide component of the ATP sensor of the invention comprises sequence corresponding to one of the polypeptides in the above table.

Example 11: Location of alternate sites for labelling ATP sensors

Figure 20 shows cartoon representations of RpMatB. The N-terminal domain (amino acids 1-399) in the apo conformation (PDB 4FUQ (Crosby, Rank et al. 2012)) is coloured grey. The C-terminal domain (amino acids 400-503) in the apo conformation is shown in pink. The C-terminal domain in the MgATP-bound conformation (PDB 4FUT (Crosby, Rank et al. 2012), after superimposing the N-terminal domain of 4FUT on the N-terminal domain of 4FUQ) is shown in green. ATP (in ball and stick

conformation and coloured by CPK convention) and Mg^{2+} (depicted as an orange sphere) are shown in the active site. The C-terminal domain rotation upon MgATP-binding is clearly visible.

- 5 The positions of mutations are shown as spheres and labeled:
 - (A) C106A in the N-terminal domain (yellow), K488A in the C-terminal domain (red), R286C in the N-terminal domain and Q457C in the C-terminal domain (both blue).
 - (B) C106A in the N-terminal domain (yellow), K488A in the C-terminal domain (red), R286C in the N-terminal domain and G461C in the C-terminal domain (both blue).
 - 10 (C) C106A in the N-terminal domain (yellow), K488A in the C-terminal domain (red), K385C in the N-terminal domain and K470C in the C-terminal domain (both blue).
 - (D) C106A in the N-terminal domain (yellow), K488A in the C-terminal domain (red) and N492C in the N-terminal domain (blue).
 - (E) C106A in the N-terminal domain (yellow), K488A in the C-terminal domain (red) and E439C in the N-terminal domain (blue).
 - 15

Example 12: Alignment and sequence similarity of ANL superfamily

A search for potentially related sequences of known structure can be performed by the
 20 `profile.build()` command of MODELLER (manual page <https://salilab.org/modeller/manual/node404.html>). A database containing all PDB sequences, as of October 16, 2014, was searched using the RpMatB wild type sequence. All the aligned sequences (166 in number) were reviewed. For clarity, the sequences / structures / alignments used in Table B and Figure 19 were extracted and information
 25 on these is given below. These all show similar lid domain closing as RpMatB.

Sequence similarity scores were calculated using the EMBOSS NEEDLE program (http://www.ebi.ac.uk/Tools/psa/emboss_needle/).

Protein	PDB ID	Sequence identity	Sequence similarity #
RpMatB, an acyl- and aryl-coenzymeA synthetases	4FUT, chain A	NA	NA
<i>Homo sapiens</i> ACSM2A, an acyl- and aryl-coenzymeA synthetases	3C5E, chain A	30	44
<i>Streptomyces sp</i> CytC1, an adenylation domain of nonribosomal peptide synthetase	3VNQ, chain A	27	43
<i>Photinus pyralis</i> luciferin 4-monooxygenase, a firefly luciferase	4G36, chain A	27	41
<i>Burkholderia xenovorans</i> benzoate-coenzyme A ligase, an acyl- and aryl-coenzymeA synthetases (light pink)	2V7B, chain A	28	41
<i>Bacillus cereus</i> DltA, an adenylation domain of nonribosomal peptide synthetase	3FCC, chain A	21	34
<i>Luciola cruciata</i> luciferin 4-monooxygenase, a firefly luciferase	2D1Q, chain A	29	44

BLOSUM62 matrix, gapopen 10.0, gapextend 0.5, endopen 10.0, endextend 0.5, pairwise alignment.
Suitably the polypeptide component of the ATP sensor of the invention comprises sequence corresponding to one of the polypeptides in the above table.

- 5 **Sequence alignment** was determined, using structure/sequence features and the **salign** command of **MODELLER** (manual page <https://salilab.org/modeller/manual/node313.html>).

10	_aln.pos	10	20	30	40	50	60			
	4futA	-----	-----	MNANL	FARLFD--	LDDPH-----	LAIET--AAGD--ISY			
	3c5eA	QWGHQ	EVPA-----	-----	KFN	FASDVL	DHWADMEKAGKRPPSPALWVWNGKGKELMWNF			
	3vnqA	-----	-----	TSDPI	ARNSDLVSL	FREVAATAP-----	ERTALSA--EDDR--ISY			
	4g36A	-----	AKNIKKGPAPFY--	PLEDGTAGEQLHKAMKRYAL---	VPGTIAFTDAHIEVN--	ITY				
15	2v7bA	-----	-----	LFNFAAYLFRLNETRA-----	-----	GKTAYID--DTGS--TTY				
	3fccA	-----	-----	MKLLEQIEKWAETP-----	-----	DQTAFFVW--RDAK--ITY				
	2dlqA	-----	E-NIVVGPKPFY--	PIEEGSAGTQLRKYMERYAK----	L-	GAIAFTNAVTVGD--YSY				
	_consvd									
20	_aln.p	70	80	90	100	110	120	130		
	4futA	AELVARAGR	RVANVLVA-RGLQVGD	RVAQTE-SVEALVLYLAT	VRAGGVYLP	LNTAYTL	HELDYFIT			
	3c5eA	RELSSENSQQAANVLSGACGLQ	RGRDRAVAVVLP	RPVPEWWLVILGCI	RAGLIFMPGTIQ	MKSTDILYRLQM				
	3vnqA	GRLDAWSDAVARTLLA-EGVRP	GDRAVALRMS	PGAEI	VAI	LA	ILKCGAAYVP	VDLRNPVSRSDFILAD		
	4g36A	AEYFEMSVRLAEAMKR-YGLNTN	HRIVVCS	ENSLQFFMPVLGAL	FIGVAVAPANDI	YNERELLNSMNI				
25	2v7bA	GELERARRFASALRT-LGVHPEERILLV	MLDVTALPVAFLGALYAGVVP	VVANTLLTPADYVYMLTH						
	3fccA	KQLKEDSDALAHWISS-EYPDDRSPIMVYGHMQPEMIINFLGCVKAGHAYI	IPVDLSIPADR	VQRI	AE					
	2dlqA	AEYLEKS--LGKALQN-YGLVVDGRIALCSENCEEFFIPVIAGLFIGVGVAPTNEIYTLREL	VHSLGI							
	_consvd					*				
30	_aln.pos	140	150	160	170	180	190	200		
	4futA	AEP-IVVCDPS----	RDGIAAIAAVGATVETL	GPDGR---	GSLT	DAAAGAS-----	EAFATIDRGA			
	3c5eA	SKAKAIVAGDEVIQEVDTVASECPSLRIKLLVS-EKSCDGWLNFKKLLNEAS-----	TTHHC	VETGS						
	3vnqA	SGASALIG--EP--HEGCA-----	VTRVVRT-----	AAVAECKD--A-----	E-----	APGPG				
	4g36A	SQPTVVFVSKKGLQKILNVQKLP	IIQKIIIMSDKTDYQGFQSMYTFV	TSHLPPGFN	EYDFVP	ESFDR				
35	2v7bA	SHARAVIASGALVQNV	TQALESAG--CQLIVSQP-----	LAPLFEELIDAAA-----	PAAKAAATGC					
	3fccA	SGAKLLLSATAV--TVTDL-----	PVRIVSE-----	DNLKDIFFTHK-----	GNTPNPEHAVK					
	2dlqA	SKPTIVFSSKKGLDKVITVQKTVT	TIKTIVILDSKVDYR	GYQCLDTFI	KRNTPPGFQASSFKT	VEVDR				
	_consvd									
40	_aln.pos	210	220	230	240	250	260	270		
	4futA	D-DLAAILYTS	GTGRSGAMLSHDN	LASNSLT	LDY--	WRFTPDDVLIHALPIYHTHGLFVAS	NSVT			
	3c5eA	Q-EASAIYFTSGTSGLPKMAEHSYSS	LGLKAKMDAG--	WTGLQASDIMWTISDTGWI	LNILCS	LMEP				
	3vnqA	AEDMAYVIYTS	GTGNPKGVPVRHANV	LALLAGAPSV--	FDFSGDDR	WLLFHSLS	FDFSV-WEIWGA			
	4g36A	DKTIALIMNS----	LPGKVALPHRTAC	VRFSHARDPI	FGNQII	PD	TAILSVVPFHHGFGMFT--	TLGY		
45	2v7bA	D-DIAFWLYS----	KPKGT	VHTHANLYWTAELYAK	PI	L	GIA--	ENDVVFSAAKLFFAYGLG	NGLTFP	
	3fccA	GDENFYIIYTSG----	PKGVQITYNCLVS	FTKWAVED--	FNLQTGQVFLN	QAPFS	FDLSV-MDIYPS			
	2dlqA	KEQVALIMNSS----	LPGKVQLTHENI	VTRFSHARDPIYGNQVSPG	TAULTVVP	FHHGFGMFT--	TLGY			
	_consvd									
50	_aln.pos	280	290	300	310	320	330	340		
	4futA	LFARGSMIFLP----	FDPD-ILD	LMA--	RATVLMGVPTFY	TRLLQSPRLT-	ETTG	HMR	L	FISGSAPLL
	3c5eA	WALGACTFVHLLP-KFDPLVILK	TLS	SYPIKSM	MGAPIVYR	MLLQ	Q--	DLSSYK	FPHLQNCVTVGESLL	
	3vnqA	FSTGAELVVLPHWAARTPEQYLAVI	IDRGVT	VINQTP	TAFLALTEAAVR	GRDRV	SLGR	YVIF	GG	
	4g36A	LICGFRVVL--Y-RFEEEL	FLRS	LQDKI	QSALLVPTL	FSFFAK	STL	SNLH	IASGGAPLS	
55	2v7bA	LSVGATAILMA-E-RPTADAI	FARLQV	HRPTV	FYGVPTLYAN	M	LVSP	NLP	PARADVAIRICTSAGEALP	
	3fccA	LVTG	GTLW	AIKDMI	ARPKDL	FASLEQSDI	QVWTSTPS	FAEMCLMEAS	FS	ESMLPNMKTFLFCGEVLP
	2dlqA	LICGFRVVM--T-KFDEET	FLKTLQ	DYKCTSVI	LVPTL	FAILNK	SELLNKYDLS	NL	VEIASGGAPLS	
	_consvd				*				*	
60	_aln.pos	350	360	370	380	390	400			
	4futA	ADTHREWSATGH----	AVLERYGMTET-NMNTSNPYD	GD	RV-----	PGAVGPALPGV	SARVTDPE			
	3c5eA	PETLENWRAQTGL----	DIRE	SYGQ	TET-GLTC	MVSKTMKIK-----	PGYMGTAASCYD	VQIID-D		
	3vnqA	APMLRPWAKAFGLDRP	RLVNGYGITET	TVFTT	FE-----	EITEAYLA-QDASI	I	GRALPSFGTRVVG-D		
	4g36A	KEVG	EAVAKRFHLP--GIRQGYGLTET	SAILITPE-GDDK-----	PGAVGKVV	PFEAK	VVDLD			
65	2v7bA	REIGERFTA	HFGC--EILDGIG	STEMLHI	FLSNRA-GAVE-----	YGT	TGRPVPGY	EIELRD-E		
	3fccA	NEVARKLIERFP--KATIMNTYGP	TEATVAVTGI--	HVTEEVLDQYK	SLPVG	GYCKSD	CRLLIMK-E			
	2dlqA	KEVG	EAVARRFNLP--GVRQGYGLTET	TS	AIITPE-GDDK-----	PGASGKVVPL	FKAKVIDLD			
	_consvd			*	*		*			
70	_aln.p	410	420	430	440	450	460	470		
	4futA	TG-ELPRGDIG	MIEV-----	GNPVF-GYWR	MP--	TSEFRDD--	GFFITGDLG-	IDERGYVHILGR		
	3c5eA	KGNVLP	PGTEGDIGIRVK	PIRPIGIFSGY	V	DNPDKTAANIRGD--	FWLLGDRGI	KEDGYFQFMGR		
	3vnqA	DGRD	VAPGETGELWLS----	GAQLAEGYLR	RP	PELTAEKFPEV--	RYRTGDLV	SELPDGRFAYEGR		
	4g36A	TGKTLGVNQR	GELCVR----	GPMIMSGYV	NNPEATNALIDK	---	GWLHSGDI	AYWDEDEHFFIVDR		

```

2v7bA    AGHAVPDGGEVGDLYIK-----GPSAAVMYWNREKSRATFL-G---EWIRSGDKYCRLPNGCYVYAGR
3fccA    DGTIAPDGEKGEIVIV-----GPSVSVGYLGSPELTEKAFTMIDGERAYKTGDAGYVE-NGLLFYNGR
2d1qA    TKKSLGPNRRGEVCVK-----GPMLMKGYVNNPEATKELIDEE---GWLHTGDIGYYDEEKHFFIVDR
   _consvd      *                               *                               **                               *
5
   _aln.pos    480          490          500          510          520          530          540
4FutA      G-DLVITGGFNVYP-EIESEIDAMPGVVESAVIGVPH---ADF-GEGVTAVVVRD--GAT--I-----
3c5eA      ADDIINSSGYRIGPSEVENALMEHPAVVETAVISSPD---PVRGEVVKAFVVLA-----SQFLS
3vnqA      ADLQIKLRGYRIELSDIETAVRRHDDV-VDAVTVREFKPGD---LRLVCAYVAREGSAT-----
10 4g36A      LKSLIKYKGYQVAPAELESILL-----DAGVAGLPD---DDA-GELPAAVVVLE-----T-----
2v7bA      SDDMLK----YVSPVEVEMVLVQHDAVLEAAVVGVDH---GG--LVKTRAFVVLKREFAP--S-----
3fccA      LDFQIKLHGYRMELEEIEHHLRACSYVEGAVIVPIKK--GEK---YDYLLAVVVPGEHSFEKE-----
2d1qA      LKSLIKYKGYQVPPAELESVLLQHPSIFDAGVAGVPD---PVA-GELPGAVVVLESCKNM--T-----
   _consvd      *
15
   _aln.pos    550          560          570          580          590          600
4FutA      -----DE--AQVLHGLDGLAFMP-----VIFVDDLPRNTMGAVQ-NVLRETYDIY
3c5eA      HDPEQLTKELQQHVKS---TAPYKYPRKIEFVL-NLPKTVTGKIQRAKLRDKEW-K
3vnqA      -----TARELRNHIKTLTPAYMHPA-RYLPLPLPRTVNGKVDRAAVARSW---
20 4g36A      -----EKEIVDYVAS---QVTTAKKLGGVVFVDEVPKGLTGKLDARKIREILIK-
2v7bA      -----EILAEELKAFVKDRLAPHKYPR-DIVFVDDLPKTATGKIQRFKLRE-----
3fccA      -----FKLTSAIKKELNERLPNYMIPR-KFMYQSSIPMTPNGKVDRKKLLSEVTA-
2d1qA      -----EKEVMDYVAS---QVSNAKRLRGVRFVDEVPKGLTGKIDGRAIREILK--
   _consvd      *           *
25

```

Structural similarity of ANL superfamily members with RpMatB

The RMSD was calculated for the same set of ANL superfamily members as above with RpMatB, using the align command in PyMOL (<http://www.pymolwiki.org/index.php/Align>). PyMOL will first do a sequence alignment, followed by a structural superposition, and then carries out cycles of refinement in order to reject structural outliers found during the fit. In the end, the minimized RMSD between the aligned residues using all atoms or backbone atoms is calculated.

Protein	PDB ID	RMSD (Å), All atoms	RMSD (Å), Backbone atoms
RpMatB, an acyl- and aryl-coenzymeA synthetases	4FUT, chain A	NA	NA
<i>Homo sapiens</i> ACSM2A, an acyl- and aryl-coenzymeA synthetases	3C5E, chain A	2.260 (1924 atoms)	2.059 (1260 atoms)
<i>Streptomyces sp</i> CytC1, an adenylation domain of nonribosomal peptide synthetase	3VNQ, chain A	1.903 (1937 atoms)	1.788 (1275 atoms)
<i>Photinus pyralis</i> luciferin 4-monooxygenase, a firefly luciferase	4G36, chain A	2.098 (2044 atoms)	1.969 (1373 atoms)
<i>Burkholderia xenovorans</i> benzoate-coenzyme A ligase, an acyl- and aryl-coenzymeA synthetases (light pink)	2V7B, chain A	1.676 (2003 atoms)	1.583 (1328 atoms)
<i>Bacillus cereus</i> DltA, an adenylation domain of nonribosomal peptide synthetase	3FCC, chain A	2.565 (1561 atoms)	2.458 (1027 atoms)
<i>Luciola cruciata</i> luciferin 4-monooxygenase, a firefly luciferase	2D1Q, chain A	2.490 (2299 atoms)	2.390 (1563 atoms)

Suitably the polypeptide component of the ATP sensor of the invention comprises sequence corresponding to one of the polypeptides in the above table.

Although illustrative embodiments of the invention have been disclosed in detail herein, with reference to the accompanying drawings, it is understood that the invention is not limited to the precise embodiment and that various changes and
 5 modifications can be effected therein by one skilled in the art without departing from the scope of the invention as defined by the appended claims and their equivalents.

References

- Berman, H. M., J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov and P. E. Bourne (2000). "The Protein Data Bank." Nucleic Acids Res **28**(1): 235-242.
- Brune, M., J. L. Hunter, J. E. Corrie and M. R. Webb (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**(27): 8262-8271.
- Burnstock, G. (2012). "Purinergic signalling: Its unpopular beginning, its acceptance and its exciting future." Bioessays **34**(3): 218-225.
- Chambers, R. W., T. Kajiwaru and D. R. Kearns (1974). "Effect of Dimer Formation of Electronic Absorption and Emission-Spectra of Ionic Dyes - Rhodamines and Other Common Dyes." Journal of Physical Chemistry **78**(4): 380-387.
- Corrie, J. E. T. and J. S. Craik (1994). "Synthesis and Characterization of Pure Isomers of Iodoacetamidotetramethylrhodamine." Journal of the Chemical Society-Perkin Transactions **1**(20): 2967-2973.
- Crosby, H. A., K. C. Rank, I. Rayment and J. C. Escalante-Semerena (2012). "Structure-guided expansion of the substrate range of methylmalonyl coenzyme A synthetase (MatB) of *Rhodopseudomonas palustris*." Appl Environ Microbiol **78**(18): 6619-6629.
- Dale, N. and B. G. Frenguelli (2012). "Measurement of purine release with microelectrode biosensors." Purinergic Signal **8**(Suppl 1): 27-40.
- Dillingham, M. S., K. L. Tibbles, J. L. Hunter, J. C. Bell, S. C. Kowalczykowski and M. R. Webb (2008). "Fluorescent single-stranded DNA binding protein as a probe for sensitive, real-time assays of helicase activity." Biophys J **95**(7): 3330-3339.
- Feng, C., S. Dai and L. Wang (2014). "Optical aptasensors for quantitative detection of small biomolecules: A review." Biosens Bioelectron **59C**: 64-74.
- Finn, R. D., A. Bateman, J. Clements, P. Coggill, R. Y. Eberhardt, S. R. Eddy, A. Heger, K. Hetherington, L. Holm, J. Mistry, E. L. Sonnhammer, J. Tate and M. Punta (2014). "Pfam: the protein families database." Nucleic Acids Res **42**(Database issue): D222-230.
- Guan, Z., N. A. Yates and R. Bakhtiar (2003). "Detection and characterization of methionine oxidation in peptides by collision-induced dissociation and electron capture dissociation." J Am Soc Mass Spectrom **14**(6): 605-613.
- Gulick, A. M. (2009). "Conformational dynamics in the Acyl-CoA synthetases, adenylation domains of non-ribosomal peptide synthetases, and firefly luciferase." ACS Chem Biol **4**(10): 811-827.
- Huang, Y., B. Niu, Y. Gao, L. Fu and W. Li (2010). "CD-HIT Suite: a web server for clustering and comparing biological sequences." Bioinformatics **26**(5): 680-682.
- Hughes, A. J. and A. Keatinge-Clay (2011). "Enzymatic extender unit generation for in vitro polyketide synthase reactions: structural and functional showcasing of *Streptomyces coelicolor* MatB." Chem Biol **18**(2): 165-176.
- Idzko, M., D. Ferrari and H. K. Eltzschig (2014). "Nucleotide signalling during inflammation." Nature **509**(7500): 310-317.

- Imamura, H., K. P. Nhat, H. Togawa, K. Saito, R. Iino, Y. Kato-Yamada, T. Nagai and H. Noji (2009). "Visualization of ATP levels inside single living cells with fluorescence resonance energy transfer-based genetically encoded indicators." Proc Natl Acad Sci U S A **106**(37): 15651-15656.
- Johnson, K. A., Z. B. Simpson and T. Blom (2009). "FitSpace explorer: an algorithm to evaluate multidimensional parameter space in fitting kinetic data." Anal Biochem **387**(1): 30-41.
- Johnson, K. A., Z. B. Simpson and T. Blom (2009). "Global kinetic explorer: a new computer program for dynamic simulation and fitting of kinetic data." Anal Biochem **387**(1): 20-29.
- Kunzelmann, S. and M. R. Webb (2009). "A biosensor for fluorescent determination of ADP with high time resolution." J Biol Chem **284**(48): 33130-33138.
- Kunzelmann, S. and M. R. Webb (2010). "A fluorescent, reagentless biosensor for ADP based on tetramethylrhodamine-labeled ParM." ACS Chem Biol **5**(4): 415-425.
- Leatherbarrow, R. J. (2009). GraFit Version 7. Horley, UK, Erithacus Software Limited.
- Maiti, R., G. H. Van Domselaar, H. Zhang and D. S. Wishart (2004). "SuperPose: a simple server for sophisticated structural superposition." Nucleic Acids Res **32**(Web Server issue): W590-594.
- Okoh, M. P., J. L. Hunter, J. E. Corrie and M. R. Webb (2006). "A biosensor for inorganic phosphate using a rhodamine-labeled phosphate binding protein." Biochemistry **45**(49): 14764-14771.
- Patergnani, S., F. Baldassari, E. De Marchi, A. Karkucinska-Wieckowska, M. R. Wieckowski and P. Pinton (2014). "Methods to monitor and compare mitochondrial and glycolytic ATP production." Methods Enzymol **542**: 313-332.
- Selwyn, J. E. and Steinfel.Ji (1972). "Aggregation Equilibria of Xanthene Dyes." Journal of Physical Chemistry **76**(5): 762-&.
- Vogt, A. D., N. Pozzi, Z. Chen and E. Di Cera (2014). "Essential role of conformational selection in ligand binding." Biophys Chem **186**: 13-21.
- Wilkins, M. R., E. Gasteiger, A. Bairoch, J. C. Sanchez, K. L. Williams, R. D. Appel and D. F. Hochstrasser (1999). "Protein identification and analysis tools in the ExPASy server." Methods Mol Biol **112**: 531-552.
- Zhou, Y., Z. Xu and J. Yoon (2011). "Fluorescent and colorimetric chemosensors for detection of nucleotides, FAD and NADH: highlighted research during 2004-2010." Chem Soc Rev **40**(5): 2222-2235.

Example 13 - Additional Mutations

The concentration range of ATP, measurable with a fluorescent reagentless biosensor, an adduct of two tetramethylrhodamines with MatB from *Rhodopseudomonas palustris*, has been increased. Mutations were introduced into the binding site to modify ATP binding but maintain the concomitant fluorescence signal. Using this signal, the effect was monitored for mutations in different parts of the binding site. Out of these, three variants were characterized, each with a single extra mutation in the phosphate-binding loop. Two variants (T167A and T303A) weakened the binding, changing the dissociation constant from the parent's 6 μM to 123 μM and 42 μM , respectively but having a fluorescence change of ~ 3 -fold on ATP binding. Kinetic measurements showed that the main effect of these mutations was as an increase in dissociation rate constants. These variants widen the range of ATP concentration that can be measured readily by this biosensor to $>100 \mu\text{M}$. In contrast, a third variant, S170A, decreases the dissociation constant of ATP to 3.6 μM and has a fluorescence change of 4.2 on binding ATP. This variant also shows while decreased ADP affinity,

thereby increasing selectivity of ATP over ADP to >200-fold. This had advantages over the parent by increasing sensitivity as well as increasing selectivity during ATP measurement in which ADP is present.

5 **Materials.** ATP, ADP and GTP were obtained from Sigma-Aldrich. 5-iodoacetamidotetramethylrhodamine (5-IATR) (Corrie and Craik 1994) was a gift from J. Corrie (NIMR, London) or obtained from Anaspec (CA).

Plasmids. Plasmid pRhoRpMatB (plasmid pTEV5 containing the coding sequence of
10 RpMatB (Genbank accession number CAE25665.1) with the point mutations C106A, R286C, Q457C, K488A and an N-terminal His₆-tag) was as described, and now termed pTEV5_ RpMatB_1 (Vancraenenbroeck and Webb 2015). The QuikChange site-directed mutagenesis protocol (Stratagene) was used for single-site mutations of the pRhoRpMatB plasmid. Plasmids were sequenced (GATC Biotech) to confirm the
15 presence of the mutations. The three variant plasmids, containing additional S170A, T167A or T303A mutations, are termed pTEV5_ RpMatB_2, pTEV5_ RpMatB_3, and pTEV5_ RpMatB_4 respectively.

Preparation of RpMatB variants, labeled with tetramethylrhodamine.
20 Protein expression in *Escherichia coli*, purification and labeling with tetramethylrhodamine were as described (Vancraenenbroeck and Webb 2015). The concentrations of labeled RpMatB variants were determined using the extinction coefficient of the protein at 280 nm, calculated from the sequence via ExPASy ProtParam (Wilkins, Gasteiger et al. 1999) and the extinction coefficients of
25 tetramethylrhodamine: $\epsilon_{280\text{nm}}$ (31000 M⁻¹ cm⁻¹) and $\epsilon_{528\text{nm}}$ (52000 M⁻¹ cm⁻¹) (Corrie and Craik 1994)

Double labeling was confirmed for the T167A and T303A variants by mass spectrometry. Unlabeled (His₆/C106A/T167A/R286C/Q457C/K488A)RpMatB had a mass of 57304.2 Da. (5-ATR)₂-(His₆/C106A/T167A/R286C/Q457C/K488A)RpMatB,
30 had a mass of 58187.3 Da, conforming to the theoretical molecular weight of the unlabeled protein with two 5-ATRs attached (57304.2 + 2 x 441.5). Unlabeled (His₆/C106A/R286C/T303A/Q457C/K488A)RpMatB had a mass of 57300.9 Da. (5-ATR)₂-(His₆/C106A/T167A/R286C/Q457C/K488A)RpMatB, had a mass of 58186.0 Da.

35

Absorbance and fluorescence measurements. Absorbance was measured on a Jasco V-550 UV-Vis Spectrophotometer. Fluorescent measurements were obtained on

a Cary Eclipse spectrofluorometer (Agilent Technologies), using a 3-mm pathlength quartz cuvette (Hellma), unless otherwise mentioned. Excitation and emission slits were set at 5 nm. Protein and nucleotide concentrations and buffer conditions are given in the figure legends. For titrations with Rho-MatB variants, excitation was at 553 nm, emission at 571 nm.

Data from titrations to measure nucleotide binding were corrected for dilution and analyzed with a quadratic binding curve using Grafit software (Leatherbarrow 2009):

$$F = F_{\min} + (F_{\max} - F_{\min}) \times \frac{(K_d + L + P - \sqrt{(K_d + L + P)^2 - 4 \times P \times L})}{2 \times P}$$

where P and L are the total concentrations of protein and ligand, respectively, K_d is the dissociation constant, and $F_{\min}$ and $F_{\max}$ are the fluorescence intensities of the free and ligand-bound protein, respectively.

Stopped-flow measurements. These were carried out using a HiTech SF-61DX2 apparatus (TgK Scientific, Bradford-upon-Avon, UK) with a xenon-mercury lamp and operated by Kinetic Studio software (TgK Scientific). The excitation wavelength was 545 nm and there was an OG570 cut-off filter (Schott Glass) on the emission. The concentrations in the text and figures are those in the mixing chamber. Data were fitted to theoretical equations using the Kinetic Studio software.

Approach

In addition to two cysteine mutations, the biosensor suitably has a C106A mutation to eliminate background labeling at that cysteine and suitably has a K488A mutation to block the adenylation half-reaction of ATP and malonate to malonyl-AMP and pyrophosphate. The resulting protein adduct, termed Rho-MatB had essentially no enzyme activity, but bound ATP with a K_d of 6 μ M. This ATP biosensor, termed Rho-MatB, couples ATP binding to a 3.7-fold increase in fluorescence intensity and measures ATP concentrations in the low micromolar range. Of importance is the fact that this biosensor is greatly selective for ATP over ADP.

We have now developed further variants of Rho-MatB to change the affinity for ATP, while maintaining a significant fluorescence change on ATP binding. In this way the measurable range ATP concentration has been changed. If the affinity is decreased to allow measurements of higher ATP concentrations, a low, sub-stoichiometric concentration of biosensor can be used, minimizing biosensor usage (Solscheid, Kunzelmann et al. 2015). Alternatively, a biosensor for measurements of lower ATP

concentrations by increasing the affinity of Rho-MatB for ATP may be increase sensitivity over the parent biosensor.

The method chosen to change affinity was to mutate amino acids identified from the crystal structure as interacting with ATP (Crosby, Rank et al. 2012). Mostly these
5 looked unlikely to affect the lid closure, that is they are not positioned at the hinge or involved in interactions across the cleft, nor might they affect the rhodamines directly. In doing this, there was essentially a survey of the active-site amino acids and the effect of (mainly) alanine mutations on ATP affinity, readily measured using the rhodamine fluorescence signal. This survey resulted in three new variants, one tighter and two
10 weaker in ATP binding, potentially resulting in the ability to measure ATP from sub-micromolar to >100 μ M.

Identification of amino acid residues around MgATP

Amino acid residues were considered that are important for the structural integrity and
15 function of RpMatB. Our first goal was the RpMatB ATP binding site. Residues were chosen that were within 0.6 nm of MgATP as determined using the crystal structure of RpMatB in the MgATP-bound conformation (Crosby, Rank et al. 2012). In addition, sequence homology amongst the ANL family was used (Figure 26 (Figure S1)), so that highly conserved amino acids were avoided as positions to mutate. Some of these
20 residues belong to the A3, A4, A5, A7, A8 and A10 core motifs, which are known to play an important role in adenylate-forming enzymes (Marahiel, Stachelhaus et al. 1997). The positions of these motifs are shown in Figure 21 and details of the motifs are in Figure 26 (Figure S1). The resulting positions that were subsequently mutated successfully are shown in Table C with their motifs along with which part of the ATP
25 structure is closest.

Preparation and testing of variants of Rho-MatB

30 To examine the effects of substituting these residues on Rho-MatB's function, the individual residues were mutated to alanine on a background of (His₆/C106A/R286C/Q457C/K488A)RpMatB (Vancraenenbroeck and Webb 2015). In addition two variants had the T167S and T303S mutations, so potentially still allowing binding to the triphosphate of ATP. The mutants were expressed and purified as for the
35 parent RpMatB. Only variants that gave reasonable expression were continued to purification and labeling with 5-iodoacetyltetramethylrhodamine (5-IATR).

The labeled proteins were then tested for the fluorescence change upon ATP binding in the presence of Mg^{2+} (Table C). For variants showing a significant fluorescence change, the affinity for ATP was determined by titration (Table C). None of the variants of Rho-MatB responded to GTP under the same conditions (data not shown).

5

Table C: Fluorescence change and affinity for ATP binding to variants of Rho-MatB.

Data are from a survey in 50 mM Tris.HCl pH 7.5, 150 mM NaCl, 10 mM $MgCl_2$, 0.3 mg ml^{-1} bovine serum albumin at 20 °C but without complete optimization of labeling and purification. The motifs are shown in Figure 21. The ratio of fluorescence at saturating ligand to that in the absence of ligand (F_+/F_-) and the ATP dissociation constant (K_d) were determined by titrations as in Figure 24. The parent Rho-MatB is the protein without any binding site mutations.

10

Variant	Motif	Position near ATP	F_+/F_-	K_d for ATP (μM)
Parent			2.9	8.8
T163A	A3	γ -P	1.8	297 ± 27
S164A	A3	α -P	2.7	17.4 ± 0.4
T166A	A3	β -P	2.0	156 ± 7
T167A	A3	γ -P	2.9	149 ± 7
T167S	A3	γ -P	3.3	52 ± 1
R169A ^b	A3	γ -P	2.8	33 ± 1
S170A	A3	γ -P	3.9	4.0 ± 0.1
H207A ^b	A4	α -P	1.4	224 ± 18
T208A		α -P	1.7	52 ± 5
S277A		adenine	0.9 ^a	--
E298A		adenine	1.1 ^a	--
R299A		adenine	1.7	2.4 ± 0.1
Y300A	A5	adenine	2.1	11.1 ± 0.5
G301A	A5	adenine	0.8 ^a	--
M302A	A5	ribose	1.5	50 ± 5
T303A	A5	α -P	2.6	42 ± 1
T303S	A5	α -P	1.7	56 ± 3

I394A		adenine	1.1 ^a	--
Q490A	A10	β-P, γ-P	2.8	66 ± 3

^a These ratios are from a single measurement at 750 μM ATP. As the fluorescence change was very low, no titration was performed.

^b Mutations in the corresponding residue of MatB from *Rhizobium trifolii* have been studied previously (An, Lee et al. 1999).

5

The Table shows that the mutations that least perturbed the fluorescence change on ATP binding were all located near the triphosphate. However, the mutations that caused the greatest decrease in affinity (T163A, T166A, T167A, S170A) were all in motif A3, the phosphate binding loop. This suggests that changes to this loop decreases

10

binding but still allows the full conformation change, which controls the interaction between rhodamines and therefore the fluorescence change. In contrast, all the mutations that cause almost complete loss of the fluorescence change are close to the adenosine.

Three variants that retained large changes in fluorescence signals but had potentially

15

Weak binding variants, T167A and T303A

Two labeled variants which showed a large fluorescent increase upon MgATP-binding, but also showed a significant reduction in ATP affinity, relative to the parent Rho-RpMatB, were T167A and T303A. Their side chains are involved in hydrogen bonding between RpMatB and the triphosphate of ATP (Figure 21B). Their fluorescence spectra with and without ATP are in Figure 22A and B. The absorbance spectra (Figure 27 (Figure S2A and B)) show the changes on ATP binding typical for tetramethylrhodamine switching between mainly stacked to unstacked configuration. This behaviour was seen previously for the parent Rho-MatB and other proteins labeled with two rhodamines (Chambers, Kajiwara et al. 1974; Hamman, Oleinikov et al. 1996; Okoh, Hunter et al. 2006; Kunzelmann and Webb 2010; Vancraenenbroeck and Webb

25
30

2015). The affinity for ATP and ADP was determined by measuring the fluorescence at different concentrations of the nucleotide (Figure 23A and B). Both show similar fluorescent changes with ATP to the parent Rho-MatB, but the dissociation constants (K_d) are much higher (Table D). In contrast, the affinity of ADP was much less affected,

remaining very weak (Table D). The change in ATP affinity presumably reflects the importance of hydrogen-bonding interactions of the triphosphate.

Table D: Fluorescence changes, dissociation constants and rate constants for ATP binding to variants of Rho-MatB.

Both the ratio of fluorescence at the saturating ligand to that in the absence of ligand (F_+/F_-) and the dissociation constants (K_d) were obtained from the data in Figure 23. Association kinetics were from Figure 25: k_{on} is the second order association rate constant, k_{off} is the dissociation rate constant and k_2 is the value obtained for the conformation change. The K_d ratio is for ADP/ATP. Values for the parent biosensor were taken at 25 °C (Vancraenenbroeck and Webb 2015). Note: the protein was purified further than that used in Table C.

Variant	F_+/F_-	K_d (ATP) (μM)	k_{on} ($\mu\text{M}^{-1} \text{s}^{-1}$)	k_{off} (s^{-1})	k_2 (s^{-1})	K_d (ADP) (μM)	K_d ratio
Parent	3.9 ± 0.2	6.4 ± 0.6	1.8 ± 0.02	8.2 ± 0.2	0.88 ± 0.13	428 ± 50	69
T167A	3.2 ± 0.1	123 ± 4	1.83 ± 0.05	87 ± 4	0.56 ± 0.01	900 ± 28	7.3
T303A	2.7 ± 0.1	42 ± 2	1.60 ± 0.06	16 ± 4	0.60 ± 0.02	1835 ± 114	44
S170A	4.2 ± 0.1	3.8 ± 0.1	1.83 ± 0.10	9 ± 6	0.64 ± 0.02	887 ± 25	233

15 Calibration curves and effect of ADP

In assays in which ATP is being measured, ADP is likely to be present also and may be at relatively high concentration. In order to show how the presence of ADP affected the response of the variants to ADP, calibration curves were constructed at different levels of ADP (Figure 24). The total concentration of nucleotide (ATP + ADP) was held constant, to mimic partial interconversion of the two nucleotides. In the absence of ADP the response was approximately linear up to a concentration of ATP, 50% of the K_d -value: of course, this is the first part of the binding curve in Figure 23. For both the weak-binding variants there was still a good response with 500 μM ADP present, the highest concentration tested. The response was ~50% of that without ADP.

25

Association kinetics of the weak-binding variants

Although a complete kinetic analysis was beyond the scope of this work, it was important to gauge how the main binding kinetics changed from the parent and so

relate that to response time with ATP. Association transients were obtained under pseudo-first order conditions by rapid mixing of each Rho-MatB variant with a large excess of ATP and following fluorescence with time (Figure 25 and Figure 28 (Figure S3)). Qualitatively the weak-binding variants were similar to the parent (Vancraenenbroeck and Webb 2015), showing two distinct phases. The rate constant obtained are in Table D.

The fast phase (Figure 25) had a rate constant that increased linearly with ATP concentration and was interpreted as binding. The association rate constant from the gradient only varied little. The dissociation rate constant could be estimated from the intercept and is the main difference between variants.

The slow phase (Figure 28 (Figure S3)) had a rate constant that did not change significantly with concentration: an average value is shown in Table D. This was interpreted, based on data with the parent Rho-MatB, as a conformation change to the apoprotein prior to binding. This rate constant was very similar to the parent, suggesting that the extra mutations did not affect the conformation change significantly.

Overall, the main effect of the active site mutations is on the dissociation kinetics, which then modify the affinity.

Properties of the tight-binding variant, S170 A

In contrast to the two mutations on the phosphate binding loop, described above, the T170A variant strengthens binding. This amino acid does not appear to interact with the triphosphate directly, unlike the hydrogen bonding of T167A and T303A. ATP and ADP binding was measured, as described above for the other variants.

The changes to the fluorescence spectra (Figure 22 C) and absorbance spectra (Figure 27 (Figure S2C)) are similar to those shown above and for the parent Rho-MatB (Vancraenenbroeck and Webb 2015). The fluorescence change is slightly higher than the parent (Table D). Titration with ATP gave a K_d of 3.8 μ M, approximately two fold tighter than the parent (Figure 23C). However, because ADP binding was apparently weakened, the relative affinity of ATP and ADP is much greater. That can be seen in Figure 23C: at 100 μ M nucleotide, ATP gives close to maximum fluorescence, ADP shows almost unchanged from the apoprotein. In the calibration curves (Figure 24C), 50 μ M ADP has no significant effect on the signal with ATP.

Like the two, weak-binding variants, the association kinetics were measured for the S170A Rho-RpMatB (Figure 25) to obtain the second order association rate constant (Table D). This rate constant was very similar for all three variants and the parent Rho-MatB, The second phase, likely to represent a conformation change prior to ATP

binding as described above, was also very similar to the other variants and parent rho-MatB (Figure 28 (Figure S3) and Table D). The main difference between these is in the intercept from the fast phase (Figure 25B), representing the dissociation rate constant, which therefore controls the size of the equilibrium dissociation constant.

5

Conclusions

The mutations in the binding site give us insights into the contributions of specific residues to ATP binding. The fluorescence labelling provides a signal to measure the effect of the amino acid side chain on the affinity of a protein-ligand complex. In particular, the effect of three mutations in the phosphate binding loop were studied in detail, as these gave desirable modifications to the biosensors of the invention for measuring ATP, such as the exemplary Rho-MatB. T167A and T303A decrease the affinity and so provide biosensors in the tens of micromolar ATP, up to >100 μ M. These can be used substoichiometrically, for example at <1 μ M, whereby the concentration of ATP is directly correlated to the fraction of the Rho-MatB in the high fluorescence, ATP-bound state.

In contrast, S170A increased ATP affinity, producing a more sensitive probe for ATP with a larger selectivity for ATP over ADP than the parent biosensor. The greater sensitivity comes from a combination of tighter binding and larger fluorescence enhancement.

20

REFERENCES TO EXAMPLE 13

- An, J. H., G. Y. Lee, et al. (1999). "Identification of residues essential for a two-step reaction by malonyl-CoA synthetase from *Rhizobium trifolii*." Biochemical Journal **344**: 159-66.
- 25 Chambers, R. W., T. Kajiwar, et al. (1974). "Effect of dimer formation on the electronic absorption and emission spectra of ionic dyes." Journal of Physical Chemistry **78**: 380-387.
- Corrie, J. E. T. and J. S. Craik (1994). "Synthesis and characterisation of pure isomers of iodoacetamidotetramethylrhodamine." Journal of the Chemical Society, Perkin Transactions I: 2967-2973.
- 30 Crosby, H. A., K. C. Rank, et al. (2012). "Structure-guided expansion of the substrate range of methylmalonyl coenzyme A synthetase (MatB) of *Rhodopseudomonas palustris*." Applied and Environmental Microbiology **78**: 6619-6629.
- Finn, R. D., A. Bateman, et al. (2014). "Pfam: the protein families database." Nucleic Acids Research **42**(Database issue): D222-30.
- 35 Gulick, A. M. (2009). "Conformational dynamics in the Acyl-CoA synthetases, adenylation domains of non-ribosomal peptide synthetases, and firefly luciferase." ACS Chemical Biology **4**: 811-827.

- Hamman, B. D., A. V. Oleinikov, et al. (1996). "Tetramethylrhodamine dimer formation as a spectroscopic probe of the conformation of *Escherichia coli* ribosomal protein L7/L12 dimers." Journal of Biological Chemistry **271**(13): 7568-7573.
- 5 Kasha, M. (1963). "Energy transfer mechanisms and the molecular exciton model for molecular aggregates." Radiation Research **20**: 55-70.
- Kasha, M., H. R. Rawls, et al. (1965). "The exciton model in molecular spectroscopy." Pure and Applied Chemistry **11**: 371-392.
- Kunzelmann, S., C. Solscheid, et al. (2014). Fluorescent biosensors: design and application to motility proteins. Fluorescent Methods Applied to Molecular Motors, Experientia Supplementum C. P. Toseland and N. Fili. Basel, Springer. **10 5**: 25-47.
- 10 Kunzelmann, S. and M. R. Webb (2010). "A fluorescent, reagentless biosensor for ADP based on tetramethylrhodamine-labeled ParM." ACS Chemical Biology **5**: 415-425.
- Leatherbarrow, R. J. (2009). GraFit Version 7. Horley, U.K., Erithacus Software Ltd.
- Marahiel, M. A., T. Stachelhaus, et al. (1997). "Modular peptide synthetases involved in nonribosomal peptide synthesis." Chemical Reviews **97**: 2651-2673.
- 15 Okoh, M. P., J. L. Hunter, et al. (2006). "A biosensor for inorganic phosphate using a rhodamine-labeled phosphate binding protein." Biochemistry **45**: 14764-14771.
- Scholes, G. D. and K. P. Ghiggino (1994). "Electronic interactions and interchromophore electron transfer." Journal of Physical Chemistry **98**: 4580-4590.
- 20 Solscheid, C., S. Kunzelmann, et al. (2015). "Development of a reagentless biosensor for inorganic phosphate, applicable over a wide concentration range." Biochemistry **54**: 5054-5062.
- Vancraenenbroeck, R. and M. R. Webb (2015). "A fluorescent, reagentless biosensor for ATP, based on malonyl-coenzyme A synthetase." ACS Chemical Biology **10**: 2650-2657.
- 25 Wilkins, M. R., E. Gasteiger, et al. (1999). "Protein identification and analysis tools in the ExPASy server." Methods in Molecular Biology **112**: 531-552.

30 All publications mentioned in the above specification are herein incorporated by reference.

CLAIMS

1. An ATP binding molecule comprising a polypeptide which undergoes a conformational change from a first conformation to a second conformation upon
5 binding of ATP, said polypeptide comprising amino acid sequence corresponding to at least amino acids 5 to 497 of SEQ ID NO:1,
said polypeptide having at least 21% sequence identity to SEQ ID NO:1,
said polypeptide having a value of RMSD $<4 \text{ \AA}$ relative to RpMatB in the ATP-bound conformation,
10 wherein said polypeptide comprises a cysteine residue for attachment of at least one reporter moiety, and comprises at least one reporter moiety attached thereto,
wherein said cysteine residue for attachment of at least one reporter moiety is positioned such that said reporter moiety undergoes a change in fluorescence upon changing from said first conformation to said second conformation.
- 15 2. An ATP binding molecule according to claim 1 wherein the polypeptide comprises a first cysteine residue at a position corresponding to a position selected from **R286**, A282, D283, H285, E287, S289, A290, **K385**, L383, G384, I386 and D287 of SEQ ID NO: 1,
20 and wherein the polypeptide comprises a second cysteine residue at a position corresponding to a position selected from
Q457, A456, V458, H460, G461, **G461**, Q457, H460, L462, G464, Q465, **K470**, L466, F469 and M471 of SEQ ID NO: 1.
- 25 3. An ATP binding molecule according to claim 2 wherein the polypeptide comprises a first cysteine residue at a position corresponding to a position selected from **R286** and **K385** of SEQ ID NO: 1,
and wherein the polypeptide comprises a second cysteine residue at a position corresponding to a position selected from **Q457**, **G461** and **K470** of SEQ ID NO: 1.
- 30 4. An ATP binding molecule according to claim 3 wherein the polypeptide comprises a first cysteine residue at a position corresponding to position R286 of SEQ ID NO: 1,
and wherein the polypeptide comprises a second cysteine residue at a position
35 corresponding to position Q457 of SEQ ID NO: 1.

5. An ATP binding molecule according to claim 3 wherein the polypeptide comprises a first cysteine residue at a position corresponding to position K385 of SEQ ID NO: 1,
and wherein the polypeptide comprises a second cysteine residue at a position
5 corresponding to position K470 of SEQ ID NO: 1.
6. An ATP binding molecule according to claim 3 wherein the polypeptide comprises a first cysteine residue at a position corresponding to position R286 of SEQ ID NO: 1,
10 and wherein the polypeptide comprises a second cysteine residue at a position corresponding to position G461 of SEQ ID NO: 1.
7. An ATP binding molecule according to any of claims 1 to 3, said molecule comprising the amino acid sequence of SEQ ID NO:4, SEQ ID NO:5 or SEQ ID NO:6.
15
8. An ATP binding molecule according to any preceding claim, said molecule further comprising at least two tetramethylrhodamine moieties attached thereto.
9. An ATP binding molecule according to claim 8 wherein each of said at least two
20 tetramethylrhodamine moieties is independently selected from the group consisting of 5-tetramethylrhodamine and 6-tetramethylrhodamine.
10. An ATP binding molecule according to claim 1 wherein said cysteine residue for attachment of a reporter moiety is at a position corresponding to a position selected
25 from **E439**, I403, P433, G438, G440, **N492**, K491, V493, R495, E496 and T497 of SEQ ID NO: 1.
11. An ATP binding molecule according to claim 10 wherein said cysteine residue for attachment of a reporter moiety is at a position corresponding to **E439** or **N492** of
30 SEQ ID NO: 1.
12. An ATP binding molecule according to any of claims 10 or 11, said molecule comprising the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:3.
- 35 13. An ATP binding molecule according to any of claims 10 to 12, said molecule further comprising at least one diethylaminocoumarin moiety attached thereto.

14. An ATP binding molecule according to claim 13 wherein said diethylaminocoumarin moiety is independently selected from the group consisting of (*N*-[2-(1-maleimidyl)ethyl]-7-diethylaminocoumarin-3-carboxamide and *N*-[2-(iodoacetamido)ethyl]-7-diethylaminocoumarin-3-carboxamide).

5

15. An ATP binding molecule according to any of claims 1 to 14 wherein said polypeptide comprises an amino acid other than Threonine at the position corresponding to T167 of SEQ ID NO: 1.

10 16. An ATP binding molecule according to claim 15 wherein said polypeptide comprises alanine or serine at the position corresponding to T167 of SEQ ID NO: 1, preferably alanine.

15 17. An ATP binding molecule according to any of claims 1 to 14 wherein said polypeptide comprises an amino acid other than Serine at the position corresponding to S170 of SEQ ID NO: 1.

18. An ATP binding molecule according to claim 17 wherein said polypeptide comprises alanine at the position corresponding to S170 of SEQ ID NO: 1.

20

19. An ATP binding molecule according to any of claims 1 to 14 wherein said polypeptide comprises an amino acid other than Threonine at the position corresponding to T303 of SEQ ID NO: 1.

25 20. An ATP binding molecule according to claim 19 wherein said polypeptide comprises alanine or serine at the position corresponding to T303 of SEQ ID NO: 1, preferably alanine.

21. An ATP binding molecule according to any of claims 1 to 20 wherein said polypeptide comprises an amino acid other than cysteine at the position corresponding to C106 of SEQ ID NO: 1.

30

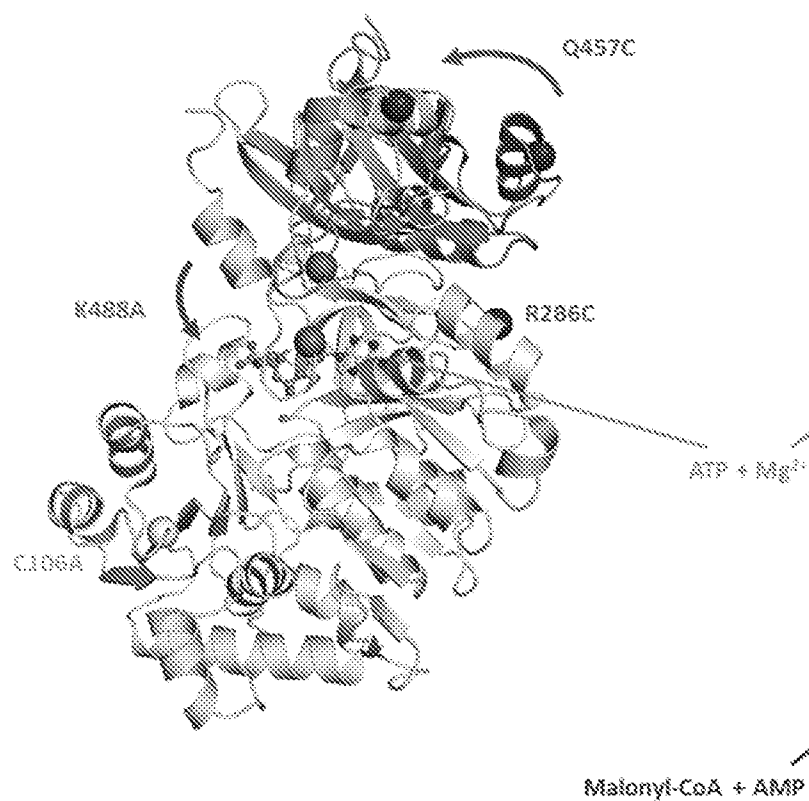
22. An ATP binding molecule according to claim 21 wherein said polypeptide comprises alanine at the position corresponding to C106 of SEQ ID NO: 1. (*C106A*)

35

23. An ATP binding molecule according to any of claims 1 to 22 wherein said polypeptide comprises an active site mutation such that the polypeptide is catalytically inactive for ATP hydrolysis.
- 5 24. An ATP binding molecule according to claim 23 wherein said active site mutation comprises an amino acid other than lysine at the position corresponding to K488 of SEQ ID NO: 1.
- 10 25. An ATP binding molecule according to claim 24 wherein said polypeptide comprises alanine at the position corresponding to K488 of SEQ ID NO: 1. (*K488A*)
26. An ATP binding molecule according to any of claims 1 to 25 further comprising the sequence of SEQ ID NO: 8.
- 15 27. A nucleic acid having a nucleotide sequence encoding the polypeptide portion of an ATP binding molecule according to any of claims 1 to 26.
- 20 28. A method for monitoring changes in ATP concentration in a sample comprising contacting said sample with an ATP binding molecule according to any of claims 1 to 20 and determining changes in conformation of said ATP binding molecule, wherein changes in conformation of said ADP binding molecule indicate changes in the concentration of ATP in said sample.
- 25 29. A method according to claim 28 wherein the conformation of said ATP binding molecule is monitored by measurement of changes in fluorescence of a fluorophore comprised by said ATP binding molecule.
- 30 30. A method according to claim 28 or claim 29 wherein the sample comprises divalent Magnesium ion (Mg^{2+}).
31. Use of an ATP binding molecule according to any of claims 1 to 26 in the determination of ATP concentration in a sample.
- 35 32. An ATP binding molecule or method substantially as disclosed herein.

FIGURE 1

(A)



(B)

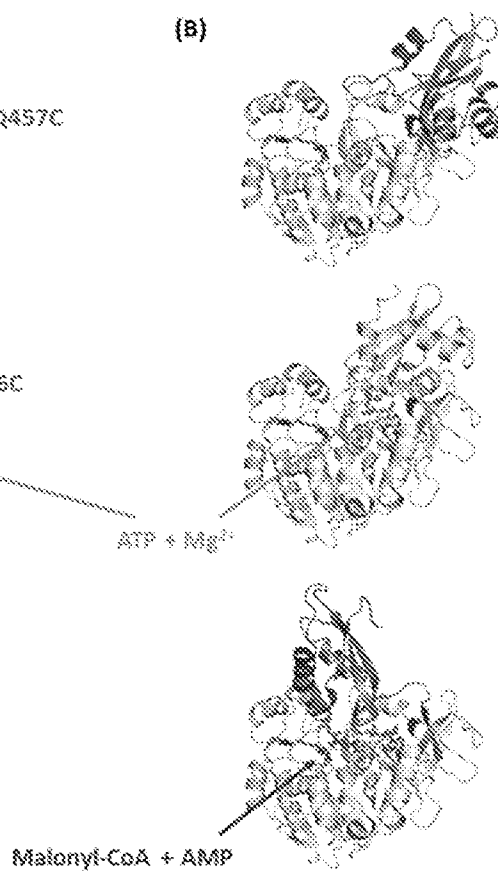
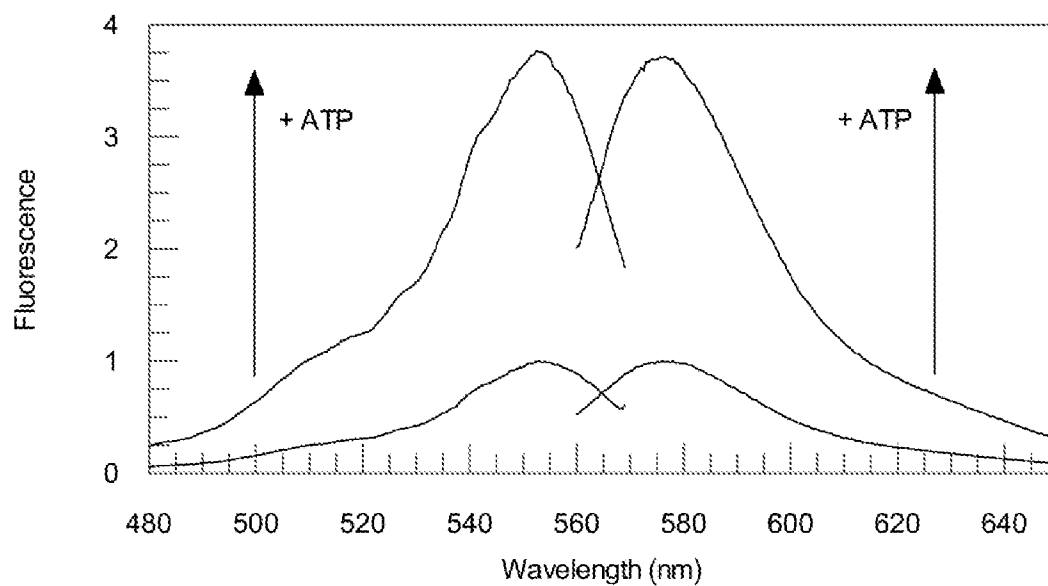


FIGURE 2

(A)



(B)

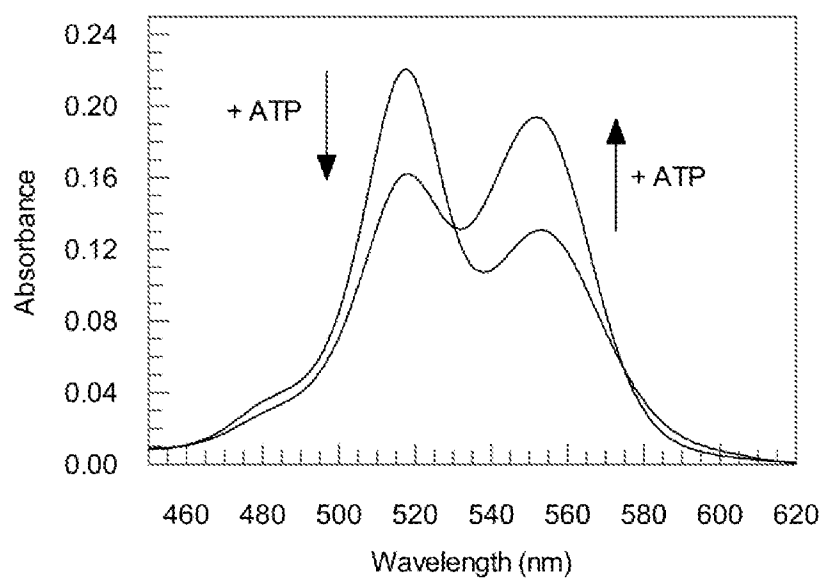
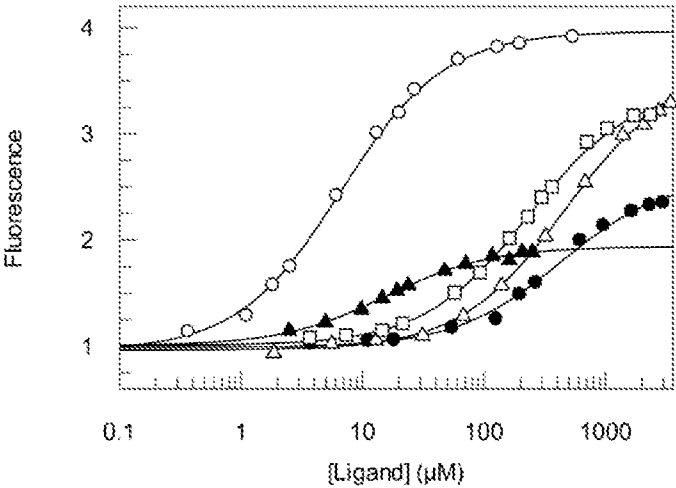
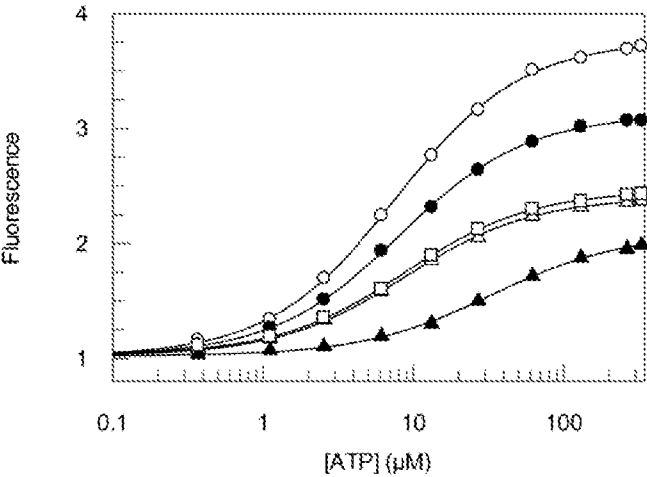


FIGURE 3

(A)



(B)



(C)

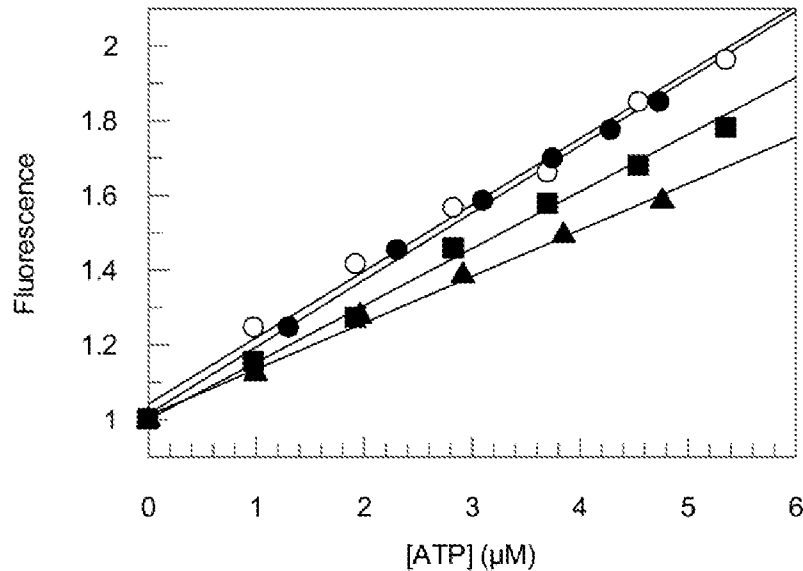


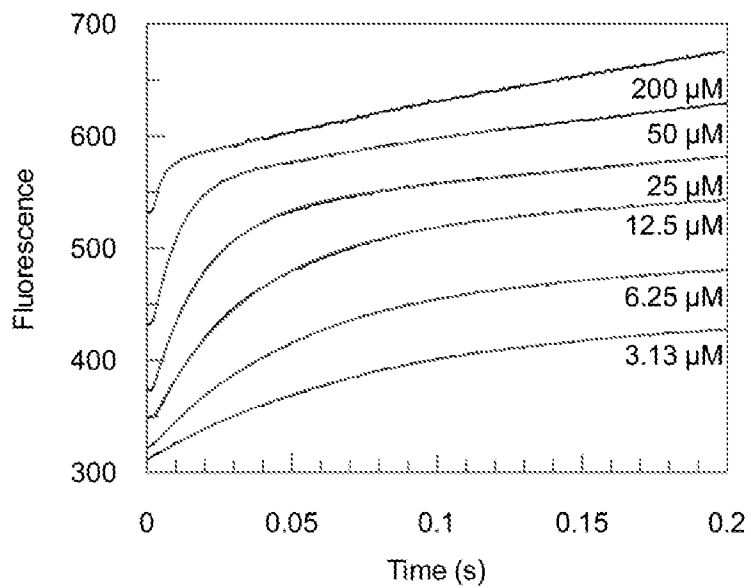
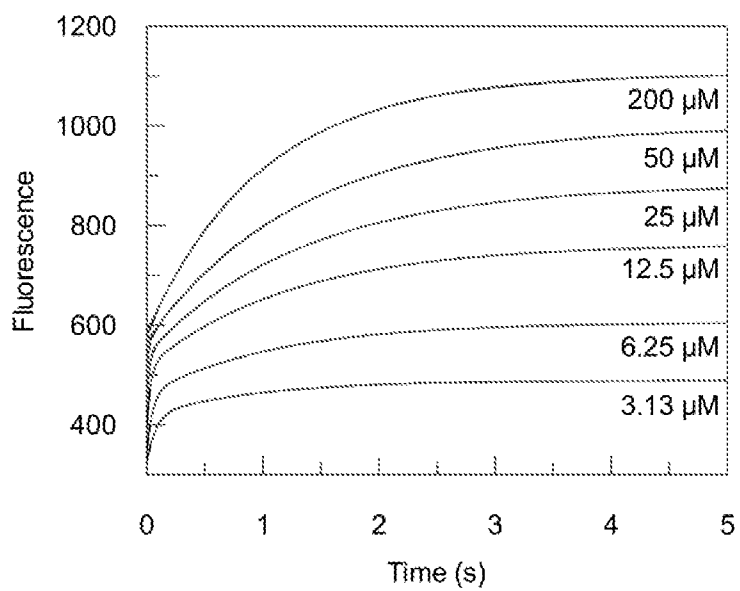
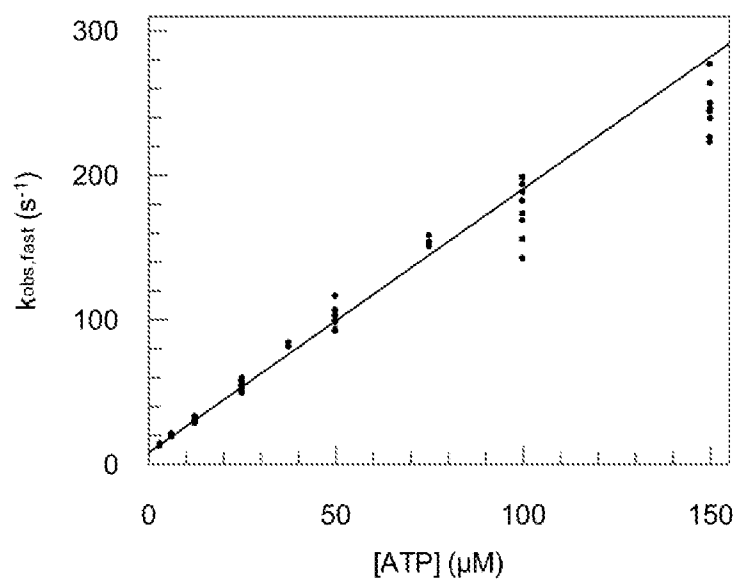
FIGURE 4**(A)****(B)**

FIGURE 4 Continued

(C)



(D)

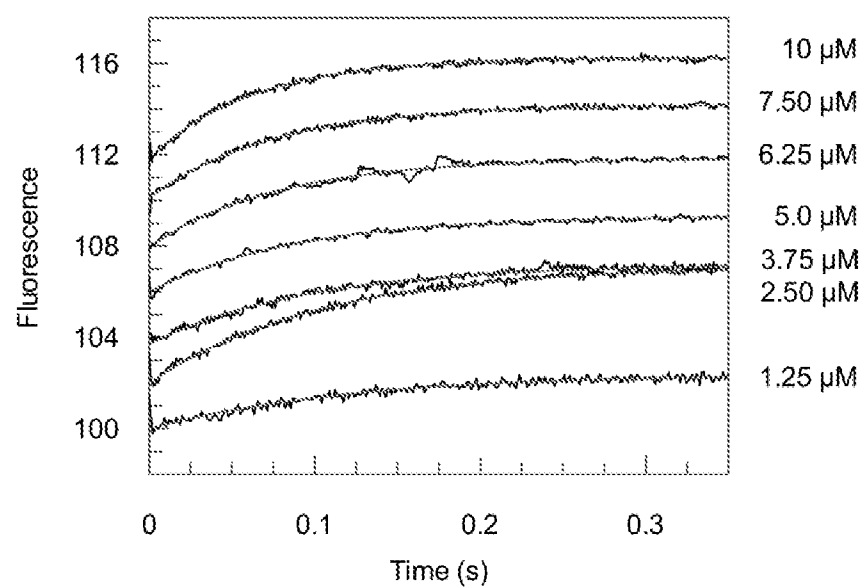
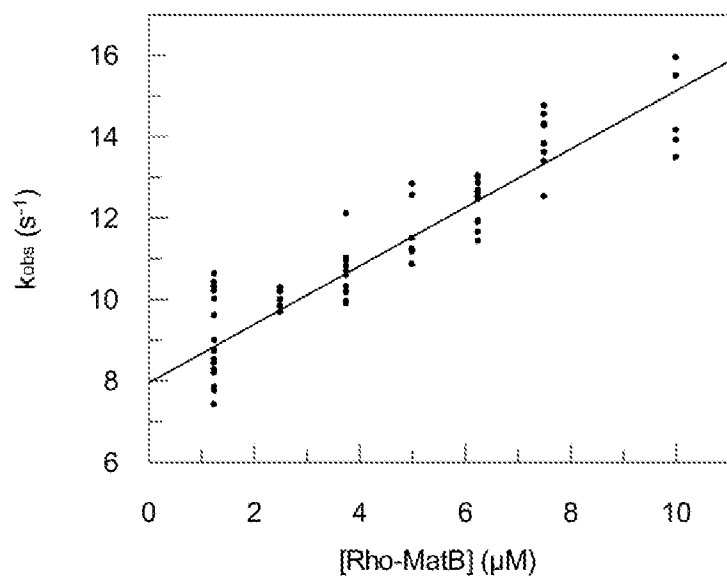


FIGURE 4 Continued

(E)



(F)

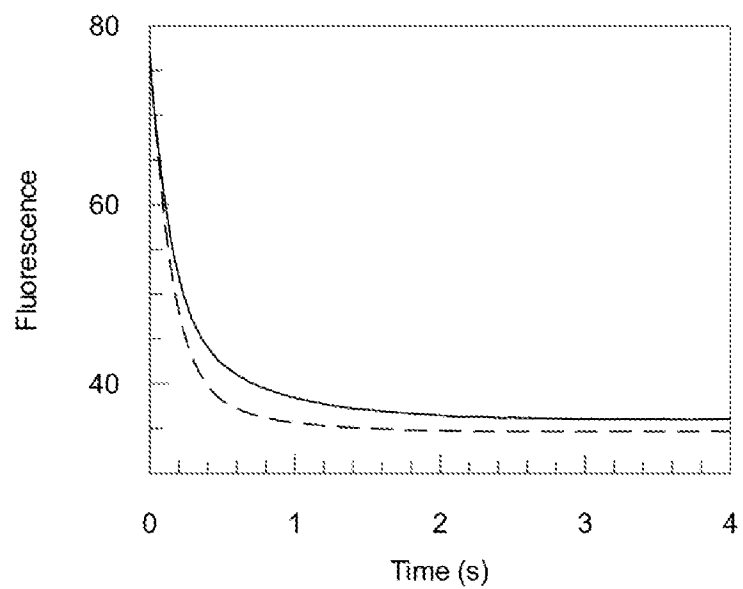
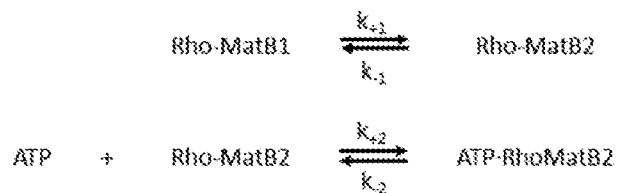


Figure 5

(A)

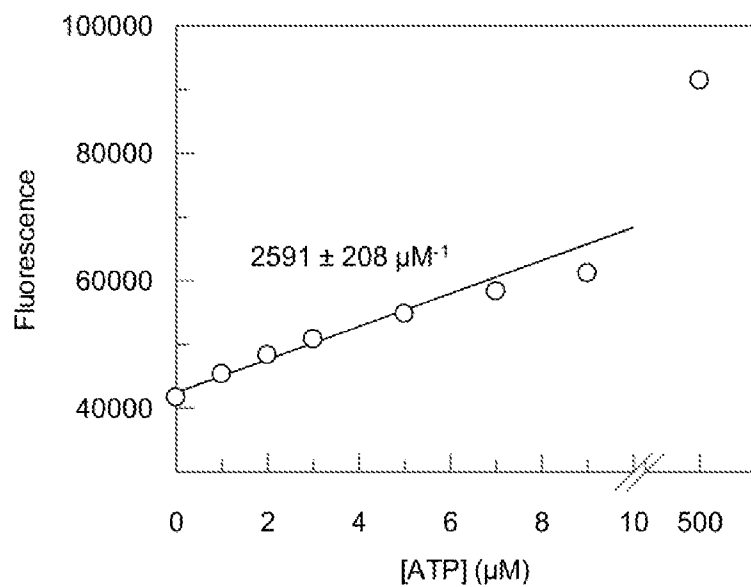


(B)



Figure 6

(A)



(B)

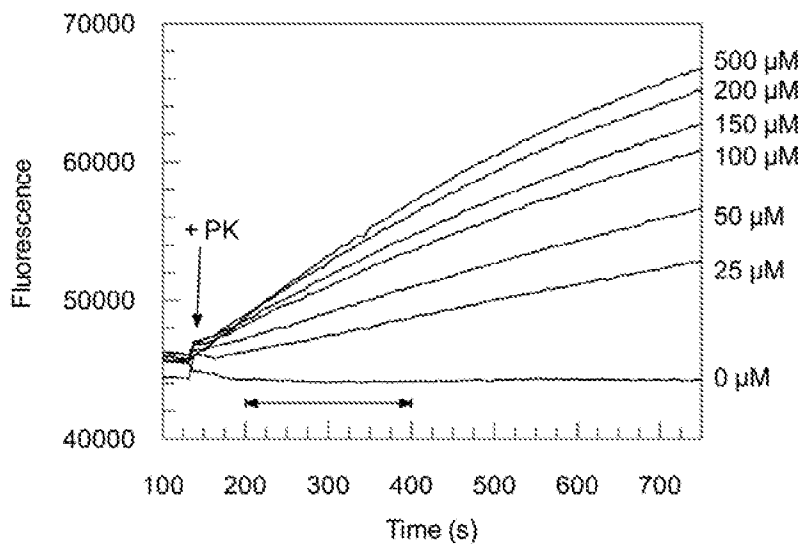
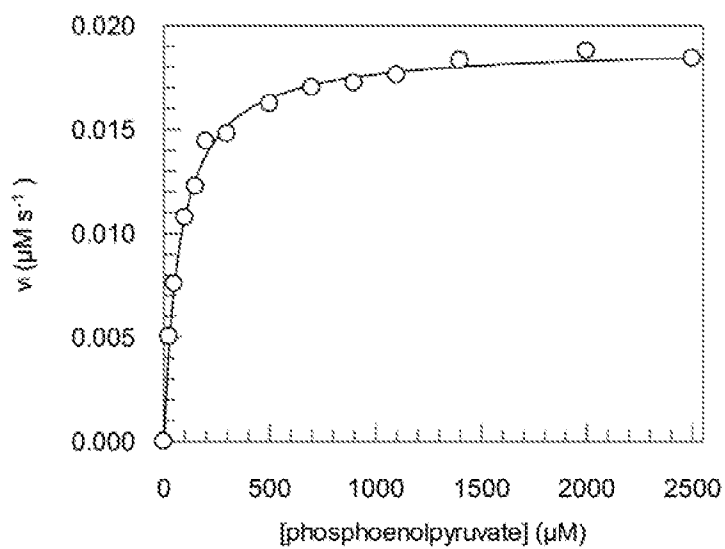
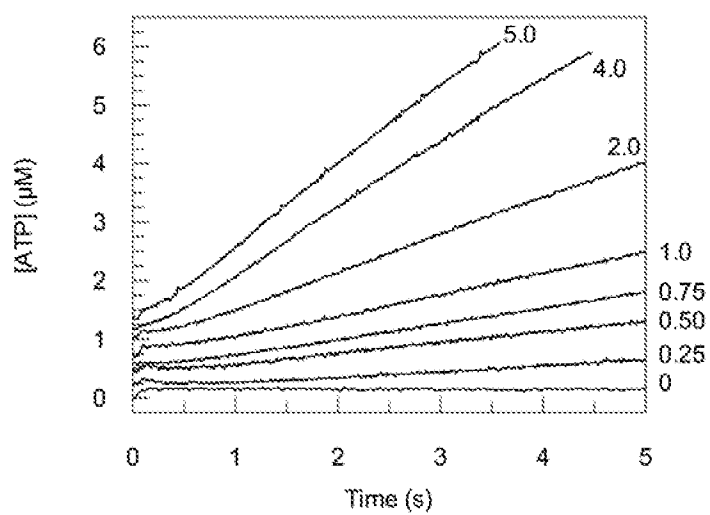


FIGURE 6 Continued

(C)



(D)



(E)

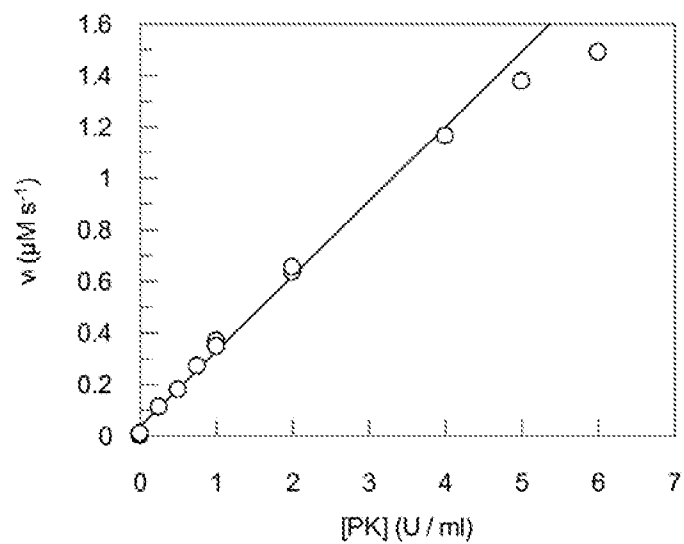
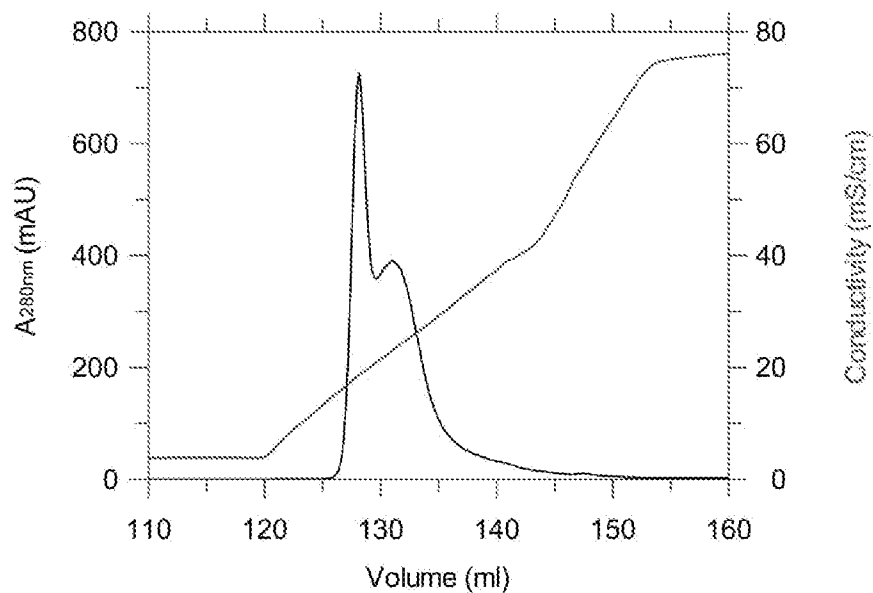
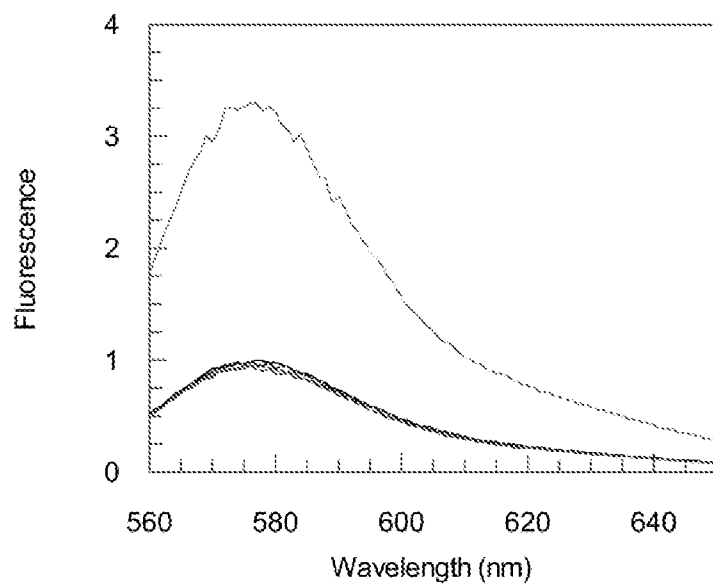


FIGURE 7 (Supplemental figure 1)**FIGURE 8 (Supplemental figure 2)****(A)**

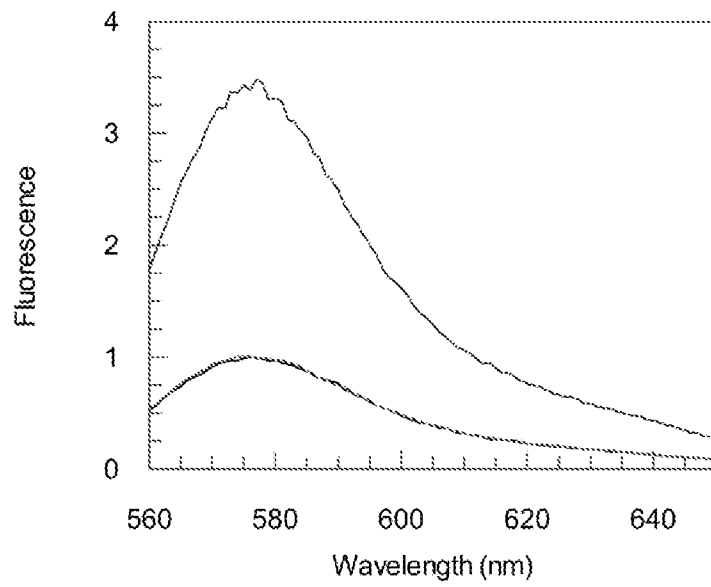
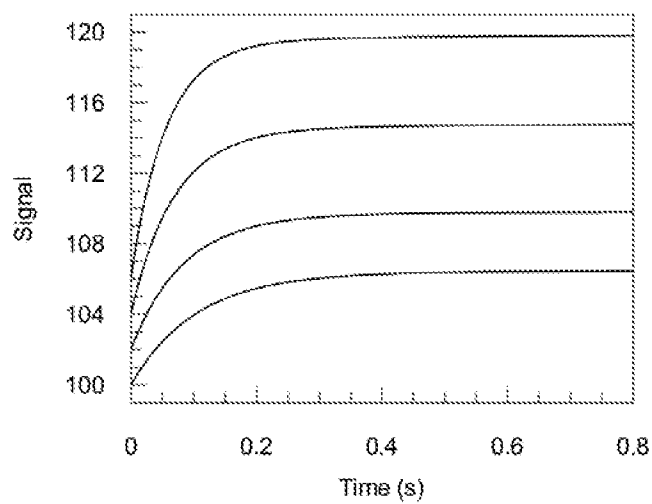
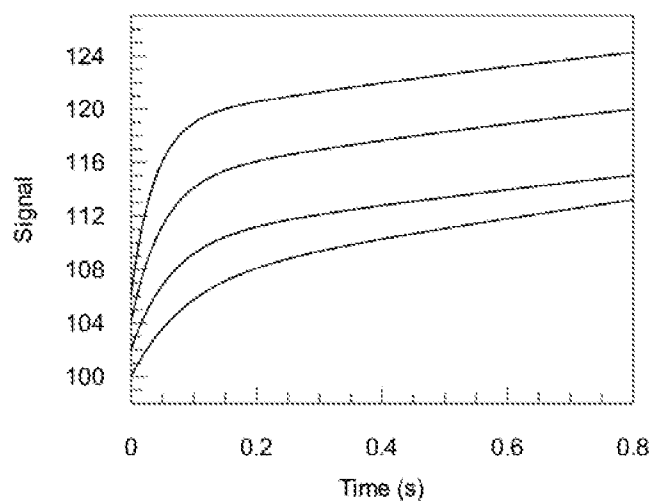
(B)**FIGURE 9 (Supplemental figure 3)****(A)****(B)**

FIGURE 10 (Supplemental figure 4)

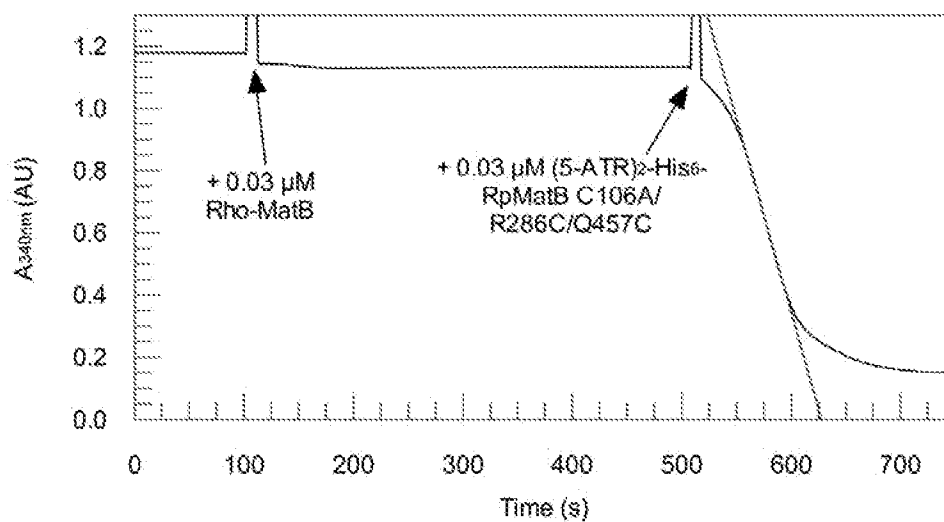
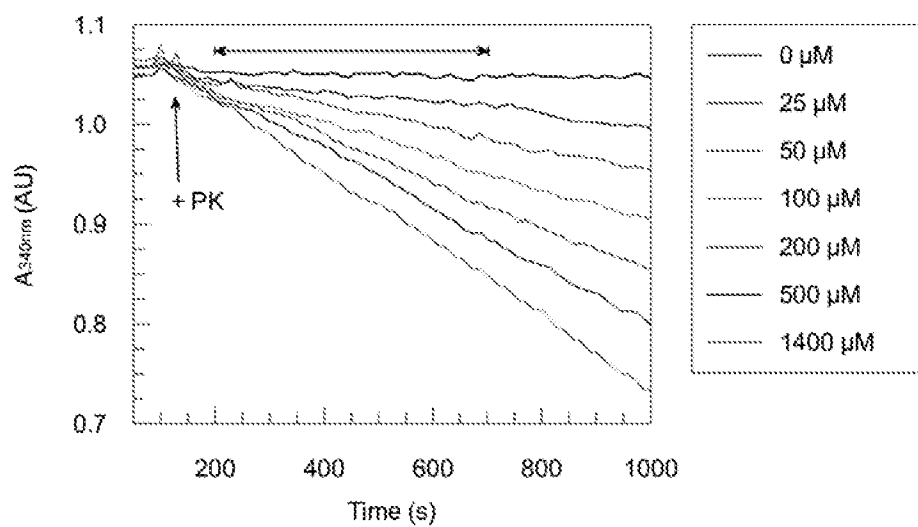


FIGURE 11 (Supplemental figure 5)

(A)



(B)

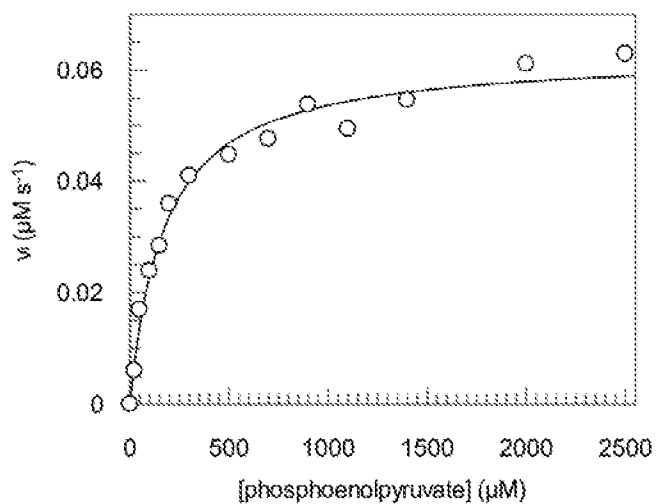


FIGURE 12 (Supplemental figure 6)

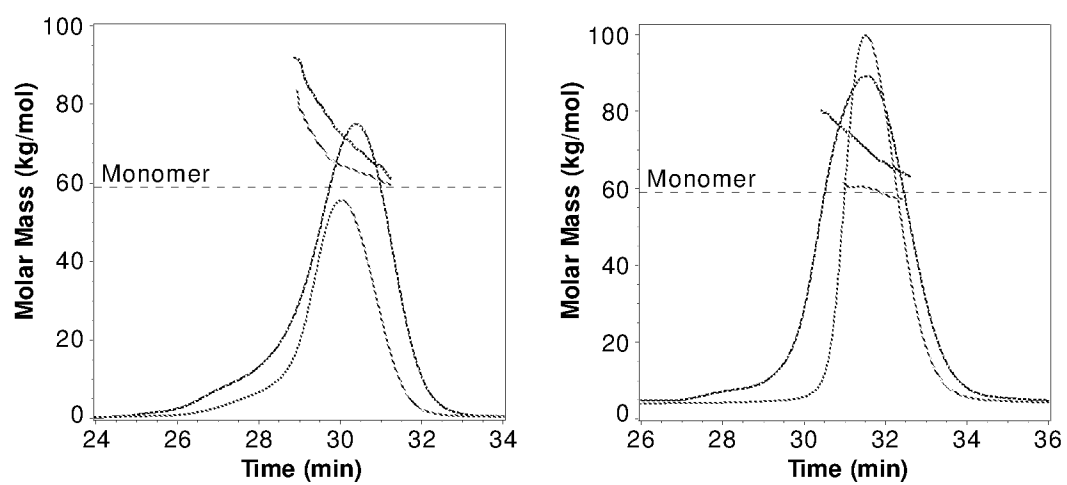


FIGURE 13 (Table 1)

ATP		ADP		dATP		ATP γ S		AMP-PNP		
F ₊ /F ₋	K _d (μM)	In the presence of 100 μM	F ₊ /F ₋	K _d (μM)	F ₊ /F ₋	K _d (μM)	F ₊ /F ₋	K _d (μM)	F ₊ /F ₋	K _d (μM)
3.88	6.37 ± 0.62	/	2.56	428 ± 50	3.87	440 ± 53	1.92	16.2 ± 0.8	3.46	253 ± 11
2.98	7.72 ± 0.57	ADP								
2.31	9.14 ± 1.24	dATP								
1.94	38.3 ± 9.2	ATP γ S								
2.39	8.24 ± 0.35	AMP-PNP								

FIGURE 14 (Supplemental table 1)

Protein	F_{+ATP} / F_{-ATP}	F_{+ADP} / F_{-ADP}	$K_{d,ATP}$ (μ M)	$K_{d,ADP}$ (μ M)
MDCC-His ₆ -RpMatB C106A/K488A/N492C	0.6	0.8	NA	NA
IDCC-His ₆ -RpMatB C106A/K488A/N492C	0.4	0.5	0.6	98
MDCC-His ₆ -RpMatB C106A/K488A/E496C	0.7	NA	NA	NA
MDCC-His ₆ -RpMatB C106A/K488A/V493C	0.9	NA	NA	NA
MDCC-His ₆ -RpMatB C106A/G398C/K488A	0.9	NA	NA	NA
MDCC-His ₆ -RpMatB C106A/K399C/K488A	1.0	NA	NA	NA
MDCC-His ₆ -RpMatB C106A/D400C/K488A	0.7	NA	NA	NA
MDCC-His ₆ -RpMatB C106A/E439C/K488A	1.6	NA	NA	NA
(6-ATR) ₂ -His ₆ -RpMatB C106A/E388C/Q465C/K488A	1.3	NA	NA	NA
(6-ATR) ₂ -His ₆ -RpMatB C106A/E388C/D463C/K488A	1.3	NA	NA	NA
(6-ATR) ₂ -His ₆ -RpMatB C106A/R286C/G461C/K488A	1.6	NA	NA	NA
(6-ATR) ₂ -His ₆ -RpMatB C106A/K385C/E417C/K488A	0.9	NA	NA	NA
(6-ATR) ₂ -His ₆ -RpMatB C106A/K385C/K470C/K488A	0.5	NA	NA	NA
(6-ATR) ₂ -His ₆ -RpMatB C106A/S289C/E417C/K488A	1.1	NA	NA	NA
(6-ATR) ₂ -His ₆ -RpMatB C106A/R286C/Q457C/K488A	2.0	1.9	2.7	283
(5-ATR) ₂ -His ₆ -RpMatB C106A/R286C/Q457C/K488A	2.9	2.8	6.0	566

FIGURE 15 (Supplemental table 2)

Buffer		$K_{d,ATP}$ (μ M)	F_{+ATP} / F_{-ATP}
50 mM Hepes pH 7.0	100 mM NaCl	6.4 ± 0.6	3.9
50 mM Mes pH 6.0	100 mM NaCl	NA	1.5
50 mM Mes pH 6.5	100 mM NaCl	26 ± 2	2.8
50 mM Hepes pH 7.5	100 mM NaCl	5.6 ± 0.6	3.1
50 mM Tris HCl pH 7.5 (*)	100 mM NaCl	6.0 ± 0.4	2.9
50 mM Hepes pH 8.0	100 mM NaCl	4.6 ± 0.2	3.3
50 mM Bicine pH 8.5	100 mM NaCl	4.4 ± 0.4	3.0
50 mM Bicine pH 9.0	100 mM NaCl	NA	2.6
50 mM Hepes pH 7.0	50 mM NaCl	5.8 ± 0.3	3.4
50 mM Hepes pH 7.0	200 mM NaCl	19 ± 1	3.3

(*) Protein used was not further purified via IEXC.

FIGURE 16 (Supplemental Table 3)

PDB ID apo (%)	Conformation apo	Rcn (+)	PDB ID holo (%)	Conformation holo	Rcn (+)	Protein	RMSD (Å) apo - holo
3L8CA (*)	Adenylation	1.4	3LGXA (#)	Adenylation	1.3	<i>Streptococcus pyogenes</i> D-alanine-poly(phosphoribitol) ligase subunit 1	0.81
3L8CB (*)	Alternate	1.0					3.94
4PZP (\$)	Adenylation	1.1	3FCC	Adenylation	1.1	<i>Bacillus cereus</i> D-alanine-poly(phosphoribitol) ligase subunit 1	0.84
3B7W	Alternate	1.4	3C5E	Adenylation	1.5	<i>Homo sapiens</i> Acyl-coenzymeA synthetase ACSM2A	9.54
4FUQA (#)	Alternate	1.3	4FUT	Adenylation	1.5	RpMatB	6.18
1ULTA (*)	Alternate	1.7	1V25A (#)	Thioester	3.4	<i>Thermus thermophilus</i> long-chain-fatty-acid-coenzymeA synthetase	8.90
1ULTB (*)	Alternate	3.1					6.38

(%) In total, we analyzed 11 holo molecules and 25 apo molecules. We found that the C-terminal domain was not present in 11 out of 25 analyzed apo crystal structures. It was either unresolved hence assumed to be flexible (like PDB 3IVR and PDB 4LGC) or not included in the purified protein (like PDB 3T5B and PDB 3WV4). For simplicity, only the 5 apo-holo pairs are shown in this table.

(+) Ratio of the average B-factor of the C-terminal domain using all atoms and the average B-factor of the N-terminal domain using all atoms.

(*) This crystal structure contained multiple chains that had a different conformation. Chains representing both conformations were used in the analysis.

(\$)

(#) This crystal structure contained multiple chains but all had the same conformation. Only one was used in the analysis.

References to publications can be found using the links at the Protein Data Bank.

FIGURE 17

SEQ ID NO: 7

```

1      11      21      1      11      21      31
MSYYHHHHHH DYDIPTSENL YFQGASMNAN LFARLFDKLD DPHKLAIETA AGDKISYAEI

41      51      61      71      81      91
VARAGR VANV LVARG LQVGD RVAAQ TEKSV EALVLYLATV RAGGVYLPLN TAYTLHELDY

101     111     121     131     141     151
FITDAEPKIV VCDPSKRDGI AAIAAKVGAT VETLGPDGRG SLTDAAAGAS EAFATIDRGA
C106A

161     171     181     191     201     211
DDLAAI LYTS GTTGR SKGAM LSHDN LASNS LTLVDYWRFT PDDVLIHALP IYHTHGLFVA

221     231     241     251     261     271
SNVTLFARG S MIPLPKFDPD KILDL MARAT VLMGVPTFYT RLLQSPRLTK ETTGHMRLFI

281     291     301     311     321     331
SGSAPLLADT NREWSAKTGH AVLERYGMTE TNMNTSNPYD GDRVPGAVGP ALPGVSARVT
R286C

341     351     361     371     381     391
DPETGKELPR GDIGMIEVKG PNVFKGYWRM PEKTKSEFRD DGFFITGDLS KIDERGYVHI
K385C

401     411     421     431     441     451
LGRGKDLVIT GGFNVYPKEI ESEIDAMPGV VESAVIGVPH ADFGEGVTAV VVRDKGATID
E432C

461     471     481     491     501
EAQVLNGLDG QIAKFKMPKK VIFVDDLPRN TMGKVGKNVL RFTYKDIYK
Q457C G461C K470C K488A N492C

```

FIGURE 18

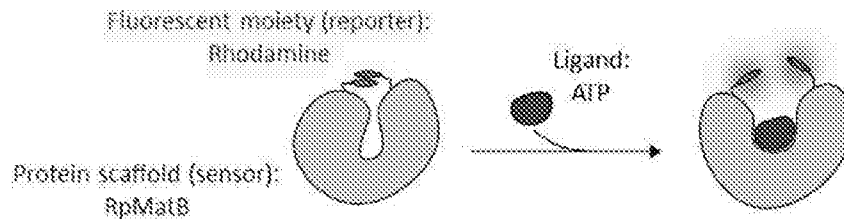


FIGURE 19

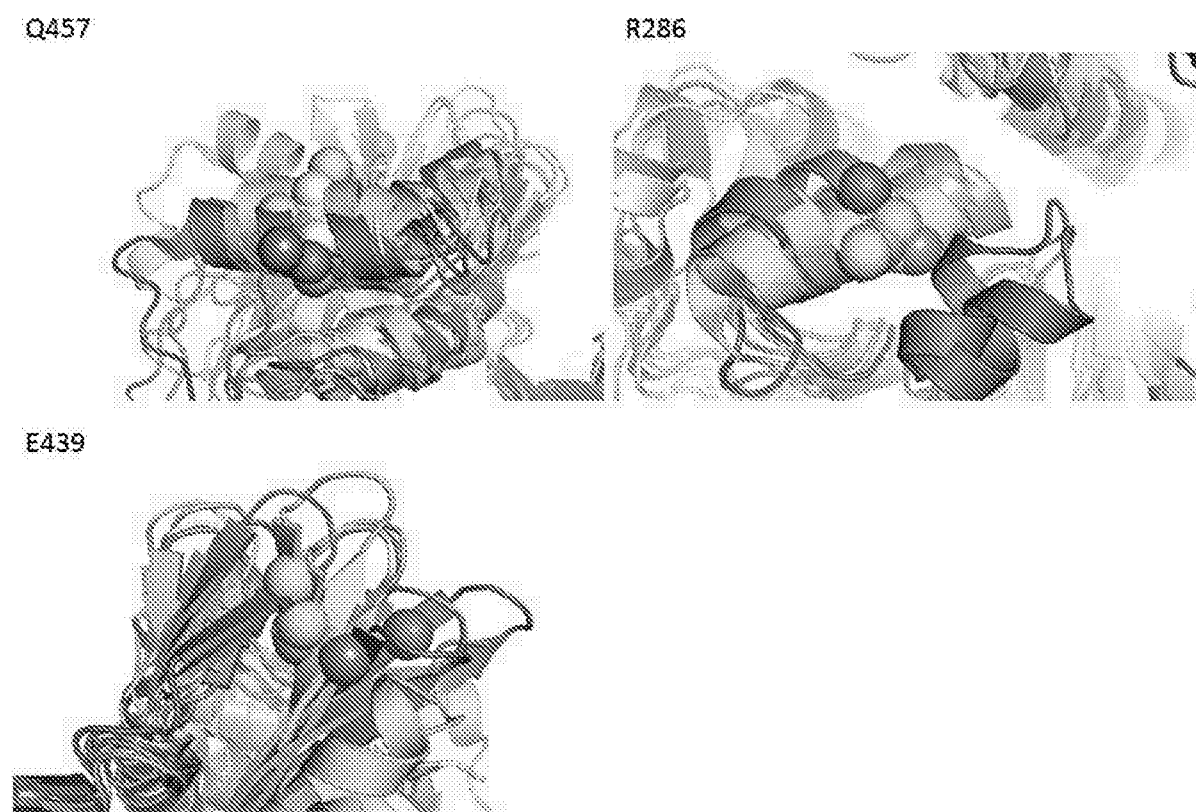


FIGURE 20

(A)

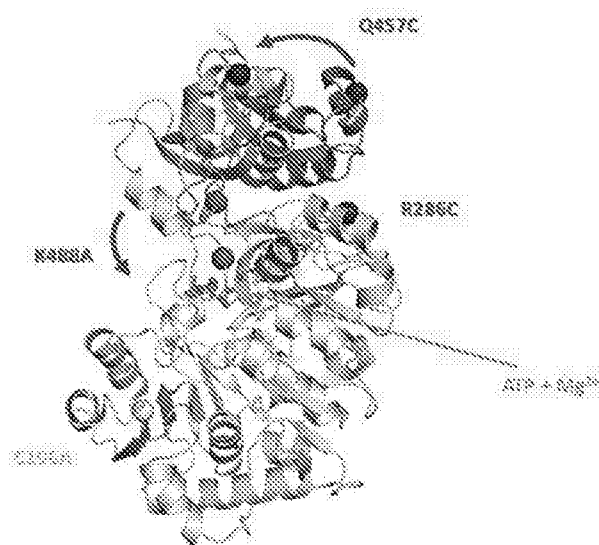
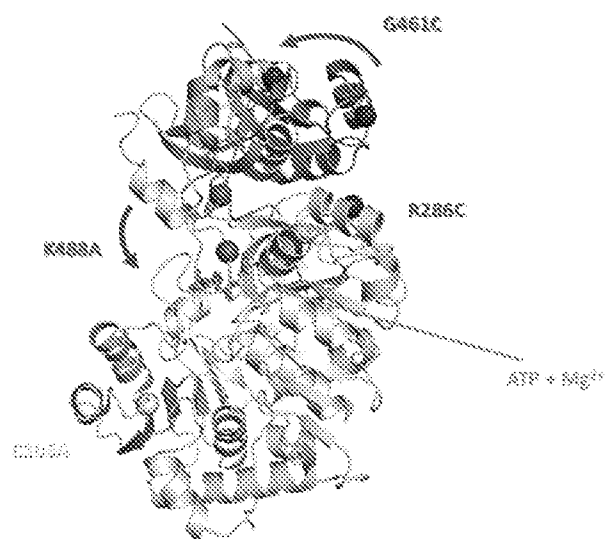
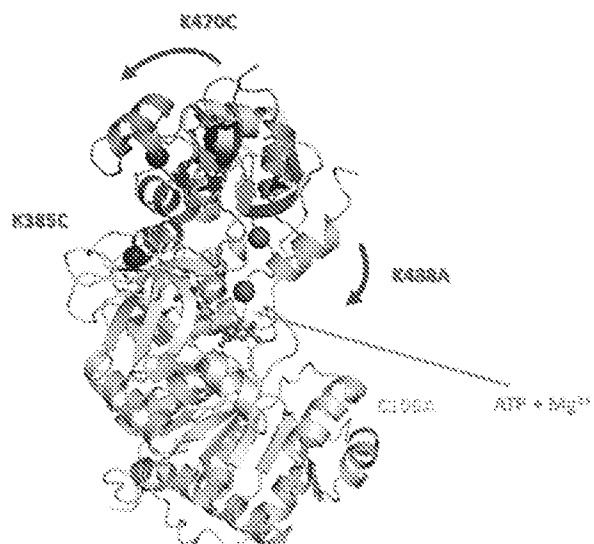


FIGURE 20 continued

(B)



(C)



(D)

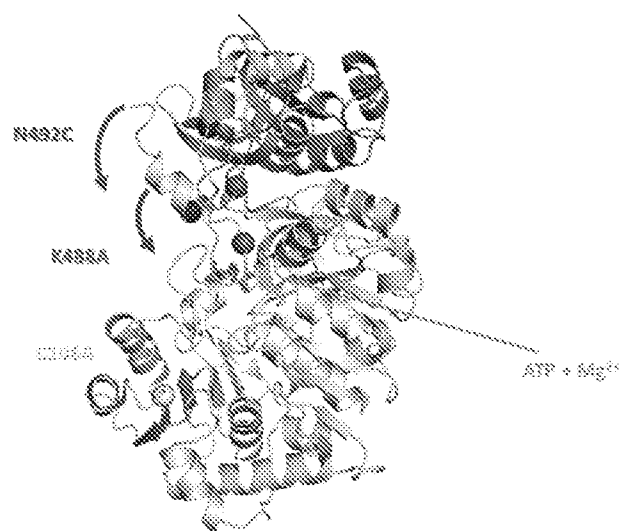


FIGURE 20 continued

(E)

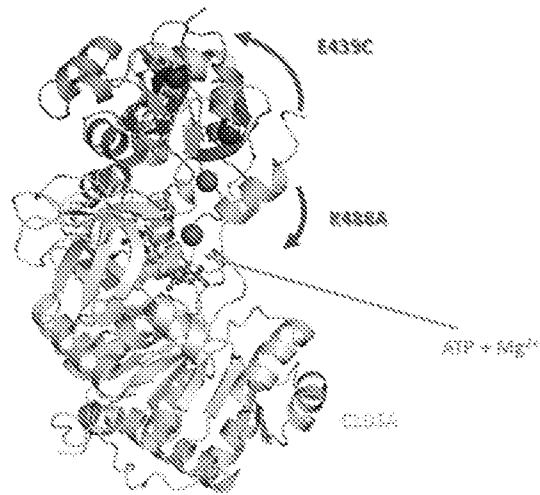


Figure 21

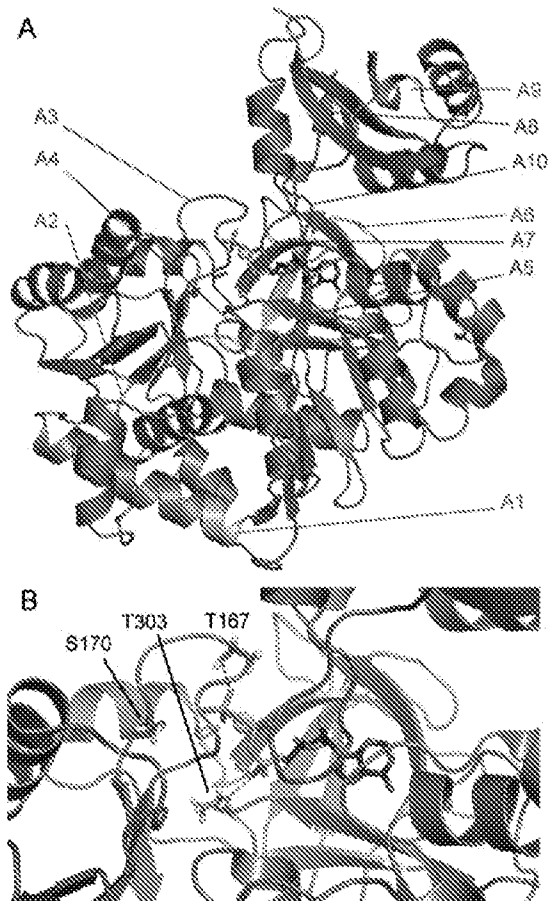


Figure 22

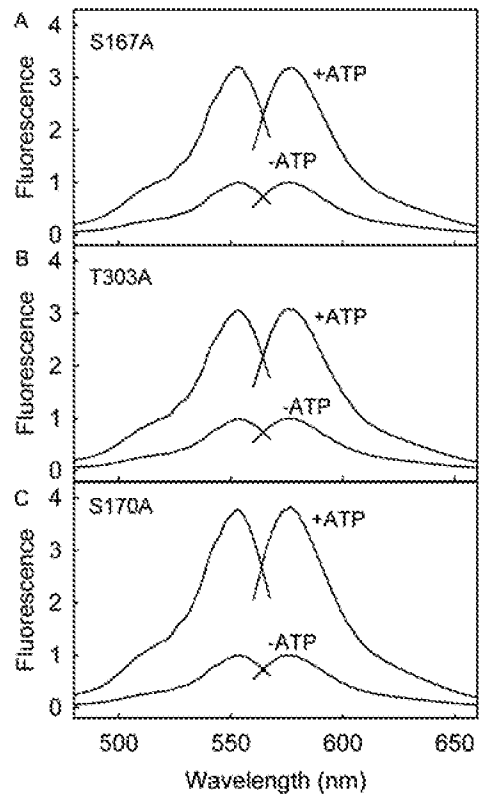


Figure 23

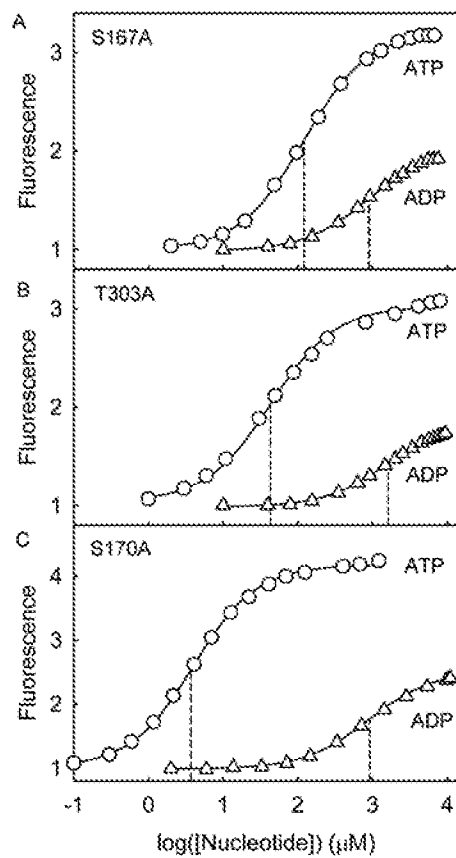


Figure 24

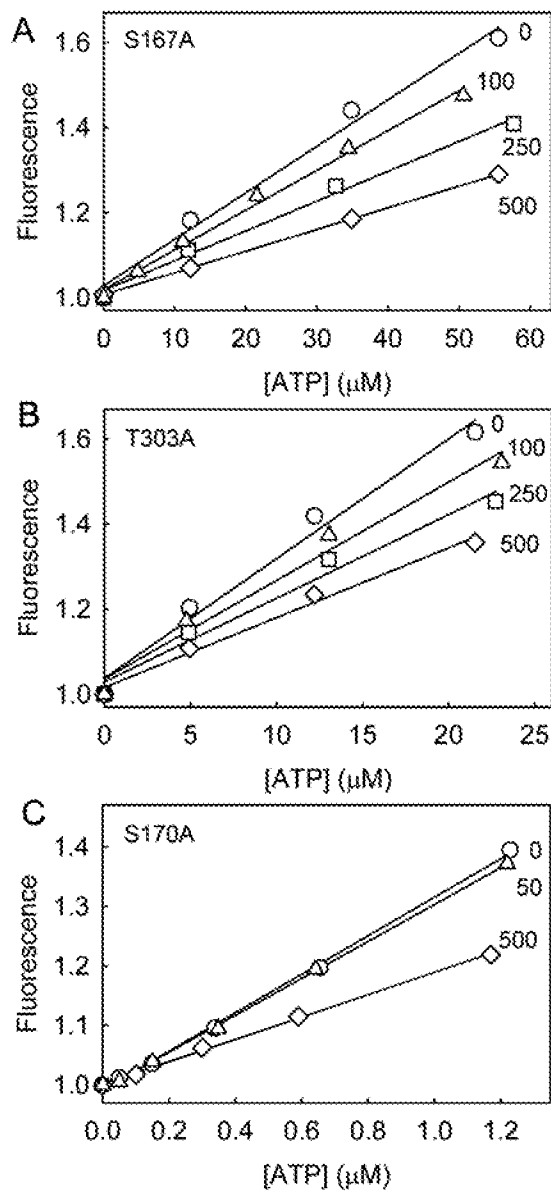


Figure 25

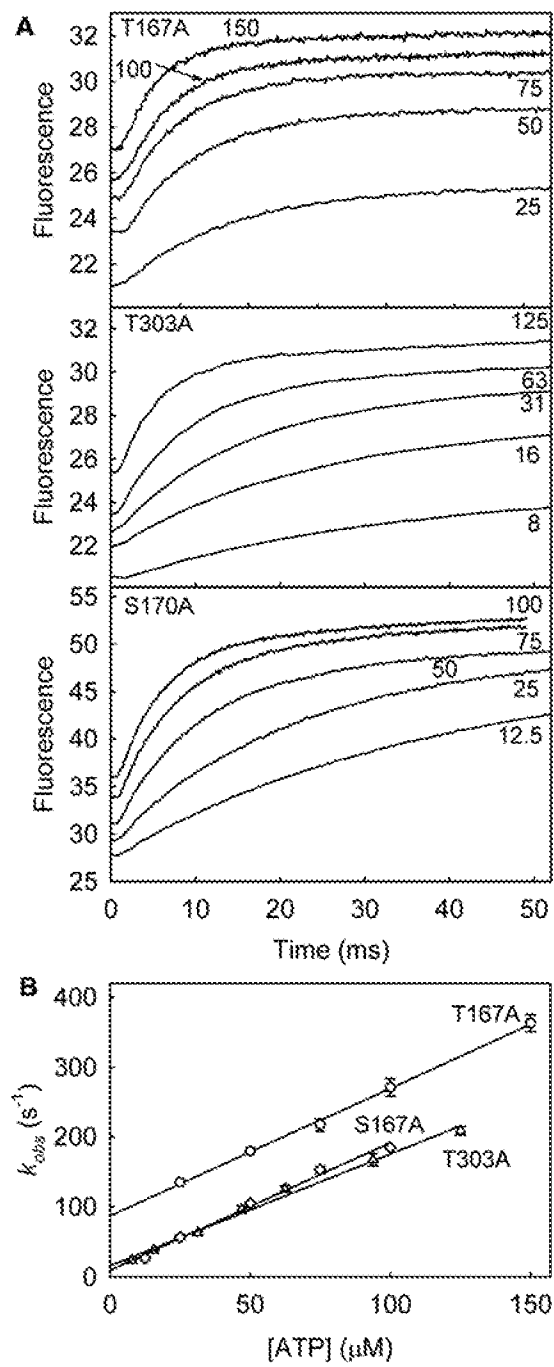


Figure 26

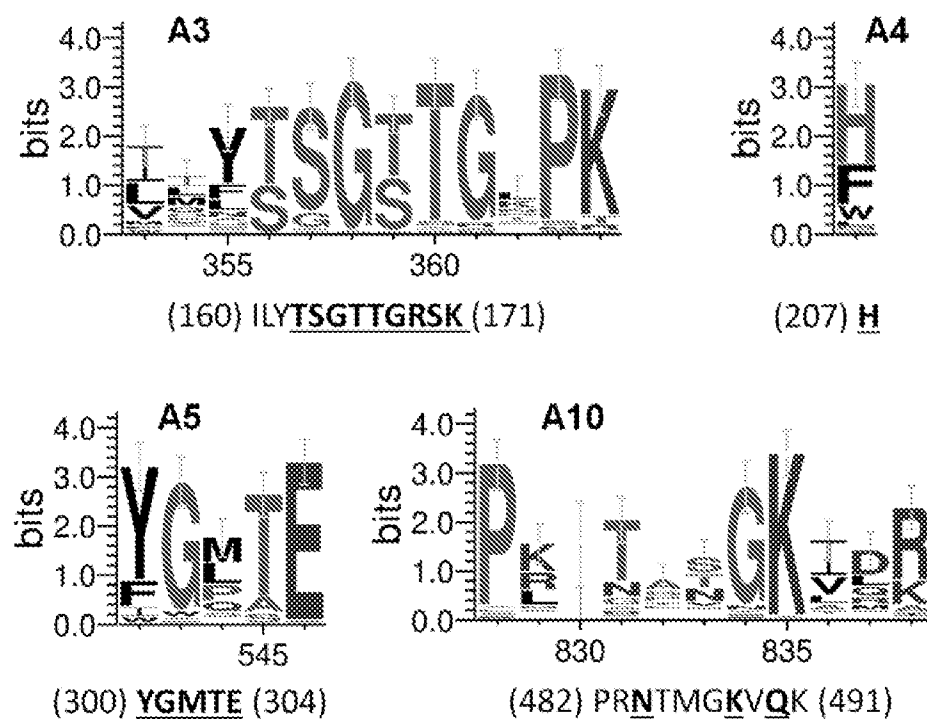


Figure 27

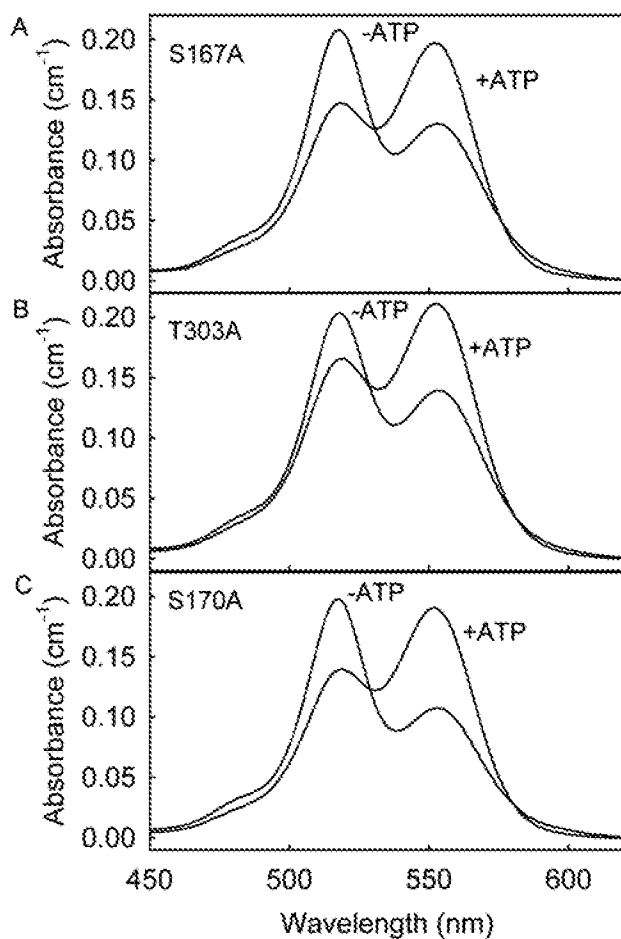
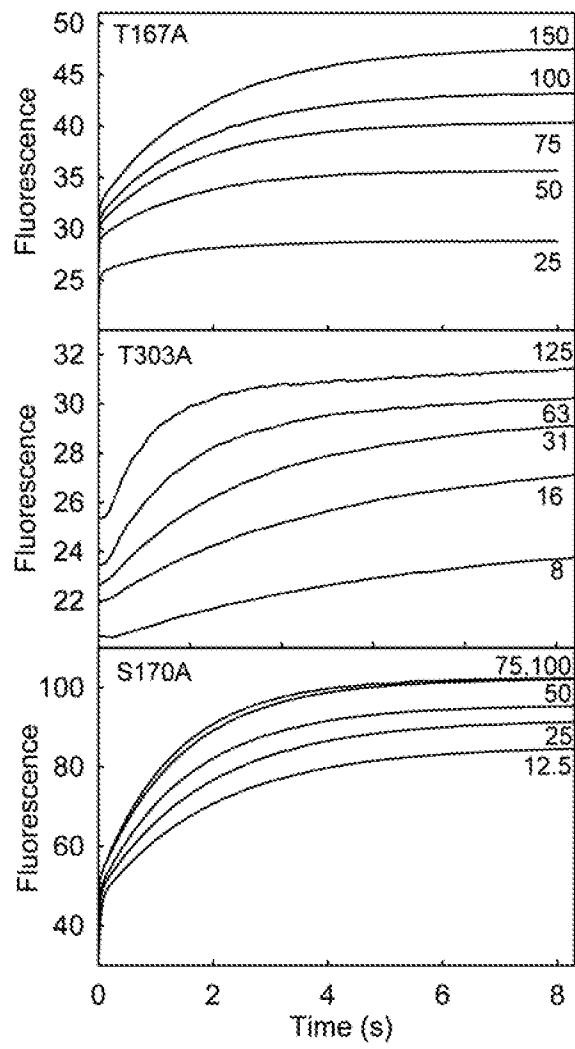


Figure 28



INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2016/050818

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12N9/00 C12Q1/00 G01N21/76
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)

C12N C12Q G01N

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, COMPENDEX, CHEM ABS Data, EMBASE, FSTA

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	IMAMURA HIROMI ET AL: "Visualization of ATP levels inside single living cells with fluorescence resonance energy transfer-based genetically encoded indicators", PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, NATIONAL ACADEMY OF SCIENCES, US, vol. 106, no. 37, 15 September 2009 (2009-09-15), pages 15651-15656, XP002575707, ISSN: 0027-8424, DOI: 10.1073/PNAS.0904764106 [retrieved on 2009-08-31] abstract; figure 2 ----- -/-	32
A		1-31



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

28 June 2016

Date of mailing of the international search report

05/08/2016

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Strobel, Andreas

INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2016/050818

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 00/68418 A1 (MEDICAL RES COUNCIL [GB]; BRUNE MARTIN HERMANN KLEMENS [GB]; CORRIE JO) 16 November 2000 (2000-11-16)	32
A	page 12, line 7 - line 20; claims 2-19 -----	1-31
X	WO 2010/032001 A1 (MEDICAL RES COUNCIL [GB]; WEBB MARTIN RONALD [GB]; KUNZELMANN SIMONE []) 25 March 2010 (2010-03-25)	32
A	page 7, line 16 - line 24 page 57, line 27 - page 62, line 11; claims 1,7-10 -----	1-31
X	KORYAKINA IRINA ET AL: "Mutant Malonyl-CoA Synthetases with Altered Specificity for Polyketide Synthase Extender Unit Generation", CHEMBIOCHEM, vol. 12, no. 15, October 2011 (2011-10), pages 2289-2293, XP002759244,	32
A	Scheme 1; page 2289, right-hand column, line 3 - last line; table 1 -----	1-31
X	CROSBY H A ET AL: "Structure-guided expansion of the substrate range of methylmalonyl coenzyme a synthetase (MatB) of rhodopseudomonas palustris", APPLIED AND ENVIRONMENTAL MICROBIOLOGY SEPTEMBER 2012 AMERICAN SOCIETY FOR MICROBIOLOGY USA, vol. 78, no. 18, September 2012 (2012-09), pages 6619-6629, XP002759245, DOI: 10.1128/AEM.01733-12	32
A	abstract; figures 1,6; table 6 page 6621, right-hand column, paragraph 4 -----	1-31
X,P	VANCRAENENBROECK RENEE ET AL: "A Fluorescent, Reagent less Biosensor for ATP, Based on Malonyl-Coenzyme A Synthetase", ACS CHEMICAL BIOLOGY, vol. 10, no. 11, November 2015 (2015-11), pages 2650-2657, XP002759246, the whole document -----	1-32

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/GB2016/050818

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0068418	A1	16-11-2000	AT 264400 T 15-04-2004
		AU 771858 B2 01-04-2004	
		AU 4592300 A 21-11-2000	
		CA 2371176 A1 16-11-2000	
		DE 60009891 D1 19-05-2004	
		DE 60009891 T2 21-04-2005	
		EP 1177314 A1 06-02-2002	
		JP 4338320 B2 07-10-2009	
		JP 2002543798 A 24-12-2002	
		US 6746849 B1 08-06-2004	
		US 2004248079 A1 09-12-2004	
		WO 0068418 A1 16-11-2000	
WO 2010032001	A1	25-03-2010	CA 2736666 A1 25-03-2010
		DK 2342216 T3 03-12-2012	
		EP 2342216 A1 13-07-2011	
		ES 2394156 T3 22-01-2013	
		US 2011294137 A1 01-12-2011	
		WO 2010032001 A1 25-03-2010	