Published in final edited form as:

Chembiochem. 2016 October 04; 17(19): 1873–1878. doi:10.1002/cbic.201600264.

Functional Site Discovery in a Sulphur Metabolism Enzyme Using Directed Evolution

Dr. Hanumantharao Paritala^[a], Dr. Prakash B. Palde^[a], and Prof. Dr. Kate S. Carroll^[a]

Kate S. Carroll: kcarroll@scripps.edu

^[a]Department of Chemistry, The Scripps Research Institute, 130 Scripps Way, 2B2, Jupiter, Florida, 33458, USA

Abstract

In human pathogens, the sulphate assimilation pathway provides reduced sulphur for biosynthesis of essential metabolites including cysteine and low-molecular weight thiol compounds. Sulfonucleotide reductases (SRs) catalyze the first committed step of sulphate reduction. In this reaction, activated sulphate in the form of adenosine-5'-phosphosulphate (APS) or 3'-phosphoadenosine 5'-phosphosulphate (PAPS) is reduced to sulphite. Gene knockouts, transcriptomic and proteomic data establish the importance of SRs in oxidative stress-inducible antimicrobial resistance mechanisms. In previous work, we have focused on rational and high-throughput design of small-molecule inhibitors that target the active site of SRs. However, another critical goal is to discover functionally important regions in SRs beyond the traditional active site. As an alternative to conservation analysis, we have used directed evolution to rapidly identify functional sites in PAPS reductase (PAPR). Four new regions are discovered that are essential to PAPR function and lie outside the substrate-binding pocket. Our results highlight the use of directed evolution as a tool to rapidly discover functionally important sites in proteins.

Graphical Abstract

We have used directed evolution to rapidly identify new functional sites in the sulphur metabolism enzyme, PAPS reductase (PAPR). These regions may have a regulatory or allosteric function and could lay the foundation for development of molecules that target regions of PAPR, which are distinct from the active site and smaller than most natural protein-protein interfaces.

Correspondence to: Kate S. Carroll, kcarroll@scripps.edu.

Dedicated to the memory of Professor David B. Knaff.

Supporting information for this article is given via a link at the end of the document.

Keywords

sulphate reduction; enzymes; directed evolution; antibiotic target

Introduction

The growing issue of antibiotic-resistant infections is viewed as an increasingly urgent public health threat.^[1] The situation is particularly acute for Gram-negative bacteria, many of which are resistant to all available antibacterial agents.^[2] For example, more than two million people are infected with antibiotic-resistant bacteria each year in the United States and more than 23,000 individuals die annually as a direct result of these infections.^[3] Frequent antibiotic use over long periods of time creates a selective pressure for resistant organisms.^[4] Gram-negative strains of *Escherichia coli*, *Klebsiella pneumoniae*, *Pseudomonas aeruginosa*, and *Acinetobacter baumannii* have been identified that are resistant to some or all of the antibiotic classes most commonly used to treat these bacteria: penicillins, quinolones, cephalosporins, carbapenems, monobactams, aminoglycosides tetracyclines, and polymyxins.^[5] Thus, there is no question that new antibiotics are needed to combat the growing problem of resistance and that targeting of new pathways will likely play an important role in their discovery.

A promising new metabolic pathway for antimicrobial drugs is the sulphate assimilation pathway,^[6] which is broadly present in bacteria, yeast, and fungi, essential for pathogen viability, and is not found in humans. In this pathway, inorganic sulphate is adenylated by the enzyme ATP sulphurylase, forming APS, which is subsequently reduced to sulphite and AMP. Sulphite is further reduced to sulphide, which is used for biosynthesis of essential biomolecules including cysteine, methionine, cofactors and low molecular weight thiol buffers. In plants, algae, and many bacteria, APS is directly reduced to sulphite by APS reductase (APR). On the other hand, in some species of Gram-negative bacteria, cyanobacteria, γ -proteobacteria, fungi, and yeast, APS must be phosphorylated on its 3'-hydroxyl to yield PAPS, which is reduced by PAPR to produce sulphite (Scheme 1). Gene knockouts, gene expression and proteomic data sets establish the importance of SRs in antimicrobial resistance mechanisms.^[7] Additionally, SRs have no human homologue, rendering them an attractive target for antibacterial therapy.

It was once believed that the APS pathway was dependent on oxygen and evolving photosynthesis, and that the pathway involving PAPS was confined to non-photosynthetic

bacteria. However, later studies in prokaryotes demonstrated that PAPS reduction is restricted to a few groups of γ -proteobacteria and cyanobacteria, while APR was also found among cyanobacteria and γ -proteobacteria.^[8] The evolution of PAPR might have been an adaptation to an iron-sulphur poor environment or to increasing concentrations of oxygen in the atmosphere since the 4Fe-4S cluster (found in APR, but absent in the great majority of PAPR) is unstable in the presence of air. However, an evolutionary advantage of one sulphate-reducing pathway over the other has not been conclusively demonstrated.

Our early studies of SRs focused on acquisition of detailed mechanistic and structural information,^[9] laying the foundation for the discovery of small-molecules that can alter the function of these enzymes. To this end, we have used rational design and high-throughput screening^[10] to discover molecules that bind to the active site of SRs. However, another goal of fundamental importance is to identify functionally important regions of SRs beyond the substrate-binding pocket that might also serve as targets for small-molecule modulators. Conservation analysis is one of the most widely used methods for predicting functionally important sites,^[11] but time consuming studies involving scanning mutagenesis must be performed to narrow down the search space and putative hot spots. Here, we use directed evolution to rapidly discover four new regions of PAPR that are essential for function *in vivo*, but are located outside the “active site box”.

Results and Discussion

To better understand structure-function relationships that govern PAPR activity, we initiated directed evolution studies. Directed evolution is a powerful synthetic biology tool to enhance or alter the activity of enzymes.^[12] In the present study, however, we employ this iterative approach to identify regions of PAPR that are essential for its cellular function. The *cysH* gene of *E. coli* encodes PAPR, a protein of 244 amino acid residues that uses PAPS as a substrate. To introduce random mutations we used the GeneMorph II system, which uses a unique error prone PCR enzyme blend to achieve equivalent rates of mutation at A's and T's vs. G's and C's.^[13] In the initial stage of this study, we varied the amount of input DNA template to generally allow mutagenesis at only a single site per round of selection. The PAPR mutant library was generated in the pBR322 plasmid encoding the *rop* gene, restricting plasmid copy number to ~15–20 per cell.^[14] Functional mutants were then selected on the basis of their ability to complement *E. coli* strain JM96, a mutant strain which is defective in PAPR activity (i.e., *cysH* mutant)^[15] grown on minimal media with sulphate as the sole source of sulphur. Amino acid positions that did not tolerate or allowed only conservative changes (i.e., an amino acid residue is replaced by a residue having similar biochemical properties) were considered “preserved” and inferred to play an important role in the function, structure, and/or cellular stability of PAPR.

Approximately 500 functional PAPR mutants were mapped through DNA sequencing. Amino acid changes in the functional PAPR mutants were distributed throughout the protein sequence (Figure S1 and Table S1). Almost 80% of selected plasmid library contained single mutations (~400 DNA sequences). With multiple mutations compensation can occur and, consequently, these variants were removed from further consideration. Of 244 residues in *E. coli* PAPR, a total of 130 amino acids or ~50% were found to tolerate single mutation. Of

these, 43 residues or ~30% underwent two or more unique substitutions. Overall, the oversampling factor (O_f) in this study was three, which provides a 95% probability of accurate prediction according to the widely accepted correlation proposed by Reetz *et al.*^[12f]

Location of functional mutations was achieved by mapping onto the available crystal structures of *E. coli* and *S. cerevisiae* PAPR.^[9d, 16] Residues that did not undergo mutation clustered around six discrete sites (Figure 1). First among these regions is the PAPS substrate-binding site. Active site conservation was expected, since mutation of such residues would directly impact the ability of PAPS to bind effectively to PAPR. Residues in an arginine-rich loop (Arg-loop; Arg157, Arg158 and Arg164) between strands $\beta 4$ and $\beta 5$ were also preserved in our screen. Upon PAPS binding, the Arg-loop folds over the active site bringing Arg164 into direct contact with oxygen that bridges the substrate phosphorous and sulphur atoms. Another critical region comprises a stretch of amino acids within a flexible C-terminal tail (residues 225–244). This peptide segment contains the essential Cys239 nucleophile, which folds over the PAPR active site upon substrate binding. With the exception of Cys239, the strongly conserved ECGLH motif in the C-terminal segment was somewhat tolerant to mutation, underscoring the view that not all highly conserved residues are essential for function.

A covalent *S*-sulfofocysteine intermediate forms during the first chemical step of the PAPR reaction, which is cleaved in the second chemical step by the small 12-kDa thiol oxidoreductase, thioredoxin (Trx), thereby generating sulphite product.^[9b] In line with the importance of this protein cofactor, our selection effectively mapped residues that make up an 800 Å² interface with Trx, as visualized in the structure of these two proteins covalently linked *via* a mixed disulfide that forms during nucleophilic attack of Trx's N-terminal cysteine on the S_γ atom of the PAPR *S*-sulfofocysteine intermediate.^[9d] The protein-protein interface in the crystal structure, and also mapped by this study, is comprised of PAPR residues 188–199 (α -helix 7) and 202–212 (ω -loop) (Figure S2). In a recent thermodynamic study many of these residues, such as Tyr191, Trp205, and Asp206, were shown to make a favorable enthalpic contribution to the Trx•PAPR interaction.^[17] Our ability to map this interface in our selection supports crystallographic and thermodynamic studies, and underscores the potential of small-molecules that target the Trx•PAPR interaction. In addition, conservation of PAPS-binding active site residues, the Arg-loop, Cys239, mapping of the PAPR-Trx interface all serve to validate our screen for enzyme function *via* directed evolution.

In addition to the highly anticipated sites described above, four additional residue clusters were identified in our functional screen (Figure 2). Site 1 is comprised of four residues: Leu28, Lys30, Leu31 and Asp32 (Figure 2A). These amino acids make up a short, flexible loop connecting α -helix 2 to α -helix 3. This segment is approximately 20 Å from the PAPS-binding site and 15 Å from the Trx-binding site. From an electrostatic perspective, Lys30 and Asp32 are strategically positioned to stabilize helix dipoles, while Leu28 and Leu31 side-chains form a hydrophobic patch between the helices. Interestingly, Leu28 and Leu31 are highly conserved among PAPRs, whereas Lys30 and Asp32 can be replaced by other polar residues (Figure S1). Residues in Site 1 establish hydrogen bonds with Arg36, Pro181, and Asp184, located on α -helix 3 and at the terminus of β -sheet 6. Positively charged Arg36

interacts with backbone carbonyls of Leu28 and Leu31. Arg36 is also stabilized by ionic interactions with Asp184 and the Pro181 main chain carbonyl. Of note, Arg36 is not strongly conserved among PAPRs and is often replaced by hydrophobic residues, like Ile (Figure S1), presumably forming a hydrophobic core with the more strongly conserved Leu28 and Leu31. Functional replacements for Arg36 identified in the present work consisted of Lys and Gln, which both preserve positive charge. Although Pro181 is conserved among PAPRs, it could be replaced by Gln or His and retain function in our screen. Among PAPRs, Asp184 is often replaced by Glu (see ref. 9c, Figure S2), a finding that was recapitulated by our functional assay. In essence, the residues located in Site 1 appear to reinforce both the active site floor and Trx-binding pocket of PAPR.

The cluster of residues in Site 2 consists of: Asp95, Lys96, Leu97, Lys98, Leu99, and Asn100 (Figure 2B). These residues form a loop, which connects α -helix 5 to β -sheet 3, opposite the active site cleft. Within Site 2, Lys96 and Lys98 are less conserved among PAPRs, but are generally replaced with a positively charged or polar residue. Together with Glu92 (which is largely conserved and replaced by Asp in our assay), these collective residues form a highly charged electrostatic patch on the surface of PAPR, which is spatially segregated from the Trx-binding domain. Residues in Site 2 make electrostatic contacts with the main chain groups of Pro73, Glu92 and Gly198, and charged Asn66 and Lys102 side chains. Leu99 participates in hydrophobic packing interactions with Leu62 and Val74. With the exception of Pro73, residues interacting with Site 2 are not substituted by other amino acids in our functional screen and are largely conserved in PAPRs. The location of Site 2, opposite the active site, and the preponderance of charge density hint at the existence of a second interface for biomolecular interaction. One candidate would be sulphite reductase (SiR). Though speculative at this stage, a direct interaction between PAPR and SiR would ensure efficient channeling of sulphite to the next enzyme in the sulphate assimilation pathway.

The third cluster of functionally important residues, Site 3, consists of: Tyr117, Gly118, Lys119, Leu120, and Trp121 (Figure 2C). These residues start at the terminus of α -helix 6 and initiate a 3_{10} -helix adjacent to α -helix 7. Residues of Site 3 are largely conserved among PAPRs. In the structural position analogous to Site 3 in *E. coli* PAPR, the family of enzymes that reduce the related metabolite, APS (APRs) coordinate an 4Fe-4S cluster, which tethers the flexible C-terminal tail to α -helix 6 and is essential for catalysis.^[9g, 9h, 9j] In this context, Trp121, His216, Tyr81 and Tyr131 occupy the space where the 4Fe-4S cluster would be located in APR, ostensibly serving the same function, namely anchoring the PAPR core to the flexible C-terminal segment containing catalytic Cys239. Although Tyr131 forms a hydrogen bond with His216 in the structure of *E. coli* PAPR, Phe effectively replaced the Tyr residue in our screen, implying that this interaction is not critical for function. Nevertheless, hydrophobic, stacking and other electrostatic interactions among Trp121, His216 and Tyr81 are clearly important as none of these residues were substituted in our screen. Interactions are also observed between the main chains of Gly118, Leu120, Gln123 and Tyr117. The side chain of Tyr117 interacts with Glu108, while the side chain of Lys199 is sandwiched in between Glu114 and Glu122, which are largely conserved among PAPRs, but could be functionally replaced by Gly and His/Val, respectively (note that in either case one Glu remains to interact with Lys199). Overall, within Site 3, Trp121 plays a key role in

stabilization of the C-terminal segment, while the remaining residues form a 3_{10} -helix, often considered as an intermediary conformation among helices.^[18]

Finally, the cluster of functional residues of Site 4 consists of Leu144, Lys145, Glu146 and Leu147, which are located at the terminus of α -helix 8 (Figure 2D). Glu146 interacts with the side chains of Tyr104 and Arg 142, which were both unchanged in our study. Main chain interactions were observed between Glu146 and Arg 142 as well as between Leu144 and Ala149. Collectively, the residues of Site 4 appear to stabilize α -helix 8 and support the parallel β stranded Rossmann-fold structure that comprises the PAPS-binding domain. To verify lack of complementation and the functional importance of each new site, we constructed the following site-directed mutants: Leu28Ala (Site 1), Leu31Ala (Site 1), Asn100Ala (Site 2), Tyr117Ala (Site 3), Trp121Ala (Site 3), and Glu146Ala (Site 4). As expected, none of these mutants was able to restore growth in the JM96 *E. coli* strain on minimal media with sulphate as the sole sulphur source (Figure S3). In aggregate, Sites 1–4 are comprised of 19 amino acids. Of these, six or ~30% were individually changed to alanine and re-tested in order to confirm their status.

Conclusions

We have used directed evolution to rapidly discover four new regions of PAPR that are essential for function *in vivo*, but are distinct from the catalytic pocket and the Trx-binding site. In addition to our observations of the lack of functional substitutions at these sites, the essential nature of these regions was further validated by generating alanine mutants, which did not restore growth of JM96. Strict conservation of critical active site residues, the Arg-loop, Cys239, and mapping of the PAPR-Trx interface also highlight the utility of our screen to rapidly identify residues that are essential for enzyme function. The primary sequence of PAPRs is highly conserved; however, we have shown that many conserved residues can undergo non-conservative amino acid exchanges. In this way, we have effectively defined the pattern of conserved residues and variable positions in PAPR, which is essential for the survival of many human pathogens. In future studies, the four novel sites identified in our screen can be probed as possible allosteric sites that regulate the activity of PAPR and also lay the foundation for development of molecules that target regions of PAPR, which are distinct from the active site and smaller than most natural protein-protein interfaces.

Experimental Section

Materials

PCR primers (Integrated DNA Technologies) were obtained as freeze-dried solids and reconstituted to the desired concentration using nuclease-free water. Mutant plasmids used in the screen were generated using the GeneMorph II EZClone Domain Mutagenesis Kit (Agilent).^[19] Plasmids were isolated using standard mini and maxi-isolation kits (Qiagen).

Cloning *E. coli* PAPR into pBR322

The gene encoding *E. coli* PAPR was amplified from *E. coli* genomic DNA obtained from ATCC 700926 (ATCC). PCR reactions contained 0.25 μ M of each primer, 10 ng of the genomic DNA template, and 2.5 units of Pfu DNA polymerase in reaction buffer supplied by

the manufacturer (Promega). Primers were designed to introduce EcoRI and HindIII restriction sites on the 5' - and 3' -ends of the resulting PCR product, respectively. The resulting PCR fragment was directly digested with EcoRI and HindIII, purified and ligated into a similarly digested pBR322 vector by T4 DNA ligase. The ligation reaction was transformed into chemically competent Top10 cells and transformed colonies were grown on LB plates containing 100 mg/ml carbenicillin. Transformants were selected and grown overnight in liquid LB media containing carbenicillin. Plasmid DNA was isolated and screened by PCR and DNA sequencing to verify incorporation of the gene encoding *E. coli* PAPR.

Error-Prone PCR

The amount of DNA template and number of amplification cycles were varied to obtain a mutation rate of ~1 mutation per 1000 bp. PCR reactions contained 0.25 μ M of each primer (forward 5'-atgtccaaactcgatctaaacgc-3' and reverse 5'-ttaccttcgtgtaaccaca-3'), 800 ng of pBR322-PAPR template, and 2.5 units of Mutazyme II DNA polymerase in the reaction buffer supplied by the manufacturer. PCR was performed in an Eppendorf thermal cycler using the following program: 2 min at 95 °C; 10 cycles of 30 s at 95 °C, 30 s at 55 °C, and 90 s at 72 °C; and a 10 min extension at 72 °C. The products of the PCR reaction were analyzed by agarose gel electrophoresis with ethidium bromide staining. The PCR products were then purified and used in the plasmid extension EZ clone reaction (see below), as described in the GeneMorph II EZClone Domain Mutagenesis Kit.^[19]

Plasmid Extension

The purified mutated PCR products serve as 'megaprimers' for the EZClone reaction during which they are denatured, annealed to the original donor plasmid, and extended with a specialized enzyme mix containing a high fidelity DNA polymerase. PCR reactions contained 250 ng of 'megaprimers', 50 ng of pBR322-PAPR template DNA and high fidelity DNA polymerase in reaction buffer supplied by the manufacturer.^[19] PCR was performed in an Eppendorf thermal cycler using the following program: 1 min at 95 °C; 25 cycles of 50 s at 95 °C, 50 s at 60 °C, and 9 mins at 68 °C. The PCR reaction products were analyzed by agarose gel electrophoresis with ethidium bromide staining. The EZ clone reaction yielded an intense band corresponding to the expected size of the *cysH* gene encoding *E. coli* PAPR. PCR products were cooled to 37 °C and treated with 1 μ l of the restriction enzyme DpnI. Reactions were then incubated at 37 °C for 2 h to digest the parental supercoiled dsDNA. A 1.5 μ l aliquot of DpnI-treated DNA was transformed into XL10-Gold ultra-competent cells. Cells were plated on LB-ampicillin agar plates containing 80 μ g/ml X-gal and 20 mM IPTG and incubated overnight at 37 °C. No blue colonies were observed, indicating the restriction of the parent wild-type PAPR plasmid. Cells were scraped from the plates and grown in 1 L cultures. Plasmid DNA was obtained using a maxi-prep kit. The resulting PAPR mutant plasmid library was obtained in a final concentration of approximately at 200 ng/ μ L.

Selection and Isolation of Mutant Constructs

100 ng of the mutant plasmid library was transformed into the *E. coli* JM96 strain and incubated for 20 h at 37 °C on LB plates containing carbenicillin. The resulting colonies were scraped into 1 ml of M9 minimal salts containing 1 mM MgSO₄, 0.1 mM CaCl₂, 0.4%

(v/v) glucose, 0.2% (v/v) mannitol, 0.001% (v/v) thiamine and all amino acids, except methionine and cysteine, and carbenicillin at 100 µg/ml. The resulting solution was dispersed onto M9 minimal media plates of the same composition and incubated at 37 °C for 60 h. The resulting colonies contained mutant PAPR plasmids that complemented the JM96 strain and were therefore defined as functional or sites that were tolerant to mutation. The resulting colonies were scraped using 1 ml of sterile LB media, cultured in 1 L of LB/carbenicillin growth media, and incubated at 37 °C for 12 hours. The culture was harvested and mutant plasmids were isolated using by maxiprep.

Deconvolution of the PAPR Library

100 ng of the mutant PAPR plasmid library was transformed into competent Top10 cells. The cells were dispersed onto LB-carbenicillin agar plates and incubated at 37 °C overnight. Each single colony obtained on the plates was selected, restreaked and transferred to 5 ml of LB-carbenicillin media and grown for 12 h. Plasmids from each colony were prepared by miniprep kit. The resulting plasmids were sequenced using the following primers: 5'-atgttcaaactcgatctaaacgc-3' (forward) and 5'-ttacccttcgtgtaaccaca-3' (reverse). Sequences were obtained for ~500 plasmids and analyzed using the BLAST NCBI sequence comparison tool.

Site Validation

Residues from each site were converted to Ala using the QuikChange II Site-Directed Mutagenesis Kit (Agilent Technologies). The resulting plasmids were sequenced to confirm successful mutation. Plasmids were transformed into JM96, plated onto LB/carbenicillin growth plates and incubated overnight at 37 °C (~1 x 10⁶ CFU/ml). The resulting colonies were scraped into 1 ml of M9 minimal media and dispersed onto M9 minimal media plates. Plates were incubated for 72 h at 37 °C. No bacterial colonies were observed, indicating that each of the selected residues were essential for PAPR function *in vivo*.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the National Institutes of Health, USA (GM087638 to K.S.C.).

References

1. Freire-Moran L, Aronsson B, Manz C, Gyssens IC, So AD, Monnet DL, Cars O. Drug Resist Updat. 2011; 14:118–124. [PubMed: 21435939]
2. a) Livermore DM. Int J Antimicrob Agents. 2012; 39:283–294. [PubMed: 22386741] b) Karaikos I, Giamarellou H. Expert Opin Pharmacother. 2014; 15:1351–1370. [PubMed: 24766095] c) Hawkey PM, Jones AM. J Antimicrob Chemother. 2009; 64(Suppl 1):i3–10. [PubMed: 19675017]
3. a) World Health Organization. World Health Organization, Antimicrobial Resistance: Global Report on Surveillance 2014. 2014. <http://www.who.int/drugresistance/documents/surveillance/en/b> b) Hampton T. JAMA. 2013; 310:1661–1663. [PubMed: 24150445]

4. a) Laxminarayan R, Duse A, Wattal C, Zaidi AK, Wertheim HF, Sumpradit N, Vlieghe E, Hara GL, Gould IM, Goossens H. *Lancet Infect Dis.* 2013; 13:1057–1098. [PubMed: 24252483] b) Perros M. *Science.* 2015; 347:1062–1064. [PubMed: 25745144]
5. Falagas ME, Bliziotis IA, Kasiakou SK, Samonis G, Athanassopoulou P, Michalopoulos A. *BMC Infect Dis.* 2005; 5:24. [PubMed: 15819983]
6. a) Paritala H, Carroll KS. *Infect Disord Drug Targets.* 2013; 13:85–115. [PubMed: 23808874] b) Campanini B, Pieroni M, Raboni S, Bettati S, Benoni R, Pecchini C, Costantino G, Mozzarelli A. *Curr Med Chem.* 2015; 22:187–213. [PubMed: 25388010]
7. a) Sassetti CM, Boyd DH, Rubin EJ. *Proc Natl Acad Sci U S A.* 2001; 98:12712–12717. [PubMed: 11606763] b) Betts JC, Lukey PT, Robb LC, McAdam RA, Duncan K. *Mol Microbiol.* 2002; 43:717–731. [PubMed: 11929527] c) Rengarajan J, Bloom BR, Rubin EJ. *Proc Natl Acad Sci U S A.* 2005; 102:8327–8332. [PubMed: 15928073] d) Provvedi R, Boldrin F, Falciani F, Palu G, Manganelli R. *Microbiology.* 2009; 155:1093–1102. [PubMed: 19332811] e) Turnbull AL, Surette MG. *Res Microbiol.* 2010; 161:643–650. [PubMed: 20600858]
8. a) Bick JA, Dennis JJ, Zylstra GJ, Nowack J, Leustek T. *J Bacteriol.* 2000; 182:135–142. [PubMed: 10613872] b) Kopriva S, Buchert T, Fritz G, Suter M, Benda R, Schunemann V, Koprivova A, Schurmann P, Trautwein AX, Kroneck PM, Brunold C. *J Biol Chem.* 2002; 277:21786–21791. [PubMed: 11940598]
9. a) Carroll KS, Gao H, Chen H, Leary JA, Bertozzi CR. *Biochemistry.* 2005; 44:14647–14657. [PubMed: 16262264] b) Carroll KS, Gao H, Chen H, Stout CD, Leary JA, Bertozzi CR. *PLoS Biol.* 2005; 3:e250. [PubMed: 16008502] c) Chartron J, Carroll KS, Shiau C, Gao H, Leary JA, Bertozzi CR, Stout CD. *J Mol Biol.* 2006; 364:152–169. [PubMed: 17010373] d) Chartron J, Shiau C, Stout CD, Carroll KS. *Biochemistry.* 2007; 46:3942–3951. [PubMed: 17352498] e) Gao H, Leary J, Carroll KS, Bertozzi CR, Chen H. *J Am Soc Mass Spectrom.* 2007; 18:167–178. [PubMed: 17023175] f) Hong JA, Bhavé DP, Carroll KS. *J Med Chem.* 2009; 52:5485–5495. [PubMed: 19678707] g) Bhavé DP, Han WG, Pazicni S, Penner-Hahn JE, Carroll KS, Noodleman L. *Inorg Chem.* 2011; 50:6610–6625. [PubMed: 21678934] h) Bhavé DP, Hong JA, Lee M, Jiang W, Krebs C, Carroll KS. *J Biol Chem.* 2011; 286:1216–1226. [PubMed: 21075841] i) Hong JA, Carroll KS. *J Biol Chem.* 2011; 286:28567–28573. [PubMed: 21673113] j) Bhavé DP, Hong JA, Keller RL, Krebs C, Carroll KS. *ACS Chem Biol.* 2012; 7:306–315. [PubMed: 22023093]
10. a) Chang MW, Belew RK, Carroll KS, Olson AJ, Goodsell DS. *J Comput Chem.* 2008; 29:1753–1761. [PubMed: 18351616] b) Cosconati S, Hong JA, Novellino E, Carroll KS, Goodsell DS, Olson AJ. *J Med Chem.* 2008; 51:6627–6630. [PubMed: 18855373] c) Paritala H, Suzuki Y, Carroll KS. *Tetrahedron Lett.* 2013; 54:1869–1872. [PubMed: 23794759] d) Paritala H, Suzuki Y, Carroll KS. *Nucleosides Nucleotides Nucleic Acids.* 2015; 34:199–220. [PubMed: 25710356] f) Palde PB, Bhaskar A, Pedro Rosa LE, Madoux F, Chase P, Gupta V, Spicer T, Scampavia L, Singh A, Carroll KS. *ACS Chem Biol.* 2016; 11:172–184. [PubMed: 26524379]
11. Panchenko AR, Kondrashov F, Bryant S. *Protein Sci.* 2004; 13:884–892. [PubMed: 15010543]
12. a) Cobb RE, Si T, Zhao H. *Curr Opin Chem Biol.* 2012; 16:285–291. [PubMed: 22673064] b) Zhenwei L, Roccatano D, Lorenz M, Schwaneberg U. *ChemBioChem.* 2012; 13:691–699. [PubMed: 22408062] c) Carter JLL, Bekhouche M, Noiriel A, Blum LJ, Doumèche B. *ChemBioChem.* 2014; 15:2710–2718. [PubMed: 25346488] d) Wechslef C, Meyer D, Loschonsky S, Funk LM, Nuemann P, Ficner R, Brodhun F, Müller M, Tittmann K. *ChemBioChem.* 2015; 16:2580–2584. [PubMed: 26488818] e) Molina-Espeja P, Cañellas M, Plou FJ, Hofrichter M, Lucas F, Guallar V, Alcalde M. *ChemBioChem.* 2016; 17:341–349. [PubMed: 26677801] f) Reetz MT, Kahakeaw D, Lohmer R. *ChemBioChem.* 2008; 9:1797–1804. [PubMed: 18567049]
13. Copp JN, Hanson-Manful P, Ackerley DF, Patrick WM. *Methods Mol Biol.* 2014; 1179:3–22. [PubMed: 25055767]
14. Watson N. *Gene.* 1988; 70:399–403. [PubMed: 3063608]
15. Jones-Mortimer MC. *Biochem J.* 1968; 110:589–595. [PubMed: 4882981]
16. a) Savage H, Montoya G, Svensson C, Schwenn JD, Sinning I. *Structure.* 1997; 5:895–906. [PubMed: 9261082] b) Yu Z, Lemongello D, Segel IH, Fisher AJ. *Biochemistry.* 2008; 47:12777–12786. [PubMed: 18991405]
17. Palde PB, Carroll KS. *Proc Natl Acad Sci U S A.* 2015; 112:7960–7965. [PubMed: 26080424]

18. Enkhbayar P, Hikichi K, Osaki M, Kretsinger RH, Matsushima N. Proteins. 2006; 64:691–699. [PubMed: 16783793]
19. Agilent Technologies. GeneMorph II EZClone Domain Mutagenesis Kit. <http://www.agilent.com/cs/library/usermanuals/Public/200552.pdf>

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 1.

Directed evolution identifies six sites on PAPR that are not tolerant to mutation: substrate-binding site (dark green), Arg-loop (light green), Trx-PAPR interface (blue), and four newly identified sites of functional importance (red). The C-terminal segment (light orange) containing the essential Cys239 nucleophile (dark orange) is shown. The figure was generated using PyMol based upon the coordinates of PDB entries 1SUR, 2O8V, and 2OQ2.

Figure 2. Directed evolution identifies four new sites of functional importance within *E. coli* PAPR. Residues corresponding to each site are colored in green, whereas interacting residues are highlighted in blue. Electrostatic interactions among residues, namely hydrogen bonds, are depicted as yellow dashes.

Scheme 1.

Sulphate assimilation pathways in bacteria. Higher plants and the majority of sulphate-reducing bacteria use adenosine-5'-phosphosulphate (APS) as their source of sulphite (upper pathway). In some organisms, APS kinase phosphorylates APS at the 3'-hydroxyl to produce 3'-phosphoadenosine 5'-phosphosulphate (PAPS) as the source of sulphite (lower pathway). The lower pathway of sulphate reduction is utilized by γ -proteobacteria such as *E. coli* and some fungi. Depending on the organism, APS or PAPS is reduced to sulphite by APS reductase (APSR) and PAPS reductase (PAPR), respectively. Sulphite is reduced to sulphide by sulphite reductase and incorporated into reduced-sulphur biomolecules, like cysteine.