Structural basis for the regulation of human 5,10-methylenetetrahydrofolate reductase by phosphorylation and S-adenosylmethionine inhibition

D. Sean Froese1, Jolanta Kopec2, Elzbieta Rembeza2, Gustavo Arruda Bezerra2, Anselm Erich Oberholzer3, Terttu Suormala1, Seraina Lutz1, Rod Chalk2, Oktawia Borkowska2, Matthias R. Baumgartner1 & Wyatt W. Yue2

The folate and methionine cycles are crucial for biosynthesis of lipids, nucleotides and proteins, and production of the methyl donor S-adenosylmethionine (SAM). 5,10-methylenetetrahydrofolate reductase (MTHFR) represents a key regulatory connection between these cycles, generating 5-methyltetrahydrofolate for initiation of the methionine cycle, and undergoing allosteric inhibition by its end product SAM. Our 2.5 Å resolution crystal structure of human MTHFR reveals a unique architecture, appending the well-conserved catalytic TIM-barrel to a eukaryote-only SAM-binding domain. The latter domain of novel fold provides the predominant interface for MTHFR homo-dimerization, positioning the N-terminal serine-rich phosphorylation region near the C-terminal SAM-binding domain. This explains how MTHFR phosphorylation, identified on 11 N-terminal residues (16 in total), increases sensitivity to SAM binding and inhibition. Finally, we demonstrate that the 25-amino-acid inter-domain linker enables conformational plasticity and propose it to be a key mediator of SAM regulation. Together, these results provide insight into the molecular regulation of MTHFR.
In humans, the folate and methionine cycles both generate products essential to cellular survival. Folate, the major cellular carrier of single carbon units, is required for the synthesis of purines and thymidine monophosphate. Within the methionine cycle, the methylation of homocysteine to methionine by methionine synthase (EC 2.1.1.13) produces an essential amino acid which may be used for protein synthesis or, crucially, be further converted to S-adenosylmethionine (SAM), a vitally important donor for the methylation of DNA, RNA and proteins as well as the creation of numerous methylated compounds. These two cycles intersect at the enzyme 5,10-methylene-tetrahydrofolate reductase (MTHFR; E.C. 1.5.1.20). MTHFR catalyzes the physiologically irreversible reduction of 5,10-methylene-tetrahydrofolate (CH₂-THF) to 5-methyl-tetrahydrofolate (CH₃-THF), a reaction requiring FAD as a cofactor and NADPH as an electron donor. Since the product CH₃-THF is exclusively used by methionine synthase, and only the demethylated form (THF) may be recycled back to the folate cycle, MTHFR commits THF-bound one-carbon units to the methionine cycle.

In accordance with this essential role, major and minor deficiencies of human MTHFR are the direct or indirect causes of human disease. Severe MTHFR deficiency (MIM #607093) is inherited in an autosomal recessive manner and is the most common inborn error of folate deficiency with ~200 patients known. To date, over 100 different clinically relevant mutations in MTHFR have been described, the majority of which are of the missense type (n = 70, >60%) and private. Milder enzyme deficiencies, due to single nucleotide polymorphisms of the MTHFR gene, have been associated with various common disorders. The most studied of these is p.Ala222Val (c.665C>T in NM_001330358, commonly annotated as c.667C>T), identified as a risk factor for an overwhelming number of multifactorial disorders, including vascular diseases, neurological diseases, various cancers, diabetes and pregnancy loss (see e.g. review by Liew and Gupta).

Human MTHFR is a 656 amino acid multi-domain protein (Fig. 1). The catalytic domain is conserved across evolution, and crystal structures of MTHFR from Escherichia (E. coli) and Thermus thermophilus, in which the catalytic domain constitutes the entire sequence (Fig. 1), have been solved. These structures reveal the catalytic domain to form a βαβα (TIM) barrel and have uncovered residues critical for binding the cofactor FAD, the electron donor NADPH (NADH in bacteria) and the product CH₃-THF. The bacterial structures, together with activity assay of trypsin cleaved porcine MTHFR, indicate that the catalytic domain is sufficient for the entire catalytic cycle. Eukaryotic MTHFR orthologs additionally possess a C-terminal regulatory domain that is connected to the catalytic domain by a linker sequence (Fig. 1). This C-terminal domain is able to bind SAM, resulting in allosteric inhibition of enzymatic activity, an effect which is very slow and can be reversed by binding to S-adenosylhomocysteine (SAH), the demethylated form of SAM.

Human MTHFR further contains a 35 amino acid serine-rich region at the very N-terminus which is not found in MTHFR orthologs of bacteria, yeast or even lower animals (Fig. 1). This region has been identified to be multiply phosphorylated following heterologous expression in insect cells and yeast, or following immunoprecipitation from human cancer cell lines. Phosphorylation has been associated with moderately decreased catalytic activity and increased total inhibition mediated by SAM. Although phosphorylation mapping of this region has been thus far unsuccessful, scanning mutagenesis has revealed substitution of alanine for threonine at position 34 (p.Thr34Ala) to almost completely block phosphorylation, suggesting Thr34 is the priming position. The cellular relevance of this modification remains unclear, although one group has suggested that phosphorylation at Thr34 can be accomplished by CDK1/cyclin B1 (ref. 16) and at Thr549 by polo-like kinase 1 (ref. 17) whereby they posit a role in histone methylation and replication. The repertoire of bacterial MTHFR structures to date does not provide any mechanistic insight into the enzymatic regulation by phosphorylation and SAM binding, because both features are absent in prokaryotes. To this end, we have combined structural, biophysical and biochemical data of human MTHFR to provide a molecular view of MTHFR function and regulation in higher eukaryotes. We have identified specific phosphorylation sites and demonstrate a distinct relationship between phosphorylation and SAM inhibition. Further, using our 2.5 Å resolution crystal structure of the almost full-length human protein, we reveal that the regulatory domain utilizes a unique topology to bind SAH/SAM and transmit a catalytic inhibition signal, likely by long-range conformational change through the linker region.

**Results**

**Phosphorylated residue identification by mass spectrometry.**

To examine the phosphorylation status of human (Hs)MTHFR, we generated full-length recombinant human MTHFR (HsMTHFR, c.656) by baculovirus expression in S9 cells. Mass spectrometry-based phosphorylation mapping (with 92% coverage) identified 16 separate phosphorylation sites in HsMTHFR, c.656, following purification from S9 cells (called here as purified) (Fig. 2a). All phosphorylation sites were considered to have partial occupancy, since no residues were phosphorylated in every tryptic peptide analysed (Supplementary Fig. 1). Of these, 11 phosphorylated amino acids (Ser9, Ser10, Ser18, Ser20, Ser21, Ser23, Ser25, Ser26, Ser29, Ser30, Thr34) were within the N-terminal serine-rich region, including the putative phosphorylation determining residue Thr34 (Fig. 2a). Additionally, we found phosphorylation of three further amino acids in the catalytic domain (Tyr90, Thr94, Ser103) and two in the regulatory domain (Ser394, Thr451). Up to ten phosphorylation sites were identified to be occupied simultaneously, whereby treatment with calf intestine alkaline phosphatase (CIP) resulted in removal of 9 (Fig. 2b) or 10 (Fig. 2c) phosphate groups, as identified by
Phosphorylation does not alter MTHFR kinetic parameters.

Phosphorylation has been described to alter MTHFR kinetics, and comparable to that of MTHFR from human fibroblast lysates (12.4 µmol min\(^{-1}\) mg\(^{-1}\)) and puriﬁed porcine MTHFR (19.4 µmol min\(^{-1}\) mg\(^{-1}\)). Importantly, CIP-treated HsMTHFR\(_{38-644}\) (dephosphorylated) retained nearly identical activity values to mock-treated HsMTHFR\(_{38-644}\) (i.e. assay without addition of CIP; phosphorylated) (Table 1), confirming that phosphorylation does not decrease enzymatic turnover. Likewise, Michaelis–Menten constants (\(K_M\)) for the substrate CH\(_2\)-THF (range: 21.3–25.5 µM) and electron donor NADPH (range: 23.5–35.5 µM) were very similar for all four proteins (Table 1), and comparable to that of MTHFR from human fibroblast lysates (CH\(_2\)-THF: 26 µM; NADPH: 30 µM\(^{18}\)). Thus, we conclude that our assay is sensitive and specific, and that phosphorylation does not signiﬁcantly alter MTHFR kinetic properties.

Interestingly, we found no increase in the speciﬁc activity of MTHFR proteins following addition of exogenous FAD to the assay buffer (Table 1). Since FAD is required as cofactor for the MTHFR reaction, this suggests the cofactor was already bound to the as puriﬁed protein, presumably acquired during cellular

---

**Fig. 2** Phosphorylation status of HsMTHFR\(_{38-644}\) and HsMTHFR\(_{656-656}\). a Phosphorylation mapping of HsMTHFR\(_{656-656}\). The protein sequence is given as amino acids in single letter code, including the C-terminal His/flag-tag (underlined). Black font represents amino acids identiﬁed by the mass spectrometer (covered), blue font represents amino acids not identiﬁed (non-covered), red font represents phosphorylated amino acids. Domains are coloured as in Fig. 1. b Dephosphorylation of HsMTHFR\(_{656-656}\) following treatment with CIP. Treatment time at 37 °C is given. Large number above peaks represents number of phosphate groups attached. Proteins were analysed by denaturing mass spectrometry, amu, atomic mass units. c Native mass spectrometry analysis of HsMTHFR\(_{656-656}\) before and after treatment with CIP. Upper panel: as puriﬁed (untreated) protein. Monomer represents protein bound to 1 FAD plus 10 phosphate groups (expected mass: 76831.16 amu); dimer represents protein bound to 2 FADs and 1 SAM plus 20 phosphate groups (expected mass: 154060.74 amu). Lower panel: protein following 180 min treatment with CIP. Monomer represents protein bound to 1 FAD (expected mass: 76031.16 amu); dimer represents protein bound to 2 FADs and 1 SAH (expected mass: 152457.47 amu). Expected sizes: protein without ﬁrst methionine, 75245.6 amu; FAD, 785.56 amu; SAM, 398.44 amu; SAH, 384.42 amu, phosphate, 80.00 amu. * indicates a truncated protein representing amino acids 353–663 (expected mass: 36136.6 amu). d HsMTHFR\(_{38-644}\) before and after treatment with CIP. Treatment time at 37 °C is given. Protein was analysed by denaturing mass spectrometry. e Native mass spectrometry of HsMTHFR\(_{656-656}\) identifying sequential binding of SAM or SAH. Graphs represent areas zoomed in on dimeric protein. Upper panel: As puriﬁed (untreated) protein. Middle panel: control protein (heated in assay buffer for 180 min without CIP). Bottom panel: treated protein (heated in assay buffer for 180 min with CIP). Expected size of protein with 2 FAD bound and 20 phosphates: 153662.3. Expected size of protein with 2 FAD bound and 0 phosphates: 152062.32. Expected size of SAM: 398.44; SAH: 384.41.
expression. This is consistent with native mass spectrometry, which identified monomeric and dimeric forms of as purified HsMTHFR1–656 which, in addition to phosphorylation, contained equivalent units of FAD (Fig. 2c). HsMTHFR38–644 also presented as monomeric and dimeric forms bound to equivalent units of FAD (Supplementary Fig. 3b), suggesting phosphorylation has no effect in this regard. Supplementation with FAD, however, helped rescue activity when provided either before, or to a lesser extent following, incubation of these proteins at 46 °C for 5 min (Table 1). Therefore, this cofactor, which is important for protein stability, may be lost under heat treatment. In our experiments, HsMTHFR1–656 was markedly more sensitive to heat inactivation than HsMTHFR38–644, but this heat sensitivity was not affected by the phosphorylation state of the protein, and therefore likely rather reflects overall protein stability.

Phosphorylation increases SAM protection and sensitivity. In addition to phosphorylation and FAD, native mass spectrometry identified the dimeric form of as purified HsMTHFR1–656, to contain 0, 1 or 2 units of SAM (Fig. 2e). Like FAD, SAM was likely acquired during cellular expression. However, following CIP treatment, the SAM bound to MTHFR was found to degrade to SAH, a chemical transition which did not occur during mock treatment of the protein (Fig. 2e). Correspondingly, as purified dimeric HsMTHFR38–644 which is not phosphorylated, was found to be bound to 0, 1 or 2 units of SAH, but not SAM (Supplementary Fig. 3c). Thus, phosphorylated MTHFR appeared to protect thermally unstable SAM from degradation to SAH, while the non-phosphorylated protein was unable to perform this function.

Phosphorylation has been identified to affect the maximum degree of inhibition of MTHFR by SAM, whereby phosphorylated protein was found to be maximally ~80% inhibited, while phosphatase treated protein was maximally ~60% inhibited44. At high concentrations of SAM (>200 μM), we were able to inhibit all recombinant HsMTHFR proteins by over 95%, regardless of the phosphorylation state (Fig. 3). However, at low SAM concentrations we found phosphorylated HsMTHFR1–656 to be more sensitive to SAM inhibition than HsMTHFR38–644 and dephosphorylated HsMTHFR1–656 (Fig. 3). Further analysis revealed as purified and mock-treated HsMTHFR1–656 to have inhibition constants (K_i) of ~3 μM, while CIP-treated HsMTHFR1–656 was approximately two-fold less sensitive to SAM inhibition, and HsMTHFR38–644 seven-fold less sensitive (Fig. 3, inset; Table 1). Thus, although phosphorylation does not directly affect MTHFR enzymatic activity, it increases the protein’s sensitivity to SAM inhibition.

An extensive linker connects and interacts with both domains. We determined the 2.5 Å resolution structure of HsMTHFR38–644 in complex with FAD and SAH by multiple-wavelength anomalous dispersion using the selenomethionine (SeMet) derivatized protein (Table 2). The identity of both ligands is guided by well-defined electron density (Supplementary Fig. 5), and in line with native mass spectrometry analysis for this construct (Supplementary Fig. 3B). The HsMTHFR38–644 protomer folds into two globular domains (Fig. 4a) to form an overall elongated molecule. As predicted from bacterial structures, the N-terminal catalytic domain (aa 40–337) consists of an 8α/8β TIM barrel, adorned with three extra α-helices (a8, a9 and a11) (Fig. 4a, Supplementary Fig. 6). The C-terminal regulatory domain (aa 363–644) makes up a novel fold of two five-stranded β-sheets arranged side-by-side in the core, flanked by a number of α-helices (Fig. 4a, Supplementary Fig. 6). The two domains do not contact each other directly, but instead are connected by an extended linker region encompassing aa 338–362 (Fig. 4a), with its amino acid sequence rich in Arg (n = 4), Pro (4), Glu (3) and Leu (3). Mediated by three β-turns, this linker makes multiple intricate contacts with both domains, and changes direction twice in traversing between the catalytic and regulatory domains.

The HsMTHFR38–644 structure allows the mapping of the 70 inherited missense mutations known to cause severe MTHFR
deficiency, which lie on 64 different residues of the polypeptide (Supplementary Fig. 7). Twice as many of the mutation sites are found in the catalytic domain \( (n = 38) \) as the regulatory domain \( (20) \), with the remainder \( (6) \) found in the linker region. By proportion, however, the linker region has a higher density \( (24\%) \) of mutation sites than the catalytic \( (11\%) \) and regulatory \( (7\%) \) domains. Additionally, a number of sites in the catalytic and regulatory domains are in direct contact with the linker region. Further, the most severe mutations, those found either homozgyously or in conjunction with a truncating mutation to result in enzymatic activity below 1.5% of control towards the interface of the two regulatory domains (Fig. 4c), and may contribute further to the dimer contacts.

The essential interfacial residues from the regulatory domain are contributed predominantly from the two central \( \beta \)-sheets, including a \( \beta \)-turn \( (\beta 11–\beta 12) \), strand \( \beta 16 \), and the loop encompassing Asn386–Asn391 (Supplementary Fig. 8), which buries in total \( ~1330 \) \( \AA^2 \) of accessible surface. Half of the sites of missense mutations in the regulatory domain causing MTHFR deficiency \( (n = 10, \) Supplementary Fig. 7) either participate in, or are within two residues of, the dimerization site.

Within the homodimer, each of the two catalytic domains is presented away from the dimeric interface and their active sites are at opposite ends of the overall shape and face away from each other (Fig. 4b). In this arrangement, the catalytic domain displays a large degree of flexibility in relative orientation with the regulatory domain. In fact, this is reflected in our structure whereby the catalytic domain of one dimer subunit (chain A) is ordered, while that of the other dimer subunit (chain B) is highly disordered, to the extent that only main chain atoms of the amino acid 40–58, 129–134 and 155–342 in chain B could be modelled.

An asymmetric MTHFR dimer with inter-domain flexibility. The \( HsMTHFR_{38-644} \) structure reveals a homodimer (Fig. 4b), consistent with native mass spectrometry (Supplementary Fig. 3b) and previous investigation of mammalian MTHFR by size exclusion chromatography and scanning transmission electron microscopy\(^9\). It was previously thought that MTHFR homodimerizes in a head-to-tail manner, where the regulatory domain of one subunit interacts with the catalytic domain of the other subunit\(^3\). Unexpectedly, in our structure dimmerization is mediated almost entirely by the regulatory domain (Fig. 4b), although the first ordered residue in chain A (Glu40) is located around 5–6 \( \AA \) from the regulatory domain of chain B (e.g. Glu553, Arg567). The N-terminal sequence that is either not present (Ser-rich phosphorylation region, aa 1–37) or present but disordered (aa 38–39) in the \( HsMTHFR_{38-644} \) structure will likely project towards the interface of the two regulatory domains (Fig. 4c), and may contribute further to the dimer contacts.

The essential interfacial residues from the regulatory domain are contributed predominantly from the two central \( \beta \)-sheets, including a \( \beta \)-turn \( (\beta 11–\beta 12) \), strand \( \beta 16 \), and the loop encompassing Asn386–Asn391 (Supplementary Fig. 8), which buries in total \( ~1330 \) \( \AA^2 \) of accessible surface. Half of the sites of missense mutations in the regulatory domain causing MTHFR deficiency \( (n = 10, \) Supplementary Fig. 7) either participate in, or are within two residues of, the dimerization site.

Within the homodimer, each of the two catalytic domains is presented away from the dimeric interface and their active sites are at opposite ends of the overall shape and face away from each other (Fig. 4b). In this arrangement, the catalytic domain is not involved in oligomerization, unlike bacterial and archaeal MTHFR proteins (Supplementary Fig. 9). This said, the N-terminus of the \( HsMTHFR_{38-644} \) construct is projecting towards the dimer interface. A direct consequence of the dimeric architecture is that the \( HsMTHFR \) catalytic domain displays a large degree of flexibility in relative orientation with the regulatory domain. In fact, this is reflected in our structure whereby the catalytic domain of one dimer subunit (chain A) is ordered, while that of the other dimer subunit (chain B) is highly disordered, to the extent that only main chain atoms of the amino acid 40–58, 129–134 and 155–342 in chain B could be modelled.

Dynamics of MTHFR observed by solution scattering. Our \( HsMTHFR_{38-644} \) crystal structure has captured the snapshot of an asymmetric dimer whereby the two catalytic domains have different orientations with respect to their own regulatory domains (Supplementary Fig. 10). We applied small-angle X-ray scattering (SAXS) to understand better the different conformational

---

### Table 2 Data collection and refinement statistics

|                      | \( HsMTHFR_{38-644} \) Native | \( \text{SeMet (MAD)} \) | \( \text{ScMET12}_{5-301} \) |
|----------------------|-------------------------------|--------------------------|-------------------------------|
| **Data collection**  |                               |                          |                               |
| Space group          | \( P_2 \times 2 \times 2 \)  | \( P_2 \times 2 \times 2 \) | \( P_2 \times 2 \times 2 \)  |
| Cell dimensions      | \( a = 127.9, 147.0 \) \( \AA \) | \( a = 127.24, 147.29 \) \( \AA \) | \( a = 110.6, 54.5, 61.9 \) \( \AA \) |
| \( \alpha, \beta, \gamma \) | \( 90.00, 90.00, 90.00 \) | \( 90.00, 90.00, 90.00 \) | \( 90.00, 90.00, 90.00 \) |
| Wavelength           | 0.9281                        | 1.00650                  | 0.99835                      |
| Resolution (\( \AA \)) | 68.53–2.50 (2.64–2.50)         | 48.62–3.00 (3.11–3.00)   | 22.75–1.56 (1.60–1.56)       |
| \( R_{\text{sym}} \) / \( R_{\text{merge}} \) | 8.6 (136.2)                   | 3.8 (58.6)               | 4.1 (60.4)                   |
| Completeness (%)     | 100 (100)                     | 100 (99.9)               | 100 (99.9)                   |
| Redundancy           | 7.5 (7.0)                     | 2.0(2.0)                 | 2.0(2.0)                     |
| **Refinement**       |                               |                          |                               |
| Resolution (\( \AA \)) | 96.51–2.50                    | 22.75–1.56               | 448.825/53.602               |
| No. of reflections (total/unique) | 479,094/64,233 | 448,825/53,602 | 15.97/19.93 |
| \( R_{\text{work}} / R_{\text{free}} \) | 21.03/24.75                   | 15.97/19.93              |                               |
| No. of atoms         |                               |                          |                               |
| Protein              | 8914                          | 2440                     |                               |
| Ligand/ion           | 184                           | 53                       |                               |
| Water                | 72                            | 258                      |                               |
| \( B \)-factors      |                               |                          |                               |
| Protein              | 92.29                         | 19.78                    |                               |
| Ligand/ion           | 98.25                         | 23.38                    |                               |
| Water                | 64.09                         | 34.65                    |                               |
| R.m.s. deviations    |                               |                          |                               |
| Bond lengths (\( \AA \)) | 0.012                        | 0.012                    | 0.012                        |
| Bond angles (\( ^\circ \)) | 1.296                      | 1.438                    | 1.438                        |
protein conformation, we next collected SAXS data for full-length HsMTHFR. a, Orthogonal views of HsMTHFR showing the catalytic domain (cyan), the linker (red) and the regulatory domain (yellow). Bound FAD (green) and SAM (pink) are shown in sticks. Dotted lines indicate disordered regions that are not modelled in the structure. α-Helices and β-sheets are labelled, and correspond to the multiple sequence alignment in Supplementary Figs. 13 and 15 and the topology depicted in Supplementary Fig. 6. b, Homodimer of HsMTHFR as seen in the crystal. Chain A represents the more ordered protomer. In chain B, the part of the catalytic domain that was poorly ordered is represented in dark blue and only the main chain was modelled. The FAD in this subunit was also partially disordered. c, Ligation of the N-terminus of HsMTHFR protomer towards both regulatory domains of the homodimer. The first residue observed in the structure, Glu40 of chain A, is shown in black spheres. Other coloured features are as described for panel a.

**Fig. 5** SAXS analysis of HsMTHFR38–644 and HsMTHFR1–656. Phosphorylated and dephosphorylated. SAXS analysis of HsMTHFR38–644. Experimental scattering profile is shown in black, theoretical scattering curve of the HsMTHFR38–644 dimer observed in the crystal is in green and that of the rigid body modelling by CORAL is in red. Chi² was determined by CRYSOL. Variations assumed by the protein in solution. Superimposition of the theoretical scattering curve back-calculated from the crystal structure dimer against experimental data obtained from HsMTHFR38–644 in solution revealed a poor fit (Chi² 14.8; Fig. 5), suggesting this is not the predominant conformation in solution. However, by employing CORAL to simulate relaxation of the relative orientations of the catalytic and regulatory domains (by allowing flexibility in residues 338–345 of the linker), and thus also permitting rigid body movement of these subunits in relative orientation to each other, we obtained a significantly improved fit (Chi² 5.5; Fig. 5). Thus, consistent with our finding from the crystal structure, HsMTHFR retains a significant degree of intra- and inter-domain conformational flexibility in solution.

To further investigate the influence of phosphorylation on protein conformation, we next collected SAXS data for full-length HsMTHFR1–656 as purified (i.e. phosphorylated and bound with SAM) and treated with CIP (i.e. dephosphorylated and bound with SAH). The experimental scattering curves for as purified and CIP-treated HsMTHFR1–656 gave rise to slightly different profiles and derived parameters (Supplementary Fig. 11, Supplementary Table 1), although both protein forms are consistent with a dimeric configuration. These data were further corroborated by charge radius analysis of native phosphorylated and dephosphorylated HsMTHFR1–656 by electrospray ionization mass spectrometry, showing that the charge-distribution of protein ions is shifted between the two protein forms. This may indicate a conformational change equivalent to a 0.5% change in radius (Supplementary Fig. 12). Together, we interpret these results to suggest that the phosphorylated SAM-bound form of the protein may present a different conformation to the dephosphorylated SAH-bound form, which merits future investigation using alternative methods.

**Subtle features provide for eukaryotic NADPH specificity.** The MTHFR catalytic domain adopts a TIM-barrel structure evolutionarily conserved across all kingdoms. In addition to HsMTHFR1–656 we further determined the catalytic domain structure of the yeast homologue MET12 (ScMET12–301) to 1.56 Å resolution (Table 2). This enables a structural comparison across mammalian (HsMTHFR), low eukaryotic (ScMET12) and bacterial (E. coli, H. influenzae, T. thermophiles) orthologues. Consistent with their sequence conservation (Supplementary Fig. 13), the catalytic domains have highly superimposable folds (main chain RMSD: 1.85 Å), although distinct local differences are found in low homology loop regions (Fig. 6a, 1–2) and helices (Fig. 6a, 3–4). Additionally, the first helix of the catalytic domain (α1) is observed in different orientations among these structures (Supplementary Fig. 14). There is sequence divergence of helix α1 among prokaryotes, lower and higher eukaryotes (Supplementary Fig. 15). In HsMTHFR (which contains, in its biological sequence but not present in the crystallized construct, the serine-rich phosphorylation region N-terminal to the catalytic domain), ScMET12 and T. thermophiles MTHFR, this helix α1 is projected towards the interface between catalytic and regulator domains.

In HsMTHFR38–644, clear electron density for FAD was observed in the TIM barrel of chain A (Supplementary Fig. 5a). However, there is high disorder in the TIM-barrel of chain B particularly around the FAD binding site, implying a low ligand occupancy of the ligand, although native mass spectrometry of
the crystallized construct indicated two FADs bound per homodimer (Supplementary Fig. 3b). Analysis of the FAD binding residues in HsMTHFR chain A (Fig. 6b) reveals perfect overlap with those predicted from the EcMTHFR structure4. These include Thr129, Arg157, Ala175 and Ala195 of HsMTHFR, which were associated with in vitro FAD responsiveness when mutated in severe MTHFR deficiency20,22 (Fig. 6b; Supplementary Fig. 7). The bi bi kinetic mechanism of MTHFR necessitates the electron donor NAD(P)H and substrate CH2-THF to interact in turn with FAD for transfer of the reducing equivalents, and hence to share the same binding site. In our structures, the FAD ligand adopts a conformation poised to expose the si face of the isalloxazine ring for the incoming NAPDH and CH2-THF. However, instead of trapping the electron donor or substrate (despite multiple attempts at co-crystallization), the binding site in ScMET121-301 and subunit A of HsMTHFR38-644 is blocked by a crystal packing interaction from a nearby symmetry mate, making π–π stacking interactions with the FAD ligand (Supplementary Fig. 16). By contrast, no crystal packing interaction is found in the chain B binding site of HsMTHFR38-644, explaining the overall mobility and disorder of its catalytic domain.

Superimposing the HsMTHFR38-644 structure with that of EcMTHFR bound with NADH (Fig. 6c) and CH2-THF (Fig. 6d) demonstrates that the human enzyme has largely preserved the same shared binding site found in prokaryotes, with Gln228, Gln267, Lys270, Leu271 and Leu323 likely to be important for interacting with both NAD(P)H and CH2-THF. EcMTHFR preferentially utilizes NADH23, and its NADH-bound structure reveals a highly uncommon bent conformation24 for the electron donor, where the nicotinamide ring stacked over the adenine base mediates π–π interactions7. Our activity assay of HsMTHFR38-644 and HsMTHFR1-656 clearly demonstrates an ~100-fold preference for NADPH compared to NADH as an electron donor (Table 1), in agreement with previous enzyme studies from pig11,23 and rat11 MTHFRs.

Within the HsMTHFR active site, we did not identify any obvious differentiating features surrounding the modelled NADH, which could indicate how the extra 2′-monophosphate group on the NADPH ribose is accommodated for the compact stacked conformation for NADPH26. Modelling an NADPH ligand with such a stacked conformation onto the HsMTHFR38-644 structure reveals severe steric clashes with helix a8 (Supplementary Fig. 17), which creates the floor of the NAD(P)H binding site (e.g. via Gln267, Lys270 and Leu271). Helix a8 is poorly aligned with bacterial and low eukaryotic orthologues in both amino acid sequence (Supplementary Fig. 13) and structural topology.
incubated with SAM

Hs

position and mobility, we propose that residue(s) on helix A novel fold for the SAM-binding regulatory domain

Hs

Ala229, respectively. (Supplementary Fig. 13). Therefore, given its Hs

conserved in

of this fold comprises two mixed

of the regulatory domain unique to eukaryotic MTHFR. The core (Supplementary Fig. 6). Strand β10 from one sheet forms a continuous segment with β11 from the other sheet, and similarly β12 from one β-sheet continues onto β13 of the other sheet. The threading of the two central β-sheets are interspersed with three loop extensions containing different numbers of α-helices (α12–α15, α16, and α17–α18). To the best of our knowledge, the MTHFR regulatory domain represents a unique SAM binding architecture distinct from the 18 known classes of SAM-dependent methyltransferases and non-methyltransferases (Supplementary Fig. 18). Further, a DALI search of this domain did not yield any structural homologue, and we found no existing annotation in PFAM/CATH/SCOP databases and no sequence for this domain beyond eukaryotic MTHFR homologues. Therefore, this appears to be a novel fold utilized only by MTHFR for SAM binding/inhibition.

In our structure, SAH is bound in an extended conformation within the part of the regulatory domain (Fig. 7a) that faces the catalytic domain. Indeed, part of the binding site is constituted by the linker region itself. The ligand is sandwiched between the loop segment preceding α15 (N465-DEPAAET468) and the first strand β10 (T481INSQ485) of the central β-sheets, where a number of conserved residues are found. For example, Thr481 (conserved in 96% of 150 orthologues; Consurf29) and Ser484 (98%) hydrogen-bond to the SAH adenine moiety, while Glu463 (99%) and Thr464 (62%) fixate the ribose hydroxyl groups. The strongest sequence conservation in the SAH binding site is found around the homocysteine moiety, including Pro348 (invariant) and Trp349 (99%) from the linker region, as well as Thr560 and Thr573 (both invariant) at the start and end of the β15-β16 turn. The SAH homocysteine sulphur atom is loosely contacted by Glu463 (3.8 Å) and Ala368 (3.7 Å). SAM is expected to bind to the same site in the regulatory domain, in a similar extended configuration as SAH and requiring the same set of binding residues. However, the additional methyl group in the sulphonium centre of SAM would create a steric clash to the Ala368 position of the structure (inter-residue distance ~2.0 Å between heteroatoms, and <1.5 Å between hydrogen atoms) (Supplementary Fig. 19). Although not strictly conserved (45% of 150 orthologues), conservation of Ala368 follows a similar evolutionary pattern as the MTHFR domain organization (Fig. 1): in higher animals alanine is invariant; lower animals may accommodate a serine; while lower eukaryotes often incorporate a bulky residue (e.g. lysine) (see Supplementary Fig. 15). Therefore, in higher organisms such as humans, SAM binding likely results in conformational rearrangement of the loop region containing Ala368 to accommodate its methyl moiety.

![SAH/SAM binding and conformational change](image-url)

**Fig. 7** SAH/SAM binding and conformational change. a The SAH binding site. Amino acids that contribute to binding are labelled and shown in black sticks. SAH is shown in green sticks. b Size exclusion chromatography of HsMTHFR with various N-terminal truncations following incubation with SAM (dashed lines), SAH (dotted lines), or buffer (apo; solid line). For each N-terminally truncated construct, the corresponding structure is shown. c Size exclusion chromatography of HsMTHFR proteins without (wt) or with (wt-SAM) pre-incubation with SAM. Mutated HsMTHFR proteins were pre-incubated with SAM.
The linker mediates SAM-dependent conformational change. Since there is no direct interface between the active site of the catalytic domain and the regulatory domain (Fig. 4), SAM binding must elicit enzymatic inhibition via a conformational change propagated from the regulatory to catalytic domain. The most likely effector of this conformational change is the extended linker region (defined as aa 338–362), since it makes multiple contacts to both the regulatory and catalytic domains (Fig. 4) and forms part of the SAM/SAH binding site (Fig. 7a). To investigate the potential of this region to elicit conformational change following SAM binding, we generated recombinant HsMTHFR proteins consisting of the regulatory domain alone attached to progressively shorter linker regions, where the N-terminus of these constructs would become Pro348 (HsMTHFR348–656), Arg357 (HsMTHFR357–656) and Arg377 (HsMTHFR377–656) (Fig. 7b; Supplementary Fig. 2). All three constructs are sufficient to bind SAM and SAH, as demonstrated by dose-dependent increases in thermostability by differential scanning fluorimetry when exposed to increasing concentrations of each ligand (Supplementary Fig. 20a). This again reinforces the catalytic and regulatory domains as separate binding modules for their cognate ligands (FAD/NADPH/CH₃-TF vs SAM/SAH respectively).

We employed analytical size exclusion chromatography (aSEC) as a means to study solution behaviour of the MTHFR regulatory domain in response to SAM/SAH binding. Exposure of MTHFR348–656 to either SAH or SAM resulted in shifts of elution volume (Vₑ) compared to as purified (apo-) protein (Fig. 7b), in contrast to HsMTHFR1–656 and HsMTHFR337–644, which did not show changes in Vₑ despite SAH or SAM binding (Supplementary Fig. 21). Importantly, for MTHFR348–656 SAM resulted in a leftward Vₑ shift (suggestive of a larger hydrodynamic volume) and SAH a rightward shift (suggestive of a smaller hydrodynamic volume) (Fig. 7b). By contrast, MTHFR357–656 showed a shift in Vₑ only when exposed to SAM, and MTHFR377–656 did not change when exposed to either ligand (Fig. 7b). A similar pattern of results were observed when using purified recombinant mouse MTHFR of the same protein boundaries (Supplementary Fig. 20a and b). Therefore, we conclude that residues within 357–377 must contribute to change of protein state upon SAM binding, which we interpret as a change in conformation.

Next we carried out site-directed mutagenesis to define residues involved in SAM binding, and/or SAM-mediated conformational change as observed in the aSEC experiment. We reasoned that mutation of Glu463 (which hydrogen-bonds a ribose oxygen) could lead to loss of SAH/SAM binding, and thus conformational change. Indeed, conservative mutation of Glu463 to either aspartate (p.E463D) or glutamine (p.E463Q) on MTHFR348–656, which did not change when exposed to either ligand (Fig. 7b). A similar pattern of results were observed when using purified recombinant mouse MTHFR of the same protein boundaries (Supplementary Fig. 20a and b). Therefore, we conclude that residues within 357–377 must contribute to change of protein state upon SAM binding, which we interpret as a change in conformation.

We mapped the entire phosphorylation landscape of HsMTHFR, revealing phosphorylated Ser/Thr not only at the far N-terminus (n = 11) as predicted from the sequence but also within the catalytic (3) and regulatory (2) domains. Many of the N-terminal phosphorylation sites identified are consistent with previous mutation analysis14, including Thr34 (refs. 14–16). The phosphorylated residues detected in the catalytic and regulatory domains were not reported before. Interestingly, two phosphorylated sites are located within the FAD binding site, although their physiological significance is currently unclear. Contrary to the recent observation of Li et al.12 we did not identify phosphorylation of Thr549.

An important finding with regard to MTHFR phosphorylation is that it does not directly alter the catalytic parameters of the enzyme, as determined by a sensitive HPLC-based activity assay. Perhaps this is not too surprising, since the first ordered residue of the structure, Glu40 (i.e. immediately following the phosphorylation region aa 1–37), is far removed from the catalytic site. Instead, MTHFR phosphorylation exerts a long-range influence on the SAM binding status at the regulatory domain some 300 amino acids away, by causing an increased sensitivity of the enzyme to SAM inhibition, but with no overall changes on total SAM inhibition. Phosphorylation likely enhances SAM sensitivity in two interdependent ways. Firstly, it enables protection of bound SAM from spontaneous degradation to SAH, a phenomenon widely observed for SAM-bound enzymes in vitro30 and in crystals31, to avoid dis-inhibition by SAH. Secondly, phosphorylation could induce a conformational change to the protein that primes an inhibition ready state. The SAM Ki differences between phosphorylated and dephosphorylated protein, while relatively small (2–3 vs. 6–7 µM), are likely to be physiologically relevant. Intracellular SAM concentrations are reported to be 1–3 µM in human cells32,33 and the mTORC1 linked starvation sensor SAMTOR, which recognizes SAM for nutrient sensing, has a SAM dissociation constant of 7 µM34.

It is not immediately clear if global phosphorylation or phosphorylation of only specific residues contributed to the results we found. While 16 phosphorylation sites of HsMTHFR are identified in this study, intact mass analysis shows that only 9–10 phosphorylations were present at a time, implying a high degree of heterogeneity. While multiply phosphorylated peptides were observed, they were not quantified using our methods. Identification of MTHFR proteoforms to assess whether phosphosites occur in combinations, perhaps using top-down MS35 would be very interesting, and may constitute the basis of future studies. Evolutionary conservation of the identified phosphorylation sites varies from absolute invariance through to yeast (e.g. Ser394), to poor conservation even among animals (e.g. Ser9 and Ser10) (Supplementary Fig. 15). Truncated recombinant HsMTHFR38–644 was not identified to be phosphorylated by mass spectrometry or crystallography, suggesting that phosphorylation at the far N-terminus primes the other phosphorylation events within the catalytic and regulatory domains. This is consistent with previous observations14–16 that removal of Thr34 results in phosphorylation or SAM binding regions) and lower eukaryotic (which do not have a phosphorylation region) counterparts. Until now, the molecular basis of how these two allosteric events modulate the catalytic machinery was entirely unknown, due to the absence of a structural context. Now, our structure-guided study has provided two major discoveries in this area: (1) identification of an extensive linker region that functionally connects SAM-binding in the regulatory domain with inhibition in the catalytic domain and (2) demonstration of the concerted effects of phosphorylation and SAM binding, individually mediated by regions more than 300 amino acids apart (Supplementary Fig. 22).

Discussion

Catalytic regulation by phosphorylation and SAM binding distinguishes human MTHFR from its bacterial (which do not have phosphorylation or SAM binding regions) and lower eukaryotic (which do not have a phosphorylation region) counterparts. Until now, the molecular basis of how these two allosteric events modulate the catalytic machinery was entirely unknown, due to the absence of a structural context. Now, our structure-guided study has provided two major discoveries in this area: (1) identification of an extensive linker region that functionally connects SAM-binding in the regulatory domain with inhibition in the catalytic domain and (2) demonstration of the concerted effects of phosphorylation and SAM binding, individually mediated by regions more than 300 amino acids apart (Supplementary Fig. 22).
concentration of deleterious disease mutations found in this fi

This is indicative of cofactor loss, in agreement with previous

with FAD enabled rescue of activity to our recombinant MTHFR.

of FAD, the essential cofactor. We observed that supplementation

unaltered? One possibility is an effect on the stability or integrity

catalytic domain, seeing that its kinetic parameters remain largely

dimensions. It remains to be determined from future studies

whether MTHFR undergoes ligand-dependent conformational changes to rearrange domain orientation for catalysis, like PAH and CBS. If this should be the case, the MTHFR linker region will likely contribute to this role, as it makes extensive contacts with both the regulatory domain (e.g. SAM binding site) and catalytic domain (e.g. helices a3 and a4). This is also consistent with the concentration of deleterious disease mutations found in this region. Additional genetic data from our lab are in accord, as patient fibroblasts homozygous for p.His354Tyr, a linker residue which contacts helix a3 in catalytic domain, exhibited a five-fold decrease in Ki for SAM.

So in what aspects could the SAM-bound signal influence the catalytic domain, seeing that its kinetic parameters remain largely unaltered? One possibility is an effect on the stability or integrity of FAD, the essential cofactor. We observed that supplementation with FAD enabled rescue of activity to our recombinant MTHFR. This is indicative of cofactor loss, in agreement with previous findings and suggestive of FAD being only loosely bound, as exemplified in chain B of our structure. Furthermore, a number of MTHFR mutations and polymorphisms are shown to affect FAD responsiveness. It is therefore possible that the inter-domain flexibility we observed, communicated by the SAM-bound signal, would alter the orientation of catalytic domains with respect to the rest of the protein, in a similar manner as the multi-domain enzymes CBS and PAH. Such structural conformations are supported by overlays between chains A and B in our structure, and between apo- and holo-subunits in TrMTHFR. In the case of MTHFR, the active site could be more sequestered (leading to FAD bound) or more exposed and mobile (leading to FAD loss) as a consequence.

The SAM/SAH ratio is regarded as an indicator of a cell’s methylation potential and is a crucial indicator of the cells’ capacity to perform DNA methylation or create compounds which require methyl groups for assimilation. In the face of a low SAM/SAH ratio, meaning methyl donor deficiency, MTHFR is dis-inhibited, increasing the production of CH3-THF to improve throughput of the methionine cycle and replenish SAM levels. Conversely, a high SAM/SAH ratio means abundant methylation capacity, in which case SAM-mediated allosteric inhibition of MTHFR turns off CH3-THF production, thereby lowering methionine cycle activity and concomitantly generation of SAM. This on/off switch is especially powerful at high SAM levels, as illustrated by almost complete inhibition of recombinant HsMTHFR at >200 µM SAM. Although these types of concentrations are unlikely to be seen inside the cell, HsMTHFR has been further outfitted with a dimer switch, whereby protein phosphorylation increases sensitivity to SAM-mediated inhibition at normal (1–10 µM) cellular SAM levels. In this regard, phosphorylation allows linkage of the methionine cycle to other cellular pathways (e.g. cell cycle) through specific kinase activities (as suggested by Zhu et al. and Li et al.)

The clear correlation we observed between phosphorylated MTHFR with SAM binding in solution (vs dephosphorylated MTHFR and SAH binding) leads us to interpret that the two regulatory properties act in concert. In fact, the architecture of the HsMTHFR homodimer is smartly tailor-made to facilitate this correlation. (1) The dimeric interface is entirely constituted by the regulatory domain to form a scaffold, while leaving each SAM binding site on a different chain for protein docking and signal transmission functions. (2) The lack of contacts between inter-monomeric catalytic domains allows for the intrinsic intramonomeric mobility with respect to the regulatory domains for signal propagation. (3) Importantly, this dimeric configuration brings the N- and C-termini of the polypeptide in proximity, projecting the phosphorylation region close to the regulatory domain dimer interface.

In summary, we provide a structural view of a eukaryotic MTHFR, pointing to the linker region playing a direct role in allosteric inhibition following SAM binding, and phosphorylation as a means to modulate SAM inhibition sensitivity. Modulating such finite control towards the level of a key metabolite may be of pharmacological interest, including in cancer metabolism. Our work here constitutes a strong starting point for future, more precise investigation by structural, biochemical and cellular studies, for example towards: identification of the kinase(s) responsible for MTHFR phosphorylation in vivo; combining different structural methods to delineate conformational changes of the entire protein, and revealing the molecular basis for its specificity over NADPH.

Methods

Recombinant production of MTHFR. Numbering of the nucleotide changes follows the nomenclature of NM_005957.4, which places the A of the ATG initiation codon as the +1 nucleotide. The protein is numbered according to NP_005948.3. For E. coli (BL21(DE3)R3) expression, DNA fragments encoding human MTHFR (IMAGE: 6374885), mouse MTHFR (IMAGE: 6834866) and yeast MET12 (clone: ScCD00096551 from Harvard Medical School) harbouring different N- and C-terminal boundaries were amplified and subcloned in pNIC28-Bsa4 vector (accession EF198106) in-frame with a tobacco etch virus protease cleavable N-terminal His6-tag. For baculovirus expression, DNA fragments encoding human MTHFR (IMAGE: 6374885) harbouring different N- and C-terminal boundaries were cloned into the pFB-CT10HF-LIC vector (Addgene plasmid: 39191) in-frame with a tobacco etch virus protease cleavable C-terminal flag/His10-tag. Site-directed mutations were constructed using the QuickChange mutagenesis kit (Stratagene) and confirmed by sequencing. All primers are available upon request. Proteins expressed in E. coli were purified by affinity (Ni-Sepharose; GE Healthcare) and size exclusion (Superdex 200; GE Healthcare) chromatography. Proteins expressed in insect cells in Sf9 media (ThermoFisher) were purified by affinity (Ni-NTA, Qiagen) and size exclusion (Superdex 200) chromatography, followed by cleavage of the C-terminal tag by His-tagged tobacco etch virus protease (/>.10 mass ratio) overnight at 4 °C and re-passaged over Ni-NTA resin. Selenomethionine (SeMet)-derivatized proteins were expressed using Seleno-Methionine Medium Complete (Molecular Dimensions) and purified as above.

Crystallization and structure determination. Purified native ScMET12,–352 as well as SeMet-derivatized and native HsMTHFR58,–244 were concentrated to 15–20 mg mL−1 and crystals were grown by sitting drop vapour diffusion at 20 °C. The protein crystallization conditions are summarized in Table 2. Crystals were cryo-protected in mother liquor containing ethylene glycol (25% v/v) and flash-cooled in liquid nitrogen. X-ray diffraction data were collected at the Diamond Light Source and
processed using XIA2 (ref. 41). The $H_{\text{MTHFR}}_{\text{ss,444}}$ structure was solved by selenium multi-wavelength anomalous diffraction phasing using autoSHARP 42 and had a nondeterminant building at 0.158 Fadero. The REMEDY model was used to solve the native structure of $H_{\text{MTHFR}}_{\text{ss,444}}$ by molecular replacement using Phaser 43. This structure was refined using PHENIX 44, followed by manual rebuilding in COOT 45. Phases for SCmET12-302 were calculated by molecular replacement using 3APY as model. Atomic coordinates and structure factors for both SCmET12-302 (PDB code: 6GUN; resolution code: 6CEx) have been deposited in the Protein Data Bank. Data collection and refinement statistics are summarized in Table 2. The final models of $H_{\text{MTHFR}}_{\text{ss,444}}$ and SCmET12-302 contain respectively 95.6% and 98.6% in the favoured region, 4.3% and 1.1%, in the allowed region, and 0.17% and 0.36% in the outlier region of the Ramachandran plot.

MTHFR assay. All enzymatic assays, including SAM inhibition and thermolability, were performed using the physiological forward assay described by Suormala et al. 46 with modifications as described by Rummel et al. 47 and Burda et al. 20.22. Younger minor adaptations were made for use with pure protein, including reducing the assay time to 7 min and the addition of boron serum albumin (BSA) to keep purified proteins stable. All purified proteins were diluted from 15–20 to 1 mg/mL in 10 mM HEPES-buffer pH 7.4, 5% glycerol and 500 mM NaCl followed by successive dilutions of 1:100 and 1:32 in 10 mM potassium phosphate, pH 6.6 plus 5 mg/mL BSA, to a final MTHFR concentration of 312.5 nM for each assay, purified proteins were diluted from 15–20 to 1 mg/mL in 10 mM HEPES-buffer pH 7.4, 5% glycerol and 500 mM NaCl followed by successive dilutions of 1:100 and 1:32 in 10 mM potassium phosphate, pH 6.6 plus 5 mg/mL BSA, to a final MTHFR concentration of 312.5 nM for each assay. The $K_m$ for NAPDH was determined by varying its concentration between 2 and 150 µM in the presence of 75 µM MethyleneTHF, Eprova AG; 200 µM saturating substrate concentrations (75 µM methyleneTHF, Eprova AG; 200 µM NAPDH; Sigma-Aldrich) with and without the addition of FAD (75 µM; Sigma-Aldrich). The $K_m$ for NADPH was determined by varying its concentration between 2.5 and 100 µM in the presence of 75 µM MethyleneTHF. All $K_m$ values were derived using a linear-nonlinear fit to a model of Michaelis–Menten kinetics with GraphPad Prism (v6.07). For SAM inhibition, purified SAM 48 was used. The $K_i$ was estimated using a plot of log(inhibitor) vs. response and a four parameter curve fit as performed by GraphPad Prism (v6.07).

**In-solution analysis.** Analytical gel filtration was performed on a Superdex 200 HiLoad 10/30 column (GE Healthcare) pre-equilibrated with 10 mM HEPES pH 7.5, 150 mM NaCl and 5% glycerol in the presence or absence of 250 µM SAH or SAM. SAXS measurements were performed at 20 °C, using an exposure time of 3 s frame −1. The temperature shifts, $\Delta T_{\text{m}}$ (ref. 1), were calculated using a Boltzmann sigmoidal fit and half-maximal effective ligand concentrations, $AC_{50}$, for each ligand was determined using GraphPad Prism (v6.07). The radius of gyration $R_g$ over the elution peaks and performed our analysis only on a selection of frames in the outlier region of the Ramachandran plot. 20

**Native mass spectrometry.** Mass spectrometry of MTHFR under native conditions was performed on the same MS instrument, and is described in detail elsewhere 21. Briefly, 50 µg of protein was desalted and exchanged into 50 mM ammonium bicarbonate pH 6.5 using three rounds of ultrafiltration against 20 mM ammonium bicarbonate pH 6.5 (phosphorylated and dephosphorylated) were performed at 0.99 Å resolution. The crystal structures of HsMTHFR and ScMET121–656 (phosphorylated and dephosphorylated) were deposited in the protein data bank (PDB) with the accession numbers 6FNU and 6DNN respectively.

**Data availability.** The crystal structures of HsMTHFR and ScMET121–656 were deposited in the protein data bank (PDB) with the identifiers 6CFX and 6FNU, respectively. The full mass spectrometry phosphorolysis dataset has been submitted to MassIVE with the accession number MSV000082179. All other data are available from the corresponding authors upon reasonable request.

References
1. Watkins, D. & Rosenblatt, D. S., in D. S. in The Online Metabolic and Molecular Bases of Inherited Disease (eds Valle, D. et al.) Ch. 155 (McGraw-Hill Medical, New York City, 2017).
2. Froese, D. S. et al. Mutation update and review of severe methylenetetrahydrofolate reductase deficiency. *Hum. Mutat.* 37, 427–438 (2016)

3. Liew, S. C. & Gupta, E. D. Methylenetetrahydrofolate reductase (MTHFR) C677T polymorphism: epidemiology, metabolism, and the associated diseases. *Eur. J. Med. Genet.* 58, 1–10 (2015)

4. Guenther, B. D. et al. The structure and properties of methylenetetrahydrofolate reductase from Thermus thermophilus HB8. *Biochim. Biophys. Acta* 133, 359–365 (1999)

5. Lee, M. N. et al. Functional role for the conformationally mobile phenylalanine 223 in the reaction of methylenetetrahydrofolate reductase from Escherichia coli. *Biochemistry* 48, 7673–7685 (2009)

6. Pejchal, R. et al. Structural perturbations in the Ala–Val polymorphism of methylenetetrahydrofolate reductase: how binding of folates may protect against inactivation. *Biochemistry* 45, 4808–4818 (2006)

7. Pejchal, R., Sargeant, R. & Ludwig, M. L. Structures of NADH and CH3-H4 tetrahydrofolate complexes with Escherichia coli methylenetetrahydrofolate reductase reveal a spartan strategy for a ping-pong reaction. *Biochemistry* 44, 11447–11457 (2005)

8. Igari, S. et al. Properties and crystal structure of methylenetetrahydrofolate reductase from Thermus thermophilus HB8. *PLoS ONE* 6, e23716 (2011)

9. Matthews, R. G., Vanoni, M. A., Hainfeld, J. F. & Wall, J. Methylenetetrahydrofolate reductase. Evidence for spatially distinct subunit domains obtained by scanning transmission electron microscopy and limited proteolysis. *J. Biol. Chem.* 259, 11647–11650 (1984)

10. Sumner, J., Jencks, D. A., Khani, S. & Matthews, R. G. Photoaffinity labeling of methylenetetrahydrofolate reductase with 8-azido-S-adenosylmethionine. *J. Biol. Chem.* 261, 7697–7700 (1986)

11. Kurzbach, C. & Stokstad, E. L. Mammalian methylenetetrahydrofolate reductase: Partial purification, properties, and inhibition by S-adenosylhomocysteine. *Biochem. Biophys. Acta* 250, 459–477 (1971)

12. Daubner, S. C. & Matthews, R. G. Purification and properties of methylenetetrahydrofolate reductase from pig liver. *J. Biol. Chem.* 257, 140–148 (1982)

13. Yamada, K., Chen, Z., Rozen, R. & Matthews, R. G. Effects of common polymorphisms on the properties of recombinant human methylenetetrahydrofolate reductase. *Proc. Natl. Acad. Sci. USA* 98, 14853–14858 (2001)

14. Yamada, K., Strahler, J. R., Andrews, P. C. & Matthews, R. G. Regulation of human methylenetetrahydrofolate reductase by phosphorylation. *Proc. Natl. Acad. Sci. USA* 102, 10454–10459 (2005)

15. Marini, N. J. et al. The prevalence of folate-remedial MTHFR enzyme variants in humans. *Proc. Natl. Acad. Sci. USA* 105, 8055–8060 (2008)

16. Zhu, B., Xiahou, Z., Zhao, H., Peng, B. & Xu, X. MTHFR promotes heterochromatin maintenance. *Biochem. Biophys. Res. Commun.* 447, 702–706 (2014)

17. Li, X. et al. Polo-like kinase 1 (PLK1)-dependent phosphorylation of methylenetetrahydrofolate reductase (MTHFR) regulates replication via histone methylation. *Cell Cycle* 16, 1933–1942 (2017)

18. Suormala, T., Ganne, G. & Fowler, B. 5,10-Methylenetetrahydrofolate reductase (MTHFR) assay in the forward direction: residual activity in MTHFR deficiency. *Clin. Chem.* 48, 855–843 (2002)

19. Matthews, R. G. Methylenetetrahydrofolate reductase from pig liver. *Methods Enzymol.* 112, 372–381 (1985)

20. Burda, P. et al. Insights into severe 5,10-methylenetetrahydrofolate reductase deficiency: molecular genetic and enzymatic characterization of 76 patients. *Hum. Mutat.* 36, 611–621 (2015)

21. Petoukhov, M. V. et al. New developments in the ATLAS program package for small-angle scattering data analysis. *J. Appl. Crystallogr.* 45, 342–350 (2012)

22. Burda, P. et al. Functional characterization of missense mutations in severe methylenetetrahydrofolate reductase deficiency using a human expression system. *J. Inherit. Metab. Dis.* 40, 297–306 (2017)

23. Sheppard, C. A., Trimmer, E. E. & Matthews, R. G. Purification and properties of NADPH-dependent 5,10-methylenetetrahydrofolate reductase (MeFt) from Escherichia coli. *J. Bacteriol.* 181, 718–725 (1999)

24. van den Heuvel, R. H. et al. Structural studies on flavin reductase PheA2 reveal binding of NAD in an unusual folded conformation and support novel mechanism of action. *J. Biol. Chem.* 279, 12860–12867 (2004)

25. Matthews, R. G. & Kaufman, S. Characterization of the dihydropterin reductase activity of pig liver methylenetetrahydrofolate reductase. *J. Biol. Chem.* 255, 6014–6017 (1980)

26. Bhattacharyya, S., Dutta, A., Dutta, D., Ghosh, A. K. & Das, A. K. Structural elucidation of the NADP(H) phosphatase activity of staphylococcal dual-specific IMPase/NADP(H) phosphatase. *Acta Crystallogr. D Struct. Biol.* 72, 281–290 (2016)
Acknowledgements

This work was supported by the Olga-Mayenfisch Stiftung (to D.S.F.), the Rare Disease Initiative Zurich, a Clinical Research Priority Program from the University of Zurich (to D. S. F. and M. R. B.) and the Swiss National Science Foundation (SNSF 31003A_156907) (to M.R.B.). The SGC is a registered charity (number 1097737) that receives funds from AbbVie, Bayer Pharma AG, Boehringer Ingelheim, Canada Foundation for Innovation, Eschelman Institute for Innovation, Genome Canada, Innovative Medicines Initiative (EU/EFPIA) [ULTRA-DD grant no. 115766], Janssen, MSD, Merck KGaA, Novartis Pharma AG, Ontario Ministry of Economic Development and Innovation, Pfizer, São Paulo Research Foundation-FAPESP, Takeda and Wellcome Trust [106169/ZZ14/Z]. We thank the Diamond Light Source for beamtime, and the staff of beamline I04-1 for assistance with data collection. We thank Merima Forny for assistance with sequencing.

Author Contributions

D.S.F., M.R.B. and W.W.Y. conceived of the study. D.S.F., J.K. and E.R. cloned, purified and crystallized MTHFR constructs. They also performed SAM binding and dephosphorylation assays. J.K., A.O. and G.A.B. performed X-ray structure determination. G.A.B. performed and analysed the SAXS measurement. T.S. and S.L. performed and analysed the activity assays. R.C. and O.B. performed and analysed the mass spectrometry assays. D.S.F. and W.W.Y. wrote the manuscript with editing and proofreading from the other co-authors.

Additional information

Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-018-04735-2.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2018