

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

Froese et al.

1 - MVNEARGNSSLNPCLEGSASSGSESSKDSSRCSTPGLDPERHERL - 45

VNEARGNSSLNPCL

EARGNSSLNPCLEGSASSGSESSK

EARGNSSLNPCLEGSASSGSE

ARGNSSLNPCLEGSASSGS

SSLNPCLEGSASSGSESSKDSSR

LNPCLEGSASSGSESSKDSS

PCLEGSASSGSESSK

CLEGSASSGSESSKD

LEGSASSGSESSKDSSRCS

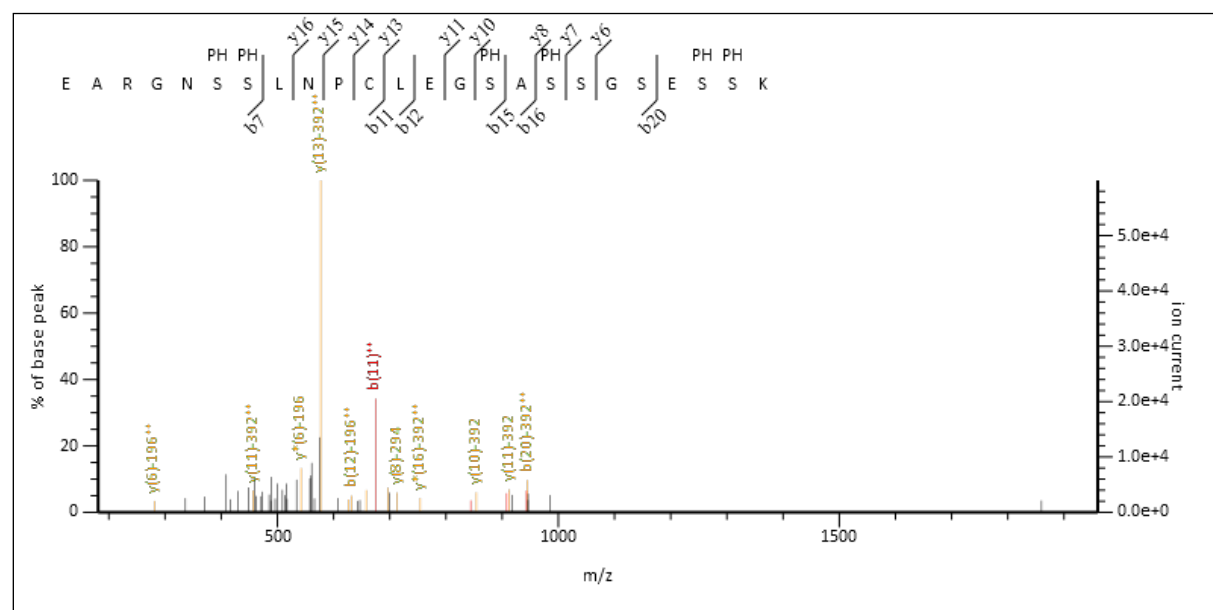
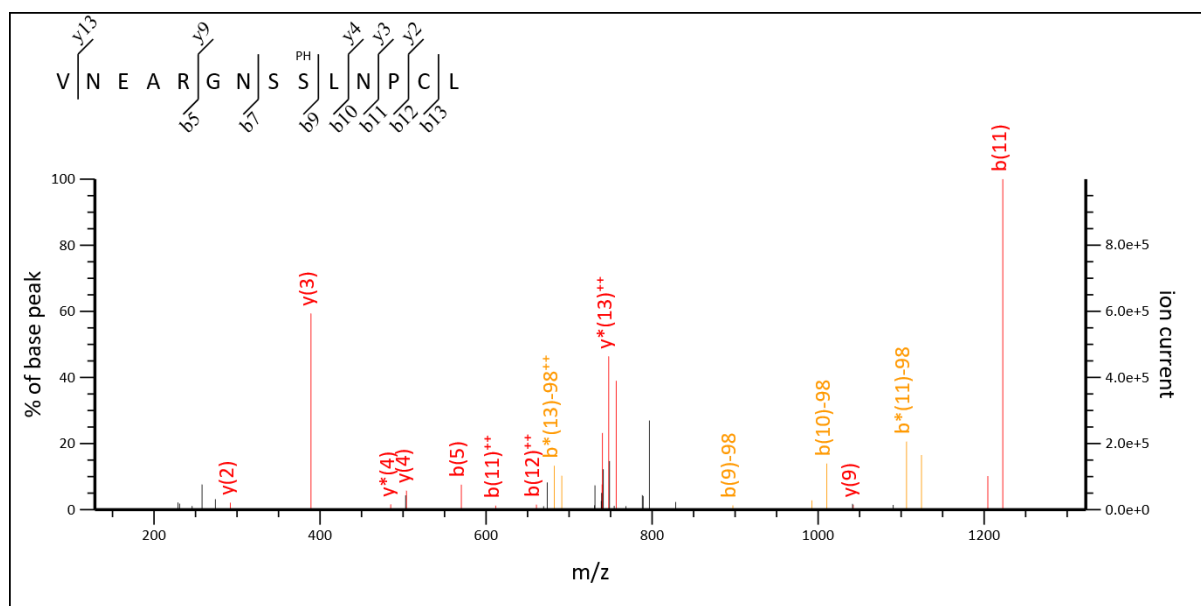
LEGSASSGSES

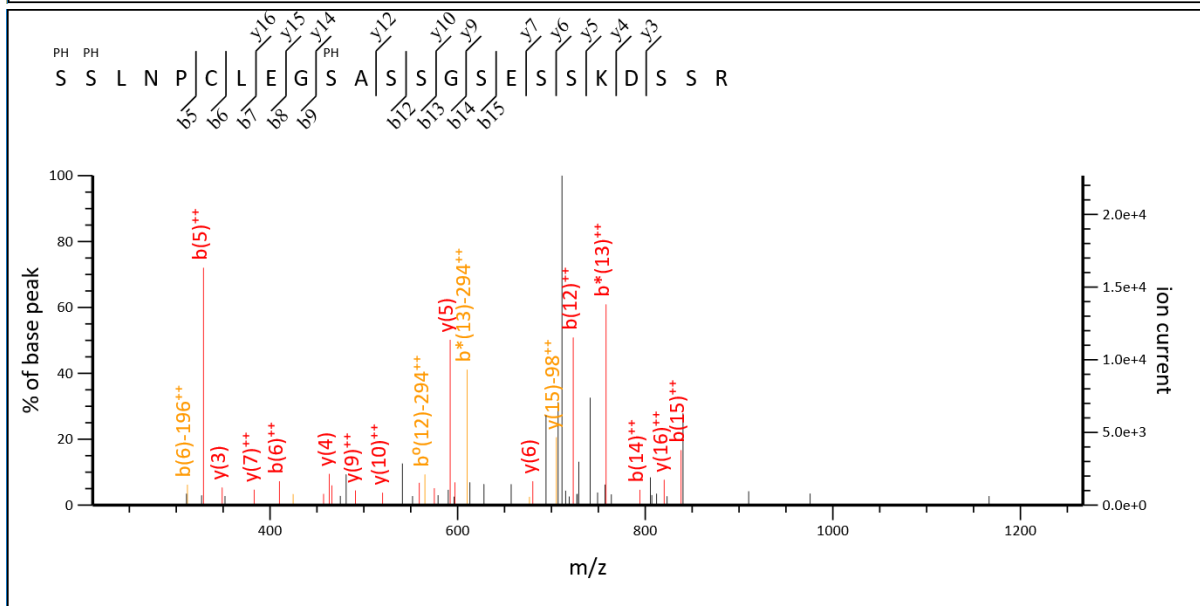
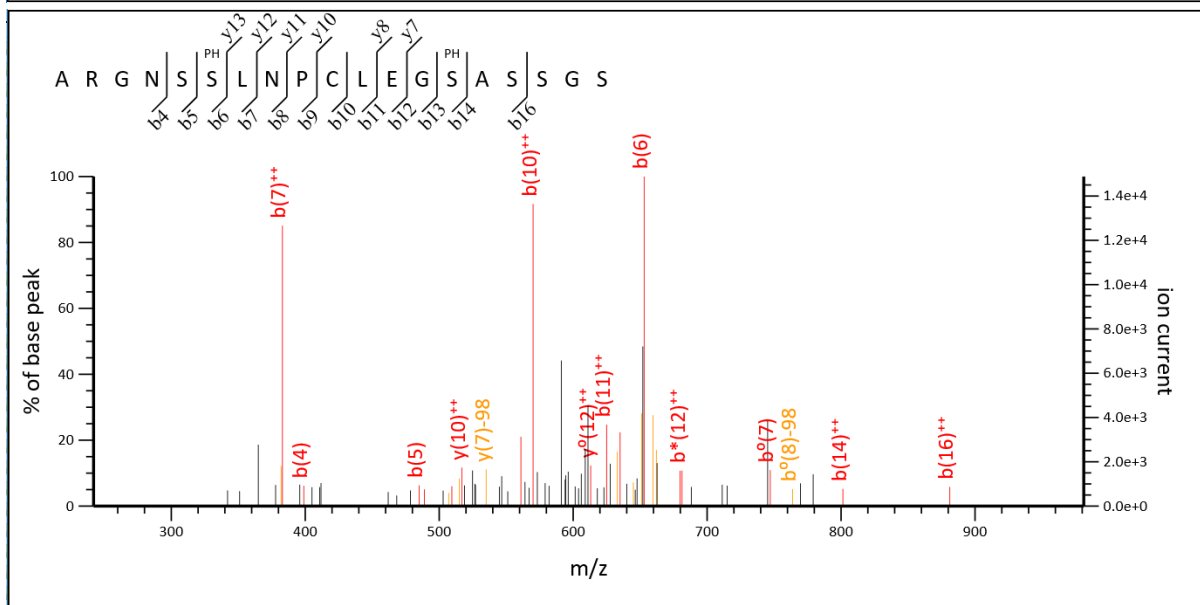
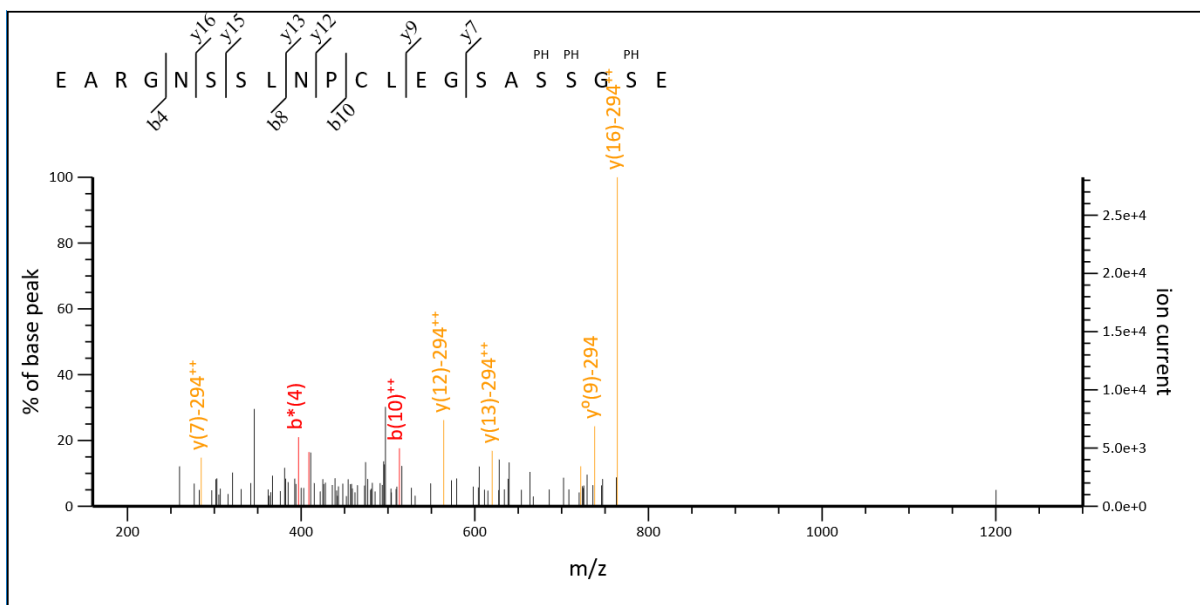
LEGSASSGSE

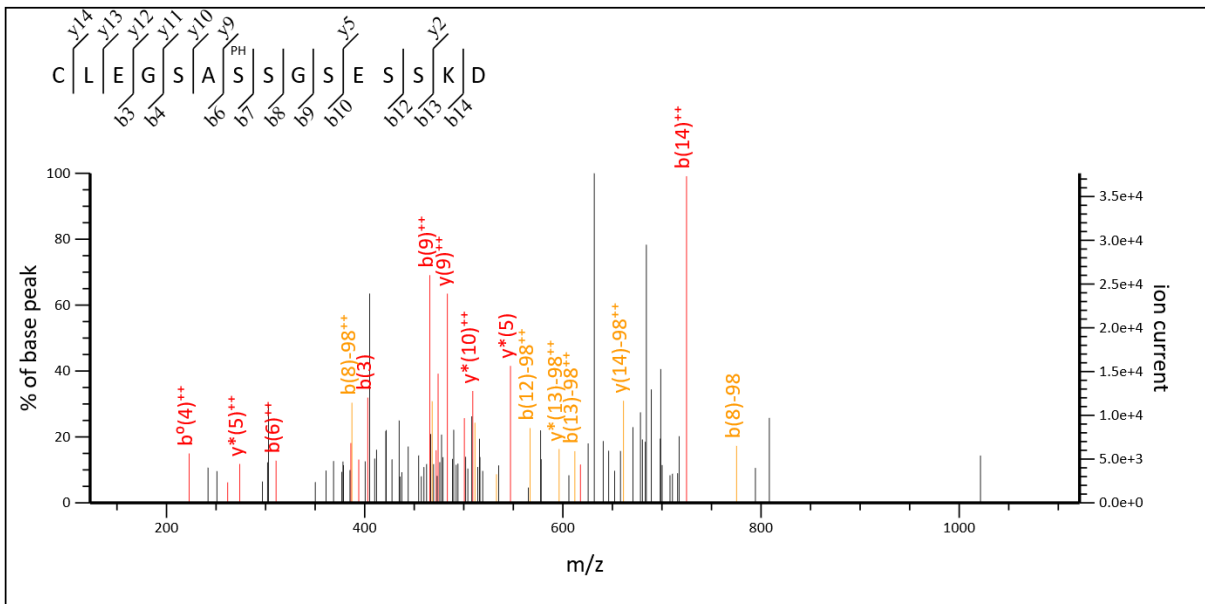
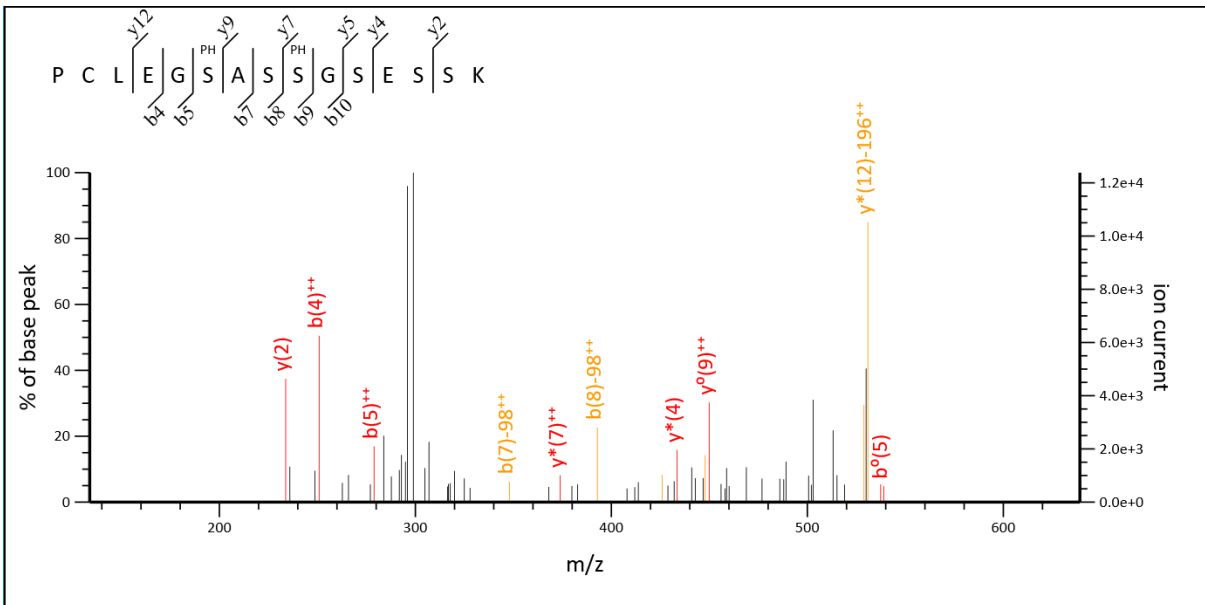
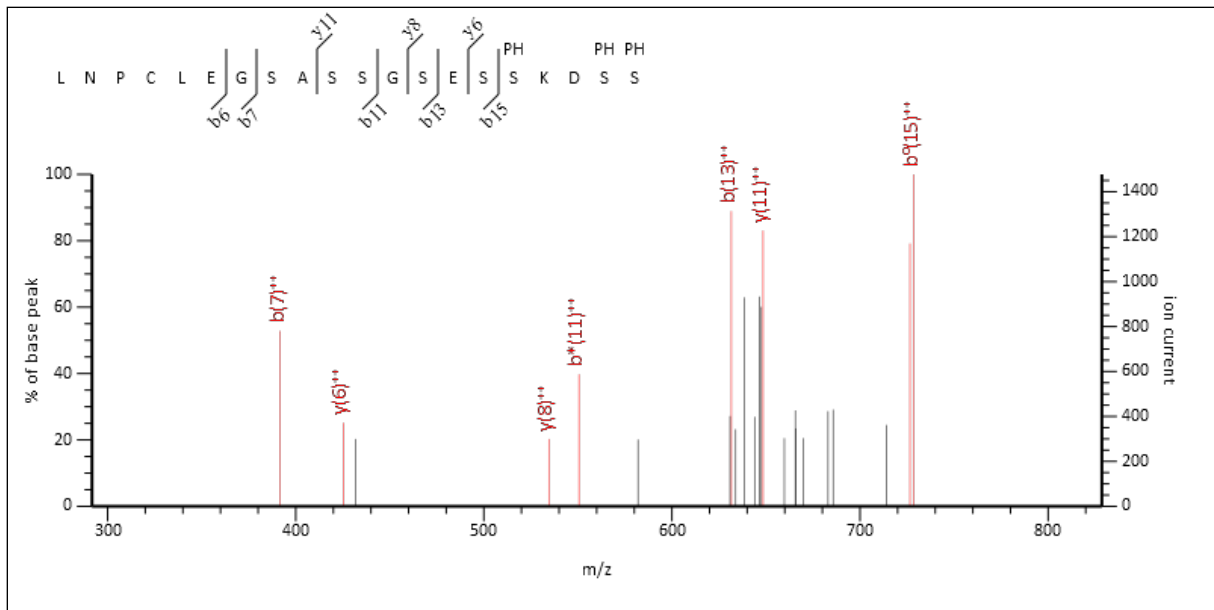
ASSGSESSKD

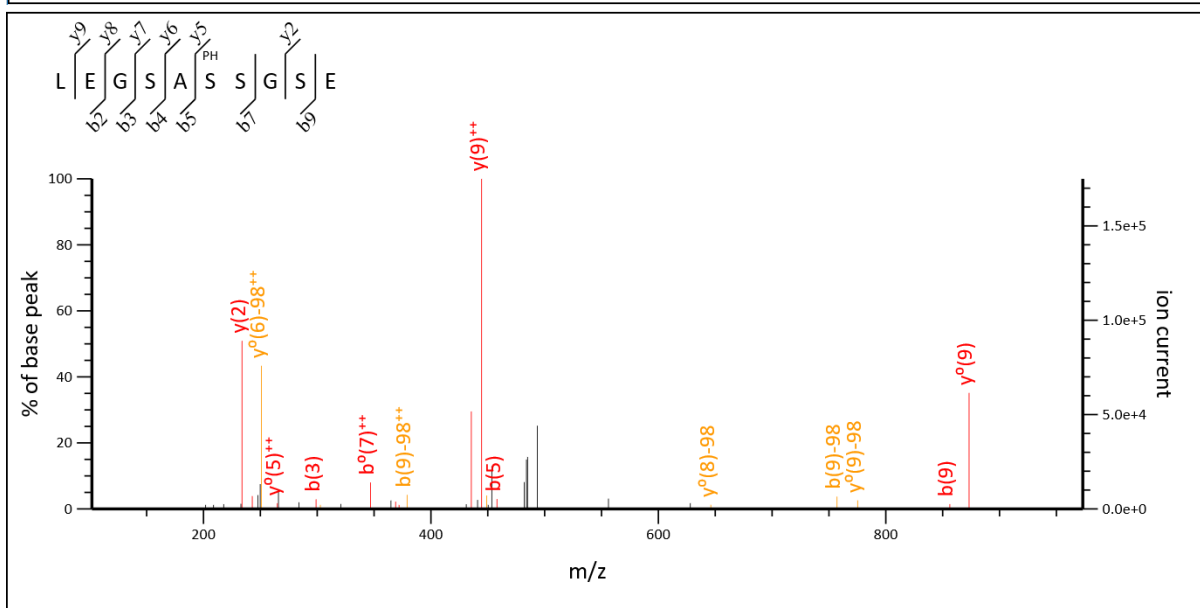
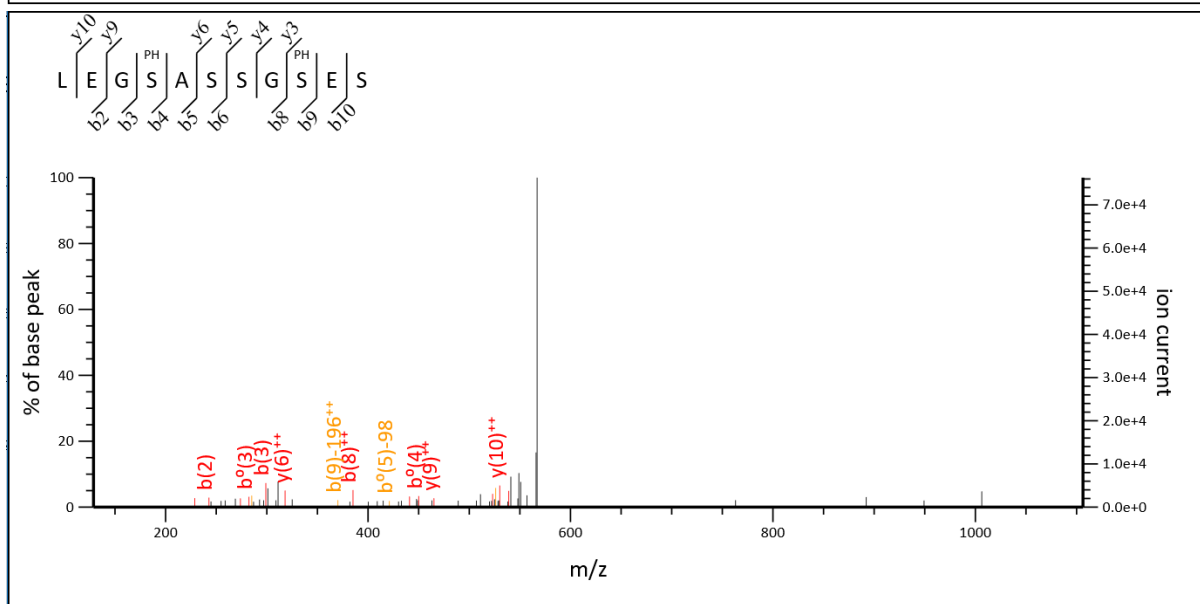
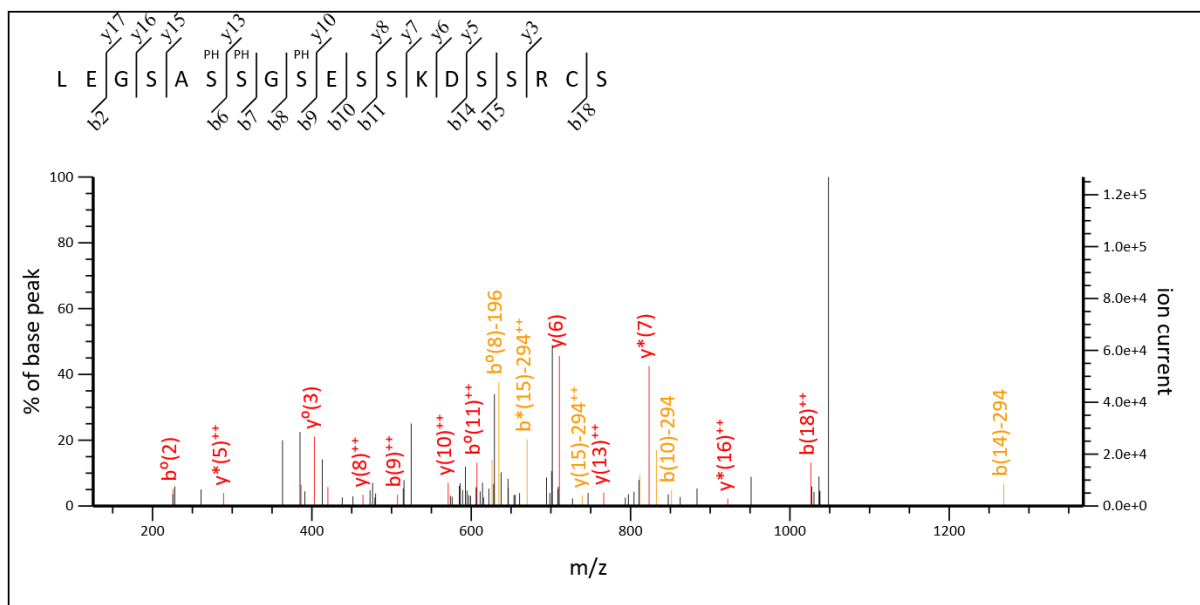
ASSGSESSK

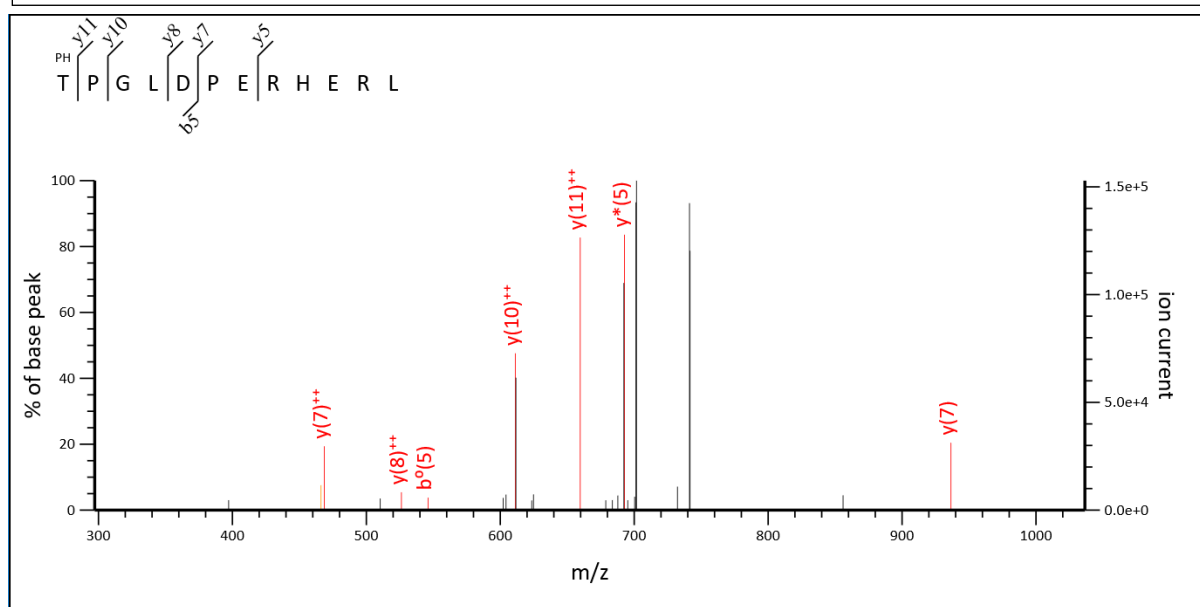
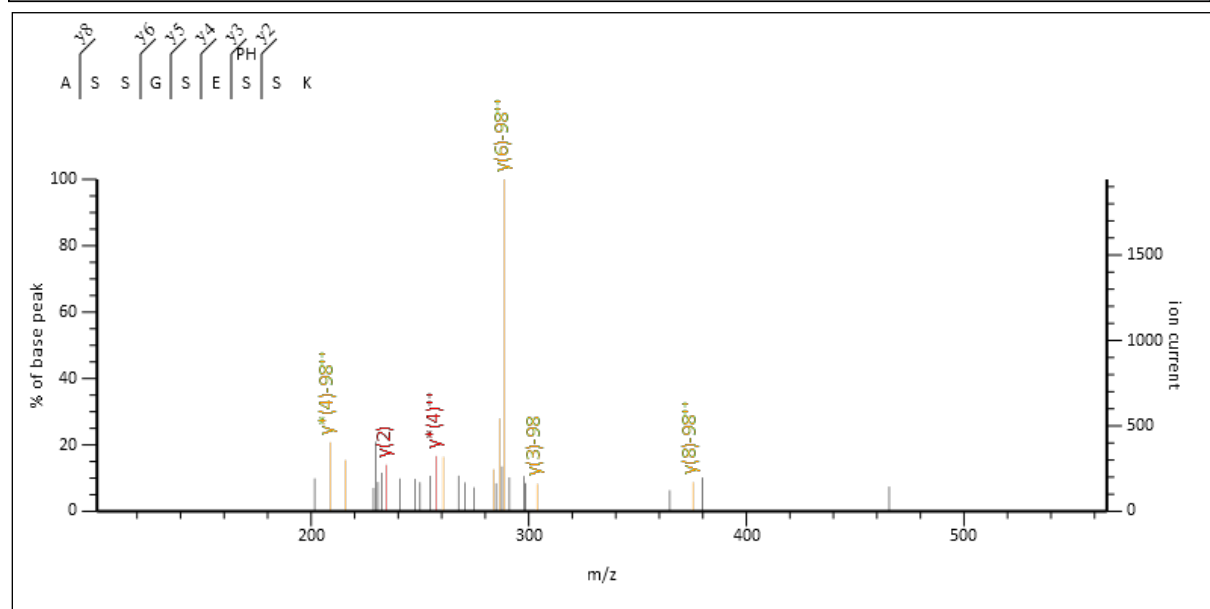
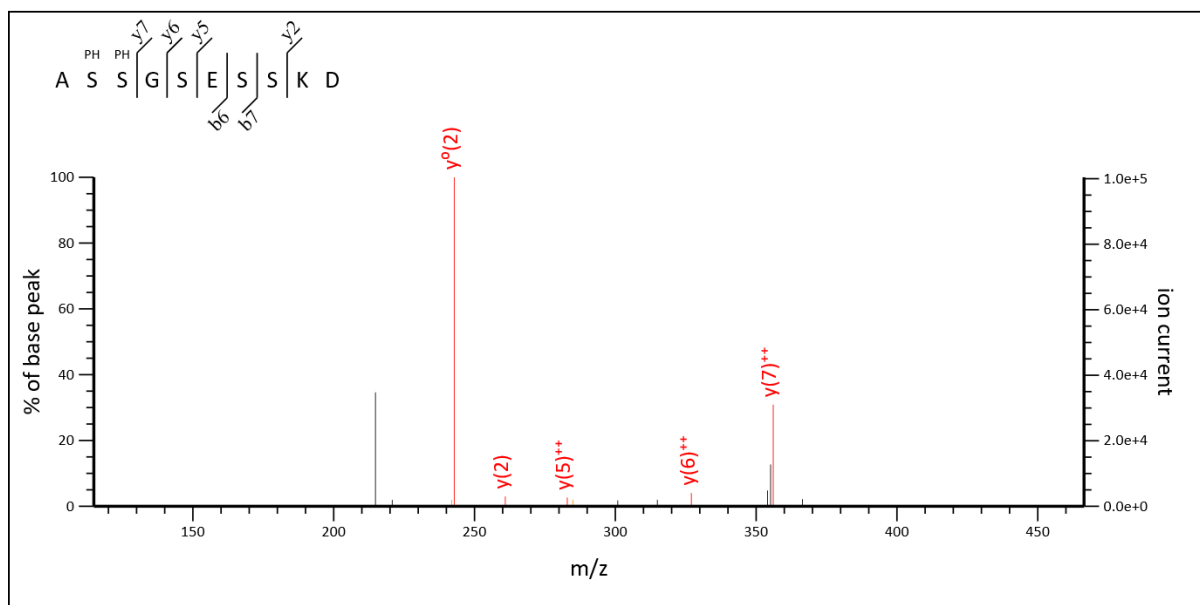
TPGLDPERHERL







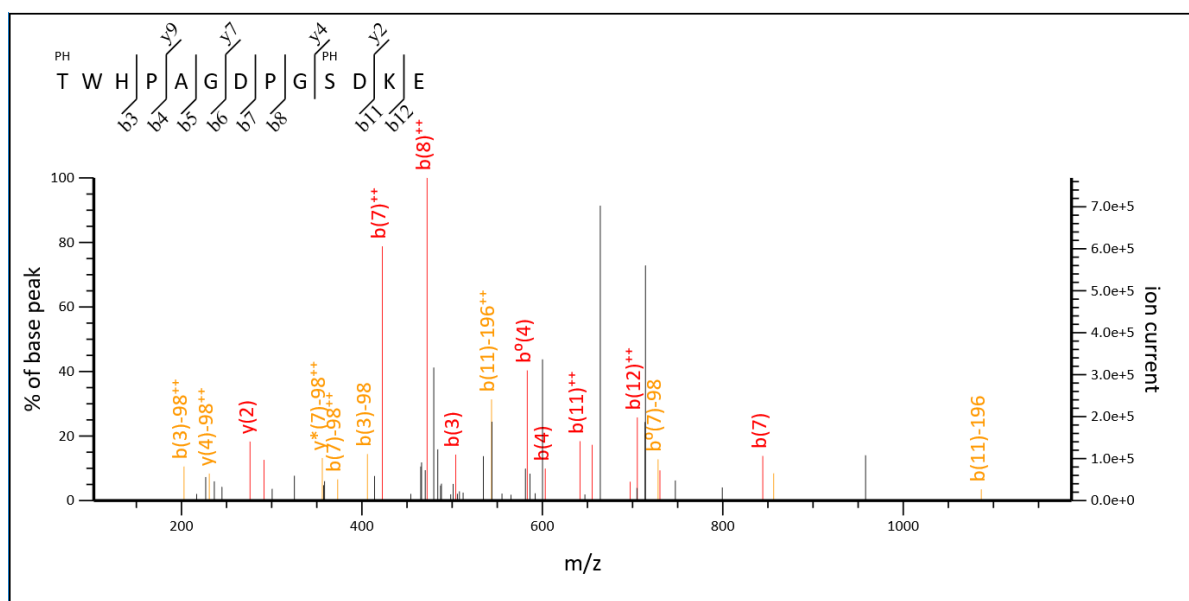
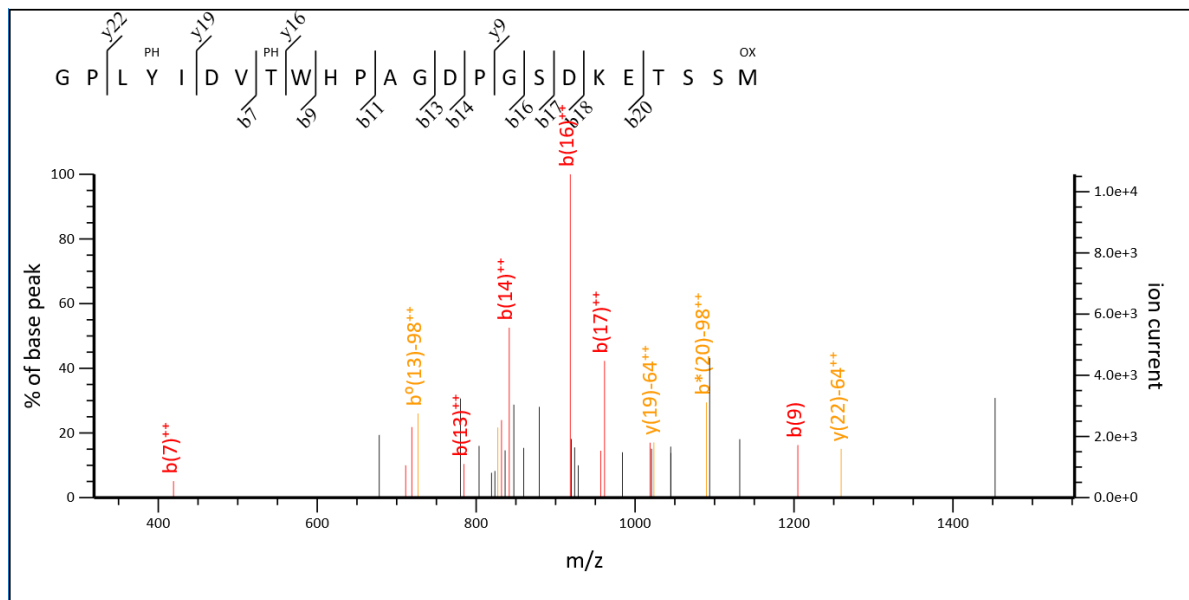




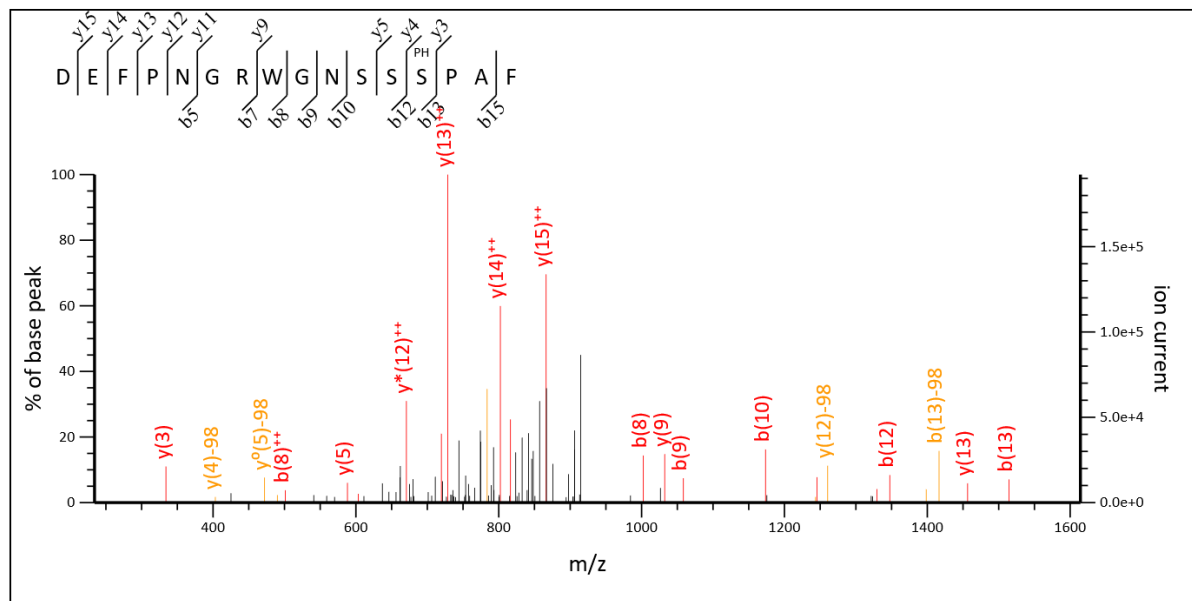
87 - GPLYIDVTWHPAGDPGSDKETSSM - 110

GPLYIDVTWHPAGDPGSDKETSSM

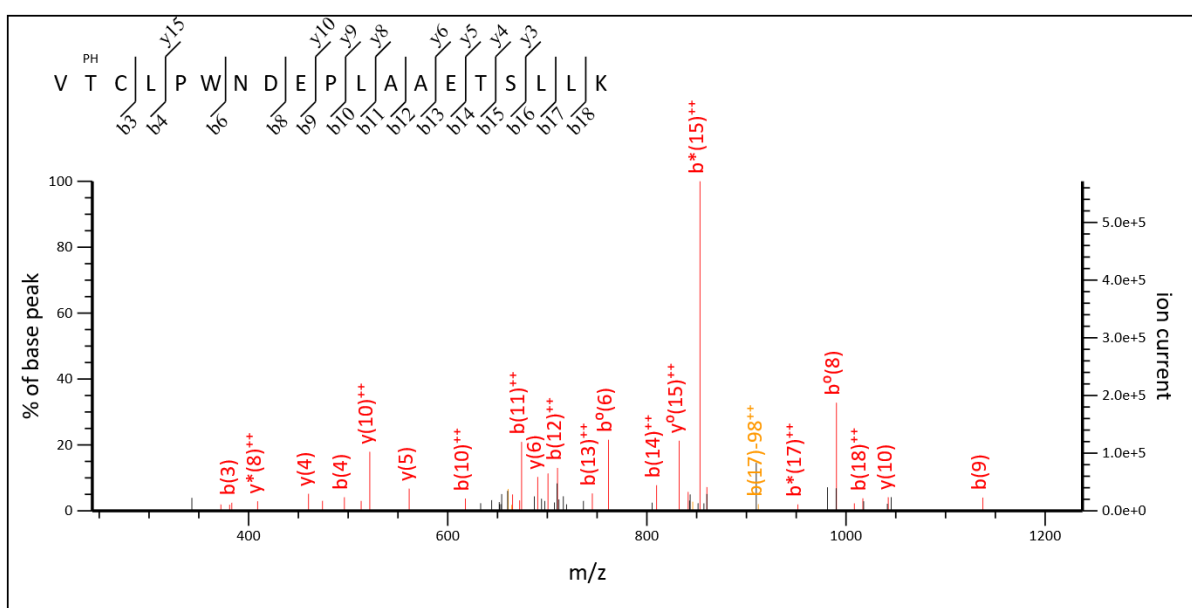
TWHPAGDPGSDKE



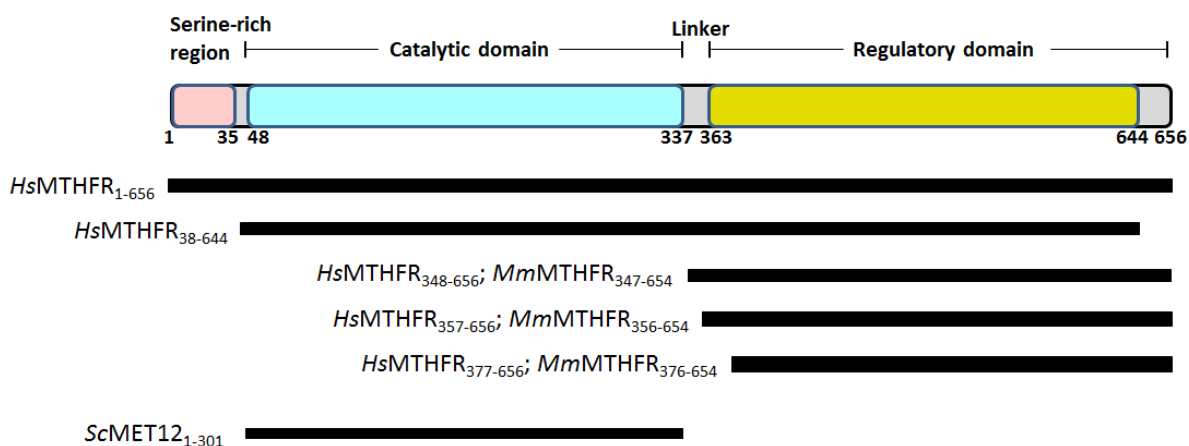
### 382 - DEFPNGRWGNSSPAF - 397



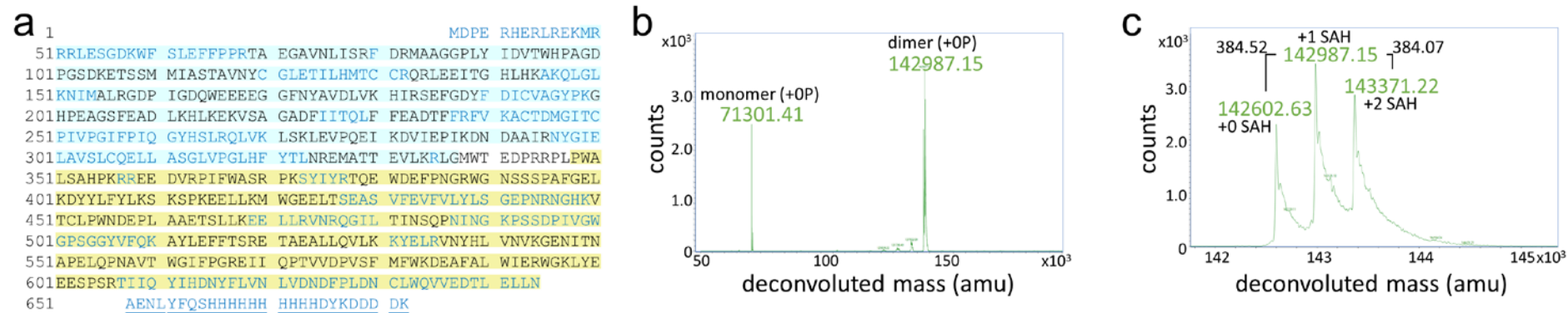
### 450 - VTCLPWNDPLAETSLLK - 468



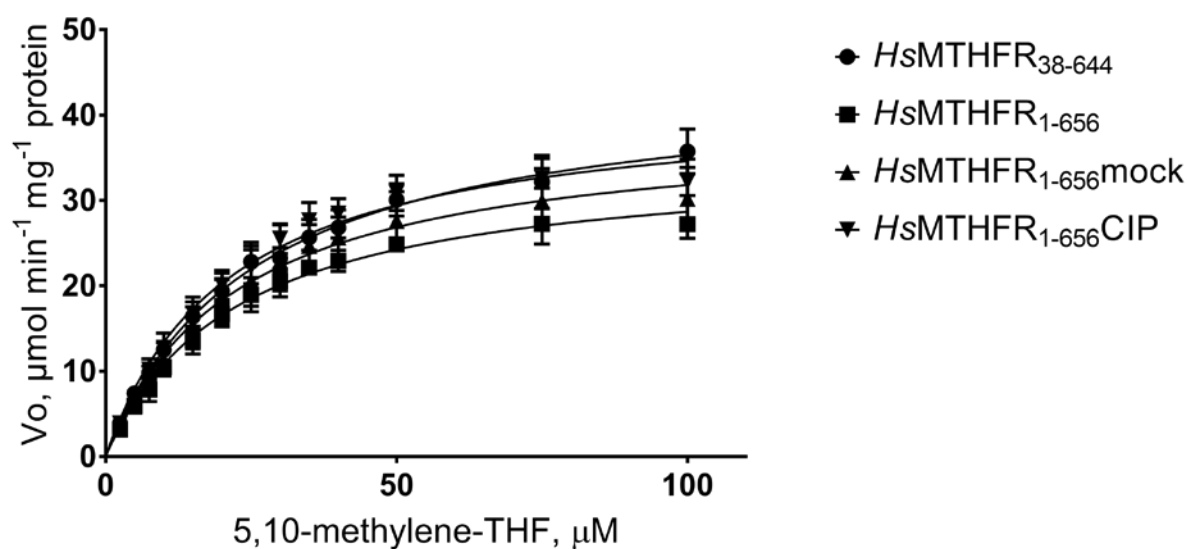
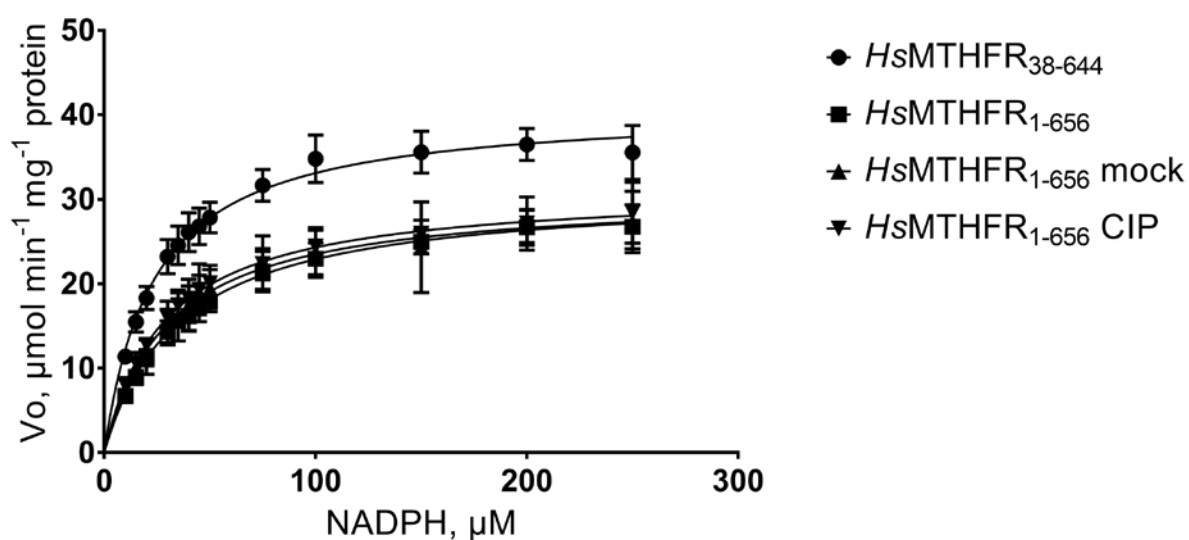
**Supplementary Figure 1. Phosphomapping of *HsMTHFR*<sub>1-656</sub>.** Annotated fragmentation spectra for 18 *HsMTHFR*<sub>1-656</sub> phosphopeptides from Mascot database searches including inset schematic showing peptide cleavages, b and y reporter ions and Mascot phosphosite assignments.



**Supplementary Figure 2. Domain organization of MTHFR and constructs used.** Domain organization of human MTHFR across evolution revealing differences in the N- and C-termini. Numbers given represent approximate amino acid boundaries in human MTHFR corresponding to NP\_005948. Constructs with the prefix *HsMTHFR* represent recombinant human protein, with the prefix *MmMTHFR* represent recombinant mouse protein, and ScMET12 represent the *Saccharomyces cerevisiae* MTHFR ortholog MET12.

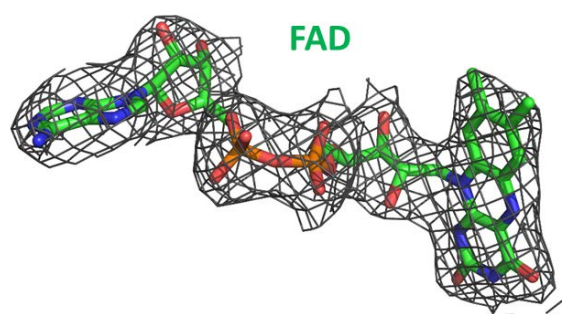


**Supplementary Figure 3. Phosphorylation status of *HsMTHFR*<sub>38-644</sub> and degradation of SAM to SAH following dephosphorylation.** **a** Phosphorylation mapping of *HsMTHFR*<sub>38-644</sub>. The protein sequence is given as amino acids in single letter code, including the C-terminal His/flag-tag (underlined). Black font represents amino acids identified by the mass spectrometer (covered), blue font represents amino acids not identified (non-covered). Domains are coloured as in Figure 1. **b** Native mass spectrometry of as purified *HsMTHFR*<sub>38-644</sub>. Broad view including monomeric and dimeric forms of the protein. Monomer represents protein bound to 1 FAD (expected mass: 71299.56 amu); dimer represents protein bound to 2 FADs and 1 SAH (expected mass: 142983.34 amu). **c** Native mass spectrometry of as purified *HsMTHFR*<sub>38-644</sub>. Zoom in on peaks corresponding to sequential addition of SAH. Expected sizes: protein, 70513.90 amu; FAD, 785.56 amu; SAH, 384.42 amu. amu: atomic mass units

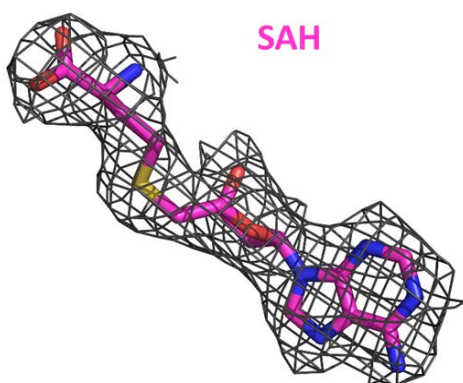
**a****b**

**Supplementary Figure 4. Kinetic characterization of *HsMTHFR*.** **a** Michaelis-Menten kinetics of *HsMTHFR* incubated with increasing concentrations of CH<sub>2</sub>-THF. **b** Michaelis-Menten kinetics of *HsMTHFR* incubated with increasing concentrations of NADPH. Both assays were performed as described in the Methods and summary data are provided in Table 1. Each value represents the results of at least 3 separate experiments and is given as  $\pm$  S.D. Lines represent non-linear fit of Michaelis-Menten kinetics as calculated by GraphPad v.6.0.

**a**

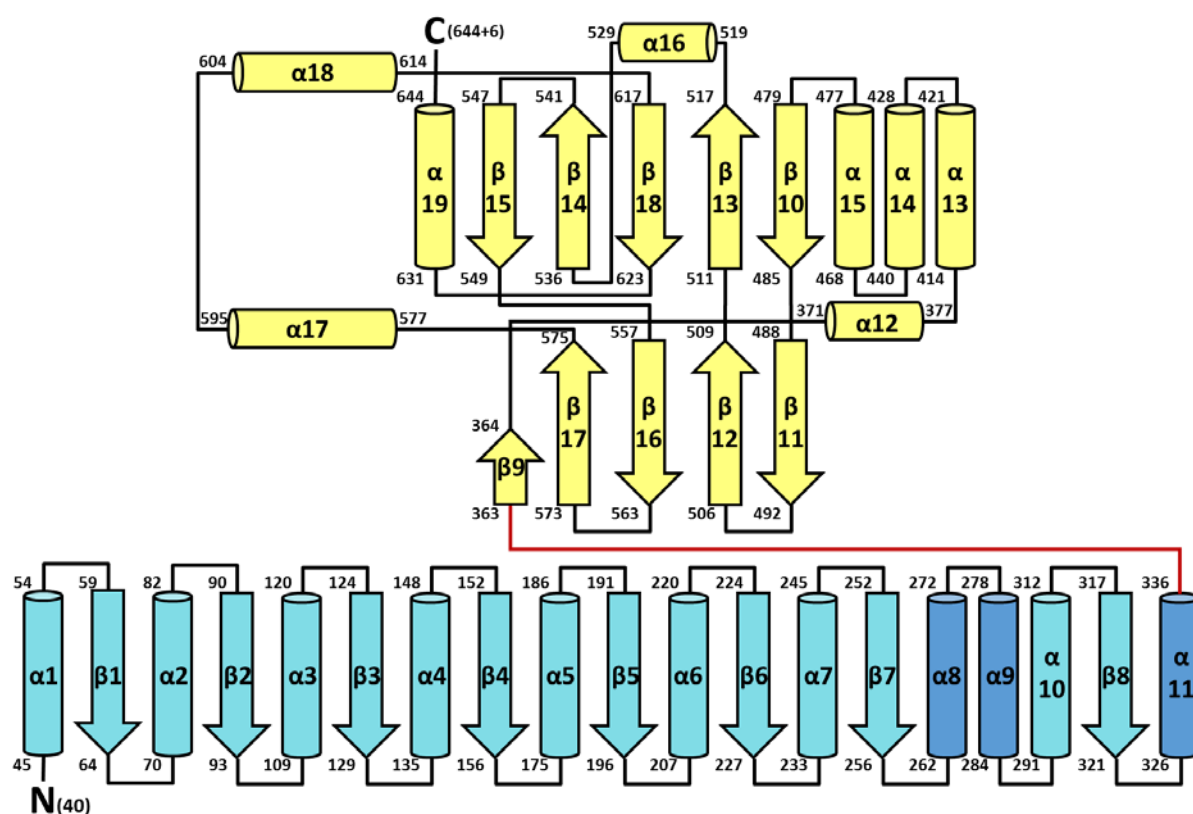


**b**

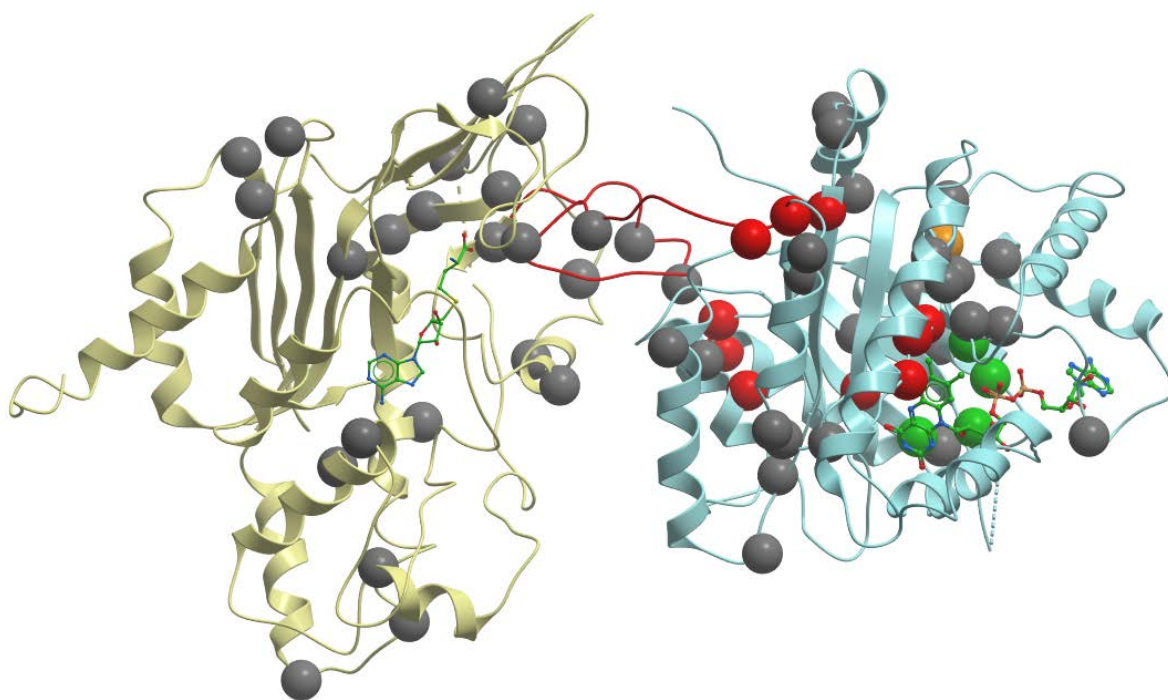


**Supplementary Figure 5. Identification of bound ligands in the *HsMTHFR*<sub>38-644</sub> crystal structure.**

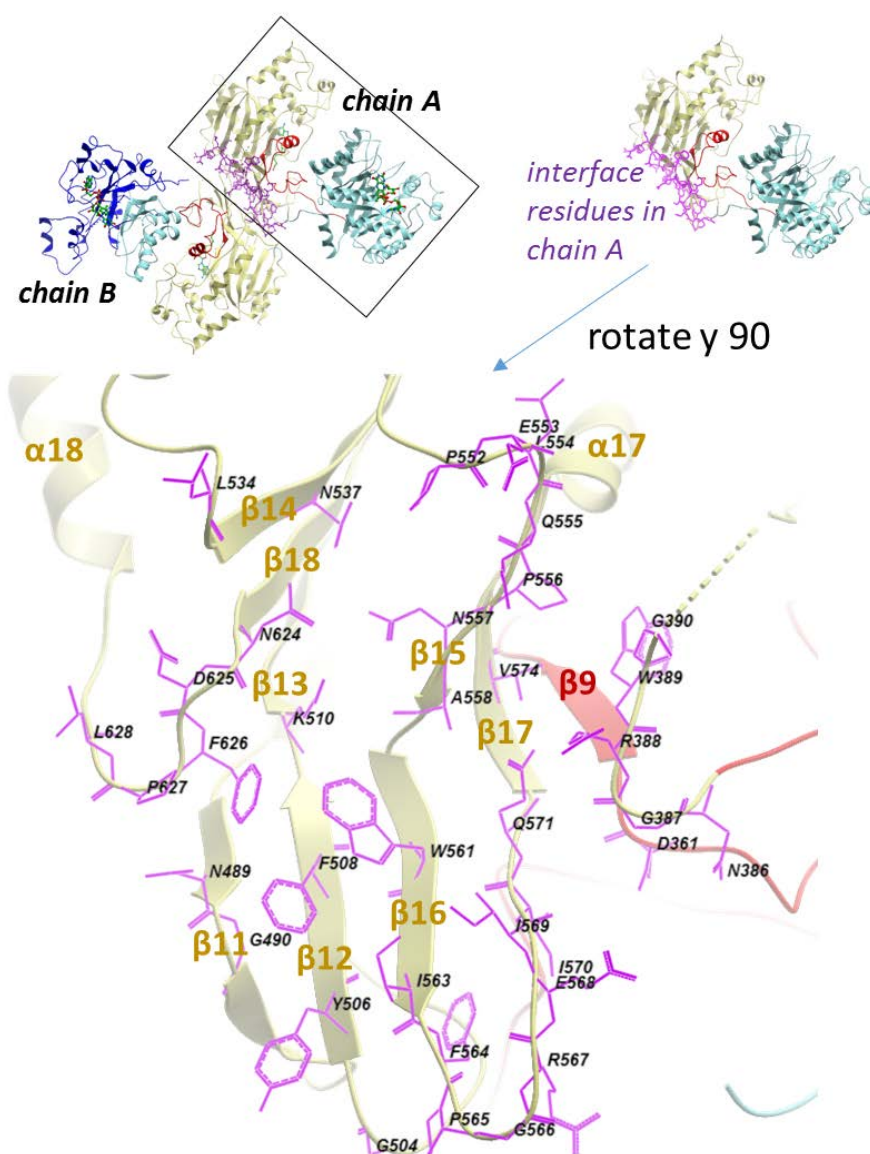
**a** 2Fo-Fc, 1.5 SigmaA-weighted map of the electron density map for FAD in the catalytic domain. **b** 2Fo-Fc, 1.5 SigmaA-weighted map of the electron density map for SAH in the regulatory domain.



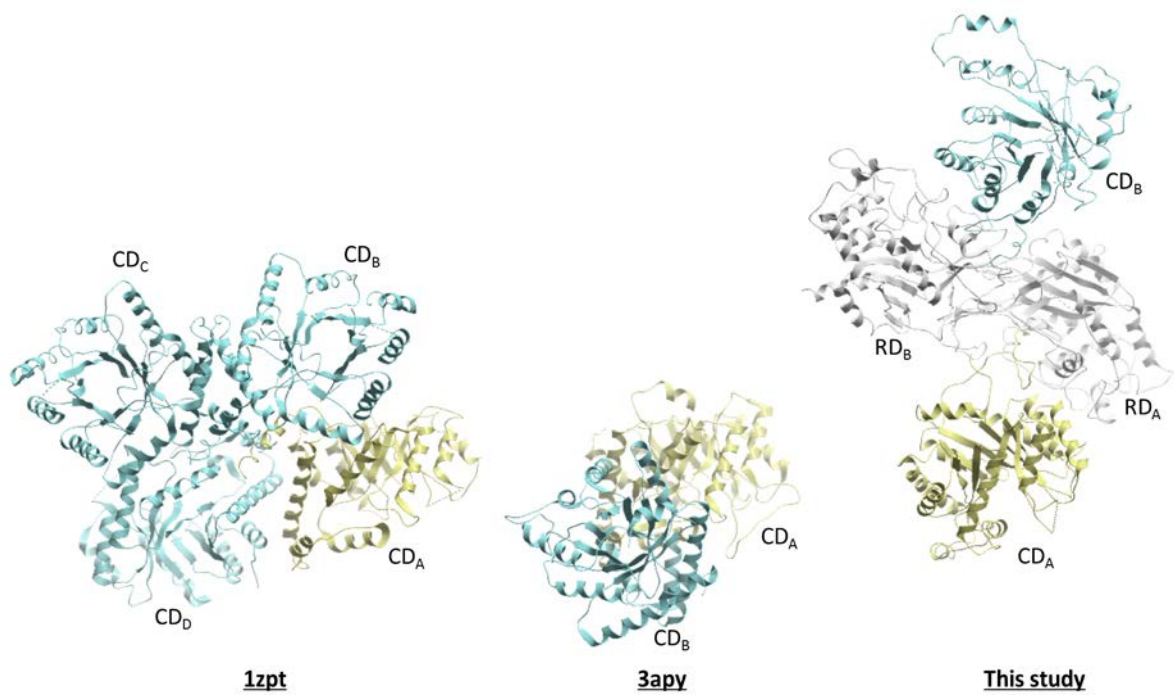
**Supplementary Figure 6. Topology diagram of HsMTHFR.** Secondary structure is shown with the same colouring and labeling as in Figure 4, with the catalytic domain in cyan, the linker in red, and the regulatory domain in yellow.  $\alpha$ -helices (cylinders) are shown in a slightly different shade as  $\beta$ -sheets (arrows).  $\alpha 8$ ,  $\alpha 9$  and  $\alpha 11$  are coloured a darker blue to indicate that they are not part of the  $(\alpha/\beta)_8$  barrel. Numbers are given to indicate amino acids at the N-terminus (40) and C-terminus (644+6 – including 6 residues incorporated from the vector) of the structure, as well as the beginning and end of each secondary structure element.



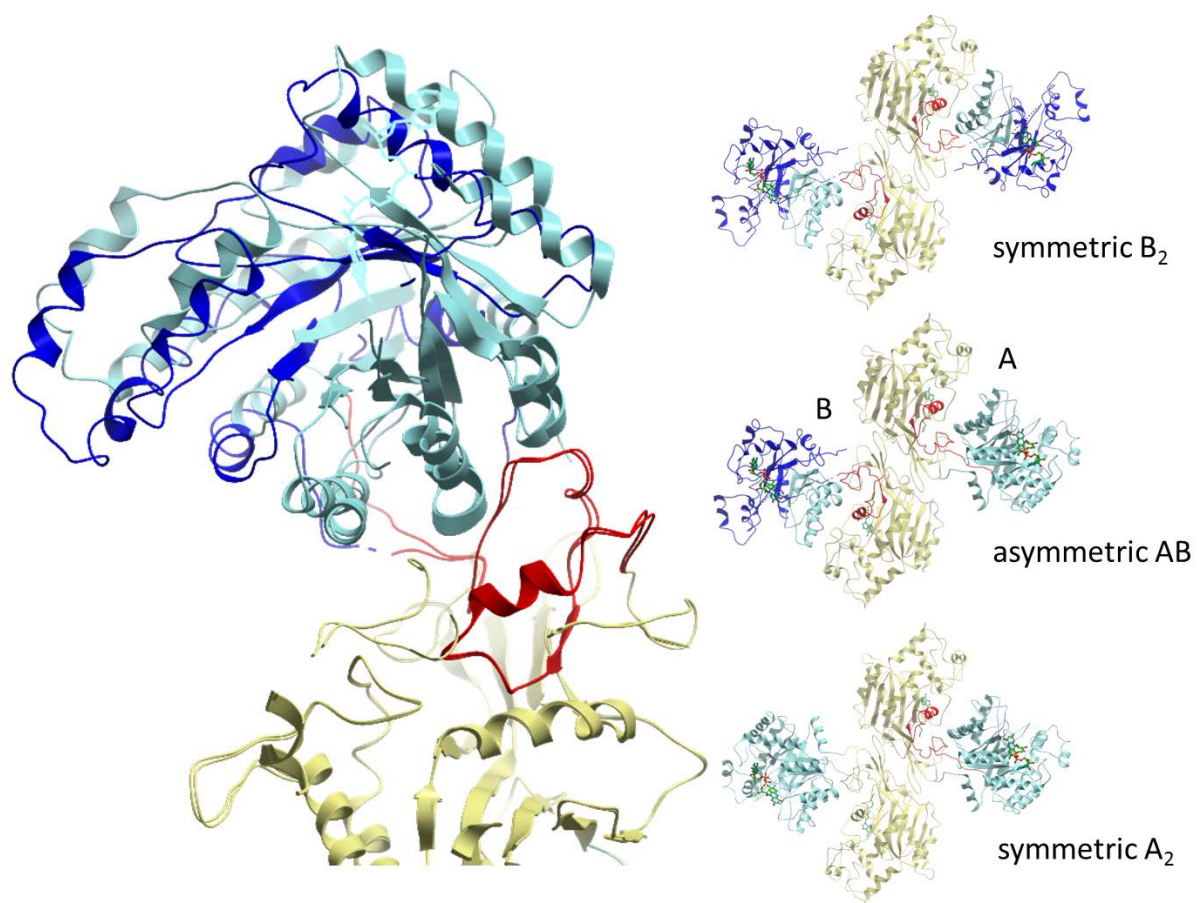
**Supplementary Figure 7. Structural analysis of MTHFR mutations.** The position of all known disease causing missense mutations represented as spheres mapped onto the *HsMTHFR*<sub>38-644</sub> structure. Green spheres represent residues involved in FAD-binding, red spheres represent mutations that result in massively decreased enzymatic activity, as defined by <sup>1</sup>, and grey spheres represent all other mutations. The orange sphere indicates the position of Ala222 found in the common variation Ala222Val.



**Supplementary Figure 8. Residues contributing to the dimeric interface in the *HsMTHFR*<sub>38-644</sub> crystal structure.** Close-up view of the dimeric interface and contributing residues (shown in purple sticks). Secondary structure is colored as in Figure 4.

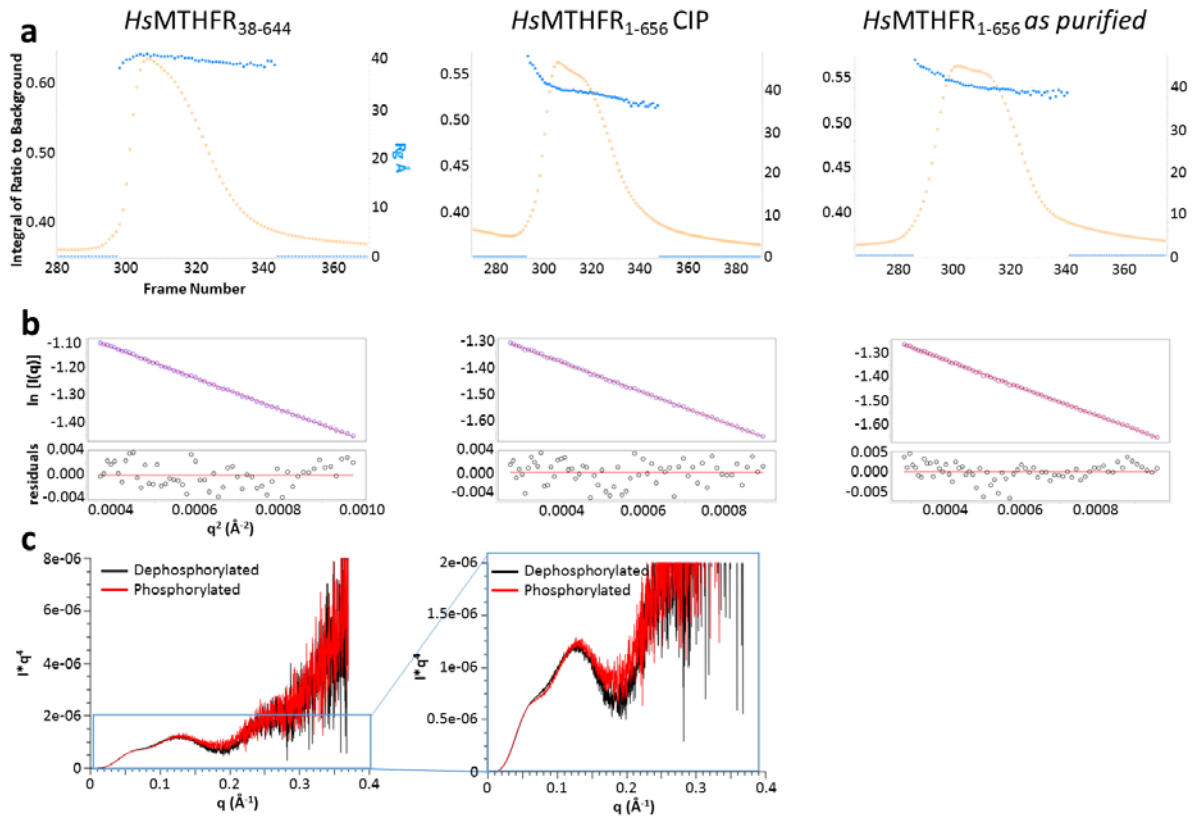


**Supplementary Figure 9. Comparison of oligomeric interfaces in MTHFR orthologs.** MTHFR from *E. coli* (PDB ID: 1ZPT) and *Thermus thermophilus* (PDB ID: 3APY) in their higher-order oligomeric states were used to compare against *HsMTHFR*<sub>38-644</sub> (this study). For each, the position of one subunit of the catalytic domain was aligned and kept in a similar orientation (CD<sub>A</sub>, shown in yellow), while other catalytic domains are shown in cyan and regulatory domains in grey. *E. coli* MTHFR forms a tetramer of four catalytic subunits. *Thermus thermophilus* forms a dimer, but also only formed of catalytic subunits. By contrast, the catalytic domains do not participate in dimerization of *HsMTHFR*<sub>38-644</sub>.

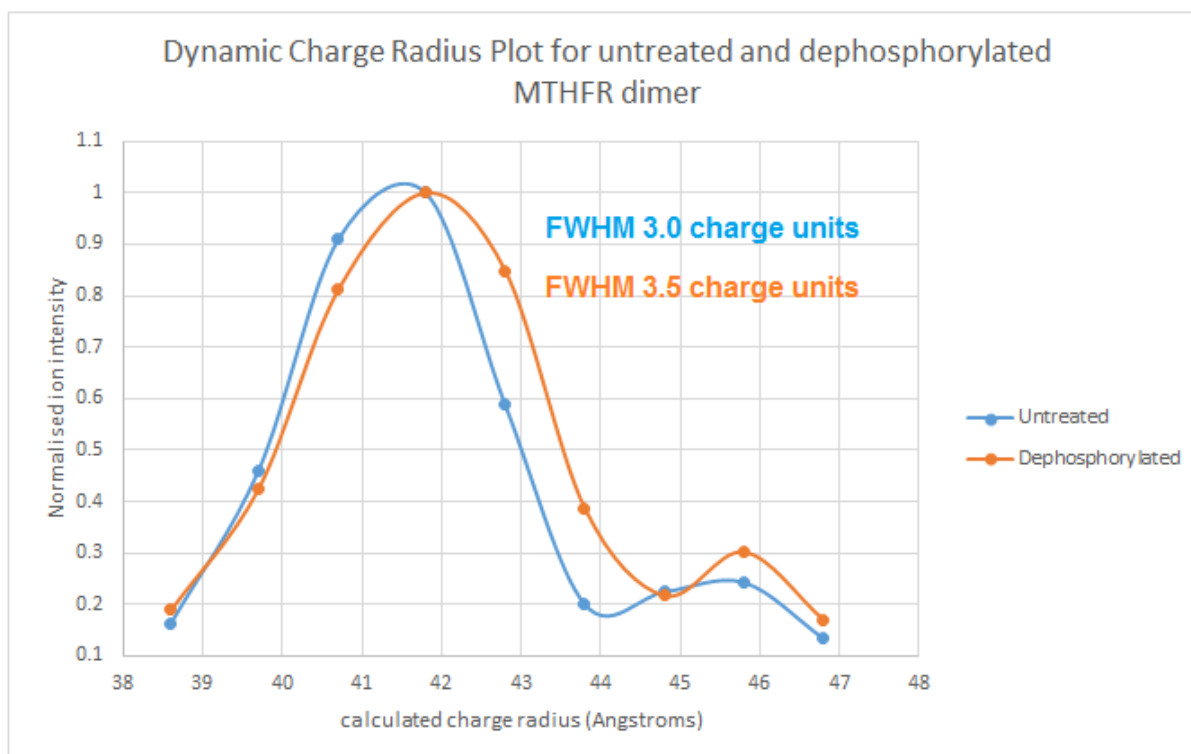


**Supplementary Figure 10. Crystal structure of *HsMTHFR*<sub>38-644</sub> depicts an asymmetric dimer.**

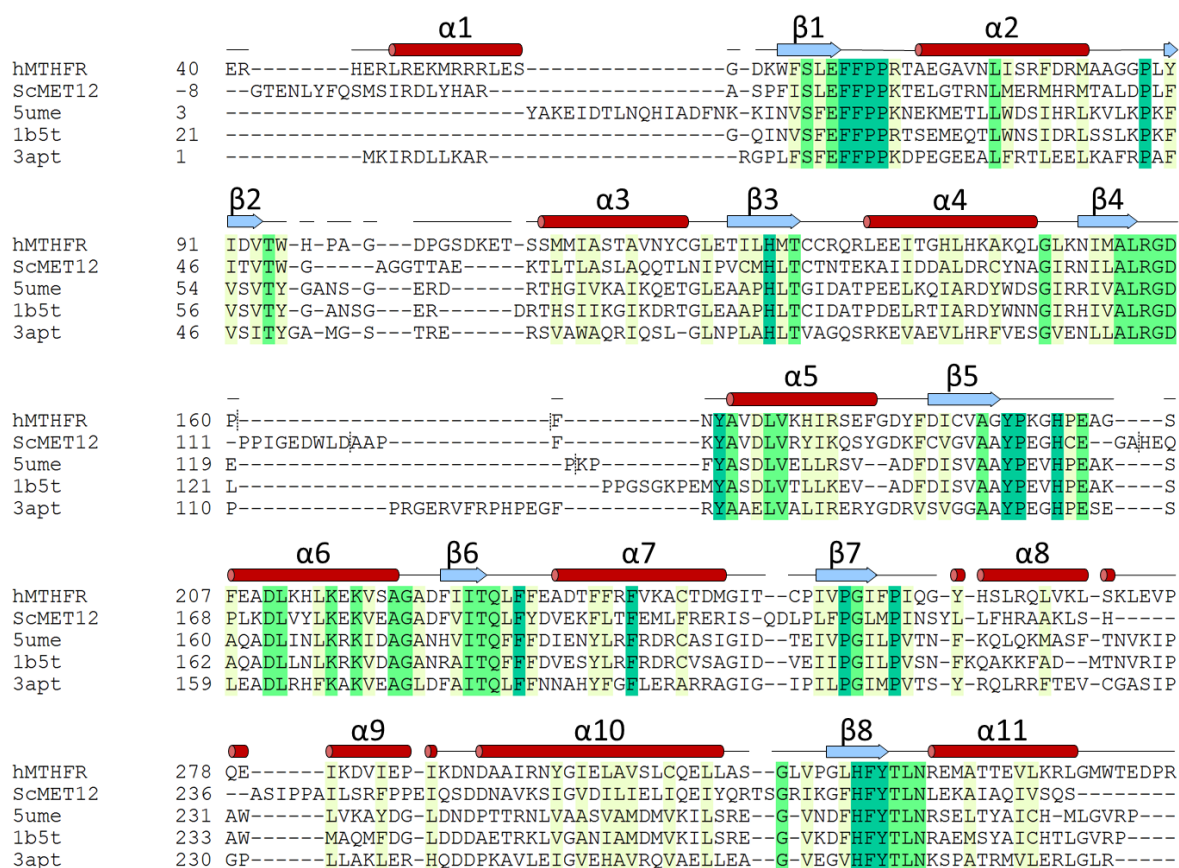
**Main figure:** Overlay and alignment of the regulatory domain (in yellow) to visualize differences in the position of the extended linker region (red) and the ordered (cyan) and disordered (dark blue) catalytic domain of subunits A and B found in the *HsMTHFR*<sub>38-644</sub> structure. FAD is shown in sticks. **Side figures:** Example of the theoretical structures of a symmetrical dimer composed only of subunit B (top) or subunit A (bottom) compared with the experimentally found asymmetric dimer (middle).



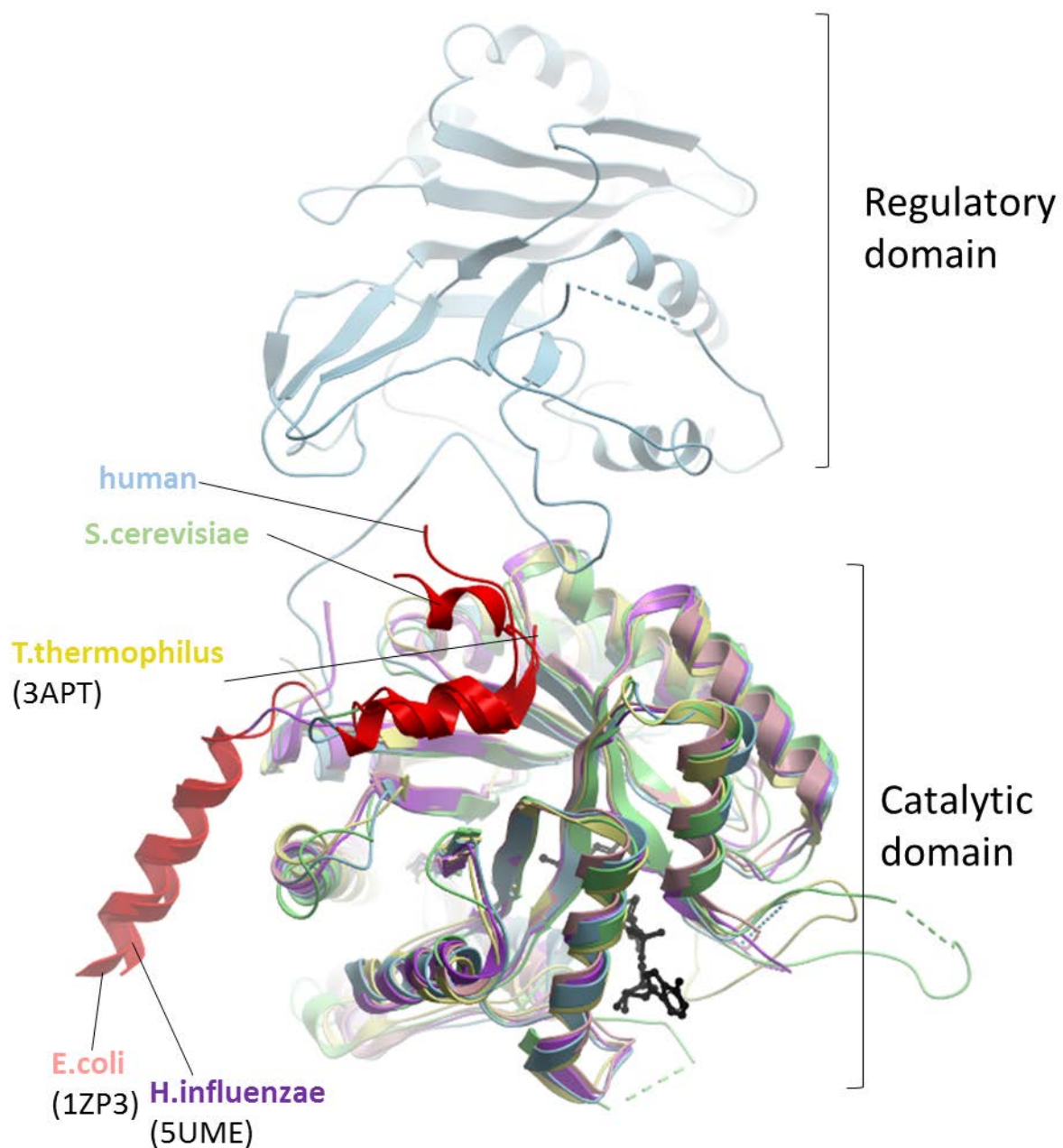
**Supplementary Figure 11. SAXS analysis of *HsMTHFR*.** **a** Analysis of the radius of gyration (R<sub>g</sub>) over elution peak in SEC-SAXS. The R<sub>g</sub> (blue spheres) and Integral of Ratio to Background (yellow spheres) are plotted versus recorded frames in SEC-SAXS profiles for either *HsMTHFR*<sub>38-644</sub>, *HsMTHFR*<sub>1-656</sub> CIP-treated or *HsMTHFR*<sub>1-656</sub> as purified. The frames used for further analysis are, respectively: 305-315, 305-320 and 300-315. Exposure of 3 seconds per frame. **b** R<sub>g</sub> (Å) values derived from Guinier are: 40.9 for *HsMTHFR*<sub>38-644</sub>, 40.8 for *HsMTHFR*<sub>1-656</sub> CIP-treated and 41.3 for *HsMTHFR*<sub>1-656</sub> as purified. **c** Porod plots of the dephosphorylated/CIP-treated (black) and phosphorylated/as purified (red) data reveal differences in the region around  $q=0.1$  Å<sup>-1</sup> and the region around  $q=0.2$  Å<sup>-1</sup>.



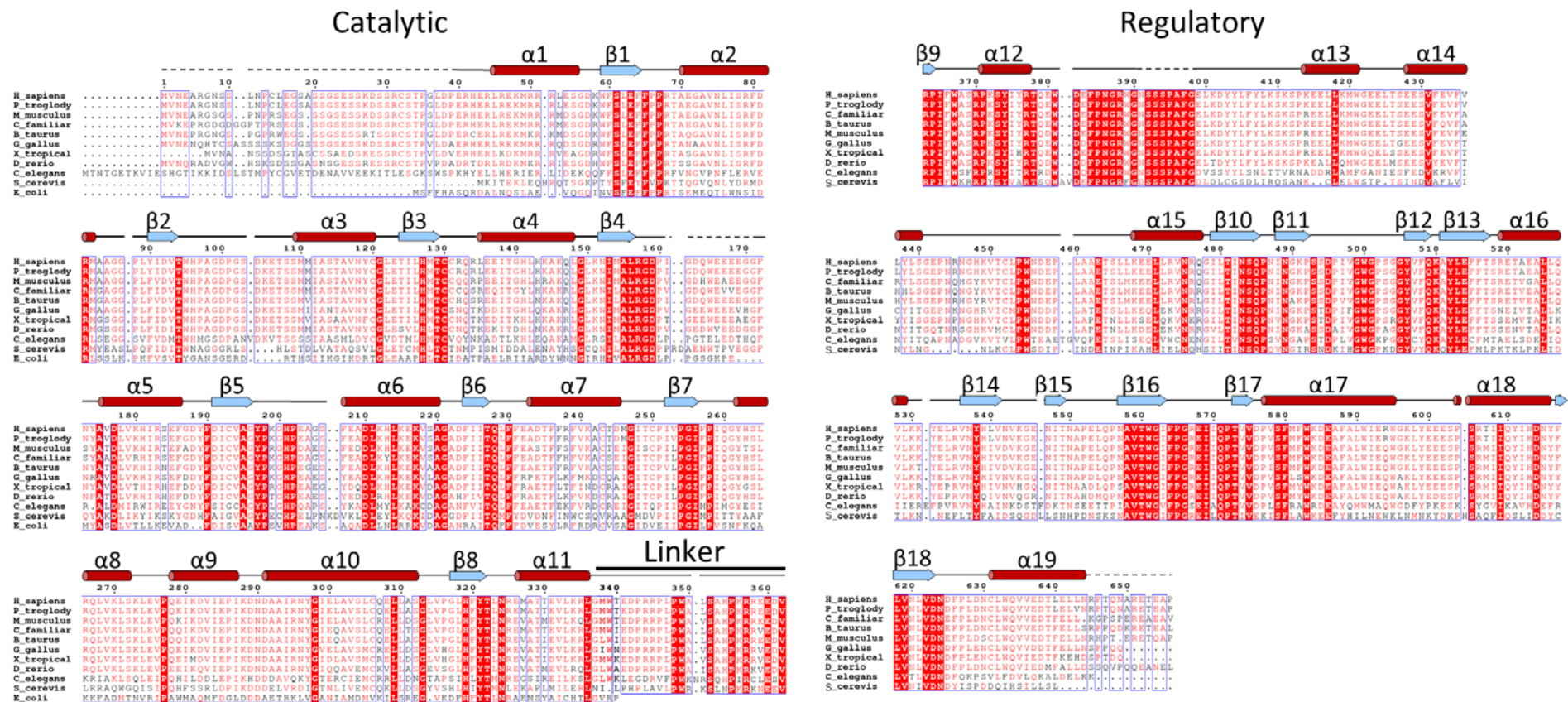
**Supplementary Figure 12. Dynamic charge radius plot for *HsMTHF*<sub>1-656</sub>.** Dynamic charge radius plot of untreated and CIP treated *HsMTHFR*<sub>1-656</sub>. The mean calculated charge radii for phosphorylated and dephosphorylated MTHFR at the modal charge state +27 are similar at 41.8 Å. However, dephosphorylation generates a shift in the charge envelope towards higher charge. Comparison of the peak widths (full width at half maximum, FWHM) suggests that dephosphorylated MTHFR is significantly more flexible than the phosphorylated form.



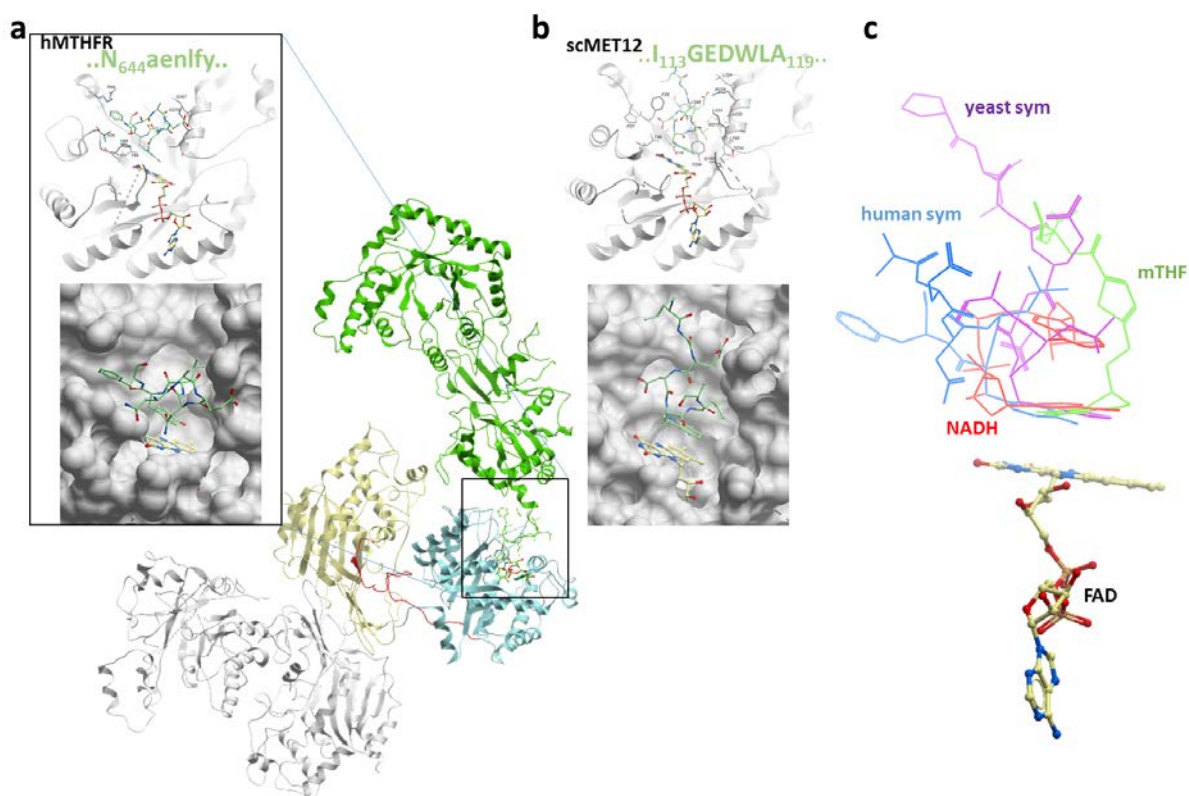
**Supplementary Figure 13. Structure-based sequence alignment of the catalytic domain of HsMTHFR<sub>38-644</sub> with the catalytic domain of ScMET12 and MetF of Haemophilus influenza (5UME), EcMTHFR (1B5T) and TtMTHFR (3APT).** Amino acid conservation is indicated by a coloured background ranging from yellow (conserved in three proteins) to green (invariant residue). Above the alignment the secondary structure of HsMTHFR is given, with red cylinders indicating α-helices and cyan arrows indicating β-sheets.



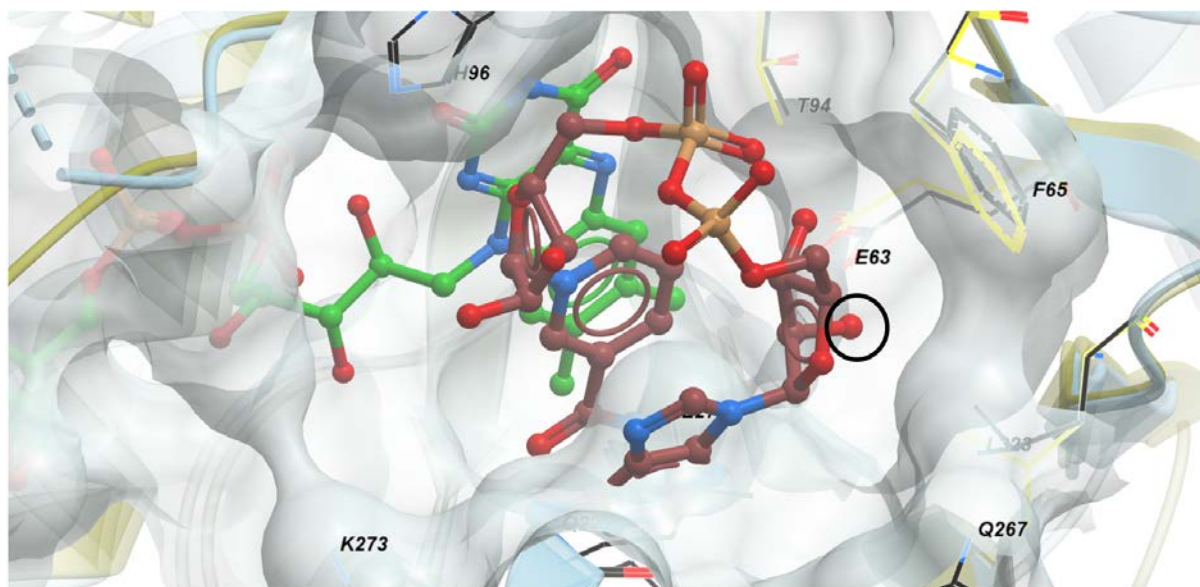
**Supplementary Figure 14. Comparison of the N-termini of various MTHFR catalytic domain structures.** They include *HsMTHFR* (blue) and *ScMET12* (green) from this study, as well as MTHFR from *T. thermophilus* (yellow), *E. coli* (pink), and *H. influenzae* (purple). In all five structures, the first helix of the catalytic domain  $\alpha 1$  is coloured red, to highlight the two different orientations observed.



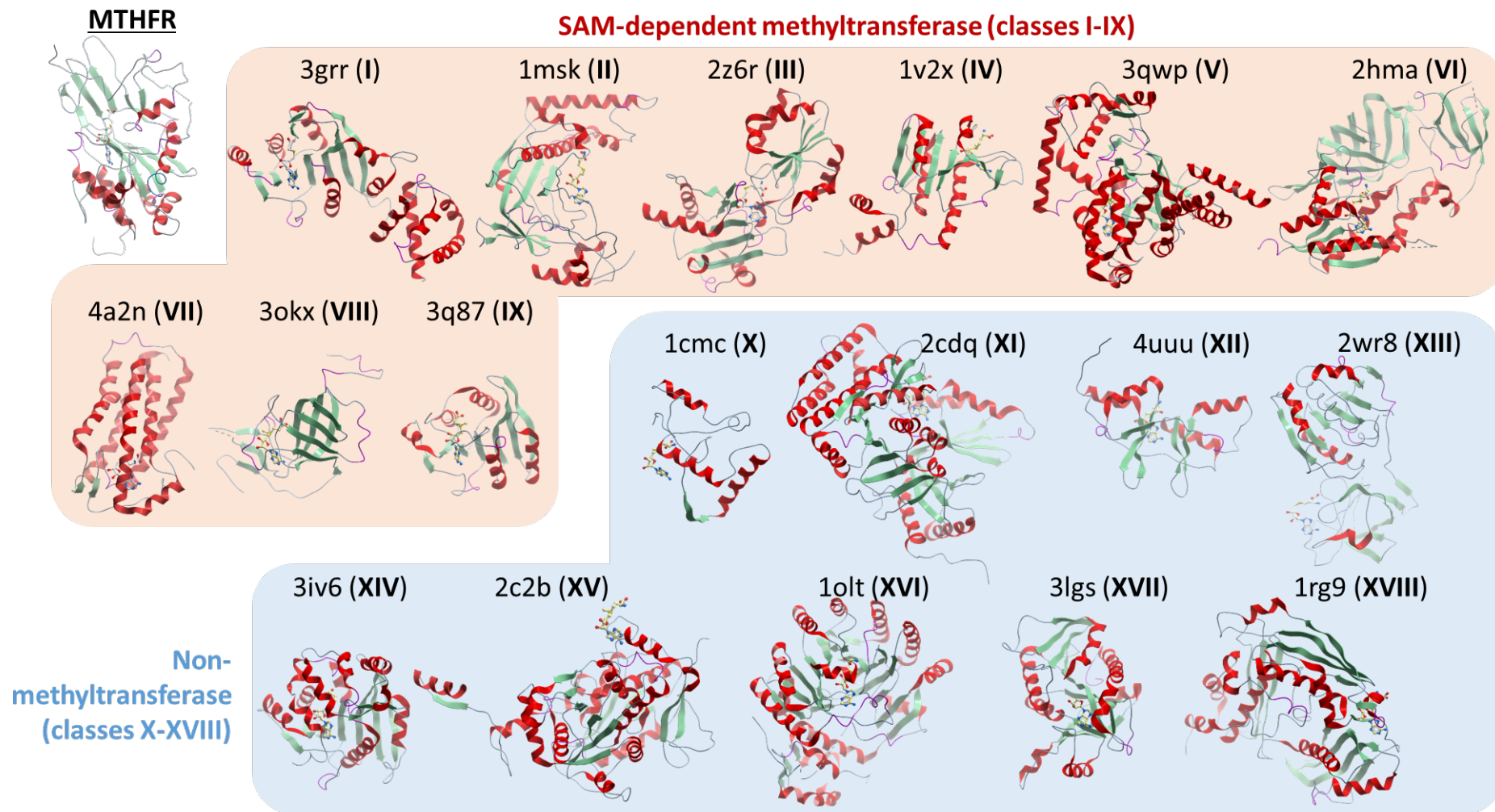
**Supplementary Figure 15. Multiple sequence alignment of MTHFR.** Multiple sequence alignment was performed for *Homo sapiens* (H\_sapiens, NP\_005948.3), *Pan troglodytes* (P\_troglody, XP\_00944660.1), *Mus musculus* (M\_musculus, NP\_034970.2), *Canus familiaris* (C\_familiar, XP\_005618050.1), *Bos taurus* (B\_taurus, NP\_001011685.1), *Gallus* (G\_gallus, XP\_417645.3), *Xenopus tropicalis* (X\_tropical, NP\_001096464.1), *Danio rerio* (D\_rerio, NP\_001268769.1), *Caenorhabditis elegans* (C\_elegans, NP\_741027.1), *Saccharomyces cerevisiae* (S\_cerevis, NP\_011390.2), and *Escherichia coli* (E\_coli, AJ086043.1) for the catalytic domain, and all but *E. coli* for the linker and regulatory domains. Absolutely conserved residues are highlighted in red shading, well-conserved residues in red font. Secondary structure elements from HsMTHFR<sub>38-644</sub> are indicated at the top of the alignment, with red cylinders indicating  $\alpha$ -helices and cyan arrows indicating  $\beta$ -sheets. Alignment was created by multalin<sup>2</sup> and visualized by ESPript<sup>3</sup>.



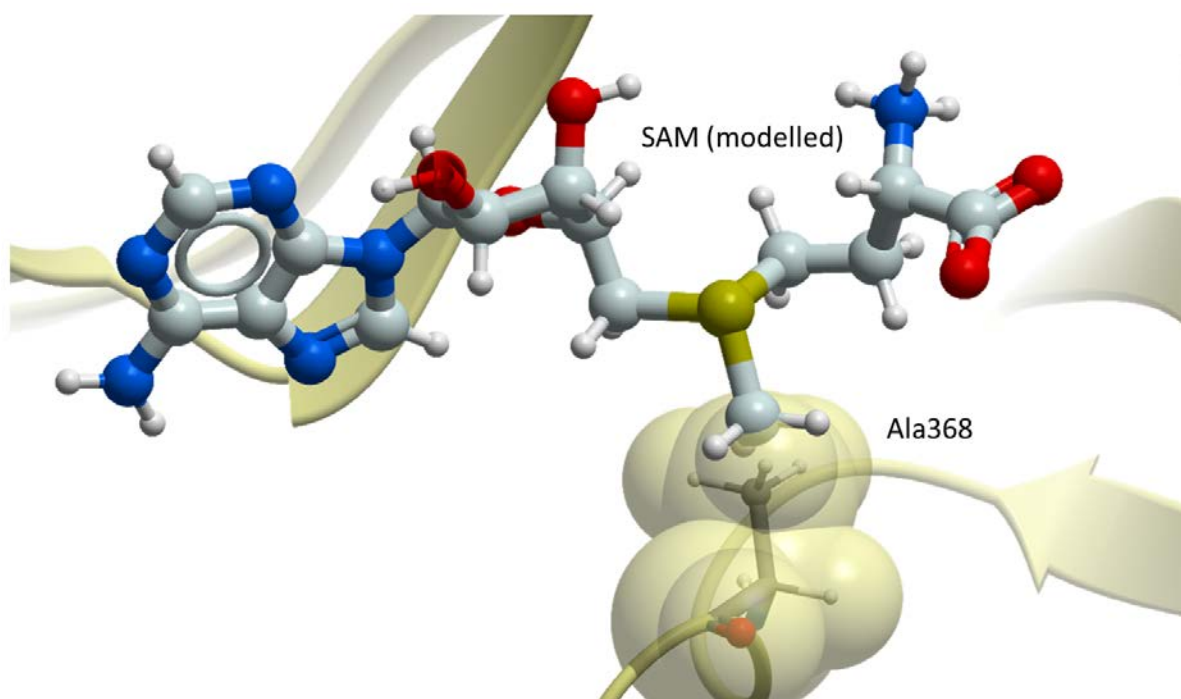
**Supplementary Figure 16. Stacking interactions from adjacent crystallographic subunits stabilize the catalytic domains of ScMET12 and HsMTHFR structures.** **a** In HsMTHFR<sub>38-644</sub>, the final phenylalanine of the uncleaved section of the C-terminal tail His-tag of one subunit stacks against the FAD of chain A of the dimer. The catalytic domain in chain A is therefore held in place by such crystal-mediated packing, resulting in an ordered catalytic domain. This also explains why the catalytic domain in chain B of the dimer, which did not accommodate packing of neighbouring molecule, was largely disordered in the crystal. The additional amino acids contributed by the adjacent unit are shown in green sticks. **b** In the MET12<sub>1-301</sub> structure, an extension of the loop p.R<sub>108</sub>-Y<sub>129</sub> performs this function, whereby p.Trp117 from one-subunit makes stacking interactions with (si face of isoalloxazine ring) the FAD from an adjacent subunit. Once again, additional contributing amino acids are depicted in green sticks. **c** Stacking interactions from adjacent subunits in HsMTHFR (blue) and ScMET12 (purple) extend into the active site to interact with FAD, and thus overlap with the NAD(P)H (red) and CH<sub>3</sub>-THF (green) binding sites.



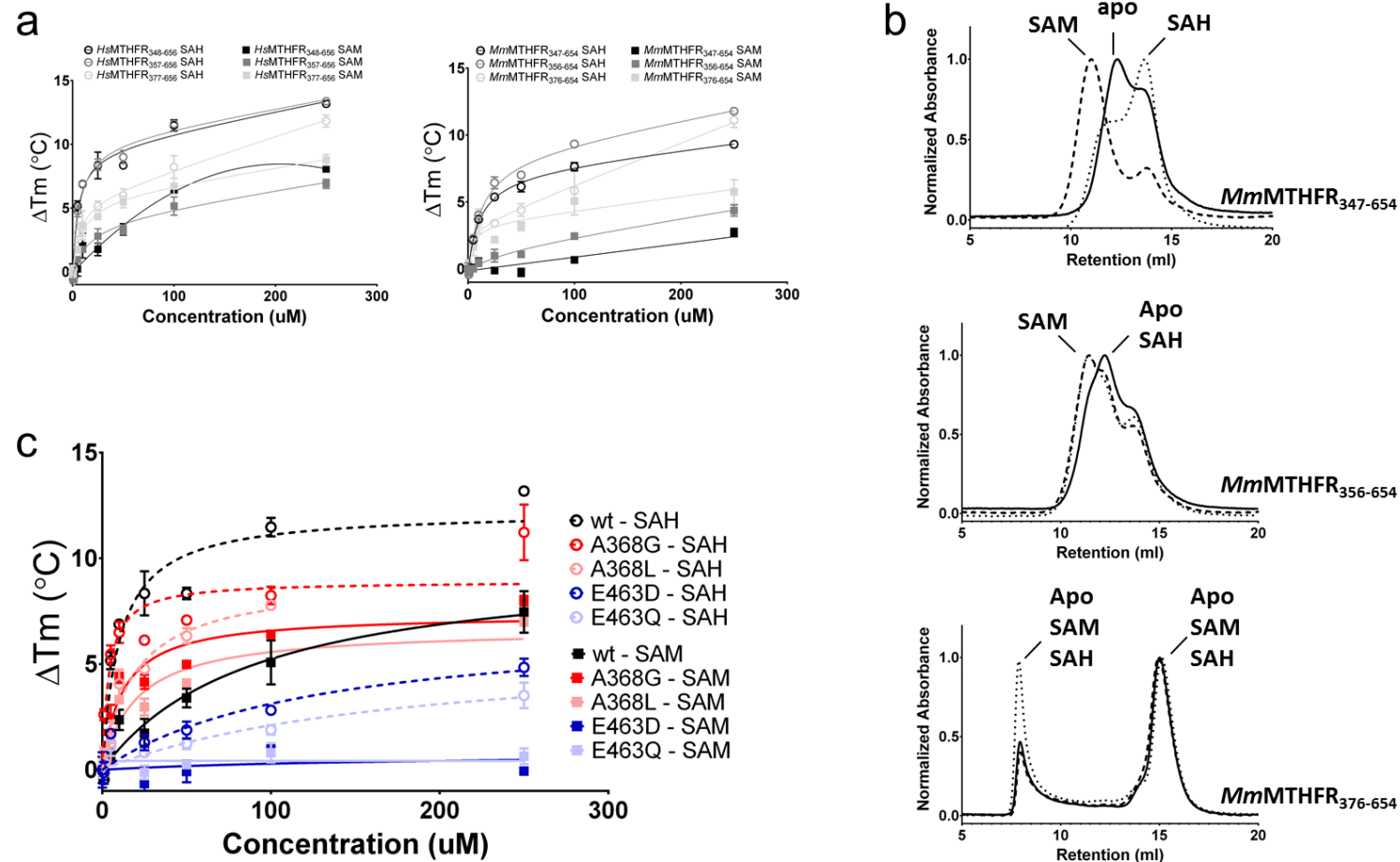
**Supplementary Figure 17. Close-up view of the NADH binding pocket indicating lack of specific accommodation for an additional ribosyl phosphate.** Structural overlay of *HsMTHFR*<sub>38-644</sub> (cyan) and *EcMTHFR* (PDB: IZRQ, yellow) with residues contributing the binding pocket shown in sticks (*HsMTHFR*: black; *EcMTHFR*: yellow) and labeled according to *HsMTHFR*. The ribosyl oxygen which accommodates the additional phosphate group in NADPH is circled in black. Notice the strong sequence and structural conservation of amino acids in the area surrounding this ribosyl oxygen.



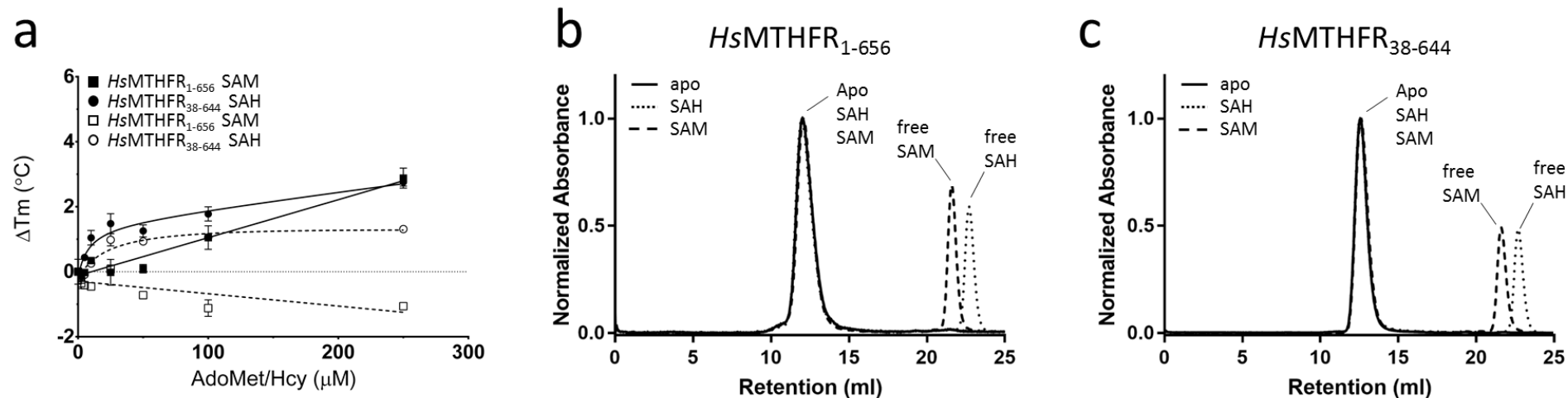
**Supplementary Figure 18. Cartoon representation of SAM binding domain of *HsMTHFR* and other known SAM-binding domains.** The *HsMTHFR* regulatory domain fold is displayed alongside with representative examples from 18 known SAM-binding folds from methyltransferases (classes I-IX) and non-methyltransferases (classes X-XVIII). All the structures are shown with the SAM ligand (sticks) in equivalent orientation.



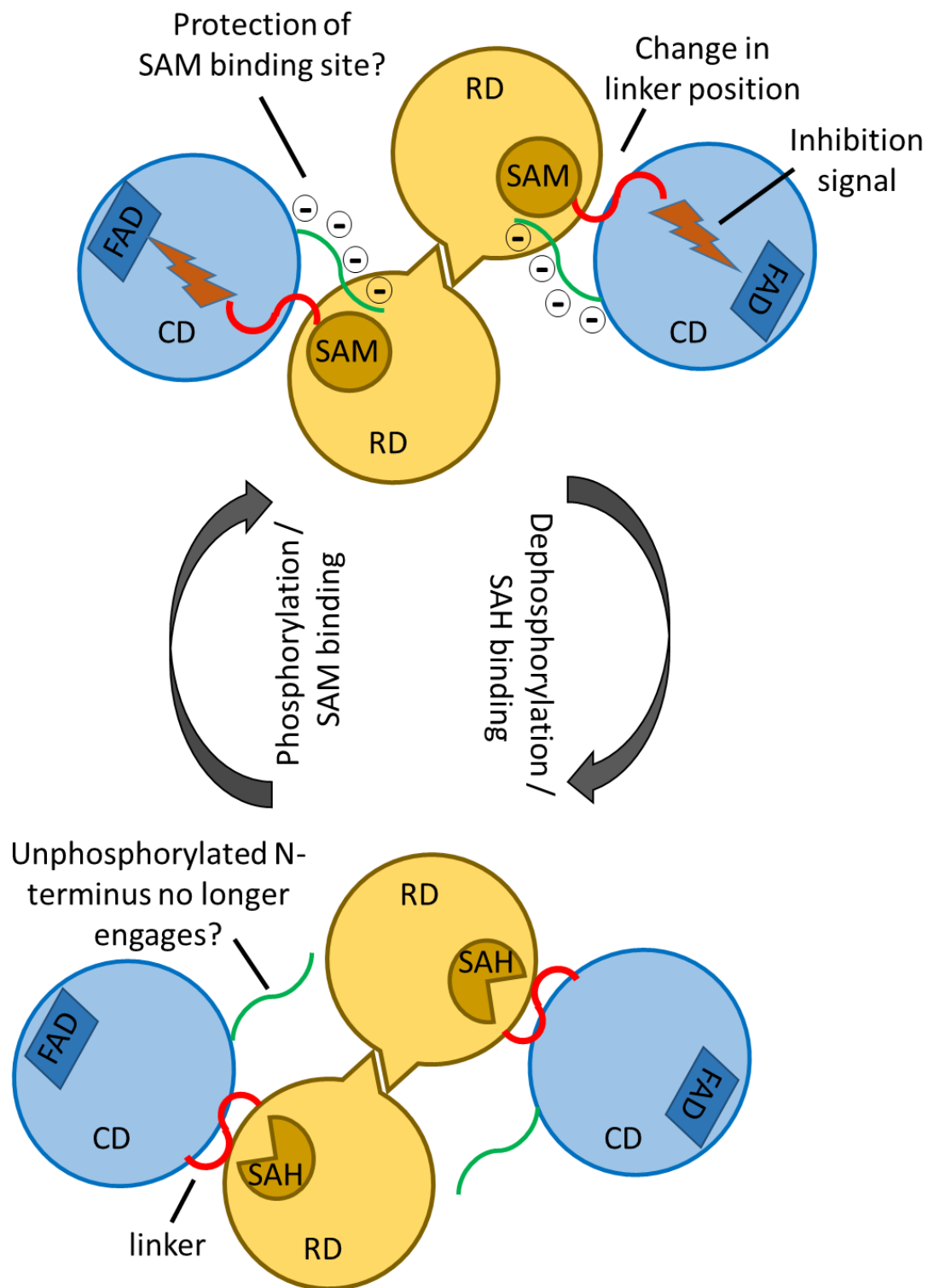
Supplementary Figure 19. Clash between modelled SAM and Ala368 of *HsMTHFR*<sub>38-644</sub>.



**Supplementary Figure 20. SAH/SAM binding and conformational change in N-terminally truncated *HsMTHFR* and *MmMTHFR*.** **a** Differential scanning fluorimetry of N-terminally truncated *HsMTHFR* (left panel) and mouse (*Mm*)*MTHFR* (right panel) following incubation with increasing concentrations of SAH or SAM. Lines represent best fit curves of one-site binding calculated by GraphPad v6.0. **b** Size exclusion chromatography of *MmMTHFR* with various N-terminal truncations following incubation with SAM (dashed lines), SAH (dotted lines), or nothing (apo; solid line). **c** Differential scanning fluorimetry of *HsMTHFR*<sub>348-656</sub> protein without (wt) or carrying the corresponding mutations following incubation with increasing concentrations of SAH or SAM. Lines represent best fit curves of one-site binding of SAH (dashed) or SAM (solid) as calculated by GraphPad v6.0. For **a** and **c**, each value represents the results of at least 3 separate experiments and is given as  $\pm$  S.D.



**Supplementary Figure 21. SAH/SAM binding does not induce conformational change in *HsMTHFR*<sub>1-656</sub> or *HsMTHFR*<sub>38-644</sub>.** **a** Differential scanning fluorimetry of *HsMTHFR* following incubation with increasing concentrations of SAH or SAM. Each value represents the results of at least 3 separate experiments and is given as  $\pm$  S.D. Lines represent best fit curves of one-site binding calculated by GraphPad v6.0. **b** Size exclusion chromatography of *HsMTHFR*<sub>1-656</sub> following incubation with SAM (dashed lines), SAH (dotted lines), or nothing (apo; solid line). **c** Size exclusion chromatography of *HsMTHFR*<sub>38-644</sub> following incubation with SAM (dashed lines), SAH (dotted lines), or nothing (apo; solid line).



**Supplementary Figure 22. Schematic mechanism of human MTHFR postulated from this study.** N-terminal phosphorylation region is coloured green, catalytic domain (CD) brown, regulatory domain (RD) blue, and inter-domain linker red.

**Supplementary Table 1. SAXS analysis of *HsMTHFR***

|                                                                                                     | <b>MTHFR<sub>1-656</sub> CIP-treated</b> | <b>MTHFR<sub>1-656</sub> as-purified</b> |
|-----------------------------------------------------------------------------------------------------|------------------------------------------|------------------------------------------|
| <b>Structural parameters</b>                                                                        |                                          |                                          |
| Rg (Å, from Guinier)                                                                                | 40.8 ± 0.5                               | 41.3 ± 0.5                               |
| Volume (Å <sup>3</sup> , from Porod)                                                                | 220                                      | 254                                      |
|                                                                                                     |                                          |                                          |
| <b>Mr determination (kDa)<sup>1</sup></b>                                                           |                                          |                                          |
| Estimated from Porod volume                                                                         | 137                                      | 158                                      |
| Estimated from volume-of-correlation (Vc)                                                           | 140                                      | 150                                      |
|                                                                                                     |                                          |                                          |
| <b>Fit to CORAL model generated from <i>HsMTHFR</i><sub>38-644</sub><sup>2</sup></b>                |                                          |                                          |
| Chi-square                                                                                          | 23.2 ± 1.2                               | 29.8 ± 1.6                               |
| <b>Fit to scattering back-calculated from <i>HsMTHFR</i><sub>38-644</sub> structure<sup>3</sup></b> |                                          |                                          |
| Chi-square                                                                                          | 5.9                                      | 11.6                                     |

<sup>1</sup>Mr calculated from sequence of a *HsMTHFR*<sub>38-644</sub> dimer: 150.7 kDa

<sup>2</sup>To clarify the difference in overall shape, we fitted the theoretical scattering curves of the *HsMTHFR*<sub>38-644</sub> rigid body CORAL models to the SAXS experimental data of dephosphorylated and phosphorylated MTHFR.

<sup>3</sup>We also fit the theoretical scattering curve of *HsMTHFR*<sub>38-644</sub> observed in the crystal, to the experimental data of dephosphorylated and phosphorylated MTHFR

### Supplementary References

1. 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).
2. Corpet F. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res* **16**, 10881-10890 (1988).
3. Gouet P, Robert X, Courcelle E. ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res* **31**, 3320-3323 (2003).