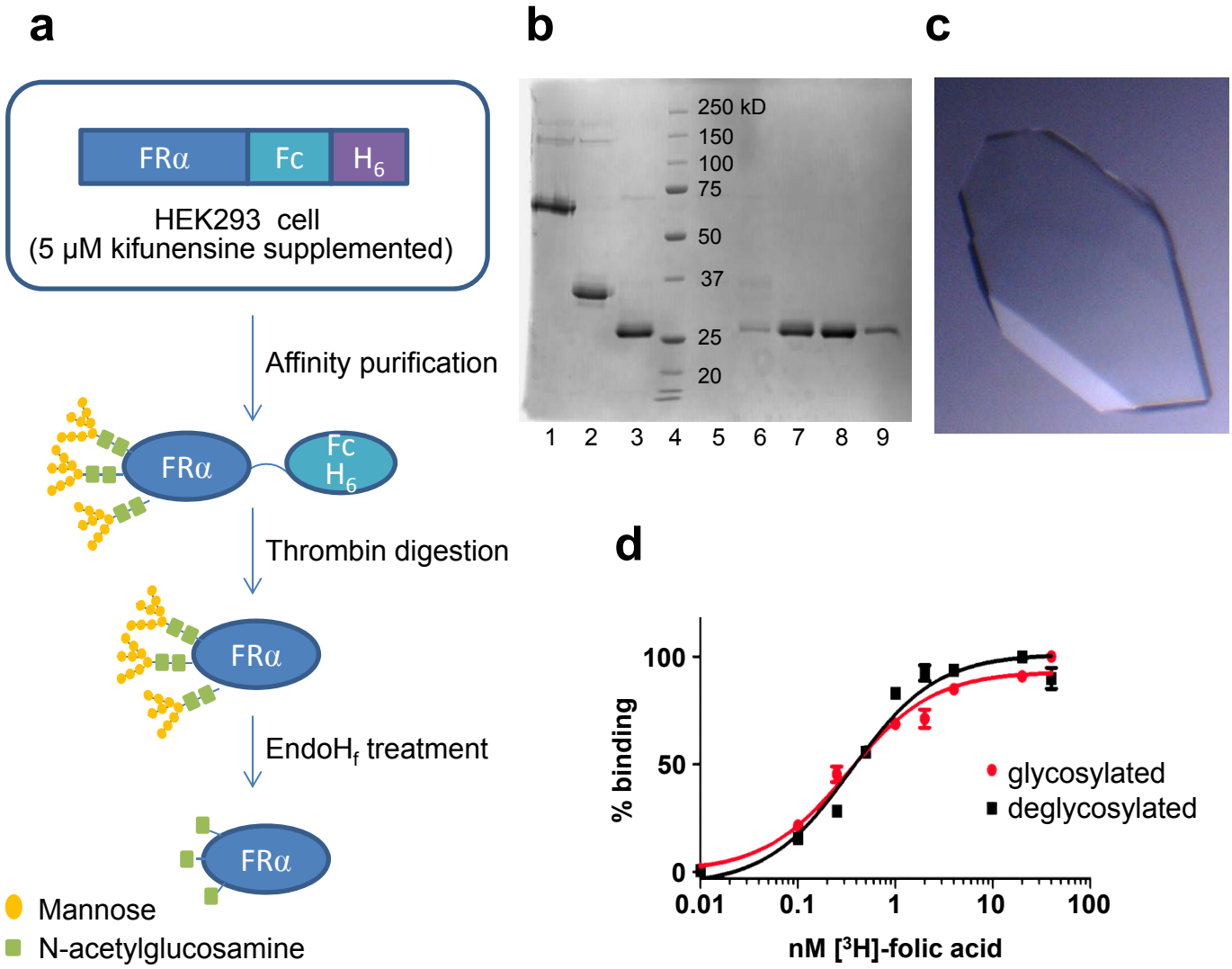
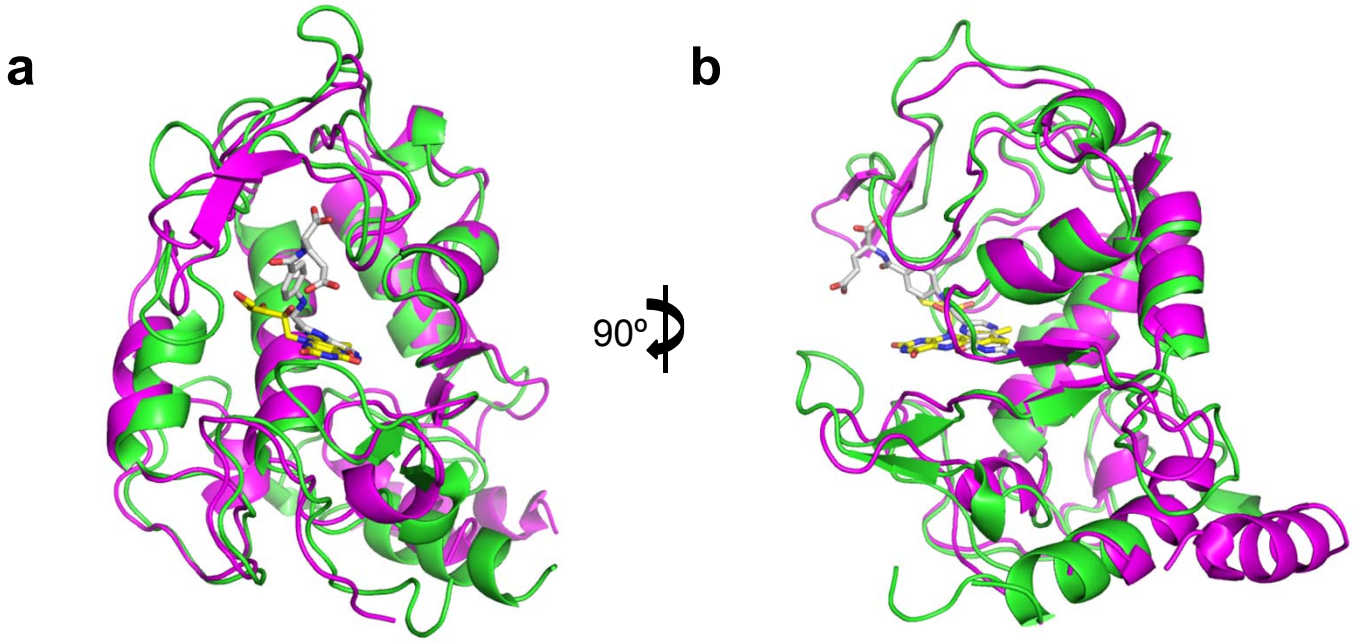
doi:10.1038/nature12327



Supplementary Fig. 1. Purification, crystallization, and ligand-binding activity of FRα. a, Purification flowchart. **b,** Gel analysis of FRα at various stages of purification. Lane 1 shows FRα-Fc protein after the first His_6 tag purification. Lane 2 shows FRα obtained after thrombin digestion and removal of the Fc tag through a second His_6 tag purification. Lane 3 shows deglycosylated FRα after Endo H_f cleavage. Lane 4 is a protein marker. Lanes 5-9 show samples from each fraction of the gel filtration column peak. **c,** FRα crystal. **d,** ^{3}[H] folic acid binding assay for K_d determination of purified protein before and after deglycosylation. Error bars indicate SD (n=2).



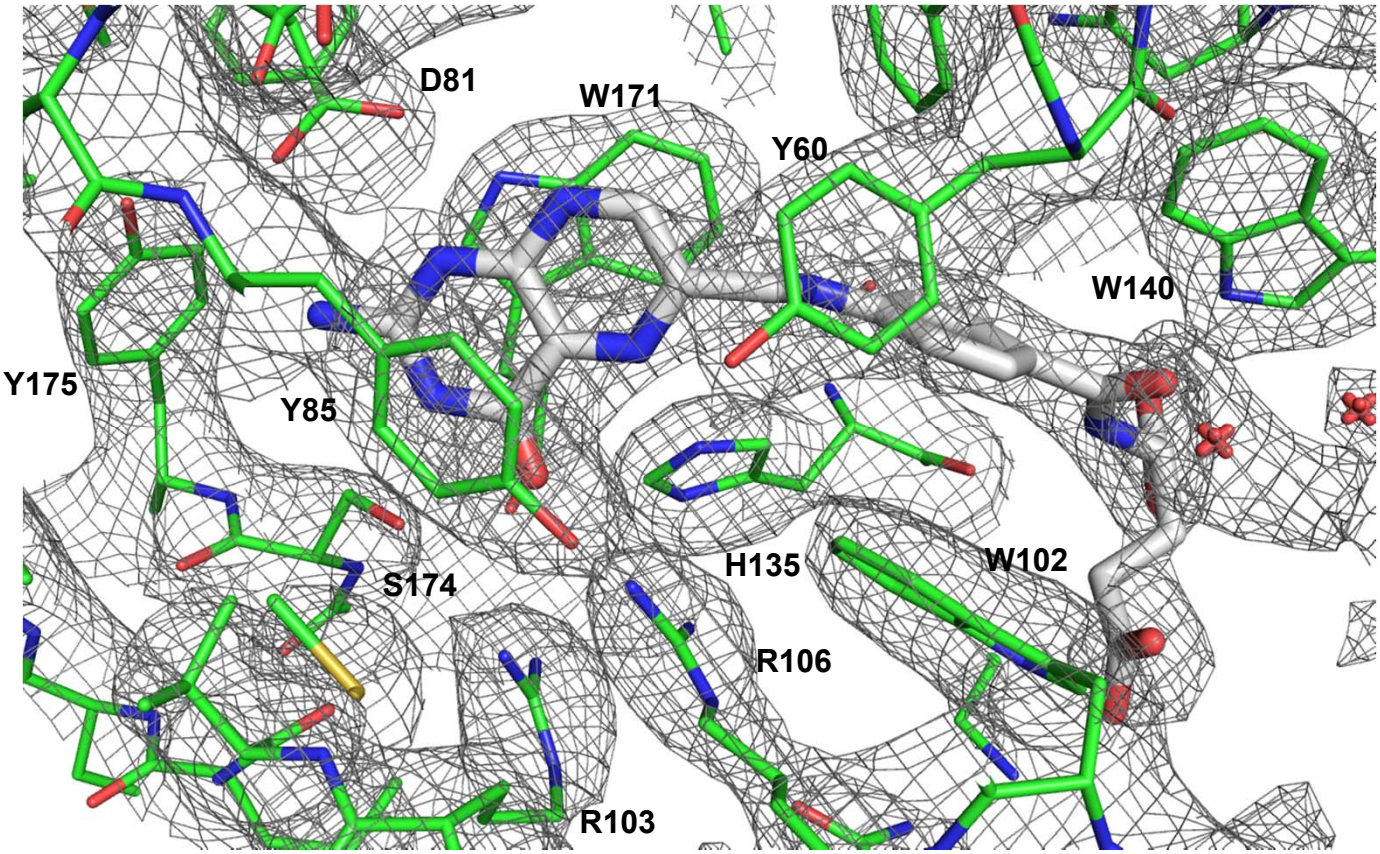
c

```
                1                  10                  20                  30
FRα     QTRIAWARTELLNVCMNAKHHKEKPGPEDKL   31
RfBP    ..........QQYGCLEGDTHKANPSPEPNM   21
                                  ★

               40                  50                  60                  70
HEQCRPWRKNACCSTN.TSQEAHKDVSYLYRFNWNHCGEMA   71
HE.CTLYSESSCCYANFTEQLAHSPIIKVSNSYWNRCGQLS   61
   ★      ★★                          ★

               80                  90                  100                 110
PACKRHFIQDTCLYECSPNLGPWIQQVDQSWRKERVLNVPL   112
KSCEDFTKKIECFYRCSPHAARWI....DPRYTAAIQSVPL   98
 ★          ★    ★

              120                 130                 140                 150
CKEDCEQWWEDCRTSYTCKSNWHKGWNW.TSGFNKCAVGAA   152
CQSFCDDWYEACKDDSICAHNWLTDWERDESGENHCKSK..   137
★    ★          ★    ★                   ★

              160                 170                 180                 190
CQPFHFYFPTPTVLCNEIWTHSYKVSNYSRGSGRCIQMW..   191
CVPYSEMYANGTDMCQSMWGESFKVSE...SSCLCLQMNKK   175
★                   ★                   ★

                  200                 210
....FDPAQGNPNEEVARFYAAAMS..............   212
DMVAIKHLLSESSEESSSMSSSEEHEHACQKKLLKFEAL   214
```

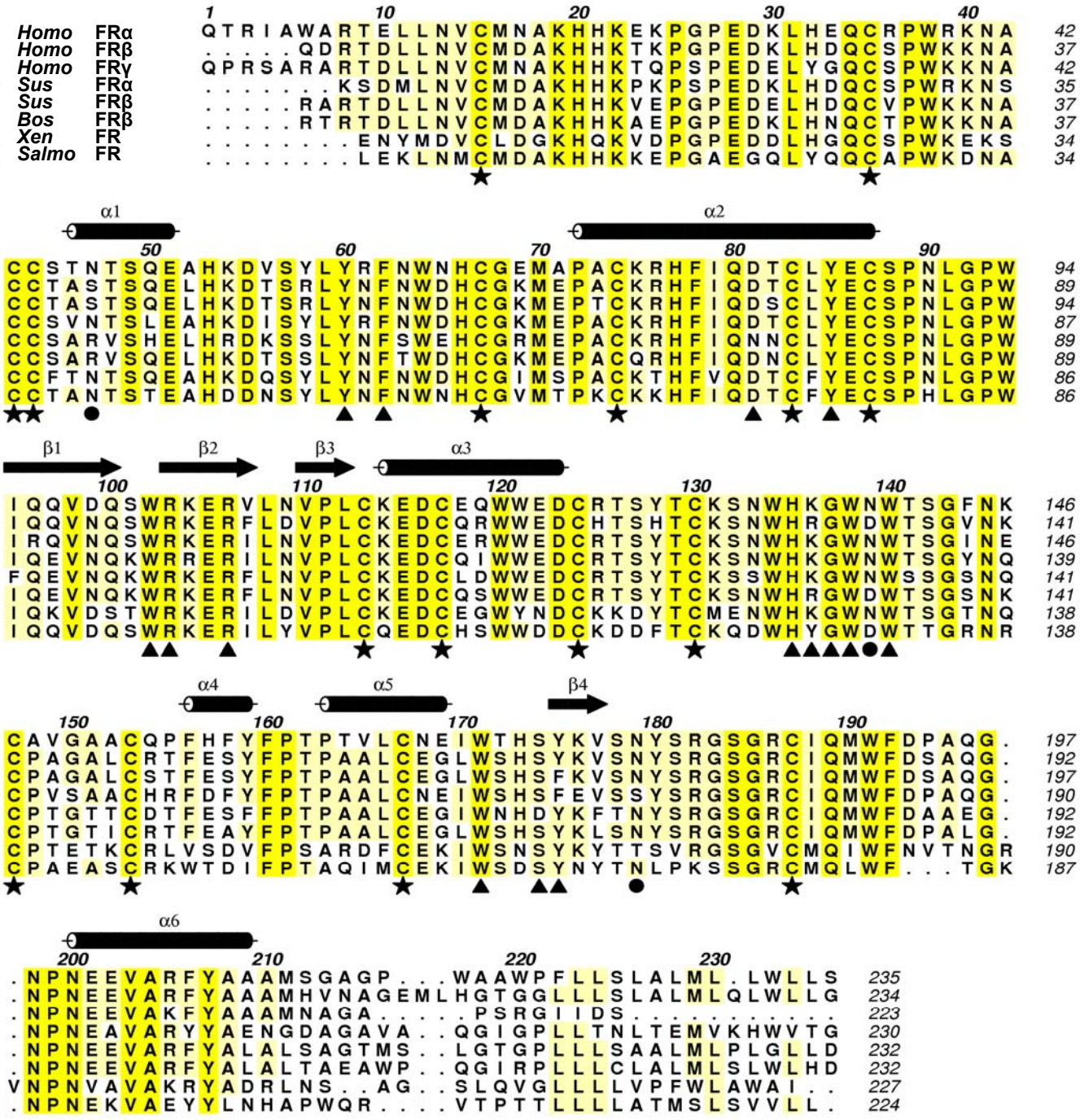
Supplementary Fig. 2: Comparison of FRα and chicken riboflavin binding protein. a, b, Two views of the structure alignment between the FRα–folic acid complex and chicken RfBP–riboflavin complex. FRα is colored in magenta, folic acid in grey, RfBP in green, and riboflavin in yellow. **c,** Sequence alignment between human FRα and chicken RfBP with conserved cysteines indicated by asterisks.

FRs were proposed to be structural homologs of chicken riboflavin binding protein[16]. Structure alignment of chicken RfBP with FRα yielded rmsd values for their 163 C_α atom positions of 1.56 Å. The four-helix bundle structure and the arrangement of eight out of nine disulfide bonds from RfBP are conserved in FRα. The ninth disulfide bond, which connects the last helix of RfBP to the ligand-binding domain, is missing in FRα, which lacks a homologous helix. Careful comparison of the ligand-binding

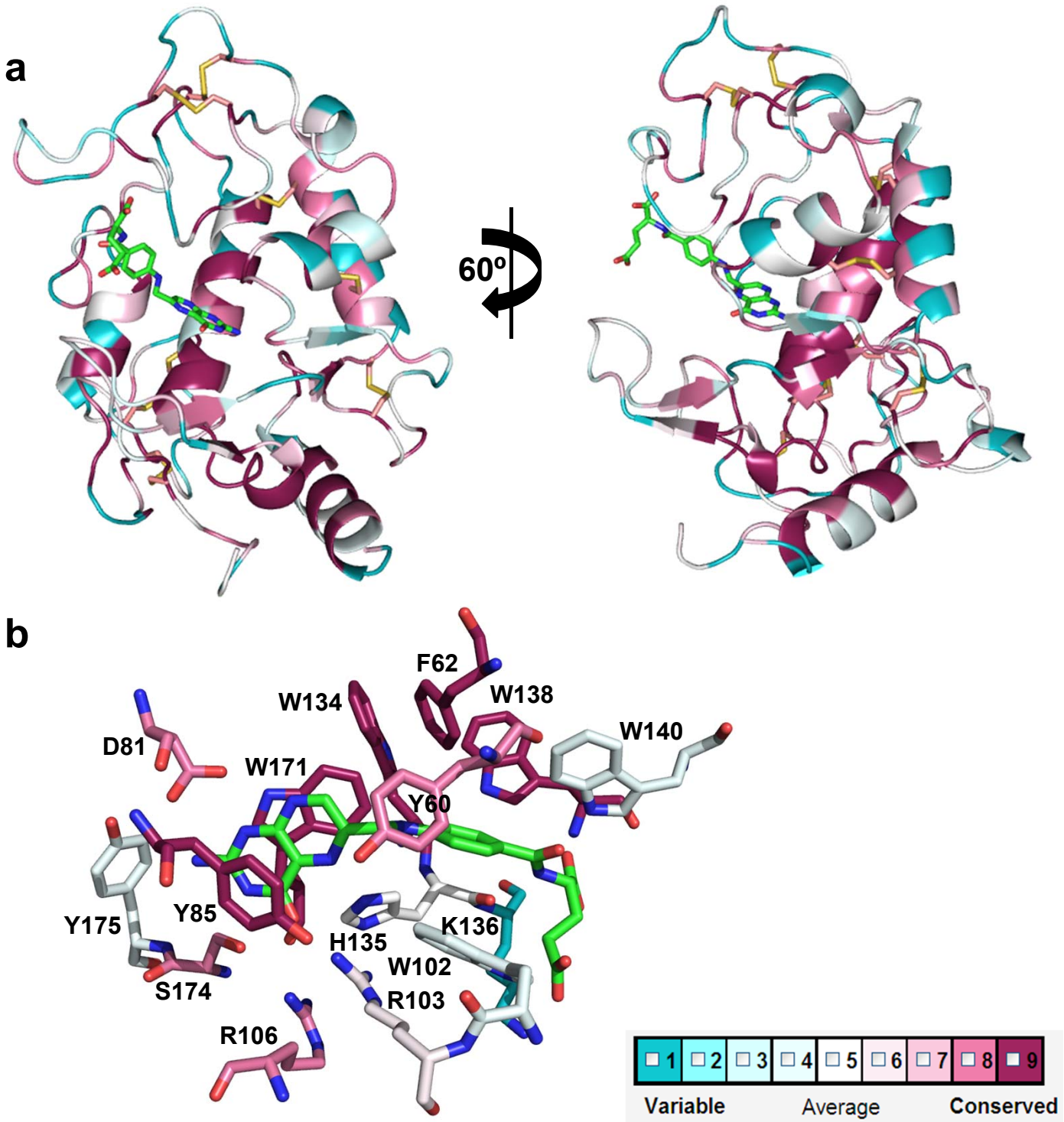
pockets revealed different ligand recognition mechanisms for the two receptors. The flat riboflavin is accommodated in a fissure-like binding pocket of RfBP with a wide opening of ~17Å. In contrast, the relatively linear folic acid molecule fits well in the FRα binding channel with limited access to solvent. Furthermore, riboflavin has an amphipathic isoalloxazine ring with its hydrophobic part deeply buried inside the hydrophobic groove of the binding cavity, while the hydrophilic pterin ring of folic acid faces inside and forms a series of hydrogen bonds with FRα. Remarkably, in spite of the distinct mechanisms to bind the ligand head groups, Y75 and W156 of RfBP are conserved in FRα (Y85 and W171) to stack the isoalloxazine/pterin ring. Finally, while the pterin ring of folic acid and the isoalloxazine ring of RfBP assume overlapping positions in their respective ligand-binding pockets, a kink at C9 of folic acid bends the molecule, which causes the ligand tail groups to point into different directions. Thus, the glutamate moiety of folic acid is stabilized by the α3-α4 loop, while the tail of riboflavin forms several H-bonds with α2 and the α1-α2 loop.



Supplementary Fig. 3. Annealed omit map of folic acid and folic acid-binding side chains. The omit map is shown within 9Å of folic acid and is contoured at 1.0 σ. The ligand and ligand binding pocket residues are shown as stick models with carbon atoms colored in grey for ligand and green for protein residues. N, O, and S atoms are colored in blue, red, and yellow, respectively. The annealed omit map excluding the folic acid ligand was calculated using CNSsolve (http://cns-online.org/v1.3/).

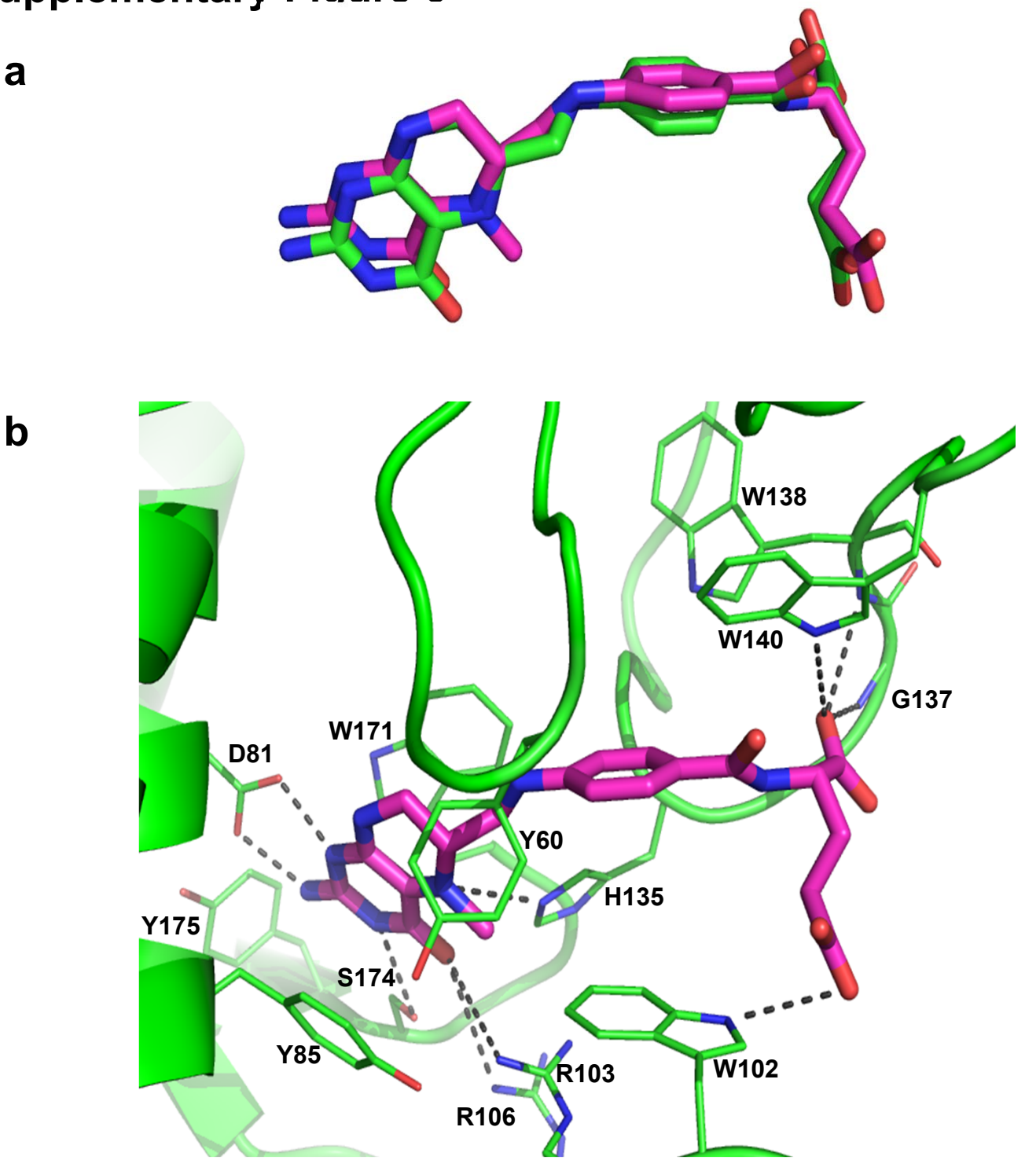


Supplementary Fig. 4. Sequence comparison of different folate receptor subtypes. The alignment includes subtypes from five different species: human (*Homo*), pig (*Sus)*, cow (*Bos*), *Xenopus* (*Xen*) and salmon (*Salmo*). Numbers (1-3) denote FR α, β, γ subtypes. Secondary structure elements of human FRα are indicated above the sequences (cylinders: α-helices, arrows: β-strands). Stars below the sequences denote cysteine residues, circles denote N-glycosylation sites, and triangles folic acid-binding residues in FRα.

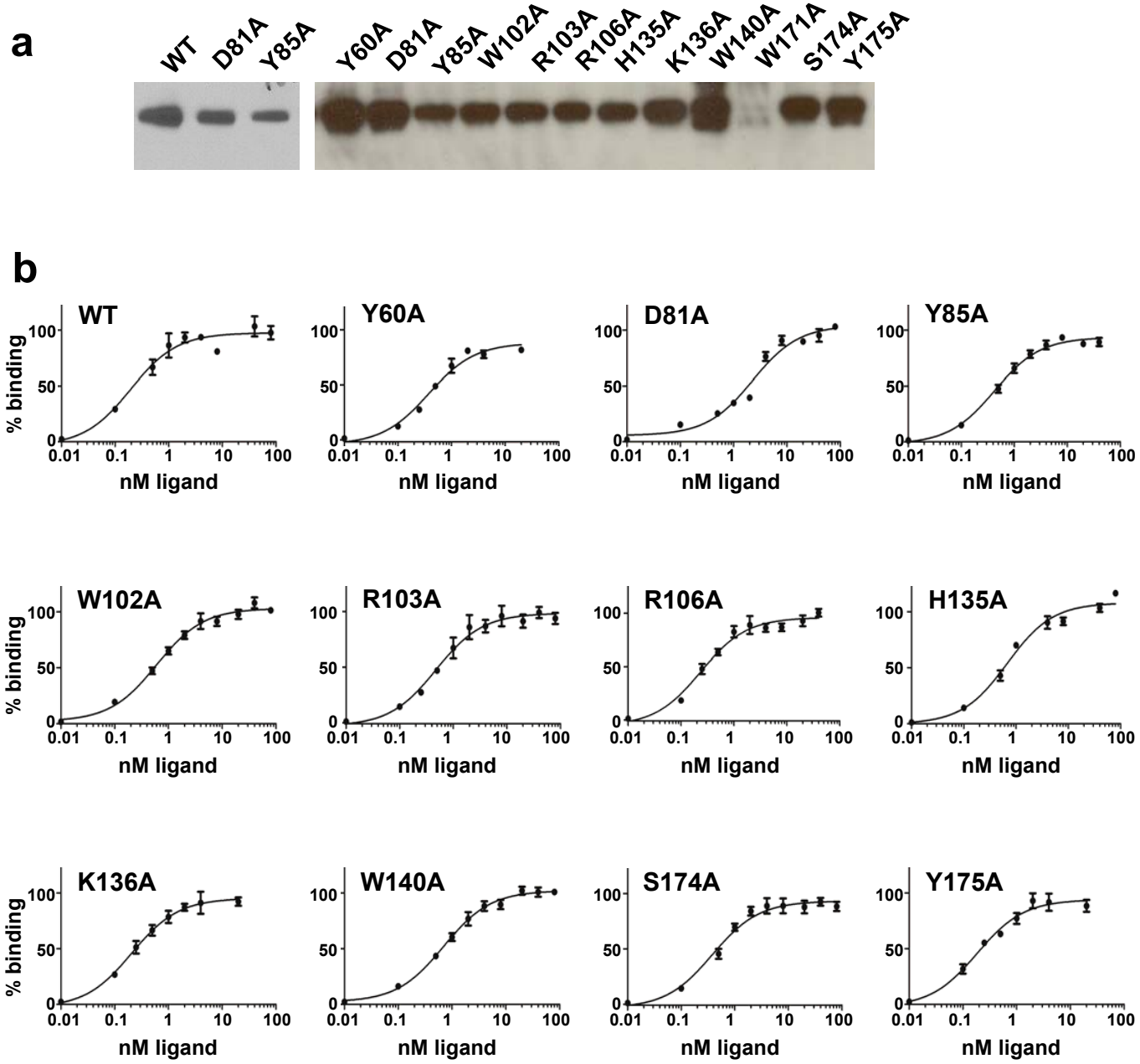


Supplementary Fig. 5. Conservation heat map overlaid on the FRα–folic acid structure. **a,** Two views of the ribbon presentation of the FRα–folic acid crystal structure overlaid with a conservation heat map. **b,** Close up of the conservation map with amino acid residues presented as stick models in the ligand-binding pocket. The conservation map is compiled from 61 sequences with 60-95% sequence identity to FRα using the program ConSurf (http://consurftest.tau.ac.il/). The heat map color code bar (bottom right) displays 9 equally-spaced bins of relative sequence conservation. Residues shown in white (bin 5) represent the average level of conservation across all FR residues, while bin 1 (blue) represents the residues with the lowest relative conservation (average of 46.2 % sequence identity), and bin 9 the residues with the highest relative conservation (100% sequence identity).

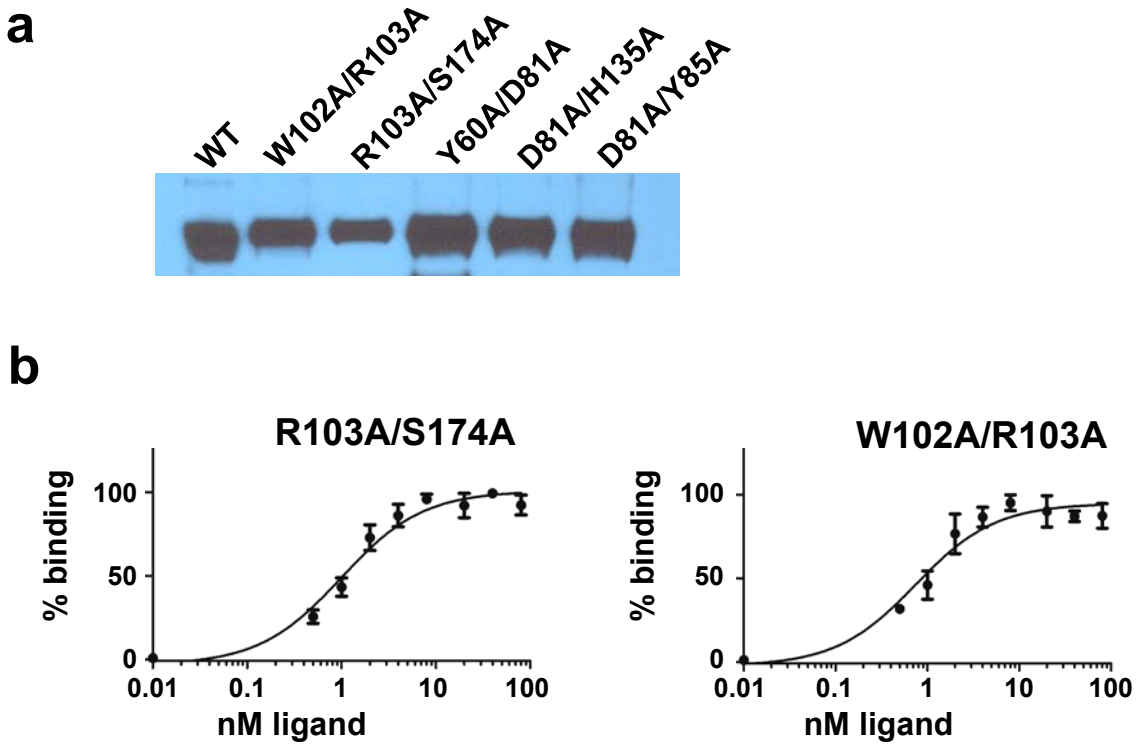
Supplementary Figure 6



Supplementary Fig. 6. Modeling of 5-mTHF in the FRα ligand-binding pocket. a, Overlay of folic acid (green) from the FRα–folic acid crystal structure with 5-mTHF (magenta) fitted into the FRα crystal structure, as shown in **b,** A 5-mTHF model was built based on the structure of tetrahydrofolate (PDB code: 2YCK). The model was manually docked into the FRα ligand binding pocket assuming similar binding orientation as folic acid. Most of the H-bonds formed with folic acid can also be formed with 5-mTHF. Note that the side chains of H135 and R103 are in close contact with the 5-methyl group of 5-mTHF (2.2 Å and 3.0 Å, respectively), which need to be moved to accommodate the 5-methyl group of 5-mTHF.

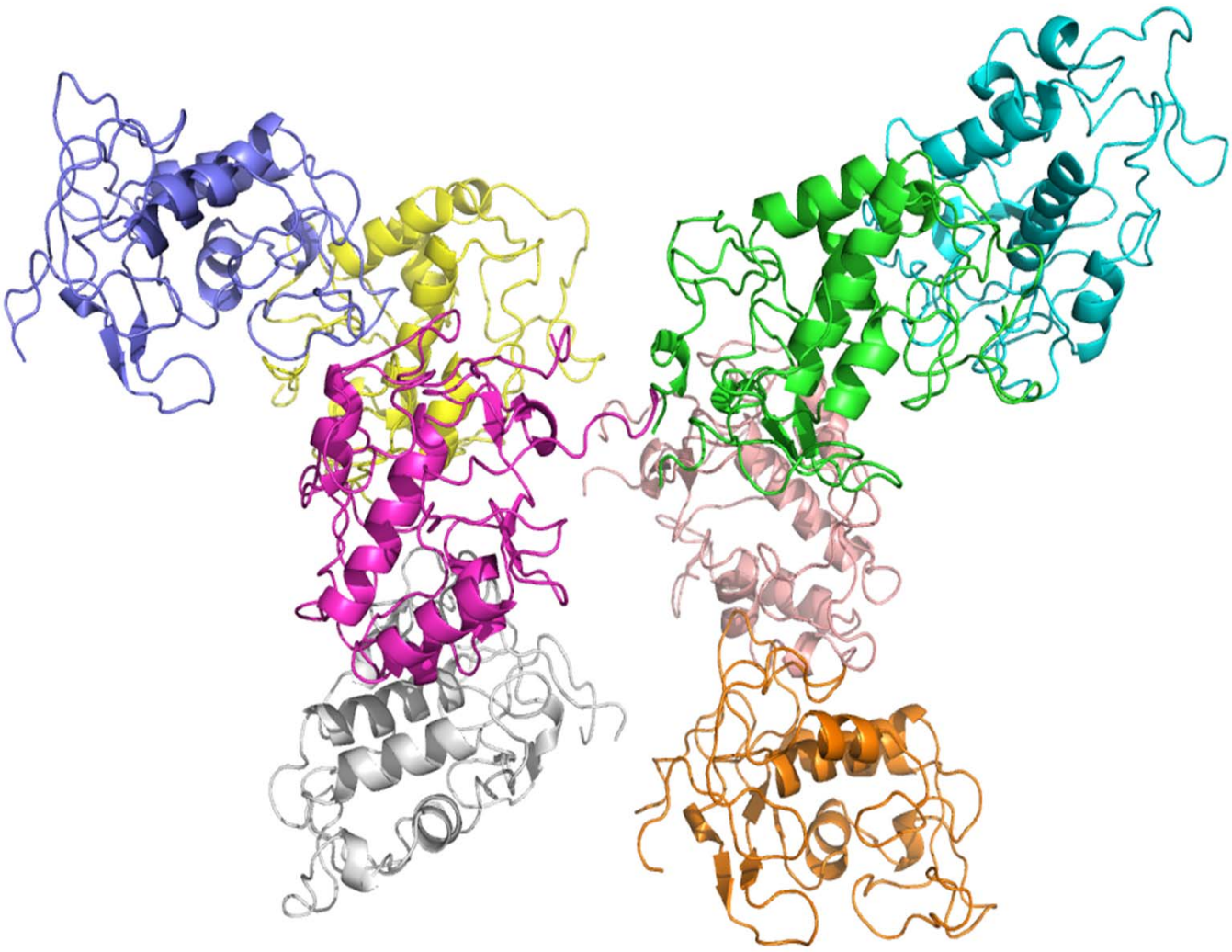


Supplementary Fig. 7. Ligand-binding assays of FRα mutant proteins with single amino acid substitutions. a, Immunoblot of secreted FRα wildtype and mutant proteins. Each lane contains equal amounts of culture medium from HEK293 cells transiently transfected with the indicated wildtype or mutant FRα expression construct. **b,** ^{3}[H]-folic acid binding isotherms using equal molar amounts of each protein. K_d values determined from the isotherms are shown in Fig. 3b. Error bars=SD (n=2).



Supplementary Fig. 8. Ligand-binding assays of FRα mutant proteins with double amino acid substitutions. a, Immunoblot of secreted FRα wildtype and mutant proteins. Each lane contains equal amounts of culture medium from HEK293 cells transiently transfected with the indicated wildtype or mutant FRα expression construct. **b,** ^{3}[H]-folic acid binding isotherms using equal molar amounts of each protein. K_d values determined from the isotherms are shown in Fig. 3b. Error bars=SD (n=2).

Note that Y60A/D81A, D81A/H135A, and D81A/Y85A have normal expression levels, but folate binding was undetectable, indicating these double mutations abolish folate binding of FRα. Y85A/W102A, H135A/W140A, and W102A/W104A are not expressed, indicating these double mutations destabilize the receptor.



Supplementary Fig. 9 Arrangement of eight FRα–folic acid complexes in each asymmetry unit. One asymmetric unit contains eight copies of FRα–folic acid complexes arranged in an overall Y shape, which looks like an IgG immunoglobulin structure. Overall structures of the eight FRα–folic acid complexes are very similar.

Supplementary Table 1: Data collection, phasing and refinement statistics (MIR)

	Native	Merged sulfur anomalous data	Pt-derivative
Data collection			
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions			
a, *b*, *c* (Å)	91.0, 144.6, 216.8	90.7, 144.8, 212.8	90.6, 145.6, 211.1
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90
Wavelength (Å)	0.9787	1.77	1.008
Resolution (Å)	50-2.8	50-3.0	50-3.55
R_{sym} or R_{merge}	0.150(1.23) *	0.311(2.77) *#	0.126(0.939) *
$I / \sigma I$	15.2(1.8) *	24.5(1.9) *	22.0(2.7) *
Completeness (%)	99.9(99.8) *	100.0(100.0) *	100.0(100.0) *
Redundancy	14.1(11.9) *	75.5(54.2) *	14.6(14.1) *
Refinement			
Resolution (Å)	50-2.80		
No. reflections	70805		
R_{work} / R_{free}	20.6%/25.6%		
No. atoms			
Protein	13568		
Ligand/ion	256		
Water	208		
B-factors			
Protein	68.5		
Ligand/ion	62.0		
Water	54.2		
R.m.s deviations			
Bond lengths (Å)	0.0114		
Bond angles (°)	1.60		

*Values in parentheses are for highest-resolution shell.
Note that the sulfur anomalous data represent the merged data from six individual crystals (see Methods)

PDB ID code 4LRH