The 6th International Conference on Biochemistry and Molecular Biology

Comparative computational analysis of substrate binding mode in bacterial and mammalian methylenetetrahydrofolate reductase

Wanisa Salaemae[1,*], Wanita Pantong[1], Manasphat Khunmud[1], Nunta Churnchow[1]

[1]Department of Biochemistry, Faculty of Science, Prince of Songkla University, Songkhla, Thailand
*E-mail: wanisa.sa@psu.ac.th

Abstract

Methylenetetrahydrofolate reductase (MTHFR) catalyzes the reduction of 5,10-methylenetetrahydrofolate to 5-methyltetrahydrofolate, which is an essential methyl donor used in the final step of methionine biosynthesis. MTHFRs require FAD as a non-covalently bound cofactor and NAD(P)H as a reducing equivalent for FAD. NADH is bound specifically to bacterial MTHFRs whereas the NADPH is preferred in mammalian enzymes. The binding mechanism of substrates and cofactor is well characterized in certain bacterial MTHFRs, but not in mammalian enzymes due to the lack of crystal structure. The modeling structure of an N-terminal catalytic domain of human MTHFR was then generated in this study using the SWISS-MODEL. The best model was predicted to be a homodimer with a 38.33% sequence identity and 1.689 Å RMSD to the FAD-bound *Thermus thermophilus* MTHFR template. Upon the superimpositions, FAD was embedded in the active groove where the *si*-face of its isoalloxazine ring was located against either NADPH or folate. The conserved catalytic residues, E63, D159 and Q228, were found at the catalytic center. Instead of conserved phenylalanine in various bacterial MTHFRs, the NADPH that was built relatively to a hairpin-like NADH was stacked by Q267 in human MTHFR.

Introduction

Methionine is a methyl- and sulfur-containing amino acid that plays a crucial role in numerous metabolic pathways in all living organisms. In protein synthesis, methionine acts as the universal initiator in the initiation step and is also required in the elongation of polypeptides.[1] Its derivative, S-Adenosylmethionine (abbreviated SAM or AdoMet), functions as a common methyl donor in the methylation of various substrates such as nucleic acids, lipids, proteins, and secondary metabolites.[2] Although methionine is requisite in all organisms, *de novo* biosynthesis of methionine can be found only in bacteria, fungi, yeast and plants.[3,4,5,6] In human and animals, methionine is classified as an essential amino acid that cannot be synthesized *de novo* inside the cells. It must be ingested from the methionine-containing protein diets.[7]

Methionine biosynthesis belongs to the same family with threonine and lysine biosynthesis that utilizes aspartate as the primary precursor.[8] Many intermediates are synthesized; however, homoserine is considered as the starting compound in the branching point of methionine biosynthesis (Figure 1). The following steps of the methionine biosynthesis involve the synthesis of *O*-succinyl-L-homoserine (or *O*-acetyl-L-homoserine), the displacement of the succinyl (or acetyl) moiety by cysteine, the β-elimination of cystathionine, and the conversion of homocysteine to methionine, respectively. To be noted, genes that encode the catalyzing enzymes are different among microorganisms.[9] In the last step of methionine biosynthesis, homocysteine is methylated by either of cobalamin (Vitamin B12)

dependent- or independent-methionine synthases, which is the product of either *metE*, *metH* or *met8* genes.[9,10,11,12] While most microorganisms possess only one methionine synthase coding gene, *Escherichia coli* has both *metE* and *metH*.[13] Apart from the formation of the homocysteine moiety from homoserine, methionine biosynthesis also involves the transfer of one-carbon units from serine via folate.[14] An intermediate 5,10-methylenetetrahydrofolate (CH_2-H_4folate) is converted to 5-methyltetrahydrofolate (CH_3-H_4folate), which acts as a co-substrate in the methylation of homocysteine. This folate branch of methionine biosynthesis is highly conserved.

Methylenetetrahydrofolate reductase (MTHFR), which is the *metF* product, provides the sole route for synthesizing the 5-methyltetrahydrofolate in microorganisms.[15] MTHFR has been proposed as a potential therapeutic target for treating infectious diseases due to its importance in infection and virulence.[16,17] However, targeting pathogenic MTHFR might raise a concern about possible side effects since the MTHFR homologue is also found in human liver cells. The human MTHFR (hMTHFR) plays an important role in the homeostasis of homocysteine.[18] The defects occurred by the polymorphisms in hMTHFR increase the risk of cardiovascular disease, Alzheimer's disease, and neural tube defects.[19,20]

MTHFRs are FAD-dependent enzymes that catalyze the transfer of a hydride ion from NAD(P)H to CH_2-H_4folate. The bacterial MTHFRs require NADH as the hydride ion donor whereas the mammalian enzymes utilize NADPH.[21,22] The binding mechanism of the cofactor and substrates is well characterized in the enzymes from certain bacteria such as *E. coli*, *Thermus thermophilus*, and *Haemophilus Influenzae*.[23,24,25] In contrast, it is still unclear for hMTHFR due to the lack of crystal structure. Therefore, this study attempts to identify the structural differences between MTHFRs from human and bacteria particularly in the binding mechanism of substrates and cofactor. The modeling structure of an N-terminal catalytic domain of hMTHFR reported in this study showed that the enzyme adopts the binding mechanism of FAD, NADPH and folate similarly to the bacterial enzymes. Instead of a phenylalanine found in various bacterial MTHFRs, Q267 was predicted to stack the hairpin-like NADPH in hMTHFR. These findings provide some useful information for further structure-based drug discovery.

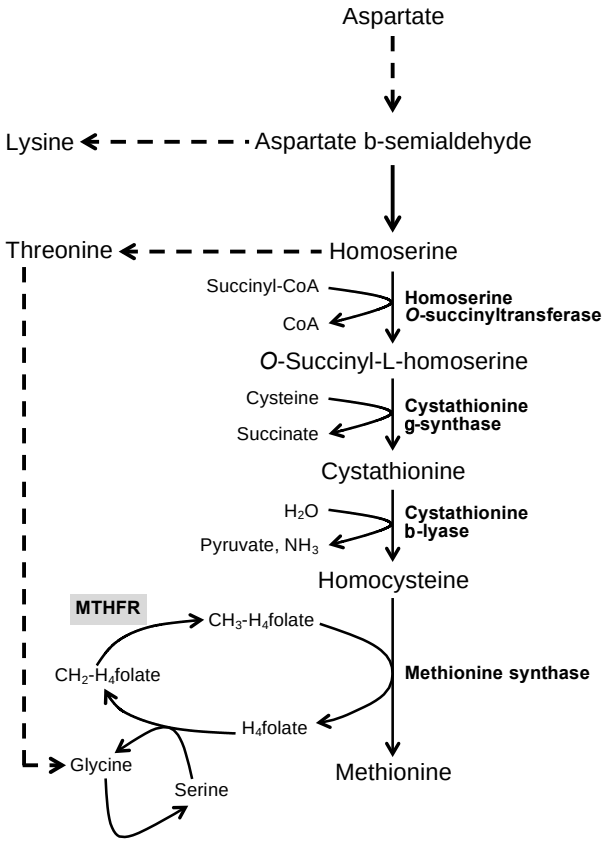

Figure 1. Methionine biosynthesis pathway in microorganisms

Methodology

Alignment of amino acid sequences

The amino acid sequences of MTHFRs from human, *E. coli*, and *T. thermophilus* were obtained from the NCBI database (Accession number: P42896, NP_418376.1, and BAD70150.1). The alignments were performed using the NCBI blastp and the Clustal Omega online software.[26,27]

Homology modeling of human MTHFR

The hMTHFR model was generated using the SWISS-MODEL online software.[28] The target (or input) sequence was the N-terminal 356 amino acid sequences of hMTHFR. Upon the automate mode, the searched template with the highest scores of GMQE (Global Model Quality Estimation), QSQE (Quaternary Structure Quality Estimation), and % identity was used to build the model.

Construction of a 3D hairpin-like NADPH

The hairpin-like NADH from the ligands bound *E. coli* MTHFR (PDB 1ZPT) was used as the starting material for generating NADPH. A phosphate group was built onto the NADH at the existing atom-by-atom using the UCSF Chimera program.[29] The 2'-OH hydrogen of the adenosyl ribose was selected (Ctrl-click on it) and was changed to P, 4 bonds, tetrahedral in the modify structure panel under Tools and Structure Editing. All three hydrogens on P was selected and then modified to O, 1 bond.

Analysis of the ligand binding

The model was superimposed on the ligands bound MTHFR structures (PDB 3APT, 1ZPT, and 1ZP4) using the UCSF Chimera program.[29] The H-bonds occurred between the model and ligands were predicted.

Results and Discussion

The similarity of bacterial and mammalian MTHFRs

MTHFRs from *E. coli* (eMTHFR) and *T. thermophilus* (tMTHFR) were selected for this study because the biochemical and structural properties of these bacterial proteins are well characterized. Both eMTHFR and tMTHFR consist of a catalytic domain with the molecular weight of ~34 kDa per subunit (296 amino acids). The hMTHFR is a large protein with the calculated molecular weight of ~75 kDa per subunit (656 amino acids).[18,30] Unlike the bacterial enzymes, mammalian MTHFRs compose of the N-terminal catalytic domain and the C-terminal domain for the allosteric regulation.[31,32] According to porcine MTHFR, the consensus KRREED is a region with the highest hydrophilicity and surface probability for the possible cleavage point between the N- and the C-terminal domains.[31] In hMTHFR, the cleavage between K356 and R357 generates the N-terminal 40 kDa and the C-terminal 35 kDa fragments. The sequence alignment showed that the catalytic domain of hMTHFR was similar to bacterial proteins with 39% and 33% identities to tMTHFR and eMTHFR, respectively (Figure 2).

Homology model of hMTHFR

In mammalian MTHFRs, the C-terminal domain plays a role in the regulation of enzyme activity upon the AdoMet/adenosylhomocyeteine ratio in biological system.[33] The allosteric binding of AdoMet causes a conformational change that decreases the affinity for NADPH substrate.[33,34] The limited proteolysis of the native porcine MTHFR into the N- and the C-terminal fragments results in the loss of inhibition by AdoMet, but the enzyme activity still remains.[31] These imply that the only catalytic domain of mammalian MTHFRs is active. The first 356 amino acids sequence of hMTHFR was then used as the target (or input) sequence for generating the homology model. The total 50 searched templates were ranked according to the GMQE score that indicates the accuracy of the tertiary structure of the resulting models. The crystal structure of tMTHFR (PDB 3APT),[25] which is an FAD-bound homodimer with 1.85 Å X-ray diffraction, was the best template as shown the highest GMQE score of 0.62 and the QSQE score of 0.40. In fact, only chain A of the tMTHFR structure (3apt.1.A) has the complete sequence with 38.33% identity to the hMTHFR target sequence. The model shown in gold was then built to be a homodimer in which each subunit was constructed relatively to the 3apt.1.A shown in blue (Figure 3A). The alignment showed that the model obtained similar folds with 1.689 Å RMSD to the tMTHFR dimer (Figure 3B).

Figure 2. Sequence alignment of MTHFRs. The eMTHFR, tMTHFR, hMTHFR represented the amino acid sequences of MTHFR from *E. coli*, *T. thermophilus*, and human, respectively. The box indicated the alignment of the catalytic domain of hMTHFR and the whole sequences of eMTHFR and tMTHFR. The arrow assigned the cleavage site of the KREED consensus sequence (underlined letters).

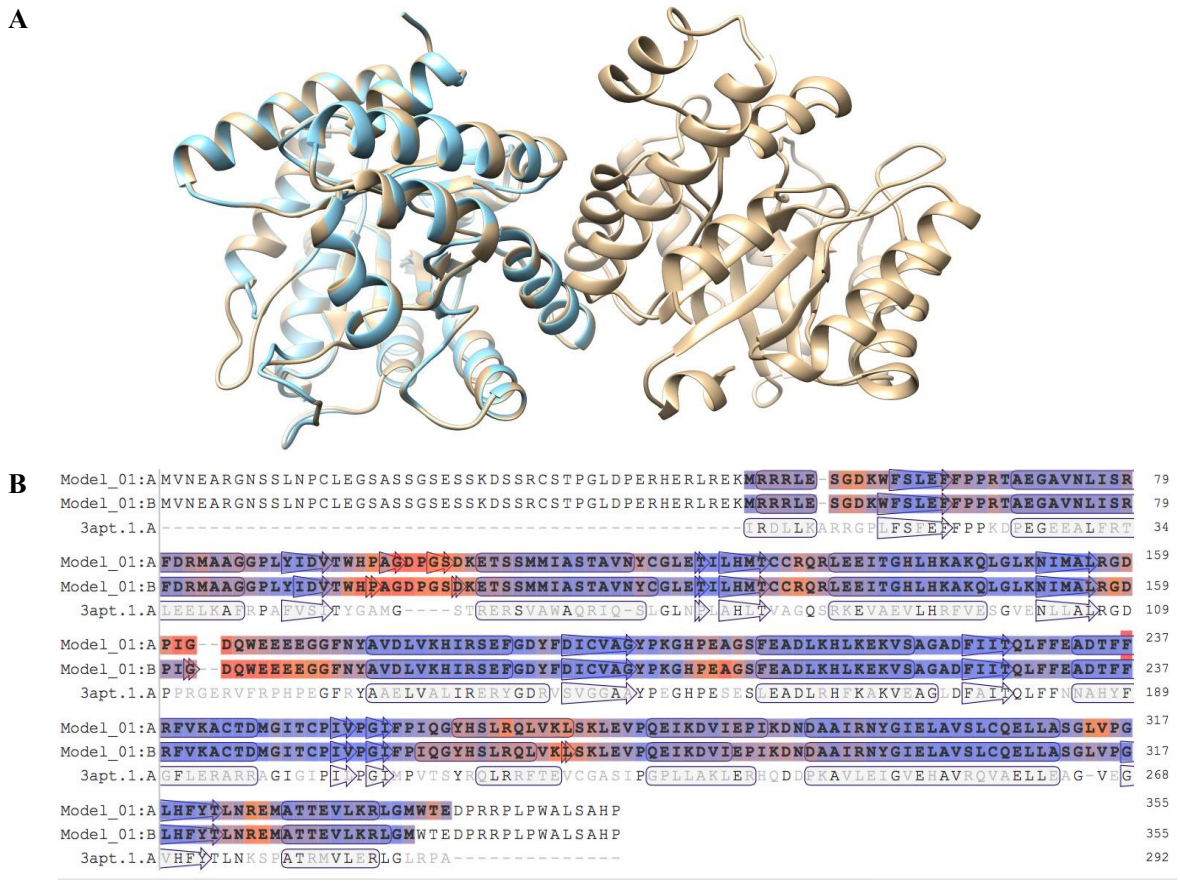

Figure 3. Homology model of hMTHFR. (A) The superimposition of the hMTHFR model (gold) on the tMTHFR chain A (3apt.1.A) template (blue). (B) The model-template alignment. The plain boxes and the arrow boxes indicate the locations of α-helixes and β-sheets, respectively.

Predicted interactions of FAD, NADPH and CH_3-H_4folate

The superimposition of the hMTHFR model on the FAD-bound tMTHFR template (PDB 3APT)[35] showed that FAD adopted a straight conformation where it formed H-bonds with backbone residues F95, R157, D159, A175 and with the side chain residues H213 and Y197. These interactions were also found in the FAD-bound tMTHFR structure. Unlike S156 in tMTHFR, the interaction between the side chain A204 and 3'–OH of the adenosyl ribose was not detected in the model (Figure 4A). The isoalloxazine ring of FAD located at the center of the catalytic binding pocket and turned its *si*-face toward NADPH and CH_3-H_4folate (Figure 4B and 4D).

To analyze the binding of NAD(P)H, the hMTHFR model was superimposed on the eMTHFR structure in complex with FAD and NADH (PDB 1ZPT).[35] The NADPH that was built relatively to the hairpin-like NADH formed a H-bond with Q228 (Q183 in eMTHFR) (Figure 4B). Unlike bacterial MTHFRs that have the conserved phenylalanine, a glutamine at the position 267 played a role in the stacking of NADPH at the end of the pocket. Surprisingly, there was no sign of specificity for the phosphate group of NADPH. The residues surrounding the phosphate group in hMTHFR were mostly identical to which were found in eMTHFR. None of these residues formed H-bond (Figure 4C). In addition, the negative charge of E63 would interfere the binding of the phosphate. These raised the mystery of NADPH affinity in the hMTHFR model.

Currently, the CH_2-H_4folate bound MTHFR structure has not been solved. The investigation of the folate binding was then performed using the available eMTHFR structure in complex with FAD and CH_3-H_4folate product (PDB 1ZP4).[35] A total of three H-bonds

occurred between the pterin ring of the folate and the conserved D159 and Q228 (Figure 4D). The conformational change in eMTHFR suggested that the possible movement of L323, Q267, and H263 side chains might occurred in order to accommodate the PABA moiety and the glutamate tail of the folate. Upon the movement, H263 in hMTHFR was predicted to form a H-bond with the glutamate tail similarly to Q219 in eMTHFR.

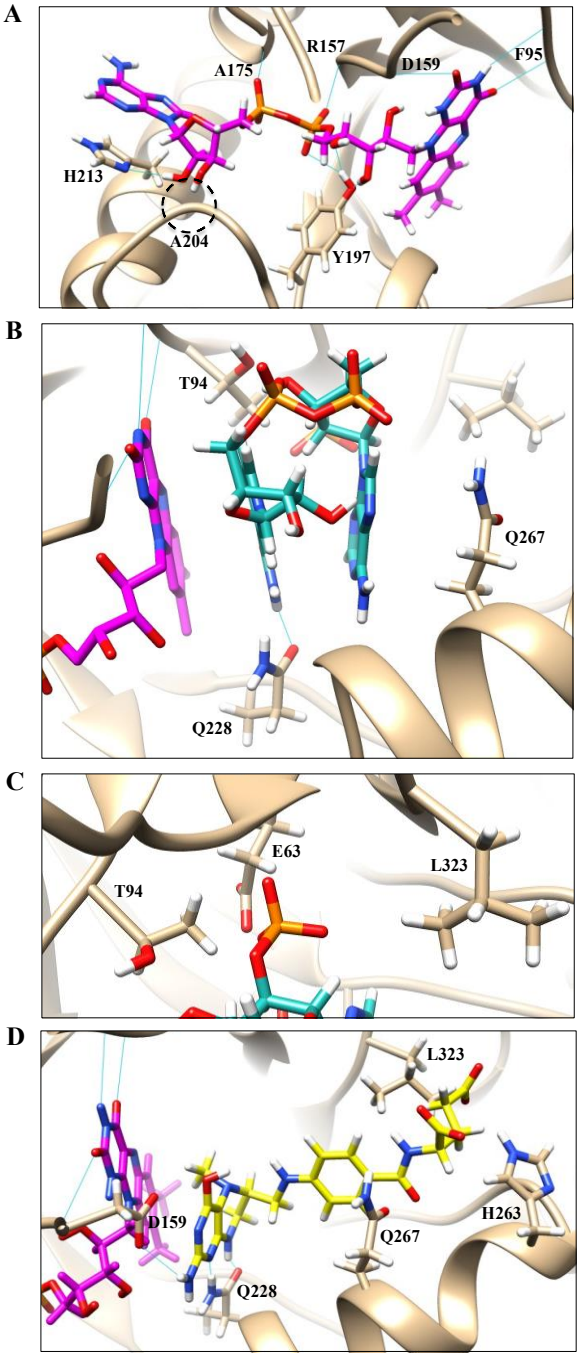

Figure 4. Predicted H-bond interactions of FAD, NADPH and CH_3-CH_4folate to the hMTHFR model (A) The FAD-bound MTHFR structure. (B) The stacking of a hairpin-like NADPH. (C) The predicted conformation of the phosphate. (D) The binding of CH_3-H_4folate.

Conclusion

In general, the predicted bindings of FAD, NADPH and CH_3-H_4folate to the hMTHFR model were almost identical to which occur in the bacterial enzymes, tMTHFR and eMTHFR. The phosphate group of NADPH did not cause the addition of H-bond formation when compared to the NADH. These observations suggest three possible explanations as follows (1) the hMTHFR had no preference for NADPH, (2) By using the FAD-bound tMTHFR structure (PDB 3APT)[25] as a template, the generated model was different from the nature hMTHFR, and (3) The C-terminal regulatory domain of hMTHFR was essential for the preference of NADPH. In addition, these also raised a concern whether using either eMTHFR or tMTHFR as a model system in many studies was a good strategy for analyzing the association of MTHFR polymorphism with human diseases or not. Therefore, it is necessary to repeat the experiments to confirm the affinity for NADH and NADPH in both bacterial and mammalian MTHFRs. It is more reasonable to solve the crystal structures of hMTHFR in complex with and without ligands than focusing on the models.

References

1. Bhattacharyya S, Varshney U. RNA Biol. 2016;19:810-819.
2. Ding W, Smulan L, Hou N, Taubert S, Watts J, Walker A. Metabolism. 2015;22:633-645.
3. Ferla M, Patrick W. Microbiol. Microbiol. 2014;160:1571-1584.
4. Joanne Pigg C, Spence K, Parks L. Arch Biochem Phys. 1962;97:491-496.
5. Saint-Macary M, Barbisan C, Gagey M, Frelin O, Beffa R, Lebrun M, Droux M. PLoS One. 2015;10:e01111108.
6. Hesse H, Hoefgen R. Trends Plant Sci. 2003;8:259-262.
7. Laurichesse H, Tauveron I, Gourdon F, Cormerals L, Champredon C, Charrier S, Rochon C, Lamain S, Bayle G, Laveran H, Thieblot P, Beytout J, Grizard J. J Nutr. 1998;128:1342-1348.
8. Usuda Y, Kurahashi O. Appl Environ Microbiol. 2005;71:3228-3234.
9. Alaminos M, Ramos J. Arch Microbiol. 2001;176:151-154.
10. Wu W, Urbanowski M, Stauffer G. J Bacteriol. 1995;177:1834-1839.
11. Urbanowsky M, Stauffer G. Gene. 1988;73:193-200.
12. Plamann L, Stauffer G. J Bacteriol. 1987;169:3932-3937.
13. Old I, Hunter M, Wilson D, Knight S, Weatherston C, Glass R. Mol Gen Genet. 1988;211:78-87.
14. Blanco J, Coque J, Martin J. J Bacteriol. 1998;180:1586-1591.
15. Shappard C, Trimmer, E, Matthews R. J Bacteriol. 1999;181:718-725.
16. Yan X, Que Y, Wang H, Wang C, Li Y, Yue X, Ma Z, Talbot N, Wang Z. PLoS One. 2013;8:e76914.
17. Vickers T, Orsomando G, de la Garza R, Scott D, Kang S, Hanson A, Beverley S. J Biol Chem. 2006;281:38150-38158.
18. Frosst P, Blom H, Milos R, *et al*. Nat Genet. 1995;10:111-113.
19. Refsum H, Ueland P, Nygard O, Vollset S. Annu Rev Medicine. 1998;49:31-62.
20. Seshadri S, Beiser A, Selhub J, Jacques P, Rosenberg I, D'Agostino R, Wilson P, Wolf P. N Engl J Med. 2002;346:476-483.
21. Daubner S, Matthew R. J Biol Chem. 1982;257:140-145.
22. Katzen H, Buchanan J. J Biol Chem. 1965;240:825-835.
23. Pejchal R, Campbell E, Guenther B, Lennon B, Matthews R, Ludwig M. Biochem. 2006;45:4808-4818.
24. Lee M, Takawira D, Nikilova A, Ballou D, Furtado V, Phung N, Still B, Thorstad M, Tanner J, Trimmer E. Biochem. 2009;48:7673-7685.

25. Igari S, Ohtaki A, Yamanaka Y, Sato Y, Yohda M, Odaka M, Noguchi K, Yamada K. PLoS One;6(8):e23716.
26. Altschul S, Gish W, Miller W, Myers E, Lipman D. J Mol Biol. 1990;215:403-410.
27. Sievers F, Higgins D. Methods Mol Biol. 2014;1079:105-116.
28. Guex N, Peitsch M, Schwede T. Electrophoresis. 2009;30;S162-S173.
29. Petterson E, Goddard T, Huang C, Couch G, Greenblatt D, Meng E, Ferrin T. J Comput Chem. 2004;25:1605-1612.
30. Matthews R, Vanoni M, Hainfeld J, Wall J. J Biol Chem. 1984;259:11647-11650.
31. Sumner J, Jencks D, Khani S, Matthews R. J Biol Chem. 1986;261:7697-7700.
32. Biasini M, Bienert S, Waterhouse A, *et al*. Nucleic Acids Res. 2014;42:W252-W258.
33. Madden T, Tatusov R, Zhang J. Methods Enzymol. 1996;266:131-141.
34. Kutzbach C, Stokstad E. Biochem Biophys Acta. 1971;250:459-477.
35. Pejchal R, Sargeant R, Ludwig M. 2005;44:11447-11457.