Crystal Structure of MJ0684 from *Methanococcus jannaschii*, a Novel Archaeal Homolog of Kynurenine Aminotransferase

Jin Kuk Yang

Department of Chemistry, Soongsil University, Seoul 156-743, Korea. E-mail: jinkukyang@ssu.ac.kr
Received September 30, 2007

MJ0684 from *Methanococcus jannaschii* is a hypothetical protein belonging to the subfamily Iγ of amino acid aminotransferases. In the present study, the crystal structure of MJ0684 has been determined at 2.2 Å resolution. It reveals that MJ0684 has an overall structure similar to subfamily Iγ aminotransferases and its active site architecture is most similar to that of kynurenine aminotransferases among several kinds of aminotransferases in the subfamily Iγ. It has two hydrophobic active site residues conserved in the kynurenine aminotransferases for recognizing hydrophobic substrates. In addition, the absence of any basic residue for recognizing the side chain carboxylic group of the aspartate in the active site rules out the possibility that MJ0684 would act as an aspartate aminotransferase. These structural observations collectively imply that MJ0684 is a novel archaeal homolog of the subfamily Iγ kynurenine aminotransferase.

Key Words : Amino acid aminotransferase, Kynurenine aminotransferase, MJ0684

Introduction

Pyridoxal 5'-phosphate (PLP)-dependent amino acid aminotransferase (AT) is a homodimeric enzyme which catalyzes the reversible transfer of an α-amino group from an α-amino acid to an α-keto acid.[1] ATs have been classified into four families, I to IV, based on the homology relationships in the primary sequence.[2] Family I, the most extensively studied one, can be further divided into seven subfamilies: Iα, Iβ, Iγ, and others.[3] Subfamily Ia mostly consists of aspartate aminotranserases (AspATs) and aromatic amino acid aminotransferases (AroATs). Subfamily Iβ includes histidine aminotransferases, some of which shows the activity toward aromatic amino acids as well as histidine. Subfamily Iγ includes AroATs such as kynurenine AT (KynAT) and tyrosine AT (TyrAT), and some set of AspATs.[3]

Structural studies on a number of ATs have presented comprehensive observations on their various strategies for recognition of various amino acids as substrates. For example, subfamily Iα AspATs from *E. coli*,[4] chicken,[5,6] pig,[7] and yeast[8] have a conserved arginine residue to recognize the side chain carboxylic group of the substrate aspartate (R292* in pig cytosolic AspAT, for instance; asterisk denotes the residue from the other subunit in dimer). In contrast, AspATs in the subfamily Iγ don't have the arginine residue in the corresponding position in the primary sequence, and instead use a lysine residue (K101 in AspAT from *Thermus thermophilus* HB8, for instance) for the same purpose.[9]

MJ0684 is a 370-residue protein encoded by the *Mj0684* (*aspB2*) gene from *Methanococcus jannaschii*. Sequence homology analysis using BLAST[10] indicates that MJ0684 is a member of AT subfamily Iγ. Among the subfamily Iγ members, it shows the highest sequence similarity to AspATs (34-41% identity in 330-370 residues overlap), and then next to AroATs such as KynATs and TyrATs (25-30% identity in 280-300 residues overlap). So BLAST results suggest that MJ0684 is most likely an AspAT in the subfamily Iγ. However, MJ0684 lacks the otherwise conserved lysine residue that plays a critical role in recognizing the side chain carboxylic group of the aspartate as a putative substrate. Thus, it may prefer other substrates with hydrophobic character such as kynurenine, tyrosine, phenylalanine and tryptophan rather than the aspartate, or it may have a different active site basic residue playing the same role of K101 of the AspAT from *Thermus thermophilus* HB8. MJ0684 structure in comparison with previously reported structures of ATs may be able to present some novel structural observations to answer this puzzling question.

Here I report the crystal structure of MJ0684 in complex with the cofactor PLP at 2.2 Å resolution. Its overall structure is more similar to AspAT than AroATs such as KynAT and TyrAT, but the details of the active site architecture imply that it may prefer hydrophobic amino acids as substrates. There is no basic residue in the active site for recognizing the side chain carboxylate of the putative substrate aspartate, and there are found the two hydrophobic residues that are conserved in KynATs for recognizing hydrophobic amino acids as substrates.

Methods

Protein expression, purification, crystallization and data collection from a native crystal have been described previously.[11] Molecular replacement calculation tried right after the native data collection was not successful, so soon Hg-derivative crystal was prepared by adding 0.2 uL of 1 mM $Hg(OAc)_2$ into 4 uL of the hanging drop. After soaking for 2 hrs in this condition, the crystal was mounted and X-ray diffraction data were collected in the same way used for the native crystal described previously.[11] The native data set was collected to 2.2 Å and the derivative to 2.5 Å. The structure

was solved by the single isomorphous replacement with anomalous scattering (SIRAS) approach. Total 4 Hg sites were located using SOLVE,[12] and the phases were refined using SHARP.[13] After density modification with SOLOMON,[14] an interpretable electron-density map was obtained for the space group $P4_32_12$ among two enantiomeric space groups. Model building and refinement were completed using the programs O[15] and CNS.[16] Atomic coordinates and structure factors were deposited in Protein Data Bank under an accession code 2Z61.

Results and Discussion

Crystal structure of MJ0684 has been determined and refined to R_{free} of 23.5% at a resolution of 2.2 Å. Asymmetric unit contains a single molecule of MJ0684 with a PLP covalently linked to Lys222 and 87 water molecules. The functional dimer can be generated by crystallographic two-fold symmetry operation. The monomer model contains 369 residues of the entire 370-residue-long chain except for the last residue K370. Ramachandran plot produced by PROCHECK[17] shows that all the residues except for only one residue, Ile253, are in the allowed region. Statistics for the data collection and the structure refinement are summarized in Table 1. The overall architecture adopts the type I fold of PLP-dependent enzymes, which is characterized by an *N*-terminal arm and two domains, one small and the other large, as observed in other ATs except for the branched-chain amino acid ATs and D-amino acid ATs.[18] The *N*-terminal arm (residues 1-11) consists of one 3_{10} helix and random coils, and the small domain (residues 12-43 and 269-369) contains four β-strands and four α-helices. The large domain (residues 44-268) adopts an $\alpha/\beta/\alpha$ structure in which a seven-stranded β-sheet is surrounded by nine α-helices (Fig. 1A). MJ0684 forms a homodimer in a manner characteristic of PLP-dependent ATs with the type I fold (Fig. 1B).

Sequence similarity search using BLAST[10] indicates that MJ0684 is a member of AT subfamily Iγ which includes AspAT, KynAT and TyrAT.[3] Among several kinds of ATs in the subfamily Iγ, MJ0684 shows the highest sequence similarity to AspATs such as AspAT (TM1255) from *Thermotoga maritima* (tmAspAT: 41% identical in 362-residue overlap)[19] and AspAT from *Thermus thermophilus* HB8 (ttAspAT: 37% identical in 369-residue overlap).[9] In addition, MJ0684 also shows high sequence similarity to AroATs such as KynAT from mouse (33% identical in 290-residue overlap) and TyrAT from *Trypanosoma cruzi* (tcTyrAT: 29% identical in 342-residue overlap).[20] Structural similarity search using DALI server[21] also shows that MJ0684 is the most similar to AspATs such as ttAspAT[9] (rmsd 1.5 Å in 363-residue overlap; PDB code 1BJW), and then next to AroATs such as KynAT from *Aedes aegypi*[22] (rmsd 2.2 Å in 347-residue overlap; PDB code 1YIY) and tcTyrAT[20] (rmsd 2.2 Å in 356-residue overlap; PDB code 1BW0). Although the similarity search in the primary sequence and tertiary structure commonly indicates MJ0684 is most likely a family Iγ AspAT, MJ0684 lacks the otherwise conserved lysine residue (for instance, K101 of ttAspAT)[9] which recognizes the carboxylate in the side chain of the putative

Table 1. Data Collection and Model Refinement

	Native	$Hg(OAc)_2$
Unit cell: $a = b, c$ (Å)	111.87, 60.86	112.01, 60.91
Resolution (Å)	50.0-2.2	50.0-2.5
Completeness (%)	99.6 (98.9)	84.4 (89.4)
$I/\sigma(I)$	36.1 (4.3)	29.1 (6.1)
R_{merge}[a]	7.0 (40.7)	6.6 (32.1)
Observed reflections	120,856	48,993
Unique reflections	20,121	11,773
R_{cryst}[b] / R_{free}[c] (%)	21.4 / 23.5	
Rmsd bond lengths (Å)	0.008	
Rmsd bond angles (°)	1.85	
Average B-factor	39.6	
Protein atoms	39.2	
Cofactor (PLP)	38.9	
Waters	50.2	

Values in parenthesis are for the outer resolution shell. [a]$R_{merge} = \Sigma_h\Sigma_i |I(h)_i - \langle I(h)\rangle| / \Sigma_h\Sigma_i I(h)_i$, where I(h) is the intensity of reflection h, Σ_h is the sum over all reflections, and Σ_i is the sum over i measurements of reflection h. [b]$R_{cryst} = \Sigma ||F_{obs}| - |F_{calc}|| / \Sigma |F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factor, respectively. [c]10% of the data was set aside for R_{free} calculation.

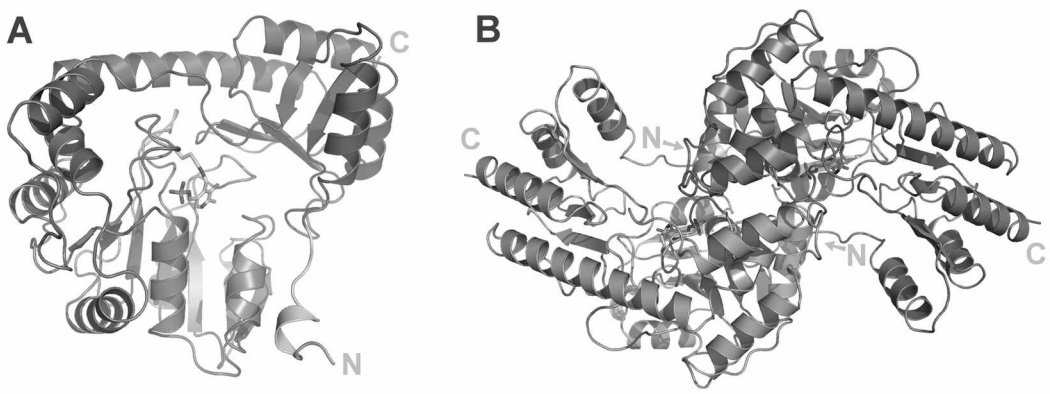

Figure 1. (A) Overall fold of a MJ0684 monomer. PLP is shown in a stick model. N-terminal arm is colored in yellow, the small domain in purple, and the large domain in blue. (B) MJ0684 dimer. One subunit in dimer is colored in purple and the other in blue.

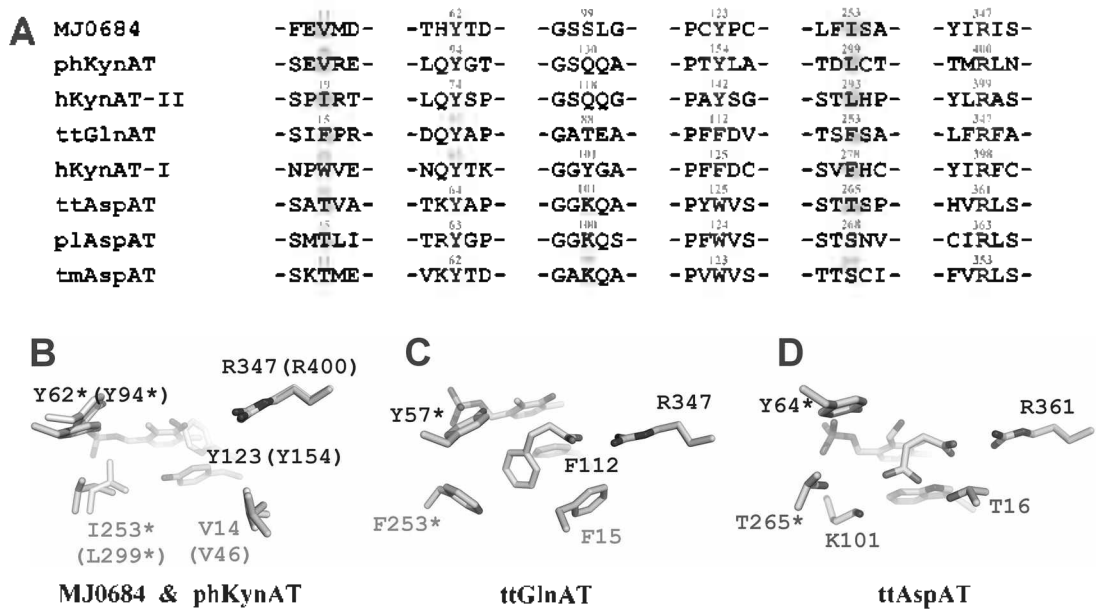

Figure 2. (A) Structure-based primary sequence alignment. MJ0684 is aligned with family Iγ KynATs and AspATs for the key residues in substrate recognition and PLP binding. Residues in red and blue boxes are critical for substrate recognition in KynATs and AspATs respectively. In cyan boxes are the residues conserved commonly in KynATs and AspATs. phKynAT is for KynAT from *Pyrococcus horikoshii* which is a bacterial homolog of human KynAT-II, hKynAT-II for human KynAT-II, ttGlnAT for GlnAT from *Thermus thermophilus* HB8 which is a bacterial homolog of human KynAT-I, hKynAT-I is for human KynAT, ttAspAT for AspAT from *Thermus thermophilus* HB8, plAspAT for AspAT from *Phormidium lapideum*, and tmAspAT for AspAT from *Thermotoga maritima*. (B) Superposition of MJ0684 and phKynAT. MJ0684 (PDB code 2Z61) is colored in green and phKynAT (PDB code 1X0M) in light cyan. (C) ttGlnAT in complex with 3-phenylpropionate (PDB code 1V2F). (D) ttAspAT in complex with malate (PDB code 1BKG). In (B)-(D), residues in red and blue are for the substrate recognition in KynATs and AspATs respectively.

substrate aspartate. Notably MJ0684 also lacks the active site arginine residue conserved in the family Iα AspATs for the same role (for example, R292* of pig cytosolic AspAT).[7] In the active site of MJ0684, there is not found any basic residue which might be able to play the same role of the lysine residue in the family Iγ AspATs or the arginine residue in the family Iα AspATs. So the absence of the key residue for recognizing the aspartate strongly raises the possibility that MJ0684 may prefer other substrates, for example aromatic amino acids such as Kyn, Phe, Tyr and Trp as BLAST and DALI search results commonly indicate as the secondly most probable substrates.

In addition to the absence of the lysine or arginine residue for recognizing the side chain carboxylic group of an aspartate, one additional important structural feature is observed in the active of MJ0684 regarding the substrate specificity. That is the presence of two hydrophobic residues conserved in KynATs, not in AspATs, for recognizing aromatic amino acids as substrates. In the crystal structure of the bacterial homolog of human KynAT-I (ttGlnAT; GlnAT from *Thermus thermophilus* HB8; PDB code 1V2F),[23] for example, it has been shown that the phenyl group of the substrate is surrounded by F15, F253*, Y57* and F112 (Fig. 2C). Among those four residues, Y57* and F112 interact with the cofactor PLP also: the former hydrogen bonds to the oxygen in the phosphate group of PLP and the latter interacts with pyridoxal ring of PLP *via* π-π stack. These two residues are strongly conserved through various kinds of ATs regardless of the substrate specificity (Fig. 2A, B, C and D). So the major role for Y57* and F112 may be to coordinate the cofactor. Therefore the key residues for recognizing aromatic amino acid substrates should be the other two, *i.e.* F15 and F253* (Fig. 2C). These two hydrophobic residues are conserved in other KynATs such as human KynAT-I (W18 & F278*; PDB code 1W7M; Fig. 2A)[24] and a bacterial homolog of human KynAT-II (V46 & L299*; PDB code 1X0M; KynAT from *Pyrococcus horikoshii*; Fig. 2A and B).[25] Just as in the active site of these KynATs, MJ0684 also has hydrophobic residues in the same position which are V14 and I253* (Fig. 2A and B). In contrast, AspATs do not have hydrophobic ones in equivalents positions. In ttAspAT, for example, those two residues are T16 and T265*, which participate in recognizing the distal carboxylic group of the substrate (Fig. 2A and D). T16 directly forms a hydrogen bond with the distal carboxylic group and T265 forms a hydrogen bond with K101 which directly interact with the distal carboxylic group. These two threonine residues are conserved in other family Iγ AspATs such as AspAT from *Phormidium lapideum* (plAspAT; PDB code 1J32)[26] and tmAspAT[19] (Fig. 2A). In summary, two residues in these positions may be the determinants of the substrate specificity, at least differentiating between KynATs and AspATs. It looks obvious from the observation that there are hydrophobic ones in those positions in KynATs, but threonine or serine in AspATs (Fig. 2A). Therefore, the presence of V14 and I253* in the same positions of MJ0684 strongly

implies that MJ0684 would prefer aromatic amino acids to dicarboxylic amino acids as substrates, just as KynATs do. In addition, it is noteworthy that, among two types of KynATs, MJ0684 resembles more KynAT-II than KynAT-I as shown in Figure 2A, 2B and 2C. Human KynAT-II and its bacterial homolog from *Pyrococcus horikoshii* (phKynAT) have I19-L293* and V46-L299* respectively just as MJ0684 has V14-I253* for substrate recognition (Fig. 2A and 2B). Meanwhile, human KynAT-I and its bacterial homolog from *Thermus thermophilus* HB8 (ttGlnAT) have W18-F278* and F15-F253* in the same positions (Fig. 2A and 2C).

In conclusion, the crystal structure of MJ0684 reveals that it doesn't have any basic residue in the active site for recognizing the side chain carboxylic group of the aspartate as a putative substrate. Moreover MJ0684 has two hydrophobic active site residues conserved in the subfamily Iγ KynATs, not in AspATs, for recognizing hydrophobic substrates. These two lines of structural observation in the active site strongly imply that MJ0684 would prefer hydrophobic amino acids to the aspartic acid as a substrate. MJ0684 is a novel example of AT whose overall structure and sequence are more similar to AspATs than any other ATs but the local structure of the active site is even more similar to KynATs than AspATs. However, it remains to be shown that the novel structural observation on MJ0684 may comply with the actual substrate specificity determined by the enzymatic assay experiment.

Acknowledgments. The author thanks Dr. Se Won Suh for supporting this work in all aspects, Dr. Soo Hyun Eom for sharing X-ray machine, and Dr. Changsoo Chang, Dr. Seung-je Cho, Dr. Jae Young Lee for invaluable comments in the course of structure analysis. This work was supported by the Soongsil University Research Fund to JKY.

References

1. Hirotsu, K.; Goto, M.; Okamoto, A.; Miyahara, I. *Chem. Rec.* **2005**, *5*, 160.
2. Mehta, P. K.; Hale, T. I.; Christen, P. *Eur. J. Biochem.* **1993**, *214*, 549.
3. Jensen, R. A.; Gu, W. *J. Bacteriol.* **1996**, *178*, 2161.
4. Jager, J.; Moser, M.; Sauder, U.; Jansonius, J. N. *J. Mol. Biol.* **1994**, *239*, 285.
5. McPhalen, C. A.; Vincent, M. G.; Jansonius, J. N. *J. Mol. Biol.* **1992**, *225*, 495.
6. Malashkevich, V. N.; Strokopytov, B. V.; Borisov, V. V.; Dauter, Z.; Wilson, K. S.; Torchinsky, Y. M. *J. Mol. Biol.* **1995**, *247*, 111.
7. Rhee, S.; Silva, M. M.; Hyde, C. C.; Rogers, P. H.; Metzler, C. M.; Metzler, D. E.; Arnone, A. *J. Biol. Chem.* **1997**, *272*, 17293.
8. Jeffery, C. J.; Barry, T.; Doonan, S.; Petsko, G. A.; Ringe, D. *Protein Sci.* **1998**, *7*, 1380.
9. Nakai, T.; Okada, K.; Akutsu, S.; Miyahara, I.; Kawaguchi, S.; Kato, R.; Kuramitsu, S.; Hirotsu, K. *Biochemistry* **1999**, *38*, 2413.
10. Altschul, S. F.; Madden, T. L.; Schaffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J. *Nucleic Acids Res.* **1997**, *25*, 3389.
11. Yang, J. K.; Chang, C.; Cho, S. J.; Lee, J. Y.; Yu, Y. G.; Eom, S. H.; Suh, S. W. *Acta Crystallogr. D Biol. Crystallogr.* **2003**, *59*, 563.
12. Terwilliger, T. C. *Methods Enzymol.* **2003**, *374*, 22.
13. Bricogne, G.; Vonrhein, C.; Flensburg, C.; Schiltz, M.; Paciorek, W. *Acta Crystallogr. D Biol. Crystallogr.* **2003**, *59*, 2023.
14. Abrahams, J. P.; Leslie, A. G. *Acta Crystallogr. D Biol. Crystallogr.* **1996**, *52*, 30.
15. Jones, T. A.; Zou, J. Y.; Cowan, S. W.; Kjeldgaard, M. *Acta Crystallogr. A* **1991**, *47*, 110.
16. Brunger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J. S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L. *Acta Crystallogr. D Biol. Crystallogr.* **1998**, *54*, 905.
17. Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. *J. Appl. Crystallogr.* **1993**, *26*, 283.
18. Grishin, N. V.; Phillips, M. A.; Goldsmith, E. J. *Protein Sci.* **1995**, *4*, 1291.
19. Schwarzenbacher, R.; Jaroszewski, L.; von Delft, F.; Abdubek, P.; Ambing, E.; Biorac, T.; Brinen, L. S.; Canaves, J. M.; Cambell, J.; Chiu, H. J.; Dai, X.; Deacon, A. M.; DiDonato, M.; Elsliger, M. A.; Eshagi, S.; Floyd, R.; Godzik, A.; Grittini, C.; Grzechnik, S. K.; Hampton, E.; Karlak, C.; Klock, H. E.; Koesema, E.; Kovarik, J. S.; Kreusch, A.; Kuhn, P.; Lesley, S. A.; Levin, I.; McMullan, D.; McPhillips, T. M.; Miller, M. D.; Morse, A.; Moy, K.; Ouyang, J.; Page, R.; Quijano, K.; Robb, A.; Spraggon, G.; Stevens, R. C.; van den Bedem, H.; Velasquez, J.; Vincent, J.; Wang, X.; West, B.; Wolf, G.; Xu, Q.; Hodgson, K. O.; Wooley, J.; Wilson, I. A. *Proteins* **2004**, *55*, 759.
20. Blankenfeldt, W.; Nowicki, C.; Montemartini-Kalisz, M.; Kalisz, H. M.; Hecht, H. J. *Protein Sci.* **1999**, *8*, 2406.
21. Holm, L.; Sander, C. *Methods Enzymol.* **1996**, *266*, 653.
22. Han, Q.; Gao, Y. G.; Robinson, H.; Ding, H.; Wilson, S.; Li, J. *FEBS J.* **2005**, *272*, 2198.
23. Goto, M.; Omi, R.; Miyahara, I.; Hosono, A.; Mizuguchi, H.; Hayashi, H.; Kagamiyama, H.; Hirotsu, K. *J. Biol. Chem.* **2004**, *279*, 16518.
24. Rossi, F.; Han, Q.; Li, J.; Li, J.; Rizzi, M. *J. Biol. Chem.* **2004**, *279*, 50214.
25. Chon, H.; Matsumura, H.; Koga, Y.; Takano, K.; Kanaya, S. *Proteins* **2005**, *61*, 685.
26. Kim, H.; Nakaoka, M.; Yagi, M.; Ashida, H.; Hamada, K.; Shibata, H.; Sawa, Y. *J. Biosci. Bioeng.* **2003**, *95*, 421.