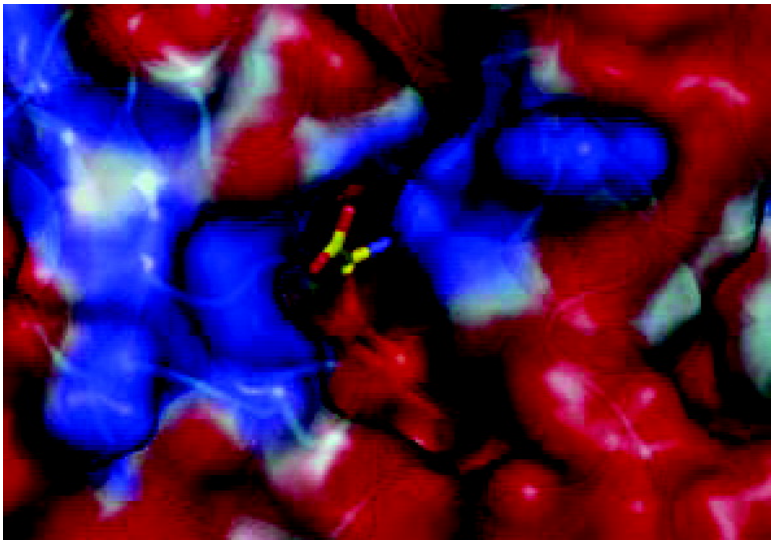
Communication

Structural Characterization of a *p*-Acetylphenylalanyl Aminoacyl-tRNA Synthetase

James M. Turner, James Graziano, Glen Spraggon, and Peter G. Schultz

J. Am. Chem. Soc., **2005**, 127 (43), 14976-14977• DOI: 10.1021/ja0549042 • Publication Date (Web): 06 October 2005



J|A|C|S
COMMUNICATIONS

Published on Web 10/06/2005

Structural Characterization of a *p*-Acetylphenylalanyl Aminoacyl-tRNA Synthetase

James M. Turner,[†] James Graziano,[†] Glen Spraggon,*[,‡] and Peter G. Schultz*[,†]

Department of Chemistry and the Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, and Genomics Institute of the Novartis Research Foundation, 3115 Merryfield Row, San Diego, California 92121

Received July 21, 2005; E-mail: gspraggon@gnf.org; schultz@scripps.edu

We have shown that the number and nature of genetically encoded amino acids can be expanded in both prokaryotes and eukaryotes by the use of orthogonal aminoacyl-tRNA synthetase (RS)/tRNA pairs and unique three- and four-base codons.[1−3] In one example, a *Methanococcus jannaschii* (*M.j.*) tyrosyl tRNA synthetase (TyrRS) was evolved, which together with a cognate *M.j.* tRNA can incorporate the amino acid *p*-acetylphenylalanine (pAcPhe) into proteins in *Escherichia coli* with good efficiency and high fidelity in response to the nonsense codon TAG. Five amino acids (Y32, D158, I159, L162, and A167) in the active site of the *M.j.* TyrRS were randomized, and alternating rounds of positive and negative selection were used to identify mutant synthetases that aminoacylate pAcPhe but do not accept endogenous host amino acids. The pAcPhe TyrRS had the following mutations: Y32L, D158G, I159C, and L162R.[4] To begin to determine the structural basis for the surprising evolvability of aminoacyl-tRNA synthetase specificity, we have determined the crystal structure of the pAcPhe-specific tyrosyl *M.j.* TyrRS.

The pAcPhe synthetase was expressed in *E. coli* with a C-terminal His_6 fusion tag. Crystals were grown in the presence of *p*-acetylphenylalanine (2 mM) and belong to space group $P4_32_12$, containing one molecule per asymmetric unit. The structure of a pAcPhe TyrRS/*p*-acetylphenylalanine complex was solved to 2.5 Å (Table 1) with $R_{cryst} = 0.21$ and $R_{free} = 0.28$. Like the wild-type *M.j.* TyrRS,[5,6] the pAcPhe synthetase is composed of five regions: the Rossmann-fold catalytic domain, the short N-terminal region, the connective polypeptide 1 (CP1) domain, the C-terminal domain, and the KMSKS loop which links the Rossmann-fold to the C-terminal domain. Although the overall structure of the AcPhe TyrRS is nearly identical to the wild-type synthetase structure, there are significant changes within the active site that result in specific recognition of the unnatural amino acid pAcPhe.

These mutations lead to both altered hydrogen bonding and packing interactions with the bound substrate. The D158G and Y32L mutations remove two hydrogen bonds to the tyrosine side-chain hydroxyl group, which would be expected to dramatically reduce binding of the natural substrate to the enzyme (Figure 1). The D158G mutation also deepens the binding pocket to accommodate the para substituent of *p*-acetylphenylalanine, while Y32L forms a suitable hydrophobic packing surface for the acetyl methyl group (Figure 1b). In addition, the side-chain carbonyl oxygen of pAcPhe forms a hydrogen bond to Gln109 Nϵ (pAcPhe carbonyl O/Gln 109 Nϵ distance = 3.2 Å). In the wild-type *M. jannaschii* TyrRS structure,[5] the side chain of Asp158 disrupts this hydrogen bond vector (Figure 1b); in the pAcPhe-specific synthetase the D158G mutation removes this intervening side chain. Interestingly, extra electron density was observed for the side chain of I159C in the structure of the pAcPhe synthetase, suggestive of disulfide formation within the crystallized protein. The only free thiol present in the crystallization conditions, β-mercaptoethanol, was modeled into this density to form a disulfide bond with I159C, as well as a hydrogen bond with the backbone nitrogen of Gly34 (3.3 Å distance). Although disulfide formation between I159C and β-mercaptoethanol is observed in the crystal structure, this is likely not representative of the oxidation state of this residue in the reducing environment of *E. coli* cytoplasm. Indeed, residue 159 is the most variable among the resulting hits from the selections,[4] suggesting that a cysteine at this position is not necessary for amino acid recognition; for example, another pAcPhe-specific synthetase[4] had mutations identical to those of this synthetase, except for I159T instead of I159C. Finally, the structure of the wild-type *M.j.* TyrRS[5,6] also shows that L162 is distal to the active site, suggesting that the L162R mutation was selected for favorable solvent interactions rather than substrate recognition.

The D158G mutation also alters the backbone structure of helix α8. The glycine residue truncates the C-terminus of helix α8 by four residues (Figure 1a), resulting in a new C-terminal cap (C-cap). Studies of amino acid preferences at specific locations within a helix have shown glycine to be strongly preferred as a C-cap for α-helices (used in one-third of the helices surveyed).[7] Several intrahelical main-chain hydrogen bonds are broken by premature termination of helix α8 (Gln155O−Ile159NH, Val156O−His160NH, Asn157O−Tyr161NH, Gly158O−Arg162NH, and Cys159O−Gly163NH). However, a short 3_{10}-helix is formed in the new

Table 1. Data and Refinement Statistics for the pAcPhe TyrRS/*p*-Acetylphenylalanine Crystal Complex

space group	$P4_32_12$
unit cell parameters (Å)	$a = b = 103.4$, $c = 70.8$
wavelength (Å)	1.0
resolution range (Å)	70.0−2.5
R_{symm} (highest resolution shell)	0.065 (0.557)
No. unique reflections	15559
completeness (%) (highest shell)	99.4 (99.8)
highest resolution shell (Å)	2.56−2.5
mean $I/\sigma(I)$	15.9 (2.3)
refinement parameters	
No. of reflections (total)	12978
No. of reflections (test)	700
R_{cryst}[a] (R_{free}[b])	0.21 (0.28)
No. of protein atoms	2462
No. of heterogen atoms	23
No. of water atoms	80
rmsd bonds (Å)	0.012
rmsd angles (deg)	1.40
ave isotropic B-value (Å^2)	37.3

[a] $R_{cryst} = \sum\sum|I_i - \langle I_i\rangle|/\sum|I_i|$ where I_i is the scaled intensity of the ith measurement, and $\langle I_i\rangle$ is the mean intensity for that reflection. [b] R_{free} = same as that for R_{cryst}, but for 5.0% of the total reflections chosen at random and omitted from refinement.

[†] The Scripps Research Institute.
[‡] Genomics Institute of the Novartis Research Foundation.

10.1021/ja0549042 CCC: $30.25

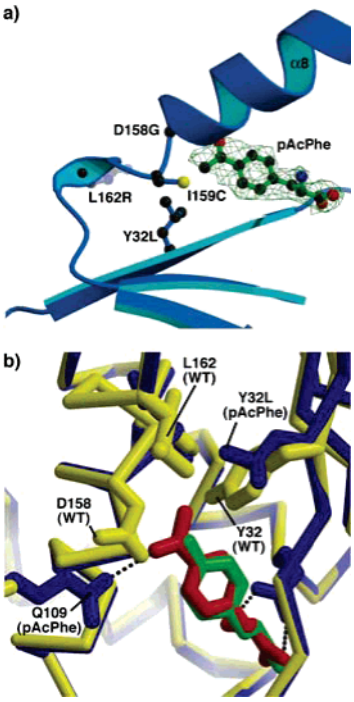


Figure 1. (a) Structure of the pAcPhe synthetase bound to *p*-acetylphenylalanine (pAcPhe). Residues Y32L, D158G, I159C, and L162R are shown. A σ-weighted $F_o - F_c$ electron density map (contoured at 3σ) for the pAcPhe amino acid is also shown. The electron density map was derived prior to the inclusion of the pAcPhe. (b) Structure of the pAcPhe synthetase (blue) bound to *p*-acetylphenylalanine (red) superimposed on the wild-type synthetase[5] (yellow) bound to tyrosine (green). Residues Y32, D158, and L162 of the wild-type synthetase and Y32L and Q109 of the pAcPhe synthetase are illustrated.

structure, stabilized by a side-chain/main-chain hydrogen bond (Asn157O−His160Nδ) and a main-chain/main-chain hydrogen bond (Gly158O−Tyr161NH). An additional main-chain/main-chain hydrogen bond (His160O−Gly163NH) forms a turn that connects to sheet β5. A comparison of the wild-type[5] and pAcPhe mutant synthetase structures superimposed at Cα atoms shows that the C-terminal residues of helix α8 adopt conformations in the mutant structure different than the wild-type structure. L162R moves 3.1 Å away from the active site ($L162_{WT}C\alpha/R162_{pAcPhe}C\alpha$ distance = 3.1 Å), presumably to alleviate steric clash with the Y32L mutation, which abrogates tyrosine recognition. As a consequence, Tyr161 moves 3.6 Å in the mutant structure ($Y161_{WT}C\alpha/Y161_{pAcPhe}C\alpha$ distance = 3.6 Å) and points away from Y32L, while His160 moves 2.7 Å ($H160_{WT}C\alpha/H160_{pAcPhe}C\alpha$ distance = 2.7 Å). This rearrangement is likely facilitated by flexibility of two flanking glycine residues (G163 and the engineered D158G). Despite these structural changes within the active site, there is little difference in the overall backbone traces of the wild-type versus pAcPhe mutant structures for the rest of the enzyme (RMSD = 0.648 Å over 297 aligned Cα atoms).

The structural differences between the pAcPhe and wild-type synthetases are more pronounced than those found in previously described mutant aminoacyl-tRNA synthetases that recognize unnatural or unusual amino acids.[8−10] Indeed, mutations that alter enzyme specificity generally have little effect on backbone configuration and are difficult to predict by theoretical methods.[11−15] The structural changes observed in the pAcPhe synthetase are likely due to their position and the method used to generate the libraries. Structural changes on the solvent-exposed α8 helix require minimal compensatory mutations since there are few steric constraints from adjacent protein domains, and bulk water can easily reorganize to solvate the exposed residues. In addition, the use of a focused library of mutated residues in close proximity increases the likelihood of selecting compensatory mutations that stabilize the new conformation.

In conclusion, we have shown that a handful of random mutations alter the substrate specificity of the *M.j*. TyrRS by changing the pattern of hydrogen bonding and packing interactions with bound substrate. These mutations lead to changes in both side-chain and backbone conformation, indicating a high degree of structural plasticity in the active site of the enzyme. We are currently attempting structural studies with the unliganded pAcPhe synthetase to determine if amino acid binding is involved in the induced conformational changes observed for the pAcPhe synthetase. In addition, we are conducting studies with synthetases specific for other unnatural amino acids to determine if structural plasticity is a general characteristic of these enzymes.

Acknowledgment. The work in this communication is based on experiments conducted at beamline 5.0.3 of the Advanced Light Source (ALS). The ALS is supported by the Director, Office of Science, Office of Basic Energy Sciences, Material Sciences Division of the U.S. Department of Energy under Contract DE-AC03-76SF00098 at the Lawrence Berkeley National Laboratory. We thank all of the staff at these beamlines for their continued support. We also thank Eileen Ambing for technical assistance in looping the crystals, Yan Zhang for helpful discussions, and Andreas Kreusch, Christian Lee, Michael Didonato, Phillip Chamberlain, and Scott Lesley for data collection. This work is supported by grants from the National Institutes of Health (GM62159), the Department of Energy (DE-FG03-00ER45812), and the Skaggs Institute for Chemical Biology. J.M.T. acknowledges support from the National Institutes of Health for a postdoctoral fellowship. Coordinates for the pAcPhe synthetase have been deposited in the Protein Data Bank as 1ZH6.

Supporting Information Available: Experimental preparation, purification, and crystallization of pAcPhe TyrRS, as well as data collection and structure determination. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Wang, L.; Schultz, P. G. *Angew. Chem., Int. Ed.* **2005**, *44*, 34−66.
(2) Chin, J. W.; Cropp, T. A.; Anderson, J. C.; Mukherji, M.; Zhang, Z.; Schultz, P. G. *Science (Washington, D.C.)* **2003**, *301*, 964−967.
(3) Wang, L.; Brock, A.; Herberich, B.; Schultz, P. G. *Science (Washington, D.C.)* **2001**, *292*, 498−500.
(4) Wang, L.; Zhang, Z.; Brock, A.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 56−61.
(5) Kobayashi, T.; Nureki, O.; Ishitani, R.; Yaremchuk, A.; Tukalo, M.; Cusack, S.; Sakamoto, K.; Yokoyama, S. *Nat. Struct. Biol.* **2003**, *10*, 425−432.
(6) Zhang, Y.; Wang, L.; Schultz, P. G.; Wilson, I. A. *Protein Sci.* **2005**, *14*, 1340−1349.
(7) Richardson, J. S.; Richardson, D. C. *Science* **1988**, *240*, 1648−1652.
(8) Kobayashi, T.; Takimura, T.; Sekine, R.; Vincent, K.; Kamata, K.; Sakamoto, K.; Nishimura, S.; Yokoyama, S. *J. Mol. Biol.* **2005**, *346*, 105−117.
(9) Kobayashi, T.; Sakamoto, K.; Takimura, T.; Sekine, R.; Vincent, K.; Kamata, K.; Nishimura, S.; Yokoyama, S. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 1366−1371.
(10) Buddha, M. R.; Crane, B. R. *Nat. Struct. Mol. Biol.* **2005**, *12*, 274−275.
(11) Hurley, J. H.; Chen, R.; Dean, A. M. *Biochemistry* **1996**, *35*, 5670−5678.
(12) Chen, R.; Greer, A.; Dean, A. M. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 12171−12176.
(13) Ballinger, M. D.; Tom, J.; Wells, J. A. *Biochemistry* **1995**, *34*, 13312−13319.
(14) Oue, S.; Okamoto, A.; Yano, T.; Kagamiyama, H. *J. Biol. Chem.* **1999**, *274*, 2344−2349.
(15) Yano, T.; Oue, S.; Kagamiyama, H. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 5511−5515.

JA0549042