Amino Acid Networks in a (β/α) 8 Barrel Enzyme Change during Catalytic Turnover

ARTICLE *in* JOURNAL OF THE AMERICAN CHEMICAL SOCIETY · APRIL 2014

Impact Factor: 12.11 · DOI: 10.1021/ja501602t · Source: PubMed

7 AUTHORS, INCLUDING:

Eric M Yezdimer
Gelita, USA
19 PUBLICATIONS **321** CITATIONS
SEE PROFILE

Chia-En A Chang
University of California, Riverside
43 PUBLICATIONS **1,347** CITATIONS
SEE PROFILE

David D Boehr
Pennsylvania State University
45 PUBLICATIONS **2,369** CITATIONS
SEE PROFILE

JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Communication
pubs.acs.org/JACS

Amino Acid Networks in a $(\beta/\alpha)_8$ Barrel Enzyme Change during Catalytic Turnover

Jennifer M. Axe,[†] Eric M. Yezdimer,[†] Kathleen F. O'Rourke,[†] Nicole E. Kerstetter,[†] Wanli You,[‡] Chia-en A. Chang,[‡] and David D. Boehr*,[†]

[†]Department of Chemistry, The Pennsylvania State University, University Park, Pennsylvania 16802, United States
[‡]Department of Chemistry, University of California, Riverside, California 92521, United States

Supporting Information

ABSTRACT: Proteins can be viewed as small-world networks of amino acid residues connected through noncovalent interactions. Nuclear magnetic resonance chemical shift covariance analyses were used to identify long-range amino acid networks in the α subunit of tryptophan synthase both for the *resting* state (in the absence of substrate and product) and for the *working* state (during catalytic turnover). The amino acid networks observed stretch from the surface of the protein into the active site and are different between the *resting* and *working* states. Modification of surface residues on the network alters the structural dynamics of active-site residues over 25 Å away and leads to changes in catalytic rates. These findings demonstrate that amino acid networks, similar to those studied here, are likely important for coordinating structural changes necessary for enzyme function and regulation.

Globular proteins are held together by complex webs of noncovalent interactions. Long-range allosteric changes may propagate across these networks of interactions, resulting in the formation and/or breakage of a series of noncovalent interactions to induce conformational and/or motional changes in the protein structure.[1] Enzymes, in particular, must often fluctuate into a variety of conformations during their catalytic cycles.[2] It is now recognized that allosteric networks may be intrinsic to all proteins.[1,3] Here, we show by using NMR chemical shift covariance analyses (CHESCA)[4] that these allosteric networks change depending on the functional state of the enzyme. Intriguingly, active-site residues known to be involved directly in chemical catalysis change networks depending if the enzyme is in a ligand-free resting state or if the enzyme is actively turning over substrate/products. These results suggest that dynamic amino acid networks play critical roles in the positioning and/or structural dynamics of catalytically relevant amino acid residues and may be important considerations in the rational (re)design of proteins and other macromolecules.

NMR spectroscopy has proven to be an especially powerful technique to experimentally delineate amino acid networks in proteins.[5−9] The chemical shift, the fundamental parameter of NMR, is a sensitive reporter of changes to the structural dynamics of proteins. In one scenario, two conformations exchange on a time scale relatively fast compared to the time evolution of the chemical shift. Under this fast exchange regime, the observed chemical shifts are weighted linear averages of the chemical shifts of the two conformations. Ligand binding or amino acid substitutions may perturb the conformational equilibrium and result in chemical shift changes both local and remote from the site(s) of perturbation. Residues with covarying chemical shift changes across a series of perturbations are likely involved in the same conformational change and are projected to be in the same amino acid network.[4] This concept forms the basis of the CHESCA approach, which has been previously used to identify amino acid networks in EPAC,[4,10] PKA,[11,12] and CFTR.[13]

Here, we have applied the CHESCA approach to delineate amino acid networks in the α subunit of tryptophan synthase (αTS), in isolation from the β subunit (βTS). αTS is a $(\beta/\alpha)_8$ barrel enzyme that catalyzes the cleavage of indole-3-glycerol phosphate (IGP) to form indole and D-glyceralde-3-phosphate (G3P), the penultimate step of tryptophan biosynthesis. The equilibrium of this reaction lies toward IGP (Figure S1). In αTS, Glu49 on the β2 strand and Asp60 on the $\beta2\alpha2$ loop (i.e., the loop between the β2 strand and the α2 helix) play direct roles in chemical catalysis[14] (Figure 1a, S1). We chose to study αTS from *Escherichia coli* considering that most of the NMR backbone resonances were previously assigned.[15,16] The αTS from *E. coli* is 85% identical to the commonly studied *Salmonella typhimurium* enzyme (Figure S2).

In our modified CHESCA approach,[16] we used Ala-to-Gly site mutants as our series of perturbations. The Ala-to-Gly modification replaces a methyl group in the protein with a hydrogen and potentially changes the conformational space accessible by the immediate protein backbone. These local changes may induce longer-range structural dynamic changes that result in chemical shift changes. All of the Ala-to-Gly site mutants (i.e., A59G, A67G, A158G, A180G, and A185G; see Figure 1a) were previously shown to perturb the conformational equilibrium between a ligand-free state and a higher energy conformation resembling that observed when G3P is bound to the protein.[16]

We have previously established conditions that allow us to make NMR measurements under dynamic chemical equilibrium conditions (i.e., starting with 10 mM indole and 10 mM G3P; K_{eq} for the reaction is 0.4, favoring formation of IGP[17]),

Received: February 14, 2014
Published: April 28, 2014

   dx.doi.org/10.1021/ja501602t | *J. Am. Chem. Soc.* 2014, 136, 6818−6821

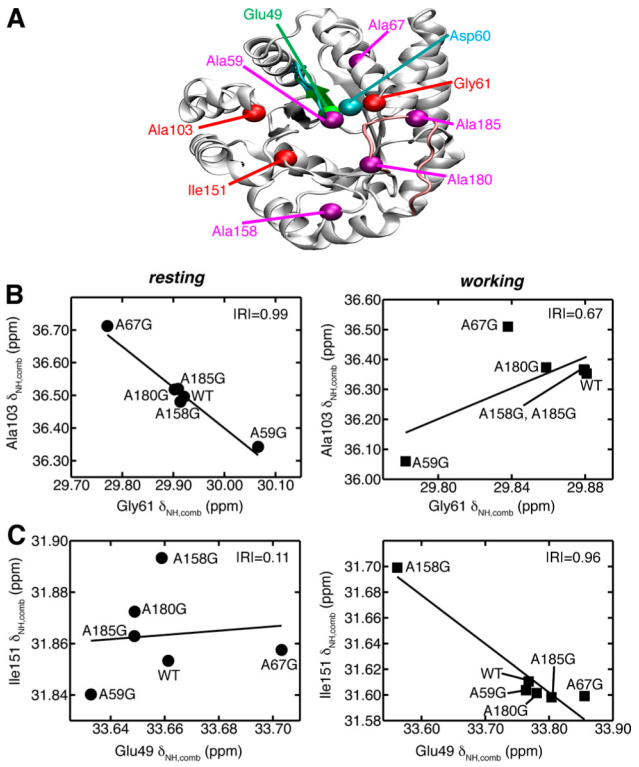

Figure 1. Chemical shift correlations are different between the *resting* and *working* states of αTS. (A) Ribbon structure of αTS (PDB 1K3U) indicating the location of Glu49 (green) on the β2 strand, Asp60 (blue) on the β2α2 loop, sites for Ala-to-Gly substitutions for CHESCA (magenta), and other probe sites (red). (B–C) Chemical shift correlations in the (left) *resting* and/or (right) *working* states of αTS. Correlation plots used a combined ^{1}H, ^{15}N chemical shift according to $\delta_{NH,comb} = \delta_H + 0.2\ \delta_N$, where δ_H and δ_N are the ^{1}H and ^{15}N chemical shifts, respectively. The Pearson correlation coefficients (R) are shown.

enabling us to compare amino acid networks delineated by CHESCA in both a ligand-free *resting* state and a ligand-bound *working* state. It should be kept in mind that the *working* state is not necessarily a distinct structural state, since it represents a sampling of both IGP-bound and indole/G3P-bound states, but rather a functional state representing an actively turning over enzyme, as similarly established for cyclophilin A[18] and adenylate kinase.[19]

Under the *working* state conditions, we would expect 80% of the indole/G3P to be converted to IGP. We previously noted that Ala59 gives rise to two resonances with an intensity ratio of ~80:20, likely reflecting the ratio between IGP and indole:G3P bound states.[16] Here, we note a similar behavior for Gly61 (Figure S3). We also note that the Ala-to-Gly substitutions neither substantially alter this ratio (Figure S3) nor affect binding affinities for indole and G3P, as reflected by wild-type (WT) NMR ligand titration behavior and steady-state kinetic parameters (Figure S3). Thus, any chemical shift changes induced by the Ala-to-Gly substitutions cannot simply reflect changes in ligand affinities. Other residues, besides Ala59 and Gly61, do not give rise to two resonances, suggesting such behavior in the *working* state must only be due to very local changes in the active site.

^{1}H–^{15}N heteronuclear single quantum coherence (HSQC) spectra were collected for WT αTS and the five protein variants under *resting* and *working* conditions. To determine correlations between chemical shifts of the backbone amides, each two-dimensional peak was first assigned a single, combined chemical shift, $s(x)_i = \delta_H + 0.2\ \delta_N$, where δ_H and δ_N are the ^{1}H and ^{15}N chemical shifts, respectively. Recognizing that the "true" experimental uncertainty could be larger than that observed in our data sets, we assigned a constant error (σ) to each $s(x)_i$, where x denotes a specific backbone amide group and i represents a specific protein variant. This uncertainty was then numerically accounted for by randomly generating sets of N points that followed a Gaussian distribution, centered on $s(x)_i$ with a standard deviation of σ. The Pearson correlation between two backbone amides within a series of protein variants was then determined (Figure S4) using N times n number of points, where n is the number of protein variants. The parameter N was set to 400 and was selected to achieve correlation coefficients that were reproducible to within two digits (i.e., following repeated calculation with a newly generated series of $N \times n$ points). Because the peaks of the backbone amides were not always clearly definable in some variants due to spectral overlap with other amide groups, or because some data points were eliminated due to "proximity" effects (i.e., much higher than expected chemical shift changes due to local perturbations around the mutation site), the number of available data points in a given correlation sometimes varied (but was always greater than four). Correlations between two amino acid residues were therefore only deemed statistically significant if the correlation met $p < 0.05$, with the corresponding Pearson coefficient criteria being dependent on the number of n data points available.

Very strikingly, the residues in αTS showing significant linear correlations are quite different between the *resting* and *working* states (Figures 1b,c, S5). An agglomerative clustering algorithm using a single-linkage (i.e., nearest-neighbor) model was used to organize the chemical shift correlation matrices (Figure S5) into dendrograms (Figure S6). Since only statistically significant correlations were used in the construction of the dendrogram, discrete networks of amino acids are clearly defined. Our analysis revealed two major clusters of residues in both the *resting* (i.e., *resting cluster 1* and *resting cluster 2*) and *working* states (i.e., *working cluster 1* and *working cluster 2*; Figure S6). Many residues are similar between *resting cluster 1* and *working cluster 1*, and likewise between *resting cluster 2* and *working cluster 2*, although there are some notable exceptions as discussed below. To better visualize these networks, we plotted the clusters onto the αTS structure (Figure 2; here we have used the X-ray crystal structure of *S. typhimurium* αTS,[20] since the *E. coli* αTS structure[21] has some disordered loops, Figure S2). For the *resting* enzyme, *cluster 1* and *cluster 2* appear to be separated across the active site (Figure 2). In contrast, *working cluster 1* includes residues on either side of the active site, potentially providing communication pathways across the protein.

One remarkable finding is that Glu49, an amino acid directly involved in chemical catalysis (Figure S1), is found in *resting cluster 2*, but changes to *cluster 1* in the *working* state (Figure 2). For the *working* state, a contiguous network of noncovalent interactions can be envisioned that involves Glu49 in the β2 strand (Figure S7[22]). However, structural analysis alone is unable to explain why Glu49 changes networks between the *resting* and *working* states, considering that the X-ray crystal structures of ligand-free and IGP-bound (from *S. typhimurium*) αTS are nearly identical, especially in those regions identified in the cluster[23,24] (Figure S7). As such, these results may reflect

dx.doi.org/10.1021/ja501602t | J. Am. Chem. Soc. 2014, 136, 6818–6821

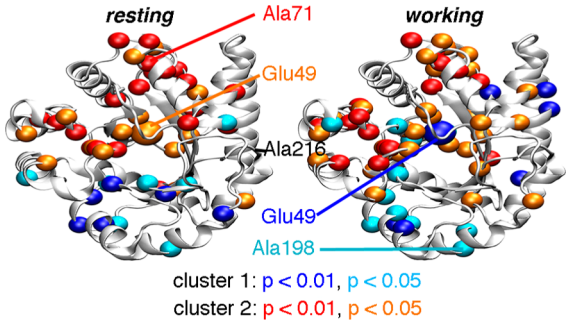

Figure 2. Amino acid networks are different between the *resting* and *working* states of αTS. Agglomerative clustering is used to identify "nearest-neighbor" clusters in the *resting* (left) and *working* (right) states. Residues in *resting/working cluster 1* and *resting/working cluster 2* are plotted as blue/cyan and red/orange spheres onto the αTS structure (PDB 1K3U). Colors also correspond to the level of statistical significance that these residues are found in their appropriate clusters ($p < 0.01$, blue/red; $p < 0.05$, cyan/orange). Dendrograms associated with these clusters are located in Figure S6.

changes in "dynamically driven allostery".[25] Indeed, we had previously noted dynamic differences in many of these regions between the *resting* and *working* states.[16] Moreover, molecular dynamics (MD) simulations for the *S. typhimurium* αTS suggest that correlated motions change substantially between the ligand-free and IGP-bound states[26,27] (Figure S8). In particular, Glu49 shows many correlations with residues identified in *working cluster 1* for the IGP-bound state, but these are all lacking in the ligand-free state. Related to this idea, it was previously noted, based on X-ray crystal structures, that Glu49 can occupy different positions, and these results suggested that the proper positioning of Glu49 helps to regulate enzyme activity.[24] It is also interesting to note that many of the cluster residues are near the interface where the β-subunit would bind (Figure S9), suggesting that βTS may influence αTS catalysis by modulating these amino acid networks. MD simulations also suggest that binding of βTS changes correlated motions in αTS.[26]

To gain a better understanding of the networks involving Glu49, we retasked our clustering algorithm to search for those residues showing significant linear chemical shift correlations with Glu49 and then identified those residues that have significant linear chemical shift correlations with the first set of residues (Figures S10, S11). To further test these clusters, we made substitutions at residues in both *resting/working cluster 1* (i.e., L162G, A167G, A198G) and *resting/working cluster 2* (i.e., A71G) and at a residue not assigned to any cluster (i.e., A216G). Remarkably, the A71G substitution induces a larger chemical shift change to Glu49 in the *resting* state (Figure 3a), and the L162G, A167G, and A198G substitutions induce larger Glu49 chemical shift changes in the *working* state (i.e., compared to WT αTS; Figures 3a, S12). These amino acid substitutions also generally induce larger chemical shift changes to other residues in their respective clusters than to those residues outside the cluster (Figure S13). Amino acid substitutions at these network residues lead to changes to the catalytic rate (k_{cat}) of αTS (Figure 3b). The A71G substitution results in a 5-fold decrease in the catalytic rate (Figure 3b). The L162G, A167G, and A198G substitutions also lead to decreases in the catalytic rate, with the largest change being an almost 2-fold decrease in the catalytic rate for the A167G variant (Figure 3b, S12). While modest, these changes are still noteworthy

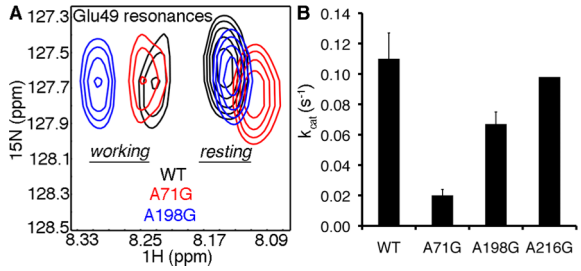

Figure 3. Testing the association of Glu49 with different clusters. (A) Comparisons of the backbone amide chemical shifts for Glu49 between WT (black), A71G (red), and A198G (blue) αTS for both *resting* (right) and *working* (left) states, based on $^{1}H-^{15}N$ HSQC spectra collected at 298 K. (B) Comparison of the catalytic rates (k_{cat}) for WT, A71G, A198G, and A216G αTS.

considering that these amino acid substitutions are >15 Å away from the active site. In contrast, the non-network amino acid substitution (i.e., A216G), a similar distance away from the active site, had no significant effect on the catalytic rate (Figure 3b).

Other methods have been developed to probe amino acid networks within proteins.[7,28−33] Here we show that these amino acid networks are not likely to be static, but change depending on the functional state of the protein. Moreover, the switching of these allosteric pathways appears to be driven by changes in conformational dynamics induced by ligand binding. While the contiguous network of interactions (e.g., Figure S7) may argue in favor of a "mechanical linkage" model of allostery,[34] an alternative view is that the allosteric changes represent a redistribution of the underlying protein conformational ensemble;[35] both models emphasize the contribution of protein dynamics to allostery. The amino acid networks identified here and elsewhere may provide molecular evolution with a means to select for both an ensemble of conformations required for function and for efficient transitions between these various states. Wiring in these dynamic amino acid networks may be critical for coordinating functionally relevant motions and structural transitions in engineered enzymes and other macromolecules.

■ ASSOCIATED CONTENT

Supporting Information

Methods, Figures S1−13. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

ddb12@psu.edu

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Drs. Philip Bevilacqua and Craig Cameron for their insightful comments. This work was supported by NSF Career Grant MCB1053993.

■ REFERENCES

(1) Csermely, P.; Sandhu, K. S.; Hazai, E.; Hoksza, Z.; Kiss, H. J.; Miozzo, F.; Veres, D. V.; Piazza, F.; Nussinov, R. *Curr. Protein Pept. Sci.* **2012**, *13*, 19−33.

(2) Villali, J.; Kern, D. *Curr. Opin. Chem. Biol.* **2010**, *14*, 636−43.

(3) Goodey, N. M.; Benkovic, S. J. *Nat. Chem. Biol.* **2008**, *4*, 474−82.
(4) Selvaratnam, R.; Chowdhury, S.; VanSchouwen, B.; Melacini, G. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 6133−8.
(5) Mayer, K. L.; Earley, M. R.; Gupta, S.; Pichumani, K.; Regan, L.; Stone, M. J. *Nat. Struct. Biol.* **2003**, *10*, 962−5.
(6) Gagne, D.; Charest, L. A.; Morin, S.; Kovrigin, E. L.; Doucet, N. *J. Biol. Chem.* **2012**, *287*, 44289−300.
(7) Rivalta, I.; Sultan, M. M.; Lee, N. S.; Manley, G. A.; Loria, J. P.; Batista, V. S. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, E1428−36.
(8) Fuentes, E. J.; Gilmore, S. A.; Mauldin, R. V.; Lee, A. L. *J. Mol. Biol.* **2006**, *364*, 337−51.
(9) Tzeng, S. R.; Kalodimos, C. G. *Nature* **2009**, *462*, 368−72.
(10) Selvaratnam, R.; Mazhab-Jafari, M. T.; Das, R.; Melacini, G. *PLoS One* **2012**, *7*, e48707.
(11) Akimoto, M.; Selvaratnam, R.; McNicholl, E. T.; Verma, G.; Taylor, S. S.; Melacini, G. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 14231−6.
(12) Cembran, A.; Kim, J.; Gao, J.; Veglia, G. *Phys. Chem. Chem. Phys.* **2014**, *16*, 6508−18.
(13) Dawson, J. E.; Farber, P. J.; Forman-Kay, J. D. *PLoS One* **2013**, *8*, e74347.
(14) Dunn, M. F. *Arch. Biochem. Biophys.* **2012**, *519*, 154−66.
(15) Vadrevu, R.; Falzone, C. J.; Matthews, C. R. *Protein Sci.* **2003**, *12*, 185−91.
(16) Axe, J. M.; Boehr, D. D. *J. Mol. Biol.* **2013**, *425*, 1527−45.
(17) Weischet, W. O.; Kirschner, K. *Eur. J. Biochem.* **1976**, *65*, 375−85.
(18) Eisenmesser, E. Z.; Millet, O.; Labeikovsky, W.; Korzhnev, D. M.; Wolf-Watz, M.; Bosco, D. A.; Skalicky, J. J.; Kay, L. E.; Kern, D. *Nature* **2005**, *438*, 117−21.
(19) Wolf-Watz, M.; Thai, V.; Henzler-Wildman, K.; Hadjipavlou, G.; Eisenmesser, E. Z.; Kern, D. *Nat. Struct. Mol. Biol.* **2004**, *11*, 945−9.
(20) Weyand, M.; Schlichting, I.; Marabotti, A.; Mozzarelli, A. *J. Biol. Chem.* **2002**, *277*, 10647−52.
(21) Jeong, M. S.; Jeong, J. K.; Lim, W. K.; Jang, S. B. *Biochem. Biophys. Res. Commun.* **2004**, *323*, 1257−64.
(22) Sobolev, V.; Eyal, E.; Gerzon, S.; Potapov, V.; Babor, M.; Prilusky, J.; Edelman, M. *Nucleic Acids Res.* **2005**, *33*, W39−43.
(23) Rhee, S.; Parris, K. D.; Ahmed, S. A.; Miles, E. W.; Davies, D. R. *Biochemistry* **1996**, *35*, 4211−21.
(24) Rhee, S.; Miles, E. W.; Davies, D. R. *J. Biol. Chem.* **1998**, *273*, 8553−5.
(25) Popovych, N.; Sun, S.; Ebright, R. H.; Kalodimos, C. G. *Nat. Struct. Mol. Biol.* **2006**, *13*, 831−8.
(26) Fatmi, M. Q.; Chang, C. E. *PLoS Comput. Biol.* **2010**, *6*, e1000994.
(27) Ai, R.; Fatmi, M. Q.; Chang, C. E. A. *J. Comput.-Aided Mol. Des.* **2010**, *24*, 819−827.
(28) Ichiye, T.; Karplus, M. *Proteins* **1991**, *11*, 205−17.
(29) Amitai, G.; Shemesh, A.; Sitbon, E.; Shklar, M.; Netanely, D.; Venger, I.; Pietrokovski, S. *J. Mol. Biol.* **2004**, *344*, 1135−46.
(30) van den Bedem, H.; Bhabha, G.; Yang, K.; Wright, P. E.; Fraser, J. S. *Nat. Methods* **2013**, *10*, 896−902.
(31) Reynolds, K. A.; Russ, W. P.; Socolich, M.; Ranganathan, R. *Methods Enzymol.* **2013**, *523*, 213−35.
(32) Suel, G. M.; Lockless, S. W.; Wall, M. A.; Ranganathan, R. *Nat. Struct. Biol.* **2003**, *10*, 59−69.
(33) Livesay, D. R.; Kreth, K. E.; Fodor, A. A. *Methods Mol. Biol.* **2012**, *796*, 385−98.
(34) Yu, E. W.; Koshland, D. E., Jr. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 9517−20.
(35) Motlagh, H. N.; Wrabl, J. O.; Li, J.; Hilser, V. J. *Nature* **2014**, *508*, 331−9.