J|A|C|S COMMUNICATIONS

Published on Web 12/10/2010

X-ray and NMR Crystallography in an Enzyme Active Site: The Indoline Quinonoid Intermediate in Tryptophan Synthase

Jinfeng Lai,[†] Dimitri Niks,[‡] Yachong Wang,[†] Tatiana Domratcheva,[§] Thomas R. M. Barends,[§] Friedrich Schwarz,[§] Ryan A. Olsen,[†] Douglas W. Elliott,[†] M. Qaiser Fatmi,[†] Chia-en A. Chang,[†] Ilme Schlichting,[§] Michael F. Dunn,*[,‡] and Leonard J. Mueller*[,†]

Departments of Chemistry and Biochemistry, University of California, Riverside, California 92521, United States, and Department of Biomolecular Mechanisms, Max Planck Institute for Medical Research, 69120 Heidelberg, Germany

Received July 23, 2010; E-mail: Michael.Dunn@ucr.edu; Leonard.Mueller@ucr.edu

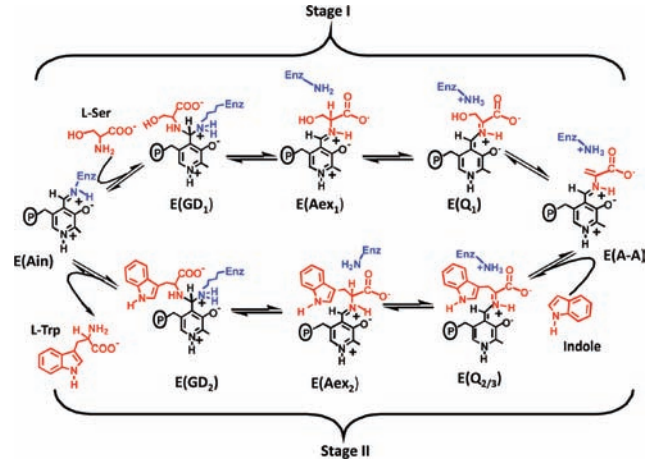
Abstract: Chemical-level details such as protonation and hybridization state are critical for understanding enzyme mechanism and function. Even at high resolution, these details are difficult to determine by X-ray crystallography alone. The chemical shift in NMR spectroscopy, however, is an extremely sensitive probe of the chemical environment, making solid-state NMR spectroscopy and X-ray crystallography a powerful combination for defining chemically detailed three-dimensional structures. Here we adopted this combined approach to determine the chemically rich crystal structure of the indoline quinonoid intermediate in the pyridoxal-5′-phosphate-dependent enzyme tryptophan synthase under conditions of active catalysis. Models of the active site were developed using a synergistic approach in which the structure of this reactive substrate analogue was optimized using ab initio computational chemistry in the presence of side-chain residues fixed at their crystallographically determined coordinates. Various models of charge and protonation state for the substrate and nearby catalytic residues could be uniquely distinguished by their calculated effects on the chemical shifts measured at specifically ^{13}C- and ^{15}N-labeled positions on the substrate. Our model suggests the importance of an equilibrium between tautomeric forms of the substrate, with the protonation state of the major isomer directing the next catalytic step.

The pyridoxal-5′-phosphate (PLP, vitamin B_6)-dependent tryptophan synthase $\alpha_2\beta_2$ bienzyme complex catalyzes the last two steps in the synthesis of L-Trp,[1,2] which are consecutive processes that require channeling of the common metabolite, indole, between the active sites of the α- and β-subunits.[2−5] Indole channeling is regulated by long-range allosteric interactions that switch the enzyme between open, low activity and closed, high activity states during the catalytic cycle.[3−6] In the α-site, 3-indole-D-glycerol-3′-phosphate (IGP) is cleaved to D-glyceraldehyde-3-phosphate and indole. In stage I of the β-reaction (Scheme 1), L-Ser reacts with the internal aldimine [E(Ain)], giving in sequence *gem*-diamine [E(GD$_1$)], L-Ser external aldimine [E(Aex$_1$)], quinonoid [E(Q$_1$)], and aminoacrylate Schiff base [E(A-A)] species and a water molecule. In stage II, indole channeled from the α-site makes a nucleophilic attack on E(A-A), giving E(Q$_{2/3}$), E(Aex$_2$), and E(GD$_2$) intermediates and finally the product, L-Trp. In nearly all PLP enzymes, proton transfer to C$^\alpha$ of the substrate is catalyzed by an active-site lysine residue.[4,5]

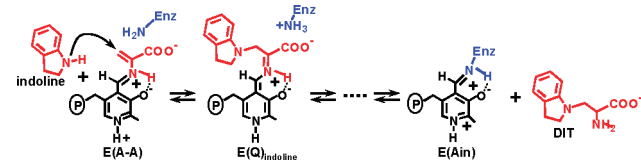
Recently, crystal structures of both the open and closed conformations have been reported.[6−8] Despite these successes, significant questions regarding the chemical mechanism of the substrate transformation remain, as the resolution of these X-ray structures does not allow for the protonation states of the substrate and nearby residues to be established. In this respect, solid-state NMR spectroscopy can be particularly useful for studying enzyme systems,[9−13] as it provides chemical-level details under essentially the same conditions used to solve the X-ray crystal structure. Moreover, solid-state NMR spectroscopy can be applied to moderately sized biomolecules such as the 144 kDa tryptophan synthase bienzyme complex. Initial work by McDowell et al.[13] showed that when labeled L-[3-^{13}C]Ser was supplied as a substrate, solid-state ^{13}C NMR spectroscopy could identify the E(A-A) intermediate in the crystalline state for the full tryptophan synthase.

Here we report on the transient quinonoid species, a defining catalytic intermediate in all PLP-dependent enzymes.[14] As shown in Scheme 2, upon reaction with indoline, the aminoacrylate intermediate gives the quasi-stable E(Q)$_{indoline}$,[15,16] a chemical and structural analogue of the short-lived quinonoid species formed in stage II of the β-reaction. We first determined at 1.85 Å resolution the crystal structure of the Cs$^+$ form of E(Q)$_{indoline}$ with an IGP analogue, *N*-(4′-trifluoromethoxybenzenesulfonyl)-2-aminoethyl phosphate (F9), bound to the α-site. The binding of Cs$^+$ and F9 and the

Scheme 1. The β-Site Reaction in Tryptophan Synthase



Scheme 2. The Indoline Reaction



[†] Department of Chemistry, University of California, Riverside.
[‡] Department of Biochemistry, University of California, Riverside.
[§] Max Planck Institute for Medical Research.

10.1021/ja106555c

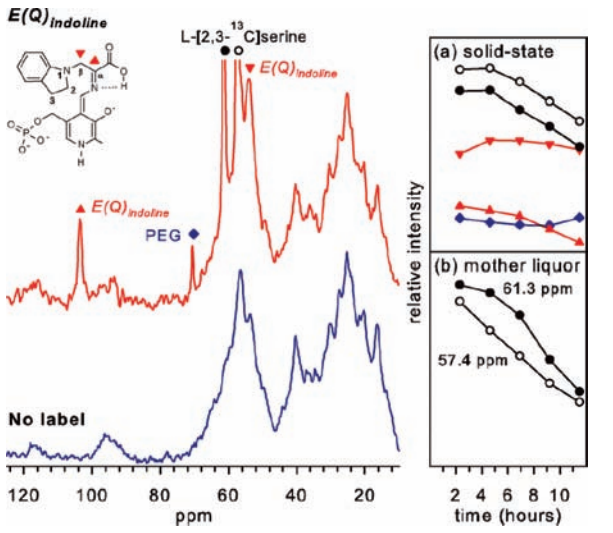


Figure 1. ^{13}C solid-state cross-polarization magic-angle-spinning (CPMAS) NMR spectra of the reaction of tryptophan synthase microcrystals with 20 mM L-[2,3-^{13}C]Ser (red spectrum) or 20 mM unlabeled L-Ser (blue spectrum) and 20 mM indoline in the presence of 3 mM F9 and 100 mM CsCl, measured at −10 °C and averaged over 4.6 h. (Inset) Time courses for the disappearance and appearance of the enriched ^{13}C signals in the (a) bound (solid) state and (b) mother liquor, acquired using (a) CPMAS with high-power ^{1}H decoupling (main panel) and (b) direct ^{13}C excitation with low-power ^{1}H decoupling (Figure S9). The data were acquired on a 9.4 T Bruker DSX spectrometer equipped with a 4 mm CPMAS probe spinning at a MAS frequency of 8 kHz.

formation of $E(Q)_{indoline}$ stabilize closed conformations of the α- and β-subunits.[5,6,8,16] The diffraction data [summarized in the Supporting Information (SI)] showed well-defined electron densities for F9 and the indoline quinonoid species (Figures S3 and S4). Except for minor structural changes at the α-site that are specific to the structure of F9, the structure of the complex was found to be nearly identical to the $E(Q)_{indoline}$ structure reported by Barends et al.[8] but significantly more stable.

We were also able to observe the $E(Q)_{indoline}$ intermediate using solid-state NMR spectroscopy (Figure 1) under conditions where the microcrystals remained catalytically active. Analysis of the time courses for the solid state and for the mother liquor (Figure 1a,b) allowed the assignment of the resonances at 103.6 ppm (red ▲) and 54.1 ppm (red ▼) as the sp^2-hybridized 2-^{13}C (C^{α}) and sp^3-hybridized 3-^{13}C (C^{β}) of $E(Q)_{indoline}$, respectively, which were observed as long as sufficient L-serine and indoline were present in solution. Resonances at 61.3 ppm and 57.4 ppm (black ● and ○) were assigned to L-[2,3-^{13}C]serine nonspecifically bound to the microcrystals. Analogous measurements using [U-^{13}C,^{15}N]serine and -indoline (Figures S5−S8) provided the additional chemical shifts in Table 1.

To interpret the experimental shifts in terms of the ionization states of the bound substrate,[9] the chemical shifts of candidate structures were predicted using ab initio electronic structure calculations[17] based on the following protocol. First, a model of the β-subunit enzyme active site (residues within 7 Å of the substrate) was constructed in which the non-hydrogen side-chain atoms were fixed at their crystallographically determined coordinates (Figure 2a). This provided a framework for optimizing the structure of the PLP−ligand complex in the presence of local charges and hydrogen-bonding interactions. Second, candidate structures for the substrate protonation states were generated and then optimized to ground-state geometries using the ONIOM method[18] [side chains were treated at the semiempirical level of theory, and the substrate was treated using density functional theory

Table 1. Experimental and Calculated Chemical Shifts [in ppm Relative to TMS (^{13}C) and $NH_3(l)$ (^{15}N)[a]] for Three Forms of $E(Q)_{indoline}$ and the Fast-Exchange Equilibrium between Two and Three Sites As Described in the Text (Site Labels Are Shown in Figure 1)

	Exp.	Calculated PSB Form	Calculated Acid Form	Two-Site Equilib.	Calculated Phenolic Form	Three-Site Equilib.
C^{α}	103.6	106.0	101.3	102.9	124.7	104.9
C^{β}	54.1	54.0	49.6	51.1	52.7	51.4
C'	173.0	172.3	169.3	170.3	175.6	170.9
N	296.5	215.1	337.2	295.7	321.7	295.7
C2	50.5	50.1	51.5	51.1	50.2	51.0
C3	28.5	31.9	32.5	32.3	31.8	32.2
N1	83.5	85.0	91.1	89.0	84.5	88.5

[a] For comparison, $\delta[NH_3(l)] = \delta[^{15}NH_4Cl(s)] + 39.2$ ppm.

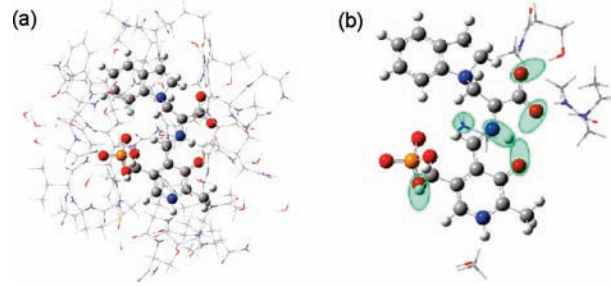
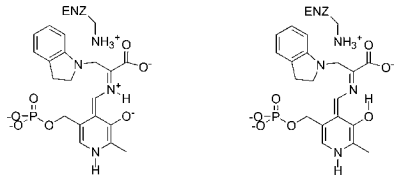


Figure 2. (a) Model of the enzyme active site in the β-subunit of tryptophan synthase, showing (wireframe) the side chains fixed at their crystallographically determined coordinates and treated at the semiempirical PM3 level of theory and (ball and stick) the PLP−substrate complex optimized using DFT at the B3LYP/6-31G(d,p) level. (b) Substructure used for calculating NMR chemical shifts [B3LYP/6-311++G(d,p) DFT]. Green ovals indicate possible sites of protonation on the PLP−substrate complex and the Lys87 side chain. The standard CPK scheme is used to designate the atom colors (H, white; C, gray; N, blue; O, red; P, orange).

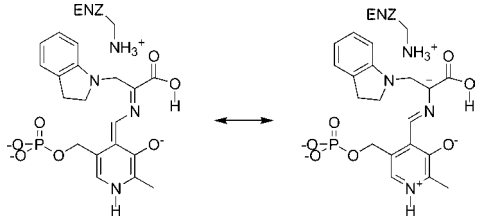
(DFT)] in Gaussian 03,[19] subject to minimal constraints as described in the SI. Finally, substrate chemical shifts were calculated using DFT for the excised substructure shown in Figure 2b, which included fragments of residues that were potentially charged or hydrogen-bonded to the substrate. Figure 2b shows possible sites of protonation used to generate the candidate structures for the PLP−ligand complex, including the phosphoryl group, the pyridoxal phenolic oxygen, the Schiff base linkage, both carboxylate oxygens, and the neighboring N^{ε} of β-Lys87. Not all of these sites can be simultaneously protonated, and we did not consider candidate structures with more than a single proton placed in the pocket defined by the pyridoxal oxygen, the Schiff base nitrogen, and the closest carboxylate oxygen; we also did not allow a doubly protonated carboxylate. This gave 28 possible candidate structures, each of which was geometrically optimized and for which chemical shifts were calculated (Tables S2 and S3) for comparison with the experimental values. This comparison was quantified using the reduced χ^2 statistic (the weighted deviation of the model and experimental shifts), which was calculated using root-mean-square deviation weights of 2 ppm for ^{13}C and 4 ppm for ^{15}N as described in the SI.

Of the candidate structures, none was found to have acceptable agreement between the calculated and experimental ^{13}C and ^{15}N chemical shifts (lowest reduced $\chi^2 = 4.9$). A large discrepancy for most of the structures occurred at the Schiff base nitrogen, which has an experimental shift of 296.5 ppm. Previous measurements in inhibited PLP-dependent alanine racemase give a protonated Schiff base nitrogen shift of 189 ppm,[12] while model compounds place a neutral Schiff base linkage with methylamine at 313 ppm[20] (both

Scheme 3. (left) Canonical Protonated Schiff Base (PSB) and (right) Phenolic Forms of the PLP−Substrate Complex



Scheme 4. Acid Form of the PLP−Substrate Complex and a Resonance Structure That Builds Up Negative Charge at C^{α}



adjusted to $\delta[NH_3(l)]$). These data suggest that the observed shifts are due to a fast proton exchange between multiple structures. The simplest model of this equilibrium is one in which only a single proton changes position. We constructed all such possible pairs for the ionizable sites in Figure 2 and then searched for the relative populations that best agreed with the experimental chemical shifts. The best-fit equilibrium, with a reduced χ^2 of 1.39, is between the canonical protonated Schiff base (PSB) substrate (Scheme 3) and, unexpectedly, the carboxylic acid form (Scheme 4), in which the proton moves from the Schiff base nitrogen to the neighboring carboxylate oxygen. A 34:66 ratio of the two forms in fast exchange gave extremely good agreement with the experimental data, as summarized by the two-site equilibrium results in Table 1. This acid form has not previously been hypothesized to play a major role in the PLP enzyme mechanism, but we note that the canonical PSB and phenolic forms (Scheme 3), alone or in equilibrium, are unable to satisfy both the Schiff base N and C^{α} chemical shifts (lowest reduced $\chi^2 = 3.2$). A three-site exchange between the PSB, acid, and phenolic forms cannot be ruled out, however, and would improve the agreement further (see the three-site equilibrium data in Table 1) with a reduced χ^2 of 1.22 for a population ratio of 33:58:9, respectively.

The structural hypothesis that emerges from the analysis of the chemical shifts is that the predominant structure is the acid form, with the protonated acid moiety forming a hydrogen bond back to the Schiff base nitrogen. This hydrogen bond is also suggested by the X-ray data, in which the distances are 2.57 and 2.73 Å from the Schiff base nitrogen to the carboxylate and phenolic oxygens, respectively (the corresponding distances for the two-site equilibrium model are 2.55 and 2.70 Å). Molecular dynamics simulations showed that the acid structure is indeed stable, with this intramolecular hydrogen bond remaining intact. In addition, the side chain of βK87 is more rigid in this complex in comparison with the side-chain motions in the free enzyme, where $E(Q)_{indoline}$ is not present. This structure has fundamental implications for the mechanism in tryptophan synthase, which at the next step involves protonation at the C^{α} site by βK87. Scheme 4 shows an additional resonance structure that builds up electron density at the C^{α} site for the acid isomer. Because of the repulsion between negative charges, this resonance is less important for the phenolic and PSB forms than for the acid form of the substrate, leading to a larger negative charge at C^{α} for the latter. The atomic charges calculated from the ab initio electronic structures using natural bond orbital analysis[21] are consistent with the resonance model, with C^{α} essentially neutral (−0.014 au) in the PSB form and having a charge of −0.097 au in the acid form. We postulate that the buildup of negative charge at C^{α} plays two roles. First, it adds an electric field component along the reaction coordinate to direct the proton from N^{ε} of βK87 to the C^{α} site. Second, it lowers the energy barrier along the reaction coordinate through charge stabilization. We also note that the buildup of negative charge at C^{α} is accompanied by a concomitant buildup of positive charge at C_4' of the PLP, disfavoring a competing transamination pathway.

The empirical correlation between chemical shift and chemical environment has been used to report on substrate−PLP protonation states in model compounds[12,20,22] and enzyme active sites.[11,12] Here the combination of solid-state NMR spectroscopy with diffraction and computational methods to build chemically detailed structural models with quantitative predictions of chemical shifts has been shown to hold remarkable promise for the structural characterization of enzyme intermediates. Predicated on the accuracy with which chemical shifts can now be calculated,[17] this approach, which is an extension of NMR crystallography[23] to a functioning enzyme system, points to a new hypothesis for the quinonoid protonation states in tryptophan synthase and their role in directing the next step in catalysis.

Acknowledgment. This work was supported by NSF Grant CHE0848607 (L.J.M.), NIH Grants S10RR023677 (L.J.M.) and GM055749 (M.F.D.), and the Max Planck Society (I.S.). Diffraction data were collected on beamline X10SA at the Swiss Light Source, Paul Scherrer Institute, Villigen, Switzerland. We thank the Dortmund−Heidelberg team and staff for their support.

Supporting Information Available: Complete ref 19, X-ray crystal structure of the $E(Q)_{indoline}$ complex, synthesis of [U-^{13}C,^{15}N]indoline, NMR experiment details, computational methods, molecular dynamics simulations, additional NMR spectra, and χ^2 analysis. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Yanofsky, C.; Crawford, I. P. In *The Enzymes*; Boyer, P. D., Ed.; Academic Press: New York, 1972; pp 1−31.
(2) Miles, E. W. *Adv. Enzymol. Relat. Areas Mol. Biol.* **1979**, *49*, 127–186.
(3) Dunn, M. F.; Aguilar, V.; Brzovic, P.; Drewe, W. F.; Houben, K. F.; Leja, C. A.; Roy, M. *Biochemistry* **1990**, *29*, 8598–8607. Kirschner, K.; Lane, A. N.; Strasser, A. W. M. *Biochemistry* **1991**, *30*, 472–478. Pan, P.; Woehl, E.; Dunn, M. F. *Trends Biochem. Sci.* **1997**, *22*, 22–27.
(4) Miles, E. W.; Rhee, S.; Davies, D. R. *J. Biol. Chem.* **1999**, *274*, 12193–12196.
(5) Dunn, M. F.; Niks, D.; Ngo, H.; Barends, T. R. M.; Schlichting, I. *Trends Biochem. Sci.* **2008**, *33*, 254–264.
(6) Ngo, H.; Kimmich, N.; Harris, R.; Niks, D.; Blumenstein, L.; Kulik, V.; Barends, T. R.; Schlichting, I.; Dunn, M. F. *Biochemistry* **2007**, *46*, 7740–7753.
(7) Rhee, S.; Parris, K. D.; Hyde, C. C.; Ahmed, S. A.; Miles, E. W.; Davies, D. R. *Biochemistry* **1997**, *36*, 7664–7680. Sachpatzidis, A.; Dealwis, C.; Lubetsky, J. B.; Liang, P. H.; Anderson, K. S.; Lolis, E. *Biochemistry* **1999**, *38*, 12665–12674. Weyand, M.; Schlichting, I.; Marabotti, A.; Mozzarelli, A. *J. Biol. Chem.* **2002**, *277*, 10647–10652. Blumenstein, L.; Domratcheva, T.; Niks, D.; Ngo, H.; Seidel, R.; Dunn, M. F.; Schlichting, I. *Biochemistry* **2007**, *46*, 14100–14116.
(8) Barends, T. R. M.; Domratcheva, T.; Kulik, V.; Blumenstein, L.; Niks, D.; Dunn, M. F.; Schlichting, I. *ChemBioChem* **2008**, *9*, 1024–1028.
(9) Mao, J. H.; Mukherjee, S.; Zhang, Y.; Cao, R.; Sanders, J. M.; Song, Y. C.; Zhang, Y. H.; Meints, G. A.; Gao, Y. G.; Mukkamala, D.; Hudock, M. P.; Oldfield, E. *J. Am. Chem. Soc.* **2006**, *128*, 14485–14497.
(10) McDermott, A.; Polenova, T. *Curr. Opin. Struct. Biol.* **2007**, *17*, 617–622.
(11) Sharif, S.; Fogle, E.; Toney, M. D.; Denisov, G. S.; Shenderovich, I. G.; Buntkowsky, G.; Tolstoy, P. M.; Huot, M. C.; Limbach, H. H. *J. Am. Chem. Soc.* **2007**, *129*, 9558–9559.
(12) Copie, V.; Faraci, W. S.; Walsh, C. T.; Griffin, R. G. *Biochemistry* **1988**, *27*, 4966–4970.
(13) McDowell, L. M.; Lee, M. S.; Schaefer, J.; Anderson, K. S. *J. Am. Chem. Soc.* **1995**, *117*, 12352–12353.
(14) Dunathan, H. C. *Proc. Natl. Acad. Sci. U.S.A.* **1966**, *55*, 712–716.
(15) Roy, M.; Keblawi, S.; Dunn, M. F. *Biochemistry* **1988**, *27*, 6698–6704.
(16) Harris, R. M.; Dunn, M. F. *Biochemistry* **2002**, *41*, 9982–9990. Harris, R. M.; Ngo, H.; Dunn, M. F. *Biochemistry* **2005**, *44*, 16886–16895.

Dierkers, A. T.; Niks, D.; Schlichting, I.; Dunn, M. F. *Biochemistry* **2009**, *48*, 10997–11010.
(17) Oldfield, E. *Philos. Trans. R. Soc., B* **2005**, *360*, 1347–1361.
(18) Vreven, T.; Morokuma, K. *J. Comput. Chem.* **2000**, *21*, 1419–1432.
(19) Frisch, M. J.; et al. *Gaussian 03*, revision B.05; Gaussian, Inc.: Wallingford, CT, 2003.
(20) Sharif, S.; Huot, M. C.; Tolstoy, P. M.; Toney, M. D.; Jonsson, K. H. M.; Limbach, H. H. *J. Phys. Chem. B* **2007**, *111*, 3869–3876. Sharif, S.; Schagen, D.; Toney, M. D.; Limbach, H. H. *J. Am. Chem. Soc.* **2007**, *129*, 4440–4455.
(21) Foster, J. P.; Weinhold, F. *J. Am. Chem. Soc.* **1980**, *102*, 7211–7218.
(22) Chan-Huot, M.; Sharif, S.; Tolstoy, P. M.; Toney, M. D.; Limbach, H.-H. *Biochemistry* [Online early access]. DOI: 10.1021/bi101061m. Published Online: Nov 10, 2010.
(23) Facelli, J. C.; Grant, D. M. *Nature* **1993**, *365*, 325–327. Olsen, R. A.; Struppe, J.; Elliott, D. W.; Thomas, R. J.; Mueller, L. J. *J. Am. Chem. Soc.* **2003**, *125*, 11784–11785. Harris, R. K. *Solid State Sci.* **2004**, *6*, 1025–1037. Salager, E.; Stein, R. S.; Pickard, C. J.; Elena, B.; Emsley, L. *Phys. Chem. Chem. Phys.* **2009**, *11*, 2610–2621.

JA106555C