Enzymes

International Edition: DOI: 10.1002/anie.201505063
German Edition: DOI: 10.1002/ange.201505063

Conversion of Anthranilate Synthase into Isochorismate Synthase: Implications for the Evolution of Chorismate-Utilizing Enzymes

*Maximilian G. Plach, Patrick Löffler, Rainer Merkl, and Reinhard Sterner**

Abstract: *Chorismate-utilizing enzymes play a vital role in the biosynthesis of metabolites in plants as well as free-living and infectious microorganisms. Among these enzymes are the homologous primary metabolic anthranilate synthase (AS) and secondary metabolic isochorismate synthase (ICS). Both catalyze mechanistically related reactions by using ammonia and water as nucleophiles, respectively. We report that the nucleophile specificity of AS can be extended from ammonia to water by just two amino acid exchanges in a channel leading to the active site. The observed ICS/AS bifunctionality demonstrates that a secondary metabolic enzyme can readily evolve from a primary metabolic enzyme without requiring an initial gene duplication event. In a general sense, these findings add to our understanding how nature has used the structurally predetermined features of enzyme superfamilies to evolve new reactions.*

Chorismate (CH) is a central metabolic branch-point molecule and the common precursor of essential primary (folate, tryptophan) and important secondary (menaquinones, siderophores, antibiotics) metabolites that are vital for plants as well as free living and infectious microorganisms[1] (Figure 1). The CH-related pathways are therefore notable targets for antimicrobials and herbicides. The enzymes catalyzing the committed steps of these pathways share a common fold and use similar reaction mechanisms. Presumably, they originated from a common ancestor and have therefore been grouped together as the MST (*m*enaquinone, *s*iderophores, *t*ryptophan) superfamily.[2] Within this superfamily, the primary metabolic anthranilate and aminodeoxychorismate synthases (AS, ADCS) employ ammonia as a nucleophile to form aminated chorismate derivatives, whereas the secondary metabolic isochorismate and salicylate synthases (ICS, SS) use water as a nucleophile to hydroxylate chorismate (Figure 1). These two subfamilies are hereafter termed ammonia-utilizing and water-utilizing MST enzymes, respectively (AMEs, WMEs). Based on the assumption that each enzyme of secondary metabolism stems from an enzyme of primary metabolism,[3] the subdivision of the MST superfamily suggests that a transition of nucleophile specificity from ammonia to water underlay the evolution of WMEs (ICS, SS) from AME (AS, ADCS) ancestors. We retraced this putative evolutionary path by identifying the residues that contribute to nucleophile specificity in MST enzymes and by subsequently establishing ICS activity on an AS scaffold.

The AS from *Salmonella typhimurium* (stAS) is a heterotetramer comprising two synthase and two glutaminase subunits (stTrpE and stTrpG, respectively).[4] Glutamine is hydrolyzed in the active site of stTrpG to yield nascent ammonia, which is subsequently channeled to the active site of stTrpE.[4,5] To identify the residues of stTrpE that are involved in this channeling and that may therefore come into contact with the ammonia nucleophile, we applied MOLE 2.0, a program for analyzing macromolecular channels.[6] We identified a 30 Å-long channel connecting the active sites in stAS (Figure 2a). This channel is similar to the channel observed in the crystal structure of the homologous aminodeoxyisochorismate (ADIC) synthase PhzE.[7] As the channel approaches the CH ligand, it is predominantly shaped by three residues: Gln263 in β-strand 11 as well as Met364 and Leu365 in α-helix 12. To estimate whether the nature of these residues correlates with nucleophile specificity in AMEs and WMEs, we computed individual multiple sequence alignments (MSAs) of ADCS, AS, ICS, and SS. Notably, these sequences formed four distinct subtrees in a phylogenetic analysis (Figure S1 in the Supporting Information), which makes them representative for their MST groups. Furthermore, all of the MSAs contained sequences of at least four major phyla. The resulting sequence logos of β-strand 11 and α-helix 12 (Figure 2b) confirmed the strict conservation of Gln263 in AS and of Lys at the corresponding position in ICS and SS, which has been noted previously.[2] It was further shown that this Lys acts as a catalytic base for water activation in ICS and SS.[2,8] In a number of ADCS sequences, Gln263 is substituted by Glu. Residues 364 and 365 are conserved to a large extent within AMEs and WMEs, but clearly differ between the two groups. In AMEs, Met364 is strictly conserved, as are Leu365 in AS and Ile365 in ADCS. In WMEs however, several hydrophobic residues (Leu, Ile, Phe, Val) are abundant at position 364, as are Val in ICS and Ser in SS at position 365. Other positions were not considered since they are either conserved throughout all four MST groups, not conserved within AMEs or WMEs, or not involved in forming the nucleophile channel in stAS.

Based on this analysis, we attempted to shift the nucleophile specificity of stAS from ammonia to water by mutating residues 263, 364, and 365 of the stTrpE subunit. For this purpose, 16 variants were generated and assayed by HPLC for the formation of reaction products starting from CH. The WME-typical catalytic Lys replaced the wild-type Gln263 in all of the variants and was combined with the different residues 364/365 found in WMEs (Table S1). The variants are hereafter denoted by their residues 263–364–365 combination

[*] M. G. Plach, P. Löffler, Prof. Dr. R. Merkl, Prof. Dr. R. Sterner
Institut für Biophysik und physikalische Biochemie
Universität Regensburg, 93040 Regensburg (Germany)
E-mail: Reinhard.Sterner@ur.de

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201505063.

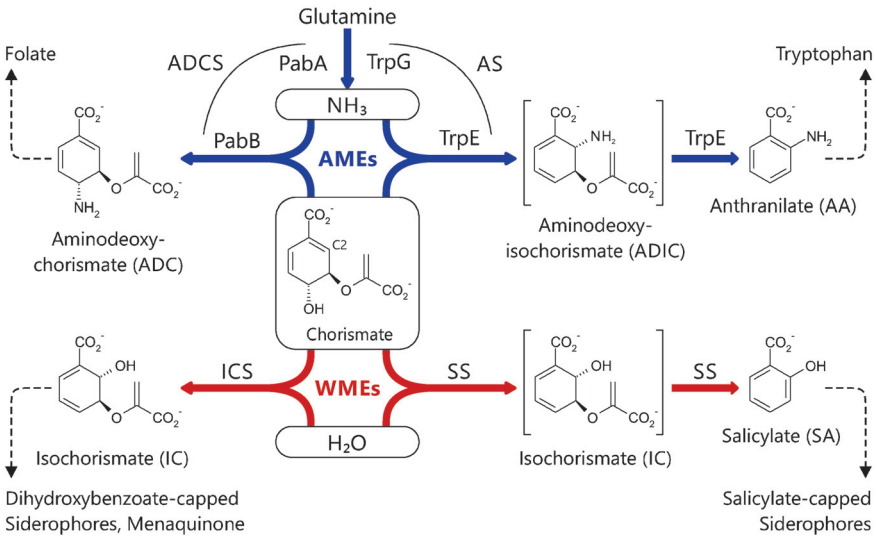

Figure 1. Chorismate is a central branch point between primary and secondary metabolism. It is converted into aminodeoxychorismate (ADC) and anthranilate (AA) by the primary metabolic ammonia-utilizing MST enzymes (AMEs) ADCS and AS. Both enzymes are heteromeric complexes comprising glutaminase (PabA, TrpG) and synthase (PabB, TrpE) subunits. ADC and AA are part of folate and tryptophan biosynthesis, respectively. The water-utilizing MST-enzymes (WMEs) ICS and SS convert chorismate into isochorismate (IC) and salicylate (SA), respectively. Both products are precursors of important secondary metabolites like iron-chelating siderophores and respiratory-chain components (menaquinone). A special characteristic of AS and SS is that pyruvate is eliminated from the reaction intermediates aminodeoxyisochorismate (ADIC) and IC to yield aromatic products.

(e.g. KML for the variant with Lys263, Met364, and Leu365). All variants expressed as soluble proteins and no adverse effects on protein stability were observed compared to wild-type stTrpE.

Twelve of these variants formed IC in the absence of an ammonia source (Figure 3a and Figure S2a). The product was unambiguously identified by ESI mass spectrometry and enzymatic conversion to SA (Figure S3). All twelve variants feature the Gln263Lys substitution. In three of these variants, the Gln263Lys substitution is combined with a mutation at position 365 (KMV, KMS, KMA). Therefore, just two mutations are sufficient for the establishment of IC specificity on the stTrpE scaffold. In nine of the variants, an additional mutation is present at position 364. Averaged over the 12 variants, 20% of supplied CH was converted to IC; the best variant, KIA, converted 37.1%. In comparison, the ICS EntC from *Escherichia coli* formed 31.5% IC, with the observed incomplete conversion being caused by equilibrium between CH and IC.[9] Upon replacing Lys263 by Ala in the IC-forming KLS variant, the resulting ALS variant became inactive, pointing to a catalytic role for Lys263 in our stAS variants similar to that in native ICS and SS.[2,8] Furthermore, no IC was formed in control reactions with wild-type stTrpE or in the absence of any enzyme.

For the KML variant, which carries only the single Gln263Lys replacement, neither IC nor SA was detected, thus indicating that Lys263 alone is not sufficient for the use of water as a nucleophile. Accordingly, the Abell group could also not detect IC or SA when characterizing the equivalent single Gln263Lys replacement in the AS from *Serratia marcescens*.[10] Ziebart and Toney reported in a comprehensive study on nucleophile specificity in MST enzymes that the Gln263Lys variant of stAS (equivalent to our KML variant) formed traces of IC and SA.[11] However, only 0.008% and 0.03% of CH were converted to IC and SA, respectively, under conditions comparable to our experimental setup. In any case, the central finding of our work, namely inversion of the nucleophile specificity of stAS through a few amino acid replacements, as exemplified by the KAS and KAA variants, is independent of whether the single Gln263Lys replacement leads to no or extremely low amounts of IC or SA.

It is worth noting that only four of the ICS-active stAS variants also formed SA, which involves the elimination of pyruvate from IC. This finding seems counterintuitive at first because the generic product of stAS is AA, which has, like SA, undergone elimination of pyruvate. However, it was recently shown that the elimination of pyruvate from chorismate-derived reaction intermediates in MST enzymes is controlled by the conformation of the chorismate ring structure.[12] Therefore, the most plausible explanation for the lack of pyruvate-elimination activity in most ICS-active stAS variants is perturbations in the IC ring conformation caused by the introduced mutations.

To examine the effects of the mutations on the use of the generic nucleophile ammonia, we investigated product formation by the stAS variants in the presence of glutamine (Figure 3b and Figure S2b). Under these conditions, all twelve ICS-active variants formed not only IC but also AA and its precursor ADIC, thus implying a broadened nucleophile specificity. Interestingly, while forming substantial amounts of AA, some variants showed only a modest reduction in IC formation compared to the absence of glutamine (KMV: 2.5-fold, KIV: 1.4-fold). Other variants, while also forming considerable amounts of AA, even showed an increase in IC formation upon the addition of glutamine (e.g., KAV: 2.2-fold).

Independent of the presence or absence of an ammonia source, a striking increase in IC formation was observed for the residue sequence Leu→Val→Ser/Ala365. Variants with Leu365 are catalytically inactive, both for IC and AA formation. The presence of Val, Ser, and Ala instead leads to average CH to IC conversions of 8.2%, 28.1%, and 24.6%, respectively.

As pointed out before, residues 263, 364, and 365 line the nucleophile channel in stTrpE. We therefore assumed that differently sized residues at these positions might reshape the channel and thus lead to the different reactivities observed. To allow a statistically sound prediction of the general

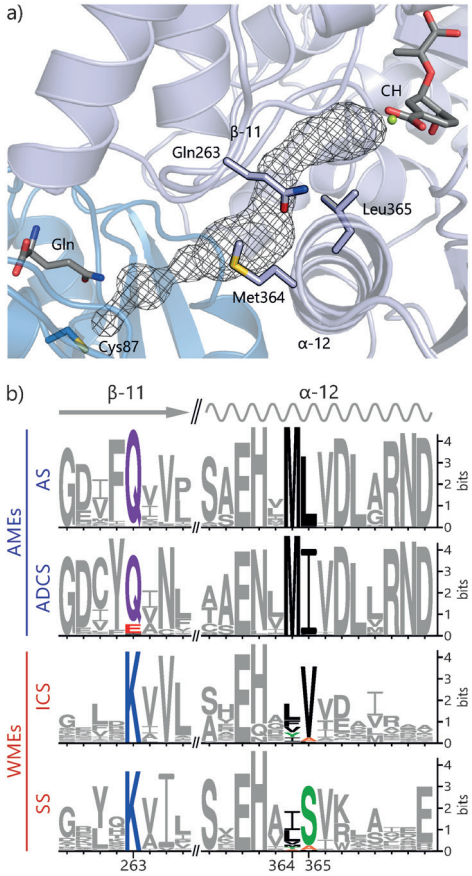

Figure 2. Location and conservation of residues characteristic for AMEs and WMEs. a) Nucleophile channel connecting the active sites of stTrpG (blue; represented by Cys87 and a glutamine ligand) and stTrpE (pale blue; represented by CH and a Mg^{2+} ion) in stAS (model based on PDB ID 1i7q). The channel-lining residues of stTrpE mutated in this work are shown as stick models. b) Sequence logos of β-strand 11 and α-helix 12. The β-strand contains a conserved Gln/Glu in AMEs and a strictly conserved Lys in WMEs. The α-helix contains the conserved Met–Leu and Met–Ile pairs in AS and ADCS, respectively. In ICS, this pair is mainly conserved as [Leu,Phe,Val]–[Val,Ala] and in SS as [Ile,Leu]–[Ser,Ala]. Residue numbering is based on stTrpE. Residues 263, 364, and 365 are colored by their chemical properties (purple: amido functionality, red: acidic, blue: basic, black: hydrophobic, green: hydroxyl functionality, orange: small).

nucleophile path in the stTrpE variants, we performed molecular dynamics (MD) simulations and applied MOLE to generate 600 putative nucleophile trajectories (PNTs) for each variant. A PNT is defined by the centerline of the corresponding MOLE channel.

We found that a reduction in the size of residue 365 leads to a shift in PNT localization in a 10 Å shell around CH (Figure 4a). In wild-type stTrpE, PNTs pass Met364 and Gln263, bend around Leu365, and end at C2 of CH, which is where the nucleophilic attack by ammonia takes place. The single Gln263Lys mutation in the KML variant does not alter this course. Since both KML and wild-type stTrpE feature Leu365, this course of PNTs was termed the L-path (Figure 4b). In the KML variant, however, the PNT-associated channels are clearly constricted between Lys263, Leu365, and Val265 owing to a change in the van der Waals volume of residue 263 from 114 $Å^3$ (Gln) to 135 $Å^3$ (Lys).[13] Ammonia or water therefore cannot access CH via these PNTs, thus resulting in the observed loss of activity in the Leu365 variants. Replacing Leu365 by Val (volume reduction from 124 $Å^3$ to 105 $Å^3$) has two consequences for the PNTs: The associated channels of L-path PNTs are enlarged and 31 % of the PNTs shift away from the L-path and approach CH from Val365 and Thr425. This effect becomes even more pronounced for the KMS variant with Ser365, the low volume (73 $Å^3$) of which allows 79 % of the PNTs to approach CH via this alternative path. Finally, Ala365 (67 $Å^3$) leads to a complete shift of PNTs to this path, which was therefore termed the A-path (Figure 4b).

For the other twelve stAS variants, MD simulations and computations of PNTs were performed in a similar manner (Figure S4, Table S2). The PNT distributions in these variants are consistent with those described above for the KML, KMV, KMS, and KMA variants. Taken together, it is evident that high CH to IC conversion rates correlate with high fractions of PNTs proceeding along the A-path (Figure 4c).

These findings provide a conclusive view on water utilization by WMEs. Ziebart and Toney suggested that two properties of WMEs promote the water-activating capacity of Lys263.[11] First, two hydrogen-bond acceptors were thought to assist in deprotonating the ε-amino group. However, one of these acceptors, a glutamic acid residue, is conserved not only in WMEs but in all but one of the 1222 MST enzymes in our dataset. The other, a serine residue corresponding to Ser365, is not conserved in SS (7 % Ala) and is not present at all in ICS. Furthermore, we could show that Ser365 is not required for IC formation (Figure 4c). The second element proposed was the residue corresponding to Met364 in stTrpE. It is conserved as a hydrophobic residue (Leu, Ile, Val, Phe) in WMEs and was thought to assist in the proper orientation of Lys263 by making van der Waals contact with the Lys methylene groups. However, variants with wild-type Met364 or Ala364 display IC specificity as well (Figure 4c), which argues against a role for residue 364 in positioning the catalytic Lys.

In summary, we have demonstrated that nucleophile specificity in MST enzymes is controlled by two factors: 1) the presence or absence of Lys263 as a catalytic base and 2) sufficient space for the nucleophile to reach CH. These findings agree well with the amino acid distribution in naturally occurring ICS and SS. Only five ICS and no SS in our MST dataset contain a residue larger than Val at position 365. Moreover, the assignment of Leu365 in those five ICS sequences is most likely due to a misalignment, since in all cases, Val directly precedes this Leu.

The straightforward establishment of ICS activity on the evolutionarily related AS scaffold indicates that the MST-superfamily represents a flexible and adaptive link between primary and secondary metabolism. It is generally accepted that each secondary metabolic pathway has its origin in mutations that accumulated in a primary metabolic gene. Consequently, these mutations should 1) enhance overall metabolic chemical diversity by yielding molecules with novel biological activities, 2) allow the retention of existing chemical variety, and 3) do so with minimum fitness costs.[3a] Our

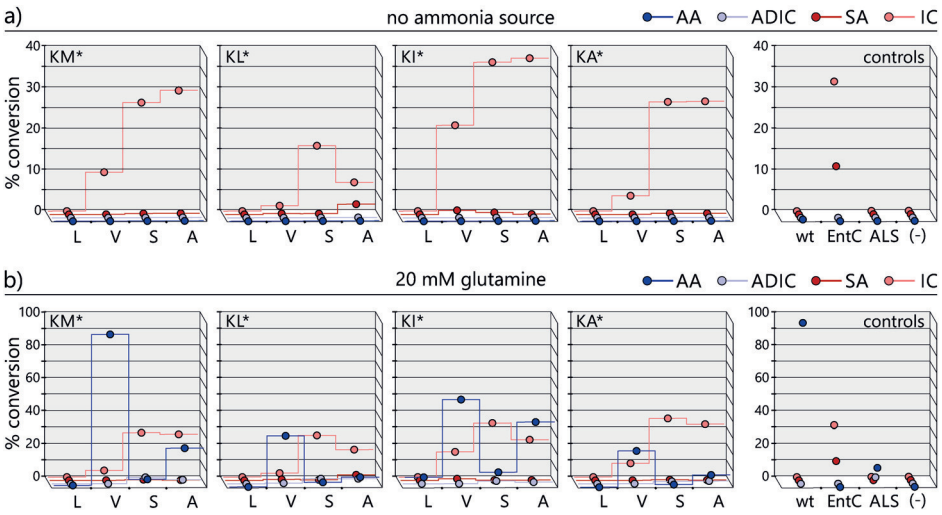

Figure 3. Product distributions of reactions of stTrpE variants with chorismate in the absence (a) and presence (b) of glutamine as an ammonia source. The variants share Lys263 and are grouped by residue 364 (e.g., KM* describes the four variants with Lys263, Met364, and either Leu, Val, Ser, or Ala at position 365). All stTrpE variants were assayed in the presence of stTrpG. The product distributions of control reactions are shown on the right: wild-type stTrpE (wt), ecEntC (EntC), Lys263Ala knockout of the KLS variant (ALS), and enzyme-free control (−). For each product, its fraction out of all of the products formed is shown (% conversion). Each data point represents the mean of at least three independent experiments. Error bars are omitted for clarity because the average and maximum absolute errors were only 0.3% and 3.7%, respectively. Step lines connecting data points were added to aid visual tracking of product-specific values.

data strongly support this hypothesis. Only two mutations are necessary to establish ICS activity and all these stTrpE variants were still able to employ ammonia as a nucleophile to form AA for tryptophan biosynthesis. Following this reasoning, the evolution of IC formation does not necessarily require gene duplication but could have proceeded via a bifunctional intermediate similar to the stTrpE variants of this work.

Substrate or nucleophile channeling, which is crucial for enzymatic function in MST enzymes, is no rare phenomenon. More than 64% of all known enzyme structures contain at least two channels reaching to their active sites.[14] Our approach may thus prove valuable for modifying the reactivities of such enzymes, which could ultimately yield novel enzymatic functions.

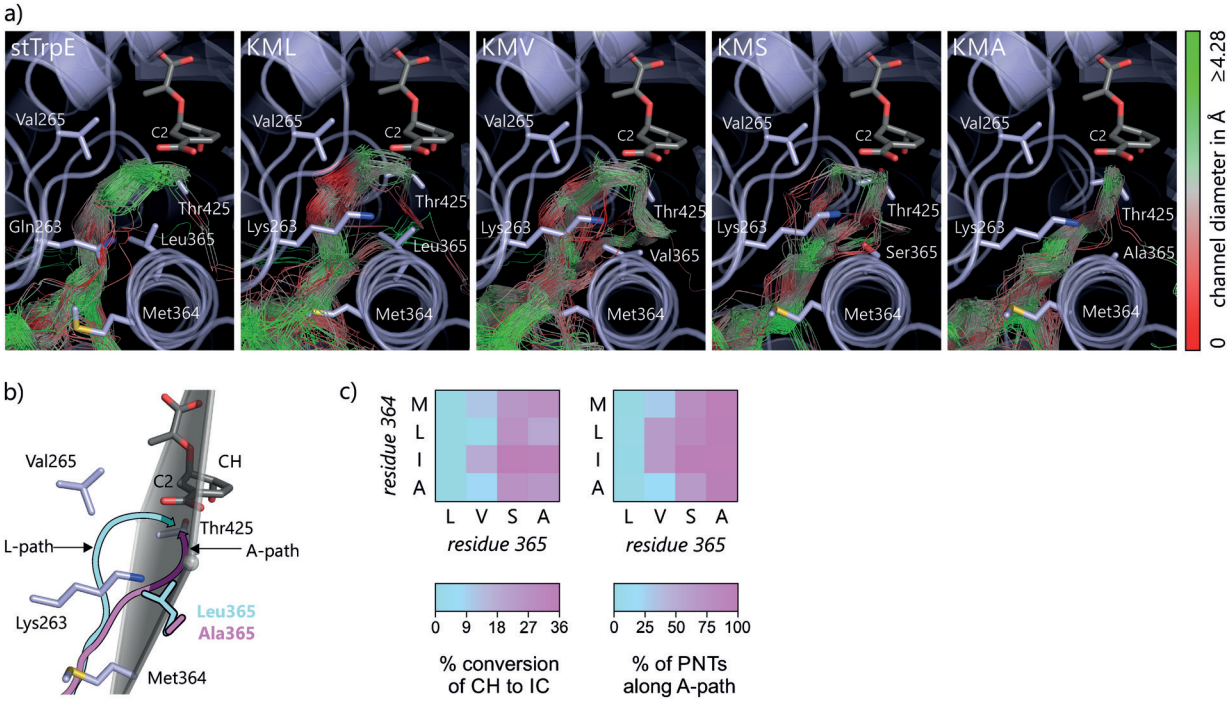

Figure 4. Visualization and quantitative analysis of PNTs in wild-type and mutant stTrpE variants. a) Comparison of PNTs in wild-type stTrpE and the KML, KMV, KMS, and KMA variants. The diameters of the PNT-associated channels are color-coded and important channel-lining residues are shown as stick models. The key observation is the correlation between decreasing size of residue 365 (Leu > Val > Ser > Ala) and a shift in PNT localization from the L-path to the A-path [see (b)]. b) The L-path (cyan) and A-path (magenta) reflect the majority of PNTs in variants with Leu365 and Ala365, respectively, and thus show the boundaries of the PNT shift. The directions from which the two paths approach CH can be separated by a plane specified by CH-C2 and the Cα-atoms of Met364 and Thr425 (for details see the methods section in the Supporting Information). c) Comparison of the average CH-to-IC conversion by all stTrpE variants with Lys263 and the fraction of PNTs proceeding along the A-path in these variants.

Acknowledgements

M.G.P. was supported by a PhD fellowship from the Fonds der Chemischen Industrie. We thank Kristina Heyn for help with phylogenetic analyses and Christiane Endres, Sonja Fuchs, and Jeannette Ueckert for expert technical assistance.

Keywords: chorismate · enzyme catalysis · enzyme evolution · nucleophile specificity · protein design

How to cite: *Angew. Chem. Int. Ed.* **2015**, *54*, 11270–11274
Angew. Chem. **2015**, *127*, 11423–11427

[1] F. Dosselaere, J. Vanderleyden, *Crit. Rev. Microbiol.* **2001**, *27*, 75–131.
[2] S. Kolappan, J. Zwahlen, R. Zhou, J. J. Truglio, P. J. Tonge, C. Kisker, *Biochemistry* **2007**, *46*, 946–953.
[3] a) R. D. Firn, C. G. Jones, *Mol. Microbiol.* **2000**, *37*, 989–994; b) L. C. Vining, *Gene* **1992**, *115*, 135–140.
[4] A. A. Morollo, M. J. Eck, *Nat. Struct. Biol.* **2001**, *8*, 243–247.
[5] A. A. Morollo, R. Bauerle, *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 9983–9987.
[6] D. Sehnal, R. Svobodova Varekova, K. Berka, L. Pravda, V. Navratilova, P. Banas, C. M. Ionescu, M. Otyepka, J. Koca, *J. Cheminf.* **2013**, *5*, 39.
[7] Q. A. Li, D. V. Mavrodi, L. S. Thomashow, M. Roessle, W. Blankenfeldt, *J. Biol. Chem.* **2011**, *286*, 18213–18221.
[8] J. Zwahlen, S. Kolappan, R. Zhou, C. Kisker, P. J. Tonge, *Biochemistry* **2007**, *46*, 954–964.
[9] J. Liu, N. Quinn, G. A. Berchtold, C. T. Walsh, *Biochemistry* **1990**, *29*, 1417–1425.
[10] O. Kerbarh, D. Y. Chirgadze, T. L. Blundell, C. Abell, *J. Mol. Biol.* **2006**, *357*, 524–534.
[11] K. T. Ziebart, M. D. Toney, *Biochemistry* **2010**, *49*, 2851–2859.
[12] J. E. Culbertson, D. H. Chung, K. T. Ziebart, E. Espiritu, M. D. Toney, *Biochemistry* **2015**, *54*, 2372–2384.
[13] N. J. Darby, T. E. Creighton, *Protein structure*, Oxford University Press, Oxford, UK, **1993**.
[14] L. Pravda, K. Berka, R. Svobodova Varekova, D. Sehnal, P. Banas, R. A. Laskowski, J. Koca, M. Otyepka, *BMC Bioinf.* **2014**, *15*, 379.

Received: June 3, 2015
Revised: June 23, 2015
Published online: August 6, 2015

