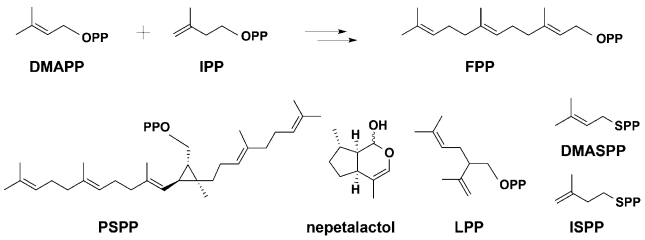
VIP Monoterpenes Very Important Paper

International Edition: DOI: 10.1002/anie.201600656
German Edition: DOI: 10.1002/ange.201600656

Structure and Function of a "Head-to-Middle" Prenyltransferase: Lavandulyl Diphosphate Synthase

Meixia Liu[+], Chun-Chi Chen[+], Lu Chen[+], Xiansha Xiao, Yingying Zheng, Jian-Wen Huang, Weidong Liu, Tzu-Ping Ko, Ya-Shan Cheng, Xinxin Feng, Eric Oldfield, Rey-Ting Guo,* and Yanhe Ma**

Abstract: *We report the first X-ray structure of the unique "head-to-middle" monoterpene synthase, lavandulyl diphosphate synthase (LPPS). LPPS catalyzes the condensation of two molecules of dimethylallyl diphosphate (DMAPP) to form lavandulyl diphosphate, a precursor to the fragrance lavandulol. The structure is similar to that of the bacterial cis-prenyl synthase, undecaprenyl diphosphate synthase (UPPS), and contains an allylic site (S1) in which DMAPP ionizes and a second site (S2) which houses the DMAPP nucleophile. Both S-thiolo-dimethylallyl diphosphate and S-thiolo-isopentenyl diphosphate bind intact to S2, but are cleaved to (thio)diphosphate, in S1. His78 (Asn in UPPS) is essential for catalysis and is proposed to facilitate diphosphate release in S1, while the P1 phosphate in S2 abstracts a proton from the lavandulyl carbocation to form the LPP product. The results are of interest since they provide the first structure and structure-based mechanism of this unusual prenyl synthase.*



Scheme 1. Structures of some isoprenoid compounds discussed in the text; OPP = diphosphate.

There are >65000 terpene or isoprene-like compounds known.[1] In general, there are three main classes of enzymes[2] involved in their biosynthesis: *trans*-prenyltransferases such as farnesyl diphosphate (FPP; Scheme 1) synthase, involved in for example, protein prenylation and quinone biosynthesis; *cis*-prenyltransferases such as undecaprenyl diphosphate synthase (UPPS), involved in bacterial cell wall biosynthesis; and a diverse range of terpene cyclases, some of which are modular proteins.[3] With the *cis* and *trans* prenyltransferases, an allylic diphosphate condenses with the homoallylic species isopentenyl diphosphate (IPP, Scheme 1) to form longer-chain isoprenoids such as FPP, many of which are then acted on by other synthases. For example, squalene synthase produces presqualene diphosphate (PSPP, Scheme 1) and then squalene, a precursor to sterols. There are also many more unusual prenyl synthases. For example, iridoid synthase converts 8-oxo-geranial to nepetalactol (Scheme 1), a key component in formation of the indole-terpenoid drugs vincristine and vinblastine.[4,5] There are also "irregular" terpene synthases such as lavandulyl diphosphate synthase (LPPS) in which two DMAPP molecules—both allylic diphosphates—condense via a so-called "head-to-middle" condensation[6] to form the (C_{10}) monterpene, lavandulyl diphosphate (LPP, Scheme 1), the precursor of the fragrances (*R*)-lavandulol and (*R*)-lavandulyl acetate. The structure of this enzyme is unknown. Based on a BLAST (basic local alignment search tool)[7] search of all reported genomes, LPPS has closest homology (ca. 47% identity over 234 residues) to (plant) dehydrodolichyl diphosphate synthases, but none of these structures have been reported. When compared with just bacterial genomes, there is also close homology to undecaprenyl diphosphate synthases (ca. 42% identity) for a catalytic "core" of 230 LPPS residues, and the structures of several UPPS proteins are known.[8,9] However, it has also been shown that LPP can be made by a chimeric terpenoid synthase with the conventional "FPPS-like" fold[10] (PDB ID code 4KK2). Here, we sought to determine the structure of LPPS (from *Lavandula x intermedia*); to determine how ligands bind; and to use this and other information to propose a structural basis for the LPPS mechanism of action.

[*] M. Liu[+]
College of Biotechnology, Tianjin University of Science and Technology, Tianjin 300457 (China)
Prof. Dr. C.-C. Chen,[+] X. Xiao, Prof. Dr. Y. Zheng, Prof. Dr. W. Liu, Y.-S. Cheng, Prof. Dr. R.-T. Guo, Prof. Dr. Y. Ma
Industrial Enzymes National Engineering Laboratory, Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences
Tianjin 300308 (China)
E-mail: guo_rt@tib.cas.cn
ma_yh@tib.cas.cn
Dr. T.-P. Ko
Institute of Biological Chemistry, Academia Sinica
Taipei 11529 (Taiwan)
Dr. J.-W. Huang
AsiaPac Biotechnology Co., Ltd.
Dongguan, 523808 (China)
L. Chen[+]
Department of Biochemistry, University of Illinois
Urbana, IL 61801 (USA)
Dr. X. Feng, Prof. Dr. E. Oldfield
Department of Chemistry, University of Illinois
Urbana, IL 61801 (USA)
E-mail: eo@chad.scs.uiuc.edu
[+] These authors contributed equally to this work.

Supporting information for this article can be found under http://dx.doi.org/10.1002/anie.201600656.

We first attempted to crystallize LPPS using a full-length (316 residue) construct, that is, containing the putative catalytic domain (261 residues) together with the chloroplast-targeting[11] N-terminus (55 residues). We were not successful in obtaining soluble protein, due perhaps to the presence of the lipid (membrane) targeting domain. We thus next expressed truncated proteins with 30, 50, or 55 N-terminal residues deleted. The Δ30- and Δ50 proteins were unstable but the Δ55 protein was successfully expressed, and crystallized. Crystals diffracted to 1.87 Å resolution and full crystallographic data acquisition and structure refinement details (for three structures) are given in Table S1 in the Supporting Information. LPPS crystallized in the orthorhombic space group $C222_1$ and the structure of the dimeric *apo*-protein (PDB ID code 5HC6) is shown in Figure 1 A.

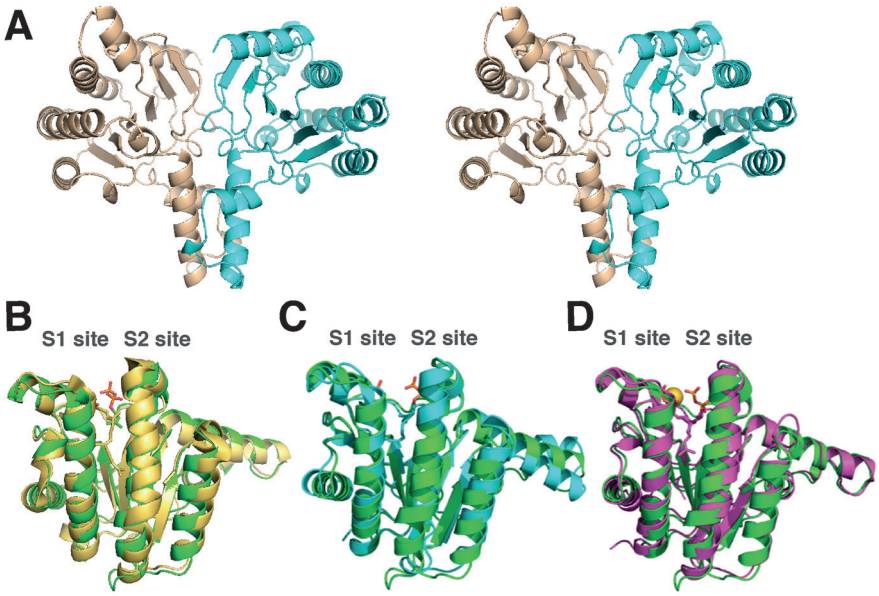


Figure 1. Structure of *Lavandula x intermedia* LPPS and comparisons with the structures of three *cis*-prenyltransferases. A) Structure of the dimeric protein in complex with sulfate (PDB ID code 5HC6). B) Structural superimposition of LPPS (green; PDB ID code 5HC6) with EcUPPS in complex with IPP and FPP (yellow; PDB ID codes 1X08 and 1X09). C) As (B) but with *M. tuberculosis cis*-farnesyl diphosphate synthase in complex with FPP, Pi (Rv1086, PDB ID code 2VG1, cyan). D) As (B) but with *cis*-decaprenyl diphosphate synthase in complex with citronellyl diphosphate (Rv2361, PDB ID code 2VG3, magenta).

Using the PDBeFold server[12] we find that this (plant) LPPS has high structural similarity to the bacterial protein UPPS from both *Staphylococcus aureus* (PDB ID code 4U82, a 1.16 Å Cα rmsd over 232 aligned residues) and *E. coli* (PDB ID code 1X06, a 1.11 Å Cα rmsd over 225 aligned residues). A structural superposition of *apo*-LPPS with (FPP and IPP-liganded) EcUPPS is shown in Figure 1 B in which the two ligand binding sites in UPPS—the allylic S1 site into which FPP binds and the IPP site S2—are indicated. Analogous sites are present in the LPPS structure and are occupied by sulfate (Figure S1) and the LPPS structure is also very similar to those of the *cis*-FPP and *cis*-decaprenyl diphosphate synthases from *Mycobacterium tuberculosis* (Figure 1 C,D).

These structural similarities are of interest since the mechanism of action of UPPS is well understood.[13] In UPPS, the allylic species (initially, FPP) binds to the allylic site S1 with both phosphates in its diphosphate group bound to a single Mg^{2+} which in turn is coordinated to a highly conserved Asp (D26 in *E. coli* UPPS), which based on sequence and structural alignments (Figure S2) is also present in LPPS (D76, due to the presence of the N-terminal plastid-targeting sequence), as well as in the homologous *M. tuberculosis* proteins, and *Staphylococcus aureus* UPPS (Figure S2). In UPPS, there is also a second, homoallylic binding site, S2, that houses the IPP used in chain elongation. Mechanistically, in UPPS the allylic diphosphate in S1 is attacked by the IPP double bond (in S2) in a concerted (that is, not sequential) manner, resulting in diphosphate loss and chain elongation. Logically, then, it would follow that in LPPS, one DMAPP should bind to the S1 allylic site with multiple interactions between its diphosphate and Mg^{2+}, facilitating PPi removal, while the second DMAPP would bind to S2. However, UPPS does not condense DMAPP to LPP, and LPPS does not condense DMAPP + IPP to longer chain species, so there must be important differences in active site structure/ligand binding between the two proteins. We thus next sought ligand-bound structures.

We soaked LPPS crystals with S-thiolo-dimethylallyl diphosphate (DMASPP, Scheme 1), and for comparison, S-thiolo-isopentenyldiphosphate (ISPP, Scheme 1). Full data acquisition and structure refinement details for the DMASPP (PDB ID code 5HC8) and ISPP (PDB ID code 5HC7) complex structures are given in Table S1. The complex structures are very similar to the *apo* structure (PDB ID code 5HC6) with ca. 0.1 Å Cα rmsd values, indicating that protein folding is not significantly altered on ligand binding. There were, however, some surprises. With DMASPP, the ligand electron density in the S2 site is well defined (Figure 2 A) and there is clearly a bond between one phosphate and Mg^{2+} (which is also coordinated to D76 and two water molecules). However, equally clearly, there is a break in the electron density for the species in the allylic site, S1, Figure 2 A. There is also a well defined diphosphate (presumably actually a thiodiphosphate) group in S1 bound via both phosphates to Mg^{2+}, together with a planar C_5 species below the diphosphate and with no connecting electron density. Only two sulfates can be modeled into the pyrophosphate binding site in *apo* LPPS crystals prior to soaking with DMASPP (Figure S1) so it appears that the DMASPP in S1 is unusually reactive, ionizing to form PPi and the dimethylallyl carbocation which then looses a proton to form the planar C_5 species,

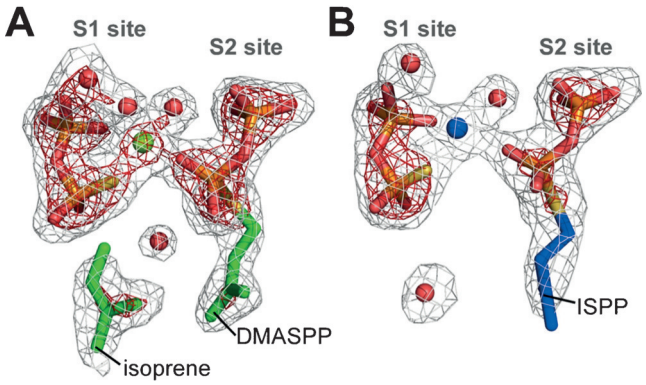


Figure 2. Electron density maps of bound ligands (stick) and Mg^{2+} (sphere), contoured at 1.0 σ (gray) and 3.0 σ (red). A) DMASPP; PDB ID code 5HC8. B) ISPP; PDB ID code 5HC7.

isoprene. With ISPP (Figure 2B) the electron density results show a well defined (ISPP) ligand bound in the S2 site but only a PPi or SPPi (bound to Mg^{2+}) in S1, together with a small additional density due, perhaps, to an H_2O (Figure 2B). This result was even more surprising than the DMASPP result because the homoallylic species ISPP is expected to be less reactive chemically than the allylic DMASPP.

We thus determined a second ISPP structure using a commercial sample of ISPP and obtained the same results (Figure S3). It thus appears that the S1 site in LPPS catalyzes diphosphate ionization with both DMASPP and ISPP, facilitating perhaps, the "head-to-middle" condensation reaction with DMAPP. There was no organic fragment observed with ISPP (on SPPi loss), suggesting formation of a polar (isopentenyl alcohol) product that would not bind tightly to the hydrophobic S1 site. Why then might DMASPP and ISPP be cleaved, in S1? And how might this relate to catalytic activity?

On inspection of the amino acid sequences (Figure S2), as well as the LPPS/UPPS X-ray structures (Figure 3), we see that there is one major difference (in the S1 active site region) between LPPS and the other *cis*-prenyltransferases. Specifically, LPPS contains an active site His (His78), but in all of the other proteins (Figure S2) the corresponding residue is a (neutral) asparagine. The other, basic, active site residues in LPPS and the structurally homologous *cis*-prenyltransferases are all located in similar spatial positions in both the S1 and S2 sites (Figure 3), suggesting that His78 might be important in catalyzing the LPPS reaction, by facilitating PPi loss. To test this hypothesis, we produced the H78N LPPS mutant (since Asn is present in UPPS). There was no lavandulyl diphosphate formation found (Figure S4). We also produced Y139F and W100V mutants, based on the UPPS structures. Y139F had 62% the activity of WT while W100V had 15% activity, so these mutations do have effects on folding/activity, but clearly only H78 is essential for catalysis (Figure S4).

The X-ray crystallographic, sequence homology and mutagenesis results thus lead to the LPPS mechanism of action shown in Figure 4. Figure 4A shows a view of the LPPS active site region in which, for clarity in explaining the

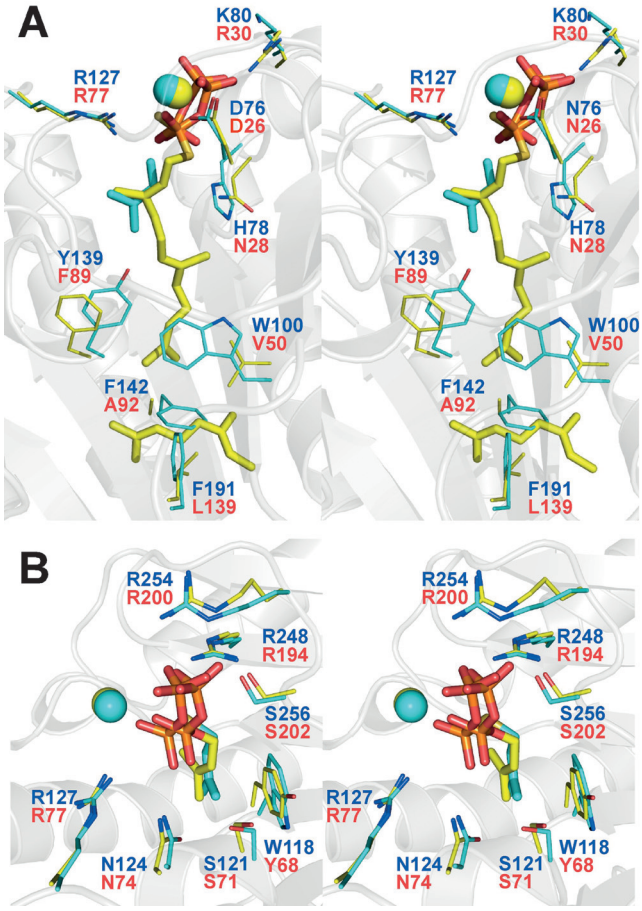


Figure 3. Stereoview of the S1 and S2 site-forming residues in LPPS (cyan, PDB ID code 5HC8, blue letters) and EcUPPS (yellow, PDB ID code 1X08, red letters). A) S1 site. B) S2 site.

catalytic mechanism, we have connected the PPi and C_5 fragments in the DMASPP structure to model DMAPP. A ChemDraw version of this substrate-bound model is shown in Figure 4B and the product-bound model is shown in Figure 4C. The proposed mechanism is as follows. First, loss of PPi from the DMAPP in S1 is facilitated by the presence of the basic residue His78, Figure 4B, since there are 2.8 Å and 4.1 Å interactions between the H78Nδ1 and (P1)O1,(P1)O2, and a 2.7 Å interaction between R127Nε and (P1)O2. The double bond in the DMAPP in S2 then attacks the carbocation in S1 forming the "lavandulyl carbocation"[6]—although a concerted reaction cannot be ruled out. There are no basic residues near the DMASPP in S2 to effect the required proton abstraction to form LPP (Figure S5) but one oxygen in the proximal phosphate is only 3 Å from one of the DMAPP methyl protons (in the DMASPP structure), suggesting this phosphate acts as the base-just as proposed for the proximal DMAPP phosphate in the isoprene synthase reaction.[14] The DMASPP (S1; C1) to DMASPP (S2; C2) distance is long (5.4 Å)—but this is actually the same distance as found between the two FPP (FSPP) molecules in dehydrosqualene synthase[15] which condense to form PSPP—so such large ligand movements do have precedent in prenyl synthases. In addition, based on this structural model, the product is (*R*)-lavandulyl diphosphate, as found experimentally.

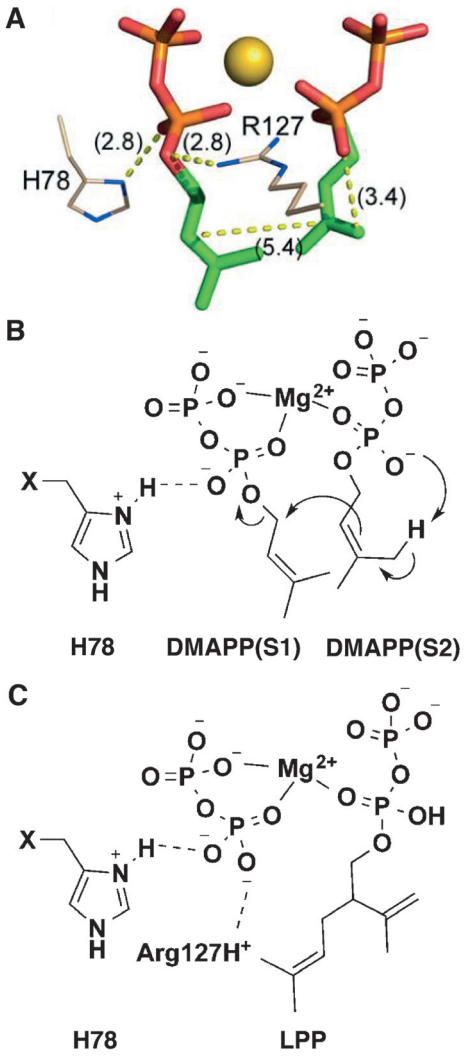


Figure 4. Proposed LPPS mechanism of action. A) A view of the LPPS active site. B) Schematic illustration of the LPPS mechanism based on the structure in (A). C) Product-bound structural model.

In summary: we have obtained the first structure of a "head-to-middle" isoprene synthase, lavandulyl diphosphate synthase, and propose a detailed structure-based mechanism of action. The structure of this plant enzyme is remarkably similar to that found in many bacterial *cis*-prenyl diphosphate synthases. The S-thiolo analogs of DMAPP and IPP bind to the "S2" site occupied by IPP in the conventional *cis*-prenyltransferases, but with both ligands the S1 site contains a diphosphate (bound to Mg^{2+}) together with, in the case of DMASPP, a planar C_5 species. The most obvious structural/sequence difference between LPPS and the other *cis*-prenyltransferases in the active site region is the presence of an active-site His (H78) in LPPS, which is Asn in the other proteins. An H78N mutant did not catalyze LPP production, consistent with a key role for this His in catalysis. There were no basic residues close to the S2 DMAPP methyl groups, suggesting that the proximal phosphate group acts as the base needed to deprotonate the "lavandulyl carbocation" intermediate—basically the same role as the DMAPP (di)phosphate in isoprene synthase. Overall, the results are of broad general interest since we have determined the first structure and structure-based mechanism of action of the irregular terpene synthase, lavandulyl diphosphate synthase, involved in formation of the commercially important fragrances, lavandulol and lavandulyl acetate.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (grants 31500642, 31570130, 31300615, 31400678 and 31470240), the United States Public Health Service (NIH grants GM065307, CA158191), by a Harriet A. Harlin Professorship, and by the University of Illinois/Oldfield Research Fund. We thank the National Synchrotron Radiation Research Center of Taiwan for beam time allocation and data collection assistance.

Keywords: biosynthesis · fragrance · monoterpenes · natural products · structure elucidation

How to cite: *Angew. Chem. Int. Ed.* **2016**, *55*, 4721–4724
Angew. Chem. **2016**, *128*, 4799–4802

[1] J. Buckingham, *Dictionary of Natural Products on DVD*, CRC Press, **2007**.
[2] E. Oldfield, F. Y. Lin, *Angew. Chem. Int. Ed.* **2012**, *51*, 1124–1137; *Angew. Chem.* **2012**, *124*, 1150–1163.
[3] M. Köksal, Y. Jin, R. M. Coates, R. Croteau, D. W. Christianson, *Nature* **2011**, *469*, 116–120.
[4] H. Kries, L. Caputi, C. E. Stevenson, M. O. Kamileen, N. H. Sherden, F. Geu-Flores, D. M. Lawson, S. E. O'Connor, *Nat. Chem. Biol.* **2016**, *12*, 6–8.
[5] Y. Hu, W. Liu, S. R. Malwal, Y. Zheng, X. Feng, T.-P. Ko, C.-C. Chen, Z. Xu, M. Liu, X. Han, J. Gao, E. Oldfield, R.-T. Guo, *Angew. Chem. Int. Ed.* **2015**, *54*, 15478–15482; *Angew. Chem.* **2015**, *127*, 15698–15702.
[6] Z. A. Demissie, L. A. E. Erland, M. R. Rheault, S. S. Mahmoud, *J. Biol. Chem.* **2013**, *288*, 6333–6341.
[7] S. F. Altschul, T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller, D. J. Lipman, *Nucleic Acids Res.* **1997**, *25*, 3389–3402.
[8] R. T. Guo, R. Cao, P. H. Liang, T. P. Ko, T. H. Chang, M. P. Hudock, W. Y. Jeng, C. K. Chen, Y. Zhang, Y. Song, C. J. Kuo, F. Yin, E. Oldfield, A. H. Wang, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 10022–10027.
[9] W. Zhu, Y. Wang, K. Li, J. Gao, C. H. Huang, C. C. Chen, T. P. Ko, Y. Zhang, R. T. Guo, E. Oldfield, *J. Med. Chem.* **2015**, *58*, 1215–1227.
[10] H. V. Thulasiram, H. K. Erickson, C. D. Poulter, *Science* **2007**, *316*, 73–76.
[11] O. Emanuelsson, H. Nielsen, G. von Heijne, *Protein Sci.* **1999**, *8*, 978–984.
[12] Protein structure comparison service PDBeFold at European Bioinformatics Institute (http://www.ebi.ac.uk/msd-srv/ssm), authored by E. Krissinel, K. Henrick.
[13] Y. P. Lu, H. G. Liu, K. H. Teng, P. H. Liang, *Biochem. Biophys. Res. Commun.* **2010**, *400*, 758–762.
[14] M. Köksal, I. Zimmer, J. P. Schnitzler, D. W. Christianson, *J. Mol. Biol.* **2010**, *402*, 363–373.
[15] C.-I. Liu, G. Y. Liu, Y. Song, F. Yin, M. E. Hensler, W.-Y. Jeng, V. Nizet, A. H. J. Wang, E. Oldfield, *Science* **2008**, *319*, 1391–1394.

Received: January 20, 2016
Published online: February 29, 2016

