

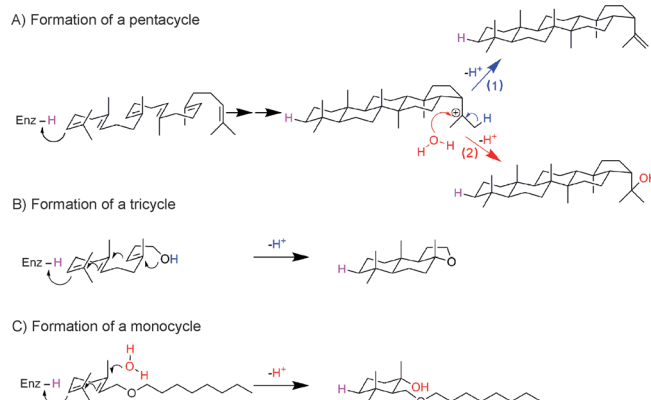
Entropy is Key to the Formation of Pentacyclic Terpenoids by Enzyme-Catalyzed Polycyclization**

Per-Olof Syrén,* Stephan C. Hammer, Birgit Claasen, and Bernhard Hauer

Abstract: Polycyclizations constitute a cornerstone of chemistry and biology. Multicyclic scaffolds are generated by terpene cyclase enzymes in nature through a carbocationic polycyclization cascade of a prefolded polyisoprene backbone, for which electrostatic stabilization of transient carbocationic species is believed to drive catalysis. Computational studies and site-directed mutagenesis were used to assess the contribution of entropy to the polycyclization cascade catalyzed by the triterpene cyclase from *A. acidocaldarius*. Our results show that entropy contributes significantly to the rate enhancement through the release of water molecules through specific channels. A single rational point mutation that results in the disruption of one of these water channels decreased the entropic contribution to catalysis by 60 kcal mol⁻¹. This work demonstrates that entropy is the key to enzyme-catalyzed polycyclizations, which are highly relevant in biology since 90 % of all natural products contain a cyclic subunit.

Terpenoids have central functions for all living organisms and are a diverse class of biologically active compounds that includes potent anticancer, antiviral, antimicrobial, and antifungal agents. The structural diversity of terpenoids originates from an enzyme-catalyzed carbocationic polycyclization cascade of linear polyisoprenes that are all derived from isopentenyl diphosphate and its regioisomer dimethylallyl diphosphate.^[1] The cyclization cascade is initiated either by the metal-dependent ionization of a labile allylic diphosphate group (class I mechanism) or by the protonation of the terminal isoprene double bond (class II mechanism). Pentacyclic chemistry based on the latter was already established during the Archean eon^[2] by triterpene cyclases. Terpene cyclases have provided access to new potential biofuels^[3] and

therapeutics for cancer^[4] and malaria^[5] and they display a high synthetic potential^[6] for the generation of novel terpenoid-based compounds. Hence it is of uttermost importance to enhance our understanding of the reaction mechanism displayed by these enzymes, which catalyze one of the most important and complex transformations known in biochemistry, one that has fascinated researchers for decades.^[7,8] It is known that these enzymes bind and stabilize the flexible substrate in the precise orientation required for catalysis through a unique cyclization sequence, trigger carbocation formation, and stabilize carbocations against premature quenching of water molecules present in the active site.^[8–11] Pre-arranging the substrate for ring closure will inevitably be associated with a high entropic penalty because of the lost internal rotational degrees of freedom.^[12] It is well known that this entropic cost plays an important role in intramolecular cyclization reactions, such as macrolactonizations,^[13] that have highly ordered transition-state structures. We were puzzled by how enzyme-catalyzed concerted polycyclizations can overcome the high entropic cost of prefolding the substrate for polycyclization. This issue is addressed in the present investigation, which is based on experimental and theoretical work with a thermophilic triterpene cyclase (Scheme 1). These enzymes function by the class II mechanism in analogy to Brønsted acid catalysis, which is the



Scheme 1. The cyclization of polyisoprene backbones of different lengths by *Alicyclobacillus acidocaldarius* triterpene cyclase. A) The polycyclization of squalene (C30) yields the pentacyclic hopenyl cation (possibly via a tricyclic core and subsequent ring expansion^[9]). This reacts to give hopene through pathway 1 (blue) or hopanol through termination by water addition (pathway 2, red). B) The cyclization of homofarnesol (C16) yields ambronan with a tricyclic C6–C6–C5 skeleton. C) The cyclization of geranyl octyl ether yields a monocyclic water addition product. Enz–H refers to the catalytic acid in the triterpene cyclase active site that initiates the polycyclization cascade by protonating the terminal isoprene group.

[*] Dr. P.-O. Syrén,^[†] S. C. Hammer, Prof. Dr. B. Hauer
Institute of Technical Biochemistry, University of Stuttgart
Allmandring 31, 70569 Stuttgart (Germany)
E-mail: per-olof.syren@biotech.kth.se

Dr. B. Claasen
Institute for Organic Chemistry
Pfaffenwaldring 55, 70569 Stuttgart (Germany)

[†] Present address: Division of Proteomics
KTH Royal Institute of Technology, AlbaNova University Center
Roslagstullsbacken 21, 10691 Stockholm (Sweden)

[**] The Alexander von Humboldt foundation, the Royal Swedish Academy of Sciences, Fonds der Chemischen Industrie, and the European Union's Seventh Framework Programme FP7/2007-2013 under grant agreement no 289646 are gratefully acknowledged for financial support.

Supporting information for this article (including experimental details) is available on the WWW under <http://dx.doi.org/10.1002/anie.201402087>.

foundation of fundamental organic transformations^[14] including biomimetic cyclization reactions.^[15] Hence the mechanistic study herein could have important implications for synthetic chemistry.

The entropic penalty associated with forming a productive precyclic substrate conformation is well recognized and amounts up to 43 cal mol⁻¹ K⁻¹ (14 kcal mol⁻¹ at 328 K) even for monocycles.^[13,16] This energetic cost would slow down the rate 10⁹-fold compared to that of a reaction with an activation entropy of zero. Nonetheless, the impact of entropy on enzyme-catalyzed concerted polycyclization cascades remains unknown. The importance of addressing entropy is corroborated by examining the enzyme-catalyzed generation of a tricyclic core (Figure S1a in the Supporting Information).

An attractive way to obtain an enzymatic polycyclization catalyst that display biologically relevant rates would be to make the entropy of catalysis more favorable in addition to stabilizing the carbocations. This would compensate for the entropic penalty associated with substrate prefolding that would otherwise significantly reduce the reaction rate. Such an entropic catalysis could be achieved by expelling ordered water molecules from the active site (Figure S1a, top) in a manner analogous to the binding of some ligands to receptors.^[17,18] In support of this idea, ¹H STD-NMR revealed that the binding of a small substrate-like probe to the active site of the triterpene cyclase from *Alicyclobacillus acidocaldarius* was driven by entropy (Figure S1b and S1c), as would be expected from the expulsion of water upon association. Moreover, seminal crystallographic work by Reinert et al.^[19] and Wendt et al.^[20] allowed us to perform an analysis of the crystal structures of this thermophilic triterpene cyclase complexed with ligands of two different sizes (PDB 1UMP^[19] and 2SQC^[20]). The results showed that the active site contained nine additional water molecules when complexed with the less bulky ligand (Figure S2 in the Supporting Information). We were faced with the problem of how such fixed water molecules in the active site could be expelled since the substrate and corresponding pentacyclic products enter and exit through the hydrophobic cell membrane.^[21] Furthermore, no conformational change occurs upon substrate binding (see the Supporting Information). The work with the extremophilic triterpene cyclase (PDB 1UMP^[19]) started with a structural analysis for water channels by using the software Caver.^[22] Interestingly, a number of different tunnels were found and we focused on the three highest-ranked channels that connect the binding site to the cytosol (Figure 1 a). Water molecules were found to reside within the identified channels and 20 ns MD-simulations in a waterbox were sufficient to observe movement of water molecules in the three tunnels in support of entropic catalysis (Figure S3).

To find experimental support for the entropically favored mechanism depicted in Figure 1 b, kinetic experiments were performed using three substrates (for mono-, tri- and pentacyclization; Scheme 1) with wild-type enzyme and several tunnel variants. The tunnel variants, in which amino acid residues in the walls of the tunnels were mutated (Figure 1a), were initially constructed in silico and MD simulations combined with Caver analysis confirmed that

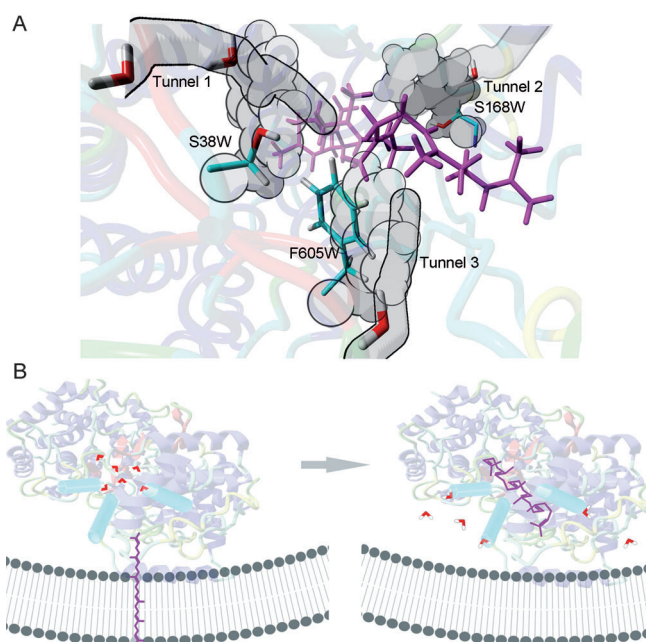


Figure 1. The existence of specific tunnels implies a central role for water in promoting the terpene cyclase catalyzed generation of multi-cyclic products. A) Water channel analysis using Caver and snapshots from MD simulations of the triterpene cyclase (PDB 1UMP).^[19] The three tunnels with highest ranking according to the software are schematically drawn with the water molecules confined within the channels shown. Key residues that were changed to block the channels are shown as stick models and the corresponding change in bulk upon introducing mutations are shown as gray space-filling representations. The natural C30 substrate squalene is shown in magenta. The catalytic Asp that initiates the cyclization cascade is located at the back of the figure and is not shown. B) Suggested reaction mechanism of entropically favored enzyme-catalyzed polycyclization. The release of ordered water molecules (red) makes the prefolding of the polyisoprenoid substrate (purple) favorable.

they blocked the water channels (Table S4). The experimental study was initiated by analyzing the temperature dependence of the second-order rate constant (apparent k_{cat}/K_M) for the enzyme-catalyzed generation of pentacycles. It was not possible to achieve saturation of the enzyme under our experimental conditions (0.2 % TritonX-100 micelles, 60 mM citrate, pH 6). This is perhaps not surprising since the triterpene cyclase is a monotopic membrane “receptor protein” and the phospholipid bilayer is different from the aqueous cellular environment that many enzymes operate in. The thermodynamic consequences of blocking the channels by the introduced mutations were studied experimentally (Figure 2) and by using transition-state theory (see Equation (1) in the Supporting Information).

It was found that the rate of pentacycle formation catalyzed by the wild-type enzyme increased two orders of magnitude when the temperature increased from 30 °C to 55 °C. The data corresponds to an activation entropy of 16 kcal mol⁻¹ at 328 K (or 50 cal mol⁻¹ K⁻¹) favoring the polycyclization reaction, whereas the high enthalpy of 31 kcal mol⁻¹ disfavors catalysis. To corroborate the very large entropic contribution to enzymatic polycyclization catalysis,

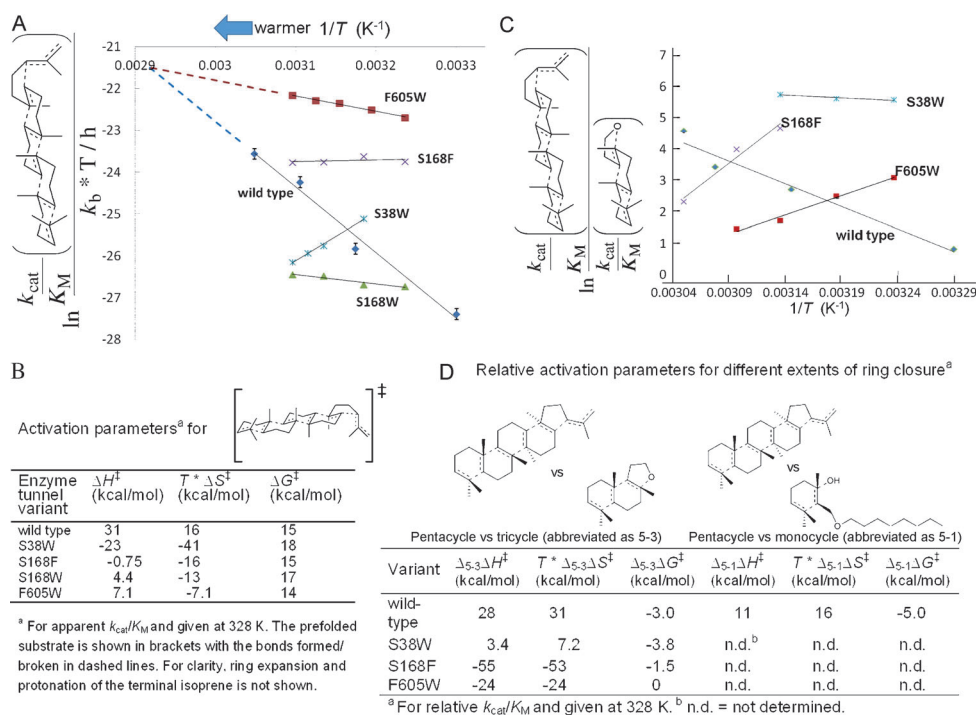


Figure 2. Thermodynamic aspects of enzyme-catalyzed polycyclization. A) Data for the generation of the natural pentacyclic scaffold (shown next to the graph with bonds formed/broken shown as dashed lines). The intersection of the extrapolated dashed blue line (wild type) and red line (F605W) on the graph corresponds to the temperature for which the two proteins have equal activities (69°C). Experiments were not conducted above 55°C because of micelle stability issues. Error bars for the wild-type enzyme from four independent experiments are shown for reference. B) Thermodynamic data from (a) were obtained by using Equation (1) in the Supporting Information. C) A competition experiment with all three substrates in one-pot to assess the relative thermodynamic consequences for different extents of cyclization. Data for pentacyclic over monocyclic formation is shown in the Supporting Information. The generation of monocycles was associated with low activity and data could only be obtained for the wild-type enzyme. Bonds formed/broken are shown as dashed lines in the structures shown to the left of the graph. D) Thermodynamic data obtained by using the data in (c) and Equation (2) in the Supporting Information.

we performed an additional thermodynamic analysis using previously published kinetic parameters for pentacyclization catalyzed by the same triterpene cyclase.^[10c] From that study, in which no thermodynamic analysis was performed (and for which k_{cat}/K_M values were only available at two temperatures), we derived an activation enthalpy of 31 kcal mol⁻¹ and an activation entropy of 19 kcal mol⁻¹ (at 328 K). The excellent agreement between the activation parameters that we obtain for the two independent systems is noteworthy.

The favorable activation entropy of 16 kcal mol⁻¹ is surprisingly high when compared to the estimated entropic cost (of opposite sign) of prefolding the C30 polyisoprene backbone (at least 22 kcal mol⁻¹ at 328 K according to the prevention of free rotation around 15 bonds).^[12] It should be noted that the entropic cost of “freezing” the rotation around a chemical bond is known experimentally.^[12] Hence the entropic contribution to enzyme-catalyzed pentacyclization corresponds to a rate enhancement of up to 10²⁸-fold, assuming that all fifteen bonds are completely frozen. The mutants harboring blocked channels were all active and displayed a different temperature dependence of catalysis. By contrast, control mutations in and around the active site had small effects (see the Supporting Information). The F605W

tunnel variant displayed up to 100-fold higher apparent k_{cat}/K_M values for pentacycle formation compared to that of the wild-type enzyme at temperatures lower than 69°C. This is a consequence of a reduced activation enthalpy (favoring catalysis) and entropy (disfavoring catalysis). By analogy, the S168F tunnel variant displayed an absolute apparent k_{cat}/K_M value that was approximately an order of magnitude higher than that of the wild-type enzyme at 40°C (Figure 2a). These facts demonstrate the important consequences of the enthalpy–entropy compensation^[23] and show how a different arrangement of water channels can affect the synthetic capabilities of triterpene cyclases at different temperatures. The S38W mutation (located 8 Å away from the polyisoprene substrate) is unusual in that it displays lower rates of polycyclization at higher temperatures (with linear kinetics) with a concomitant massive decrease in entropy of almost 60 kcal mol⁻¹ compared to that of the wild-type enzyme at

328 K. The entropic contribution to the rate enhancement was thus decreased 10⁴⁰-fold by one point mutation. This value corresponds to “freezing” 17 water molecules in a salt crystal.^[24] Such very large entropic effects, which are furthermore not associated with large structural rearrangements,^[25] were previously not acknowledged in enzyme catalysis.^[26,27] Moreover, the association of ligands with receptors and enzymes is not driven by entropy per se.^[18] The negative activation enthalpies are in accordance with a difference in exposure of the hydrophobic active site to water going from the ground state to the transition state and have previously been observed for the kinetics of protein folding.^[28]

Kinetic isotope effect analysis in D₂O (Figure S8) resulted in an isotope effect of around two, which demonstrates that protonation of the terminal isoprene unit of the substrate is rate limiting (except for in S168W, which gave no isotope effect). The measured entropic effect is thus concomitant with the actual chemical reaction since cyclization occurs in concert with the protonation that generates a transient reactive carbocationic species.^[11] A close to perfect linear (Figure 3a, R² 0.9986) enthalpy–entropy compensation^[23] for the four proteins that display a kinetic isotope effect indicates

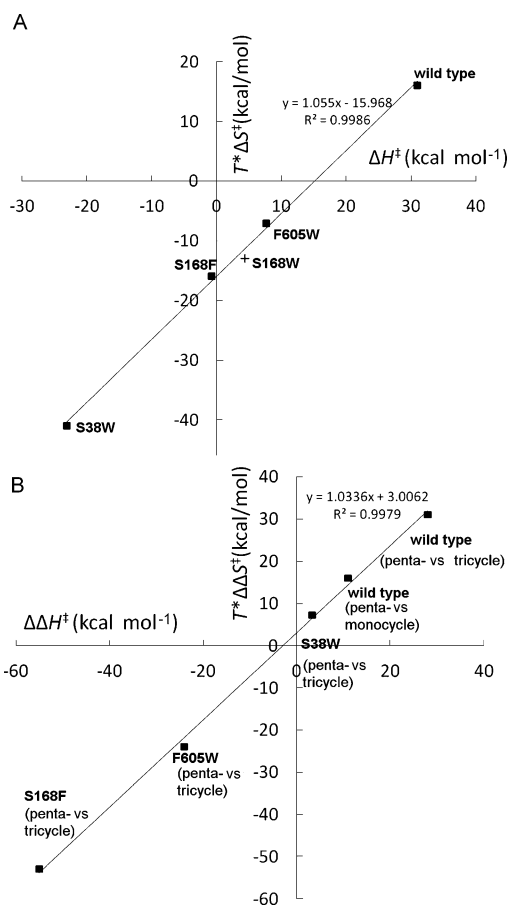


Figure 3. Enthalpy–entropy compensations with the entropic contributions given at 328 K. A) The activation parameters for pentacycle formation were taken from Figure 2b. B) The $T^*\Delta\Delta S^\ddagger$ values are plotted against the corresponding $\Delta\Delta H^\ddagger$ values (taken from Figure 2d) for the competition experiments performed using three different substrates with the wild-type enzyme and the variants.

that they could operate by the same reaction mechanism and share the same transition state.

Mono-, tri-, and pentacyclic ring closures (Scheme 1) were analyzed in a one-pot reaction in order to investigate the impact of entropy on different extents of cyclization (Figure 2c,d and Equation (2) in the Supporting Information). The rate for the one-ring closure was 2000-fold lower than for pentacycle formation and could only be measured for the wild type cyclase. Our results show that pentacyclization catalyzed by the wild-type enzyme is entropically favored over monocyclization (by 16 kcal mol^{−1} at 328 K), although the inherent entropic constraints of formation of a productive precyclic conformation for the latter are much less severe. By analogy, the formation of the pentacyclic scaffold is entropically favored over the tricyclic core by close to 30 kcal mol^{−1} for the wild-type enzyme, as illustrated by an increase in the relative amount of pentacyclic product with temperature (Figure 2c,d). In terms of the hydrophobic effect,^[29] this energy difference would correspond to a massive change in accessible surface area of around 1500 Å² for the two substrates that differ by 13 atoms. By contrast, the difference in the accessible surface area between the substrates that yield

the pentacyclic and the tricyclic product is calculated to be only around 200 Å² (by using the modeling software YASARA^[30]). This result indicates that the entropic contribution to catalysis is more complicated than what would perhaps be expected, a fact that highlights the prerequisite of prefolding of the substrate for cyclization. Moreover, the formation of pentacycles was entropically disfavored compared to the generation of tricycles for the S168F and F605W tunnel variants (Figure 2d). The largest effect was observed for the S168F variant, which showed a massive change in the relative activation entropy for pentacycle over tricycle formation of 84 kcal mol^{−1} when compared to that of the wild-type enzyme ($\Delta_{S168F-wild\ type} \Delta_{5-3} \Delta S^\ddagger$ at 328 K; Figure 2d). These effects together indicate that the large entropic contribution to catalysis is more complex than what would be expected. The enthalpy–entropy compensation between $\Delta\Delta H^\ddagger$ and $\Delta\Delta S^\ddagger$ was found to be close to perfect, which indicates that different extents of cyclization could operate by the same reaction mechanism (Figure 3b). Moreover, the analyzed tunnel variants displayed a shift in the multicyclic product spectrum in relation to that of the wild-type enzyme (Figure 2c and Table S10 in the Supporting Information).

In summary, this study on a thermophilic cyclase provides new mechanistic insight for a very important enzyme class. The expulsion of water molecules through specific channels assists in the prefolding of the linear polyisoprene chain and entropy is the key to polycyclizations catalyzed by terpene cyclases. This process allows remarkable chemistry through a carbocationic polycyclization cascade initiated by the rate-limiting transfer of a proton to the terminal isoprene group of the prefolded substrate. Our results demonstrate how an important class of enzymes has evolved to benefit from entropy rather than enthalpy to perform efficient catalysis in contrast to virtually all other described enzymes.^[25,27,31] The findings herein indicate that the use of entropy as a strategy in enzyme catalysis likely appeared early^[2] and could be of high importance and more general than previously thought.^[32] Moreover, the entropy-driven cyclization strategy described herein is expected to be of high importance in biology since 90 % of all natural products contain a cyclic subunit.^[33]

Received: February 4, 2014

Revised: February 21, 2014

Published online: April 7, 2014

Keywords: biosynthesis · enzyme catalysis · polycyclization · reaction mechanisms · thermodynamics

- [1] Y. Gao, R. B. Honzatko, R. J. Peters, *Nat. Prod. Rep.* **2012**, *29*, 1153–1175.
- [2] J. J. Brocks, G. A. Logan, R. Buick, R. E. Summons, *Science* **1999**, *285*, 1033–1036.
- [3] P. P. Peralta-Yahya, F. Zhang, S. B. del Cardayre, J. D. Keasling, *Nature* **2012**, *488*, 320–328.
- [4] P. K. Ajikumar, W.-H. Xiao, K. E. J. Tyo, Y. Wang, F. Simeon, E. Leonard, O. Mucha, T. H. Phon, B. Pfeifer, G. Stephanopoulos, *Science* **2010**, *330*, 70–74.
- [5] D.-K. Ro, et al., *Nature* **2006**, *440*, 940–943.
- [6] a) S. C. Hammer, P.-O. Syren, M. Seitz, B. M. Nestl, B. Hauer, *Curr. Opin. Chem. Biol.* **2013**, *17*, 293–300; b) M. Seitz, P.-O.

- Syren, L. Steiner, J. Klebensberger, B. M. Nestl, B. Hauer, *ChemBioChem* **2013**, *14*, 436–439.
- [7] a) D. C. Hibbit, R. P. Linstead, *J. Chem. Soc.* **1936**, 470–476; b) C. D. Upper, C. A. West, *J. Biol. Chem.* **1967**, *242*, 3285–3292; c) D. J. Miller, R. K. Allemann, *Nat. Prod. Rep.* **2012**, *29*, 60–71; d) B. T. Greenhagen, P. E. O'Maille, J. P. Noel, J. Chappell, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 9826–9831.
- [8] a) L. Ruzicka, *Experientia* **1953**, *9*, 357–367; b) G. Stork, A. W. Burgstahler, *J. Am. Chem. Soc.* **1955**, *77*, 5068–5077; c) R. Rajamani, J. Gao, *J. Am. Chem. Soc.* **2003**, *125*, 12768–12781; d) Y. Jin, D. C. Williams, R. Croteau, R. M. Coates, *J. Am. Chem. Soc.* **2005**, *127*, 7834–7842.
- [9] K. U. Wendt, *Angew. Chem.* **2005**, *117*, 4032–4037; *Angew. Chem. Int. Ed.* **2005**, *44*, 3966–3971.
- [10] a) E. E. Van Tamelen, J. D. Willett, R. B. Clayton, *J. Am. Chem. Soc.* **1967**, *89*, 3371–3373; b) T. Hoshino, T. Sato, *Chem. Commun.* **2002**, 291–301; c) N. Morikubo, Y. Fukuda, K. Ohtake, N. Shinya, D. Kiga, K. Sakamoto, M. Asanuma, H. Hirota, S. Yokoyama, T. Hoshino, *J. Am. Chem. Soc.* **2006**, *128*, 13184–13194; d) D. W. Christianson, *Curr. Opin. Chem. Biol.* **2008**, *12*, 141–150; e) E. Oldfield, F.-Y. Lin, *Angew. Chem.* **2012**, *124*, 1150–1163; *Angew. Chem. Int. Ed.* **2012**, *51*, 1124–1137; f) D. J. Tantillo, *Nat. Prod. Rep.* **2011**, *28*, 1035–1053.
- [11] a) L. Smentek, B. A. Hess, *J. Am. Chem. Soc.* **2010**, *132*, 17111–17117; b) D. H. Aue, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 487–508.
- [12] M. I. Page, W. P. Jencks, *Proc. Natl. Acad. Sci. USA* **1971**, *68*, 1678–1683.
- [13] G. Illuminati, L. Mandolini, *Acc. Chem. Res.* **1981**, *14*, 95–102.
- [14] I. Čorić, B. List, *Nature* **2012**, *483*, 315–319.
- [15] K. Ishihara, S. Nakamura, H. Yamamoto, *J. Am. Chem. Soc.* **1999**, *121*, 4906–4907.
- [16] a) K. Manna, W. C. Everett, G. Schoendorff, A. Ellern, T. L. Windus, A. D. Sadow, *J. Am. Chem. Soc.* **2013**, *135*, 7235–7250; b) W. D. Huntsman, T. H. Curry, *J. Am. Chem. Soc.* **1958**, *80*, 2252–2254.
- [17] a) T. Young, R. Abel, B. Kim, B. J. Berne, R. A. Friesner, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 808–813; b) P. Ball, *Chem. Rev.* **2008**, *108*, 74–108; c) P. C. Weber, J. J. Wendoloski, M. W. Pantoliano, F. R. Salemme, *J. Am. Chem. Soc.* **1992**, *114*, 3197–3200.
- [18] P. W. Snyder, J. Mecinovic, D. T. Moustakas, S. W. Thomas III, M. Harder, E. T. Mack, M. R. Lockett, A. Heroux, W. Sherman, G. M. Whitesides, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 17889–17894.
- [19] D. J. Reinert, G. Balliano, G. E. Schulz, *Chem. Biol.* **2004**, *11*, 121–126.
- [20] K. U. Wendt, A. Lenhart, G. E. Schulz, *J. Mol. Biol.* **1999**, *286*, 175–187.
- [21] K. U. Wendt, K. Poralla, G. E. Schulz, *Science* **1997**, *277*, 1811–1815.
- [22] E. Chovancova et al., *PLoS Comput. Biol.* **2012**, *8*, e1002708.
- [23] L. Liu, Q.-X. Guo, *Chem. Rev.* **2001**, *101*, 673–695.
- [24] J. D. Dunitz, *Science* **1994**, *264*, 670.
- [25] M. Wang, K. Warncke, *J. Am. Chem. Soc.* **2013**, *135*, 15077–15084.
- [26] a) M. Garcia-Viloca, J. Gao, M. Karplus, D. G. Truhlar, *Science* **2004**, *303*, 186–195; b) D. Ringe, G. A. Petsko, *Science* **2008**, *320*, 1428–1429.
- [27] A. Sievers, M. Beringer, M. V. Rodnina, R. Wolfenden, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 7897–7901.
- [28] M. Oliveberg, Y.-J. Tan, A. R. Fersht, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 8926–8929.
- [29] D. Chandler, *Nature* **2005**, *437*, 640–647.
- [30] E. Krieger, T. Darden, S. B. Nabuurs, A. Finkelstein, G. Vriend, *Proteins Struct. Funct. Bioinf.* **2004**, *57*, 678–683.
- [31] a) R. Wolfenden, *Annu. Rev. Biochem.* **2011**, *80*, 645–667; b) S. S. Licht, C. C. Lawrence, J. Stubbe, *Biochemistry* **1999**, *38*, 1234–1242.
- [32] R. B. Stockbridge, C. A. Lewis, Jr., Y. Yuan, R. Wolfenden, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 22102–22105.
- [33] K. Gilmore, I. V. Alabugin, *Chem. Rev.* **2011**, *111*, 6513–6556.