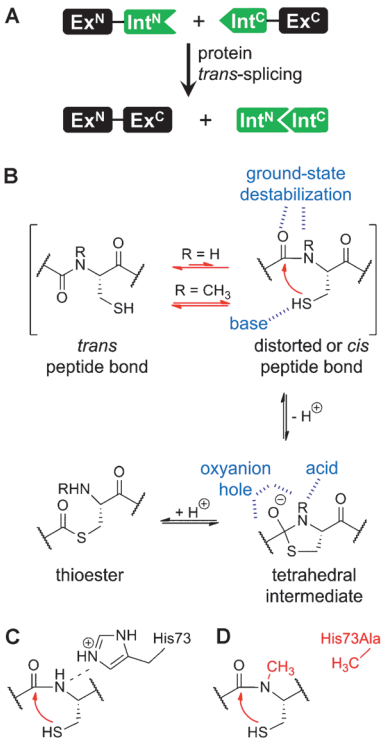
Protein Chemistry

DOI: 10.1002/anie.201208863

Chemical Bypass of Intein-Catalyzed N–S Acyl Shift in Protein Splicing**

*Jens Binschik and Henning D. Mootz**

Inteins are self-processing protein domains that excise themselves out of a precursor polypeptide chain in a multi-step pathway termed protein splicing. In the course of this reaction, the sequences flanking the intein, termed N- and C-terminal exteins, are linked with a peptide bond (Scheme 1 A). Inteins perform only a single turn-over, however, they employ catalytic strategies similar to those of enzymes. While inteins have found widespread use in many applications in biotechnology and protein chemistry, important aspects of the mechanism of protein splicing are still not understood.[1] In particular, the N–S (or N–O) acyl shift of the upstream scissile peptide bond into a thioester (or oxoester) remains intriguing (Scheme 1 B). This rearrangement represents the first step of protein splicing in standard inteins[2] and is widely exploited for the generation of protein thioesters.[3] Although thermodynamically unfavored, it does not require any cofactors or energy sources. Several catalytic mechanisms have been proposed for this reaction, including destabilization of the ground state of the scissile peptide bond, general acid–base catalysis to increase the nucleophilicity of the cysteine (or serine) side chain at position 1 of the intein, and stabilization of the tetrahedral intermediate by means of an oxyanion hole (Scheme 1 B). Some of these mechanisms, if not all, can likely be used in combination and may contribute to catalysis to different degrees in different inteins. Here, we pursued a novel chemical approach to directly probe the importance of ground-state destabilization by introducing an alkyl substituent at the amide nitrogen of the scissile peptide bond (Scheme 1 B); a strategy inspired by the recent development of N–S switch devices for the chemical synthesis of peptide thioesters.[4] This chemical manipulation indeed supported the N–S acyl shift, even to the extent that it could complement an otherwise essential part of the catalytic framework of the intein, the highly conserved block B histidine. Together, our findings reveal the role of the histidine in a ground-state destabilization mechanism and rule out other roles previously proposed.



Scheme 1. Ground-state destabilization for N–S acyl shift. The same mechanistic considerations apply for *cis*-splicing inteins. A) Scheme of protein trans-splicing. (Ex^N, Ex^C = N- and C-terminal flanking extein polypeptides; Int^N, Int^C = N- and C-terminal intein fragments) B) Pathway of thioester formation with proposed catalytic contributions of the intein. C) Proposed role of the block B histidine (His 73 in the *Ssp* DnaB intein) to destabilize the scissile peptide bond. D) Chemical activation of the scissile peptide bond by N-methylation.

The ground-state destabilization of the upstream scissile peptide bond has been proposed on the basis of structural analyses, but its direct investigation remains challenging. In two intein structures, deviations of this peptide bond from the common *trans* conformation to the energetically less favored *cis* conformation[5] or a distorted conformation[6] were observed (Scheme 1 B). A solution NMR study by Muir and co-workers on an *Mxe* GyrA intein precursor that is active in the N–S acyl shift suggested a significant distortion of the peptide bond.[7] Importantly, the highly conserved histidine residue in the block B signature motif (His 75 im *Mxe* GyrA Intein) was essential for this effect, as a His75Ala mutant was inactive in thioester formation and did not induce any measurable conformational strain on the peptide bond.[7] Block B is one of the conserved sequence motives in inteins.[1a] A key role of the block B histidine in the N–S

[*] Dr. J. Binschik, Prof. Dr. H. D. Mootz
Institute of Biochemistry, University of Muenster
Wilhelm-Klemm-Str. 2, 48149 Münster (Germany)
E-mail: Henning.Mootz@uni-muenster.de

[**] We thank Alexandra Wasmuth and Julian Matern for discussions. Financial support from the Deutsche Forschungsgemeinschaft (grant MO 1073/3-1,2) is acknowledged.

Supporting information for this article (including detailed experimental protocols and full gels of SDS-PAGE analysis in Figure 2) is available on the WWW under http://dx.doi.org/10.1002/anie.201208863.

acyl shift was already suggested from many crystal structures (for example, Refs [5,6,8]), because it is usually found in hydrogen-bonding distance to the amide nitrogen atom of the scissile bond (Scheme 1C). Mutation of the histidine to alanine blocks the N–S acyl shift in most inteins and significantly inhibits it in others,[9] suggesting that its contribution does not have to be essential when other catalytic mechanisms, for example, oxyanion stabilization of the tetrahedral intermediate,[9,10] are more dominant. Nevertheless, this residue is the most highly conserved amino acid in all intein sequences.[10,11] The widely postulated model for the role of the block B histidine is a contribution to the N–S acyl shift by destabilization of the scissile peptide bond. This is believed to be facilitated through polarization and ultimately protonation of the amine leaving group of the oxythiazolidine anion intermediate (Scheme 1B).[7,10] However, except for the above-mentioned study,[7] direct biochemical evidence for this model is scarce, and in most studies only loss-of-function mutagenesis was carried out. Furthermore, Belfort, Wang, and co-workers recently suggested a dual role for the block B histidine based on pK_a measurements in the *Mtu* RecA intein. From a pK_a shift seen between a precursor lacking the C-extein and the spliced intein they proposed that the histidine first deprotonates the catalytic cysteine side chain for nucleophilic attack in the precursor and then protonates the amine leaving group.[12] Stabilization of the thiolate anion may then be further supported by the reasonably well conserved block F aspartate playing a secondary role.[13] Another proposed role for the block B histidine is the stabilization of the oxyanion in the tetrahedral intermediate of the N–S acyl shift in the *Sce* VMA intein.[14] Thus, the precise role for the block B histidine remains unclear.

A further complication with the conclusions drawn from structural studies arises from the fact that in most reported intein structures, unstrained, regular *trans* peptide bonds were found at the position of the N-terminal scissile bond,[8c,14,15] thus questioning the requirement for peptide bond distortion. However, a general problem in interpreting all structural data so far is that in these experiments mutated or truncated intein versions were used to block splicing or cleavage activities of the self-processing proteins. Also in the above-mentioned solution NMR study,[7] the intein fusion protein was manipulated by deletion of the C-extein and mutation of the ultimate asparagine, which is part of the catalytic center, thus making the protein incapable of splicing. Such changes could also affect the integrity of the catalytic center and the orientation of the amino acids under investigation. Together, these considerations underline the difficulty of rigorously studying a single step or an individual mechanistic contribution in the splicing pathway in isolation. Clearly, a minimally invasive perturbation would be desirable, with only slight structural changes that keep the entire protein splicing pathway intact.

Here, we aimed to study the contribution of ground-state destabilization to the N–S acyl shift by introducing a subtle chemical perturbation in the active site of an intein. We hypothesized that an *N*-methyl group at the N-terminal splice junction (Scheme 1B) would weaken the scissile peptide bond by increasing its ground-state destabilization. We reasoned that if ground-state destabilization is an important part of the mechanism of intein, then a positive effect on the N–S acyl shift would be expected and should be measurable.

N-alkylation is known to shift the thermodynamic equilibrium between *trans* and *cis* peptide bonds in favor of the energetically higher *cis* isomer, because the energy of the *trans* form is increased (Scheme 1B). This effect is well known for proline, the only proteinogenic imino acid. In aminoacyl-prolyl peptide bonds, the *trans*:*cis* ratio is typically only about 4:1 compared to 1000:1 for a regular non-proline peptide bond. Because of steric effects, the energy of the *trans* isomer is increased by N-alkylation, leading to more similar interconversion rates between the two isomers.[16] *N*-alkyl cysteine has been exploited as a switch device for the chemical synthesis of peptide thioesters used in native chemical ligation.[4,17] The N–S acyl shift of the weakened amide bond is initiated by a low pH value, probably because protonation of the more basic, secondary amine is favored under these conditions, thus inhibiting the back reaction.

To incorporate this unnatural backbone methylation into the active-site framework of an intein, we resorted to a semisynthetic split *Ssp* DnaB intein in which the scissile bond is part of a synthetic peptide (see Figure 1 for all constructs used in this study). This intein variant was

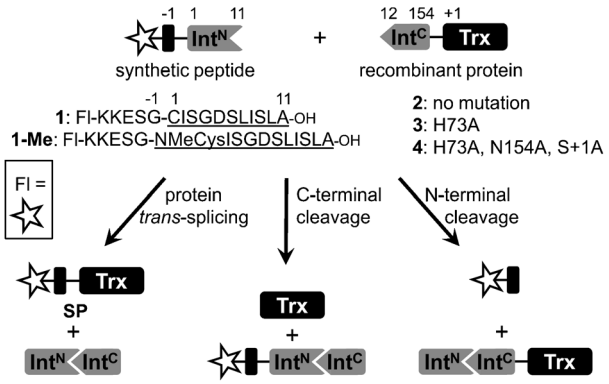


Figure 1. Constructs used in this study and reaction schemes. The artificially split semisynthetic *Ssp* DnaB intein consists of a short Int^N fragment of 11 aa, which harbors the catalytic Cys1, and a longer Int^C(12-154) fragment of 143 aa. The underlined peptide sequence corresponds to the Int^N part. Trx served as a model fusion protein. Possible reaction outcomes are illustrated. Calculated masses are: proteins **2**, **3** = 30.3 kDa; protein **4** = 30.2 kDa; Int^C = 16.7 kDa; Int^C (H73A) = 16.6 kDa; Trx = 13.7 kDa. (Fl: 5(6)-carboxyfluorescein; SP = splice product; Trx = thioredoxin with hexahistidine tag); S + 1A = mutant with alanine at position + 1 replacing serine in the native protein.

previously created by artificially splitting the mini-intein between amino acids 11 and 12 to give rise to a short N-terminal fragment, Int^N, and a longer C-terminal fragment, Int^C, comprising 11 and 143 amino acids (aa), respectively.[18] Int^N peptides with a short flanking sequence are readily accessible by solid-phase peptide synthesis. They spontaneously associate with Int^C-protein fusions, fold into the active conformation, and give the semisynthetic splice product (SP) as well as the liberated Int^N and Int^C fragments (Figure 1).

The unmethylated control peptide Fl-KKES*G*-*C*ISGDSLI-SLA-OH (**1**; scissile peptide bond italicized; Fl = 5(6)-carboxyfluorescein) was synthesized using standard Fmoc protecting group strategy. Incubation with recombinantly produced Int^C-Trx-His_6 (construct **2**) led to splicing in 76 % yield (Figure 2), similar to previous reports.[18b] Mutation of the highly conserved block B histidine (His73Ala; construct **3**) virtually completely abrogated splice product formation, as expected, whereas the side reaction of C-terminal cleavage (Figure 1), a consequence of premature asparagine cyclization at the C-terminal splice junction, occurred at moderately increased levels (Figure 2). Importantly, the latter finding indicated that the mutated Int^C fragment retained a correct overall structure in the reconstituted intein complex. The lost splicing activity thus only resulted from the missing histidine side chain. In agreement with this observation, we had previously reported that a His73Ala mutant was inactivated in the N–S acyl shift in the context of an Int^C-construct with two additional mutations (His73Ala, Asn154Ala, Ser + 1Ala; construct **4**) to block C cleavage reactions.[18b] Together, these findings support the notion that the block B His73 plays an essential role in activating the N-terminal scissile bond, consistent with most other inteins.

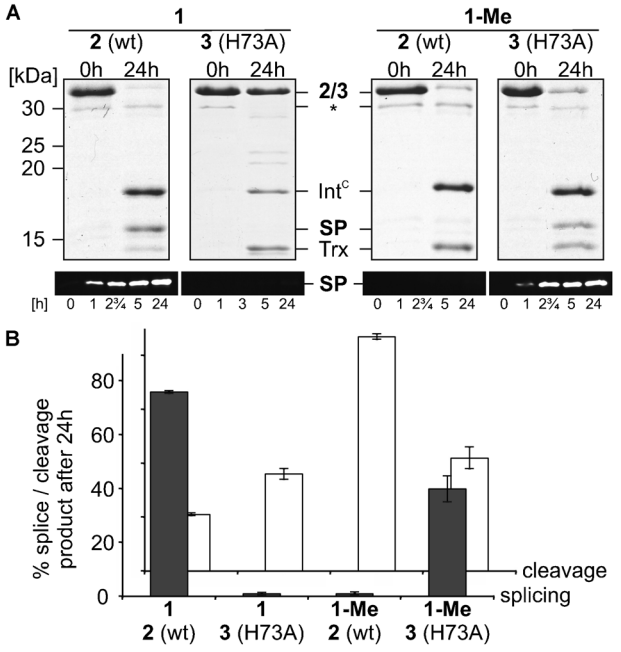


Figure 2. Protein trans-splicing reactions. Reactions were carried out with 60 μM Int^N peptides **1** or **1-Me** and 20 μM Int^C constructs **2** or **3** at pH 7.0 and incubated for 24 h at 25 °C A) Shown are SDS-PAGE gels stained with coomassie brilliant blue (top) and illuminated under UV light (sections at bottom). SP indicates splice product Fl-KKESG-Trx-His_6 and Trx indicates C-terminal cleavage product Trx-His_6. B) Quantification of SP (front) and C-terminal cleavage product (back) formation by densitometric analysis. (* = protein contamination).

Next, we synthesized peptide **1-Me** with *N*-methyl cysteine substituting the catalytic Cys1 (Figure 1). The unnatural building block was incorporated using Fmoc-*N*-MeCys(Trt)-OH during peptide assembly. However, upon incubation of peptide **1-Me** with protein **2** under the same conditions as for peptide **1**, only marginal amounts of splice product (< 1 %) could be detected. Efficient formation of the C-terminal cleavage product (> 85 % yield) was observed and indicated an overall correctly folded intein domain (Figure 2). We therefore speculated that the loss of splicing activity is a consequence of an undesired steric hindrance caused by the extra methyl group. The prime candidate for such an unfavorable interaction was the catalytically essential His73 itself, because it is found in hydrogen-bond-forming distance of 2.7 Å to the amide nitrogen of the Cys1 residue.[8c] We therefore incubated peptide **1-Me** with the His73Ala mutant construct **3** (Scheme 1 D). Indeed, in this combination, robust protein trans-splicing with 40 % yield was observed. Also the rates decreased only modestly compared to the initial system (lower panels in Figure 2 A). About 42 % C-terminal cleavage was still observed, giving rise to a total amount of active intein of 80–85 %; similar to the other functional combination **1** + **2** (Figure 2 B). Thus, Int^N peptides **1** and **1-Me** behaved reciprocally in their splicing activities with native Int^C fusion protein **2** and the corresponding block B histidine mutant **3** (Figure 2 B). These findings show that N-methylation of the scissile peptide bond is indeed beneficial for the N–S acyl shift, even to the extent that it complements the function of the block B histidine.

We carried out control reactions to answer the question at which stage the N-methylation supported the activation of the scissile peptide bond for the N–S acyl shift. Given the known effect of the *N*-alkyl cysteine building block as switch device, it was conceivable that the peptide entered the active site already in the thioester form, in which case the intein would not need to contribute to the N–S acyl shift anymore. To address this possibility, we performed control experiments in which peptide **1-Me** was incubated at pH 7.0 with various thiols to cleave any thioester by thiolysis, either in the absence or in the presence of the Int^C triple mutant **4**. As mentioned above, this mutant is blocked in all catalytic steps following intein fragment association and thus allows investigation of uncatalyzed thioester formation and cleavage in isolation. In fact, peptide **1-Me** alone remained mostly stable (> 90 %) in the presence of 50 mM DTT, whereas in the presence of construct **4**, 2 mM DTT was already sufficient to effect complete cleavage at the scissile bond (Figure 3 A). The small amounts of cleavage observed in the absence of the protein partner might be attributed to a low level of a spontaneous N–S acyl shift with the equilibrium on the side of the peptide bond. In contrast, the unmethylated peptide **1** was completely stable under all these conditions (Figure 3 B). Taken together, it can be concluded that the scissile peptide bond of the free peptide **1-Me** was in the amide form. The context of the folded intein is still required for the efficient N–S acyl shift to take place.

The known chemical *N*-alkyl switch devices are triggered at low pH instead of neutral pH; typically with 20 % (*v*/*v*) 3-mercaptopropionic acid (MPA) at a final pH of approximately 2.[4b] To confirm that the sequence of **1-Me** is not prohibitive for this kind of induction, we also tested these conditions. Indeed, complete cleavage of peptide **1-Me** was

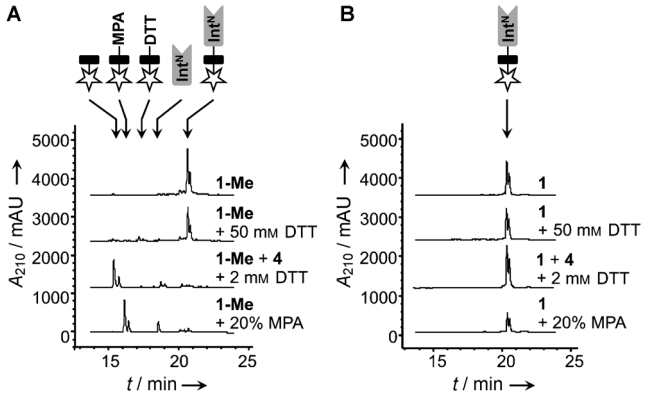


Figure 3. Thioester formation and cleavage. Shown are reversed-phase HPLC traces of reaction mixtures. Indicated combinations of Int^N peptides (15 μM) and Int^C mutant constructs (30 μM) were incubated for 24 h at pH 7.0 and 25 °C in presence of the indicated thiols and subsequently boiled for the removal of protein **4**. The marked peaks were identified by subsequent MALDI-TOF analysis (data not shown). While the Ex^N-signals are clearly visible, only traces of Int^N were detected, likely because of co-precipitation with **4**. Note that all peptides bearing Fl (= 5(6)-carboxyfluorescein) appear as double peak because of the two regioisomers.

observed in the absence of **4**, while the unmethylated peptide **1** was completely stable (Figure 3).

Our results have several implications for the protein splicing mechanism of the *Ssp* DnaB intein. 1) N-methylation of the scissile peptide bond clearly promoted the N–S acyl shift, even in the absence of the block B histidine, which is otherwise critical for this reaction. This finding supports the model of ground-state destabilization being an important part of catalysis, because the N-methylation energetically favors a distorted scissile peptide bond or a *cis* conformer relative to an unmodified peptide bond. It corroborates the idea that the block B histidine plays an important role to bring about the ground-state destabilization, probably by coordinating the amide nitrogen atom (Scheme 1C). 2) Our observations also indicate that this is the only important role of the histidine in the N–S acyl shift. An additional role of this residue as a general base to deprotonate the thiol side chain of the catalytic Cys1 was proposed for the *Mtu* RecA intein.[12] However, this can be ruled out as an important part of the mechanism, at least for the *Ssp* DnaB intein that was studied here, if we assume that deprotonation of *N*-MeCys thiol side chain would require the same base. This notion is consistent with our previous finding that a substitution of Cys1 to homocysteine still resulted in an efficient N–S acyl shift, despite the perturbed positioning of the thiol and the imidazole groups because of the extra methylene group in the side chain.[19] Our model is also in agreement with the unusually low pK_a measured for the histidine in a spliced intein,[12] which would be expected from the proximity to the protonated free amino terminus of the intein. 3) Our data show that the block B histidine is also not critical for the downstream steps of the protein splicing pathway or the coordination of the individual reactions, because the His73Ala mutant catalyzed robust protein splicing with peptide **1-Me** and thus reconstituted the nearly fully active intein. The proposed role of the block B histidine appears to be general, given its high conservation in the primary sequence and the high structural conservation of the active site of inteins. However, it cannot be excluded that other inteins, such as the *Mtu* RecA intein, employ slightly different mechanisms.

Interestingly, in a very recent study, Perler and co-workers investigated the *Tko* CDC21-1 intein as a representative of the only small group of inteins lacking the block B histidine. Mutational analysis and structural modeling led them to propose that in this case the loss of reactivity that is engendered by the missing histidine is mitigated by stabilization of the oxyanion in the tetrahedral intermediate (Scheme 1B) through a specific lysine residue.[10] Our results similarly suggest an alternative pathway to compensate for the loss of the histidine, however, by an artificial, chemical bypass to facilitate the breaking of the scissile peptide bond.

More generally, also the chemical synthesis protocols for the generation of peptide thioesters through N–S acyl shift reactions have to work without a protein-mediated catalytic support.[20] These preparative routes include the aforementioned *N*-alkyl cysteine switch devices,[4] as well as cysteinyl–prolyl ester units to trap the formed thioester,[21] and even particularly reactive peptide bonds, such as Gly–Cys, Cys–Cys, and His–Cys.[22] They all work best at low pH, indicating that the high proton concentration helps to slow down the reverse S–N acyl shift. Strikingly, under these conditions, the cysteine side chain is protonated, consistent with our proposal that a general base for thiol deprotonation may likewise not be required for the catalytic pathway of inteins. The activation of the scissile peptide bond by the block B histidine is thus likely achieved through conformational strain by electronic effects, from polarizing the amide to protonating the amino leaving group (Scheme 1B–C), but may as well be brought about by distortion of the amide through steric effects.

In summary, we have shown that the catalytically essential block B histidine can be complemented by N-methylation of the cysteine residue in the scissile peptide bond. Importantly, and in contrast to other studies, the perturbed intein was still capable of protein splicing. We have thus collected strong evidence for the importance of the histidine for promoting the N–S acyl shift through ground-state destabilization and could rule out other potential roles of this residue in the protein splicing pathway. Our work also underlines the power of subtle chemical modifications in the active site of a protein for mechanistic studies.

Received: November 5, 2012
Revised: December 13, 2012
Published online: March 6, 2013

Keywords: enzymology · peptide chemistry · peptide ligation · protein semisynthesis · thioesters

[1] a) L. Saleh, F. B. Perler, *Chem. Rec.* **2006**, *6*, 183; b) S. Elleuche, S. Poggeler, *Appl. Microbiol. Biotechnol.* **2010**, *87*, 479; c) H. D. Mootz, *ChemBioChem* **2009**, *10*, 2579; d) N. H. Shah, T. W. Muir, *Isr. J. Chem.* **2011**, *51*, 854; e) H. Paulus, *Chem. Soc. Rev.* **1998**, *27*, 375; f) Y. Anraku, Y. Satow, *Proc. Jpn. Acad. Ser. B*

2009, *85*, 409; g) T. L. Aboye, J. A. Camarero, *J. Biol. Chem.* **2012**, *287*, 27026; h) G. Volkmann, H. D. Mootz, *Cell. Mol. Life Sci.* **2012**, DOI: 10.1007/S00018-012-1120-4.
[2] K. Tori, B. Dassa, M. A. Johnson, M. W. Southworth, L. E. Brace, Y. Ishino, S. Pietrokovski, F. B. Perler, *J. Biol. Chem.* **2010**, *285*, 2515.
[3] T. W. Muir, *Annu. Rev. Biochem.* **2003**, *72*, 249.
[4] a) H. Hojo, Y. Onuma, Y. Akimoto, Y. Nakahara, Y. Nakahara, *Tetrahedron Lett.* **2007**, *48*, 25; b) L. A. Erlich, K. S. Kumar, M. Haj-Yahya, P. E. Dawson, A. Brik, *Org. Biomol. Chem.* **2010**, *8*, 2392.
[5] T. Klabunde, S. Sharma, A. Telenti, W. R. Jacobs, Jr., J. C. Sacchettini, *Nat. Struct. Biol.* **1998**, *5*, 31.
[6] B. W. Poland, M. Q. Xu, F. A. Quiocho, *J. Biol. Chem.* **2000**, *275*, 16408.
[7] A. Romanelli, A. Shekhtman, D. Cowburn, T. W. Muir, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 6397.
[8] a) X. Duan, F. S. Gimble, F. A. Quiocho, *Cell* **1997**, *89*, 555; b) K. Ichiyanagi, Y. Ishino, M. Ariyoshi, K. Komori, K. Morikawa, *J. Mol. Biol.* **2000**, *300*, 889; c) Y. Ding, M. Q. Xu, I. Ghosh, X. Chen, S. Ferrandon, G. Lesage, Z. Rao, *J. Biol. Chem.* **2003**, *278*, 39133.
[9] Z. Du, J. Liu, C. D. Albracht, A. Hsu, W. Chen, M. D. Marieni, K. M. Colelli, J. E. Williams, J. N. Reitter, K. V. Mills, C. Wang, *J. Biol. Chem.* **2011**, *286*, 38638.
[10] K. Tori, M. Cheriyan, C. S. Pedamallu, M. A. Contreras, F. B. Perler, *Biochemistry* **2012**, *51*, 2496.
[11] F. B. Perler, *Nucleic Acids Res.* **2002**, *30*, 383.
[12] Z. Du, P. T. Shemella, Y. Liu, S. A. McCallum, B. Pereira, S. K. Nayak, G. Belfort, M. Belfort, C. Wang, *J. Am. Chem. Soc.* **2009**, *131*, 11581.
[13] Z. Du, Y. Zheng, M. Patterson, Y. Liu, C. Wang, *J. Am. Chem. Soc.* **2011**, *133*, 10275.
[14] R. Mizutani, S. Nogami, M. Kawasaki, Y. Ohya, Y. Anraku, Y. Satow, *J. Mol. Biol.* **2002**, *316*, 919.
[15] a) E. Werner, W. Wende, A. Pingoud, U. Heinemann, *Nucleic Acids Res.* **2002**, *30*, 3962; b) P. Sun, S. Ye, S. Ferrandon, T. C. Evans, M. Q. Xu, Z. Rao, *J. Mol. Biol.* **2005**, *353*, 1093; c) M. A. Johnson, M. W. Southworth, T. Herrmann, L. Brace, F. B. Perler, K. Wuthrich, *Protein Sci.* **2007**, *16*, 1316; d) J. S. Oeemig, A. S. Aranko, J. Djupsjobacka, K. Heinamaki, H. Iwai, *FEBS Lett.* **2009**, *583*, 1451; e) B. P. Callahan, N. I. Topilina, M. J. Stanger, P. Van Roey, M. Belfort, *Nat. Struct. Mol. Biol.* **2011**, *18*, 630; f) J. S. Oeemig, D. Zhou, T. Kajander, A. Wlodawer, H. Iwai, *J. Mol. Biol.* **2012**, *421*, 85.
[16] G. Fischer, *Chem. Soc. Rev.* **2000**, *29*, 119.
[17] P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. Kent, *Science* **1994**, *266*, 776.
[18] a) C. Ludwig, M. Pfeiff, U. Linne, H. D. Mootz, *Angew. Chem.* **2006**, *118*, 5343; *Angew. Chem. Int. Ed.* **2006**, *45*, 5218; b) C. Ludwig, D. Schwarzer, H. D. Mootz, *J. Biol. Chem.* **2008**, *283*, 25264.
[19] D. Schwarzer, C. Ludwig, I. V. Thiel, H. D. Mootz, *Biochemistry* **2012**, *51*, 233.
[20] a) J. Kang, D. Macmillan, *Org. Biomol. Chem.* **2010**, *8*, 1993; b) F. Mende, O. Seitz, *Angew. Chem.* **2011**, *123*, 1266; *Angew. Chem. Int. Ed.* **2011**, *50*, 1232.
[21] T. Kawakami, S. Aimoto, *Tetrahedron Lett.* **2007**, *48*, 1903.
[22] J. Kang, J. P. Richardson, D. Macmillan, *Chem. Commun.* **2009**, 407.

