A conserved threonine spring-loads precursor for intein splicing

Albert K. Dearden,¹ Brian Callahan,² Patrick Van Roey,³
Zhong Li,³ Utsav Kumar,¹ Marlene Belfort,^{2*} and Saroj K. Nayak^{1,4*}

¹Department of Physics, Applied Physics, and Astronomy, Rensselaer Polytechnic Institute, Troy, New York 12180

²Department of Biological Sciences and RNA Institute, University at Albany, Albany, New York 12222

³Wadsworth Center, New York State Department of Health, Albany, New York 12201

⁴School of Basic Sciences, Indian Institute of Technology, Bhubaneswar, India

Received 14 December 2012; Revised 12 February 2013; Accepted 13 February 2013

DOI: 10.1002/pro.2236

Published online 19 February 2013 proteinscience.org

Abstract: Protein splicing is an autocatalytic process where an “intein” self-cleaves from a precursor and ligates the flanking N- and C-“extein” polypeptides. Inteins occur in all domains of life and have myriad uses in biotechnology. Although the reaction steps of protein splicing are known, mechanistic details remain incomplete, particularly the initial peptide rearrangement at the N-terminal extein/intein junction. Recently, we proposed that this transformation, an N-S acyl shift, is accelerated by a localized conformational strain, between the intein’s catalytic cysteine (Cys1) and the neighboring glycine (Gly-1) in the N-extein. That proposal was based on the crystal structure of a catalytically competent trapped precursor. Here, we define the structural origins and mechanistic relevance of the conformational strain using a combination of quantum mechanical simulations, mutational analysis, and X-ray crystallography. Our results implicate a conserved, but largely unstudied, threonine residue of the Ssp DnaE intein (Thr69) as the mediator of conformational strain through hydrogen bonding. Further, the strain imposed by this residue is shown to position the splice junction in a manner that enhances the rate of the N-S acyl shift substantially. Taken together, our results not only provide fundamental understanding of the control of the first step of protein splicing but also have important implications in various biotechnological applications that require precursor manipulation.

Keywords: intein; autocatalytic proteins; first principles quantum mechanics; crystal structure; reaction rate

Additional Supporting Information may be found in the online version of this article.

Brian Callahan’s current address is State University of New York at Binghamton Department of Chemistry, 4400 Vestal Parkway East Binghamton, NY 13902.

Grant sponsors: The Interconnect Focus Center funded by the MARCO program of SRC and the State of New York; Grant sponsor: NSF CBET; Grant number: 0931717; Grant sponsor: NIH; Grant number: GM44844.

*Correspondence to: Saroj K. Nayak, Department of Physics, Applied Physics, & Astronomy, Rensselaer Polytechnic Institute, 110 Eighth Street, Troy, NY 12180-3590. E-mail: nayaks@rpi.edu or Marlene Belfort, Biological Sciences, University at Albany, State University of New York, 1400 Washington Avenue, Albany, NY 12222. E-mail: mbelfort@albany.edu

Introduction

Inteins are protein self-splicing elements that cleave themselves from precursors soon after translation. The polypeptide products, called N- and C-exteins, are concomitantly ligated by the intein, thereby generating a functional protein.^{1,2} In the first step of splicing, the backbone peptide bond at the N-extein/intein junction is attacked by the intein’s first residue, a cysteine (or serine) [Fig. 1(A)], leading to a tetrahedral intermediate followed by the formation of a thioester (or ester).^{3,4} Substantial kinetic and thermodynamic barriers are associated with such S/N-O acyl shifts⁵ and their molecular triggers remain undefined.

Figure 1. Protein splicing. (A) Canonical pathway. The intein (red) self-cleaves from a precursor polypeptide and ligates the two cleaved fragments, called exons (green), in three sequential, autocatalytic reactions: (1) N-S acyl shift (2) transthiostereification; and (3) asparagine cyclization. (B) Step 1, the N-S acyl shift. Splicing initiates with the catalytic cysteine residue of the intein (Cys1) attacking the peptide bond to the adjacent N-extein residue (Gly-1), generating a thioester.

Recently, we reported the structure of a fused Ssp DnaE intein precursor that had been trapped in its elusive ground state. The trap comprised an engineered disulfide bond between the catalytic Cys1 residue of the intein and a cysteine at the -3 position of the N-extein.⁶ We noted that Gly-1, which corresponds to the N-extein junction residue in the precursor, adopts an unusual geometry, which is consistent with similar studies involving splicing reactions in which distortions are present.^{7–11} The (ϕ , Ψ) angles of Gly-1 fall outside the typical range for this residue. We speculated that the structural distortion at Gly-1 arises from hydrogen bonding with a conserved threonine, Thr69 [Fig. 2(A)]. Although largely unstudied, Thr69 is highly conserved and forms part of the signature TXXH motif, where the histidine is a catalytic residue in N-terminal cleavage.¹³ Here, we define the structural origins and mechanistic relevance of the conformational strain using a combination of quantum mechanical (QM) simulations, mutational analysis, and X-ray crystallography.

Results and Discussion

Computational validation of the role of Thr69 in the distortion of the Gly-1 conformation

Our initial calculation for our model system in the presence of Thr69 was performed using a full QM

Figure 2. Ground state glycine angles and reaction scheme. (A) Interaction of Thr69 with the intein-extein junction residues as observed in the crystal structure of the trapped DnaE intein precursor. Thr69 is shown interacting with the amide nitrogen of Cys1, the first intein residue, and the carbonyl oxygen of Pro-2, bridging Gly-1, the last N-extein residue. (B) Section of the Ramachandran plot for glycine. Darker regions are allowed, white is nonallowed.¹² Experimental and computational results are shown for different residues at position 69: threonine (dark green, calculated; light green observed); Alanine (red, calculated; pink, observed); Ser(1) (yellow) was derived directly from the crystal structure and Ser(2) (dark blue) after computational relaxation of the starting structure involving threonine; Glycine (light blue). (C) Scheme of the N-S acyl shift between cysteine and glycine. The reactant is on the left and the tetrahedral intermediate is on the right. The dashed lines indicate the hydrogen bonds between the Thr69 hydroxyl and the Cys1 peptide nitrogen and the Pro-2 carbonyl that are important for distorting the Gly-1 conformation.

Table I. Computational Values for the Reaction Barriers

	φ	Ψ	Region	Barrier (kcal/mol)
Thr	-123.2	77.8	NA	14.8
	-124.6	53.0	A	34.6
Ala	-119.6	50.7	A	24.9
	-123.2	77.8	NA	48.3

Computational values for the reaction barriers involving the native Thr69 and the T69A mutant. Each mutant is grouped with its corresponding allowed (A) or nonallowed (NA) region in the Ramachandran plot and its energy barrier.

method. Despite the large computational cost, this was a necessary step to appropriately construct our mixed QM and molecular mechanical (MM) system. Following the full QM calculation, we moved to our QM/MM calculation, again in which Thr69 is present, and were able to obtain Gly-1 (φ, Ψ) dihedral angles of $(-123.2^\circ, 77.8^\circ)$ which compare well with the experimentally derived angles of $(-122.5^\circ, 86.9^\circ)$ (Table I). Increasing the system size to include more atoms did not change the results presented here, which suggested convergence of system size for the properties of interest. This QM/MM system with Thr69 provided similar structural parameters such as bond angles and bond distances to those obtained using the full QM scheme, particularly around the catalytic region, suggesting that QM/MM could provide accurate results. To explicitly probe the influence of Thr69-mediated hydrogen bonding on the Gly-1 conformation, we substituted alanine for Thr69 (T69A) and again optimized the precursor structure. In this ground state structure with the T69A substitution, the dihedral angles of Gly-1 returned to allowed values, with (φ, Ψ) angles of $(-119.6^\circ, 50.7^\circ)$ (see Table I). These results support our hypothesis that hydrogen bonding interactions between Thr69 of the TXXH motif and the N-extein-intein junction are responsible for the distortion at Gly-1 [interatomic distances are shown in Fig. 2(A)].¹⁴ Specifically, we propose that hydrogen bonding between Thr69 and both the amide nitrogen of Cys1 and the carbonyl oxygen of Pro-2 that flanks the glycine is responsible for the unusual backbone angles at Gly-1 [Fig. 2(A) and (B)].

Next, we expanded our *in silico* mutagenic study of Thr69 to include substitutions with serine, valine, cysteine, glycine, and aspartic acid using our QM/MM regime. Could these residues, which occur in place of threonine in the TXXH motif in a limited number of inteins,^{12,15} produce the same distortion at the N-extein/intein junction? Of the five residues tested, only the serine mutation (T69S) was able to generate a distorted conformation at Gly-1 [Fig. 2(B)]. Interestingly, serine was able to produce both distorted and nondistorted Gly-1 dihedral angles. The calculation in which serine provided nondis-

torted dihedral angles, Ser(1) in Figure 2(B), began in the same manner as the native Thr69 calculation, that is, from the provided crystal structure. Using the result of the calculation in which serine provided nondistorted angles, we reintroduced threonine and performed another geometry optimization. This calculation resulted in nondistorted angles as well. However, the system resulted in a higher energy than the calculation resulting in distorted angles, indicating that the case in which threonine is present with distorted dihedral angles, is more stable. For the calculation in which we obtained the distorted dihedral angles, Ser(2), we implemented the T69S mutation after our computational geometry relaxation on the native Thr69 system. Comparing their ground state energies showed a difference below our computational accuracy (~ 0.1 eV) and thus our conclusion is that serine may produce both distorted and nondistorted dihedral angles. These results further implicate hydrogen bonding interactions of the side chain hydroxyl group as the mediator of structural distortion in the precursor. All mutations and their resulting dihedral angles in Gly-1 are presented in Table S1 in Supporting Information.

Computational analysis of role of the Gly-1 distortion on intein splicing

The distorted geometry at Gly-1 is reflected in this residue in the nonallowed region of the Ramachandran plot [Ref. ⁶ and Fig. 2(B)]. We furthered our investigation by considering the extent to which this distorted geometry impacts the first step of the splicing reaction, the N-S acyl shift. As seen in Figure 2(C), we started our reaction as the sulfur attacks the carbonyl carbon. The computed energy barrier for Thr69, starting with the Gly-1 nonallowed dihedral angles, was found to be 14.8 kcal/mol (Table I). In contrast, this energy barrier was 34.6 kcal/mol when starting with Gly-1 allowed dihedral angles. These angles were adapted from the T69S calculations as serine is the only mutant that imposed both allowed and nonallowed angles as seen in Figure 2(B). Although the computed energy barrier for the allowed case represents the upper bound of the value, the results nevertheless provided an estimate of the reduction of the energy barrier due to the strained configuration.

Armed with this understanding, we hypothesized that due to the absence of a hydrogen bond to the Cys1 amide, the T69A mutant should have a higher energy barrier. Indeed, our computed energy barrier for the T69A mutant for the N-S acyl shift reaction is increased by 10.1 kcal/mol compared to the wild type (from 14.8 to 24.9 kcal/mol) (Table I) suggesting that the rate of splicing for T69A should be slower by a factor of $\sim 10^7$. The differences between the Thr69 and T69A energy barriers may

Figure 3. Effect of conserved Thr69 on the activity of the fused DnaE intein precursor. Cleavage of the thioester intermediate, depicted left, by added DTT (0.02 M), liberates an 18 kDa intein-C-extein fragment from a 22 kDa precursor. Reaction progress was monitored by SDS-PAGE followed by staining with Coomassie Blue (right). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

be explained by the computed electronic charges for the corresponding atoms shown in Table S2 in Supporting Information.¹⁴ Specifically, comparing the charges of the attacking sulfur atom on the catalytic cysteine for the allowed cases of threonine and alanine, we see a difference in the values. This suggests that threonine reduces the charge on the attacking sulfur through its hydrogen bonding interaction.

Intein splicing activity of Thr69 mutants

To test our predictions, we assayed the activity of point mutants of the DnaE intein in which T69 was replaced with alanine or serine. Inteins were expressed in the oxidizing *Escherichia coli* origami strain¹⁶ and purified by Ni-NTA chromatography. Activity was assessed by chemical cleavage of the intein-generated thioester by strong nucleophiles. For our experiments, we incubated purified precursor in Tris buffer with added dithiothreitol (DTT) as the cleavage agent. At selected intervals, aliquots from the cleavage reactions were removed and analyzed by SDS-PAGE. As shown in Figure 3, DTT at 20 mM completely cleaved the N-extein from the wild type DnaE intein after 18 h (lane 2). A $t_{1/2}$ of ~40 min was determined from reaction progress curves, in reasonable accord with earlier studies.⁶ The two mutants displayed slower kinetics. Less than 5% of the N-extein had been cleaved from the T69A variant after 18 h (lane 4) and after 48 h (not shown). A faster reaction was observed with the T69S mutant, but the rate was still considerably reduced relative to the wild type (compare lane 2 to lane 6). These results are in accord with our simulations, indicating that the threonine plays a crucial role in accelerating the first step of protein splicing.

Structural confirmation of the role of Thr69 in distorting Gly-1

As a final test for our simulations, we determined the crystal structure of the T69A mutant of the DnaE intein precursor. The tertiary structures of the T69A

mutant and the wild-type intein are superimposable [Fig. 4(A)]. However, it was gratifying to find deviations at the splice junction that were consistent with the QM/MM calculations. In the T69A structure, the Gly-1 conformation ($\phi = -88.0^\circ$, $\Psi = 29.8^\circ$) [Fig. 4(B)] falls in the allowed region of the Ramachandran plot as seen in Figure 2(B). The conformation of the disulfide trap, residues -3 through 1, was similar to those observed for the same Cys-Pro-Gly-Cys loop in other structures in which the glycine conformation is not strained (Fig. S1 in Supporting Information). Additionally, the smaller side chain at position 69 allows for the insertion of a water molecule adjacent to Ala69 (Fig. S2 in Supporting Information). This water molecule is in hydrogen bonding contact with the carbonyls of Pro-2 and Cys1, the nitrogen of Ala69, and the N δ 1 side chain nitrogen of conserved His72. The latter interaction requires flipping of the His side chain and alters the interaction of the catalytic His72 with residues at the C terminal of the protein. The observation that the T69A mutation perturbs the conformation of the coconserved His72 residue may be indicative of their cooperative action during catalysis, a possibility that we are currently investigating. Regardless, relief of the conformational strain of Gly-1 by the T69A mutation underscores the role of Thr69 in activating the scissile bond for splicing. Not only does this conclusion enhance our fundamental insight into the control of the first step of splicing but also it has implications for biotechnological applications of inteins. Many such applications require control of splicing and precursor accumulation to be optimally useful. These include peptide ligation as well as protein modification, cyclization, and affinity purification.^{17–19}

Conclusions

In summary, using first principles quantum mechanics (QM) and molecular mechanics (MM), mutational analysis, and X-ray crystallography, we provided the

Figure 4. Structural comparison of wild-type intein and T69A mutant. (A) DnaE intein T69A mutant retains the native tertiary structure. Structural alignment of T69A (green, 4GIG) and WT DnaE inteins (grey, 3NZM). (B) Conformation of the T69A mutant (green) with the native intein (grey) at the active site. The bifurcated hydrogen bonds from the threonine hydroxyl group to the nitrogen atom of Cys1 and to the carbonyl oxygen atom of Pro-2 that contribute to the distortion of the conformation of Gly-1 are absent in the T69A mutant.

first atomic level understanding of the role of the threonine of the TXXH motif in protein splicing. We found that ground state interactions involving Thr69 lead to a backbone distortion that “prepared” Gly-1 for efficient N-S acyl shift in the first step of protein splicing. In particular, hydrogen bonding to Thr69 helped to stabilize the transition state thereby linking structural and mechanistic features. Our study not only provides a fundamental understanding of splicing but also could have implications in various biotechnological applications that require precursor isolation for subsequent manipulations.

Materials and Methods

Computational methods

To analyze the role of Thr69 in stabilizing the structural distortion at Gly-1, we used first principles quantum mechanics based on the Becke hybrid functional B3LYP²⁰ with the 6-31G(d,p) basis set as well as a mixed QM and molecular mechanics method using the AMBER force field^{21,22} as implemented in the Gaussian 03 package.²³ For our simulations, we analyzed a minimized DnaE intein based on the published precursor structure.⁶ The model system consisted of an engineered disulfide loop that trapped the intermediate structure and all molecular fragments within an area extending 10 Å outward radially at the full QM level. To help simulate the natural constraints on the system that would otherwise not be present due to the removal of the surrounding intein, the outer most structures were held in place during all geometry optimizations.

For investigations involving the N-S acyl reaction, we could not use the crystal structure as the

structure represents a trapped intermediate of an intein that does not splice. To overcome this difficulty, we chose a small model system localized around the active catalytic region. This model is similar to the one used in the past to study intein splicing and consists of the Pro–Gly–Cys tripeptide and threonine or alanine depending on the mutation. The model system was studied using a full QM method with the Synchronous Transit and Quasi-Newton^{24,25} technique which was then refined with the rational function optimization using the Berny algorithm. Additionally, single-point energy calculations confirmed the presence of only a single negative frequency for each transition structure, confirming a transition state. To simulate the solvent, the polarized continuum model²⁶ was implemented, which creates a spherical cavity in a dielectric with a specified constant in which the atoms are placed. The three different constants used were $\epsilon = 1$ (vacuum), $\epsilon = 78.3553$ (water), and $\epsilon = 10$, which is in the range of calculated dielectric constants within protein structures.^{27,28} Because of the reduced systems in which there is no surrounding protein to hold the reaction site together, the respective amino acids, threonine or alanine, were held at fixed distances from the reaction site whose values were obtained from the previous QM/MM calculations. We have considered the proton transfer²⁹ for the reaction in a concerted manner similar to the mechanism proposed by Anraku and Satow.³⁰ Although other possible schemes for proton transfer during N-S acyl shift have been proposed³¹ and considering our structural information that shows proximity of the carbonyl group and the sulfhydryl of cysteine, we believe that our choice is reasonable. In addition,

our focus is to compute any energy difference between various mutants and this choice should not affect our conclusion.

Mutagenesis, protein production, and purification

The T69A and T69S point mutants of the fused DnaE intein (Ming Xu, NEB Biolabs) were generated by overlap extension PCR using the primers in Table S3 in Supporting Information with plasmid, His₆-^{CPGC}DnaE-spl¹ as the template. Briefly, the T69A variant was created in two steps: first we amplified overlapping fragments with primers 2845, 3086, 2846, and 3085; next, the two amplified fragments were fused together by PCR using primers 3085 and 3086. The T69S construct was prepared in a similar manner. The final PCR products were digested with PmlI and XhoI, and cloned into pET45b for expression with an N-terminal His₆ tag. We used the oxidizing *E. coli* origami DE3 strain (Novagen) as the host for protein expression. For crystallography experiments, the His₆ tag was cleaved off from the purified precursor by treatment with thrombin and then separated by gel filtration. Details of the primers used are listed in Supporting Information Table S3.

Crystallography

The protein sample was concentrated to approximately 8 mg/ml in a buffer containing 20 mM Tris.HCl, pH 8.0, and 200 mM NaCl. Crystals were obtained by hanging drop vapor diffusion methods at room temperature, using a precipitation buffer containing 52% ammonium sulfate, 2% isopropanol, 100 mM HEPES, and pH 7.0. The crystals belong to space group I4 with cell dimensions $a = b = 93.63$ Å, and $c = 41.24$ Å. Prior to data collection, all crystals were transferred to a cryosolvent consisting of the precipitation buffer complemented with 20% (v/v) glycerol. Crystals were flashcooled in a cold nitrogen gas stream (100 K). All X-ray diffraction data were collected using an in-house R-axis IV++ (Rigaku) detector and processed using CrystalClear (Rigaku). Data collection statistics are summarized in Table S4 in Supporting Information section.

The structure was determined by molecular replacement with the structure of the native DnaE trapped precursor intein (PDB entry 3NZM) as the search model, and then was refined using Crystallography & NMR System³² combined with model building using COOT.³³ The final refinement statistics are summarized in Supporting Information Table S4. All but one residue (Gln145) are within the preferred region of the Ramachandran plot as defined in the program MOLPROBITY.³⁴ Atomic coordinates and structure factors have been deposited with the Protein Data Bank, entry 4GIG.

Acknowledgment

The authors thank Matt Stanger for technical assistance and for help preparing figures. This work is supported partly by an anonymous gift from Rensselaer Computational Center for Nanotechnology Innovations at Rensselaer which is partly funded by State of New York was used for our work. Use of the facilities of the Wadsworth Center's Crystallography Core is gratefully acknowledged. SKN would like to thank Professor Mike Payne and Cavendish laboratory for their hospitality. The authors declare no conflict of interest in this work.

References

1. Kane PM, Yamashiro CT, Wolczyk DF, Neff N, Goebel M, Stevens TH (1990) Protein splicing converts the yeast TFP1 gene product to the 69-kD subunit of the vacuolar H⁺-adenosine triphosphatase. *Science* 250: 651–657.
2. Hirata R, Ohsumi Y, Nakano A, Kawasaki H, Suzuki K, Anraku Y (1990) Molecular-structure of a gene, VMA1, encoding the catalytic subunit of H⁺-translocating adenosine-triphosphatase from vacuolar membranes of *Saccharomyces cerevisiae*. *J Biol Chem* 265: 6726–6733.
3. Perler FB (1998) Protein splicing of inteins and hedgehog autoproteolysis: structure, function, and evolution. *Cell* 92: 1–4.
4. Paulus H (2000) Protein splicing and related forms of protein autoprocessing. *Ann Rev Biochem* 69: 447–496.
5. Lewis CA, Jr, Wolfenden R (2011) Amide bonds to the nitrogen atoms of cysteine and serine as “weak points” in the backbones of proteins. *Biochemistry* 50: 7259–7264.
6. Callahan BP, Topilina NI, Stanger MJ, Van Roey P, Belfort M (2011) Structure of catalytically competent intein caught in a redox trap with functional and evolutionary implications. *Nat Struct Mol Biol* 18: 630–633.
7. Xu Q, Buckley D, Guan C, Guo HC (1999) Structural insights into the mechanism of intramolecular proteolysis. *Cell* 98: 651–661.
8. Poland B, Xu MQ, Quirocho FA (2000) Structural insights into the protein splicing mechanism of Pl-Scel. *J Biol Chem* 275: 16408–16413.
9. Romanelli A, Shekhtman A, Cowburn D, Muir TW (2004) Semisynthesis of a segmental isotopically labeled protein splicing precursor: NMR evidence for an unusual peptide bond at the N-extein-intein junction. *Proc Natl Acad Sci USA* 101: 6397–6402.
10. Johansson DG A, Wallin G, Sandberg A, Macao B, Åqvist J, Hård T (2009) Protein autoproteolysis: conformational strain linked to the rate of peptide cleavage by the pH dependence of the N→O acyl shift reaction. *J Am Chem Soc* 131: 9475–9477.
11. Klabunde T, Sharma S, Telenti A, Jacobs WR, Jr, Sacchettini JC (1998) Crystal structure of GyrA intein from *Mycobacterium xenopi* reveals structural basis of protein splicing. *Nat Struct Biol* 5: 31–36.
12. Chen VB, Arendall BIII, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* 66: 12–21.

13. Du Z, Shemella PT, Liu Y, McCallum SA, Pereira B, Nayak SK, Belfort G, Belfort M, Wang C (2009) Highly conserved histidine plays a dual catalytic role in protein splicing: a pK(a) shift mechanism. *J Am Chem Soc* 131: 11581–11589.
14. Arunan E, Desiraju GR, Klein RA, Sadlej J, Scheiner S, Alkorta I, Clary DC, Crabtree RH, Dannenberg JJ, Hobza P, Kjaergaard HG, Legon AC, Mennucci B, Nesbitt DJ (2011) Definition of the hydrogen bond (IUPAC recommendations 2011). *Pure Appl Chem* 83: 1637–1641.
15. Perler FB (2002) InBase: the intein database. *Nucleic Acids Res* 30: 383–384.
16. Bessette PH, Aslund F, Beckwith J, Georgiou G (1999) Efficient folding of proteins with multiple disulfide bonds in the *Escherichia coli* cytoplasm. *Proc Natl Acad Sci USA* 96: 13703–13708.
17. Cheriyan M, Perler FB (2009) Protein splicing: a versatile tool for drug discovery. *Adv Drug Del Rev* 61: 899–907.
18. Vila-Perello M, Muir TW (2010) Biological applications of protein splicing. *Cell* 143: 191–200.
19. Callahan BP, Stanger M, Belfort M (in press) A redox trap to augment the intein toolbox. *Biotechnol Bioeng*. doi: 10.1002/bit.24821.
20. Becke AD (1993) Density-functional thermochemistry. III. The role of exact exchange. *J Chem Phys* 98: 5648–5642.
21. Maseras F, Morokuma K (1995) IMOMM—a new integrated ab-initio plus molecular mechanics geometry optimization scheme of equilibrium structures and transition-states. *J Comput Chem* 16: 1170–1179.
22. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA (1995) A 2nd generation force-field for the simulation of proteins, nucleic-acids, and organic-molecules. *J Am Chem Soc* 117: 5179–5197.
23. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery Jr. JA, Vreven T, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson GA, Nakatsuji H, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox J. E, Hratchian HP, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain MC, Farkas O, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Gonzalez C, and Pople JA. (2004) Gaussian 03, Rev. C.02. Wallingford, CT; Gaussian.
24. Peng C, Schlegel HB (1993) Combining synchronous transit and quasi-Newton methods to find transition-states. *Isr J Chem* 33: 449–454.
25. Peng C, Ayala PY, Schlegel HB, Frisch MJ (1996) Using redundant internal coordinates to optimize equilibrium geometries and transition states. *J Comput Chem* 17: 49–56.
26. Miertus S, Scrocco E, Tomasi J (1981) Electrostatic interaction of a solute with a continuum—a direct utilization of ab initio molecular potentials for the prevision of solvent effects. *J Chem Phys* 55: 117–129.
27. Schutz CN, Warshel A (2001) What are the dielectric “constants” of proteins and how to validate electrostatic models? *Proteins* 44: 400–417.
28. Jordanides XJ, Lang MJ, Song X, Fleming GR (1999) Solvation dynamics in protein environments studied by photon echo spectroscopy. *J Phys Chem B* 103: 7995–8005.
29. Blomberg MRA, Siegbahn PEM (2006) Different types of biological proton transfer reactions studied by quantum chemical methods. *Biochim Biophys Acta* 1757: 969–980.
30. Anraku Y, Satow Y (2009) Reflections on protein splicing: structures, functions, and mechanisms. *Proc Jpn Acad Ser B* 85: 409–421.
31. Paulus H (1998) The chemical basis of protein splicing. *Chem Rev* 27: 375–386.
32. Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* 54: 905–921.
33. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60: 2126–2132.
34. Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, Wang X, Murray LW, Arendall WB, III, Snoeyink J, Richardson JS, Richardson DC (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* 35: W375–W383.