VIP Protein Toxins Very Important Paper

Deutsche Ausgabe: DOI: 10.1002/ange.201603420
Internationale Ausgabe: DOI: 10.1002/anie.201603420

Elucidation of the Covalent and Tertiary Structures of Biologically Active Ts3 Toxin

*Bobo Dang, Tomoya Kubota, Kalyaneswar Mandal, Ana M. Correa, Francisco Bezanilla, and Stephen B. H. Kent**

Abstract: *Ts3 is an alpha scorpion toxin from the venom of the Brazilian scorpion Tityus serrulatus. Ts3 binds to the domain IV voltage sensor of voltage-gated sodium channels (Na_v) and slows down their fast inactivation. The covalent structure of the Ts3 toxin is uncertain, and the structure of the folded protein molecule is unknown. Herein, we report the total chemical synthesis of four candidate Ts3 toxin protein molecules and the results of structure–activity studies that enabled us to establish the covalent structure of biologically active Ts3 toxin. We also report the synthesis of the mirror image form of the Ts3 protein molecule, and the use of racemic protein crystallography to determine the folded (tertiary) structure of biologically active Ts3 toxin by X-ray diffraction.*

Ts3 is an alpha scorpion toxin protein natural product isolated from the venom of the Brazilian scorpion *Tityus serrulatus*.[1] Alpha scorpion toxins belong to the family of gating-modifier proteins that act on voltage-gated sodium channels (Na_v), and they mainly bind to the loop connecting S3 and S4 in Na_v domain IV.[2,3] Upon binding to the Na_v channel proteins, these toxins inhibit the fast inactivation of Na_v channels without dramatically affecting activation of the channels.[4–8]

Prior to the work in this paper, there was no X-ray or NMR structure of the Ts3 protein. Furthermore, there is confusion in the literature regarding the C terminus of the Ts3 polypeptide chain: some papers state that the active Ts3 toxin has 62 amino acid residues,[9] while others state it has 64 amino acid residues.[10,11] From analysis of the gene sequence, Martin-Eauclaire et al. proposed that the 64-residue Ts3 toxin polypeptide chain should terminate as the carboxamide.[10] In the same paper, they also suggested there could be three different C-terminal candidates for the correct sequence of the Ts3 toxin protein.[10] To resolve this confusion and to determine the amino acid sequence and covalent structure of active Ts3 toxin, our first task was to prepare the three proposed Ts3 polypeptide chains, together with the precursor polypeptide chain Ts3(1–67) (Scheme 1a), in order to determine which if any of these polypeptides folds to form biologically active Ts3 toxin.

We have previously reported total synthesis of the Ts1 toxin protein molecule,[12] but that experience was of limited use for the current work. Although isolated from the same scorpion venom, Ts3 and Ts1 toxins are very different protein molecules: they have distinct biological activities—Ts1 is a beta scorpion toxin while Ts3 is an alpha scorpion toxin; there is little sequence homology between the two polypeptide chains (only 16/53 conserved non-Cys residues; Figure 1) and the Ts3 polypeptide chain has much greater structural ambiguity at its C terminus than was the case with the Ts1 toxin.[12] Although the Ts3 and Ts1 toxins do share a conserved pattern of eight Cys residues in the amino acid sequence of their polypeptide chains (Figure 1), it has been shown that, even where the pattern of Cys residues is conserved, scorpion venom proteins can have distinct tertiary structures.[13]

```
Ts1:  KEGYLMDHEGCKLSCFIRPSGYCGRECGIKKGSSGYCAWP
Ts3:  KKDGYPVEYDNCAYICWNYDNAYCDKLCKDKKADSGYCYWV
      1        10        20        30        40
Ts1:  --ACYCYGLPNWVKVWDRATNKC
Ts3:  HILCYCYGLPDSEPTKTNG--KC(KSGKK)
               50        60
```

Figure 1. Comparison of the amino acid sequences of Ts3 and Ts1 toxins. The alignment is based on Cys residues (bold); conserved amino acids are underlined. Until now, the presence in mature biologically active Ts3 protein of the C-terminal residues enclosed in parentheses has been in dispute.

Total chemical synthesis of the candidate Ts3 polypeptide chains was carried out by using a modular approach as shown in Scheme 1b. Native chemical ligation (NCL) sites were chosen at -Tyr^{23}-Cys^{24}- and -Tyr^{46}-Cys^{47}- so that, after preparation of the C-terminal peptide variants, the peptide segments used were of similar length (23; 23; and 16, 18, 20, or 21 amino acids). Peptide segments were synthesized as described in the Supporting Information (Figures S1—S9). Four distinct C-terminal peptide segments were prepared through native chemical ligation of the Thz^{47}-Lys^{61}-thioester to the corresponding C-terminal Cys-peptide or cysteine (Scheme 1b). Analytical data for the preparation by native chemical ligation of the full-length Ts3(1–64)-$CONH_2$ polypeptide chain are shown in Figure 2a. Synthetic manipula-

[*] Dr. B. Dang, Dr. K. Mandal, Prof. Dr. S. B. H. Kent
Department of Chemistry, University of Chicago
Chicago, IL 60637 (USA)
E-mail: skent@uchicago.edu
Dr. B. Dang, Dr. T. Kubota, Dr. K. Mandal, Prof. Dr. A. M. Correa, Prof. Dr. F. Bezanilla, Prof. Dr. S. B. H. Kent
Department of Biochemistry & Molecular Biology
University of Chicago, Chicago, IL 60637 (USA)
Dr. B. Dang, Dr. K. Mandal, Prof. Dr. F. Bezanilla, Prof. Dr. S. B. H. Kent
Institute for Biophysical Dynamics, University of Chicago
Chicago, IL 60637 (USA)

Supporting information (including experimental details) for this article can be found under: http://dx.doi.org/10.1002/anie.201603420.

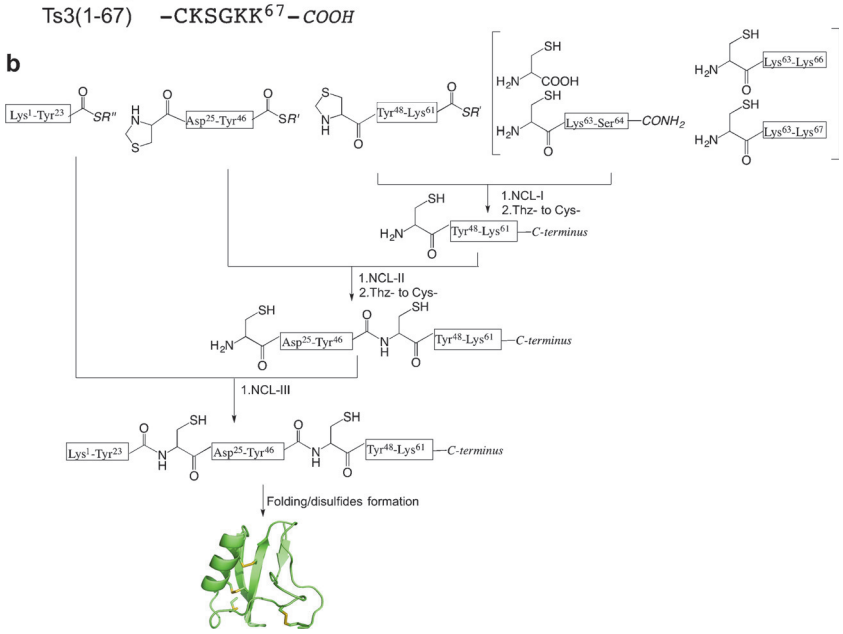
a Residue 1 – Residue 61:

KKDGYPVEYD NCAYICWNYD NA**YC**DKLCKD^{30}
KKADSGYCYW VHILC**YC**YGL PDSEPTKTNG^{60} K–

Four C-terminus candidates:

Ts3(1-62) –C^{62}–*COOH*
Ts3(1-64) –CKS^{64}–*CONH$_2$*
Ts3(1-66) –CKSGK^{66}–*COOH*
Ts3(1-67) –CKSGKK^{67}–*COOH*

b



Scheme 1. Synthesis of candidate Ts3 proteins. a) Sequences of candidate Ts3 polypeptide chains.[9,10] b) A modular synthetic strategy based on native chemical ligation for preparation of the four candidate Ts3 polypeptide chains. "C-terminus" is one of the four candidates listed in Scheme 1 a.

tions (see the Supporting Information) for making each of the four full-length candidate Ts3 polypeptide chains were essentially similar (Figures S10–12).

Folding the synthetic Ts3 polypeptide chains turned out to be extremely challenging. Conventional conditions with a thiol/disulfide redox couple, together with a low concentration of the chaotrope Gu·HCl to keep misfolded/mispaired disulfide intermediates in solution, did not produce any discrete folded product. Nor did the folding conditions we had previously developed for Ts1.[12] Addition of 20% *v*/*v* DMSO to the redox buffer gave acceptable yields of discrete folded products of mass 8 Da less than that of the corresponding polypeptide chains, which indicates the formation of four disulfide bonds. After careful optimization, the final folding conditions used were: [polypeptide] 0.02 mg mL^{-1}; 2.0 mM GSH/1.0 mM GSSG; 0.5 M L-arginine·HCl; 1.0 mM EDTA; 0.1 M Tris, 20% DMSO, pH 8.5; room temperature. Data for the folding of Ts3(1–64)-$CONH_2$ are shown in Figure 2 b. Using these conditions, we were able to fold all four Ts3 candidate polypeptide chains, with concomitant formation of 4 disulfides in each case as shown by the loss of 8 Da. Analytical LCMS data for the four variant synthetic candidate Ts3 proteins are shown in Figure S13.

Structure–activity profiles of the four candidate Ts3 protein molecules were evaluated by the fast inactivation decay rate of sodium currents recorded from heterologously expressed rat skeletal muscle voltage-gated sodium channels (rNa_v1.4) in the presence of 210 nM of each protein (5 min incubation at a holding potential of −90 mV; Figure 3). The Ts3(1–64)-$CONH_2$ protein showed obvious disruption of rNa_v1.4 fast inactivation (Figure 3 a), while the other three protein molecules showed minimal effects. On treatment with Ts3(1–64)-$CONH_2$ protein, fast inactivation decay of rNav1.4 became slower in a dose-dependent manner (Figure 3 b). The effect saturated at concentrations between 210–2100 nM, which is consistent with previous reports using Ts3 toxin isolated from scorpion venom.[8] The Ts3(1–64)-$CONH_2$ protein can thus be presumed to be the principal toxic component of natural Ts3 toxin isolated from scorpion venom.

With the covalent structure of biologically active Ts3 protein in hand, we set out to determine its folded structure by X-ray crystallography. Crystallization of L-Ts3(1–64)-$CONH_2$ toxin protein was problematic, so racemic protein crystallization was used.[14–16] Preparation of the D-Ts3(1–64)-$CONH_2$ polypeptide chain and its folding were as described for the L-protein enantiomer (Figures S14–16). Crystallization trials using the Hampton Index (HR2-144) screen and a racemic mixture at 10 mg mL^{-1} (5 mg mL^{-1} of L-Ts3(1–64)-$CONH_2$ toxin plus 5 mg mL^{-1} of D-Ts3(1–64)-$CONH_2$ protein) produced diffraction-quality crystals from a single condition [0.1 M BIS-TRIS pH 6.5, 28% *w*/*v* Polyethylene glycol monomethyl ether 2000 at 19 °C] after three weeks.

Diffraction data were collected from the racemic D/L-protein crystal at an energy level of 12.6 keV using the NE-CAT beamline 24-ID-E at the Advanced Photon Source at Argonne National Laboratory. Reasonably complete diffraction data were collected to a resolution of 1.93 Å. Diffraction intensity statistics revealed that the racemic D/L-protein crystallized in space group $P\bar{1}$, with one enantiomer in the asymmetric unit. The X-ray structure of racemic D/L-Ts3(1–64)-$CONH_2$ protein was solved by molecular replacement[17] using the scorpion toxin LQH-alpha-IT (PDB ID: 2ASC) as a search model. The final model was refined to a crystallographic *R* factor of 0.26 (*R* free 0.30) using Phenix.[18] Note that this centrosymmetric *R* factor corresponds to a conventional *R* factor that is approximately 33% smaller (i.e. *R* factor ≈ 0.20).[19] The packing of the D-Ts3(1–64)-$CONH_2$ and L-Ts3(1–64)-$CONH_2$ proteins in the unit cell is shown in Figure 4 a, with the two enantiomeric molecules in the unit cell.

As shown by comparison with the structures of the prototypical alpha scorpion toxins LQH-α-IT and Aah2 (PDB ID: 1AHO), the Ts3(1–64)-$CONH_2$ protein has a typ-

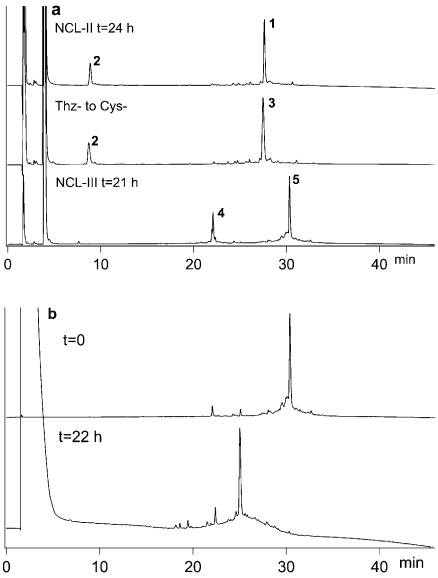


Figure 2. Chemical synthesis of the Ts3 (1–64)-$CONH_2$ protein. a) Analytical LCMS data for the synthesis of polypeptide Lys^{1}-Ser^{64}-$CONH_2$. **1** Thz^{24}-Ser^{64}-$CONH_2$; **2** residual Cys^{47}-Ser^{64}-$CONH_2$ from NCL-II; **3** Cys^{24}-Ser^{64}-$CONH_2$; **4** Lys^{1}-Tyr^{23}-COS thiolactone; **5** desired product Lys^{1}-Ser^{64}-$CONH_2$. b) Analytical HPLC data for folding/disulfides formation of crude Ts3(1–64)-$CONH_2$ polypeptide. Folding conditions are given in the text. Analytical LCMS conditions are described in the Supporting Information. GSH = reduced glutathione, GSSG = oxidized glutathione, EDTA = ethylendiaminetetraacetate, Tris = tris(hydroxymethyl)aminomethane, DMSO = dimethyl sulfoxide.

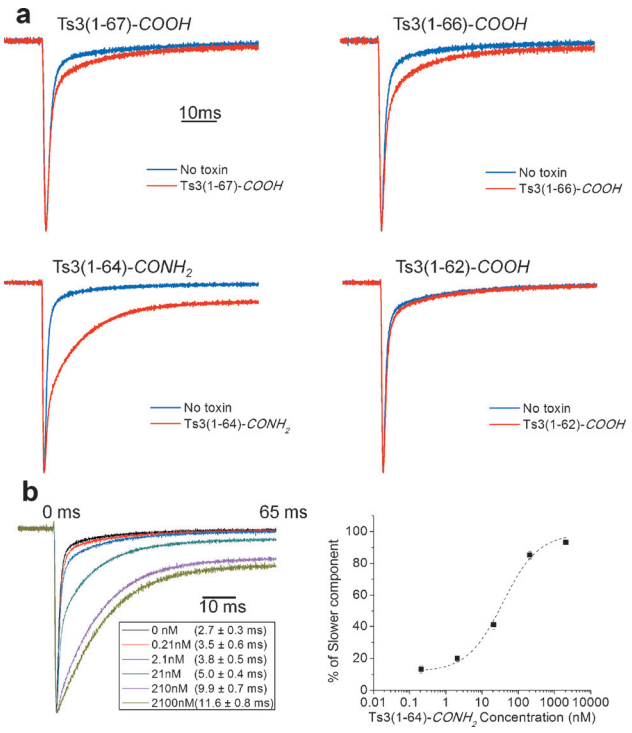


Figure 3. Biological activities of the Ts3 protein candidates. a) Disruption of $rNa_v1.4$ fast inactivation by the four candidate Ts3 proteins. Na^+ ionic current, elicited by a step-pulse to −10 mV from a holding potential of −90 mV, was measured without (blue traces) and with (red traces) 210 nM of each candidate protein. b) Dose–response assay of $rNa_v1.4$ fast inactivation modification by the Ts3(1–64)-$CONH_2$ protein. The right panel shows the proportion of the slower component of the fit as a function of protein concentration. Error bars indicate standard error of the mean (number of cells = 3–6).

ical alpha scorpion toxin fold made up of an α-helix (residues 21–31) packed against a three-stranded antiparallel β-sheet (residues 2–4, 35–39, 44–50; Figure 4b). Alpha scorpion toxins have two surfaces that interact with the Na_v channel: a "core domain", formed by loop residues that connect the secondary structure elements, and an "NC domain", formed by a five-residue turn near the N terminus (residues 8–12) and the C-terminal region.[21] In the core domain of Ts3, the Tyr^{19} side chain was found to overlap with the side chain of the functionally important Trp^{38} of other alpha scorpion toxins (Figure 4c). The NC domain of Ts3 is quite similar to the NC domain of the other members of the alpha scorpion toxin family in terms of both structure (Figure 4d), and surface electrostatics. The C-terminal region of Ts3 and similar alpha scorpion toxins is heavily populated with multiple positively charged residues. The side chains of the conserved Asn^{11} and Lys^{57} of Ts3 form a hydrogen-bonding network with the main-chain amide at the C terminus which, together with a disulfide bond at Cys^{62}, holds the C terminus in position for interacting with the sodium channel (Figure 4e).

A robust total synthesis of the Ts3 protein using a modular synthetic strategy enabled us to prepare the four suggested covalent structures for the Ts3 toxin. Correlating biological activity measurements with the known structures of these synthetic proteins verified the covalent structure of biologically active Ts3 toxin. By using racemic protein crystallization, diffraction-quality crystals of racemic Ts3 protein were obtained, and from the resulting X-ray diffraction data we determined the three-dimensional structure of biologically active Ts3 toxin. The covalent and tertiary structures of biologically active Ts3 toxin reported herein constitute essential information for future work aimed at developing the Ts3 protein as a biophysical probe of the Na_v channel[22] in order to complement and extend the information obtained by using labeled Ts1 neurotoxin.[23]

Acknowledgements

This research was supported in part by funds from NIH Grants U54 GM087519, and R01-GM030376 to F. Bezanilla and by the American Heart Association (13POST14800031) to T. Kubota. Use of NE-CAT beamline 24-ID at the Advanced Photon Source is supported by the National Institute of General Medical Sciences from the National Institutes of Health (P41 GM103403). Use of the Advanced Photon Source is supported by the U.S. Department of Energy, Office of Basic Energy Sciences, under Contract No. DE-AC02-06CH11357.

Keywords: proteins · protein structures · chemical protein synthesis · toxins · X-ray crystallography

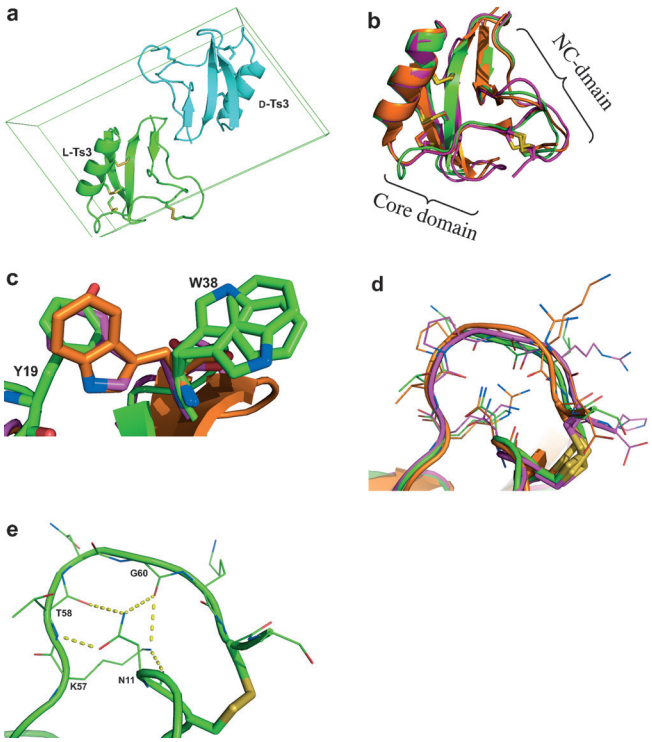


Figure 4. Crystal structure of Ts3 toxin. a) Crystal structure of the D/L racemate of Ts3(1–64)-$CONH_2$ protein (PDB ID: 5CY0). Cartoon representation of the packing of D-Ts3(1–64)-$CONH_2$ (cyan) and L-Ts3(1–64)-$CONH_2$ (green) proteins in the unit cell in $P\bar{1}$. b) Structure alignment of Ts3 toxin (green), alpha scorpion toxin LQH-α-IT (yellow), and Aah2 (magenta).[20] c) Core-domain hydrophobic-surface comparison between Ts3 and other alpha scorpion toxins. d) NC-domain comparison between Ts3 and other alpha scorpion toxins. e) Hydrogen bonding of Asn^{11} and Lys^{57} with residues in the C-terminal region of Ts3 toxin.

How to cite: *Angew. Chem. Int. Ed.* **2016**, *55*, 8639–8642
Angew. Chem. **2016**, *128*, 8781–8784

[1] L. D. Possani, B. M. Martin, J. Mochcamorales, I. Svendsen, *Carlsberg Res. Commun.* **1981**, *46*, 195.
[2] J. C. Rogers, Y. Qu, T. N. Tanada, T. Scheuer, W. A. Catterall, *J. Biol. Chem.* **1996**, *271*, 15950.
[3] G. R. Benzinger, J. W. Kyle, K. M. Blumenthal, D. A. Hanck, *J. Biol. Chem.* **1998**, *273*, 80.
[4] W. A. Catterall, *J. Biol. Chem.* **1977**, *252*, 8660.
[5] F. Couraud, H. Rochat, S. Lissitzky, *Biochem. Biophys. Res. Commun.* **1978**, *83*, 1525.
[6] G. E. Kirsch, A. Skattebol, L. D. Possani, A. M. Brown, *J. Gen. Physiol.* **1989**, *93*, 67.
[7] F. V. Campos, F. I. Coronas, P. S. L. Beirão, *Br. J. Pharmacol.* **2004**, *142*, 1115.
[8] F. V. Campos, B. Chanda, P. S. Beirao, F. Bezanilla, *J. Gen. Physiol.* **2008**, *132*, 251.
[9] C. T. Cologna, S. Marcussi, J. R. Giglio, A. M. Soares, E. C. Arantes, *Protein Pept. Lett.* **2009**, *16*, 920.
[10] M. F. Martin-Eauclaire, B. Céard, A. M. Ribeiro, C. R. Diniz, H. Rochat, P. E. Bougis, *FEBS Lett.* **1994**, *342*, 181.
[11] C. E. Teixeira, D. R. Ifa, G. Corso, V. Santagada, G. Caliendo, E. Antunes, G. De Nucci, *FASEB J.* **2003**, *17*, 485.
[12] B. Dang, T. Kubota, A. M. Correa, F. Bezanilla, S. B. H. Kent, *Angew. Chem. Int. Ed.* **2014**, *53*, 8970; *Angew. Chem.* **2014**, *126*, 9116.
[13] A. L. Saucedo, D. Flores-Solis, R. C. Rodríguez de la Vega, B. Ramírez-Cordero, R. Hernández-López, P. Cano-Sánchez, R. Noriega Navarro, J. García-Valdés, F. Coronas-Valderrama, A. de Roodt, L. G. Brieba, L. Domingos Possani, F. del Río-Portilla, *J. Biol. Chem.* **2012**, *287*, 12321.
[14] B. L. Pentelute, Z. P. Gates, V. Tereshko, J. L. Dashnau, J. M. Vanderkooi, A. A. Kossiakoff, S. B. H. Kent, *J. Am. Chem. Soc.* **2008**, *130*, 9695.
[15] B. Dang, T. Kubota, K. Mandal, F. Bezanilla, S. B. H. Kent, *J. Am. Chem. Soc.* **2013**, *135*, 11911.
[16] T. O. Yeates, S. B. H. Kent, *Annu. Rev. Biophys.* **2012**, *41*, 41.
[17] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, *J. Appl. Crystallogr.* **2007**, *40*, 658.
[18] P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, *Acta Crystallogr. Sect. D* **2010**, *66*, 213.
[19] P. V. Luzzati, *Acta Crystllogr.* **1952**, *5*, 802.
[20] G. D. Smith, R. H. Blessing, S. E. Ealick, J. C. Fontecilla-Camps, H. A. Hauptman, D. Housset, D. A. Langs, R. Miller, *Acta Crystallogr. Sect. D* **1997**, *53*, 551.
[21] I. Karbat, F. Frolow, O. Froy, N. Gilles, L. Cohen, M. Turkov, D. Gordon, M. Gurevitz, *J. Biol. Chem.* **2004**, *279*, 31679.
[22] P. R. Selvin, *Annu. Rev. Biophys. Biomol. Struct.* **2002**, *31*, 275.
[23] T. Kubota, P. Brugarolas, B. Dang, R. K. Finolurdaneta, L. Frezza, R. J. French, S. B. H. Kent, F. Bezanilla, A. M. Correa, *Biophys. J.* **2014**, *106*, 133a.

Received: April 9, 2016
Published online: May 31, 2016