Protein Chemistry

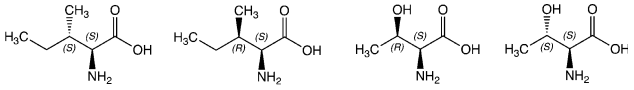
International Edition: DOI: 10.1002/anie.201612398
German Edition: DOI: 10.1002/ange.201612398

Inversion of the Side-Chain Stereochemistry of Indvidual Thr or Ile Residues in a Protein Molecule: Impact on the Folding, Stability, and Structure of the ShK Toxin

Bobo Dang, Rong Shen, Tomoya Kubota, Kalyaneswar Mandal, Francisco Bezanilla, Benoit Roux, and Stephen B. H. Kent*

Abstract: *ShK toxin is a cysteine-rich 35-residue protein ion-channel ligand isolated from the sea anemone Stichodactyla helianthus. In this work, we studied the effect of inverting the side chain stereochemistry of individual Thr or Ile residues on the properties of the ShK protein. Molecular dynamics simulations were used to calculate the free energy cost of inverting the side-chain stereochemistry of individual Thr or Ile residues. Guided by the computational results, we used chemical protein synthesis to prepare three ShK polypeptide chain analogues, each containing either an allo-Thr or an allo-Ile residue. The three allo-Thr or allo-Ile-containing ShK polypeptides were able to fold into defined protein products, but with different folding propensities. Their relative thermal stabilities were measured and were consistent with the MD simulation data. Structures of the three ShK analogue proteins were determined by quasi-racemic X-ray crystallography and were similar to wild-type ShK. All three ShK analogues retained ion-channel blocking activity.*

Proteins are chiral molecules. Natural proteins have the L-configuration since they are made by ribosomal translation and thus consist only of L-amino acids and the achiral amino acid glycine. Of the 20 principal genetically-encoded proteinogenic amino acids, isoleucine and threonine uniquely have a second chiral center in their side chains. Inversion of the alpha-carbon chiral center of most natural L-amino acids will generate the mirror-image D-amino acids. In the case of isoleucine and threonine, both chiral centers must be inverted to generate the mirror-image amino acids D-isoleucine and D-threonine. Inversion of stereochemistry only at the alpha carbon will generate the diastereomeric amino acids D-alloisoleucine (D-allo-Ile) and D-allothreonine (D-allo-Thr). Conversely, inversion of the alpha carbon stereochemistry in D-Ile and D-Thr generates L-alloisoleucine (L-allo-Ile) and L-allothreonine (L-allo-Thr;[1] Scheme 1). There has been only very limited work on the effect of inversion of Thr and/or Ile side-chain stereochemistry on the properties of globular protein molecules.[2]



Scheme 1. Stereochemistry of L-Ile, L-allo-Ile and L-Thr, L-allo-Thr.[1].

The small protein toxin ShK was originally isolated from the sea anemone *Stichodactyla helianthus*.[3] It binds to the Kv1.3 ion channel with very high affinity.[3] There are 4 Thr and 2 Ile residues in the 35 amino acid ShK toxin polypeptide chain (Scheme 2).

RSCIDTIPKS[10] RCTAFQCKHS[20] MKYRLSFCRK[30] TCGTC[35]

Scheme 2. Amino acid sequence of the ShK toxin protein. The native chemical ligation site used in the synthesis is underlined.

Recently, we reported a convergent total synthesis of wild-type ShK toxin and its mirror-image form (D-ShK) for determination of the ShK protein structure by racemic protein X-ray crystallography.[4] Among other things, we showed that D-ShK protein does not bind to the Kv1.3 ion channel. Surprisingly, it had been reported that a diastereomer of the mirror-image ShK protein, "D-allo-ShK", retained its biological activity and could bind to the Kv1.3 ion channel.[5] This "D-allo-ShK" was described as being made up of D-amino acids (and Gly) but with the natural stereochemistry of the chiral side chains of the Ile and Thr residues, that is, D-allo-Ile and D-allo-Thr.

In this work, In the work reported here, as a prelude to a systematic reinvestigation of the structure and properties of the D-allo-ShK protein molecule, we set out to explore the effects of inverting the side chain stereochemistry of *individual* Thr or Ile residues on the foldability/stability/structure/biological activity of the ShK protein, and to correlate biological activities with atomic-resolution structures of ShK analogues containing allo-amino acids.

First, we calculated the free-energy cost of substituting individual Thr or Ile residues in ShK protein with an allo-Thr or allo-Ile residue, respectively, using bidirectional alchemical free energy perturbation molecular dynamics (FEP/MD) simulations. The FEP/MD calculations allowed us to estimate

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

Supporting information for this article can be found under: http://dx.doi.org/10.1002/anie.201612398.

the influence of inverting the side-chain stereochemistry of each of the six Thr/Ile amino acid residues on the stability of the ShK protein. The molecular systems were set up using the VMD program[6] and simulated with the NAMD program.[7] The results of the forward and the backward FEP/MD simulations were then combined using the Bennett acceptance ratio method[8] to calculate the free energy change for each of the Thr/allo-Thr or Ile/allo-Ile mutations within the VMD plugin ParseFEP.[9] The results are shown in Figure 1.

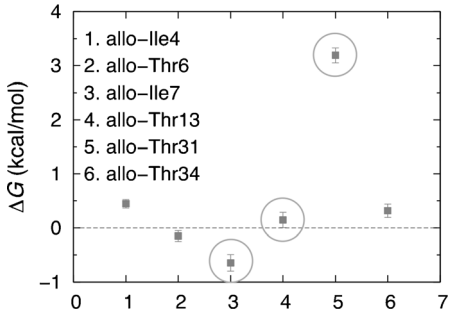


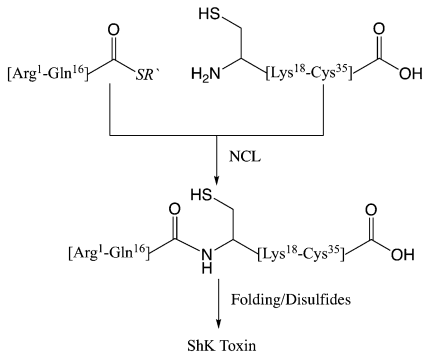
Figure 1. Calculated free-energy differences between ShK protein and allo-amino acid ShK analogues. Error bars represent standard deviations of the means.

To be rigorous, it would be necessary to close the thermodynamic cycle of FEP/MD by including the free-energy cost of the chirality change in the unfolded protein. However, because the aqueous solvent is not a chiral environment, the solvation free energy of the side-chain moiety is invariant upon such a change. Thus, it can be reasonably assumed that the free-energy difference is negligible in the unfolded protein.

From these MD calculations, we found that replacing Ile 7 with allo-Ile7 will *stabilize* the folded structure by a free energy of 0.7 kcal mol^{-1}; and replacing Thr 31 with allo-Thr 31 will *destabilize* the folded structure by a free energy of 3.2 kcal mol^{-1}. Other Thr/Ile residues did not show significant effects upon substitution with allo-Thr/allo-Ile. (Figure 1) Therefore, we chose three amino acid sites (Ile 7, Thr 13, and Thr 31) for further investigation to determine the impact of side-chain chirality inversion of individual Ile and Thr residues on the stability of the ShK protein.

Authentication of the identity and stereochemistry of the protected L-allo-Ile and L-allo-Thr amino acids used is described in the Supporting Information. All of the ShK polypeptide chains were synthesized from two synthetic peptide segments by following the method that was used for ShK toxin (Scheme 3).[4] First, we chemically synthesized peptides [allo-Ile 7]Arg 1–Gln 16-*COSR*, [allo-Thr 13]Arg 1–Gln 16-*COSR*, and [allo-Thr 31]Cys 17–Cys 35 using highly optimized Boc chemistry "in situ neutralization" solid phase peptide synthesis (SPPS)[10] The two appropriate peptide segments were then covalently condensed by native chemical ligation[11] at the -Gln 16–Cys 17- site.

The resulting synthetic polypeptides were purified by HPLC on a C4 reverse-phase semi-prep column; LCMS data for the purified allo-amino acid containing polypeptide chains are shown in Figure 2.



Scheme 3. Convergent synthesis of ShK toxin by native chemical ligation of two unprotected peptide segments, followed by folding and formation of disulfides. The figure is adapted from Ref. [4].

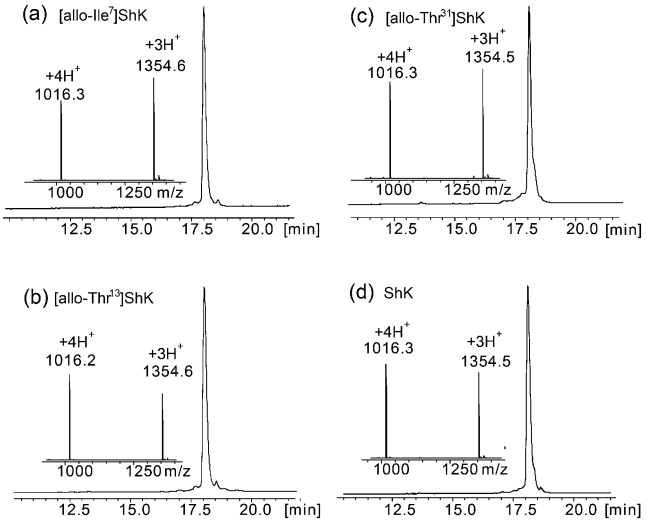


Figure 2. LCMS data for the three allo-amino acid containing ShK analogues and the wild-type ShK polypeptide chain. a) [allo-Ile 7]ShK polypeptide. Mass: Obsd. 4061.0 ± 0.2 Da, calcd 4060.8 Da (av. isotope composition). b) [allo-Thr 13]ShK polypeptide. Mass: Obsd. 4060.8 ± 0.2 Da, calcd 4060.8 Da (av. isotope composition). c) [allo-Thr 31]ShK polypeptide. Mass: Obsd. 4060.9 ± 0.2 Da, calcd 4060.8 Da (av. isotope composition). d) ShK polypeptide. Mass: Obsd. 4060.9 ± 0.2 Da, calcd 4060.8 Da (av. isotope composition). The MS data shown [insets] were collected across the entire UV-absorbing main peak in each chromatogram.

Once we had obtained the three allo-amino acid containing ShK analogues, we carried out separate folding reactions for each synthetic polypeptide under the conditions that were used to fold wild-type ShK toxin: 50 mM $AcONH_4$, pH 8.0, [polypeptide] 0.4 mg mL^{-1}, air oxidation (without stirring). Each of the four folding reactions was monitored at chosen time points by HPLC analysis. Data for the folding reactions at time points of 3 hours and 18 hours are shown in Figure 3.

Folded [allo-Ile 7]ShK protein, which has mass 6 Da less than the [allo-Ile 7]ShK polypeptide, corresponding to the formation of three disulfide bonds, was observed to form most rapidly (ca. 30 minutes), followed by wild-type L-ShK (ca. 1 hour), [allo-Thr 13]ShK (ca. 2 hours), and [allo-Thr 31]ShK (ca. 3 hours). For preparative-scale folding reactions, the yield

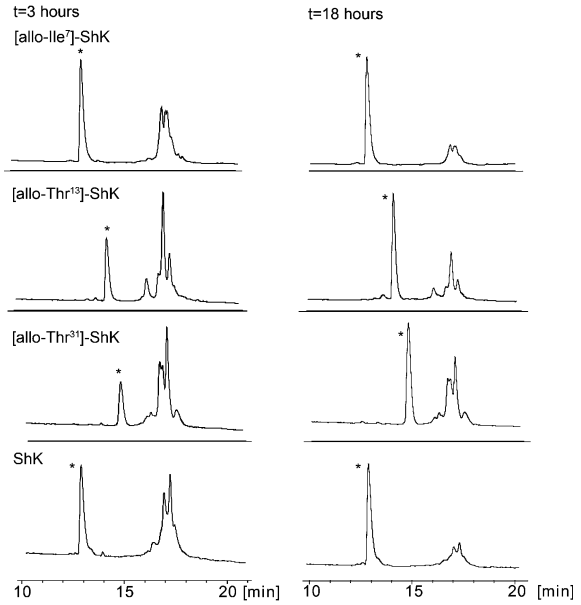


Figure 3. Folding the three allo-amino acid containing ShK analogues and wild-type ShK. The folding reactions were monitored by LCMS (MS data not shown). Analytical HPLC data are shown for each folding reaction at time points 3 hours (left panel) and 18 hours (right panel). Folding conditions: 50 mM $AcONH_4$, pH 8.0, [polypeptide] 0.4 mg mL^{-1}, air oxidation (without stirring). * indicates the folded protein products.

of each folded protein product was calculated based on HPLC analysis (peak integrations). The folded product yields were: [allo-Ile 7]ShK 67%, ShK 61%, [allo-Thr 13]ShK 51% and [allo-Thr 31]ShK 40%.

In the HPLC analysis of these folded products (Figure 3), we observed that under identical reverse-phase analytical HPLC conditions, the folded protein molecules [allo-Ile 7]ShK and ShK eluted at approximately the same time, while [allo-Thr 13]ShK eluted later, and [allo-Thr 31]ShK eluted later still. Our structural studies (see below) indicated that the crystal structures of these four protein molecules were essentially the same. The reason for the difference in retention times might be that the folded products have different stability properties under the denaturing reverse-phase HPLC conditions: if the less stable ShK analogues were partially denatured so that more hydrophobic surfaces were exposed, it would lead to the later retention times observed on HPLC analysis.

To study the thermal stability of these ShK protein analogues, we performed CD experiments to measure the thermal melting temperature. Each purified ShK analogue (see the Supporting Information) was dissolved in 10 mM PBS buffer at pH 7.4, at a protein concentration of 0.3 mg mL^{-1}. Peak absorptions were measured at a wavelength of 220 nm. After fitting the experiment data with a sigmoid function, the approximate thermal melting temperatures were: wild type ShK ca. 120 °C; [allo-Ile 7]ShK ca. 120 °C; [allo-Thr 13]ShK ca. 100 °C; and, [allo-Thr 31]ShK ca. 80 °C (see the Supporting Information).

This order of ShK protein thermal stabilities matches our observations on the retention times in HPLC analysis, where we hypothesized that less stable analogues were partially denatured and thus had later retention times, and with the rates of formation of the folded protein molecules in our analytical folding studies, where the most stable protein folded first and gave the highest yields.

Next, we set out to determine the crystal structures of the synthetic ShK protein diastereomers. We have reported the use of quasi-racemic protein mixtures to facilitate protein crystallization.[12,13] Since we already had the D-ShK protein from previous studies,[4] we attempted to crystallize the three new ShK analogues by using quasi-racemic protein crystallization under the seven conditions that produced true racemate ShK protein crystals. We performed crystallization trials under identical conditions using both a quasi-racemic protein mixture and conventional L-protein alone.

For [allo-Ile 7]ShK, none of the conditions that we screened produced any crystals after one week. At that point, when we checked for crystals, we accidentally spilled well solution into the quasi-racemic mixture hanging drop in one condition, and this one condition produced crystals after another week. Later, we tried to mimic this adventitious condition by mixing 1.6 μL well solution and 0.8 μL protein solution to form the hanging drop, and this method indeed produced crystals after two weeks. In the case of [allo-Thr 13]ShK, two conditions produced crystals overnight from the quasi-racemic mixture. For [allo-Thr 31]ShK, six of the seven conditions (both quasi-racemic mixture and L-protein alone) produced needle-shaped crystals overnight. Conditions used to produce diffraction-quality crystals for each analogue are given in the Supporting Information. Diffraction resolution and space group data are summarized in Table 1.

Table 1: X-ray diffraction data for the allo-amino acid ShK analogues.

	Resolution [Å]	Space group	Molecules in asymmetric unit
[allo-Ile 7]ShK/D-ShK	1.20	*C*2	1 [allo-Ile 7]ShK and 1 D-ShK
[allo-Thr 13]ShK/D-ShK	0.90	*P*1	1 [allo-Thr 13]ShK and 1 D-ShK
[allo-Thr 31]ShK/D-ShK	1.30	$P2_1$	3 [allo-Thr 31]ShK
[allo-Thr 31]ShK	1.56	$P2_1$	3 [allo-Thr 31]ShK

The X-ray structure of quasi-racemic crystalline L-[allo-Ile 7]ShK/D-ShK was solved by molecular replacement[14] using the true racemate L/D-ShK (PDB ID: 4LFS) as a search model. The asymmetric unit contained one L-[allo-Ile 7]ShK protein and one D-ShK protein. The final model was refined to a crystallographic R-factor of 0.14 (R-free 0.20) using CCP4.[15] The X-ray structure of quasi-racemic crystalline L-[allo-Thr 13]ShK/D-ShK was solved by molecular replacement[14] using PDB ID: 4LFS as a search model. The asymmetric unit contained one L-[allo-Thr 13]ShK protein and one D-ShK protein. The final model was refined to a crystallographic R-factor of 0.11 (R-free 0.13) using CCP4.[15]

For [allo-Thr 31]ShK, the fact that all of the crystals have three molecules in each asymmetric unit, and all of the crystals (from both quasi-racemic crystallization and L-

protein crystallization) obtained have the same needle shape, while racemic ShK protein and the other quasi-racemic ShK proteins produced cube-shaped crystals, thus suggested that the crystals from both the quasi-racemic mixture and the L-protein alone might be the same and that they might only contain [allo-Thr31]ShK. The protein structures in both crystals were solved by molecular replacement[14] using L-ShK (PDB ID: 4LFQ) as the search model. Indeed, neither of the [allo-Thr31]ShK crystals contained the D-ShK molecule; that is, the quasi-racemic mixture gave crystals of the L-[allo-Thr31]ShK protein, not quasi-racemate crystals. The final model of the L-[allo-Thr31]ShK protein was refined to a crystallographic R-factor of 0.14 (R-free 0.20) using CCP4.[15]

The structures of the three allo-amino acid containing ShK protein analogues are all very similar to wild-type ShK protein. The structure comparisons are shown in Figure 4.

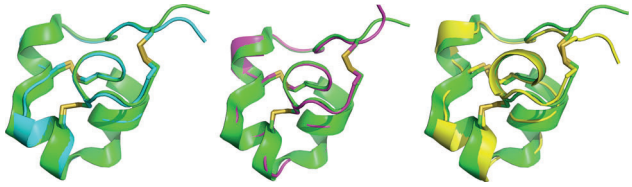


Figure 4. Comparison of the crystal structures of L-ShK protein and the allo-amino acid containing analogues. Left: structure alignment of L-ShK (green) and [allo-Ile7]ShK (cyan) with RMSD 0.38 Å. Center: structure alignment of L-ShK (green) and [allo-Thr13]ShK (magenta) with RMSD 0.34 Å. Right: structure alignment of L-ShK (green) and [allo-Thr31]ShK (yellow) with RMSD 0.76 Å.

Crystal structure data for [allo-Ile7]ShK, [allo-Thr13]ShK, and [allo-Thr31]ShK have been deposited in the Protein Data Bank with PDB IDs 5I5A, 5I5B, and 5I5C, respectively.

To examine the channel-blocking abilities of allo-ShK analogues, we used the cut-open oocyte voltage clamp method to measure potassium ionic currents from *Xenopus laevis* oocytes expressing human Kv1.3 (hKv1.3) ion channels, before and after addition of the toxin. All of the allo-ShK analogues retain Kv1.3-blocking activity. However, their activities are 4–6-times lower than wild type ShK toxin, with no significant difference between the analogues (Table 2).

In conclusion, we have explored the impact of threonine and isoleucine side-chain chirality on the folding and structural stability of the ShK protein. Free energy perturbation molecule dynamics (FEP/MD) simulations helped us to estimate the thermodynamic cost of inverting the stereochemistry of individual Thr/Ile side chains. Chemical protein synthesis enabled us to prepare the allo-Thr/allo-Ile-containing ShK polypeptide analogues and to then perform folding and stability studies. The structures of the synthetic proteins were determined by X-ray crystallography; each of the three allo-amino acid ShK analogues proteins had essentially the same folded structure as wild-type ShK protein. The three allo-amino acid ShK analogues all retained Kv1.3 ion channel blocking activity.

The three allo-Thr/allo-Ile-containing ShK polypeptides have distinct folding properties, and the folded protein products have different stabilities. Indeed, the experimental folding and thermal stability data we obtained were in good agreement with the computational data. These results illustrate the utility of combining computational calculations of protein properties and total protein synthesis enabled by modern chemical ligation methods for the systematic investigation of the molecular basis of protein structure and function.

Table 2: Ability of the ShK analogues to block the Kv1.3 ion channel.

	IC_{50} [pM]	Hill coefficient
Wild type	140 ± 30	1.1 ± 0.1
[Allo-Ile7]ShK	440 ± 220	0.7 ± 0.2
[Allo-Thr13]ShK	540 ± 200	0.8 ± 0.1
[Allo-Thr31]ShK	860 ± 260	0.8 ± 0.1

Experimental uncertainties indicate standard errors of the mean.

Acknowledgements

This research was supported in part by funds from NIH Grants U54 GM087519, and R01-GM030376 to F. Bezanilla. Use of NE-CAT beamline 24-ID-C at the Advanced Photon Source is supported by the National Institute of General Medical Sciences of the National Institutes of Health (P41 GM103403). The Pilatus 6M detector on 24-ID-C beam line is funded by a NIH-ORIP HEI grant (S10 RR029205). 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.

Conflict of interest

The authors declare no conflict of interest.

Keywords: allo-isoleucine · allo-threonine · protein crystallography · protein synthesis · stereochemistry

[1] Nomenclature and Symbolism for Amino Acids and Peptides, *Pure. Appl. Chem.* **1984**, *56*, 595–624.
[2] a) B. Xu, Q. X. Hua, S. H. Nakagawa, W. Jia, Y. C. Chu, P. G. Katsoyannis, M. A. Weiss, *J. Mol. Biol.* **2002**, *316*, 435–441; b) H. H. Chung, D. R. Benson, V. W. Cornish, P. G. Schultz, *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 10145–10149.
[3] O. Castañeda, V. Sotolongo, A. M. Amor, R. Stocklin, A. J. Anderson, A. L. Harvey, A. Engstrom, C. Wernstedt, E. Karlsson, *Toxincon* **1995**, *33*, 603–613.
[4] B. Dang, T. Kubota, K. Mandal, F. Bezanilla, S. B. H. Kent, *J. Am. Chem. Soc.* **2013**, *135*, 11911–11919.
[5] C. Beeton, B. J. Smith, J. K. Sabo, G. Crossley, D. Nugent, I. Khaytin, V. Chi, K. G. Chandy, M. W. Pennington, R. S. Norton, *J. Biol. Chem.* **2008**, *283*, 988–997.
[6] W. Humphrey, A. Dalke, K. Schulten, *J. Molec. Graphics* **1996**, *14*, 33–38.

[7] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, K. Schulten, *J. Comput. Chem.* **2005**, *26*, 1781 – 1802.
[8] C. H. Bennett, *J. Comput. Phys.* **1976**, *22*, 245 – 268.
[9] P. Liu, F. Dehez, W. Cai, C. Chipot, *J. Chem. Theory Comput.* **2012**, *8*, 2606 – 2616.
[10] M. Schnölzer, P. Alewood, A. Jones, D. Alewood, S. B. H. Kent, *Int. J. Pept. Protein Res.* **1992**, *40*, 180 – 193.
[11] P. E. Dawson, T. W. Miur, I. Clark-Lewis, S. B. H. Kent, *Science* **1994**, *266*, 776 – 779.
[12] 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 – 9701.
[13] B. Dang, T. Kubota, K. Mandal, A. M. Correa, F. Bezanilla, S. B. H. Kent, *Angew. Chem. Int. Ed.* **2016**, *55*, 8639 – 8642; *Angew. Chem.* **2016**, *128*, 8781 – 8784.
[14] 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 – 674.
[15] G. N. Murshudov, P. Skubak, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long, A. A. Vagin, *Acta. Crystallogr. Sect. D* **2011**, *67*, 355 – 367.

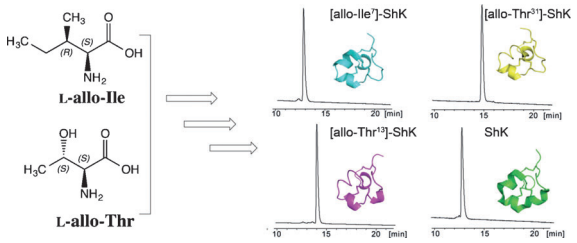
Manuscript received: December 21, 2016
Final Article published: ■■ ■■, ■■■■

Communications

Protein Chemistry

B. Dang,* R. Shen, T. Kubota, K. Mandal, F. Bezanilla, B. Roux, S. B. H. Kent _____ ■■■■–■■■■

Inversion of the Side-Chain Stereochemistry of Indvidual Thr or Ile Residues in a Protein Molecule: Impact on the Folding, Stability, and Structure of the ShK Toxin



Ch-ch-ch-changes: Guided by molecular dynamics calculations, chemical protein synthesis was used to explore the impact of allo-Thr and allo-Ile substitutions for individual Thr and Ile residues on the folding, crystal structure, and stability of the ShK protein. The experimental folding and thermal stability data matched well with the computational results.

