Published in final edited form as:

Angew Chem Int Ed Engl. 2017 March 13; 56(12): 3324–3328. doi:10.1002/anie.201612398.

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

Dr. Bobo Dang,

Department of Chemistry, Department of Biochemistry & Molecular Biology, Institute for Biophysical Dynamics, University of Chicago Chicago, IL 60637

Dr. Rong Shen,

Department of Biochemistry & Molecular Biology, Institute for Biophysical Dynamics, University of Chicago Chicago, IL 60637

Dr. Tomoya Kubota,

Department of Biochemistry & Molecular Biology, Institute for Biophysical Dynamics, University of Chicago Chicago, IL 60637

Dr. Kalyaneswar Mandal,

Department of Chemistry, Department of Biochemistry & Molecular Biology, Institute for Biophysical Dynamics, University of Chicago Chicago, IL 60637

Prof. Dr. Francisco Bezanilla,

Department of Biochemistry & Molecular Biology, Institute for Biophysical Dynamics, University of Chicago Chicago, IL 60637

Prof. Dr. Benoit Roux, and

Department of Biochemistry & Molecular Biology, Institute for Biophysical Dynamics, University of Chicago Chicago, IL 60637

Prof. Dr. Stephen B. H. Kent

Department of Chemistry, Department of Biochemistry & Molecular Biology, Institute for Biophysical Dynamics, University of Chicago Chicago, IL 60637

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 impact of inverting the side chain stereochemistry of individual Thr and Ile residues on the properties of the ShK protein molecule. Molecular dynamics simulations were used to calculate the free energy cost of inverting the side chain stereochemistry of individual Thr/Ile residues. Guided by the computational results, we used chemical protein synthesis to prepare three ShK polypeptide chain analogs each containing either an allo-Thr or an allo-Ile residue. The three allo-Thr/allo-Ile containing ShK polypeptides were each able to fold into a defined protein product, but with different folding propensities. Relative thermal stabilities of these ShK analogs were measured and were in agreement with the MD

simulation data. Structures of the three ShK analogs were determined by quasi-racemic X-ray crystallography and were similar to wild type ShK protein. All three ShK analogs retained Kv1.3 ion channel blocking activity.

Abstract

Text for table of content: Guided by molecular dynamics calculations, we used chemical protein synthesis to explore the impact of allo-Thr and allo-Ile substitutions for Thr and Ile residues on the folding, crystal structure, and stability of the ShK protein molecule. The experimental folding and thermal stability data we obtained matched well with the computational results.

Keywords

ShK toxin; allothreonine; alloseuleucine; folding; chemical protein synthesis; Kv1.3 ion channel; quasi-racemic protein crystallography

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. The 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 the 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 impact of inversion of Thr and/or Ile side chain stereochemistry on the properties of globular protein molecules.^[2]

The small protein toxin ShK is 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 residue ShK toxin polypeptide chain (Scheme 2).

Recently we reported a convergent total synthesis of wild type ShK toxin and its mirror image protein form, D-ShK, for the determination of the ShK protein structure by racemic protein X-ray crystallography.^[4] Inter alia, we showed that D-ShK protein did 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, i.e. D-allo-Ile and D-allo-Thr.

In the work reported here, as a prelude to the 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, Ile residues on the foldability/stability/structure/biological activity of the ShK protein molecule, and to correlate biological activities with atomic resolution structures of allo-amino acid-containing ShK protein analogues.

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

Rigorously, it would be necessary to close the thermodynamical 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 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 Ile⁷ with allo-Ile⁷ will *stabilize* the folded structure by a free energy of 0.7 kcal/mol; and, replacing Thr³¹ with allo-Thr³¹ will *destabilize* the folded structure by a free energy of 3.2 kcal/mol. Other Thr/Ile residues did not show significant effects upon substitution by allo-Thr/allo-Ile. (Figure 1) Therefore, we chose three amino acid sites (Ile⁷, Thr¹³ and Thr³¹) for further investigation to determine the impact of side chain chirality inversion in individual Ile and Thr residues on the stability of the ShK protein molecule.

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 the ShK polypeptide chains were synthesized from two synthetic peptide segments, following the method that was used for ShK toxin (Scheme 3).^[4] First, we chemically synthesized peptides [allo-Ile⁷]Arg¹-Gln¹⁶-COSR, [allo-Thr¹³]Arg¹-Gln¹⁶-COSR, and [allo-Thr³¹]Cys¹⁷-Cys³⁵ 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 –Gln¹⁶-Cys¹⁷- site.

The 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.

Once we obtained the three allo-amino acid containing ShK analog polypeptides, we carried out separate folding reactions for each synthetic polypeptide under the conditions that were used to fold wild type ShK toxin: 50 mM AcONH₄, pH 8.0, [polypeptide] 0.4 mg/ml, air oxidation (without stirring). Each of the four folding reactions was monitored at chosen time points by HPLC analysis. Folding reactions at time points of 3 hours and 18 hours are shown in Figure 3.

Folded [allo-Ile⁷]ShK protein with mass 6 Da less than [allo-Ile⁷]ShK polypeptide, corresponding to the formation of three disulfide bonds, was observed to form most rapidly (~30 minutes), followed by wild-type L-ShK (1 hour), [allo-Thr¹³]ShK (~2 hours), and [allo-Thr³¹]ShK t (~3 hours). For preparative-scale folding reactions, the yield of each folded protein product was calculated based on HPLC analysis (peak integrations). The folded product yields were: [L-allo-Ile⁷]ShK 67%, L-ShK 61%, [L-allo-Thr¹³]ShK 51% and [L-allo-Thr³¹]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 [L-allo-Ile⁷]ShK and L-ShK eluted at approximately the same time, while [L-allo-Thr¹³]ShK later, and [L-allo-Thr³¹]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 retention time difference might be that the folded products have different stability properties under the denaturing reverse phase HPLC conditions: thus, if the less stable ShK analogs 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 analogs, we performed CD experiments to measure the thermal melting temperature. Each purified ShK protein analog (see Supporting Information) was dissolved in 10 mM PBS buffer at pH 7.4, at a protein concentration of 0.3 mg/mL. Peak absorptions were measured at 220 nm wavelength. After fitting the experiment data with a sigmoid function, the approximate thermal melting temperatures were: wild type L-ShK, ~120 °C; [L-allo-Ile⁷]ShK, ~120 °C; [L-allo-Thr¹³]ShK, ~100 °C; and, [L-allo-Thr³¹]ShK, ~80 °C (see Supporting Information).

This order of ShK protein stabilities matches our observations on the retention times in HPLC analysis, where we hypothesized that less stable analogs 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 molecule from previous studies,^[4] we attempted to crystallize the three new ShK analogs 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 screen.

For [allo-Ile⁷]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 μ L well solution and 0.8 μ L protein solution to form the hanging drop, and this method indeed produced crystals after two weeks. In the case of [allo-Thr¹³]ShK, two conditions produced crystals overnight from the quasi-racemic mixture. For [allo-Thr³¹]ShK, six of the seven conditions (both quasi-racemic mixture and L-protein alone) produced needle shape crystals overnight. Conditions used to produce diffraction quality crystals for each analog are given in the Supporting Information. Diffraction resolution and space group data are summarized in Table 1.

The X-ray structure of quasi-racemic crystalline L-[allo-Ile⁷]ShK/D-ShK was solved by molecular replacement^[14] using as a search model the true racemate L/D-ShK (PDB ID: 4LFS). The asymmetric unit contained one L-[allo-Ile⁷]ShK protein molecule and one D-ShK protein molecule. 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-Thr¹³]ShK/D-ShK was solved by molecular replacement^[14] using PDB ID: 4LFS as a search model. The asymmetric unit contained one L-[allo-Thr¹³]ShK protein molecule and one D-ShK protein molecule. The final model was refined to a crystallographic R-factor of 0.11 (R-free 0.13) using CCP4.^[15]

For [allo-Thr³¹]ShK, the fact that all the crystals have 3 molecules in each asymmetric unit, and all 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 cubic shape crystals, 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 structures of 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 0.19 (R-free 0.23) using CCP4.^[15]

The structures of the three allo-amino acid containing ShK protein analogs are all very similar to wild type ShK protein. The structure comparisons are shown in Figure 4. Crystal structure data for [allo-Ile⁷]ShK, [allo-Thr¹³]ShK, and [allo-Thr³¹]ShK have been deposited in the Protein Data Bank with PDB codes 5I5A, 5I5B and 5I5C, respectively.

To examine the channel-blocking abilities of allo-ShK analogs, we used the cut-open oocyte voltage clamp method to measure potassium ionic currents from *Xenopus laevis* oocytes that expressed human Kv1.3 (hKv1.3) channels, before and after addition of toxin. All allo-ShK

analogs retain the Kv1.3 ion channel blocking activity. However, their activities are 4-to-6 times lower than wild type ShK toxin with no significant difference between the analogs (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 molecule. 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 chain analogs 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 containing ShK analog proteins had essentially the same folded structure as wild type ShK protein. The three allo-amino acid containing ShK analogs all retained Kv1.3 ion channel blocking activity.

The three allo-Thr/allo-Ile containing ShK polypeptides indeed 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.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

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.

References

1. Pure. Appl. Chem. 1984; 56:595–624.
2. (a) Xu B, Hua QX, Nakagawa SH, Jia W, Chu YC, Katsoyannis PG, Weiss MA. J. Mol. Biol. 2002; 316:435–441. [PubMed: 11866509] (b) Chung HH, Benson DR, Cornish VW, Schultz PG. Proc. Nat. Acad. Sci. USA. 1993; 90:10145–10149. [PubMed: 8234268]
3. Castaneda O, Sotolongo V, Amor AM, Stocklin R, Anderson AJ, Harvey AL, Engstrom A, Wernstedt C, Karlsson E. Toxincon. 1995; 33:603–613.
4. Dang B, Kubota T, Mandal K, Bezanilla F, Kent SBH. J. Am. Chem. Soc. 2013; 135:11911–11919. [PubMed: 23919482]
5. Beeton C, Smith BJ, Sabo JK, Crossley G, Nugent D, Khaytin I, Chi V, Chandy KG, Pennington MW, Norton RS. J. Biol. Chem. 2008; 283:988–997. [PubMed: 17984097]
6. Humphrey W, Dalke A, Schulten K K. J. Molec. Graphics. 1996; 14:33–38.
7. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K. J. Computational Chemistry. 2005; 26:1781–1802.
8. Bennett CH. J. Computational Physics. 1976; 22:245–268.

9. Liu P, Dehez F, Cai W, Chipot C, Chem J. Theor. Comput. 2012; 8:2606–2616.
10. Schnölzer M, Alewood P, Jones A, Alewood D, Kent SBH. Int. J. Pept. Protein Res. 1992; 40:180–193. [PubMed: 1478777]
11. Dawson PE, Miur TW, Clark-Lewis I, Kent SBH. Science. 1994; 266:776–779. [PubMed: 7973629]
12. Pentelute BL, Gates ZP, Tereshko V, Dashnau JL, Vanderkooi JM, Kossiakoff AA, Kent SBH. J. Am. Chem. Soc. 2008; 130:9695–9701. [PubMed: 18598029]
13. Dang B, Kubota T, Mandal K, Correa AM, Bezanilla F, Kent SBH. Angew. Chem. Int. Ed. Engl. 2016; 55:8639–8642. [PubMed: 27244051]
14. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. J. Appl. Crystallogr. 2007; 40:658–674. [PubMed: 19461840]
15. Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F, Vagin AA. Acta. Crystallogr. 2011; D67:355–367.

Figure 1.

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

Figure 2.

LCMS data for the three allo-amino acid containing ShK and wild type ShK polypeptide chains. (a) [allo-Ile⁷]ShK polypeptide. Mass: Obsd. 4061.0 ± 0.2 Da, Calc. 4060.8 Da (av. isotope composition). (b) [allo-Thr¹³]ShK polypeptide. Mass: Obsd. 4060.8 ± 0.2 Da, Calc. 4060.8 Da (av. isotope composition). (c) [allo-Thr³¹]ShK polypeptide. Mass: Obsd. 4060.9 ± 0.2 Da, Calc. 4060.8 Da (av. isotope composition). (d) ShK polypeptide. Mass: Obsd. 4060.9 ± 0.2 Da, Calc. 4060.8 Da (av. isotope composition). The MS data shown [Insets] were collected across the entire UV absorbing main peak in each chromatogram.

Figure 3.

Folding the three allo-amino acid containing ShK analogs 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 point 3 hours (left panel) and 18 hours (right panel). Folding conditions: 50 mM AcONH₄, pH 8.0, [polypeptide] 0.4 mg/ml, air oxidation (without stirring). * indicates the folded products.

Figure 4.

Comparison of the crystal structures of L-ShK protein and different allo-amino acid containing ShK protein analogs. Left: structure alignment of L-ShK (green) and [allo-Ile⁷]ShK (cyan) with RMSD 0.38 Å; Middle: 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.

Scheme 1.

Stereochemistries of L-Ile, L-allo-Ile and L-Thr, L-allo-Thr.^[1]

RSCIDTIPKS¹⁰ RCTAFQCKHS²⁰ MKYRLSFCRK³⁰ TCGTC³⁵

Scheme 2.

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

Scheme 3.

Convergent synthesis of ShK toxin by native chemical ligation of two unprotected peptide segments, followed by folding and formation of disulfides. Figure is adapted from reference [4].

Table 1

X-ray diffraction data for allo-amino acid containing ShK analogs

	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	1.22	P 2 ₁	3 [allo-Thr ³¹]ShK
[allo-Thr ³¹]ShK	1.56	P 2 ₁	3 [allo-Thr ³¹]ShK

Table 2

ShK analogs blocking ability of Kv1.3 ion channel

	IC ₅₀ (pM)	Hill co-efficient
Wild type	140 ± 30	1.1 ± 0.1
[Allo-Ile ⁷]ShK	440 ± 220	0.7 ± 0.2
[Allo-Thr ¹³]ShK	540 ± 200	0.8 ± 0.1
[Allo-Thr ³¹]ShK	860 ± 260	0.8 ± 0.1