Appendix D

Recombinant expression and purification of a xylanase from the thermophilic fungus *T. aurantiacus*

Abstract

The *Thermoascus aurantiacus* xylanase 10A (TAX) is of interest for industrial and enzyme engineering purposes because of its ability to hydrolyze polymers of xylose at elevated temperatures. Previous expression of this xylanase has relied on natural production from fungal cell cultures, limiting the quantities that can be produced and hindering genetic manipulation that could be used to optimize existing activity or engineer novel activities. Here, we optimize the gene sequence of TAX for expression in *E. coli* and successfully express the protein in quantities up to 150 mg per liter of culture. In addition, a quick single-step purification protocol is developed that results in highly pure protein.

Introduction

Endo-β-1,4-xylanases are of industrial interest because of their hydrolytic activity against the internal glycosidic bonds of xylan, which is the major polysaccharide component of plant cell walls. The optimization of the stability and catalytic activity of such enzymes has potential applications in paper bleaching as well as in food and animal feed processing. Thus, thermostable xylanases such as xylanase 10A from the thermophilic fungus *Thermoascus aurantiacus* (TAX) are particularly well suited for industrial applications at elevated temperatures. Previously, TAX had only been obtained from the endogenous expression in *T. aurantiacus* found in samples of Indian soil.[1] Although natural expression gave sufficient quantities for crystallization studies, cultivation of fungal cells was necessary, requiring between 3 and 13 days of incubation to produce the protein.[2]

Our interest in TAX focused on its potential use as a scaffold for computational enzyme design. Three features of TAX make it attractive for these purposes: (1) The protein is thermostable, which generally indicates a robust scaffold more amenable to multiple mutations than its mesophilic homologs.[3] (2) TAX has a large natural binding pocket that could accommodate large substrates and multiple catalytic residues. (3) A high-resolution crystal structure is available at 1.7 Å with the natural ligand bound in the active site.[4]

The engineering of any protein requires an expression system amenable to genetic manipulation, thus the natural fungal system is not appropriate for the expression of designed TAX variants. Here, we present an efficient, recombinant *E. coli* expression

and purification protocol that yields up to 150 mg of active enzyme per liter of culture and is readily transferable to a high-throughput format.

Materials and Methods

Optimization of the TAX-His$_6$ gene

The protein sequence for TAX was taken from the FASTA sequence of the crystal structure of TAX (PDB:1GOR).[4] A hexahistidine purification tag was added to the C-terminus of the protein sequence preceded by a Factor Xa cleavage site resulting in the C-terminal sequence: GSIEGRGHHHHHH. The gene for TAX-His$_6$ was designed from the protein sequence using DNA 2.0 Gene Designer and optimized for expression in *E. coli*.[5] A unique *Nde*I restriction site was incorporated at the 5′ end of the gene, and a stop codon (TAG) along with a unique *Bam*HI restriction site were incorporated immediately following the 3′ end of the gene. The gene and protein sequences for TAX-His$_6$ are shown in Figure D-1.

Design of the oligonucleotides

Forty-eight overlapping oligonucleotides for the construction of the TAX-His$_6$ gene were designed using Assembly PCR Oligo Maker.[6] The overlapping oligonucleotides spanned the length of the gene and were less than 40 base pairs long with 16 to 20 base-pair overlapping regions (Table D-1). All oligonucleotides were synthesized by Integrated DNA Technologies at a 20 mmol scale with no additional purification.

Gene construction

The full-length TAX-His_6 gene was constructed by recursive PCR based on the method of Stemmer *et al*.[7] The protocol for gene construction is outlined in Chapter III, Materials and Methods. After construction, the gene was amplified using the primers TAX_forward 5′-(TAAGAAGGAGATATACATATGG)-3′ and TAX_reverse 5′-(AACTCAGCTTCCTTTCGGG)-3′, which were calculated to have melting temperatures of 63.9 and 67.6°C, respectively. The purified gene fragment was digested using *Bam*HI/*Nde*I (New England Biolabs) and was then ligated into a similarly digested pET11a plasmid (Novagen) yielding the TAX-His6-pET11a plasmid. The plasmid was then transformed into *E. coli* XL-1 Blue cells (Stratagene) and the gene sequence was confirmed by DNA sequencing.

Overexpression of TAX-His_6

TAX-His_6-pET11a was transformed into *E. coli* BL-21(DE3) cells for expression. A single colony was then used to inoculate 50 mL of LB containing 100 μg/mL ampicillin, which was then grown overnight at 37°C with shaking. 30 mL of the pre-culture was used to innoculate 1 L LB/ampicillin. The culture was grown at 37°C with shaking to an OD_{600} of approximately 0.3. The temperature was then reduced to 25°C and expression was induced at an OD_{600} of 0.6 with the addition of 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). The culture was grown 18 hours and the cells were harvested by centrifugation.

Purification of TAX-His_6

The harvested cells were resuspended in lysis buffer (10 mM imidazole, 100 mM Tris pH 7.4, 300 mM NaCl), lysed mechanically with an Emulsiflex-C5 (Avestin), and then clarified by centrifugation at 10,000 × g for 40 min. The supernatant was then applied to 3 mL of Ni-NTA resin (Qiagen), which had been equilibrated with lysis buffer. The column was washed with 10 column volumes of lysis buffer and 10 column volumes of wash buffer (20 mM imidazole, 100 mM Tris pH 7.4, 300 mM NaCl). The protein was eluted with 3 column volumes of elution buffer (200 mM imidazole, 100 mM Tris pH 7.4, 300 mM NaCl). The eluate was dialyzed exhaustively against 50 mM sodium citrate, 50 mM NaCl, pH 5.5 and then concentrated using Amicon 10,000 MWCO centrifugal concentrators (Millipore).

Protein concentration determination

Protein concentrations were determined by UV absorbance after protein denaturation in 8 M guanidinium hydrochloride for 10 min with a dilution of at least 10×. An extinction coefficient at 280 nm of 55,280 $M^{-1}cm^{-1}$ was used.

Protein characterization

CD and MS were carried out as in Chapter III, Materials and Methods. CD spectra and thermal denaturation curves were obtained using 10 μM protein in 50 mM MES, pH 5.5, 100 mM NaCl.

Xylanase activity assays

The specific activity of TAX-His_6 was assayed using pNP-β-D-glucopyranoside as in Lo Leggio *et al*.[4] 2.5 mM substrate in 50 mM sodium citrate, pH 5.3 was incubated at 50°C for 5 min. The reaction was initiated with the addition of purified TAX-His_6 to a final volume of 300 μL. Protein concentrations between 170 μM and 1.7 μM were used. The reactions were mixed and incubated at 50°C for 15 min. The reaction was quenched with the addition of 400 uL of 20% Na_2CO_3 and the amount of liberated pNP was determined by measuring the absorbance at 410 nm. An extinction coefficient at 410 nm of 18,400 $M^{-1}cm^{-1}$ was used for pNP at pH 10.0.

Results and Discussion

We observed very high levels of TAX-His_6 expression in the BL-21 (DE3) *E. coli* cells (Figure D-2, lane 2) and were able to obtain 150 mg of highly pure protein per liter of culture after a single Ni-NTA affinity chromatography purification step (Figure D-2, lane 6). A majority of TAX-His_6 was expressed into the soluble fraction (Figure D-2, lane 3), but a significant amount of protein was found in the insoluble fraction, perhaps a result of the very high total amount of expressed protein. The mass was confirmed with electrospray mass spectrometry (Figure D-3), and is within 2 amu of the expected mass (34,433 kDa).

CD analysis shows that TAX-His_6 is folded and has a Tm of 75.2°C (Figure D-4). Unfortunately, no CD data is available for direct comparison with fungally expressed TAX. One source reports that TAX retains 40% of its activity at temperatures as high as

80°C for 1 hour. In the recombinant TAX-His_6, the protein is fully and irreversibly unfolded at this temperature. One subtle difference between TAX from fungal sources and from *E. coli* is the pyroglutamic acid modification of the N-terminal glutamate that is seen in the crystal structure of fungal TAX that is probably not present in the recombinant protein.[4] However, it is not clear that this small modification is enough to cause significant destabilization. The hexahistidine tag and Factor Xa cleavage site that were added to the protein sequence in TAX-His_6 may also contribute to the differences in stability, although we did not attempt to cleave the tag to determine if this was the cause of the destabilization. In addition, the strain of *T. aurantiacus* that was used to produce the xylanase in the crystal structure, which was the source of our protein sequence, is not explicitly stated. If the xylanase produced by the strain used in the temperature studies (C436) differed from the one that was used to produce the protein for the crystal structure, the protein sequences could be different, perhaps resulting in variations in thermal stability.

TAX-His_6 has a specific activity for pNP-β-D-glucopyranoside of 13.6 ± 3.9 U/mg. This is similar to the reported value from the natural expression system (12.7 U/mg), indicating that the enzyme is functionally unchanged when expressed in *E. coli*.[4]

Conclusions

We were able to express TAX-His_6 recombinantly, and because of our optimization of the gene for expression in *E. coli*, we obtained up to 150 mg of highly pure protein per liter of culture. The recombinantly expressed TAX-His_6 is folded, stable, and functionally indistinguishable from the protein obtained from the fungal

expression system. The high bacterial expression levels, ease of purification, and thermostability of TAX-His_6 make it an ideal scaffold for *de novo* computational enzyme design and other enzyme engineering studies. In addition, the method of constructing the TAX-His_6 gene through recursive PCR of overlapping 40-mer oligonucleotides facilitates future enzyme engineering projects on this scaffold. Multiple mutations can easily be introduced by simply substituting a subset of oligos that encode the mutations and carrying out the gene assembly and cloning steps to create the gene for the modified protein.

Acknowledgements

We would like to thank Leila Lo Leggio for discussion about the fungal TAX crystal structure and expression system.

References

1. Lo Leggio, L.; Kalogiannis, S.; Bhat, M. K.; Pickersgill, R. W., High resolution structure and sequence of T. aurantiacus xylanase I: implications for the evolution of thermostability in family 10 xylanases and enzymes with (β)α-barrel architecture. *Proteins* **1999**, *36*, 295-306.

2. Kalogiannis, S.; Owen, E.; Beever, D. E.; Bhat, M. K., Screening of ten strains of *Thermoascus aurantiacus* and characterization of a major xylanase. *Med. Fac. Landbouww. Univ. Gent.* **1995**, *60*, 1995-1998.

3. Besenmatter, W.; Kast, P.; Hilvert, D., Relative tolerance of mesostable and thermostable protein homologs to extensive mutation. *Proteins* **2007**, *66*, 500-506.

4. Lo Leggio, L.; Kalogiannis, S.; Eckert, K.; Teixeira, S. C.; Bhat, M. K.; Andrei, C.; Pickersgill, R. W.; Larsen, S., Substrate specificity and subsite mobility in *T. aurantiacus* xylanase 10A. *FEBS Lett* **2001**, *509*, 303-308.

5. Villalobos, A.; Ness, J.; Gustafsson, C.; Minshull, J.; Govindarajan, S., Gene Designer: a synthetic biology tool for constructing artificial DNA segments. *BMC Bioinformatics* **2006**, *7*, 285.

6. Rydzanicz, R.; Zhao, X. S.; Johnson, P. E., Assembly PCR oligo maker: a tool for designing oligodeoxynucleotides for constructing long DNA molecules for RNA production. *Nucleic Acids Res* **2005**, *33*, W521-525.

7. Stemmer, W. P.; Crameri, A.; Ha, K. D.; Brennan, T. M.; Heyneker, H. L., Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides. *Gene* **1995**, *164*, 49-53.

Table D-1. Assembly oligonucleotides for the construction of the TAX-His$_6$ gene. Overlapping regions are indicated with bold text and same-colored sections of adjacent oligonucleotides are complementary. All oligonucleotides are written from 5′ to 3′.

name	sequence	length
TAX-1t	TAAGAAGGAGATATACATATGGCAGAAGCGG	31
TAX-1b	TCAGTTGGTCAACGGATTGAGCCGCTTCTGCCATATGTAT	40
TAX-2t	CAATCCGTTGACCAACTGATTAAAGCTCGTGGTAAAGTGT	40
TAX-2b	GTCGGTTGCAACACCGAAATACACTTTACCACGAGCTTTA	40
TAX-3t	TTCGGTGTTGCAACCGACCAGAACCGCCTGACCACTGGC	39
TAX-3b	CTGAATGATCGCCGCGTTCTTGCCAGTGGTCAGGCGGT	38
TAX-4t	AACGCGGCGATCATTCAGGCAGATTTCGGTCAGGTTACCC	40
TAX-4b	CCCATTTCATGGAGTTCTCTGGGGTAACCTGACCGAAATC	40
TAX-5t	GAGAACTCCATGAAATGGGACGCGACCGAGCCTTCTCAA	39
TAX-5b	CGCCAGCGAAGTTGAAGTTGCCTTGAGAAGGCTCGGTCGC	40
TAX-6t	AACTTCAACTTCGCTGGCGCAGACTACCTGGTGAACTGG	39
TAX-6b	CAGCTTACCGTTCTGCTGCGCCCAGTTCACCAGGTAGTC	39
TAX-7t	CAGCAGAACGGTAAGCTGATCCGCGGTCATACGCTGGTT	39
TAX-7b	AGACGGCAGCTGGCTGTGCCAAACCAGCGTATGACCGCG	39
TAX-8t	ACAGCCAGCTGCCGTCTTGGGTGTCTTCCATCACCGATA	39
TAX-8b	TCACGTTGGTCAGAGTGTTTTTATCGGTGATGGAAGACAC	40
TAX-9t	AACACTCTGACCAACGTGATGAAGAACCATATCACTACCC	40
TAX-9b	TGCCTTTATAACGGGTCATCAGGGTAGTGATATGGTTCTTC	41
TAX-10t	GATGACCCGTTATAAAGGCAAAATCCGCGCGTGGGATGTT	40
TAX-10b	CCTCGTTGAATGCTTCATTCACAACATCCCACGCGCGGA	39
TAX-11t	GAATGAAGCATTCAACGAGGACGGCAGCCTGCGCCAGAC	39
TAX-11b	CACCGATCACATTCAGAAAAACGGTCTGGCGCAGGCTGC	39
TAX-12t	TTTTCTGAATGTGATCGGTGAAGATTACATCCCGATCGCAT	41
TAX-12b	GCAGCACGGGCGGTCTGGAATGCGATCGGGATGTAATC	38
TAX-13t	AGACCGCCCGTGCTGCAGATCCAAACGCTAAGCTGTACA	39
TAX-13b	AGTCCAGGTTGTAATCGTTAATGTACAGCTTAGCGTTTGG	40
TAX-14t	TAACGATTACAACCTGGACTCTGCGTCTTATCCGAAAACCC	41
TAX-14b	AACACGGTTCACGATGGCCTGGGTTTTCGGATAAGACGC	39
TAX-15t	GCCATCGTGAACCGTGTTAAACAGTGGCGTGCGGCTGG	38
TAX-15b	GCCGATGCCGTCGATCGGAACGCCAGCCGCACGCCACT	38
TAX-16t	CGATCGACGGCATCGGCTCCCAGACGCATCTGTCTGCA	38
TAX-16b	AGCACGCCAGCGCCCTGGCCTGCAGACAGATGCGTCTG	38
TAX-17t	AGGGCGCTGGCGTGCTGCAGGCCCTGCCGCTGCTGGC	37
TAX-17b	ACCTCCGGAGTGCCGGCGCTTGCCAGCAGCGGCAGGG	37
TAX-18t	CCGGCACTCCGGAGGTTGCAATCACCGAGCTGGATGTAG	39
TAX-18b	ATCAGTCGGGCTTGCGCCCGCTACATCCAGCTCGGTGAT	39
TAX-19t	GCGCAAGCCCGACTGATTATGTCAACGTCGTGAACGCG	38
TAX-19b	CACAAGACTGCACATTCAGGCACGCGTTCACGACGTTGAC	40
TAX-20t	CTGAATGTGCAGTCTTGTGTGGGCATTACCGTATGGGGT	39
TAX-20b	CCAAGAATCAGGATCGGCAACACCCCATACGGTAATGCC	39
TAX-21t	GCCGATCCTGATTCTTGGCGCGCATCCACTACCCCGCTG	39
TAX-21b	GGTTGAAATTACCGTCGAACAGCAGCGGGGTAGTGGATG	39
TAX-22t	GTTCGACGGTAATTTCAACCCGAAACCAGCTTACAACGCTA	41
TAX-22b	CTGTTGCAGGTCCTGAACGATAGCGTTGTAAGCTGGTTTC	40
TAX-23t	GTTCAGGACCTGCAACAGGGCAGCATCGAGGGTCGTGGT	39
TAX-23b	CCTAATGGTGGTGGTGATGGTGACCACGACCCTCGATGC	39
TAX-24t	CATCACCACCACCATTAGGGATCCGGCTGCTAACAAAGC	39
TAX-24b	AACTCAGCTTCCTTTCGGGCTTTGTTAGCAGCCGGA	36

T. aruantiacus xylanase (TAX-His$_6$)

```
    TAA GAA GGA GAT ATA CAT ATG GCA
                        NdeI    A

  1 GAA GCG GCT CAA TCC GTT GAC CAA CTG ATT AAA GCT CGT GGT AAA GTG TAT TTC GGT GTT
  1  E   A   A   Q   S   V   D   Q   L   I   K   A   R   G   K   V   Y   F   G   V

 61 GCA ACC GAC CAG AAC CGC CTG ACC ACT GGC AAG AAC GCG GCG ATC ATT CAG GCA GAT TTC
 21  A   T   D   Q   N   R   L   T   T   G   K   N   A   A   I   I   Q   A   D   F

121 GGT CAG GTT ACC CCA GAG AAC TCC ATG AAA TGG GAC GCG ACC GAG CCT TCT CAA GGC AAC
 41  G   Q   V   T   P   E   N   S   M   K   W   D   A   T   E   P   S   Q   G   N

181 TTC AAC TTC GCT GGC GCA GAC TAC CTG GTG AAC TGG GCG CAG CAG AAC GGT AAG CTG ATC
 61  F   N   F   A   G   A   D   Y   L   V   N   W   A   Q   Q   N   G   K   L   I

241 CGC GGT CAT ACG CTG GTT TGG CAC AGC CAG CTG CCG TCT TGG GTG TCT TCC ATC ACC GAT
 81  R   G   H   T   L   V   W   H   S   Q   L   P   S   W   V   S   S   I   T   D

301 AAA AAC ACT CTG ACC AAC GTG ATG AAG AAC CAT ATC ACT ACC CTG ATG ACC CGT TAT AAA
101  K   N   T   L   T   N   V   M   K   N   H   I   T   T   L   M   T   R   Y   K

361 GGC AAA ATC CGC GCG TGG GAT GTT GTG AAT GAA GCA TTC AAC GAG GAC GGC AGC CTG CGC
121  G   K   I   R   A   W   D   V   V   N   E   A   F   N   E   D   G   S   L   R

421 CAG ACC GTT TTT CTG AAT GTG ATC GGT GAA GAT TAC ATC CCG ATC GCA TTC CAG ACC GCC
141  Q   T   V   F   L   N   V   I   G   E   D   Y   I   P   I   A   F   Q   T   A

481 CGT GCT GCA GAT CCA AAC GCT AAG CTG TAC ATT AAC GAT TAC AAC CTG GAC TCT GCG TCT
161  R   A   A   D   P   N   A   K   L   Y   I   N   D   Y   N   L   D   S   A   S

541 TAT CCG AAA ACC CAG GCC ATC GTG AAC CGT GTT AAA CAG TGG CGT GCG GCT GGC GTT CCG
181  Y   P   K   T   Q   A   I   V   N   R   V   K   Q   W   R   A   A   G   V   P

601 ATC GAC GGC ATC GGC TCC CAG ACG CAT CTG TCT GCA GGC CAG GGC GCT GGC GTG CTG CAG
201  I   D   G   I   G   S   Q   T   H   L   S   A   G   Q   G   A   G   V   L   Q

661 GCC CTG CCG CTG CTG GCA AGC GCC GGC ACT CCG GAG GTT GCA ATC ACC GAG CTG GAT GTA
221  A   L   P   L   L   A   S   A   G   T   P   E   V   A   I   T   E   L   D   V

721 GCG GGC GCA AGC CCG ACT GAT TAT GTC AAC GTC GTG AAC GCG TGC CTG AAT GTG CAG TCT
241  A   G   A   S   P   T   D   Y   V   N   V   V   N   A   C   L   N   V   Q   S

781 TGT GTG GGC ATT ACC GTA TGG GGT GTT GCC GAT CCT GAT TCT TGG CGC GCA TCC ACT ACC
261  C   V   G   I   T   V   W   G   V   A   D   P   D   S   W   R   A   S   T   T

841 CCG CTG CTG TTC GAC GGT AAT TTC AAC CCG AAA CCA GCT TAC AAC GCT ATC GTT CAG GAC
281  P   L   L   F   D   G   N   F   N   P   K   P   A   Y   N   A   I   V   Q   D

901 CTG CAA CAG GGC AGC ATC GAG GGT CGT GGT CAC CAT CAC CAC CAC CAC CAT TAG
301  L   Q   Q   G   S   I   E   G   R   G   H   H   H   H   H   H

    GGA TCC GGC TGC TAA CAA AGC CCG AAA GGA AGC TGA GTT
     BamHI
```

Figure D-1. Protein and DNA sequences for TAX-His$_6$. The hexahistidine tag is shown in pink, the Factor Xa cleavage site is in red, and the restriction endonuclease cleavage sites are in green.

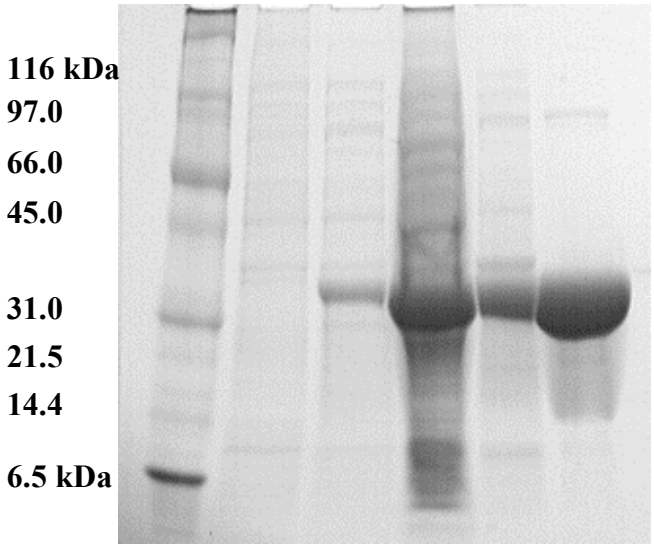

Figure D-2. SDS-PAGE analysis of TAX-His$_6$ expression and purification. Approximate molecular weights are indicated. Lane 1: Molecular weight marker. Lane 2: uninduced cells. Lane 3: cells after 18 hours of induction. Lane 4: lysate supernatant. Lane 5: lysate pellet. Lane 6: Ni-NTA affinity column eluate.

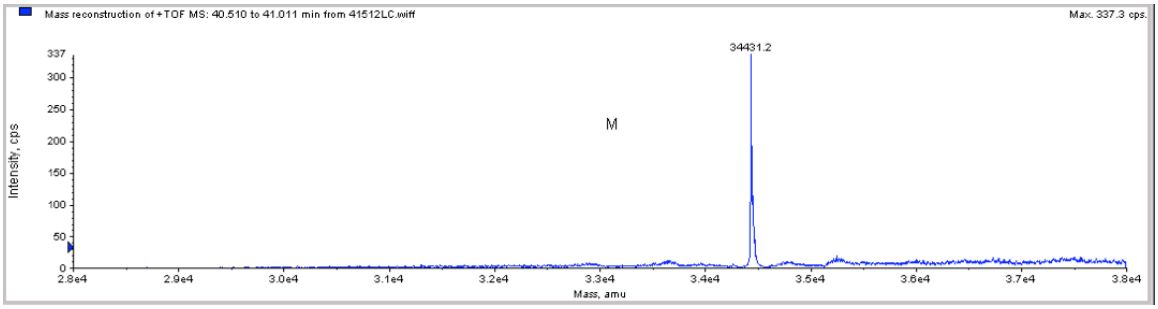

Figure D-3. Mass spectrometry analysis of TAX-His$_6$. The actual mass (34431.2) is within 2 amu of the expected mass (34433 amu).

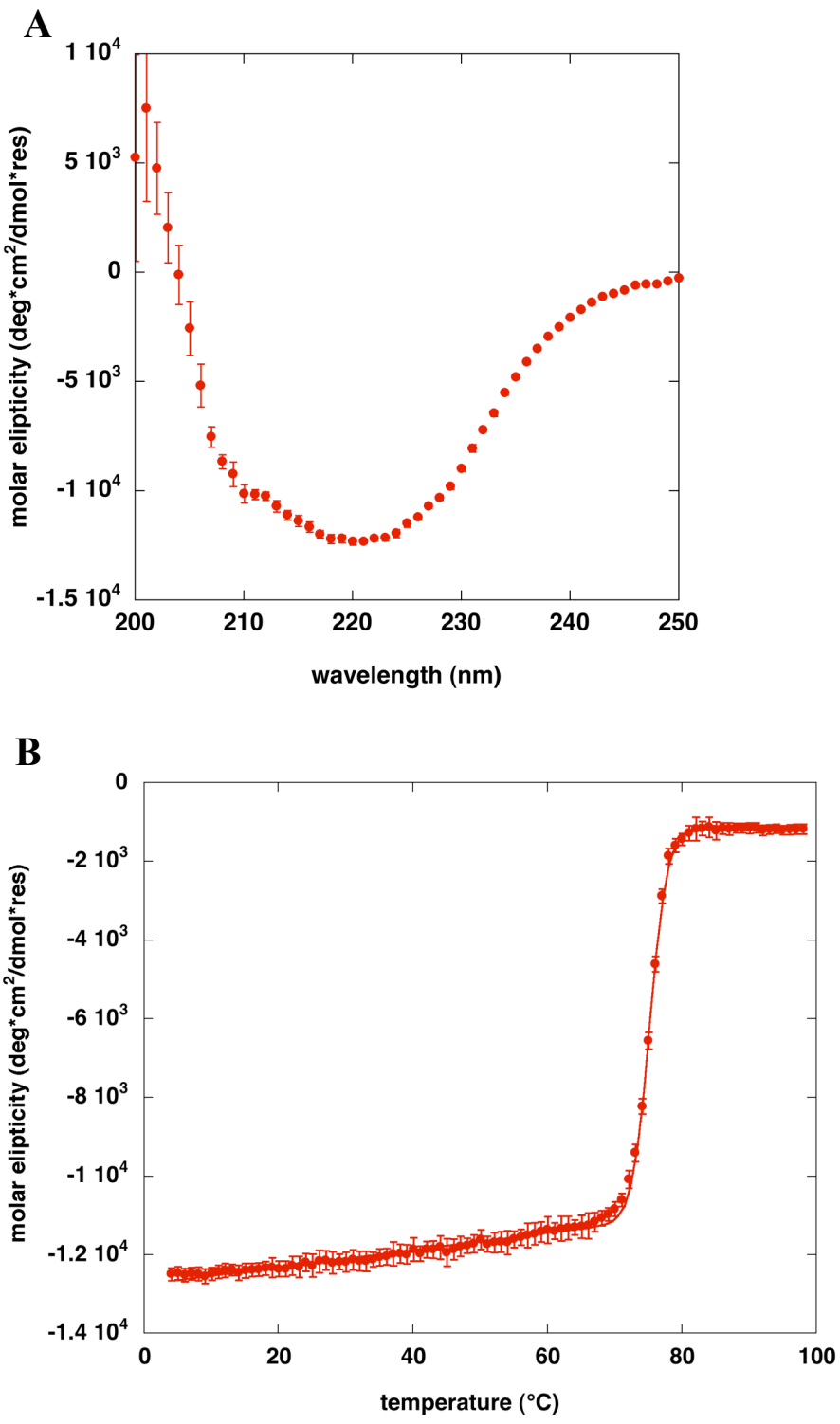

Figure D-4. CD analysis of TAX-His$_6$. **(A)** Far-UV wavelength scan, 25°C. **(B)** Thermal denaturation monitored at 222 nm. All experiments were carried out with 10 μM protein in 50 mM MES, pH 5.5, 50 mM NaCl.