



Crystal Structure of Human Protein N-Terminal Glutamine Amidohydrolase, an Initial Component of the N-End Rule Pathway

Mi Seul Park^{1,9}, Eduard Bitto^{2,9}, Kyung Rok Kim¹, Craig A. Bingman³, Mitchell D. Miller⁴, Hyun-Jung Kim⁵, Byung Woo Han^{1*}, George N. Phillips Jr^{3,4*}

1 Research Institute of Pharmaceutical Sciences, College of Pharmacy, Seoul National University, Seoul, Korea, **2** Department of Chemistry and Biochemistry, Georgian Court University, Lakewood, New Jersey, United States of America, **3** Department of Biochemistry, Center for Eukaryotic Structural Genomics, University of Wisconsin Madison, Madison, Wisconsin, United States of America, **4** Department of Biochemistry and Cell Biology, Rice University, Houston, Texas, United States of America, **5** College of Pharmacy, Chung-Ang University, Seoul, Korea

Abstract

The N-end rule states that half-life of protein is determined by their N-terminal amino acid residue. N-terminal glutamine amidohydrolase (Ntaq) converts N-terminal glutamine to glutamate by eliminating the amine group and plays an essential role in the N-end rule pathway for protein degradation. Here, we report the crystal structure of human Ntaq1 bound with the N-terminus of a symmetry-related Ntaq1 molecule at 1.5 Å resolution. The structure reveals a monomeric globular protein with alpha-beta-alpha three-layer sandwich architecture. The catalytic triad located in the active site, Cys-His-Asp, is highly conserved among Ntaq family and transglutaminases from diverse organisms. The N-terminus of a symmetry-related Ntaq1 molecule bound in the substrate binding cleft and the active site suggest possible substrate binding mode of hNtaq1. Based on our crystal structure of hNtaq1 and docking study with all the tripeptides with N-terminal glutamine, we propose how the peptide backbone recognition patch of hNtaq1 forms nonspecific interactions with N-terminal peptides of substrate proteins. Upon binding of a substrate with N-terminal glutamine, active site catalytic triad mediates the deamination of the N-terminal residue to glutamate by a mechanism analogous to that of cysteine proteases.

Citation: Park MS, Bitto E, Kim KR, Bingman CA, Miller MD, et al. (2014) Crystal Structure of Human Protein N-Terminal Glutamine Amidohydrolase, an Initial Component of the N-End Rule Pathway. *PLoS ONE* 9(10): e111142. doi:10.1371/journal.pone.0111142

Editor: Titus J. Boggon, Yale University School of Medicine, United States of America

Received: June 14, 2014; **Accepted:** September 28, 2014; **Published:** October 30, 2014

Copyright: © 2014 Park et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability: The authors confirm that all data underlying the findings are fully available without restriction. Atomic coordinates and structure factors are available from the Protein Data Bank database, accession number 4W79 (<http://www.pdb.org/pdb/explore/explore.do?structureId=4W79>).

Funding: This study was supported by GM064598, National Institutes of Health Protein Structure Initiative, (<http://www.nigms.nih.gov/research/specificareas/PSI/pages/default.aspx>), GNP GM074901, National Institutes of Health Protein Structure Initiative, (<http://www.nigms.nih.gov/research/specificareas/PSI/pages/default.aspx>), GNP GM094816, National Institutes of Health Protein Structure Initiative, (<http://www.nigms.nih.gov/research/specificareas/PSI/pages/default.aspx>), GNP1231306, The BioXFEL Science and Technology Center under National Science Foundation, (<https://www.bioxfel.org/>), MDM2013-043695, Ministry of Science, ICT and Future Planning of Korea, (www.msip.go.kr), BWH 2014-001848, Ministry of Science, ICT and Future Planning of Korea, (www.msip.go.kr), BWH A092006, Ministry of Health and Welfare of Korea, (www.mw.go.kr), BWH. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* Email: bwhan@snu.ac.kr (BWH); georgep@rice.edu (GNP)

⁹ These authors contributed equally to this work.

Introduction

Aberrant polypeptides or proteins should be accurately removed in many physiological processes. Intracellular protein degradation is mainly conducted through the ubiquitin-proteasome pathway (UPP) or the lysosomal proteolysis [1]. The UPP is required for degradation of short-lived proteins in eukaryotic cells. In the UPP, ubiquitin first attaches to target proteins or polypeptides, which leads to their recognition by the 26S proteasome [2]. On the other hand, lysosomal proteolysis leads to breakdown of unnecessary proteins or polypeptides by lysosomes [3].

The N-end rule is related to the ubiquitin-dependent proteolytic system [4]. The N-end rule is one of common pathways for the degradation of polypeptides and proteins in prokaryotes and eukaryotes, which determines the stability of a protein by its N-terminal residue [1]. Val, Gly, and Pro are classified as stabilizing

residues in mammals whereas Asp, Gln, Cys, and Arg are known as destabilizing residues [5]. In the N-end rule pathway, N-terminal glutamine and asparagine are tertiary destabilizing residues and these residues are converted into secondary destabilizing N-terminal glutamate and aspartate by deamidation [6,7]. Arginine is then conjugated to glutamate and aspartate residues by Arg-tRNA-protein transferase, converting target proteins into ones possessing a primary destabilizing residues [6–8]. The N-end rule mechanism is involved in degradation of misfolded proteins, regulation of DNA repair, apoptosis and meiosis [9–12]. The N-terminal amidohydrolases are classified into the N-terminal glutamine amidohydrolase (Ntaq) and the N-terminal asparagine amidohydrolase (Ntan), which share low amino acid sequence identity and mediate the deamidation of N-terminal glutamine and asparagine, respectively [6].

In this work, we present the crystal structure of hNtaq1 bound with the N-terminus of a symmetry-related Ntaq1 molecule at 1.5 Å resolution. The structure contains the catalytic triad (Cys-His-Asp) in the active site, which is well conserved among Ntaq proteins and transglutaminases from diverse organisms. Additionally, we conducted docking study with all the tripeptides containing N-terminal glutamine to elucidate how N-termini of proteins with N-terminal glutamine are recognized and positioned by hNtaq1 into catalytically conducive conformations. We also propose a catalytic mechanism of hNtaq1 based on the crystal structure of hNtaq1 and docking study.

Materials and Methods

Cloning, protein expression, and purification

The standard Center for Eukaryotic Structural Genomics pipeline protocols were used for cloning [13], protein expression [14], protein purification [15]. In summary, using Gateway cloning (Life Technologies, USA), hNtaq1 gene was cloned into pVP16 plasmid (Clontech, USA) containing N-terminal fusion (His)₆-Maltose Binding Protein (MBP) and a linker region with the TEV protease site for cleavage of target proteins. It results in the hNtaq1 construct mutated with Ser for the initial Met. The hNtaq1 construct was transformed into B834 *E. coli* cells (Novagen, USA) to express Se-Met labeled protein. Cells were cultured with auto-induction medium adapted from the work of Studier [16] and incubated in a shaker at 250 rpm, 25°C for 22–24 hours. Cells were harvested by centrifugation at 5000×g for 20 min and suspended in cell lysis buffer (20 mM sodium phosphate, pH 7.5, 500 mM sodium chloride, 20% ethylene glycol, 35 mM imidazole, 0.3 mM tris(2-carboxyethyl) phosphate (TCEP), and E64 protease inhibitor cocktail (Sigma Aldrich, USA). Cells were disrupted by sonication on ice and the cell lysate was centrifuged at 75,600×g for 30 min twice. The supernatant was collected and filtered through a 0.8 μm pore size filter.

The sample was loaded on the Ni-charged HiTrap chelating 5 ml HP column (GE Healthcare, UK) and washed with IMAC-washing buffer (20 mM sodium phosphate, pH 7.5, 500 mM sodium chloride, and 0.3 mM TCEP). The protein was eluted by applying a 30 column volume linear gradient from 10% to 80% IMAC-elution buffer (20 mM sodium phosphate, pH 7.5, 350 mM imidazole, 500 mM sodium chloride, and 0.3 mM TCEP) and buffer was exchanged to TEV proteolysis buffer (20 mM sodium phosphate, pH 7.5, 100 mM sodium chloride, and 0.3 mM TCEP) using a HiPrep 26/10 desalting column (GE Healthcare, UK). The (His)₆-MBP fusion hNtaq1 protein was treated with TEV protease (1:100 w/w) at 25°C for overnight. After the TEV protease treatment, the protein was loaded to HiTrap chelating 5 ml HP column (GE Healthcare, UK) and eluted with IMAC-washing buffer. The eluted sample was desalted with crystallization buffer (50 mM sodium chloride, 3 mM sodium azide, 0.3 mM TCEP, and 100 mM Bis-Tris, pH 7.0) using a HiPrep 26/10 desalting column (GE Healthcare, UK). For crystallization, the purified Se-Met hNtaq1 protein was concentrated to 10 mg/ml.

Crystallization, X-ray data collection, structure determination, and model evaluation

Crystals of the hNtaq1 were obtained by hanging-drop vapor diffusion method at 291 K by mixing the protein solution (10 mg/ml Se-Met protein, 50 mM sodium chloride, 3 mM sodium azide, 0.3 mM TCEP, and 100 mM Bis-Tris, pH 7.0) and the well solution (1% ethylene glycol, 1.8 M ammonium sulfate, 100 mM MES, pH 6.0) in 1:1 ratio. Crystals were cryoprotected in four

stages with well solution using containing 0 to 25% ethylene glycol and were flash-frozen in liquid nitrogen gas at 100 K. X-ray diffraction data were collected using synchrotron beam line 23-ID-D at the Advanced Photon Source of the Argonne National Laboratory. The crystal structure of the hNtaq1 was solved by SAD phasing at 1.5 Å resolution. SHARP [17] was used to solve experimental phase information, which was improved by density modification using DM [18]. Crystals of hNtaq1 belong to the space group $P2_12_12_1$ with unit cell parameters $a = 34.3$ Å, $b = 64.0$ Å, and $c = 113.6$ Å. Subsequent manual model building and refinement were carried out using *Coot* [19] and *REFMAC* [20] from CCP4 program suite [20]. All refinement steps were monitored using an R_{free} value based on 5.0% of the independent reflections. The stereochemical quality of the final model was assessed using PROCHECK [21] and *MolProbity* [22]. The data collection, phasing, and refinement statistics are summarized in Table 1.

Docking studies of tripeptides with N-terminal glutamine

AutoDock Vina program [23] was used for the docking studies of hNtaq1 with all the possible 400 tripeptides containing N-terminal glutamine (Gln-X-X, X is any amino acid residue, thus 20×20 candidates). Coordinates for the 400 tripeptides were generated using *Coot* [19] and converted to *pdbqt* files using *AutoDockTools4* [24]. The grid maps for docking studies were centered on C_α of Ser1 of the bound N-terminus from the symmetry-related hNtaq1 molecule in the substrate binding cleft and the maps comprised 30×30×30 points. AutoDock Vina program was run with four-way multithreading and the default settings were used for others computational parameters. Figures are generated using PyMol [25].

Data deposition

Atomic coordinates and structure factors have been deposited in the RCSB Protein Data Bank, accession code 4W79.

Results and Discussion

Overall structure of hNtaq1

The human C8orf32 gene encodes human N-terminal glutamine amidohydrolase isoform 1 (hNtaq1). Ntaq is an initial component of the N-end rule pathway and converts N-terminal glutamine to glutamate. In order to understand the relationship between the structure and function of hNtaq1, we determined the crystal structure of recombinant hNtaq1 bound with the adjacent N-terminus of a symmetry-related hNtaq1 molecule at 1.5 Å resolution. The hNtaq1 protein contains 205 amino acids, of which 202 have been successfully modeled in the presented hNtaq1 structure. Three C-terminal residues (Lys203, Asn204, and Cys205) were disordered in the crystal and could not be modeled. The R and R_{free} values of the final refined model were 14.4% and 17.0%, respectively. hNtaq1 is a monomeric globular protein with a novel structural fold of alpha-beta-alpha three-layer sandwich architecture (Figure 1A). To our surprise, the N-terminus of a symmetry-related hNtaq1 molecule was captured in the substrate binding cleft, even though Ser1 is the N-terminal residue, not an anticipated glutamine residue (Figure 1B). The core region of the protein shows antiparallel beta-sheets surrounded by helices. The catalytic triad (Cys28, His81, and Asp97) is highly conserved among Ntaq proteins, transglutaminases, and cysteine proteases of diverse organisms (Figure 1C). Ntaq and Ntan in human share only 13.1% sequence identity and the structure of Ntan has yet not been determined. Elucidation of subtle differences in the catalytic mechanism between Ntaq and

Table 1. Statistics for data collection, phasing, and model refinement.

Data collection and phasing ^a	
Space group	$P 2_1 2_1 2_1$
Cell dimensions	
a, b, c (Å), α , β , γ (°)	34.32, 64.04, 113.66, 90, 90, 90
Data set	Se λ 1 (peak)
X-ray wavelength (Å)	0.9794
Resolution range (Å) ^b	32.86–1.50 (1.53–1.50)
$\langle I/\sigma(I) \rangle$	12.1 (2.5)
Multiplicity	12.2 (6.9)
Unique reflections	40,943 (2,566)
Completeness (%)	99.5 (95.9)
R_{merge} (%) ^c	0.5 (54.3)
Figure of merit ^d for SAD phasing: 0.44	
Refinement	
R_{work}^e/R_{free}^f	0.144/0.170
No. of protein atoms	1,666
No. of water atoms	332
No. of Non-water atoms	75
Mean B value (Å ²)	18.9
Ramachandran plot analysis (for Chain A)	
Most favored regions	198 (96.6%)
Additional allowed regions	7 (3.4%)
Disallowed regions	0 (0%)
R.m.s. deviations from ideal geometry	
Bond lengths (Å)	0.017
Bond angles (°)	1.91

^aData collected at the Sector 23-ID-D of the Advanced Photon Source.

^bNumbers in parentheses indicate the highest resolution shell of 20.

^c $R_{merge} = \sum_i \sum_h |I(h)_i - \langle I(h) \rangle| / \sum_i \sum_h I(h)_i$, where $I(h)$ is the observed intensity of reflection h , and $\langle I(h) \rangle$ is the average intensity obtained from multiple measurements.

^dFigure of merit = $\langle |\sum P(\alpha)e^{i\alpha} / \sum P(\alpha)| \rangle$, where α is the phase angle and $P(\alpha)$ is the phase probability distribution.

^e $R_{work} = \sum | |F_o| - |F_c| | / \sum |F_o|$, where $|F_o|$ is the observed structure factor amplitude and $|F_c|$ is the calculated structure factor amplitude.

^f R_{free} = R-factor based on 5.0% of the data excluded from refinement.

doi:10.1371/journal.pone.0111142.t001

Ntan in the N-end rule pathway awaits the structural information of Ntan.

To find out structural features of hNtaq1, we analyzed of structure similarity using the DALI server [26]. The result reveals that hNtaq1 shows low level of sequence conservation with other structurally similar proteins and the amino acid sequence identities are in the range of 9 to 15%. The result shows that protein-glutaminase from *Chryseobacterium proteolyticum* (PDB ID: 3A54) is the closest structural homolog of hNtaq1 (Z score 7.0 and RMSD distance 3.6 Å for 123 equivalent C α positions out of 280 residues) and shares only 15% sequence identity. The protein-glutaminase from *Chryseobacterium proteolyticum* converts a glutamine residue to a glutamate and has catalytic triad comprising Cys156, His197, and Asp217 [27] (Figure 1D). The secreted effector protein SseI from *Sallionella typhimurium* (PDB ID: 4G2B) is the second best structural homolog of hNtaq1 (Z score 6.3 and RMSD distance 3.7 Å for 116 equivalent C α positions out of 169 residues) and shares only 9% sequence identity. The secreted effector protein SseI from *Sallionella typhimurium* belongs to cysteine protease superfamily and also contains catalytic triad with Cys178, His216, and Asp 231 [28] (Figure 1E). LapG from *Legionella pneumophila* (PDB ID: 4FGP),

a bacterial transglutaminase-like cysteine protease (BTLCP) containing catalytic triad, also shares structural similarity to hNtaq1 (Z score 6.1 and RMSD distance 2.7 for 96 equivalent C α positions out of 186 residues) [29] (Figure 1F). Even though structural homologs of hNtaq1 do not share high sequence identities, the catalytic triad is structurally well conserved, mediating similar type of reactions with different substrates. Uniqueness of substrate binding region except the catalytic triad could contribute to the specificity with respect to each substrates.

Active site with catalytic triad of hNtaq1

The structure of hNtaq1 shows the precise conformation of the active site and the catalytic triad, Cys28, His81, and Asp97. Previous researches showed that Cys and His residues play a vital role in the activity of Ntaq. In the case of mouse Ntaq Cys30Ala and His83Ala mutants, the mutations almost abolish activities as an N-terminal glutamine amidohydrolase [8]. In the structure of hNtaq1, Asp97 forms hydrogen bonds with His81, Tyr111 and a water molecule 1 (water403 in PDB) and His81 interacts with another water molecule 2 (water444 in PDB) via hydrogen bond. The distances from the sulfhydryl group of Cys28 to water2 and His81 are 3.4 Å and 3.8 Å, respectively. After proper conformational

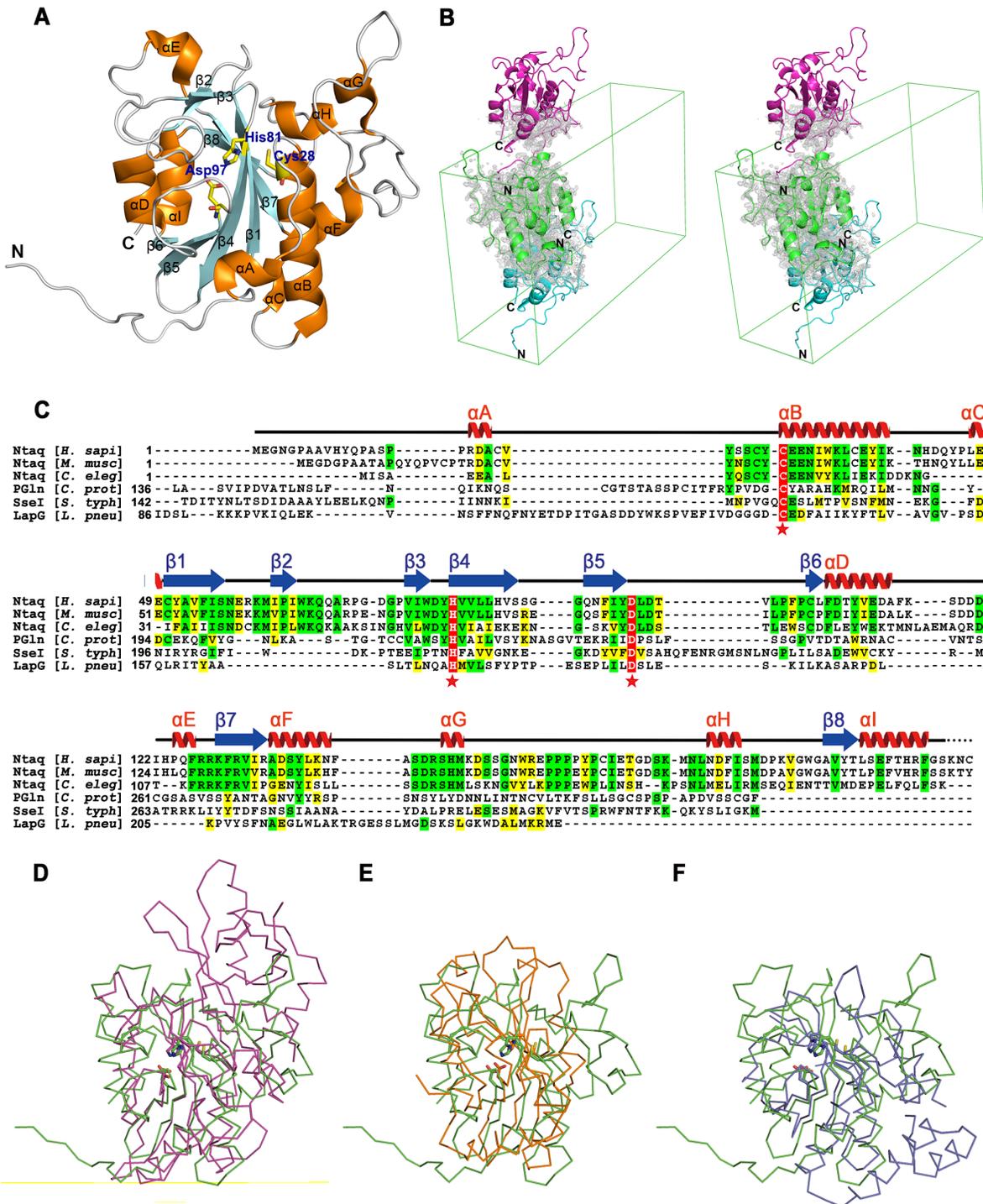


Figure 1. Overview of the crystal structure of hNtaq1. (A) Overall structure of hNtaq1. α -helices, β -strands, and loops are colored in orange, cyan, and white, respectively. The representative amino acid residues in the active site are shown as stick model (carbon, oxygen, nitrogen, and sulfur in yellow, red, blue, and gold color, respectively). (B) Stereo view of the crystallographic contact of hNtaq1 with a symmetry-related molecule. hNtaq1 and symmetry-related molecules are represented as green, cyan, and magenta cartoon, respectively. Unit cell is shown with green line and electron density map is shown as gray cloud. (C) Sequence alignment of hNtaq1 with Ntaq proteins from *Mus musculus*, *Caenorhabditis elegans*, protein-glutaminase from *Chryseobacterium proteolyticum*, secreted effector protein Ssel from *Sallonella typhimurium*, and periplasmic protease LapG from *Legionella pneumophila*. Residues of the catalytic triad are represented by red asterisk below the amino acid sequences. Completely conserved, identical, moderately conserved residues are highlighted with red, green, and yellow shaded boxes, respectively. The secondary structure of hNtaq1 is shown on top of the sequence alignment. α -helix, β -sheet, and connecting region are represented by red spiral, blue arrow, and black line, respectively. Structural comparison of hNtaq1 with protein-glutaminase from *C. proteolyticum* (D), secreted effector protein Ssel from *S. typhimurium* (E), and periplasmic protease LapG from *L. pneumophila* (F). The catalytic triads are shown as stick models and main chains of hNtaq1, protein-glutaminase, Ssel, and LapG are represented with ribbon diagram in green, magenta, orange, and blue, respectively.

doi:10.1371/journal.pone.0111142.g001

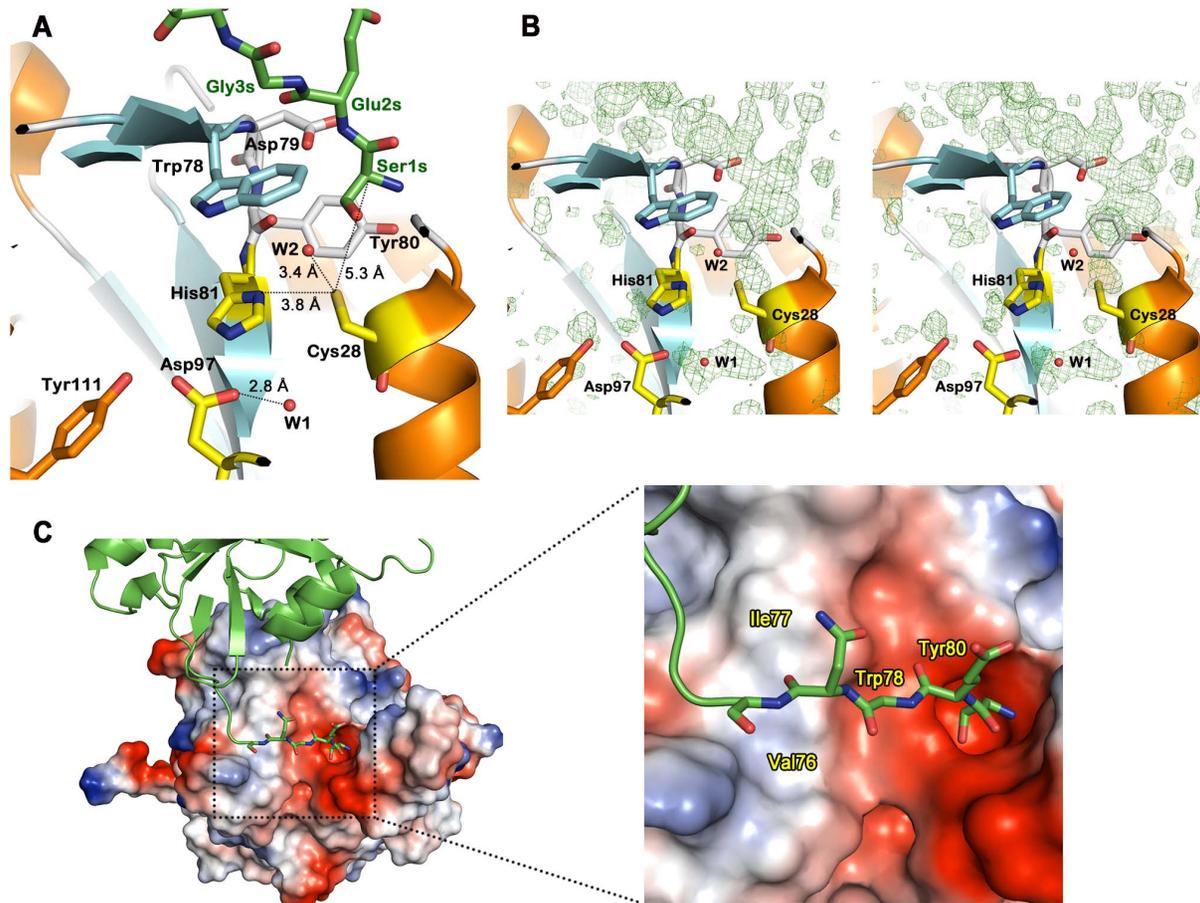


Figure 2. Active site and electrostatic potential surface charge of hNtaq1. (A) Substrate binding cleft of hNtaq1. Carbon in the substrate-mimicking peptide, catalytic triad, α -helices, β -strands, and loops are colored in green, yellow, orange, cyan, and white, respectively. Oxygen, nitrogen, and sulfur atoms are represented as red, blue, and gold, respectively. Two water molecules are shown as red sphere and labeled as W1 and W2. (B) Electron density map from an $F_o - F_c$ omit map calculated without the bound substrate-mimicking peptide. Positive electron density are shown as a green mesh contoured at 2.0σ , in a stereo view. (C) Electrostatic potential surface and substrate binding cleft region of hNtaq1. Negatively and positively charged surfaces are represented as red and blue shade, respectively. Residues interacting with the substrate-mimicking peptide molecule are labeled.
doi:10.1371/journal.pone.0111142.g002

changes that shorten the distance between His81 and Cys28 during a catalytic conversion, the sulfhydryl group of Cys28 is deprived of proton and increases its nucleophilicity for a successive attack on substrates (Figure 2A).

Interestingly, the crystal structure of hNtaq1 revealed an unusual crystallographic contact between symmetry-related molecules. The N-terminus of a symmetry-related monomer is anchored into the active site of hNtaq1 with the side chain of the N-terminal residue (Ser1) in close proximity of the catalytic Cys28. The electron density of the bound N-terminus of a symmetry-related monomer could be clearly observed in an $F_o - F_c$ omit map when calculated without the N-terminus residues as shown in Figure 2B. This phenomenon was similarly observed in the case of crystal structure of another N-end rule pathway protein, UBR box of ubiquitine ligases (PDB ID: 3NIS) [30]. The N-terminus of the symmetry-related monomer (from now on, the “substrate-mimicking” peptide) forms an antiparallel beta-strand segment that adheres to the surface exposed beta-strand of hNtaq1 (residues 76–79, β_3 in Figure 1A). The strands are stabilized by several interactions: 1) three peptide backbone hydrogen bonds typical for beta-sheet structures (two between Ile77 and Asn4s, and one between Asp79 and Glu2s); 2) a hydrogen bond between side

chain of Asp79 and peptide backbone of the N-terminal residue Ser1s; 3) a hydrogen bond between Tyr80 and the amino group of the substrate-mimicking peptide; 4) a direct van der Waals contact of Trp78 with the substrate-mimicking peptide. The surface charge of the binding cleft is predominantly negative, and several hydrophobic and aromatic residues (Val76, Ile77, Trp78, and Tyr80) are well oriented to recognize main chain backbone of the substrate-mimicking peptide and to stabilize aliphatic part of the bound substrate-mimicking peptide (Figure 2C). Overall, the active site bound with the substrate-mimicking peptide directly suggests possible interaction mode of hNtaq1 with substrates.

Docking study with anticipated substrate peptides

In order to further elucidate the interaction between hNtaq1 and anticipated substrate peptides with N-terminal glutamine, we conducted docking experiments with all the possible 400 tripeptides with N-terminal glutamine (Gln-X-X), using AutoDock Vina program [23]. Around the binding cleft, there exist three regions that are predominantly negatively-charged, positively-charged, and nonpolar. The docking study was performed to see how the three regions contributes to the recognition of its anticipated substrate. Control docking experiments performed

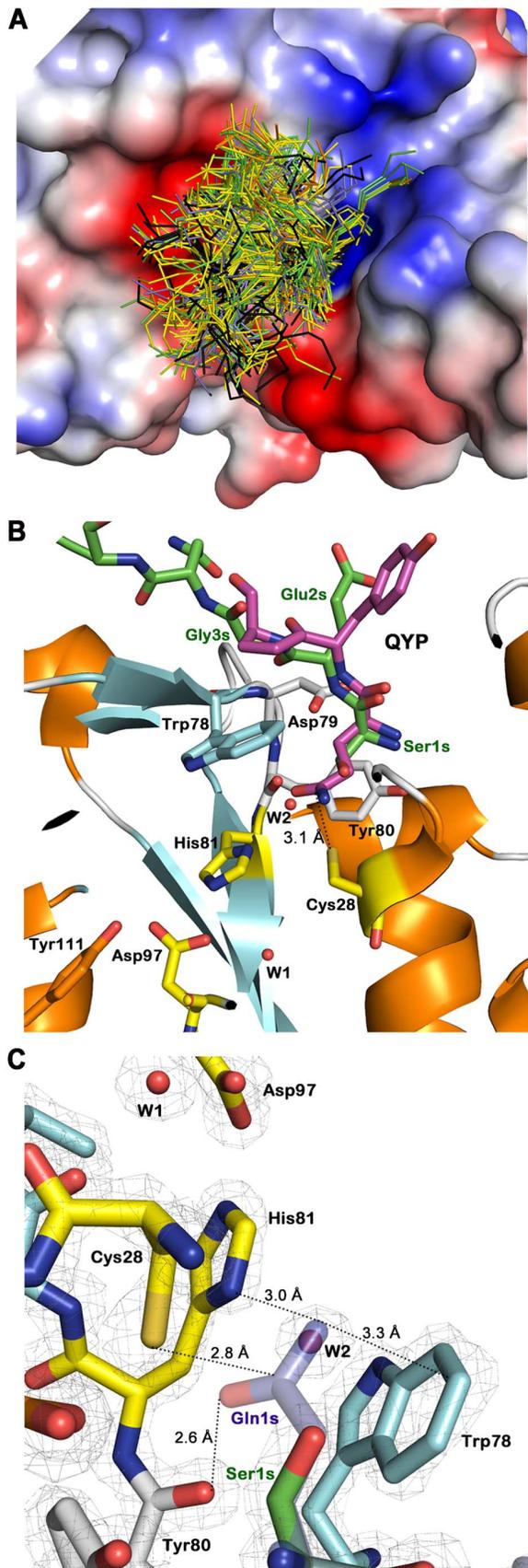


Figure 3. Docking study and suggested binding pose of N-terminal glutamine peptide in hNtaq1. (A) Tripeptides docking study of hNtaq1. Categorized by features of its second residue, the

backbone of Ala, Gly, Ile, Leu, Met, Pro, Val tripeptides are colored in yellow, the backbone of Cys, Asn, Gln, Ser, Thr tripeptides are colored in green, the His, Lys, Arg tripeptides are colored in blue, the Phe, Trp, Tyr tripeptides are colored in black, and the Asp, Glu tripeptides are colored in orange. (B) The nearest tripeptide in docking study. The substrate-mimicking peptide is shown as green and predicted docking tripeptide Gln-Tyr-Pro is colored in magenta. (C) Binding mode of refined Ser1Gln hNtaq1 mutant on electron density map of hNtaq1. Carbon in substrate-mimicking peptide, catalytic triad, β -strands, loops, and Ser1Gln mutant are colored in green, yellow, cyan, white, and blue, respectively. Oxygen, nitrogen, and sulfur are represented as red, blue, and gold, respectively. Two water molecules are shown as red sphere and electron density map is represented as gray mesh contoured at 2.0 σ .
doi:10.1371/journal.pone.0111142.g003

with the three N-terminal residues, Ser-Glu-Gly, showed similar binding poses to that from the crystal structure of hNtaq1 bound with the substrate-mimicking peptide. In the control docking study, stabilization energy of the best predicted binding mode was -4.8 kcal/mol. In the experimental docking study, the average stabilization energy of 400 docking results was 5.0 kcal/mol similar to that of the control docking. From the docking results, most stable tripeptides were QWF, QWW, QWQ, QHW, QWV, QQW, QHF, QGW, QEW, and QWS in the order of stabilization energy ranging from -5.8 to -5.3 kcal/mol. Backbones of all the 400 tripeptides seemed to be recognized by the predominantly negatively-charged patch around the substrate binding cleft, which confers nonspecific interaction with substrates regardless of side chains of second and thereafter residues or with very minor influences (Figure 3A). In the case of the closest tripeptide from the catalytic triad, Gln-Tyr-Pro, C $_8$ of glutamine residue is located 3.15 Å away from sulfhydryl group of Cys28 in the catalytic triad and stabilization energy was -4.6 kcal/mol (Figure 3B).

Interestingly, when we refined our structure with N-terminal Ser1Gln mutation to get a clue of reaction mechanism, modified first serine residue and water2 exactly overlapped with mutated glutamine in the refined structure, the actual substrate of hNtaq1, thus they are reminiscent of an actual substrate (Figure 3C). Sulfur of catalytic Cys28 is 2.8 Å away from the C $_8$ of Gln1s and amide nitrogen of Gln1s is 3.0 Å away from the spot on His81, optimal distance for protonation of a leaving amine group. Carbonyl oxygen of Gln1s seems to lock the N-terminal glutamine via a hydrogen bond to peptide carbonyl of Tyr80 (2.6 Å). We suggest that coordinates of Ser1 and water2 in our crystal structure mimic the pose of an N-terminal glutamine and this binding mode would represent the substrate binding step in hNtaq1 mechanism.

Proposed mechanism of hNtaq1

Based on our crystal structure of hNtaq1 and docking study with all the possible anticipated substrate tripeptides, we suggest a catalytic mechanism of hNtaq1 as shown in Figure 4. In the first step, nucleophilic sulfhydryl group of Cys28 approaches C $_8$ of the amide group of the N-terminal glutamine and becomes deprotonated by His81 as shown in Figure 4 step 1. The sulfhydryl group of Cys28 plays a crucial role in the nucleophilic attack on acyl group in the N-terminal glutamine side chain of substrates, which results in formation of a tetrahedral intermediate (Figure 4 step 2). Asp97 facilitates the process by forming a hydrogen bond and electrostatic interactions with His81. The ammonia is released upon productive collapse of the tetrahedral intermediate and a water molecule enters the active site cavity and attacks S-acyl intermediate to convert glutamine to a glutamate (Figure 4 step 3 and 4). As the final step, the glutamate side chain is cleaved from S-acyl of Cys28 (Figure 4 step 5 and 6). In these steps, His81 first

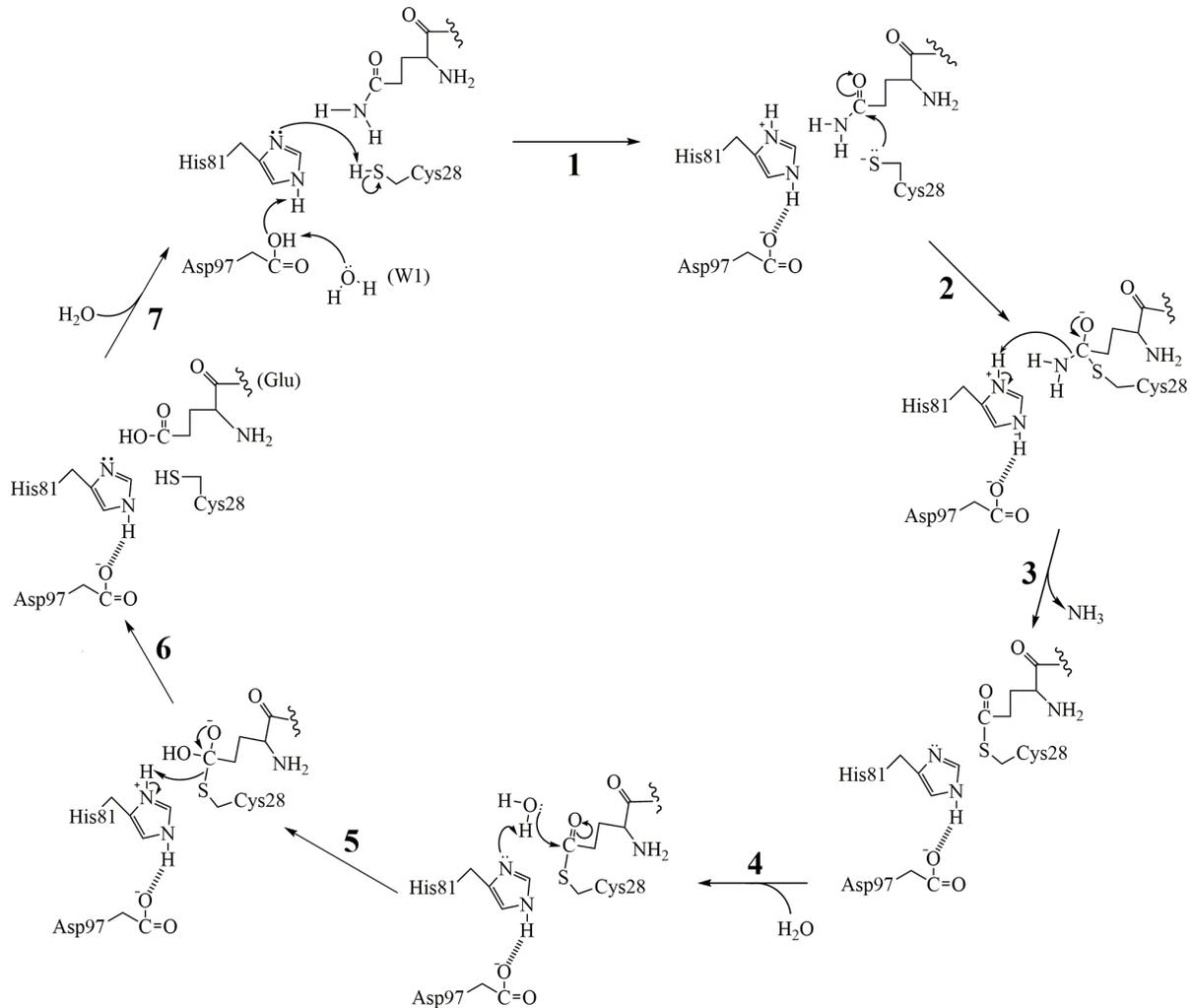


Figure 4. Proposed catalytic mechanism of hNtaq1.
doi:10.1371/journal.pone.0111142.g004

acts as a general base activation water for a nucleophilic attack on the S-acyl intermediate, and then upon collapse of the tetrahedral intermediate acts a general acid to protonate the leaving group, i.e. the thiolate of Cys28. The substrate peptide with newly formed N-terminal glutamate is released from the binding cleft at this stage and the enzyme is ready for another round of catalysis (Figure 4 step 7). The proposed reaction mechanism of hNtaq1 and structural information from our study will provide valuable information for understanding the N-end rule pathway and the interaction between hNtaq1 and its protein substrate.

Acknowledgments

The authors thank all members of the Center for Eukaryotic Structural Genomics for various important contributions. GM/CA@APS has been

References

- Mogk A, Schmidt R, Bukau B (2007) The N-end rule pathway for regulated proteolysis: prokaryotic and eukaryotic strategies. *Trends Cell Biol* 17: 165–172.
- Lecker SH, Goldberg AL, Mitch WE (2006) Protein degradation by the ubiquitin-proteasome pathway in normal and disease states. *J Am Soc Nephrol* 17: 1807–1819.
- Knop M, Schiffer HH, Rupp S, Wolf DH (1993) Vacuolar/lysosomal proteolysis: proteases, substrates, mechanisms. *Curr Opin Cell Biol* 5: 990–996.
- Bachmair A, Finley D, Varshavsky A (1986) In vivo half-life of a protein is a function of its amino-terminal residue. *Science* 234: 179–186.
- Tasaki T, Kwon YT (2007) The mammalian N-end rule pathway: new insights into its components and physiological roles. *Trends Biochem Sci* 32: 520–528.
- Baker RT, Varshavsky A (1995) Yeast N-terminal amidase. A new enzyme and component of the N-end rule pathway. *J Biol Chem* 270: 12065–12074.
- Kwon YT, Balogh SA, Davydov IV, Kashina AS, Yoon JK, et al. (2000) Altered activity, social behavior, and spatial memory in mice lacking the NTAN1p

- amidase and the asparagine branch of the N-end rule pathway. *Mol Cell Biol* 20: 4135–4148.
8. Wang H, Piatkov KI, Brower CS, Varshavsky A (2009) Glutamine-specific N-terminal amidase, a component of the N-end rule pathway. *Mol Cell* 34: 686–695.
 9. Hwang CS, Shemorry A, Varshavsky A (2009) Two proteolytic pathways regulate DNA repair by cotargeting the Mgt1 alkylguanine transferase. *Proc Natl Acad Sci U S A* 106: 2142–2147.
 10. Eisele F, Wolf DH (2008) Degradation of misfolded protein in the cytoplasm is mediated by the ubiquitin ligase Ubr1. *FEBS Lett* 582: 4143–4146.
 11. Kwon YT, Xia Z, An JY, Tasaki T, Davydov IV, et al. (2003) Female lethality and apoptosis of spermatocytes in mice lacking the UBR2 ubiquitin ligase of the N-end rule pathway. *Mol Cell Biol* 23: 8255–8271.
 12. Ditzel M, Wilson R, Tenev T, Zachariou A, Paul A, et al. (2003) Degradation of DIAP1 by the N-end rule pathway is essential for regulating apoptosis. *Nat Cell Biol* 5: 467–473.
 13. Thao S, Zhao Q, Kimball T, Steffen E, Blommel PG, et al. (2004) Results from high-throughput DNA cloning of *Arabidopsis thaliana* target genes using site-specific recombination. *J Struct Funct Genomics* 5: 267–276.
 14. Sreenath HK, Bingman CA, Buchan BW, Seder KD, Burns BT, et al. (2005) Protocols for production of selenomethionine-labeled proteins in 2-L polyethylene terephthalate bottles using auto-induction medium. *Protein Expr Purif* 40: 256–267.
 15. Jeon WB, Aceti DJ, Bingman CA, Vojtik FC, Olson AC, et al. (2005) High-throughput purification and quality assurance of *Arabidopsis thaliana* proteins for eukaryotic structural genomics. *J Struct Funct Genomics* 6: 143–147.
 16. Studier FW (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* 41: 207–234.
 17. Bricogne G, Vornrhein C, Flensburg C, Schiltz M, Paciorek W (2003) Generation, representation and flow of phase information in structure determination: recent developments in and around SHARP 2.0. *Acta Crystallogr D Biol Crystallogr* 59: 2023–2030.
 18. Cowtan KD, Zhang KY (1999) Density modification for macromolecular phase improvement. *Prog Biophys Mol Biol* 72: 245–270.
 19. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60: 2126–2132.
 20. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 53: 240–255.
 21. Laskowski RA, Moss DS, Thornton JM (1993) Main-chain bond lengths and bond angles in protein structures. *J Mol Biol* 231: 1049–1067.
 22. Lovell SC, Davis IW, Arendall WB 3rd, de Bakker PI, Word JM, et al. (2003) Structure validation by Calpha geometry: phi, psi and Cbeta deviation. *Proteins* 50: 437–450.
 23. Trott O, Olson AJ (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* 31: 455–461.
 24. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, et al. (2009) AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem* 30: 2785–2791.
 25. The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.
 26. Holm L, Rosenstrom P (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res* 38: W545–549.
 27. Hashizume R, Maki Y, Mizutani K, Takahashi N, Matsubara H, et al. (2011) Crystal structures of protein glutaminase and its pro forms converted into enzyme-substrate complex. *J Biol Chem* 286: 38691–38702.
 28. Bhaskaran SS, Stebbins CE (2012) Structure of the catalytic domain of the *Salmonella* virulence factor SseI. *Acta Crystallogr D Biol Crystallogr* 68: 1613–1621.
 29. Chatterjee D, Boyd CD, O'Toole GA, Sondermann H (2012) Structural characterization of a conserved, calcium-dependent periplasmic protease from *Legionella pneumophila*. *J Bacteriol* 194: 4415–4425.
 30. Choi WS, Jeong BC, Joo YJ, Lee MR, Kim J, et al. (2010) Structural basis for the recognition of N-end rule substrates by the UBR box of ubiquitin ligases. *Nat Struct Mol Biol* 17: 1175–1181.