Biodesign

pISSN 2288-6982 | eISSN 2288-7105

CRYSTALLIZATION P 71-74

Purification, crystallization, and X-ray crystallographic analysis of histidine triad nucleotide-binding protein from *Candida albicans*

Ahjin Jung, Shinae Kim and Jeong Ho Chang*

Department of Biology Education, Kyungpook National University, Daegu 41566, Republic of Korea

*Correspondence: jhcbio@knu.ac.kr

Histidine triad nucleotide-binding protein (HINT) is a member of the histidine triad (HIT) superfamily, and exhibits dinucleotide hydrolase activity via a histidine triad motif. HITs are divided into five classes that have various functions including galactose metabolism, DNA repair, and tumor suppression. Although multiple crystal structures have been reported, limited structural information is available for HINT proteins of fungal species. In this study, to understand the structural features and functional diversity of HINT proteins, we crystallized HINT from four fungal species and obtained X-ray diffraction data from the pathogenic fungus *Candida albicans* at a resolution of 2.5 Å. The crystal of the *C. albicans* HINT (CaHINT) belonged to the space group of P2221, with unit cell parameters a = 40.4, b = 101.9, c = 175.2 Å, $\alpha = \beta = \gamma = 90°$. The crystal contained four macromolecules in asymmetric units.

INTRODUCTION

Histidine triad nucleotide binding protein (HINT) is a member of the histidine triad (HIT) protein superfamily that contains three highly conserved histidine residues involved in catalysis (Bardaweel et al., 2010). A conserved HXHXHXX sequence motif (where X represents a hydrophobic residue) in the active site supports several catalytic activities such as the hydrolysis of nucleotides and phosphoramidate, and nucleotidyl transfer (Krakowiak et al., 2004; Martin et al., 2011). In the active site, while the first histidine stabilizes the second histidine by its main chain, the well-conserved second and third histidine participate in catalytic reactions by forming a covalent nucleotidyl-phosphohistidyl intermediate (Maize et al., 2013). In addition, a conserved CXXC motif (where X represents a hydrophobic residue) coordinates a zinc ion via cysteine residues, together with the first histidine residue (Klein et al., 1998; Lima et al., 1996).

Based on the catalytic specificities, sequence composition, and structural similarities, members of the HIT superfamily have been grouped into five classes: HINT, galactose-1-phosphate uridylyltransferase (GalT), fragile HIT protein (Fhit), DcpS, and aprataxin (Martin et al., 2011). The HINT proteins are classified as the oldest branch of the superfamily, and its members are extensively distributed among eukaryotes and archaea. In human, three isotypes of HINT (termed hHint family) are currently known; they are Hint1, Hint2, and Hint3. Hint4 was also initially identified but was found to be identical to Hint3 (Chou et al., 2007; Maize et al., 2013). Hint1 plays a role as a tumor suppressor by inducing apoptosis without its enzymatic activity (Weiske and Huber, 2006). Further, Hint1 hydrolyzes aminoacyl-adenylates because lysyl-adenylate, produced by lysyl-tRNA synthetase, is a substrate of Hint1; this suggests a possible role of Hint1 as a regulator of translation (Butland et al., 2005; Chou and Wagner, 2007). In addition, a mutation in Hint1 causes heritable axonal neuropathy with neuromyotonia (Zimon et al., 2012). Both hHint2 and hHint3 are involved in the development of cancers such as breast, pancreatic, and colon cancers via their upregulation relative to hHint1 expression (Martin et al., 2011).

More than 20 HINT structures from a wide variety of organisms have been deposited in the Protein Data Bank (PDB). However, limited structural information for HINT from fungal/yeast species are currently available. In this study, to understand the specific structural features and phylogenetic relationship of fungal HINT, we overexpressed, purified, and crystallized full-length HINT from the pathogenic fungus *Candida albicans*.

RESULTS AND DISCUSSION

The HINT gene from *Candida albicans* was cloned into the pET26b vector, and the recombinant plasmid was transformed into *Escherichia coli* BL21 (DE3) Rosetta cells (Table 1). The C-terminal hexa-histidine tagged HINT from *Candida albicans* (CaHINT) was purified by Ni-NTA affinity chromatography followed by size-exclusion chromatography (SEC). The SEC elution peak profile showed a single symmetric peak at an elution volume of 76.2 mL, which represented a molecular weight of approximately 60 kDa (Figure 1a). As the calculated molecular

TABLE 1 | Protein production information

Source organism	*Candida albicans*
DNA source	Genomic DNA
Forward primer* (5′-3′)	GGGAAACATATGGCTTCTCATGCTTCCTGT ATATT
Reverse primer* (5′-3′)	GGACCTCGAGTAATTTTTCATTATCTACCT TAGCCAATTCTTC
Cloning vector	pET26b
Expression host	*Escherichia coli* BL21 Rosetta (DE3)
Complete amino acid sequence of the construct produced (5′-3′)	MASHASCIFCKIIKGEIPSFKLIETAKTYSFLDIQPIA EAHVLIIPKHHGAKLHNIPDDYLSDILPVVKKLTK VLKLDENNTPEGEGYNVLQNNGRIAHQVVDHVH FHLIPKKDEATGLGVGWPAEATDFDKLGKLHEKL KEELAKVDNEKL

* Underlined sequence: restriction-enzyme site.

TABLE 2 | Crystallization

Method	Vapor diffusion
Plate type for screening	96-well sitting-drop crystallization plate, Art Robbins Instruments
Plate type for optimization	24-well plate, Hyundai Micro
Temperature (°C)	4
Protein concentration (mg/mL)	50
Composition of protein solution	20 mM Tris-HCl pH 7.5, 150 mM NaCl, 2 mM DTT
Composition of reservoir solution	0.2M $MgCl_2$, 20% (w/v) PEG3350
Volume and ratio of drop	2 μL, 1:1
Volume of reservoir (μL)	500

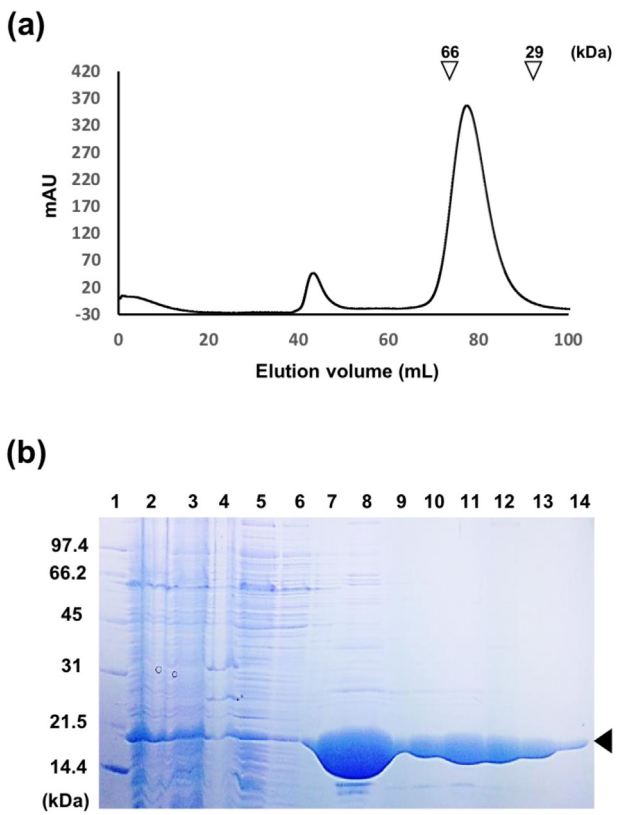

FIGURE 1 | **Preparation of CaHINT protein.** (a) SEC profile of C-terminal His6-HINT from *Candida albicans*. The standards were calibrated using albumin (66 kDa) and carbonic anhydrase (29 kDa) as indicated above the SEC profile. (b) SDS-PAGE analysis of CaHINT. The black arrowhead indicates the size of the target protein. Molecular weights of the standard size markers are shown on the left-hand side. Lanes are as follows: 1, marker; 2, cells induced with IPTG; 3, supernatant; 4, cell pellet; 5, flow-through; 6, wash with 20 mM imidazole; 7, wash with 25 mM imidazole; 8, elution with 250 mM imidazole; 9–14, fractions from the gel filtration peaks on the SEC profile.

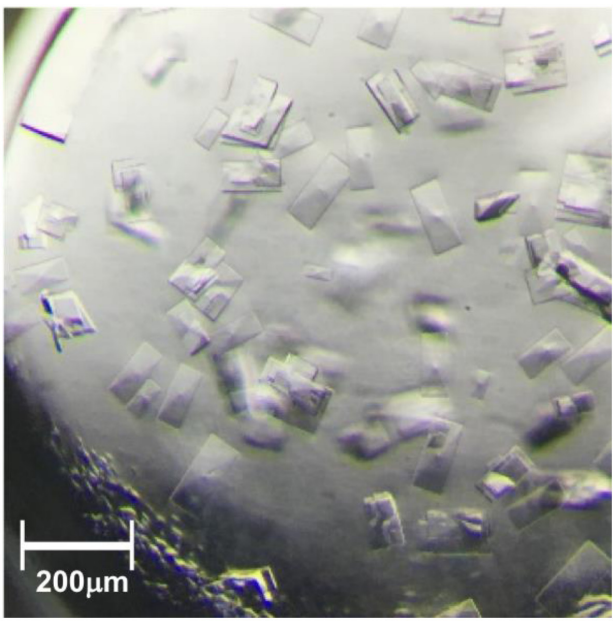

FIGURE 2 | **Crystals of CaHINT protein.** Plate-shaped crystals of CaHINT were obtained in a solution of 0.2 M magnesium chloride and 20% PEG 3350 at 4°C after 2 days. The scale bar is presented in the left bottom side.

weight of monomeric CaHINT was 16.9 kDa, this indicated that CaHINT exists as a tetramer in solution. The purified protein was subsequently concentrated to 50 mg/mL based on UV absorbance at 280 nm, with a high purity of over 95% as determined via SDS-PAGE analysis (Figure 1b).

Crystallization trials were performed at 4°C using the hanging-drop vapor diffusion method with over 600 different conditions from sparse-matrix screening solution kits. The plate-shaped crystals of CaHINT usually appeared in 2 days with the condition of 0.2 M magnesium acetate and 20% (w/v) polyethyleneglycol (PEG); the size of the crystals was approximately 200 × 100 × 15 μm (Figure 2). Optimal crystals were obtained in a solution of 0.2 M magnesium chloride and 20% (w/v) PEG 3350 (Table 2). The optimized CaHINT protein crystals were flash-cooled in liquid nitrogen in the presence of 15% (v/v) glycerol with the crystallized condition. X-ray diffraction data of the optimized

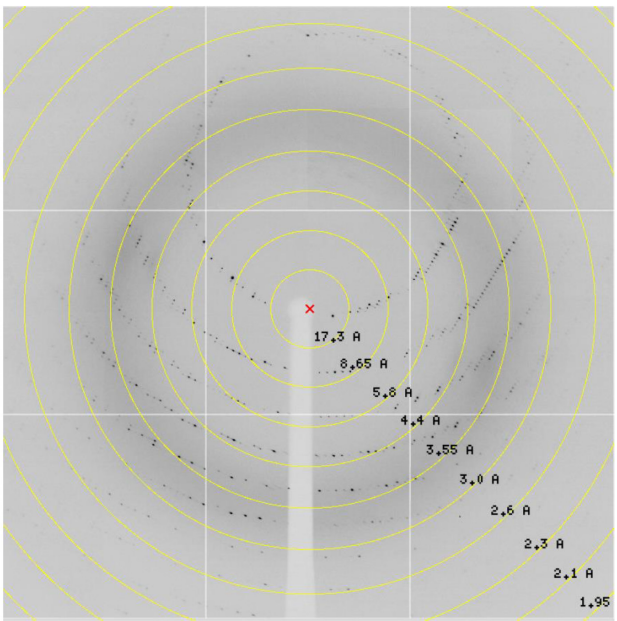

FIGURE 3 | A representative diffraction pattern of the CaHINT crystal. Diffraction image of a CaHINT crystal. The yellow circles and black text represent resolution ranges.

crystals were collected to 2.5 Å resolution on the Beamline 5C at the Pohang Accelerator Laboratory (Pohang, Republic of Korea) (Figure 3). The CaHINT crystals belonged to the primitive orthogonal space group $P222_1$ with the following unit cell parameters: a = 40.354, b = 101.907, c = 175.175, α = β = γ = 90°. The V_M and Matthews coefficient were 720381.812 and 2.53, respectively, which indicated that there were four molecules in the asymmetric unit, with 51.38% solvent content (Matthews, 1968). The crystallographic tetrameric state of CaHINT was consistent with that of the SEC results (Figure 1a).

We obtained initial phase information by molecular replacement using the monomeric structure of HINT from *Leishmania major* (PDB code: 3KSV) which shows 43% sequence identity as a search model. Molecular replacement was performed using *Phaser-MR* in the *PHENIX* crystallographic software package (PHENIX; USA) (Adams et al., 2010). The top MR solution showed that the scores of LLG and TFZ were 8691 and 50.3, respectively. The initial round of refinement using the best model obtained by molecular replacement was subjected to rigid-body refinement, resulting in convergence to an R factor of 24.53% and an R_{free} of 29.65% with four chains of CaHINT in the asymmetric unit. Currently, further refinement and model building are in progress using PHENIX.refine and Wincoot softwares (Adams et al., 2010; Emsley and Cowtan, 2004). The detailed crystallographic information is summarized in Table 3.

TABLE 3 | Data collection and processing

	CaHINT
Diffraction source	Beamline 5C, PAL
Wavelength (Å)	1.2818
Temperature (°C)	-180
Detector	ADSC Quantum 315
Crystal-to-detector distance (mm)	270
Rotation range per image (°)	1
Total rotation range (°)	360
Exposure time per image (s)	1
Space group	$P2_12_12$
a, b, c (Å)	40.354, 101.907, 175.175
α, β, γ (°)	90, 90, 90
Resolution range (Å)	50.0–2.5
Total no. of reflections	759049
No. of unique reflections	26068
Completeness (%)	100 (100)[a]
Multiplicity	7.5
I / σI	35.2 (7.8)
R_{merge} (%)[b]	10.9 (41.3)

[a]The numbers in parentheses are statistics from the highest resolution shell.
[b]R_{merge} = Σ $|I_{obs} - I_{avg}|$ / I_{obs}, where I_{obs} is the observed intensity of individual reflections and I_{avg} is averaged over symmetry equivalents.

METHODS

Preparation of the CaHINT protein

Genomic DNA from the *Candida albicans* was obtained from the Korean Collection for Type Cultures (KCTC, Republic of Korea). Full-length HINT was designed to contain a C-terminal His_6-tag, and amplified by polymerase chain reaction (PCR) using Pfu-X DNA polymerase (Solgent, Republic of Korea). Oligonucleotide primers used in the study were purchased from Cosmogenetech Inc. The amplified fragment was digested with the restriction enzymes NdeI and XhoI (R006S and R007S, respectively; Enzynomics, Republic of Korea) for 3 h at 37°C in a water bath, and ligated into the pET26b vector using T4 ligase (M0202S; Roche, Germany) overnight at 20°C. The ligated plasmid was then transformed into the *Escherichia coli* strain DH5α, and transformants were confirmed by colony PCR. The CaHINT plasmid was transformed into *E. coli* Rosetta cells, which were grown at 37°C to an optical density at 600 nm of approximately 0.6 in LB medium (Ambrothia, Republic of Korea) containing 50 mg/L kanamycin (Applichem, USA) and 50 mg/L Chloroamphenicol (Applichem, USA). Following induction with 0.3 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) (Calbiochem, Germany), and the addition of 100 mM zinc chloride (Junsei Chemical, Tokyo), the cells were further grown for 16 h at 20°C. The cultured cells were harvested by centrifugation at 4500×g for 20 min at 4°C. The cell pellet was resuspended in buffer containing 20 mM Tris (pH 8.0) (Sigma-Aldrich, USA), 250 mM NaCl (Applichem, USA), 5% glycerol (Affymetrix, USA), 0.2% Triton X-100 (Sigma-Aldrich, USA), 10 mM β-mercaptoethanol (BioBasic, Canada), and 0.2 mM phenylmethylsulfonyl fluoride (Sigma-Aldrich, USA). Cells were disrupted by ultrasonication with 3 sec on/off pulses for 20 min. Cell debris was removed by centrifugation at 13000×g for 40 min, and the supernatant was bound to Ni-NTA agarose (Qiagen, Germany) for 90 min at 4°C. After washing with buffer A (200 mM NaCl, 50 mM Tris, pH 8.0) containing 20 mM imidazole (Sigma-Aldrich, USA), the bound proteins were eluted with 250 mM imidazole. SEC was performed on the purified proteins using a HiPrep 16/60 Sephacryl S-300 HR column (GE Healthcare, Canada). The SEC buffer contained 20 mM Tris (pH 7.5), 150 mM NaCl, and 2 mM dithiothreitol (DTT; Calbiochem, Germany).

Following SEC, proteins were stored at −80°C pending crystallization trials. The purified proteins were assessed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis using a 18% acrylamide gel, which showed a single band corresponding to the calculated molecular weight of the target protein.

Crystallization

All crystallization trials were performed at 4°C using the sitting-drop vapor diffusion method in 96-well sitting-drop plates (Art Robbins Instruments, USA). Over 600 different conditions from sparse-matrix screening solution kits were tested to identify optimal crystallization conditions. The kits used included Crystal Screen 1/2 (HR2-110 and -112), Index (HR2-144), Salt Rx 1/2 (HR2-107 and -109), and PEG/Ion (HR2-126 and -098) from Hampton Research (USA), in addition to Wizard 1/2 (CS-311, Jena Bioscience, Germany) and SG1 screen (MD1-88, Molecular dimensions, UK). Crystals of CaHINT grew within two days in drops containing equal volumes (1 μL) of protein sample (50 mg/mL in 150 mM NaCl, 2 mM DTT, and 20 mM Tris [pH 7.5]) and reservoir solution (20% w/v PEG 3350, 0.2 M magnesium acetate). To improve the crystals, additional screening was performed using the Additive (HR2-428, Hampton Research, USA) and Detergent (HR2-406, from Hampton Research, USA) screening kits. The optimized crystallization condition was 20% w/v PEG 3350 and 0.2 M magnesium chloride.

Data collection and processing

Prior to data collection, 15% glycerol was added to the reservoir solutions as a cryoprotectant, and crystals were flash-cooled in liquid nitrogen. All diffraction datasets were collected at 100 K on the Beamline 5C at the Pohang Accelerator Laboratory (PAL; Republic of Korea) using a Quantum 315 CCD detector (ADSC, USA)(Park et al., 2017). Data were processed using the HKL2000 software suite. Experimental electron density maps were obtained by Molecular replacement methods using PHENIX software, version 1.9, and were interpreted using the *WinCoot* program (PHENIX; USA) (Adams et al., 2010; Emsley and Cowtan, 2004).

CONFLICT OF INTEREST

The authors declare no potential conflicts of interest.

ACKNOWLEDGEMENTS

We would like to thank beamline staff Yeon-Gil Kim at beamline 5C of the Pohang Accelerator Laboratory (Pohang, Korea) for data collection. This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT to JHC (2016R1C1B2009691).

Original Submission: Aug 30, 2018
Revised Version Received: Sep 6, 2018
Accepted: Sep 6, 2018

REFERENCES

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., McCoy, A.J., Moriarty, N.W., Oeffner, R., Read, R.J., Richardson, D.C., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221.

Bardaweel, S., Pace, J., Chou, T.F., Cody, V., and Wagner, C.R. (2010). Probing the impact of the echinT C-terminal domain on structure and catalysis. *J Mol Biol* **404**, 627-638.

Brenner, C. (2002). Hint, Fhit, and GalT: function, structure, evolution, and mechanism of three branches of the histidine triad superfamily of nucleotide hydrolases and transferases. *Biochemistry* **41**, 9003-9014.

Butland, G., Peregrin-Alvarez, J.M., Li, J., Yang, W., Yang, X., Canadien, V., Starostine, A., Richards, D., Beattie, B., Krogan, N., Davey, M., Parkinson, J., Greenblatt, J., and Emili, A. (2005). Interaction network containing conserved and essential protein complexes in Escherichia coli. *Nature* **433**, 531-537.

Chou, T.F., Cheng, J., Tikh, I.B., and Wagner, C.R. (2007). Evidence that human histidine triad nucleotide binding protein 3 (Hint3) is a distinct branch of the histidine triad (HIT) superfamily. *J Mol Biol* **373**, 978-989.

Chou, T.F., and Wagner, C.R. (2007). Lysyl-tRNA synthetase-generated lysyl-adenylate is a substrate for histidine triad nucleotide binding proteins. *J Biol Chem* **282**, 4719-4727.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132.

Klein, M.G., Yao, Y., Slosberg, E.D., Lima, C.D., Doki, Y., and Weinstein, I.B. (1998). Characterization of PKCI and comparative studies with FHIT, related members of the HIT protein family. *Exp Cell Res* **244**, 26-32.

Krakowiak, A., Pace, H.C., Blackburn, G.M., Adams, M., Mekhalfia, A., Kaczmarek, R., Baraniak, J., Stec, W.J., and Brenner, C. (2004). Biochemical, crystallographic, and mutagenic characterization of hint, the AMP-lysine hydrolase, with novel substrates and inhibitors. *J Biol Chem* **279**, 18711-18716.

Lima, C.D., Klein, M.G., Weinstein, I.B., and Hendrickson, W.A. (1996). Three-dimensional structure of human protein kinase C interacting protein 1, a member of the HIT family of proteins. Proc Natl Acad Sci *U S A* **93**, 5357-5362.

Maize, K.M., Wagner, C.R., and Finzel, B.C. (2013). Structural characterization of human histidine triad nucleotide-binding protein 2, a member of the histidine triad superfamily. *FEBS J* **280**, 3389-3398.

Martin, J., St-Pierre, M.V., and Dufour, J.F. (2011). Hit proteins, mitochondria and cancer. *Biochim Biophys Acta* **1807**, 626-632.

Matthews, B.W. (1968). Solvent content of protein crystals. *J Mol Biol* **33**, 491-497.

Park, S.Y., Ha, S.C., and Kim, Y.G. (2017). The Protein Crystallography Beamlines at the Pohang Light Source II. *Biodesign* **5**, 30-34.

Weiske, J., and Huber, O. (2006). The histidine triad protein Hint1 triggers apoptosis independent of its enzymatic activity. *J Biol Chem* **281**, 27356-27366.

Zimon, M., Baets, J., Almeida-Souza, L., De Vriendt, E., Nikodinovic, J., Parman, Y., Battaloglu, E., Matur, Z., Guergueltcheva, V., Tournev, I., Auer-Grumbach, M., De Rijk, P., Petersen, B.S., Muller, T., Fransen, E., et al. (2012). Loss-of-function mutations in HINT1 cause axonal neuropathy with neuromyotonia. *Nat Genet* **44**, 1080-1083.