Acta Crystallographica Section F
**Structural Biology and Crystallization Communications**
ISSN 1744-3091

# Sample preparation, crystallization and structure solution of HisC from *Mycobacterium tuberculosis*

**Nazia Nasir, Rajan Vyas‡ and Bichitra K. Biswal***

Protein Crystallography Laboratory, National Institute of Immunology, Aruna Asaf Ali Marg, New Delhi 110 067, India

‡ Current address: Department of Biochemistry, 484 West 12th Avenue, Ohio State University, Columbus, OH 43210, USA.

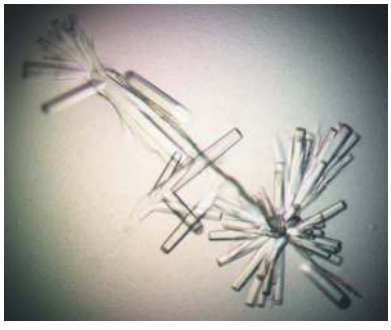
Correspondence e-mail: bbiswal@nii.res.in

Received 22 January 2013
Accepted 4 March 2013

Histidinolphosphate aminotransferase (HisC; Rv1600) from *Mycobacterium tuberculosis* was overexpressed in *M. smegmatis* and purified to homogeneity using nickel–nitrilotriacetic acid metal-affinity and gel-filtration chromatography. Diffraction-quality crystals suitable for X-ray analysis were grown by the hanging-drop vapour-diffusion technique using 30% polyethylene glycol monomethyl ether 2000 as the precipitant. The crystals belonged to the hexagonal space group $P3_221$, with an unusual high solvent content of 74.5%. X-ray diffraction data were recorded to 3.08 Å resolution from a single crystal using in-house Cu $K\alpha$ radiation. The structure of HisC was solved by the molecular-replacement method using its *Corynebacterium glutamicum* counterpart as a search model. HisC is a dimer in the crystal as well as in solution.

## 1. Introduction

Aminotransferases or transaminases are ubiquitous enzymes that are involved in amino-acid biosynthesis, carbohydrate metabolism, carbon assimilation *etc*. These enzymes catalyze a reaction involving the transfer of an amino group from an amino donor to an acceptor group. The *Mycobacterium tuberculosis* (*Mtb*) genome contains about 15 different aminotransferases (Cole *et al.*, 1998), of which two, HisC and HisC2, encoded by the genes Rv1600 and Rv3772, respectively, are involved in histidine metabolism as histidinol-phosphate aminotransferases. Both enzymes catalyze the conversion of imidazole acetol-phosphate to histidinol phosphate. The histidine-biosynthesis pathway, which provides an enzymatic synthesis of histidine from 5-phosphoribosyl-1-pyrophosphate in ten steps, is conserved in archaea, bacteria, fungi and plants, but is absent in mammals (Alifano *et al.*, 1996; Nasir *et al.*, 2012). Moreover, a high-density mutagenesis study suggests that *Mtb* requires histidine for optimal growth (Sassetti *et al.*, 2003). The molecules of this pathway are therefore potential targets for antituberculosis drug discovery. In this respect, studying the structure–function relationship of the enzymes of this pathway is of interest as it will help in understanding the molecular basis of histidine biosynthesis in *Mtb* and, importantly, their three-dimensional structures will be helpful for the design of inhibitors of these enzymes using a structure-based approach. With this aim, we have been pursuing a program aiming to determine the crystal structures of as many enzymes as possible in the pathway. Previously, we have reported the details of the molecular cloning, overexpression in *M. smegmatis*, crystallization and preliminary X-ray diffraction studies of two His-pathway enzymes: HisB and HisC2 (Ahangar *et al.*, 2011; Nasir *et al.*, 2012). In the present study, we report the details of the enzyme preparation, crystallization and structure-solution studies of histidinol-phosphate aminotransferase (HisC), which shares 29% sequence identity with HisC2.





## 2. Materials and methods

### 2.1. Preparation of the expression construct and overexpression

The gene encoding HisC was cloned into the expression vector pYUB1062 *via* an entry vector, pENTR, using a previously described protocol (Nasir *et al.*, 2012). Briefly, the gene was amplified by PCR

**Table 1**
Sequences of the primers used in PCR and of the recombinant HisC.

| Sequence description | Sequence |
|---|---|
| Forward primer | 5′-**CACC**CATATGATGACCAGGTCCGGACACCC-3′ |
| Reverse primer | 5′-**TAT**AAGCTTTGGCGCTCCTACAGGACTGC-3′ |
| Recombinant protein sequence | MMTRSGHPVTLDDLPLRADLRGKAPYGAPQLAVPVRL-NTNENPHPPTRALV DDVVRSVREAAIDLHRYPDRDA-VALRADLAGYLTAQTGIQLGVENIWAAN GSNEILQQ-LLQAFGGPGRSAIGFVPSYSMHPIISDGTHTEWIEASR-ANDFGLDVDVAVAAVVDRKPDVVFIASPNNPSGQSVS-LPDLCKLLDVAPGIAIV DEAYGEFSSQPSAVSLVEEYP-SKLVVTRTMSKAFAFAGGRLGYLIATPAV IDAMLLVR-LPYHLSSVTQAAARAALRHSDDTLSSVAALIAERERV-TTSLNDMGFRVIPSDANFVLFGEFADAPAAWRRYLEA-GILIRDVGIPGYLRATTG LAEENDAFLRASARIATDL-VPVTRSPVGAP*KLAAALE*HHHHHH |

using the gene-specific forward primer [with four directional cloning-specific nucleotides (shown in bold), an *Nde*I restriction site (underlined) and the first 20 nucleotides of the open reading frame (ORF) Rv1600] and reverse primer [with a *Hin*dIII restriction site (underlined), flanking nucleotides (bold) and the reverse complement of last 20 nucleotides of the ORF] shown in Table 1, Phusion polymerase, *Mtb* H37Rv genomic DNA, dNTPs and $MgCl_2$. The PCR product was cloned into the entry vector pENTR using the pENTR/TOPO directional cloning kit (Invitrogen) as per the manufacturer's protocol. The entry clone, pENTR1600 and the final vector pYUB1062 were then digested using the *Nde*I and *Hin*dIII restriction enzymes. Ligation was carried out overnight at 289 K using T4 DNA ligase, the reaction mixture was transformed into DH5$\alpha$ cells and the transformed colonies were selected on hygromycin B (150 µg ml$^{-1}$) Luria–Bertani (LB) agar plates.

The expression plasmid pYUB1600 was electroporated into the *M. smegmatis* mc$^2$4517 expression host at 2500 V, 1000 Ω and 25 µF using a 2 mm diameter electroporation cuvette (Bio-Rad). The cells were then plated on 7H10 agar with oleic acid–albumin–dextrose–catalase (OADC) nutrient supplements and the antibiotics hygromycin B (100 µg ml$^{-1}$) and kanamycin (25 µg ml$^{-1}$) for selection of transformed cells. Colonies appeared after approximately 72 h, one of which was inoculated in a 50 ml flask containing 10 ml LB broth with 0.05% Tween-80 and 0.2% glycerol (LBTG) to revive the colony. The culture was grown for about 24 h at 310 K at 180 rev min$^{-1}$. A primary culture was inoculated from the revival culture in 50 ml LBTG. The culture was incubated for 12 h at 310 K with shaking until the $OD_{600}$ reached around 1. A 2 l secondary culture was set up from the primary culture and induced with 0.2% acetamide at an $OD_{600}$ of about 0.7. The cells were harvested 20 h after induction by centrifugation at 277 K and were frozen at 253 K until further processing.

### 2.2. Purification

Since the recombinant protein has a 6×His tag at the C-terminus, purification using an Ni–NTA column was the choice for affinity chromatography. All purification steps were carried out at 277 K. The protein was purified using a protocol similar to that reported previously (Nasir *et al.*, 2012) on an ÄKTAexplorer chromatographic system (GE Healthcare Life Sciences, USA). The cells were resuspended in buffer *A* (20 m*M* Tris, 50 m*M* NaCl pH 8.5) containing a Complete EDTA-free protease-inhibitor tablet (Roche). The resuspension was then homogenized at 241 MPa by passing it through a cell disrupter twice (Constant Systems Ltd, England). The soluble protein fraction was collected by centrifugation at 10 000*g*. After loading the lysate onto a column pre-equilibrated with buffer *A*, the unbound and nonspecifically bound proteins were washed out with buffer *A*; this was followed by an additional wash with buffer *A* containing 2 *M* NaCl. The column was further washed with 10 m*M* imidazole in buffer *A* and the protein was eluted using buffer *A* with 300 m*M* imidazole. The pooled eluate was then further purified and resolved by gel-filtration chromatography on a Superdex 200 prep-grade 16/60 column pre-equilibrated in buffer *A*. The purity of the protein was analysed on SDS–PAGE and the identity of the protein was confirmed by mass-spectrometric analysis (Technoconcept, India). A 2 l culture yielded approximately 10 mg recombinant HisC.

### 2.3. Crystallization and data collection

A pre-crystallization test was performed at 293 K with a protein concentration varying from 3 to 15 mg ml$^{-1}$ in buffer *A* to determine the appropriate protein concentration for setting up crystallization

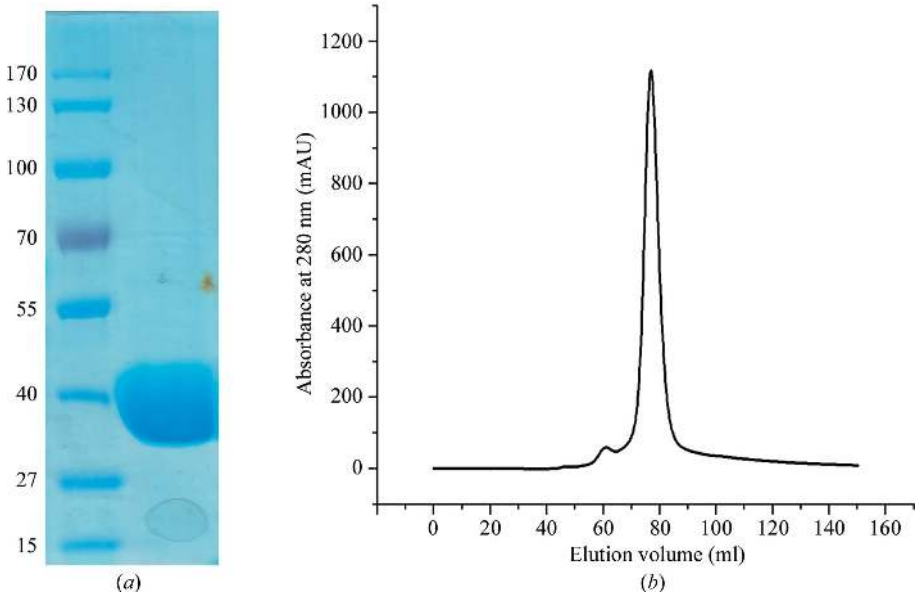



**Figure 1**
Purification profile of HisC. (*a*) Coomassie Brilliant Blue R-250 stained SDS–PAGE of the purified recombinant HisC. Left lane, molecular-mass marker (labelled in kDa); right lane, purified HisC. (*b*) Gel-filtration profile of HisC showing that HisC exists as a dimeric form in solution.

screens. The protein concentration was optimized to 7 mg ml$^{-1}$ and high-throughput crystallization trials using 480 different conditions from Hampton Research and Jena Biosciences were set up using the hanging-drop method at 296 K. Drops with a volume of 1 µl and a 1:1 protein:precipitant ratio were set up using a Mosquito robot (TTP LabTech, England) in 96-well plates and were equilibrated against 100 µl reservoir solution. Microcrystals appeared in condition No. 62 of Index (Hampton Research) after three weeks. Diffraction-quality rod-shaped crystals with a hexagonal cross-section were grown in two weeks in an optimized condition consisting of 0.2 *M* trimethylamine *N*-oxide dehydrate, 0.1 *M* Tris pH 8.5, 30% PEG MME 2000, 20 m*M* EDTA.

A single crystal was mounted on a Cryo-Loop and aligned in a Cu $K\alpha$ X-ray beam generated by an in-house X-ray generator (Rigaku FR-E+ SuperBright microfocus rotating anode). A complete native diffraction data set was collected on an R-AXIS IV$^{++}$ detector at 100 K. The data set was processed with *HKL*-2000 (Otwinowski & Minor, 1997). The data-collection statistics are given in Table 2.

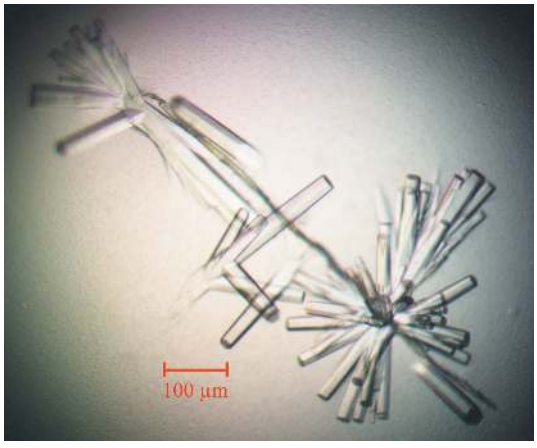



**Figure 2**
Crystals of HisC. The approximate dimensions of the HisC crystals were 150 × 40 × 40 µm. The scale bar is 100 µm in length.

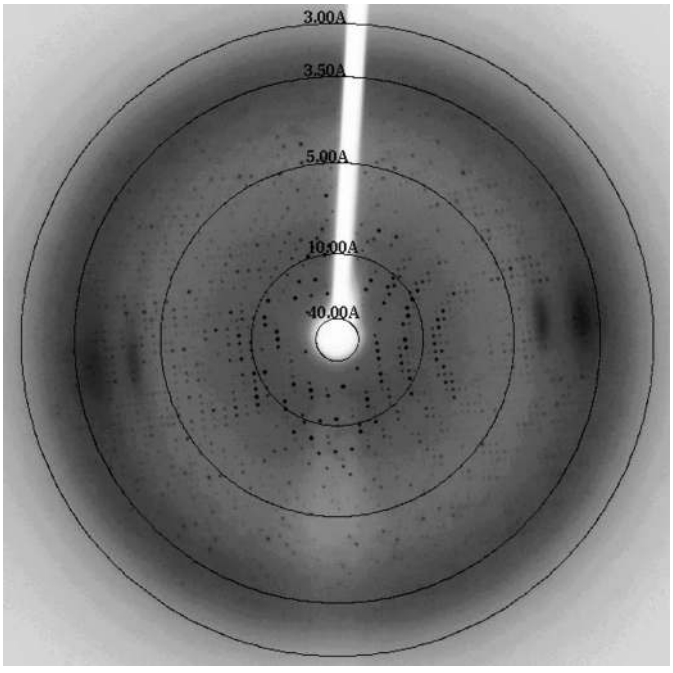



**Figure 3**
A representative diffraction image collected at 1.0° oscillation range from a single HisC crystal. The resolution shells are shown by concentric circles.

**Table 2**
Data-collection statistics.

Values in parentheses are for the highest resolution shell.

| | |
|---|---|
| Space group | $P3_221$ |
| Unit-cell parameters (Å, °) | $a$ = 159.985, $b$ = 159.985, $c$ = 110.221, $\alpha = \beta = 90$, $\gamma = 120$ |
| Matthews coefficient (Å$^3$ Da$^{-1}$) | 4.82, 3.21, 2.41† |
| Solvent content (%) | 74.50, 61.73, 48.97† |
| Temperature (K) | 100 |
| Detector | R-AXIS IV$^{++}$ |
| Wavelength (Å) | 1.5418 |
| Resolution (Å) | 50.0–3.08 (3.19–3.08) |
| Unique reflections | 30506 (2997) |
| Multiplicity | 6.9 (6.8) |
| $\langle I/\sigma(I)\rangle$ | 11.5 (1.9) |
| Completeness (%) | 99.3 (99.1) |
| $R_{\rm merge}$‡ (%) | 18.3 (98.5) |

† Values are given for two, three and four molecules in the crystal asymmetric unit, respectively. ‡ $R_{\rm merge}(I) = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$ for $i$ observations of a given reflection $hkl$. $\langle I(hkl)\rangle$ is the average intensity of the $i$ observations.

## 3. Results and discussion

Milligram quantities of HisC were obtained by overexpression in *M. smegmatis*. Protein of crystallographic grade purity was obtained by Ni–NTA metal-affinity and gel-filtration chromatography (Fig. 1). Crystals suitable for crystallographic analysis were grown by the vapour-diffusion technique using the precipitant 30% polyethylene glycol monomethyl ether 2000 (Fig. 2). The crystals diffracted to 3.08 Å resolution (Fig. 3). The structure of HisC was solved by the molecular-replacement method using the crystal structure of a monomeric molecule of a homologous aminotransferase from *C. glutamicum* (PDB entry 3cq5; Marienhagen *et al.*, 2008), which shares 59% sequence identity with HisC, as the search model. The program *Phaser* (McCoy *et al.*, 2007) from *CCP*4 (Winn *et al.*, 2011) was used to solve the structure and yielded a model comprised of two subunits (a dimer) in the crystal asymmetric unit. The corresponding Matthews coefficient and solvent content are 4.82 Å$^3$ Da$^{-1}$ and 74.50%, respectively (Matthews, 1968). The model was subjected to 50 cycles of rigid-body refinement followed by 100 cycles of positional refinement using the program *REFMAC*5 (Murshudov *et al.*, 2011) from *CCP*4 (Winn *et al.*, 2011). At this stage, the values of $R_{\rm work}$ and $R_{\rm free}$ were 0.435 and 0.494, respectively. Subsequently, replacement of the *C. glutamicum* sequence by the corresponding specific amino acids of *Mtb* HisC, where necessary, was initiated using the program *Coot* (Emsley & Cowtan, 2004). After every round of model building 50 cycles of maximum-likelihood restrained refinement were carried out. The current values of $R_{\rm work}$ and $R_{\rm free}$ are 0.224 and 0.278, respectively. Final rounds of refinement, including incorporation of water molecules, are in progress.

We thank Professor William R. Jacobs of the Department of Microbiology and Immunology and the Howard Hughes Medical Institute, Albert Einstein College of Medicine, Bronx, New York, USA for providing us with the *M. smegmatis* expression system. *Mtb* H37Rv genomic DNA was obtained through the Biodefense and Emerging Infections Research Resources Repository (BEI Resources), NIAID, NIH. BKB received funding from the Indian Council of Medical Research (Reference No. 5/8/5/4/2010-ECD-I) and the National Institute of Immunology (NII), New Delhi, India. The in-house X-ray diffraction facility used for data collection was

established with financial support from the Department of Biotechnology (DBT), Government of India. We acknowledge Ravikant Pal for his help during data collection. NN is a senior research fellow of the Council of Science and Industrial Research, Government of India.

## References

Ahangar, M. S., Khandokar, Y., Nasir, N., Vyas, R. & Biswal, B. K. (2011). *Acta Cryst.* F**67**, 1451–1456.
Alifano, P., Fani, R., Liò, P., Lazcano, A., Bazzicalupo, M., Carlomagno, M. S. & Bruni, C. B. (1996). *Microbiol. Rev.* **60**, 44–69.
Cole, S. T. *et al.* (1998). *Nature (London)*, **393**, 537–544.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Marienhagen, J., Sandalova, T., Sahm, H., Eggeling, L. & Schneider, G. (2008). *Acta Cryst.* D**64**, 675–685.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Nasir, N., Vyas, R., Chugh, C., Ahangar, M. S. & Biswal, B. K. (2012). *Acta Cryst.* F**68**, 32–36.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Sassetti, C. M., Boyd, D. H. & Rubin, E. J. (2003). *Mol. Microbiol.* **48**, 77–84.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.