

**Azusa Seto,<sup>a</sup> Hiroaki Ikushima,<sup>b</sup> Toshiyasu Suzuki,<sup>a</sup> Yusuke Sato,<sup>a</sup> Shuya Fukai,<sup>c</sup> Keiko Yuki,<sup>b</sup> Keiji Miyazawa,<sup>b</sup> Kohei Miyazono,<sup>b</sup> Ryuichiro Ishitani<sup>d</sup> and Osamu Nureki<sup>d,\*</sup>**

<sup>a</sup>Department of Biological Information, Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, 4259 Nagatsuta-cho, Midori-ku, Yokohama-shi, Kanagawa 226-8501, Japan, <sup>b</sup>Department of Molecular Pathology, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan, <sup>c</sup>Structural Biology Laboratory, Life Science Division, Synchrotron Radiation Research Organization, The University of Tokyo, 211 General Research Building, Institute of Molecular and Cellular Biosciences, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-0032, Japan, and <sup>d</sup>Division of Structural Biology, Department of Basic Medical Science, The Institute of Medical Science, The University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan

Correspondence e-mail:  
nureki@ims.u-tokyo.ac.jp

Received 18 September 2008  
Accepted 18 November 2008



© 2009 International Union of Crystallography  
All rights reserved

## Crystallization and preliminary X-ray diffraction analysis of GCIP/HHM transcriptional regulator

GCIP/HHM is a human nuclear protein that is implicated in regulation of cell proliferation. Its primary structure contains helix-loop-helix and leucine-zipper motifs but lacks a DNA-binding basic region. Native and selenomethionine-derivatized (SeMet) crystals of full-length GCIP/HHM were obtained using the hanging-drop vapour-diffusion method. The crystals were greatly improved by adding tris(2-carboxyethyl)phosphine as a reducing reagent and diffracted to 3.5 Å resolution. Preliminary phase calculations using the data set obtained from the SeMet crystal suggested that the crystal belonged to space group  $P3_21$  and contained one molecule per asymmetric unit. Structure determination by the multiple-wavelength anomalous dispersion method using the SeMet crystals is in progress.

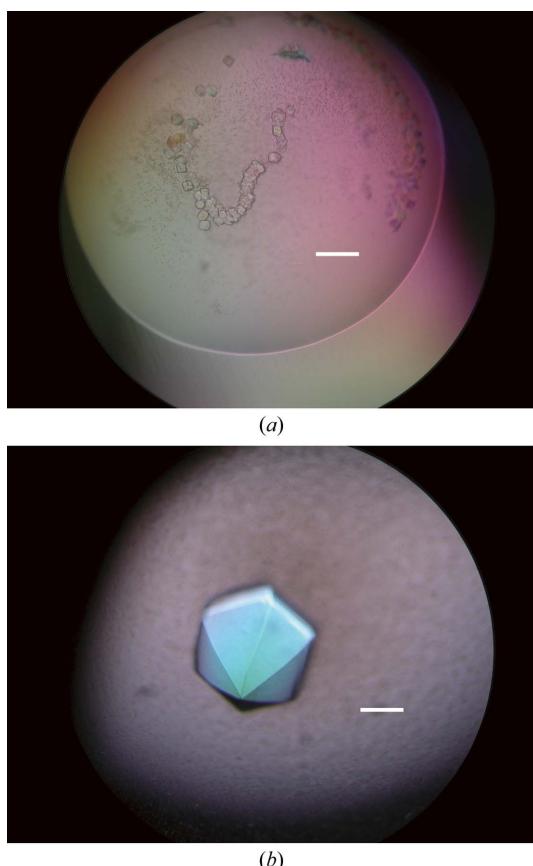
### 1. Introduction

Cell growth and differentiation are complicated processes that are controlled by cellular signalling pathways that execute the intrinsic genetic programme. Basic helix-loop-helix (bHLH) transcriptional factors are key regulators of cell growth, proliferation and differentiation in embryonic and adult tissues. They have been demonstrated to regulate gene expression by binding to an 'E box' in the promoter region of tissue-specific genes through homodimerization and heterodimerization (Blackwell *et al.*, 1990; Kreider *et al.*, 1992; Zebedee & Hara, 2001). Based on the presence or absence of a DNA-binding domain as well as other functional domains, HLH proteins can be classified into five major classes: basic HLH (bHLH) proteins, basic HLH Per-AhR-Arnt-Sim (bHLH-PAS) proteins, basic HLH leucine-zipper (bHLH-LZ) proteins, dominant-negative HLH (dnHLH) proteins and HLH leucine-zipper (HLH-LZ) proteins. The class A bHLH proteins, such as the E2 gene products, are ubiquitously expressed in many tissues, while the class B bHLH proteins, such as MyoD, NeuroD and Hes, exhibit tissue-specific expression and regulate tissue-specific cell growth and differentiation through heterodimerization with the class A bHLH proteins. The bHLH-LZ proteins include the *Myc*-family proteins, which have a leucine-zipper motif at the C-terminus and regulate multiple cellular functions, including proliferation, differentiation, transformation and apoptosis (Blackwell *et al.*, 1990). The dnHLH proteins are Id-family proteins (Id1, Id2, Id3 and Id4) that lack a basic region prior to the HLH domain. Id-family proteins form complexes with bHLH proteins to inhibit their functions in a dominant-negative manner, thus acting as negative regulators of cell differentiation and positive regulators of cell growth (Benezra *et al.*, 1990; Sun *et al.*, 1991; Peppelenbosch *et al.*, 2005; Yokota & Mori, 2002).

GCIP is a 361-residue human HLH leucine-zipper protein (molecular weight of 40 kDa) which contains a putative leucine-zipper motif at its N-terminus, an HLH domain in the centre and an acidic C-terminal region. GCIP was independently identified as a cyclin D1-binding protein (DIP1) or a human homologue of murine

maternal Id-like molecule (a human homologue of MAID or HHM). Like the Id-family proteins, GCIP/HHM lacks a basic DNA-binding region, but it contains leucine-zipper domains as in the *Myc*-family proteins and thus represents a new class of dnHLH proteins. GCIP/HHM appears to be involved in the regulation of liver-cell growth and the progression of hepatocellular carcinomas, although GCIP/HHM can also exert the opposite effect depending on the cellular context. In GCIP-deficient mice, cell proliferation after partial hepatectomy was suppressed compared with that in wild-type mice (Sonnenberg-Riethmacher *et al.*, 2007). In addition, GCIP/HHM expression is upregulated in the early stages of hepatocarcinogenesis and GCIP/HHM promotes S-phase entry in HepG2 hepatocellular carcinoma cells (Terai *et al.*, 2000). These findings suggest positive regulatory roles for GCIP/HHM in liver-cell proliferation. In contrast, GCIP-deficient mice more frequently develop liver tumours (Sonnenberg-Riethmacher *et al.*, 2007) and hepatic overexpression of GCIP/HHM in transgenic mice decreases their susceptibility to chemical hepatocarcinogenesis (Ma *et al.*, 2006). These findings suggest that GCIP/HHM has a tumour-suppressor function in the liver. The mechanism of these dual roles of GCIP/HHM in the control of cell proliferation and tumour progression remains to be elucidated.

Here, we crystallized full-length GCIP/HHM and performed preliminary phasing of the crystal, which suggested that the crystals were promising for structural determination. The crystal structure of GCIP/HHM may clarify the regulatory mechanisms of cell proliferation as well as tumorigenesis by GCIP/HHM, enabling us to design new therapeutic agents to suppress hepatocellular carcinoma as well as other tumours.



**Figure 1**

(a) Native crystals of GCIP/HHM observed in the initial screening. The white scale bar is 100  $\mu\text{m}$  in length. (b) Native crystals of GCIP/HHM under the refined conditions including TCEP as an additive. The white scale bar is 100  $\mu\text{m}$  in length.

**Table 1**

Data-collection statistics of GCIP/HHM crystals.

Values in parentheses are for the last shell.

	Native	Peak
Wavelength ( $\text{\AA}$ )	1.0000	0.97924
Space group	$P3_21$	$P3_21$
Unit-cell parameters ( $\text{\AA}$ , $^\circ$ )	$a = b = 105.42$ , $c = 122.43$ , $\alpha = \beta = 90$ , $\gamma = 120$	$a = b = 106.26$ , $c = 118.36$ , $\alpha = \beta = 90$ , $\gamma = 120$
Resolution ( $\text{\AA}$ )	50–3.5 (3.63–3.50)	50–4.0 (4.14–4.00)
Redundancy	9.6 (3.6)	6.7 (4.7)
Unique reflections	10281 (972)	6735 (624)
Completeness (%)	98.2 (94.6)	98.4 (95.9)
$I/\sigma(I)$	37.1 (2.1)	34.5 (3.0)
$R_{\text{merge}}^{\dagger}$	0.073 (0.339)	0.088 (0.318)

†  $R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$ , where  $I_i(hkl)$  is the observed intensity and  $\langle I(hkl) \rangle$  is the average intensity over symmetry-equivalent measurements.

## 2. Materials and methods

### 2.1. Cloning, expression and purification

The gene encoding GCIP/HHM (40 kDa) was cloned into the pGEX-6P-1 vector (GE Healthcare). Native proteins were over-expressed in *Escherichia coli* Rosetta (DE3) strain cells (Novagen), which were grown in LB medium containing ampicillin (50  $\mu\text{g ml}^{-1}$ ), induced at an absorbance at 600 nm ( $A_{600}$ ) of  $\sim 0.8$  with 1.0 mM isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) and cultivated for a further 20 h at 293 K. Selenomethionine-substituted (SeMet) protein was expressed using similar procedures in methionine-auxotrophic B834 (DE3) cells, which were grown in Core medium (Wako) with 10  $\mu\text{g ml}^{-1}$  L-selenomethionine (Nacalai Tesque).

Cells were harvested by centrifugation at 7000 rev  $\text{min}^{-1}$  at 277 K, resuspended in 50 mM Tris buffer pH 7.5 containing 200 mM NaCl and 5 mM  $\beta$ -mercaptoethanol ( $\beta$ ME) and disrupted by sonication. The crude cell extract was centrifuged for 20 min at 14 000 rev  $\text{min}^{-1}$  at 277 K to remove cell debris. The supernatant was applied onto a glutathione-Sepharose FF column (GE Healthcare). The GST tag was cleaved by digestion with 10 units  $\text{ml}^{-1}$  PreScission Protease (GE Healthcare) for 20 h at 297 K on the column. The cleaved proteins in the flowthrough fraction were collected and were further purified by chromatography on a MonoQ anion ion-exchange column (GE Healthcare) and a HiLoad 16/60 Superdex 200 pg gel-filtration column (GE Healthcare). The purity of the protein was monitored by 12.0% SDS-PAGE followed by staining with Coomassie Brilliant Blue. The purified protein was concentrated to 10 mg  $\text{ml}^{-1}$  using an Amicon Ultra 10K filter (Amicon). All of the purification procedures described above were conducted at 277 K. All of the purification buffers included 5 mM  $\beta$ ME.

### 2.2. Crystallization

Initial screening for crystallization conditions was performed using several commercial screening kits from Hampton Research, JB Screen kits (Jena Bioscience) and MemSys and MemStart kits (Molecular Dimensions). A Hydra II Plus One (Matrix Technologies) crystallization robot was used with the sitting-drop vapour-diffusion method at 293 K in IntelliPlates (Art Robbins Instruments). In the initial crystallization screen, 0.2  $\mu\text{l}$  protein solution (10 mg  $\text{ml}^{-1}$  in 10 mM Tris-HCl buffer containing 50 mM NaCl and 5 mM  $\beta$ -mercaptoethanol) was mixed with 0.2  $\mu\text{l}$  mother liquor. Native crystals of GCIP/HHM were first obtained in 2 d from condition No. 47 of PEG/Ion Screen (Hampton Research). After optimizing the conditions, crystals were grown using the hanging-drop vapour

diffusion method by manually mixing 2  $\mu$ l protein solution with an equal volume of reservoir solution containing 8–10% (v/v) polyethylene glycol (PEG) 3350, 200 mM potassium citrate and 10 mM tris(2-carboxyethyl)phosphine (TCEP). The hanging drops were equilibrated against 500  $\mu$ l reservoir solution and the crystals grew in 2 d. SeMet crystals were obtained under similar conditions to those used for the native protein.

### 2.3. Data collection and processing

All diffraction data sets were collected on beamline BL41XU at SPring-8 (Hyogo, Japan). Before measurement, the crystals were soaked for a few seconds in a cryoprotection solution containing 35% (v/v) xylitol, 200 mM potassium citrate, 13% (v/v) PEG 3350 and 10 mM TCEP and were flash-cooled to 100 K in a cold nitrogen-gas stream. A total of 180 diffraction images were collected with 1° rotation per image using an ADSC Quantum 315 detector for the native, SeMet peak and SeMet edge data sets. The resulting data sets were processed with the program HKL-2000 (Otwinowski & Minor, 1997) and the statistics are summarized in Table 1.

## 3. Results and discussion

In the GST-affinity chromatography step, we first eluted the GST-fused protein with glutathione and then cleaved the GST tag with PreScission protease (GE Healthcare; an extra Gly-Pro sequence remained at the N-terminus after this cleavage). After the proteolytic digestion we repurified the GCIP/HHM with a glutathione Sepharose FF column to exclude the contaminating GST, but not all of the cleaved GST proteins were retained on the column and some were included in the flowthrough fraction containing the GCIP/HHM protein. However, we obtained crystals that diffracted very weakly to 8.0  $\text{\AA}$  resolution using a synchrotron X-ray source.

We then improved the purification procedure for the GCIP/HHM protein by on-column digestion of the GST tag. All the cleaved GST

tag was retained on the glutathione Sepharose FF resin, resulting in successful purification of the GCIP/HHM protein with high quality, which further improved the reproducibility of the crystals. Finally, 5.8 mg purified GCIP/HHM protein was obtained from 4 l of *E. coli* culture.

In the initial crystallization screen, multiple GCIP/HHM crystals with dimensions of 50  $\times$  50  $\times$  50  $\mu\text{m}$  were obtained (Fig. 1a). After optimization of the crystallization conditions, bipyramidal crystals with dimensions of 300  $\times$  300  $\times$  300  $\mu\text{m}$  were obtained (Fig. 1b). The addition of TCEP drastically improved the quality of the crystals. However, the crystals were extremely sensitive to the cryoprotection procedure. After optimization of the cryoprotection conditions, we succeeded in collecting 3.5  $\text{\AA}$  resolution data from the GCIP/HHM crystals using 35% (v/v) xylitol as the cryoprotectant (Table 1).

A diffraction image is shown in Fig. 2. Crystals of SeMet GCIP/HHM were obtained in a similar manner as the native crystals and diffracted X-rays to 4.0  $\text{\AA}$  resolution (Table 1). The two crystals belonged to the same space group,  $P3_121$  or  $P3_221$ , with slightly different unit-cell parameters, as shown in Table 1. The calculated Matthews coefficients ( $V_M$ ) of 4.8 and 2.4  $\text{\AA}^3 \text{Da}^{-1}$  suggested the presence of one and two molecules in the asymmetric unit, with solvent contents of about 75% and 49%, respectively (Matthews, 1968).

*SHELXC* and *SHELXD* (Sheldrick, 2008) were used to identify the Se sites using the SAD data set collected at the peak wavelength. Examinations of the anomalous signal-to-noise ratio and of the correlation coefficients between the signed anomalous differences of the peak data set as a function of resolution (not shown) suggested the presence of a useful anomalous signal for a selenium-site search to approximately 6.0  $\text{\AA}$ . As the numbers of molecules and selenium sites per asymmetric unit were initially unknown, we set up many parallel *SHELXD* processes searching for variable numbers of selenium sites ranging from 11 to 22. The result clearly revealed a consistent and significant occupancy drop after 11 selenium sites, indicating the existence of one molecule per asymmetric unit. Given

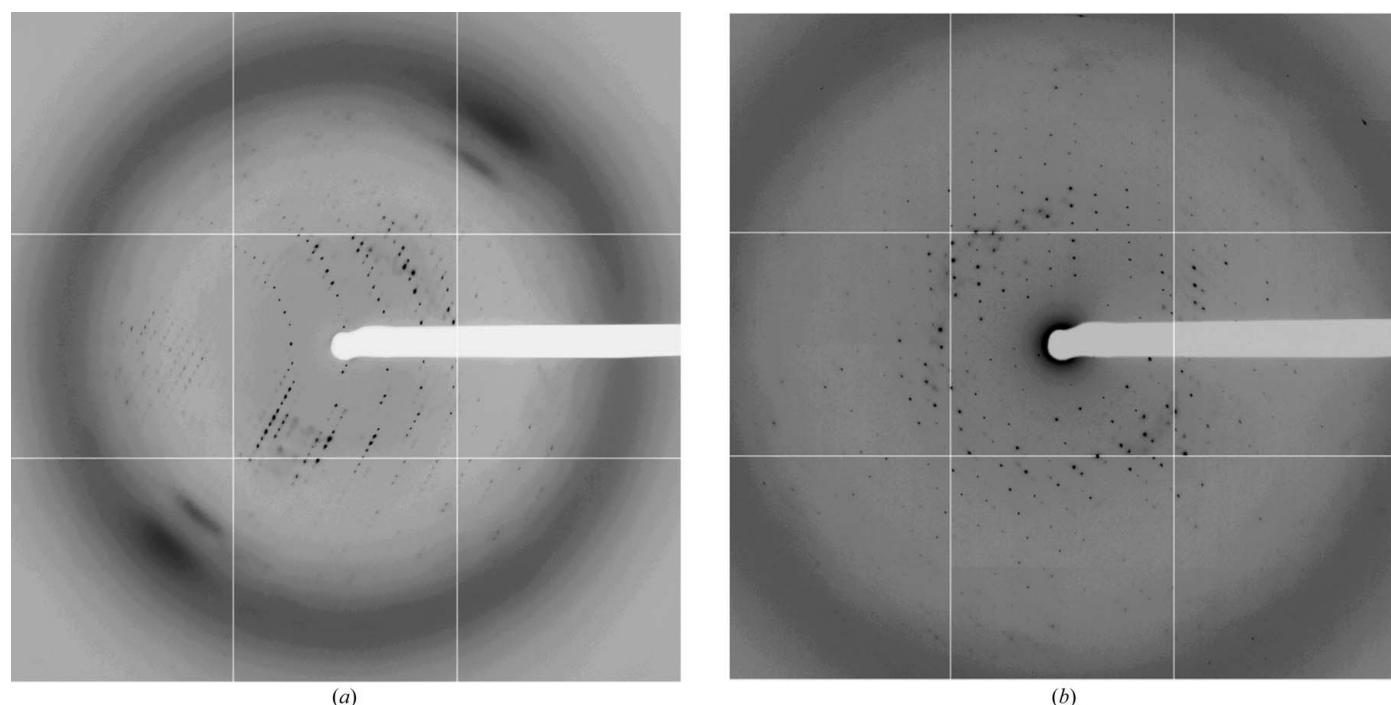


Figure 2

(a) Diffraction pattern of a native crystal of GCIP/HHM. (b) Diffraction pattern of an SeMet crystal of GCIP/HHM.

the fragility and the medium resolution of the crystal, this high solvent content (75%) seems reasonable.

Phase calculation and refinement with the programs *SHARP* (de La Fortelle & Bricogne, 1997) and *SOLOMON* (Abrahams & Leslie, 1996) were performed for both space groups (*i.e.*  $P3_121$  and  $P3_221$ ) using the identified Se sites. Preliminary phase information obtained in space group  $P3_221$  showed a clear solvent–protein boundary and the existence of a characteristic ten-helix bundle per molecule, which is consistent with the results of the previous secondary-structure prediction. Further improvement of the crystal resolution and model building are now in progress.

We thank the beamline staff at BL41XU of SPring-8 for technical help during data collection. This work was supported by a grant for the National Project on Protein Structural and Functional Analyses from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) to ON and by grants from MEXT to ON.

## References

Abrahams, J. P. & Leslie, A. G. W. (1996). *Acta Cryst. D* **52**, 30–42.

Benezra, R., Davis, R. L., Lockshon, D., Turner, D. L. & Weintraub, H. (1990). *Cell*, **61**, 49–59.

Blackwell, T. K., Kretzner, L., Blackwood, E. M., Eisenman, R. N. & Weintraub, H. (1990). *Science*, **250**, 1149–1151.

Kreider, B. L., Benezra, R., Rovera, G. & Kadesch, T. (1992). *Science*, **255**, 1700–1702.

La Fortelle, E. de & Bricogne, G. (1997). *Methods Enzymol.* **276**, 472–494.

Ma, W., Xia, X., Stafford, L. J., Yu, C., Wang, F., LeSage, G. & Liu, M. (2006). *Oncogene*, **25**, 4207–4216.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Otinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Peppelenbosch, M. P., Qiu, R. G., deVries-Smits, A. M., Tertoolen, L. G., de Laat, S. W., Perk, J., Iavarone, A. & Benezra, R. (2005). *Nature Rev. Cancer*, **5**, 603–614.

Sheldrick, G. M. (2008). *Acta Cryst. A* **64**, 112–122.

Sonnenberg-Riethmacher, E., Wüstefeld, T., Miehe, M., Trautwein, C. & Riethmacher, D. (2007). *Hepatology*, **45**, 404–411.

Sun, X. H., Copeland, N. G., Jenkins, N. A. & Baltimore, D. (1991). *Mol. Cell. Biol.* **11**, 5603–5611.

Takami, T., Terai, S., Yokoyama, Y., Tanimoto, H., Tajima, K., Uchida, K., Yamasaki, T., Sakaida, I., Nishina, H., Thorgeirsson, S. S. & Okita, K. (2005). *Gastroenterology*, **128**, 1369–1380.

Terai, S., Aoki, K., Ashida, K. & Thorgeirsson, S. S. (2000). *Hepatology*, **32**, 357–366.

Yokota, Y. & Mori, S. (2002). *J. Cell. Physiol.* **190**, 21–28.

Zebedee, Z. & Hara, E. (2001). *Oncogene*, **20**, 8317–8325.