Crystal structure of GCN5 PCAF N-terminal domain reveals atypical ubiquitin ligase structure

Received for publication, March 12, 2020, and in revised form, July 22, 2020 Published, Papers in Press, August 19, 2020, DOI 10.1074/jbc.RA120.013431

Sachiko Toma-Fukai[1,‡], Ryota Hibi[1,‡], Takao Naganuma[2], Mashito Sakai[2], Shinya Saijo[3], Nobutaka Shimizu[3], Michihiro Matsumoto[2], and Toshiyuki Shimizu[1,*]

From the [1]Graduate School of Pharmaceutical Sciences, The University of Tokyo, Bunkyo-ku, Tokyo, Japan, the [2]Department of Molecular Metabolic Regulation, Diabetes Research Center, Research Institute, National Center for Global Health and Medicine, Shinjyuku-ku, Tokyo, Japan, and the [3]Photon Factory, Institute of Materials Structure Science, High Energy Accelerator Research Organization (KEK), Tsukuba, Ibaraki, Japan

Edited by Joseph M. Jez

General control nonderepressible 5 (GCN5, also known as Kat2a) and p300/CBP-associated factor (PCAF, also known as Kat2b) are two homologous acetyltransferases. Both proteins share similar domain architecture consisting of a PCAF N-terminal (PCAF_N) domain, acetyltransferase domain, and a bromodomain. PCAF also acts as a ubiquitin E3 ligase whose activity is attributable to the PCAF_N domain, but its structural aspects are largely unknown. Here, we demonstrated that GCN5 exhibited ubiquitination activity in a similar manner to PCAF and its activity was supported by the ubiquitin-conjugating enzyme UbcH5. Moreover, we determined the crystal structure of the PCAF_N domain at 1.8 Å resolution and found that PCAF_N domain folds into a helical structure with a characteristic binuclear zinc region, which was not predicted from sequence analyses. The zinc region is distinct from known E3 ligase structures, suggesting this region may form a new class of E3 ligase. Our biochemical and structural study provides new insight into not only the functional significance of GCN5 but also into ubiquitin biology.

Post-translational modification of proteins regulates many biological processes. Acetyltransferases transfer acetyl groups to lysine residues on target proteins and are one of the major types of post-translational enzymes. General control nonderepressible 5 (GCN5) is one of the best characterized histone acetyltransferases that promote transcriptional activity (1, 2). GCN5 is the enzyme subunit of the SAGA (Spt-Ada-Gcn5-acetyltransferase) complex, whose architecture was visualized by cryo-EM analyses (3, 4), and modifies multiple lysine residues on histone H3 *in vitro* (5). Recently, it was reported that GCN5 is phosphorylated by protein kinase A in a manner dependent on the transcriptional coregulator Cbp/p300-interacting transactivator 2 (CITED2) (6), thereby increasing its acetyltransferase activity for histone and attenuating that for peroxisome proliferator-activated receptor-γ coactivator-1α (PGC-1α). This suggests that GCN5 functions in multiple situations.

Metazoans possess two GCN5 isoforms that arise from alternative splicing (7). The lower molecular weight isoform (isoform 2) is similar in size and function to yeast GCN5, consisting of an acetyltransferase (AT) domain and a bromodomain at the N and C termini, respectively (2, 8) (Fig. 1*A*). The higher molecular weight isoform (isoform 1) contains an N-terminal extension that has high similarity to the N-terminal domain of PCAF (8), termed PCAF_N domain, which is conserved only among vertebrate (Fig. 1*A*) (9). According to the Pfam database (10), there are 695 eukaryotic proteins harboring the PCAF_N domain, categorized into 18 architectures.

Ubiquitination is also a post-translational modification that targets lysine residues. This modification regulates many cellular processes, including cell division and immune responses, among others. Ubiquitination is achieved by the sequential reaction of ubiquitin-activating enzyme (E1), ubiquitin-conjugating enzyme (E2), and ubiquitin ligase (E3), which is responsible for the ligation of ubiquitin onto a substrate in conjunction with the E2 (11). The human genome encodes two E1s, ~38 E2s, and more than 600 E3s (12–14). E3s plays a pivotal role in selecting substrates and to date, three classes of E3 have been identified (15): RING (really interesting new gene), HECT (homologous to E6AP C terminus), and RBR (RING-between-RING).

It has been reported that in addition to acetyltransferase activity, PCAF also harbors ubiquitination activity (16) and it has been demonstrated that PCAF acts as an E3 ligase targeting human Hdm2, human Gli1, and human CIITA and promotes self-ubiquitination (16–18). PCAF_N is identified as a domain containing ubiquitin E3 ligase activity and the longer isoform of GCN5 possesses PCAF_N domain. However, whether GCN5 functions as an E3 enzyme remains to be determined.

Here, we demonstrated that GCN5 exhibits ubiquitination activity in a similar manner to PCAF. We also performed a structural based study focusing on PCAF_N domain to elucidate any functions of GCN5 as an E3 enzyme. Our findings provide new avenues for both the functional study of GCN5 as well as that of ubiquitin biology.

Results

GCN5 has ubiquitin ligase activity

GCN5 and PCAF are highly homologous with ~75% sequence identity (Fig. 1*A*) and both share the same domain architecture

This article contains supporting information.
‡These authors contributed equally to this work.
* For correspondence: Toshiyuki Shimizu, shimizu@mol.f.u-tokyo.ac.jp.
Present address for Sachiko Toma-Fukai: Graduate School of Science and Technology, Nara Institute of Science and Technology, Takayama-cho, Ikoma, Nara, Japan.

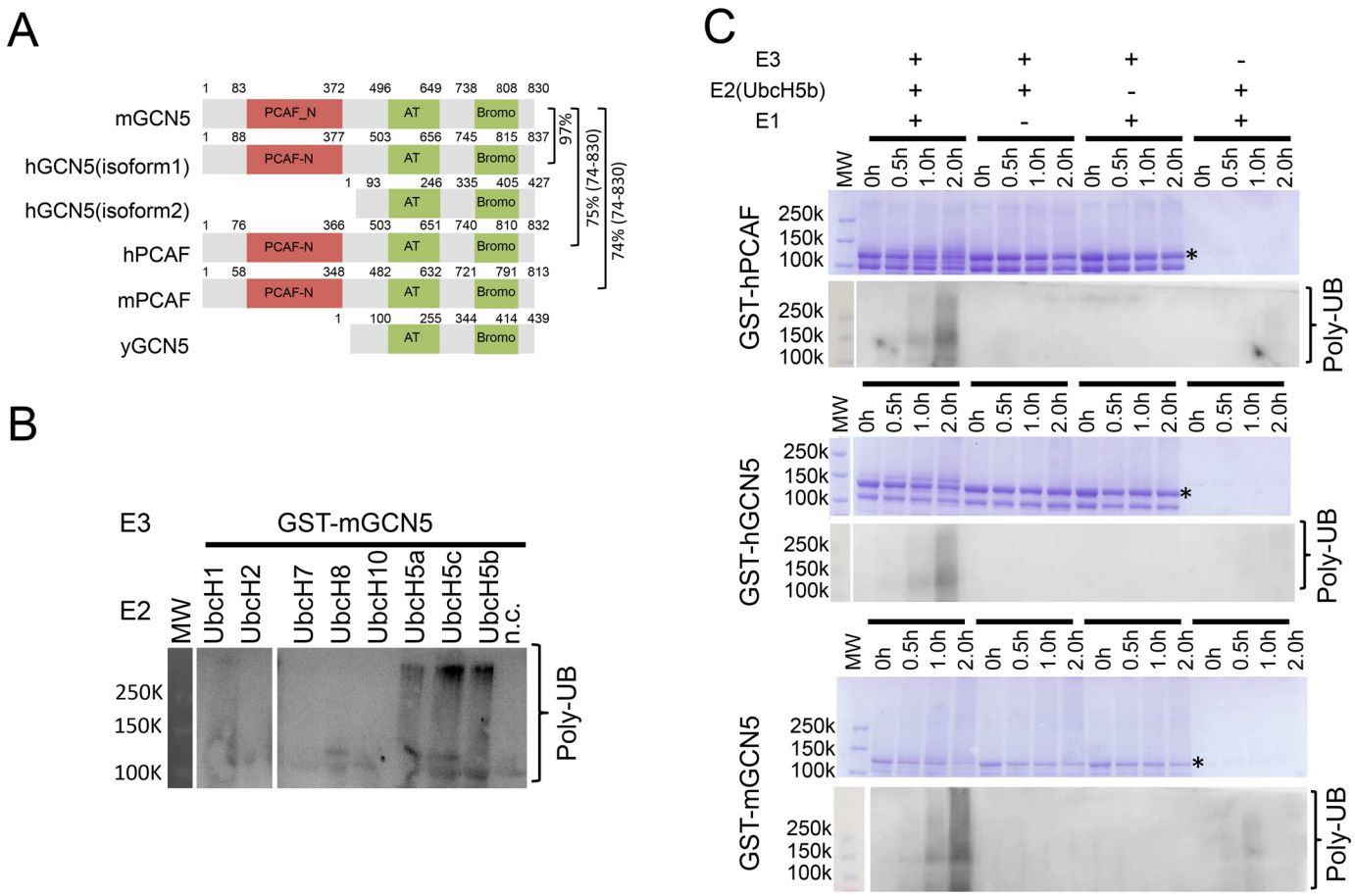

Figure 1. Ubiquitin ligase activities of GCN5 and PCAF. *A,* domain architecture of GCN5 and PCAF. The identity between the different molecules is indicated in % on the *right of the horizontal lines*, representing the pairwise comparisons. The value in *parentheses* indicates the amino acid sequence region employed to calculate sequence identity. PCAF_N domain is colored in *pink*, and acetyltransferase domain (*AT*) and bromo domain (*Bromo*) are colored in *green*. The *numbers over the boxes* indicate amino acid positions. *B,* E2 screening results for GST-mGCN5. Polyubiquitination chains were detected by Western blotting analysis using an anti-ubiquitin (*anti-Ub*) antibody. The negative control (*n.c.*) means that there was no E2 enzyme in the reaction solution. The reaction time was 1 h. *C,* ubiquitination assays of hPCAF, hGCN5, and mGCN5. The GST-tagged protein indicated by an *asterisk* were used as E3 enzymes. UbcH5b was used as E2 enzyme. The experiments were performed in the presence (+) or absence (−) of E1, E2, and/or E3 samples. The condition of each reaction is described on each lane. Results of both SDS-PAGE analysis (*upper*) and Western bot analysis using an anti-ubiquitin (anti-Ub) antibody (*lower*) were shown. The reaction time is indicated on each lane.

(Fig. 1*A*, Fig. S1). This strongly suggested that GCN5 was able to act as an E3 enzyme in a similar manner to that of PCAF. E3 capacity for ubiquitination is utilized as a means of assessing their potential to function with E2. Therefore, we evaluated whether GCN5 demonstrated autoubiquitination activity. Full-length human PCAF (hPCAF), full-length mouse PCAF (mPCAF), hGCN5 with 80 residues truncated at the N-terminal (hGCN5), and full-length mGCN5 (mGCN5) were each expressed in *Escherichia coli*. As a previous study used UbcH5b as an E2 enzyme in ubiquitination assays of PCAF (16), we further examined several E2 enzymes with a variety of identities ranging from 31 (UbcH2) to 97% (UbcH5c). Eight E2s (UbcH1, UbcH2, UbcH5a, UbcH5b, UbcH5c, UbcH7, UbcH8, and UbcH10) were tested to assess ubiquitination activity, UbcH5a and UbcH5c were likely to work as E2 enzymes for PCAF and GCN5, probably because UbcH5a, UbcH5b, and UbcH5c are highly homologous proteins. Next, we assessed ubiquitination activity using UbcH5b more precisely. Polyubiquitinated bands were increased as time passed for both GCN5 and PCAF, whereas the bands of purified proteins were decreased, demonstrating that GCN5 as well as PCAF exhibited E3 ligase activity (Fig. 1*C*). The result also revealed that UbcH5b was an effective E2 for PCAF and GCN5 (Fig. 1*C*).

Structure of mGCN5 PCAF_N domain

Linares *et al.* (16) identified two regions critical for PCAF autoubiquitination with amino acid (aa) residues 121–242 corresponding to the active ubiquitination domain of hPCAF and residues 350–445 corresponding to a motif or residues important for autoubiquitination. PCAF_N domain included residues 121–242, but based on sequence analysis could not be categorized with other known E3s. Although many crystal structures of ATs and bromodomains have previously been determined and are deposited in the Protein Data Bank (19, 20), no PCAF_N domain structure has been reported. We crystallized mGCN5 (aa 67–378) and determined its crystal structure at 1.8 Å resolution. Residues 67–82 and 125–133 were disordered and unable to be modeled. Pfam database

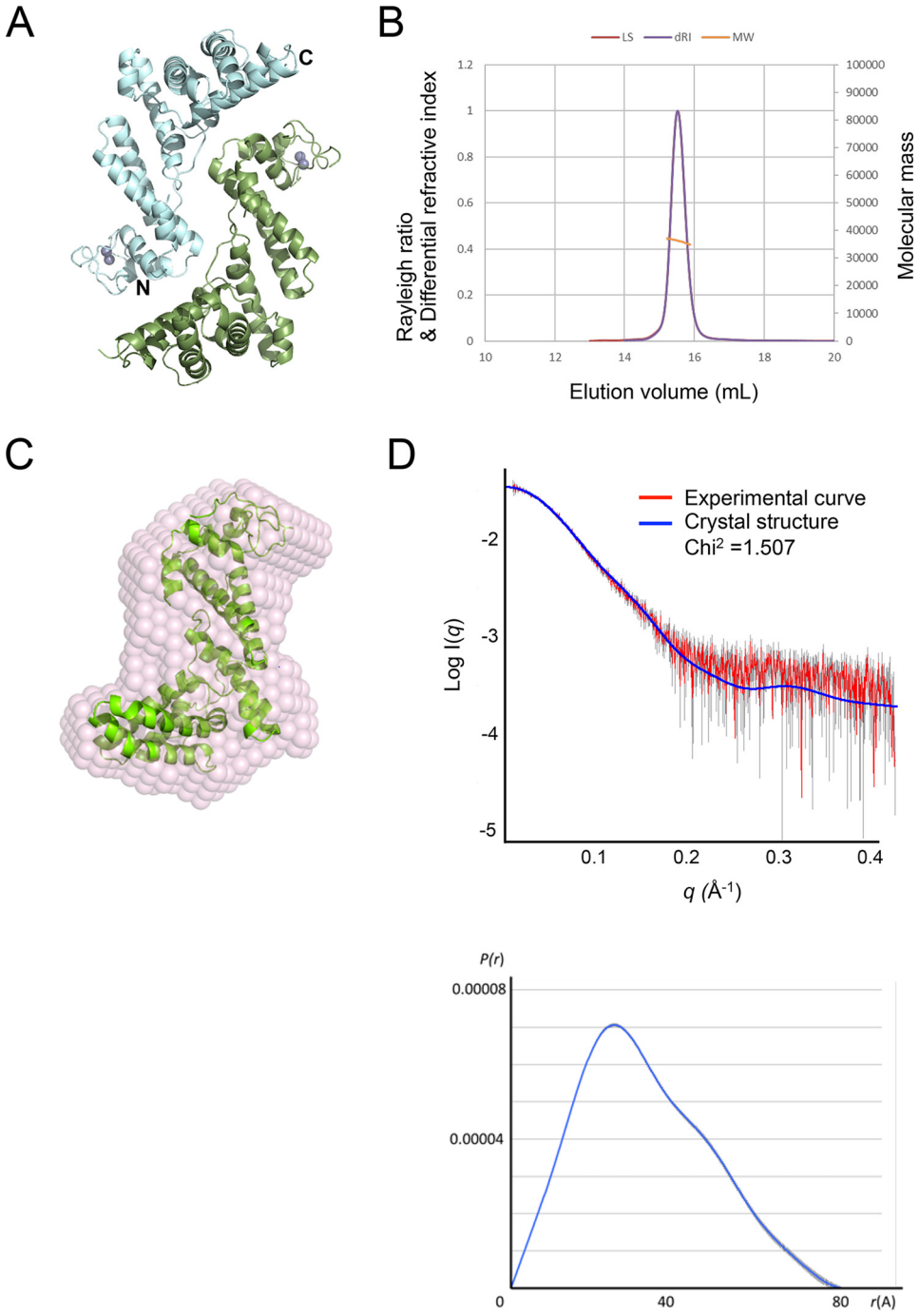

Figure 2. Crystal and solution structures of mGCN5 PCAF_N domain. *A*, crystal structure of mGCN5 PCAF_N domain. There are two molecules in an asymmetric unit. Two molecules are shown in ribbon diagram and colored in *green* or *blue*. *B,* result of SEC-MALS analysis. Rayleigh ratio (*LS*), differential refractive index (*dRI*), and molecular mass (*MW*) are indicated with *red*, *purple*, and *orange lines*, respectively. *C*, SAXS dummy atom model. Crystal structure of PCAF_N is superposed onto dummy atom model. Dummy atom model of PCAF_N is colored in *pink* and its crystal structure is drawn in ribbon diagram. *D,* scattering curves and *P*(*r*) function. The experimental curve and theoretical curve derived from crystal structure are shown in *red* and *blue lines*, respectively. The calculated χ^2-value was indicated.

defined mGCN5(81-331) as PCAF_N domain, but our crystal structure revealed mGCN5(83-378) folds into a single domain. Therefore, this region can be redefined as PCAF_N. There were two PCAF_N domain molecules in an asymmetric unit (Fig. 2*A*) that were essentially the same with a root mean square deviation (r.m.s. deviation) of 0.3 Å. To determine the solution structure and oligomerization state in solution for PCAF_N domain, we performed size exclusion chromatography (SEC)-small angle X-ray scattering (SAXS)/multiangle light scattering (MALS) analyses. The molecular weight derived from SEC-MALS analysis was ~37,000, which coincided with the theoretical molecular weight of 35,457.8 (Fig. 2*B*), demonstrating that mGCN5 PCAF_N domain existed as a monomer in solution. Consistent with the SEC-MALS analysis, SEC-SAXS solution structure fitted well with the monomer crystal structure (Fig. 2, *C* and *D*).

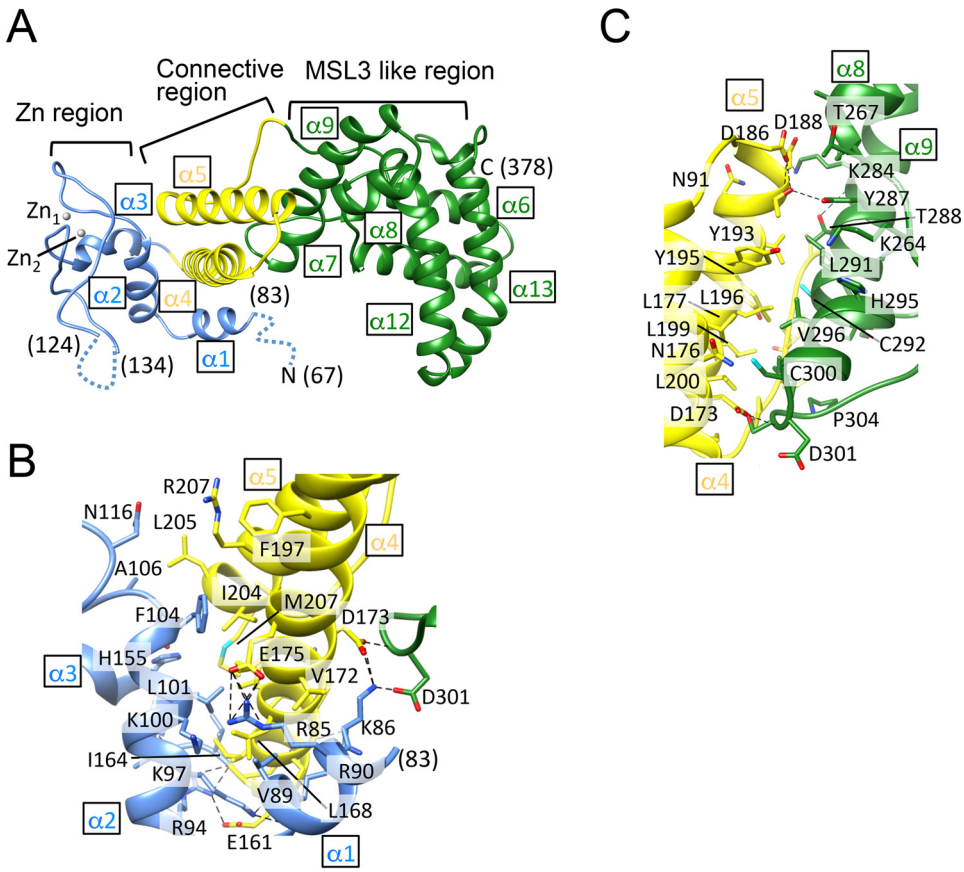

Figure 3. Crystal structure of mGCN5 PCAF_N domain. Overall structure of mGCN5 PCAF_N domain. The zinc region, connective region, and MSL3-like region are colored in *blue*, *yellow*, and *green*, respectively. Two zinc ions are drawn using a sphere model and colored *gray*. Disorder regions are indicated by the *dotted line*. *B,* close-up view of the interface between zinc and connective regions. Hydrogen bonds and electrostatic interactions are shown as *dashed lines*. *C,* close up view of the interface between connective and MSL3-like regions. Hydrogen bonds are shown as *dashed lines*.

PCAF_N domain consists of three regions

The mGCN5 PCAF_N domain folds into a compact structure comprising three regions. The N-terminal region (aa 83–159) forms an unexpected binuclear zinc region described below in detail (Fig. 3, Fig. S1). The second region connects the N-terminal and C-terminal regions with an antiparallel coiled-coil structure associated with a helix (α1) in the N-terminal region, forming a bundle structure. This second region was termed a connective region. The C-terminal region (aa 216–372) folds into an α-helix–rich structure. Analysis using the Dali server (21) revealed that this region shares structural homology to the MORF4-related gene (MRG) domain of male-specific lethal 3 (MSL3; PDB code 2y0n) (22). The Cα atoms from residues 219 to 367 of the MRG domain superposed well onto those of the C-terminal region of the PCAF_N domain (aa 216–372) with a Z-score of 8.4 and r.m.s. deviation of 3.3 Å according to the Dali server (Fig. S2*A*); however, its sequence identity is low, being ~22% (Fig. S2*B*). This C-terminal region was termed a MSL3-like region. Three regions are packed with hydrogen bonds and hydrophobic interactions (Fig. 3, *B* and *C*). The zinc region interacts with the connective region and has a contact area of 814 Å^2. The MSL3-like region also interacts with the connective region and has a contact area of 711 Å^2.

PCAF_N domain has an atypical zinc domain

Notably, two clear and strong electron densities in the N-terminal region were observed between the α2 and α3 helices, indicating that metal ions were bound. To identify the specific metal ions, we performed the following experiments. First, XAFS analysis was performed to evaluate whether the strong electron densities were zinc ions. Weak but clear absorption edges caused by zinc ions were detected from PCAF_N domain crystal. Next, we collected two datasets with a higher energy wavelength (1.28310 Å) and lower energy wavelength (1.28410 Å) relative to the K absorption edge of zinc (1.2837 Å) and generated an anomalous Fourier map. Two strong densities were clearly observed in the higher energy data set but the densities were completely absent in the lower energy data set (Fig. 4*A*). This indicated that the two densities were zinc ions and revealed that the PCAF_N domain had a zinc region. This structure was not predicted from the amino acid sequence. Moreover, a Dali search using this domain identified no structures homologous to the well-known RING E3 ligase.

The zinc region has a binuclear zinc-coordination structure ($Zn_2Cys_5His_2$). The two zinc ions coordinate with seven residues (Cys-107, Cys-113, Cys-115, Cys-142, Cys-145, His-147, and His-151). The sulfur atom of Cys-145 is used to coordinate both zinc ions (Fig. 4, *B* and 4*D*, Fig. S1). All the coordinating residues of the zinc ions are conserved between GCN5 and

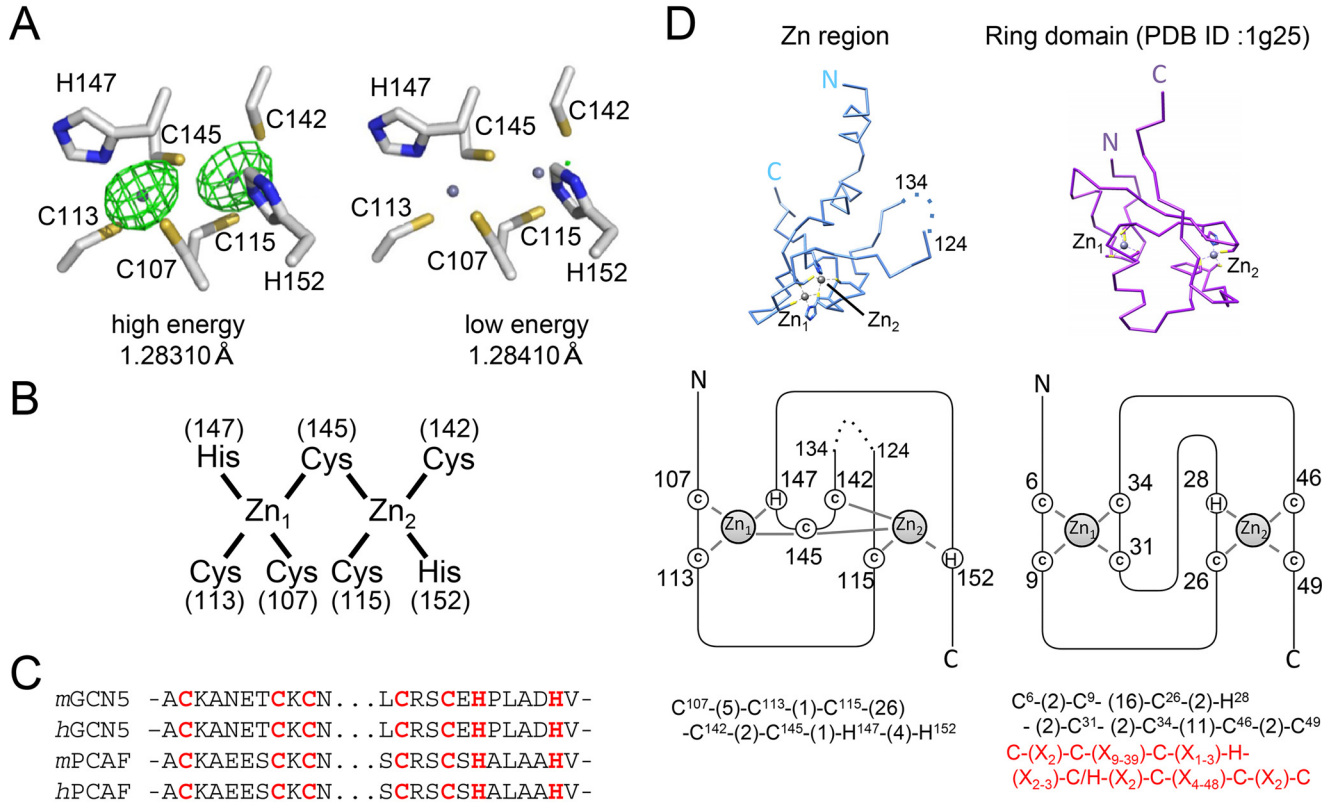

Figure 4. Zinc coordination in the zinc region and comparison of the zinc region and RING domain. *A,* structures around zinc-binding sites of mGCN5 PCAF_N domain. Anomalous Fourier maps (contoured at 3σ) colored *green* are shown using the diffraction data collected with high energy (*left*) and low energy (*right*) X-rays. Coordination residues are drawn using a ball-and-stick model. Zinc ions are shown in the *gray sphere*. *B,* schematic diagram of zinc coordination. *C,* amino acid sequence alignment of zinc-coordinated residues in GCN5 and PCAF. The residues that coordinated with the zinc atom are indicated in *red*. *D,* Cα-trace models and topological diagram of the zinc region of mGCN5 PCAF_N (*blue*) (*left*) and typical RING domain (sky *blue*) (*right*). *C* and *H* denote the zinc-coordinating Cys and His, respectively. The zinc region sequence of mGCN5 PCAF_N domain, and the RING domain sequence are shown below. The consensus sequence of the RING domain are shown in *red*.

PCAF (Fig. 4*C*). The zinc region of PCAF_N domain is unique in its zinc coordination pattern and ternary structure compared with the RING domain. The RING domains are typically characterized with the RING motif bound to two ions of zinc using a cross-brace arrangement (Fig. 4*D*). In this arrangement, the first and third pairs of zinc ligands share Zn1 and the second and fourth pairs of zinc ligands share Zn2. PCAF_N sequence barely aligns with the RING domain due to the Cys/His spacing and because the two domains coordinate with the zinc ions in a different manner (Fig. 4*D*). The zinc region of PCAF_N domain utilizes a cross-brace arrangement, but the ternary structure of PCAF_N domain is distinct from that of the RING domain. Thereby, sequence and structural similarity search programs failed to detect the motif.

PCAF_N domain has ubiquitin ligase activity

PCAF(1-241) is reported to have ubiquitin ligase activity, but this region insufficiently covers PCAF_N domain, lacking most of MSL region, according to the crystal structure. We conducted ubiquitin ligase assay using PCAF_N domain. PCAF_N domain clearly exhibited ligase activity (Fig. 5*A*). To further evaluate whether or not the zinc region or the zinc coordination is required for E3 ligase activity, we tried to prepare the zinc region alone or Ala mutant of mGCN5 zinc-coordination residues, but we failed to obtain the protein probably due to the instability. Alternatively, we prepared the deletion mutant lacking the zinc region (GST-GCN5(dZn)). This mutant exhibited no ligase activity (Fig. 5*B*).

Binding analysis used purified His-tagged mGCN5(67-378) and His-tagged hUbcH5b (Fig. S3). His-tagged UbcH5b alone was eluted around 17 ml. When the mixture of His-tagged mGCN5(67-378) and His-tagged UbcH5b was applied, two proteins were eluted at the same volume and His-tagged UbcH5b was eluted earlier (Fig. S3), suggesting that both proteins were bound.

Discussion

This is the first report to describe a ligase activity of PCAF_N domain of GCN5 and the first crystal structure of PCAF_N domain among PCAF_N family proteins. Unexpectedly, PCAF_N domain harbored a zinc region that adopted a cross-brace motif, reminiscent of RING E3 ligase. RING E3s are highly diverse in oligomerization state. Some RING domains are active as monomers, whereas others are active as oligomers, or multisubunit assembles. GCN5 is a multidomain protein and each domain (PCAF_N, AT (20), and bromodomain (19)) is solved as a monomer. Moreover, the PCAF_N domain, which exists as a monomer in solution, exhibited ligase activity. Taking it into account, monomeric GCN5 could act as E3 ligase. All E3 ligases harbor an E2-ubiquitin–binding domain. In E2-

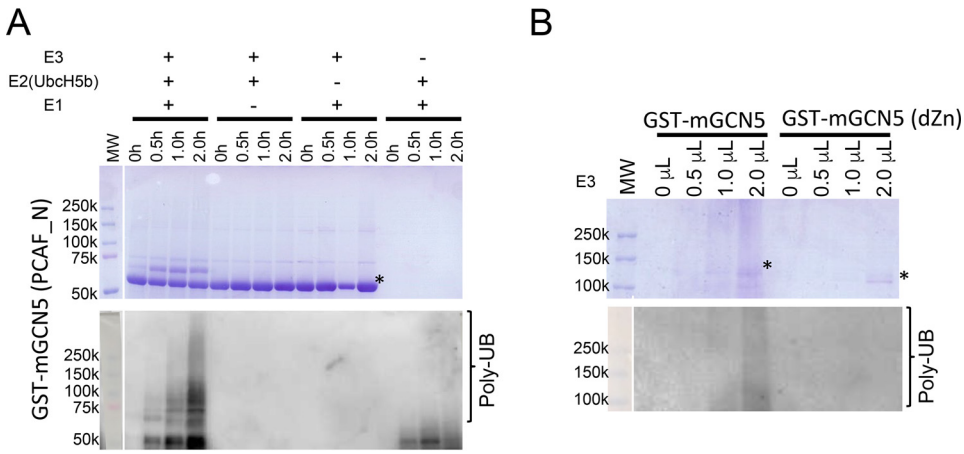

Figure 5. Ubiquitin ligase activities of mGCN5 PCAF_N domain and mGCN5 (dZn). GST-tagged protein indicated by *asterisk* was used as E3 enzymes. UbcH5b was used as E2 enzyme. The experiments were performed in the presence (+) or absence (−) of E1, E2, and/or E3 samples. The condition of each reaction is described on each lane. Results of both SDS-PAGE analysis (*upper*) and Western blot analysis by using an anti-ubiquitin (*anti-Ub*) antibody (*lower*) were shown. *A,* ubiquitination assay of mGCN5 (PCAF_N). The reaction time is indicated on each lane. *B,* ubiquitination assays of mGCN5 (*left*) and mGCN5 (dZn) (*right*). The volume of E3 enzyme in reaction solution is indicated on each lane. The reaction time was 2 h.

RING structures, RINGs generally have a common mode of interaction with E2: the two loop-like regions, which coordinate zinc surround a shallow groove formed by α-helix. These elements serve as a platform for interaction with E2. In PCAF_N domain, two loops coordinating zinc forms a narrow groove together with α-helices (α2 and α3), although its ternary structure and coordination pattern of PCAF_N domain were completely distinct from that of RING E3 ligases. However, the binding surface remains unknown and further study is necessary. Homologous to E6AP C terminus and RING-between-RING E3 ligases among three classes of E3s contain a catalytic cysteine that receives ubiquitin from E2-ubiquitin. GCN5 PCAF_N domain and PCAF have three conserved cysteines in addition to the zinc coordinating cysteines, but they form a hydrophobic core. Therefore, these residues are unlikely to receive ubiquitin. The E2-ubiquitin–binding site and ubiquitin translation mechanism, including chain specificity, remain to be elucidated.

The C-terminal region of the PCAF_N domain is structurally similar to the MRG domain in MSL3, which is reported to enhance activity of the histone acetyltransferase MOF through interaction with MSL1 (22). Considering the MRG domain is reported to interact with a diverse group of proteins, PCAF_N domain is likely to engage diverse protein targets as well, including CBP/p300, which was reported previously (8). It is conceivable that function of the RING E3 is to recruit both the E2 and the substrate protein. MRG domain at the C termini would work as a substrate-binding site.

To date, no GCN5 substrate has been identified; however, three substrates for PCAF have been reported. Interestingly, the dual enzymatic functions of ubiquitination and acetylation by GCN5 and PCAF target Lys residues as substrate. This additional function of GCN5 and PCAF conserved in vertebrates may provide a specific signal in higher organisms. To address these issues, further structural and functional studies are required.

Experimental procedures

Preparation of GST-fused hPCAF, hGCN5, mGCN5, and mGCN5 mutants

The gene of PCAF was provided by the RIKEN BRC through the National BioResource Project of the MEXT/AMED, Japan. For preparation of GST-fused proteins, all genes (hPCAF, hGCN5, mGCN5, and mGCN5 mutant) were subcloned into pGEX-6P-1 vector. The *E. coli* BL21 (DE3)-RIPL cells were transformed with the vectors and were cultured at 37 °C with LB medium to a suitable cell density (A_{600} 0.4∼0.5), and then induced protein expression with isopropyl 1-thio-β-D-galactopyranoside (final concentration 0.5 mM) overnight at 18 °C. Cells were harvested and lysed with the cell lysis buffer (50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 10% glycerol, 100 μM zinc acetate, 0.2 mg/ml of lysozyme, Protease Inhibitor mixture (EDTA free) (Nachalai Tesque) for 20 min on ice, and then added Triton X-100 and DTT (final concentrations 0.2% and 1 mM, respectively). After incubation on ice for 20 min, the lysate was centrifuged at 20,000 × *g* for 20 min at 4 °C. The supernatant was applied to GS4B resin (GE Healthcare) and purified with a batch method by using elution buffer (50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 10% glycerol, 100 μM zinc acetate, 1 mM DTT, 100 mM reduced GSH). The buffer of each sample was exchanged with Amicon ultra (Merck Millipore) to the stock buffer (50 mM Tris-HCl, pH 8.0, 500 mM NaCl), and the samples were concentrated up to 10∼30 of A_{280}. The expression and purification of all samples were confirmed by the mobility change on SDS-PAGE by PreScission protease digestion.

Protein preparation of hUbcH5b and hUb

The genes of hUbcH5b were provided by the RIKEN BRC through the National BioResource Project of the MEXT/AMED, Japan. For preparing His-UbcH5b, the UbcH5b gene was subcloned into pET44a vector. Expressed protein harbors His_6 and PreScission protease recognition sequences at the N-terminal. Protein expression was performed by the same procedure as described above. Harvested cells were sonicated with a

sonication buffer (50 mM Tris-HCl, 500 mM NaCl, 1 mM DTT, pH 8.0). The proteins were bound to Ni-NTA resin (Qiagen) and eluted by elution buffer (50 mM Tris-HCl, 500 mM NaCl, 500 mM imidazole, 1 mM DTT, pH 8.0) with batch method. His-hUbcH5b proteins used for the binding experiment were purified with Superdex 200 Increase (10/300) (GE Healthcare) gel filtration chromatography before the binding experiment. The N-terminal His tag of hUbcH5b was cleaved with PreScission protease. Then the proteins were purified with Superdex 200 Increase (10/300) gel filtration chromatography. The mobile phase of gelfiltration chromatography contained 25 mM Tris-HCl, 50 mM NaCl, 1 mM DTT, pH 7.5.

The expression vector of human ubiquitin (hUb) was provided by Prof. Shuya Fukai. Protein expression and purification were performed by the same procedure reported previously (23). Prior to ubiquitination assay, the hUb proteins were purified with Superdex 200pg gel filtration chromatography. The mobile phase of gel filtration chromatography contained 10 mM Tris-HCl, 50 mM NaCl, 5 mM 2-mercaptoethanol (pH 7.2).

Preparation of mGCN5(67-378) for structure analysis

The mGCN5(67-378) gene was subcloned into pET44a vector. Protein expression was performed by the same procedure as described above. LB medium containing 10 μM zinc acetate was used. Expressed mGCN5(67-378) harbors His_6 and PreScission protease recognition sequences at the N-terminal. Harvested cells were sonicated with sonication buffer (50 mM Tris-HCl, 500 mM NaCl, 1 mM DTT, pH 8.0). The proteins were bound to Ni-NTA resin (Qiagen) and eluted by elution buffer (50 mM Tris-HCl, 500 mM NaCl, 500 mM imidazole, 1 mM DTT, pH 8.0) with the batch method. After the first affinity chromatography purification, the proteins were purified using HisTrap (GE Healthcare) with a gradient method using the same purification buffers used in Ni-NTA purification. Then, His-tagged mGCN5(67-378) was purified using Superdex 200pg (GE Healthcare) size exclusion chromatography with the buffer (20 mM Tris-HCl, 200 mM NaCl, 1 mM DTT, pH 7.0).

Ubiquitination assay

The FLAG-hUBE1 was purchased from Sigma-Aldrich. For E2 screening, purchased E2 proteins (Ubiquitinylation kit, Enzo Life Sciences) and purchased hUb (R&D Systems, Inc.) were used. Home-made hUbcH5b and hUb were used for other ubiquitination assays. GST-fused hPCAF, hGCN5, mGCN5, and mGCN5 mutants were used as E3 and substrate. The amount of each protein was confirmed with SDS-PAGE.

For E2 screening experiment, the ubiquitination reaction was performed for 1 h at 37 °C. The reaction solution contained 25 mM Tris-Cl (pH 7.5), 50 mM NaCl, 1 mM DTT, 2.5 mM ATP, 2.5 mM $MgCl_2$, 0.07 μM FLAG-hUBE1, 3∼6 μM E2 enzymes, 5 μM (estimated) E3, and 58 μM hUb.

In other ubiqutination assays, reaction solutions contained 25 mM Tris-HCl (pH 7.5), 50 mM NaCl, 1 mM DTT,10 mM ATP, 10 mM $MgCl_2$, 0.1 μM FLAG-hUBE1, 65 μM hUBCH5b as E2, and 140 μM hUb. An amount of E3 was estimated with absorbance at 280 nm or SDS-PAGE. The reaction solution contained 8∼0.1 μM (estimated) E3 proteins. An ubiquitination reaction was performed for 0∼2 h at 37 °C, the reaction was terminated by adding nonreduced SDS-PAGE sample loading buffer into the reaction solution, and then the solution was applied to SDS-PAGE. To detect the polybiquitination chain in Western blotting analysis, ubiquitin antibody (P4D1) (Ub(P4D1) mouse number sc-8017 Santa Cruz) was used as 1st antibody. Horseradish peroxidase-conjugated anti-mouse IgG was used as second antibody.

Protein crystallization

mGCN5(67-378) was used for crystallization. Prior to crystallization, the N-terminal His tag was cleaved with PreScission protease, and then the protein was purified with size exclusion chromatography. The protein solution was exchanged to the crystallization solution (20 mM Tris-HCl, pH 8.0, 200 mM NaCl, 1 mM DTT) and concentrated to 10 mg/ml with Amicon Ultra 30k (Merck Millipore). The crystals were grown by vapor diffusion method with the reservoir solution (100 mM imidazole/MES, pH 6.5, 20% PEGMME 500, 10% PEG 20,000, 20 mM sodium formate, 20 mM ammonium acetate, 20 mM sodium citrate tribasic dehydrate, 20 mM sodium potassium tartrate tetrahydrate, 20 mM sodium oxamate) at 20 °C. Heavy atom derivatives were prepared with the solution in which 1 mM ethylmercurithiosalicylate was added to the reservoir solution. After soaking for 1 h, they were transferred to the harvesting solution with the same composition as the reservoir solution.

Data collection and structure determination

The diffraction data set for the native crystal was collected on beamline BL-1A at PF (Tsukuba, Japan) and for the derivative on beamline NE3A at PF-AR (Tsukuba, Japan). Collected data were integrated, merged, and scaled with XDS (24). Phase determination was performed by a SAD method with the program SHARP/autoSHARP (25). After several cycles of rebuilding with Coot (26) and refinement with REFMAC (27), the model structure was well converged at 1.8 Å resolution. XAFS analysis and data collection to evaluate the existence of zinc atoms were performed at beamline BL44XU at SPring-8 (Hyogo, Japan). Geometry of the final structure was checked with the program PROCHECK (28). Coordinates and the structure factor had been deposited in the Protein Data Bank (PDB code 7BY1). Measurement summary and statistics of crystallographic data are summarized in Table 1.

Size exclusion chromatography/multiple angle light scattering analysis (SEC-MALS)

SEC-MALS was performed to determine the molecular weight. Superdex 200 Increase 10/300 column was used. The crystallization solution was used as mobile phase. 50 μl of 5 mg/ml (141 μM) of mGCN5(67-378) protein solution was injected to a Superdex 200 Increase column. The flow rate was 0.5 ml/min. Light scattering and differential refractive index were measured using a multi angle light scattering DAWN HELEOS II detector (Wyatt Technology), a 2414 Refractive Index Detector (Waters) respectively.

Table 1
Crystallographic statistics of mGCN5(67-378)

	Native	Ethylmercurithiosalicylate	Low energy	High energy
Data collection				
Beamline	PF BL-1A	PF-AR NE3A	SPring-8 BL44XU	SPring-8 BL44XU
Wavelength (Å)	1.1000	1.0000	1.2841	1.2831
Space group	$P2_1$	$P2_1$	$P2_1$	$P2_1$
Cell dimensions				
a, *b*, *c* (Å)	43.1, 148.9, 59.0	58.2, 149.5, 85.4	43.0, 149.2, 58.9	43.0, 149.0 59.0
β (°)	101	101	101	101
Resolution (Å)	50.0-1.80 (1.84-1.80)[a]	50.0-3.00 (3.18-3.00)	2.20 (2.32-2.20)	2.00 (2.11-2.00)
$R_{p.i.m.}$[b] or $R_{meas.}$[c]	4.6 (47.4)[b]	11.0 (32.4)[b]	20.8 (103.5)*d*	21.0 (109.7)*d*
$<I/\sigma(I)>$	14.7 (2.2)	9.2 (3.6)	6.14 (1.44)	5.05 (1.18)
Completeness (%)	100.0 (100.0)	98.6 (99.7)	99.8 (99.6)	99.8 (99.4)
Refinement				
Resolution (Å)	50.0-1.80		50.0-2.19	50.0-1.99
No. reflections	63,909		280,189	374,300
R_{work} (%)*d*	19.6		20.1	19.9
R_{free} (%)[e]	23.1		25.1	24.3
No. atoms				
Protein	4,658		4,648	4,646
Water	415		306	296
B-factors (Å^2)				
Protein	29.2		29.6	33.0
Water	38.5		34	39.3
R.m.s deviations				
Bond lengths (Å)	0.018		0.015	0.016
Bond angles (°)	1.762		1.583	1.656
Ramachandran plots				
Favored	563(99.8%)		562 (100%)	560(99.6%)
Allowed	1(0.2%)		0 (0%)	2(0.4%)
Outlier	0 (0.0%)		0 (0%)	0 (0%)

[a]The numbers in parentheses represent statistics in the highest resolution shell.
[b]$R_{p.i.m.} = \sum\sum_j|<I(h)> - I(h)_j|/(n\text{-}1)^{1/2}\sum\sum_j|<I(h)>$, where $<I(h)>$ is the mean intensity of symmetry-equivalent reflections.
[c]$R_{meas} = \sum(n/n\text{-}1)^{1/2}\sum_j|<I(h)> - I(h)_j|/\sum\sum_j|<I(h)>$, where $<I(h)>$ is the mean intensity of symmetry-equivalent reflections.
[d]$R_{work} = \sum||F_o| - |F_c|/\sum|F_o|$, where F_o and F_c are the observed and calculated structure factors for data used for refinement, respectively.
[e]$R_{free} = \sum||F_o| - |F_c|/\sum|F_o|$ for 5% of the data not used at any stage of structural refinement.

SEC-SAXS data measurement and analysis

SEC-SAXS data for mGCN5(67-378) were collected on beamline BL-10C at PF. A Superdex 200 Increase 10/300 column was used. The crystallization solution was used as mobile buffer. 150 μl of 5 mg/ml (141 μM) of mGCN5(67-378) protein solution was injected to a Superdex 200 Increase column. The flow rate was changed from 0.5 to 0.05 ml/min when the protein elution was started. The elution profile was evaluated using UV-visual installed at the irradiation position. Background data were collected at a position before the sample was eluted. 10 images were collected and its average data were uses as background data. Scattering data processing and R_g calculation were performed with program package SAngler (29). The measurement details of collected data are summarized in Table 1.

For solution structural determination by SEC-SAXS, only one diffraction data at elution top peak was used. The radius of gyration (R_g) was derived by the Guinier approximation by using the program PRIMUS (30) and the pair distance distribution functions ($P(r)$ function) using the program GNOM (31). To evaluate measurement conditions, the agreement of R_g values calculated by Guinier approximation and $P(r)$ function was confirmed. The $P(r)$ functions were also used for determining the maximum dimension of the macromolecules (D_{max}) and used for estimation of the shape of them. The dummy atom model was calculated using the program DAMMIN (32). After 30 times calculation, independent 30 dummy atom models were generated. All models were averaged and selected using program DAMAVER (33), and then second model calculation using DAMMIN was performed with damstart model derived from DAMAVER as the starting model. Regenerated 30 models were selected and averaged with the same procedure and the calculated damstart model was used for final dammin calculation. After the 2nd selection, the model structure was well converged and all models were selected for averaging. To gain a final model structure, one last DAMMIN calculation was performed using the damstart model derived from 2nd round averaging using DAMAVER. The final model was fitted well against the monomer crystal structure. Superposition was performed with the program SUPCOMB (33). Comparison of the experimental scattering curve with a theoretical curve derived from the crystal structure was performed by using program CRYSOL (34). Detailed structural parameters are summarized in Table S1.

Binding assay

The binding assay was performed with Superdex 200 Increase 10/300 at 4 °C. The mobile phase contained 25 mM Tris-HCl, 50 mM NaCl, 1 mM DTT (pH 7.5). 500 μl of 150 μM His-tagged mGCN5(67-378), 500 μl of 150 μM His-tagged hUbcH5b or 500 μl of the mixture of which the final concentration was same was injected onto the column. The flow rate was 0.5 ml/min and elution was collected (1.0 ml/tube).

Data availability

Coordinates and the structure factor had been in the Protein Data Bank (PDB code 7BY1) .

Acknowledgments—We thank all beamline staffs at BL44XU (SPring-8), BL-1A, NE3A, BL-10C (Photon Factory) for help and kind suggestions about experiment. We thank Prof. Shuya Fukai and Dr. Kei Okatsu of the Kyoto University for providing a ubiquitin expression vector. A part of this work was performed using synchrotron beamline BL44XU at SPring-8 under the Cooperative Research Program of the Institute for Protein Research, Osaka University Grant 2016A6626. This work was performed under the approval of the Photon Factory Program Advisory Committee (Proposal No. 2014G712, 2015G706, 2016G597).

Author contributions—S. T. F. and T. S. conceptualization; S. T. F., R. H., T. N., M. S., S. S., N. S., and M. M. data curation; S. T. F. and T. S. writing-original draft; T. S. supervision; T. S. writing-review and editing.

Funding and additional information—This work was supported by a Grant-in-Aid from the Japanese Ministry of Education, Culture, Sports, Science, and Technology (to S. T. F. and T. S.) and CREST JST (to T. S).

Conflict of interest—All authors declare no conflict of interest in this work.

Abbreviations—The abbreviations used are: AT, acetyltransferase; PCAF, p300/CBP-associated factor; RING, really interesting new gene; aa, amino acid(s); r.m.s. deviation, root mean square deviation; SEC, size exclusion chromatography; SAXS, small angle X-ray scattering; MALS, multiangle light scattering; MRG, MORF4-related gene; GST, glutathione *S*-transferase; CBP, CREB-binding protein; CREB, cAMP-response element-binding protein; Ni-NTA, nickel-nitrilotriacetic acid.

References

1. Brownell, J. E., Zhou, J., Ranalli, T., Kobayashi, R., Edmondson, D. G., Roth, S. Y., and Allis, C. D. (1996) Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell* **84,** 843–851 CrossRef Medline
2. Sterner, D. E., and Berger, S. L. (2000) Acetylation of histones and transcription-related factors. *Microbiol. Mol. Biol. Rev.* **64,** 435–459 CrossRef Medline
3. Papai, G., Frechard, A., Kolesnikova, O., Crucifix, C., Schultz, P., and Ben-Shem, A. (2020) Structure of SAGA and mechanism of TBP deposition on gene promoters. *Nature* **577,** 711–716 CrossRef Medline
4. Wang, H., Dienemann, C., Stutzer, A., Urlaub, H., Cheung, A. C. M., and Cramer, P. (2020) Structure of the transcription coactivator SAGA. *Nature* **577,** 717–720 CrossRef Medline
5. Grant, P. A., Eberharter, A., John, S., Cook, R. G., Turner, B. M., and Workman, J. L. (1999) Expanded lysine acetylation specificity of Gcn5 in native complexes. *J. Biol. Chem.* **274,** 5895–5900 CrossRef Medline
6. Sakai, M., Tujimura-Hayakawa, T., Yagi, T., Yano, H., Mitsushima, M., Unoki-Kubota, H., Kaburagi, Y., Inoue, H., Kido, Y., Kasuga, M., and Matsumoto, M. (2016) The GCN5-CITED2-PKA signaling module controls hepatic glucose metabolism through a cAMP-induced substrate switch. *Nat. Commun.* **7,** 13147 CrossRef Medline
7. Smith, E. R., Belote, J. M., Schiltz, R. L., Yang, X. J., Moore, P. A., Berger, S. L., Nakatani, Y., and Allis, C. D. (1998) Cloning of *Drosophila* GCN5: conserved features among metazoan GCN5 family members. *Nucleic Acids Res.* **26,** 2948–2954 CrossRef Medline
8. Xu, W., Edmondson, D. G., and Roth, S. Y. (1998) Mammalian GCN5 and P/CAF acetyltransferases have homologous amino-terminal domains important for recognition of nucleosomal substrates. *Mol. Cell. Biol.* **18,** 5659–5669 CrossRef Medline
9. Nagy, Z., and Tora, L. (2007) Distinct GCN5/PCAF-containing complexes function as co-activators and are involved in transcription factor and global histone acetylation. *Oncogene* **26,** 5341–5357 CrossRef Medline
10. El-Gebali, S., Mistry, J., Bateman, A., Eddy, S. R., Luciani, A., Potter, S. C., Qureshi, M., Richardson, L. J., Salazar, G. A., Smart, A., Sonnhammer, E. L. L., Hirsh, L., Paladin, L., Piovesan, D., Tosatto, S. C. E., *et al.* (2019) The Pfam protein families database in 2019. *Nucleic Acids Res.* **47,** D427–D432 CrossRef Medline
11. Pickart, C. M., and Eddins, M. J. (2004) Ubiquitin: structures, functions, mechanisms. *Biochim. Biophys. Acta* **1695,** 55–72 CrossRef Medline
12. Jin, J., Li, X., Gygi, S. P., and Harper, J. W. (2007) Dual E1 activation systems for ubiquitin differentially regulate E2 enzyme charging. *Nature* **447,** 1135–1138 CrossRef Medline
13. Li, W., Bengtson, M. H., Ulbrich, A., Matsuda, A., Reddy, V. A., Orth, A., Chanda, S. K., Batalov, S., and Joazeiro, C. A. (2008) Genome-wide and functional annotation of human E3 ubiquitin ligases identifies MULAN, a mitochondrial E3 that regulates the organelle's dynamics and signaling. *PLoS One* **3,** e1487 CrossRef Medline
14. Ye, Y., and Rape, M. (2009) Building ubiquitin chains: E2 enzymes at work. *Nat. Rev. Mol. Cell Biol.* **10,** 755–764 CrossRef Medline
15. Buetow, L., and Huang, D. T. (2016) Structural insights into the catalysis and regulation of E3 ubiquitin ligases. *Nat. Rev. Mol. Cell Biol.* **17,** 626–642 CrossRef Medline
16. Linares, L. K., Kiernan, R., Triboulet, R., Chable-Bessia, C., Latreille, D., Cuvier, O., Lacroix, M., Le Cam, L., Coux, O., and Benkirane, M. (2007) Intrinsic ubiquitination activity of PCAF controls the stability of the oncoprotein Hdm2. *Nat. Cell Biol.* **9,** 331–338 CrossRef Medline
17. Mazzà, D., Infante, P., Colicchia, V., Greco, A., Alfonsi, R., Siler, M., Antonucci, L., Po, A., De Smaele, E., Ferretti, E., Capalbo, C., Bellavia, D., Canettieri, G., Giannini, G., Screpanti, I., *et al.* (2013) PCAF ubiquitin ligase activity inhibits Hedgehog/Gli1 signaling in p53-dependent response to genotoxic stress. *Cell Death Differ.* **20,** 1688–1697 CrossRef Medline
18. Morgan, J. E., and Greer, S. F. (2017) Pulling a ligase out of a "HAT": pCAF mediates ubiquitination of the class II transactivator. *Int. J. Cell Biol.* **2017,** 8093813 CrossRef Medline
19. Filippakopoulos, P., Picaud, S., Mangos, M., Keates, T., Lambert, J. P., Barsyte-Lovejoy, D., Felletar, I., Volkmer, R., Muller, S., Pawson, T., Gingras, A. C., Arrowsmith, C. H., and Knapp, S. (2012) Histone recognition and large-scale structural analysis of the human bromodomain family. *Cell* **149,** 214–231 CrossRef Medline
20. Schuetz, A., Bernstein, G., Dong, A., Antoshenko, T., Wu, H., Loppnau, P., Bochkarev, A., and Plotnikov, A. N. (2007) Crystal structure of a binary complex between human GCN5 histone acetyltransferase domain and acetyl coenzyme A. *Proteins* **68,** 403–407 CrossRef Medline
21. Holm, L., and Laakso, L. M. (2016) Dali server update. *Nucleic Acids Res.* **44,** W351–W355 CrossRef Medline
22. Kadlec, J., Hallacli, E., Lipp, M., Holz, H., Sanchez-Weatherby, J., Cusack, S., and Akhtar, A. (2011) Structural basis for MOF and MSL3 recruitment into the dosage compensation complex by MSL1. *Nat. Struct. Mol. Biol.* **18,** 142–149 CrossRef Medline
23. Sato, Y., Yoshikawa, A., Yamagata, A., Mimura, H., Yamashita, M., Ookata, K., Nureki, O., Iwai, K., Komada, M., and Fukai, S. (2008) Structural basis for specific cleavage of Lys 63-linked polyubiquitin chains. *Nature* **455,** 358–362 CrossRef Medline
24. Kabsch, W. (2010) Xds. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 125–132 CrossRef Medline
25. Bricogne, G., Vonrhein, C., Flensburg, C., Schiltz, M., and 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 CrossRef Medline

26. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60,** 2126–2132 CrossRef Medline
27. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol. Crystallogr.* **53,** 240–255 CrossRef Medline
28. Laskowski, R. A., Macarthur, M. W., Moss, D. S., and Thornton, J. M. (1993) Procheck: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26,** 283–291 CrossRef
29. Shimizu, N., Yatabe, K., Nagatani, Y., Saijyo, S., Kosuge, T., and Igarashi, N. (2016) Software development for analysis of small-angle x-ray scattering data. *AIP Conf. Proc.* **1741,** 050017 CrossRef
30. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J., and Svergun, D. I. (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.* **36,** 1277–1282 CrossRef
31. Svergun, D. I. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* **25,** 495–503 CrossRef
32. Svergun, D. I. (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys. J.* **77,** 2896–2896 CrossRef
33. Volkov, V. V., and Svergun, D. I. (2003) Uniqueness of ab initio shape determination in small-angle scattering. *J. Appl. Crystallogr.* **36,** 860–864 CrossRef
34. Svergun, D., Barberato, C., and Koch, M. H. J. (1995) CRYSOL: a program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr.* **28,** 768–773 CrossRef
35. Touw, W. G., Baakman, C., Black, J., te Beek, T. A., Kreiger, E., Joosten, R. P., and Vriend, G. (2015) A series of PDB-related databanks for everyday needs. *Nucleic Acids Res.* 43, D364–D368 CrossRef Medline
36. Kabsch, W., and Sander, C. (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22,** 2577–2637 CrossRef Medline
37. Kadlec, J., Hallaci, E., Lipp, M., Holz, H., Sanchez-Weatherly, J., Cusack, S., and Akhtar, A. (2011) Structural basis for MOS and MSI.3 recruitment into the dosage compensation complex by MSL.1. *Nat. Struct. Mol. Biol.* **18,** 142–149 CrossRef Medline
38. Sievers, F., Wilm, A., Dineen, D., Bibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D., and Higgins, D. G. (2011) Fast, scalable generation of high quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7,** 539 CrossRef Medline