Assessment of a western blot signal for the Bcnt/Cfdp1, a tentative component of Srcap chromatin remodeling complex; trial to overcome off-target problems

Shintaro Iwashita[1*], Takehiro Suzuki[2], Yoshimitsu Kiriyama[1], Naoshi Dohmae[2], Yoshiharu Ohka[1], Si-Young Song[1], Kentaro Nakashima[1,*]

Affiliations:

1. Kagawa School of Pharmaceutical Sciences, Tokushima Bunri University, Shido 1314-1, Sanuki, Kagawa 769-2193, Japan
2. Biomolecular Characterization Unit, RIKEN Center for Sustainable Resource Science, Saitama 351-0198, Japan

* **Corresponding authors:**

Shintaro Iwashita

E-mail: siwast@kph.bunri-u.ac.jp

ORCID: 0000-0001-8086-9192

Kentaro Nakashima

E-mail: nkentaro@kph.bunri-u.ac.jp

ORCID: 0000-0002-7881-5686

Competing Interests: All authors declare no competing interests.

Funding: No applicable funding.

All authors approved the final version of the manuscript.

Author Contributions:

Conceived, designed, performed the experiments, and wrote the manuscript: SI NK. Performed mass analysis and interpretation: TS ND. Performed the experiments and discussion: YK. Contributed reagents and discussion: YO. Performed the experiments and wrote the manuscript: SYS.

Abstract

The BCNT (Bucentaur) protein family is characterized by a conserved amino acid sequence at the C-terminus (BCNT-C domain) and plays an essential role in gene expression and chromosomal maintenance in fungi, fly, and chicken. The mammalian Bucentaur/Craniofacial developmental protein 1 (Bcnt/Cfdp1) is also a tentative component of the Srcap (SNF2-Related CBP Activator Protein) chromatin remodeling complex, but little is known about its properties, partly because there are few suitable antibodies to detect the endogenous protein. We used multiple anti-Bcnt/Cfdp1 antibodies against unrelated immunogens derived from BCNT-C domain and mouse-specific N-terminal peptide. To assign western blot signals and evaluate these antibodies, we utilized a stem cell line from mutant embryos of mouse *Bcnt/Cfdp1*, whose mRNA expression levels were reduced to 75% of the parental cells. In western blotting of these mutant and parental cell extracts with the anti-Bcnt/Cfdp1 antibodies, mouse Bcnt/Cfdp1 was detected as a doublet of approximately 45 kDa. LC-MS/MS analysis of the corresponding doublet for the Flag-tagged mouse Bcnt/Cfdp1 constitutively expressed in T-REx 293 cell (a HEK293 derivative) exhibited that the

upper band was much more phosphorylated than the lower band and that there was preferential Ser phosphorylation in the WESF motif in the BCNT-C domain. Western blot with these validated antibodies indicated a preferential expression of Bcnt/Cfdp1 in the early stages of brain development in mouse and rat, which is consistent with the expression of Bcnt/Cfdp1 mRNA. This article describes the evaluation of anti-Bcnt/Cfdp1 antibodies, including a scheme to prepare a potential negative control for western blot, and discusses immune-cross reactions with off-target proteins, particularly immunoreaction probabilities.

Introduction

The BCNT family members in yeast, *Drosophila*, and chicken have been shown to play essential functions in gene expression and chromosomal maintenance [1]. Mammalian Bcnt/Cfdp1 is also presumed to be a tentative component of the SRCAP (SNF2-Related CBP Activator Protein) chromatin remodeling complex (Human soluble protein complexes; http://human.med.utoronto.ca/php/search_complex.php?clusterid=595) based on the

analysis of Swc5, a budding yeast ortholog of Bcnt/Cfdp1, in Swr1 (yeast Srcap) chromatin complex [2-4]. Although Swc5 is not integrated with the Swr1 complex, it participates in activation of the remodeler ATPase and in the ATP-dependent histone exchange reaction, which replaces nucleosomal H2A–H2B with H2A.Z–H2B dimers by recruiting the variant H2A.Z in transcription and DNA repair [5-7]. *Swc5* is not essential for survival, but its deletion mutant *swc*5Δ cells result in genetic instability, hypersensitivity to drugs, and transcriptional misregulation because Swr1 binds to chromatin but lacks histone replacement activity [8][The *Saccharomyces* Genome Database https://www.yeastgenome.org/].

The protein structure of the Bcnt family members generally consists of an acidic N-terminal region, a highly conserved C-terminal region with about 80 amino acids (BCNT-C domain), and a hydrophilic region between them (S1 Fig and [1]). Recently, the BCNT-C domain of Swc5 was found to be essential for the histone exchange reaction [9]. On the other hand, a point mutant of the *Drosophila* ortholog, *Yeti*, which entirely lacks the BCNT-C domain, shows substantial chromosomal abnormalities, resulting in lethality before pupation [10]. Thus, the metazoan *Bcnt/Cfdp1* is essential

for survival in contrast to the yeast *swc5*. Furthermore, the chicken ortholog CENP-29 has been identified to be a kinetochore-associated protein [11]. Given a report that CENP-B protects centromere chromatin integrity by promoting histone deposition [12], these results imply that the Bcnt members may play a broader role in the maintenance of the structure and function of the chromosome.

A RNA sequence analysis, as shown by the dramatic influence caused by genetic mutations of *swc5* in fission yeast [13], has revealed complex mechanisms and dynamic processes such as embryonic development and stress adaptation. However, recent studies have shown that there exists discordance between mRNA and protein expression in such dynamic processes and argue that analysis at the transcriptional level is insufficient to predict protein levels [14]. While these processes are quite complex, involving both noncoding RNAs and antisense RNAs, the western blot analysis is the most straightforward way to examine changes in target molecules at the protein level. Because the Bcnt family members may function preferentially in these dynamic processes, analysis of their protein dynamics is essential. However, most of the currently available antibodies against Bcnt/Cfdp1 are challenging to assign the correct

signal in western blot analysis.

We previously characterized human BCNT/CFDP1 (hBCNT/CFDP1) using a constitutively expressed His-tag molecule [15] and the custom-made antibody generated against an 18-mer peptide (EELAIHNRGKEGYIERKA) located in the BCNT-C domain (anti-BCNT-C Ab) [16]. Despite a calculated mass of 34.9 kDa (33.6 kDa plus His-tag) of His-hBCNT/CFDP1, the immunoreactive signal was detected around 50 kDa as a doublet band on SDS-polyacrylamide gel electrophoresis (SDS/PAGE), and we showed that the difference in its apparent molecular mobility is mainly due to the acidic stretch located in the N-terminal region and Ser^{250} phosphorylation in the BCNT-C domain [15]. However, we failed to identify the endogenous hBCNT/CFDP1 due to high background caused by anti-His Ab reactive proteins and to accurately assess the specificity of the anti-BCNT-C Ab to endogenous hBCNT/CFDP1. Furthermore, we recently found that the anti-BCNT-C Ab cross-reacts with a completely unrelated target, glutamine synthetase (NP_001035564.1; EC 6.3.1.2, which is also known as γ-glutamate: ammonia ligase)[17]. In this paper, we evaluate and validate the anti-Bcnt/Cfdp1 Abs and assign western blot signals using various target-related

materials, including *Bcnt/Cfdp1* knockdown cells. We also present a scheme to prepare a potential negative control for western blot to detect Bcnt/Cfdp1. Then, we demonstrate high expression of Bcnt/Cfdp1 at an early developmental stage of the brains of mice and rats by using the above-evaluated Abs. We also discuss off-target proteins in terms of immune reaction probability based on the analyses to solve the tasks of the present study.

Results

Detection of Flag-tagged mBcnt as a doublet band

To characterize mammalian endogenous Bcnt/Cfdp1, we first expressed Flag-tagged mouse Bcnt/Cfdp1 (F-mBcnt) in T-REx 293 cells (a derivative of HEK 293) as a reference. We did this because although we previously expressed exogenous His-tagged hBCNT/CFDP1, the relatively high background by cross-reacting proteins with anti-His tag Ab made it difficult to assign the signals in western blot [15]. Therefore, we used Flag-tagging to expect the lower tag specific background, though having negatively charged sequence (DYKDDDDK).

The mBcnt/Cfdp1 is composed of 295 amino acids, which is four amino acids less than the human counterpart. The N-terminal region has low homology between mouse and human (75%) and can be used as species-specific immunogens, while the C-terminal 82 amino acid sequence of the BCNT-C domain is identical except for two amino acid residues (S1 Fig, [1]). We isolated T-REx cell colonies that constitutively expressed F-mBcnt using G418 selection. Both the number and size of the antibiotic-resistant colonies from F-mBcnt transfectants were significantly lower and smaller than those from F-multi-cloning site (F-MCS) transfectants as a control. After growing each colony in the presence of G418, the extracts were prepared and evaluated by western blot using either anti-Flag Ab or anti-BCNT-C Ab (Fig 1, S1 Fig). Whereas all signals with anti-Flag Ab showed doublet bands, the anti-BCNT-C Ab detected one and three bands in the extracts from the F-MCS-derived and F-mBcnt derived colonies, respectively (S1 Fig). Compared with the doublet pattern between transiently expressing cells and constitutively expressing cells, the upper band (Upper) in the transient expression was significantly stronger than that of the constitutive expression (Fig 1). These features are similar to those of His-tagged hBCNT/CFDP1, as previously

reported [15].

Fig 1. F-mBcnt expression as a doublet band and isolation of each band.

(A) Flag-tagged mBcnt was detected as a doublet in both transient and constitutive expression. The F-mBcnt or F-multi-cloning site (MCS) in the vector was expressed in T-REx cells, and the extracts were prepared after culturing for 46 h as transiently expressed samples (Lanes: F-MCS and F-mBcnt). In addition, the extract of the E3 colony that constitutively expressed F-mBcnt was prepared (~5 x 10^4 cells). These proteins were subjected to western blot analysis with anti-Flag Ab. The image in this figure is a replica of part of S1 Fig. (B) Isolation of the upper and lower bands from F-mBcnt doublet expressed in the E3 colony. The supernatant of the E3 extract isolated by centrifugation was applied to anti-Flag-tag agarose beads, and the adsorbed fraction was eluted, as shown in S2 Fig. After evaluation of the chromatogram, the more massive Eluate #1 and #2, which had been isolated by sequential elution with Flag peptide, were separated on SDA/PAGE and detected by Coomassie Brilliant Blue staining. The arrows indicate a doublet band.

Phosphorylation with serine^{246} preference in the upper band.

To reveal the molecular differences between the upper and lower bands, the bands were isolated from lysates of T-Rex-derived E3 cells that constitutively express F-mBcnt using anti-Flag Ab conjugated agarose beads (S2 Fig, Fig 1). Each band excised from the gel was digested with three different proteases—that is, *Achromobacter* protease I (API), AspN, and chymotrypsin—and each digest was subjected to LC-MS/MS analysis. In the analysis where each digested fragment covered 60-67% of the entire F-mBcnt/Cfdp1 sequence, the ratios of the upper to lower bands for each fragment amounts were listed (S1 Table). Furthermore, we focused on phosphorylation sites and presented them systematically (Fig. 2). The upper band is much more phosphorylated than the lower band, and, in particular, the serine 246th mBcnt/Cfdp1 that corresponds to the S250 of hBCNT/CFDP1 is apparently the preferential site of phosphorylation. Their characteristics are very similar to those of His-tagged hBCNT/CFDP1, as previously described [15].

Fig 2. Differential phosphorylation between the upper and lower bands of F-mBcnt.

Top panel: Molecular architecture of Flag-tagged mouse Bcnt/Cfdp1. The protein shown in a large outline comprises the acidic N-terminal region, Lys/Glu/Pro-rich 40 amino acids (named intramolecular repeat [IR], white box), and a highly conserved C-terminal region (BCNT-C domain, blue box) in addition to the Flag-tag at the N-terminus (yellow box). The numbers above or below the outline show the amino acid residues of the protein with (below) or without (above) the Flag tag, respectively. Three black bars A, B, and C indicate the regions of focused phosphorylation sites. Each upper and lower band of F-mBcnt from Fig. 1B were digested with three proteases (API, AspN, and Chymotrypsin) and subjected to LC-MS/MS analysis. All of the identified peptides are listed in Table S1, and their typical phosphorylated fragments and their unphosphorylated counterparts are represented. Each top amino acid sequence in A, B, and C represents each focused region, and the numbers at the N-terminus and the C-terminus correspond to the amino acid residue of mBcnt, respectively. **X**s are different amino acid residues from human BCNT/CFDP1. Red letters show the

identified phosphorylated sites. Red bars indicate the ratios of the amounts of identified phosphorylated fragments in the upper band compared to those in the lower band, and blue bars show the corresponding ratios of their unphosphorylated fragments, respectively. The numbers of peptide-spectrum match (PSM) values are presented in parentheses.

Evaluation of the anti-BCNT-C antibody

We had assumed that the ~50-kDa signal above the F-mBcnt doublet detected by anti-BCNT-C Ab corresponds to hBCNT/CFDP1 from previous results [15]. However, the band migration was judged to be too slow on SDS/PAGE, because F-mBcnt is expected to be a slow migration due to the acid Flag tag. Therefore, we reexamined the previous data of His-hBCNT/CFDP1 [15] in the western signal by introducing anti-Bcnt/Cfdp1 Abs. We chose the two commercial anti-BCNT/CFDP1Abs from the following criteria; the description of the immunogen is clear and the candidate signal is detected in a region significantly smaller than 50 kDa in western blot. They are 26636-1-AP9 (Proteintech,

https://www.ptglab.com/products/CFDP1-Antibody-26636-1-AP.htm) and A305-624A-M (Bethyl, https://www.bethyl.com/product/A305-624A-M/CFDP1+Antibody). According to each catalog, the former Ab has been generated using the larger immunogen (172-299 hBCNT/CFDP1); it detects a single band with 45-50 kDa, and the specificity is validated by siRNA knockdown. The latter Ab recognizes the region of 249-299 hBCNT/CFDP1, has immunoprecipitation ability and detects ~48, ~37, and ~19 kDa signals. However, we found that A305-624A-M detected only a major signal around 48 kDa, whereas 26636-1-AP9 revealed other signals, including a signal of a ~50 kDa (as shown later). Based on the results, we compared western signal patterns between the anti-BCNT-C Ab and A305-624A-M using cell lysates of parent T-REx and its derivative G11 clone that constitutively expresses His-tagged hBCNT/CFDP1 [15]. Both Abs efficiently recognized exogenously expressed His-hBCNT/CFDP1 in the G11 extract but showed distinctly different patterns in the T-REx extract. It is noteworthy that the anti-BCNT-C Ab reacted to a band above the doublet detected with A305-624A-M (Fig 3, left panel). The difference between the two patterns was

confirmed by reprobing with each of the replaced Abs (Fig 3, right panel). Whereas the anti-BCNT-C Ab revealed a few other bands, A305-624A-M detected a doublet but not the bands of ~37 and ~19 kDa demonstrated in the catalog. These results strongly suggest that the distinct ~ 50-kDa band detected with anti-BCNT-C Ab as shown in S1 Fig. is an off-target signal.

Fig 3. Comparative assessment of western blot signals between the anti-BCNT-C antibody and A305-624A-M.

The supernatants (SUPs) and their pellets (PPTs) of T-REx or its G11 cells (His-tagged human BCNT/CFDP1 constitutively expressing clone) were prepared from each cell lysate by centrifugation at 25,000 x *g*. Equal amounts of protein (20 μg) were subjected to western blotting analysis with either anti-BCNT-C Ab or A305-624A-M (left panel). After obtaining their images, each filter was stripped and re-probed with the exchanged Abs (right panel).

Assigning the candidate signal of mBcnt/Cfdp1 using ES

mutant cells

To assign the appropriate western blot signal of endogenous Bcnt/Cfdp1 and evaluate its Abs, we utilized a mouse embryonic stem (ES) cell line that is listed as homozygous m*Bcnt/Cfdp1* mutant cells (i.e., double-knockout cell line) (Cfdp1-K1)[18]. The gene trap vector is inserted in m*Bcnt/Cfdp1* intron 5 (Fig 4A) (GenBank: accession number AG999723.1). As the BCNT-C domain is encoded by exons 6 and 7 (Fig 4A), the Cfdp1-K1 cell lysate could be used as a potential negative control for the validation of Abs generated with the BCNT-C domain as an immunogen, assuming the cells are homozygous m*Bcnt/Cfdp1* mutants.

Fig 4. Assigning western blot signals of endogenous Bcnt/Cfdp1 using mutant ES cells.

(A) Location of the inserted gene trap vector in the Cfdp1-K1 cell. Mouse *Bcnt/Cfdp1* consists of 7 exons, and the dashed lines show their corresponding regions to Bcnt/Cfdp1. A red box indicates the gene trap vector. Two black bars under Bcnt/Cfdp1 predict each location of immunogens for Ab production of anti-BCNT-C Ab (1) and

A305-624A-M (2), respectively. (B) RT-PCR analysis of *Bcnt/Cfdp1* mRNA from Cfdp1-K1 and its parental cells. Using each cDNA from Cfdp1-K1 (Mutant), vdR2-4 (WT), or mouse brain (Contl), RT-PCR was carried out to examine the products that correspond to (a) the full ORF of m*Bcnt/Cfdp1* (928 bp) and (b) the fused region of m*Bcnt* exons 1-5 and a part of *hygromycin phosphotransferase* in the gene trap vector (948 bp), respectively. The products and DNA size ladder markers were accessed by separation in agarose gel, followed by staining with ethidium bromide. (C) Assessment of western blot signal of endogenous Bcnt/Cfdp1 using extracts of Cfdp1-K1 and its parental cells. Cell extracts from equal number (~2 x 10^5 cells) of vdR2-4 (WT) or Cfdp1-K1 (Mutant) cells that had been serially passaged in the presence [Mutant (2)] or absence [Mutant (1)] of G418 and puromycin were subjected to western blot analysis with either anti-BCNT-C antibody (left filter) or 305-624A-M (right filter), respectively. A red arrow indicates a candidate signal of endogenous mouse Bcnt/Cfdp1.

First, we confirmed whether Cfdp1-K1 is knocked out on the expression of *Bcnt/Cfdp1* by reverse transcription-polymerase chain reaction (RT-PCR). After

subculturing in the presence or absence of G418 and puromycin, which can delete the feeder layer cells, we prepared cDNAs from Cfdp1-K1 and vdR2-4 cells; the latter is a parent cell line of Cfdp1-K1. We then examined the target m*Bcnt/Cfdp1* mRNA by RT-PCR and analyzed their products by DNA sequencing. In the cDNA from Cfdp1-K1, we detected PCR products corresponding to not only the fusion gene coding *Bcnt/Cfdp1* exons 1-4 and *hygromycin phosphotransferase* derived from the gene trap vector but also the full-length *mBcnt/Cfdp1* ORF (Fig 4B). This result indicated that the gene trap vector was inserted adequately into intron 5, but Cfdp1-K1 cells were not *Bcnt/Cfdp1* double knockout. Then, a comparative transcriptome analysis by NovaSeq 6000 was carried out between Cfdp1-K1 (mutant) and vdR2-4 (wild-type), and it has been shown that 694 genes were differentially expressed with a 2-fold difference, 188 genes were upregulated, and 506 genes were downregulated (S1 Appendex and S2 Table). Among them, *Bcnt/Cfdp1* mRNA in the Cfdp1-K1 cells was reduced to 74.4% compared to the parent cells (101.75 vs. 136.79 FPKM). Besides, the mRNAs of several housekeeping genes that are frequently used as internal controls, the flanking genes *Bcar1/Cas* and *Tmem 170A*, were also significantly altered (S3 Table.).

Next, we examined whether anti-BCNT-C and A305-624A-M Abs detected the differential expression of Bcnt/Cfdp1 between mutant and parental cells in western blot analysis. Of the several bands detected by the anti-BCNT-C Ab, one ~45-kDa band was significantly reduced in the Cfdp1-K1 cells compared to the band intensity in vdR2-4 cells by both Abs (Fig 4C). These results suggest that the ~45-kDa band is a candidate signal derived from endogenous mBcnt/Cfdp1 in mouse ES cells.

Validating anti-mBcnt-N Ab and assigning the mBcnt/Cfdp1 signal

As mentioned earlier, whereas the C-terminal region of BCNT members is highly conserved, the N-terminal region is variable among species. To further investigate whether the ~45-kDa band was a valid signal, we generated an anti-mBcnt/Cfdp1 Ab using a mouse-specific N-terminal peptide consisting of 16 amino acids as an immunogen (named anti-mBcnt-N Ab, Fig 5A).

Fig 5. Validation of anti-mBcnt-N antibody.

(A) Amino acid sequence alignment of the immunogen for anti-mBcnt-N Ab generation and its counterparts of rat and bovine. Mouse Bcnt/Cfdp1 is schematically shown in Fig 4A. The underlined two red bars present each location of the immunogens for the generation of anti-mBcnt-N Ab and A305-624A-M, respectively. The rat and bovine counterparts are aligned, and red letters indicate the different amino acids from the immunogen peptide. (B) Equal amounts of mouse and rat brain extracts (15 or 20 μg) and the enriched bovine placenta extract (30 or 40 μg) were separated on SDS/PAGE followed by western blot analyses with either anti-mBcnt-N Ab (left filter) or A305-624A-M (right filter). Each first Ab was used at a final concentration of 500 ng/mL and 1 μg/mL, respectively.

While the rat Bcnt/Cfdp1 has the same sequence as the immunogenic peptide derived from mBcnt/Cfdp1, the bovine counterpart has an entirely different amino acid sequence (Fig 5A), implying that the probability of cross-reactivity with anti-mBcnt-N Ab was expected to be very low. Indeed, we could use rat and bovine tissue extracts as potential positive or negative controls, respectively, to evaluate anti-mBcnt-N Ab

specificity concerning endogenous Bcnt/Cfdp1 in western blot analysis. However, since it was difficult to obtain a bovine source containing high Bcnt/Cfdp1 content, we concentrated an extract of bovine placenta with Phos-tag agarose [19], which allows enrichment of phosphorylated protein (S3 Fig), and used it for the western blot analysis (Fig 5).

As a result, while A305-624A-M detected a ~45-kDa signal in both tissue extracts, anti-mBcnt-N Ab detected the band in mice and rats but not in cattle. This result strongly suggests that anti-mBcnt-N Ab specifically recognizes endogenous mBcnt/Cfdp1 despite the fact that a nonspecific cross-reaction with unknown proteins was observed. We further confirmed that the ~45-kDa signal was a valid target signal using another anti-hBCNT/CFDP1 Ab, 26636-1-AP (S4 Fig). Finally, we performed western blot analysis with two Cfdp1-K1 cell lysates prepared after passage in the presence or absence of G418 / puromycin using anti-mBcnt-N Ab or A305-624A-M. The results showed that both Abs recognized a ~45-kDa band with significantly reduced intensity in both Cfdp1-K1 lysates as compared to that in the parental cell lysate (Fig 6).

From these results, we concluded that the ~45-kDa signal is derived from the endogenous mBcnt/Cfdp1.

Fig 6. Assignment of western blot mBcnt/Cfdp1 signals.

Equal amounts of cell extract (20 μg protein) of vdR2-4 (WT) or Cfdp1-K1 (Mutant) cells serially passaged in the presence [Mutant (2)] or absence [Mutant (1)] of G418 and puromycin were subjected to western blot analysis with anti-mBcnt-N Ab (left panel) or A305-624A-M (right filter). A red arrow indicates a signal of mouse Bcnt/Cfdp1. A band detected with anti-mBcnt-N Ab at ~75 kDa (shown as *) is probably the fusion protein of a part of mBcnt (derived from exon 1-5) and hygromycin phosphotransferase. It is 63.9 kDa as a calculated molecular mass but may run slowly on SDS/PAGE due to the acid stretch located in the N-terminal region of the mouse Bcnt/Cfdp1.

To confirm whether the ~45-kDa band detected by the both Abs is identical, the filter that had been detected with anti-mBcnt-N Ab was reprobed with A305-624A-M. The result indicated that the signal was identical.

Expression of mBcnt/Cfdp1 in the early stage of brain development

RNA profiling data of the mouse and rat ENCODE (The Encyclopedia of DNA Elements) projects show that Bcnt/Cfdp1 mRNA expresses ubiquitously and preferentially in the early stage of development. However, recent studies have revealed pervasive discordance between mRNA levels and protein levels, especially in embryonic development [14]. Therefore, we examined mBcnt/Cfdp1 expression in the cerebrum of mouse and rat, focusing on developmental stages using the evaluated anti-Bcnt/Cfdp1 Abs above. The results showed that mBcnt/Cfdp1 preferentially expresses in the early stages and significantly decreased according to the postnatal stages in the rat cerebrum (Fig 7).

Fig 7. Bcnt/Cfdp1 expression of rodent brain in the early stage of development. Equal amounts (20 μg of protein) of cerebrum extracts of the embryo around 17 days (E17) and P0 mouse and rat postnatal day samples (denoted by P# on the top of each

lane) were loaded and subjected to western blotting analysis with anti-mBcnt-N Ab (left panel) or A305-624A-M (middle panel). The latter filter was re-probed with another anti-BCNT/CFDP1 Ab, 26636-1-AP (right panel). The filters were finally stained with Coomassie Brilliant Blue (CBB) to check the amounts of loading proteins (bottom panel).

Discussion

In this paper, we assigned the endogenous mBcnt/Cfdp1 signal in western blot, validated anti-mBcnt/Cfdp1Abs by utilizing various target-related materials, and showed that mouse and rat Bcnt/Cfdp1 expressed preferentially at early stages of brain development. Moreover, based on the problems encountered when attempting to solve these tasks in the present study, we discuss immune-cross reactions with off-target proteins from the viewpoint of immune reaction probability.

We generated F-mBcnt instead of His-tagged hBCNT/CFDP1 because the latter made it challenging to evaluate the specificity of anti-Bcnt/Cfdp1 Abs due to the difficulty to distinguish specific immune-positive signals from the high background of

anti-His tag positive proteins [15]. Indeed, Nono (p54rnb)[20], which contains the HHQHHH sequence in the N-terminus region, could be isolated using anti-His-tag Ab-conjugated beads (S6 Fig). F-mBcnt appeared as a doublet, and the upper band was much more phosphorylated than the lower band-in particular, the phosphorylation of the Ser^{246} residue in the WESF motif of the BCNT-C domain was characteristic. The mBcnt/Cfdp1 doublet was probably caused by the presence or absence of Ser^{246} phosphorylation because this characteristic phosphorylation is very similar to that of His-tagged hBCNT/CFDP1, as previously described [15]. Though its biological significance is not yet clear, one interesting possibility is that this phosphorylation might be a determinant of the intracellular localization of Bcnt/Cfdp1.

We first evaluated a custom-made anti-BCNT-C Ab that detected a distinct band from those detected by anti-Flag Ab in F-mBcnt-expressing cell lysates (S1 Fig). We initially assumed that the band was hBCNT/CFDP1. By introducing another commercially available anti-hBCNT/CFDP1 Ab (i.e., A305-624A-M from Betyl laboratory), we compared the immune-positive signals detected with ant-BCNT-C Ab and found that their patterns were distinctively different in the parental cell extract. On

397 the other hand, the exogenous hBCNT/CFDP1 in the extract of G11 clone were well
398 recognized by both Abs. Therefore, we utilized a Cfdp1-K1 cell line that has been
399 produced by a library of random mutations introduced by gene trap vector insertion in
400 Bloom-deficient ES cells, selected for populations of homozygous mutant cells
401 following mitotic recombination [18], and was listed as a cell line of mouse *Bcnt/Cfdp1*
402 homozygous mutant (Japanese Collection of Research Bioresources (JCRB) Cell
403 Bank_AyuK7D01). However, RT-PCR analysis of cDNA from Cfdp1-K1 revealed the
404 presence of mRNA corresponding to the full-length ORF of the m*Bcnt/Cfdp1*. Indeed, a
405 comparative analysis of transcriptome RNA sequencing between Cfdp1-K1 and its
406 parent cells showed that *Bcnt/Cfdp1* mRNA levels were 74.4%. Splicing may efficiently
407 occur by skipping the acceptor site in the trap gene vector, which is located in the intron
408 5 spanning over 50 kb (Fig 4A). Although the Cfdp1-K1 cell line was not *Bcnt/Cfdp1*
409 double knockout, it is useful as *Bcnt/Cfdp1* knockdown mutant cells to detect a strong
410 candidate signal at ~45 kDa. On the other hand, an attempt has been made to generate
411 dog *Bcnt/Cfdp1* knockout MDCK (Madin-Darby Canine Kidney) cells by targeting its
412 exon 1 according to the same strategy as the production of β- and γ-catenin double-

knockout cells [21], but the expected cells were not obtained so far (W. Kobayashi, personal communication). Finally, we were able to assign the signal of endogenous mBcnt/Cfdp1 detected by two Abs raised against unrelated immunogens. One of the antigens is a mouse-specific N-terminal peptide, and the other is a peptide of BCNT-C domain, which is highly conserved in mammalian Bcnt/Cfdp1.

The following is evidences that the ~45-kDa band is the signal of endogenous Bcnt/Cfdp1 in western blot. First, among several signals detected by the anti-BCNT-C Ab, the ~45-kDa signal was significantly reduced in the western blot of the m*Bcnt/Cfdp1* mutant cell extract compared to the signal of the parent cell (Fig 4C). Second, the ~45-kDa signal was detected by two other antibodies (i.e., anti-mBcnt-N Ab and A305-624A-M), each of which was generated using mutually unrelated immunogens (Fig 6) as well as another anti-hBCNT/CFDP1 Ab, 26636-1-AP (S4 Fig). We confirmed the specificity of the anti-mBcnt-N Ab concerning the ~45-kDa signal by preparing enriched bovine Bcnt/Cfdp1 with Phos-tag agarose (Fig 5) as a potential negative control in a western blot.

The target band at ~45-kDa appears significantly smaller than signals reported by many available anti-Bcnt/Cfdp1 Abs, including our custom-made anti-BCNT-C Ab. The molecular behavior of ~50-kDa protein reported with many anti-hBCNT/CFDP1 Abs was significantly different from that of endogenous Bcnt/Cfdp1 (S1 Fig and Figs 3 and 4C); therefore, the 50-kDa signal is probably a common non-specific band(s). Since we used this anti-BCNT-C Ab, our initial conjectures regarding the 50-kDa signal (refer to abstract in [15]), the 43-kDa signal (glutamine synthetase, [17]) (refer to Fig 3C in [16] and Fig 1B in [22]) determined by western blot analysis, and the intracellular localization by immunostaing (refer to Fig 6b in [23]) were all misdirected. On the other hand, we did not recognize two bands at ~37 and ~19 kDa, which are shown in the catalog of A305-624A-M. Furthermore, although five isoforms of m*Bcnt/Cfdp1* have been reported (*Mus musculus*, NCBI Accession No. ID: 23837), we could not identify any isoforms during our western blot analyses.

It was evident that the anti-BCNT-C Ab detected a weak signal to the target molecule (Fig 4C). A similarly difficult situation must occur with many Abs, which is considered to be problematic (e.g., [24], [25]). Concerning the time-consuming

validation of anti-Bcnt/Cfdp1 Ab using western blot, we consider the intrinsic issues that may have led to its inappropriate assignment and how these could be resolved. First, we should use the information of preferential expression of Bcnt/Cfdp1 mRNA–for example, mouse embryo brain is more suitable than the adult brain as screening samples for its Ab evaluation. Second, although we needed other independent anti-BCNT-C Abs and attempted to prepare them using the N-terminal region as immunogens, we failed to obtain the appropriate Abs that could be utilized at that time. Third, although non-specific band patterns detected by Abs raised against different immunogens are generally not identical, many available anti-Bcnt/Cfdp1 Abs including the anti-BCNT-C Ab and 26636-1-AP, commonly detected a relatively strong signal(s) near 50-kDa, and thus resulted in an incorrect assignment. Regarding the second and third points, the properties of Bcnt/Cfdp1 may be related since it mainly consists of structurally disordered regions. It is noteworthy that the epitopes from disordered antigens are smaller than those from the ordered counterparts and that they interact more efficiently with their Abs [26]. Thus, the anti-Bcnt/Cfdp1 Abs generated against most parts of the molecule as immunogens may cause cross-reaction with various proteins with high

affinity, resulting in the poor quality of antisera. Fourth, complex migration of Bcnt/Cfdp1 on SDS/PAGE made it challenging to assign a valid signal. BCNT/CFDP1 as well as mBcnt/Cfdp1 is expressed as a doublet band and migrates slower on SDS/PAGE than expected from the calculated molecular mass (33.6 kDa of human BCNT/CFDP1 and 32.7 kDa of mouse Bcnt/Cfdp1). This feature may be mainly due to the acid stretch of the N-terminal region and the Ser phosphorylation of the BCNT-C domain [15]. Lastly, despite a lack of substantial evidence that *Bcnt/Cfdp1* is directly involved in craniofacial development, its attractive but misleading nomenclature (i.e., craniofacial development protein 1) may have caused confusion regarding its function in Ab providers and researchers. As a result, this may have prevented prior careful analysis of the molecule.

Off-target problems have been widely discussed in many Ab validation studies (e.g., [24], [25]), including specific in-depth efforts for Ab evaluation (e.g., [27]). Moreover, a strategy for Ab validation has been proposed [28]. However, the focus of this discussion appears to be blurred, at least regarding western blot analysis.

Immune cross-reactivity of Abs is based on a general chemical reaction determined by the reaction probability. Epitopes are conventionally divided into two categories: linear or sequential and discontinuous or conformational epitopes [29]. However, Abs do not recognize even linear epitopes as a series of amino acid residues but rather the physicochemical and stereochemical states that they constitute, and these properties as a whole constitute epitope [30, 31]. For example, the following two cases may reflect topological or stereochemical similarity of small environments that determine the common epitopes between completely different proteins: Bcnt/Cfdp1 and glutamine synthetase [17] and phosducin and β-actin [32], respectively. Thus, in principle, it is impossible to exactly match the properties of the epitope with their amino acid sequences, even in linear epitopes. Furthermore, it has been shown that linear epitope peptides that reveal apparent off-target binding at the peptide levels have a strict conformational component at the protein levels [33]. The classification of two types of epitopes is therefore not easily producible.

Western blot signals detected by a certain Ab are strongly influenced by the target concentration in test samples, as shown in Fig 3. The anti-BCNT-C Ab recognized the

exogenously expressed His-tag hBCNT/CFDP1 but scarcely detected an endogenous counterpart. The result implies that the extract of cells overexpressing a target protein does not qualify as a positive control for Ab evaluation in some experiments. However, a good Ab means that it is useful to reveal new evidence regardless of the type of experiment, which depends mainly on the target concentration to be analyzed and the extract preparation method (Figs 3 and 5 and [25]). The experimental materials are quite different with respect to their species, tissues, and developmental stages. Besides, when the target gene has a strong influence on other genes or when the target is a lowly expressed protein, some efforts may be required to assign their appropriate target signals even using lysates from knockout or knockdown cells as a negative control. As shown by the effect of Bcnt/Cfdp1 knockdown in Cfdp1-K1 cells, it is not uncommon that a single gene mutation dramatically alters the expression levels of other genes (S3 Table and S4 Table). Furthermore, it is possible that many proteins showing similar mobility with the target molecule on SDS/PAGE overlap the migration region of the target molecule, making the signal assignment challenging [34]. There is no guarantee that the once validated Ab will work with other samples.

Of course, what we can do is generate antibodies using many unrelated immunogens and also eliminate troublesome off-targets by rigorously screening candidate antibodies (e.g., [28]). However, it is not necessary to be concerned with a single signal, and even if a few off-target signals are detected, these kinds of Abs can be useful if we have evaluated their limitation correctly, as shown in Fig 5 and in many reports (e.g., [35]). Abs do not act within the confines of all-or-nothing modes via specific reactions; therefore, it is critical to understand the efficacy and limitations of the antibody used in any experiments.

Although sensitivity of western blot analysis to crude extract is much better than that of mass spectroscopy, these tools are fundamentally different–that is, individual Ab is not a tool for identifying a molecule but rather a tool for checking a contradiction. Recently, mass analyses have become remarkably advanced and widely available. Thus, it is now much easier to identify molecules that have been considered to be false targets due to wrong-cross reaction with Abs. These trials may provide byproducts of the excellent Abs, such as conformation-specific antibodies against the proteins.

As a reliable anti-Bcnt/Cfdp1 Ab (i.e., A305-624A-M) becomes clearer at present, it is possible to characterize endogenous Bcnt/Cfdp1 more accurately, including subcellular fractionation and tissue distribution. On the other hand, the comparative

transcriptome analysis of Cfdp1-K1 cells revealed that only a 25% decrease in Bcnt/Cfdp1 mRNA resulted in a marked up-regulation or down-regulation of many genes (S1 Appendix, S2 Table), though the confirmation is required by generating their revertant cells by removing the mutagenic vector sequences through Flp-FRT recombination [18]. Among them, it is noteworthy that mRNA expression of intermediate filaments of keratins 8, as well as 18 and 19, was dramatically suppressed –that is, 1% relative to the parental cells (S4 Table, [36]). Given the function of Swc5, which is required under stressed conditions and conditions requiring rapid transcription ([*Saccharomyces* Data Base], [6], and [7]), Bcnt/Cfdp1 may play an essential role for maintaining cell homeostasis, especially in processes such as developmental stage, cell differentiation, and DNA damage repair. The Cfdp1-K1 cell line may serve as a stable Bcnt/Cfdp1 knockdown cell in elucidating the functional role of Bcnt/Cfdp1 in the early developmental stage by using western blotting analysis with further reliable anti-Bcnt/Cfdp1 Abs.

Materials and Methods

All of the reagents and materials, and primers used are listed as S5 Table and S6 Table, respectively.

Ethical approval

All of the genetic recombination experiments and all of the animal experiments in the present study were approved by the Genetic Recombination Experiment Safety Committee and the Animal Care and Use Committee, respectively, of Tokushima Bunri University. All of the experiments were performed in accordance with NIH Guidelines for the Care and Use of Laboratory Animals.

Cell culture

T-REx-293 cells (a HEK293 cell derivative, T-REx) and all clone/subcolonies including G11 clone cells that constitutively expressed His-tag hBCNT/CFDP1 [15] were routinely maintained in DMEM-GlutaMAX-1 (DMEM), which contained 10% Fetal calf serum, 50 μg/mL gentamicin, and with or without G418 (0.5 mg/mL) in 5% CO_2 incubator. For subculturing, the medium was aspirated completely, and the cell layers (2 ml culture on 35-mm dish) were incubated with 0.75 mL accutase for 5 min at

557 room temperature. Then, they were homogenized by pipetting using a 1000-μL pipet tip,
558 and 40-60 μL of the suspension was directly plated on the dish preincubated with the 2
559 mL culture medium for at least 30 min, and G418 was added the next day when needed.
560 For transfection or preparation of cellular protein extract, the cell layers (5 mL culture
561 on 60-mm dish) were washed with 5 mL prewarmed Hepes buffered saline (HBS, [10
562 mM Hepes-NaOH, pH 7.5, 150 mM NaCl], treated with 1.5 mL accutase, homogenized
563 as described above, and then 1 mL medium was added. The cell suspension was
564 transferred to a 15-mL tube, and the dish was washed with another 1-mL medium
565 followed by combining the suspension within a total 10 min (total ~2.5 mL). After
566 taking out 20 μL to estimate the cell number with a disposable Hemocytometer（Watson
567 Bio Lab）, the suspension was centrifuged (Sakuma Model RSL-IV; 1, 000 rpm, 1 min,
568 room temperature) and resuspended in the culture medium for further study. TrypLE
569 Express was also used in the earlier stage of the study in isolation of T-Rex-derived
570 colonies that constitutively expressed Flag-MCS or Flag-tagged mBcnt (see below).
571

572 **ES cell culture**

Cfdp1-K1 and vdR2-4 cells were obtained from Japanese Collection of Research Bioresources (JCRB) Cell Bank and grown in ESGRO Complete Clonal Grade Medium plus GSK3β Inhibitor (50 μl/100 mL) supplemented with gentamicin (50 μg/mL) in a 35-mm dish that was precoated by incubation with recombinant human Laminin (iMatrix-511) at a final concentration of 5 μg/mL in PBS for either 2 h at room temperature or 1 h at 37 □. For subculturing, the medium was aspirated completely, and the cell layers were washed with 2 mL of prewarmed HBS and incubated with 0.75 mL accutase at room temperature. After 5 min, the cell layer was homogenized by pipetting using a 1000-μL pipet tip, and 1-mL DMEM containing 0.1% polyvinyl alcohol (DMEM-PVA) was added. The suspension was then transferred into a 15-mL tube and the dish was washed with another 1 mL of DMEM-PVA and the suspension was combined (total ~2.7 mL) within a total of 10 min. After taking out 20 μl for counting the cell number, the suspension was centrifuged (700 rpm, 3 min, room temperature) and resuspended in the new medium and seeded at the density of ~ 2.5 x 10^5 cells per 35-mm dish. For the preparation of protein extract and total RNA, the cell suspension obtained above were twice washed with chilled HBS, dispensed at ~1 x 10^6

589 cells per 1.5-mL tube, and centrifuged (100 x *g*, 5 min, 4°C). After removing the buffer
590 completely, the pelleted cells were softly vortexed, snap-frozen in liquid N_2, and
591 stocked at -80 □ until use.
592

593 Generation of T-REx colonies expressing Flag-mBcnt

594 T-REx cells (~2 x 10^6 per 100-mm dish in 10 mL medium) were cultured for 20 h,
595 and each 5 μg of pcDNA3.1 plasmid carrying Flag-MCS or Flag-*Bam*H1-mBcnt cDNA
596 was added in the culture using Lipofectamine 3000 (5 μL in 250 μL Opti-MEM)
597 according to the manufacturer protocol. Just before transfection, each 5 mL medium
598 was once removed, transfected, and the medium that had been saved was back to the
599 culture 4-6 h after transfection. After culturing for a total of 44-48 h, each cell layer was
600 washed with 10 mL prewarmed HBS and treated with 1.5 mL of TrypLE Express for 10
601 min at 37 °C, harvested using 1 mL of medium in a 15-mL tube, centrifuged and
602 resuspended in a 1-mL medium.
603 The number of each cultured cell was counted, and ~60 cells in 50 μL were seeded in a
604 48-well plate preincubated with 150-μL medium. After 4 hours, G418 was added to a

final concentration of 0.75 mg/mL. Colonies growing in 48-well plates were sequentially expanded to 12-well plates (on day 15 after media change twice) and 35-mm dishes (on day 21 after media change twice). Each colonies was then maintained in the presence of 0.5 mg / mL G418.

Protein extracts of T-REx cells and their colonies

After washing with chilled HBS (10 mL/100-mm dish) followed by removing the buffer completely using a piece of filter paper, the cells were homogenized with a cell scraper (17-mm width) in 0.5 mL lysis buffer [20 mM Hepes-NaOH, pH 7.5, 150 mM NaCl, 1 mM 2-mercaptoethanol, designated L_buffer] supplemented with both inhibitors of proteinases and phosphatases. The homogenate was transferred into a 1.5-mL tube, sonicated by a Bioruptor (BM Equipment) in an ice-water bath (10-s pulses repeated 15 times at 10-s intervals) and centrifuged (25,000 x g, 30 min, 4 °C) using a centrifuge (Kubota 3780, rotor AF-2536A). After taking out five μL for measurement of protein concentration, the supernatants were aliquoted, snap-frozen in liquid nitrogen, and stored at -80 °C until use. On the other hand, the pellets were

dissolved in 50-μL lysis buffer containing SDS [1%SDS, 1 mM EDTA in 10-mM Hepes-NaOH, pH 7.5, designated LS_buffer], sonicated (3 x 10-s pulses at 10-s intervals) and centrifuged （10,000 x *g*, 1 min, 4℃）. After taking out 5 μl for measurement of protein concentration, the sample was boiled in SDS/PAGE sample buffer.

Isolation of F-mBcnt by anti-Flag Ab-conjugated agarose beads

The frozen supernatant of E3 colony was thawed, and Nonidet P-40 (NP-40) was added to a final concentration of 0.05%. After sonication for 30 s in the ice-water bath (3 x 10-s pulses at 10-s intervals) followed by centrifuging at 10, 000 x *g* for 1 min at 4 □, the sup (1 mL of 1.2 mg) was mixed with anti-Flag-tag agarose beads (20 μL settled volume) in a 1.5-mL siliconized tube and incubated in a rotary shaker for 2 h at 4 □. The mixture was centrifuged for 30 s using a swing type centrifuge (Swing Man, Type ATT-101), and the supernatant was saved as the unbound fraction. The pellet was suspended in 100-μL of L_buffer containing 0.05% NP-40 plus inhibitors and

637 transferred to a spin column (0.8 mL size) using a 200-μL wide-bore tip, and the tube
638 was once more washed with 100 μL of L_buffer plus NP-40, and the suspension was
639 recovered to the spin column. The through-flow fraction obtained by centrifugation was
640 stocked as the first wash fraction. The agarose in the column was washed twice with the
641 same buffer, followed by being washed once more with HBS, and then the bound
642 proteins were eluted with 50 μL of Flag (DYKDDDK) peptide (150 μg in HBS) by
643 incubation at 4 ℃ for 30 min (Eluate #1) and another 5 min (Eluate #2), sequentially.
644 The agarose in a column was further treated with 50 μL of glycine-HCl (50 mM, pH
645 2.5), and its eluate and the agarose in the column were immediately neutralized with 2
646 M Tris. Finally, 50 μL of 1 x SDS/PAGE sample buffer was added to the column,
647 vortexed, and boiled for 5 min. All of the fractionated samples except the fraction eluted
648 with SDS/PAGE sample buffer were boiled in 1 x SDS/PAGE sample buffer by mixing
649 with 4 x SDS/PAGE buffer. Each five μL (3.75 μL of the net eluate) were separated on
650 1.25 % SDA/PAGE and followed by western blotting analysis using anti-Flag Ab. For
651 LC-MS/MS analysis, Eluates #1 and #2 described above (each ~40 μL) were
652 concentrated with acetone according to the protocol

(http://tools.thermofisher.com/content/sfs/brochures/TR0049-Acetone-precipitation.pdf). The dry pellet was once dissolved in 6 μL of 10-fold diluted LS_buffer, and two μL of 4 × SDS/PAGE sample buffer was added followed by boiling for 3 minutes. After separating the sample for 10 minutes longer than usual on 12.5% gel SDS/PAGE, the gel was fixed with 50% MeOH-10% acetic acid solution for 20 minutes, stained with 0.25% CBB for 10 minutes, and then de-stained in 10% MeOH-7% acetic acid solution.

Protein extracts from mouse and rat brain

Cerebrum and cerebral cortices were dissected from mice (P0 of C57BL/6J, male) and rats (P0-P56 of Wistar rat, male) after euthanasia under anesthesia, respectively. These isolated samples were snap-frozen in liquid N_2 and stored at -80 □ until use. Frozen samples were crushed with a hammer on dry ice and immediately transferred to a glass-Teflon homogenizer with 1 mL of LS_buffer per 100 mg of samples. Then, tissues were homogenized at 600 rpm using a digital homogenizer and boiled for 5 min. These homogenates were sonicated (12 x 10-s pulses at 20-s interval) and centrifuged (28,000 x *g*, 30 min, 20 ℃). In the case of a mouse embryo on day around 17, a frozen

669 piece (~40 mg) was wrapped with aluminum foil, crushed with pliers with the help of
670 liquid nitrogen, transferred to a 1.5-mL BioMasher, and soaked in chilled 200 μL of
671 LS_buffer. Then, the tissues were homogenized in ice, boiled for 5 min, sonicated for
672 2.5 min (15 x 10-s pulses at 10-s intervals), and centrifuged (15,000 x *g*, 10 min, 20 ℃).
673 The supernatants were used for western blot analysis. Protein concentrations in the
674 supernatants were determined using a Bicinchoninic Acid (BCA) protein assay kit with
675 bovine serum albumin as a standard.
676

677 Enrichment of bovine Bcnt/Cfdp1 by Phos-tag agarose

678 Bovine placenta (from Holstein Day 116) is a gift from Dr. Kazuyuki Hashizume
679 (Iwate University, Morioka) and was stored in liquid nitrogen until use. Each piece (~30
680 mg) was further subdivided with a razor blade and extracted in two ways: one for the
681 whole extract preparation and the other for the concentration of Bcnt/Cfdp1 content. The
682 minced pieces were placed in 1.5 mL of BioMasher, soaked in cold 200 μL of L_buffer
683 and homogenized, and then 10 μl of 20% SDS was mixed, followed by boiling for 5
684 minutes. The extract was sonicated for 2.5 minutes (15 x 10-s pulses at 10-s intervals),

685 centrifuged (15,000 × *g*, 10 minutes, 22 °C.), and the supernatant was used as a whole
686 extract. For enrichment of Bcnt/Cfdp1, the fined tissue piece in the 1.5-mL BioMasher
687 was homogenized in 100 μL of chilled RIPA buffer [20 mM Tris-HCl, pH 7.5, 150 mM
688 NaCl, 0.5% sodium deoxycholate, 1% NP-40, 1 mM EDTA] plus both inhibitors of
689 proteinases/phosphatases. After transferring the homogenate into a new 1.5-mL tube, the
690 homogenizer was washed with another 100 μL of RIPA buffer, and the solution was
691 combined and centrifuged (15, 000 x *g*, 10 min, 4 □). After taking out 5 μL for
692 measurement of protein concentration, the supernatant was aliquoted, snap-frozen in
693 liquid N_2 and stored in -80 □ until use (total 1.4 mg protein). After dilution of the
694 supernatant with RIPA buffer, 100 μL (200 μg) was enriched by Phos-tag agarose (100
695 μL settled volume in a spin column, 0.8 mL size) according to the manufactured protocol
696 with the following modification; use of a swing-type centrifuge (Swing Man, Type
697 ATT-101) and softly tapping the spin column during washing (200 μL, three times) and
698 elution (100 μL per tube three times). Each eluate was precipitated with TCA at a final
699 concentration of 20% and washed twice with 50 μL of acetone according to the protocol
700 (http://www.its.caltech.edu/~bjorker/TCA_ppt_protocol.pdf). After heating at 95 □, 50

μL of SDS/PAGE sample buffer was added and solubilized by a mixer (Tomy MT-360) for 30 min at room temperature. For larger preparation, the above 700 μg of the extract was diluted to 350 μL with RIPA buffer, applied to a 350-μL settled volume of Phos-tag agarose in a spin column. After rinsing with 0.5 ml of washing buffer three times, the bound proteins were eluted with 0.5 and 0.45 mL of elution buffer sequentially into one tube, as described above. After precipitation with TCA, the pellets were mixed with 50 μL of LS_buffer for one hour. The total recovered protein was 112 μg, with a yield of 16 %. On the other hand, to estimate the approximate Bcnt/Cfdp1 content in the pellet (~20 μL) of the above centrifugation (15,000 x *g*, 10 min, 4□), 6.7 μL of 4 x SDS/PAGE sample buffer was added, mixed vigorously, boiled for 5 min, and sonicated for 2. 5 min (15 x 10-s pulsed at 10-s intervals).

Immunoblotting

Procedures of SDS/PAGE of 12.5% or 15% gel and blotting onto membranes were mostly the same as previously described [15, 17]. The blotted PVDF membrane was blocked in 5% skim milk in TBT buffer [10 mM Tris-HCl, pH 7.6, 150 mM NaCl,

0.1 % Tween 20] for 2 h at room temperature or overnight at 4 □. The first Ab was incubated for 2 h at room temperature or overnight at 4 □ and the second Ab was treated for one h at room temperature. All immunoreactivity was visualized by chemiluminescent using horseradish peroxidase (HRP)-conjugated secondary antibodies that are listed in the S5 Table. Their image was detected by a scanner (GeneGenome, Syngene BioImaging) using ImmunoSTAR Zeta or LD as a substrate. For re-probing, the bound antibodies on the used membrane were stripped by incubation of the filters in a solution of 62.5 mM Tris-HCl (pH 6.8), 2% SDS, and 100 mM 2-mercaptoethanol with stirring for 30 min at ~60 □ on hot block, followed by washing three times with TBT buffer.

Mass spectroscopy analysis

The upper and lower bands of a doublet band that were detected by CBB staining were cut out from the gel and digested with API, AspN, and chymotrypsin, and each digest was analyzed by nano-LC-MS/MS using a Q Exactive mass spectrometer as described previously [15]. The quantification of peptides derived from the upper and

lower bands was carried out using a label-free quantification method using Proteome Discoverer Ver 2.2.0.388 (Thermo Fisher Scientific).

Isolation of total RNA

The frozen ES cells were homogenized with Trizol reagent, and total RNAs were purified using PureLink RNA Mini kit according to the manufacturer protocol [17]. The purified total RNAs were treated with TURBO DNase to eliminate contaminating genomic DNA, extracted with phenol/chloroform/isoamyl alcohol (pH 5.2), and re-purified using RNA Clean & Concentrator -25 kit. The concentration of the total RNAs was determined by the absorbance at 260 and 280 nm using NanoDrop One (Thermo Fisher Scientific), and the quality was estimated using a 2100 Bioanalyzer (Agilent). Two RNAs with the high quality shown below were subjected to RNA sequencing (Macrogen Japan Corp., Kyoto).

RNA source	Concentration (ng/μL)	A260/A280	RIN
Cfdp1-K1 (Bcnt/Cfdp1 Mutant)	454.6	2.07	9.9
vdR2-4 (Wild type)	366.8	2.03	9.6～9.9

Reverse transcription-PCR and plasmid construction

cDNAs were synthesized by Superscript III SuperMix according to the manufacturer protocol using oligo-(dT) 20 primer from purified total RNAs of Cfdp1-K1, vdR2-4 cells, and whole mouse brain of a P56 C57BL/6J male as previously described [17] . The full-length ORF of m*Bcnt* or the fragment of m*Bcnt* exons 1-5 fused with *h*ygromycin phosphotransferase were amplified from each cDNA (1 ng as a total RNA) by PCR using KAPA HiFi HotStart DNA polymerase under the following cycling conditions: denaturation at 95 °C for 3 min, followed by 35 cycles of 98 °C for 10 sec, 68 °C for 15 sec, and 72 °C for 90 sec. The primer sequences are listed in S6 Table. After confirmation of their amplicon size by 0.8% Tris-Acetate EDTA agarose gel electrophoresis and purification by Wizard SV Gel and PCR Clean-Up System (Promega), the sequences of their ORFs were confirmed using BigDye Terminator V3.1 Cycle Sequencing Kit (Thermo Fisher Scientific). For the construction of Flag-mBcnt expression plasmid, the PCR product was inserted into a mammalian

expression vector, Flag-MCS-pcDNA3.1 (Accession No. LC311018), using restriction enzymes *Bam* HI and *Xho* I as previously described [17].

Transcriptome analysis

The following is from a report of HN00101712 (Macrogen Corp Japan, S1 Appendix). The two cDNA libraries from RNAs of Cfdp1-K1 or dvR2-4 cells were prepared using the TruSeq Stranded mRNA LT Sample Prep Kit (Illumina). Their sequences were obtained using NovaSeq 6000 S4 Reagent Kit by a Nova Sequencing system (S1 Appendix, S2 Table). Paired-end reads (read length 101) were mapped to a mouse reference genome (UCSC GRCm38.p4/mm10, annotation RefSeq_2017_06_12). After trimming, 98.81% were mapped on 45,262,214 cleaned reads from Cfdp1-K1 RNA, while 99.09 % were mapped on 52,245,252 cleaned reads from vdR2-4 RNA.

Acknowledgments

We wish to thank Drs. Kyoji Horie and Eiji Kinoshita for providing technical suggestions and useful discussions on ES cell culture and Phos-tag agarose handling,

respectively. We thank Drs. Kazuyoshi Hashizume and Takeharu Masaki for giving us bovine tissues and *Achromobacter* protease I, respectively, Dr. Wakako Kobayashi for sharing the interest in generating Bcnt/Cfdp1 knockout cells, and Dr. Sang Wan Kim for conducting the transcriptome analysis. We are grateful to Drs. Salvador Eugenio Caoili, Christopher A. MacRaild, and Brian McWilliams for their particularly useful discussions and opinions. We also thank Drs. Ed Luk, Takashi Yasuda for discussions on the chromatin remodeling complex and on repair of damaged DNA, respectively. We were indebted to Professor Hiromi Nochi for her encouragement.

References

1. Messina G, Celauro E, Atterrato MT, Giordano E, Iwashita S, Dimitri P. The Bucentaur (BCNT) protein family: a long-neglected class of essential proteins required for chromatin/chromosome organization and function. Chromosoma. 2015; 124(2):153-62. doi: 10.1007/s00412-014-0503-8
2. Krogan NJ, Keogh MC, Datta N, Sawa C, Ryan OW, Ding H, et al. A

Snf2 family ATPase complex required for recruitment of the histone H2A variant Htz1. Mol Cell. 2003; 12(6):1565-76.

3. Mizuguchi G, Shen X, Landry J, Wu WH, Sen S, Wu C. ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. Science. 2004; 303(5656):343-8.

4. Nguyen VQ, Ranjan A, Stengel F, Wei D, Aebersold R, Wu C, et al. Molecular Architecture of the ATP-Dependent Chromatin-Remodeling Complex SWR1. https://doi.org/10.1016/j.cell.2013.08.018.

5. Wu WH, Alami S, Luk E, Wu CH, Sen S, Mizuguchi G, et al. Swc2 is a widely conserved H2AZ-binding module essential for ATP-dependent histone exchange. Nat Struct Mol Biol. 2005; 12 (12):1064-71.

6. Tramantano M, Sun L, Au C, Labuz D, Liu Z, Chou M, et al. Constitutive turnover of histone H2A.Z at yeast promoters requires the preinitiation complex. Elife. 2016; 5. pii: e14243. doi: 10.7554/eLife.14243.

7. Lin CL, Chaban Y, Rees DM, McCormack EA, Ocloo L, Wigley DB.

Functional characterization and architecture of recombinant yeast SWR1 histone exchange complex. Nucleic Acids Res. 2017; 45(12): 7249-7260. doi: 10.1093/nar/gkx414.

8. Morillo-Huesca M, Clemente-Ruiz M, Andújar E, Prado F. The SWR1 histone replacement complex causes genetic instability and genome-wide transcription misregulation in the absence of H2A.Z. PLoS One. 2010; 5(8): e12143. doi: 10.1371/journal.pone.0012143.

9. Sun L, Luk E. Dual function of Swc5 in SWR remodeling ATPase activation and histone H2A eviction. Nucleic Acids Res. 2017; 45(17): 9931-9946. doi: 10.1093/nar/gkx589.

10. Messina G, Damia E, Fanti L, Atterrato MT, Celauro E, Mariotti FR, et al. *Yeti*, an essential *Drosophila melanogaster* gene, encodes a protein required for chromatin organization. J Cell Sci 2014; 127: 2577-2588; doi: 10.1242/jcs.150243.

11. Ohta S, Bukowski-Wills JC, Sanchez-Pulido L, Alves Fde L, Wood L, Chen ZA, et al. The protein composition of mitotic chromosomes

determined using multiclassifier combinatorial proteomics. Cell. 2010; 142(5): 810-21. doi: 10.1016/j.cell.2010.07.047.

12. Morozov VM, Giovinazzi S, Ishov AM. CENP-B protects centromere chromatin integrity by facilitating histone deposition via the H3.3-specifne Daxx. Epigenetics Chromatin. (2017) 10(1):63. doi: 10.1186/s13072-017-0164-y.

13. Clément-Ziza M, Marsellach FX, Codlin S, Papadakis MA, Reinhardt S, Rodríguez-López M, Martin S, Marguerat S, et al. Natural genetic variation impacts expression levels of coding, non-coding, and antisense transcripts in fission yeast. Mol Syst Biol. 2014;10: 764. doi: 10.15252/msb.20145123.

14. Liu Y, Beyer A, Aebersold R. On the Dependency of Cellular Protein Levels on mRNA Abundance. Cell. 2016;165(3): 535-550. doi: 10.1016/j.cell.2016.03.014.

15. Iwashita S, Suzuki T, Yasuda T, Nakashima K, Sakamoto T, Kohno T, et al. Mammalian Bcnt/Cfdp1, a potential epigenetic factor characterized

by an acidic stretch in the disordered N-terminal and Ser250 phosphorylation in the conserved C-terminal regions. Biosci Rep. 2015; 35(4). pii: e00228. doi: 10.1042/BSR20150111.

16. Iwashita S, Osada N, Itoh T, Sezaki M, Oshima K, Hashimoto E, et al. A transposable element-mediated gene divergence that directly produces a novel type bovine Bcnt protein including the endonuclease domain of RTE-1. Mol Biol Evol. 2003; 20(9):1556-63. doi: 10.1093/molbev/msg168

17. Nakashima K, Iwashita S, Suzuki T, Katoh C, Kohno T, Kamei Y, et al. A spatial similarity of stereochemical environments formed by amino acid residues defines a common epitope between two non-homologous proteins. Scientific Rep (in press).

18. Horie K, Kokubu C, Yoshida J, Akagi K, Isotani A, Oshitani A, et al. A homozygous mutant embryonic stem cell bank applicable for phenotype-driven genetic screening. Nat Methods. 2011; 8(12): 1071-1077. doi: 10.1038/nmeth.1739.

19. Kinoshita-Kikuta E, YamadaA, Inoue C, Kinoshita E, Koike E. A novel phosphate-affinity bead with immobilized Phos-tag for separation and enrichment of phosphopeptides and phosphoproteins. OMICS 2011;1(1) 157-169.

20. Passon DM, Lee M, Rackham O, Stanley WA, Sadowska A, Filipovska A, et al. Structure of the heterodimer of human NONO and paraspeckle protein component 1 and analysis of its role in subnuclear body formation. Proc Natl Acad Sci U S A. 2012;109(13):4846-50. doi: 10.1073/pnas.1120792109.

21. Kobayashi W, Ozawa M. The epithelial-mesenchymal transition induced by transcription factor LEF-1 is independent of β-catenin. Biochem Biophys Rep. 2018; 15: 13-18. doi: 10.1016/j.bbrep.2018.06.003.

22. Iwashita S1, Ueno S, Nakashima K, Song SY, Ohshima K, Tanaka K, et al. A tandem gene duplication followed by recruitment of a retrotransposon created the paralogous bucentaur gene (bcntp97) in the ancestral ruminant. Mol Biol Evol. 2006; 23(4): 798-806. doi: 10.1093/molbev/msj088

23. Iwashita S, Nakashima K, Sasaki M, Osada N, Song SY. Multiple duplication of the bucentaur gene family, which recruits the APE-like domain of retrotransposon: Identification of a novel homolog and distinct cellular expression. Gene. 2009;435(1-2):88-95. doi: 10.1016/j.gene.2009.01.012.

24. Vanli G, Cuesta-Marban A, Widmann C. Evaluation and validation of commercial antibodies for the detection of Shb. PLoS One. 2017; 12(12): e0188311. doi: 10.1371/journal.pone.0188311.

25. Qi W, Davidson BA, Nguyen M, Lindstrom T, Grey RJ, Burnett R, et al. Validation of anti-glucocerebrosidase antibodies for western blot analysis on protein lysates of murine and human cells. Biochem J. 2019; 476(2): 261-274. doi: 10.1042/BCJ20180708.

26. MacRaild CA, Richards JS, Anders RF, Norton RS. Antibody Recognition of Disordered Antigens. Structure. 2016; 24(1): 148-157. doi: 10.1016/j.str.2015.10.028.

27. McIntush EW. Response: Antibody crossreactivity between the tumour

suppressor PHLPP1 and the proto-oncogene β-catenin. EMBO Rep. 2013; 14(6): 494-6. doi: 10.1038/embor.2013.68.

28. Uhlen M, Bandrowski A, Carr S, Edwards A, Ellenberg J, Lundberg E, et al. A proposal for validation of antibodies. Nat Methods. 2016; 13(10): 823-7. doi: 10.1038/nmeth.3995.

29. Caoili SE, Expressing Redundancy among Linear-Epitope Sequence Data Based on Residue-Level Physicochemical Similarity in the Context of Antigenic Cross-Reaction. Adv Bioinformatics. 2016; 1276594. doi: 10.1155/2016/1276594.

30. Dalkas GA, Rooman M. SEPIa, a knowledge-driven algorithm for predicting conformational B-cell epitopesfrom the amino acid sequence. BMC Bioinformatics 2017;18(1):95. doi: 10.1186/s12859-017-1528-9.

31. Piotrowska U, Adler G. Phosducin and monomeric β-actin have common epitope recognized by anti-phosducin antibodies. Immunol Lett. 2010;134(1):62-68. doi: 10.1016/j.imlet.2010.08.010.

32. Forsström B, Axnäs BB, Stengele KP, Bühler J, Albert TJ, Richmond TA, et al. Proteome-wide epitope mapping of antibodies using ultra-dense peptide arrays. Mol Cell Proteomics. 2014; 13(6): 1585-97. doi: 10.1074/mcp.M113.033308.

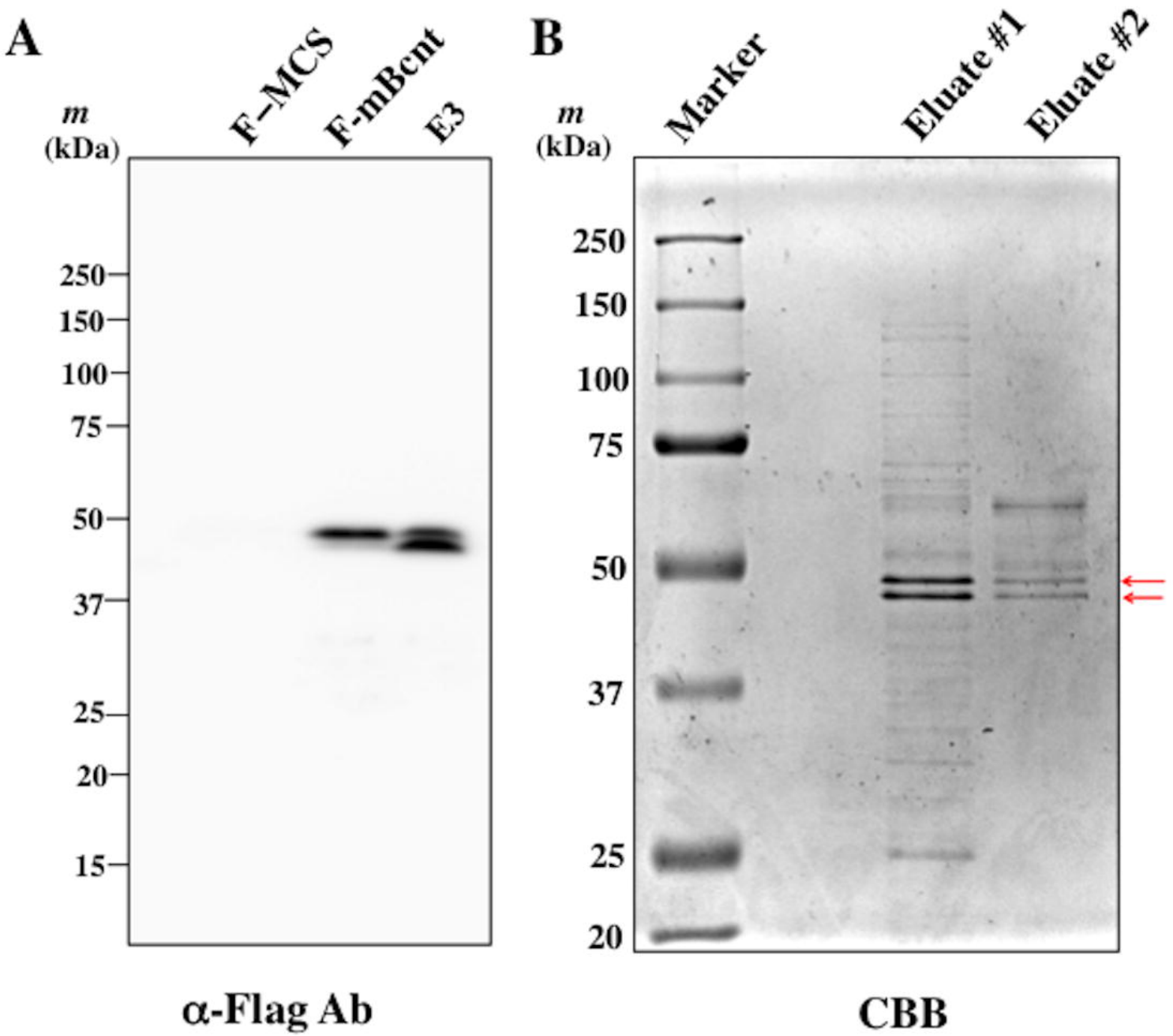
33. Forsström B, Axnäs BB, Rockberg J, Danielsson H, Bohlin A, Uhlen M (2015) Dissecting antibodies with regards to linear and conformational epitopes. PLoS One 10 (3) :e0121673.

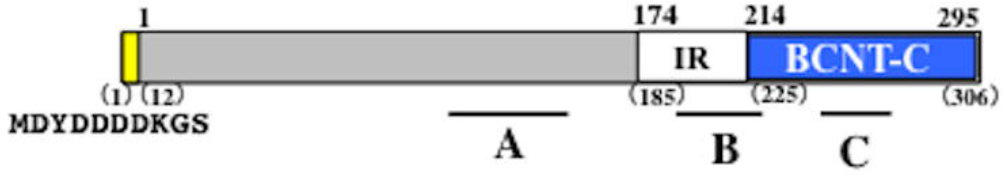
34. McLane MW, McCann U, Ricaurte G. Identifying the serotonin transporter signal in Western blot studies of the neurotoxic potential of MDMA and related drugs. Synapse. 2011; 65(12):1368-72. doi: 10.1002/syn.20958.

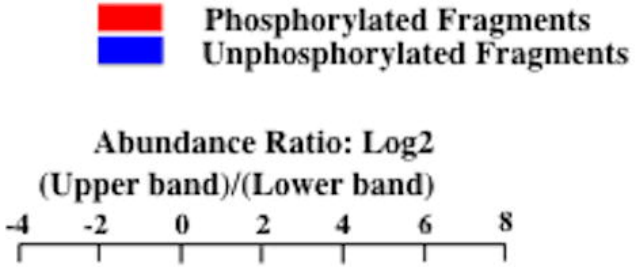
35. Jank T, Bogdanović X, Wirth C, Haaf E, Spoerner M, Böhmer KE, et al. A bacterial toxin catalyzing tyrosine glycosylation of Rho and deamidation of Gq and Gi proteins. Nat Struct Mol Biol. 2013; 20(11):1273-80. doi: 10.1038/nsmb.2688.

36. Saykali B, Mathiah N, Nahaboo W, Racu ML, Hammou L, Defrance M,

et al. Distinct mesoderm migration phenotypes in extra-embryonic and embryonic regions of the early mouse embryo. Elife. 2019;8. pii: e42434. doi: 10.7554/eLife.42434.


A
m
(kDa)
F-MCS
F-mBcnt
E3
250
150
100
75
50
37
25
20
15
α-Flag Ab
B
m
(kDa)
Marker
Eluate #1
Eluate #2
250
150
100
75
50
37
25
20
CBB





A

104 KKEDELWASFLNDVGPKSKAAPGSQTKVAEETEEISSNKPL 144
(115) (155)

KKEDELWASFLNDVGPK	API	(137)	(4)
DELWASFLN	AspN	(7)	(2)
DVGPKSKAAPGSQTKVA	AspN	(0)	(3)
DVGPKSKAAPGSQTKVAEET	AspN	(6)	(2)
LNDVGPKSKAAPGSQTKVAEETEEISSNKPL	Chymotrypsin	(55)	(4)

B

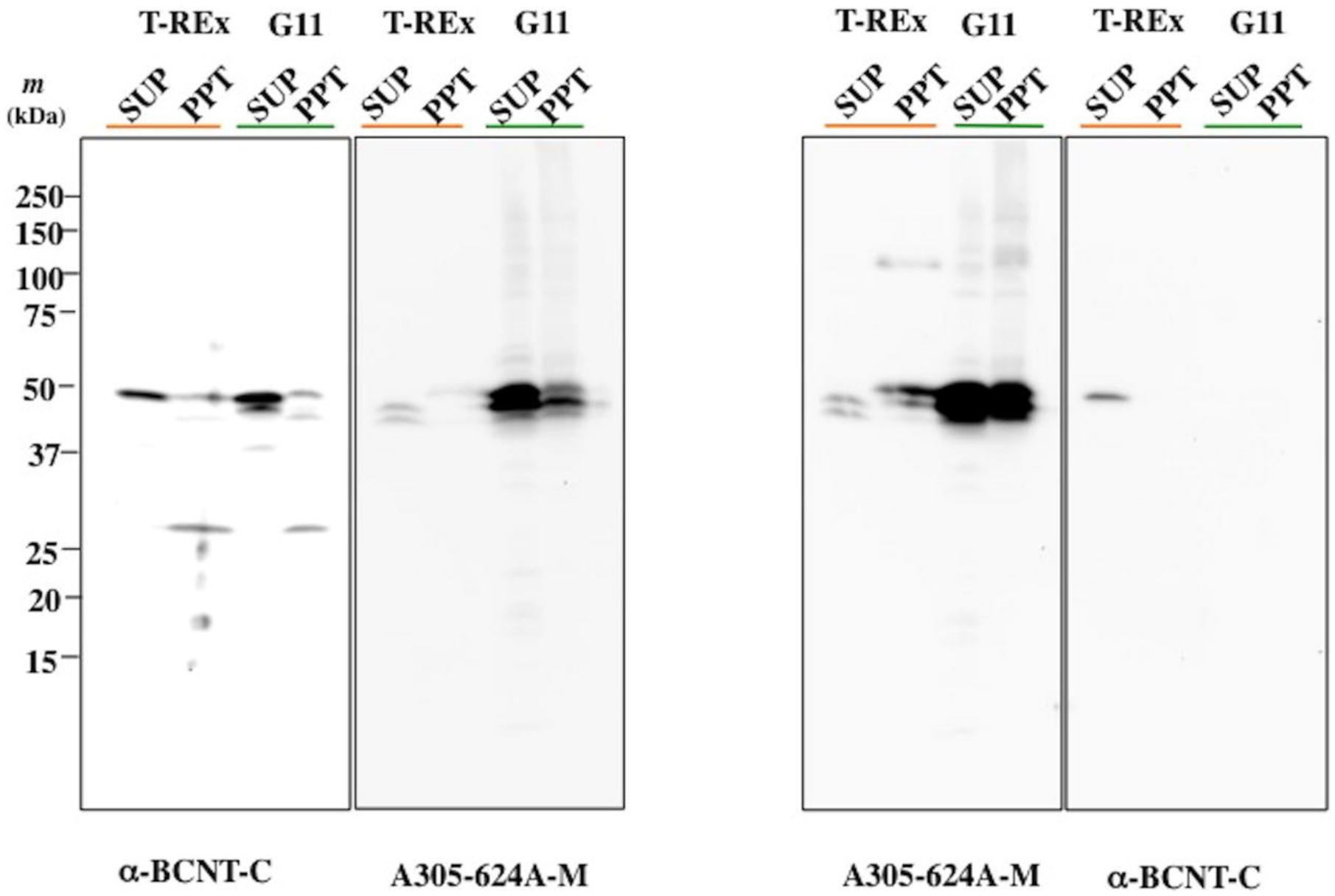
191 QTEREKPQALVTSPATPLPAGSGIKRASGMSSLL 224
(202) (235)

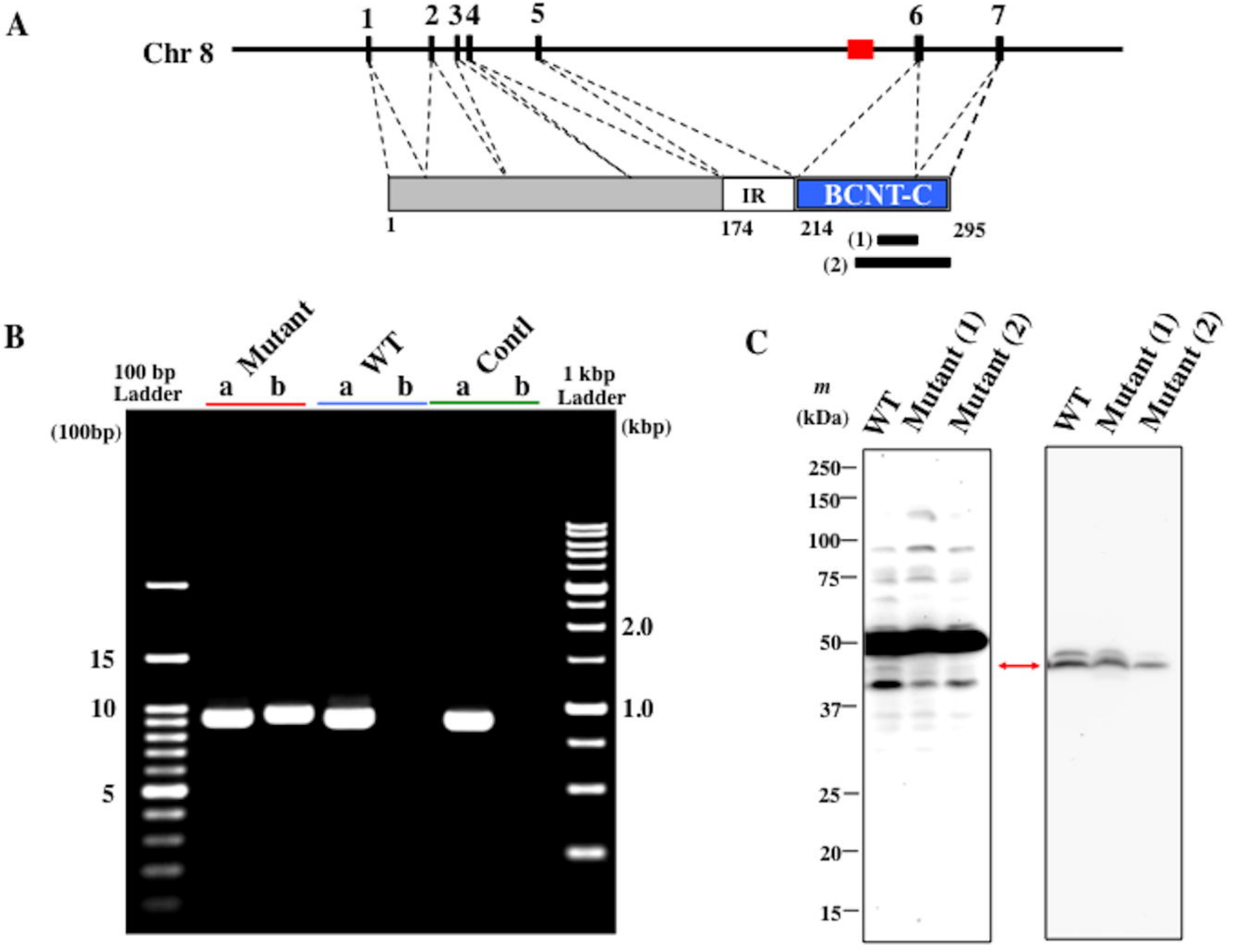
PyrQTEREKPQALVTSPATPLPAGSGIK	API	(4)	(5)
QTEREKPQALVTSPATPLPAGSGIK	API	(36)	(10)
PQALVTSPATPLPAGSGIK (2xP T/S)	API	(79)	(5)
PQALVTSPATPLPAGSGIK	API		(6)
VTSPATPLPAGSGIKRASGMSSLL	Chymotrypsin	(4)	(7)

C

238 EKSKLDWESFKEEEGIGEELAIHNRGK 264
(249) (275)

SKLDWESFKEEEGIGEELAIHNRGK	API	(19)	(71)
DWESFKEEEGIG	AspN	(3)	(4)
EKSKLDWESFKEEEGIGEEL	Chymotrypsin	(11)	(13)


T-REx
G11
T-REx
G11
m
(kDa)
SUP
PPT
SUP
PPT
SUP
PPT
SUP
PPT
250
150
100
75
50
37
25
20
15
α-BCNT-C
A305-624A-M
T-REx
G11
T-REx
G11
SUP
PPT
SUP
PPT
SUP
PPT
SUP
PPT
A305-624A-M
α-BCNT-C

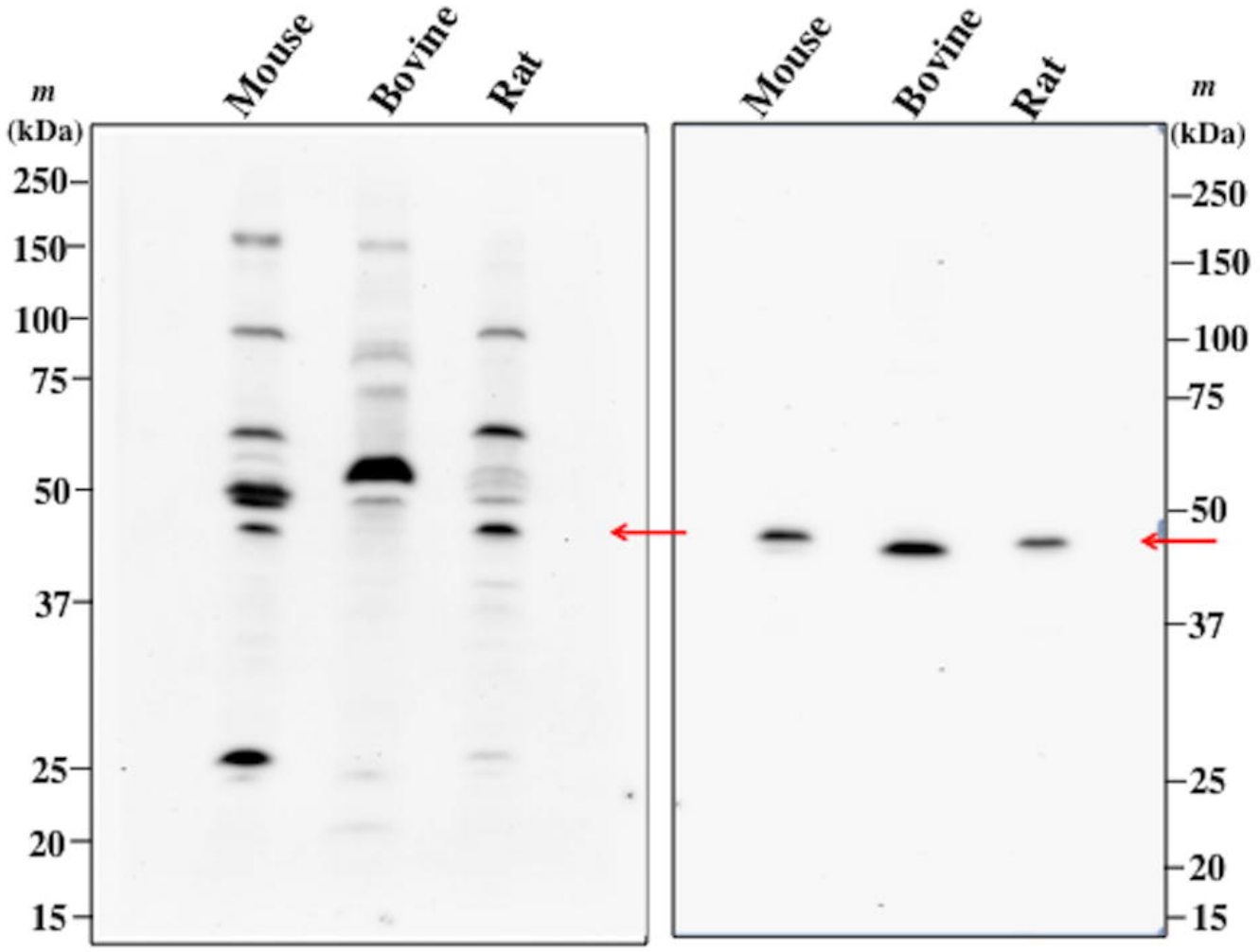

A
Chr 8
1
2 3 4
5
6
7
IR
BCNT-C
1
174
214
295
(1)
(2)
B
Mutant
WT
Contl
100 bp Ladder
a
b
a
b
a
b
1 kbp Ladder
(100bp)
(kbp)
15
10
5
2.0
1.0
C
m (kDa)
WT
Mutant (1)
Mutant (2)
WT
Mutant (1)
Mutant (2)
250
150
100
75
50
37
25
20
15

41 56

Mouse GEEQAEKTKGKRRKAQ

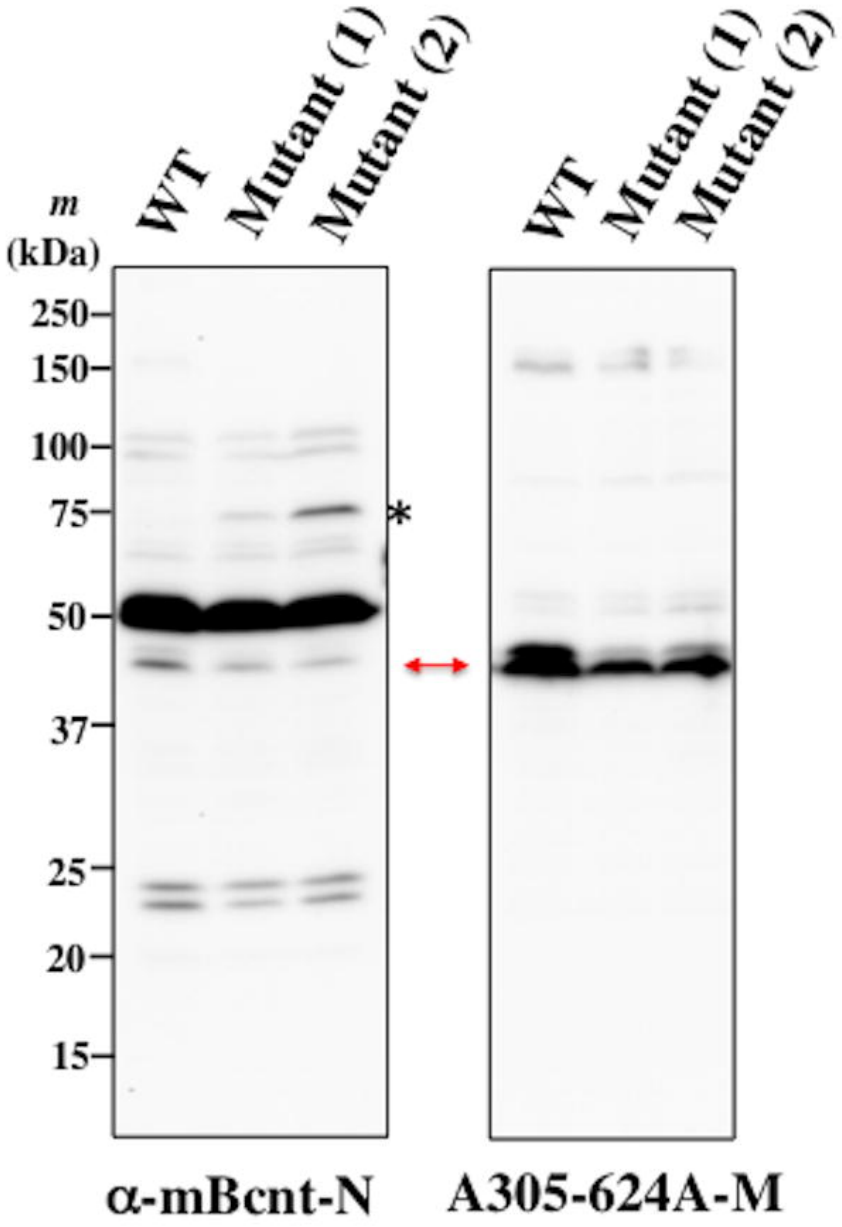
Rat GEEQAEKTKGKRRKAQ

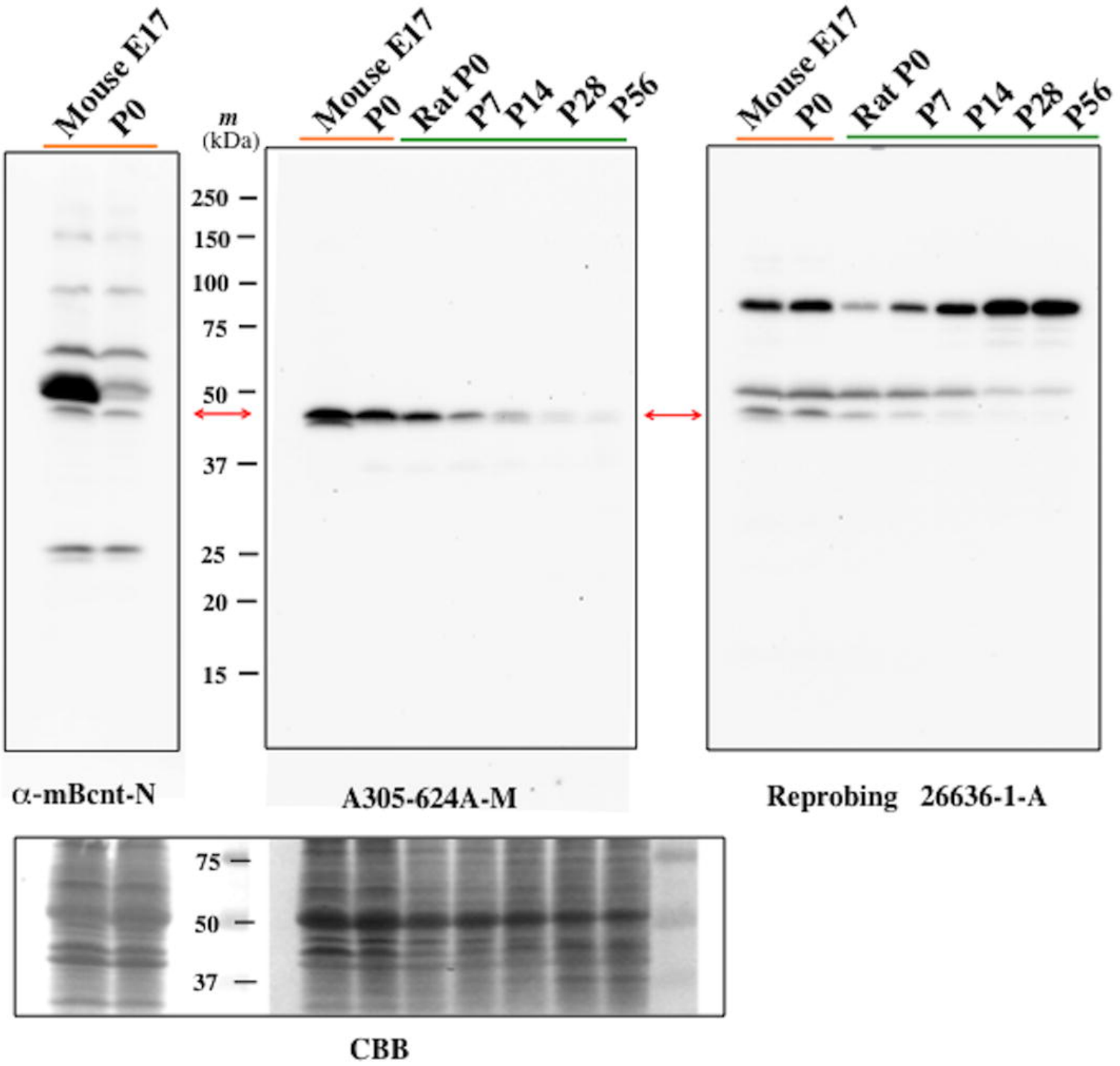
Bovine GEEETQKTKGTKRKAE



α-mBcnt-N

A305-624A-M


WT
Mutant (1)
Mutant (2)
WT
Mutant (1)
Mutant (2)
m
(kDa)
250
150
100
75
50
37
25
20
15
*
α-mBcnt-N
A305-624A-M


Mouse E17
P0
m
(kDa)
Mouse E17
P0
Rat P0
P7
P14
P28
P56
Mouse E17
P0
Rat P0
P7
P14
P28
P56
250
150
100
75
50
37
25
20
15
α-mBcnt-N
A305-624A-M
Reprobing 26636-1-A
75
50
37
CBB

TITLE:

A spatial similarity of stereochemical environments formed by amino acid residues defines a common epitope of two non-homologous proteins

AUTHORS:

Kentaro Nakashima[1*], Shintaro Iwashita[1*, 2], Takehiro Suzuki[3], Chieko Kato[1], Toshiyuki Kohno[4], Yasutomi Kamei[5], Motoki Sasaki[6], Osamu Urayama[7], Yoshiko Ohno-Iwashita[2], Naoshi Dohmae[3] and Si-Young Song[1*]

AFFILIATIONS:

1. Institute of Neuroscience, Tokushima Bunri University, Kagawa 769-2193, Japan
2. Department of Pharmacy, Iwaki Meisei University, Fukushima 970-8551, Japan
3. Biomolecular Characterization Unit, RIKEN Center for Sustainable Resource Science, Saitama 351-0198, Japan
4. Department of Biochemistry, Kitasato University School of Medicine, Kanagawa 242-0374, Japan
5. Laboratory of Molecular Nutrition, Graduate School of Environmental and Life Science, Kyoto Prefectural University, Kyoto 606-0823, Japan
6. Department of Basic Veterinary Medicine, Obihiro University of Agriculture and Veterinary Medicine, Hokkaido 080-8555, Japan
7. Faculty of Health Sciences, Tsukuba International University, Ibaraki 300-0051, Japan

*** CORRESPONDING AUTHORS:**

Kentaro Nakashima (sections of molecular biology and biochemistry)

E-mail: nkentaro@kph.bunri-u.ac.jp

Phone: (81) 87-899-7484

Fax: (81) 87-894-0181

Shintaro Iwashita (section of biochemistry)

E-mail: siwast@kph.bunri-u.ac.jp

Phone: (81) 87-899-7484

Fax: (81) 87-894-0181

Si-Young Song (section of immunohistochemistry)

E-mail: sysong@kph.bunri-u.ac.jp

Phone: (81) 87-899-7484

Fax: (81) 87-894-0181

Institute of Neuroscience,

Tokushima Bunri University

1314-1 Sido, Sanuki, Kagawa, 769-2193, Japan

ABSTRACT

It is critical for development of high-quality antibodies in research and diagnostics to predict accurately their cross-reactivities with "off-target" molecules, which potentially induce false results. Herein, we report a good example of such a cross-reactivity for an off-target due to a stereochemical environment of epitopes, which does not simply depend on amino acid sequences. We found that significant subpopulation of a polyclonal peptide antibody against Bcnt (Bucentaur) (anti-BCNT-C antibody) cross-reacted with a completely different protein, glutamine synthetase (GS), and identified four amino acids, GYFE, in its C-terminal region as the core amino acids for the cross-reaction. Consistent with this finding, the anti-BCNT-C antibody strongly recognized endogenously and exogenously expressed GS in tissues and cultured cells by Western blotting and immunohistochemistry. Furthermore, we elucidated that the cross-reaction is caused by a spatial similarity between the stereochemical environments formed by amino acid residues, including the GYFE of GS and the GYIE of Bcnt, rather than by their primary sequences. These results suggest it is critical to comprehensively analyze antibody interactions with target molecules including off-targets with special attention to the physicochemical environments of epitope-paratope interfaces to decrease the risk of false interpretations of results using antibodies in science and clinical applications.

INTRODUCTION

Antibodies (Abs) are very powerful tools not only in basic research but also in clinical application of molecular targeted therapies for cancer (1-3) and autoimmune diseases such as rheumatoid arthritis (4, 5). Antibody (Ab) drugs are a rapidly growing class of biopharmaceuticals and a large number of novel promising Abs such as nivolumab and pembrolizumab have been generated during the last few years. To efficiently develop high quality Abs for research use as well as Ab drugs and vaccines, it is critical not only to design immunogens that generate Abs with quite specific immunoreactivity for on-target molecules but also to predict more accurately cross-reactivities of Abs with "off-targets", which potentially induce side effects in therapies using Abs and vaccines (6). The amino acid sequences of epitopes have been considered to be a primary basis for specific immunoreactions of Abs with target molecules. However, in addition to Abs against proteins, those against nonprotein molecules, such as anti-DNA and anti-phospholipid Abs, are also well known. On the other hand, recent reports demonstrated that a multispecific monoclonal Ab recognizes completely different linear epitope sequences with high affinity (7), and progress is proceeding toward development of therapies by intentionally using such a dual-specific Ab (8). These examples are difficult to understand from a simple grasp of epitopes based on only their amino acid sequences. The epitopes are conventionally classified into two groups, linear or sequential and discontinuous or conformational (9, 10). Such a conventional classification based on amino acid sequences is useful to simply explain epitopes of proteins but insufficient to more accurately understand epitopes in antigen-Ab interactions such as the above-described examples because these examples indicate that Ab recognition of target molecules is based on the physicochemical

environment of epitopes and does not necessarily depend on their linear amino acid sequences. Accordingly, it is necessary to have a more comprehensive understanding of Ab specificities on the basis of physicochemical features of epitopes. In this report, we present a didactic example to consider these points.

We identified the Bcnt (Bucentaur) / Cfdp1 (craniofacial developmental protein 1) protein family, which probably function as a component of chromatin remodeling complex by the study of yeast ortholog Swc5 (11). Previously, we raised an Ab against a peptide of a C-terminal region of BCNT (called BCNT-C), EELAIHNRGKEGYIERKA, which is evolutionarily conserved from yeast to humans (12). This anti-BCNT-C Ab detected doublet bands of 43/45 kDa by Western blot analysis of proteins extracted from bovine and rat brains. We tried to isolate them by immunoprecipitation and identify them by mass spectrometry. Surprisingly, we obtained the unexpected result that the 43 kDa protein is glutamine synthetase (GS: NP_001035564.1); EC 6.3.1.2, also known as γ-glutamate: ammonia ligase (13, 14), an enzyme entirely unrelated to Bcnt, which catalyzes the ATP-dependent condensation of glutamate and ammonia to form glutamine. The anti-BCNT-C Ab can specifically detect not only recombinant GS but also endogenous GS in rat brain. This quite specific reactivity of the Ab with an off-target molecule prompted us to seek a common epitope between GS and Bcnt to elucidate a chemical basis for the cross-reactivity of the anti-BCNT-C Ab with GS. We report here that the core amino acids of the common epitope causing the cross-reaction were identified as GYFE in GS and GYIE in Bcnt. Also, we discuss here the significance of physico- or stereochemical environments in epitope–paratope (10) or protein–protein interactions (15) and findings on the proteome-wide analyses of epitopes (9).

RESULTS

Detection of the 43 kDa protein by anti-BCNT-C Ab

Previously, we reported detection of strong doublet bands of 43/45 kDa by Western blot analysis of proteins extracted from bovine and rat brains with the anti-BCNT-C Ab (12). This Ab reproducibly detected the 43/45 kDa proteins in mouse and bovine brain as well as the mouse Bcnt around 45 kDa exogenously expressed in HEK293T cells (Fig. 1 and see Supplementary Fig. S1 and S2 online). To examine the specificity of the anti-BCNT-C Ab, we generated two other anti-Bcnt/Cfdp1 Abs using the C-terminus 10-mer peptide, which is evolutionarily conserved, (anti-Bcnt-Cter Ab) and the 14-mer peptide of the N-terminal region in mouse Bcnt (anti-mBcnt-N Ab), respectively, as different antigens from the anti-BCNT-C Ab (see Supplementary Table S1 online). The polyclonal anti-Bcnt-Cter Ab and the polyclonal anti-mBcnt-N Ab detected the recombinant Bcnt similarly as the anti-BCNT-C Ab, whereas they did not detect the 43/45 kDa proteins (Fig. 1). These results clearly indicate that 43/45 kDa proteins are specifically detected by only the anti-BCNT-C Ab, not the anti-Bcnt-Cter and the anti-mBcnt-N Ab. In addition to this discrepancy and different molecular sizes of these immunoreactants from the recombinant Bcnt (Fig. 1 and see Supplementary Fig. S2 online), their very high immunoreactivity prompted us to identify these proteins at the molecular level.

Identification of the 43 kDa protein

We found that we could immunoprecipitate the 43 kDa protein, but not the 45 kDa protein, from the extracts of bovine brain and rat olfactory bulb (16) when the extracts were boiled once in SDS, whereas we failed using a conventional Radioimmunoprecipitation assay

(RIPA) buffer (see Supplementary Fig. S3 online). Using this method, we isolated the 43 kDa protein from the extracts by ammonium sulfate precipitation, phenyl-sepharose chromatography (see Supplementary Fig. S4A online), and immunoprecipitation using agarose beads coupled with the anti-BCNT-C Ab (see Supplementary Fig. S4B online). Three broad Coomassie Brilliant Blue (CBB)-positive bands with a molecular mass of 43, 55 or 65 kDa, respectively, were detected on SDS-Polyacrylamide gel electrophoresis (SDS-PAGE) of immunoprecipitated proteins (Fig. 2). Protein mixtures contained in each separate band were subjected to LC-MS/MS analysis. Surprisingly, results from the analyses showed that glutamine synthetase (GS) was detected in all three bands, whereas no Bcnt fragment was detected in any analyzed band except for the GKEGYIER that was derived from the immunogen peptide used for elution of bound proteins from the affinity beads. In particular, GS having a calculated molecular mass of 42 kDa was detected as a major component in the 43 kDa band together with gamma actin, alpha skeletal muscle actin, and tubulin alpha-1B chain, where the coverage of GS by the tryptic peptides was 76.9% (see Supplementary Fig. S5A, S5B and Table S2 online). Detection of GS in these three distinct bands was probably caused by a mobility shift due to its modification by ubiquitination (17). These data suggest that GS, a protein not homologous to Bcnt, was the most promising candidate for the 43 kDa protein recognized by the polyclonal anti-BCNT-C Ab. It should be noted that the anti-BCNT-C Ab with cross-reactivity to GS was generated in seven of eight guinea pigs immunized with the 18-mer BCNT-C peptide in two independent immunizations, making it unlikely that generation of the anti-BCNT-C Ab with such a cross-reactivity was accidental.

Recognition of glutamine synthetase by the anti-BCNT-C Ab

To examine whether the anti-BCNT-C Ab really cross-reacts with GS, we carried out Western blot analysis using two distinct GS sources: a purified recombinant His-tagged human GS expressed in *Escherichia coli* (*E. coli*) and extracts of skeletal muscle of FOXO1-transgenic mice, which preferentially increase the mRNA and protein expressions of GS (18). These results demonstrate that the Ab detected the same size band at a comparable level as a canonical anti-GS Ab does (see Supplementary Fig. S6 online). Considering the effect of a His-tag added to a recombinant protein, we also expressed tag-free recombinant mouse GS and Bcnt or those with various tags, such as Flag, enhanced green fluorescent protein (EGFP), and mCherry, in HEK293T cells to examine multilaterally the cross-reactivity of the anti-BCNT-C Ab with GS, and then these recombinant proteins were analyzed by Western blotting and immunocytochemistry. Because the anti-BCNT-C Ab recognized GS regardless of type or presence or absence of tags, representative results of tag-free and Flag-tagged mGS and mBcnt are shown in Fig. 3A, which clearly show that the anti-BCNT-C Ab specifically recognized both Bcnt and GS. Additionally, we further examined by immunocytochemistry whether the cross-reactivity to GS *in situ* is specific for the anti-BCNT-C Ab using the anti-Bcnt-Cter Ab and the anti-mBcnt-N Ab. While the anti-BCNT-C Ab can specifically immunostain exogenously expressed GS in HEK cells, but not the two other anti-Bcnt Abs (Fig. 3B). The anti-BCNT-C Ab shows a strong immunoreactivity *in situ* as shown by positive immunostaining of Bergmann glia, a type of astrocyte, in the rat cerebellum (see Supplementary Fig. S7 online), which is consistent with previous reports that GS is a marker of astrocytes (19). Other data are described in supporting information, which show

that some characteristics of the immunoreactant recognized by the anti-BCNT-C Ab are also consistent with those previously reported for GS (see Supplementary Fig. S8 online, ref. 20, 21). Because we confirmed that the anti-BCNT-C Ab really recognizes GS, a protein not homologous to Bcnt, we conducted epitope analysis of GS to investigate the mechanisms for such a peculiar cross-reactivity.

Epitope analysis of GS for cross-reactivity of the anti-BCNT-C Ab

First, we prepared several deletion mutants of GS protein in *E. coli* to identify the region recognized by the anti-BCNT-C Ab on a GS molecule. Representative results of Western blotting are shown in Fig. 4A. While a clear positive reaction was detected for a GS mutant lacking almost half of the N-terminal amino acids (D2), no reaction was observed for a GS mutant lacking C-terminal amino acids 332–373 (D1), strongly suggesting that the epitope is localized in the C-terminal region of GS. To further narrow down the predicted epitope region, we employed competition experiments using four peptides with short overlapping amino acids (P1–P4 in Fig. 4B, upper panel, see Supplementary Fig. S9 online) derived from the C-terminal region of GS. Western blot analysis using the preincubated anti-BCNT-C Ab with each peptide or the antigen peptide showed that only P1 among the four peptides strongly inhibited the cross-reactivity of the Ab with GS similar to the antigen peptide (Fig. 4B, lower panel). These results strongly suggest that the epitope must be included in the P1 peptide sequence. Furthermore, to identify key amino acids involved in the cross-reaction, we prepared graded deletion mutants of GS and examined their reactivity with the anti-BCNT-C Ab by Western blotting. As shown in Fig. 4C, the cross-reactivity was clearly observed for the Del-338 mutant but was conspicuously absent for the Del-336

mutant. This result indicates that key amino acids for the cross-reaction are limited to RTVGQEKKGYFE in the P1 peptide sequence. We noticed that KKGYFE of GS is similar to KEGYIE of antigen peptide for anti-BCNT-C Ab, especially GYFE of GS and GYIE of BCNT (Fig. 4C, lower left), and we examined the cross-reactivity with a focus on GYFE of GS using alanine substitution mutants of GS. Whereas weak but significant reactivity was detected for the F337A mutant, no reactivity was detected for other alanine substitution mutants including E338A (Fig. 4D). These results clearly show that GYFE (hereinafter called "the core amino acids") is required for the cross-reactivity of the anti-BCNT-C Ab with GS. As shown in Fig. 4D, mutant A4A also lacked reactivity with the Ab, suggesting the possibility that K333 and/or K334 may be included in the epitope.

Evaluation of the affinity of the anti-BCNT-C Ab to GS and BCNT

Next, we evaluated the affinity of the anti-BCNT-C Ab to GS and BCNT by semi-quantitative Western blotting of Flag-tagged mouse GS (F-mGS) and Bcnt (F-mBcnt) normalized with an anti-Flag-tag Ab (Fig. 5). A subpopulation of the polyclonal anti-BCNT-C Ab, which cross-reacts with GS, was fractionated by a column of agarose resin cross-linked with GS(332-346) peptide including the core amino acids, GYFE (shown in Fig. 5A). The binding fraction (GYFE fraction) was isolated in about 20% yield as a protein weight, indicating that the subpopulation of the polyclonal anti-BCNT-C Ab cross-reacting with the core amino acids, GYFE, of GS occupied about one fifth. Consistent with these results, the normalized immunoreactivity of the anti-BCNT-C Ab for F-mGS is estimated to be approximately one fifth compared with that for F-mBcnt (Fig. 5A, left). Furthermore, the GYFE fraction from the anti-BCNT-C Ab reacted to F-mGS and F-mBcnt

with equivalent immunoreactivity (Fig. 5A, right), whereas the flow-through fraction did not react to F-mGS (Fig. 5A, middle). These results suggest that the subpopulation of the polyclonal anti-BCNT-C Ab cross-reacting with GS, derives from only Ab molecules recognizing epitopes including the core amino acids, GYFE. On the other hand, the immunoreactivity of the GYFE fraction was inhibited with a peptide EELAIHNRGKEGYIERKA-NH_2, BCNT-C (GYIE) in Fig. 5B, but not with its mutant replacing GYIE with Alanine, BCNT-C (All-Ala) in Fig. 5B. This result shows that the subpopulation of the polyclonal anti-BCNT-C Ab cross-reacting with GYFE of GS specifically reacts to GYIE of Bcnt. We also isolated the subpopulation from the anti-BCNT-C Ab by affinity purification on the membrane using full-length recombinant proteins of wild-type GS or All-Ala GS mutant expressed in *E. coli* (see Supplementary Fig. S10 online). These results indicate that a certain subpopulation exists in the polyclonal anti-BCNT-C Ab that reacts to both GS and Bcnt with similar affinities, and that GYFE of GS composes a core sequence of the epitope required for the cross-reactivity of the anti-BCNT-C Ab. It should be noted that the Ab recognized both GS and Bcnt with similar immunoreactivities in spite of the difference between Phe (GYFE) for GS and Ile (GYIE) for Bcnt.

Spatial similarity of stereochemical environment between GS epitope and BCNT-C antigen

To understand the similar immunoreactivities of the anti-BCNT-C Ab against both GS and Bcnt in spite of the one amino acid difference of the core sequence, we considered the possibility that the cross-reactivity was caused by a spatial similarity of the stereochemical

environment between the BCNT-C antigen peptide and a region of GS including GYFE. As described above, weak but significant immunoreactivity was detected for an alanine substitution mutant of GS (F337A) but not for other alanine-substitution mutants (Fig. 4D). Taken together, we conjectured that F337 of GS is not directly involved in the interface of the antigen-Ab interaction between GS and the anti-BCNT-C Ab. To investigate how F337 actually participates in the antigen-Ab interaction, we prepared substitution mutants of F337 of GS with various amino acids having different steric bulks (the amount of space that side chain atoms of amino acids occupy) or polarities and tried to semi-quantitatively analyze the effect of those substitutions on stereochemical and physiochemical changes of GS by Western blotting. As shown in Fig. 6A, when four hydrophobic and branched-chain amino acids were substituted in the order of steric bulk from isoleucine to alanine, the cross-reactivity of the anti-BCNT-C Ab was attenuated with an accompanying decrease in the steric bulk of the substituted amino acid, but it was not completely abolished. This result further supports our conjecture that F337 of GS is not directly engaged in the contact site of the antigen-Ab interaction but importantly constitutes the stereochemical environment of the epitope. Fig. 6B shows a 3D layout of a region around GYFE of the common epitope of GS, which was drawn based on X-ray crystallography data of GS (PDB code 2OJW, ref. 22). Note that Phe (F337) resides on the opposite side of neighboring Y336 and E338, occupying rather a space by its aromatic ring. This arrangement of F337 supports our speculation above that this amino acid is not directly involved in the contact site of the antigen-Ab interaction, but has a steric bulk effect on the stereochemical environment of the epitope. On the other hand, the cross-reactivity was approximately equal to the F337Y mutant as compared with the wild-type GS (F337) (Fig. 6A). These data suggest that the polarity of

tyrosine, a hydroxyl group adduct of phenylalanine, does not significantly affect the stereochemical environment of the epitope. In conclusion, the results obtained strongly suggest that the spatial similarity of stereochemical environments formed by the core amino acids of common epitopes of GS and Bcnt is critically important in the cross-reaction of the anti-BCNT-C Ab with GS.

DISCUSSION

In this paper, we showed that the anti-BCNT-C Ab quite specifically cross-reacted with GS, which is a protein not homologous to Bcnt, in addition to the original target molecule. GS is present predominantly in the brain where it participates in the metabolic regulation of glutamate (14, 23). The anti-BCNT-C Ab recognized exogenous GS expressed in HEK cells and also showed strong immunoreactivity in astrocyte of adult rat cerebellum (see Supplementary Fig. S7 online), consistent with the localization of GS in brain previously reported (20). It is most likely that strong immunoreactivities of the anti-BCNT-C Ab in Leydig cells of bovine testis (24) and astrocytes of rat brain, we had conjectured as Bcnt/Cfdp1 so far, are immunoreactivities for endogeneous GS (20, 25). In addition to its strong immunoreactivity, this Ab detects GS under different denaturing conditions: thermally and reductively denatured GS in Western blotting, chemically fixed GS *in situ* in immunocytochemistry and immunohistochemistry, and a more naïve form of GS in aqueous solution in immunoprecipitation. These data strongly suggest that the epitope including GYFE, which is reactive to the anti-BCNT-C Ab, localizes on the surface of the native GS.

Although no significant homology of amino acid sequences was found between Bcnt and

GS, we were able to identify the core amino acid sequence, GYFE, localized in a limited region of the C-terminal of GS, that is necessary for the cross-reactivity of the anti-BCNT-C Ab (Fig. 4C). We conjectured that this sequence has an epitope in common with GYIE in the immunogen peptide for the anti-BCNT-C Ab. This conjecture implies that replacement of an aromatic amino acid, Phe, of GYFE with a branched chain amino acid, Ile, maintains the epitope environment. This assumption is consistent with the results shown in Fig. 6A. GS mutants with a substitution of Phe with Ile, Leu, or Tyr showed comparable cross-reactivity of the anti-BCNT-C Ab, whereas replacement of Phe with a smaller amino acid, Val or Ala, resulted in less reactivity with the Ab, indicating that the steric bulk of each amino acid is critical for the reactivity of the epitope with the Ab. Moreover, the result also indicates that the steric bulk of the aromatic ring of F337 or Tyr of the F337Y mutant of GS, but not the polarity of amino acids, is important to maintain the stereochemical environment for the cross-reactivity of the anti-BCNT-C Ab with GS. Therefore, this result strongly suggests that Phe in GYFE (GS) and Ile in GYIE (Bcnt) play a role in maintaining the spatial arrangement of the common epitope but are not directly involved in the contact site of the antigen-Ab interaction. This speculation is consistent with a previous report that F337 resides on the opposite side of neighboring Y336 and E338 (22), and that each is involved in binding ATP and glutamate, a substrate of GS, respectively, in the active site of the GS complex (Fig. 6B). It is also reported that though F337 in the active site of GS is highly conserved among species, it is replaced by Leu in *Mycobacterium tuberculosis* and Ile in *Salmonella typhimurium* (13, 22), suggesting that both the function and stereochemical environment of the active site of GS are maintained even if F337 is replaced with Leu or Ile. These data lead to the speculation that the core amino acids of the common epitopes, GYFE

located in this site of GS and GYIE of Bcnt, constitute similar stereochemical environments.

It is also noted that the conjectured core amino acids of the common epitope are bracketed by charged amino acid residues such as lysine and arginine (Fig. 4C, lower left). The hydrophilicity of these charged residues may bring the core amino acids of each epitope to the surface of GS or BCNT. Elucidation of this characteristic of epitopes awaits further studies.

It is well known now that whereas cross-reactions of Abs with structurally unrelated peptides of off-targets frequently occur in binding assays of Abs using small peptides, such cross-reactions rarely occur in those using protein fragments and in Western blot analysis (9). It has been generally believed that Abs recognize linear epitopes of target molecules under the denatured condition of SDS-PAGE, but some Abs can also react to proteins via conformational epitopes, probably depending on the protein renaturation on transferred membranes after denaturation in SDS-PAGE (26, 27). The anti-BCNT-C Ab reacted to each GS mutant for Phe of GYFE in a steric bulk-dependent manner (Fig. 6A), in contrast to the loss of immunoreactivities in alanine-substituted mutants for Gly, Tyr, and Glu of GYFE (Fig. 4D). These results reflect differences in whether these amino acids exist at the contact site for the Ab (Gly, Tyr, and Glu) or not (Phe). This result, which is consistent with the configuration of the core amino acids displayed in the 3D structure of GS (Fig. 6B), suggests that Bcnt and GS are partially renatured and that their conformations are partially maintained on transferred membranes in Western blot analysis. Such renaturation of Bcnt and GS is inferred to be one of the reasons why we were able to elucidate the similarity of stereochemical environments between GYIE of Bcnt and GYFE of GS by Western blot

analysis.

Recent reports including this paper indicate that cross-reactions of Abs with off-target molecules cannot be avoided, even when choosing a very specific amino acid sequence as an immunogen (9, 28). Therefore, we have to pay attention to the fact that Abs do not necessarily recognize only “on-targets” but potentially react with “off-targets” that comprise similar stereochemical or physicochemical environments as that of on-target molecules. Recent advances in computational and bioinformatic analyses using extensively accumulated data of epitope-paratope interfaces make it possible to group T cell receptors of common specificity using grouping algorithms of lymphocyte interactions by paratope hotspots (29). Furthermore, these advances have also revealed the functional redundancy of epitopes based on physicochemical similarities at a level of amino acid residues involved in the antigenic cross-reaction (30). In addition to these accumulated data, further comprehensive understanding of the physicochemical environment of epitope-paratope interfaces including both “on-targets” and “off-targets” may make it possible to evaluate and predict more accurately antigenicity and immunogenicity, and these advances will bring about therapeutic Abs and vaccines with high specificity and without adverse effects.

MATERIALS AND METHODS

Ethical approval

All the genetic recombination experiments and all the animal experiments in the present study were approved by the Genetic Recombination Experiment Safety Committee and the Animal Care and Use Committee, respectively, of Tokushima Bunri University.

Reagents

Generation of anti-BCNT-C Ab, anti-mBcnt-N Ab, and anti-Bcnt-Cter Ab is described in Supplementary Materials and Methods. Abs against Flag tag (DYKDDDDK tag; 014-22383) and GS (GTX109121) were obtained from Wako Pure Chemical Industries and GeneTex, respectively. Unstained or prestained SDS-PAGE molecular weight standard markers (Bio-Rad) were used differently for each experiment.

Construction of expression vectors for mGS and mBcnt

A plasmid carrying Myc-tagged mGS was a gift from Dr. T. Araki (Natl. Inst. Neurosci., Tokyo) (17). Using cloning primers (see Supplementary Table S3-1 online), mGS and mBcnt cDNAs were amplified from cDNA, which was reverse transcribed from the total RNA of the brain of an adult male C57BL/6J mouse using dT-primer and inserted into mammalian expression vectors, as described in Supplementary Materials and Methods. All the expression vectors for GS mutants were constructed and used as described in Supplementary Materials and Methods. Among GS mutants, F337A and an all-Ala replacement mutants of 355-358 (GYFE) were confirmed by MALDI-TOF MS and LC-MS/MS after digestion with Asp-N protease or trypsin.

Cell culture and transfection

HEK-293T cells were maintained at 37°C, 5% CO_2, in high glucose (4.5 g D-glucose/L) DMEM-GlutaMAX-I (Thermo Fisher Scientific) supplemented with 10% Fetal bovine serum and 50 mg/L gentamicin. HEK-293T cells were plated in 6-well plates at ~60%

confluency and transfected with an expression vector DNA (0.5 μg/well) using PLUS reagent (0.5 μL/well) and Lipofectamine LTX reagent (1.2 μL/well) (Thermo Fisher Scientific).

Preparation of extracts of *E. coli* expressing GS and its mutants

Culture of *E. coli* and induction of coded molecules by cDNAs in pCold II vector were described previously (11). *E. coli* were harvested and washed with PBS by centrifugation at 5,000 x g for 5 min. After addition of 50 μL H_2O per pellet derived from 1 mL culture, resuspended cells were mixed with an equal volume of 2-fold concentrated lysis buffer (40 mM Tris-HCl buffer (pH 7.6) containing 2% SDS) and immediately boiled for 3 min. Extracts were sonicated by a Bioruptor (BM Equipment) in an ice-water bath (15 x 10 s pulses at 10 s intervals) and followed by centrifugation at 28,000 x *g* for 10 min. Supernatants were subjected to Western blot analysis.

Immunoblotting

Procedures of SDS-PAGE and blotting onto membranes were essentially the same as previously described (11) except that a low fluorescence membrane (Immobilon-FL, Merck) was used in the fluorescent method. Immunodetections were carried out by chemiluminescent, colorimetric, and fluorescent methods using horseradish peroxidase (HRP)-, alkaline phosphatase (AP)- and fluorescent dye-conjugated secondary Abs, respectively. The secondary IgGs used are listed in Supplementary Materials and Methods. Tris-buffered saline supplemented with 0.1% Tween 20 (TBT) was used as a wash buffer in all three methods. In the chemiluminescent method, immunoreaction was carried out in

TBT containing 5% or 10% skim milk and the immunoreactivity was detected by a scanner (GeneGenome, Syngene Bio Imaging) using Immobilon (Merck), ImmunoSTAR Zeta, or LD (Wako Chem.) as a substrate. In the colorimetric method, immunoreaction was carried out in TBT containing 0.2% casein (I-Block, Thermo Fisher Scientific), and the immunoreactivity was detected by an LAS-3000 Imager (Fuji-Film) using 5-bromo-4-chloro-3-indolyl phosphate/nitro blue tetrazolium (Sigma) as a chromogenic substrate. In the fluorescent method, the immunoreaction was carried out in TBT containing 0.2% casein, and the immunoreactivity was detected by an Odyssey 700 CLX Infrared Imaging system (LI-COR Biosciences). Membranes of bound Abs for reprobing were stripped by shaking membranes in 62.5 mM Tris-HCl (pH 6.8), 2% SDS, and 100 mM 2-mercaptoethanol for 1 h at 70°C. To avoid inhibition by former enzymatic reaction products, reprobing was carried out only after detection by the fluorescent method.

Isolation of a subpopulation of the polyclonal anti-BCNT-C Ab, which cross-reacts with GYFE of GS

GS(332-346) peptide, acetylated-EKKGYFEDRRPSANC-COOH (0.75 mg, 91.2% purity, obtained from AnyGen), was coupled to 1 mL of SulfoLink Coupling Resin (Thermo Fisher Scientific) packed in a column (MoBiTec) according to a manufacture's protocol. The anti-BCNT-C Ab of 4.45 mg diluted in 10 mL with Binding/Wash buffer (20 mM PBS (pH 7.0) containing 0.05% sodium azide) passed through the column 5 times and a flow-through fraction was obtained. After washing the resin with 50 mL of Binding/Wash buffer, the bound fraction (GYFE fraction) was eluted with 6 mL of 100 mM Glycine (pH 3.0) and immediately neutralized with Neutralization buffer (1 M Tris-HCl (pH 8.5), 1.5 M NaCl, 5

mM EDTA). The protein concentrations of the binding fraction (GYFE fraction) and the flow-through fraction were determined by Bicinchoninic acid (BCA) assay kit (Thermo Fisher Scientific) using bovine serum albumin as a standard after buffer exchange and concentration by ultrafiltration using Amicon® Ultra-4, 30kDa (Merck Millipore). Their protein concentrations were 0.78 mg and 2.80 mg, respectively, with total recovery of 80.7%.

Competitive experiments for evaluation of the specificity of the GYFE fraction to GYIE of Bcnt

Extracts of HEK293T cells expressing Flag-mGS (F-mGS) or Flag-mBcnt (F-mBcnt) were adjusted to show an equal immunoreactivity to anti-Flag Ab, then subjected to Western blotting. The membrane was first reacted with anti-Flag Ab to confirm loading amounts. After stripping, the membranes were reporobed with the GYFE fraction (1.33 nM as a IgG) pre-incubated with BCNT-C peptide (NH2-EELAIHNRGKEGYIERKA-NH2, 92.3% purity, from AnyGen) or BCNT-C peptide replacing GYIE with Alanine (NH2-EELAIHNRGKEAAAARKA-NH2, 98.3% purity, from AnyGen) at a final concentration of 25 μM.

Mass spectroscopy analysis

The CBB stained protein bands in the gel were excised and de-stained. The protein bands were digested with trypsin (TPCK-treated, Worthington Biochemical) at 37°C for 12 h in 50 mM Tris-HCl (pH 8.0). The digests were analyzed by nano LC–MS/MS using a Q Exactive mass spectrometer exactly as described previously (11).

Immunocytochemistry

Transfected HEK-293T cells (see "*Cell culture and transfection*") were plated on glass slides (Millicell EZ 4-well glass slides, Merck Millipore) coated with poly-D-lysine (Sigma-Aldrich). After 30 h, the cultured cells were fixed with ice-cold 4% paraformaldehyde (PFA) in 0.1 M phosphate-buffer (PB, pH 7.2) for 15 min and subjected to double immunostaining as described in the "*Immunohistochemistry*" section of Supplementary Materials and Methods. First, cells were stained with guinea pig anti-BCNT-C Ab (1:500) and visualized by FITC-conjugated donkey anti-guinea pig IgG (1:50, Jackson ImmunoResearch). The second staining was done with rabbit anti-GS Ab (1:1000, GeneTex) visualized by Cy3-conjugated donkey anti-rabbit IgG (1:200, Jackson ImmunoResearch). Immunostained cells were examined by confocal laser scanning microscopy (FV1000, Olympus).

ACKNOWLEDGEMENT

We thank Dr. Toshiyuki Araki for gift of GS-myc plasmid and Drs. Mathias Uhlen, Howard W. Davidson, Eiji Kinoshita, Ichiro N. Maruyama, Daniel M. Roberts, and Masami Kusunoki for discussion. We also thank Rie Fujii and Dr. Tomoko Warita for supporting experimental works. We are grateful to Dr. Makoto Iwata for critical reading the manuscript and comments.

AUTHOR CONTRIBUTIONS

N.K., S.I. and S.Y.S. conceived the study and wrote the manuscript. K.N., S.I., C.K. and S.Y.S. executed experiments and T.S. and N.D. conducted MS analysis. T.K. performed 3D analysis of epitope and Y.K., M.S., O.U. and Y.O.I. contributed reagents. All authors reviewed the manuscript.

ADDITIONAL INFORMATION

Competing interests: The authors declare no competing interests.

REFERENCES

1. Chiavenna, S. M., Jaworski, J. P. & Vendrell, A. State of the art in anti-cancer mAbs. *J Biomed Sci* **24**, 15 (2017).
2. Sharma, P. & Allison, J. P. The future of immune checkpoint therapy. *Science* **348**, 56-61 (2015).
3. Thomas, A., Teicher, B. A. & Hassan, R. Antibody-drug conjugates for cancer therapy. *Lancet Oncol* **17**, e254-e262 (2016).
4. Rossi, D., Modena, V., Sciascia, S. & Roccatello, D. Rheumatoid arthritis: Biological therapy other than anti-TNF. *Int Immunopharmacol* **27**, 185-188 (2015).
5. Zampeli, E., Vlachoyiannopoulos, P. G. & Tzioufas, A. G. Treatment of rheumatoid arthritis: Unraveling the conundrum. *J Autoimmun* **65**, 1-18 (2015).
6. Feyen, O. *et al.* Off-target activity of TNF-alpha inhibitors characterized by proteinbiochips. *Anal Bioanal Chem* **391**, 1713-1720 (2008).
7. Mahmud, M. N. *et al.* A multispecific monoclonal antibody G2 recognizes at least three completely different epitope sequences with high affinity. *Protein Sci* **26**, 2162-2169 (2017).
8. Koenig, P., Sanowar, S., Lee, C. V. & Fuh, G. Tuning the specificity of a Two-in-One Fab against three angiogenic antigens by fully utilizing the information of deep mutational scanning. *MAbs* **9**, 959-967 (2017).
9. Forsström, B. *et al.* Proteome-wide epitope mapping of antibodies using ultra-dense peptide arrays. *Mol Cell Proteomics* **13**, 1585-1597 (2014).
10. Peng, H. P., Lee, K. H., Jian, J. W. & Yang, A. S. Origins of specificity and affinity in antibody-protein interactions. *Proc Natl Acad Sci USA* **111**, E2656-2665 (2014).

11. Iwashita, S. *et al.* Mammalian Bcnt/Cfdp1, a potential epigenetic factor characterized by an acidic stretch in the disordered N-terminal and Ser250 phosphorylation in the conserved C-terminal regions. *Biosci Rep* **35**, e00228 (2015).
12. Iwashita, S. *et al.* A transposable element-mediated gene divergence that directly produces a novel type bovine Bcnt protein including the endonuclease domain of RTE-1. *Mol Biol Evol* **20**, 1556-1563 (2003).
13. Eisenberg, D., Gill, H. S., Pfluegl, G. M. & Rotstein, S. H. Structure-function relationships of glutamine synthetases. *Biochim Biophys Acta* **1477**, 122-45 (2000).
14. Rose, C. F., Verkhratsky, A. & Parpura, V. Astrocyte glutamine synthetase: pivotal in health and disease. *Biochem Soc Trans* **41**, 1518-1524 (2013).
15. Peri, C., Morra, G. & Colombo, G. Surface energetics and protein-protein interactions: analysis and mechanistic implications. *Sci Rep* **6**, 24035 (2016).
16. Urayama, O., Murakoshi, T. & Ikawa, Y. K rev-1 protein is abundantly expressed in the rat spinal cord. *Biochem Biophys Acta* **1243**, 446-452 (1995).
17. Saitoh, F. & Araki, T. Proteasomal degradation of glutamine synthetase regulates Schwann cell differentiation. *J Neurosci* **30**, 1204-1212 (2010).
18. Kamei, Y. *et al.* FOXO1 activates glutamine synthetase gene in mouse skeletal muscles through a region downstream of 3'-UTR: possible contribution to ammonia detoxification. *Am J Physiol Endocrinol Metab* **307**, E485–E493 (2014).
19. Norenberg, M. D. & Martinez-Hernandez, A. Fine structural localization of glutamine synthetase in astrocytes of rat brain. *Brain Res* **161**, 303-310 (1979).
20. Kulijewicz-Nawrot, M., Syková, E., Chvátalm, A., Verkhratsky, A., & Rodríguez, J. J. Astrocytes and glutamate homoeostasis in Alzheimer's disease: a decrease in glutamine

synthetase, but not in glutamate transporter-1, in the prefrontal cortex. *ASN Neuro* **5**, 273-282 (2013).

21. Huyghe, D. *et al.* Glutamine synthetase stability and subcellular distribution in astrocytes are regulated by γ-aminobutyric type B receptors. *J Biol Chem* **289**, 28808-28815 (2014).
22. Krajewski, W. W. *et al.* Crystal structures of mammalian glutamine synthetases illustrate substrate-induced conformational changes and provide opportunities for drug and herbicide design. *J Mol Biol* **375**, 217-228 (2008).
23. Jayakumar, A. R. & Norenberg, M. D. Glutamine synthetase: role in neurological disorders. *Adv Neurobiol* **13**, 327-350 (2016).
24. Iwashita, S., Nakashima, K., Sasaki, M., Osada, N. & Song S. Y. Multiple duplication of the bucentaur gene family, which recruits the APE-like domain of retrotransposon: Identification of a novel homolog and distinct cellular expression. *Gene* **435**, 88-95 (2009).
25. Holash, J. A., Harik, S. I., Perry, G. & Stewart, P. A. Barrier properties of testis microvessels. *Proc Natl Acad Sci USA* **90**, 11069-11073 (1993).
26. Forsström, B. *et al.* Dissecting antibodies with regards to linear and conformational epitopes. *PLoS One* **10**, e0121673 (2015).
27. Zhou, Y. H., Chen, Z., Purcell, R. H. & Emerson, S. U. Positive reactions on Western blots do not necessarily indicate the epitopes on antigens are continuous. *Immunol Cell Biol* **85**, 73-78 (2007).
28. McIntush, E. W. Response: 'Antibody crossreactivity between the tumour suppressor PHLPP1 and the proto-oncogene β-catenin' *EMBO reports* **14**, 494-496 (2013).

29. Glanville, J. *et al.* Identifying specificity groups in the T cell receptor repertoire. *Nature* **547**, 94-98 (2017).

30. Caoili, S. E. Expressing redundancy among linear-epitope sequence data based on residue-level physicochemical similarity in the context of antigenic cross-reaction. *Adv Bioinformatics* 1276594 (2016).

31. Kurien, B.T. Affinity Purification of Autoantibodies from an Antigen Strip Excised from a Nitrocellulose Protein Blot. *Protein Blotting and Detection. Methods in Molecular Biology (Methods and Protocols)* (eds. Kurien, B., Scofield, R.) vol 536 (Humana Press, Totowa, NJ, 2009)

32. Fang, L. Antibody purification from Western blotting. *Bio-protocol* **2**, e133 (2012).

33. Kaspersen, J. D., Søndergaard, A., Madsen, D. J., Otzen, D. E. & Pedersen, J. S. Refolding of SDS-unfolded proteins by nonionic surfactants. *Biophys J* **112**, 1609-1620 (2017).

34. Yamakuni, T. *et al.* Isolation and immunohistochemical localization of a cerebellar protein. *Neurosci Lett* **45**, 235-240 (1984).

FIGURE LEGENDS

Figure 1: The 43 kDa protein as a major reactant in Western blotting with the anti-BCNT-C Ab.

Western blotting was conducted with the polyclonal anti-BCNT-C, mBcnt-N and Bcnt-Cter Abs, as shown under each figure, using extracts of mouse brain and HEK293T cells transfected with mouse bcnt expression vector, as shown on the top of each lane. Note that a strong signal of the 43 kDa in mouse brain extracts was detected only by anti-BCNT-C Ab,while the recombinant mouse Bcnt protein was detected by all Abs.

Figure 2: The isolated 43 kDa protein from bovine brain.

The 43 kDa protein was isolated from bovine brain (see Supplementary Materials and Methods and Fig. S4 online). The eluted fractions from anti-BCNT-C Ab-linked agarose with the most positive immunoreactivities detected by a colorimetric method (left panel shows a representative fraction) were pooled, subjected to acetone precipitation, separated by SDS-PAGE and stained with Coomassie brilliant blue (right panel).

Figure 3: Specific recognition of GS by the anti-BCNT-C Ab.

(A) HEK239T cells were transfected with plasmids encoding Flag-tagged or tag-free mouse GS (F-mGS, mGS), mouse Bcnt/Cfdp1 (F-mBcnt, mBcnt), respectively, and their extracts were subjected to Western blotting by double probing with the anti-BCNT-C Ab (upper panel) and the anti-GS Ab (lower panel). Both signals were detected by a fluorescent method. None and Vector indicate a vehicle and a blank vector control, respectively. ***(B)***

HEK293T cells were transfected with plasmids encoding mGS (A-C, E-G, I-K) or mBcnt (D, H, L) and immunostained with the anti-BCNT-C Ab (A, D), anti-mBcnt-N Ab (E, H), anti-Bcnt-Cter Ab (I, L), or anti-GS Ab (B, F, J). C, G, and K are merged figures of A with B, E with F, and I with J, respectively. Scale bars indicate 50 μm. The asterisk indicates non-specific band.

Figure 4: Core sequence of the epitope recognized by the anti-BCNT-C Ab on GS.

(A) Upper panel: Schematic structures of full-length mGS (GS-WT) and its two deletion mutants (GS-D1 and GS-D2). Orange and light green boxes indicate His-tag and the amino acid residues of GS, respectively. Figures above each box represent the number of amino acids constituting the coding region. Lower panels: Western blotting of extracts of *E. coli* expressing GS-WT (WT), GS-D1 (D1), or GS-D2 (D2) and non-induced *E. coli* (NI) by the anti-BCNT-C Ab. The membrane was reprobed with anti-His-tag Ab to confirm the amounts of applied GS mutants. ***(B)*** Upper panel: Amino acid sequences of the antigen peptide (Ag) used to raise the anti-BCNT-C Ab, the C-terminal region of mouse GS and four peptide fragments (P1–P4) used for competition experiment. To avoid dimer formation, two cysteine residues of GS peptides were replaced with serine (red letters). Lower panel: Western blotting of extracts of *E. coli* expressing GS-WT with the anti-BCNT-C Ab preincubated with 5 μM of Ag, each GS peptide (P1-P4) or DMSO for negative control (-) as indicated on the top of each lane. An anti-His-tag Ab was used in the same way as in *(A)*. ***(C, D)*** Western blotting of extracts of *E. coli* expressing GS-WT, six deletion GS mutants *(C)* or alanine substitution mutants of one or four amino acids (in red letters) *(D)* with the anti-BCNT-C or anti-His-tag Ab in the same way as in *(A)*. Red letters in *(C)* indicate an

amino acid sequence in antigen peptide for anti-BCNT-C Ab, which is similar to KKGYFE in P1 peptide of GS.

Figure 5: Isolation of the subpopulation of the polyclonal anti-BCNT-C Ab, which cross-reacts with GYFE of GS, and its affinity to GYFE of GS and GYIE of Bcnt.

The subpopulation of the polyclonal anti-BCNT-C Ab cross-reacting with GYFE of GS was isolated by an affinity column using GS(332-346) peptide (see MATERIALS AND METHODS). Extracts of HEK293T cells expressing Flag-mGS (F-mGS) or Flag-mBcnt (F-mBcnt), which were adjusted to show an equal immunoreactivity to anti-Flag Ab, were subjected to Western blotting. The membrane was first reacted with anti-Flag Ab to confirm loading amounts (lower panels in *A* and *B*). After stripping, ***(A)*** the membranes were reprobed with untreated (Whole), flow-through (Flow Through) or affinity-purified (GYFE fraction) Abs from the polyclonal anti-BCNT Ab by GS(332-346) peptide affinity column. ***(B)*** The membranes were reporobed with the GYFE fraction pre-incubated using BCNT-C peptide (BCNT-C(GYIE)) or BCNT-C peptide replacing GYIE with Alanine (BCNT-C (All-Ala)). The asterisk indicates non-specific band.

Figure 6: Role of Phe-337 in stereochemical environment of core amino acids of common epitope.

(A) Western blotting was carried out as described in Fig. 4A using the extracts *of E. coli* expressing wild type GS (GS-WT) or its substitution mutants of F337 with various amino acids as shown in upper left panel and on the top of each lane. Blue letters indicate the core amino acids and red letters denote the respective substituted amino acids. ***(B)*** The inserted

picture at the left top shows top-view of one pentameric ring for human GS consisted of two pentameric rings (PDB code 2OJW). The red circle depicts the location of the one of the core amino acids, GYFE, in the five identical subunits. 3D-layout of the KKGYFEDR peptide in GS displaying the spatial arrangement of the common epitope region. This layout was drawn using Waals (Altif Laboratories) based on a data of X-ray crystallography (PDB code 2OJW, ref. 22) of human GS.

Fig 1

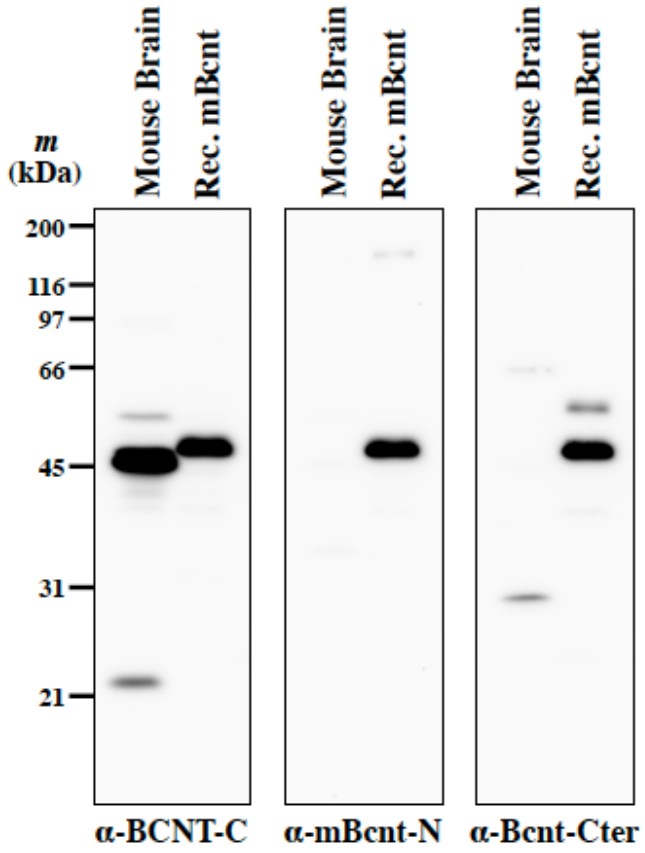

m
(kDa)
Mouse Brain
Rec. mBcnt
Mouse Brain
Rec. mBcnt
Mouse Brain
Rec. mBcnt
200
116
97
66
45
31
21
α-BCNT-C
α-mBcnt-N
α-Bcnt-Cter

Fig 2

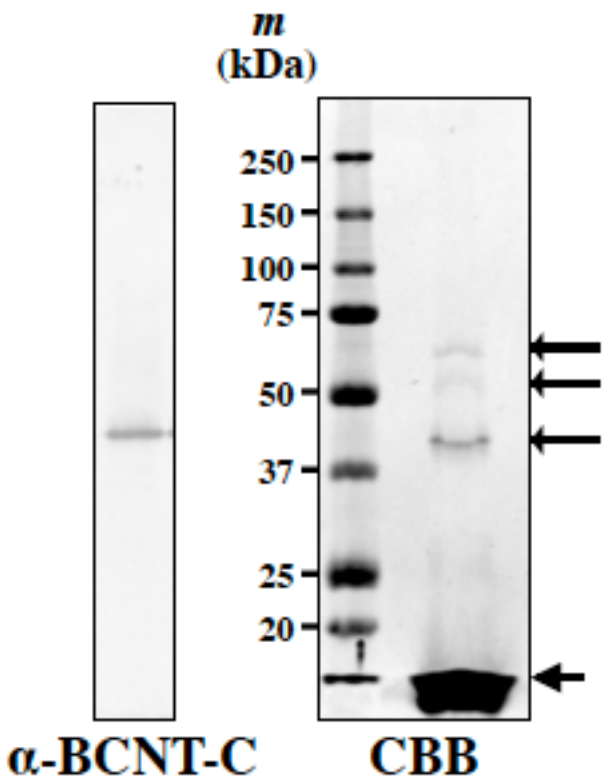

m
(kDa)
250
150
100
75
50
37
25
20
α-BCNT-C
CBB

Fig 3

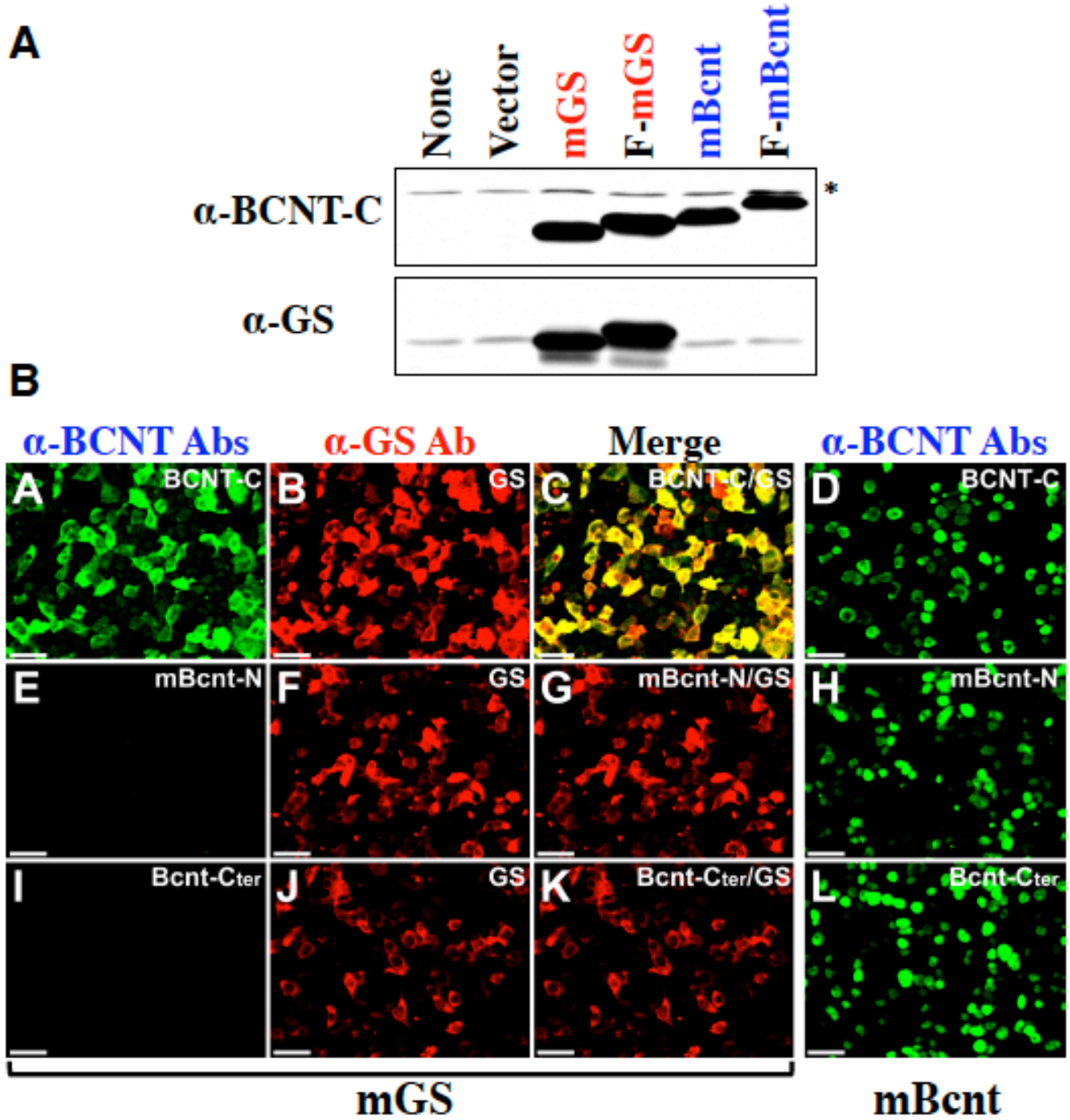

A
None
Vector
mGS
F-mGS
mBcnt
F-mBcnt
α-BCNT-C
*
α-GS
B
α-BCNT Abs
α-GS Ab
Merge
α-BCNT Abs
A BCNT-C
B GS
C BCNT-C/GS
D BCNT-C
E mBcnt-N
F GS
G mBcnt-N/GS
H mBcnt-N
I Bcnt-Cter
J GS
K Bcnt-Cter/GS
L Bcnt-Cter
mGS
mBcnt

Fig 4

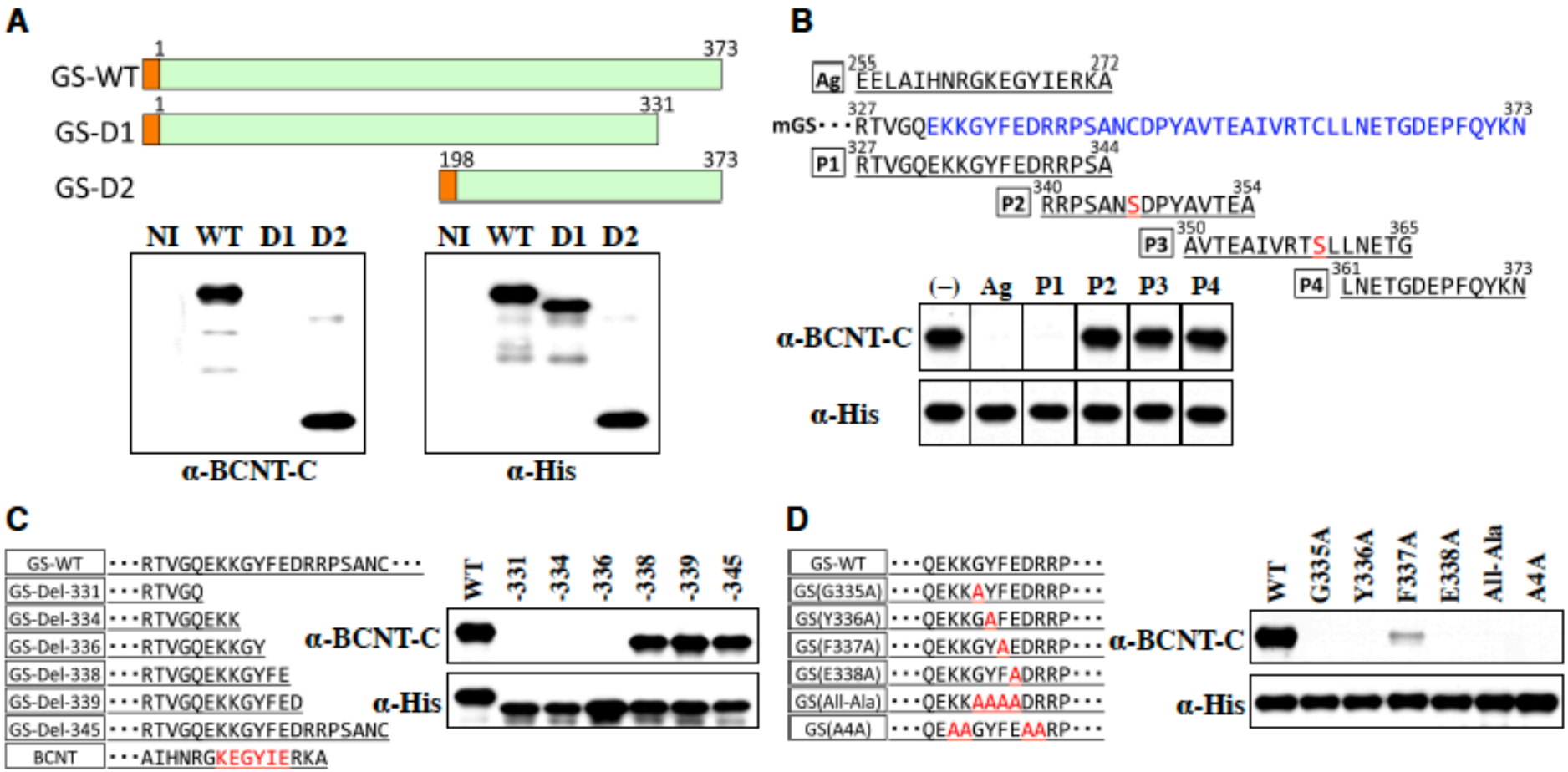
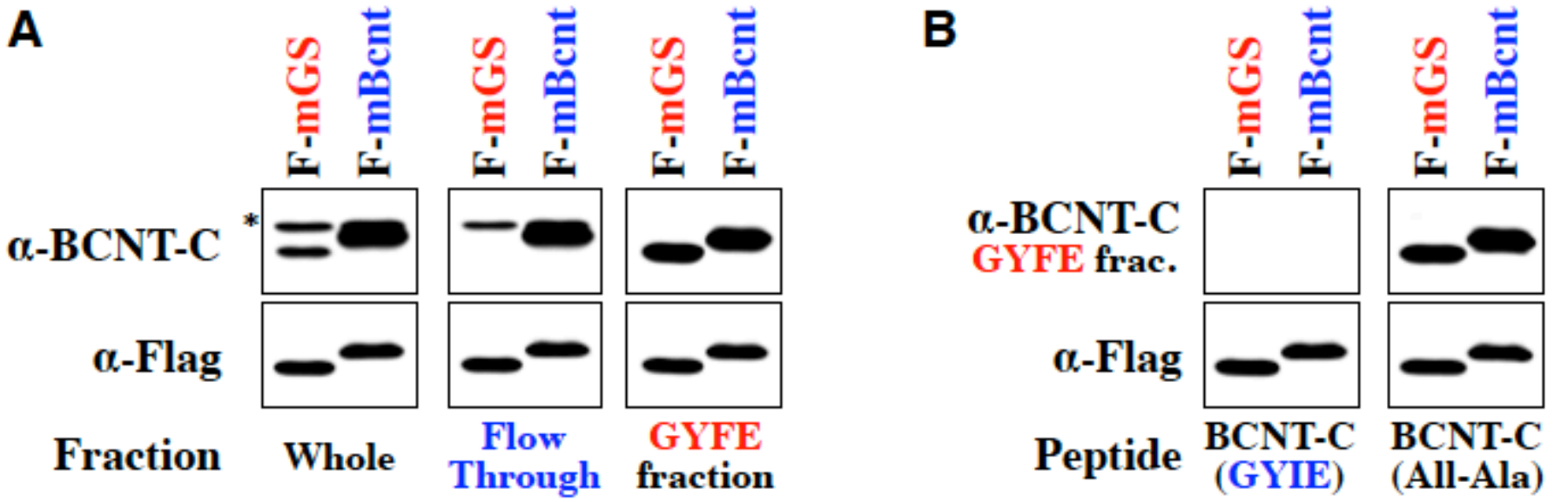

A
GS-WT
GS-D1
GS-D2
1
373
331
198
NI WT D1 D2
α-BCNT-C
α-His
B
Ag 255 EELAIHNRGKEGYIERKA 272
mGS···RTVGQEKKGYFEDRRPSANCDPYAVTEAIVRTCLLNETGDEPFQYKN
P1 327 RTVGQEKKGYFEDRRPSA 344
P2 340 RRPSANSDPYAVTEA 354
P3 350 AVTEAIVRTSLLNETG 365
P4 361 LNETGDEPFQYKN 373
(–) Ag P1 P2 P3 P4
α-BCNT-C
α-His
C
GS-WT ···RTVGQEKKGYFEDRRPSANC···
GS-Del-331 ···RTVGQ
GS-Del-334 ···RTVGQEKK
GS-Del-336 ···RTVGQEKKGY
GS-Del-338 ···RTVGQEKKGYFE
GS-Del-339 ···RTVGQEKKGYFED
GS-Del-345 ···RTVGQEKKGYFEDRRPSANC
BCNT ···AIHNRGKEGYIERKA
WT -331 -334 -336 -338 -339 -345
α-BCNT-C
α-His
D
GS-WT ···QEKKGYFEDRRP···
GS(G335A) ···QEKKAYFEDRRP···
GS(Y336A) ···QEKKGAFEDRRP···
GS(F337A) ···QEKKGYAEDRRP···
GS(E338A) ···QEKKGYFADRRP···
GS(All-Ala) ···QEKKAAAADRRP···
GS(A4A) ···QEAAGYFEAARP···
WT G335A Y336A F337A E338A All-Ala A4A
α-BCNT-C
α-His

Fig 5

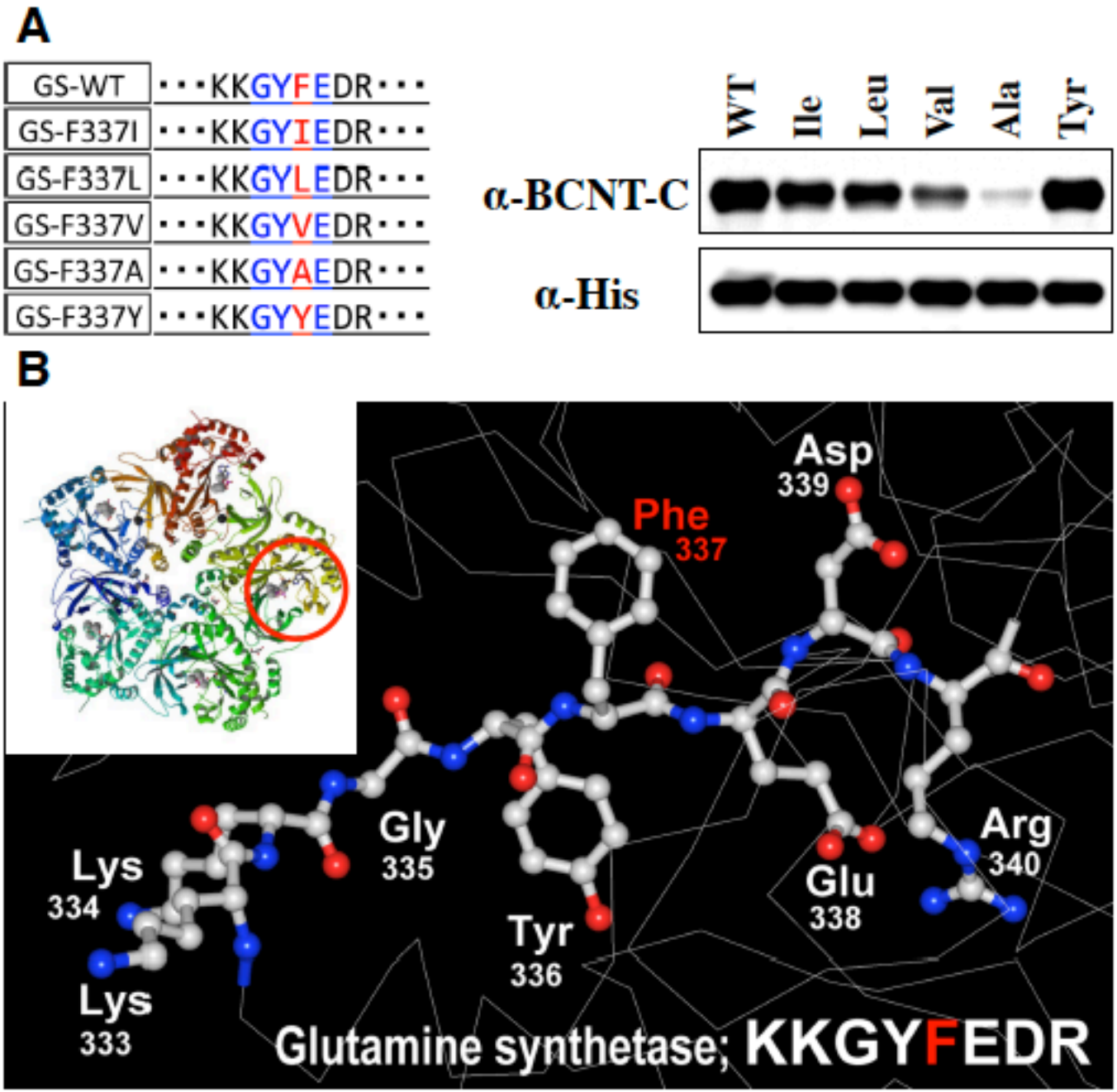


GS(332-346): CH_3CO-EKKGYFEDRRPSANC-COOH

GYIE: NH_2-EELAIHNRGKEGYIERKA-NH_2

All-Ala: NH_2-EELAIHNRGKEAAAARKA-NH_2

Fig 6


A
GS-WT ···KKGYFEDR···
GS-F337I ···KKGYIEDR···
GS-F337L ···KKGYLEDR···
GS-F337V ···KKGYVEDR···
GS-F337A ···KKGYAEDR···
GS-F337Y ···KKGYYEDR···
WT
Ile
Leu
Val
Ala
Tyr
α-BCNT-C
α-His
B
Asp
339
Phe
337
Gly
335
Lys
334
Lys
333
Tyr
336
Glu
338
Arg
340
Glutamine synthetase; KKGYFEDR