Crystal Structure of the Human Centromere Protein B (CENP-B) Dimerization Domain at 1.65-Å Resolution*

Received for publication, September 21, 2003, and in revised form, September 29, 2003
Published, JBC Papers in Press, September 30, 2003, DOI 10.1074/jbc.M310388200

Maki S. Tawaramoto‡, Sam-Yong Park‡, Yoshinori Tanaka‡, Osamu Nureki††, Hitoshi Kurumizaka***‡‡, and Shigeyuki Yokoyama****§§

From the §RIKEN Genomic Sciences Center, 1-7-22 Suehiro-cho, Tsukuru, Yokohama 230-0045, Japan, the ‡Protein Design Laboratory, Yokohama City University, 1-7-29 Suehiro-cho, Tsukuru, Yokohama 230-0045, Japan, the †Department of Biophysics and Biochemistry, Graduate School of Science, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan, **RIKEN Harima Institute at SPring8, 1-1-1 Koito, Miki-suku-cho, Sayo, Hyogo 679-5148, Japan, and the ††Waseda University School of Science and Engineering, 3-4-1 Okubo, Shinjuku-ku, Tokyo 169-8555, Japan

The human centromere protein B (CENP-B), a centromeric heterochromatin component, forms a homodimer that specifically binds to a distinct DNA sequence (the CENP-B box), which appears within every other α-satellite repeat. Previously, we determined the structure of the human CENP-B DNA-binding domain, CENP-B-(1-129), complexed with the CENP-B box DNA. In the present study, we determined the crystal structure of its dimerization domain (CENP-B-(540–599)), another functional domain of CENP-B, at 1.65-Å resolution. CENP-B-(540–599) dimer contains two α-helices, which are folded into an antiparallel configuration. The CENP-B-(540–599) dimer formed a symmetrical, antiparallel, four-helix bundle structure with a large hydrophobic patch in which 23 residues of one monomer form van der Waals contacts with the other monomer. In the CENP-B-(540–599) dimer, the N-terminal ends of CENP-B-(540–599) are oriented on opposite sides of the dimer. This CENP-B dimer configuration may be suitable for capturing two distant CENP-B boxes during centromeric heterochromatin formation.

The centromere, an essential chromosomal region required for the proper segregation of chromosomes during cell division, is embedded within heterochromatin in most eukaryotes (1–3). This centromere-specific heterochromatin contains specific DNA-binding proteins that target the centromeric region of chromosomes. In humans, the centromere-specific DNA-binding proteins CENP-A,1 CENP-B, and CENP-C were identified as antigens for autoimmune sera from scleroderma patients (4, 5).

CENP-A is the centromere-specific histone H3 variant (6), and CENP-C is a fundamental component of the inner kinetochore plate (7–11). CENP-A and CENP-C do not seem to have any obvious sequence specificity in DNA binding. In contrast, CENP-B specifically binds a 17-base pair sequence (the CENP-B box), which appears in every other 171-base pair α-satellite repeat in human centromeres (12), and induces nucleosome positioning in the vicinity of the CENP-B box (13). These results suggest that CENP-B may function as a transacting factor, which induces the formation of the centromere-specific heterochromatin. Analyses with cultured human cells revealed that the existence of the CENP-B box within the α-satellite sequence is required for the formation of a functional centromere containing CENP-B in minichromosomes in vivo (14, 15). Biochemical and genetic studies also showed that CENP-A, CENP-B, and CENP-C cooperatively constitute functional centromeres in mammalian cells (16, 17).

However, CENP-B null mice appeared to be normal (18–20). This discrepancy about the CENP-B requirement for centromere formation may be explained by presuming the existence of functional homologue(s) of CENP-B. In fact, CENP-B-like proteins of unknown function have been identified in humans, such as the jerkly-like protein (21, 22) and the transposases encoded by the human Tigger1 and Tigger2 transposable elements (23, 24). The functional redundancy of CENP-B homologues has also been found in the fission yeast Schizosaccharomyces pombe. Three fission yeast proteins, Abp1, Chb1, and Chb2, are homologues of CENP-B (25–27) and associate with centromeric heterochromatin (28, 29). None of these proteins is essential for S. pombe viability like human CENP-B (30). However, the double disruptants of Abp1 and Chb1 showed a synergistic reduction of centromere function in fission yeast (29, 30).

Human CENP-B is a dimeric protein composed of 80-kDa subunits (31). It contains DNA-binding domains and a dimerization domain at the N terminus and the C terminus, respectively (32). Electron microscopic observations showed that CENP-B bundles the two distant CENP-B boxes by its dimer formation and DNA binding activities (13). The bundling activity suggests that, in the centromere, CENP-B may function in the higher ordered chromatin formation, in addition to the centromere-specific nucleosome assembly. To reveal the molecular mechanism of the CENP-B functions in centromeric chromatin formation, we determined previously the crystal structure of the CENP-B N-terminal DNA-binding domain complexed with the CENP-B box DNA (33). In the present study, we determined the crystal structure of another CENP-B...
functional domain, the dimerization domain, at 1.65-Å resolution. The structure provides new insights into CENP-B dimerization that may play an essential role in the centromere-specific heterochromatin formation.

EXPERIMENTAL PROCEDURES

Purification of the Dimerization Domain of CENP-B—A DNA fragment containing the CENP-B-(540–599) portion of the human CENP-B gene was amplified by polymerase chain reaction and cloned into the Ndel site of the pET15b vector. Selenomethionine-labeled CENP-B-(540–599) was overexpressed in Escherichia coli BL21 CodonPlus (DE3)-RP-X cells under the control of the T7 promoter as a His6-tagged protein. The E. coli cultures carrying the CENP-B-(540–599) expression vector were grown at 30 °C, and isopropyl-β-D-thiogalactopyranoside (IPTG; 80 μM) was added at an OD600 of 0.4 to induce protein expression. After overnight cultivation at 18 °C, the cells were harvested and disrupted by sonication in 50 mM Tris-HCl buffer (pH 8.0) containing 50 mM NaCl, 5% glycerol, and 1 mM phenylmethylsulfonyl fluoride (PMSF). The His6 tag was uncoupled from the His6-tagged protein in the 250 mM CHES buffer (pH 9.5) containing 10% glycerol for a few minutes and transferred again to 50 mM CHES buffer (pH 9.5) containing 20% glycerol. The crystals were soaked in a cryo-protectant solution containing the CENP-B-(540–599) portion of the human CENP-B-(540–599) was eluted in a 12-column volume linear gradient from 50 to 800 mM NaCl in this buffer. The CENP-B-(540–599) protein was concentrated and further purified by HiLoad 26/60 Superdex 75 gel filtration column chromatography (Amersham Biosciences).

Crystal Structure of the human CENP-B Dimerization Domain

Crystal Growth and Data Collection—The purified CENP-B-(540–599) was concentrated to 12 mg of protein per milliliter, and crystals were obtained by the hanging drop method with nickel-nitrilotriacetic acid (Ni-NTA)-agarose beads (Qiagen) at 4 °C for 1 h. The CENP-B-(540–599)-bound Ni-NTA-agarose beads were then packed into an Econo-Column (Bio-Rad) and washed with 30-column volumes of 50 mM Tris-HCl buffer (pH 8.5) containing 500 mM NaCl. The cell debris was removed by centrifugation for 20 min at 30,000 × g, and the lysate was mixed gently by the batch method with nickel-nitrilotriacetic acid (Ni-NTA)-agarose beads (Qiagen) at 4 °C for 1 h. The CENP-B-(540–599) crystals were obtained by the hanging drop method after mixing an equal volume of 12 mg/ml CENP-B-(540–599) crystals were obtained by the hanging drop method after mixing an equal volume of 12 mg/ml CENP-B-(540–599) and cycled in 50 mM Tris-HCl buffer (pH 8.0) containing 50 mM NaCl, 5% glycerol, and 1 mM phenylmethylsulfonyl fluoride. The His6 tag was uncoupled from CENP-B-(540–599) by a digestion with 10 units of thrombin protease (Amersham Biosciences) per milligram of CENP-B-(540–599). Then, the fractions containing CENP-B-(540–599) were mixed with 5 ml of heparin-Sepharose (Amersham Biosciences) at 4 °C for 2 h and then packed into an Econo column (Bio-Rad). The heparin-Sepharose beads with CENP-B-(540–599) were washed with 6-column volumes of 50 mM Tris-HCl buffer (pH 8.5) containing 50 mM NaCl, and CENP-B-(540–599) was eluted in a 12-column volume linear gradient from 50 to 800 mM NaCl in this buffer. The CENP-B-(540–599) crystal structure was solved by the multiple wavelength anomalous dispersion (MAD) method using crystals of the selenomethionine-substituted sample (Table 1). A diffraction data set was collected at 1.65-Å resolution at beamline BL44B2 at SPring-8, Harima, Japan.

Protein Structure—In the crystal structure, CENP-B-(540–599) is composed of two α-helices (α1 and α2) (Fig. 2A), which are folded into an antiparallel configuration (Fig. 2B). The human CENP-B-(540–599) region is perfectly conserved among the mammalian CENP-Bs, except for the hamster CENP-B Trp-595 residue, indicating that the dimerization domain structures of mammalian CENP-Bs are identical (Fig. 2A). The α1 and α2 helices are amphipathic, and the two CENP-B-(540–599) monomers, which interact via their hydrophobic surfaces, form a dimer with a symmetrical, antiparallel, four-helix bundle structure (Fig. 2, B–D). The last 16 residues of each CENP-B-(540–599) C terminus, which is the natural C terminus of the full-length CENP-B, are disordered. Therefore, they are not required for dimer formation. This may reflect the flexible nature of the CENP-B C-terminal region, which is rich in glycine residues. Interestingly, the amino acid residues in the CENP-B region containing the amphipathic helices are significantly conserved in the human Tigger1 transposase (Fig. 2A), which also shares sequence similarity with CENP-B in the N-terminal DNA-binding and central transposase-like domains. Therefore, the Tigger1 transposase may be a CENP-B functional homologue, which can complement the CENP-B function in CENP-B-deficient mice. Furthermore, the amino acid residues in the structured region of CENP-B-(540–599) are also conserved in the S. pombe CENP-B homologues Abp1, Cbb1, and Cbb2 (Fig. 2A), suggesting that the dimerization mechanisms of the human and fission yeast CENP-Bs are similar.

The crystals used in the present study contain two dimers per asymmetric unit in which the CENP-B-(540–599) dimer dimers interacted well with each other (Fig. 2E). The root mean square deviation value between these CENP-B-(540–599) dimers was 0.7 Å for the Ca atoms, indicating that these
CENP-B-(540–599) dimers share almost identical structures.

**The Dimer Interface**—As shown in Fig. 3A, the two CENP-B-(540–599) monomers dimerize to form an antiparallel, four-helix bundle. This type of dimerization is also present in the bacterial ROP protein (root mean square deviation value of 2.6 Å for the Cα atoms) (Fig. 3B). The two α-helices of CENP-B-
FIG. 2. Crystal structure of CENP-B-(540–599). A, the secondary structure of CENP-B-(540–599). The secondary structures of the CENP-B-(540–599) monomers in the dimer are presented in the top and second rows. The N-terminal and C-terminal α-helices are denoted as α1 and α2, respectively, and the N-terminal loop and the loop between α1 and α2 are denoted as L1 and L2, respectively. The two molecules are colored blue and pink, respectively, and dashed lines indicate the C-terminal disordered regions. The dimerization domain sequence of human CENP-B is aligned to those of the African green monkey (AGM) CENP-B, hamster CENP-B, mouse CENP-B, S. pombe Abp1, S. pombe Cbh1, and S. pombe Cbh2 and is independently aligned to the corresponding region of the human Tigger1 transposase (bottom row). These amino acid sequences were aligned by the ClustalW program (43). Fully conserved residues, moderately conserved residues, and weakly conserved residues are colored red, orange, and yellow, respectively. Numbers indicate the amino acid position from the N terminus. B–D, three views of the CENP-B-(540–599) structure. Colors correspond to those described above for panel A. E, Overall structures of the two CENP-B-(540–599) dimers in the asymmetric unit.
The topologies in the dimeric forms of CENP-B-(540–599) and ROP are the same; the N-terminal helices interact with those of the other monomer. Similar four-helix bundle structures were also found in the human telomere-binding proteins TRF1 and TRF2 (Fig. 3, C and D). However, these four helices are only parts of the large dimerization domains of TRF1 and TRF2 (Fig. 3E). Therefore, the dimerization domains of these telomere-binding proteins differ from that of CENP-B. The structural similarities between the CENP-B-(540–599) dimer and the dimerization interfaces of TRF1 (65–111 residues) and TRF2 (43–89 residues) were not significant (root mean square deviation values of 3.5 and 3.3 Å, respectively, for the Ca atoms), and the two helices of the TRF1 or TRF2 monomer within the interface are non-parallel (the angles between these helices are ~15°) (Fig. 3, C and D). Thus, these results indicate that the CENP-B-(540–599) dimer structure is unique.

The dimerization interface of CENP-B-(540–599) is a large, buried surface area of 1,386 Å² (~34% of the monomer surface) (Fig. 4A), which is large enough to confirm the natural dimerization of CENP-B. On the other hand, the area buried in the dimer-dimer surface in the asymmetric unit (Fig. 2D) is as small as 711 Å², indicating that this interaction is non-physiological. In the functional dimerization interface, 23 residues of one monomer form van der Waals contacts with the other monomer (Fig. 4B). Specific hydrogen bonds were also found in the large hydrophobic interface between the monomers, indicating that the CENP-B-(540–599) dimer formation is structure-specific. The side chain OH group of Tyr-557 forms a hydrogen bond with the Thr-582 O atom of the other monomer (Fig. 4C, top). The NH group of His-570 forms a hydrogen bond with the Asp-577 O atom of the other monomer (Fig. 4C, bottom). These interactions are nearly symmetrical between the monomers because of their non-crystallographic symmetry.

**Centromeric Chromatin**—CENP-B reportedly bundles two distant CENP-B boxes through its dimer formation and DNA-binding abilities (13). This DNA-bundling activity should be important for the centromeric chromatin formation. To reveal the functional significance of the CENP-B DNA-bundling activity, structural analyses of the DNA-binding and dimerization domains of CENP-B would be useful. We determined previously the crystal structure of the CENP-B N-terminal DNA-binding domain (CENP-B-(1–129)) in a complex form with the
CENP-B box DNA (33). In the present study, we determined the crystal structure of another CENP-B functional domain, the dimerization domain, which revealed that the CENP-B dimerization domains form a homodimer with a symmetrical, antiparallel, four-helix bundle structure. Many hydrophobic residues face the monomer-monomer interface and form van der Waals contacts between the monomers. In fact, no CENP-B-(540–599) monomer was detected in solution by the gel filtration analysis. Therefore, the monomeric form of the CENP-B dimerization domain cannot exist in solution, because the hydrophobic surfaces, which correspond to about 34% of the monomer surface, are exposed to the solvent.

In the crystal structure, the N-terminal loops of the CENP-B-(540–599) dimer were located on opposite sides of the dimer. Black and blue lines indicate van der Waals contacts formed between the CENP-B-(540–599) monomers, and red lines indicate hydrogen bonds formed between the CENP-B-(540–599) monomers. C, stereo views of the \( (2F_o - F_c) \) electron density maps for the hydrogen bonds formed between the CENP-B-(540–599) monomers. Top, \( (2F_o - F_c) \) electron density maps for the hydrogen bond between Tyr-557 and Thr-582; bottom, \( (2F_o - F_c) \) electron density maps for the hydrogen bond between His-570 and Asp-577 (lower panel). The red dashed line indicates the hydrogen bond. Amino acid residues denoted by blue and black letters are those from different monomers.

**Fig. 4. Interactions between the CENP-B-(540–599) monomers.** A, molecular surface of the CENP-B-(540–599) monomer. Blue and red represent positive and negative electrostatic potential, respectively. The buried surface between the CENP-B-(540–599) monomers (left) and the solvent-accessible surface (right) are represented. B, schematic representation summarizing the interactions between the CENP-B-(540–599) monomers in the dimer. The α-helices and loops are presented as cylinders and tubes, respectively. Black and blue lines indicate van der Waals contacts formed between the CENP-B-(540–599) monomers, and red lines indicate hydrogen bonds formed between the CENP-B-(540–599) monomers. C, stereo views of the \( (2F_o - F_c) \) electron density maps for the hydrogen bonds formed between the CENP-B-(540–599) monomers. Top, \( (2F_o - F_c) \) electron density maps for the hydrogen bond between Tyr-557 and Thr-582; bottom, \( (2F_o - F_c) \) electron density maps for the hydrogen bond between His-570 and Asp-577 (lower panel). The red dashed line indicates the hydrogen bond. Amino acid residues denoted by blue and black letters are those from different monomers.
These N-terminal positions of the CENP-B-(540–599) dimer may be suitable for capturing two distant CENP-B boxes with its N-terminal DNA-binding domains, as shown in the model presented in Fig. 5. Because the CENP-B box sequence exists in every other α-satellite repeat (171 base pairs), CENP-B may accommodate a pair of centromeric nucleosomes that contain histones H2A, H2B, H4, and a centromere-specific histone H3 variant, CENP-A, between two CENP-B boxes tethered by the CENP-B dimer (Fig. 5).

Acknowledgments—We thank Drs. N. Kamiya and S. Adachi for data collection at SPring8 (BL44B2).

REFERENCES

1. Sullivan, B., and Karpen, G. (2001) J. Cell Biol. 154, 683–690
2. Choo, K. H. A. (2001) Dev. Cell 1, 165–177
3. Cleveland, D. W., Mao, Y., and Sullivan, K. F. (2003) Cell 112, 407–421
4. Earnshaw, W. C., and Rothfield, N. (1985) Chromosoma 91, 313–321
5. Valdivia, M. M., and Brinkley, B. R. (1985) J. Cell Biol. 101, 1124–1134
6. Sullivan, K. F., Hechenberger, M., and Masri, K. (1994) J. Cell Biol. 127, 581–592
7. Saitoh, H., Tomkiel, J., Cooke, C. A., Rattrie, H., III, Maurer, M., Rothfield, N. F., and Earnshaw, W. C. (1992) Cell 10, 115–125
8. Tomkiel, J., Cooke, C. A., Saitoh, H., Bernat, R. L., and Earnshaw, W. C. (1994) J. Cell Biol. 125, 531–545
9. Fukagawa, T., and Brown, W. R. (1997) Hum. Mol. Genet. 6, 2301–2308
10. Kalitsis, P., Fowler, K. J., Earle, E., Hill, J., and Choo, K. H. A. (1998) Proc. Natl. Acad. Sci. U. S. A. 95, 1136–1141
11. Fukagawa, T., Penden, C., Morris, J., and Brown, W. (1999) EMBO J. 18, 1146–14209
12. Masumoto, H., Muro, Y., Nosaki, N., and Okazaki, T. (1989) J. Cell Biol. 109, 1963–1973
13. Yoda, K., Ando, S., Muro, Y., Nosaki, N., and Okazaki, T. (1998) Genes Cells 3, 533–548
14. Ikeda, M., Grimes, B., Nakano, M., Saitoh, K., Hoshino, H., McGill, N. L., Cooke, H., and Masumoto, H. (1998) Nat. Biotechnol. 16, 431–439
15. Ohzeki, J.-I., Nakano, M., Okada, T., and Masumoto, H. (2002) J. Cell Biol. 159, 765–775
16. Howman, E. V., Fowler, K. J., Nelson, A. J., Redward, S., MacDonald, A. C., Kalitsis, P., and Choo, K. H. A. (2000) Proc. Natl. Acad. Sci. U. S. A. 97, 1148–1153
17. Ando, S., Yang, H., Nosaki, N., Okazaki, T., and Yoda, K. (2002) Mol. Cell. Biol. 22, 2229–2241
18. Hudson, D. F., Fowler, K. J., Earle, E., Saffery, R., Kalitsis, P., Trowell, H., Hill, J., Wreford, N. G., de Kretser, D. M., Cancilla, M. R., Howman, E., Hui, L., Cuts, S. M., Irvine, D. V., and Choo, K. H. A. (1998) J. Cell Biol. 141, 309–319
19. Perez-Castro, A. V., Shamanski, F. L., Meneses, J. J., Lovato, T. L., Vogel, K. G., Moyzis, R. K., and Pedersen, R. (1998) Dev. Biol. 201, 155–164
20. Fowler, K. J., Hudson, D. F., Salamonsen, L. A., Edmondson, S. R., Earle, E., Sibson, M. C., and Choo, K. H. A. (2000) Genes Cells 5, 30–41
21. Toth, M., Grimes, B., Buzsaki, G., and Donovan, J. P. (1995) Genet. 11, 71–75
22. Liu, W., Seto, J., Sibille, E., and Toth, M. (2003) Mol. Cell. Biol. 23, 4083–4093
23. Smit, A. F., and Biggs, A. D. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 1403–1414
24. Kipling, D., and Warburton, P. E. (1997) Trends Genet. 13, 141–145
25. Murakami, Y., Huberman, J. A., and Hurwitz, J. (1996) Proc. Natl. Acad. Sci. U. S. A. 9, 502–507

FIG. 5. A model for DNA bundling by the CENP-B dimer in centromeric chromatin. Orange ribbons with arrowheads indicate 171-base pair α-satellite repeats. Green circles show the centromeric nucleosomes, which may contain histones (H2A, H2B, and H4) and CENP-A. The crystal structures of the dimerization and DNA-binding domains of CENP-B are connected by the Pro-rich region (yellow dashed lines), the central transposase-like domains (pink dashed circles), and the Asp/Glu-rich region (blue dashed lines).
Crystal Structure of the human CENP-B Dimerization Domain

51461

26. Lee, J. K., Huberman, J. A., and Hurwitz, J. (1997) *Proc. Natl. Acad. Sci. U. S. A.* 94, 8427–8432
27. Irelan, J. T., Gutkin, G. I., and Clarke, L. (2001) *Genetics* 157, 1191–1203
28. Halverson, D., Baum, M., Stryker, J., Carbon, J., and Clarke, L. (1997) *J. Cell Biol.* 136, 487–500
29. Nakagawa, H., Lee, J. K., Hurwitz, J., Allshire, R. C., Nakayama, J., Grewal, S. I., Tanaka, K., and Murakami, Y. (2002) *Genes Dev.* 16, 1766–1778
30. Baum, M., and Clarke, L. (2000) *Mol. Cell. Biol.* 20, 2852–2864
31. Earnshaw, W. C., Machlin, P. S., Bordwell, B. J., Rothfield, N. F., and Cleveland, D. W. (1987) *Proc. Natl. Acad. Sci. U. S. A.* 84, 4979–4983
32. Kitagawa, K., Masumoto, H., Ikeda, M., and Okazaki, T. (1995) *Mol. Cell. Biol.* 15, 1602–1612
33. Tanaka, Y., Nuroki, O., Kurumizaka, H., Fukai, S., Kawaguchi, S., Ikuta, M., Iwahara, J., Okazaki, T., and Yokoyama, S. (2001) *EMBO J.* 20, 6612–6618
34. Adachi, S., Oguchi, T., Tanida, H., Park, S.-Y., Shimizu, H., Miyatake, H., Kamiya, N., Shiro, Y., Inoue, Y., Ueki, T., and Izuka, T. (2001) *Nucl. Intr. Meth. Phys. Res. Sect. A* 467/468, 711–714
35. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* 276, 307–326
36. Collaborative Computational Project Number 4 (1994) *Acta Crystallogr. Sect. D Biol. Crystallogr.* 50, 769–763
37. Terwilliger, T. C., and Berendzen, J. (1999) *Acta Crystallogr. Sect. D Biol. Crystallogr.* 55, 849–861
38. Roussel, A., and Cambillau, C. (1989) *Silicon Graphics Geometry Partner Directory*, pp. 77–78, Silicon Graphics, Mountain View, CA
39. Perrakis, A., Merris, R., and Lamzin, V. S. (1999) *Nat. Struct. Biol.* 6, 458–463
40. Brunger, A. T. (1996) *X-FLOR Version 3.851 Manual*, Yale University Press, New Haven, CT
41. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* 26, 283–291
42. Fairall, L., Chapman, L., Moss, H., de Lange, T., and Rhodes, D. (2001) *Mol. Cell* 8, 351–361
43. Higgins, D., Thompson, J., and Gibson, T. J. (1994) *Nucleic Acids Res.* 22, 4673–4680