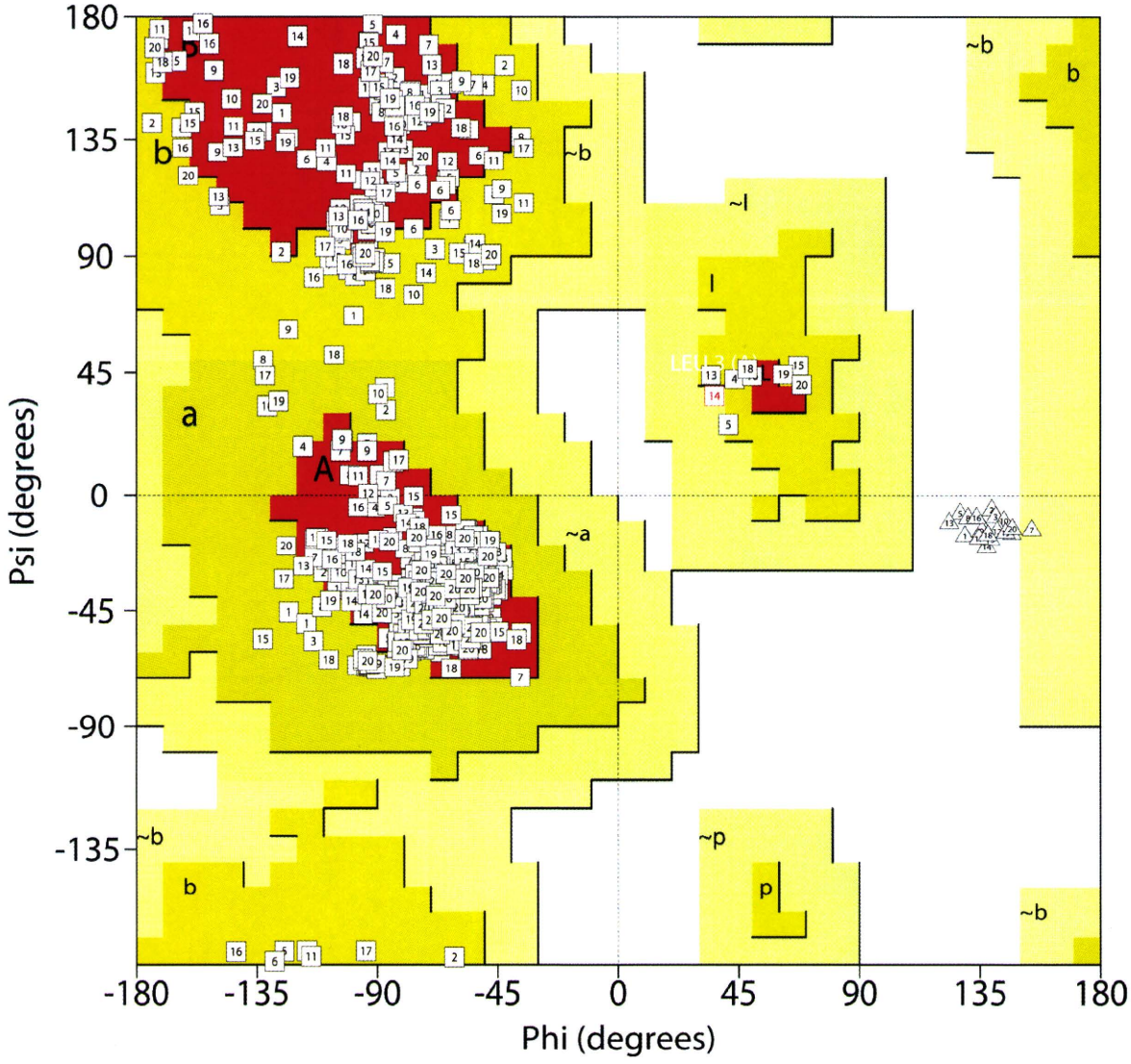
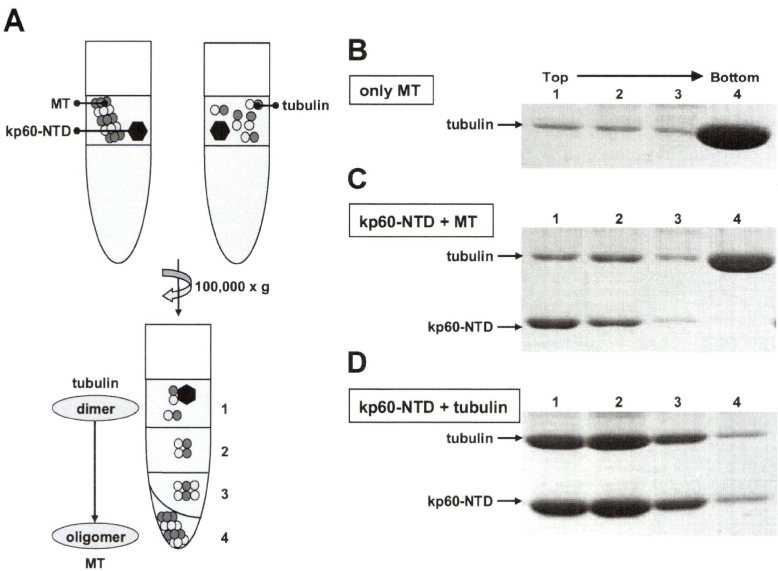
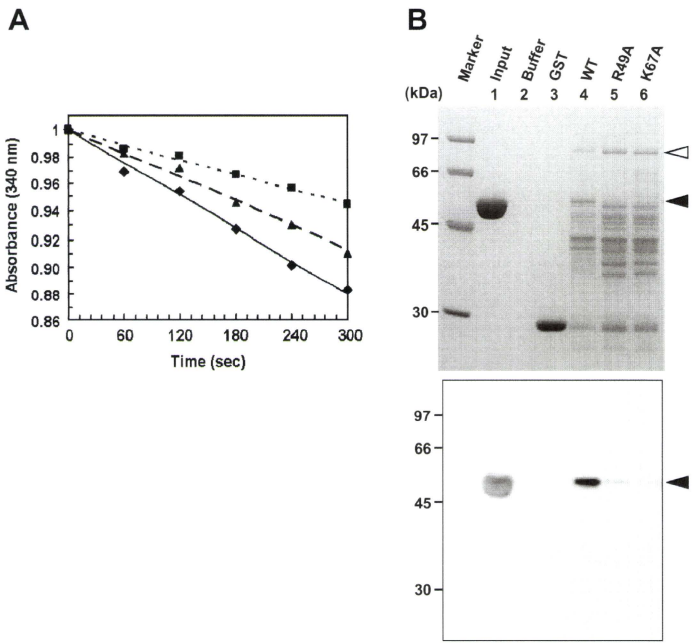


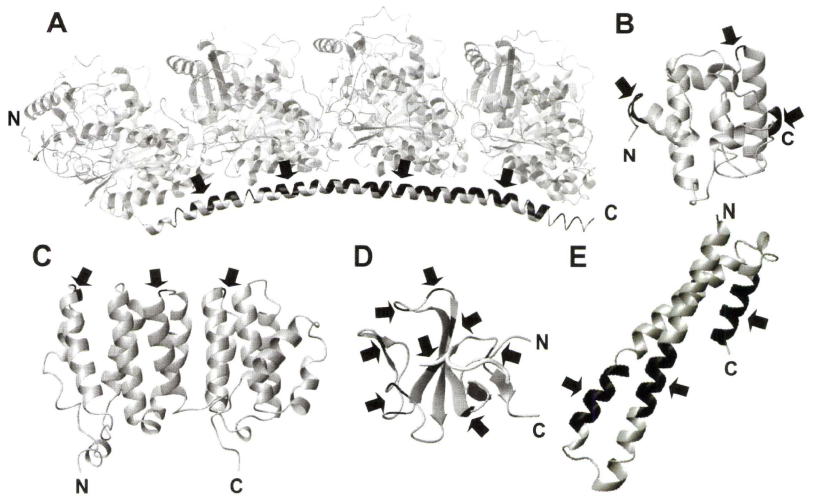

Supplementary Figure 2



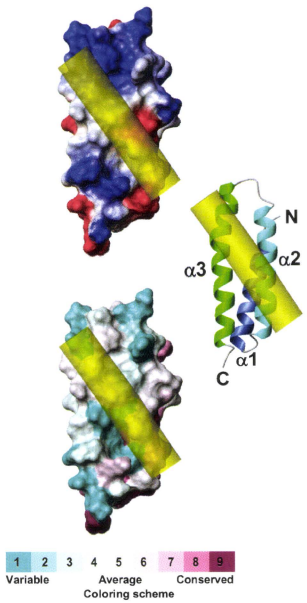

Supplementary Figure 3



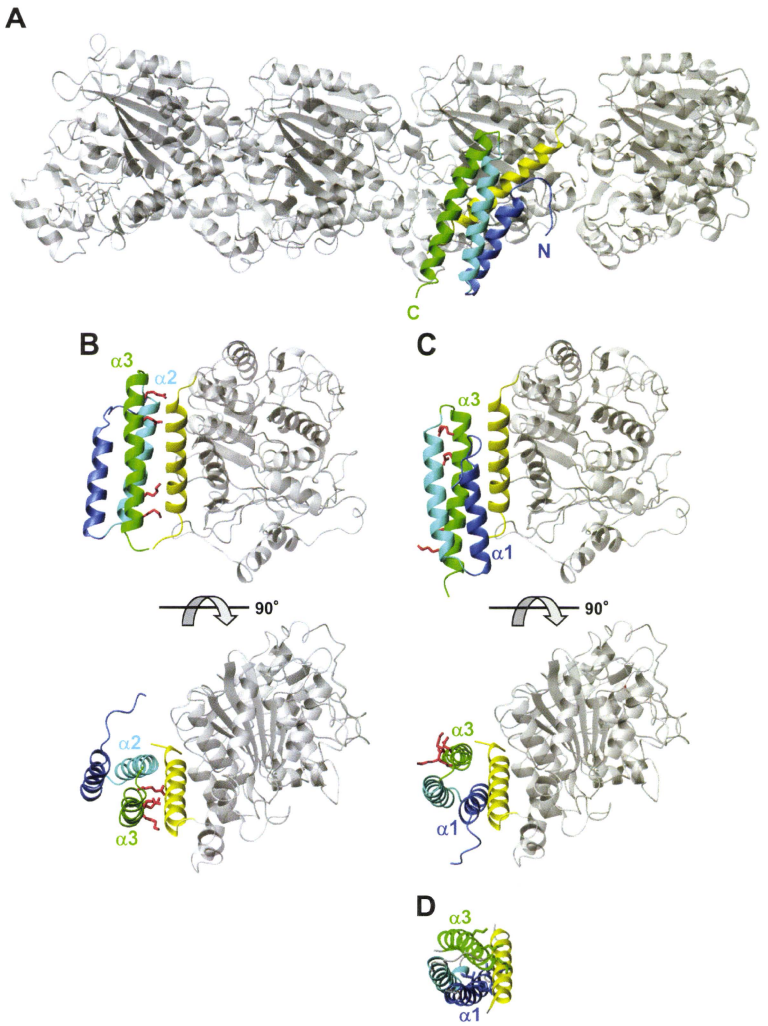

Supplementary Figure 4



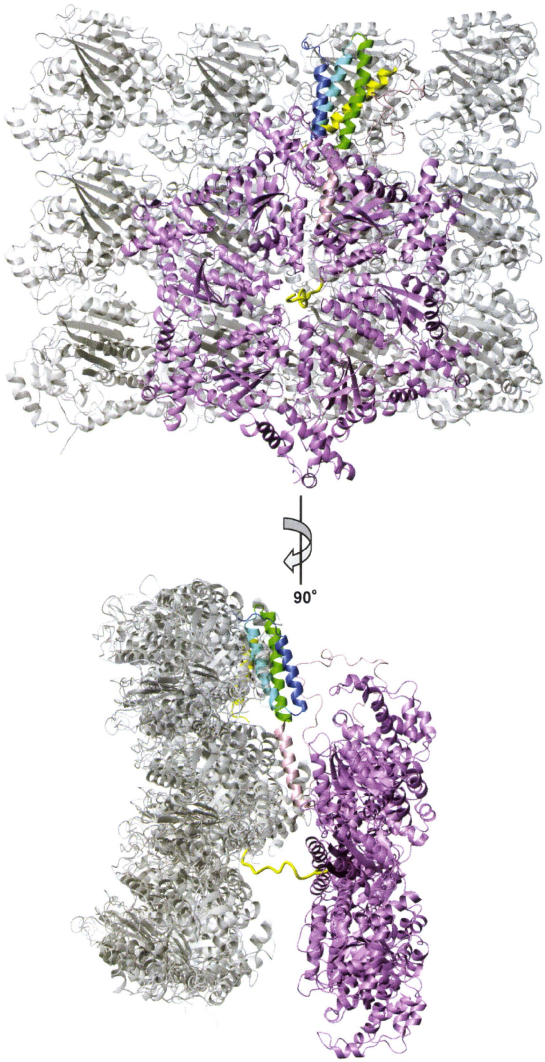

Supplementary Figure 5




Supplementary Figure 6




Supplementary Figure 7




Supplementary Figure 8

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 285, NO. 21, pp. 15931–15940, May 21, 2010


# Structure and Mutagenesis Studies of the C-terminal Region of Licensing Factor Cdt1 Enable the Identification of Key Residues for Binding to Replicative Helicase Mcm Proteins

Received for publication, October 18, 2009, and in revised form, March 8, 2010 Published, JBC Papers in Press, March 24, 2010, DOI 10.1074/jbc.M109.075333

**JunGoo Jee[‡1,2], Takeshi Mizuno[§1,3], Katsuhiko Kamada[§4], Hidehito Tochio[¶], Yasumasa Chiba[‖], Ken-ichiro Yanagi[§], Gentaro Yasuda[§], Hidekazu Hiroaki[**], Fumio Hanaoka[§‡‡5], and Masahiro Shirakawa[¶]**

*From the [‡]Center for Priority Areas, Tokyo Metropolitan University, 1-1 Minami-Osawa, Hachioji, Tokyo 192-0397, [§]Cellular Physiology Laboratory, RIKEN Discovery Research Institute, Wako 351-0198, the [¶]Department of Molecular Engineering, Graduate School of Engineering, Kyoto University, Nishikyo, Kyoto 615-8510, [‖]Ajinomoto Company, Kawasaki, Kawasaki 210-8681, the [**]Division of Structural Biology, Graduate School of Medicine, Kobe University, 7-5-1 Kusunoki, Chuo, Kobe, Hyogo 650-0017, and the [‡‡]Faculty of Science, Gakushuin University, 1-5-1 Mejiro, Toshima, Tokyo 171-8588, Japan*

**In eukaryotes, DNA replication is fired once in a single cell cycle before cell division starts to maintain stability of the genome. This event is tightly controlled by a series of proteins. Cdt1 is one of the licensing factors and is involved in recruiting replicative DNA helicase Mcm2–7 proteins into the pre-replicative complex together with Cdc6. In Cdt1, the C-terminal region serves as a binding site for Mcm2–7 proteins, although the details of these interactions remain largely unknown. Here, we report the structure of the region and the key residues for binding to Mcm proteins. We determined the solution structure of the C-terminal fragment, residues 450–557, of mouse Cdt1 by NMR. The structure consists of a winged-helix domain and shows unexpected similarity to those of the C-terminal domain of Cdc6 and the central fragment of Cdt1, thereby implying functional and evolutionary relationships. Structure-based mutagenesis and an *in vitro* binding assay enabled us to pinpoint the region that interacts with Mcm proteins. Moreover, by performing *in vitro* binding and budding yeast viability experiments, we showed that ~45 residues located in the N-terminal direction of the structural region are equally crucial for recognizing Mcm proteins. Our data suggest the possibility that winged-helix domain plays a role as a common module to interact with replicative helicase in the DNA replication-licensing process.**

In eukaryotes, DNA replication is highly coordinated to retain the integrity of the genome. Whereas DNA replication in prokaryotes begins at a single site and stops at the end of the genome, eukaryotic genomes consist of multiple replication origins where DNA replication starts. These origins are synchronized so that they are activated only once in a single division cycle. A series of proteins, the origin recognition complex (ORC),[6] cell division cycle 6 homolog (Cdc6), chromatin licensing and DNA replication factor 1 (Cdt1), and minichromosome maintenance 2–7 (Mcm2–7) are known to play correlated roles in licensing (1–5). Oncogenic proliferation often causes abnormal expression of the proteins involved in the DNA-licensing process, thus emphasizing the importance of harmonious adjustments between these proteins (6). A complicated interaction network between these proteins has been reported (7), although the details at residue and atom levels remain largely unknown.

Formation of a pre-replication complex at each origin is the first event in the replication process. ORC proteins bind initially to each replication origin of DNA. The DNA sequences of the origins where ORC binds have not been identified, except for those in *Saccharomyces cerevisiae* (8), suggesting that there may be other factors involved in addition to the sequences (9). The DNA strand at the origin needs to be unpaired to begin replication; therefore, the existence of replicative helicase is essential. Mcm2–7 proteins are believed to function as the replicative helicase in eukaryotes. Each Mcm2–7 protein consists of a conserved AAA+ ATPase type C-terminal helicase domain and a rather diverse N-terminal domain. The stoichiometry of Mcm2–7 while functioning remains unclear. Both hetero-hexameric and several additional complexes with different combinations of Mcm proteins have been found *in vivo* and *in vitro* (10). Meanwhile, reconstituted Mcm4/6/7 complex has helicase activity *in vitro* (11, 12). The proteins that bridge between ORC and Mcm2–7 are Cdc6 and Cdt1. Cdc6, like Mcm proteins, belongs to the AAA+ ATPase family and shows similarity to clamp loader proteins that load ring-shaped sliding clamps onto DNA (13). The result, Mcm2–7 forming a donut-like ring shape, led to the postulation that Cdc6 behaves as a clamp loader for Mcm2–7 (14, 15).

Cdt1 was originally found in fission yeast (16), and its function as a factor involved in replication licensing was first characterized in *Xenopus* and fission yeast (17, 18). Several subsequent experiments have shown that Cdt1 interacts with

*The atomic coordinates and structure factors (code 2RQQ) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).*

[1] Both authors contributed equally to this work.

[2] To whom correspondence may be addressed. Tel.: 81-42-677-4873; Fax: 81-42-677-4873; E-mail: jee-jungoo@tmu.ac.jp.

[3] Present address: Cellular Dynamics Laboratory, RIKEN.

[4] Present address: Chromosome Dynamics Laboratory, RIKEN.

[5] To whom correspondence may be addressed. Tel.: 81-3-3986-0221; Fax: 81-3-5992-1029; E-mail: fumio.hanaoka@gakushuin.ac.jp.

[6] The abbreviations used are: ORC, origin recognition complex; Cdc6, cell division cycle 6 homolog; Cdt1, chromatin licensing and DNA replication factor 1; Mcm2, minichromosome maintenance 2; GST, glutathione *S*-transferase; WHD, winged-helix domain; r.m.s.d., root mean square deviation.

Mcm2–7 both *in vitro* and *in vivo* (19–22). Cdt1 can be divided into three functional regions, namely, N-terminal, central, and C-terminal regions, which have DNA-, geminin-, and Mcm2–7-binding activity, respectively (20, 22). In addition to its role in binding to DNA, the N-terminal region is employed to interact with proliferating cell nuclear antigen (23) and Cdc7 and contains ubiquitination and acetylation sites. The function of Cdt1 is negatively controlled by a small protein termed geminin (24) and by the ubiquitin-mediated degradation pathway (25). In metazoans, the degree of sequence conservation in the C-terminal fragment of Cdt1 is higher than in the N-terminal and central regions (22). However, even in the C-terminal region, homology decreases abruptly in budding yeast, showing sequence identity of <20% compared with the mouse sequence. In fact, *S. cerevisiae* Cdt1 could not be uncovered using only a simple database search (21). Of the three functional regions in Cdt1, structural information is available for the central region, which was determined by x-ray crystallography as a complex with geminin by Lee *et al.* (26). They also showed the long helical region of geminin is involved in hindering the interaction between Cdt1 and Mcm2–7.

Because of a resemblance in DNA replication initiation between archaea and eukaryotes, the data from archaea, particularly structural information, have assisted our understanding of the events in eukaryotes (27). The structures of ORC (28, 29), Cdc6 (13), and Mcm (14) from archaea are very informative; however, it is impossible to deduce the function of Cdt1 using this strategy because of a lack of orthologues in archaea. So far, the regions of Cdt1 for Mcm2–7 binding have been characterized mainly by using deletion mutants (20, 22). At a residue-specific level, to date, only one study has reported the importance of two conserved lysines (30).

To understand the function of Cdt1 in detail, we carried out a structural and mutagenesis study. In this study, we used multidimensional NMR spectroscopy to determine the structure of the C-terminal fragment (residues 450–557) of mouse Cdt1. We then performed an *in vitro* binding assay with Mcm proteins with mutants prepared based on the NMR structure. Finally, we carried out *in vitro* and *in vivo* assays to demonstrate that the additional residues in the N-terminal direction from the structural region are equally important for the recognition of Mcm proteins.

## EXPERIMENTAL PROCEDURES

*DNA Cloning and Protein Purification for Structure Determination*—DNA corresponding to the C-terminal fragment of mouse Cdt1 (residues 450–557, mCdt1$^{C\text{-}WHD}$) was cloned into a pGEX-6p-1 vector and then transformed into *Escherichia coli* BL21 strain. The fusion proteins were purified by GSH affinity chromatography. The N-terminal GST was removed by cleavage with recombinant rhinovirus 3C protease. Further purification steps using Q-Sepharose and Sephadex 75 column chromatography were applied. Finally, the buffer was replaced with 50 mM phosphate buffer, pH 6.5, for NMR measurement. Site-directed mutagenesis was performed using the QuikChange mutagenesis kit (Stratagene) according to the manufacturer's instructions. All the mutants and the Mcm4/6/7 complex for the *in vitro* binding assay were cloned and purified using methods identical to those described previously (22). We performed densitometric analysis for GST pulldown data by using Multi Gauge (Fujifilm) and quantified binding affinities between Cdt1 mutants and Mcm4/6/7 proteins.

*Yeast Strains*—Strains were constructed using standard genetic techniques. W303-1A (*MATa ade2-1 ura3-1 his3–11, 15 trp1-1 leu2–3 112can1–100*) and YIp128-based strains (*ade2-1 ura3-1 his3–11, 15 trp1-1 112can1–100, TAH11::GAL-TAH11 HIS3, leu2::LEU2 TAH11-FLAG*) were grown at 30 °C in YP (1% (w/v) yeast extract (Difco); 2% (w/v) Bacto-peptone (Difco)) containing the indicated carbon source at a final concentration of 2% (w/v). cDNA for Tah11 was amplified by PCR and subcloned into pYIp128.

*NMR Spectroscopy and Resonance Assignment*—NMR spectra were acquired with Bruker DRX 500-, DRX 800-, or Varian 900-MHz spectrometers equipped with pulse-field gradient, triple resonance probes. Assignment of the $^{1}H$, $^{15}N$, and $^{13}C$ resonances of mCdt1$^{C\text{-}WHD}$ was made by a series of triple resonance experiments, CBCA(CO)HN, HNCACB, HNCO, C(CCO)NH, H(CCCO)NH, and HCCH-TOCSY (31). The $^{15}N$-edited NOESY-HSQC (100- and 150-ms mixing times) and $^{13}C$-edited NOESY-HSQC (150-ms mixing time) spectra were used to derive distance restraints for structure determination and assignment. All the data were processed with NMRPipe (32) and analyzed with NMRView (33).

*Structure Calculation*—Structures were calculated by CYANA 2.0 (34) and further refined by the AMBER 7 package (35). All the NOE cross-peaks were checked manually using NMRView (33) and assigned using the CANDID algorithm (36) of CYANA. Several tens of CANDID runs were carried out according to the criteria reported (37). A total of 1,826 meaningful NOE upper distance restraints was obtained by CANDID (415 intra-residual, 512 sequential, 523 medium range, and 375 long range). Backbone torsion angle restraints of 159 were derived by the TALOS program (38) and used for all the calculations. Finally, 100 structures that did not show significant violations against experiment restraints were generated using 15,000 steps torsion angle dynamics by CYANA and further refined by AMBER with an all-atom force field. To approximate solvent effects, the generalized Born implicit solvent model was used (39). The best 20 structures were selected and analyzed using AQUA and PROCHECK-NMR software (40). The structure closest to the mean coordinate was selected as a representative. The coordinates and NMR restraints for structure calculation were deposited in the PDB data base (ID: 2RQQ). All the figures were created by using PyMOL software (41).

## RESULTS

*Preparation of Stable Protein Fragment for NMR Study*—Initially the region containing residues 407–557 of mCdt1 (mCdt1$^{CTF}$) was chosen for the NMR study. However, the residues 407–449 (mCdt1$^{C\text{-}N}$) of the region were sensitive to protease and easily degraded at the protein concentration needed to measure three-dimensional NMR experiments in the required time. Therefore, we focused our attention on the region containing residues 450–557 (mCdt1$^{C\text{-}WHD}$) for structure determination, even though the region containing residues 407–557 is needed for binding to Mcm proteins. The construct

**TABLE 1**
**Statistics of mCdt1$^{C\text{-}WHD}$ final 20 structures**
All variations are S.D. unless shown otherwise.

| | |
|---|---|
| NOE restraints | 1826 |
| Intra (\|i-j\| = 0) | 416 |
| Sequential (\|i-j\| = 1) | 512 |
| Medium range (2≤\|i-j\|≤4) | 523 |
| Long range (\|i-j\|≥5) | 375 |
| Dihedral angle restraints | 79/80 ($\phi/\varphi$) |
| **AMBER energies (kcal/mol)**[a] | |
| Total energy | −4712 ± 19 |
| Constraint energy | 24 ± 2 |
| **Numbers of violations** | |
| Distance violation (>0.5 Å) | 0 |
| Angle violation (>5.0°) | 0 |
| **Maximum violations** | |
| Distance violation (Å) | 0.406 |
| Angle violation (°) | 4.842 |
| **Coordinate precision**[b] | |
| Backbone atoms (Å) | 0.35 |
| Heavy atoms (Å) | 0.87 |
| **Ramachandran plot (%)** | |
| Most favored | 92.7 |
| Additionally favored | 6.5 |
| Generously favored | 0.4 |
| Disallowed | 0.4 |

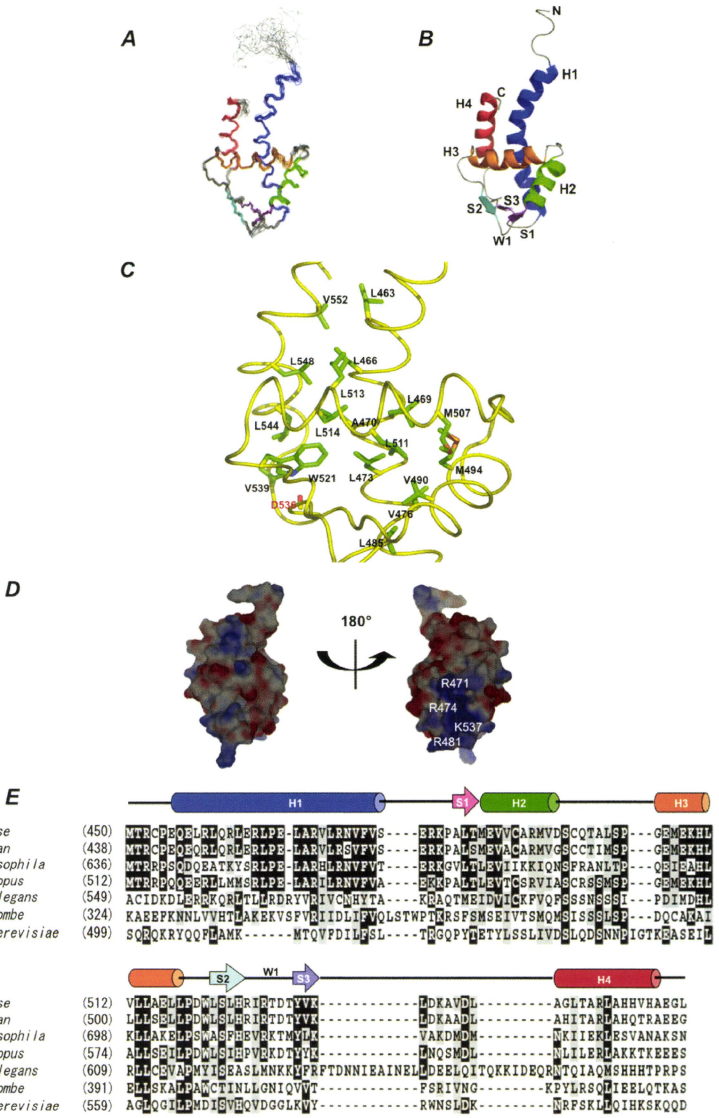
[a] Represents the generalized Born model energy of AMBER 7.
[b] Residues 454–555 are considered.

of residues 407–557 was used instead for the *in vitro* binding assay. The sample of mCdt1$^{C\text{-}WHD}$ was stable for several weeks at high concentration. The results of size-exclusion chromatography indicated that mCdt1$^{C\text{-}WHD}$ exists as a monomer (data not shown).

*Winged-helix Domain in C-terminal Fragment of Cdt1*—Final structures are well refined, with a backbone and heavy atom root mean square deviation (r.m.s.d.) in residues Pro-454 to Glu-555 of 0.35 Å and 0.87 Å, respectively. The structures also have good geometry with 92.6% of residues in the most favored region of the Ramachandran plot (Table 1 and Fig. 1*A*). N-terminal residues to Cys-453 and C-terminal residues of Gly-556 and Leu-557 are not well converged in the ensemble due to absence of long range NOE. The secondary structures comprise four α-helices (H1, residues 455–480; H2, 487–496; H3, 504–517; and H4, 541–552) and three β-strands (S1, 485–486; S2, 522–526; and S3, 531–535). Pro-467 is positioned in the middle of H1, causing a slight kink of H1 around the residue. The topology of the secondary structures, including the wing region between S2 and S3 (W1, 527–530) is H1-S1-H2-H3-S2-W1-S3-H4 and constitutes a typical winged-helix domain (hereafter WHD) (Fig. 1*B*). Residues located at the interior layer of each helix contribute to stabilization of the helices by forming hydrophobic interactions (Fig. 1*C*). Between helices H1 and H4, pairs of residues, Leu-463 and Val-552, and Leu-466 and Leu-548, are in contact with each other, respectively. For helices H1, H3, and H4, three residues, Leu-513, Leu-514, and Leu-517, of H3 interact with Leu-466 of H1 and Leu-548 of H4, respectively. Between H1 and H2, residues Leu-469 and Met-494 also form hydrophobic interactions. The Trp-521 on the loop connecting H3 and S2 stabilizes the secondary structures by contact with Ala-470 and Leu-473 of H1, and Asp-536, Val-539, and Leu-544 of H4. In addition to the hydrophobic interactions described, a hydrogen bond between Nε1 of Trp-521 and backbone carbonyl of Asp-536 participates in tightening of the secondary structures. The presence of this bond is supported by its observation in all the final 20 NMR structures. Intriguingly, the mCdt1$^{C\text{-}WHD}$ sequence contains 19 leucines, with many of these residues being involved in hydrophobic interactions between secondary structures. Surface electrostatic charges reveal that, on the outward face from the end of helix H1 and the next loop connecting helices H1 and H2 and the loop between S3 and H3, there is a positively charged patch consisting of Arg-471, Arg-474, Arg-481, and Lys-537 of mCdt1$^{C\text{-}WHD}$. With the exception of this region, there is no noticeable charged patch (Fig. 1*D*).

*Sequence Alignment and Conservation of Cdt1$^{C\text{-}WHD}$*—To know sequence conservation from the viewpoint of function and evolution, the regions corresponding to mCdt1$^{C\text{-}WHD}$ (residues 450–557) were extracted from six representative species and aligned using MUSCLE software (42). The alignment was adjusted manually on the basis of structural information. The extracted sequences are as follows. Mouse, Met$^{450}$–Leu$^{557}$; Human, Met$^{438}$–Leu$^{546}$; *Xenopus*, Met$^{512}$–Ser$^{619}$; *Drosophila*, Met$^{636}$–Asn$^{743}$; *Caenorhabditis elegans*, Ala$^{546}$–Ser$^{665}$; *S. pombe*, Lys$^{324}$–Ser$^{437}$; and *S. cerevisiae*, Ser$^{499}$–Asp$^{604}$ (Fig. 1*E*). The sequence identities against mCdt1$^{C\text{-}WHD}$ are 83% for human Cdt1$^{C\text{-}WHD}$, 59% for *Xenopus*, 44% for *Drosophila*, 24% for *C. elegans*, 20% for *S. pombe*, and 19% *for S. cerevisiae*. Two regions between S3 and H4 in *C. elegans* are not well aligned, and two yeast sequences show lower similarity than those from metazoan. In particular, the H1 region of *S. cerevisiae* is not well aligned with the others. Nonetheless, those residues that have important roles in forming the tertiary structure are well converged in ensemble conformers of mCdt1$^{C\text{-}WHD}$ and are also conserved in sequence alignment. This suggests all the Cdt1$^{C\text{-}WHD}$s have similar structures.

*Structures Sharing Similarities with Cdt1$^{C\text{-}WHD}$*—The DALI server (43) was used to search for the protein structures similar to that of mCdt1$^{C\text{-}WHD}$. This search retrieved 562 WHD structures with Z-scores higher than 2.0 from the PDB data base in March of 2009 (44). Strikingly, the closest similarity was found in the structure of the C-terminal domain of the archael Cdc6 orthologue (PDB: 1FNN, Z score: 6.8, r.m.s.d.: 3.0 Å (over 81 residues)). WHDs in proteins, zinc uptake regulation protein Furb (2O03, Z score: 6.6, r.m.s.d.: 2.7 (76 residues)), *Staphylococcus* Sars protein (1P4X, Z score: 6.6, r.m.s.d.: 4.2 (76 residues)), and viral protein F93 (2CO5, Z-score: 6.5, r.m.s.d.: 2.8 (76 residues)), were followed. Despite not being included in the structures of closest similarity, it should be noted that the central region of mouse Cdt1 (residues 179–365, mCdt1$^{M\text{-}WHD}$) from the complex structure with geminin (PDB: 1WLQ) were also retrieved from the DALI server and had a Z-score of 4.9 and r.m.s.d. of 3.4 Å in 93 residues. To obtain detailed information from similar structures, we extracted 21 sequences in the structures with Z-scores higher than 6.0. Of these, considering sequence redundancies, we selected 16 and aligned the sequences based on the structural overlays. For comparison, we included the sequence from mCdt1$^{M\text{-}WHD}$. In each sequence any amino acid that had a corresponding mCdt1$^{C\text{-}WHD}$ amino acid was extracted (Fig. 2*A*). The alignment revealed that residues in the loop regions of mCdt1$^{C\text{-}WHD}$ were less aligned, suggesting that the loop conformations in mCdt1$^{C\text{-}WHD}$ and other



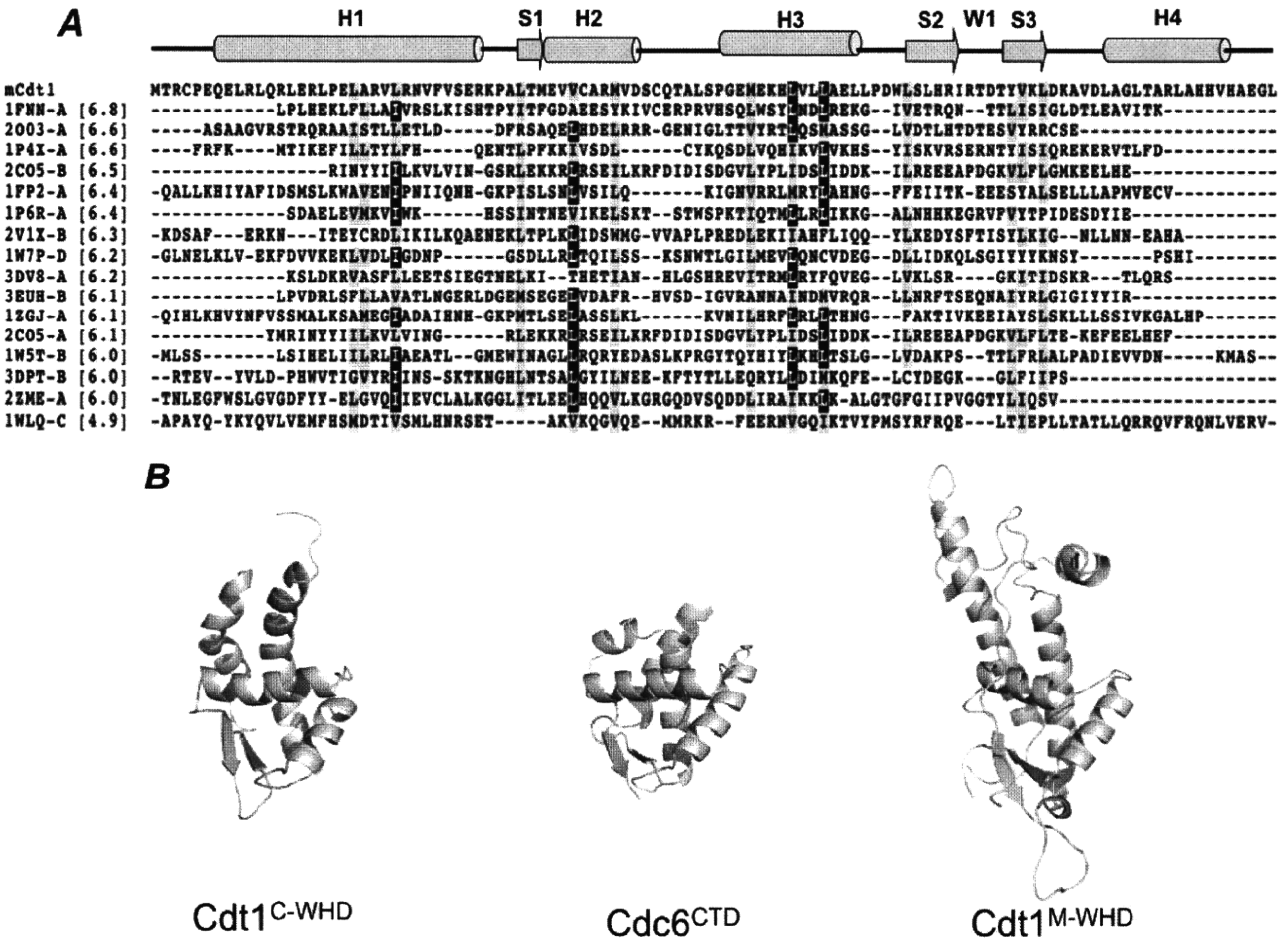



FIGURE 2. **Comparison of sequences and structures between Cdt1$^{C\text{-}WHD}$ and other winged-helix domains.** *A*, the sequences from DALI (43)-driven homologous structures ($Z \geq 6.0$) are aligned when there is a corresponding sequence in mCdt1$^{C\text{-}WHD}$. For comparison, the sequence from the central fragment from mouse Cdt1 (1WLQ) is included. PDB codes and the chain names of WHD are shown. The values in *parentheses* represent the Z-scores. When the portion of an amino acid is >50%, the amino acid is colored *black*. The residues with similar properties are shown in *gray*. *B*, winged-helix domains found in DNA replication licensing. mCdt1$^{C\text{-}WHD}$ in the current study (*left*), C-terminal WHD of Cdc6 (Cdc6$^{CTD}$, PDB: 1FNN) (*center*) and mCdt1$^{M\text{-}WHD}$ (PDB: 1WLQ) (*right*) are drawn using a *ribbon diagram*. Coordinates are aligned to have the same direction with mCdt1$^{C\text{-}WHD}$.

WHDs are diverse. Of the secondary structures, the helix H4 is hardly aligned even in the C-terminal domain structure of Cdc6 (1FNN) (Fig. 2). It is clear that the properties of hydrophobic residues involved in maintaining the tertiary fold in mCdt1$^{C\text{-}WHD}$ are still conserved in the other WHDs despite overall lower sequence similarities, implying that their roles are to maintain tertiary folds. In the next section, we discuss WHDs in detail. mCdt1$^{M\text{-}WHD}$ (1WLQ) has additional residues in N and C termini and longer inserted loop between S2 and S3 (Fig. 2). However, secondary structures in winged-helix regions are similar to each other and other WHDs in length and positions (Fig. 2). Structure-based sequence alignment reveals 16% sequence identity between the two domains mCdt1$^{M\text{-}WHD}$ and mCdt1$^{C\text{-}WHD}$, whereas sequence-only alignment could not align correctly even the secondary structural regions. Under "Discussion," we provide a detailed review of these findings from an evolutionary viewpoint.

*In Vitro Binding Assays with C-terminal Regions of Cdt1 and Mcm4/6/7*—In a previous experiment, we showed that the C-terminal fragment of Cdt1 binds to Mcm proteins (22). In the present study, we characterized the region in greater detail (Fig. 3, *A–C*). Because reconstituting Mcm2–7 *in vitro* is not possi-

FIGURE 1. **NMR structures of mCdt1$^{C\text{-}WHD}$.** *A*, superimposed final 20 structures. The structures are overlaid by backbone atoms of residues Pro$^{454}$ to Glu$^{555}$. Four helices are drawn with different colors, H1 (residues 455–480) in *blue*, H2 (487–496) in *green*, H3 (504–517) in *brown*, and H4 (541–552) in *red*. In three $\beta$-strands, S1 (485–486) is drawn in *pink*, S2 (522–526) in *purple*, and S3 (531–535) in *cyan*. The wing region between S2 and S3 is labeled as *W1*. *B*, a *ribbon diagram* of the representative NMR structure. The secondary structures are colored using the same code as in *A*. *C*, residues involved in stabilizing inter secondary structures are drawn as *stick models*. Asp-536 whose backbone carbonyl forms hydrogen bonds with N$\epsilon$1 of Trp$^{521}$ is labeled in *red*, and the others are in *black*. *D*, electrostatic surface calculated from APBS (56) is drawn in the range of $-4k_B/T$ (*red*) to $4k_B/T$ (*blue*). All the figures of the structure were created by using PyMOL (41). All the coordinates were aligned to have the same direction. *E*, sequence alignment of Cdt1$^{C\text{-}WHD}$. Residues that correspond to 450–557 of mouse Cdt1 were extracted in human, *Xenopus*, *Drosophila*, *C. elegans*, *S. pombe*, and *S. cerevisiae* and aligned using MUSCLE (42). The number of starting residues in each sequence line is written in *parentheses*. Secondary structures are *drawn* on the sequences. Residues showing identities of >70% are colored *black*, and those >50% are *gray*.

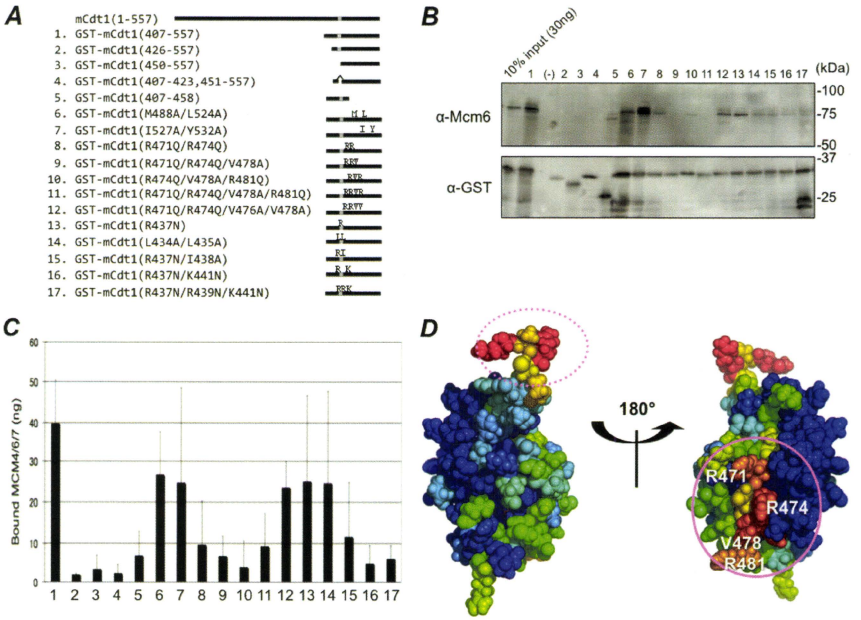



FIGURE 3. **In vitro binding assay with the fragments and the structure-based point mutants of mCdt1$^{CTF}$ and Mcm4/6/7.** *A*, peptide constructs used in this experiment. *B*, pulldown analysis. GST-tagged mCdt1$^{CTF}$ peptides were mixed with purified Mcm4/6/7 complex and incubated with glutathione-Sepharose to precipitate complexes containing mCdt1 peptides. The complexes were washed and visualized by Western blotting using anti-Mcm6 (*α-Mcm6*) and anti-GST (*α-GST*) antibodies. The reaction was done under in 200 mM NaCl, 20 mM Tris HCl, pH 7.5, 0.1% Nonidet P-40, 0.25% gelatin. The constructs having point mutations were generated with GST-mCdt1 (407–557). *C*, the amounts of bound Mcm4/6/7 and errors were quantified by repeating the reaction under identical conditions. *D*, surface-exposed and conserved residues that form a patch are shown. The *colors* in the order of *blue* to *red* indicate the lower to higher degrees of conservation. The first putative protein-protein interaction site is marked by a *dashed circle*, and the second site is marked by a *continuous circle*.

ble, we selected the Mcm4/6/7 complex based on reports that reconstituted Mcm4/6/7 has DNA helicase activity (11, 12, 30). Initially, the region of mCdt1 that acted as a Mcm4/6/7 binding site was narrowed to a fragment of residues 407–557. This region showed comparable binding affinity toward Mcm4/6/7 to that of wild type. However, the fragment, whose structure was determined in this study (mCdt1$^{C\text{-}WHD}$, residues 450–557), hardly bound to Mcm4/6/7 (Fig. 3, *B* and *C*, *lane* and *bar 3*). We selected only the protease-sensitive extra region of 407–458, but it also did not show binding affinity (Fig. 3, *B* and *C*, *lane* and *bar 5*). We further prepared two additional constructs, the former or the latter half of mCdt1$^{C\text{-}N}$ and the region of mCdt1$^{C\text{-}WHD}$ (407–423/451–557 and 426–450/451–557), and investigated their activities. Both constructs bound little with Mcm4/6/7 (Fig. 3, *B* and *C*, *lanes* and *bars 2* and *5*). Altogether, the existence of both residues 407–449 and 450–557 are indispensable for binding to Mcm4/6/7. The PSIPRED software (45) predicted the secondary structure of residues 424–448 as a helix, whereas the first part (407–423) that includes many prolines is likely to form neither helix nor strand. Judging from the sensitivity to protease, the region of mCdt1$^{C\text{-}N}$ probably does not make stable interaction with the rest of the region in the construct 450–557. Ferenbach *et al.* reported that the region required for *Xenopus* Cdt1 to bind to Mcm2/4/6/7 consists of residues 447–620 (20). However, residues 407–557 that we identified in mouse Cdt1 correspond to residues 469–619 in *Xenopus*, suggesting that the residues 447–468 of *Xenopus* Cdt1 are not necessary for binding to Mcm2/4/6/7.

*Regions of Winged-helix Domain for Binding to Mcm Proteins*—Both BLAST and PSI-BLAST searches when queried with residues 450–557 of mouse Cdt1 could not retrieve meaningful homologous sequences other than those derived from Cdt1 of various species. Currently available bioinformatics tools do not list noticeable consensus motifs from the sequence (data not shown). As a result, knowing only the sequence is not informative for speculating the functional residues of Cdt1$^{C\text{-}WHD}$. Therefore, we combined structural information with sequence data. This procedure is based on the idea that the residues important for scaffold or function in protein are restrained during evolution. Therefore, searching for residues that are identi-

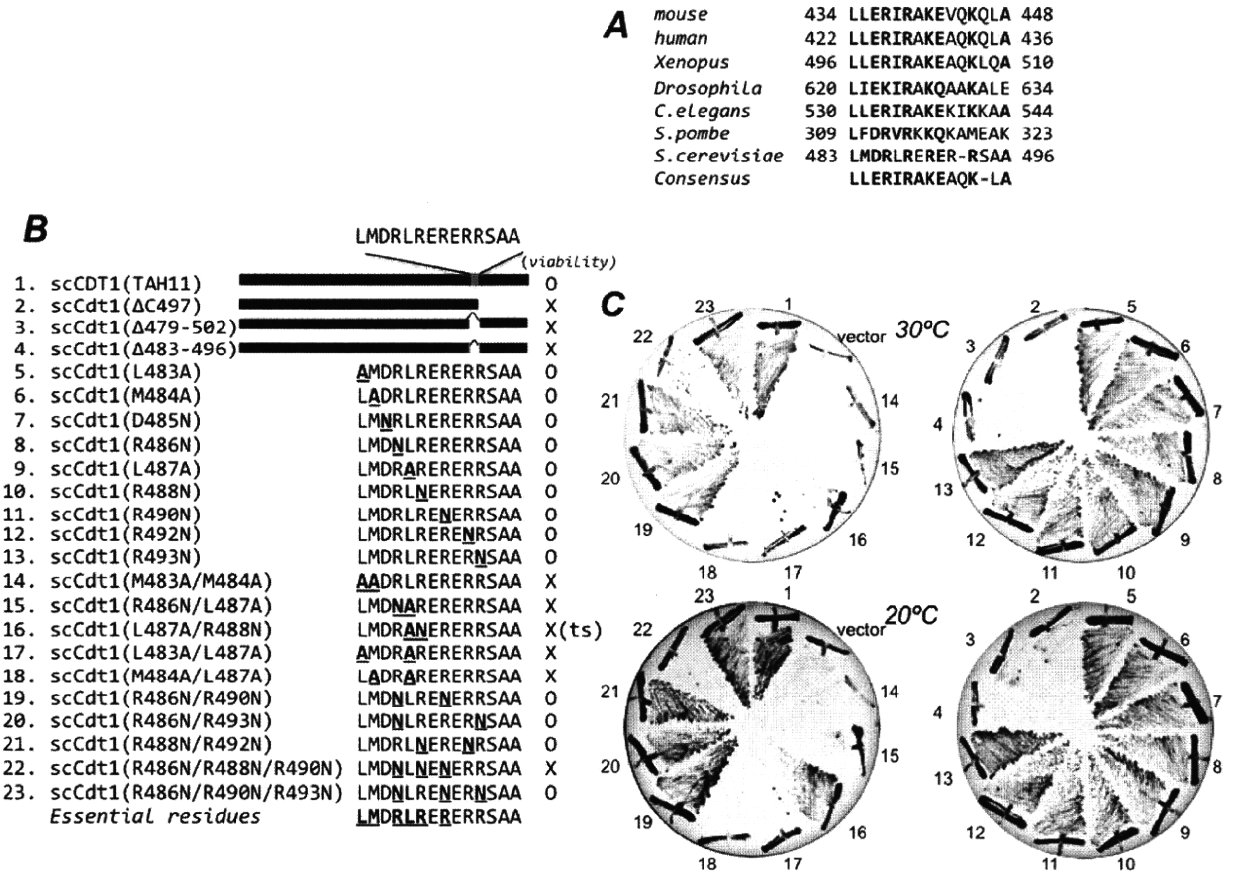



FIGURE 4. **Yeast viability assay.** *A*, the alignment of sequences corresponds to residues 434–448 in mCdt1. The consensus sequence is shown in the *bottom line*. *B*, summary of yeast viability assay. *Lines 1–4* contain vectors for deletion mutations. *5–23* are prepared for point mutations at the residues that are *underlined* with *bold characters*. *C*, yeast expression in glucose media lacking Leu. The same residue numbers are used in *B* and *C*. The yeast harboring vector 16 was viable at 30 °C, but nonviable at 20 °C, indicating temperature sensitivity (*ts*).

cal or have conserved properties on the exposed surface is a reasonable approach for deducing the functional portions from a clueless sequence. For this purpose, we used the WHISCY server (46). It selects only patches formed on conserved and solvent-accessible residues. In addition to the sequences used for alignment (Fig. 1*E*) and the current structure, we added nine more sequences from metazoans to increase precision. Two patches were scored as potential protein-protein interfaces (Fig. 3*D*). The first (*dashed circle*) was the N-terminal three residues of $Met^{450}$, $Thr^{451}$, and $Arg^{452}$, and the other (*continuous circle*) was the region containing the end of helix H1 and its consecutive loop, which include $Arg^{471}$, $Arg^{474}$, $Val^{478}$, and $Arg^{481}$. Considering protease sensitivity around the region of residues 450–452, the N-terminal three residues will be probably flexible even in the intact condition. The latter region has both hydrophilic and hydrophobic characteristics, overlapped partially with the positively charged surface of $Cdt1^{C\text{-}WHD}$, and is more conserved (Fig. 1). To identify the functional residues for Mcm2–7 proteins, we carried out an *in vitro* binding assay with Mcm4/6/7 proteins with a series of mutants, each of which has two to four amino acid substitutions in the region of $Cdt1^{C\text{-}WHD}$. Indeed, triple and quadruple mutants, R471Q/R474Q/V478A, R474Q/V478A/R481Q, and R471Q/R474Q/V478A/R481Q, showed markedly reduced binding activities to Mcm4/6/7 (Fig. 3, *B* and *C*, *lanes* and *bars 9–11*). Considering that R471Q/R474Q influenced a little (Fig. 3*B*, *lane 8*), $Val^{478}$ will probably play the central role in interacting with Mcm proteins, further supplemented by electrostatic contacts with basic residues. In contrast, other mutants containing mutations in residues $Met^{488}$, $Leu^{524}$, $Ile^{527}$, and $Tyr^{532}$, chosen simply from sequence conservation, resulted in marginal changes in binding affinity compared with that observed in the wild type (Fig. 3, *B* and *C*, *lanes* and *bars 6* and *7*).

*Yeast Viability Assay with Mutants at Additional Fragment*—To better characterize the flanking region, $Cdt1^{C\text{-}N}$ (residues 407–449 in mCdt1), especially the highly conserved second half part, $Leu^{434}$ to $Ala^{448}$, we performed a yeast viability assay. The alignment of eukaryotic sequences, including yeast sequences, reveals high sequence conservation of the region (Fig. 4*A*), which is distinct from $Cdt1^{C\text{-}WHD}$ where sequence similarity is low between mammals and yeast (Fig. 1*E*). Hydrophobic residues (Leu, Met, and Ile) and basic residues (Arg) are particularly conserved. We undertook a systematic mutagenesis of the region for Cdt1, Tah11 (hereafter *scCDT1*), in *S. cerevisiae* and then observed cell viability. Note that *scCDT1* is an essential gene. The promoter of endogenous *scCDT1* gene in W303a was replaced with the GAL1 promoter. Integration of the mutant *scCDT1* gene with the wild-type *scCDT1* promoter and cyc1

terminator at the LEU2 locus allowed the strain to produce mutant proteins in the glucose medium. At first, a C-terminal truncated mutant (Fig. 4*B*, *row 2*) or mutants that lack the conserved region (Fig. 4*B*, *rows 3* and *4*) were checked for growth defects. Both deletion constructs were unable to support growth in glucose medium, while full-length *scCDT1* (Fig. 4*B*-1) could be substituted for endogenous protein. These results indicate again that both regions of scCDT1, which correspond to $Cdt1^{C\text{-}WHD}$ and highly conserved $Cdt1^{C\text{-}N}$, are essential for scCDT1 function. To examine the functional effects of substitutions in the conserved region, 19 alleles of *scCDT1*, each of which contains amino acid substitution(s), were then generated by site-directed mutagenesis (Fig. 4*B*, *rows 5–23*). We substituted Ala for conserved Leu and Met residues, and charged residues, Arg and Asp, were replaced with neutral amino acid Asn. All the single amino acid substitutions proved to be viable. In contrast and intriguingly, five of eight double amino acid substitutions were nonviable. Also one of two triple amino acid changes for Arg was shown to be nonviable. Taken together, the existences of the hydrophobic residues at the positions of $Leu^{483}$/$Met^{484}$/$Leu^{487}$ in scCdt1 and $Leu^{434}$/$Leu^{435}$/$Ile^{438}$ in mCdt1, and the basic residues at $Arg^{486}$/$Arg^{488}$/$Arg^{490}$ in $scCdt1^{CTF}$ and $Arg^{437}$/$Arg^{439}$/$Lys^{441}$ in $mCdt1^{CTF}$ would be important for Cdt1 function. This was confirmed by *in vitro* GST pulldown experiments too. Whereas a single point mutant showed a nominal change (Fig. 3, *B* and *C*, *lanes* and *bars 13*), double and triple mutations in these residues caused the reduced affinities toward Mcm4/6/7 proteins (Fig. 3, *B* and *C*, *lanes* and *bars 15–17*), which is qualitatively in agreement with the results of yeast viability assay. You and Masai (30) reported that a double point mutant of mCdt1, K441A/K445A, failed to generate a complex with Mcm proteins. $Lys^{441}$ is one of those residues that we identified as an important residue in this study. Interestingly, the combined feature of hydrophobic and basic residues ($Leu^{434}$, $Leu^{435}$, $Arg^{437}$, $Ile^{438}$, $Arg^{439}$, and $Lys^{441}$) is reminiscent of the residues pinpointed in $mCdt1^{C\text{-}WHD}$ ($Arg^{471}$, $Arg^{474}$, $Phe^{477}$, $Val^{478}$, and $Arg^{481}$) for binding to Mcm proteins, likely suggesting similar contacting mode. On the other hand, the underlying role of the proline-rich region, residues 407–423, remains a subject for future studies.

## DISCUSSION

WHD is a typical DNA recognition motif (47). When WHD binds to double strand DNA, the interactions occur at H3, S2, S3, and W1 in the majority of cases. H3 is the primary contact site and W1 acts as an additional and supportive site. To the best of our knowledge, a reversal in these roles of H3 and W1 has been reported once in an exceptional case (47). In addition to H3 and W1, the existence of positively charged residues at H3, S2, S3, and W1 are indispensable for DNA recognition. $Cdt1^{CTF}$ binds to neither double nor single strand DNA (22). To determine the differences between $mCdt1^{C\text{-}WHD}$ and other WHDs, we superimposed the coordinate of $mCdt1^{C\text{-}WHD}$ into each WHD position of known DNA-protein complex structures and investigated the hypothetical contact sites between $mCdt1^{C\text{-}WHD}$ and DNA. In contrast to WHDs that bind to DNA, the lengths of H1 or H4 of $mCdt1^{C\text{-}WHD}$ are considerably longer, with these helices being more inclined to the direction of the plane that consisted of H3 and S2, S3, and W1. Because of these features, one or both of H1 and H4 of $mCdt1^{C\text{-}WHD}$ would encounter steric clashes with DNA, if it binds to DNA in the same manner as the other WHDs do. In fact, as shown in the structure-based sequence alignment (Fig. 2*A*), the residues in H1 and H4 are not fully aligned, meaning that they are unique to $mCdt1^{C\text{-}WHD}$. Furthermore, there is no particularly charged patch on the surface regions of H3, S2, S3, and W1 (Fig. 1*C*). It should be noted that the WHDs that share the closest similarities with $mCdt1^{C\text{-}WHD}$, those from PDB codes of 1FNN, 2O03, 1P4X, and 2CO5, are known to participate in protein-protein interactions, indicating that the structural features of WHD in $Cdt1^{C\text{-}WHD}$ are distinct from other DNA binding domains. Taken together, the inability of $Cdt1^{C\text{-}WHD}$ in DNA binding is consistent with its structural features, including long H1 and H4 helices and surface charge.

Accumulated data have revealed that some WHDs serve as a protein-protein interaction module. We need more examples to characterize in detail the region involved in protein-protein interaction, but at least in two such cases, the apoptotic protease-activating factor 1 (48) and ESCRT-II–III system (49), protein-protein interactions occur at similar areas of WHD to that identified as the Mcm2–7 binding site. The question then is, do other WHDs involved in DNA replication licensing use the same region to interact with Mcm2–7? The WHD of archael Cdc6 engages in interaction with Mcm proteins as well as DNA (50). As shown with typical DNA binding WHDs, H3 is a major DNA recognition site in Cdc6 WHD (51, 52), but the region that interacts with Mcm proteins has not been characterized. The structural resemblance between WHDs of archael Cdc6 and $mCdt1^{C\text{-}WHD}$ and the high degree of sequence homology between archael and eukaryotic Cdc6s (13) tempt us to presume that similar regions in WHDs are employed for Mcm proteins. In yeast, the Cdc6 WHD can interact with both DNA and Mcm2–7 simultaneously (53). When considering the steric effects and the large molecular sizes in DNA replication licensing protein, the prime candidate site of Cdc6 WHD for binding to Mcm2–7 will be placed at the opposite side from the H3 DNA binding site and likely include the region identified as the Mcm binding site in $Cdt1^{C\text{-}WHD}$. Yet there has been no report that $Cdt1^{Central}$ has a helicase-binding activity. Nevertheless, both $Cdt1^{Central}$ and $Cdt1^{CTF}$ are required to load Mcm2–7 for licensing activity (20). Moreover, Lee *et al.* reported that the interaction between Cdt1 and Mcm2–7 is inhibited by steric hindrance caused by a contact between geminin that binds to the central region of Cdt1 and Mcm proteins (26). This raises the possibility that two fragments of $Cdt1^{Central}$ and $Cdt1^{CTF}$ are located in close proximity, at least when Cdt1 interacts with Mcm2–7, implying potential cooperation between the two regions.

Similarities in folding, sequence, and interacting site with partner protein led us to postulate that the central and C-terminal regions of Cdt1 were duplicated in tandem from a domain and have evolved together. This type of duplication occurs frequently in eukaryotic proteins and confers functional diversities (54). In addition to the region of WHD in the crystal structure (residues 179–365 of mouse Cdt1), $Cdt1^{Central}$ needs its flanking sequence to bind to geminin *in vivo* (20). Interest-

ingly, this sequence is located in the N-terminal direction from the structural region of $Cdt1^{M\text{-}WHD}$ and corresponds to residues 193–243 in *Xenopus* and 121–181 in mouse Cdt1, respectively. These positions are reminiscent of the relative positions between the protease-sensitive fragment, residues 407–448, and the structural part, residues 450–557, of $Cdt1^{C\text{-}WHD}$ in the current study. However, the central regions of Cdt1 are under weaker evolutionary pressure, because $Cdt1^{Central}$ has lower sequence homology across species than those found in $Cdt1^{CTF}$ (22). Evolutionary trends in Cdt1 fragments are also reflected by the presence or diverseness of their binding partners. Geminin, the partner of $Cdt1^{Central}$, is missing or elusive in budding yeast, whereas the presence of Mcm2–7 with which $Cdt1^{CTF}$ mainly interacts is highly conserved in eukaryotes, consistent with evolutionary conservation in the fragments. In addition to the clues and questions required to understand the interaction between Cdt1 and Mcm2–7 raised by our data, it would be worthwhile studying the roles of WHDs in $Cdt1^{CTF}$, $Cdt1^{Central}$, and Cdc6 as a common evolutionary module.

In summary, we have shown that the C-terminal fragment of Cdt1 forms a WHD, but its structural features are different from other typical DNA binding WHDs. Although the structural region determined in the current study is insufficient to bind to Mcm proteins alone, we were able to characterize the residues of the WHD that are involved in interactions with replicative helicase. We also showed an additional region in the N-terminal direction of the WHD is equally important. How these two parts work cooperatively remains unanswered, and further experiments are required to elucidate this mechanism. Eukaryotic DNA replication initiation machinery is large and complicated, and presently, it is not possible to reconstitute the system *in vitro*. Therefore, dissecting each component and examining the parts is a reasonable strategy. We believe that our data will be a valuable addition to understanding the process and evolution of DNA replication licensing in eukaryotes.

*Acknowledgments—We thank Drs. Toshihiko Eki for constructing yeast strains and Naoko Imamoto for encouragement and support.*

*Addendum*—While we were in preparation of the manuscript, a report describing the solution and crystal structures of C-terminal regions of mouse Cdt1 was published (55).

## REFERENCES


1. Bell, S. P. (2002) *Genes Dev.* **16,** 659–672
2. Bell, S. P., and Dutta, A. (2002) *Annu. Rev. Biochem.* **71,** 333–374
3. Blow, J. J., and Dutta, A. (2005) *Nat. Rev. Mol. Cell Biol.* **6,** 476–486
4. Nishitani, H., and Lygerou, Z. (2004) *Front. Biosci.* **9,** 2115–2132
5. Takeda, D. Y., and Dutta, A. (2005) *Oncogene* **24,** 2827–2843
6. Blow, J. J., and Gillespie, P. J. (2008) *Nat. Rev. Cancer* **8,** 799–806
7. Ramachandran, N., Hainsworth, E., Bhullar, B., Eisenstein, S., Rosen, B., Lau, A. Y., Walter, J. C., and LaBaer, J. (2004) *Science* **305,** 86–90
8. Bell, S. P., and Stillman, B. (1992) *Nature* **357,** 128–134
9. Cvetic, C., and Walter, J. C. (2005) *Semin. Cell Dev. Biol.* **16,** 343–353
10. Tye, B. K., and Sawyer, S. (2000) *J. Biol. Chem.* **275,** 34833–34836
11. You, Z., Komamura, Y., and Ishimi, Y. (1999) *Mol. Cell. Biol.* **19,** 8003–8015
12. You, Z., and Masai, H. (2005) *Nucleic Acids Res.* **33,** 3033–3047
13. Liu, J., Smith, C. L., DeRyckere, D., DeAngelis, K., Martin, G. S., and Berger, J. M. (2000) *Mol. Cell* **6,** 637–648
14. Fletcher, R. J., Bishop, B. E., Leon, R. P., Sclafani, R. A., Ogata, C. M., and Chen, X. S. (2003) *Nat. Struct. Biol.* **10,** 160–167
15. Pape, T., Meka, H., Chen, S., Vicentini, G., van Heel, M., and Onesti, S. (2003) *EMBO Rep.* **4,** 1079–1083
16. Hofmann, J. F., and Beach, D. (1994) *EMBO J.* **13,** 425–434
17. Maiorano, D., Moreau, J., and Méchali, M. (2000) *Nature* **404,** 622–625
18. Nishitani, H., Lygerou, Z., Nishimoto, T., and Nurse, P. (2000) *Nature* **404,** 625–628
19. Cook, J. G., Chasse, D. A., and Nevins, J. R. (2004) *J. Biol. Chem.* **279,** 9625–9633
20. Ferenbach, A., Li, A., Brito-Martins, M., and Blow, J. J. (2005) *Nucleic Acids Res.* **33,** 316–324
21. Tanaka, S., and Diffley, J. F. (2002) *Nat. Cell Biol.* **4,** 198–207
22. Yanagi, K., Mizuno, T., You, Z., and Hanaoka, F. (2002) *J. Biol. Chem.* **277,** 40871–40880
23. Arias, E. E., and Walter, J. C. (2006) *Nat. Cell Biol.* **8,** 84–90
24. Wohlschlegel, J. A., Dwyer, B. T., Dhar, S. K., Cvetic, C., Walter, J. C., and Dutta, A. (2000) *Science* **290,** 2309–2312
25. O'Connell, B. C., and Harper, J. W. (2007) *Curr. Opin. Cell Biol.* **19,** 206–214
26. Lee, C., Hong, B., Choi, J. M., Kim, Y., Watanabe, S., Ishimi, Y., Enomoto, T., Tada, S., Kim, Y., and Cho, Y. (2004) *Nature* **430,** 913–917
27. Cunningham, E. L., and Berger, J. M. (2005) *Curr. Opin. Struct. Biol.* **15,** 68–76
28. Dueber, E. L., Corn, J. E., Bell, S. D., and Berger, J. M. (2007) *Science* **317,** 1210–1213
29. Gaudier, M., Schuwirth, B. S., Westcott, S. L., and Wigley, D. B. (2007) *Science* **317,** 1213–1216
30. You, Z., and Masai, H. (2008) *J. Biol. Chem.* **283,** 24469–24477
31. Cavanagh, J., Fairbrother, W. J., Skelton, N. J., and Palmer, A. G. (1996) *Protein NMR Spectroscopy*, Academic Press, San Diego, CA
32. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR* **6,** 277–293
33. Johnson, B. A. (2004) *Methods Mol. Biol.* **278,** 313–352
34. Güntert, P., Mumenthaler, C., and Wüthrich, K. (1997) *J. Mol. Biol.* **273,** 283–298
35. Pearlman, D. A., Case, D. A., Caldwell, J. W., Ross, W. S., Cheatham, T. E., III, DeBolt, S., Ferguson, D., Seibel, G., and Kollman, P. (1995) *Comp. Phys. Commun.* **91,** 1–41
36. Herrmann, T., Güntert, P., and Wüthrich, K. (2002) *J. Mol. Biol.* **319,** 209–227
37. Jee, J., and Güntert, P. (2003) *J. Struct. Funct. Genomics* **4,** 179–189
38. Cornilescu, G., Delaglio, F., and Bax, A. (1999) *J. Biomol. NMR* **13,** 289–302
39. Bashford, D., and Case, D. A. (2000) *Annu. Rev. Phys. Chem.* **51,** 129–152
40. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) *J. Biomol. NMR* **8,** 477–486
41. DeLano, W. L. (2002) *The PyMOL Molecular Graphics System*, DeLano Scientific LLC, San Carlos, CA
42. Edgar, R. C. (2004) *Nucleic Acids Res.* **32,** 1792–1797
43. Holm, L., and Sander, C. (1993) *J. Mol. Biol.* **233,** 123–138
44. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) *Nucleic Acids Res.* **28,** 235–242
45. Jones, D. T. (1999) *J. Mol. Biol.* **292,** 195–202
46. de Vries, S. J., van Dijk, A. D., and Bonvin, A. M. (2006) *Proteins* **63,** 479–489
47. Gajiwala, K. S., and Burley, S. K. (2000) *Curr. Opin. Struct. Biol.* **10,** 110–116
48. Riedl, S. J., Li, W., Chao, Y., Schwarzenbacher, R., and Shi, Y. (2005) *Nature* **434,** 926–933
49. Im, Y. J., Wollert, T., Boura, E., and Hurley, J. H. (2009) *Dev. Cell* **17,** 234–243
50. Kasiviswanathan, R., Shin, J. H., and Kelman, Z. (2005) *Nucleic Acids Res.* **33,** 4940–4950
51. Capaldi, S. A., and Berger, J. M. (2004) *Nucleic Acids Res.* **32,** 4821–4832
52. Wilce, J. A., Vivian, J. P., Hastings, A. F., Otting, G., Folmer, R. H.,

Duggin, I. G., Wake, R. G., and Wilce, M. C. (2001) *Nat. Struct. Biol.* **8,** 206–210
53. Randell, J. C., Bowers, J. L., Rodriguez, H. K., and Bell, S. P. (2006) *Mol. Cell* **21,** 29–39
54. Orengo, C. A., and Thornton, J. M. (2005) *Annu. Rev. Biochem.* **74,** 867–900
55. Khayrutdinov, B. I., Bae, W. J., Yun, Y. M., Lee, J. H., Tsuyama, T., Kim, J. J., Hwang, E., Ryu, K. S., Cheong, H. K., Cheong, C., Ko, J. S., Enomoto, T., Karplus, P. A., Güntert, P., Tada, S., Jeon, Y. H., and Cho, Y. (2009) *Protein Sci.* **18,** 2252–2264
56. Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98,** 10037–10041

Gene Therapy and Regulation
Vol. 5, No. 1 (2010) 57–65

DOI: 10.1142/S1568558610000173

World Scientific
www.worldscientific.com

# GENE THERAPY FOR PARKINSON'S DISEASE: STRATEGIES FOR THE LOCAL PRODUCTION OF DOPAMINE*

SHIN-ICHI MURAMATSU[†,§], SAYAKA ASARI[†], KEN-ICHI FUJIMOTO[†], KEIYA OZAWA[‡] and IMAHARU NAKANO[†]

[†]*Division of Neurology, Department of Medicine*
*and*
[‡]*Genetic Therapeutics, Center for Molecular Medicine*
*Jichi Medical University, Tochigi, Japan*
[§]*muramats@jichi.ac.jp*

Received 17 July 2010
Accepted 2 August 2010

The cardinal motor symptoms of Parkinson's disease (PD) are associated with the profound depletion of dopamine in the striatum. The replacement of dopamine is the most straightforward strategy to improve motor performance in PD. Researchers have been developing gene therapy aimed at local production of dopamine via the introduction of dopamine-synthesizing enzyme genes into the putamen. Two phase I clinical studies have used recombinant adeno-associated virus (AAV) vectors to transfer the aromatic L-amino acid decarboxylase (AADC) gene into the putamen to restore efficient conversion of orally administered L-3,4-dihydroxyphenylalanine (L-dopa). The initial results of these studies have not only confirmed the safety of AAV vectors, but have also demonstrated the alleviation of motor symptoms associated with PD. Interestingly motor performance in the "off" medication state was improved after gene therapy, suggesting long-term modulation of dopaminergic signals in the striatal neurons was induced by gene transfer. Gene delivery of tyrosine hydroxylase (TH) and guanosine triphosphate cyclohydrolase I (GCH) in addition to AADC may help to avoid motor fluctuations associated with intermittent intake of L-dopa by continuously supplying dopamine in the putamen. A clinical study of such triple gene transfer is presently underway using equine infectious anemia virus (EIAV) vector.

*Keywords*: Adeno-associated virus; aromatic L-amino acid decarboxylase; L-dopa; guanosine triphosphate cyclohydrolase I; positron emission tomography.

## Parkinson's Disease

Parkinson's disease (PD) is second only to Alzheimer disease as the most common neurodegenerative disorder among the elderly, with an estimated 1% of the population over 60 years old suffering from PD and a lifetime risk of 6.7% in men (Driver *et al.*, 2009). The pathological hallmarks of PD are the presence of Lewy bodies,

*Invited review article.
§Corresponding author.

cytoplasmic inclusions, in the substantia nigra pars compacta (SNc) neurons that project to the striatum and the loss of these neurons. The main protein component of Lewy bodies is α-synuclein, which accumulates in a phosphorylated and aggregated form (Dickson *et al.*, 2009).

The causes of PD remain largely unknown, although genetic causes have been elucidated in some familial cases including mutations in the gene encoding α-synuclein, leucine-rich repeat kinase 2 (LRRK2), or ubiquitin carboxy-terminal hydrolase-L1 (UCH-L1) in families with an autosomal dominant pattern of inheritance, and mutations in the genes encoding parkin, PTEN-induced putative kinase 1 (PINK1), or DJ-1 in families with an autosomal recessive pattern of inheritance (Nuytemans *et al.*, 2010). Duplication or triplication of the *SNCA* gene, which encodes α-synuclein, gives rise to late-onset and early-onset familial PD, thus suggesting the expression level of α-synuclein might be an important determinant of disease onset and severity. Although genome-wide association studies successfully revealed some susceptibility genes (Tsuji, 2010), purely genetic causes probably account for only a small number of PD patients and multiple factors including environmental factors may contribute to the development of sporadic PD. Postmortem investigations demonstrate that the rate of decrease of nigral neurons is fast in the initial stage of the disease, namely about 40–50% are lost in the first decade, with possibly a slower rate of degeneration occurring thereafter and finally approaching a normal age-related linear decline (Fearnley and Lees, 1991). Recent imaging studies using radiotracers for nigrostriatal nerve terminals support this progression pattern, thus suggesting the mechanisms underlying PD initiation and progression are probably different (Bruck *et al.*, 2009; Nandhagopal *et al.*, 2009). The reason for the selective susceptibility of nigral dopaminergic neurons and the temporal sequence of events leading to cell loss in PD, however, remain to be elucidated.

## Dopamine Synthesis in the Striatum

Dopamine is synthesized almost exclusively in the terminals of nigrostriatal neurons in the normal striatum. Three enzymes are necessary for efficient dopamine synthesis (Fig. 1): tyrosine hydroxylase (TH), aromatic L-amino acid decarboxylase (AADC), and guanosine triphosphate cyclohydrolase I (GCH). L-Tyrosine is converted to L-3,4-dihydroxyphenylalanine (L-dopa) by TH in the first rate-limiting step. AADC then converts L-dopa to dopamine. GCH is the rate-limiting enzyme for synthesis of the essential TH co-factor tetrahydrobiopterin ($BH_4$). Since low levels of endogenous $BH_4$ do not yield sufficient TH activity, GCH is considered to regulate TH activity via regulation of $BH_4$ biosynthesis, thus indirectly controlling dopamine production in TH-containing neurons (Nagatsu *et al.*, 1987). These three enzymes are transported from the SNc to the striatum in an anterograde manner.

The cardinal symptoms in PD including resting tremor, muscular rigidity, and bradykinesia become apparent after the 40–50% of the SNc neurons are lost and striatal dopamine is reduced to about 20% of normal levels. A severe loss of

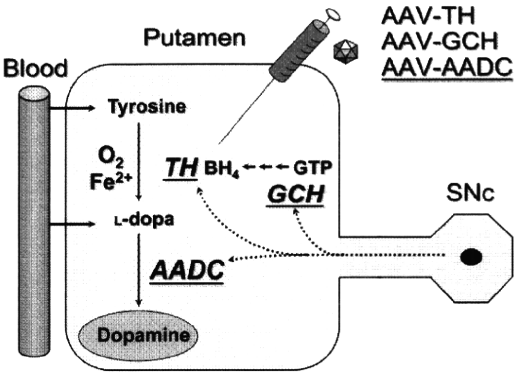



**Fig. 1. Biosynthesis pathway of dopamine.** Three enzymes are necessary for the efficient production of dopamine. Dopamine precursor, L-dopa is a standard drug for treating Parkinson disease. AADC, aromatic L-amino acid decarboxylase; $BH_4$, tetrahydrobiopterine; GCH, guanosine triphosphate cyclohydrolase I; TH, tyrosine hydroxylase; SNc, substantia nigra pars compacta.

dopaminergic nerve terminals in advanced PD is associated with an 80–95% depletion of the striatal TH and AADC activity (Zhong *et al.*, 1995; Nagatsu and Sawada, 2007) thus leading to a profound decrease of dopamine. There are several types of AADC-containing cells in the striatum, such as serotonin neurons, intrinsic dopamine neurons, AADC-containing "D" neurons, and glial cells. These cells may act as a local source of dopamine. However, endogenous AADC activity in the striatum is thought to be insufficient and the functional efficacy of dopamine produced from exogenous L-dopa in these cells may be limited at least in primates. The activity of GCH in the striatum has also been reported to decrease in PD (Nagatsu and Sawada, 2007). The restoration of the GCH activity is necessary to supply sufficient $BH_4$ in advanced PD, since the uptake of exogenous $BH_4$ from the blood is low (Hoshiga *et al.*, 1993) and the primary source of $BH_4$ in the brain is intracellular biosynthesis.

## Complications of Long-Term L-dopa Therapy

The current accepted therapeutic strategy for PD is the replacement of dopamine in the striatum to alleviate motor dysfunction. Unlike dopamine, which does not cross the blood-brain barrier, the dopamine precursor, L-dopa can be transported into the brain and is the most effective drug in pharmacotherapy for PD. Virtually all patients experience a clinically meaningful benefit after receiving L-dopa treatment. However, as the disease progresses, the loss of AADC activity and the decreased capacity for dopamine storage in the synaptic vesicles lead to the failure of L-dopa therapy. Frequent systemic administration of high doses of L-dopa causes oscillations in motor performance with a variety of abnormal involuntary movements or dyskinesia (Fox and Lang, 2008). After 4–6 years of L-dopa treatment, 40 to 50% of patients are estimated to have motor complications (Ahlskog and Muenter, 2001). The diagnosis of idiopathic PD may be incorrect if a patient does