December 2009 *Chem. Pharm. Bull.* **57**(12) 1309—1334 (2009)

Review

Structural Studies of Specific Intermolecular Interactions and Self-Aggregation of Biomolecules and Their Application to Drug Design

Toshimasa Ishida

Osaka University of Pharmaceutical Sciences; 4–20–1 Nasahara, Takatsuki, Osaka 569–1094, Japan.

Received August 3, 2009

Information on the structural basis of intermolecular recognition or self-aggregation of biomolecules at the atomic level is important to understand biological functions and to develop devices for treating disorders caused by abnormal functions. Thus structural analysis of specific intermolecular or intramolecular interactions of biomolecules has been performed using various physicochemical approaches. Herein, the following three subjects are reviewed: (1) structural analyses of mRNA cap structure recognition by eukaryotic initiation factor 4E and its functional regulation by endogenous 4E-binding protein; (2) structural studies of self-aggregation mechanism of microtubule-binding domain in tau protein and aggregation inhibitor; and (3) molecular design of cathepsin B-specific inhibitor.

Key words recognition; aggregation; inhibition; initiation factor 4E; tau protein; cathepsin B

1. Introduction

For a biomolecule in a living cell to fulfill its function, exact interaction with its partner is generally required. Intermolecular interaction is generally achieved through strict recognition based on 3D complementarity, and ambiguities cause unusual biological events that may lead to serious diseases. Therefore structural studies of the conformation and interaction of a biomolecule at the atomic level are indispensable for understanding its biological function and developing clinically usable drugs to control abnormal activity in living cells. Various physicochemical approaches are now available to clarify the 3D features of biomolecules and their interactions. Among them, the combination of X-ray crystallography and solution NMR analyses, with the help of circular dichroism (CD) and fluorescence spectroscopy and energy calculations, is most effective for such investigations. Making the best use of these methods, our group has been performing structural analyses on biologically important molecules and their complexes with partners, hoping to contribute to molecular design for controlling biologically abnormal functions. Among several projects performed thus far, the following are reviewed: (i) mRNA cap structure recognition by eukaryotic initiation factor 4E and its functional regulation by endogenous 4E-binding protein; (ii) the self-aggregation mechanism of the microtubule-binding domain in tau protein; and (iii) molecular design of cathepsin B-specific inhibitor.

2. Structural Analyses of mRNA Cap Structure Recognition by Eukaryotic Initiation Factor 4E (eIF4E) and Its Functional Regulation by Endogenous 4E-Binding Protein (4EBP)

The initiation of translation in eukaryotes involves a complex biochemical pathway catalyzed by a large number of initiation factors.[1] The first step in the translation process involves the binding of a small ribosomal subunit to mRNA. This is the rate-limiting step in translation initiation and is also a target for translational control. Eukaryotic mRNA has a common cap structure at its 5′-terminal end.[2] The cap structure is important for stabilizing mRNA and facilitating its binding to ribosomes in the translation process. In cap-dependent translation, interaction is required between the cap structure and the eukaryotic initiation factor 4F for efficient translation of mRNA. The smallest subunit in eIF4F, eIF4E, binds specifically to the mRNA cap structure and plays an important role in the initiation of protein synthesis (Fig. 1). Elucidation of the selective mRNA cap recognition mechanism by eIF4E is therefore important to understand the initiation mechanism of protein synthesis.

2.1. Background: Prominent Stacking Interaction of Tryptophan (Trp) with N-Protonated Guanine Base

From a series of X-ray structural studies on the interactions of Trp with various biomolecules such as thiamin, flavin, and pyridine coenzymes,[3—6] it was clarified that the π-electron-rich indole ring of Trp can specifically catch the aromatic chromophores of these coenzymes through π–π stacking interactions, suggesting the importance of Trp residues in enzymes as having catalytic rather than structural function.[7] On the other hand, the functional analysis of Trp in the specific interaction with a nucleic base is of special importance,

e-mail: ishida@gly.oups.ac.jp

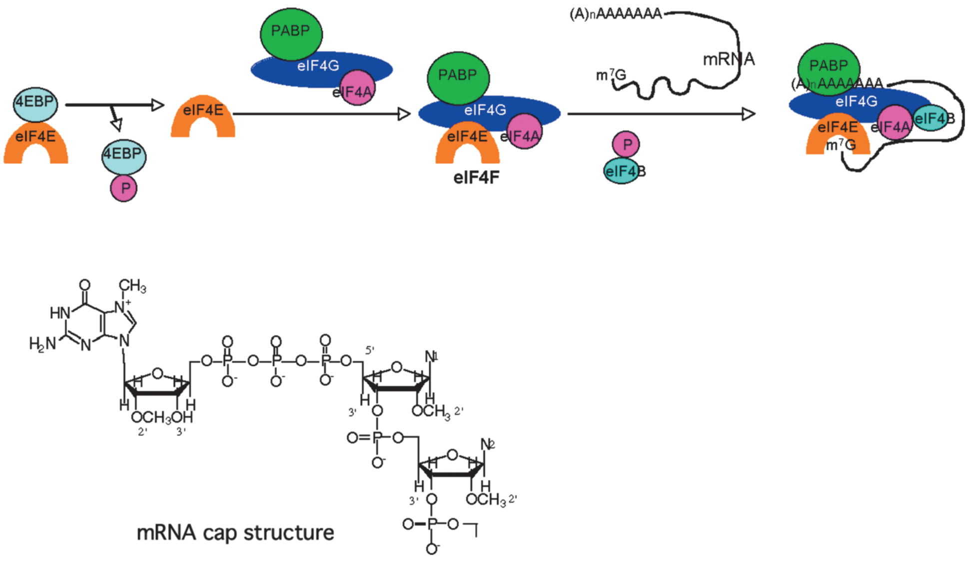


Fig. 1. Scheme of Eukaryotic Cap-Dependent Translation Initiation Pathway[1)]

mRNA is activated by the binding of eIF4F (eIF4E, eIF4G, eIF4A) to the cap at the 5′-terminal and poly A-binding protein (PABP) to the poly A tail at the 3′-terminal. The eIF4F assembly is blocked by 4EBP binding to eIF4E. Phosphorylation of 4EBP dissociates 4EBP from eIF4E, leading to association of eIF4F.

because the selective recognition of a specific nucleotide or nucleic acid sequence by an amino acid residue or its sequence in an enzyme is an essential biological event in the life cycle.

Quantum chemical analysis clarified that the N-quaternization of nucleic acid bases (adenine, guanine, and cytosine) lowers the lowest unoccupied molecular orbital (LUMO) energy level to close to the level of the highest occupied molecular orbital (HOMO) energy of the Trp indole ring, reinforcing the HOMO–LUMO interaction; an example of a guanine base is shown in Fig. 2a. Because this implies that Trp can selectively recognize the N-alkylated nucleic bases through prominent π–π-stacking interactions, much experimental evidence was accumulated by the structural analyses of related model complexes to clarify the function of the Trp residue in the interaction with nucleotide bases.[8—19)] The X-ray crystal structure of the Trp-Glu–N7-methyl guanosine 5′-monophosphate (m^7GMP) complex[20,21)] and the solution NMR structure of the Trp-Leu-Glu–m^7GpppA complex[22,23)] are shown in Fig. 2b. The results suggest the importance of cooperative stacking and hydrogen bond pairing interactions of Trp and Glu residues for the selective recognition of the mRNA cap structure,[24)] in which the cap structure of mRNA is characterized by N7-methyl guanosine linked with the 5′-terminal portion of mRNA *via* a 5′,5′-triphosphate (Fig. 1). Further structural and spectroscopic refinements showed that the sequence –Trp-$(Gly)_m$-Asp/Glu-$(Gly)_n$-Trp– ($m=0$—2, $n=0$—2) is most preferred for the N7-methyl guanine base (m^7G).[25—27)] These model interaction studies enabled us to confirm whether this specific recognition mode of the m^7G base by the cooperation of Trp and acidic amino acid residues is actually applied to the recognition of the mRNA cap structure by eIF4E.

2.2. Gene Construction, Expression, and Purification and Cap-Binding Ability of Recombinant Human eIF4E Human eIF4E (*ca.* 25 kDa) consists of a 217 amino acid sequence containing eight Trp residues (Fig. 3a). To obtain a large amount of recombinant human eIF4E in *Escherichia coli*, an artificial gene encoding human eIF4E was chemically synthesized using the codon used most frequently in *E. coli*, in which the DNA duplex of 662 base pairs was designed then constructed from 44 oligodeoxynucleotide fragments of typically 30 nucleotides in length.[28)] The gene was then cloned into an expression vector and direct expression

Toshimasa Ishida is currently Professor of Osaka University of Pharmaceutical Sciences. He was born in 1946 and graduated from Osaka University of Pharmaceutical Sciences in 1969. After obtaining his M.Sc. degree from Osaka University under the supervision of Professor Emeritus Ken-ichi Tomita in 1975, he joined Osaka University of Pharmaceutical Sciences as an Assistant in 1976. He received his Ph.D. degree from Osaka University under the supervision of the late Professor Emeritus Masatoshi Inoue and Professor Emeritus Ken-ichi Tomita in 1979. He was promoted to Lecturer in 1979 and Assistant Professor in 1982. After working as a postdoctoral fellow of Alexander von Humboldt-Stiftung with Professor Wolfram Saenger at Berlin Free University during the mid 1980s, he returned to Osaka University of Pharmaceutical Sciences and was promoted to full Professor in 1989. He received The Pharmaceutical Society of Japan Award for Young Scientists in 1988 and The Pharmaceutical Society of Japan Award for Divisional Scientific Contributions in 2009. Professor Ishida's research interests are situated in the areas of structural physical chemistry and biochemistry.

Toshimasa Ishida

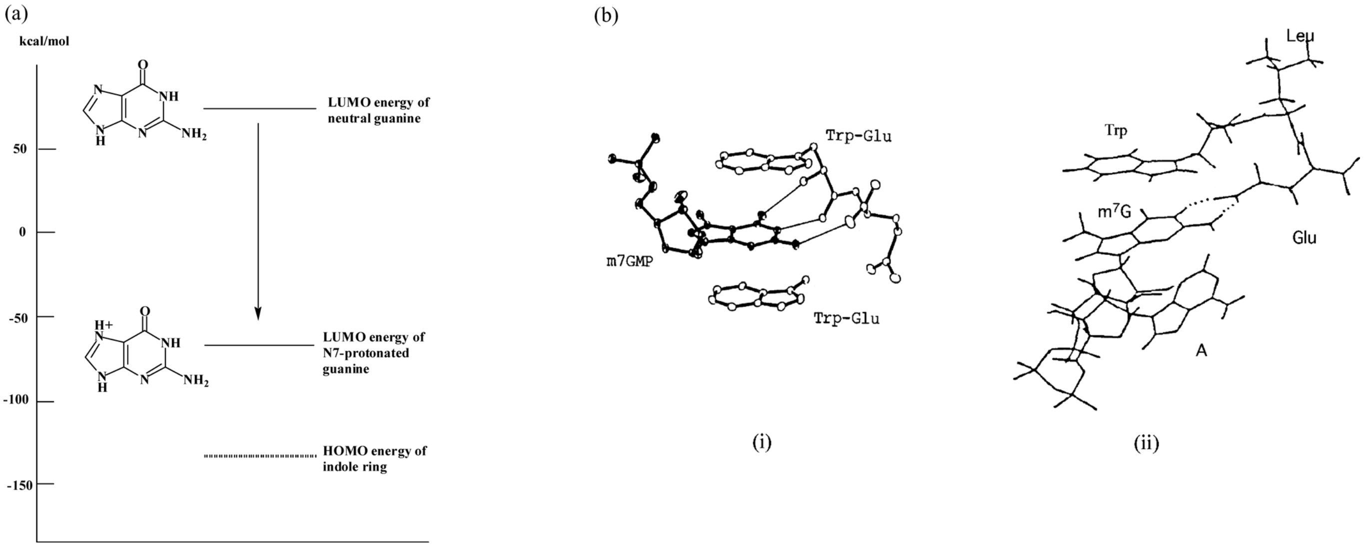
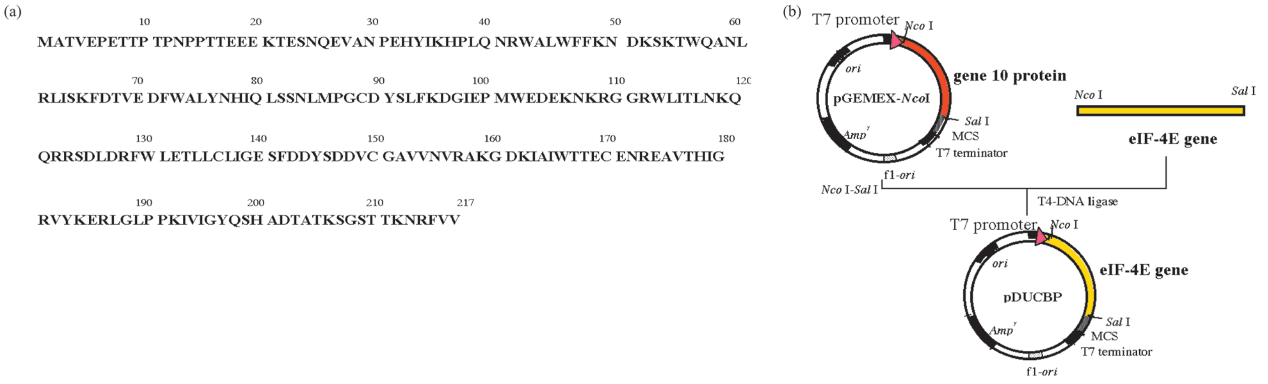


Fig. 2. (a) LUMO Energy Levels of Neutral and N-Quarternized Guanine Bases and HOME Energy Level of Indole Ring, Calculated by MNDO Method and (b) X-Ray Crystal Structure of Trp-Glu–m^7GMP (1 : 1) Complex (i) and Solution ^{1}H-NMR Structure of Trp-Leu-Glu–m^7GpppA (1 : 1) Complex in Deuteriated 25 mM Phosphate Buffer (ii)



MATVEPETTP TPNPPTTEEE KTESNQEVAN PEHYIKHPLQ NRWALWFFKN DKSKTWQANL
RLISKFDTVE DFWALYNHIQ LSSNLMPGCD YSLFKDGIEP MWEDEKNKRG GRWLITLNKQ
QRRSDLDRFW LETLLCLIGE SFDDYSDDVC GAVVNVRAKG DKIAIWTTEC ENREAVTHIG
RVYKERLGLP PKIVIGYQSH ADTATKSGST TKNRFVV

Fig. 3. (a) Amino Acid Sequence of Human eIF4E and (b) Construction of Direct Expression pDUCBP Plasmid

Plasmid pGEMEX-*Nco*I was digested with *Nco*I and *Sal*I, and synthetic human eIF4E gene (*Nco*I–*Sal*I fragment) was inserted into the vector by ligation.

was attempted in *E. coli* [BL21(DE3)] under the control of the T7 promoter (Fig. 3b).[29,30] The active gene product was purified to homogeneity by m^7GTP-affinity chromatography and confirmed to be recombinant human eIF4E by amino acid composition and sequence analyses and isoelectric focusing measurements. Identity in 3D structures between the recombinant and native human eIF4Es was confirmed by comparing their CD spectra and association constants with m^7G cap analogues by fluorescence measurements.[31,32)]

The amino acid residues required for the mRNA cap binding of eIF4E were analyzed by a series of amino acid mutations of recombinant eIF4E. Each of the 15 functional residues conserved among various eIF4E species was replaced with a nonpolar amino acid (Ala or Leu), and their cap-binding abilities were compared with that of the wild type. Consequently, all eight Trp residues, Glu103, and two histidine residues at positions 37 and 200 were shown important for mRNA cap binding.[32—34)]

2.3. Crystallization and X-Ray Crystal Structure of Human–m^7GTP and –m^7GpppA Complexes After concentrating the recombinant human eIF4E (7 mg/ml), the solution was subjected to crystallization. The eIF4E–m^7GpppA and eIF4E–m^7GTP complexes were finally crystallized at 15 °C by the hanging drop vapor diffusion method, in which the protein solution was equilibrated against a reservoir solution containing 100 mM *N*-(2-hydroxyethyl)piperazine-*N*′-2-ethanesulfonic acid (HEPES)–KOH (pH 7.5), 27% polyethylene glycol (PEG)-6000, and 0.2 M $(NH_4)_2SO_4$. Long needle-shaped crystals were obtained after 1 week.[35)] The crystals diffracted ≤1.9 Å resolution. The initial structures of both complex crystals were determined by the molecular replacement method using the starting structure of N-terminal region-deleted mouse eIF4E.[36)] The *R*-factors for the final structure refinement for eIF4E–m^7GTP and eIF4E–m^7GpppA complexes were 19.2% and 19.6%, respectively.[37,38)]

The overall tertiary structure of the eIF4E–m^7GpppA complex and the mutual disposition of secondary structural elements of α-helices (H1—H3) and β-sheets (S1—S8) are shown in Fig. 4. Because of the conformational flexibility of the N-terminal region, it was not possible to determine the structure from the N-terminal to Gln-26, and no specific secondary structural element was observed in the Glu-27–His-37 sequence. Human full-length eIF4E consists of an $\alpha+\beta$ structural motif similar to RNA-binding protein and resem-

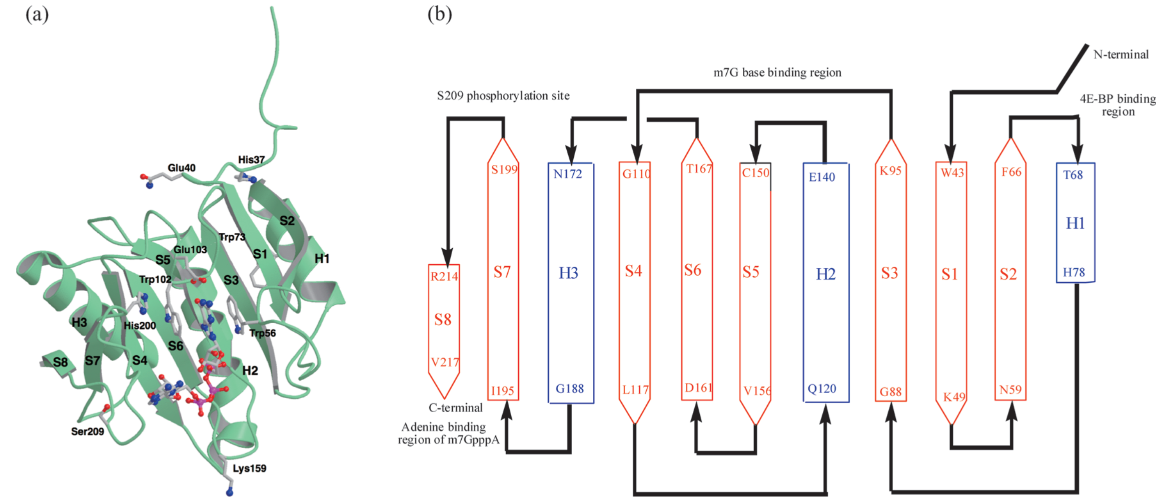


Fig. 4. Overall Tertiary Structure of eIF4E–m^{7}GpppA Complex (a) and Mutual Disposition of Secondary Structure Elements of α-Helices (H1—H3) and β-Sheets (S1—S8)

m^{7}GpppA in (a) is depicted with the stick-bond model.

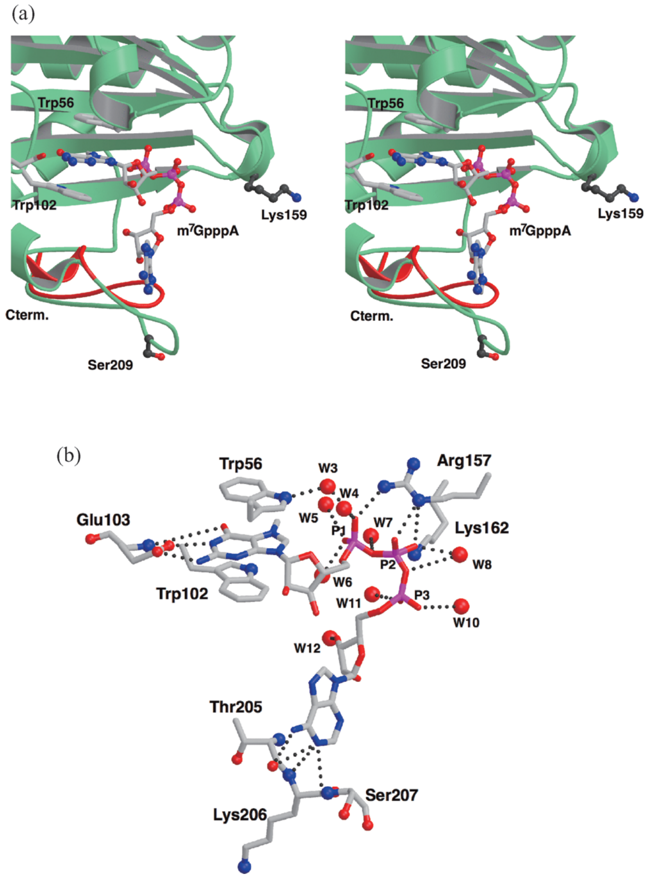


Fig. 5. (a) Stereo View of Cap-Binding Pocket of eIF4E and Cap Structure-Dependent Flexibility of C-Terminal Loop and (b) Binding Mode of m^{7}GpppA

The green rope of the C-terminal moiety in (a) represents the location in the complex with m^{7}GpppA, while the red rope shows the location in the complex with m^{7}GTP. Hydrogen bonds or short contacts in (b) are shown by dotted lines. The first, second, and third phosphorus atoms of m^{7}GpppA are numbered P1, P2, and P3, respectively.

bles a bowl-shaped structure with a concave surface as a whole. The cap-binding pocket of eIF4E and the detailed interaction mode of m^{7}GpppA at the cap-binding pocket are shown in Fig. 5; nearly the same interaction mode was also observed in the m^{7}GTP complex, except for the adenosine moiety. It is remarkable to note that the recognition mechanism of the mRNA cap structure by eIF4E is almost the same as that proposed from the model studies. The m^{7}G moiety of the cap structure is sandwiched between two aromatic side chains of Trp-56 and Trp-102. Furthermore, two (m^{7}G)NH$\cdots$O(Glu-103 carboxyl) hydrogen bonds stabilize the stacking interaction; this result clearly indicates that the model study is a useful approach to analyze the mechanism of a complicated biological function. The first P1 and second P2 phosphate groups of m^{7}GpppA are linked to the basic amino acid residues located at the cap-binding pocket (Arg-112, Arg-157, and Lys-162) through direct or water-mediated hydrogen bonds. The third phosphate group P3 of m^{7}GpppA forms a hydrogen bond with the Arg-112 side chain *via* a water molecule.

In contrast to the well-conserved core structure of the cap-binding region, the C-terminal region showed cap structure-dependent conformational flexibility. As compared with the m^{7}GTP complex, the adenine nucleoside located in the C-terminal loop region in the m^{7}GpppA complex is hydrogen-bonded to Thr-205 and Thr-211 (*via* a water molecule) and further stabilized by van der Waals contacts with Ser-207 and its neighboring residues, indicating that the nucleotide or its sequence linked to the m^{7}Gppp moiety of the cap structure could regulate the flexibility and orientation of the C-terminal loop region. For example, the spatial disposition of this region shifted to widen the cap-binding slot of eIF4E by the adenosine moiety of m^{7}GpppA (Fig. 5a).

It is generally believed that the phosphorylation of eIF4E takes place mainly at the O$^\gamma$ atom of Ser-209 and plays a role in controlling the switch-on/off of association between eIF4E and mRNA.[39)] The molecular dynamics (MD) simulation[38)] clearly showed that the phosphorylated Ser-209 could form a retractable salt bridge with Lys-159 to clamp the cap moiety of mRNA because the loop region and Lys-159 side chain are both flexible, suggesting that the charge change around the cap-binding site caused by Ser-209 phosphorylation increases the ability to bind to capped mRNA.

2.4. X-Ray Crystal Structure of 4EBP1 Fragment–

Human eIF4E–m^{7}GpppA Ternary Complex The biological function of eIF4E is regulated by association and dissociation with endogenous 4EBP (Fig. 1).[40] Three isoforms have been isolated for 4EBP (Fig. 6). Although each isoform binds eIF4E, the functional difference between the isoforms has not yet been fully elucidated. To investigate the regulation mechanism of the eIF4E function by 4EBP at the atomic level, the crystal structure of the m^{7}GpppA–eIF4E–4EBP1 peptide (residues 36—70) ternary complex was analyzed by X-ray diffraction[41]; similar crystal structures have been reported for m^{7}GDP-bound murine eIF4E (residues 28—217)–4EBP1 peptide (residues 51—67) and –human eIF4GII peptide (residues 621—637) ternary complexes.[42] Crystal structure analysis at 2.1 Å resolution (Fig. 7a) revealed that the 4EBP1 peptide (Pro47–Pro66) is located at the N-terminal dorsal molecular surface of the eIF4E cap-binding pocket. The eIF4E-binding moiety of the 4EBP1 peptide adopted a distorted 'reverse L'-shaped conformation (Fig. 7b), in which the Arg51–Asp55 and Arg56–Cys62 sequences form β-strand and α-helical secondary structures, respectively. The 4EBP1 peptide was mainly fixed by three structural parts of eIF4E, *i.e.*, the His37–Gln40 (N-terminal loop), Val69–Asn77 (H1 helix), and Glu132–Ser141 (H2 helix) regions. The Trp73 residue of eIF4E plays an important role in the interaction: (indole)NH$\cdots$O(Leu59) hydrogen bond and (indole)$\pi\cdots$C=O (Arg63) edge-to-face interaction. The Arg56–Cys62 helical region of the peptide was further stabilized by the Arg56 N$^{\varepsilon}$H$\cdots$Gly132 O$^{\varepsilon 2}$ hydrogen bond and the C–H$\cdots\pi$ interactions between Phe58 benzene and His37 imidazole rings. Tyr54 also played an important role in the interactions *via* the Tyr54 O$^{\eta}$H$\cdots$O Pro38 and Tyr54 NH$\cdots$O Gly139 hydrogen bonds and the edge-to-face C–H$\cdots\pi$ interactions with the Val69 side chain.

4EBP1 MSGGSSCSQTPS--RAIPATRRVVLGDGVQLPPGDYSTTPGGTLFSTTPGGT 50
4EBP2 MSSSAGSGHQPSQSRAIP-TRTVAISDAAQLPH-DYCTTPGGTLFSTTPGGT 50
4EBP3 MSTSTSC--------PIP-------GGRDQLPD-CYSTTPGGTLYATTPGGT 36

4EBP1 RII**YDRKFLM**ECRNSPVTKTPPRDLPTIPGVTSPSS--DEPPMEASQSHLRN 100
4EBP2 RII**YDRKFLL**DRRNSPMAQTPPCHLPNIPGVTSPGTLIEDSKVEVNNLNNLN 102
4EBP3 RII**YDRKFLL**ECKNSPIARTPPCCLPQIPGVTTPPT------APLSKLEELK 82

4EBP1 SPEDKRAGGEESQFEMDI 118 (12.6kDA)
4EBP2 NHDRKHAVGDDAQFEMDI 120 (12.9kDA)
4EBP3 EQETEEEIPDDAQFEMDI 100 (10.9kDA)

human eIF4GI(609–660)
KKR**YDREFLL**GFQFIFASMQKPEGLPHISDVVLDKANKTPLRPLDPTRLOOI

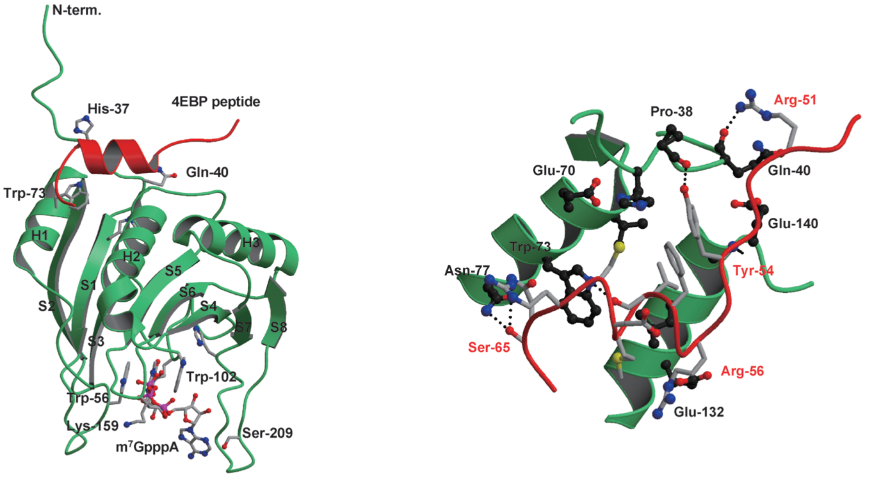
Fig. 6. Amino Acid Sequences of Three 4EBP Isoforms

The sequences of 4EBP1—3 isoforms are aligned on the basis of their functional similarities. The eIF4E-binding Y(X)$_4$Lϕ sequence and His74–Glu89 of 4EBP2 or its corresponding sequence of 4EBP1 or 4EBP3 are shown with bold letters and rectangular boxes, respectively. The 609—660 sequence of human eIF4GI is also shown for comparison.

The Tyr-X-X-X-X-Leu-ϕ sequence (X is any amino acid and ϕ is Leu, Met, or Phe), which is known as the eIF4E recognition motif of eIF4Gs and 4EBPs,[43] corresponds to the sequence of Tyr54-Asp55-Arg56-Lys57-Phe58-Leu59-Met60 in this peptide. Tyr54 and Leu59 contributed importantly to the interaction with eIF4E by providing the structural basis for the recognition motif.

On the other hand, it is known that phosphorylation of Ser65 and Thr70 of 4EBP weakens the binding ability to eIF4E, leading to dissociation from the complex. X-ray analysis showed hydrogen bond formation between Ser65 O$^{\gamma}$H (4EBP1) and Asn77 N$^{\varepsilon}$H$_2$(eIF4E). Phosphorylation of Ser65 will cause electrostatic repulsion between these two residues in eIF4E and thus lead to a decrease in binding ability with eIF4E (Fig. 7b).

To investigate the structural effect of the 4EBP1-binding region in eIF4E on the cap binding of eIF4E, molecular dynamics (MD) simulations were performed for eIF4E–4EBP1 peptide binary and m^{7}GpppA–eIF4E–4EBP1 peptide ternary complexes in aqueous solution.[41] Comparison of the time profiles of the selected atomic distances and torsion angles suggested that: (i) the β strand and α helical conformations



(a) (b)

Fig. 7. (a) Overall View of m^{7}GpppA–eIF4E–4EBP1 Peptide Ternary Complex and (b) Binding Mode of 4EBP1 Peptide with eIF4E

(a) View along β-sheets composing the cap-binding pocket. The eIF4E and 4EBP1 peptides are shown by the ribbon model with green and red colors, respectively, and m^{7}GpppA is shown by the stick-bond model. Three α-helices and eight β-sheets of eIF4E are labeled H1—H3 and S1—S8, respectively. Side chain orientations of His37, Trp56, Trp73, Trp102, Glu103, Lys159, and Ser209 residues are shown by the stick-bond model. 4EBP1 peptide in (b) is shown by red-colored wire. The binding region of eIF4E is constructed by the N-terminal extended (His37–Gln40, green-colored wire), H1 (Trp73–Asn77) and H2 (Glu132–Gly139) helical regions (green-colored ribbons). The functional residues of eIF4E and 4EBP1 peptides are shown by the ball-and-stick models. Dotted lines represent possible hydrogen bonds.

Table 1. SPR Kinetic Parameters of Full-Length 4EBP2 and Its N-, C-, or Both-Terminal-Residue-Deleted Mutants with m^{7}GTP-Bound Full-Length eIF4E at 25 °C$^{a,b)}$

4EBP2 mutant	k_a (M$^{-1}\cdot$s^{-1})	k_d (s^{-1})	K_D (M)
Full-length	$2.42\pm0.05\times10^5$	$6.63\pm0.42\times10^{-4}$	$2.74\pm0.23\times10^{-9}$
Nd12-4EBP2	$3.18\pm0.04\times10^5$	$3.38\pm0.25\times10^{-4}$	$1.20\pm0.10\times10^{-9}$
Nd32-4EBP2	$1.63\pm0.02\times10^5$	$5.43\pm0.40\times10^{-4}$	$3.33\pm0.28\times10^{-9}$
Nd45-4EBP2	$1.70\pm0.02\times10^5$	$7.61\pm0.28\times10^{-4}$	$4.48\pm0.22\times10^{-9}$
Cd9-4EBP2	$3.07\pm0.03\times10^5$	$2.84\pm0.41\times10^{-4}$	$9.23\pm1.41\times10^{-10}$
Cd31-4EBP2	$1.15\pm0.03\times10^5$	$3.75\pm0.23\times10^{-4}$	$3.26\pm0.28\times10^{-9}$
Cd47-4EBP2	—	—	$1.88\pm0.30\times10^{-6}$
Cd55-4EBP2	—	—	$4.21\pm0.79\times10^{-6}$
Nd45/Cd31-4EBP2	$1.55\pm0.05\times10^5$	$8.79\pm0.44\times10^{-4}$	$5.65\pm0.45\times10^{-9}$
Peptide I	—	—	$2.21\pm0.21\times10^{-6}$
Peptide II	n.d.	n.d.	n.d.
Peptide III	n.d.	n.d.	n.d.

Peptide I: Thr46–Thr70 (4EBP2), peptide II: His74–Glu89 (4EBP2), peptide III: Gly632–Thr647 (eIF4GI). *a*) The kinetic parameters were estimated by 1 : 1 (Langmuir) binding model. The K_D (M) values of Cd47-4EBP2, Cd55-4EBP2, and peptide I were estimated by static Scatchard affinity model. *b*) The K_D (M) value of the full-length 4EBP2–eIF4E pair estimated using a Scatchard model was $3.4\pm1.9\times10^{-9}$, and no notable difference was observed from the value determined using the Langmuir model.

of 4EBP are important for the binding to eIF4E; (ii) the cap structure stabilizes the binding of eIF4E with 4EBP; (iii) the 4EBP–eIF4E complex regulates the structural fluctuation at the cap-binding slot of eIF4E and consequently facilitates cap binding; and (iv) phosphorylation of Ser67 alone, although it is insufficient to release 4EBP from eIF4E, significantly increases fluctuation of the 4EBP structure. These insights suggest that the flexible N- and C-terminal moieties of eIF4E change their structures easily so as better to accommodate their biological partners, *i.e.*, 4EBP/eIF4G and mRNA cap structures, and to create a functionally active supramolecule.

2.5. Role of 4EBP C-Terminal Region in Interaction with eIF4E The N- and C-terminal sequences remain unstructured or structurally flexible in the crystal structure of the m^{7}GpppA–eIF4E–4EBP1 peptide ternary complex. However, the N-terminal region of eIF4E has been reported to participate primarily in the conformational coupling between eIF4E and eIF4G to enhance ribosome loading onto the mRNA cap[44)] and also to be necessary for the optimal regulation of insulin-simulated phosphorylation[45)] and for the functional regulation by 4EBP isoform.[46)] To clarify the roles of the N- and C-terminal flexible regions in the interaction with eIF4E, the binding features of the sequentially N-, C-, or both-terminal-residue-deleted 4EBP2 mutants were investigated by surface plasmon resonance (SPR).[47)] The constants of association rate, k_{on}, dissociation rate, k_{off}, and the equilibrium constant for dissociation, K_D ($=k_{off}/k_{on}$) are given in Table 1, where the data for the peptides of the Thr46–Thr70 (including Y(X)$_4$Lϕ binding motif) and His74–Glu89 sequences of 4EBP2 and of the Gly632–Thr647 sequence of eIF4GI are also given for comparison (see also Fig. 6).

Concerning the role of the N-terminal flexible region (1—45 sequence) of 4EBP2 in the interaction with eIF4E, the SPR kinetic parameters showed no notable contribution to the interaction, because all SPR profiles showed similar patterns despite the difference in the length of deletion, where the K_D values of Nd12-, Nd32-, and Nd45-4EBP2 were almost on the same order as that of full-length 4EBP2. The contrast SPR profiles were observed for the C-terminal-residue-deleted 4EBP2s. Characteristic differences in the interaction with full-length eIF4E were observed between the SPR profiles of Cd31- and Cd47-4EBP2s (Fig. 8) and nearly identical differences were also observed between the SPR profiles of full-length 4EBP2 and Cd55-4EBP2 (data not shown). The SPR analyses show that deleting more than the C-terminus 47 residues of 4EBP2 weakens the interaction with eIF4E by more than three orders of magnitude compared with the full-length 4EBP2. In contrast, the C-terminal deletion of <31 residues showed nearly the same kinetic parameters as those of full-length 4EBP2. This marked difference indicates that the 32—47 sequence from the C-terminus (corresponding to His74–Glu89 sequence in Fig. 6) participates primarily in the interaction with eIF4E.

To examine how strongly this sequence interacts with eIF4E and to compare its strength with that of the Y(X)$_4$Lϕ sequence, the kinetic parameters of peptide I (Thr46–Thr70) and peptide II (His74–Glu89) for the interaction with eIF4E were measured (Table 1). Consequently, the peptide I–eIF4E pair showed a dissociation constant on the order of *ca.* 10^{-6}, the association of which is about three orders of magnitude weaker than that of Cd31-4EBP2. On the other hand, peptide II showed no notable SPR signal, indicating a lack of binding ability of the His74–Glu89 sequence with eIF4E. Therefore, there is clearly a supportive role of the His74–Glu89 sequence in the primary binding of the Thr46–Thr70 sequence including the Y(X)$_4$Lϕ binding motif to eIF4E, where the auxiliary effect of the His74–Glu89 sequence is sufficiently effective to elevate the interaction to a K_D of *ca.* 10^{-9}.

To estimate the possible binding site of the 4EBP2 His74–Glu89 sequence to eIF4E, MD simulation of m^{7}GTP–eIF4E–4EBP2 peptide (Thr46–Glu89) ternary complex in aqueous solution was carried out (Fig. 9).[47)] Consequently, it was shown that the Cys73–Glu89 sequence of 4EBP2 binds stably to the root (Ile35–Pro38) of the N-terminal flexible region of eIF4E through relatively weak hydrophobic interactions. Since the Arg51–Ser65 sequence including the Y(X)$_4$Lϕ consensus region binds stably to the 4EBP-binding site of eIF4E *via* hydrogen bonds and hydrophobic interactions, as observed in the X-ray crystal structure of the complex,[41)] the Thr46–Glu89 sequence of 4EBP2 results in a right-handed helical ring formation; a similar bracelet structure in solution was observed for the 393—490 sequence of yeast eIF4G.[44)] It is noteworthy that similar auxiliary role could be expected for the C-terminal sequence in 4EBP1 and 4EBP3 and the eIF4E-binding region of human eIF4GI

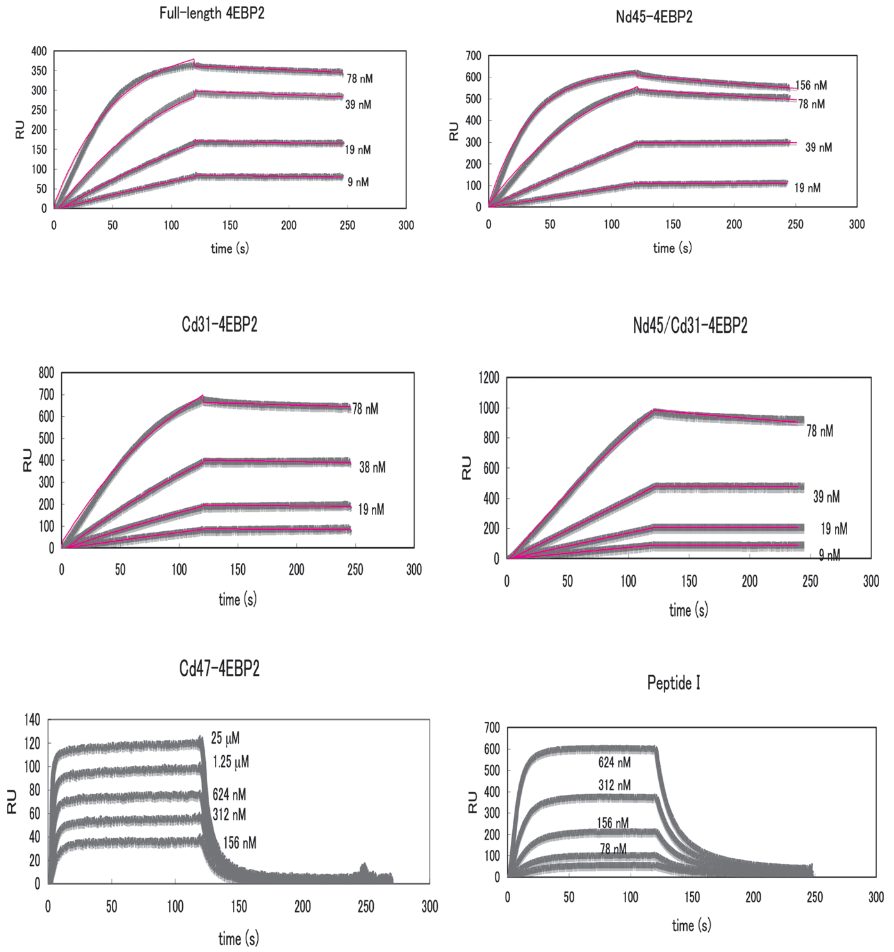


Fig. 8. Experimental SPR Sensorgrams (Grey-Colored Wide Lines) and Calculated Fits (Red Lines) of Full-Length, Nd45-, Cd-31-, and Nd45/Cd31-4EBP2–m^{7}GTP-bound eIF4E Pairs, Together with Sensorgrams of Cd47-4EBP2– and peptide 1–eIF4E Pairs

The right and left parts of each sensorgram indicate the phases of injection of eIF4E and exposure of the chip to eIF4E-free buffer, respectively.

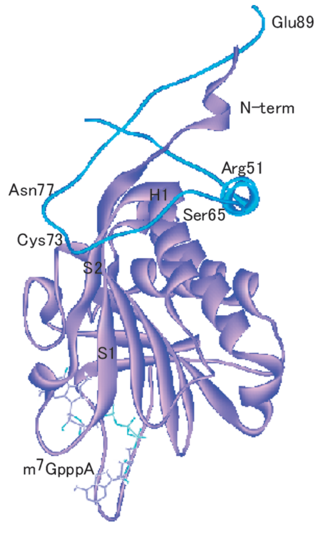


Fig. 9. Possible Interaction between eIF4E (Purple) and 4EBP2 Thr46–Glu89 Sequence (Blue) by MD Simulation (Snapshot at 3 ns)

The molecules are shown using the ribbon and wire model, and m^{7}GpppA is shown using the stick-bond model. H1 and S2 on eIF4E denote the S2-sheet and H1-helix secondary structures (see Fig. 7). N-term indicates the N-terminal side of eIF4E.

(shown with rectangular box in Fig. 6). Indeed, a similar auxiliary role has been confirmed for the C-terminal sequence in 4EBP1 and 4EBP3 (Park K., *et al.*, unpublished data).

3. Structural Studies of Self-Aggregation Mechanism of Microtubule-Binding Domain in Tau Protein and Aggregation Inhibitor

Alzheimer's disease (AD) is the most common cause of dementia in the elderly population. AD is accompanied by a number of structural and metabolic alterations in the brain and is characterized by two histopathological hallmarks: extracellular deposits of β-amyloid in neuritic plaques and intracellular neurofibrillary tangles (NFTs).[48] The latter are composed of bundles of paired helical filaments (PHFs) resulting from the abnormal aggregation of tau protein.

Tau protein, a neuronal microtubule-associated protein (MAP) in the mammalian brain, has six isoforms consisting of 352—441 amino acids in length; these isoforms are produced from a single human tau gene by the alternative mRNA splicing of exons 2, 3, and 10 on chromosome 17.[49] The microtubule binding domain (MBD) of tau protein is located in the C-terminal half (Fig. 10). It is composed of three or four structural repeats (3RMBD or 4RMBD, respectively), where each single-repeat peptide (named R1 to R4, from the first to the fourth repeat, respectively) consists of relatively

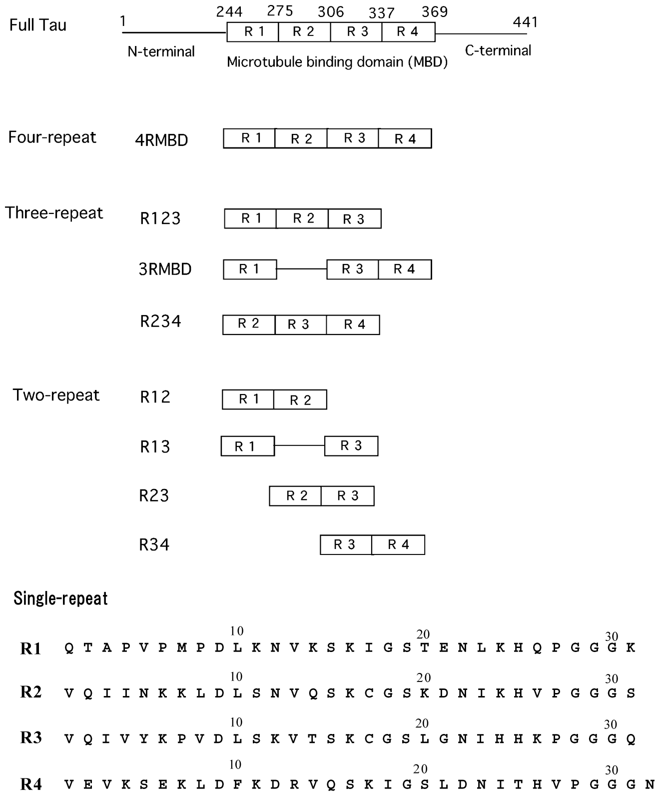


Fig. 10. Schematic Diagram of Full-Length Tau Protein, Wild-Type Four- and Three-Repeat MBDs (4RMBD and 3RMBD), and Non-wild Type Three- and Two-Repeat MBDs, Together with Amino Acid Sequence of Each Single-Repeat Peptide

similar and conserved amino acid sequences of 31—32 amino acids.

Tau protein, by nature, is a highly soluble protein with a random conformation in aqueous solution and barely shows any tendency to assemble under physiological conditions. In the brains of AD patients, however, it dissociates from axonal microtubules as a result of hyperphosphorylation and abnormally aggregates with itself to form insoluble PHF, which is implicated in neurodegeneration.[50] Therefore one priority in AD research is to elucidate the PHF formation mechanism of tau protein. Thus many *in vivo* and *in vitro* studies have been performed to uncover any clues for developing anti-AD drugs. However, satisfactory success has not yet been achieved, and some important questions still remain unanswered. For example, is tau aggregation based on formation of β-sheet structure or some other structures? What is the building block of filament formation? What is different in the filament formation pathway among tau isoforms? These issues must be further investigated.

As a useful approach to achieve this end, it appears important to clarify which parts and what structural changes in the tau protein are most responsible for the transformation from its native random structure to PHF. Because the MBD repeat structure is known to constitute the core moiety of the PHF structure,[51] we have studied the functional and conformational contribution of each repeat peptide to the aggregation of MBD, although the filament of each repeat peptide may not directly reflect the intact feature of the tau PHF structure.

It is now known that (i) PHF formation by tau is accelerated to an experimentally allowable time with the help of polyanions such as heparin,[52] and (ii) the fluorescence of a dye such as thioflavin S (ThS) permits quantitative monitoring of PHF formation by tau.[53] To clarify the functional and conformational contributions of each single-repeat peptide (R1—R4) to the PHF formation of wild-type 3RMBD (R134) or 4RMBD (R1234), we investigated heparin-induced aggregation and conformational transition behaviors of single-, two-, and three-repeat peptides by a combination of ThS fluorescence, dynamic light scattering (DLS), circular dichroism (CD), and electron microscopy (EM) in neutral buffer (pH=7.6).[54—59] The relation among these experimental approaches and the filament formation process is shown in Fig. 11. The MBD single-repeat peptides (*i.e.*, R1, R2, R3, R4) were chemically synthesized by solid-phase peptide synthesizer. The remaining MBD samples were prepared as His-tagged recombinant proteins by their gene expressions in *E. coli* and were purified by chromatography.

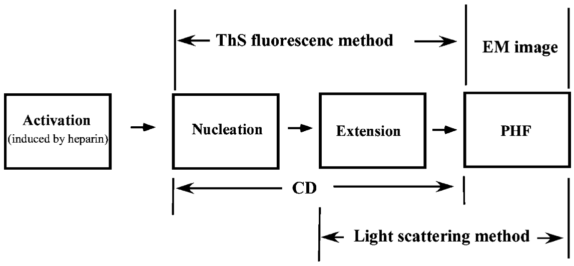


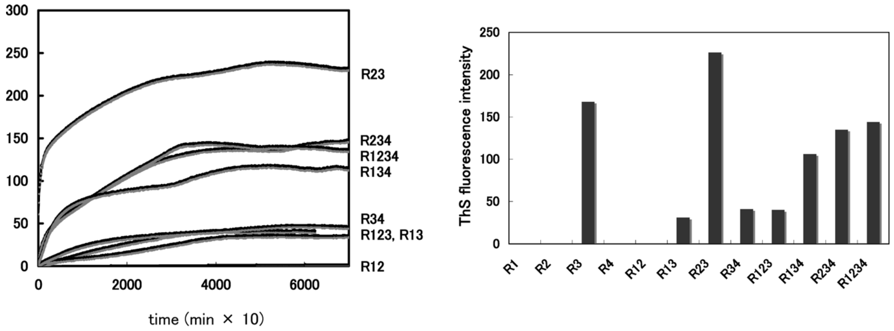
Fig. 11. Relationship between Filament Formation Process of Tau, MBD, or Repeat Peptide and ThS Fluorescence, LS, CD, or EM Measurement

3.1. Self-Aggregation Behaviors of Single- and Two-Repeat Peptides Monitored by ThS Fluorescence Intensity The ThS fluorescence profiles of the MBD repeat peptides were measured as a function of time after addition of heparin (Fig. 12), in which the ThS intensity after 1 h of initiating the reaction is also shown. The kinetic parameters characterizing the aggregation profiles are given in Table 2.[58] The self-aggregation rate and quantity showed the following order:

R3≫R2=R1=R4=0	for single-repeat peptide
R23≫R34>R13≫R12	for two-repeat peptide
R234>R134≫R123	for three-repeat peptide
R23≫R3≫4RMBD=R234>3RMBD≫R34=R123>R13>R12>R2 =R1=R4=0	for all repeat peptides

Concerning the single-repeat peptides, we conclude the following: (i) R3 has the highest rate and fastest self-aggregation ability, suggesting its main role in filamentous aggregation of MBD; (ii) although R2 takes approximately 6 h to aggregate, the quantity of aggregation increases by nearly the same extent as that of R3 after about 28 h; and (iii) R1 and R4 do not self-aggregate in neutral solution. It is noteworthy that R4 forms a filament similar to those of R2 and R3 in acidic solution, whereas R1 still shows resistance to filament formation.[57]

On the other hand, it is remarkable that the two-repeat R23 shows the fastest rate and largest aggregation volume among all the peptides. Although the aggregation behavior of the single-repeat peptide clearly suggests that R3 plays a main role in filamentous aggregation, the observed order of



(a) (b)

Fig. 12. (a) ThS Fluorescence Profiles of MBD Repeat Peptides, except for Single-Repeat Peptides, as a Function of Time and (b) Comparison of Their ThS Fluorescence Intensities at 1 h after Addition of Heparin

ThS fluorescence intensities of repeat peptides (25 μM concentration in 50 mM Tris–HCl buffer, pH 7.6) were monitored as a function of time (h) at 25 °C after addition of 6.25 μM heparin, where the excitation and emission wavelengths were set at 440 and 550 nm, respectively. The aggregated quantities of repeat peptides were estimated by their ThS fluorescence intensities 1 h after addition of heparin.

Table 2. Parameters of Defining Characteristic of ThS Fluorescence Increases and CD Spectral Changes Accompanying Aggregations of Repeat Peptides[a]

Repeat peptide	ThS fluorescence measurement			CD measurement		
	Time of starting filament formation	Aggregation rate (min^{-1})[b]	Saturation time (h)	Lag time for showing CD change	Time for completion of transition	Isochromatic point
R1[c]	—	—	—	—	—	—
R2	*ca.* 6 h	0.1	*ca.* 28	*ca.* 10 h	*ca.* 38 h	n.d.
R3	*ca.* 0 min	0.5	*ca.* 9	*ca.* 0.3 h	*ca.* 2.5 h	209 nm
R4[c]	—	—	—	—	—	—
R12	1.5 h	0.1	*ca.* 20	*ca.* 6 h	*ca.* 30 h	n.d.
R13	*ca.* 0 min	0.8	0.5	*ca.* 1 h	*ca.* 9 h	219 nm
R23	*ca.* 0 min	Biphasic: 155.5, 5.9	0.4	*ca.* 0.3 h	*ca.* 5 h	214 nm
R34	*ca.* 0 min	3.87	0.6	*ca.* 0.7 h	*ca.* 5 h	209 nm
R123	*ca.* 0 min	2.25	0.4	*ca.* 0.3 h	*ca.* 8 h	n.d.
3RMBD(R134)	*ca.* 0 min	26.93	0.2	*ca.* 0.3 h	*ca.* 5 h	202 nm
R234	*ca.* 0 min	Biphasic: 27.8, 9.2	0.4	*ca.* 0.5 h	*ca.* 6 h	209 nm
4RMBD(R1234)	*ca.* 0 min	Biphasic: 32.0, 5.5	0.4	*ca.* 0.2 h	*ca.* 3 h	n.d.

a) The same experiments were repeated at least three times using fresh samples, and the reproducibility of all given values was confirmed to be within the error of ±10%. *b*) The value equals the slope of ThS fluorescence intensity as a function of reaction time after heparin addition. *c*) No fluorescence intensity and no CD spectral change is observed.

$R23 \gg R3 \gg R2$ $(=0)$ means that the linkage of R3 and R2 produces a novel ability as a trigger for promotion of aggregation. The aggregation ability of R23 could not be explained by the simple summation of the abilities of R2 and R3. Therefore it is reasonable to conclude that R23 itself behaves as an independent and elemental aggregation unit for MBD.[58)]

3.2. Role of Each Repeat Peptide in Aggregation of Wild-Type 3RMBD and 4RMBD 1) R3 and R23 Repeats Are Aggregation Units of 3RMBD and 4RMBD, Respectively: The aggregation of single repeat peptides was observed for R2 and R3, but not R1 and R4, in neutral solution, which suggests that intermolecular disulfide bond formation between the Cys residues of R2 or R3 is a major step for initiating filament formation. To confirm this possibility, the effect of dithiothreitol (DTT), a reducing agent, on aggregation was investigated.[54,59)] The aggregation kinetics of 3RMBD and 4RMBD were investigated by the DLS method, because this method shows the effect of DTT more directly than the ThS fluorescence method. Figure 13A shows the change in the average hydrodynamic radius (R_H) accompanied by aggregation of 3RMBD or 4RMBD as a function of time after addition of heparin. When the changes in the presence and absence of DTT are compared, the difference between the aggregation processes of 3RMBD and 4RMBD becomes more prominent. Aggregation of 4RMBD is markedly enhanced by addition of DTT, suggesting that the oxidative 4RMBD is not as suitable as its reductive form for aggregation; this is in contrast to the marked suppression of 3RMBD aggregation by DTT. The DTT effect on changes in the average hydrodynamic radii of other repeats is shown in Fig. 13B. It is obvious from this figure that DTT-induced increase in 4RMBD aggregation is commonly observed for the R23-containing repeats (R23, R123, R234), indicating that the filamentous behavior of 4RMBD is dominated by the R23 repeat, in which the intramolecularly disulfide-bonded monomer between the R2 and R3 repeats is the starting structure of 4RMBD aggregation under physiological conditions; this is also suggested by SDS-PAGE analysis (Fig. 14). On the other hand, the DTT-induced decrease in 3RMBD aggregation is similarly observed for R3 and R13; a similar decrease by DTT was also observed for R2 aggregation (data not shown). This shows that the aggregation of 3RMBD is dependent on the R3 single repeat, in which the intermolecularly disulfide-bonded dimer between the neighboring two R3 repeats is the fundamental unit structure for 3RMBD aggregation.

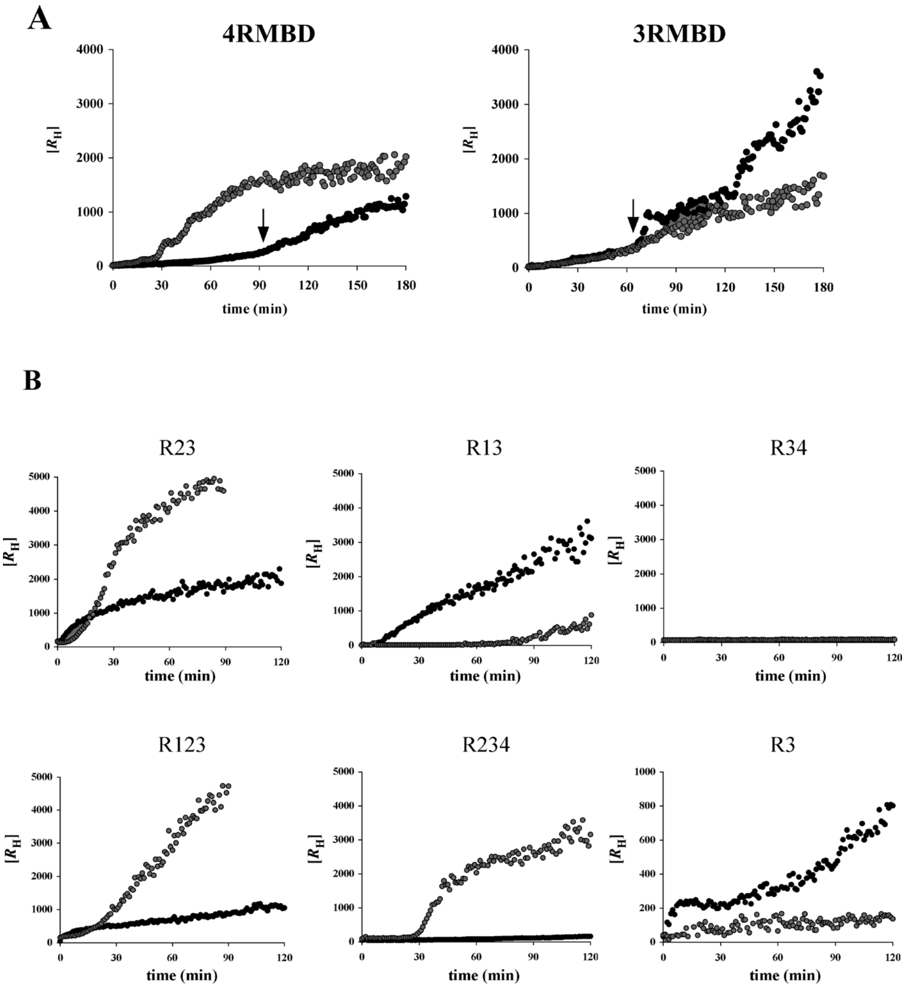


Fig. 13. (A) Changes in Z-Average Hydrodynamic Radius (R_H, nm) of 4RMBD and 3RMBD and (B) Other Repeat Peptide Aggregates as a Function of Aggregation Time in Absence (Black) and Presence (Grey) of DTT

The arrows in (A) indicate the inflection points of aggregation curves.

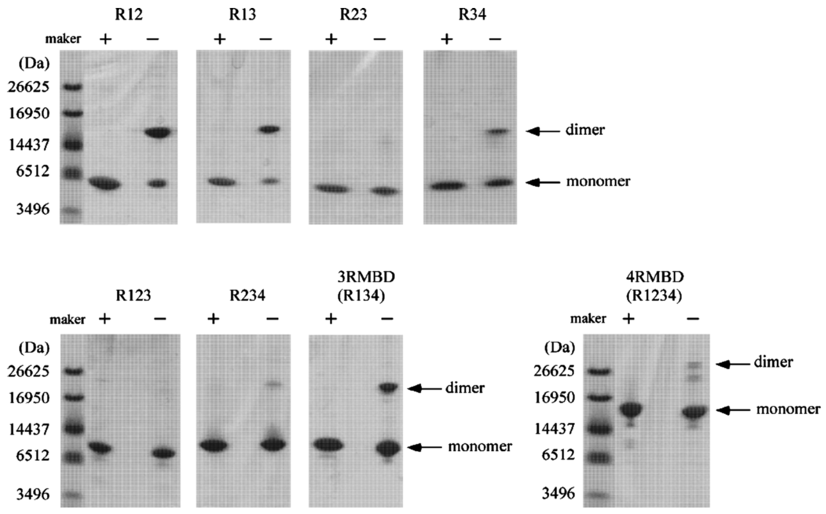
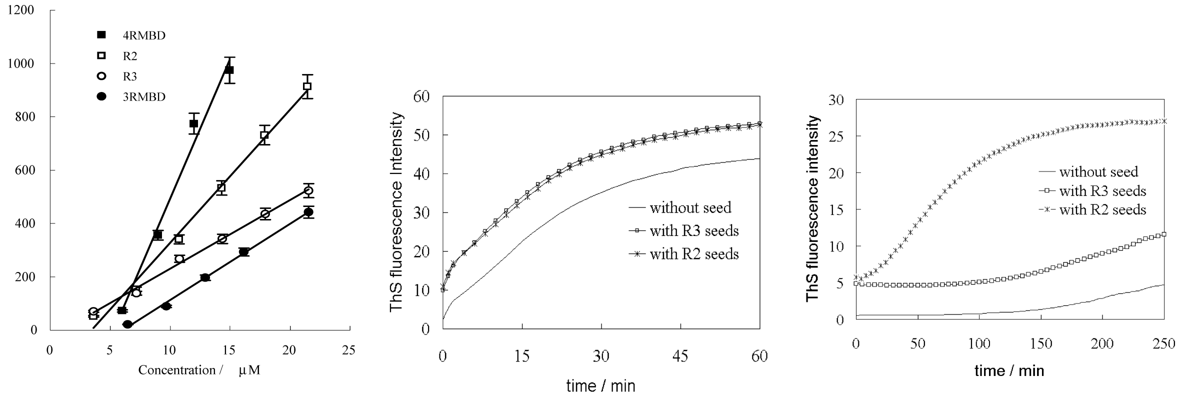


Fig. 14. Dimer and Monomer Distributions of MBD Repeats by SDS-PAGE

The signs (+) and (−) indicate the presence and absence of DTT, respectively.

Note that DTT enhances 4RMBD aggregation to nearly the same extent as that of oxidative 3RMBD (Fig. 13), meaning that the structure of the 4RMBD monomer unrestricted by scission of the intramolecular disulfide bond gains potent ability to induce noncovalent aggregation.

2) Repressive Function of R1 and R4 for Aggregation of 3RMBD and 4RMBD: Although the aggregation abilities of R1 and R4 are absent in neutral solution, they exert different



(a) (b) (c)

Fig. 15. (a) Concentration Dependence of LS Intensities for R2 (Open Squares), R3 (Open Circles), 3RMBD (Filled Circles), and 4RMBD (Filled Squares) and Aggregation Profiles of (b) Seeded R3 and (c) R2 as a Function of Reaction Time

Each repeat peptide in (a) was prepared using 50 mM Tris–HCl buffer (pH 7.5). After adding 3.8 μM heparin, the peptide was incubated at 37 °C overnight. The intensity of LS was plotted as a function of concentration, with the excitation and emission wavelengths both set at 550 nm. The observed intensities were fitted to straight lines with $r>0.98$ by linear regression.

repressive effects on overall aggregation of MBD. Comparison of R13 with R34 and of R123 with R234 in the absence of DTT (Fig. 13) shows prominent inhibitory effect of R4 on aggregation, while the inhibitory effect of R1 is less pronounced. On the other hand, the coexistence of R1 and R4 in 3RMBD and 4RMBD generates a unique behavior inherent in their filament formation, in which the presence of R1 compensates largely for the inhibitory effect of R4.

3) Seeding and Filament Extension Effects of R2 Repeat on Self-Aggregation: The filament extension effects of four single repeats and wild-type 3RMBD and 4RMBD on self-aggregation were examined on the basis of their concentration-dependent LS intensities.[54,55] The results are given in Fig. 15a and Table 3, where the *x*-intercepts correspond to the critical concentrations of the respective MBD repeats for initiating the filament extension process within the sensitivity of the LS method. It is remarkable that the slope of R2 is two-fold larger than that of R3, clearly indicating potent filament extension ability of the R2 repeat. This tendency is observed for 4RMBD including the R2 repeat, but not for 3RMBD. As is obvious from these results, R2 and R3 repeat peptides show that the filament extension mechanism of R2 is different from that of R3. Further to investigate the biological/structural implication of this difference, the seeding effect on filament formation was investigated. The results are shown in Figs. 15b and c. Because the nucleation step is skipped in this seeding experiment, the result reflects the extension processes of R2 and R3. This figure shows a notable difference between the two peptides. The R2 peptide shows a marked R2-dependent seeding effect, whereas R3 is only slightly affected by the seeding. From these results, it is reasonable to consider that R3 easily self-aggregates without the assistance of any template. In contrast, although the R2 peptide itself does not self-aggregate easily, its filament formation progresses significantly *via* a nucleation step, in which a template of homogeneous aggregates is required. Figure 15c indicates that the R2 seed-dependent aggregation (filament formation) of the R2 peptide is a first-order reaction, *i.e.*, R2 peptide+R2 seed→R2 aggregate with $k=(3.2\pm0.4)\times 10^{6}\,\mathrm{M}^{-1}\,\mathrm{min}^{-1}$. In addition, it may be concluded that the large slope for 4RMBD (Fig. 15a) is due to the seeding effect of the R2 repeat.

Table 3. Slopes and *x*-Intercepts of Repeat Peptides, 3RMBD, and 4RMBD in Fig. 15a

Tau construct	Slope	*x*-Intercept (critical conc. μM)
▨ R1	0	—
▦ R2	49.8±1.3	3.40±0.18
▭ R3	25.7±0.9	1.05±0.21
▥ R4	0	—
▨▭▥ 3RMBC	28.7±0.9	6.06±0.07
▨▦▭▥ 4RMBC wild	104.4±1.6	5.26±0.10

3.3. Conformational Change of MBD Repeat Accompanied by Aggregations To investigate the conformational transitions of each single repeat and its linked peptides accompanied by their heparin-induced filament formation in neutral solution, the CD spectra were measured as a function of time.[56,58] The time-dependent CD spectra for single-, two-, three-, and four-repeat peptides, together with their α-helix, β-sheet, and random structure contents calculated according to the literature,[60] are shown in Figs. 16 and 17. The decrease in the large negative ellipticity $[\theta]$ at 200 nm over time indicates the transition from random to α-helix or β-sheet structure. The parameters for characterizing the conformational transitions of the respective peptides are summarized in Table 2.

According to the changes in the CD spectra, the repeat peptides may be classified into three groups: (I) R3, R23, R34, and R234, which have an isochromatic point at *ca.* 209 nm and show a two-state transition from random to β-sheet structure without any contribution of α-helix structure; (II) R2, R12, R13, R123, and 4RMBD (R1234), which do not have an isochromatic point and show the concomitant transition of random and α-helix structures to β-sheet structure; and (III) R134 (3RMBD), which has an isochromatic point at approximately 202 nm and shows two-state transition from random to α-helix structure without any contribution of β-sheet structure. Since group (I) includes R3, it may be suggested that the overall conformational transition is to some

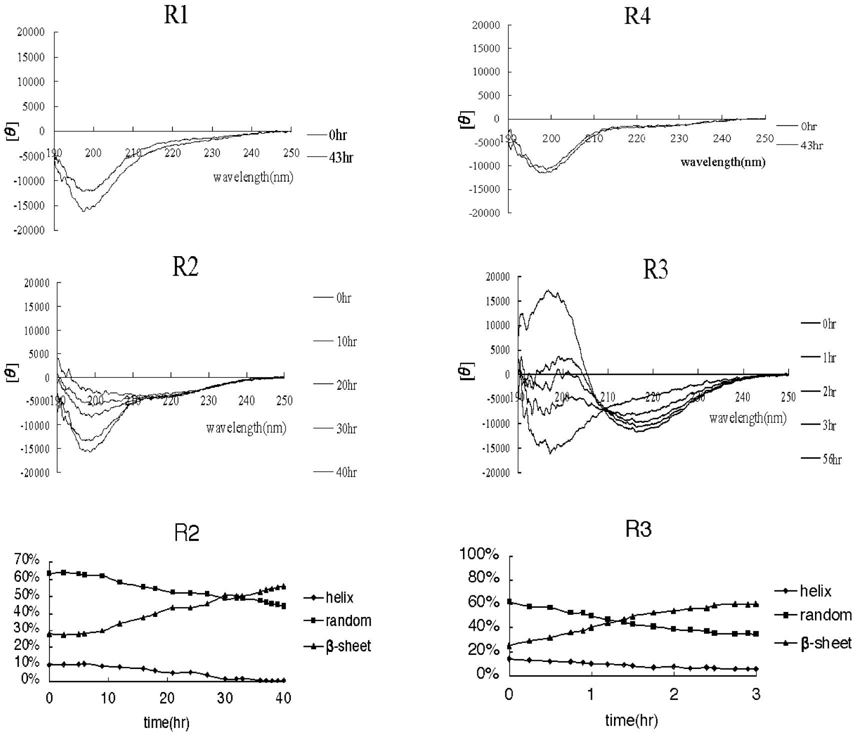


Fig. 16. Time-Dependent CD Spectral Change in R1— R4 and Secondary Structural Variations of R2 and R3 Peptides

The CD spectra for each repeat peptide (from lower to upper negative ellipticity [θ] at 200 nm) correspond to the time indicated in the right side (from upper to lower). The α-helix, random, and β-sheet contents were calculated according to the reference[60] and are marked by diamonds, squares, and triangles, respectively. The CD spectra of each repeat peptide (25 μM concentration in 20 mM phosphate buffer, pH 7.6) were measured as a function of time (h) at 25 °C after addition of 6.25 μM heparin.

extent dominated by the transition of R3. On the other hand, group (II) maintains a large negative peak at around 200 nm and lacks an isochromatic point, and the [θ] value is almost constant in the region of 210—250 nm. The CD comparison of peptides belonging to this group suggests that these spectral features are mainly dependent on the conformational behavior of R1 or R2; R1 shows a similar CD spectrum to R2 at the initial stage and remains unchanged throughout the experiment because it does not form filaments (Fig. 16). Only 3RMBD (R134) could be classified into group (III) on the basis of the following features: an isochromatic point at approximately 202 nm and no contribution of β-structure to filament formation.

Since R3 and R23 were demonstrated the building units for aggregation of wild-type 3RMBD and 4RMBD, respectively, it is important to investigate the effects of R1 and R4 on the conformational transitions of R3 and R23 accompanied by their aggregation. Comparison of the CD spectra of R34 and R13 with R3 and of R234 and R123 with R23 clearly shows significant contribution of R1, but not R4, to the conformational transition. However, the coexistence of R1 and R4 exerts a collaborative contribution, as shown in the CD spectra of 4RMBD and 3RMBD.[56,58]

3.4. Morphologic Effect of Repeat Peptide on MBD Filament EM images of heparin-induced fibrils of the repeat peptides in neutral buffer (pH=7.6) are shown in Fig. 18. The R3 peptide forms short and segmented filaments, whereas R2 forms thin and relatively long filaments; this behavior is due to the difference between the self-assembly speeds and to the behavior inherent in both single-repeat peptides. Comparison of the filamentous features of R2 and R3 with two-, three-, and four-repeat peptides suggests that the morphology of the repeat including R1 (R12, R13, R123, R134, R1234) has a relatively long and thick filament, whereas the morphology of the repeat including R4 (R34, R234) has a relatively short and thin filament; the coexistence of R1 and R4 (R134, R1234) reveals the predominant characteristic of R1. On the other hand, it is noteworthy that R23 forms a fairly thick and long filament, the morphology of which may be explained by the R23-unified characteristics, and not by summation of the characteristics of R2 and R3.[58]

In conclusion, R3 and R23 are the starting repeats for causing nucleation of 3RMBD and 4RMBD, respectively. Coexistence of R1 and R4 repeats is necessary for the aggregation behavior inherent in 3RMBD and 4RMBD, whereas the R1 or R4 repeat alone functions as a repressor or modifier of filament formation. The R2 repeat acts as a promoter for filament extension among the R2 repeats of neighboring MBDs.

3.5. Repeat Sequence and Its Conformation Responsible for Initiating MBD Aggregation[61—66] All the CD spectra of the nonaggregated R1—R4 and their linked peptides show random conformations in aqueous solution, indicating that the wild-type MBD assumes a random structure. Thus it is important for the inhibition of tau PHF formation to clarify which part of MBD and what types of conformation are most responsible for initiating the aggregation. To investigate the relation between the conformational change and aggregation, the filament formation behavior of 4RMBD was investigated as a function of trifluoroethanol (TFE) content, because TFE is known to be an α-helix inducer for peptides. The results are shown in Fig. 19. It is remarkable that 4RMBD is most extensively aggregated at 10—20% TFE.

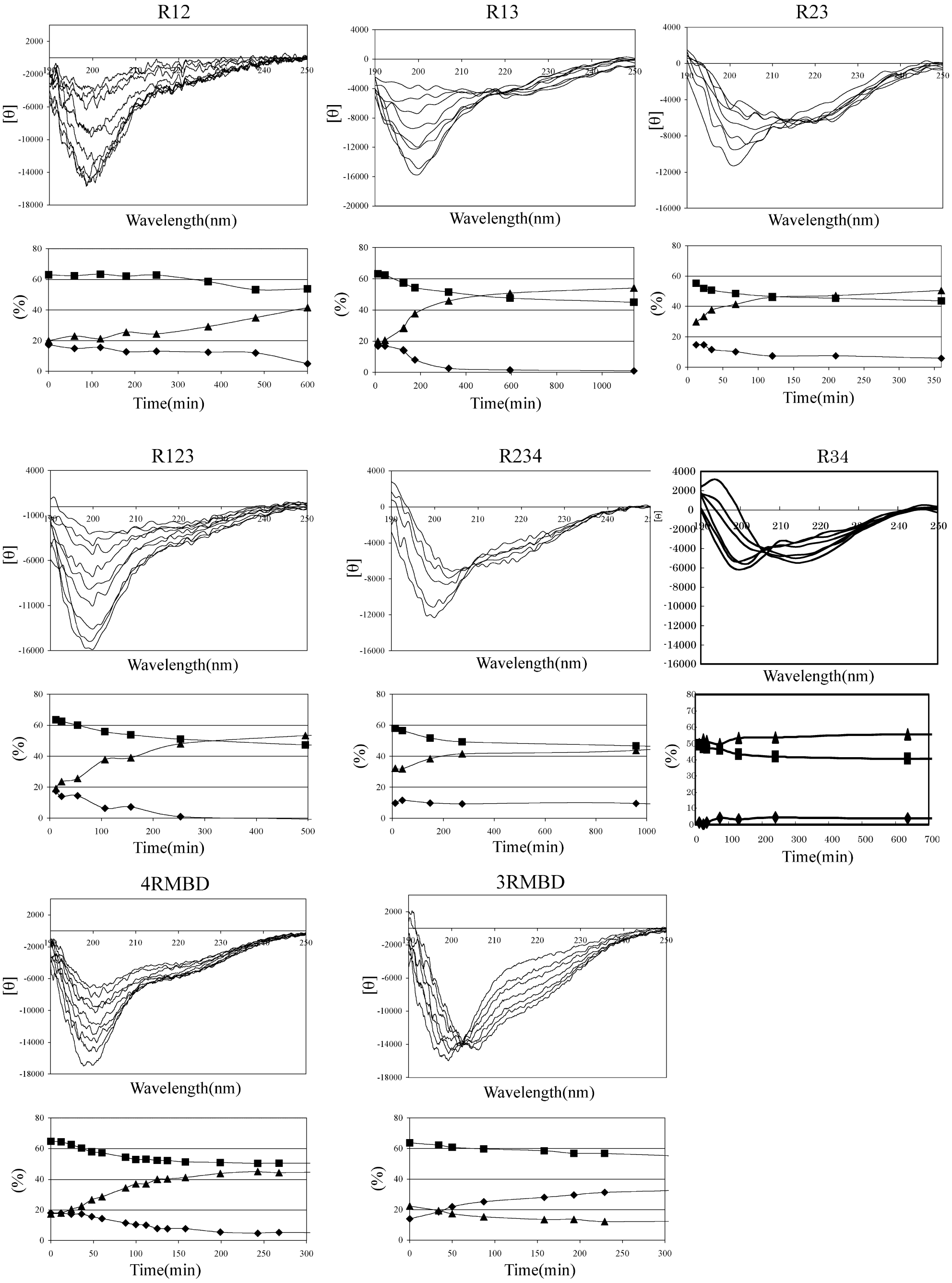


Fig. 17. Time-Dependent CD Spectral Changes (Upper Spectra) and Secondary Structural Variations (Lower Graphs) for Two-, Three-, and Four-Repeat Peptides

The CD spectra for each repeat peptide correspond to the time (horizontal axis) indicated in the secondary structure content profiles. The α-helix, random, and β-sheet contents are marked by diamonds, squares, and triangles, respectively.

The relation of α-helix content to TFE % indicates that the initial stage of transforming the conformation from random to helical structure is most responsible for filament formation.[61,64)]

To investigate how each single-repeat peptide transforms its starting random conformation to the helical structure, the conformation in 100% TFE was analyzed by both ^{1}H-NMR spectroscopy and molecular modeling calculations,[61,63)] because this information is important for extracting the conformational feature responsible for filament formation, although its conformation does not directly reflect the active form for initiating molecular aggregation. Because the 1D and 2D NMR spectral data reflect the ensemble average of various dynamic conformers, conformational comparison of R1—R4 peptides was possible only in terms of the structural features observed commonly in various NMR-constructed conformers. The common features of their backbone conformers and the helical wheel drawings of 10—20 sequences of the most stable conformers, viewed from the N-terminal side, are shown in Fig. 20.

Concerning the R2 and R3 peptides, similar conformational features may be described as follows. The Leu10–Lys20 sequence of R2 forms an α-helical structure, and each amino acid residue of this sequence shows an amphipathic

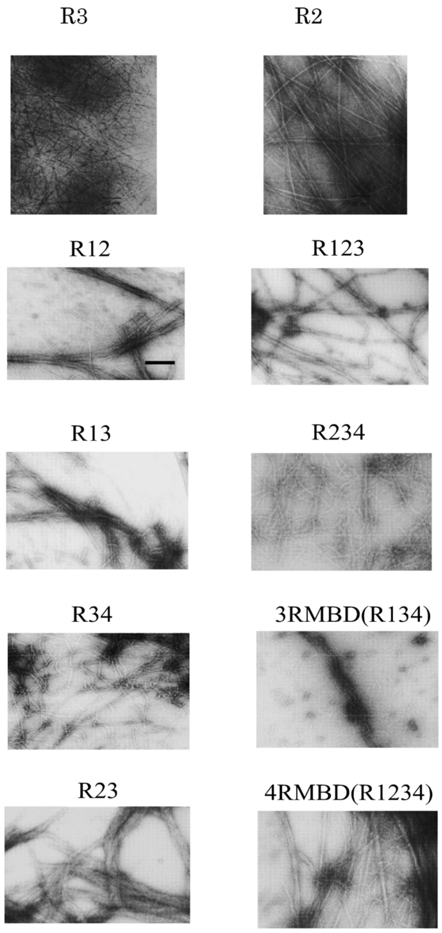


Fig. 18. Negative-Staining EM Images of Filamentous Repeat Peptides (25 μM) Induced by Addition of 6.25 μM Heparin

After incubation at 25 °C for 24 h in neutral buffer (pH 7.6), the samples were negatively stained with 2% uranyl acetate. The length of the bar in R12 corresponds to 200 nm, and the other images are shown at the same scale.

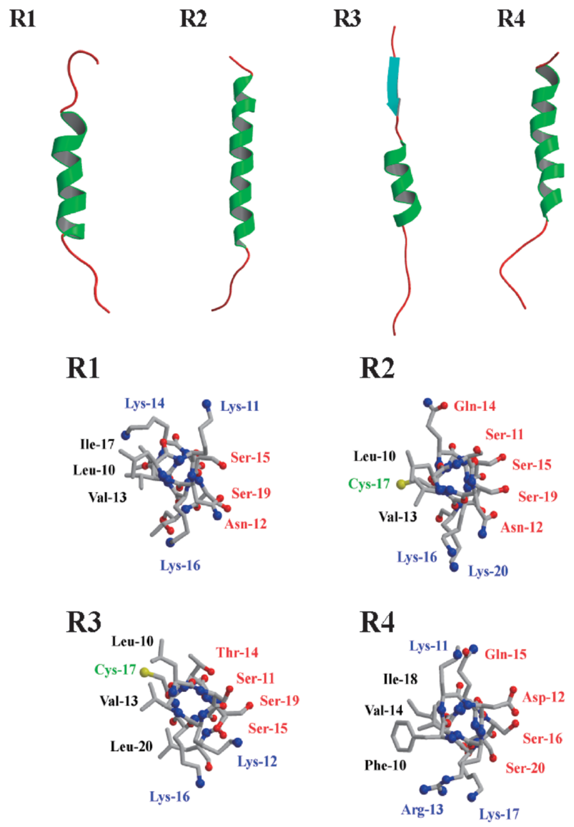


Fig. 20. Averaged Backbone Conformations and Helical Wheel Drawings (10—20 Sequence) of R1—R4 in TFE Solution

The hydrophobic, polar, and basic residues are marked in black, red, and blue, respectively. The Cys residue is shown in green.

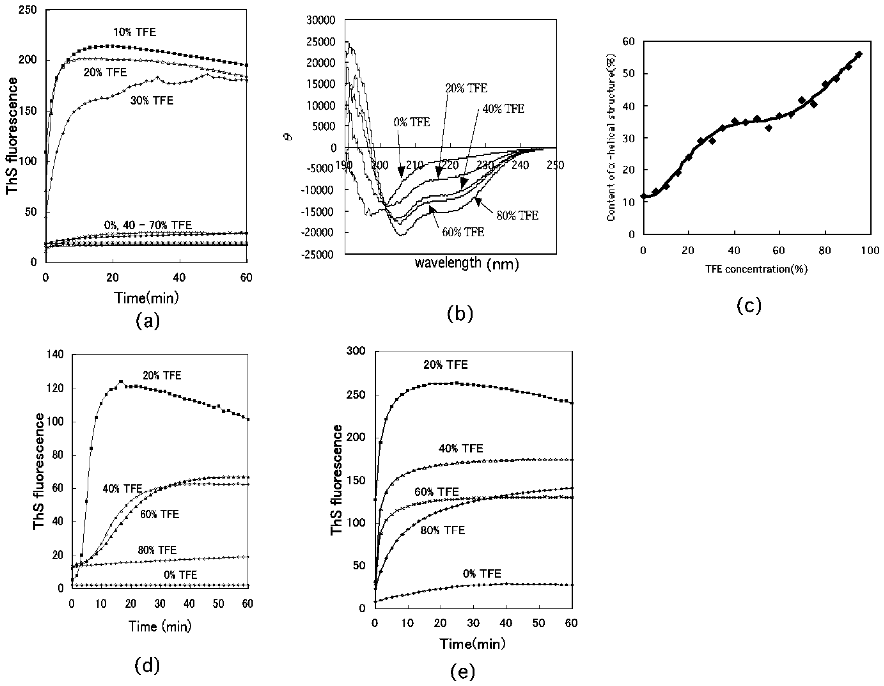


Fig. 19. (a) ThS Fluorescence–Time Profiles, (b) CD Spectra, and (c) α-Helical Content of 4RMBD, and ThS Fluorescence–Time Profiles of (d) R2 and (e) R3 Under Different TFE % in 50 mM Tris–HCl Buffer (pH 7.6)

distribution, in which the hydrophobic residues Leu10 and Val13 and the hydrophilic residues Ser11, Ser15, and Ser19 are arranged on both sides of the helix axis, and the polar residues Asn12, Gln14, Lys16, and Lys20 are located at the interface between the two sides. Similar features can also be observed in the conformation of R3, where the Cys17 residue is characteristically located at the hydrophobic side. This indicates that the disulfide bond formation of the neighboring R2–R2, R3–R3, or R2–R3 repeats or of the intramolecular R2–R3 repeats is promoted by the hydrophobic interactions. On the other hand, marked discrepancy between R2 and R3 conformations can also be observed. The N-terminal Val1–Lys6 sequence of R3 shows an extended conformation, while that of R2 assumes a typical α-helical structure; this structural feature of R3 has also been observed in aqueous solution.[62] This indicates that the extended structure of the N-terminal VQIVYK sequence in R3 is responsible for the rapid self-aggregation without any template and is in contrast with the helical structure of the N-terminal VQIINK sequence of R2, which shows very slow self-aggregation and rapid template-dependent filament formation, as stated in section 3.2.3). Thus concerning the conformation–filament formation relationship of MBD, we propose β-strand-mediated association of the N-terminal VQIVYK sequence in R3, together with the amphipathic helical structures of R2 and R3 repeats.

Concerning R1 and R4 peptides, the constructed conformers exhibit the α-helical structures of the Leu10–Lys24 sequence for R1 and the Ser5–Ser20 sequence for R4. Although both N- and C-terminal moieties are flexible and do not assume any definite 3D structure, their regions are considerably different. Great flexibility was observed in the N-terminal region of R1, which contrasts the large flexibility in the C-terminal region of R4. Both helical regions of R1 and R4 showed amphipathic residual side chain distributions similar to those of the R2 and R3 peptides.

The TFE content-dependent aggregation behavior of the R2 and R3 peptides is shown in Figs. 19d and e. These figures suggest that the starting region for 4RMBD aggregation is in either R3 or R2. Because R3 shows the most significant filament formation ability among the four repeat peptides, the possible conformational feature of R3 in an optimal TFE aqueous solution (10—20%) was estimated on the basis of the TFE content-dependent CαH proton chemical shift changes.[65] It is known that a positive deviation of Cα protons from their chemical shifts in water correlates with the structural propensity towards β-strand structure, whereas a negative deviation correlates with the α-helical structure.[67,68] Therefore we estimated the TFE content-dependent changes of CαH proton chemical shifts of R3 from those in water (Fig. 21).

Each residue of the N-terminal Gln2–Lys6 sequence exhibited a small, but nearly identical positive deviation for the chemical shift in ≧20% TFE, indicating the preference of this region for a β-strand structure. Characteristically, this deviation profile was almost independent of TFE content >20% and was dissimilar to that of the C-terminal His25–Gln31 region. In contrast, the Leu10–Cys17 sequence, except Lys16, shows a negative deviation for the chemical shift in >20% TFE. Although the deviation is not necessarily proportional to TFE content, this profile indicates the preference of this sequence for an α-helical structure. In particular, the relationship of the TFE content–chemical shift deviation shows a large slope for the Leu10–Cys17 sequence, suggesting the α-helical formation in this region at an early stage of the TFE-induced conformational transition from a random structure in water.

Generally, filament formation by the tau protein can be divided into the following four steps: activation→nucleation→extension→PHF formation. Herein we discuss the possible structural/functional role of each repeat unit in the initial *in vitro* filament formation process of 4RMBD on the basis of the results obtained thus far concerning the conformational, associational, and morphological features of the four repeat peptides (R1—R4). A possible model of 4RMBD is proposed in Fig. 22.[55,59]

As for 4RMBD, it is considered that the step of activation→nucleation is dependent on the R23 repeat moiety that constitutes the building unit for initiating the aggregation reaction. The Cys residues between the intramolecular R2 and R3 repeats form a disulfide bond in an oxidative state, resulting in a folded form. The R4 repeat at the C-terminal side of the R23 repeat acts as a regulator for aggregate formation, while the aggregation is negligibly affected by the R1 repeat on the N-terminal side. This suggests the association order of R23→R123→4RMBD to induce nucleation. Previously, we reported that 4RMBD usually exists as a noncovalent dimer in neutral buffer solution.[54] Thus it is reasonable to consider that aggregation of the folded monomer is processed *via* an intermolecular noncovalent dimer, in which the amphipathic Val1–Lys6 and Lys12–Leu20 regions of R3 provide an important driving force (hydrophobic interaction)

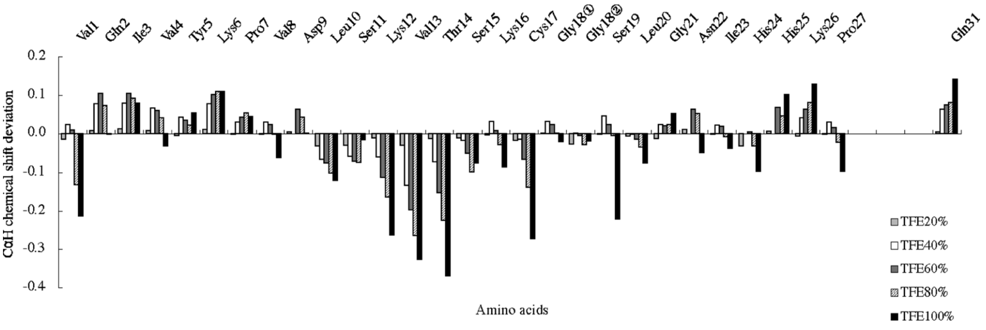


Fig. 21. TFE Content-Dependent Deviation of CαH Proton Chemical Shifts of R3 Peptide Compared with Those in Water

The bars arranged from left to right for each residue represent the differences (in ppm) in chemical shifts in 20%, 40%, 60%, 80%, and 100% TFE compared with those in water. The horizontal line corresponds to the chemical shift change of each residue from that in water. The upper and lower sides of the base line represent the downfield and upfield shifts of the proton, respectively.

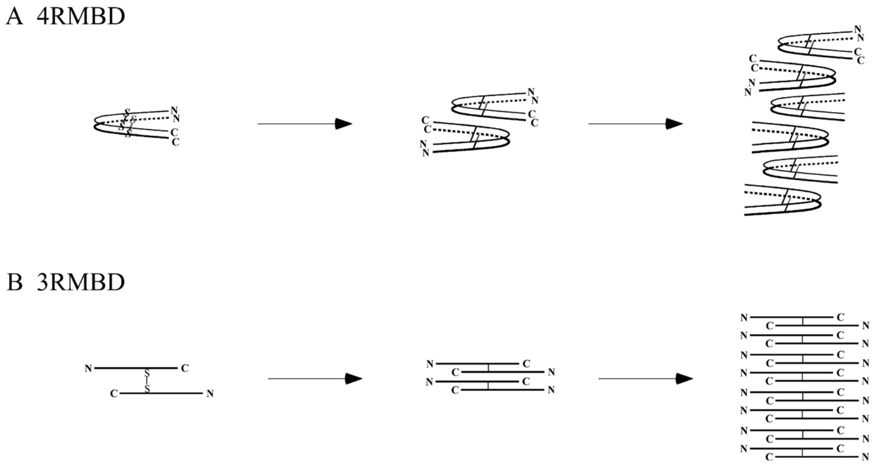


Fig. 22. Possible Schematic of Initial Aggregation of 4RMBD and 3RMBD

The lines represent the R1–R2–R3–R4 and R1–R3–R4 sequences for 4RMBD and 3RMBD, respectively. The thin lines represent intramolecular or intermolecular disulfide bonds. The N and C symbols represent the N- and C-termini of MBD, respectively.

to form the dimer structure of 4RMBD. In the reduced condition, the 4RMBD monomer takes an open form suitable for aggregation. As a possible reason why 4RMBD and 3RMBD under reducing conditions exhibit opposite aggregation behavior, it may be considered that the potent R2-dependent seeding and nucleation→extension promoting abilities of R2 are missing for R3.

On the other hand, the aggregation process of 3RMBD (Fig. 22) is initiated by an intermolecular dimer involving disulfide bonds between the neighboring two R3 cysteine residues, where the amphipathic R3 facilitates disulfide bond dimer formation through hydrophobic interaction, because the Cys residue of R3 is located at the hydrophobic side. Furthermore, the amphipathic Val1–Lys6 and Lys12–Leu20 regions of R3 stabilize the dimer structure of 3RMBD for the same reason as with 4RMBD. The functional role of R1 and R4 in the filament formation of 3RMBD is the same as that of 4RMBD.

3.6. Inhibition of MBD Aggregation Because the degree of dementia in AD patients correlates significantly with the appearance and distribution of tau tangle[69] and tau aggregation correlates with toxicity in cells,[70] it is important for the prevention and treatment of AD to find a method of inhibiting such abnormal aggregation of the tau protein. The results clarified thus far are presented in this section.[71—73]

1) Tyr Residue in R3 Repeat as Target for Inhibition of MBD Aggregation[73]: To clarify why the self-aggregation ability of the R3 repeat is much higher than that of the R2 repeat despite their similar sequences, we focused on the function of the Tyr310 residue in the Val306—Lys311 sequence for MBD filament formation, because this sequence has been identified as being related to filament formation,[53] and Tyr is the only aromatic residue in MBD. Thus we investigated the aggregation behavior of a series of Tyr-substituted, -deleted, or -inserted MBD mutants (Fig. 23) by ThS fluorescence, CD, and EM measurements.

As shown in Figs. 24a and b, both 4RMBD and 3RMBD completely lost aggregation ability as a result of Tyr→Ala substitution. This complete loss was observed under acidic and reductive conditions, although DTT significantly increases aggregation of 4RMBD (Fig. 13). This shows the importance of the Tyr residue for filament formation of MBD. The CD spectra show that the inhibition of heparin-induced filament formation of MBD results from stabilization of the native random conformation by Tyr→Ala substitution (Fig. 24d). Similarly, the EM pictures of 3RMBD and 4RMBD show the importance of the Tyr310 residue for filament formation (Fig. 25).

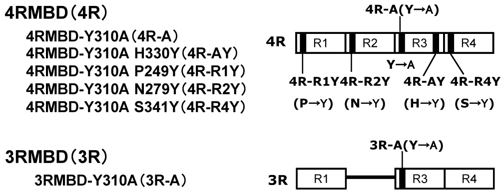


Fig. 23. Tyr310-Substituted MBD Mutants

The letters in parentheses indicate the abbreviation for the corresponding MBD mutants. The number of tau amino acid residues refers to the longest isoform of the human tau protein (441 residues).

To investigate the positional importance of Tyr residues for the MBD filament formation, four double-substituted mutants were prepared: Y310A/H330Y (4R-AY), Y310A/P249Y (4R-R1Y), Y310A/N279Y (4R-R2Y), and Y310A/S341Y (4R-R4Y); their aggregation profiles were monitored by fluorescence measurements (Fig. 24c). From these results, it is obvious that the Tyr residue located at the R3 N-terminal 310 position is essential for filament formation. The 4R-AY mutant shows no contribution of the Tyr residue positioned at the R3 C-terminal side to aggregation, although the sequence around the mutated region is similar to the N-terminal sequence. The location of the Tyr residue at the R4 N-terminal side also shows no contribution to aggregation. The CD spectra of these mutants show random conformations and no notable changes as a result of the addition of heparin. This again indicates the close relationship between the Tyr residue and filament formation of MBD. In contrast, notable self-aggregation was observed for the 4R-R2Y mutant. Because the R2 N-terminal sequence of 4R-R2Y is similar to that of the R3 repeat in 4R, this indicates that the six residues of the N-terminal of R3 participate importantly in initiating aggregation of MBD. A slight, but definite, recovery was also observed for 4R-R1Y, although R1 itself has no aggregation

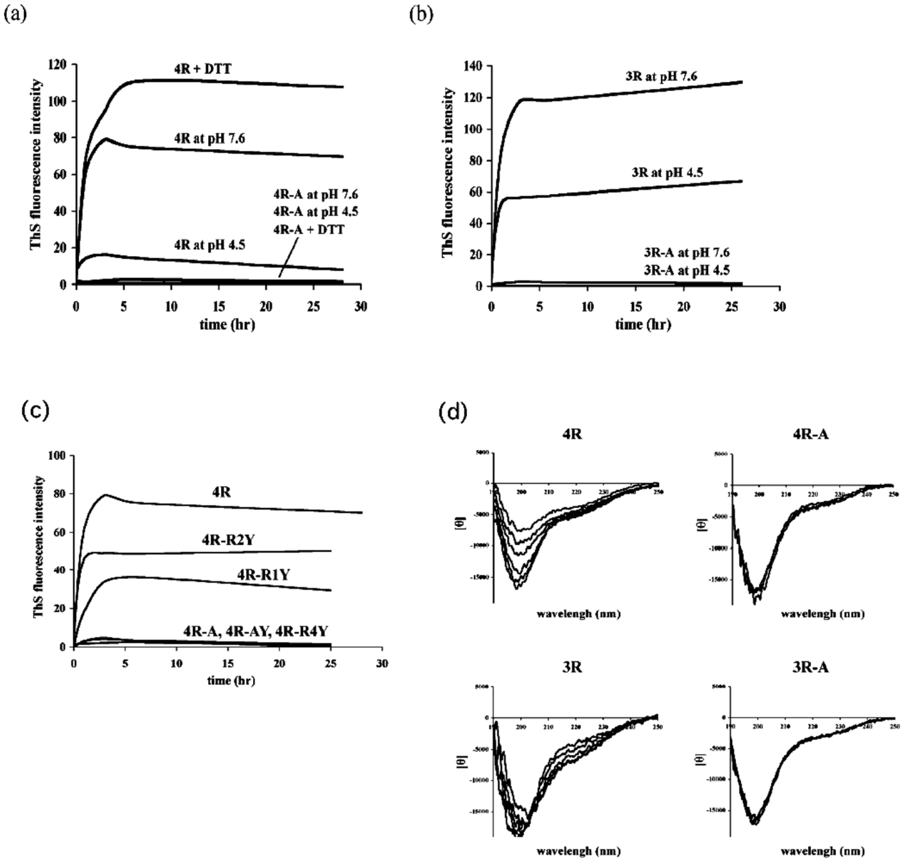


Fig. 24. ThS Fluorescence Intensity Profiles of (a) 4R and 4R-A and (b) 3R and 3R-A under Neutral, Acidic, and Reducing Conditions and of (c) 4R, 4R-A, 4R-AY, 4R-R1Y, 4R-R2Y, and 4R-R4Y under Neutral Condition, and (d) Time-Dependent CD Changes of 4R, 4R-A, 3R, and 3R-A under Neutral Condition

The intensities under acidic and reducing conditions were measured in 50 mM acetic acid buffer (pH 4.5) and in 50 mM Tris–HCl buffer (pH 7.6) containing 1 mM DTT, respectively. The respective CD spectra of 4R and 3R correspond to 0 and 20 min, and 1, 3, 6, and 23 h after the addition of heparin to the solution (horizontal axis from bottom to top at 200 nm).

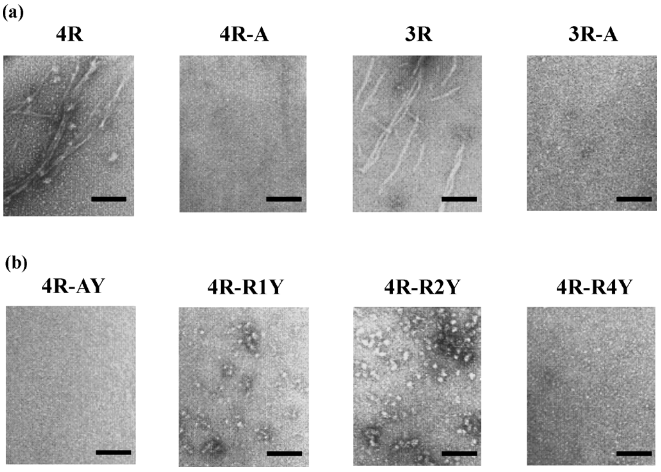


Fig. 25. Negative-Staining EM Images of (a) 4R, 4R-A, 3R, and 3R-A and (b) 4R-AY, 4R-R1Y, 4R-R2Y, and 4R-R4Y

The length of the bar corresponds to 500 nm.

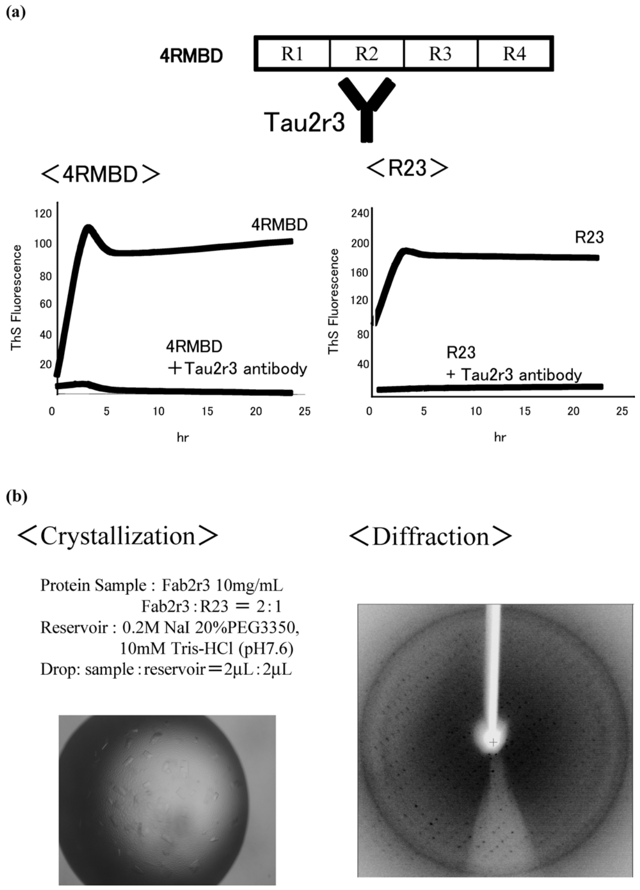
ability.

The EM images of these mutants reflect well the fluorescence and CD spectral results. No notable filaments are formed for any mutant. However, 4R-R1Y and 4R-R2Y form globular assemblies, probably the intermediate state of filament formation, although such assemblies are not formed for 4R-AY and 4R-R4Y (Fig. 25).

These results indicate that the Tyr310 residue plays a vital role in forming the extended structure of the R3 N-terminal sequence, which is indispensable for the aggregation of MBD. Therefore a compound that can disturb the function of the tyrosine residue may be expected to be a potent inhibitor of tau PHF formation.

2) MBD-Recognizing Antibody[71]: To investigate whether MBD repeat-recognizing antibody inhibits tau PHF formation, a series of monoclonal antibodies were prepared. The monoclonal antibody (Tau2r3) recognized the second repeat R2 of MBD. Its inhibition of aggregation of 4RMBD and R23 is shown in Fig. 26a; it also inhibited PHF formation of full-length tau protein. Recently, we succeeded in crystallizing a Fab2r3 (Tau2r3 variable region)–R23 complex (Fig. 26b). X-ray analysis will clarify the recognition mode of R23 by an antibody at the atomic level, which will be very useful for developing drugs for the treatment of AD.

3) Low-Molecular-Weight Inhibitor of MBD Aggregation: Cyanidine and Methylene Blue[72]: To search for useful inhibitors from foods, we investigated the inhibitory effects of a series of foodstuffs on heparin-induced filament formation of 3RMBD and 4RMBD by ThS fluorescence, CD, and EM measurements in neutral solution. To date, the following foodstuffs were shown to inhibit aggregation and promote filament disaggregation of MBD: tea, green tea, bananas,



Protein Sample : Fab2r3 10mg/mL
Fab2r3 : R23 = 2 : 1
Reservoir : 0.2M NaI 20%PEG3350,
10mM Tris-HCl (pH7.6)
Drop: sample : reservoir = 2µL : 2µL

Fig. 26. (a) Inhibition of Tau2r3 Antibody for Aggregations of 4RMBD and R23 and (b) Crystallization of Fab2r3–R23 Complex and X-Ray Diffraction Pattern

mushrooms, red wine, and perilla. Among them, it is interesting to note that red wine significantly inhibits filament formation of MBD, whereas this effect is not observed for white wine. Because this difference is probably due to the absence of polyphenols, the inhibitory behaviors of cyanidin (Cy), a representative polyphenol, and its derivatives [(−)-catechin and (−)-epicatechin] (Fig. 27) were investigated to clarify the relationship between chemical structure and inhibitory modes. Furthermore, the inhibitory effects of methylene blue (MB), thionin, and chlorpromazine were also investigated, because such π-electron-poor phenothiazine compounds appear to inhibit aggregation through π–π interaction with π-electron-rich Tyr310 residue, which is important for aggregation of MBD.

As shown in Figs. 27a and b, both Cy and MB inhibit heparin-induced filament formation of 4RMBD and 3RMBD in a concentration-dependent manner, although the IC_{50} values of the compounds (Table 4) are considerably different. They inhibit filament formation of 3RMBD more efficiently than that of 4RMBD and form sphere-like aggregates more frequently in 3RMBD than in 4RMBD; such granular tau oligomers have been considered intermediates of tau protein en route to filament formation.[74)]

To investigate the effects of Cy and MB on heparin-induced conformational transition of 4RMBD, the rates of the time-dependent CD spectral changes were measured in the absence and presence of inhibitor (Table 4). It was clarified that both Cy and MB inhibit the conformational transition of 4RMBD, owing to stabilization of MBD random structure; similar changes were also observed for 3RMBD.

To examine the structural requirements of Cy and MB for inhibiting filament formation of MBD, the inhibitory effects of structurally related compounds were measured by ThS fluorescence assay (Figs. 27a, b). Compared with Cy or MB-treated 3RMBD and 4RMBD, the inhibiting effects of these

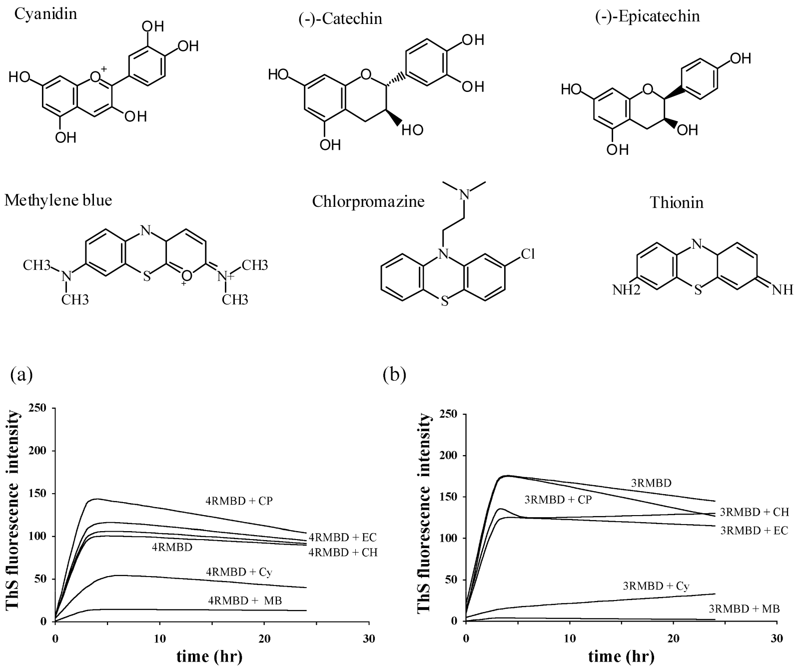


Fig. 27. Chemical Structures of Cyanidin, Methylene Blue and Their Analogues, and ThS Fluorescence Intensity Changes of (a) 4RMBD and (b) 3RMBD (25 mM) in Presence of Chlorpromazine (CP) and MB (12.5 mM), and (−)-Cathechin (CH), (−)-Epicatechin (EC), and Cy (25 mM) in 50 mM Tris–HCl Buffer (pH 7.6)

Table 4. Parameters for Defining Time-Scanned ThS Fluorescence and CD Spectral Changes of 4RMBD and 3RMBD (25 μM) by Inhibitors

Repeat peptide	IC_{50} value by ThS filament assay[a]		Conformational transition rate by CD ([$\lvert\Delta\theta\rvert$]/h)[b,c]		
Inhibitor	Cy	MB	Control	Cy (25 μM)	MB (25 μM)
4RMBD	25 μM	1 μM	420	100	110
3RMBD	12 μM	0.3 μM	700	350	160

a) Concentration producing 50% decrease in ThS fluorescence intensity. *b*) Absolute change in the value of molar ellipticity $\lvert[\theta]\rvert$ at 198 nm per unit (h). *c*) The variation is in the range of ±15%.

compounds were considerably lower, indicating their poor inhibitory abilities. Because EM measurements gave similar results, it is clear that the N-unsubstituted phenothiazine ring of MB and the planar anthocyanidin ring of Cy are important for the inhibitory activity; importance of the N-unsubstituted phenothiazine ring has also been reported.75)

Note that Cy and MB showed different responses toward the aggregations of R2 and R3, *i.e.*, inhibition by Cy but not by MB.72) Nevertheless, MB exhibited concentration-dependent inhibition of filament formations of 3RMBD and 4RMBD, indicating important contribution of the R1 and R4 repeats to the inhibition. As R1 and R4 themselves have no ability to form self-aggregates in neutral solution, their behavior of promoting inhibition of MBD filament formation is remarkable. Our results indicate supportive roles of R1 and R4 repeats in inhibition of MB for filament formation of MBD, and this finding contrasts with the aggregation-promotive role of the R2 and R3 repeats.

4. Molecular Design of Cathepsin B-Specific Inhibitor

Lysosomal cysteine proteases are abundant in living cells and play important roles in intracellular proteolysis. Cathepsin B (CB) is a major component of lysosomal enzymes and has been implicated in muscular dystrophy, bone resorption, pulmonary emphysema, and tumor metastasis. Because abnormal elevation of CB activity leads to such serious diseases, development of CB-selective low-molecular-weight inhibitors is important for the development of therapeutic drugs for these diseases and as reagents for further clarification of the physiological functions of CB.

E64 (Fig. 28) was isolated from a culture of *Aspergillus japonicus* TPR-64[76) as a potent cysteine protease inhibitor, and its derivative E64c was developed as a clinical drug, which consists of epoxysuccinyl, leucyl, and isoamylamide groups. Although E64c is a potent inhibitor of the cysteine protease family, it has no specificity for particular proteases, *i.e.*, IC_{50} (nM)=40 for papain, 13.4 for cathepsin B, 0.01 for cathepsin L, 1649 for cathepsin H, and 3000 for calpain. To design a CB-specific inhibitor from E64c by molecular alteration, we carried out X-ray crystal structure analyses of papain–E64c and CB–E64c complexes and clarified the difference between their E64c-binding pockets at the atomic level, and designed a series of CB inhibitors on the basis of these structural scaffolds. Finally, we developed CA074 [*N*-(L-3-*trans*-propylcarbamoyloxirane-2-carbonyl)-L-isoleucyl-L-proline, Fig. 28] as a CB-specific covalent-type inhibitor both *in vivo* and *in vitro*.77—94) Herein, the developmental methodology from E64c to CA074 is presented.

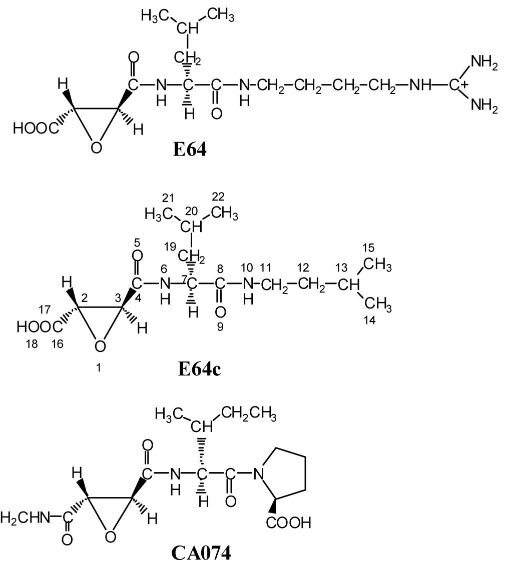


Fig. 28. Chemical Structures of E64, E64c, and CA074

The atomic numbering used in this work is given for E64c.

4.1. X-Ray Crystal Structure of CB–E64c Complex

To clarify the binding mode and inhibitory mechanism of E64c and to characterize the substrate-binding pocket of CB at the atomic level, we analyzed the crystal structure of the CB–E64c complex84,93); the crystal structure of the related CB–inhibitor complex was also reported by Turk *et al.*95) The structure was determined using phase information from two heavy-atom derivatives, where the initial multiple isomorphous replacement map showed clear electron density for the helix domains of the protein.87) Two residues, Gly48 and Arg49, were not visible in the electron density map and therefore were deleted from the refinement. Thus CB consists of a light chain (residues 1—47) and a heavy chain (residues 50—253).

CB is an $\alpha+\beta$-type protein with four α-helices and seven β-sheets (Fig. 29) and can be divided into two domains, designated as L- (residues 13—147 and 251—253) and R- (residues 1—12 and 148—250) domains. An extended polar active site cleft (substrate-binding pocket) is formed at the interface between these two domains; the Cys29 residue, the active center of CB, is positioned at the middle of this active site. The regions located at the molecular surface assumed the relatively large temperature factors of 20—65 Å^2, indicating high conformational flexibility. Among them, the region called the ‘occluding loop’ (residues 105—125) is a structural feature of CB, because it is missing in other lysosomal enzymes and has been proposed important for the peptidyldipeptidase activity of CB.96)

The overall structure of CB–E64c complex is shown in Fig. 30a. The active site pocket of CB is defined according to the concept of Schechter and Berger,97) and the *N*-(3-methylbutyl)-aminoleucylcarbonyl moiety of E64c is located at the S*n* subsite (*n*=1—3) of CB (Fig. 30b). The direct hydrogen bonds between E64c and CB are formed at pairs of O5–Gly74NH, O17–Gln23N$^{\varepsilon 2}$H, and O17–Cys29NH. The latter two hydrogen bonds strengthen the oxyanion hole of O17 (E64c), Cys29NH, and Gln23N$^{\varepsilon 2}$, which plays an important role in the catalytic reaction. The hydrogen bond of O5–Gly74NH reflects a characteristic of cysteine protease,

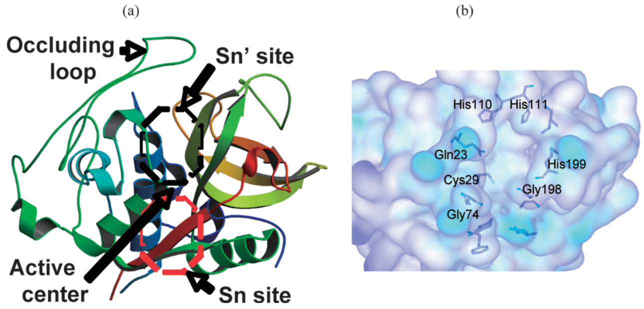


Fig. 29. (a) Overall Structure of CB and (b) Substrate-Binding Pocket

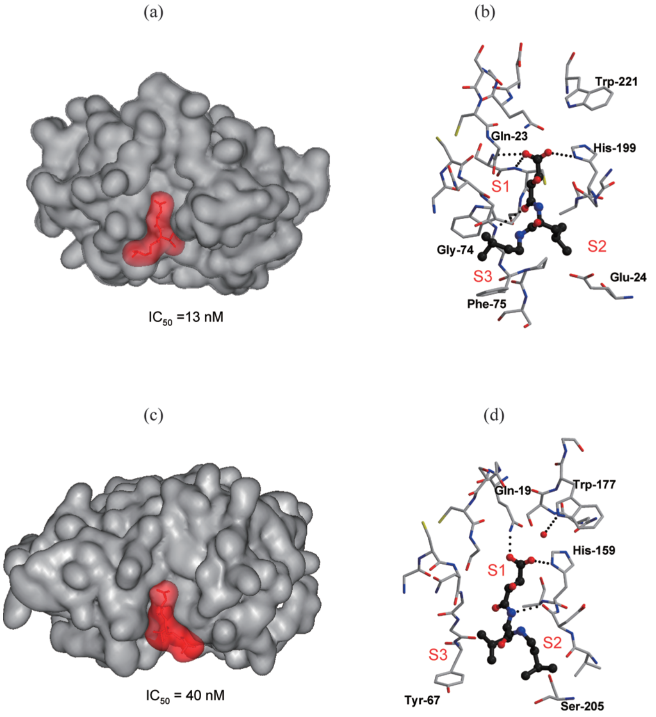


Fig. 30. Docking of E64c (Red) to Active Site Cleft of (a) CB and (c) Papain and Interaction Mode of Epoxysuccinyl, Leucyl, and Isoamylamide Groups of E64c at Substrate-Binding Pocket of (b) CB and (d) Papain

Dotted lines represent hydrogen bonds. The red dot represents a water molecule.

because Gly74 is a conserved residue in the cysteine protease family and always participates in the formation of such a hydrogen bond. The Leu residue of E64c is located at the hydrophobic S2 pocket constructed by Pro76, Ala173, Gly198, and Ala200, although Glu24 is located close by. The methylbutyl moiety of E64c is tightly locked by hydrophobic interaction with Phe75.

4.2. X-Ray Crystal Structure of Papain–E64c Complex To clarify the difference in the substrate-binding pocket of papain from that of CB at the atomic level, the crystal structure of the papain–E64c complex was analyzed by X-ray crystallography. Two polymorphous complex crystals (forms I and II) were determined.[77,82,85,86] Although these crystals differ concerning the binding mode of the E64c isoamylamide group at the substrate-binding pocket of papain, the binding patterns of the remaining groups are essentially the same.

The structure of form I is shown in Figs. 30c, d. Papain has 212 amino acid residues and is folded into two domains:

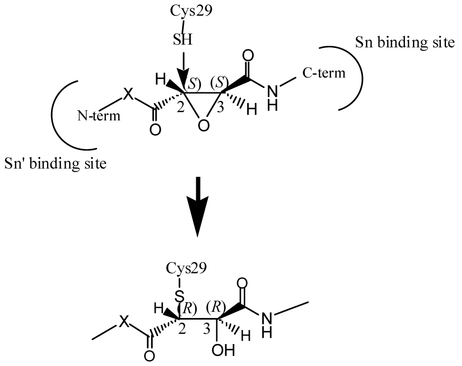


Fig. 31. Scheme of Covalent Bond Formation between *trans*-Epoxysuccinyl Fragment of E64c and Cys29 $S^{\gamma}H$ in CB

The same covalent bond is formed between papain Cys25 $S^{\gamma}H$ and the epoxysuccinyl group of E64c.

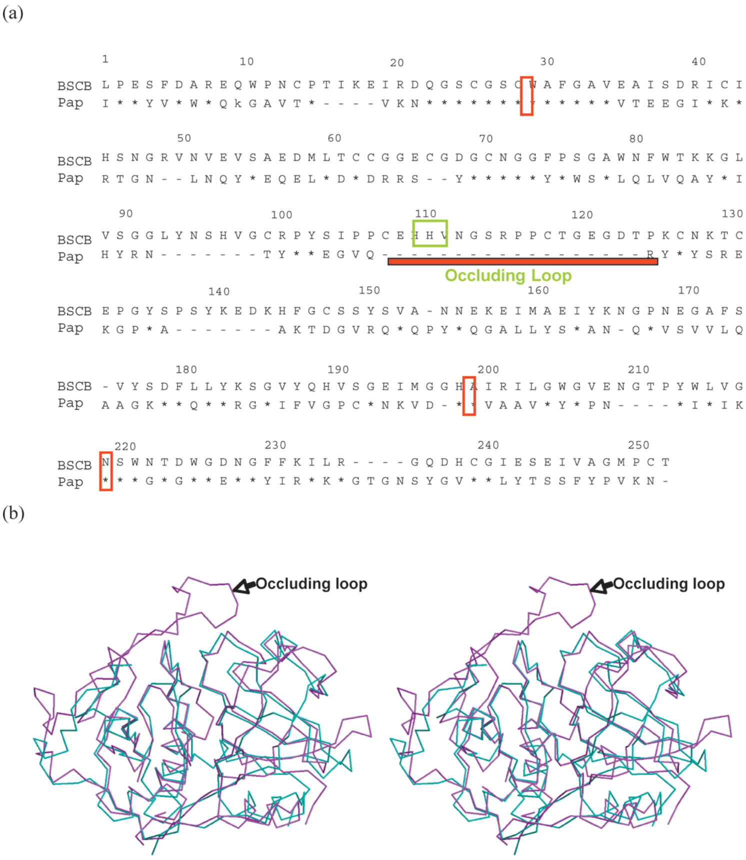
an L-domain consisting of residues 10—111 and 207—212 and an R-domain consisting of the remaining residues. Cys25, the active center of papain, is positioned at the interface between both domains. X-ray analysis of papain at 1.65 Å resolution has already been performed[98] and showed that the secondary structure of papain is negligibly affected by binding of E64c.

E64c is bound to Cys25 *via* a covalent bond between the epoxy C2 atom of E64c and the Cys25 $S^{\gamma}H$ group (discussed in section 4.3). Such covalent bond formation was also ascertained in form II. Each of the two oxygen atoms of the terminal carboxy group participates bifurcatively in hydrogen bonds or short contacts with the polar atoms of Gln19, Cys25, and His159. The hydroxy O1 atom, formed by covalent bond formation of the E64c epoxy ring with the Cys25 $S^{\gamma}H$ group, is located in the vicinity of Asp158 O (main chain), and a weak hydrogen bond is possible between them.

The leucyl side chain and isoamylamide moiety of E64c are held by hydrophobic interactions (van der Waals contacts) with the residues of Gly66 and Tyr67 and of Val157 and Asp158, respectively. These hydrophobic groups on the inhibitor are located at the entry of the hydrophobic pocket corresponding to the S_2 and S_3 subsites.

4.3. Covalent Bond Formation between E64c Epoxy Ring and Active Cysteine Residue and Binding Feature to CB S1—S3 Subsites The covalent bond formation between the E64c epoxy ring and the Cys29 $S^{\gamma}H$ of CB is shown in Fig. 31; an analogous covalent bond was formed by the Cys25 $S^{\gamma}H$ of papain. As for the covalent bond formation of E64c at the active center of CB or papain, the oxirane C2 atom is covalently bonded to the Cys29 S^{γ} atom ($C2\text{-}S^{\gamma}=1.81$ Å) in CB and to the Cys25 S^{γ} atom ($C2\text{-}S^{\gamma}=1.80$ Å) in papain, thereby leading to formation of a free OH group at the C3 atom. The configuration at the oxirane C3 atom is maintained during the σ-covalent bond formation, although this is not definitive because of the nearly planar density around the C3 atom; the chirality changes of *S* to *R* at the C2 and C3 atoms are only a matter of form, owing to the change in ranking order of the substituents.

The major binding sites of E64c are located in S*n* (*n*=1—3) subsites. These are characterized as follows: (i) the S1 subsite including the Cys29 active center corresponds to an oxyanion hole consisting of the inhibitor's carbonyl O5, Cys29N, and Gln23$N^{\varepsilon 2}$, and hydrogen bonds with polar



(a)

```
      1           10              20              30              40
BSCB  L P E S F D A R E Q W P N C P T I K E I R D Q G S C G S C W A F G A V E A I S D R I C I
Pap   I * * Y V * W * Q k G A V T * - - - - V K N * * * * * * * * * * * * * V T E E G I * K *

              50              60              70              80
BSCB  H S N G R V N V E V S A E D M L T C C G G E C G D G C N G G F P S G A W N F W T K K G L
Pap   R T G N - - L N Q Y * E Q E L * D * D R R S - - Y * * * * * Y * W S * L Q L V Q A Y * I

        90              100             110             120             130
BSCB  V S G G L Y N S H V G C R P Y S I P P C E H H V N G S R P P C T G E G D T P K C N K T C
Pap   H Y R N - - - - - - - T Y * * E G V Q - - - - - - - - - - - - - - - - - - R Y * Y S R E

                  140             150             160             170
BSCB  E P G Y S P S Y K E D K H F G C S S Y S V A - N N E K E I M A E I Y K N G P N E G A F S
Pap   K G P * A - - - - - - - A K T D G V R Q * Q P Y * Q G A L L Y S * A N - Q * V S V V L Q

              180             190             200             210
BSCB  - V Y S D F L L Y K S G V Y Q H V S G E I M G G H A I R I L G W G V E N G T P Y W L V G
Pap   A A G K * * Q * * R G * I F V G P C * N K V D - * * V A A V * Y * P N - - - - * I * I K

        220             230             240             250
BSCB  N S W N T D W G D N G F F K I L R - - - - G Q D H C G I E S E I V A G M P C T
Pap   * * * G * G * * E * * Y I R * K * G T G N S Y G V * * L Y T S S F Y P V K N -
```

Fig. 32. (a) Alignment of Amino Acid Sequences of CB and Papain, Based on Their Tertiary Structures and (b) Stereoscopic Superimposition of Cα Backbones of Tertiary Structures of CB (Red) and Papain (Blue)

BSCB in (a) represents bovine spleen cathepsin B.

atoms of the inhibitor are smoothly formed in this hole; (ii) the S2 subsite provides the hydrophobic pocket formed by Pro76, Ala173, Gly198, and Ala200 and interacts with the hydrophobic P2 site of the inhibitor; and (iii) the S3 subsite interacts with a hydrophobic P3 site through hydrophobic interactions with Tyr75 and Phe75.

4.4. Difference between Substrate-Binding Pockets of CB and Papain: Design of CB-Specific Inhibitor To clarify the difference between the substrate-binding pockets of CB and papain and to define the substrate specificity of CB, the Cα-chain foldings of CB and papain were superimposed on the basis of their crystal structures and their amino acid sequences were also aligned (Fig. 32). As the occluding loop (residues 105—125) constitutes the structural feature of CB, it may be essential to utilize this loop for designing a CB-specific inhibitor. Because the tertiary structure of the occluding loop is located at the S*n*′ subsite ($n \geqq 3$) (Fig. 29), the structural alteration from E64c is in the S*n*′-binding side chain. A strategy for the molecular design of a CB-specific E64c analog is shown in Fig. 33. As is obvious from this figure, it is necessary to construct the side chain at the C2 atom of tES (L-*trans*-epoxysuccinyl ring), which is on the opposite side of the Leu-methylbutyl chain of E64c (Fig. 31). It is characteristic that the His110 and His111 residues of the CB occluding loop are located to block the extension of the S*n*′ subsite ≧3 (Fig. 29b). Thus it was designed that the P1′ and P2′ side chains correspond to the S1′ and S2′ positions of CB, where it is important that the P2′ side chain is able to interact with the imizadole N atom and/or NH group of His110/111.

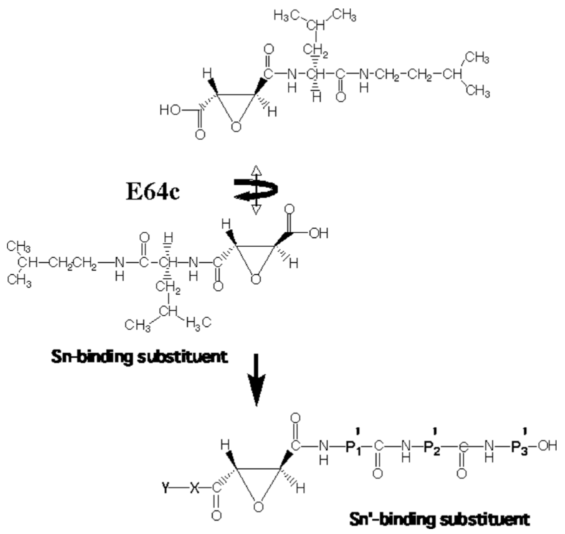


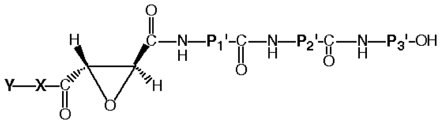
Fig. 33. Strategy for Molecular Design of CB-Specific E64c Analogue

4.5. Molecular Design of CA074 The structural modifications of P1′—P3′ subsites are summarized in Table 5.[83] The design began with a survey of the residues that could be best matched at the pocket in the CB S1′ subsite. Because this pocket has a hydrophobic environment (Leu181, Val176, Ileu195, Met196, Gly198, and His199), five amino acids

Table 5. Structure–Inhibitory Activity Relationship of E64c Analogues at (a) S*n*′ and (b) S*n* Subsites

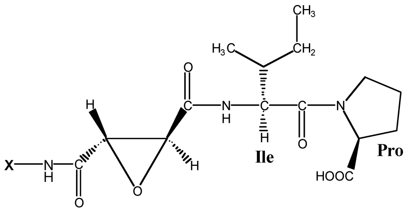
(a)

Compound			Cathepsin B (nM)	Papain (nM)
P1′	P2′	P3′		
Ile			24000	132000
Gly	Pro		15300	95400
Pro	Pro		50	>100000
Ile	Ala		23	5800
Ile	Pro		3	16000
Ile	Pro	OBzl	930	2000



(b)

Compound X	Cathepsin B (nM)	Papain (nM)
CH_3CH_2	7	53000
$(CH_3)_2CH$	5	55000
$CH_3CH_2CH_2$ CA074	2	57400
$(CH_3)_2CH_2CH_2$	2	60000
Cyclo-(C_6H_{12})	2	14200



(Gly, Ser, Thr, Ile, and Trp) were considered as P1′ and their inhibitory activities were measured. Although a significant preference was not observed for their inhibitory activities, Ile was suggested as the best match by the energy-minimized MD calculation. Thus Ile was selected as P1′ (Fig. 34a). The S2′ subsite of CB is constructed with the hydrophobic Trp221, Gly24, Gly121, and His199 residues, and two imidazole rings of His110 and His111 in the CB occluding loop are close to the S2′ subsite. Thus it appears important to design the P2′ residue, the side and C-terminal groups of which are hydrophobic and imidazole N and/or NH-interactive polar groups, respectively. Taking this into consideration, several derivatives were synthesized and their inhibitory activities were measured. The energy-minimized MD calculations showed a linear relationship between the binding fitness and inhibitory activity of the inhibitors that form the hydrogen bonds between the C-terminal COOH of the inhibitor and the imidazole N and NH of His110 and His111.

Concerning the substrate specificity of CB S*n* subsites, the X-ray crystal structure of the CB–E64c complex (Figs. 30a, b) shows that the molecular surfaces of S*n* subsites have a wider groove than those of the S*n*′ subsite and no notable specificity is formed for the CB–E64c interaction, except for blockage of the substrate to access to the catalytic site. This suggests that inhibitory activity of CB is primarily determined by the S*n*′–P*n*′ interaction. Thus the *n*-propyl-NH group was selected as the side chain of the P*n* subsite because of the stability of the structure and absorption in living cells. Thus CA074 (*N*-(L-3-*trans*-propylcarbamoyl-oxirane-2-carbonyl)-L-isoleucyl-L-proline) was designed as a CB-specific inhibitor *in vitro* and *in vivo*: IC_{50} (nM)$=1.8$ for CB, cathepsin L$=1.72\times10^5$, cathepsin H$=4.20\times10^5$, and calpain II$=>2.0\times10^5$.

4.6. X-Ray Crystal Structure of CB–CA074 Complex To clarify the inhibitory mechanism of CA074 at the atomic level, the crystal structure of the CB–CA074 complex was determined by X-ray diffraction.[84,87,92)] The electron density map of CA074 at the CB active site is shown in Fig. 35. The L-isoleucyl-L-proline and propylcarbamoyl moieties of CA074 are situated at the S′ and S subsites, respectively. This binding mode is almost identical to what we assumed in the molecular design of CA074. It is obvious that the L-proline, L-isoleucyl, carbamoyl, and propyl moieties of CA074 correspond to the P2′, P1′, P1, and P2 substituents, respectively. As the main reason why the L-isoleucyl-L-proline moiety of CA074 prefers to bind to the S′ subsites, although CB has sufficient space to accommodate this moiety at S subsites, the following may be proposed: (i) the same *S* configuration (Fig. 31) of the oxirane C2 and C3 atoms makes the collinear arrangement between the L-isoleucyl-L-proline and propylcarbamoyl moieties possible, which is suitable for the tight binding of the inhibitor to the CB binding cleft[91)]; and (ii) double hydrogen bonds are possible between the C-terminal Pro carboxyl Os and the imidazole NH and N of His-110 and His-111 residues. On the other hand, the propylcarbamoyl moiety of CA074 is located at the S1 and S2 subsites. The carbonyl O5 is hydrogen-bonded to Gly74 NH, and the amide NH is within the range of a hydrogen bond with Gly198 O [=4.02 Å]. The electron density map shows covalent bond formation between the oxirane C3 atom of CA074 and the Cys29 S^γ atom (C3-S^γ=1.79 Å). Concerning the covalent bond formation, on the other hand, the nucleophilic attack of the Cys S^γ atom on the oxirane carbon atom appears to occur according to a universal rule that the Cys S^γ atom attacks the oxirane carbon atom possessing the P*n*′ substituents (C3 for CA074–CB and C2 for E64c–CB and –papain).[93)]

4.7. Quantitative Structure–Inhibitory Activity Relationship of CB Catalytic Subsite by L-*trans*-Epoxysuccinyl Inhibitors[94)] To estimate quantitatively the inhibitory effect of each substrate-binding subsite of CB, a series of L-*trans*-epoxysuccinyl derivatives with different functional groups bound to both carbon atoms of the epoxy ring was synthesized, and the relationships between their inhibitory activities and binding modes at CB subsites were evaluated on the basis of the X-ray crystal structure analyses of eight complexes. Under the common reaction in which the epoxy ring of an inhibitor is opened to form a covalent bond with the S^γH group of the active center Cys29, the observed binding modes of the substituents of inhibitors at the binding subsites of CB enabled quantitative assessment of the inhibitory effect of each subsite. Although single blockage of the S1′ or S2′ subsite exerts only the inhibitory effect of IC_{50}=*ca.* 24 μM (k_2=*ca.* 1250 $\mathrm{M}^{-1}\cdot\mathrm{s}^{-1}$) or *ca.* 15 μM (k_2=*ca.* 1800 $\mathrm{M}^{-1}\cdot\mathrm{s}^{-1}$),[94)] respectively, synchronous block of both subsites leads to IC_{50}=*ca.* 23 nM (k_2=153000—185000 $\mathrm{M}^{-1}\cdot\mathrm{s}^{-1}$) (Fig. 34) under the condition that (i) the inhibitor possesses a P1′= hydrophobic residue such as Ile and a P2′=hydrophobic residue such as Ala, Ile, or Pro; and (ii) the C-terminal

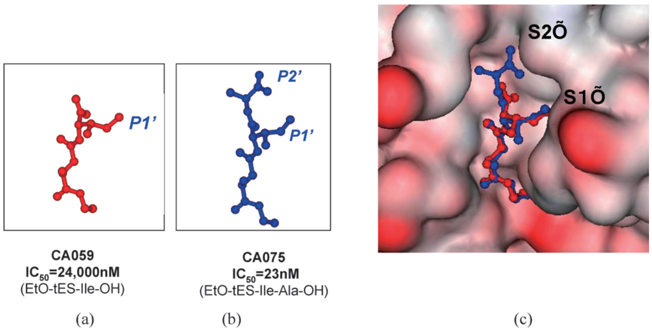


Fig. 34. Structure–Inhibitory Activity Relationship of (a) S1′ and (b) S1′ and S2′ Subsites, and (c) Their X-Ray-Analyzed Docking Modes at CB Active Site

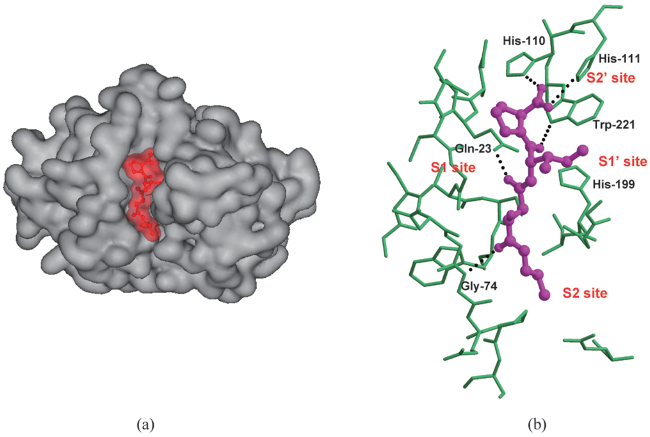


Fig. 35. (a) Docking of CA074 (Red) to Active Site Cleft of CB and (b) Interaction Mode of CA074 (Purple) at CB Binding Pocket (Green)

Dotted lines represent hydrogen bonds.

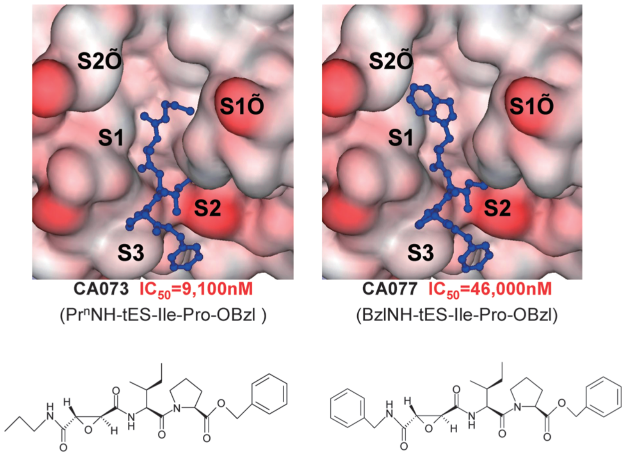


Fig. 36. Structure–Inhibitory Activity Relationship of E64c Analogues of Pn' Side Chain of $n' \geqq 3$

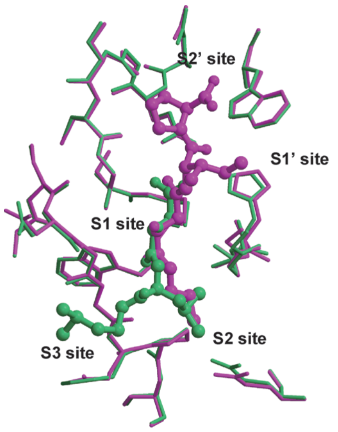


Fig. 37. Superimposition of Binding Modes of Substituents in CA074 (Purple) and E64c (Green) on CB Active Site

carboxyl group of the P2′ residue is able to form paired hydrogen bonds with the imidazole NH of His110 and the imidazole N of His111 of CB (Fig. 35). The inhibitor with Pn' ($n \geqq 3$) substituent decreases its activity by collision with the occluding loop (Fig. 36).

4.8. Noncovalent Inhibitor Based on Inhibition Modes of CA074 and E64c[93] Comparison of the crystal structures of CB–CA074 and CB–E64c complexes suggests the design of noncovalent CB-specific inhibitors by adopting the continuous P2′–P1′–P1–P2–P3 sequence of both inhibitors, because their inhibition sites are different, *i.e.*, S1′ and S2′ subsites for CA074 and S1—S3 subsites for E64c (Fig. 37). Thus a series of residues corresponding to Pn' ($n=1$, 2) and Pn ($n=1$—3) was designed by replacing the L-*trans*-epoxysuccinyl group of E64c with –CH(OH)–CH$_2$–, –CH$_2$CH$_2$– or –C(=O)–CH(OH)– groups (Fig. 38), and their docking simulations to the CB active site were performed by MD calculation. As a compound that satisfied the stated conditions, CAA0445 was selected. The inhibitory activity of CAA0445 against CB showed an initial IC$_{50}$ value of $<10\,\mu$M, and it was not sufficiently strong to be a CB-specific inhibitor. However, this is a lead compound for further development of the noncovalent inhibitor (Fig. 39).

5. Summary

Information on the structure-based mechanism of biomolecular recognition or self-aggregation at the atomic level is important to understand the behavior and develop a drug for controlling the abnormal function of biologically essential molecules. Under the theme of "Structural Studies of Specific Intermolecular Interactions and Self-aggregation of Biomolecules and their Application to Drug Design", three topics were reviewed. First, the structural scaffold for the selective recognition mechanism of the mRNA cap structure by eIF4E and the regulation mechanism of eIF4E by 4EBP were clarified at the atomic level. Next, the structural and conformational contributions of MBD repeat peptides to aggregation of tau protein were clarified and some approaches to inhibit the aggregation proposed. Finally, a CB-selective covalent inhibitor CA074 was developed from the cysteine protease inhibitor E64c using structure-based drug design and a noncovalent CB inhibitor was proposed as a lead compound. These results contribute to the application of the physicochemical approaches to target-directed drug design.

Acknowledgements I would like to express my sincere appreciation to numerous coworkers whose excellent contributions made the performance of this research possible; the names of these coworkers are provided in the references. In particular, I am greatly indebted to Dr. Koji Tomoo and Dr. Yasuko In (Osaka University of Pharmaceutical Sciences) for the patient performance of all studies, Professor Emeritus Seiichi Uesugi (Yokohama National University), Professor Mitsunobu Doi (Osaka University of Pharma-

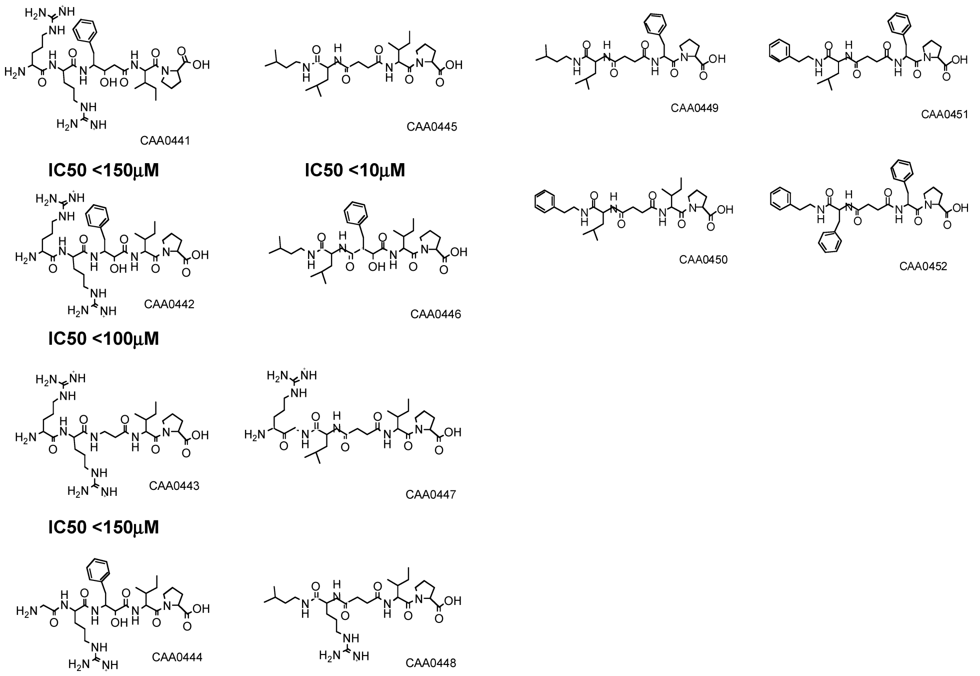
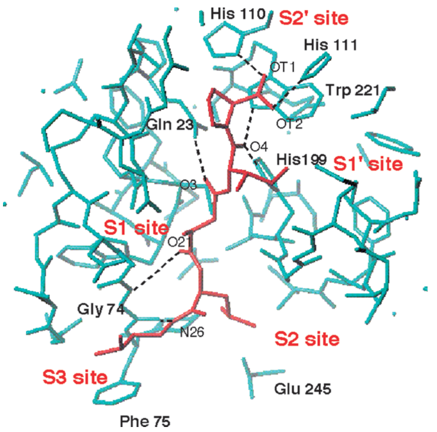


Fig. 38. Chemical Structures of Designed Noncovalent Inhibitors



Possible interaction between BSCB and CAA0445

BSCB	CAA0445	distance (Å)
Hydrogen bonding interactions		
His110 $N^{\delta1}H$	OT1	2.62
His111 $N^{\delta1}H$	OT2	3.21
Trp221 $N^{\varepsilon1}H$	O4	2.95
His199 $N^{\delta2}H$	O4	3.23
Gln23 $N^{\varepsilon2}H$	O3	2.80
Gly74 NH	O21	3.15
Gly74 O	N26	2.79
Hydrophobic interactions		
Phe75 $C^{\varepsilon2}H$	C27	4.22

Fig. 39. A Snapshot (at 85 ps) of Binding Mode of CAA0445 (Red) to S and S′ Subsites in BSCB (Blue)

Dotted lines represent hydrogen bonds.

ceutical Sciences), Dr. Hitoshi Ueda (Bayer Yakuhin, Ltd.) and Dr. Shigenobu Morino (Astellas Pharma Inc.) for the structural studies of mRNA cap structure recognition by eukaryotic initiation factor 4E, Professor Taizo Taniguchi (Himeji Dokkyou Univ.) and Dr. Miho Sumida (Behavioral and Medical Sciences Research Consortium) for studies of the self-aggregation mechanism of tau protein, and the Taisho Pharmaceutical Company group (Dr. K. Kitamura, Dr. H. Miyagawa, Dr. M. Murata, Dr. S. Sumiya, and Dr. K. Matsumoto), Associate Professor Daisuke Yamamoto (Osaka Medical College), and Dr. Atsushi Yamamoto (Mitsubishi Tanabe Pharma Corp.) for molecular design of the cathepsin B-specific inhibitor.

References

1) Sonenberg N., Hinnebusch A. G., *Cell*, **136**, 731—745 (2009).
2) Miura K., *Adv. Biophys.*, **14**, 205—238 (1981).
3) Ishida T., Tomita K., Inoue M., *Arch. Biochem. Biophys.*, **200**, 492—502 (1980).
4) Inoue M., Shibata M., Kondo Y., Ishida T., *Biochemistry*, **20**, 2936—2945 (1981).
5) Ishida T., *J. Cryst. Soc. Jpn.*, **25**, 157—167 (1983).
6) Ishida T., Matsui M., Inoue M., Hirano H., Yamashita M., Sugiyama K., Sugiura M., Tomita K., *J. Am. Chem. Soc.*, **107**, 3305—3314 (1985).
7) Ishida T., *Yakugaku Zasshi*, **108**, 506—527 (1988).
8) Ishida T., Shibata M., Fujii K., Inoue M., *Biochemistry*, **22**, 3571—3581 (1983).
9) Ishida T., Katsuta M., Inoue M., Yamagata Y., Tomita K., *Biochem. Biophys. Res. Commun.*, **115**, 849—854 (1983).
10) Kamiichi K., Danshita M., Minamino N., Doi M., Ishida T., Inoue M., *FEBS Lett.*, **195**, 57—60 (1986).
11) Kamiichi K., Doi M., Nabae M., Ishida T., Inoue M., *J. Chem. Soc. Perkin Trans. 2*, **1987**, 1739—1745 (1987).
12) Ishida T., Doi M., Ueda H., Inoue M., Scheldrick G. M., *J. Am. Chem. Soc.*, **110**, 2286—2294 (1988).
13) Ishida T., Doi M., Inoue M., *Nucleic Acid Res.*, **16**, 6175—6190 (1988).
14) Ishida T., Ueda H., Segawa K., Doi M., Inoue M., *Arch. Biochem. Biophys.*, **278**, 217—227 (1990).
15) Tarui M., Furumura H., Kafuku Y., Ishida T., Inoue M., *Biochem. Biophys. Res. Commun.*, **183**, 577—583 (1992).
16) Ishida T., *The Research Foundation for Pharmaceutical Sciences*, **1992**, 55—82 (1992).
17) Ishida T., Tarui M., In Y., Ogiyama M., Doi M., Inoue M., *FEBS Lett.*, **333**, 214—216 (1993).
18) Kawai H., Tarui M., Doi M., Ishida T., *FEBS Lett.*, **370**, 193—196 (1995).
19) Tarui M., Nomoto N., Hasegawa Y., Minoura K., Doi M., Ishida T.,

Chem. Pharm. Bull., **44**, 1998—2002 (1996).

20) Ishida T., Iyo H., Ueda H., Doi M., Inoue M., *J. Chem. Soc., Chem. Commun.*, **1990**, 217—218 (1990).
21) Ishida T., Iyo H., Ueda H., Doi M., Inoue M., Nishimura S., Kitamura S., *J. Chem. Soc., Perkin Trans. 1*, **1991**, 1847—1853 (1991).
22) Ueda H., Doi M., Inoue M., Ishida T., Tanaka T., Uesugi S., *Biochem. Biophys. Res. Commun.*, **154**, 199—204 (1988).
23) Ueda H., Iyo H., Doi M., Inoue M., Ishida T., Morioka H., Tanaka T., Nishikawa S., Uesugi S., *FEBS Lett.*, **280**, 207—210 (1991).
24) Ueda H., Iyo H., Doi M., Inoue M., Ishida M., *Biochim. Biophys. Acta*, **1075**, 181—186 (1991).
25) Kafuku Y., Matsui Y., Ohtani J., Usami Y., Ueda H., Doi M., Inoue M., Ishida T., *Chem. Pharm. Bull.*, **39**, 2487—2490 (1991).
26) Kafuku Y., Ohtani J., Doi M., Inoue M., Ishida T., *Chem. Pharm. Bull.*, **41**, 231—234 (1993).
27) Ishida T., Toda Y., Tarui M., Doi M., Inoue M., *Chem. Pharm. Bull.*, **42**, 674—676 (1994).
28) Ueda H., Maruyama H., Doi M., Inoue M., Ishida T., Morioka H., Tanaka T., Nishikawa S., Uesugi S., *J. Biochem.*, **109**, 882—889 (1991).
29) Morino S., Yasui M., Doi M., Inoue M., Ishida T., Ueda H., Uesugi S., *J. Biochem.*, **116**, 687—693 (1994).
30) Morino S., Teraoka Y., Doi M., Ishida T., Ueda H., Uesugi S., *Biol. Chem. Pharm.*, **18**, 372—376 (1995).
31) Ishida T., *Yakugaku Zasshi*, **115**, 401—419 (1995).
32) Morino S., Hazama H., Ozaki M., Teraoka Y., Shibata S., Doi M., Ueda H., Ishida T., Uesugi S., *Eur. J. Biochem.*, **239**, 597—601 (1996).
33) Teraoka Y., Morino S., Tomoo K., Ishida T., *Biochem. Biophys. Res. Commun.*, **228**, 704—708 (1996).
34) Shen X., Tomoo K., Uchiyama S., Kobayashi Y., Ishida T., *Chem. Pharm. Bull.*, **49**, 1299—1303 (2001).
35) Morino S., Tomoo K., Nishi N., Okabe K., Doi M., Ishida T., Kitamura K., *J. Biochem.*, **119**, 224—225 (1996).
36) Marcotrigiano J., Gingras A. C., Sonenberg N., Burley S. K., *Cell*, **89**, 951—961 (1997).
37) Tomoo K., Shen X., Okabe K., Nozoe Y., Fukuhara S., Morino S., Ishida T., Taniguchi T., Hasegawa H., Terashima A., Sasaki M., Katsuya Y., Kitamura K., Miyashi H., Ishikawa M., Miura K., *Biochem. J.*, **362**, 539—544 (2002).
38) Tomoo K., Shen X., Okabe K., Nozoe Y., Fukuhara S., Morino S., Sasaki M., Taniguchi T., Miyagawa H., Kitamura K., Miyashi H., Miura K., Ishida T., *J. Mol. Biol.*, **328**, 365—383 (2003).
39) Shibata S., Morino S., Tomoo K., In Y., Ishida T., *Biochem. Biophys. Res. Commun.*, **247**, 213—216 (1998).
40) Sonenberg N., *Curr. Opin. Struct. Biol.*, **13**, 56—63 (2003).
41) Tomoo K., Matsushita Y., Fujisaki H., Abiko F., Shen X., Taniguchi T., Miyagawa H., Kitamura K., Miura K., Ishida T., *Biochim. Biophys. Acta*, **1753**, 191—208 (2005).
42) Marcotrigiano J., Gingras A. C., Sonenberg N., Burley S. K., *Mol. Cell*, **3**, 707—716 (1999).
43) Mader S., Lee H., Pause A., Sonenberg N., *Mol. Cell. Biol.*, **15**, 4990—4997 (1995).
44) Gross J. D., Moerke N. J., Von der Haar T., Lugovskoy A. A., Sachs A. B., McCarthy J. E. G., Wagner G., *Cell*, **115**, 739—750 (2003).
45) Tee A. R., Proud C. G., *Mol. Cell. Biol.*, **22**, 1674—1683 (2002).
46) Abiko F., Tomoo K., Mizuno A., Morino S., Imataka H., Ishida T., *Biochem. Biophys. Res. Commun.*, **355**, 667—672 (2007).
47) Mizuno A., In Y., Fujita Y., Abiko F., Miyagawa H., Kitamura K., Tomoo K., Ishida T., *FEBS Lett.*, **582**, 3439—3444 (2008).
48) Lee V. M., Goedert M., Trojanowski L. Q., *Annu. Rev. Neurosci.*, **24**, 1121—1159 (2001).
49) Goedert M., Spillantini M. G., Jakes R., Rutherford D., Crowther R. A., *Neuron*, **3**, 519—526 (1989).
50) Goedert M., Spillantini M. G., *Biochim. Biophys. Acta*, **1502**, 110—121 (2000).
51) Friedhoff F., Von Bergen M., Mandelkow E. M., Mandelkow E., *Proc. Natl. Acad. Sci. U.S.A.*, **95**, 15712—15717 (1998).
52) Goedert M., Jakes R., Spillantini M. G., Hasegawa M., Smith M. J., Crowther R. A., *Nature* (London), **383**, 550—553 (1996).
53) Friedhoff P., Schneider A., Mandelkow E. M., Mandelkow M., *Biochemistry*, **37**, 10223—10230 (1998).
54) Yao T. M., Tomoo K., Ishida T., Hasegawa H., Sasaki M., Taniguchi T., *J. Biochem.*, **134**, 91—99 (2003).
55) Tomoo K., Yao T. M., Minoura K., Hiraoka S., Sumida M., Taniguchi T., Ishida T., *J. Biochem.*, **138**, 413—423 (2005).
56) Mizushima F., Minoura K., Tomoo K., Sumida M., Taniguchi T., Ishida T., *Biochem. Biophys. Res. Commun.*, **343**, 712—718 (2006).
57) Mizushima F., Minoura K., Tomoo K., Sumida M., Taniguchi T., Ishida T., *J. Biochem.*, **142**, 49—54 (2007).
58) Okuyama K., Nishiura C., Mizushima F., Minoura K., Sumida M., Taniguchi T., Tomoo K., Ishida T., *FEBS J.*, **275**, 1529—1539 (2008).
59) Sugino E., Nishiura C., Minoura K., In Y., Sumida M., Taniguchi T., Tomoo K., Ishida T., *Biochem. Biophys. Res. Commun.*, **385**, 236—2400 (2009).
60) Sreerama N., Venyaminov S. Y., Woody R. W., *Anal. Biochem.*, **299**, 271—274 (2001).
61) Minoura K., Tomoo K., Ishida T., Hasegawa H., Saasaki M., Taniguchi T., *Biochem. Biophys. Res. Commun.*, **294**, 210—214 (2002).
62) Minoura K., Tomoo K., Ishida T., Hasegawa H., Saasaki M., Taniguchi T., *Bull. Chem. Soc. Jpn.*, **76**, 1617—1624 (2003).
63) Minoura K., Yao T. M., Tomoo K., Sumida M., Sasaki M., Taniguchi T., Ishida T., *Eur. J. Biochem.*, **271**, 545—552 (2004).
64) Hiraoka S., Yao T. M., Minoura K., Tomoo K., Sumida M., Taniguchi T., Ishida T., *Biochem. Biophys. Res. Commun.*, **315**, 659—663 (2004).
65) Minoura K., Mizushima F., Tokimasa M., Hiraoka S., Tomoo K., Sumida M., Taniguchi T., Ishida T., *Biochem. Biophys. Res. Commun.*, **327**, 1100—1104 (2005).
66) Tokimasa M., Minoura K., Hiraoka S., Tomoo K., Sumida M., Taniguchi T., Ishida T., *FEBS Lett.*, **579**, 3481—3486 (2005).
67) Wishart D. S., Sykes B. D., Richards F. M., *J. Mol. Biol.*, **222**, 311—333 (1991).
68) Spera S., Bax S., *J. Am. Chem. Soc.*, **113**, 5490—5492 (1991).
69) Braak H., Braak E., *Acta Neuropathol.*, **82**, 239—259 (1991).
70) Bandyopadhyay B., Li G., Yin H., Kuret J., *J. Biol. Chem.*, **282**, 16454—16464 (2007).
71) Taniguchi T., Sumida M., Hiraoka S., Tomoo K., Kakehi T., Minoura K., Sugiyama S., Inaka K., Ishida T., Saito N., Tanaka C., *FEBS Lett.*, **579**, 1399—1404 (2005).
72) Hattori M., Sugino E., Minoura K., In Y., Sumida M., Taniguchi T., Tomoo K., Ishida T., *Biochem. Biophys. Res. Commun.*, **374**, 158—163 (2008).
73) Nishiura C., Takeuchi K., Minoura K., Sumida M., Taniguchi T., Tomoo K., Ishida T., *J. Biochem.*, in press.
74) Maeda S., Sahara N., Saito Y., Murayama M., Yoshiike Y., Kim H., Miyasaka T., Murayama S., Ikai A., Takashima A., *Biochemistry*, **46**, 3851—3861 (2007).
75) Taniguchi S., Suzuki N., Masuda M., Hisanaga S., Iwatsubo T., Goedert M., Hasegawa M., *J. Biol. Chem.*, **280**, 7614—7623 (2005).
76) Hanada K., Tamai M., Yamagishi S., Ohmura S., Sawada J., Tanaka I., *Argric. Biol. Chem.*, **42**, 523—528 (1978).
77) Matsumoto K., Yamamoto D., Ohishi H., Tomoo K., Ishida T., Inoue M., Sadatome T., Kitamura K., Mizuno H., *FEBS Lett.*, **245**, 177—180 (1989).
78) Yamamoto D., Matsumoto K., Ishida T., Inoue M., Sumiya S., Kitamura K., *Chem. Pharm. Bull.*, **37**, 2577—2581 (1989).
79) Yamamoto D., Matsumoto K., Ohishi H., Ishida T., Inoue M., Kitamura K., Hanada K., *FEBS Lett.*, **263**, 134—136 (1990).
80) Yamamoto D., Ohishi H., Ishida T., Inoue M., Sumiya S., Kitamura K., *Chem. Pharm. Bull.*, **38**, 2339—2343 (1990).
81) Yamamoto D., Ishida T., Inoue M., *Biochem. Biophys. Res. Commun.*, **171**, 711—716 (1990).
82) Yamamoto D., Matsumoto K., Ohishi H., Ishida T., Inoue M., Kitamura K., Mizuno H., *J. Biol. Chem.*, **266**, 14771—14777 (1991).
83) Sumiya S., Yoneda T., Kitamura K., Murata M., Yokoo C., Tamai M., Yamamoto A., Inoue M., Ishida T., *Chem. Pharm. Bull.*, **40**, 299—303 (1992).
84) Yamamoto A., Kaji T., Tomoo K., Ishida T., Inoue M., Murata M., Kitamura K., *J. Mol. Biol.*, **227**, 942—944 (1992).
85) Kim M. J., Yamamoto D., Matsumoto K., Inoue M., Ishida T., Mizuno H., Sumiya S., Kitamura K., *Biochem. J.*, **287**, 797—803 (1992).
86) Matsumoto K., Murata M., Sumiya S., Kitamura K., Ishiuda T., *Biochim. Biophys. Acta*, **1208**, 268—276 (1994).
87) Yamamoto A., Hara T., Tomoo K., Ishida T., Fujii T., Hata Y., Murata M., Kitamura K., *J. Biochem.*, **121**, 974—977 (1997).
88) Matsumoto K., Murata M., Sumiya S., Mizoue K., Kitamura K., Ishida T., *Biochim. Biophys. Acta*, **1383**, 93—100 (1998).
89) Yamamoto A., Ishida T., *Biomed. Therapeu.*, **32**, 985—993 (1998).
90) Matsumoto K., Mizoue K., Kitamura K., Tse W. C., Huber C. P.,

Ishida T., *Biopolymers*, **51**, 99—107 (1999).

91) Yamamoto A., Tomoo K., Miyagawa H., Takaoka Y., Sumiya S., Kitamura K., Ishida T., *Chem. Pharm. Bull.*, **48**, 480—485 (2000).

92) Yamamoto A., Tomoo K., Hara T., Murata M., Kitamura K., Ishida T., *J. Biochem.*, **127**, 635—643 (2000).

93) Yamamoto A., Tomoo K., Matsugi K., Hara T., In Y., Murata M., Kitamura K., Ishida T., *Biochim. Biophys. Acta*, **1597**, 244—251 (2002).

94) Watanabe D., Yamamoto A., Tomoo K., Matsumoto K., Murata M., Kitamura K., Ishida T., *J. Mol. Biol.*, **362**, 979—993 (2006).

95) Turk D., Podobnik M., Popovic T., Katunuma N., Bode W., Huber R., Turk V., *Biochemistry*, **34**, 4791—4797 (1995).

96) Illy C., Quraishi O., Wang J., Purisima E., Vernet T., Mort J. S., *J. Biol. Chem.*, **272**, 1197—1202 (1997).

97) Schechter I., Berger A., *Biochem. Biophys. Res. Commun.*, **27**, 157—162 (1967).

98) Kamphuis I. G., Kalk K. H., Swarte M. B. A., Drenth J., *J. Mol. Biol.*, **179**, 233—256 (1984).