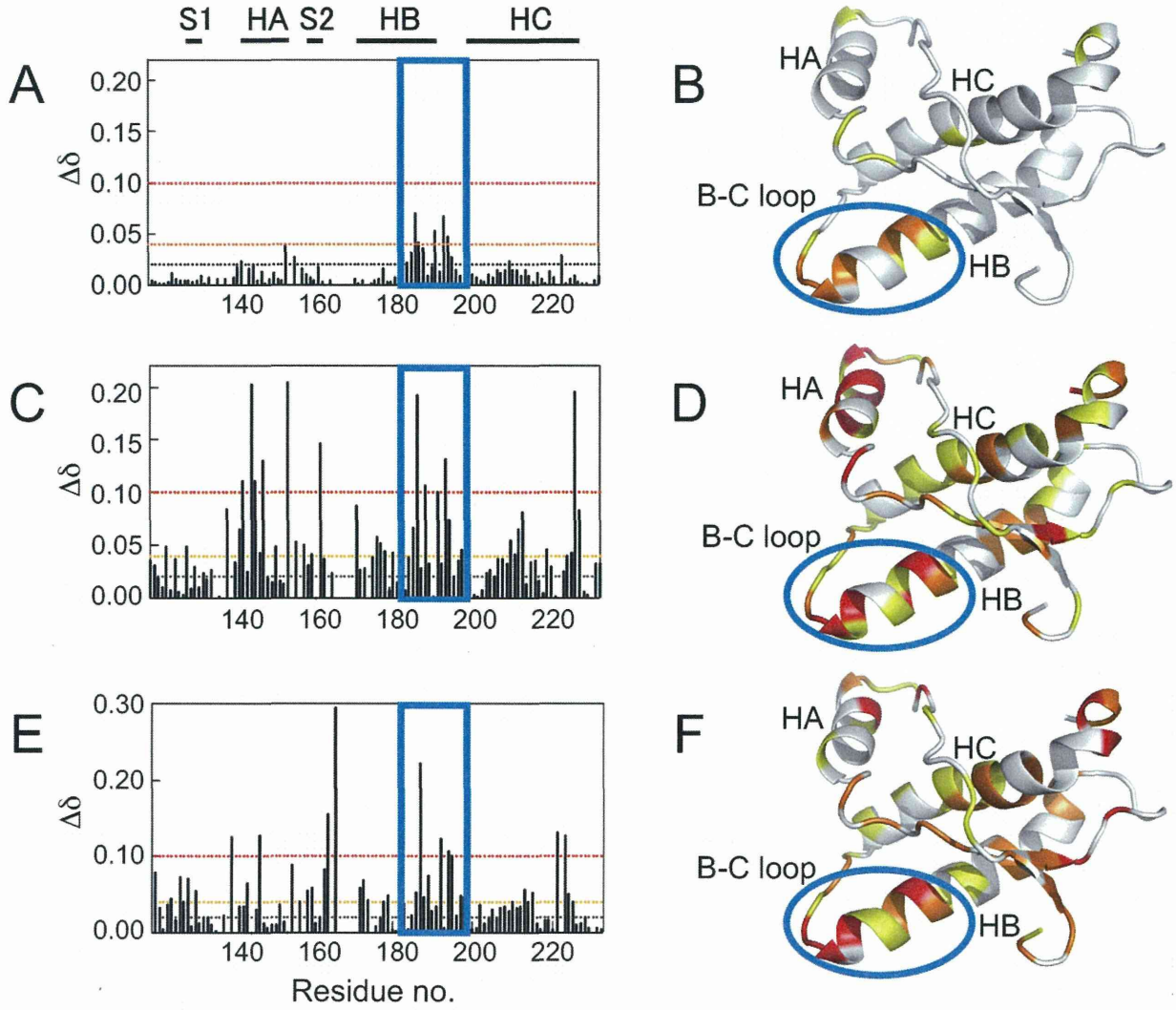


Figure 3. Chemical shift perturbations (Δδ) as a function of residue number and the mapping of Δδ onto the three dimensional structure of recombinant mouse PrP(121–231) (PDB ID: 1AG2). (A, B) GJP49, (C, D) quinacrine, and (E, F) epigallocatechin gallate (EGCG). Final concentrations of the protein and compounds were, respectively, 31 μ*M* and 500 μ*M* for GJP49 and EGCG binding experiments and 32 μ*M* and 5.3 m*M* for quinacrine binding experiment.

Quinacrine binds to PrP^C in a nonspecific manner

We analyzed SPR responses to examine the specificity of the interaction between quinacrine and PrP^C [Fig. 2(A)]. The SPR response increased as the concentration of test compound increased, without any trend toward saturation [inset in Fig. 2(A)]. The SPR response for quinacrine could not be fit using a simple binding model, which suggested nonspecific adhesion possibly due to strong hydrophobic interactions.

It has been reported that quinacrine accumulates in the brain after long-term administration.[22] In addition to PrP^C binding, quinacrine may also bind to other proteins without being degraded, which would cause the observed accumulation in the brain and significant side effects. To characterize the interaction sites between quinacrine and PrP^C, we obtained the chemical shift perturbations (Δδ) from the ^1H-^{15}N HSQC spectra [Fig. 3(C)] and mapped these onto the three dimensional structure [Fig. 3(D)]. When quinacrine was added to a 30.6 μ*M* solution of PrP^C at a final concentration of 500 μ*M*, no significant change in chemical shift was observed. When we increased the concentration of quinacrine (to 5.3 m*M*), the chemical shift perturbations increased to >0.04 for residues covering the entire molecular surface, particularly around parts of helices A, B, and C, including residues Y225, Y226, and D227. This result strongly suggested a nonspecific interaction.

Based on the results of an *ex vivo* assay using the GT + FK cell line,[23] the IC_{50} values for GN8 and quinacrine were 1.40 μ*M* and 1.11 μ*M*, respectively.[8] The mean Δδ values with GN8 at 1.0 m*M* and PrP^C at 26 μ*M* roughly corresponded to those for quinacrine at 5.3 m*M* and PrP^C at 32.4 μ*M*, respectively. Although their IC_{50} values were close, quinacrine required an approximately fourfold

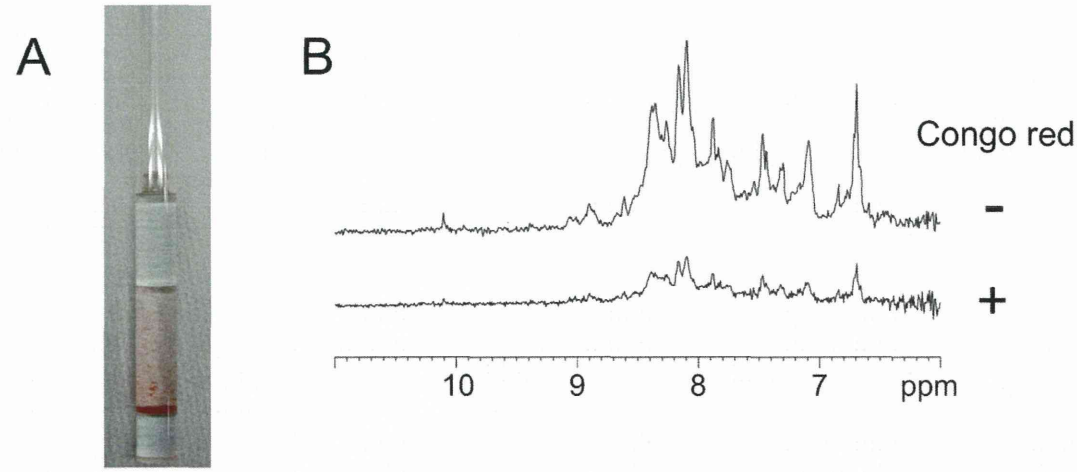


Figure 4. Aggregation and precipitation of mouse PrP(121–231) by Congo red. (A) Precipitation of mouse PrP(121–231) by Congo red in a Shigemi tube. (B) 1D ^{1}H NMR spectra with (lower trace) and without (upper trace) Congo red.

greater concentration than GN8 to reach the same $\Delta\delta$ level in its NMR spectrum.

This discrepancy can be explained by the differences in the specificities of the ligand–protein interaction (i.e., the binding regions for quinacrine are broadly distributed over the entire protein surface). In addition, the local structural stability of the ligand–protein complex may be important because structural fluctuations may reduce the $\Delta\delta$ values through an averaging effect. Quinacrine has also been reported to be transported into the intracellular space by endocytosis; lysosomes have a 10,000 times greater concentration of quinacrine than what is found in the extracellular space.[24] Thus, locally concentrated quinacrine may have been able to interact with PrP^{C} in the *ex vivo* experiment.

EGCG binds to PrP^{C} in a nonspecific manner and partially aggregates PrP^{C}

Although we measured the SPR responses of PrP^{C} upon binding with EGCG [Fig. 2(B)], the interaction was too strong to reverse the sensorgram to its original level, even after an exhaustive washing and subsequent recovery procedure. Before NMR determinations, we found that mixing PrP^{C} and EGCG at a molar ratio of 1:10 resulted in precipitating 80% of the PrP^{C} (data not shown). When we measured the ^{1}H-^{15}N HSQC spectrum for PrP^{C} and EGCG at a molar ratio of 1:8.2, we observed large chemical shift perturbations ($\Delta\delta$) for the regions that diffusely covered the entire molecule [Fig. 3(F)], which indicated nonspecific binding. Although EGCG binds to PrP^{C}, which forms aggregates and causes the precipitation of PrP^{C}, a significant fraction of this complex remains in the soluble phase.

Congo red promotes aggregation and reduces the concentration of PrP^{C}

Affinity analysis showed that increased concentrations of Congo red resulted in significantly increased SPR responses [Fig. 2(C)] without any significant saturation effect [inset in Fig. 2(C)]. Thus, Congo red nonspecifically adheres to PrP^{C}. In addition, PrP^{C} was precipitated as the concentration of Congo red was increased in the NMR tube [Fig. 4(A)]. To characterize this interaction at the atomic level, we measured the NMR spectrum for a sample of PrP^{C} and Congo red at a molar ratio of 1:2, but we could not observe any significant chemical shift perturbation ($\Delta\delta$; data not shown). When we increased the molar ratio of PrP^{C} and Congo red to 1:16.3, PrP^{C} was precipitated [Fig. 4(A)] and the signal heights were shifted by 20% [Fig. 4(B)].

Consequently, we could not obtain a high quality two-dimensional spectrum for the chemical shift perturbation. Thus, under *in vivo* conditions, Congo red binds nonspecifically to PrP^{C} and promotes its aggregation. In addition, even at a lower PrP^{C} concentration, this would reduce the amount of available PrP^{C} required for the conversion reaction. Congo red also induced aggregation of hen lysozyme and reduced the protein concentration (Supporting Information Fig. 1), indicating that the aggregation by Congo red is not specific to PrP^{C}.

PPS promotes aggregation and reduces the concentration of PrP^{C}

Affinity analysis showed that increased concentrations of PPS resulted in significantly increased SPR responses, and incomplete dissociation [Fig. 2(D)]. The amount of PPS that was bound was greater for full-length PrP, but PPS did significantly bind to the C-terminal domain of this protein (data not shown). Thus, PPS adheres strongly and nonspecifically to PrP^{C}. In addition, PrP^{C} was precipitated as the concentration of PPS was increased (data not shown), as was observed with PrP^{C} and Congo red. Thus, under *in vivo* conditions, PPS nonspecifically binds to PrP^{C} and promotes its aggregation. PPS induced

Table II. *Classification of the Antiprion Mechanisms*

Class	Mechanism	Biacore	NMR
I	Specific conformational stabilization of PrP^C	Specific binding	Specific binding
II	Nonspecific stabilization of PrP^C	Nonspecific binding	Nonspecific binding
III	Aggregation and precipitation to reduce PrP^C population	Nonspecific binding	Not detectable
IV	Interaction with molecules other than PrP^C	No binding	No interaction

aggregation of hen lysozyme and reduced the protein concentration (Supporting Information Table I), indicating that the aggregation by PPS is not specific to PrP^C.

CP-60, Edaravone derivative 13, D-PEN, and indole-3-glyoxylamide derivatives do not bind to PrP^C

The responses on the SPR sensorgrams using mouse PrP(23–231) and CP-60 [Fig. 2(E)], edaravone derivative 13 [Fig. 2(F)], D-PEN [Fig. 2(G)], indole-3-glyoxylamide derivative 10 [Fig. 2(H)], and indole-3-glyoxylamide derivatives 2, 12, and 13 (data not shown) were all minimal. Thus, these data suggested that these compounds did not directly interact with either the C- or N-terminal domain of PrP^C. As shown in Table I, CP-60, edaravone derivative 13, D-PEN, and indole-3-glyoxylamide derivatives did not inhibit PrP^{Sc} formation. The mean PrP^{Sc} levels using 10 μM of each of these compounds were 94.3, 93.0, 95.5, 81.8, 110, and 103% for CP-60, edaravone derivative 13, D-PEN, and indole-3-glyoxylamide derivatives 2, 10, and 12, respectively. Indole-3-glyoxylamide derivative 13 was toxic at a concentration of 10 μM. Thus we could not confirm the antiprion activities for these compounds. This could depend on the strain used for the *ex vivo* assay.

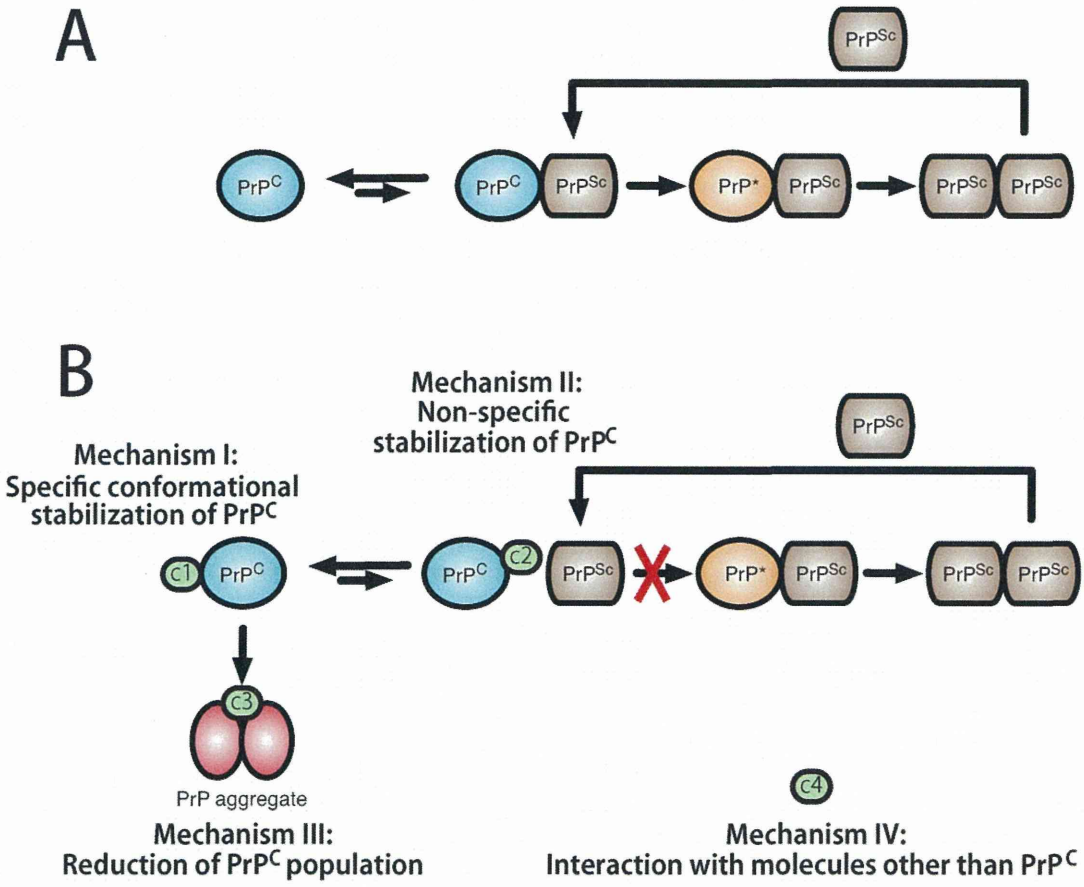


Figure 5. Illustration of the pathogenic conversion process from PrP^C to PrP^{Sc} in the absence (A) and presence (B) of antiprion compounds and the classification of antiprion mechanisms based on interactions with PrP^C. Mechanism I molecules (c1), designated "medical chaperones," stabilize the PrP^C conformation. Mechanism II molecules (c2) bind to PrP^C nonspecifically to stabilize the PrP^C and interfere with the interaction between PrP^C and PrP^{Sc}. Mechanism III molecules (c3) induce prion protein aggregation and cause precipitation, which reduces the amount of PrP^C. Mechanism IV molecules (c4) interact with molecules other than PrP^C, such as PrP^{Sc}, PrP*, or membrane proteins.

In this study, we used GT + FK cells that produce significant amounts of PrP^{res} as well as PrP^{C}. Hence the assay system using GT + FK cells is reliable, reproducible, and less sensitive to minor perturbations such as those due to the culture conditions.[9] However, it has been reported that administering these compounds to prion-infected mice days after the onset of their symptoms extended their life spans.[19] Although these compounds are reported to have antiprion activities, we could not confirm them. In other systems, they may possibly interact with the intermediate prion protein (PrP^{*}), PrP^{Sc}, membrane proteins, or other relevant proteins,[25] and may regulate a prion's toxicity.

Discussion

Classifying antiprion mechanisms

Although there have been many reports on antiprion compounds, their mechanisms of action have not been clearly defined. In particular, comparisons of the mechanisms between various antiprion compounds are not available. Antiprion compounds can be classified as shown in Table II and Figure 5 based on their binding properties to PrP^{C}. We classified antiprion activities into four categories: I, specific conformational stabilization of PrP^{C};[6] II, nonspecific stabilization including the interference with the interaction between PrP^{C} and PrP^{Sc} as well as the contribution of binding to the hot spots; III, aggregation and precipitation to reduce the amount of available PrP^{C}; and IV, interactions with molecules other than PrP^{C}. Here, we examined the activities of representative antiprion compounds (GJP49, GJP14, quinacrine, EGCG, Congo red, PPS, CP-60, edaravone derivative 13, D-PEN, and indole-3-glyoxylamide derivatives) and characterized their antiprion activities based on their binding properties with PrP^{C}.

Characterizing antiprion compounds

GN8 and GJP49 directly bound to PrP^{C} at specific binding sites on PrP^{C} and inhibited its pathogenic conversion. These binding regions are shown in Figure 1(B,E), and are essentially the same as those for GN8 (i.e., hot spot).[6] Specific binding to this region is associated with mechanisms I.

In contrast, quinacrine and EGCG bound nonspecifically to PrP^{C} as shown in Figure 3(C,E). To check correlation between the $\Delta\delta$ values for each compound, we compared $\Delta\delta$ values for each residue between each compound (Supporting Information Fig. 2). However, except GJP14-GJP49, clear correlation could not be observed. This also indicates the nonspecific binding property of quinacrine and EGCG. This nonspecific binding may be attributed to mechanism II. For conformational stabilization, binding to the specific hot spot that is responsible for conformational instability is required, while for interference between PrP^{C} and PrP^{Sc}, nonspecific binding to PrP^{C} would be allowed to some extent because they mutually interact with large contact areas that form oligomers,[26] amyloid fibrils,[27] or nonspecific aggregates[28] depending on the characteristics of the relevant strains. Hence, small compounds can interfere with these interactions at various nonspecific sites. However, these binding regions of these compounds also included the hot spots, therefore, there may be partially attributed to mechanism I.

Congo red and PPS induced PrP^{C} aggregation and reduced its effective concentration. In general, many small compounds inhibit amyloid formation by binding with amyloidogenic proteins via π–π interactions and form aggregates.[29,30] Thus, their mechanisms of action are basically described by mechanism III. However, the nonspecific binding properties of these compounds indicate that they may also act by mechanism II, especially at low concentrations of these compounds and protein.

Finally, CP-60, Edaravone derivative 13, D-PEN, and indole-3-glyoxylamide derivatives do not bind to PrP^{C}, and we could not confirm their antiprion activities. Though in other system, they may possibly interact with other proteins or membranes and thus, indirectly reduce the amount of PrP^{Sc}. Thus, mechanism IV could be assigned.

Compounds associated with mechanism I share a common binding site

We compared the binding sites for the compounds that bound directly to PrP^{C} and found that GJP49, GJP14, quinacrine, and EGCG bound to a common site, the C-terminal region of helix B and the B-C loop, which is shown in Figure 3(A,C,E) (blue boxes) and Figure 3(B,D,F) (blue oval shapes). In a previous study, we demonstrated that GN8 also bound to this region.[6,8] This region is known to undergo global fluctuations on a time scale of micro- to milliseconds.[6] Each of the compounds examined here with antiprion activity via direct binding to PrP^{C} had affinity for this region. Thus, these compounds can be referred to as "medical chaperones" that stabilize the native conformation of the target protein and inhibit its transition to the abnormal conformation, PrP^{Sc}.

Medical chaperones can be considered to suppress the seismic fluctuations of a protein (i.e., protein quakes),[31,32] which results in destroying its native structure. Among the compounds examined in this study, GJP49 specifically bound to the common binding site. Thus, GJP49 may not induce any side effects and could be a candidate therapeutic agent for prion diseases.

It has been assumed that factor X is required for the pathogenic conversion of PrP,[25] although factor X has not yet been identified. It is reported that

quinacrine binds to Tyr225, Tyr226, and Gln227 (Asp227 for mouse)[24] which is located near the factor X" binding sites.[33] As shown in Figure 3(D), Δδ values in the presence of 2.67 m*M* quinacrine indicated that Tyr225, Tyr226, and Gln227 were also involved in the binding with quinacrine. Thus, quinacrine may inhibit the pathogenic conversion of PrP by competitive binding with factor X. However, it must be noted that even the maximum Δδ values for Tyr226 at this site were comparable to that at other quinacrine binding sites [Fig. 3(C)], and that this site was not a common binding site for the other compounds (Fig. 3).

Medical chaperones

It is considered that a high-energy barrier exists between PrP^C and PrP^{Sc},[34,35] and that prion proteins in the PrP^C state rarely overcome this barrier; they only achieve the PrP^{Sc} state when triggered by some unknown causes. Compounds in the first and second classes (GN8, GJP49, GJP14, quinacrine, and EGCG) bound to the residues surrounding the major binding pocket, which may be the cause of the seismic fluctuations that lead to pathogenic conversion. Medical chaperones prevent these protein quakes and suppress these fluctuations. Thus, these would stabilize the PrP^C conformation (i.e., reduce the free energy level of PrP^C), which would result in an increase in the energy barrier between PrP^C and PrP^{Sc}.[36] Therefore, these would suppress the pathogenic conversion of PrP^C to PrP^{Sc}.

Although the PrP^C conformations are nearly identical between species, PrP^{Sc} conformations are quite heterogeneous and are referred to as different "strains." Strain conformation is conserved after transmission between individuals.[37,38] Antiprion compounds that vary significantly in their efficiency, depending on the strain,[39] may directly affect the PrP^{Sc} conformation rather than the PrP^C conformation. A great advantage of a "medical chaperone" is that its effect is strain-independent[6,40] because it acts on PrP^C which has a common structure in nearly all mammals.

Medical chaperones also act by mechanism II, interference with the interaction between PrP^C and PrP^{Sc}, as described in the following section.

Interference with the interaction between PrP^C and PrP^{Sc} during the template-dependent self-replication process

PrP^C interacts with PrP^{Sc} during the process of pathogenic conversion. Compounds that bind to PrP^C nonspecifically can interfere with the interaction between PrP^C and PrP^{Sc}. Although there have been no evidence for the direct interference of the interaction between PrP^C and PrP^{Sc}, this mechanism constructs the critical difference from the originally proposed chemical chaperone for the misfolding diseases.[41] Specific binding of a medical chaperone with this function could be advantageous in clinical use by reducing side effects.

The possibility of administering a cocktail of antiprion compounds

We classified 13 different antiprion compounds into four classes based on their antiprion mechanisms of action. Compounds associated with different mechanisms of action can be administered together (e.g., GJP49 and indole-3-glyoxylamide derivative 13). GJP49 binds specifically to PrP^C and suppresses the available population in the intermediate state (PrP*), while indole-3-glyoxylamide derivative 13 does not bind to PrP^C, which indicates that this compound interacts with molecules other than PrP^C. Indole-3-glyoxylamide derivative 13 reduced PrP^{Sc} formation at the nanomolar concentration range when it was screened using a scrapie infected mouse cell line (SMB).[42] Cocktail administration of antiprion compounds might possibly reduce the concentration of each compound, which could decrease their associated side effects. Cocktail administration has also been reported for HIV[43,44] and yeast prions,[45] and could be useful for mammalian prion diseases.

Toward further optimization of chemical structures

GJP49, GJP14, quinacrine, and EGCG bind to the major pocket of PrP^C and contribute to its structural stabilization. We need to quantitatively evaluate the degree of stabilization using relaxation measurements for PrP^C, particularly around the major pocket in the presence of each compound. We also need to determine the structures of complexes using X-ray crystallography. This dynamical and structural information will be important for the rational optimization of antiprion compounds.

A specific interaction at the common binding site is critical for antiprion efficiency according to mechanism I. Specific interactions may also be advantageous for reducing side effects. Therefore, selecting specifically interacting derivatives may be important for compound optimization. For this purpose, we may able to use the Δδ values from NMR as shown in Figure 3. For example, quinacrine binds to PrP^C in a nonspecific manner [Figs. 2(A) and 3(C,D)] and has been reported to be toxic[22] (Table I). By selecting specifically interacting derivatives of quinacrine, we may be able to find compounds that have fewer side effects than quinacrine.

It should be noted that the above discussion was focused entirely on the interactions between compounds and PrP^C. When the interactions of compounds with PrP^{Sc} or the cell surface can be monitored, we will obtain additional insights into the

detailed mechanisms of prion diseases, which will aid in the development of therapeutics.

Conclusion

Elucidating the molecular mechanisms of action of previously identified antiprion compounds is one of the most important steps for developing novel antiprion drugs and optimizing the most promising compounds. When this strategy (i.e., "medical chaperones") has been established for prion diseases, it could also be applied to the rational design of drugs for other neurodegenerative diseases, such as Alzheimer's disease, Parkinson's disease, and amyotrophic lateral sclerosis.

Materials and Methods

Compounds

Compounds GJP49, GJP14, GJP55, and Indole-3-glyoxylamide derivatives were purchased from ASINEX (Moscow, Russia). Quinacrine, Congo red, and EGCG were purchased from Sigma-Aldrich (Japan, Tokyo). PPS, CP-60, Edaravone derivative 13, and D-PEN were obtained from bene-Arzneimittel GmbH (Munich, Germany), Ambinter (Orléans, France), Labotest (Niederschöna, Germany), and Wako Pure Chemical Industries (Osaka, Japan), respectively.

Recombinant mouse PrP

An expression plasmid for recombinant mouse PrP residues 23–231 [PrP(23–231)] was prepared according to a previously described protocol.[6] An expression plasmid for mouse PrP residues 121–231 [PrP(121–231)] was a kind gift from Professor Kurt Wüthrich and Dr. Simone Hornemann.[46] Recombinant PrP was prepared as previously described.[6] The concentrations of mouse PrP(23–231) and PrP(121–231) were estimated by absorbance at 280 nm using specific absorbances (ε_{280}) = 2.68 and 1.49 $(mg/mL)^{-1}$ cm^{-1}, respectively.

SPR measurements

Interactions between prion proteins and the different compounds were analyzed using the Biacore T200 system (GE Healthcare, Buckinghamshire, UK). Recombinant mouse PrP was immobilized on a sensor chip (CM5) according to the manufacturer's instructions. Varying concentrations of each compound were injected into the running buffer (10 m*M* HEPES buffer, pH 7.4, containing 0.15 *M* NaCl, 0.1% Surfactant P20, and 5% DMSO) for 1 to 2 min at a flow rate of 30 mL/min. Running buffer without compounds was then injected for 10 min at the same flow rate. Data were corrected by subtracting the response of blank sensor chip from that of protein-bound sensor chip, so the contribution of nonspecific binding of the compounds to SPR chip surface was removed.

Nuclear magnetic resonance (NMR) measurements and data analysis

For NMR measurements, mouse PrP(121–231) uniformly labeled with ^{15}N was prepared in 30 m*M* acetate-d_3 buffer (pH 4.5) containing 1 m*M* NaN_3, 4.5 μ*M* AEBSF, 20 μ*M* EDTA, 0.4 μ*M* Bestatin, 0.06 μ*M* pepstatin, 0.06 μ*M* E-64, and 1 n*M* DSS dissolved in 90% H_2O/10% D_2O. NMR spectra were recorded at 20.0°C on a Bruker Avance600 spectrometer (Bruker BioSpin, Rheinstetten, Germany) at Gifu University. The spectrometer operated at a ^1H frequency of 600.13 MHz and a ^{15}N frequency of 60.81 MHz. A 5-mm ^1H inverse detection probe with triple-axis gradient coils was used for all measurements. ^1H-^{15}N HSQC spectra were acquired with 2048 complex points covering 9600 Hz for ^1H and 256 complex points covering 1200 Hz for ^{15}N. NMR data were processed using the TOPSPIN software package (Bruker BioSpin, Rheinstetten, Germany) and NMR assignment and integration software Sparky.[47] Resonance frequencies in these spectra were identified using the chemical shift lists for mouse PrP(121–231).[48] The backbone ^1H and ^{15}N chemical shifts for a compound-bound protein were assigned by tracing the corresponding peaks in ^1H-^{15}N HSQC spectra determined with varying concentrations of the compounds. The protein structures were generated using PyMOL Molecular Graphics System Version 0.99rc9 (Schrödinger, LLC, NY).

Ex vivo *assay for antiprion activity*

An *ex vivo* assay was performed as described previously.[7] Briefly, we used an immortalized neuronal mouse cell line that was persistently infected with a human TSE agent (Fukuoka-1 strain),[23] which was designated GT + FK. This cell line was grown and maintained in 5% CO_2 at 37°C in D-MEM (Invitrogen, Carlsbad, CA) supplemented with 10% fetal bovine serum (Equitech-bio, Kerrville, TX), 50 U/mL penicillin G sodium, and 50 μg/mL streptomycin sulfate (Invitrogen, Carlsbad, CA). Stock solutions of the test compounds were prepared fresh in 100% DMSO at a concentration of 10 m*M* and stored at 4°C. Before use, the test compounds were diluted with medium to a concentration of 10 μ*M*. Control cells were treated with medium containing solvent alone (0.1%). Approximately 3.0×10^5 cells were plated in each well of a six-well plate, and treatment with a test compound was started 15 h later. After 72 h of treatment, cells were lysed in 150 μL of 1× Triton X-100/DOC lysis buffer.[49] Western blotting for PrP^{Sc} was done as described previously.[7]

Acknowledgments

The authors thank Professor Kurt Wüthrich and Dr. Simone Hornemann for kindly providing them with the expression plasmid for mouse PrP(121–231).

Pentosan polysulfate was kindly provided by Professor Masami Niwa and Professor Noriyuki Nishida. The authors thank Dr. Tomoharu Matsumoto, Ms. Junko Matsubara, and Ms. Miki Horii for their assistance with protein sample preparation. The authors thank Ms. Tomomi Saeki for technical help with evaluating antiprion activity.

References

1. Prusiner SB (1998) Prions. Proc Natl Acad Sci USA 95: 13363–13383.
2. Chesebro B (1999) Prion protein and the transmissible spongiform encephalopathy diseases. Neuron 24: 503–506.
3. Prusiner SB (1982) Novel proteinaceous infectious particles cause scrapie. Science 216:136–144.
4. O'Riordan JF, Goldstick TK, Ditzel J, Ernest JT (1983) Characterization of oxygen-hemoglobin equilibrium curves using nonlinear regression of the Hill equation: parameter values for normal human adults. Adv Exp Med Biol 159:435–444.
5. Harrison PM, Bamborough P, Daggett V, Prusiner SB, Cohen FE (1997) The prion folding problem. Curr Opin Struct Biol 7:53–59.
6. Kuwata K, Nishida N, Matsumoto T, Kamatari YO, Hosokawa-Muto J, Kodama K, Nakamura HK, Kimura K, Kawasaki M, Takakura Y, Shirabe S, Takata J, Kataoka Y, Katamine S (2007) Hot spots in prion protein for pathogenic conversion. Proc Natl Acad Sci USA 104:11921–11926.
7. Hosokawa-Muto J, Kamatari YO, Nakamura HK, Kuwata K (2009) Variety of antiprion compounds discovered through an in silico screen based on cellular-form prion protein structure: Correlation between antiprion activity and binding affinity. Antimicrob Agents Chemother 53:765–771.
8. Kimura T, Hosokawa-Muto J, Kamatari YO, Kuwata K (2011) Synthesis of GN8 derivatives and evaluation of their antiprion activity in TSE-infected cells. Bioorg Med Chem Lett 21:1502–1507.
9. Kimura T, Hosokawa-Muto J, Asami K, Murai T, Kuwata K (2011) Synthesis of 9-substituted 2,3,4,9-tetrahydro-1H-carbazole derivatives and evaluation of their anti-prion activity in TSE-infected cells. Eur J Med Chem 46:5675–5679.
10. Doh-Ura K, Iwaki T, Caughey B (2000) Lysosomotropic agents and cysteine protease inhibitors inhibit scrapie-associated prion protein accumulation. J Virol 74: 4894–4897.
11. Perrier V, Wallace AC, Kaneko K, Safar J, Prusiner SB, Cohen FE (2000) Mimicking dominant negative inhibition of prion replication through structure-based drug design. Proc Natl Acad Sci USA 97:6073–6078.
12. Kocisko DA, Baron GS, Rubenstein R, Chen J, Kuizon S, Caughey B (2003) New inhibitors of scrapie-associated prion protein formation in a library of 2000 drugs and natural products. J Virol 77:10288–10294.
13. Korth C, May BC, Cohen FE, Prusiner SB (2001) Acridine and phenothiazine derivatives as pharmacotherapeutics for prion disease. Proc Natl Acad Sci USA 98: 9836–9841.
14. Caughey B, Ernst D, Race RE (1993) Congo red inhibition of scrapie agent replication. J Virol 67:6270–6272.
15. Rambold AS, Miesbauer M, Olschewski D, Seidel R, Riemer C, Smale L, Brumm L, Levy M, Gazit E, Oesterhelt D, Baier M, Becker CF, Engelhard M, Winklhofer KF, Tatzelt J (2008) Green tea extracts interfere with the stress-protective activity of PrP and the formation of PrP. J Neurochem 107:218–229.
16. Doh-ura K, Ishikawa K, Murakami-Kubo I, Sasaki K, Mohri S, Race R, Iwaki T (2004) Treatment of transmissible spongiform encephalopathy by intraventricular drug infusion in animal models. J Virol 78: 4999–5006.
17. Tsuboi Y, Doh-Ura K, Yamada T (2009) Continuous intraventricular infusion of pentosan polysulfate: clinical trial against prion diseases. Neuropathology 29:632–636.
18. Kimata A, Nakagawa H, Ohyama R, Fukuuchi T, Ohta S, Doh-ura K, Suzuki T, Miyata N (2007) New series of antiprion compounds: pyrazolone derivatives have the potent activity of inhibiting protease-resistant prion protein accumulation. J Med Chem 50:5053–5056.
19. Sigurdsson EM, Brown DR, Alim MA, Scholtzova H, Carp R, Meeker HC, Prelli F, Frangione B, Wisniewski T (2003) Copper chelation delays the onset of prion disease. J Biol Chem 278:46199–46202.
20. Thompson MJ, Borsenberger V, Louth JC, Judd KE, Chen B (2009) Design, synthesis, and structure-activity relationship of indole-3-glyoxylamide libraries possessing highly potent activity in a cell line model of prion disease. J Med Chem 52:7503–7511.
21. Kuwata K, Kamatari YO, Akasaka K, James TL (2004) Slow conformational dynamics in the hamster prion protein. Biochemistry 43:4439–4446.
22. Nakajima M, Yamada T, Kusuhara T, Furukawa H, Takahashi M, Yamauchi A, Kataoka Y (2004) Results of quinacrine administration to patients with Creutzfeldt-Jakob disease. Dement Geriatr Cogn Disord 17: 158–163.
23. Milhavet O, McMahon HE, Rachidi W, Nishida N, Katamine S, Mange A, Arlotto M, Casanova D, Riondel J, Favier A, Lehmann S (2000) Prion infection impairs the cellular response to oxidative stress. Proc Natl Acad Sci USA 97:13937–13942.
24. Vogtherr M, Grimme S, Elshorst B, Jacobs DM, Fiebig K, Griesinger C, Zahn R (2003) Antimalarial drug quinacrine binds to C-terminal helix of cellular prion protein. J Med Chem 46:3563–3564.
25. Telling GC, Scott M, Mastrianni J, Gabizon R, Torchia M, Cohen FE, DeArmond SJ, Prusiner SB (1995) Prion propagation in mice expressing human and chimeric PrP transgenes implicates the interaction of cellular PrP with another protein. Cell 83:79–90.
26. Masel J, Genoud N, Aguzzi A (2005) Efficient inhibition of prion replication by PrP-Fc(2) suggests that the prion is a PrP(Sc) oligomer. J Mol Biol 345:1243–1251.
27. Wille H, Zhang GF, Baldwin MA, Cohen FE, Prusiner SB (1996) Separation of scrapie prion infectivity from PrP amyloid polymers. J Mol Biol 259:608–621.
28. Pham N, Yin S, Yu S, Wong P, Kang SC, Li C, Sy MS (2008) Normal cellular prion protein with a methionine at position 129 has a more exposed helix 1 and is more prone to aggregate. Biochem Biophys Res Commun 368:875–881.
29. Lamberto GR, Binolfi A, Orcellet ML, Bertoncini CW, Zweckstetter M, Griesinger C, Fernandez CO (2009) Structural and mechanistic basis behind the inhibitory interaction of PcTS on alpha-synuclein amyloid fibril formation. Proc Natl Acad Sci USA 106:21057–21062.
30. Nicoll AJ, Trevitt CR, Tattum MH, Risse E, Quarterman E, Ibarra AA, Wright C, Jackson GS, Sessions RB, Farrow M, Waltho JP, Clarke AR, Collinge J (2010) Pharmacological chaperone for the structured domain of human prion protein. Proc Natl Acad Sci USA 107: 17610–17615.

31. Ansari A, Berendzen J, Bowne SF, Frauenfelder H, Iben IE, Sauke TB, Shyamsunder E, Young RD (1985) Protein states and proteinquakes. Proc Natl Acad Sci USA 82:5000–5004.
32. Itoh K, Sasai M (2004) Dynamical transition and proteinquake in photoactive yellow protein. Proc Natl Acad Sci USA 101:14736–14741.
33. Cohen FE, Prusiner SB (1998) Pathologic conformations of prion proteins. Annu Rev Biochem 67:793–819.
34. Cohen FE, Pan KM, Huang Z, Baldwin M, Fletterick RJ, Prusiner SB (1994) Structural clues to prion replication. Science 264:530–531.
35. Cohen FE, Prusiner SB, Structural studies of prion proteins. In: Prusiner SB, Ed. (1999) Prion biology and diseases. New York: Cold Spring Harbor, pp191–228.
36. Yamamoto N, Kuwata K (2009) Regulating the conformation of prion protein through ligand binding. J Phys Chem B 113:12853–12856.
37. Telling GC, Parchi P, DeArmond SJ, Cortelli P, Montagna P, Gabizon R, Mastrianni J, Lugaresi E, Gambetti P, Prusiner SB (1996) Evidence for the conformation of the pathologic isoform of the prion protein enciphering and propagating prion diversity. Science 274: 2079–2082.
38. Safar J, Wille H, Itri V, Groth D, Serban H, Torchia M, Cohen FE, Prusiner SB (1998) Eight prion strains have PrP(Sc) molecules with different conformations. Nat Med 4:1157–1165.
39. Kawasaki Y, Kawagoe K, Chen CJ, Teruya K, Sakasegawa Y, Doh-ura K (2007) Orally administered amyloidophilic compound is effective in prolonging the incubation periods of animals cerebrally infected with prion diseases in a prion strain-dependent manner. J Virol 81:12889–12898.
40. Li J, Browning S, Mahal SP, Oelschlegel AM, Weissmann C (2010) Darwinian evolution of prions in cell culture. Science 327:869–872.
41. Sawkar AR, Cheng WC, Beutler E, Wong CH, Balch WE, Kelly JW (2002) Chemical chaperones increase the cellular activity of N370S beta -glucosidase: a therapeutic strategy for Gaucher disease. Proc Natl Acad Sci USA 99:15428–15433.
42. Kanu N, Imokawa Y, Drechsel DN, Williamson RA, Birkett CR, Bostock CJ, Brockes JP (2002) Transfer of scrapie prion infectivity by cell contact in culture. Curr Biol 12:523–530.
43. Montagnier L (2010) 25 years after HIV discovery: prospects for cure and vaccine. Virology 397:248–254.
44. Shafer RW, Vuitton DA (1999) Highly active antiretroviral therapy (HAART) for the treatment of infection with human immunodeficiency virus type 1. Biomed Pharmacother 53:73–86.
45. Roberts BE, Duennwald ML, Wang H, Chung C, Lopreiato NP, Sweeny EA, Knight MN, Shorter J (2009) A synergistic small-molecule combination directly eradicates diverse prion strain structures. Nat Chem Biol 5:936–946.
46. Hornemann S, Korth C, Oesch B, Riek R, Wider G, Wuthrich K, Glockshuber R (1997) Recombinant full-length murine prion protein, mPrP(23-231): purification and spectroscopic characterization. FEBS Lett 413: 277–281.
47. Goddard TD, Kneller DG (2001) SPARKY, version 3. San Francisco: University of California.
48. Riek R, Hornemann S, Wider G, Billeter M, Glockshuber R, Wuthrich K (1996) NMR structure of the mouse prion protein domain PrP(121-231). Nature 382: 180–182.
49. Nishida N, Harris DA, Vilette D, Laude H, Frobert Y, Grassi J, Casanova D, Milhavet O, Lehmann S (2000) Successful transmission of three mouse-adapted scrapie strains to murine neuroblastoma cell lines overexpressing wild-type mouse prion protein. J Virol 74:320–325.

THE JOURNAL OF PHYSICAL CHEMISTRY Letters

Letter

pubs.acs.org/JPCL

RI-MP2 Gradient Calculation of Large Molecules Using the Fragment Molecular Orbital Method

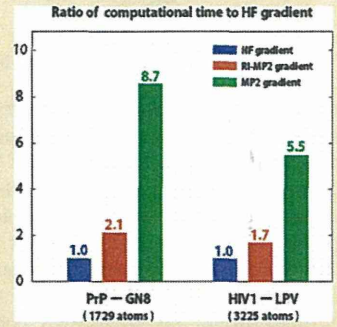
Takeshi Ishikawa[†,‡] and Kazuo Kuwata*[,†,‡]

[†]Division of Prion Research, Center for Emerging Infectious Disease, Gifu University, 1-1 Yanagido, Gifu 501-1194, Japan

[‡]CREST Project, Japan Science and Technology Agency, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan

Supporting Information

ABSTRACT: The second-order Møller–Plesset perturbation theory (MP2) gradient using resolution of the identity approximation (RI-MP2 gradient) was combined with the fragment molecular orbital (FMO) method to evaluate the gradient including electron correlation for large molecules. In this study, we adopted a direct implementation of the RI-MP2 gradient, in which a characteristic feature of the FMO scheme was utilized. Test calculations with a small peptide presented a computational advantage of the RI-MP2 gradient over the canonical MP2 gradient. In addition, it was shown that the error of the RI-MP2 gradient, caused by RI approximation, was negligible. As an illustrative example, we performed gradient calculations for two biomolecules—a prion protein with GN8 and a human immunodeficiency virus type 1 (HIV1) protease with lopinavir (LPV). These calculations demonstrated that the gradient including the correlation effect could be evaluated with only about twice the computational effort of the Hartree–Fock (HF) gradient.



SECTION: Molecular Structure, Quantum Chemistry, General Theory

Resolution of the identity (RI) approximation[1,2] has been used for the speed-up of several quantum chemical methods (e.g., Hartree–Fock (HF),[3,4] second-order Møller–Plesset perturbation theory (MP2),[5−8] and coupled cluster theory[9]). The original study of the first derivative of the RI-MP2 correlation energy with respect to the atomic coordinates was reported by Weigend and Häser,[6] in which significant computational effort was saved while maintaining accuracy. An extension of the RI-MP2 gradient to restricted open-shell systems was successfully performed[7] and was used to evaluate the potential energy surface of a tetracyanoethylene dimer anion. A parallel implementation of the RI-MP2 gradient for large-scale calculations was reported by Hättig et al.,[8] where not only the operation count but also the communication issue were considered. Some demonstrative calculations including the geometry optimization of a fullerene were reported using their parallel algorithm.

The fragment molecular orbital (FMO) method[10,11] is one of the most efficient approaches for ab initio quantum chemical calculation of large molecules including biomolecular systems. This method has been extended to geometry optimization[12−14] and molecular dynamics (MD) simulation[15−18] for large systems, where the gradient including electron correlation is strongly required. Recently, Mochizuki et al. reported the MP2 gradient calculations in the FMO method, in which an MD simulation of water cluster was performed.[19] Nagata et al. presented a high-accuracy MP2 gradient method based on the FMO scheme, in which the orbital response contribution is fully considered.[20] As seen above, the MP2 gradient with the FMO method is now available, but studies using RI approximation have been reported only for energy calculations.[21−23] It is clear that the speed-up of the MP2 gradient calculation using the FMO scheme will be very useful for the geometry optimization or ab initio MD simulation. In this Letter, we report an implementation of the RI-MP2 gradient into the FMO scheme to reduce the computational cost for large molecules including biomolecular systems. The following text describes the theory and implementation and several test and demonstrative calculations.

The electron correlation energy in the RI-MP2 method is written as[6]

$$E^{\mathrm{MP2}} = \sum_{ij}\sum_{ab} \frac{2(ialjb)^{\mathrm{RI}} - (iblja)^{\mathrm{RI}}}{\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b}(ialjb)^{\mathrm{RI}} \tag{1}$$

where ij and ab indicate the valence orbital and virtual orbital, respectively. The four-center electron-repulsion integrals are approximated using the two- and three-center electron-repulsion integrals, as shown in the following equations:

$$(ialjb)^{\mathrm{RI}} = \sum_{P} B_{ia}^{P} B_{jb}^{P} \tag{2}$$

$$B_{ia}^{P} = \sum_{Q} (ia|Q) V^{-1/2}{}_{QP} \tag{3}$$

Received: December 26, 2011
Accepted: January 15, 2012
Published: January 15, 2012

ACS Publications © 2012 American Chemical Society

$$V_{PQ} = (P|Q) \tag{4}$$

where PQ indicates the auxiliary basis functions. The first derivative of eq 1 with respect to an atomic coordinate ξ can be written as[6]

$$\begin{aligned}\frac{\partial}{\partial\xi}E^{\mathrm{MP2}} = {} & 4\sum_{\mu\nu P}\Gamma^{P}_{\mu\nu}(P|\mu\nu)^{\xi} - 2\sum_{PQ}\gamma_{PQ}(P|Q)^{\xi} \\ & - 2\sum_{\mu\nu}D^{\mathrm{MP2}}{}_{\mu\nu}h^{\xi}{}_{\mu\nu} \\ & - 2\sum_{\mu\nu}W^{\mathrm{MP2}}{}_{\mu\nu}S^{\xi}_{\mu\nu} \\ & + 2\sum_{\mu\nu\lambda\sigma}(2D^{\mathrm{MP2}}{}_{\mu\nu}D^{\mathrm{SCF}}{}_{\lambda\sigma} \\ & - D^{\mathrm{MP2}}{}_{\mu\sigma}D^{\mathrm{SCF}}{}_{\lambda\nu})(\mu\nu|\lambda\sigma)^{\xi}\end{aligned} \tag{5}$$

where superscript ξ indicates the derivative, and $\mu\nu\lambda\sigma$ indicates the normal basis function. This equation is equivalent to the theory reported by Weigend and Häser.[6] The definition of each variable in eq 5 is given in the Supporting Information.

In the FMO method, a target molecule is divided into small fragments, and its total properties are obtained from calculations of individual fragments (referred to as monomers) and pairs of the fragments (referred to as dimers).[10,11] These calculations include the electrostatic potential from the other fragments, called "environmental electrostatic potential (ESP)". The ESP can be determined by an iterative calculation, the so-called "monomer self-consistent charge (SCC)" process. The total energy at the HF level of theory is given as

$$E^{\mathrm{HF}} = \sum_{I}E'^{\mathrm{HF}}{}_{I} + \sum_{I<J}\Delta E^{\mathrm{HF}}{}_{IJ} \tag{6}$$

where $E'^{\mathrm{HF}}{}_{I}$ is the monomer energy without contribution of the ESP, and $\Delta E^{\mathrm{HF}}{}_{IJ}$ is the interaction energy between two fragments.[24] For the dimers constructed with largely separated fragments, $\Delta E^{\mathrm{HF}}{}_{IJ}$ is approximated only from the electrostatic interaction. This treatment is called "dimer-es" approximation.[24] The HF gradient is obtained as the derivative of this equation.[25] The MP2 correlation energy is written as[26,27]

$$\begin{aligned}E^{\mathrm{MP2}} = {} & \sum_{I}E^{\mathrm{MP2}}{}_{I} + \sum_{I<J}(E^{\mathrm{MP2}}{}_{IJ} - E^{\mathrm{MP2}}{}_{I} \\ & - E^{\mathrm{MP2}}{}_{J})\end{aligned} \tag{7}$$

where $E^{\mathrm{MP2}}{}_{I}$ and $E^{\mathrm{MP2}}{}_{IJ}$ are the correlation energy of the monomers and dimers, respectively. Consequently, gradient of the correlation energy is obtained as the derivative of this equation with respect to the atomic coordinates.[19,20] In this study, the derivative of each term of eq 7 is evaluated using eq 5. Note that for dimer calculations, additional ESP terms are included in eq 5. This is an important difference between the conventional quantum chemical method and the FMO method.[19,20]

In our implementation of the RI-MP2 gradient, a direct algorithm was adopted, that is, no disk storage was used. Such a direct implementation is available by utilizing a characteristic feature of the FMO scheme. Figures 1 and 2 illustrate the structure of our RI-MP2 gradient program. As shown in Figure 1, it is assumed that $B^{P}{}_{ia}$ can be completely stored on the main

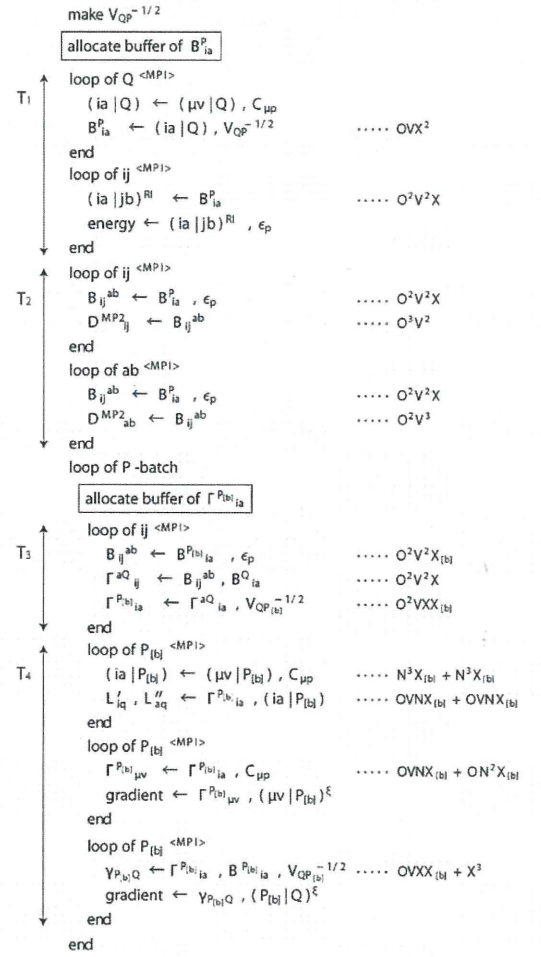


Figure 1. Outline of our RI-MP2 gradient program. (The variables and indexes in this figure correspond to the definitions in the main text or Supporting Information.) The computational costs are given using O, V, N, and X, representing the number of occupied orbital, virtual orbital, basis function, and auxiliary basis function, respectively. The superscripts ⟨MPI⟩ indicate that the loop can be parallelized with the MPI and the subscripts [b] show the length of the auxiliary basis function for each batch.

memory. At a glance, this assumption dose not appear to be realistic because the memory requirement for a large molecule is huge. However, in the FMO scheme, the calculation size is limited within the fragment pairs; thus, the memory requirement is not an issue. This is the most important benefit of the FMO scheme for this study. Although the memory requirement of $\Gamma^{P}{}_{ia}$ is similar to that of $B^{P}{}_{ia}$, a batch treatment is available (the definition of $\Gamma^{P}{}_{ia}$ is given in the Supporting Information). Figure 1 also shows the computational cost for several time-consuming parts, which are coded using a DGEMM function of the basic linear algebra subroutines (BLAS). Superscripts ⟨MPI⟩ indicate that the loop is parallelized with the message passing interface (MPI). However, in the FMO method, the parallelization of the RI-MP2 gradient calculation is not that important. This is because the parallelization with the index of a monomer or a dimer is more efficient; thus, a large amount of parallelization is not applied to a single calculation of the RI-

dx.doi.org/10.1021/jz201697x | J. Phys. Chem. Lett. 2012, 3, 375–379

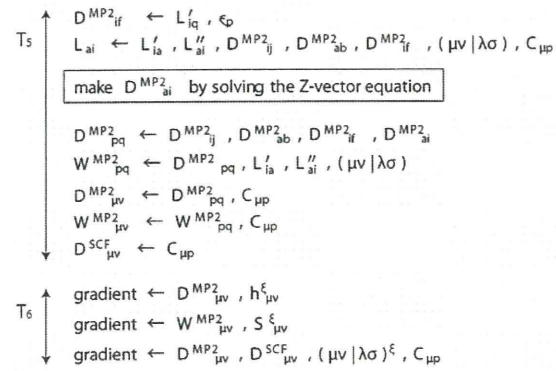


Figure 2. Outline of our RI-MP2 gradient program. (The variables and indexes in this figure correspond to the definitions in the main text or Supporting Information.)

MP2 gradient. This is an important feature of the FMO scheme. In Figure 2, the remaining parts of our program are illustrated. In our program, the so-called Z-vector equation[28] is solved with the direct inversion iterative sequence (DIIS).[29] A formulation involving the Z-vector equation is given in the Supporting Information. The time-consuming parts in Figure 2 are easily parallelized with MPI in the same manner as that in the standard SCF program. In this study, new programs were developed within PAICS.[30]

In this Letter, we will first present test calculations to check the performance of our implementation and will then present demonstrative calculations of the two biomolecular systems. A small peptide constructed with two tryptophans (Trp−Trp) was selected for the test. In a typical FMO calculation of proteins, each amino acid residue is treated as a single fragment; thus, a pair of tryptophans is one of the largest dimer. This is the reason why the Trp−Trp was selected. The structure used here is the same as that used in our previous study.[21] We employed cc-pVDZ and cc-pVTZ,[31] together with auxiliary basis functions developed by Weigend et al.[32] For comparison, we also performed gradient calculations with the canonical MP2 method using the MP2 gradient program of PAICS, in which an integral direct algorithm with a batch treatment[19] is adopted.

In Table 1, we summarize the computational times for energy and gradient of the Trp−Trp. In calculations with cc-pVDZ, timing of the HF gradient is 63.4 min, and that of the RI-MP2 gradient is 165.6 min; thus, the ratio is 2.60. In the case of cc-pVTZ, the ratio of the computational time of the RI-MP2 gradient (2431.2 min) to that of the HF gradient (1181.5 min) is 2.06. Thus, the gradient including the electron correlation is obtained with two or three times the effort of the HF gradient. The memory requirement of B^P_{ia} was 450.5 MB for cc-pVDZ and 1826.2 MB for cc-pVTZ, which is reasonable requirement for modern computer systems. Timing of the canonical MP2 gradient with cc-pVDZ is much larger, and the calculation with cc-pVTZ could not be completed. This is due to the large number of batches, i.e., 5 and 37, for cc-pVDZ and cc-pVTZ, respectively. On the other hand, the number of batches of the RI-MP2 gradient is 1 and 2 for cc-pVDZ and cc-pVTZ, respectively. This is an important advantage of the RI-MP2 gradient. In Table 2, we summarize the differences in the maximum gradient and those in the correlation energy between the RI-MP2 method and the canonical MP2 method. For the

Table 1. Computational Times (min) of Energy and Gradient for the Trp−Trp Using the HF, RI-MP2, and Canonical MP2 Methods[a]

	cc-pVDZ	cc-pVTZ
HF (eng)	44.9	594.9
HF (eng+gra)	63.6	1181.5
HF+RI-MP2 (eng)	52.6	636.0
HF+RI-MP2 (eng+gra)	165.6[b]	2431.2[c]
HF+MP2 (eng)	200.4[d]	9723.2[e]
HF+MP2 (eng+gra)	941.5[f]	[g]

[a]Here (eng) and (eng+gra) indicate the energy calculation and the energy and gradient calculations, respectively. All calculations were performed with a single core (Core2 Quad Q9650) and 4.0 GB memory per core (i.e., these calculations are not parallelized). The calculation of the canonical MP2 gradient with cc-pVTZ could not be completed because of the large number of batches. [b]Number of batch is 1. [c]Number of batch is 2. [d]Number of batch is 2. [e]Number of batch is 15. [f]Number of batch is 5. [g]Number of batch is 37.

Table 2. Correlation Energy and Maximum Gradients of the RI-MP2 and Canonical MP2 Methods[a]

		correlation energy	maximum gradient
Trp−Trp	RI-MP2	−4.125599	0.109558
	MP2	−4.125857	0.109567
	diff.	0.000258	0.000009
PrP−GN8	RI-MP2	−135.369267	0.062030
	MP2	−135.375404	0.062109
	diff.	0.006137	0.000079
HIV1−LPV	RI-MP2	−237.295825	−0.083449
	MP2	−237.309615	−0.083483
	diff.	0.013790	0.000034

[a]All values are results obtained using cc-pVDZ. Energy and gradient are presented in hartree and hartree/bohr, respectively.

Trp−Trp, the difference in the maximum gradient was only 0.000009 hartree/bohr. Thus, the RI-MP2 method can be used instead of the canonical MP2 method for gradient calculations. The time profile and parallel efficiency of our program are also given in the Supporting Information.

As illustrative examples of the RI-MP2 gradient based on the FMO method, we performed gradient calculations for two biomolecules. One is a prion protein (PrP) complexed with GN8. The PrP is a key molecule of prion diseases, and the GN8 is a potential curative agent that was developed in our laboratory.[33] The structure and fragmentation used here are the same as those used in our previous study,[21] and the total number of residues, fragments, atoms, and basis functions are 103, 106, 1729, and 16736, respectively. The other biomolecule is a human immunodeficiency virus type 1 (HIV1) protease with a lopinavir (LPV) molecule. The HIV1 protease is an important molecule in the infection of acquired immune deficiency syndrome (AIDS), and LPV is an inhibitor. Detailed information about the atomic coordinate and fragmentation can be found in a previous study.[34] The total number of residues, fragments, atoms, and basis functions are 198, 203, 3225, and 30224, respectively. In these calculations, cc-pVDZ was employed. We performed the calculations of the energies and gradients with the HF, RI-MP2, and canonical MP2 methods using 128 cores, in which parallelization of the index of a monomer or a dimer was mainly adopted.

Table 3 summarizes the computational times of the energy and gradient. In the calculation of the PrP−GN8, computa-

Table 3. Computational Times (min) of Energy and Gradient for the PrP−GN8 and HIV1−LPV Using the HF, RI-MP2, and Canonical MP2 Methods[a]

	PrP−GN8	HIV1−LPV
HF (eng)	132.7	225.1
HF (eng+gra)	293.6	660.1
HF+RI-MP2 (eng)	177.6	285.6
HF+RI-MP2 (eng+gra)	624.2[b]	1118.8[c]
HF+MP2 (eng)	546.7[d]	783.1[e]
HF+MP2 (eng+gra)	2514.2[f]	3618.1[g]

[a]In these calculations, cc-pVDZ was employed. Here (eng) and (eng +gra) indicate the energy calculation and the energy and gradient calculations, respectively. All calculations were performed with 128 cores (Xeon E5429) and a 2.0 GByte memory per core. [b]Number of batch of the largest dimer is 1. [c]Number of batch of the largest dimer is 1. [d]Number of batch of the largest dimer is 4. [e]Number of batch of the largest dimer is 4. [f]Number of batch of the largest dimer is 9. [g]Number of batch of the largest dimer is 10.

tional times of the HF and RI-MP2 gradients were 293.6 and 624.2 min, respectively; thus, the ratio was only 2.13. Similarly, the ratio between the HF and RI-MP2 methods was 1.69 in the calculation of the HIV1−LPV (660.2 min for the HF gradient and 1118.8 min for the RI−MP2 gradient). The memory requirement of B^P_{ia} for the largest dimer was 404.5 MB for PrP−GN8 and 403.9 MB for HIV1−LPV. Consequently, these calculations show that the gradient including the electron correlation can be obtained with only about twice the computational effort of the HF gradient for large molecules. The computational times of the canonical MP2 gradient were much larger than those of the RI-MP2 gradient (2514.2 and 3618.1 min for the PrP−GN8 and HIV1−LPV, respectively). The reason was similar to that of the previous test calculations, i.e., the number of batches was large for the canonical MP2 gradient. Time profiles of these FMO calculations are given in the Supporting Information. Here, we focus on the error caused by RI approximation. In Table 2, the maximum gradients of the RI-MP2 and canonical MP2 methods are presented. The differences in the maximum gradients were 0.000079 and 0.000034 hartree/bohr for the PrP−GN8 and HIV1−LPV, respectively, indicating that the error is sufficiently small to allow the use of the RI-MP2 gradient instead of the canonical MP2 gradient in the FMO method. The maximum gradients were those of the α carbon in Pro158 and the oxygen in the water molecule for PrP−GN8 and HIV1−LPV, respectively. These atoms are illustrated in the Supporting Information.

In this Letter, we report an implementation of the RI-MP2 gradient into the FMO scheme. One of the key points of this study is that the memory requirement of B^P_{ia} is not an issue because the calculation size can be limited within only a pair of the fragments. In addition, note that the parallel efficiency of the RI-MP2 gradient calculation is not as important because parallelization with the index of a monomer or a dimer is mainly adopted in the FMO scheme. Utilizing these features of the FMO method, we could easily use a direct algorithm for the RI-MP2 gradient. In the test calculations with the Trp−Trp, the RI−MP2 method showed a computational advantage over the canonical MP2 method. Finally, we performed some illustrative calculations of the two biomolecular systems, i.e., PrP−GN8 and HIV1−LPV, which were constructed with 1729 and 3225 atoms, respectively. These calculations demonstrated that the gradient including the electron correlation was obtained with only about twice the computational effort of the HF gradient. In addition, the difference between the maximum gradients in the RI-MP2 and canonical MP2 methods was negligible. Thus, we conclude that, in the context of the FMO scheme, the RI-MP2 method can be used instead of the canonical MP2 method for gradient calculations of large systems. Our next endeavor is to use our RI-MP2 gradient program for geometry optimization of biomolecules and ab initio MD simulation of condense systems including chemical reactions. In addition to the usage in theoretical studies, our program will be utilized in several fields of research, for example, a research in which determination of accurate atomic coordinates of biomolecules is required (some additional discussions are given in the Supporting Information).

■ ASSOCIATED CONTENT

Supporting Information

Theoretical details and additional data of the calculations performed here. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: kuwata@gifu-u.ac.jp.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Calculations necessary in this study were performed using the computational resource at Gifu University.

■ REFERENCES

(1) Whitten, J. L. Coulombic Potential Energy Integrals and Approximations. *J. Chem. Phys.* **1973**, *58*, 4496−4501.

(2) Dunlap, B. I.; Connolly, J. W. D.; Sabin, J. R. On Some Approximations in Applications of $X\alpha$ Theory. *J. Chem. Phys.* **1979**, *71*, 3396−4402.

(3) Vahtras, O.; Almlöf, J.; Feyereisen, M. W. Integral Approximations for LCAO-SCF Calculations. *Chem. Phys. Lett.* **1993**, *213*, 514−518.

(4) Früchtl, H. A.; Kendall, R. A.; Harrison, R. J.; Dyall, K. G. An Implementation of RI-SCF on Parallel Computers. *Int. J. Quantum Chem.* **1997**, *64*, 63−69.

(5) Feyereisen, M.; Fitzgerald, G.; Komornicki, A. Use of Approximate Integrals in Ab Initio Theory. An Application in MP2 Energy Calculations. *Chem. Phys. Lett.* **1993**, *208*, 359−363.

(6) Weigend, F.; Häser, M. RI-MP2: First Derivatives and Global Consistency. *Theor. Chem. Acc.* **1997**, *97*, 331−340.

(7) Rhee, Y. M.; DiStasio, R. A. Jr.; Lochan, R. C.; Head-Gordon, M. Analytical Gradient of Restricted Second-Order Møller−Plesset Correlation Energy with the Resolution of the Identity Approximation, Applied to the TCNE Dimer Anion Complex. *Chem. Phys. Lett.* **2006**, *426*, 197−203.

(8) Hättig, C.; Hellweg, A.; Löhn, A. Distributed Memory Parallel Implementation of Energies and Gradients for Second-Order Møller−Plesset Perturbation Theory with the Resolution-of-the-Identity Approximation. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1159−1169.

(9) Rendell, A. P.; Lee, T. J. Coupled-Cluster Theory Employing Approximate Integrals: An Approach to Avoid the Input/Output and Storage Bottlenecks. *J. Chem. Phys.* **1994**, *101*, 400−408.

(10) Kitaura, K.; Ikeo, E.; Asada, T.; Nakano, T.; Uebayasi, M. Pair Interaction Molecular Orbital Method: An Approximate Computational Method for Molecular Interactions. *Chem. Phys. Lett.* **1999**, *312*, 319−324.

(11) Kitaura, K.; Ikeo, E.; Asada, T.; Nakano, T.; Uebayasi, M. Fragment Molecular Orbital Method: An Approximate Computational Method for Large Molecules. *Chem. Phys. Lett.* **1999**, *313*, 701−706.
(12) Fedorov, D. G.; Ishida, T.; Uebayasi, M.; Kitaura, K. The Fragment Molecular Orbital Method for Geometry Optimizations of Polypeptides and Proteins. *J. Phys. Chem. A* **2007**, *111*, 2722−2732.
(13) Ishikawa, T.; Yamamoto, N.; Kuwata, K. Partial Energy Gradient Based on the Fragment Molecular Orbital Method: Application to Geometry Optimization. *Chem. Phys. Lett.* **2010**, *500*, 149−154.
(14) Fedorov, D. G.; Alexeev, Y.; Kitaura, K. Geometry Optimization of the Active Site of a Large System with the Fragment Molecular Orbital Method. *J. Phys. Chem. Lett.* **2011**, *2*, 282−288.
(15) Mochizuki, Y.; Komeiji, Y.; Ishikawa, T.; Nakano, T.; Yamataka, H. A Fully Quantum Mechanical Simulation Study on the Lowest n−π^* State of Hydrated Formaldehyde. *Chem. Phys. Lett.* **2007**, *437*, 66−72.
(16) Sato, M.; Yamataka, H.; Komeiji, Y.; Mochizuki, Y.; Ishikawa, T.; Nakano, T. How Does an SN_2 Reaction Take Place in Solution? Full Ab Initio MD Simulations for the Hydrolysis of the Methyl Diazonium Ion. *J. Am. Chem. Soc.* **2008**, *130*, 2396−2397.
(17) Komeiji, Y.; Ishikawa, T.; Mochizuki, Y.; Yamataka, H.; Nakano, T. Fragment Molecular Orbital Method-Based Molecular Dynamics (FMO-MD) as a Simulator for Chemical Reactions in Explicit Salvation. *J. Comput. Chem.* **2009**, *30*, 40−50.
(18) Komeiji, Y.; Mochizuki, Y.; Nakano, T.; Fedorov, D. G. Fragment Molecular Orbital-Based Molecular Dynamics (FMO-MD), A Quantum Simulation Tool for Large Molecular Systems. *J. Mol. Struct.* **2009**, *898*, 2−7.
(19) Mochizuki, Y.; Nakano, T.; Komeiji, Y.; Yamashita, K.; Okiyama, Y.; Yoshikawa, H.; Yamataka, H. Fragment Molecular Orbital-Based Molecular Dynamics (FMO-MD) Method with MP2 Gradient. *Chem. Phys. Lett.* **2011**, *504*, 95−99.
(20) Nagata, T.; Fedorov, D. G.; Ishimura, K.; Kitaura, K. Analytic Energy Gradient for Second-Order Møller−Plesset Perturbation Theory Based on the Fragment Molecular Orbital Method. *J. Chem. Phys.* **2011**, *135*, 044110.
(21) Ishikawa, T.; Kuwata, K. Fragment Molecular Orbital Calculation with RI-MP2 Method. *Chem. Phys. Lett.* **2009**, *474*, 195−198.
(22) Okiyama, Y.; Nakano, T.; Yamashita, K.; Mochizuki, Y.; Taguchi, N.; Tanaka, S. Acceleration of Fragment Molecular Orbital Calculations with Cholesky Decomposition Approach. *Chem. Phys. Lett.* **2010**, *490*, 84−89.
(23) Katouda, M. Application of Resolution of Identity Approximation of Second-Order Møller−Plesset Perturbation Theory to Three-Body Fragment Molecular Orbital Method. *Theor. Chem. Acc.* **2011**, *130*, 449−453.
(24) Nakano, T.; Kaminuma, T.; Sato, T.; Fukuzawa, K.; Akiyama, Y.; Uebayasi, M.; Kitaura, K. Fragment Molecular Orbital Method: Use of Approximate Electrostatic Potential. *Chem. Phys. Lett.* **2002**, *351*, 475−480.
(25) Nagata, T.; Fedorov, D. G.; Kitaura, K. Derivatives of the Approximated Electrostatic Potentials in the Fragment Molecular Orbital Method. *Chem. Phys. Lett.* **2009**, *475*, 124−131.
(26) Mochizuki, Y.; Nakano, T.; Koikegami, S.; Tanimori, S.; Abe, Y.; Nagashima, U.; Kitaura, K. A Parallelized Integral-Direct Second-Order Møller−Plesset Perturbation Theory Method with a Fragment Molecular Orbital Scheme. *Theor. Chem. Acc.* **2004**, *112*, 442−452.
(27) Fedorov, D. G.; Kitaura, K. Second Order Møller−Plesset Perturbation Theory Based upon the Fragment Molecular Orbital Method. *J. Chem. Phys.* **2004**, *121*, 2483−2490.
(28) Handy, N. C.; Schaefer, H. F. III On the Evaluation of Analytic Energy Derivatives for Correlated Wave Functions. *J. Chem. Phys.* **1984**, *81*, 5031−5033.
(29) Pulay, P. Convergence Acceleration of Iterative Sequences. The Case of SCF Iteration. *Chem. Phys. Lett.* **1980**, *73*, 393−398.
(30) Ishikawa, T.; Ishikura, T.; Kuwata, K. Theoretical Study of the Prion Protein Based on the Fragment Molecular Orbital Method. *J. Comput. Chem.* **2009**, *30*, 2594−2601.
(31) Dunning, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron Through Neon and Hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007−1023.
(32) Weigend, F.; Köhn, A.; Hättig, C. Efficient Use of the Correlation Consistent Basis Sets in Resolution of the Identity MP2 Calculations. *J. Chem. Phys.* **2002**, *116*, 3175−3183.
(33) Kuwata, K.; Nishida, N.; Matsumoto, T.; Kamatari, Y. O.; Hosokawa-Muto, J.; Kodama, K.; Nakamura, H. K.; Kimura, K.; Kawasaki, M.; Takakura, Y.; et al. Hot Spots in Prion Protein for Pathogenic Conversion. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 11921−11926.
(34) Ishikawa, T.; Mochizuki, Y.; Amarai, S.; Nakano, T.; Tokiwa, H.; Tanaka, S.; Tanaka, K. Fragment Interaction Analysis Based on Local MP2. *Theor. Chem. Acc.* **2007**, *118*, 937−945.

ELSEVIER

Institut Pasteur

Microbes and Infection 14 (2012) 740–747

www.elsevier.com/locate/micinf

Original article

Identification of a novel compound with antiviral activity against influenza A virus depending on PA subunit of viral RNA polymerase

Kazunori Yamada [a,b], Hiroko Koyama [c], Kyoji Hagiwara [a], Atsushi Ueda [a], Yutaka Sasaki [a], Shin-nosuke Kanesashi [a], Ryuki Ueno [c], Hironori K. Nakamura [d], Kazuo Kuwata [d], Kazufumi Shimizu [e], Masaaki Suzuki [f], Yoko Aida [a,b,*]

[a] *Viral Infectious Diseases Unit, RIKEN, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan*

[b] *Laboratory of Viral Infectious Diseases, Department of Medical Genome Sciences, Graduate School of Frontier Science, The University of Tokyo, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan*

[c] *Division of Regeneration and Advanced Medical Science, Gifu University Graduate School of Medicine, 1-1 Yanagido, Gifu 501-1194, Japan*

[d] *Division of Prion Research, Center for Emerging Infectious Diseases, Gifu University, 1-1 Yanagido, 1-1 Yanagido, Gifu 501-1194, Japan*

[e] *Nihon University School of Medicine, 30-1 Oyaguchi-kami-cho, Itabashi-ku, Tokyo, 173-8610, Japan*

[f] *RIKEN Center for Molecular Imaging Science, 6-7-3 Minatojima-minamimachi, Chuo-ku, Kobe, Hyogo 650-0047, Japan*

Received 19 October 2011; accepted 29 February 2012
Available online 8 March 2012

Abstract

Influenza viruses have developed resistance to current drugs, creating a need for new antiviral targets and new drugs to treat influenza virus infections. In this study, computational and experimental screening of an extensive compound library identified THC19, which was able to suppress influenza virus replication. This compound had no cytotoxic effects and did not disrupt cell cycle progression or induce apoptosis in MDCK cells as confirmed by WST-1 assays, flow cytometry analysis, and caspase-3 assays. Time-of-addition experiments showed that THC19 acts at a relatively early stage of the viral lifecycle. Subsequent mini-genome assays revealed that THC19 inhibited viral genome replication and/or transcription, suggesting that it interferes with one or more of the viral components that form the ribonucleoprotein complexes, namely polymerase basic 2 (PB2), polymerase basic 1 (PB1), polymerase acidic (PA), nucleoprotein (NP) and viral RNA. Finally, mini-genome assays where PB2, PB1, PA or NP from A/WSN/33 (H1N1) virus were replaced with those from A/Udorn/307/1972 (H3N2) virus effectively demonstrated that THC19 inhibited viral multiplication in a manner dependent upon the PA subunit. Taken together, these results suggest that influenza virus PA protein is a potential target for, and may aid the development of, novel compounds that inhibit influenza A virus replication.
© 2012 Institut Pasteur. Published by Elsevier Masson SAS. All rights reserved.

Keywords: Influenza A virus; Polymerase acidic (PA); Antiviral agents; Mini-genome assay

1. Introduction

In the face of the persistent threat posed by human influenza A infections and the possibility of a highly pathogenic avian influenza (H5N1) pandemic, there is great concern regarding the shortage of effective anti-influenza virus agents. In fact, viral M2 channel blockers and neuraminidase inhibitors, the drugs currently available for treating influenza virus infection, have problems in terms of safety and/or the emergence of viral resistance [1]. For this reason, novel anti-influenza virus agents are urgently needed.

Promising targets for novel drugs include influenza virus RNA polymerases, nucleoprotein (NP), and viral RNA, which form the viral ribonucleoprotein (vRNP) complex. The three influenza viral polymerase proteins, polymerase basic 2 (PB2), polymerase basic 1 (PB1) and polymerase acidic (PA), are mainly responsible for viral genome replication and

* Corresponding author. Viral Infectious Diseases Unit, RIKEN, 2-1 Hirosawa, Wako, Saitama 351-0198 Japan. Tel.: +81 48 462 4420; fax: +81 48 462 4399.
E-mail address: aida@riken.jp (Y. Aida).

1286-4579/$ - see front matter © 2012 Institut Pasteur. Published by Elsevier Masson SAS. All rights reserved.
doi:10.1016/j.micinf.2012.02.012

transcription [2]. Recently, a therapeutic drug against these polymerases, called favipiravir, was developed [3], which inhibits the replication and transcription of influenza viruses. A similar drug, ribavirin, is also capable of inhibiting influenza virus replication and transcription [4]. Furthermore, other anti-influenza compounds, which target the PB1 and PA protein, respectively, were recently identified [5,6]. The NP forms the major component of vRNP complexes [7]. Recently, a few novel antiviral compounds, including mycalamide analogues and nucleozin and its analogue FA-2, both targeting NP, were identified as influenza inhibitors [5,8,9]. These observations suggest that targeting components of the vRNP complexes, such as PB2, PB1, PA and NP, may provide a new strategy for influenza virus therapy and that this promising strategy may be complemented by identifying more drugs that target vRNP complexes.

Therefore, in this study, we computationally screened approximately 290,000 compounds from an extensive compound library using parameters common among anti-influenza drugs such as molecular weight, H-bond donors, H-bond acceptors and XLogP values. We then randomly selected 340 compounds and experimentally identified a compound that inhibited influenza viruses. Finally, we showed that the compound inhibited viral multiplication in a manner dependent upon the PA subunit. Taken together, these results demonstrate that the viral PA subunit may be a feasible antiviral target for the development of novel small-molecule therapies.

2. Materials and methods

2.1. Compounds

All compounds used in the first screening were purchased from ASINEX Ltd (Moscow, Russia). Compounds S48, S50, S75, S92, T-081269 and T-117976 were kind gifts from Dr. Honma (RIKEN Systems and Structural Biology Center).

The 8-Methyl-2,3,4,9-tetrahydro-1*H*-carbazole intermediate was synthesized from 1-cyclohexylidene-2-o-tolylhydrazine via the Fischer indole reaction [10]. GIF-1004 was then synthesized from the carbazole intermediate product by treating it with epichlorhydrin followed by piperidine as previously described [11]. GIF-1005, GIF-1030 and GIF-1031 were synthesized in the same manner as GIF-1004. GIF-1029 was obtained via the oxidation of THC19.

2.2. Cells and viruses

MDCK and HEK293T cells were cultured in Dulbecco's modified Eagle's medium (DMEM, Invitrogen, Carlsbad, CA) containing Pen-Strep Glutamine (Invitrogen) and 10% foetal bovine serum (FBS, Sigma–Aldrich, St. Louis, MO).

Influenza strains A/Wilson-Smith Neurotropic/1933 (H1N1, A/WSN/33) [12], A/Udorn/307/1972 (H3N2), and A/Sydney/5/1997 (H3N2) were grown in MDCK cells. The Influenza A/Puerto Rico/8/1934 (H1N1) and A/duck/Hokkaido/Vac-3/2007 (H5N1) viruses were gifts from Dr. H. Kida (Graduate School of Veterinary Medicine, Hokkaido University).

2.3. Construction of expression plasmids

RNA from the A/Udorn/307/1972 (H3N2) virus was purified from infected MDCK cells using an RNeasy Mini Kit (Qiagen, Hamburg, Germany). Reverse transcription was performed at 42 °C for 1 h using the SuperScript First-Strand Synthesis System (Invitrogen) according to the manufacturer's instructions with reverse primers (F)-*Xho*I-PB1: 5′-GGCTCGAGAGCGAAAGCAGGCAAACCATTTGAAT-3′, (F)-*Xho*I-PB2: 5′-GGCTCGAGAGCAAAAGCAGGTCAATTAT-3′, (F)-*Xho*I-PA: 5′-GGCTCGAGAGCAAAAGCAGGTACTGATTCGAGAT-3′, and (F)-*Xho*I-NP: 5′-GGCTCGAGAGCAAAAGCAGGGTTAATAATC-3′, which are complementary to the genomic RNA of PB1, PB2, PA and NP, respectively. Restriction enzyme sites are underlined. The cDNA corresponding to each genome was amplified by PCR using primers (F)-*Xho*I-PB1 and (R)-PB1-*Not*I, 5′-AAGCGGCCGCAGTAGAAACAAGGCATTTTTTCA-3′ for PB1; (F)-*Xho*I-PB2 and (R)-PB2-*Not*I, 5′-AAGCGGCCGCAGTAGAAACAAGGTCGTTTTTA-3′ for PB2; (F)-*Xho*I-PA and (R)-PA-*Not*I, 5′-AAGCGGCCGCAGTAGAAACAAGGTACTTTTTT-3′ for PA; and (F)-*Xho*I-NP and (R)-NP-*Not*I, 5′-AAGCGGCCGCAGTAGAAACAAGGGTATTTTTCC-3′ for NP, and the DNA polymerase KOD FX (Toyobo, Osaka, Japan). After purification using a MinElute PCR Purification Kit (Qiagen), the PCR products were digested with *Xho*I (CTCGAG) and *Not*I (GCGGCCGC) and cloned into the pCAGGS vector [13]. The DNA sequences were confirmed to match the previously reported sequences (GenBank ID: CY009642.1, CY009643.1, CY009641.1 and CY009639.1, respectively). The expression plasmids PB1/pCAGGS, PB2/pCAGGS, PA/pCAGGS, NP/pCAGGS and NP-luc/pPolI derived from the A/WSN/33 (H1N1) virus were a gift from Dr. Y Kawaoka, (Department of Microbiology and Immunology, Institute of Medical Science, the University of Tokyo).

2.4. Computational screening

The compound library ZINC, provided by the Shoichet Laboratory at the University of California, San Francisco, was used for screening. The parameters used to screen the compounds were as follows: molecular weight from 150 to 450 g/mol, less than three H-bond donors, less than six H-bond acceptors, and an XLogP range from −2 to 4.

2.5. Plaque assay

Six-well tissue culture plates (Techno Plastic Products AG; TPP, Trasadingen, Switzerland) were seeded with 5.0×10^5 MDCK cells/well in 10% FBS/DMEM and incubated at 37 °C for 17 h. The cells were then washed twice with 0.3% bovine serum albumin (BSA)/MEM and inoculated with approximately 200 plaque forming units (PFU) of each virus at 37 °C for 1 h. After washing three times with 0.3% BSA/MEM, the

cells were incubated with THC19 in 2 ml of 0.3% BSA/MEM containing 1% agarose and 1 μg/ml of tosylsulfonyl phenylalanyl chloromethyl ketone-treated trypsin (Worthington Biochemical Corporation, Lakewood, NJ) at 37 °C for 48 h. After 48 h, uninfected living cells were stained with 0.1% crystal violet (Wako, Osaka, Japan). The number of plaques was counted and the half-maximal inhibitory concentration (IC_{50}) was determined by plaque number reduction. In addition, the size of each plaque was measured at 0.1 mm intervals and a histogram was constructed. The IC_{50} of a compound was determined by plaque size reduction, as follows. First, an image of the plaque was converted into a matrix file of intensity values (8-bit, 0-255) followed by two-hierarchization using Scion image (Scion Corporation, Frederick, MD). Each pixel in the lower hierarchy was counted and the total area was calculated as the plaque region. The total area was then divided by the total number of plaques and the size of the plaques was calculated.

2.6. Water soluble tetrazolium salt-1 (WST-1) assay

MDCK cells (1.0×10^5 cells/well) in 10% FBS/DMEM were incubated at 37 °C for 17 h in 24-well tissue culture plates (TPP). THC19 (diluted 0–100 μM) was added to each well and the plate was incubated at 37 °C for 48 h. After the incubation, 30 μl of WST-1 reagent (Takara Bio, Otsu, Japan) was added and incubated at 37 °C for 1.5 h. Finally, 100 μl of the culture supernatant was transferred to a 96-well optical bottom plate (Thermo Fisher Scientific, Rochester, NY) and the absorbance was measured at 450 nm.

2.7. Cell cycle analysis

MDCK cells were cultured in the presence or absence of 50 μM THC19 for 48 h and harvested by trypsinization. The cells were fixed with 70% ethanol and DNA was stained with 40 μg/ml propidium iodide, whereas RNA was catabolized by 50 μg/ml RNaseA for 20 min at 37 °C. The fluorescence generated by 10,000 cells was analyzed by flow cytometry (Becton–Dickinson, Franklin Lakes, NJ) using CELL Quest software (Becton–Dickinson). The percentage of cells in the G1 phase was calculated using ModFit LT Software (Verity Software House, Topsham, ME).

2.8. Caspase-3 assay

MDCK cells were cultured in the presence or absence of 50 μM THC19 for 48 h, and caspase-3 activity was measured using the Caspase-3 Fluorometric Assay kit (BioVision, Mountain View, CA) according to the manufacturer's protocol. Briefly, the cells were washed twice with phosphate-buffered saline (PBS) followed by lysis in the provided cell lysis buffer. The supernatants were collected and 50 μl of each supernatant was mixed with an equal amount of reaction buffer containing 10 mM dithiothreitol and 5 μl of 1 mM Diethylpyrocarbonate-7-amino-4-trifluoromethyl coumarin (AFC) followed by incubation at 37 °C for 1.5 h. The AFC fluorescence (505 nm) was measured in a Wallac ARVO.SX 1420 Multilabel Counter (Perkin Elmer, Waltham, MA) and the caspase-3 activity was calculated.

2.9. Determination of the inhibitory activity of THC19

Time-of-addition experiments were performed as previously described [3]. Briefly, test medium containing 40 μM THC19 was added (at −1–1, 1–3, 3–5, 5–7, or 7–9 h post-infection) to confluent MDCK cells preliminarily inoculated with approximately 80 PFU of A/WSN/33 (H1N1) viruses for 1 h at 37 °C on 24-well tissue culture plates (TPP) in 0.3% BSA/MEM. After incubation, the compound was washed out with 0.3% BSA/MEM and the cells were incubated in fresh medium. At 13 h post-infection, culture supernatants were collected and viral production was determined using the plaque assay.

2.10. Mini-genome assay

PB1/pCAGGS, PB2/pCAGGS, PA/pCAGGS, NP/pCAGGS and NP-luc/pPolI (0.5 μg in Opti-MEMI (Invitrogen)) was transfected into HEK293T cells (2.0×10^5 cells/well in 10% FBS/DMEM) in 24-well plates, which were then incubated at 37 °C for 48 h with TransIT-293 (Takara Bio) according to the manufacturer's protocol. At 48 h post-transfection, cells were washed with PBS and lysed in Piccagene luciferase lysis buffer (Toyo Ink, Tokyo, Japan), and luciferase activity was measured with a luminometer using the Piccagene luciferase assay system (Toyo Ink).

2.11. Alignment

Alignment of the PA subunit sequences was performed using GENETYX ver. 10.0.0 (GENETYX Corporation, Tokyo, Japan).

3. Results

3.1. Screening for compounds, and the antiviral activity of THC19 and its effect on cell growth

Commonly used anti-influenza viral drugs share certain properties such as molecular weight, the number of H-bond donors and acceptors, and XLogP values (the computationally predicted partition coefficient) [14]. These parameters can be used as an indicator of whether a compound may be a viable therapeutic drug or not. Thus, these parameters were used to screen the ZINC compound library (the values are described above) in the present study. Following computational screening, 295,086 compounds were extracted and 340 compounds were randomly selected for further biological analysis. Plaque assays using influenza A/WSN/33 (H1N1) identified a single compound, THC19, as inhibiting influenza viruses with no cytotoxicity (according to the WST-1 assay and total cell counts from Trypan Blue exclusion tests; data not shown). As shown in Fig. 1A, THC19 (1-(3,4-dihydro-1*H*-

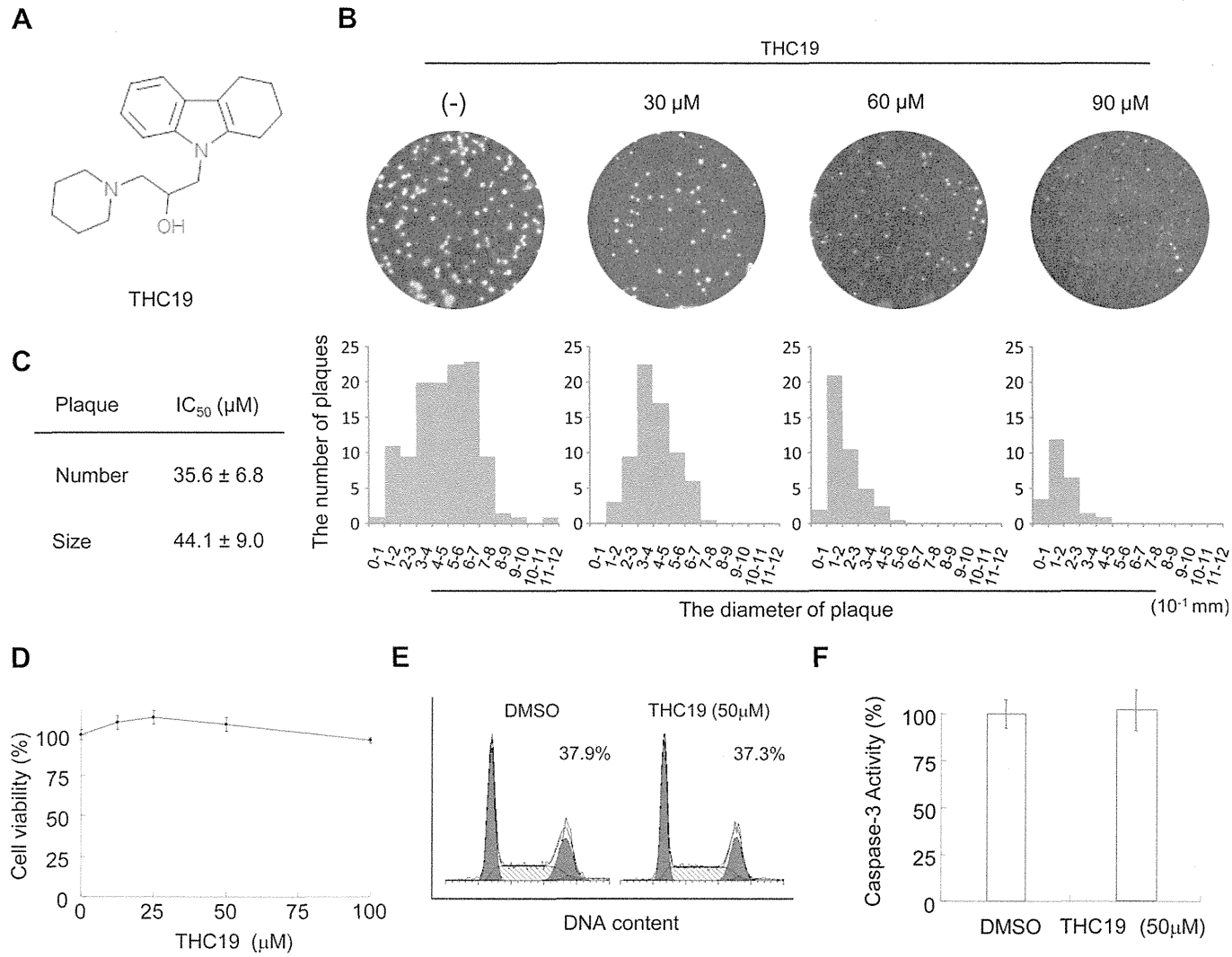


Fig. 1. Screening of compounds, and the antiviral activity of THC19 and its effect on cell growth. (A) Chemical structure of THC19. (B) The effect of THC19 on viral plaque formation by A/WSN/33 (H1N1) in MDCK cells. Upper panels show plaques formed in the absence (−) or presence of 30, 60 or 90 μM THC19. Lower panels show the number of plaques and sizes at 0.1 mm intervals. (C) IC_{50} of THC19. IC_{50} determined by plaque number reduction (Number) and plaque size reduction (Size), respectively. (D) The viability of MDCK cells treated with serially diluted THC19 was measured using the WST-1 assay system. (E) Effect of THC19 on cell cycle progression. Cell cycle profiles of MDCK cells treated with or without 50 μM THC19 were analyzed by flow cytometry. The percentage of cells in G1 phase was calculated (shown on the right of each panel). (F) The effect of THC19 on apoptosis. Caspase-3 activity in MDCK cells treated with or without 50 μM THC19 was determined using a caspase-3/CPP32 fluorometric assay kit. DMSO was used as a negative control. Values represent the mean ± SD of two independent experiments; each experiment was performed in triplicate.

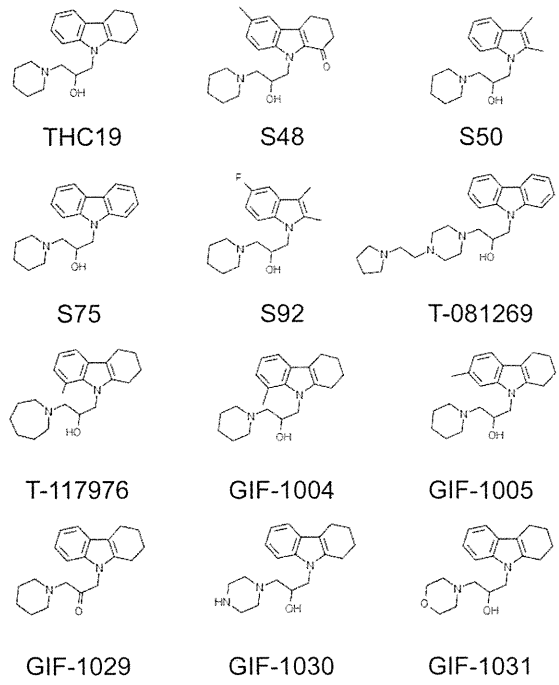
carbazol-9(2H)-yl)-3-(piperidin-1-yl)propan-2-ol) is a derivative of 1,2,3,4-tetrahydrocarbazole piperidine. In the first screen (the plaque assay), THC19 inhibited viral multiplication by approximately 70% at 100 μM. Subsequent plaque assays showed that THC19 inhibited viral production in a dose-dependent manner as determined by both the number (Fig. 1B; upper panel) and size (Fig. 1B; lower panel) of the plaques, with an IC_{50} of 35.6 μM for plaque number reduction and 44.1 μM for plaque size reduction (Fig. 1C).

The cytotoxicity of THC19 was then examined in MDCK cells using the WST-1 assay. As shown in Fig. 1D, cell viability was not reduced at any concentration up to 100 μM; even at 200 μM no cytotoxicity was observed. Therefore, the 50% cytotoxic concentration (CC_{50}) was >200 μM. In addition, THC19 had no effect on host cell cycle progression as assayed by flow cytometry (Fig. 1E), or on apoptosis as determined by caspase-3 activity (Fig. 1F). Thus, it appears that THC19 inhibits viral replication by affecting the influenza virus directly, rather than through non-specific effects on the infected cell.

3.2. Antiviral effects of THC19 derivatives

To identify compounds that inhibit influenza virus replication even more effectively, we next examined the antiviral effects of THC19 analogues (S48, S50, S75, S92, T-081269 and T-117976; Fig. 2A and B). The results showed that the introduction of substitutions into the 1,2,3,4-tetrahydrocarbazole ring of THC19 (S48, CH3 and ketone; S92, F; T-117976, Cl)

A



B

Derivative	IC_{50} (μM)	CC_{50} (μM)
THC19	35.6 ± 6.8	> 200
S48	> 50	> 200
S50	> 50	> 200
S75	> 50	> 200
S92	> 50	> 200
T-081269	> 50	45.2 ± 3.1
T-117976	> 50	49.6 ± 7.8
GIF-1004	> 50	162.9 ± 7.8
GIF-1005	> 50	189.7 ± 3.1
GIF-1029	> 50	> 200
GIF-1030	> 50	155.0 ± 9.0
GIF-1031	> 50	> 200

Fig. 2. Antiviral effect of THC19 derivatives. (A) Chemical structures of the THC19 analogues. (B) IC_{50} and CC_{50} of THC19 analogues against the influenza A/WSN/33 (H1N1) virus in MDCK cells. Values represent the mean ± SD of two independent experiments; each experiment was performed in triplicate.

and/or the conversion of the 1,2,3,4-tetrahydrocarbazole ring to indole or carbazole (S50 and S92, indole; S75 and T-081269, carbazole) reduced the inhibitory effects (IC_{50} > 50 μM). Notably, the extension of the side chain via piperidine (T-081269) or ring expansion (T-117976) resulted in cellular cytotoxicity (CC_{50} = 40.5 and 41.4 μM, respectively). Next, derivatives of THC19 were produced to identify the key structures required for inhibition of the influenza virus. Because the four analogues described above contained more than two points of modification, we synthesized derivatives of THC19 containing only a single modification (GIF-1004, GIF-1005, GIF-1029, GIF-1030, and GIF-1031; Fig. 2A). As shown in Fig. 2B, substitutions within the 1,2,3,4-tetrahydrocarbazole ring of THC19 resulted in reduced inhibitory activity (IC_{50} > 50 μM), even when the substitution was only a single CH_3 modification. Replacement of the hydroxy group by a hydrogen bond acceptor (keto group; GIF-1029) resulted in decreased activity (IC_{50} > 50 μM). Piperazine and morpholine analogues also showed no activity (IC_{50} > 50 μM). These results demonstrated that the 1,2,3,4-tetrahydrocarbazole ring, the propanol spacer, and the piperidine moiety of THC19 are necessary for antiviral activity without cytotoxicity.

3.3. Stage of the viral lifecycle targeted by THC19

To determine the point of action of THC19 during the influenza virus lifecycle, a time-of-addition experiment was conducted following the scheme illustrated in Fig. 3A. As shown in Fig. 3B, THC19 interfered with viral replication (approximately 60% reduction) when added before 1 h post-infection, and yielded a 40% reduction when added up to 3 h post-infection. When THC19 was added later than 3 h post-infection, it had no inhibitory effect on the virus. These results suggest that THC19 acts at a relatively early stage of the viral lifecycle rather than during the later stages (virion assembly or viral budding).

Although there were some candidates for the most likely early stage events affected by THC19, such as viral adsorption, uncoating, or replication/transcription, we first explored the possibility that THC19 affected viral genome replication and/or transcription. Therefore, the effect of THC19 on vRNP activity was examined using a mini-genome assay incorporating five expression plasmids for influenza A/WSN/33 (H1N1): PB1/pCAGGS, PB2/pCAGGS, PA/pCAGGS and NP/pCAGGS, which control replication and transcription of the viral genome, and the luciferase-containing plasmid NP-luc/pPolI, which encodes a viral-like genome. As shown in Fig. 3C, treatment with THC19 decreased luciferase activity in a dose-dependent manner; luciferase activity would normally be enhanced by the vRNP during the replication and transcription stages of the viral lifecycle. This result indicates that THC19 inhibits viral genome replication and/or transcription. Also, haemagglutination assays confirmed that THC19 did not affect viral adsorption; viruses pre-incubated with THC19 adsorbed onto MDCK cells to the same degree as control viruses (data not shown).

3.4. Antiviral effect of THC19 on various strains of influenza virus and identification of the target protein

To identify the antiviral effects of THC19 against other strains of influenza, plaque assays were performed using influenza strains A/WSN/33 (H1N1), A/Puerto Rico/8/1934

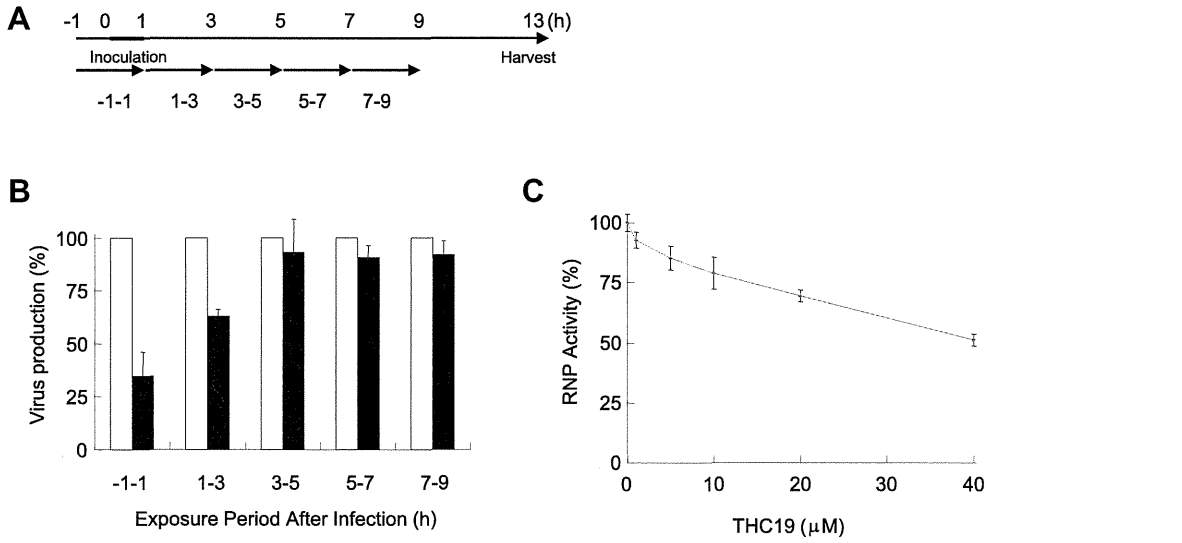


Fig. 3. Stage of the viral life cycle targeted by THC19. (A) THC19 was added at the indicated times and cell supernatants were collected at 13 h post-infection. (B) Viral yields were determined at 13 h post-infection by plaque assays. Open columns represent the mean number of plaques for control cells (0.8% w/w DMSO); closed columns represent the mean number of plaques for cells treated with THC19 (40 μM). (C) Mini-genome assay. HEK293T cells were transfected with plasmids containing PB1, PB2, PA, or NP derived from the influenza A/WSN/33 (H1N1) and NP-luc/pPolI. THC19 (0–40 μM) was added to the cells. The inhibitory effect of THC19 on replication was determined by measuring the reduction in luciferase activity. Values represent the mean ± SD of two independent experiments; each experiment was performed in triplicate.

(H1N1), A/Udorn/307/1972 (H3N2), A/Sydney/5/1997 (H3N2) and A/duck/Hokkaido/Vac-3/2007 (H5N1). As shown in Fig. 4A, THC19 inhibited multiplication of both of the H1N1 viruses, A/Sydney/5/1997 (H3N2), and the A/duck/Hokkaido/Vac-3/2007 (H5N1) virus, but surprisingly did not inhibit the A/Udorn/307/1972 (H3N2) virus.

To examine why THC19 did not affect multiplication of the A/Udorn/307/1972 (H3N2) strain, a modified mini-genome assay was performed. First, a normal mini-genome assay was performed with PB1/pCAGGS, PB2/pCAGGS, PA/pCAGGS and NP/pCAGGS derived from A/Udorn/307/1972 (H3N2). As expected, treatment with THC19 hardly impaired the luciferase activity generated by A/Udorn/3071972 (H3N2)-derived components (Fig. 4B; first column). By contrast, THC19 decreased the luciferase activity generated by A/WSN/33-derived components by more than 40% (Fig. 4B; second column), similar to the results shown in Fig. 3C. Next, to identify the viral protein affected by THC19, PB1, PB2, PA and NP expression plasmids derived from A/Udorn/307/1972 (H3N2) were exchanged for their A/WSN/33 (H1N1) counterparts and subjected to a mini-genome assay (modified mini-genome assay). In the modified assay, if treatment with THC19 decreased the level of luciferase activity in cells containing only one foreign plasmid, it could be inferred that the foreign protein was affected by THC19. Throughout the experiment, we confirmed that the replacement of a plasmid derived from A/Udorn/307/1972 (H3N2) with its counterpart derived from A/WSN/33 (H1N1) did not affect normal vRNP activity. The results showed that when the PB1, PB2 or NP A/Udorn/307/1972 (H3N2) virus plasmids were replaced with corresponding plasmids derived from A/WSN/33 (H1N1), THC19 had no effect on luciferase activity. However, replacement of PA resulted in a reduction in luciferase activity to the same degree as that in A/WSN/33 (Fig. 4B; second and fourth columns). Taken together, although THC19 had no effect on vRNP activity in the A/Udorn/307/1972 (H3N2) virus, when PA was exchanged with that from A/WSN/33 (H1N1), THC19 recovered its ability to inhibit the virus, indicating that THC19 inhibits viral genome replication and/or transcription in a manner dependent upon the PA protein in A/WSN/33 (H1N1).

Thus, to identify the amino acid differences between these groups and to further understand the mechanism of action of THC19, we compared the coding region sequence of PA protein from the compound-ineffective strain (A/Udorn/307/1972 (H3N2)) with that of the compound-effective strains. As shown in Fig. 4C, there were only two amino acid differences, namely codon position 44 and 165, which encoded Ile and Asp at positions 44 and 165, respectively, in A/Udorn/307/1972 (H3N2). By contrast, A/WSN/33 (H1N1), A/Puerto Rico/8/1934 (H1N1), A/Sydney/5/1997 (H3N2) and A/duck/Hokkaido/Vac-3/2007 (H5N1) encoded Val and Glu at positions 44 and 165, respectively.

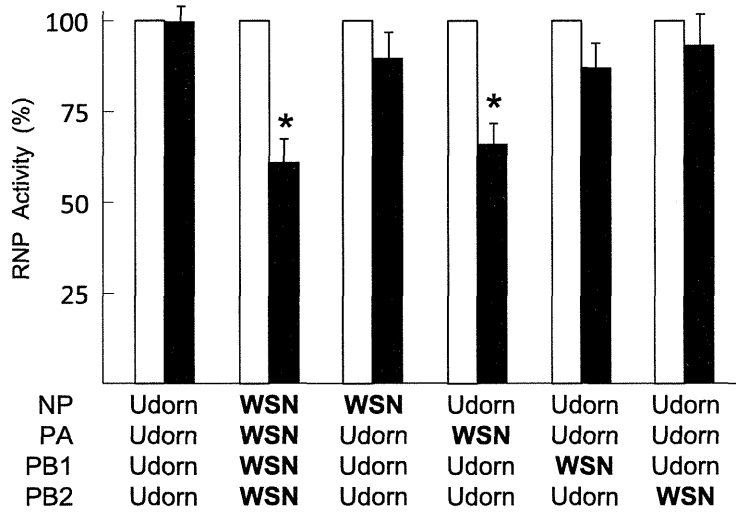
4. Discussion

In this study, we identified a compound that inhibited viral multiplication in a manner dependent upon the viral PA subunit. PA is a component of influenza virus RNA polymerase and plays an important role in both the transcription and replication of the RNA genome [15]. To date, siRNA targeting of PA has been shown to specifically inhibit the replication of the highly pathogenic avian influenza A H5N1 virus [16], and novel antiviral compounds, which may inhibit PA, have been identified [6]. In combination with these reports, the identification in the present study of a compound

A

Strain	IC_{50} (μM)
A/Wilson-Smith N/1933 (H1N1)	35.6 ± 8.8
A/Puerto Rico/8/1934 (H1N1)	30.7 ± 2.5
A/Udorn/307/1972 (H3N2)	> 90
A/Sydney/5/1997 (H3N2)	44.7 ± 1.0
A/duck/Hokkaido/Vac-3/2007 (H5N1)	39.8 ± 5.5

B



C

Strain	Amino acid position	
	44	165
A/Udorn/307/1972 (H3N2)	I	D
A/Wilson-Smith N/1933 (H1N1)	V	E
A/Puerto Rico/8/1934 (H1N1)	V	E
A/Sydney/5/1997 (H3N2)	V	E
A/duck/Hokkaido/Vac-3/2007 (H5N1)	V	E

Fig. 4. Antiviral effect of THC19 on various strains of influenza virus and the identification of the target protein of THC19. (A) MDCK cells were infected with viruses in the presence of THC19 (0–40 μM), and the IC_{50} of each strain was determined by plaque assay. (B) Effect of antiviral activity on vRNP activity assessed using a modified mini-genome assay. A single gene from the A/Udorn/307/1972 (H3N2) virus was replaced with its counterpart from the A/WSN/33 (H1N1) virus and the inhibitory effect of THC19 (40 μM) was determined. The components of the mini-genome assay are represented below the horizontal axis (WSN or Udorn). Open and closed columns represent vRNP activity in the presence of DMSO or THC19, respectively. Values are presented as the mean ± SD of two independent experiments, and each experiment was performed in triplicate. *Results differ significantly from the control (Dunnett's test ($P < 0.01$)). (C) Comparison of amino acid residues at codon positions 44 and 165 in the THC19-effective strain A/Udorn/307/1972 (H3N2) with those in the -ineffective strains.

that inhibits viruses in a PA subunit-dependent manner strongly supports the notion that PA may be a target for antiviral drugs.

PA binds to PB1 to form the vRNP complex with NP, PB2 subunit and vRNA. Although the vRNP complex is transported from the cytoplasm to the nucleus via its interaction with the cellular karyopherin nuclear import signal (NLS), the PA subunit also possesses its own NLS. Also, the PA subunit is involved in multiple functions of the polymerase, including protein stability, endonuclease activity, cap binding, and promoter binding [17]. Thus, THC19 might inhibit viral multiplication via one and/or more of the mechanisms described above. As shown in Fig. 3B, THC19 inhibited viral multiplication at a relatively early phase of the viral lifecycle. Given that mRNA levels generally reach a plateau in infected cells a little faster than those of vRNA, THC19 might inhibit