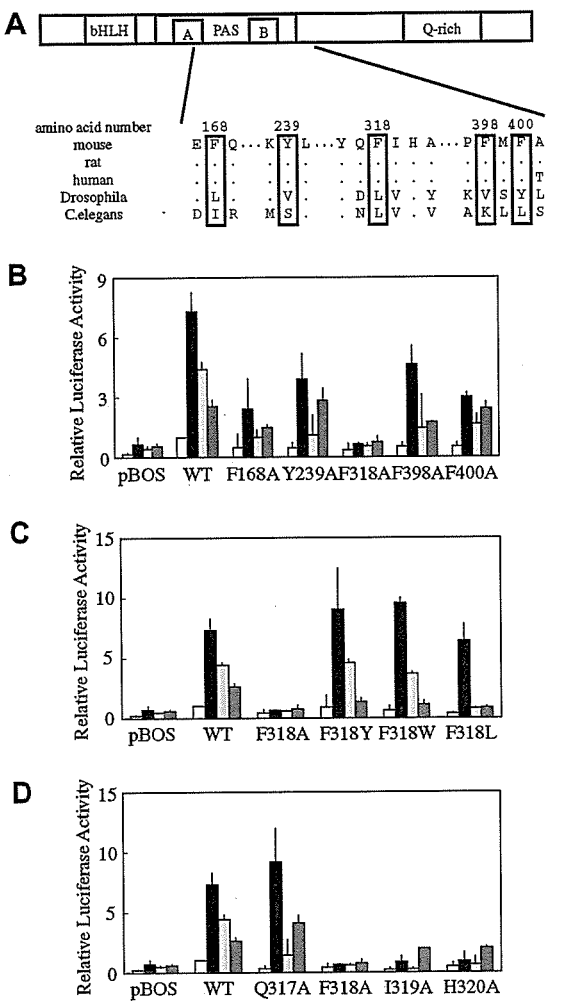


Fig. 1. Transactivation activity of the AhR and its mutants. (A) Alignment of the PAS-B sequences of vertebrate and invertebrate AhR. Structure of the mouse AhR is schematically shown above. Aromatic amino acid residues that were conserved within and around the PAS-B domain of mouse, human, and rat AhRs and that were not conserved in the domain of *Drosophila* and *C. elegans* AhRs were boxed. Dots indicate the same amino acids as those of the mouse AhR. (B) Transactivation activity of the AhRs with mutation of selected amino acids shown in (A). Selected aromatic amino acids were mutated to Ala, and cotransfected into HeLa cells with a reporter plasmid. Four hours after transfection, MC (1 μM), β-NF (0.5 μM), or TCDD (10 nM) were added to the culture medium and cells were further incubated for 40 h. Luciferase activity driven by the AhRs is shown. The values represent means ± SD of at least three separate determinations, and were normalized using the value(s) of wild type AhR treated with DMSO. Open bars, DMSO (vehicle); filled bars, MC; light gray bars, β-NF; dark gray bars, TCDD. (C,D) Transactivation activity of the AhRs with mutations of Phe318 to Ala, Tyr, Trp, and Leu, and transactivation activity of the AhRs with mutation of amino acids neighboring Phe318. Experimental procedures are shown in (B).

probably because the aromatic nature of the side chain was preserved, although induction by TCDD was weak. Mutation to Leu showed a luciferase activity in response to MC with an induction ratio similar to that of wild type, although induced activity was somewhat lower than that of wild type. Interestingly, this mutant exhibited no induction of luciferase activity by the addition of β-NF or TCDD, indicating that this mutation caused a ligand-binding specificity different from the wild type and suggesting that Phe318 may have contact with ligands. Three amino acid residues neighboring Phe318 were changed to Ala. The mutation of Gln317 had no effect on activity although induction by β-NF was weak (Fig. 1D). The mutation of Ile319 or His320 resulted in complete loss of activity, suggesting that these two amino acids also play an important role for ligand binding.

### *Nuclear translocation of AhRs in response to inducers*

Chimeric proteins of mutated AhRs fused to YFP were expressed in CHO-K1 cells, and subcellular localization of the chimeric proteins was observed. These chimeric proteins were evenly expressed and showed transactivation activity similar to the AhR without YFP tag (data not shown). Fluorescence from the YFP moiety of the wild-type AhR fusion protein was diffused over the cell, and treatment of cells with MC caused accumulation of the signal in the nucleus as shown in Fig. 2. Approximately 50% of the fluorescent cells showed nuclear localization at the maximal concentration of MC. The nuclear accumulation was accomplished within 2 h and dependent on the concentration of MC. Nuclear translocation was also observed by the addition of β-NF, although the rate of the translocation was slow, and 4 h was required for completion. The reason is not clear as to why nuclear localization of expressed AhR–YFP did not occur in all fluorescent cells even at high concentrations of inducers. Nuclear localization of mutant AhR(Phe318Ala) was similarly examined. When neither MC nor β-NF was added, nuclear accumulation of the mutant did not occur. Mutant AhR(Phe318Leu) was translocated into the nucleus by the addition of MC similar to the level of the wild type. However, the mutant remained in the cytosol with the addition of β-NF, in accordance with the result of transactivation activity of the mutant. Taken together, these results strongly suggest that stimulus-dependent nuclear localization of mutated AhRs is the causal event for their transactivation activity.

### *Ligand-binding activity of mutated AhRs*

*In vitro* binding activity of mutated AhRs to [$^3$H]-labeled MC was examined using cytosolic extracts of COS-7 cells transfected with expression plasmids for the AhRs. A clear peak at around the 9S position of the [$^3$H]MC-AhR complex appeared in the glycerol gradient as shown in Fig. 3A. This binding of the radioactive ligand was competed out with 22 times the molar excess of unlabeled MC or β-NF. A similar binding signal was also observed when cytosol containing AhR(Phe318Leu) was used. This signal was competed with unlabeled MC. However, unlabeled β-NF could not compete with the [$^3$H]MC bound to the mutated AhR, reflecting the results of transactivation and

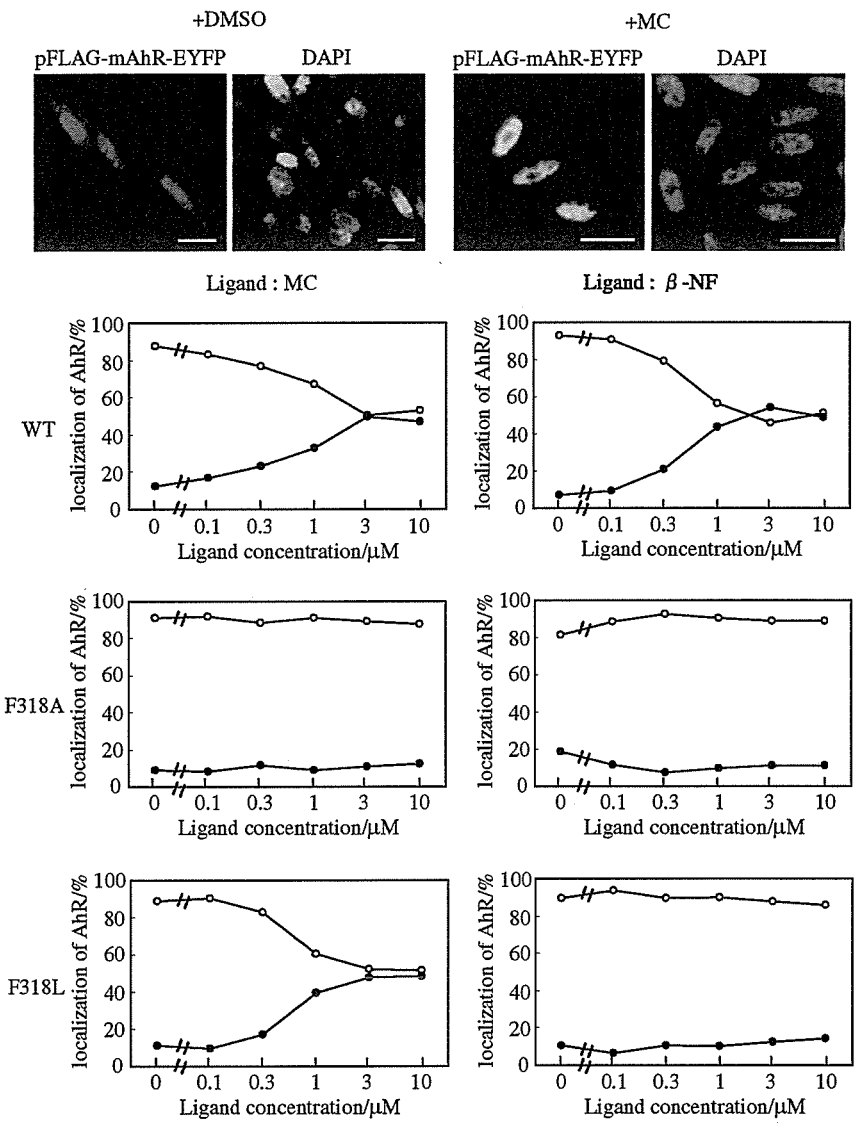


Fig. 2. Nuclear translocation of the AhR and its mutants. CHO-K1 cells were transfected with AhR–YFP-fusion plasmids, and treated with inducers for 2 h (MC) or 4 h (β-NF) before fixation in 4% paraformaldehyde. The fixed cells were counterstained with DAPI. Typical images of fluorescent cells after treatment with DMSO (vehicle) and MC are shown above. Scale bars, 20 μm. Approximately 300 cells were randomly selected, and percentages of cells with only nuclear localization (shown by closed circles) and cells with both nuclear and cytosolic localization (shown by open circles) of the chimeric protein are shown.

nuclear translocation experiments. When cytosol fraction containing expressed AhR(Phe318Ala) was used, no signal of ligand binding was detected. The mutated AhRs were evenly expressed in COS-7 cells as shown in Fig. 3B.

*Modeling the AhR ligand-binding domain*

Since the three-dimensional (3D) structure of AhR has not been elucidated so far, and previously reported 3D models of the ligand-binding domain have failed to identify Phe318 as a ligand-recognition amino acid [16], we concentrated on the active site of AhR, and obtained a 3D model for it using comparative modeling techniques. A combined FASTA and PSI-BLAST search of the protein data bank (PDB) [17] reveals a high number of matches between mouse AhR PAS-B and other PAS proteins, including HLF (HIF-2α), several histidine kinases, and other light receptors as well as sensor proteins (oxygen/redox sensors) and ion channels (data not shown). Sequences such as HLF, PHY3, HERG, FixL, EC Dos, and PAS kinase were found that were based on a moderate sequence similarity, characterized by $E$ values of less than $10^{-3}$. Fig. 4A illustrates the multiple sequence alignment of AhR PAS-B with these sequences. A further alignment of the secondary structure predicted for AhR PAS-B and the secondary structures for the 3D structures extracted from PDB

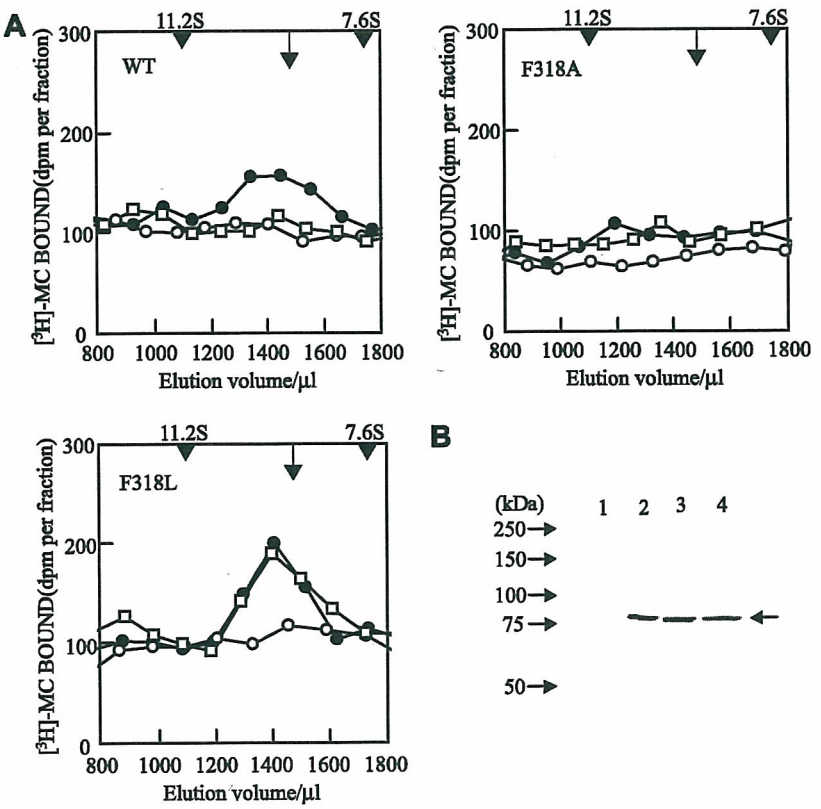


Fig. 3. Ligand-binding activity of the AhR and its mutants. (A) Binding of mutant AhRs to [$^3$H]-labeled MC. AhRs were expressed in COS-7 cells and cytosolic extracts were prepared. Glycerol gradient centrifugation was performed as described under Materials and methods. Fractions (about 100 μl) were collected and radioactivity was counted on a liquid scintillation counter. Catalase (11.2S) and fibrinogen (7.6S) were used as size markers. Arrows show the position of 9S. Filled circle, [$^3$H]MC with DMSO; open circle, [$^3$H]MC with unlabeled MC; open square, [$^3$H]MC with unlabeled β-NF. (B) Expression of mutant AhRs. Western blotting analysis using 20 μg of cell extracts was performed by the ECL plus Western blotting detection system kit. Lane 1, whole cell extracts without transfection; lanes 2–4, cytosolic extracts of cells transfected with expression plasmid for AhR, AhR(Phe318Ala), and AhR(Phe318Leu), respectively. An arrow shows the bands of the AhR.

indicate that the fold of HLF is the optimal template on which to model AhR PAS-B (Fig. 4B). The threading process of the AhR sequence into the template was performed using Swiss PDB viewer (spdbv) software. The structure was further minimized using the GROMOS force field embedded in spdbv to optimize the position of the lateral chains of the amino acids constituting the receptor.

Assisted by the docking module in Cerius$^2$, we first mapped plausible ligand-binding pockets for the model of AhR PAS-B. A unique deep cavity was recognized by the system, the boundaries of which are constituted by the amino acids in Table 1. The model was then used to dock three ligand molecules, MC, β-NF, and TCDD. The docking process was also performed using the Cerius$^2$ software. Orientations for the ligands within the binding pocket ranked as the highest by the docking software were further minimized so as to obtain reliable 3D structures for the receptor–ligand complexes. Amino acids in contact with the ligand are identified by computing the fraction of SASA (solvent accessible surface area) buried by each of the amino acids on the ligand. We performed this calculation for β-NF, and Table 1 illustrates the decrement in SASA of β-NF when docked to the cavity of the model of AhR by each of the amino acids composing the cavity. The SASA is computed using Richards' algorithm [18], and a radius for the solvent molecule (water) of 1.4 Å. The buried SASA is calculated as the difference of the SASA of β-NF at the isolated state minus the SASA of β-NF when it is in contact with each of the amino acids listed in Table 1. Docking models of the complex of AhR with TCDD, MC or β-NF, show their extensive contact with Phe318, Ile319, and His320 (Fig. 4C). Extensive hydrophobic interaction can also be observed with Ala328, Met342, Leu347, and Leu348. Two lysine residues (Lys284 and Lys286) suggest the formation of hydrogen bonds with the oxygens on the aromatic rings of β-NF.

## Discussion

From the results shown in Fig. 1, it is suggested that Phe318 plays a critical role in ligand binding to AhR. The importance of this amino acid is also demonstrated by the complex model of AhR PAS-B that we built by comparative modeling and docking simulations. The decrement

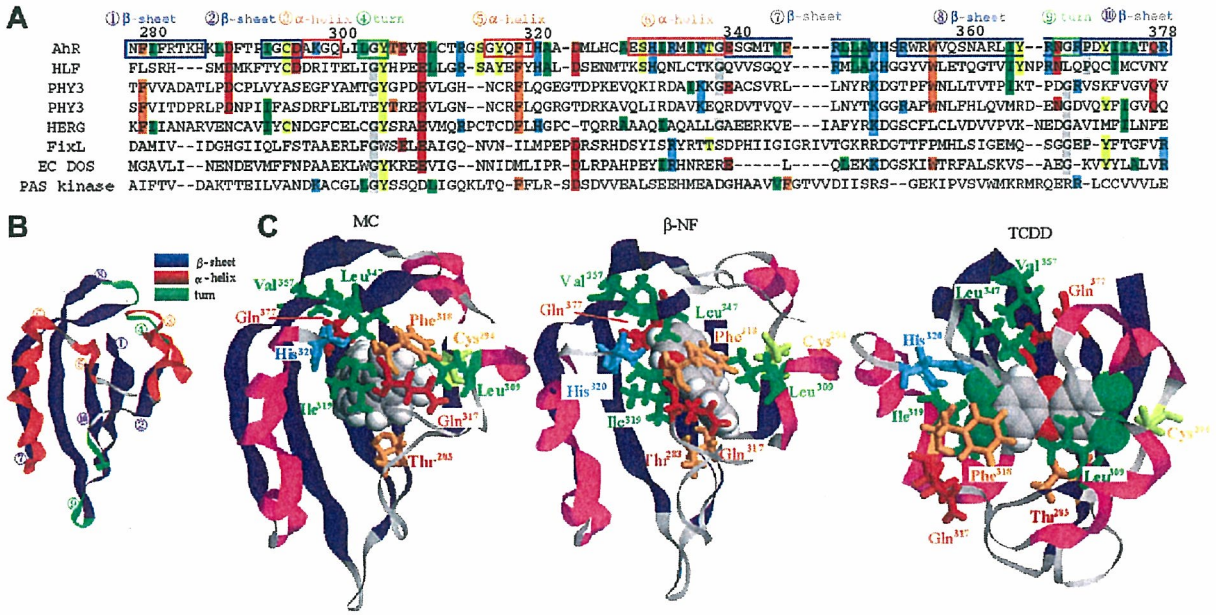


Fig. 4. Model of the AhR PAS-B domain and contacts between ligands and amino acid residues in the binding pocket. (A) Sequence alignment of AhR-related sequences. Sequences of Mouse AhR, human HLF (HIF-2α), PHY3 from *Chlamydomonas reinhardtii* (upper sequence) and *Adiantum capillus-veneris* (lower sequence), human HERG, *Bradyrhizobium japonicus* FixL, *Escherichia coli* EC DOS and human PAS kinase were aligned. Secondary structure regions are shown above the sequences. (B) Ribbon-style drawing of the AhR PAS-B domain. Numbers show the secondary structure regions as shown in (A). (C) Model of MC, β-NF, and TCDD in the binding site of AhR PAS-B domain. Secondary structure elements are colored blue (strands) and red (helices).

Table 1
Calculated buried SASA of β-NF by amino acids in the binding region of AhR

| Amino acid | Difference in SASA (in Å$^2$) for β-NF |
|---|---|
| THR283 | 0.07 |
| LYS284 | 73.37 |
| LYS286 | 49.73 |
| CYS294 | 0.07 |
| GLN299 | 72.54 |
| GLY303 | 47.77 |
| TYR304 | 29.22 |
| LEU309 | 0.07 |
| CYS310 | 27.07 |
| GLY315 | 46.93 |
| GLN317 | 0.07 |
| PHE318 | 84.37 |
| ILE319 | 77.81 |
| HIS320 | 60.62 |
| ALA328 | 81.26 |
| SER330 | 43.35 |
| MET342 | 73.18 |
| LEU347 | 88.59 |
| LEU348 | 87.90 |
| ALA349 | 42.20 |
| VAL357 | 37.66 |
| SER359 | 61.20 |
| ALA375 | 0.00 |
| GLN377 | 41.40 |

in the SASA of β-NF due to contact with Phe318 is large (about 84 Å$^2$, Table 1), suggesting that extensive interaction exists between the residue and the ligand. Mutation of Phe318 by Leu completely eliminated responsiveness of the mutant towards β-NF, although responsiveness to MC remained unchanged. Binding experiments using [$^3$H]MC corroborate the geometry of our modeled complex, since the location of Phe318 at the ligand-binding surface of the receptor plays a key role in ligand-binding specificity. This is clearly shown by the mutations performed on Phe318 that caused changes in the binding activity of the receptor. Moreover, two amino acids, Ile319 and His320, neighboring Phe318, were both found to play an essential role in the binding activity of the receptor. Although contact surface areas of these amino acids with the docked ligands are less than that of Phe318, their relevance in the binding affinity of the receptor cannot be neglected. In fact, Ile319 has been reported to be important in avian AhRs [19]. The Ile residue (Ile324 corresponding to Ile319 in the mouse AhR) of the chicken AhR that showed high affinity towards TCDD was changed to Val in tern AhR that exhibited low TCDD-binding activity. As shown in Fig. 1D, a large decrement in the activity of the receptor was observed when they were mutated to Ala. Therefore, the importance of these two amino acids in the binding of the ligands can be rationalized in terms of their bulky lateral chains and polar characteristics. The imidazole group in His confers it a polar property absent in Ala while Ile has a larger chain and is more strongly hydrophobic. Thus, mutations of these amino acids have strong repercussions in the activity of the receptor. Ala375, whose allelic mutation was demonstrated to be

responsible for the different ligand-binding affinity between C57BL/6 and DBA/2 mouse strains [8] was also exposed into the ligand-binding pocket, although it does not appear to be in direct contact with β-NF in the model. Further mutagenesis studies are necessary to confirm the amino acids composing ligand-binding domain of the AhR.

## Acknowledgments

This work was supported in part by Grant-in-Aid for research from the Ministry of Education, Culture, Sports, Science and Technology of Japan, and by funds for Research for the Future Program of JSPS.

## References

[1] J. Mimura, Y. Fujii-Kuriyama, Molecular mechanisms of AhR functions in the regulation of cytochrome P450 genes, Biochem. Biophys. Res. Commun. 338 (2005) 311–317.

[2] L. Bergander, N. Wahlstrom, T. Alsberg, J. Bergman, A. Rannug, U. Rannug, Characterization of in vitro metabolites of the aryl hydrocarbon receptor ligand 6-formylindolo[3,2-*b*]carbazole by liquid chromatography-mass spectrometry and NMR, Drug Metab. Dispos. 31 (2003) 233–241.

[3] E.M. Dioum, J. Rutter, J.R. Tuckerman, G. Gonzalez, M.A. Gilles-Gonzalez, S.L. Mcknight, NPAS2: a gas-responsive transcription factor, Science 298 (2002) 2385–2387.

[4] K.M. Dolwick, H.I. Swanson, C.A. Bradfield, In vitro analysis of Ah receptor domains involved in ligand-activated DNA recognition, Proc. Natl. Acad. Sci. USA 90 (1993) 8566–8570.

[5] J.A. Powell-Coffman, C.A. Bradfield, W.B. Wood, *Caenorhabditis elegans* orthologs of the aryl hydrocarbon receptor and its heterodimerization partner the aryl hydrocarbon receptor nuclear translocator, Proc. Natl. Acad. Sci. USA 95 (1998) 2844–2849.

[6] R.B. Emmons, D. Duncan, P.A. Estes, P. Kiefel, J.T. Mosher, M. Sonnenfeld, M.P. Ward, I. Duncan, S.T. Crews, The spineless-aristapedia and tango bHLH-PAS proteins interact to control antennal and tarsal development in *Drosophila*, Development 126 (1999) 3937–3945.

[7] J. Mimura, M. Ema, K. Sogawa, Y. Fujii-Kuriyama, Identification of a novel mechanism of regulation of Ah (dioxin) receptor function, Genes Dev. 13 (1999) 20–25.

[8] M. Ema, N. Ohe, M. Suzuki, J. Mimura, K. Sogawa, S. Ikawa, Y. Fujii-Kuriyama, Dioxin binding activities of polymorphic forms of mouse and human arylhydrocarbon receptors, J. Biol. Chem. 269 (1994) 27337–27343.

[9] W. Gong, B. Hao, S.S. Mansy, G. Gonzalez, M. Gilles-Gonzalez, M.K. Chan, Structure of a biological oxygen sensor: a new mechanism for heme-driven signal transduction, Proc. Natl. Acad. Sci. USA 95 (1998) 15177–15182.

[10] J.H.M. Cabral, A. Lee, S.L. Cohen, B.T. Chait, M. Li, R. Mackinnon, Crystal structure and functional analysis of the HERG potassium channel N terminus: a eukaryotic PAS domain, Cell 95 (1998) 649–655.

[11] S. Crosson, K. Moffat, Photoexcited structure of a plant photoreceptor domain reveals a light-driven molecular switch, Plant Cell 14 (2002) 1067–1075.

[12] T. Kinoshita, M. Doi, N. Suetsugu, T. Kagawa, M. Wada, K. Shimazaki, Phot1 and phot2 mediate blue light regulation of stomatal opening, Nature 414 (2001) 656–660.

[13] H. Kurokawa, D.S. Lee, M. Watanabe, I. Sagami, B. Mikami, C.S. Raman, T. Shimizu, A redox-controlled molecular switch revealed by the crystal structure of a bacterial heme PAS sensor, J. Biol. Chem. 279 (2004) 20186–201893.

[14] P.J. Erbel, P.B. Card, O. Karakuzu, R.K. Bruick, K.H. Gardner, Structural basis for PAS domain heterodimerization in the basic helix–loop–helix-PAS transcription factor hypoxia-inducible factor, Proc. Natl. Acad. Sci. USA 100 (2003) 15504–15509.

[15] J. Rutter, C.H. Michnoff, S.M. Harper, K.H. Gardner, S.L. McKnight, PAS kinase: an evolutionarily conserved PAS domain-regulated serine/threonine kinase, Proc. Natl. Acad. Sci. USA 98 (2001) 8991–8996.

[16] M. Procopio, A. Lahm, A. Tramontano, L. Bonati, D. Pitea, A model for recognition of polychlorinated dibenzo-*p*-dioxins by the aryl hydrocarbon receptor, Eur. J. Biochem. 269 (2002) 13–18.

[17] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, The protein data bank, Nucleic Acids Res. 28 (2000) 235–242.

[18] B. Lee, F.M. Richards, Interaction of protein structures: estimation of static accessibility, J. Mol. Biol. 55 (1971) 379–400.

[19] S.I. Karchner, D.G. Franks, S.W. Kennedy, M.E. Hahn, The molecular basis for differential dioxin sensitivity in birds: role of the aryl hydrocarbon receptor, Proc. Natl. Acad. Sci. USA 103 (2006) 6252–6257.

# Rapid SNP diagnostics using asymmetric isothermal amplification and a new mismatch-suppression technology

Yasumasa Mitani[1,2], Alexander Lezhava[1], Yuki Kawai[1,2], Takeshi Kikuchi[1,2], Atsuko Oguchi-Katayama[1], Yasushi Kogo[1], Masayoshi Itoh[3], Toru Miyagi[4], Hideki Takakura[1,5], Kanako Hoshi[1,5], Chiaki Kato[2], Takahiro Arakawa[1], Kazuhiro Shibata[3], Kenji Fukui[6,7], Ryoji Masui[6,7], Seiki Kuramitsu[6,7], Kazuma Kiyotani[8], Alistair Chalk[9], Katsuhiko Tsunekawa[10], Masami Murakami[10], Tetsuya Kamataki[8], Takanori Oka[4], Hiroshi Shimada[5], Paul E Cizdziel[1] & Yoshihide Hayashizaki[1,3]



**We developed a rapid single nucleotide polymorphism (SNP) detection system named smart amplification process version 2 (SMAP 2). Because DNA amplification only occurred with a perfect primer match, amplification alone was sufficient to identify the target allele. To achieve the requisite fidelity to support this claim, we used two new and complementary approaches to suppress exponential background DNA amplification that resulted from mispriming events. SMAP 2 is isothermal and achieved SNP detection from whole human blood in 30 min when performed with a new DNA polymerase that was cloned and isolated from *Alicyclobacillus acidocaldarius* (*Aac* pol). Furthermore, to assist the scientific community in configuring SMAP 2 assays, we developed software specific for SMAP 2 primer design. With these new tools, a high-precision and rapid DNA amplification technology becomes available to aid in pharmacogenomic research and molecular-diagnostics applications.**

The availability of the human genome sequence[1,2] and genome diversity databases[3–5] at the beginning of the 21st century are causing a paradigm shift away from the standard protocol of medical care toward genotyped medicine. This new type of medicine is based on the accumulating knowledge of gene polymorphisms (SNPs) and their relationship to specific phenotypes, such as disease predisposition, drug metabolism and disease development. A key step for the development of individualized medicine is the ability to rapidly test patients for these SNPs and/or other mutations correlated to diseases and disease predisposition. Supporting this point, the US Food and Drug Administration has required the drug industry to publicly provide SNP data examined in the process of procuring a drug license. Today SNP genotyping technologies[6–9] are still a bottleneck in drug discovery research and clinical applications. But high-throughput gene analysis and SNP detection technologies will inevitably become both cheaper and faster in the future. Besides SNP genotyping, these improved sequence-detection technologies would also allow and advance studies in other disciplines such as population genetics, the global surveillance of infectious disease and the study of somatic mutations in human cancer.

Almost all previously developed SNP-detection systems consist of two steps: amplification (usually by PCR) and detection of SNP (using DNA fragments amplified in the first step). This approach is reasonably fast, but to shorten the time required and simplify the detection, it is ideal to develop a one-step method, in which the amplification itself can be the SNP detection signal. The difficulty in developing such a technology is in the suppression of the background amplification. For example, primers for allele-specific primer PCR are designed with the nucleotide mismatch at the 3′ end of the PCR primers, but the misamplified PCR products primed from mismatched primers are still exponentially amplified, producing background signals that must be addressed.

Here we report SMAP 2, the first rapid one-step SNP detection technology in which the amplification of the targeted DNA is the signal of the target SNP itself.

[1]Genome Exploration Research Group (Genome Network Project Core Group), RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan. [2]K.K. Dnaform, 75-1, Ono-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0046, Japan. [3]Genome Science Laboratory, Discovery Research Institute, RIKEN Wako Institute, 2-1 Hirosawa, Wako, Saitama, 351-0198, Japan. [4]Wakunaga Pharmaceutical Co. Ltd., 1624 Shimokotachi, Koda-cho, Akitakata-shi, Hiroshima, 739-1195, Japan. [5]Department of Gastroenterological Surgery, Yokohama City University Graduate School of Medicine, 3-9, Fukuura Kanazawa-ku, Yokohama, 236-0004, Japan. [6]Department of Biological Sciences, Graduate School of Science, Osaka University, 1-1, Machikaneyama-cho, Toyonaka, Osaka 560-0043, Japan. [7]RIKEN SPring-8 Center, RIKEN Harima Institute, 1-1-1 Kohto, Sayo-cho, Sayo-gun, Hyogo 679-5148, Japan. [8]Faculty of Pharmacy, Takasaki University of Health and Welfare, 60 Nakaorui-machi, Takasaki-shi, Gunma 370-0033, Japan. [9]Center for Molecular Medicine, Karolinska Institutet, 171 76 Stockholm, Sweden. [10]Department of Clinical Laboratory Medicine, Gunma University Graduate School of Medicine, 3-39-22 Showa-machi, Maebashi-shi, Gunma, 371-8511, Japan. Correspondence should be addressed to Y.H. (yosihide@gsc.riken.go.jp).

RECEIVED 22 SEPTEMBER 2006; ACCEPTED 28 DECEMBER 2006; PUBLISHED ONLINE 18 FEBRUARY 2007; DOI:10.1038/NMETH1007

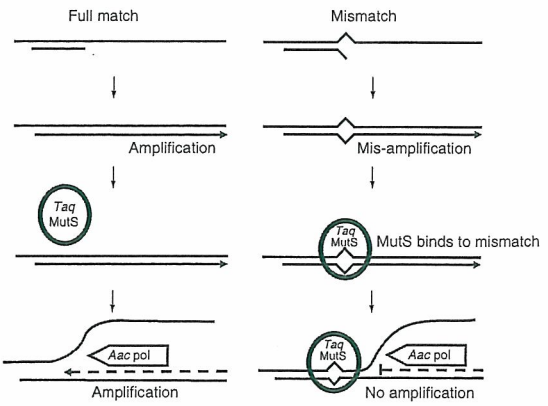


**Figure 1** | The mechanism of allele discrimination as mediated by *Taq* MutS. SNP typing with a wild-type allele-specific primer, using the wild-type allele (left) and the mutant allele (right) as templates. The wild-type allele discrimination primer is designed to encompass the SNP nucleotide site at a position close to the 3′ end. Hybridization of the wild-type discrimination primer to the wild-type allele promotes amplification. On occasion, however, misamplification occurs using the wild-type discrimination primer on a mutant allele. This mismatch amplification event generates a dsDNA with a single mismatched base pair. In SMAP 2, *Taq* MutS tightly binds to mismatched nucleotides in dsDNA and *Aac* DNA polymerase cannot strand-displace or extend through the nucleoprotein complex. Hence mismatch amplification is inhibited.

## RESULTS

### Suppression of background amplification

To avoid exponential background amplification, we developed two fundamental technologies. Together these technologies ensure that primer sets designed to amplify highly related sequences, such as a SNP variant and a wild-type allele, do not misamplify the incorrect sequence. The first technology uses *Thermus aquaticus* MutS (*Taq* MutS)[10] in combination with an isothermal amplification procedure (**Fig. 1**). In a SMAP 2 assay for SNP genotyping, one primer is designed with homology to the SNP sequence and flanking DNA. We refer to this primer as the discrimination primer. As with any DNA polymerase *in vitro*, occasionally mispriming occurs. In SMAP 2, however, the *Taq* MutS protein binds to the nucleotides consisting of a mismatched duplex between the target DNA and the extended discrimination primer. This protein binding blocks the disassociation of the mismatched DNA duplex by the strand-displacing DNA polymerase and thus prevents exponential background amplification. The effectiveness of the *Taq* MutS is a consequence of using it in an isothermal amplification protocol rather than in a thermal cycling system like PCR, in which the *Taq* MutS dissociates at high temperatures.

The second strategy to suppress the background signals in SMAP 2 is the asymmetric design of the primers. Two asymmetric primers with unique sequences flanking the target sequence, a folding primer (FP) and a turn-back primer (TP), are used to amplify a specific genomic region.

### Primer design and amplification principle

The entire process of amplification by SMAP 2 requires five primers: TP, FP, boost primer (BP) and two outer primers (OP1 and OP2; **Figs. 2** and **3**). The TP 3′-end sequence is complementary to the target sequence and the 5′-end sequence is complementary to a sequence approximately 15–40 base pairs downstream on the same strand of DNA. The FP 3′-end sequence is complementary to a target genome sequence, and the 5′-end sequence is self-annealing and creates a hairpin structure. OP1 and OP2 strand-displace the DNA synthesized in FP and TP primer extension. The discrimination primer can be FP, TP or BP. For assay design any choice may suffice, but occasionally one specific primer design strategy may be more effective. At present this can only be determined empirically. The variable-nucleotide position in the discrimination primer can be at the 3′-end *n*, *n*–1 or *n*–2 position of FP, TP or BP, and the 5′-end *n*, *n*–1 or *n*–2 position of TP. The role of BP is to boost the speed of amplification, but it can also be used as a discrimination primer. Coordinated melting temperature ($T_m$) values and proper distancing between primers annealing sites are important in SMAP 2 assay design (**Fig. 2**). When designing primers, one should avoid sequences that have the potential to form secondary structures and follow the same rules that apply to the design of PCR primers. We have been successful in designing primers with 40–60% G+C content. Furthermore, an amplicon length of less than 200 bp is required to achieve the highest-efficiency (fastest) amplification.

The genomic sequence between and including the TP and FP primers is the target region to be amplified by the SMAP 2 reaction (**Figs. 2** and **3**). In the first step, FP and TP hybridize to the template genomic DNA. Both products primed from the FP and TP are then dissociated from the template genomic DNA by the strand-displacement activity of the DNA polymerase, whose extension is primed by OP1 and OP2. These displaced single-stranded products then become templates in the second step for the opposing FP and TP. Single-stranded displaced DNA products, which we refer to as intermediate products (IM1 and IM2), are then generated by the strand-displacement activity of the DNA

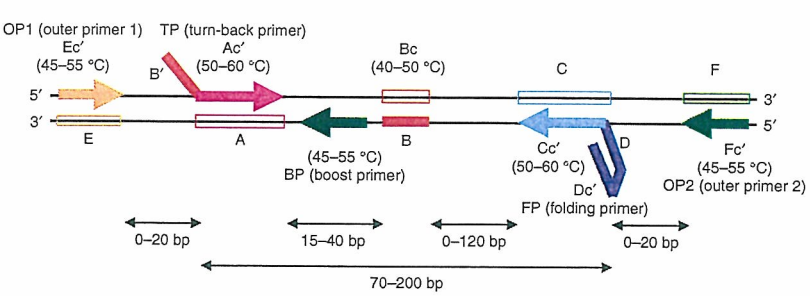


**Figure 2** | SMAP 2 primer design. Primers for SMAP 2 amplification with the recommended distance between primer regions and a $T_m$ range for each primer type. Genomic DNA regions of unique sequence are designeated A, B, C, F and E. Primers are indicated by solid arrows and given the corresponding capital letter designation followed by c′, indicating it is a complementary sequence. The D and Dc′ region of FP are self-complementary and have no homology to genomic DNA. The region B′ of the TP is complementary to the Bc region on the same strand and is responsible for 'loop-back' self-priming that occurs later on intermediate species. In SMAP 2, the target SNP (or mutation) must be present between the 5′-end of the Ac′ region of the TP and the 5′-end of the Cc′ region of the FP, as these define the outer regions of the genomic DNA to be amplified in the SMAP 2 amplicon.

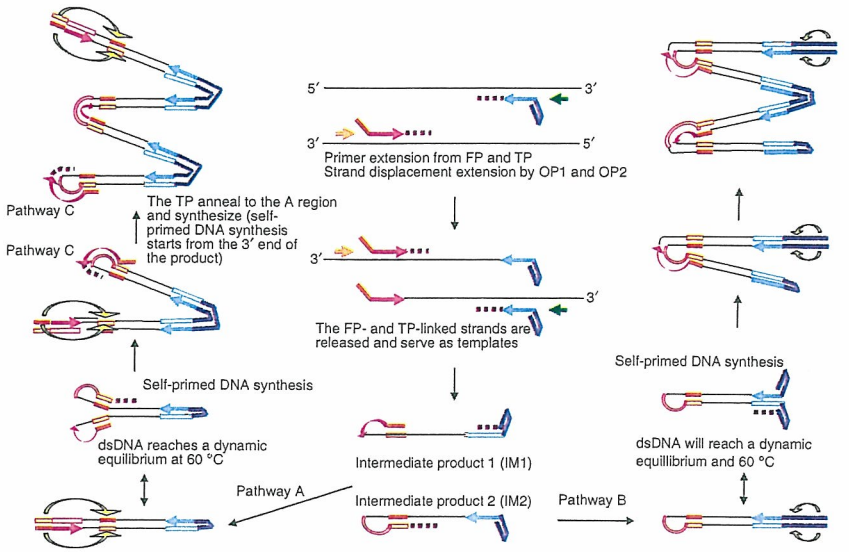


**Figure 3** | SMAP amplification process. The initial priming events of the SMAP amplification process generate two intermediates, IM1 and IM2. Next self-priming DNA synthesis from each of these intermediates create hairpin molecules via pathway A or B. These structures lead to further self-primed DNA synthesis to create a dimeric amplicon and then subsequent larger species. Pathway C is a third possibility, in which a free TP anneals to the amplicon at an internal complementary site (A region; see **Fig. 2**) and primes strand-displacement DNA synthesis. Amplification proceeds until reaction components are exhausted. Typical yields of SMAP 2 exceed those of PCR by 100-fold.

polymerase primed from the flanking OP regions adjacent to the target sequence. IM1 and IM2 are milestone products for the subsequent amplification steps. IM1 has the TP sequence at the 5′ end and the FP complementary sequence at the 3′ end, and IM2 is complementary to IM1. The initial self-priming elongation site on IM1 is the 3′-end of the FP sequence of IM1, and this priming also occurs later on other multiple-unit–length (more mature) products. Concatenated products of IM1 are synthesized by an elongation process termed pathway A. The characteristic feature of the products of pathway A is that the free 5′ and 3′ ends carry TP and its complementary sequence, forming long double-stranded hairpin DNA. The initial self-priming elongation site on IM2 is located at the 3′ end of the TP sequence of IM2, and this priming also occurs later on other multiple-unit-length (more mature) products. Long concatenated DNA products are synthesized as in pathway A, but denoted as pathway B because the end products are different. These long-hairpin DNA products carry FP and its complementary sequence at the free 5′ and 3′ ends, respectively. There is one other elongation event that is defined as pathway C. The elongation via pathway C starts from the 3′ end of a free TP-primer that hybridizes to the looping structure of the TP complementary sequence, which is located at the intermediate region of the long products of pathway A.

One key feature responsible for the suppression of the background is the asymmetric design of FP and TP. On comparison with other isothermal amplification technologies, SMAP 2 is most similar to LAMP[11]. LAMP technology, however, uses a symmetrical primer design with two TP primers, such that a dumbbell form is the milestone amplification product. A consequence of this LAMP primer design strategy is that additional opportunities exist for free-primer hybridization and priming events that otherwise do not occur in SMAP because self-priming events are thermodynamically favored by the FP when incorporated in the amplicon. Hence the primer design of SMAP 2 minimizes the alternative misamplification pathways and consequential background products (**Supplementary Fig. 1** online).

To test and observe the effect of these primer design and technique differences, we performed SMAP 2 and LAMP with closely matched primer sets on an identical target. We performed the assays on whole blood from a homozygous wild-type individual, with wild-type and mutant primer sets designed to amplify the human type 2 iodothyronine deiodinase[12] gene (*DIO2*). We performed the amplification reaction directly from lysed blood samples without purifying the DNA. We achieved real-time monitoring of the amplification by using intercalating SYBR Green I to detect the generation of double-stranded DNA at 60 °C. Both SMAP 2 and LAMP used the same BP, OP1, OP2 and TP. The TP was designed as the discrimination primer with the SNP detection nucleotide on the 5′ end. The other 'reversing' primer, which is TP for LAMP and FP for SMAP 2, had 3′-end sequence homologous to the target, but differed on the 5′ end consistent with design requirements for the respective technique (**Supplementary Fig. 2** online). Furthermore, the reaction condition and enzymes for each of the techniques were identical, and experiments were performed with both full-match and mismatched primer sets, with and without *Taq* MutS. The data showed that LAMP amplification proceeds slightly faster than SMAP 2, and that *Taq* MutS slightly delayed full-match amplification in both techniques. On mismatched targets, however, the effect of *Taq* MutS on LAMP was substantially less than compared to the total suppression of mismatch amplification that was evident with SMAP 2 under the same conditions (**Supplementary Fig. 3** online). This is due to the greater abundance of mismatch product in the LAMP reaction that was generated early in the process via other priming pathways. Analysis of the assay products on an agarose gel revealed a dissimilar banding pattern for SMAP 2 and LAMP, indicating different self-priming mechanisms of amplification (**Supplementary Fig. 4** online).

### The effect of *Taq* MutS on SNP typing of *ALDH2*

To demonstrate the amplification efficiency and specificity of SMAP 2, we used the SNP (G1543A) of aldehyde dehydrogenase 2 (*ALDH2*), which is the gene involved in alcohol sensitivity[13], as the test model system. On a blood sample from an individual known to be homozygous wild-type (as determined by PCR sequencing; data not shown), amplification occurred when using a wild-type primer, BP(W), both in the presence of *Taq* MutS and in its absence (**Fig. 4**). When using the SNP-variant primer, BP(M), on the wild-type blood sample, we observed misamplification of the wild-type target after 25 min in the absence of *Taq* MutS. But the amplification was completely suppressed by inclusion of *Taq* MutS in the reaction, with no signal observed even after 255 min of

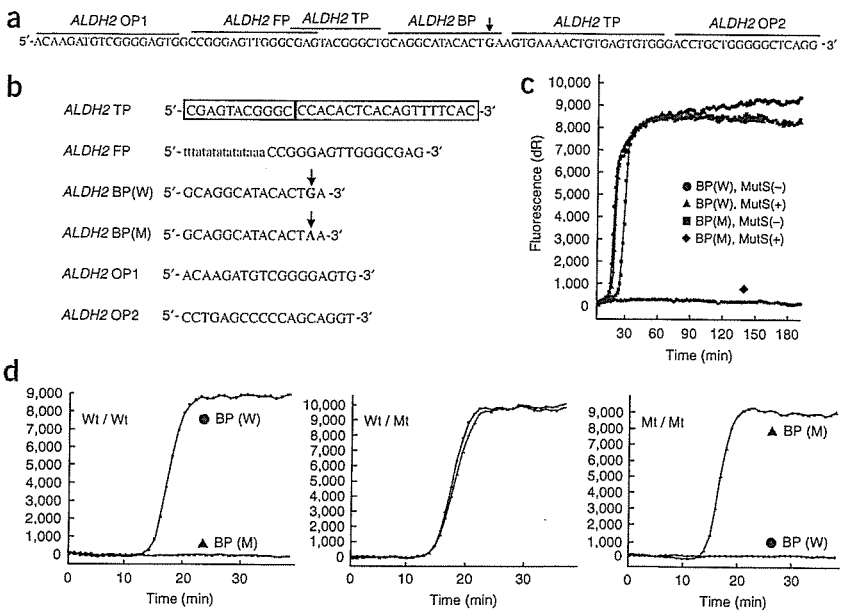


**Figure 4** | SNP typing of *ALDH2* allele. (**a**) Sequence of *ALDH2* gene. The arrowhead indicates the position of the SNP nucleotides of the *ALDH2* allele (G1543A). DNA sequences used for primer design are marked. (**b**) Primer set for SNP typing of *ALDH2*. BP(W) and BP(M) were designed to be wild-type allele–specific and mutant allele–specific, respectively. The arrowheads indicate the nucleotides corresponding to SNP site. (**c**) The effect of *Taq* MutS on SNP detection of *ALDH2* gene. The intensity of fluorescence of SYBR Green I dye monitored during SMAP 2 reaction with or without *Taq* MutS. The blood sample of a homozygous wild-type DNA was used as a template for all reactions. dR, baseline-subtracted fluorescence reading. (**d**) Amplification time course of SMAP 2 reaction with *ALDH2* allele–specific primers using human blood specimens. Three possible diploid genotypes of *ALDH2* alleles are shown. The left, center and right graphs are time courses for a homozygous wild type, heterozygote and a homozygous mutant, respectively.

incubation. These data indicated that *Taq* MutS was capable of completely suppressing the mismatch background amplification.

To demonstrate the accuracy and clinical utility of this assay for handling multiple samples, we next tested 63 blood samples of healthy candidates. The fluorescence of both wild-type and mutant alleles could be detected at 17 min after incubation at 60 °C. Absolutely no background signal was evident, indicating that mismatch amplification was completely suppressed. To validate these results, we also detected the same SNP by using PCR–restriction fragment length polymorphism (PCR-RFLP)[14]. Of the SMAP 2 assays performed on 63 individuals, 41 tested homozygous for wild-type, 20 were heterozygous and 2 were homozygous mutant for the *ALDH2* alleles. All the data for the 63 samples obtained by PCR-RFLP and SMAP 2 showed perfect concordance (data not shown).

## Accurately genotyping gene-family members with SMAP 2

The cytochrome *P450* proteins comprise a large gene family and are challenging to distinguish with hybridization or amplification techniques. SMAP 2 was capable of discrimination of not only these gene family members, but also of unique SNP variants of a specific subtype. We designed primer sets for amplification of the wild-type cytochrome P450, family 2, subfamily C, polypeptide 19 gene (*CYPZC19*1*) and an allelic variant known as *CYP2C19*2*(G681A)[15]. We chose this example to illustrate the amplification specificity and discrimination power of SMAP 2. In the case of *CYP2C19*, it is necessary to discriminate the signals from highly related subtypes *CYP2C8*, *CYP2C9* and *CYP2C18*, as all four of these genes have a high degree of sequence similarity (**Fig. 5**). We selected SMAP 2 primer sequences (BP, FP and OP) that are unique to the *CYP2C19* subtype to allow the amplification of it alone, by

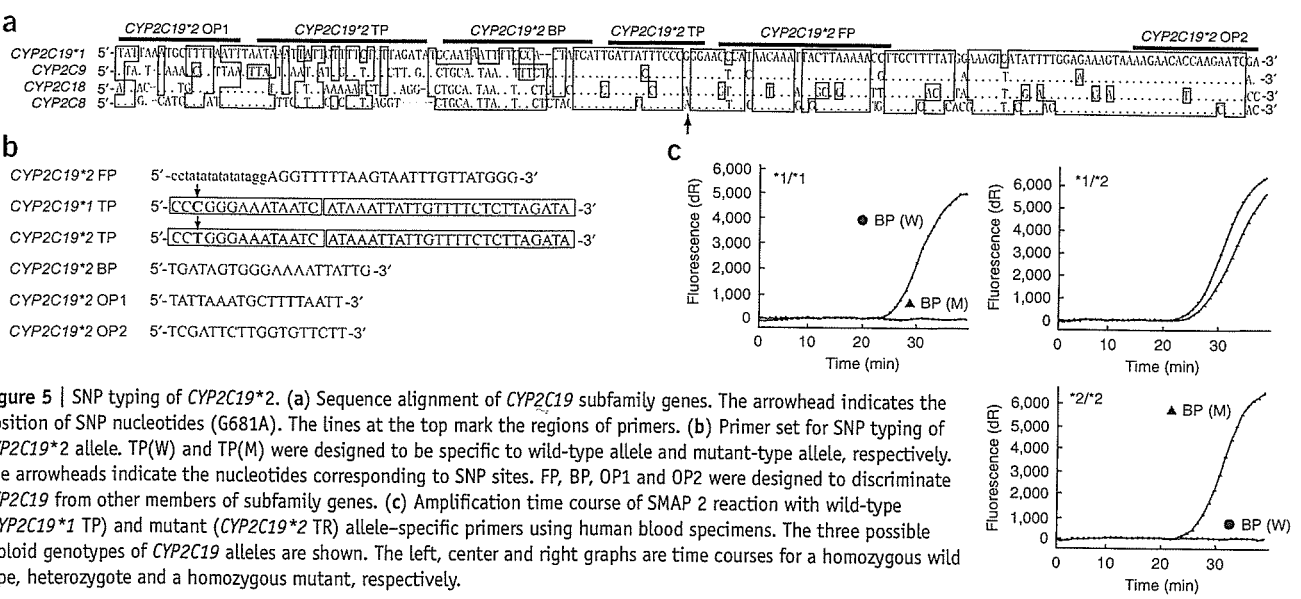


**Figure 5** | SNP typing of *CYP2C19*2*. (**a**) Sequence alignment of *CYP2C19* subfamily genes. The arrowhead indicates the position of SNP nucleotides (G681A). The lines at the top mark the regions of primers. (**b**) Primer set for SNP typing of *CYP2C19*2* allele. TP(W) and TP(M) were designed to be specific to wild-type allele and mutant-type allele, respectively. The arrowheads indicate the nucleotides corresponding to SNP sites. FP, BP, OP1 and OP2 were designed to discriminate *CYP2C19* from other members of subfamily genes. (**c**) Amplification time course of SMAP 2 reaction with wild-type (*CYP2C19*1* TP) and mutant (*CYP2C19*2* TR) allele–specific primers using human blood specimens. The three possible diploid genotypes of *CYP2C19* alleles are shown. The left, center and right graphs are time courses for a homozygous wild type, heterozygote and a homozygous mutant, respectively.

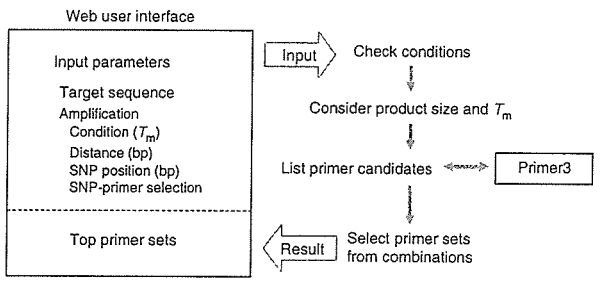


**Figure 6** | SMAP primer design software version 1. Web user input variables and design process are outlined for the SMAP primer design software. The software generates primers designed in both forward and reverse directions with respect to the position of BP, TP and FP on the template.

being an imperfect match to all other family members (*CYP2C8*, *CYP2C9* and *CYP2C18*). We used TP as a discrimination primer to base pair with the target guanine in the wild type (*CYP219C*1* TP) and the adenine in the variant (*CYP2C19*2* TP). We verified all the data for the human blood samples by PCR-sequencing, and found perfect concordance with the SMAP 2 results, indicating that the SMAP 2 method is versatile and that the use of specific primers for differentiation of highly related targets was achievable.

### SMAP Primer design software version 1

To facilitate the design of primer sets for SMAP applications we developed the SMAP server and algorithms specifically for SMAP 2 primer design (**Fig. 6**). The program, developed in Java and Struts uses Primer3 [Primer3][16] to generate a set of potential sequences for each of the primers required. Primer selection is based on the Primer3 scoring algorithms using optimal $T_m$ and product-size range. Also, the requirements for TP and FP as well as all other primers are generated and combined to form all possible sets of primers where the SNP is in a valid location on either the FP, BP or TP. Primer sets are then ranked based on the Primer3 scoring for each of the primers. The SMAP 2 primer design software is freeware, available to assist researchers in generating unique primer sets for amplification of target sequences of interest. Presently there is not enough information to fully understand what constitutes the best SMAP 2 primers or primer combinations. Typical PCR rules applied by the Primer3 software are relevant, but other poorly understood SMAP-specific requirements clearly exist as not all primer sets that are theoretically designed for SNP detection, can faithfully discriminate the SNP from wild-type sequence. Hence the SMAP 2 primer design software cannot rank-order the primer sets based on theoretical performance, nor purposefully generate the best primer combinations for the target sequence. We found, however, that the software is an effective tool for initial experiments. It is recommended that the primer sets be refined through empirical experimentation (modification of primer length, position and other parameters), if an increase in speed and specificity of the SMAP reaction is desired.

We tested the primer-design software on two different gene targets, iodothyronine deiodinase type 2 (*DIO2*) and β3 adrenergic receptor (*ADRB3*)[17]. We used the software to design several primer sets for each gene and SNP variant, which we then synthesized and ran SMAP 2 trial experiments. We selected the best primer sets based on specificity and speed of the reaction. Then we optimized these primer sets by further rational design until a set providing optimal speed and specificity was configured. These optimized primers performed well and yielded accurate SNP typing data that was evident in 30 min (**Supplementary Figs. 2** and **5** online). Information on how to use the software and how to optimize the primers by incorporating minor changes (length changes, repositioning), is available on the RIKEN SMAP server.

## DISCUSSION

SMAP 2 is based on the concept that DNA amplification is itself the signal for detection of a specific target sequence. It is conceivable that SMAP 2 background suppression technologies can also benefit other types of isothermal amplification, although we observed minimal impact when MutS was used in LAMP assays (**Supplementary Fig. 3**). The technique is simple, requires no DNA purification and is performed in a closed tube, which reduces the risk of contamination. We included data generated on real-time PCR detection systems in this report; however, end-point detection, including colorimetric or spectrophotometric analysis may be possible because of the background suppression and high yields (>100× PCR yields in equivalent assay volumes). End-point determination is essentially digital and may enable simple and cost-effective detection methodologies that could be used in field applications or in countries with limited financial resources for health-care diagnostics. Furthermore, owing to the high yields and isothermal nature, very small-scale microfluidic designs for high-throughput SNP screening applications may be feasible.

In SMAP 2 flexibility of primer design to optimize performance is greater than in many other traditional isothermal amplification methods[18–20]. We demonstrated that each of the primers (FP, TP and BP) can be engineered as a discrimination primer, facilitating the ability to detect any type of nucleotide change. In addition to single-base alterations, insertions and deletions of 1–18 bases have also been detected in clinical samples from human cancer and other tissues (unpublished data). The SMAP 2 assay is sensitive enough to detect a few molecules (**Supplementary Fig. 6** online) making it potentially useful for mutation-screening of tumor samples that frequently display defined mutations in specific genes.

SMAP 2 likely will not replace PCR as a basic research tool for amplification because of the complexity of its primer design and optimization. Furthermore, the concatenated form of the amplified material is not useful for most research applications such as cloning, sequencing and expression. We have demonstrated, however, that SMAP 2 has strong potential for SNP genotyping with high accuracy, and unlike any other method, gives a reliable diagnostic result based exclusively on amplification alone.

## METHODS

**Preparation of blood sample.** We collected blood samples from volunteers after obtaining their written informed consent for each collection. All donors were researchers and employees of the institutions represented by the authors of this paper. Institutional approval for conducting research using human material was obtained from the RIKEN Ethical Advisory Committee before initiating the study. We collected ~30 μl of blood by pricking the finger of volunteers. We divided the collected samples and used one aliquot directly for SMAP 2 reactions; we diluted the

remaining blood samples approximately threefold with 50 mM NaOH and heated them at 98 °C for 3 min. For each 25-μl SMAP reaction, we added 1 μl of the diluted and heated blood sample directly to the assay. We purified genomic DNA from a second aliquot of each sample with the QuickGene-mini80 (FUJIFILM) and used this for PCR-RFLP of *ALDH2*. To examine the accuracy of the other SMAP 2 data reported in this paper, we PCR-amplified the SNP targets and directly sequenced the amplified DNA to verify the SNP genotype results.

**Reaction mix for SMAP (asymmetrical primer amplification).** We typically performed SMAP reactions in a 25-μl volume. Enzymatic components (*Aac* DNA polymerase and *Taq* MutS) for the SMAP 2 assays were supplied by DNAFORM K.K. in a prototype SMAP 2 core kit. Each reaction contained 3.2 μM each of FP and TP, 0.4 μM each of OP1 and OP2, 1.6 μM BP, 1.4 mM dNTPs, 5% DMSO, 20 mM Tris-HCl (pH 8.8), 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 8 mM $MgSO_4$, 0.1% Tween 20, 1/100,000 diluted original SYBR Green I, 40 units of *Aac* DNA polymerase, 1.5–2.4 μg of *Taq* MutS and 1 μl of prepared blood sample. We performed SMAP 2 reactions at 60 °C.

**Reaction mix for LAMP (symmetrical primer amplification).** We performed LAMP reactions in a 25-μl volume using *Aac* DNA polymerase (DNAFORM K.K.). The assay composition was identical to the reaction mix for SMAP 2 reactions with the exception of the FP. For LAMP we designed a second TP primer (TP2; **Supplementary Figs. 2** and **3**) and used at the same concentration (3.2 μM) as the other TP (TP1). We prepared the LAMP and SMAP 2 comparative samples identically and carried out the reactions at 60 °C.

**Data analysis and diagnostic judgment.** We evaluated the application of SMAP 2 in practical diagnostics according the principle of amplification versus non-amplification compared to threshold values. The presence of intercalating SYBR Green I dye during the reaction was monitored with Mx3000P system (Stratagene) and SNP typing was determined according to the fluorescence intensity. Generally, we considered amplification to be positive if the fluorescence (dR: baseline-subtracted fluorescence reading) strength was higher than a value of 1,000, and that no amplification occurred if the signal was less than 1,000.

**Additional methods.** Descriptions of the SMAP 2 assay design strategy, sample preparation, amplification-mix setup and additional information are available in **Supplementary Methods** online.

**URLs.** Information about availability of reagents and licenses for conducting SMAP technology assay development and commercial use is available from the DNAFORM K.K. website (http://www.dnaform.jp/index_e.html). The RIKEN SMAP server web site for SMAP 2 primer design software and suggestions for further optimization can be found at (http://www.smapdna.com).

*Note: Supplementary information is available on the Nature Methods website.*

**ACKNOWLEDGMENTS**

We thank A. Wada for support and encouragement, T. Ishikawa (Tokyo Institute of Technology) for his constructive discussion about the application of this technology and future prospects, M. Matsunaga, J. Nakashima, M. Matsushita and S. Uno for technical assistance. We acknowledge K. Nakano (NTT Software Corporation) for assistance in web support and software design. We also thank H. Daub and M. Nishikawa (RIKEN) for their editorial and coordination efforts. This study was mainly supported by the research grant for the RIKEN Genome Exploration Research Project from the Ministry of Education, Culture, Sports, Science and Technology of the Japan (MEXT) to Y. Hayashizaki, and RIKEN "Research Collaborations with Industry" Program to K. Shibata. S. Kuramitsu is supported by the Research Grant for National Project on Protein Structure and Functional Analysis from MEXT.

**COMPETING INTERESTS STATEMENT**

The authors declare competing financial interests (see the *Nature Methods* website for details).

**Published online at http://www.nature.com/naturemethods/**
**Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions**

1. Venter, J.C. *et al.* The sequence of the human genome. *Science* **291**, 1304–1351 (2001).
2. Lander, E.S. *et al.* Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).
3. Miller, R.D. *et al.* High-density single-nucleotide polymorphism maps of the human genome. *Genomics* **86**, 117–126 (2005).
4. Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. *Nature* **431**, 931–945 (2004).
5. Sachidanandam, R. *et al.* A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* **409**, 928–933 (2001).
6. Hardenbol, P. *et al.* Multiplexed genotyping with sequence-tagged molecular inversion probes. *Nat. Biotechnol.* **21**, 673–678 (2003).
7. Kennedy, G.C. *et al.* Large-scale genotyping of complex DNA. *Nat. Biotechnol.* **21**, 1233–1237 (2003).
8. Lyamichev, V. *et al.* Polymorphism identification and quantitative detection of genomic DNA by invasive cleavage of oligonucleotide probes. *Nat. Biotechnol.* **17**, 292–296 (1999).
9. Shen, R. *et al.* High-throughput SNP genotyping on universal bead arrays. *Mutat. Res.* **573**, 70–82 (2005).
10. Biswas, I. & Hsieh, P. Identification and characterization of a thermostable MutS homolog from *Thermus aquaticus*. *J. Biol. Chem.* **271**, 5040–5048 (1996).
11. Iwasaki, M. & Yonokawa, T. Validation of the loop-mediated isothermal amplification method for single nucleotide polymorphism genotyping with whole blood. *Genome Lett.* **2**, 119–126 (2003).
12. Mentucci, D. *et al.* Association between a novel variant of the human type 2 deiodinase gene Thr92Ala and insulin resistance. *Diabetes* **51**, 880–883 (2002).
13. Harada, S., Okubo, T., Tsutsumi, M., Takase, S. & Muramatsu, T. Investigation of genetic risk factors associated with alcoholism. *Alcohol. Clin. Exp. Res.* **20**, 293A–296A (1996).
14. Harada, S. & Zhang, S. New strategy for detection of *ALDH2* mutant. *Alcohol Alcohol. (Suppl.)* **1A**, 11–13 (1993).
15. de Morais, S.M. *et al.* The major genetic defect responsible for the polymorphism of *S*-mephenytoin metabolism in humans. *J. Biol. Chem.* **269**, 15419–15422 (1994).
16. Rozen, S. & Skaletsky, H. Primer3 on the WWW for general users and for biologist programmers. *Methods Mol. Biol.* **132**, 365–386 (2000).
17. Walston, J. *et al.* Time of onset of non-insulin-dependent diabetes mellitus and genetic variation in the beta 3-adrenergic-receptor gene. *N. Engl. J. Med.* **333**, 343–347 (1995).
18. Berard, C., Cazalis, M.A., Leissner, P. & Mougin, B. DNA nucleic acid sequence-based amplification-based genotyping for polymorphism analysis. *Biotechniques* **37**, 680–686 (2004).
19. Pickering, J. *et al.* Integration of DNA ligation and rolling circle amplification for the homogeneous, end-point detection of single nucleotide polymorphisms. *Nucleic Acids Res.* **30**, e60 (2002).
20. Wang, S.S., Thornton, K., Kuhn, A.M., Nadeau, J.G. & Hellyer, T.J. Homogeneous real-time detection of single-nucleotide polymorphisms by strand displacement amplification on the BD ProbeTec ET system. *Clin. Chem.* **49**, 1599–1607 (2003).

*Xenobiotica*, January 2007; 37(1): 30–43

# CYP2C76-mediated species difference in drug metabolism: A comparison of pitavastatin metabolism between monkeys and humans

Y. UNO[1,3], T. KUMANO[1], G. KITO[1,3], R. NAGATA[1,3], T. KAMATAKI[2,5], & H. FUJINO[4]

[1]*Laboratory of Translational Research,* [2]*Laboratory of Drug Metabolism, Graduate School of Pharmaceutical Sciences, Hokkaido University, Sapporo, Japan,* [3]*Shin Nippon Biomedical Laboratories, Tokyo, Japan,* [4]*Tokyo New Drug Laboratories I, Kowa Co., Tokyo, Japan, and* [5]*Faculty of Pharmacy, Takasaki University of Health and Welfare, Takasaki, Japan*

*(Received 30 March 2006; accepted 7 July 2006)*

**Abstract**

The monkey is often used to predict metabolism of drugs in humans since it generally shows a metabolic pattern similar to humans. However, metabolic profiles different from humans are occasionally seen in monkeys for some drugs including pitavastatin. Recently, we have successfully identified a monkey-specific cytochrome P450 (CYP) 2C76, which possibly accounts for a species difference between monkeys and humans because of its sequence and functional uniqueness. The present study on the role of CYP2C76 and other monkey CYP2Cs in pitavastatin metabolism, as an example, has revealed that CYP2C76 is important for the metabolism of the lactone form, indicating a major role of CYP2C76 for the difference in the metabolism of pitavastatin and possibly other drugs between monkeys and humans. The current investigation on the involvement of CYP2C76 in the metabolism of other drugs is expected to reveal further the further importance of this monkey-specific drug-metabolizing enzyme.

**Keywords:** *CYP2C76, monkey-specific, species difference, drug metabolism, pitavastatin*

## Introduction

Animal models are regularly used during drug development to predict the metabolic fate of drugs in humans (Nedelcheva and Gut 1994; Bogaards et al. 2000).

Correspondence: H. Fujino, Tokyo New Drug Research Laboratories I, Kowa Co. Ltd, Noguchicho 2-17-43, Higashimurayama, Tokyo 189-0022, Japan. Tel: +81-42-391-6211. Fax: +81-42-395-0312. E-mail: fujino@pk2.so-net.ne.jp

ISSN 0049-8254 print/ISSN 1366-5928 online 
DOI: 10.1080/00498250600968275

However, numerous reports have documented the difference in the metabolic fate of drugs between animal species (Jacqz et al. 1988; Ohi et al. 1989; Stevens et al. 1993; Weaver et al. 1994, 1999; Sharer et al. 1995; Prueksaritanont et al. 1996; Narimatsu et al. 2000). Consequently, the clinical relevance of the results obtained from animals is difficult to assess. Hence, to predict the metabolism of drugs and the influence of drug–drug interactions in humans, species differences in drug-metabolizing enzymes including cytochrome P450s (CYPs) should be clarified and taken into account when selecting animal species.

Our and other groups have identified and characterized cDNAs for CYP2C20, CYP2C43, CYP2C75 and CYP2C76 from cynomolgus or rhesus monkeys (Komori et al. 1992; Matsunaga et al. 2002; Uno et al. 2006). Failure to detect expression of both the mRNA and the protein in human liver and absence of the gene in the corresponding region of the human genome strongly suggest that CYP2C76 does not have the orthologue in humans (Uno et al. 2006). The abundant expression of CYP2C76 in the primate liver and the unique metabolic profile further support the possibility that the species–specific CYP2C76 could account for the difference in drug metabolism between monkeys and humans.

Pitavastatin, a novel synthetic HMG-CoA reductase inhibitor (statin), has a persistent effect on serum lipids (Suzuki et al. 1999; Aoki et al. 2002). Our analysis of pitavastatin showed that the profiles of plasma concentration after oral administration indicated a relatively rapid absorption from intestine and that the half-life was approximately 4 h in rats, rabbits, monkeys, dogs and humans (Fujino et al. 1999a). The apparent bioavailability of pitavastatin was greater than 80% in these latter species except for monkeys that showed a relatively low plasma concentration and low bioavailability (less than 20%). Moreover, faecal and urinary excretions of pitavastatin (unchanged form) were negligible in monkeys compared with other animals after intravenous administration, indicating a marked species differences of pitavastatin metabolism between monkeys and other animal species including humans (Fujino et al. 1999b).

Pitavastatin is administered as the acid form; however, the area under the concentration–time curve (AUC) of its lactone form was similar in clinical studies (Fujino et al. 1999a; Kojima et al. 1999). We previously demonstrated the involvement of UDP-glucuronosyl transferase (UGT) for the lactonization of pitavastatin in human and animals (Fujino et al. 2003). Pitavastatin lactone was mainly hydrolysed to its acid form in human hepatic microsomes, whereas a remarkable increase in metabolic clearance was noted for the lactone form in monkey hepatic microsomes as compared to the acid form (Yamada et al. 2003). Moreover, metabolic clearance of the acid and lactone forms was much greater in monkeys than in humans. These results demonstrate that both forms of pitavastatin undergo different metabolic processes between monkeys and humans. In humans, the acid form is mainly metabolized by CYP2Cs while the lactone form is metabolized by CYP3A4 (Fujino et al. 2004b), raising the possibility that CYP2Cs might be responsible for species difference in pitavastatin metabolism. Based on these lines of evidence, we decided to examine the involvement of CYP2C76 in pitavastatin metabolism to examine the hypothesis that CYP2C76 accounts for species difference in drug metabolism between monkeys and humans.

In the current study, we reveal that CYP2C76 is a critical factor in the species difference of pitavastatin metabolism between monkeys and humans. The lactone form of pitavastatin was metabolized differently between the two species, for which CYP2C76 was responsible.

## Materials and methods

### *Chemicals*

Pitavastatin and its metabolites such as pitavastatin lactone, M-3 (5-keto pitavastatin) and M-13 (8-hydroxylated pitavastatin) were synthesized by Nissan Chemical (Chiba, Japan). 4-Hydroxytolbutamide, 6$\alpha$-hydroxypaclitaxel, 3-hydroxypaclitaxel, 2$\alpha$-hydroxytestosterone, 16$\alpha$-hydroxytestosterone and 6$\beta$-hydroxytestosterone were purchased from Ultrafine Chemicals (Manchester, UK). The chemical purity of these parent compound and its metabolites were >98% during the experimental period. Benzoflavone, quercetin, quinidine, sulfaphenazole, ketoconazole, gemfibrozil and tranlycypromine were obtained from Ultrafine Chemicals, Wako Pure Chemical (Osaka, Japan), or Sigma-Aldrich (St Louis, MO, USA).

[Fluorobenzene-U-$^{14}$C] pitavastatin (specific radioactivity: 981 kBq mg$^{-1}$) was synthesized by Amersham Bioscience (Little Chalfort, UK). $^{14}$C-pitavastatin lactone and M-14 (5′,6′ or 7′,8′-dihydrodiol pitavastatin) were isolated according to our previous reports (Fujino et al. 2003; Yamada et al. 2003). [Ring-U-$^{14}$C] tolbutamide (2.26 GBq mmol$^{-1}$) and [4-$^{14}$C] testosterone (2.11 GBq mmol$^{-1}$) were purchased from Amersham Bioscience, whereas [2-benzoyl ring-U-$^{14}$C] paclitaxel (2.23 MBq mg$^{-1}$) was from Sigma-Aldrich. The radiochemical purity of the $^{14}$C-labelled chemicals was >99% during the experimental period.

### *Microsomes and antiserum for CYPs*

Pooled hepatic microsomes from human subjects (male and female mixed) and those of male cynomolgus monkeys were both purchased from BD-GENTEST (Woburn, MA, USA). The renal and intestinal microsomes of male cynomolgus monkeys were purchased from KAC (Kyoto, Japan). The human recombinant proteins derived from baculovirus expressing human CYP1A2, CYP2A6, CYP2C8, CYP2C9, CYP2C19, CYP2E1 and CYP3A4 were purchased from BD-GENTEST. Monkey CYP2C20, CYP2C43, CYP2C75 and CYP2C76 were expressed in *Escherichia coli* and subsequently purified as described before (Iwata et al. 1998; Daigo et al. 2002). Antisera for human CYP1A2, CYP2A6, CYP2C8, CYP2C9, CYP2C19, CYP2D6 and CYP3A4 were purchased from Nosan (Yokohama, Japan). $\beta$-NADP$^+$, glucose-6-phosphate, glucose-6-phosphate dehydrogenase and UDP-glucronic acid were obtained from Wako Pure Chemical.

### *Metabolic clearance of pitavastatin in hepatic microsomes*

The incubation conditions for NADPH-mediated metabolism or UGT-mediated lactonization were the same as described in our previous reports (Fujino et al. 1999b; 2004a). The experiments were performed with 2.5 μM of $^{14}$C-pitavastatin or its lactone form. The remaining acid and lactone forms in the incubation mixtures were measured for up to 60 min after the addition of hepatic microsomes to evaluate metabolic clearance ($CL_{int}$). *In vitro* $CL_{int}$ was determined using the following equation:

$$CL_{int} = k \times (\text{ml incubation})/(\text{mg protein of microsomes}),$$

where $k$ is the slope of the linear regression from the log concentration vs. incubation time relationships.

*Inhibition studies using anti-CYP antisera and chemical inhibitors*

An immuno-inhibition study was conducted according to our previous report (Fujino et al. 2004a). Monkey hepatic microsomes were pre-incubated at room temperature for 20 min with 40–100 $\mu l\,mg^{-1}$ microsomes of the goat antiserum for CYP1A2, CYP2A6, CYP2C8, CYP2C9, CYP2C19, CYP2D6, or CYP3A4. For control reactions, normal serum was used. Benzoflavone, quercetin, sulfaphenazole/gemfibrozil, tranylcypromine, qunidine and ketoconazole were used at 1–100 μM as relatively specific inhibitors of human CYP1A2, CYP2C8, CYP2C9, CYP2C19, CYP2D6 and CYP3A4, respectively. $^{14}$C-pitavastatin or its lactone form was co-incubated at 2.5 μM as a substrate with these inhibitors. The inhibitory effect was estimated from the remaining amount of the substrates in the presence or absence of the inhibitors. The $IC_{50}$ values were calculated by linear interpolation from the chemical inhibition study.

*Identification of pitavastatin-metabolizing CYPs*

To investigate the CYPs participating in pitavastatin metabolism, the partially purified recombinant CYP proteins for humans and monkeys were incubated with pitavastatin or its lactone form (both concentrations: 2.5 μM) for 30 min. Incubation with control microsomes from native insect cells was performed in parallel. The incubation conditions were the same as described above for the hepatic microsomes.

*Kinetic analysis of pitavastatin and CYP2C model substrates*

The CYP2C recombinant proteins and hepatic microsomes were used to perform the kinetic analysis of pitavastatin and several CYP2C model substrates in cynomolgus monkeys. The incubation conditions were essentially the same as the previous reports (Ludwig et al. 1998; Fujino et al. 2001; 2003). The Michaelis–Menten parameters of the apparent $K_m$ and $V_{max}$ were estimated by the Lineweaver–Burk reciprocal plots. The incubation was carried out for 30 min using the acid or lactone form at 1–80 μM as a substrate. The apparent $K_m$ and $V_{max}$ values were also calculated for 6α-hydroxypaclitaxel, 3-hydroxypaclitaxel, 4-hydroxytolbutamide, 2α-hydroxytestosterone, 16α-hydroxytestosterone and 6β-hydroxytestosterone. Paclitaxel, tolbutamide and testosterone were used at 2.5–40, 20–800 and 25–400 μM, respectively. Incubation with human hepatic microsomes was performed in parallel.

*Analytical methods*

The measurement of $^{14}$C-labelled pitavastatin and its metabolites was performed according to the HPLC-radioluminography (RLG) method (Kimata et al. 1998) with some modifications. Briefly, using the HPLC systems (1100 Series, Agilent Technologies, Palo Alto, CA, USA) equipped with a fraction collector, the HPLC eluate was fractionated at an interval of 15 s up to 25 min into polystyrene flat-bottomed microplates (RLG plate Beta48, Toyobo, Osaka, Japan). The plates were allowed to stand for 12 h at room temperature for evaporation of the solvent. The analysis of 6α-hydroxypaclitaxel and 3-hydroxypaclitaxel was performed according to our previous report (Fujino et al. 2001). Aliquots (approximately 2 μl) of the supernatant were spotted onto the TLC plates (Silicagel $60F_{254}$, 20 × 20 cm; Merck, Darmstadt, Germany) and developed with toluene–acetone–formic acid (60:39:1, v/v/v) to 12 cm in a horizontal TLC chamber that was saturated with solvent vapour.

The analysis of testosterone metabolites was carried out as follows: supernatant was applied to TLC plates and developed with dichloromethane–acetone (4:1, v/v) to 16 cm. The measurement of 4-hydroxytolbutamide was performed according to a previous report (Ludwig et al. 1998). Briefly, the supernatant was spotted and developed with toluene–acetone–formic acid (60:39:1, v/v/v) to 10 cm. The RLG and TLC plates were dried and placed in contact with a phosphor imaging plate for 12 h. The amounts of unchanged drug and metabolites were determined using the BAS-2500 (Fuji Photo Film, Tokyo, Japan) system. The radioactive metabolites were positively identified by a comparison of retention times or $R_f$ values using authentic unlabeled standard.

## Results

*Metabolic characterization of monkey CYP2Cs*

We have previously reported the involvement of monkey CYP2C20, CYP2C43, CYP2C75 and CYP2C76 in the metabolism of model substrates for human CYPs such as paclitaxel, tolbutamide, *S*-mephenytoin and testosterone (Uno et al. 2006). In the current study, we further characterized the metabolic properties of CYP2Cs by kinetic analysis using monkey and human hepatic microsomes (Table I) and the recombinant proteins for the monkey CYP2Cs (Table II).

We first evaluated the metabolic constants for paclitaxel, tolbutamide, *S*-mephenytoin and testosterone using human and monkey hepatic microsomes (Table I). The apparent $K_m$ values for tolbutamide 4-hydroxylation in humans and monkeys were 130 and 866 μM, respectively, indicating a weaker affinity for tolbutamide metabolism in monkeys than humans. In contrast, the apparent $V_{max}$ for *S*-mephenytoin 4-hydroxylation was 41 and 167 pmol $min^{-1}$ $mg^{-1}$ protein in human and monkey hepatic microsomes, respectively, suggesting a higher activity for *S*-mephenytoin 4-hydroxylation in monkeys than humans.

To identify which CYPs are responsible for the metabolism of each substrate in monkeys, a similar analysis was carried out using the recombinant monkey CYP2Cs (Table II).

Table I. Kinetic constants of testosterone, tolbutamide, paclitaxel and *S*-mephenytoin metabolism in human and monkey hepatic microsomes.

| Species | Substrate | Metabolite | Km (μM) | Vmax (pmol $min^{-1}$ $mg^{-1}$ protein) |
|---|---|---|---|---|
| Human | Testosterone | 6β-hydroxy testosterone | 109 | 4535 |
| | | 2α-hydroxy testosterone | 105 | 363 |
| | | 16α-hydroxy testosterone | 60.0 | 140 |
| | Tolbutamide | 4-hydroxy tolbutamide | 130 | 169 |
| | Paclitaxel | 6α-hydroxy paclitaxel | 6.3 | 83.0 |
| | | 3-hydroxy paclitaxel | 9.7 | 101 |
| | *S*-mephenytoin | 4-hydroxy mephenytoin | 97.2 | 41.0 |
| Monkey | Testosterone | 6β-hydroxy testosterone | 136 | 5643 |
| | | 2α-hydroxy testosterone | 687 | 3420 |
| | | 16α-hydroxy testosterone | 154 | 601 |
| | Tolbutamide | 4-hydroxy tolbutamide | 866 | 497 |
| | Paclitaxel | 6α-hydroxy paclitaxel | 9.6 | 60.0 |
| | | 3-hydroxy paclitaxel | 16.1 | 128 |
| | *S*-mephenytoin | 4-hydroxy mephenytoin | 75.8 | 167 |

Data are derived from duplicate determinations.

Table II. Kinetic constants of paclitaxel, tolbutamide and testosterone metabolism in the recombinant proteins for monkey CYP2Cs.

| Substrate | Enzyme | Metabolite | *K*m (μM) | $V_{max}$ (pmol $min^{-1}$ $nmol^{-1}$ P450) |
|---|---|---|---|---|
| Paclitaxel | CYP2C20 | 6α-Hydroxy paclitaxel | 2.2 | 15.1 |
| | | 3-Hydroxy paclitaxel | – | – |
| Tolbutamide | CYP2C75 | 4-Hydroxy tolbutamide | 775 | 863 |
| | CYP2C76 | 4-Hydroxy tolbutamide | 866 | 283 |
| Testosterone | CYP2C75 | 6β-Hydroxy testosterone | – | – |
| | | 2α-Hydroxy testosterone | 34.0 | 864 |
| | | 16α-Hydroxy testosterone | – | – |
| | CYP2C76 | 6β-Hydroxy testosterone | – | – |
| | | 2α-Hydroxy testosterone | 47.4 | 56.1 |
| | | 16α-Hydroxy testosterone | 42.4 | 152 |

Data are derived from duplicate determinations.

The apparent $K_m$ values for paclitaxel 6α-hydroxylation in monkey hepatic microsomes and CYP2C20 were 9.6 and 2.2 μM, respectively. For tolbutamide metabolism, the apparent $K_m$ values for tolbutamide 4-hydroxylation in monkey hepatic microsomes, CYP2C75 and CYP2C76 were 866, 755 and 866 μM, respectively. For testosterone metabolism, the apparent $K_m$ values for 2α-hydroxytestosterone in monkey hepatic microsomes, CYP2C75 and CYP2C76 were 687, 34.0 and 47.4 μM, respectively. Moreover, the apparent $K_m$ values for 16α-hydroxytestosterone in monkey hepatic microsomes and CYP2C76 were 154 and 42.4 μM, respectively. These $K_m$ values suggest that CYP2C20 is critically involved in the metabolism of paclitaxel, whereas CYP2C75 and CYP2C76 play major roles in the metabolism of tolbutamide and testosterone, respectively. However, the metabolic activity toward tolbutamide was 2.2-fold lower in monkeys than humans. We conclude that CYP2C76 apparently catalyses a different metabolic pattern from any other monkey CYP2Cs, but this does not explain the species difference observed in the metabolism of model substrates for human CYP2Cs between monkeys and humans. To show clearly the involvement of CYP2C76 in the species difference in drug metabolism, we therefore investigated pitavastatin metabolism.

*Metabolism of pitavastatin in hepatic microsomes from monkey and human hepatic microsomes*

NADPH- and UGT-mediated metabolic clearance of pitavastatin was determined using monkey and human hepatic microsomes (Table III). A larger clearance of the acid and/or lactone forms was seen in monkeys compared with humans. In addition, the appearance of the acid form hydrolysed from the lactone form was approximately 30% of the total amount in humans as compared with monkey hepatic microsomes (data not shown).

NADPH-mediated metabolism of the acid and lactone forms of pitavastatin is shown in Figure 1. The acid form underwent hydroxylation in monkey microsomes to afford at least two metabolic products, M-13 and M-14. In the case of the lactone form, several metabolites such as M-3, M-13 and M-14 were detected including unknown metabolites. Moreover, NADPH- and UGT-mediated metabolism were negligible in renal and intestinal microsomes of monkeys (data not shown).

Table III. Metabolic clearance of pitavastatin in monkey and human hepatic microsomes.

| | $CL_{int}$ ($\mu l\,min^{-1}\,mg^{-1}$ protein) | | |
|---|---|---|---|
| | Acid form | | Lactone form |
| Species | NAPDH-mediated | UGT-mediated | NAPDH-mediated |
| Monkey | 17.1 | 62.8 | 33.2 |
| Human | 3.3 | 5.6 | 3.5 |

Data are derived from duplicate determinations.

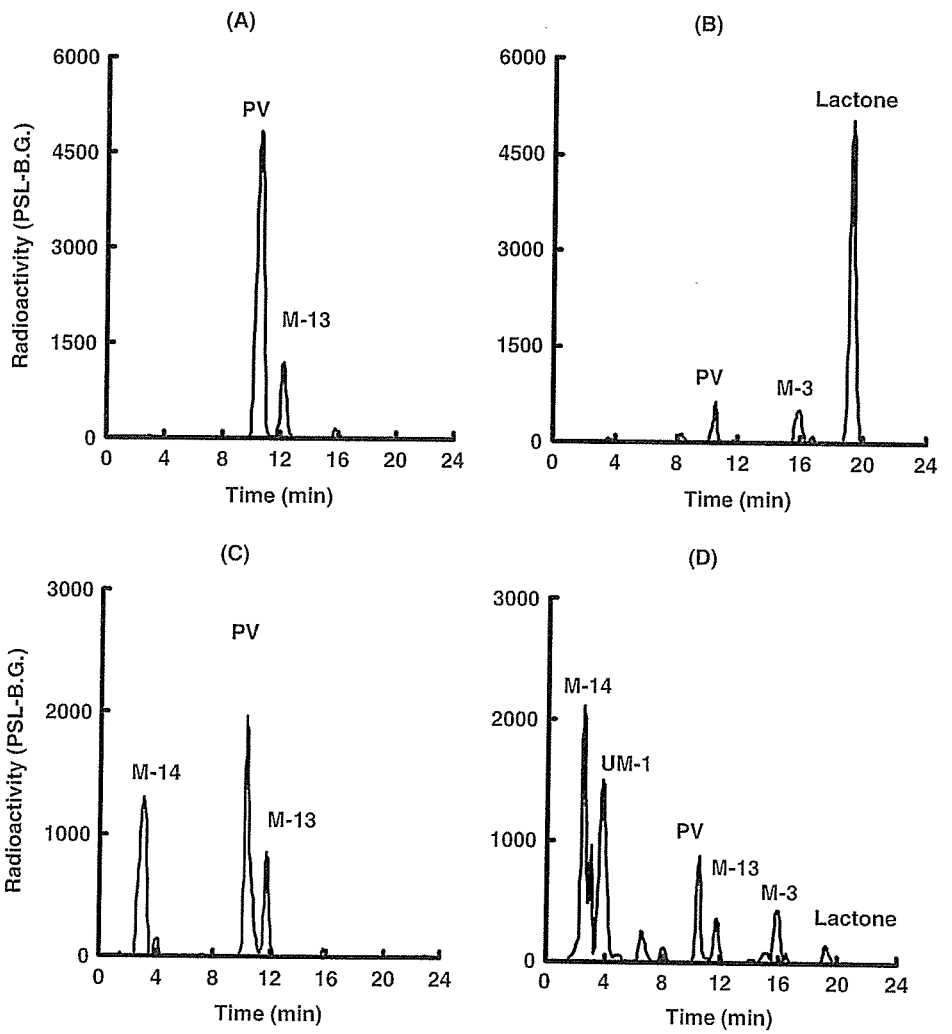


Figure 1. HPLC radiochromatograms of pitavastatin and its metabolites in human and monkey hepatic microsomes. The metabolic assays were performed with the acid (A and C) and lactone (B and D) forms of pitavastatin using human (upper) and monkey (lower) hepatic microsomes as described in the Materials and methods section. Results are representative of two separate experiments with each microsome sample.

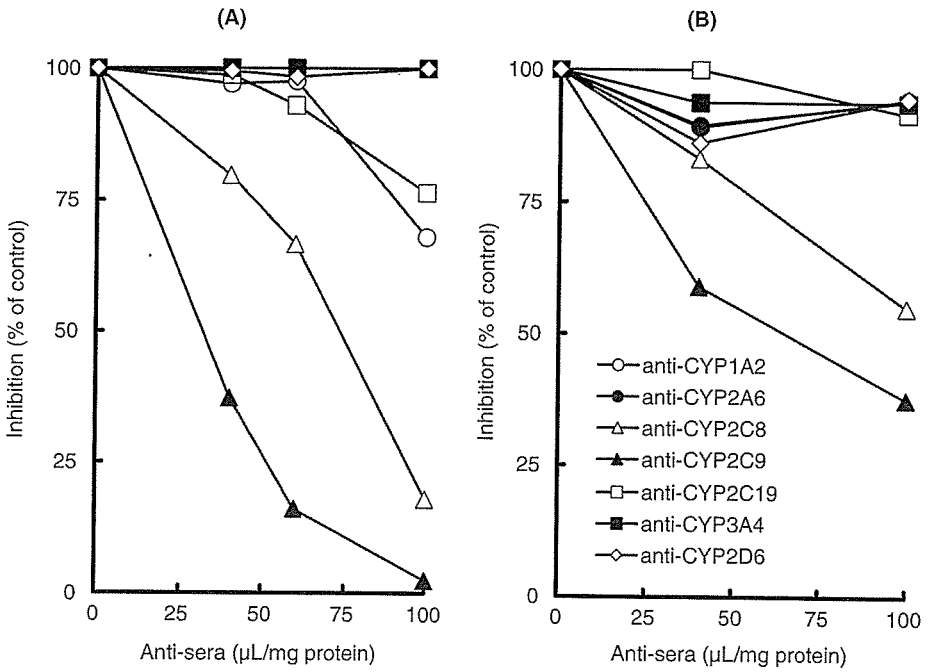


Figure 2. Immuno-inhibition of pitavastatin metabolism in monkey hepatic microsomes. Inhibition of the acid (A) and lactone (B) forms of pitavastatin was carried out using the antibodies for human CYP1A2, CYP2A6, CYP2C8, CYP2C9, CYP2C19, CYP2D6, or CYP3A4. Monkey hepatic microsomes were pre-incubated at room temperature for 20 min with 100 μl $mg^{-1}$ microsomes of goat antiserum for the CYPs. The acid or lactone form (2.5 μM) was then added and co-incubated in the reaction mixture. The inhibitory effect was estimated from the remaining pitavastatin and its lactone form in the presence and absence of the antibodies. Data are from duplicate experiments.

*Inhibition of pitavastatin metabolism in monkey hepatic microsomes by antibodies to human CYPs and typical CYP inhibitors*

The involvement of CYPs in the metabolism of the acid and lactone forms of pitavastatin was investigated in monkey hepatic microsomes using antisera and chemical inhibitors of human CYPs. Among antisera for human CYP1A2, CYP2A6, CYP2C8, CYP2C9, CYP2C19, CYP2D6 and CYP3A4 used with the acid form, only anti-CYP2C8 or CYP2C9 antisera strongly inhibited metabolite formation in monkey microsomes (Figure 2A). For the lactone form, a moderate or strong inhibition was observed in monkey microsomes by anti-CYP2C8 or CYP2C9 antisera, respectively (Figure 2B). These results indicate that monkey CYP2Cs are substantially involved in pitavastatin metabolism; however, critical site- or substrate-selective differences exist in the activities of CYP2Cs between monkeys and humans. Next, the inhibitory effects of several chemical inhibitors of human CYPs on the metabolism of pitavastatin were determined using benzoflavone, tranylcypromine, quercetin, sulfaphenazole, ketoconazole, gemfibrozil and quinidine (Table IV). For the acid form, the $IC_{50}$ values for tranylcypromine and ketoconazole were 22 and 64 μM, respectively, indicating the inhibitory effect of these inhibitors. In contrast, the metabolism of the lactone form was inhibited by benzoflavone and ketoconazole because the $IC_{50}$ values for these

Table IV. Chemical inhibition of pitavastatin metabolism in monkey hepatic microsomes.

| Inhibitors | $IC_{50}$ (μM) | |
|---|---|---|
| | Acid form | Lactone form |
| Benzoflavone | >100 | 29 |
| Tranylcypromine | 22 | 95 |
| Quercetin | >100 | >100 |
| Sulfaphenazole | >100 | >100 |
| Ketoconazole | 64 | 25 |
| Gemfibrozil | >100 | >100 |
| Quinidine | >100 | >100 |

Data are derived from duplicate determinations

inhibitors were 29 and 25 μM, respectively. A substantial difference in the inhibitory profiles of these chemical inhibitors supports the involvement of different CYP2Cs in the metabolism of each form of pitavastatin.

*Role of CYPs in the metabolism of pitavastatin in humans and monkeys*

The catalytic roles of each recombinant CYP on the metabolism of the acid and lactone forms of pitavastatin were examined using the recombinant proteins for human CYP1A2, CYP2A6, CYP2C8, CYP2C9, CYP2C19, CYP2E1 and CYP3A4 (Figure 3A). For monkeys, CYP2C20, CYP2C43, CYP2C75 and CYP2C76 were used (Figure 3B). The monkey CYP2E and CYP3A subfamilies were not included because the recombinant proteins for these CYPs were not available. In humans, CYP2C8 and CYP2C9 were shown to catalyse the metabolism of the acid form. In contrast, CYP3A4, not CYP2Cs, supported the metabolism of the lactone form. In monkeys, the acid form was metabolized by CYP2C43 and CYP2C75 while the lactone form was by CYP2C43, CYP2C75 and CYP2C76. CYP2C20 was not involved in pitavastatin metabolism. These observations indicate a major role of CYP2C76 in the metabolism of the lactone form, for which CYP3A4 is essential in humans, raising the possibility of CYP2C76 in the species difference in pitavastatin metabolism.

*Kinetic analysis of pitavastatin metabolism*

The metabolism of pitavastatin in monkey hepatic microsomes and the recombinant monkey CYP2Cs followed Michaelis–Menten kinetics as demonstrated by Lineweaver–burk plots, and the curves were all linear (data not shown). In the metabolism of the acid form, the apparent $K_m$ of hepatic microsomes, CYP2C43 and CYP2C75 for M-13 formation were 13.5, 129.1 and 8.8 μM, respectively, while the apparent $K_m$ of hepatic microsomes and CYP2C75 for M-14 formation was 13.4 and 1.2 μM, respectively (Table V). In the case of the lactone form, the apparent $K_m$ for M-13 formation was 7.1 μM for recombinant CYP2C75. On the other hand, CYP2C76 was involved in M-3, M-13 and M-14 formation with apparent $K_m$ values of 4.5, 24.9 and 36.6 μM, respectively. These results suggest that CYP2C75 is critically involved in the metabolism of the acid form and for M-13 formation from the lactone form. In contrast, CYP2C76 plays a major role in the formation of the several metabolites with relatively low $K_m$ values.