Dahl, P., Lindstrom, G., Wiberg, K., Rappe, C., 1995. Absorption of PCBs, PCDDs, and PCDFs by breast-fed infants. Chemosphere 30, 2297–2306.

Environment Agency. Report of survey on the exposure of dioxins in human (2000). http://www.env.go.jp/press/press.php3?serial=3160 (in Japanese.)

Faqi, A.S., Dalsenter, P.R., Merker, H.J., Chahoud, I., 1998. Reproductive toxicity and tissue concentrations of low doses of 2,3,7,8-tetrachlorodibenzo-p-dioxin in male offspring rats exposed throughout pregnancy and lactation. Toxicol. Appl. Pharmacol. 150, 383–392.

Fries, G.F., Paustenbach, D.J., 1990. Evaluation of potential transmission of 2,3,7,8-tetrachlorodibenzo-p-dioxin-contaminated incinerator emissions to humans via foods. J. Toxicol. Environ. Health 29, 1–43.

Giesy, J.P., Jude, D.J., Tillit, D.E., Gale, R.W., Meadows, J.C., Zajieck, J.L., Peterman, P.H., Verbrugge, D.A., Sanderson, J.T., Schwartz, T.R., Tuchman, M.L., 1997. Polychlorinated dibenzo-p-dioxins, dibenzofurans, biphenylles and 2,3,7,8-tetrachlorodibenzo-p-dioxin equivalents in fishes from Saginaw bay Michigan. Environ. Toxicol. Chem. 16, 713–724.

Guo, X., Longnecker, M.P., Michalek, J.E., 2001. Relation of serum tetrachlorodibenzo-p-dioxin concentration to diet among veterans in the Air Force Health Study with background-level exposure. J. Toxicol. Environ. Health A 63, 159–172.

Gray Jr., L.E., Kelce, W.R., Monosson, E., Ostby, J.S., Birnbaum, L.S., 1995. Exposure to TCDD during development permanently alters reproductive function in male Long Evans rats and hamsters: reduced ejaculated and epididymal sperm numbers and sex accessory gland weights in offspring with normal androgenic status. Toxicol. Appl. Pharmacol. 131, 108–118.

Gray, L.E., Ostby, J.S., Kelce, W.R., 1997. A dose-response analysis of the reproductive effects of a single gestational dose of 2,3,7,8-tetrachlorodibenzo- p-dioxin in male Long Evans hooded rat offspring. Toxicol. Appl. Pharmacol. 146, 11–20.

Hoover, S.M., 1999. Exposure to persistent organochlorines in Canadian breast milk: a probabilistic assessment. Risk Anal. 19, 527–545.

House, R.V., Lauer, L.D., Murray, M.J., Thomas, P.T., Ehrlich, J.P., Burleson, G.R., Dean, J.H., 1990. Examination of immune parameters and host resistance mechanisms in B6C3F1 mice following adult exposure to 2,3,7,8-tetrachlorodibenzo-p-dioxin. J. Toxicol. Environ. Health 31, 203–215.

Hurst, C.H., DeVito, M.J., Setzer, R.W., Birnbaum, L.S., 2000. Acute administration of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) in pregnant Long Evans rats: association of measured tissue concentrations with developmental effects. Toxicol. Sci. 53, 411–420.

Iida, T., Hirakawa, H., Matsueda, T., Takenaka, S., Nagayama, J., 1999. Polychlorinated dibenzo-p-dioxins and related compounds in breast milk or Japanese primiparas and multiparas. Chemosphere 38, 2461–2466.

Kirman, C.R., Hays, S.M., Kedderis, G.L., Gargas, M.L., Strother, D.E., 2000. Improving cancer dose-response characterization by using physiologically based pharmacokinetic modelling: an analysis of pooled data for acrylonitrile-induced brain tumors to assess cancer potency in the rat. Risk Anal. 20, 135–151.

Kitamura, K., Sunaga, M., Watanabe, S., Yamada, T., Hata, J., 1999. Concentration of polychlorinated dibenzo-p-dioxins and their related compounds in the human bile in relation to those in the liver and blood. Organohalogen Comp. 44, 165–168.

Kissel, J.C., Robarge, G.M., 1988. Assessing the elimination of 2,3,7,8-TCDD from humans with a physiologically based pharmacokinetic model. Chemosphere 17, 2017–2027.

Kociba, R.J., Keyes, D.G., Beyer, J.E., Carreon, R.M., Wade, C.E., Dittenber, D.A., Kalnins, R.P., Frauson, L.E., Park, C.N., Barnard, S.D., Hummel, R.R.A., Humiston, C.G., 1978. Results of a two-year chronic toxicity and oncogenicity study of 2,3,7,8-tetrachlorodibenzo-p-dioxin in rats. Toxicol. Appl. Pharmacol. 46, 279–303.

Kohn, M.C., Lucier, G.W., Clark, G.C., Sewall, C., Trischer, A.M., Portier, C.J., 1993. A mechanistic model of effects of dioxin on gene expression in the rat liver. Toxicol. Appl. Pharmacol. 120, 138–154.

Korner, W., Dawidowsky, N., Hagenmaier, H., 1993. Fecal excretion of PCDDs and PCDFs in two breast-fed infants. Chemosphere 27, 157–162.

Kreuzer, P.E., Csanady, G.A., Baur, C., Kessler, W., Papke, O., Greim, H., Filser, J.G., 1997. 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) and congeners in infants. A toxicokinetic model of human lifetime body burden by TCDD with special emphasis on its uptake by nutrition. Arch. Toxicol. 71, 383–400.

Lakind, J.S., Berlin, C.M., Park, C.N., Naiman, D.Q., Gudka, N.J., 2000. Methodology for chalacterizing distributions of incremental body burdens of 2,3,7,8-TCDD and DDE from breast milk in North American nursing infants. J. Toxicol. Environ. Health A 59, 605–639.

LaKind, J.S., Berlin, C.M., Naiman, D.Q., 2001. Infant Exposure to Chemicals in Breast Milk in the United States: what we need to learn From a breast milk monitoring program. Environ. Health. Perspect. 109, 75–88.

Lawrence, G.S., Gobas, F.A.P.C., 1997. A pharmacokinetic analysis of interspecies extrapolation in dioxin risk assessment. Chemosphere 35, 427–452.

Leung, H.W., Ku, R.H., Paustenbach, D.J., Andersen, M.E., 1988. A physiologically based pharmacokinetic model for 2,3,7,8-tetrachlorodibenzo-p-dioxin in C57BL/6J and DBA/2J mice. Toxicol. Lett. 42, 15–28.

Leung, H.W., Paustenbach, D.J., Andersen, M.E., 1989. A physiologically based pharmacokinetic model for 2,3,7,8-tetrachlorodibenzodioxin. Chemosphere 18, 659–664.

Liem, A.K.D, Theelen, R.M.C., 1997. Dioxins: chemical analysis, exposure and risk assessment. Ph.D. Dissertation, National Institute of Public Health and the Environment, P.O.Box 3720 BA, Bilthoven, The Netherlands.

Luebke, R.W., Copeland, C.B., Diliberto, J.J., Akubue, P.I., Andrews, D.L., Riddle, M.M., Williams, W.C., Birnbaum, L.S., 1994. Assessment of host resistance to *Trichinella spiralis* in mice following preinfection exposure to 2,3,7,8-TCDD. Toxicol. Appl. Pharmacol. 125, 7–16.

Maruyama, W., Yoshida, K., Tanaka, T., Nakanishi, J., 2002. Possible range of dioxin concentration in human tissues: simulation with a physiological based model. J. Toxicol. Environ. Health A 65, 2053–2073.

Maruyama, W., Yoshida, K., Tanaka, T., Nakanishi, J., 2003. Simulation of dioxin accumulation in human tissues and analysis of reproductive risk. Chemosphere 53, 301–313.

McLachlan, M., 1993. Digestive tract absorption of PCDDs, PCDFs, and PCBs in a nursing infant. Toxicol. Appl. Pharmacol. 123, 68–72.

McClelland, D.B., McGrath, J., Samson, R.R., 1978. Antimicrobial factors in human milk. Studies of concentration and transfer to the infant during the early stages of lactation. Acta Paediatr. Scand. Suppl. 271, 1–20.

National Nutrition Survey, Japan (J-NNS) 1998. http://www1.mhlw.go.jp/toukei/k-eiyou_11/t0225-1.html.

Nagayama, J., Okamura, K., Iida, T., Hirakawa, H., Matsueda, T., Tsuji, H., Hasegawa, M., Sato, K., Ma, H.Y., Yanagawa, T., Igarashi, H., Fukushige, J., Watanabe, T., 1998. Postnatal exposure to chlorinated dioxins and related chemicals on thyroid hormone status in Japanese breast-fed infants. Clearance of PCDD/Fs via gastrointestinal tract in occupationally exposed persons. Chemosphere 37, 1789–1793.

Neal, R.A., Olson, J.R., Gasiewicz, T.A., Geiger, L.E., 1982. The toxicokinetics of 2,3,7,8-tetrachlorodibenzo-p-dioxin in mammalian systems. Drug. Metab. Rev. 13, 355–385.

Ogawa, Y., Tada, H., Nakamura, H., Nishida, H., 2000. New Neonatal Study, second ed. Medicus Shuppan, Osaka, pp. 233–234. (In Japanese).

Ohsako, S., Miyabara, Y., Nishimura, N., Kurosawa, S., Sakaue, M., Ishimura, R., Sato, M., Takeda, K., Aoki, Y., Sone, H., Tohyama, C., Yonemoto, J., 2001. Maternal exposure to a low dose of

2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) suppressed the development of reproductive organs of male rats: dose-dependent increase of mRNA levels of 5α-reductase type 2 in contrast to decrease of androgen receptor in the pubertal ventral prostate. Toxicol. Sci. 60, 132–143.

Patandin, S., Dagnelie, P.C., Mulder, P.G., Op de Coul, E., van der Veen, J.E., Weisglas-Kuperus, N., Sauer, P.J., 1999. Dietary exposure to polychlorinated biphenyls and dioxins from infancy until adulthood: a comparison between breast-feeding, toddler, and long-term exposure. Environ. Health Perspect. 107, 45–51.

Paustenbach, D.J., 2000. The practice of exposure assessment: a state-of-the-art review. J. Toxicol. Environ Health B Crit. Rev. 3, 179–291.

Pitot, H.C., Campbell, H.A., Poland, A., 1980. Quantitative evaluation of the promotion by 2,3,7,8-tetrachlorodibenzo-p-dioxin of hepatocarcinogenesis from diethylnitrosoamine. Cancer Res. 40, 3616–3620.

Pluim, H.J., Koppe, J.G., Olie, K., van der Slikke, J.W., Slot, P.C., van Boxtel, C.J., 1994. Clinical laboratory manifestations of exposure to background levels of dioxins in the perinatal period. Acta Paediatr. 83, 583–587.

Portier, C.J., Sherman, C.D., Kohn, M., Edler, L., Kopp-Schneider, A., Maronpot, R.M., Lucier, G., 1996. Modeling the number and size of hepatic focal lesions following exposure to 2,3,7,8-TCDD. Toxicol. Appl. Pharmacol. 138, 20–30.

Rohde, S., Moser, A., Papke, O., McLachlan, M.S., 1999. Chemosphere 38, 3397–3410.

Rogan, W.J., Blanton, P.J., Portier, C.J., Stallard, E., 1991. Should the presence of carcinogens in breast milk discourage breast feeding? Regul. Toxicol. Pharmacol. 13, 228–240.

Sato, M., Araga, N., Suzuki, J., Hattori, R., Yagihashi, M., 1999. Child Health: Its Theory and Practice, fifth ed. Jusonbo Press, Tokyo, pp. 92–94. (In Japanese).

Schecter, A.J., Papke, O., Lis, A., Olson, J.R., 1995. Chlorinated dioxin, dibenzofuran and PCB levels in human fetal tissue at 8-18 weeks gestational age, compared with placental, newborn and adult tissue level. Organohalogen Comp. 25, 167–171.

Schecter, A., Ryan, J.J., Papke, O., 1998. Decrease in levels and body burden of dioxins, dibenzofurans, PCBs, DDE, and HCB in blood and milk in a mother nursing twins over a thirty-eight month period. Chemosphere 37, 1807–1816.

Schrey, P., Wittsiepe, J., Mackrodt, P., Selenka, F., 1998. Human fecal PCDD/F-excretion exceeds the dietary intake. Chemosphere 37, 1825–1831.

Snyder, W.S., Cook, M.J., Nasset, E.S., Karhausen, L.R., Howell, G.P., Tipton, I.H., 1974. Report of the task group on reference man. ICRP publication 23. The International Commission on Radiological Protection (ICRP). Pergamon, Elsevier Science, U.K., pp. 8–234.

Tada, H., Oda, S., Kitajima, T., Morita, M., Nakamura, K., 1999. Intake of PCDDs, PCDFs and CoPCBs in breast fed infants of Japan. Organohalogen Comp. 44, 271–273.

Teeguarden, J.G., Dragan, Y.P., Singh, J., Vaughan, J., Xu, Y.H., Goldsworthy, T., Pitot, H.C., 1999. Quantitative analysis of dose- and time-dependent promotion of four phenotypes of altered hepatic foci by 2,3,7,8-tetrachlorodibenzo-p-dioxin in female Sprague-Dawley rats. Toxicol. Sci. 51, 211–223.

Toyoda, M., Iida, T., Uchibe, H., Yanagi, T., Kono, Y., Hori, T., 1999. Total diet study report of daily intake of dioxins from food materials. Ministry of Health and Welfare of Japan. (In Japanese).

Tritscher, A.M., Goldstein, J.A., Portier, C.J., McCoy, Z., Clark, G.C., Lucier, G.W., 1992. Dose-response relationships for chronic exposure to 2,3,7,8- tetrachlorodibenzo- p-dioxin in a rat tumor promotion model - quantification and immunolocalization of CYP1A1 and CYP1A2 in the liver. Cancer Res. 52, 3436–3442.

US Environmental Protection Agency (US EPA). 2000. Exposure and human health reassessment of 2,3,7,8-Tetrachlorodibenzo-p-dioxin (TCDD) and related compounds. Part III: Integrated summary and risk characterization for 2,3,7,8-Tetrachlorodibenzo-p-dioxin (TCDD) and related compounds. EPA/600/P-00/001Bg.

US Environmental Protection Agency (US EPA). 2001. Dioxin reassessment—an SAB review of the office of research and development's reassessment of dioxin. Review of the revised sections (dose response modeling, integrated summary, risk characterization, and toxicity equivalency factors) of the EPA's reassessment of dioxin by the dioxin reassessment review subcommittee of the EPA science advisory board (SAB). EPA-SAB-EC-01-006

Van den Berg, M., Birnbaum, L., Bosveld, A.T.C., Brunstrom, B., Cook, P., Feeley, M., Giesy, J.P., Hanberg, A., Hasegawa, R., Kennedy, S.W., Kubiak, T., Larsen, J.C., Rolaf van Leeuwen, F.X., Liem, A.K.D., Nolt, C., Peterson, R.E., Poellinger, L., Safe, S., Schrenk, D., Tillitt, D., Tysklind, M., Younes, M., Waern, F., Zacharewski, T., 1998. Toxic equivalency factors (TEFs) for PCBs, PCDDs, PCDFs for humans and wildlife. Environ. Health Perspect. 106, 775–792.

Yang, Y.G., Lebrec, H., Burleson, G.R., 1994. Effect of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) on pulmonary influenza virus titer and natural killer (NK) activity in rats. Fundam. Appl. Toxicol. 23, 125–131.

Wang, X., Santostefano, M.J., Evans, M.V., Richardson, V.M., Diliberto, J.J., Birnbaum, L.S., 1997. Determination of parameters responsible for pharmacokinetic behavior of TCDD in female Sprague-Dawley ratrs. Toxicol. Appl. Pharmacol. 147, 151–168.

Weisglas-Kuperus, N., Patandin, S., Berbers, G.A., Sas, T.C., Mulder, P.G., Sauer, P.J., Hooijkaas, H., 2000. Immunologic effects of background exposure to polychlorinated biphenyls and dioxins in Dutch preschool children. Environ. Health Perspect. 108, 1203–1207.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biochimica et Biophysica Acta 1700 (2004) 151–159
BIOCHIMICA ET BIOPHYSICA ACTA
BBA
www.bba-direct.com

Identification, characterization, and site-directed mutagenesis of recombinant pentachlorophenol 4-monooxygenase☆

Takashi Nakamura[a,1], Takayuki Motoyama[a,1], Shuichi Hirono[b,2], Isamu Yamaguchi[a,*]

[a] *Laboratory for Remediation Research, Environmental Plant Research Group, Plant Science Center, RIKEN Institute, 1-7-22 Suehiro, Tsurumi, Yokohama City, Kanagawa 230-0045, Japan*
[b] *Laboratory of Physical Chemistry for Drug Design, School of Pharmaceutical Sciences, Kitasato University, 5-9-1 Shirokane, Minato, Tokyo 108-8641, Japan*

Received 19 December 2003; received in revised form 7 April 2004; accepted 13 April 2004
Available online 31 May 2004

Abstract

In a previous study, we constructed a three-dimensional (3D) structure of pentachlorophenol 4-monooxygenase (PcpB). In this study, further analyses are performed to examine the important amino acid residues in the catalytic reaction by identification of the proteins with mass spectrometry, circular dichroism (CD) and UV spectrometry, and determination of kinetic parameters. Recombinant histidine-tagged PcpB protein was produced and shown to have a similar activity to the native protein. Mutant proteins of PcpB were then produced (F85A, Y216A, Y216F, R235A, R235E, R235K, Y397A and Y397F) on the basis of the proposed 3D structure. The CD spectra of the proteins showed that there were no major changes in the structures of the mutant proteins, with the exception of R235E. Steady-state kinetics showed a 20-fold reduction in k_{cat}/K_m and a ninefold increase in K_m for Y216F and a threefold reduction in k_{cat}/K_m and a sixfold increase in K_m for Y397F compared to the wild type. On the other hand, the value of k_{cat}/K_m of R235K mutant was the same as that of wild type. As a result, it was confirmed that Y216 and Y397 play an important role with respect to the recognition of the substrate.

Keywords: Pentachlorophenol 4-monooxygenase (PcpB); Homology modeling; Site-directed mutagenesis; Mass spectrometry; Circular dichroism (CD) spectrometry; Steady-state kinetics

1. Introduction

Pentachlorophenol (PCP) is a polychlorinated aromatic compound that has been used as a pesticide and herbicide for many years, and its widespread use has caused contamination of soil and groundwater [1]. In addition, paper pulp mill bleach effluent contains large amounts of chlorinated phenols, including PCP [2]. Consequently, PCP is listed as a 'Priority Pollutant' by the Environmental Protection Agency because of its toxicity and widespread distribution in the environment [3]. Bioremediation of PCP contamination has therefore become an important focus of research.

Pentachlorophenol 4-monooxygenase (PcpB) has been purified from the PCP-degrading bacterium *Sphingobium chlorophenolicum* (formerly *Sphingomonas chlorophenolica*) strain ATCC39723 [4–6]. This enzyme is a flavin monooxygenase that converts PCP to tetrachlorobenzoquinone (TCBQ) in the presence of oxygen and a reduced form of nicotinamide adenine dinucleotide phosphate (NADPH) [7]. Although the nucleotide sequence of the *pcpB* gene has already been determined and recombinant protein produced [8,9], its three-dimensional (3D) structure has not yet been solved. In fact, little information is currently available regarding the structure and catalytic mechanism of this enzyme.

Computational chemistry and molecular modeling offer effective tools to study the mechanisms of biodegradation reactions at the molecular level [10–12], and can provide guidelines for mutagenesis experiments. PcpB was selected firstly because it appears to catalyze the rate-limiting step in

☆ Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bbapap.2004.04.008.

* Corresponding author. Tel.: +81-45-503-9488; fax: +81-45-503-9489.
E-mail addresses: tnakamu@postman.riken.go.jp (T. Nakamura), tmotoyam@postman.riken.go.jp (T. Motoyama), hironos@pharm.kitasato-u.ac.jp (S. Hirono), yamaism@postman.riken.go.jp (I. Yamaguchi).
[1] Tel.: +81-45-503-9488; fax: +81-45-503-9489.
[2] Tel.: +81-3-3443-7780; fax: +81-3-3440-5246.

1570-9639/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbapap.2004.04.008

the biodegradation of PCP [3,13] and therefore the possibility exists of producing an enzyme which can degrade PCP more effectively. Secondly, PcpB is strikingly nonspecific, unlike most flavin monooxygenases. The enzyme can turn over a variety of substituted phenols and can replace hydrogen, nitro, amino, and cyano groups, as well as halogens, with a hydroxyl group [14]. Consequently, it appears that PcpB could be altered to break down synthetic organic pollutants such as polychlorinated biphenyls (PCBs) and 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) that natural microorganisms currently are unable to degrade.

In the previous study [15], we constructed the 3D structure of PcpB by homology modeling, where the X-ray crystal structures of phenol hydroxylase [16] and *p*-hydroxybenzoate hydroxylase (PHBH) [17] were used as templates (Fig. 1). However, experimental data for the model, such as identification of the proteins with mass spectrometry, biophysical characterization with CD and UV spectrometry, and determination of kinetic parameters, were not shown sufficiently. In this article, we show the results of these experimental data and examined the important amino acid residues in the catalytic reaction.

2. Materials and methods

2.1. Bacterial strains, plasmids, and culture conditions

The vector pET-19b (Novagen, Madison, WI, USA), which carries an N-terminal histidine tag sequence, followed by an enterokinase site and three cloning sites, was used for the cloning and expression of recombinant PcpB. The vector pKF18K (TaKaRa Shuzo, Kyoto, Japan) was used for construction of PcpB mutants. *Escherichia coli*

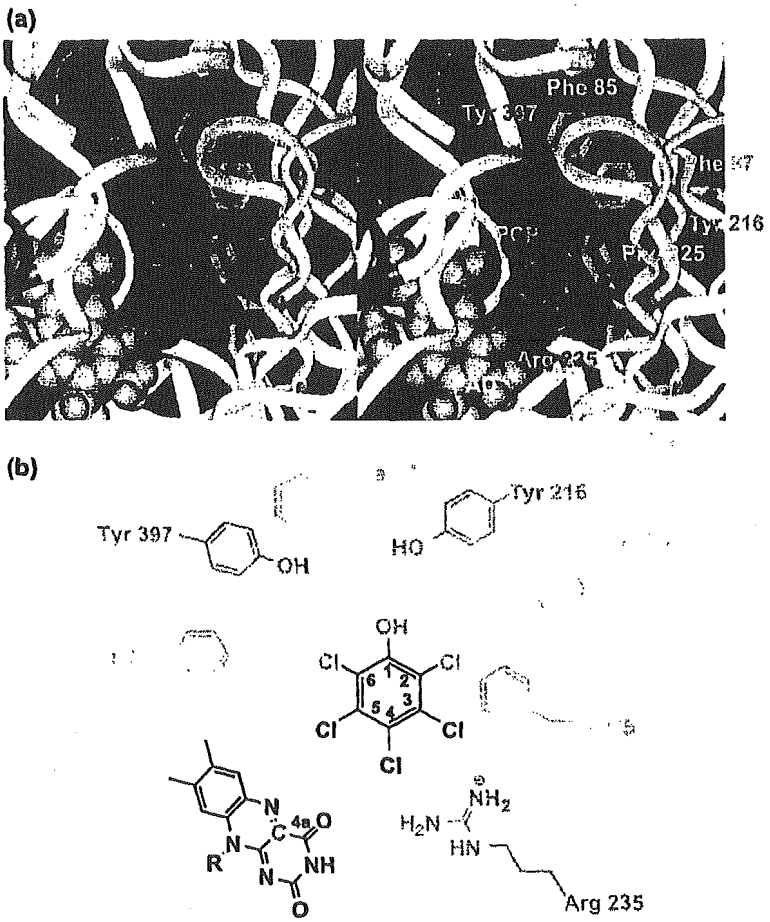


Fig. 1. The 3D structure of PcpB, constructed by molecular modeling [15]. (a) Stereo view of the active site of PcpB. Important residues, predicted from inspection, are shown and distinguished with different colors: tyrosine, phenylalanine, and arginine are shown in pink, light green, and blue, respectively. PCP and FAD are shown in orange and yellow CPK models, respectively. The red atom in FAD shows the C4a atom to which molecular oxygen is likely added. The color of the ribbon corresponds to the secondary structure elements calculated by the Kabsch–Sander method [27]. The regions of red, yellow, and light blue refer to α-helix, β-sheet, and random coil, respectively. (b) A schematic of the active site of PcpB shown in (a). The blue atoms in PCP are important for PcpB recognition. See on-line version of figure for colour. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

strain DH5α (Toyobo, Osaka, Japan) was cultured in Luria-Bertani (LB) medium [18] and *E. coli* strain BL21 (DE3), pLysS (Novagen) was cultured in Terrific Broth (TB) [18] at 37 °C with vigorous shaking. Media were solidified with 1.5% Bacto agar (Difco Laboratories, Detroit, MI, USA) and, where appropriate, supplemented with kanamycin or carbenicillin at final concentrations of 200 and 50 μg/ml, respectively. Restriction enzymes were purchased from TaKaRa Shuzo. Primers for site-directed mutagenesis were purchased from Sigma Genosys Japan (Hokkaido, Japan).

2.2. Expression, purification, and determination of molecular weight of recombinant enzymes

Cloning and generation of recombinant plasmids was carried out using standard methods and as detailed previously [15,18]. Mutant constructs were prepared by Mutan®-Express Km kit (TaKaRa Shuzo). Following sequencing, the correct mutant PcpB expression plasmids were transformed into *E. coli* BL21 (DE3) pLysS strain for protein expression. The cells containing expression vectors were grown in TB medium in the presence of carbenicillin (200 μg/ml) at 37 °C until the optical density (OD_{600}) of the medium reached 0.6. The expression of the desired protein was induced by adding isopropyl β-D-thiogalactopyranoside and carbenicillin to final concentrations of 100 and 500 μM, respectively. The cells were incubated at 20 °C overnight, and all purification steps were performed at 4 °C. Purification of the proteins was achieved by affinity chromatography using Ni-NTA agarose (Qiagen K. K., Tokyo, Japan). Details of these procedures were reported previously [15].

DNASIS-Mac (v3.7, Hitachi Software, Engineering, Co, Ltd, Tokyo, Japan) was used for the determination of molecular weight of recombinant enzymes. The peptide sequence predicted from the His-tag component gene, which is MGHHHHHHHHHHSSGHIDDDDKH, was added to the component of PcpB. Considering removal of the N-terminal fMet by methionyl aminopeptidase, the averaged molecular weights of the recombinant enzymes were calculated with and without the N-terminal methionine. The calculated molecular weights are as follows: Native PcpB (without His-tag), 59989.9, Wild type, 62756.6, F85A, 62680.5, Y216F, Y397F, 62740.6, Y216A, Y397A, 62664.5, R235A, 62671.5, R235K, 62728.6, and R235E, 62729.6.

2.3. Digestion of PcpB and its mutants by trypsin or lysyl endopeptidase (LysC)

Trypsin (Sequencing Grade Modified Trypsin) and LysC from *Achromobacter lyticus* M497-1 were purchased from Promega (Madison, WI, USA) and Wako Pure Chemicals (Osaka, Japan), respectively.

Fragmentation of PcpB or its mutants was performed using proteases (trypsin and LysC) [19,20]. 23 μl of the solution containing 120 μg of the protein in 100 mM ammonium hydrogen carbonate (NH_4HCO_3) buffer (pH 8.0) was incubated with 50 μl of 10 mM DTT in 100 mM NH_4HCO_3 buffer at 50 °C for 15 min. The solution was then cooled to room temperature and 2 μl of 250 mM iodoacetamide in 100 mM NH_4HCO_3 was added to the solution. The mixture was incubated in the dark at the room temperature for 15 min. Digestion of the proteins by trypsin was started with the addition of 5 μl of 100 ng/μl trypsin in 50 mM acetic acid. The proteins were digested at 37 °C overnight. In the case of LysC digestion, Tris–HCl buffer (pH 9.5) was used instead of NH_4HCO_3 buffer.

2.4. Mass spectrometric analysis of peptides

Monoisotopic mass values of all peptides were measured by a Biflex III matrix-assisted laser desorption/ionization (MALDI)-time-of-flight (TOF) mass spectrometer (Bruker daltonics, Bremen, Germany) operated in the reflector mode when the objective mass is less than 3000, otherwise averaged mass values were measured in the linear mode. Between 60 and 160 accumulations were analyzed for each spectrum, and peptides were observed as $(M+H)^+$ in the positive mode. Peptides produced by digestion of PcpB or its mutants with LysC were desalted and fractionated by ZipTip C18 (Millipore, Billerica, MA, USA) as described below. A 100 pmol sample was adsorbed in the ZipTip C18 at first. The tip was washed with 2.5% acetonitrile/0.1% trifluoroacetic acid (TFA) in water (HPLC grade, Kanto Kagaku, Tokyo, Japan) for deionization, and then the peptides were eluted utilizing from 10% to 80% acetonitrile and 0.1% TFA and fractionated using 10% acetonitrile. Peptide solution (0.5 μl) was co-crystallized with equal volumes of a saturated α-cyano-4-hydroxycinnamic acid matrix in 50% acetonitrile/0.1% TFA. All MALDI spectra were calibrated externally using a peptide standard.

2.5. Sequence analysis of the peptides by liquid chromatography/electrospray ionization ion trap mass spectrometry (LC/ESI-MS/MS)

LC/ESI-MS/MS analysis was performed using LCQ-DECA XP ion trap mass spectrometer (Thermo Finnigan, San Jose, CA, USA) and QSTAR quadrupole orthogonal acceleration TOF mass spectrometer (AB/MDS Sciex, Toronto, Ontario, Canada). QSTAR was used for R235K and R235E due to their extremely close sizes (only 1-Da difference between them). The analysis of the other enzymes was performed by LCQ. The LC/ESI-MS/MS data acquisition program was set up to collect ions signal from the eluted peptides using an automatic data-dependent scan procedure with a cyclic series of two different scan modes (full scan and MS/MS scan). First, the minimum MS signal ($>1-5\times10^5$ counts) in a full scan (m/z 400–2000) was selected as the precursor ion, or a list of parent ions was selected and then an MS/MS scan was performed to confirm the sequence of the precursor ion

using collision-induced dissociation (CID) with a relative collision energy of 35%. MASCOT (Matrix Science Inc., London, UK) was used for sequence analysis of the peptides. The LC/MS analysis was conducted using a MAGIC 2002 HPLC system (Michrom Bioresources Inc., Aubum, CA, USA) coupled to the LCQ, with a capillary Cadenza C18 custom-packed column (50 × 0.2-mm i.d., Michrom), in which Cadenza CD-C18 resin (Imtakt Corporation, Kyoto, Japan) was packed, at a flow rate of 1 μl/min. The enzymatically digested peptides were separated using mobile phase A and B with a linear gradient of 5% to 75% over 50 min (F85A, Y216A, R235A, and Y397F), 5% to 65% over 20 min (R235K and R235E), or 15% to 85% over 35 min (Y216F). Mobile phases A and B were 2% acetonitrile and 0.1% formic acid in water and 90% acetonitrile and 0.1% formic acid in water, respectively.

2.6. Circular dichroism (CD) spectra measurement

The CD spectra of the wild-type PcpB and its mutants in their substrate-free state were measured using a Jasco J-720 spectropolarimeter (Jasco, Tokyo, Japan) at 25 °C using quartz cells (Tosoh Quartz Corp., Yamagata, Japan) with a pathlength of 0.2 cm. Low noise CD spectra were measured by averaging 10 scans and the final spectra were corrected by subtracting the corresponding baseline of buffer (80 mM potassium phosphate buffer (pH 7.0)). The enzymes were dissolved at a concentration of 3.6 mM (200 μg, the mean residue molar concentration) in a 500-μl solution containing 80 mM potassium phosphate buffer (pH 7.0). CD data were transformed into molar ellipticity [θ] in the units of degree cm^2/dm of mean residue molar concentration. The mean residue molecular weight of the enzymes was calculated as 112.

2.7. UV–visible spectra measurement

The UV–visible spectra of the wild-type PcpB and its mutants in their substrate-free state were measured using a UV-2500 UV/VIS spectrophotometer (Shimadzu, Kyoto, Japan) at 25 °C. The final spectra were corrected by subtracting the corresponding baseline of buffer (80 mM potassium phosphate buffer (pH 7.0)). The enzymes were dissolved at a concentration of 6.4 μM (400 μg) in a 500-μl solution containing 80 mM potassium phosphate buffer (pH 7.0).

2.8. Determination of steady-state kinetic parameters

The activities of wild-type PcpB and its mutants were assayed by measuring the production of tetrachlorohydroquinone (TCHQ), as previously reported [5,15]. One milliliter of enzyme solution was made from wild-type PcpB or its mutants, and stabilized with 25% glycerol (v/v), and 10 mg/ml BSA (Fraction V, crystalline, Calbiochem-Novabiochem Co., La Jolla, CA, USA). The concentration of the enzymes was as follows: Wild type, Y397F, and R235K, 2 μM, and Y216F, 8 μM. The assay system contained 80 mM potassium phosphate (KPi) buffer (pH 7.0), 100 μM NADPH, and 5 μl of enzyme solution in a volume of 100 μl. The final concentration of the enzymes was as follows: Wild type, Y397F, and R235K, 0.1 μM, and Y216F, 0.4 μM. The PCP degradation reaction was carried out for 90 s in the case of wild-type, for 60 s in the case of R235K, and 120 s in the case of Y216F and Y397F, at a constant temperature of 25 °C to determine the initial velocities. The extent of reaction by these enzymes was less than 10% under these conditions. The values of K_m and V_{max} for PCP were determined with a variable concentration of PCP, from 0.4 to 20 μM. After the reaction, the solutions were acidified with 2 mol/l HCl, extracted with ethyl acetate, acetylated with pyridine and acetic anhydride, and analyzed by GC-MS as reported previously [15]. Experiments were repeated three times.

Initial velocities were fitted to Eq. (1) where v is the observed initial velocity of the reaction, V_{max} is maximum velocity, S is the substrate concentration, and K_m is the Michaelis constant. The curve fitting was carried out by using Delta Graph (ver 4.5, SPSS Inc., Chicago, IL, USA).

$$v = V_{max}S/(K_m + S) \quad (1)$$

2.9. Chemicals

PCP, pentachlorothioanisole (PCTA), external standard for GC-MS analysis, and ethyl acetate were purchased from Wako. These chemicals were of residual pesticide analysis grade, and the purities of PCP and PCTA were more than 99%. All other chemicals were of analytical grade and of the highest purity available.

3. Results

3.1. Comparison of enzyme activity of the recombinant wild-type PcpB with that of the previously reported native PcpB

Enzyme activity of the recombinant wild-type PcpB in this study was compared to that of PcpB from *S. chlorophenolicum* ATCC39723 [5] to show that the enzyme was catalytically competent [9]. In the previous report [5], native PcpB converted PCP to TCHQ completely. Recombinant PcpB, which has a histidine tag at the N-terminal, converted 72 ± 4.9% and 89 ± 2.6% of PCP in the absence and presence of 10 mg/ml of BSA, respectively. The reaction was confirmed to be enzymatic and NADPH-dependent (supporting information Fig. S1). As a result, it was found that the recombinant protein has qualitatively same activity as wild-type enzyme in the presence of BSA and that the N-terminal histidine tag did not affect recombinant-enzyme activity.

Table 1
ESI-MS analysis of trypsin-digested PcpB mutants

Enzyme	Sequence[a]	Calculated mass	Measured mass	Expected mass[b]	ΔM (Da)
Wild type	SNGFTFNFENTDAKPLLDFSVLPGR	2787.08	930.10 (3+)	2787.27	0.19
F85A	SNGATFNFENTDAKPLLDFSVLPGR	2710.98	1356.48 (2+)	2710.95	−0.03
Wild type	ECLGIAYEGEDYEENVLQMMDVGIQDFEAGDDWIHYFIGQDK	4944.36	1648.86 (3+)	4943.54	−0.82
Y216A	ECLGIAYEGEDYEENVLQMMDVGIQDFEAGDDWIHAFIGQDK	4852.26	1618.59 (3+)	4852.73	0.47
Y216F	ECLGIAYEGEDYEENVLQMMDVGIQDFEAGDDWIHFFIGQDK	4928.36	1643.81 (3+)	4928.39	0.03
Wild type	FVFVTK	739.91	740.67	739.66	−0.25
R235A	LPGSNYAVIISDLGGANK	1789.02	895.52 (2+)	1789.03	0.01
R235K	LPGSNYK	777.40	778.40	777.39	−0.01
R235E	LPGSNYEVIISDLGGANK	1845.95	923.90 (2+)	1845.79	−0.16
Wild type	VSGMSYNYR	1076.19	539.16 (2+)	1076.30	0.11
Y397F	VSGMSYNFR	1060.19	531.10 (2+)	1060.19	0.00

[a] The sequences mutated in the experiment are shown in boldface.
[b] The expected mass was calculated based on the measured mass.

3.2. *Identification of wild-type PcpB and its mutants by mass spectrometry*

From inspection of the active site residues in PcpB, seven residues seemed to be in direct contact with either the substrate molecule or with each other, as shown in Fig. 1. Y216 and Y397 were selected because of the potential of these tyrosines to make hydrogen bonds between the enzyme and the hydroxyl group of the substrate. The positively charged guanidinium group of R235 seemed to stabilize the chloride ion released from PCP during the enzymatic reaction. F225 possibly made a $\pi-\pi$ interaction with PCP. The phenyl ring edge of F87 seemed to interact with the π face of Y397. Finally, F85 and F223 seemed to be involved in the recognition of the *ortho* chloride of PCP, and was therefore selected, because PcpB could convert the substrate for both dichlorinated and trichlorinated phenols as long as one of the *ortho* positions was chlorinated [14]. Alanine mutants of these seven residues were constructed to observe the effect of the residue substitution. Y216F and Y397F were constructed to examine the importance of hydroxy group of tyrosine. R235E and R235K were constructed to observe the effect of the arginine positive charge. In the previous report, F223A, F225A, R235K, and Y397F had greater activity than wild type and the others had no (F87A, Y216A, R235A, R235E, Y397A) or weaker (F85A and Y216F) activity than wild type [15]. It was therefore not deemed necessary to derive further analysis of F223A and F225A in the light of previous data showing that they were not important for catalytic reaction. F87A was not also analyzed in this study because it does not seem to be correlated with catalytic reaction directly, although it may be important for stabiliza-

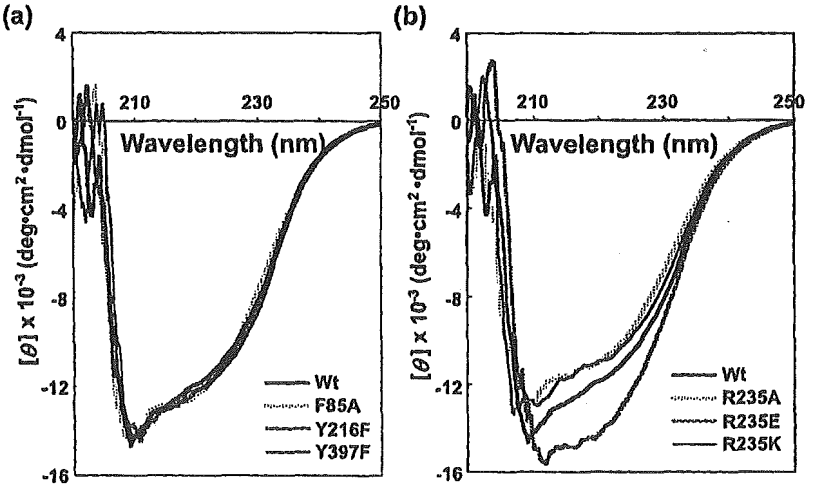


Fig. 2. Far-UV circular dichroism (CD) spectra of recombinant wild-type PcpB and its mutants. Low noise CD spectra were measured by averaging 10 scans and the final spectra were corrected by subtracting the corresponding baseline of buffer (80 mM potassium phosphate buffer (pH 7.0)). The enzymes were dissolved at a concentration of 3.6 mM (200 μg, the mean residue molar concentration) in a solution (500 μl) containing 80 mM potassium phosphate buffer (pH 7.0) and measured at 25 °C. CD data were transformed into molar ellipticity $[\theta]$ in the units of degree cm^2/dm of mean residue molar concentration. The mean residue molecular weight was estimated as 120. (a) CD spectra of Wild type (Wt), F85A, Y216F, and Y397F. (b) CD spectra of Wild type (Wt), R235A, R235E, and R235K.

tion of the enzyme structure. As a result, eight mutants (F85A, Y216A, Y216F, R235A, R235E, R235K, Y397A, and Y397F) were selected for further analysis.

To examine whether the recombinant wild-type PcpB and its mutants were expressed correctly, we analyzed the purified proteins by MALDI-TOF mass spectrometry and LC/ESI MS/MS. The purified proteins were digested with trypsin or LysC and the resultant peptides were analyzed. Of the whole digested fragments with LysC, 88–96% could be identified and the mutated sites were included in them by MALDI-TOF mass spectrometry (Supporting information Tables S1 and S2.

Sequence coverage of PcpB digested with trypsin by LC/ESI-MS/MS analysis was 72%. The mass of all peptides containing the mutated sites could be identified as shown in Table 1 and, moreover, almost all kinds of peptides of the mutants (Y216A, Y216F, R235A, R235K, R235E, and Y397F) could also be identified by sequence analysis (Supporting information Figs. S2–S7). Therefore, we confirmed that wild-type and all mutants of PcpB except Y397A are expressed correctly and contained in the purified fraction.

3.3. CD spectra of PcpB and its mutants

To examine whether the conformational change of the proteins is responsible for the difference among relative activities of PcpB and its mutants, CD spectra in the

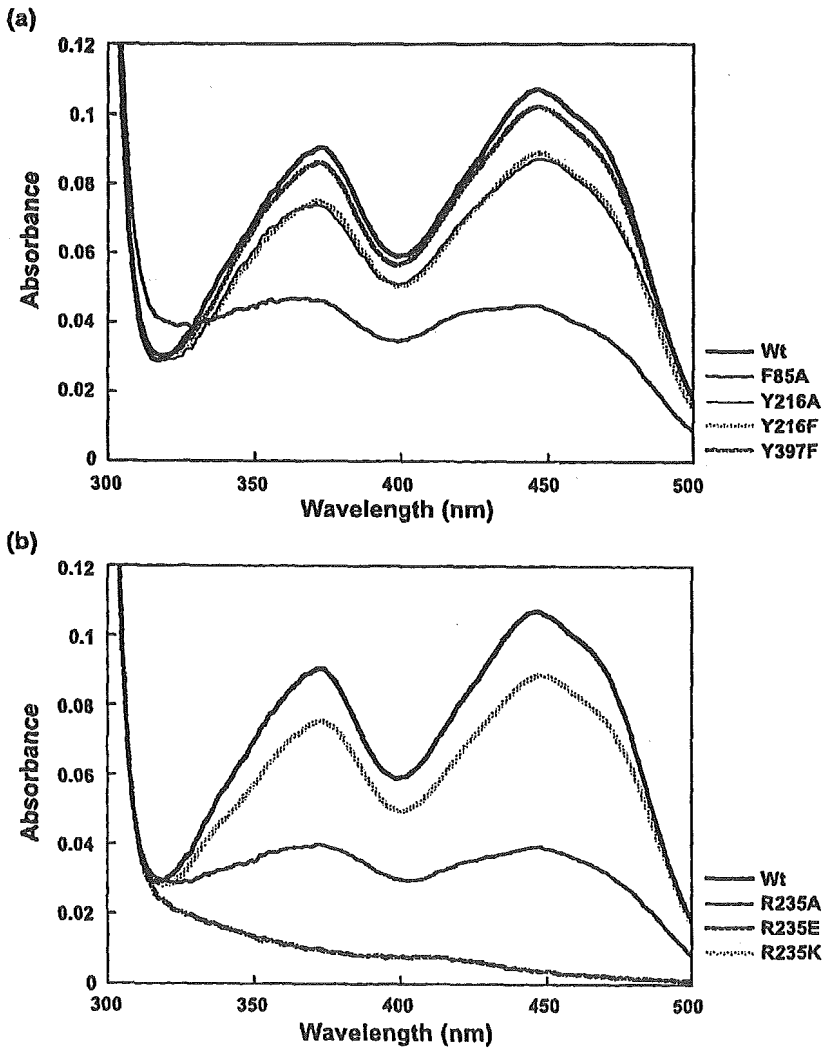


Fig. 3. Absorption spectra of recombinant wild-type PcpB and associated mutants. The final spectra were corrected by subtracting the corresponding baseline of buffer (80 mM potassium phosphate buffer (pH 7.0)). The enzymes were dissolved at a concentration of 6.4 M (400 μg) in a solution (1 ml) containing 80 mM potassium phosphate buffer (pH 7.0). Spectra were measured at 25 °C. (a) Absorption spectra of Wild type (Wt), F85A, Y216A, Y216F, and Y397F. (b) Absorption spectra of Wild type (Wt), R235A, R235E, and R235K.

substrate-free state were obtained. The spectra were classified by the amino acid residues such as phenylalanine and tyrosine (Fig. 2a), and arginine (Fig. 2b), respectively. Apparent changes in the CD spectrum (210–230 nm) were only observed in the case of R235E. This result shows that with the exception of R235E, there were no major changes in the CD spectra between wild type and mutants (including Y216A, of which the spectrum was shown in Fig. S8a (supporting information)). The near UV-CD spectrum (250–300 nm) of Y397A was quite different from that of wild type (Fig. S8b (supporting information)), however, no definitive conclusions can be drawn due to the problems involved in obtaining an accurate mass for Y397A.

3.4. Absorption spectra of PcpB and its mutants

To investigate whether FAD, the co-factor of PcpB, is correctly positioned in the protein, absorption spectra of the proteins were measured. The spectra were classified same as shown in Fig. 2. The typical absorption spectra of PcpB derived from FAD at 446 and 370 nm [5] were not seen in the case of inactive PcpB mutants such as R235E (Fig. 3b) and Y397A (Fig. 8c (supporting information)). Mutants, which show a weaker absorption spectra derived from FAD, had no (R235A, Fig. 3b) or weaker activity than wild-type (F85A, Fig. 3a). On the other hand, the absorption spectra of Y216A and Y216F, which have no or weaker activity than wild type, as well as R235K and Y397F, which have stronger activity than wild type, were similar to that of the wild type. These results suggest that the strength of absorption spectra derived from FAD corresponds well to the activity of the mutants except Y216 mutants.

3.5. Steady-state kinetic parameters

Steady-state kinetic parameters for the PcpB-catalyzed oxidation of PCP are listed in Table 2 for the wild-type and mutant PcpB enzymes analyzed by GC-MS. Enzyme activity assays were also performed spectrophotometrically, in a similar manner to that reported for PHBH, a homologue of PcpB [21]. However, we were unable to obtain the correct plot of the initial velocity versus the substrate concentration.

Constants could not be derived for R235E and Y397A due to the inactivity of the proteins, even if the reaction time was extended up to 30 min and the enzyme concentration was increased fivefold (i.e. five times greater than stated in Table 2). In the case of Y216A and R235A, PCP degradation was observed only when the enzyme concentration and reaction time were increased. However, the kinetics of the reaction did not conform to the classical Michaelis–Menten equation (Eq. (1)).

In the case of F85A, the kinetic parameters were obtained only when the amount of substrate was not in excess with respect to the enzyme. Under these conditions, the apparent K_m, k_{cat} and k_{cat}/K_m of wild type and F85A for PCP were 0.6 μM, 5.5 s^{-1}, and 8.8 μM^{-1} s^{-1} for wild type and 3.3 μM, 0.9 s^{-1}, and 0.3 μM^{-1} s^{-1} for F85A. Thus, there was an observed sixfold increase in K_m, and k_{cat} and k_{cat}/K_m decreased by sixfold and 29-fold, respectively, as a consequence of replacing phenylalanine with alanine at position 85. As shown in Table 2, replacing tyrosine with phenylalanine at position 216 results in a 20-fold reduction in k_{cat}/K_m and a ninefold increase in K_m, whilst the Y397F mutant showed k_{cat}/K_m reduced by threefold and a sixfold increase in K_m. On the other hand, the value of k_{cat}/K_m of R235K mutant was as same as that of wild type.

Table 2
Steady-state kinetic parameters for PcpB and mutants with substrate PCP at pH 7.0 and 25 °C[a]

Enzyme	$k_{cat} \times 10^3$ (s^{-1})	$k_{cat}/K_m \times 10^3$ (μM^{-1} s^{-1})	K_m (μM)
Wild type	9.3 ± 0.4	19.0 ± 1.1	0.5 ± 0.05
Y216F	4.1 ± 0.3	0.9 ± 0.05	4.5 ± 0.47
R235K	15.2 ± 1.4	19.0 ± 1.6	0.8 ± 0.15
Y397F	16.0 ± 0.3	5.5 ± 0.2	2.9 ± 0.17

[a] Data are shown as means standard deviation and reported values were determined for triplicate experiments.

4. Discussion

It has previously been reported that the K_m, V_{max}, and k_{cat}/K_m of the native PcpB were 30 ± 7 μM, 16 ± 5 μmol/min/mg of protein, and 405 s^{-1} mM^{-1}, respectively [5]. The authors state that the concentration of PcpB was 5 μM (15 μg PcpB in 50 μl of reaction mixture) whilst the concentration of PCP ranged from 20 to 50 μM. Therefore, since the substrate was not in excess of the enzyme, this could raise doubt over the reliability of the kinetic parameters obtained. Thus, these conditions are not strictly steady state and therefore would produce inaccurate results. In fact, we examined the kinetic analysis under the same conditions, but a steep gradient at the start of the curve was not observed at all. On the other hand, the K_m, V_{max}, and k_{cat}/K_m of the recombinant PcpB constructed by Wang et al. [9] were 50 ± 7 μM, 30 ± 3 nmol/min/mg of protein, and 0.7 s^{-1} mM^{-1}, respectively. In this case, the concentration of PcpB was 5 nM and PCP ranged from 20 to 80 μM. These conditions are strictly steady state, but the majority of the enzyme in the reaction mixture seemed to be apoprotein without the required cofactor FAD. Moreover, dilution of recombinant PcpB, without a stabilizer such as BSA, may result in the inactivation of the enzyme. Indeed, it was shown in this study that the activity of recombinant PcpB was decreased in the absence of BSA. The values for K_m, V_{max}, and k_{cat}/K_m of the recombinant wild-type PcpB obtained in this study were 0.5 ± 0.05 μM, 8.9 ± 0.4 nmol/min/mg of protein, and 19 ± 1.1 s^{-1} mM^{-1}, respectively. These values are lower than those of native PcpB. However, the recombinant protein had a qualitatively similar activity to the wild-type protein in the presence of BSA

(see Results). Moreover, the flavin content and the molar ratio of flavin to protein were calculated to be 9.9 μM and 0.77, respectively (using 0.1120 as the value of absorbance of the recombinant PcpB at 446 nm and a molar extinction coefficient of 11300 for FAD [22]). This shows that the recombinant wild-type PcpB contains a comparable amount of FAD to that of the native PcpB because the molar ratio of flavin to protein of the native PcpB was calculated to be 0.64 [5]. Therefore, we believe that recombinant PcpB and associated mutants can be used to generate data representative of native protein.

In the previous paper, proteins were identified by their molecular weight as determined by SDS-PAGE. It is clear that wild-type and seven mutant forms of PcpB except Y397 are expressed correctly. The reduced amount of the desired protein in the soluble fraction, and contamination with impurities, might be responsible for the failure of identification in the case of Y397A [15]. The CD spectra illustrate that there are no major changes in the secondary structure between the wild-type PcpB and PcpB mutants with the exception R235E. When examined in conjunction with the absorption spectra (see below), it seems that the conformational changes in R235E prevented FAD from binding with a concomitant loss of activity. Hence, the charge of the amino acid at this position may greatly affect the protein structure. However, further analysis is needed to investigate the role of R235 to stabilize the chloride ion released from PCP during the enzymatic reaction.

The enzyme activities of PcpB and its mutants were correlated with the absorbance of the cofactor, FAD (Fig. 3). For example, no FAD-characteristic spectrum was seen in R235E (Fig. 3b) and Y397A (Fig. 8c (supporting information)), which have no activity toward PCP degradation [15]. Moreover, the decrease in the amount of FAD in the protein may also be responsible for the reduced activity of mutants such as F85A. The decreases in FAD content may be due to local conformational changes of the FAD binding site induced by the substitution of the amino-acid residues, which may not be detectable by CD spectra. Thus, more experimental data are needed to confirm that F85A is related to the recognition of PCP although the result of apparent kinetic parameters of F85A showed the importance of F85A for the catalytic reaction of PcpB.

The roles of Y216 and Y397 can be interpreted in the light of attained mass spectrometry data, CD and UV spectra. The increase in k_{cat} of Y397F and R235K compared to the wild-type PcpB was consistent with the previous report [15]. The increase in K_m, however, affects the decrease in k_{cat}/K_m and therefore it was found in this study that enzyme activity of these two mutants was not necessarily improved. Previously it was suggested that Y216 appears to be more important for the mechanism of action of PCP than Y397 [15]. This suggestion was confirmed in the present study (Table 2) due to the observation of ninefold and sixfold increases in the K_m of Y216F and Y397F, respectively. This also confirms the importance of both of these residues with respect to the recognition of the substrate.

In PHBH it has been shown that residues Y201 and Y385 form a hydrogen bond network with the 4-hydroxy group of *p*-hydroxybenzoate, the substrate of PHBH. Residues Y201 and Y385 of PHBH correspond to Y216 and Y397 in PcpB, respectively. The role of Y201 is to activate the substrate by stabilizing the phenolate and Y385 is essential for lowering the pK_a of the hydroxy group of Y201, thereby stimulating the ionization of the substrate upon binding. Y385 is also crucial for controlling the substrate specificity [23–26]. Thus, it could be very important for the elucidation of the roles of Y216 and Y397 to examine the pH dependence of the kinetic parameters and the reaction with the TCHQ or TCBQ, the reaction product of PcpB, in these mutants. In conclusion, analysis of recombinant histidine-tagged PcpB protein mutants is performed to examine the important amino acid residues in the catalytic reaction by identification of the proteins with mass spectrometry, circular dichroism (CD) and UV spectrometry, and determination of kinetic parameters. As a result, it was confirmed Y216 and Y397 play an important role with respect to the recognition of the substrate. 3D models are the result of continual refinement based upon experimental data. Limitations of homology modeling and molecular dynamics make it is impossible to predict the position of the side chains in the protein correctly. Thus, it will be essential to obtain data from X-ray crystallography in order to determine the positions of the side chains in PcpB, which will allow further mutagenesis of PcpB with a view towards improved catalytic properties.

Acknowledgements

We are indebted to the Research Resources Center at the RIKEN Brain Science Institute for DNA sequencing and mass spectrometric analysis. This work was supported by The Ministry of Education, Culture, Sports, Science, and Technology (MEXT) through a Grant-in-Aid for Young Scientists (B) (grant no. 14760067).

References

[1] K.A. McAllister, H. Lee, J.T. Trevors, Biodegradation 7 (1996) 1–40.
[2] I.S. Thakur, P.K. Verma, K.C. Upadhaya, Biochem. Biophys. Res. Commun. 286 (2001) 109–113.
[3] D.L. McCarthy, A.A. Claude, S.D. Copley, Appl. Environ. Microbiol. 63 (1997) 1883–1888.
[4] D.L. Saber, R.L. Crawford, Appl. Environ. Microbiol. 50 (1985) 1512–1518.
[5] L. Xun, C.S. Orser, J. Bacteriol. 173 (1991) 4447–4453.
[6] M. Takeuchi, K. Hamana, A. Hiraishi, Int. J. Syst. Evol. Microbiol. 51 (2001) 1405–1417.
[7] M.H. Dai, J.B. Rogers, J.R. Warner, S.D. Copley, J. Bacteriol. 185 (2003) 302–310.

[8] C.S. Orser, C.C. Lange, L. Xun, T.C. Zahrt, B.J. Schneider, J. Bacteriol. 175 (1993) 411–416.
[9] H. Wang, M.A. Tiirola, J.A. Puhakka, M.S. Kulomaa, Biochem. Biophys. Res. Commun. 289 (2001) 161–166.
[10] M. Bohac, Y. Nagata, Z. Prokop, M. Prokop, M. Monincova, M. Tsuda, J. Koca, J. Damborsky, Biochemistry 41 (2002) 14272–14280.
[11] L. Ridder, A.J. Mulholl, I.M.C.M. Rietjens, J. Vervoort, J. Am. Chem. Soc. 122 (2000) 8728–8738.
[12] L. Ridder, B.A. Palfey, J. Vervoort, I.M.C.M. Rietjens, FEBS Lett. 478 (2000) 197–201.
[13] S.D. Copley, Trends Biochem. Sci. 25 (2000) 261–265.
[14] L. Xun, E. Topp, C.S. Orser, J. Bacteriol. 174 (1992) 2898–2902.
[15] T. Nakamura, T. Motoyama, T. Hirokawa, S. Hirono, I. Yamaguchi, Chem. Pharm. Bull. 51 (2003) 1293–1298.
[16] C. Enroth, H. Neujahr, G. Schneider, Y. Lindqvist, Structure 6 (1998) 605–617.
[17] H.A. Schreuder, P.A.J. Prick, R.K. Wierenga, G. Vriend, K.S. Wilson, W.G.J. Hol, J. Drenth, J. Mol. Biol. 208 (1989) 679–696.
[18] J. Sambrook, E.F. Fritsch, T. Maniatis, Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989.
[19] J. Rosenfeld, J. Capdevielle, J.C. Guillemot, P. Ferrara, Anal. Biochem. 203 (1992) 173–179.
[20] A. Shevchenko, M. Wilm, O. Vorm, M. Mann, Anal. Chem. 68 (1996) 850–858.
[21] H. Shoun, T. Beppu, K. Arima, J. Biol. Chem. 254 (1979) 899–904.
[22] S. Strickland, V. Massey, J. Biol. Chem. 248 (1973) 2944–2952.
[23] B. Entsch, B.A. Palfey, D.P. Ballou, V. Massey, J. Biol. Chem. 266 (1991) 17341–17349.
[24] K. Eschrich, F.J.T. van der Bolt, A. de Kok, W.J.H. van Berkel, Eur. J. Biochem. 216 (1993) 137–146.
[25] M.S. Lah, B.A. Palfey, H.A. Schreuder, M.L. Ludwig, Biochemistry 33 (1994) 1555–1564.
[26] F.J.T. van der Bolt, R.H.H. van den Heuvel, J. Vervoort, W.J.H. van Berkel, Biochemistry 36 (1997) 14192–14201.
[27] W. Kabsch, C. Sander, Biopolymers 22 (1983) 2577–2637.

Studies of Binding Modes of (S)-Mephenytoin to Wild Types and Mutants of Cytochrome P450 2C19 and 2C9 Using Homology Modeling and Computational Docking

Akifumi Oda[1,2,3], Noriyuki Yamaotsu[2], and Shuichi Hirono[2]

Received June 9, 2004; accepted August 22, 2004

Purpose. This study investigated the structural features of CYP2C19 complexed with (S)-mephenytoin, using computational methods. In addition to wild-type CYP2C19 proteins (1A and 1B), which have selective 4 -hydroxylase activities of (S)-mephenytoin, CYP2C19 mutants were also studied, together with a wild type and artificial mutants of CYP2C19.

Methods. Three-dimensional structures of wild-type and mutant proteins of CYP2C19 and CYP2C9 were estimated from homology modeling using the crystal structure of rabbit CYP2C5 as a reference. The binding mode of (S)-mephenytoin to CYP2C19 was investigated using computational docking.

Results. The results reproduced the specific bindings between (S)-mephenytoin and the wild types of CYP2C19. Our findings suggest that Asp293 of CYP2C19 plays an important role in the binding of (S)-mephenytoin, which was surrounded by Val113 and Ala297, and points the phenyl ring at the heme iron. In addition the wild types of CYP2C19, the computational docking studies also accounted for the experimental activities of CYP2C19 mutants, and wild-type and mutant CYP2C19 proteins.

Conclusions. These results confirm that the predicted three-dimensional structure of the CYP2C19-(S)-mephenytoin complex is reasonable, and that this strategy is useful for investigating complex structures. Virtual screening for drug discovery can also be carried out using these methods.

KEY WORDS: binding mode; computational docking; cytochrome P450 2C19; homology modeling.

INTRODUCTION

Cytochrome P450 (CYP) [1]plays important roles in the metabolism of a wide variety of xenobiotic and endogenous compounds, including clinically important drugs. CYP2C19 (1) is one of the major enzymes that metabolizes drugs in the human liver and has an important role in the metabolism of (S)-mephenytoin (2,3). In 1984, Küpfer and co-workers observed that urinary recovery of 4′-hydroxylated (S)-mephenytoin was not detected for approximately 3–5% of Caucasians, which attracted attention to the identification of polymorphisms of CYP2C19 (4). A number of gene mutations in CYP2C19 were found by analyzing the genes of poor metabolizers (PMs) (5–11) (Table I). Almost all Oriental PMs can be described by two defective alleles (*CYP2C19*2A* and *CYP2C19*3*), but other mutant alleles are also important for Caucasian PMs. CYP2C9 is another enzyme that plays a significant role in drug metabolism (12). Although it metabolizes a variety of drugs, its 4′-hydroxylase activity of (S)-mephenytoin is negligible. In order to identify the key residues of CYP2C19 for 4′-hydroxylase activity, Tsao and co-workers constructed chimeras by replacing portions of CYP2C9 with those of CYP2C19, mutating individual residues by site-directed mutagenesis and assessing (S)-mephenytoin 4′-hydroxylase activity (13). They found that the mutation of six residues of CYP2C9 to imitate CYP2C19 resulted in 6% of the activity of wild-type CYP2C19.

To date, the three-dimensional structure of CYP2C19 has not yet been determined. The mechanism by which (S)-mephenytoin complexes with CYP2C19 is also unknown. However, the ligand-free crystal structure of one of the enzymes belonging to the CYP2C subfamily (rabbit CYP2C5) has been obtained (14). This provides a model upon which to base structurally unknown human CYPs and their ligands using computational methods.

The homology modeling method is the most powerful computational approach to predict the three-dimensional structures of proteins based on sequence similarities with structurally known proteins. This method estimates a three-dimensional model of a protein from the known structures of homologous proteins, and it is based on the assumption that a three-dimensional structure of one protein is analogous to the others, which have similar sequences. In addition to the structure prediction, computational ligand docking might also be used for structure-based drug design (15). This is one of the optimization problems, in which atoms of the ligand molecule are positioned into points in ligand-binding pockets of the target biomolecule. The difficulties of these problems exponentially increase in terms of the number of ligand atoms. Even if only the grid points of three-dimensional space are considered, these problems are NP-hard. For this problem, several approaches—for example, probabilistic search methods and sophisticated data structures—were proposed (16–18).

Previous structures that have been proposed for CYP2C19 complexed with (S)-mephenytoin using homology modeling and computational docking methods (19,20) were based on CYP102 (P450-BM3). The homology between CYP102 and CYP2C19 is lower than between CYP2C19 and CYP2C5. Although the homology modeling and docking with ligands for CYP2C19 were carried out using CYP2C5 as template (21–23), the docking of (S)-mephenytoin has not yet been examined. Furthermore, these studies only looked at wild-type CYP2C19 and not related mutants.

In this study, a three-dimensional structure of wild-type CYP2C19 complexed with (S)-mephenytoin was constructed using homology modeling and computational docking procedures. In addition to wild-type CYP2C19, homology modeling and docking studies were also attempted for mutant forms of the protein. Docking studies of a wild type and mutants of CYP2C9 were also used to investigate important structural properties for (S)-mephenytoin docking. Homology modeling calculations were generally carried out only for wild-type proteins, and this is the first study of the homology modeling and

[1] Discovery Laboratories, Toyama Chemical Co. Ltd., 2-4-1 Shimookui, Toyama 930-8508, Japan

[2] School of Pharmaceutical Sciences, Kitasato University, 5-9-1 Shirokane, Minato-ku, Tokyo 108-8641, Japan

[3] To whom correspondence should be addressed. (e-mail: AKIFUMI_ODA@toyama-chemical.co.jp)

0724-8741/04/1200-2270/0 © 2004 Springer Science+Business Media, Inc.

Table I. CYP2C19 Alleles

Allele	Enzyme activity	Effect of nucleotide changes	Name of protein
*2C19*1A*	Active	—	2C19.1A
*2C19*1B*	Active	Ile331Val	2C19.1B
*2C19*2A*	Inactive	Splicing defect	—
*2C19*2B*	Inactive	Splicing defect	—
*2C19*3*	Inactive	Stop codon	—
*2C19*4*	Inactive	Initial codon	—
*2C19*5A*	Inactive	**Arg433Trp**[b]	2C19.5A
*2C19*5B*	Inactive	Ile331Val, **Arg433Trp**[b]	2C19.5B
*2C19*6*	Inactive	**Arg132Gln**,[b] Ile331Val	2C19.6
*2C19*7*	Inactive	Exon skipping	—
*2C19*8*	Active/inactive[a]	**Trp120Arg**[b]	2C19.8

[a] 2C19.8 is active *in vitro* although it is inactive *in vivo*.
[b] Mutations without Ile331Val, which does not affect enzyme activities in 2C19.1B, are described by bold letter.

docking of mutant CYPs in order to elucidate the binding mode of the substrate. The availability of the modelings of mutants for the binding mode prediction is also discussed.

MATERIALS AND METHODS

Proteins

There are two ways to validate the availability of docking methods. One is the way in which several proteins are used, and another is the way in which several small molecules are used. Because one of the central purposes of this study is the proposal of the method for predicting the complex structures by using various mutant proteins, identical ligand molecule, (S)-mephenytoin, was docked into various target proteins. Furthermore, multi proteins were desired also for investigations of the availability of homology modeling for mutants.

Two wild-type and four mutant CYP2C19 proteins, along with one wild type and four artificial mutants of CYP2C9 constructed by Tsao and co-workers, were used for homology modeling and computational docking studies (Tables I and II). Wild-type enzymes that are expressed from alleles *CYP2C19*1A* and *CYP2C19*1B* are described in Roman script (that is, CYP2C19.1A and CYP2C19.1B). The proteins expressed from the mutant alleles *CYP2C19*5A*, *CYP2C19*5B*, *CYP2C19*6* and *CYP2C19*8* are described as CYP2C19.5A, CYP2C19.5B, CYP2C19.6, and CYP2C19.8, respectively. Four artificial mutants of CYP2C9 were used: protein **1**, which had four mutant residues (2C9/I99H, S220P, P221T and S286N); protein **2**, which had five mutant residues (2C9/I99H, S220P, P221T, S292A and F295L); protein **3**, which had five mutant residues (2C9/I99H, S220P, P221T, S286N and F295L); and protein **4**, which had six mutant residues (2C9/I99H, S220P, P221T, S286N, V292A and F295L). All mutants of CYP2C9 are not naturally occurring. In Tables I and II, amino-acid mutations are also described (note that although CYP2C19.1B has an Ile331 to Val mutation, there is no loss of (S)-mephenytoin 4′-hydroxylase activity. In Table I, bold type letters illustrate mutations without Ile331Val). (S)-Mephenytoin 4′-hydroxylase activities were experimentally observed only for CYP2C19.1A, CYP2C19.1B, CYP2C19.8, and protein **4**. Although the enzyme activity of protein **4** is 6% of CYP2C19, it is much larger than the other mutants (e.g. enzyme activity of CYP2C19.5A is about 0.37% of that of CYP2C19.1A). In this study, because only binding modes are discussed, we treat the protein **4** as active enzyme.

Homology Modeling

In this study, we used the crystal structure of rabbit CYP2C5 (PDB code: 1DT6) (14) as the template for homology modeling of CYP2C19 and CYP2C9. The basic local alignment search tool (BLAST) based on finite automaton was used for the sequence alignments of the target (CYP2C19 and CYP2C9) against template (CYP2C5). BLOSUM62 was used as alternate scoring matrix for BLAST. The alignment produced by BLAST was used as the input for the FAMS program (24). In FAMS, the modeling of C_α atoms, main-chain atoms and side-chain atoms are carried out sequentially. For modeling calculations, FAMS can fully automatically perform both the database search and optimization steps by using the simulated annealing (SA) method. Even if the positions of side-chain atoms extracted from databases are not adequate—for example, if they have a conflict of main chain atoms—FAMS can refine the structure by using SA. Although homology models for wild types and mutants of CYP2C19 and CYP2C9 were obtained by FAMS, there were no hydrogen atoms in these model structures because hydrogens were not observed in the template of modeling (the experimental structure of CYP2C5 produced by X-ray crystallography). In this study, constructed homology models were protonated by the "protonate" module in AMBER 6.0 (25), and their structures were refined by AMBER PARM94 force fields. Only 1000 cycle calculations were carried out for refinements, because these calculations were performed only for reducing structural distortions. The explicit water molecules and cyclic-boundary conditions were not adopted for these refinements; a cutoff distance of 12 Å for non-bonded interactions and a distance-dependent dielectric coefficient $\varepsilon = 80R$ (R: distance between two interacting atoms) were used. For these calculations, the steepest-descent minimizer was used in the first 10 steps, and the rest of the calculations were carried out using the conjugate-gradient method.

Table II. CYP2C9 Mutants

Construct	Enzyme activity	Abbreviated name of protein
CYP2C9	Inactive	CYP2C9
CYP2C9/I99H, S220P, P221T, S286N	Inactive	**1**
CYP2C9/I99H, S220P, P221T, V292A, F295L	Inactive	**2**
CYP2C9/I99H, S220P, P221T, S286N, F295L	Inactive	**3**
CYP2C9/I99H, S220P, P221T, S286N, V292A, F295L	Active	**4**

Generation of a Conformer Set for the Ligand Molecule

The structure of the ligand molecule (S)-mephenytoin optimized by *ab initio* HF/6-31G* method was used for computational docking studies. Atomic charges of the ligand molecule were obtained by ESP method, which can reproduce the electrostatic potential calculated by *ab initio* HF/6-31G* (26). Gaussian 98 was used for quantum chemical calculations (27). In the complex structure, because the conformation of a ligand molecule is not always the most stable conformer, we considered many metastable states of ligands for computational docking studies. For this purpose, a calculation of only the most stable conformation is not sufficient, and a conformer set consisting of many metastable conformers of a ligand is required. In this study, CAMDAS, which is a program developed by this laboratory (28) for the generation of ligand conformer sets, was used to calculate a conformer set for (S)-mephenytoin. Molecular dynamics (MD) simulations at high temperatures were carried out by this program, and the conformers obtained in the trajectory of this simulation were sampled. The clustering procedure was carried out for all sampled conformers, and the representative conformers were picked up from the clusters. In the CAMDAS calculation, a modified MMFF force field was used (29), in which electrostatic interactions were omitted and the weights of both angle bending and dihedral terms equal to 0.8 were included. The temperature of MD simulation by CAMDAS was 1200 K and 1 ns (1,000,000 steps) calculation was carried out. Chiral inversions of (S)-mephenytoin were not allowed through the MD simulation. The two dihedral angles shown in Fig. 1 were used as explanatory variables for the clustering. For (S)-mephenytoin, the conformer set in which 31 conformers were included was obtained by CAMDAS calculation.

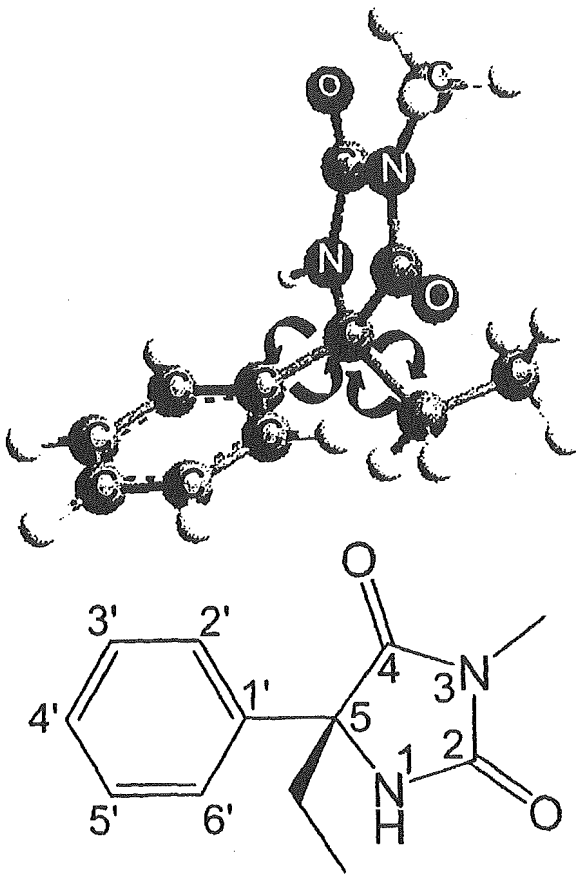


Fig. 1. Two torsional angles for clustering of (S)-mephenytoin conformers. The two torsional angles illustrated in this figure were used as criteria for the clustering of conformers.

Computational Docking

For computational ligand docking, ligand-binding pockets in the homology models of the proteins were explored using the SiteID module in the SYBYL program package (Tripos Inc., St. Louis, MO, USA). SiteID can search for potential binding sites within or on the proteins by using the floodfill-solvation technique. The grid method with default parameters was used in this study. MOLCAD surfaces for the amino-acid residues within a 6 Å radius of explored binding pockets were calculated using the MOLCAD program (Tripos Inc.). These surfaces were used as steric constraints for computational docking trials. In the obtained ligand-binding pocket of the proteins, hydrogen-bond acceptor and donor site queries were defined by the UNITY program (Tripos Inc.). In this procedure, some queries were deleted because of VDW bumps and we used only the remaining queries for computational docking in the next step. We adopted tolerances of 1.5 Å for the spatial constraints of all hydrogen-bond queries. Using the steric constraint and hydrogen-bond queries, each conformer from the conformer set of (S)-mephenytoin generated by CAMDAS was docked into the homology models of the CYPs by the UNITY 3D program. For wild-type CYP2C19.1A, docking models that fulfilled at least one of the hydrogen-bond queries ("partial match") were adopted. For the other proteins, only one query that plays an important role in docking models of wild-type CYP2C19.1A was considered. The steric constraint was constructed from the MOLCAD surface with default parameters. Using these strategies, many model structures of the complex were obtained for one ligand-protein system; for example, 31 models were generated by the docking calculation of (S)-mephenytoin with CYP2C19.1A. In computational docking studies, we need to determine which models can be adopted as candidates for the complex structure. Usually, the scoring functions that calculate the binding free energies of ligand-protein systems were used for this purpose, and the models that had smaller scores (that is, smaller binding free energy) were accepted as more reasonable models. Although many types of scoring functions were introduced (16,18,30–32), each type has different strengths and weaknesses. Recently, a method for evaluating docking models called "consensus scoring" has been introduced (33,34). In consensus scoring, multiple scoring functions were used to cover the shortcomings of each method. In this study, the validities of the calculated docking models were evaluated using CScore (Tripos Inc.) (34), which is one of the consensus-scoring programs. It includes five scoring functions: FlexX score (16), Gold score (18), PMF score (30), Dock score (31), and ChemScore (32). These five scores were independently calculated for docking models, and awarded marks out of five for the validities of the models. The models that are given high scores by CScore are

the possible candidates of the structures of protein–ligand complexes. However, when docking models with high CScore values were not obtained, we assumed that the computational docking trials for the ligand–protein systems could not be successfully carried out.

MD Simulation

The homology model the docking results of which were not reasonable were further analyzed by structural refinement by the MD simulation with explicit water molecules. The rectangular solid box in which the complex was included was set up and the box was filled with TIP3P water molecules (35). Sodium ions were used as counter ions for setting the total charge of the system to 0. The 1 ns MD simulation at 300 K was carried out using the particle-mesh Ewald method for calculations of electrostatic interactions under cyclic-boundary conditions by the AMBER 6.0 program. The bond lengths between two atoms were constrained by SHAKE (36). The cutoff distance for van der Waals interactions was 9.0 Å, and the MD simulation was carried out under isothermal and isochoric conditions. After the MD calculation, 100,000 steps of structure refinement were performed using steepest-descent and conjugate-gradient optimizers. The first 5000 steps of refinement were calculated by the steepest-descent method, and the rest of the calculation was carried out using the conjugate-gradient method, until the energy gradient become less than 0.1 kcal/molÅ. Explicit TIP3P waters were used in this refinement with the particle-mesh Ewald method under the cyclic-boundary condition. A cutoff distance of 9.0 Å for van der Waals interactions and no constraint of bond lengths were used for the refinement. After this calculation, an additional computational docking study was carried out for refined structure. The same computational conditions as previous docking calculations were used in the additional docking.

For FAMS calculations we used a personal computer with one Pentium3 600 MHz processor and Red Hat Linux 6.1. For the other calculations, an Octane2 workstation by SGI with dual MIPS R12000 400 MHz processors and IRIX release 6.5 was used.

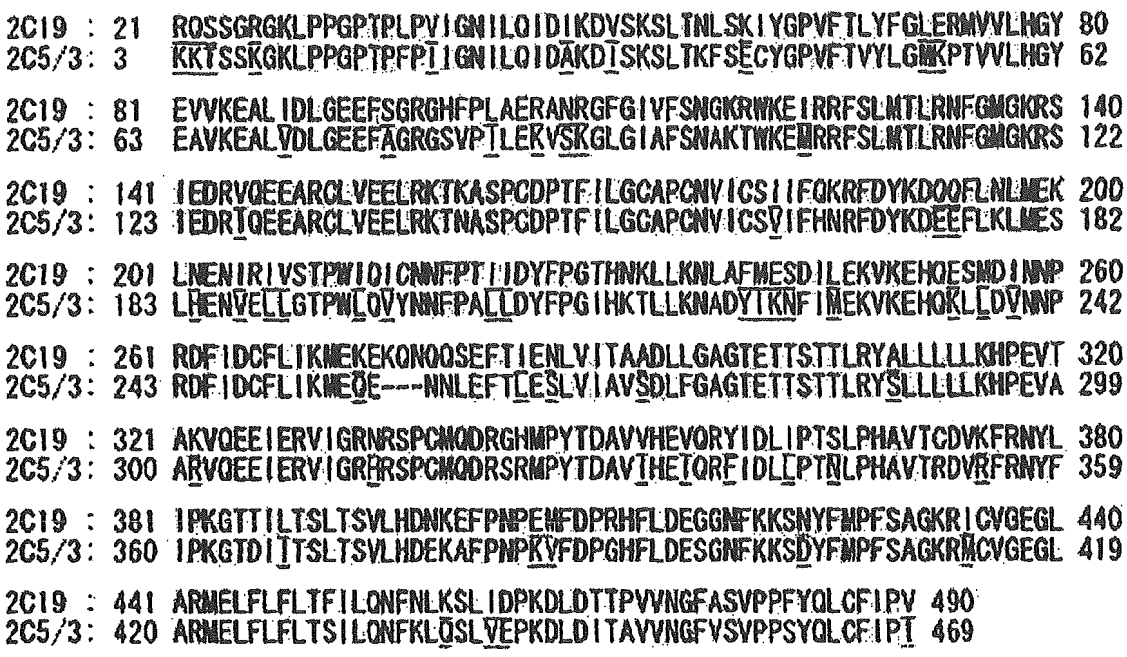
RESULTS

Sequence alignments of CYP2C19.1A and CYP2C9 against CYP2C5/3LVdH using BLAST are illustrated in Fig. 2. Similarities in amino-acid sequences are shown in Table III. In aligned sequences, identical amino acids were defined as "identity", and the pairs of residues whose score were positive in BLOSUM62 were defined as "positive". Although there are three insertions between helix H and I in both CYP2C19.1A and CYP2C9, they have a high degree of homology with CYP2C5/3LVdH. In this study, these alignments were used for homology modeling calculations of not only wild types but also mutants of CYP2C19 and CYP2C9. In Fig. 3, structurally optimized homology models for wild types and mutants of CYP2C19 are illustrated. The wild-type CYP2C9 and artificial mutant protein **4** are also shown in Fig. 4. In these figures, mutated residues in proteins without wild-type CYP2C19.1A and CYP2C9 are also shown.

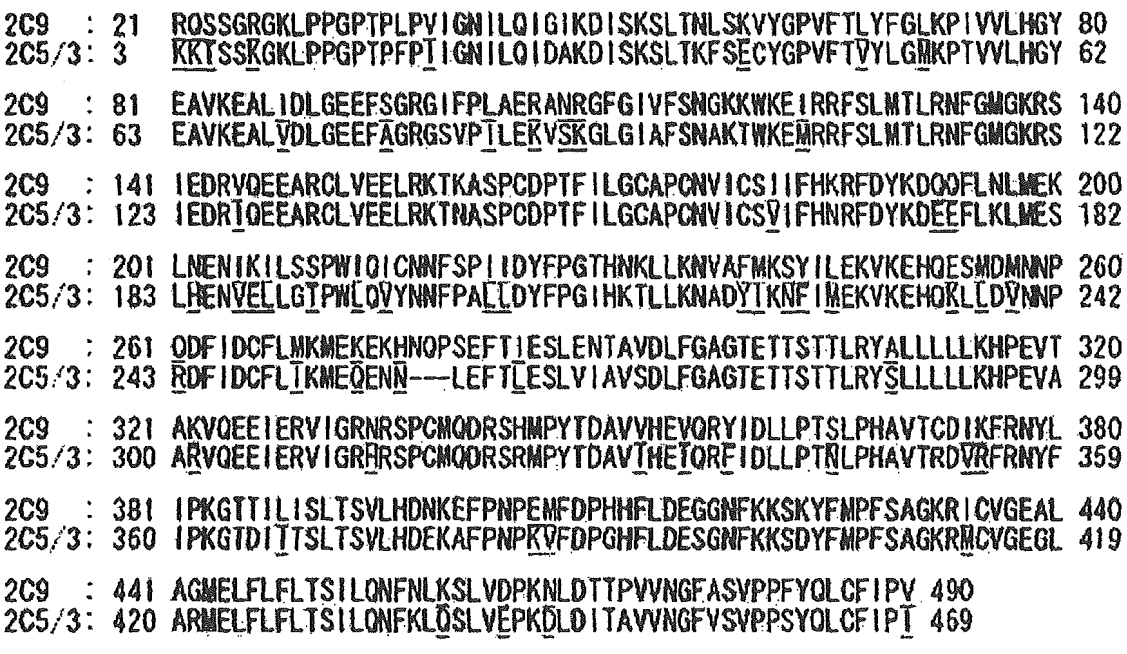
The ligand-binding pocket found by SiteID for the CYP2C19.1A homology model is illustrated in Fig. 5. Although several candidates of binding pockets were obtained by SiteID, the adjacent pocket of heme is shown in this figure because we know that active sites of CYPs are near the heme region. As shown in this figure, the ligand-binding pocket in the homology model was mainly composed of hydrophobic amino acids such as Val, Ile, Phe, Leu and Ala. The hydrophobic feature of the pocket of CYP2C19.1A was consistent with the known crystallographic structures of other CYPs and it is intimately related with the property of CYPs, which bind with hydrophobic substrates and oxidatively metabolize them. The hydrogen-bond queries in this region, which were three hydrogen-bond acceptor sites, were also illustrated. They were associated with the main-chain and side-chain oxygen atoms of Asp293 and the main-chain carbonyl oxygen of Gly296, respectively. Although all of them were directed to the heme moiety of CYP2C19.1A, the query from the carbonyl oxygen atom in the main chain of Asp293 was closest to the heme iron. Some of these queries might play important roles in the docking of (S)-mephenytoin in CYP2C19.

Using the hydrogen-bond queries and steric constraints calculated from the MOLCAD surface, a computational docking trial of (S)-mephenytoin was carried out for wild-type CYP2C19.1A. The conformer set of (S)-mephenytoin included 31 conformers, which were used for docking studies. Figure 6 shows one of the docking models with CScore values equal to 5 (that is, full points). The hydrogen bond between the amide hydrogen adjacent to the nitrogen atom of (S)-mephenytoin and the carbonyl oxygen in the main chain of Asp293 in CYP2C19.1A played a significant role in forming this docking model. The 4'-hydrogen of (S)-mephenytoin in this model was positioned near the heme iron. (S)-Mephenytoin is surrounded by hydrophobic amino acids such as Val113, Ile205, Ala292, Ala297, Leu366, and Phe476.

Computational docking trials for proteins other than CYP2C19.1A with (S)-mephenytoin were also carried out by using SiteID and UNITY (Table IV). These protein models contained ligand-binding pockets that were mainly composed of hydrophobic amino acids, which were similar to CYP2C19.1A. Although Asp 293 was conserved in all proteins, hydrogen-bond queries constructed from the carbonyl oxygen of Asp293 in CYP2C19.6, CYP2C9, proteins **1** and **3** were omitted by the VDW bump check. The bumps were caused by the differences of the shapes of ligand-binding sites between CYP2C19.1A and these proteins, and the shapes of the sites were affected by the mutations (CYP2C19.6) or differences in amino-acid sequences between CYP2C19 and CYP2C9 (CYP2C9, proteins **1** and **3**). These proteins were all experimentally observed to be inactive for (S)-mephenytoin 4'-hydroxylation. Computational docking calculations for another six proteins, including the hydrogen-bond queries of Asp293, were carried out using the same procedure of CYP2C19.1A. The obtained docking models were evaluated by CScore. When no models had high CScore values for a certain complex system, we considered that the docking trial had failed, even if docking models were obtained. For four of these six proteins (CYP2C19.1B, CYP2C19.8, proteins **2** and **4**) computational docking models were obtained in which 4'-hydrogens of (S)-mephenytoin were positioned nearby heme irons and with CScore values of 5 (that is, full points). For three of these four proteins (CYP2C19.1B, CYP2C19.8 and protein **4**) the enzyme activities were experimentally observed and computational docking studies seemed able to account for the experimental results. In these complex models,



(a) CYP2C19



(b) CYP2C9

Fig. 2. Sequence alignments of CYP2C19 and CYP2C9 against CYP2C5/3LVdH. The identical residues were expressed by meshed letters and the underlined residues were positives.

the ligand (S)-mephenytoin was surrounded by hydrophobic amino acids such as Val113, Ile205, Ala297 and Phe476, which were conserved in all wild types and mutants in this study. Furthermore, proteins for which the solutions of CScore = 5 were not obtained by computational docking trials based on the hydrogen-bond query from carbonyl oxygen of Asp293 (that is, CYP2C19.5A and CYP2C19.5B) were experimentally inactive mutants.

Table III. Identities and Positives[a] Between Amino-Acid Sequences of CYP2Cs

	CYP2C5/3LVdH	CYP2C19	CYP2C9
CYP2C5/3LVdH	100%	75%	75%
CYP2C19	87%	100%	92%
CYP2C9	87%	96%	100%

[a] Identities and positives were described in upper right and lower left portions, respectively.

By contrast, for the experimentally inactive protein 2, docking models with CScore values of 5 were also obtained. Thus, we carried out more structure refinement for ligand-free protein 2 using MD simulation, steepest-descent and conjugate-gradient optimizers. After the refinement, we also performed additional computational docking for the refined structure with (S)-mephenytoin. Figure 7 shows the three-dimensional structures of protein 2 before and after the refinement. This reveals that structural features, such as the position of heme and the shape of the ligand-binding pocket around Asp293, were changed by MD calculations. After the structure refinement, no solutions of the docking trials of (S)-mephenytoin with protein 2 were obtained under the assumption that the hydrogen bond between carbonyl oxygen of Asp293 and (S)-mephenytoin are indispensable to ligand docking.

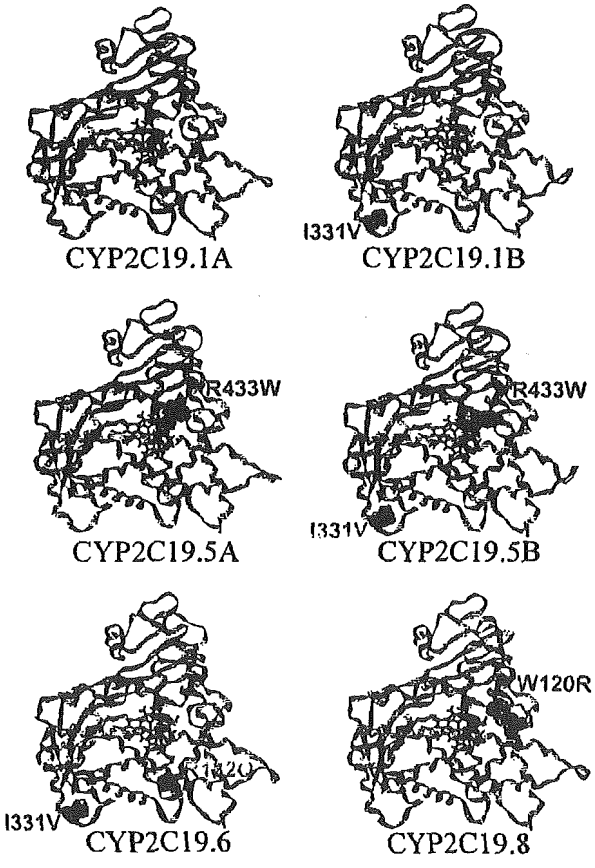


Fig. 3. Homology model for wild types and mutants of CYP2C19.

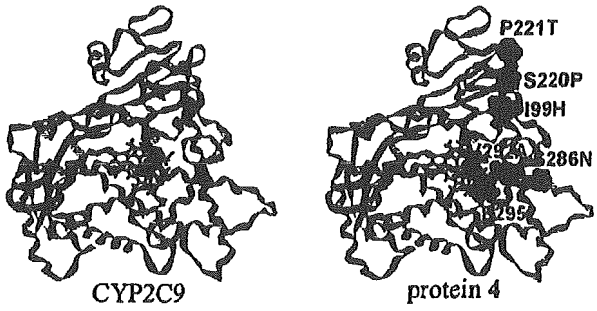


Fig. 4. Homology model for wild types and mutants of CYP2C9.

DISCUSSION

Homology models illustrated in Figs. 3 and 4 suggest that even if the mutated portions are not near the active sites (i.e., in adjacent regions of heme) mutations of one or two residues are still capable of inactivating CYP2C19. For example, Ile331Val, which does not affect the enzyme activity in CYP2C19.1B, and also mutations that play important roles for inactivities of mutants of CYP2C19 (Arg132Gln in CYP2C19.6), were positioned far from the ligand-binding pocket. Furthermore, although Arg433Trp in both CYP2C19.5A and CYP2C19.5B is located near by heme moiety, this residue is far away from the iron atom and does not seem to be directly implicated in the ligand docking. This indicates that mutations change the conformations of ligand-binding pockets of proteins, and affect the enzyme activities even if the mutations are located far from the ligand-binding pocket. By contrast, Ile99His, Ser286Asn, Val292Ala and Phe295Leu in protein 4 existed near the active site in our homology model. In particular, Ser286Asn, Val292Ala and Phe295Leu belong to the helix I, and these positions are also members of the substrate-recognition site 4 (SRS-4) (37). Therefore, these mutations might be directly implicated in (S)-mephenytoin docking for protein 4.

In the model of CYP2C19.1A complex with (S)-

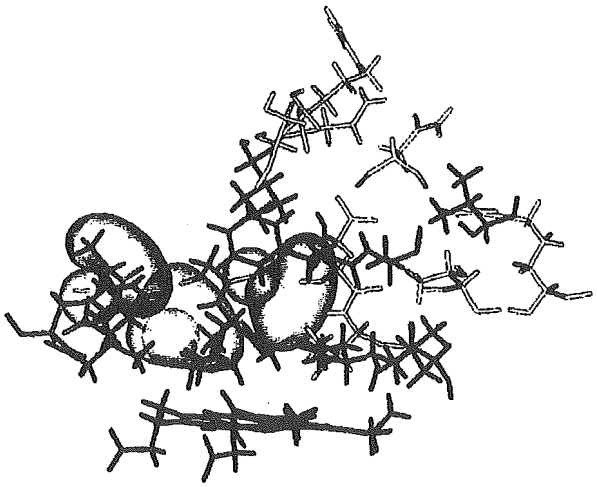


Fig. 5. Active-site residues explored by Site ID and hydrogen-bond queries in CYP2C19.1A. Hydrophobic amino acids are colored green; Asp293 and Gly296 are illustrated in red. Three hydrogen-bond queries were described with spatial constraint.

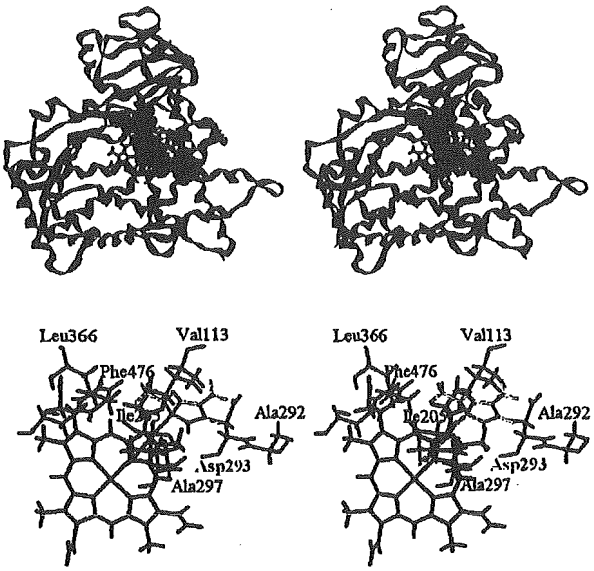


Fig. 6. CYP2C19.1A complex with (S)-mephenytoin. (S)-Mephenytoin, Asp293, and hydrophobic amino acids were illustrated in yellow, red, and green, respectively.

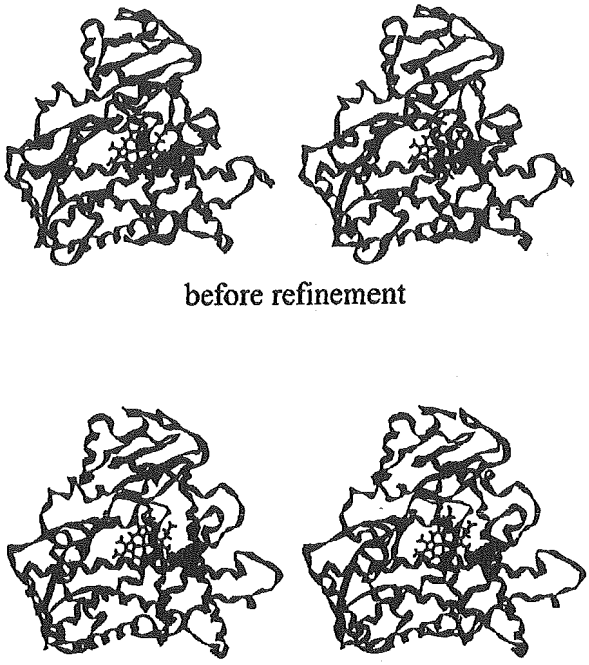


Fig. 7. Three-dimensional structures of protein 2 before and after refinement by MD simulation, steepest-decent and conjugate-gradient optimizers. The two structures were aligned by the fitting of their main chains. The RMSD of the protein 2 structure before and after refinement was 3.524 Å. The red colored residue was Asp293.

mephenytoin shown in Fig. 6, the orientation of (S)-mephenytoin whose 4′-hydrogen was located near the heme iron is consistent with (S)-mephenytoin 4′-hydroxylase activity of CYP2C19.1A and suggests that the three-dimensional structure of the calculated docking model is reasonable. The result that (S)-mephenytoin was surrounded by hydrophobic amino acids indicates that this hydrophobicity plays an important role for (S)-mephenytoin docking in the binding pocket of CYP2C19.1A. Significantly, Ala297 and Phe476 were located nearby the phenyl ring of ligand, suggesting that the CH-π interaction between the phenyl ring and Ala297, and/or the π-π stacking between ligand and Phe476, are important for the location of (S)-mephenytoin in CYP2C19.1A. Furthermore, Asp293 and Ala297, which seem to be important residues for ligand docking, are members of helix I and SRS-4. This is consistent with the results of Tsao *et al.* (11)

Table IV. The Results of Computational Docking

Protein	Experimental 4′-hydroxylase activity	Hydrogen bond query on Asp293[a]	Maximum CScore value[b]
2C19.1A	Active	OK	5
2C19.1B	Active	OK	5
2C19.5A	Inactive	OK	3
2C19.5B	Inactive	OK	4
2C19.6	Inactive	Bump	—
2C19.8	Active (*in vitro*)	OK	5
2C9	Inactive	Bump	—
1	Inactive	Bump	—
2	Inactive	OK	5
3	Inactive	Bump	—
4	Active	OK	5

[a] The result of VDW bump checks for hydrogen-bond query on Asp293.
[b] The best CScore value of those for all docking models.

For other proteins than CYP2C19.1A, hydrogen-bond queries constructed from the carbonyl oxygen of Asp293 in CYP2C19.6, CYP2C9, proteins 1 and 3 were omitted by the VDW bump check, and there were no docking models whose CScore values were 5 in CYP2C19.5A and CYP2C19.5B. For the refined model of protein 2, no solutions were obtained by the computational docking. Because all these proteins were experimentally observed enzyme inactivities without exceptions, these results support that Asp293 plays an important role in CYP2C19–(S)-mephenytoin complex. Furthermore, for CYP2C19.1B, CYP2C19.8 and proteins 4 computational docking calculations gave the solutions of CScore = 5, and these proteins were experimentally active. These results suggest that our methods applied to this study are reasonable for binding mode predictions.

Only recently, the three-dimensional structures of CYP2C5 complex with sulfaphenazole (38), ligand-free CYP2C9 and CYP2C9 complexed with warfarin (39), were determined by X-ray crystallographic analyses. Because the sizes of the ligands in these complexes were different from (S)-mephenytoin, these structures were not able to give critical clues for the investigation of the complex structure of CYP2C19 with (S)-mephenytoin. Furthermore, for the observed structure of the CYP2C9 complex with warfarin, the ligand molecule was not located in the usual position of CYP substrate-binding pockets. Although this is useful for investigating the mechanisms of inhibitions and allosteric effects of CYPs, it does not aid research into the binding modes of normal CYP substrates, such as (S)-mephenytoin. Nonethe-

less, when these data are used together with the results of computational docking studies in this paper, it might give some support to the predictions of complex structures. Comparing the structures of CYP2C19.1A with (S)-mephenytoin in this study, and CYP2C5 with sulfaphenazole in crystallographic data, it was found that Asp290 of CYP2C5 (the residue corresponding to Asp293 in CYP2C19.1A) was in the substrate-binding pocket and interacted with the ligand sulfaphenazole. This indicates that both Asp290 of CYP2C5 and Asp293 of CYP2C19.1A play significant roles in binding of substrates, and supports the validity of our model structure of CYP2C19 complex with (S)-mephenytoin.

CONCLUSIONS

In this study, we found that models produced by homology modeling and computational docking procedures of CYP complexed with (S)-mephenytoin were obtained for experimentally active enzymes but not for inactive proteins. The binding modes of all active enzymes were similar to wild-type CYP2C19.1A. For inactive proteins, hydrogen bonds between the carbonyl oxygen in the main chain of Asp293 and the ligand were not able to form, or the scores of the resulting docking models were low. These results indicate that the computational methods used in this study are useful for the investigation of the structures of ligand–protein complexes. Although more detailed confirmations of the importance of Asp293 for (S)-mephenytoin binding into CYP2C19 by using structural biologic methods such as X-ray diffraction or NMR spectroscopy were preferable, the role of Asp293 was supported by the results that all calculations for wild types and mutants are consistent with experimentally observed enzyme activities without exception.

Three-dimensional structures of biopolymers are very important for the rational drug design. Though the determinations of structures by experimental methods (i.e., X-ray diffraction or NMR) give the useful information for drug design, these experiments are expensive and time-consuming. Therefore, it is preferable that the three-dimensional structures of structurally unknown proteins were computationally predicted using known structures of homologous proteins. Especially, when many complexes have to be considered for drug design (i.e., the situation that there are many mutants such as CYP2C19), homology modeling play an important role in drug design studies because all structures can not be determined by experiments within reasonable costs. In this study, even if the mutated parts of the mutants were far from the substrate-binding pockets, such as CYP2C19.6, the computational results were consistent with the experimental activities. This suggests that the computational methods used in this study, especially for FAMS, are promising not only for wild-type enzymes but also for mutants. We expect that virtual screening trials can be carried out for several proteins using these methods.

REFERENCES

1. M. Romkes, M. B. Faletto, J. A. Blaisdell, J. L. Raucy, and J. A. Goldstein. Cloning and expression of complementary DNAs for multiple members of the human cytochrome P450IIC subfamily. *Biochemistry* **30**:3247–3255 (1991).
2. S. A. Wrighton, J. C. Stevens, G. W. Becker, and M. VandenBranden. Isolation and characterization of human liver cytochrome P450 2C19: correlation between 2C19 and S-mephenytoin 4′-hydroxylation. *Arch. Biochem. Biophys.* **306**: 240–245 (1993).
3. J. A. Goldstein, M. B. Faletto, M. Romkes-Sparks, T. Sullivan, S. Kitareewan, J. L. Raucy, J. M. Lasker, and B. I. Ghanayem. Evidence that CYP2C19 is the major (S)-mephenytoin 4′-hydroxylase in humans. *Biochemistry* **33**:1743–1752 (1994).
4. A. Küpfer, P. Desmond, R. Patwardhan, S. Schenker, and R. A. Branch. Mephenytoin hydroxylation deficiency: kinetics after repeated doses. *Clin. Pharmacol. Ther.* **35**:33–39 (1984).
5. S. M. F. de Morais, G. R. Wilkinson, J. Blaisdell, K. Nakamura, U. A. Meyer, and J. A. Goldstein. The major genetic defect responsible for the polymorphism of S-mephenytoin in humans. *J. Biol. Chem.* **269**:15419–15422 (1994).
6. S. M. F. de Morais, G. R. Wilkinson, J. Blaisdell, K. Nakamura, U. A. Meyer, and J. A. Goldstein. Identification of a new genetic defect responsible for the polymorphism of S-mephenytoin metabolism in Japanese. *Mol. Pharmacol.* **46**:594–598 (1994).
7. Z. S. Xiao, J. A. Goldstein, H.-G. Xie, J. Blaisdell, W. Wang, C.-H. Jiang, F.-X. Yan, N. He, S.-L. Huang, Z.-H. Xu, and H.-H. Zhou. Differences in the incidence of the CYP2C19 polymorphism affecting the S-mephenytoin phenotype in Chinese Han and Bai populations and identification of a new rare *CYP2C19* mutant allele. *J. Pharmacol. Exp. Ther.* **281**:604–609 (1997).
8. R. J. Ferguson, S. M. F. de Morais, S. Benhamou, C. Bouchardy, J. Blaisdell, G. Ibeanu, G. R. Wilkinson, T. C. Sarich, J. M. Wright, P. Dayer, and J. A. Goldstein. A novel defect in human *CYP2C19*: mutation of the initiation codon is responsible for poor metabolism of S-mephenytoin. *J. Pharmacol. Exp. Ther.* **284**:356–361 (1998).
9. G. C. Ibeanu, J. A. Goldstein, U. Meyer, S. Benhamou, C. Bouchardy, P. Dayer, and B. I. Ghanayem. and J. Blaisdell. Identification of new human *CYP2C19* alleles (*CYP2C19*6* and *CYP2C19*2B*) in a Caucasian poor metabolizer of mephenytoin. *J. Pharmacol. Exp. Ther.* **286**:1490–1495 (1998).
10. G. C. Ibeanu, J. Blaisdell, B. I. Ghanayem, C. Beyeler, S. Benhamou, C. Bouchardy, G. R. Wilkinson, P. Dayer, A. K. Daly, and J. A. Goldstein. An additional defective allele, *CYP2C19*5*, contributes to the S-mephenytoin poor metabolizer phenotype in Caucasians. *Pharmacogenetics* **8**:129–135 (1998).
11. G. C. Ibeanu, J. Blaisdell, R. J. Gerguson, B. I. Ghanayem, K. Brosen, S. Benhamou, C. Bouchardy, G. R. Wilkinson, P. Dayer, and J. A. Goldstein. A novel transversion in the intron 5 donor splice junction of *CYP2C19* and a sequence polymorphism in exon 3 contribute to the poor metabolizer phenotype for the anticonvulsant drug S-mephenytoin. *J. Pharmacol. Exp. Ther.* **290**:635–640 (1999).
12. S. Kimura, J. Pastewka, H. V. Gelboin, and F. J. Gonzalez. cDNA and amino acid sequences of two members of the human P450IIC gene subfamily. *Nucleic Acids Res.* **15**:10053–10054 (1987).
13. C.-C. Tsao, M. R. Wester, B. Ghanayem, S. J. Coulter, B. Chanas, E. F. Johnson, and J. A. Goldstein. Identification of human CYP2C19 residues that confer S-mephenytoin 4′-hydroxylation activity to CYP2C9. *Biochemistry* **40**:1937–1944 (2001).
14. P. A. Williams, J. Cosme, V. Sridhar, E. F. Johnson, and D. E. McRee. Mammalian microsomal cytochrome P450 monooxygenase: structural adaptations for membrane binding and functional diversity. *Mol. Cell* **5**:121–131 (2000).
15. R. D. Taylor, P. J. Jewsbury, and J. W. Essex. A review of protein-small molecule docking methods. *J. Comput. Aided Mol. Des.* **16**:151–166 (2002).
16. M. Rarey, B. Kramer, T. Lengauer, and G. Klebe. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* **261**:470–489 (1996).
17. T. J. Ewing. DOCK 4.0: search strategies for automated molecular docking of flexible molecule databases. *J. Comput. Aided Mol. Des.* **15**:411–428 (2001).
18. G. Jones, P. Willet, R. C. Glen, A. R. Leach, and R. Taylor. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **267**:727–748 (1997).
19. D. F. V. Lewis, M. Dickins, R. J. Weaver, P. J. Eddershaw, P. S. Goldfarb, and M. H. Tarbit. Molecular modeling of human CYP2C subfamily enzymes CYP2C9 and CYP2C19: rationalization of substrate specificity and site-directed mutagenesis experiments in the CYP2C subfamily. *Xenobiotica* **28**:235–268 (1998).
20. V. A. Payne, Y.-T. Chang, and G. H. Loew. Homology modeling

and substrate binding study of human CYP2C18 and CYP2C19 enzymes. *Proteins* 37:204–217 (1999).
21. M. Ridderström, I. Zamora, O. Fjellström, and T. B. Andersson. Analysis of selective regions in the active sites of human cytochromes P450, 2C8, 2C9, 2C18 and 2C19 homology models using GRID/CPCA. *J. Med. Chem.* 44:4072–4081 (2001).
22. D. F. V. Lewis. Modeling human cytochromes P450 involved in drug metabolism from the CYP2C5 crystallographic template. *J. Inorg. Biochem.* 91:502–514 (2002).
23. M. J. de Groot, A. A. Alex, and B. C. Jones. Development of a combined protein and pharmacophore model for cytochrome P450 2C9. *J. Med. Chem.* 45:1983–1993 (2002).
24. K. Ogata and H. Umeyama. An automatic homology modeling method consisting of database searches and simulated annealing. *J. Mol. Graphics Mod.* 18:258–272 (2000).
25. D. A. Case, D. A. Pearlman, J. W. Caldwell, T. E. Cheatham, W. S. Ross, C. Simmerling, Y. Duan, J. Pitera, I. Massova, G. L. Seibel, U. C. Singh, P. Weiner, and P. A. Kollman. *AMBER 6*, University of California, San Francisco, CA 1999.
26. U. C. Singh and P. A. Kollman. An approach to computing electrostatic charges for molecules. *J. Comput. Chem.* 5:129–145 (1984).
27. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, A. G. Baboul, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. G. Johnson, W. Chen, M. W. Wong, J. L. Andres, M. Head-Gordon, E. S. Replogle, and J. A. Pople. *Gaussian 98 Revision A.7*, Gaussian, Inc, Wallingford, CT 1998.
28. H. Tsujishita and S. Hirono. CAMDAS: an automated conformational analysis system using molecular dynamics. *J. Comput. Aided Mol. Des.* 11:305–315 (1997).
29. T. A. Halgren. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *J. Comput. Chem.* 17:490–519 (1996).
30. I. Muegge and Y. C. Martin. A general and fast scoring function for protein–ligand interactions: a simplified potential approach. *J. Med. Chem.* 42:791–804 (1999).
31. I. D. Kuntz, J. M. Blaney, S. J. Oatley, R. Langridge, and T. E. Ferrin. A geometric approach to macromolecule–ligand interactions. *J. Mol. Biol.* 161:269–288 (1982).
32. M. D. Eldridge, C. W. Murray, T. R. Auton, G. V. Paolini, and R. P. Mee. Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *J. Comput. Aided Mol. Des.* 11: 425–445 (1997).
33. P. S. Charifson, J. J. Corkery, M. A. Murcko, and W. P. Walters. Consensus scoring: a method for obtaining improved hit rates from docking databases of three-dimensional structures into proteins. *J. Med. Chem.* 42:5100–5109 (1999).
34. R. D. Clark, A. Strizhev, J. M. Leonard, J. F. Blake, and J. B. Matthew. Consensus scoring for ligand/protein interactions. *J. Mol. Graphics Mod* 20:281–295 (2002).
35. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935 (1983).
36. J. P. Ryckaert, G. Cicotti, and H. J. C. Berensden. Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes. *J. Comput. Phys.* 23:327–341 (1977).
37. O. Gotoh. Substrate recognition sites in cytochrome P450 family 2 (CYP2) proteins inferred from comparative analyses of amino acid and coding nucleotide sequences. *J. Biol. Chem.* 267:83–90 (1992).
38. M. R. Wester, E. F. Johnson, C. Marques-Soares, P. M. Dansette, D. Mansuy, and C. D. Stout. Structure of a substrate complex of mammalian cytochrome P450 2C5 at 2.3 Å resolution: evidence for multiple substrate binding modes. *Biochemistry* 42:6370–6379 (2003).
39. P. A. Williams, J. Cosme, A. Ward, H. C. Angove, D. M. Vinkovi, and H. Jhoti. Crystal structure of human cytochrome P450 2C9 with bound warfarin. *Nature* 424:464–468 (2003).