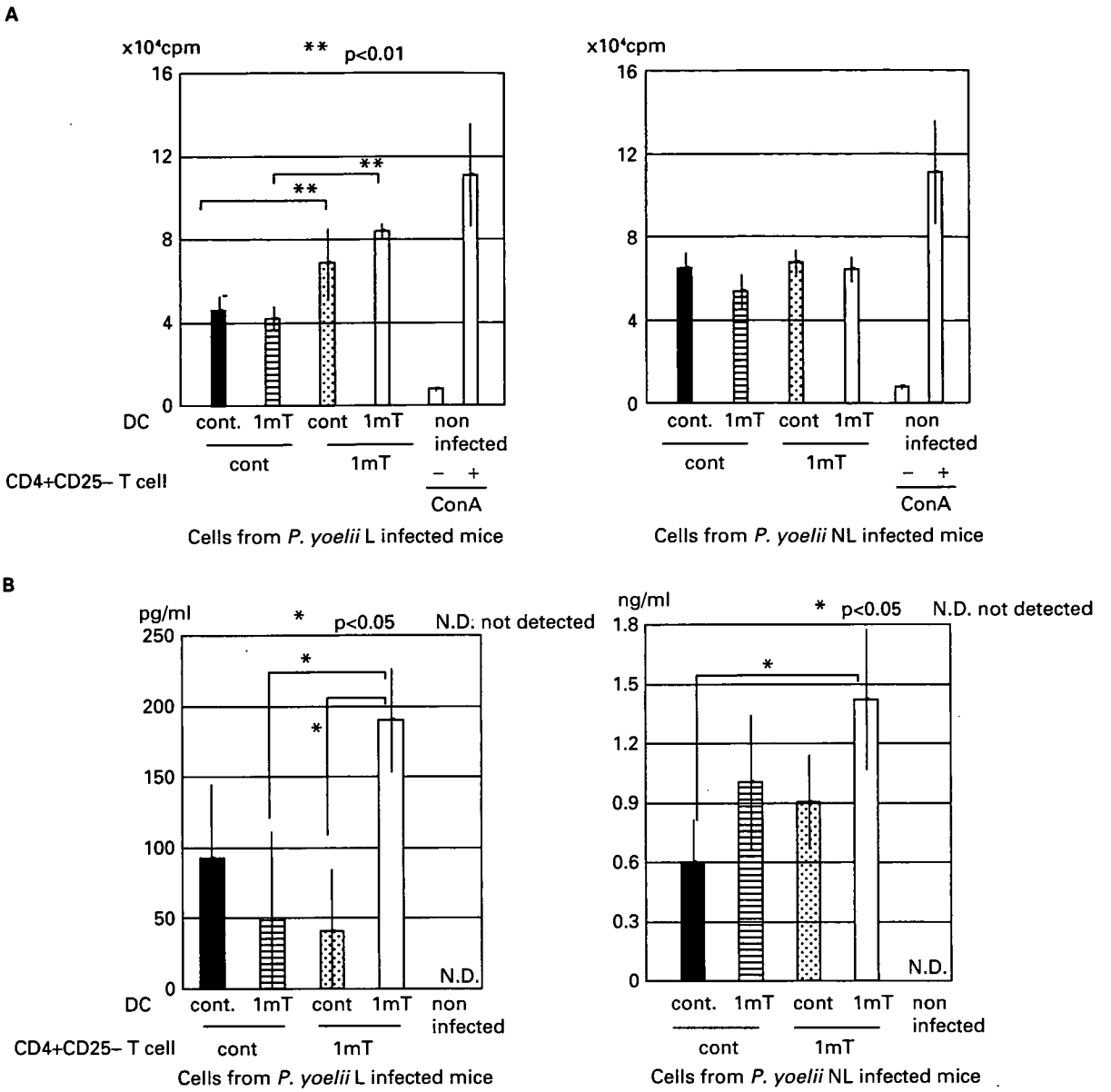


Fig. 4. Enhanced immune responses in PyL-infected mice treated with 1mT. (A) Proliferation of spleen $CD4^+CD25^-$ T cells isolated from mice infected with PyL (left panel) or PyNL (right panel), and treated with or without (control) 1mT. T cells and DCs were obtained 5 days after infection, and T cells ($1 \times 10^5$) were cultured with PyL-parasitized RBCs ($1 \times 10^5$) in the presence of DCs ($1 \times 10^4$) from the indicated mice. One of the non-infected mouse-derived samples was cultured with ConA as an assay positive control. Results represent the mean ± S.D. of triplicate cultures. Asterisks represent statistical significance ($P < 0{\cdot}01$) with Student's *t*-test. (B) $CD4^+CD25^-$ T cells isolated from mice infected with PyL (left panel) or PyNL (right panel), and treated with or without (control) 1mT were cultured as in (A). IFN-$\gamma$ in the supernatant was measured by using ELISA. Data are mean ± S.D. ($n = 3$). Asterisks indicate statistical significance ($P < 0{\cdot}05$) with Student's *t*-test. N.D., Means not detected.

IFN-$\gamma$ secreted from $CD4^+$ T cells is very important for protection against blood-stage malaria (Shear *et al.* 1989). We quantified IFN-$\gamma$ concentration in culture supernatants obtained from $CD4^+$ T cells and DC co-culture, as above. $CD4^+$ T cells from PyL-infected mice produced more IFN-$\gamma$ in the presence of DC from the same mice than did cells from control mice (Fig. 4B). In agreement with the *in vivo* infection experiment, IFN-$\gamma$ secretion in PyNL-infected mouse-derived cells did not show a significant difference between 1mT-treated and control mice. When these $CD4^+$ T cells were mixed with DCs from control mice, they did not produce IFN-$\gamma$. These results indicate that when infection-induced hyper-catabolism of tryptophan was inhibited, $CD4^+$ T cells were effectively primed for proliferation and production of IFN-$\gamma$. However, DCs with impaired IDO still played a crucial role

in the recall reaction of IFN-$\gamma$ production by $CD4^+$ T cells.

## DISCUSSION

Here, we provide evidence that infection with malaria parasites resulted in systemic activation of tryptophan catabolism mediated by IDO. Furthermore, inhibition of IDO *in vivo* partially protected mice from lethal infection, in association with enhanced lymphocyte proliferation and IFN-$\gamma$ production by $CD4^+$ T cells specific for malaria parasites. This enhancement was also observed when parasitaemia of cell-origin mice was almost the same between 1mT treated and control mice (data not shown). Interestingly, the effect of IDO inhibition on protective immunity was more obvious in infection with strain PyL than that with strain PyNL, although strain PyNL could activate IDO in a manner comparable to strain PyL. Concerning the non-lethality of the PyNL strain, we observed that *in vitro* IFN-$\gamma$ production was suppressed by IDO, but that parasite growth was not enhanced *in vivo*. This suggested that even with a competent IDO mouse host, the dose of anti-malarial IFN-$\gamma$ that is produced is approximately 3 times more than with IDO-incompetence with PyL strain infection. These results suggest that strain PyL evades host immunity more competitively by utilizing IDO which, in turn, could explain the higher lethality of strain PyL.

The tryptophan catabolism mediated immune suppression was found to be induced in our malaria mouse model which increased parasite load but did not shorten infected mice survival (data not shown). We did not observe any suggestive findings for reasonable interpretation on this rather contradictory result, but it is worthy to be considered that immune pathology is probably accelerated under administration of 1mT. Hyper-catabolism of tryptophan suppresses not only pathogen-reactive immunity but also host-damaging immune responses and both were activated with 1mT treatment and might have cancelled effects on survival with each other.

Although we confirmed neither protein expression nor mRNA transcription for IDO in malaria-infected mice, IDO was apparently activated in our malaria model, for the following reasons: inducible catabolism of tryptophan via the kynurenine pathway is mediated only by IDO, and IDO-specific inhibition reduces tryptophan catabolism. IDO is known to be detected in various organs and tissues such as gut, epididymis, lung and lymph node, especially antigen-presenting cells (Takikawa *et al.* 1986; Munn *et al.* 2004). We speculate that expression of IDO in DCs is responsible for the immunomodulation observed in our model. IDO-positive DCs, especially plasmacytoid DCs (pDCs), have been reported to play the main role in IDO-mediated immunosuppression. They induce suppression of not only cellular, but also humoral immunity (Adikari *et al.* 1988). Interestingly, our data show that antigen-specific proliferation of $CD4^+$ T cells in response to *in vitro* reactivation was not affected by IDO activity in DCs. These results can be interpreted as IDO-expressing DCs induce immunosuppression during T cell priming, and T cells primed by $IDO^+$ DCs are fated to be unresponsive, with less functional plasticity than DCs.

Another important point is how the malaria parasite stimulates IDO activation. Various signals, such as IFN-$\alpha$, IFN-$\gamma$, cytotoxic T lymphocyte antigen (CTLA)-4, and Toll-like receptor 9 (TLR9) ligation (CpG-oligonucleotides), are known to induce IDO in DCs (Mellor *et al.* 2004). During malaria, powerful inflammatory responses occur, and several cytokines like IFNs are secreted, which could be IDO-inducing factors. More directly, malaria parasites might stimulate, with their TLR9 ligands, the malaria pigment haemozoin (Coban *et al.* 2005) and some protein molecules (Pichyangkul *et al.* 2004). Our results and those of previous studies (Hisaeda *et al.* 2004) reveal that malaria parasites induce regulatory T cells, one of whose suppressive mechanisms is CTLA4 molecules. It is possible that malaria-induced regulatory T cells activate IDO in DCs via CTLA4 stimulation, and that IDO-expressing DCs might function as suppressive effector mechanisms.

In addition to immunosuppression, IDO has been reported to be involved in the host–parasite relationship in malaria. In a cerebral malaria model, IDO has been reported to be associated with pathology (Mitchell *et al.* 2005). Indeed, the expression of IDO has been confirmed in the brain vascular endothelial cells in cases of cerebral malaria, using immunohistological techniques (Hansen *et al.* 2000). On the other hand, one of the metabolites of tryptophan, xanthurenic acid, has been found to be a gametocyte activating factor (Billker *et al.* 1998) which enhances the development of gametocytes into sporozoites, the infectious form of the malaria parasite, in mosquito vectors. It will be important to determine the importance of IDO in the pathology of cerebral malaria, and in transmission of malaria parasites to mosquito vectors.

In conclusion, our results demonstrate that IDO-mediated immunosuppression contributes to immune evasion by malaria parasites, providing a new insight into the host–parasite relationship. As IDO inhibition is a potent method for reducing parasite load and pathological complications, understanding the molecular mechanisms underlying IDO activation may be helpful in establishing therapeutic strategies against malaria.

REFERENCES

**Adikari, S. B., Lian, H., Link, H., Huang, Y. M. and Xiao, B. G.** (1988). Interferon-gamma-modified dendritic cells suppress B cell function and ameliorate the development of experimental autoimmune myasthenia gravis. *Journal of Interferon Research* **8**, 691–702.

**Billker, O., Lindo, V., Panico, M., Etienne, A. E., Paxton, T., Dell, A., Rogers, M., Sinden, R. E. and Morris, H. R.** (1998). Identification of xanthurenic acid as the putative inducer of malaria development in the mosquito. *Nature, London* **392**, 289–292.

**Cady, S. G. and Sono, M.** (1991). 1Methyl-DL tryptophan, beta-(3-benzofuranyl)-DL-alanine (the oxygen analog of tryptophan), and beta-[3-benzo(b)thienyl]-DL-alanine (the sulfur analog of tryptophan)are competitive inhibitors for indoleamine 2,3 dioxygenase. *Archives of Biochemistry and Biophysics* **291**, 326–333.

**Coban, C., Ishii, K. J., Kawai, T., Hemmi, H., Sato, S., Uematsu, S., Yamamoto, M., Takeuchi, O., Itagaki, S., Kumar, N., Horii, T. and Akira, S.** (2005). Toll-like receptor 9 mediates innate immune activation by the malaria pigment hemozoin. *Journal of Experimental Medicine* **201**, 19–25.

**Fujioka, H., Kato, N., Fujita, M., Fujimura, K. and Nishiyama, Y.** (1990). Activities of new acridone alkaloid derivatives against *Plasmodium yoelii in vitro*. *Arzneimittel Forschung* **40**, 1026–1029.

**Gurtner, G. J., Newberry, R. D., Schloemann, S. R., McDonald, K. G. and Stenson, W. F.** (2003). Inhibition of indoleamine 2,3-dioxygenase augments trinitrobenzene sulfonic acid colitis in mice. *Gastroenterology* **125**, 1762–1773.

**Hansen, A. M., Driussi, C., Turner, V., Takikawa, O. and Hunt, N. H.** (2000). Tissue distribution of indoleamine 2,3-dioxygenase in normal and malaria-infected tissue. *Redox Report: Communications in Free Radical Research* **5**, 112–115.

**Helleberg, M., Goka, B. Q., Akanmori, B. D., Obeng-Adjei, G., Rodriquies, O. and Kurtzhals, J. A. L.** (2005). Bone marrow suppression and severe anaemia associated with persistent *Plasmodium falciparum* infection in African children with microscopically undetectable parasitaemia. *Malaria Journal* **4**, 56.

**Hisaeda, H., Maekawa, Y., Iwakawa, D., Okada, H., Himeno, K., Kishihara, K., Tsukumo, S. and Yasutomo, K.** (2004). Escape of malaria parasites from host immunity requires CD4+CD25+ regulatory T cells. *Nature Medicine* **10**, 29–30.

**Hisaeda, H., Yasutomo, K. and Himeno, K.** (2005). Malaria: immune evasion by parasites. *International Journal of Biochemistry and Cell Biology* **37**, 700–706.

**Ing, R., Segura, M., Thawani, N., Tam, M. and Stevenson, M. M.** (2006). Interaction of mouse dendritic cells and malaria-infected erythrocytes: uptake, maturation, and antigen presentation. *Journal of Immunology* **176**, 441–450.

**Mellor, A. L. and Munn, D. H.** (2004). IDO expression by dendritic cells: tolerance and tryptophan catabolism. *Nature Reviews Immunology* **4**, 762–774.

**Millington, O. R., Di Lorenzo, C., Phillips, R. S., Garside, P. and Brewer, J. M.** (2006). Suppression of adaptive immunity to heterologous antigens during plasmodium infection through hemozoin-induced failure of dendritic cell function. *Journal of Biology (Online)* **5**, 5.

**Mitchell, A. J., Hansen, A. M., Hee, L., Ball, H. J., Potter, S. M., Walker, J. C. and Hunt, N. H.** (2005) Early cytokine production is associated with protection from murine cerebral malaria. *Infection and Immunity* **73**, 5645–5653.

**Munn, D. H., Zhou, M., Attwood, J. T., Bondarev, I., Conway, S. J., Marshall, B., Brown, C. and Mellor, A. L.** (1998). Prevention of allogeneic faetal rejection by tryptophan catabolism. *Science* **281**, 1191–1193.

**Munn, D. H., Sharma, M. D., Hou, D., Baban, B., Lee, J. R., Antonia, S. J., Messina, J. L., Chandler, P., Koni, P. A. and Mellor, A. L.** (2004). Expression of indoleamine 2,3-dioxygenase by plasmacytoid dendritic cells in tumour-draining lymph nodes. *Journal of Clinical Investigation* **114**, 280–290.

**Ocana-Mongner, C., Mota, M. M. and Rodriguez, A.** (2003). Malaria blood stage suppression of liver stage imunity by dendritic cells. *Journal of Experimental Medicine* **197**, 143–151.

**Pawlak, D., Tankiewicz, A., Mysliwiec, P. and Buczko, W.** (2002). Tryptophan metabolism via the kynurenine pathway in experimental chronic renal failure. *Nephron* **90**, 328–335.

**Pichyangkul, S., Yongvanitchit, K., Kum-arb, U., Hemmi, H., Akira, S., Krieg, A. M., Heppner, D. G., Stewart, V. A., Hasegawa, H., Looareesuwan, S., Shanks, G. D. and Miller, R. S.** (2004). Malaria blood stage parasites activate human plasmacytoid dendritic cells and murine dendritic cells through a Toll-like receptor 9-dependent pathway. *Journal of Immunology* **172**, 4926–4933.

**Rescigno, M. and Borrow, P.** (2001). The host-pathogen interaction: new themes from dendritic cell biology. *Cell* **106**, 267–270.

**Sakurai, K., Zou, J. P., Tschetter, R., Ward, J. M. and Shearer, G. M.** (2002). Effect of indoleamine 2,3-dioxygenase on induction of experimental autoimmune encephalomyelitis. *Journal of Neuroimmunology* **129**, 186–196.

**Sanni, L. A., Thomas, S. R., Tattam, B. N., Moore, D. E., Chaudhri, G., Stocker, R. and Hunt, N. H.** (1998). Dramatic change in oxidative tryptophan metabolism along the kynurenine pathway in experimental cerebral and noncerebral malaria. *American Journal of Pathology* **152**, 611–619.

**Shear, H. L., Srinivasan, R., Nolan, T. and Ng, C.** (1989). Role of IFN-gamma in lethal and nonlethal malaria in susceptible and resistant murine hosts. *Journal of Immunology* **143**, 2038–2044.

**Takikawa, O., Yoshida, R., Kido, R. and Hayaishi, O.** (1986). Tryptophan degradation in mice initiated by indoleamine 2,3 dioxygenase. *Journal of Biological Chemistry* **261**, 3648–3653.

**Tosta, C. E., Sedegah, M., Henderson, D. C. and Wedderburn, N.** (1980). *Plasmodium yoelii* and *Plasmodium berghei*: isolation of infected erythrocytes

from blood by colloidal silica gradient centrifugation. *Experimental Parasitology* **50**, 7–15.

**Urban, B. C., Ferguson, D. J. P., Pain, A., Willcox, N., Plebanski, M., Austyn, J. M. and Roberts, D. J.** (1999). *Plasmodium falciparum*-infected erythrocytes modulate the maturation of dendritic cell. *Nature, London* **400**, 73–77.

**Uyttenhove, C., Pilotte, L., Theate, I., Stroobant, V., Colau, D., Parmentier, N., Boon, T. and van den Eynde, B. J.** (2003). Evidence for a tumoral immune resistance mechanism based on tryptophan degradation by indoleamine 2,3 dioxygenase. *Nature Medicine* **9**, 1269–1274.

**Xu, H., Wipasa, J., Yan, H., Zeng, M., Makobongo, M. O., Finkelman, F. D., Kelso, A. and Good, M. F.** (2002). The mechanism and significance of deletion of parasite-specific CD4+ T cells in malaria infection. *Journal of Experimental Medicine* **195**, 881–892.

JOURNAL OF BACTERIOLOGY, Nov. 2007, p. 8241–8249 Vol. 189, No. 22
0021-9193/07/$08.00+0 doi:10.1128/JB.00550-07


# Control of Cell Wall Assembly by a Histone-Like Protein in Mycobacteria[▽]

Tomoya Katsube,[1,2] Sohkichi Matsumoto,[1]* Masaki Takatsuka,[1,2] Megumi Okuyama,[1] Yuriko Ozeki,[3] Mariko Naito,[4] Yukiko Nishiuchi,[1] Nagatoshi Fujiwara,[1] Mamiko Yoshimura,[1] Takafumi Tsuboi,[5] Motomi Torii,[6] Nobuhide Oshitani,[2] Tetsuo Arakawa,[2] and Kazuo Kobayashi[7]

*Department of Host Defense[1] and Department of Gastroenterology,[2] Osaka City University Graduate School of Medicine, 1-4-3 Asahi-machi, Abeno-ku, Osaka 545-8585, Japan; Sonoda Women's University, 7-29-1 Minamitsukaguchi-cho, Amagasaki, Hyogo, Japan[3]; Division of Microbiology and Oral Infection, Department of Molecular Microbiology and Immunology, Nagasaki University Graduate School of Biomedical Sciences, 1-7-1 Sakamoto, Nagasaki 852-8588, Japan[4]; Cell-Free Science and Technology Research Center and Venture Business Laboratory, Ehime University, 3 Bunkyo-cho, Matsuyama, Ehime 790-8577, Japan[5]; Department of Molecular Parasitology, Ehime University Graduate School of Medicine, Toon, Ehime 791-0295, Japan[6]; and Department of Immunology, National Institute of Infectious Diseases, Toyama 1-23-1, Shinjuku-ku, Tokyo 162-8640, Japan[7]*

Received 11 April 2007/Accepted 1 September 2007

**Bacteria coordinate assembly of the cell wall as well as synthesis of cellular components depending on the growth state. The mycobacterial cell wall is dominated by mycolic acids covalently linked to sugars, such as trehalose and arabinose, and is critical for pathogenesis of mycobacteria. Transfer of mycolic acids to sugars is necessary for cell wall biogenesis and is mediated by mycolyltransferases, which have been previously identified as three antigen 85 (Ag85) complex proteins. However, the regulation mechanism which links cell wall biogenesis and the growth state has not been elucidated. Here we found that a histone-like protein has a dual concentration-dependent regulatory effect on mycolyltransferase functions of the Ag85 complex through direct binding to both the Ag85 complex and the substrate, trehalose-6-monomycolate, in the cell wall. A histone-like protein-deficient *Mycobacterium smegmatis* strain has an unusual crenellated cell wall structure and exhibits impaired cessation of glycolipid biosynthesis in the growth-retarded phase. Furthermore, we found that artificial alteration of the amount of the extracellular histone-like protein and the Ag85 complex changes the growth rate of mycobacteria, perhaps due to impaired down-regulation of glycolipid biosynthesis. Our results demonstrate novel regulation of cell wall assembly which has an impact on bacterial growth.**

Bacteria organize biogenesis of the cell wall as well as synthesis of cellular components depending on the growth state. However, factors linking the growth state and cell wall biogenesis have not been identified.

*Mycobacterium tuberculosis* is a top killer among bacterial pathogens and is responsible for 2 million deaths annually (6). *M. tuberculosis* can be quiescent in host cells for a long period of time, growing very slowly or present in a dormant state without multiplication, and it latently infects one-third of the world's human population (22, 43). In 5 to 10% of infected hosts the bacterium reactivates and causes progressive disease during their lifetimes. Most cases of active tuberculosis do not result from the initial infection but instead represent reactivation of previously implanted bacteria (22, 43).

The cell wall is critical for long-term persistence of *M. tuberculosis* in the hostile environment in the host cells and for progression of tuberculosis (3, 7). Mycobacteria are gram-positive bacilli, but the cell wall structures are different from those of other gram-positive bacteria. Approximately one-half of the cell wall mass is comprised of large ($C_{70}$ to $C_{90}$) branched-chain fatty acids called mycolic acids. Mycolic acids are distributed in acid-fast positive bacteria, such as *Mycobacterium*, *Corynebacterium*, *Rhodococcus*, and *Nocardia*, although mycobacterial mycolic acids are the longest mycolic acids and have the largest side chains ($C_{20}$ to $C_{24}$). The cell wall outer layer is composed of extractable glycolipids containing mycolic acids, such as trehalose-6-monomycolate (TMM) and trehalose-6,6′-dimycolate (TDM) (also called cord factor), while the inner layer is composed of mycolic acids covalently linked to the distal portion of the arabinogalactan (AG) moiety (7). TMM- or TDM-derived mycolic acids are transferred to other TMM to synthesize TDM and also to peptidoglycan-linked AG to construct the inner layer of the envelope. Antigen 85 (Ag85) complex proteins (Ag85A, Ag85B, and Ag85C) are mycolyltransferases and catalyze transfer of mycolic acids to free trehalose, TMM, and TDM (4). Ag85 complex proteins are also believed to catalyze the transfer of mycolic acids from TMM or TDM to peptidoglycan-linked AG, because inactivation of Ag85C reduced the level of AG-linked mycolic acids (16).

Regulatory proteins involved in cell wall assembly should localize in the mycobacterial cell wall. We and other groups found that a histone-like DNA-binding protein, which was designated mycobacterial DNA-binding protein 1 (MDP1), laminin-binding protein, histone-like protein (HLP), or HupB, not only localizes in the cytoplasmic space but also occurs externally or is in the mycobacterial cell wall (26, 36, 38).

* Corresponding author. Mailing address: Department of Host Defense, Osaka City University Graduate School of Medicine, 1-4-3 Asahi-machi, Abeno-ku, Osaka 545-8585, Japan. Phone: 81-6-6645-3746. Fax: 81-6-6645-3747. E-mail: sohkichi@med.osaka-cu.ac.jp.

▽ Published ahead of print on 14 September 2007.

MDP1 is mycobacterium-specific histone-like protein. *M. tuberculosis* has a single *mdp1* gene (Rv2986c, also called *hupB*) (9, 13), and the *mdp1* gene is conserved even in *Mycobacterium leprae*, which lost many genes during evolution (10). MDP1 likely plays a significant role in DNA functions in mycobacteria, as a transposon-based screen suggested that *mdp1* is essential in *M. tuberculosis* (34). However, the *mdp1* gene can be knocked out in *Mycobacterium smegmatis*, suggesting that another DNA-binding protein may compensate for loss of MDP1 in *M. smegmatis*. The *M. smegmatis* MDP1 knockout (KO) strain exhibited growth kinetics similar to those of the wild type in anaerobic culture (19) but was unable to resume growth at 10°C (37), suggesting that MDP1 plays an important role during stress responses in *M. smegmatis*.

Accumulation of MDP1 in stationary phase or under anaerobic conditions implies that MDP1 is a possible factor that participates in growth-state-dependent regulation of cell wall assembly through binding to sacchariferous components, such as glycolipids, in the cell wall. Based on this hypothesis, here we examined the physiological role of MDP1 in the cell wall.

## MATERIALS AND METHODS

**Extraction and purification of glycolipids.** *M. tuberculosis* H37Rv and *Mycobacterium bovis* bacillus Calmette-Guérin (BCG) were cultivated on Sauton medium at 37°C. *M. smegmatis* was cultured in Luria-Bertani (LB) medium at 37°C. Bacteria were autoclaved for 10 min, disrupted ultrasonically, and then suspended in chloroform-methanol (4:1, 3:1, or 2:1, vol/vol) to extract lipids. The chloroform layer was collected and dried. TDM was first partially purified by precipitation with acetone, chloroform-methanol (2:1, vol/vol), and tetrahydrofuran-methanol (1:2, vol/vol), followed by passage through a column of silica gel (Wakogel C-200; Wako Pure Chemical, Osaka, Japan) with chloroform-methanol (4:1, vol/vol). The purity of TDM was demonstrated by a single spot on a thin-layer chromatogram. TMM was separated by preparative thin-layer chromatography (TLC) on a silica gel plate (Uniplate; 20 by 20 cm; 250 mm; Analtech, Inc., Newark, DE) using a chloroform-methanol-acetone-acetic acid (80:20:6:1, vol/vol/vol/vol) solvent system. Glycolipids were visualized with a 20% $H_2SO_4$ spray, followed by charring at 200°C for analytical purposes or with iodine vapor for a few minutes for preparative purposes. TMM was recovered from the plate immediately after the iodine color had disappeared by passing the plate through a small glass column with the solvent chloroform-methanol (2:1, vol/vol). Finally, TMM was purified until a single spot was obtained by repeating TLC.

**Fluorescence microscopy.** MDP1 was purified from BCG by using a method described previously (26). Egg white lysozyme was purchased from Wako. Bovine histone H1 was purchased from Roche Diagnostics. Proteins were labeled with 5(6)-carboxyfluorescein-*N*-hydroxysuccinimide ester (FLUOS) by using a fluorescein protein labeling kit (Roche Diagnostics) according to the manufacturer's instructions. BCG was grown in Middlebrook 7H9-ADC medium at 37°C until the optical density at 600 nm ($OD_{600}$) was 1.5; then it was collected by centrifugation and washed three times with phosphate-buffered saline (PBS) with 10% fetal bovine serum (FBS). FLUOS-labeled proteins, such as MDP1 (50 μg), egg white lysozyme (38 μg), and bovine histone H1 (50 μg), were added to the BCG suspension and incubated at 37°C for 30 min. BCG was washed three times with PBS with 10% FBS and then mounted on a microscope slide and viewed with a confocal scanning laser microscope (LSM510; Carl Zeiss).

**MAb preparation.** Ten micrograms of MDP1 and 10 ng of *M. tuberculosis* DNA were mixed in PBS, emulsified in Freund's incomplete adjuvant, and injected into BALB/c mice subcutaneously (24). Two weeks later, the mice were boosted in the same way; after an additional 2 weeks, the mice were again boosted by injection of 10 μg of MDP1 in PBS into the tail vein. Three days after the final boost, mice were sacrificed and splenocytes were obtained. A monoclonal antibody (MAb) was prepared essentially as described by Harlow and Lane (15) and was screened by an enzyme-linked immunosorbent assay (ELISA) as described below. After an initial screening, several hybridoma cell lines were cloned by two rounds of limiting dilution. A selected subclone was expanded for freezing and for ascites production in pristane-primed mice. The subclass of the hybridoma subclone was determined with a mouse MAb isotyping kit (Amersham). Immunoglobulins (Ig) were purified from ascites fluid by using an Ampure PA kit (Amersham).

**ELISA to detect interactions between MDP1 and TDM, TMM, mycolic acids, or Ag85 complex proteins.** TMM, TDM, and mycolic acid methyl esters (MAMEs) purified from *M. tuberculosis* Aoyama B were purchased from Nakarai and dissolved in *N*-hexane at a concentration of 50 μg/ml. Glycolipids were immobilized on a 96-well ELISA plate (Sumitomo, Osaka, Japan) by adding 100 μl of glycolipid solution and dried. Ag85A, Ag85B, and Ag85C derived from *M. tuberculosis* H37Rv and bovine serum albumin (BSA) were immobilized on the 96-well ELISA plate (Sumitomo, Osaka, Japan) by incubation of serial twofold dilutions of the protein solutions in sodium bicarbonate buffer (pH 9.6) at 4°C overnight. MDP1 at a concentration of 1 μg/ml in PBS containing 0.05% Tween 20 (PBS-T) was added to wells and incubated for 1 h at 37°C. The wells were washed with PBS-T four times, and then MAb 3A was added. After incubation at 37°C for 1 h, the wells were washed, peroxidase-conjugated anti-mouse antibody (Dako) diluted 1:2,000 was added, and the plate was incubated for 1 h. After the wells were washed, the level of MDP1 binding was detected by color development with *o*-phenylenediamine dihydrochloride (Wako, Tokyo, Japan) and measuring the $OD_{492}$.

**Preparation of subcellular fractions.** Subcellular fractions were prepared by a method described previously (1, 27). Briefly, after BCG and *M. smegmatis* were cultured in Sauton medium and LB medium, respectively, bacteria were collected by centrifugation at 10,000 × *g* and suspended in ice-cold PBS. The bacteria were then disrupted with a Bioruptor UCD-200T sonicator (Toso), and each suspension was centrifuged at 3,000 × *g* for 5 min to remove unbroken bacteria. The supernatant was used as the total cellular fraction. The total cellular fraction was fractionated further to obtain a cell wall fraction and a non-cell-wall fraction using the following procedure. The total cellular fraction was centrifuged at 10,000 × *g* for 10 min. The pellet was rinsed with cold PBS and centrifuged again at 10,000 × *g* for 10 min. The supernatants were designated the non-cell-wall fraction. The cell wall-containing pellet was suspended again in ice-cold PBS, and then Percoll (Amersham Biosciences) was added to a concentration of 60% and mixed by vortexing. Next, the cell wall-containing fraction was centrifuged at 27,000 × *g* for 1 h to separate the cell walls from the unbroken cells completely. The cell wall band was collected and washed twice with PBS, and it was designated the cell wall fraction.

**Immunoprecipitation assay.** Cell walls derived from BCG as described above were precleaned by using 25 μg of mouse IgG (Chemicon International Temecula) or rabbit Ig in 100 μl of protein G-coupled Sepharose (Amersham Pharmacia Biotech), incubated at 4°C for 2 h, and centrifuged. The supernatants were colleted and incubated with 250 μg of anti-MDP1 MAb 3A, control mouse IgG, anti-Ag85 Ig, or control rabbit Ig at 4°C for 16 h, and then 300 μl of protein G-coated Sepharose was added and the preparation was incubated for 5 h at 4°C. Then the beads were washed with PBS-T three times. After washing, bead-bound proteins were eluted by boiling the preparations in 40 μl of 2× sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) sample buffer (0.125 M Tris-HCl [pH 6.8], 4% SDS, 20% glycerol, 10% mercaptoethanol). The samples were then fractionated on a 12.5% SDS-polyacrylamide gel and transferred to a polyvinylidene difluoride membrane. The membrane was then blocked for 30 min at room temperature by incubating it in PBS containing 5% skim milk. Then the membrane was probed with anti-MDP1 MAb 3A or anit-Ag85 antibodies overnight at 4°C. After probing, the membrane was washed four times with PBS-T. Next, the membrane was incubated for 4 h at room temperature with peroxidase-conjugated anti-mouse IgG or anti-rabbit IgG (Dakopatts A/S, Denmark) diluted 1:10,000. The membrane was then washed as described above, and immunoreactive bands were visualized by using an ECL Western blot detection reagent (Amersham Bioscience, Buckinghamshire, United Kingdom) according to the manufacturer's instructions.

Bead-bound glycolipids were eluted with chloroform-methanol (3:1, vol/vol). Thirty microliters of the chloroform layer was spotted on a TLC plate (HPTLC plate; 10 by 10 cm; Silica Gel 60; Merck, Darmstadt, Germany) and developed with the chloroform-methanol-acetone-acetic acid (80:20:6:1, vol/vol/vol/vol) solvent system. The TLC plate was exposed overnight to an immunoprecipitation plate (BAS-MS2025 or BAS-SR2025; Fujifilm, Japan) and visualized with the BAS system (BAS-5000). The radioactivity of separated spots was quantified by using the BAS system's software.

**Analysis of mycolyltransferase activity.** The mycolyltransferase catalytic reaction was analyzed by using the method developed by Belisle et al. (4). Twenty-five micrograms of TMM purified from *M. tuberculosis* H37Rv was immobilized in each glass vial, and then PBS containing 4 μl of a 1-mg/ml dithiothreitol solution and 0.5 μCi of [$^{14}$C]trehalose (American Radiolabeled Chemicals Inc.) were added. Two hundred micrograms of culture filtrate or 20 μg of Ag85 complex protein and various amounts of MDP1 or BSA were mixed to obtain a total

volume of 200 μl and incubated for 30 min at 37°C. Glycolipids were eluted with chloroform-methanol (2:1, vol/vol), 10 μl of the chloroform layer was spotted on a TLC plate (HPTLC plate; 10 by 10 cm; Silica Gel 60; Merck, Darmstadt, Germany), and the plate was developed with the chloroform-methanol-acetone-acetic acid (80:20:6:1, vol/vol/vol/vol) solvent system. TLC plates were analyzed by using the BAS system described above.

**Construction of HLP/MDP1 complemented strain.** Based on the *mdp1/hlp* nucleotide sequences, two oligonucleotide primers, forward primer 5'-GGGAA GCTTATTCGCCGCCCACCTAGT-3' and reverse primer 5'-TAACGCACCA ACGCGAAA-3', were purchased from Sigma Genosys. A PCR was carried out by targeting 10 ng chromosomal DNA of *M. smegmatis* $MC^2155$ with an automated thermal sequencer (Perkin Elmer). The samples were first denatured by heating them at 94°C for 5 min; then they were subjected to 30 cycles of 94°C for 1 min, 58°C for 1 min, and 72°C for 3 min and were finally incubated for 5 min at 72°C. An amplified 0.9-kb DNA fragment, which contained the promoter and structural gene of HLP/MDP1, was cloned into pGEM-T (Promega) utilizing a ligation kit (version 1, Takara), sequenced, excised by digestion with HindIII and NotI, and ligated to the same site of pMV306-Hyg, a one-copy integrated vector for the phage attachment site. pMV306-Hyg was provided by H. I. Boshoff (National Institutes of Health, Bethesda, MD). The resulting plasmid was introduced into an *M. smegmatis* HLP/MDP1 KO strain by a standard electroporation procedure (17, 39), and then a hygromycin-resistant colony was selected. Complementation was confirmed by protein expression with Western blotting using anti-MDP1 MAb 3A (data not shown).

**SEM analysis.** Five-milliliter portions of culture aliquots were concentrated by centrifugation (3,000 × *g*) before suspension in fresh Middlebrook 7H9-ADC medium. The $OD_{600}$ was adjusted to 0.1, and then the preparations were placed on poly-L-lysine-coated Thermanox coverslips in 24-well tissue culture plates. The bacteria were allowed to settle for 30 min before gentle decanting and addition of 1 ml of a solution containing 2% paraformaldehyde and 2.5% glutaraldehyde in 0.1 M sodium cacodylate buffer (pH 7.4) with 0.2 M sucrose. The samples were incubated at 4°C overnight before treatment with 2% $OsO_4$ in 0.1 M cacodylate buffer for 2 h at room temperature. A series of sequential ethanol dehydration steps were performed (50, 70, 95, and 100% ethanol, 10 min each) before samples were dried under $CO_2$ using a critical point drier (HCP-2; Hitachi). Samples were Pt-Pd sputter coated (E-1030; Hitachi) and imaged with an Hitachi S-4700 scanning electron microscope (SEM). The SEM analysis was performed in triplicate using three independently grown cultures.

**Chase of glycolipid synthesis.** *M. smegmatis* strains, including the wild type, the MDP1 KO mutant, and the complemented strain, were precultured at 37°C in LB broth (Sigma) containing 1-mm-diameter glass beads. First, bacterial clumps were disrupted with the beads by vortexing, and then the $OD_{600}$ was adjusted to 0.1. Then 5 μl of each bacterial suspension was added to 5 ml of fresh medium containing glass beads. [1-$^{14}$C]acetic acid (sodium salt; 37 MBq/ml; PerkinElmer Life & Analytical Sciences, Massachusetts) was added at a final concentration of 1 μCi/ml on day 3 or7. Each bacterial medium was incubated for 16 h and diluted to obtain 6.0 × $10^7$CFU/ml. Two milliliters of this bacterial suspension was collected by centrifugation and washed with pure water three times. After the supernatants were discarded, we added 1 ml of chloroform-methanol (3:1, vol/vol) and sonicated the preparation for 30 min. Thirty microliters of the chloroform layer was spotted on a TLC plate. The radioactivity of the separated spots was quantified by using the BAS system software as described above.

**MALDI-TOF mass spectrometry analysis.** A matrix-assisted laser desorption ionization–time of flight (MALDI-TOF) mass spectrometry analysis was carried out by using an Ultraflex mass spectrometer (Bruker Daltonics) in the reflectron mode. Samples were dissolved in chloroform-methanol (2: 1, vol/vol) at a concentration of 1 mg/ml and applied to the sample plate as droplets. 2,5-Dihydroxybenzoic acid was used as the matrix. The accelerating voltage was 20 kV.

**Construction of BCG-Luc and assessment of its growth.** Linker DNAs including the Shine-Dalgarno sequence (AGCTTAGTACTGGATCCGAGGACC TGCC and GATCGGCAGGTCCTCGGATCCAGTACTA) were synthesized by Sigma Genosys. pGEM-Luc (Promega) was digested with both BamHI and HindIII, and annealed linker DNA was inserted by ligation utilizing a ligation kit (version 1; Takara). The construct was then digested with HindIII and StuI, and the gene fragment containing the Shine-Dalgarno sequence and the luciferase gene was inserted into pSO246 (25), which had been digested with BamHI, blunt ended with T4 DNA polymerase, and digested with HindIII. The resulting plasmid was designated pSO-Luc. pMV261 (39) was digested with KpnI and HindIII, and the Hsp60 promoter region was inserted into the same site of pSO-Luc. The final construct was introduced into BCG by electroporation by the method described previously (25), and the kanamycin-resistant BCG-Luc strain was obtained. BCG-Luc was grown at 37°C in Middlebrook 7H9-ADC medium until mid-log phase and collected by centrifugation. Bacteria were suspended in RPMI 1640 (Sigma) containing 10% FBS with or without MDP1 or Ag85 complex proteins in a 96-well tissue culture plate (Becton Dickinson) and incubated at 37°C. Luciferase activity was determined at each time point as described previously (12).

**Statistical analyses.** Data were analyzed by using a Power Macintosh G5 and StatView 5.0 (SAS Institute Inc.) and were expressed as means ± standard deviations. Data that appeared to be statistically significantly different were compared by using an analysis of variance for comparing the means of multiple groups and were considered significantly different if the *P* value was less than 0.05.

## RESULTS

**MDP1 binds externally to the cell wall of BCG.** MDP1 acts as an adhesin on the envelope through interaction with glycosaminoglycans on the host cell surface (1, 38). Because MDP1 is retained in the cell wall, we hypothesized that MDP1 is tightly bound to some unknown targets in the cell wall. In order to assess this possibility, first we examined the binding of MDP1 to the cell wall. We incubated FLUOS-labeled MDP1 with BCG and investigated the interaction by fluorescence microscopy. The results revealed that MDP1 (pI 12.4; molecular mass, 21 kDa) bound to the surface of BCG (Fig. 1A). In contrast, other FLUOS-labeled basic proteins, such as egg white lysozyme (pI 9.2; molecular mass, 16 kDa) (Fig. 1B) and bovine histone H1 (pI 11.5; molecular mass, 21 kDa) (Fig. 1C), did not interact with BCG. Bovine serum albumin (pI 5.4) did not interact either (data not shown). Thus, MDP1 can specifically bind externally to the cell wall.

**MDP1 binds to TMM and TDM but not to free mycolic acids.** Mycolic acids form an ordered structure in the envelope and are believed to be some of the outermost covalently linked elements (3, 20). *M. tuberculosis* produces three kinds of mycolic acids, the alpha-, methoxy-, and keto-mycolates. Three types of MAMEs (alpha-, methoxy-, and keto-mycolates), as well as TMM and TDM derived from *M. tuberculosis* strain Aoyama B, were immobilized on an ELISA plate and reacted with MDP1, and the interaction was detected with anti-MDP1 MAb 3A (subclass IgG1, light chain κ). The affinity of MAb 3A for MDP1 was 1.39 $e^{-9}$ M as determined by surface plasmon resonance analysis with a Biacore biosensor (Biacore) (data not shown). This MAb binds to neither TMM nor TDM (data not shown). The results showed that MDP1 bound to TMM and TDM but not to any type of MAME (Fig. 1D). This indicates that MDP1 specifically recognizes the covalent linkage of mycolic acids to the 6-hydroxyl group of trehalose, because MDP1 does not bind to any mycolic acid (Fig. 1D) or free trehalose (1).

**Physiological interaction between MDP1 and glycolipid in the mycobacterial cell wall.** Next we examined whether MDP1 actually associated with TMM and TDM in the mycobacterial envelope by using an immunoprecipitation assay. A cell wall fraction derived from BCG by the method described previously (27) was incubated with MAb 3A or control mouse IgG. Then MDP1-bound glycolipids were precipitated with protein G-coupled beads, extracted with an organic solvent (chloroform-methanol, 3:1 [vol/vol]), and analyzed by TLC. Both TMM and TDM were precipitated by MAb 3A but not by control IgG (Fig. 1E), showing that MDP1 associated with TDM and TMM in the cell wall. The data also showed that MDP1 bound to other unknown lipids that migrated above TDM, as well as below TMM (Fig. 1E).

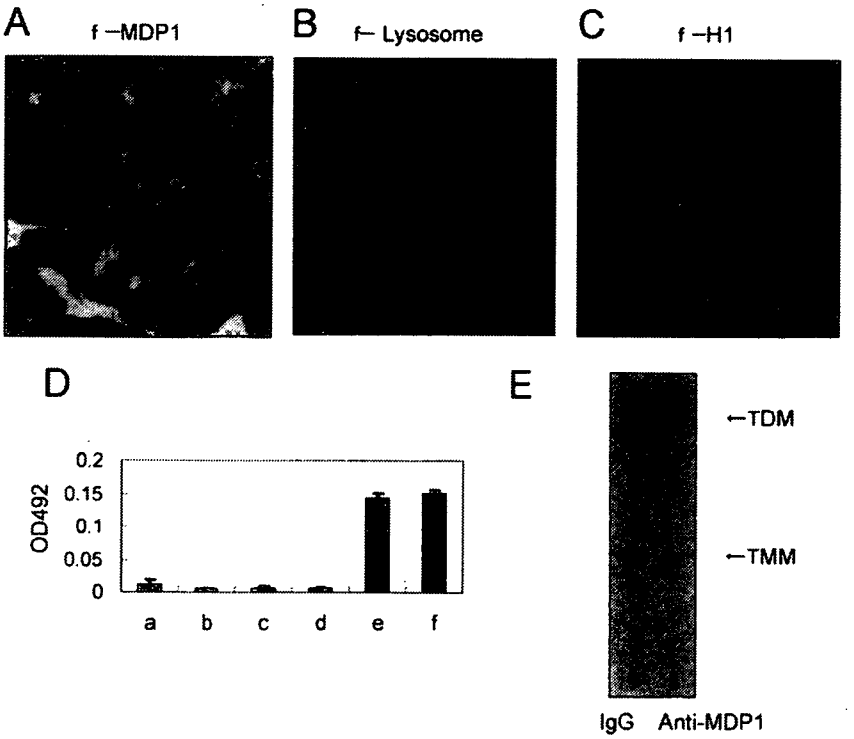


FIG. 1. MDP1 binds to the cell wall of BCG. (A to C) BCG was incubated with FLUOS-labeled MDP1 (A), egg white lysozyme (B), and bovine histone H1 (C). Bacteria were viewed with a confocal scanning laser microscope. (D) ELISA to detect MDP1-lipid or -glycolipid interactions. Mycolic acids (bar b, alpha-MAME; bar c, methoxy-MAME; bar d, keto-MAME) and glycolipids (bar e, TMM; bar f, TDM) were immobilized on an ELISA plate and reacted with MDP1, and the level of binding was detected with anti-MDP1 MAb 3A. Bar a shows the results for a negative control without immobilization of either lipids or glycolipids. OD492, optical density at 492 nm. (E) Immunoprecipitation assay to detect physiological interaction of MDP1 and glycolipid. Cell wall derived from BCG was incubated with anti-MDP1 MAb 3A or control mouse IgG in the presence of protein G-coated Sepharose. Samples were spotted on a TLC plate and developed with the chloroform-methanol-acetone-acetic acid (80:20:6:1, vol/vol/vol/vol) solvent system.

**MDP1 regulates transfer of mycolic acids by Ag85 complex proteins in vitro.** We next examined the effect of MDP1 on transfer of mycolic acids by Ag85 complex proteins, because MDP1 can bind TMM and TDM, which are substrates of Ag85 complex proteins. We purified Ag85 complex proteins, as well as Ag85A, Ag85B, and Ag85C individually, from *M. tuberculosis* H37Rv by the method described previously (28). Purified Ag85 complex proteins transferred mycolic acids to $^{14}$C-labeled trehalose as described previously (4). A molar amount of MDP1 equivalent to that of the Ag85 complex reduced synthesis of TDM and TMM by 44.4% ± 18.4% and 57.4% ± 25.2%, respectively (Fig. 2A and 2B). In contrast, a molar ratio of MDP1 to Ag85 proteins of 1/1,000 enhanced synthesis of TDM and TMM by 36.8% ± 32.4% and 36.4% ± 22.1%, respectively (Fig. 2A and 2B). The same results were obtained when total culture filtrates were used (Fig. 2C, upper chromatogram) or when purified Ag85B was used (Fig. 2C, lower chromatogram). These results demonstrate that MDP1 possesses an activity to control the function of mycolyltransferases.

**MDP1 binds to Ag85 complex proteins.** Because MDP1 controlled transfer of mycolic acids in the presence of excess amounts of substrate (TMM) (Fig. 2), we considered the possibility that MDP1 interacts not only with TMM but also with Ag85 complex proteins, which we examined next. Ag85A, Ag85B, or Ag85C was immobilized on ELISA plates by serial concentration and then incubated with MDP1. The level of binding of MDP1 to Ag85 complex proteins was detected by anti-MDP1 MAb. The results showed that the interaction between MDP1 and each Ag85 complex protein produced standard binding curves (Fig. 3A). BSA did not bind to MDP1. These results suggested that MDP1 binds to all Ag85 proteins.

Ag85 complex proteins bind to fibronectin by the motif conserved in the complex, which is comprised of 11 amino acid residues (29). We examined if this region also participated in the MDP1-Ag85 complex protein interaction. However, neither human fibronectin nor synthetic peptide corresponding to the region from position 98 to position 108 of Ag85B, which inhibits an Ag85 complex-fibronectin interaction, inhibited an MDP1-Ag85 complex protein interaction (data not shown). Thus, Ag85 complex proteins associate with MDP1 through a region other than the fibronectin-binding site.

We next examined whether MDP1 interacted with Ag85 complex proteins in the mycobacterial cell wall. A cell wall fraction derived from BCG was immunoprecipitated with MAb 3A or control mouse IgG and separated by SDS-PAGE. One gel was stained with Coomassie brilliant blue R250, and another gel was used for Western blot analysis. Although we observed only faint bands corresponding to Ag85 complex proteins (Ag85A and Ag85C at 32 kDa and Ag85B at 30 kDa) in the stained gel, in the Western blot analysis we observed that the bands reacted with anti-Ag85 complex protein Ig in the

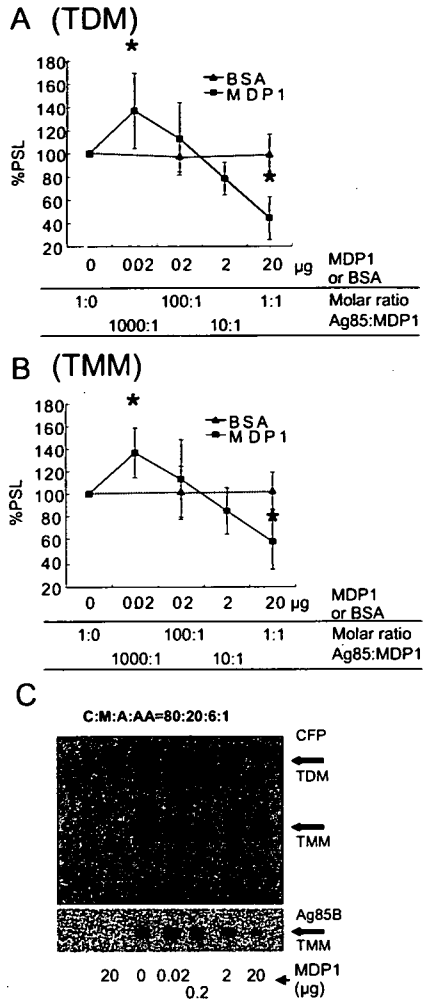


FIG. 2. MDP1 regulates mycolyltransferase activity in vitro. (A and B) Radioactivity of $^{14}$C-labeled TDM (A) and TMM (B) synthesized by Ag85 complex proteins in vitro was quantified by using the BAS system software. Radioactivity was expressed as a percentage of the value for the sample without MDP1, which was defined as 100%. Similar experiments were performed utilizing BSA instead of MDP1, and the results were compared. PSL, photo-stimulated luminescence. Asterisk, $P < 0.05$ for a comparison with control BSA (0 or 20 μg) (as determined by analysis of variance). (C) Culture filtrate (CFP) (upper chromatogram) or Ag85B (lower chromatogram) was added to each vial and incubated for 30 min at 37°C in the presence of various doses of MDP1 and [$^{14}$C]trehalose, and then glycolipids were eluted with chloroform-methanol (2:1, vol/vol). Biosynthesis of glycolipids was analyzed by using the BAS system after TLC plates were developed with the chloroform-methanol-acetone-acetic acid (C:M:A:AA) (80:20:6:1, vol/vol/vol/vol) solvent system. The lane on the left contained a negative control without culture filtrate and Ag85B.

precipitates when anti-MDP1 MAb was used but not when control IgG was used (Fig. 3B). Similar experiments were performed with anti-Ag85 Ig. Anti-Ag85 Ig, but not control Ig, precipitated MDP1 from the cell wall fraction of BCG (Fig. 3C). Taken together, these results suggest that the Ag85 complex proteins associate with MDP1 in the cell wall.

**Subcellular localization of MDP1 in the course of culture.** We analyzed the MDP1 content of the cell wall during the course of culture. We cultured *M. smegmatis* in LB medium for

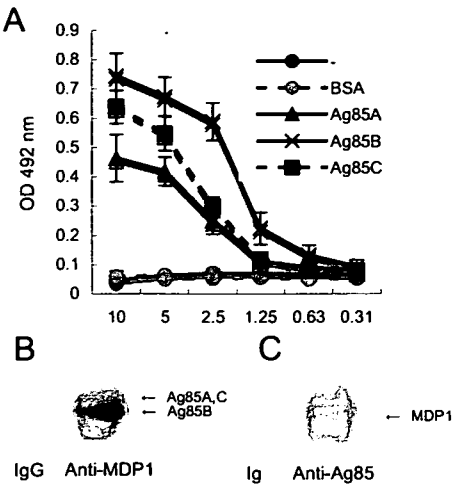


FIG. 3. MDP1 binds to Ag85 complex proteins. (A) ELISA to detect the MDP1-Ag85 complex protein interaction. BSA, Ag85A, Ag85B, and Ag85C were immobilized on an ELISA plate by incubating various concentrations of proteins and reacted with MDP1, and the binding was detected with anti-MDP1 MAb. −, results without immobilized proteins. OD 492 nm, optical density at 492 nm. (B and C) Immunoprecipitation assay to detect a physiological interaction between MDP1 and Ag85 complex proteins. Cell walls derived from BCG were incubated with (B) anti-MDP1 MAb 3A (Anti-MDP1) or control mouse IgG (IgG) or (C) with anti-Ag85 Ig (Anti-Ag85) or control rabbit Ig (Ig) in the presence of protein G-coated Sepharose. Samples for Western blotting were probed with rabbit anti-Ag85 Ig (B) or anti-MDP1 MAb (C).

3, 5, and 7 days, and each subcellular fraction was purified. In this experiment, bacteria grew to stationary phase by day 4 (Fig. 4A). The MDP1 content in each fraction was analyzed by Western blotting. The results showed that MDP1 accumulated in both cell wall and other cellular fractions (membrane, ribosome, and cytoplasmic fractions) at the stationary growth phase (Fig. 5A).

We next carried out a similar experiment with BCG. Total cellular proteins, the cell wall fraction, and other cellular fractions were obtained from BCG after growth for 10, 20, 30, and 60 days on Sauton medium. A Western blot analysis showed that the cellular content of MDP1 increased in both the cell wall and other cellular fractions (Fig. 5B), while the levels of Ag85 complex proteins in the cell wall decreased with time (Fig. 5C).

**Role of MDP1 in glycolipid biosynthesis.** MDP1 is presumed to be essential in slow growers, such as *M. tuberculosis* (34). However, the *mdp1/hlp* gene can be knocked out in *M. smegmatis* (19). In order to determine the physiological role of MDP1 in assembly of the cell wall, we employed an *M. smegmatis* MDP1/HLP KO strain constructed by the group of Thomas Dick (19). We additionally generated an MDP1-complemented strain by insertion of a single copy of the *M. smegmatis* MDP1 gene into the genome of the MDP1/HLP KO strain.

We first analyzed the growth kinetics of the wild type, the MDP1 KO strain, and the complemented strain in LB medium. In this experiment, all strains reached stationary phase on day 4, but the bacterial density of the MDP1 KO strain was lower than that of the wild-type strain during the stationary growth

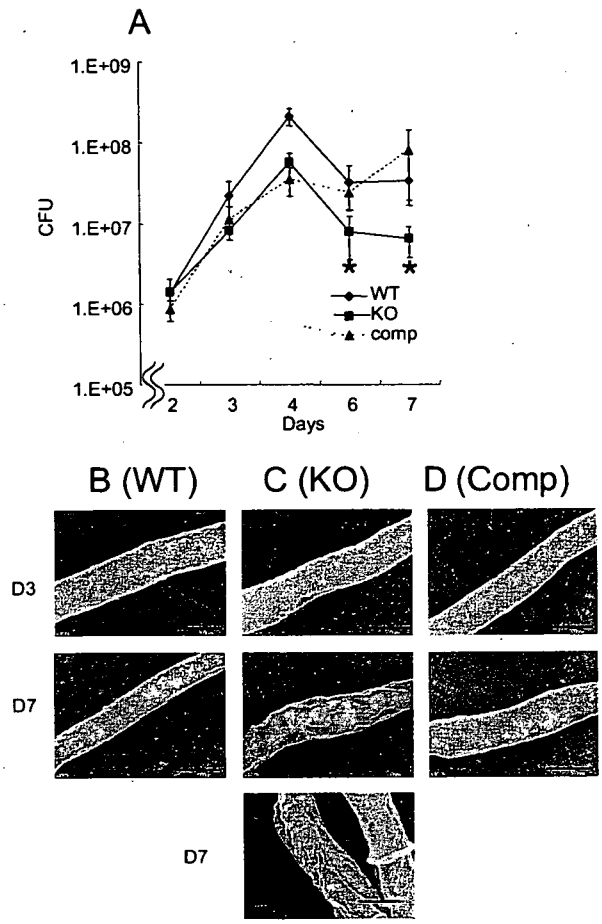


FIG. 4. Effect of MDP1 deficiency on growth kinetics and cell surface morphology. (A) Wild-type *M. smegmatis* mc$^2$155, MDP1 KO, and complemented strains were cultured in LB broth, and bacterial numbers at various time points were determined and expressed as CFU after serially diluted samples on LB agar were harvested. Asterisk, $P < 0.05$ for a comparison with the wild-type strain (as determined by analysis of variance). The results are the results of one representative experiment of seven experiments in which similar results were obtained. (B to D) Visualization of cell surface structure of the wild-type (WT) (B), MDP1 KO (KO) (C), and complemented (Comp) (D) strains by SEM. Bars = 0.5 μm. Bacteria in both exponential (day 3 [D3]) and stationary (day 7 [D7]) phases were analyzed.

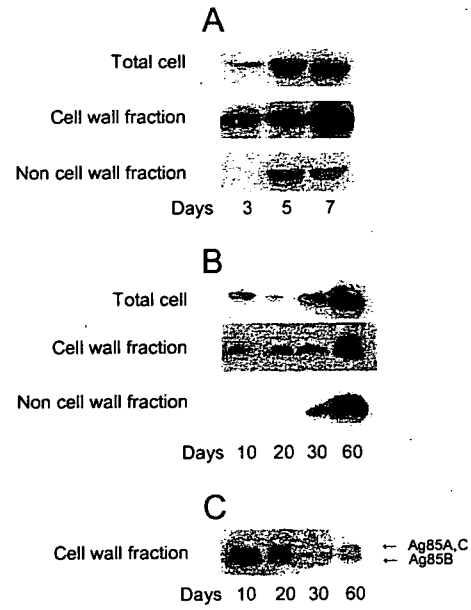


FIG. 5. Subcellular localization of MDP1 and Ag85 in the course of culture. (A to C) Total cellular protein (Total cell), the cell wall fraction (Cell wall fraction), and the residual material after isolation of the cell wall fraction (Non cell wall fraction) were obtained from *M. smegmatis* (A) and BCG (B and C) at each time point indicated. The samples were analyzed by Western blotting with anti-MDP1 MAb 3A (A and B) or anti-Ag85B Ig (C). The data are representative of three to five experiments.

phase (Fig. 4A). This phenotype was almost completely reversed by complementation. Next, we analyzed bacterial surface morphology by SEM. All strains produced similar smooth structures at exponential phase (Fig. 4B, C, and D, upper panels), and both the wild-type and complemented strains were normal with a smooth shape even in the stationary growth phase (Fig. 4B and D, middle panels). However, at the same time point, the MDP1 KO strain displayed a crenellated structure (Fig. 4C, middle and lower panels), implying that MDP1 influences cell envelope structure during stationary phase in *M. smegmatis*.

In order to analyze the effect of MDP1 disruption on glycolipid biosynthesis, we chased synthesis of mycolates and TMM by adding $^{14}$C-labeled acetic acid to cultures of the wild-type, MDP1 KO, and complemented strains. Incorporation of radioactivity into TMM and MAMEs was analyzed by using the BAS system after fractionation by high-performance TLC (Fig. 6). TLC analysis revealed two TMM spots (Fig. 6A). *M. smegmatis* produces three types of mycolic acids, alpha, alpha′, and epoxy mycolates. To determine which types of mycolates were present in each TMM spot, we determined the molecular mass of each spot by MALDI-TOF mass spectrometry. The major peak of the upper TMM spot showed a pseudomolecular ion $[M+Na]^+$ at *m/z* 1556, which was identified as alpha-$C_{83}$ TMM, while the major peak of the lower spot showed a pseudomolecular ion $[M+Na]^+$ at *m/z* 1306, which was presumed to be alpha′-$C_{65}$ TMM (data not shown) (14). Thus, high-performance TLC analysis can separate alpha TMM and alpha′ TMM.

The three strains synthesized similar amounts of TMM and MAMEs during exponential growth (day 3). However, during stationary phase (day 7), in both the wild-type and complemented strains mycolate synthesis was reduced strongly (day 7). By contrast, the MDP1-deficent strain continued production of both mycolates and TMM, synthesizing 9.9-fold-higher amounts of TMM (Fig. 6B), 4.3-fold-higher amounts of alpha-MAME (Fig. 6C), and 5.9-fold-higher amounts of alpha′-MAME (Fig. 6D) than the wild type. This phenotype was completely reversed by complementation, indicating that a lack of MDP1 impaired the down-regulation of biosynthesis of mycolates and TMM in stationary phase.

**Altering the growth rate with Ag85 complex proteins and MDP1.** Because cell wall biogenesis is a biological event involved in multiplication of bacteria, we next examined whether regulation of the transfer of mycolic acids by MDP1 influences mycobacterial growth. We assessed the effects of exogenously

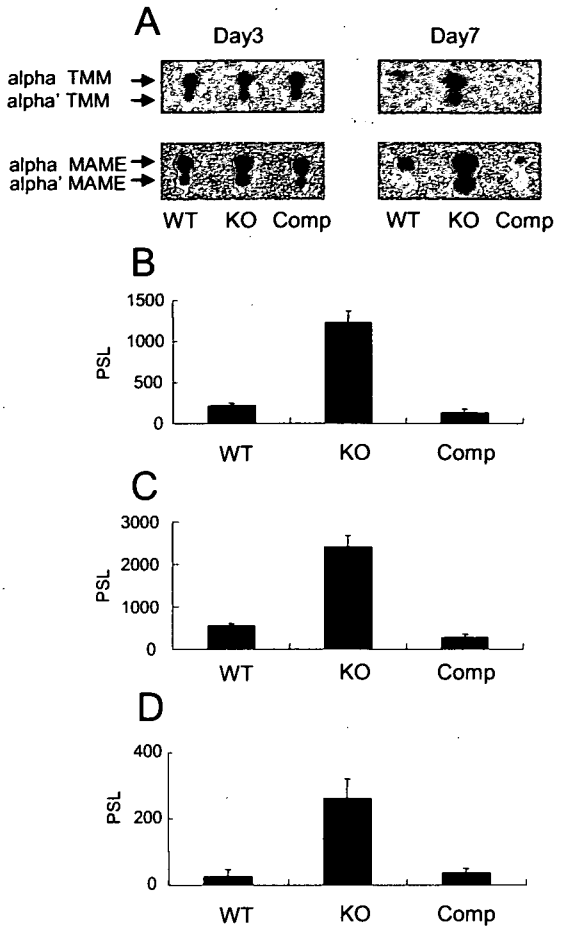


FIG. 6. MDP1 mediates cessation of glycolipid biosynthesis at the stationary growth phase. (A) Lipid synthesis was measured by addition of [$^{14}$C]acetate to a bacterial culture. The TMM and MAMEs from $10^8$ bacteria, including the *M. smegmatis* parent strain (wild type [WT]), the MDP1 KO strain (KO), and the complemented strain (Comp), were extracted on days 3 and 7 and fractionated by TLC. The radioactivities of synthesized TMM and MAMEs were visualized by using the BAS system. (B to D) Radioactivities of TMM (B), alpha-MAME (C), and alpha'-MAME (D) extracted from a bacterial culture after 7 days were quantified by using the BAS system software and compared for the wild-type, MDP1 KO, and complemented strains. The values are means ± standard deviations of five experiments. PSL, photostimulated luminescence.

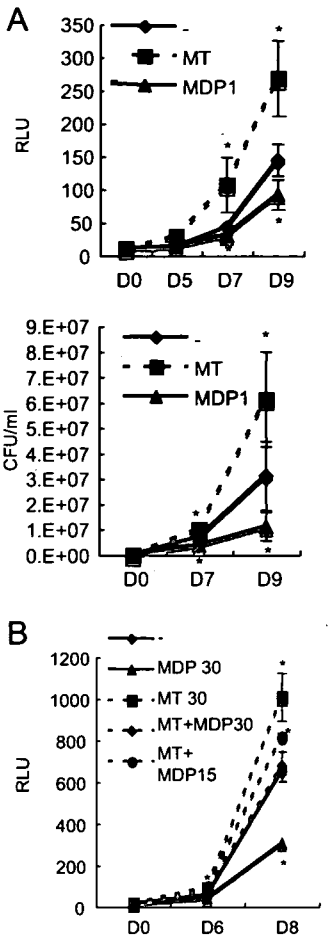


FIG. 7. Growth control of BCG by exogenously added MDP1 and Ag85 complex proteins. BCG-Luc was grown in RPMI 1640 containing 10% FBS in the presence of various doses of MDP1, Ag85 complex, or a mixture of MDP1 and the Ag85 complex in a 96-well tissue culture plate. The total culture volume was 200 μl in each well. Representative results of three independent experiments are presented. (A) Luciferase activity (expressed in relative luciferase units [RLU]) was measured on days 0, 5, 7, and 9 (upper graph), and CFU were counted on days 0, 7, and 9 (lower graph). MT, 30 μg/well Ag85 complex; MDP1, 30 μg/well MDP1; −, BCG-Luc alone. The values are the means ± standard deviations of three experiments. (B) Luciferase activity was measured on days 0, 6, and 8. MT 30, 30 μg/well Ag85 complex; MDP 30, 30 μg/well MDP1; MT+MDP30, 30 μg/well Ag85 complex plus 30 μg/well MDP1; MT+MDP15, 30 μg/well Ag85 complex plus 15 μg/well MDP1; −, BCG-Luc alone. The values are the means ± standard deviations of three experiments. Asterisk, $P < 0.05$ for a comparison with the no-protein control (as determined by analysis of variance).

added MDP1 in culture media on this growth. We generated a luciferase-producing BCG strain (BCG-Luc) to estimate the growth of BCG. The activity of luciferase paralleled the CFU assay results up to $10^4$ CFU/ml (data not shown). We cultured BCG-Luc in the presence or absence of MDP1 or Ag85 complex proteins. Both the luciferase-based assay and determination of the CFU showed that exogenously added MDP1 suppressed growth of BCG, while Ag85 complex proteins enhanced growth (Fig. 7A). We next added serial doses of MDP1 to the culture of BCG-Luc in the presence of Ag85 complex proteins. We found that a low dose of MDP1 further boosted Ag85 complex-induced growth enhancement, while a high dose of MDP1 suppressed Ag85-dependent growth enhancement (Fig. 7B).

## DISCUSSION

In this study, we analyzed the role of the mycobacterial histone-like protein MDP1 in the cell wall. We found that MDP1 plays an important role in tuning cell wall assembly by controlling transfer of mycolic acids to sugars by Ag85 complex proteins.

We showed that the cellular content of MDP1 was increased in advanced cultures of both *M. smegmatis* and BCG (Fig. 5). Previously, we and other groups showed that MDP1/HLP was accumulated in growth-retarded phases, including dormant bacilli of *M. tuberculosis* and *M. smegmatis* (19, 26, 37), by one-

dimensional SDS-PAGE analysis. MDP1 is resistant to analysis by two-dimensional gel electrophoresis because of its strong positive charge (pI of BCG, 12.4; pI of *M. tuberculosis*, 12.45). However, in spite of the extensive analysis of gene expression profiles with DNA microarrays, increased expression of *mdp1* mRNA in a stationary or anaerobic culture has not been observed (5, 18, 30, 35, 42). These results imply that accumulation of MDP1 in growth-retarded phases is largely due to posttranscriptional regulation. Pethe et al. found that lysines of the C-terminal region of laminin-binding protein/MDP1 are methylated by an unknown enzyme, which is present at a high level in the cell wall and confers resistance to proteolysis (31). Thus, posttranslational modification stabilizes MDP1 and might be involved in accumulation of MDP1 during a growth-retarded phase. In addition, methylation of basic charged amino acids may help the association of MDP1 with glycolipids by negating the charge of the protein. Posttranslational modifications, resembling eukaryotic histones, may control the cellular function and stability of MDP1. This issue should be clarified by additional study. By contrast, Ag85 complex proteins were localized in the cell wall during culture, but the contents gradually decreased with time in BCG (Fig. 5C) and *M. smegmatis* cultures (data not shown). It is likely that mycobacteria regulate transfer of mycolic acids by altering amounts of both MDP1 and Ag85 complex proteins in the cell wall.

The density of the MDP1 KO strain was lower in stationary phase (Fig. 4A). Lee et al. reported that an HLP/MDP1 KO strain displayed the same growth kinetics in Dubos Tween-albumin broth (19). We confirmed that an *M. smegmatis* MDP1 KO strain exhibited similar levels of cell density and survival at stationary phase when it was cultured in Dubos Tween-albumin broth (data not shown). Furthermore, alteration of the surface structure at stationary phase was not revealed by SEM analysis of the surface of the MDP1 KO strain cultured in Dubos Tween-albumin broth (data not shown).

TMM-derived mycolic acids are a major source of TDM and AG-linked mycolic acids (40, 41). Thus, the amount of TMM is an important factor for determining the level of cell wall assembly. The MDP1 KO strain exhibited continued synthesis of TMM other than MAMEs (Fig. 6) and TDM (data not shown) at the stationary growth phase. TMM could be synthesized from TMM and TDM by Ag85 complex proteins, once the substrates (TMM and TDM) were synthesized. However, the primary enzyme that catalyzes synthesis of TMM is still not known. Takayama et al. proposed that TMM could be synthesized in the cytoplasm (40). However, recently, Tropis et al. demonstrated that TMM is synthesized outside the plasma membrane but not inside the cytoplasm (41). The Ag85 complex is the most abundant protein secreted by *M. tuberculosis* (around 10 to 30%) and is a possible candidate enzyme for transfer of mycolic acids to trehalose to synthesize the TMM precursor (β-keto-acyl trehalose), as deduced from the structure of Ag85C (33). MDP1 suppresses Ag85 complex-dependent secondary synthesis of TMM (Fig. 2B), but we cannot eliminate the possibility that MDP1 also inhibits the primary synthesis of the TMM precursor whenever it is catalyzed by Ag85 complex proteins or undetermined enzymes.

An activity controlling glycolipid biosynthesis prompted us to examine whether exogenously added MDP1 and Ag85 complex proteins influence bacterial growth. We showed that growth of BCG could be altered by a combination of Ag85 complex proteins and MDP1 (Fig. 7). The MDP1 content in the total proteins of the cell wall increased at the stationary growth phase, while that of the Ag85 complex proteins decreased (Fig. 3). It can be speculated that a change in the ratio of MDP1 to Ag85 complex proteins in the cell wall is involved in determining the growth rate.

Inhibition of specific cellular metabolism causes cell death, as many antibiotics kill bacteria. However, expression of MDP1 does not kill mycobacteria; instead, it just suppresses growth (23). This is probably due to suppression of whole cellular metabolism, including macromolecular biosynthesis of DNA, RNA, and proteins (23) and cell wall assembly, as shown in this study. This can be caused by multiple interacting activities of different classes of macromolecules. The mechanism of such multiple binding activities should be resolved by structure-based studies, like those done on bacterial proteins such as SecB (32, 44). Recently, it has been reported that the histone-like protein HN-S mediates silencing of global gene expression in growth-retarded *Salmonella* (2, 8, 21). In humans, histone H2A seems to be involved in X chromosome inactivation (11). It can be speculated that use of a nonspecific DNA-binding protein to inactivate cellular metabolism is a general strategy. Here we found an alternative role of a histone-like protein in bacterial metabolism. Our data suggest that MDP1-medaited control of glycolipid biosynthesis is involved in the mechanism linking the growth state and cell wall biogenesis. Although we conducted experiments using nonpathogenic or slightly pathogenic mycobacteria, such as *M. smegmatis* and BCG, these bacteria share the basic structure of the cell wall with pathogenic mycobacteria, including *M. tuberculosis*. Taken together, our data provide significant information for understanding both the coordination of bacterial growth and virulence.

## ACKNOWLEDGMENTS

We are grateful to Keizou Oka, Department of Bioscience, INCS, Ehime University, for technical assistance with preparation of MAbs and to Thomas Dick, Novartis Institute for Tropical Diseases, for providing the *M. smegmatis* HLP/MDP1 KO strain. We also thank Todd P. Primm and Charles Scanga for editing the manuscript and Sara Matsumoto for heartfelt encouragement.

This work was supported by grants from the Ministry of Health, Labor and Welfare (Research on Emerging and Re-emerging Infectious Diseases, Health Sciences Research Grants), The Japan Health Sciences Foundation, the Ministry of Education, Culture, Sports, Science and Technology, and The United States-Japan Cooperative Medical Science Program against Tuberculosis and Leprosy.

We have no competing interests.

## REFERENCES

1. **Aoki, K., S. Matsumoto, Y. Hirayama, T. Wada, Y. Ozeki, M. Niki, P. Domenech, K. Umemori, S. Yamamoto, A. Mineda, M. Matsumoto, and K. Kobayashi.** 2004. Extracellular mycobacterial DNA-binding protein 1 participates in *Mycobacterium*-lung epithelial cell interaction through hyaluronic acid. J. Biol. Chem. **279**:39798–39806.
2. **Asakura, H., K. Kawamoto, T. Shirahata, and S. Makino.** 2004. Changes in *Salmonella enterica* serovar Oranienburg viability caused by NaCl-induced osmotic stress is related to DNA relaxation by the H-NS protein during host infection. Microb. Pathog. **36**:147–151.
3. **Barry, C. E., III, R. E. Lee, K. Mdluli, A. E. Sampson, B. G. Schroeder, R. A. Slayden, and Y. Yuan.** 1998. Mycolic acids: structure, biosynthesis and physiological functions. Prog. Lipid Res. **37**:143–179.
4. **Belisle, J. T., V. D. Vissa, T. Sievert, K. Takayama, P. J. Brennan, and G. S. Besra.** 1997. Role of the major antigen of *Mycobacterium tuberculosis* in cell wall biogenesis. Science **276**:1420–1422.

5. **Betts, J. C., P. T. Lukey, L. C. Robb, R. A. McAdam, and K. Duncan.** 2002. Evaluation of a nutrient starvation model of *Mycobacterium tuberculosis* persistence by gene and protein expression profiling. Mol. Microbiol. **43:** 717–731.
6. **Bloom, B. R.** 2002. Tuberculosis—the global view. N. Engl. J. Med. **346:** 1434–1435.
7. **Brennan, P. J., and H. Nikaido.** 1995. The envelope of mycobacteria. Annu. Rev. Biochem. **64:**29–63.
8. **Chen, C. C., H. Y. Wu, M. Y. Chou, C. H. Huang, and A. Majumder.** 2005. LeuO protein delimits the transcriptionally active and repressive domains on the bacterial chromosome. J. Biol. Chem. **280:**15111–15421.
9. **Cole, S. T., R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S. V. Gordon, K. Eiglmeier, S. Gas, C. E. Barry III, F. Tekaia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, A. Krogh, J. McLean, S. Moule, L. Murphy, K. Oliver, J. Osborne, M. A. Quail, M. A. Rajandream, J. Rogers, S. Rutter, K. Seeger, J. Skelton, R. Squares, S. Squares, J. E. Sulston, K. Taylor, S. Whitehead, and B. G. Barrell.** 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature **393:**537–544.
10. **Cole, S. T., K. Eiglmeier, J. Parkhill, K. D. James, N. R. Thomson, P. R. Wheeler, N. Honore, T. Garnier, C. Churcher, D. Harris, K. Mungall, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. M. Davies, K. Devlin, S. Duthoy, T. Feltwell, A. Fraser, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, C. Lacroix, J. Maclean, S. Moule, L. Murphy, K. Oliver, M. A. Quail, M. A. Rajandream, K. M. Rutherford, S. Rutter, K. Seeger, S. Simon, M. Simmonds, J. Skelton, R. Squares, S. Squares, K. Stevens, K. Taylor, S. Whitehead, J. R. Woodward, and B. G. Barrell.** 2001. Massive gene decay in the leprosy bacillus. Nature **409:**1007–1011.
11. **Costanzi, C., and J. R. Pehrson.** 1998. Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. Nature **393:**599–601.
12. **De Voss, J. J., K. Rutter, B. G. Schroeder, H. Su, Y. Zhu, and C. E. Barry III.** 2000. The salicylate-derived mycobactin siderophores of *Mycobacterium tuberculosis* are essential for growth in macrophages. Proc. Natl. Acad. Sci. USA **97:**1252–1257.
13. **Fleischmann, R. D., D. Alland, J. A. Eisen, L. Carpenter, O. White, J. Peterson, R. DeBoy, R. Dodson, M. Gwinn, D. Haft, E. Hickey, J. F. Kolonay, W. C. Nelson, L. A. Umayam, M. Ermolaeva, S. L. Salzberg, A. Delcher, T. Utterback, J. Weidman, H. Khouri, J. Gill, A. Mikula, W. Bishai, W. R. Jacobs, Jr., J. C. Venter, and C. M. Fraser.** 2002. Whole-genome comparison of *Mycobacterium tuberculosis* clinical and laboratory strains. J. Bacteriol. **184:**5479–5490.
14. **Fujita, Y., T. Naka, T. Doi, and I. Yano.** 2005. Direct molecular mass determination of trehalose monomycolate from 11 species of mycobacteria by MALDI-TOF mass spectrometry. Microbiology **151:**1443–1452.
15. **Harlow, E., and D. Lane.** 1988. Antibodies: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
16. **Jackson, M., C. Raynaud, M. A. Laneelle, C. Guilhot, C. Laurent-Winter, D. Ensergueix, B. Gicquel, and M. Daffe.** 1999. Inactivation of the antigen 85C gene profoundly affects the mycolate content and alters the permeability of the *Mycobacterium tuberculosis* cell envelope. Mol. Microbiol. **31:**1573–1587.
17. **Jacobs, W. R., Jr., M. Tuckman, and B. R. Bloom.** 1987. Introduction of foreign DNA into mycobacteria using a shuttle phasmid. Nature **327:**532–535.
18. **Karakousis, P. C., T. Yoshimatsu, G. Lamichhane, S. C. Woolwine, E. L. Nuermberger, J. Grosset, and W. R. Bishai.** 2004. Dormancy phenotype displayed by extracellular *Mycobacterium tuberculosis* within artificial granulomas in mice. J Exp. Med. **200:**647–657.
19. **Lee, B. H., B. Murugasu-Oei, and T. Dick.** 1998. Upregulation of a histone-like protein in dormant *Mycobacterium smegmatis*. Mol. Gen. Genet. **260:** 475–479.
20. **Liu, J., C. E. Barry III, G. S. Besra, and H. Nikaido.** 1996. Mycolic acid structure determines the fluidity of the mycobacterial cell wall. J. Biol. Chem. **271:**29545–29551.
21. **Lucchini, S., G. Rowley, M. D. Goldberg, D. Hurd, M. Harrison, and J. C. Hinton.** 2006. H-NS mediates the silencing of laterally acquired genes in bacteria. PLoS Pathog. **2:**e81.
22. **Manabe, Y. C., and W. R. Bishai.** 2000. Latent Mycobacterium tuberculosis—persistence, patience, and winning by waiting. Nat. Med. **6:**1327–1329.
23. **Matsumoto, S., M. Furugen, H. Yukitake, and T. Yamada.** 2000. The gene encoding mycobacterial DNA-binding protein 1 (MDP1) transformed rapidly growing bacteria to slowly growing bacteria. FEMS Microbiol. Lett. **182:**297–301.
24. **Matsumoto, S., M. Matsumoto, K. Umemori, Y. Ozeki, M. Furugen, T. Tatsuo, Y. Hirayama, S. Yamamoto, T. Yamada, and K. Kobayashi.** 2005. DNA augments antigenicity of mycobacterial DNA-binding protein 1 and confers protection against *Mycobacterium tuberculosis* infection in mice. J. Immunol. **175:**441–449.
25. **Matsumoto, S., M. Tamaki, H. Yukitake, T. Matsuo, M. Naito, H. Teraoka, and T. Yamada.** 1996. A stable *Escherichia coli*-mycobacteria shuttle vector 'pSO246' in *Mycobacterium bovis* BCG. FEMS Microbiol. Lett. **135:**237–243.
26. **Matsumoto, S., H. Yukitake, M. Furugen, T. Matsuo, T. Mineta, and T. Yamada.** 1999. Identification of a novel DNA-binding protein from *Mycobacterium bovis* bacillus Calmette-Guerin. Microbiol. Immunol. **43:**1027–1036.
27. **Mikusova, K., M. Mikus, G. S. Besra, I. Hancock, and P. J. Brennan.** 1996. Biosynthesis of the linkage region of the mycobacterial cell wall. J. Biol. Chem. **271:**7820–7828.
28. **Naito, M., N. Ohara, S. Matsumoto, and T. Yamada.** 1998. Immunological characterization of alpha antigen of *Mycobacterium kansasii*: B-cell epitope mapping. Scand. J. Immunol. **48:**73–78.
29. **Naito, M., N. Ohara, S. Matsumoto, and T. Yamada.** 1998. The novel fibronectin-binding motif and key residues of mycobacteria. J. Biol. Chem. **273:**2905–2909.
30. **Park, H. D., K. M. Guinn, M. I. Harrell, R. Liao, M. I. Voskuil, M. Tompa, G. K. Schoolnik, and D. R. Sherman.** 2003. Rv3133c/dosR is a transcription factor that mediates the hypoxic response of *Mycobacterium tuberculosis*. Mol. Microbiol. **48:**833–843.
31. **Pethe, K., P. Bifani, H. Drobecq, C. Sergheraert, A. S. Debrie, C. Locht, and F. D. Menozzi.** 2002. Mycobacterial heparin-binding hemagglutinin and laminin-binding protein share antigenic methyllysines that confer resistance to proteolysis. Proc. Natl. Acad. Sci. USA **99:**10759–10764.
32. **Randall, L. L., and S. J. Hardy.** 1995. High selectivity with low specificity: how SecB has solved the paradox of chaperone binding. Trends Biochem. Sci. **20:**65–69.
33. **Ronning, D. R., T. Klabunde, G. S. Besra, V. D. Vissa, J. T. Belisle, and J. C. Sacchettini.** 2000. Crystal structure of the secreted form of antigen 85C reveals potential targets for mycobacterial drugs and vaccines. Nat. Struct. Biol. **7:**141–146.
34. **Sassetti, C. M., D. H. Boyd, and E. J. Rubin.** 2003. Genes required for mycobacterial growth defined by high density mutagenesis. Mol. Microbiol. **48:**77–84.
35. **Schnappinger, D., S. Ehrt, M. I. Voskuil, Y. Liu, J. A. Mangan, I. M. Monahan, G. Dolganov, B. Efron, P. D. Butcher, C. Nathan, and G. K. Schoolnik.** 2003. Transcriptional adaptation of *Mycobacterium tuberculosis* within macrophages: insights into the phagosomal environment. J. Exp. Med. **198:**693–704.
36. **Shimoji, Y., V. Ng, K. Matsumura, V. A. Fischetti, and A. Rambukkana.** 1999. A 21-kDa surface protein of *Mycobacterium leprae* binds peripheral nerve laminin-2 and mediates Schwann cell invasion. Proc. Natl. Acad. Sci. USA **96:**9857–9862.
37. **Shires, K., and L. Steyn.** 2001. The cold-shock stress response in *Mycobacterium smegmatis* induces the expression of a histone-like protein. Mol. Microbiol. **39:**994–1009.
38. **Soares de Lima, C., L. Zulianello, M. A. Marques, H. Kim, M. I. Portugal, S. L. Antunes, F. D. Menozzi, T. H. Ottenhoff, P. J. Brennan, and M. C. Pessolani.** 2005. Mapping the laminin-binding and adhesive domain of the cell surface-associated Hlp/LBP protein from *Mycobacterium leprae*. Microbes Infect. **7:**1097–1109.
39. **Stover, C. K., V. F. de la Cruz, T. R. Fuerst, J. E. Burlein, L. A. Benson, L. T. Bennett, G. P. Bansal, J. F. Young, M. H. Lee, G. F. Hatfull, et al.** 1991. New use of BCG for recombinant vaccines. Nature **351:**456–460.
40. **Takayama, K., C. Wang, and G. S. Besra.** 2005. Pathway to synthesis and processing of mycolic acids in *Mycobacterium tuberculosis*. Clin. Microbiol. Rev. **18:**81–101.
41. **Tropis, M., X. Meniche, A. Wolf, H. Gebhardt, S. Strelkov, M. Chami, D. Schomburg, R. Kramer, S. Morbach, and M. Daffe.** 2005. The crucial role of trehalose and structurally related oligosaccharides in the biosynthesis and transfer of mycolic acids in Corynebacterineae. J. Biol. Chem. **280:**26573–26585.
42. **Voskuil, M. I., D. Schnappinger, K. C. Visconti, M. I. Harrell, G. M. Dolganov, D. R. Sherman, and G. K. Schoolnik.** 2003. Inhibition of respiration by nitric oxide induces a *Mycobacterium tuberculosis* dormancy program. J. Exp. Med. **198:**705–713.
43. **Wayne, L. G., and C. D. Sohaskey.** 2001. Nonreplicating persistence of *Mycobacterium tuberculosis*. Annu. Rev. Microbiol. **55:**139–163.
44. **Zhou, J., and Z. Xu.** 2005. The structural view of bacterial translocation-specific chaperone SecB: implications for function. Mol. Microbiol. **58:**349–357.

Available online at www.sciencedirect.com
ScienceDirect
Molecular & Biochemical Parasitology 158 (2008) 11–21
MOLECULAR & BIOCHEMICAL PARASITOLOGY

# Diversity and evolution of the *rhoph1/clag* multigene family of *Plasmodium falciparum*☆

Hideyuki Iriko[a,b,1], Osamu Kaneko[a,*], Hitoshi Otsuki[a], Takafumi Tsuboi[b,c], Xin-zhuan Su[d], Kazuyuki Tanabe[e], Motomi Torii[a]

[a] *Department of Molecular Parasitology, Ehime University Graduate School of Medicine, Shitsukawa, Toon, Ehime 791-0295, Japan*
[b] *Venture Business Laboratory, Ehime University, Matsuyama, Ehime 790-8577, Japan*
[c] *Cell-free Science and Technology Research Center, Ehime University, Matsuyama, Ehime 790-8577, Japan*
[d] *Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892, USA*
[e] *Laboratory of Malariology, International Research Center for Infectious Diseases, Research Institute for Microbial Diseases, Osaka University, Suita, Osaka 565-0871, Japan*

Received 4 June 2007; received in revised form 5 November 2007; accepted 6 November 2007
Available online 17 November 2007

**Abstract**

A complex of high-molecular-mass proteins (*Pf*RhopH) of the human malaria parasite *Plasmodium falciparum* induces host protective immunity and therefore is a candidate for vaccine development. Understanding the level of polymorphism and the evolutionary processes is important for advancements in both vaccine design and knowledge of the evolution of cell invasion in this parasite. In the present study, we sequenced the entire open reading frames of seven genes encoding the proteins of the *Pf*RhopH complex (*rhoph2*, *rhoph3*, and five *rhoph1/clag* gene paralogs). We found that four *rhoph1/clag* genes (*clag2*, *3.1*, *3.2*, and *8*) were highly polymorphic. Amino acid substitutions and indels are predominantly clustered around amino acid positions 1000–1200 of these four *rhoph1/clag* genes. An excess of nonsynonymous substitutions over synonymous substitutions was detected for *clag8* and *9*, indicating positive selection. The McDonald–Kreitman test with a *Plasmodium reichenowi* orthologous sequence also supports positive selection on *clag8*. Based on the ratio of interspecific genetic distance to intraspecific distance, the time to the most recent common ancestor of the *clag2* and *8* polymorphisms was estimated to be 1.89 and 0.87 million years ago, respectively, assuming divergence of *P. falciparum* and *P. reichenowi* 6 million years ago. In addition to a copy number polymorphism, gene conversion events were detected for the *rhoph1/clag* genes on chromosome 3, which likely play a role in increasing the diversity of each locus. Our results indicate that a high diversity of the *Pf*RhopH1/Clag multigene family is maintained by diversifying selection forces over a considerably long period.


*Keywords:* Gene conversion; Malaria; Polymorphism; Rhoptry; Selection

## 1. Introduction

Malaria infects more than 300 million people and kills 1–2 million each year. Efforts have been made to develop effective malaria vaccines, but none is available so far. Malaria is caused by the obligate intracellular protozoan *Plasmodium* parasites; and entry into erythrocytes is prerequisite for the growth in the mammalian host. After contact with the erythrocyte surface, parasite discharge the content of the microorganelles called the micronemes to establish a tight junction with the erythrocyte surface molecules; parasite then invaginates into a nascent parasitophorous vacuole (PV) [1,2]. During formation of the PV, the parasite discharges the contents of another pair of microorganelles, the rhoptries [3]. The molecules located within these

*Abbreviations:* aa, amino acid; chr, chromosome; *clag*, cytoadherence-linked asexual gene(s); CI, confidence interval; mya, million years ago; nt, nucleotide; ORF, open reading frame(s); PCR, polymerase chain reaction; PV, parasitophorous vacuole; PVM, PV membrane; S.E., standard error; TMRCA, time to the most recent common ancestor; UTR, untranslated region(s).

☆ *Note:* Sequence data from this article have been deposited with the GenBank™/EMBL/DDBJ databases under accession numbers AB250801–AB250912.

* Corresponding author. Present address: Department of Protozoology, Institute of Tropical Medicine, Nagasaki University, 1-12-4 Sakamoto, Nagasaki 852-8523, Japan. Tel.: +81 95 819 7838; fax: +81 95 819 7805.

[1] Current address: Division of Medical Zoology, Department of Microbiology and Immunology, Faculty of Medicine, Tottori University, Yonago, Tottori 683-8503, Japan.

0166-6851/$ – see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.molbiopara.2007.11.004

organelles play a key role in erythrocyte invasion and have been studied as vaccine targets, with the aim to induce antibodies to block invasion. One erythrocyte-binding molecule in the rhoptry is a complex of high-molecular-mass proteins called the RhopH complex [4,5]. The RhopH complex is distributed throughout the erythrocyte and PV membrane (PVM) and has been detected in ring-stage parasites [6], suggesting an important role during PV establishment. The importance of the complex has further been emphasized by the failure of attempts to disrupt the *pfrhoph3* gene locus, suggesting its necessity for parasite survival [7].

The RhopH complex comprises three distinct components: RhopH1, RhopH2, and RhopH3 [8–12]. The genes encoding RhopH1 are members of the *rhoph1/clag* gene family, which was originally defined by the cytoadherence-linked asexual gene (*clag*) on chromosome 9 in *Plasmodium falciparum* (*clag9*) and consists of at least three members; *clag2*, *3.1*, and *9* [13–15]. Although not yet determined experimentally, molecules encoded by *clag3.2* and *8* are likely parts of the RhopH complex as judged by their similarity in amino acid sequence and transcription pattern with other members [15]. Because only one RhopH1/Clag participates to form a single RhopH complex [15,16], five types of *Pf*RhopH complex are expected to exist, each of which contains one *rhoph1/clag* gene product. In this report we employ 'RhopH1/Clag' (protein) and '*rhoph1/clag*' (gene) as the family name, and 'Clag' (protein) and '*clag*' (gene) for each member.

Erythrocyte-binding proteins discharged from *P. falciparum* merozoites are considered to be targets of host immune responses. Strong diversifying selections on microneme proteins have been detected (e.g., AMA-1 and EBA-175), suggesting that polymorphism of these proteins has been maintained to evade host immunity in parasite populations [17,18]. Antibodies against the *Pf*RhopH complex partially inhibit the growth of *P. falciparum* in vitro and in vivo, consistent with its potential as a vaccine target [19–21]. Although the RhopH complex has been shown to induce host protective immunity and is likely to be under host immune pressure, the genetic diversity and immunologic characteristics of this complex are not fully understood. Here, we analyzed sequence polymorphism in five *rhoph1/clag* members, *rhoph2*, and *rhoph3* and show that some of the *rhoph* genes are under positive/diversifying selection. In addition, we assessed a population genetic mechanism that might drive the evolution of the *rhoph1/clag* multigene family.

## 2. Materials and methods

### 2.1. Malaria parasites

All cloned lines of *P. falciparum* were maintained in vitro, essentially as described previously [22]. The parasite lines examined originated from Southeast Asia (Dd2, FVO, Camp, T9/96, T9/102, K1, and Thai838), Papua New Guinea (MAD20), Central and South America (HB3, 7G8, DIV17, DIV29, DIV30, PC49, PC54, Santa Lucia, and Haiti), and Africa (RO33, 123/5, 128/4, SL/D6, LF4/1, 102/1, M2, M5, Fab9, 713, P13, and KMWII) and have been previously described [23–25]. Their geographic origins have also been previously described [26].

### 2.2. DNA and RNA isolation

Genomic DNA was obtained as described previously [24]. Total RNA was isolated from schizont stage-enriched HB3 and Dd2 parasite lines using the RNeasy mini kit (Qiagen, Valencia, CA). Complementary DNA was synthesized using random hexamers and an Omniscript reverse transcription kit (Qiagen) after DNase treatment.

### 2.3. Polymerase chain reaction (PCR) amplification and sequencing

Nucleotide sequences corresponding to open reading frame (ORF) were determined for five *pfrhoph1/clag* genes, *rhoph2*, and *rhoph3* in four parasite lines: Dd2, HB3, 7G8, and FVO. DNA fragments were PCR amplified with KOD-Plus DNA polymerase (Toyobo, Japan) using a panel of oligonucleotides specific for the genes (Supplemental Table 1) and sequenced directly using an ABI PRISM® 310 genetic analyzer (Applied Biosystems, Foster City, CA) or sequenced after cloning into pGEM-T Easy® plasmid (multiple plasmid clones sequenced for each DNA fragment; Promega, Madison, WI). To PCR amplify DNA fragments including the entire ORF of *clag3.1* or *3.2*, LA *Taq* DNA polymerase (TaKaRa, Japan) was used with oligonucleotide primers 3.1F (5′-TGTGCAATATATCAAAGTGTACATGC-3′) and 3.1R (5′-TAGAAAATATTAGAATTGCTATTATGTAC-3′) or 3.2F (5′-AATAGTTGAGTACGCACTAATATGTC-3′) and 3.2R (5′-ACACAAATTCTTAATAATTATATAAAACC-3′), respectively. A highly polymorphic region identified in *clag2*, *3.1*, *3.2*, and *8* in this study was further analyzed by increasing the number of parasite lines ($n = 25$) from different geographic areas.

### 2.4. Plasmodium reichenowi sequences

A TBLASTN search was performed against the *P. reichenowi* preliminary genome shotgun database (Dennis strain; Sanger Centre, UK) using Clag2, 3.1, 3.2, and 8 amino acid sequences as queries. For *prclag2* and *prclag8*, sequences were assembled using SeqMan II accompanied with Lasergene software (DNASTAR Inc., Madison, WI) with manual corrections. Regions covered by at least two independent reads and showing identical sequences were selected and used for analysis (Supplemental Figs. S1 and S2). The generated sequences were 3273 bp long for *prclag2*, corresponding to nucleotide (nt) positions 193–870, 1021–1902, 2458–3432, and 3448–4185 of *pfclag2* (3D7), and 2175 bp long for *prclag8*, corresponding to nt positions 1459–4173 of *pfclag8* (3D7). For Clag3 orthologs in the *P. reichenowi* genome, only sequences possessing homology with the 5′ untranslated region (UTR) (reich908g11.plk) or 3′ UTR (reich1194c08.plk and reich289f06.qlk) were used.

### 2.5. Sequence alignment and analysis

The entire ORFs for the 7 *Pf*RhopH complex-related genes (5 *rhoph1/clag* genes, *rhoph2*, and *rhoph3*) in four culture-adapted

*P. falciparum* lines—Dd2 (Southeast Asia), 7G8 (Brazil), HB3 (Honduras), and FVO (Vietnam)—were aligned with those retrieved from a genome database (3D7 line, presumably African in origin) using a CLUSTAL W program [27] with manual corrections; nucleotide diversity ($\pi$) and its standard error (S.E.) were computed with the Jukes and Cantor method using MEGA 3.1 software [28] after excluding insertions/deletions (indels) and highly polymorphic regions. The mean numbers of synonymous substitutions per synonymous site ($d_S$) and nonsynonymous substitutions per nonsynonymous site ($d_N$) and their standard errors were computed using the Nei and Gojobori method [29] with the Jukes and Cantor correction, implemented in MEGA 3.1. The statistical difference between $d_S$ and $d_N$ was tested using a one-tailed Z-test with 500 bootstrap pseudosamples using MEGA 3.1. A value of $d_N$ significantly higher or lower than $d_S$ at the 95% confidence level was taken as evidence for positive or purifying selection, respectively. The $d_N$:$d_S$ ratio was evaluated using a sliding window method (50 bases with a step size of 10 bases) in DnaSP 4.0 [30]. Positive selection was also evaluated using the McDonald–Kreitman test [31]. Before estimating the time to the most recent common ancestor (TMRCA) for *P. falciparum clag2* and *8* polymorphism, the evolutionary rate constancy of *clag2* and *8* between *P. falciparum* and *P. reichenowi* was validated using a *Plasmodium yoelii* ortholog *Py*RhopH1A (accession number AB060734) as an outgroup using Tajima's relative rate test [32] implemented in MEGA 3.1. Mean and 95% confidence intervals (CI) for estimated TMRCA were computed based on the model assuming the distribution of the distance and the substitution rate were Gamma-distributed [33]. Gene conversion was evaluated for each exon using an algorithm by Betrán et al. [34] implemented in DnaSP 4.0.

Unrooted dendrograms of the *pfrhoph1/clag* members were constructed using the neighbor-joining and maximum parsimony methods in MEGA 3.1, and Tajima's relative rate test was used to evaluate the evolutionary rate among members. Indels and highly polymorphic regions could not be satisfactorily aligned and were therefore excluded from the analysis. The sequences (3D7 parasite line) used to construct trees and the evolutionary rate were as follows: nt positions 154–312, 331–573, 727–1122, 1207–1266, 1324–1560, 1609–2988, 3004–3288, and 3382–3924 for *clag3.2*; nt positions 160–318, 337–579, 733–1128, 1213–1272, 1330–1566, 1615–2994, 3010–3294, and 3388–3930 for *clag3.1*; nt positions 223–381, 400–642, 799–1194, 1279–1338, 1390–1626, 1696–3075, 3091–3375, and 3553–4095 for *clag2*; nt positions 130–288, 307–549, 706–1101, 1186–1245, 1300–1536, 1606–2985, 3001–3285, and 3415–3957 for *clag8*; and nt positions 82–240, 265–507, 652–1047, 1132–1191, 1276–1512, 1582–2961, 2977–3261, and 3394–3936 for *clag9*.

## 3. Results

### 3.1. Polymorphism of the PfRhopH complex-related genes

All seven *Pf*RhopH complex-related genes showed greater nucleotide diversity levels than the average (+2 S.E.) of 204 ORFs on *P. falciparum* chromosome (chr) 3 [35] (Table 1). Among the seven genes, *clag2*, *3.1*, *3.2*, and *8* are highly polymorphic with nucleotide diversity ($\pi = 0.0053$–$0.0164$) comparable to malaria vaccine candidate antigen protein genes such as *eba-175* ($\pi = 0.0030$) and *ama-1* ($\pi = 0.0166$) [17,18]. The observed nucleotide diversity levels of *clag2*, *3.1*, and *3.2* should be taken as minimum estimates, because indels and highly polymorphic regions were excluded from this analysis to obtain reliable alignments. The highly polymorphic nature of four *rhoph1/clag* genes at the nucleotide level extends to the amino acid level, which is represented by high $d_N$ values (Table 1). Thus, the genes encoding RhopH1/Clag are more polymorphic than RhopH2 and RhopH3.

Among the four RhopH1/Clag showing high polymorphism (Clag2, 3.1, 3.2, and 8), the majority of polymorphic sites are clustered in a region at amino acid (aa) positions 1000–1200

Table 1
Nucleotide diversity of the *Pf*RhopH complex genes[a]

| Gene | *n* | Indel | Sites | $\pi$ | $\pi$ (S.E.) | $d_N$ | $d_N$ (S.E.) | $d_S$ | $d_S$ (S.E.) | $d_N/d_S$ | $P^b$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *clag2*[c] | 5 | (+) | 4,317 | 0.0053 | (0.0008) | 0.0032 | (0.0007) | 0.0133 | (0.0028) | 0.24 | (0.0003) |
| *clag3.1*[c] | 5 | (+) | 4,140 | 0.0164 | (0.0015) | 0.0062 | (0.0011) | 0.0582 | (0.0058) | 0.11 | ($<10^{-10}$) |
| *clag3.2*[c] | 5 | (+) | 4,134 | 0.0138 | (0.0011) | 0.0063 | (0.0011) | 0.0445 | (0.0050) | 0.14 | ($<10^{-10}$) |
| *clag8* | 5 | (−) | 4,182 | 0.0066 | (0.0007) | 0.0065 | (0.0011) | 0.0069 | (0.0020) | 0.94 | ns |
| *clag9* | 5 | (−) | 4,020 | 0.0009 | (0.0003) | 0.0011) | (0.0004) | 0.0000 | (0.0000) | ∞ | 0.002 |
| *rhoph2* | 5 | (−) | 4,134 | 0.0009 | (0.0003) | 0.0010 | (0.0004) | 0.0005 | (0.0005) | 2.00 | ns |
| *rhoph3* | 5 | (−) | 2,691 | 0.0013 | (0.0004) | 0.0012 | (0.0005) | 0.0015 | (0.0010) | 0.80 | ns |
| *clag2*[d] | 24 | (+) | 522 | 0.0131 | (0.0032) | 0.0114 | (0.0042) | 0.0192 | (0.0074) | 0.60 | ns |
| *clag8*[d] | 26 | (−) | 585 | 0.0267 | (0.0042) | 0.0305 | (0.0060) | 0.0132 | (0.0061) | 2.31 | 0.020 |
| *Chr 3*[e] | 5 | | 202,069 | 0.00044 | (0.00006) | 0.00039 | (0.0060) | 0.00068 | (0.00010) | 0.57 | |

[a] *n*, Number of sequences sampled; sites, sites analyzed excluding noncoding sequences and alignment gaps; $\pi$, pairwise nucleotide diversity; $d_N$, number of nonsynonymous substitutions over numbers of nonsynonymous sites; $d_S$, number of synonymous substitutions over numbers of synonymous sites; S.E., standard error computed using the Nei-Gojobori method with Jukes-Cantor correction. S.E. was estimated using the bootstrap method with 500 replication.

[b] *P*-value indicates that $d_N$ is significantly greater than $d_S$. Those shown in parenthesis indicate that $d_S$ are significantly greater than $d_N$. The statistical difference between $d_S$ and $d_N$ was tested using an one-tail Z-test with 500 bootstrap pseudosamples implemented in MEGA 3.1. ns indicate not significant ($P > 0.05$).

[c] For optimal sequence alignment, nt 3433–3435 was excluded from *clag2*, nt 3337–3447 from *clag3.1*, and nt 88–99 and 3343–3444 from *clag3.2* for the analysis. Nucleotide numbering are after the 3D7 line sequences.

[d] nt 3022–3606 of *clag8* and nt 3106–3420 and 3436–3642 of *clag2* were used.

[e] Data from 204 ORF on *P. falciparum* chr 3 using five parasite lines [35].

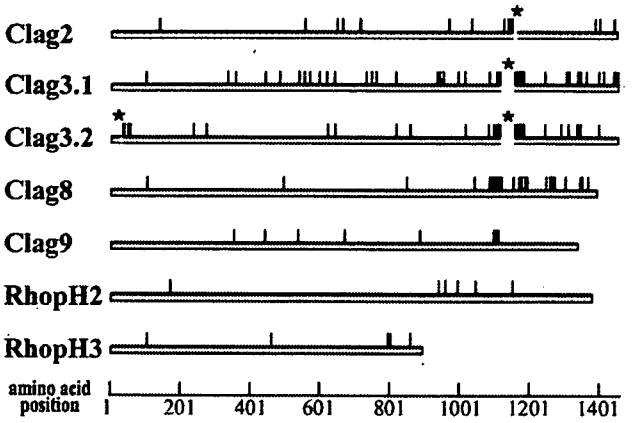


Fig. 1. Locations of amino acid polymorphism of seven components of the *Pf*RhopH complex among five *P. falciparum* parasite lines (3D7, HB3, Dd2, FVO, and 7G8). Indels are shown as gaps with asterisks (aa position 1145 for Clag2, aa positions 1113–1149 for Clag3.1, and aa positions 30–33 and 1115–1148 for Clag3.2). Numbers are those of 3D7 line sequences.

(Fig. 1). In addition, numerous polymorphic sites in this region have more than one amino acid substitution, whereas most polymorphisms in the other regions are dimorphic (at both nucleotide and amino acid levels). Most indels are also located in this region (Fig. 1, asterisks). Thus, the region at aa positions 1000–1200 of RhopH1/Clag is the most highly polymorphic region of the *Pf*RhopH complex.

### 3.2. Gene conversion between clag3.1 and 3.2

Of interest, *clag3.1* and *3.2* share some polymorphic sites. Because *clag3.1* and *3.2* have 96.7% nucleotide identity (3D7 parasite line) and are located on chr 3 and separated by only 10 kb harboring one putative ORF (PFC0115c) (Fig. 2A), we assessed gene conversion between these two loci. Using an algorithm by Betrán et al. [34], we identified multiple gene conversion tracts located at nt positions 1314–1353, 1447–1452, 1612–1659, 1702–1785, 1852–1983, and 2148–2208 in 3D7 *clag3.1*; nt positions 3824–4240 in HB3 *clag3.1*; nt positions 189–247 in 7G8 *clag3.1*; nt positions 813–817 and 3821–4182 in 3D7 *clag3.2*; nt positions 88–151 in HB3 *clag3.2*; and nt positions 3320–3755 in 7G8 *clag3.2* (Fig. 3). The detected conversion tracts had less than 5% informative nucleotides showing a mosaic origin, indicating that the probability of these tracts being involved in a recombination event more than once is negligible [34]. No gene conversion was detected between the other *rhoph1/clag* genes.

Because gene conversion potentially accelerates nucleotide diversity, we evaluated the evolutionary rates of *clag3.1* and *3.2*. Results showed that *clag2*, *3.1*, and *3.2* form a single clade and *clag8* another (Fig. 4); thus we performed Tajima's relative rate test using *clag8* as an outgroup and found that the evolutionary rates between *clag3.1* and *2* and between *clag3.2* and *2* were significantly different for all combinations of the sequences from five parasite lines. Because *clag3.1* and *3.2* were more diverse than *clag2*, *clag3.1* and *3.2* appear to have evolved more rapidly than *clag2*.

### 3.3. Amino acid polymorphism of the region around aa positions 1000–1200 of Clag2, 3.1, 3.2, and 8

Because extensive polymorphisms were observed around aa positions 1000–1200 in Clag2, 3.1, 3.2, and 8, we further analyzed polymorphism in this region with additional sequences from parasite lines originating worldwide. Alignment of Clag2 sequences showed multiple amino acid substitutions per site at multiple sites, e.g., five amino acids at aa position 1139 (K, R, S, G, and I). Indels were also observed (Supplemental Fig. S3). Clag8 has even higher levels of amino acid substitutions at between 1077 and 1136; five different amino acids (I, S, R, G, and N) at 1100, seven at 1101 (D, S, T, E, N, I, and K), six at 1104 (S, N, I, K, R, and T), and five at 1105 (G, D, T, S, and N) (Supplemental Fig. S4). Clag3.1 and 3.2 are also highly polymorphic (Fig. 5), which will be discussed later.

### 3.4. Copy number polymorphism of rhoph1/clag genes on chr 3

Notably, when PCR amplification was performed to obtain DNA fragments of the entire ORFs of *clag3.1* or *3.2*, 17 parasite

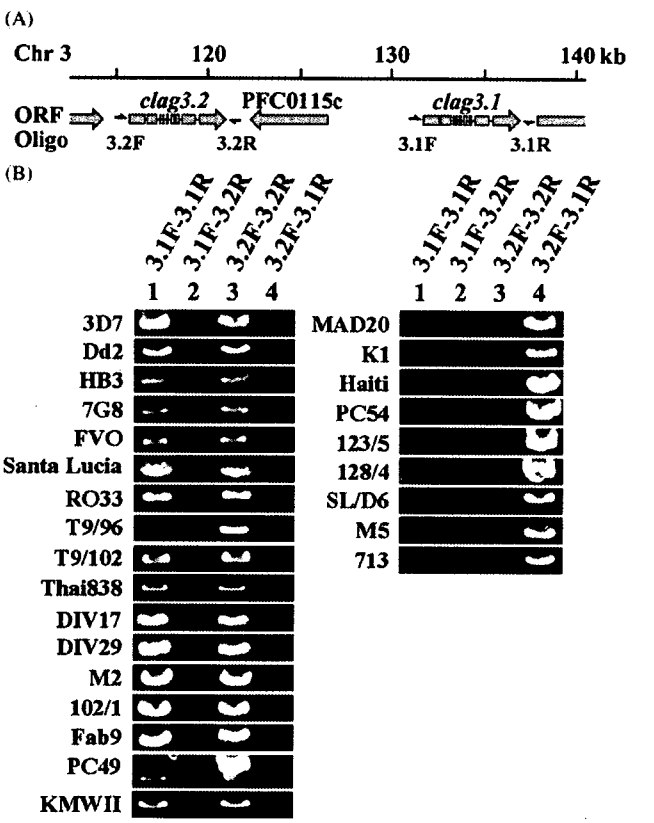


Fig. 2. Copy number polymorphism of *rhoph1/clag* genes on chr 3. (A) Genome organization around *clag3.2* and *3.1* gene loci on chr 3. The locations of the oligonucleotide primers are indicated. Oligonucleotide 3.2F and 3.1F were designed on the 5′ UTR of *clag3.2* and *3.1*, respectively. Oligonucleotide 3.2R and 3.1R were designed on 3′ UTR of *clag3.2* and *3.1*, respectively. (B) PCR-amplified DNA fragments of 26 *P. falciparum* lines with different combinations of oligonucleotides.

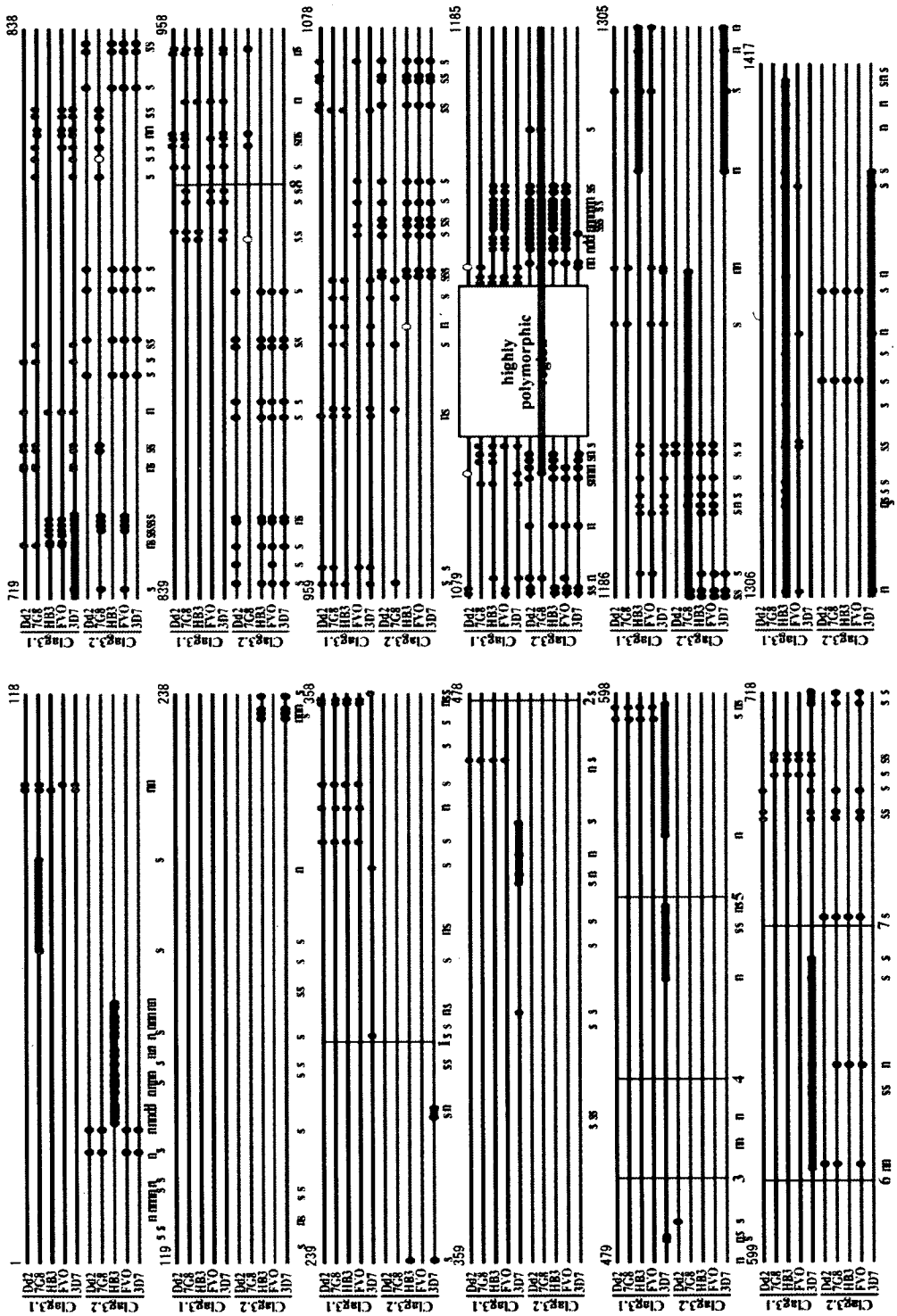


Fig. 3. Gene conversion tracts in *clag3.1* and *clag3.2*. Polymorphic codons (circles) in the coding sequences of Clag3.1 and 3.2 were compared in five *P. falciparum* lines. Clag3.1, black bar; Clag3.2, gray bar. Polymorphisms matching the paralogous sequence are shown in gray or black circles, respectively, and rare polymorphisms by an open circle. Exons are separated by vertical bar with the intron number at the bottom. Polymorphic sites that differ between consensus sequences are shown below the line classified as nonsynonymous (n), synonymous (s), and deletion (d). Gene conversion tracts identified using algorithm by Betrán et al. [34], wide gray bars.

lines showed the 2 expected positive bands with the primer sets 3.1F–3.1R and 3.2F–3.2R, whereas 9 parasite lines showed a positive band only with the primer set 3.2F–3.1R, which suggests that these 9 parasite lines possessed a hybrid gene with *clag3.2* sequence at the 5′ UTR and *clag3.1* sequence at the 3′ UTR (Fig. 2B). DNA fragments were not amplified with other primer combinations, indicating that artificial amplification due to primer mispairing was negligible. This is consistent with a

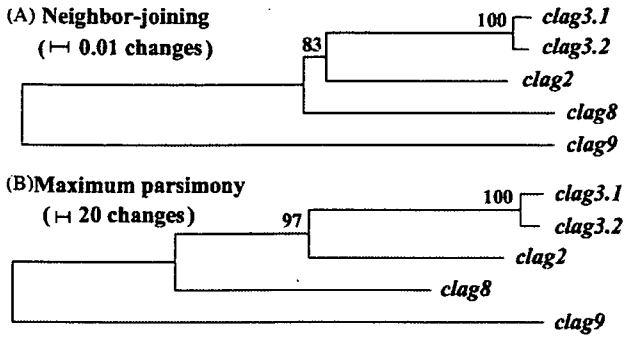


Fig. 4. Unrooted dendrograms of *pfrhoph1/clag* genes using nucleotide sequences from the 3D7 parasite line. The trees were constructed by the neighbor-joining and maximum parsimony methods using MEGA 3.1. Numbers on branches indicate bootstrap values (500 pseudoreplicates).

recent report by Chung et al. [36], who found that some parasite lines possess only a single *rhoph1/clag* on chr 3 by Southern blot hybridization. We here designate this *clag3* gene as *clag3h* (*clag3 hybrid*; Clag3H for protein). In addition, we obtained two distinct sequences for *clag3.1* from the KMWII parasite line (Fig. 5), using several cloned plasmids after experiencing difficulty in direct sequencing of PCR products. Sequences for *clag2*, *3.2*, and *8* were easily obtained from the KMWII line by direct sequencing of the PCR products, supporting the assumption that this line was a clone. Thus, the KMWII line appears to possess at least three *clag3*-related sequences in the genome. This data suggests that the number of *clag3*-related sequences in *P. falciparum* varies from one to at least three.

To deduce the direction of the one-gene to two-gene (or vice versa) change, we searched *P. reichenowi* orthologs in the genome database and found one sequence read (reich908g11.plk) showing high similarity with the sequence around the start codon of *clag3.1* and *3.2*. We also found two reads (reich289f06.qlk and reich1194c08.plk) showing strong homology with the sequence around the stop codons of *clag3.1* and *3.2*. Comparison of the nucleotide sequences at the UTR revealed that reich908g11.plk and reich289f06.qlk were similar to the *pfclag3.2* sequence and that reich1194c08.plk was similar to the *pfclag3.1* sequence (Fig. 6). Thus, duplication of *clag3.1* and *3.2* gene loci appears to predate the divergence of *P. falciparum* and *P. reichenowi*, suggesting that a single *rhoph1/clag* (*clag3h*) found in some *P. falciparum* lines is likely a result of an unequal crossover between two closely related genes. Notably, Clag3H had characteristic amino acids that were not observed in Clag3.1 and 3.2. For example, Ala at 1116 was found in three of nine Clag3H (30%). If Clag3H originated recently, for example during culture, the amino acid allele observed in Clag3H would also exist in Clag3.1 or 3.2; however, Ala at 1116 was not found in a total of 36 sequences of non-Clag3H protein sequences. Three in nine Clag3Hs is a significant excess compared to zero Ala at 1116 in 36 non-Clag3H sequences by Fisher's exact test ($P = 0.013$). This suggests that at least some Clag3H have accumulated some unique amino acid substitutions since their creation.

### 3.5. Selection on the PfRhopH complex

Positive selection was evaluated by comparing synonymous and nonsynonymous substitutions (Table 1). A significant excess of $d_N$ over $d_S$ was observed for *clag9* (entire ORF of five parasite lines) and for *clag8* (highly polymorphic region at nt positions 3022–3606 of 26 parasite lines), suggesting positive selection acting on these genes. A sliding window plot of $d_N$:$d_S$ ratios revealed that *clag2* and *8* had the highest peaks, around nt positions 3000–3600 (Fig. 7). It should be noted that the corresponding regions of *clag3.1* and *3.2* are the regions showing highly extensive polymorphism with indels (asterisks in Fig. 7), thereby preventing evaluation of $d_N$:$d_S$ ratios in this region. The peak at the N-terminus of *clag3.2* is due to introduction of part of the *clag3.1* sequence into the HB3 line *clag3.2* by gene conversion (see Fig. 3).

Positive selection was further evaluated by the McDonald–Kreitman test using *P. reichenowi* orthologs for *clag2* and *8*. Significant excess of intraspecific nonsynonymous substitutions over synonymous substitutions was observed in *clag8* as compared with interspecies fixed differences of nonsynonymous and synonymous changes, suggesting positive selection (Table 2).

### 3.6. Early origin of the clag2 and 8 polymorphism

We estimated the TMRCA for *clag2* and *8* polymorphism using aligned regions. Distances of synonymous single-nucleotide polymorphisms are $0.0139 \pm 0.0031$ for *clag2* and $0.0106 \pm 0.0030$ for *clag8*. Distances between *P. falciparum* and *P. reichenowi* are $0.0455 \pm 0.0082$ and $0.0748 \pm 0.0120$ for *clag2* and *8*, respectively. Assuming that the divergence time of *P. falciparum* and *P. reichenowi* was 6 million years ago (mya) [37,38], the estimated TMRCA of the polymorphism of *clag2*

Table 2
The McDonald–Kreitman test of selection for *Plasmodium falciparum clag2* and *8*

| Locus | $n$[a] | No. of sites | Fixed differences between species | | Polymorphic sites within *P. falciparum* | | |
|---|---|---|---|---|---|---|---|
| | | | Syn | Nsyn[b] | Syn | Nsyn | $P$[c] |
| *clag2* | 5 | 3273 | 23 | 47 | 18 | 12 | (0.011) |
| *clag8* | 5 | 2715 | 36 | 55 | 11 | 38 | 0.030 |

[a] *n*, Number of *P. falciparum* lines used.
[b] Syn, synonymous; Nsyn, nonsynonymous substitutions.
[c] Fisher's exact test (one-tailed) was used. *P*-value indicates that Nsyn are significantly greater than Syn. Value in parenthesis indicates that Syn are significantly greater than Nsyn.

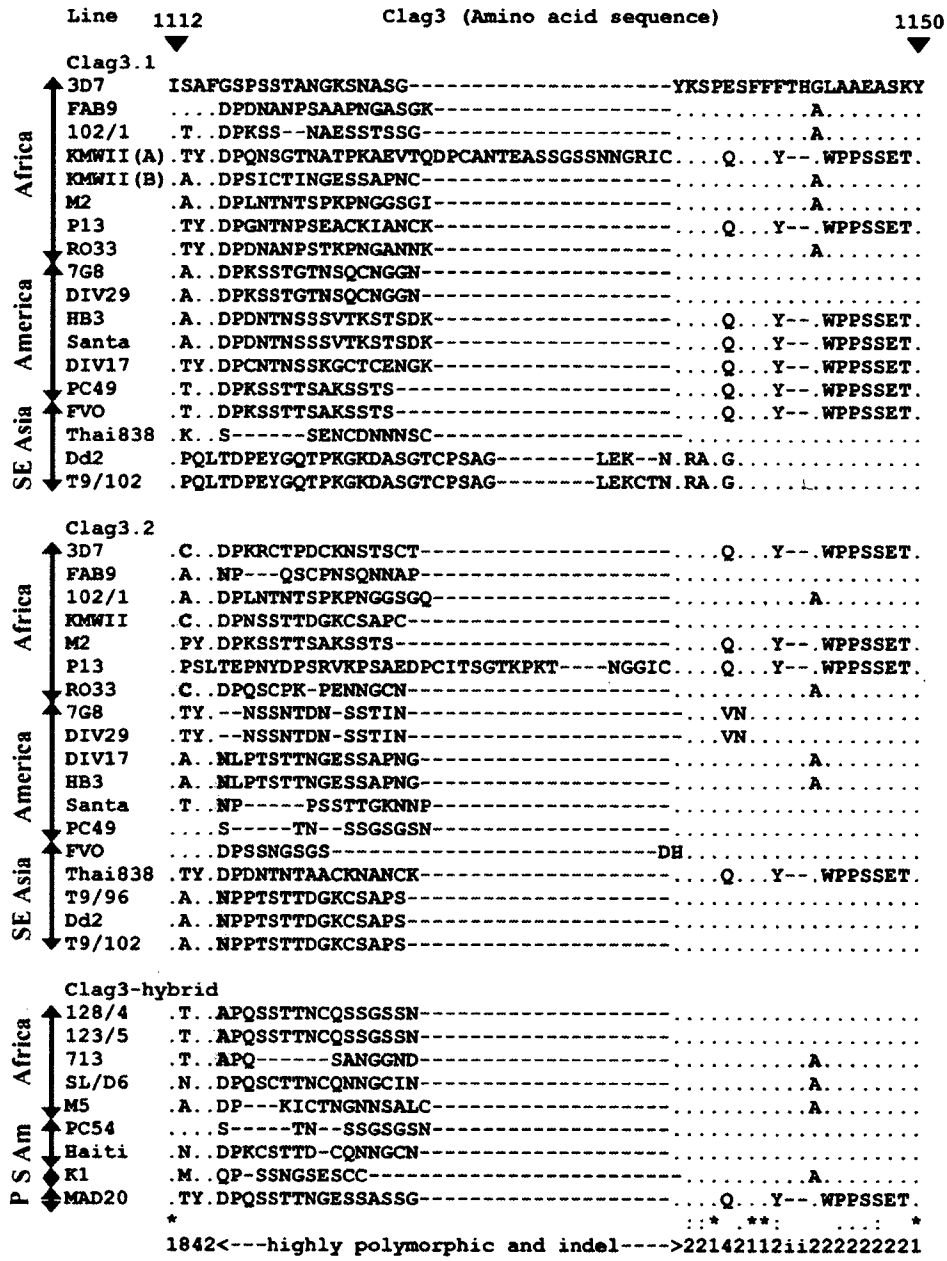


Fig. 5. Polymorphism of Clag3. An amino acid region 1112–1150 (after 3D7 line Clag3.1 sequence) of *P. falciparum* (27 lines) was aligned. Geographic origins are shown at left: SE Asia or S, Southeast Asia; Am, America; P, Papua New Guinea; Santa, Santa Lucia cloned line. Identical, conserved, or semiconserved residues in the alignment are indicated with asterisk, colon, or period, respectively. The number of amino acid replacements at each position and the region with indels are shown at the bottom. Cys residue at aa position 1113 and Asn and Ala residues at aa position 1116 are masked.

and *8* are 1.89 (95% CI, 1.02–3.18) and 0.87 (95% CI, 0.42–1.54) mya, respectively.

## 4. Discussion

### *4.1. Diversifying selection on the rhoph1/clag gene loci*

The present study revealed that the RhopH1/Clag-encoding genes *clag2*, *3.1*, *3.2*, and *8* contain a highly polymorphic region, particularly at nt positions 3000–3600. Diversifying selection increases nucleotide diversity ($\pi$), and an excess of $d_N$ to $d_S$ is indicative of positive selection favoring amino acid replacement [39]. Thus, the observed excess of $d_N$ to $d_S$ at nt positions 3000–3600 of *clag8* suggests that the polymorphism in *clag8* is positively maintained. An excess of $d_N$ to $d_S$ was also observed for *clag9*, indicating that this gene is also under positive selection. The most polymorphic region, in which positive selection was detected for *clag8*, was excluded from *clag2*, *3.1*, and *3.2* due to extensive sequence variation that made sequence alignment unreliable. Further analysis is required to evaluate positive selection on these three *rhoph1/clag* genes.