mation following primary infection with *Schistosoma mansoni*. *J Immunol 158:* 294–300.
20. Owhashi M, Nawa Y, Watanabe N, 1989. Granulomatous response in selective IgE-deficient SJA/9 mice infected with *Schistosoma japonicum*. *Int Arch Allergy Appl Immunol 90:* 310–312.
21. Silveira AM, Bethony J, Gazzinelli A, Kloos H, Fraga LA, Alvares MC, Prata A, Guerra HL, Loverde PT, Correa-Oliveira R, Gazzinelli G, 2002. High levels of IgG4 to *Schistosoma mansoni* egg antigens in individuals with periportal fibrosis. *Am J Trop Med Hyg 66:* 542–549.
22. Hirsch C, Zouain CS, Alves JB, Goes AM, 1997. Induction of protective immunity and modulation of granulomatous hypersensitivity in mice using PIII, an anionic fraction of *Schistosoma mansoni* adult worm. *Parasitology 115:* 21–28.
23. Hirsch C, Carvalho-Queiroz C, Franco GR, Pena SD, Simpson AJ, Goes AM, 1997. Evidentiation of paramyosin (Sm-97) as a modulating antigen on granulomatous hypersensitivity to *Schistosoma mansoni* eggs. *Mem Inst Oswaldo Cruz 92:* 663–667.
24. Ohmae H, Tanaka M, Hayashi M, Matsuzaki Y, Kurosaki Y, Blas BL, Portillo GG, Sy OS, Irie Y, Yasuraoka K, 1992. Improvement of ultrasonographic and serologic changes in *Schistosoma japonicum*-infected patients after treatment with praziquantel. *Am J Trop Med Hyg 46:* 99–104.
25. Ohmae H, Tanaka M, Hayashi M, Matsuzaki Y, Kurosaki Y, Blas BL, Portillo GG, Sy OS, Irie Y, Yasuraoka K, 1992. Ultrasonographic and serologic abnormalities in *Schistosoma japonicum* infection in Leyte, the Philippines. *Am J Trop Med Hyg 46:* 89–98.
26. Ohmae H, Sy OS, Chigusa Y, Portillo GP, 2003. Imaging diagnosis of schistosomiasis japonica: the use in Japan and application for field study in the present endemic area. *Parasitol Int 52:* 385–393.
27. Li YS, Sleigh AC, Ross AG, Li Y, Williams GM, Tanner M, McManus DP, 2000. Two-year impact of praziquantel treatment for *Schistosoma japonicum* infection in China: reinfection, subclinical disease and fibrosis marker measurements. *Trans R Soc Trop Med Hyg 94:* 191–197.
28. Tsukamoto T, Yamamoto T, Ikebe T, Takemura S, Shuto T, Kubo S, Hirohashi K, Kinoshita H, 2004. Serum markers of liver fibrosis and histologic severity of fibrosis in resected liver. *Hepatogastroenterology 51:* 777–780.
29. Acosta LP, McManus DP, Aligui GD, Olveda RM, Tiu WU, 2004. Antigen-specific antibody isotype patterns to *Schistosoma japonicum* recombinant and native antigens in a defined population in Leyte, the Philippines. *Am J Trop Med Hyg 70:* 549–555.
30. Jankovic D, Kullberg MC, Dombrowicz D, Barbieri S, Caspar P, Wynn TA, Paul WE, Cheever AW, Kinet JP, Sher A, 1997. Fc epsilonRI-deficient mice infected with *Schistosoma mansoni* mount normal Th2-type responses while displaying enhanced liver pathology. *J Immunol 159:* 1868–1875.
31. Deng J, Gold D, LoVerde PT, Fishelson Z, 2003. Inhibition of the complement membrane attack complex by *Schistosoma mansoni* paramyosin. *Infect Immun 71:* 6402–6410.
32. Laclette JP, Shoemaker CB, Richter D, Arcos L, Pante N, Cohen C, Bing D, Nicholson-Weller A, 1992. Paramyosin inhibits complement C1. *J Immunol 148:* 124–128.
33. Parizade M, Arnon R, Lachmann PJ, Fishelson Z, 1994. Functional and antigenic similarities between a 94-kD protein of *Schistosoma mansoni* (SCIP-1) and human CD59. *J Exp Med 179:* 1625–1636.
34. Loukas A, Jones MK, King LT, Brindley PJ, McManus DP, 2001. Receptor for Fc on the surfaces of schistosomes. *Infect Immun 69:* 3646–3651.
35. Jankovic D, Cheever AW, Kullberg MC, Wynn TA, Yap G, Caspar P, Lewis FA, Clynes R, Ravetch JV, Sher A, 1998. CD4+ T cell-mediated granulomatous pathology in schistosomiasis is downregulated by a B cell-dependent mechanism requiring Fc receptor signaling. *J Exp Med 187:* 619–629.

Transactions of the Royal Society of Tropical Medicine and Hygiene (2007) 101, 34–39

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevierhealth.com/journals/trst

Control of *Schistosoma mekongi* in Cambodia: results of eight years of control activities in the two endemic provinces

M. Sinuon[a], R. Tsuyuoka[b], D. Socheat[a], P. Odermatt[c], H. Ohmae[d], H. Matsuda[e], A. Montresor[f,*], K. Palmer[g]

[a] *National Center for Parasitology, Entomology and Malaria Control, Ministry of Health of Cambodia Phnom Penh, Cambodia*
[b] *WHO Office, Phnom Penh, Cambodia*
[c] *Swiss Tropical Institute, Basel, Switzerland*
[d] *National Institute of Infectious Diseases, MoHLW, Shinjuku-ku, Tokyo, Japan*
[e] *Dokkyo University, School of Medicine, Mibu, Tochigi, Japan*
[f] *World Health Organization, 63 Tran Hung Dao Street, Mail P.O. Box 52, Hanoi, Vietnam*
[g] *WHO Western Pacific Regional Office, Manila, Philippines*

Received 9 December 2005; received in revised form 6 April 2006; accepted 6 April 2006
Available online 9 October 2006

KEYWORDS
Schistosomiasis;
Schistosoma mekongi;
Helminths;
Ascaris lumbricoides;
Trichuris;
Control;
Cambodia

Summary In Cambodia, schistosomiasis is transmitted in the provinces of Kratie and Stung Treng where approximately 80 000 individuals are estimated to be at risk of infection. The baseline prevalence of infection was estimated to be between 73% and 88%, and cases of severe morbidity (hepatosplenomegaly, puberty retardation) and mortality were very common. In 1994, the Ministry of Health of Cambodia started schistosomiasis control applying universal chemotherapy with praziquantel (40 mg/kg). The coverage of the programme was between 62% and 86% for 8 years. This simple control measure resulted in the control of the disease: no cases were reported in 2004 and only three cases were reported in 2005. In addition, there are no longer reports of cases of severe morbidity due to schistosomiasis. Since the beginning of the control programme, a single dose of mebendazole (500 mg) has been combined with praziquantel during the mass chemotherapy; as a result the prevalence of *Ascaris lumbricoides* and hookworms dropped from 74.5% to 10% and from 86% to 40% respectively. The experience in Cambodia demonstrates that, with political commitment, control of parasitic diseases is achievable even in a situation of minimal resources. The programme represents a successful model for other developing countries.

* Corresponding author. Tel.: +84 4 943 3734/5/6x29; fax: +84 4 943 3740.
E-mail address: montresora@vtn.wpro.who.int (A. Montresor).

0035-9203/$ – see front matter © 2006 Royal Society of Tropical Medicine and Hygiene. Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.trstmh.2006.04.011

1. Introduction

Schistosomiasis is one of the most prevalent parasitic infections in the world. It is endemic in 76 countries and continues to be a public health concern in the developing world (Engels et al., 2002). In Cambodia the disease is caused by *Schistosoma mekongi* (Voge et al., 1978), and the intermediate host is *Neotricula aperta*, a river snail that lives in the fissures of partially submerged rocks (Mouchet, 1995). Pigs and dogs have been found to be animal reservoirs of the parasite (Strandgaard et al., 2001). Symptoms and signs associated with *S. mekongi* infection include cachexia, hepatosplenomegaly, stunting and retardation of puberty, portal hypertension, ascites and rupture of oesophageal varices (Biays et al., 1999). Pathology associated with the infection consists in periportal thickening and portal vein enlargement (Hatz, 2001).

In 1994, 20 villages in Kratie Province were identified as the origin of severe cases of schistosomiasis (Stich et al., 1999). In the same year, the Ministry of Health started control activities with the initial support of Médicins Sans Frontières (MSF) and subsequently of the WHO. The control measures consisted mainly of periodic administration of praziquantel (40 mg/kg) to the entire population, except for children under 2 years of age and pregnant women.

At the same time, surveys were conducted to accurately establish the extent of the endemic area. From 2000, the campaigns covered districts in Stung Treng Province, which were inaccessible in the past due to lack of security. Since 2001, the campaign has covered 80 000 people in Kratie and Stung Treng (Figure 1). No new endemic areas have been identified since then.

The parasitological stool surveys carried out for schistosomiasis monitoring also revealed high infection rates for soil-transmitted helminths (STH). This finding led to the simultaneous administration of mebendazole (single dose 500 mg) to all the individuals treated with praziquantel.

The data presented in this paper were collected for monitoring purposes. Despite some gaps and lacks of continuity in the intervention, and in the data collection and recording (due to temporary lack of funds, armed conflicts and other unpredictable events), we consider the present data to be interesting for managers of similar programmes.

The paper aims to demonstrate that a major reduction of schistosomiasis prevalence can be obtained with mass distribution of anthelminthics and that this can be obtained also in a situation of minimal resources. The paper also shows the feasibility of integrating different anthelminthics in the same distribution system.

2. Materials and methods

2.1. Identification of the endemic area

Epidemiological assessments were conducted in the zones previously reported as endemic: Kratie (Audebaud et al., 1968; Jolly et al., 1970), Strung Treng (Urbani and Socheat, 1997), Rattanakiri and Kampong Cham (Ijima, 1970).

Three methods were used in order to delimit the area of intervention:

- questionnaires
- stool surveys in households and schools
- serological surveys.

Normally the first investigation was done with a questionnaire and then the positive data confirmed with parasitological or serological methods.

Over 30 000 questionnaires were distributed in the provinces of Kratie, Stung Treng, Rattanakiri and Kompong Cham (Urbani et al., 2002), over 1300 individuals were investigated using the Kato-Katz method (WHO, 1980) during household surveys and over 1200 during school surveys (Stich et al., 1999), and 12 villages were screened using

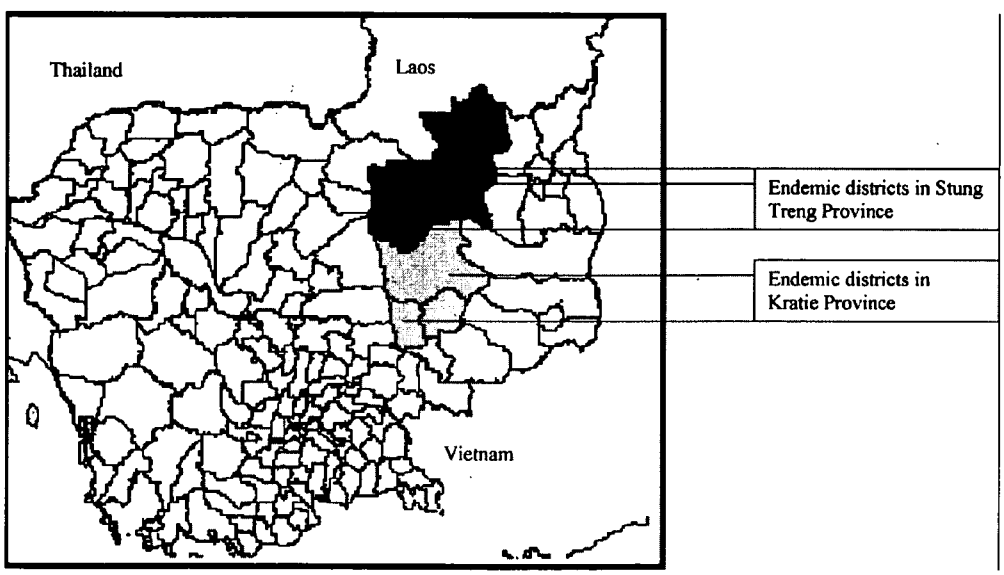


Figure 1 Location of districts endemic for *Schistosoma mekongi* in Cambodia.

an ELISA with soluble egg antigen of *S. japonicum* (Homae, 2004).

The data collected with the different assessments were analyzed, checked for consistency, assembled and then used to prepare maps showing the areas to be covered with mass distribution of anthelminthics.

2.2. Control intervention

Every year, between February and April, the campaign was conducted by a team from the National Center for Parasitology, Entomology and Malaria Control, and local health staff. Drugs were administered under direct observation of health staff. Treatments were combined with health education, in order to increase health awareness and to reduce contact with water on the rocky banks of the river.

From 1994 to 2005, the area of intervention increased from the 20 villages where schistosomiasis was originally reported to include all new areas that were identified successively. Since 2001, the campaign has covered 80 000 people. The target population for each campaign and details of the type of mass treatment applied in different years are presented in Table 1.

2.2.1. Coverage

The programme coverage has been evaluated through the drug distribution reports: on the occasion of each drug distribution campaign, forms recording age, sex, weight and number of tablets of praziquantel given to each individual were compiled by health staff. The coverage was calculated as the number of people treated divided by population estimates (census data periodically updated by the health centres).

2.2.2. Impact on schistosomiasis prevalence

The programme's impact on schistosomiasis prevalence has been periodically estimated with three methods:

- Parasitological surveys were conducted in randomly selected villages: between 2000 and 3000 individuals in an average of 20 villages were selected every year. The sample size in each village was between 70 and 150. Stool samples were examined with the Kato-Katz method and data on STHs were also collected.
- Annual parasitological surveys were conducted in primary schools of four sentinel villages in Kratie Province (Achen, Chatnaol, Srekheurn and Sambok) (from 1995) and in Sdau village in Stung Treng Province (from 1997). After the year 2000 in each site, three consecutive stool samples were examined with the Kato-Katz method.
- There was active search and follow up of cases of severe schistosomiasis during the mass treatment and the parasitological survey.

The monitoring surveys have been conducted in school-aged children because this age group is easily accessible and can provide indications of the situation in other age groups (Guyatt et al., 1999).

3. Results

3.1. Epidemiology

The epidemiological assessment identified 114 villages along the Mekong River and its tributaries Sesan and Sekong as endemic for schistosomiasis: the villages were in two districts of Kratie Province and in five districts of Stung Treng Province (Figure 1). About 80 000 individuals living along the Mekong River were estimated to be at risk of infection (Urbani et al., 2002). The area between Kratie and Sambo, on the left bank of the Mekong River, was identified as the most affected (over 70% positive answers to the questionnaire). No evidence of *S. mekongi* transmission came from Kampong Cham and Rattanakiri.

Table 1 Mass drug administration campaigns for the control of schistosomiasis in Cambodia, 1994–2005

Year	Province covered	Type of mass treatment conducted	Organizing institution	Approximate target population
1994/95	Kratie	Universal	MoH-MSF	45 000
1996	Kratie	Universal	MoH-MSF	45 000
1997	Kratie and Strung Treng	Universal	NMC-MSF	60 000
1998	Kratie and Strung Treng	No MDA[a]	NMC-MSF	–
1999	Kratie and Strung Treng	Universal	NMC-MSF	60 000
2000	Kratie and Strung Treng	Universal	NMC-MSF-WHO	70 000
2001	Kratie and Strung Treng	Universal	NMC-WHO	80 000
2002	Kratie and Strung Treng	Universal	NMC-WHO	80 000
2003	Kratie and Strung Treng	Targeted[b]	NMC-WHO	78 500
	Kratie and Strung Treng	Universal[c]	NMC-WHO	1 500
2004	Kratie and Strung Treng	Universal	NMC-WHO	80 000
2005	Kratie and Strung Treng	Universal	NMC-WHO	80 000

MoH: Ministry of Health; MSF: Médicins Sans Frontières; NMC: National Center for Parasitology, Entomology and Malaria Control.
[a] Mass drug administration was not conducted because of lack of funds.
[b] In 112 villages.
[c] In three villages.

Table 2 Prevalence of *Schistosoma mekongi* in Stung Treng (estimated in villages randomly selected every year)

Year	Coverage	Prevalence survey (by village)			
		No. of villages surveyed	No. of exams	No. of villages with active transmission (%)	Range of prevalence in positive villages (%)
1997	64%	13	1033	11 (85)	6–88%
1998	–[a]	6	401	9 (83)	5.1–43.1%
1999	63%	8	676	5 (62)	1.8–7.7%
2000	64%	13	849	10 (76)	2–19.6%
2001	67%	13	1029	5 (39)	1.4–7.2%
2002	62%	13	999	5 (39)	1.3–4.5%
2003	–[b]	14	1573	0	–
		5	502	0	–
2004	83.6%	8	905	0	–
2005	81.5%	10	1298	3 (30%)	0.7–3.5%

[a] Mass drug administration was not conducted because of lack of funding.
[b] Coverage was not estimated.

3.2. Mass treatment

The coverage of mass treatment was maintained between 62% and 74% between 1996 and 2002 with the exception of 1998 when universal drug administration was not conducted because lack of funds. In 2000 and 2001, the campaigns covered new areas, previously inaccessible due to lack of security. Since 2001 the campaign has covered 50 000 people in 56 villages in two districts of Kratie Province, and about 30 000 people in 58 villages in five districts of Stung Treng Province.

3.3. Programme impact on schistosomiasis

The initial surveys performed by MSF in Kratie Province between 1994 and 1995 revealed a prevalence of schistosomiasis in primary school children of 72.9% (Stich et al., 1999). Table 2 shows the decline in the number of villages presenting cases of schistosomiasis in Stung Treng and the range of prevalence. For Kratie, only the data on declining schistosomiasis prevalence in the four sentinel sites are available (Figure 2).

3.4. Clinical morbidity

Before the start of the intervention several new cases of severe schistosomiasis morbidity (hepatosplenomegaly, cachexia, anaemia, ascites, haematoemesis) and mortality were reported every year at the Provincial Hospital of Kratie and the District Hospital of Sambour as well as in the surrounding communities (Biays et al., 1999). In a survey in 1999, the number of cases identified with signs and symptoms of severe schistosomiasis reached 124 in Kratie Province. After treatment with praziquantel, 101 of them improved and were then able to perform light labour, four patients died and in eight cases the disease remained severe. Of the 11 patients who received surgical treatment between 2000 and 2002 at the National Calmette Hospital, 10 patients recovered and one patient died a few days after the operation. In Stung Treng Province, the registration and follow-up could not be done because people move more frequently from their home villages to field huts on farms.

Every year during the stool survey, clinical examination has been conducted to assess signs and symptoms including liver and spleen enlargement. Since 2003 no new symptomatic case has been observed in the provinces' health facilities.

3.5. Soil-transmitted helminths

The prevalence of STHs was examined in a sample of villages in Kratie and Stung Treng in 1997.

In Kratie, the prevalence of *Ascaris lumbricoides* was between 9.5% and 74.5%, *Trichuris trichiura* between 0% and 15.9%, and hookworms between 18.2% and 86%. Since 1997, mebendazole 500 mg has been combined with praziquantel during the mass drug administration. The prevalence of *A. lumbricoides* and *T. trichiura* decreased dramatically while

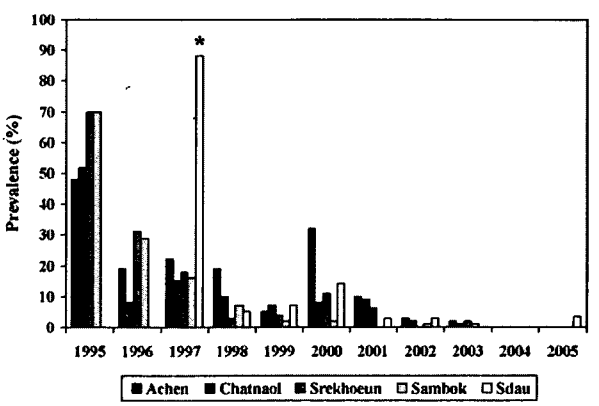


Figure 2 Changes in schistosomiasis prevalence in sentinel sites during the control programme implementation. * In Stung Treng Province the treatment started in 1998: Sdau is the only sentinel site in Stung Treng.

Table 3 Changes in prevalence of soil-transmitted helminths in selected villages of Kratie and Stung Treng between 1997 and 2005

	1997			2005		
	Ascaris lumbricoides	*Trichuris trichiura*	Hookworms	*Ascaris lumbricoides*	*Trichuris trichiura*	Hookworms
Achen	25	1.6	54.7	0	0	14.3
Chartnol	9.5	2.4	46.4	0.9	0	6.1
Sambok	16.2	2.1	45.1	0	0	9
Srekhoeun	13.8	3.4	50	0	0	5.7
Sdau	51.3	9.5	86	0	0.7	18.4
Koh Sneng	34	2.1	55.3	5.4	2.0	15.6
K. Chanh Tuk	69.6	2.2	67.4	0	0	26

the prevalence of hookworms has halved but remains relatively high. Table 3 presents the STH prevalence in villages for which data for both 1997 and 2005 are available.

4. Discussion

According to the UNDP Human Development Report (2003), Cambodia ranks 130th out of 175 countries on the human development index. Despite this low ranking and because decision makers are strongly committed, Cambodia was able to organize universal administration of praziquantel and mebendazole for 8 years resulting in a reduction of schistosomiasis and STH infection, so that:

- No new cases of severe morbidity due to schistosomiasis have been reported in the past four years in the health facilities in the area.
- Relevant nutritional benefits are expected after the reduction of STH (WHO, 2002).

We interpret the presence of new cases in three villages in 2005 as probably due to contamination by animal reservoirs (Strandgaard et al., 2001), immigrants from Laos or infected people that were not covered by the control programme. The new cases demonstrate also that all the conditions are in place for the resurgence of the disease if the control measures are interrupted. Experiences from Lao PDR demonstrated that, after a drastic reduction in prevalence, if the drug pressure is not maintained the parasite could easily return to original levels (Urbani et al., 2002). We suggest that the drug distribution should be maintained until sanitation standards have improved significantly. A possible option to maintain the disease under control and to reduce the cost of the yearly mass distribution would be to increase the intervals between mass administrations (i.e. every 2 years) and associate this with a very sensitive (ELISA) monitoring of the prevalence of the infection (Homae et al., 2004).

The experience in Cambodia demonstrates that with political commitment, parasitic disease control is achievable even in situation of minimal resources and that different anthelminthic drug can be provided by the same delivery system. The programme represents a successful model for other developing countries.

Conflicts of interest statement
The authors have no conflicts of interest concerning the work reported in this paper.

Acknowledgements

This article is dedicated to Dr Carlo Urbani. An indispensable contribution was made by MSF. We would like to thank all the staff from the National Center for Parasitology, Entomology and Malaria Control, Provincial Health Department in Kratie and Stung Treng, Ministry of Health, Calmette Hospital, Ministry of Education Youth and Sports. The programme has been funded by Médecins Sans Frontières (1994–1999) and the Sasakawa Memorial Health Foundation and the Japanese Government through the WHO (1999 to present).

References

Audebaud, G., Tournier-Lasserve, C., Brumpt, V., Jolly, M., Mazaud, R., Imbert, X., Bazillio, R., 1968. 1st case of human schistosomiasis observed in Cambodia (Kratie area). Bull. Soc. Pathol. Exot. Filiales 61, 778–784.

Biays, S., Stich, A.H.R., Odermatt, P., Chan, L., Yersin, C., Chan, M., Chaem, S., Lormand, J.-D., 1999. Foyer de bilharziose à *Schistosoma mekongi* redécouvert au Nord du Cambodge: I. Perception culturelle de la maladie; description et suivi de 20 cas cliniques graves. Trop. Med. Int. Health 4, 662–673.

Engels, D., Chitzulo, L., Montresor, A., Savioli, L., 2002. The global epidemiological situation of schistosomiasis and new approaches to control and research. Acta Trop. 82, 139–146.

Guyatt, H.L., Brooker, S., Dönnelly, C.A., 1999. Can prevalence of infection in school-aged children be used as an index for assessing community prevalence? Parasitology 118, 257–268.

Hatz, C., 2001. The use of ultrasounds in schistosomiasis. Adv. Parasitol. 48, 225–284.

Homae, H., Sinuon, M., Kirinoki, M., Matsumoto, J., Chigusa, Y., Socheat, D., Matsuda, H., 2004. *Schistosoma mekongi*: from discovery to control. Parasitol. Intern. 53, 135–142.

Ijima, T., 1970. Enquete sur la schistosomiase dans le basin du mekong. Rapport de mission WPR/059/70. World Health Organization, Geneva.

Jolly, M., Bazillio, R., Audebaud, G., Brumpt, V., Sophinn, B., 1970. Existence of a focus of human bilharziosis, in Cambodia in Kratie area. II. Epidemiologic survey. Preliminary results. Med. Trop. 30, 462–471.

Mouchet, F., 1995. Malacological data on the transmission of *Schistosoma mekongi* in Cambodia. Abstract for the European Conference on Tropical Medicine, Hamburg, October 1995.

Stich, A.H.R., Biays, S., Odermatt, P., Chan, M., Cheam, S., Kiev, S., Chuong, S.L., Legros, P., Philips, M., Lormand, J.-D., Tanner, M., 1999. Foci of schistosomiasis mekongi, Northern Cambodia: II. Distribution of infection and morbidity. Trop. Med. Int. Health 4, 674–685.

Strandgaard, H., Johansen, M.V., Pholsena, K., Teixayavong, K., Christensen, N.O., 2001. The pig as a host for *Schistosoma mekongi* in Laos. J. Parasitol. 87, 708–709.

UNDP, 2003. Human Development Report. Oxford University Press, Oxford.

Urbani, C., Socheat, D., 1997. Schistosomiasis control project: activity report. Mèdecins Sans Frontieres Report, Phnom Penh.

Urbani, C., Sinoun, M., Socheat, D., Pholsena, K., Strandgaard, H., Odermatt, P., Hatz, C., 2002. Epidemiology and control of mekongi schistosomiasis. Acta Trop. 82, 157–168.

Voge, M., Bruckner, D., Bruce, J.I., 1978. *Schistosoma mekongi* sp. n. from man and animals, compared with four geographic strains of *Schistosoma japonicum*. J. Parasitol. 64, 577–584.

WHO, 1980. Manual of Basic Techniques for a Health Laboratory. World Health Organization, Geneva, pp. 377–378.

WHO, 2002. Prevention and control of schistosomiasis and soil-transmitted helminthiasis. Report of a WHO Expert Committee. World Health Organization, Geneva, Technical Report Series No. 912.

Available online at www.sciencedirect.com
ScienceDirect
Gene 394 (2007) 78–86
GENE
SECTION
EVOLUTIONARY GENOMICS
www.elsevier.com/locate/gene

Occurrence of multiple, independent gene fusion events for the fifth and sixth enzymes of pyrimidine biosynthesis in different eukaryotic groups☆

Takashi Makiuchi [a], Takeshi Nara [a,*], Takeshi Annoura [a], Tetsuo Hashimoto [b], Takashi Aoki [a]

[a] *Department of Molecular and Cellular Parasitology, Juntendo University School of Medicine, Hongo 2-1-1, Bunkyo-ku, Tokyo 113-8421, Japan*
[b] *Institute of Biological Sciences, University of Tsukuba, Tennoudai 1-1-1, Tsukuba, Ibaraki 305-8572, Japan*

Received 22 November 2006; received in revised form 13 February 2007; accepted 13 February 2007
Available online 24 February 2007
Received by S. Yokoyama

Abstract

The genes encoding orotate phosphoribosyltransferase (OPRT) and orotidine-5′-monophosphate decarboxylase (OMPDC), the fifth and sixth enzymes in the *de novo* pyrimidine biosynthetic pathway, are fused as *OPRT–OMPDC* in most eukaryotic groups. On the other hand, the inversely linked *OMPDC–OPRT* fusion is present in trypanosomatids, belonging to kinetoplastids together with bodonids in a supergroup, Euglenozoa. Here, we show the presence of *OMPDC–OPRT* in the bodonid, *Bodo caudatus*, while *OPRT–OMPDC* in *Euglena gracilis*, another euglenozoan species belonging to euglenoids. These results suggest that the *OMPDC–OPRT* fusion event occurred in a common ancestor of kinetoplastids. Genome sequence database searches further revealed the presence of *OMPDC–OPRT* in stramenopiles and cyanobacteria. Phylogenetic reconstruction of OPRT and OMPDC rejected statistically the monophyly of the *OPRT* domains of stramenopile and kinetoplastid *OMPDC–OPRT*, demonstrating that these gene fusions do not share a common evolutionary origin, despite the identical gene order. Thus, the *OMPDC–OPRT* fusion is likely to have emerged independently in these eukaryotic groups. Phylogenetic analyses also suggested that cyanobacterial *OMPDC–OPRT* arose *via* lateral transfer. We conclude that gene fusion events occur more frequently than previously thought and that lateral gene transfer has made a marked contribution to establishment of the rearranged structure of *OPRT* and *OMPDC* genes in eukaryotes.

Keywords: Gene fusion; *De novo* pyrimidine biosynthesis; Lateral gene transfer; Evolutionary marker

1. Introduction

Gene fusion, yielding formation of multidomain proteins, is one of the major driving forces in protein evolution and is believed to be a very rare evolutionary event in comparison to the frequency of point mutations. Gene fusion events were recently utilized as an evolutionary marker to pinpoint the root of the eukaryotic tree. Based on the unique distribution of a gene fusion of dihydrofolate reductase (DHFR) and thymidylate synthase (TS) in bikonts (Plantae, Alveolata/stramenopiles, Heterolobosea/Euglenozoa, *etc.*) and a three-gene fusion of the first three enzymes, carbamoyl–phosphate synthetase II (CPS II), aspartate carbamoyltransferase (ACT), and dihydroorotase (DHO), in the *de novo* pyrimidine biosynthetic pathway in unikonts (Metazoa, Fungi, and Amoebozoa), it was assumed that the root of the eukaryote tree lies in bikonts and unikonts (Stechmann and Cavalier-Smith, 2002, 2003).

The validity of a derived gene fusion as a phylogenetic marker is, however, a matter of debate for a number of reasons. First, it is difficult to exclude the possibility of lateral gene transfer (LGT). There is ample evidence of eukaryote-to-eukaryote LGT events (Keeling and Palmer, 2001, Andersson et al., 2003, Bergthorsson et al., 2003, Andersson et al., 2005),

Abbreviations: OPRT, orotate phosphoribosyltransferase; OMPDC, orotidine-5′-monophosphate decarboxylase; LGT, lateral gene transfer; BP, bootstrap proportion; DM, distance matrix; ML, maximum likelihood; MP, maximum parsimony.

☆ Sequence availability: The sequences reported in this paper appear in the GenBank™, EMBL, and DDBJ databases with the accession numbers AB185845 and AB185846 for *Euglena gracilis* orotate phosphoribosyltransferase (OPRT)-orotidine-5′-monophosphate decarboxylase (OMPDC) fused enzyme and *Bodo caudatus* OMPDC–OPRT fused enzyme, respectively.

* Corresponding author. Tel.: +81 3 5802 1043; fax: +81 3 5800 0476.
E-mail address: tnara@med.juntendo.ac.jp (T. Nara).

0378-1119/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2007.02.009

and LGT between distantly related groups is regarded as an important evolutionary mechanism in eukaryotes, as well as in prokaryotes (Boucher and Doolittle, 2000, de Koning et al., 2000, Field et al., 2000, Henze et al., 2001, Andersson and Roger, 2002, Nixon et al., 2002, Striepen et al., 2002).

Other factors affecting the validity of gene fusion as a phylogenetic marker include the possibilities of independent fusion events in different groups and of rearrangement of gene fusion including secondary splitting of fused genes and subsequent re-fusion. For example, the green alga *Mougeotia scalaris* has a chimeric photoreceptor gene similar to that of ferns, while the fern and algal genes seem to have arisen independently in plant evolution (Suetsugu et al., 2005). Stover et al. (2005) recently reported reciprocal fusions of formaldehyde dehydrogenase and *S*-formylglutathione hydrolase genes in ciliates and diatoms.

Orotate phosphoribosyltransferase (OPRT) and orotidine-5′-monophosphate decarboxylase (OMPDC) are the fifth and sixth enzymes of the *de novo* pyrimidine biosynthetic pathway, which consists of a series of six enzymatic reactions. In eukaryotes, the organization of *OPRT* and *OMPDC* genes is polytypic, and is therefore suitable for investigation of whether multiple gene fusion/fission events have occurred in this set of genes.

The *OPRT–OMPDC* gene fusion is present in the majority of eukaryotic groups, *i.e.*, Metazoa, Amoebozoa, Plantae, and Heterolobosea. On the other hand, they exist as separate genes in Fungi. Special attention should be paid to the inversely linked *OMPDC–OPRT* fusion in the parasitic protists, trypanosomatids (Gao et al., 1999), because the existence of the "reverse fusion" itself represents clear evidence of an independent fusion event. Phylogenetic analyses showed that the separate fungal *OPRT* and the *OMPDC* domain of the trypanosomatid *OMPDC–OPRT* have a different origin from other eukaryotes, suggesting the acquisition of these genes *via* LGT (Nara et al., 2000). Thus, the gene fusion of *OPRT* and *OMPDC* is not necessarily a stable evolutionary character and cannot be regarded as a decisive cladistic marker.

In the present study, we addressed 1) the timing of the *OMPDC–OPRT* fusion event on the line leading to trypanosomatids and 2) the evolutionary origin of *OMPDC–OPRT*. Trypanosomatids form a monophyletic group and clusters within bodonids in the kinetoplastid clade, which is included in a larger supergroup, Euglenozoa, together with the euglenoids and diplonemids. (Fernandes et al., 1993, Simpson et al., 2002). We demonstrated that the bodonid, *Bodo caudatus*, and the euglenoid, *Euglena gracilis*, possess *OMPDC–OPRT* and *OPRT–OMPDC*, respectively. Our results suggest that a common ancestor of kinetoplastids acquired *OMPDC–OPRT* after separation from the line leading to euglenoids.

We searched for *OMPDC–OPRT* in the various genome sequence databases and found *OMPDC–OPRT* in the diatom, *Thalassiosira pseudonana*, which belongs to the stramenopiles, and in a subset of cyanobacteria. Phylogenetic inference on the *OPRT* and *OMPDC* trees indicated that the *OMPDC–OPRT* fusion event occurred independently in stramenopiles and kinetoplastids. Cyanobacterial *OMPDC–OPRT* may have been acquired *via* LGT. We propose an evolutionary scenario of the multiple events of gene fusion and fission of *OPRT* and *OMPDC* genes in eukaryotes, paying special attention to the importance of LGT events in establishing gene fusion.

2. Materials and methods

2.1. Organisms, cultivation, and RNA extraction

E. gracilis Klebs (NIES-48 strain from Microbial Culture Collection at the National Institute for Environmental Studies, Tsukuba, Japan) and *B. caudatus* (ATCC 50361, from American Type Culture Collection, Rockville, MD, USA) were maintained and total RNA was extracted from freshly prepared cells of *E. gracilis* or *B. caudatus*, as described essentially (Annoura et al., 2005).

2.2. cDNA cloning of OPRT and OMPDC from Euglenozoa

We selected the conserved regions of the amino acid sequences of OPRT–OMPDC and OMPDC–OPRT and designed degenerate primers to amplify the cDNA fragments from *E. gracilis* and *B. caudatus* by nested PCR. The PCR products for *E. gracilis* were obtained using primers specific for OPRT–OMPDC and those for *B. caudatus* were obtained using the OMPDC–OPRT-specific primers. To prepare the full-length cDNA, we carried out rapid amplification of the 5′- and 3′-ends of cDNA (5′- and 3′-RACE, respectively). For 5′-RACE, cDNA were synthesized using antisense primers and then PCR was performed using a spliced leader (SL)-specific primer and antisense primer. 3′-RACE was performed using 3′-anchor primer and specific primer as described previously (Annoura et al., 2005). Finally, the full-length cDNA was amplified using the specific primer sets and KOD -plus- DNA polymerase (High-fidelity type DNA polymerase, TOYOBO Co., Ltd., Osaka, Japan), subcloned, and sequenced using an automated DNA sequencer. Primers used in this study are listed in Supplementary Table S1.

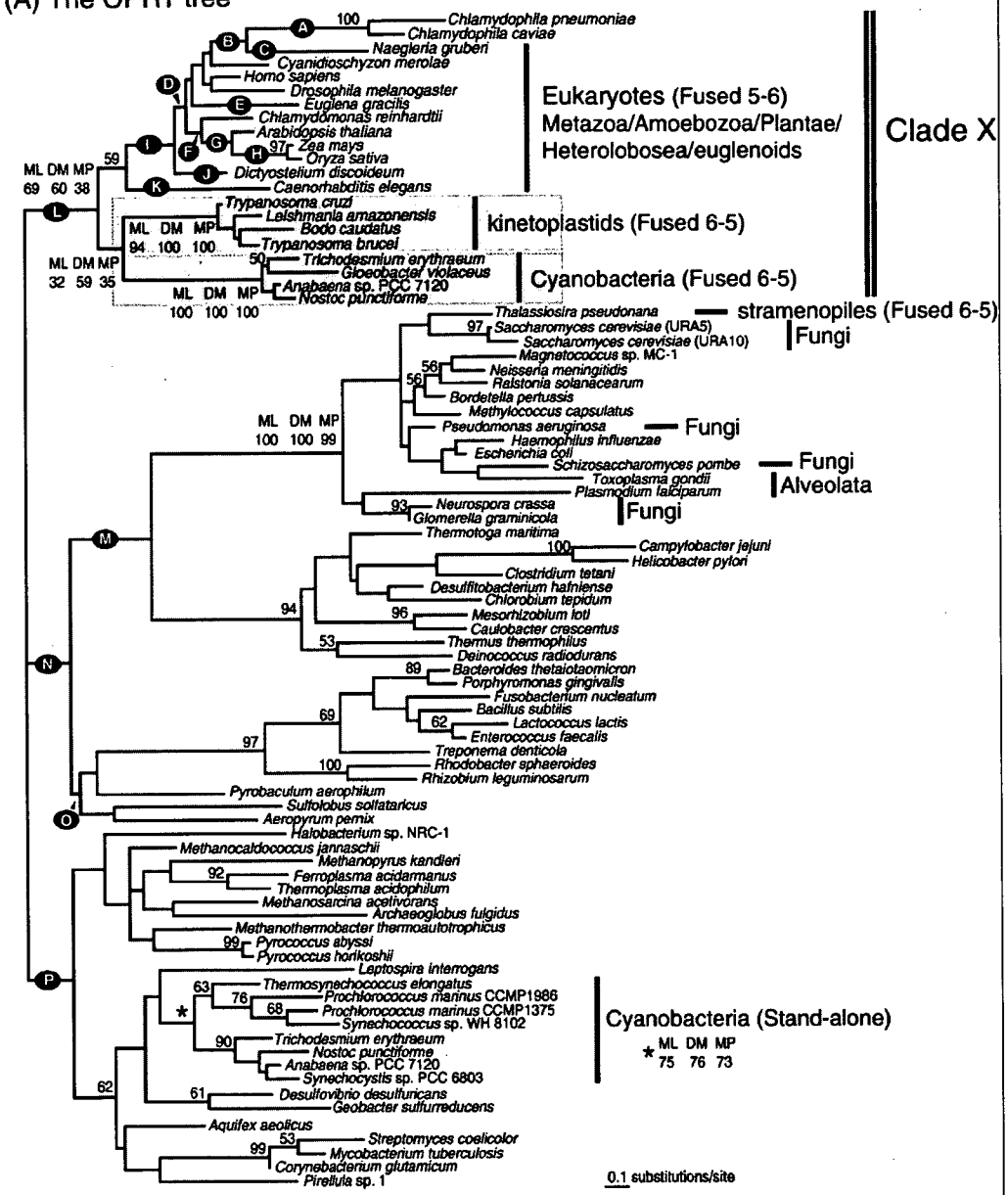
2.3. Phylogenetic analysis

All sequence data, with the exception of *E. gracilis* OPRT–OMPDC and *B. caudatus* OMPDC–OPRT first reported here, were collected from public sequence databases by taxonomic and BLAST searches. Unpublished OPRT and OMPDC sequences were obtained from various genome project databases (see Supplementary Table S2). Multiple alignments for OPRT and OMPDC sequences were obtained using SAM version 2.0 (Hughey and Krogh, 1996). The alignments were corrected by manual inspection, and the unambiguously aligned positions were selected from OPRT and OMPDC, respectively, and used for phylogenetic analyses. The original alignments and selected sites are given in Supplementary Tables S3 and S4.

The maximum likelihood (ML), distance matrix (DM), and maximum parsimony (MP) methods for protein phylogeny were applied to the data sets using the CODEML program in PAML3.1 (Yang, 1997) and PROML, PROTDIST, NEIGHBOR, PROTPARS, SEQBOOT, and CONSENSE programs in PHYLIP3.6a distributed by Dr. Joseph Felsenstein, University of Washington.

In ML analysis, an initial tree search was performed by applying PROML with the JTT model for amino acid substitution, assuming homogeneous rates across sites. Based on the best tree obtained, the Γ shape parameter (α) of the discrete Γ-distribution with four categories that approximated site rates was estimated using CODEML. Using this α value, a further tree

(A) The OPRT tree



(B) The probability (p) values of alternative trees

	Branch position	A	B	C	D	E	F	G	H	I	J	K
Groups grafted	Kinetoplastids (K)	$p = 0.000$	0.004	0.000	0.000	0.025	0.019	0.022	0.035	0.448	0.012	0.450
	Cyanobacteria (C)	0.063	0.044	0.101	0.023	0.022	0.023	0.004	0.006	0.078	0.099	0.081
	K plus C	0.025	0.070	0.036	0.052	0.135	0.093	0.089	0.051	0.666	0.192	0.344

	Branch position	L	M	N	O	P
Groups grafted	Kinetoplastids (K)	0.217	0.008	0.036	0.002	0.032
	Cyanobacteria (C)	0.921	0.059	0.174	0.059	0.169
	K plus C	0.786	0.053	0.281	0.053	0.050

search was performed with the JTT model with four site-rate categories using PROML with the global rearrangement option, producing the final best tree. In DM analysis, ML estimates for pairwise distances among the sequences analyzed were calculated using PROTDIST, based on the JTT model with rate variation allowed among sites. Then the neighbor-joining (NJ) tree was reconstructed from the distances using NEIGHBOR. In MP analysis, the MP tree was searched using PROTPARS. Bootstrap analysis for each of the three methods was performed in the same way by applying PROML, PROTDIST + NEIGHBOR, or PROTPARS to the resampled data sets produced by SEQBOOT. One hundred and 1000 resamplings were performed for ML and for DM and MP analyses, respectively. A consensus tree was generated using the CONSENSE program based on the bootstrap analysis of the ML method. Branch lengths of the consensus tree were estimated by using CODEML program.

The AU test (Shimodaira, 2002) in the CONSEL program (Shimodaira and Hasegawa, 2001) was used for statistical comparisons among the alternative trees.

3. Results

3.1. cDNA cloning of OPRT and OMPDC in Euglenozoa

We examined whether the *OMPDC–OPRT* fusion is a derived character in trypanosomatids or is common to kinetoplastids or Euglenozoa. cDNA cloning and sequence analysis clearly demonstrated that the bodonid, *B. caudatus*, possesses *OMPDC–OPRT* (GenBank™ AB185846), while the euglenoid, *E. gracilis,* has *OPRT–OMPDC* (GenBank™, AB185845). These results indicated that the *OMPDC–OPRT* fusion occurred in a common ancestor of kinetoplastids after diversification from the line leading to other euglenozoan groups.

B. caudatus OMPDC–OPRT possesses an SKL tripeptide at its C-terminus. The SKL sequence is a targeting signal to glycosomes, the peroxisome-like, kinetoplastid-specific cellular compartments (Michels et al., 2000). The presence of this tripeptide in *B. caudatus OMPDC–OPRT* suggests that *B. caudatus* OMPDC–OPRT is localized in glycosomes, similar to the trypanosomatid OMPDC–OPRT.

3.2. Separation of kinetoplastid and stramenopile OMPDC–OPRT in the OPRT tree

In conjunction with these findings, the organization of *OPRT* and *OMPDC* genes is classified into three different forms: 1) *OPRT–OMPDC* fusion in Metazoa, Amoebozoa, Plantae, Heterolobosea, and euglenoids, 2) *OMPDC–OPRT* fusion in kinetoplastids, in a diatom *T. pseudonana* (representative stramenopiles in the following sections), and in a subset of cyanobacteria, and 3) stand-alone genes in Fungi, in Apicomplexa, *Plasmodium* and *Toxoplasma* (representative Alveolata), and in prokaryotes. Interestingly, Apicomplexa possess the stand-alone *OPRT* and *OMPDC* genes, similar to Fungi, although they are phylogenetically closer to stramenopiles than to Fungi (Baldauf et al., 2000, Arisue et al., 2002, Bapteste et al., 2002).

We analyzed protein phylogeny to clarify the origin of *OMPDC–OPRT* in the phylogenetically distant groups and to trace the evolutionary history of *OPRT* and *OMPDC* genes in eukaryotes. The OPRT phylogeny reconstructed a strongly supported grouping of *Thalassiosira* (stramenopile) *OMPDC–OPRT* with fungal, alveolate, and α- and β-proteobacterial *OPRT* to the exclusion of kinetoplastid and cyanobacterial *OMPDC–OPRT* (100% bootstrap support in ML analysis, Fig. 1). Clear separation of stramenopile *OMPDC–OPRT* from kinetoplastid/cyanobacterial *OMPDC–OPRT* indicates that, despite the same order of the gene fusion, the *OPRT* domains of stramenopiles and kinetoplastid/cyanobacterial *OMPDC–OPRT* do not share a common ancestor.

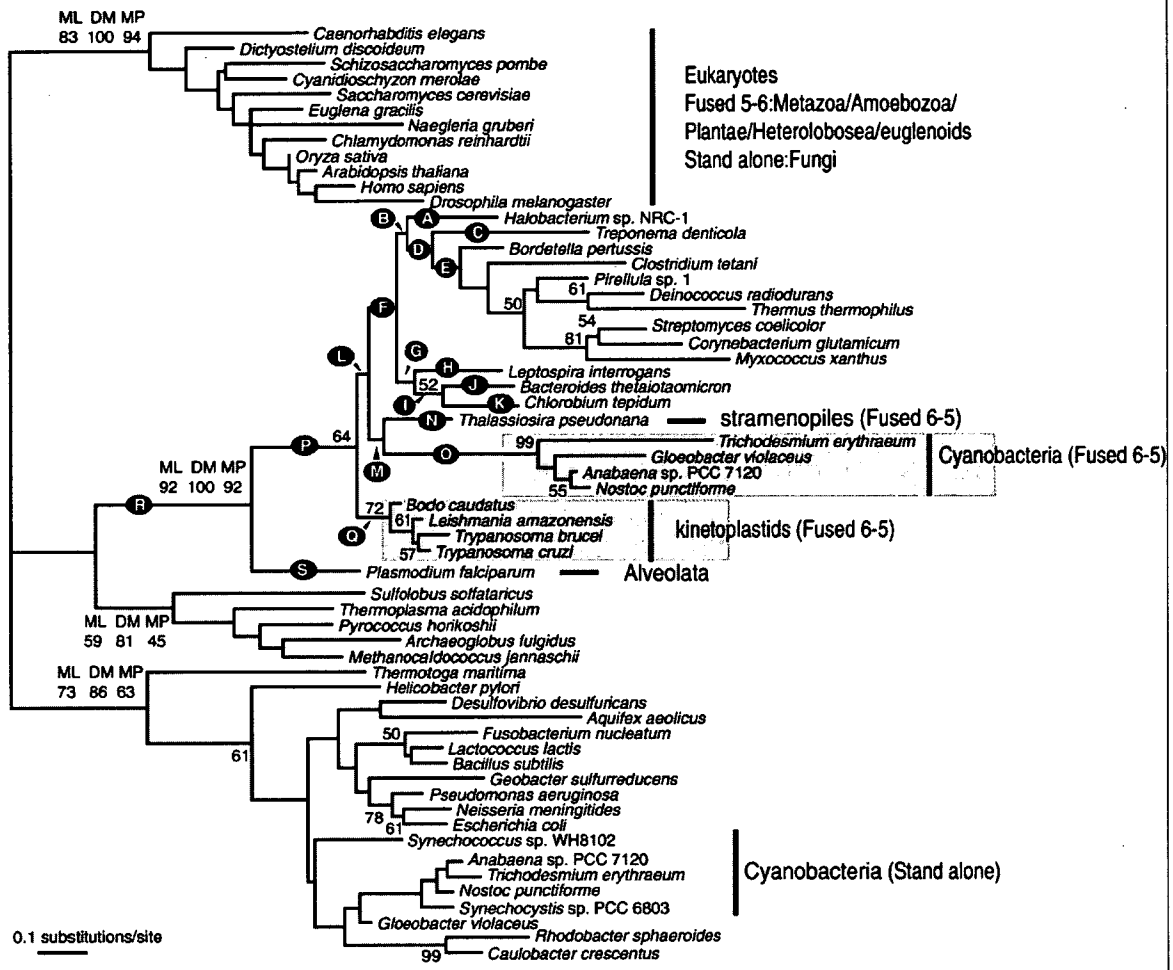
3.3. The phylogenetic relationship between kinetoplastid and cyanobacterial OMPDC–OPRT

In the OPRT tree, eukaryotic *OPRT–OMPDC* formed a moderately supported clade including chlamydial *OPRT* with 59% bootstrap support in the ML analysis (Fig. 1A). Monophyly of kinetoplastid and cyanobacterial *OMPDC–OPRT* was reconstructed, while its BP support was very weak (32% in the ML analysis, Fig. 1A). Kinetoplastid and cyanobacterial *OMPDC–OPRT* was located at the position of an immediate outgroup to the eukaryotic *OPRT–OMPDC* and chlamydial *OPRT* comprising a moderately supported clade (69% bootstrap support in the ML analysis, Fig. 1A). We referred to this clade as "Clade X" in the following sections.

Robustness of the positioning of kinetoplastid and cyanobacterial *OMPDC–OPRT* in the OPRT tree was further examined by the AU test. In the AU test, kinetoplastid, cyanobacterial, or their combined *OMPDC–OPRT* branch was grafted to all possible branches in the backbone tree (the OPRT tree, Fig. 1A) to know whether kinetoplastid and cyanobacterial *OMPDC–OPRT* were originated within Clade X or from other clade. Most of the alternative trees were statistically significantly rejected at a probability level of $p<0.05$ (Supplementary Fig. S1 and Table S5). Of 453 alternative trees examined, only

Fig. 1. Phylogenetic analyses of orotate phosphoribosyltransferase (OPRT). (A) The consensus maximum likelihood (ML) tree of OPRT inferred by the JTT model taking across-site-rate heterogeneity into consideration. The α-value of the Γ shape parameter used in the analysis was 1.2075. Bootstrap proportions (BPs) by the ML method are attached to the internal branches. Branches with less than 50% BP support are unmarked. For the six nodes of interest, BP values determined by the distance matrix (DM) and maximum parsimony (MP) methods are also shown. The length of each branch is proportional to the estimated number of substitutions. Ninety-three unambiguously aligned amino acid sites with 85 taxa were used for the analysis. Fused 5–6 and fused 6–5 indicate *OPRT–OMPDC* and *OMPDC–OPRT* fusions, respectively. Monophyletic clades of both kinetoplastid and cyanobacterial *OMPDC–OPRT* are shaded. "Clade X" represents a clade that includes eukaryotic *OPRT–OMPDC* and kinetoplastid and cyanobacterial *OMPDC–OPRT*. (B) The probability (p) value of alternative OPRT trees that are not rejected by the AU test at the significance level of $p<0.05$. The test was done by grafting kinetoplastid, cyanobacterial or combined kinetoplastid/cyanobacterial branch (Panel A, shaded) to all of the candidate branches on the backbone tree. "Branch position" corresponds to the branch labeled alphabetically in Panel A. The p value for the backbone tree is underlined. The detail results of the AU test are given in the Supplementary Fig. S1 and Supplementary Table S5.

(A) The OMPDC tree



(B) The probability (p) values of alternative trees

	Branch position	A	B	C	D	E	F	G	H	I	J	K
Groups grafted	Kinetoplastids (K)	$p = 0.123$	0.151	0.194	0.261	0.158	0.790	0.087	0.063	0.039	0.041	0.029
	Cyanobacteria (C)	0.301	0.407	0.304	0.318	0.139	0.597	0.421	0.252	0.247	0.108	0.180
	K plus C	0.006	0.042	0.033	0.047	0.008	0.240	0.043	0.013	0.012	0.016	0.006

	Replaced positions	L	M	N	O	P	Q	R	S
Groups grafted	Kinetoplastids (K)	**0.860**	0.902	0.514	0.101	**0.860**	**0.860**	0.151	0.065
	Cyanobacteria (C)	0.231	**0.860**	**0.860**	**0.860**	0.101	0.097	0.019	0.019
	K plus C	0.097	0.101	0.101	0.101	0.097	0.097	0.019	0.019

Fig. 2. Phylogenetic analyses of orotidine-5′-monophosphate decarboxylase (OMPDC). (A) The best maximum likelihood (ML) tree of OMPDC inferred by the JTT model taking across-site-rate heterogeneity into consideration. The α-value of the Γ shape parameter used in the analysis was 0.8205. Seventy-seven unambiguously aligned amino acid sites with 59 taxa were used for the analysis. (B) The probability (p) value of alternative OMPDC trees that are not rejected by the AU test at the significance level of $p < 0.05$. The grafting analysis was done as in Fig. 1. The p value for the trees having the same topology as the backbone tree is underlined. The detail results of the AU test are given in the Supplementary Fig. S2 and Supplementary Table S6. Methods and labeling are as in Fig. 1.

16 trees were not rejected, 12 of which included the *OMPDC–OPRT* branch within Clade X (position no. A–P, Fig. 1B). The other four alternative trees included cyanobacterial and the combined *OMPDC–OPRT* branches outside Clade X, but these *OMPDC–OPRT* clades were found to branch immediately to Clade X (Fig. 1B).

Cyanobacterial stand-alone *OPRT* genes are nested within the bacterial clade (Fig. 1A). Monophyly of the two versions of cyanobacterial sequences was assessed by the AU test, in which the OMPDC–OPRT-type sequences were grafted from the original positions (see Supplementary Fig. S1 and Table S5). The alternative trees that showed their monophyly were rejected

significantly by the AU test ($p<0.01$), indicating that the two versions of cyanobacterial *OPRT* have a different origin.

3.4. The OMPDC domain of OMPDC–OPRT has a lateral gene transfer-origin

The *OMPDC* tree reconstructed monophyly of eukaryotic *OPRT–OMPDC* (83% in ML, Fig. 2A). In conjunction with the similar results in the *OPRT* phylogeny, it was suggested that the *OPRT–OMPDC* fusions in these eukaryotic groups have the same origin. Fungi possess separate *OMPDC* but were positioned within this clade as expected from the results of previous analyses, suggesting again a secondary split of *OPRT–OMPDC* in Fungi (Nara et al., 2000).

Plasmodium (Alveolata) *OMPDC* and *Thalassiosira* (stramenopile), kinetoplastid, and cyanobacterial *OMPDC–OPRT* were nested within the bacterial clade with strong bootstrap support (≥92% for all methods, Fig. 2A). These results imply that the *OMPDC* gene was transferred laterally to these organisms from bacteria, although identification of the donor lineage(s) is difficult. Clear BP support was obtained neither for a monophyletic origin of the *Plasmodium*, stramenopiles, kinetoplastid, and cyanobacterial sequences nor for the multiple LGT events in these groups.

To know whether the *OMPDC* domain of kinetoplastid and cyanobacterial *OMPDC–OPRT* have the same origin, we performed the AU test by grafting kinetoplastid, cyanobacterial, or their combined *OMPDC–OPRT* branch to all of the possible branches in the backbone tree. This grafting analysis did not reject statistically the alternative trees showing monophyly of kinetoplastid and cyanobacterial *OMPDC–OPRT* (branch positions O and Q, Fig. 2), thus leaving a possibility of the same origin of these *OMPDC–OPRT*.

Cyanobacterial *OMPDC–OPRT* and the stand-alone genes are separated in the OMPDC tree, again suggesting different origins between the two versions (Fig. 2A). Likewise, monophyly of the two versions of cyanobacterial sequences was rejected significantly by the AU test ($p<10^{-20}$, see Supplementary Fig. S2 and Supplementary Table S6). These results indicate that the two versions of cyanobacterial *OMPDC*, as well as those of *OPRT*, have a different origin.

4. Discussion

4.1. Acquisition of the OMPDC–OPRT gene fusion in a common ancestor of kinetoplastids

In the present study, we demonstrated that the bodonid, *B. caudatus*, and the euglenoid, *E. gracilis*, possess *OMPDC–OPRT* and *OPRT–OMPDC*, respectively. The shared possession of *OMPDC–OPRT* in kinetoplastids, including trypanosomatids and bodonids, indicates that a common ancestor of kinetoplastids acquired *OMPDC–OPRT* after separation from the line leading to euglenoids.

This particular gene fusion in kinetoplastids can be argued in relation to the localization of OMPDC–OPRT in glycosomes, containing enzymes for glycolysis, fatty acid β-oxidation, purine salvage, and *de novo* pyrimidine biosynthesis (Michels et al., 2000). Glycosomes are found in both trypanosomatids and bodonids and the C-terminal tripeptide SKL is the targeting signal of the glycosomal proteins (Sommer et al., 1992, Wiemer et al., 1995). The putative *B. caudatus* OMPDC–OPRT possesses SKL at its C-terminus and thus may be localized within the glycosomes. It should be noted that euglenoids lack glycosomes (Parsons et al., 2001). Thus, it is assumed that establishment of *OMPDC–OPRT* was related to the evolutionary history of glycosomes in kinetoplastids.

Euglenozoa includes another flagellate protistan group, the diplonemids. Phylogenetic analyses demonstrated that diplonemids and kinetoplastids constitute a monophyletic group to the exclusion of euglenoids (Simpson et al., 2002). Further phylogenetic and gene-order analyses of *OPRT* and *OMPDC* are necessary to pinpoint the timing of the acquisition of *OMPDC–OPRT* in Euglenozoa.

4.2. The independent OMPDC–OPRT fusion events in eukaryotes

Unfortunately, it was difficult to obtain clear resolution in either OPRT or OMPDC tree, and bootstrap supports within major groups were very weak. This was probably due to the lack of sufficient alignable positions (only 93 and 77 positions with 85 and 59 taxa in *OPRT* and *OMPDC*, respectively) to allow reconstruction of clearly resolved trees. However, the important nodes critical for inferring LGT events were resolved and were supported well by bootstrapping.

Phylogenetic reconstruction of OPRT indicated the different origin of the *OPRT* domain between kinetoplastid and *Thalassiosira OMPDC–OPRT*. This finding provides evidence of occurrence of independent gene fusion events for *OPRT* and *OMPDC* in eukaryotes. Very recently, we found the *OMPDC–OPRT* in the genome sequence database of an oomycete, *Phytophthora sojae*, which belongs to stramenopiles (http://genome.jgi-psf.org/sojae1/sojae1.home.html). Both *OPRT* and *OMPDC* domains of the *Phytophthora OMPDC–OPRT* clustered with that of *Thalassiosira* (data not shown), suggesting that a common ancestor of stramenopiles might have acquired the *OMPDC–OPRT* fusion.

Andersson and Roger (2002) reported multiple events of gene fusion in the small subunit of the glutamate synthase gene family in eukaryotes, in which a different set of bacterial operons carrying two genes may have been transferred and fused subsequently in distantly related eukaryotic groups. Several distantly related prokaryotes, including *Bacillus subtilis*, *Thermus thermophilus*, and *Aeropyrum pernix*, have been shown by the respective genome sequence projects to possess operons carrying *OMPDC* and *OPRT* genes, a gene structure similar to the *OMPDC–OPRT* fusion. In our analyses, none of the prokaryotic operons were shown to form a monophyletic grouping with the *OMPDC–OPRT* fusions, thus excluding the possibility that the *OMPDC–OPRT* fusions in eukaryotes originated from a bacterial operon.

The chlamydial *OPRT* were nested within a clade of eukaryotic *OPRT–OMPDC*. The inability of *de novo* pyrimidine synthesis in chlamydia has been suggested previously

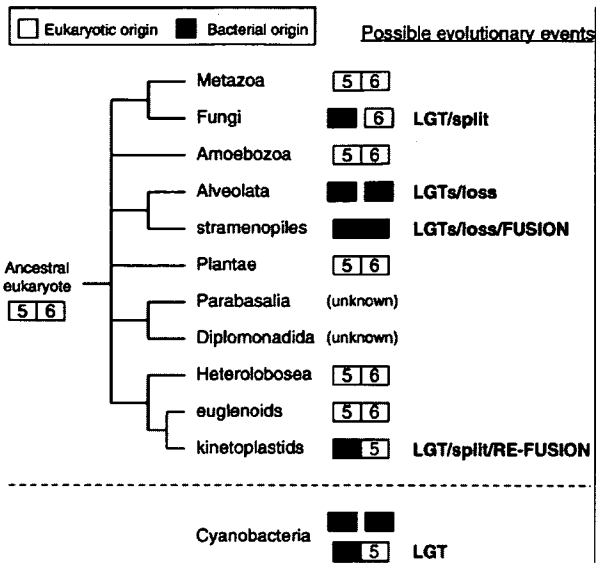


Fig. 3. Schematic representation of an evolutionary scenario of *OPRT* and *OMPDC* genes in eukaryotes. The deep-level structure of the tree of eukaryotes remains controversial, and the unresolved eukaryotic groups are shown in multifurcation (Simpson and Roger, 2004, Arisue et al., 2005). *OPRT* and *OMPDC* genes encode the fifth and sixth enzymes of the *de novo* pyrimidine biosynthetic pathway and are indicated as 5 and 6, respectively. Open and gray boxes represent *OPRT* (5) and *OMPDC* (6) genes of eukaryotic origin and bacterial origin, respectively. The presence and organization of *OPRT* and *OMPDC* genes in Parabasalia and Diplomonadida are unknown. The presumed evolutionary events, such as lateral gene transfer (LGT), secondary gene splitting, and subsequent gene fusion, are labeled in the corresponding groups. In this scenario, the cyanobacterial *OMPDC–OPRT* fusion was subsequently acquired *via* LGT as shown below the dashed line.

(McClarty and Qin, 1993) and we found no other pyrimidine-biosynthetic genes in the chlamydial genome databases. Therefore, the affinity of chlamydia to eukaryotes in the OPRT tree may suggest a eukaryote-to-prokaryote LGT event that may have occurred under parasitic conditions.

4.3. The LGT origin of OMPDC–OPRT in cyanobacteria

In the present study, a subset of cyanobacteria was found to possess not only separate *OPRT* and *OMPDC* genes but also an additional *OMPDC–OPRT*. In both OPRT and OMPDC phylogenies, cyanobacterial *OMPDC–OPRT* formed a strongly supported monophyletic cluster. In addition, the fused and separate versions of the genes of cyanobacteria were clearly distantly positioned in both OPRT and OMPDC trees. These observations were confirmed by the AU test that rejected statistically monophyly of the two versions of the cyanobacterial genes. These results indicate that cyanobacterial *OMPDC–OPRT* has a monophyletic origin and that this fused gene was not derived from fusion of the cyanobacterial stand-alone genes but was acquired *via* LGT.

Regarding to the origin of cyanobacterial *OMPDC–OPRT*, identification of the donor lineage is difficult. Our phylogenetic reconstruction of OPRT showed affinity of cyanobacterial *OMPDC–OPRT* not only with kinetoplastid *OMPDC–OPRT* but also with eukaryotic *OPRT–OMPDC* with moderate bootstrap support (69% in ML analysis, in Fig. 1A). In the grafting analysis of these *OMPDC–OPRT* branches, most of the alternative trees that were not rejected by the AU test showed the limited distributions of their branching to Clade X in the OPRT tree (Fig. 1B). It is noteworthy that no other prokaryotes were nested within Clade X with the exception of parasitic chlamydial species. Thus, the most likely explanation would be that cyanobacterial *OMPDC–OPRT* have a eukaryotic origin and were derived from kinetoplastid *OMPDC–OPRT via* LGT. Nevertheless, we cannot entirely exclude the possibility of the independent fusion event or the LGT event of the possible prokaryotic fusions in kinetoplastids and cyanobacteria.

4.4. An evolutionary scenario of OPRT and OMPDC genes in eukaryotes

An evolutionary scenario of *OPRT* and *OMPDC* genes in eukaryotes inferred from the present analyses is illustrated in Fig. 3. The deep-level structure of the tree of eukaryotes remains contentious, and the unresolved parts of the eukaryotic tree are shown in multifurcation (Simpson and Roger, 2004, Arisue et al., 2005). As the diverse eukaryotic groups, Metazoa, Amoebozoa, Plantae, Heterolobosea, and euglenoids, share *OPRT–OMPDC* and form monophyletic clusters in both OPRT and OMPDC trees, the most likely explanation is that the *OPRT–OMPDC* gene fusion was established once in the common ancestor of eukaryotes. Nevertheless, we cannot exclude the possibility that the *OPRT–OMPDC* fusion occurred independently or was transferred subsequently *via* LGT among distantly related eukaryotic groups.

The eukaryotic groups that do not possess *OPRT–OMPDC* are Fungi and Alveolata (stand-alone genes) and kinetoplastids and stramenopiles (*OMPDC–OPRT*). The results of the present study indicate that the *OPRT* genes in Fungi, Alveolata, and stramenopiles and the *OMPDC* genes in Alveolata, kinetoplastids, and stramenopiles were acquired from bacteria *via* LGT. The possible evolutionary scenario in these groups is as follows. In Fungi, *OPRT–OMPDC* was split and the bacterial-type *OPRT* gene was acquired instead of the split gene as suggested previously (Nara et al., 2000).

Kinetoplastids established *OMPDC–OPRT* through LGT-based acquisition of the *OMPDC* gene, splitting of the original *OPRT–OMPDC*, and re-fusion between the acquired *OMPDC* and the resident *OPRT* genes in the reverse order. This scenario is based on the assumption that the *OPRT* domain of kinetoplastid *OMPDC–OPRT* has a eukaryotic origin. Regarding to cyanobacterial *OMPDC–OPRT*, we speculated that the kinetoplastid *OMPDC–OPRT* was transferred subsequently to cyanobacteria.

Alveolata and stramenopiles thoroughly lost the original *OPRT–OMPDC* and acquired the laterally transferred *OPRT* and *OMPDC* genes. Due to the difficulties in reconstructing clear monophyly of Alveolata and stramenopiles and in establishing the evolutionary origin of the transferred genes in both OPRT and OMPDC trees, it is unclear both how a common ancestor of Alveolata and stramenopiles acquired *OPRT* and *OMPDC* genes and when the *OMPDC–OPRT* gene fusion emerged.

4.5. Evolutionary implications of lateral gene transfer in establishing gene fusions

To our knowledge, this is the first well-documented case of multiple, independent events of gene fusion in eukaryotes. Our findings counter the assumption that correctly ordered re-fusion, as well as secondary gene split, are improbable in eukaryotes (Stechmann and Cavalier-Smith, 2002, 2003), and thus lead to the new concept that multiple events of gene fusion and fission are possible. It is important to note that, without exception, all of the gene fusion/fission events accompany LGT events. This implies that LGT has made a marked contribution to the establishment of the rearranged structure of *OPRT* and *OMPDC* genes.

In the present study, we could not examine *OPRT* and *OMPDC* in monophyletic supergroups, Parabasalia and Diplomonadida, because of the lack of the *de novo* pyrimidine biosynthetic pathway examined so far (Wang et al., 1983, Aldritt et al., 1985, Arisue et al., 2005). Detection of the *OPRT* and *OMPDC* genes in Parabasalia/Diplomonadida and further taxon sampling in diverse groups of eukaryotes should clarify the precise evolutionary history of the *de novo* pyrimidine biosynthetic pathway and the biological significance of their gene rearrangement in eukaryotes.

Acknowledgments

We thank Yuji Inagaki at the Center for Computational Sciences, University of Tsukuba, for helpful discussions and critical reading of the manuscript. This work was supported in part by Grants-in-Aid for Scientific Research (Nos.17370086, 17390123, 17590377, and 18890188) from the Ministry of Education, Science, Sports, Culture, and Technology of Japan, from Ohyama Health Foundation (to TN), and from Kampou Science Foundation (to TN). T.AN and T.AO were supported by a Grant-in-Aid for the 21st Century COE Research from the Ministry of Education, Science, Sports, Culture, and Technology of Japan.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.gene.2007.02.009.

References

Aldritt, S.M., Tien, P., Wang, C.C., 1985. Pyrimidine salvage in *Giardia lamblia*. J. Exp. Med. 161, 437–445.

Andersson, J.O., Roger, A.J., 2002. Evolutionary analyses of the small subunit of glutamate synthase: gene order conservation, gene fusions, and prokaryote-to-eukaryote lateral gene transfers. Eukaryot. Cell 1, 304–310.

Andersson, J.O., Sjögren, Å.M., Davis, L.A., Embley, T.M., Roger, A.J., 2003. Phylogenetic analyses of diplomonad genes reveal frequent lateral gene transfers affecting eukaryotes. Curr. Biol. 13, 94–104.

Andersson, J.O., Sarchfield, S.W., Roger, A.J., 2005. Gene transfers from nanoarchaeota to an ancestor of Diplomonads and Parabasalids. Mol. Biol. Evol. 22, 85–90.

Annoura, T., Nara, T., Makiuchi, T., Hashimoto, T., Aoki, T., 2005. The origin of dihydroorotate dehydrogenase genes of kinetoplastids, with special reference to their biological significance and adaptation to anaerobic, parasitic conditions. J. Mol. Evol. 60, 113–127.

Arisue, N., et al., 2002. Phylogenetic position of *Blastocystis hominis* and of stramenopiles inferred from multiple molecular sequence data. J. Eukaryot. Microbiol. 49, 42–53.

Arisue, N., Hasegawa, M., Hashimoto, T., 2005. Root of the Eukaryota tree as inferred from combined maximum likelihood analyses of multiple molecular sequence data. Mol. Biol. Evol. 22, 409–420.

Baldauf, S.L., Roger, A.J., Wenk-Siefert, I., Doolittle, W.F., 2000. A kingdom-level phylogeny of eukaryotes based on combined protein data. Science 290, 972–977.

Bapteste, E., et al., 2002. The analysis of 100 genes supports the grouping of three highly divergent amoebae: *Dictyostelium, Entamoeba*, and *Mastigamoeba*. Proc. Natl. Acad. Sci. U. S. A. 99, 1414–1419.

Bergthorsson, U., Adams, K.L., Thomason, B., Palmer, J.D., 2003. Widespread horizontal transfer of mitochondrial genes in flowering plants. Nature 424, 197–201.

Boucher, Y., Doolittle, W.F., 2000. The role of lateral gene transfer in the evolution of isoprenoid biosynthesis pathways. Mol. Microbiol. 37, 703–716.

de Koning, A.P., Brinkman, F.S., Jones, S.J., Keeling, P.J., 2000. Lateral gene transfer and metabolic adaptation in the human parasite *Trichomonas vaginalis*. Mol. Biol. Evol. 17, 1769–1773.

Fernandes, A.P., Nelson, K., Beverley, S.M., 1993. Evolution of nuclear ribosomal RNAs in kinetoplastid protozoa: perspectives on the age and origins of parasitism. Proc. Natl. Acad. Sci. U. S. A. 90, 11608–11612.

Field, J., Rosenthal, B., Samuelson, J., 2000. Early lateral transfer of genes encoding malic enzyme, acetyl-CoA synthetase and alcohol dehydrogenases from anaerobic prokaryotes to *Entamoeba histolytica*. Mol. Microbiol. 38, 446–455.

Gao, G., Nara, T., Nakajima-Shimada, J., Aoki, T., 1999. Novel organization and sequences of five genes encoding all six enzymes for de novo pyrimidine biosynthesis in *Trypanosoma cruzi*. J. Mol. Biol. 285, 149–161.

Henze, K., et al., 2001. Unique phylogenetic relationships of glucokinase and glucosephosphate isomerase of the amitochondriate eukaryotes *Giardia intestinalis, Spironucleus barkhanus* and *Trichomonas vaginalis*. Gene 281, 123–131.

Hughey, R., Krogh, A., 1996. Hidden Markov models for sequence analysis: extension and analysis of the basic method. Comput. Appl. Biosci. 12, 95–107.

Keeling, P.J., Palmer, J.D., 2001. Lateral transfer at the gene and subgenic levels in the evolution of eukaryotic enolase. Proc. Natl. Acad. Sci. U. S. A. 98, 10745–10750.

McClarty, G., Qin, B., 1993. Pyrimidine metabolism by intracellular *Chlamydia psittaci*. J. Bacteriol. 175, 4652–4661.

Michels, P.A.M., Hannaert, V., Bringaud, F., 2000. Metabolic aspects of glycosomes in trypanosomatidae — new data and views. Parasitol. Today 16, 482–489.

Nara, T., Hashimoto, T., Aoki, T., 2000. Evolutionary implications of the mosaic pyrimidine–biosynthetic pathway in eukaryotes. Gene 257, 209–222.

Nixon, J.E., et al., 2002. Evidence for lateral transfer of genes encoding ferredoxins, nitroreductases, NADH oxidase, and alcohol dehydrogenase 3 from anaerobic prokaryotes to *Giardia lamblia* and *Entamoeba histolytica*. Eukaryot. Cell 1, 181–190.

Parsons, M., Furuya, T., Pal, S., Kessler, P., 2001. Biogenesis and function of peroxisomes and glycosomes. Mol. Biochem. Parasitol. 115, 19–28.

Shimodaira, H., Hasegawa, M., 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. Bioinformatics 17, 1246–1247.

Shimodaira, H., 2002. An approximately unbiased test of phylogenetic tree selection. Syst. Biol. 51, 492–508.

Simpson, A.G., Lukes, J., Roger, A.J., 2002. The evolutionary history of kinetoplastids and their kinetoplasts. Mol. Biol. Evol. 19, 2071–2083.

Simpson, A.G., Roger, A.J., 2004. The real 'kingdoms' of eukaryotes. Curr. Biol. 14, R693–R696.

Sommer, J.M., Cheng, Q.L., Keller, G.A., Wang, C.C., 1992. In vivo import of firefly luciferase into the glycosomes of *Trypanosoma brucei* and mutational analysis of the C-terminal targeting signal. Mol. Biol. Cell 3, 749–759.

Stechmann, A., Cavalier-Smith, T., 2002. Rooting the eukaryote tree by using a derived gene fusion. Science 297, 89–91.

Stechmann, A., Cavalier-Smith, T., 2003. The root of the eukaryote tree pinpointed. Curr. Biol. 13, R665–R666.

Stover, N.A., Cavalcanti, A.R., Li, A.J., Richardson, B.C., Landweber, L.F., 2005. Reciprocal fusions of two genes in the formaldehyde detoxification pathway in ciliates and diatoms. Mol. Biol. Evol. 22, 1539–1542.

Striepen, B., et al., 2002. Genetic complementation in apicomplexan parasites. Proc. Natl. Acad. Sci. U. S. A. 99, 6304–6309.

Suetsugu, N., Mittmann, F., Wagner, G., Hughs, J., Wada, M., 2005. A chimeric photoreceptor gene, *NEOCHROME*, has arisen twice during plant evolution. Proc. Natl. Acad. Sci. U. S. A. 102, 13705–13709.

Wang, C.C., Verham, R., Tzeng, S.F., Aldritt, S., Cheng, H.W., 1983. Pyrimidine metabolism in *Tritrichomonas foetus*. Proc. Natl. Acad. Sci. U. S. A. 80, 2564–2568.

Wiemer, E.A., Hannaert, V., van den, I.P.R., Van Roy, J., Opperdoes, F.R., Michels, P.A., 1995. Molecular analysis of glyceraldehyde-3-phosphate dehydrogenase in *Trypanoplasma borelli*: an evolutionary scenario of subcellular compartmentation in kinetoplastida. J. Mol. Evol. 40, 443–454.

Yang, Z., 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. Comput. Appl. Biosci. 13, 555–556.

Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Biochemical and Biophysical Research Communications 358 (2007) 253–258
BBRC
www.elsevier.com/locate/ybbrc

Dihydroorotate dehydrogenase arises from novel fused gene product with aspartate carbamoyltransferase in *Bodo saliens*

Takeshi Annoura [a], Idalia Sariego [a], Takeshi Nara [a], Takashi Makiuchi [a], Tsutomu Fujimura [b], Hikari Taka [b], Reiko Mineki [b], Kimie Murayama [b], Takashi Aoki [a,*]

[a] *Department of Molecular and Cellular Parasitology, Juntendo University School of Medicine, 2-1-1 Hongo, Bunkyo-ku, Tokyo 113-8421, Japan*
[b] *Division of Proteomics and Biomolecular Science, Biomedical Research Center, Juntendo University Graduate School of Medicine, 2-1-1 Hongo, Bunkyo-ku, Tokyo 113-8421, Japan*

Received 12 April 2007
Available online 24 April 2007

Abstract

The *ACT–DHOD* gene in the kinetoplastid *Bodo saliens* encodes aspartate carbamoyltransferase and dihydroorotate dehydrogenase, the second and fourth enzymes of pyrimidine biosynthesis. Although the single mRNA species yielded a 70-kDa ACT–DHOD protein, Western blotting with anti-DHOD-peptide antibody showed a major band of 35-kDa and minor bands. In-gel digestion and liquid chromatography–tandem mass (MS/MS) spectrometry showed that the 35-kDa band contained DHOD-specific polypeptides and an ACT-specific polypeptide, suggesting the occurrence of independent DHOD and ACT. Immunoprecipitation and MS/MS analysis identified a 70-kDa ACT–DHOD and a 35-kDa DHOD independently, and the N-terminal amino acid of 35-kDa DHOD was blocked. In vitro processing assay showed that recombinant ACT–DHOD was decreased by the *B. saliens* lysate, accompanying the appearance of 35-kDa DHOD and 35-kDa ACT. These results indicate that fused ACT–DHOD is the precursor to mature DHOD. Large amount of 35-kDa DHOD in *B. saliens* is discussed from a viewpoint of its physiological roles.

Keywords: Post-translational processing; Aspartate carbamoyltransferase; Dihydroorotate dehydrogenase; Fumarate reductase; Kinetoplastid; *Trypanosoma*; Bodonid; *Bodo saliens*; In vitro translation; Liquid chromatography–tandem mass spectrometry

The trypanosomatids and bodonids belong to the kinetoplastids, some of which cause human diseases such as Chagas' disease (*Trypanosoma cruzi*), African sleeping sickness (*T. brucei* group), and leishmaniasis (*Leishmania* spp.). Most bodonids are free-living organisms, while there is no free-living trypanosomatid reported so far. Phylogenetic studies revealed that trypanosomatids and bodonids may have evolved from a common euglenozoan ancestor [1,2]. Such a closer relationship suggested that a free-living bodonid could be used as a good model for studying an initial stage of establishing parasitism.

The de novo pyrimidine biosynthetic pathway consists of six enzymes for the production of uridine 5′-monophosphate (UMP). Previously, we reported a unique pyrimidine biosynthetic gene cluster that contains all the six genes [3], as a polycistronic transcription unit, in *Trypanosoma* and *Leishmania*. Dihydroorotate dehydrogenase (DHOD) is the fourth enzyme in the pathway catalyzing the oxidation of dihydroorotate to orotate. In these trypanosomatids, DHOD is classified as a family 1A enzyme, localized in the cytosol, and serves as a major soluble fumarate reductase [4]. The *DHOD* gene lies upstream to aspartate carbamoyltransferase (*ACT*) gene with a non-cording region

Abbreviations: ACT, aspartate carbamoyltransferase; DHOD, dihydroorotate dehydrogenase; UMP, uridine 5′-monophosphate; CAD, carbamoyl phosphate synthetase II–aspartate carbamoyltransferase–dihydroorotase; FMN, flavin mononucleotide; FRD, fumarate reductase; DIG, digoxigenin; ESI-Q-TOF, electrospray ionization-quadrupole-time-of-flight; LC–MS/MS, liquid chromatography–tandem mass spectrometry; CNBr, cyanogen bromide.

* Corresponding author. Fax: +81 3 5800 0476.
E-mail address: tksaoki@med.juntendo.ac.jp (T. Aoki).

0006-291X/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2007.04.102

between them within the pyrimidine biosynthetic gene cluster in *T. cruzi* [5–7]. The DHOD is essential for fumarate reductase activity in *T. cruzi* and may be important in its adaptation to anaerobic conditions [8]. Thus, *T. cruzi* DHOD would function in pyrimidine biosynthesis and in fumarate redox homeostasis.

In general, eukaryotic enzymes catalyzing continuous or coupled reactions occur as multifunctional proteins, including dihydrofolate reductase (DHFR) and thymidylate synthase (TS) (DHFR–TS) [9,10]; carbamoyl phosphate synthetase II (CPS II), ACT, and dihydroorotase (DHO) (CAD) [11,12]; and orotate phosphoribosyltransferase (OPRT) and orotidine-5′-monophosphate decarboxylase (OMPDC) (OPRT–OMPDC) [13,14]. These proteins should have the advantage of channeling substrates and products in continuous enzymatic reactions. Exceptionally, however, we found a novel fused gene *ACT–DHOD* in *Bodo saliens* that may encode a multifunctional protein, ACT and DHOD (ACT–DHOD) [8], the second and fourth enzymes of de novo pyrimidine biosynthesis, lacking the third enzyme.

In the present study, we show that the *B. saliens ACT–DHOD* gene is transcribed to *ACT–DHOD* mRNA, translated to the single protein, ACT–DHOD, and finally converted to mature independent DHOD. The physiological roles of the mature enzyme that resembles the trypanosomatid DHOD are also discussed.

Materials and methods

Materials. Bodo saliens (ATCC 50358) was cultured at 25 °C in 500-cm^2 flasks (No. 132867, Nalge Nunc International, Denmark), each containing 250 ml of artificial seawater. They were fed with *Klebsiella pneumoniae* subsp. (ATCC 27889) every other day, and charged with mixed gas (5% O_2 and 5% CO_2 in N_2) in an anaerobic chamber every day. They were harvested and washed as described [8].

Northern blotting. Total RNA was prepared from *B. saliens* using TRIZOL reagent (Invitrogen, San Diego, CA) according to the manufacturer's protocol. Aliquots (10 μg RNA) were electrophoresed on 1% agarose/2% formaldehyde gels, blotted to nylon membranes (Roche, Mannheim, Germany), prehybridized for 2 h at 42 °C in DIG Easy Hyb solution (Roche), and hybridized for 16 h in the same solution containing DIG-labeled DNA probes (10 ng/ml) corresponding to full-length *ACT–DHOD* (nucleotides 1–1944), *ACT* (nucleotides 1–933), and *DHOD* (nucleotides 1000–1944) [3,8,15]. Membranes were washed and signal was detected using CSPD chemiluminescence detection system (Roche) [15].

Antigens and antibodies. We synthesized polypeptides, N-CPLPRNEE LSTDVDGDRR-C and N-KSCTAQQRDGNPAPR-C, specific for the *B. saliens* ACT and DHOD domains, respectively, which had been cysteine cross-linked to keyhole limpet hemocyanin. *B. saliens* recombinant DHOD (V334-E648) gene was amplified by PCR using the primers, 5′-CA CCGTGGACCTGAGCGTGAGC-3′ (sense) and 5′-TTACTCGATGAC CTTGAGCTT-3′ (antisense), and KOD-Plus-DNA polymerase. The PCR product was inserted in Champion pET100 Directional TOPO Expression system (Invitrogen). The recombinant DHOD was produced and affinity-purified as described [5]. Female Japanese white rabbits (for anti-peptide antibody) or female BALB/c mice (for anti-whole-DHOD antibody) were subcutaneously injected with individual antigens (200 μg peptide per rabbit or 100 μg DHOD per mouse) emulsified with Freund's complete adjuvant and boosted several times at 2-week intervals with 100 μg of corresponding antigens emulsified with Freund's incomplete adjuvant.

Immunoprecipitation and Western blotting. Immunoprecipitation was performed against the *B. saliens* cytosolic fraction (200 μg protein), prepared as described [8], using Protein G Magnetic Beads (New England Biolabs, Beverly, MA). Aliquots of the cytosolic fraction were mixed with equal volumes of SDS sample loading buffer, incubated at 95 °C for 3 min and applied to 4–20% gradient SDS–polyacrylamide gels. After the electrophoresis, the proteins were transferred to nitrocellulose membranes, blocked in 2% fetal calf serum (FCS) plus 0.05% Tween-PBS (PBST), incubated overnight at 4 °C with purified polyclonal antibody (1:1000 in blocking solution), and washed with PBST. The membranes were incubated with alkaline phosphatase-conjugated anti-rabbit or anti-mouse IgG antibody (1:3000 each, Bio-Rad Laboratories, CA) and washed with PBST, and signals were visualized with CSPD chemiluminescence detection system (Roche).

In-gel digestion and electrospray ionization-quadrupole-time-of-flight tandem mass spectrometry (ESI-Q-TOF MS/MS). In-gel digestion by trypsin of SDS–PAGE bands was performed as described [16]. Tryptic peptides were extracted, the solution was evaporated, and the peptides were dissolved in 10 μl 1% formic acid. Mass spectrometry was performed in AB-QSTAR pulsar *i* hybrid (Applied Biosystems, Framingham, CA) combined with a microliquid chromatograph (Mi 2002; Michrom Bioresources, Abum, CA) equipped with a 0.2-mm ID × 50 mm Magic C18 column. Amino acid sequences of tryptic peptides were examined by ESI-Q-TOF MS/MS using Mascot search engine (www.matrixscience.com/search_form_select.html) for all peptide mass mapping and MS/MS ion searching, against the proteome data base, NCBInr.

N-terminal sequence analysis. Electrophoresed proteins were transferred to nitrocellulose membranes and subjected to automated Edman degradation using a Hewlett Packard model G1005A protein sequencer

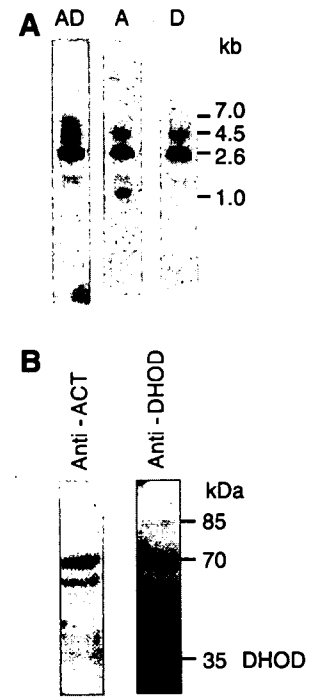


Fig. 1. Northern and Western blot analyses of *Bodo saliens* transcripts and proteins. (A) Northern blot analysis of *B. saliens* transcsripts. Aliquots of 10 μg total RNA were electrophoresed on 1% agarose gels, transferred to nylon membranes, and hybridized with DIG-labeled probes specific for *ACT* (lane A), *DHOD* (lane D), and full-length *ACT–DHOD* (lane AD). (B) Detection of 35-kDa DHOD and 35-kDa ACT proteins in *B. saliens* extract. Cytosolic fraction (50 μg protein) was resolved on SDS–PAGE and Western blotted with anti-ACT-peptide or anti-DHOD-peptide antibody. kDa, kilodaltons.

(Palo Alto, CA) attached to a HP1090 M liquid chromatograph as a PTH (phenylthiohydantoin) amino acid analyzer [16]. The same membrane was used for cyanogen bromide (CNBr)-digestion sequence analysis. The membrane was washed twice with 10% acetonitrile plus 0.1% TFA, and digestion was carried out in 100% formic acid containing 0.03 g/200 μl of CNBr at room temperature for 90 min. Samples were washed 3 times with distilled water, and amino acid sequences were determined as described.

In vitro expression and processing assay. The recombinant FLAG–ACT–DHOD–MYC was prepared in vitro using TNT SP6 High-Yield Protein Expression System (Promega). The DNA fragment carrying the *FLAG–ACT–DHOD–MYC* gene was amplified by PCR using the primers, 5′-GGGTGCGATCGCCATGGACTACAAAGACGATGACGACAA GATGCAGGGCACCCTCCAGGCGC-3′ (sense) and 5′-TTCTGTTT AAACCAGATCCTCTTCTGAGATGAGTTTTTGTTCTTGCTCGA TGACCTTGAGCTTGCC-3′ (antisense), and KOD-Plus-DNA polymerase. The PCR product was digested with *SgfI* and *PmeI*, and cloned into the protein-coding region (barnase site) of pF3A WG (BYDV) Flexi vector. The resulting recombinant plasmid, designated pF3A/FLAG–ACT–DHOD–MYC, was used for the in vitro cell-free expression system. Protein synthesis was initiated by adding the appropriate DNA template and allowed to proceed for 4 h at 25 °C in the presence or absence of flavin mononucleotide (FMN: 100 μM). The synthesized recombinant protein was incubated with *B. saliens* lysate, with or without a high concentration of protease inhibitor cocktail (Complete Mini Roche, with 1% Triton-X 100) at 25 °C, and aliquots (25 μg protein) were withdrawn at indicated times, followed by SDS–PAGE and Western blotting. FLAG-tagged proteins were detected by incubation with 1:2000 dilution of the primary FLAG M2 mouse monoclonal antibody (Sigma). MYC-tagged proteins were detected following the manufacturer's protocol (Sigma).

Results

Northern blot analysis of *B. saliens* total RNA, using DIG-labeled DNA probes specific for full-length *ACT–DHOD*, *ACT*, and *DHOD* showed that all three probes bound strongly to a 2.6-kb band, probably representing full-length *ACT–DHOD* mRNA (Fig. 1A). The open reading frame of the *ACT–DHOD* gene consists of 1944 bp [8], and thus, the other region, the sum of the spliced leader sequence and 3′ UTR may become approximately 650 bp. Prolonged development did not show any smaller bands, indicating the absence of independent *ACT* and *DHOD* transcripts. A quite faint band of 1.1-kb in lane A did not seem to be an independent *ACT* transcript, since the putative *ACT* gene (933 bp; [8]) plus above 650 bp should exhibit a 1.6-kb band. These results strongly suggested that there is neither independent *ACT* nor *DHOD* gene.

Western blotting of the *B. saliens* cytosolic fraction with anti-ACT-peptide antibody yielded a strong band at 70-kDa and faint bands at 60- and 35-kDa, whereas blotting with anti-DHOD-peptide antibody resulted in a strong band at 35-kDa and faint bands at 70- and 85-kDa (Fig. 1B). Extensive searches for ACT- or DHOD-specific tryptic peptides in the 60-, 70-, and 85-kDa bands by LC-MS/MS did not detect these polypeptides, which may have been due to the crude extract used in this experiment. Surprisingly, MS/MS analysis of tryptic peptides from the 35-kDa band detected a polypeptide, D592-K612, specific for the *B. saliens* DHOD domain (Table 1), as well as an ACT-specific polypeptide in the same 35-kDa band (Fig. 1B and Table 1). Although signal intensities of the spectra were weak, delta values (experimental minus calculated M_r) were 0.04 and 0.07 for the ACT- and DHOD-inner peptides, respectively (Table 1). These results suggest that *B. saliens* possesses independent ACT and DHOD proteins, each of molecular mass 35 kDa.

Immunoprecipitation of *B. saliens* DHOD protein with anti-DHOD-peptide antibody resulted in a 70-kDa band, visualized by both anti-ACT-peptide and anti-DHOD-peptide antibodies (Fig. 2A), likely representing full-length ACT–DHOD protein. MS/MS analysis of the 70-kDa protein showed seven polypeptides, six specific for ACT and one for DHOD (Table 1). This discrepancy in peptide number may have been due to the presence of many hydrophobic amino acids in ACT [8], resulting in their higher affinity binding to the liquid chromatography column. Typical MS/MS spectra of the DHOD- and ACT-specific polypeptides had higher intensity and lesser delta values of M_r

Table 1
Amino acid sequences of polypeptides, identified by liquid chromatography–tandem mass spectrometry, specific for *Bodo saliens* DHOD and ACT in the cytosolic fraction and in the immunoprecipitated 35-, 70-, and 85-kDa proteins

Fraction	Molecular mass	Enzyme	Sequence	M_r (expt)	M_r (calc)
Cytosolic	35-kDa	ACT	^{22}TILEDLTALALHLK^{35}	1549.87	1549.91
		DHOD	^{592}DAYQHLLAGASLVQVGTQLWK^{612}	2297.15	2297.22
Immunoprecipitated with anti-DHOD-peptide antibody	70-kDa	ACT	^{9}GAHIAGASQYNR^{21}	1243.63	1243.61
			^{22}TILEDLTALALHLK^{35}	1549.93	1549.91
			^{74}LGGSVVALPIEASSVSK^{90}	1612.89	1612.9
			^{99}TMDAYSDVIVLR^{111}	1381.74	1381.69
			^{161}TVVLVGDLK^{169}	942.59	942.57
			^{173}TVHSLAR^{179}	782.46	782.44
		DHOD	^{579}VIIGCGGVLCGR^{591}	1287.71	1287.68
Immunoprecipitated with anti-whole-DHOD antibody	35-kDa	DHOD	^{485}YLEAVTAVYPRPFGVK^{500}	1808.98	1808.98
			^{592}DAYQHLLAGASLVQVGTQLWK^{612}	2297.2	2297.22
			^{621}IRDELQAHLAR^{631}	1320.72	1320.73
	85-kDa	DHOD	^{579}VIIGCGGVLCGR^{591}	1287.74	1287.68

M_r (expt) and M_r (calc) denote the experimentally determined and calculated molecular masses, respectively. Delta values (experimental minus calculated) of less than 0.1 are considered highly specific.

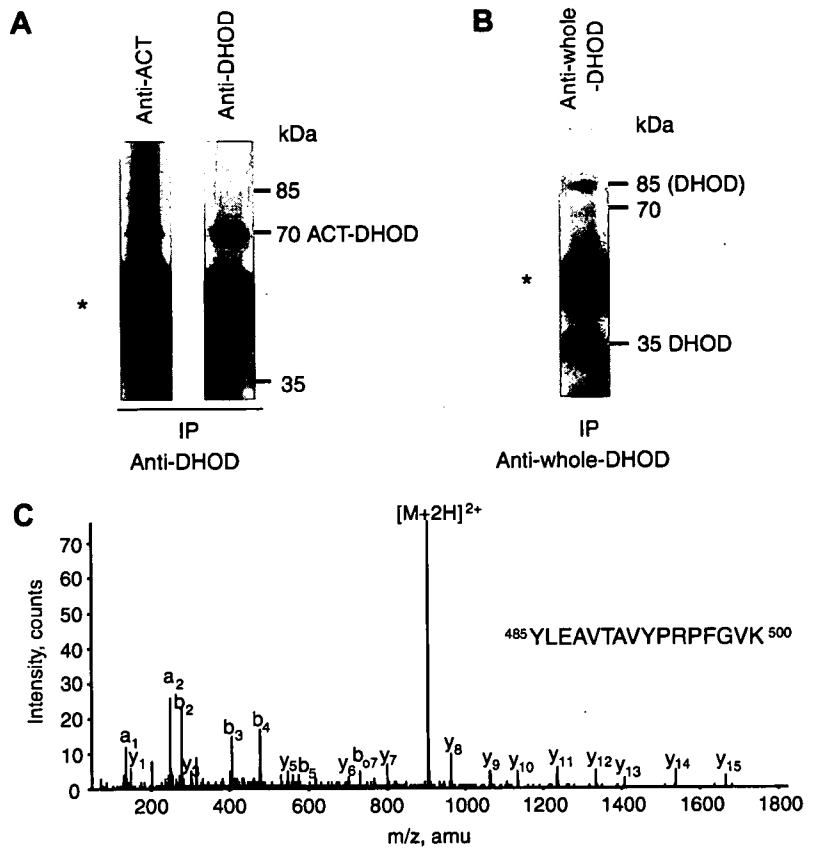


Fig. 2. Identification of the 70-kDa ACT–DHOD and the 35-kDa DHOD protein from *Bodo saliens*. (A) Immunoprecipitation of the 70-kDa ACT–DHOD protein from *B. saliens*. Cytosolic fraction (200 μg protein) was incubated with anti-DHOD-peptide antibody (5 μg) and the immunoprecipitated proteins were Western blotted using anti-ACT-peptide or anti-DHOD-peptide antibody. The asterisk, with a strong signal, was confirmed as the precipitated antibody used. (B) Immunoprecipitation of the 35-kDa DHOD protein was carried out using the anti-whole-DHOD antibody. The precipitated proteins were Western blotted using anti-whole-DHOD antibody. (C) Identification of a DHOD-specific amino acid sequence from the immunoprecipitated 35-kDa band (B), in MS/MS spectra of tryptic peptides Y485-K500.

(Table 1 and Supplemental Fig. 4S, A and B). These results indicate that the 2.6-kb *ACT–DHOD* mRNA (Fig. 1A) may be translated to a single, full-length ACT–DHOD protein of molecular mass 70-kDa. This 70-kDa primary translation product may be converted into 35-kDa DHOD in *B. saliens*.

We prepared a polyclonal antibody against the purified recombinant DHOD (V334–E648) to immunoprecipitate the 35-kDa DHOD. This antibody revealed two specific bands, at 35- and 85-kDa, and a faint band at 70-kDa (Fig. 2B). MS/MS analysis of the 35-kDa band demonstrated three tryptic polypeptides specific for DHOD with strong intensities and precise MS/MS spectra (Fig. 2C and Table 1), with one of these polypeptides, D592–K612, also seen in the cytosolic 35-kDa band (Table 1). In tryptic peptides from 85-kDa band (Fig. 2B), we detected only one MS/MS spectrum (^{579}VIIGCGGV LCGR^{591}) (Table 1 and Fig. 4S, C), the same polypeptide detected in the DHOD domain in immunoprecipitated 70-kDa ACT–DHOD (Table 1 and Fig. 4S, A). Thus, we expect that the 85-kDa protein may consist of 70-kDa ACT–DHOD protein and an additional protein (or polypeptide), of molecular mass 10- to 15-kDa. However, we could not exclude the possibility that the 85-kDa protein contains DHOD and a moiety of higher molecular mass ~50 kDa.

To determine the N-terminal sequence of 35-kDa DHOD, we analyzed proteins immunoprecipitated with anti-whole-DHOD antibody (Fig. 2B) using automated Edman degradation, but we detected no amino acid, indicating that the N-terminal amino acid of 35-kDa DHOD is blocked and that the 35-kDa protein is not an artificial degradation product of ACT–DHOD. Using the CNBr-digestion method, sequence analysis of the same membrane harboring 35-kDa band detected two independent DHOD-inner sequences (^{404}GLPNEGYEYY^{413} and ^{508}AHFD-QAAEVL^{517}). It is thus highly likely that the 35-kDa DHOD is a mature enzyme in vivo in *B. saliens*.

To confirm experimentally that ACT–DHOD is the precursor to mature DHOD, we prepared the recombinant ACT–DHOD, using an in vitro expression system in the presence and absence of FMN, the cofactor for DHOD. The FLAG- and MYC-tagged recombinant protein, FLAG–ACT–DHOD–MYC, was incubated with *B. saliens*