域をRT-PCRで増幅し、直接またはクローニング後、塩基配列およびアミノ酸配列を決定した。薬剤耐性変異の有無とその種類をデータベースと比較し決定した。また、末梢血リンパ球からゲノムDNAを抽出し、SBT (sequence based typing)法を用いてHLA型を決定した。

倫理面への配慮

本研究はケニア国の医学研究に関する研究審査委員会および倫理審査委員会と金沢大学医学部の倫理委員会の承認を受けている。

C. 研究結果

(1) 抗レトロウイルス療法(ART)有効率

追跡調査されている児100名中84名がART（第一次選択薬）を受けており、68名（81%）では治療開始24ヶ月の時点でVLが検出限界以下であった。尤も、その内10名ではVLが検出限界以下になるのに長期間（18ヶ月以上）を要した。

ART有効群では無効群に比べ、ベースライン CD4+T 細胞数が多い傾向が見られたが、ART開始年齢、ベースラインのVLとともにこれらの群間に有意差は認められなかった（表1）。

(2) 薬剤耐性変異の出現

治療中の84名中16名（19%）は治療開始後24ヶ月以内にVLのリバンウンド、または24ヶ月間VL無反応(治療無効)であったが、そのうち4名は治療開始前から薬剤耐性ウイルスを有していた。これらの4名を除外した12名中4名の児に感染しているHIV-1では、チミジン誘導体関連変異（TAM）が、また他の4名のウイルスではM184Vが最初のヌクレオシド系逆転者阻害剤（NRTI）耐性変異として検出された。残りの4名の児ではM184VとTAMが同時に検出された。NNRTI耐性変異では、K103NとG190Aが共通してみられた（表2）。

(3)HLAタイピング

追跡調査しているHIV感染児のうち、エイズ長期未発症群/遅発症群（30例：10歳までART不要の児）は、早期発症群（17例：5歳までにARTを開始した児）に比較し、HLAクラス1のA74をより高率に保有することが明らかとなった(表3)。また、エイズ遅発症群はその他の児に比べ、HLAクラス 1/B72 をより高率に保有する可能性が示唆された（表4)。

D. 考察

HIV母子感染児におけるART開始前のHIV-1薬剤耐性検査の重要性が明らかとなった。非Bサブタイプと Bサブタイプ HIV-1では、NRTI耐性変異の出現過程が少なくとも小児においては異なっていることが示唆された。

HLAクラス1のA74とB72がケニアの小児エイズ発症遅延に関与している可能性が示唆された。これらのHLA型は、従来、エイズ発症促進・遅延に関与するといわれている亜型ではない。今後、残りの症例の解析を行い、この結果を確認する予定である。」

E. 研究発表

1. 論文発表

(1) Tanimoto T, Ichimura H, et al. : Multiple routes of hepatitis C virus transmission among injection drug users in Hai Phong, Northern Vietnam. J Med Virol (in press).

(2) Phan TTC, Ichimura H, et al. : Characterization of HIV-1 Genotypes and Drug Resistance Mutations among Drug-Naïve HIV-1-Infected Patients in Northern Vietnam. AIDS Res Hum Retroviruses (in press).

(3) Agdamag DM, Ichimura H, et al. : Prediction of Response to Pegylated Interferon Treatment of Chronic Hepatitis B in the Philippines. J Med Virol 82(2): 213-219, 2010.

(4) Lhihana RW, Ichimura H, et al. : HIV-1 subtype diversity and drug resistance among

HIV-1-infected Kenyan patients initiating antiretroviral therapy. AIDS Res Hum Retroviruses 25(12):1211-1217, 2009.
(5) Kageyama S, Ichimura H, et al.: Tracking the entry routes of hepatitis C virus as a surrogate of HIV in an HIV-low prevalence country, the Philippines. J Med Virol 81(7):1157-62, 2009.
(6) Mwangi J, Ichimura H, et al.: Molecular Genetic Diversity of Hepatitis B Virus in Kenya. Intervirology 51(6):417-421, 2009.
(7) Hosaka N, Ichimura H, et al.: Rapid Detection of Human Immunodeficiency Virus Type 1 group M by a Reverse Transcription-Loop-Mediated Isothermal Amplification Assay. J Virol Methods 157(2):195-199, 2009.
(8) Miyashita M, Ichimura H, et al.: High-risk HPV types for uterine abnormal cervices of female commercial sex workers in the Philippines. J Med Virol 81(3):545-551, 2009.
(9) Ishizaki A, Ichimura H, et al.: Profile of HIV-1 infection and genotypic resistance mutations to antiretroviral drugs in treatment-naïve HIV-1–infected individuals in Hai Phong, Viet Nam. AIDS Res Hum Retroviruses 25(2):175-182, 2009.
(10) Lwembe R, Ichimura H, et al.: Changes in the HIV-1 envelope gene from non-subtype B HIV-1-infected children in Kenya. AIDS Res Hum Retroviruses 25(2):141-147, 2009.
(11) 市村宏：ヒト免疫不全ウイルス．臨床と微生物 37(2), 2010（発行予定）.

2. 学会発表

・HIV-1 subtype diversity and drug resistance among HIV-1-infected Kenyan patients initiating antiretroviral therapy. Lihana R, 市村宏、他．第 23 回日本エイズ学会学術集会．2009．11，名古屋．

・ Genotypic resistance mutations to antiretrovirals among treatment-naïve HIV-1-infected individuals in Northern Vietman. Pham TTC, 市村 宏、他．第 23 回日本エイズ学会学術集会．2009．11，名古屋．

・ Non-B Subtype HIV-1 Drug Resistance-associated Mutations among Vertically-infected Kenyan Children on ART. Ichimura H. et al., 5th IAS Conference on HIV Pathogenesis, Treatment and Prevention 19-22 July 2009, Cape Town, South Africa.

F. 知的財産権の出願・登録状況

特に予定なし。

Table 1. Comparison of demographic, virologic and immunologic characteristics of children with successful and failed antiretroviral therapy

Variable		First line ART	
		Treatment Success (n=68)	Treatment Failure (n=12)
Age at ART start	Mean (range)	7.7 (0.2 - 17 yrs)	7.2 (1-13 years)
Baseline Viral load (log_{10} copies/ml)	Mean (range)	5.03[a] (4.65 - 5.53)	5.14 [b] (4.28 – 6.01)
Baseline $CD4^+$ T-cell counts (cells/mm^3)	Median (range)	464 [a] (93 - 1442)	236 (6 - 979)

ART: anti-retroviral Therapy; [a] Data for 48 children. [b] Data for 11 children

Table 2. Characteristics of children who failed therapy and time (months) to emergence of first RTI resistance-associated mutations.

ID	Gender	Age at ART start (years)	Regimen	Duration to Emergence of mutations (months)	HIV-1 subtype	First Drug Resistance Mutations	
						NRTI	NNRTI
88	M	6	AZT/3TC/EFV	11.9	D	M184V	K103N
41	M	11	AZT/3TC/NVP	17	D	M184V	K103N
103	M	3	AZT/3TC/NVP	5.1	A1	M184V	K103N
48	M	12	AZT/3TC/NVP	11.6	D	M184V	K103N
79	M	5	ddI/3TC/EFV	9.1	A1	V75M, M184V	K101E, G190A
89	F	1	AZT/3TC/NVP	7.7	A1	M184V, L210W, T215K	K101E, G190A
85	F	5	AZT/3TC/NVP	8.6	CRF_02AG	D67N, M184V	K103N
36	M	8	AZT/3TC/EFV	5.7	A1	M184V, T215F	K103N
10	M	13	AZT/ddI/NVP	19.8	A1	M41L	NONE
38	M	8	ddI/3TC/EFV	6	C	D67N	NONE
62	M	6	AZT/ddI/NVP	7	A1	D67N, K70R	G190A
33	F	8	AZT/ddI/EFV	21	A1	K219Q, D218E	K101Q
Children with drug resistance mutations before ART: Retrospective drug resistance analysis							
44	M	7	AZT/ddI/EFV	-3.4	A1	NONE	K103N
56	M	9	AZT/3TC/NVP	-3.1	D	NONE	K103N
90	F	1	AZT/3TC/NVP	-0.2	D	NONE	Y181C
69	M	6	AZT/3TC/NVP	-0.1	A1	NONE	K103N

ART: Antiretroviral therapy. NRTI: Nucleoside reverse transcriptase inhibitor; DRM: Drug resistance mutations; AZT: zidovudine; 3TC: lamivudine; ddI: didanosine; NVP: nevirapine; EFV: efavirenz.

Table 3. HLA class 1A allele frequency among HIV-1-infected Kenyan children

HLA class 1 allele	RP (n=34)	SP (n=60)	Others (n=52)
A1	4 (11.8%)	3 (5.0%)	3 (5.8%)
A2	8 (23.5%)	12 (20.0%)	13 (25.0%)
A3	1 (2.9%)	3 (5.0%)	2 (3.8%)
A11	0	0	1 (1.9%)
A23	2 (5.9%)	8 (13.3%)	1 (1.9%)
A24	0	2 (3.3%)	1 (1.9%)
A26	0	2 (3.3%)	1 (1.9%)
A29	4 (11.8%)	3 (5.0%)	5 (19.2%)
A30	9 (26.5%)	6 (10.0%)	4 (7.7%)
A31	1 (2.9%)	0	0
A32	0	0	2 (3.8%)
A33	0	1 (1.7%)	1 (1.9%)
A34	1 (2.9%)	2 (3.3%)	3 (5.8%)
A36	0	2 (3.3%)	0
A66	1 (2.9%)	1 (1.7%)	1 (1.9%)
A68	3 (8.8%)	6 (10.0%)	10 (19.2%)
A74	**0**	**9 (15.0%)***	**4 (7.7%)**

SP: HIV-1 infected children who did not need to start ART by the age of ten years, RP: HIV-1 infected children who needed to start antiretroviral therapy by the age of five years.
* P<0.05 (SP vs. RP)

Table. 4. HLA class 1B allele frequency among HIV-1-infected Kenyan children

HLA class 1 allele	RP (n=34)	SP (n=56)	Others (n=52)
B7	2 (5.9%)	1 (1.8%)	4 (7.7%)
B8	2 (5.9%)	0	1 (1.9%)
B12	0	1 (1.8%)	0
B13	1 (2.9%)	1 (1.8%)	2 (3.8%)
B18	1 (2.9%)	0	2 (3.8%)
B35	1 (2.9%)	3 (9.4%)	4 (7.7%)
B38	0	1 (1.8%)	0
B39	1 (2.9%)	1	1 (1.9%)
B41	1 (2.9%)	2 (3.6%)	0
B42	4 (11.8%)	3 (5.4%)	2 (3.8%)
B44	1 (2.9%)	1 (1.8%)	2 (3.8%)
B45	3 (11.8%)	6 (10.7%)	6 (11.5%)
B47	1 (2.9%)		1 (1.9%)
B49	2 (5.9%)	4 (7.1%)	2 (3.8%)
B49/B15	1	0	0
B50	0	1 (1.8%)	0
B51	1 (2.9%)	0	2 (3.8%)
B53	2 (5.9%)	8 (14.3%)	5 (9.6%)
B57	1 (2.9%)	1 (1.8%)	2 (3.8%)
B58	5 (14.7%)	8 (14.3%)	7 (13.5%)
B60	0	0	1 (1.9%)
B60/B48	0	1 (1.8%)	0
B61	1 (2.9%)	0	0
B63	1 (2.9%)	0	1 (1.9%)
B64	0	1 (1.8%)	0
B65	0	0	1 (1.9%)
B39/B67	1 (2.9%)	0	0
B71	0	1 (1.8%)	2 (3.8%)
B72	**1 (2.9%)**	**10 (17.9%)***	**2 (3.8%)**
B75	0	0	1 (1.9%)
B81	0	1 (1.8%)	1 (1.9%)

SP: HIV-1 infected childrern who did not need to start ART by the age of ten years, RP: HIV-1 infected children who

別紙４

研究成果の刊行に関する一覧表

書籍

著者氏名	論文タイトル名	書籍全体の編集者名	書　籍　名	出版社名	出版地	出版年	ページ

雑誌

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Zong, L., Chen, Y., Peng, H., Gao, F., Iwamoto, A., and Gao, GF.	Rehsus macaque: a tight homodimeric CD8alpha/alpha	Proteins:Structure, Function and Bioinformatics	75	241-244	2009
Kondo, N., Ebihara, A., Ru, H., Kuramitsu, S., Iwamoto. A., Rao, Z., and Matsuda, Z.	Thermus thermophilus-derived protein tags that aid in preparation of insoluble viral proteins.	*Anal Biochem*	385	278-85	2009
Mizukoshi, F., Yamamoto, T., Mitsuki, Y., Terahara, K., Kawana-Tachikawwa, A., Kobayashi, K., Iwamoto, A., Morikawa, Y. Tsunetsugu-Yokota, Y.	Activation of HIV-1 Gag-specific CD8+ T cells by yeast-derived VLP-pulsed dendritic cells is influenced by the level of mannose on the VLP antigen.	Microbes and Infection	11	191-197	2009
Miyazaki, E., Kawana-Tachikawa, A., Tomizawa, M., Nunoya, J., Odawara, T., Fujii, T., Shi, Y., Gao, G.F., and Iwamoto, A	Highly restricted TCR repertoire in the CD8-positive T cell response against an HIV-1 epitope with a stereotypic amino acid substitution.	AIDS	23	651-660	2009
Nunoya, J., Nakashima, T., Kawana-Tachikawa, A., Kiyotani, K., Ito, Y., Sugimura, K., and Iwamoto, A.	Generation of recombinant monoclonal antibodies against an immunodominant HLA-A*2402-restricted HIV-1 CTL epitope.	AIDS Research and Human Retroviruses	25	897-904	2009

Wang, J., Kondo, N., Long, Y., Iwamoto, A., and Matsuda, Z.	Monitoring of HIV-1 envelope-mediated membrane fusion using modified split green fluorescent proteins.	J. Virol. Methods.	161	216-222	2009
Koga, M., Kawana-Tachikawa, A., Heckerman, D., Odawara, T., Nakamura, H., Koibuchi, T., Fujii, T., Miura, T., and Iwamoto, A.	Transition of impact of HLA class I allele expression on HIV-1 plasma virus loads at a population level over time.	Microbiol. Immunol.	in Press		2010
Zhu, D., Kawana-Tachikawa, A., Iwamoto, A., and Kitamura, Y.	Influence of polymorphism in dendritic cell-specific intercellular adhesion molecule-3-grabbing nonintegrin-related (DC-SIGNR) gene on HIV-1 trans-infection.	Biochem Biophys Res Commun	in Press		2010

別紙４

研究成果の刊行に関する一覧表

書籍

著者氏名	論文タイトル名	書籍全体の編集者名	書　籍　名	出版社名	出版地	出版年	ページ

雑誌

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Kono K, Bozek K, Domingues FS, Shioda T, Nakayama EE.	Impact of a single amino acid in the variable region 2 of the old world monkey TRIM5a SPRY (B30.2) domain on anti-human immunodeficiency virus type 2 activity.	Virology	388(1)	160-8	2009
Kuroishi A, Saito A, Shingai Y, Shioda T, Nomaguchi M, Adachi A, Akari H, Nakayama EE.	Modification of a loop sequence between alpha-helices 6 and 7 of virus capsid (CA) protein in a human immunodeficiency virus type 1 (HIV-1) derivative that has simian immunodeficiency virus (SIVmac239) vif and CA alpha-helices 4 and 5 loop improves replication in	Retrovirology	6	70	2009
Nakajima T, Nakayama EE, Kaur G, Terunuma H, Mimaya JI, Ohtani H, Mehra N, Shioda T, Kimura A.	Impact of novel TRIM5alpha variants, Gly110Arg and G176del, on the anti-HIV-1 activity and the susceptibility to HIV-1 infection.	AIDS	23(16)	2091-100	2009

Likanonsakul S, Rattanatham T, Feangvad S, Uttayamakul S, Prasithsirikul W, Tunthanathip P, Nakayama EE, Shioda T.	HLA-Cw*04 allele associated with nevirapine-induced rash in HIV-infected Thai patients.	AIDS Research and Therapy	6	22	2009
Onyango CO, Leligdowicz A, Yokoyama M, Sato H, Song H, Nakayama EE, Shioda T, Silva T, Townend J, Jaye A, Whittle H, Rowland-Jones S, Cotton M.	HIV-2 Capsids Distinguish High and Low Virus Load Patients in a West African Community Cohort.	Vaccine	in press		2010
Nakayama EE, Shioda T.	Anti-retroviral activity of TRIM5 (alpha).	Reviews in Medical Virology	in press		2010
Maegawa H, Miyamoto T, Sakuragi J, Shioda T, Nakayama EE.	Contribution of RING domain to retrovirus restriction by TRIM5α depends on combination of host and Virus.	Virology	in press		2010

別紙４

研究成果の刊行に関する一覧表

書籍

著者氏名	論文タイトル名	書籍全体の編集者名	書　籍　名	出版社名	出版地	出版年	ページ

雑誌

発表者氏名	論文タイトル	発表誌名	巻号	ページ	出版年
Yamada, Y., Ochiai, C., Yoshimura, K., Tanaka, T., Ohashi, N., Narumi T., Nomura, W., Harada, S., Matsushita, S., Tamamura, H.	CD4 mimics targeting the mechanism of HIV entry	Bioorganic & Medicinal Chemistry Letters	20	354–358	2010
Hatada, M., Yoshimura, K., Harada, S., Kawanami, Y., Shibata, J., Matsushita S.	HIV-1 evasion of a neutralizing anti-V3 antibody involves acquisition of a potential glycosylation site in V2.	J. Gen. Virol	in press		2010

別紙４

研究成果の刊行に関する一覧表

書籍

著者氏名	論文タイトル名	書籍全体の編集者名	書　籍　名	出版社名	出版地	出版年	ページ

雑誌

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
市村　宏	ヒト免疫不全ウイルス	臨床と微生物	37巻2号	3月25日発行予定	2010
Tanimoto T, Ichimura H, et al.	Multiple routes of hepatitis C virus transmission among injection drug users in Hai Phong, Northern Vietnam	J Med Virol	in press		
Phan TTC, Ichimura H, et al.	Characterization of HIV-1 Genotypes and Drug Resistance Mutations among Drug -Naïve HIV -1-Infected Patients in Northern Vietnam	AIDS Res Hum Retroviruses	in press		
Agdamag DM, Ichimura H, et al.	Prediction of Response to Pegylated Interferon Treatment of Chronic Hepatitis B in the Philippines	J Med Virol	2(2)	213-219	2010
Lhihana RW, Ichimura H, et al.	HIV-1 subtype diversity and drug resistance among HIV-1-infected Kenyan patients initiating antiretroviral therapy	AIDS Res Hum Retroviruses	25(12)	1211-1217	2009
Kageyama S, Ichimura H, et al.	Tracking the entry routes of hepatitis C virus as a surrogate of HIV in an HIV-low prevalence country, the Philippines	J Med Virol	81(7)	1157-1162	2009
Mwangi J, Ichimura H, et al.	Molecular Genetic Diversity of Hepatitis B Virus in Kenya	Intervirology	51(6)	417-421	2009

Hosaka N, Ichimura H, et al.	Rapid Detection of Human Immuno-deficiency Virus Type 1 group M by a Reverse Transcription-Loop-Mediated Isothermal Amplification Assay	J Virol Methods	157(2)	195-199	2009
Miyashita M, Ichimura H, et al.	High-risk HPV types for uterine abnormal cervices of female commercial sex workers in the Philippines	J Med Virol	81(3)	545-551	2009
Ishizaki A, Ichimura H, et al.	Profile of HIV-1 infection and genotypic resistance mutations to antiretroviral drugs in treatment-naïve HIV-1–infected individuals in Hai Phong, Viet Nam	AIDS Res Hum Retroviruses	25(2)	175-182	2009
Lwembe R, Ichimura H, et al.	Changes in the HIV-1 envelope gene from non-subtype B HIV-1-infected children in Kenya	AIDS Res Hum Retroviruses	25(2)	141-147	2009

PROTEINS
STRUCTURE • FUNCTION • BIOINFORMATICS

STRUCTURE NOTE

Rhesus macaque: A tight homodimeric CD8αα

Lili Zong,[1,2] Yong Chen,[1,3,4] Hao Peng,[1] Feng Gao,[1] Aikichi Iwamoto,[3,5,6] and George F. Gao[1,3]*

[1] CAS Key Laboratory of Pathogenic Microbiology and Immunology (CASPMI), Institute of Microbiology, Chinese Academy of Sciences (CAS), Beijing 100101, China

[2] Department of Obstetrics and Gynaecology, Zhujiang Hospital, Nanfang Medical University, Guangzhou 510280, China

[3] China-Japan Joint Laboratory of Molecular Immunology and Molecular Microbiology, Institute of Microbiology, Chinese Academy of Sciences (CAS), Beijing 100101, China

[4] College of Life Sciences, Graduate University, Chinese Academy of Sciences (GUCAS), Beijing 100049, China

[5] Research Center for Asian Infectious Diseases, The Institute of Medical Science, The University of Tokyo, Tokyo, Japan

[6] Division of Infectious Diseases, Advanced Clinical Research Center, The Institute of Medical Science, The University of Tokyo, Tokyo, Japan

Key words: rhesus macaque; crystal structure; MHC binding; CD8; dimer; HIV; vaccine.

INTRODUCTION

Simian immunodeficiency virus (SIV) infection of rhesus macaque (*macaca mulatta*) is widely used as an animal model for human immunodeficiency virus (HIV) infection[1–3] as well as other human diseases. It is known that the host cytotoxic T lymphocyte (CTL) responses provide powerful protection against HIV infection, and CTL-based immunization is currently believed to be the most promising approach toward vaccine development.[4]

As a coreceptor of T cell receptor (TCR) on the surface of CTLs, CD8 molecules stabilize the interaction of the TCR with major histocompatibility complex (MHC) by binding to the MHC class I (MHCI) molecule on the surface of antigen-presenting cells. In the absence of CD8 interaction, MHCI-restricted immune responses are hampered.[5] In addition, recent data indicate that CD8 has the ability to bind to a nonclassical MHC class I-like molecule, TL antigen, independently of TCR and CD3, expanding the function of CD8 further to an immunomodulator.[6–8] Moreover, soluble forms of CD8 can disrupt activation of some T cell clones with higher efficacy than anti-CD8 antibodies.[9,10]

In both human and mouse, the functions of CD8 involved in immune responses have been extensively studied.[11,12] The crystal structures of the human HLA-A*0201-CD8αα complex,[13,14] murine MHC H-2K^b-CD8αα complex,[15] TL antigen-CD8αα complex,[16] and the murine CD8αβ heterodimer[17] have been solved. For macaque monkeys, however, little is known on the structures of the CTL-related molecules (e.g., TCR, MHC, and CD8). Only the structure of MHC allele Mamu-A*01 has been recently solved in our laboratory.[18]

In this article, we present the crystal structure of rhesus macaque CD8αα (rCD8αα) homodimer and discuss the relatedness and uniqueness of rCD8αα structure with that of human/mouse CD8αα homodimer. Strikingly, with two Thr43 residues in C–C′ loop, rCD8αα shows a unique extra hydrogen bond in the homodimeric interface indicating a tighter homodimeric interaction.

Grant sponsor: Ministry of Science and Technology (MOST), China (Basic Research Program 973); Grant number: 2006CB504204; Grant sponsor: National Natural Science Foundation (NSFC), China; Grant number: 30671903; Grant sponsor: Chinese Academy of Sciences (CAS, Knowledge Innovation Project); Grant number: KSCX2-SW-227; Grant sponsor: NSFC; Grant number: 30525010; Grant sponsor: Postdoctoral Fund of China; Grant number: 20070410649; Grant sponsor: Japan Ministry of Education, Culture, Sports, Science and Technology (MEXT, The China-Japan Joint Laboratory of Molecular Immunology and Molecular Microbiology).

Lili Zong and Yong Chen contributed equally to this work.

*Correspondence to: George F. Gao, Center for Molecular Immunology, Institute of Microbiology, Chinese Academy of Sciences (CAS), Datun Road, ChaoYang District, Beijing 100101, China. E-mail: gaof@im.ac.cn

Received 28 September 2008; Revised 10 November 2008; Accepted 12 November 2008

Published online 19 November 2008 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.22331

© 2008 WILEY-LISS, INC.

MATERIALS AND METHODS

Expression and purification of rCD8αα homodimer

Rhesus macaque CD8 (rCD8) alpha chain nucleotides covering amino acids 1–120 of the ectodomain were synthesized based on the sequence of Indian origin rhesus (GeneBank ID: 698329). Inclusion bodies of rCD8α were prepared, and rCD8αα homodimer was renatured and purified by using the protocols described earlier.[13,14]

Crystallization, data collection, and processing

All crystallization attempts were performed at 18°C by the hanging drop vapor diffusion method. Ideal rCD8αα crystals grew from a 1:1 mixture of the protein solution (10 mg/mL) with crystallization reagent of 0.05*M* potassium phosphate monobasic, 20% w/v polyethylene glycol 8000. Data were collected using a Rigaku MicroMax007 rotating-anode X-ray generator (Cu Kα; λ = 1.5418 Å) equipped with an R-AXIS VII++ image-plate detector. Data were processed and scaled using HKL2000.[19]

Structure solution, refinement, and analysis

Data were analyzed by molecular replacement[20] using Molrep in the CCP4 package,[21] taking human CD8αα as the search probe (PDB code: 1AKJ).[14] Final rounds of refinement resulted in a final Rcryst of 21.3% (R_{free} = 25.7%) for all data between 35.0 and 2.20 Å.

Buried surface areas were calculated using SURFACE[21] with a 1.4 Å probe radius. The PyMOL Molecular Graphics System (DeLano Scientific, http://www.pymol.org) was used to prepare figures. Geometry of the refined structure was validated according to Ramachandran plot criteria.[22] The data collection and refinement statistics of the structure are shown in Table I.

Accession number

Atomic coordinates of rhesus macaque CD8αα homodimer have been deposited in the Protein Data Bank (PDB, http://www.rcsb.org/pdb) under accession code: 2Q3A.

Table I
X-Ray Diffraction Data Processing and Refinement Statistics

Data collecting	
Space group	$P2_12_12_1$
Unit cell dimensions (*a*, *b*, *c*)	46.54, 56.26, 82.31
Unit cell dimensions (α, β, γ)	90.00, 90.00, 90.00
Resolution range (Å)	40.00–2.20 (2.32–2.20)[a]
Total number of reflections	138,752
Number of unique reflections	17,807
Number of molecule in the asymmetric unit	2
Average redundancy	6.60 (5.90)
Completeness (%)	99.6 (100.0)
R_{merge} (%)	9.6 (29.5)
I/σ	12.8 (5.7)
Refinement	
Resolution (Å)	35.85–2.20
R-factor (%)	21.3
R_{free}[b] (%)	25.7
RMS deviations from restraint target values:	
Bond lengths (Å)	0.006
Bond angles (°)	1.33
Ramachandran plot Quality:	
Residues in most favored regions	165 [84.2%]
Residues in additional allowed	27 [13.8%]
Residues in generously allowed	4 [2.0%]
Residues in disallowed regions	0 [0%]

[a]Values in parentheses are given for the highest resolution shell.
[b]R_{free} is calculated over reflections in a test set (5%) not included in atomic refinement.

RESULTS AND DISCUSSION

Overall structure of rhesus macaque CD8αα homodimer

The crystals contained two CD8αα molecules as a dimer in a hand-shaking mode per crystallographic asymmetric unit. Belonging to the V set of Ig folds,[14] the overall structure comparison of rCD8αα homodimer with the human counterpart is shown in Figure 1(A). Each CD8α molecule is primarily composed of β structure arranged into two antiparallel β sheets. Short regions of 3_{10} helix are found between the E and F strands, which are not commonly found in CD8β molecules.

All residues corresponding to the HLA-A2-CD8αα interface remain the same in MHC Mamu-A*01 and rCD8αα.[18] The interaction of HLA-CD8 is mainly based on charge complementarity and exhibits relatively low affinity (KD = 100–223 μ*M*) and rapid kinetics.[8,23] The molecular surfaces of Mamu-A*01 and rCD8αα show similar complementarities (data not shown), indicating their interaction.

Structural comparison of rhesus macaque with human or murine CD8αα homodimer

Superposition of the final structure of rCD8αα dimer shows closer resemblance to human CD8αα homodimer (1AKJ, in complex with HLA-A2)[14] than to murine CD8αα (mCD8αα) (1BQH, in complex with Kd)[15] [Fig. 1(A,B)]. Root mean square deviation (RMSD) of rCD8αα and hCD8αα dimer results in 0.831 E for all Cα atoms, which is much smaller than that of rCD8αα with mCD8 dimer of 1.634 E.

The comparison of r/h/m CD8αα reveals some differences in loop regions, especially in complementarity

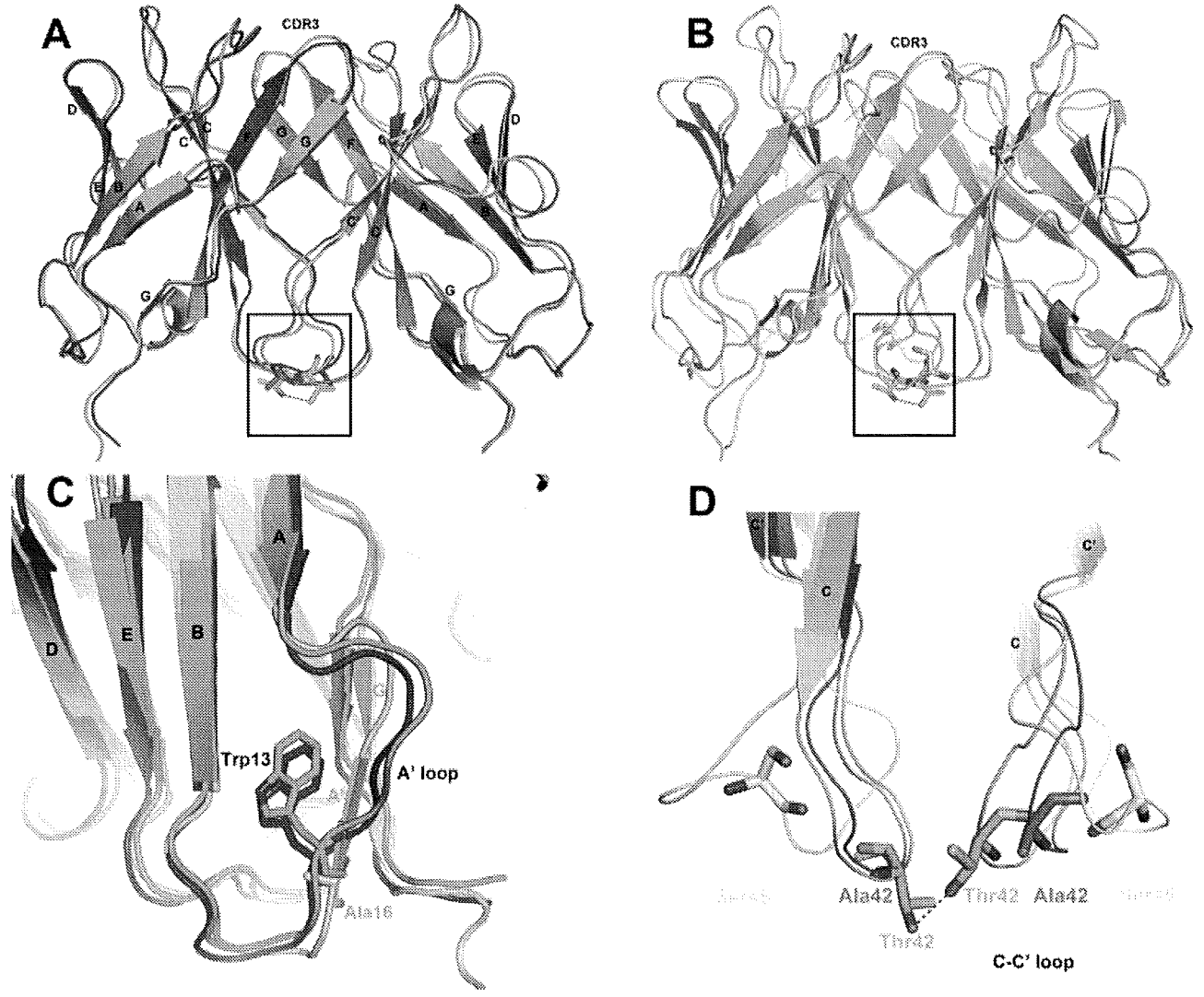

Figure 1

Crystal structure of rCD8αα homodimer and its superposition with human and mouse CD8αα structures. The ribbon diagram of each CD8αα was drawn and color-coded as: rCD8αα, green and cyan; hCD8αα, blue and magenta; mCD8αα, yellow and orange. (**A**) Superposition of rCD8αα and hCD8αα. The extra hydrogen bond region of rCD8αα was boxed. (**B**) Superposition of rCD8αα and mCD8αα. The extra hydrogen bond region of rCD8αα was boxed. (**C**) The transition of A'-β strand in mCD8αα and A' loop in both rCD8αα/hCD8αα as the residue changes (Ala16 to Trp13). (**D**) Thr43 residues in C-C' loop of rCD8αα homodimer (cyan) form extra hydrogen bonds. Hydrogen bond between two main-chains is shown as red dashed line.

determining regions (CDRs), which are involved in MHCIs binding. Interestingly, though CDR1 and CDR2 are structurally variable, CDR3 has almost the same conformation among r/h/m CD8αα molecules [Fig. 1(A,B)].

As one of the elements commonly found in many Ig domains, the first β strand of each domain is split into two shorter strands (A and A′).[20] rCD8αα homodimer can also be characterized by a *cis*-proline (Pro7) at the transition point at the A strand. However, at the place of the A′ strand, a loop is located in rhesus macaque as well as human CD8α molecules as the result of the big side chain of the residue Trp13, which is equivalent to Ala16 in mCD8α molecules [Fig. 1(C)].

Corresponding to Ala42 in hCD8α and Ser48 in mCD8α, Thr43 residues in rCD8α form an additional hydrogen bond to each other. This extra H-bond enhances the dimeric interaction and brings the C–C′ loop of the two molecules into a much closer position [Fig. 1(D)]. As a result, the interface size of the two rCD8α subunits (~2274 Å^2, total buried solvent-accessible surface area) is significantly larger than that of human CD8α interface (~2038 Å^2) indicating a tighter homodimeric interaction of CD8αα in rhesus macaque.

ACKNOWLEDGMENTS

The authors thank Dr. Fuliang Chu for his valuable suggestions for this project and Mr. Christopher Pannell for his critical reading of the manuscript.

REFERENCES

1. Allen TM, Sidney J, del Guercio MF, Glickman RL, Lensmeyer GL, Wiebe DA, DeMars R, Pauza CD, Johnson RP, Sette A, Watkins DI. Characterization of the peptide binding motif of a rhesus MHC class I molecule (Mamu-A*01) that binds an immunodominant CTL epitope from simian immunodeficiency virus. J Immunol 1998;160:6062–6071.
2. Matano T, Shibata R, Siemon C, Connors M, Lane HC, Martin MA. Administration of an anti-CD8 monoclonal antibody interferes with the clearance of chimeric simian/human immunodeficiency virus during primary infections of rhesus macaques. J Virol 1998;72:164–169.
3. Schmitz JE, Kuroda MJ, Santra S, Sasseville VG, Simon MA, Lifton MA, Racz P, Tenner-Racz K, Dalesandro M, Scallon BJ, Ghrayeb J, Forman MA, Montefiori DC, Rieber EP, Letvin NL, Reimann KA. Control of viremia in simian immunodeficiency virus infection by CD8+ lymphocytes. Science 1999;283:857–860.
4. Kaufmann SH, McMichael AJ. Annulling a dangerous liaison: vaccination strategies against AIDS and tuberculosis. Nat Med 2005;11: S33–S44.
5. Fung-Leung WP, Schilham MW, Rahemtulla A, Kundig TM, Vollenweider M, Potter J, van Ewijk W, Mak TW. CD8 is needed for development of cytotoxic T cells but not helper T cells. Cell 1991;65:443–449.
6. Cole DK, Gao GF. CD8: adhesion molecule, co-receptor and immuno-modulator. Cell Mol Immunol 2004;1:81–88.
7. Leishman AJ, Naidenko OV, Attinger A, Koning F, Lena CJ, Xiong Y, Chang HC, Reinherz E, Kronenberg M, Cheroutre H. T cell responses modulated through interaction between CD8αα and the nonclassical MHC class I molecule, TL. Science 2001;294:1936–1939.
8. Wyer JR, Willcox BE, Gao GF, Gerth UC, Davis SJ, Bell JI, van der Merwe PA, Jakobsen BK. T cell receptor and coreceptor CD8 αα bind peptide-MHC independently and with distinct kinetics. Immunity 1999;10:219–225.
9. Cole DK, Rizkallah PJ, Boulter JM, Sami M, Vuidepot AL, Glick M, Gao F, Bell JI, Jakobsen BK, Gao GF. Computational design and crystal structure of an enhanced affinity mutant human CD8 αα coreceptor. Proteins 2007;67:65–74.
10. Sewell AK, Gerth UC, Price DA, Purbhoo MA, Boulter JM, Gao GF, Bell JI, Phillips RE, Jakobsen BK. Antagonism of cytotoxic T-lymphocyte activation by soluble CD8. Nat Med 1999;5:399–404.
11. Gao GF, Jakobsen BK. Molecular interactions of coreceptor CD8 and MHC class I: the molecular basis for functional coordination with the T-cell receptor. Immunol Today 2000;21:630–636.
12. Gao GF, Rao Z, Bell JI. Molecular coordination of αβ T-cell receptors and coreceptors CD8 and CD4 in their recognition of peptide-MHC ligands. Trends Immunol 2002;23:408–413.
13. Gao GF, Gerth UC, Wyer JR, Willcox BE, O'Callaghan CA, Zhang Z, Jones EY, Bell JI, Jakobsen BK. Assembly and crystallization of the complex between the human T cell coreceptor CD8α homodimer and HLA-A2. Protein Sci 1998;7:1245–1249.
14. Gao GF, Tormo J, Gerth UC, Wyer JR, McMichael AJ, Stuart DI, Bell JI, Jones EY, Jakobsen BK. Crystal structure of the complex between human CD8α(α) and HLA-A2. Nature 1997;387:630–634.
15. Kern PS, Teng MK, Smolyar A, Liu JH, Liu J, Hussey RE, Spoerl R, Chang HC, Reinherz EL, Wang JH. Structural basis of CD8 coreceptor function revealed by crystallographic analysis of a murine CD8αα ectodomain fragment in complex with H-2Kb. Immunity 1998;9:519–530.
16. Liu Y, Xiong Y, Naidenko OV, Liu JH, Zhang R, Joachimiak A, Kronenberg M, Cheroutre H, Reinherz EL, Wang JH. The crystal structure of a TL/CD8αα complex at 2.1 A resolution: implications for modulation of T cell activation and memory. Immunity 2003;18:205–215.
17. Chang HC, Tan K, Ouyang J, Parisini E, Liu JH, Le Y, Wang X, Reinherz EL, Wang JH. Structural and mutational analyses of a CD8αβ heterodimer and comparison with the CD8αα homodimer. Immunity 2005;23:661–671.
18. Chu F, Lou Z, Chen YW, Liu Y, Gao B, Zong L, Khan AH, Bell JI, Rao Z, Gao GF. First glimpse of the peptide presentation by rhesus macaque MHC class I: crystal structures of Mamu-A*01 complexed with two immunogenic SIV epitopes and insights into CTL escape. J Immunol 2007;178:944–952.
19. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol 1997;276:307–326.
20. Chen Y, Chu F, Gao F, Zhou B, Gao GF. Stability engineering, biophysical, and biological characterization of the myeloid activating receptor immunoglobulin-like transcript 1 (ILT1/LIR-7/LILRA2). Protein Expr Purif 2007;56:253–260.
21. CCP4. The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr 1994;50:760–763.
22. Lovell SC, Davis IW, Arendall WB, III, de Bakker PI, Word JM, Prisant MG, Richardson JS, Richardson DC. Structure validation by Cα geometry: π,ψ and Cβ deviation. Proteins 2003;50:437–450.
23. Gao GF, Willcox BE, Wyer JR, Boulter JM, O'Callaghan CA, Maenaka K, Stuart DI, Jones EY, Van Der Merwe PA, Bell JI, Jakobsen BK. Classical and nonclassical class I major histocompatibility complex molecules exhibit subtle conformational differences that affect binding to CD8αα. J Biol Chem 2000;275:15232–15238.

Analytical Biochemistry 385 (2009) 278–285

Contents lists available at ScienceDirect

Analytical Biochemistry

journal homepage: www.elsevier.com/locate/yabio

Thermus thermophilus-derived protein tags that aid in preparation of insoluble viral proteins

Naoyuki Kondo [a,b], Akio Ebihara [c,1], Heng Ru [b], Seiki Kuramitsu [c,d], Aikichi Iwamoto [a,e], Zihe Rao [b,f], Zene Matsuda [a,b,*]

[a] *Research Center for Asian Infectious Diseases, Institute of Medical Science, University of Tokyo, 4-6-1, Shirokanedai, Minato-ku, Tokyo 108-8639, Japan*
[b] *China–Japan Joint Laboratory of Structural Virology and Immunology, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, People's Republic of China*
[c] *RIKEN SPring-8 Center, Harima Institute, Hyogo 679-5148, Japan*
[d] *Department of Biological Sciences, Graduate School of Science, Osaka University, Osaka 560-0043, Japan*
[e] *Division of Infectious Diseases, Advanced Clinical Research Center, Institute of Medical Science, University of Tokyo, Tokyo 108-8639, Japan*
[f] *National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, People's Republic of China*

ARTICLE INFO

Article history:
Received 1 September 2008
Available online 19 November 2008

Keywords:
Thermus thermophilus
HIV
Protein tag
Solubilization
Vpr
Membrane-spanning domain

ABSTRACT

The expression and solubilization of insoluble proteins have been facilitated by the introduction of protein tags. In our analyses of viral protein R (Vpr) of human immunodeficiency virus 1 (HIV-1), however, several conventional tag proteins enhanced its expression but failed to solubilize it. Therefore, we decided to explore whether proteins derived from *Thermus thermophilus* HB8 (*T. th.*), a highly heat-stable bacterium, could be used as tag proteins to enhance the solubilization of Vpr. Based on the data accumulated during the recent structural genomics project of *T. th.*, we selected 15 *T. th.* proteins with high expression levels and solubilities. From this group, we identified a *T. th.* tag protein that expressed Vpr in a soluble form. Furthermore, two *T. th.* tag proteins, including the identified one, were found to solubilize the extremely insoluble membrane-spanning domain of the envelope protein of HIV-1. When green fluorescent protein (GFP) was used as a passenger protein of *T. th.* tags, the brightness and stability of GFP were similar to those of untagged GFP, suggesting that the *T. th.* tags do not negatively affect the function of the passenger protein. Thus, data of structural genomics can be applied to generate a customized versatile protein tag for protein analyses.

© 2008 Elsevier Inc. All rights reserved.

The structures of various proteins have been determined during recent years, owing mainly to the progress of structural genomics (SG)[2] research. This information not only enriches our knowledge of certain aspects of basic protein science, such as protein-folding patterns and specific protein–protein interactions, but also provides practical contributions for structure-oriented drug design against pathogens [1–5]. The structures of some proteins remain unknown, however, because the proteins have unfavorable properties that hinder attempts to analyze their structures. Such properties include low expression level, low solubility, difficulty of purification, poor crystallizability, and instability of formed crystals. Adaptor proteins that interact with multiple other proteins, or membrane proteins with highly hydrophobic segments, often possess these properties.

The use of certain tag proteins can overcome some of these problems. A small peptide tag such as a polyhistidine tag, for example, can facilitate both the detection and purification of a target protein [6,7]. Some larger peptide tags (or protein tags) can likewise improve detection and purification while also enhancing the expression level and solubility of a protein of interest. Indeed, maltose binding protein (MBP) [8], glutathione *S*-transferase (GST) [9], thioredoxin (TRX) [10], and N utilization substance A (NusA) [11] have been widely used, and several expression vectors are commercially available. Several groups have compared the efficiency of conventional tag proteins for expression [12–15]. Although each of these studies used different expression vectors and passenger proteins, it seems reasonable to conclude that one should try several tag proteins for a particular passenger protein because the potency of a tag can be affected by multiple factors,

* Corresponding author. Address: Research Center for Asian Infectious Diseases, Institute of Medical Science, University of Tokyo, 4-6-1, Shirokanedai, Minato-ku, Tokyo 108-8639, Japan. Fax: +81 3 6409 2208.
E-mail address: zmatsuda@ims.u-tokyo.ac.jp (Z. Matsuda).
[1] Present address: Laboratory of Applied Biochemistry, Faculty of Applied Biological Sciences, Gifu University, Gifu 501-1193, Japan.

[2] *Abbreviations used:* SG, structural genomics; MBP, maltose binding protein; GST, glutathione *S*-transferase; TRX, thioredoxin; NusA, N utilization substance A; HIV, human immunodeficiency virus; SIV, simian immunodeficiency virus; AIDS, acquired immunodeficiency syndrome; Vpr, virus protein R; MSD, membrane-spanning domain; *T. th.*, *Thermus thermophilus;* GFP, green fluorescent protein; Vpx, viral protein X; PCR, polymerase chain reaction; LB, Luria–Bertani; IPTG, isopropylthiogalactoside; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; PBS, phosphate-buffered saline; GdmHCl, guanidium hydrochloride; CD, circular dichroism; UV, ultraviolet; FACS, fluorescence-activated cell sorting.

0003-2697/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.ab.2008.10.050

including the specific properties of the passenger proteins as well as the linker between the tag and passenger proteins.

In many cases, tag proteins serve as molecular chaperons during the synthesis and folding of the passenger protein [12], only to be removed after production and purification. For an extremely insoluble passenger protein, a tag protein is retained so as to facilitate solubilization. If a proper tag protein is selected, the passenger protein's function is well maintained even with a fused tag form [16]. Moreover, the remaining tag proteins, if their structures are known, can serve as models of molecular replacement for phase determination [16,17]. Lysozyme [18], MBP [19], and GST [20] have been successfully used for this purpose.

In our laboratory, we are studying the several proteins of primate lentiviruses, human immunodeficiency virus (HIV), and simian immunodeficiency virus (SIV) that are linked with acquired immunodeficiency syndrome (AIDS). One of them is virus protein R (Vpr), and another is gp41, a subunit of envelope protein. Vpr is a highly insoluble 14-kDa accessory gene product of HIV-1. It works as a pleiotropic adaptor protein in the life cycle of HIV-1 [21–26]. The gp41 subunit plays a critical role for membrane fusion and contains highly hydrophobic membrane-spanning domain (MSD). The membrane-spanning domain of gp41 (gp41-MSD) anchors the HIV-1 envelope protein to lipid bilayers [27]. During our study of these insoluble proteins, we tried to improve solubilization by using several conventional tag proteins but had limited success. Therefore, we tried to identify new tag proteins that would facilitate the production and structure determination of proteins of interest. We sought candidate tag proteins from *Thermus thermophilus* (*T. th.*), one of the well-characterized model organisms of the SG project [28,29]. Because *T. th.* survives at 85 °C and its proteins are heat-stable [30], the structures of more than 400 of approximately 2000 known gene products have been determined during the past 5 years [28]. Drawing on information that has accumulated from SG research on *T. th.* proteins, we tested proteins with high expression and solubility characteristics and known high-resolution three-dimensional structures for their utility as candidate tag proteins.

We tested 15 candidate tag proteins on Vpr of HIV-1 and identified a *T. th.* tag protein that could express a soluble form of Vpr. This tag and another tag were also able to help expression of a gp41-MSD. Analyses using green fluorescent protein (GFP) as a fusion passenger protein suggested that these *T. th.* tags did not significantly affect the biophysical properties of the passenger proteins. Our results suggest that the information accumulated during the SG project, as exemplified here for *T. th.*, can be a versatile resource for the identification of customized tag proteins that may facilitate biophysical and structural analyses of highly insoluble proteins.

Materials and methods

Selection of candidate tag proteins from T. th

We used the accumulated data from expression, purification, and crystallization trials in the Whole Cell Project of *T. th.* [29,31] to select candidate proteins. From the more than 400 proteins with known structures, we selected proteins that showed good expression and solubility. In the Whole Cell Project of *T. th.*, we arbitrarily defined seven classes to indicate a protein's expression level and solubility fitness: (i) no expression, (ii) low expression and low solubility, (iii) high expression and low solubility, (iv) high expression and half solubility, (v) low expression and high solubility, (vi) high expression and half solubility, and (vii) high expression and high solubility. We selected proteins that fell in the high expression and high solubility category. Among these selected proteins, we further narrowed the field by choosing candidates whose resolution of the determined structure is higher than 2 Å.

Construction of expression plasmids

All of the proteins were expressed in *Escherichia coli* BL21 Star (DE3) (Invitrogen) using the same plasmid derived from pET-47b (Novagen). pET-47b contains a T7 promoter, a six polyhistidine tag-coding sequence, and an HRV 3C protease recognition sequence upstream of the multiple cloning site. A tag protein was added at either the N or C terminus of a passenger protein (Fig. 1). To add the tag at the N terminus, we used a modified pET-47b plasmid called pET-47md that had been produced by cloning an oligonucleotide cassette with *Asc*I, *Pac*I, *Bam*HI, and *Bsr*GI sites, in that order, between the *Kpn*I and *Avr*II sites of pET-47b. To fuse a candidate tag protein at the C terminus of a passenger protein, pET-47b was used. The pET-47b vector has *Bam*HI, *Bsr*GI, *Asc*I, and *Pac*I sites, in that order, in the multiple cloning site.

The genes for tag proteins were amplified by polymerase chain reaction (PCR) using primers that contained *Asc*I and *Pac*I sites at the 5′ and 3′ termini of the genes while also using KOD Plus (Toyobo), Pfu Turbo (Stratagene), and Ex Taq (Takara) DNA polymerases. The template DNAs we used included 15 *T. th.* expression plasmids for *T. th.* genes [31,32] (RIKEN Bioresource Center DNA bank), pMALp2e (New England Biolabs) for *mbp*, pGEX5x-1 (GE Healthcare) for *gst*, pET-50b (Novagen) for *nusA*, pSP65HXB2gpt for *vpr*, and the synthetic *vpx* (viral protein X) gene whose sequence is based on that of SIVmac251 and optimized for yeast codon use. Modified GFPs, which were used as passenger proteins, were also amplified by PCR using primers with *Bam*HI and *Bsr*GI sites at the 5′ and 3′ termini, respectively [33]. The MSD portion of gp41 [27] was synthesized as a complementary pair of oligode-

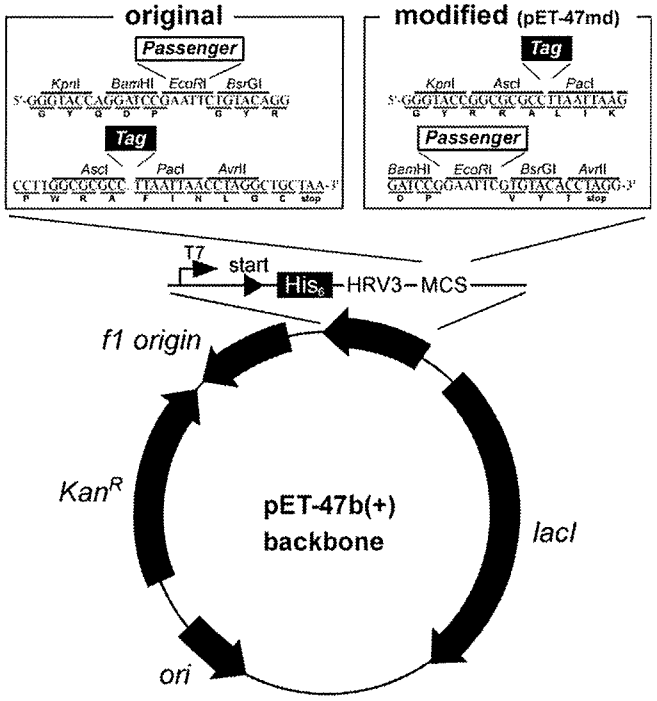

Fig. 1. Schematic representation of plasmids used in this study. The upper panel shows the details of the multiple cloning site of pET-47b and pET-47md. The lower panel shows the structure of the backbone. T7, T7 promoter; His_6, polyhistidine tag; HRV3, region coding amino acid sequences recognized by HRV 3C protease; MCS, multiple cloning site; Tag, genes of tag proteins; Passenger, genes of passenger proteins.

oxyribonucleotides with *Bam*HI and *Bsr*GI sites at the 5′ and 3′ termini, respectively. All amplicons were first cloned into pCR4blunt-TOPO vectors using the TOPO Cloning Kit (Invitrogen), and their sequences were verified before being cloning into pET-47b or pET-47md. After cloning the individual passenger gene between the *Bam*HI and *Bsr*GI sites of pET-47b or pET-47md, the respective tag genes were cloned using the *Asc*I and *Pac*I sites. As a control, we constructed expression vectors containing tag genes but without passenger genes; this was accomplished by cloning the tag genes to pET-47md with the *Asc*I and *Pac*I sites. We verified all sequences of the constructed expression vectors.

Analysis of protein expression on a small-scale

BL21 Star (DE3) (Invitrogen) was transformed by the constructed expression plasmids and cultured on Luria–Bertani (LB) plates containing 50 μg/ml kanamycin. Fresh colonies were picked and inoculated into 2 ml of LB medium containing 50 μg/ml kanamycin. All *E. coli* cells were grown at 37 °C throughout the experiments. When the cells had reached a log phase, isopropylthiogalactoside (IPTG) was added in a final concentration of 1 mM. After a 5-h growth period, the cells were harvested by centrifuging at 15,000*g* for 10 min at 4 °C using an MX-301 centrifuge (Tomy). We extracted the expressed proteins using 300 μl of BugBuster (Novagen) that contained 0.3 μl of Benzonase (Novagen). The total lysate and supernatant was subjected to sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) using a 5% to 20% SDS gradient gel (DRC). We confirmed the expressed protein with the peptide mass fingerprinting method using AXIMA-CFR Plus (Shimadzu).

Measurement of fluorescent activity of GFP

BL21 Star (DE3) cells that had been transformed using the GFP expression vectors (GFP alone or tag-fused GFP at the N terminus) were grown for 5 h in LB medium containing 1 mM IPTG and then harvested by centrifugation (15,000*g*, 4 °C). The cell pellets were suspended in 0.7 ml of phosphate-buffered saline (PBS), and the optical densities at 600 nm were determined. The fluorescence of the suspended cells was analyzed using a FACSCalibur flow cytometer (BD Biosciences). We normalized the measured mean fluorescence signal by dividing it by the obtained optical densities at 600 nm (OD_{600}).

Analysis of stability of tagged protein

A 20-μl aliquot of the supernatant of tag-free or tag-fused GFP-expressing cell lysate prepared as described above was added to a 500-μl buffer solution containing 20 mM Tris–HCl (pH 8.0) and increasing concentrations (1, 2, 3, 4, and 5 M) of guanidium hydrochloride (GdmHCl). After incubating the preparations for 1 day at room temperature, we measured the emission spectra from 500 to 600 nm using an F-4500 fluorescence spectrophotometer (Hitachi) set to an excitation wavelength of 470 nm.

Large-scale expression and purification

BL21 Star (DE3) cells transformed with individual expression plasmids were grown in 1 L of LB medium containing 50 μg/ml kanamycin. When OD_{600} reached 0.5 to 0.7, IPTG was added to produce a final concentration of 1 mM so as to induce the expression of target proteins. After a 5-h growth period, cells were pelleted down and frozen at −30 °C. To extract the protein, the cell pellets were resuspended in buffer A (20 mM NaPi [pH 7.4], 0.5 M NaCl, and 20 mM imidazole). The suspended cells were placed on ice and subjected to ultrasonication for 1 h. The disrupted cell extract was centrifuged at 5000*g* for 30 min at 4 °C to remove insoluble materials. The supernatants were loaded onto Ni Sepharose 6 Fast Flow columns (GE Healthcare) and eluted with a second buffer (20 mM NaPi [pH 7.4], 0.5 M NaCl, and 0.5 M imidazole). HRV 3 C protease (Novagen) was added to the eluted fractions, and the preparations were incubated at 4 °C for more than 16 h to remove the polyhistidine tags. After exchanging the second buffer for buffer A using VivaSpin (Sartorius), the solutions again were loaded onto Ni Sepharose 6 Fast Flow columns. The flow-through fraction was collected and concentrated to 2 to 5 ml in a buffer containing 50 mM Tris–HCl (pH 8.0) and 0.5 M NaCl. The concentrated fraction was subjected to Superdex 200 size exclusion chromatography using an AKTA Purifier (GE Healthcare). We collected the highest peaks and verified their content using SDS–PAGE.

Secondary structure analyses

Purified proteins were placed in buffer (20 mM NaPi [pH 7.0] and 0.5 M NaCl), and their circular dichroism (CD) in the far-UV (ultraviolet) region of 200 to 250 nm was measured at 25 °C using a Pistar-180 Spectrometer (Applied Photophysics).

Results

Expression of Vpr protein with conventional tag fusion

To obtain soluble Vpr protein for future analysis, we tried to express Vpr in *E. coli*. When Vpr was expressed without any tag, it did not express very well. Even when it was expressed, it did not come to a soluble fraction (Fig. 2A). This is consistent with previous studies [26,34]. The conventional tag genes *mbp, gst,* and *nusA* were introduced upstream or downstream of the HIV *vpr* gene in the expression plasmid (Fig. 1). Among the conventional tag proteins, GST and NusA facilitated expression of fusion protein when added at the N or C terminus, whereas MBP failed to achieve a high expression level (Fig. 2B). Unfortunately, all of the tagged Vpr proteins that were expressed were observed in the insoluble fraction (Fig. 2B).

Selection of candidate T. th. tag protein

To seek other effective tags for Vpr protein expression and solubilization, we decided to introduce proteins from *T. th.*; the SG project of *T. th.* [28] has provided many data sets, including those for protein expression, purification, and crystallization [29], that aided in our selection of viable candidates for this effort. We selected 59 candidates from more than 400 proteins based on their expression level and the solubility (see Materials and methods for details). From these 59, we selected 31 candidates with resolutions of structures higher than 2 Å. By coincidence, the molecular weight of each of the 31 candidates was less than 40 kDa. This may reflect the fact that smaller molecular weight proteins are easier to express and crystallize [35].

We further narrowed the candidate pool by selecting 14 proteins for which information on small-scale protein preparation was available. To these 14 candidates, we added a 57.9-kDa protein, the largest in the first-pass group of 59 proteins. We included this protein so as to have a larger molecular weight tag protein that could serve as a favorable model of molecular replacement for larger passenger proteins. The complete list of 15 candidate proteins appears in Table 1. As a group, there seemed to be no common functional or structural correlations, probably because we selected the candidates solely for their practical properties of expression and solubility and on their structural analyses.