

A NEW DENSOVIRUS ISOLATED FROM *Tx. SPLENDENS*

<i>Ts</i> DNV	AGGAACAACA AACCAACAA CATGGGACTT TGAAACAAGT CAAAACATGT	3038
<i>Aae</i> DNV	G..G..... ..T..... ..T..... ..T..... ..T..... ..T.....	
<i>Aa</i> DNV	G..... ..T..... ..T..... ..T..... ..T..... ..T.....	
	TCATCGCAGA TGCAGACAGA GAACCAGAAA ACTTCAACTT AACAAACAGCA	3088
	T..... GG.....	
	 C..... ..G..... ..T.....C.. ..GA.....TT	
	GCAGCAACTG GACCACTTGC ACAGCAAACA GCACAAACAT TACTTTTCAA	3138
	A..... A.....C ..A.....	
	C.T.A.. ..A..... A..... ..A.....	
	TGCAACAAT GACAGATATA CAAAATATGA ATTACCACAA AGAAACCAAT	3188
	C ..T.TC..... ..C..... ..C..... ..C..... ..G.....	
	ATACAAGAGA AATTGACTTC CAACAATAA CAAACAATTA TATGTGGAAA	3238
	TA..... ..T..... ..T..C.. C.....	
	G.. ..CA..... ..G..T... ..T..C.. ..G.....	
	CCATTGGACA TCAGCGCTGC AACAACTTT AGAAAATTAA TCCCAATGGC	3288
	...ACA.... ..T..... ..G..... ..G...G.	
	A..T.AAA.AC ...TCCC... ..T.....T.	
	AGAAGGAATA TACACAACAT CCAATGCTAC AAGTGATATG ACCGAATTAA	3338
	G.....G.. ..T.....A ..AGCG..A.. ..C.A.A... G.A.....	
	G.. ..T....A.. ..GG-A-GA.. TT.A..C-AA ..G.....C.	
	CACAACAAAA TACAACATGG GCCACATCAG GAAAAACAAC CCAAGCAACA	3388
	..G..... AT..GT..AT ..AGG..... ..C..... AG.....T..	
	CATG..... ..GC.....-A ..G.G.A...C	
	CTATTCAGAA AT	3400
		
	G.. ..C	

Fig 4—Nucleotide sequence of the PCR amplicon from the new densovirus isolated from *Toxorhynchites splendens* mosquitoes (*Ts*DNV). Shown is the 412 bp nucleotide sequence of the 451 bp amplicon with the primer sequences removed from each end. It is aligned with sequences of *Aae*DNV (M 37899) and the *Aa*/DNV (X 74945). Only the nucleotide residues of *Aae*DNV and *Aa*/DNV that are different from *Ts*DNV sequence are shown. The nucleotide position numbers of the *Aae*DNV gene are presented at the right side. Dots indicate positions where sequences are the same as in *Ts*DNV. Nucleotide deletions and insertions are represented as dashes and as base letters over the sequence, respectively. The nucleotide sequence of *Ts*DNV shown is the representative sequence of 5 separate amplicons derived by PCR from different individual mosquitoes.

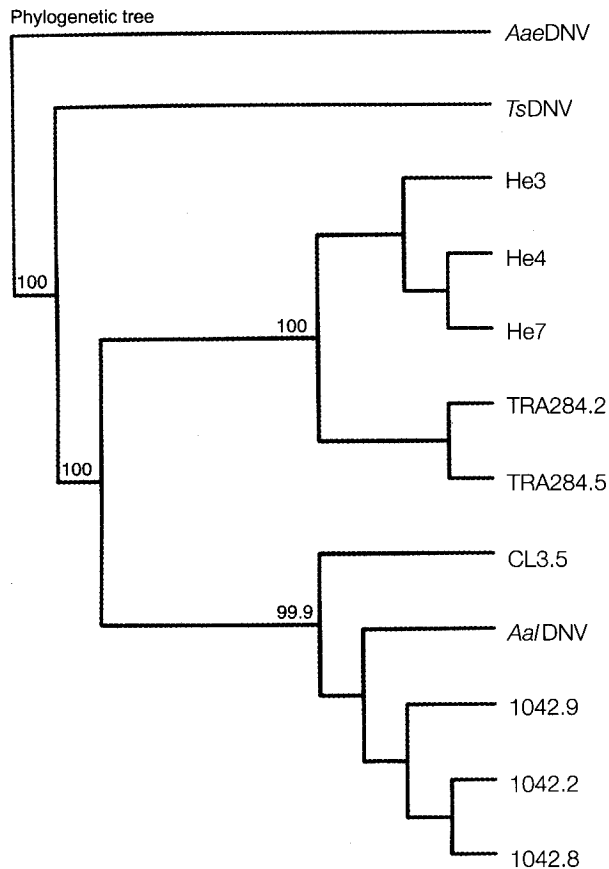


Fig 5—Phylogenetic tree analysis of TsDNV and other densoviruses found from mosquitoes. The phylogenetic tree is based on a comparison of a 301-bp portion of the sequence shown in Fig 4. It was analyzed using PHYLIP package version 3.57c and shows the relationship of TsDNV to other described densoviruses. The sources of densoviruses were AaeDNV (M 37899) and Aa/DNV (X 74945) from the GenBank database and others from mosquito densoviruses from cell lines including clones 1042.2, 1042.8 and 1042.9 from *Aedes aegypti* cell line 1042; CL3.5 from *Culex theileri* cell line CL3; He3, He4 and He7 from *Haemagogus equinus* cell lines; and TRA284.2 and TRA284.5 from *Toxorhynchites amboinensis* cell lines (O'Neill *et al*, 1995). Numbers above the branches refer to the percentage of bootstrap values for 1000 replicates and only the percentage over 50% is indicated.

ruses have been described as widespread in mosquito cell lines distributed in several laboratories (Boublik *et al*, 1994b; O'Neill *et al*, 1995; Chen *et al*, 2004; Paterson *et al*, 2005), suggesting that chronic infection is a common phenomenon that results from contamination by culture manipulation and by exchanges between laboratories. Densoviruses have been reported in laboratory strains of *Cx. pipiens* (Jousset *et al*, 2000) and in indigenous *Ae. aegypti*, *Ae. albopictus* and *An. minimus* S.L. from Thailand (Kittayapong *et al*, 1999; Rwegoshora *et al*, 2000).

The virus we describe here was found unexpectedly in laboratory strains of *Tx. splendens*. This mosquito colony had been used extensively for isolation of dengue virus by inoculation with serum from hospital patients as a standard method for routine virologic confirmation of dengue hemorrhagic fever (DHF) and dengue shock syndrome (DSS) cases (Rosen and Gubler, 1974; Gubler *et al*, 1979; Sumarmo *et al*, 1983; Vaughn *et al*, 1997). The virus was discovered using primers based on conserved densovirus nucleic acid sequences and its uniformity demonstrated by SSCP. Thus, these techniques might be adopted as appropriate techniques in broader screening for the prevalence of densoviruses in indigenous mosquitoes.

Our original inclination was to suspect that *Tx. splendens* had become infected with a densovirus derived from experimental cell lines. However, results from DNA analysis showed that this was not the case, since its DNA sequence differed mostly from the viruses associated with cell lines. It was closer to sequences of viruses derived from captured mosquitoes, suggesting that it came with the original *Tx. splendens* isolate or from contact with other mosquitoes. Indeed, the *Tx. splendens* colony was maintained by feeding with larvae of *Ae. aegypti* or *Cx. quinquefasciatus* and it is possible that they could have been the source of the original infection.

The fact that the virus in all the sampled mosquitoes was identical and that positive PCR results were uniformly obtained within the laboratory for a span of 2 years, suggests the virus had been maintained in the laboratory by horizontal, venereal or vertical transmission throughout the period. Although *Tx. splendens* was found to be infected by predated *Ae. aegypti* larvae in the laboratory, these two mosquito species are also found together in the same breeding places in nature. Unfortunately, the new densovirus did not replicate well in the mosquito cell lines used. It may require mosquito larvae or adults for good replication or it may require primary cultures derived from a specific mosquito. This phenomenon was described previously for its failure to cultivate *Cx. pipiens* densovirus beyond the second passage in *Ae. albopictus* (C6/36), *An. gambiae*, *Cx. quinquefasciatus*, *Cx. tarsalis*, *Drosophila melanogaster* dipteran cells and *Spodoptera littoralis* (SPC SL52) lepidopteran cells (Jousset *et al*, 2000). As far as we could determine, this densovirus did not have any pathological effect on its host during observation period of 2 years. This phenomenon may be similar to that of infectious hypodermal and hematopoietic necrosis virus (IHHNV) that can survive and replicate in some species and varieties of shrimp without noticeable harmful effect (Shike *et al*, 2000).

Many arthropod-borne viral diseases, such as dengue virus, Japanese encephalitis virus, West Nile virus and St. Louis encephalitis virus are transmitted by mosquitoes and cause serious illness in humans (Gubler, 1996; Rigau-Perez *et al*, 1998). Thus, control of these diseases has relied on several intervention strategies that focus on infected mosquitoes. Densoviruses are expected to offer possibilities as potential tools for biological control of arthropod-borne diseases based on experiments performed by several investigators (Belloncik, 1990; Kimmick *et al*, 1998; Afanasiev *et al*, 1999; Allen-Miura *et al*, 1999;

Carlson *et al*, 2006). Before this goal can be achieved however, much more work is needed on the molecular biology, physiology and pathology of these viruses in their hosts.

ACKNOWLEDGEMENTS

We are grateful to Dr Pattamaporn Kittayapong, Department of Biology, Faculty of Science, Mahidol University for providing Aa/DNV, Prof TW Flegel, Department of Biotechnology, Faculty of Science, Mahidol University for reviewing the manuscript and Mr Pravech Ajawatanawong for phylogenetic tree analysis. This study received financial support from the National Center for Genetic Engineering and Biotechnology (BIOTEC), National Science and Technology Development Agency (NSTDA), Thailand. PM was supported by a Senior Research Scholar Grant (no. RTA 4680017) from the Thailand Research Fund.

REFERENCES

- Afanasiev BN, Galyov EE, Buchatsky LP, Kozlov YV. Nucleotide sequence and genomic organization of *Aedes* densovirus. *Virology* 1991; 185: 323-36.
- Afanasiev BN, Kozlov YV, Carlson JO, Beaty BJ. Densovirus of *Aedes aegypti* as an expression vector in mosquito cells. *Exp Parasitol* 1994; 79: 322-39.
- Afanasiev BN, Ward TW, Beaty BJ, Carlson JO. Transduction of *Aedes aegypti* mosquitoes with vectors derived from *Aedes* densovirus. *Virology* 1999; 257: 62-72.
- Allen-Miura TM, Afanasiev BN, Olson KE, Beaty BJ, Carlson JO. Packaging of AeDNV-GFP transducing virus by expression of densovirus structural proteins from a sindbis virus expression system. *Virology* 1999; 257: 54-61.
- Bachmann PA, Hoggan MD, Melnick JL, Pereira HG, Vago C. *Parvoviridae*. *Intervirology* 1975; 5: 83-92.
- Bando H, Choi H, Ito Y, Kawase S. Terminal structure of a densovirus implies a hairpin transfer replication which is similar to the model for

- AAV. *Virology* 1990; 179: 57-63.
- Bando H, Choi H, Ito Y, Nakagaki M, Kawase S. Structural analysis on the single-stranded genomic DNAs of the virus newly isolated from silkworm: the DNA molecules share a common terminal sequence. *Arch Virol* 1992; 124: 187-93.
- Bando H, Hayakawa T, Asano S, Sahara K, Nakagaki M, Iizuka T. Analysis of the genetic information of a DNA segment of a new virus from silkworm. *Arch Virol* 1995; 140: 1147-55.
- Bannai M, Tokunaga K, Lin L, et al. Discrimination of human HLA-DRB1 alleles by PCR-SSCP (single-strand conformation polymorphism) method. *Eur J Immunogenet* 1994; 21: 1-9.
- Belloncik S. Potential use of densovirus as biological control agents of insect pests. In: Tijssen P, ed. Handbook of Parvoviruses, Vol.2. Boca Raton: CRC Press, 1990: 285-9.
- Boublik Y, Jousset FX, Bergoin M. Complete nucleotide sequence and genomic organization of the *Aedes albopictus* parvovirus (AaPV) pathogenic for *Aedes aegypti* larvae. *Virology* 1994a; 200: 752-63.
- Boublik Y, Jousset FX, Bergoin M. Structure, restriction map and infectivity of the genomic and replicative forms of AaPV DNA. *Arch Virol* 1994b; 137: 229-40.
- Buchatsky LP. Densovirus of bloodsucking mosquitoes. *Dis Aquat Org* 1989; 6: 145-50.
- Burivong P, Pattanakitsakul S-N, Thongrunskiat S, Malasit P, Flegel TW. Markedly reduced severity of Dengue virus infection in mosquito cell cultures persistently infected with *Aedes albopictus* densovirus (Aa/DNV). *Virology* 2004; 329: 261-9.
- Carlson J, Suchman E, Buchatsky L. Densoviruses for control and genetic manipulation of mosquitoes. *Adv Virus Res* 2006; 68: 361-92.
- Chen S, Cheng L, Zhang Q, et al. Genetic, biochemical, and structural characterization of a new densovirus isolated from a chronically infected *Aedes albopictus* C6/36 cell line. *Virology* 2004; 318: 123-33.
- Dumas B, Jourdan M, Pascaud AM, Bergoin M. Complete nucleotide sequence of the cloned infectious genome of *Junonia coenia* densovirus reveals an organization unique among parvoviruses. *Virology* 1992; 191: 202-22.
- Gubler DJ. Arboviruses as imported disease agents: the need for increased awareness. *Arch Virol* 1996; 11 (suppl): 21-32.
- Gubler DJ, Suharyono W, Sumarmo WH, Jahja E, Sulianti SJ. Virological surveillance for dengue haemorrhagic fever in Indonesia using the mosquito inoculation technique. *Bull World Health Organ* 1979; 57: 931-6.
- Huang YM. The mosquitoes of Polynesia with a pictorial key to some species associated with filariasis and/or dengue fever. *Mosq Syst* 1977; 9: 289-322.
- Jourdan M, Jousset FX, Gervais M, Skory S, Bergoin M, Dumas B. Cloning of the genome of a densovirus and rescue of infectious virions from recombinant plasmid in the insect host *Spodoptera littoralis*. *Virology* 1990; 179: 403-9.
- Jousset FX, Baquerizo E, Bergoin M. A new densovirus isolated from the mosquito *Culex pipiens* (Diptera: Culicidae). *Virus Res* 2000; 67: 11-6.
- Jousset FX, Barreau C, Boublik Y, Cornet M. A parvo-like virus persistently infecting a C6/36 clone of *Aedes albopictus* mosquito cell line and pathogenic for *Aedes aegypti* larvae. *Virus Res* 1993; 29: 99-114.
- Kimmick MW, Afanasiev BN, Beaty BJ, Carlson JO. Gene expression and regulation from the p7 promoter of *Aedes* densovirus. *J Virol* 1998; 72: 4364-70.
- Kittayapong P, Baisley KJ, O'Neill SL. A mosquito densovirus infecting *Aedes aegypti* and *Aedes albopictus* from Thailand. *Am J Trop Med Hyg* 1999; 61: 612-7.
- Kurstak E. Small DNA densovirus (DNV). *Adv Virus Res* 1972; 17: 207-41.
- Lebedeva OP, Kuznetsova MA, Zelenko AP, Gudzenko AP. Investigation of a virus disease of the densovirus type in a laboratory culture of *Aedes aegypti*. *Acta Virol* 1973; 17: 253-6.
- Molitor TW, Joo HS, Collett MS. Porcine parvovirus

- DNA: characterization of the genomic and replicative form DNA of two virus isolates. *Virology* 1984; 137: 241-54.
- O'Neill SL, Kittayapong P, Braig HR, Andreadis TG, Gonzalez JP, Tesh RB. Insect densoviruses may be widespread in mosquito cell lines. *J Gen Virol* 1995; 76: 2067-74.
- Orita M, Suzuki Y, Sekiya T, Hayashi K. Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. *Genomics* 1989; 5: 874-9.
- Paterson A, Robinson E, Suchman E, Afanasiev B, Carlson J. Mosquito densovirus cause dramatically different infection phenotypes in the C6/36 *Aedes albopictus* cell line. *Virology* 2005; 337: 253-61.
- Rigau-Perez JG, Clark GG, Gubler DJ, Reiter P, Sanders EJ, Vorndam AV. Dengue and dengue haemorrhagic fever. *Lancet* 1998; 352: 971-7.
- Rosen L, Gubler D. The use of mosquitoes to detect and propagate dengue viruses. *Am J Trop Med Hyg* 1974; 23: 1153-60.
- Rwegoshora RT, Baisley KJ, Kittayapong P. Seasonal and spatial variation in natural densovirus infection in *Anopheles minimus* S.L. in Thailand. *Southeast Asian J Trop Med Public Health* 2000; 31: 3-9.
- Shike H, Dhar AK, Burns JC, et al. Infectious hypodermal and hematopoietic necrosis virus of shrimp is related to mosquito brevidensoviruses. *Virology* 2000; 277: 167-77.
- Siegl G, Bates RC, Berns KI, et al. Characteristics and taxonomy of Parvoviridae. *Intervirology* 1985; 23: 61-73.
- Sumarmo WH, Jahja E, Gubler DJ, Suharyono W, Sorensen K. Clinical observations on virologically confirmed fatal dengue infections in Jakarta, Indonesia. *Bull World Health Organ* 1983; 61: 693-701.
- Tijssen P, van den Hurk J, Kurstak E. Biochemical, biophysical, and biological properties of densovirus. I. Structural proteins. *J Virol* 1976; 17: 686-91.
- Vaughn DW, Green S, Kalayanarooj S, et al. Dengue in the early febrile phase: viremia and antibody responses. *J Infect Dis* 1997; 176: 322-30.

Analysis of published *PKD1* gene sequence variants

To the Editor:

We retrospectively reviewed published variants in the *PKD1* genes and detected errors in 39 of 771 variants (5.06% (95% c.i., 3.62–6.85)). All arose from human processing mistakes. As peer-reviewed publication is no safeguard for those considering the clinical significance of an unknown variant, we suggest that reporting of new variants for the proposed Human Variome Project should employ both automated reporting and expert scrutiny.

Measurement of the biological activity of a mutant gene provides the best indication of the functional effect of a gene variant¹. As this is seldom possible, less direct measures are used to assess the likely clinical significance of a variant. These include observing the degree of nucleotide sequence conservation in orthologous genes, the nature and position of resultant amino acid changes, its frequency within cases and control populations and whether it exhibits familial segregation with the disease. With the exception of the latter, this information is often present within published mutation reports, providing grounds for the clinical interpretation of a gene variant.

Not surprisingly, guidance from professional organizations representing medical geneticists emphasizes the important role of locus-specific mutation databases (LSDBs)² (<http://www.hgvs.org/dblist/dblist.html>) and peer-reviewed publications in the evaluation of unknown variants³ (see http://www.acmg.net/Pages/ACMG_Activities/stds-2002/stdsmenu-n.htm and <http://www.cmgs.org/sequencing.htm>). At this early stage in the development of LSDBs, professional bodies also caution against overreliance on databases to interpret the meaning of an observed variant³. Similar cautionary warnings are not issued about published mutation reports, presumably because the rigor of the peer review process instills a stronger sense of confidence in published clinical mutation reports. Studies of quality control in genetic diagnostic laboratories have so far focused on the quality of DNA sequencing⁴ and genotypic allele calling⁵. To our knowledge, there has been no systematic study of the accuracy of variants in peer-reviewed publications.

A recent upgrading of the autosomal dominant polycystic kidney disease (ADPKD) mutation database (PKDB; see <http://pkdb.mayo.edu>)⁶ presented an opportunity to

evaluate the accuracy of published reports of variants in the ADPKD-associated gene *PKD1* (16p13.3). A large number of disease-causing mutations, a large number of polymorphisms not associated with disease and a considerable number of gene variants with an unknown effect are now reported for this gene. This allowed us to assess the accuracy with which 771 variants were reported in 55 peer-reviewed publications (**Supplementary Table 1** online). In order to evaluate the accuracy of these variant reports, the numbering of each was first standardized to the ATG start codon of the *PKD1* NCBI RefSeq mRNA sequence (NM_000296.2). The nomenclature was also altered where necessary, to comply with current nomenclature standards⁷ (<http://www.hgvs.org/mutnomen/>). The accuracy of each reported gene variant and its associated amino acid effect were assessed via an in-house mutation checker tool⁶ in conjunction with the Artemis DNA sequence visualization software⁸. Inconsistencies between the reported nucleotide changes and associated reference sequence or reported amino acid change were subjected to closer investigation until resolved. Corresponding authors were invited to recheck their own publication as well as our suggested amendments. Twenty-two of thirty-four corresponding authors (65%) responded. Excluding the time contributed by corresponding authors, approximately 170 h of curator time was required to standardize and check all the gene variant reports.

As a number of reference sequences and numbering conventions had been used to describe the gene variant reports, the nucleotide numbering of 542 (70.3%) variants required updating. Twenty publications (36%) did not provide details of the reference sequence used or indicate the numbering convention employed. This required the empiric determination of these details. A total of 39 errors (**Supplementary Table 2** online) were identified (5.06% (95% c.i. 3.62–6.85)), 36 of which were detected using the procedure describe above. The remaining three errors were reported by a corresponding author after re-examination of the primary data. The errors, which have been categorized as either miscounting, misassignment or typographical (**Fig. 1**), were identified in both the nucleotide and amino acid descriptions of variants. While it was evident the majority of errors identified in

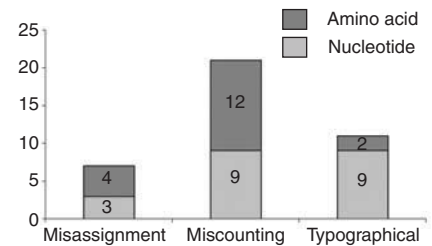


Figure 1 Errors were grouped into three categories: misassignment, miscounting and typographical. The number of instances of each category at the nucleotide (light gray) and amino acid (dark gray) levels is shown.

this study arose from human copying, the discovery of errors in reporting primary sequence data indicates that this vital first step is also prone to error. As many primary data were not reviewed in this study, it is highly likely the error rate reported here is an underestimate. The 10% administrative error rate detected among diagnostic laboratories reported in ref. 5 and the 13% rate of genotyping and nomenclature errors among 64 diagnostic laboratories participating in an international external quality assessment of sequence-based genetic testing⁴ could more closely approximate the error rate in peer-reviewed publications.

Our experience of checking published variants, correcting their nomenclature and adjusting their correspondence to a common reference sequence provides some insight into the likely cost of completing the task for other LSDBs. At present, there is little scope for automating the initial steps required to standardize numbering and nomenclature. However, the later steps of checking for inconsistencies in the description of variants are being aided increasingly by mutation-checking and visualization software^{6,8} (<http://www.ebi.ac.uk/cgi-bin/mutations/check.cgi>).

Much is known about the processes that contribute to human errors, as well as the proofreading and feedback processes required to identify mistakes⁹. Although the most effective way of reducing human error is to introduce automated processes that reduce or eliminate the need for human characterization and transcription of gene variants, software introduces new sources of error: for example, gene name errors introduced through the use of Microsoft Excel¹⁰.

In the interim, journal editors and LSDB curators should consider a range of strategies,

including encouraging authors of genomic variant reports to recheck gene variant data before releasing it into the public domain. Sequence data may now be checked prospectively with mutation checking software.

We must point out that the scope of our analysis did not extend to scrutinizing the accuracy of clinical interpretation of the reported variants, which adds another dimension of risk to the clinical use of published gene variants. The consequences of being unable to access a report of a rare disease-causing variant may not, in practice, be a major issue because its rarity alone should prompt suspicions about its potential pathogenicity. However, failure to access a report of a rare polymorphic variant, perhaps found only in a specific population, would seem likely to increase the risk of wrongly concluding that the variant is associated with disease. This, in turn, could prompt inappropriate clinical decisions. By contrast, incorrect assignment of a novel pathogenic variant as non-disease associated may prompt unnecessary expenditure on additional molecular screening when the primary disease-causing lesion has already been characterized. Finally, erroneous reporting of a rare non-disease associated polymorphism as disease-causing could result in inappropriate clinical decisions because of the mistake. Potential ramifications of these scenarios are poignantly illustrated in a hypothetical scenario presented in ref. 11.

The error rate detected among these gene variants published in peer-reviewed journals shows that caution must be exercised—particularly by curators of LSDBs, genetic diagnostic laboratories, genetic counselors and other health care practitioners—when relying on published reports to evaluate the likely clinical significance of an unknown variant. The nature of the mistakes demonstrates that steps toward reducing human single-entry recording of sequence variants into peer-reviewed publications or LSDBs will enhance the accuracy and clinical utility of the Human Variome Project.

Alexander M Gout^{1–3}, the ADPKD Gene Variant Consortium⁴ & David Ravine^{3,5}

¹The Walter and Eliza Hall Institute of

Medical Research, 1G Royal Parade, Parkville, Victoria 3050, Australia. ²Department of Medical Biology, The University of Melbourne, Parkville, Victoria 3010, Australia. ³School of Medicine and Pharmacology, The University of Western Australia, Nedlands 6009, Australia.

⁴Corresponding authors of publications describing variants within the PKD1 gene (a full list of authors is given at the end of the paper). ⁵Western Australian Institute for Medical Research, Centre for Medical Research, The University of Western Australia, Nedlands 6009, Australia.

e-mail: david.ravine@uwa.edu.au

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

The members of the ADPKD Gene Variant Consortium have been critical in this collaborative effort of verifying the accuracy of the peer-reviewed PKD1 gene variants included in this retrospective audit. This work was supported by funding from the PKD Foundation. We are grateful to J. Crowhurst and H. Scott for their careful scrutiny of the manuscript. The encouragement of R. Cotton in the preparation of this report is appreciated.

COMPETING INTERESTS STATEMENT

The authors declare no competing financial interests.

1. Cotton, R.G. & Scriver, C.R. *Hum. Mutat.* **12**, 1–3 (1998).
2. Horaitis, O. & Cotton, R.G. *Hum. Mutat.* **23**, 447–452 (2004).
3. Maddalena, A., Bale, S., Das, S., Grody, W. & Richards, S. & the ACMG Laboratory Quality Assurance Committee. *Genet. Med.* **7**, 571–583 (2005).
4. Patton, S.J., Wallace, A.J. & Elles, R. *Clin. Chem.* **52**, 728–736 (2006).
5. Dequeker, E. & Cassiman, J.J. *Nat. Genet.* **25**, 259–260 (2000).
6. Gout, A.M., Martin, N.C., Brown, A.F. & Ravine, D. *Hum. Mutat.* (in the press).
7. den Dunnen, J.T. & Antonarakis, S.E. *Hum. Mutat.* **15**, 7–12 (2000).
8. Rutherford, K. et al. *Bioinformatics* **16**, 944–945 (2000).
9. Pilotti, M., Chodorow, M. & Thornton, K.C. *J. Gen. Psychol.* **131**, 242–266 (2004).
10. Zeeberg, B.R. et al. *BMC Bioinformatics* **5**, 80–85 (2004).
11. den Dunnen, J.T. & Paalman, M.H. *Hum. Mutat.* **22**, 181–182 (2003).

Members of the ADPKD Gene Variant Consortium include the following: Peter C Harris⁶, Sandro Rossetti⁶, Dorien Peters⁷, Martijn Breuning⁷, Elizabeth Petri Henske⁸, Akio Koizumi⁹, Sumiko Inoue⁹, Yoshiko Shimizu¹⁰, Wonna Thongnoppakhun¹¹,

Pa-thai Yenchitsomanus¹¹, Constantinos Deltas¹², Richard Sandford¹³, Roser Torra¹⁴, Alberto E Turco¹⁵, Steve Jeffery¹⁶, Michel Fontes¹⁷, Stefan Somlo¹⁸, Laszlo M Furu¹⁸, Yvo M Smulders¹⁹, Bernard Mercier²⁰, Claude Ferec²⁰, Stéphane Burtet¹⁷, York Pei²¹, Luba Kalaydjieva²², Nadja Bogdanova²³, Marie McCluskey²⁴, Lee Jung Geon²⁵, C H Wouters²⁶, Jana Reiterova²⁷, Jitka Stekrova²⁷, Jose L San Millan²⁸, Gianluca Aguiari²⁹ & Laura Del Senno²⁹

⁶Mayo Clinic College of Medicine, Rochester, New York, USA. ⁷Leiden University Medical Centre, Leiden, The Netherlands. ⁸Fox Chase Cancer Center, Philadelphia, Pennsylvania, USA. ⁹Department of Health and Environmental Sciences, Kyoto University, Japan. ¹⁰Kyoto University School of Health Sciences, Japan. ¹¹Faculty of Medicine, Siriraj Hospital, Mahidol University, Bangkok, Thailand. ¹²Department of Biological Sciences, University of Cyprus, Lefkosia, Cyprus. ¹³Department of Medical Genetics, Cambridge Institute of Medical Research, Addenbrooke's Hospital, Cambridge, UK. ¹⁴Nephrology Department, Fundació Puigvert, Barcelona, Spain. ¹⁵Department of Mother & Child - Genetics, University of Verona School of Medicine, Verona, Italy. ¹⁶Medical Genetics Unit, St. George's Hospital Medical School, London, UK. ¹⁷Génétique médicale et développement, INSERM, Marseille, France. ¹⁸Section of Nephrology, Yale University School of Medicine, New Haven, Connecticut, USA. ¹⁹Free University Medical Center Department of Internal Medicine, Amsterdam, The Netherlands. ²⁰Laboratoire de génétique moléculaire, CHU Brest, France. ²¹Division of Nephrology, Department of Medicine, University of Toronto, Canada. ²²Western Australian Institute for Medical Research, Centre for Medical Research, The University of Western Australia, Perth, Australia. ²³Institute of Human Genetics, University of Muenster, Germany. ²⁴Centre for Human Genetics, Edith Cowan University, Joondalup, Australia. ²⁵Department of Internal Medicine, Eulji Medical College, Seoul National University, Seoul, Korea. ²⁶Department of Clinical Genetics, Rotterdam, The Netherlands. ²⁷Department of Nephrology & Department of Biology and Medical Genetics, Charles University, Prague, Czech Republic. ²⁸Departments of Endocrinology and Molecular Genetics, Hospital Ramo, Madrid, Spain. ²⁹Dipartimento di Biochimica e Biologia Molecolare, Università degli Studi, Ferrara, Italy.

IgA nephropathy associated with Hodgkin's disease in children: a case report, literature review and urinary proteome analysis

Sookkasem Khositseth · Nonglak Kanitsap ·
Naree Warnnissorn · Visith Thongboonkerd

Received: 21 July 2006 / Revised: 17 October 2006 / Accepted: 24 October 2006 / Published online: 2 December 2006
© IPNA 2006

Abstract We report herein a rare case of IgAN associated with Hodgkin's disease in a 14-year-old boy. Clinical manifestations and laboratory parameters were improved after treatment with CHOP chemotherapy and enalapril. Urinary proteins were isolated and examined using state-of-the-art proteomic technology, before and during the treatment course. Two-dimensional gel electrophoresis showed obvious alterations in the urinary proteome profile in response to such therapy. Quantitative intensity analysis and gel mapping revealed 14 altered proteins with reduced

excretion levels during the treatment course, including albumin, albumin complexed with decanoic acid, α -1 antitrypsin, cadherin-11 precursor, collagen α 1 (VI) chain precursor, complement C1q tumor necrosis factor-related protein, Ig heavy chain, Ig light chain, kininogen, inter- α -trypsin inhibitor (α -1 microglobulin), inter- α -trypsin inhibitor heavy chain, leucine-rich α -2 glycoprotein, β -2 microglobulin, and transferrin precursor. Their potential roles in the pathogenesis and pathophysiology of IgAN are discussed.

There are no potential conflicts of interest in this article.

S. Khositseth
Department of Pediatrics, Faculty of Medicine,
Thammasat University,
Bangkok, Thailand

N. Kanitsap
Department of Medicine, Faculty of Medicine,
Thammasat University,
Bangkok, Thailand

N. Warnnissorn
Department of Pathology, Faculty of Medicine,
Thammasat University,
Bangkok, Thailand

V. Thongboonkerd (✉)
Medical Molecular Biology Unit,
Office for Research and Development,
Faculty of Medicine, Siriraj Hospital,
Mahidol University,
12th Fl Adulyadej Vikrom Bldg, 2 Prannok Rd,
Bangkoknoi,
Bangkok 10700, Thailand
e-mail: thongboonkerd@dr.com

Keywords IgA nephropathy · Hodgkin's lymphoma ·
Urine · Proteome · Proteomics · Proteins

Introduction

IgA nephropathy (IgAN) is the most common cause of glomerulonephritis, and is characterized by mesangial proliferation and deposition of IgA [1]. Secondary IgAN is associated with several diseases, including immunological disorders, infections and cancers. Malignancies that have been reported to be associated with IgAN include renal cell carcinoma [2], cancers of the lung [3], larynx [4] and esophagus [5], cutaneous T-cell lymphoma [6], Hodgkin's disease [7–9], and non-Hodgkin lymphoma [10]. Secondary IgAN in patients with Hodgkin's lymphoma is very rare, and its pathogenic mechanisms remain unclear. To our knowledge, this is the first report of these two diseases coinciding in a pediatric patient. We also describe, for the first time, serial changes in the urinary proteome profile during the treatment course. The altered

proteins were subsequently identified by gel mapping using human urinary proteome databases, which have been recently available [11–13].

Case report

A 14-year-old boy who had suffered cough and right chest pain for 6 months prior to hospitalization. His past history was unremarkable and there were no episodes of hematuria prior to this admission. Computerized tomography (CT) of the chest showed multiple, thick-walled, cavitary lesions and pulmonary nodules in the right lung, as well as enlarged mediastinal lymph nodes (Fig. 1). Open lung biopsy was performed with an informed consent, and histological examination revealed nodular sclerosing, Hodgkin's disease (Fig. 2a and b). His blood pressure, temperature and pulse rate were normal. Physical examination revealed decreased breath sound over right upper lung and pitting edema at both lower extremities. Routine laboratory tests showed normal serum levels of electrolytes, calcium and phosphate. Blood urea nitrogen was 9 mg/dl, serum creatinine was 0.9 mg/dl, serum albumin was 2.2 g/dl, and hemoglobin was 7 g/dl. Liver function parameters were within their normal limits. Urinalysis showed RBC > 100 per high power field (HPF), and urinary protein excretion was 917 mg/day. Antinuclear antibody, anti-DNA, anti-streptolysin O, anti-neutrophil cytoplasmic antibody, anti-glomerular basement membrane antibody, anti-HBs, and anti-HCV were all negative. Complement factors C3 and C4 were within normal ranges. Renal biopsy was performed, and in total 27 glomeruli were obtained and evaluated. Two glomeruli had global sclerosis, whereas diffuse mesangial proliferation with increased cellularity and expanded matrix was observed in the remaining glomeruli (Fig. 2c). The interstitial area had mild-degree

fibrosis, and renal tubules were mildly atrophic. Renal arterioles and small arteries appeared normal. Immunohistochemical staining showed mesangial deposition of IgA (Fig. 2d) and, to a much less extent, IgM.

The patient received 5 cycles of CHOP chemotherapy, which included cyclophosphamide (day 1; 750 mg/m²), vincristine (day 1; 1.2 mg/m²), doxorubicin (day 1; 40 mg/m²) and prednisone (days 1–5; 0.75 mg/kg/day). Enalapril (20 mg/day) was also prescribed. Proteinuria was progressively decreased after starting chemotherapy and enalapril. Urine protein/creatinine ratio was 2.2 at the beginning of treatment, and declined to 1.6, 0.5 and 0.2 after the treatment for 1, 4 and 5 months, respectively. Urine RBC counts were >100/HPF, 50–100/HPF, and <50/HPF after the treatment for 1, 4 and 5 months, respectively.

Urinary proteome analysis

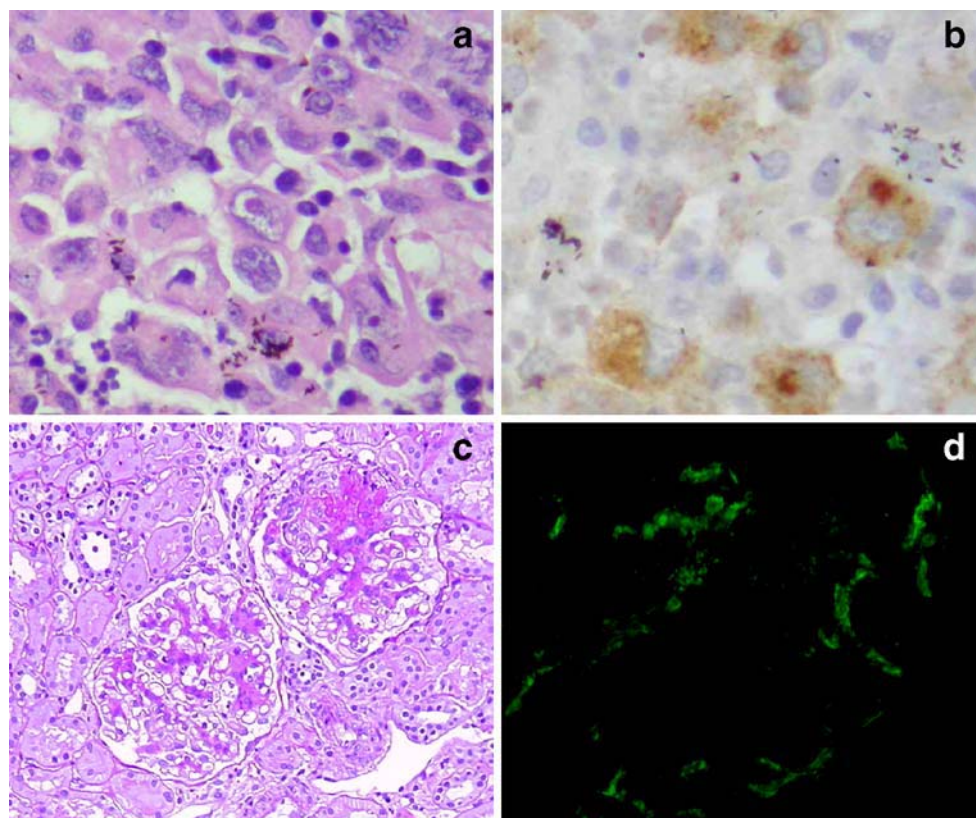
Morning (either first or second void) urine samples (10 ml each) were collected before and during the treatment course. Cells, debris and other particulate matters were removed immediately after the collection using low-speed centrifugation (1,000 × *g* for 5 min). Supernatants were then saved at –20°C until use. Sample preparation at the time of proteome analysis was performed using 50% acetone precipitation method, as previously described [14]. The isolated proteins (200 µg for each sample) were then resolved by two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) (linear pH gradient 3–10) and visualized with Coomassie Brilliant Blue R-250 stain. Each visualized spot represented individual protein excreted into the urine. The 2D spot patterns of urinary proteins obtained from different time-points (before and after treatment for 1 and 4 months) were compared using Image Master 2D Platinum software (Amersham Biosciences, Uppsala, Sweden). This 2D-image analysis software was used for matching and analysis of protein spots on 2D gels. A reference gel was artificially created by combining all of the spots presenting in different gels into one image. This reference gel was then used for matching of corresponding protein spots between gels. The quantity of each protein spot was measured based on the relative intensity level (which determined the intensity volume of each spot normalized to the total intensity volume of all spots visualized within the same gel) and reported as arbitrary unit.

Figure 3 clearly illustrates that the pattern of protein spots (urinary proteome profile) of the reported case with secondary IgAN associated with Hodgkin's disease was significantly altered after the treatment with CHOP regimen and ACEI, when clinical manifestations and proteinuria were improved. Quantitative intensity analysis revealed 14



Fig. 1 Chest computerized tomography revealed multiple, thick-walled, cavitary lesions and nodules in the right lung

Fig. 2 Histopathological findings. **a** Reed-Sternberg cells (mirror image) and mononuclear variants are found in a polymorphous background of small lymphocytes, granulocytes and macrophages in lung biopsy tissue (Hematoxylin and Eosin; original magnification=6000 $\times$). **b** Membrane and perinuclear (golgi apparatus) immunoreactivity for CD30 in Reed-Sternberg cells (Immunoperoxidase; original magnification=6000 $\times$). **c** Mesangial proliferation and segmental sclerosis were observed in renal biopsy specimen (Hematoxylin and Eosin; original magnification=200 $\times$). **d** Granular IgA deposits localized at mesangial areas, of the glomeruli (Original magnification=200 $\times$)



differentially expressed protein spots, which had obvious changes in their relative quantities (intensity levels) after the treatment. These altered proteins were then identified by gel mapping using the available proteome maps of human urine [11–13], and their identities as well as quantitative data are shown in Table 1.

We also compared the urinary proteome profile of this reported case to the profiles of a patient with primary IgAN and a normal healthy boy with comparable age. Figure 4 clearly shows that the urinary proteome profile of secondary IgAN (in this reported case) obviously differed from the patterns/profiles of urinary proteomes of primary IgAN and a normal healthy individual.

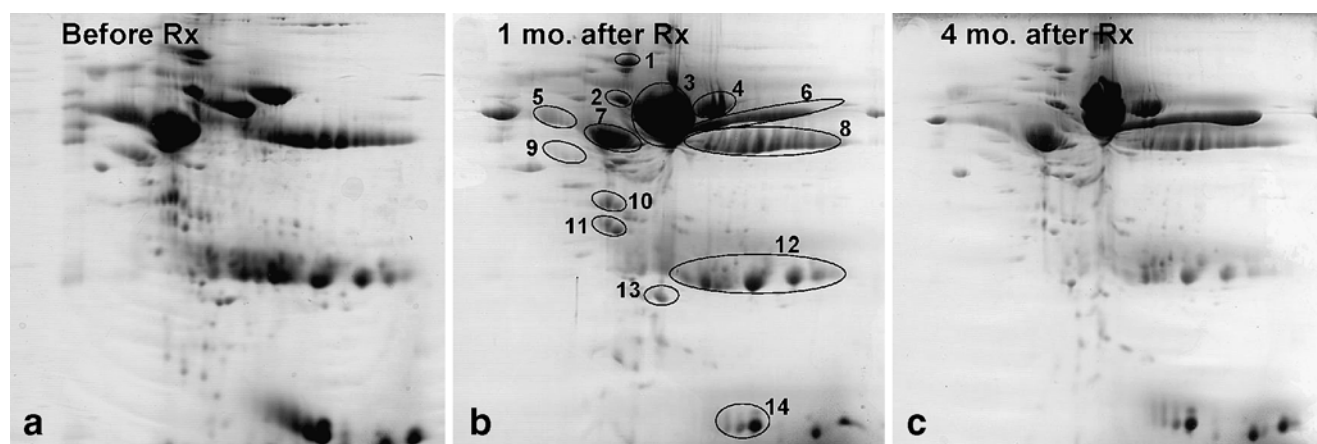


Fig. 3 Urinary proteome analysis. Proteins (200 μ g for each sample) isolated with 50% acetone precipitation method were resolved by two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) on the bases of differential pHs or charges of proteins for the 1st dimensional separation (x-axis) and of differential molecular masses for the 2nd dimension (y-axis). During the treatment, when clinical symptoms and laboratory parameters were improved (protein/creatinine ratios were 2.2 before treatment, 1.6 after treatment for 1 month, and 0.5 after

treatment for 4 months), the 2D spot pattern of urinary proteins was clearly changed. The protein spots labeled with numbers are those significantly altered after the treatment, some of which might play important role(s) in the pathogenesis and/or pathophysiology of secondary IgAN associated with Hodgkin's disease. These altered proteins were subsequently identified by gel mapping using currently available human urinary proteome maps, and their identities as well as quantitative data are shown in Table 1

Table 1 Urinary proteins with obvious changes in their relative quantities during treatment

Spot no.	Protein	Relative intensity (arbitrary unit)		
		Before Rx	After 1 month of Rx	After 4 months of Rx
1	Collagen α 1 (VI) chain precursor	2.2557	0.9672	0.4176
2	Cadherin-11 precursor	1.4349	0.8640	0.8987
3	Albumin	3.8356	29.4095	26.3580
4	Transferrin precursor	6.2259	4.6295	3.9640
5	Kininogen	4.5545	2.2577	2.1885
6	Albumin complexed with decanoic acid	0.0000	9.5297	18.7047
7	α -1 antitrypsin	12.4364	7.6334	9.1724
8	Ig heavy chain	17.6034	14.4169	13.5228
9	Leucine-rich α -2 glycoprotein	1.4343	0.7459	0.3757
10	Inter- α -trypsin inhibitor heavy chain	1.6352	1.2922	0.4895
11	Inter- α -trypsin inhibitor (α -1 microglobulin)	1.0898	1.5999	0.6712
12	Ig light chain	32.3378	18.8605	14.8984
13	Complement C1q tumor necrosis factor-related protein	0.7584	0.8432	0.2912
14	β -2 microglobulin	9.7680	2.6651	3.5858

Discussion

We report a first pediatric patient with secondary IgAN associated with Hodgkin's disease. The association between Hodgkin's lymphoma, particularly the mixed-cellularity type, and glomerular diseases is well recognized [15]. Minimal change nephropathy is the most common histological finding [16]. Other lesions, which have been previously reported, include membranous nephropathy [17], focal segmental glomerulosclerosis [18], mesangiocapillary glomerulonephritis [19], anti-glomerular basement membrane nephritis [20], and acute necrotizing glomerulonephritis [21]. The association between IgAN and Hodgkin's disease has been previously described in only four adult patients. Ng et al. [22] described a 39-year-old man who had typical symptoms of Henoch–Schönlein purpura (HSP), including abdominal pain, bloody diarrhea, arthral-

gia and purpura of both legs. The patient had proteinuria (1 g/day), and chest CT revealed a mass in the mediastinum. Histopathology diagnosed nodular sclerosing Hodgkin's disease. Unfortunately, kidney biopsy was not performed in this case report.

Blanco et al. [7] reported a 29-year-old male who presented with fever, night sweating, cervical mass and typical symptoms of HSP. Skin biopsy showed leukocytoclastic vasculitis without any immunofluorescence deposit. Urinalysis showed nephrotic-range proteinuria. Renal biopsy revealed focal and segmental glomerulonephritis, with extracapillary crescent in 30% of the glomeruli. The immunofluorescence study showed granular deposition of IgA in mesangial areas. Lymph node biopsy revealed mixed-cellularity Hodgkin's disease. After three courses of chemotherapy (adriamycin, bleomycin, vincristine and

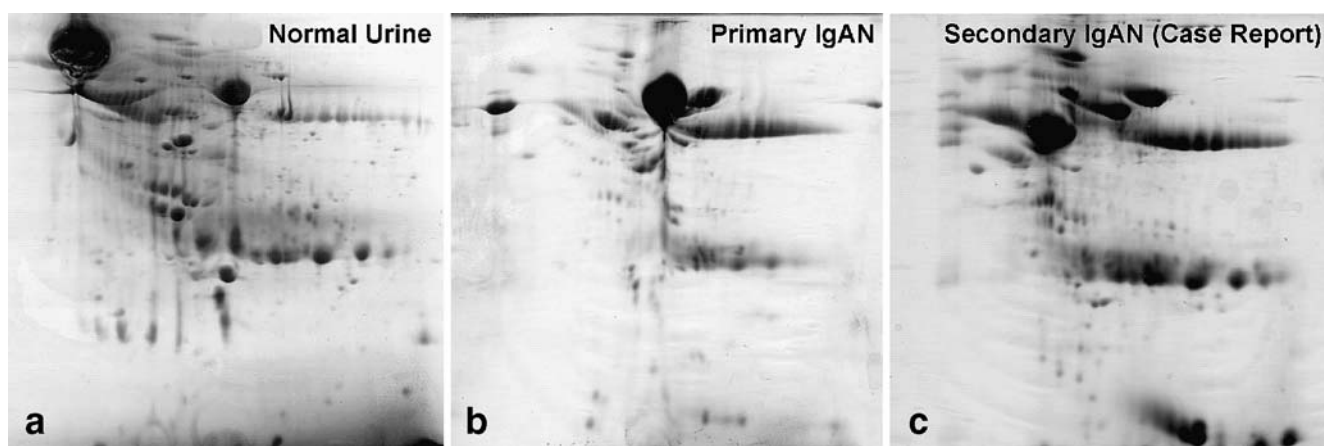


Fig. 4 Differential urinary proteome profiles of primary versus secondary IgAN. Equal amount of proteins (200 μ g for each sample) were resolved with 2D-PAGE and visualized with Coomassie Brilliant

Blue-R250 stain. **a** Normal urine. **b** Urine obtained from a patient with primary IgAN. **c** Urine sample of secondary IgAN (this reported case)

DTIC) and radiation, proteinuria declined progressively and the patient was in complete remission.

Recently, Cherubini et al. [8] reported a 44-year-old man who presented with severe hypertension, nephrotic syndrome, chronic sacroiliac osteomyelitis, normosideremic anemia and mediastinal mass. His renal function rapidly declined within only 1 week, and hemodialysis was finally required. Biopsy of the mediastinal mass showed nodular sclerosing, Hodgkin's disease. Renal biopsy demonstrated diffuse extracapillary glomerulonephritis with IgA deposits. The patient received pulse methylprednisolone for 3 days, and MOPP/ABVD chemotherapy. The renal function was completely recovered after 6 months of such therapy.

More recently, Bergmann et al. [9] reported a 60-year-old woman who had cervical lymphadenopathy and proteinuria. Lymph-node biopsy showed mixed-cellularity type, Hodgkin's disease. Renal biopsy revealed diffuse mesangial proliferation with increased cellularity. Ten percent of the glomeruli had crescentic lesions, and one glomerulus had global sclerosis. The immunofluorescence study showed predominately mesangial granular deposition of IgA. The patient received pulse methylprednisolone treatment and eight cycles of chemotherapy. Thereafter, she was in complete hematologic remission, and serum creatinine gradually declined to 1.3 mg/dl at the end of therapy.

IgAN could be present at the same time as Hodgkin's disease, or following the onset of Hodgkin's disease, as in our and Cherubini's reports [8]. Interestingly, all previously reported patients with secondary IgAN associated with Hodgkin's disease, including our patient, had nodular sclerosing of the mixed-cellularity type. The remission of renal involvement after chemotherapy in all patients may imply the relationship between lymphoma and IgAN. However, effect from ACEIs must be accounted for in the renal remission as they were used in all of the mentioned studies.

Proteome analysis revealed differential urinary proteome profiles of primary versus secondary IgAN (Fig. 4). This difference might indicate that spectra of pathogenic mechanisms, pathophysiology, and/or molecular basis of primary IgAN and secondary IgAN associated with Hodgkin's disease are somewhat different. Recent advances in proteomic technologies have allowed identification of several proteins simultaneously, leading to biomarker discovery and better understanding of the disease processes. When the therapeutic outcome in our case was satisfactory, and laboratory parameters were improved or recovered, the 2D spot pattern of urinary proteins (proteome profile) was also significantly altered. Proteins of which relative quantities were reduced during the treatment course might be the ones that play important role(s) in the pathogenesis and/or pathophysiology of secondary IgAN. These proteins include type VI collagen ($\alpha 1$ chain), cadherin-11, transferrin,

kininogen, α -1 antitrypsin, Ig heavy and light chains, leucine-rich α -2 glycoprotein, α -1 microglobulin, complement C1q tumor necrosis factor-related protein, and β -2 microglobulin. Interestingly, these proteins were down-regulated as early as 1 month after the treatment, prior to the reduction of the relative abundance of albumin, which was significantly decreased 4 months after the treatment, implicating that the decrease of these proteins occurred prior to the improvement of glomerular leakage, and that they might play important role(s) in the pathogenesis and/or pathophysiology of secondary IgAN associated with Hodgkin's disease. The potential roles for some of these proteins are highlighted as follows.

The collagen superfamily has been found to be associated with several glomerular diseases. The most common type of collagens found in glomerular diseases is type IV (basement membrane-associated) [23]. The role of type VI (ubiquitous, bead filaments) collagens in IgAN remains unclear, and needs to be further elucidated. Cadherins are type I membrane, calcium-dependent, cell-adhesion proteins. They preferentially interact with themselves and other molecules (e.g., catenins [24] and nephrin [25], which are the key components of podocyte slit diaphragm that control glomerular permeability), and contribute to the sorting of heterogeneous cell types. Cell-adhesive mechanisms have an impact on cell migration and proliferation, and thus are involved in the pathogenesis of glomerulonephritides. Alpha-1 antitrypsin (AAT) is a member of the serine protease inhibitor (SERPIN) family. It is one of the abundant proteins in human urine, and it has been suggested that it serves as a marker for glomerulopathy [26]. Recently, it has been proposed that another potential role of AAT is an association with the regulation of inflammation and matrix accumulation [27, 28]. Our data also demonstrated that the decrease in AAT occurred before the recovery of glomerular leakage, suggesting that AAT not only serves as a biomarker for glomerulopathy, but also has another important role in the regulation of glomerular injury in IgAN.

It should be noted that the data reported in Table 1 were the "relative" intensities of individual proteins normalized with the "total" intensity of all protein spots visualized within the same gel. During the treatment course, while the total protein excretion was reduced, the "relative" intensity levels of albumin (spots #3 and #6) were increased after 1 month of therapy as compared to basal levels (before treatment). This was not unexpected, as we used equal amount of "total" protein in each 2D gel. The relative abundance of albumin at 1 month after the treatment was increased could be simply explained by the fact that the amounts of other proteins, which have been thought to be inflammatory mediators or to play important pathogenic/pathophysiologic roles, were dramatically reduced at that

time, resulting in an increase in the “relative” amount of albumin. At the later stage, when glomerular leakage was improved while the inflammatory processes had been readily controlled or stable, the relative amount of albumin finally declined.

In summary, our report has demonstrated the potential use of proteomic technology in clinical medicine to monitor therapeutic response and to explore the pathogenic mechanisms of human diseases. Further functional analyses of the altered proteins mentioned above, and characterization of IgA subsets (i.e., aberrantly glycosylated IgA involved in the mesangial deposition and disease progression) as well as phenotyping of malignant T lymphocytes, would result in a better understanding of the pathogenesis and pathophysiology of secondary IgAN associated with Hodgkin’s lymphoma, and might also lead to biomarker discovery.

Acknowledgements We thank Dr. Somchai Chutipongtanate, Ms. Theptida Semangoen, Dr. Adis Tasanarong, and Dr. Thipaporn Tharavanij for their assistance. This study was supported by Mahidol University and by support from the National Research Council of Thailand to V. Thongboonkerd.

References

- Barratt J, Feehally J (2005) IgA nephropathy. *J Am Soc Nephrol* 16:2088–2097
- Magyarlaki T, Kiss B, Buzogany I, Fazekas A, Sukosd F, Nagy J (1999) Renal cell carcinoma and paraneoplastic IgA nephropathy. *Nephron* 82:127–130
- Schutte W, Ohlmann K, Koall W, Rosch B, Osten B (1996) Paraneoplastic IgA nephritis as the initial symptom of bronchial carcinoma. *Pneumologie* 50:494–495
- Mustonen J, Pasternack A, Helin H (1984) IgA mesangial nephropathy in neoplastic diseases. *Contrib Nephrol* 40:283–291
- Lam KY, Law SY, Chan KW, Yuen MC (1998) Glomerulonephritis associated with basaloid squamous cell carcinoma of the oesophagus. A possible unusual paraneoplastic syndrome. *Scand J Urol Nephrol* 32:61–63
- Moe SM, Baron JM, Coventry S, Dolan C, Umans JG (1993) Glomerular disease and urinary Sezary cells in cutaneous T-cell lymphomas. *Am J Kidney Dis* 21:545–547
- Blanco P, Denisi R, Rispoli P, Deminiere C, Pellegrin JL, Leng B, Aparicio M (1999) Henoch-Schönlein purpura associated with segmental and focal proliferative glomerulonephritis in a patient with Hodgkin’s disease. *Nephrol Dial Transplant* 14:179–180
- Cherubini C, Barbera G, Di Giulio SD, Muda AO, Faraggiana T (2001) Lymphomas and IgA nephropathy. *Nephrol Dial Transplant* 16:1722–1723
- Bergmann J, Buchheidt D, Waldherr R, Maywald O, van der Woude FJ, Hehlmann R, Braun C (2005) IgA nephropathy and Hodgkin’s disease: a rare coincidence. Case report and literature review. *Am J Kidney Dis* 45:e16–e19
- Zahner J, Bach D, Malms J, Schneider W, Diercks K, Grabensee B (1997) Glomerulonephritis and malignant lymphoma. Mostly men with low-grade lymphoma with various forms of glomerulonephritis. *Med Klin (Munich)* 92:712–719
- Thongboonkerd V, McLeish KR, Arthur JM, Klein JB (2002) Proteomic analysis of normal human urinary proteins isolated by acetone precipitation or ultracentrifugation. *Kidney Int* 62:1461–1469
- Pieper R, Gatlin CL, McGrath AM, Makusky AJ, Mondal M, Seonarain M, Field E, Schatz CR, Estock MA, Ahmed N, Anderson NG, Steiner S (2004) Characterization of the human urinary proteome: a method for high-resolution display of urinary proteins on two-dimensional electrophoresis gels with a yield of nearly 1,400 distinct protein spots. *Proteomics* 4:1159–1174
- Oh J, Pyo JH, Jo EH, Hwang SI, Kang SC, Jung JH, Park EK, Kim SY, Choi JY, Lim J (2004) Establishment of a near-standard two-dimensional human urine proteomic map. *Proteomics* 4:3485–3497
- Thongboonkerd V, Chutipongtanate S, Kanlaya R (2006) Systematic evaluation of sample preparation methods for gel-based human urinary proteomics: quantity, quality, and variability. *J Proteome Res* 5:183–191
- Alpers CE, Cotran RS (1986) Neoplasia and glomerular injury. *Kidney Int* 30:465–473
- Peces R, Sanchez L, Gorostidi M, Alvarez J (1991) Minimal change nephrotic syndrome associated with Hodgkin’s lymphoma. *Nephrol Dial Transplant* 6:155–158
- Eagen JW (1977) Glomerulopathies of neoplasia. *Kidney Int* 11:297–303
- Watson A, Stachura I, Fragola J, Bourke E (1983) Focal segmental glomerulosclerosis in Hodgkin’s disease. *Am J Nephrol* 3:228–232
- Buyukpamukcu M, Hazar V, Tinaztepe K, Bakkaloglu A, Akyuz C, Kutluk T (2000) Hodgkin’s disease and renal paraneoplastic syndromes in childhood. *Turk J Pediatr* 42:109–114
- Ma KW, Golbus SM, Kaufman R, Staley N, Londer H, Brown DC (1978) Glomerulonephritis with Hodgkin’s disease and herpes zoster. *Arch Pathol Lab Med* 102:527–529
- Wolf G, Krenz I, Hegewisch-Becker S, Hossfeld DK, Helmchen U, Stahl RA (2001) Necrotizing glomerulonephritis associated with Hodgkin’s disease. *Nephrol Dial Transplant* 16:187–188
- Ng JP, Murphy J, Chalmers EM, Hogg RB, Cumming RL, Peebles S (1988) Henoch-Schönlein purpura and Hodgkin’s disease. *Postgrad Med J* 64:881–882
- Akazawa H, Nakajima M, Nishiguchi M, Yamoto Y, Sado Y, Naito I, Yoshioka A (2005) Quantitative immunoelectron-microscopic analysis of the type IV collagen alpha1-6 chains in the glomerular basement membrane in childhood thin basement membrane disease. *Clin Nephrol* 64:329–336
- Nakopoulou L, Lazaris AC, Boletis IN, Michail S, Giannopoulou I, Zeis PM, Stathakis CP, Davaris PS (2002) Evaluation of E-cadherin/catenin complex in primary and secondary glomerulonephritis. *Am J Kidney Dis* 39:469–474
- Khoshnoodi J, Sigmundsson K, Ofverstedt LG, Skoglund U, Obrink B, Wartiovaara J, Tryggvason K (2003) Nephron promotes cell-cell adhesion through homophilic interactions. *Am J Pathol* 163:2337–2346
- Thongboonkerd V, Klein JB, Jevans AW, McLeish KR (2004) Urinary proteomics and biomarker discovery for glomerular diseases. *Contrib Nephrol* 141:292–307
- Sharma K, Lee S, Han S, Lee S, Francos B, McCue P, Wassell R, Shaw MA, RamachandraRao SP (2005) Two-dimensional fluorescence difference gel electrophoresis analysis of the urine proteome in human diabetic nephropathy. *Proteomics* 5:2648–2655
- Machii R, Sakatume M, Kubota R, Kobayashi S, Gejyo F, Shiba K (2005) Examination of the molecular diversity of alpha1 antitrypsin in urine: deficit of an alpha1 globulin fraction on cellulose acetate membrane electrophoresis. *J Clin Lab Anal* 19:16–21

Enrichment of the Basic/Cationic Urinary Proteome Using Ion Exchange Chromatography and Batch Adsorption

Visith Thongboonkerd^{*,†} Theptida Semangoen^{†,‡} and Somchai Chutipongtanate[†]

Medical Molecular Biology Unit, Office for Research and Development, and Department of Immunology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand

Received October 31, 2006

Abstract: Anionic or acidic proteins are the main compositions of normal urinary proteome. Efforts to characterize human urinary proteome, thus, have focused mainly on the anionic compartment. The information of cationic or basic proteins present in the normal urine is virtually unknown. In the present study, we applied different methods to enrich cationic urinary proteome. Efficacies of these methods were compared using equal volume (1 L) of urine samples from the same pool obtained from 8 normal healthy individuals. Cation exchange chromatography using RESOURCE-S column provided the least amount of the recovered proteins, whereas batch adsorption using SP Sepharose 4 Fast Flow beads equilibrated with acetic acid (pH 4.8) provided the greatest yield of protein recovery. The recovered proteins were then resolved with 2-DE (pI 7–11) and identified by peptide mass fingerprinting using MALDI-TOF MS. There were several isoforms of immunoglobulin heavy and light chains enriched by these methods. In addition, three isoforms of interferon alpha-3 (IFN α 3) and six isoforms of eosinophil-derived neurotoxin (EDN), were also enriched. The enrichment of IFN α 3 and EDN was particularly effective by batch adsorption using SP Sepharose 4 Fast Flow beads equilibrated with acetic acid (pH 6.0). Initial depletion of anionic components using DEAE batch adsorption reduced the recovery yield of these two proteins and did not improve recovery of any other cationic urinary proteins. We conclude that batch adsorption using SP Sepharose Fast Flow beads equilibrated with acetic acid (pH 6.0) is the method of choice to examine the basic/cationic urinary proteome, as this protocol provided the satisfactory yield of protein recovery and provided the greatest amount as well as maximal number of IFN α 3 and EDN isoforms. Our data will be useful for further highly focused study targeting on cationic/basic urinary proteins. Moreover, the techniques described herein may be applicable for enrichment of cationic proteomes in other body fluids, cells, and tissues.

Keywords: Proteomics • Proteome • Urine • Cationic urinary proteins • Ion exchange chromatography • Interferon alpha • Eosinophil-derived neurotoxin

Introduction

Proteomics has been widely applied onto clinical samples with the ultimate goal of discovery of novel biomarkers for more accurate and earlier diagnosis of human diseases, as well as for prediction and monitoring of therapeutic outcome.^{1,2} The urine is one of the most interesting resources for such study, as it is available in almost all of patients and its collection is noninvasive and very simple. However, one issue of concern in clinical applications of urinary biomarkers is that concentration and volume of the urine have a high degree of variation, based on time, activity, exercise, sweating, humidity, water drinking, caffeine, and so forth. Thus, normalization of the quantitative data with urine creatinine or using actual 24-h excretion levels may be required. Even with this sophistication in data interpretation, urinary proteomics has been extensively applied during the past 4–5 years.^{3,4}

An initial attempt, in parallel to technical development, has been identification or characterization of numerous proteins (proteome) that are present in the normal human urine. Two-dimensional gel electrophoresis (2-DE), although labor intensive, provides an advantage of the feasibility to create a reference proteome map to be used as a guide for subsequent 2-DE experiments. Recently, some partial 2-D proteome maps of the normal human urine have been constructed and are readily available.^{5–7} All of these maps, however, do not cover the entire urinary proteome, as they identified mostly anionic or acidic proteins. Most of these studies even used acidic-range immobilized pH gradient (IPG) strips (e.g., pI 4–7). Even using the broad-range IPG strips (e.g., pI 3–10), anionic proteins, which are the major abundant components, obscured the identification of cationic or basic urinary proteins. Although a previous study by Heine et al.⁸ applied solid-phase extraction and cation exchange chromatography to isolate urinary polypeptides/proteins, almost all of their identified proteins were acidic or anionic proteins. The information of cationic urinary proteins therefore remains mysterious, and unraveling this compartment may provide additional knowledge or better understanding of renal (patho)physiology.

The present study was conducted to enrich and characterize cationic proteins in the normal human urine using different methods, including RESOURCE-S ion exchange chromatography and batch adsorption using SP Sepharose 4 Fast Flow

* To whom correspondence should be addressed. Visith Thongboonkerd, MD, FRCPT, Medical Molecular Biology Unit, Office for Research and Development, 12th Floor Adulyadej Vikrom Building, Siriraj Hospital, 2 Prannok Road, Bangkoknoi, Bangkok 10700, Thailand. Phone/Fax: +66-2-4184793. E-mail: thongboonkerd@dr.com. Second e-mail: vthongbo@yahoo.com.

[†] Medical Molecular Biology Unit, Office for Research and Development, Faculty of Medicine Siriraj Hospital, Mahidol University.

[‡] Department of Immunology, Faculty of Medicine Siriraj Hospital, Mahidol University.

beads with various equilibration buffers. The recovered cationic proteins were then resolved with 2-DE (using pI 7–11 IPG strips) and identified by peptide mass fingerprinting using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS). Efficiencies (quantitative and qualitative) of these different methods were also compared. The results clearly showed that SP Sepharose Fast Flow batch adsorption equilibrated with acetic acid (pH 6.0) appeared to be the method of choice, as this protocol provided the satisfactory yield of protein recovery, and provided the greatest amount as well as maximal number of isoforms of interferon alpha-3 (IFN α 3) and eosinophil-derived neurotoxin (EDN).

Materials and Methods

Urine Collection. Urine samples were collected from 8 normal healthy individuals: 4 males (age 28.8 ± 3.9 years) and 4 females (age 25.8 ± 1.0 years), who had no recent medication. All females had no menstrual cycle at the time of collection. Cell debris and particulate matters were removed using low-speed centrifugation (1000g for 5 min), and supernatants were then pooled. This pool was then used for all subsequent experiments, each of which utilized equal volume (1 L) of the pooled urine.

Cation Exchange Chromatography Using RESOURCE-S Column. The pooled urine (1 L) was lyophilized until the volume was reduced to 4–5 mL. Cation exchange chromatography was then performed using RESOURCE S column (Amersham Biosciences, Uppsala, Sweden), which was run with 0.02 M 2-morpholinoethanesulfonic acid (MES) (pH 6.0) (Sigma-Aldrich, Inc., St. Louis, MO) at a flow rate of 1 mL/min. The bound proteins were eluted with 0.6 M NaCl, and the eluate was then desalted with HiTrap desalting column (Amersham Biosciences). Thereafter, the resultant was lyophilized to reduce the volume and to concentrate the recovered proteins.

Batch Adsorption Using SP Sepharose 4 Fast Flow Beads. Fresh pooled urine (1 L) was diluted (1:4) with 18.2 M Ω ·cm (dI) water, and its pH was adjusted to be the same as a buffer used in subsequent steps (4.8 or 6.0). SP Sepharose 4 Fast Flow beads (Amersham Biosciences), which had been initially equilibrated with 50 mM glacial acetic acid (pH 4.8) (Carlo Erba Reagenti, Rodano, Italy), 50 mM glacial acetic acid (pH 6.0), or 50 mM MES (pH 6.0), were added into the urine and stirred at 4 °C for 30 min. The urine–beads mixture was filtered through a sintered glass filter, and the cake was washed with the same equilibration buffer used initially until the filtrate became colorless. The adsorbed proteins were then eluted with 200 mL of the same buffer in the presence of 1 M NaCl. Thereafter, the eluates were dialyzed against 5 L of dI water (with two changes) for 24 h, and then lyophilized to reduce the volume and to concentrate the recovered proteins.

DEAE Adsorption To Eliminate Anionic Urinary Proteins Prior to Cationic Exchange. Because anionic proteins are the major abundant components in the urine, we also removed them by DEAE adsorption before cationic exchange to evaluate whether the preceding exclusion of anionic components can enhance the efficacy of cation exchange. Briefly, fresh pooled urine (1 L) was diluted (1:4) with dI water, and its pH was adjusted to 7.3. DEAE cellulose beads (DE-52, Whatman, Inc., Florham Park, NJ), which had been initially equilibrated with a buffer containing 0.01 M Tris-HCl and 0.05 M NaCl (pH 7.3), were added into the urine and stirred at room temperature for 30 min. The slurry was filtered through a sintered glass filter, and the cake was washed with 0.01 M Tris-HCl in 0.05 M NaCl

(pH 7.3) until the filtrate became colorless. The filtrate (or excluded solution) was then collected for further cationic exchange using SP Sepharose 4 Fast Flow beads as described above.

Two-Dimensional Gel Electrophoresis (2-DE). Immobiline DryStrip, pH 7–11, 7 cm long (Amersham Biosciences, Uppsala, Sweden) was rehydrated overnight with 50 μ g of total protein (equal to 5.59–53.41 mL of urine, based on protocols used) that was premixed with rehydration buffer containing 7 M urea, 2 M thiourea, 2% CHAPS (3-[(3-cholamidopropyl)dimethylamino]-1-propanesulfonate), 2% (v/v) ampholytes, 120 mM DTT, 40 mM Tris-base, and bromophenol blue (to make the final volume of 150 μ L per strip). The first-dimensional separation (IEF) was performed in Ettan IPGphor II IEF System (Amersham Biosciences) at 20 °C, using stepwise mode to reach 9000 Vh. After completion of the IEF, proteins on the strip were equilibrated with a buffer containing 6 M urea, 130 mM DTT, 30% glycerol, 112 mM Tris base, 4% SDS, and 0.002% bromophenol blue for 10 min, and then with another buffer containing 6 M urea, 135 mM iodoacetamide, 30% glycerol, 112 mM Tris base, 4% SDS, and 0.002% bromophenol blue for 10 min. The IPG strip was then transferred onto 12% acrylamide slab gel (8 $\times$ 9.5 cm), and the second-dimensional separation was performed in SE260 Mini-Vertical Electrophoresis Unit (Amersham Biosciences) with the current 20 μ A/gel for 1.5 h. Separated protein spots were then visualized using EMBL silver stain.⁹

Analysis and Matching of Protein Spots. Image Master 2D Platinum (Amersham Biosciences) software was used for matching and analysis of protein spots on 2-D gels. Parameters used for spot detection were (i) minimal area was 10 pixels; (ii) smooth factor was 2.0; and (iii) saliency was 2.0. A reference gel was created from an artificial gel combining all of the spots presenting in different gels into one image. The reference gel was then used for matching of corresponding protein spots between gels. Quantity of each protein was measured and reported as relative intensity volume (%Vol) (which was the intensity volume of each spot compared to the total intensity volume of all spots visualized within the same gel).

Protein Identification by Peptide Mass Fingerprinting Using MALDI-TOF MS. Protein spots were manually excised from the 2-D gels, and in-gel tryptic digestion was performed using techniques described previously.^{10,11} MALDI-TOF MS was performed on a reflectron MALDI-TOF mass spectrometer (Reflex IV; Bruker Daltonics; Leipzig, Germany). Peptide mass fingerprinting was performed to match resulting masses to theoretical peptide masses using the Mascot search tool (www.matrixscience.com). The National Center for Biotechnology Information (NCBI) protein database was utilized and restricted to mammalian entries, and peptides were assumed to be monoisotopic, oxidized at methionine residues, and carbamidomethylated at cysteine residues. Up to one missed trypsin cleavage was allowed. A mass tolerance of up to 100 ppm was allowed for matching peptide mass values. Probability-based MOWSE scores were calculated by a comparison of search results against an estimated random match population [scores = $-10 \cdot \log_{10}(P)$, where P was the absolute probability]. Scores >69 were considered statistically significant ($P < 0.05$).

2-D Western Blotting. Totally, 50 μ g of urinary proteins, enriched by batch adsorption using SP Sepharose Fast Flow beads equilibrated and run with acetic acid (pH 4.8), was resolved by 2-DE as described above. After completion of the

Table 1. Amounts of Cationic Urinary Proteins Obtained from Different Protocols

protocol	volume of urine used	amount of cationic proteins obtained
A. RESOURCE-S ion exchange chromatography (with MES, pH 6.0)	1 L	111.70 μ g
B. SP Sepharose Fast Flow batch adsorption (with MES, pH 6.0)	1 L	300.20 μ g
C. SP Sepharose Fast Flow batch adsorption (with acetic, pH 4.8)	1 L	1,068.14 μ g
D. SP Sepharose Fast Flow batch adsorption (with acetic, pH 6.0)	1 L	315.55 μ g
E. SP Sepharose Fast Flow batch adsorption (with acetic, pH 4.8) with prior DEAE anionic exclusion (1 cycle)	1 L	386.76 μ g
F. SP Sepharose Fast Flow batch adsorption (with acetic, pH 4.8) with prior DEAE anionic exclusion (3 cycles)	1 L	262.50 μ g

2-DE separation, the resolved proteins were transferred onto a nitrocellulose membrane, and nonspecific bindings were blocked with 5% milk in PBS for 1 h. The membrane was then incubated with mouse anti-human EDN (eosinophil-derived neurotoxin) polyclonal antibody (1:2500 in 5% milk in PBS) (Abnova, Taipei, Taiwan) at 4 °C overnight. After a washing step, the membrane was further incubated with goat anti-mouse IgG conjugated with horseradish peroxidase (1:4000 in 5% milk in PBS) (Dako, Glostrup, Denmark) at room temperature for 1 h. Reactive protein spots were then visualized with SuperSignal West Pico chemiluminescence substrate (Pierce Biotechnology, Inc., Rockford, IL). Equal amount of the nonenriched urine sample (concentrated by lyophilization) was also subjected to 2-D Western blotting using the identical protocol as described above to compare with the enriched sample.

Results and Discussion

Totally, 6 protocols (one used RESOURCE-S cation exchange chromatography, whereas other five utilized SP Sepharose Fast Flow batch adsorption) were performed to enrich cationic urinary proteins from normal healthy individuals. Urine samples from these eight individuals were pooled to avoid interindividual variabilities. The efficacies for recovering cationic proteins of these protocols were compared (Table 1). RESOURCE-S ion exchange chromatography with MES (pH 6.0) provided the least amount of the recovered proteins, whereas batch adsorption using SP Sepharose Fast Flow beads equilibrated and run with acetic acid (pH 4.8) provided the greatest amount of proteins recovered from equal volume (1 L) of the pooled urine. When buffers other than MES were used in RESOURCE-S ion exchange chromatography, recovery yields were unsatisfactory and much less than when using MES. Therefore, only one protocol of RESOURCE-S ion exchange chromatography was further investigated. For batch adsorption using SP Sepharose Fast Flow beads, equilibration and running with MES (pH 6.0) provided a comparable yield compared with using acetic acid at the same pH level (6.0). DEAE adsorption reduced the yield of protein recovery.

While Table 1 shows only quantitative data, Figure 1 demonstrates both quantitative and qualitative data of individual enrichment protocols. Proteins recovered from individual protocols (with the equal amount of 50 μ g per gel) were resolved in each 2-D gel using basic-range IPG strip (pI 7–11).

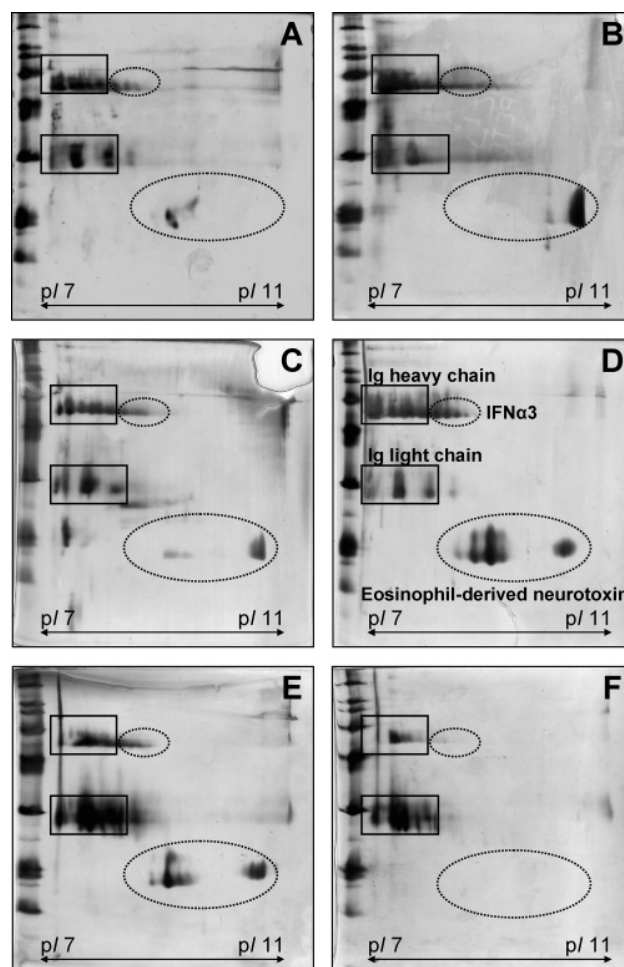


Figure 1. 2-D proteome profile of cationic urinary proteins enriched with different protocols. Proteins derived from the pooled urine using each enrichment protocol (with the equal amount of 50 μ g per gel) were resolved in each 2-D gel using 7-cm-long basic IPG strips (pI 7–11). The separated protein spots were then visualized with EMBL silver stain. (A) RESOURCE-S ion exchange chromatography (with MES, pH 6.0); (B) SP Sepharose Fast Flow batch adsorption (with MES, pH 6.0); (C) SP Sepharose Fast Flow batch adsorption (with acetic acid, pH 4.8); (D) SP Sepharose Fast Flow batch adsorption (with acetic acid, pH 6.0); (E) SP Sepharose Fast Flow batch adsorption (with acetic acid, pH 4.8) following DEAE adsorption (1 cycle); and (F) SP Sepharose Fast Flow batch adsorption (with acetic acid, pH 4.8) following DEAE adsorption (3 cycles).

The protein spots were visualized with EMBL silver stain and were excised for subsequent in-gel tryptic digestion and identification by peptide mass fingerprinting using MALDI-TOF MS. These protein spots were identified as several isoforms of four unique proteins, including immunoglobulin heavy chain, immunoglobulin light chain, interferon alpha 3 (IFN α 3) (gi|53680582; theoretical pI = 7.59), and eosinophil-derived neurotoxin (EDN) (gi|790266; theoretical pI = 9.23). Although immunoglobulin heavy and light chains have been readily identified in most of 2-DE studies of normal human urine, their enrichment using our techniques may also be beneficial for the investigation of diseases such as gammopathies, in which immunoglobulin heavy chain or light chain, or both, are involved in the pathogenic mechanisms.

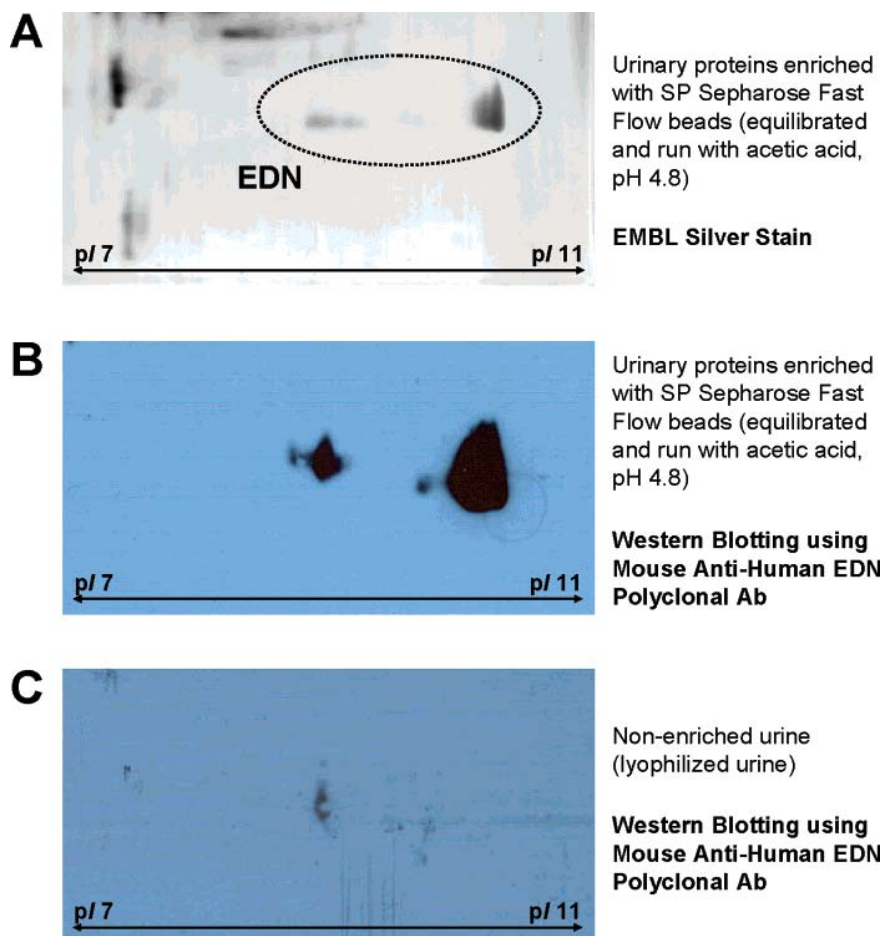


Figure 2. 2-D Western blot analysis of EDN. Totally, 50 μ g of proteins was resolved in each 2-D gel. (A and B) Proteins enriched by batch adsorption using SP Sepharose Fast Flow beads equilibrated and run with acetic acid (pH 4.8). They were either visualized with silver stain (A) or subjected to Western blot analysis for EDN (B) (detailed in Materials and Methods). Western blot analysis revealed immunoreactive EDN spots, which corresponded to the EDN spots identified by peptide mass fingerprinting using MALDI-TOF MS. (C) 2-D Western blot of the nonenriched urine simply concentrated by lyophilization was also performed to compare with that of the enriched sample using an equal amount of total protein loaded in each gel. The data clearly confirmed that EDN was successfully enriched by batch adsorption using SP Sepharose Fast Flow beads.

To confirm the proteomic data, we performed 2-D Western blot analysis of urinary cationic proteins, which were enriched by batch adsorption using SP Sepharose Fast Flow beads equilibrated and run with acetic acid (pH 4.8). When mouse anti-human EDN polyclonal antibody is used, Figure 2B clearly illustrates the immunoreactive spots of EDN, which corresponded to the EDN spots identified by peptide mass fingerprinting using MALDI-TOF MS in a 2-D gel stained with EMBL silver (Figure 2A). Additionally, we compared the 2-D Western blot of the enriched sample (Figure 2B) to that of the nonenriched urine simply concentrated by lyophilization (Figure 2C), using equal amount (50 μ g) of total protein loaded in each gel. The data clearly confirmed that EDN was successfully enriched by batch adsorption using SP Sepharose Fast Flow beads.

Interestingly, summation of relative intensity volume ($\Sigma\%$ Vol) of all three isoforms of IFN α 3 was greatest ($\Sigma\%$ Vol = 9.53) in the sample derived from the protocol using SP Sepharose Fast Flow batch adsorption with acetic acid (pH 6.0), but was lowest ($\Sigma\%$ Vol = 0.78) when DEAE (3 cycles) was applied prior to SP Sepharose Fast Flow batch adsorption with acetic acid (pH 4.8) (Figure 3). For EDN, the protocol using SP Sepharose Fast Flow

batch adsorption with acetic acid (pH 6.0) yielded six isoforms with the highest $\Sigma\%$ Vol (41.66), whereas only two isoforms were detected when MES was used instead of acetic acid, and only three isoforms were recovered from RESOURCE-S ion exchange chromatography (Figure 4). The recovery of this protein was also reduced by DEAE anionic exclusion prior to SP Sepharose Fast Flow cationic adsorption.

IFN α is a member of Type I IFNs family (which includes IFN α , IFN β , IFN ω , and IFN τ). Type I IFNs have been shown to induce expression of a large number of genes involved in regulating a wide variety of important biological responses, including antiviral, antiproliferative, and immunomodulating activities.¹² The classical function of IFN α is the protection against viral infections. However, after the ability of IFN α to inhibit tumor cell proliferation has been established, recombinant IFN α (the first cytokine produced by recombinant DNA technology) has been emerged as a widely used medicine in the therapy of several solid tumors and hematologic malignancies (i.e., multiple myeloma, hairy cell leukemia, and chronic myeloid leukemia). Even with the wide applications, its anti-cancer effects remain unclear. One of the potential mechanisms has been thought to involve induction of apoptosis in cancer

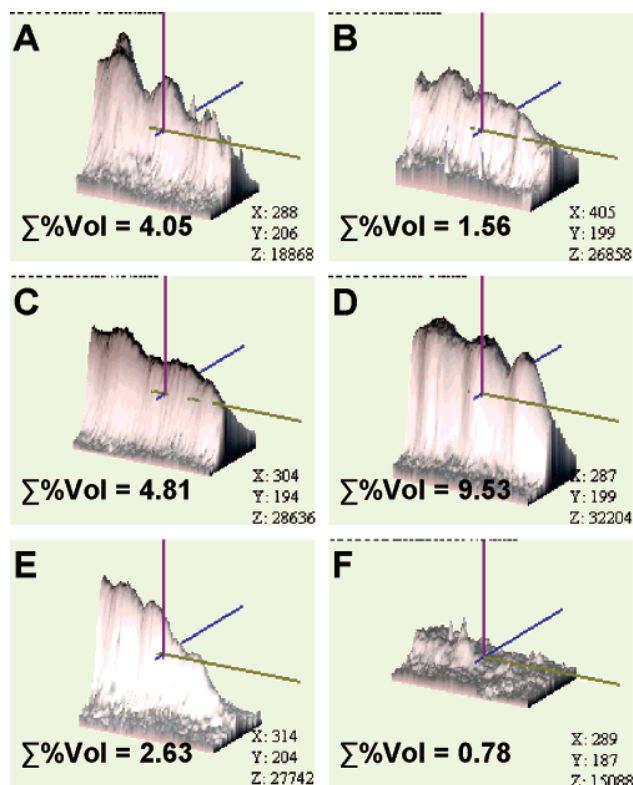


Figure 3. Zoom-in, 3-D view of IFN α spots. (A) RESOURCE-S ion exchange chromatography (with MES, pH 6.0); (B) SP Sepharose Fast Flow batch adsorption (with MES, pH 6.0); (C) SP Sepharose Fast Flow batch adsorption (with acetic acid, pH 4.8); (D) SP Sepharose Fast Flow batch adsorption (with acetic acid, pH 6.0); (E) SP Sepharose Fast Flow batch adsorption (with acetic acid, pH 4.8) following DEAE adsorption (1 cycle); and (F) SP Sepharose Fast Flow batch adsorption (with acetic acid, pH 4.8) following DEAE adsorption (3 cycles). $\Sigma\%Vol$ denotes summation of the relative intensity volumes, which indicate the quantity of all isoforms of IFN α .

cells via signal transducer and activator of transcription (STAT)-dependent mechanisms.¹³ In contrast to the therapeutic properties, IFN α can also induce autoimmunity,¹⁴ particularly in systemic lupus erythematosus (SLE) and IgA nephropathy.^{15,16} IFN α had not previously been identified in the normal human urine in any of the previous urinary proteome studies. The enrichment of IFN α may allow rapid evaluation of urinary IFN α as a novel biomarker or a new therapeutic target in a wide variety of kidney and nonkidney diseases.

EDN, also known as eosinophil protein X (EPX) and eosinophilic cationic protein (ECP), is a nonsecretory ribonuclease that can possess a wide variety of biological activities. Other synonyms include ribonuclease 2, ribonuclease US, RNase 2, and RNase Upl-2. Its locations include liver, lungs, spleen, leukocytes, and body fluids. EDN has been widely accepted as a marker for eosinophil-induced allergy and eosinophilic inflammatory diseases, such as inflammatory bowel disease.¹⁷ A recent study by Ye et al.¹⁸ has demonstrated the increase in glycosylated form of EDN in the urine of ovarian cancer patients as compared to the urine from normal healthy individuals and from patients with benign ovarian cysts or tumors. Interestingly, urinary EDN has been shown to have antitumor and antiangiogenic activities.¹⁹ All of the previous studies that characterized urinary EDN utilized conventional

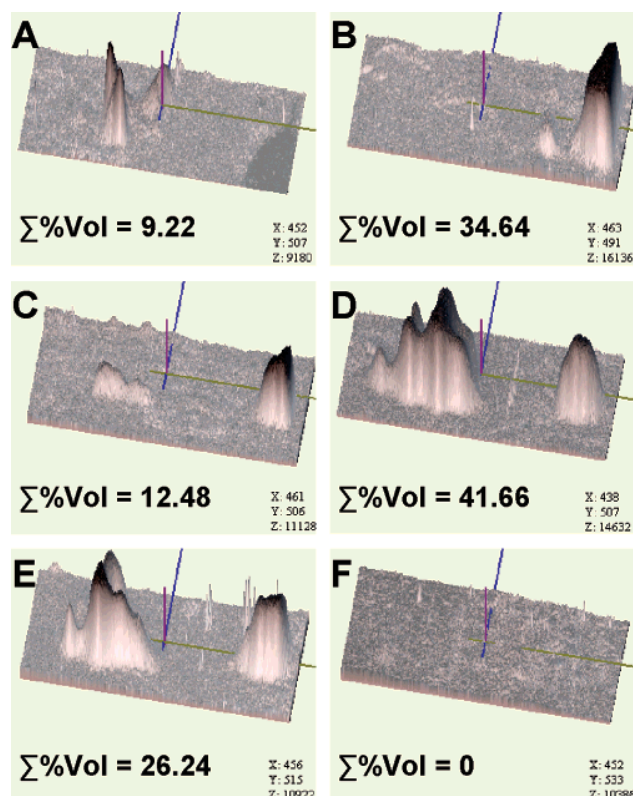


Figure 4. Zoom-in, 3-D view of EDN spots. (A) RESOURCE-S ion exchange chromatography (with MES, pH 6.0); (B) SP Sepharose Fast Flow batch adsorption (with MES, pH 6.0); (C) SP Sepharose Fast Flow batch adsorption (with acetic acid, pH 4.8); (D) SP Sepharose Fast Flow batch adsorption (with acetic acid, pH 6.0); (E) SP Sepharose Fast Flow batch adsorption (with acetic acid, pH 4.8) following DEAE adsorption (1 cycle); and (F) SP Sepharose Fast Flow batch adsorption (with acetic acid, pH 4.8) following DEAE adsorption (3 cycles). $\Sigma\%Vol$ denotes summation of the relative intensity volumes, which indicate the quantity of all isoforms of EDN.

biochemical methods or immunologic assays. EDN has also been identified in the normal human urine by Spahr and colleagues²⁰ using liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS). However, EDN had not previously been identified in the normal human urine using 2-DE-based proteomics approach. Identification of EDN in a 2-D proteome map would be beneficial, as further sophisticated experiments using mass spectrometry may not be required for subsequent study of EDN using 2-DE. We identified, for the first time, six isoforms of urinary EDN in the normal urine using 2-DE followed by peptide mass fingerprinting. This finding was a result of enrichment of EDN, particularly by SP Sepharose Fast Flow batch adsorption with acetic acid (pH 6.0). The successful enrichment of EDN may allow rapid evaluation of urinary EDN as a novel biomarker or new therapeutic target in allergy, inflammatory diseases, cancers, and angioproliferative disorders. The presence of six isoforms of EDN in the normal urine may imply its post-translational modifications and deserves further investigation, which may lead to the identification of modification-specific EDN marker.

It should be noted that the total number of cationic proteins successfully enriched by our methods was quite few. A question has been raised whether these proteins represent all the cationic components in the normal human urinary proteome.

Recently, Soldi et al.²¹ utilized 2-D liquid-phase fractionation, based on chromatofocusing for the first dimension and hydrophobicity for the second, to fractionate human urinary proteome. Although several spectra were revealed at the fractions with pH > 7.0, only one protein (prostaglandin-H2 D-isomerase precursor) was a “real” basic protein identified in these fractions. Other two proteins in these cationic fractions that were successfully identified by subsequent peptide mass fingerprinting using MALDI-TOF MS, including CD59 fragment and apolipoprotein A-I, are indeed the “acidic” proteins.²¹ These results suggest that there are only few cationic proteins in the normal human urine. Enrichment of these proteins may be beneficial to the highly focused study targeting on urinary biomarkers or new therapeutic targets in some illnesses, particularly gammopathies as well as autoimmune and inflammatory diseases.

In summary, we evaluated and compared the efficacies of different methods to enrich cationic urinary proteome. Batch adsorption using SP Sepharose Fast Flow beads equilibrated with acetic acid (pH 6.0) appeared to be the method of choice, as this protocol provided the satisfactory yield of protein recovery and provided the greatest amount as well as maximal number of IFN α 3 and EDN isoforms. Our methods may be applicable also for enrichment of cationic proteomes in other body fluids, cells, and tissues.

Abbreviations: 2-DE, two-dimensional gel electrophoresis; CHAPS, 3-[(3-cholamidopropyl)dimethylamino]-1-propanesulfonate; EDN, eosinophil-derived neurotoxin; IFN α 3, interferon alpha-3; IPG, immobilized pH gradient; MES, 2-morpholinoethanesulfonic acid.

Acknowledgment. We thank Rattiyaporn Kanlaya and Napat Songtawee for their technical assistance. This study was supported by Mahidol University, the National Research Council of Thailand, and the National Center for Genetic Engineering and Biotechnology (BIOTEC) (Project No. BT-B-02-MG-B4-4908) to V. Thongboonkerd.

References

- Colantonio, D. A.; Chan, D. W. The clinical application of proteomics. *Clin. Chim. Acta* **2005**, *357*, 151–158.
- Hanash, S. HUPO initiatives relevant to clinical proteomics. *Mol. Cell. Proteomics* **2004**, *3*, 298–301.
- Thongboonkerd, V.; Malasit, P. Renal and urinary proteomics: Current applications and challenges. *Proteomics* **2005**, *5*, 1033–1042.
- Thongboonkerd, V. Proteomics in Nephrology: Current status and future directions. *Am. J. Nephrol.* **2004**, *24*, 360–378.
- Thongboonkerd, V.; McLeish, K. R.; Arthur, J. M.; Klein, J. B. Proteomic analysis of normal human urinary proteins isolated by acetone precipitation or ultracentrifugation. *Kidney Int.* **2002**, *62*, 1461–1469.
- Oh, J.; Pyo, J. H.; Jo, E. H.; Hwang, S. I.; Kang, S. C.; Jung, J. H.; Park, E. K.; Kim, S. Y.; Choi, J. Y.; Lim, J. Establishment of a near-standard two-dimensional human urine proteomic map. *Proteomics* **2004**, *4*, 3485–3497.
- Pieper, R.; Gatlin, C. L.; McGrath, A. M.; Makusky, A. J.; Mondal, M.; Seonarain, M.; Field, E.; Schatz, C. R.; Estock, M. A.; Ahmed, N.; Anderson, N. G.; Steiner, S. Characterization of the human urinary proteome: a method for high-resolution display of urinary proteins on two-dimensional electrophoresis gels with a yield of nearly 1400 distinct protein spots. *Proteomics* **2004**, *4*, 1159–1174.
- Heine, G.; Raida, M.; Forssmann, W. G. Mapping of peptides and protein fragments in human urine using liquid chromatography-mass spectrometry. *J. Chromatogr., A* **1997**, *776*, 117–124.
- Shevchenko, A.; Wilm, M.; Vorm, O.; Mann, M. Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal. Chem.* **1996**, *68*, 850–858.
- Thongboonkerd, V.; Luengpailin, J.; Cao, J.; Pierce, W. M.; Cai, J.; Klein, J. B.; Doyle, R. J. Fluoride exposure attenuates expression of *Streptococcus pyogenes* virulence factors. *J. Biol. Chem.* **2002**, *277*, 16599–16605.
- Thongboonkerd, V.; Gozal, E.; Sachleben, L. R.; Arthur, J. M.; Pierce, W. M.; Cai, J.; Chao, J.; Bader, M.; Pesquero, J. B.; Gozal, D.; Klein, J. B. Proteomic analysis reveals alterations in the renal kallikrein pathway during hypoxia-induced hypertension. *J. Biol. Chem.* **2002**, *277*, 34708–34716.
- Caraglia, M.; Vitale, G.; Marra, M.; Del Prete, S.; Lentini, A.; Budillon, A.; Beninati, S.; Abbruzzese, A. Translational and post-translational modifications of proteins as a new mechanism of action of alpha-interferon: review article. *Amino Acids* **2004**, *26*, 409–417.
- Caraglia, M.; Marra, M.; Pelaia, G.; Maselli, R.; Caputi, M.; Marsico, S. A.; Abbruzzese, A. Alpha-interferon and its effects on signal transduction pathways. *J. Cell. Physiol.* **2005**, *202*, 323–335.
- Conrad, B. Potential mechanisms of interferon-alpha induced autoimmunity. *Autoimmunity* **2003**, *36*, 519–523.
- Crow, M. K.; Kirou, K. A. Interferon-alpha in systemic lupus erythematosus. *Curr. Opin. Rheumatol.* **2004**, *16*, 541–547.
- Wardle, E. N. Is IgA nephropathy induced by hyperproduction of interferon-alpha? *Med. Hypotheses* **2004**, *62*, 625–628.
- Inamura, H.; Tomita, M.; Okano, A.; Kurosawa, M. Serial blood and urine levels of EDN and ECP in eosinophilic colitis. *Allergy* **2003**, *58*, 959–960.
- Ye, B.; Skates, S.; Mok, S. C.; Horick, N. K.; Rosenberg, H. F.; Vitonis, A.; Edwards, D.; Sluss, P.; Han, W. K.; Berkowitz, R. S.; Cramer, D. W. Proteomic-based discovery and characterization of glycosylated eosinophil-derived neurotoxin and COOH-terminal osteopontin fragments for ovarian cancer in urine. *Clin. Cancer Res.* **2006**, *12*, 432–441.
- Pati, S.; Lee, Y.; Samaniego, F. Urinary proteins with pro-apoptotic and antitumor activity. *Apoptosis* **2000**, *5*, 21–28.
- Spahr, C. S.; Davis, M. T.; McGinley, M. D.; Robinson, J. H.; Bures, E. J.; Beierle, J.; Mort, J.; Courchesne, P. L.; Chen, K.; Wahl, R. C.; Yu, W.; Luethy, R.; Patterson, S. D. Towards defining the urinary proteome using liquid chromatography-tandem mass spectrometry. I. Profiling an unfractionated tryptic digest. *Proteomics* **2001**, *1*, 93–107.
- Soldi, M.; Sarto, C.; Valsecchi, C.; Magni, F.; Proserpio, V.; Ticozzi, D.; Mocarelli, P. Proteome profile of human urine with two-dimensional liquid phase fractionation. *Proteomics* **2005**, *5*, 2641–2647.

PR0605771

Advances in Urinary Proteome Analysis and Biomarker Discovery

Danilo Fliser,^{*} Jan Novak,[†] Visith Thongboonkerd,[‡] Àngel Argilés,[§] Vera Jankowski,^{||} Mark A. Girolami,[¶] Joachim Jankowski,^{||} and Harald Mischak^{**}

^{*}Medical School Hannover, Hannover, Germany; [†]University of Alabama at Birmingham, Birmingham, Alabama;

[‡]Faculty of Medicine at Siriraj Hospital, Mahidol University, Bangkok, Thailand; [§]Department of Nephrology, University Hospital, Néphrologie et Dialyse St Guilhem and SAS RD, Néphrologie, Montpellier, France; ^{||}Nephrology Department, Charité Hospital, Berlin, Germany; [¶]Department of Computing Science, University of Glasgow, Glasgow, Scotland; and ^{**}Mosaiques-Diagnostics and Therapeutics AG, Hannover, Germany

Noninvasive diagnosis of kidney diseases and assessment of the prognosis are still challenges in clinical nephrology. Definition of biomarkers on the basis of proteome analysis, especially of the urine, has advanced recently and may provide new tools to solve those challenges. This article highlights the most promising technological approaches toward deciphering the human proteome and applications of the knowledge in clinical nephrology, with emphasis on the urinary proteome. The data in the current literature indicate that although a thorough investigation of the entire urinary proteome is still a distant goal, clinical applications are already available. Progress in the analysis of human proteome in health and disease will depend more on the standardization of data and availability of suitable bioinformatics and software solutions than on new technological advances. It is predicted that proteomics will play an important role in clinical nephrology in the very near future and that this progress will require interactive dialogue and collaboration between clinicians and analytical specialists.

J Am Soc Nephrol 18: 1057–1071, 2007. doi: 10.1681/ASN.2006090956

As early as in the 17th century, physicians were already performing a sort of “proteomics” when looking at the urine and appreciating the foaming in the supernatant as an indirect sign for pathologic proteinuria (Figure 1). A new era began with the introduction of electrophoresis, which allowed separation and detection of potentially distinct proteins (1). In 1975, O’Farrell (2) described the separation of proteins of *Escherichia coli* by two-dimensional gel electrophoresis, but he was still unaware that he was performing proteomics. With the availability of modern mass spectrometers for the analysis of proteins and the development of sophisticated means to evaluate and compare the vast amount of information generated, the era of proteomics started with soaring hope, following the footsteps of “genomics.” Soon, many challenges became evident (e.g., as a result of the complexity of the proteins that are generated by alternative processing of mRNA and posttranslational modifications). These issues and potential solutions were recently addressed (3).

Regardless of these obstacles, proteomics has been used in clinical medicine, including nephrology, since the early developmental stages. Several of the basic considerations with respect to the use of proteomics in nephrology and the discovery of protein biomarkers for kidney diseases were summarized

recently (4–6). In this review, we provide an overview of new developments, cover the most promising technical aspects of different approaches to proteome analysis, and examine the inherent technical advantages and limitations. In the second part of the review, we focus on proteome analysis of the urine and its expectations and challenges and the drive behind the efforts: An unfulfilled need for clinically useful biomarker assays that allow the early diagnosis of kidney diseases and assessment of the patient’s prognosis and response to therapy.

What Do Proteome and Proteomics Entail?

Proteins and peptides within a particular body compartment comprise a proteome. In contrast to the genome, which is unique and relatively stable (most cells of an organism possess the total genomic capital during life), proteomes are cell and tissue specific and change over time in response to different situations. Proteomics (*i.e.*, the assessment of these proteomes) may contribute to the elucidation of the proteome of a given cell type (e.g., epithelial, mesangial, endothelial), tissue (e.g., renal cortex), or even specific parts of the tissue (e.g., glomerulus). Furthermore, proteomics can reveal proteome changes in biologic fluids (e.g., urinary or plasma proteome). The term “peptidome” has been used for a “subset” of the proteome, the lower molecular weight peptides in a sample. Because of these somewhat ill-defined differences and because the invention of additional terms does not improve comprehension of an already complicated matter, in this review we use the term “protein” for all naturally occurring polypeptides (poly-amino acids).

Published online ahead of print. Publication date available at www.jasn.org.

Address correspondence to: Dr. Harald Mischak, Mosaiques Diagnostics & Therapeutics AG, Mellendorfer Strasse 7-9, D-30625 Hannover, Germany. Phone: +49-511-554744-13; Fax: +49-511-554744-31; E-mail: mischak@mosaiques-diagnostics.com



The Physician

Gerrit DOU,
Leiden (1613–1675)

Figure 1. The Physician. Painting by the Dutch painter Gerrit Dou (Leiden 1613 to 1675), describing the meticulous observation of urine by a physician of the 17th century.

Two basic sources of material are available for proteomic studies: Body fluids (*e.g.*, urine, blood) and tissue. Proteome research related to nephrology has generally focused on the examination of urine because it is easily accessible in a large quantity without the use of invasive procedures. Furthermore, as a rule, pathophysiologic changes in the genitourinary tract and the kidney are reflected by changes in the urinary proteome. Although many studies have shown that proteins in biologic fluids may degrade rapidly when handled inappropriately, urinary proteins have been shown to remain stable long enough to perform reliable proteome analysis. In two independent sets of experiments, Schaub *et al.* (7) and Theodorescu *et al.* (8) showed that the urinary proteome did not undergo significant changes when urine was stored for 3 d at 4°C or 6 h at room temperature, respectively. In addition, urine can be stored for several years, even at –20°C, without significant alterations in its proteome. Although these reports suggest a much greater stability of the urinary proteome compared with the blood proteome, it must be clearly noted that other issues will influence data quality and comparability. Given the complexity of the urinary proteome that we can currently only estimate, there are likely to be important changes associated with differences among samples as a result of variations in procedures for collection, storage, and, of course, processing. As outlined in more detail elsewhere (3), these issues must be taken into account, and standardized protocols for urine sampling and for handling of the samples should be adopted.

In contrast to urine, collection of blood is invasive and requires meticulous preanalytical handling. Its proteomic analysis is prone to analytical artifacts. A detailed comparison of serum and plasma proteomes revealed that an array of proteases are activated immediately upon clotting, resulting in the generation of many degradation products (9). As a conse-

quence, the human proteome consortium has recommended that blood be examined as plasma rather than as serum and established a standardized sample collection protocol (10). Similar and absolutely essential efforts for standardization of collection and processing are under way for urine and would certainly be desirable for various types of tissue.

Technical Aspects of Proteomics

Because of the inherent complexity of a proteome, all approaches for its examination generally rely on a separation step, followed by ionization and subsequent mass spectrometry (MS) analysis. This complexity is further increased by posttranslational modifications (PTM), which may well serve as biomarkers for disease (*e.g.*, advanced glycation in diabetes [11]), hence must be considered. Because of their ability to analyze mass, in general all MS-based proteomics technologies will identify PTM, because these result in a change in mass. Furthermore, PTM frequently result in changes in the migration in any of the pre-MS separation approaches described in more detail herein. It is beyond the scope of this review to outline the differences in the ionization processes and the modern mass spectrometers; these topics have been summarized elsewhere (9). In general, quadrupole (Q), ion-trap, time-of-flight (TOF), and Fourier transform-ion cyclotron resonance (FT-ICR) instruments or their combinations (*e.g.*, hybrid instruments such as Q-TOF, combining quadrupole and time-of-flight detectors) are currently used for proteome analysis. To obtain sequence information (as well as information on PTM), sequential use of these techniques, termed tandem mass spectrometry (MS/MS), is used. In general, the first MS instrument serves as a mass filter, “collecting” only the ions with the mass of interest (“parent ions”), and the second MS instrument is used to analyze fragmentation products (“daughter ions”) that may be generated by collision with other molecules (collision-induced dissociation) or transfer of electrons (electron transfer dissociation). Frequently, individual advantages of the different mass detectors are combined (*e.g.*, the precision of quadrupoles and the high accuracy of the measurements of mass with TOF in a Q-TOF instrument). This approach cannot be easily applied to sequencing large proteins (top-down approach); instead, these proteins are usually digested by proteases, such as trypsin. The resulting smaller fragments are then sequenced with good success. Recent technical advances enabled sequencing of large native peptides and proteins up to 66 kD (12–14). In general, every pre-MS separation technology (with the exception of surface-enhanced laser desorption/ionization [SELDI]) can be combined with any MS approach. The highest resolution and the best mass accuracy (deviation of the measured mass from the theoretical mass) can be obtained by using FT-ICR instruments (<1 ppm). Unfortunately, the high cost of these instruments has limited their widespread use. Alternatively, any other mass spectrometer that is capable of delivering high-quality data, with mass deviation <50 ppm and resolution >5000 would be acceptable for proteomic applications. However, these values should probably be regarded as the lowest standards for accurate proteome analysis.

All MS instruments analyze a mass/charge ratio (m/z) and

consequently require ions for analysis. Ionization of the compound of interest can be achieved using matrix-assisted laser desorption ionization (MALDI) or electro-spray ionization (ESI). For MALDI, the sample is mixed with a matrix, spotted onto a target, and ionization is achieved with a pulsed laser. Matrix absorbs the energy of the laser, transfers it to the analyte, and thereby assists in its ionization. This method is generally applied off-line (separation is not coupled to a mass spectrometer). It is technically less demanding, but it is susceptible to “signal suppression,” a phenomenon whereby certain analytes are preferentially ionized and more easily detected at the expense of other compounds that may even become undetectable (15,16). ESI generates charged droplets in a high-voltage field; solvent from these charged droplets evaporates, giving rise to multiply charged ions of the analyte. This approach is generally used on-line, is less stable, but also is less susceptible to signal suppression.

Proteomics that aim toward biomarker discovery can be considered a sophisticated comparative analysis. Consequently, quantification of proteins is an essential consideration. For gels, protein stains are generally used for quantification. The newer fluorescence stains seem to provide superior results, because they have large linear dynamic ranges and similar or better sensitivity compared with the older Coomassie blue and silver stains (17). There are also good MS methods for relative quantification on the basis of ion counting (18), but care in instrument and technical consistency is essential. Absolute quantification usually requires previous identification of the biomarker sequence and/or chemical derivation that consequently may become restrictive (19). Relative quantification of biomarker abundances with reference to constant peaks generally seems sufficient, especially when considering biologic variation.

Four different approaches that frequently are used for pre-MS fractionation are briefly reviewed because they provide very different types and quality of data. For a comparative evaluation of the potential of these MS-based approaches for clinical applications, the most prominent advantages and disadvantages are listed in Table 1.

Two-Dimensional Gel Electrophoresis Followed by Mass Spectrometry

SDS-PAGE reported by Laemmli (20) and subsequently the two-dimensional gel electrophoresis (2DE) method reported by O’Farrell in 1975 (2) laid the foundation for what we understand today as proteomics. The technique reproducibly separates proteins according to two intrinsic characteristics: Isoelectric point and molecular mass. Separation is accomplished in two steps; proteins are first fractionated by electrofocusing (proteins migrate to their isoelectric point in a pH gradient) and then in a perpendicular dimension by SDS-PAGE (proteins migrate on the basis of their molecular mass). Examples for such separations are given in Figure 2A. A characteristic protein pattern (composed of many “spots”) is produced. After separation in a 2DE gel, the challenge of identification of these spots remains. Whenever possible, immune detection (using antibodies in Western blot assays) is used, as described by Burnette (21).

The gel separation and blotting techniques have been successfully used for more than 30 yr and are still widely popular. A major advancement in the identification of proteins was achieved with the implementation of MS. The first step in the process of identifying proteins in the spots (2DE gel) or bands (1DE gel) by MS is a proteolytic in-gel digestion (*e.g.*, by exposing the excised pieces of gel to trypsin) (22,23) followed by extraction of the proteolytic fragments from the gel. Masses of at least three proteolytic fragments are needed for identification of a protein from a database of proteins (Figure 2B). The identity of a match can be subsequently verified by MS/MS sequencing or by other techniques, such as Western blotting (if a specific antibody is available). Limitations of the 2DE approach include low reproducibility, the considerable time for the analysis, and the difficulty to automate the process. Recently, the concept of 2D difference gel electrophoresis (2D-DIGE) was introduced to reduce gel-to-gel variability. Briefly, two samples are differentially labeled with fluorescence dyes (Cy3 and Cy5), and the two samples are then resolved simultaneously within the same 2DE gel (Figure 2C). This technique also allows intro-

Table 1. Advantages and disadvantages of various mass spectrometry–based proteomic techniques for clinical applications^a

Technology	Advantages	Limitations
2DE-MS	Applicable to large molecules	Not applicable to peptides <10 kD, no automation, time-consuming, quantification difficult, expensive
LC-MS	Automation, multidimensional, high sensitivity	Time-consuming, sensitive toward interfering compounds, restricted mass range
SELDI-MS	Easy-to-use system, high throughput, automation, low sample volume	Restricted to selected proteins, low-resolution MS, low information content, lack of comparability
CE-MS	Automation, high sensitivity, fast, low sample volumes, multidimensional, low cost	Not suited for larger proteins (>20 kD)

^a2DE-MS, two-dimensional gel electrophoresis followed by mass spectrometry; LC-MS, liquid chromatography coupled to mass spectrometry; SELDI-MS, surface-enhanced laser desorption/ionization coupled to mass spectrometry; CE-MS, capillary electrophoresis coupled to mass spectrometry.

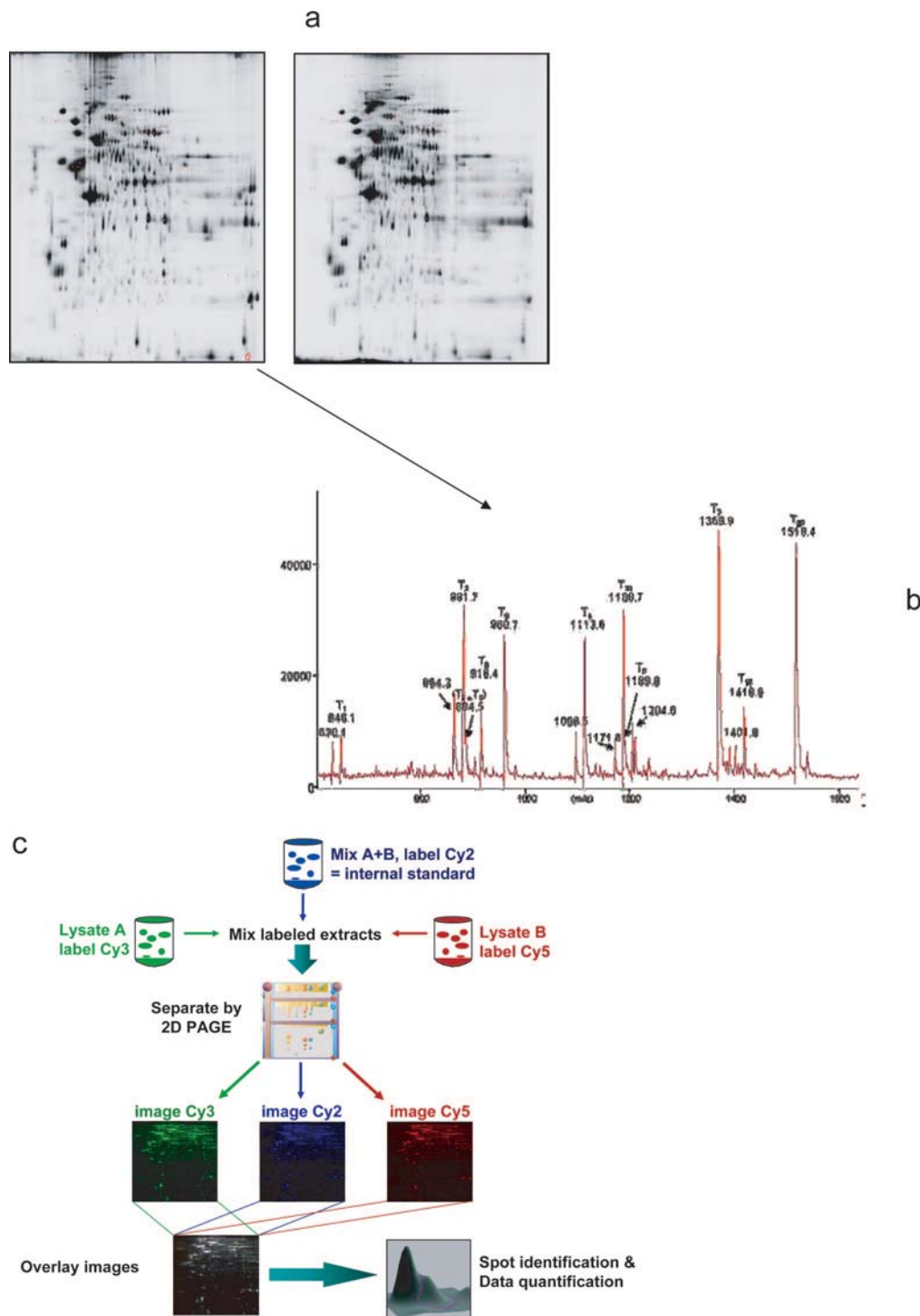


Figure 2. Two-dimensional gel electrophoresis coupled to mass spectrometry (2DE-MS). (A) Proteins in samples from different individuals are separated in two dimensions according to the isoelectric point and molecular weight. The resulting 2DE gels are stained and compared. (B) Protein spots that seem to differ between the two gels are excised and digested with trypsin, and the resulting peptides are analyzed by mass spectrometry. (C) Derivation of the samples before analysis using different fluorescence dyes allows analysis of two different samples in the same gel.

duction of an internal standard labeled with a third dye (Cy2), thereby allowing quantitative analysis. Although the comparison of two samples with 2D-DIGE has been satisfactory (24), the comparison of several different experiments remains chal-

lenging. Furthermore, the technique is generally limited to proteins between 10 and 200 kD. 2DE is certainly the method of choice for the comparative analysis of unmodified medium-size or large proteins in the discovery phase of biomarker definition

but, in part because of its demand in skills and time, has not been adapted for clinical applications.

Liquid Chromatography Coupled to Mass Spectrometry

Liquid chromatography (LC) provides a powerful fractionation method that is compatible with virtually any mass spectrometer. This method separates large amounts of analytes on an LC column (15,25,26) and offers high sensitivity. A sequential separation using different media in two independent steps provides a multidimensional fractionation that can generate vast amounts of information. For example, the multidimensional protein identification technology (MudPIT) (27,28) and a 2D liquid-phase fractionation approach (29) are well suited for an in-depth analysis of body fluids and tissues (Figure 3). Limitations of LC-MS include difficulties with comparative analysis, in part because of the variability in multidimensional separations and the substantial time required for the analysis of a single sample (generally, in days). Larger protein (above approximately 10 kD) cannot generally be analyzed by this technique; instead, they have to be cleaved by a protease (*e.g.*, trypsin), and the resulting fragments can be subsequently analyzed. Furthermore, the method suffers from its sensitivity to interfering compounds (*e.g.*, lipids, detergents). When data from tryptic digests that were analyzed by LC-MS were compared with 2DE-MS analysis, mostly different proteins were detected by each technique (30). Therefore, both techniques either individually or in combination have the potential to identify valid biomarkers that may not be detected with any other technology that is available. However, neither seems to be suitable for the comparative analysis of hundreds of samples or application in clinical diagnostic processes. Therefore, it cannot be overemphasized that it is imperative to have a strategy for subsequent validation and application of an assay in the clinical laboratory based on another technology (see Mischak *et al.* [3]).

Two alternative MS-based approaches have been used in “clinical proteomics.” These techniques have been found appli-

cable not only to the discovery phase but also to the validation and subsequent application phases.

SELDI-MS

SELDI technology (31–34) reduces the complexity of a sample by selective adsorption of proteins to different active surfaces. Proteins bind to a specific surface (hydrophilic matrix, reverse-phase material, or affinity reagents, *e.g.*, lectins, antibodies) with varying degrees of selectivity while the unbound sample is washed away. A matrix that absorbs energy and allows vaporization and ionization of the sample by laser is added. The sample is subsequently analyzed by MS that provides a low-resolution mass “fingerprint” (Figure 4). Advantages of SELDI include its capacity to analyze multiple samples in a short time. SELDI has been used in numerous attempts to define biomarkers for a variety of diseases (32,35,36). Although the technology is easy to use, it is also prone to generating artifacts (37,38). This may be due, in part, to difficulties with calibration and lack of precision of the determined molecular masses of the analytes. Furthermore, only a very small fraction of all proteins in a sample binds to the chip surface, and the binding varies depending on sample concentration, pH, salt content, presence of interfering compounds such as lipids, *etc.* Therefore, because most of the information contained in a biologic sample is eliminated during the sample preparation, comparability of data sets is limited.

Capillary Electrophoresis Coupled to MS

This approach is based on capillary electrophoresis (CE) at the front end coupled to a mass spectrometer (Figure 5). CE separates proteins in a single step with high resolution based on their migration through a gel in the electrical field (300 to

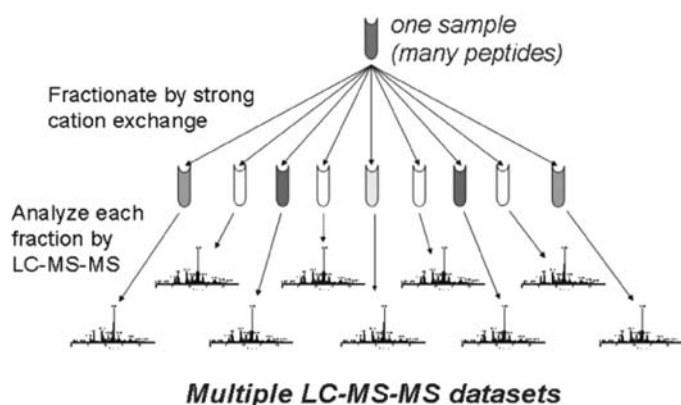


Figure 3. Multidimensional liquid chromatography coupled to mass spectrometry (LC-MS). Proteins are digested and fractionated in first dimension, using cation- or anion-exchange chromatography. Each of these individual fractions is subsequently analyzed in depth (*e.g.*, by reversed-phase LC coupled to MS/MS instruments).

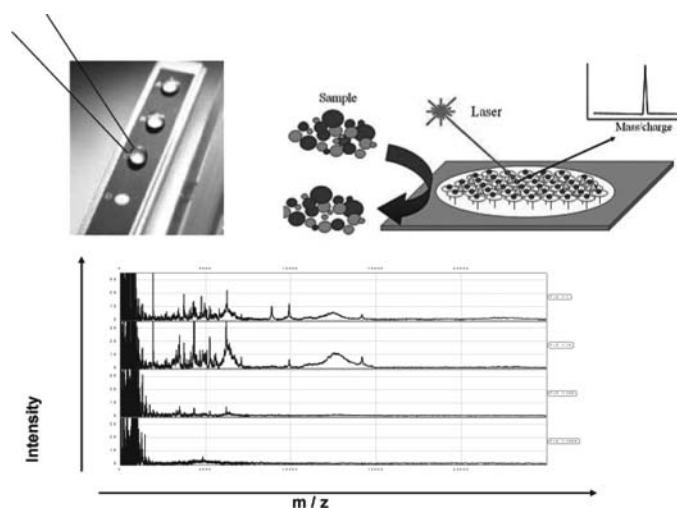


Figure 4. Surface-enhanced laser desorption/ionization mass spectrometry (SELDI-MS). The sample is deposited on the active chip surface (top left). After several washing steps, only a few proteins stay bound to the surface; these are subsequently analyzed using low-resolution MS. (Bottom) A typical SELDI-MS spectrum from urine. Reprinted from Neuheff *et al.* (76), with permission.

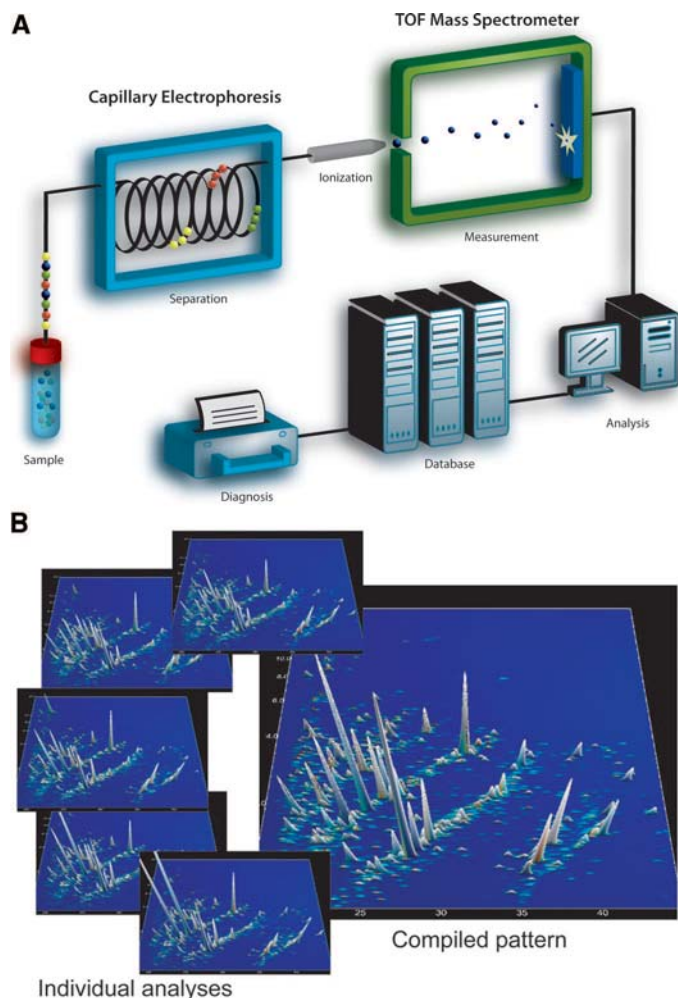


Figure 5. (A) Schematic drawing of the on-line coupling of capillary electrophoresis to the mass spectrometer (CE-MS). Capillary electrophoresis separates polypeptides according to their charge and size. After the electrophoretic separation, the polypeptides are ionized on-line by the application of high voltage and analyzed in the mass spectrometer (ESI-TOF). The combination of the two instruments yields a mass spectrogram of mass per charge plotted against migration time. Subsequently, specialized software solutions allow automated data interpretation. (B) Urine samples of different individuals are analyzed by CE-MS. The small panels on the left are data of five different measurements from samples that were obtained from healthy volunteers. Proteins are displayed as peaks defined by migration time, molecular weight, and signal amplitude (color coded). (Right) Data can be compiled to generate a typical pattern. The migration time (in min) and the mass (in kD, on a logarithmic scale) are indicated.

500 V/cm). CE-MS offers several advantages: (1) It provides fast separation and high resolution (39); (2) it is robust and uses inexpensive capillaries instead of expensive LC columns (9); (3) it is compatible with most buffers and analytes (40); and (4) it provides a stable constant flow, thus avoiding elution gradients that may otherwise interfere with MS detection (41). As shown in Figure 5B, CE-MS enables the generation of comparable high-resolution data sets. The data sets from individual analy-

ses can be compiled to generate a typical proteome pattern that can be based on >100 individual analyses. A disadvantage of CE (although to a lesser degree than LC) is that it cannot be easily used for the analysis of high molecular weight proteins because the large proteins tend to precipitate at the low pH that generally is used in the running buffer. However, such proteins can be digested and the resulting fragments can be analyzed (as described for LC-MS). Furthermore, they can be effectively removed by ultrafiltration (42). In addition, certain proteomes, such as the urinary proteome of healthy individuals, contain mostly low molecular weight proteins; in such cases, the restricted ability to analyze large native proteins does not constitute a severe drawback. Another limitation of CE-MS is the relatively small sample volume that can be loaded onto the capillary, leading to a potentially lower sensitivity of detection. This obstacle has been resolved by improved methods of ionization and by better delivery of the separated protein from the end of a capillary to the MS instrument in a small stream of liquid by nano-ion spray. Also, improvements in the detection limits of mass spectrometers render the issue of sensitivity less important (43–45). Sequencing of potential protein biomarkers that are defined by CE-MS analysis can be achieved by directly interfacing CE with MS/MS instruments (46) or by subsequent targeted sequencing using LC-MS/MS (12,15). Consequently, CE-MS has become a viable alternative to the commonly used proteomic technologies and recently was successfully applied in several clinical studies (8,47,48).

Protein Arrays

As a non-MS-based approach, protein arrays can be used to detect specific proteins ("targeted proteomics"). For more detailed information, we refer the reader to a recent review by Kozarova *et al.* (49). In general, this technique can be considered the modern version of immune detection of multiple proteins. Specific antibodies or antigens are printed on a surface (generally a slide or membrane). A single sample is hybridized to the array that may contain several hundred targets; the captured antigens or antibodies are then detected. As an example, such arrays have been successfully used for detection of antibodies that are directed against specific glomerular antigens (50). The targeted proteome analysis has uncovered novel associations of autoantibodies with disease. Investigators have been able to differentiate between patients with lupus nephritis and healthy and disease controls and to categorize further the severity of the lupus nephritis on the basis of differential IgG and IgM autoreactivity. Similar to other approaches, protein arrays show several limitations (49): The need for a specific probe for every protein to be analyzed (in contrast to nucleic acid arrays that can use the antisense sequence), the generally low density that allows detection of only a few proteins, and posttranslational modifications are usually not detected.

Applications of Proteomics in Nephrology

The main focus of proteome analysis in nephrology is on detection and identification of (urinary) proteins that significantly change (in abundance, distribution, *etc.*) during (patho) physiologic changes of the kidney structure and/or function.

To allow specific and early assessment of disease, at least some of these proteins should be biomarkers that are independent from proteinuria. These biomarkers may be directly related to the disease (*e.g.*, IgA immune deposits in IgA nephropathy) or may result from secondary events (*e.g.*, generation of specific cleavage products by metalloproteases that are upregulated during inflammatory processes in the kidney). After validation, some of these changes may potentially be new therapeutic targets or novel biomarkers for disease detection and/or prognosis.

Although this review is focused on the urine, kidney tissue certainly contains relevant proteomic information. However, analysis of its proteome encompasses several disadvantages: The kidney is composed of different cell types (all of which express different and specialized proteomes) and tissue samples must be obtained invasively, rendering proteome analysis especially of the normal human kidney (which is required as the control) ethically difficult or even impossible. Therefore, most research has focused on tissue that is obtained from experimental animals. The proteome of the rat kidney was described recently by Arthur *et al.* (51). The authors showed differential expression of proteins in the renal cortex and medulla. 2DE resolved 1095 spots from the cortex and 885 spots from the medulla. By MALDI-TOF MS, 54 unique proteins were identified. Nine of them were differentially expressed in the cortex and medulla, and four were expressed in only one region. Xu *et al.* (52) examined glomeruli that were obtained by laser capture dissection and subsequently analyzed the proteome of tissue in the five-sixths nephrectomy rat model of FSGS. They identified thymosin β 4 as a marker of glomerulosclerosis. Recently, the first report of such “protein maps” from murine tissue has also been published (53).

Urine Proteomics

As already outlined, the urine seems to be an ideal source of potential biomarkers. Urinary proteins can be analyzed directly or separated by centrifugation into distinct fractions. For example, supernatants from low-speed centrifugation contain proteins that are derived from filtered plasma proteins and secreted by tubular epithelial cells. This supernatant can be further centrifuged at high speed (ultracentrifugation), yielding pellet-containing exosomes, small vesicles (with diameter <80 nm) with cell membrane, and cytosolic proteins. These exosomes are derived from epithelial cells that line the urinary tract with a contribution from filtered exosomes from blood cells (54,55).

Before the “proteomics” era, many investigators have sought to define better the urinary proteome in a variety of clinical situations. In this respect, one of the first attempts to define proteins in the urine was published by Spahr and co-workers (56,57). Using LC-MS, they analyzed pooled urine samples after tryptic digestion and identified 124 proteins. Although this study did not attempt to define any urinary biomarkers for a disease, it clearly highlighted the plethora of information in the urinary proteome and also a possible approach toward its mining. This conclusion was underscored by Pang *et al.* (58), who used not only 2DE but also 1D- and 2D-LC to define

potential biomarkers for inflammation. Using acetone-precipitated urine samples from healthy volunteers, Thongboonkerd *et al.* (59) defined the first human urinary proteome map, consisting of 67 proteins and their isoforms, that could be used as a reference. In a subsequent study by Oh *et al.* (60), pooled urine samples from 20 healthy volunteers were used to annotate 113 proteins on a 2DE by peptide mass fingerprinting. Additional experiments that further expanded the knowledge of the normal urinary proteome were reported by Pieper *et al.* (61), Sun *et al.* (62), and Castagna *et al.* (63). Taken together, these approaches have identified approximately 800 proteins and laid the foundation for the subsequent discovery of biomarkers in the urinary proteome. In a recent study on urine that was obtained from healthy individuals, Adachi *et al.* (64) identified more than 1500 proteins (or fragments) in the urine of healthy individuals, further underlining the complexity of the human urine proteome. A large proportion of the proteins that were identified in this study were represented by membrane proteins. This may be due to the presence of exosomes (55). Recently, exosomal fetuin-A was proposed as biomarker of acute kidney injury, based on data from a rat model (65), which were further supported by Western blots on three patients. Although these data and the concept of exosomes are very promising, these preliminary observations need to be verified and further explored.

Identification of “Biomarkers” for Kidney Diseases

The definition of disease-specific biomarkers in the urine is complicated by significant changes in the urinary proteome during the day, most likely as a result of exercise, variations in the diet, circadian rhythms, *etc.* (66). As a consequence, the reproducibility of the assay is reduced as a result of these physiologic changes, even if the analytical method shows high reproducibility. In addition, clear differences between first-void and midstream samples can be noted (Mischak *et al.*, unpublished data), further highlighting the importance of standardized protocols for urine sample collection. In one of the first reports on specific urinary proteomic biomarkers in 1979, the identification of several urinary proteins was reported by Anderson *et al.* (67). 2DE analyses of urine from living kidney donors identified several potential proteins related to compensatory kidney growth (68). This work was followed by studies that showed a significant effect of retinoids on renal cells (69,70). Furthermore, 2DE of urine from patients with various biopsy-proven (primary) kidney diseases displayed distinct differences. For example, α 1-antitrypsin was increased in a group of patients with FSGS (71); these results are in line with more recent findings that were obtained with modern proteomic technologies.

On the basis of three patients with diabetic nephropathy, Sharma *et al.* (72) used 2D-DIGE and identified urinary proteins that were differentially present in disease. α 1-Antitrypsin was identified as a potentially upregulated biomarker; this finding was confirmed by Western blotting. More recently, Park *et al.* (73) examined pooled urine from 13 patients with IgA nephropathy and compared the data with those from 12 normal control

subjects. The authors found an array of differentially present proteins and used the data to initiate the establishment of a human urinary proteome map of IgA nephropathy. This study also outlined the limitations of 2DE: The method is tedious and time-consuming; therefore, only a few samples and controls can be analyzed with reasonable effort. However, it is evident that such approaches will enable the definition of several potential biomarkers, and it will be essential in the future to develop a strategy for their evaluation.

A procedure that increases throughput by reducing complexity is SELDI technology. As reviewed already, this technique was recently used by several researchers (*e.g.*, Clarke *et al.* [32] and Schaub *et al.* [74]) to detect potential biomarkers for allograft rejection in kidney transplant patients. Clusters of five and three urinary proteins, respectively, were sufficient for correct classification of 34 and 50 patients with high sensitivity and specificity. It is interesting that these two groups defined completely different biomarkers for the same disorder, and neither found differences between patients with transplanted and native kidneys, certainly an unexpected observation. Urine from patients with acute rejection of a renal allograft has also been examined using CE-MS. Wittke *et al.* (75) found several proteins that revealed substantial differences in concentration in the urine of patients who received a transplant and healthy individuals. Recent unpublished data strongly indicate that these are mostly due to immunosuppression with cyclosporin A. Moreover, several potential biomarkers for acute rejection could be defined. These findings subsequently were verified in a small, blinded study (75): One of 10 controls was misclassified as rejection, whereas six of seven biopsy-proven rejections were correctly identified.

A direct comparison of SELDI with CE-MS technology by Neuheff *et al.* (76) using identical urine samples from control subjects and patients with membranous glomerulonephritis resulted in the definition of three potential biomarkers using SELDI and 200 potential biomarkers from the CE-MS analysis. The authors concluded that it is necessary to characterize any disease with a panel of well-defined biomarker proteins rather than a few peaks that are not too well defined. Mischak and co-workers (77–79) established CE coupled to MS together with appropriate software solutions with the goal to analyze urine (and other body fluids) and develop well-defined protein patterns for diagnosis of various kidney disorders. The urine samples were analyzed individually, and the data from individual CE-MS runs were combined. This feature allowed compilation of data sets and their subsequent comparisons (*e.g.*, patients with a specific kidney disease compared with patients with other types of kidney disease or healthy control subjects). This comparison permitted the definition of an array of biomarkers that differentiate healthy subjects from patients and other markers that define the specific (kidney) disease or clinical condition. The latter type of biomarkers can be useful for differential diagnosis.

One of the first applications of CE-MS for the analysis of a specific urinary proteome was in patients with type 2 diabetes. A total of 168 urinary proteins were present in >90% of the samples, suggesting the existence of a consistent urinary pro-

teome. Panels of 20 to 50 protein markers allowed not only the diagnosis of a specific (primary) kidney disease but also the discrimination with high sensitivity and specificity between different kidney diseases, such as IgA nephropathy, FSGS, membranous glomerulonephritis, minimal-change disease, and diabetic nephropathy (79–81). These findings were recently validated in blinded assessments (82) (J.N. *et al.*, Haubitz *et al.*, and Rossing *et al.*, manuscripts in preparation). As shown in Figure 6, the comparison of compiled data sets that were obtained from control subjects or patients with different renal diseases permits the definition of an array of biomarkers that differentiate healthy subjects from patients and additional markers that define the specific (kidney) disease or clinical condition. The latter type of biomarkers is useful for differential diagnosis.

It is of interest to note that many of the identified biomarkers in the urine of patients with renal diseases are proteolytic fragments of larger proteins. Apparently, specific proteases in the urine may cleave these excreted proteins, as suggested by a recent study of patients with nephrotic syndrome as a result of several glomerular diseases. The authors demonstrated specific urinary proteases that cleaved albumin and α 1-antitrypsin and generated almost 100 proteolytic fragments; these polypeptides appeared as distinct spots on 2D gels (83).

In a recent study, Decramer *et al.* (47) used CE-MS-based urinary proteome analysis to define specific biomarker patterns for different grades of ureteropelvic junction obstruction, a frequently encountered pathology in newborns (of note, these patients do not have significant proteinuria). In a blinded prospective study, the biomarker patterns predicted with 94% accuracy the clinical outcome of these newborns 9 mo in ad-

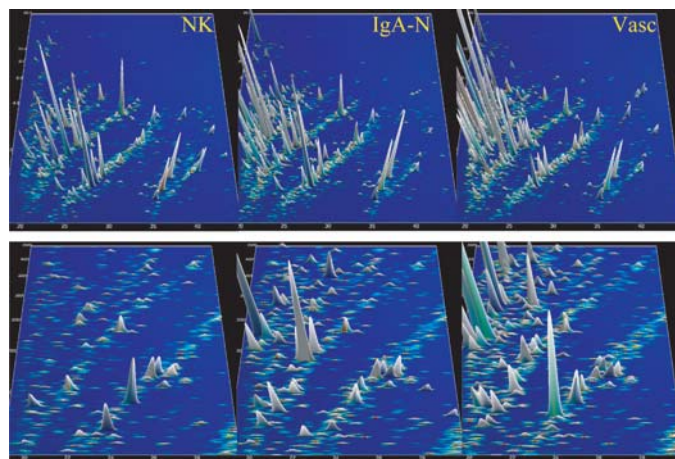


Figure 6. Protein patterns of healthy individuals (NK) and patients with IgA nephropathy (IgA-N) or vasculitis (Vasc). (Top) Patterns that consist of 20 to 100 single measurements, molecular mass (0.8 to 20 kD, on a logarithmic scale) against normalized migration time (18 to 45 min), peak height, and color encode the signal intensity. (Bottom) Zoom of the top patterns (1.5 to 5 kD, 19 to 30 min). As evident, an array of general biomarkers for kidney disease that are present both in IgA-N and vasculitis can be defined. In addition, biomarkers that are specific only for IgA-N or vasculitis can be detected.

vance (Figure 7). These results indicated the potential of urinary proteomics not only to diagnose the renal disease accurately but also to predict its prognosis correctly.

Urinary proteome analysis may also be an excellent tool for fast, noninvasive, and unbiased monitoring of disease progression or response to therapy. In a randomized, double-blinded study, Rossing *et al.* (84) evaluated the treatment of macroalbuminuric patients with daily doses of 8, 16, and 32 mg of candesartan or placebo for 2 mo. Candesartan treatment resulted in a significant change in 15 of 113 proteins that are characteristic for diabetic renal damage. Similar data have been obtained in patients with vasculitis (Haubitz *et al.*, manuscript in preparation) (85), where therapy improved/changed the vasculitis-specific protein pattern toward a normal urinary proteome.

Most current proteomic approaches were used to define new biomarkers for disease. The (patho)physiologic relevance of these biomarkers, although initially unknown, can be clarified by sophisticated bioinformatic approaches. Recently, an alternative approach was used for the molecular phenotyping of human samples (86,87). First, samples are fractionated by chromatographic methods. Specific physiologic effects that are caused by the resultant fractions are then analyzed by an appropriate bioassay after each chromatographic step. The identity of the underlying substance is identified by MS methods such as TOF-TOF MS or FT-ICR MS. After identification of the compound, the (patho)physiologic effect(s) of the substance can

be validated by analysis of the effects of the substance that was obtained by chemical synthesis. Finally, the concentration of the substance in samples of individuals is determined. Using this approach, not only the biomarker but also its (patho)physiologic relevance can be identified, if an appropriate *in vitro* model is available.

Urine Proteomics for Nonrenal Diseases

Proteome analysis of the urine has revealed biomarkers for several nonrenal diseases. As in the case of ureteropelvic junction obstruction, these diseases generally do not result in significant proteinuria. Not surprising, biomarkers for urothelial cancer have been found in urine. Whereas the first studies that were based on SELDI technology analyzed only a few samples and reported different biomarkers for the same disease (88,89), Theodorescu *et al.* (8) recently used CE-MS to assay more than 600 samples, including 180 samples that were examined in a blinded manner, as a validation set. The discovered biomarkers correctly classified all blinded urothelial cancer samples and normal controls; however, nine of 138 patients with various chronic kidney diseases or nephrolithiasis were incorrectly classified as having urothelial cancer.

Kaiser *et al.* (48) defined biomarkers for graft-versus-host disease after bone marrow transplantation using CE-MS-based urine proteomics. This preliminary observation was validated in a recent prospective multicenter study with more than 600 urine samples from more than 100 patients (Weissinger *et al.*, manuscript submitted). In recent studies, we were able to define several biomarkers that are indicative of cardiovascular disease (von zur Muhlen *et al.* and Zimmerli *et al.*, manuscripts submitted). Although these observations at first sight seem intriguing, they may be explained by the microvascular architecture of the kidney. Graft-versus-host disease and cardiovascular disease cause or are the consequence of endothelial dysfunction that may also alter kidney structure and/or function. This complication may, in turn, influence glomerular filtration and/or tubular function and subsequently add disease-specific proteins to the urine.

Identification of Uremic Toxins Using Proteomics

Another application of proteomics that has gained considerable interest is the examination and definition of potential uremic toxins. Spent dialysate and hemofiltrate fluid is an excellent source for proteomic analysis because it contains little albumin and other interfering large proteins (90). In 1994, Forssmann and colleagues (90,91) used an advanced LC-MS approach to identify proteins from hemofiltrate fluid using a “peptide bank” with up to 300 different chromatographic fractions that were prepared from 10,000 L of human hemofiltration fluid. Several additional peptides with various biochemical functions were isolated (*e.g.*, human peptide hormone guanylin, endostatin and resistin as angiogenesis inhibitors, a proopiomelanocortin-derived peptide with lipolytic activity) (92,93).

Ward and Brinkley (94) recently used a proteomic approach that was based on 2DE and MALDI-TOF-MS to identify uremic

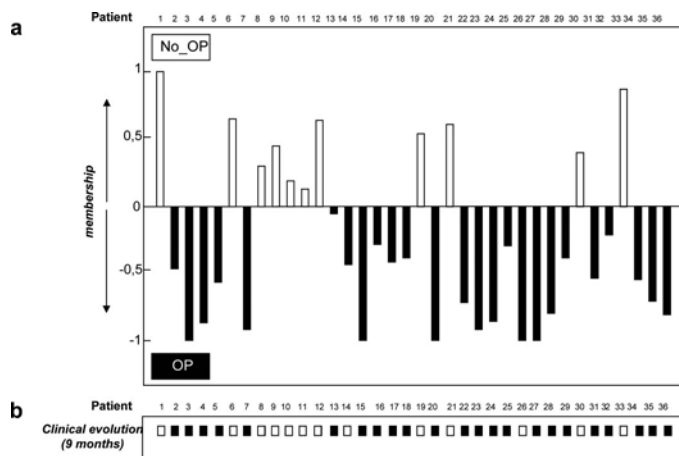


Figure 7. Predictive potential of urinary proteome analysis in a prospective blinded study. The figure shows membership to a specific urinary proteomic pattern in newborns with congenital unilateral ureteropelvic junction (UPJ) obstruction and the clinical outcome of this condition 9 mo after membership prediction. (A) A negative membership value (■) predicted the need for a surgical correction (OP) in the course of the disease, whereas a positive value (□) predicted the evolution toward spontaneous resolution of the UPJ obstruction. (B) Clinical outcome of the OP-positive patients 9 mo after sample analysis. □, the patient had evolved toward the No_OP-negative group (No_OP; spontaneous resolution of the UPJ); ■, the patient needed surgical treatment (OP). The prediction was correct for 34 of the 36 newborns, resulting in a correct prediction in 94% of cases. Reprinted from Decramer *et al.* (47), with permission.

toxins from an ultrafiltrate. Six proteins that harbored several posttranslational modifications (thus presenting as multiple spots for the same protein) were identified. The proteins included α 2-microglobulin (95), as well as α 1-antitrypsin, albumin (mature and complexed), complement factor D, cystatin C, and retinol-binding protein. Molina *et al.* (96) performed a proteome analysis of human hemodialysis fluid applying 1D gel electrophoresis in combination with LC-MS/MS. With this approach, 292 proteins were identified; 205 had not been previously found in serum or plasma. Additional Western blot analysis of a subset of these proteins confirmed their presence in normal serum. This observation indicates that a low sensitivity of detection may explain why most of these proteins had not been previously identified in serum or plasma. The authors concluded that this discrepancy may have resulted from enrichment of the low molecular weight proteins in the hemodialysis fluid.

CE-MS represents a supplement to these proteomic techniques, enabling the analysis of molecules in the low molecular mass range, from 1 up to 10 kD ("middle molecules"). In an initial approach, the effect of different dialysis membranes (low-flux *versus* high-flux) on the number of polypeptides in the dialysate was investigated (77). More than 600 polypeptides have been analyzed in a single sample. Larger polypeptides (>10 kD) were present mostly in dialysates that were obtained with high-flux membranes, whereas most of the polypeptides in dialysates that were obtained with low-flux membranes were smaller than 10 kD. Another study assessed the potential of CE-MS and CE-MS/MS to identify uremic retention molecules in dialysis fluids that were obtained with low-flux and high-flux membranes (97). For obtaining further insight into the uremic toxins within a mass range of 800 to 15,000 Da, the same CE-MS setup was used in combination with a different sample preparation procedure. CE-MS analysis detected 1394 polypeptides in the spent dialysate samples that were obtained with high-flux membranes, whereas 1046 polypeptides were recovered in the dialysate from the same patients that was obtained after hemodialysis with low-flux membranes. In an unrelated study, the same technology was used to identify polypeptides in the plasma of dialysis patients that are generally absent from normal control subjects (98). A combination of data from the study of human plasma and hemodialysate fluid should identify multiple previously unknown uremic toxins.

Bioinformatic Approaches in Proteomics

The results of most, if not all, proteomic studies indicated that a single biomarker does not allow reliable diagnosis, staging, or prognosis of a kidney disease. This finding immediately raised the question how to combine several biomarkers to provide a precise diagnostic or predictive pattern. Although a definitive answer is probably still on the horizon, a number of approaches emerged, which we discuss only in brief.

Hierarchical decision tree–based classification methods, such as classification and regression trees (99), were among the first algorithms to analyze the available information on multiple biomarkers. However, these approaches were not too successful because the number of incorrect predictions increased with

the number of biomarkers (and, consequently, the complexity of the decision tree). Support Vector Machines (SVM; for an example, see Burges [100]) seemed to be an excellent way to overcome this problem. Indeed, reliable results have been obtained when the number of variables (biomarkers) was less than 20 and substantial differences between the data sets (biomarker panel) existed. However, when the differences were more subtle, the precision decreased considerably, particularly when blinded data sets were analyzed (H.M. *et al.*, unpublished data), indicating once more the importance of the blinded samples in any clinical proteomics study.

An important caveat for the use of biomarker patterns for a predicted diagnosis with a classification algorithm is the level of confidence in the prediction. In other words, a classification such as "this urine sample is from an individual with type 2 diabetes" should have a numeric score indicating how likely it is to be a correct classification (*e.g.*, "with 90% confidence, this urine sample is from an individual with type 2 diabetes"). Unfortunately, SVM are generally unable to provide levels of confidence to any classification. Therefore, the clinician is left with no information on the statistical significance of such a prediction. A promising classification method that shares many of the positive characteristics of the SVM but in addition provides the levels of confidence with each classification prediction is based on the Gaussian process (see Rasmussen and Williams [101]). An efficient Gaussian process–based classification method was recently developed (102) and successfully applied to the problem of correct prediction of BRCA1 and BRCA2 heterozygous genotypes (103). No matter which of these mathematical approaches is used, two basic considerations apply: (1) The number of independent variables should be kept to a minimum and should certainly be below the number of samples investigated, and (2) an approach is valid only when it is tested with a blinded validation set; it should be imperative to include such a blinded data set in any report on potential biomarkers.

Limitations of Proteome Analysis

Major limitations of proteomics are related to the type of biologic material to be analyzed and the sensitivity of the methods that are available for the analyses. For example, the role of some proteins in (patho)physiologic processes is not necessarily proportional to the concentration of that protein in the biologic compartment. Therefore, one of the main challenges is to identify scarce compounds and determine their changes between samples. Another limitation of proteome analysis—even more important than the analytical limits—is the stability of the proteome from the time of collecting/processing until completion of the preparation for analysis. This is especially evident when examining blood, as outlined previously.

The evident lack of standards and, subsequently, of comparability of results is another major limitation. The vast majority of the existing reports cannot be compared, thereby greatly reducing their relevance. The situation can be improved by using standard protocols for sample collection, storage, and preparation as well as by using standard analytical perfor-

mance (e.g., mass resolution and accuracy). The establishment of reliable 2DE-, LC-, and CE-MS databases using such standardized protocols would benefit this field. In addition, stricter rules for publication (e.g., mandatory blinded data sets) in peer-reviewed journals may improve the situation. Most of these issues have been outlined recently elsewhere (3), and adherence to the proposed guidelines will hopefully culminate in commonly accepted standards for clinical proteomics.

Lack of appropriate and user-friendly bioinformatics software for data evaluation also hinders development of clinical applications. So far, no standard has been developed for data evaluation, resulting in a set of different solutions that may work well only for a particular problem. However, because the different groups use highly divergent approaches, the data are generally not comparable. A data repository using a specific format, together with certain software solutions that would be universally available, may be an excellent first step toward establishing databases that can be directly compared.

Conclusion

Since the very first clinical observations of kidney diseases, it has been apparent that urinary proteins imply pathologic changes in the kidney. In the past, personal skills (simple observation, smelling, or even tasting of urine) were required and were skillfully performed by our predecessors. Presently, advanced technologies are available to improve the analytical description of the protein content of urine. During what we call the modern era, the contribution of proteomics to the understanding of the pathogenesis, diagnosis, and treatment of kidney disease has already been significant. However, its impact is modest in comparison with the expectations that have been generated by the more than 25 yr of technological progress.

Essentially all studies indicate that the “perfect biomarker” (i.e., a single molecule that clearly defines one disease) does not exist. A panel of distinct biomarkers may be better suited for disease detection (diagnosis) and also for assessment of disease progression and response to therapy. However, this panel of biomarkers must consist of individual biomarkers that are clearly defined and subsequently sequenced (in clear contrast to an ill-defined “pattern”). Furthermore, it is imperative to validate the clinical utility of such biomarkers using a blinded set of samples. Adherence to these simple guidelines should greatly improve the value of future proteome-based studies. The reports based on different technologies, albeit promising, clearly indicate an urgent need for standardization and show that a “common platform” that allows comparison of data sets from different laboratories is needed. Otherwise, these bits of information will never paint a “big picture” that is essential for the full expansion/realization of the potential of proteomics. Given the complexity of the task, it is essential that thousands of comparable data sets be available for data evaluation and validation. Because this cannot be accomplished separately by each laboratory, the establishment of comparability and standards for quality control (e.g., minimal requirements for mass accuracy and resolution of the choice of the mass spectrometers) is essential. First steps in that direction are the definition of guidelines for clinical proteomics (3) and the establishment of

the Human Kidney and Urine Proteome Project” (HKUPP; <http://hkupp.kir.jp>).

Proteome analysis is still far from displaying its full potential as a routine tool in clinical examination, assessment of disease progression, etc. However, first studies with several hundred patients clearly reveal its utility for accurate noninvasive clinical diagnosis (8,47). It may take years or even decades until the entire urinary (or any other) proteome is explored. The question is whether this should be our primary goal or we should take full advantage of a subset of the proteome that contains highly valuable information for clinical use today.

Acknowledgments

J.N. was supported in part by grants DK61525 and DK71802 from the National Institutes of Health and by the General Clinical Research Center of the University of Alabama at Birmingham (M01 RR00032).

A.A., V.J., J.J., and H.M. are members of the European Uraemic Toxin Group of the ESAO (EuTox).

We are grateful to Bruce A. Julian, Eric Schiffer, Jochen Ehrlich, Tadashi Yamamoto, and Joost Schanstra for critically reviewing the manuscript.

Disclosures

None.

References

- Seitz W, Zimmer E, Alberti PE: Paper-electrophoretic studies on proteins in urine of normal and sick individuals [in German]. *Z Klin Med* 152: 196–201, 1953
- O'Farrell PH: High resolution two-dimensional electrophoresis of proteins. *J Biol Chem* 250: 4007–4021, 1975
- Mischak H, Apweiler R, Banks RE, Conaway M, Coon JJ, Dominizak A, Ehrlich JH, Fliser D, Girolami M, Hermjakob H, Hochstrasser DF, Jankowski V, Julian BA, Kolch W, Massy Z, Neususs C, Novak J, Peter K, Rossing K, Schanstra JP, Semmes OJ, Theodorescu D, Thongboonkerd V, Weissinger EM, Van Eyk JE, Yamamoto T: Clinical proteomics: A need to define the field and to begin to set adequate standards. *Proteomics Clin Appl* 1: 148–156, 2007
- Thongboonkerd V: Proteomics in nephrology: Current status and future directions. *Am J Nephrol* 24: 360–378, 2004
- Thongboonkerd V, Malasit P: Renal and urinary proteomics: Current applications and challenges. *Proteomics* 5: 1033–1042, 2005
- Hewitt SM, Dear J, Star RA: Discovery of protein biomarkers for renal diseases. *J Am Soc Nephrol* 15: 1677–1689, 2004
- Schaub S, Wilkins J, Weiler T, Sangster K, Rush D, Nickerson P: Urine protein profiling with surface-enhanced laser-desorption/ionization time-of-flight mass spectrometry. *Kidney Int* 65: 323–332, 2004
- Theodorescu D, Wittke S, Ross MM, Walden M, Conaway M, Just I, Mischak H, Frierson HF: Discovery and validation of new protein biomarkers for urothelial cancer: A prospective analysis. *Lancet Oncol* 7: 230–240, 2006
- Kolch W, Neususs C, Pelzing M, Mischak H: Capillary electrophoresis-mass spectrometry as a powerful tool in clinical diagnosis and biomarker discovery. *Mass Spectrom Rev* 24: 959–977, 2005
- Omenn GS, States DJ, Adamski M, Blackwell TW, Menon

- R, Hermjakob H, Apweiler R, Haab BB, Simpson RJ, Edde JS, Kapp EA, Moritz RL, Chan DW, Rai AJ, Admon A, Aebersold R, Eng J, Hancock WS, Hefta SA, Meyer H, Paik YK, Yoo JS, Ping P, Pounds J, Adkins J, Qian X, Wang R, Wasinger V, Wu CY, Zhao X, Zeng R, Archakov A, Tsugita A, Beer I, Pandey A, Pisano M, Andrews P, Tammen H, Speicher DW, Hanash SM: Overview of the HUPO Plasma Proteome Project: Results from the pilot phase with 35 collaborating laboratories and multiple analytical groups, generating a core dataset of 3020 proteins and a publicly-available database. *Proteomics* 5: 3226–3245, 2005
11. Lapolla A, Tubaro M, Fedele D, Reitano R, Arico NC, Ragazzi E, Seraglia R, Vogliardi S, Traldi P: A matrix-assisted laser desorption/ionization mass spectrometry study of the non-enzymatic glycation products of human globins in diabetes. *Rapid Commun Mass Spectrom* 19: 162–168, 2005
 12. Chalmers MJ, Mackay CL, Hendrickson CL, Wittke S, Walden M, Mischak H, Fliser D, Just I, Marshall AG: Combined top-down and bottom-up mass spectrometric approach to characterization of biomarkers for renal disease. *Anal Chem* 77: 7163–7171, 2005
 13. Coon JJ, Ueberheide B, Syka JE, Dryhurst DD, Ausio J, Shabanowitz J, Hunt DF: Protein identification using sequential ion/ion reactions and tandem mass spectrometry. *Proc Natl Acad Sci U S A* 102: 9463–9468, 2005
 14. Good DM, Coon JJ: Advancing proteomics with ion/ion chemistry. *Biotechniques* 40: 783–789, 2006
 15. Zurbig P, Renfrow MB, Schiffer E, Novak J, Walden M, Wittke S, Just I, Pelzing M, Neuss C, Theodorescu D, Root C, Ross M, Mischak H: Biomarker discovery by CE-MS enables sequence analysis via MS/MS with platform-independent separation. *Electrophoresis* 27: 2111–2125, 2006
 16. Annesley TM: Ion suppression in mass spectrometry. *Clin Chem* 49: 1041–1044, 2003
 17. Shaw J, Rowlinson R, Nickson J, Stone T, Sweet A, Williams K, Tonge R: Evaluation of saturation labelling two-dimensional difference gel electrophoresis fluorescent dyes. *Proteomics* 3: 1181–1195, 2003
 18. Wang W, Zhou H, Lin H, Roy S, Shaler TA, Hill LR, Norton S, Kumar P, Anderle M, Becker CH: Quantification of proteins and metabolites by mass spectrometry without isotopic labeling or spiked standards. *Anal Chem* 75: 4818–4826, 2003
 19. Ong SE, Mann M: Mass spectrometry-based proteomics turns quantitative. *Nat Chem Biol* 1: 252–262, 2005
 20. Laemmli UK: Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227: 680–685, 1970
 21. Burnette WN: “Western blotting”: Electrophoretic transfer of proteins from sodium dodecyl sulfate–polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. *Anal Biochem* 112: 195–203, 1981
 22. Henzel WJ, Billeci TM, Stults JT, Wong SC, Grimley C, Watanabe C: Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases. *Proc Natl Acad Sci U S A* 90: 5011–5015, 1993
 23. Henzel WJ, Watanabe C, Stults JT: Protein identification: The origins of peptide mass fingerprinting. *J Am Soc Mass Spectrom* 14: 931–942, 2003
 24. Wu TL: Two-dimensional difference gel electrophoresis. *Methods Mol Biol* 328: 71–95, 2006
 25. Aebersold R, Mann M: Mass spectrometry-based proteomics. *Nature* 422: 198–207, 2003
 26. Issaq HJ, Conrads TP, Janini GM, Veenstra TD: Methods for fractionation, separation and profiling of proteins and peptides. *Electrophoresis* 23: 3048–3061, 2002
 27. Ru QC, Katenhusen RA, Zhu LA, Silberman J, Yang S, Orchard TJ, Brzeski H, Liebman M, Ellsworth DL: Proteomic profiling of human urine using multi-dimensional protein identification technology. *J Chromatogr A* 1111: 166–174, 2006
 28. Kislinger T, Gramolini AO, MacLennan DH, Emili A: Multidimensional protein identification technology (MudPIT): Technical overview of a profiling method optimized for the comprehensive proteomic investigation of normal and diseased heart tissue. *J Am Soc Mass Spectrom* 16: 1207–1220, 2005
 29. Soldi M, Sarto C, Valsecchi C, Magni F, Proserpio V, Ticozzi D, Mocarelli P: Proteome profile of human urine with two-dimensional liquid phase fractionation. *Proteomics* 5: 2641–2647, 2005
 30. Komatsu S, Zang X, Tanaka N: Comparison of two proteomics techniques used to identify proteins regulated by gibberellin in rice. *J Proteome Res* 5: 270–276, 2006
 31. Hampel DJ, Sansome C, Sha M, Brodsky S, Lawson WE, Goligorsky MS: Toward proteomics in uroscopy: Urinary protein profiles after radiocontrast medium administration. *J Am Soc Nephrol* 12: 1026–1035, 2001
 32. Clarke W, Silverman BC, Zhang Z, Chan DW, Klein AS, Molmenti EP: Characterization of renal allograft rejection by urinary proteomic analysis. *Ann Surg* 237: 660–664, 2003
 33. Dare TO, Davies HA, Turton JA, Lomas L, Williams TC, York MJ: Application of surface-enhanced laser desorption/ionization technology to the detection and identification of urinary parvalbumin- α : A biomarker of compound-induced skeletal muscle toxicity in the rat. *Electrophoresis* 23: 3241–3251, 2002
 34. Issaq HJ, Conrads TP, Prieto DA, Tirumalai R, Veenstra TD: SELDI-TOF MS for diagnostic proteomics. *Anal Chem* 75: 148A–155A, 2003
 35. Woroniecki RP, Orlova TN, Mendelev N, Shatat IF, Hailpern SM, Kaskel FJ, Goligorsky MS, O’Riordan E: Urinary proteome of steroid-sensitive and steroid-resistant idiopathic nephrotic syndrome of childhood. *Am J Nephrol* 26: 258–267, 2006
 36. Schaub S, Rush D, Wilkins J, Gibson IW, Weiler T, Sangster K, Nicolle L, Karpinski M, Jeffery J, Nickerson P: Proteomic-based detection of urine proteins associated with acute renal allograft rejection. *J Am Soc Nephrol* 15: 219–227, 2004
 37. Check E: Running before we can walk. *Nature* 429: 496–497, 2004
 38. Baggerly KA, Morris JS, Coombes KR: Reproducibility of SELDI-TOF protein patterns in serum: Comparing datasets from different experiments. *Bioinformatics* 20: 777–785, 2002
 39. Johannesson N, Wetterhall M, Markides KE, Bergquist J: Monomer surface modifications for rapid peptide analysis by capillary electrophoresis and capillary electrochromatography coupled to electrospray ionization-mass spectrometry. *Electrophoresis* 25: 809–816, 2004

40. Hernandez-Borges J, Neususs C, Cifuentes A, Pelzing M: On-line capillary electrophoresis-mass spectrometry for the analysis of biomolecules. *Electrophoresis* 25: 2257–2281, 2004
41. Neususs C, Pelzing M, Macht M: A robust approach for the analysis of peptides in the low femtomole range by capillary electrophoresis-tandem mass spectrometry. *Electrophoresis* 23: 3149–3159, 2002
42. Theodorescu D, Fliser D, Wittke S, Mischak H, Krebs R, Walden M, Ross M, Eltze E, Bettendorf O, Wulfiging C, Semjonow A: Pilot study of capillary electrophoresis coupled to mass spectrometry as a tool to define potential prostate cancer biomarkers in urine. *Electrophoresis* 26: 2797–2808, 2005
43. Sassi AP, Andel F 3rd, Bitter HM, Brown MP, Chapman RG, Espiritu J, Greenquist AC, Guyon I, Horch-Alegre M, Stults KL, Wainright A, Heller JC, Stults JT: An automated, sheathless capillary electrophoresis-mass spectrometry platform for discovery of biomarkers in human serum. *Electrophoresis* 26: 1500–1512, 2005
44. Ullsten S, Danielsson R, Backstrom D, Sjoberg P, Bergquist J: Urine profiling using capillary electrophoresis-mass spectrometry and multivariate data analysis. *J Chromatogr A* 1117: 87–93, 2006
45. Klampfl CW: Recent advances in the application of capillary electrophoresis with mass spectrometric detection. *Electrophoresis* 27: 3–34, 2006
46. Wittke S, Mischak H, Walden M, Kolch W, Radler T, Wiedemann K: Discovery of biomarkers in human urine and cerebrospinal fluid by capillary electrophoresis coupled to mass spectrometry: Towards new diagnostic and therapeutic approaches. *Electrophoresis* 26: 1476–1487, 2005
47. Decramer S, Wittke S, Mischak H, Zurbig P, Walden M, Bouissou F, Bascands JL, Schanstra JP: Predicting the clinical outcome of congenital unilateral ureteropelvic junction obstruction in newborn by urinary proteome analysis. *Nat Med* 12: 398–400, 2006
48. Kaiser T, Kamal H, Rank A, Kolb HJ, Holler E, Ganser A, Hertenstein B, Mischak H, Weissinger EM: Proteomics applied to the clinical follow-up of patients after allogeneic hematopoietic stem cell transplantation. *Blood* 104: 340–349, 2004
49. Kozarova A, Petrinac S, Ali A, Hudson JW: Array of informatics: Applications in modern research. *J Proteome Res* 5: 1051–1059, 2006
50. Li QZ, Xie C, Wu T, Mackay M, Aranow C, Putterman C, Mohan C: Identification of autoantibody clusters that best predict lupus disease activity using glomerular proteome arrays. *J Clin Invest* 115: 3428–3439, 2005
51. Arthur JM, Thongboonkerd V, Scherzer JA, Cai J, Pierce WM, Klein JB: Differential expression of proteins in renal cortex and medulla: A proteomic approach. *Kidney Int* 62: 1314–1321, 2002
52. Xu BJ, Shyr Y, Liang X, Ma LJ, Donnert EM, Roberts JD, Zhang X, Kon V, Brown NJ, Caprioli RM, Fogo AB: Proteomic patterns and prediction of glomerulosclerosis and its mechanisms. *J Am Soc Nephrol* 16: 2967–2975, 2005
53. Kislinger T, Cox B, Kannan A, Chung C, Hu P, Ignatchenko A, Scott MS, Gramolini AO, Morris Q, Hallett MT, Rossant J, Hughes TR, Frey B, Emili A: Global survey of organ and organelle protein expression in mouse: Combined proteomic and transcriptomic profiling. *Cell* 125: 173–186, 2006
54. Zhou H, Yuen PS, Pisitkun T, Gonzales PA, Yasuda H, Dear JW, Gross P, Knepper MA, Star RA: Collection, storage, preservation, and normalization of human urinary exosomes for biomarker discovery. *Kidney Int* 69: 1471–1476, 2006
55. Pisitkun T, Shen RF, Knepper MA: Identification and proteomic profiling of exosomes in human urine. *Proc Natl Acad Sci U S A* 101: 13368–13373, 2004
56. Davis MT, Spahr CS, McGinley MD, Robinson JH, Bures EJ, Beierle J, Mort J, Yu W, Luethy R, Patterson SD: Towards defining the urinary proteome using liquid chromatography-tandem mass spectrometry. II. Limitations of complex mixture analyses. *Proteomics* 1: 108–117, 2001
57. Spahr CS, Davis MT, McGinley MD, Robinson JH, Bures EJ, Beierle J, Mort J, Courchesne PL, Chen K, Wahl RC, Yu W, Luethy R, Patterson SD: Towards defining the urinary proteome using liquid chromatography-tandem mass spectrometry. I. Profiling an unfractionated tryptic digest. *Proteomics* 1: 93–107, 2001
58. Pang JX, Ginanni N, Dongre AR, Hefta SA, Opitke GJ: Biomarker discovery in urine by proteomics. *J Proteome Res* 1: 161–169, 2002
59. Thongboonkerd V, McLeish KR, Arthur JM, Klein JB: Proteomic analysis of normal human urinary proteins isolated by acetone precipitation or ultracentrifugation. *Kidney Int* 62: 1461–1469, 2002
60. Oh J, Pyo JH, Jo EH, Hwang SI, Kang SC, Jung JH, Park EK, Kim SY, Choi JY, Lim J: Establishment of a near-standard two-dimensional human urine proteomic map. *Proteomics* 4: 3485–3497, 2004
61. Pieper R, Gatlin CL, McGrath AM, Makusky AJ, Mondal M, Seonarin M, Field E, Schatz CR, Estock MA, Ahmed N, Anderson NG, Steiner S: Characterization of the human urinary proteome: A method for high-resolution display of urinary proteins on two-dimensional electrophoresis gels with a yield of nearly 1400 distinct protein spots. *Proteomics* 4: 1159–1174, 2004
62. Sun W, Li F, Wu S, Wang X, Zheng D, Wang J, Gao Y: Human urine proteome analysis by three separation approaches. *Proteomics* 5: 4994–5001, 2005
63. Castagna A, Cecconi D, Sennels L, Rappsilber J, Guerrier L, Fortis F, Boschetti E, Lomas L, Righetti PG: Exploring the hidden human urinary proteome via ligand library beads. *J Proteome Res* 4: 1917–1930, 2005
64. Adachi J, Kumar C, Zhang Y, Olsen JV, Mann M: The human urinary proteome contains more than 1500 proteins including a large proportion of membrane proteins. *Genome Biol* 7: R80, 2006
65. Zhou H, Pisitkun T, Aponte A, Yuen PS, Hoffert JD, Yasuda H, Hu X, Chawla L, Shen RF, Knepper MA, Star RA: Exosomal fetuin-A identified by proteomics: A novel urinary biomarker for detecting acute kidney injury. *Kidney Int* 70: 1847–1857, 2006
66. Fliser D, Wittke S, Mischak H: Capillary electrophoresis coupled to mass spectrometry for clinical diagnostic purposes. *Electrophoresis* 26: 2708–2716, 2005
67. Anderson NG, Anderson NL, Tollaksen SL: Proteins of human urine. I. Concentration and analysis by two-dimensional electrophoresis. *Clin Chem* 25: 1199–1210, 1979
68. Argiles A, Mourad G, Basset N, Xelrud-Cavadore C,

- Haiech J, Mion C, Cavadore JC, Demaille JG: Acute adaptive changes to unilateral nephrectomy in humans. *Kidney Int* 32: 714–720, 1987
69. Argiles A, Kraft NE, Hutchinson P, Senes-Ferrari S, Atkins RC: Retinoic acid affects the cell cycle and increases total protein content in epithelial cells. *Kidney Int* 36: 954–959, 1989
70. Argiles A, Kraft N, Ootaka T, Hutchinson P, Atkins RC: Epidermal growth factor and transforming growth factor alpha stimulate or inhibit proliferation of a human renal adenocarcinoma cell line depending on cell status: Differentiation of the two pathways by G protein involvement. *Cancer Res* 52: 4356–4360, 1992
71. Argiles A, Mourad G, Mion C, Atkins RC, Haiech J: Two-dimensional gel electrophoresis of urinary proteins in kidney diseases. *Contrib Nephrol* 83: 1–8, 1990
72. Sharma K, Lee S, Han S, Lee S, Francos B, McCue P, Wassell R, Shaw MA, RamachandraRao SP: Two-dimensional fluorescence difference gel electrophoresis analysis of the urine proteome in human diabetic nephropathy. *Proteomics* 5: 2648–2655, 2005
73. Park MR, Wang EH, Jin DC, Cha JH, Lee KH, Yang CW, Kang CS, Choi YJ: Establishment of a 2-D human urinary proteomic map in IgA nephropathy. *Proteomics* 6: 1066–1076, 2006
74. Schaub S, Wilkins JA, Antonovici M, Krokshin O, Weiler T, Rush D, Nickerson P: Proteomic-based identification of cleaved urinary beta2-microglobulin as a potential marker for acute tubular injury in renal allografts. *Am J Transplant* 5: 729–738, 2005
75. Wittke S, Haubitz M, Walden M, Rohde F, Schwarz A, Mengel M, Mischak H, Haller H, Gwinner W: Detection of acute tubulointerstitial rejection by proteomic analysis of urinary samples in renal transplant recipients. *Am J Transplant* 5: 2479–2488, 2005
76. Neuhoof N, Kaiser T, Wittke S, Krebs R, Pitt A, Burchard A, Sundmacher A, Schlegelberger B, Kolch W, Mischak H: Mass spectrometry for the detection of differentially expressed proteins: A comparison of surface-enhanced laser desorption/ionization and capillary electrophoresis/mass spectrometry. *Rapid Commun Mass Spectrom* 18: 149–156, 2004
77. Kaiser T, Hermann A, Kielstein JT, Wittke S, Bartel S, Krebs R, Hausadel F, Hillmann M, Golovko I, Koester P, Haller H, Weissinger EM, Fliser D, Mischak H: Capillary electrophoresis coupled to mass spectrometry to establish polypeptide patterns in dialysis fluids. *J Chromatogr A* 1013: 157–171, 2003
78. Wittke S, Fliser D, Haubitz M, Bartel S, Krebs R, Hausadel F, Hillmann M, Golovko I, Koester P, Haller H, Kaiser T, Mischak H, Weissinger EM: Determination of peptides and proteins in human urine with capillary electrophoresis-mass spectrometry, a suitable tool for the establishment of new diagnostic markers. *J Chromatogr A* 1013: 173–181, 2003
79. Mischak H, Kaiser T, Walden M, Hillmann M, Wittke S, Herrmann A, Kneuppel S, Haller H, Fliser D: Proteomic analysis for the assessment of diabetic renal damage in humans. *Clin Sci (Lond)* 107: 485–495, 2004
80. Haubitz M, Wittke S, Weissinger EM, Walden M, Rupprecht HD, Floege J, Haller H, Mischak H: Urine protein patterns can serve as diagnostic tools in patients with IgA nephropathy. *Kidney Int* 67: 2313–2320, 2005
81. Weissinger EM, Wittke S, Kaiser T, Haller H, Bartel S, Krebs R, Golovko I, Rupprecht HD, Haubitz M, Hecker H, Mischak H, Fliser D: Proteomic patterns established with capillary electrophoresis and mass spectrometry for diagnostic purposes. *Kidney Int* 65: 2426–2434, 2004
82. Haubitz M, Fliser D, Rupprecht H, Floege J, Haller H, Rossing K, Walden M, Wittke S, Mischak H: Defining renal diseases based on proteome analysis. *Nephrol Dial Transplant* 20: V20, 2005
83. Candiano G, Musante L, Bruschi M, Petretto A, Santucci L, Del BP, Pavone B, Perfumo F, Urbani A, Scolari F, Ghiggeri G: Repetitive fragmentation products of albumin and alpha1-antitrypsin in glomerular diseases associated with nephrotic syndrome. *J Am Soc Nephrol* 17: 3139–3148, 2006
84. Rossing K, Mischak H, Parving HH, Christensen PK, Walden M, Hillmann M, Kaiser T: Impact of diabetic nephropathy and angiotensin II receptor blockade on urinary polypeptide patterns. *Kidney Int* 68: 193–205, 2005
85. Schiffer E, Mischak H, Novak J: High resolution proteome/peptidome analysis of body fluids by capillary electrophoresis coupled with MS. *Proteomics* 6: 5615–5627, 2006
86. Jankowski V, Vanholder R, van der GM, Henning L, Tolle M, Schonfelder G, Krakow A, Karadogan S, Gustavsson N, Gobom J, Webb J, Lehrach H, Giebing G, Schluter H, Hilgers KF, Zidek W, Jankowski J: Detection of angiotensin II in supernatants of stimulated mononuclear leukocytes by matrix-assisted laser desorption ionization time-of-flight/time-of-flight mass analysis. *Hypertension* 46: 591–597, 2005
87. Jankowski V, Tolle M, Vanholder R, Schonfelder G, van der Giet M, Henning L, Schluter H, Paul M, Zidek W, Jankowski J: Uridine adenosine tetraphosphate: A novel endothelium-derived vasoconstrictive factor. *Nat Med* 11: 223–227, 2005
88. Liu W, Guan M, Wu D, Zhang Y, Wu Z, Xu M, Lu Y: Using tree analysis pattern and SELDI-TOF-MS to discriminate transitional cell carcinoma of the bladder cancer from non-cancer patients. *Eur Urol* 47: 456–462, 2005
89. Vlahou A, Schellhammer PF, Mendrinos S, Patel K, Kondylis FI, Gong L, Nasim S, Wright JG Jr: Development of a novel proteomic approach for the detection of transitional cell carcinoma of the bladder in urine. *Am J Pathol* 158: 1491–1502, 2001
90. Schepky AG, Bensch KW, Schulz-Knappe P, Forssmann WG: Human hemofiltrate as a source of circulating bioactive peptides: Determination of amino acids, peptides and proteins. *Biomed Chromatogr* 8: 90–94, 1994
91. Schulz-Knappe P, Schrader M, Standker L, Richter R, Hess R, Jurgens M, Forssmann WG: Peptide bank generated by large-scale preparation of circulating human peptides. *J Chromatogr A* 776: 125–132, 1997
92. John H, Radtke K, Standker L, Forssmann WG: Identification and characterization of novel endogenous proteolytic forms of the human angiogenesis inhibitors restin and endostatin. *Biochim Biophys Acta* 1747: 161–170, 2005
93. Fricke K, Schulz A, John H, Forssmann WG, Maronde E: Isolation and characterization of a novel proopiomelano-

- cortin-derived peptide from hemofiltrate of chronic renal failure patients. *Endocrinology* 146: 2060–2068, 2005
94. Ward RA, Brinkley KA: A proteomic analysis of proteins removed by ultrafiltration during extracorporeal renal replacement therapy. *Contrib Nephrol* 141: 280–291, 2004
95. Gejyo F, Yamada T, Odani S, Nakagawa Y, Arakawa M, Kunitomo T, Kataoka H, Suzuki M, Hirasawa Y, Shirahama T: A new form of amyloid protein associated with chronic hemodialysis was identified as beta 2-microglobulin. *Biochem Biophys Res Commun* 129: 701–706, 1985
96. Molina H, Bunkenborg J, Reddy GH, Muthusamy B, Scheel PJ, Pandey A: A proteomic analysis of human hemodialysis fluid. *Mol Cell Proteomics* 4: 637–650, 2005
97. Weissinger EM, Kaiser T, Meert N, De Smet R, Walden M, Mischak H, Vanholder RC: Proteomics: A novel tool to unravel the patho-physiology of uraemia. *Nephrol Dial Transplant* 19: 3068–3077, 2004
98. Weissinger EM, Nguyen-Khoa T, Fumeron C, Saltiel C, Walden M, Kaiser T, Mischak H, Drueke TB, Lacour B, Massy ZA: Effects of oral vitamin C supplementation in hemodialysis patients: A proteomic assessment. *Proteomics* 6: 993–1000, 2006
99. Breimann L, Friemann J, Olshen RA, Stone JC: *Classification and Regression Trees*, Belmont, CA, Wadsworth, 1984
100. Burges CJC: A tutorial on support vector machines for pattern recognition. *Knowledge Discovery and Data Mining* 2: 121–167, 1998
101. Rasmussen CE, Williams CKI: *Gaussian Processes for Machine Learning*, Cambridge, MIT Press, 2006
102. Girolami M, Rogers S: Variational Bayesian multinomial probit regression with Gaussian process priors. *Neural Computation* 18: 1790–1817, 2006
103. Kote-Jarai Z, Matthews L, Osorio A, Shanley S, Giddings I, Moreews F, Locke I, Evans DG, Eccles D, Williams RD, Girolami M, Campbell C, Eeles R: Accurate prediction of BRCA1 and BRCA2 heterozygous genotype using expression profiling after induced DNA damage. *Clin Cancer Res* 12: 3896–3901, 2006

Altered Proteome in *Burkholderia pseudomallei* *rpoE* Operon Knockout Mutant: Insights into Mechanisms of *rpoE* Operon in Stress Tolerance, Survival, and Virulence

Visith Thongboonkerd,^{*,†} Muthita Vanaporn,[‡] Napat Songtawee,[†] Rattiyaporn Kanlaya,^{†,‡} Supachok Sinchaikul,[§] Shui-Tein Chen,^{§,||} Anna Easton,[⊥] Karen Chu,[⊥] Gregory J. Bancroft,[⊥] and Sunee Korbsrisate[‡]

Medical Molecular Biology Unit, Office for Research and Development, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand, Department of Immunology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand, Institute of Biological Chemistry and Genomic Research Center, Academia Sinica, Taipei, Taiwan, Institute of Biochemical Sciences, College of Life Science, National Taiwan University, Taipei, Taiwan, and Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, Keppel Street, London, United Kingdom

Received September 6, 2006

We have previously shown that the alternative sigma factor σ^E (RpoE), encoded by *rpoE*, is involved in stress tolerance and survival of *Burkholderia pseudomallei*. However, its molecular and pathogenic mechanisms remain unclear. In the present study, we applied gel-based, differential proteomics to compare the cellular proteome of an *rpoE* operon knockout mutant (RpoE Mut) to that of wild-type (K96243 WT) *B. pseudomallei*. Quantitative intensity analysis ($n = 5$ gels from 5 individual culture flasks in each group) revealed significantly differential expression of 52 proteins, which were subsequently identified by Q-TOF MS/MS. These included oxidative, osmotic, and other stress response proteins; chaperones; transcriptional/translational regulators; metabolic enzymes; proteins involved in cell wall synthesis, fatty synthesis, glycogen synthesis, and storage; exported proteins; secreted proteins; adhesion molecule; protease/peptidase; protease inhibitor; signaling proteins; and other miscellaneous proteins. The down-regulation of several stress response proteins, chaperones, transcriptional/translational regulators, and proteins involved in cell wall synthesis in RpoE Mut provided some new insights into the mechanisms of the *rpoE* operon for the stress tolerance and survival of *B. pseudomallei*. In addition, the proteomic data and *in vivo* study indicated that the *rpoE* operon is also involved in the virulence of *B. pseudomallei*. Our findings underscore the usefulness of proteomics for unraveling pathogenic mechanisms of diseases at the molecular level.

Keywords: Proteome • Proteomics • Melioidosis • Burkholderia • RpoE • Stress tolerance • Survival • Virulence

Introduction

Burkholderia pseudomallei is a Gram-negative bacillus found in soil and water and is the causative agent of melioidosis, a disease of which clinical manifestations can be acute or chronic. Organ involvement in melioidosis ranges from local to systemic, and its severity varies from mild to fatal. Almost all organs, with the exception of hair and nails, can be affected

by melioidosis.¹ Mortality rates in patients with septic shock caused by melioidosis are approximately 80–95% despite adequate treatment.¹ Previous studies on melioidosis have focused not only on the improvement of therapeutic outcome, but also on the understanding of the pathogenic mechanisms of this infectious disease. Moreover, *B. pseudomallei* has been considered as a potential bioterrorism weapon.² Better understanding of the molecular basis and pathogenic mechanisms of this organism is, therefore, critically required for the discovery of new therapeutic targets and vaccine development for disease prevention.

As *B. pseudomallei* is a saprophyte found in soil and water, it is a difficult microorganism to kill, and it can survive in these environments for years. Additionally, this microorganism is resistant to several antibiotics, chemicals, organic compounds, and other stressful conditions.³ Moreover, it can survive within different eukaryotic cell types, including mammalian phagocytes. Thus, stress tolerance has been thought to be one of the

* To whom correspondence should be addressed. Visith Thongboonkerd, MD, FRCPT, Medical Molecular Biology Unit, Office for Research and Development, 12th Floor Adulyadej Vikrom Building, 2 Prannok Road, Siriraj Hospital, Bangkoknoi, Bangkok 10700, Thailand. Phone/fax, +66-2-4184793; e-mail, thongboonkerd@dr.com or vthongbo@yahoo.com.

[†] Office for Research and Development, Faculty of Medicine Siriraj Hospital, Mahidol University.

[‡] Department of Immunology, Faculty of Medicine Siriraj Hospital, Mahidol University.

[§] Academia Sinica.

^{||} National Taiwan University.

[⊥] London School of Hygiene and Tropical Medicine.

important factors for the survival of *B. pseudomallei* both inside and outside of the human body. While much progress has been made regarding its virulence factors, that is, secretory proteins and cell-associated antigens,³ little is known about the stress tolerance of this bacterial pathogen.

In *Escherichia coli*, the alternative sigma factor σ^E (RpoE), encoded by *rpoE*, plays an important role in maintaining the integrity of the cell envelope (by controlling the transcription of several genes associated with cell envelope integrity) and is, thus, essential for viability of the bacterium.⁴ During stresses (heat stress, chemical exposure, etc.), RpoE is activated and transcribes genes in its regulon, including those encoding chaperones and proteases, which subsequently refold and degrade misfolded proteins, respectively. Recently, RpoE has been shown to play a critical role in survival, stress response, and virulence of several other bacteria, that is, *Azotobacter vinelandii*,^{5,6} *Bacillus subtilis*,^{7,8} *Haemophilus influenzae*,⁹ *Mycobacterium tuberculosis*,¹⁰ *Pseudomonas aeruginosa*,^{11,12} *Pseudomonas fluorescens*,¹³ *Salmonella enterica*,¹⁴ *Streptomyces antibioticus*,¹⁵ and *Vibrio cholerae*.¹⁶ More recently, we have demonstrated that the *rpoE* operon also plays a pivotal role in stress tolerance and biofilm formation in *B. pseudomallei*.¹⁷ These data underscore the significance of RpoE in various bacteria. However, the available information is limited and does not provide sufficient insights into the mechanisms of how the *rpoE* operon controls such functions in these bacteria.

In the present study, we explored further the mechanisms of the *rpoE* operon in controlling the stress tolerance and survival of *B. pseudomallei*. Fortunately, the complete genome sequence and annotation of *B. pseudomallei* (K96243 strain) have recently been made available.¹⁸ We, thus, performed a proteomic analysis of RpoE-associated proteins in *B. pseudomallei*. The cellular proteome of K96243 wild-type (K96243 WT) was compared with that of an *rpoE* operon knockout mutant (RpoE Mut) using a gel-based, differential proteomics strategy. The results showed that several components of stress response proteins were down-regulated in RpoE Mut. Additional findings were the down-regulation of two potential virulence factors. *In vivo* experiments using BALB/c mice showed that animals infected with RpoE Mut had a marked delay in time to death, indicating that the *rpoE* operon is also involved in the virulence of *B. pseudomallei*.

Materials and Methods

Bacterial Culture. *B. pseudomallei* K96243 WT (kindly provided by Prof. T. Dharakul) and *rpoE* operon mutant (RpoE Mut)¹⁷ were maintained in Luria–Bertani (LB) broth at 37 °C until the stationary phase was reached. The RpoE Mut was constructed as described previously by Korbsrisate et al.¹⁷ Briefly, a 270-bp internal fragment of the putative *rpoE* coding sequence was PCR-amplified from *B. pseudomallei* K96243 genomic DNA using the primers ALG36 (5' CTC CAA ATA CCA CCG CAA GAT 3') and ALG37 (5' TAT CCC TTA GTT GGT CCG 3'), which correspond to *B. pseudomallei* *rpoE* nucleotides at the positions of 78–98 and 332–349, respectively. The 270-bp PCR product was cloned into the *EcoRV* restriction site of the pKNOCK-Cm vector¹⁹ to create pPK-1. This construct was introduced from *E. coli* S17–1 λ pir²⁰ into *B. pseudomallei* K96243 by conjugation. An insertion mutant was selected on *Pseudomonas* agar supplemented with SR103 and 30 μ g/mL chloramphenicol.

Southern Blot Analysis. Southern blot hybridization was done according to the method described by Southern.²¹ Briefly,

chromosomes of both strains were digested with restriction enzymes, including *Xho*I and *Xho*I/*EcoRV*, and separated by agarose gel electrophoresis. DNAs in the gel were denatured and transferred onto a nylon membrane using a capillary blotting system. The blot was fixed by baking at 80 °C for 2 h, then hybridized with a 270-bp *rpoE* homologue DNA labeled probe. After eliminating the nonspecific binding of probe, the hybridized bands were detected by radiography.

Protein Extraction for Proteomic Analysis. At the stationary phase with a comparable bacterial count, bacteria were collected using 1000g centrifugation for 5 min and washed three times with phosphate-buffered saline (PBS). Bacterial proteins were extracted using a buffer containing 7 M urea, 2 M thiourea, 4% 3-[(3-cholamidopropyl)dimethylamino]-1-propanesulfonate (CHAPS), 2% (v/v) ampholytes (pH 3–10), 120 mM dithiothreitol (DTT), and 40 mM Tris-base and incubated at 4 °C for 30 min. After centrifugation at 12 000g for 5 min, the supernatant was saved and the protein concentration was measured by spectrophotometry using the Bio-Rad Protein Assay (Bio-Rad Laboratories, Hercules, CA) based on Bradford's method. Because urea, thiourea, CHAPS, and other compositions in the sample/lysis buffer can interfere with the protein estimation, we generated the standard curve using bovine serum albumin at concentrations of 0, 2, 5, 7, and 10 μ g/ μ L in the same sample/lysis buffer to ensure that the standards and the samples had the same background that might occur due to chemical interference. Proteins extracted from each cultured flask were further resolved in individual 2-D gels; $n = 5$ gels (from 5 cultured flasks) for each group; total $n = 10$ gels.

Two-Dimensional Electrophoresis (2-DE) and Staining. Immobiline DryStrip (nonlinear pH 3–10, 7 cm long; Amersham Biosciences, Uppsala, Sweden) was rehydrated overnight with 200 μ g of total protein (equal loading for each sample) that was premixed with a rehydration buffer containing 7 M urea, 2 M thiourea, 2% CHAPS, 2% (v/v) ampholytes (pH 3–10), 120 mM DTT, 40 mM Tris-base, and bromophenol blue (to make the final volume of 150 μ L per strip). The first dimensional separation (IEF) was performed in an Ettan IPGphor II IEF System (Amersham Biosciences) at 20 °C, using a stepwise mode to reach 9000 Vh. After completion of the IEF, proteins on the strip were equilibrated in a buffer containing 6 M urea, 130 mM DTT, 30% glycerol, 112 mM Tris-base, 4% sodium dodecyl sulfate (SDS), and 0.002% bromophenol blue, for 10 min, and then with another buffer containing 6 M urea, 135 mM iodoacetamide, 30% glycerol, 112 mM Tris-base, 4% SDS, and 0.002% bromophenol blue for 10 min. The IPG strip was then transferred onto a 12% acrylamide slab gel (8 $\times$ 9.5 cm), and the second-dimensional separation was performed in an SE260 Mini-Vertical Electrophoresis Unit (Amersham Biosciences) with the current of 20 μ A/gel for 1.5 h. Separated protein spots were then visualized with Coomassie Brilliant Blue R-250 stain (Fluka Chemica AG, Buchs, Switzerland).

Spot Analysis and Matching. Image Master 2D Platinum (Amersham Biosciences) software was used for matching and analysis of protein spots on 2-D gels. Parameters used for spot detection were (i) minimal area = 10 pixels; (ii) smooth factor = 2.0; and (iii) saliency = 2.0. A reference gel was created from an artificial gel combining all of the spots presenting in different gels into one image. The reference gel was then used for matching of corresponding protein spots between gels. Background subtraction was performed, and the intensity volume of each spot was normalized with total intensity volume (summation of the intensity volumes obtained from all spots

within the same 2-D gel). The variability of the 2-D spot pattern was evaluated by determining the coefficient of variation (CV) of the normalized intensity of corresponding spots across different gels using the following formula: $CV = \text{Standard deviation} / \text{Mean}$.

In-Gel Tryptic Digestion. Differentially expressed protein spots were excised from 2-D gels, and the gel pieces were washed with 200 μL of 50% acetonitrile (ACN)/25 mM $\text{NH}_4\text{-HCO}_3$ buffer (pH 8.0) for 15 min twice. The gel pieces were then washed once with 200 μL of 100% ACN and dried using a Speed Vac concentrator (Savant, Holbrook, NY). Dried gel pieces were swollen with 10 μL of 1% (w/v) trypsin (Promega, Madison, WI) in 25 mM NH_4HCO_3 . The gel pieces were then crushed with a siliconized blue stick and incubated at 37 $^\circ\text{C}$ for at least 16 h. Peptides were subsequently extracted twice with 50 μL of 50% ACN/5% trifluoroacetic acid (TFA); the extracted solutions were then combined and dried with the Speed Vac concentrator. The peptide pellets were then resuspended in 10 μL of 0.1% TFA, and the resuspended solutions were purified using ZipTip_{C18} (Millipore, Bedford, MA). Ten microliters of sample was drawn up and down in the ZipTip 10 times and then washed with 10 μL of 0.1% formic acid by drawing up and expelling the washing solution three times. The peptides were finally eluted with 5 μL of 75% ACN/0.1% formic acid.

Protein Identification by Q-TOF MS/MS. The proteolytic samples were premixed 1:1 with the matrix solution (5 mg/mL α -cyano-4-hydroxycinnamic acid (CHCA) in 50% ACN, 0.1% v/v TFA, and 2% w/v ammonium citrate) and spotted onto the 96-well sample stage. The samples were analyzed using the Q-TOF Ultima mass spectrometer (Micromass, Manchester, U.K.), which was fully automated with predefined probe motion pattern and the peak intensity threshold for switching over from MS survey scanning to MS/MS, and from one MS/MS to another. Within each sample well, parent ions that met the predefined criteria (any peak within the m/z 800–3000 range with intensity above 10 count $\pm$ include/exclude list) were selected for CID MS/MS using argon as the collision gas and a mass dependent ± 5 V rolling collision energy until the end of the probe pattern was reached (all details are available at <http://proteome.sinica.edu.tw>). The MASCOT (<http://www.matrixscience.com>) search engine was used for obtaining protein identities and peptide sequences, based on the assumptions that peptides were monoisotopic, oxidized at methionine residues, and carbamidomethylated at cysteine residues. The search was performed using the entire protein databases of the Swiss-Prot and TrEMBL and MSDB. A mass tolerance of 50 ppm was used, and up to 1 missed trypsin cleavage was allowed.

In Vivo Virulence Study. Evaluation of the virulence of RpoE Mut was performed using a pulmonary model of melioidosis in BALB/c mice as described previously.²² In summary, 1000 CFU of either K96243 WT or RpoE Mut was administered *via* the intranasal route ($n = 6$ per group), and the mice were then monitored twice daily for signs of infection and mortality.

Statistical Analysis. Comparisons between groups were performed using unpaired *t* test. For the virulence study, differences in survival of the infected animals were analyzed using a log rank test. *P*-values less than 0.05 were considered statistically significant.

Results and Discussion

Validation of the *rpoE* Operon Knockout Mutant. Southern hybridization was performed to confirm the integration of the

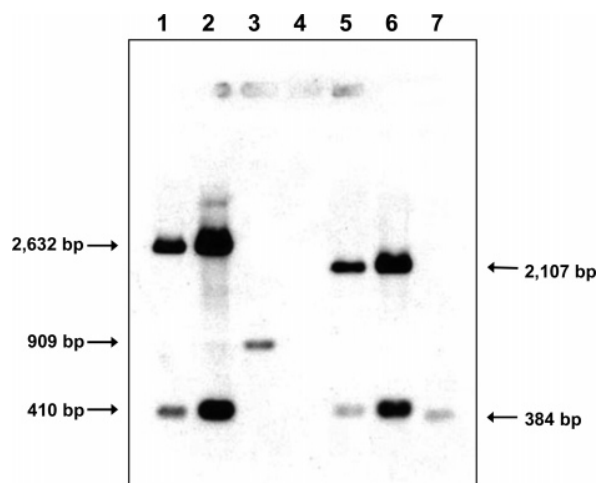


Figure 1. Southern blot analysis of digested *B. pseudomallei* genomic DNA hybridized with a 270-bp *rpoE* specific DNA probe. Lanes 1 and 2, *XhoI*-digested genomic DNA from RpoE Mut; lane 3, *XhoI*-digested genomic DNA from K96243 WT; lane 4, standard marker (1 kb DNA ladder); lanes 5 and 6, *XhoI/EcoRV*-digested genomic DNA from RpoE Mut; lane 7, *XhoI/EcoRV*-digested genomic DNA from K96243 WT.

pKNOCK suicide vector on the *rpoE* operon knockout mutant (RpoE Mut). The genomic DNAs from the K96243 WT and RpoE Mut of *B. pseudomallei* were prepared. Both genomic DNAs were digested with restriction enzymes, including *XhoI* and *EcoRV/XhoI*. Southern blot of the digested genomic DNAs was hybridized with an *rpoE* DNA probe. As expected, one DNA hybridization fragment (909 bp) was detected in the *XhoI* genomic DNA of K96243 WT, whereas two DNA hybridizing fragments (2632 bp and 410 bp) were detected in *XhoI* genomic DNA of RpoE Mut (Figure 1; lanes 1–3). As also predicted, a 384 bp *EcoRV/XhoI* DNA fragment from K96243 WT was detected (Figure 1; lane 7), whereas 2107 bp and 410 bp *EcoRV/XhoI* DNA fragments from RpoE Mut were shown (Figure 1; lanes 5 and 6). These results indicated that there was an insertion of the pKNOCK vector at the *rpoE* gene on the chromosome of *B. pseudomallei* RpoE Mut.

Altered Proteome in RpoE Mut *B. pseudomallei*. As RpoE is a transcriptional factor, it is expected that production of several proteins (gene products) are controlled by RpoE. To explore such gene products and to further investigate the molecular mechanisms of RpoE, we performed a proteomic analysis of differentially expressed proteins in RpoE Mut compared to K96243 WT. Proteins extracted from bacteria in each cultured flask were resolved in individual 2-D gels; $n = 5$ gels (from 5 cultured flasks) for each group; total $n = 10$ gels. Figure 2 shows representative 2-D gels of cellular proteins extracted from K96243 WT and RpoE Mut. Up to 450 protein spots were visualized in each gel. Among these, quantitative intensity analysis and statistics revealed 52 differentially expressed protein spots (with $p < 0.05$) between the K96243 WT and RpoE Mut groups.

These differentially expressed proteins were subsequently identified by Q-TOF MS/MS, and their identities as well as quantitative data are shown in Table 1. The amino acid sequences of peptides identified are provided in Supporting Information (Table S1). These identified proteins were classified into functional groups based on their functions provided in the Swiss-Prot and TrEMBL protein databases and on literature

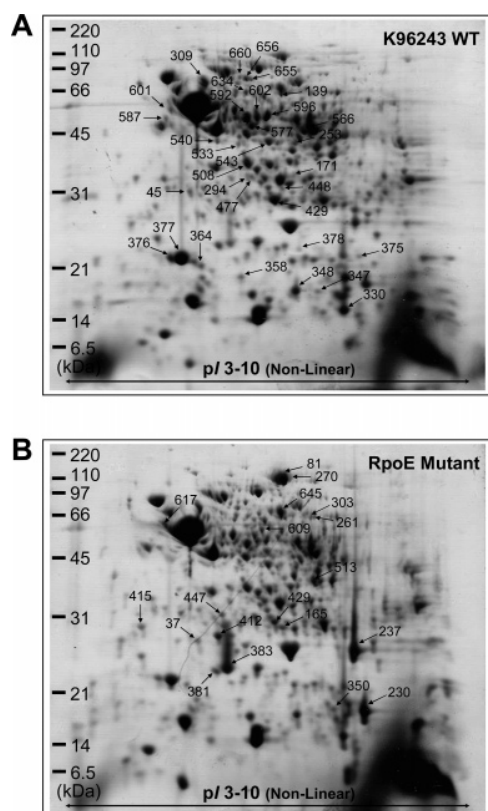


Figure 2. Proteome maps of differentially expressed proteins. (A) Representative 2-D gel for K96243 WT and (B) for RpoE Mut ($n = 5$ gels for each group; total $n = 10$). Quantitative intensity analysis revealed 52 differentially expressed protein spots between the two groups. Down-regulated proteins are labeled with numbers in panel A, whereas up-regulated proteins are labeled in panel B. These differentially expressed proteins were subsequently identified by Q-TOF MS/MS (see Table 1 and Supporting Information).

search in PubMed. Some of the identified proteins were 'hypothetical proteins' of which the functions are unknown or have not previously been determined. However, sequences of some of these hypothetical proteins were identical (100% similarity) or almost identical (89–94% similarity) to the known proteins, of which functions have been clearly defined. Therefore, these hypothetical proteins were functionally classified based on their respective homologues.

Impaired Stress Tolerance and Decreased Intracellular Survival of the *rpoE* Operon Knockout Mutant. Because *B. pseudomallei* is a saprophyte found in soil and can survive in eukaryotic cells as well as in phagocytes, it is expected that this bacterium must have regulatory mechanisms for adaptation in these stressful environments, particularly the enrichment of free radicals or reactive oxygen intermediates and high osmolality. In our previous report [see Korbsrisate et al.¹⁷], we had examined the effects of *rpoE* operon knockout on the susceptibility of *B. pseudomallei* to oxidative (using 100 mM menadione, 1 M H₂O₂, and 4 M H₂O₂) and osmotic stresses (using 2 M NaCl), and evaluated the viability of RpoE Mut inside mammalian (murine) macrophages. The results clearly showed that the *rpoE* operon knockout caused impaired tolerance to the oxidative stress (as the zone of growth inhibition was greater in RpoE Mut compared to K96243 WT) and to the osmotic stress.¹⁷ For the intracellular survival, the

viability of RpoE Mut in mammalian phagocytes was significantly reduced.¹⁷ These data indicated that the *rpoE* operon is crucial for stress tolerance and intracellular survival of *B. pseudomallei*.

Interestingly, 16 out of 52 differentially expressed proteins identified in our present study were oxidative stress response proteins, osmotic stress response proteins, chaperones, and other stress response proteins (Table 1). Almost all of these stress response proteins were down-regulated (RpoE Mut/K96243 WT ratios ranged from 0.24 to 0.79; average = 0.47). The down-regulation (approximately 50% from the baseline) of these stress associated proteins, particularly the oxidative stress response group (i.e., AhpC/Tsa family antioxidant protein, ferritin-like domain protein, flavohemoprotein, and peroxidase/catalase) and chaperones (i.e., 60 kDa chaperonin, GroEL, heat shock protein HtpG, PspA/IM30 family protein, and universal stress protein family), in RpoE Mut was concordant with the phenotype of the *rpoE* operon knockout mutant, of which the stress tolerance was impaired. Osmotically inducible Y domain protein was one among the three up-regulated stress associated proteins. While it was up-regulated, the phenotype of RpoE Mut showed impaired osmotic stress tolerance. These disparate results were not surprising as osmotic stress in several other models can be regulated by various chaperones and other stress associated proteins,^{23–25} some of which were down-regulated in our present study (Table 1). These data strengthened the information that the *rpoE* operon is required for stress tolerance in *B. pseudomallei* via controlling transcription/translation of various stress associated proteins.

Other groups of the differentially expressed proteins were transcriptional/translational regulators and proteins involved in cell wall synthesis. Almost all of these altered proteins were down-regulated or absent (or might be under the detection limit of our experimental procedures) in RpoE Mut. The down-regulation of these two groups of proteins together with the down-regulation of several stress response proteins, as mentioned above, most likely led to the decreased survival of RpoE Mut in mammalian (murine) phagocytes.¹⁷

There were a few groups of proteins (i.e., metabolic enzymes and exported proteins) in which the number of the up-regulated proteins was comparable to that of the down-regulated proteins within the same group. This pattern of changes implicates that there was a balance between the up-regulations and the down-regulations, similar to changes almost always observed in several other models of experimental interventions or gene manipulation in our previous studies.^{26–30} Further examination of their roles and association with RpoE is required to better understand the functional significance of these altered proteins in *B. pseudomallei*.

Impaired Virulence of the *rpoE* Operon Knockout Mutant.

Potential virulence factors of *B. pseudomallei* (which include both well-known and possible ones) are secretion type I–IV systems, surface components (i.e., lipopolysaccharide, capsular polysaccharide, and potential surface polysaccharide biosynthesis), fimbriae/pili, exoproteins (phospholipase C, metalloproteases A, collagenase, and other proteases), and adhesins or adhesive molecules that modulate host-cell interaction.^{3,18} In the present study, we identified the down-regulation of a protease/peptidase (serine-type carboxypeptidase family protein; spot no. 634) and of an adhesion molecule (phospholipid-binding protein; spot no. 364) (Table 1) that may be the potential virulence factors of *B. pseudomallei*. Although there is no direct evidence demonstrating that carboxypeptidase is

Table 1. The Altered Proteins in RpoE Mut

altered proteins	spot no.	Quantity (Intensity) (Mean ± SEM)			ratio (Mut/WT)	change	ID	ion score	%cov	pI	MW
		K96243 WT	RpoE Mut	p							
Oxidative Stress Response Proteins											
Antioxidant, AhpC/Tsa family [Burkholderia mallei] ^a	376	0.5071 ± 0.0466	0.2176 ± 0.0311	0.001	0.43	Down	Q62JI8_BURMA	197	31	5.1	20.46
Antioxidant, AhpC/Tsa family [Burkholderia mallei] ^a	377	2.6473 ± 0.4435	1.3248 ± 0.1059	0.020	0.50	Down	Q62JI8_BURMA	187	31	5.1	20.46
Antioxidant, AhpC/Tsa family [Burkholderia mallei] ^a	429	0.8098 ± 0.0580	0.3810 ± 0.0662	0.001	0.47	Down	Q62I24_BURMA	294	30	5.8	23.90
Ferritin-like domain protein [Burkholderia mallei]	348	1.1054 ± 0.0998	0.6615 ± 0.0450	0.004	0.60	Down	Q62H71_BURMA	153	22	6.0	18.14
Flavoheмоprotein [Burkholderia pseudomallei]	566	0.2350 ± 0.0156	0.1357 ± 0.0061	<0.001	0.58	Down	Q63R34_BURPS	157	16	6.1	43.64
Peroxidase/catalase (EC 1.11.1.6 (Catalase-peroxidase) [Burkholderia mallei]	655	0.0832 ± 0.0089	0.0227 ± 0.0095	0.002	0.27	Down	Q62H74_BURMA	248	6	5.7	79.45
Putative phenylacetic acid degradation oxidoreductase [Burkholderia pseudomallei]	645	0.6072 ± 0.0358	0.7921 ± 0.0618	0.032	1.30	Up	Q63QI2_BURPS	335	10	5.9	60.85
Osmotic Stress Response Proteins											
Osmotically inducible protein Y domain protein [Burkholderia mallei]	412	0.2610 ± 0.0387	0.3757 ± 0.0211	0.032	1.44	Up	Q62E87_BURMA	135	12	5.3	23.15
Other Stress Response Proteins											
60 kDa chaperonin [Burkholderia pseudomallei K96243]	601	0.3586 ± 0.0571	0.1402 ± 0.0121	0.006	0.39	Down	CAH36705	157	10	5.1	57.14
GroEL (Fragment) [Burkholderia pseudomallei]	540	0.1241 ± 0.0093	0.0572 ± 0.0049	<0.001	0.46	Down	Q83WK0_BURPS	195	10	5.2	56.49
GroEL (Fragment) [Burkholderia pseudomallei]	577	0.2034 ± 0.0154	0.0996 ± 0.0109	0.001	0.49	Down	Q83WK0_BURPS	175	10	5.2	56.49
GroEL (Fragment) [Burkholderia pseudomallei]	587	0.1323 ± 0.0202	0.0410 ± 0.0048	0.002	0.31	Down	Q83WK0_BURPS	190	10	5.2	56.49
Heat shock protein HtpG [Burkholderia mallei]	309	0.9076 ± 0.0816	0.6645 ± 0.0665	0.049	0.73	Down	Q62ID1_BURMA	274	14	5.1	71.22
PspA/IM30 family protein (phage shock protein A) (suppresses sigma54-dependent transcription) [Burkholderia mallei]	45	0.2709 ± 0.0174	0.2152 ± 0.0072	0.018	0.79	Down	Q62IH7_BURMA	100	18	5.1	24.48
Trigger factor (EC 5.2.1.8) [Burkholderia mallei]	617	0.2335 ± 0.0191	0.4046 ± 0.0418	0.006	1.73	Up	Q62JK6_BURMA	38	6	5.0	49.71
Universal stress protein family [Burkholderia mallei]	340	0.2697 ± 0.0095	0.0647 ± 0.0066	<0.001	0.24	Down	Q62EI9_BURMA	320	44	5.8	16.62
Transcriptional/Translational Regulators											
Adenosylhomocysteinase [Burkholderia pseudomallei K96243]	596	0.5861 ± 0.0413	0.3618 ± 0.0269	0.002	0.62	Down	CAH37303	140	7	5.7	52.51
Adenosylhomocysteinase [Burkholderia pseudomallei K96243]	602	0.4094 ± 0.0555	0.1955 ± 0.0383	0.013	0.48	Down	CAH37303	186	11	5.7	52.51
Arginine deiminase [Burkholderia pseudomallei K96243]	592	0.7188 ± 0.0774	0.3946 ± 0.0434	0.006	0.55	Down	CAH35742	273	16	5.6	46.42
DNA-directed RNA polymerase alpha chain [Burkholderia pseudomallei K96243]	543	0.3397 ± 0.0050	0.3037 ± 0.0064	0.002	0.89	Down	CAH37198	295	26	5.8	35.78
Endoribonuclease, L-PSP family [Burkholderia mallei]	347	0.1282 ± 0.0075	0.0850 ± 0.0055	0.002	0.66	Down	Q62HN3_BURMA	122	22	6.2	15.95
Hypothetical protein [Burkholderia pseudomallei] ^b	237	1.3144 ± 0.1586	3.0324 ± 0.2261	<0.001	2.31	Up	Q63NT6_BURPS	119	16	6.7	21.65
Putative methyltransferase [Burkholderia pseudomallei]	448	0.0540 ± 0.0137	0.0000 ± 0.0000	0.004	0.00	Absent	Q63L99_BURPS	155	18	5.9	29.52
Ribonuclease PH [Burkholderia pseudomallei]	171	0.0926 ± 0.0109	0.0571 ± 0.0095	0.040	0.62	Down	CAH36573	51	4	6.0	26.21
Transcriptional regulator, AsnC family [Burkholderia mallei]	350	0.0392 ± 0.0025	0.1066 ± 0.0248	0.027	2.72	Up	Q62M96_BURMA	63	11	6.5	19.08
YbaK/prolyl-tRNA synthetases-associated domain family protein [Burkholderia mallei]	358	0.1446 ± 0.0096	0.0758 ± 0.0080	0.001	0.52	Down	Q62C75_BURMA	159	21	5.5	18.14
Metabolic Enzymes											
3-Oxoadipate CoA-succinyl transferase beta subunit [Burkholderia mallei]	415	0.0749 ± 0.0078	0.2232 ± 0.0343	0.003	2.98	Up	Q62KH2_BURMA	236	23	4.7	22.33
Aldehyde dehydrogenase family protein [Burkholderia mallei]	609	0.1036 ± 0.0128	0.2361 ± 0.0371	0.010	2.28	Up	Q62FN8_BURMA	235	16	5.7	55.72
Carbamate kinase (EC 2.7.2.2) [Burkholderia pseudomallei]	508	0.5604 ± 0.0263	0.4655 ± 0.0098	0.010	0.83	Down	Q63U71_BURPS	225	17	5.5	33.51
GMP synthase [glutamine-hydrolyzing] (EC 6.3.5.2) [Burkholderia pseudomallei]	139	0.0845 ± 0.0056	0.0063 ± 0.0063	<0.001	0.07	Down	Q63T42_BURPS	196	11	5.9	60.48
Spermidine n(1)-acetyltransferase (EC 2.3.1.57) [Burkholderia pseudomallei]	378	0.0699 ± 0.0098	0.0000 ± 0.0000	<0.001	0.00	Absent	Q63YU3_BURPS	169	14	5.9	22.22
Urocanate hydratase (EC 4.2.1.49) [Burkholderia mallei]	261	0.0717 ± 0.0042	0.1054 ± 0.0055	0.001	1.47	Up	Q62LJ4_BURMA	125	6	6.1	61.86
Urocanate hydratase (EC 4.2.1.49) [Burkholderia mallei]	303	0.0535 ± 0.0079	0.0849 ± 0.0082	0.025	1.59	Up	Q62LJ4_BURMA	297	11	6.1	61.86

Table 1 (Continued)

altered proteins	spot no.	Quantity (Intensity) (Mean ± SEM)		p	ratio (Mut/WT)	change	ID	ion score	%cov	pI	MW
		K96243 WT	RpoE Mut								
Cell Wall Synthesis Pathway											
Glucose-1-phosphate thymidyltransferase (EC 2.7.7.24) (dTDP-glucose pyrophosphorylase) (dTDP-glucose synthase) [Burkholderia mallei]	477	0.1201 ± 0.0095	0.0742 ± 0.0033	0.002	0.62	Down	Q9AEV6_BURMA	51	6	5.6	33.07
Putative dTDP-D-glucose 4, 6-dehydratase (EC 4.2.1.46) [Burkholderia pseudomallei]	253	0.1977 ± 0.0178	0.1521 ± 0.0059	0.041	0.77	Down	Q9AEV7_BURPS	127	7	6.0	39.34
Glycogen Biosynthesis and Storage Pathway											
Inorganic pyrophosphatase (EC 3.6.1.1) [Burkholderia mallei]	381	0.2477 ± 0.0067	0.4036 ± 0.0300	0.001	1.63	Up	Q62LB4_BURMA	42	17	5.2	19.26
Fatty Acid Synthesis											
Putative beta-ketoacyl-ACP synthase [Burkholderia pseudomallei]	533	0.1011 ± 0.0062	0.0573 ± 0.0019	<0.001	0.57	Down	Q63N14_BURPS	96	10	5.4	40.75
Exported Proteins											
Putative exported protein [Burkholderia pseudomallei]	330	1.2433 ± 0.0680	0.9420 ± 0.0898	0.028	0.76	Down	Q63VW9_BURPS	110	59	9.2	13.75
Putative porin related exported protein [Burkholderia pseudomallei]	513	0.0619 ± 0.0099	0.3803 ± 0.0648	0.001	6.14	Up	Q63JN8_BURPS	188	15	7.8	40.14
Secreted Proteins											
Hypothetical protein [Burkholderia pseudomallei] ^c	447	0.0595 ± 0.0014	0.1156 ± 0.0079	<0.001	1.94	Up	Q63IV8_BURPS	177	22	5.3	23.46
Adhesion Molecule											
Hypothetical protein [Burkholderia mallei] ^d	364	0.1374 ± 0.0121	0.1078 ± 0.0043	0.050	0.78	Down	Q62FL1_BURMA	274	33	5.3	18.34
Protease/Peptidase											
Serine-type carboxypeptidase family protein [Burkholderia mallei]	634	0.1359 ± 0.0104	0.0730 ± 0.0062	0.001	0.54	Down	Q62AX8_BURMA	224	14	5.6	60.18
Protease Inhibitors											
Ecotin [Burkholderia mallei]	375	0.1133 ± 0.0081	0.0000 ± 0.0000	<0.001	0.00	Absent	Q62FK6_BURMA	33	6	8.5	19.47
Signaling Proteins											
Hypothetical protein [Burkholderia pseudomallei] ^e	656	0.1637 ± 0.0064	0.0805 ± 0.0016	<0.001	0.49	Down	Q63TY9_BURPS	101	5	5.5	73.72
Hypothetical protein [Burkholderia pseudomallei] ^e	660	0.0516 ± 0.0043	0.0179 ± 0.0079	0.006	0.35	Down	Q63TY9_BURPS	118	7	5.5	73.72
Miscellaneous and Unknown Function											
Chitin binding protein, putative [Burkholderia pseudomallei]	165	0.1212 ± 0.0344	0.2440 ± 0.0327	0.032	2.01	Up	Q63PN3_BURPS	66	11	6.2	25.89
Hypothetical bacteriophage protein [Burkholderia pseudomallei]	230	1.0938 ± 0.1523	2.5481 ± 0.2037	<0.001	2.33	Up	Q63LD3_BURPS	313	40	8.4	19.48
Hypothetical protein [Burkholderia pseudomallei] ^f	294	0.0998 ± 0.0020	0.0796 ± 0.0067	0.021	0.80	Down	Q63KM8_BURPS	132	15	5.6	31.14
Hypothetical protein [Burkholderia pseudomallei]	37	0.1138 ± 0.0277	0.2363 ± 0.0215	0.008	2.08	Up	Q63UP7_BURPS	285	37	5.1	23.38
Hypothetical protein [Burkholderia pseudomallei]	81	0.8615 ± 0.0449	1.8358 ± 0.0917	<0.001	2.13	Up	Q63KK6_BURPS	224	6	5.9	124.63
Hypothetical protein [Burkholderia pseudomallei]	270	0.0857 ± 0.0155	0.1621 ± 0.0198	0.016	1.89	Up	Q63KK6_BURPS	279	7	5.9	124.63
Hypothetical protein [Burkholderia pseudomallei]	383	0.6012 ± 0.0935	1.7693 ± 0.2578	0.003	2.94	Up	Q63NT7_BURPS	90	16	5.3	22.45

^a Protein is 100% identical to peroxiredoxin (gi|67763109; ZP_00501806). ^b Protein is 94% identical to methylase of polypeptide chain release factors (gi|67760879; ZP_00499595). ^c Protein is 89% identical to predicted periplasmic or secreted lipoprotein [*Burkholderia pseudomallei* S13] (gi|67762369; ZP_00501070). ^d Protein is 91% identical to phospholipid-binding protein [*Burkholderia pseudomallei* S13] (gi|67760526; ZP_00499246). ^e Protein is 100% identical to putative Ser protein kinase [*Burkholderia pseudomallei* S13] (gi|67758379; ZP_00497148). ^f Protein is 100% identical to *O*-methyltransferase involved in polyketide biosynthesis [*Burkholderia pseudomallei* Pasteur] (gi|67754005; ZP_00492924).

a virulence factor of *B. pseudomallei*, several lines of references have demonstrated its potential role in the virulence of *Naegleria fowleri*,³¹ *Porphyromonas gingivalis*,^{32,33} *Brucella abortus*,³⁴ and so forth. Similarly, there are some indirect evidence demonstrating that proteins with a phospholipid-binding domain may play role in the virulence of bacteria.^{35–38}

On the basis of these data, we therefore hypothesized that the *rpoE* operon is also involved in the virulence of *B. pseudomallei*. An *in vivo* experiment was performed to address whether *rpoE* operon knockout affects the virulence of *B. pseudomallei* (Figure 3). All mice challenged with K96243 WT (*n* = 6) succumbed within 7 days, whereas all RpoE Mut infected mice (*n* = 6) survived for the duration of the experiment (terminated at 40 days; *p* < 0.002). At this time, no bacteria were detected in the spleen or lungs of the RpoE Mut

infected mice (limit of detection $\leq$ 100 CFU/organ), suggesting that this mutant is severely attenuated *in vivo*. The decreased virulence of the RpoE Mut was probably due to the down-regulation of potential virulence factors and also from the down-regulation of several various stress response proteins, as well as transcriptional/translational regulators and proteins involved in cell wall synthesis, all of which are very important for the bacteria to survive in the host. These novel findings underline the usefulness of proteomics as a screening tool to generate a new hypothesis from a set of candidate proteins that can finally be confirmed by conventional functional methods.

Technical Concerns. Some technical issues in the present study need to be discussed. First, there are four genes in the *B. pseudomallei* *rpoE* operon, including *rpoE*, *bprE*, *rseB*, and *mucD*, making the *rpoE* operon a four-gene cluster type. In

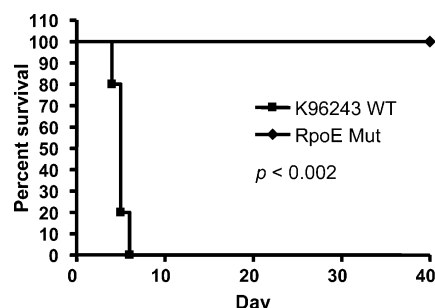


Figure 3. Impaired virulence of *rpoE* operon knockout mutant (RpoE Mut). BALB/c mice (6 per group) were infected intranasally with 1000 cfu of K96243 WT or RpoE Mut, and the survival of the infected mice was monitored.

the present study, we generated an *rpoE* operon knockout mutant. Therefore, the effects of this knockout on the other three genes could also contribute to the altered proteome observed in the present study.

Second, we examined the altered proteome of RpoE Mut only at the stationary phase. Analysis of the altered proteins at the other growth phases should provide more complete information of the molecular mechanisms of the *rpoE* operon in *B. pseudomallei*.

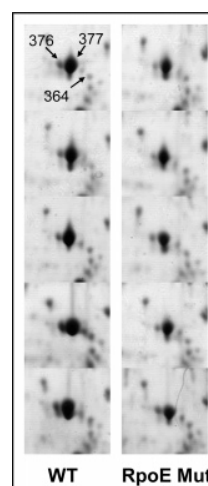
Third, we examined only the cellular proteins of *B. pseudomallei* in the present study. As we identified a periplasmic or secreted lipoprotein as one of the altered proteins, we believe that the production of other secreted proteins is also regulated by the *rpoE* operon. Hence, proteomic analysis of proteins secreted into the culture medium is of potential interest.

Fourth, the total number of protein spots visualized in each gel in the present study was relatively small for cellular proteins extracted from a bacterium. This limitation might be due to the gel-size and staining used in the present study. Many more spots should have been detected using the larger format 2-D gel and a more sensitive stain (i.e., silver or fluorescence).

Finally, we performed proteomic analysis of cellular proteins obtained from 10 different cultures to address the reproducibility and comparability of 2-D spot patterns across different gels, one of the most important issues in gel-based, differential proteomics studies. We calculated the coefficient of variation (CV) on selected spots (nos. 364, 376, and 377) across different gels. Figure 4 illustrates such selected areas and demonstrates the CVs of individual spots that ranged from 0.08 to 0.37 within the K96243 WT or RpoE Mut groups. When all these 3 spots were considered together, the CVs were reduced to 0.29 and 0.13 for K96243 WT and RpoE Mut groups, respectively. Moreover, when all spots detected in each gel were evaluated together, the CVs were only 0.0029 for K96243 WT and 0.0056 for RpoE Mut group. The CVs in our present study were in acceptable ranges of CVs previously observed in standard, gel-based, differential proteomics studies for analyzing tissues, cell lines, and body fluids.^{30,39–41} In addition, the summation of normalized intensity volumes of all spots detected in each 2-D gel was comparable between the two different groups. Therefore, quantitative intensity analysis in our present study was justified.

Conclusions

We have identified a number of *B. pseudomallei* cellular proteins that were altered in an *rpoE* operon knockout mutant using proteomic technology. The down-regulation of several



Coefficient of Variation (CV)		
Spot No.	K96243 WT	RpoE Mut
364	0.1974	0.0889
376	0.2057	0.3196
377	0.3746	0.1787
364+376+377	0.2851	0.1294
All spots in each gel	0.0029	0.0056

Figure 4. Reproducibility and variability of gel-based proteomics approach. Coefficients of variation (CVs) of selected spots were obtained using the formula [CV = SD/Mean]. The results show that CVs in our present study were in acceptable ranges of CVs previously detected in standard, gel-based, differential proteomics studies analyzing human tissues, cell lines, and body fluids.^{30,39–41}

stress response proteins, chaperones, transcriptional/translational regulators, and proteins involved in cell wall synthesis in this mutant provided some new insights into the mechanisms of *rpoE* operon for stress tolerance and survival of *B. pseudomallei*. In addition to the stress tolerance and survival, the proteomic data and *in vivo* study indicated that the *rpoE* operon is also involved in the virulence of *B. pseudomallei* in the mammalian host.

Abbreviations: 2-DE, two-dimensional electrophoresis; ACN, acetonitrile; CHAPS, 3-[(3-cholamidopropyl)dimethylamino]-1-propanesulfonate; CHCA, α -cyano-4-hydroxycinnamic acid; CV, coefficient of variation; DTT, dithiothreitol; IEF, isoelectric focusing; K96243 WT, *B. pseudomallei* wild-type; MN, menadione; Q-TOF MS/MS, quadrupole time-of-flight tandem mass spectrometry; RpoE, sigma factor σ^E ; RpoE Mut, *rpoE* operon knockout mutant of *B. pseudomallei*; SDS, sodium dodecyl sulfate; TFA, trifluoroacetic acid

Acknowledgment. We thank Drs. Pattarachai Kiratisin, Paiboon Vattanaviboon, and Ganjana Lertmemongkolchai for their valuable advice, and are grateful to Jon Cuccui for his assistance in verifying the mutant line for the *in vivo* study. This study was supported by Siriraj Grant for Research and Development, Commission on Higher Education, and The Thailand Research Fund (Grant RMU 5080015) to S. Korbsrisate and V. Thongboonkerd. M. Vanaporn was supported by the Royal Golden Jubilee Ph.D. Program (PHD/0044/2545).

Supporting Information Available: Table listing the sequences of the identified peptides using Q-TOF MS/MS. This material is available free of charge via the Internet at <http://pubs.acs.org>.

References

- (1) Leelarasamee, A. Recent development in melioidosis. *Curr. Opin. Infect. Dis.* **2004**, *17*, 131–136.
- (2) White, N. J. Melioidosis. *Lancet* **2003**, *361*, 1715–1722.
- (3) Cheng, A. C.; Currie, B. J. Melioidosis: epidemiology, pathophysiology, and management. *Clin. Microbiol. Rev.* **2005**, *18*, 383–416.

- (4) Ades, S. E. Control of the alternative sigma factor sigmaE in *Escherichia coli*. *Curr. Opin. Microbiol.* **2004**, *7*, 157–162.
- (5) Moreno, S.; Najera, R.; Guzman, J.; Soberon-Chavez, G.; Espin, G. Role of alternative sigma factor algU in encystment of *Azotobacter vinelandii*. *J. Bacteriol.* **1998**, *180*, 2766–2769.
- (6) Martinez-Salazar, J. M.; Moreno, S.; Najera, R.; Boucher, J. C.; Espin, G.; Soberon-Chavez, G.; Deretic, V. Characterization of the genes coding for the putative sigma factor AlgU and its regulators MucA, MucB, MucC, and MucD in *Azotobacter vinelandii* and evaluation of their roles in alginate biosynthesis. *J. Bacteriol.* **1996**, *178*, 1800–1808.
- (7) Henriques, A. O.; Beall, B. W.; Roland, K.; Moran, C. P., Jr. Characterization of *cotJ*, a sigma E-controlled operon affecting the polypeptide composition of the coat of *Bacillus subtilis* spores. *J. Bacteriol.* **1995**, *177*, 3394–3406.
- (8) Shazand, K.; Frandsen, N.; Stragier, P. Cell-type specificity during development in *Bacillus subtilis*: the molecular and morphological requirements for sigma E activation. *EMBO J.* **1995**, *14*, 1439–1445.
- (9) Craig, J. E.; Nobbs, A.; High, N. J. The extracytoplasmic sigma factor, final sigma(E), is required for intracellular survival of nontypeable *Haemophilus influenzae* in J774 macrophages. *Infect. Immun.* **2002**, *70*, 708–715.
- (10) Manganelli, R.; Fattorini, L.; Tan, D.; Iona, E.; Orefici, G.; Altavilla, G.; Cusattelli, P.; Smith, I. The extra cytoplasmic function sigma factor sigma(E) is essential for *Mycobacterium tuberculosis* virulence in mice. *Infect. Immun.* **2004**, *72*, 3038–3041.
- (11) Hershberger, C. D.; Ye, R. W.; Parsek, M. R.; Xie, Z. D.; Chakrabarty, A. M. The algT (algU) gene of *Pseudomonas aeruginosa*, a key regulator involved in alginate biosynthesis, encodes an alternative sigma factor (sigma E). *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 7941–7945.
- (12) Schurr, M. J.; Yu, H.; Boucher, J. C.; Hibler, N. S.; Deretic, V. Multiple promoters and induction by heat shock of the gene encoding the alternative sigma factor AlgU (sigma E) which controls mucoidy in cystic fibrosis isolates of *Pseudomonas aeruginosa*. *J. Bacteriol.* **1995**, *177*, 5670–5679.
- (13) Schnider-Keel, U.; Lejbolle, K. B.; Baehler, E.; Haas, D.; Keel, C. The sigma factor AlgU (AlgT) controls exopolysaccharide production and tolerance towards desiccation and osmotic stress in the biocontrol agent *Pseudomonas fluorescens* CHA0. *Appl. Environ. Microbiol.* **2001**, *67*, 5683–5693.
- (14) Kenyon, W. J.; Sayers, D. G.; Humphreys, S.; Roberts, M.; Spector, M. P. The starvation-stress response of *Salmonella enterica* serovar Typhimurium requires sigma(E)-, but not CpxR-regulated extracytoplasmic functions. *Microbiology* **2002**, *148*, 113–122.
- (15) Bralley, P.; Jones, G. H. Transcriptional analysis and regulation of the sigma-E gene of *Streptomyces antibioticus*. *Biochim. Biophys. Acta* **2001**, *1517*, 410–415.
- (16) Kovackova, G.; Skorupski, K. The alternative sigma factor sigma(E) plays an important role in intestinal survival and virulence in *Vibrio cholerae*. *Infect. Immun.* **2002**, *70*, 5355–5362.
- (17) Korbsrisate, S.; Vanaporn, M.; Kerdsuk, P.; Kespichayawattana, W.; Vattanaviboon, P.; Kiatpapan, P.; Lertmemongkolchai, G. The *Burkholderia pseudomallei* RpoE (AlgU) operon is involved in environmental stress tolerance and biofilm formation. *FEMS Microbiol. Lett.* **2005**, *252*, 243–249.
- (18) Holden, M. T.; Titball, R. W.; Peacock, S. J.; Cerdano-Tarraga, A. M.; Atkins, T.; Crossman, L. C.; Pitt, T.; Churcher, C.; Mungall, K.; Bentley, S. D.; Sebahia, M.; Thomson, N. R.; Bason, N.; Beacham, I. R.; Brooks, K.; Brown, K. A.; Brown, N. F.; Challis, G. L.; Cherevach, I.; Chillingworth, T.; Cronin, A.; Crossett, B.; Davis, P.; DeShazer, D.; Feltwell, T.; Fraser, A.; Hance, Z.; Hauser, H.; Holroyd, S.; Jagels, K.; Keith, K. E.; Maddison, M.; Moule, S.; Price, C.; Quail, M. A.; Rabinowitsch, E.; Rutherford, K.; Sanders, M.; Simmonds, M.; Songsivilai, S.; Stevens, K.; Tumapa, S.; Vesaratchavest, M.; Whitehead, S.; Yeats, C.; Barrell, B. G.; Oyston, P. C.; Parkhill, J. Genomic plasticity of the causative agent of melioidosis, *Burkholderia pseudomallei*. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 14240–14245.
- (19) Alexeyev, M. F. The pKNOCK series of broad-host-range mobilizable suicide vectors for gene knockout and targeted DNA insertion into the chromosome of gram-negative bacteria. *Bio-Techniques* **1999**, *26*, 824–826, 828.
- (20) de, L. V.; Timmis, K. N. Analysis and construction of stable phenotypes in gram-negative bacteria with Tn5- and Tn10-derived minitransposons. *Methods Enzymol.* **1994**, *235*, 386–405.
- (21) Southern, E. M. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **1975**, *98*, 503–517.
- (22) Stevens, M. P.; Haque, A.; Atkins, T.; Hill, J.; Wood, M. W.; Easton, A.; Nelson, M.; Underwood-Fowler, C.; Titball, R. W.; Bancroft, G. J.; Galyov, E. E. Attenuated virulence and protective efficacy of a *Burkholderia pseudomallei* bsa type III secretion mutant in murine models of melioidosis. *Microbiology* **2004**, *150*, 2669–2676.
- (23) Hennequin, C.; Collignon, A.; Karjalainen, T. Analysis of expression of GroEL (Hsp60) of *Clostridium difficile* in response to stress. *Microb. Pathog.* **2001**, *31*, 255–260.
- (24) Beck, F. X.; Grunbein, R.; Lugmayr, K.; Neuhofer, W. Heat shock proteins and the cellular response to osmotic stress. *Cell. Physiol. Biochem.* **2000**, *10*, 303–306.
- (25) Yang, X. X.; Maurer, K. C.; Molanus, M.; Mager, W. H.; Siderius, M.; Vies, S. M. The molecular chaperone Hsp90 is required for high osmotic stress response in *Saccharomyces cerevisiae*. *FEMS Yeast Res.* **2006**, *6*, 195–204.
- (26) Thongboonkerd, V.; Luengpailin, J.; Cao, J.; Pierce, W. M.; Cai, J.; Klein, J. B.; Doyle, R. J. Fluoride exposure attenuates expression of *Streptococcus pyogenes* virulence factors. *J. Biol. Chem.* **2002**, *277*, 16599–16605.
- (27) Thongboonkerd, V.; Gozal, E.; Sachleben, L. R.; Arthur, J. M.; Pierce, W. M.; Cai, J.; Chao, J.; Bader, M.; Pesquero, J. B.; Gozal, D.; Klein, J. B. Proteomic analysis reveals alterations in the renal kallikrein pathway during hypoxia-induced hypertension. *J. Biol. Chem.* **2002**, *277*, 34708–34716.
- (28) Thongboonkerd, V.; Klein, J. B.; Pierce, W. M.; Jevans, A. W.; Arthur, J. M. Sodium loading changes urinary excretion: A proteomic analysis. *Am. J. Physiol. Renal Physiol.* **2003**, *284*, F1155–F1163.
- (29) Thongboonkerd, V.; Barati, M. T.; McLeish, K. R.; Benarafa, C.; Remold-O'Donnell, E.; Zheng, S.; Rovin, B. H.; Pierce, W. M.; Epstein, P. N.; Klein, J. B. Alterations in the renal elastin-elastase system in Type 1 diabetic nephropathy identified by proteomic analysis. *J. Am. Soc. Nephrol.* **2004**, *15*, 650–662.
- (30) Thongboonkerd, V.; Chutipongtanate, S.; Kanlaya, R.; Songtawee, N.; Sinchaikul, S.; Parichatikanond, P.; Chen, S. T.; Malasit, P. Proteomic identification of alterations in metabolic enzymes and signaling proteins in hypokalemic nephropathy. *Proteomics* **2006**, *6*, 2273–2285.
- (31) Hu, W. N.; Kopachik, W.; Band, R. N. Cloning and characterization of transcripts showing virulence-related gene expression in *Naegleria fowleri*. *Infect. Immun.* **1992**, *60*, 2418–2424.
- (32) Chen, Y. Y.; Cross, K. J.; Paolini, R. A.; Fielding, J. E.; Slakeski, N.; Reynolds, E. C. CPG70 is a novel basic metallo-carboxypeptidase with C-terminal polycystic kidney disease domains from *Porphyromonas gingivalis*. *J. Biol. Chem.* **2002**, *277*, 23433–23440.
- (33) Veith, P. D.; Chen, Y. Y.; Reynolds, E. C. *Porphyromonas gingivalis* RgpA and Kgp proteinases and adhesins are C terminally processed by the carboxypeptidase CPG70. *Infect. Immun.* **2004**, *72*, 3655–3657.
- (34) Kikuchi, H.; Kim, S.; Watanabe, K.; Watarai, M. *Brucella abortus* -alanyl-D-alanine carboxypeptidase contributes to its intracellular replication and resistance against nitric oxide. *FEMS Microbiol. Lett.* **2006**, *259*, 120–125.
- (35) Fox, D. S.; Cox, G. M.; Heitman, J. Phospholipid-binding protein Ctl1 controls septation and functions coordinately with calcineurin in *Cryptococcus neoformans*. *Eukaryotic Cell* **2003**, *2*, 1025–1035.
- (36) Khursigara, C.; Abul-Milh, M.; Lau, B.; Giron, J. A.; Lingwood, C. A.; Barnett, Foster, D. E. Enteropathogenic *Escherichia coli* virulence factor bundle-forming pilus has a binding specificity for phosphatidylethanolamine. *Infect. Immun.* **2001**, *69*, 6573–6579.
- (37) Naylor, C. E.; Eaton, J. T.; Howells, A.; Justin, N.; Moss, D. S.; Titball, R. W.; Basak, A. K. Structure of the key toxin in gas gangrene. *Nat. Struct. Biol.* **1998**, *5*, 738–746.
- (38) Titball, R. W. Gas gangrene: an open and closed case. *Microbiology* **2005**, *151*, 2821–2828.
- (39) Hunt, S. M.; Thomas, M. R.; Sebastian, L. T.; Pedersen, S. K.; Harcourt, R. L.; Sloane, A. J.; Wilkins, M. R. Optimal replication and the importance of experimental design for gel-based quantitative proteomics. *J. Proteome Res.* **2005**, *4*, 809–819.
- (40) Terry, D. E.; Desiderio, D. M. Between-gel reproducibility of the human cerebrospinal fluid proteome. *Proteomics* **2003**, *3*, 1962–1979.
- (41) Molloy, M. P.; Brzezinski, E. E.; Hang, J.; McDowell, M. T.; VanBogelen, R. A. Overcoming technical variation and biological variation in quantitative proteomics. *Proteomics* **2003**, *3*, 1912–1919.

PR060457T

Urinary Proteome Profiling Using Microfluidic Technology on a Chip

Visith Thongboonkerd,^{*,†} Napat Songtawee,[†] and Suchai Sritippayawan[‡]

*Medical Molecular Biology Unit, Office for Research and Development, and Division of Nephrology,
Department of Medicine, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand*

Received November 7, 2006

Abstract: Clinical diagnostics and biomarker discovery are the major focuses of current clinical proteomics. In the present study, we applied microfluidic technology on a chip for proteome profiling of human urine from 31 normal healthy individuals (15 males and 16 females), 6 patients with diabetic nephropathy (DN), and 4 patients with IgA nephropathy (IgAN). Using only 4 μ L of untreated urine, automated separation of proteins/peptides was achieved, and 1–7 (3.8 ± 0.3) spectra/bands of urinary proteins/peptides were observed in the normal urine, whereas 8–16 (11.3 ± 1.2) and 9–14 (10.8 ± 1.2) spectra were observed in urine samples of DN and IgAN, respectively. Coefficient of variations of amplitudes of lower marker (1.2 kDa), system spectra (6–8 kDa), and upper marker (260.0 kDa) were 22.84, 24.92, and 32.65%, respectively. ANOVA with Tukey post-hoc multiple comparisons revealed 9 spectra of which amplitudes significantly differed between normal and DN urine (DN/normal amplitude ratios ranged from 2.9 to 3102.7). Moreover, the results also showed that 3 spectra (with molecular masses of 12–15, 27–28, and 34–35 kDa) were significantly different between DN and IgAN urine (DN/IgAN amplitude ratios ranged from 3.9 to 7.4). In addition to the spectral amplitudes, frequencies of some spectra could differentiate the normal from the diseased urine but could not distinguish between DN and IgAN. There was no significant difference, regarding the spectral amplitude or frequency, observed between males and females. These data indicate that the microfluidic chip technology is applicable for urinary proteome profiling with potential uses in clinical diagnostics and biomarker discovery.

Keywords: proteomics • proteome • urine • microfluidics • chip • arrays • diagnostics • biomarker discovery

Introduction

Clinical diagnostics and biomarker discovery (for more accurate and earlier diagnosis of human diseases, as well as

for prediction or monitoring of therapeutic outcome) are the major focuses in current clinical proteomics.^{1,2} The urine is one of the most useful resources for such study, as it is available in most of patients and its collection is noninvasive and very simple. Hence, urinary proteome profiling has been extensively applied to human diseases during the past 4–5 years.^{3,4} To date, the most commonly used methodology in urinary proteomics has been gel-based proteomic technology; i.e., using two-dimensional electrophoresis (2-DE).⁵ This technique is not difficult to perform and is available in most of proteomic laboratories. However, some limitations prevent its use in clinical settings: (i) experimental procedures and spot analyses are labor intensive and impractical for a large number of urine samples; (ii) proteins need to be concentrated and salts must be removed prior to analysis;⁶ (iii) a considerable amount (volume) of urine from each individual is required, thus, space occupying for storage at -20°C to -80°C ; (iv) recovery of hydrophobic proteins is problematic;⁷ and (v) visualization of resolved proteins depends on sensitivity and linear dynamic range of the stain.^{8,9} Although some of these limitations can be reduced with a less sophisticated gel-based technique (i.e., SDS-PAGE or 1-DE), sample preparation to concentrate urinary proteins prior to analysis is still required. Not only the long duration of sample processing but also the loss of some protein components due to an incomplete recovery that may generate a bias for data analysis remain the major disadvantages of both 1-DE and 2-DE. It is almost impossible to analyze normal human urinary proteome from the untreated urine using gel-based proteomic technology. Therefore, an appropriate and rapid method for proteome analysis of a small volume of untreated urine is crucially required.

During the past decade, miniaturization with microfluidics has gained a wide acceptance for clinical applications.^{10,11} Advantages of microfluidic technology include smaller volume of samples required, reduced reagent consumption, higher sensitivity, decreased analytical time, higher throughput and automation, and smaller footprints of analytical devices.^{12,13} The microfluidic technology on a chip is one among currently available microfluidic devices that require a tiny amount (down to 10^{-18} liters) of biological fluids to be analyzed on multichannels chip with dimension of tens to hundreds of micrometers.^{14,15} Therefore, this lab-on-a-chip is feasible for bedside applications, particularly for clinical diagnosis, prognosis, and prediction of therapeutic response. In addition, the microfluidic chip technology can be applied also to biomarker and drug discovery.¹⁶ Previously, a number of studies had applied microfluidic technology on a chip to examine urine. However, most of these studies had evaluated only on urine chemistry

* To whom correspondence should be addressed. Visith Thongboonkerd, MD, FRCPT, Medical Molecular Biology Unit, Office for Research and Development, 12th Floor Adulyadej Vikrom Building, Siriraj Hospital, 2 Prannok Road, Bangkoknoi, Bangkok 10700, Thailand; Phone/Fax, +66-2-4184793; E-mail, thongboonkerd@dr.com or vthongbo@yahoo.com.

[†] Medical Molecular Biology Unit.

[‡] Division of Nephrology.

and urinary levels of chemical compounds, drug metabolites, microalbuminuria, and total protein.^{15,17–21} None of these studies had analyzed urinary protein pattern (proteome profile) using microfluidic technology on a chip.

In the present study, we evaluated the applicability of microfluidic technology on a chip for urinary proteome profiling. Untreated urine samples obtained from 31 normal healthy individuals, 6 patients with diabetic nephropathy (DN), and 4 patients with IgA nephropathy (IgAN) were analyzed with an automated microfluidic system, and their proteome profiles were compared. The data indicated that urinary proteome profiles of the three different groups could be differentiated and the potential disease-specific urinary biomarkers could be determined.

Materials and Methods

Urine Collection. Urine samples were collected from 31 normal healthy individuals (15 males, age 27.62 ± 0.66 years; and 16 females, age = 26.29 ± 0.52 years), 6 DN patients (3 males and 3 females, age = 55.3 ± 5.1 years, serum creatinine = 5.08 ± 3.81 mg/dL, urine protein/creatinine ratio = 5.28 ± 6.02), and 4 IgAN patients (4 females, 51.0 ± 3.3 years, serum creatinine = 1.97 ± 0.94 mg/dL, urine protein/creatinine ratio = 1.73 ± 0.49). All 6 DN patients had diabetic retinopathy and overt proteinuria without urinary tract infection, whereas all 4 IgAN patients were diagnosed by renal biopsy with characteristic IgA deposition in the glomeruli. All normal healthy individuals had no recent illness or medication. All females had no menstrual cycle at the time of collection. This study was approved by the Institutional Ethical Committee. The collection of patients' urine was done prior to treatment and before the biopsy or at least 1 month after the biopsy. Cell debris and particulate matters were removed immediately after the collection using low-speed centrifugation ($1000 \times g$ for 5 min), and supernatants were then saved at -70°C until analyzed.

Microfluidic System, Analytical Procedures, and Data Analysis. The Experion microfluidic separation system (Bio-Rad Laboratories, Hercules, CA) and Pro260 analysis kit (Bio-Rad Laboratories) were used in the present study. This LabChip microfluidic-based system (Figure 1) has integrated separation (based on differential molecular masses), detection (using fluorescence dye), and data analysis within a single platform. The microfluidic chip, which contained a series of plastic wells bonded over a small glass plate, was placed at the center of this automated separation system. The chip's glass plate was etched with an optimized network of microchannels, each of which intersects with others of the plastic wells. Once these channels were primed with a gel matrix and the samples were applied to the appropriate wells, the separation station could direct the samples through these microchannels by specifically controlling the applied voltages and currents.

The Pro260 molecular weight ladder and urine samples (4 μL each) were mixed with 2 μL of reducing sample buffer (with β -mercaptoethanol) and heated at 95°C for 5 min. After cooling at room temperature, 84 μL of deionized (18.2 megaOhm·cm) water was added to each sample. The chip was placed onto the chip platform and prepared by carefully loading the filtered gel-stain solution (12 μL) into the top-right well (priming well) as detailed in the manufacturer's instruction. Bubble was avoided or eliminated prior to the next step. The priming was automatically done within 60 s, and any remaining gel-stain solution at the priming well was discarded. Thereafter, the filtered gel-stain solution was loaded into all four wells labeled

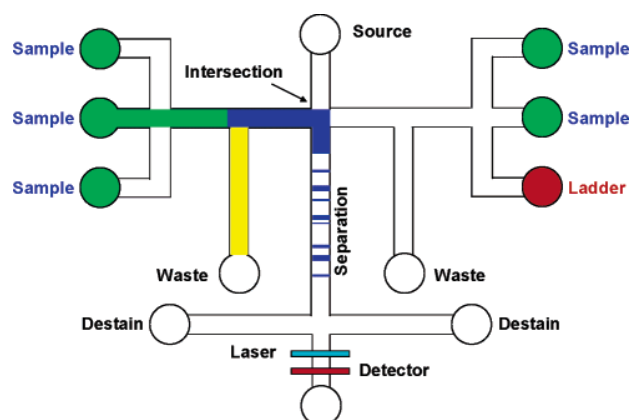


Figure 1. Microfluidics-based separation system. This microfluidic system uses electrokinetic force (as a result of current and voltage) to control motion of the charged species through the microfluidic chip via a polymer gel/sieving matrix. Prior to sample loading, the gel solution from the source well is introduced into the microchannels by a pressure-driven process called “priming”. Thereafter, a preprogrammed series of voltage changes is applied at the sample wells, allowing a small amount of the sample to be directed from one channel toward the intersection and then separated (electrophoresis). A fluorescent dye added to the gel is then incorporated into the sample components along the separation or electrophoresis path. After destaining, the resolved proteins pass through the laser at which the dye is excited and emits a fluorescent signal that can be detected by the detector. (Kindly provided by Laura Madia).

with “GS” (gel-stain) (12 μL each), whereas another well labeled with “G” (gel) was loaded with 12 μL of the filtered gel solution. The urine samples diluted with reducing buffer were then loaded into individual sample wells, labeled with #1–10 (6 μL each), whereas the diluted ladder solution (6 μL) was loaded into a well labeled with “L” (ladder). The loaded chip was then run immediately in the Experion microfluidic system. The data were obtained and analyzed with the Experion analytical software (Bio-Rad Laboratories). More details of this microfluidic separation system can be found in the manufacturer's instructions.

Statistical Analysis. Comparisons among groups were performed using ANOVA with Tukey post-hoc test for multiple comparisons. *p* values less than 0.05 were considered statistically significant, and the data are reported as mean $\pm$ SEM. The difference in frequency of any protein (spectral data) was arbitrarily determined when the presence of a protein was greater than 80% in one group but was less than 20% in the other group. The coefficient of variation (CV) was calculated using the formula: $\%CV = \text{standard deviation}/\text{mean} \times 100\%$.

Results and Discussion

The microfluidic technology on a chip, applied herein, used precision-engineered lower and upper internal alignment markers to provide clean baselines and accurate microfluidic-based molecular mass sizing and protein quantitation. The molecular weight ladder containing 9 highly purified recombinant proteins, with molecular masses from 10 to 260 kDa, was used for molecular mass determination of urinary proteins and/or polypeptides (the word “proteins” will be used for both meanings throughout the manuscript). The sample buffer contained a lower marker (1.2 kDa) and an upper marker (260 kDa), which were used for the proper alignment of samples to

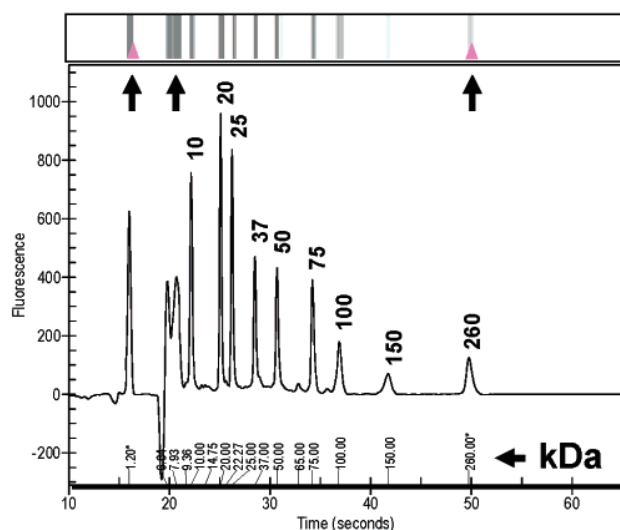
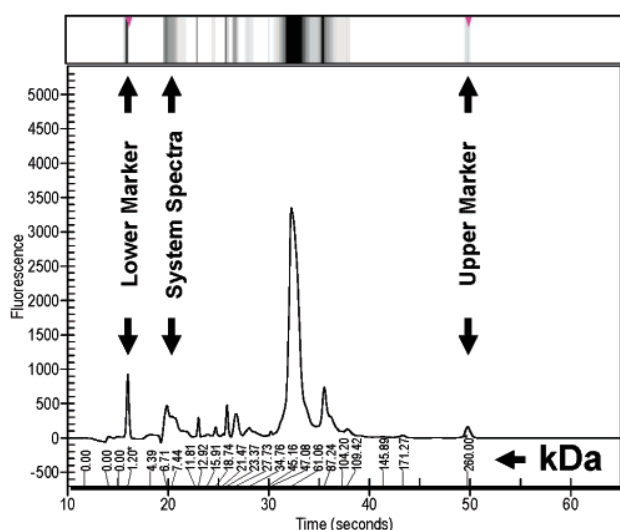
A Molecular Weight Ladder**B Urinary Proteome Profile**

Figure 2. Examples of spectral and band formats of data obtained from the Experion microfluidic chip technology. (A) Spectra and bands of the molecular weight ladder containing 9 highly purified proteins (indicated with 9 numbers of their molecular masses). (B) Spectra and bands of proteins present in a urine sample obtained from a patient with proteinuria. The arrows indicate lower marker (1.2 kDa), system spectra (6–8 kDa), and upper marker (260 kDa), respectively. x-axis represents running time, whereas y-axis shows fluorescence intensity or quantitative data (arbitrary unit).

the molecular weight ladder. The upper marker was also used for normalization of quantitative data of individual proteins. Using the reagents and supplies included with the analysis kit, this microfluidic system could achieve separation and analysis of proteins with molecular masses from 10 kDa to approximately 260 kDa.

Figure 2A shows the data, in both spectral and band formats, of the molecular weight ladder. The arrows indicate lower marker, system spectra, and upper marker. System spectra (6–8 kDa) appeared in every sample following the 1.2 kDa lower

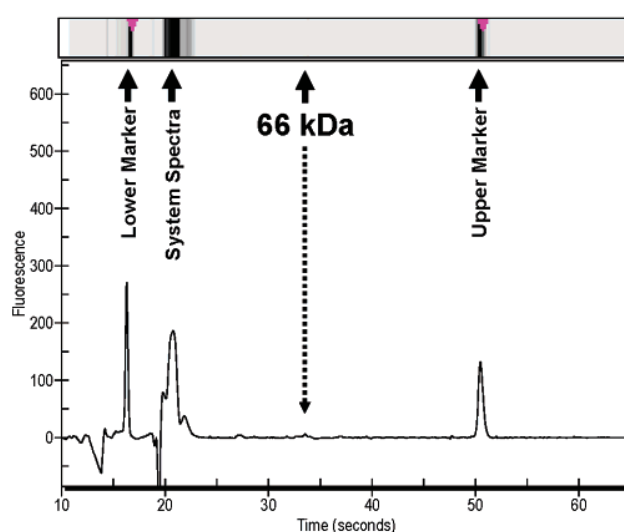
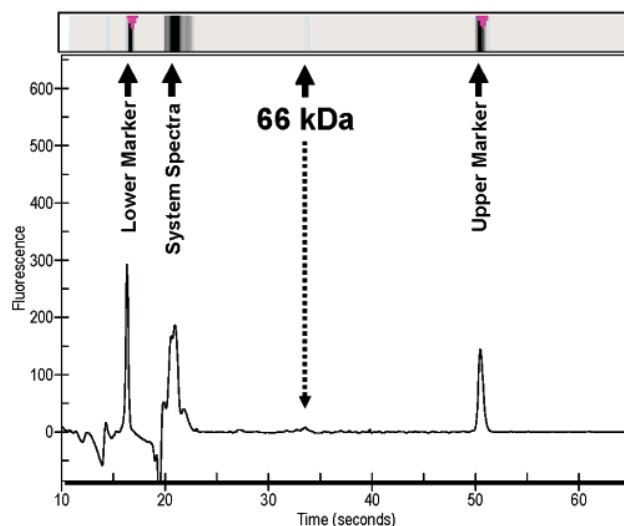
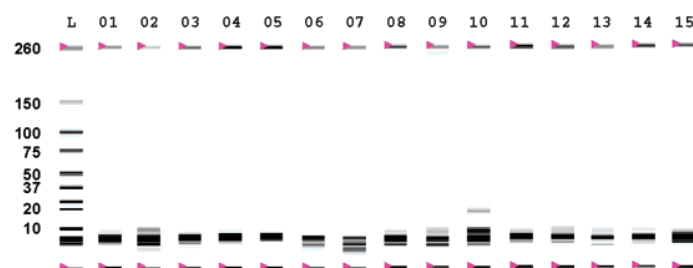
A Normal Urine 1A**B Normal Urine 1B**

Figure 3. Reproducibility. The data obtained from two different runs, (A) and (B), of the same sample were reproducible. There was a faint band at a molecular mass of 66 kDa, which was expected to be albumin, detected in this urine sample.

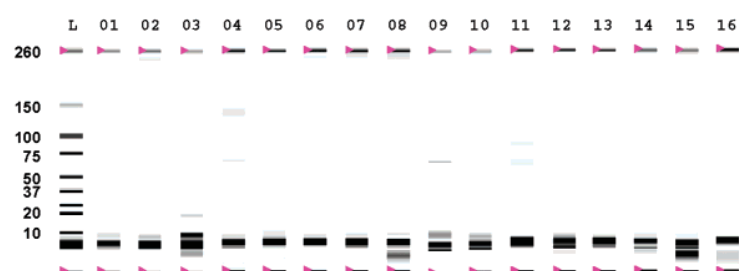
marker. This area consisted of low molecular weight detergent micelle complexes. Also, spectral peaks of the nine highly purified proteins used in the molecular weight ladder (10, 20, 25, 37, 50, 75, 100, 150, and 260 kDa) are clearly illustrated. Figure 2B demonstrates an example of data, in both spectral and band formats, obtained from urine sample of a patient with proteinuria with the prominent band, which was expected to be albumin. Figure 3 shows the data obtained from urine sample of a normal healthy individual. In addition to lower marker, system spectra, and upper marker, there was a faint band or small peak detected at the molecular mass of 66 kDa that was expected to be albumin. The figure also shows the reproducibility of the results obtained from two different runs of the same sample.

We applied this microfluidic technology on a chip to examine urinary proteome profiles of normal healthy individuals ($n =$

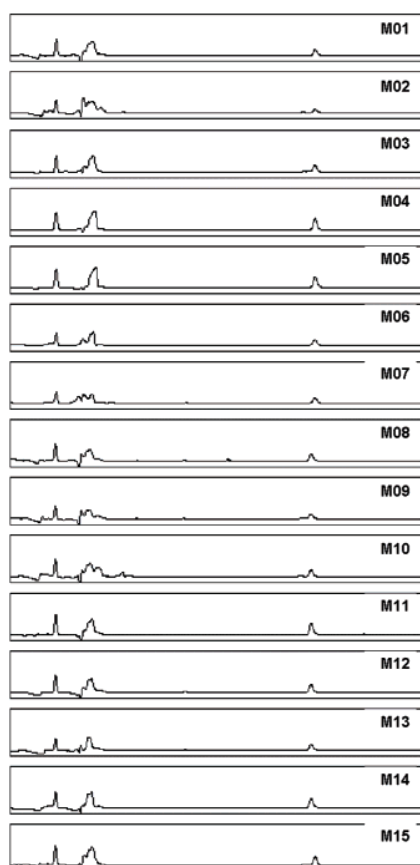
A. Normal Males



B. Normal Females



C. Normal Males



D. Normal Females

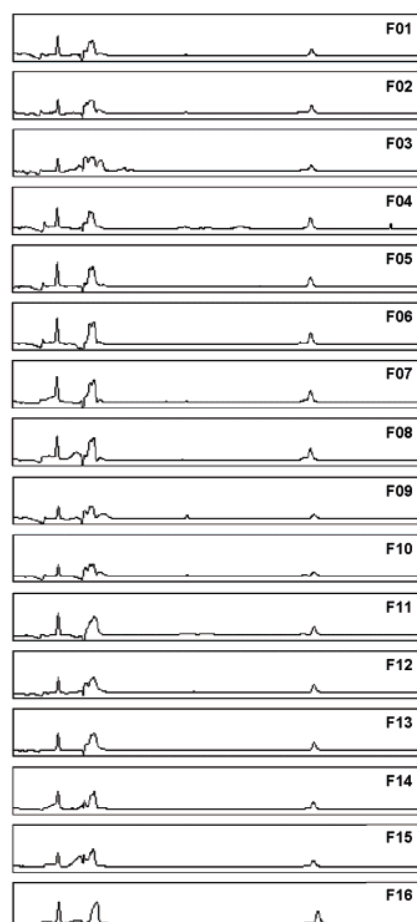


Figure 4. Normal urinary proteome profile. (A) and (B) show band format of the results obtained from 15 normal healthy males and 16 normal healthy females, respectively, whereas (C) and (D) illustrate spectral data of all these normal healthy individuals.

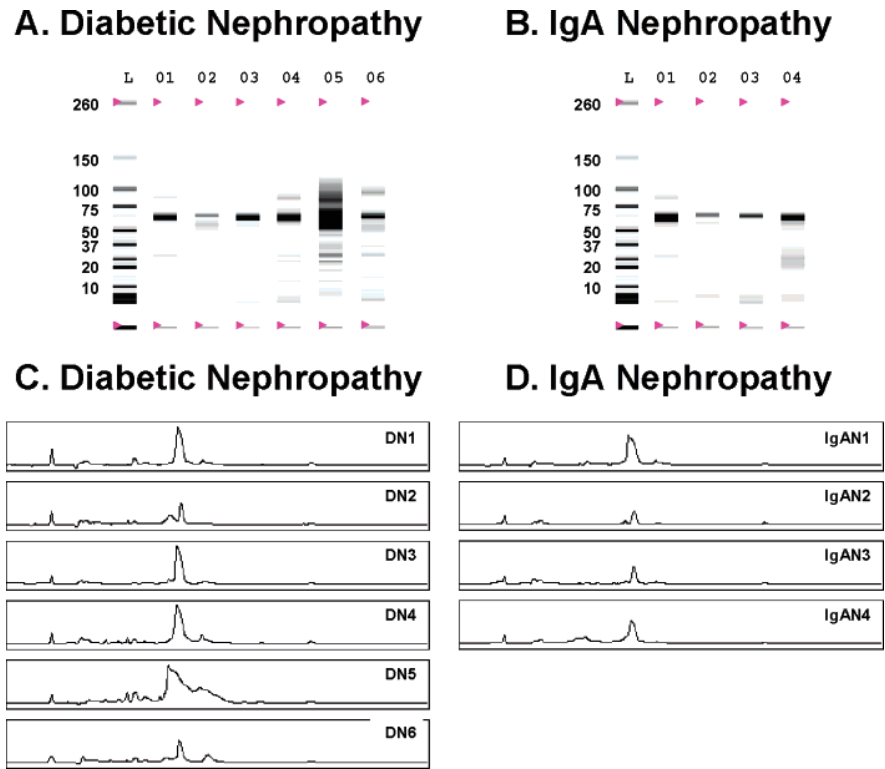


Figure 5. Urinary proteome profiles of diabetic nephropathy (DN) and IgA nephropathy (IgAN). (A) and (B) show band format of the results obtained from six patients with DN and four patients with IgAN, respectively, whereas (C) and (D) illustrate their respective spectral data.

Table 1. Amplitudes and Frequencies of Protein Spectra/Bands Detected in Urine Samples of Normal Healthy Individuals

band	normal males (<i>N</i> = 15)		normal females (<i>N</i> = 16)		<i>p</i> values	amplitude ratio
	amplitude (mean ± SEM)	frequency (%)	amplitude (mean ± SEM)	frequency (%)		male/female
9–10 kDa	143.6 ± 56.9	9/15 (60.0)	160.2 ± 52.9	11/16 (68.8)	0.832	1.12
12–15 kDa	3.4 ± 3.4	1/15 (6.7)	0.2 ± 0.2	1/16 (6.3)	0.361	0.06
18–20 kDa	12.3 ± 12.3	1/15 (6.7)	7.8 ± 7.8	2/16 (12.5)	0.758	0.63
22–24 kDa	3.5 ± 3.3	3/15 (20.0)	0.8 ± 0.8	1/16 (6.3)	0.428	0.22
34–35 kDa	0.6 ± 0.6	1/15 (6.7)	0.0 ± 0.0	0/16 (0.0)	0.334	0.00
61–68 kDa	4.1 ± 1.4	8/15 (53.3)	19.6 ± 9.2	14/16 (87.5)	0.119	4.73
86–90 kDa	0.9 ± 0.7	2/15 (13.3)	6.9 ± 6.7	3/16 (18.8)	0.385	7.93
	band count for males (mean ± SEM)		band count for females (mean ± SEM)		<i>p</i> values	ratio of band count male/female
all bands	3.4 ± 0.4		4.2 ± 0.4		0.195	1.24

Table 2. Coefficients of Variations (CVs) of Amplitudes of Lower Marker, System Spectra, and Upper Marker

band	males (<i>N</i> = 15)		females (<i>N</i> = 16)		males and females (<i>N</i> = 31)	
	amplitude (mean ± SD)	CV (%)	amplitude (mean ± SD)	CV (%)	amplitude (mean ± SD)	CV (%)
lower marker (1.2 kDa)	366.83 ± 69.74	19.01	388.00 ± 100.62	25.93	377.75 ± 86.30	22.84
system peak (6–8 kDa)	341.29 ± 97.54	28.58	361.19 ± 79.19	21.92	351.56 ± 87.62	24.92
upper marker (260 kDa)	167.81 ± 59.81	35.64	168.02 ± 51.68	30.76	167.92 ± 54.82	32.65

31; 15 males and 16 females). Figure 4A and B shows their band data, whereas Figure 4C and D demonstrate their spectral data. There were 1–7 bands/spectra detected in each normal urine sample. Table 1 summarizes amplitudes and frequencies of

these bands/spectra. When we compared the data of males to those of females, band counts were comparable (3.4 ± 0.4 in males versus 4.2 ± 0.4 in females; *p* was not significant), and there was no significant difference observed in either ampli-

Table 3. Comparisons of Amplitude and Frequency Data of Protein Spectra/Bands among Different Groups (Normal vs DN vs IgAN)

band (kDa)	normal (<i>N</i> = 31)		DN (<i>N</i> = 6)		IgAN (<i>N</i> = 4)		ANOVA	Tukey's post-hoc tests of amplitude			amplitude ratio	
	amplitude (mean ± SEM)	frequency (%)	amplitude (mean ± SEM)	frequency (%)	amplitude (mean ± SEM)	frequency (%)		normal vs DN	normal vs IgAN	DN vs IgAN	DN/normal	DN/IgAN
9–10	152.2 ± 38.1	20/31 (64.5)	446.7 ± 245.5	3/6 (50.0)	534.8 ± 245.5	3/4 (75.0)	<i>p</i> = 0.024	<i>p</i> = 0.048	NS	NS	2.9	0.8
12–15	1.7 ± 1.6	2/31 (6.5)	364.7 ± 177.7	4/6 (66.7)	49.3 ± 48.6	2/4 (50.0)	<i>p</i> < 0.001	<i>p</i> < 0.001	NS	<i>p</i> = 0.011	209.3	7.4
18–20	9.9 ± 7.0	3/31 (9.7)	215.3 ± 139.8	4/6 (66.7)	24.0 ± 24.0	1/4 (25.0)	<i>p</i> = 0.004	<i>p</i> = 0.003	NS	NS	21.7	9.0
22–24	2.1 ± 1.6	4/31 (12.9) ^a	622.8 ± 366.5	5/6 (83.3) ^a	969.3 ± 868.0	3/4 (75.0)	<i>p</i> = 0.003	<i>p</i> = 0.048	<i>p</i> = 0.010	NS	301.7	0.6
27–28	0.0 ± 0.0	0/31 (0.0)	1240.0 ± 547.9	6/6 (100.0)	320.3 ± 200.5	3/4 (75.0)	<i>p</i> < 0.001	<i>p</i> < 0.001	NS	<i>p</i> = 0.019	NA	3.9
34–35	0.3 ± 0.3	1/31 (3.2) ^a	900.8 ± 439.5	6/6 (100.0) ^a	212.3 ± 126.8	2/4 (50.0)	<i>p</i> < 0.001	<i>p</i> < 0.001	NS	<i>p</i> = 0.028	3102.7	4.2
61–68	12.1 ± 4.9	22/31 (71.0)	12526.3 ± 4716.0	6/6 (100.0)	6971.8 ± 1902.3	4/4 (100.0)	<i>p</i> < 0.001	<i>p</i> < 0.001	<i>p</i> = 0.012	NS	1035.5	1.8
86–90	4.0 ± 3.5	5/31 (16.1) ^a	3709.8 ± 2354.5	6/6 (100.0) ^a	768.5 ± 274.4	4/4 (100.0)	<i>p</i> = 0.001	<i>p</i> = 0.001	NS	NS	935.0	4.8
170–175	0.0 ± 0.0	1/31 (3.2)	49.7 ± 33.4	3/6 (50.0)	7.8 ± 3.5	3/4 (75.0)	<i>p</i> = 0.003	<i>p</i> = 0.002	NS	NS	1539.7	6.4

band	normal	DN	IgAN	ANOVA	Tukey's post-hoc tests of band count			ratio of band count	
	band count (mean ± SEM)	band count (mean ± SEM)	band count (mean ± SEM)		normal vs DN	normal vs IgAN	DN vs IgAN	DN/normal	DN/IgAN
all bands	3.8 ± 0.3	11.3 ± 1.2	10.8 ± 1.2	<i>p</i> < 0.001	<i>p</i> < 0.001	<i>p</i> < 0.001	NS	3.0	1.1

^a Spectra with significant difference in their frequencies between groups.

tudes or frequencies of individual bands/spectra although they tended to differ (Table 1). Not only the reproducibility between runs (shown in Figure 3) but also the variability among different wells on chips were of concern in our present study. We used amplitudes of lower marker (1.2 kDa), system spectra (6–8 kDa), and upper marker (260 kDa) in all 31 samples of normal healthy individuals to calculate for such variability. CVs of amplitudes of lower marker, system spectra, and upper marker were 22.84, 24.92, and 32.65%, respectively (Table 2).

We then compared the normal urinary proteome profile to the urinary proteome profiles of DN (*n* = 6) and IgAN (*n* = 4). Figure 5A and B shows the band patterns of DN and IgAN, respectively, whereas Figure 5C and D demonstrates their respective spectral patterns. Obviously, the band and spectral patterns of diseased urine (DN and IgAN) (Figure 5) were distinguishable from the patterns in normal urine (Figure 4). There were 8–16 (11.3 ± 1.2) and 9–14 (10.8 ± 1.2) bands/spectra observed in urine samples of DN and IgAN, respectively. Considering the amplitudes of these individual bands/spectra, ANOVA with Tukey post-hoc multiple comparisons revealed 9 spectra (with molecular masses of 9–10, 12–15, 18–20, 22–24, 27–28, 34–35, 61–68, 86–90, and 170–175 kDa) of which amplitudes significantly differed between normal and diseased urine (particularly normal versus DN; DN/normal amplitude ratios ranged from 2.9 to 3102.7) (Table 3). Moreover, the results also showed that the amplitudes of 3 spectra (with molecular masses of 12–15, 27–28, and 34–35 kDa) were significantly different between DN and IgAN urine (DN/IgAN

amplitude ratios ranged from 3.9 to 7.4). It should be also noted that serum creatinine levels and urine protein/creatinine ratios in DN patients were somewhat different from those in IgAN group. One should concern that our preliminary data need to be confirmed in a validation set of patients' urine samples using a larger number of patients with various degrees of proteinuria and impaired renal function.

In addition to the spectral amplitudes, we also considered differences in frequencies of their presence. The difference in frequency of any protein was arbitrarily determined when the presence of one protein was greater than 80%, whereas its presence was less than 20% in the other group. Frequencies of 3 spectra (with molecular masses of 22–24, 34–35, and 86–90 kDa) could differentiate the normal from the diseased urine, but could not distinguish between DN and IgAN urine (Table 3).

Our data underscore the potential use of microfluidic technology on a chip in clinical diagnostics using untreated urine samples. The amount or volume of urine required for the analysis was only 4 μL. Additionally, tricky preceding procedures of sample preparation to concentrate urinary proteins and to remove salts or other interfering substances were not required. Moreover, the nature of robustness to analyze several samples (up to 10) within a relatively short period (30 min) and the ease of use as well as automation make the clinical applicability of this system feasible. The data interpretation may be less biased compared to gel-based techniques, which require sample preparation that may selec-

tively recover some of protein components, whereas other proteins are lost (e.g., precipitation favors recovery of hydrophilic proteins but, on the other hand, is problematic for highly hydrophobic proteins; thus, the interpretation of protein spots/bands on 2-D/1-D gel relies mainly on hydrophilic proteins and the use of relative abundance compared to the total protein may be erroneous as the occupancy or proportion of hydrophobic proteins is ignored).

The principles of this LabChip microfluidics-based system are somehow different from another chip-based technology so-called "ProteinChip" or SELDI-TOF MS (surface-enhanced laser desorption/ionization-time-of-flight mass spectrometry), which has been previously applied to urinary proteome profiling.^{22–25} Whereas the LabChip microfluidic system is based on electrophoretic separation, fluorescence detection, and quantification of all proteins in the protein mixture or biofluid, SELDI-TOF MS combines MALDI-TOF MS with surface retentate chromatography. For SELDI, a biological sample is applied onto a chip surface carrying a functional group (e.g., normal phase, hydrophobic, cation- or anion-exchange). After incubation, unbound proteins are removed and the bound proteins are analyzed by a TOF mass spectrometer. The detection and quantitation of proteins by SELDI-TOF MS is based on their concentration, binding capacity to the chromatographic surface, and ionization process within the mass spectrometer. This approach reduces the complexity of proteins in the sample being analyzed by selecting only a group of particular functionality or property. Unlike the LabChip microfluidic system, which analyzes all proteins present in the sample, SELDI-TOF MS analyzes only a subset of proteins retained on the chip.

Even with the feasibility in clinical applications, as demonstrated by our convincing results, technical concerns and limitations of the microfluidic chip technology need to be pointed out. First, bubbles during loading of the sample, gel, and gel-stain solutions can be problematic as remaining macro- or micro-bubbles will definitely generate artificial spectra, which may be misinterpreted as proteins' spectra. Second, any dust or other particulate contaminants during preparation and loading of the chip can interfere in analysis, with the same reason as for bubbles. Third, this system is not free of chemical interference. Some chemicals, reagents, detergents, and reducing agents can also affect the results. Additionally, high salt concentrations may reduce the sensitivity of the protein detection. Fourth, this microfluidic separation system seemed to detect only major abundant proteins in the urine, indicating an insufficient sensitivity for analyzing untreated urinary proteome to discover all possible biomarkers, which may be low-abundance components. Enrichment of low-abundance proteins in the urine may be required prior to analysis using this system. Finally, this microfluidic chip technology can provide only the profile or pattern of proteins in the urine, other body fluids, or other biological samples. It is incapable of proceeding to identification or further analysis of the protein/peptide of interest. In the case that the identity of the indicative spectrum (i.e., the potential biomarker) needs to be identified, additional nonchip-based experiments must be performed to isolate and identify the protein/peptide with the same molecular mass of the indicative spectrum. After the identification or sequencing, the validation (e.g., using immunological methods) of its quantitative data (amplitude or abundance) and frequency of its presence is crucially required.

Abbreviations: 2-DE, two-dimensional electrophoresis; CV, coefficient of variation; DN, diabetic nephropathy; IgAN, IgA

nephropathy; SELDI-TOF, surface-enhanced laser desorption/ionization-time-of-flight.

Acknowledgment. We thank Duangporn Chuawattana for her technical assistance. This study was supported by Mahidol University, the National Research Council of Thailand, and the National Center for Genetic Engineering and Biotechnology (BIOTEC) (Project #BT-B-02-MG-B4-4908) to V.T.

References

- Colantonio, D. A.; Chan, D. W. The clinical application of proteomics. *Clin. Chim. Acta* **2005**, *357*, 151–158.
- Hanash, S. HUPO initiatives relevant to clinical proteomics. *Mol. Cell Proteomics* **2004**, *3*, 298–301.
- Thongboonkerd, V.; Malasit, P. Renal and urinary proteomics: Current applications and challenges. *Proteomics* **2005**, *5*, 1033–1042.
- Thongboonkerd, V. Proteomics in Nephrology: Current Status and Future Directions. *Am. J. Nephrol.* **2004**, *24*, 360–378.
- Gorg, A.; Weiss, W.; Dunn, M. J. Current two-dimensional electrophoresis technology for proteomics. *Proteomics* **2004**, *4*, 3665–3685.
- Thongboonkerd, V.; Klein, E.; Klein, J. B. Sample preparation for 2-D proteomic analysis. *Contrib. Nephrol.* **2004**, *141*, 11–24.
- Rabilloud, T. Solubilization of proteins in 2-D electrophoresis. An outline. *Methods Mol. Biol.* **1999**, *112*, 9–19.
- Lopez, M. F.; Berggren, K.; Chernokalskaya, E.; Lazarev, A.; Robinson, M.; Patton, W. F. A comparison of silver stain and SYPRO Ruby Protein Gel Stain with respect to protein detection in two-dimensional gels and identification by peptide mass profiling. *Electrophoresis* **2000**, *21*, 3673–3683.
- Klein, E.; Klein, J. B.; Thongboonkerd, V. Two-dimensional gel electrophoresis: a fundamental tool for expression proteomics studies. *Contrib. Nephrol.* **2004**, *141*, 25–39.
- Tudos, A. J.; Besselink, G. J.; Schasfoort, R. B. Trends in miniaturized total analysis systems for point-of-care testing in clinical chemistry. *Lab Chip* **2001**, *1*, 83–95.
- Verpoorte, E. Microfluidic chips for clinical and forensic analysis. *Electrophoresis* **2002**, *23*, 677–712.
- Whitesides, G. M. The origins and the future of microfluidics. *Nature* **2006**, *442*, 368–373.
- Lion, N.; Reymond, F.; Girault, H. H.; Rossier, J. S. Why the move to microfluidics for protein analysis? *Curr. Opin. Biotechnol.* **2004**, *15*, 31–37.
- Srinivasan, V.; Pamula, V. K.; Fair, R. B. An integrated digital microfluidic lab-on-a-chip for clinical diagnostics on human physiological fluids. *Lab Chip* **2004**, *4*, 310–315.
- Minas, G.; Wolffenbuttel, R. F.; Correia, J. H. A lab-on-a-chip for spectrophotometric analysis of biological fluids. *Lab Chip* **2005**, *5*, 1303–1309.
- Dittrich, P. S.; Manz, A. Lab-on-a-chip: microfluidics in drug discovery. *Nat. Rev. Drug Discovery* **2006**, *5*, 210–218.
- Yang, Y.; Kameoka, J.; Wachs, T.; Henion, J. D.; Craighead, H. G. Quantitative mass spectrometric determination of methylphenidate concentration in urine using an electrospray ionization source integrated with a polymer microchip. *Anal. Chem.* **2004**, *76*, 2568–2574.
- Liu, B. F.; Ozaki, M.; Hisamoto, H.; Luo, Q.; Utsumi, Y.; Hattori, T.; Terabe, S. Microfluidic chip toward cellular ATP and ATP-conjugated metabolic analysis with bioluminescence detection. *Anal. Chem.* **2005**, *77*, 573–578.
- Llorent-Martinez, E. J.; Ortega-Barrales, P.; Molina-Diaz, A. Multicommuted flow-through fluorescence optosensor for determination of furosemide and triamterene. *Anal. Bioanal. Chem.* **2005**, *383*, 797–803.
- Miyaguchi, H.; Tokeshi, M.; Kikutani, Y.; Hibara, A.; Inoue, H.; Kitamori, T. Microchip-based liquid-liquid extraction for gas-chromatography analysis of amphetamine-type stimulants in urine. *J. Chromatogr., A* **2006**, *1129*, 105–110.
- Chan, O. T.; Herold, D. A. Chip Electrophoresis as a Method for Quantifying Total Microalbuminuria. *Clin. Chem.* **2006**.
- Wright, G. L., Jr. SELDI proteinchip MS: a platform for biomarker discovery and cancer diagnosis. *Expert Rev. Mol. Diagn.* **2002**, *2*, 549–563.

- (23) Fung, E.; Diamond, D.; Simonsesn, A. H.; Weinberger, S. R. The use of SELDI ProteinChip array technology in renal disease research. *Methods Mol. Med.* **2003**, *86*, 295–312.
- (24) Tang, N.; Tornatore, P.; Weinberger, S. R. Current developments in SELDI affinity technology. *Mass Spectrom. Rev.* **2004**, *23*, 34–44.
- (25) Seibert, V.; Wiesner, A.; Buschmann, T.; Meuer, J. Surface-enhanced laser desorption ionization time-of-flight mass spectrometry (SELDI TOF-MS) and ProteinChip technology in proteomics research. *Pathol., Res. Pract.* **2004**, *200*, 83–94.

PR060586+

Use of immunoblotting as an alternative method for serogrouping *Leptospira*

Galayanee Doungchawee,¹ Worachart Sirawaraporn,²
Albert Icksang-Ko,^{3,4} Suraphol Kongtim,¹ Pimjai Naigowit⁵
and Visith Thongboonkerd⁶

Correspondence

Galayanee Doungchawee
scgdu@mahidol.ac.th

¹Department of Pathobiology, Faculty of Science, Mahidol University, Bangkok 10400, Thailand

²Department of Biochemistry, Faculty of Science, Mahidol University, Bangkok 10400, Thailand

³Gonçalo Moniz Research Center, Oswaldo Cruz Foundation/Brazilian Ministry of Health, Rua Waldemar Falcão, 12140295-001 Salvador, Bahia, Brazil

⁴Division of International Medicine and Infectious Disease, Weill Medical College of Cornell University, New York, NY 10021, USA

⁵Research Center for Leptospira Laboratory, National Institute of Health, Nonthaburi, Thailand

⁶Medical Molecular Biology Unit, Office for Research and Development, Department of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand

Leptospirosis is a worldwide zoonotic disease caused by a spirochaete bacterium, *Leptospira*. Serological detection of this micro-organism basically relies on a conventional microscopic agglutination test (MAT), which has some limitations and disadvantages. In the present study, immunoblotting has been applied as an alternative method for differentiating serogroups and serovars of leptospires. Leptospiral whole-cell lysates from a total of 26 serovars were subjected to immunoblotting using rabbit antisera against individual serovars. The findings clearly demonstrated that the pattern of immunoreactive bands could be used to differentiate between leptospires of different serogroups, consistent with MAT results. There was a multi-band pattern that was unique for the pathogenic *Leptospira* antigens and was not observed in the non-pathogenic *Leptospira biflexa* and non-leptospiral bacteria (i.e. *Escherichia coli*, *Burkholderia pseudomallei* and *Helicobacter pylori*). For pathogenic *Leptospira* species, a prominent smear-like band at approximately 19–30 kDa was present when the antigens were probed with the homologous antisera. The molecular size of the prominent band, although it showed a cross-reaction between members within the same serogroup, differed among different serovars. The results obtained from polyclonal antibodies (antisera) were confirmed using mAb. With its simplicity and safety of experimental procedures, it is proposed that immunoblotting may potentially be useful as an alternative method for differentiating between serogroups of leptospires.

Received 30 December 2006

Accepted 5 January 2007

INTRODUCTION

Leptospirosis is one of the most important zoonotic diseases worldwide, and is caused by pathogenic *Leptospira* species, which are not readily distinguishable from saprophytic leptospire strains on the basis of morphology and biochemical characteristics. The disease varies from subclinical infection to a severe illness with multi-organ involvement (Bharti *et al.*, 2003), which makes the diagnosis difficult – sometimes it is diagnosed as other febrile illnesses. Therefore, confirmation of the diagnosis made by

specific microbiological tests is necessary (Levett, 2001). Initially, the genus *Leptospira* was divided into two species, *Leptospira interrogans*, comprising all pathogenic strains, and *Leptospira biflexa*, containing saprophytic strains isolated from the environment (Johnson & Faine, 1984). Both *L. interrogans* and *L. biflexa* are divided into numerous serovars defined by agglutination after cross-absorption with homologous antigen (Dikken & Kmety, 1978; Johnson & Faine, 1984; Kmety & Dikken, 1993). Currently, the phenotypic classification of leptospires has been replaced by a genomic one, in which genomospecies, including *L. interrogans sensu stricto* and *L. biflexa sensu stricto*, do not correspond to the previous two species (*L. interrogans*

Abbreviation: MAT, microscopic agglutination test.

and *L. biflexa*), and indeed, pathogenic and non-pathogenic serovars occur within the same species (Levett, 2001). However, the molecular classification is incompatible with the system of serogroups that has been familiar to clinicians and epidemiologists for a long time. Most clinical laboratories find it necessary to retain serological classification of pathogenic leptospires for epidemiological purposes and for clinical diagnosis, of which microscopic agglutination test (MAT) is the most widely used and meets the requirements (Ahmad *et al.*, 2005). But, the performance of the MAT has been associated with some disadvantages. Here, the immunoblotting of whole-cell bacteria is attractive in comparison with the reference standard MAT, because it is simple, inexpensive, less burden and suitable for laboratory diagnosis.

METHODS

Bacterial culture. Twenty six leptospiral serovars, representing twenty serogroups (Table 1), were obtained from the National Leptospirosis Reference Center, National Institute of Health (NIH), Thailand and maintained by weekly subculture at 28–30 °C in liquid Difco *Leptospira* medium base EMJH (Becton Dickinson).

Burkholderia pseudomallei (K96243; kindly provided by the NIH, Thailand), *Escherichia coli* (DH5 α ; kindly provided by Department of Microbiology, Faculty of Medicine Siriraj Hospital, Mahidol University, Thailand) and *Helicobacter pylori* (kindly provided by Dr Anuchai Niwetpathomwat, Department of Veterinary Medicine, Faculty of Veterinary Science, Chulalongkorn University, Thailand) were used as the non-leptospiral antigens (negative controls).

Rabbit antisera (polyclonal antibodies) and mAbs against *Leptospira*. Each New Zealand white rabbit (8–10 weeks old) was immunized with an individual serovar of live leptospires by weekly intravenous injection for 4–6 weeks, as described elsewhere (Doungchawee *et al.*, 2005; Sitprija *et al.*, 1980). The serovar-specific antisera were then obtained and tested for MAT titre. All experimental procedures with animals were approved by the Animal Research Committee of the National Laboratory Animal Center, Thailand. For mAb, purified murine IgG specific to Bratislava and Bataviae serovars were kindly provided by Dr Pattama Ekpo, Department of Microbiology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand.

MAT. MAT was performed according to a modified method (Adler & Faine, 1978). Briefly, 50 μ l each antiserum was incubated at room temperature with an equal volume of a suspension of live leptospires (approx. 1×10^8 cells ml $^{-1}$) in separate wells of microtitre plates. After 2 h incubation, agglutination in each well was examined under a

Table 1. Representative leptospires among pathogenic and non-pathogenic serogroups, serovars, strains and species used

Serogroup	Serovar	Strain	Species
Australis	Australis	Ballico*	<i>Leptospira interrogans</i>
Australis	Bangkok	Bangkok D 92*	<i>Leptospira interrogans</i>
Australis	Bratislava	Jez Bratislava*	<i>Leptospira interrogans</i>
Autumnalis	Autumnalis	Akiyami A	<i>Leptospira interrogans</i>
Autumnalis	Rachmati	Rachmat	<i>Leptospira interrogans</i>
Ballum	Ballum	Mus 127*	<i>Leptospira borgpetersenii</i>
Bataviae	Bataviae	Swart	<i>Leptospira interrogans</i>
Canicola	Canicola	Hond Utrech IV*	<i>Leptospira interrogans</i>
Cellidoni	Cellidoni	Celledoni	<i>Leptospira weilii</i>
Djasiman	Djasiman	Djasiman	<i>Leptospira interrogans</i>
Grippotyphosa	Grippotyphosa	Moskva V	<i>Leptospira kirscheneri</i>
Hebdomadis	Hebdomadis	Hebdomadis	<i>Leptospira interrogans</i>
Icterohaemorrhagiae	Copenhageni	M 20	<i>Leptospira interrogans</i>
Icterohaemorrhagiae	Icterohaemorrhagiae	RGA	<i>Leptospira interrogans</i>
Javanica	Javanica	Veldrat Batavia 46*	<i>Leptospira borgpetersenii</i>
Louisiana	Saigon	L 79	<i>Leptospira interrogans</i>
Panama	Panama	CZ 214 K	<i>Leptospira noguchii</i>
Pomona	Pomona	Pomona*	<i>Leptospira interrogans</i>
Pyrogenes	Pyrogenes	Salinem*	<i>Leptospira interrogans</i>
Pyrogenes	Zanoni	Zanoni*	<i>Leptospira interrogans</i>
Ranarum	Ranarum	ICF	<i>Leptospira meyeri</i> †
Sarmin	Sarmin	Sarmin	<i>Leptospira weilii</i>
Sejroe	Hardjo	Hardjoprajitno	<i>Leptospira interrogans</i>
Sejroe	Sejroe	M84	<i>Leptospira borgpetersenii</i>
Tarassovi	Tarassovi	Perepelicin	<i>Leptospira borgpetersenii</i>
Semarang	Patoc	Patoc I*	<i>Leptospira biflexa</i>

*Species that were used for immunization.

†This species has been classified as a pathogenic species according to the International Committee on Systemic Bacteriology, Subcommittee on Taxonomy of *Leptospira* (Gravekamp *et al.*, 1993).

dark-field microscope (Olympus DP70 BX51; Shinjuku). The test was considered positive when >50 % agglutination was observed, and the most diluted titre with positivity was reported.

SDS-PAGE and immunoblotting. *Leptospira* were harvested at the mid-exponential phase and approximately 2×10^7 cells were used for each strain. The bacteria were washed with PBS three times for 5 min each, and then lysed with a standard Laemmli buffer ($1 \times$) and heated in boiling water for 5 min. After removal of the remaining particulate matter using microcentrifugation, the supernatant was loaded onto a 12.5 % acrylamide gel. SDS-PAGE was performed in a Hoefer Mighty Small II mini-gel apparatus (Amersham Biosciences) using a constant voltage of 200 V for 1 h (Kelson *et al.*, 1988). After completion, the resolved antigens were transferred onto a 0.45 μ m thick polyvinylidene fluoride (PVDF) membrane using a semi-dry system (TE70; Amersham Biosciences) with a constant current density of 1.5 mA cm^{-2} for 60 min.

The blotted membrane was washed three times (5 min each) with PBST (PBS with 0.05 %, v/v, Tween 20), and then incubated with 1 : 1000 rabbit antisera or 1 : 300 murine mAbs diluted in 2 % skimmed milk in PBST. The membrane was then washed as above and subsequently incubated, for 45 min, with horseradish peroxidase-conjugated goat anti-rabbit immunoglobulins (1 : 2000; Dakopatt) or sheep anti-mouse IgG (1 : 2500; Dakopatt), respectively. Immunoreactive bands were visualized using 3,3'-diaminobenzidine (Sigma-Aldrich) as the substrate. Molecular masses of the immunoreactive bands were estimated using Amersham Bioscience standard protein markers.

RESULTS AND DISCUSSION

This study was designed to identify immunogenic *Leptospira* protein antigens, valuable in serology, with a high test specificity. The reactivities of rabbit antisera raised against individual serovars were assessed in both MAT and immunoblot assays (Table 2). The sera reacted strongly to the corresponding leptospiral strains, as demonstrated in serovars Pomona, Pyrogenes and Canicola individually by a major smear-like banding (Fig. 1a, b, c; lanes 8, 9 and 6, respectively). Such smear-like banding was not observed when heterologous antisera were used. However, there was a cross-reaction producing the smear-like band when

heterologous antisera against members within the same serogroup were used (as shown in Fig. 1b for the cross-reactivity in Zanoni, when anti-Pyrogenes antiserum was used – Zanoni and Pyrogenes serovars are in the same serogroup, namely Pyrogenes). Similar findings were also observed for serovars Australis, Bangkok and Bratislava of the Australis serogroup (data not shown). This cross-reactivity among members within the same serogroup suggested that the serovar-specific epitope(s) might be similar within the same serogroup and that these antigenic determinants could be responsible for the agglutination when the MAT typing method was employed. Although there was a cross-reaction among the serovars within the same serogroup, molecular masses of such prominent smear-like bands were distinguishable among different serovars. Estimated molecular masses of the smear-like bands (within the range of 19–30 kDa), which were specific for individual serovars, are shown in Table 2. Obviously, the immunoblot pattern obtained with the antisera to pathogenic strains was different from that of non-pathogenic *L. biflexa* (Fig. 1d), which shared a characteristic pattern of multiple reactive bands, ranging from 10 to 90 kDa. This multi-band pattern, which was characteristic for the pathogenic leptospiral strains, was not observed when the non-pathogenic *L. biflexa* (Patoc) antigen was probed with homologous or heterologous antiserum (Fig. 1, lane 11).

Additional testing was carried out with antisera against serovars other than Pomona, Pyrogenes and Canicola, the lower molecular mass protein components at 14–20 and the flagella proteins of 35–36 kDa were found to be antigenically unique to *Leptospira*. Fig. 2 illustrates that immunoblotting using serovar-specific mAbs also provided the same prominent smear-like band, consistent with the results obtained from polyclonal antibodies (antisera). Evidence from a number of studies has suggested that the serovar-specific and/or serogroup-specific antigens might be outer-membrane glycolipids and lipopolysaccharides of

Table 2. Approximate molecular size of the smear-like band and MAT results of 11 representative antisera

Antiserum to serovar (strain)	Serogroup	Prominent smear-like band (kDa)	Corresponding MAT level
Australis (Ballico)	Australis	19–22	1 : 25 600
Bangkok (Bangkok D92)	Australis	19–20	1 : 12 800
Bratislava (Jez Bratislava)	Australis	24–30	1 : 6 400
Bataviae (Swart)	Bataviae	22–30	1 : 51 200
Ballum (Mus 127)	Ballum	21–30	1 : 6 400
Canicola (Hond Utrecht IV)	Canicola	21–28	1 : 25 600
Javanica (Veldrat Batavia 46)	Javanica	20–21	1 : 12 800
Pomona (Pomona)	Pomona	20–21, 22–24, 25–28	1 : 6 400
Pyrogenes (Salinem)	Pyrogenes	22–28	1 : 1 600
Zanoni (Zanoni)	Pyrogenes	19–21	1 : 3 200
Patoc (Patoc 1)	Semarang	17–18	1 : 25 600

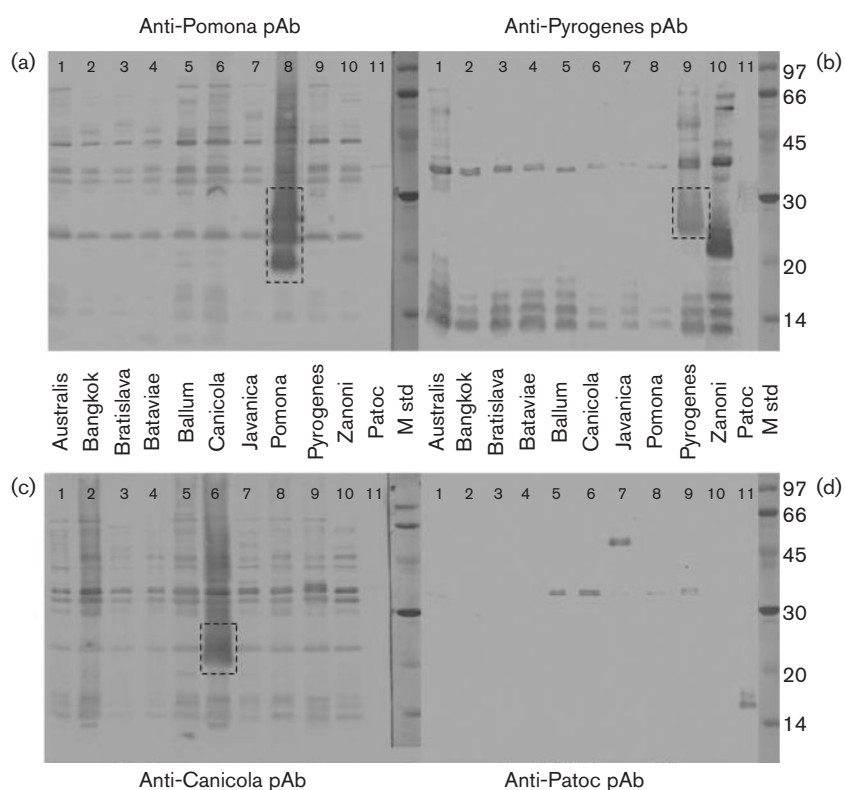


Fig. 1. Immunoblotting of antigens derived from whole-cell lysates of 11 leptospiral serovars. The membranes were probed with rabbit antisera (pAb) raised against individual serovars. This figure shows representative immunoblots of 11 serovars reacted with anti-Pomona (a), anti-Pyrogenes (b), anti-Canicola (c) and anti-Patoc (d) pAb. The multi-band pattern (~10–90 kDa) was unique for the pathogenic *Leptospira* antigens (lanes 1–10) and was not observed in the non-pathogenic *L. biflexa* (Patoc; lane 11). For pathogenic *Leptospira* species, a prominent smear-like band at approximately 19–30 kDa was present when the antigens were probed with the homologous antisera. Proteins of the prominent band, although showing a cross-reaction between members within the same serogroup, had molecular sizes that differed among different serovars. M std, molecular mass standard marker.

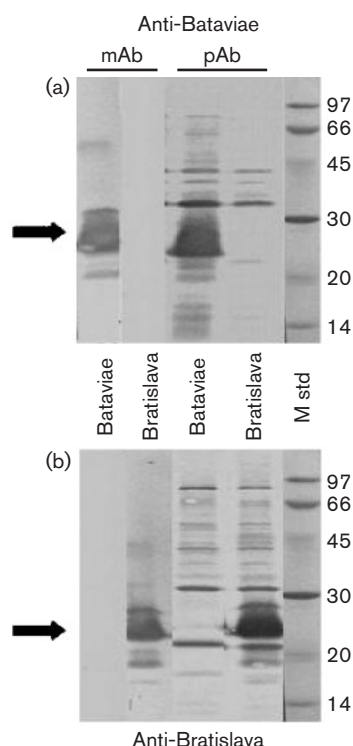


Fig. 2. Consistent results obtained from polyclonal antibodies (antisera) and mAbs. The prominent smear-like band (with a molecular mass of approximately 19–30 kDa) was also detectable when homologous mAb were used. M std, molecular mass standard marker.

L. interrogans (Barnett *et al.*, 1999; Brown *et al.*, 1991; Cho *et al.*, 1992; Shinagawa & Yanagawa, 1972). Glycolipid antigens have been suggested to play a major role in immunity, and to contribute to the production of agglutinating and opsonic antibodies (Adachi & Yanagawa, 1977; Farrelly *et al.*, 1987; Jost *et al.*, 1986; Masuzawa *et al.*, 1990; Midwinter *et al.*, 1994). Although, several published reports have described various methods for assessing size variation of the lipopolysaccharide antigens of *L. interrogans* (Cho *et al.*, 1992; Gitton *et al.*, 1992; Masuzawa *et al.*, 1990; Zuerner *et al.*, 1991), the utility of these tests remains controversial, such as the variable degree of serovar specificity of 21–31 kDa antigens determined by SDS-PAGE and ELISA (Cho *et al.*, 1992), the detection of 21–26 kDa as serovar-specific or serogroup-specific antigens among seven leptospiral strains by immunoblotting (Gitton *et al.*, 1992), and the identification of 23–30 kDa antigens of *L. interrogans* serovar Canicola with silver stain on SDS-PAGE (Masuzawa *et al.*, 1990).

Fig. 3(a) shows that levels of the immunoreactivity were varied by the different amounts (10^4 , 10^5 and 10^6 cells per assay) of bacterial antigens used for blotting. We suggest using at least 10^5 cells per assay to ensure the high quality of results using our method. Fig. 3(b) shows the results on non-leptospiral bacteria, i.e. *E. coli*, *B. pseudomallei* and *H. pylori*, which demonstrated minimal banding compared to that of leptospiral origin, which is indicative of the specificity of our technique for detecting *Leptospira*. Our immunoblot data show that the characteristics of some of the immunoreactive bands were similar, whereas the others

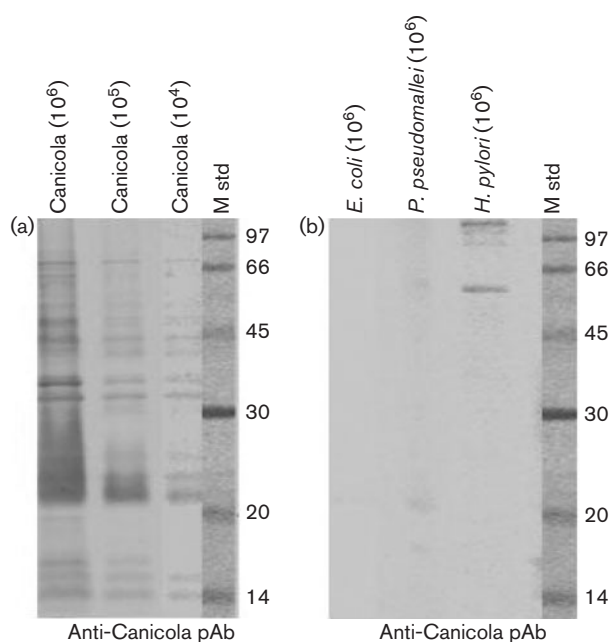


Fig. 3. Feasibility of the assay. (a) Various amounts of *Leptospira* serovar Canicola (10^4 , 10^5 and 10^6 cells per lane) were used for immunoblotting with anti-Canicola pAb (antisera). The degree of immunoreactivity was varied by the amount of antigens used. (b) Negative results were obtained when antigens derived from whole-cell lysates of *E. coli*, *P. pseudomallei* and *H. pylori* were used to react with anti-Canicola pAb. M std, molecular mass standard marker.

differed from published data obtained using whole-cell and/or outer-envelope protein fractions (Brown *et al.*, 1991; Cho *et al.*, 1992; Gitton *et al.*, 1992). It is likely that several *Leptospira* protein antigens have been characterized as minor bands on SDS-PAGE and these were clearly recognized when antiserum to a homologous strain was applied (Gitton *et al.*, 1992; Zuerner *et al.*, 1991). In addition, *Leptospira* species-associated antigens have been recognized and characterized, including flagellar components (35 or 33–36 kDa bands) (Chapman *et al.*, 1988; Kelson *et al.*, 1988), outer-membrane proteins and carbohydrate components (14.4–26.5 kDa bands) (Chapman *et al.*, 1988), outer-membrane-associated antigens (defined as LipL32, LipL36, LipL41 and LipL48) of leptospiral strains (Cullen *et al.*, 2002), a novel 48 kDa outer-membrane lipoprotein (designated LipL48) (Haake & Matsunaga, 2002), and two non-agglutinating protein antigens (p12 and p20), which are conserved for the genus *Leptospira* (Doherty *et al.*, 1989).

In our study, several antigens were predominantly detected in pathogenic leptospires (shown as multiple immunoreactive bands, ranging from 10 to 90 kDa when heterologous antisera were used) that could be used as the markers to discriminate from the non-pathogenic *L. biflexa*. Immunoblotting allows for the analysis of the

immune response to a number of defined antigens and has confirmed that the concept of serovar specificity of *Leptospira* species is confined to the 19–30 kDa epitopes. Characterizations of serovar-specific antigens, i.e. using MS, would be very interesting and deserves further studies. Extending the study to other reference strains and to human isolates may lead to further use of this test in epidemiological survey and/or in clinical diagnosis of leptospirosis as well.

ACKNOWLEDGEMENTS

This study was supported by a grant from the Department of Communicable Disease Control, Ministry of Public Health, Thailand.

REFERENCES

- Adachi, Y. & Yanagawa, R. (1977). Inhibition of leptospiral agglutination by the type specific main antigens of leptospires. *Infect Immun* 17, 466–467.
- Adler, B. & Faine, S. (1978). The antibodies involved in the human immune response to leptospiral infection. *J Med Microbiol* 11, 387–400.
- Ahmad, S. N., Shah, S. & Ahmad, F. M. (2005). Laboratory diagnosis of leptospirosis. *J Postgrad Med* 51, 195–200.
- Barnett, J. K., Barnett, D., Bolin, C. A., Summers, T. A., Wagar, E. A., Chevillat, N. F., Hartskeerl, R. A. & Haake, D. A. (1999). Expression and distribution of leptospiral outer membrane components during renal infection of hamsters. *Infect Immun* 67, 853–861.
- Bharti, A. R., Nally, J. E., Ricaldi, J. N., Matthias, M. A., Diaz, M. M., Lovett, M. A., Levett, P. N., Gilman, R. H., Willig, M. R. & other authors (2003). Leptospirosis: a zoonotic disease of global importance. *Lancet Infect Dis* 3, 757–771.
- Brown, J. A., LeFebvre, R. B. & Pan, M. J. (1991). Protein and antigen profiles of prevalent serovars of *Leptospira interrogans*. *Infect Immun* 59, 1772–1777.
- Chapman, A. J., Adler, B. & Faine, S. (1988). Antigens recognised by the human immune response to infection with *Leptospira interrogans* serovar Hardjo. *J Med Microbiol* 25, 269–278.
- Cho, S. N., Uhm, J. R. & Kim, J. D. (1992). Comparative analysis of lipopolysaccharide and lipid antigens of *Leptospira interrogans* serovars. *Yonsei Med J* 33, 24–31.
- Cullen, P. A., Cordwell, S. J., Bulach, D. M., Haake, D. A. & Adler, B. (2002). Global analysis of outer membrane proteins from *Leptospira interrogans* serovar Lai. *Infect Immun* 70, 2311–2318.
- Dikken, H. & Kmety, E. (1978). Serological typing methods of leptospires. In *Methods in Microbiology*, vol. 11. pp. 259–307. Edited by T. Bergan & J. R. Norris. London: Academic Press.
- Doherty, J. P., Adler, B., Rood, J. I., Billington, S. J. & Faine, S. (1989). Expression of two conserved leptospiral antigens in *Escherichia coli*. *J Med Microbiol* 28, 143–149.
- Doungchawee, G., Phulsuksombat, D., Naigowit, P., Khoaprasert, Y., Sangjun, N., Kongtim, S. & Smythe, L. (2005). Survey of leptospirosis of small mammals in Thailand. *Southeast Asian J Trop Med Public Health* 36, 1516–1522.
- Farrelly, H. E., Adler, B. & Faine, S. (1987). Opsonic monoclonal antibodies against lipopolysaccharide antigens of *Leptospira interrogans* serovar Hardjo. *J Med Microbiol* 23, 1–7v.

- Gitton, X., Andre-Fontaine, G., Andre, F. & Ganiere, J. P. (1992). Immunoblotting study of the antigenic relationships among eight serogroups of *Leptospira*. *Vet Microbiol* **32**, 293–303.
- Gravekamp, C., van de Kemp, H., Franzen, M., Carrington, D., Schoone, G. J., van Eys, G. J. J. M., Everard, C. O. R., Hartskeerl, R. A. & Terpstra, W. J. (1993). Detection of seven species of pathogenic leptospires by PCR using two sets of primers. *J Gen Microbiol* **139**, 1691–1700.
- Haake, D. A. & Matsunaga, J. (2002). Characterization of the leptospiral outer membrane and description of three novel leptospiral membrane proteins. *Infect Immun* **70**, 4936–4945.
- Johnson, R. C. & Faine, S. (1984). *Leptospira*. In *Bergey's Manual of Systematic Bacteriology*, vol. 1, pp. 62–67. Edited by N. R. Krieg & J. G. Holt. Baltimore, MD: Williams & Wilkins.
- Jost, B. H., Adler, B., Vinh, T. & Faine, S. (1986). A monoclonal antibody reacting with a determinant on leptospiral lipopolysaccharide protects guinea pigs against leptospirosis. *J Med Microbiol* **22**, 269–275.
- Kelson, J. S., Adler, B., Chapman, A. J. & Faine, S. (1988). Identification of leptospiral flagellar antigens by gel electrophoresis and immunoblotting. *J Med Microbiol* **26**, 47–53.
- Kmety, E. & Dikken, H. (1993). *Classification of the Species Leptospira interrogans and History of its Serovars*. Groningen: University Press Groningen.
- Levett, P. N. (2001). Leptospirosis. *Clin Microbiol Rev* **14**, 296–326.
- Masuzawa, T., Matsumoto, T., Nakamura, R., Suzuki, R., Shimizu, T. & Yanagihara, Y. (1990). Protective activity of glycolipid antigen against infection by *Leptospira interrogans* serovar Canicola. *J Gen Microbiol* **136**, 327–330.
- Midwinter, A., Vinh, T., Faine, S. & Adler, B. (1994). Characterization of an antigenic oligosaccharide from *Leptospira interrogans* serovar Pomona and its role in immunity. *Infect Immun* **62**, 5477–5482.
- Shinagawa, M. & Yanagawa, R. (1972). Isolation and characterization of a leptospiral type specific antigen. *Infect Immun* **5**, 12–19.
- Sitprija, V., Pipatanagul, V., Mertowidjojo, K., Boonpucknavig, V. & Boonpucknavig, S. (1980). Pathogenesis of renal disease in leptospirosis: clinical and experimental studies. *Kidney Int* **17**, 827–836.
- Zuercher, R. L., Knudtson, W., Bolin, C. A. & Trueba, G. (1991). Characterization of outer membrane and secreted proteins of *Leptospira interrogans* serovar Pomona. *Microb Pathog* **10**, 311–322.

Humidity control as a strategy for lattice optimization applied to crystals of HLA-A*1101 complexed with variant peptides from dengue virus

Pojchong Chotiyarnwong, Guillaume B. Stewart-Jones, Michael J. Tarry, Wanwisa Dejnirattisai, Christian Siebold, Michael Koch, David I. Stuart, Karl Harlos, Prida Malasit, Gavin Screaton, Juthathip Mongkolsapaya and E. Yvonne Jones

Copyright © International Union of Crystallography

Author(s) of this paper may load this reprint on their own web site provided that this cover page is retained. Republication of this article or its storage in electronic databases or the like is not permitted without prior permission in writing from the IUCr.

Pojchong Chotiyarnwong,^{a,b}
Guillaume B. Stewart-Jones,^c
Michael J. Tarry,^{c,‡} Wanwisa
Dejnirattisai,^{a,b} Christian
Siebold,^c Michael Koch,^c
David I. Stuart,^c Karl Harlos,^c
Prida Malasit,^{b,d} Gavin
Screaton,^a Juthathip
Mongkolsapaya^{a,b,*} and
E. Yvonne Jones^{c,*}

^aDepartment of Immunology, Division of
Medicine, Hammersmith Hospital, Imperial
College, London, England, ^bMedical Molecular
Biology Unit, Faculty of Medicine, Siriraj
Hospital, Mahidol University, Thailand,

^cDivision of Structural Biology and Oxford
Protein Production Facility (OPPF), The Henry
Wellcome Building for Genomic Medicine,
Roosevelt Drive, Headington, Oxford OX3 7BN,
England, and ^dMedical Biotechnology Unit,
National Center for Genetic Engineering and
Biotechnology (BIOTEC), National Science and
Technology Development Agency,
Pathumthani, Bangkok, Thailand

‡ Current address: Department of Biochemistry,
University of Oxford, South Parks Road,
Oxford OX1 3QU, England.

Correspondence e-mail:
j.mongkolsapaya@imperial.ac.uk,
yvonne@strubi.ox.ac.uk

Received 20 February 2007
Accepted 21 March 2007



© 2007 International Union of Crystallography
All rights reserved

Humidity control as a strategy for lattice optimization applied to crystals of HLA-A*1101 complexed with variant peptides from dengue virus

T-cell recognition of the antigenic peptides presented by MHC class I molecules normally triggers protective immune responses, but can result in immune enhancement of disease. Cross-reactive T-cell responses may underlie immunopathology in dengue haemorrhagic fever. To analyze these effects at the molecular level, the functional MHC class I molecule HLA-A*1101 was crystallized bound to six naturally occurring peptide variants from the dengue virus NS3 protein. The crystals contained high levels of solvent and required optimization of the cryoprotectant and dehydration protocols for each complex to yield well ordered diffraction, a process that was facilitated by the use of a free-mounting system.

1. Introduction

Dengue is an emerging viral disease that is now responsible for frequent epidemics in a large number of tropical and subtropical countries (Gubler, 1998). Dengue virus is an arthropod-borne flavivirus that can cause haemorrhagic fever (World Health Organization, 2002). There are four serotypes of the virus and immunity to one serotype does not provide protection against the others (World Health Organization, 2002). Indeed, there is evidence that immune enhancement of disease may be a consequence of cross-reactive T-cell responses (Mongkolsapaya *et al.*, 2003, 2006). The four dengue virus serotypes differ by up to 30% in amino-acid sequence and are almost as divergent as dengue virus is from other members of the flavivirus family. Since immunity to one serotype does not confer protection to other serotypes, serial infections are common and these can lead to life-threatening haemorrhagic fever. Immune enhancement is implicated since severe cases occur in individuals who have previously encountered the virus (Guzman *et al.*, 2000).

The T-cell response in dengue has been characterized in detail and we have described responses to an epitope (termed GTS in accordance with the single-letter code for the first three residues of the epitope) in nonstructural protein 3 (NS3), which has six different sequence variants among the dengue-virus serotypes (Mongkolsapaya *et al.*, 2003). This epitope is presented in the form of a 10-mer peptide complexed by the MHC class I molecule HLA-A*1101 (HLA-A11). A variety of T-cell responses can be made as a result of T-cell receptor recognition of these peptide–MHC class I (pMHC) complexes; some T cells are highly specific to a single variant, whilst other T cells can be highly promiscuous. In order to understand the molecular basis of this specificity and cross-reactivity, we set out to solve the structure of HLA-A11 in complex with each of the six peptides. The six pMHC complexes crystallized under closely related conditions which, in their absence of PEG-type precipitants, differed markedly from those typically reported for pMHC complexes in general and peptide–HLA-A11 complexes in particular. Bragg diffraction to sufficient resolution for structure determination of the full set of complexes was only obtained after extensive optimization of the cryoprotectant and dehydration procedures. Possibly as a result

Table 1
Dengue-peptide variants.

Variant	Dengue serotype	Sequence
GTS1.1	Den 1	GTSGSPIVNR
GTS2.1	Den 2	GTSGSPIIDK
GTS2.2	Den 2	GTSGSPIVDR
GTS2.3	Den 2	GTSGSPIVDK
GTS2.4	Den 2	GTSGSPIADK
GTS3.1	Den 3 or Den 4	GTSGSPIINR

of their high solvent content (~70%), these crystals therefore provide a clear example of the utility of humidity-controlled free-mounting techniques.

2. Materials and methods

2.1. Protein expression and purification

Soluble pMHC samples of all six variant peptides (termed GTS1.1, GTS2.1, GTS2.2, GTS2.3, GTS2.4 and GTS3.1 as defined in Table 1) in complex with HLA-A*1101 (HLA-A11) were produced using previously described methods (Reid *et al.*, 1996). Briefly, the extracellular region of the HLA-A11 heavy chain (residues 1–275; with and without the BirA-recognizing sequence) and β_2 microglobulin (β_2 m) were expressed separately in *Escherichia coli* strain BL21 using the pET23d+ and pGMT7 expression systems, respectively, and isopropyl β -D-thiogalactoside (Melford, UK) as an inducer. The peptides were synthesized in-house on a 396 MPS automated peptide synthesizer (Advanced ChemTech) and their purity was characterized by reverse-phase HPLC (Gilson).

Inclusion bodies of HLA-A11 heavy chain and β_2 m were washed (in 0.5% Triton X-100, 50 mM Tris pH 8.0, 100 mM NaCl, 0.1% sodium azide and 2 mM DTT), solubilized in urea (8 M urea, 0.1 M NaH₂PO₄, 0.01 M Tris pH 8.0, 0.1 M EDTA and 0.1 mM DTT) and refolded in the presence of peptide at a molar ratio of 2.5:10 heavy chain: β_2 m:peptide at 277 K for 48 h using a refolding buffer (100 mM Tris pH 8.0, 400 mM L-arginine, 2 mM EDTA pH 8.0, 5 mM reduced glutathione, 0.5 mM oxidized glutathione and 100 μ M PMSF protease inhibitor). pMHC complexes with the BirA-recognizing sequence were biotinylated using the BirA enzyme (Avidity, Denver, CO, USA) following the company protocol. Biotinylated and nonbiotinylated samples were purified by FPLC size-exclusion gel filtration on a Sephadex 75 26/60 column (Pharmacia, UK) using 20 mM Tris pH 8.0 and 25 mM NaCl buffer, further purified on a 1 ml HiTrap Q HP anion-exchange column (Pharmacia, UK) and kept at 193 K until required. The nonbiotinylated pMHC complexes were buffer-exchanged to 7 mM Tris pH 7.0 and concentrated to 10–12 mg ml⁻¹ prior to crystallization. Phycoerythrin-labelled pMHC tetrameric complexes (PE-tetramers) were prepared by mixing biotinylated pMHC complexes with fluorochrome-conjugated streptavidin as described previously (Altman *et al.*, 1996).

2.2. Conformational and functional analysis of pMHC complexes

The refolded pMHCs were tested for correct folding by dot enzyme-linked immunoassay (DEIA) using the monoclonal antibody W6/32 (which recognizes a conformational epitope of a pMHC complex; Shields & Ribaud, 1998) and their functionality was verified using a competitive binding assay as described previously (Mongkolsapaya *et al.* 2003) with slight modification. GTS epitope-specific T-cell lines which recognize all six variants used in this study were washed (with PBS, 1% BSA, 0.5 mM EDTA) and then stained with PE-tetramer in the presence of a 100-fold excess of nonbio-

tinylated pMHC and incubated at 310 K for 30 min. Cells were washed and stained with fluorescein isothiocyanate (FITC) labelled anti-CD8 monoclonal antibody at 277 K for 20 min, washed again, fixed and analyzed by flow cytometry.

2.3. Protein crystallization

Initial screens to establish crystallization conditions used nanolitre-scale sitting drops (100 nl protein solution plus 100 nl reservoir solution) dispensed by a Cartesian Technologies Microsys MIC4000 (Genomic Solutions; Walter *et al.*, 2003, 2005). These experiments used CrystalQuick square-well plates (Greiner) and 90 μ l reservoir solution. A total of 672 concentrations were tested from a range of commercial kits: Hampton Crystal Screens 1 and 2, Emerald Wizard 1 and 2 (DeCode Genetics), Hampton PEG/Ion Screen, Hampton Grid Screen PEG 6000, Hampton Grid Screen Ammonium Sulfate, Hampton Matrix, Hampton Crystal Screen Cryo, Hampton Grid Screen PEG/LiCl, Hampton Grid Screen NaCl, Hampton Grid Screen MPD, Hampton Quik Screen (sodium/potassium phosphate) and Hampton Index Screen (Hampton Research). Promising conditions were repeated by hand in larger scale (1 μ l protein solution plus 1 μ l reservoir solution) sitting-drop screens [using microbridges (Harlos, 1992) in VDX plates (Hampton) containing 500 μ l of reservoir], varying the concentrations of salt, buffer, precipitant and pH. After a condition capable of yielding well ordered crystals had been found for one pMHC complexes (pMHC variant GTS1.1; 0.63 M NaH₂PO₄, 1.17 M K₂HPO₄ pH 8.2), optimal crystallization conditions for the other variants were established using standard sitting-drop screens (Harlos, 1992) to fine-tune the salt concentration and pH. Crystals of all the HLA-A11-peptide complexes grew at room temperature in 1.17 M K₂HPO₄ and 0.63 M NaH₂PO₄ pH 8.0–8.3.

2.4. Cryoprotection, crystal dehydration and data collection

Cryo methods had to be developed independently for each complex. The cryoprotectants screened included perfluoropolyether oil, sodium malonate and a range of concentrations of glycerol. For cryoprotectant screening, crystals were soaked briefly (1–5 min) in 2 μ l cryoprotectant solution, flash-cooled in a cold gas stream and maintained at 100 K. The resultant diffraction quality was assessed using an in-house microfocus X-ray generator (Rigaku, Micromax-007) with either MAR 345 or MAR dtb imaging-plate detectors.

Only one of the pMHC complexes, GTS1.1, yielded diffraction of sufficient quality for data collection from flash-cooled cryoprotected crystals without further manipulation. A second, GTS2.4, was cryoprotected by transfer into a solution of the crystal mother liquor plus 30% glycerol, flash-cooled directly in the cryostream at station BM 10.1 of the Synchrotron Radiation Source (SRS, Daresbury, UK) and, after initial diffraction characterization, annealed by blocking off the cryostream for 2 s. The remaining four complexes, GTS2.1, GTS2.2, GTS2.3 and GTS3.1, had to be mounted on a free-mounting system (FMS, Rigaku, The Woodlands, USA; Kiefersauer *et al.*, 2000) to allow fine control of humidity prior to flash-cooling. This device allows the crystal to be maintained at a chosen relative humidity at room temperature in a standard cryo-loop. A simple graphical user interface then allows the relative humidity to be manipulated *in situ*. The 'start' humidity for each crystal was defined as the humidity at which a drop of the corresponding mother liquor was maintained at constant size. Following dehydration analysis, crystals were mounted on the FMS either at the start humidity or at a defined humidity below 94.5%, allowed to equilibrate and then covered in 2.9 M sodium malonate cryosolution before being frozen in a cold gas

Table 2

Crystallographic data-collection statistics.

Values in parentheses are for the highest resolution shell.

Peptide	GTS2.1	GTS2.2	GTS2.3	GTS2.4	GTS1.1	GTS3.1
Crystallization pH	8.0	8.3	8.1	8.3	8.2	8.2
Cryoprotectant	2.9 M sodium malonate pH 7.0	2.9 M sodium malonate pH 7.0	2.9 M sodium malonate pH 7.0	30% glycerol with mother liquor	20% glycerol with mother liquor	2.9 M sodium malonate pH 7.0
Method of improving crystal diffraction	FMS†	FMS†	FMS†	Annealing	—	FMS†
Humidity at which crystal was frozen (%)	94.5	93.5	93.5	—	—	93.5
FMS† start humidity (%)	94.5	94.5	93.5	—	—	94.5
Solvent content (%)	70.4	67.9	70.2	69.9	70.1	67.9
Resolution range (Å)	30–2.80 (2.90–2.80)	20–2.60 (2.66–2.60)	25–2.40 (2.47–2.40)	20–3.20 (3.31–3.20)	30–2.60 (2.69–2.60)	20–2.80 (2.86–2.80)
No. of reflections collected	213283	206211	209829	117774	222499	146030
Unique reflections	50399 (5025)	58379 (3648)	74962 (6075)	32447 (1567)	62767 (5948)	46370 (2872)
Redundancy (%)	4.2	3.5	2.8	3.6	3.5	3.1
Completeness (%)	100.0 (100.0)	99.9 (99.9)	93.9 (92.0)	94.4 (45.7)	99.3 (94.3)	99.1 (98.1)
$R_{\text{merge}}^{\ddagger}$ (%)	9.9 (52.7)	10.7 (51.9)	7.5 (52.8)	14.1 (61.8)	9.7 (52.1)	10.5 (47.9)
$\langle I/\sigma(I) \rangle$	13.9 (2.6)	10.5 (2.4)	16.8 (1.7)	10.9 (1.8)	12.6 (1.8)	10.4 (2.3)
Space group	C2	C2	C2	C2	C2	C2
Unit-cell parameters (Å)						
<i>a</i> (Å)	158.5	151.8	157.7	145.4	157.1	152.2
<i>b</i> (Å)	145.7	143.6	146.3	143.6	147.1	143.9
<i>c</i> (Å)	105.0	102.6	104.6	113.5	104.1	101.7
β (°)	120.2	119.8	120.2	122.3	120.1	119.7
Unit-cell volume (10 ⁶ Å ³)	2.096	1.941	2.086	2.003	2.081	1.935
Source	ESRF ID14-EH2	SRS BM14.1	SRS BM14.1	SRS BM10.1	ESRF ID23-EH1	SRS BM14.1
Wavelength (Å)	0.933	1.488	1.488	1.381	0.979	1.488
Detector	ADSC Q4	ADSC Q4	ADSC Q4	MAR 225 CCD	MAR 225 CCD	ADSC Q4

† FMS, free-mounting system (Kiefersauer *et al.*, 2000). $\ddagger R_{\text{merge}} = \sum_h \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_h \sum_i I_i(hkl)$, where $I_i(hkl)$ is the i th measurement of reflection hkl and $\langle I(hkl) \rangle$ is the weighted mean of all measurements of reflection hkl .

stream and checked for diffraction. Crystals with improved diffraction were stored in liquid nitrogen for subsequent synchrotron data collection. For most crystals the best diffraction was obtained at 93.5% relative humidity, which was slightly below the usual start humidity (Table 2).

Diffraction data for the six complexes were recorded at the beamlines of the European Synchrotron Radiation Facility (ESRF, Grenoble, France) and the UK Synchrotron Radiation Source (SRS, Daresbury, UK), as detailed in Table 2. Diffraction data sets were autoindexed and integrated with *DENZO* and scaled using *SCALEPACK* (Otwinowski & Minor, 1997; <http://www.hkl-xray.com>) (Table 2).

3. Results and discussion

3.1. Protein production and functional assays

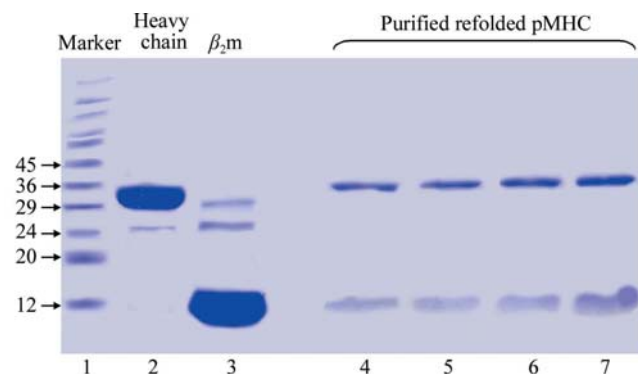
The yields of refolded complex after size-exclusion chromatography purification were 3–5 times higher for nonbiotinylated pMHCs compared with biotinylated pMHCs, indicating typical refolding efficiencies of 12 and 3%, respectively (data not shown), the former being in line with previously reported values (10–20% depending on peptide and MHC heavy chain) for pMHCs refolded for structural studies (Reid *et al.*, 1996). Gel electrophoresis indicated that the samples were over 98% pure (Fig. 1).

The conformation of the refolded pMHCs was verified by DEIA using monoclonal antibody W6/32 that recognizes a conformational epitope. Furthermore, the T-cell receptor (TCR) binding ability of the refolded pMHCs was tested by a competitive binding assay as shown in Fig. 2. The refolded pMHCs inhibited PE-tetramers binding to GTS-specific T-cell lines, resulting in reduction of the fluorescent intensity. These demonstrated that the refolded pMHCs retained the functionality of the physiological membrane-bound pMHCs since they had the ability to interact with GTS epitope-specific T-cell lines.

3.2. Crystallization

672 crystallization conditions were screened in a series of experiments using several batches of the GTS1.1 and GTS3.1 variant pMHCs. Crystallization occurred at room temperature in two major groups of screen conditions; namely, those containing ammonium sulfate or phosphate salts (Table 3). Initial efforts to collect suitable X-ray data focused on optimized crystals from the ammonium sulfate condition (which was the first of the two crystallization precipitants to be identified). These crystals typically diffracted to 7 Å resolution using both in-house and synchrotron sources. Extensive efforts to improve the diffraction quality (varying the cryoprotection and use of dehydration protocols) provided no consistent route forward. Attention was therefore shifted to the second group of crystallization precipitants to be identified: the phosphate salts.

A reservoir solution consisting of 0.63 M NaH₂PO₄ and 1.17 M K₂HPO₄ pH 8.2 was judged on the basis of visual inspection to be the

**Figure 1**

Production of pMHC. pMHCs purified by anion-exchange chromatography were run on 15% SDS-PAGE under reducing conditions. Standard molecular-weight markers (kDa) were run in lane 1, while inclusion bodies of MHC heavy chain and β_2m were in lane 2 and 3. Lanes 4–7 show purified GTS1.1, GTS2.2, GTS2.1 and GTS3.1, respectively.

Table 3

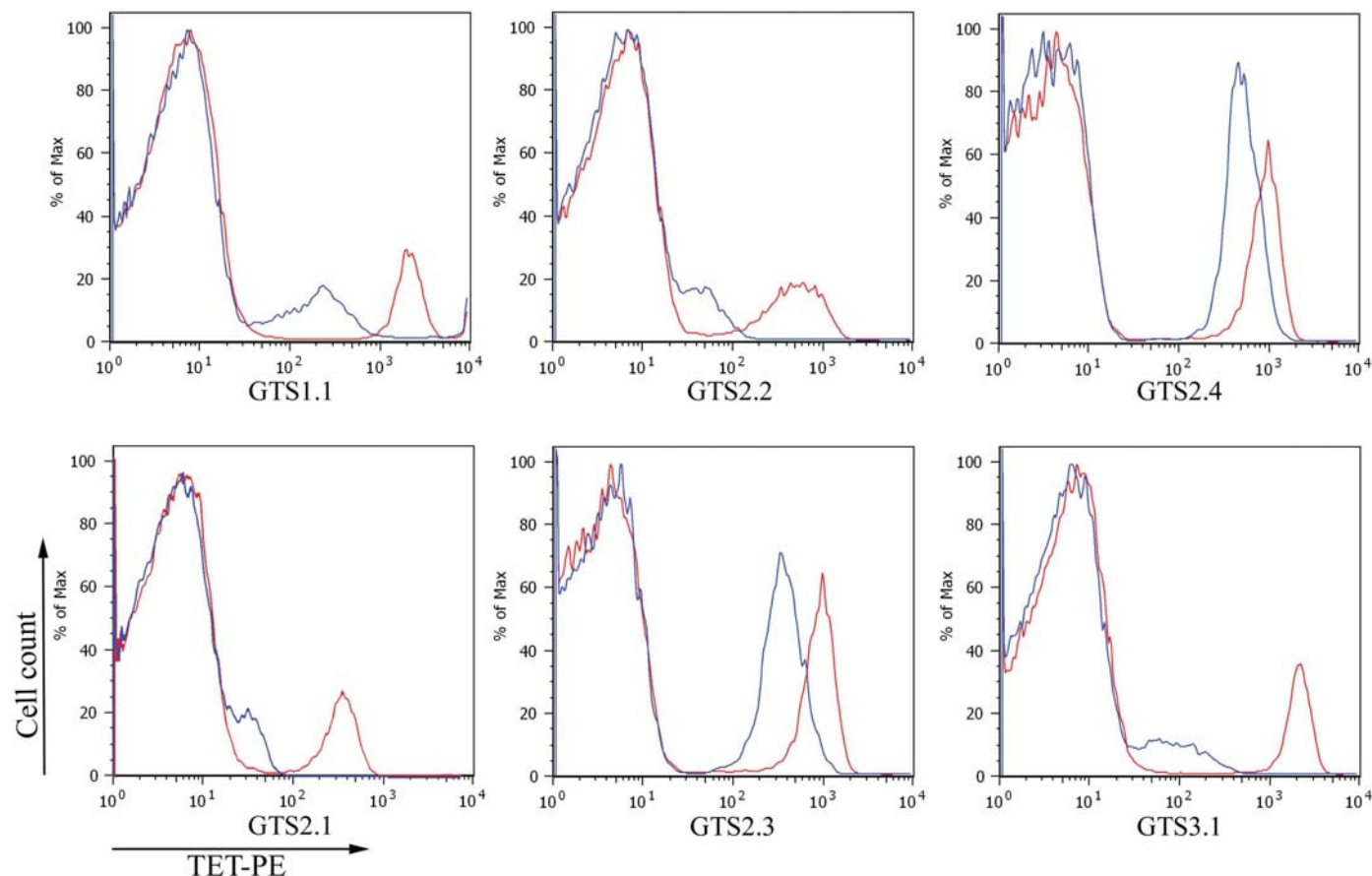
Crystallization conditions that produced pMHC crystals.

Conditions			
Salt	Buffer	Additive	Crystallization time
2 M (NH ₄) ₂ SO ₄	100 mM citrate pH 5.5	None	2 d
2 M (NH ₄) ₂ SO ₄ /200 mM NaCl	100 mM cacodylate pH 6.5	None	2 d
2 M (NH ₄) ₂ SO ₄ /200 mM Li ₂ SO ₄	100 mM Tris pH 7.0	None	2 d
2.4 M (NH ₄) ₂ SO ₄	1 M MES pH 6.0	None	1 d
1.26 M trisodium citrate	0.09 M Na HEPES pH 7.5	10% glycerol	13 d
1.17 M K ₂ HPO ₄ /0.63 M NaH ₂ PO ₄	None	None	1 d
1.53 M K ₂ HPO ₄ /0.27 M NaH ₂ PO ₄	None	None	1 h
1.73 M K ₂ HPO ₄ /0.072 M NaH ₂ PO ₄	None	None	1 d
None	1.8 M triammonium citrate pH 7.0	None	2 d

most promising of the phosphate conditions identified in the nano-litre-scale screens. This condition was therefore used as the basis for crystal optimization in standard microlitre-scale sitting drops. The GTS1.1 variant crystallized within 16 h using this condition and showed Bragg diffraction to better than 3.5 Å resolution when tested on the in-house rotating-anode X-ray source (after cryoprotection by transfer to mother liquor containing 20% glycerol).

Having established a potentially suitable condition for the growth of crystals of suitable quality for structure determination, microlitre-scale crystallization optimizations were carried out for all six pMHC variants. For each pMHC a series of conditions were tested: these conditions were centred on 0.63 M NaH₂PO₄, 1.17 M K₂HPO₄ pH 8.2, but varied in salt concentration and pH. Identical salt concentrations but slight differences in pH proved to be essential for the crystallization of the full set of pMHC variants (Table 2 and Fig. 3).

Crystallization conditions for the MHC class I molecule HLA-A11 have been reported previously by two groups (Li *et al.*, 2002; Li & Bouvier, 2004; Blicher *et al.*, 2005, 2006). Bouvier and coworkers reported data for HLA-A11 in complex with two unrelated peptide epitopes from HIV-1 proteins (AIFQSSMTK from reverse transcriptase and QVPLRPMTYK from Nef; Li & Bouvier, 2004), while Blicher and coworkers (Blicher *et al.*, 2005, 2006) detail crystals of two further complexes, those of HLA-A11 with a peptide (KTFPPTEPK) derived from the SARS nucleocapsid and with AIMPARYFPK, a sequence homologue of an epitope in hepatitis B virus DNA polymerase. The crystallization conditions in all these examples were PEG-based and were closely related (the crystallization precipitant was either PEG MME 5000 or PEG 4000 at a concentration of 20–30%). Indeed, PEG-based conditions are typical for the crystallization of pMHC complexes (Garboczi *et al.*, 1994;

**Figure 2**

Tetramer competition assay. The GTS-specific T-cell lines were stained with 1 µg PE-conjugated tetrameric pMHC in the presence (blue) or absence (red) of a 100-fold excess of the refolded pMHC and analyzed by FACS. The histograms show the fluorescent intensity of cells stained with the PE-conjugated tetrameric pMHC complexes.

Reid *et al.*, 1996). The nonstandard conditions required for the crystallization of our HLA-A11–peptide complexes indicate that a broad crystallization screen should be considered if initial attempts centred on PEG-based conditions fail to yield pMHC crystals. In particular, the 1.8 M phosphate conditions that yielded useful crystals came from the Hampton Quik Screen (block 5 in our standard set of 480 conditions and usually one of the less successful screens; Walter *et al.*, 2003, 2005).

3.3. Crystal lattice-order optimization and X-ray data collection

It was difficult to establish cryo methods for the collection of well ordered Bragg diffraction. Optimal diffraction data for each complex were obtained using different protocols (Table 2). Two cryoprotectants proved useful, glycerol and sodium malonate, but X-ray data to a suitable resolution for structure determination could only be collected directly from cryoprotected crystals for the GTS1.1 variant pMHC (2.6 Å; Table 2). The crystals of all five other complexes required annealing/fine control of humidity to optimize the diffraction quality. The results of a simple annealing experiment carried out at station BM 10.1 of the SRS (Daresbury, UK) by blocking off the cryostream for 2 s first alerted us to the significant potential for improvement in diffraction quality (Fig. 4). In this experiment, the resolution limit for Bragg diffraction from a crystal of the GTS2.4 variant complex was improved dramatically from ~ 7 to 3.2 Å. Subsequently, the diffraction quality for crystals of the four remaining complexes, GTS2.1, GTS2.2, GTS2.3 and GTS3.1, was systematically characterized at defined humidity levels using a combination of FMS (Kiefersauer *et al.*, 2000) and in-house X-ray source/detector. The FMS is still a relatively new device, the principle of which is explained in §2.4. Essentially, by maintaining the crystal in an air stream of defined humidity, the crystal can be dehydrated or rehydrated over time. The FMS goniometer head can be mounted onto standard X-ray equipment so that the diffraction at each humidity level can be recorded.

In this case, we performed full dehydration analyses on only a few crystals and then used fresh crystals to repeat the dehydration steps which resulted in the best diffraction. Crystals were then frozen

directly in a standard cold nitrogen-gas stream without testing their diffraction at room temperature and were only exposed to X-rays once they had been frozen (suitable crystals were then stored in liquid nitrogen for subsequent synchrotron data collection). This approach avoids any of the radiation damage which the crystals experience when they are exposed to X-rays at room temperature. Typically, out of ten crystals tested, four showed improved diffraction. This rather low success rate may partly be a consequence of the difficulty in reproducing the freezing protocol exactly (for example, in covering the crystal with the freezing solution and transferring it to the cold gas stream always within the same time interval). It is clear that even with the FMS there is scope for improving the crystal-handling procedures in order to achieve greater reproducibility for sensitive crystals such as these. From Table 2, it can be seen that for two of the complexes (GTS2.1 and GTS2.3) the optimal humidity level is the start humidity (*i.e.* that of the crystallization mother liquor); thus, the use of the FMS appears to have allowed us to circumvent residual problems with cryoprotection. For the other two complexes (GTS2.2 and GTS3.1), the optimal humidity is 1% below the start humidity, implying that a slight dehydration of the crystal promotes a beneficial lattice rearrangement.

As expected given their closely related crystallization conditions, the crystals of all six pMHC complexes had essentially identical lattices. The crystals belong to space group *C2* and contain three pMHC complexes in the crystallographic asymmetric unit plus 68–70% solvent (the Matthews coefficients are 3.83–4.16 Å³ Da^{−1}; further details including unit-cell parameters are given in Table 2). It is noteworthy that the percentage of solvent in the crystals does indeed correlate with the protocols applied before data collection. The three crystals subjected to annealing or controlled dehydration all show lower solvent content than those simply cryoprotected or flash-frozen at start humidity on the FMS. In all cases the solvent content is relatively high for a protein crystal, particularly when compared with the values of 48–58% for other peptide–HLA-A11 complexes (Li & Bouvier, 2004; Blicher *et al.*, 2005, 2006) and pMHC complex crystals in general (Garboczi *et al.*, 1994; Reid *et al.*, 1996). The diffraction data reported here have resulted in successful structure determinations (phased by molecular replacement) for all six

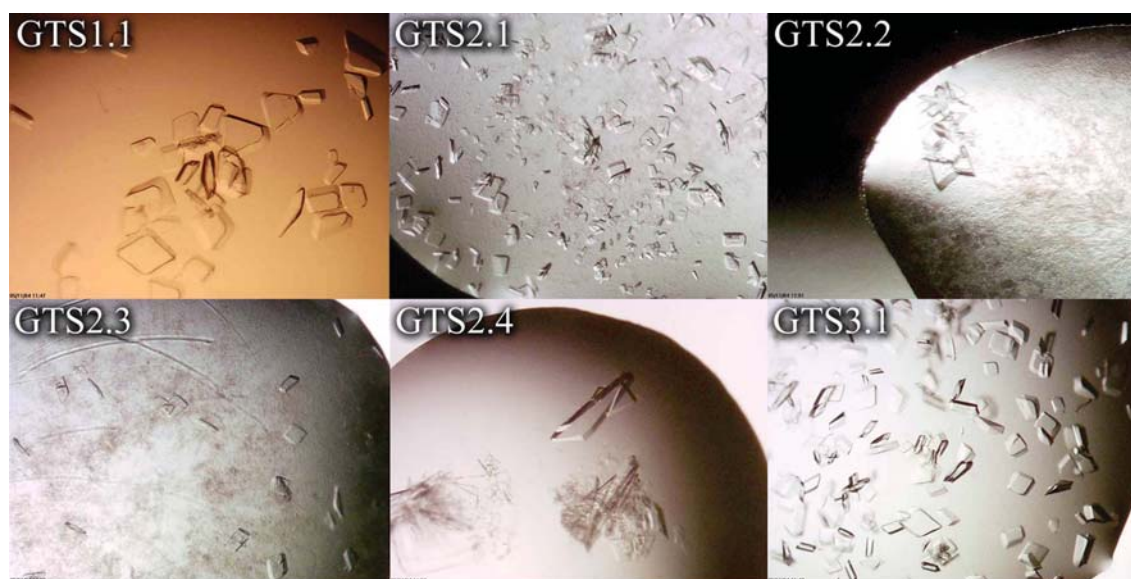


Figure 3

GTS pMHC crystals. Purified pMHCs were crystallized in 0.63 M NaH₂PO₄, 1.17 M K₂HPO₄ buffer using sitting-drop vapour diffusion. The buffer pH was varied for each variant as described in Table 2. All crystals grew at room temperature within 24 h.

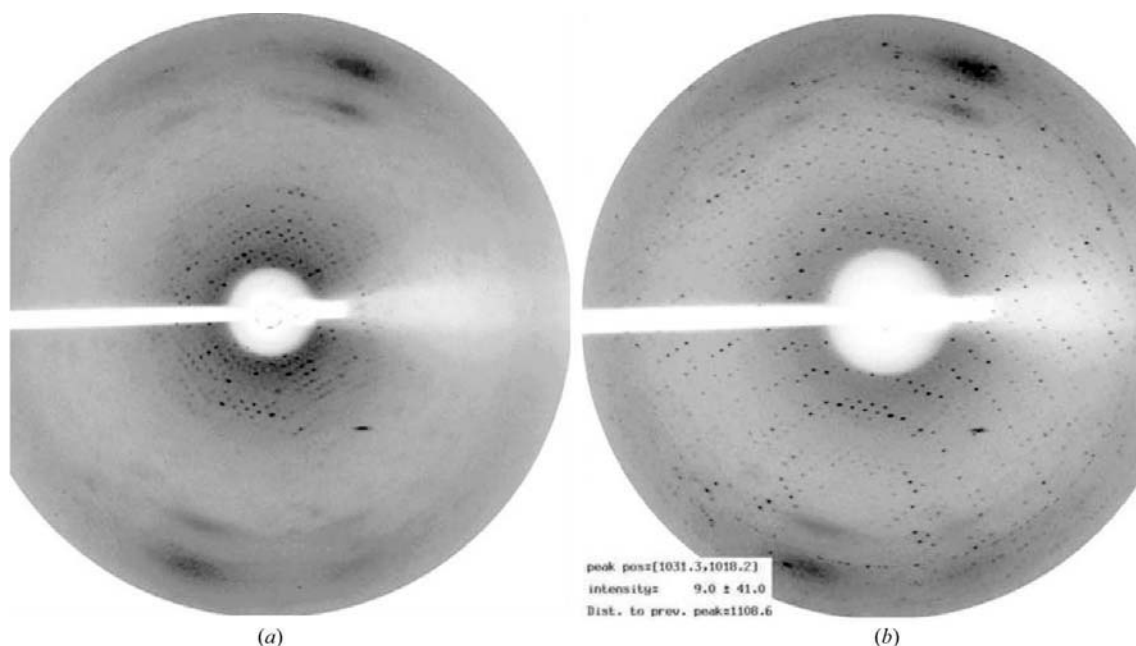


Figure 4

Improvement of X-ray diffraction after annealing experiment. The cooling flow from the cryostream to the GTS2.4 crystal was blocked off for 2 s between the measurement of the X-ray diffraction patterns shown in (a) and (b). The images are from a MAR CCD detector on beamline 10.1 of the SRS (Daresbury), with the crystal-to-detector distance set such that the resolution at the edge of the detector was 3.5 Å. The experimental conditions (exposure time, oscillation range etc.) are identical between images, with the exception that the backstop position has altered (the backstop was manually moved out of position to allow access to the crystal for annealing and then repositioned). For subsequent data collection, the detector position was altered to measure data to a somewhat higher resolution limit (3.2 Å, Table 2).

variant complexes. A full analysis of the crystal structures and their implications for the immune response to dengue virus is ongoing and will be reported elsewhere.

4. Conclusions

In our experience, as in many other laboratories, dehydration of protein crystals can provide essential improvements in diffraction quality (reviewed by Heras & Martin, 2005). For example, a reduction of crystal solvent content from 56 to 48% and a consequent increase in diffraction limit from 3.7 to 2.2 Å allowed the structure of HIV-1 reverse transcriptase to be refined at sufficient resolution to facilitate antiviral drug design (Esnouf *et al.*, 1998). Similarly significant improvements in resolution (from ~3.5 to 2.0 Å) were achieved at rather higher solvent content (63% reduced to 56%) for crystals of human semaphorin 4D (Esnouf *et al.*, 2006). In both these previous examples, the dehydration protocol was based on stepwise increases in the reservoir concentration of a PEG-based precipitant prior to crystal mounting. The development of humidity-controlling systems such as the FMS now provide an alternative route by which to manipulate crystal hydration levels (see, for example, Koch *et al.*, 2004). The very high level of solvent content in the dengue peptide–HLA-A11 complexes may well have contributed to the difficulties we encountered in optimizing crystal-preparation protocols for cryo X-ray data collection; however, for these examples a systematic investigation of the effects of humidity, facilitated by an FMS, was successful in bringing problematic crystal lattices into sufficient order to allow useful data collection.

We are grateful to T. S. Walter for help with the crystallization setups and to the staff of the ESRF and EMBL in Grenoble, France as well as the SRS in Daresbury, UK for assistance with X-ray data collection. We acknowledge use of the crystallization facilities

provided by the MRC-funded Oxford Protein Production Facility (OPPF) and The European Commission Integrated Programme (SPINE, grant code QL2-CT-2002-00988). EYJ is a Cancer Research UK Principal Research Fellow. PM is supported by Senior Research Scholar Program of the Thailand Research Fund (TRF). PC is supported by The Medical Scholar Program, Mahidol University and The Split Mode PhD program (BIOTEC, Thailand). This work was funded by the UK Medical Research Council and The Wellcome Trust.

References

- Altman, J. D., Moss, P. A., Goulder, P. J., Barouch, D. H., McHeyzer-Williams, M. G., Bell, J. I., McMichael, A. J. & Davis, M. M. (1996). *Science*, **274**, 94–96.
- Blicher, T., Kastrop, J. S., Buus, S. & Gajhede, M. (2005). *Acta Cryst.* **D61**, 1031–1040.
- Blicher, T., Kastrop, J. S., Pedersen, L. O., Buus, S. & Gajhede, M. (2006). *Acta Cryst.* **F62**, 1179–1184.
- Esnouf, R. M., Love, C. A., Harlos, K., Stuart, D. I. & Jones, E. Y. (2006). *Acta Cryst.* **D62**, 108–115.
- Esnouf, R. M., Ren, J., Garman, E. F., Somers, D. O'N., Ross, C. K., Jones, E. Y., Stammers, D. K. & Stuart, D. I. (1998). *Acta Cryst.* **D54**, 938–953.
- Garboczi, D. N., Madden, D. R. & Wiley, D. C. (1994). *J. Mol. Biol.* **239**, 581–587.
- Gubler, D. J. (1998). *Clin. Microbiol. Rev.* **11**, 480–496.
- Guzman, M. G., Kouri, G., Valdes, L., Bravo, J., Alvarez, M., Vazquez, S., Delgado, I. & Halstead, S. B. (2000). *Am. J. Epidemiol.* **152**, 793–799.
- Harlos, K. (1992). *J. Appl. Cryst.* **25**, 536–538.
- Heras, B. & Martin, J. L. (2005). *Acta Cryst.* **D61**, 1173–1180.
- Kiefersauer, R., Than, M. E., Dobbek, H., Gremer, L., Melero, M., Strobl, S., Dias, J. M., Soulimane, T. & Huber, R. (2000). *J. Appl. Cryst.* **33**, 1223–1230.
- Koch, M., Breithaupt, C., Kiefersauer, R., Freigang, J., Huber, R. & Messerschmidt, A. (2004). *EMBO J.* **23**, 1720–1728.
- Li, L. & Bouvier, M. (2004). *J. Immunol.* **172**, 6175–6184.
- Li, L., Promadej, N., McNicholl, J. M. & Bouvier, M. (2002). *Acta Cryst.* **D58**, 1195–1197.
- Mongkolsapaya, J., Dejnirattisai, W., Xu, X. N., Vasanawathana, S., Tangthawornchaikul, N., Chairunsri, A., Sawasdivorn, S., Duangchinda, T., Dong,

- T., Rowland-Jones, S., Yenchitsomanus, P. T., McMichael, A., Malasit, P. & Screaton, G. (2003). *Nature Med.* **9**, 921–927.
- Mongkolsapaya, J., Duangchinda, T., Dejnirattisai, W., Vasanawathana, S., Aviruthnan, P., Jairungsri, A., Khemnu, N., Tanghawornchaikul, N., Chotiyarnwong, P., Sae-Jang, K., Koch, M., Jones, Y., McMichael, A., Xu, X., Malasit, P. & Screaton, G. (2006). *J. Immunol.* **176**.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Reid, S. W., Smith, K. J., Jakobsen, B. K., O'Callaghan, C. A., Reyburn, H., Harlos, K., Stuart, D. I., McMichael, A. J., Bell, J. I. & Jones, E. Y. (1996). *FEBS Lett.* **383**, 119–123.
- Shields, M. J. & Ribaldo, R. K. (1998). *Tissue Antigens*, **51**, 567–570.
- Walter, T. S., Diprose, J., Brown, J., Pickford, M., Owens, R. J., Stuart, D. I. & Harlos, K. (2003). *J. Appl. Cryst.* **36**, 308–314.
- Walter, T. S., Diprose, J. M., Mayo, C. J., Siebold, C., Pickford, M. G., Carter, L., Sutton, G. C., Berrow, N. S., Brown, J., Berry, I. M., Stewart-Jones, G. B. E., Grimes, J. M., Stammers, D. K., Esnouf, R. M., Jones, E. Y., Owens, R. J., Stuart, D. I. & Harlos, K. (2005). *Acta Cryst. D* **61**, 651–657.
- World Health Organization (2002). *Dengue and Dengue Haemorrhagic Fever*. World Health Organization Fact Sheet 117. <http://www.who.int/mediacentre/factsheets/fs117/en/>.

Characterization of dengue virus NS1 stably expressed in 293T cell lines

Sansanee Noisakran^{a,b}, Thanyaporn Dechtawewat^b, Panthip Rinkaewkan^b,
Chunya Puttikhunt^{a,b}, Amornrat Kanjanahaluethai^c, Watchara Kasinrer^{a,d},
Nopporn Sittisombut^{a,c,*}, Prida Malasit^{a,b,**}

^a Medical Biotechnology Unit, National Center for Genetic Engineering and Biotechnology,
National Science and Technology Development Agency, Bangkok 10400, Thailand

^b Medical Molecular Biology Unit, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand

^c Department of Microbiology, Faculty of Medicine, Chiang Mai University, Chiang Mai 50200, Thailand

^d Department of Clinical Immunology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai 50200, Thailand

Received 16 August 2006; received in revised form 29 December 2006; accepted 15 January 2007

Available online 28 February 2007

Abstract

Dengue virus NS1 is a viral nonstructural protein detected in sera of infected individuals and in infected cells. Multiple NS1 structural forms have been reported but the functional characteristics of these forms remain unknown. In this study, a set of 293T cell lines stably expressing recombinant dengue NS1 without additional C-terminal sequence (rNS1s), with a heterologous transmembrane segment (rNS1tm), or with the 26-residue N-terminal portion of NS2A (rNS1v1) was established to aid in the characterization of different NS1 forms. Each NS1 protein form had distinct phenotypes and the following properties were documented: (1) dissipated expression in the cytoplasm, dimerization, and *N*-glycosylation were observed, regardless of the forms of NS1 expressed; (2) the rNS1v1 and rNS1tm forms, but not the rNS1s, were observed prominently on the surface membrane; (3) only the rNS1v1 form incorporated ethanolamine, a precursor of the glycosylphosphatidylinositol moiety, and was partially sensitive to digestion with phosphatidylinositol-specific phospholipase C. The stable 293T transfectants expressing multiple forms of dengue NS1 may be a useful model to investigate the function of NS1 and the mechanism by which NS1 associates with membrane.

© 2007 Elsevier B.V. All rights reserved.

Keywords: Dengue virus; NS1; Recombinant protein; Transfection; 293T cell

1. Introduction

Illnesses due to infection with dengue virus are one of the most important arthropod-borne human diseases. A spectrum of dengue diseases has been recognized, ranging from mild febrile illness, dengue fever, to more severe and potentially fatal dengue hemorrhagic fever (DHF) and dengue shock syndrome (DSS) (Halstead, 1997). There are four dengue serotypes and each one of them is capable of inducing dengue illness.

The genome of dengue virus is a single-stranded RNA encoding one polypeptide, which is cleaved into at least 10 known proteins (C, prM, E, NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5) by viral encoded or host proteases. Dengue NS1 is an approximately 45-kDa protein with two N-linked glycosylation sites. It exists in multiple forms (i.e., monomer, dimer, and hexamer) in different compartments of infected mosquito and mammalian cells (Flamand et al., 1992; Jacobs et al., 2000; Leblais and Young, 1995; Mason, 1989). Soluble NS1 can be detected in the culture media and acute sera from dengue virus-infected patients (Alcon et al., 2002; Libraty et al., 2002; Young et al., 2000).

NS1 is synthesized initially as a portion of the polypeptide with the C-terminal hydrophobic sequence of E serving as the signal peptide for its translocation into the lumen of endoplasmic reticulum (ER). Cleavages at the N-terminus and C-terminus by host signalase and an unknown ER-resident protease, respectively, generate the monomeric NS1 molecules.

* Corresponding author at: Department of Microbiology, Faculty of Medicine, Chiang Mai University, 110 Intawaroros Street, Chiang Mai 50200, Thailand. Tel.: +66 53 945334; fax: +66 53 217144.

** Corresponding author at: Medical Molecular Biology Unit, Adulyadejvirogrom Building (12th floor), Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand. Tel.: +66 2 418 4793; fax: +66 2 418 4793.

E-mail addresses: nsittiso@mail.med.cmu.ac.th (N. Sittisombut), sipml@mahidol.ac.th (P. Malasit).

The C-terminal octapeptide sequence of NS1 is required for efficient cleavage at the NS1-NS2A junction (Falgout et al., 1989; Falgout and Markoff, 1995; Hori and Lai, 1990). Within the ER lumen, NS1 monomers are modified by an addition of the high-mannose type glycans at both N-linked glycosylation sites, followed by rapid dimerization (Pryor and Wright, 1993; Winkler et al., 1988). Further modifications of the carbohydrate moieties in NS1 dimers occur in Golgi apparatus just prior to their transport to the cell surface and the release from infected cells. Extracellular NS1 molecules are detected as hexamers (Flamand et al., 1999; Winkler et al., 1989; Winkler et al., 1988).

The functional role of NS1 in dengue virus infection has not been clearly elucidated. Several lines of evidence suggest that NS1 is involved in viral RNA replication, assembly and maturation (Lindenbach and Rice, 1997, 1999; Mackenzie et al., 1996; Muylaert et al., 1996; Muylaert et al., 1997). NS1 may play a role in the pathogenesis of DHF/DSS (Jacobs et al., 2000), possibly by activating complement and forming circulating immune complexes (Avirutnan et al., 2006; Brandt et al., 1970; Ruangjirachuporn et al., 1979; Schlesinger et al., 1990). Antibodies to NS1 may cause the reduction of blood coagulation factors, endothelial cell damage and vascular leakage in DHF/DSS patients (Chang et al., 2002; Falconar, 1997). However, the contributions of the soluble versus the membrane-associated forms of NS1 to these observations remain unclear. The present study was undertaken to generate human 293T cell lines stably expressing NS1 in a membrane-anchorless form (rNS1s), a transmembrane form (rNS1tm), and a glycosylphosphatidylinositol (GPI)-linked form (rNS1v1) by using calcium phosphate-mediated transfection. The association with cell surface as well as pertinent post-translational modifications of these recombinant NS1 forms was then compared in representative highly expressing clones.

2. Materials and methods

2.1. Cell line, virus, and antibodies

A human embryonic kidney epithelial cell line, 293T, was cultured in RPMI 1640 (Invitrogen, Grand Island, NY, USA) supplemented with 10% fetal bovine serum (FBS, Hyclone, Logan, UT, USA), 37 $\mu\text{g ml}^{-1}$ of penicillin, 60 $\mu\text{g ml}^{-1}$ of streptomycin and 2 mM L-glutamine at 37 °C in the 5% CO₂, humidified atmosphere. Dengue serotype 2 virus strains 16681 and NGC were prepared as previously described (Avirutnan et al., 1998). For virus infection, the monolayer of 293T cells was incubated with dengue virus at the multiplicity of infection (MOI) of 1 at 37 °C for 3 h. The culture medium was then replaced with RPMI 1640-10% FBS and incubated for 2–3 days. Mouse monoclonal antibodies recognizing linear epitopes (1B2, NS1-1F.1, and NS1-3F.1) (Puttikhunt et al., 2003) or conformational epitopes (NS1-8.2 and 1A4) (Puttikhunt et al., 2003; Sittisiri, 1994) on dengue virus NS1 were employed in the forms of culture media or ascitic fluid. An anti-dengue E monoclonal antibody (clone 3H5; Henchal et al., 1982) was produced in our laboratory. Biotin-conjugated anti-human CD55 mono-

clonal antibody was purchased from Pharmingen, San Diego, CA, USA.

2.2. Construction of plasmid cDNA clones expressing recombinant dengue NS1

Six cDNA cassettes for the expression of recombinant NS1 forms were constructed using an expression vector, pcDNA3.1/Hygro (Invitrogen) (Fig. 1). Briefly, the genomic RNA was extracted from strain NGC-infected C6/36 mosquito cells and used as a template for reverse transcription (RT) and RT-PCR with the primer pairs listed in Table 1. The rNS1v1 cassette, encoding the whole NS1 connected to the 26-residue N-terminal segment of NS2A, was generated by amplifying the cDNA template with the primer pair A-B and cloning the PCR product into the Bgl II-Not I sites of pDisplay (Invitrogen). The Kpn I-Not I fragment containing a murine Ig κ leader sequence and a hemagglutinin epitope located upstream of the NS1-NS2A sequence was then excised and inserted into pcDNA3.1/Hygro. The rNS1tm cassette, encoding NS1 connected to the transmembrane domain of platelet derived growth factor receptor, was constructed by transferring the Kpn I-Not I fragment from the plasmid pDisplay/sNS1tm (Puttikhunt et al., 2003) into the Kpn I- and Not I-digested pcDNA3.1/Hygro. In the generation of the rNS1s cassettes, pDisplay/sNS1tm was employed as a template for PCR using the primer pair C-D (Table 1). The amplified product containing only the NS1 sequence was ligated into pDisplay at the Bgl II-Not I sites. The Kpn I-Not I fragment excised from this subclone was ligated into pcDNA3.1/Hygro in order to generate pcDNAhygro/soluble NS1 (NGC). A sim-

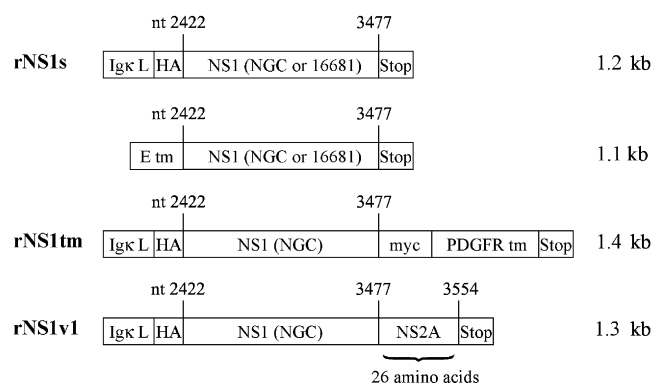


Fig. 1. A diagram illustrating the construction of plasmid cDNA clones encoding various forms of dengue virus NS1. cDNA of dengue type 2 virus, strains NGC and 16681, served as templates for the amplification of NS1 segment using different sets of oligonucleotide primers. In four constructs, the amplified products were inserted into pDisplay to generate the NS1 subclones as described in the Materials and methods. The NS1 fragments flanked by the coding sequences for the murine Ig κ leader sequence (Igk L), hemagglutinin epitope (HA), and stop codon were then excised from the subclones and ligated into pcDNA3.1/Hygro at the Not I and Kpn I sites to generate the following recombinant forms of NS1: (i) the NS1 sequence only, rNS1s (NGC or 16681); (ii) the NS1 sequence connected to a myc epitope and the transmembrane domain of platelet derived growth factor receptor (PDGFR), rNS1tm; (iii) the NS1 sequence connected to the 26-amino acid region at the N-terminus of NS2A, rNS1v1. In two constructs, the amplified products containing the C-terminal 28 residues sequence of E (E tm) and the whole NS1 sequence were ligated directly into pcDNA3.1/Hygro at the Not I and Kpn I sites.

Table 1
Oligonucleotide primers for cloning of recombinant NS1

Designation	Orientation	Sequence
A	Sense	5' CCGGCCAGATCT ^a GATAGTGGTTGCGTTGTGAGC 3'
B	Antisense	5' CAAGCTTCGCGGCCGC ^b GTAAAC ^c TIA ^d TCGGGTCCT GAGCATTCTCTCCAG 3'
C	Sense	5' CCGGCCAGATCT ^a GATAGTGGTTGCGTTGTGAGC 3'
D	Antisense	5' GATCGATCGCGGCCGC ^b TIA ^d GGCTGTGACCAAG GAGTTAAAC ^c CAAATCTCTCTTTCTC 3'
E	Sense	5' GCGGTACC ^e ACCATGAATTCACGCAGCACC 3'

^a Bgl II site.

^b Not I site.

^c Hpa I site (a bold letter represents a silent mutation introduced to generate Hpa I site).

^d Termination codon.

^e Kpn I site.

ilar construct, pcDNAhygro/soluble NS1 (16681) was derived in the same manner, but with a full-length cDNA clone of strain 16681 (Sriburi et al., 2001) serving as a template in the amplification reaction. Expression of these four NS1 constructs was based on the murine Igκ leader sequence derived from pDisplay. Two additional rNS1s cassettes, pcDNAhygro/Etm-NS1 (NGC) and pcDNAhygro/Etm-NS1 (16681), were constructed by including the coding sequence for 28 hydrophobic residues at the C-terminus of E, the native signal peptide for NS1 translocation. The cDNA of strain NGC and a full-length cDNA clone of strain 16681 were utilized as the templates in PCR employing the primer pair D–E (Table 1). The resultant PCR products were digested with Kpn I and Not I, and ligated directly into pcDNA3.1/Hygro.

Following cloning, all expected changes were verified by nucleotide sequence analysis of the pcDNAhygro-based plasmid clones. In addition, the rNS1tm cassette contained a silent mutation TTG → CTG at the base position 2677 (base position according to Irie et al., 1989) and a missense mutation TTG → TGG at the position 3455, resulting in the leucine-to-tryptophan substitution in the NS1 ectodomain just proximal to the transmembrane portion. The NS1 sequence in pcDNAhygro/soluble NS1 (NGC) contained only a silent mutation at the base position 2677. The missense mutation at the position 3455 was not incorporated into this clone as the reverse primer D (Table 1) used in the construction spanned the positions 3439–3477.

2.3. Transfection and selection of cells expressing recombinant NS1 forms

293T cell monolayer in a 35-mm² dish was transfected with pcDNAhygro-based NS1 expressing cassettes by using calcium phosphate-DNA precipitates, as follows. One hundred μl of 0.25 M CaCl₂ were mixed with 5 μg of plasmid followed by drop wise addition of an equal volume of 50 mM HEPES (pH 7.0), 274 mM NaCl and 1.5 mM Na₂HPO₄, and the 5-min incubation on ice. DNA-CaCl₂ mixture was layered onto the monolayer, incubated at 37 °C for 12 h, and then replaced with fresh RPMI 1640-10% FBS. Transfected cells were thereafter propagated in the presence of 100–400 μg ml⁻¹ of hygromycin B and observed daily for the appearance of colonies of surviving

cells. Individual colonies were selected and maintained in RPMI 1640-10% FBS medium containing 100 μg ml⁻¹ of hygromycin B. pcDNA3.1/Hygro-transfected 293T cells were also selected with hygromycin B for use as a negative control.

2.4. Determination of NS1 expression by indirect immunofluorescence assay

The expression of recombinant NS1 on transfected cell surface was examined by an indirect immunofluorescence assay (IFA). Transfected cells were detached with 0.5 mM EDTA in phosphate buffered saline (PBS) and washed once with PBS containing 1% bovine serum albumin and 10 mM NaN₃. The cell pellet (5 × 10⁵ cells) was resuspended with 50 μl of 10% human AB serum in washing buffer for 30 min at 4 °C to block non-specific binding. Thereafter, 50 μl of either anti-NS1 antibodies or isotype-matched control antibodies were added and the cells were incubated for 1 h at 4 °C followed by washing twice. Stained cells were further incubated with 25 μl of 8 μg ml⁻¹ FITC-conjugated rabbit anti-mouse Ig antibody (Dako, Glostrup, Denmark) for 30 min at 4 °C in the dark, washed twice, and resuspended in 400 μl of washing buffer.

In the detection of intracellular NS1, detached cells (1 × 10⁶) were fixed with 4% paraformaldehyde for 20 min, and permeabilized with 0.2% triton X-100 (Sigma, St. Louis, MO, USA) for 10 min at room temperature (RT). Cells were washed once with 0.1% triton X-100 in PBS and then stained with the same set of antibodies as described above. Enumeration of transfected cells with the cell surface or intracellular expression of NS1 was performed by using a flow cytometer (FACScan, Becton Dickinson, San Jose, CA, USA) with 10,000-gated events per sample. The pattern of NS1 expression was visualized with a fluorescence microscope (Axioplan2, Carl Zeiss, Germany) and a laser scanning confocal microscope (LSM 510 Meta, Zeiss, Germany).

2.5. Immunoblot analysis

Transfected 293T cells (1 × 10⁷) were lysed by incubating with 350 μl of 1% triton X-100, 150 mM NaCl, 50 mM Tris-HCl, pH 7.5, 1 mM EDTA and 1 mM PMSF for 20 min on ice, followed by centrifugation at 5500 × g for 5 min. The

clear lysate was mixed with non-reducing loading buffer (50 mM Tris-Cl pH 6.8, 2% SDS, 0.1% bromophenol blue, and 10% glycerol), boiled (95 °C, 5 min) and subjected to electrophoresis in 10% or 12% SDS-polyacrylamide gel. Proteins were transferred to 0.45- μ m nitrocellulose membrane (Schleicher & Schuell, Keene, NH, USA) by using a semi-dry electroblotter. The non-specific binding sites on the membrane were blocked with 5% skim milk in PBS for 1 h and then the membrane was incubated with an anti-NS1 antibody at 4 °C overnight. Following three washes with PBS, rabbit anti-mouse Ig antibody conjugated with horseradish peroxidase (HRP) (Dako) at the dilution of 1:1000 was added for 1 h at RT. The membrane was again washed three times and the immunoreactive proteins were visualized with 0.06% diethylaminobenzidine, 0.04% NiCl₂, and 0.01% H₂O₂. The size of protein bands was determined by comparing with the broad-range protein markers (Bio-Rad, Hercules, CA, USA).

2.6. Analysis of N-glycosylation modification of NS1

The transfected cell lysates or culture media were denatured by boiling for 10 min in a 20- μ l reaction containing 0.5% SDS and 1% β -mercaptoethanol and then subjected to digestion with endoglycosidase H (Endo H) or peptide: N-glycosidase F (PNGase F) (New England Biolabs, Beverly, MA, USA) at 37 °C for 18 h according to the manufacturer's protocol. Endo H removes the high-mannose glycans from glycoproteins whereas PNGase F cleaves all N-linked carbohydrate groups (Tarentino et al., 1985). Control sample contained the same amount of lysates in the digestion buffer but without enzyme. Following enzymatic digestion, samples were electrophoresed in 10% SDS-polyacrylamide gel, transferred to nitrocellulose membrane, and reacted with anti-NS1 antibody in an immunoblot format as described above.

2.7. Radiolabeling, triton X-114 phase separation and immunoprecipitation

NS1 transfectants and dengue virus-infected cells that had been cultured for 2 days were starved with cysteine- and methionine-free DMEM (Invitrogen) containing 10% dialyzed FBS (Hyclone) for 1 h and then metabolically labeled with [³⁵S]cysteine/methionine (Redivue cell labeling mix, Amersham Biosciences, Buckinghamshire, England) at 50 μ Ci ml⁻¹ for 16–18 h. For radiolabeling NS1 with [³H]ethanolamine, cells were incubated with 1.5 ml of DMEM-10% dialyzed FBS for 1 h and then [1-³H]ethan-1-ol-2-amine hydrochloride (Amersham) was added to the final concentration of 100 μ Ci ml⁻¹. At the end of the labeling period, cells were washed twice with PBS, incubated with lysis buffer (25 mM Tris pH 7.5, 150 mM NaCl, 5 mM EDTA, and protease inhibitor cocktail) containing 1% (v/v) triton X-100 or triton X-114 on ice for 30 min. The cell lysates were passed through a 25-gauge needle and then clarified by centrifugation at 17,000 $\times$ g for 5 min. In some experiments, clear lysate was subjected to triton X-114 phase separation (Bordier, 1981) prior to immunoprecipitation. The samples were then diluted with RIPA buffer (0.5% triton X-100,

0.1% SDS, 300 mM NaCl, 50 mM Tris-HCl pH 7.4 and 4 mM EDTA) and subsequently pre-cleared by incubating with protein G sepharose 4 fast flow beads (Amersham) at 4 °C for 4 h. The pre-cleared samples were incubated at 4 °C for 12 h with each of the following antibodies: 2.5 μ g of a mouse plasmacytoma protein (MOPC 21, Sigma), 2.5 μ g of a purified mouse anti-NS1 antibody (NS1-8.2), or an anti-E antibody (3H5) at the dilution of 1:500. The samples were incubated with protein G sepharose beads at 4 °C for 4 h followed by centrifugation at 17,000 $\times$ g for 5 min. The beads were washed twice with RIPA buffer. Bound proteins were eluted with gel loading buffer containing 5% mercaptoethanol and separated by electrophoresis in 10% SDS-polyacrylamide gel. In the visualization of labeled proteins, the gel was fixed with isopropanol:water:acetic acid (25:65:10) mixture for 30 min and then soaked in an enhancer solution (Amplify fluorographic reagent, Amersham) for 15 min prior to drying and exposure to X-ray film (Kodak X-Omat AR, Sigma) at -70 °C.

2.8. Digestion of surface NS1 with phosphatidylinositol-specific phospholipase C

NS1 transfectants (1 $\times$ 10⁶) were detached with 0.5 mM EDTA and washed once in 100 mM phosphate buffer containing protease inhibitor cocktail (Complete, Roche, Germany). The cells were suspended in 100 μ l of the same buffer in the presence or absence of 1.5 U ml⁻¹ of recombinant phosphatidylinositol-specific phospholipase C (PI-PLC) from *Bacillus cereus* (Sigma) at 37 °C for 1 h. Subsequently, IFA and flow cytometry were employed to assess the remaining NS1 as described above. As a positive control for PI-PLC digestion, pcDNA3.1/Hygro-transfected cells that had been digested similarly were stained successively with a biotin-conjugated anti-human CD55 antibody and the streptavidin-phycoerythrin conjugate (Pharmingen) at the final dilutions of 1:20 and 1:10, respectively.

3. Results

3.1. Establishment of 293T lines expressing recombinant NS1

293T cells were transfected with the expression vectors encoding the rNS1s, rNS1v1, and rNS1tm forms derived from strain NGC together with the murine Ig κ leader sequence. On day 3 after transfection, the expression of recombinant proteins was assessed by IFA. Approximately 10–30% of cells transfected with these plasmids expressed cytoplasmic NS1 antigen as compared with none of the pcDNA3.1/Hygro-transfected cells (data not shown). During serial passages of transfected cells in the presence of hygromycin, cell death became apparent within 7 days and the concentration of 100 μ g ml⁻¹ seemed to be most suitable for the selection (data not shown). Four weeks after starting the selection, isolated colonies were identified visually and 24 colonies or more were picked for the second round screening of NS1 expression by IFA within a period of three weeks. Individual clones were selected based on the relative level and

the proportion of cells expressing NS1. Eventually, two clones expressing high levels of NS1 from each recombinant form were obtained.

3.2. Intracellular and cell surface expression

The expression of recombinant NS1 was determined by IFA and flow cytometry within the 30–50th passages of selected clones. Using antibodies specific for a linear epitope (1B2) and a conformational epitope (NS1-8.2), up to 50–90% of the transfectants were found to express NS1 intracellularly at any time regardless of the antibody specificity (Fig. 2B–D and 3). Staining of fixed and permeabilized cells revealed that, for all NS1 transfectants, NS1 proteins dispersed throughout the cytoplasm,

while in the rNS1v1 and rNS1tm transfectants, NS1 could be observed on the plasma membrane as well (Fig. 2B–D). The generalized cytoplasmic expression of recombinant NS1 was distinct from the localized, perinuclear pattern of native NS1 in dengue virus-infected cells (Fig. 2E). Cell lines stably transfected with pcDNA3.1/Hygro were used as control. No specific staining was observed (Fig. 2A).

Staining of surface-associated NS1 on unfixed, detached cells with the antibodies 1B2 or NS1-8.2 showed that the rNS1s form, which was constructed without additional C-terminal sequence, was rarely detectable on the cell surface (2–4%) by flow cytometric analysis (Figs. 2C and 3). In contrast, 30–50% of the rNS1v1 and rNS1tm transfectants were positive for surface-associated NS1 (Fig. 3). The spatial pattern of surface NS1

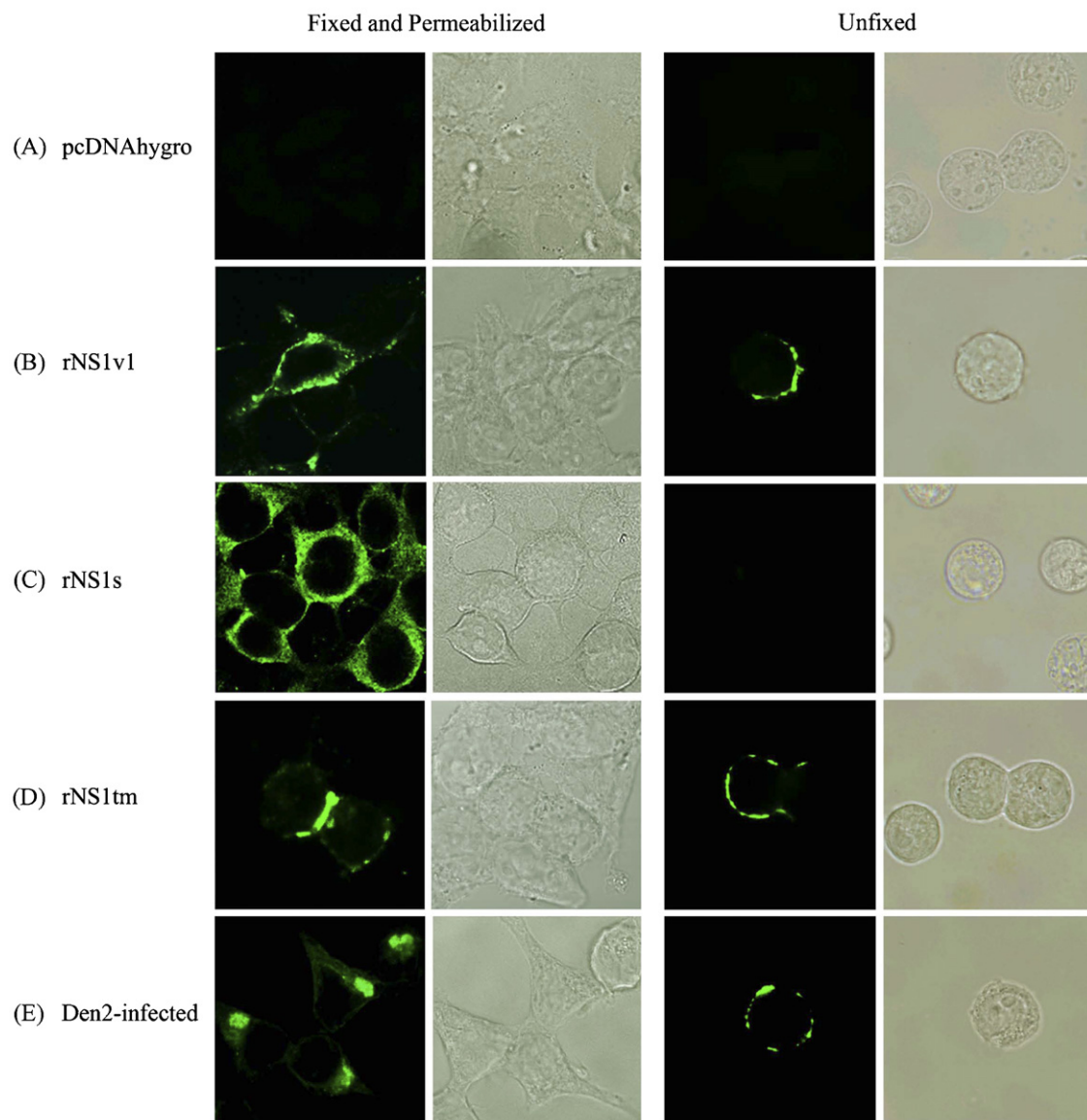


Fig. 2. Localization of NS1 in transfectants. 293T cells were transfected with either pcDNA3.1/ Hygro (vector control, A) or plasmid constructs expressing recombinant NS1 forms, rNS1v1 (B), rNS1s (C), and rNS1tm (D). The transfectants were subsequently selected for stable expression of NS1 protein by hygromycin treatment and the stable cells were verified for cell surface (unfixed, right panel) and intracellular (fixed and permeabilized, left panel) expression of NS1 by IFA using an anti-NS1 antibody (1B2) and FITC-conjugated rabbit anti-mouse Ig antibody. Virus-infected 293T cells (E) served as the positive control. The stained cells were observed under a laser-scanning confocal microscope (left panel) and a fluorescence microscope (right panel).

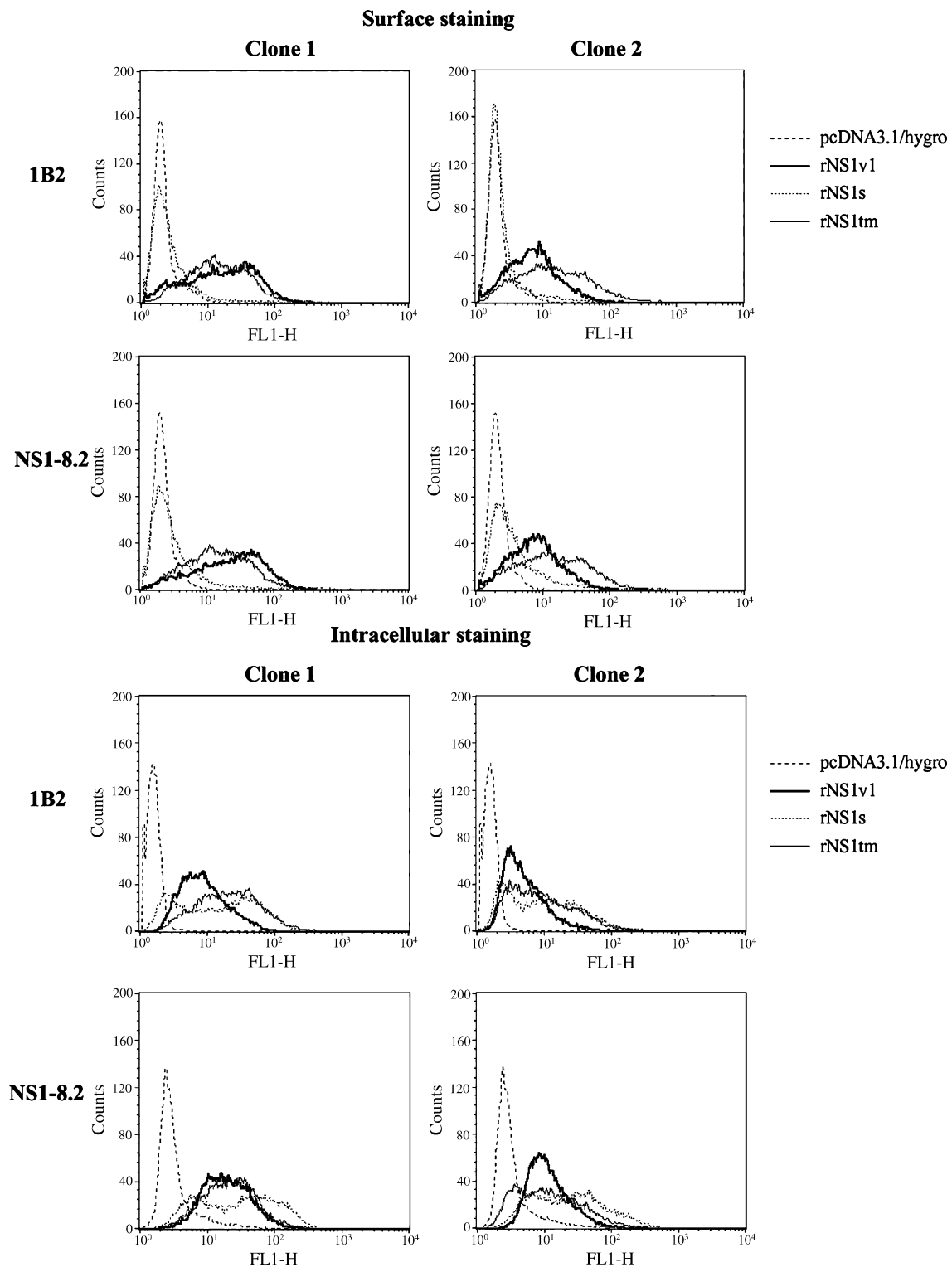


Fig. 3. Determination of intracellular and cell surface NS1 expression by FACS analysis. 293T cells that had been transfected with pcDNA3.1/Hygro vector or different NS1-expressing plasmid constructs were propagated under hygromycin selection. At the passages 9 and 12, two representative clones were subjected to IFA for intracellular and surface NS1 expression by using either an isotype-matched control antibody, or anti-NS1 antibodies specific for a linear epitope (1B2) or a conformational epitope (NS1-8.2). The percentage of NS1-expressing cells was analyzed by FACS from 10,000-gated events.

on the rNS1v1 and rNS1tm transfectants was indistinguishable from that of virus-infected cells (Fig. 2B, D-E), but the intensity of NS1 on the cell surface or inside the cells as analyzed by flow cytometry was a little bit different among the transfectants

(Fig. 3). Nevertheless, monoclonal antibodies specific for both linear and conformational epitopes recognized the recombinant NS1 proteins from all transfectants, suggesting that the three recombinant NS1 forms are folded appropriately.

3.3. Dimer formation

In dengue virus-infected cells, newly synthesized NS1 appears in the ER lumen as the 46 kDa monomer and then under-

goes dimerization within 20–40 min after translation (Winkler et al., 1988,1989). It was of interest to determine whether the recombinant NS1 forms were able to assemble into dimeric molecules in the absence of other viral proteins. The presence of

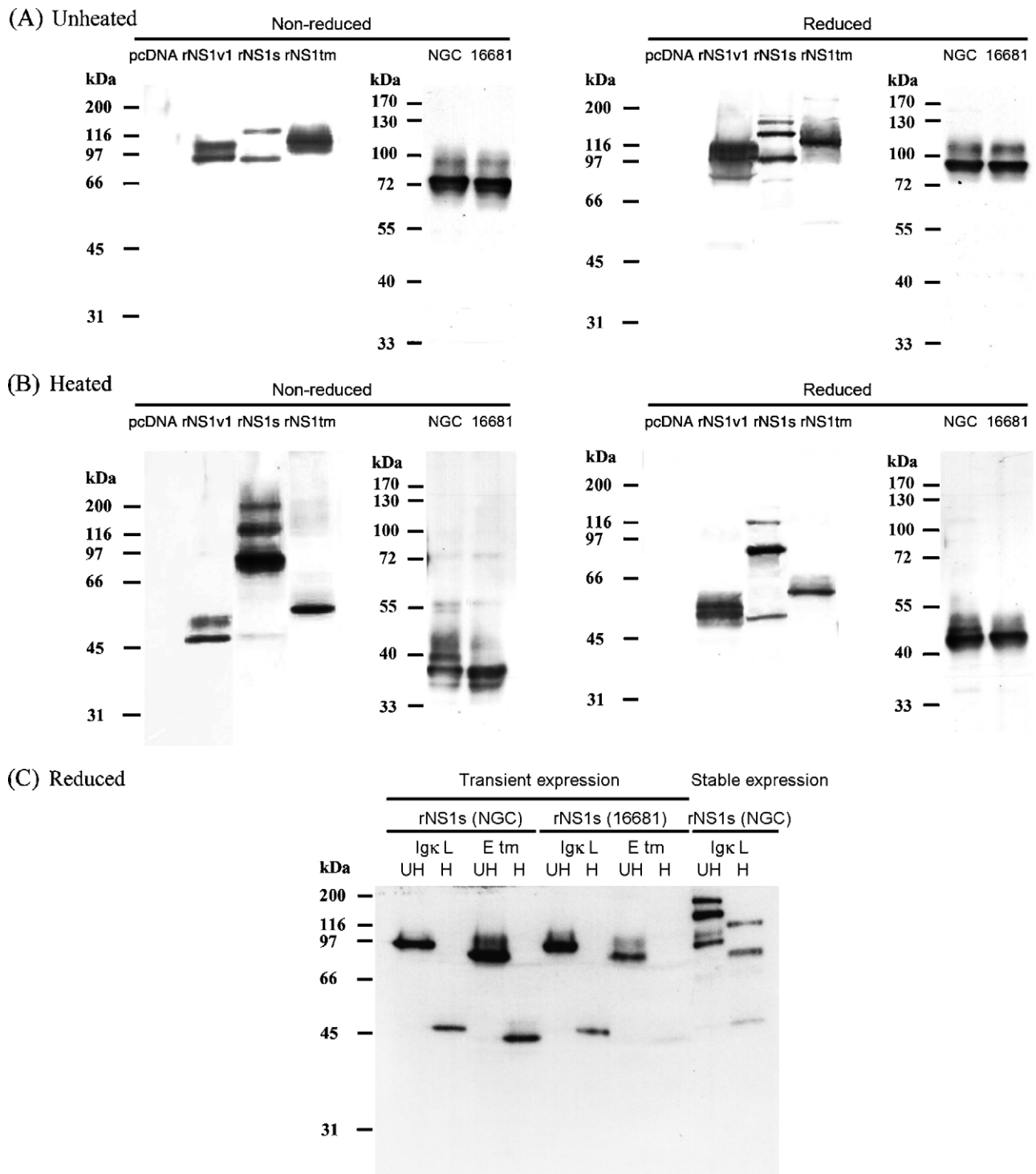


Fig. 4. Immunoblot analysis of NS1 in transfectants. Cell lysates were prepared from stable transfectants expressing the rNS1v1, rNS1s, and rNS1tm forms under non-reducing or reducing conditions. The samples were either unheated (A) or heated (B) prior to immunoblot analysis using an anti-NS1 antibody (NS1-3F.1) and HRP-conjugated rabbit anti-mouse Ig antibody. Lysate of pcDNA3.1/Hygro transfectants (pcDNA) and virus-infected 293T cells (stains NGC and 16681) served as the negative and positive controls, respectively. (C) Immunoblot analysis of rNS1s (NGC and 16681) in lysates of transiently transfected cells prepared under the reducing condition. The samples were either unheated (UH) or heated (H) prior to electrophoresis. The migration patterns of the rNS1s transiently expressed with a heterologous (Igκ L) signal peptide or a native (E tm) signal peptide were compared with that of the rNS1s from the stable transfectant. The positions of the protein standards are shown on the left.

NS1 dimers was verified by immunoblot analysis without heating the cell lysate as NS1 dimers readily dissociate when heated (Winkler et al., 1988,1989).

Immunoblot analysis of dengue virus-infected cell lysate revealed mainly the NS1 doublet of about 86 and 90 kDa, which was reduced to 36 and 38 kDa doublet after heating (Fig. 4A and B, lanes NGC and 16681). In both dengue serotype 2 strains, the 86 kDa lower NS1 band was consistently more intense than the 90 kDa upper band. Under the same conditions, the rNS1tm form migrated as a broad 108 kDa band in the absence of heat treatment and the 60 kDa band after heat treatment (Fig. 4A and B, lane rNS1tm). In both conditions, an increase in the apparent molecular weight of rNS1tm was likely due to the additions of an N-terminal hemagglutinin epitope and the C-terminal transmembrane portion. The reduction of NS1 molecular weight following heat treatment observed in the lysates of virus-infected cells and the rNS1tm transfectant indicated that these NS1 forms were non-covalently linked dimers. In the rNS1v1 transfectant lysate, two distinct bands, 90 and 103 kDa, were observed under unheated conditions (Fig. 4A, lane rNS1v1) and there was a shift of both bands to 51 and 56 kDa after heat treatment (Fig. 4B, lane rNS1v1). Heat-induced changes in the migration of the rNS1v1 form also indicated the ability of this recombinant protein to form dimeric molecules. The presence of two NS1 bands in the rNS1v1 transfectant lysate were likely due either to differential glycosylation, or to partial cleavage of NS2A from the C-terminus of rNS1v1.

Similar to the rNS1v1 form, two distinct NS1 bands were observed in the lysates from the rNS1s transfectant without heat treatment: a faster migrating 90-kDa band, which migrated to the same position as the lower dimeric band of the rNS1v1 form, and a slower migrating 120-kDa band, which was larger than the upper rNS1v1 band (Fig. 4A, lane rNS1s). Heat treatment resulted in an expected monomeric 50-kDa band and also three other slower migrating bands of about 90, 120, and 170 kDa (Fig. 4B, lane rNS1s; Fig. 5A). The three high molecular weight bands were observed without using cross-linking agent and they were resistant to 5% 2-mercaptoethanol (Fig. 4A and B) or 100 mM dithiothreitol treatment (data not shown), suggesting that aggregation occurs without the formation of intermolecular disulfide bond. As the rNS1s form was expressed with a heterologous signal peptide and a hemagglutinin epitope, the possibility that these high molecular weight bands were associated with extra N-terminal sequences was tested by generating additional rNS1s cassettes with the transmembrane domain of E protein (Fig. 1). Transient expression of the NGC-derived rNS1s form with the native signal peptide did not yield high molecular weight bands (Fig. 4C). Surprisingly, the rNS1s construct with the Ig κ leader sequence also did not produce extra bands when tested in transiently expressing cells (Fig. 4C). Similar results were observed when the NS1 sequence was derived from strain 16681 (Fig. 4C). These findings suggested that the observed propensity of the NGC-derived rNS1s to aggregate was related to its expression in stable transfectants, but not a general property of dengue NS1.

3.4. N-glycosylation

Previous studies on the post-translational modifications of dengue virus NS1 in mammalian cells revealed that intracellular NS1 contains two N-linked, high-mannose glycans and that one of these glycans is further modified into the complex structure in Golgi apparatus just prior to release (Smith and Wright, 1985; Winkler et al., 1988). Whether the recombinant NS1 forms were modified similarly with N-linked glycans was examined. NS1 in the stable transfectant lysates and the culture media were subjected to digestion with Endo H or PNGase F, and an immunoblot analysis. The N-glycosylation in all three forms of recombinant NS1 was compatible to that of NS1 from virus-infected cells but with subtle differences (Fig. 5A). Cell-associated NS1 from infected cells, the rNS1s and rNS1tm transfectants contained

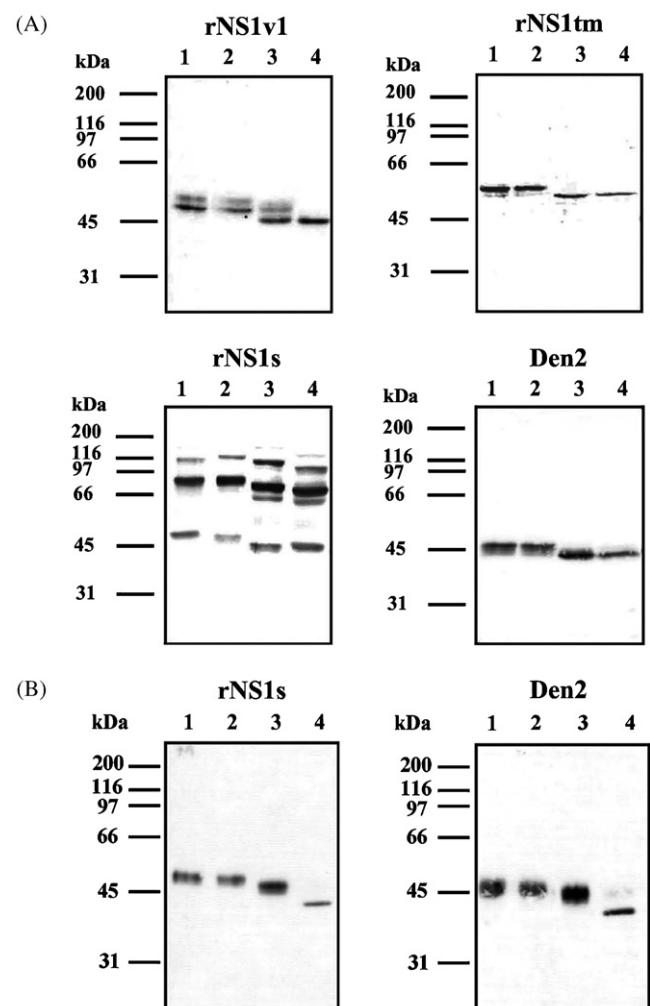


Fig. 5. Analysis of N-glycosylation of NS1 by glycosidase treatment. Cell lysate (A) and supernatant (B) prepared from transfectants expressing the rNS1v1, rNS1s, and rNS1tm forms and strain 16681-infected cells were treated with Endo H (lane 3) or PNGase F (lane 4) under reducing and heated conditions. Lysate that had been processed in the same conditions without the enzymes served as negative control (lanes 1 and 2 for Endo H and PNGase F treatment, respectively). Following an 18-h digestion, samples were subjected to 10% SDS-PAGE and immunoblot analysis using an anti-NS1 antibody (NS1-3F.1) and HRP-conjugated rabbit anti-mouse Ig antibody. The positions of the protein standards are shown on the left.

only the high mannose-type glycan as treatment with Endo H or PNGase F caused a similar increase in the mobility of NS1 monomers (Fig. 5A). Digestion of extracellular NS1 with Endo H resulted in an enhanced electrophoretic mobility of native NS1 and the rNS1s form, but not to the same extent as PNGase F treatment (Fig. 5B). The levels of rNS1v1 and rNS1tm forms in the culture media were below the sensitivity of immunoblot analysis (data not shown). Partial sensitivity to Endo H indicated the presence of complex type glycan at one of the two *N*-glycosylation sites in NS1 released from rNS1s transfectant and virus-infected cells. Our detection of only the high mannose-type glycan on cell-associated native NS1 and rNS1s form, therefore, was consistent with previous findings, (Despres et al., 1991; Fan and Mason, 1990; Flamand et al., 1992; Mason, 1989; Muylaert et al., 1996), suggesting a rapid release of NS1 molecules with the complex type glycans from an intracellular pool. A striking difference between cell-associated NS1 and extracellular NS1 from NGC-derived rNS1s transfectant was also evident: extracellular rNS1s migrated as single monomeric band after heat treatment as opposed to multiple bands of its intracellular counterpart (Fig. 5A and B), possibly reflecting selective release of NS1 dimers.

In contrast to the rNS1s and rNS1tm forms, Endo H digestion of NS1 in the lysate of rNS1v1 stable transfectant revealed that the two NS1 bands were not equally sensitive to deglycosylases. The high molecular weight band was partially resistant to Endo H digestion, but sensitive to PNGase F, whereas the low molecular weight band was sensitive to both Endo H and PNGase F (Fig. 5A). This finding was suggestive of two species of rNS1v1 with differential modifications of *N*-linked glycans. The low molecular weight band contained the high-mannose type glycans at both *N*-glycosylation sites while the high molecular weight band contained the complex type glycan at one *N*-glycosylation site and the high mannose type at the other. The appearance of fully deglycosylated NS1 as a sharp band following PNGase F digestion was incompatible with the partial cleavage of NS2A segment from the C-terminus of NS1 in the generation of distinct electrophoretic species of the rNS1v1 form. Similar to the rNS1tm form, the rNS1v1 form was scarcely detectable in the culture media by immunoblot analysis.

Analysis of *N*-glycosylation, thus, revealed that these recombinant NS1 forms underwent initial *N*-glycosylation in 293T cells independently of membrane association, but subsequent modifications were detected only with the cell-associated rNS1v1 and extracellular rNS1s forms. The lack of complex type glycan in the cell-associated rNS1tm molecules cannot be explained by its rapid release as in the cases of NS1 from the rNS1s transfectant and virus-infected cells. The two membrane-associated recombinant NS1 forms may be present in different environments and/or acquire distinctive spatial orientations that affect their interaction with the *N*-glycan modifying enzymes.

3.5. Association of the rNS1v1 and rNS1tm forms with surface membrane

In dengue virus-infected cells, the association of NS1 with cell surface is unusual as this protein lacks the membrane-

spanning hydrophobic domain (Wright et al., 1989). The rNS1v1 form was constructed based on a previous study that revealed membrane association of NS1 via GPI-anchoring when NS1 was expressed with the 26-residue hydrophobic segment at the N-terminus of NS2A (Jacobs et al., 2000). The linkage of NS1 in the rNS1v1 transfectant with GPI was determined by subjecting transfected 293T cells to digestion with PI-PLC and flow cytometric analysis. A cellular protein known to associate with cell membrane via the GPI-linkage, CD55, was used as a control to show that a GPI-linked protein could be cleaved completely and removed from the surface of 293T cells with PI-PLC treatment (Fig. 6B). Without the PI-PLC treatment, NS1 could be detected readily on the surface of rNS1v1 and rNS1tm transfectants but not on the rNS1s transfectant (Fig. 6A). Following PI-PLC treatment, reduction of a large proportion of surface-associated NS1 was observed in the rNS1v1 transfectant (Fig. 6A); increasing the concentration of PI-PLC did not result in complete elimination of surface-associated rNS1v1 (data not shown). These results indicated that, in the rNS1v1 transfectant, some NS1 molecules associated with cell surface in the PI-PLC-sensitive manner, either directly through their own GPI anchors or indirectly through binding with other GPI-linked cellular protein(s). However, our failure to detect NS1 on the surface of rNS1s transfectant suggested that the possibility of peripheral association of NS1 with GPI-linked cellular protein was unlikely. Under the same conditions, NS1 on the cell surface of rNS1tm transfectant was resistant to PI-PLC digestion (Fig. 6A), consistent with the expectation that rNS1tm is an integral membrane protein.

A possibility that the recombinant NS1 forms associate non-covalently with cellular protein(s) was examined by metabolic labeling newly synthesized proteins with ³⁵S-labeled precursors and immunoprecipitation using anti-NS1, anti-E, and isotype-matched control antibodies. Immunoprecipitation of SDS lysate of ³⁵S-labeled, virus-infected cells with anti-E antibody resulted in four bands corresponding to E doublet and prM (the expected non-covalently linked, intracellular E + prM heterodimer) and also an unknown protein (Fig. 7). An anti-E antibody did not pull down any specific band from the transfectant lysate (Fig. 7). An anti-NS1 antibody precipitated NS1 dimers from lysate of virus-infected cells and transfectants in the patterns closely similar to the immunoblot analysis of NS1 in the absence of heat treatment (Figs. 4 and 7). As compared with isotype-matched control and anti-E antibodies, an anti-NS1 antibody did not reveal NS1-coprecipitated protein in the transfectants and virus-infected cells (Fig. 7). Similar results were observed when infected cells were lysed with 1% triton X-114, 1% NP40 or 1% Brj 58 (data not shown). Thus, association of the recombinant NS1 from transfectants as well as the native NS1 molecules with cellular protein could not be detected by immunoprecipitation, suggesting that dengue NS1 may not need to associate with cellular and other viral proteins to be localized on the cell membrane.

3.6. Changes in hydrophobicity and selective GPI linkage of surface-associated recombinant NS1

The presence of C-terminal sequences in the rNS1tm and rNS1v1 constructs allowed NS1 to associate with the cell sur-

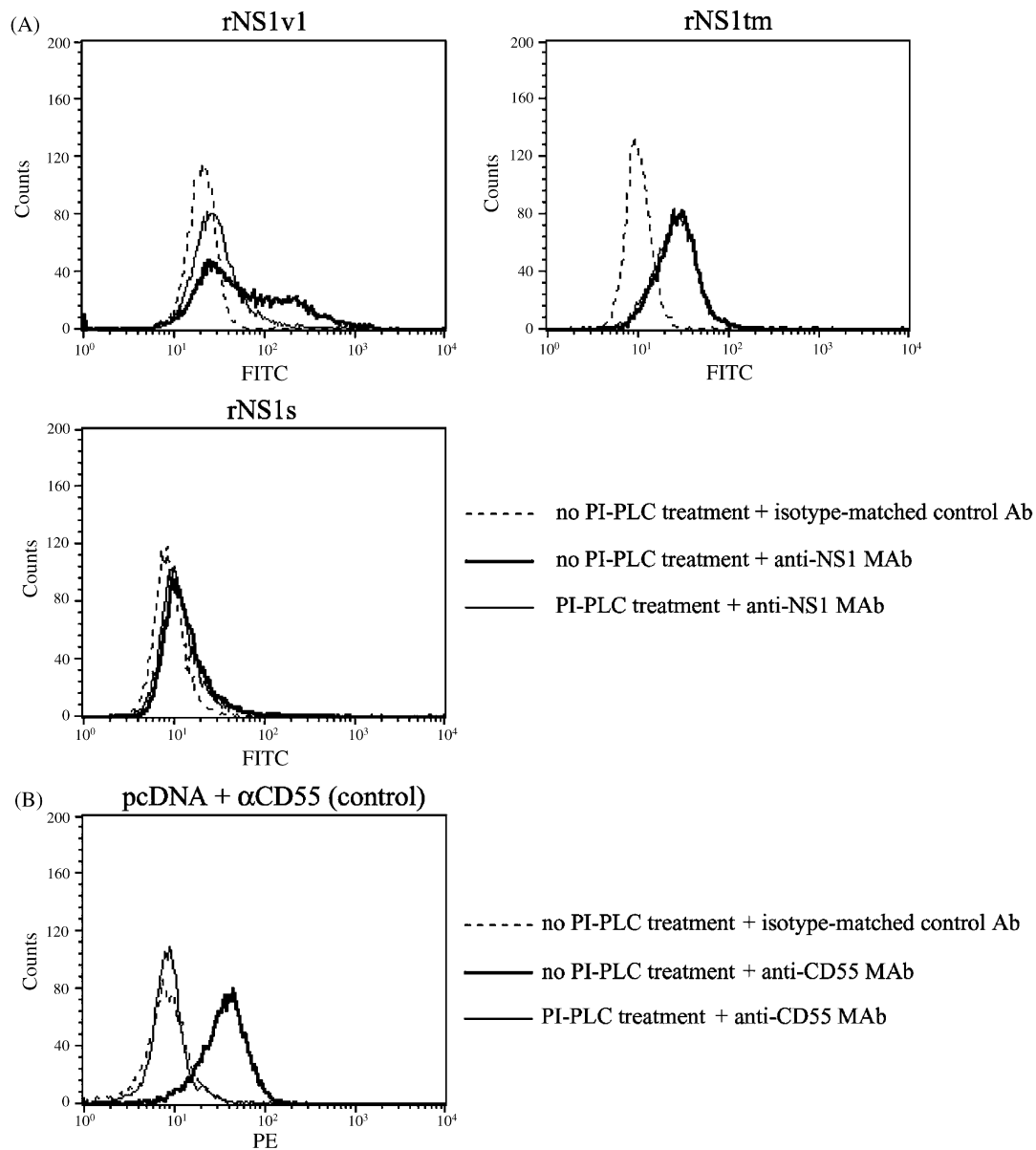


Fig. 6. Digestion of recombinant NS1 on the cell surface with PI-PLC. (A) Stable NS1 transfectants and infected 293T cells were digested with PI-PLC at 37 °C for 1 h and the remaining NS1 were assessed by IFA using a control antibody or an anti-NS1 antibody (1B2) followed by FITC-conjugated rabbit anti-mouse Ig antibody. (B) As a positive control for PI-PLC digestion, pcDNA3.1/Hygro transfectants that had been subjected to PI-PLC treatment were stained with a biotin-conjugated anti-human CD55 antibody and the streptavidin-phycoerythrin conjugate. Stained cells were analyzed for cell surface CD55 or NS1 by flow cytometric analysis using 10,000-gated events. Similar results were observed in two other separate experiments.

face (Figs. 2, 3 and 6). We next sought to determine whether membrane association was accompanied by an increase in the hydrophobicity of recombinant NS1 proteins. The three NS1 transfectants and dengue virus-infected cells were metabolically labeled with [³⁵S] cysteine/methionine, lysed, and subjected to the triton X-114 phase separation (Bordier, 1981). Extracted proteins in the aqueous and detergent phases were then precipitated with anti-NS1 or anti-E antibodies and analyzed by SDS-polyacrylamide gel electrophoresis and fluorography. As an indicator of complete separation, E and prM from infected cells were detected exclusively in the detergent phase as expected of proteins with hydrophobic membrane spanning regions (Fig. 8A). In accordance with previous findings, NS1

dimers from virus-infected cells distributed in both of the aqueous and detergent phases (Despres et al., 1991; Fan and Mason, 1990; Flamand et al., 1992; Leblois and Young, 1995; Mason, 1989; Winkler et al., 1989). NS1 from the rNS1s transfectant were detected similarly in the two phases, indicating that the amphipathicity intrinsic to dimeric NS1 was retained in the rNS1s form. However, NS1 dimers from the rNS1v1 and rNS1tm transfectants were found only in the detergent phase, suggesting that the C-terminal sequence additions may result in an increase in the hydrophobicity of these two recombinant NS1 forms (Fig. 8A). No specific ³⁵S-labeled band was observed in the precipitation of NS1 transfectants with anti-E and isotype-matched control antibodies (Fig. 8A).

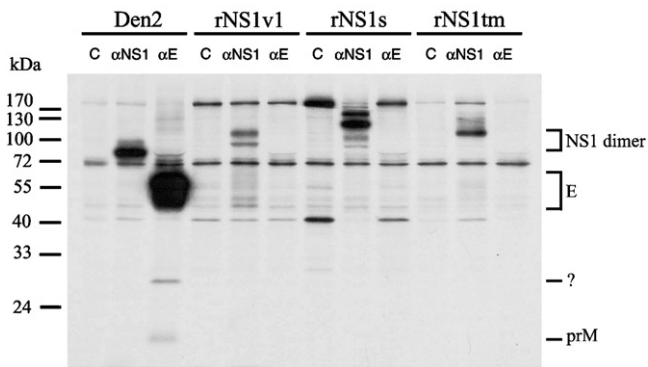


Fig. 7. Immunoprecipitation of ^{35}S -labeled NS1 proteins. NS1 transfectants and infected 293T cells were metabolically labeled with $50\ \mu\text{Ci ml}^{-1}$ of [^{35}S]cysteine/methionine for 16–18 h, and then lysed with 1% triton X-100. Cell lysate was subjected to immunoprecipitation using a control antibody (lane C), or the antibodies specific for NS1 (lane αNS1) and E (lane αE). Precipitated proteins were separated by electrophoresis in 10% SDS-polyacrylamide gel in the presence of 2-mercaptoethanol without heat treatment and visualized by fluorography. The positions of the protein standards are shown on the left.

Whether the presence of a C-terminal sequence in the rNS1v1 transfectant resulted in an incorporation of the GPI anchor was further examined by metabolically labeling the rNS1v1 and rNS1tm transfectants with ^3H -ethanolamine, a precursor of the GPI moiety. The labeled cells were lysed and subjected to the triton X-114 phase separation and immunoprecipitation. As expected, an anti-NS1 antibody precipitated the ^3H -labeled NS1 dimers present in the detergent phase from the rNS1v1 transfectant, but not the rNS1tm transfectant (Fig. 8B). No specific protein band was observed from these samples when an anti-E antibody was employed (Fig. 8B). It should be noted that the ^3H -labeled NS1 bands from the rNS1v1 transfectant became visible after a rather prolonged exposure (2 months) to a high sensitivity X-ray film. These results were compatible with an incorporation of the GPI moiety into the rNS1v1 form.

4. Discussion

NS1 remains one of the poorly characterized non-structural proteins of dengue virus. Previously, several protein expression systems have been used in the investigation of the physical and

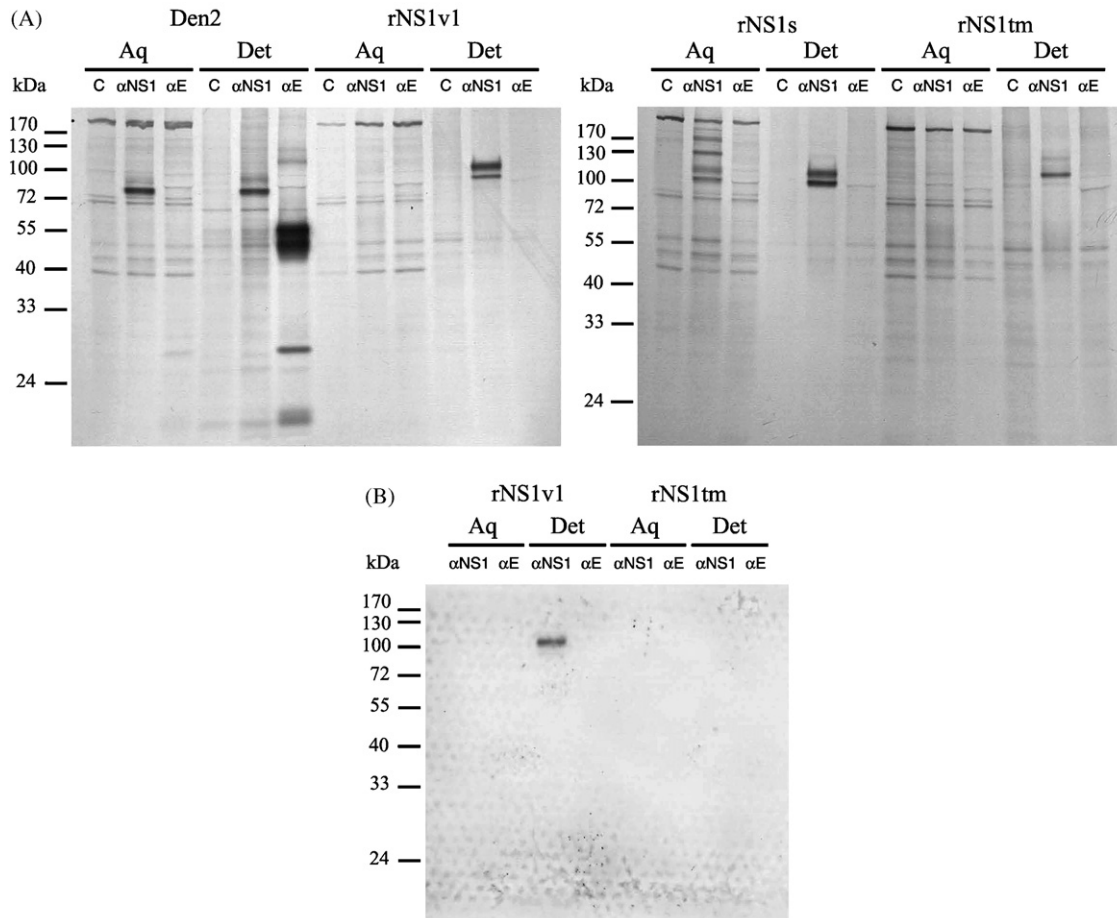


Fig. 8. Triton X-114 phase separation and incorporation of a GPI precursor. NS1 transfectants and infected 293T cells were metabolically labeled with [^{35}S]cysteine/methionine (A), or [^3H] ethanolamine (B) for 16–18 h, lysed with buffer containing 1% triton X-114 and then subjected to temperature-induced phase separation. Aqueous (Aq) and detergent (Det) phases were employed in the immunoprecipitation using a control antibody (lane C), an anti-NS1 antibody (lane αNS1), or an anti-E antibody (lane αE). Precipitated proteins were separated by electrophoresis in 10% SDS-polyacrylamide gel in the presence of 2-mercaptoethanol and visualized by fluorography.

biological properties of dengue NS1 and its possible contribution in dengue virus infection and the pathogenesis of DHF/DSS. These systems rely on the use of vaccinia virus- and baculovirus-based gene constructs for protein expression in mammalian or insect cells (Falgout et al., 1989; Falgout and Markoff, 1995; Hori and Lai, 1990; Leblois and Young, 1995; Parrish et al., 1991; Pethel et al., 1992; Zhao et al., 1987). More recently, a virus-free system, in which a eukaryotic expression vector directed the transient expression of NS1 in HeLa cell line, was reported (Jacobs et al., 2000). In this study, a prolonged expression of distinct recombinant NS1 forms was achieved by employing an expression vector with a selectable marker and identifying drug resistant 293T transfectants that expressed high levels of intracellular NS1 during subsequent passages.

In this long-term system, NS1 expression has been detected for up to 50 passages under the continuous presence of hygromycin. Although the level of NS1 protein in individual transfected cell is variable, the distribution of recombinant NS1 forms remains constant and reflects differences at the C-termini. The rNS1s form, deprived of additional C-terminal sequence, was released into the culture medium, but undetected on the cell surface. The two recombinant forms with C-terminal sequences prominently associated with the cell membrane while being scarcely detectable in the culture medium. Similar to the native NS1 from dengue virus-infected cells, all three NS1 forms were able to undergo dimerization and *N*-glycosylation. However, each recombinant NS1 form exhibited unique phenotypic characteristics. An initial *N*-glycosylation with the high-mannose glycans occurred in these recombinant forms, but additional modifications resulting in the complex glycans were found only in the cell-associated rNS1v1 and extracellular rNS1s molecules. The latter closely resembles the modification of NS1 generated in virus-infected cells. Different *N*-glycosylation patterns of intracellular NS1 in the two membrane-associated forms (rNS1v1 and rNS1tm) possibly reflects subtle influences in which the C-terminal hydrophobic sequence of the rNS1tm and the GPI moiety of the rNS1v1 form differentially affect the interaction of NS1 with *N*-glycan modifying enzymes. Dimerization and *N*-glycosylation of dengue NS1 are, therefore, independent of membrane association, but the modification of *N*-linked glycans on membrane-associated NS1 may not be achieved similarly in the stable transfectants.

Characterization of the two membrane-associated NS1 forms supports the presence of distinct mechanisms underlying the association of NS1 molecules in the rNS1tm and rNS1v1 transfectants with membrane. The C-terminal additional sequence in the rNS1tm form likely functions directly as the membrane-spanning domain as this form appeared to be more hydrophobic than NS1 without additional sequence while being insensitive to PI-PLC digestion and unable to incorporate a GPI precursor. On the contrary, NS1 in the rNS1v1 transfectant incorporated a GPI precursor and became sensitive to PI-PLC treatment, strongly indicating that the NS2A-derived C-terminal sequence serves as a signal for linking GPI moiety to the C-terminus of NS1. Our results confirm a previous study on the role of the 26-residue segment at the N-terminus of NS2A as a GPI-linking signal that was performed in HeLa cells using different expression con-

struct (Jacobs et al., 2000), and suggest that GPI modification of dengue virus NS1 is not cell type-specific. However, it is notable that PI-PLC treatment did not remove NS1 completely from the surface of rNS1v1 transfectants. The partial sensitivity of rNS1v1 form to PI-PLC digestion may be due to the failure to link all surface-associated NS1 molecules with the GPI anchor in the rNS1v1 transfectant. Alternatively, further modification of the GPI moiety, such as an acylation of the inositol group (Ferguson, 1999), may render some NS1 molecules resistant to cleavage. Another possible, but unlikely, explanation is an association of the rNS1v1 form with cellular GPI-linked protein(s) in the cell surface; this is incompatible with the failure to precipitate additional protein band with anti-NS1 antibody in the immunoprecipitation experiment.

An intriguing finding from the characterization of recombinant NS1 forms is the lack of detectable NS1 on the surface of rNS1s transfectant. It has been shown that NS1 dimers in dengue virus-infected cells become more hydrophobic and membrane-associated after dimerization (Winkler et al., 1989). The rNS1s transfectant generated NS1 dimers that were indistinguishable from the native NS1 from infected cells in the triton X-114 phase separation. The failure to detect the rNS1s form on cell surface suggests that an increase in the hydrophobicity after NS1 dimerization by itself is not adequate for an efficient association of NS1 with the surface membrane. Also, the binding of extracellular NS1 to the cell surface appears to be negligible in the rNS1s transfectant. This finding suggests that the association of NS1 with the surface membrane of infected cells requires additional posttranslational modification of NS1. This could possibly involve the GPI-linkage as proposed previously by Jacobs et al. (2000) or other unknown mechanism.

A comparison of the intracellular rNS1s molecules in transiently transfected cells and stably expressing cells revealed a difference that stipulates caution in the interpretation of data derived from cells producing an exogenous protein under selective pressure. Heat treatment of the lysate of rNS1s stable transfectants, but not transiently transfected cells, generated high molecular weight aggregates, which appeared to be unrelated to the additional N-terminal sequence or a heterologous signal peptide. The failure to observe these heat-induced aggregates in the lysate of transiently transfected cells expressing NS1 from another dengue strain, irrespective of the source of signal peptide, further strengthens a restricted occurrence of the rNS1s aggregates. The conditions that are favorable to the formation of aggregates remain unclear, but may involve an altered level of partially folded molecules in the highly expressing cells and/or an unusual propensity of the soluble NS1 form to aggregate. These results serve to deter a direct extrapolation of the results obtained from stable transfectants to that of dengue virus-infected cells.

In summary, the established multiple forms of NS1 transfectants could be a useful *in vitro* model for dissecting the intrinsic property of particular NS1 protein forms and their potential role in dengue virus infection and host responses. These NS1 transfectants had proved valuable in a recent study assessing the role of soluble and membrane-associated forms of dengue NS1 in complement activation (Avirutnan et al., 2006).

Acknowledgements

We thank Dr. Guey Chuen Perng, Department of Virology, USAMC-AFRIMS for helpful discussion and editing of the manuscript. This work was supported by the Thailand-Tropical Diseases Research (T2) Program (01-1-DEN-02-006 to SN, and 99-2-DEN-03-008 and 02-2-DEN-03-003 to N.S.), the Senior Research Scholarship Program from Thailand Research Fund (to P.M.), and the National Center for Genetic Engineering and Biotechnology, National Science and Technology Development Agency, Thailand.

References

- Alcon, S., Talarmin, A., Debruyne, M., Falconar, A., Deubel, V., Flamand, M., 2002. Enzyme-linked immunosorbent assay specific to dengue virus type 1 nonstructural protein NS1 reveals circulation of the antigen in the blood during the acute phase of disease in patients experiencing primary or secondary infections. *J. Clin. Microbiol.* 40, 376–381.
- Avirutnan, P., Malasit, P., Seliger, B., Bhakdi, S., Husmann, M., 1998. Dengue virus infection of human endothelial cells leads to chemokine production, complement activation, and apoptosis. *J. Immunol.* 161, 6338–6346.
- Avirutnan, P., Punyadee, N., Noisakran, S., Komoltri, C., Thiemmea, S., Auethavornanan, K., Jairungsri, A., Kanlaya, R., Tangthawornchaikul, N., Puttikhunt, C., Pattanakitsakul, S., Yenchitsomanus, P., Mongkolsapaya, J., Kasinrer, W., Sittisombut, N., Husmann, M., Blettner, M., Vasanawathana, S., Bhakdi, S., Malasit, P., 2006. Vascular leakage in severe dengue virus infections: a potential role for the non-structural viral protein NS1 and complement. *J. Infect. Dis.* 193, 1078–1088.
- Bordier, C., 1981. Phase separation of integral membrane proteins in Triton X-114 solution. *J. Biol. Chem.* 256, 1604–1607.
- Brandt, W.E., Chiewslip, D., Harris, D.L., Russell, P.K., 1970. Partial purification and characterization of a dengue virus soluble complement-fixing antigen. *J. Immunol.* 105, 1565–1568.
- Chang, H.H., Shyu, H.F., Wang, Y.M., Sun, D.S., Shyu, R.H., Tang, S.S., Huang, Y.S., 2002. Facilitation of cell adhesion by immobilized dengue viral non-structural protein 1 (NS1): arginine–glycine–aspartic acid structural mimicry within the dengue viral NS1 antigen. *J. Infect. Dis.* 186, 743–751.
- Despres, P., Girard, M., Bouloy, M., 1991. Characterization of yellow fever virus proteins E and NS1 expressed in Vero and Spodoptera frugiperda cells. *J. Gen. Virol.* 72, 1331–1342.
- Falconar, A.K., 1997. The dengue virus nonstructural-1 protein (NS1) generates antibodies to common epitopes on human blood clotting, integrin/adhesin proteins and binds to human endothelial cells: potential implications in haemorrhagic fever pathogenesis. *Arch. Virol.* 142, 897–916.
- Falgout, B., Chanock, R., Lai, C.J., 1989. Proper processing of dengue virus nonstructural glycoprotein NS1 requires the N-terminal hydrophobic signal sequence and the downstream nonstructural protein NS2a. *J. Virol.* 63, 1852–1860.
- Falgout, B., Markoff, L., 1995. Evidence that flavivirus NS1-NS2A cleavage is mediated by a membrane-bound host protease in the endoplasmic reticulum. *J. Virol.* 69, 7232–7243.
- Fan, W.F., Mason, P.W., 1990. Membrane association and secretion of the Japanese encephalitis virus NS1 protein from cells expressing NS1 cDNA. *Virology* 177, 470–476.
- Ferguson, M.A., 1999. The structure, biosynthesis and functions of glycosylphosphatidylinositol anchors, and the contributions of trypanosome research. *J. Cell Sci.* 112, 2799–2809.
- Flamand, M., Deubel, V., Girard, M., 1992. Expression and secretion of Japanese encephalitis virus nonstructural protein NS1 by insect cells using a recombinant baculovirus. *Virology* 191, 826–836.
- Flamand, M., Megret, F., Mathieu, M., Lepault, J., Rey, F.A., Deubel, V., 1999. Dengue virus type 1 nonstructural glycoprotein NS1 is secreted from mammalian cells as a soluble hexamer in a glycosylation-dependent fashion. *J. Virol.* 73, 6104–6110.
- Halstead, S.B., 1997. Epidemiology of dengue and dengue hemorrhagic fever. In: Gubler, D.J., Kuno, G. (Eds.), *Dengue and Dengue Hemorrhagic Fever*. CAB International, New York, pp. 23–44.
- Henchal, E.A., Gentry, M.K., McCown, J.M., Brandt, W.E., 1982. Dengue virus-specific and flavivirus group determinants identified with monoclonal antibodies by indirect immunofluorescence. *Am. J. Trop. Med. Hyg.* 31, 830–836.
- Hori, H., Lai, C.J., 1990. Cleavage of dengue virus NS1-NS2A requires an octapeptide sequence at the C terminus of NS1. *J. Virol.* 64, 4573–4577.
- Irie, K., Mohan, P.M., Sasaguri, Y., Putnak, R., Padmanabhan, R., 1989. Sequence analysis of cloned dengue virus type 2 genome (New Guinea-C strain). *Gene* 75, 197–211.
- Jacobs, M.G., Robinson, P.J., Bletchly, C., Mackenzie, J.M., Young, P.R., 2000. Dengue virus nonstructural protein 1 is expressed in a glycosylphosphatidylinositol-linked form that is capable of signal transduction. *FASEB J.* 14, 1603–1610.
- Leblois, H., Young, P.R., 1995. Maturation of the dengue-2 virus NS1 protein in insect cells: effects of downstream NS2A sequences on baculovirus-expressed gene constructs. *J. Gen. Virol.* 76, 979–984.
- Libraty, D.H., Young, P.R., Pickering, D., Endy, T.P., Kalayanarooj, S., Green, S., Vaughn, D.W., Nisalak, A., Ennis, F.A., Rothman, A.L., 2002. High circulating levels of the dengue virus nonstructural protein NS1 early in dengue illness correlate with the development of dengue hemorrhagic fever. *J. Infect. Dis.* 186, 1165–1168.
- Lindenbach, B.D., Rice, C.M., 1997. trans-Complementation of yellow fever virus NS1 reveals a role in early RNA replication. *J. Virol.* 71, 9608–9617.
- Lindenbach, B.D., Rice, C.M., 1999. Genetic interaction of flavivirus nonstructural proteins NS1 and NS4A as a determinant of replicase function. *J. Virol.* 73, 4611–4621.
- Mackenzie, J.M., Jones, M.K., Young, P.R., 1996. Immunolocalization of the dengue virus nonstructural glycoprotein NS1 suggests a role in viral RNA replication. *Virology* 220, 232–240.
- Mason, P.W., 1989. Maturation of Japanese encephalitis virus glycoproteins produced by infected mammalian and mosquito cells. *Virology* 169, 354–364.
- Muylaert, I.R., Chambers, T.J., Galler, R., Rice, C.M., 1996. Mutagenesis of the N-linked glycosylation sites of the yellow fever virus NS1 protein: effects on virus replication and mouse neurovirulence. *Virology* 222, 159–168.
- Muylaert, I.R., Galler, R., Rice, C.M., 1997. Genetic analysis of the yellow fever virus NS1 protein: identification of a temperature-sensitive mutation which blocks RNA accumulation. *J. Virol.* 71, 291–298.
- Parrish, C.R., Woo, W.S., Wright, P.J., 1991. Expression of the NS1 gene of dengue virus type 2 using vaccinia virus. Dimerisation of the NS1 glycoprotein. *Arch. Virol.* 117, 279–286.
- Pethel, M., Falgout, B., Lai, C.J., 1992. Mutational analysis of the octapeptide sequence motif at the NS1-NS2A cleavage junction of dengue type 4 virus. *J. Virol.* 66, 7225–7231.
- Pryor, M.J., Wright, P.J., 1993. The effects of site-directed mutagenesis on the dimerization and secretion of the NS1 protein specified by dengue virus. *Virology* 194, 769–780.
- Puttikhunt, C., Kasinrer, W., Srisa-ad, S., Duangchinda, T., Silakate, W., Moon-som, S., Sittisombut, N., Malasit, P., 2003. Production of anti-dengue NS1 monoclonal antibodies by DNA immunization. *J. Virol. Meth.* 109, 55–61.
- Ruangjirachuporn, W., Boonpucknavig, S., Nimmanitya, S., 1979. Circulating immune complexes in serum from patients with dengue haemorrhagic fever. *Clin. Exp. Immunol.* 36, 46–53.
- Schlesinger, J.J., Brandriss, M.W., Putnak, J.R., Walsh, E.E., 1990. Cell surface expression of yellow fever virus non-structural glycoprotein NS1: consequences of interaction with antibody. *J. Gen. Virol.* 71, 593–599.
- Sittisiri, K., 1994. M.Sc. Thesis, Mahidol University.
- Smith, G.W., Wright, P.J., 1985. Synthesis of proteins and glycoproteins in dengue type 2 virus-infected vero and Aedes albopictus cells. *J. Gen. Virol.* 66, 559–571.
- Sriburi, R., Keelapang, P., Duangchinda, T., Pruksakorn, S., Maneekarn, N., Malasit, P., Sittisombut, N., 2001. Construction of infectious dengue 2 virus cDNA clones using high copy number plasmid. *J. Virol. Meth.* 92, 71–82.

- Tarentino, A.L., Gomez, C.M., Plummer Jr., T.H., 1985. Deglycosylation of asparagine-linked glycans by peptide: *N*-glycosidase F. *Biochemistry* 24, 4665–4671.
- Winkler, G., Maxwell, S.E., Ruemmler, C., Stollar, V., 1989. Newly synthesized dengue-2 virus nonstructural protein NS1 is a soluble protein but becomes partially hydrophobic and membrane-associated after dimerization. *Virology* 171, 302–305.
- Winkler, G., Randolph, V.B., Cleaves, G.R., Ryan, T.E., Stollar, V., 1988. Evidence that the mature form of the flavivirus nonstructural protein NS1 is a dimer. *Virology* 162, 187–196.
- Wright, P.J., Cauchi, M.R., Ng, M.L., 1989. Definition of the carboxy termini of the three glycoproteins specified by dengue virus type 2. *Virology* 171, 61–67.
- Young, P.R., Hilditch, P.A., Bletchly, C., Halloran, W., 2000. An antigen capture enzyme-linked immunosorbent assay reveals high levels of the dengue virus protein NS1 in the sera of infected patients. *J. Clin. Microbiol.* 38, 1053–1057.
- Zhao, B.T., Prince, G., Horswood, R., Eckels, K., Summers, P., Chanock, R., Lai, C.J., 1987. Expression of dengue virus structural proteins and nonstructural protein NS1 by a recombinant vaccinia virus. *J. Virol.* 61, 4019–4022.

Distal Renal Tubular Acidosis Associated With *Anion Exchanger 1* Mutations in Children in Thailand

Sookkasem Khositseth, MD,¹ Apiwan Sirikanerat, MD,² Kulruedee Wongbenjarat, MD,³
Sauwalak Opastirakul, MD,⁴ Siri Khoprasert, MD,⁵ Ratikorn Peuksungnern, MD,⁶
Duangrurdee Wattanasirichaigoon, MD,⁷ Wanna Thongnoppakhun, PhD,⁸
Vip Viprakasit, MD, PhD,⁹ and Pa-thai Yenchitsomanus, PhD¹⁰

Background: Mutations in the *anion exchanger 1* (AE1) gene encoding the erythroid and kidney anion (chloride-bicarbonate) exchanger 1 may result in hereditary distal renal tubular acidosis (dRTA). Hemoglobinopathies are common in Thailand. We analyzed AE1 and hemoglobin mutations in children in Thailand with dRTA to evaluate their association with clinical manifestations.

Study Design: Case series.

Setting & Participants: 17 patients were recruited from 6 referral hospitals in 4 regions of Thailand.

Predictors: AE1 mutations were detected by means of nucleotide sequence alterations. Hemoglobin E (HbE) was detected by means of hemoglobin typing, and thalassemia, by means of analysis of globin genes. Hemolytic anemia was indicated by decreased hemoglobin and hematocrit values in the presence of reticulocytosis.

Outcomes & Measurements: Leading clinical manifestations in patients were failure to thrive and muscle weakness. Compensated or overt anemia was identified in some cases. Coexistence of AE1 mutations with HbE or α^+ -thalassemia was present in a number of patients.

Results: 12 of 17 patients (70%) carried AE1 mutations, 7 patients (41%) had HbE, and 1 patient (6%) had α^+ -thalassemia. Patients with AE1 mutations presented with compensated hemolysis when they had metabolic acidosis. A patient with compound heterozygous Southeast Asian ovalocytosis/G701D and heterozygous α^+ -thalassemia showed severe hemolytic anemia.

Limitations: 5 patients (30%) without detectable AE1 mutation also were unknown for other genetic abnormalities.

Conclusions: Most of the patients with dRTA studied carried autosomal recessive AE1 mutations. Metabolic acidosis, which could be alleviated by adequate alkaline therapy, induced variable degrees of hemolysis in patients with dRTA associated with autosomal recessive AE1 mutations, especially in the presence of thalassemia.

Am J Kidney Dis 49:841-850. © 2007 by the National Kidney Foundation, Inc.

INDEX WORDS: Distal renal tubular acidosis; *anion exchanger 1*; Southeast Asian ovalocytosis; Thai.

Distal renal tubular acidosis (dRTA) is a disease characterized by an incapability of the distal nephron to secrete acid into urine,

leading to persistent metabolic acidosis.¹ Children with dRTA have hyperchloremic metabolic acidosis, metabolic bone disease, and nephrocal-

From the ¹Department of Pediatrics, Faculty of Medicine, Thammasat University; ²Khonken Hospital; ³Maharat Nakornrachsririma Hospital; ⁴Department of Pediatrics, Chaing-mai University; ⁵Surajthani Hospital; ⁶Saraburi Hospital; ⁷Division of Medical Genetics, Department of Pediatrics, Ramathibodi Hospital; ⁸Division of Molecular Genetics, Department of Research and Development; ⁹Division of Hematology/Oncology, Department of Pediatrics; and ¹⁰Division of Medical Molecular Biology and BIOTEC-Medical Biotechnology Unit, Department of Research and Development, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand.

Received December 3, 2006; accepted in revised form March 5, 2007.

Originally published online as doi:10.1053/j.ajkd.2007.03.002 on May 3, 2007.

Support: This study was supported by grants from the Thailand Research Fund (TRG4880005) to Drs Khositseth, Viprakasit, and Yenchitsomanus. In addition, Dr Wattanasirichaigoon is supported by the Ramathibodi Hospital for Research Career Development Fund, and Dr Yenchitsomanus is supported by the Pre-clinic Staff Development Fund from the Faculty of Medicine at Siriraj Hospital, Mahidol University. Potential conflicts of interest: None.

Address correspondence to Pa-thai Yenchitsomanus, PhD, Division of Medical Molecular Biology, Department of Research and Development, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand. E-mail: grpye@mahidol.ac.th

© 2007 by the National Kidney Foundation, Inc.

0272-6386/07/4906-0015\$32.00/0

doi:10.1053/j.ajkd.2007.03.002

cinosis.² Mutations of at least 3 genes responsible for hereditary dRTA include *ATP6V1B1* (MIM267300) and *ATP6V0A4* (MIM605239) encoding B1 and $\alpha 4$ subunit of H^+ ATPase, respectively, and *SLC4A1* (MIM109270) encoding *anion exchanger 1* (AE1).³⁻⁵

SLC4A1 (AE1) encodes both erythroid (eAE1) and kidney (kAE1) isoforms of *anion* (chloride-bicarbonate) *exchanger 1* (AE1) protein. eAE1 is present in red blood cell membranes, whereas kAE1 is found at the basolateral membrane of type A intercalated cells of the renal collecting duct and is indirectly involved in hydrogen secretion.^{6,7} Thus, AE1 mutations show pleiotropic effects resulting in 2 distinct disorders; namely, red blood cell abnormalities and dRTA.

eAE1 is found in erythrocyte membranes as dimer or higher oligomer. In addition to its anion exchange function, it serves as an anchor protein of the cytoskeleton network, binding to ankyrin, bands 4.1 and 4.2, and a number of cytoplasmic proteins.^{8,9} Many AE1 mutations are associated with hemolytic anemia, particularly those causing hereditary spherocytosis, but these rarely cause dRTA.⁹ Another phenotype of AE1 mutation caused by a 27-bp deletion in exon 11 and affecting red blood cell morphological characteristics is Southeast Asian ovalocytosis (SAO).

A defect of kAE1 in type A intercalated cells results in failure to establish a cell-to-lumen hydrogen gradient, leading to dRTA.⁷ Mutations of AE1 can give rise to autosomal dominant^{5,10-14} or autosomal recessive (AR)¹⁵⁻²¹ dRTA. The latter was found especially in children in Southeast Asian populations.¹⁵⁻²¹ Mutations associated with autosomal dominant dRTA are R589H, R589S, R589C, S613F, and R901X,^{5,10-14} whereas those found in AR dRTA include G701D/G701D, SAO/G701D, SAO/ Δ V850, SAO/A858D, Δ V850/ Δ V850, Δ V850/A858D, V488M/V488M, G701D/S773P, and SAO/Q759H.¹⁵⁻²¹

Interestingly, there are several reports of hemolytic anemia with dRTA in patients with compound heterozygosity of SAO/G701D, SAO/ Δ V850, SAO/A858D, and Δ V850/A858D.^{15,18} Because hemoglobinopathy (such as hemoglobin [Hb] E [HbE]: $\beta 26 \text{ Glu} \rightarrow \text{Lys}$) and thalassemia are prevalent in Southeast Asian countries where AE1 mutations with heterogeneous genotypes causing dRTA are found, it is possible that AE1 and globin gene mutations may coexist in some

individuals and potentially could increase the severity of hematologic diseases. Tanphaichitr et al¹⁶ reported xerocytic anemia in 2 siblings in Thailand with dRTA who carried double homozygosities of AE1 (G701D/G701D) mutation and β -globin (*HBB*) gene mutation (HbEE). The objectives of this study thus are to examine the causative association between AE1 mutations and dRTA in children in Thailand and investigate the effect of coexisting HbE or thalassemia on anemia severity.

METHODS

Patients

Pediatric patients with a clinically confirmed diagnosis of dRTA were recruited from 6 referral hospitals in different regions of Thailand. Medical records were retrospectively reviewed and blood samples were collected for analyses of biochemical, hematologic, and genetic profiles, after giving informed consent. The study protocol was approved by the pertinent institutional review board. Inclusion criteria were: (1) the presence of persistent hyperchloremic metabolic acidosis with serum bicarbonate level less than 18 mEq/L (mmol/L), observed on at least 2 occasions; (2) inability to acidify urine, shown by urine pH greater than 5.5 during metabolic acidosis and positive urine anion gap, and (3) absence of glucosuria. Exclusion criteria were: (1) the presence of creatinine clearance less than 60 mL/min/1.73 m² (<1 mL/s/1.73 m²) at the time of diagnosis of dRTA, and (2) the presence of glucosuria indicating Fanconi syndrome. All patients were evaluated for weight, height, serum electrolyte levels, serum calcium and phosphate levels, spot urinary calcium-creatinine ratio, and nephrocalcinosis by means of roentgenogram. Hypercalciuria is defined as a spot urinary calcium-creatinine ratio greater than the normal value for age.²² Creatinine clearance was calculated using the Schwartz formula.²³

Hematologic Studies

Hematologic studies of patients and their family members were carried out using standard techniques. Red blood cell indices and complete blood counts were analyzed using an automatic cell counter (Sysmex, Kakogawa, Japan). Hb typing was performed using an LPLC automated Hb analyzer (HB Gold; Drew Scientific Ltd, Cumbria, UK). Serum ferritin and reticulocyte levels were determined using standard methods. Mutations in the α -globin gene were identified by means of multiplex Gap-polymerase chain reaction (PCR) method for 7 deletional and 6 nondeletional α -thalassemia types reported in the Thai population (primers and conditions available on request).²⁴

Genetic Analysis

Genomic DNA was prepared from peripheral-blood leukocytes. Sequences of all coding exons for kAE1 (exons 4 to 20), as well as the kidney promoter in intron 3, were amplified by means of PCR and subjected to initial mutation

screening by denaturing high-performance liquid chromatography (DHPLC) using the Wave DNA Fragment Analysis System (Transgenomic, Crewe, UK). Fragments showing abnormal elution profiles on DHPLC were reamplified by PCR, purified, and sequenced as previously described.^{17,19} Examinations of *AE1* G701D and SAO mutations were conducted using PCR-based methods as previously described.^{17,19} For details of DHPLC and PCR-based analysis, see supplementary data with this article at www.ajkd.org.

RESULTS

Clinical and Biochemical Findings

Seventeen patients were recruited into the study. Of these, 11 patients were residents of northeastern; 2 patients, northern; 2 patients, central; and 2 patients, southern regions of Thailand. Mean age at diagnosis was 4.5 ± 3.0 (SD) years (Table 1). Leading clinical manifestations were failure to thrive (length/height less than the third percentile; 100% of patients) and muscle weakness (94%). Consanguinity was present in 1 family. Anemia and hepatosplenomegaly were identified in 13% and 0.7% of patients, respectively. Characteristics and baseline blood chemistries for each patient are listed in Table 2. All patients had serum bicarbonate levels less than 18 mEq/L (mmol/L). Nephrocalcinosis was detected by means of roentgenogram in 7 patients (41%). Ten patients (59%) had rickets diagnosed by means of physical examination and roentgenograms.

Clinical reevaluation at the time of study entry showed that 12 (80%) and 7 patients (41%) had gained body weight and height above the 25th percentile, respectively. Glomerular filtration rates

of most patients remained normal, except for 1 patient (L:II-1) who had extensive nephrocalcinosis at presentation and a decreasing creatinine clearance from 70 to 25 mL/min/1.73 m² (1.2 to 0.4 mL/s/1.73 m²) during 8 years of treatment with poor compliance.

Patients received various types of alkali therapy, including potassium citrate, sodium citrate, and sodium bicarbonate tablets. After excluding 2 patients with chronic renal failure and age older than 15 years, 15 growing patients received alkaline treatment of 3.2 ± 2.0 (1.0 to 6.6) mEq/kg/d. Urinary calcium-creatinine ratios were less than normal value for age in all patients.²² Patients also received potassium supplements at an average dose of 1.75 ± 0.8 mEq/kg/d.

Hematologic Findings

Hb and ferritin levels were normal in 15 patients (88%). Selected hematologic data after correcting iron deficiency are listed in Table 2. Macrocytosis/ovalocytosis (mean corpuscular volume > 85 fL) and microcytosis (mean corpuscular volume < 78 fL) were found in 3 and 6 patients, respectively. Six patients had HbE trait (HbAE) and 1 patient was homozygous for HbE (HbEE; Table 2). Four patients (A:II-1, E:II-1, F:II-1, and H:II-1) had reticulocytosis without anemia when they received inadequate alkaline therapy (Table 3), indicating that acidotic control affected the red blood cell phenotype, as shown in patient A:II-1, whose peripheral-blood smear before alkali therapy contained many spherocytes with numerous polychromasia (Fig 1A).

Table 1. Summary of Clinical Manifestations of Patients With dRTA

Patient Clinical Data	All Patients With dRTA	Patients With dRTA With <i>AE1</i> Mutations	Patients With dRTA Without <i>AE1</i> Mutations
No. of patients	17 (100)	12 (70)	5 (30)
Age at diagnosis (y)	4.5 ± 3.0	3.9 ± 2.9	5.4 ± 4.0
Sex (male/female)	10/7	10/2	0/5
Failure to thrive	17 (100)	12 (100)	5 (100)
Muscle weakness	16 (94)	11 (92)	5 (100)
Polyuria	11 (65)	7 (58)	4 (80)
Bone pain	6 (35)	6 (50)	0 (0)
Bowlegs/knock knee	4 (23)	4 (33)	0 (0)
Flaying of wrist	4 (23)	4 (33)	0 (0)
Rachitic rosary	4 (23)	4 (33)	0 (0)
Alkaline therapy (mEq/kg/d)	3.0 ± 2.1	3.6 ± 2.0	1.4 ± 0.6

Note: Data presented as number (percent) or mean $\pm$ SD.

Abbreviations: dRTA, distal renal tubular acidosis; *AE1*, anion exchanger 1.

Table 2. Clinical and Laboratory Data for Patients With dRTA

Patient	Sex	Age* (y)	Serum		Creatinine Clearance* (mL/1.73 m ² /min)	Treatment Duration (y)	Hemoglobin (g/dL)	Hematocrit† (%)	MCV† (fL)	Reticulocyte Count† (%)	Alkaline Therapy† (mEq/kg/d)	Hemoglobin Typing	AE1 Genotype
			Potassium* (mEq/L)	Bicarbonate* (mEq/L)									
A:II-1	M	1.5	3.1	16.1	93	3.5	13.3	36.8	73	2.1	5.3	A/A2	G701D/G701D
B:II-1	M	3.5	3.0	10.3	118	5.5	15.3	43.3	86	2.1	2.8	A/A2	G701D/G701D
C:II-1	M	2.5	2.2	9.0	113	7.1	12.6	37.5	82	3.0	7.0	A/A2	G701D/G701D
D:II-2	M	2.5	3.5	9.0	110	12.5	14.0	39.0	80	1.3	1.0	A/A2	G701D/G701D
E:II-1	M	3.5	3.2	7.0	116	5.5	14.0	39.0	69	1.0	3.0	A/E	G701D/G701D
F:II-1	M	8.5	3.8	14.0	103	0.8	14.2	40.0	70	1.2	5.0	A/E	G701D/G701D
G:II-1	M	4.0	3.0	10.5	98	8.5	13.0	38.0	74	1.2	1.0	A/E	G701D/G701D
H:II-1	M	1.5	3.2	8.0	97	17.1	12.0	37.0	91	1.0	1.6	A/A2	SAO/G701D
I:II-1	M	1.8	3.3	14.0	85	0.3	11.2	36.7	92	3.0	6.6	A/A2	SAO/G701D
J:II-2	M	9.0	2.5	13.0	90	0.5	10.3	31.0	90	1.4	4.0	A/A2	SAO/G701D
K:II-1	F	5.0	3.1	12.0	79	10.0	12.0	35.0	80	4.0	4.0	A/E	G701D/A858D
K:II-2	F	5.0	3.5	15.0	113	9.0	13.9	38.0	82	3.0	2.5	A/A2	G701D/A858D
L:II-3	F	1.0	1.3	11.0	90	12.2	13.0	37.5	81	2.0	1.0	A/A2	Normal
M:II-2	F	7.8	2.0	13.0	93	0.8	12.4	37.0	78	2.5	2.3	A/A2	Normal
N:II-1	F	1.0	2.2	14.0	70	9.0	12.3	36.4	67	1.5	1.4	A/E	Normal
O:II-2	F	8.0	1.9	9.9	100	3.8	13.0	40.5	74	2.0	1.4	A/E	Normal
P:II-1	F	9	2.9	15.0	120	1.5	11.5	31.6	57	1.5	1.0	E/E	Normal
Mean ± SD		4.5 ± 3.0	2.8 ± 0.6	11.8 ± 2.7	99.3 ± 14.3	6.4 ± 4.9	12.8 ± 1.2	37.3 ± 3.0	77.9 ± 9.3	1.98 ± 0.8	3.0 ± 2.0		

Note: Data presented as number (percent) or mean ± SD. To convert potassium and bicarbonate in mEq/L to mmol/L, multiply by 1; hemoglobin in g/dL to g/L, multiply by 10; creatinine clearance in mL/min/1.73 m² to mL/s/1.73 m², multiply by 0.01667.

Abbreviations: dRTA, distal renal tubular acidosis; MCV, mean corpuscular volume; AE1, anion exchanger 1.

*At the time of diagnosis of dRTA.

†At the time of the study.

Table 3. Hematologic Data for Selected Patients

Patients	AE1 Genotype	Hb Typing	With Metabolic Acidosis				Without Metabolic Acidosis				Urine Calcium-Creatinine Ratio
			Hb/Hct (g/dL/%)	Reticulocyte Count (%)	Serum Bicarbonate (mEq/L)	Alkaline Therapy (mEq/kg/d)	Hb/Hct (g/dL/%)	Reticulocyte Count (%)	Serum Bicarbonate (mEq/L)	Alkaline Therapy (mEq/kg/d)	
A:II-1	G701D/G701D	A/A2	12.9/36	14	14	2.9	13.0/38	2.1	25	5.3	0.04
E:II-1	G701D/G701D	A/E	13.0/36	7	19	2.5	14.0/39	1.0	25	3	0.03
F:II-1	G701D/G701D	A/E	12.3/35	9	19	2.4	14.2/40	1.2	23	5	0.1
H:II-1	SAO/G701D	A/A2	12.0/37	19	20	3.0	13.0/38	1.0	25	5	0.2
J:II-2	SAO/G701D	A/A2	8.6/24	7	9	0	10.3/31	1.4	24	4	0.1

$-\alpha^{3.7/\alpha\alpha}$

Note: To convert Hb in g/dL to g/L, multiply by 10; bicarbonate in mEq/L to mmol/L, multiply by 1.

Abbreviations: AE1, anion exchanger 1; Hb, hemoglobin; Hct, hematocrit.

Interestingly, reticulocytosis normalized after receiving an adequate dose of alkaline therapy, shown by normal serum bicarbonate levels and urinary calcium-creatinine ratios (Table 3). The peripheral-blood smear of patient A:II-1 after receiving adequate treatment showed normochromic normocytosis without evidence of hemolysis (Fig 1B). Remarkably, 1 patient (J:II-2) who presented with hemolytic anemia and hepatosplenomegaly had a normal-size liver and spleen with mild anemia after adequate alkaline therapy (Table 3).

Molecular Genetic Study

AE1 mutations were detected in 12 patients (70%). Seven patients carried the homozygous AE1 G701D/G701D mutation (Table 2), detected by means of *Hpa*II digestion of amplified exon 17 DNA fragment and DNA sequencing, showing a substitution of G by A in codon 701 (Fig 2A).

Three patients with compound heterozygous AE1 SAO/G701D mutations (Table 2) were identified by means of *Hpa*II digestion and DNA sequencing of exon 17, showing the presence of the heterozygous AE1 G701D mutation (Fig 2B), and by analysis of amplified fragment of exon 11 using agarose-gel electrophoresis and DNA sequencing that showed the presence of the heterozygous AE1 SAO mutation (Fig 2B).

Two patients carried compound heterozygous AE1 G701D/A858D mutations (Table 2), shown by means of *Hpa*II digestion and DNA sequencing of amplified exon 17 (Fig 2C). DHPLC analysis of amplified exon 19 DNA fragments from both patients showed a biphasic pattern indicating nucleotide variations. Subsequent DNA sequencing of exon 19 showed C to A transition in codon 858 resulting in the replacement of alanine by aspartate (A858D; Fig 2C). AE1 mutations could not be found in 5 patients (30%).

Clinical Manifestations of dRTA Associated With AE1 Mutations and HbE or Thalassemia

Seven patients with dRTA from the northern and northeastern regions of Thailand carried the AE1 G701D/G701D mutation. They presented with failure to thrive, polyuria ($n = 6$), and muscle weakness ($n = 6$) and received alkaline therapy (4.0 ± 2.1 mEq/kg/d). Of these 7 patients, 4 had normal Hb typing without anemia,

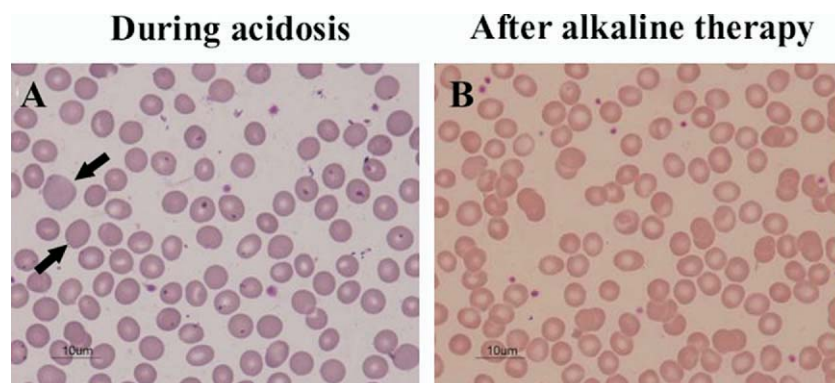


Figure 1. Peripheral-blood smears of A:II-1 during metabolic acidosis show (A) numerous spherocytes and polychromasia (arrows) and (B) red blood cell morphological characteristics were changed after alkali therapy.

with peripheral-blood smears showing prominently normal red blood cells with few ovalocytes. Three patients (E:II-1, F:II-1, and G:II-1) with HbAE (HbE ~28% to 30%) had normal Hb levels, but some hypochromia and ovalocytosis.

Three patients with dRTA from the northern, northeastern, and central parts of Thailand carried *AE1* SAO/G701D mutations. They presented with failure to thrive and muscle weakness and received alkaline therapy (4.1 ± 2.5

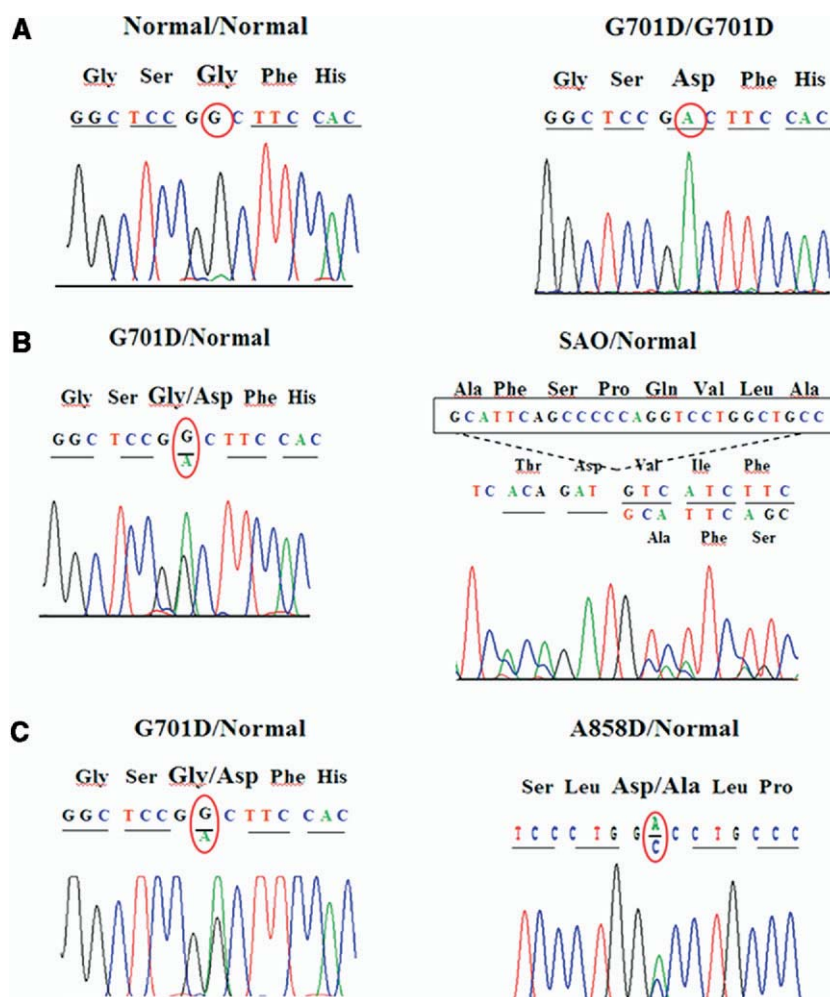


Figure 2. DNA sequencing of the *anion exchanger 1* (*AE1*) gene from 3 patients. (A) Normal sequence of exon 17 shows nucleotides GGC in (left) codon 701 and (right) a homozygous single-nucleotide substitution from G to A at the second base of codon 701 corresponding to homozygous *AE1* G701D/G701D; (B) a (left) heterozygous *AE1* G701D and a (right) 27-nucleotide deletion in codon 400 to 408 of exon 11 corresponding to heterozygous *AE1* Southeast Asian ovalocytosis (SAO) mutation, superimposing the normal sequence; and (C) a (left) heterozygous *AE1* G701D mutation and (right) a heterozygous nucleotide substitution from C to A at the second base of codon 858 in exon 19 corresponding to heterozygous *AE1* A858D mutation.