The Journal of Microbiology, September 2001, p.236-239 Vol. 39, No. 3

NOTE

Association of a Provisional New *emm* Type Opacity Factor-Negative Group A Streptococci Strain ST4529 with Septicemia

R.R. Rantty[1], M. Eshaghi[2], A.M. Ali[2], F. Jamal[3], and K. Yusoff[1]*

[1]*Department of Biochemistry and Microbiology,* [2]*Department of Biotechnology,*
[3]*Department of Clinical Laboratory Science, Universiti Putra Malaysia 43400 UPM, Serdang, Selangor, Malaysia*

(Received June 12, 2001 / Accepted August 22, 2001)

Group A *Streptococcus* strain ST4529 is a provisional new *emm* type which has been recently reported in Malaysia (Jomal, *et al.* 1999. Energ. Infect. Dis. 5, 10-14). This strain was found to be opacity factor (OF) negative with a T1 phenotype. Usually, OF negative strains with T1 phenotypes are associated with acute rheumatic fever. However, strain ST4529 was isolated from the blood of a patient with septicemia. Comparison of the deduced amino acid sequence of the mature hypervariable N-terminus of ST4529 showed only 43% identity with that of M5, the closest matched OF negative strain with a T1 phenotype. Thus, ST4529 most probably encodes a new serospecifically unique M protein which is associated with septicemia rather than pharyngitis infections. The strains with these phenotypes are very important because their sequences should be considered for developing any anti-streptococcal vaccines.

Key words: GAS, provisional *emm* type, OF-negative T1 phenotype, septicemia, ARF

The group A Streptococci (GAS) is responsible for a variety of diseases in humans worldwide such as rheumatic fever and toxic shock syndrome (4). They have been divided into two distinct groups, OF positive and OF negative strains (1) based on their ability to produce apoproteinase that causes mammalian serum to increase in opacity. They are also typed based on T and M antigenic phenotypes. GAS express a range of cell surface and extracellular products which have the potential to act as virulence factors of which the M protein which is encoded by *emm* gene is the most important one. *emm* genes are located in the *mga* regulon locus and are flanked by the *mga* and *scpA* genes. The M protein blocks antiphagocytosis *via* the alternative complement pathway (15). Based on the antigenic specificity of the M proteins, GAS can be divided into more than 100 M types, provisional types and *emm* types (5). Complete sequences of these *emm* and *emm*-like genes show that they all possess a similar overall structure while relationships between these genes vary in detail (13). It has been shown that there are significant differences in OF positive and OF negative streptococcal strains based on variation in the M protein signal region as well as C repeat regions (8, 14).

Bessen *et al.* (3) reported that isolates of most M serotypes readily fall into one of two antigenic M-associated proteins (MAP groups I and II), based on test antisera reactive with one group and not the other. These proteins can be correlated with particular properties; for example, MAP I antigens are associated with acute rheumatic fever (ARF), and MAP II antigens are associated with opacity factor (OF) production.

Here, we report the cloning and sequencing of the *emm* gene of a provisional new M type from Malaysia which is OF negative with a T1 phenotype. Interestingly, this strain was isolated from the blood of a patient with septicemia rather than from pharyngitis. We show that the *emm* gene sequence of this isolate is similar to those of other OF negative strains.

The strain ST4529 (also known as isolate D1323) was collected from the blood of a 25 year-old male Malay patient in 1996 in Kuala Lumpur, Malaysia. Initial identification of this isolate was done by haemolysis and sensitivity to bacitracin. The Lancefield group antigen was identified using "Streptex" grouping kit (Wellcome Diagnostics, UK).

* To whom correspondence should be addressed.
(Tel) 03-89489101, 3560; (Fax) 603-89430913
(E-mail) khatijah@fsas.upm.edu.my

Opacity factor was detected essentially as described by Johnson and Kaplan (10). Briefly, 10 µl of culture supernatants was mixed with 100 µl of horse serum (Gibco-BRL, USA) in a round-bottom 96-well microtitre plate. The plate was sealed and incubated overnight at 37°C. Opacity was determined by visual examination. T agglutination patterns were determined by slide agglutination with antisera obtained from Prague Streptococcal Laboratory (Czech Republic).

Streptococcal genomic DNA was prepared according to Bert *et al.* (2) and subjected to PCR using a sense primer (5-GGGGGGGGATCCATAAGGAGCATAAAAATGGCT-3') and an antisense primer (5'-GGGGGGGAATTCAGCTTAGTTTTCTTCTTTGCG-3'). These two primers, also known as 'all-M' primers, were earlier shown to be useful for amplification of the *emm* gene of OF positive and OF negative GAS strains (12).

Amplification of the *emm* gene of strain ST4547 (*emmst4547*) was carried out in a final volume of 25 µl containing 2.5 mM of $MgCl_2$, 0.2 mM of dNTPs mixture, 1 × PCR buffer [(10 mM Tris-HCl, pH 9.0 at 25°C), 50 mM KCl and 0.1% (v/v) Triton X-100; Promega, USA], 1 µM of each primer, 2.5 units of *Taq* DNA polymerase (Promega, USA), and 2 µl of template DNA (100 ng). The amplicon produced was cloned into $pCR^{R}2.1$-TOPO vector using TOPO TA Cloning Kit (Invitrogen, USA) for sequencing purposes according to the manufacturer's instructions. High Pure Plasmid Isolation Kit (Boehringer-Meinheim, Germany) was used according to the manufacturers instruction's to extract plasmid from the *Escherichia coli* TOP 10 (Invitrogen, USA). The *E. coli* was grown aerobically at 37°C on LB agar (Pronudisa, Spain) plates or in LB broth (Pronudisa, Spain) in the presence of 50 µg/ml ampicillin, with broth culture being shaken at 250 rpm in an orbital incubator.

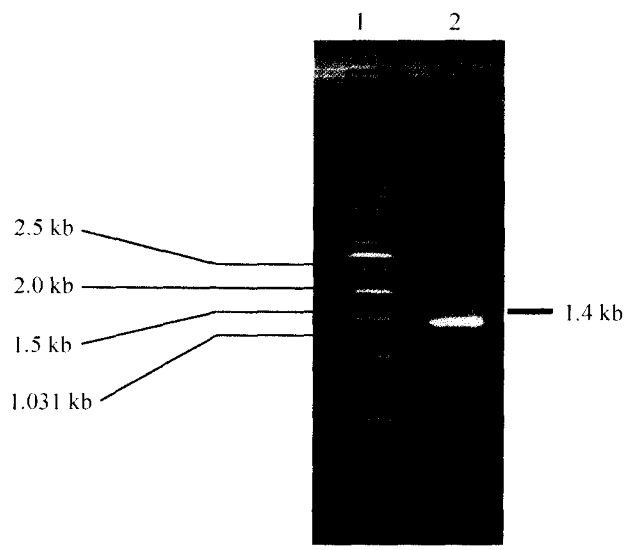

Fig. 1. PCR amplification product of *emm* gene of ST4529. PCR products were electrophoresed on a 1% ethidium bromide stained agarose gel. Lane 1, DNA Ladder marker (Fermentas, USA); Lane 2, 1.4 kb product using all M primer.

Plasmid DNA of the positive recombinants containing *emmST4529* gene was subjected to automated sequencing on the ABI sequencer. Nucleotide sequence editing, analysis, and prediction of amino acids sequences were conducted by using the Biology WorkBench at the web site address http://workbench.sdsc.edu. The nucleotide sequences of the isolates were initially identified by BLAST searches in the GenBank database.

The nucleotide sequence data presented here was submitted to GenBank under accession number AY033333.

Sequencing of the *emmst4529* gene was performed by primer walking. Basically, 7 primers were used to sequence the entire gene. The amplicon produced by Podbielski's primers (Fig. 1) which comprised the entire sequence of *emmst45429* gene was cloned into TOPO TA Cloning vector (Invitrogen, USA) and sequenced bi-directionally by primer walking.

The entire sequence of ST4529 *emm* gene is shown in Fig. 2. The start codon of *emmst4529* gene begins at nucleotide 1 and a stop codon at nucleotide 1414. It contains an open reading frame of 1416 nucleotides which encodes a precursor protein of 471 amino acids with an molecular mass of 52.6 kDa. If this predicted precursor protein was processed by cleavage between residue 42 and the following residue, then the processed mature protein would be about 48.3 kDa. The predicted *emm* gene product shows similar features to other M-like proteins with a well-conserved N-terminal region containing a predicted 42-residue signal peptide. Similarly the C-terminal half of the *emmst4529* gene product is highly homologous with conserved C-terminal regions of other M-like proteins including a proline-glycine-threonine-serine (PGTS)- rich region, an LPXTGX motif and a hydrophobic domain which is followed by a charged tail at the extreme C-terminus. The central portion of this gene possesses three 72 bp repeats (24 residues) which are designated as C repeats. Each of theses repeats are separated by a 33 nucleotide spacer region. Amino acid sequence alignment of ST4529 M protein showed a higher homology with the amino acid sequence of OF negative strains than OF positive strains (data not shown). The C repeats of the *emmst4529* gene contained the typical OF negative class-I specific amino acid (6). The deduced amino acid sequence of the mature hypervariable N terminus of ST4529 showed 43% identity to M5 protein which is the closest matched OF negative published strain. Therefore, there is a high possibility of this strain being designated as a new M type. Furthermore, these results reveal that this strain has very similar structure at the DNA level with other OF negative GAS strains.

From our typing, it was found that strain ST4529 is an OF negative and T1 phenotype. It should be noted that nearly all OF negative strains with T1 phenotypes have

	atg	gct	aga	aaa	gat	acg	aat	aaa	cag	tat	tcg	ctt	aga	aaa	tta	aaa	48
1	M	A	R	K	D	T	N	K	Q	Y	S	L	R	K	L	K	
	aaa	ggc	act	gct	tca	gta	gca	gtg	gct	ttg	agt	gta	ata	ggg	gca	gga	96
17	K	G	T	A	S	V	A	V	A	L	S	V	I	G	A	G	
	tta	gtt	gtc	aat	act	aat	gaa	gtt	agt	gca	gaa	gtg	aat	act	agg	agc	144
33	L	V	V	N	T	N	E	V	S	A	E	V	N	T	R	S	
	cgg	gca	caa	gat	gcg	ggc	tac	caa	aaa	ggc	cgt	gct	gac	aag	ctt	gag	192
49	R	A	Q	D	A	G	Y	Q	K	G	R	A	D	K	L	E	
	aca	gaa	aac	cat	ggg	tta	aaa	ttt	cag	aat	gag	aag	tta	caa	aat	cag	240
65	T	E	N	H	G	L	K	F	Q	N	E	K	L	Q	N	Q	
	aat	aat	gac	tta	aaa	act	cag	act	gct	act	tta	aca	agt	gag	aat	aaa	288
81	N	N	D	L	K	T	Q	T	A	T	L	T	S	E	N	K	
	acg	ctt	caa	gga	caa	gta	gca	gca	ggc	cag	aaa	gaa	cta	gaa	gaa	caa	336
97	T	L	Q	G	Q	V	A	A	G	Q	K	E	L	E	E	Q	
	aaa	gaa	caa	aat	aaa	gct	ctt	gaa	aaa	aaa	gca	gcg	gaa	aag	gaa	caa	384
113	K	E	Q	N	K	A	L	E	K	K	A	A	E	K	E	Q	
	gat	aat	aaa	gcg	tta	aga	caa	cgg	ggt	gat	acg	tta	ttt	aat	cag	aga	432
129	D	N	K	A	L	R	Q	R	G	D	T	L	F	N	Q	R	
	gta	aga	ctt	gaa	aaa	cag	gta	cag	gaa	aag	gaa	cac	aat	aat	aaa	acg	480
145	V	R	L	E	K	Q	V	Q	E	K	E	H	N	N	K	T	
	tta	aaa	att	gag	aat	ggt	gag	tta	aaa	act	gag	aat	ggt	gac	tta	act	528
161	L	K	I	E	N	G	E	L	K	T	E	N	G	D	L	T	
	aaa	aag	ttg	gat	gaa	act	cga	caa	gaa	tta	gca	aat	aaa	cag	caa	gag	576
177	K	K	L	D	E	T	R	Q	E	L	A	N	K	Q	Q	E	
	agt	aaa	gaa	aat	gaa	aag	acc	ctt	aat	gaa	ctc	ttg	gaa	aag	aca	gta	624
193	S	K	E	N	E	K	T	L	N	E	L	L	E	K	T	V	
	aaa	gat	aaa	att	gct	aag	gag	caa	aaa	agt	aaa	caa	gac	ttt	ggt	gcc	672
209	K	D	K	I	A	K	E	Q	K	S	K	Q	D	F	G	A	
											C1						
	ctt	gaa	caa	gaa	tta	gct	aaa	aaa	gaa	gaa	caa	aac	aag	att	tca	gac	720
225	L	E	Q	E	L	A	K	K	E	E	Q	N	K	I	S	D	
	gca	agt	cgt	caa	ggt	ctt	cgc	cgt	gac	ttg	gac	gca	tcg	cgt	gaa	gct	768
241	A	S	R	Q	G	L	R	R	D	L	D	A	S	R	E	A	
														C2			
	aag	aaa	caa	tta	gaa	gct	gaa	cac	caa	aaa	ctt	gaa	gaa	caa	aac	aag	816
257	K	K	Q	L	E	A	E	H	Q	K	L	E	E	Q	N	K	
	atc	tca	gaa	gca	agc	cgc	aaa	ggc	ctt	cgc	cgt	gac	ttg	gac	gca	tcg	864
273	I	S	E	A	S	R	K	G	L	R	R	D	L	D	A	S	
	cgt	gaa	gct	aag	aaa	caa	tta	gaa	gct	gaa	cac	caa	aaa	ctt	gaa	gaa	912
289	R	E	A	K	K	Q	L	E	A	E	H	Q	K	L	E	E	
	C3																
	caa	aac	aag	atc	tca	gaa	gca	agc	cgc	aaa	ggc	ctt	cgc	cgt	gac	ttg	960
305	Q	N	K	I	S	E	A	S	R	K	G	L	R	R	D	L	
	gac	gca	tca	cgt	gaa	gct	aag	aaa	caa	gtt	gaa	aaa	gct	tta	gaa	gaa	1008
321	D	A	S	R	E	A	K	K	Q	V	E	K	A	L	E	E	
	gca	aac	agc	aaa	tta	gct	gct	ctt	gaa	aat	ctt	aac	aaa	gag	ctt	gaa	1056
337	A	N	S	K	L	A	A	L	E	N	L	N	K	E	L	E	
	gaa	agc	atg	aaa	tta	aca	gaa	aaa	gaa	aaa	gct	gag	cta	caa	gca	aaa	1104
353	E	S	M	K	L	T	E	K	E	K	A	E	L	Q	A	K	
	ctt	gaa	gca	gaa	gca	aaa	gca	ctc	aaa	gaa	caa	tta	gcg	aaa	caa	gct	1152
369	L	E	A	E	A	K	A	L	K	E	Q	L	A	K	Q	A	
	gaa	gaa	ctt	gca	aaa	cta	aga	gct	gga	aaa	gca	tca	gac	tca	caa	acc	1200
385	E	E	L	A	K	L	R	A	G	K	A	S	D	S	Q	T	
	cct	gat	gca	aaa	cca	gga	aac	aaa	gct	gtt	cca	ggt	aaa	gct	caa	gca	1248
401	P	D	A	K	P	G	N	K	A	V	P	G	K	A	Q	A	
	cca	caa	gca	ggt	aca	aaa	cct	aac	caa	aac	aaa	gca	cca	atg	aag	gaa	1296
417	P	Q	A	G	T	K	P	N	Q	N	K	A	P	M	K	E	
	act	aag	aga	cag	tta	cca	tca	aca	ggt	gaa	aca	gct	aac	cca	ttc	ttc	1344
433	T	K	R	Q	L	P	S	T	G	E	T	A	N	P	F	F	
	ggt	gaa	aca	gct	aac	cca	ttc	ttc	ggt	gaa	aca	gct	gta	gca	gca	gtt	1392
449	G	E	T	A	N	P	F	F	G	E	T	A	V	A	A	V	
	gta	aaa	cgc	aaa	gaa	gaa	aac	taa									1416
465	V	K	R	K	E	E	N	·									

Fig. 2. Nucleotide sequence of *emm* gene of ST4529 and the deduced amino acid sequence. The repeated regions described in the text are boxed in different repeat units.

been reported to be associated with ARF (3, 7, 8, 11). However, in our study ST4529 strain was isolated from the blood of a patient with septicemia.

In conclusion, ST4529 is a provisional new *emm* type. At the DNA level, this strain had a structure very similar to other OF negative GAS strains. Our results also suggest the possibility of OF negative T1 GAS strains being associated with septicemia.

Acknowledgments

This work has been supported by the Ministry of Environmental Science and Technology of Malaysia (IRPA Grant No. 06-02-04-0052) and Japan International Cooperation Agency (JICA After Care Project for the Department of Biotechnology).

References

1. Beall, B., G. Gherardi, M. Lovgren, R.R. Facklam, B.A. Forwick, and G.J. Tyrrell. 2000. *emm* and *sof* gene sequence variation in relation to serological typing of opacity-factor-positive group A streptococci. *Microbiol.* 146, 1195-1209.
2. Bert, F., B. Picard, C. Branger, and N. Lambert-Zechovsky. 1996. Analysis of genetic relationships among strains of group A, C and G streptococci by random amplified polymorhic DNA analysis. *J. Med. Microbiol.* 45, 278-284.
3. Bessen, D., K.F. Jones, and V.A. Fischetti. 1989. Evidence for two distinct classes of streptococcal M protein and their relationship to rheumtic fever. *J. Exp. Med.* 169, 269-283.
4. Bisno, A.L. 1991. Group A streptococcal infections and acute rheumatic fever. *New England Medical* 325, 783-793.
5. Facklam, R., B. Beall, A. Efstratiou, V. Fischetti, D. Johnson, E. Kaplan, P. Kriz, M. Lovgren, D. Martin, B. Schwartz, A. Totolian, D. Bessen, S. Hollingshead, F. Rubbin, J. Scott, and G. Tyrrell. 1999. *emm* typing and validation of provisional M types for group A streptococci. *Emerg. Infect. Dis.* 5, 1-9.
6. Fischetti, V.A., V. Pancholi, and O. Schneewind. 1990. Conservation of hexapeptide sequence in the anchor region of surface protein from gram-positive cocci. *Mol. Microbiol.* 4, 1603-1605.
7. Gardiner, D.L. and K.S. Spiprakash. 1996. Molecular epidemiology of impetiginous group A streptococcal infections in Aboriginal communities of Northen Australia. *J. Clinic. Mirobiol.* 34, 1448-1452.
8. Haanes, E.J., D.G. Heath, and P.P. Cleary. 1992. Architecture of the Vir regulons of group A streptococci parallels opacity factor phenotype and M protein class. *J. Bacteriol.* 174, 4967-4976.
9. Jamal, F., S. Pit, R. Facklam, and B. Beall. 1999. New *emm* (M protein gene) sequences of group A streptococci isolated from Malaysian patients. *Emerg. Infect. Dis.* 5, 10-14.
10. Johnson, D.R. and E.L. Kaplan. 1988. Microtechnique for serum opacity Factor characterization of group A streptococci adaptable to the use of human sera. *J. Clin. Microbiol.* 26, 2025-2030.
11. Norgren, M., A. Norrby, and S.E. Holm. 1992. Genetic diversity in T1M1 group A streptococci in relation to clinical outcome of infection. *J. Infect. Dis.* 166, 1014-1020.
12. Podbielski, A., B. Melzer, and R. Lutticken. 1991. Application of the polymerase chain reaction to study the M protein(-like) gene family in beta-haemolytic streptococci. *Med. Microbiol. Immunol.* 180, 213-227.
13. Whatmore, A.M. and M.A. Kehoe. 1994. Horizontal gene transfer in the evolution of group A streptococcal *emm*-like genes: gene mosaics and variation in Vir regulons. *Mol. Microbiol.* 11, 363-374.
14. Whatmore, A.M., V. Kapur, D.J. Sullivan, J.M. Musser, and M.A. Kehoe. 1994. Non-congruent relationships between variation in *emm* gene sequences and the population genetic structure of group A streptococci. *Mol. Microbiol.* 14, 619-631.
15. Whitnack, E. and E.H. Beachey. 1985. Inhibition of complement-mediated opsonisation and phagocytosis of *Streptococcus pyogenes* by D fragments of fibrinogen and fibrin bound to cell surface M protein. *J. Exp. Med.* 162, 1983-1997.