

Clonal Distribution and Phase-Variable Expression of a Major Histocompatibility Complex Analogue Protein in *Staphylococcus aureus*

Angus Buckling,¹ James Neilson,² Jodi Lindsay,³ Richard French-Constant,² Mark Enright,² Nicholas Day,⁴ and Ruth C. Massey^{1*}

Department of Zoology, University of Oxford, Oxford,¹ Department of Biology and Biochemistry, University of Bath, Bath,² and Cellular and Molecular Medicine, St. George's Medical School, London,³ United Kingdom, and Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand⁴

Received 17 November 2004/Accepted 13 January 2005

The *mapW* gene of *Staphylococcus aureus* strain N315 contains a poly(A) tract which truncates translation of the protein. This study demonstrates that *mapW* is an allelic variant of the *map/eap* genes found in other strains and that the variation in the length of this poly(A) tract suggests that it is a contingency locus.

Phase variation is one of the many strategies employed by pathogenic bacteria to avoid detection by a host's immune systems and involves the ability to switch on the expression of proteins when they are needed and to switch it off when they are likely to trigger immune responses (5). Several mechanisms have evolved to accomplish this strategy, one of which is known as slip-strand mispairing, where the lengths of long-repeat tracts of nucleotides vary (11, 13). This mechanism can have the effect of either altering the binding efficiency of transcriptional machinery to promoter regions or shifting the triplet coding out of frame.

Staphylococcus aureus is a major human pathogen causing severe community- and hospital-acquired infections (2, 15). This study examines the MapW protein of *S. aureus*, identified when the first genome, strain N315, was sequenced (gene reference, SA1751 [9]). The gene encoding this protein contains a tract of ten adenine nucleotides [poly(A) tract] within the coding region which introduces a stop codon immediately after the tract. This process shifts translation of the triplet-coding sequence out of frame for the final section of the protein (9), where 8 or 11 adenines would result in the full-length translation of the protein, and in so doing, is reminiscent of the slip-strand mispairing regions described above.

The MapW protein of strain N315 is 69% identical at the amino acid level to a protein called Map from *S. aureus* strain FDA574, which has several analogues identified in different strains of *S. aureus* (e.g., Eap from strain Newman) (8, 12). These proteins contain regions with high degrees of identity to major histocompatibility complex (MHC) class II molecules (8) and have been shown to act as extracellular matrix-binding proteins, invasins, and modulators of T-cell response to infection, facilitating immune evasion (4, 10). To examine the diversity of these proteins, a phylogenetic (neighbor-joining) tree was constructed (Fig. 1) using the predicted amino acid se-

quences of the MapW protein of strain N315, Map of strain FDA574, Eap of strain Newman, MapW of strain Mu50, MapW of strain MW2, and the homologues in sequenced strains MRSA252 and MSSA476 (gene reference, SAR2030 [6]). Two distinct groups, consisting of (i) MapW (of strains N315, MW2, and Mu50) and the homologue in MSSA476 and (ii) Map of FDA574, Eap of Newman, and the homologue in strain MRSA252, can be seen. This grouping illustrates that the Map/Eap/MapW proteins can be separated into two distinct lineages, suggesting the Map/Eap and MapW proteins are allelic variants of one another.

A previous study examining the distribution of the *map/eap* genes found them to be present in 97.6% of isolates examined but did not differentiate between *map/eap* and *mapW* (7). To examine the distribution of each allele in a natural population of *S. aureus*, a large, well-characterized strain collection (312 strains) isolated from hospital-acquired infections, community-acquired infections, and healthy colonized volunteers was surveyed. Four pairs of specific oligonucleotide primers were designed for both the 3' and 5' ends of the *mapW* gene from *S. aureus* strain N315 and for those of the *map* gene from strain FDA574 (Map3' forward [F], GAGGTATCAATGATATTG AATTG; Map3' reverse [R], GTAACGCTAGTCTAAGTT TAG; Map5' F, CTGTGAATGGTACAAGCCAAAA; Map5' R, TCCAAATCACGTTCACTAACGA; MapW3' F, CCTTG ATCATTGGCCATTGC; MapW3' R, GAGGTATCAGTGA TCTTGAC; MapW5' F, TTTAAATCTTGTTCACTAATACC; and MapW5' R, CTGTGAACGGTACAAGCCAAAA). These primers were used in PCRs to determine the presence of both ends of both genes in each of 312 *S. aureus* isolates. Genomic DNA was isolated as described previously (14), the PCRs were performed using a PTC-200 DNA engine with Taq polymerase (Promega) and 2 mM MgCl₂, and the products were analyzed by 0.8% agarose gel electrophoresis. The vast majority of strains in the collection (307 of 312 strains) had either one or the other allele, as determined by a positive PCR for both pairs of primers for the given allele. Fifty-four percent of the strains in the collection were found to contain the *mapW* allele, and 45% contained the *map/eap* allele. This finding concurs with the results of the

* Corresponding author. Mailing address: Department of Zoology, University of Oxford, South Parks Rd., Oxford OX1 3PS, United Kingdom. Phone: 44-1865-271288. Fax: 44-1865-310447. E-mail: ruth.massey@zoo.ox.ac.uk.

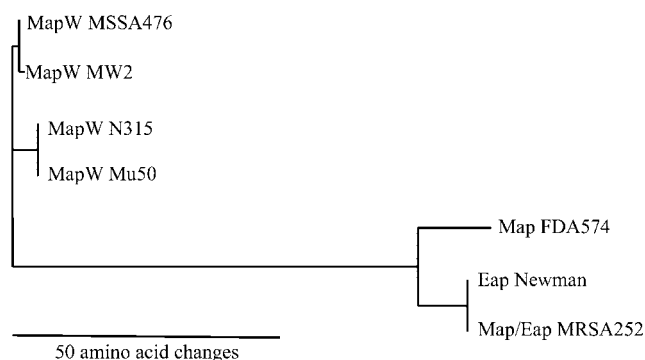


FIG. 1. Relationship between the Map-like proteins in several *S. aureus* strains. This neighbor-joining tree illustrates that there are two distinct groups of Map-like proteins. The proteins in strains N315, Mu50, MW2, and MSSA476 all have the poly(A) tract. The proteins in strains FDA574, Newman, and MRSA252 do not contain a poly(A) tract. All the proteins were truncated to the length of the shortest protein for alignment and clustering.

phylogenetic tree, suggesting that two major alleles of this gene exist in natural populations of *S. aureus*. These PCR results were verified by hybridizing a subset of these strains (161 strains) to the *S. aureus* DNA microarray constructed by the B μ GS group at St. George's Medical School, London, United Kingdom.

The *S. aureus* strain collection used in this study has been characterized by multilocus sequence typing, which groups the isolates into sequence types and clonal complexes (CCs) (3). The distribution of the two alleles within these groups was examined where the presence of either allele was found to correlate exactly with sequence type and CC (Table 1). All isolates of CC30, CC39, and CC45 were positive for the *map/eap* allele, and all isolates from the other CC groups were positive for the *mapW* allele.

The number of adenines in the poly(A) tracts of the six sequenced strains of *S. aureus* found to contain the *mapW* gene was examined to determine whether there was any variation in the lengths of the poly(A) tracts. The *mapW* genes of strains N315, MW2, and Mu50 contained 10 adenines, while those of Col, 8325, and MSSA476 each contained 9. The effect of having 9 adenines is equivalent to that of having 10 in that a stop codon is introduced shortly after the poly(A) tract and the translational-triplet coding remains out of frame. To examine

TABLE 1. Distribution of *mapW* and *map/eap* alleles within CCs (defined by multilocus sequence typing)

CC	No. of isolates	Allele
1	16	<i>mapW</i>
5	17	<i>mapW</i>
8	12	<i>mapW</i>
9	6	<i>mapW</i>
12	11	<i>mapW</i>
15	29	<i>mapW</i>
16	21	<i>mapW</i>
22	23	<i>mapW</i>
25	23	<i>mapW</i>
30	87	<i>map/eap</i>
39	24	<i>map/eap</i>
45	29	<i>map/eap</i>
51	9	<i>mapW</i>

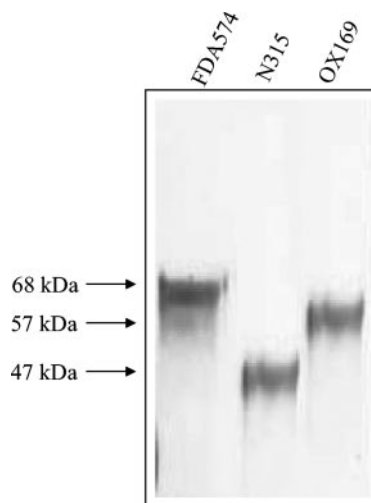


FIG. 2. Secretion of Map and MapW proteins in long and short forms. A Western immunoblot of secreted proteins from strain FDA574 (expressing the Map protein), strain N315 (expressing the short form of MapW), and strain OX169 [predicted to express the longer form of MapW due to the presence of eight adenines in the poly(A) tract] is shown.

whether many strains of *S. aureus* express the full-length protein, the lengths of the poly(A) tracts in 45 randomly selected strains positive for *mapW* by PCR were examined by using the forward and reverse primers *MapW* 3' F and R described above. The amplified products were precipitated with 20% polyethylene glycol and 2.5 M NaCl and washed with 70% ethanol. The sequences of both strands were determined with BigDye fluorescent terminators and the primers used in the initial PCR amplification. A large amount of variation was observed, with 44% of strains containing 9 adenines, 36% containing 10 adenines, 18% containing 8 adenines, and 2% containing 11 adenines. This variation was independent of CC (logistic regression of association between long or short forms of protein and CC; data not overdispersed [1]; $\chi^2_7 = 1.58$; $P > 0.3$). These data suggest that variation occurs at this poly(A) tract, resulting in variation in the lengths of the MapW proteins expressed.

To verify that the MapW protein is expressed in both the long and short forms in a similar manner to the Map protein, a Western immunoblot was performed as described previously (12) using antisera raised against a recombinant form of the Map protein of strain FDA574 (a generous gift from Eric Brown, Texas A&M University). Secreted proteins were harvested from *S. aureus* strains FDA574 and N315 and from one of the *mapW* isolates from the strain collection that had 8 adenines in the poly(A) tract (and thus was predicted to allow full-length translation of the MapW protein). In Fig. 2 we can see that the anti-Map antisera react with a protein in the strains predicted to express the MapW protein. A band of approximately 68 kDa can be seen in the extracts of strain FDA574, a band of approximately 47 kDa can be seen in strain N315, and one of approximately 57 kDa can be seen in the strain predicted to express the full-length MapW protein. This visualization demonstrates that the MapW protein is secreted in a similar manner to the Map/Eap proteins and that the

length of the poly(A) tract does affect the size of the MapW protein expressed.

In summary, our data suggest that Map/Eap and MapW are allelic variants of each other and that the *mapW* gene is a phase-variable gene that can express both long and short forms of the MapW protein. This is the first putative phase variation mechanism that makes use of slip-strand mispairing identified in *S. aureus* and the first of any phase variation mechanism that results in variation in the length of the protein expressed rather than switching it off and on. Further work is under way to determine both the role this phase variation plays in the pathogenicity of this organism and its adaptive significance.

REFERENCES

1. **Crawley, M. J.** 1993. GLIM for ecologists. Blackwell Science, Oxford, United Kingdom.
2. **Emori, T. G., and R. P. Gaynes.** 1993. An overview of nosocomial infections, including the role of the microbiology laboratory. *Clin. Microbiol. Rev.* **6**:428–442.
3. **Feil, E. J., J. E. Cooper, H. Grundmann, D. A. Robinson, M. C. Enright, T. Berendt, S. J. Peacock, J. M. Smith, M. Murphy, B. G. Spratt, C. E. Moore, and N. P. Day.** 2003. How clonal is *Staphylococcus aureus*? *J. Bacteriol.* **185**:3307–3316.
4. **Harraghy, N., M. Hussain, A. Haggart, T. Chavakis, B. Sinha, M. Herrmann, and J. I. Flock.** 2003. The adhesive and immunomodulating properties of the multifunctional *Staphylococcus aureus* protein Eap. *Microbiology* **149**:2701–2707.
5. **Henderson, I. R., P. Owen, and J. P. Nataro.** 1999. Molecular switches—the ON and OFF of bacterial phase variation. *Mol. Microbiol.* **33**:919–932.
6. **Holden, M. T., et al.** 2004. Complete genomes of two clinical *Staphylococcus aureus* strains: evidence for the rapid evolution of virulence and drug resistance. *Proc. Natl. Acad. Sci. USA* **101**:9786–9791.
7. **Hussain, M., K. Becker, C. von Eiff, G. Peters, and M. Herrmann.** 2001. Analogs of Eap protein are conserved and prevalent in clinical *Staphylococcus aureus* isolates. *Clin. Diagn. Lab. Immunol.* **8**:1271–1276.
8. **Jonsson, K., D. McDevitt, M. H. McGavin, J. M. Patti, and M. Höök.** 1995. *Staphylococcus aureus* expresses a major histocompatibility complex class II analog. *J. Biol. Chem.* **270**:21457–21460.
9. **Kuroda, M., et al.** 2001. Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*. *Lancet* **357**:1225–1240.
10. **Lee, L. Y., Y. J. Miyamoto, B. W. McIntyre, M. Höök, K. W. McCrea, D. McDevitt, and E. L. Brown.** 2002. The *Staphylococcus aureus* Map protein is an immunomodulator that interferes with T cell-mediated responses. *J. Clin. Investig.* **110**:1461–1471.
11. **Martin, P., T. van de Ven, N. Mouchel, A. C. Jeffries, D. W. Hood, and E. R. Moxon.** 2003. Experimentally revised repertoire of putative contingency loci in *Neisseria meningitidis* strain MC58: evidence for a novel mechanism of phase variation. *Mol. Microbiol.* **50**:245–257.
12. **McGavin, M. H., D. Krajewska-Pietrasik, C. Ryden, and M. Höök.** 1993. Identification of a *Staphylococcus aureus* extracellular matrix-binding protein with broad specificity. *Infect. Immun.* **61**:2479–2485.
13. **Moxon, E. R., P. B. Rainey, M. A. Nowak, and R. E. Lenski.** 1994. Adaptive evolution of highly mutable loci in pathogenic bacteria. *Curr. Biol.* **4**:24–33.
14. **Peacock, S. J., C. E. Moore, A. Justice, M. Kantzanou, L. Story, K. Mackie, G. O'Neill, and N. P. Day.** 2002. Virulent combinations of adhesin and toxin genes in natural populations of *Staphylococcus aureus*. *Infect. Immun.* **70**:4987–4996.
15. **Steinberg, J. P., C. C. Clark, and B. O. Hackman.** 1996. Nosocomial and community-acquired *Staphylococcus aureus* bacteremias from 1980 to 1993: impact of intravascular devices and methicillin resistance. *Clin. Infect. Dis.* **23**:255–259.