Journal of Medical Microbiology (2005), **54**, 113–117 DOI 10.1099/jmm.0.45810-0

Revised nomenclature of *Clostridium difficile* toxins and associated genes

Maja Rupnik,[1] Bruno Dupuy,[2] Neil F. Fairweather,[3] Dale N. Gerding,[4,5] Stuart Johnson,[4,5] Ingo Just,[6] David M. Lyerly,[7] Michel R. Popoff,[8] Julian I. Rood,[9] Abraham L. Sonenshein,[10] Monica Thelestam,[11] Brendan W. Wren,[12] Tracy D. Wilkins[7] and Christoph von Eichel-Streiber[13]

Correspondence
Maja Rupnik
maja.rupnik@bf.uni-lj.si

[1]University of Ljubljana, Department of Biology, Vecna pot 111, 1000 Ljubljana, Slovenia

[2,8]Unité de Génétique Moléculaire Bactérienne[2] and Unité des Toxines Microbiennes[8], Institut Pasteur, Paris, France

[3]Centre for Molecular Microbiology and Infection, Imperial College, London, UK

[4]Hines VA Hospital, Hospital/151 Fifth Avenue & Roosevelt Road, Hines, IL 60141, USA

[5]Loyola University Medical Center, 2160 South First Avenue, Maywood, IL 60153, USA

[6]Institute of Toxicology, Hannover Medical School, Hannover, Germany

[7]TechLab Inc., Blacksburg, VA, USA

[9]Australian Bacterial Pathogenesis Program, Department of Microbiology, Monash University, Victoria 3800, Australia

[10]Department of Molecular Biology and Microbiology, Tufts University School of Medicine, Boston, MA, USA

[11]Microbiology and Tumorbiology Center, Karolinska Institute, Stockholm, Sweden

[12]London School of Hygiene and Tropical Medicine, University of London, UK

[13]Institute of Medical Microbiology and Hygiene, Johannes Gutenberg University of Mainz, D-55101 Mainz, Germany

Received 3 July 2004
Accepted 6 August 2004

Several different nomenclatures have been applied to the *Clostridium difficile* toxins and their associated genes. This paper summarizes the new nomenclature that has been agreed to by the research groups currently active in the field. The revised nomenclature includes *C. difficile* toxins and other related large clostridial toxins produced by *Clostridium sordellii* and *Clostridium novyi*, and corresponding toxin genes, as well as toxin production types of *C. difficile* strains.

INTRODUCTION

Clostridium difficile is the causative agent of antibiotic-associated diarrhoea and pseudomembranous colitis. Two large protein cytotoxins, toxin A (TcdA) and toxin B (TcdB), are recognized as the main virulence factors. Some *C. difficile* strains also produce a third, unrelated toxin (binary toxin CDT).

The molecular analysis of *C. difficile* toxins started in the 1980s with the first attempts to clone fragments of toxin genes and has progressed with the sequencing of both toxin genes (von Eichel-Streiber & Sauerborn, 1990; von Eichel-Streiber *et al.*, 1990, 1992; Dove *et al.*, 1990; Johnson *et al.*, 1990), with the definition of the region encoding the toxins (Hammond & Johnson, 1995; Braun *et al.*, 1996), and with studies of toxin gene regulation (Moncrief *et al.*, 1997; Hammond *et al.*, 1997; Hundsberger *et al.*, 1997; Dupuy & Sonenshein, 1998; Mani & Dupuy, 2001). Unfortunately, in parallel with developments on the molecular biology and biochemistry of *C. difficile*, several different nomenclature systems have been applied to the toxins and their associated genes (Table 1).

With an increasing number of research groups working on the molecular biology of *C. difficile*, or using its toxins as tools in cell biology, and with the imminent finalization of the first

This paper was presented at the round-table discussion on *Clostridium difficile* nomenclature at the First International *Clostridium difficile* Symposium, Kranjska Gora, Slovenia, 5–7 May 2004.

Abbreviation: LCT, large clostridial toxin.

Table 1. Names previously used for *C. difficile* toxins and toxin genes (see also Fig. 1)

From references: Braun *et al.* (1996); Dupuy & Sonenshein (1998); Fiorentini *et al.* (1998); Green *et al.* (1995); Hammond & Johnson (1995); Hammond *et al.* (1997); Hofmann *et al.* (1995); Johnson *et al.* (1990); Dove *et al.* (1990); Karlsson *et al.* (2003); Mani & Dupuy (2001); Moncrief *et al.* (1997); Spigaglia & Mastrantonio (2002); von Eichel-Streiber *et al.* (1990, 1992).

Genes of the *C. difficile* toxin coding region (PaLoc)				
tcdD	*tcdB*	*tcdE*	*tcdA*	*tcdC*
txeR	*toxB*	*utxA*	*toxA*	*dtxA*
ORFtxe1	Toxin B	*ORFtxe2*	Toxin A	*ORFtxe3*
		txe2		*txe3*
Names used for the gene products				
TcdD	TcdB	TcdE	TcdA	TcdC
TxeR	Toxin B		Toxin A	
	CdB		CdA	
	Cd Cyt			

C. difficile genome sequence (http://www.sanger.ac.uk/Projects/C_difficile/), the need for a unified nomenclature has become apparent. A similar approach was taken on naming the clostridial neurotoxins, where a proposed unified nomenclature (Niemann, 1992) has resulted in consistent and logical citation in the literature.

At the recent First International *C. difficile* Symposium (FICDS, Kranjska Gora, Slovenia, May 2004), a round-table discussion on *C. difficile* toxins and toxin gene nomenclature was held. This paper summarizes the new nomenclature that was agreed to by the research groups currently active in the field.

REVISED NOMENCLATURE

Large clostridial toxins (LCTs) and toxin genes

C. difficile produces three toxins: toxin A (TcdA), toxin B (TcdB) and binary toxin (CDT). Toxins A and B are similar to one another, and belong, together with three other toxins produced by *Clostridium sordellii* and *Clostridium novyi*, to the well-defined group of LCTs (von Eichel-Streiber *et al.*, 1996). The structural genes encoding four of the five LCTs have been sequenced (von Eichel-Streiber & Sauerborn, 1990; von Eichel-Streiber *et al.*, 1990, 1992; Dove *et al.*, 1990; Johnson *et al.*, 1990; Green *et al.*, 1995; Hofmann *et al.*, 1995), and in 1995 a nomenclature for LCTs and their genes was proposed that combines information on the bacterial species and the identity of the toxin (Hofmann *et al.*, 1995): Tcd for *C. difficile*, Tcs for *C. sordellii* and Tcn for *C. novyi*; and TcdA and TcdB for toxin A and toxin B, TcsL and TcsH for toxin LT and HT, and Tcnα (Tcn-alpha) for the alpha toxin of *C. novyi*.

At the Slovenia meeting it was agreed that a unified nomenclature should be based on the *tcd–tcs–tcn* system (Table 2) as it is recommended that homologous genes present in different organisms receive the same or similar names (Demerec *et al.*, 1966).

The genomic region encoding the *C. difficile* toxins

A 19 kb region encoding two related toxins produced by *C. difficile* (toxin A or TcdA and toxin B or TcdB) has been defined as the toxigenic element (Hammond & Johnson, 1995) or as the pathogenicity locus, PaLoc (Braun *et al.*, 1996). As the latter name has been used predominately in recent publications, we suggest accepting it for the new *C. difficile* nomenclature.

However, it is important to stress that although the use of the term 'pathogenicity locus' or 'PaLoc' is correct, the term 'pathogenicity island' is not an accurate description of this gene region since there is no evidence that the PaLoc region is horizontally acquired and it does not fit the generally accepted definition of a pathogenicity island (Hacker *et al.*, 2004).

C. difficile LCT genes and associated genes

For the *C. difficile* genes found in the PaLoc the nomenclature shown in Table 2 and Fig. 1 was agreed. The genes for toxins A and B should be called *tcdA* and *tcdB*, respectively. The other three genes were previously defined as *tcdC*, *D* and *E* by the von Eichel-Streiber group according to the predicted sizes of the encoded proteins and not according to their relative positions in the PaLoc (Braun *et al.*, 1996). As the names *tcdE* and *tcdC* have been used in recent publications by several different groups, these two names will remain.

However, the majority of papers dealing with the *tcdD* gene use the name *txeR*. To be consistent with the *tcd* nomenclature, which indicates that the gene is a part of the PaLoc, but to show that the gene product is involved in regulation, we have decided to rename this gene *tcdR*.

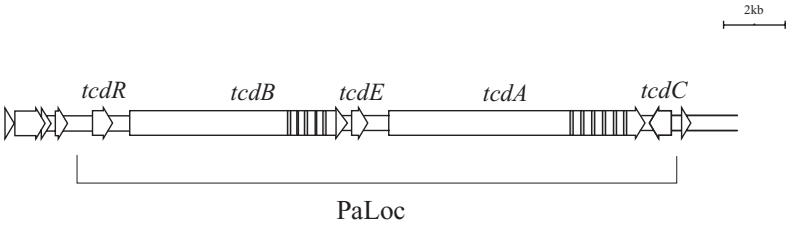

Fig. 1. Revised nomenclature for five genes of the *C. difficile* pathogenicity locus PaLoc.

Table 2. Revised unified nomenclature for *C. difficile* and related large clostridial toxins

Proposed name to be used	Synonyms (not to be used further)	Description
Genetic locus		
PaLoc (pathogenicity locus)	Toxinogenic element	19 kb well-defined region consisting of five genes including the genes encoding *C. difficile* toxins A and B
Gene names		
tcdR	*txeR*, *tcdD*	Encodes an alternative sigma factor that is a positive regulator of toxin production
tcdB	*toxB*	Gene encoding toxin B
tcdE	*txe2*	Gene encoding holin-like protein
tcdA	*toxA*	Gene encoding toxin A
tcdC	*txe3*	Putative negative regulator of toxin production
$tcdB_{1470}$		Add strain number in subscript to indicate gene from the strain 1470
Protein names		
Toxin A or TcdA	ToxA	
Toxin B or TcdB	ToxB	
$TcdB_{1470}$	TcdB-1470	Add strain number in subscript to indicate strain 1470 as origin of toxin
Toxin production*		
A^+B^+, A^-B^-, A^-B^+	A+B+, A−B−, A−B+ A−ve B+ve	Describes toxin production of *C. difficile* strain, if only LCT group (TcdA and TcdB) was tested
$A^+B^+CDT^-$, $A^+B^+CDT^+$, etc.	A+B+CDT−, A+B+CDT+	Describes toxin production of *C. difficile* strain, if all toxins (TcdA, TcdB and binary toxin CDT) were tested
$A^+B^+(CDT^+)$		Designation is included in parentheses if toxin production is inferred from genetic testing only. Additional textual comments should indicate the basis for such designation.
Other large clostridial toxins (gene name; protein name)		
tcsL; toxin LT or TcsL		*C. sordellii* toxin genes and toxins
tcsH; toxin HT or TcsH		
tcnA; alpha toxin or TcnA	*tcnα*; Tcnα	*C. novyi* toxin gene and toxin

*Superscript is preferred for description of phenotype (Demerec *et al.*, 1966).

C. difficile toxins

Various names have been used for *C. difficile* toxins in recent years (Table 1), but 'toxin A' and 'toxin B' are the most prevalent. After the first establishment of the *tcd–tcs–tcn* nomenclature, the toxins were also called 'TcdA' or 'TcdB'. At the round table it was agreed that both prefixes, 'Tcd' and 'toxin', could be used for the toxins (Table 2).

Designation of strain-specific (variant) toxins and corresponding genes

Many *C. difficile* strains are known to have toxin genes with marked sequence changes in comparison to the toxin genes from the reference strain VPI 10463, resulting in changed properties of the resultant toxins. If applicable, the strain identification should therefore be added to the toxin or to the

toxin gene. However, the number after the locus (e.g. '*tcdB*-1470') is used as a designation of *mutation site* in standard genetic nomenclature (Demerec *et al.*, 1966). To distinguish between the same genes from different strains the use of subscripts was recommended. Therefore, the designation '$tcdB_{1470}$' for the gene and '$TcdB_{1470}$' or 'toxin B_{1470}' for the protein should be used, when appropriate.

Other LCTs and corresponding genes

As already mentioned, the *tcd–tcs–tcn* nomenclature applies to the entire group of related LCTs. Only one change was suggested by the nomenclature discussion group. According to standard rules (Demerec *et al.*, 1966), genes should have a three-letter designation followed by a Latin capital letter (not Greek). Therefore the *C. novyi* alpha toxin gene was renamed *tcnA* (Table 2).

Binary toxin

The third known toxin produced by *C. difficile*, the binary toxin CDT (Perelle *et al.*, 1997), is not related to TcdA and TcdB and the nomenclature of this toxin and the corresponding genes was not discussed. The names currently used are '*C. difficile* binary toxin' or 'binary toxin CDT'. The term 'CDT' should not be used alone as it could be confused with cytolethal distending toxins produced by several other bacteria. The genes are already designated *cdtB* for the binding component and *cdtA* for the enzymic component, and we suggest that these gene names be retained.

Strains – description of toxin production

In early studies, *C. difficile* strains always produced either both toxins, TcdA and TcdB, or neither of them, making the discrimination between toxigenic and nontoxigenic strains easy. During the last decade, however, this definition has become unclear for two reasons. First, strains producing only one of the toxins have been described and have been repeatedly isolated all over the world (A^-B^+ or 'toxin A-negative, toxin B-positive strains'). Secondly, a third toxin (binary toxin CDT) has been found in some *C. difficile* strains and, although the majority of binary toxin positive strains still produce TcdA and TcdB ($A^+B^+CDT^+$ strains), up to 2 % are estimated to produce only binary toxin CDT but not TcdA and TcdB ($A^-B^-CDT^+$ strains). Therefore, the nomenclature discussion group agreed on the following definition of toxigenic and nontoxigenic strains.

Toxigenic strains must produce at least one of the three known toxins. At the present time, five toxin production patterns can be differentiated (Table 3). Toxin production types should be described as A^+B^+, A^-B^-, A^-B^+, etc. The term 'toxin-positive' should be avoided unless it is specifically defined in the text to which of the three toxins it refers. If the production of binary toxin CDT was also tested, toxin production type would be $A^+B^+CDT^+$, etc. If the toxin designation is based on gene identification only we suggest using the toxin designation in parentheses, with an additional text commentary to indicate the basis of presumptive toxin production, e.g. $A^+B^+(CDT^+)$ based on PCR amplification of *cdtB*. The term 'nontoxigenic strain' should be reserved for strains known not to produce any of these three toxins.

REFERENCES

Braun, V., Hundsberger, T., Leukel, P., Sauerborn, M. & von Eichel-Streiber, C. (1996). Definition of the single integration site of the pathogenicity locus in *Clostridium difficile*. *Gene* **181**, 29–38.

Demerec, M., Adelberg, E. A., Clark, A. J. & Hartman, P. E. (1966). A proposal for a uniform nomenclature in bacterial genetics. *Genetics* **54**, 61–76.

Dove, C. H., Wang, S.-Z., Price, S. B., Phelps, C. J., Lyerly, D. M., Wilkins, T. D. & Johnson, J. L. (1990). Molecular characterization of the *Clostridium difficile* toxin A gene. *Infect Immun* **58**, 480–488.

Dupuy, B. & Sonenshein, L. (1998). Regulated transcription of *Clostridium difficile* toxin genes. *Mol Microbiol* **27**, 107–120.

Fiorentini, C., Fabbri, A., Falzano, L., Fattorossi, A., Matarrese, P., Rivabene, R. & Donelli, G. (1998). *Clostridium difficile* toxin B induces apoptosis in intestinal cultured cells. *Infect Immun* **66**, 2660–2665.

Green, G. A., Schué, V. & Monteil, H. (1995). Cloning and characterization of the cytotoxin L-encoding gene of *Clostridium sordellii*: homology with *Clostridium difficile* cytotoxin B. *Gene* **161**, 57–61.

Hacker, J., Hochhut, B., Middendorf, B., Schneider, G., Buchrieser, C., Gottschalk, G. & Dobrindt, U. (2004). Pathogenomics of mobile genetic elements of toxigenic bacteria. *Int J Med Microbiol* **293**, 453–461.

Hammond, G. A. & Johnson, J. L. (1995). The toxinogenic element of *Clostridium difficile* strain VPI 10463. *Microb Pathog* **19**, 203–213.

Table 3. *C. difficile* toxin production types

+ and − refer to the toxin production and not to the presence of toxin genes *tcdA*, *tcdB*, *cdtA* and *cdtB*.

Toxin production type	TcdA	TcdB	CDT	Toxinotype*
$A^+B^+CDT^-$	+	+	−	0 Minor toxinotypes (I, II, XII, XIII, XVIII–XXII)
$A^+B^+CDT^+$	+	+	+	III–VII, IX, XIV, XV, XXIII, (XXIV)
$A^-B^+CDT^-$	−	+	−	VIII
$A^-B^+CDT^+$	−	+	+	X, XVI, XVII
$A^-B^-CDT^+$	−	−	+	XI and some PaLoc-negative strains
$A^-B^-CDT^-$	−	−	−	Nontoxigenic strains

*Toxinotyping is based on genetic alterations of the PaLoc.

Hammond, G. A., Lyerly, D. M. & Johnson, J. L. (1997). Transcriptional analysis of the toxigenic element of *Clostridium difficile*. *Microb Pathog* **22**, 143–154.

Hofmann, F., Herrmann, A., Habermann, E. & von Eichel Streiber, C. (1995). Sequencing and analysis of the gene encoding the α-toxin of *Clostridium novyi* proves its homology to toxins A and B of *Clostridium difficile*. *Mol Gen Genet* **247**, 670–679.

Hundsberger, T., Braun, V., Weidman, M., Leukel, P., Sauerborn, M. & von Eichel-Streiber, C. (1997). Transcription analysis of the genes *tcdA-E* of the pathogenicity locus of *Clostridium difficile*. *Eur J Biochem* **244**, 735–742.

Johnson, J. L., Phelps, C., Barroso, L., Roberts, M. D., Lyerly, D. M. & Wilkins, T. D. (1990). Cloning and expression of the toxin B gene of *Clostridium difficile*. *Curr Microbiol* **20**, 397–401.

Karlsson, S., Dupuy, B., Mukherjee, K., Norin, E., Burman, L. G. & Akerlund, T. (2003). Expression of *Clostridium difficile* toxins A and B and their sigma factor TcdD is controlled by temperature. *Infect Immun* **71**, 1784–1793.

Mani, N. & Dupuy, B. (2001). Regulation of toxin synthesis in *Clostridium difficile* by an alternative RNA polymerase sigma factor. *Proc Natl Acad Sci U S A* **98**, 5844–5849.

Moncrief, J. S., Barroso, L. A. & Wilkins, T. D. (1997). Positive regulation of *Clostridium difficile* toxins. *Infect Immun* **65**, 1105–1108.

Niemann, H. (1992). Clostridial neurotoxins – proposal of a common nomenclature. *Toxicon* **30**, 223–225.

Perelle, S., Gibert, M., Bourlioux, P., Corthier, G. & Popoff, M. R. (1997). Production of a complete binary toxin (actin-specific ADP-ribosyltransferase) by *Clostridium difficile* CD196. *Infect Immun* **65**, 1402–1407.

Spigaglia, P. & Mastrantonio, P. (2002). Molecular analysis of the pathogenicity locus and polymorphism in the putative negative regulator of toxin production (TcdC) among *Clostridium difficile* clinical isolates. *J Clin Microbiol* **40**, 3470–3475.

von Eichel-Streiber, C. & Sauerborn, M. (1990). *Clostridium difficile* toxin A carries a C-terminal repetitive structure homologous to the carbohydrate binding region of streptococcal glycosyltransferases. *Gene* **96**, 107–113.

von Eichel-Streiber, C., Laufenberg-Feldmann, R., Sartringen, S., Schulze, J. & Sauerborn, M. (1990). Cloning of *Clostridium difficile* toxin B gene and demonstration of high N-terminal homology between toxin A and B. *Med Microbiol Immunol* **179**, 271–279.

von Eichel-Streiber, C., Laufenberg-Feldmann, R., Sartingen, S., Shulze, J. & Sauerborn, M. (1992). Comparative sequence analysis of the *Clostridium difficile* toxins A and B. *Mol Gen Genet* **233**, 260–268.

von Eichel-Streiber, C., Boquet, P., Sauerborn, M. & Thelestam, M. (1996). Large clostridial cytotoxins – a family of glycosyltransferases modifying small GTP-binding proteins. *Trends Microbiol* **4**, 375–382.