Available online at www.sciencedirect.com
ScienceDirect
Current Opinion in Chemical Biology

Discovery of virulence factors of pathogenic bacteria

Hsing-Ju Wu[1], Andrew H-J Wang[1] and Michael P Jennings[2]

Discovering virulence factors of pathogenic bacteria is a key in understanding pathogenesis and for identification of targets for novel drugs and design of new vaccines. Comparative genomics, transcriptomics, and proteomics have become the popular tools in discovering the virulence factors in bacterial pathogens, such as *Neisseria meningitidis*, *Yersinia pestis*, *Mycobacterium tuberculosis*, and *Staphylococcus aureus*. In addition, proteomics has been employed successfully in the study of the mechanism of post-translationally modified proteins of bacterial pathogens. Once the putative virulence factors are identified by genomics and/or proteomics, their functions and mechanisms can be further investigated by phenotypic analyses including mutagenesis and biochemical methods and/or structural biology. Combination of these techniques will accelerate the developments of therapeutic drugs and vaccines in combating bacterial diseases.

Addresses
[1] Institute of Biological Chemistry, Academia Sinica, Taipei 115, Taiwan
[2] School of Molecular and Microbial Sciences & Centre for Metals in Biology, The University of Queensland, Brisbane 4072, Australia

Corresponding author: Wang, Andrew H-J (ahjwang@gate.sinica.edu.tw) and Jennings, Michael P (jennings@uq.edu.au)

Current Opinion in Chemical Biology 2008, **12**:93–101

This review comes from a themed issue on
Proteomics and Genomics
Edited by Natalie Ahn and Andrew H.-J. Wang

Available online 5th March 2008

1367-5931/$ – see front matter

DOI 10.1016/j.cbpa.2008.01.023

Introduction

Discovering virulence factors is important in understanding bacterial pathogenesis and their interactions with the host, which may also serve as novel targets in drug and vaccine development. In the pregenomic era, systematic identification of virulence factors were typically done either by biochemical approaches or through genetic screens for genes expressed under *in vivo* conditions or essential for survival in the host (e.g. *In Vivo* Expression Technology (IVET) [1]; Signature-Tagged Mutagenesis (STM) [2]). The development of post-genomic approaches, including genomics, transcriptomics, and proteomics, has accelerated the virulence factor discovery over the past decades. Bacterial genome sequences rapidly add candidate virulence genes to databases. Beyond this rather static description of the cell are the dynamic 'transcriptomic' and 'proteomic' analyses, often referred to as functional genomic studies.

In spite of many newly developed techniques, the gel-based proteomics is still the most frequently used technique in investigation of pathogenic bacteria. Notably, the advantage of proteomics over genomics is the capacity of analyzing post-translational protein modifications that may not be apparent from the analysis of nucleotide sequence data. Post-translational protein modifications have been demonstrated to play an important role in virulence factors; consequently, efforts using proteomic techniques have been devoted to resolve their contribution to bacterial pathogenesis. A further key role for proteomic approaches is the definition of the proteomes from distinct cellular compartment, most notably the cell surface. Once identified by proteomics, the remaining challenge is in deciphering the precise role and function of virulence factor mechanisms of pathogenesis and their interactions with host cells. High-throughput structural analyses such as X-ray crystallography and nuclear magnetic resonance (NMR) are keys in this respect.

In this review, we will discuss the major virulence factors of pathogenic bacteria and the genomic, transcriptomic, and proteomic techniques applied in the field of pathogenic bacteria with the prime focus on proteomic approaches taken and the virulence factors discovered.

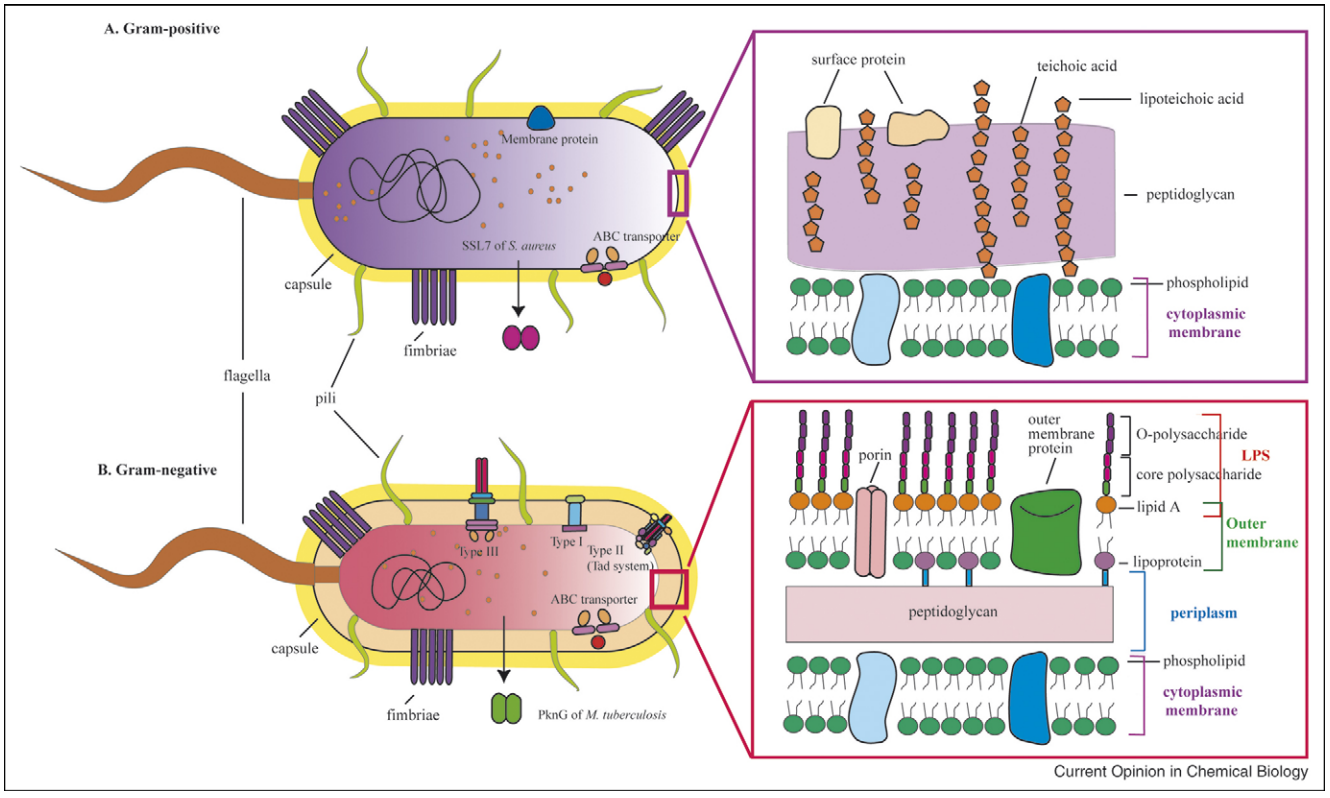
Bacterial virulence factors

In spite of advances in treatment and prevention, bacterial pathogens still pose a major threat on public health worldwide. To understand how pathogenic bacteria interact with their hosts to produce clinical disease is a fundamental issue. A key first step in this process is the identification of novel virulence determinants that may serve as targets for vaccine and drug development.

In essence, the ability of pathogenic bacteria to cause disease in a susceptible host is determined by multiple virulence factors acting individually or together at different stages of infection. Virulence factors are often involved in direct interactions with the host tissues or in concealing the bacterial surface from the host's defense mechanisms. The virulence factors of bacterial pathogens were discussed in an earlier review [3•]. Also, Chen *et al.* [4••] created a database called virulence factor database (VFDB) (http://www.mgc.ac.cn/VFs/).

Bacterial virulence factors can be divided into several groups on the basis of the mechanism of virulence and function [3•] (Figure 1 and Table 1). These are (1)

Figure 1



The schematic diagram showing the major virulence factors of pathogenic bacteria. **(A)** Gram-positive and **(B)** Gram-negative bacteria.

membrane proteins, which play roles in adhesion, colonization, and invasions, promote adherence to host cell surfaces, are responsible for resistance to antibiotics, and promote intercellular communication. (2) Polysaccharide capsules that surround the bacterial cell and have antiphagocytic properties. (3) Secretory proteins, such as toxin, which can modify the host cell environment and are responsible for some host cell–bacteria interactions. Bacterial pathogens use distinct secretion systems, most commonly types I–IV [5] (Figure 1 and Table 1), to transport protein toxins from their cytoplasm into the host or extracellular matrix [6]. Autotransporters (ATs) are virulence proteins translocated by a variety of pathogenic Gram-negative bacteria across the cell envelope to the cell surface or extracellular environment. ATs comprise a family of proteins collectively secreted by the type V pathway [7]. The structure and proposed mechanism of ATs have been reviewed by Dautin and Bernstein [8••]. (4) Cell wall and outer membrane components, such as lipopolysaccharide (LPS or endotoxin) and lipoteichoic acids. Gram-positive bacteria are naturally surrounded by a thick cell wall that has a low permeability to the surrounding environment, while in Gram-negative bacteria the major outer membrane glycolipid, LPS, can protect against complement-mediated lysis. LPS activates the host complement pathway and is a potent inducer of inflammation [3•]. (5) Other virulence factors, such as biofilm forming proteins and siderophores (Table 1). Some bacteria form biofilm, such as *Pseudomonas aeruginosa*, *Mycobacterium*, *Streptococcus pneumoniae*, and *Staphylococcus aureus* [9]. Biofilm formation confers a selective advantage for persistence under environmental conditions and for resistance to antimicrobial agents and also facilitates colonization in the host by the bacteria. In addition, some bacterial virulence factors act as mimics of mammalian proteins to subvert normal host cell processes. Newman *et al.* [10•] identified a novel virulence factor from *Salmonella enterica* serovar Enteritidis, TlpA (TIR-like protein A), which modulates host defense mechanisms.

Genomic and transcriptomic strategies for virulence factor discovery

The continuing reports of complete genome sequences for a variety of bacteria have fuelled the rapid developments in microbial genomics. In 2005, Fraser and Rappuoli [11•] provided a comprehensive list of the microbial genome published. Since then, this list has increased by more than 300 new genome sequences, including at least one strain of every major human pathogen (http://www.tigr.org/tigr-scripts/CMR2/CMRHomePage.spl and http://www.genomesonline.org/). The genomic techniques

Table 1

The classification of the virulence factors of pathogenic bacteria including newly identified virulence factors

Classification	Subclassification	Examples	Reference
1. Membrane proteins	Adhesion	Pilus-associated proteins: microbial surface cell recognition adhesion matrix molecules (MSCRAMMs), for example, Cpa, PrtF1, and PrtF2 of *S. pyogenes*, FnBPA of *S. aureus*	[13,21•]
		Pla and pH 6 fimbriae antigen (PsaA) of *Y. pestis*	[3•]
		Fimbrial adhesins (type I, P and S/F1C) of uropathogenic *E. coli*	[16•]
		LraI family of proteins of *S. pyogenes* and *S. pneumoniae*	[28,30•]
		PsaA of *S. pneumoniae*, ScaA of *S. gordonii*, SsaB from *S. sanguis* and FimA of *S. parasanguis*	[33–36]
	Invasion	Hyaluronidase, lecithinase, and phospholipase of *Clostridium* and Gram-positive cocci	[3•]
	Colonization	Type IV pilus of *N. gonorrhoeae*, *N. menigitidis*, *V. cholerae*, *P. aeruginosa* and entero-pathogenic strains of *E. coli*	[3•,47•]
		Urease of *H. pylori*	[27]
	Surface components	Spa (surface protein A) of *S. aureus*	[17•]
		Surface protein A (SpsA), pneumococcal surface protein A (PspA), choline-binding protein A (CbpA), LytA amidase and pneumococcal surface antigen A (PsaA) of *S. pneumoniae*	[3•]
		LipL32, LipL21 and LipL41 of *Leptospira* spp.	[22]
		Spy0416 of Group A *Streptococcus*	[29]
		VI antigen of *Salmonella typhi*	[3•]
	Outer membrane proteins	YaeT of *E. coli*	[50•]
		FhaC of *B. pertussis*	[51•]
2. Capsule		poly-γ-D-glutamic acid of *B. anthracis*	[3•]
		F1 capsule antigen of *Y. pestis*	[3•]
3. Secretory proteins	Immune response inhibitors	TlpA of *S. enterica* serovar Enteritidis	[10•]
		AvrA of *S. enterica* serovar Typhimurium	[26]
		YopJ of *Yersinia*	[3•]
		Protein kinase G (PknG) and phosphatase (MptpB) of *M. tuberculosis*	[52]
		SSL7 of *S. aureus*	[53•]
	Toxins	Exotoxins: for example, (1) Ymt of *Y. pestis*; (2) Lethal toxin (zinc metalloprotease, Npr599 and InhA) of *B. anthracis*; (3) Protective antigen (PA) and edema toxin of *B. anthracis*; (4) α-Toxin of *S. aureus*; (5) α-Hemolysin (Hly) of uropathogenic *E. coli*; (6) Exotoxin A of *P. aeruginosa*; (7) Diphtheria exotoxin (DT) of *Corynebacterium diptheriae*; (8) Vacoulating toxin of *H. pylori*; (9) Superantigens of *S. pyogenes* and *S. aureus*	[3•,16•,21•,25,29]
	Transport of toxins	Type I: for example, haermolysin of *E. coli* Type II: for example, (1) Pseudopilin XcpT of *Pseudomonas aeruginosa*; (2) The Tad system Type III: for example, (1) Yop of *Y. pestis*; (2) SptP, SgD/SopB and PrgI of *S. typhimurium*; (3) BsaL of *B. peudomallei*; (4) MxiH and Ipa of *S. flexneri* Type V: Autotransporter, for example, (1) AusI of *N. meningitidis*; (2) YapA, C, E-H and K-N of *Y. pestis*	[5,7,15•]
4. Cell wall and outer membrane components		Peptidoglycan, LPS or endotoxin or teichoic acid	[3•]

Table 1 (*Continued*)

Classification	Subclassification	Examples	Reference
5. Others	Biofilm	α-Acetolactate decarboxylase (AlsD) of *S. aureus*	[17•]
		acetolactate synthase of *S. aureus*	[17•]
	Iron acquisition	Siderophore receptor, for example, FrpB, LbpA/B of *N. meningitidis*	[3•]
		Siderophore, for example, (1) Ybt system in *Y. pestis*; (2) Aerobactin, enterobactin, IroN and yersiniabactin of urogenic *E. coli*; (3) Enterochelin of *Salmonella*	[16•,26]
		ABC transport system, for example, YfeABCDE of *Y. pestis*	[3•]
	PhoP/PhoQ two-component system	*Y. pestis*	[3•]

have been widely applied in many pathogenic bacteria [12••], such as *Streptococcus pyogenes* [13] and *Mycobacterium* [14]. Genome sequencing has led to the development of other 'high-throughput' approaches to defining essentiality of genes on the genomic scale. A key example is a process called 'reverse vaccinology' in which *in silico* identification of candidate outer membrane proteins is followed by individual analysis to assess its suitability as a vaccine antigen. The first example of this approach was reported in serogroup B *Neisseria meningitidis* [11•]. Yen *et al.* applied *in silico* screening of the *Yersinia pestis* KIM genome, which led to the identification of 10 putative ATs and reported their possible roles in the *Y. pestis* pathogenesis [15•] (Table S1).

Comparative genomics is a popular tool to identify virulence factors and genes involved in environmental persistence of pathogens. The goal is to correlate those differences to biological function and to gain insight into selective evolutionary pressures and patterns of gene transfer or loss, particularly within the context of virulence in pathogenic species. Comparisons can be performed either with genome sequence or by using microarray-based methods. Ribeiro-Guimaraes and Passolani's study [14] is a good example (Table S1). They compared the protease-coding genes present in the genome of four species of *Mycobacterium* and identified 38 well-conserved proteases that are probably essential for pathogenesis [14]. Similarly, Lloyd *et al.* [16•] utilized comparative genomic hybridization (CGH) analysis on investigating the virulence factors of uropathogenic *Escherichia coli* (UPEC) (Table S1). They were able to conclusively identify 131 genes that were exclusively found in UPEC relative to commensal and fecal isolates. However, half of these genes are annotated as hypothetical or have little functional characterization. Thus, improving the genome annotation and more functional and structural biology studies for characterizing these hypothetical proteins are needed.

Comparison of transcriptomes has been applied in the bacterial pathogen, *Staphylococcus aureus* [17•] (Table S1). Cassat *et al.* [17•] compared the *S. aureus* clinical isolate UAMS-1 with the prototype laboratory strain RN6390 in order to exploit the genes involved in the biofilm formation and virulence (Table 1). The overall profile in RN6390 had the relatively high expression level of genes encoding exotoxins and low expression level of genes encoding surface protein. Conversely, UAMS-1 had the opposite profile [17•]. In this sense, the capacity to efficiently bind host proteins makes an important contribution to staphylococcal infection, and that exotoxin production may be less important. However, they have focused their effort on UAMS-1 and there was a considerable variability among clinical isolates; therefore, there is a need to extend analyses to other staphylococcal clinical strains.

Proteomic strategies for virulence factor discovery

Compared with genomics and transcriptomics, proteomics has the advantage of defining proteins that are differentially expressed, not just purely transcriptional regulation. Also, it can define proteins that are differentially located or secreted to outside of the cell (i.e. to the media or host cell), namely, the surfaceome. In many cases, genomics can predict the proteins that fall into these classes, but proteomics always shows some that are not predicted. Moreover, only proteomics can define proteins that are post-translationally modified. The application of proteomics in pathogenic bacteria on some particular pathogens, such as *Chlamydia* [18], *S. aureus* [19], and *Porphyromonas gingivalis* [20] has been reviewed recently. Therefore, this review will only discuss progresses in the past two years.

Two-dimensional gel electrophoresis and mass spectrometry (2-DE-MS) have been used extensively to characterize and compare proteomes of pathogenic bacteria. Virulence factors are largely membrane, surface, cell wall, or secreted proteins. Therefore, the general approach employs the separation of membrane and cell wall fractions from the cytoplasmic fraction before identification of proteins by 2-DE (Tables S2 and S3). Gatlin *et al.* [21•] present the most comprehensive cell envelope proteome analysis of *S. aureus* so far (Tables S2 and S3). However, one-third of the 48 identified proteins are uncharacterized [21•]. Cullen *et al.* [22] investigated the surfaceome of

Leptospira by biotin labeling of viable leptospires, affinity capture of the biotinylated proteins, 2-DE, and mass spectrometry (Table S3). They showed that the surfaceome consists predominantly of a relatively small number of proteins, most of which have been previously identified, for example, LipL41 and LipL21 [22].

Gram-negative bacteria constitutively secrete outer membrane vesicles (OMVs) into the extracellular milieu, and OMVs are recently proven to be essential for bacterial survival and pathogenesis [23•]; however, the mechanism of vesicle formation and the biological roles of OMVs have not been clearly defined. Therefore, studies using proteomics on OMVs of Gram-negative pathogens have been carried out, for example, in *N. meningitidis* [23•] (Table S3) and *E. coli* [24]. Like comparative genomics, comparative proteomics is a powerful tool to investigate bacterial pathogenesis, such as the studies in *N. meningitidis* [23•], *Bacillus antrhacis* [25], *Salmonella typhimurium* [26], and *S. aureus* [9]. Ferrari *et al.* [23•] compared the proteome of detergent-derived outer membrane vesicles (DOMVs) of group B *N. meningitides* with that of outer membrane vesicles (m-OMVs) of *N. meningitidis* delta *gna33* mutant, in which the gene coding for a lytic transglycosylase was deleted. They presented the first detailed proteomic analysis of DOMVs obtained from the New Zealand epidemic strain NZ98/254, currently under evaluation in clinical trials [23•] (Table S3). Intriguingly, this study demonstrates that the accurate selection of specific mutations represents an effective way to obtain highly enriched membrane fractions (Table S2). Such Δ*gna33*-derived m-OMVs represent a promising alternative vaccine to DOMVs.

Other examples of using comparative proteomics are the studies by Chitlaru *et al.* [25] and Chuang *et al.* [27]. Chitlaru *et al.* [25] investigated the Gram-positive pathogen, *B. anthracis* but focused on secreted proteins rather than outer membrane proteins. *B. anthracis* is the causative agent of anthrax, a lethal disease sporadically affecting humans and animals and the biological warfare agents. They compared the secretomes of a virulent strain Vollum and avirulent strains and identified many putative virulence factors [25] (Tables S2 and S3). Furthermore, this indicates that *B. anthracis* evolved its own set of secreted factors as it is different from the closely related *B. cereus* or *B. thuringiensis* and thus these putative virulence factors particularly involved in anthrax pathogenesis are present in the *B. anthracis* secretome [25].

Host defense system, such as polymorphonuclear leukocytes (PMN), producing substantial amounts of superoxide anion ($O_2^{\bullet -}$) and hydrogen peroxide (H_2O_2) as part of their oxygen-dependent bactericidal mechanisms and thus, oxidative stress had a great effect on bacterial virulence. To address the influence of oxidative stress on *Helicobacter pylori*, Chuang *et al.* [27] compared the protein expression profiles of *H. pylori* under normal and oxidative stress conditions by 2-DE and MALDI-MS. Notably, the protein expression levels of urease accessory protein E (UreE, an essential metallochaperone for urease activity), one of the major virulence factors, and alkylhydroperoxide reductase (AhpC) with antioxidant potential are greatly decreased under stress conditions [27]. Conceivably, UreR and AhpC may thus be potential drug targets against *H. pylori*.

Another important factor needed to be considered is the limitation in using laboratory conditions on defined culture media. Therefore, a large number of studies employed *in vitro* model systems in which they infected eukaryotic cells with bacteria. For example, Zhang *et al.* [28] grew group A streptococcus (GAS), which causes uncomplicated pharyngitis, impetigo, pneumonia, sepsis, necrotizing fasciitis, and streptococcal toxic shock syndrome, in the hyaluronic acid-enrich media in the attempt to create a simple biological system that could reflect some elements of GAS pathogenesis.

Reliable methods capable of providing detailed pictures of surface protein organization in pathogenic bacteria are still unavailable. Recently, Rodriguez-Ortega *et al.* [29] described a new procedure using proteolytic enzymes to 'shave' the GAS surface and the peptides generated are separated and identified (Tables S2 and S3). This approach provides the most extensive map of the surface antigens of GAS strain M1-SF370, including a new possible vaccine target, Spy0416 [29] (Table 1).

S. pneumoniae is a leading cause of bacterial pneumoniae, meningitis, otitis media, and bacteraemia in children and adults worldwide. Encheva *et al.* [30•] developed an extraction method combining the use of detergent, enzyme and a step of mechanical homogenization that allows the characterization and evaluation of a large number of proteins for *S. pneumoniae* through the use of 2-DE and a more sensitive technology, surface-enhanced laser desorption ionization time-of-flight MS with the ProteinChip® arrays, perhaps the most established chip-based proteomics available at present (Tables S2 and S3). As a result, more than 800 protein spots were identified on a single 2-D gel. This was the first proteomic investigation for the characterization of the cytosolic protein fraction of *S. pneumoniae*, and the result was used subsequently to create an expression reference map of this pathogen. Furthermore, they demonstrated that this method does not require high protein yield and can be used in a complementary manner to 2-DE [30•].

Quantitative proteomics

In contrast to numerous 2-DE studies, there are limited studies of using the quantitative proteomic techniques on bacterial pathogenesis. Cho *et al.* [31] presented the first use of a second generation of Isotope-Coded Affinity

Tags (ICAT), that is, cleavable ICAT (cICAT) for comparative proteomics analysis of *M. tuberculosis*. In their study, 586 and 628 proteins were unambiguously identified in the early and later stage non-replicating persistent (NRP-1 and NRP-2) *M. tuberculosis*, respectively [31]. Furthermore, the expression ratio of each protein between log phase vs. NRP-1 and log phase vs. NRP-2 was determined [31] (Table S3). Similarly, Nanduri *et al.* [32] applied the cICAT technology to analyze the *Pasteurella mulocida* proteome response to subminimum inhibitory concentrations (MICs) of amoxicillin, chlortetracycline, and enrofloxacin (Table S3), demonstrating that antibiotics cause secondary effects in addition to the primary target effects.

Oxidative stress proteins and manganese transporter are starting to get recognized as the virulence factors. Metal ions, like Fe^{2+} and Mn^{2+}, are involved in oxidative stress; unlike Fe^{2+}, however, Mn^{2+} and its transporters play important roles in protecting cells against reactive oxygen species. The importance of Mn^{2+} transporter in virulence has been demonstrated in *S. typhimurium* [33], *B. anthracis* [34], *S. pyogenes* [35] and *S. pneumoniae* [36]. Also, we have proven that the accumulation of intracellular Mn^{2+} and ABC-type Mn^{2+} transporters play the important roles in the protection against $O_2^{\bullet-}$ and H_2O_2 in the bacterial pathogens, *Neisseria gonorrhoeae* [37], *N. meningitidis* [38] and *S. pneumoniae* [36,39]. Therefore, there is a three-way interlocking relationship among Mn^{2+}/Mn^{2+} transporter, oxidative stress, and virulence. We further characterized the Mn regulation globally using one-dimensional sodium dodecyl sulfate–polyacrylamide gel electrophoresis with one-dimensional liquid chromatography–tandem mass spectrometry (MS/MS), and cICAT with MS/MS. We showed that 98 proteins are differentially regulated at the post-transcriptional level by Mn^{2+}, helping to resolve the mechanism underlying a complex phenotype (Wu *et al.*, unpublished).

The alternative chemical labeling of quantitative shotgun proteomics uses isobaric tags [40], a technique becoming popular for bacterial pathogenesis study. Recently, Radosevich *et al.* [41] demonstrated the differences in protein expression in *Mycobacterium avium* subsp. *paratuberculosis* laboratory-adapted strain K-10 and the clinical strain 187 using the iTRAQ technology (Table S3). This bacterium is the causative agent of paratuberculosis (Johne's disease) in cattle and sheep [41]. These data may provide insights into the proteins whose expression is important in natural infection but are modified once the pathogen is adapted to laboratory cultivation.

Proteomic analyses for detection of post-translational modifications (PTMs)

For a long time, PTMs have been considered to be restricted to eukaryotes; but recently, PTMs have been proposed in several bacterial models. The functions of PTMs include stability, protection from proteases and signal transduction. PTMs of surface proteins in microbial pathogens are now a well-established phenomenon. Consequently, efforts have been devoted into the role of PTMs in parasite–host interactions. Moreover, PTMs provide effective means to generate diversity and to influence antigenicity. For example, N- and O-linked carbohydrates appear more and more as common features of proteins of bacterial pathogens [42].

Most of our knowledge on microbial protein glycosylation has been obtained from studies on S-layers of archaea and bacteria [42,43]. During the past decade, microbial glycosylation model has been proposed in the surface structures, such as flagella (*P. aeruginosa* and *C. jejuni*) [44] and pili (*N. meningitidis*, *N. gonorrhoeae*, and *P. aeruginosa*) [45,46,47•]. As many of the proposed bacterial glycoproteins are surface-exposed, these modified proteins may play important roles in pathogenicity and antigenicity. The elucidation of the structure of glycosylated peptides, particularly identification of the sugars and their specific sites of attachment, can be made by comparative MS [43]. Linton *et al.* [48] reported mutational and MS/MS analyses for providing the first direct evidence for the function of five glycosyltransferases, that is, PglA, PglC, PglH, PglI, and PglJ, involved in the biosynthesis of the *Campylobacter jejuni* N-linked heptasaccharide glycan (Table S3). ATs has also been reported to have considerable diversity in the post-translational processing of passenger domains, such as cleavage by a variety of mechanisms, lipidation, glycosylation, and oligomerization; these undoubtedly contribute to the functional diversity of the AT superfamily [8••].

During the process of pathogenesis, protein phosphorylation occurs at different stages, including cell–cell interaction and adherence, translocation of bacterial effectors into host cells, and changes in host cellular structure and function induced by infection. A major obstacle in our understanding of protein kinase biology in prokaryotes is the identification of physiologically relevant kinase substrates. Villarino *et al.* [49•] reported the identification of GarA, a Forkhead-associated (FHA) domain-containing protein, as a putative physiological substrate of an essential protein kinase, PknB, in *M. tuberculosis* (Table S3).

Conclusions

Recent advances in bacterial pathogenesis research by genomics, proteomics, and transcriptional profiling have been impressive. The roles of glycosylation and phosorylation in bacteria are only now starting to emerge and other types of PTMs will surely follow. The new field of proteomics is concerned with structural and functional properties of large sets of proteins. The complete characterization of the primary structure of large populations of proteins, however, remains a challenging area for proteomics. As a result, structural studies including X-ray crys-

tallography and NMR [8••,50•,51•,52,53•,54] of the newly discovered virulence factors by genomic or proteomic techniques play an important role in characterizing their functions and interactions with their hosts. Moreover, Craig *et al.* [55•] combined the techniques of X-ray crystallography and 3D cryo-electron microscopy in order to solve the type IV pilus assembly of *N. gonorrhoeae* that cannot be solved by the individual technique otherwise. We expect that combining different genomic, proteomic, and structural results will substantially increase our understanding of complex biological processes associated with virulence factors and assist the development of antibacterial drugs and vaccines.

Acknowledgements

Hsing-Ju Wu is supported by the CJ Martin NHMRC fellowship. We thank Ming-Chin Shih for the help in preparing Figure 1.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.cbpa.2008.01.023.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Mahan MJ, Slauch JM, Mekalanos JJ: **Selection of bacterial virulence genes that are specifically induced in host tissues**. *Science* 1993, **259**:686-688.

2. Hensel M, Shea JE, Gleeson C, Jones MD, Dalton E, Holden DW: **Simultaneous identification of bacterial virulence genes by negative selection**. *Science* 1995, **269**:400-403.

3. • Finlay BB, Falkow S: **Common themes in microbial pathogenicity revisited**. *Microbiol Mol Biol Rev* 1997, **61**:136-169.

The authors elucidated the various definitions of microbial pathogenicity and the idea that pathogens can be distinguished from their non-virulent counterparts by the presence of such virulence genes.

4. •• Chen L, Yang J, Yu J, Yao Z, Sun L, Shen Y, Jin Q: **VFDB: a reference database for bacterial virulence factors**. *Nucleic Acids Res* 2005, **33**:D325-D328.

VFDB provides a comprehensive database with in-depth coverage of the major virulence factors from various best-characterized bacterial pathogens.

5. Tomich M, Planet PJ, Figurski DH: **The *tad* locus: postcards from the widespread colonization island**. *Nat Rev Microbiol* 2007, **5**:363-375.

6. China B, Goffaux F: **Secretion of virulence factors by *Escherichia coli***. *Vet Res* 1999, **30**:181-202.

7. van Ulsen P, Adler B, Fassler P, Gilbert M, van Schilfgaarde M, van der Ley P, van Alphen L, Tommassen J: **A novel phase-variable autotransporter serine protease, AusI, of Neisseria meningitidis**. *Microbes Infect* 2006, **8**:2088-2097.

8. •• Dautin N, Bernstein HD: **Protein secretion in Gram-negative bacteria via the autotransporter pathway**. *Annu Rev Microbiol* 2007, **61**:89-112.

This review summarizes the protein structures of a variety of the autotransporters and discusses each stage of atutotransporter biogenesis.

9. Resch A, Leicht S, Saric M, Pasztor L, Jakob A, Gotz F, Nordheim A: **Comparative proteome analysis of *Staphylococcus aureus* biofilm and planktonic cells and correlation with transcriptome profiling**. *Proteomics* 2006, **6**:1867-1877.

10. • Newman RM, Salunkhe P, Godzik A, Reed JC: **Identification and characterization of a novel bacterial virulence factor that shares homology with mammalian Toll/interleukin-1 receptor family proteins**. *Infect Immun* 2006, **74**:594-601.

Genetic and biochemical characterization of TlpA, a novel virulence factor, suggests that it is important for bacterial virulence *in vivo* and modulates host defense mechanism involved in regulation of NF-κB and caspase activation.

11. • Fraser CM, Rappuoli R: **Application of microbial genomic science to advanced therapeutics**. *Annu Rev Med* 2005, **56**:459-474.

The authors comment on the genomic techniques applied in developing antimicrobial agents and vaccines.

12. •• Raskin DM, Seshadri R, Pukatzki SU, Mekalanos JJ: **Bacterial genomics and pathogen evolution**. *Cell* 2006, **124**:703-714.

The authors critically review genomic techniques to study bacterial pathogenesis, such as transposon site hybridization (TraSH), saturating transposon mutagenesis, comparative genomics, and transcriptional profiling.

13. Musser JM, DeLeo FR: **Toward a genome-wide systems biology analysis of host-pathogen interactions in group A *Streptococcus***. *Am J Pathol* 2005, **167**:1461-1472.

14. Ribeiro-Guimaraes ML, Pessolani MC: **Comparative genomics of mycobacterial proteases**. *Microb Pathog* 2007, **43**:173-178.

15. • Yen YT, Karkal A, Bhattacharya M, Fernandez RC, Stathopoulos C: **Identification and characterization of autotransporter proteins of Yersinia pestis KIM**. *Mol Membr Biol* 2007, **24**: 28-40.

Ten putative ATs were identified and their possible roles were proposed in this study in order to understand more about the *Y. pestis* pathogenesis.

16. • Lloyd AL, Rasko DA, Mobley HL: **Defining genomic islands and uropathogen-specific genes in uropathogenic *Escherichia coli***. *J Bacteriol* 2007, **189**:3532-3546.

This is a very recent example utilizing comparative genomic hybridization (CGH) analysis to investigate the virulence factors of UPEC. In this study, the genomes of three pyeolonephritis strains, four cystitis strains, and three fecal/commensal *E. coli* isolates were hybridized against the *E. coli* CFT073 microarray.

17. • Cassat J, Dunman PM, Murphy E, Projan SJ, Beenken KE, Palm KJ, Yang SJ, Rice KC, Bayles KW, Smeltzer MS: **Transcriptional profiling of a *Staphylococcus aureus* clinical isolate and its isogenic *agr* and *sarA* mutants reveals global differences in comparison to the laboratory strain RN6390**. *Microbiology* 2006, **152**:3075-3090.

The genome-scale transcriptional profiling was carried out for comparing the *S. aureus* clinical isolate UAMS-1 with the prototype laboratory strain RN6390.

18. Vandahl BB, Birkelund S, Christiansen G: **Genome and proteome analysis of *Chlamydia***. *Proteomics* 2004, **4**:2831-2842.

19. Hecker M, Engelmann S, Cordwell SJ: **Proteomics of *Staphylococcus aureus*—current state and future challenges**. *J Chromatogr B Analyt Technol Biomed Life Sci* 2003, **787**:179-195.

20. Lamont RJ, Meila M, Xia Q, Hackett M: **Mass spectrometry-based proteomics and its application to studies of *Porphyromonas gingivalis* invasion and pathogenicity**. *Infect Disord Drug Targets* 2006, **6**:311-325.

21. • Gatlin CL, Pieper R, Huang ST, Mongodin E, Gebregeorgis E, Parmar PP, Clark DJ, Alami H, Papazisi L, Fleischmann RD *et al.*: **Proteomic profiling of cell envelope-associated proteins from *Staphylococcus aureus***. *Proteomics* 2006, **6**:1530-1549.

The study applied multiple proteomic techniques including 2-DE, liquid chromatography–tandem mass spectrometry (LC–MS/MS) of cell surface biotinylated proteins and two-dimensional LC–MS/MS (2-D LC–MS/MS) on two isogenic vancomycin-intermediate *S. aureus* (VISA) strains, HIP5827 and VP32.

22. Cullen PA, Xu X, Matsunaga J, Sanchez Y, Ko AI, Haake DA, Adler B: **Surfaceome of *Leptospira* spp.**. *Infect Immun* 2005, **73**:4853-4863.

23. • Ferrari G, Garaguso I, Adu-Bobie J, Doro F, Taddei AR, Biolchi A, Brunelli B, Giuliani MM, Pizza M, Norais N *et al.*: **Outer membrane vesicles from group B *Neisseria meningitidis* delta *gna33* mutant: proteomic and immunological comparison with detergent-derived outer membrane vesicles**. *Proteomics* 2006, **6**:1856-1866.

The Δ*gna33* mutant is a useful tool for producing a massive amount of outer membrane proteins for proteomic technique.

24. Lee EY, Bang JY, Park GW, Choi DS, Kang JS, Kim HJ, Park KS, Lee JO, Kim YK, Kwon KH *et al.*: **Global proteomic profiling of native outer membrane vesicles derived from *Escherichia coli***. *Proteomics* 2007, **7**:3143-3153.

25. Chitlaru T, Gat O, Gozlan Y, Ariel N, Shafferman A: **Differential proteomic analysis of the *Bacillus anthracis* secretome: distinct plasmid and chromosome CO_2-dependent cross talk mechanisms modulate extracellular proteolytic activities**. *J Bacteriol* 2006, **188**:3551-3571.

26. Adkins JN, Mottaz HM, Norbeck AD, Gustin JK, Rue J, Clauss TR, Purvine SO, Rodland KD, Heffron F, Smith RD: **Analysis of the *Salmonella typhimurium* proteome through environmental response toward infectious conditions**. *Mol Cell Proteomics* 2006, **5**:1450-1461.

27. Chuang MH, Wu MS, Lin JT, Chiou SH: **Proteomic analysis of proteins expressed by *Helicobacter pylori* under oxidative stress**. *Proteomics* 2005, **5**:3895-3901.

28. Zhang M, McDonald FM, Sturrock SS, Charnock SJ, Humphery-Smith I, Black GW: **Group A *Streptococcus* cell-associated pathogenic proteins as revealed by growth in hyaluronic acid-enriched media**. *Proteomics* 2007, **7**:1379-1390.

29. Rodriguez-Ortega MJ, Norais N, Bensi G, Liberatori S, Capo S, Mora M, Scarselli M, Doro F, Ferrari G, Garaguso I *et al.*: **Characterization and identification of vaccine candidate proteins through analysis of the group A *Streptococcus* surface proteome**. *Nat Biotechnol* 2006, **24**:191-197.

30. • Encheva V, Gharbia SE, Wait R, Begum S, Shah HN: **Comparison of extraction procedures for proteome analysis of *Streptococcus pneumoniae* and a basic reference map**. *Proteomics* 2006, **6**:3306-3317.

This paper provides a cell lysis and protein solubilization method that minimizes protein losses and allows for maximal coverage of the proteome of *S. pneumoniae*.

31. Cho SH, Goodlett D, Franzblau S: **ICAT-based comparative proteomic analysis of non-replicating persistent *Mycobacterium tuberculosis***. *Tuberculosis (Edinb)* 2006, **86**:445-460.

32. Nanduri B, Lawrence ML, Boyle CR, Ramkumar M, Burgess SC: **Effects of subminimum inhibitory concentrations of antibiotics on the *Pasteurella multocida* proteome**. *J Proteome Res* 2006, **5**:572-580.

33. Zaharik ML, Cullen VL, Fung AM, Libby SJ, Kujat Choy SL, Coburn B, Kehres DG, Maguire ME, Fang FC, Finlay BB: **The *Salmonella enterica* serovar typhimurium divalent cation transport systems MntH and SitABCD are essential for virulence in an Nramp1G169 murine typhoid model**. *Infect Immun* 2004, **72**:5522-5525.

34. Gat O, Mendelson I, Chitlaru T, Ariel N, Altboum Z, Levy H, Weiss S, Grosfeld H, Cohen S, Shafferman A: **The solute-binding component of a putative Mn(II) ABC transporter (MntA) is a novel *Bacillus anthracis* virulence determinant**. *Mol Microbiol* 2005, **58**:533-551.

35. Janulczyk R, Ricci S, Bjorck L: **MtsABC is important for manganese and iron transport, oxidative stress resistance, and virulence of *Streptococcus pyogenes***. *Infect Immun* 2003, **71**:2656-2664.

36. McAllister LJ, Tseng HJ, Ogunniyi AD, Jennings MP, McEwan AG, Paton JC: **Molecular analysis of the *psa* permease complex of *Streptococcus pneumoniae***. *Mol Microbiol* 2004, **53**:889-901.

37. Tseng HJ, Srikhanta Y, McEwan AG, Jennings MP: **Accumulation of manganese in *Neisseria gonorrhoeae* correlates with resistance to oxidative killing by superoxide anion and is independent of superoxide dismutase activity**. *Mol Microbiol* 2001, **40**:1175-1186.

38. Seib KL, Tseng HJ, McEwan AG, Apicella MA, Jennings MP: **Defenses against oxidative stress in *Neisseria gonorrhoeae* and *Neisseria meningitidis*: distinctive systems for different lifestyles**. *J Infect Dis* 2004, **190**:136-147.

39. Tseng HJ, McEwan AG, Paton JC, Jennings MP: **Virulence of *Streptococcus pneumoniae*: *psaA* mutants are hypersensitive to oxidative stress**. *Infect Immun* 2002, **70**:1635-1639.

40. Aggarwal K, Choe LH, Lee KH: **Quantitative analysis of protein expression using amine-specific isobaric tags in *Escherichia coli* cells expressing *rhsA* elements**. *Proteomics* 2005, **5**:2297-2308.

41. Radosevich TJ, Reinhardt TA, Lippolis JD, Bannantine JP, Stabel JR: **Proteome and differential expression analysis of membrane and cytosolic proteins from *Mycobacterium avium* subsp. paratuberculosis strains K-10 and 187**. *J Bacteriol* 2007, **189**:1109-1117.

42. Szymanski CM, Wren BW: **Protein glycosylation in bacterial mucosal pathogens**. *Nat Rev Microbiol* 2005, **3**:225-237.

43. Cordwell SJ: **Exploring and exploiting bacterial proteomes**. *Methods Mol Biol* 2004, **266**:115-135.

44. Szymanski CM, Logan SM, Linton D, Wren BW: ***Campylobacter*—a tale of two protein glycosylation systems**. *Trends Microbiol* 2003, **11**:233-238.

45. Parge HE, Forest KT, Hickey MJ, Christensen DA, Getzoff ED, Tainer JA: **Structure of the fibre-forming protein pilin at 2.6 A resolution**. *Nature* 1995, **378**:32-38.

46. Stimson E, Virji M, Makepeace K, Dell A, Morris HR, Payne G, Saunders JR, Jennings MP, Barker S, Panico M *et al.*: **Meningococcal pilin: a glycoprotein substituted with digalactosyl 2,4-diacetamido-2,4,6-trideoxyhexose**. *Mol Microbiol* 1995, **17**:1201-1214.

47. • Aas FE, Egge-Jacobsen W, Winther-Larsen HC, Lovold C, Hitchen PG, Dell A, Koomey M: ***Neisseria gonorrhoeae* type IV pili undergo multisite, hierarchical modifications with phosphoethanolamine and phosphocholine requiring an enzyme structurally related to lipopolysaccharide phosphoethanolamine transferases**. *J Biol Chem* 2006, **281**:27712-27723.

The top-down MS approach, in which PTMs can be detected directly from intact proteins, was applied in this paper and it can alleviate the problems of a bottom-up approach in which proteolytically derived peptides were examined by MS/MS.

48. Linton D, Dorrell N, Hitchen PG, Amber S, Karlyshev AV, Morris HR, Dell A, Valvano MA, Aebi M, Wren BW: **Functional analysis of the *Campylobacter jejuni* N-linked protein glycosylation pathway**. *Mol Microbiol* 2005, **55**:1695-1703.

49. • Villarino A, Duran R, Wehenkel A, Fernandez P, England P, Brodin P, Cole ST, Zimny-Arndt U, Jungblut PR, Cervenansky C *et al.*: **Proteomic identification of *M. tuberculosis* protein kinase substrates: PknB recruits GarA, a FHA domain-containing protein, through activation loop-mediated interactions**. *J Mol Biol* 2005, **350**:953-963.

The authors report the identification of GarA, as a putative physiological substrate of an essential protein kinase, PknB, in *M. tuberculosis* using a global proteomic approach that combined 2-DE, autoradiography, and MS identification. They further investigated protein kinase–substrate interactions by MS, enzymological, and binding studies of wild-type and mutant proteins.

50. • Kim S, Malinverni JC, Sliz P, Silhavy TJ, Harrison SC, Kahne D: **Structure and function of an essential component of the outer membrane protein assembly machine**. *Science* 2007, **317**:961-964.

This is an excellent work solving the crystal structure of a fragment of the Omp85-family member YaeT from *E. coli*. The fragment encompasses four complete polypeptide transport-associated (POTRA) domains and a short segment of the fifth one at the carboxyl terminus.

51. • Clantin B, Delattre AS, Rucktooa P, Saint N, Meli AC, Locht C, Jacob-Dubuisson F, Villeret V: **Structure of the membrane protein FhaC: a member of the Omp85-TpsB transporter superfamily**. *Science* 2007, **317**:957-961.

This is an excellent study solving the crystal structure of FhaC, a member of the Omp85 superfamily involved in the secretion of filamentous hemagglutinin (FHA) in *Bordetella pertussis*. The structure shows a 16-stranded β barrel that is occluded by an N-terminal α helix and an extracellular loop and two periplasmic POTRA domains structurally resembling those of YaeT.

52. Scherr N, Honnappa S, Kunz G, Mueller P, Jayachandran R, Winkler F, Pieters J, Steinmetz MO: **Structural basis for the specific inhibition of protein kinase G, a virulence factor of *Mycobacterium tuberculosis***. *Proc Natl Acad Sci U S A* 2007, **104**:12151-12156.

53. • Ramsland PA, Willoughby N, Trist HM, Farrugia W, Hogarth PM, Fraser JD, Wines BD: **Structural basis for evasion of IgA immunity by *Staphylococcus aureus* revealed in the complex of SSL7 with Fc of human IgA1**. *Proc Natl Acad Sci U S A* 2007, **104**:15051-15056.

The authors report the X-ray crystal structure of protein kinase G (PknG) in complex with the inhibitor, AX20017. The structure of PknG consists of a central kinase domain that is flanked by N- and C-terminal rubredoxin and tetratrico–peptide repeat domains, respectively, and the rubredoxin domain is essential for the PknG activity.

54. Zhang L, Wang Y, Picking WL, Picking WD, De Guzman RN: **Solution structure of monomeric BsaL, the type III secretion needle protein of *Burkholderia pseudomallei***. *J Mol Biol* 2006, **359**:322-330.

55. • Craig L, Volkmann N, Arvai AS, Pique ME, Yeager M, Egelman EH, Tainer JA: **Type IV pilus structure by cryo-electron microscopy and crystallography: implications for pilus assembly and functions**. *Mol Cell* 2006, **23**:651-662.

The Type IV pilus (T4P) structure of *N. gonorrhoeae* was determined by quantitative fitting of a 2.3 Å full-length pilin crystal structure into a 12.5 Å resolution native T4P solved by cryo-electron microscopy.