Pathogen virulence genes – implications for vaccines and drug therapy

Christoph Tang* and David Holden*

*Department of Paediatrics, John Radcliffe Hospital, Oxford, UK and †Department of Infectious Diseases, Imperial College of Science, Technology and Medicine, Hammersmith Hospital, London, UK

The emergence and spread of bacteria resistant to antimicrobial drugs is a major public health problem with a growing number of infections becoming virtually untreatable. There is a need to develop interventions both to prevent and to treat diseases caused by multi-resistant microbes. We review some recently developed methods (including whole genome nucleotide sequencing projects) to study bacterial pathogenesis, and discuss how knowledge gained from understanding the molecular mechanisms of disease can be applied to combat the threat of infectious diseases.

Correspondence to: Prof. David Holden, Department of Infectious Diseases, Imperial College of Science, Technology and Medicine, Hammersmith Hospital, Du Cane Road, London W12 0NN, UK

Our ability to treat and prevent bacterial infections is becoming increasingly inadequate. Over the past decade, bacterial pathogens have continued to be leading causes of mortality in developing countries, and have re-emerged as important public health problems in developed countries. This represents a significant change from the position in the 1970s and early 1980s. Until that time, bacterial infections appeared to be waning in the face of improvements in living conditions and the implementation of vaccination programmes. Furthermore, there were sufficient antimicrobial drugs available to withhold classes of compounds for life-threatening infections. However, no new class of antimicrobial has become available in the past 20 years, and the therapeutic reserve has now run dry. We are currently faced with the prospect of untreatable systemic bacterial infections caused by important human pathogens such as *Staphylococcus aureus* and *Enterococcus faecium*. Therefore there is an urgent need to identify new therapies and vaccines to treat and/or prevent bacterial infections.

The global importance of the emergence and spread of antimicrobial resistance has been recognised by the US Centers for Disease Control and the World Health Organization, resulting in the International Surveillance Program for Emerging Antimicrobial Resistance in Hospitals (INSPEAR). This collaborative venture is designed to monitor antimicrobial susceptibility from 1999 in hospitals across 25 countries.
Table 1  Bacterial pathogens causing increased concern

<table>
<thead>
<tr>
<th>Problem</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiply drug resistant</td>
<td>Enterococcus spp.</td>
</tr>
<tr>
<td></td>
<td><em>Mycobacterium tuberculosis</em></td>
</tr>
<tr>
<td></td>
<td><em>Salmonella typhi</em></td>
</tr>
<tr>
<td></td>
<td><em>Salmonella typhimurium</em></td>
</tr>
<tr>
<td></td>
<td><em>Staphylococcus aureus</em></td>
</tr>
<tr>
<td></td>
<td><em>Streptococcus pneumoniae</em></td>
</tr>
<tr>
<td></td>
<td><em>Yersinia pestis</em></td>
</tr>
<tr>
<td>Newly Identified pathogens</td>
<td><em>Borrelia burgdorferi</em></td>
</tr>
<tr>
<td></td>
<td><em>Ehrlichosis</em></td>
</tr>
<tr>
<td></td>
<td><em>Escherichia coli O157</em></td>
</tr>
<tr>
<td></td>
<td><em>Legionella pneumophila</em></td>
</tr>
<tr>
<td>Nosocomial pathogens</td>
<td><em>Acinetobacter baumannii</em></td>
</tr>
<tr>
<td></td>
<td><em>Clostridium difficile</em></td>
</tr>
<tr>
<td></td>
<td><em>Coagulase-negative Staphylococci</em></td>
</tr>
<tr>
<td></td>
<td><em>Enterococcus spp.</em></td>
</tr>
<tr>
<td></td>
<td><em>Pseudomonas aeruginosa</em></td>
</tr>
<tr>
<td></td>
<td><em>Serratia marcescens</em></td>
</tr>
</tbody>
</table>

The bacteria responsible for the current precarious situation can be broadly classified into three groups (Table 1). Firstly, several common pathogens have become multiply drug resistant, in part through the increased and inappropriate medical and agricultural use of antimicrobials. Secondly, ‘new’ pathogens have been described that result from exposure of individuals through travel or changes in lifestyle (e.g. animal husbandry, food processing and urbanisation). Furthermore, increasing medical intervention has rendered people susceptible to a wide range of organisms by reducing physical (e.g. cannulation and catheterisation) and immunological barriers to infection.

In the last few years, our understanding of the molecular basis of virulence of certain well-studied bacterial pathogens has increased dramatically. This has resulted from the application of recombinant DNA technology and cell biology to investigate bacterial infections, and the development of genetic techniques for identifying virulence genes. More recently, information about bacterial pathogens has become available from bacterial genome sequencing projects. Many who work in the field hope this understanding will be exploited to develop new interventions against bacterial infections. We will discuss what has been learnt from recent progress in our knowledge of the biology of bacterial pathogens, and how this might impact on both drug discovery and vaccine design.
Virulence genes

The traditional view of virulence has focused on microbial cell surface structures and secreted toxins that mediate interactions with host cells or tissue damage. However, through genetic analysis it has become apparent that bacterial pathogenesis is a much more complex process with many examples of specific, interdependent interactions between hosts and pathogens. Bacteria can sense and respond to intracellular and extracellular habitats in their hosts, modulate cell surface structures that are the target of immune responses, and scavenge nutrients from their environment to allow replication. The broad view of virulence encompasses all those attributes that contribute to bacterial survival and replication in vivo, as well as determinants of tissue tropism and damage.

Microbial genomics

The rapid completion of the Haemophilus influenzae genome sequencing project in 1995 demonstrated that bacterial sequencing projects can be undertaken by single academic or commercial institutes through improvements in DNA sequencing technology (including sequence assembly), and by avoiding the labour-intensive construction of high-density genetic maps. Since then, there has been an exponential increase in completed genomes with the possibility that the nucleotide sequence for most major bacterial pathogens will be known by the year 2000. The relevance of sequencing projects for drug and vaccine discovery are obvious. The complete sequence includes information on every virulence gene and all potential vaccine candidates, and the sequence databases will become indispensable for research in microbiology. The next challenge is to understand the function of the genes. Unfortunately, Bacillus subtilis apart, there do not appear to be any collaborative projects aimed at systematically analysing the data generated from bacterial sequencing projects.

Once the nucleotide sequencing is completed and the sequence assembled, potential open reading frames (orfs) are identified and used to perform database searches for genes of known function in other organisms. In this way, functions are ascribed to orfs and the sequence is annotated. Annotation can also be based on comparing the predicted folding of putative proteins, though this is a relatively insensitive method given current knowledge about protein structure. For highly conserved gene products involved in fundamental cellular processes (such as DNA replication and protein synthesis), the assignment of
Impact of genomics on healthcare

Table 2: Proportion of predicted proteins of unknown function in bacterial pathogens for which the complete nucleotide sequence is known

<table>
<thead>
<tr>
<th>Species</th>
<th>Genome size (Mbp)</th>
<th>No. of proteins</th>
<th>Genes without predicted function (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Borrelia burgdorferi</em></td>
<td>0.91</td>
<td>853</td>
<td>41</td>
</tr>
<tr>
<td><em>Escherichia coli</em></td>
<td>4.64</td>
<td>4288</td>
<td>38</td>
</tr>
<tr>
<td><em>Haemophilus influenzae</em></td>
<td>1.83</td>
<td>1703</td>
<td>40</td>
</tr>
<tr>
<td><em>Helicobacter pylori</em></td>
<td>1.7</td>
<td>1590</td>
<td>31</td>
</tr>
<tr>
<td><em>Mycobacterium tuberculosis</em></td>
<td>4.41</td>
<td>3942</td>
<td>60</td>
</tr>
<tr>
<td><em>Staphylococcus aureus</em></td>
<td>2.8</td>
<td>2400</td>
<td>&gt;50</td>
</tr>
<tr>
<td><em>Streptococcus pneumoniae</em></td>
<td>2.5</td>
<td>2300</td>
<td>&gt;50</td>
</tr>
</tbody>
</table>

function from database searches is relatively secure. However, for the majority of orfs, information from database searches is far less certain as annotation frequently ascribes only potential biochemical functions to the gene products. To establish unambiguously the role of gene products in living bacteria requires genetic rather than genomic analysis. Overinterpretation of sequence data is more of a problem in those organisms that contain many genes whose products perform biochemically similar reactions. For example, 79 potential ABC transporters have been identified in *Escherichia coli*, but the function of most of these have not been defined. Indeed, 38% of all genes in *E. coli* belong to such paralogous gene families. The frequency of paralogues in bacterial genomes varies from species to species, and is lower in *H. influenzae* and *Helicobacter pylori* than *E. coli*. Given the lack of specific information from database searches, attempts to generate computer models of living cells, based largely on the results of nucleotide sequence matches, are over ambitious and premature.

The bacterial sequencing projects have underscored our limited understanding of the biology of bacteria. In *E. coli*, which has been studied for over 30 years and for which there is more accumulated biochemical and structural information than for any other free-living organism, the function of over 30% of genes is not known. A significant proportion of orfs in every sequenced microbial genome is of unknown function (Table 2). The function of these genes will have to be determined by experimentation, and the importance of mutational analysis cannot be overstated. Genetic research is moving from phenotype-based gene finding to gene-based function finding. The relevant phenotypes relating to drug discovery are bacterial survival and proliferation in the host, and/or the ability to induce tissue damage. In vaccine development, genes must encode proteins involved in the synthesis of immunogenic structures (both of protein and non-protein antigens) or be required for virulence (for live attenuated vaccines).
There is no shortage of potential applications of genome sequence data for the development of drugs and vaccines. These include exploiting the sequence data directly by DNA immunisation. With the sequence data, orfs can be amplified by PCR and ligated directly into the DNA immunisation vectors. DNA vaccination should prove most valuable for bacteria that are predominantly intracellular in the host and that elicit cellular immune responses (e.g. *Chlamydia*, *Mycobacterium tuberculosis* and *Salmonella* spp.). The validity of this approach was demonstrated by expression library immunisation developed with *Mycoplasma pneumoniae*. This method involves cloning random fragments of bacterial DNA in a vector downstream of a promoter active in eukaryotic cells, vaccinating animals with libraries of recombinant plasmids, then challenging animals with *Mycoplasma*. Libraries that confer protection can then be further analysed to identify the clones responsible for protection. Alternative delivery systems for DNA vaccines include live attenuated *Salmonella typhimurium* strains that have been shown to be highly effective carriers of DNA vaccines against *Listeria monocytogenes*, and mucosal immunisation with liposomes containing plasmid vectors.

Analysis of complete nucleotide sequence for potentially membrane-spanning proteins can identify surface expressed antigens for incorporation in vaccines. Several computer programmes such as TopPred and Moment have been devised to identify membrane-spanning domains; in *Borrelia burgdorferi*, the causative agent of Lyme disease, a total of 526 proteins have one or more putative membrane-spanning domain with 183 having more than one domain. The *Borrelia* sequence also illustrates the potential antigenic diversity of this organism. *Borrelia* can express an extensive variety of lipoproteins (VMPs) on its cell surface, and the genetic mechanism underlying this antigenic switching is well understood; structural genes are translocated from 'silent' loci to an 'expression' site on linear plasmids. The complete sequence has demonstrated the full antigenic repertoire of VMPs, confirming that VMP-based vaccines would be difficult to develop. Many microorganisms have developed ways to generate antigenic variation which subverts host immune responses. Epitopes that are usually non-immunogenic would be expected to be conserved, and vaccination strategies could be employed (e.g. improved presentation or mode of delivery) that lead to enhanced responses against these epitopes.

Although the genome sequences do not offer such a direct route to non-protein based vaccines, they can provide invaluable information for the analysis of non-protein vaccine candidates as the basis for the rational design and production of candidate immunogens. An excellent example is the study of *H. influenzae* lipopolysaccharide (LPS). LPS is important both for the interaction of this organism with epithelial cells.
Impact of genomics on healthcare

in the nasopharynx, and as a mediator of septic shock that is sometimes seen during systemic infection. The *Haemophilus* genome was searched for homologues of gene products involved in the biosynthesis of LPS in other organisms. The *H. influenzae* genes responsible for the biosynthesis of terminal LPS structures were identified and mutations introduced into the orfs. The specific mutants were used to analyse the function of each orf by examining their reactivity with a selection of monoclonal antibodies, their LPS structure by mass spectroscopy, and their behaviour in an animal model of infection. By this approach it should be possible to determine firstly precise structure:function relationships for this important molecule, and secondly LPS structures that are safe and potentially immunogenic.

Another valuable aspect of complete sequence information is the identification of chromosomal substrates for the transfer and acquisition of genes involved in either virulence or resistance to antibiotics. The entire repertoire of transposons, phages, and retromers that can harbour cassettes encoding resistance to antimicrobials will be shown by the nucleotide sequence. Important determinants of resistance such as the *vanA* operon, responsible for high-level vancomycin resistance in *E. faecium* and *Staph. aureus*, reside on the transposon Tn1546. Furthermore, genome analysis will identify whether specific virulence determinants lie on transferable genetic elements. Examples of virulence genes located on resident bacteriophages include those encoding cholera and the diphtheria toxins. Additionally, a number of pathogens possess large regions of DNA, often in excess of 10 kb, containing clusters of virulence related genes. These ‘pathogenicity islands’ were first identified by Hacker in two clinical isolates of uropathogenic *E. coli*. The occurrence of pathogenicity islands is of considerable interest, not only because they contain many virulence-related genes, but also because their G+C content usually differs from the remainder of the resident chromosome and the finding of insertion sequences at their termini both suggest they were acquired by horizontal transfer. Not only do these islands provide a means of rapidly identifying virulence gene clusters but also can help understand the intra- and interspecies transfer of these traits. Comparison of the complete nucleotide sequence of closely related bacteria that behave differently in hosts should reveal genes responsible for their difference in behaviour. For instance, *E. coli* O157 (an important cause of bloody diarrhoea and haemolytic uraemic syndrome) has 1 Mb more genetic information than the sequenced *E. coli* K12 strain. One pathogenicity island has already been found in *E. coli* O157 that is absent from *E. coli* K12, and the remainder of the O157-specific sequence is located in about 300 regions scattered throughout the genome, some of which will contribute to the pathogenicity of this serotype.
Screens for genes

Virulence genes can be identified using assays that replicate one or more stages of infection. Much work has focused on the interaction of pathogens with monolayers of cells in tissue culture; these screens have facilitated the analysis of bacterial determinants of host cell invasion or intracellular survival. The advantages of working with such in vitro systems are that they are easily controlled and amenable to detailed biochemical analysis.

However, in vitro screens will only identify a subset of the virulence genes required for pathogenesis because they do not reflect the diverse environments that bacteria encounter in a host. More complex assays are required for a fuller appreciation of bacterial virulence. Recently, two genetic schemes have been developed in which pools of bacterial strains can be screened in animal models of infection. In vivo expression technology (IVET) was designed to identify promoters of genes that are specifically induced in host tissues, and signature-tagged mutagenesis (STM) allows a large number of insertional mutant strains to be screened simultaneously for loss of virulence in a single animal.

IVET is a promoter trap in which random fragments of bacterial DNA are ligated upstream of a promoterless gene whose activity is easily monitored (Fig. 1A). A pool of bacteria harbouring plasmids with promoter fusions is inoculated into an animal, and promoters that are specifically active in the host identified. Once these promoters have been isolated, specific mutations are introduced into the corresponding genes. Evaluation of the virulence of the mutant strains is used to confirm whether the genes are required for pathogenesis. IVET of S. typhimurium, a Gram-negative bacterium that causes a systemic illness in mice similar to human typhoid fever, resulted in the isolation of three genes that were subsequently shown to be important virulence determinants by mutational and LD analysis.

Application of IVET to Pseudomonas aeruginosa led to the identification of genes induced after intraperitoneal inoculation of neutropenic mice. A total of 22 in vivo induced loci were partially sequenced. These encode products involved in iron acquisition, amino acid biosynthesis, signal transduction, gene regulation and several with unknown function. Inactivation of a gene with significant similarity to members of the bacterial ferric uptake regulator family resulted in a 100-fold reduction in LD, demonstrating the importance of this locus in Pseudomonas virulence. IVET was also used to isolate virulence genes in Staph. aureus including agr, a locus that controls the expression of several virulence genes. Mutagenesis of 11 in vivo induced loci produced seven strains significantly attenuated in virulence. These studies show that IVET is a useful approach to identify subsets of genes that are involved in bacterial pathogenesis (Fig. 2).
Fig. 1 Principles of IVET and STM. In STM, an insertional mutagen (transposon) is modified by incorporation of DNA signature tags. A collection of different insertional mutants of a bacterial pathogen, each carrying a different tag, is assembled in a microtitre dish. The mutants are pooled and used as the inoculum for an appropriate animal model of infection. After a period of time in which virulent bacteria have multiplied within the host, bacteria are recovered from the host. Signature tags representing strains in the inoculum and strains recovered from the animal are separately amplified using primers that anneal to invariant sequences flanking the signature tags, then labelled and used to probe nylon membranes carrying DNA from the mutants in the microtitre dish. An avirulent mutant is identified by the failure to yield a signal on the membrane hybridized with the tags recovered from the animal. In the original version of IVET, small chromosomal DNA fragments from the bacterial pathogen are cloned upstream of a promoterless operon consisting of purA and lacZY. Introduction of these fusion plasmids into a strain of the pathogen carrying a purA mutation results in their chromosomal integration by homologous recombination. A pool of chromosomal fusion strains is injected into an animal host and, because purine auxotrophy greatly attenuates growth in vivo, only those strains in which the purA gene is driven by a bacterial promoter that is active in vivo will be pathogenic. The presence of the lacZY gene in the operon allows the transcriptional activity of the promoter to be monitored. Strains that are virulent in the host but Lac- in vitro represent in vivo-induced (ivi) genes, which can be cloned for further analysis.
Signature-tagged mutagenesis (STM) is a technique that allows the fate of a large number of different mutant strains to be monitored simultaneously in a living host (Fig. 1B). In STM, every mutation that is generated carries a different sequence tag that allows mutants to be differentiated from each other. The tags are short segments of DNA containing a 40-bp variable central region flanked by invariant ‘arms’ that facilitate the co-amplification and labelling of the central portions by PCR. Tagged mutant strains are combined and inoculated into an animal. After infection is established, bacteria are isolated from the animal; mutants that are attenuated will not be recovered from the animals. Comparison of tags that are present in the inoculum but absent from the recovered bacteria identifies attenuated mutants. STM is, therefore, a negative selection technique based on the fact that the majority of random insertional mutations cause a loss-of-function phenotype.

STM of S. typhimurium identified both known virulence genes and a novel type III secretion system essential for the systemic growth of this pathogenic microbe. Type III secretion systems consist of a large number of proteins involved in the secretion and translocation of virulence determinants into eukaryotic cells in a contact-dependent manner. 

**Fig. 2** Categories of bacterial genes identified by STM and IVET. The outer circle encloses all the genes in a bacterial pathogen, and the inner solid circle encloses all genes required for virulence that are potential targets for antimicrobials. STM (pale grey area) identifies a subset of the genes required for virulence; mutations in essential genes (dark grey area) will not be viable, and some mutations, although in genes involved in pathogenesis, will not be attenuated in mixed infections. IVET (dashed outline) identifies genes that are transcribed in the host but not under laboratory conditions; essential genes are transcribed both on laboratory media and in the host.
Components of type III secretion systems are conserved in *Yersinia*, *Shigella*, *Salmonella*, *Pseudomonas*, and *E. coli*, and are often located on pathogenicity islands. The type III systems are of special interest for drug and vaccine development as they constitute a potential pathogen-specific target present in a range of disease-causing bacteria. Pathogen-specific targets may be preferable to targets also found in commensal species as selection against a sub-set of the microbial flora should make resistance less likely. Furthermore, type III systems could be adapted to deliver heterologous antigens for vaccination or inactivated to generate avirulent bacterial strains as live, attenuated vaccines.

The broader applicability of STM was assessed using the Gram-positive pathogen *Staph. aureus*. *Staph. aureus* causes a wide range of diseases in humans resulting from either localised or systemic infections, leading to abscesses, endocarditis, pneumonia and other life-threatening diseases. For methicillin resistant *Staph. aureus* that is prevalent in both Europe and the US, the only available forms of treatment are the glycopeptides. Reports of the emergence of vancomycin (the principal glycopeptide in clinical use) resistant strains are of great concern.

STM was used to identify 50 genes of *Staph. aureus* required for virulence in a murine model of bacteraemia. Analysis of the transposon insertion points in the attenuated mutants revealed that approximately half had insertions in genes of unknown function. Most of the remaining insertion points were in genes involved in nutrient biosynthesis and cell surface metabolism, including *femA* and *femB*; a further mutation occurred in a previously unknown gene that shares significant similarity to *femB* but whose function in cell wall structure is unknown. Both *femA* and *femB* are involved in the formation of cell wall peptidoglycan pentaglycine cross-bridges. It is not surprising that such mutants are affected in pathogenesis, because a null mutation in *femAB* results in a strain that has exclusively monoglycine cross-bridges, a slower growth rate, and aberrant cross-walls. Furthermore, many surface proteins are anchored via the pentaglycine cross-bridges, and could be affected in *femA* or *femB* mutants.

Both IVET and STM have their limitations as genetic screens. In the IVET screens performed to date, only promoters that are active in the host and not on laboratory media have been analysed. However, this is an arbitrary exclusion criterion. There is no a priori reason why genes transcribed both during infection and on laboratory media are not of interest in bacterial virulence; indeed, the adhesins expressed by *Shigella* and *Yersinia*, and the type III secretion systems expressed by *Salmonella* fall into this category. Furthermore, once a promoter with the desired pattern of activity has been isolated, it is still necessary to perform mutagenesis and animal studies to evaluate whether the gene is involved in virulence. STM combines the power of mutational analysis with the ability to screen many...
mutant strains in a single assay. However, two categories of virulence genes will not be isolated by STM. Firstly, it will not be possible to obtain complete loss-of-function mutations in essential genes whose products might be useful drug targets. Secondly, some mutant strains, for instance those with insertions in genes encoding secreted bacterial toxins, will be attenuated when tested individually but not when inoculated in a pool of toxin producing strains. STM primarily identifies mutant strains that are viable on laboratory media but defective for growth in a host. IVET and STM are therefore complementary technologies that offer the prospect of rapid identification of bacterial virulence genes (Fig. 2).

**From virulence genes to drug and vaccine discovery**

Mutation of virulence genes frequently produces non-pathogenic strains that are, nevertheless, viable when grown on laboratory media, making them candidates as live, attenuated vaccines. The best examples of live attenuated bacterial vaccines are the two *Salmonella typhi* vaccines that are both immunogenic in human volunteers after administration of a single dose\(^{38,39}\). One vaccine strain has a deletion in the *phoP/phoQ* sensor kinase system that modulates the expression of a large number of virulence genes\(^{38}\). The other vaccine strain carries deletions in genes encoding part of the purine biosynthetic pathway\(^{39}\). Therefore, an understanding of the nutritional requirements of bacteria during infection as well as the microbial regulatory networks of virulence genes have proved valuable in vaccine development. Not only can attenuated strains act as vaccines in their own right, they can also be used to deliver heterologous antigens. An attractive goal is to construct live attenuated vaccines carrying antigens from many pathogens that can be administered as a single dose mucosally\(^{40}\).

The ideal drug target should fulfil the following criteria. The target should: (i) be conserved across a range of pathogenic bacteria; (ii) not be present in eukaryotic cells; (iii) have an activity that can be assayed by *in vitro* tests; and (iv) show an inhibition profile that leads to microbial death or loss of pathogenicity.

Much interest continues to focus on the metabolic pathways that bacteria require during growth in the host. A good example of this is the aromatic amino acid biosynthetic pathway that is present in microbes but not in eukaryotic cells. A number of antimicrobial compounds (e.g. sulphonamides, trimethoprim, dapsone) are active against enzymes involved in terminal steps in this pathway. Many useful antimicrobials, such as the β-lactams and the glycopeptides, act on components unique to the bacterial cell membrane, and are selectively active against
Impact of genomics on healthcare

microbes. There are additional cell surface structures and enzymes that catalyse the synthesis of these structures that could be targets for effective antimicrobials. The significantly greater β-lactam susceptibility of femA and femB mutants has suggested that the FemA and FemB proteins are potential targets for drugs that could restore sensitivity to β-lactam antibiotics, though such drugs may prove to be useful antibiotics in their own right. There is also interest in developing agents that act against bacterial LPS.

Most of the work on drug development has relied upon the ability to assay the activity of potential drug targets and then to screen chemical libraries for compounds that inhibit the activity. This approach has proved extremely valuable in the past, and is the basis for all the classes of antimicrobials in current use. New compounds have been derived by chemical modification of the structure of these classes and re-testing. However, in recent years, repeated screening of compound libraries has led to the re-isolation of the same classes of inhibitors. A new strategy is being developed that relies upon determining the detailed structure of the intended drug target to design specific inhibitors. Ligand binding approaches have not as yet resulted in the licensing of any new drugs, but there is hope in the pharmaceutical industry that this approach will yield dividends in terms of new drugs. A good example of the application of ligand binding is the recent solving of the structure of AroM bound to an inhibitor. AroM is a pentavalent enzyme in aromatic amino acid biosynthesis that is not present in eukaryotic cells. Detailed analysis of the structure of this molecule has led to the synthesis of inhibitors that can block its activity. The demonstration of inhibition in vitro is a long way from having an effective antibiotic; compounds must reach the site of infection in the host, enter the bacterial cell, and pass all the stringent safety testing laid down for new compounds. However, the approach marks a way forward in generating novel compounds for evaluation.

Conclusion

The race between drug development and bacterial evolution is currently being led by the microbes. However, the rapid increase in our understanding of molecular mechanisms of bacterial pathogenesis, made possible by the genome sequencing projects and new genetic methods, should provide an abundance of information for drug and vaccine discovery. The next challenge will be to convert this information into practical applications to reduce the burden of bacterial diseases.
References


6. Fischer D, Eisenberg D. Assigning folds to the proteins encoded by the genome of *Mycoplasma genitalium*. *Proc Natl Acad Sci USA* 1997; 94: 11299-34


Impact of genomics on healthcare

29 Mahan MJ, Slauoch JM, Mekalanos JJ. Selection of bacterial virulence genes that are specifically induced in host tissues. *Science* 1993; 259: 686–8
38 Hohmann EL, Oletta CA, Killeen KP, Miller SI. *phoP/phoQ*-deleted *Salmonella typhi* (Ty800) is a safe and immunogenic single-dose typhoid fever vaccine in volunteers. *J Infect Dis* 1996; 173: 1408–14