Anti-Bacterial Drug Discovery Using Systems Biology

K.D. Kumble*

Genesis Research and Development Corporation Limited, One Fox Street, Parnell, Auckland, New Zealand

Abstract: The pipeline for new antibacterials is bleak despite the fact that infectious diseases account for a quarter of all worldwide deaths due to disease. Bacteria are ideal organisms for a systems biology approach to understanding pathogenesis by combined use of genomic technologies and computer algorithms. This approach can be applied to identify control points in molecular networks, which could be targets for novel drugs.

Key Words: Genomics, systems biology, molecular networks, antibacterials, pathogens.

DISEASE TRENDS

Infectious diseases are the second leading cause of death with an estimated 14 million deaths each year, ninety percent of which occur in third world countries [1]. This is largely due to the breakdown of public health measures in these countries resulting in an uneven supply of antibiotics. Mortality due to infectious disease has been increasing even in developed countries, with an annual rate of 4.8% in the United States [2]. It has been reported that nearly 2 million people contract infections in hospitals and nearly five percent of them die as a result of this [3]. In 1998, the World Health Organization (WHO) estimated that three bacterial diseases accounted for half the deaths due to infectious disease worldwide- pneumonia, 3.5 million; diarrhoeal disease, 2.2 million and tuberculosis, 1.5 million [1].

The problem is being worsened by resistance among pathogens previously controlled by antibiotics. Initially, drug-resistant bacteria developed in hospital settings where the use of antibiotics was higher than in the general population, such as resistance of *Streptomyces pyogenes* to sulfonamides in military hospitals during the 1930s [4]. Since then resistance has been growing, with pathogens resistant to multiple drugs such as vancomycin-resistant *Staphylococcus aureus* and multi-drug resistant tuberculosis emerging as infections that will give rise to major public health concerns [5].

The last four decades has seen the recognition of new infectious diseases as well as diseases not identified previously to have a pathogenic etiology: Lyme disease caused by *Borrelia burgdorferii*, and *Helicobacter pylori* as the cause of peptic ulcers. Also, diseases believed to be under control in most parts of the world such as tuberculosis and dengue fever are re-emerging [6]. The Institute of Medicine identified six factors (Table **1**) influencing the emergence of infectious disease all of which relate to change- societal and technological, which will continue in this century. One of the important lessons from these trends is that infectious diseases will continue to emerge and strategies must be developed to respond to these threats rapidly.

*Address correspondence to this author at the Genesis Research and Development Corporation Limited, One Fox Street, Parnell, Auckland, New Zealand; E-mail: a.kumble@genesis.co.nz

Table 1. Factors Giving Rise to Emergence of Infectious Disease

Microbial adaptation and change
Breakdown of public health measures
Environmental change and land use
Changes in technology and industry
International trade and travel
Changes in demographics and behaviour

Increasing threat of infectious disease is due to increase in host susceptibility and emergence of drug resistant bacteria. Adapted from Ref. [6].

TREATMENT OPTIONS

Most of the currently used antibiotics were discovered prior to 1970. These belong to eleven main structural classes each of which target one of four main bacterial biosynthetic pathways: cell wall, DNA, protein or folic acid synthesis. During that period the emphasis was on development of broad spectrum agents which offer the advantage of immediate treatment without the need for identification of the causative agent. Since the 70's two new structural classes of antibacterials (oxazolidinones and lipopeptides) have been approved by the FDA, of which only one acts on a new target-daptomycin, a lipopeptide, binds to membrane components disrupting efflux pumps. Also, a drug belonging to a new class of macrolides, ketolides (Ketek), has been put on the market. Table **2** lists currently marketed structural classes of antibacterials and their cellular targets.

Resistance problems have seriously limited antibiotic options for a number of highly prevalent infections. Over seventy percent of all infectious bacteria are resistant to at least one commonly used antibiotic. Moreover, resistance is predicted to increase by four to seven percent annually. Resistance has developed to every class of antibiotic over a course of more than a decade or sometimes within a year of being marketed such as for linezolid [8, 9].

With the rising incidence of antibiotic-resistant infections, it might be practical to discover and develop narrow-spectrum antibacterials especially with the development of molecular diagnostic techniques that can rapidly identify the

1389-5575/06 $50.00+.00

Table 2. Structural Classes of Antibiotics with Their Targets and the Synthetic Process that they Affect

Structural class	Example	Target	Cellular process affected by action
β-lactam	Penicillin	Transpeptidase	Cell wall biosynthesis
Cephalosporin	Cexifime	Transpeptidase	Cell wall biosynthesis
Glycopeptides	Vancomycin	Transglycosylase	Cell wall biosynthesis
Phenylpropanoids	Chloramphenicol	Peptidyl transferase	Protein synthesis
Aminoglycosides	Gentamicin	30S ribosomal subunit	Protein synthesis
Macrolides	Erythromycin	50S ribosomal subunit	Protein synthesis
Oxazolidinone	Linezolid	RNA polymerase	Protein synthesis
Rifampicin	Rifampin	RNA polymerase	Protein synthesis
Quinolone	Ciprofloxacin	DNA gyrase	DNA synthesis
Nitroimidazole	Metronidazole	DNA strands	DNA synthesis
Sulfonamide	Sulfamethoxazole	Dihydropteroate synthase	Folate synthesis
Diaminopyrimidine	Trimethoprim	Dihydrofolate reductase	Folate synthesis
Phospholipid	Daptomycin		Cell membrane

causative agent. These agents might also be used in combination therapies, precedents for which are certainly increasing in the area of viral infections and have also been successfully used for bacterial pathogens. Augmentin, a combination of amoxicillin and clavulanate, has annual sales of over $1 billion.

ANTIBACTERIAL DISCOVERY PIPELINE IS BLEAK

The antibiotic era, coupled with immunization programs, significantly decreased mortality from infectious diseases especially in developed countries through the mid parts of the last century. This decrease has been coincident with increasing mortality due to chronic diseases such as cancer, stroke and heart disease causing most large pharmaceuticals to refocus their efforts on developing drugs for these diseases. Efforts in antibacterial development, meanwhile, began to be centred towards making chemical modifications in structural scaffolds of the existing eleven classes of antibacterials resulting in the development of over hundred antibacterial agents [9]. No new antibacterial class was discovered throughout the 70s, 80s and 90s.

Currently antibiotics are the second most frequently prescribed category of drugs with total sales of just over $27 billion. However, this group represents only 10.6% of the total pharmaceutical market [10, 11]. The short acute nature of antibacterial therapy lowers the return on investment compared to drugs used for treatment of chronic conditions, thus reducing the incentives for pharmaceutical companies to invest in this area.

Out of over 600 drugs currently in clinical trials, eighteen new antibacterials are being tested of which only four belong to new classes [3]. Additionally, none of the new antibiotics-those approved in the last decade and those currently under development- show improved activity against Gram-negative, multi-drug resistant pathogens [12]. According to a joint report by the WHO and the pharmaceutical industry, one of the main barriers limiting Research and Development on infectious diseases is a lack of basic understanding of these diseases making the prospect of finding new medicines difficult [13].

GENOMIC TECHNOLOGIES

The synergistic use of "omic" tools- genomic, transcriptomic, proteomic and metabolomic- provide an opportunity for understanding not only the functioning of bacterial cells, but also how these functions are integrated. These tools are especially suited to study bacteria due to the relatively small size of the bacterial genome, and the tractability of these organisms to biochemical and genetic analysis.

DNA sequencing has provided information on more than three hundred bacterial genomes, over a hundred of which are human pathogens [14]. These gene catalogs are an important first step in understanding the particularity of pathogens. In addition to predicting the number of coding sequences in the genome, sequence information can be used to identify virulence proteins by searching for motifs such as membrane-spanning domains and secretion signals. Searches can also be done for tandem repeats, changes in the number of copies of which alter expression of virulence genes leading to antigenic or other types of variation. Comparison of sequences between strains of the same species that infect different tissues can identify genes that are different, providing insights into mechanisms of infectivity or help in identifying regions of variability. This information is valuable in target selection for therapy. Accessibility to sequence information and tools for sequence analysis, which for the most part are freely available, makes this a routine tool in studies of microbial pathogens [15].

Sequence information can also be used to create DNA microarrays for measuring total cellular transcript expression levels. This approach permits investigation of up- or down-regulation of genes in a variety of contexts enabling assignment of putative function to novel genes. Information obtained from expression data can be used to infer molecular networks by identifying co-regulatory genes. This has been used to investigate various aspects of host-pathogen interaction, such as mechanisms of host invasion, and pathogen survival strategies in the infected host [16].

A growing transcript profile database of pathogen response to host infection under a variety of conditions and at different stages of disease progression is providing insights into understanding the molecular basis of infection. The mechanisms by which bacteria evade immune recognition by the host, such as interference with antigen processing pathways, have been studied in infection models using macrophages and dendritic cells [17]. Other tissue culture models of infection have been used to obtain expression patterns of transcripts and proteins during various stages of infection. This information has then been used to map out molecular mechanisms that underlie the infection process [18, 19].

Although most microarray data has been used for univariate statistical analysis, more recently integration of expression data with gene ontology functional classification is being increasingly used for pathway analysis [20]. These methods can also be used for interpreting proteomic and metabolomic data.

In addition to transcript profiling, protein expression data also needs to be considered since the amount and activity of proteins cannot be directly inferred from determining the amount of mRNA [21]. The complexity of the bacterial proteome is a few logs of magnitude less than the species of proteins in mammalian cells making it possible to obtain global protein expression profiles using current technologies.

Two-dimensional electrophoresis (2-DE) has been the most commonly used method for profiling protein expression levels, and in combination with mass spectrometry (MS) can reveal the identity of expressed proteins. This technology has been enhanced by the use of robotics which run multiple gels, and pick and analyze hundreds of spots each day. Relatively new technologies such as isotope coded affinity tag (ICAT)-liquid chromatography/mass spectrometry (LC/MS) and intact cell MALDI-TOF (ICM) are also starting to be used to generate protein expression profiles [22, 23].

Protein-protein interactions are a method by which proteins exert their influence in the cellular milieu. These interactions are being identified by initially isolating multi-protein complexes using affinity tags, followed by MS analysis to obtain the identities of the proteins in the complex. Compared with two-hybrid assays, this strategy has the advantage of obtaining complexes from the native environment [21].

In order to understand the molecular mechanisms that facilitate bacterial infection of host cells, the system has to be studied in the context of the infection process. This has been done using relevant tissue culture and animal models of infection, or in the natural host. Techniques for visualizing infected cells and tissues can enable isolation of pathogens from infected cells for molecular analysis. Constitutively expressing Green Fluorescent Protein has been used as a marker to separate cells infected with *Bartonella henselae* from non-infected host cells using a fluorescent-activated cell sorter [24], and more recently to obtain Salmonella from the spleen and caecum of infected mice for proteome analysis [25]. The ability to transform pathogenic bacteria with constitutively expressing reporter genes is extremely useful to identify and isolate pathogens from infected hosts for protein and transcript expression analysis during various stages of infection.

SYSTEMS BIOLOGY

Molecular Networks

There is undoubtedly a need for new classes of antibacterials with novel mechanisms of action rendering them insusceptible to existing mechanisms of resistance. An appropriate combination of these technologies, together with proper experimental design and bioinformatic tools can transform the discovery process by studying, in parallel, complex relationships among molecular pathways and networks [26]. The Holy Grail in systems biology is the integration of information obtained from "top-down" approaches in which mechanistic information is obtained from computer models of disease with that from the "bottom-up" approach where "omics" data is obtained by identification of molecular components of an organism. Systems biology provides information on molecular components, and their functional interplay in biological networks. Phenotypic properties of the biological systems are linked to node degree distribution, mean path length and clustering coefficients- parameters that identify the significance of relationships between various molecules.

This will provide models of pathogenesis, which can be used to obtain an increased understanding of the infection process. These resulting networks provide a testable hypothesis that can be used to determine if the identified proteins are important in pathogen establishment, and disease induction and progression. Key control points in pathways important in pathogenesis could then serve as novel antibacterial targets. These control points are likely to be pathogen-dependent based on the distinct mode in which each pathogen infects the host and then establishes infection by evading the immune system. Drugs to targets identified in this manner are likely to have relatively narrow-spectrum reducing the probability of resistance emergence.

Even though the estimated number of genes based on analysis of pathogen genome sequences can vary 10-fold between various species (480 for *Mycoplasma genitalium* to 4,800 for *Salmonella typhii*), it is still roughly 10-fold lower than the estimated expressed genes in humans. Moreover, the complexity of the human proteome as defined by the number of protein species is at least 2 to 3-logs higher than the bacterial proteome. The comparatively reduced complexity of bacterial genomes, as well as expressed transcripts and proteins, make them suitable for studying molecular networks and dynamics. In the context of a good infection model, systems biology is a reasonable approach for linking molecular networks to the infection process thereby identifying nodes and control points which could be tested as targets for therapeutic intervention.

Animal models are extremely useful in this context since the pathophysiology of infectious disease involves complicated interactions between bacterial products and various kinds of host cells. Also, the pathogen will respond to mediators like inflammatory cytokines produced by the host upon infection [27]. Additionally, virulence is partly determined by the host genetic environment, which makes the host either susceptible, or resistant to infection. Mice make good models since not only does their immune system closely resemble that of humans, but they can often be infected with the same pathogens [28]. Efforts to chart out molecular networks should incorporate the use of animal models early in the process so that the data approximates the human infection processes as closely as possible. In a recent report, Becker *et al.* [25] used mass spectrometry to determine protein expression profiles of Salmonella retrieved from infected mice. This data was analyzed in combination with mutant phenotype analysis and genome sequence comparison to determine metabolic networks in Salmonella during infection. Thirteen previously unexploited targets were identified as a result of this analysis.

Molecular network models have been constructed using data from single platforms- either transcript or protein- obtained from different experimental setups and from different laboratories [29, 30]. These network maps show that most proteins are separated by only a few links, and network hierarchy suggests local clusters to be coordinated by hubs, which are likely to correspond to essential proteins. This observation appears to be true for cells with a range of complexity- from bacteria to human [31].

Most attempts at integrating data from various sources, whether technology platform-derived or from published literature, has been made at the computer level [32]. The challenge is in rethinking experimental design so that data from multiple platforms, including transcript and protein expression, obtained at various stages of infection can be integrated and analyzed together to build network models that describe the dynamic behavior of the cell. Datasets from different technology platforms can be integrated and analyzed if the experiment is initially designed in a manner that provides an appropriate context for comparison. This would require the incorporation of proper controls at the outset into the experimental design so that normalized data from each of the platforms can be used for data integration. Developing techniques for integrating data from various experimental and observational approaches are crucial for progress in developing molecular network models [33]. These models can provide a hypothetical framework for experimental validation, whereby one or more members in identified pathways can be inactivated using one of several available techniques, followed by measuring the effects on other members of the pathway.

Such approaches can identify pathways that are involved in the synthesis of virulence factors or in biofilm production, both of which are important in the maintenance of infection in the host. A recent report has demonstrated the feasibility of screening a small molecule library and identifying an inhibitor of virulence regulation in *Vibrio cholerae* [34]. This molecule, 4-[N-(1,8-naphthalimide)]-n-butyric acid (Virstatin) reduced colonization of the pathogen in mice with established infection by three orders of magnitude compared to untreated controls.

Case Study- *Plasmodium falciparum* Systems Biology

To date, no example of a systems approach to understanding bacterial pathogenicity has been reported wherein genomic data was integrated with transcript and protein expression information to obtain an understanding of the molecular systems involved in the infection process. Such an approach has been initiated for the malaria parasite, the protozoan *Plasmodium falciparum* with an initial study that integrated information obtained from a number of technology platforms.

The complete genome of the malaria parasite, *Plasmodium falciparum* was published in 2002 [35]. Useful insights into Plasmodium gene evolution could be obtained by comparative genomics made possible due to the availability of sequences from other strains of Plasmodium, especially those that are pathogenic to rodents- *P. yeolii*, *P. chabaudi* and *P. berghei* [36].

A first generation microarray representing approximately 6000 open treading frames from the *P. falciparum* genome was constructed and used for characterizing the gene expression profiles of the intraerythrocytic trophozoite and schizont stages of the Plasmodium lifecycle. Data revealed extensive regulation of genes associated with stage-specific processes [37, 38]. Expression profiles of human and mosquito stages of Plasmodium life cycle using high-density oligonucleotide arrays showed highly correlated levels and temporal patterns of expression for genes involved in similar functions or cellular processes [39]. Transcriptome data were obtained from three time points during the G1 phase and from two time points during the S/M phase, as well as from purified mature and immature gametocytes. The profiles at these stages were compared through pair-wise hybridizations to *P. berghei* genomic DNA microarrays [40].

Large-scale mass spectrometric proteome analysis of some stages during *Plasmodium falciparum* life cycle revealed over 1200 proteins, 931 and 645 of which were identified in gametocytes and gametes, respectively. Proteins in these two stages, likely to provide biological insights into the sexual stages of the parasite, include stage-specific, secreted and membrane-associated proteins. Some of these proteins contain domains suggestive of their role in cell-cell interactions, and therefore can be evaluated as potential components of a malaria vaccine formulation. In another proteomic study, four stages of the parasite life cycle (sporozoites, merozoites, trophozoites and gametocytes) were characterized by multidimensional protein identification technology. This study revealed chromosomal clusters encoding co-expressed proteins suggesting a likely mechanism for controlling gene expression during the Plasmodium life cycle [41, 42].

Initial integration of these diverse data sets was performed by Hall *et al.* [40] and their analysis provided useful insights into evolution of stage-specific gene expression due to selective pressures, identification of 4500 conserved genes in the central regions of the 14 chromosomes, and post-transcriptional gene regulation mechanisms.

These studies from Plasmodium demonstrate the feasibility of using a systems approach to understand the molecular data obtained using post-genomic technologies. This data can be integrated and analyzed in a form that provides insight into the biological processes involved during host infection. It also reveals that in order to take advantage of these technologies, efforts and expertise from various research groups needs to be coordinated at the stage of experimental design so that appropriate controls can be put in place to ensure data quality and permit the integration of information from diverse technology platforms. The tractability of bacteria for genetic and biochemical manipulation, in addition to their relatively smaller genomes make them organisms of choice for using such a systems approach for understanding infectivity, identifying proteins critical to that process and thereby identifying drug targets.

CHOICE OF TARGETS

Traditionally identification of potential antibacterial targets has been based on knowledge of bacterial physiology and biochemistry. More recently, high-throughput genomic technologies offer the ability to use a Systems Biology approach to understanding the molecular networks critical for establishment of infection can help in identification of novel targets. An understanding of these molecular events that trigger the initiation and maintenance of the disease process is extremely useful in identification and selection of drug targets.

A widely accepted view has been that for a protein to be a good target it should be (1) universally present among pathogens, (2) essential for growth and, (3) have no human homolog. Typically such targets were families of essential proteins that could be inhibited by a single compound or had invariant active site residues to minimize development of resistance [43]. A systems approach would broaden the quailties of a good target to include proteins that regulate important pathways that help the pathogen establish and maintain infection in the host. Such a target might be pathogen-specific resulting in identification of drugs that have a narrow spectrum of activity. Such treatment options may be feasible in light of advances in molecular diagnostic technologies that can identify the infective agent in the early stages of infection.

However, targets identified by a systems approach would still have to meet the criteria for (1) Novelty. These should not be targeted by currently used antibacterials and have low frequency of resistance among pathogens (2) Selectivity. This will allow specific pathogens to be treated without affecting the normal flora, and (3) Amenable to high throughput screening [44].

The identification of potential anti-microbial targets has to be followed by validation of these targets. This can be done by knockdown strategies such as targeted mutagenesis [45, 46], or by using siRNA [47]. Conditional requirement for a gene can also be tested by growing the bacteria under a variety of conditions to analyze growth behavior [48]. Validation of targets also has the potential for identifying other targets downstream of the same pathway enabling synergy and possibly reducing the development of resistance.

The driving force behind genomics-based target validation is the identification of genes whose loss of function coincides with loss of viability or attenuation of virulence. Gene targets involved in virulence functions are likely to lead to development of narrow-spectrum compounds, which would reduce the probability of resistance emerging because of reduced impact of antibiotic selection on normal flora [49]. Signature-tagged mutagenesis is one of the methods successfully used for identification of virulence genes by *in vivo* analysis [50, 51, 52].

Pathogens in which the target gene has been inactivated can be used for assessing the effect of loss-of-function of that gene on virulence and infectivity. Targets can be further analyzed using focused DNA, and possibly protein arrays to determine the effects of target inactivation on other members of the pathway to which the targets belong using sample sets obtained from mouse models infected with various strains of the pathogen, or from pathogens isolated from infected human patients.

ANTIBACTERIAL SCREENING

Assays

Although the platform technologies described in previous sections have identified a significant number of targets the effort has not yet resulted in discovery of new antibacterials. The limitation is partly due to the difficulty of target screening approaches usng synthetic libraries to identify compounds that can reach targets within bacterial cells. Almost all currently used antibacterials are compounds that were identified using whole cell screening assays. These compounds had no known cellular targets but were selected for development on the basis of inhibiting a broad spectrum of pathogens. This method of screening resulted in the same classes of antibacterials being repeatedly identified [43]. The obvious advantage of whole cell assays is that identified compounds have the ability to enter cells and exert the desired effect, but is followed by a labour intensive target identification effort.

In the 1970s there was a shift towards target-directed screening approaches which were initially applied towards identification of new compounds for old targets, and for enhancing activity against common resistance determinants [53]. Target-based assays are amenable to high throughput and can test hundreds of compounds for inhibitors of enzymes or biochemical pathways. Advantages of these assays which include increased sensitivity for detection of weakly interactive compounds, and facilitation of rational drug design have to be offset by issues such as penetration and insufficient accumulation into bacterial cells [54].

A better understanding of the fundamental underlying biology of systems at the molecular level will result in finding targets whose physiological function can be inhibited by novel classes of drugs with greater specificity, potency and with fewer side effects. After proteins have been identified as potential drug targets, an understanding of their function is important in order to develop target-specific screening assays. Discovery of new antibacterial compounds will depend on the reliability and sensitivity of the assays, which for targets identified using a systems approach, are likely to be

based on cell-free assays, measuring either binding or activity [55]. Upon identification of candidate compounds, specificity of binding can be tested on proteins belonging to the same family using protein microarray-based assays in which multiple proteins have been spotted in the same well.

Hits obtained from high throughput screening of compound libraries can be used for designing an active chemical series to establish lead compounds. Alternatively, structural information for target proteins may already be available for use in rational drug design or for virtual screening [48]. Medicinal chemistry can then used to increase potency, bioavailability and pharmacokinetic properties. These steps can be geared towards improving the relationship between target inhibition and antimicrobial action by increasing cell penetration while reducing drug efflux.

Novel Antibacterial Compounds

Screening compound collections is the first step in selecting drug candidates for development, since these collections contain compounds with disparate properties. Selected compounds can then be optimized by medicinal chemistry to yield a drug with the desired pharmacokinetic and pharmacodynamic properties [56]. An ideal antibacterial compound should be (i) target-specific (ii) unaffected by current bacterial resistance mechanisms, and (iii) have no adverse effects on the host. In addition, the compound must be (iv) able to access the target within the bacteria by traversing the cell wall and membrane(s), and (v) achieve high concentrations within the bacterial cell by overcoming active efflux systems.

To date, a vast majority of antibacterials have been identified from natural products [57], and currently of the nineteen antibacterials in clinical development, thirteen are of natural product origin [58]. It is possible that natural antibiotics occur for only twenty or so gene products since evolution selected targets that could only inhibit growth of competitive organisms. For targets avoided by antibiotic-producing organisms, chemically synthesized compounds may prove to be a better source of inhibitors. It would therefore be preferable to screen non-natural compounds to obtain the next generation of antibacterials, since these would be foreign to the pathogens and the chances of a microbe having enzymes to destroy these compounds are minimal, although drug efflux may be an issue for some of these compounds [49]. Such a strategy has resulted in the development of a new class of compounds-diarylquinolines- identified in a screen using live mycobacteria [59].

In fact the reverse can also be done- a structurally defined target can be used to design an inhibitor- using a strategy such as heteronuclear-NMR-based screening. In this method, target proteins are screened against a compound library to identify binding sites on the protein surface. Computational analysis then provides a quantitative assessment of the affinity and specificity of a binding site to a compound [60]. Identification of a set of compounds can then be followed by combinatorial chemistry for synthesizing a small library based around the inhibitor for screening. Such an approach has resulted in development of novel peptide deformylase inhibitors [61, 62], which were initially identified by using a combination of comparative genomics for target validation and rational drug design for identification of a therapeutic candidate on which to base medicinal chemistry for improving pharmacokinetic properties.

ATP Synthase Inhibitors

A medium-throughput screen for new anti-tuberculosis compounds was set up with selected prototypes of different chemical series using live mycobacteria. This screen resulted in a number of hits, each of which belonged to the diarylquinoline (DARQ) family [59]. Chemical optimization of a lead compound (Fig. (**1A**)) led to a series of DARQs with potent *in vitro* activity against several mycobacteria, including *M. tuberculosis*, three of which were shown to have *in vivo* activity. The most active compound of this class, R207910 (Fig. **(1B)**), was isolated from a mixture of four diasterioisomers prepared in five steps. No DARQ resistant strain was identified from screening 30 isolates of multi-drug resistant *M. tuberculosis*. DARQs had much higher minimum inhibitory concentrations (MICs) for a wide variety of other pathogenic bacteria tested [59].

In order to identify the drug mechanism of action, genome analysis of mycobacterial strains resistant to the drug was carried out. This analysis highlighted the gene coding for the F_0 subunit of ATP synthase, indicating that R207910 inhibits the proton pump of *M. tuberculosis* ATP synthase. Complementation studies have verified that this protein is responsible for resistance to R207910. The target and mechanism of action of R207910 are different from those of other anti-tuberculosis agents. Comparison of ATP synthase sequence from different bacteria and eukaryotes provides a rationale for the specificity and the safety profile of R207910 as an anti-tuberculosis agent. Furthermore the distinct target of R207910 makes it unlikely for the compound to have cross-resistance with other existing tuberculosis drugs.

Peptide Deformylase Inhibitors

Bacterial protein synthesis is initiated with N-formylmethionine, which is then subsequently removed by sequential action of peptide deformylase (PDF) and methionine amino peptidase. This formylation-deformylation cycle is important for bacterial growth and is conserved among all bacterial species [61]. Mammalian protein synthesis initiates with methionine and although mammalian gene sequences with homology to PDF exist their exact role remains unclear. This metalloenzyme essential for bacterial growth has been crystallized from several bacterial species and its structure determined. Briefly, *E. coli* PDF contains 3 major helices, 3 β-sheet regions and a short helix [62]. The geometry and chemistry around the active-site metal of the enzyme has similarities to proteases in the thermolysin family making it a promising selective target for antibacterial drug development [62, 63].

Since PDF is a metalloenzyme rational mechanism based strategies used successfully for other matrix metalloproteases were adopted for designing PDF inhibitors. A generic PDF inhibitor structure shown in Fig. (**1C**) was proposed containing a chelating pharmacophore and an *n*-butyl group that mimics the methionine side chain of the substrate [64]. Several types of potent PDF inhibitors were identified using this

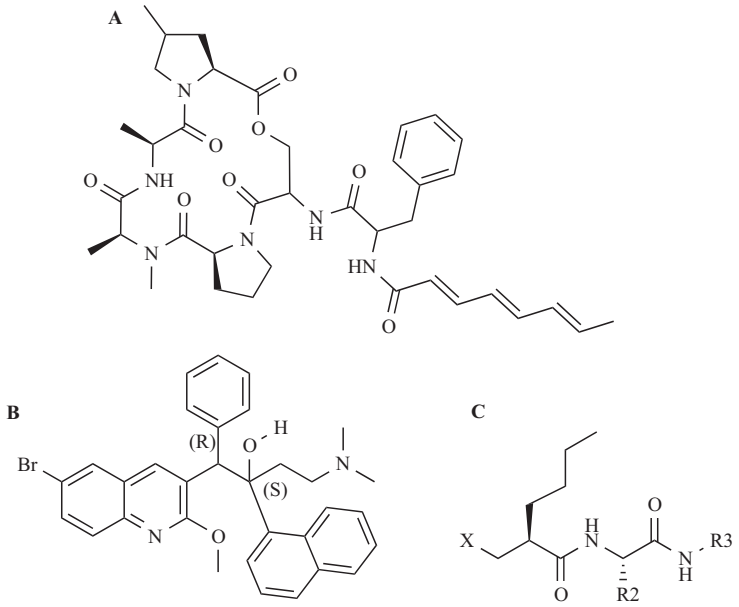

Fig. (1). Structures of novel drug classes for three new antibacterial targets. (A) ADEP1 or 'Factor A' is the main component of the A54556 complex isolated from fermentation broth of *S. hawaiiensis*. (B) R207910, a diarylquinoline isomer isolated from a synthetic mixture of four isomers. (C) Generic peptide deformylase inhibitor structure with X representing a chelating pharmacophore, and R2 and R3 are inhibitor regions that provide favorable pharmacokinetic properties.

structural and mechanistic information by high throughput screening. A significant number of inhibitors had no antibacterial activity either because of their inability to penetrate the bacterial cell, or due to poor pharmacokinetic properties. The frequency of mutations resistant to PDF inhibitors was high in the laboratory mainly due to inactivation of the bacterial transformylation gene that bypasses the need for PDF activity. However these mutants are attenuated for virulence *in vivo* suggesting that resistance may not be an issue for PDF to be a valid antibacterial therapeutic target.

Three groups of PDF inhibitors have been shown to have *in vivo* activity- alkyl succinate-proline hydroxamates, N-formyl hydroxylamines with tert-butyl at P2', and N-alkyl urea hydroxamates- the former two are currently in clinical trials [62]. All three groups display excellent selectivity for PDF and have been demonstrated to inhibit bacterial growth through inhibition of PDF activity. Also, preliminary data suggests that bacteria resistant to other antibacterial classes do not show cross resistance to these PDF inhibitors. These drugs represent an example of applying mechanism-based rational drug design in screening focused libraries for the identification of therapeutically valid antibacterial agents [62].

PERSPECTIVES

The key to future discovery of infectious disease treatment lies in application of new technologies that utilize global approaches, integrating information available from sequenced pathogen genomes, with transcript and protein profiling and functional genomics of microbial infections. Such a systems approach can be used to understand the interactions between pathogens and their hosts during the infection process. Computer-based translation of information into knowledge, will offer insights leading to identification of new targets in bacteria as antibacterials. Focus on a few high-quality targets screened with novel compound libraries can help in development of new antibacterial classes.

The old paradigm of developing broad-spectrum or cost-cutting drugs that could be used against multiple pathogens must give way to niche markets that develop highly selective drugs taking advantage of the particularity of individual pathogens. This could be especially significant in organisms such as *M. tuberculosis*, a major problem in increasing numbers of HIV-positive patients. The identification of diarylquinolines as specific inhibitors of mycobacterial ATP synthase demonstrates the feasibility of successfully killing specific pathogens by inhibiting targets that, although ubiquitous, have sufficient sequence diversity to be specific to the organism [65]. This may be less likely to contribute to the general rising level of antibiotic resistance of unrelated pathogens by not affecting normal flora [66, 67]. Additionally, combinations of drugs that target control points in two or more pathogenesis-specific pathways would delay development of resistance.

Infectious diseases will continue to remain a public health problem with a continual evolution of emerging and reemerging infectious diseases. In order to effectively address this problem a common strategy is required that uses the latest technologies which can be broadly applied to all pathogens for rapid identification of novel antibacterial agents. Sequencing, transcript and protein profiling, combinatorial chemistry and high-throughput screening are being established as technologies that can be used to link targets of interest to compounds. These can transform conventional anti-

bacterial screening into rational systems-based drug discovery programs with wide applicability. This would make the development of new antibacterials cheaper and bring us closer to the aim of providing these drugs to third world countries at affordable prices.

REFERENCES

[1] Haden, A. The World Health Report 2000- Health systems: Improving performance, World Health Organization: Geneva, **2000**.
[2] Armstrong, G. L.; Conn, L. A.; Pinner, Y. *J. Am. Med. Assoc*., **1999**, *281*, 61.
[3] Barrett, J. F. *Curr. Opin. Microbiol*., **2005**, *8*, 498.
[4] Levy, S. B.; Marshall, B. *Nat. Med*., **2004**, *10*, S122.
[5] Hughes, D. *Nat. Rev. Genet*., **2003**, *4*, 432.
[6] Cohen, M. *Nature*, **2000**, *406*, 762.
[7] Lederberg, J.; Shope, R. E.; Oaks, S. C. *Emerging Infections: Microbial Threats to Health in the United States*. National Academy Press: Washington D.C., **1992**.
[8] Wenzel, R.P. *N. Engl. J. Med*., **2004**, *351*, 523.
[9] Walsh, C. *Nat. Rev. Microbiol*., **2003**, *1*, 65.
[10] McCaig, L. F.; Hughes, J. M. *JAMA*, **1995**, *273*, 214.
[11] Trouiller, P.; Torreele, E.; Olliaro, P.; White, N.; Foster, S.; Wirth, D.; Pecoul, B. *Trop. Med. Int. Health*, **2001**, *6*, 945.
[12] Norrby, S. R.; Nord, C. E.; Finch, R. *Lancet Infect. Dis*., **2005**, *5*, 115.
[13] Webber, D.; Kremer, M. *Bull. WHO*, **2001**, *79*, 735.
[14] http://ncbi.nlm.nih.gov/genomes/lproks.cgi
[15] Salzberg, S. L.; Delcher, A. L.; Kasif, S.; White, D. *Nucleic. Acids Res*., **1998**, *26*, 544.
[16] Bryant, P. A.; Venter, D.; Robins-Browne, R.; Curtis, N. *Lancet*, **2004**, *4*, 100.
[17] Harding, C. V.; Ramachandra, L.; Wick, M. J. *Curr. Opin. Immunol*., **2003**, *15*, 112.
[18] Huang, Q.; Liu D., Majewski, P.; Schulte, L. C.; Korn, J. M.; Young, R. A.; Lander, E. S.; Hacohen, N. *Science*, **2001**, *294*, 870.
[19] Florens, L.; Washburn, M. P.; Raine, J. D.; Anthony, R. M.; Grainger, M.; Haynes, J. D.; Moch, J. K.; Muster, N.; Sacci, J. B.; Tabb, D. L.; Witney, A. A.; Wolters, D.; Wu, Y.; Gardner, M. J.; Holder, A. A.; Sinden, R. E.; Yates, J. R.; Carucci, D. J. *Nature*, **2002**, *419*, 520.
[20] Curtis, R. K.; Oresic, M.; Vidal-Puig, A. *Trends Biotechnol*., **2005**, *23*, 429.
[21] Griffin, T. J.; Gygi, S. P.; Ideker, T.; Rist, B.; Eng, J.; Hood, L.; Aebersold, R. *Mol. Cell Proteomics*, **2002**, *1*, 323.
[22] Aebersold, R., Mann, M. *Nature*, **2003**, *422*, 198.
[23] Fenselau, C.; Demirev, P. A. *Mass Spectrom. Rev*., **2001**, *20*, 157.
[24] Lee, A. K.; Falkow, S. *Infect. Immun*., **1998**, *66*, 3964.
[25] Becker, D.; Selbach, M.; Rollenhagen, C.; Ballmaier, M.; Meyer, T. F.; Mann, M.; Bumann, D. *Nature*, **2006**, *440*, 303.
[26] Kumble, K. D. *Drug Disc. Dev*., **2006**, *9*, 30.
[27] Knudsen, J. D.; Frimodt-Møller, N. *Animal Testing in Infectiology*, Karger Publishers: Basel, **2001**.
[28] Flajnik, M. F. *Nat. Rev. Immunol*., **2002**, *2*, 688.
[29] Starck, J.; Callard, R.; Hubank, M. *Trends Biotechnol*., **2003**, *21*, 290.
[30] Steffen, M.; Petti, A.; Aach, J.; D'haeseleer, P.; Church, G. *BMC Bioinformatics*, **2002**, *3*, 34.1.
[31] Hegde, P. S.; White, I. R.; Debouck, C. *Curr. Opin. Biotechnol*., **2003**, *14*, 647.
[32] Skipper, M. *Nat. Rev. Genet*., **2005**, *6*, 800.
[33] Searls, D. B. *Nat. Rev. Drug Discov*., **2005**, *4*, 45.
[34] Hung, D. T.; Shakhnovich, E. A.; Pierson, E.; Mekalanos, J. J. *Science*, **2005**, *310*, 670.
[35] Gardner, M.J.; Hall, N.; Fung, E.; White, O.; Berriman, M.; Hyman, R.W.; Carlton, J.M.; Pain, A.; Nelson, K.E.; Bowman, S.; Paulsen, I.T.; James, K.; Eisen, J.A.; Rutherford, K.; Salzberg, S.L.; Craig, A.; Kyes, S.; Chan, M.S.; Nene, V.; Shallom, S.J.; Suh, B.; Peterson, J.; Angiuoli, S.; Pertea, M.; Allen, J.; Selengut, J.; Haft, D.; Mather, M.W.; Vaidya, A.B.; Martin, D.M.; Fairlamb, A.H.; Fraunholz, M.J.; Roos, D.S.; Ralph, S.A.; McFadden, G.I.; Cummings, L.M.; Subramanian, G.M.; Mungall, C.; Venter, J.C.; Carucci, D.J.; Hoffman, S.L.; Newbold, C.; Davis, R.W.; Fraser, C.M.; Barrell, B. *Nature*, **2002**, *419*, 498.
[36] Carlton, J.M.; Angiuoli, S.V.; Suh, B.B.; Kooij, T.W.; Pertea, M.; Silva, J.C.; Ermolaeva, M.D.; Allen, J.E.; Selengut, J.D.; Koo, H.L.; Peterson, J.D.; Pop, M.; Kosack, D.S.; Shumway, M.F.; Bidwell, S.L.; Shallom, S.J.; van Aken, S.E.; Riedmuller, S.B.; Feldblyum, T.V.; Cho, J.K.; Quackenbush, J.; Sedegah, M.; Shoaibi, A.; Cummings, L.M.; Florens, L.; Yates, J.R.; Raine, J.D.; Sinden, R.E.; Harris, M.A.; Cunningham, D.A.; Preiser, P.R.; Bergman, L.W.; Vaidya, A.B.; van Lin, L.H.; Janse, C.J.; Waters, A.P.; Smith, H.O.; White, O.R.; Salzberg, S.L.; Venter, J.C.; Fraser, C.M.; Hoffman, S.L.; Gardner, M.J.; Carucci, D.J. *Nature*, **2002**, *419*, 512.
[37] Bozdech, Z.; Llinas, M.; Pulliam, B. L.; Wong, E. D.; Zhu, J.; DeRisi, J. L. *PLoS Biol.*, **2003**, *1*, E5.
[38] Bozdech, Z.; Zhu, J.; Joachimiak, M. P.; Cohen, F. E.; Pulliam, B.; DeRisi, J. L. *Genome Biol*., **2003**, *4*, R9.
[39] Le Roch, K. G.; Zhou, Y.; Blair, P. L.; Grainger, M.; Moch, J. K.; Haynes, J. D.; De La Vega, P.; Holder, A. A.; Batalov, S.; Carucci, D. J.; Winzeler, E. A. *Science*, **2003**, *301*, 1503.
[40] Hall, N.; Karras, M.; Raine, J.D.; Carlton, J.M.; Kooij, T.W.; Berriman, M.; Florens, L.; Janssen, C.S.; Pain, A.; Christophides, G.K.; James, K.; Rutherford, K.; Harris, B.; Harris, D.; Churcher, C.; Quail, M.A.; Ormond, D.; Doggett, J.; Trueman, H.E.; Mendoza, J.; Bidwell, S.L.; Rajandream, M.A.; Carucci, D.J.; Yates, J.R. 3rd; Kafatos, F.C.; Janse, C.J.; Barrell, B.; Turner, C.M.; Waters, A.P.; Sinden, R.E. *Science*, **2005**, *307*, 82.
[41] Lasonder E.; Ishihama Y.; Andersen J. S.; Vermunt A. M.; Pain A.; Sauerwein R. W.; Eling W. M.; Hall N.; Waters A. P.; Stunnenberg H. G.; Mann M. *Nature*, **2002**, *419*, 537.
[42] Florens, L.; Washburn, M. P.; Raine, J. D.; Anthony, R. M.; Grainger, M.; Haynes, J. D.; Moch, J. K.; Muster, N.; Sacci, J.B.; Tabb, D. L.; Witney, A. A.; Wolters, D.; Wu, Y.; Gardner, M. J.; Holder, A. A.; Sinden, R. E.; Yates, J. R.; Carucci, D. J. *Nature*, **2002**, *419*, 520.
[43] Miesel, L.; Greene, J.; Black, T. A. *Nat. Rev. Genet*., **2003**, *4*, 442.
[44] Chan, P.F.; Macarron, R.; Payne, D. J.; Zalacain, M.; Holmes, D. J. *Curr. Drug Targets Infect. Disord.*, **2002**, *2*, 291.
[45] Lacks, S. A. *Microb. Drug Resist*., **1997**, *3*, 327.
[46] Cregg, K. M.; Wilding, I.; Black, M. T. *J. Bacteriol*., **1996**, *178*, 5712.
[47] Yanagihara, K.; Tashiro, M.; Fukuda, Y.; Ohno, H.; Higashiyama, Y.; Miyazaki, Y.; Hirakata, Y.; Tomono, K.; Mizuta, Y.; Tsukamoto, K.; Kohno S. *J. Antimicrob. Chemother*., **2006**, *57*, 122.
[48] Black, M. T.; Hodgson, J. *Adv. Drug Deliv. Rev*., **2005**, *57*, 1528.
[49] Knowles, D. J. *Trends Microbiol*., **1997**, *5*, 379.
[50] Hensel, M.; Shea, J. E.; Gleeson, C.; Jones, M. D.; Dalton, E.; Holden, D.W. *Science*, **1995**, *269*, 400.
[51] Hava, D. L.; Camilli, A. *Mol. Microbiol*., **2002**, *45*, 1389.
[52] Sun, Y. H.; Bakshi, S.; Chalmers, R.; Tang, C. M. *Nat. Med*., **2000**, *6*, 1269.
[53] Imada, A.; Hotta, K. *Emerging Targets in Antibacterial and Antifungal Chemotherapy*. Chapman and Hall: New York, **1992**.
[54] Rosamund, J.; Allsop, A. *Science*, **2000**, *287*, 1973.
[55] Donadio, S.; Carrano, L.; Brandi, L.; Serina, S.; Soffientini, A.; Raimondi, E.; Montanini, N.; Sosio, M.; Gualerzi, C. O. *J. Biotechnol*., **2002**, *99*, 175.
[56] Thomson, C. J.; Power, E.; Ruebsamen-Waigmann, H.; Labischinski, H. *Curr. Opin. Microbiol*., **2004**, *7*, 445.
[57] Newman, D. J.; Cragg, G. M.; Snader, K. M. *J. Nat. Prod*., **2003**, *66*, 1022.
[58] Bush, K.; Macielag, M.; Weidner-Wells, M. *Curr. Opin. Microbiol.*, **2004**, *7*, 466.
[59] Andries, K.; Verhasselt, P.; Guillemont, J.; Gohlmann, H. W.; Neefs, J. M.; Winkler, H.; Van Gestel, J.; Timmerman, P.; Zhu, M.; Lee, E.; Williams, P.; de Chaffoy, D.; Huitric, E.; Hoffner, S.; Cambau, E.; Truffot-Pernot, C.; Lounis, N.; Jarlier, V. *Science*, **2005**, *307*, 223.
[60] Clements, J. M.; Beckett, R. P.; Brown, A.; Catlin, G.; Lobell, M.; Palan, S.; Thomas, W.; Whittaker, M.; Wood, S.; Salama, S.; Baker, P. J.; Rodgers, H. F.; Barynin, V.; Rice, D. W.; Hunter, M. G. *Antimicrob. Agents Chemother*., **2001**, *45*, 563.
[61] Hajduk, P. J., Huth, J. R., Fesik, S. W. *J. Med. Chem*., **2005**, *48*, 2518.
[62] Jain, R.; Chen, D.; White, R. J.; Patel, D. V.; Yuan, Z. *Curr. Med. Chem*., **2005**, *12*, 1607.

[63] Lee, M. D.; Antczak, C.; Li, Y.; Sirotnak, F. M.; Bornmann, W. G.; Scheinberg, D. A. *Biochem. Biophy. Res. Commun*., **2003**, *312*, 309.
[64] Chan, M. K.; Gong, W.; Rajagopalan, P. T.; Hao, B.; Tsai.; C. M.; Pei, D. *Biochemistry*, **1997**, *36*, 13904.
[65] Dardel, F.; Ragusa, S.; Lazennec, C.; Blanquet, S.; Meinnel, T. *J. Mol. Biol*., **1998**, *280*, 501.
[66] Yuan, Z.; Trias, J.; White, R. J. *Drug Discov. Today*, **2001**, *6*, 954.
[67] Gootz, T. D. *Clin. Microbiol. Rev*., **1990**, *3*, 13.

Received: April 13, 2006 **Revised: May 30, 2006** **Accepted: May 31, 2006**