CMLS, Cell. Mol. Life Sci. 56 (1999) 779–787
1420-682X/99/100779-09 $ 1.50 + 0.20/0
© Birkhäuser Verlag, Basel, 1999

CMLS Cellular and Molecular Life Sciences

Bacterial targets and antibiotics: genome-based drug discovery

C. P. Gray* and W. Keck

F. Hoffmann-La Roche Ltd, Pharma Research Preclinical, Infectious Diseases, CH-4070 Basel (Switzerland), Fax +41 61 688 27 29, e-mail: christopher.gray@roche.com

Abstract. The requirement for novel classes of antibiotics to combat the emergence of resistant and multi-resistant bacteria has coincided with the completion sequencing of a number of bacterial genomes. The in silico analysis of these genomes coupled with innovative genetic manipulation has already led to the identification of conserved essential (either in vitro or in vivo, depending on the methodology) genes that are potential targets for antibacterial research. New technologies, made possible by access to the genomic sequences, are capable of simultaneously quantifying almost the entire complement of gene products synthesised by bacterial cells. These technologies are opening up the way for the analysis of expression patterns elicited in cells in response to changes in their environment. The integration of these technologies into the drug discovery process is still in its infancy and the potential wealth of information, some of it already available, has yet to be fully realised.

Key words. Antibiotics; genomics; essential genes; proteomics; transcriptional analysis; DNA microarrays.

Introduction

The ever-increasing frequency and spectrum of infections resistant to currently available antibacterials raise the question as to the feasibility of keeping pace with the rate at which drugs become ineffective. With the appearance and spread of multi-resistant and, in fact, untreatable tuberculosis [1], we are effectively returning to the pre-antibiotic era. Risk assessment for a number of infectious diseases such as cholera, diphtheria and respiratory tract infections, including tuberculosis, will have to be reconsidered and the strategies to combat the causative pathogens re-evaluated [2].
Factors contributing to the emergence of resistance and its rapid spread are multifaceted and differ depending upon geographic location. The opening of borders, brought about by the changing political scene, together with the ever-increasing ease of travel are major contributors to this rapid spread.
The continuous exposure of micro-organisms to antibiotics used as 'growth promoters' selects for general resistance mechanisms against classes of antibiotics [3]. Paradoxically, improvements in medical technologies have also led to an increase in resistance, as there has been a substantial increase in the number of immune-compromised patients, including AIDS patients, who rely on the effectiveness of antimicrobials [4]. This pool of patients acts as a reservoir, selecting for resistant organisms. Chronic or persistent infections involving intracellular organisms or resulting from situations that encourage 'slow growth' such as bio-films on catheters are also sources for resistance development [5, 6].
The lack of innovation in antibacterial treatment is clearly reflected by the diminishing interval in time between when a new drug enters the market and the appearance of resistance to it in one or more different strains of bacteria. An intrinsic problem here is that the new antibiotics are all, in fact, modifications of an already relatively limited number of existing classes of compounds. This phenomenon and the lack of new antibiotics with novel mechanisms of action can partly be explained by the reduced efforts in antibacterial research in the 1980s. It was believed that the existing repertoire of antibiotics would result in the extinction of a number of infectious diseases and thus bacterial infections were no longer considered a problem. Research efforts and financing by both the major pharmaceutical

* Corresponding author.

companies and public health organisations were shifted to other more pressing medical problems.

Fortuitously, the revival of interest in this problem has coincided with an explosion of technological advances originally initiated by the programmes to sequence genomic DNA. Miniaturisation, automation and bioinformatics are changing our concepts regarding the way we can approach antimicrobial research. Instead of studying the expression of individual genes, operons or even stimulons, we are now able to follow the entire expression pattern of an organism. The major stimulus for this change of concept in the prokaryotic field was the publication of the fully assembled genome sequence of *Haemophilus influenzae* [7].

The starting point: the genomic sequence

To date, more than 17 fully assembled bacterial genomic sequences, including strains of at least 10 pathogenic species, have been published. Analyses of these sequences together with a further 50 that are in varying stages of completion (for a comprehensive list see: http://www.tigr.org/tdb/mdb/mdb.html) serve to demonstrate our lack of understanding of the physiology of bacteria. Even in the smallest genome sequenced, *Mycoplasma genitalium*, approximately 35% of the encoded proteins cannot be annotated with a function [8]. Many of the annotations assigned to the others are also of little relevance, e.g. 'ATPase', reflecting the specific interests of individual researchers rather than addressing the true physiological role of the protein in the cell. Studying the functions and interactions of all the gene products encoded by a cell is not only an interesting academic exercise but is very important for identifying and validating new targets. The genomic sequence is, therefore, also the starting point for a new strategy for drug discovery (fig. 1).

For our research we have chosen to work with two model organisms, *H. influenzae* Rd KW20 and *Streptococcus pneumoniae* R6, as representatives of a Gram − and Gram + pathogenic species, respectively. Our reasons for selecting these two organisms, besides their clinical relevance as pathogens, were the availability of their sequences, their relatively small genomes and their amenability to genetic manipulation. As our focus is the development of broad-spectrum antibiotics, targets must be conserved between species. This allows us to use the more appropriate of the two organisms for any particular technology, while taking the second organism for confirmation of our results. For example, insertional gene disruption was relatively more efficient using *S. pneumoniae*, whereas *H. influenzae* was more amenable to proteomics due to the early accessibility to sequence information and the existence of a minimal medium for metabolic labelling.

In silico analysis

Comparative genome analysis can be an extremely useful tool for both assessing the degree of conservation and for providing a starting point for addressing the function of a potentially encoded protein. Many of the unassigned proteins can be grouped by phylogenetic approaches into protein families and superfamilies, which also provides an indication as to their functions [9]. A problem in this analysis is that proteins within a family have often evolved such that their functions are now very different. Clearly the results of this kind of comparison are only suggestive and must be verified experimentally. However, by using such approaches and also by extending the capabilities of the available software, the superposition of the *H. influenzae* enzyme repertoire upon the known *Escherichia coli* metabolic pathways was recently demonstrated [10]. The information derived from this analysis allowed the reconstruction of similar and alternative pathways in *H.*

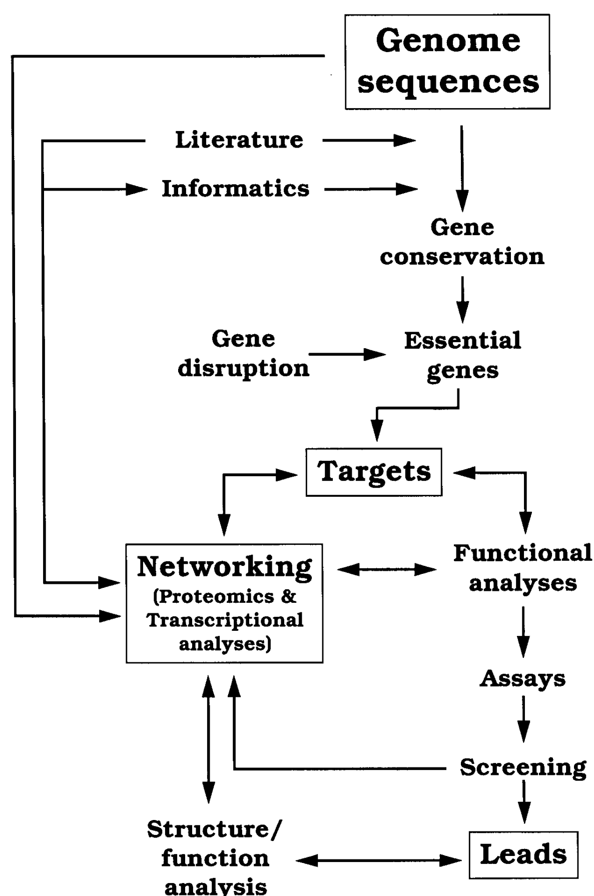

Figure 1. Schema showing the integration of 'functional genomics' into drug discovery.

influenzae, which provide a plausible explanation for its nutritional requirements.
Within drug discovery, this sort of detailed analysis is, at present, only of interest in the case of genes conserved in many species. It should be noted that the term 'conserved' is relative as it is based on a predetermined value for the degree of homology and usually does not take phylogenetic distance into account. 'Conservation' also refers here to sequence, and sometimes structure, but there is also conservation of function, which is practically impossible to decipher from the sequence itself. Superposing the enzyme repertoire deduced by homology onto the known metabolic pathways reveals gaps for which the bacteria must possess a function. This would indicate that proteins of completely different families with no homology in their sequences or their general structure exist that are able to carry out the particular activity. An example for this is the dehydrogenase/reductase activity of an enzyme in the protozoan parasite *Leishmania* which, although belonging to a very different protein superfamily, can take over the activity of dihydrofolate reductase, thereby conferring resistance to methotrexate [11].
Comparative studies have proven useful in our own research in an analysis of the database for *S. pneumoniae*. For example, as mentioned below, 13 potential members of the bacterial phosphotransfer signalling mechanisms could be identified.

Identifying essential genes

The search for essential genes is an obvious starting point for identifying new targets, which, in turn, will result in the discovery of novel antimicrobial compounds. From the experimental point of view, eukaryotic models have the advantage that disruptions can be created in the diploid cell and homozygotes selected after a round of meiosis. Disruption of genes required for the viability of the cell or organism will result in only half the progeny compared to those within non-essential genes. In a microbe, such as *Saccharomyces cerevisiae*, this methodology can be developed into a screen allowing the identification of essential genes in one single experiment [12]. Such a system is inapplicable to bacteria in which only the absences of colonies in a selectable transformation indicate the requirement for the gene. However, coupled with the knowledge of fully assembled genome sequences, the recently introduced technologies using DNA arrays have allowed screens to be run whereby all disrupted non-essential genes can be identified, thereby identifying the essential genes by default [13]. Clearly, in such experiments, care must be taken to obtain total coverage of the genome. Also, each potential candidate essential gene must be confirmed in a second round of experiments to ensure exclusion of artefacts such as polar effects.
In our department, a screen for essential genes was run until well over 100 potentially essential genes had been identified. Confirmation of the essentiality, using constructs which integrate into the genome so that downstream genes in operons are under the control of tetracycline-regulated promoters [14], is still in progress.

Finding virulence factors

Each stage of an infection will require different patterns of gene expression depending upon the conditions encountered by the bacteria. The genes, which are essential for these specific conditions, are termed virulence genes. The corresponding gene products, termed virulence factors, are not necessarily essential in vitro and can often only be assessed in an animal model or in the appropriately stressed cell. Two methodologies that have proven useful in the large-scale identification of new virulence genes in different bacterial species are in vivo expression technology, (IVET) [15] and signature-tagged mutagenesis, (STM) [16]. IVET was originally developed for use with *Salmonella typhimurium* and has since been applied to *Vibrio cholerae* [17] and *Pseudomonas aeruginosa* [18]. STM uses the concept originally developed by Walsh and Cepko [19] to monitor the fate of clonally related neocortical cells during development. Transposons are tagged, each with a unique DNA sequence, allowing the detection of the individual transposon insertions. Analysis of a pool of clones both before and after passage within an animal identifies those insertions that have proved detrimental for the bacteria. This methodology was also originally developed for use in *S. typhimurium* [16] but has been extended to *Staphylococcus aureus* [20], *V. cholerae* [21] and *S. pneumonia* [22].
As well as the more classical virulence genes, such as those involved in adhesion or capsule biosynthesis, many of the genes identified using these methods are involved in metabolism or in the transport of metabolites, indicating a requirement for these specific metabolites at some stage during the infection. The differences between mammalian and bacterial metabolism can be exploited as a source of targets for antibiotics, as in the case of the sulphonamides (inhibitors of the folate pathway), perhaps the classical example of this principle.
Most classic virulence factors are specific for one species of bacteria and are therefore not of use as targets for broad-spectrum antibiotics. A search through all sequenced genomes with sequences annotated with the term 'virulence' does, however, indicate that the phosphotransfer signalling mechanisms [23] could be appropriate targets (fig. 2). One member of the

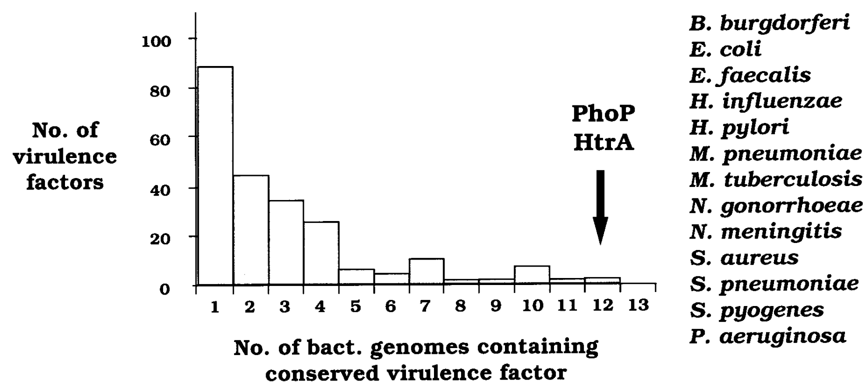

Figure 2. A serial BLAST of all genes annotated with the term 'virulence' found in the databases for the pathogens listed in the figure was run using a *p* value of 10^{-10}. Only the one phosphotransfer signalling system and one serine protease were conserved in twelve of the organisms. None of the genes was conserved in all thirteen.

phosphotransfer signalling mechanisms, which is regulated by phosphate concentration, is represented in all the species sequenced to date except for *Mycoplasma*. This complex also has the advantage that it is possible to develop both cellular and in vitro enzyme assays which can then be used in screening.

From stimulus to physiological function

Networking

The above mentioned topics do not rely on but are clearly facilitated by access to the genome sequence. Two technologies, which would be totally impractical without cognisance of the complete genome sequence, are proteomics and global transcriptional analysis. Together these technologies form the basis of an approach toward the understanding of cellular networking, a step towards possibly achieving a goal long sought after by biologists, namely, of showing how a cell functions as an entirety.

Proteomics

The term proteomics is defined as the study of the protein complement expressed by a genome or, perhaps more realistically, the protein complement expressed by a genome at a given time. In practice, this study consists of two parts: (1) separating and identifying the expressed proteins and (2) analysis of the changes in their expression levels, which reflect the response of the cell to its environment. Apart from the availability of genome sequences, the usefulness of this technology has grown with the improvements in gels reproducibility and increase in the sensitivity of mass spectrometry, which now makes it possible to identify the polypeptides separated in the gels. Visualisation of the separated polypeptides is now the limiting factor.

Exposure of cells to sub-lethal levels of antibiotics elicits survival responses that are characteristic for any particular antibiotic. This property is being utilised in an attempt to characterise the mode of action of antimicrobials identified in various whole-cell assays and to validate the target for compounds synthesised in the chemical programmes. Proteomics is used as an aid in understanding which response is being induced. Using the appropriate software, it is possible to compare the patterns of unknown responses with a database of patterns produced using antibiotics for which the mode of action has been established (fig. 3). (In this context the term 'pattern' refers not only to the presence or absence of a spot but also to the relative intensity of the spots.) The 'established' responses in the database, although showing a common pattern of changes for different antibiotics acting upon the same target, also contain superimposed secondary responses which cannot be explained with reference to the 'known' target. These secondary responses most probably represent the mechanistic differences in the action of the antibiotics but perhaps also indicate secondary targets. We have recently encountered an example of the latter situation in a compound selected in vitro for its activity against DNA gyrase. A secondary in vivo response could be discerned from a comparison of the patterns that indicated inhibition of protein synthesis. This activity could be confirmed in an in vitro transcription/translation assay.

This comparison is a useful first step in determining the response, as it does not rely on the identification of the spots and therefore uses all the information contained

within the gel image. However, as the goal is to discover and validate new targets, it is clear that there will be responses for which there is no comparable pattern in the database. Although the number of spots that have been identified is limited, it is still possible to visualize, albeit with gaps, the changes in the enzyme levels within the major metabolic pathways. The information gained from such analyses coupled to what can be learnt from the literature about the regulation of these pathways is proving a useful tool to indicate potential candidate enzymes as targets of antimicrobial compounds. In these cases, a detailed analysis of the expression patterns with respect to the identified spots is required. In a recent study, we examined the response of *H. influenzae* to the inhibitors of dihydrofolate reductase (trimethoprim) and dihydropteroate synthase (sulphamethoxazole) [24]. We chose these particular antibiotics as a proof of concept as the downstream effects of tetrahydrofolate starvation, for example: inhibition of dTMP and L-methionine biosynthesis, have been examined in detail for *E. coli* [25, 26]. Most of the enzymes involved in L-methionine biosynthesis, are sufficiently abundant that they can be identified with confidence in the gel pattern. The changes in the intensity of the representative spots could be measured and are shown for trimethoprim in figure 4 where the degree of change is colour coded. Such a graphical representation could be extended to the enzymes involved in gluconeogenesis and pyruvate biosynthesis. It was not, however, possible to identify, with any degree of confidence, enough of the spots representing the enzymes involved in pyrimidine biosynthesis to construct any meaningful figure. Proteins that were found to be diagnostic for the two responses, most notably several of the stress proteins (fig. 5), as previously noted for *E. coli* [27], cannot be assigned to any particular pathway. This study demonstrated that responses showing similarity in one or more pathways to an already determined antibiotic target pattern, but deviating in others, can indicate a second target related to that in the database. An explanation for all the observed changes is at present not possible. These studies are still in their infancy and much more experience is required to enable the differentiation of the secondary from the primary effects of inhibition. A further consideration when examining the response of the cell to an antibiotic is the specificity of the antibiotic. For example, quinolones, inhibitors of DNA gyrase, also inhibit topoisomerase IV [28] and at higher concentrations, translation [29].

The scarcity of information concerning the pyrimidine pathway is indicative of the limitations in the sensitivity of the methodology. On a preparative scale, it is possible to assign spots representing almost half of the polypeptide products encoded within the genome. However, the difference in the migration of the polypeptides in preparative gels when compared with analytical gels

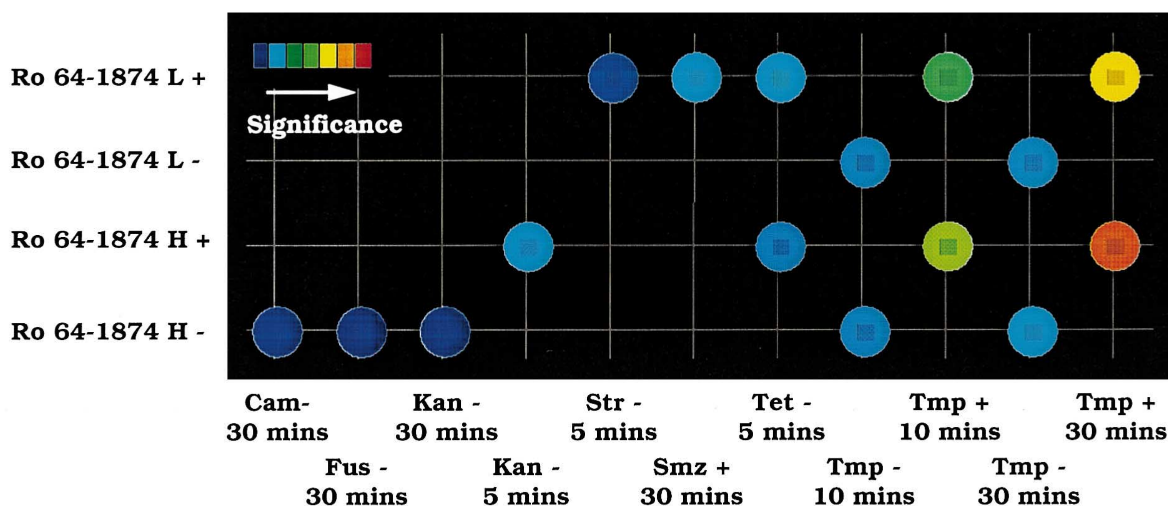

Figure 3. 'MineSet' software was used to make a pairwise comparison between the responses of *H. influenzae* to 8 mg/l (L) and 150 mg/l (H) of a 2,4-diaminopyrimidine, Ro 64-1874, and those to other antibiotics [inhibitors of: translation (seven compounds), transcription (two), gyrase (three), folate pathway (two), cell wall synthesis (two), tRNA synthetases (one) and fatty acid biosynthesis (one)]. The significance value is calculated as the ratio of the percentage of spots induced (repressed) by the second antibiotic that are also induced (repressed) by Ro 64-1874 to the total percentage of spots induced (repressed) by Ro 64-1874. + and − designate sets of spots synthesised at an induced or repressed rate (induction ratio > 1.25 and < 0.8), respectively. Cam, chloramphenicol; Kan, kanamycin; Str, streptomycin; Tet, tetracycline; Trip, trimethoprim; Fus, fusidic acid; Smz, sulphamethoxazole.

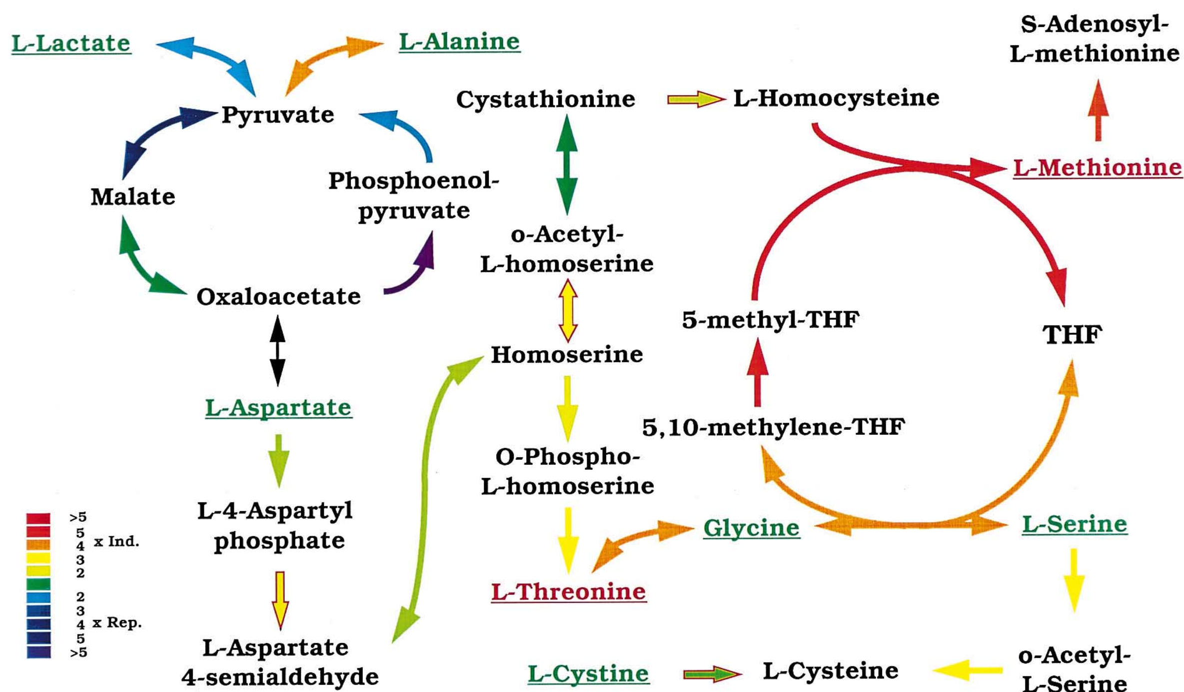

Figure 4. Cultures of *H. influenzae* were induced with trimethoprim (Tmp) for 10 min before labelling for 3 min with ^{35}S. Induction ratios, compared to non-induced parallel cultures, were calculated for the intensities of the spots representing the respective enzymes after 2D-PAGE analysis. Values for arrows outlined in red are taken from measurements of RNA levels whereas the value for the black arrow was not determined. Supplements present in excess in the medium are written in green whereas red indicates limiting concentrations. The figure demonstrates that, as would be expected, the enzymes involved in the biosynthesis of L-methionine are regulated in a similar fashion to those of *E. coli*. It also demonstrates that the response to Tmp is more complex than previously thought (cf. the enzymes involved in pyruvate biosynthesis).

excludes, in many cases, a direct transfer of this information. Pooling the excised spots from many analytical gels is at present the only reliable method of verifying the assignation of a spot. Undoubtedly, the sensitivity of mass spectrometry will be further improved so that proteomics will become a routine and reliable tool for studying the protein ratios within cells.

Transcriptional analysis

The second basic tool for networking, transcriptional analysis, addresses the same questions as tackled in proteomics. As the transcriptional response is more rapid than the translational response, this technology is more suitable for determining primary responses. It is important to realise that the level of a transcript determined for any particular gene is a measure of the total level of that transcript in the culture at the time of harvesting. The metabolic labelling of proteins on the other hand measures the protein synthesised in the specified period of labelling. From our own experiments in proteomics we know that continuous labelling of the cells produces a less sensitive, with respect to the response, and often different result compared to pulse labelling. This can be explained in part by the accumulation of proteins each with a different half-life, but is most probably also the result of a heterogeneous response within the cell population. Such a heterogeneous response was recently demonstrated for *S. typhimurium*/ F\lac+ cells upon sub-optimal induction of the lac operon [30].

DNA microarrays

Confirming the presence of an RNA transcript within a mixture using the complementary DNA sequence fixed to a solid support is a methodology developed at the end of the 1970s. The recent advances made in this methodology involve the technology of fixing hundreds and up to thousands of DNA fragments within a small area and identifying the individual signals. The first experiments in this direction were in the expression analysis of *E. coli* using overlapping lambda clones and ^{32}P-labelled cDNA [31]. As more sequences became

available there was a need to have many smaller and more specific DNA fragments for such an analysis. In the case of bacterial genomes, it is still possible to consider using a membrane as a support as the number of fragments to be spotted, such that each gene is represented at least once, is relatively small. The drive in this technology, as in most of genomics, is towards the analysis of the human genome where, compared to bacteria, a further two orders of magnitude in the number of genes and therefore spots are to be expected. The first high-density arrays of fragments were again lambda clones, this time from *S. cerevisiae*, which were spotted on a glass substrate and probed with fluorescent-labelled cDNA [32]. In these initial experiments, over 1700 spots were placed within 1.8 × 1.8 cm. The number of spots on such an array has now been increased to 10,000.

The GeneChip, manufactured by Affymetrix, is a high-density array of DNA oligomers, the sequences of which are defined with respect to the genome sequence of the relevant organism. These oligomers are synthesised directly onto the chip in a defined area termed a 'feature' using technology developed in the computer industry. Within an agreement between Roche and Affymetrix, our department is optimising methods for detecting prokaryotic mRNA. The alternatives to the Affymetrix chips, i.e. fragment arrays as discussed above, were considered but rejected. As it is not feasible to enrich for mRNA in an unbiased manner, any labelling procedure will result in > 95% of the label being incorporated into rRNA. This, depending upon the stringency used and the choice of fragment, will cross-hybridise to varying degrees. The design of the Affymetrix chip can, however, be directed to exclude any oligomers that are not unique to any given gene thus reducing the problem of cross-hybridisation. Our first attempts using a test chip embodying 65,000 features led to the successful detection of a selected set of transcripts using direct labelling of the RNA [33]. A yet higher-density chip (200,000 features), encompassing oligomers representing the total genomes of both *H. influenzae* Rd KW20 and *S. pneumoniae* 1313, has proven problematic because it has been difficult to obtain a sensitivity that is high enough for reproducible analyses. The problems have been successfully resolved and we are now able to gather information which, coupled with the proteome analysis, also allows us to compare protein and RNA ratios.

Combining proteomics with transcriptional analysis

Although it would appear that transcriptional analysis is less complicated and provides information on more genes than the analysis of the proteome, there are good reasons for determining the levels of both the mRNA and protein complements of the cell. The number of transcripts for a particular gene only indicates one level of regulation of the final functional product, namely the protein. This is clearly demonstrated by the observation that one protein can be represented by several spots, each of which varies in its relative intensity, in the 2D gel. This variation in the number of spots is the result of

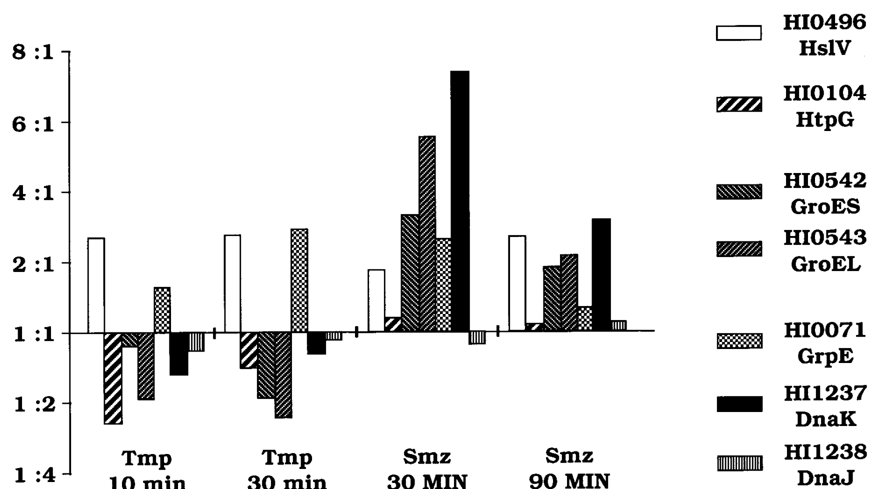

Figure 5. Induction ratios for the intensities of the spots identified as representing stress proteins in *H. influenzae* after induction (see fig. 4) with trimethoprim and sulphamethoxazole. Although these gene products cannot be assigned to a specific pathway, they are diagnostic for the response to trimethoprim. It is of particular interest to note that the regulation of the synthesis of GrpE, DnaK and DnaJ, which are considered to form a single complex, is divergent in this response.

post-translational processing, which, whether it be proteolytic cleavage or the modification of specific amino acid residues, regulates the activity of the protein. What the modifications are and which of the spots observed in the gel represents the active protein is at present only known in a few cases. Further regulation of the quantity of a protein is achieved not only at the level of translation (via, for example, codon usage) but also by the rate of degradation. Therefore, in summary, the response of the cell to any stimulus, although resulting in the activation of several regulons, which would be detectable in transcriptional analysis, may only manifest itself in the induction/repression of a specific set of proteins.

The future

The initial phase in post-genomics could be rapidly implemented into antimicrobial research combining informatics with well-established methods. This phase has already led to the discovery of unexploited targets. The second phase will take longer as it involves not only development and adaptation of technologies that will be able to cope with large numbers of genes or gene products varying in quantity by two to three orders of magnitude but also a better understanding of cell physiology. At present we do not know enough about the biochemistry of the cell to understand all that we observe. Further experimentation with the above-mentioned methodologies, examining already well studied physiological conditions, is vital. Together with the determination of other parameters such as metabolite levels and metabolic flux, we will, with time, extend our knowledge to a point where we will attain at least a basic understanding of how a bacterial cell responds to changes in its environment. This knowledge will be crucial for both understanding the activity of, and developing, new classes of antibacterial compounds.

1 Rattan A., Kalia A. and Ahmad N. (1998) Multidrug-resistant *Mycobacterium tuberculosis*: molecular perspectives. Emerg. Infect. Dis. **4:** 195–209
2 Domin M. A. (1998) Highly virulent pathogens – a post antibiotic era? Br. J. Theatre Nurs. **8:** 14–18
3 Aarestrup F. M., Bager F., Jensen N. E., Madsen M., Meyling A. and Wegener H. C. (1998) Surveillance of antimicrobial resistance in bacteria isolated from food animals to antimicrobial growth promoters and related therapeutic agents in Denmark. Apmis **106:** 606–622
4 Kilbourne E. D. (1996) The emergence of "emerging diseases": a lesson in holistic epidemiology. Mt. Sinai J. Med. **63:** 159–166
5 Millar M. R. and Pike J. (1992) Bactericidal activity of antimicrobial agents against slowly growing *Helicobacter pylori*. Antimicrob. Agents Chemother. **36:** 185–187
6 Chuard C., Vaudaux P. E., Proctor R. A. and Lew D. P. (1997) Decreased susceptibility to antibiotic killing of a stable small colony variant of *Staphylococcus aureus* in fluid phase and on fibronectin-coated surfaces. J. Antimicrob. Chemother. **39:** 603–608
7 Fleischmann R. D., Adams M. D., White O., Clayton R. A., Kirkness E. F., Kerlavage A. R. et al. (1995) Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science **269:** 496–512
8 Fraser C. M., Gocayne J. D., White O., Adams M. D., Clayton R. A., Fleischmann R. D. et al. (1995) The minimal gene complement of *Mycoplasma genitalium*. Science **270:** 397–403
9 Saier M. H. (1996) Phylogenetic approaches to the identification and characterization of protein families and superfamilies. Microb. Comp. Genomics **1:** 129–150
10 Tatusov R. L., Mushegian A. R., Bork P., Brown N. P., Hayes W. S., Borodovsky M. et al. (1996) Metabolism and evolution of *Haemophilus influenzae* deduced from a whole-genome comparison with *Escherichia coli*. Curr. Biol. **6:** 279–291
11 Wang J., Leblanc E., Chang C. F., Papadopoulou B., Bray T., Whiteley J. M. et al. (1997) Pterin and folate reduction by the *Leishmania tarentolae* H locus short-chain dehydrogenase/reductase PTR1. Arch. Biochem. Biophys. **342:** 197–202
12 Goebl M. G. and Petes T. D. (1986) Most of the yeast genomic sequences are not essential for cell growth and division. Cell **46:** 983–992
13 Smith V., Chou K. N., Lashkari D., Botstein D. and Brown P. O. (1996) Functional analysis of the genes of yeast chromosome V by genetic footprinting. Science **274:** 2069–2074
14 Stieger M., Wohlgensinger B., Kamber M., Rolf L. and Keck W. (1999) Integrational plasmids for the tetracycline-regulated expression of genes in *Streptococcus pneumoniae*. Gene **226:** 243–251
15 Slauch J. M., Mahan M. J. and Mekalanos J. J. (1994) In vivo expression technology for selection of bacterial genes specifically induced in host tissues. Methods Enzymol. **235:** 481–492
16 Hensel M., Shea J. E., Gleeson C., Jones M. D., Dalton E. and Holden D. W. (1995) Simultaneous identification of bacterial virulence genes by negative selection. Science **269:** 400–403
17 Camilli A. and Mekalanos J. J. (1995) Use of recombinase gene fusions to identify *Vibrio cholerae* genes induced during infection. Mol. Microbiol. **18:** 671–683
18 Wang J., Mushegian A., Lory S. and Jin S. (1996) Large-scale isolation of candidate virulence genes of *Pseudomonas aeruginosa* by in vivo selection. Proc. Natl. Acad. Sci. USA **93:** 10434–10439
19 Walsh C. and Cepko C. L. (1992) Widespread dispersion of neuronal clones across functional regions of the cerebral cortex. Science **255:** 434–440
20 Mei J. M., Nourbakhsh F., Ford C. W. and Holden D. W. (1997) Identification of *Staphylococcus aureus* virulence genes in a murine model of bacteraemia using signature-tagged mutagenesis. Mol. Microbiol. **26:** 399–407
21 Chiang S. L. and Mekalanos J. J. (1998) Use of signature-tagged transposon mutagenesis to identify *Vibrio cholerae* genes critical for colonization. Mol. Microbiol. **27:** 797–805
22 Polissi A., Pontiggia A., Feger G., Altieri M., Mottl H., Ferrari L. et al. (1998) Large-scale identification of virulence genes from *Streptococcus pneumoniae*. Infect. Immun. **66:** 5620–5629
23 Mizuno T. (1998) His-Asp phosphotransfer signal transduction. J. Biochem. (Tokyo) **123:** 555–563
24 Evers S., Di Padova K., Meyer M., Fountoulakis M., Keck W. and Gray C. P. (1998) Strategies towards a better understanding of antibiotic action: folate pathway inhibition in *Haemophilus influenzae* as an example. Electrophoresis **19:** 1980–1988
25 Ahmad S. I., Kirk S. H. and Eisenstark A. (1998) Thymine metabolism and thymineless death in prokaryotes and eukaryotes. Annu. Rev. Microbiol. **52:** 591–625
26 Weissbach H. and Brot N. (1991) Regulation of methionine synthesis in *Escherichia coli*. Mol. Microbiol. **5:** 1593–1597

27 Gage D. J. and Neidhardt F. C. (1993) Modulation of the heat shock response by one-carbon metabolism in *Escherichia coli*. J. Bacteriol. **175:** 1961–1970
28 Tanaka M., Onodera Y., Uchida Y., Sato K. and Hayakawa I. (1997) Inhibitory activities of quinolones against DNA gyrase and topoisomerase IV purified from *Staphylococcus aureus*. Antimicrob. Agents Chemother. **41:** 2362–2366
29 Hooper D. C. and Wolfson J. S. (1993) Mechanisms of quinolone action and bacterial killing. In: Quinolone Antimicrobial Agents, pp. 53–75, Hooper D. C. and Wolfson J. S. (eds), American Society for Microbiology, Washington, DC
30 Tolker-Nielsen T., Holmstrom K., Boe L. and Molin S. (1998) Non-genetic population heterogeneity studied by in situ polymerase chain reaction. Mol. Microbiol. **27:** 1099–1105
31 Chuang S. E., Daniels D. L. and Blattner F. R. (1993) Global regulation of gene expression in *Escherichia coli*. J. Bacteriol. **175:** 2026–2036
32 Shalon D., Smith S. J. and Brown P. O. (1996) A DNA microarray system for analyzing complex DNA samples using two-color fluorescent probe hybridization. Genome Res. **6:** 639–645
33 De Saizieu A., Certa U., Warrington J., Gray C., Keck W. and Mous J. (1998) Bacterial transcript imaging by hybridization of total RNA to oligonucleotide arrays. Nat. Biotechnol. **16:** 45–48