SIXTH REVIEW CONFERENCE OF THE STATES PARTIES TO THE CONVENTION ON THE PROHIBITION OF THE DEVELOPMENT, PRODUCTION AND STOCKPILING OF BACTERIOLOGICAL (BIOLOGICAL) AND TOXIN WEAPONS AND ON THEIR DESTRUCTION

BWC/CONF.VI/INF.4
28 September 2006

Original: ENGLISH

Geneva, 20 November – 8 December 2006
Item 10 of the provisional agenda
Review of the operation of the Convention as provided for in its Article XII

BACKGROUND INFORMATION DOCUMENT ON NEW SCIENTIFIC AND TECHNOLOGICAL DEVELOPMENTS RELEVANT TO THE CONVENTION

Prepared by the Secretariat

Introduction

1. In paragraph 22 of its report (BWC/CONF.VI/PC/2), the Preparatory Committee for the Sixth Review Conference decided to request the Secretariat to compile a background information document on new scientific and technological developments relevant to the Convention, to be compiled from information submitted by States Parties as well as from information provided by relevant international organisations. The Secretariat has prepared this document in accordance with that request.

2. The Second, Third and Fourth Review Conferences[1], *conscious of apprehensions arising from relevant scientific and technological developments,* inter alia, *in the fields of microbiology, genetic engineering and biotechnology, and the possibilities of their use for purposes inconsistent with the objectives and the provisions of the Convention,* reaffirmed that *the undertaking given by the States Parties in Article I applies to all such developments.* The Fourth Review Conference supplemented the list of scientific and technological developments with *molecular biology… and any applications resulting from genome studies*.

3. This document discusses significant developments in these fields since the Fifth Review Conference, as well as the evolution of new disciplines. It covers: biotechnology; genomics; proteomics; bioinformatics and computational biology; systems biology; drug discovery, design and delivery; synthetic biology and biological engineering; as well as a number of other relevant developments. Annexed to this document are an overview on identifying experiments of concern (Annex I), and a list of actual experiments often quoted as being particularly relevant to the Convention (Annex II).

4. The following States Parties submitted information to the Secretariat for the preparation of this document: Australia, Czech Republic, the Netherlands, Portugal, Sweden, the United

[1] BWC/CONF.II/13, BWC/CONF.III/23, and BWC/CONF.IV/9.

Kingdom, and the United States of America. The full texts of the submissions from these States Parties, as well as any subsequent submissions received too late to include in this document, are available online at http://www.unog.ch/bwc in the Sixth Review Conference section. Information in this document has also been drawn from a variety of documents made available by intergovernmental, international and professional scientific organizations.

5. An inclusive approach has been taken in determining which developments may be of relevance to the Convention. Although the advances discussed in this document have obvious applications for prophylactic, protective or other peaceful purposes, they may also have the potential to be applied in contravention of the objectives and provisions of the Convention. Inclusion of a development in this document does not imply any assessment by the Secretariat of its permissibility or otherwise under the Convention.

Biotechnology

6. Biotechnology has yielded public health, agricultural, and economic benefits as well as feeding back into the life sciences to improve development. Numerous commercial applications have been found and the last few years have seen dramatic increases in reliance on this technology. The benefits derived from biotechnology can be found increasingly in developing countries. A paper presented at the United Nations Industrial Development Organization's Global Partners Symposium in Austria in March 2005, records that both the number of papers published and patents awarded, to various developing States for developments in health biotechnology, rose dramatically between 1991 and 2002.[2]

7. The high turnover of staff at small biotechnology companies, especially those attached to academic institutions, has increased the rate at which developments spread between institutions. These companies are often dependent on a small number of potential products. If their efforts are successful, they (and their intellectual property) tend to be bought up by larger companies. If they fail, they tend to be liquidated. This dynamic employment environment is conducive for the spreading of knowledge (intangible biotechnology).

8. There are clear indications that biotechnology, both tangible and intangible, is indeed spreading. Disposable bioreactors are one example of a technology which enables this dissemination. These are self-contained production devices delivered with all the necessary fixtures and fittings as ready for use. The come in sizes from 1 to 500 litres and when used can be disposed of, eliminating the need for cleaning, sterilisation and validation.

Bioprospecting

9. *Bioprospecting* is the search for previously unrecognised, naturally occurring, biological diversity that might serve as a source of material for medicine, agriculture and industry. Material gathered in bioprospecting may include genetic sequences, proteins, complex biological compounds or entire organisms. Bioprospecting has also led to the identification of many new species, especially different microbes. In recent years, bioprospecting has produced candidates for new antibiotics, antiviral compounds, anticancer agents, antioxidants, anti-diabetic agents, immunosuppressive compounds, insecticides, as well as genetic sequences encoding high and

[2] *Health Biotechnology Innovation in Developing Countries*, UNIDO, Global Partners Symposium, Vienna, Austria, 3-4 March 2005. For more information see: http://www.unido.org/file-storage/download/?file_id=35240.

low temperature stability, high or low pH tolerance and high or low salt tolerance. Surveys for microbial agents which might act as pathogens in the future have also been carried out. Vectors and natural reservoirs have been examined for previously unknown pathogens or microbes related to known pathogens. This improves risk management programmes, provides a degree of early warning for future disease outbreaks and enhances understandings of microbial diversity and inferred microbial function.

10. A number of advances have underpinned the development of bioprospecting. Without developments in microbial cultivation, seriological surveys, extraction and purification techniques, abilities to amplify genetic material, *genomics* techniques, and *drug discovery, design and delivery*, current bioprospecting efforts would not be possible.

High-Throughput Screening

11. Since the Fifth Review Conference there has been considerable progress in automating and miniaturising repetitive laboratory tasks for examining biologically active compounds for properties of interest. Application of these advances led to significant improvements in the efficiency and speed of processing samples. *High-throughput screening* technologies allow large numbers of compounds (libraries – see *combinatorial biochemistry*) to be screened for specific activities, such as looking for a compound to bind to a particular receptor, or inactivate a specific enzyme. They also allow a single compound to be tested for numerous possible activities. Such an ability was developed for DNA chips (oligonucleotide microarrays) which are commonly used in basic and applied research, to monitor the levels of expression of genes, in identifying the functions of genes, assessing genetic variation, and elucidating new targets for therapeutic drugs. They therefore enable numerous other fields of scientific and technological endeavour.

12. Advances in high-throughput technologies were produced by an amalgamation of developments in a wide number of fields, including miniaturised screening formats, liquid handling, signal detection, robotics, *bioinformatics* and conducting biological assays. These advances now allow a user to screen more than 100,000 compounds per day for a desired activity. As a result research projects can now regularly screen in excess of 1 million compounds, a task, if conducted by hand, which would take millions of laboratory-hours.

Biological Microprocessing

13. Advances in miniaturisation and automation have also been applied to produce discrete functional devices capable of conducting entire biological tests, previously requiring a laboratory. They have been described as being a "lab-on-a-chip". Advances in microfluidics and microfabrication have allowed the production of devices ranging in size from a microscope slide to a compact disc. It is no longer necessary to work on a macro scale, such as a laboratory bench with flasks of material. The same processes can be replicated with picolitres (10^{-12} of a litre) of initial sample and equally small volumes of reagents. These devices are generally produced to carry out one specific activity, such as DNA analysis, an immunoassay, cell analysis, or for measuring enzyme-activity. They can be fully automated and can conduct numerous steps to complete their prescribed activity, such as breaking down the sample, diluting it, adding reagents, mixing it and detecting reactions.

14. Advanced forms of the "lab-on-a-chip" can be fully integrated and carry out all the stages from sample introduction, through to interpreting the results. More commonly, they can be connected to standard laboratory equipment, such as external detectors. Reducing the amount of specialist equipment and training required to undertake these tests allows devices to be developed for use in new environments. Such devices have great potential for epidemiology and, therefore, for detecting, diagnosing, characterising and responding to disease outbreaks (see *detection technology*).

Genomics

15. If a gene is a working sub-unit of DNA which encodes a specific product, such as a protein, then the genome is the entire collection of genes within an organism. Advances in technology have dramatically increased the speed at which it is possible to identify, characterise and manipulate genes. Advances in sequencing genes underlie many of the scientific and technological developments relevant to the Convention. Due to the large amounts of genetic information being generated, it is now possible to talk in terms of genomes as opposed to individual genes. Progress in *genomics* (the study of genomes) has been assisted by concerted efforts to ensure genomic information remains open source.

DNA Sequencing

16. Sequencing is the identification of the order of the nucleotides which make up genetic information. In other words, it involves turning the physical material into abstract information. It is not a new technology. Experienced scientists have been conducting sequencing for some time. Prior to the Fifth Review Conference there had been significant advances in automation and efficiency. This permitted entire genomes to be sequenced: the first eukaryotic genome (a yeast) in 1997; the first animal genome in 1998; and the human genome in 2001. Advances made during the sequencing of the human genome illustrate the impact of this trend toward increasing automation. The Human Genome Project (the international collaborative effort to elucidate the sequence) aimed to accomplish the task in 15 years using a relatively large number of highly specialised facilities staffed by experienced scientists. As technology improved, the limiting factor became the associated staffing costs. Automation and miniaturisation have allowed post-doctoral researchers to be replaced by master's students; master's students to be replaced by undergraduates; and technicians, with only minimal higher learning, to replace the undergraduates. Adoption of sequencing advances allowed a commercial competitor to publish a draft of the human genome at the same time as the international coalition, despite have started on the project almost ten years later.

17. DNA sequencing technology has continued to evolve since the last Review Conference increasing the ability to identify and characterise previously unknown organisms. Efficiency and levels of automation have continued to increase over the last four years. The associated costs are falling by a factor of two every 12-18 months. Current developments, such as capillary sequencing machines and DNA Chips, are enabling studies of sequence variation within species by sequencing large numbers of strains, including pathogens, in parallel. Existing work also focuses on creating machines capable of reading a single copy of a sequence. This will significantly reduce the numbers of errors incorporated into a sequence (compared to the current approach of amplifying multiple copies), allowing the genome of an individual cell to be read more accurately and enable advances in functional genomics and *proteomics*.

DNA Synthesis

18. *DNA synthesis* is the reverse process to DNA sequencing. It involves turning the sequence data back into physical material. The ability to generate physical DNA to match a sequence of information is also not new. There have, however been significant improvements in the efficiency and automation of the process. In the 1970s it was possible to generate DNA sequences by hand. During the 1980s advanced methods were developed to allow short strings of DNA to be synthesised much more easily. In the 1990s, automated machines began to appear which allowed a technician to feed raw sequence data in at one end and receive short fragments of DNA from the other. Advances in the interim have increased the length of the stands which can be produced to around 40,000 base pairs, reduced the time taken to produce the strands, decreased the number of errors which appear in the finished strands and allowed for strands to be joined together to form entire genomes. Experiments in 2002 and 2003 demonstrated that it was possible to assemble the entire genomes of viruses from scratch and that these viruses could then function as well as their natural counterparts (see, for example, the polio virus experiment in Annex II).

19. Semiautonomous DNA synthesis machines can currently produce long strands of DNA sequence and have had their error rate reduced to 1 in 10,000 base pairs. Older versions of these machines are already appearing for sale on Internet auction sites for between $5,000 and $10,000. Simple DNA sequencers can be built from scratch from commonly available components using instructions available on the Internet for around $10,000. Current restraining factors on DNA synthesis are the cost and time taken to produce the strands of DNA. Recent statistical analyses of cost/time trends indicate these factors halve every 12 to 18 months. Currently, the cost for DNA fragments is around $0.10 per base pair. For context, if the entire structure of the smallpox virus were to be produced (in numerous segments) it would cost in the region of $18,600. This progress has been helped in part by the commercialisation of the technology. Gene sequencing companies, from which DNA fragments can be purchased online and express delivered, have appeared throughout the world. An exercise which would have taken a dedicated laboratory numerous human-years to produce by hand at the time of the Second Review Conference can now be achieved at a moderate cost almost immediately.

DNA Silencing

20. Plants, fungi and animals (including humans) all share an ancient defence against certain types of virus. The presence of certain viral genetic information (known as dsRNA) in a cell starts a mechanism (known as RNA interference or RNAi) to stop the dsRNA replicating, by interfering with the process by which genetic material is read and converted into a product. This process was described as recently as 2001, and researchers soon realised it could be adapted for use as a laboratory tool. By creating dsRNA which corresponded to a specific DNA sequence they could fool the defence mechanism into stopping the sequence from being translated into a product. In other words it is possible to selectively switch off the operation of a given sequence. This capability became increasingly important as the amount of genetic sequence information, with unknown function, increased dramatically thanks to developments in *DNA sequencing*. Being able to switch off a sequence or gene at will allowed scientists to see the effect of its absence on a biological system. This permits them to identify its function. As an example of the power of this tool, by May 2003 researchers had used this approach to determine the function of

1,722 genes of a species of worm, many of which were previously unknown. A project is already underway to use this technology to determine the function of every gene in the human genome.

21. *DNA silencing* also has therapeutic applications. Genes can be associated with diseases. DNA silencing can switch off these genes, alleviating symptoms, preventing a disease from taking hold or curing it. Efforts are already underway to use DNA silencing to counter HIV, hepatitis and cancer. An experiment in 2004 also used these techniques to reduce cholesterol in mice. Clinical delivery remains a problem. Recent research suggests scientists are well on their way to overcoming this hurdle. One approach is to use a virus to deliver the RNAi.

DNA Shuffling

22. In order to create a sequence with enhanced properties, traditional genetic engineering relied on being able to cut a DNA sequence from one place and insert it into another location. This involved a directed attempt to combine properties from two separate sequences to create a third with properties more efficient than either of its antecedents. This process was known as directed evolution. Using genetic engineering it was necessary to develop each new sequence one at a time, combining sequences by hand and then screening for combinations with the desired characteristics. The process was repeated several times with successive sets of offspring to optimise the process. *DNA shuffling*, on the other hand, takes a library of related versions of the same sequence (such as genes from related species), breaks them apart and then recombines them into new versions of the basic sequence. It effectively allows for the simultaneous mating of numerous species. It produces a higher yield of functional offpring than can be achieved from the older approach. DNA shuffling improves the efficiency with which a wide diversity of genetic sequences can be derived.

23. DNA shuffling was used in 2002 to combine sequences from four microbes to produce a new sequence with 270 to 540 times greater activity than the best parental sequence. Results from DNA shuffling also suggest that the best combination of parents might not be those most closely resembling the offspring. (Counterintuitively, if you want a sequence with a good part A and a good part B, it may not be best to start off with one parent with a good part A and another with a good part B.) This complicates attempting the same process using the directed evolution. Even if it is possible to obtain the optimised sequence using older techniques, DNA shuffling produces the same results significantly faster.

24. DNA shuffling has progressed to the point where it is now possible to shuffle entire genomes. Experiments have already been completed using related bacteria. The results from a single shuffle were comparable to 20 generations of directed evolution. Work has been carried out on improving human molecules. Researchers in 2003 succeeded in producing a human cytokine (a group of molecules involved in signalling and the immune system) with ten times the activity of the one that appears naturally. These techniques have also been employed to optimise viruses for use in *gene therapy*. The limiting factor on DNA shuffling remains the ability to screen for and isolate offspring with the most enhanced desired properties. Developments in *high-throughput screening* are slowly addressing these shortcomings.

Genomic Medicine

25. Advances in the understanding of genomics have demonstrated that genetic sequences play an important role in disease, both for the pathogen and those infected. The genome of a pathogen provides information about its infectivity, virulence and other disease-determining factors. This allows novel detection technologies, diagnosis mechanisms, prophylactics and therapeutics to be devised. The genetic sequence of those infected confers a disposition to certain diseases and also explains why certain therapies are less effective in specific individuals and why some patients suffer unusual or extreme side effects. This offers the possibility of developing prophylactic or therapeutic regimes to match a patient's specific genetic make-up. The required research is already underway. A pre-requisite of effective *genomic medicine* is a catalogue of human genetic diversity. Such an effort is underway in the form of the HapMap Project.[3] Information generated by the international effort is freely available in the public domain. Developments in genomic medicine have already indicated that certain drugs are more efficient in certain geographic regions, offering a capacity for ethnic or geographic specificity.

26. Genomic medicine will not develop its full utility until the costs and time taken to sequence a genome are considerably reduced. There may be some shorter-term benefits from observations based upon common ethnic or geographic traits. This would allow drugs to be optimised for a sub-population. Tangible uses are already being found for genomic medicine in detecting, diagnosing, preventing and treating disease.

Proteomics

27. If *genomics* is the study of all the genes in an organism, then *proteomics* is the study of all the proteins encoded by those genes. It incorporates elements of both their structure and function as well as how they interact to regulate biological systems. Proteomics is related to functional genomics, as it examines the function of genes, specifically those encoding proteins. Proteins form the basis of the majority of biological functions. They are closely related to disease – either because they cause it (for example, the anthrax toxin is composed of three proteins), or because they are targeted in the host (equally the receptor the anthrax toxins binds to on human cells is a protein). It has also been postulated that certain proteins can act as self-replicating, infectious pathogens, better known as prions. Prions are believed to cause a variety of neurodegenerative disorders in animals and humans.

28. One rapidly expanding area within proteomics is that of comparative studies. Proteins from different growth conditions, strains or species can be labelled and detected. This approach allows the identification of proteins that have a role in virulence, interaction with the host or the environment, and antibiotic resistance. Proteomics also provides information which can be used to improve detection systems, diagnoses, vaccines and therapeutics. Studies have already led to the identification of novel drug and vaccine targets, including for the parasite which causes malaria. They are also contributing to characterisation of pathogenicity, the study of host-pathogen interactions, including the humoral immune response, and the evaluation of mechanisms of action for anti-microbials.

[3] See: www.hapmap.org.

29. Traditionally, proteomics has involved a combination of gel electrophoresis and mass spectrometry. Advances in these and other fields, including nucleotide sequencing, complex mixture analysis, chip-based approaches, and algorithms, have enhanced proteomics capabilities. New tools have become available for studying the functioning and interactions of proteins, most noticeably *high-affinity binding reagents*. There have also been considerable advances made in isolating small molecules from complex environments, even if they are present in very small quantities. Handling proteins from dangerous organisms has also progressed. It is now possible to chemically synthesise the relevant gene, and express and purify it from a host cell, negating the need for the presence of the organism itself, possibly reducing the level of the safety and security precautions required for the work. Other developments include progress in non-gel techniques for creating, separating and analysing protein mixtures, even those it had been previously difficult to express. It is now possible to mark proteins with tags so small that they do not need to be removed after production, and to recover secreted proteins directly from feedstock without need for centrifugation.

High-Affinity Binding Reagents

30. Recent advances have provided a capacity to inhibit or modulate the expression of specific protein targets. Two classes of these high-affinity binding reagents are currently used: aptamers and tadpoles. Aptamers are short, single-stranded nucleic acid or peptidic ligands. They have been used in target validation, detection reagents and as functional proteomics tools. Their use as therapeutics has also been explored. They have been tested in animal models to inhibit blood clot formation and to treat age-related degenerative changes to the eye. Tadpoles are comprised of a protein head with a DNA (oligonucleotide) tail. Their ability to bind to specific targets (such as one of the three proteins which comprise the anthrax toxin) combined with their easy quantification (due to their DNA tails) has led to their development for use in disease diagnosis, surveillance and environmental detection (see *detection technology*).

Bioinformatics and Computational Biology

31. Developments in *genomics* and *proteomics* have produced a huge amount of information. For example, in August 2005 the three largest repositories of gene sequence information contained 100 billion bases of sequence data from 165,000 organisms. The contents of one of these repositories, Genbank, doubles in size every 18 months. The synergy between increasing knowledge and the open exchanges of ideas and information is accelerating advances in medicine, industry, and agriculture.

32. The amount of data available and the complexity of biological interactions means that manipulating it by hand is no longer feasible. The decreasing cost and increasing power of computers, as well as the development of specific platforms for analysis and data management, have provided a medium for handling this information. The application of large-scale data analysis techniques in this field has become known as *bioinformatics*. Bioinformatics is creating new scientific and commercial opportunities. The ability to combine bioinformatics with *high-throughput screening* technologies offers the potential to reduce the time it takes to conduct research, as well as the time taken to convert a discovery into a viable commercial product.

33. *Computational biology* goes beyond data analysis to look at the broader interface between computing and biology. Four key points of overlap have been identified. First, there are

computational tools (software or hardware) which allow biologists to acquire, store, manage, query and analyse biological data to solve very specific and precisely defined problems. Second, there are computational models which can be used to test insights, make quantitative predictions and help interpret experimental data. Third, computational perspective or abstraction can provide well-understood constructs which can be used to characterise the biological function of interest. Fourth, scientists increasingly rely on high-end general computing centres; well-managed, accessible data repositories; digital libraries; high-speed networks; and data acquisition technologies, such as genome sequencers.

34. The types and format of the information becoming available varies widely and includes: sequences; graphs; geometric information; scalar and vector fields; patterns of organization; constraints; images; and prose. Advances in bioinformatics and computational biology have provided tools to deal with this, including:

(i) Increased capacity for the storage and analysis of large amounts information on low-cost platforms;

(ii) Enhanced efficiency for data distribution and communications technology permitting the sharing and operation of large, complex data sources with a wide geographic distribution;

(iii) Internet-based tools enabling simple world-wide access to biological information;

(iv) Common data formats that permit the integration of multiple data streams; and

(v) Enhanced search methods capable of dealing with a variety of different types of information stored in geographically disparate locations.

Systems Biology

35. *Systems biology* has been described as the expansion of physiology into unprecedented levels of complexity. Instead of studying how the body works on the visible scale, it has been expanded to the molecular scale. It is based upon the premise that observable biological behaviour is caused by a complex system of dynamically interacting molecular events. These interactions add another layer of complexity to a biological system. For example, the number of genes found in humans (such as review conference delegates) is not significantly larger than the number found in simpler organisms (such as worms). How can the obvious difference in complexity between a conference delegate and a worm be explained if they have similar numbers of genes, and therefore similar numbers of components? The answer is that it is a complexity of regulation as opposed to a complexity of structure. Interactions between the various components were found to be more complex in humans than in worms. Systems biology is the study of the complex interactions between networks of molecules in a discrete biological system.

36. Previously scientists may have examined a single facet of a biological pathway, for example, the control of a cellular response to infection. Systems biology permits them to look more broadly at the affect of a particular stimulus on multiple different pathways, for example other cascades initiated by the cellular response to infection which in turn have a counter-regulatory effect on the one identified first. These studies are demonstrating that there are many

molecular interactions which have not been previously identified and that there exist entirely new regulatory mechanisms using novel signalling pathways.

37. Systems biology entails a four-stage process. First, information about the system is gathered using variety of tools, including *high-throughput screening* technology, *genomics* advances, *proteomics* devices and data mining of *bioinformatics* databases. Second, the amount of information involved is too large to manipulate by hand and is, therefore, manipulated using *computational biology* techniques in an attempt to quantify all the molecular elements that make up the system and map them into a single graphical network model. Third, the model can be used to establish how the manipulation of the system affects its functioning. Finally, the computational predictions can be checked against empirical experimentation, and the data produced can then be used to improve the model. The ultimate goal would be to produce a model of the system which accurately reproduced the entire system and which allowed accurate simulated experimentation.

38. Systems biology, therefore, contributes to *drug discovery, design and delivery*, especially through *rational drug design*, by assisting the modelling of effects of molecules interacting with the system. Systems biology also has wider medical implications, as it is possible to frame almost all disease as a manipulation of biological systems by genetic, molecular or environmental factors. By studying how a system suffering from a disease differs from a healthy counterpart, information on how disease-related processes interact and are controlled can be inferred. This is providing possibilities for new diagnostic and therapeutic approaches as well as opening the door to improved *genomic medicine*.

39. Fully integrated systems biology is still in its infancy. There are still shortcomings in the ability of computational tools to handle the variety and quantity of information currently available in an efficient manner. There has, however, been significant progress in examining the interactions of complex regulatory mechanisms in *bioregulation*, especially in neurobiology and immunology.

Bioregulation (Neurobiology and Immunology)

40. Scientific publications have shown that biologically-active biochemicals and bioregulatory peptides can modulate physiological systems and processes such as the brain and the immune system in very precise ways. There is considerable commercial interest in these bioregulatory compounds as they offer novel opportunities for the relief of pain, depression, and a wide range of mental disorders. There are indications that through their use it will be possible to manipulate perception, sensation, cognition, emotion, mood, volition, bodily control and alertness. Use of bioregulators has been considered limited in the past because the compounds involved tend to be environmentally unstable. Developments in *microencapsulation* are allowing the commercial development of these agents.

41. Efforts to identify the molecular circuits and control systems which regulate the functioning of the body, in addition to studies designed to show what perturbations cause various alterations and disease states, have identified a wide range of targets for bioregulators. Considerable efforts have been made to study what products are produced in various disease states (transcriptional profiling). Such studies extend to how pathogens overcome various immune responses or treatments (such as antibiotics). An increasing understanding of how the

structure of biologically active compounds affects their affinity for and reactivity with specific molecular targets provides capabilities for optimising their development. For example, it is becoming increasingly apparent that the three-dimensional folding of biologically active compounds plays a key role in their function. Advances in the production of these molecules allow for more complex structures to be produced. The availability of large libraries of biologically active compounds allows for *high-throughput screening*. As a result, it is becoming increasingly feasible to identify a compound to perturb a specific biological process. Industry has generated much of the relevant information to date, and views it as commercial proprietary information. Current efforts are underway to make more of this information publicly available to assist researchers around the world in their work and in the identification of new drug targets.

Drug Discovery, Design and Delivery

42. There have been numerous developments over the last five years with respect to identifying, creating and making use of biologically active substances (drugs). This is leading to new prophylactics, such as vaccines, and therapeutics, such as anti-microbial agents. Significant funding has been injected into the field for biodefence work. Recent efforts have also made progress on sub-unit and DNA vaccines.

Combinatorial Biochemistry

43. The development of *high-throughput screening* technologies has made it possible to assess, in a short time, the potential for use as drugs of large numbers of biochemical structures. A parallel capability to produce numerous different biochemicals to be screened was therefore desirable. Traditional biochemistry allowed proteins to be built amino acid by amino acid, in a process known as solid-phase synthesis. New techniques contain the individual constructions within "mesh bags", allowing their contents to be combined after each step to significantly increase the diversity of the amino acid sequence (solution-phase parallel synthesis). It has also become possible to tag each construction so that the order of the amino acids can be read easily and their structure understood. Solution-phase parallel synthesis also increases the range of chemical reactions which can be performed, significantly increasing variations of structures which can be created. Recent developments have also refined the purity of products. These advances, when combined with automation, miniaturization and *bioinformatics* permit the rapid creation of large numbers of synthetic compounds (libraries).

44. Early developments in drug design suggested that if a large enough library was screened, a biochemical structure with desirable characteristics would eventually be found. Advances in *combinational biochemistry* provide for libraries of an unparalleled scale. Rewards from this approach appear to have been few, however, and recent advances have seen a move away from the use of large, general libraries towards more focused, less diverse libraries where variations on a theme are used to screen for optimisation against a specific characteristic. This has proven particularly useful for the further development of promising (or "lead") compounds in drug discovery. It allows variations of the lead compound to be assessed for efficiency and other desirable characteristics. Information generated from *combinatorial biochemistry* effectively describes how chemical modulation of compounds affects biological activity. It therefore also provides data relevant to *systems biology* and *rational drug design*.

Rational Drug Design

45. Advances in the understanding of biological systems, especially in the interactions of the various components, have highlighted that a number of molecules are important for the maintenance of health or play a role in the onset of disease (see *bioregulation*). Developments, such as those in x-ray crystallography and nuclear magnetic resonance imaging (NMR), have also permitted the structure of target molecules and the structures which interact with them to be mapped. Understanding how these interactions occur opens the door to designing molecules which will interact with specific targets to produce a desired effect. If the structure of a target molecule is known, it is possible to draw upon known interactions to predict what a drug that will interact with it will look like. The more data there is to draw upon, the more accurate the prediction is likely to be.

46. Developments in *bioinformatics and computational biology* have allowed the construction of computer hardware and software to model these interactions. These computers allow virtual screening to be run on massive libraries far faster than can be achieved using even the most advanced biochemical techniques. The product is a rationally-designed molecule which when designed can be fed into the traditional drug development programme. In practice, this computerised rational drug design is used in parallel with *combinatorial biochemistry* and *high-throughput screening*. Traditional approaches may be used to identify a lead compound, which can then be optimised using rational drug design techniques.

Drug Targeting

47. The efficiency of a drug depends upon its ability to reach the part of the body it needs to interact with. The efficiency of a drug can be further improved by ensuring that it only interacts with its desired target. This means that there is more of it to act, because it has not been wasted interacting with molecules which were not targeted. It also minimises the risk of unwanted reactions or adverse effects.

48. Although a truly efficient, selective targeted delivery system remains elusive, considerable progress has been made. There is a growing library of targeting systems for different targets. Three approaches form the basis of current efforts to improve drug targeting. First, drugs have been encapsulated with a structure to enhance targeting (see *microencapsulation*). Second, numerous viruses and bacteria have been manipulated for use as delivery devices, taking advantage of their natural ability to selectively infect specific cells (for example, see *gene therapy*). Third, drugs have been joined to carrier molecules designed to recognise specific targets. All of these systems rely upon molecular recognition systems. They selectively recognise and bind to their targets before releasing the drug. These delivery systems can be further enhanced to increase the take-up of the drug by the target.

Microencapsulation

49. Coating biologically active agents can protect them from environmental factors, such as evaporation, oxidation and contamination. It can also improve target recognition and therefore specificity. Coatings can be made from a variety of materials including organic polymers, hydrocolloids, sugars, waxes, fats, metals or inorganic oxides. The coatings are designed to preserve the functionality of their contents until they arrive where they are needed. This requires

a variety of release mechanisms, to ensure that they are able to interact when they reach the desired location. Release mechanisms developed to date include controlled release, delayed release, targeted release (see *drug targeting*), biodegradable release, and salt-induced release. There are two common approaches to microencapsulation: the physical and the chemical. Physical microencapsulation can include spray drying, fluid bed coating, co-extrusion and rotary-disk atomisation; while chemical microencapsulation commonly uses polymerisation, phase separation, solvent evaporation and coacervation.

50. This technology is not new but has found many new uses since the Fifth Review Conference. As a result, it is becoming increasingly available commercially. It is currently used in water treatment, food production, agriculture and the cosmetics industry, as well as in bioremediation and hazardous waste management. It is also being developed as part of a cancer treatment and for the treatment of damaged skin.

Biopharming and Bioproduction

51. Biopharming is the use of genetically altered plants which can be grown in large numbers to produce complex biologically active molecules without the need for industrial facilities. It offers the possibility of a low-technology, cost-efficient form of mass production for biological compounds. The gene for a desired substance is inserted into the plant, which then can be grown naturally and either used as a delivery device itself (in the case of food plants) or have the molecule harvested and processed. Using plants in this manner reduces the cost of production, lowers the technical threshold for production (once the plant has been engineered) and allows for the construction of complex biologically active structures (such as vaccines or anti-bodies) which could not be produced using traditional approaches, or for which such production was prohibitively expensive.

52. Transgenic (genetically engineered) plants - including rice, potatoes, maize, fruits, vegetables, and tobacco - have been developed to produce: beta carotene; human milk proteins; cholera antigens; antigens for diarrhoeal pathogens; the hepatitis B vaccine; AIDS antigens; sub-unit vaccines for rabies; human glycoproteins; human haemoglobin; and hepatitis B antigens. Plants are also in development for the production of antibodies to protect against biological weapons. Relatively long lead times and high regulatory costs have meant that licensed applications are uncommon, and biopharming may not become a standardised approach for some time. Protein expression systems for artificial genetic sequences have also been developed in bacteria, yeast, filamentous fungi, insects and mammalian tissue.

Drug Delivery

53. There are the three common routes of entry for a drug into an organism: inhalation, oral, and trans-dermal. Inhalation offers advantages, including the rapid speed of onset, a more even distribution, and the potential for wide-area coverage. Delivery by inhalation relies upon aerobiology and aerosol technology and has been developed for administering prophylactics and therapeutics to humans, animals and plants. Advances have been made in two important areas: the preparation of biological structures to make them suitable for aerosol delivery, and the efficiency of the delivery devices themselves. There have been important developments in powder technology and particle engineering which have improved particle dispersibility, control of particle morphology, and physical and chemical stability. For example, the development of

supercritical fluids (SCF) provides compounds with the properties of both liquids and gasses, which also overcomes complications in purifying the active agent. Equally, advances in the construction of large porous particles allows optimised aerosol delivery of substances larger than those traditionally considered suitable for effective inhalation. There are three different aerosol delivery systems in current usage: propellant metered-dose inhalers, dry powder inhalers, and nebulizers. Each of these three systems has had to face hurdles in utility. Advances since the Fifth Review Conference have addressed many of these issues.

54. Aerosol technologies are increasingly applied to the treatment of disease. They are in widespread use for the treatment of asthma and chronic obstructive pulmonary disease. Similar models are in development for the treatment of diabetes, human growth hormone deficiency, prostate cancer and endometriosis. Applications have not been limited to human delivery. Advances in this field have also led to the wide-area dissemination of agents as pesticides, such as the aerial distribution of aerosols of *Bacillus turingiensis* to protect forests from the spruce bud-worm (see *biological pest control*). Animal studies have also been carried out to assess the biological impact of inhalation of toxic particles. There have also been aerosolization studies conducted for disseminating bacteria over large areas of water as part of water treatment. Aerobiology has also been used to assess the plume characteristics of biological aerosols for use in farming and agriculture. These technologies are, therefore, more widely available than they were at the time of the Fifth Review Conference.

55. Progress has also been made in oral delivery techniques. Problems with oral delivery, or ingestion, stem from the ability of the stomach and intestines to break down biological structures. Biologically active compounds are normally denatured or even digested before they can be taken up and travel to the areas where they should act. Advances in *microencapsulation* have, to a large extent, overcome these problems. Recent research has provided the ability to coat proteinaceous molecules so that they can pass through the stomach unaffected and bind to the mucosal lining of the intestines, where they can enter the blood stream.

56. Until recently, trans-dermal penetration has not been a feasible delivery mechanism. It required a vector, such as an insect, to break the skin and deliver the biologically active compounds into the bloodstream. The approach has become more sophisticated both through the development of transgenic (genetically manipulated) insects and through an increased understanding of how certain vectors, such as the mosquitoes which spread malaria and West Nile Virus, overcome insecticides. There have also been significant improvements in chemical solutions for crossing the skin barrier. Certain technologies have already been proven and found commercial applications, such as the nicotine-administering patches to help smokers give up. Work carried out with chemical penetration enhancers has increased 100-fold the trans-dermal penetration of relatively large proteins.

Synthetic Biology and Biological Engineering

57. Many of the advances discussed in this paper have involved the development of novel enabling technologies and the application of engineering principles to biology, such as those required to enhance automation and reduce the time and cost of often-repeated activities. This growing overlap between biology and engineering has facilitated a new approach to the life sciences, *synthetic biology*, which focuses on using knowledge of biological systems to begin to construct them from scratch. Key to synthetic biology is a requirement for biological components

which can be combined to produce a biological system in a manner reminiscent of the way in which a circuit board is compiled from pre-packaged electronic components.

58. Synthetic biology has attracted both biologists and engineers, who tend to view it in very different lights. Engineers see synthetic biology as a way to fabricate biological devices to do what no current technology is able to. Biologists see it as a powerful new way to learn about the principles underlying biological function. Both come together to model biological systems with desirable properties, create these systems in reality, test them for functionality and adjust them until they work properly. The empirical refinement required is currently considerable, although it is likely to decrease with experience. Throughout the process the knowledge gained is fed back into design and construction. As a result, understanding of the principles of the operation and design of biological systems is advancing rapidly.

59. Synthetic biology has absorbed concepts from the engineering sciences, such as standardisation and abstraction. These concepts underlie any attempt to establish a process for function-oriented design and are important for the development of biological components. In other engineering disciplines, it is possible to take a component off the shelf and know that it is compatible with those already used in a device. For synthetic biology, efforts are underway to ensure that the various biological components constructed around the world are interoperable and use consistent, standardised parameters.[4] Abstraction is also important for the development of component-driven biological systems. An individual does not have to understand every step of a process to contribute to it. A complex task can be broken down into different strata and it is possible to become an expert at any level without an in depth knowledge of other levels. For example, it is possible to design an electrical circuit board without being familiar with how each component can be produced. This makes the process more accessible, helping it to be commercialised.

60. Synthetic biology has developed as a concept since the Fifth Review Conference. There have been notable successes in using it to design biological systems. For example, in 2003 a bacterium was re-engineered to change colour when it grew in the presence of an explosive. This was developed for use in locating unexploded ordnance or landmines. The re-engineered bacteria could be sprayed over an area and where it grew in the presence of TNT, it would glow fluorescent green. Similarly, in 2004 researchers produced a DNA computer to search for early signs of prostate and lung cancer and control the delivery of biologically active drugs.

Other Relevant Developments

Nanotechnology

61. *Nanotechnology* has been described as "a clever means of making incredibly small things". Although many of the advances relate to the manipulation of non-organic matter and therefore fall outside the life sciences, structural elements of biological systems do fall within the size requirements for nanotechnology. Manipulating biological systems to create devices with a specific purpose might therefore be considered nanotechnology.

[4] For example: Registry of Standard Biological Parts, the Endy Lab, Massachusetts Institute of Technology, see: http://parts.mit.edu.

62. Of all the properties found in biological systems, self-assembly has received particular attention from nanotechnologists. DNA and its synthetic versions have been used to make objects, lattices and devices. These have been used both as components and as joining agents in the construction of complex structures. Current nanotechological applications for biological systems include molecular imaging and detection, reporters for therapy efficiency determination, multifunctional therapeutics, disease prevention and control and various enabling technologies. Under development are architectural controls and scaffolding, nanomechanical devices, and self-replicating nano-systems. Devices built to date include a synthetic DNA structure which cuts up RNA molecules, and contact lenses which release precise dosages of medication to treat glaucoma.

Gene Therapy

63. There is a link between the functioning of genes and disease. Faulty genes working in an abnormal manner cause many diseases. Gene therapy is the attempt to replace these faulty genes with a healthy copy. It is built upon the ability of certain viruses to copy DNA back into a host's genome. In gene therapy, a vector is used to transport the healthy gene to the target cells. Vectors under development include a range of engineered viruses, including retroviruses, adenoviruses, adeno-associated viruses and herpes simplex viruses, which have had their disease causing genes removed and the space made available for insertion of the healthy gene. Viruses are often selected for use as vectors because of their ability to target certain cells (see *drug targeting*). A considerable library of vectors targeting different tissues has been developed. Numerous non-viral delivery mechanisms have also been created. They require both direct administration and a large amount of DNA, and are only compatible for use with certain tissue types.

64. Although considerable progress has been made on gene therapy since the Fifth Review Conference, especially in the development of targeted vectors, it remains an unproven technology. Notable successes have been achieved in animal models, including restoring hearing to deaf guinea pigs. The few human clinical trials which have taken place have been, however, less successful than hoped. There are still significant issues to be overcome for delivery systems and gene expression rates.

Genetic Engineering of Viruses

65. Many of the viruses used in *gene therapy* contain genetic material similar to that found in other organisms, including humans (DNA). Other viruses, however, use an alternate form of genetic material (RNA). While the ability to manipulate and genetically engineer DNA is well established and in common use, manipulating viruses which use RNA poses more of a challenge. RNA is less stable and is less amenable to recombinant genetic techniques. Alternative strategies to deal with RNA viruses have been developed. RNA can be copied into a complementary DNA version through reverse genetic engineering (sometimes requiring the presence of certain viral proteins or helper viruses) which can then be inserted into bacteria where they can be manipulated using traditional techniques. There has been significant progress since the Fifth Review Conference in the size of sequence with which this can be accomplished. It is now possible to reverse-genetic-engineer the very largest RNA viruses, such as the coronavirus which causes SARS. After modification, the DNA can then be removed and inserted into a system to convert it back to RNA, where it can be inserted into a permissive cell (again sometimes with certain viral proteins or helper viruses) where it will be read and the resulting viruses

constructed. This mechanism has improved understanding of how viruses replicate and has presented options for the development of new vaccines and vectors.

66. Similar technology has recently been used to recreate the influenza virus which caused the 1918-1919 pandemic. The resulting manipulated virus included surface structures from the original virus which had been artificially reintroduced. The addition of these structures converted formerly non-pathogenic viruses into a pathogenic strain in animal models. Furthermore, those infected with the virus demonstrated symptoms characteristic of the 1918-1919 strain (see the 1918 influenza experiments in Annex II).

Anti-Viral Drugs

67. Effective, safe anti-viral drugs, with properties similar to antibiotics, remain elusive. Although progress has been made since the Fifth Review Conference, such as the development of a number of drugs reportedly effective against pox virus infection, their side-effects may well preclude these drugs from prophylactic use. Alternative strategies, such as the use of monoclonal or polyclonal antibodies, for use against viruses have also advanced, with reports of effective treatments for both pox viruses and Venezuelan Equine Encephalitis. There has also been progress in developing non-specific immunostimulators, which prompt a general defensive response through immunomodulators or cytokines. One successful example of this approach is interferon-alpha used in the treatment of HIV-AIDS.

Detection Technology

68. Advances and enabling technologies drawn from a number of fields have led to significant improvements in detection and identification technologies. Relevant developments include: an increased range of fluorescent detection molecules; more rapid polymerase chain reaction (PCR) amplification strategies; refinement of gene probe systems and enhanced specificity; advances in microarray technologies; freeze-drying or lyophilisation of reagents; more robust antibodies; the advent of aptamers and antigen recognition; nanotechnolgies, including the use of quantum dots and gold nanoparticles; evanescent wave detection technologies; light scattering surface plasmon resonance; metal clad leaky waveguide technology; improved detection limits through the use of ultrasound, electrophoresis and dielectrophoresis; bioluminescence technology; and auto-fluorescent detection techniques.

69. Since the Fifth Review Conference, biological detection equipment has become more sensitive, easier to use and cheaper. It is increasingly miniaturised and autonomous. There are increasing numbers of commercially available technologies suitable for use in field environments. Hand-held devices for rapid diagnosis and near real-time environmental sampling have been developed. Numerous approaches to detection are currently being pursued, including through the use of: antibodies; *high-affinity binding reagents*; optical detection; bioluminescence; dip sticks; and *nanotechnology*.

Biological Pest Control

70. Advances in a number of fields have stimulated research into biological pest control systems (biopesticides). It is unlikely that such systems will replace chemical approaches in the near future, as there are still problems with formulation, speed of action and efficiency. The most

commonly cited biological pest control system currently developed is the use of *Bacillus thuringiensis* to control plant pests. Other toxins under development for use as biological pest control agents include those found in *Photorhabdus luminescens*, *Pseudomonas entomophila*, and *Bacillus nematocida*. Research to produce transgenic crops incorporating the pesticidal toxin found in *Bacillus thuringiensis* has been undertaken. Current work centres upon finding more powerful toxins and expanding the range of targets. The toxin found in *Bacillus thuringiensis* has also been inserted into other microorganisms, such as baculovirus or other bacteria, for use as biological control agents.

Annex I

Identifying Experiments of Concern

1. There have been numerous attempts, including by professional scientific bodies and by States Parties, to characterise experiments of concern. Some of these experiments have already been conducted and published in peer-reviewed scientific literature; others are considered theoretically possible. All are technically challenging and require well-funded and coordinated research programmes.

2. The United States National Academy of Sciences included in its report *Biotechnology Research in the Age of Terrorism*, published in 2004, a list of seven experiments of concern, namely those which would:

(i) Demonstrate how to render a vaccine ineffective;

(ii) Confer resistance to therapeutically useful antibiotics or antiviral agents;

(iii) Enhance the virulence of a pathogen or render a non-pathogen virulent;

(iv) Increase transmissibility of a pathogen;

(v) Alter the host range of a pathogen;

(vi) Enable evasion of diagnostic and detection modalities;

(vii) Enable the weaponization of a biological agent or toxin.

3. In its contribution for the preparation of this document, Australia provided the following list of experiments of concern:

(i) Rendering a vaccine ineffective;

(ii) Conferring resistance to therapeutically useful antibiotics or antiviral agents in pathogenic organisms;

(iii) Enhancing the virulence of a pathogen or rendering a non-pathogen virulent;

(iv) Increasing the transmissibility of a pathogen;

(v) Altering the host range of a pathogen;

(vi) Enabling the evasion of diagnosis and/or detection by established methods;

(vii) Undertaking genetic sequencing of pathogens;

(viii) Synthesising pathogenic microorganisms;

(ix) Large-scale protein production employing heterologous expression systems (and associated production technology);

(x) Optimisation of live attenuated vaccine production processes;

(xi) Enabling the weaponisation of a biological agent or toxin;

(xii) Any experiment with the smallpox virus.

Annex II

Actual Experiments Often Quoted as Being Particularly Relevant to the Convention

1. Four experiments are often discussed when scientific and technological developments relevant to the Convention are considered:

(i) *Expression of Mouse Interleukin-4 by a Recombinant Ectromelia Virus Suppresses Cytolytic Lymphocyte Responses and Overcomes Genetic Resistance to Mousepox*, by Ronald J. Jackson, Alistair J. Ramsay, Carina D. Christensen, Sandra Beaton, Diana F. Hall, and Ian A. Ramshaw. Published in the *Journal of Virology*, Vol. 75, No. 3, February 2001, pp.1205-1210;

(ii) *Chemical Synthesis of Poliovirus cDNA: Generation of Infectious Virus in the Absence of Natural Template*, by Jeronimo Cello, Aniko V. Paul, and Eckard Wimmer. Published in *Science*, Vol. 297. No. 5583, 9 August 2002: pp.1016-1018;

(iii) *Characterization of the Reconstructed 1918 Spanish Influenza Pandemic Virus*, by Terrence M. Tumpey, Christopher F. Basler, Patricia V. Aguilar, Hui Zeng, Alicia Solórzano, David E. Swayne, Nancy J. Cox, Jacqueline M. Katz, Jeffery K. Taubenberger, Peter Palese, and Adolfo García-Sastre. Published in *Science*, Vol. 310, 7 October 2005, pp. 77-80; and

(iv) *Characterization of the 1918 influenza virus polymerase genes, by* Jeffery K. Taubenberger, Ann H. Reid, Raina M. Lourens, Ruixue Wang, Guozhong Jin and Thomas G. Fanning. Published in *Nature*, Vol. 437, 6 October 2005, pp. 889-893.

In addition, work is currently underway on experiments involving avian influenza.

The Mousepox Experiment

2. The researchers were attempting to produce a virus which could be used to control a mice population which had gotten out of control causing significant damage to grain production. They were attempting to alter a mouse pathogen, mousepox, by inserting a protein found in eggs in female mice. It was an attempt to provoke the immune system of the mice to sterilise themselves. To enhance antibody response in the mice, they also inserted another gene, that encoding IL-4. The virus produced was 100% lethal in infected mice, including those which were either genetically immune to natural mousepox and those which had been vaccinated against it. The experiment raised concerns that a similar effect could be reproduced in related viruses, including smallpox.

The Polio Virus Experiment

3. In 2002, a group of researchers managed to artificially create a live and pathogenic polio virus from its genetic sequence. The sequence information was obtained from an open on-line repository. It was broken down into a number of smaller segments. The sequences of the segments were submitted over the internet to commercial DNA synthesising companies, who

mailed the physical DNA back to the researchers, who were then able to join the segments back together to make the genome of the pathogen. The genome was then used to generate the actual pathogen. Follow-up experiments have started to use this process to re-create larger viruses, and plans to synthesise a bacteria have also been revealed. Despite advances in technology over the last four years, it is still not possible to use this technique to reproduce all viruses.

The 1918 Influenza Experiments

4. The research reported in these 2005 publications indicated that two research teams had succeeded in recreating the strain of influenza virus which caused the 1918-19 pandemic. They reassembled this extinct pathogen from its sequence, which they compiled from tissue samples from those infected. Analysis of its sequence and various structural elements indicate that it was likely a mutated avian influenza virus. Research has also been undertaken to characterise what structural properties allowed it to infect humans and kill them so efficiently.

Avian Influenza Experiments

5. As an extension of the increasing knowledge of influenza viruses, efforts have already begun to force the strain responsible for the current pandemic of bird deaths to be able to infect and become transmissible between humans. This work is aimed at improving the understanding about how this host species change occurs in the wild and what it takes to make a pandemic human super-flu. A number of laboratories have reported trying to mix the genes of the avian influenza virus with human influenza viruses.

_____