SCIENTIFIC AND TECHNOLOGICAL DEVELOPMENTS RELEVENT TOTHE BIOLOGICAL WEAPONS CONVENTION

Submitted by the United States of America

1. The rapid advances being achieved in understanding living systems contribute significantly to healthcare, agriculture, industry, and a number of other areas. However, many of these same advances could be used malevolently for biological and chemical weapons purposes. At the same time, these advances are also critical for developing better countermeasures against the biological and chemical weapons threat. Virtually all developments in science and technology with relevance to the Biological Weapons Convention have this "dual use" character.

2. This paper draws heavily on the findings and conclusions of a recent Institute of Medicine and National Research Council study, *Globalization, Biosecurity, and the Future of the Life Sciences*, published under the auspices of the U.S. National Academy of Sciences in January 2006.[1] The goal of this paper is to summarize advances in techniques that are relevant to nonproliferation policy. Of special interest to the BWC are applications in directed molecular evolution (i.e., genetic modification), proteomics, bioinformatics, synthetic and systems biology. The number of countries which are developing and enhancing their biotechnology capabilities continues to grow as the applications continue to expand into commercial sectors and the resulting industry has expanded in both scope and products developed and marketed. The purpose of this paper is to stimulate discussion of these technologies and their relevance to nonproliferation issues.

3. Advancing technologies with relevance to the life sciences enterprise can be grouped into four general categories: (I) Many technologies are developed to create or to acquire collections of molecules with greater breadth of diversity than found so far in nature, as well as types of diversity that may not exist in nature; (II) Another useful set of technologies, characterized as "directed design," are used to deliberately generate pre-determined and specific biological or molecular diversity; (III) Some technologies are driven by efforts to understand and manipulate complex biological systems; (IV) Finally, some technologies are driven by efforts to improve capabilities for producing, re-engineering, or delivering biological or biology-derived products, and miniaturizing these processes. Developments in these four areas are equally important for the overall advancment of biotechnology.

4. The pace of advances in biotechnology will continue to increase and certain advances in areas such as genomics, proteomics, computational biology, nanotechnology, and industrial biotechnology remain a BWC concern. Nations must remain cognizant of and carefully monitor for potential abuse of the constantly evolving technologies in biotechnology.

Acquisition of Novel Biological or Molecular Diversity

5. *DNA synthesis* is a widely-used and rapidly developing technique for the *de novo* generation of genetic sequences that specifically program cells for the expression of a given protein. It is not new, but technical enhancements continue to increase the speed, ease, and

[1] The material in this paper is derived from Chapter 3, "Advances in Technology with Relevance to Biology: The Future Landscape," pages 103-158 of the report.

accuracy with which larger and larger sequences can be generated chemically. In recent developments scientists have synthesized genomes of several key viruses and bacteria, including poliovirus and the 1918 influenza virus. This raises concerns that more complex dangerous organisms, such as the smallpox virus, may be synthesized *de novo* someday and be reintroduced to an immunologically naïve population.

6. *DNA Shuffling*: DNA shuffling is a powerful technique for directed molecular evolution that is used to accelerate the rate at which one can evolve genes and proteins. It is a method for in vitro homologous recombination of genes by random fragmentation and polymerase chain reaction (PCR) reassembly and amplification. This technique generates a range of recombined or "progeny" genes that encode proteins with unique (or improved) functionality. If a series of alleles or mutated genes is used as a starting point for DNA shuffling, the result is a library of recombined genes that can be translated into novel proteins, which can in turn be screened for novel functions. Genes with beneficial mutations could be generated by serial mutagenesis and screening; however, DNA shuffling is a faster process, generating more functional mutant recombinants. Ultimately, this rapid molecular method of directed evolution will allow biologists to generate novel proteins, viruses, bacteria, and other organisms in a cost-effective manner and in a fraction of the time required with classical breeding. Application of DNA shuffling technology has been extended to small molecule pharmaceuticals, pharmaceutical proteins, gene therapy vehicles and transgenes, bacterial and viral vaccines, and laboratory animals. For example, virologists are using DNA shuffling to optimize viruses for gene therapy and vaccine applications. However, new, potentially lethal, viruses and bacteria could also be inadvertently released, or deliberately created and used for non-peaceful purposes.

7. *Bioprospecting* is the search for previously-unrecognized, naturally-occurring, biological materials for use in medicine, agriculture, and industry. Scientists collect environmental samples, extract materials (be it genetic, proteinacious, or other biologically active compounds), and then scan and search for novel genes and gene pathways with specific applications. Additionally, bioprospecting could be used to search for previously unrecognized microbes in the environment, identifying microbes that might serve as pathogens and providing an early warning for potential disease-causing agents. Conversely, bioprospecting could be used to search for dangerous pathogens and novel agents with the intent of introducing them into an immunologically-naïve population. Malign use of the technique would be limited, however, by the requirement for an assay for pathogenicity in the target population.

8. *Combinatorial chemistry* is used for the rapid creation of large numbers of synthetic compounds in order to screen them for activity against biological drug targets. These techniques generate huge libraries of chemical compounds with a diversity of shapes, sizes, and charge characteristics, and thus varied abilities to interact with and alter the properties of biologically-active proteins and macromolecular complexes. Combinatorial chemistry techniques contribute to drug discovery and development, as well as the search for better superconductors and biosensors for the detection of medically important molecules and environmental toxins. The libraries generated by combinatorial chemistry could also be quickly and easily searched by those with malign intent for compounds with the potential to interact with endogenous physiological pathways for use as biochemical weapons. Malign use of the technique would be limited by the requirement for an assay for toxicity in the target population.

9. *High-throughput (HTP) screening,* as used in this context, is used to examine large numbers of diverse biomolecular or chemical compounds quickly for properties of specific interest, in particular for the efficient identification of compounds with therapeutic potential. In recent years, the advancement of high-throughput (HTP) technologies has moved analytical research to the forefront of biology. High-throughput technologies, such as automated sequencing, mass spectrometry proteomics, transcriptomics, proteomics and metabolomics are offering approaches to capturing information about an organism at a global scale. However, this paradigm shift from single molecule experimentation to studying the entire collection of molecules concurrently is generating massive amounts of heterogeneous data sources that require advanced computational approaches to analyze and build predictive models - a substantial contributor to the rising new field of systems biology. Malign use of the technique would be limited by the requirement for an assay for harmful properties.

10. Advances in the capabilities and efficiencies of high-throughput technologies contribute greatly to the productivity and speed of advances in other sectors within the life sciences. For example, they allow for the rapid identification of biological or chemical agents that could be employed for hostile purposes; however, they also allow rapid screening of new and potentially dangerous agents. Additionally, high-throughput technologies have contributed to a striking change in the amount of genetic information available worldwide. In considering the BW implications of the availability of this information, we must assume that researchers and others worldwide will be able to access basic DNA, protein, and functional informational resources.

Directed Design

11. *Rational drug design* uses structural knowledge of drug targets to design novel chemical compounds that bind to selected sites on the surface of target molecules or mimic the structure of the target molecule, thereby competing for a receptor molecule's binding site(s). Many thousands of virtual compounds can be rapidly and effectively assessed for potential target molecule complementarity, as a prerequisite for biological activity, prior to any actual chemistry being carried out or biological assays being performed. This technique is more specific than combinatorial chemistry in that it allows the scientist to target a desired site and function, and design a drug with particular properties rather than screen through a large library of compounds looking for those properties. Due to the complex infrastructure required for research and development, rational drug design is likely to be limited to legitimate pharmaceuticals for some time, although rapid advances and increasing amounts of information in the public domain raise the dual-use implications significantly.

12. A new field, known as *synthetic biology*, incorporates elements from such diverse disciplines as engineering, biology and chemistry, and draws upon technologies in DNA synthesis (see above), bioinformatics, and reverse genetic engineering (see below). In synthetic biology, scientists seek to fabricate useful microbes and learn about underlying principles of cellular function by reducing biological systems to their simplest components and by creating models of genetic circuits. The ultimate objective is to have the *complete ability to design and produce* biological systems from component elements without a supporting organism system (i.e., *in vitro*, in a test tube). Challenges include synthesizing larger genomes while maintaining a low error rate, developing appropriate expression systems for synthetic DNA, and managing mutations of synthetic DNA. Synthetic biology will be facilitated by an improved understanding of gene function and gene-to-gene interactions. Synthetic biology may allow researchers to

develop a registry of biological parts and essentially create tiny programmable computers from living organisms. Synthetic biology has already been used to reengineer a bacterial protein to bind to TNT and then activate a gene circuit to emit a luminescent signal in the presence of TNT. Thus, bacteria could be engineered to detect chemical or biological agent signatures, clean up environmental pollutants, and possibly even "fix" faulty genes. The tools and knowledge embedded in synthetic biology (i.e. DNA synthesis) could also be used to recreate virulent pathogens from scratch (*de novo*), without access to the original organism, or to synthesize the toxin sequence of a pathogen and insert it into a benign host. Long-term, synthetic biology could allow creation of a new or existing pathogen that could be used for malevolent purposes, without accessing controlled substances or technologies. Concerns about potential harm to human health or the environment should engineered organisms survive outside the laboratory have prompted studies of the implications of much research which is now underway.

13. *Genetic engineering* of viruses is closely related to synthetic biology. Both DNA viruses and RNA viruses can be manipulated and engineered by inserting specific genes or nucleotide sequences into a "host" genome. For example, using a process known as "reverse genetic engineering," researchers can introduce viral RNA into bacterial cells, where the RNA can then be manipulated much more easily. Scientists have constructed clones for such viruses as poliovirus, yellow fever virus, H1N1 influenza A, rabies, and others. Recent advances have even made it possible to reverse engineer coronaviruses (causative agent of SARS). While the ability to reverse engineer dangerous viruses such as H1N1 influenza A (the cause of the devastating 1918-19 Spanish influenza pandemic) and H5N1 (avian flu) obviously have enormous beneficial potential for development of vaccines and therapeutics, the same techniques could be used to manipulate the viruses maliciously. In particular, avian flu potentially could be genetically modified to render it transmissible to humans and to induce resistance to currently available antivirals and vaccines.

Understanding and Manipulating Biological Systems

14. A broader, more holistic understanding of complex biological systems is emerging through techniques that allow collection and analysis of comprehensive data on a large variety of biological processes. "Systems biology" and genomic medicine are but two examples of such techniques. Examples of the tools that could be used to manipulate complex biological systems include gene silencing, novel binding reagents (e.g., nucleic acid and peptide aptamers, engineered antibodies), and small-molecule modulators of physiological systems. In many ways this category of technologies opens up entirely novel aspects of the future biodefense and biothreat agent landscapes and changes the fundamental paradigm for future discussions on the topic.

15. *Systems biology*, or integrative biology, is the simultaneous study of complex interactions involving networks of molecules, including DNA, RNA, and proteins using high-throughput tools. It is, in a sense, classical physiology taken to a new level of complexity and detail. Systems biology involves the application of systems- and signal-oriented approaches to understanding inter- and intracellular dynamic processes. Its purpose is to understand inter- and intra-cellular dynamic processes through systems- and signal-oriented approaches. Systems biology tools are useful for analyzing cellular regulatory networks and pathways, as well as genomic and proteomic data sets. Advances in these tools will improve the predictive accuracy

of models of biological systems and will allow physicians to better identify preventative and therapeutic measures based on an individual's genotype and phenotype. The same advances could also make it easier to identify ways to maliciously manipulate biological systems.

16. *Genomic medicine* refers to potential, patient-tailored treatment of diseases based on analysis of the person's genetic makeup. Scientists have known for a long time that human genetic variation is associated with many diseases and questions. With recent advances in technology that allow for quick, affordable genotypic assessments (i.e., from polymerase chain reaction, PCR, to high throughput sequencing) researchers have begun to understand the implications of human genetic variation for the treatment of disease. The applications of these technologies will enable physicians to identify genetic variations associated with disease in order to develop more effective and specific treatments. The same techniques might also be used to identify populations most vulnerable to certain diseases. Such knowledge might – in turn -- be used to target those populations (e.g., ethnic group, gender, etc.) in an attack. It should be noted, however, that existing techniques already could be exploited for such purposes. Researchers have frequently used adenoviruses to target tumor cells in individuals and are refining techniques for introducing the viruses into the cells by modifying the mechanisms that the viruses use to attach to particular cell types. These adenoviruses could be used to target distinct ethnic or gender-specific groups without even identifying population-specific markers to cause disease, debilitating illness, or reduce fecundity.

17. Scientists are also mapping the molecular signatures of the body's bioregulatory circuits and of the disease-induced disturbances in these regulatory pathways. Critical components can then serve as targets for therapeutic and preventative intervention or manipulation; they can also serve as targets for malevolent manipulation and as the basis for novel kinds of biological attack. Technologies that facilitate a better understanding of intracellular, organ, and whole-animal control "circuitry" will enhance the ability of scientists to manipulate these complex systems. A tool known as *RNA interference (RNAi)* is a common antiviral defense mechanism in plants and is also commonly observed in animals' regulation of gene expression by micro-RNA molecules (miRNA), which interact with cellular RNAs to suppress certain proteins and silence the expression of target genes. Work is underway to develop this tool to target disease-causing genes and proteins that are inaccessible to conventional drugs. But just as RNAi promises new therapeutic options for treating cancer and other diseases, it could also be used to manipulate gene expression with the intent to do harm.

18. Aptamers (short, single-stranded nucleic acids) and tadpoles (protein-DNA chimeras) are *high affinity binding reagents* that bind to and inhibit or modulate the expression of specific protein targets. Aptamers have already been explored in animal models for their potential to be used in the inhibition of blood clot formation and as a treatment for age-related degenerative changes in the eye. Tadpoles and aptamers are potentially useful for disease diagnosis and environmental detection, including identification of pathogens or biological agents in the event of a naturally-occurring or deliberate biological attack.

19. The rapid advances in the life sciences are increasingly dependent on, and enabled by, *computational biology and bioinformatics*. The term *bioinformatics* is used to describe the application of large-scale data analysis techniques to the life sciences, encompassing such areas as biology and medicine, computer science, statistics, mathematics and physics. Bioinformatics has become an essential component of modern biology in academic, government and industry

research sectors due to several advances: 1) the availability of large volumes of genetic information from high through-put (HTP) technologies; 2) advanced computing platforms for analytics and data management; 3) high market value of potential products created by this type of R&D; 4) availability of large amounts of venture capital; and 5) the potential for shortening the time from discovery to market. Bioinformatics is creating new scientific and commercial opportunities, as well as working with the high throughput technologies to make it easier to create novel structures and substances from biomaterials.

20. Core information technologies that have a fundamental impact on bioinformatics include: 1) the dramatic increase in capacity of commodity hardware allowing storage and analysis of large amounts of bio-information on low-cost platforms; 2) significant improvement in efficiency of data sharing and communication technology making it possible to share and operate on very large, complex data sources which are dispersed over a wide geographical area but which frequently do not emphasize security; 3) a suite of modular technologies such as web-based applications that has enabled worldwide access to tools for analyzing easily accessible biological and genetic information; 4) common data formats that allow integration of data streams from multiple sources; 5) emerging novel methods for searching through massive complex and geographically dispersed information stores allowing almost anyone to rapidly find core information which could potentially be used in design of biomaterials, engineered organisms or strategies for detection or intervention against biological agents; and, 6) multi-dimensional visual analysis technologies. Such technologies enable a bio-designer or analyst to quickly browse huge amounts of data for evaluation and product design or engineering using not only sequence information, but also all the derived and associated data about the sequences and structures, functions, and properties of end products.

21. The first and most striking change in the last 5 years has been the amount of genetic information on a wide range of organisms that is available worldwide. For instance, the Genbank public repository containing records of organism genetic content doubles in size every 18 months.

Production, Delivery, and "Packaging"

22. Technologies that allow biological systems to be manipulated in a defined, deliberate manner are evolving very quickly to serve the pharmaceutical, agricultural and health care fields. Some of these technologies have not been traditionally viewed as having relevance to future biological threats. A prime example is use of nanoparticle science for creation of novel and highly efficient delivery systems for difficult-to-deliver biologically active compounds.

23. *"Biopharming"* is the use of plants to produce bioactive molecules intended for industrial products and pharmaceuticals. Biopharming could potentially enable the production of vaccines and antibodies that would be too expensive or inefficient to produce using conventional methods. Scientists are also experimenting with plant-derived microbiocides and are exploring plants as a cost-effective way to produce antibodies for use against potential biowarfare agents. These same methods could, however, be used malevolently to produce large quantities of bioregulatory or other toxic proteins to be used as biological agents. Using plants as production platforms would also help eliminate the risk of discovery, as genetically modified plants are not easily distinguishable from non-transgenic crops. However, there is an obvious hazard with such

transgenic crops becoming mixed with non-transgenic crops, as has occurred with non-malevolent transgenic materials, or if they are eaten by wild animals.

24. *Microfluidics and microfabrication* cover a wide variety of processes and manipulations carried out at miniaturized scales, usually automated. Microfluidic, or "lab-on-a-chip" technology, underlies many recent advances in point-of-care diagnostics, including DNA analysis, immunoassays, cell analysis, and enzyme-activity measurements. Sophisticated, miniaturized diagnostic systems can greatly enhance the ability to identify and respond to disease outbreaks, whether naturally-occurring or deliberately caused.

25. *Nanotechnology* is the science of making very small things, to the order of one-billionth of a meter. Advances in the ability to produce and characterize materials with smaller and smaller dimensions have resulted in materials with novel and useful properties – to nanomaterials and nanotechnology. Application of nanotechnology to biological systems currently focuses on using custom nanoscale materials for *in vivo* applications, such as molecular imaging and detection, reporters for therapy efficacy determination, multifunctional therapeutics, disease prevention and control, and enabling technology such as nanofluidics and massively parallel systems for rapid discovery.
As with other previously discussed technologies, nanotechnology is potentially dual-use and could be used to deliver toxic agents maliciously.

26. *Aerosol technology* is a rapidly developing field to deliver inhaled therapeutic particles as a means to treat human disease. Propellant metered-dose inhalers (pMDI), dry powder inhalers (DPI), and nebulisers are examples of aerosol technology commonly used to deliver drugs directly to the lungs and circulatory system. The efficacy of an aerosol technology depends greatly on the particle size of the substance to be aerosolized and its ability to be delivered directly into the bloodstream. Rapid progress in this area is being made through supercritical fluid processing. New advances in microencapsulation technology, discussed below, raise concerns about the delivery of bioregulators and other toxic chemicals through nefarious use of aerosol technology.

27. *Microencapsulation* is the envelopment of small solid particles, liquid droplets, or gas bubbles with a protective coating comprised of any of a number of compounds (organic polymer, hydrocolloid, sugar, wax, fat, metal or inorganic oxide.) The capsules protect their contents from evaporation, oxidation, and contamination, thus maintaining stability in a wide variety of environments. Such capsules have countless applications in commerce and industry, including for controlled and delayed drug delivery. Researchers are investigating ways in which the technology can be used to deliver microencapsulated, engineered, live cells for therapeutic purposes. Alternatively, nefarious use of microencapsulation technology could deliver pathogens or toxins that would be inactivated by normal environmental conditions (e. g., UV light, oxidation) in sufficiently lethal concentrations that would be unachievable using other delivery mechanisms.

28. Considerable attention is also being devoted to *targeting biologically active materials to specific locations within the body.* The delivery method of medical drugs, vaccines, and other biologically-active materials is important, but most important is the ability to deliver precisely and accurately to the cells of interest. Researchers are exploring using carrier particles containing encapsulated drugs, exploiting the "homing" ability of microorganisms to bind selectively to

specific cells (e.g., viruses or bacteria as vectors for targeted delivery of genes and proteins), and coupling drugs to carrier molecules that recognize only the desired cell type before releasing their cargo. One delivery method of particular dual-use concern is the use of viral delivery vectors to insert genes into chromosomal DNA. Research has shown that it is feasible to insert "foreign" genes into multiple types of cells, although, as discussed above, gene therapy using this method is not yet clinically useful.

29. *Gene therapy* is a technique that uses "healthy" genes to treat or prevent disease by inserting a "normal" gene into the genome to replace an abnormal, disease-causing gene. The technique typically uses a carrier molecule or vector, such as a "harmless" virus, to deliver the "healthy" gene to the target cell. Gene therapy is still experimental and has yet to be successfully applied in human clinical trials. However, when it becomes refined, the technology could theoretically be used to introduce harmful genes into cells and tissues with unknown consequences on individuals and populations.

The Complementarity and Synergy of Technologies

30. Biotechnology, nanotechnology, and information technology are converging in ways that will enable life processes to be manipulated with far-reaching implications and great potential for nefarious and disastrous outcomes. The various tools discussed above are interacting and converging – both additively and synergistically – and creating unanticipated opportunities for these technologies to be used for good and for ill. While naturally-occurring threat agents, such as anthrax, and "conventionally" genetically engineered pathogenic organisms are the near-term threats we must confront today; the emerging threat spectrum will become much wider and will include biologically active agents such as bioregulators. On an optimistic note, the rapid advances discussed in this paper will empower improvement in human health and well-being and development of better defenses against the current and new threats.

_____