Perspectives for the new peptide millennium

Bruce Merrifield,[1] George Barany,[2] Charles M. Deber,[3] Murray Goodman,[4] Robert S. Hodges,[5] Victor J. Hruby,[6] Tom W. Muir,[1] Robin Offord,[7] Arno F. Spatola,[8] Daniel F. Veber,[9] and Gregg B. Fields[10]

[1]The Rockefeller University, New York, NY 10021, U.S.A.; [2]University of Minnesota, Minneapolis, MN 55455, U.S.A.; [3]University of Toronto, Toronto M5G 1X8, Canada; [4]University of California at San Diego, La Jolla, CA 92093, U.S.A.; [5]University of Alberta, Edmonton T6G 2S2, Canada; [6]University of Arizona, Tucson, AZ 85721, U.S.A.; [7]Université de Genève, CH-1211 Genève 4, Switzerland; [8]University of Louisville, Louisville, KY 40292, U.S.A.; [9]SmithKline Beecham Pharmaceuticals, King of Prussia, PA 19406, U.S.A.; and [10]Florida Atlantic University, Boca Raton, FL 33431, U.S.A.

George Barany and **Gregg Fields** (Overview): When we were discussing an appropriate way to bring this Symposium to an exciting scientific conclusion, we recognized that the calendar had given us a once in a millennium opportunity. We asked Professor Bruce Merrifield of The Rockefeller University, the 1984 Nobel laureate in Chemistry, to convene a high powered panel of top peptide scientists to summarize the Symposium and provide a vision of where the field is headed in the 21st century. Here follows a minimally edited reconstruction of what was said.

Bruce Merrifield (Introduction): This brings us to the end of our Symposium. I think it has been a very exciting, important meeting and it is clear that the peptide field is alive and well. This final session is entitled "Perspectives for the New Millennium." Our purpose is to examine what has been said this week, to draw it together, and, based on this evidence, to try to extend these predictions farther into the next millennium.

Peptide Science can be divided into many sub-categories, but for the purpose of this discussion we have selected three broad areas (Fig. 1). These three – Chemistry, Physics, Biology – are clearly not separate, sharply divided disciplines, but overlap in important ways: chemistry flows into physics and biology; biology into physics, etc. They are dependent on each other and are complementary. Each panelist will focus on one segment, particularly those that are pointing to future directions of research.

Arno F. Spatola (Chemistry): This is an exciting time to be engaged in organic peptide synthesis. There were many presentations and posters at this meeting that illustrate the incredible diversity of reactions, products, and matrices used in modern peptide science. These products are highly sought after for drug lead discovery and in proteomics. In fact, a colleague interested in discovering new proteins and modified proteins suggested to me that "We've got the targets and you have the bullets!" While I'm not sure that in this era of gun control that I can endorse his analogy, it is clear that those of us engaged in synthesis will be very busy at the start of this new millennium. One of the clearest themes at this meeting was the broad use of organic reactions to create new modified amino acids, peptidomimetics, and protein derivatives. Many of these are attempts to expand structural diversity with glycopeptides, lipopeptides, and even nucleopeptides. However, there is also increasing emphasis on replicating phosphorylated peptides and their analogs, as well as farnesylated derivatives or peptides with various branched carbohydrates, in an effort to duplicate the wide range of post-translational modifications being catalogued in humans and other organisms.

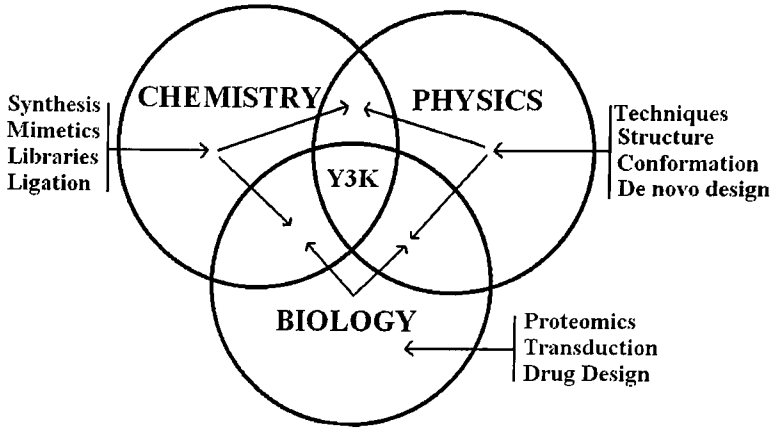

Fig. 1. Venn diagram of the broad areas encompassing peptide science.

Another trend evident at this pivotal conference was the emphasis on improved methods of synthesis. These ranged from increased use of solid-phase reactions, new and improved coupling reagents (including perhaps the rediscovery of Fischer's amino acid chlorides), and further investigation of novel multifaceted solid supports that could prove equally useful for synthesis and analysis. To this admittedly biased observer, there seemed to be even more interest and examples of the use of cyclization for constrained peptide analog synthesis.

Finally, we also learned of the need to consider the synthesis of structures that could survive the many different types of traps that hinder bioavailability as peptide analogs meander to their ultimate sites of action. The aforementioned glycopeptides can help, as can various PEG-ylated derivatives. Synthetic chemists need to consider not only new strategies of pro-drugs, but must also be well informed about alternate delivery methods and the special requirements that accompany these modes.

Murray Goodman (Chemistry): Our field of chemistry has been and will continue to be centered on M^3: Molecules, Medicinals, and Materials. Organic chemists involved in peptide research have focused their efforts on bond making and bond breaking in syntheses and structure determinations. It must be stressed that the synthesis of peptides involves much more than amide bond formation. The vast majority of molecules that have been made contain multiple stereogenic sites. Therefore, synthetic strategies (protections, deprotections, and activations) which avoid epimerization have been a maxim for peptide chemists. Synthetic efficiency and stereopurity of target peptides and peptidomimetics are critically important issues and will remain major concerns for the foreseeable future.

We now enter an era of molecular diversity which includes combinatorial syntheses of libraries, *de novo* design of protein mimetics, and the synthesis of dendrimers and other macrostructures. To accomplish the syntheses of these molecular systems, peptide chemists will be required to design new reactions and novel building blocks. As part of molecular diversity, researchers in our field must devise scaffolds and templates on which to array peptide and peptidomimetic pharmacophores, sensors, catalysts, and complexing

agents. Other structures will be based on glyco-, nucleo-, and lipopeptides. These peptide conjugates will be the basis for the design of structures with novel properties. In addition, new structures will be created for specific medicinal targets including antimicrobial, antiviral, and anticancer agents.

Peptide chemists have established collaborations with molecular biologists, biophysicists, and material scientists. These collaborations will be expanded in the future. Thus, the molecular aspects of peptide chemistry remain exciting and peptide researchers will continue to be in the center of molecular discoveries. It is difficult to be a prognosticator, but I am certain that peptides are here to stay and will form the basis of major new applications of M^3: Molecules, Medicinals, and Materials.

Daniel Veber (Chemistry): A clear trend at this meeting is a reconvergence of peptide chemistry with organic chemistry. There is a bridging of a gap that had developed over the course of the 20th century from the time of Emil Fischer – who, of course, perceived no such gap. An annealing factor has been combinatorial chemistry. Traditional organic chemists are learning to use solid-phase chemistry and peptide chemists are broadening the scope of their reaction base well beyond the formation of amide bonds. The diversity of chemical properties that can be achieved by combinatorial chemistry will continue to enrich our design of bioactive molecules.

De novo design has and will continue to improve in a qualitative sense, but precision of design will remain an elusive goal. This will be a consequence of something like an uncertainty principle, never allowing us to precisely define the complex molecular properties of the components of the living system that is constantly changing as we observe it. The subtlety of molecular interactions with a protein and the changes in interactions on even a single mutation are outside the resolution of our physical methods. Fortunately, we are now learning how to handle these design issues by using combinatorial chemistry. Solid, well-conceived design concepts tend to fail when only a single or a few compounds are made. This is a simple consequence of probability. Combinatorial chemistry gives us hundreds or thousands of chances to succeed with a good new scaffold or mechanism-based idea. Dan Rich referred to the convergence of design and combinatorial in his excellent Merrifield Award lecture [see page 1], and I concur that this is an inevitable outcome.

The microbes that attack us – viruses, bacteria and parasites – have long understood the power of combinatorics. They have used it to move ahead of 20th century medicines. They threaten our very survival in the new millennium. Knowledge of genome sequences will be used to show the way to new drug targets that are unique to infective microbes. Genome sequences will also reveal the structures and allow us to prepare quantities of the proteins that limit our ability to direct new drugs to the places in the body where they can act on infective agents. The transporters and metabolizing enzymes that limit duration and oral availability of drugs are now being identified, cloned and expressed for *in vitro* studies. Orthologs of these proteins from the species that serve as animal models are also becoming available for *in vitro* studies. Proper understanding of these proteins that influence drug action will have enormous impact on the drugs that will become available in the new millennium. The outcome should be more rapid drug discovery, safer new drugs, and greater assurance of success for the molecules that enter clinical studies in humans. The challenges of the new knowledge covered at this meeting highlight the dynamic nature of our field. The challenges are especially directed to the younger scientists whose insights will make advances that I can hardly project today.

Charles Deber (Physics): Structure is the bridge between chemistry and biology. Because the central 'mantra' of our field has been rational drug design, the need to deduce structure in turn relates to the need for new knowledge of the drug targets, *viz.*, the proteins. The limitations to this have always been technical, but two themes at the Minneapolis meeting have emerged in confluence. First, the line is blurring – becoming elastic – between peptides and proteins. Peptides are getting larger, proteins are getting 'smaller'. This is because modern peptide chemistry – and I would suggest that chemistry be considered in conjunction with molecular biology/mutagenesis techniques – means that the models used to ferret out physical principles of structure can now be much more complex than ever before. Yet at the same time, research reported at this meeting indicates that the array of biophysical techniques for structure deduction, and their capabilities, have been vastly improving and expanding. From talks and posters, it was apparent that the established techniques, including CD, NMR, X-ray crystallography, fluorescence, and MS, along with computational chemistry and several developing techniques, are being put to novel and important uses.

Several examples from the meeting illustrate this situation. Fluorescent-labeled lipopeptides (palmitoyl/farnesyl) were used to study insertion and selective targeting to membranes, and surface plasmon resonance was employed to gain additional insights into the peptide/lipid system. Fluorescent probes capable of detecting tumor-associated protease activity *in vivo* were described. Tandem mass spectrometry was used, in conjunction with computer searches, to analyze peptide and protein expression profiles (an application of 'proteomics'). MALDI-TOF mass spectrometry was used in a small molecule library to identify individual components within a mixture via their molecular weight differences versus an invariant core. Electrospray mass spectrometry was employed to determine the rate and extent of H/D exchange in purine nucleoside phosphorylase systems. Segmental isotopic labeling of proteins for TROSY NMR structural studies on tyrosine kinase receptor pathways was carried out in conjunction with ligation of domains of folded recombinant proteins. Isotopic (^{15}N) labeling of peptides was also used for NMR monitoring of the folding kinetics of collagen triple helices. Transfer NOE NMR experiments were used to obtain the conformations of protease-bound inhibitors. Magic angle spinning NMR was employed to help optimize reaction profiles of resin-bound peptides. CD spectroscopy was used to measure β-sheet stability; to examine β-promoting cassette segments within helical coiled-coils; to measure the extent of peptide insertion into membranes; and to study thermal denaturation of collagen-mimetic triple helices. We saw a novel use of CD spectroscopy for distinguishing 3_{10} helices from α-helices via the asymmetric appearance of 222/208 nm bands. Atomic force microscopy was used to study head-to-tail self-assembly of synthetic peptides into monolayers on graphite surfaces. Molecular dynamics simulations in a water/decane/water cell were employed to mimic a membrane environment for studies of parathyroid hormone and its G-coupled receptor.

From just this tip-of-the-iceberg sampling, we see that the future of biophysical analysis of peptides and their protein targets is bright indeed. As our field moves toward the 21st century, it is also clear that some disciplinary boundaries which may have formed in the '80s and early '90s are breaking down. Now, specialties are coming together again, such that peptide approaches to structural biology should be expected to have an ever-increasing impact, and become indispensable to our basic understanding of peptide/protein structure and function.

Robert Hodges (Physics): I would like to focus on areas where I believe peptide chemists can have a major impact in the future. First, understanding protein folding and protein stability is critical in the prediction of protein structure. It is obvious that even with the massive expansion in structural biology (NMR and X-ray crystallography) which is taking place around the world, we will not keep pace with the hundreds of thousands of new protein sequences available from the human genome project. Thus, protein structure prediction remains the key problem to be solved in the biological sciences. The question becomes: how can peptide chemists stay at the leading edge? The answer is to actively expand our involvement in the research discussed at this meeting, where we heard about design, folding and stability of monomeric α-helices, two-stranded α-helical coiled-coils, four helix bundles and β-sheet proteins. Understanding how small regions of sequence can switch conformation from α-helix to β-sheet is vital in order to tackle a number of fatal neurodegenerative diseases such as Alzheimer's disease, Creutzfeldt-Jakob disease, and bovine spongiform encephalopathy. Second, *de novo* design of small proteins with catalytic activities is still in its infancy, and we need a massive influx of scientists in this area for the future. Third, studying protein-protein interactions, in particular in multi-protein complexes, is an area that can benefit from synthetic peptide approaches where a vast number of sites of interactions between proteins involves small regions of sequences. This is an area where X-ray and NMR techniques are having extreme difficulties (either in crystallization, or the protein systems are too large for NMR analysis). Preparing synthetic fragments to pinpoint protein-protein interactions is a fundamental requirement for success in protein structure and function studies, and will simplify X-ray and NMR determination of the smaller peptide-protein complexes. Fourth, we all know that there is a growing problem with antibiotic resistance. This opens up the field of antimicrobial peptides to the peptide chemist where such peptides can avoid the resistance problem and exert their effect in the lipid bilayer. Interestingly, antimicrobial peptides exist in cyclic or linear forms, containing either different secondary structures (α-helix, β-sheet) or negligible secondary structure.

Robin Offord (Biology): A working draft of the human genome will be ready much sooner than anyone originally expected, probably in the spring of next year (2000). We will then have sequences corresponding to the many tens of thousands of human proteins that we expect to find there. We won't know just from looking at them what the majority of them do. This will clearly be a staggering opportunity and challenge, but it is only the starting point. Increasingly, the study of post-translational modification has led us to the realization of how widespread such mechanisms are, and how inadequate a mere knowledge of the structure of, say, messenger RNA, is for our understanding. We must not forget that *in vivo* fragmentation is increasingly recognized as an immensely rich source of additional diversity. We each of us have a vast range of protein fragments in our natural constitution, and many of them are much more than just junk. Even a conservative estimate of the likely number of post-translational modifications and biologically significant fragments (and I don't personally feel conservative) shows us that we won't any longer have to think of tens of thousands of significant target structures, but of more than a million. We are concerned not with the genome, but with the proteome.

All of these possibilities come up just at the moment when we begin to be equipped to deal with them. Do not let the size and number of these potential targets daunt you. I submit to you as a take-home message that, whatever the size of the protein concerned, the largest things that we will normally need to synthesize will be the functional domains, typically around 200 residues each. To put these together, if we need to, we have all the ligation techniques developed over a couple of decades for semisythesis, complemented more recently by the natural ligation methods that we have heard about so often at this

meeting, and which have transformed our ideas of the possible in terms of total synthesis. I repeat: one or two labs could *already*, with some effort, make domain-sized proteins, and everything suggests that methods will become simpler and more accessible. We *already* know how to stick domains together: the largest controlled-structure semisynthetic construction that I know of has a molecular mass of 180 kDa.

My second, personal take-home message is that the huge advantage of chemical synthesis is the total control that we have over the structure of the products. We have heard during the meeting how we can now, covalently, and at chosen sites only, place the exact lipid structure that we need, the exact carbohydrate, the exact PEG-like structure, or polyamide. We can place the exact complex cofactor-like structure where we want it, we can introduce regions of molecular diversity, or incorporate at will any one of the thousands of non-coded amino acids now available to us.

The analytical and bioinformatics techniques are evolving at the same rate as our biological insights as to what is important. I would say to any younger scientists present (or even older ones!) who are wondering whether to stay in, or enter this field that, if this sort of thing interests you at all, stay with it. We in this room, with the ability to have total control over structure which is the hallmark of what we do, are uniquely placed to exploit to the full the fantastic situation which is developing around us.

Tom Muir (Biology): I would like to comment on the important role that I believe synthetic peptide and protein chemistry will play in the post-genomic (i.e., proteomic) era. It should again be stressed that the so-called "proteome project" is a hugely daunting undertaking since it involves the chemical and biological characterization of perhaps hundreds of thousands of polypeptides (*cf.* previous comments), and is further complicated by the emerging picture of complexity in biological processes. Clearly, both established and novel technologies will have to be brought to bear on this problem. I believe that chemistry, and in particular organic chemistry, will have an critical role to play in this endeavor as it evolves, both through the synthesis of small molecule probes of biological processes and through the direct chemical manipulation of peptide and protein structures – in other words, *peptide chemistry.*

As many of you may know, the last several years have seen the emergence of the so-called peptide ligation approach to protein chemical synthesis, that is to say the chemo- and regioselective assembly of large protein targets from constituent unprotected peptide building blocks. There have, over the years, been several key contributions to this area, which can be traced back to the pioneering work of Wieland and Brenner in the 1950's and 1960's, and we have been fortunate enough to hear exciting new ideas from many of the world leaders in this field during this meeting. The highlights have included:

- the development of solid-phase peptide ligation strategies;
- the application of chemical ligation principles to the synthesis of neoglycopeptides/proteins, and peptide/protein conjugates;
- the development of approaches for the synthesis of thioester peptides using the Fmoc SPPS strategy;
- the development of novel auxiliary approaches for use in chemical ligation strategies;
- semisynthetic ligation strategies which allow synthetic peptides and recombinant peptides to be freely intermixed in chemical ligation approaches.

The field of peptide ligation has so blossomed in recent years that I would submit that the routine application of organic chemistry to the synthesis of large proteins is a reality. Chemistries are now in place which allow the practical synthesis or semi-synthesis

of proteins of largely unlimited size, and possessing ever more complex patterns of chemical modification.

All this having been said, what are the opportunities for the peptide/protein chemist in the next millennium? While there are still several outstanding technical problems in the peptide ligation field and important refinements of the strategies will undoubtedly continue, I believe that in the long term the field must be fueled by the manifold challenges posed by the proteome. The opportunities are staggering and too numerous to list, or realistically to even imagine at the present time. Consider, as an example, the post-translational modification of peptide and protein structure. We already know (or at least suspect) that post-translational modification (e.g., phosphorylation, prenylation, glycosylation) is Nature's way of conferring functional diversity onto the same translated sequence, and is crucial to the way proteins are regulated, localized, and stabilized *in vivo*. Despite this, the generation of proteins possessing precise and homogeneous patterns of post-translational modification has been extremely problematic, and in many cases impossible, using standard biotechnology approaches. In contrast, peptide chemistry offers the ability to precisely introduce such modifications into synthetic peptides, and thus, with the aid of ligation strategies, into larger proteins. This will allow the biochemical and structural consequences of these modifications to be studied in detail, in most cases for the first time. I suspect that this one slice of the proteomic pie could sustain the entire field for a great many years!

In conclusion, I believe that this is a great time to be a chemist interested in how peptides and proteins work. The next millennium is paved with opportunities.

Victor J. Hruby (Biology): To comment on peptide and peptidomimetic drug design in the new millennium following the excellent discussions which preceded me is a daunting task indeed. What the previous speakers have pointed out is extremely exciting for our field and those interested in ligand design and in drug design. These are not necessarily the same thing, but both will be essential for our ability to understand the chemical and physical basis for living systems, and for the diagnosis and treatment of disease. Essentially, the human genome and its use and applications is up for grabs, and those of us interested in ligand/drug design have enormous opportunities to make seminal contributions. The wave of the future will be collaboration, so that the structural, chemical, biological, and behavioral effects of our designed ligands and drugs can be more rapidly designed and evaluated.

In the case of the design of bioactive peptides and peptidomimetics, the intersection of chemistry, physics, and biology is obvious to anyone who heard the many talks and saw the many posters at this Symposium which examined ligand and drug design and evaluation. What is not always obvious is whether the usual hierarchy of science (mathematics → physics → chemistry → biology → behavior) is extant. The tools in all of these fields are under rapid development for ligand/drug design, and molecular biology (broadly defined) is at the intersection of all of these areas. Many aspects of the problem have been well covered by this panel, in terms of the development of new ligands with unique chemical, physical, and biological properties. I would like to emphasize three areas which pose significant problems, but for which outstanding progress will be made in the next millennium. The first is understanding non-covalent bond interactions in biological systems. While covalent bonds are clearly essential for construction of biological compounds and building blocks (and thus synthetic chemistry will continue to make great strides), it can be argued that biology depends largely on non-covalent bond interactions for its manifestations of life. Of special importance are the interactions of membranes/proteins/ligands, and our increased ability to understand the properties of these complex systems in terms of structure-activity relationships will be critical. A second area is information transduction, which was discussed in many talks and posters at this

Symposium. The understanding of the transduction pathways and their interactions will require design and synthesis of specific transduction-controlling agonists and antagonists. Finally, the third area, delivery of peptides and peptidomimetics. To mimic bodily functions in terms of ligand distribution for the cure and treatment of disease will require enormous progress in the design of chemical delivery systems. We heard a number of excellent presentations at this Symposium in this direction, and I believe very significant progress will continue because the need is great. In all of these areas, the differences between humans and our current animal and cellular models has become apparent, and being able to understand and utilize these differences as part of our design will become increasingly important.

The human genome project, and its implications for peptide and protein sciences, has been emphasized in several talks at this Symposium. The knowledge of the entire human genome, and our ability to distinguish genetic differences which are related to disease, pose enormous opportunities, but also enormous philosophical, social, and ethical questions. For the long-term future, I would like especially to point to an area which has enormous philosophical, cultural, and human implications, and which we in peptide and protein chemistry and biology will become intimately involved in, namely that of cognition and behavior. Though not much was said at this meeting about this area, we are in a unique position to begin to make significant contributions to understanding the age-old problems related to behavior, from feeding behavior to sexual behavior, from addiction to depression, to anxiety, to joy, and many more. All have chemical and biochemical correlates related to peptide ligands and protein receptors, ion channels, enzymes, and regulatory components. Already ligands have been discovered in which small changes in structure can significantly affect behavior. The applications of this to Society, and the ethical issues which are raised, require our most serious and thoughtful examination, and one that we as scientists must take responsibility for.

I cannot imagine a more exciting time to be in peptide science. The challenges and opportunities are enormous, and will require a change in our behavior as scientists to maximize our creativity by cooperation and collaboration. We already have seen several examples of this new paradigm, especially from our industrial colleagues. We should move forward with tremendous enthusiasm and confidence in our field and the central role we can play in the science of the new millennium.

Bruce Merrifield (Conclusion): I would like to add one long-term prediction of my own. When all of these disciplines in the Figure intersect perfectly, perhaps in Y3K or maybe much sooner, I think it will be possible to produce a totally synthetic system that will self-replicate. Manfred Eigen has provided a theoretical background involving a hypercycle that allows the system to evolve at each turn of the cycle. New experimental capabilities may eventually combine with theory to achieve this goal.

What I get from this meeting and this session is that peptide science is growing and advancing rapidly. There is much to be done and there are not enough people to do it all. I think there are exciting times ahead.