Essays

Israel Journal of Chemistry

DOI: 10.1002/ijch.201800171

A Personal Perspective on Chemical Biology: Before the Beginning

Peter B. Dervan*[a]

Dedicated to Claude Hélène

Abstract: This perspective represents a brief personal account of early days before "chemical biology" emerged as a field of inquiry. Imagine a time when oligomers of DNA could not be synthesized and the order of the TACG letters in DNA could not be sequenced. Even the high resolution structure of the DNA double helix was not yet determined. 1975 was a time when there was a deep chasm between chemistry and biology. Chemists with precise knowledge of all the atoms in natural product architectures looked with dismay at the imprecise messy world of biology. Water was to be avoided! My view was that the power of synthetic organic chemistry should be used to create function, synthesis with a purpose. Our organic group at Caltech would embrace molecular recognition of biologics in water as a frontier for chemistry. We dreamed of inventing small molecules that would control the activity of macromolecules such as DNA, proteins and carbohydrates in living cells. We chemists would sky dive into the messy world of biology.

Keywords: chemical biology · history · molecular recognition · water

1. Introduction: Exploring the Interface of Chemistry and Biology at Caltech (1975)

My early training in chemistry as a graduate student in Jerry Berson's group occurred at two institutions, the University of Wisconsin, Madison (1967–69) and Yale (1969–72). Those were exciting days in physical organic chemistry with the publication of the Woodward-Hoffman orbital symmetry rules for predicting stereochemical outcomes for certain thermal hydrocarbon rearrangements. The core of our inquiry was physical, the properties (structure, lifetimes) of transient intermediates in the gas phase, yet I spent most of my years as a graduate student doing synthetic chemistry. Clearly, in physical organic chemistry, synthesis was important for the study of reaction mechanisms underpinning stereochemical relationships between reactants and products and as well as kinetics.

In early 1973 I moved to Stanford University for postdoctoral research to broaden my horizons in synthetic chemistry. At Stanford, Gene Van Tamelen was interested in an audacious venture to synthesize tRNA. The size of my ignorance for biological macromolecules was vast. Sadly, I had never taken a course in biochemistry or biology in college or graduate school. At Stanford, I attended a few biochemistry lectures by a brilliant young professor Paul Berg and felt a new magical universe unfolding before my eyes. I had an offer to start at Caltech that Fall of 1973 so I headed to Pasadena with a 6-month postdoc under my belt. I was 28 years old and had no idea what I would do. During my first two years (1973–75) we initiated several key experiments to reveal the properties of 1,4-biradicals which settled some arguments between theory and experiment at the time.[1–4] The work was considered a technical *tour de force* and it is fair to say most coworkers in the academic world belived I would follow in the footsteps of the great physical organic masters Roberts, Winstein, Doering and Berson. However, teaching advanced organic chemistry at Caltech made me realize that much of the breakthrough papers in physical organic had been written in the preceding 20 years (1952–72). With forty years of research in front of me, I needed to move in a new direction! Believing the *goal of synthetic chemistry is the discovery or invention of new properties,* we turned our attention to the interface of chemistry and biology. In design of *"molecules with function"* we would emphasize simplicity over complexity. If one designed a molecule for a purpose, a function never achieved before, the chance of failure was high! In order to overcome repetitive failures over time, the synthesis must be efficient and complexity minimized. Not unlike soccer, we may need multiple shots on goal.

There were early role models in organic chemistry that set the stage for this new direction. Ronald Breslow at Columbia was deeply interested in understanding the physical organic principles of catalysis. His biomimetic approach was to draw lessons from Nature's enzymes and build *artificial systems* that were mimics of nature's catalysts. One particular paper in 1973 that inspired me was the concept of *remote functionaliza-*

[a] Prof. P. B. Dervan
Division of Chemistry & Chemical Engineering
California Institute of Technology
1200 E. California Blvd., Pasadena, CA 91125, USA
E-mail: dervan@caltech.edu

tion.[5] Another role model was Nelson Leonard at Illinois who synthesized fluorescent base analogs of nucleosides.[6] Nelson was an early pioneer regarding the value of *"synthesis with a purpose,"* the design and synthesis of materials with useful properties. Organic chemists Jean Marie Lehn (Strasbourg) and Don Cram (UCLA) were demonstrating the principle of synthesis with a purpose with pioneering studies of host-guest chemistry and the intellectual richness of studying the non-covalent bond.[7] In 1975 our research group at Caltech decided to focus on double helical DNA and molecular recognition in water. This was before DNA sequencing[8] or reliable methods for the synthesis of DNA were available.[9] An x-ray crystal structure of the right handed double helix would be published in 1981.[10]

1.1 Reflections on Two Cultures: Chemistry and Biology (1975–87)

Founding chemistry graduate students Mark Mitchell, Tadhg Begley, Michael Becker, Richard Ikeda, Robert Hertzberg, Michael Van Dyke were joining my group at Caltech to work at the interface of chemistry and biology. It may be valuable for the readers of this essay, many of whom were not born in 1975, to grasp the cultural divide between chemistry and biology. An article appeared in *Biochemistry* in 1987 by Arthur Kornberg (Stanford) that is so thoughtfully written and that I urge all interested in the history of science and the founding of chemical biology to read his essay *The Two Cultures: Chemistry and Biology.*[11] This was written 12 years after our launch, yet resonated perfectly with my Ph.D. experience in organic chemistry. In 1987 Kornberg wrote:

"Much of life can be understood in rational terms if expressed in the language of chemistry. It is an international language, a language for all of time, and a language that explains where we came from, what we are, and where the physical world will allow us to go. Chemical language has great esthetic beauty and links the physical sciences to the biological sciences. Unfortunately, the full use of this language to understand life processes is hindered by a gulf that separates chemistry from biology. This gulf is not nearly as wide as the one between the humanities and sciences on which C. P. Snow focused attention. Yet, chemistry and biology are two distinctive cultures and the rift between them is serious, generally unappreciated, and counterproductive."

"It might have been expected that the separation between chemistry and biology in the 19th century would have been bridged by the emergence and growth of biochemistry in this century. In biochemistry, one could fulfill the wish to understand the chemical basis of cellular function in fermentation and photosynthesis, in muscle contraction and digestion, and in vision and heredity. Despite its enormous success in solving these and other problems, biochemistry has nevertheless failed to fill the gulf between chemistry and biology. Instead, as I shall discuss, biochemistry itself is being pulled apart by the separate drifts of the two cultures for which it was assembled."

Kornberg laid the problem at the feet of organic chemists.

"By 1950, organic chemistry had been enriched by more than a century of impressive achievements. Organic chemists had prepared and characterized the sugar and amino acid substrates and products of enzyme reactions. They belonged to a proud and venerable science. Insights into the structure and reactivity of carbon compounds had made possible the synthesis of the extraordinarily complex alkaloids and antibiotics found in nature. The problem was that organic chemists placed arbitrary boundaries on their science. Although, in their pursuit of natural products, they might still eagerly seek the challenge of an Amazonian butterfly pigment, they would not accept nucleic acids, proteins, and enzymes as proper natural products."

"Nucleic acids, proteins, and enzymes had been excluded from the province of legitimate chemistry, and chemists had excluded themselves from the essence of biology."

In his 1987 paper Kornberg wrote the following:

"Increasingly, young professors in chemistry departments are pursuing problems of biologic significance."

I had just published an invited review in *Science* 1986 on our work at Caltech.[12] Could he be including our group at Caltech in this sentence? I don't know.

Kornberg faults genetic chemistry/molecular biology for not bridging the gap.

"Molecular biology appears to have broken into the bank of cellular chemistry, but for lack of chemical tools and training, it is still fumbling to unlock the major vaults."

"We now have the paradox of the two cultures, chemistry and biology, growing farther apart even as they discover more common ground."

Peter B. Dervan was born in Boston, Massachusetts on 28 June 1945. He obtained his B.S. degree at Boston College in 1967, his Ph.D. with J. A. Berson at Yale in 1972. After a postdoctoral at Stanford in 1973 with Van Tamelen he joined the faculty at the California Institute of Technology that same year. He was promoted to Associate Professor in 1979, Full in 1972 and Bren Professor in 1988. He was Chair of the Division of Chemistry and Chemical Engineering 1994–99. His research interests are the interface of chemistry and biology, in particular, understanding the chemical principles for the sequence specific recognition of DNA. Over 60 of Dervan's former graduate students and postdoctoral coworkers hold academic research positions around the world, many of whom are leaders in chemistry, biology and medicine. Dervan is a member of the National Academy of Sciences, the American Philosophical Society, the French Academy of Science and the German National Academy of Sciences.

Israel Journal of Chemistry

2. Interhelical DNA-DNA Crosslinking, A Probe of Packaged Nucleic Acid

An early first project was to ask the question how is a long strand of double strand DNA folded in a virus. What is the 3-D structure of condensed DNA? Could we determine whether DNA packaging in phage was a ball of yarn, coaxial spool or chain-folded structure? Cross-linking reagents are useful tools for probing the higher order structure of macromolecules. We set the criteria that useful DNA-DNA interhelical crosslinkers should be bifunctional, nucleic acid specific, water soluble, chemically inert, and photoactivated. In 1976 graduate student Mark Mitchell synthesized an azido-methidium dimer with a polyether tether (Figure 1). As a test case, he found that upon irradiation BAMO crosslinks packaged nucleic acid (48,000 bp) in bacteriophage λ.[13] The multi-crossslinked DNA could be digested with a restriction enzyme to give an ensemble of crossed fragments. We could examine unique covalently crossed DNA restriction fragments by electron microscopy.[13] In a formal sense, this afforded locations of *"interhelical nearest neighbors"* to create a 3-D map (not unlike multi-dimensional NMR analysis). At the time there was no way to sequence the identity of the four arms of the crossed fragments necessary to assign *a unique nearest neighbor address* (although in 2018 that would be trivial). One could argue that we failed to complete the 3-D project. This happened many times in our research program. We were often years and decades ahead of a problem where technologies in chemical biology did not exist to move the project to completion. In the year 2018, it remains a compelling research problem in biology to map the 3-D structure and dynamics of folding/unfolding of chromatin in the nucleus!

2.1 Molecular Recognition of DNA by Small Molecules – Multivalency for High Affinity

Some ligands bind noncovalently to duplex DNA by intercalation, the insertion of a flat heterocycle between the base pairs of the double helix. Graduate student Michael Becker designed a bis(intercalator), bis(methidium) spermine (BMSP) which had a DNA binding site size of four base pairs and affinity 10,000 times stronger for DNA than the monomer (Figure 2).[14] For a bis-intercalated species the polyamine spacer would lie in the minor groove of the DNA. We imagined that modification of the linker with respect to charge, chirality, length, flexibility and functionality could be expected to play a role in controlling the specificity of small molecule-DNA complexes. It was a first step launching our program on molecular recognition of DNA and a time-point regarding limitations of biophysical methods for characterizing ligand-DNA binding. For example, the method for characterizing the binding site size and affinity was Scatchard plots to calf thymus DNA and viscometric titrations to phage DNA. There was an outstanding group of biophysical chemists in the nucleic acid field and I found inspiration from the pioneering work in ligand-DNA interactions by Don Crothers (Yale) and Claude Hélène (Paris).

2.2 Recognition & Reaction: Cleavage of Double Helical DNA by MPE·Fe(II)

The analysis of ligands binding to nucleic acids by Scatchard plots was labor intensive, time consuming and the DNA target limited to calf thymus DNA and a few homopolymers e.g. poly (dG-dC)·poly (dG-dC). We despaired at the lack of precise biophysical methods that would reveal *in an unbiased*

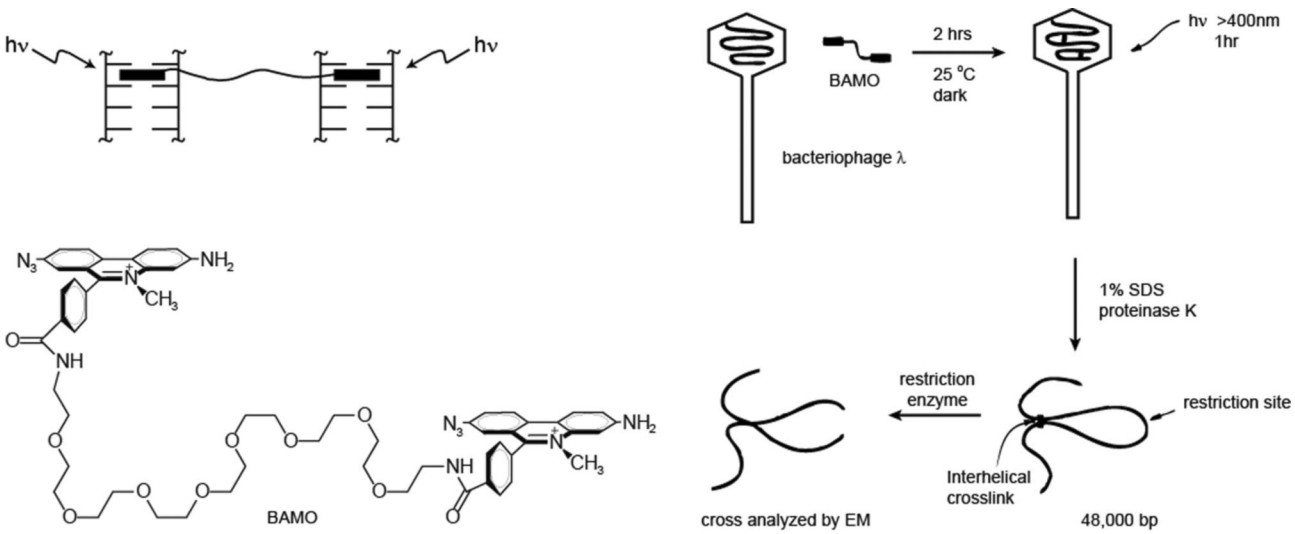


Figure 1. (left) Design of interhelical DNA–DNA crosslinker BAMO (right) Experimental protocol for photocrosslinking folded DNA in bacteriophage λ, followed by restriction digestion and visualization of unique covalent crosslinks by electron microscopy.

Israel Journal of Chemistry

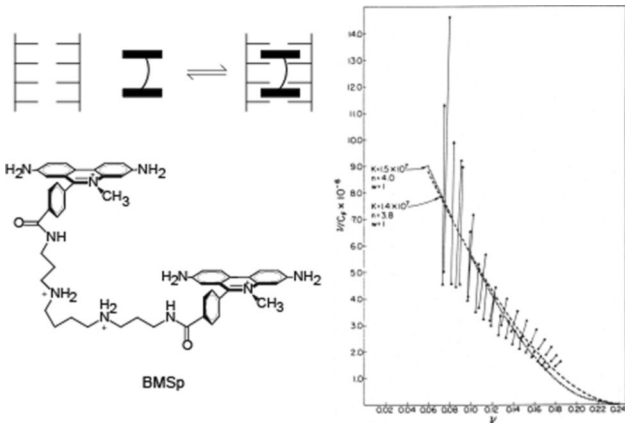


Figure 2. (left) Molecular recognition of DNA by a bisintercalator BMSp (right) Scatchard plot of binding of BMSp to calf thymus DNA.

experiment the sequence specificity of ligands designed at Caltech for sequence specific DNA recognition. For example, for a ligand that would bind 5 bp of DNA (one half turn of the helix), there are 512 different sequences of double helical DNA, combinations comprised of the four Watson Crick base pairs G·C, C·G, T·A, A·T. Ideally, we would want a rank order of sequence preferences if we had any hope of revealing general principles for DNA recognition.

There was one experiment in the literature that caught my attention. The natural product *bleomycin*, a glycopeptide that binds and cleaves DNA in a reaction that depends on ferrous ion and dioxygen. The sequence specificity of this binding and cleaving reaction could be determined by polyacrylamide gel electrophoresis. Remarkably, one could "know" the sequence specificity without any information how this glycopeptide bound iron and activated dioxygen at two base pair (5′-3′GC or GT) sites on DNA! Here was an inspiration from nature, the concept of using a DNA binding molecule to deliver a metal ion to a site on the DNA helix where activation of molecular oxygen results in cleavage of the DNA backbone. We needed to invent a simple version of the the bleomycin metal binding

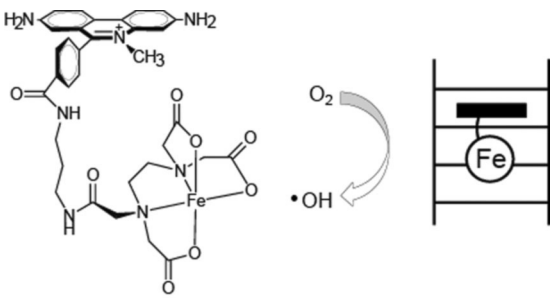


Figure 3. (left) Separate modules for recognition and reaction. MPE·Fe(II) contains two domains: intercalator for binding DNA and metal chelator EDTA for iron mediated reduction of O_2 to $^{\bullet}OH$, a short-lived diffusible highly reactive oxidant.

domain. With the guiding principle to keep it simple, we chose the iron chetator EDTA and the non-specific intercalator methidum as the DNA binder. Would a synthetic molecule with *two separate functional domains, recognition and reaction,* bind and cleave double helical DNA (Figure 3)? Graduate student Robert Hertzberg synthesized methidium tethered to EDTA (MPE) and showed that this molecule efficiently cleaved double helial DNA in a reaction dependent on iron and dioxygen.[15] Remarkably, reducing agents such as dithiothreitol (DTT) allowed the reaction to occur at very low concentrations of MPE and we interpreted this to mean that the reaction is *catalytic* and regenerates Fe(II) from Fe(III) for multiple turnovers. Unlike bleomycin·Fe(II), MPE·Fe(II) cleaves DNA non-sequence specifically, consistent with studies demonstrating that the intercalator methidum has no overall base composition specificity. Graduate student Michael Van Dyke would exploit then non-sequence specificity of MPE·Fe(II) to introduce the method of footprinting to our program to detect specificity of DNA binding by natural products in sequencing gels.[16]

2.3 Affinity Cleaving: Distamycin-EDTA·Fe(II)

In principle, attachment of EDTA·Fe(II) to a sequence specific DNA binding molecule could create a sequence specific DNA cleaving molecule. The crescent-shaped *distamycin* is a natural product containg three N-methylpyrrole caboxamides that was shown in biophysical experiments to bind in the minor groove of DNA with a preference for A/T rich polymers. Graduate student Peter Schultz and postdoctoral fellow John Taylor (Stork Ph.D., Columbia) synthesized the new bifunctional molecule, distamycin-EDTA, and tested the sequence specific cleavage on 32P end-labeled on DNA restriction fragments 167 and 381 nucleotides in length.[17] DNA cleavage by DE·Fe(II) was revealed on a denaturing polyacrylamide gel capable of resolving DNA fragments differing in length by one nucleotide. The predominant cleavage was centered around four-five base pair A/T rich sites. The three-four strand scissons flanking this site was consistent with a short-lived diffusible oxidizing species, most likely hydroxyl radical (Figure 4). In 1982 we speculated that attachment of EDTA·Fe to other sequence-specific DNA binding molecules such as antibiotics, polypeptides, oligonucleotides or proteins should provide a new class of DNA affinity cleaving molecules that may form a primitive basis for the design and construction of artificial restriction endonucleases with defined target sequences and binding site sizes. In retrospect, our frustration with Scatchard plots with lower information output caused us to invent a new method to explore the sequence specificity of synthetic designed DNA (and RNA) binding ligands. The homeogeneous DNA restriction fragments typically 150–250 bp in size is a *linear library of DNA sites* addressed by the specific 32P label on one end. Affinity cleavage allows detection and rank ordering (as a function of changing concentration) of DNA binding

Israel Journal of Chemistry

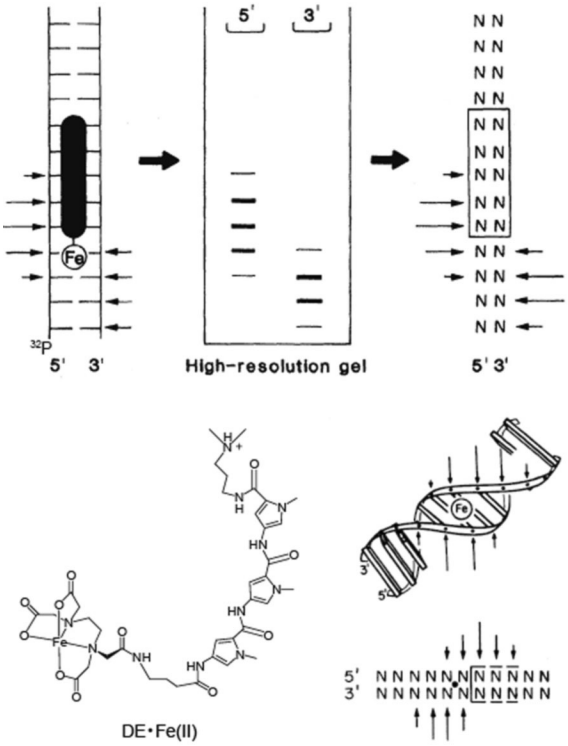


Figure 4. Affinity cleavage: Three pyrrolecarboxamide distamycin analog is tethered to the iron chelator EDTA. DE·Fe(II) binds long fragments of DNA and oxidatively cleaves at specific sites. The specificity of binding can be analyzed on sequencing gels. Information on binding orientation and groove location is revealed by the location and the cleavage pattern by short live diffusible oxidant, hydroxyl radical (*Science*, 1986).

sites by a synthetic ligand in an unbiased experiment (Figure 4).

In addition, the local cleavage pattern on opposite DNA strands reveals the *DNA groove location of the bound ligand*, asymmetric to the 3′ side is minor groove and 5′ side is major groove occupancy. In addition, since we know from synthesis the precise location of the EDTA, the cleavage pattern reveals the *orientation of the molecule at each DNA site.* Taken together, specific binding site sequences, groove location and orientation for multiple sites on DNA is revealed in a single unbiased screening experiment. We were off to the races. With the power of synthetic organic chemistry and the logic of incremental change afforded by synthesis combined with the powerful affinity cleavage method we believed we could move in earnest with our early vision to learn some chemical principles for the molecular recognition of DNA.

Subsequent structural studies of the natural products netropsin and distamycin revealed the structural basis for preferred specificity for A/T tracts. Richard Dickerson (UCLA) showed by x-ray crystallography that the C–H on the corner of the 5 membered N-methylpyrrole amino acid pointed to the floor of the minor groove.[18] The C–H was sterically permissive at the edge of a TA/AT base pair but there would be an energetic penalty for a steric clash with the exocyclic-NH_2 on the minor groove edge of a GC base pair. Dickerson's structure for the ligand:DNA complex was 1 : 1 stoichiometry. Several years later David Wemmer (UC Berkeley) showed by NMR methods in solution that distamycin could adopt a 2 : 1 stoichiometry, antiparallel side-by-side binding in the minor groove.[19] This complexity of two stoichiometries 1 : 1 and 2 : 1 could be explained by the likely difference in sequence dependent minor groove width (and dynamics) for various A/T rich sequences.

We will return to the chemical modification of the aromatic amino acid rings in distamycin in search of new sequence specificities for minor groove recognition, but to maintain chronological order, our first major breakthrough for synthetic ligands that can bind a large repertoire of DNA sequence occurred *in the major groove* and was enabled by the analytical power of the affinity cleaving method. The fact that hydroxyl radical is a relatively nonspecific cleaving species is useful when studying DNA recognition because cleavage specificity is due to the binding moiety alone, not some combination of cleavage specificity superimposed on binding specificity.

2.4 Major Groove Recognition by Triple Helix Formation: Site-Specific Genome Cleavage

Triple-stranded structures of polynucleotides were discovered by Felsenfeld, Davies, and Rich in 1957, shortly after the double helix paper by Watson and Crick. Poly(U) and poly(A) were found to form stable 2 : 1 complex in the presence of $MgCl_2$.[20] Poly(C) forms a triple-stranded complex at pH 6.2 with guanine oligonucleotides. In principle, isomorphous base triplets (T-A-T and C-G-C) can be formed between any homopurine·homopyrimidine duplex site and a corresponding homopyrimidine strand. A homopyrimidine oligodeoxyribonucleotide-EDTA should recognize the complementary sequence of double helical homopurine·homopyrimidine DNA in the major groove and yield a strand break at the target sequence. The affinity cleaving method with oligo DNA-EDTA·Fe allows the effects of reaction conditions, probe length and single base mismatches to be analyzed on high resolution sequencing gels. In addition, the orientation of the third strand can be determined as well as the identity of the groove occupied by the bound oligo DNA-EDTA probe. That said, an oligodeoxynucleotide is a polyanion and the energetics for binding a local region of anionic duplex DNA was unknown. Postdoctoral coworker Heinz Moser (Ph.D. Eschenmoser) undertook this project in 1985–86 and wrote an important paper on sequence specific cleavage of double helical DNA by triple helix formation (*Science* 1987).[21] He demonstrated sensitivity to single base mismatches for a 15mer probe and double strand site specific cleavage of plasmid DNA 4.06 kb in size (Figure 5). In the same year, my colleague Claude Hélène and coworkers in France reported that α-oligonucleotides could bind in the major groove of DNA by triple-helix

Israel Journal of Chemistry

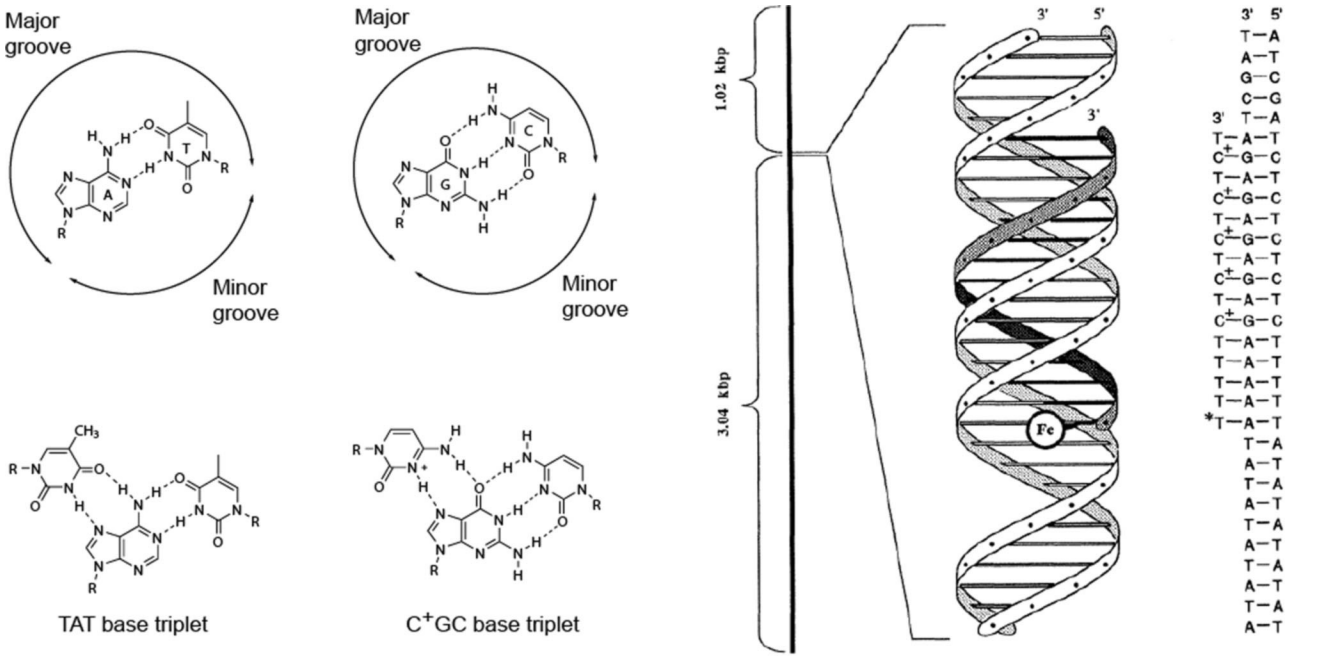


Figure 5. (Top) Watson-Crick base pairs, Isomorphous base triplets TAT and CGC. The additional pyrimidine strand is bound by Hoogsteen hydrogen bonds in the major groove to the complementary purine strand in the Watson-Crick double helix (Bottom) Site specific double strand cleavage of plasmid DNA analyzed on a nondenaturing agerose gel (*Science*, 1987).

formation.[22] The Pasadena and Paris groups recognized the potential to create "artificial repressors" based on triple helix forming analogs for "gene silencing", although the size of an anionic 15mer suggested hurdles regarding cell uptake and tissue penetration would need to be addressed.[23]

For oligonucleotides targeting 15 bp sites and with sensitivity to single base mismatches, one could imagine that pyrimidine probes equipped with DNA cleaving moieties could be useful tools *for mapping and editing chromosomes.* Graduate student Scott Strobel took up the task to ask whether this chemistry would be useful for *single site cleavage of human chromosomes.*[24–27] In a clever experimental design, Strobel combined the oligo third strand for single site binding/protection followed by exhaustive methylase modification and restriction enzyme cleavage. He reported the near quantitative single site enzymatic cleavage of *S. cerevisie* genome mediated by triple helix formation.[26] The 340-kilobase yeast chromosome III was cut uniquely at an overlapping homopurine-EcoRI target site 27 base pairs long to produce two expected cleavage products of 110 and 230 kb (*Nature* 1991).[26] Strobel then demonstrated this triple helix forming guide strand/enzymatic cleavage protocol can liberate a specific segment of a *human chromosome*, the tip of human chromosome 4 (Figure 6).[27] A 16-base pyrimidine oligodeoxyriboneucleotide was able to locate a 16-base pair purine target within more than 10 gigabase pairs of genomic DNA (*Science* 1991).

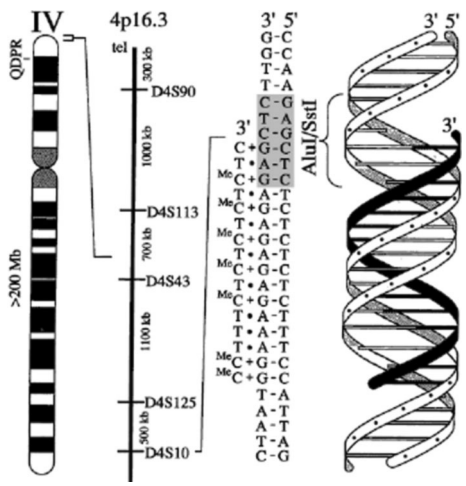


Figure 6. (left) Schematic diagram of human chromosome 4 and physical map of targeted cleavage site recognized by a 16 base ^{Me}CT oligodeoxyribonucleotide. (right) Schematic diagram of the triple helix complex. The pyrimidine oligodeoxyribonucleotide is bound in the major groove, parallel to the purine strand of the DNA duplex and covers the purine half of the AluI methylation site (*Science*, 1991).

2.5 Py–Im Polyamides – Minor Groove Recognition and the Pairing Rules

Pyrrole–Imidazole (Py–Im) polyamides evolved in our group over the course of 20 years (1982–2002) from the natural

Israel Journal of Chemistry

product distamycin to a class of programmable DNA binding oligomers with high sequence specificity and affinity for the minor groove of DNA.[28] The development of unbiased methods such as footprinting and affinity cleaving allowed new heterocycle amino acids to be screened for sequence specificity. A modular set of aromatic amino acids can be combined as side-by-side antiparallel pairs in the minor groove to distinguish the four Watson-Crick base pairs (Figures 7& 8).[29–32]

In a formal sense the edges of the four Watson Crick base pairs can be differentiated on the minor groove floor by the specific positions of hydrogen bond donors and acceptors, by differences in shape, and by electronic potential surfaces. Footprinting and affinity cleaving experiments by graduate student Warren Wade revealed that the imidazole containing polyamide ImPyPy bound to the five bp sequence 5′-WGWCW-3′ (where W=A or T) as an antiparallel dimer in the minor groove of DNA.[29] The data were consistent with an Im/Py pair targeting G·C, Py/Im pair targeting C·G and Py/Py pair specifying both T·A and A·T. NMR studies with David Wemmer[30] and x-ray crystal structures with Doug Rees (Caltech)[31] confirmed this specificity is due to hydrogen bonding between the imidazole nitrogen lone pair of the polyamide and the exocyclic 2-amino group of guanine on the minor groove edge of a G·C base pair.

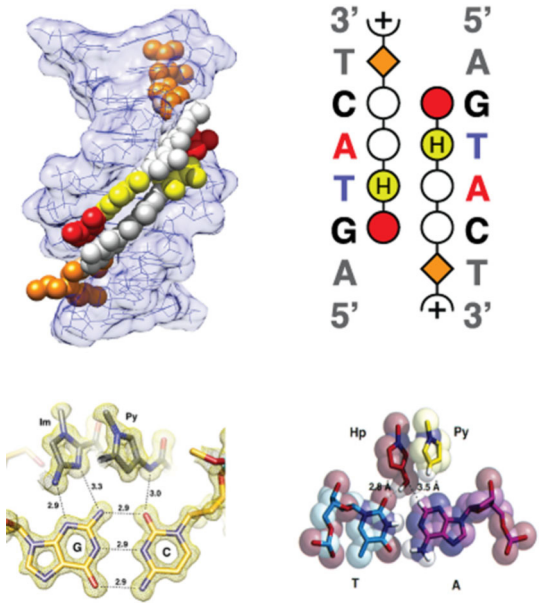
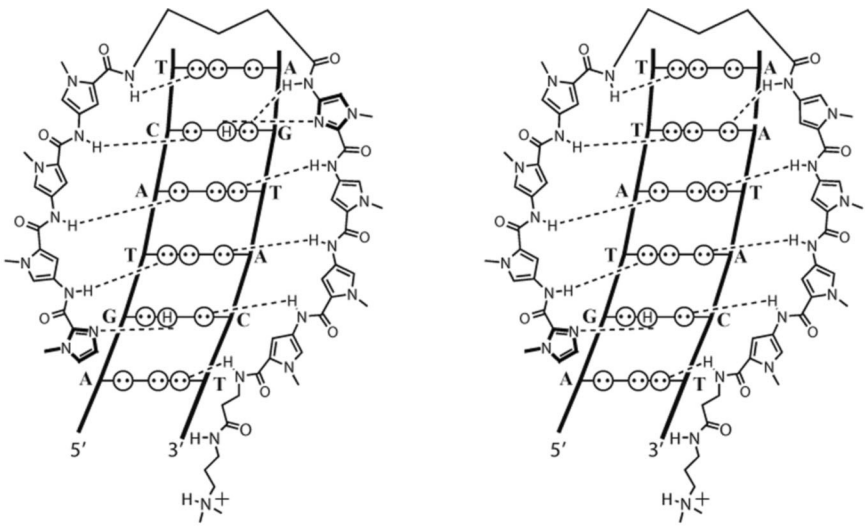


Figure 8. (left) Space filling model of $(ImHpPyPy)_2$·5′-CCAG-TACTGG-3′. Four ring pairs Im/Py, Py/Im, Hp/Py, Py/Hp bind and distinguish the four Watson-Crick base pairs. (Bottom) Space filling model of Im/Py interacting with the minor groove edge of GC base pair and Hp/Py with minor groove edge of TA base. The Hp–OH tightly fits the cleft formed by the adenine C2H (*Science*, 1998).



Equilibrium Association Constants (M^{-1})

Binding Site	ImPyPyPy-γ-ImPyPyPy-β-Dp	ImPyPyPy-γ-PyPyPyPy-β-Dp
5'- AGTACT -3'	3.7×10^{10}	5.0×10^{8}
5'- AGTATT -3'	4.1×10^{8}	3.5×10^{9}

Conditions: 10 mM Tris•HCl, 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$, pH 7.0, 22°C

Figure 7. (Top) Binding models of a six base pair sequences 5′-AGTACT-3′ and 5′-AGTATT-3′ in complex with 8-ring hairpin Py–Im polyamides coding for those sequences by the pairing rules. (Bottom) The equlibium association constants reveal the sensitivity for single base mismatch as well as affinities comparable to transcription factors (*Nature*, 1996).

Israel Journal of Chemistry

In contrast to G·C with a steric bump, the edge of an A·T base pair presents hydrogen bond acceptors on the floor of the minor groove and an asymmetric cleft between the thymine O2 and adenine C2. Guided by our simple bump and hole model, 3-hydroxypyrrole (Hp) when paired with Py was designed by graduate students Sarah White, Jason Szewczyk and Eldon Baird to specifically bind a T·A base pair.[32] X-ray crystallographic studies by graduate student Clara Kielkopf on polyamides containing the Hp/Py pair have shown that specificity arise from a combination of hydrogen bonding between hydroxyl and the thymine O2 together with shape-selective recognition of the asymmetric cleft by the hydroxyl group as predicted.[33] *Three aromatic amino acids – Py, Im and Hp – can be combined as four unsymmetrical ring pairs to recognize and distinguish each of the four Watson-Crick base pairs.* We refer to these as pairing rules. They are considered guidelines only. The pairing code Im/Py for G·C is agnostic regarding sequence context. Undoubtedly the sequence-dependent microstructure of DNA would be expected to play a role regarding the energetics of binding, some sequences being a more ideal steric fit than others.

2.6 The Hairpin Motif

In principle, crescent-shaped dimers of multi-ring Py–Im polyamides could bind in the minor groove antiparallel and in perfect register with complete overlap. Alternatively, antiparallel oligomers in the minor groove might adopt slipped dimer structures affording some ambiguity in predicting optimum sequence targeted. In order to align ring pairs in a predictable sense, covalent linkage of the carboxy and amino termini with an aliphatic tether should afford a hairpin structure where aromatic ring pairs in a polyamide oligomer are *unambiguously paired when folded in the minor groove.* Graduate student Milan Mirksich found that a diaminobutyric acid linker was the optimal "turn unit" for hairpin binding.[34] The γ-turn of a hairpin polyamide is specific for a A·T/T·A base pairs for steric reasons, the bump of a G·C base pair is an energetically unfavorable steric clash with the aliphatic turn. The hairpin structures resulted in ligands that bound DNA with 100 fold higher affinity compared with unlinked dimers. Several other linking approaches were explored beyond the hairpin, H-pin, U-pin motifs and cycles. In addition, tandem hairpin dimers were explored for binding larger binding site sizes.[35]

Eight-ring hairpin polyamides were shown by graduate student John Trauger to bind 6-bp sequences with equilibrium association constants $K_a = 10^8$ to $10^{10}\,M^{-1}$ not unlike the affinity of natural DNA binding transcription factors which typically bind 4–6 bp of DNA.[36] Py–Im oligomers of more than four contiguous ring pairs no longer match the curvature of the DNA helix. To allow the polyamide to relax and adjust to the DNA curvature John Trauger introduced the strategic incorporation of β-alanine/β -alanine pairs (β/β) in place of Py/Py pairs.[37] By extension β/Im and Im/β pairs could be used for targeting C·G and G·C, respectively.[38]

2.7 Disruption of Transcription Factor – DNA Interfaces and Inhibition of Pol II Elongation

In collaboration with Joel Gottesfeld (Scripps) transcription factor-DNA binding was found to be disrupted by hairpin Py–Im polyamides.[39] How could this be? Most transcription factors bind in the major groove of DNA (large shallow surface) whereas our hairpin oligomers bound exclusively in the narrow minor groove. A direct steric blockade is insufficient to explain this result. An explanation was discovered by graduate student David Chenoweth who determined a high resolution x-ray crystal structure of a cyclic polyamide bound to duplex DNA.[40] In the polyamide-DNA structure he observed a minor groove widening of up to 4 Å with a simultaneous compression of the major groove and bending of the DNA helix towards the major groove by $>15^\circ$. The *local structural alteration of DNA* caused by polyamide binding enforces a major groove surface geometry incompatible with transcription factor-DNA binding. In effect we have discovered an indirect *allosteric mechanism* for inhibition of transcription factor-DNA binding by cyclic (and by extension hairpin) Py–Im polyamides. Many human diseases, such as cancer, are caused by the overactivity of transcription factors and this raises the question whether cell permeable Py–Im polyamides could modulate (reprogram) dysregulated gene expression in human disease. Our program would need to evolve from synthetic organic/biophysical chemistry to cell biology, animal toxicity and xenograft experiments to follow the problem.

DNA targeting agents have been shown to inhibit DNA dependent enzymes including RNA polymerase, DNA polymerase, topoisomerase and helicases. DNA minor groove binders such as distamycin and actinomycin have been shown to inhibit RNA polymerase II mediated transcription in enzymatic studies and in cell culture. Similarly, Py–Im polyamides have been shown to inhibit the elongation reaction on DNA catalyzed by RNA polymerase II. In an in vitro transcription assay using isolated eukaryotic RNA pol II with a DNA template containing a single polyamide binding site Dong Wang and coworkers at UC San Diego were able to show robust transcription inhibition at nanomolar IC_{50}. The RNA pol II enzyme was stalled on the DNA template for up to 20 hours. Polyamide induced transcription inhibition was abolished when a single mismatch mutation was introduced into the polyamide binding site, demonstrating the sequence specificity of the effect.[41]

Medicinal chemistry issues such as cell permeation, nuclear localization, gene regulation in cell culture, global sequence analysis of sequence specificity, as well as preclinical studies including pharmacokinetics, toxicity in animals and xenograft cancer models led by graduate students Nick Nickols and Fei Yang have been recently summarized.[41]

Israel Journal of Chemistry

3. Recognition and Recruitment – Activation of Gene Expression

Eukaryotic transcription activators are minimally comprised of a DNA binding domain and a separable activation domain. In 1999, postdoctoral coworker Anna Mapp (Ph.D. Heathcock, UC Berkeley) replaced natural protein modules with synthetic counterparts to create *artificial transcription factors.* The engineered molecule contains a sequence-specific DNA binding polyamide covalently linked to a 15 residue peptide amphipathic helix activating region (AH) (Figure 9). This two domain hybrid molecule (4.2 kd) was show by postdoctoral Aseem Ansari (Sloan Kettering) to mediate high levels of DNA site-specific transcriptional activation *in vitro.*[42] Based on this success, barbell-shaped molecules with separate domains for a separate functions, *recognition and recruitment*, were designed as protein–DNA dimerizers (Figure 10). Postdoctoral coworker Hans Dieter Arndt and graduate student Ryan Stafford constructed a protein–DNA dimerizer comprised of a DNA binding polyamide linked to the peptide FYPWMKG which facilitates the binding of the natural transcription factor Exd to an adjacent DNA site (Figure 11). Amazingly, the Exd binding domain domain can be reduced to a *dipeptide* WM attached to the hairpin polyamide.[43] In collaboration with Aseem Ansari (Wisconsin) we hope to provide design principles for artificial transcription factors that may function in concert with the cellular regulatory circuitry.

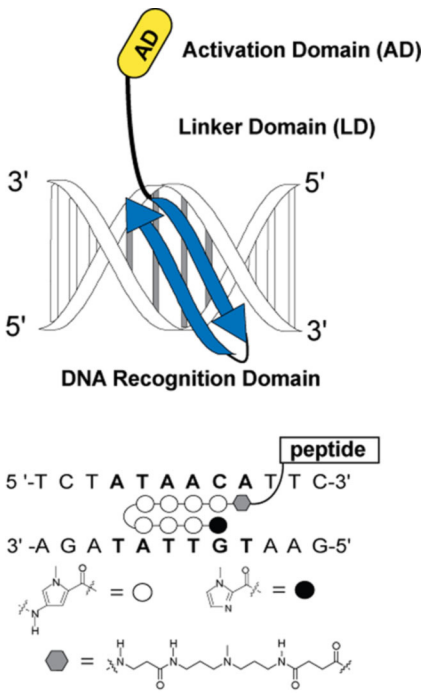


Figure 9. Design of a synthetic transcription activator comprised of peptide activation domain (AD), a linker domain (LD) and a hairpin Py–Im polyamide DNA binding domain (DBD).

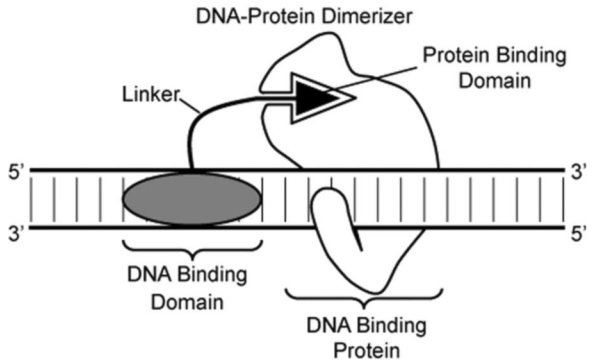


Figure 10. Design of a protein–DNA dimerizer constructed from a sequence specific DNA binding Py–Im hairpin polyamide and a short peptide which binds a pocket in a transcription factor.

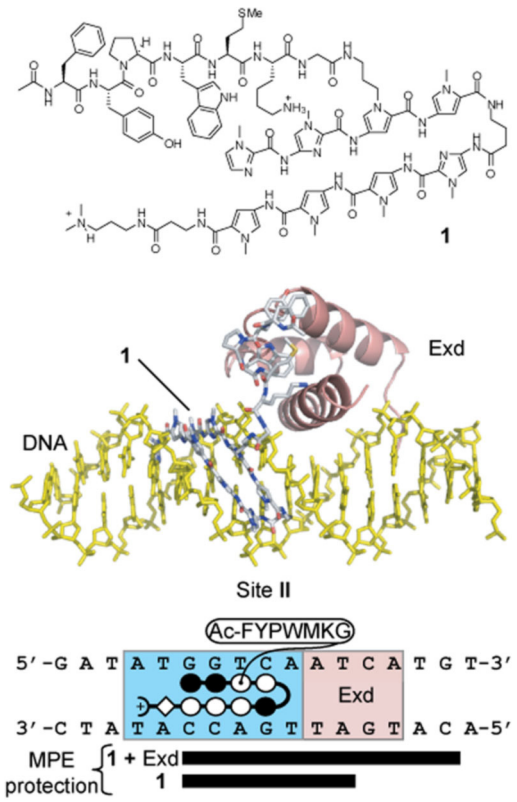


Figure 11. A protein–DNA dimerizer, hairpin polyamide-YPWM conjugate facilitates binding of the natural transcription factor Exd to an adjacent DNA site.

4. How do You View the Historical Development of Chemical Biology?

Our research on DNA recognition at Caltech was initiated in 1975, a decade before we had a first glimpse in 1987 from x-ray structures by Harrison and Ptashne how proteins might bind DNA (repressor-operator complex) sequence specifically.[44] One could argue we began our program too early

before sufficient analytical and biochemical methods were in place to facilitate our work. That said, chemical sequencing of DNA by Maxim Gilbert and high resolution separation of DNA strands by gel electrophoresis appeared in 1977 and methods to synthesize oligiodeoxynucleoties were emerging from Caruthers laboratory in 1981. We were lucky to ride the wave. A generation of creative researchers filled the literature with important papers. The field at the interface of chemistry and biology, rebranded chemical biology, had a critical mass of researchers and it was time for new journals to be created to spotlight the research. Just as an array of analytical techniques powered the explosive growth of organic chemistry in 1950–70 (IR, UV, NMR, mass spectrometry, x-ray crystallography), separation technologies (HPLC, gel electrophoreses), automated solid phase synthetic methods for DNA, RNA, peptides, and analytical methods (high resolution mass spec, multidimensional NMR, cryo-EM, single molecule fluorescence spectroscopy) as well as engineering (microfabrication of chips, robotics for HTS) has powered advances in chemical biology.

5. What have been the Most Significant Contributions?

It is difficult to rank the most significant contributions to science because the field has evolved over four decades 1975–2018. Advances in methodology and instrumentation during the past 20 years (e. g. multidimensional NMR, cryo-EM, mass spectrometry, single molecule detection, robotics) were not available in early days. In the 1980s the early pathfinders were trained as organic chemists, not biochemists, and applied chemical thinking to biologic processes. That said, first movers such as our group at Caltech were satisfying a culture that required quantitative chemical rigor. Indeed the beginning of chemical biology was CHEMICAL not quite full biology. Certainly there is a value following the increasing sophistication of the scientific contributions. Another metric is to track the evolution of generations of researchers. I believe there are four generations. How do I arrive at four? I was speaking at an ACS National meeting in Dallas (2014) and scheduled to be the final speaker on the afternoon program in Chemical Biology. The person before me was a young assistant professor Zev Gartner (UCSF) who gave a lecture on tissue engineering. I realized I was following my academic *great grandchild* (Dervan-Schultz-Liu-Gartner)!

In my view, a culture for the first movers that valued chemical insight over biological expertise was acceptable. These founder labs became *transition laboratories, incubator training* for the next generation of chemists who would balance the chemical-biology portfolio. Chemistry graduate students at Caltech such as Peter Schultz, Tadhg Begley, Brent Iverson, Scott Strobel, Milan Mirksich, Peter Beal, Clay Wang, Adam Urbach, David Chenoweth readily embraced graduate research in chemical biology. It is perhaps instructive to examine the backgrounds of postdoctoral fellows Ph.D.'s in chemistry who found their way to Pasadena to join a group grounded in physical organic and synthetic chemistry, but were committed to the biophysical analysis of biopolymers such as nucleic acids and proteins. From Breslow's physical organic group at Columbia, Sam Gellman, Alanna Schepartz and Eric Kool came to Pasadena. From Berson (Yale), came Mark Greenberg. From Koji Nakanishi (Columbia) came John Termini. In synthetic organic chemistry, John Taylor from Stork, David Horne from George Buchi (MIT), Laura Kiessling from Schreiber (Yale), Erick Carreira from Evans (Harvard), Yitzhak Tor from Shanzer (Israel), Heinz Moser from Eschenmoser (ETH), Alex Heckel from Seebach (ETH), Anna Mapp from Heathcock (Berkeley), Thomas Minehan from Kishi (Harvard). I believe it is a mistake to diminish the early evolution of a new field by saying the mission of chemical biology is to impress the biology customer. Let's create a large tent and look outward in all directions. As the third generation of leaders emerged (1990–), they trained as Ph.D. in chemistry labs but followed with a *postdoctoral in biology.* This third generation were true hybrid chemists-biologists.

Two contributions that were highly influential emerged from Schultz-Lerner (UC Berkeley, Scripps) and Schreiber-Crabtree (Harvard, Stanford). Schultz-Lerner's *catalytic antibodies* taught us new chemistry about the mechanisms of enzyme catalysis and how the binding energy of proteins can facilitate chemical transformations.[45,46] Schreiber-Crabtree's papers (1993–96) on *chemical inducers of proximity* (CIP) were a remarkable contribution in the field of chemical biology.[47] This chemical contribution (first described as chemically induced dimerization) has informed areas of research in biology ranging from fundamental advances to the development of cellular and molecular therapeutics. There are others. Dennis Dougherty's beautiful physical organic/ab initio quantum mechanics papers led him to the discovery of the cation-π binding site in the nicotine receptor and the importance of cation-π interactions in neurobiology. Who would have guessed that quarternary ammonium group of acetylcholine binds an aromatic box in the nicotinic acetylcholine receptor?[48] Sam Gellman and Dieter Seebach's vision that beyond natural peptide oligomers there is a huge field of *"foldamers"* and physical organic principles can be developed to understand the intra- and intermolecular forces connecting foldamer sequence to 3-D structure and function.[49] Orgel-Eschenmoser-Szostak-Sutherland asked profoundly important chemical questions: how did life begin? What were the first possible molecules that could carry information, replicate and catalyze reactions? Advances in synthetic chemistry of natural polymers were important. The beautiful complex carbohydrate field unfolded with synthetic leaps by Wong-Seeberger-Danishefsky paving the way for new vaccines.[50] The manipulation of cell surfaces by Bertozzi and Kiessling for chemical modifications that will be useful in cell therapy and understanding the immune response.[51,52] For drug discovery with undruggable targets, there is a bump and hole strategy by

Kevan Shokat (UCSF) engineering unnatural nucleotide-specificity for tyrosine kinase[53] and activity based protein profiling by Ben Cravatt (Scripps) and his chemical strategies for the global analysis of enzyme function.[54] Finally the discovery by Doudna and coworkers (Berkeley) of the CRISPR/cas genome cutting machinery[55] followed by the protein engineering of CRISPR by David Liu (Harvard) for programmable editing of a target base in genomic DNA is driving new discovery in biology.[56]

6. Is Chemical Biology a Service Technology or Independent Science?

The answer is both. The phrase service technology sounds somewhat subservient and suggests chemical biology makes tools and reagents which biologists use to make the important and significant discoveries. In my view there is no problem here! New methods drive and enable new discoveries often in a neighboring interdisciplinary field. The example in chemistry is the profound influence of physics on chemistry, creating tools (NMR, x-ray) that powered discovery in chemistry and later biochemistry. It is the case that new methods in chemistry, chemical inducers of proximity and CRISPR/cas influenced profoundly discovery of mechanisms in biology.

Some of the key achievements in chemical biology were predictable and some were not. Rumfeld would call these *known unknowns* and the *unknown unknowns.* I believe the most exciting discoveries will be the *unknown unknowns. What I mean here is that one cannot even phrase the scientific question.* The ignorance is profound, that something chemical is "out there" in the biological universe we don't know about. Who would have guessed there is *endogenous antisense* (RNAi/siRNA)? Who would have guessed there would be CRISPR/cas in bacteria with the power to be *gene editing tools*? Who would have guessed that *DNA replication* in a cell manages the complex multiprotein machinery by *DNA mediated electron transfer* for protein–protein signaling on DNA?! The (4Fe4S) cluster of human DNA primase functions as a redox switch using DNA charge transport.[57]

7. What are the Open Questions?

I believe that solving the 3-D structure of chromatin in a living cell is an important problem which will take the combined effort of chemical biology, chemical physics, bioinformatics and biology. In fact, there is not one genome structure but *multiple folded states* in *dynamic movement* that are important in controlling gene expression in different cells and tissues in time. How does all this work? The physics of chromatin fibers imposes a range of constraints on communicating between proper expression of genes and regulatory elements.

In the area of discovery in human therapeutics, I am not surprised by the recent emergence of core chemical biology groups in big pharma. There was early acceptance of decades ago regarding the value of chemical biology groups in academic research universities, as well as early stage biotech discovery programs. But the hiring of core chemical biology groups in pharma is a big deal. This signals change. For the past century modern drug discovery (antibiotics, antivirals, statins, etc.) changed our lives for the better. The platform has been chemistry, small molecules targeted to a single protein target, a paradigm of single molecule-protein (or DNA) interface. In the last 10 years, the world of drug discovery has changed with the breakthrough blockbusters being biologics (e. g. antibodies, RNA). The emergence of biology as a quantitative science, protein engineering, synthetic biology, and whole genome sequencing suggests *biology and biologics will shape the future of drug discovery!* What does this mean for chemical biology and how does this connect to the emergence of chemical biology groups in big pharma?

Chemical biologists will create new biology to *reprogram disease states.* Take the example of a blockbuster drug *revlimid*, a small molecule important in cancer therapy, in particular multiple myeloma, whose mechanism of action was poorly understood. Only recently it was discovered how this works. It is a small dimerizer that recruits E3 ubiquitin ligases on one end and a transcription factor on the other for protein degradation. This reveals a paradigm shift for the field of drug discovery…invent small molecule chemical dimerizers that degrade dysregulated proteins by the cell's proteasome machinery. I believe chemical biologists such as Craig Crews (Yale) and Jay Bradner (Novartis) are leading this field and one can predict new pharma CIP-based therapeutics degrading dysregulated proteins such as androgen receptor (AR) in prostate cancer.

Another area of future advances will be *regenerative medicine.* Chemical biologists with a materials engineering bent will learn to replace dysfunctional tissue with new biocompatible materials or perhaps, even better, invent small molecules that reprogram stem cells to regenerate new natural tissue. Chemical biologist Zev Gartner (UCSF) is working to understand how cells assemble into multicellular tissues, how the structure of tissues controls the behavior of individual cells and how changes to tissue structure drive the progression of diseases like cancer. Finally, we have seen the dramatic power of immunotherapy, biologics enabling human T-cells to attack (self) cancer. Early amazing success has been seen with remission in aggressive hot cancers such as melanoma. That said, can chemical biology partner with immunotherapy and invent co-drugs administered to turn cold cancers such as prostate cancer into a hot cancer for immunotherapy efficacy?

Finally, as we migrate to a biologic therapeutic universe of large molecules (proteins, RNA) for use in humans, how do we get them inside the body? I believe cross functional teams of chemical biologists, applied physics, and device engineers will create means to deliver biologics with large volumes and viscosities to patients who want to be pain-free, motivated to

stay on therapy and self-administered without inconvenient visits to a hospital for infusions.

Acknowledgements

I wish to thank the Editors in particular Ehud Keinan for the invitation to participate in this Rosarium Philosophorum on Chemical Biology. I am grateful to the National Institutes of Health (General Medical) for 45 years of uninterrupted financial support and, more recently, the Prostate Cancer Foundation. I am grateful for the contributions by the 192 graduate students and postdoctoral fellows who passed through my laboratory at Caltech 1973–2018. It was difficult for them to have an advisor who was always the learner and never the expert in the room as we migrated from biophysical chemistry to cell biology to xenograft mouse experiments. There was risk and many projects failed but I believe the experience fortified my former coworkers for success later in their careers. I am grateful to my collaborators, David Wemmer, Doug Rees, Joel Gottesfeld, Karolin Luger, Aseem Ansari, Michael Phelps and Dong Wang. I am indebted for the support of our program and encouragement over the years by Michael Waring, Hiroshi Sugiyama and Bengt Nordén. This personal essay is dedicated to Claude Hélène who shared our passion for discovery in nucleic acids and who passed away too young. I am grateful to my coworker Alexis Kurmis who helped prepare graphics.

References

[1] P. B. Dervan, T. Uyehara, *J. Am. Chem. Soc.* **1976**, *98*, 1262–1264.
[2] P. B. Dervan, T. Uyehara, D. S. Santilli, *J. Am. Chem. Soc.* **1979**, *101*, 2069–2075.
[3] D. S. Santilli, P. B. Dervan, *J. Am. Chem. Soc.* **1979**, *101*, 3663–3664.
[4] P. B. Dervan, D. S. Santilli, *J. Am. Chem. Soc.* **1980**, *102*, 3863–3870.
[5] R. Breslow, S. Baldwin, T. Flichtner, P. Kalicky, S. Liu, W. Washburn, *J. Am. Chem. Soc.* **1973**, *95*, 3251–3262
[6] J. R. Barrio, J. A. Secrist, N. J. Leonard, *Biochem. Biophys. Res. Commun.* **1972**, *46*, 597–603.
[7] J.-M. Lehn. *Angew. Chem. Int. Ed.* **1990**, *29*, 1304–1319; *Angew. Chem.* **1990**, *102*, 1347–1362.
[8] A. M. Maxam, W. A. Gilbert, *Proc. Natl. Acad. Sci. USA* **1977**, *74*, 560–564.
[9] S. L. Beaucage, M. H. Caruthers, *Tet. Lett.* **1981**, *22*, 1859–1862.
[10] H. R. Drew, R. M. Wing, T. Takano, C. Broka, S. Tanaka, K. Itakura, D. E. Dickerson, *Proc. Natl. Acad. Sci. USA* **1981**, *78*, 2179–2183.
[11] A. Kornberg, *Biochemistry* **1987**, *26*, 6888–6891.
[12] P. B. Dervan, *Science* **1986**, *232*, 464–471.
[13] M. M. Mitchell, P. B. Dervan, *J. Am. Chem. Soc.* **1982**, *104*, 4265–4266.
[14] a) P. B. Dervan, M. M. Becker, *J. Am. Chem. Soc.* **1978**, *100*, 1968–1970. b) M. M. Becker, P. B. Dervan, *J. Am. Chem. Soc.* **1979**, *101*, 3664–3666.
[15] R. P. Hertzberg, P. B. Dervan, *J. Am. Chem. Soc.* **1982**, *104*, 313–315.
[16] a) M. W. Van Dyke, R. P. Hertzberg, P. B. Dervan, *Proc. Natl. Acad. Sci. USA* **1982**, *79*, 5470–5474. b) M. W. Van Dyke, P. B. Dervan, *Science* **1984**, *225*, 1122–1127.
[17] P. G. Schultz, J. S. Taylor, P. B. Dervan, *J. Am. Chem. Soc.* **1982**, *104*, 6861–6863; J. S. Taylor, P. G. Schultz, P. B. Dervan, *Tetrahedron* **1984**, *40*, 457–465.
[18] M. L. Kopka, C. Yoon, D. Goodsell, P. Pjura, R. E. Dickerson, *Proc. Natl. Acad. Sci. USA*. **1985**, *82*, 1376–1380.
[19] J. G. Pelton, D. E. Wemmer, *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 5723–5727.
[20] G. Felsenfeld, D. R. Davies, A. Rich, *J. Am. Chem. Soc.* **1957**, *79*, 2023–2024.
[21] H. E. Moser, P. B. Dervan, *Science* **1987**, *238*, 645 –650.
[22] T. L. Doan, L. Perrouault, D. Praseuth, N. Hahhoub, J.-L. Decout, J. Lhomme, C. Hélène, *Nucleic Acids Res.* **1987**, *15*, 7749–7760.
[23] a) L. J. Maher III, B. J. Wold, P. B. Dervan, *Science* **1989**, *245*, 725–730. b) L. J. Maher III, P. B. Dervan, B. J. Wold, *Biochemistry* **1992**, *31*, 70–81.
[24] S. A. Strobel, H. E. Moser, P. B. Dervan, *J. Am. Chem. Soc.* **1998**, *110*, 7927–7929.
[25] S. A. Strobel, P. B. Dervan, *Science* **1990**, *249*, 73–75.
[26] S. A. Strobel, P. B. Dervan, *Nature* **1991**, *350*, 172–174.
[27] S. A. Strobel, L. A. Doucette-Stamm, L. Riba, D. E. Housman, P. B. Dervan, *Science* **1991**, *254*, 1639–1642.
[28] P. B. Dervan, *Bioorg. Med. Chem.* **2001**, *9*, 2215–2235.
[29] a) W. S. Wade, M. Mrksich, P. B. Dervan, *J. Am. Chem. Soc.* **1992**, *114*, 8783–8794. b) W. S. Wade, M. Mrksich, P. B. Dervan, *Biochemistry* **1993**, *32*, 11385–11389. c) M. Mrksich, P. B. Dervan, *J. Am. Chem. Soc.* **1995**, *117*, 3325–3332.
[30] M. Mrksich, W. S. Wade, T. J. Dwyer, B. H. Geierstanger, D. E. Wemmer, P. B. Dervan, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 7586–7590.
[31] C. L. Kielkopf, E. E. Baird, P. B. Dervan, D. C. Rees, *Nature Struc. Biol.* **1998**, *5*, 104–109.
[32] S. White, J. W. Szewczyk, J. M. Turner, E. E. Baird, P. B. Dervan, *Nature* **1998**, *391*, 468–471.
[33] C. L. Kielkopf, S. White, J. W. Szewczyk, J. M. Turner, E. E. Baird, P. B. Dervan, D. C. Rees, *Science* **1998**, *282*, 111–115.
[34] M. Mrksich, M. E. Parks, P. B. Dervan, *J. Am. Chem. Soc.* **1994**, *116*, 7983–7988.
[35] a) D. M. Herman, E. E. Baird, P. B. Dervan, *Chem. Eur. J.* **1999**, *5*, 975–983. b) P. Weyermann, P. B. Dervan, *J. Am. Chem. Soc.* **2002**, *124*, 6872–6878. c) I. Kers, P. B. Dervan. *Bioorg. Med. Chem.* **2002**, *10*, 3339–3349.
[36] J. W. Trauger, E. E. Baird, P. B. Dervan, *Nature* **1996**, *382*, 559–561.
[37] a) J. W. Trauger, E. E. Baird, M. Mrksich, P. B. Dervan, *J. Am. Chem. Soc.* **1996**, *118*, 6160–6166. b) J. W. Trauger, E. E. Baird, P. B. Dervan, *Chem. Biol.* **1996**, *3*, 369–377.
[38] a) S. White, E. E. Baird, P. B. Dervan, *Chem. Biol.* **1997**, *4*, 569–578. b) J. M. Turner, S. E. Swalley, E. E. Baird, P. B. Dervan, *J. Am. Chem. Soc.* **1998**, *120*, 6219–6226.
[39] J. M. Gottesfeld, L. Neely, J. W. Trauger, E. E. Baird, P. B. Dervan, *Nature* **1997**, *387*, 202–205.
[40] D. M. Chenoweth, P. B. Dervan, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 13175–13179
[41] P. B. Dervan A. A. Kurmis, P. B. Finn in *DNA-targeting Molecules as Therapeutic Agents* (Ed.: M. J. Waring), Royal

Society of Chemistry, CPI Group (UK) LTD, Coydon, United Kingdom, **2018**, pp. 298–325.
[42] A. Z. Ansari, A. K. Mapp, D. H. Nguyen, P. B. Dervan, M. Ptashne, *Chem. Biol.* **2001**, *8*, 583–592.
[43] H. D. Arndt, K. E. Hauschild, D. P. Sullivan, K. Lake, P. B. Dervan, A. Z. Ansari, *J. Am. Chem. Soc.* **2003**, *125*, 13322–13323.
[44] J. E. Anderson, M. Ptashne, S. C. Harrison, *Nature* **1987**, *326*, 846–852
[45] R. A. Lerner, S. J. Benkovic, P. G. Schultz, *Science.* **1991**, *252*, 659–667
[46] P. G. Schultz, R. A. Lerner, *Science.* **1995**, *269*, 1835–1842.
[47] a) D. M. Spencer, T. J. Wandless, S. L. Schreiber, G. R. Crabtree, *Science* **1993**, *262*, 1019–1024. b) P. J. Belshaw, D. M. Spencer, G. R. Crabtree, S. L. Schreiber, *Chem. Biol.* **1996**, *3*, 731–738. c) S. N. Ho, S. R. Biggar, D. M. Spencer, S. L. Schreiber, G. R. Crabtree, *Nature* **1996**, *382*, 822–826. d) B. Z. Stanton, E. J. Chory, G. R. Crabtree. *Science* **2018**, *359*, 1109–1117.
[48] a) R. A. Kumpf, D. A. Dougherty, *Science.* **1993**, *261*, 1708–1710. b) W. Zhong, J. P. Gallivan, Y. Zhang, L. Li, H. A. Lester, D. A. Dougherty, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 12088–12093.
[49] a) S. H. Gellman *Acc. Chem. Res.* **1998**, *31*, 173–180. b) D. Seebach, S. Abele, K. Gademann, B. Jaun, *Angew. Chem. Int. Ed.* **1999**, *38*, 1595–1597.
[50] a) K. M. Koeller, C-H. Wong, *Chem. Rev.* **2000**, *100*, 4465–4494. b) O. J. Plante, E. R. Palmacci, P. H. Seeberger, *Science.* **2001**, *291*, 1523–1527. c) S. J. Danishefsky, M. T. Bilodeau, *Angew. Chem. Int. Ed.* **1996**, *35*, 1380–1419.
[51] a) L. K. Mahal, K. J. Yarema, C. R. Bertozzi, *Science.* **1997**, *276*, 1125–1128. b) N. J. Agard, C. R. Bertozzi, *Acc. Chem. Res.* **2009**, *42*, 788–797.
[52] L. L. Kiessling, J. E. Gestwicki, L. E. Strong, *Curr. Opin. Chem. Biol.* **2000**, *4*, 696–703.
[53] a) K. Shah, Y. Liu, C. Deimengian, K. M. Shokat. *Proc. Nat. Acad. Sci. USA.* **1997**, *94*, 3565–3570. b) M. S. Cohen, C. Zhang, K. M. Shokat, J. Taunton, *Science.* **2005**, *308*, 1318–1321.
[54] L. Bar-Peled, E. K. Kemper, R. M. Suciu, E. V. Vinogradova, K. M. Backus, B. D. Horning, T. A. Paul, T.-A. Ichu, R. U. Svensson, J. Olucha, M. W. Chang, B. P. Kok, Z. Zhou, N. Ihle, M. M. Dix, P. Jiang, M. M. Hayward, E. Saez, R. J. Shaw, B. F. Cravatt, *Cell* **2017**, *171*, 696–709.
[55] J. A. Doudna, E. Charpentier, *Science* **2014**, *346*, 1258096-1–1258096-9.
[56] A. C. Komar, Y. B. Kim, M. S. Packer, J. A. Zuris, D. R. Liu, *Nature* **2016**, *533*, 420–424.
[57] E. O'Brien, M. E. Holt, M. K. Thompson, L. E. Salay, A. C. Ehlinger, W. J. Chazin, J. K. Barton, *Science* **2017**, *355*, 813–822.

Manuscript received: November 30, 2018
Version of record online: January 23, 2019