

The shaping of a molecular linguist: How a career studying DNA energetics revealed the language of molecular communication

<https://doi.org/10.1016/j.jbc.2021.100522>

Kenneth J. Breslauer*

From the Department of Chemistry and Chemical Biology, Rutgers University, Piscataway, New Jersey, USA and The Rutgers Cancer Institute of New Jersey, New Brunswick, New Jersey, USA

Edited by F. Peter Guengerich

My personal and professional journeys have been far from predictable based on my early childhood. Owing to a range of serendipitous influences, I miraculously transitioned from a rebellious, apathetic teenage street urchin who did poorly in school to a highly motivated, disciplined, and ambitious academic honors student. I was the proverbial “late bloomer.” Ultimately, I earned my PhD in biophysical chemistry at Yale, followed by a postdoc fellowship at Berkeley. These two meccas of thermodynamics, coupled with my deep fascination with biology, instilled in me a passion to pursue an academic career focused on mapping the energy landscapes of biological systems. I viewed differential energetics as the language of molecular communication that would dictate and control biological structures, as well as modulate the modes of action associated with biological functions. I wanted to be a “molecular linguist.” For the next 50 years, my group and I used a combination of spectroscopic and calorimetric techniques to characterize the energy profiles of the polymorphic conformational space of DNA molecules, their differential ligand-binding properties, and the energy landscapes associated with mutagenic DNA damage recognition, repair, and replication. As elaborated below, the resultant energy databases have enabled the development of quantitative molecular biology through the rational design of primers, probes, and arrays for diagnostic, therapeutic, and molecular-profiling protocols, which collectively have contributed to a myriad of biomedical assays. Such profiling is further justified by yielding unique energy-based insights that complement and expand elegant, structure-based understandings of biological processes.

I was born in Jönköping, Sweden, the third child of parents who were refugees from Nazi Germany and who eventually settled in New York City. My mother was a Columbia University-trained social worker who always provided me with gentle nurturing and encouragement. My father was a frustrated academician who became a businessman in New York. In stark contrast with my mother, my father was strict and demanding. Much to my father's delight, my older sister and brother both were academic superstars. By contrast, much to

my father's consternation, I was a “street urchin” who was far more interested in sports than academics.

I was “saved” by my high school baseball coach, Dr Edson Scudder, a Princeton PhD, who taught history. He told me that unless I improved my grades, he would not allow me on the baseball team. Each day, Dr Scudder would stay after school and tutor me. My grades soared, as I became the only person in the high school to go in one semester from academic probation to the academic honor roll! For the rest of my high school years, I started in left field and won all-conference honors. Despite my relatively weak arm, I was a very good hitter. I went on to college and majored in history because of Dr Scudder's influence. I also majored in chemistry because I wanted eventually to become employed. I believe I was the only such double major that year at the university.

From these somewhat uneven beginnings, I never could have envisioned a day when I would write a *Reflections* piece of my nearly 5 decades of DNA scientific research. My personal story proves that, in the race of life, not all winners take the lead immediately out of the starting gate. I am convinced this is why I never give up on any of my students, knowing that they also may prove to be late bloomers. I am humbled and honored to be able to share with you my personal and professional reflections.

As a teenager growing up in New York City, I had three passions: baseball, history, and science. Brief stints playing college and minor league baseball convinced me to focus my attention on the other two passions. As noted above, I earned a BA in history and a B.S. in chemistry from the University of Wisconsin at Madison, graduating in 1968. The Vietnam War and the associated draft sharply shaped the next prioritization of my career. Science would become my vocation, and history would be my avocation, as they both remain so to this day.

In graduate school at Yale, I began as a physical organic chemist. Under the mentorship of Professor Jerome A. Berson, I investigated terpene rearrangements. The goal was to test the limits of the orbital symmetry properties that formed the underpinnings of the Nobel Prize-winning Woodward-Hoffmann rules. At the same time, I took an elective course in physical biochemistry, taught by Joe Fruton and Fred Richards. This course had a profound impact on my scientific trajectory. I was fascinated to learn how the three-dimensional shapes

* For correspondence: Kenneth J. Breslauer, kjbdna@rutgers.edu.

and functions of large, complex biological compounds, including polymers, could be understood in terms of fundamental thermodynamic and kinetic properties. The lecturers frequently underscored that all biochemical events are controlled and regulated by differential energetics, which are manifest in maps of the isolated and interacting energy landscapes. In essence, favorable and unfavorable interaction energies constitute the words of the molecular language by which molecules communicate and orchestrate their structures, stabilities, and functions, or dysfunctions. I was smitten!

I went on to pursue my PhD studies in biothermodynamics and kinetics at Yale University with Julian Sturtevant (1), followed by a postdoc at the University of California, Berkeley, with Ignacio “Nacho” Tinoco Jr. (2), and finally, in 1974, a faculty position at Rutgers University, where I remain to this day.

During my first 5 or 6 years at Rutgers, I split my time between intramural softball with the students and setting up my laboratory, where I conducted research in pursuit of nature’s secrets of life, as well as in pursuit of tenure! I have been at Rutgers for 47 years, during which I became a husband, a father, the Linus C. Pauling distinguished university professor, the founding dean of the Division of Life Sciences, and vice president for health science partnerships. Currently enabled by an outstanding group of colleagues in my laboratory, I continue my biophysical research programs, which collectively focus on characterizing the forces that modulate the regulation and dysregulation of biological processes crucial to human health.

My laboratory at Rutgers has spent nearly 5 decades creating DNA databases that systematically characterize the intramolecular and intermolecular forces, which collectively constitute the words, sentences, paragraphs, and chapters in the book of molecular communication. As such, one could reasonably characterize me as a molecular linguist.

This scientific focus was shaped within me by my early training at Yale from 1968 to 1972. Julian Sturtevant (1) and Donald Crothers (3) unrelentingly emphasized to me, and others, the importance of quantitatively characterizing the energy profiles of molecular assemblies to understand the origins of the structures, as well as the modes of action of their functions. I was surprised by how quickly I became a convert to their mantras because as a student I observed that no single word in the scientific lexicon could more rapidly empty a classroom than “thermodynamics.” However, conversations and classes with them, as well as with Lars Onsager, Oktay Sinanoğlu, Kenneth Wiberg, and others, made energetic profiling my passion rather than my fear. My conversion was further facilitated by Don Crothers asking me to proofread and propose edits for a textbook manuscript that he, Ignacio Tinoco, and Victor Bloomfield were writing (4). I was mesmerized by the quantitative rigor and clarity of explanations presented in the text, and I felt privileged to have such a sneak peek before publication. This opportunity was coupled with the 1973 publication of the Tinoco/Crothers landmark paper “Improved Estimation of Secondary Structure in

Ribonucleic Acids” (5), and I became even more convinced of the need for further energy studies on biological systems. In 1973, Douglas Turner and I were postdocs in the Tinoco laboratory, during which time we became close friends; although I still cannot get over the fact that he bested me 1-0 in a game of stickball! It is interesting to note that when we accepted independent faculty appointments, Doug built his stellar academic career on an impressive and much-needed elaboration of RNA energy profiling (referred to as the “Turner Numbers”) (6), while my academic career has focused on establishing DNA energy databases, including their drug- and protein-binding properties.

My choice of research programs as well as my deep commitment to data collection were profoundly influenced by my years in Sturtevant’s laboratory at Yale (1). I fondly recall a conversation with Julian in which I showed him my voluminous tabulation of calorimetric data on a biological system and lamented that I could not discern any pattern that would glean new insights. He gently smiled and said, “That only means you haven’t yet collected enough data, since the patterns will become self-evident when you reach the critical data density.” From that moment forward, I committed myself to a career of databasing, of ever-increasing density, on biologically relevant molecules. It therefore should come as no surprise that as a dean and vice president, I recruited Helen Berman, who established at Rutgers the international structural database (Research Collaboratory for Structural Bioinformatics Protein Data Bank, or RCSB EDB) (7, 8) and subsequently recruited her successor, Stephen Burley, as the director (9). During my early years, in the mid-1970s to early 1980s, I also was influenced by the published works of and conversations with Walter Kauzmann, Rufus Lumry, Peter Privalov, Harold Scheraga, Charles Tanford, and other luminaries. These interactions further convinced me of the importance of an energy database focus for my own research. While Tinoco, Crothers, Turner, and others focused on RNA energetics, I decided that direct, model-independent, calorimetric characterizations of DNA energetics was needed. My postdoc project, published in the *Journal of Molecular Biology* in 1975 (10), introduced the Tinoco group to the added insights derived from a differential scanning calorimetric mapping of the energetics of an oligomeric ribo duplex. Based on this demonstration, coupled with improvements in DNA synthesis and commercially available, high-sensitivity calorimeters, Tinoco, Crothers, and Sturtevant encouraged me initially to focus my own laboratory work on DNA energetics, as elaborated below.

During my first decade at Rutgers (1974–1984), I pursued a range of calorimetric studies on specially designed and synthesized DNA oligomers of systematically varying sequences. From these measurements, and inspired by the work of Tinoco and Crothers (5, 11, 12), I was able to derive the first direct, model-independent, calorimetrically derived DNA thermodynamic library of sequence-dependent, nearest-neighbor interactions (13)—a library that quickly became used throughout molecular biology to rationally design DNA primers and probes. During my second decade and a half at Rutgers

(1984–2000), my laboratory increased the data density of this DNA library to include many noncanonical, biologically relevant altered features of DNA forms, such as triplexes, tetraplexes, bent DNA, crosslinked DNA structures, ligand-perturbed DNA structures, and drug-bound structures. Overlapping this period up to the present, we have focused on mapping the rough energy landscapes of metastable DNA states, as well as characterizing lesion-induced alterations in DNA properties and how these perturbations correlate with biological recognition, processing, repair, mutagenesis, cytotoxicity, and other biologically relevant properties.

In my filing cabinet at Rutgers is an artifact reflecting my laboratory's growing recognition in biothermodynamics. The folder is titled *Kauzmann v. Lumry*.

They were very close friends; in fact, Rufus was Walter's best man at his wedding. However, their friendship was tested in the late 1970s, as they exchanged heated arguments about thermodynamics. They finally asked me to serve as an intermediary for their correspondence to avoid making their letters too abrasive. I was both honored and intimidated. I refrain from sharing further substance of this file, given my deep respect for both individuals and my confidential role as an impartial arbitrator! I will say, however, the experience validated me in the eyes of two giants in the field, further cementing my confidence to pursue a career dedicated to mapping the energetics of biological systems.

In this reminiscing *Reflections* article, I provide select examples of how, starting in 1974, my laboratory at Rutgers, and, previously and/or contemporaneously, the laboratories of others (e.g., Ignacio Tinoco, Don Crothers, Victor Bloomfield, Peter von Hippel, David Mathews, Doug Turner) were demonstrating that nucleic acid energy profiles, both thermodynamic and kinetic, provide the missing links between DNA and/or RNA biological structures, differential stabilities, and biological functions.

Energy profiling: The missing link

The experimental and computational mapping of energy profiles that bridge the conceptual and practical gaps between biopolymer structures and biological functions is motivating a generation of young scientists to pursue bioenergetic studies. As elaborated in this article, such profiling is justified by yielding unique energy-based insights that complement and expand elegant, structure-based understandings of biological processes.

Enhanced reporting of energy profiling, both thermodynamic and kinetic, reflects the expanding recognition that such fundamental data are essential for foundational understandings of molecular recognition, as well as for the establishment of new biomedically relevant applications. In fact, DNA energy databases have enabled a revolution in quantitative molecular biology through the rational design of primers, probes, and arrays for diagnostic, therapeutic, and molecular-profiling protocols that have contributed to a myriad of biomedical assays. The future is bright for those physical, computational, and structural chemists who also

choose to focus on characterizing the energy landscapes of complex biological assemblies.

As noted above, energetics provides the bridge between biological structure and biochemical function. International genomics and proteomics initiatives, empowered by technological and methodological advances, have produced a wealth of crucial structural and functional data (7, 9, 14). In contrast, the mapping of the energy landscapes that link these structures with their biological functions has lagged. This deficiency is manifest in and inhibitory of numerous biomedical research efforts, including, but not limited to, energy-based drug design (EBDD), hybridization-based gene regulation, mechanisms of DNA repair, differential protein stabilities and their folding/misfolding pathways, rational protein engineering, and many other examples. This article is not intended to discuss all such areas but rather to select several DNA systems that exemplify the value of energetics as the essential link between biological structure and function. To this end, this article primarily will use DNA examples from my own laboratory, while acknowledging that many other laboratories have made important contributions to building such energy bridges, both computationally and experimentally, for other biological systems.

The impact of collaborative science without borders

Over the years, my laboratory's studies have been greatly influenced and enriched by interactions with visiting luminaries, as part of distinguished lecture series, as well as by national and international collaborations with visiting scientists on sabbatical, followed by the hiring of some of their best students. To be specific, in the late 1960s to early 1970s, the Peruvian government sent some of their best young scientists to the United States for training. I was fortunate enough to have one, Dr Luis Marky, join my group and take the lead on several projects, including calorimetric determination of our original, nearest-neighbor DNA database compilation. True to the collaborative nature of my group, we were able to achieve the requisite data density and diversity for these studies by studying specially designed oligomeric and polymeric DNA samples produced and provided by a German research team led by Helmut Blöcker and Ronald Frank. Clearly, it took an international scientific village to produce the first calorimetrically determined, nearest-neighbor DNA database (13).

In this same collaborative spirit of science without borders, we pursued additional energy profiling, including drug–DNA mapping, empowered by cooperative, sabbatical-based joint ventures with Slovenian scientists, Professors Gorazd Vesnaver and Ciril Pohar, as well as Natasa Poklar, all from the University of Ljubljana; with Czech Republic scientists, Professors Bohdan Schneider and Ctirad Hofr; with German and South African scientists, Professors Horst Klump and Jens Völker; with Armenian and Russian scientists (from the former Union of Soviet Socialist Republics), Professors Armen Sarvazyan, Tigran Chalikian, and Vera Gindikina; with Chinese scientists, Dr Renzhe Jin (a Tiananmen Square refugee) and Wan-Yin Chou; with a South Korean scientist, Dr Young-Whan Park; with Polish scientists, Dr Danuta Szwajkajzer and Slawomir S.

REFLECTIONS: *The shaping of a molecular linguist*

Mielewczyk; with a Brazilian scientist, Professor Cica Minetti; as well as with United States–based scientists, including Drs. Eric Plum, Craig Gelfand, William Braunlin, and Scott Law and Professors David Remeta, Dorothy Erie, Arthur Grollman, Francis Johnson, Dani Pilch, Ron Breslow, Peter Dervan, Roger Jones, Barbara Gaffney, Gary Glick, Nick Geacintov, Doug Turner, Jerry Manning, Wilma Olson, and many others.

One of the goals of this article is to illustrate selectively the need for fundamental thermodynamic measurements on complex DNA systems, while also acknowledging the significance of kinetic factors that further shape energy landscapes; particularly, as manifest in suboptimal, metastable states that can serve as regulatory switches, a feature reflected in the work of Jens Völker, who is now a research professor in my group at Rutgers. In short, DNA energy profiling enables the establishment of essential linkages between high-profile DNA structures and crucial biological functions/pathways. By underscoring the enhanced insights derived from such DNA energy mapping, a generation of classically trained physical chemists may be motivated to apply their invaluable expertise to clarify the complexities of biological systems in terms of fundamental physiochemical principles. In so doing, they would greatly enrich the scientific arsenal of the biophysical community.

Calorimetry: The experimental method of choice

My introduction to calorimetric instrumentation was less than encouraging; in fact, it bordered on the traumatic, as I explained in a published tribute to the memory of Julian Sturtevant (1). It was late 1968/early 1969, and I was beginning my PhD studies at Yale in the Sturtevant laboratory. One evening, during my first few months there, I detected some erratic behavior by the flow calorimeter. As I wrote in the tribute to Julian (1):

It appeared that some moisture had breached the sealed submarine in which the heat sink and flow tubing were embedded. With some trepidation, I called Julian at his home, anticipating that he would tell me to stop and that he would look at it in the morning. Instead, Julian directed me to disassemble the instrument, place the heat sink in a drying oven, and then reassemble the unit in the morning. I politely listened, but I was terrorized by this directive, having never even seen the immersed calorimeter beyond the motor-driven syringes I used to deliver the reactant solutions. When I shared my fear with Dr Sturtevant, he chuckled and said ‘You should know what the insides of the instrument looks like and how it works. Only then will you understand what you really are measuring so you can interpret your data and design new experiments in an informed manner.’ I proceeded to disassemble the instrument, with the screwdriver shaking in my hand. I finally got to the heat sink, disconnected the flow tubing, and I placed the aluminum block in the drying oven. I then returned to my desk to enter all this information into my laboratory notebook. An hour later, I opened the drying oven and looked inside the heat sink. To my amazement and terror, I saw a pool of a reflective, silvery liquid that looked like melted metal. Had I just destroyed the instrument by melting the flow tubing? My terror prevented me from realizing that the temperature of the gentle drying oven was very far below the melting point of platinum or any other material used for construction of the instrument. I decided that I had to call and tell Dr Sturtevant what transpired, despite fearfully envisioning next day headlines proclaiming the shortest tenure of a graduate student in the history of Yale University. When I apologetically confessed to Dr Sturtevant what had happened, he roared with laughter and told me that the melted material was just ‘woods metal’ used to enhance thermal

contact and that everything was fine. I instantly became a Julian Sturtevant fan for life.

Given this maiden voyage with calorimetric instrumentation, it is a true testament to Julian’s sensitivity and humanity that I have made this class of instruments the cornerstone of my research career.

It has been long recognized that calorimetry represents the most direct, model- and label-free experimental approach for obtaining the energy data of interest. However, for decades, most commercially available calorimeters lacked the sensitivity required for detecting the small amount of heat accompanying biological studies. A few laboratories, most notably those of Julian Sturtevant, Peter Privalov, Rufus Lumry, and later John Brandts and Rod Biltonen built one-off, more sensitive microcalorimeters to conduct their bi-thermodynamic measurements, yet these instruments initially were not commercially available, and they lacked the throughput desired for many applications. Motivated by these limitations, specialty companies such as MicroCal, Calorimetry Sciences Corp., Tronac, Malvern Panalytical, and Thermal Analysis began to commercialize highly sensitive microcalorimeters, thereby enabling an increasing number of laboratories capable of performing bio-level measurements. The next technical hurdle was to increase throughput, a challenge that has been surmounted *via* automation and multicell configurations, thereby making the methodology far more appealing to drug discovery/material science screening efforts.

Armed with this new generation of microcalorimeters, one now has the ability to map the subtleties of energy landscapes, including metastable states and conformational switches that are at the foundation of regulatory events in biology and thus are essential for the rational design of physiochemical properties in material sciences. Physical chemists know that energetics represent the universal language of molecular recognition. Likewise, biophysical chemists know that energetics are required to bridge the conceptual and practical gaps between structure and function. Through the ongoing efforts of both overlapping communities, energy databases are beginning to become sufficiently populated to allow for fundamental and predictive correlations to be established between biophysical properties and biological function. What follows is a sampling of examples that illustrate the need for and enhanced insights produced by energy profiling of DNA biological systems.

DNA thermodynamic databases: Rational design of primers and probes

PCR is a method that allows one to make billions of copies of originally minute amounts of target DNA samples of interest (15). This amplifying technology revolutionized a myriad of DNA-based applications, including molecular diagnostics and therapeutics, forensic molecular fingerprinting, DNA cloning, and mapping of DNA-based phylogenies, while enabling multiple other fundamentally important areas of inquiry. A key reagent in PCR is known as a primer. Primers are

relatively short pieces of single-stranded DNA that are Watson–Crick complementary to a target region(s) on one strand of the DNA to be amplified. The resultant hybridized primer–DNA target strand mini-duplex, with the overhanging single strand, then serves as a template for a DNA polymerase enzyme to produce a DNA copy. As this chain reaction cycles, it yields billions of copies of the original minute DNA sample. To optimize the error-free fidelity of this amplification technique requires a rational basis for the primer design that goes beyond simple Watson–Crick hybridization and which includes defining PCR conditions that maximize stringency. To this end, the need for sequence and solution-dependent thermodynamic data became self-evident. The requisite nearest-neighbor thermodynamic databases for Watson–Crick domains (13, 16–23) were determined by multiple laboratories, including those of Tinoco (5, 11), Crothers (12), Turner and Mathews (6, 24, 25), as well as my own, and they continue to be improved upon. The resultant data, including more recently reported heat capacity effects (26–31), have been incorporated into commercial and proprietary algorithms used worldwide to enable differential stability-based design of hybridization-based diagnostic and therapeutic protocols (32–36).

These energy databases also characterize DNA structures that contain biologically consequential mismatches (37, 38), modified bases/lesions (37, 39–47), as well as nonduplex secondary structures, including the energetics of triplexes (48–52), tetraplexes (53–56) (associated with modulating telomere stability and epigenetic modifications) (57, 58), and higher order, functionally relevant DNA assemblies (59–64). My laboratory's triplex DNA studies were performed in collaboration with Peter Dervan's group, whom I knew well from our time together as graduate students at Yale. He sent one of his students, Scott Singleton, to my laboratory at Rutgers to learn calorimetric techniques for measuring the energetics of triplex formation, an assessment he needed to tune the stringency of his designed, groove-binding synthetic endonucleases (49). My laboratory's subsequent focus on mapping the energetics of tetraplex complexes (53, 54) was stimulated by Alex Rich meeting with me and my group as part of his seminar visit to Rutgers. By the end of the day, he had convinced most members of my group to drop their projects and pursue energy mapping of tetraplexes and Z-DNA! Likewise, a seminar visit from Peter Schultz resulted in a postdoc offer for a graduate student in my laboratory, accompanied by lobbying me to allow my student to complete his thesis expeditiously so he could get to Berkeley as soon as possible. In 1991, the Merck lecture at Rutgers was given by Ron Breslow of Columbia University. In a separate meeting with me and my group during his visit, Ron asked if any of us could explain why DNA "evolved" to use 5'-3' rather than 5'-2' phosphodiester linkages. Without waiting for our speculation, he proclaimed that he had developed a synthetic pathway for the 5'-2' links and wanted us to conduct a comparative thermodynamic profile of it with the base sequence homologous 5'-3' counterpart. I agreed that it would be a very interesting comparison. Ron bellowed, "Do you think the world is ready for a Breslow–Breslauer Proceedings of the National Academy of Sciences

paper?" I responded, "Probably not, but I am confident that they are ready for a Breslauer–Breslow paper." He roared with laughter, we performed the work, and we published the fascinating results in Proceedings of the National Academy of Sciences in 1993 (65). We scored another victory for collaborative science.

To date, the most important applications of the expanding DNA databases include the following: the differential stability-based rational design of nucleic acid primers and probes with enhanced stringency that minimizes false positives and false negatives, in both clinical and forensic assays (66, 67); the rational modification and conversion of oligonucleotides into sequence-specific antisense and antigene agents, including synthetic endonucleases (68–75); and the rational design/configuration of DNA oligonucleotide arrays on high-throughput chips for molecular profiling (76, 77). In sum, these DNA energy databases have helped revolutionize quantitative molecular biology through the energy-based design of primers, probes, and arrays for diagnostic, therapeutic, and molecular-profiling protocols that have contributed to the rational design of a myriad of biomedical assays.

"Energy recognition": isostructural does not necessarily mean isoenergetic

Throughout the 1980s and 1990s, my laboratory was part of a National Cancer Institute–funded program with Arthur Grollman's group at Stony Brook University, on DNA recognition, repair, and mutagenesis. A central component of our project focused on comparing the properties of oligomeric DNA duplexes both with and without lesion sites. A consistent observation from our studies revealed that DNA lesions can have significant consequences on DNA energetics, even in the absence of detectable levels of lesion-induced structural/conformational alterations. In other words, *isostructural does not necessarily imply isoenergetic*. Two examples of this phenomenon are illustrated by the data in Figure 1 (42, 44). Note that the canonical and corresponding lesion-containing duplexes (with either an 8-oxo-dG or abasic "damaged" site) exhibit indistinguishable duplex forms, insofar as the global conformations as defined by CD, NMR, and/or X-ray measurements are concerned, yet they manifest substantial differences in their energy profiles.

The examples in Figure 1 underscore the need for comparative structural studies in parallel to the corresponding energy characterizations. Furthermore, the magnitudes of these structurally/conformationally "silent," lesion-induced energy perturbations were consistent with their destabilizing impact beyond the lesion site to create domains of genome instability (37, 39, 40, 42, 78). We proposed that such energetically perturbed domains may serve as initial DNA recognition elements for the machinery of repair, a phenomenon we referred to as *energetic recognition* (79). Such initial sensing of an energy-perturbed domain, as the processing repair enzyme surveils DNA for damage, is envisioned to occur before the more localized, site-specific protein interactions with distinct structural features of the lesion. Consequently, mapping energy landscapes of lesion-containing DNA substrates,

REFLECTIONS: The shaping of a molecular linguist

The 8-oxo-dG Lesion (**G***) Alters Duplex Energetics Absent Structural Perturbations

<u>Canonical DNA Duplex</u>		<u>Damaged DNA</u>
CGCATGCGTACGC	versus	CGCATGCGTACGC
GCGTACGCATGCG		GCGTACG*CATGCG
<u>GC Duplex</u>		<u>G*C Duplex</u>

Structurally—“identical”: NMR, CD, X-ray characterize these two duplexes as being “equivalent”

Thermodynamically—different

$$\begin{aligned}\Delta\Delta H &= -7.2 \text{ kcal/mol} \\ \Delta\Delta G_{25^\circ\text{C}} &= -1.3 \text{ kcal/mol}\end{aligned}$$

The Abasic Lesion (**F**) Alters Duplex Energetics Absent Structural Perturbations

<u>Canonical DNA Duplex</u>		<u>Damaged DNA Duplex</u>
CGCATGAGTACGC	versus	CGCATGAGTACGC
GCGTACTCATGCG		GCGTACFCATGCG
<u>AT Duplex</u>		<u>AF Duplex</u>

Structurally—“identical”: NMR and CD characterize these two duplexes as being “equivalent”

Thermodynamically—very different

$$\begin{aligned}\Delta\Delta H &= -28.7 \text{ kcal/mol} \\ \Delta\Delta G_{25^\circ\text{C}} &= -7.6 \text{ kcal/mol}\end{aligned}$$

Figure 1. Examples of isostructural does not necessarily mean isoenergetic.

and their protein-binding properties, is essential for understanding lesion recognition, repair, and translesion DNA synthesis—including the failure of these pathways that leads to mutagenesis. Knowledge of energy profiles/mechanistic pathways of repair (e.g., base excision repair, or BER) is required for the rational design of intervention strategies that target specific steps/species/checkpoints encountered in such processing pathways, such as abasic sites, damaged/modified bases, altered secondary structures, bulges, and mismatches.

Repair enzyme differential binding affinities: Preferential binding for product over substrate

A logical next step in our contribution to the program was to probe differential repair enzyme binding to duplexes with the lesion and with the lesion excised by a DNA glycosylase. Before our promulgation of the concept of energetic recognition, it was broadly assumed that an enzyme would exhibit the greatest affinity for its nascent substrate. However, our energy studies revealed, counter to this expectation, that an endonuclease can bind more strongly to the highly destabilizing abasic site product generated by lesion processing, in contrast with its relatively weak binding to the only marginally destabilizing 8-oxo guanine nascent substrate target (80–82). Figure 2 illustrates the relevant data for one system (79).

Differential Binding (at 5°C) of Repair Enzyme (Fpg) to 8-oxo-dG “Substrate” versus Abasic Site “Product”

DNA TARGET	$\Delta\Delta G \text{ kcal}\cdot\text{mol}^{-1}$	Binding ΔG_{REL}
8-oxo-dG	-2.0 ± 0.7	< 0.01
Abasic	-8.1 ± 1.4	1

Figure 2. Example of preferential binding of a processing enzyme for the product over substrate.

This unexpected differential affinity required energy mapping to discern and would not have been evident from structural studies alone. Such an energy-based insight supports a novel hypothesis that enzymatic excision of a damaged base can be driven by a repair enzyme binding more strongly to the product than to the substrate. One can view this differential affinity as another example of *energy recognition*, as the product-containing abasic site is far more destabilizing than the substrate-containing 8-oxo guanine site. This insight is consistent with the hypothesis that lesions can induce domains of altered stability that may serve as recognition sites for repair proteins (46).

In summary, under nonturnover conditions, a DNA glycosylase can bind more strongly to a highly destabilized duplex containing the base excision-derived abasic product analog than to the less-destabilized duplex containing the nascent lesion substrate analog (79, 81). This finding suggests that protein–DNA interactions during catalysis can favor transition state/product binding/stabilization over substrate binding/stabilization, thereby driving the catalytic event. The preferred binding to the abasic product/intermediate also serves to sequester mutagenic abasic sites before subsequent repair steps. Beyond providing insights into protein–DNA binding and catalysis, such *energetic recognition* has implications in the rational design of inhibitors and antagonists of DNA-processing pathways.

Energetics as signatures of differential biological consequences

The connectivity between DNA thermodynamic studies and carcinogenesis is reflected in the energy profiles/signatures of DNA mutagenic lesions and their recognition and processing by the cellular repair machinery (80, 83–87). Secondly, these studies also reinforce the concept of *energy recognition* in which DNA-repair enzymes sense regions of lesion-perturbed DNA stability before contacting structurally localized lesion sites (39, 42, 46). This concept is consistent with the hypothesis that nearly all lesions induce energy penalties that propagate several base pairs beyond the 5' and 3' sides of the lesion. For the

processing repair enzyme, this propagated perturbation creates recognizable islands/domains of instability distal from the lesion site, before the more direct and structurally dictated encounter with the localized lesion site.

Here is an additional example of the enhanced insights derived from thermodynamic characterizations: it has been postulated that perturbed energetics of damaged duplexes serve as predictors of the biological consequences exhibited by the DNA lesion (*e.g.*, mutagenicity *versus* cytotoxicity) (88, 89). Specifically, the replication repair machinery engages template-primer substrates and may either bypass/misread a damaged base, such as the mutagen 8-oxo-dG (42, 79, 90), or stall at a potentially cytotoxic site, such as thymine glycol (40, 91). As the requisite energy data density and diversity increases, more biological outcomes should be predictable based on the energetic signatures of the DNA lesions embedded in their respective host duplexes. Again, in these measurements, parallel structural studies reveal that *isostructural does not always imply isoenergetic*, a fundamental insight that requires energy profiling to avoid missing DNA lesion-induced energy perturbations that may be structurally “silent.” This maxim has been demonstrated to be true in both nucleic acid and protein systems. Taken together, these results enable energy- and product-based design of exogenous agents that can modulate DNA repair, a major pathway that is dysregulated in cancer initiation and progression. Of course, it also is true that *isoenergetic does not necessarily imply isostructural*, an observation that may reflect significant enthalpy–entropy compensations that compress free energy distinctions. In such cases, one detects the energy differences in one or more of the other thermodynamic parameters.

There are many other examples of energetics as putative signatures of differential biological consequences. For instance, frameshift mutagenesis frequency correlates with the differential stabilities of bulge-looped duplex structures that serve as intermediates for insertion and deletion mutations (92, 93). Consequently, energy profiling of single-base bulged structures (37) is of direct biological relevance, while also providing an energy database for predicting potential hotspots in the genome. It has been postulated that the differential energy profiles of lesions as a function of the lesion type, sequence context, and counterbase correlate with repair success and failure (39, 43, 44, 94), thus providing a basis for rationalizing differential mutagenicity frequencies. Further energy profiling of lesion-containing DNA constructs is required to more robustly test this hypothesis.

Energy profiling of DNA recognition, repair, and replication

DNA recognition, repair, and replication (the three R’s of DNA) represent interrelated DNA processing events of central biomedical significance, particularly because errors in any of these three steps can result in mutational events that might trigger physiological disorders. In advocating for parallel energy and structural studies, I previously have emphasized that DNA lesions can significantly alter DNA energetics, even in

the absence of detectable levels of lesion-induced structural/conformational perturbations. Consequently, mapping energy landscapes of lesion-containing DNA substrates, and their protein-binding properties, is essential for understanding lesion recognition, repair, and translesion synthesis—including the failure of these pathways, which can lead to mutagenesis.

In this context, it has been proposed, and in some cases demonstrated, that the differential stabilities of lesions as a function of their opposite bases dictate and control error-free *versus* error-prone translesion synthesis (95). As a result, energy profiling of lesion-induced destabilizations as a function of the opposite base relative to the corresponding undamaged pair/mispair (46, 96) is of direct biological relevance for predicting relative insertion frequencies, as well as predicting lesion-mediated mutagenicity *versus* adduct-induced genotoxicity. Knowledge of energy profiles/mechanistic pathways of repair (*e.g.*, BER) also is required to guide the design of intervention strategies to target specific steps/species/checkpoints encountered in such processing pathways. Intriguingly, under the leadership of Research Professors David Remeta, Cica Minetti, Eric Plum, and Craig Gelfand, we have demonstrated a correlation between polymerase-mediated insertion/extension enthalpies (80) and the corresponding nearest-neighbor values in DNA databases (13, 16–23), as illustrated in Figure 3.

Thus, despite the making and breaking of covalent bonds during the replication process, due to large compensatory effects, only modest noncovalent nearest-neighbor forces control replication efficiency and fidelity in a predictable sequence-dependent manner. In short, base insertion discriminations are dictated by modest energy differences in noncovalent interactions at the template-primer level that correlate with the corresponding nearest-neighbor data. In addition to providing the framework for the design of sequencing platforms that employ thermoelectric high-throughput microfluidics initiatives (96), such data also set the stage for studies of sequence-dependent energetics for polymerase-template DNA interactions and catalysis, both during normal replication and during translesion synthesis. The resultant energy profiles allow predictions of the relative probabilities of specific base insertion as a function of the polymerase and the canonical base or lesion in the template.

DNA-repair pathways, energetics, and cancer

There is an emerging literature on DNA-repair pathways as targets for cancer therapy (97–99), which underscores how mechanistic/energetic mapping of such pathways provide the basis for developing such therapy (86, 100) and cancer treatment protocols (101) [as reviewed (102)]. In fact, DNA-targeting drugs represent a large proportion of the available anticancer pharmacopeia (103). DNA-interacting molecules share the ability of certain DNA-binding proteins to perturb both DNA structure and energetics, some of which display helix-destabilizing properties. The formation of such locally destabilized DNA domains is likely to interfere with

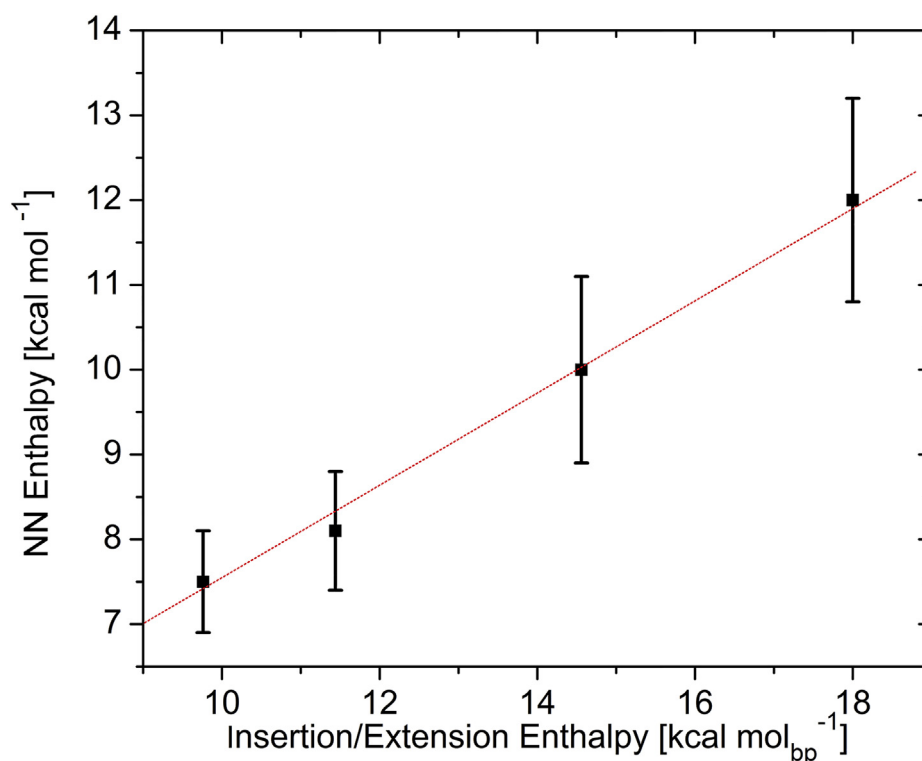


Figure 3. Average nearest-neighbor enthalpies versus base insertion/extension enthalpies.

protein–DNA recognition, thereby potentially affecting DNA repair, DNA replication, and transcription (103). This property may be exploited in antitumor therapeutics. To enable such strategies, energetic profiles are required to map oxidative, exocyclic, and drug-induced DNA lesion recognition/repair pathways, as a prerequisite for defining the appropriate targets for modulating biological outcomes. This work in my laboratory, originally inspired by interactions with Professors Dinshaw Patel and Arthur Grollman, has been, and continues to be, led and conducted by Remeta and Minetti at Rutgers.

Existing thermodynamic data suggest that sequence-dependent energetic signatures of DNA lesions may serve as diagnostic probes of promutagenic *versus* cytotoxic propensities of specific adducts, the latter feature being a desirable property in some anticancer interventions. Furthermore, as a probe of miscoding and/or lesion-blocking properties, energy data help define the basis for mechanisms of chemical-induced carcinogenesis (104) and/or radiation-induced cell death (105). Other applications of energy data include delineating the origins for failure in anticancer therapeutic interventions (*e.g.*, chemical- or radiation-induced DNA adduct formation and consequent damage persistence) (106, 107), as well as mechanisms associated with metabolic imbalances and age-dependent neurodegenerative processes (108, 109).

Sequence context–dependent lesion energetics are also valuable for identifying and defining the origins of biologically consequential hotspots within the genome. Differential energy profiles may provide a metric for risk assessment of

environmental exposures or metabolic anomalies. Calorimetric evaluations of the requisite energetics associated with template-directed polymerase-mediated synthesis at single base resolution (80) already has inspired investigators within the chemical engineering field to design high-throughput microfluidics DNA-sequencing methods and instrumentation (96). The more complete mapping of differential energetic profiles for repair and replication pathways and their error-free and error-prone consequences remain to be determined.

Collectively, the above examples illustrate that additional energy profiling is required to more robustly bridge the gap between microscopic structural insights and macroscopic/cellular biological function/mechanisms. These energy profiles map the nature and relative strengths of the forces that regulate and control biological functions. Such characterizations create databases that can be mined to establish unique macroscopic signatures/correlations for biologically relevant properties (*e.g.*, mutagenicity *versus* genotoxicity) and, more generally, for energetically profiling successful and unsuccessful recognition, repair, and replication events. Failure in any of these events can result in mutations and, ultimately, cancers.

Energy profiling of DNA triplet repeats: Rough DNA energy landscapes and dynamics, metastable substrate ensembles

Another example of connectivity between fundamental energy studies and biological structure and function is the

determination of rough energy landscapes of metastably populated triplet repeat DNA systems. Such repeats are associated with families of neurodegenerative complex diseases, such as Huntington's disease, fragile X syndrome, and so forth (110–114). My laboratory has taken an interest in this topic, primarily because of the recruitment of Dr Jens Völker from Germany, by way of Professor Horst Klump's laboratory at the University of Cape Town in South Africa. Rather than focusing exclusively on the most stable states, Dr Völker developed calorimetric methods and designed DNA systems that would probe metastability/suboptimal states, a feature central to biological regulatory switches. His first interest focused on mapping the rough energy landscape of DNA triplet conformational gymnastics. To this end, he conducted calorimetric and spectroscopic studies on strategically designed and synthesized families of long oligonucleotides with different triplet repeat lengths and disease-inducing sequence repeats, flanked by duplex domains (59, 115–117), as illustrated in Figure 4.

These measurements revealed that triplet repeats are extruded out of the duplex as bulge loops and can traverse along the helix to form a dynamic ensemble of metastable “rollamers” (115, 118), whose properties depend on the size and sequence (59) (Fig. 4).

Beyond the intrinsic value of such studies, this family of unique DNA constructs provides an opportunity to map the interplay between kinetic and thermodynamic factors that collectively shape the rough energy landscapes of these molecules (119) as well as their recognition (120). Especially revealing is the dynamic nature of the DNA substrate, which differentially impacts the protein processing of such substrate ensembles. As such, beyond the thermodynamically most stable substrates, energy profiling also must consider kinetic factors that can trap metastable species of biological relevance.

By strategically inserting fluorescent probes and modifying mutagenic lesions at selective sites within the loops and the adjacent duplex domains, it was possible to probe and map intriguing crosstalk between DNA domains/regions (115, 121). These biologically consequential results required energy profiling and would not have been gleaned from structural studies alone. Significantly, it was revealed that rollamers with lesions within the loop were more stable than those with lesions in the duplex domains, thereby

favoring the former family of “rollamers” within the ensemble.

Such *dynamic metastable substrate ensembles* (115, 118) present a challenge to the repair and transcription machinery that evolved to process fewer dynamic substrates (122–124). To this point, it has been demonstrated that CAG repeat bulge loops trap nonproductively a critical mismatch repair complex, MSH2–MSH3 (124, 125), which in turn stimulates components of the BER machinery, not normally associated with mismatch repair, causing these repeat domains to expand (126). Results from our laboratory provided a biophysical rationale for intriguing reports that activation of DNA repair induces expansion and pathogenicity of proximal DNA repeat domains, thereby yielding a basis for the rationale design of strategies for inhibiting triplet expansion associated with developmental disorders.

Energy-based drug design

The concept of EBDD is an approach that further empowers and complements the paradigm of experimental and computational *structure-based drug design*. The missing structure/energy synergy was emphasized by Ray Salemme's quote in the cover story of Chemical & Engineering News (127), stating that “The initial expectation of structure-based design—that you were going to design molecules and they would work right out of the box—was unrealistic. We didn't understand the thermodynamics well enough.”

In other words, as cartooned in Figure 5, just because a drug fits topologically into its biological receptor, it does not mean that it functionally “sticks.” The latter assessment requires thermodynamic and kinetic binding profiles.

A number of laboratories have produced the binding energetics for a range of drugs that target protein and/or nucleic acid “receptors,” parsing the contributions of specific interactions by characterizing calorimetrically homologous series of ligands (128–132). Many of these studies further revealed that binding affinities alone do not provide the requisite insights for rational drug design needed to enhance lead optimization. Complete binding profiles (ΔG , ΔH , ΔS , ΔC_p), preferably determined calorimetrically, are needed to elucidate the nature of the driving and/or inhibitory forces, including binding-induced and coupled changes in conformation and/or solvation (131–133). Changes in solvation have

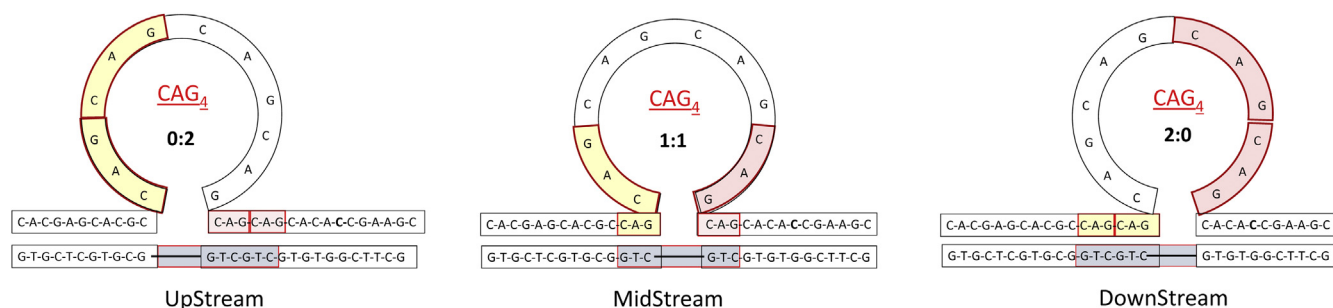


Figure 4. Schematic of DNA triplet repeats being extruded out of a duplex as bulge loops that traverse along the helix to form a dynamic ensemble of metastable “rollamers” (115, 118), whose properties depend on the size and sequence (59).

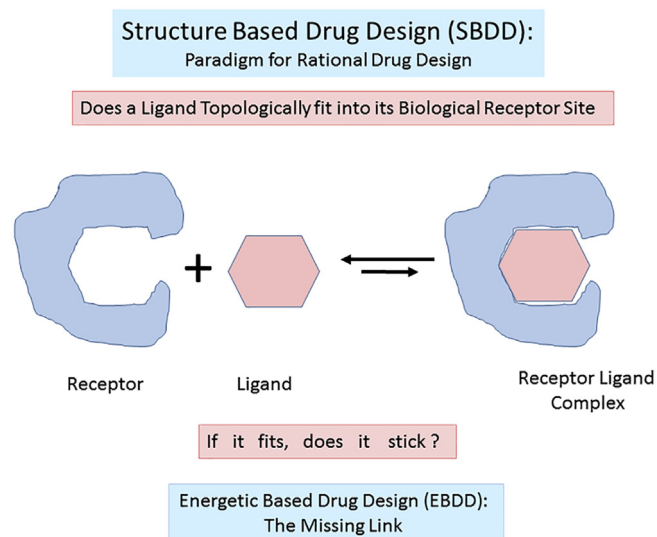


Figure 5. Schematic of structure-based drug design complemented by energetic-based drug design.

long been acknowledged as important features that strongly influence drug binding and target stability. However, not all thermodynamically perturbed solvent is sufficiently ordered so as to be detectable by structural techniques, ergo the need for parallel energy measurements that relate to alterations in solvation (134, 135), particularly heat capacity and volumetric assessments.

Fortunately, Armen Sarvazyan and Tigran Chalikian had joined my group from Armenia (in the former Union of Soviet Socialist Republics). They had pioneered the use of acoustic and densitometric techniques to define the solvation/desolvation impact on ligand affinity and target stability (136–141). Their volumetric profiles provided unique insights into the impact of solvation changes as significant modulators of binding energetics (129, 142, 143). These data also provide an empirical basis for refining computational tools that predict and characterize binding energetics and dynamics with explicit incorporation of solvent contributions. Significantly, these collective energy characterizations correlate with the success or failure of a designed ligand to target its site and to express its biological activity (144–146).

In short, the enhanced power of parallel structure-based drug design and EBDD is now widely accepted by the scientific community. EBDD is now on an equal footing with the more traditional *structure-based drug design*, thereby enabling more effective drug discovery/screening protocols to help drive the development of new drugs. The importance of this reality is reflected in the large number of pharmaceutical and biotech companies that have incorporated isothermal titration and differential scanning calorimeters into their drug design and screening protocols (147–150).

Energy profiling of the genetic code: Implications for code evolution and molecular Darwinism

For many years, Klump, Völker, and I had been pondering how one could link fundamental DNA energetics with the

shaping of the genetic code as an energy code (an *energy genome map*), ultimately correlating the resultant sequence-dependent genomic energy profiles with evolutionary trends. A very recent article from my laboratory addresses these interrelationships, revealing empirical correlations between codon energetics, genetic code evolution, and the concept of molecular Darwinism. In it, we demonstrate that one can employ a thermodynamic framework to rationalize the origins and evolution of the genetic code, and, more broadly, the genomic energy landscape (151). Specifically, we showed that differential DNA energy stability profiles correlate with differential biological outcomes. This observation is suggestive of the code functioning as an energy ladder comprised of linked thermodynamic cycles that evolved *via* sequential transitions and transversions, which selectively were driven by differential energy factors. In this context, we also propose that, in addition to the traditional Darwinian perspective of evolution, which is driven by natural selection due to phenotypic advantages, a form of “molecular Darwinism” also is at work, in which evolutionary pressures derive from energy optimization of a biophysical property. These recent results provide additional, more global examples of the unique insights revealed by fundamental energy profiling of complex biological processes.

To be specific, using nearest-neighbor DNA stability data, we translated the iconic chemical genetic code—the blueprint of life—into an energy code, revealing unique insights into two major universal enigmas in the life sciences. First, the energy mapping of the code explained how, out of an astronomical number of possible blueprints of life, the DNA genetic code evolved *via thermodynamic selection* into a nearly identical, singular blueprint for all living organisms.

Second, as alluded to above, this energy mapping of the genetic code expands the underpinnings of Darwin’s natural selection/survival-of-the-fittest theory of evolution to include an additional component of molecular Darwinism. This concept of molecular Darwinism posits that some species characteristics evolutionarily persist across generations because their DNA energy codes are unusually stable, independent of whether or not the characteristics provide a survival advantage in a given environment, as proposed in the classic Darwinian evolutionary theory. Going forward, such energy mapping of the DNA chemical genome to create an *energy genome* will produce an enhanced roadmap for rationally designing DNA molecular targets that enable new diagnostic and therapeutic protocols.

Concluding remarks

Mapping of energy profiles bridges the conceptual and practical gaps between biopolymer structures and biological functions. As elaborated in this *Reflections* article, such energy profiling for DNA is justified by unique energy-based insights that complement and expand elegant, structure-based understandings of DNA biological processes. Enhanced reporting of DNA energy profiling reflects the expanding recognition that such fundamental data are essential for foundational understanding of molecular recognition—and for the establishment of new biomedical insights and applications. DNA

energy databases have enabled a revolution in quantitative molecular biology through the rational design of primers, probes, and arrays for diagnostic, therapeutic, and molecular-profiling protocols that have contributed to a myriad of biomedical assays that collectively impact human health.

I am very pleased to see how the foundational DNA energy data produced by my laboratory over nearly 5 decades have been, and continue to be, used for both fundamental understandings of molecular communication and for biomedically relevant, practical applications. Going forward, I am further encouraged by our DNA energy profiles/landscapes being used to empower the interpretations of complex new datasets produced by a myriad of emerging methodologies. The advent of increasingly sophisticated structural techniques (e.g., single particle cryo-EM), some of which even provide time-resolved motion picture mapping of the dynamics associated with cellular processes (cryo-EM tomography), require enhanced computational methods to interpret the volumes of new primary structural/dynamic/functional data they yield. These computational methods are empowered when calibrated with experimental energetics such as those my laboratory is producing. Thus, aside from the intrinsic and application-based value of characterizing the intermolecular and intramolecular forces that modulate relative stabilities and molecular communications, calorimetry is having an increasing role in enabling the intensive computational methods required to interpret the rich information embedded within the primary data produced by these new methodologies (131).

Beyond enabling computational methods, we also now are focusing our energy profiling on complex, multicomponent biological machines—preferably machines in action, such as repair complexes or polymerase activities. These studies already include determining the comprehensive landscapes of complex dynamic systems, inclusive of kinetic effects, that modulate regulatory conformational switches (117). The methodologies required to accomplish such goals are at, or beyond, the limits of what one can do with existing instrumentation and analyses tools. Our pursuit of these goals will require the incorporation of methods in addition to calorimetry, as well as the development of different analysis methods to do justice to the complexities of such biological machines. Yet, the potential insights one can derive from comprehensive energy landscapes on complex biological machines more than justify the effort.

It is my hope that this *Reflections* article illustrates the personal and professional influences that shaped and motivated my five decades of contributions toward establishing *DNA Energy Profiling* as the co-equal sibling of *DNA Biological Structure and Biological Function*. As a “molecular linguist,” it has been and remains my fervent desire to establish DNA energy databases that characterize the words, sentences, paragraphs, and chapters that collectively constitute the language used in the book of molecular communication, thereby enabling the elucidation of the forces that dictate and control fundamental and complex biological processes. In the year

2071, I look forward to sharing with my colleagues the contributions from my second half-century of work!

Acknowledgments—Kenneth J. Breslauer wishes to acknowledge the outstanding scientists in his laboratory who, over the years, have designed and expertly executed many of the studies used as examples in this *Reflections* article. Specifically noteworthy, in alphabetical order, are Drs. Tigran Chalikian, Wan-Yin Chou, Dorothy Erie, Craig Gelfand, Vera Gindikina, Renzhe Jin, Scott Law, Yi Liu, Luis Marky, Cica Minetti, Young-Whan Park, Dani Pilch, G. Eric Plum, David P. Remeta, Armen Sarvazian, Mary Senior, Danuta Szwajkajzer, and Jens Völker, as well as my collaborating colleagues, Professors Arthur Grollman, Francis Johnson, Roger Jones, Horst Klump, and Gorazd Vesnaver.

Funding and additional information—This work was supported by grants from the NIH (United States) to K.J.B.: General Medicine 23509 and 34469 and NIH (United States): Cancer 47995. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflict of interest—The author declares that he has no conflicts of interest with the contents of this article.

References

- Breslauer, K. J. (2007) Julian Sturtevant: Scientific giant, warm humanist, social activist, nature lover, gentle teacher, kind friend. *Biophys. Chem.* **126**, 9–10
- Tinoco, I. (2014) Fun and games in Berkeley: The early years (1956–2013). *Annu. Rev. Biophys.* **43**, 1–17
- Tinoco, I., Jr. (2015) The ethical scientist: An old-fashioned view. *Biopolymers* **103**, 424–431
- Bloomfield, V. A., Victor, B., Crothers, D. M., Tinoco, I., and Crothers, D. (1974) *Physical Chemistry of Nucleic Acids*, Harper & Row, New York, NY
- Tinoco, I., Jr., Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M., and Gralla, J. (1973) Improved estimation of secondary structure in ribonucleic acids. *Nat. New Biol.* **246**, 40–41
- Turner, D. H., Sugimoto, N., and Freier, S. M. (1988) RNA structure prediction. *Annu. Rev. Biophys. Chem.* **17**, 167–192
- Berman, H. M., Burley, S. K., Kleywegt, G. J., Markley, J. L., Nakamura, H., and Velankar, S. (2016) The archiving and dissemination of biological structure data. *Curr. Opin. Struct. Biol.* **40**, 17–22
- Young, J. Y., Westbrook, J. D., Feng, Z., Peisach, E., Persikova, I., Sala, R., Sen, S., Berrisford, J. M., Swaminathan, G. J., Oldfield, T. J., Gutmanas, A., Igarashi, R., Armstrong, D. R., Baskaran, K., Chen, L., et al. (2018) Worldwide Protein Data Bank biocuration supporting open access to high-quality 3D structural biology data. *Database (Oxford)* **2018**, bay002
- Burley, S. K., Berman, H. M., Bhikadiya, C., Bi, C., Chen, L., Di Costanzo, L., Christie, C., Dalenberg, K., Duarte, J. M., Dutta, S., Feng, Z., Ghosh, S., Goodsell, D. S., Green, R. K., Guranovic, V., et al. (2019) RCSB protein Data Bank: Biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy. *Nucleic Acids Res.* **47**, D464–D474
- Breslauer, K. J., Sturtevant, J. M., and Tinoco, I., Jr. (1975) Calorimetric and spectroscopic investigation of the helix-to-coil transition of a ribooligonucleotide: rA7U7. *J. Mol. Biol.* **99**, 549–565
- Aboul-ela, F., Koh, D., Tinoco, I., Jr., and Martin, F. H. (1985) Base-base mismatches. Thermodynamics of double helix formation for dCA3XA3G + dCT3YT3G (X, Y = A, C, G, D). *Nucleic Acids Res.* **13**, 4811–4824
- Gralla, J., and Crothers, D. M. (1973) Free energy of imperfect nucleic acid helices. *J. Mol. Biol.* **78**, 301–319

13. Breslauer, K. J., Frank, R., Blocker, H., and Marky, L. A. (1986) Predicting DNA duplex stability from the base sequence. *Proc. Natl. Acad. Sci. U. S. A.* **83**, 3746–3750
14. Goodsell, D. S., Burley, S. K., and Berman, H. M. (2013) Revealing structural views of biology. *Biopolymers* **99**, 817–824
15. Mullis, K. B. (1994) The polymerase chain reaction (Nobel lecture). *Angew. Chem. Int. Ed. Engl.* **33**, 1209–1213
16. SantaLucia, J. (1998) A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 1460–1465
17. Huguet, J. M., Bizarro, C. V., Forns, N., Smith, S. B., Bustamante, C., and Ritort, F. (2010) Single-molecule derivation of salt dependent base-pair free energies in DNA. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 15431–15436
18. Owczarzy, R., Vallone, P. M., Gallo, F. J., Paner, T. M., Lane, M. J., and Benight, A. S. (1997) Predicting sequence-dependent melting stability of short duplex DNA oligomers. *Biopolymers* **44**, 217–239
19. Doktycz, M. J., Goldstein, R. F., Paner, T. M., Gallo, F. J., and Benight, A. S. (1992) Studies of DNA dumbbells. I. Melting curves of 17 DNA dumbbells with different duplex stem sequences linked by T4 endloops: Evaluation of the nearest-neighbor stacking interactions in DNA. *Biopolymers* **32**, 849–864
20. Gotoh, O., and Tagashira, Y. (1981) Locations of frequently opening regions on natural DNAs and their relation to functional loci. *Biopolymers* **20**, 1043–1058
21. Delcourt, S. G., and Blake, R. D. (1991) Stacking energies in DNA. *J. Biol. Chem.* **266**, 15160–15169
22. Vologodskii, A. V., Amirikyan, B. R., Lyubchenko, Y. L., and Frank-Kamenetskii, M. D. (1984) Allowance for heterogeneous stacking in the DNA helix-coil transition theory. *J. Biomol. Struct. Dyn.* **2**, 131–148
23. Ghosh, S., Takahashi, S., Ohyama, T., Endoh, T., Tateishi-Karimata, H., and Sugimoto, N. (2020) Nearest-neighbor parameters for predicting DNA duplex stability in diverse molecular crowding conditions. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 14194–14201
24. Turner, D. H., and Mathews, D. H. (2010) Nndb: The nearest neighbor parameter database for predicting stability of nucleic acid secondary structure. *Nucleic Acids Res.* **38**, D280–D282
25. Spasic, A., Berger, K. D., Chen, J. L., Seetin, M. G., Turner, D. H., and Mathews, D. H. (2018) Improving RNA nearest neighbor parameters for helices by going beyond the two-state model. *Nucleic Acids Res.* **46**, 4883–4892
26. Chalikian, T. V., Völker, J., Plum, G. E., and Breslauer, K. J. (1999) A more unified picture for the thermodynamics of nucleic acid duplex melting: A characterization by calorimetric and volumetric techniques. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 7853–7858
27. Rouzina, I., and Bloomfield, V. A. (1999) Heat capacity effects on the melting of DNA. 1. General aspects. *Biophys. J.* **77**, 3242–3251
28. Rouzina, I., and Bloomfield, V. A. (1999) Heat capacity effects on the melting of DNA. 2. Analysis of nearest-neighbor base pair effects. *Biophys. J.* **77**, 3252–3255
29. Holbrook, J. A., Capp, M. W., Saecker, R. M., and Record, M. T.,J. (1999) Enthalpy and heat capacity changes for formation of an oligomeric DNA duplex: Interpretation in terms of coupled processes of formation and association of single-stranded helices. *Biochemistry (N.Y.)*. **38**, 8409–8422
30. Jelesarov, I., Crane-Robinson, C., and Privalov, P. L. (1999) The energetics of HMG box interactions with DNA: Thermodynamic description of the target DNA duplexes. *J. Mol. Biol.* **294**, 981–995
31. Völker, J., Plum, G. E., and Breslauer, K. J. (2020) Heat capacity changes (ΔC_p) for interconversions between differentially-ordered DNA states within physiological temperature domains: Implications for biological regulatory switches. *J. Phys. Chem. B.* **124**, 5614–5625
32. Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31**, 3406–3415
33. Markham, N. R., and Zuker, M. (2008) UNAFold: Software for nucleic acid folding and hybridization. *Methods Mol. Biol.* **453**, 3–31
34. Kibbe, W. A. (2007) OligoCalc: An online oligonucleotide properties calculator. *Nucleic Acids Res.* **35**, 43
35. Panjkovich, A., and Melo, F. (2005) Comparison of different melting temperature calculation methods for short DNA sequences. *Bioinformatics* **21**, 711–722
36. Panjkovich, A., Norambuena, T., and Melo, F. (2005) dnaMATE: a consensus melting temperature prediction server for short DNA sequences. *Nucleic Acids Res.* **33**, 570
37. Minetti, C. A., Remeta, D. P., Dickstein, R., and Breslauer, K. J. (2010) Energetic signatures of single base bulges: Thermodynamic consequences and biological implications. *Nucleic Acids Res.* **38**, 97–116
38. SantaLucia, J., and Hicks, D. (2004) The thermodynamics of DNA structural motifs. *Annu. Rev. Biophys. Biomol. Struct.* **33**, 415–440
39. Minetti, C. A., Remeta, D. P., Johnson, F., Iden, C. R., and Breslauer, K. J. (2010) Impact of alpha-hydroxy-propanodeoxyguanine adducts on DNA duplex energetics: Opposite base modulation and implications for mutagenicity and genotoxicity. *Biopolymers* **93**, 370–382
40. Minetti, C. A., Remeta, D. P., Iden, C. R., Johnson, F., Grollman, A. P., and Breslauer, K. J. (2015) Impact of thymine glycol damage on DNA duplex energetics: Correlations with lesion-induced biochemical and structural consequences. *Biopolymers* **103**, 491–508
41. Plum, G. E., Grollman, A. P., Johnson, F., and Breslauer, K. J. (1992) Influence of an exocyclic guanine adduct on the thermal stability, conformation, and melting thermodynamics of a DNA duplex. *Biochemistry* **31**, 12096–12102
42. Plum, G. E., Grollman, A. P., Johnson, F., and Breslauer, K. J. (1995) Influence of the oxidatively damaged adduct 8-oxodeoxyguanosine on the conformation, energetics, and thermodynamic stability of a DNA duplex. *Biochemistry (N.Y.)*. **34**, 16148–16160
43. Gelfand, C. A., Plum, G. E., Grollman, A. P., Johnson, F., and Breslauer, K. J. (1996) The impact of a bistrand abasic lesion on DNA duplex properties. *Biopolymers* **38**, 439–445
44. Gelfand, C. A., Plum, G. E., Grollman, A. P., Johnson, F., and Breslauer, K. J. (1998) Thermodynamic consequences of an abasic lesion in duplex DNA are strongly dependent on base sequence. *Biochemistry (N.Y.)*. **37**, 7321–7327
45. Pilch, D. S., Plum, G. E., and Breslauer, K. J. (1995) The thermodynamics of DNA structures that contain lesions or guanine tetrads. *Curr. Opin. Struct. Biol.* **5**, 334–342
46. Plum, G. E., and Breslauer, K. J. (1994) DNA lesions. A thermodynamic perspective. *Ann. N. Y. Acad. Sci.* **726**, 45–46
47. Vesnaver, G., Chang, C. N., Eisenberg, M., Grollman, A. P., and Breslauer, K. J. (1989) Influence of abasic and nucleosidic sites on the stability, conformation, and melting behavior of a DNA duplex: Correlations of thermodynamic and structural data. *Proc. Natl. Acad. Sci. U. S. A.* **86**, 3614–3618
48. Plum, G. E., and Breslauer, K. J. (1995) Thermodynamics of an intramolecular DNA triple helix: A calorimetric and spectroscopic study of the pH and salt dependence of thermally induced structural transitions. *J. Mol. Biol.* **248**, 679–695
49. Plum, G. E., Park, Y. W., Singleton, S. F., Dervan, P. B., and Breslauer, K. J. (1990) Thermodynamic characterization of the stability and the melting behavior of a DNA triplex: A spectroscopic and calorimetric study. *Proc. Natl. Acad. Sci. U. S. A.* **87**, 9436–9440
50. Völker, J., and Klump, H. H. (1994) Electrostatic effects in DNA triple helices. *Biochemistry* **33**, 13502–13508
51. Sugimoto, N. (2014) Chapter eight - noncanonical structures and their thermodynamics of DNA and RNA under molecular crowding: Beyond the Watson–Crick double helix. *Int. Rev. Cell Mol. Biol.* **307**, 205–273
52. Sun, J., Carestier, T., and Hélène, C. (1996) Oligonucleotide directed triple helix formation. *Curr. Opin. Struct. Biol.* **6**, 327–333
53. Jin, R. Z., Breslauer, K. J., Jones, R. A., and Gaffney, B. L. (1990) Tetraplex formation of a guanine-containing nonameric DNA fragment. *Science* **250**, 543–546
54. Jin, R., Gaffney, B. L., Wang, C., Jones, R. A., and Breslauer, K. J. (1992) Thermodynamics and structure of a DNA tetraplex: A spectroscopic and calorimetric study of the tetramolecular complexes of d(TG3T) and d(TG3T2G3T). *Proc. Natl. Acad. Sci. U. S. A.* **89**, 8832–8836

55. Lane, A. N., Chaires, J. B., Gray, R. D., and Trent, J. O. (2008) Stability and kinetics of G-quadruplex structures. *Nucleic Acids Res.* **36**, 5482–5515
56. Lane, A. N. (2012) The stability of intramolecular DNA G-quadruplexes compared with other macromolecules. *Biochimie* **94**, 277–286
57. Redstone, S. C. J., Fleming, A. M., and Burrows, C. J. (2019) Oxidative modification of the potential G-quadruplex sequence in the PCNA gene promoter can turn on transcription. *Chem. Res. Toxicol.* **32**, 437–446
58. Fleming, A. M., and Burrows, C. J. (2020) Interplay of guanine oxidation and G-quadruplex folding in gene promoters. *J. Am. Chem. Soc.* **142**, 1115–1136
59. Völker, J., Plum, G. E., Gindikin, V., Klump, H. H., and Breslauer, K. J. (2014) Impact of bulge loop size on DNA triplet repeat domains: Implications for DNA repair and expansion. *Biopolymers* **101**, 1–12
60. McMurray, C. T. (1999) DNA secondary structure: A common and causative factor for expansion in human disease. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 1823–1825
61. Wells, R. D. (2009) Discovery of the role of non-B DNA structures in mutagenesis and human genomic disorders. *J. Biol. Chem.* **284**, 8997–9009
62. Bacolla, A., Wojciechowska, M., Kosmider, B., Larson, J. E., and Wells, R. D. (2006) The involvement of non-B DNA structures in gross chromosomal rearrangements. *DNA Repair (Amst)* **5**, 1161–1170
63. Zhao, J., Bacolla, A., Wang, G., and Vasquez, K. M. (2009) Non-B DNA structure-induced genetic instability and evolution. *Cell. Mol. Life Sci.* **67**, 43–62
64. Cooper, D. N., Bacolla, A., Férec, C., Vasquez, K. M., Kehrer-Sawatzki, H., and Chen, J. (2011) On the sequence-directed nature of human gene mutation: The role of genomic architecture and the local DNA sequence environment in mediating gene mutations underlying human inherited disease. *Hum. Mutat.* **32**, 1075–1099
65. Jin, R., Chapman, W. H., Srinivasa, A., Olson, W. K., Breslow, R., and Breslauer, K. J. (1993) Comparative spectroscopic, calorimetric, and computational studies of nucleic acid complexes with 2',5'- versus 3',5'-phosphodiester linkages. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 10568–10572
66. Braunlin, W., Völker, J., Plum, G. E., and Breslauer, K. J. (2013) DNA meter: Energy tunable, quantitative hybridization assay. *Biopolymers* **99**, 408–417
67. Matson, R. S. (2007) Hybridization analysis using oligonucleotide probe arrays. *Methods Mol. Biol.* **381**, 279–298
68. Crooke, S. T., Liang, X. H., Crooke, R. M., Baker, B. F., and Geary, R. S. (2020) Antisense drug discovery and development technology considered in a pharmacological context. *Biochem. Pharmacol.* **114196**
69. Crooke, S. T. (2004) Progress in antisense technology. *Annu. Rev. Med.* **55**, 6195
70. Cathomen, T., and Weitzman, M. D. (2005) Gene Repair: Pointing the finger at genetic disease. *Gene Ther.* **12**, 1415–1416
71. Shigi, N., Sumaoka, J., and Komiyama, M. (2017) Applications of PNA-based artificial restriction DNA cutters. *Molecules* **22**, 1586
72. Dervan, P. B. (1986) Design of sequence-specific DNA-binding molecules. *Science* **232**, 464–471
73. Maher, L. J., Wold, B., and DERVAN, P. B. (1991) Oligonucleotide-directed DNA triple-helix formation: An approach to artificial repressors? *Antisense Res. Dev.* **1**, 277–281
74. Huang, A., Garraway, L. A., Ashworth, A., and Weber, B. (2020) Synthetic lethality as an engine for cancer drug target discovery. *Nat. Rev. Drug Discov.* **19**, 23–38
75. Halpin-Healy, T., Klompe, S. E., Sternberg, S. H., and Fernández, I. S. (2020) Structural basis of DNA targeting by a transposon-encoded CRISPR-Cas system. *Nature* **577**, 271–274
76. Baskerville, S., and Bartel, D. P. (2005) Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. *RNA* **11**, 241–247
77. Lan, X., and Jin, V. X. (2012) Experimental and computational challenges from array-based to sequence-based CHIP techniques. *Curr. Bioinf.* **7**, 447–453
78. Jin, Q., Fleming, A. M., Ding, Y., Burrows, C. J., and White, H. S. (2013) Structural destabilization of DNA duplexes containing single-base lesions investigated by nanopore measurements. *Biochemistry (N.Y.)* **52**, 7870–7877
79. Minetti, C. A., Remeta, D. P., Zharkov, D. O., Plum, G. E., Johnson, F., Grollman, A. P., and Breslauer, K. J. (2003) Energetics of lesion recognition by a DNA repair protein: Thermodynamic characterization of formamidopyrimidine-glycosylase (Fpg) interactions with damaged DNA duplexes. *J. Mol. Biol.* **328**, 1047–1060
80. Minetti, C. A., Remeta, D. P., Miller, H., Gelfand, C. A., Plum, G. E., Grollman, A. P., and Breslauer, K. J. (2003) The thermodynamics of template-directed DNA synthesis: Base insertion and extension enthalpies. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 14719–14724
81. Esadze, A., Rodriguez, G., Cravens, S. L., and Stivers, J. T. (2017) AP-endonuclease 1 accelerates turnover of human 8-oxoguanine DNA glycosylase by preventing retrograde binding to the abasic-site product. *Biochemistry* **56**, 1974–1986
82. Stivers, J. T., and Jiang, Y. L. (2003) A mechanistic perspective on the chemistry of DNA repair glycosylases. *Chem. Rev.* **103**, 2729–2759
83. Zharkov, D. O., Shoham, G., and Grollman, A. P. (2003) Structural characterization of the Fpg family of DNA glycosylases. *DNA Repair (Amst)* **2**, 839–862
84. Shibutani, S., Bodepudi, V., Johnson, F., and Grollman, A. P. (1993) Translesional synthesis on DNA templates containing 8-oxo-7,8-dihydrodeoxyadenosine. *Biochemistry (N.Y.)* **32**, 4615–4621
85. Silvestri, R., and Landi, S. (2019) DNA polymerases in the risk and prognosis of colorectal and pancreatic cancers. *Mutagenesis* **34**, 363–374
86. Pilzecker, B., Buoninfante, O. A., and Jacobs, H. (2019) DNA damage tolerance in stem cells, ageing, mutagenesis, disease and cancer therapy. *Nucleic Acids Res.* **47**, 7163–7181
87. Akbari, M., and Krokan, H. E. (2008) Cytotoxicity and mutagenicity of endogenous DNA base lesions as potential cause of human aging. *Mech. Ageing Dev.* **129**, 353–365
88. Ma, X., Tang, T., and Guo, C. (2020) Regulation of translesion DNA synthesis in mammalian cells. *Environ. Mol. Mutagen.* **61**, 680–692
89. Nohmi, T., and Tsuzuki, T. (2016) Possible mechanisms underlying genotoxic thresholds. *Thresholds Genotoxic Carcinog.* 49–66
90. Yudkina, A. V., Shilkin, E. S., Endutkin, A. V., Makarova, A. V., and Zharkov, D. O. (2019) Reading and misreading 8-oxoguanine, a paradigmatic ambiguous nucleobase. *Crystals* **9**, 269
91. Yuan, B., Jiang, Y., Wang, Y., and Wang, Y. (2010) Efficient formation of the tandem thymine glycol/8-oxo-7,8-dihydroguanine lesion in isolated DNA and the mutagenic and cytotoxic properties of the tandem lesions in *Escherichia coli* cells. *Chem. Res. Toxicol.* **23**, 11–19
92. Datta, K., Johnson, N. P., LiCata, V. J., and von Hippel, P. H. (2009) Local conformations and competitive binding affinities of single- and double-stranded primer-template DNA at the polymerization and editing active sites of DNA polymerases. *J. Biol. Chem.* **284**, 17180–17193
93. Streisinger, G., and Owen, J. (1985) Mechanisms of spontaneous and induced frameshift mutation in bacteriophage T4. *Genetics* **109**, 633–659
94. Minetti, C. A., Sun, J. Y., Jacobs, D. P., Kang, I., Remeta, D. P., and Breslauer, K. J. (2018) Impact of bistrand abasic sites and proximate orientation on DNA global structure and duplex energetics. *Biopolymers* **109**, e23098
95. Liang, J. J., Wang, H., Rashid, A., Tan, T. H., Hwang, R. F., Hamilton, S. R., Abbruzzese, J. L., and Evans, D. B. (2008) Expression of MAP4K4 is associated with worse prognosis in patients with stage II pancreatic ductal adenocarcinoma. *Clin. Cancer Res.* **14**, 7043–7049
96. Nestorova, G. G., Adapa, B. S., Kopparchy, V. L., and Guilbeau, E. J. (2016) Lab-on-a-chip thermoelectric DNA biosensor for label-free detection of nucleic acid sequences. *Sens. Actuators B: Chem.* **225**, 174–180
97. Klinakis, A., Karagiannis, D., and Rampias, T. (2020) Targeting DNA repair in cancer: Current state and novel approaches. *Cell Mol. Life Sci.* **77**, 677–703
98. Gavande, N. S., VanderVere-Carozza, P. S., Pawelczak, K. S., and Turchi, J. J. (2016) Chapter 5 - Targeting the nucleotide excision repair

- pathway for therapeutic applications. In: Kelly, M. R., Fishel, M. L., eds. *DNA Repair in Cancer Therapy*, 2nd Ed., Academic Press, Boston: 135–150
99. Burdak-Rothkamm, S., Mansour, W. Y., and Rothkamm, K. (2020) DNA damage repair deficiency in prostate cancer. *Trends Cancer* **11**, 974–984
 100. Limp-Foster, M., and Kelley, M. R. (2000) DNA repair and gene therapy: Implications for translational uses. *Environ. Mol. Mutagen.* **35**, 71–81
 101. Ghosal, K., Agatemor, C., Han, R. I., Ku, A. T., Thomas, S., and Mukherjee, S. (2020) Fanconi anemia DNA repair pathway as a new mechanism to exploit cancer drug resistance. *Mini Rev. Med. Chem.* **20**, 779–787
 102. Hansen, W. K., and Kelley, M. R. (2000) Review of mammalian DNA repair and translational implications. *J. Pharmacol. Exp. Ther.* **295**, 1–9
 103. Lenglet, G., and David-Cordonnier, M. H. (2010) DNA-destabilizing agents as an alternative approach for targeting DNA: Mechanisms of action and cellular consequences. *J. Nucleic Acids* **2010**, 290935
 104. Suzuki, N., Ohashi, E., Kolbanovskiy, A., Geacintov, N. E., Grollman, A. P., Ohmori, H., and Shibutani, S. (2002) Translesion synthesis by human DNA polymerase kappa on a DNA template containing a single stereoisomer of dG-(+)- or dG-(-)-anti-N(2)-BPDE (7,8-dihydroxy-anti-9,10-epoxy-7,8,9,10-tetrahydrobenzo[a]pyrene). *Biochemistry* **41**, 6100–6106
 105. Rosenquist, T. A., Zaika, E., Fernandes, A. S., Zharkov, D. O., Miller, H., and Grollman, A. P. (2003) The novel DNA glycosylase, NEIL1, protects mammalian cells from radiation-mediated cell death. *DNA Repair (Amst)* **2**, 581–591
 106. Boothman, D. A. (1994) Enhanced malignant transformation is accompanied by increased survival recovery after ionizing radiation in Chinese hamster embryo fibroblasts. *Radiat. Res.* **138**, 121
 107. Boothman, D. A., Greer, S., and Pardee, A. B. (1987) Potentiation of halogenated pyrimidine radiosensitizers in human carcinoma cells by beta-lapachone (3,4-dihydro-2,2-dimethyl-2H-naphtho[1,2-b]pyran- 5, 6-dione), a novel DNA repair inhibitor. *Cancer Res.* **47**, 5361–5366
 108. Szczesny, B., Tann, A. W., and Mitra, S. (2010) Age- and tissue-specific changes in mitochondrial and nuclear DNA base excision repair activity in mice: Susceptibility of skeletal muscles to oxidative injury. *Mech. Ageing Dev.* **131**, 330–337
 109. Hegde, M. L., Hegde, P. M., Holthausen, L. M., Hazra, T. K., Rao, K. S., and Mitra, S. (2010) Specific inhibition of NEIL-initiated repair of oxidized base damage in human genome by copper and iron: Potential etiological linkage to neurodegenerative diseases. *J. Biol. Chem.* **285**, 28812–28825
 110. Ashley, C. T., J., and Warren, S. T. (1995) Trinucleotide repeat expansion and human disease. *Annu. Rev. Genet.* **29**, 703–728
 111. Orr, H. T., and Zoghbi, H. Y. (2007) Trinucleotide repeat disorders. *Annu. Rev. Neurosci.* **30**, 575–621
 112. Cummings, C. J., and Zoghbi, H. Y. (2000) Trinucleotide repeats: Mechanisms and pathophysiology. *Annu. Rev. Genomics Hum. Genet.* **1**, 281–328
 113. Sutherland, G. R., and Richards, R. I. (1995) Simple tandem DNA repeats and human genetic disease. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 3636–3641
 114. Khristich, A. N., and Mirkin, S. M. (2020) On the wrong DNA track: Molecular mechanisms of repeat-mediated genome instability. *J. Biol. Chem.* **295**, 4134–4170
 115. Völker, J., Gindikina, V., Klump, H. H., Plum, G. E., and Breslauer, K. J. (2012) Energy landscapes of dynamic ensembles of rolling triplet repeat bulge loops: Implications for DNA expansion associated with disease states. *J. Am. Chem. Soc.* **134**, 6033–6044
 116. Völker, J., Makube, N., Plum, G. E., Klump, H. H., and Breslauer, K. J. (2002) Conformational energetics of stable and metastable states formed by DNA triplet repeat oligonucleotides: Implications for triplet expansion diseases. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 14700–14705
 117. Völker, J., Klump, H. H., and Breslauer, K. J. (2007) DNA metastability and biological regulation: Conformational dynamics of metastable omega-DNA bulge loops. *J. Am. Chem. Soc.* **129**, 5272–5280
 118. Völker, J., Plum, G. E., Gindikina, V., and Breslauer, K. J. (2019) Dynamic DNA energy landscapes and substrate complexity in triplet repeat expansion and DNA repair. *Biomolecules* **9**, 709
 119. Völker, J., Klump, H. H., and Breslauer, K. J. (2008) DNA energy landscapes via calorimetric detection of microstate ensembles of metastable macrostates and triplet repeat diseases. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 18326–18330
 120. Li, M., Völker, J., Breslauer, K. J., and Wilson, D. M. (2014) APE1 incision activity at abasic sites in tandem repeat sequences. *J. Mol. Biol.* **426**, 2183–2198
 121. Lee, B. J., Barch, M., Castner, E. W., Völker, J., and Breslauer, K. J. (2007) Structure and dynamics in DNA looped domains: CAG triplet repeat sequence dynamics probed by 2-aminopurine fluorescence. *Biochemistry* **46**, 10756–10766
 122. Polyzos, A. A., and McMurray, C. T. (2017) Close encounters: Moving along bumps, breaks, and bubbles on expanded trinucleotide tracts. *DNA Repair (Amst)* **56**, 144–155
 123. McMurray, C. T. (2010) Mechanisms of trinucleotide repeat instability during human development. *Nat. Rev. Genet.* **11**, 786–799
 124. McMurray, C. T. (2008) Hijacking of the mismatch repair system to cause CAG expansion and cell death in neurodegenerative disease. *DNA Repair (Amst)* **7**, 1121–1134
 125. Liu, Y., Prasad, R., Beard, W. A., Hou, E. W., Horton, J. K., McMurray, C. T., and Wilson, S. H. (2009) Coordination between polymerase beta and FEN1 can modulate CAG repeat expansion. *J. Biol. Chem.* **284**, 28352–28366
 126. Lai, Y., Budworth, H., Beaver, J. M., Chan, N. L., Zhang, Z., McMurray, C. T., and Liu, Y. (2016) Crosstalk between MSH2-MSH3 and polbeta promotes trinucleotide repeat expansion during base excision repair. *Nat. Commun.* **7**, 12465
 127. Henry, C. M. (2001) Structure-based drug design. *Chem. Eng. News* **79**, 69–78
 128. Haq, I. (2001) Calorimetric techniques in the study of high-order DNA-drug interactions. *Methods Enzymol.* **340**, 109–149
 129. Breslauer, K. J., Remeta, D. P., Chou, W. Y., Ferrante, R., Curry, J., Zaunczkowski, D., Snyder, J. G., and Marky, L. A. (1987) Enthalpy-entropy compensations in drug-DNA binding studies. *Proc. Natl. Acad. Sci. U. S. A.* **84**, 8922–8926
 130. Haq, I. (2000) Parsing free energies of drug-DNA interactions. *Methods Enzymol.* **323**, 373–405
 131. Chaires, J. B. (2008) Calorimetry and thermodynamics in drug design. *Annu. Rev. Biophys.* **37**, 135–151
 132. Chaires, J. B. (2018) Thermal denaturation of drug-DNA complexes. *Chem. Biol. (Cambridge, U.K.)* **7**, 74–95
 133. Garbett, N. C., Garbett, N. C., and Chaires, J. B. (2012) Thermodynamic studies for drug design and screening. *Expert Opin. Drug Discov.* **7**, 299–314
 134. Yu, H., Ren, J., Chaires, J. B., and Qu, X. (2008) Hydration of drug–DNA complexes: Greater water uptake for adriamycin compared to daunomycin. *J. Med. Chem.* **51**, 5909–5911
 135. Qu, X., and Chaires, J. B. (2001) Hydration changes for DNA intercalation reactions. *J. Am. Chem. Soc.* **123**, 1–7
 136. Chalikian, T. V., and Breslauer, K. J. (1996) Compressibility as a means to detect and characterize globular protein states. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 1012–1014
 137. Son, I., Shek, Y. L., Dubins, D. N., and Chalikian, T. V. (2014) Hydration changes accompanying helix-to-coil DNA transitions. *J. Am. Chem. Soc.* **136**, 4040–4047
 138. Chalikian, T. V. (2008) On the molecular origins of volumetric data. *J. Phys. Chem. B.* **112**, 911–917
 139. Chalikian, T. V., and Breslauer, K. J. (1996) On volume changes accompanying conformational transitions of biopolymers. *Biopolymers* **39**, 619–626
 140. Chalikian, T. V. (2003) Volumetric properties of proteins. *Annu. Rev. Biophys. Biomol. Struct.* **32**, 207–235
 141. Chalikian, T. V., and Breslauer, K. J. (1998) Volumetric properties of nucleic acids. *Biopolymers* **48**, 264–280
 142. Han, F., Taulier, N., and Chalikian, T. V. (2005) Association of the minor groove binding drug Hoechst 33258 with d(CGCGAATTCGCG)₂: Volumetric, calorimetric, and spectroscopic characterizations. *Biochemistry* **44**, 9785–9794

143. Chalikian, T. V., Plum, G. E., Sarvazyan, A. P., and Breslauer, K. J. (1994) Influence of drug binding on DNA hydration: Acoustic and densimetric characterizations of netropsin binding to the poly(-dAdT).poly(dAdT) and poly(dA).poly(dT) duplexes and the poly(dT).poly(dA).poly(dT) triplex at 25 degrees C. *Biochemistry* **33**, 8629–8640
144. Garbett, N. C., and Chaires, J. B. (2008) Binding: A polemic and rough guide. *Methods Cell Biol.* **84**, 1–23
145. Freire, E. (2008) Do enthalpy and entropy distinguish first in class from best in class? *Drug Discov. Today* **13**, 869–874
146. Kawasaki, Y., and Freire, E. (2011) Finding a better path to drug selectivity. *Drug Discov. Today* **16**, 985–990
147. Holdgate, G., Geschwindner, S., Breeze, A., Davies, G., Colclough, N., Temesi, D., Ward, L., Williams, M. A., and Daviter, T. (2013) Biophysical methods in drug discovery from small molecule to pharmaceutical. *Methods Mol. Biol.* **1008**, 327–355
148. Ciulli, A., Williams, M. A., and Daviter, T. (2013) Biophysical screening for the discovery of small-molecule ligands. *Methods Mol. Biol.* **1008**, 357–388
149. Fulroth, B. (2017) Modern biophysical methods for screening and drug discovery. In: Bittker, J. A., Ross, N. T., eds, *High Throughput Screening Methods: Evolution and Refinement. Chemical Biology*. Royal Society of Chemistry, London, England: **1**, 58–86
150. Hevener, K. E., Pesavento, R., Ren, J., Lee, H., Ratia, K., and Johnson, M. E. (2018) Chapter 12 — Hit-to-lead: Hit validation and assessment. *Methods Enzymol.* **610**, 265–309
151. Klump, H. H., Völker, J., and Breslauer, K. J. (2020) Energy mapping of the genetic code and genomic domains: Implications for code evolution and molecular Darwinism. *Q. Rev. Biophys.* **53**, e11