

Redesigning the Architecture of the Base Pair: Toward Biochemical and Biological Function of New Genetic Sets

Andrew T. Krueger¹ and Eric T. Kool^{1,*}

¹Department of Chemistry, Stanford University, Stanford, CA 94305-5080, USA

*Correspondence: kool@stanford.edu

Chemistry & Biology invites your comments on this topic. Please write to the editors at chembiol@cell.com.

DOI 10.1016/j.chembiol.2008.12.004

Recognition of the nucleic acid bases within the DNA scaffold comprises the basis for transmission of genetic information, dictating protein and cell assembly, organismal development, and evolution. Driven in part by the need to test our current understanding of this information transfer, chemists have begun to design and synthesize nonnatural bases and base pair structures to mimic the function of DNA without relying on Nature's purine-pyrimidine base pair scaffold. Multiple examples have been recently described that self-assemble stably and sequence specifically *in vitro*, and some isolated unnatural base pairs can be replicated *in vitro* as well. Moreover, recent experiments with unnatural bases in bacterial cells have demonstrated surprisingly efficient reading of the chemical information. This suggests the future possibility of redesigning and replacing the chemical information of an evolving cell while retaining biological function.

What I cannot build, I cannot understand.

—Richard P. Feynman

The exquisitely organized complexity of biological systems has been a longstanding curiosity both of biologists and chemists. However, the two disciplines have commonly taken different approaches to studying the machinery that gives cells their function. Biologists have often taken a “top down” approach, for example by taking a cell type and removing or altering genes, observing the effect on living cell function. In contrast, chemists are often drawn to a “bottom up” strategy, in which the smallest molecular components (such as amino acids or enzyme cofactors) are altered and examined outside the cell; this approach started as the field called biomimetic chemistry (Breslow, 1972). This dichotomy between biologists and chemists is also evident in the emerging discipline of synthetic biology (Rawls, 2000; Benner, 2005), in which biological engineers design and swap pathways between different living systems, whereas chemists have been working on design of new components, building upward toward biological function. However, chemistry laboratories are increasingly adopting biological methods, and so the ability to probe/mimic the machinery of a living cell on a chemical level is being increasingly realized. A meeting in the middle, using

approaches from both fields, ultimately seems likely, because the approaches and goals are often the same: altering natural components in living (or lifelike) systems, and exploiting their properties for biotechnology and medicine. At the same time, as Feynman would have recognized, we test and refine our knowledge of how biological systems work.

The biology-inspired approach has long been integral in furthering our understanding of the natural genetic molecule (DNA) and its interaction with proteins. Since Watson and Crick, chemists have been inspired to ask what it is that makes DNA so unique, and explore whether alternative architectures for transmitting biological information can exist. With advances in organic synthesis, analytical techniques, and solid phase chemistry, and in combination with modern biochemical and biological methods, an increasing number of research labs have synthesized nonnatural nucleic acid base and base pair architectures designed to mimic, complement, or even replace the natural structures. The work has enabled the exploration of these fundamental questions, and application of the chemical and biological insight is leading to biotechnological advances.

Why Modify DNA Bases?

By designing new replacements for DNA bases, and observing their functions, we

gain a better understanding of natural nucleic acids, and their crucial roles in the context of cellular function. Beyond this basic science, the knowledge gained from understanding a genetic system on a chemical level allows for design of components that might have modified or enhanced functionality, such as a means of reporting on a particular biological event, or a better design of DNA recognition scaffolds, or an important assay for detection of disease.

Although the chemical modification of DNA has long been studied, research on DNA base modifications is a newer field. The design of replacements for the sugar-phosphate backbone of DNA has proceeded for three decades and has yielded many successes, which have been reviewed elsewhere (Eschenmoser, 1999, and references therein). In contrast, the topic at hand here is on the chemical information encoded by the structures of DNA bases, leaving the sugar-phosphate backbone alone. In this review we outline recent accomplishments in creating nonnatural nucleic acid bases and base pairs, and describe efforts toward alternative genetic systems that contain them. We also outline both the basic scientific information and the technological benefits that have resulted from observing their function *in vitro* and more recently, in living systems.

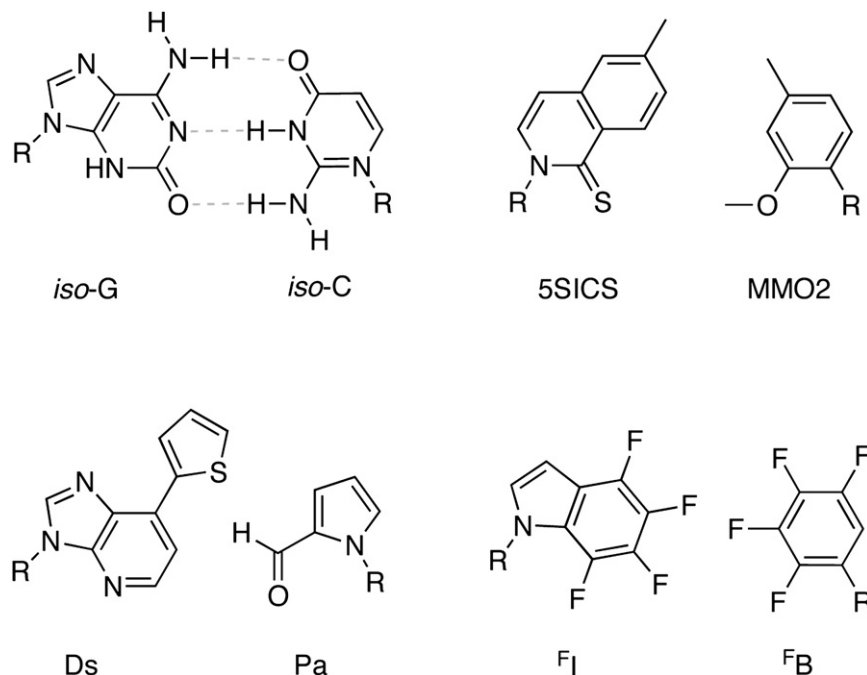


Figure 1. Examples of Base Pairs Designed to Augment the Natural Genetic Set
Their dimensions are close to those of natural Watson-Crick pairs. R represents deoxyribose.

Reason #1: Expanding the Genetic Lexicon

One goal of redesigning the DNA base pair architecture is to expand the information-encoding capability of DNA. To date this has been done, for the most part, in the context of DNA, with new base pairs being designed to function alongside the natural pairs (Figure 1). Such new pairs can function as selective hybridizing groups that do not pair with natural bases, or as selective encoders for incorporation of reporters into RNA, or as additions to the genetic code, leading to proteins with novel functionality. Design and synthesis of new, orthogonal base pairs capable of functioning similarly to, and in context of, the natural DNA scaffold has been a topic of extensive research since the late 1980s.

In this regard, pioneering work from the Benner laboratory has yielded new base pairs in which the hydrogen bonding arrangement is different than the canonical Watson-Crick hydrogen bonding motif of the natural bases (Switzer et al., 1989). The earliest of these designs, the isoG-isoC base pair, was shown by Benner to be both replicated and translated correctly in vitro (Piccirilli et al., 1990). With practical applications in mind,

Prudent demonstrated this as a third base pair for the polymerase chain reaction (PCR) (Johnson et al., 2004), which demands high base pair fidelity. In these experiments, up to 96% fidelity per cycle was attained, and errors from mispairing during replication were found to be due to oxidation and tautomerization of the unnatural analogs (Switzer et al., 1993). To ameliorate this, Benner designed thymidine analogs that still base pair efficiently with natural adenine, but due to steric bulk of a thiocarbonyl no longer mispair with the undesired isoG enol tautomer (Sismour and Benner, 2005). Recently, retaining the same shuffled hydrogen bonding pattern, Benner and colleagues have designed a new base pair between a nitropyridone (dZ) and a triaza-isoG (dP) that improves upon these problems as well (Yang et al., 2007).

The aforementioned analogs have yet to be incorporated into living cells and evaluated for their potential to expand the genetic code for in vivo translation. However, the orthogonality of such unnatural base pairs to natural DNA has been exploited elegantly in biotechnology. Currently, the isoC-isoG base pair is used as part of an amplification assay in clinics to monitor viral load of hepatitis

patients (Collins et al., 1997). The high signal-to-background of the assay relies on the design of capture probes containing unnatural bases, thereby eliminating false positives caused by accidental binding to cellular DNA, which does not recognize the unnatural bases. Prudent has also applied the isoC/isoG bases in sensitive real-time PCR quantification (Moser et al., 2003) using endonuclease activity to report on sequence.

In contrast to the approach of rearranging the Watson-Crick hydrogen bonds, others have eliminated base pair hydrogen bonding altogether. Over a decade ago, Kool and coworkers described the concept of nonpolar nucleoside isosteres (Schweitzer and Kool, 1994); examples included difluorotoluene (dF) and 4-methylbenzimidazole (dZ), which are hydrophobic, non-hydrogen-bonding base analogs of natural thymidine and adenine, respectively. These analogs were shown to form selective pairs with one another, and helped to reveal the important roles of base-base and base-solvent interactions in the stability and selectivity of pairing in DNA (Schweitzer and Kool, 1995). Results of pairing studies of nonpolar base analogs demonstrated that selective base pairing can also be achieved by shape complementarity and hydrophobic effects (Lai et al., 2004). These nonpolar isosteres have been extensively studied both in DNA alone and in relation to DNA-protein recognition in multiple contexts, including topoisomerases (Yakovleva et al., 2006), DNA repair enzymes (Chepanoske et al., 2000; Francis et al., 2003; Kirby et al., 2005), and DNA polymerases (Moran et al., 1997; Morales and Kool, 1998, 2000; Potapova et al., 2006; Mizukami et al., 2006; Kim et al., 2006; Lee et al., 2008).

The roles of stacking, hydrophobic effects, solvation, and steric effects were more recently demonstrated in the Kool laboratory with pairs that were both selective and in some cases, as stable as natural Watson-Crick pairs. Examples include the “pairing” of a synthetic pyrene nucleoside and an abasic site (Matray and Kool, 1998), as well as superhydrophobic fluorous pairs (Lai et al., 2004; Lai and Kool, 2005). In both cases, the unnatural bases preferred pairing with their designed partners relative to any of the natural bases within the framework of

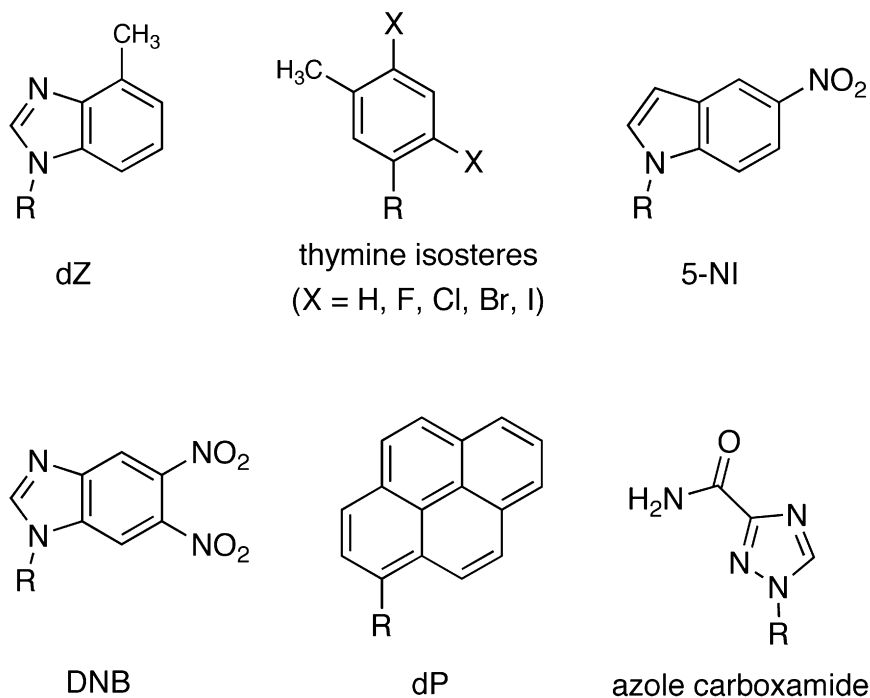


Figure 2. Examples of DNA Base Replacements that Have Been Used for Probing Nucleobase Recognition Mechanisms

DNA. Interestingly, the Klenow fragment (Exo-) of *E. coli* DNA polymerase I was able to selectively and efficiently process such pairs; for example, the nucleoside triphosphate derivative of pyrene nucleoside was efficiently incorporated opposite an abasic sugar (Matray and Kool, 1999).

In the last decade, several laboratories have explored hydrophobic nucleobase analogs as a viable solution for expansion of the genetic alphabet. Not only can such pairs be stable, but they also commonly have the advantage that their pairing with the more polar natural bases is unfavorable as a result of strong desolvation costs. Prominent examples have come from the laboratories of Romesberg and of Hirao, both of which have made multiple successful contributions on this front. Romesberg has synthesized an impressive array of hydrophobic bases, a number of which show selective base pair formation relying on hydrophobic packing, and a few of which are exceptionally stable (Leconte et al., 2006; Matsuda et al., 2006). An important recent example is a methyl methoxybenzene (MMO2)5-methylthioisocarbostyryl (5SICS) base pair, which could be efficiently inserted and extended by the Klenow fragment, making MMO2-5SICS a potentially

useful hydrophobic third base pair for DNA (Leconte et al., 2008).

Also utilizing the principles of shape complementarity and close hydrophobic packing, Hirao and coworkers were able to show highly efficient PCR amplification of an unnatural pair added to DNA using a 7-(2-thienyl)-imidazo(4,5-*b*)pyridine (Ds) and a pyrrole-2-carbaldehyde (Pa) base pair (Hirao et al., 2006). Even more impressively, Hirao designed an additional unnatural RNA nucleotide that selectively pairs with Pa, thereby allowing the amplified DNA to be transcribed. This method offers great potential for selective incorporation of unnatural bases and other functionalities into RNA. In one example, Hirao successfully incorporated a fluorescent unnatural base at specific sites into RNA as a tool for studying RNA secondary structure (Kimoto et al., 2007).

Although the Ds-Pa base pair functions well in *in vitro* amplification and transcription, the modified amidotriphosphates required for optimal base pair selectivity were found to slow the PCR process. To address this, a 2-nitropyrrole base analog (Pn) was designed, which better pairs with Ds, thereby outcompeting the previously observed mispair with natural adenine. After 20 cycles of PCR amplification using

Vent Exo- polymerase, less than 1% mutation rate was observed, making the Ds-Pn base pair even more promising as a third base pair for DNA (Kimoto et al., 2008).

Reason #2: Building Tools for Studying Biological Mechanisms

Unnatural nucleic acid bases have been useful not only in imparting novel function to DNA, but also in probing interactions within DNA itself. Unnatural nucleosides have been integral in determining the forces and mechanisms essential in base pairing, which is of course required for transfer of genetic information in a cell. Because many diseases can be traced to errors in this information transfer, it is important to understand the mechanisms of base pair recognition in this process, not only *in vitro*, but also *in vivo*.

Several dozens of base analogs have been reported to date as tools for investigating these steric, shape, and electrostatic effects in both DNA and RNA (Figure 2). For example, Berdis has synthesized substituted indole analogs and found pi-surface area to be a prime factor in base pair insertion during DNA polymerization and translesion synthesis by bacteriophage T4 polymerase (Zhang et al., 2005, 2006). Using substituted benzimidazole analogs, Engels and Kuchta concluded that hydrophobics, electronics, and sterics cannot alone regulate polymerization of DNA by polymerase α and Klenow fragment (Kincaid et al., 2005). Engels also described a number of fluorobenzenes and fluorobenzimidazoles that were used to investigate electrostatic effects of base pairing in RNA (Parsch and Engels, 2000). In one study, it was found that the fluorinated bases prefer to pair with themselves relative to the natural bases, similar to results seen by Kool in DNA (Schweitzer and Kool, 1995; Lai et al., 2004). Davisson collected evidence on the roles of sterics and electrostatics in DNA synthesis using azole carboxamide base analogs (Paul et al., 2003). McLaughlin has designed analogs lacking minor groove hydrogen bond acceptors to explore the roles of minor groove solvation in DNA stability (Woods et al., 2003).

Also in this light, ongoing work with nonpolar nucleoside analogs of varying shape and size has shown that steric effects can be critical in the biology of DNA replication. Kim and Kool reported

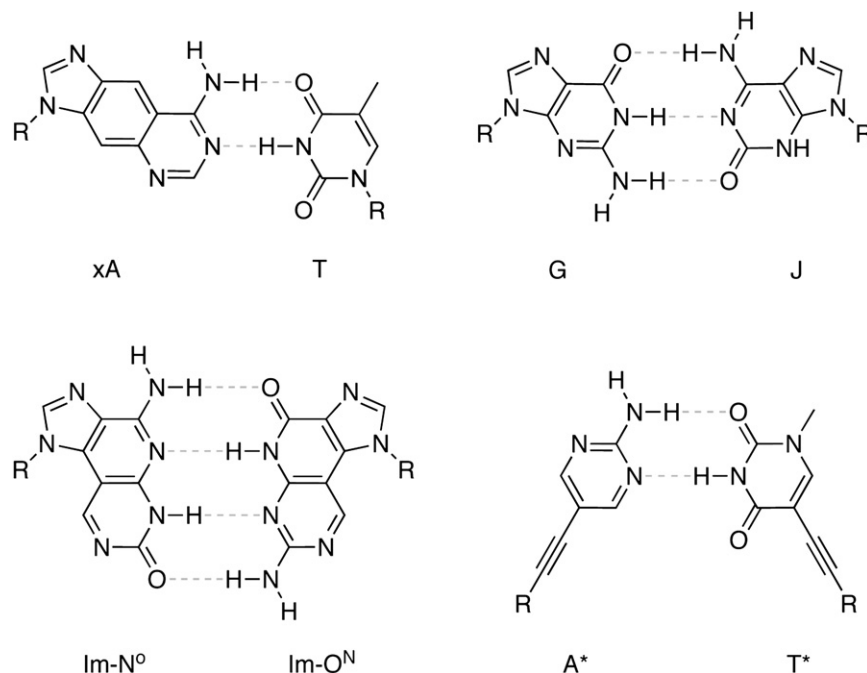


Figure 3. Novel Base-Pairing Systems Designed to Form New, Non-Purine-Pyrimidine Genetic Sets

All have sizes larger than Watson-Crick pairs.

a series of nonpolar isostere analogs of thymidine with systematically varying size, where the 2,4 carbonyl oxygens were replaced by -H, -F, -Cl, -Br, and -I (Kim and Kool, 2005). Studies with this series pointed out a marked size preference during replication by Klenow fragment, with efficiency reaching a sharp maximum at the Cl analog (Kim et al., 2005). Sintim tested related analogs with similar size but with varying shapes, and also found large kinetic effects of shape during DNA polymerization (Sintim and Kool, 2006). The studies showed that steric effects around the incipient base pair are large enough in magnitude to explain most of the selectivity that occurs during DNA replication.

Although modifying the chemical parts of DNA to probe DNA structure and recognition in vitro has proven beneficial, these studies are not necessarily a good predictor for how unnatural nucleobases will be tolerated in a living cell, where multiple polymerases, repair mechanisms, and biological pathways are at play. To date there are very few studies in which unnatural bases have been introduced into a living system while retaining functionality. Essigmann and Kool achieved this in 2003 when they per-

formed replication bypass experiments in *E. coli* with a single-stranded M3 phage containing a synthetic insert having the nonpolar nucleoside isosteres dQ (a deoxyadenosine isostere) or difluorotoluene dF (thymidine isostere) (Delaney et al., 2003). By counting phage progeny, it was found that dF and dQ are moderately to modestly bypassed by polymerases under normal conditions, and when an SOS response was induced, both isosteres were efficiently bypassed. More remarkably, a base composition assay showed that replication occurred with insertion of the correct Watson-Crick partner opposite the unnatural bases under both SOS induced and noninduced conditions. These observations show that it is possible to retain biological function with unnatural nucleobase components, and also suggested the importance of DNA base shape (in opposed to hydrogen bonding alone) as a prime determinant of fidelity in vivo.

Essigmann and Kool followed this study with an in vivo evaluation of steric effects, using the variably sized thymidine nonpolar isostere series previously reported by Kim. Phage replication experiments in *E. coli* demonstrated that the cellular replicative machinery preferred

the slightly larger-than-natural dichlorotoluene analog, with a steep drop off in efficiency and fidelity for analogs containing substituents smaller (H, F) or larger (Br, I) than Cl. This showed a marked steric dependence for DNA replication in vivo, and also led to the suggestion that high-fidelity polymerases have more steric room than necessary, perhaps to allow for evolutionary variation (Kim et al., 2005).

Nonpolar nucleoside isosteres continue to prove useful as tools for probing biological mechanisms and interactions. More recently, two laboratories at Alnylam and at Stanford have independently used difluorobenzene ribonucleoside derivatives to test electrostatic requirements of sequence recognition in RNA interference (Xia et al., 2006; Somoza et al., 2006). Interestingly, the Watson-Crick hydrogen bonding requirement varied strongly along the sequence, and at some sites it was found that this bonding is not required to maintain activity in RNA interference (RNAi) in HeLa cells. The origins of selectivity and activity in RNA interference are complex and are still not well understood; thus future research involving a variety of nucleic acid analogs will be invaluable in understanding RNAi.

Reason #3: Designing Alternative Genetic Systems

Most if not all of the unnatural bases mentioned above were designed to function within the context and geometry of the natural genetic system. Thus they maintain a general size and shape that fits well next to purine-pyrimidine pairs. However, recently an alternative strategy has begun to be explored. In this new approach, researchers are asking whether *all* of the natural base pairs of DNA can be replaced, giving rise to an entirely new genetic set not confined by the geometry of the natural Watson-Crick base pairs (Figure 3). Designing an alternative genetic set (and ultimately, a functioning genetic system) can be useful in exploring nature's evolutionary arrival at the structure of DNA for storage of genetic information. This allows one to address some very basic questions: Can other base pair architectures assemble and function similarly to DNA? Do we know enough about DNA to design a similar information-storing system from novel chemical parts?

In this context, Kool and coworkers have extensively explored base pairing,

assembly, and structural properties of an alternative genetic set with size-expanded geometry (xDNA) (Liu et al., 2003; Krueger et al., 2007). This genetic set is composed of eight bases with four types of pairings. Each of the pairs contains a size-expanded base that resembles a natural purine or pyrimidine, but has its Watson-Crick pairing edge shifted outward by the insertion of a benzene ring (Leonard et al., 1976; Liu et al., 2004, 2005a). This gives rise to base pairs with a 2.4 Å larger C1'-C1' backbone distance than natural DNA. In xDNA, benzopurines pair with natural purines, and benzopyrimidines pair with natural purines, making the eight-letter set orthogonal to natural DNA. Because there are eight letters, the information encoding potential of xDNA is exponentially larger than that of DNA. Kool has found that whereas a single expanded base pair in natural DNA is destabilizing, helices composed entirely of size-expanded base pairs are more stable than natural DNA, because of more favorable base stacking (Liu et al., 2005b; Gao et al., 2005). In addition, the expanded DNA bases are brightly fluorescent (Krueger and Kool, 2008), a property virtually absent in natural DNA.

Although the xDNA genetic set is only one component of a potentially alternative genetic system (which also would require machinery for replication), the insight gained from these experiments has shed light on the tolerance of the natural phosphodiester backbone toward alternative pairing geometries that might be capable of genetic information storage. Recently, Kool has explored other expanded geometries (such as one termed "wide DNA" [yDNA]) (Lu et al., 2004; Lee and Kool, 2005) and has even shown that duplex assembly is possible with base pairs containing naphtho-homologated pyrimidines paired with purines (yyDNA) (Lee and Kool, 2006). Thus it appears that the natural phosphodiester backbone of DNA can be quite tolerant of different base pair sizes, as long as they are all congruent in their geometry.

Other laboratories have also experimented with altering the intermolecular backbone distance in DNA. Battersby has observed antiparallel, selective Watson-Crick-like duplex formation of oligonucleotides composed entirely of purines (Battersby et al., 2007). In this scaffold,

the pairing partner for natural adenine is hypoxanthine (H), and for guanine, an N³-H tautomer of isoG (J). The demonstration of an all-purine duplex is interesting, considering all the building blocks could have potentially originated from prebiotic mixtures. This raises interesting questions about nature's criteria for an evolving system of nucleic acids.

Taking a related expanded-size approach, Inouye has recently designed an artificial genetic set composed entirely of nonnatural C-glycosides in which added length is conferred by an ethynyl group (Doi et al., 2008). The recognition of these bases relies on six-member heterocycles with Watson-Crick-like hydrogen-bonding faces (iG^{*}:iC^{*}, three hydrogen bonds, A^{*}:T^{*} two hydrogen bonds), linked to the C1' of deoxyribose by the ethynyl spacer. Oligomers composed of these unnatural bases formed highly stable, antiparallel duplexes and triplexes, and displayed selectivity against mismatches as well.

Matsuda and coworkers have synthesized base pairs that increase not only the C1'-C1' glycosidic distance relative to natural DNA, but also the number of hydrogen bonds (Minakawa et al., 2003; Hikishima et al., 2005). It was found that the addition of one imidazopyridopyrimidine (N^N, O^O, O^N, N^O) base pair in DNA did not add any thermal stabilization to the duplex relative to an A:T or G:C base pair. However, with the addition of consecutive imidazopyridopyrimidine base pairs (up to three), substantial duplex stabilization relative to the natural sequence could be seen, likely due to favorable aromatic stacking and relief of backbone strain (Minakawa et al., 2003). Although helices composed entirely of imidazopyridopyrimidine base pairs have not yet been reported, such bases and pairs seem quite promising.

The work to date aimed at developing alternative genetic systems has clearly yielded base-pairing systems that function very well in assembly, affording highly stable helices with selective base pairing. Thus they have the potential to store information (in some cases, large amounts of information). However, it is not yet clear whether this information can be read and transferred, which are clearly required for a functioning genetic system. One challenge in this regard is to find enzymes that can accept the unnatural geometries

of such pairs and helices (Gloekner et al., 2007; Xia et al., 2002); as it stands now, such large systems are expected to be alien to natural polymerases, which for the most part evolved to process the natural geometry and exclude molecules that do not fit.

Future Outlook

Although the chemical and biophysical knowledge of DNA is becoming more complete by use of nonnatural bases and pairs, the implications on the larger biological system of replacing these chemical parts is relatively new, and requires more research. In addition, the redesign of more complex biological components such as the natural polymerases is a challenging task that clearly deserves much more work. As more research is done with unnatural base pair design, alternative genetic systems, and polymerases, synthetic biologists will come closer to the goal of understanding and redesigning living systems to further benefit science and technology.

ACKNOWLEDGMENTS

We thank the U.S. National Institutes of Health (GM63587 and GM072705) for support of work on unnatural bases and base pairs for DNA and RNA.

REFERENCES

- Battersby, T.R., Albalos, M., and Friesenhahn, M.J. (2007). An unusual mode of DNA duplex association: Watson-Crick interaction of all-purine deoxyribonucleic acids. *Chem. Biol.* 14, 525–531.
- Benner, S.A. (2005). Synthetic biology: act natural. *Nature* 427, 118.
- Breslow, R. (1972). Centenary lecture. Biomimetic chemistry. *Chem. Soc. Rev.* 1, 553–580.
- Chepanoske, C.L., Langelier, C.R., Chmiel, N.H., and David, S.S. (2000). Recognition of the nonpolar base 4-methylindole in DNA by the DNA repair adenine glycosylase MutY. *Org. Lett.* 2, 1341–1344.
- Collins, M.L., Irvine, B., Tyner, D., Fine, E., Zayati, C., Chang, C., Horn, T., Ahle, D., Detmer, J., Shen, L., et al. (1997). A branched DNA signal amplification assay for quantification of nucleic acid targets below 100 molecules/ml. *Nucleic Acids Res.* 25, 2979–2984.
- Delaney, J.C., Henderson, P.T., Helquist, S.A., Morales, J.C., Essigmann, J.M., and Kool, E.T. (2003). High-fidelity *in vivo* replication of DNA base shape mimics without Watson-Crick hydrogen bonds. *Proc. Natl. Acad. Sci. USA* 100, 4469–4473.
- Doi, Y., Chiba, J., Morikawa, T., and Inouye, M. (2008). Artificial DNA made exclusively of

- nonnatural C-nucleosides with four types of nonnatural bases. *J. Am. Chem. Soc.* **130**, 8762–8768.
- Eschenmoser, A. (1999). Chemical etiology of nucleic acid structure. *Science* **284**, 2118–2124.
- Francis, A.W., Helquist, S.A., Kool, E.T., and David, S.S. (2003). Probing the requirements for recognition and catalysis in Fpg and MutY with nonpolar adenine isosteres. *J. Am. Chem. Soc.* **125**, 16235–16242.
- Gao, J., Liu, H., and Kool, E.T. (2005). Assembly of the complete eight-base artificial genetic helix, xDNA, and its interaction with the natural genetic system. *Angew. Chem. Int. Ed. Engl.* **44**, 3118–3122.
- Gloekner, C., Sauter, K.B.M., and Marx, A. (2007). Evolving a thermostable DNA polymerase that accurately amplifies highly-damaged templates. *Angew. Chem. Int. Ed. Engl.* **46**, 3115–3117.
- Hikishima, S., Minakawa, N., Kuramoto, K., Fujisawa, Y., Ogawa, M., and Matsuda, A. (2005). Synthesis of 1,8-naphthyridine C-nucleosides and their base-pairing properties in oligodeoxynucleotides: thermally stable naphthyridine:imidazopyridopyrimidine base-pairing motifs. *Angew. Chem. Int. Ed. Engl.* **44**, 596–598.
- Hirao, I., Kimoto, M., Mitsui, T., Fujiwara, T., Kawai, R., Sato, A., Harada, Y., and Yokoyama, S. (2006). An unnatural hydrophobic base pair system: site-specific incorporation of nucleotide analogs into DNA and RNA. *Nat. Methods* **3**, 729–735.
- Johnson, S.C., Sherrill, C.B., Marshall, D.J., Moser, M.J., and Prudent, J.R. (2004). A third base pair for the polymerase chain reaction: inserting isoC and isoG. *Nucleic Acids Res.* **32**, 1937–1941.
- Kim, T.W., and Kool, E.T. (2005). A series of nonpolar thymine analogues of increasing size: DNA base pairing and stacking properties. *J. Org. Chem.* **70**, 2048–2053.
- Kim, T.W., Delaney, J.C., Essigmann, J.M., and Kool, E.T. (2005). Probing the active site tightness of DNA polymerase in sub-angstrom increments. *Proc. Natl. Acad. Sci. USA* **102**, 15803–15808.
- Kim, T.W., Brieba, L.G., Ellenberger, T., and Kool, E.T. (2006). Functional evidence for a small and rigid active site in a high fidelity DNA polymerase: probing T7 DNA polymerase with variably-sized base pairs. *J. Biol. Chem.* **281**, 2289–2295.
- Kimoto, M., Mitsui, T., Harada, Y., Sato, A., Yokoyama, S., and Hirao, I. (2007). Fluorescent probing for RNA molecules by an unnatural base-pair system. *Nucleic Acids Res.* **35**, 5360–5369.
- Kimoto, M., Kawai, R., Mitsui, T., Yokoyama, S., and Hirao, I. (2008). Efficient PCR amplification by an unnatural base pair system. *Nucleic Acids Symp. Ser. (Oxf.)* **52**, 469–470.
- Kincaid, K., Beckman, J., Zivkovic, A., Halcomb, R.L., Engels, J.W., and Kuchta, R.D. (2005). Exploration of factors driving incorporation of unnatural dNTPs into DNA by Klenow fragment (DNA polymerase I) and DNA polymerase α . *Nucleic Acids Res.* **33**, 2620–2628.
- Kirby, T.W., DeRose, E.F., Beard, W.A., Wilson, S.H., and London, R.E. (2005). A thymine isostere in the templating position disrupts assembly of the closed DNA polymerase β ternary complex. *Biochemistry* **44**, 15230–15237.
- Krueger, A.T., and Kool, E.T. (2008). Fluorescence of size-expanded DNA bases: reporting on DNA sequence and structure with an unnatural genetic set. *J. Am. Chem. Soc.* **130**, 3989–3999.
- Krueger, A.T., Lu, H., Lee, A.H.F., and Kool, E.T. (2007). Synthesis and properties of size-expanded DNAs: toward designed, functional, genetic systems. *Acc. Chem. Res.* **40**, 141–150.
- Lai, J.S., and Kool, E.T. (2005). Fluorous base-pairing effects in a DNA polymerase active site. *Chemistry* **11**, 2966–2971.
- Lai, J.S., Qu, J., and Kool, E.T. (2004). Selective pairing between polyfluorinated DNA bases. *J. Am. Chem. Soc.* **126**, 3040–3041.
- Leconte, A.M., Matsuda, S., and Romesberg, F.E. (2006). An efficiently extended class of unnatural base pairs. *J. Am. Chem. Soc.* **128**, 6780–6781.
- Leconte, A.M., Hwang, G.T., Matsuda, S., Capek, P., Hari, Y., and Romesberg, F.E. (2008). Discovery, characterization, and optimization of an unnatural base pair for expansion of the genetic alphabet. *J. Am. Chem. Soc.* **130**, 2336–2343.
- Lee, A.H.F., and Kool, E.T. (2005). A new four-base genetic helix, yDNA, composed of widened benzopyrimidine-purine pairs. *J. Am. Chem. Soc.* **127**, 3332–3338.
- Lee, A.H.F., and Kool, E.T. (2006). Exploring the limits of DNA base pair size: naphtho-homologous DNA bases and pairs. *J. Am. Chem. Soc.* **128**, 9219–9230.
- Lee, H.R., Helquist, S.A., Kool, E.T., and Johnson, K.A. (2008). Importance of hydrogen bonding for efficiency and specificity of the human mitochondrial DNA polymerase. *J. Biol. Chem.* **283**, 14402–14410.
- Leonard, N.J., Sprecker, M.A., and Morrice, A.G. (1976). Defined dimensional changes in enzyme substrates and cofactors. Synthesis of *lin*-benzoadenosine and enzymatic evaluation of derivatives of the benzopurines. *J. Am. Chem. Soc.* **98**, 3987–3994.
- Liu, H., Gao, J., Lynch, S.R., Saito, D.Y., Maynard, L., and Kool, E.T. (2003). A four-base paired genetic helix with expanded size. *Science* **302**, 868–871.
- Liu, H., Gao, J., Maynard, L., Saito, Y.D., and Kool, E.T. (2004). Toward a new genetic system with expanded dimensions: size-expanded analogs of deoxyadenosine and thymidine. *J. Am. Chem. Soc.* **126**, 1102–1109.
- Liu, H., Gao, J., and Kool, E.T. (2005a). Size-expanded analogues of dG and dC: synthesis and pairing properties in DNA. *J. Org. Chem.* **70**, 639–647.
- Liu, H., Gao, J., and Kool, E.T. (2005b). Helix-forming properties of size-expanded DNA, an alternative four-base genetic form. *J. Am. Chem. Soc.* **127**, 1396–1402.
- Lu, H., He, K., and Kool, E.T. (2004). yDNA: a new geometry for size-expanded base pairs. *Angew. Chem. Int. Ed. Engl.* **43**, 5834–5836.
- Matray, T.J., and Kool, E.T. (1998). Selective and stable DNA base pairing without hydrogen bonds. *J. Am. Chem. Soc.* **120**, 6191–6192.
- Matray, T.J., and Kool, E.T. (1999). A specific partner for abasic damage in DNA. *Nature* **399**, 704–708.
- Matsuda, S., Henry, A.A., and Romesberg, F.E. (2006). Optimization of unnatural base pair packing for polymerase recognition. *J. Am. Chem. Soc.* **128**, 6369–6375.
- Minakawa, N., Kojima, N., Hikishima, S., Sasaki, T., Kiyosue, A., Atsumi, N., Ueno, Y., and Matsuda, A. (2003). New base pairing motifs. *J. Am. Chem. Soc.* **125**, 9970–9982.
- Mizukami, S., Kim, T.W., Helquist, S.A., and Kool, E.T. (2006). Varying DNA base-pair size in sub-angstrom increments: evidence for a loose, not large, active site in low-fidelity Dpo4 polymerase. *Biochemistry* **45**, 2772–2778.
- Morales, J.C., and Kool, E.T. (1998). Efficient replication of a DNA base pair between non-hydrogen-bonded nucleoside analogues. *Nat. Struct. Biol.* **5**, 950–954.
- Morales, J.C., and Kool, E.T. (2000). Functional hydrogen-bonding map of the minor groove binding tracks of six DNA polymerases. *Biochemistry* **39**, 12979–12988.
- Moran, S., Ren, R.X.-F., Rumney, S., and Kool, E.T. (1997). Difluorotoluene, a nonpolar isostere of thymine, codes specifically and efficiently for adenine in DNA replication. *J. Am. Chem. Soc.* **119**, 2056–2057.
- Moser, M.J., Marshall, D.J., Grenier, J.K., Kieffer, C.D., Killeen, A.A., Ptacin, J.L., Richmond, C.S., Roesch, E.B., Scherrer, C.W., Sherrill, C.B., et al. (2003). Exploiting the enzymatic recognition of an unnatural base pair to develop a universal genetic analysis system. *Clin. Chem.* **49**, 407–414.
- Parsch, J., and Engels, J.W. (2000). Synthesis of fluorobenzene and benzimidazole nucleic acid analogues and their influence on stability of RNA duplexes. *Helv. Chim. Acta* **83**, 1791–1808.
- Paul, N., Nashine, V.C., Hoops, G., Zhang, P., Zhou, J., Bergstrom, D.E., and Davisson, V.J. (2003). DNA polymerase template interactions probed by degenerate isosteric nucleobase analogs. *Chem. Biol.* **10**, 815–825.
- Piccirilli, J.A., Krauch, T., Moroney, S.E., and Benner, S.A. (1990). Enzymatic incorporation of a new base pair into DNA and RNA extends the genetic alphabet. *Nature* **343**, 33–37.
- Potapova, O., Chan, C., DeLucia, A.M., Helquist, S.A., Kool, E.T., Grindley, N.D.F., and Joyce, C.M. (2006). DNA polymerase catalysis in the absence of Watson-Crick hydrogen bonds: analysis by single-turnover kinetics. *Biochemistry* **45**, 890–898.
- Rawls, R. (2000). 'Synthetic Biology' makes its debut. *Chem. Eng. News* **78**, 49–53.
- Schweitzer, B.A., and Kool, E.T. (1994). Nonpolar aromatic nucleosides as hydrophobic isosteres of DNA nucleosides. *J. Org. Chem.* **59**, 7238–7242.
- Schweitzer, B.A., and Kool, E.T. (1995). Hydrophobic, non-hydrogen-bonding bases and base pairs in DNA. *J. Am. Chem. Soc.* **117**, 1863–1872.
- Sintim, H.O., and Kool, E.T. (2006). Remarkable sensitivity to DNA base shape in the DNA polymerase active site. *Angew. Chem. Int. Ed. Engl.* **45**, 1974–1979.

- Sismour, A.M., and Benner, S.A. (2005). The use of thymidine analogs to improve the replication of an extra DNA base pair: a synthetic biological system. *Nucleic Acids Res.* **33**, 5640–5646.
- Somoza, A., Chelliserrykattil, J., and Kool, E.T. (2006). The roles of hydrogen bonding and sterics in RNA interference. *Angew. Chem. Int. Ed. Engl.* **45**, 4994–4997.
- Switzer, C.Y., Moroney, S.E., and Benner, S.A. (1989). Enzymatic incorporation of a new base pair into DNA and RNA. *J. Am. Chem. Soc.* **111**, 8322–8323.
- Switzer, C.Y., Moroney, S.E., and Benner, S.A. (1993). Enzymatic recognition of the base-pair between isocytidine and isoguanosine. *Biochemistry* **32**, 10489–10496.
- Woods, K.K., Lan, T., McLaughlin, L.W., and Williams, L.D. (2003). The role of minor groove functional groups in DNA hydration. *Nucleic Acids Res.* **31**, 1536–1544.
- Xia, J., Noronha, A., Toudjarska, I., Li, F., Akinc, A., Braich, R., Frank-Kamenetskii, M., Rajeev, K.G., Egli, M., and Manoharan, M. (2006). Gene silencing activity of siRNAs with a ribo-difluorotoluyll nucleotide. *ACS Chem. Biol.* **1**, 176–183.
- Xia, G., Chen, L., Takashi, S., Fa, M., Schultz, P.G., and Romesberg, F.E. (2002). Directed evolution of novel polymerase activities: mutation of a DNA polymerase into an efficient RNA polymerase. *Proc. Natl. Acad. Sci. USA* **99**, 6597–6602.
- Yakovleva, L., Lai, J.S., Kool, E.T., and Shuman, S. (2006). Nonpolar nucleobase analogs illuminate requirements for site-specific DNA cleavage by vaccinia topoisomerase. *J. Biol. Chem.* **281**, 35914–35921.
- Yang, Z., Sismour, A.M., Sheng, P., Puskar, N.L., and Benner, S.A. (2007). Enzymatic incorporation of a third nucleobase pair. *Nucleic Acids Res.* **35**, 4238–4249.
- Zhang, X., Lee, I., and Berdis, A.J. (2005). The use of nonnatural nucleotides to probe the contributions of shape complementarity and pi-electron surface area during DNA polymerization. *Biochemistry* **44**, 13101–13110.
- Zhang, X., Lee, I., Zhou, X., and Berdis, A.J. (2006). Hydrophobicity, shape, and pi-electron contributions during translesion DNA synthesis. *J. Am. Chem. Soc.* **128**, 143–149.