Discovery, Vol. 52, No. 242, February 1, 2016

PERSPECTIVE

ISSN 2278–5469 EISSN 2278–5450

Discovery

Future of Designing Artificial Genetic Code

Tiwari RK[1,*], Ojha RP[2]

1. Assistant Professor, Department of Physics, DDU Gorakhpur University, Gorakhpur 273009, India
2. Professor, Reader, Department of Physics, DDU Gorakhpur University, Gorakhpur 273009, India

*Corresponding author: Department of Physics, DDU Gorakhpur University, Gorakhpur 273009, India, e-mail: drrkt@yahoo.com

Article History
Received: 7 December 2015
Accepted: 18 January 2016
Published: 1 February 2016

Citation
Tiwari RK, Ojha RP. Future of Designing Artificial Genetic Code. *Discovery*, 2016, 52(242), 182-185

ABSTRACT

The genetic bases of nucleic acid are constrained by the five natural bases A, T, G, C and U. A and G make pair with T(U) and C respectively. The Watson-Crick and Hoogsteen H- bindings are fundamental basis to explain their function in cell. The laboratory expansion of the genetic base pair would serve as the foundation of a semi-synthetic organism with an expanded genetic code. Recently it is proposed that H-Bonds in base pairs is not absolute requirement for DNA function rather the acceptable shape and packing energy is the key for cellular functions in molecular biology. The general principle to select these bases is that "a crucial position in the base has to be hydrophobic for enzymes to insert the base into DNA, yet it also has to accept hydrogen bonds if enzymes are replicating the strand". Here we discuss the possibilities and challenges in designing artificial genetic code using new unnatural base-pairs so that it can be used as to synthesize new organism.

Keywords: H-bond, Modified Hydrophobic base (MHB), Unnatural Base pair (UBP), and DNA function

1. INTRODUCTION

The genetic bases of nucleic acid are constrained by the five natural bases A, T, G, C and U. A and G make pair with T(U) and C respectively. Building a new base pair by modified hydrophobic base pair is an emerging technique in the field of genetic engineering (Amos, 8 May 2014). It is known that natural A-T and G-C are only nucleic acid base pairs inside nuclear genome and have all necessary genetic information of building organism. These two genetic base pairs are responsible for formation of 20 natural amino acids, necessary component of organism, through cellular mechanism in natural cell. Indeed DNA function is responsible for all events inside the cell.

Organism is defined by the information encoded by their genome (Denis A. Malyshev K. D., 2014). There are 20,500 genes in human beings almost same range as in mice. The human genome has nearly segmental duplication, nearly identical (99%), repeated section of other mammalian genomes ("Finishing the euchromatic sequence of the human genome"., 2004), (http://www.ensembl.org), (http://www.genome.gov/12011238). These sections may underline the creation of new primate-specific gene. The unit of gene is known as codon, consists of 3 nucleotide units are actually provide information for protein synthesis. It is possible to make artificial organism by using artificial base-pairs, if they mimic to DNA-function inside the cell. These new base-pairs may explore some new unnatural amino acids other than 20 natural amino acids. Similar kinds of effort are running in many laboratories around the world on the basis of theoretical and experimental facts (Kwok, 2012).

2. SEMI-SYNTHETIC ORGANISM IS NOW REALITY

Floyed Romesberg (www.scripps.edu/romesberg) has successfully replicated a sequence of nucleic acid using hydrophobic base pairs. His team found 60 bases out of about 3,600 combinations which are proved on this principle to some extent. Out of these only MNO2 and 5IICS, called X and Y base pair was found very near to goal (Young Jun Seo, 2009). This base-pair effectively shows transcription and PCR (Polymerase Chain Reaction) in laboratory. This success has excited scientific community.

It is known that the mechanism of formation of protein by flow of genetic information through DNA and RNA signals are important and essential mechanism of DNA functions in cell. Union of these nucleic acid functions are called central dogma in molecular biology. By these natural functions RNA transcript 20 natural amino acids in presence of different enzymes and other biomolecules. Today it is possible to localize different DNA-functions through florescent spectroscopy by help of electron microscope and computer. In DNA replication, DNA makes its replica in presence of some important enzymes (DNA polymerase). The starting points for DNA replication in gene are called ori. This DNA sequence is pyrimidine rich and combination of about 250 nucleotides. A gene may have a lot of ori. The essential energy to start this process comes from hydrolysis of high energy phosphate bond among three phosphate combined with each un-incorporated bases. The speed of replication in cell is about 749 per second. The accepted mechanism of starting DNA replication and transcription begins from binding of the base sequence with DNA by H-bonds (Christope Rocher, 2001). Formation of helix integrity through selectivity of hydrogen bonds in base pairs of DNA strands is basis of this model and it is considered as a necessity for DNA-functions (Rakesh K. Tiwari, 2011).

3. DESIGN OF ARTIFICIAL BASE PAIR MAY HAVE KEY FOR SYNTHESIS OF SUCH TYPE OF ORGANISM

Artificial base pairs are already used in detecting virus and have potential to be used as a drug by forming third strand in major groove of DNA. These bases have many uses as a bio-technological tool by the virtue of triplex formation property through hydrogen bond. (Maria Duca, 2008). But scientists are hopeful to code more information by using it, moreover even to make new artificial organism.

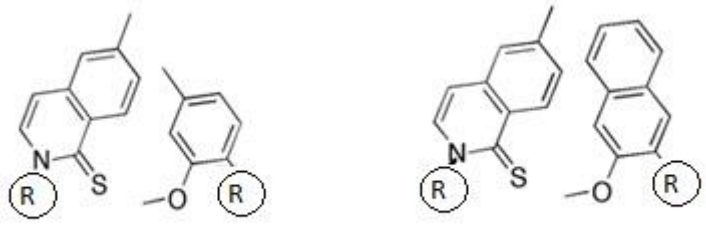
Repulsion of negative charges in phosphate in DNA strands, hydrogen bonds in base pairs and base stacking forces are prime factors of helix integrity in DNA structure. Using these facts, in 1996, Banner (Christopher Switzer, 1989) replaced natural bases with new orthogonal pairs i) iso-C - iso-G and ii) k – xanthosine in DNA and showed that polymerase recognize these pairs in a surprising way. They successfully showed replication and transcription of these sequences, moreover ribosomes, the cellular machines, could translate unnatural amino acid to growing protein (J. D. Bain, 1992). Base pairing through H-bond immersed as most crucial property of genetics. But soon, researchers encountered a problem that hydrogen atom of iso-G often morphed into different form and paired with T instead of iso-C, proved again superiority of nature's selectivity. Therefore genetic engineers started to consider other type of unnatural bonds instead of hydrogen bond in between base pairs.

4. HYDROGEN BOND IS NOT ABSOLUTELY REQUIRED CONDITION TO START CELLULAR MECHANISM

A hydrophobic base, diflurotoluene (designated F) (Barbara A. Schweitzer, 1995) was synthesized by replacing F atom instead of O atom in thymine (T). This mimic of T was capable to recognize DNA polymerase. During replication it faithfully made pair with adenine (A) and vice-versa and even could maintain integrity in DNA-function (Sean Moran, 1997). Therefore, DNA polymerase recognizes right shape of base and hydrogen bond is not absolutely required condition to start replication in nucleic acid, relies on harnessing hydrophobic and packing forces between the nucleobase analogs. It is concluded on the basis of hypothesis "a right

shape is right key", although this conclusion emerges with many new questions in central dogma in molecular biology. The chemical property of Diflurotolune (F) is hydrophobic and makes pair with adenine (A) and perhaps fulfils above conditions.

Figure 1 Base pair MMO2 –SICS and DNAM-5SICS



Romesberg (Denis A. Malyshev K. D., 2014) (www.scripps.edu/romesberg)used similar principle and screened three thousand six hundred combinations of sixty bases for pair which were capable to do DNA replication effectively. Out of these the base pair MMO2 and SICS (Figure 1) was most appropriate for pairing as they are hydrophobic for enzymes to insert the base into DNA and accept hydrogen bond if enzymes are replicating the strand, two contradictory characteristics. Therefore they have to show hydrophobic characteristics to maintain helix integrity and hydrophilic to do central dogma function, just like switching device at the key position of mechanism of central dogma. The success of this pair excited the scientific community and elaborated it as alien DNA (Amos, 8 May 2014), (Callaway, 2014).

Another group in Japan (Rie Yamashige, 2012) has designed the shape of bases by adding electron-rich chemical group that repel the corresponding part of natural bases to balance the force just like natural base-pair. They could replicate an unnatural hydrophobic base pair Ds and Diol1-Px with rate of more than ninety nine percent fidelity. Similar types of work are running in many laboratories around the world (Kwok, 2012)

Unnatural bases still have a lot to prove, however Researchers haven't shown more than four bases in row during replication in-vitro (Kwok, 2012).

Over all the general principle to select these bases is (Rie Yamashige, 2012) that a crucial position in the base has to be hydrophobic for enzymes to insert the base into DNA, yet it also has to accept hydrogen bonds if enzymes are replicating the strand".

5. EXPENDED CODE'S IN CELLULAR MECHANISM

In the refining process of these bases the base pair DNAM-D5SICS was synthesized by modifying base pair MMO2-SICS (Denis A. Malyshev K. D., 2012). It showed faithfulness in PCR-amplification and transcription in-vitro. Moreover triphosphates of both DNAM and D5SICS were efficiently imported by an exogenously expressed algal nucleotide triphosphate transporter into Escherichia coli, and that the endogenous replication machinery used them to accurately replicate a plasmid containing DNAM-D5SICS. They showed that UBP was not efficiently excised by DNA repair pathway. Neither the replication of the UBP nor presence of unnatural triphosphate introduced growth burden. This bacterium is claimed as first organism by an expanded nucleic code (Denis A. Malyshev K. D., 2014). This laboratory bacterium was made in following laboratory conditions. i) The UBP must be stable in the presence of pathways that maintain DNA integrity. ii) The unnatural nucleoside triphosphates must be available inside the cell and endogenous polymerases must be able to use the unnatural triphosphates to faithfully replicate DNA within the complex cellular milieu. The base pair DNAM and D5SICS is large hydrophobic and does not form hydrogen bond with pair (Figure 1). The DNA duplex formed by this base pair is partially intercalated. This structure is very similar to DA-DA mismatch and does not follow the accepted mechanism of DNA replication.

In the crystal structure of DNA polymerase with this pair it was observed that they maintain helix integrity and accept specific planer Watson-Crick geometry and reject others identical to natural base pair (Karin Betz, 2013). The phosphate-phosphate distance in UBP duplex in presence of DNA polymerase is about 11 A, almost similar to natural DNA. And without polymerase it is 9.1 A less than 0.4 A to natural DNA. Therefore shape selectivity and energy selectivity of DNA strand are two important key factors for entry in cellular mechanism if unnatural bonds mimic to natural bonds (Denis A. Malyshev K. D., 2012). There electivity produces selectivity of DNA polymerase in replication, the first challenge to acceptance of foreign molecule in nucleic acid function. After that, nucleic acids

have to retain the unnatural base pairs during central dogma for very large time domain. This new code may encode new amino acids totally different than twenty natural amino acids probably encode more information.

Therefore synthesis of E. coli by using single pair of UNB in DNA strands out of millions a new beginning of creation of new organism by foreign DNA in the field of genetic engineering. This technique may be used to synthesize more complex systems. However expansion of an organism's genetic code presents new and unprecedented challenges. Perhaps scientist would extend this technique to make an organism by using only UNB DNA strand. Moreover, the careful examination of acceptance of foreign molecule during central dogma may provide data to explore UNB as a drug.

ACKNOWLEDGEMENT

Authors acknowledge UGC, New Delhi for financial support.

REFERENCES

1. http://www.ensembl.org.
2. http://www.genome.gov/12011238.
3. "Finishing the euchromatic sequence of the human genome". (2004). *Nature 431*, 931–945.
4. Highly specific unnatural base pair systems as a third base pair amplication. (2012, Vol 40, No. 6). *Nuclic Acid Research*, 2793-2806.
5. Amos, J. (8 May 2014). *Semi-synthetic bug extends 'life's alphabet'*. Science correspondent, BBC News.
6. Anna K. Shchyolkina, E. N. (1994). The R-form of DNA does exist. *FEBS letters, 339*, 113-118.
7. Barbara A. Schweitzer, E. T. (1995). Hydrophobic, Non-Hydrogen-Bonding Bases and Base Pairs in DNA. *J. Am. Chem. Soc., 117 (7)*, 1863–1872.
8. Callaway, E. (2014). First life with 'alien' DNA. *Nature*.
9. Christope Rocher, R. D. (2001). Initiation of DNA replication by DNA polymerases from primers forming triple helix. *Nucleic Acids Research, Vol. 29, No. 16*, 3320-3326.
10. Christopher Switzer, S. E. (1989). Enzymatic incorporation of a new base pair into DNA and RNA. *J. Am. Chem. Soc.*, 8322–8323.
11. Denis A. Malyshev, K. D. (2012). Efficient and sequence-independent replication of DNA containing a third base pair establishes a functional six-letter genetic alphabet. *PNAS, Vol 109*, 12005-12010.
12. Denis A. Malyshev, K. D. (2014). A semi-syntheic organism with an expanded genetic alphabet. *509*.
13. J. D. Bain, C. S. (1992). Ribosome-mediated incorporation of a non-standard amino acid into a peptide through expansion of the genetic code. *Nature 356,*, 537 - 539.
14. Karin Betz, D. A. (2013). Structural Insight into DNA Replication without Hydrogen Bonds. *J. Am. Chem. Soc. 135*, 18367-18643.
15. Kwok, R. (2012). DNA's New Alphabet. *Nature, Vol 491*, 516-519.
16. Maria Duca, P. V. (2008). The triple helix: 50 years later, the outcome. *Nuclic Acid Research, Vol 36*, 5123-5138.
17. Rakesh K. Tiwari, R. P. (2011). *Molecular Modeling of unusual DNA*. LAP Lambert Academic Publishing.
18. Rie Yamashige, M. K. (2012). Highly specific unnatural base pair systems as a third base pair for PCR amplification. *Nuclic Acid Research Vol 40*, 2793-2806.
19. Sean Moran, R. X.-F. (1997). Difluorotoluene, a Nonpolar Isostere for Thymine, Codes Specifically and Efficiently for Adenine in DNA Replication. *J. Am. Chem. Soc.*, 2056-2057.
20. *www.scripps.edu/romesberg*. (n.d.). Retrieved from Institute, The Scripps Research.
21. Young Jun Seo, S. e. (2009). Transcription of an Expanded Genetic Alphabet. *J. Am. Chem. Soc. 131(14)*, 5046-5047.