

Word Design for Biomolecular Information Processing

J. Ackermann and F.-U. Gast^a

Fraunhofer Gesellschaft, Schloss Birlinghoven, D-53754 Sankt Augustin

^a Justus-Liebig-Universität Gießen, Institut für Anorganische und Analytische Chemie,
Schubertstraße 60, Haus 16, D-35392 Gießen

Reprint requests to Dr. J. A.; Fax: -141511, E-mail: joerg.ackermann@gmd.de

Z. Naturforsch. **58a**, 157 – 161 (2003); received October 10, 2002

The design of DNA sequences plays a fundamental role for many biomolecular applications and is one of the most important theoretical tasks to fathom the potential of molecular information processing. Optimization strategies have been based on the model of stiff “digital” polymers by counting the number of base mismatches (Hamming distance and related distances). In this work we show the limitation of such a combinatorial approach because of the ability of DNA to build more complex structures. We develop a model platform to optimize word sets according to all possible secondary structures occurring for the relevant word-word interactions. The fidelity of the hybridization reactions can be improved significantly and as an example of a set of 24 words of 16-mers we show that the optimal set has unique physical properties, such as binding energy, melting temperature, and G+C content.

Key words: DNA Library; DNA Computing; Hybridization; Folding; Partition Function.