Random Number Generation by a Two-Dimensional Crystal of Protein Molecules
Yasuhiro Ikezoe, Song-Ju Kim, Ichiro Yamashita, and Masahiko Hara

Langmuir, Article ASAP • Publication Date (Web): 10 March 2009
Downloaded from http://pubs.acs.org on March 22, 2009

More About This Article

Additional resources and features associated with this article are available within the HTML version:

• Supporting Information
• Access to high resolution figures
• Links to articles and content related to this article
• Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML
Random Number Generation by a Two-Dimensional Crystal of Protein Molecules

Yasuhiro Ikezoe,*† Song-Ju Kim,‡ Ichiro Yamashita,§‖ and Masahiko Hara†‡§

†Flucto-Order Functions Asian Collaboration Team, RIKEN Advanced Science Institute, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan; ‡Flucto-Order Functions Asian Collaboration Team, RIKEN Advanced Science Institute, Fusion Technology Center, Hanyang University, 17 Haengdong-dong, Seongdong-gu, Seoul 133-791, Korea; §Advanced Technology Research Laboratories, Panasonic Corporation, 3-4 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0237, Japan; ‖Graduate School of Materials Science, Nara Institute of Science and Technology, 4259 Sagamihara, Midori-ku, Yokohama 226-8502, Japan.

Received January 6, 2009. Revised Manuscript Received February 26, 2009.

We discuss 2D and binary self-assemblies of protein molecules using apo-ferritin and holo-ferritin, which have identical outer-shell structures but different inner structures. The assemblies do not show any phase separation but form 2D monomolecular-layer crystals. Statistical analyses showed a random molecular distribution in the crystal where the molar ratio was conserved as it was in the solution. This molecular pattern is readily prepared, but it is neither reproducible nor predictable and hence can be used as a nanometer-scale cryptographic device or an identification tag.

Introduction

Random number sequences play key roles in our daily lives in a highly complicated telecommunication environment and are an important element in computer simulations in the natural sciences, mathematics, economics, and so on. To ensure security in the information technology or the validity of simulation results, it is important to use an unpredictable and aperiodic random number sequence. Many kinds of methods have been known to obtain such sequences. Some are based on physical phenomena such as radiation,† lasers,‡§ and quantum effects.§ Others are based on mathematical algorithms such as the linear congruential algorithm,‖ the Blum−Blum−Shub algorithm based on prime factorization,† the Mersenne twister algorithm based on the maximum length sequence,§ the cellular automaton,‖ and so on. Although the methods based on physical phenomena should be random in principle, unfortunately the observation processes or instruments adopted often accompany nonrandom factors. However, the methods based on algorithms implemented on digital computers produce intrinsically deterministic and periodic sequences. Therefore, these random number sequences are called pseudorandom numbers.

Here we report a new random number generator based on the self-assembly of molecules. The molecule used here is ferritin, a protein molecule that has a quasi-spherical hollow shell structure (12 nm outer diameter and 7 nm inner diameter) composed of 24 subunits (~450 kDa) and is capable of storing both iron hydroxide in vivo and a variety of inorganic compounds, such as magnetic materials1−15 or semiconductors,16,17 catalysts,18−20 and diagnostics.21 Ferritin is roughly classified into two species: apo-ferritin (AF), without any core materials in the molecular cage, and holo-ferritin (HF), which stores an inorganic nanoparticle inside. Except for the inner difference, these molecules have exactly identical chemical structures. Now we consider the difference in the intermolecular interaction between these molecules. The presence of the nanoparticle in HF would induce the conformation change and electric field modification within the protein shell. However, such a discrepancy between HF and AF should be relaxed along the outward direction in the thick protein shell. Therefore, it

DOI: 10.1021/la9000413

Langmuir XXXX, XXX(XX), 000–000
would mostly disappear at the molecular surface. The intermolecular interaction is governed mainly by the outermost surface of the molecules, and thus there would be a small discrepancy in the interaction between HF and AF. In this letter, we show some 2D crystals of ferritin molecules obtained from solution including both HF and AF molecules in a certain ratio. Our experiments revealed no phase separations. We obtained 2D monodomain-layer crystals where the molar ratio was conserved as it was in the solution. In addition, statistical analyses found the microscopic molecular distribution to be random. Especially for crystals obtained from solution including equimolar AF and HF, we obtained a uniform random number sequence by indexing 0 and 1 for AF and HF, respectively. We know of many excellent reports on the binary self-assembly of nanometer-scale or micrometer-scale particles in 3D or 2D in which the ordered structures, superlattices, or phase separations have been well discussed. However, we first discuss the randomness in the ordered structure in the binary assembly of nanoparticles, and we show a fundamental Monte Carlo simulation utilizing the random molecular distribution.

Experimental Section

In our experiments, we used a genetically synthesized ferritin called N1-LF. The HF molecule with an indium oxide nanoparticle was obtained from the AF molecules using a method previously reported (Supporting Information S1). N1-LF has a more hydrophobic surface than the native ferritin, which facilitates crystallization as previously reported. The crystallization procedure is summarized as follows. A 300 μL droplet of an aqueous solution of N1-LF (2 μg/mL) is formed on a hydrophilic SiO₂ surface surrounded by a hydrophobic surface and then is slowly dried. The adsorbed molecules on the water surface are concentrated in the vicinity of the perimeter of the droplet by outward convective flow during the drying process, and eventually the 2D crystal grows at the water surface perpendicularly to the contact line. After the water is entirely evaporated, the 2D crystal is left on the SiO₂ surface (Supporting Information S2). The random molecular mixture in the 2D crystal cannot be realized unless the molecular distribution in the solution or at the water surface is also random. Therefore, for AF and HF molecules to be mixed sufficiently evenly on the microscopic scale, the solution was stirred with a magnetic stirrer for more than 30 min. Under some unfavorable conditions for crystallization such as the fast evaporation of water, only a disordered or 3D aggregation of molecules was observed, in which we can no longer discuss the molecular distribution. However, whenever the crystal was formed, we obtained the random molecular distribution shown in the following section. The crystal obtained was robust, and its structure has not been broken for a few weeks as long as it was kept in a dry environment. However, when the crystal was exposed to a humid atmosphere, dew drops readily destroyed the crystal because it was fixed by physical adsorption onto the Si substrate, not by covalent bonds. It has been known that some thermal processes, for example, heating the 2D crystal to remove the protein shell, are effective at fixing inorganic nanoparticles directly onto a Si substrate. The details of the crystallization conditions are described in our previous paper. Figure 1 shows a scanning electron microscope (SEM) image of the 2D molecular crystal obtained from the solution, including equimolar HF and AF molecules. The scale bar is 50 nm, and the white spots are indium oxide nanoparticles in the HF's; one white spot corresponds to one HF molecule. The fast Fourier transform (FFT) analysis of this image, as shown in the lower left, reveals a 6-fold-symmetrical pattern with a lattice constant of approximately 12 nm that is equal to the molecular size. This FFT pattern is similar to that of the hexagonally close-packed (HCP) molecular crystal obtained from the pure HF solution, and hence it is thought that the HF molecules in Figure 1 are also on the lattice points of the HCP molecular crystal. Interestingly, the FFT pattern showed no other patterns except for the above 6-fold one, which indicates that the crystal has no superlattice structures. However, many dark lattice points were simultaneously found in this image. We can surmise that these dark points are occupied by the AF molecules because the AF molecules are invisible in the SEM observation as a result of the absence of inorganic elements. To verify this hypothesis, we investigated other crystals obtained from the solution including HF and AF molecules in a molar ratio of N/1. We call the crystal simply the N/1 crystal hereafter. When the N/1 crystal was fabricated, the total concentration of ferritin molecules was adjusted to 2 μg/mL.

Results and Discussions

Figure 2a–d shows a series of SEM images of the N/1 crystal. The images are the results under the conditions where the value of N is 1, 2, 3, and 4, respectively. The yellow spots in the right-hand image of each Figure correspond to the dark lattice points of the left-hand image. These yellow spots and the HF molecules cover the whole area without any defects in the hexagonally symmetric manner. Besides, the number of dark spots obviously decreases as the value of N increases. We also investigated the ratio of the number of HF molecules to whole molecules using much larger single-crystal domains composed of more than 1000 molecules (Supporting Information, S3). As a result, the number ratios of HF in the N/1 crystal with N = 1, 2, 3, and 4 were 0.498, 0.642, 0.739, and 0.818, respectively. These results are nearly equal to the molar ratio in the solution, N/(N + 1). Here, we should emphasize that the dark spots are not vacancies. If there is a vacancy, then it looks very bright because of the emission of the secondary electrons from the bare silicon surface (Supporting Information, S4). From all of these observations and analyses, we can reasonably conclude that the crystal was composed of HF and AF molecules. The absence of phase separation implies that the intermolecular interaction is identical between molecules. Under such a condition, the probability distribution of the number of HF


DOI: 10.1021/la9000413 Langmuir XXX, XXX(X), 000–000
molecules surrounding one HF molecule should obey the binomial distribution as described below.

\[ P_{N,k} = C_6^k R_N^k (1 - R_N)^{6-k} \]  

Here, \( P_{N,k} \), \( C_6^k \), and \( R_N \) are the probability that the number of HF's surrounding one HF is \( k \) in the \( N/1 \) crystal, the binomial coefficient expressed as \( 6!/(k!(6-k)) \), and the number ratio of the HF molecules to the whole molecules in the \( N/1 \) crystal, respectively. Figure 3a–d shows a probability distribution series with respect to the number of HF's surrounding one HF. The blue diamonds and red circles show calculated data from eq 1 and experimental data, respectively. In all cases, the experimental data are in good agreement with the calculated data. These results serve as evidence that the intermolecular interactions are identical, which means that the molecular distribution is truly random and has the potential to generate a random number sequence.
one of the results of the statistical test—the poker test, in which the frequency distribution of the 16 possible 4-bit values (from 0 to 15) is shown. The calculated data were obtained under the condition that 4-bit values appear equally. In addition to the results of Seq1–Seq3, another result with 4-bit values is simultaneously shown. These values were obtained from the molecular matrix with $4 \times 2$ molecules shown in the dotted-line square in Figure 4a. This is a unique point in our system; the crystal has two dimensions, and the random number can also be generated in a 2D manner. These experimental data show small, unbiased deviations from the calculated data. Table 1 shows all of the results of the statistical tests. The requirement for each test was recalculated to adjust to our 2500-bit sequence. From these results, it was found that all of the sequences—Seq1, Seq2, Seq3, and the $(2 \times 2)$ matrix-based sequence—pass the requirements and can be concluded to be random. Strictly speaking, no bit sequences that pass the statistical tests can be concluded to be random. We should conclude that non-randomness was not detected for any of the bit sequences. However, in this letter, we call the sequence simply a random number sequence in order to avoid the above complex expression.

Let us now look at the term “random” from the point of view of materials science. For example, a phrase such as “random mixture of molecules” is often found on a variety of occasions; however, whether it is truly random has been discussed very little. Considering the different interactions among different species, a truly random mixture is impossible. A binary or multicomponent mixture of molecules, atoms, or nanoparticles basically shows phase separation or an alternative ordered structure. However, we first and quantitatively discussed the random molecular distribution in the crystal on the molecular scale with the aid of statistical tests. In the present case, randomness was realized because the differences in intermolecular interactions between the different kinds of molecules were removed by the unique molecular structure: a thick protein shell that conceals the inner difference. Furthermore, another important and notable point here is that the molecular assembly retains the HCP crystal structure regardless of the molar ratio. This result also explains that the AF and HF molecules are indistinguishable from each other by their outer surface characteristics.

Finally, we show that the obtained random number can be used effectively in computer simulations. We chose a simple and fundamental simulation: a problem regarding squaring a circle. The expected convergent value here is $\pi$ (Supporting Information, S6). We prepared four 16-bit random number sequences from the above 1D 0 and 1 sequences and a 2D matrix. The average and standard deviations of the results of 100 experiments are shown in Table 2, where all of the average values obtained are within a margin of 1% error compared to the true $\pi$ value. The probability distribution in the simulation should be described simply by a binomial distribution with the probabilities $p$ and $q (= 1 - p)$. Hence, the expected value, $E$, and the standard deviation, $\sigma$, derived from $N$ points are $p$ and $(pq/N)^{1/2}$, respectively. Under the condition that $p$ and $N$ are equal to $\pi/4$ and 2500, respectively, the value of $\sigma/E$ is approximately 0.01 ($= 1\%$). Therefore,

---

Footnotes:


Although the four statistical tests used here have been omitted from the document, they are still used frequently to estimate randomness of a bit sequence.


each experimental error is in good agreement with the calculated one. From this analysis, we can conclude that the simulation was successfully executed and, as a result, the simulated $\pi$ value reasonably converged to the true $\pi$ value. These results indicate that the random number sequence obtained from the binary protein crystal has no unfavorable bias and is capable of practical use.

Conclusions

We have successfully fabricated 2D monomolecular-layer crystals composed of binary kinds of protein molecules—apo-ferritin (AF) and holo-ferritin (HF)—and demonstrated a nanometer-scale random number generator (RNG) originating from the random molecular distribution in the 2D binary crystal. RNGs based on physical phenomena or mathematical algorithms are well known. However, our study is the first demonstration of RNGs based on chemical characteristics.

We not only investigated the randomness of the molecular distribution from some statistical tests but also showed the validity of the use of our RNG for a Monte Carlo simulation. We have so far succeeded in fabricating a single crystal with a size of approximately 100 $\mu$m$^2$, incorporating one million molecules. Now let us consider that such a large crystal consists of equimolar AF and HF molecules. The number of molecular distribution patterns in the crystal becomes astronomically large because it reaches the $2^{10^6}$. Although the binary 2D crystal can be readily prepared, it is neither reproducible nor predictable. Therefore, this molecular pattern can be regarded as a kind of microscopic fingerprint and thus would be used for a nanometer-scale cryptographic device or an identification tag.

Acknowledgment. We thank K. Tamada, T. Matsui, and N. Matsukawa for discussions and comments regarding this work. This study was supported in part by the Leading Project of the Ministry of Education, Culture, Sports, Science, and Technology, Japan.

Supporting Information Available: Description of ferritin, 2D crystals, statistical tests, and a Monte Carlo simulation. This material is available free of charge via the Internet at http://pubs.acs.org.

| Table 1. Statistical Analyses of the Randomness of the Molecular Distribution |
|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
| requirement of X            | monobit test                | poker test                  | runs test                   | long-run test               |
| seq1                        | $1185 < X < 1315$           | $4.60 < X < 32.80$          | $1186 < X < 1314$           | $X < 17$                    |
| seq2                        | 1245                        | 13.64                       | 1251                        | 16                          |
| seq3                        | 1245                        | 18.81                       | 1245                        | 12                          |
| $2 \times 2$ matrix         | 1245                        | 27.67                       | 1245                        | 16                          |
|                             |                             | 15.44                       | N. A.                       | N. A.                       |

Table 2. Results of Monte Carlo Simulations

<table>
<thead>
<tr>
<th>obtained $\pi$ value</th>
<th>average</th>
<th>standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>seq1</td>
<td>3.137</td>
<td>0.022</td>
</tr>
<tr>
<td>seq2</td>
<td>3.151</td>
<td>0.018</td>
</tr>
<tr>
<td>seq3</td>
<td>3.120</td>
<td>0.019</td>
</tr>
<tr>
<td>$4 \times 4$ matrix</td>
<td>3.150</td>
<td>0.019</td>
</tr>
</tbody>
</table>

Langmuir XXXX, XXX(XX), 000–000  DOI: 10.1021/la9000413