

# Materials Design and Prediction Using Computational Science

Sumitomo Chemical Co., Ltd.  
Advanced Materials Research Laboratory  
Masaya ISHIDA  
Yasuyuki KURITA  
Akiko NAKAZONO

Recent progress in simulation and computer technology has helped the growth of computational materials science. Computer aided design of materials with high functionality based on theories of materials science has become an important and essential part of research and development in our company. Here, we introduce our work on the computational analysis of inorganic fluorescent and organic semiconductor materials. Also, we show how recent progress in this field can influence the directions of our research in the future.

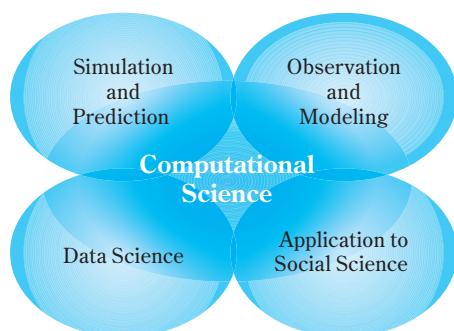
This paper is translated from R&D Report, “SUMITOMO KAGAKU”, vol. 2015.

## Introduction

Computational science has, in recent years, been referred to as the “third science” in addition to experimental and theoretical approaches. It consists of four areas, as shown in Fig. 1.<sup>1)</sup>

Sumitomo Chemical has positioned computational science as one of the useful methods for our various research and development. It has been used to present the direction of research and development, such as in proposing indices for efficient screening of chemical compounds and proposing structures of high-performance materials.

As described in later sections, we at Sumitomo Chemical have a long history and experience in molecular design to predict molecular structures and functions.



**Fig. 1** Schematic overview of four functions of computational science pictured by the author based on reference 1

Initially, we applied it to the design of agrochemicals. However, along with technology innovation and the significant evolution of computers, it has developed into materials design, as intended for larger, more complex structures (hereinafter referred to as “computational materials science”, which corresponds to “Simulation and Prediction” in Fig. 1), and it now occupies an important position in our materials development.

The situation surrounding computational science is now undergoing significant change. Particularly, since 2000 many computational science software programs have become compatible with parallel computation due to the dissemination of general-purpose parallel computers. This trend has accelerated the speed-up of scientific computation and the expansion of its scale in industry. Parallel to such a trend, the industrial use of domestic large-scale computers has been promoted, thus accelerating the industrial use of the “Earth Simulator” of the Japan Agency for Marine-Earth Science and Technology, as well as “TSUBAME” of the Tokyo Institute of Technology. Moreover, September 2012 brought the full operation of a supercomputer called “K” (hereinafter referred to as the K computer) developed by the National Research and Development Agency RIKEN, thus allowing industrial users to take advantage of its tremendously improved computational capability.<sup>2)</sup> We have also used the K computer since its initial operation in order to establish the fundamental techniques for the future of large-scale calculation of material structures

and electronic states. In this paper we will introduce our computational approaches from molecular design to materials design, presenting practical examples, and will discuss how computational materials science is applied in industrial materials development and the future perspectives.

This paper consists of as follows: The next section will review the history of computational science in Sumitomo Chemical as well as some practical examples of computational materials science in our Advanced Materials Research Laboratory in Tsukuba. Lastly, we will review the recent trend in computational materials science and present our goals in computational science in our company based on the above reviews and examples.

## Our Approach to Computational Science

### 1. History of Computational Science in Sumitomo Chemical

The origin of computational science in our company is research in the structure-activity relationship of agrochemicals, which began at the laboratory in the Takarazuka district during the 1970s. The Hansch-Fujita method<sup>3)</sup>— which is used for multiple regression analyses between agrochemical activities (insecticidal, fungicidal, herbicidal activities) and the hydrophobic, electrical and steric parameters of a substituent group that substitutes for a hydrogen atom on the parent chemical structure of an agrochemical — was applied to the development of agrochemicals such as the herbicide named Sumiherb<sup>®</sup>.<sup>4)</sup> When there are multiple substitution sites, it is difficult to find the combination of substituents that will optimize the agrochemical activities. In order to solve this issue, Takayama and Yoshida et al. have developed a computer program named PREHAC, which puts parameters stored in the substituent-parameter database into the correlation equation obtained from a structure-activity relationship analysis, predicts the activities of various substitution combinations, and outputs the activities in descending order.<sup>5)</sup> It can be said that the PREHAC is a pioneering software program for materials informatics which is now attracting global attention. Subsequently, during the 1980s Sumitomo Chemical jointly developed the computer chemistry system named ACACS<sup>6)</sup> together with NEC Corporation by expanding PREHAC and other similar programs. The ACACS system was then applied not only to structure-activity correlation analysis of agrochemicals<sup>7)–11)</sup>, but also to characteristics analysis of synthetic dyes and the

like. In 1990, Yoshida, one of the ACACS developers, moved to the Tsukuba Laboratory with Zempo (currently working as a professor at Hosei University) et al., and further expanded the scope of computer chemistry to encompass materials design and device design, and consequently it evolved to computational materials science.<sup>12)</sup>

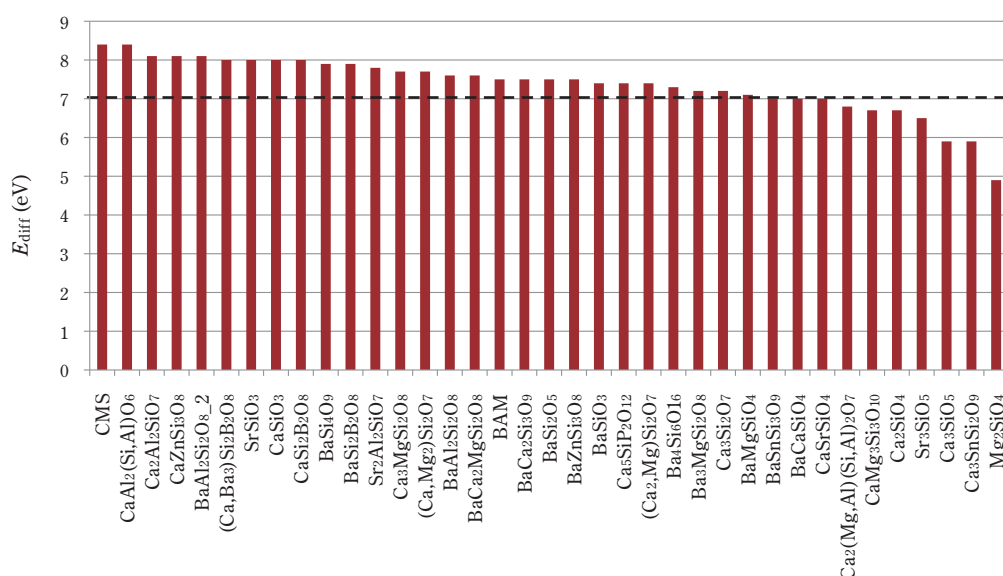
### 2. Examples of Our Approach to Computational Materials Science

The process of industrial materials development consists not only of exploring molecules and materials having outstanding characteristics and functionalities as raw materials, but also of a series of efforts to optimize various parameters of the material so that it will demonstrate the desired characteristics and functions in systems such as electric devices and launch it to the market as a final product. This process is closely related to the hierarchical structure of materials. For example, in electronics-related materials, atoms and molecules aggregate and form a more complex higher-order structure, then the aggregates are incorporated into devices and finally the device properties are demonstrated.

One of the required roles of computational materials science is to improve the efficiency of research and development through the following activities: revealing the nature of phenomena derived from various forms of matter during such a successive development stages; based on understanding of the nature, inspiring new ideas through the prediction of characteristics and functions that each material and its aggregate structure may demonstrate; and contributing to the creation of new or highly functional materials and innovative processes.

This section will introduce four practical examples of our approach to computational materials science. First, we will introduce the computational material search using optical band gap prediction for inorganic phosphors, followed by the results of investigation into the methodologies regarding the conduction characteristics and optical properties of organic semiconductor materials through the use of TSUBAME of the Tokyo Institute of Technology and the K computer. Lastly, we will introduce the result of investigation into the solubility prediction of chemical compounds, such prediction being crucial for the coating process.

- (1) Computational Material Search Using the Optical Band Gap Prediction of Inorganic Phosphors  
Inorganic phosphors are used for plasma displays and



**Fig. 2** The host excitation energies of various non-doped silicate hosts, estimated by using the energy difference ( $E_{diff}$ ) between the VB top and the first maximum intense d band calculated using the VASP code<sup>17)</sup>. The scissors operator of 1.4 eV for all unoccupied bands is uniformly applied. The dashed line at 7 eV indicates the uncorrected value of  $E_{diff}$  for CMS. © IOP Publishing. Reproduced by permission of IOP Publishing. All rights reserved.

scintillation materials, as well as in lighting such as fluorescent lights. The luminescent colors of red, green and blue can be obtained from inorganic phosphors, depending on the kinds of doped rare-earth elements and their valences. The inorganic phosphors introduced here are  $\text{Eu}^{2+}$  doped silicate materials, which show blue luminescent color due to vacuum UV excitation (wavelengths of 147nm and 172nm).<sup>13)</sup> It is generally considered that during the light absorption/emission process of inorganic phosphors, once a host crystal of an oxide or the like has absorbed vacuum UV rays, its energy is transferred to a doped rare-earth element, thereby causing the element to illuminate.<sup>14)</sup> Therefore, from the perspective of material search, the question of how to find host crystals having the optical band gap, which corresponds to two types of the excitation wavelength of vacuum UV rays, is crucially important.

The optical band gap of a host crystal is calculated using the first-principles band calculation method. Although the calculated optical band gap is often smaller than the experimental value due to the approximation contained in the original theory, it is well known that its qualitative tendency derived from the difference in materials is often reproduced.<sup>15)</sup> In our search for silicate host crystals suitable for the excitation wavelengths of 147nm and 172nm, we calculated the optical band gaps of various types of silicates using the band calculation.<sup>16)</sup>

In this study we extracted approximately 30 crystal structure data for silicates having different compositions from the inorganic crystal structure database, then we carried out a structure-optimization calculation via the band calculation in order to obtain the optical band gap. The optical band gap was estimated using the energy difference ( $E_{diff}$ ) between the valence band edge of the density-of-state and the first peak of the d band. **Fig. 2** shows the results. Because it is known that CMS ( $\text{CaMgSi}_2\text{O}_6$ ) (shown at the far-left of **Fig. 2**) has 147nm absorption, all the calculated optical band gaps are shifted by adding the energy difference between the experimental value 8.4 eV(147nm) and the calculated value of CMS.

The aluminate host crystal BAM ( $\text{BaMgAl}_{10}\text{O}_{17}$ ) known as a commercialized blue color inorganic phosphor has 172nm absorption, and the calculated value of the optical band gap of BAM corresponds well to its excitation energy (7.2eV). Based on this calculation result, we predict that  $\text{BaCa}_2\text{MgSi}_2\text{O}_8$  and  $\text{BaCa}_2\text{Si}_3\text{O}_9$ , which are present in the vicinity of BAM, should be promising candidate silicates corresponding to 172nm excitation. In fact,  $\text{BaCa}_2\text{MgSi}_2\text{O}_8$  was developed as an inorganic phosphor corresponding to 172nm excitation. The calculation result is consistent with the experiment.

We use the term “computational material screening” for the method of reducing the number of candidates by calculating characteristics of chemical compounds

and materials using theoretical methods. Currently, although there is an issue of calculation accuracy, it can be expected that if computational material screening is applied to an even broader search for chemical compounds and materials, it will improve the efficiency of material search. For more detailed analysis of the mechanism of photo-excitation and luminescence, please see the reference.<sup>16)</sup>

## (2) Design of Organic Semiconductor Materials

Organic EL (electroluminescence) displays, transistors and solar cells can be manufactured through relatively inexpensive process methods such as inkjet and printing methods. They have garnered attention as next-generation devices having favorable characteristics such as light weight, thinness and flexibility. Organic semiconductors are the key materials in the above devices.

The important factors when an organic semiconductor demonstrates its functions are carrier (electrons or holes) injection from the electrodes and carrier conduction within the organic semiconductor. Additionally, light emission (which occurs due to the recombination of electrons and holes) is important for organic EL displays, and the electron-hole pair generation from the excited states (which occurs due to light absorption) is crucial for organic solar cells. Furthermore, it is known that the characteristics (such as crystallinity) of the aggregated structure of organic-semiconductor molecules greatly affect the above process. The following

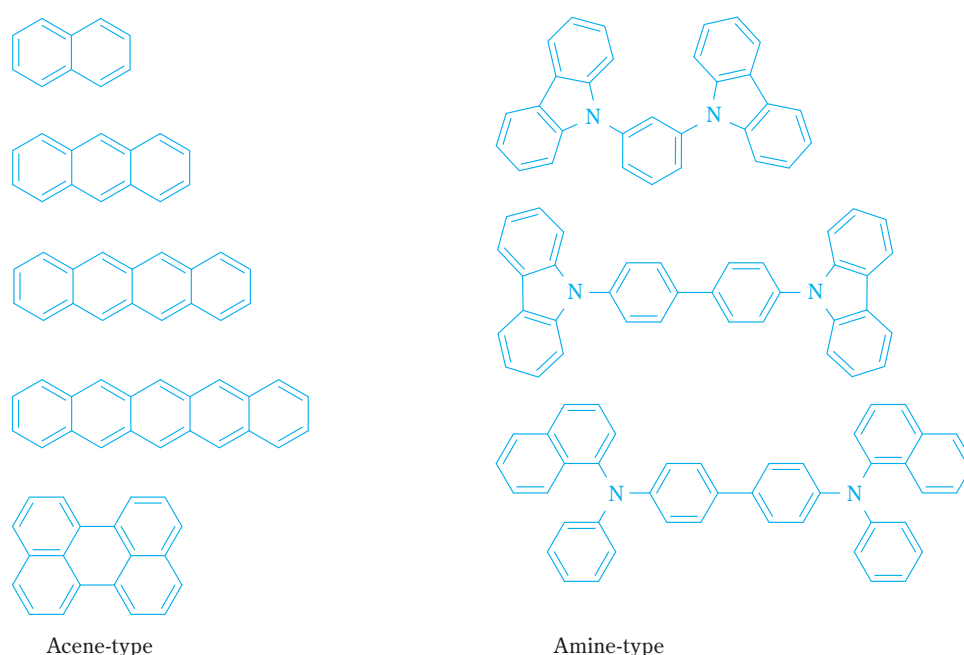
describes our investigation of the hole-injection property and hole conductivity.

### (i) Hole-Injectability from Electrode to Organic Semiconductor

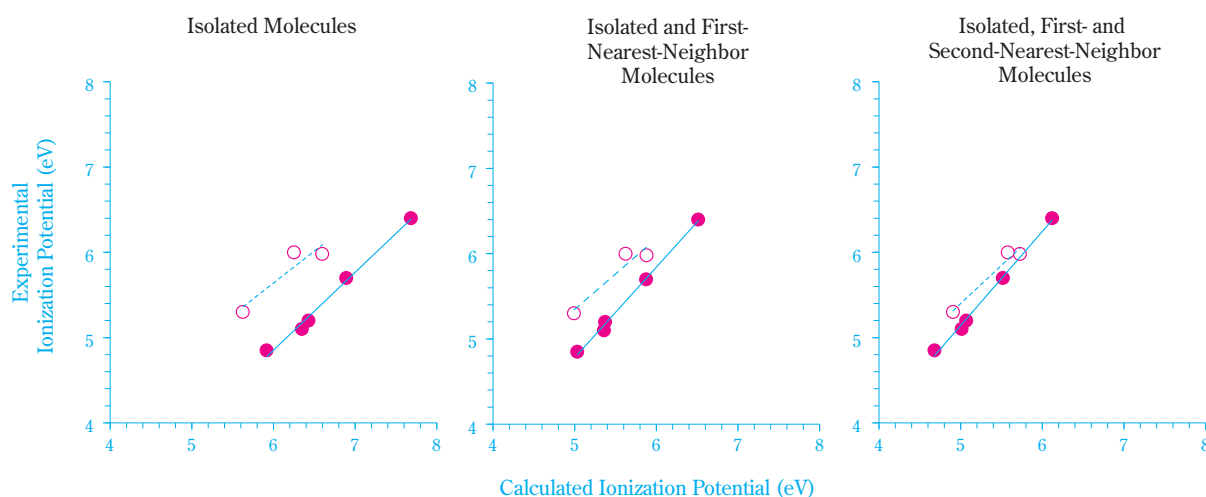
One of the factors that affect the injection of holes from an electrode to an organic semiconductor is the difference between the Fermi energy of the electrode material and the valence band top energy of the organic semiconductor. The smaller this energy difference is, the more often hole injection will occur, thus causing the electric resistance to become smaller. This is ideal from the perspective of economical power consumption.

In an experiment the valence band top energy of the organic semiconductor is evaluated as a negative value of the ionization potential, which is measured using photoelectron spectroscopy. In a theoretical calculation it is often evaluated by calculating the HOMO (Highest Occupied Molecular Orbital) levels of single molecules. However, because an organic semiconductor is an aggregate composed of numerous organic semiconductor molecules, it is affected by the intermolecular interactions such as  $\pi - \pi$  interaction, and consequently the tendency of the measured ionization potential may not be explained merely by the calculated values of the HOMO levels of single molecules. An example is shown below.<sup>18)</sup>

We examined the ionization potentials of low-molecular organic semiconductors as a foundation for predict-



**Fig. 3** Molecular structures of organic semiconductors for ionization potential calculations



**Fig. 4** Experimental and calculated ionization potentials

ing the aforementioned hole-injectability. Regarding single molecules, which constitute each of the low-molecular organic semiconductors shown in **Fig. 3** and were isolated from measured single crystal structures, the left graph in **Fig. 4** depicts the relationship between the measured ionization potential and the calculated ionization potential as a difference between the energy of the neutral state and that of the cation radical state using the density functional method (B3LYP/6-31G\* level) through the Gaussian09 program.<sup>20)</sup> The acene-type low-molecular compounds (●) and amine-type low-molecular compounds (○) are described using different correlation lines. However, when calculating the ionization potential including the first-nearest-neighbor molecules in the measured single crystal structure<sup>21)</sup>, the correlation lines of the acene-based and amine-based low-molecular compounds become closer to one another. Furthermore, if this calculation includes the second-nearest-neighbor molecules, the correlation lines of both molecules will nearly overlap. When this occurs, the correlation between the total experimental value and the total calculated value of the acene-based and amine-based molecules are 0.73, 0.89 and 0.98 for the isolated molecule model, first-nearest-neighbor model and second-nearest-neighbor model, respectively.

The above results indicate that in order to comprehensively understand the tendencies of the ionization potentials of various organic semiconductors, the calculation for single molecules is insufficient. It is therefore necessary to use a calculation model for a molecular aggregate that includes at least the first- and the second-nearest-neighbor molecules. It can be surmised that this is due to the fact that the delocalization of spin den-



**Fig. 5** Spin density calculated for isolated, first- and second-nearest-neighbor pentacenes

sity occurs in multiple molecules when ionized, as seen in the example of pentacenes in **Fig. 5**.

#### (ii) Hole Conductivity in Organic Semiconductors

Of the three devices listed in a previous section, the improvement of carrier mobility in an organic semiconductor is important in order to improve the performance of organic transistors in particular. When the voltage between the electrodes is constant, electric current can be increased by increasing the carrier mobility. The application of organic transistors to adjustment of the brightness of pixels of flexible organic EL displays has been examined. However, because a certain amount of electric current is required at low voltage, the improvement of carrier mobility is necessary from the standpoint of economical power consumption.

The theories and methods to calculate the carrier conductivity in organic semiconductors include the Marcus theory<sup>22)</sup>, band theory<sup>23)</sup>, non-equilibrium Green's function method<sup>24)</sup> and time-dependent Schrödinger equation.<sup>25)</sup> Below are shown examples of our application of the intra-chain current calculated through the non-equilibrium Green's function method as an index for the electric conductivity in polymer chains, and the resonance integral between the HOMOs of polymer chains as an index for the electric conductivity between the polymer chains.<sup>26)</sup>

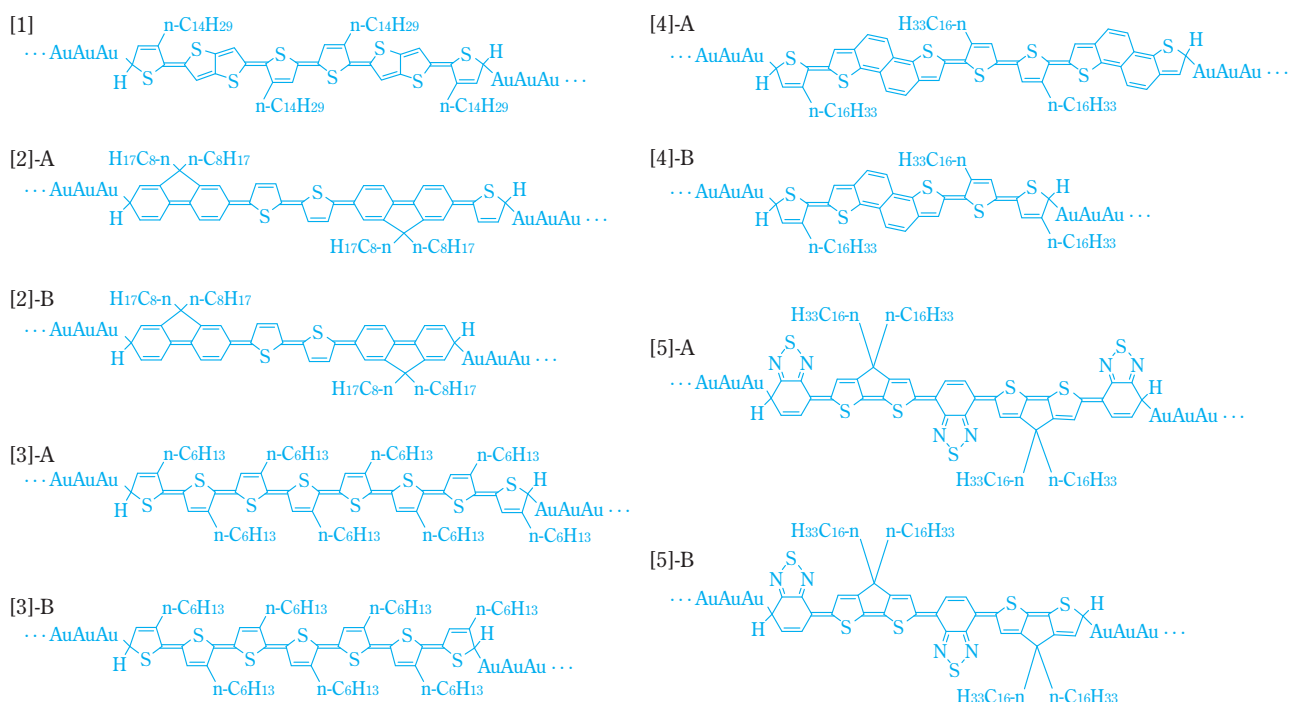
The application of the non-equilibrium Green's function method to molecules was initiated by Datta et al. in 1997.<sup>24)-a)</sup> The description of quantum mechanics using the Green's function is mathematically equivalent to the description using the wave function, i.e., the Schrödinger equation. The Green's function is suitable for describing the electric conduction phenomena because it expresses the probability that a certain phenomenon occurred at a certain time and a certain point propagates to another point after that time.

We calculated the intra-chain current by conducting the following procedures: First, using the Gaussian09 program<sup>19)</sup>, we optimized the geometries of the structure created by adding a gold atom to each of the carbon atoms at the edges of a polymer model through the density functional method (B3LYP/3-21G\*, Lan12DZ (Au) level). Next, we drew a straight line connecting the

above two atoms at both ends and other straight lines parallel to this straight line passing through the gold atoms bonded to the atoms at the edges. Then a one-dimensional gold electrode was created on the above straight line by bonding the gold atomic row that continues infinitely toward the opposite direction of the other edge. We created a one-dimensional gold electrode for each of the two atoms at the edges, and set the distance between the adjacent gold atoms at 2.884Å, which was the measured value in the gold crystal. Lastly, using the PBE density functional, DZP basis function and Troullier-Martins pseudo-potential through the Transiesta program, we calculated the electric current when a voltage of 0.3V was applied.<sup>24)-c)</sup> **Fig. 6** shows model structures for which intra-chain current values were calculated. The reason for establishing two model types (A and B) for each structure except for [1] is to calculate  $I_{A-B}$ , which is the current value in the same length chain as that of [1], using equation (1), considering the dependence of the current on the chain length.

$$I_{A-B} = \exp\{[\ln(I_A) - \ln(I_B)](L_S - L_B)/(L_A - L_B) + \ln(I_B)\} \quad (1)$$

In the above equation,  $I_A$  and  $I_B$  are the calculated current values in models A and B, respectively, when the voltage of 0.3 V is applied.  $L_A$ ,  $L_B$  and  $L_S$  express the distances between the edge atoms of models A, B and [1], respectively.



**Fig. 6** Model structures for calculations of the intra-chain current

When the carrier conduction is classified as the “hopping” type, the Marcus theory can be applied and the probability of hole hopping between the molecules is proportional to the second power of the resonance integral between HOMOs. Because the hole mobility is proportional to the hole-hopping probability, the larger the resonance integral between HOMOs is, the greater the hole mobility will be. Moreover, when the carrier conduction is classified as the “band conduction” type, the band theory can be applied, and the greater the resonance integral between HOMOs is, the smaller the effective mass of holes will be. Because the hole mobility is inversely proportional to the effective mass of holes, the larger the resonance integral between HOMOs is, the greater the hole mobility will be in this case as well.

The resonance integral between HOMOs, which is expressed as  $H_{AB}$ , was calculated using equation (2).

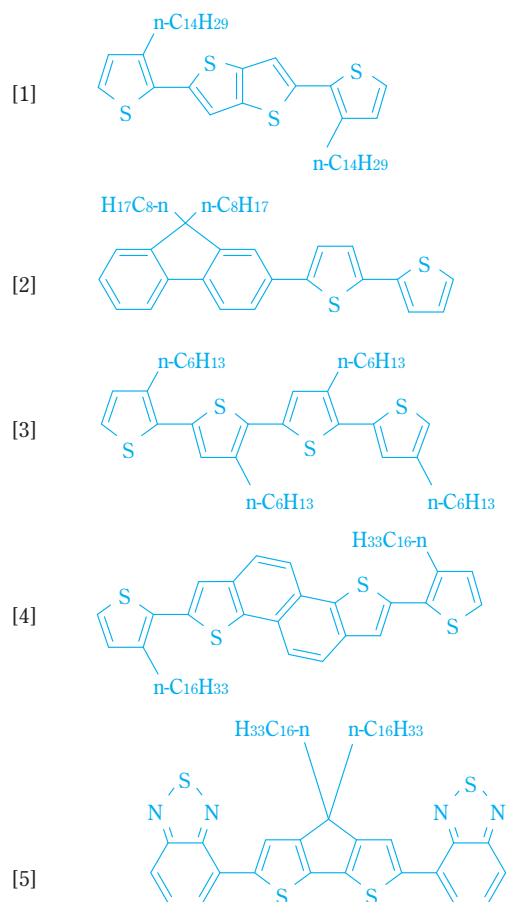
$$\begin{aligned} H_{AB} &= \langle \psi_{HOMO}^A | KS^{AB} | \psi_{HOMO}^B \rangle \\ &= \left| \sum_i \sum_j C_{i,HOMO}^A C_{j,HOMO}^B \langle \psi_i^{AB} | KS^{AB} | \psi_j^{AB} \rangle \right| \\ &= \left| \sum_i C_{i,HOMO}^A C_{i,HOMO}^B \varepsilon_i^{AB} \right| \end{aligned} \quad (2)$$

$KS^{AB}$  indicates the Kohn-Sham operator in a stacking stable structure of molecules A and B.  $\Psi_{HOMO}^A$  and  $\Psi_{HOMO}^B$  are HOMOs of molecules A and B, respectively, isolated from the stacking stable structure conserving their geometries and orientations. They are expressed by equation (3) using the molecular orbital  $\Psi_i^{AB}$  in the stacking stable structure.

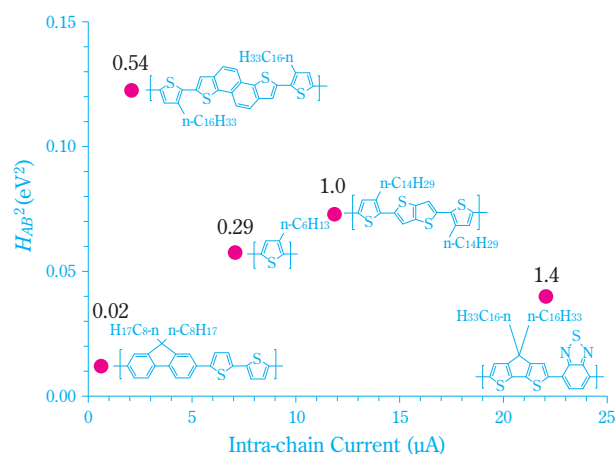
$$\begin{aligned} \psi_{HOMO}^A &= \sum_i \psi_i^{AB} C_{i,HOMO}^A \\ \psi_{HOMO}^B &= \sum_j \psi_j^{AB} C_{j,HOMO}^B \end{aligned} \quad (3)$$

$C_{i,HOMO}^A$  and  $C_{j,HOMO}^B$  indicate expansion coefficients, while  $\varepsilon_i^{AB}$  indicates the energy of the  $i$ -th molecular orbital. The Gaussian09 program<sup>19)</sup> was used for obtaining the stacking stable structures and resonance integral between HOMOs of two molecules through the density functional method (MPWB1K/6-31G\* level). Fig. 7 shows the model structure used for the stacking structure calculation.

Fig. 8 depicts the results of the two-dimensional plot of the second power of the resonance integral between the HOMOs of polymer organic semiconductor model structures and the intra-chain current calculated for the

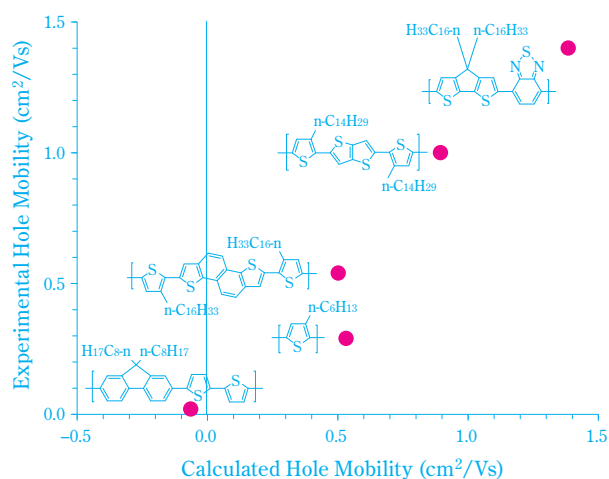


**Fig. 7** Model structures for calculations of stacking structures and  $H_{AB}$



**Fig. 8**  $H_{AB}^2$  and intra-chain current calculated for polymer models

model structures. The value shown at each plotting point is a measured value of the hole mobility.<sup>27)</sup> Viewing this plot, one can see that if either the squared value of the resonance integral between HOMOs or the intra-chain current is large, the hole mobility will tend to be greater. Fig. 9 depicts the plot of comparisons between the measured hole mobility and the hole mobility calcu-



**Fig. 9** Experimental and calculated hole mobilities

lated using the correlation equation obtained by conducting a multiple-regression analysis on the measured hole mobility using the above two parameters. The multiple regression coefficient was high at 0.97, thus proving the aforementioned tendency.

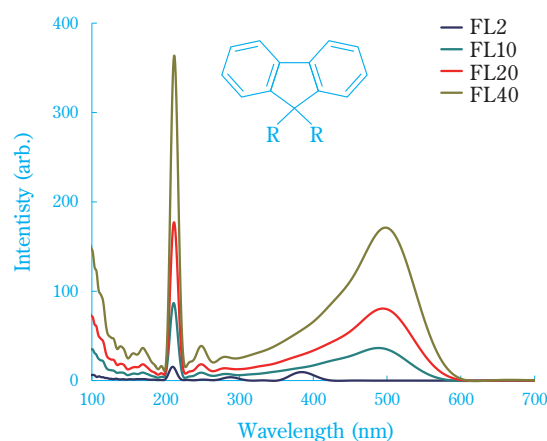
### (3) Optical Spectrum Simulation of Organic Semiconducting Polymers Using the K Computer

Since characteristics of aggregated structures of organic semiconductor molecules (such as crystallinity) greatly affect the device characteristics as described in the previous section, it is necessary to use a molecular-aggregate model in calculation as well. It can be assumed that the same is true in the calculation of organic semiconducting polymers.

However, with respect to polymer models, even only a single main chain consists of at least 10 monomer units (several hundreds of atoms). Therefore, an aggregate model containing several such main chains will have more than a few thousands or even more than ten thousand constituent atoms. In the current situation, it is not realistic to calculate the electronic structures of such large systems using our in-house computers. For such a large-scale electronic structure calculation, we have used large computers outside the company, such as the Earth Simulator and the K computer, in order to acquire the state-of-the-art technologies and establish fundamental technologies for the future. In this section we will introduce the analysis regarding the optical spectrum of organic semiconducting polymers using the K computer. First, we installed a real-space and real-time program based on time-dependent density functional theory (TDDFT)<sup>28), 29)</sup> herein called RSRT<sup>30), 31)</sup>, which would

be used for spectrum calculations, into the K computer. Subsequent to the installation, we performed a tuning of the RSRT in collaboration with the industrial-use support staff, while the RSRT demonstrated excellent performance on the Earth Simulator<sup>32)</sup>, by which a speed-up in program execution by a factor of 1.4 was accomplished without making any significant change in the program.<sup>33)</sup>

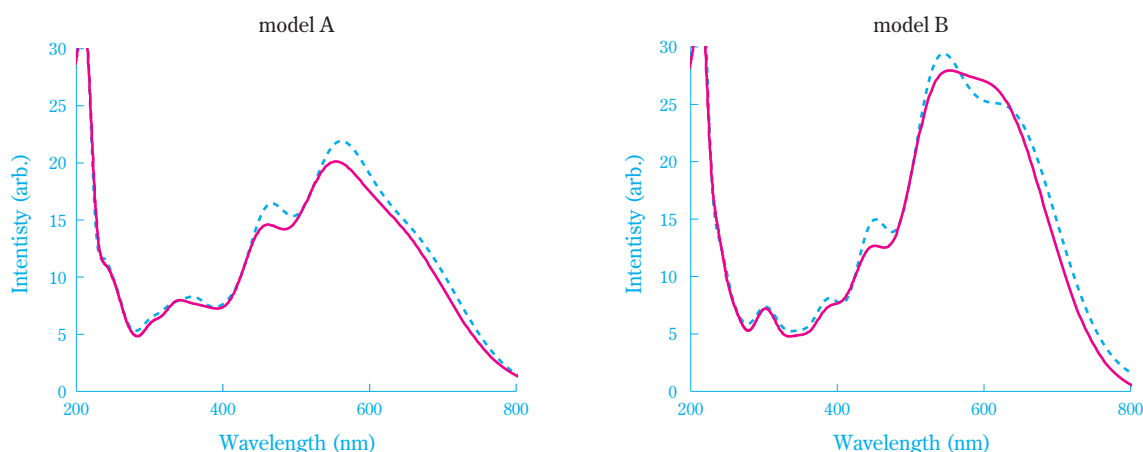
Next, we examined the effect of the polymer chain length on the absorption spectrum using a simplified single straight chain model having a methyl group at the ninth position of fluorine (FL), which is one of the typical organic EL materials, using the RSRT program.<sup>34)</sup> Fig. 10 shows the results of the spectrum calculation.



**Fig. 10** Calculated optical spectra of fluorene models using TDDFT/RSRT program. FL2, FL10, FL20 and FL40 stand for dimer, decamer, icosamer and tetracontamer of the models, respectively. Inset picture: chemical structure of fluorene monomer.

The use of the K computer enabled us to successfully perform the first-principles calculation of a polymer chain model having a size of 40 monomer units (approximately 30 nm). As the molecular chain length was extended from 2 monomer units to 10, 20 and 40 monomer units, the longest-wavelength absorption peak indicated a red shift. However, if the length was 10 monomer units or longer, the peak positions remained nearly the same. In the case of an ideal straight-chain structure, we suppose that the electronic state of a conjugated polymer can be expressed using a model having a molecular length of approximately 10 monomer units. Based on studies of solitons and polarons in PPV, we believe this assumption qualitatively agrees with the





**Fig. 11** Calculated optical spectra using TDDFT/RSRT for disordered aggregation model A (left) and oriented aggregation model B (right). Solid line: spectrum of aggregation model including three molecular chains; Dashed line: sum of three spectrum from individual molecular chain.

result that unpaired  $\pi$  electrons and the charge expand to approximately four phenyl rings of the PPV.<sup>35)</sup>

We are also attempting to associate the electronic structure of the aggregated polymers with that of each component polymer chain in the aggregates. In this study, using an aggregate model in which there were several 10 monomer-unit single chains containing an octyl group at the ninth position of the FL, we calculated the absorption spectrum of the model structure A having disordered molecular chains and that of the model structure B having relatively oriented molecular chains through the RSRT. We then compared those absorption spectra to the sum of absorption spectra of the components, which are separately calculated using the RSRT to individual polymer chain models.

The aggregate-structure model was used for an absorption spectrum calculation through the RSRT after conducting the following procedures: The molecular dynamics calculation was performed under the periodic boundary condition (PBC) after arranging three molecular-chain models having FL10 monomer units within a cell; and an aggregated stable structure of the three chains was extracted from the cell under PBC. **Fig. 11** shows the results.

In the model structure A having disordered molecular chains, the shape of the absorption spectrum of the molecular aggregate turned out to be similar to that of the sum of absorption spectra of the components, comprising three individual polymer chains.

Contrastingly, in model structure B, in which molecular chains were relatively oriented, those absorption spectra had different peak shapes on the longest wave-

length side. Presuming a difference in strength of molecular interaction in those disordered and molecular oriented models, it suggests that while the characteristics of an isolated molecular chain are reflected in the spectrum in the aggregate structure of disordered model A, the effect of intermolecular interaction is reflected in molecular oriented model B in addition to the characteristics of an isolated molecular chain.

Although the improvement of calculation accuracy and more detailed analyses of the structural characteristics and electronic structure in each model are future challenges, we believe that understanding the relationship between a material structure and the electronic state using such a large-scale calculation is crucial to the future design and development of organic electronic materials, because experimental research on single molecule spectroscopy has progressed in recent years<sup>36)</sup>. It has been reported that the conformation such as bending structures and the contact between segments of isolated molecular chains are related to energy transfer and the emission mechanism.<sup>37)</sup>

#### (4) Prediction of Solubility

In the development of organic device materials which are characterized by a solution coating process, the control of solubility of chemical compounds is an important technique. Therefore, we pay attention to the solubility parameter (hereinafter referred to as the “SP”) as an index for the prediction of solubility. The solubility parameter is a value defined by the regular solution (i.e., non-electrolytic solution having no chemical interaction or association) theory introduced by Hildebrand et al.<sup>38)</sup> In a regular solution, the mixing energy of two compo-

nents A and B (expressed as  $\Delta E^{mix}$ ) is expressed by molar evaporation energy (cohesive energy)  $\Delta E_A^V$  and  $\Delta E_B^V$ , molar volume  $V_A$  and  $V_B$ , and molar numbers  $n_A$  and  $n_B$ , as shown in equation (4).

$$\Delta E^{mix} = \frac{n_A V_A \cdot n_B V_B}{n_A V_A + n_B V_B} \left[ \left( \frac{\Delta E_A^V}{V_A} \right)^{1/2} - \left( \frac{\Delta E_B^V}{V_B} \right)^{1/2} \right]^2 \quad (4)$$

The SP is an amount expressed as the square root of cohesive energy density. It can be considered that the closer the SP's are, the more completely the two components will mix (according to the equation (4)). Although the SP is not always theoretically accurate for the actual material, it is a convenient parameter that can be used for a simple solubility prediction. Because the SP is determined on the basis of substances, a combination can be selected merely through intercomparison of the list of solute and solvent values without calculating it for each individual solute-solvent combination. The group contribution method is generally used for estimation of the SP.<sup>39)–41)</sup> In this method the value of the atomic group is determined on the basis of the measured value and the SP is obtained through summation. If there are many similar chemical compounds that have measured values, such as general-purpose polymers, the SP can be estimated using the group contribution method. However, chemical compounds within a conjugate framework for use in organic devices have only minimal numbers of measured values. Consequently, in some cases the group contribution method cannot be applied to predict the solubility of a new chemical compound.

We therefore examined the SP estimation method using an approach in which the cohesive energy density is calculated directly from the nonbonded intermolecular interaction in a molecular aggregate through the molecular dynamics method (MD). In an estimation using MD, it is possible to reflect all the effects that are hard to be taken into account when using the group contribution method, such as shape-related effects, including the position of a substituent, the packing effect between the molecules and temperature dependence.

The SP estimation procedure using the MD calculation method is as follows: (1) create initial geometries of randomly arranged molecular aggregates of a single chemical compound for which SP must be predicted in a unit cell; (2) perform MD calculation until the energy and volume of the simulation cell reach the equilibrium state assuming the NPT ensemble; (3) fix the volume and shape of the simulation cell, and perform MD cal-

ulation through the NVT ensemble for sampling; (4) obtain the  $\delta$  value of SP through statistical processing using equation (5), in which  $V$  indicates the volume of the simulation cell of a molecular aggregate,  $E_{coh}$  indicates the cohesive energy,  $E_{aggregate}$  is the energy of the aggregate and  $E_{isolated\ molecule}$  is the energy of each molecule in the aggregate.

$$\delta = \sqrt{\frac{E_{coh}}{V}}, E_{coh} = -(E_{aggregate} - \sum E_{isolated\ molecule}) \quad (5)$$

Although there are numerous types of intermolecular interactions in actual materials, they can be roughly classified into two types: vdW interaction (dispersion force) and Coulomb's interaction (electrostatic force). Because it can be assumed that chemical compounds having closer cohesive energy density will be more miscible with regard to both interactions, we decided to use the two-dimensional indication (hereinafter referred to as the "2D-SP") by dividing the SP into the vdW and Coulomb's interaction as shown in equation (6). In that equation ( $E_{coh}$ )<sub>vdW</sub> indicates the vdW interaction of the cohesive energy of the molecular aggregate and ( $E_{coh}$ )<sub>Coulomb</sub> indicates the Coulomb's interaction.

$$\delta_{vdW} = \left[ \frac{(E_{coh})_{vdW}}{V} \right]^{1/2}, \delta_{Coulomb} = \left[ \frac{(E_{coh})_{Coulomb}}{V} \right]^{1/2} \quad (6)$$

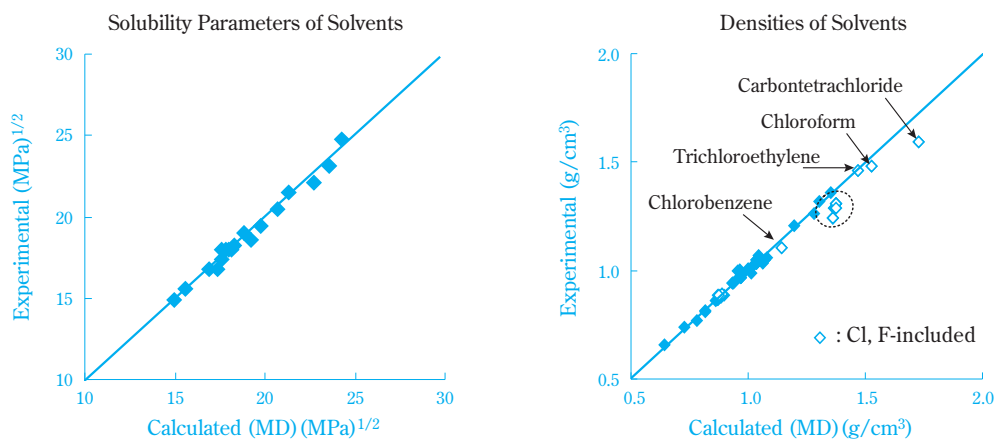
Here, by indicating  $\delta_{vdW}$  and  $\delta_{Coulomb}$  using the two-dimensional graph, the degrees of solubility and affinity of the materials can be expressed by the distance on the graph. As other similar methods to express the solubility by taking types of intermolecular interactions into account, the Hansen<sup>42)</sup> and Hoy<sup>43)</sup> methods are known. In those methods the SP is indicated using the three interactions; dispersion forces, dipole forces and hydrogen bond forces. However, we believe the 2D-SP can serve as an effective index for the solubility due to the following reasons: There are actual cases in which the solubility was adequately predicted through the aforementioned method (which divides the SP into two interactions) for materials having a large hydrogen bonding strength; and it is simply easier to plot two types of interactions in a graph than it is to plot three types of interactions.

The SP estimated through such methods have been used to find the optimum ratio of two chemical compounds in a mixing system and then select solvents taking into account the wettability, as well as other applications, including prediction of the solubility of new chemical compounds and selection of soluble solvents.

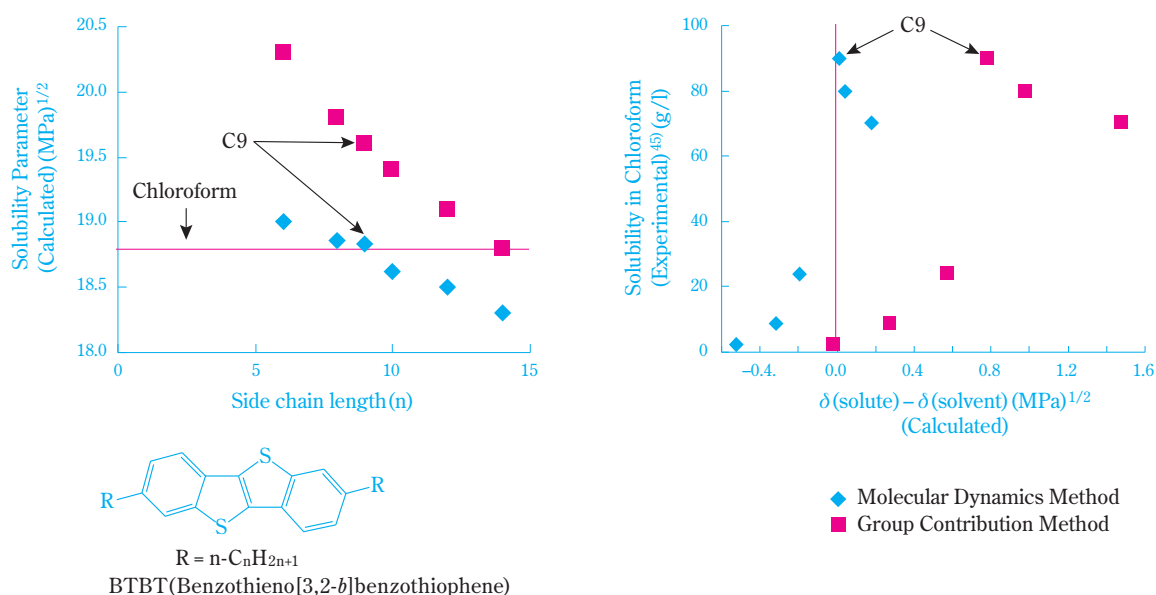
Fig. 12 shows examples of calculation of the SP and densities of solvents. With regard to general-purpose low-molecular compounds such as solvents, measured values can be accurately predicted by the MD calculation. However, the calculated densities of solvents containing Cl and F atoms will be slightly greater than the measured values.

For materials which have a high orientation, such as organic transistor materials, we can carry out a calculation that takes into account the molecular shape and packing property by setting a nematic initial structure in which the molecular axes are aligned toward a single direction within the reasonable range of angle, assuming the NPT ensemble and allowing the cell shape to be vari-

able. As an example, the application of the SP estimation method to BTBT, which is a low-molecular organic transistor material, is shown here. The tendency—which is that the longer the alkyl side-chain length is, the smaller the SP will be—is the same in both the MD and group contribution methods (Fig. 13, left). However, while by the group contribution method it is predicted that the longer the side-chain length is, the smaller the difference in the SP of solute and solvent will be, thereby causing the chemical compound to be more soluble, by the MD method it is predicted that when the number of carbon atoms of the alkyl side chain is nine (C9), the chemical compound will become most soluble in chloroform, thereby the tendency is consistent with experi-



**Fig. 12** Solubility parameters (left) and densities (right) of solvents calculated by MD (Experimental data: Polymer Handbook 4th<sup>44</sup>)

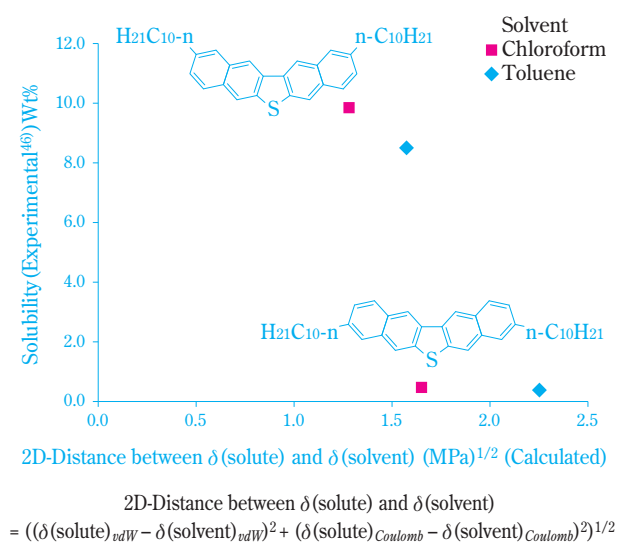


**Fig. 13** Solubility parameters vs. side-chain lengths of BTBT (left) Experimental solubilities in chloroform vs. Calculated solubility parameters of solute ( $\delta(\text{solute})$ ) and solvent ( $\delta(\text{solvent})$ ) (right)

ments<sup>45)</sup> (Fig. 13, right). This is because the characteristics of the chemical compound as an aggregate, such as orientation and packing property, have been taken into account during the MD calculation.

While it is impossible to predict the effect on the substituent position to the solubility from the group contribution method, it is possible from the MD method.

The solubility of low-molecular organic transistor materials greatly varies even only by changing the side-chain positions. The SP of two types of DNT having different side-chain positions obtained through the MD cal-



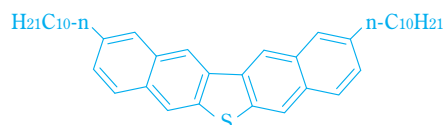
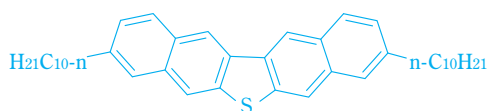
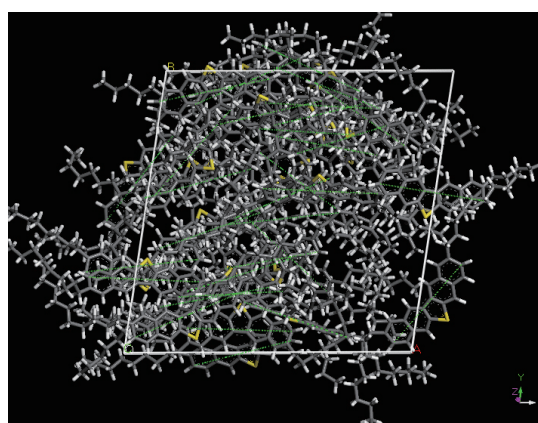
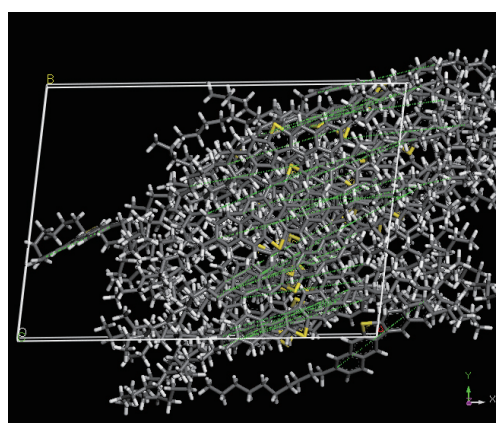
**Fig. 14** Experimental solubilities vs. Calculated solubility parameters for 2 structures of DNT (dinaphtho [2,3-*b*:2',3'-*d*] thiophene)

culcation are significantly different, and the results of the prediction of solubility with respect to solvents were consistent with the tendencies of measured values<sup>46)</sup> (Fig. 14). Viewing the snapshot of MD calculation of both DNTs, it can be observed that the isomer having the greater solubility tends not to pack, which suggests that the packing property has a correlation with solubility (Fig. 15).

Additionally, we have examined methods for selecting the optimum combination that would enhance the transistor characteristics of polymer and low-molecular compound mixed type organic semiconductor materials by using the SP of the ingredients. We have discovered that the difference in the SP obtained under certain conditions between polymers and low-molecular compounds has a correlation with the transistor characteristics.<sup>47)</sup>

Regarding the SP estimation using the MD calculation as a solubility prediction method, there are the following challenges:

- (i) With regard to polymers, it is difficult to estimate the absolute value because the calculated value varies according to the chain length of the model. Although intercomparison between the same types of polymers is possible, it is hard to compare the SP of polymers with those of low-molecular materials.
- (ii) The results of calculation depend on the initial state.
- (iii) It is time-consuming.



**Fig. 15** MD Snap-Shot of DNTs

For the SP estimation, we will improve the calculation speed and accuracy and examine the MD calculation with more consideration for structural parameters such as an ordered structure. Moreover, we will investigate the solubility predictive equation including solubility-related parameters other than the SP.

## Conclusion

U.S. President Barack Obama, in 2011, issued the Materials Genome Initiative under the slogan of “2X faster & 2X cheaper” in order to enhance the competing power of the U.S. industry.<sup>48)</sup> In regard to the initiative it was pointed out that it usually takes approximately 10 to 20 years from developing a material to launching it into the market as a final product. In order to shorten this process as much as possible, the materials informatics approach is attracting much attention. The materials informatics procedures are as follows: Characteristics that are unique to the material are calculated using a computer simulation beforehand; the results of this calculation are accumulated as a data set together with the corresponding experimental data; and the structure that realizes the target characteristics and functions is derived from the above data set.<sup>49)</sup> While in the deductive approach one attempts to design a material having the target functions by understanding and predicting the material properties through experiments and theoretical calculations using model structures, in materials informatics an inductive approach is taken as follows: An enormous data set of material characteristics is analyzed through the informatics method, and a material having the target functions is discovered based on the knowledge obtained from that analysis.

Although in the beginning of this paper we stated that computational science is considered as a “third science,” such data-oriented science is currently advocated as a “fourth science,” which covers a broader area. In the field of materials science, this data-oriented science has become mainstream in accelerating the movement of materials informatics in terms of how (and how quickly) we can discover the optimal combination from numerous material-composition combinations.

In the previous chapter we introduced four cases with a focus on our practical approach toward computational materials science in electronics materials development at Sumitomo Chemical. Our efforts remain within the framework of the deductive approach. We believe that in order to achieve an inductive approach such as mate-

rials informatics, we must consciously practice the manner of thinking based on data-oriented science. Not only must we continue to improve our elemental technologies that we have accumulated thus far, but we must also accelerate our efforts in new frontiers such as materials informatics.

## Acknowledgment

The organic semiconductor analysis cases described in the first half of Section 2 of Chapter 2 of this paper were obtained from the 2013 industrial use of TSUB-AME of the Tokyo Institute of Technology, subject number 13IBG. Also, the application examples of the K computer described in Section 3 of Chapter 2 of this paper were obtained through the use of the K computer of the National Research and Development Agency RIKEN (subject numbers: hp120028 and hp140089).

## References

- 1) “Keisan no Kagaku in Iwanami Lecture Series Keisan Kagaku”, Vol.1, A. Ukawa, A. Oshiyama, Y. Oyanagi, M. Sugihara, A. Sumi and H. Nakamura eds., Iwanami Shoten, p.3 (2013).
- 2) a) T. Nishikawa, *Joho Kanri*, **55**(12), 882 (2013).  
b) Research Organization for Information Science and Technology, General Incorporated Foundation, “Guide to HPCI for Beginner Users”, [http://www.hpci-office.jp/pages/e\\_guide](http://www.hpci-office.jp/pages/e_guide) (refer at April. 16, 2015).
- 3) C. Hansch and T. Fujita, *J. Am. Chem. Soc.*, **86**, 1616 (1964).
- 4) O. Kirino, *J. Pestic. Sci.*, **9**, 571 (1984).
- 5) C. Takayama, Y. Miyashita, S. Sasaki and M. Yoshida, *Quant. Struct.-Act. Relat.*, **2**, 121 (1983).
- 6) M. Yoshida, *J. Syn. Org. Chem., Jpn.*, **42**, 743 (1984).
- 7) Y. Kurita, K. Tsushima and C. Takayama, “Probing Bioactive Mechanisms (ACS Symposium Series No.413)”, American Chemical Society (1989), p.183.
- 8) T. Katagi and Y. Kurita, *J. Pestic. Sci.*, **14**, 93 (1989).
- 9) K. Watanabe, K. Umeda, Y. Kurita, C. Takayama and M. Miyakado, *Pestic. Biochem. Physiol.*, **37**, 275 (1990).
- 10) Y. Kurita, C. Takayama, T. Katagi and M. Yoshida, “Computer Aided Innovation of New Materials”, Elsevier Science (1991), p.455.
- 11) C. Takayama, N. Meki, Y. Kurita and H. Takano, “Classical and Three-Dimensional QSAR in Agro-

- chemistry (ACS Symposium Series No. 606)", American Chemical Society (1995), p.154.
- 12) F. Goto, M. Ishitobi, T. Endo, M. Ishida, A. Nakazono, Y. Zenpo and M. Yosida, *SUMITOMO KAGAKU*, **1994-II**, 50 (1994).
  - 13) T. Kunimoto, R. Yoshimatsu, K. Ohmi, S. Tanaka and H. Kobayashi, *IEICE Trans. Electron.*, **E85-C**, 1888 (2002).
  - 14) G. Bizarri and B. Moine, *J. Lumin.*, **115**, 53 (2005).
  - 15) M. Städele, M. Moukara, J. A. Majewski, P. Vogl and A. Görling, *Phys. Rev. B*, **59**, 10031 (1999).
  - 16) a) M. Ishida, Y. Imanari, T. Isobe, S. Kuze, T. Ezuhara, T. Umeda, K. Ohno and S. Miyazaki, *J. Phys.: Condens. Matter*, **22**, 384202 (2010).  
b) M. Ishida, Y. Imanari, T. Isobe, S. Kuze, T. Ezuhara, T. Umeda, K. Ohno and S. Miyazaki, *J. Phys.: Conf. Ser.*, **454**, 012062 (2013).
  - 17) G. Kresse and J. Hafner, *Phys. Rev. B*, **47**, 558 (1999).
  - 18) Y. Kurita, "Proceedings of Industrial Use Symposium of TSUBAME, Tokyo Institute of Technology", **2014**, 203.
  - 19) Gaussian 09, Revision C.01, Gaussian, Inc., Wallingford CT, 2009.
  - 20) a) P. K. Nayak and N. Periasamy, *Org. Elec.*, **10**, 532 (2009).  
b) P. K. Nayak and N. Periasamy, *Org. Elec.*, **10**, 1396 (2009).
  - 21) a) S. C. Capelli, A. Albinati, S. A. Mason and B. T. M. Willis, *J. Phys. Chem. A*, **110**, 11695 (2006).  
b) C. P. Brock and J. D. Dunitz, *Acta Cryst.*, **B46**, 795 (1990).  
c) D. Holmes, S. Kumaraswamy, A. J. Matzger and K. P. C. Vollhardt, *Chem. Eur. J.*, **5**, 3399 (1999).  
d) S. Haas, B. Batlogg, C. Besnard, M. Schiltz, C. Kloc and T. Siegrist, *Phys. Rev. B*, **76**, 205203 (2007).  
e) C. Näther, H. Bock, Z. Havlas and T. Hauck, *Organometallics*, **17**, 4707 (1998).  
f) Y.-H. Sun, X.-H. Zhu, Z. Chen, Y. Zhang and Y. Cao, *J. Org. Chem.*, **71**, 6281 (2006).  
g) P. J. Low, M. A. J. Paterson, D. S. Yufit, J. A. K. Howard, J. C. Cherryman, D. R. Tackley, R. Brook and B. Brown, *J. Mater. Chem.*, **15**, 2304 (2005).  
h) J.-A. Cheng and P.-J. Cheng, *J. Chem. Cryst.*, **40**, 557 (2010).
  - 22) a) R. A. Marcus, *J. Chem. Phys.*, **24**, 966 (1956).  
b) R. A. Marcus and N. Sutin, *Biochimica et Biophysica Acta*, **811**, 265 (1985).  
c) W.-Q. Deng and W. A. Goddard III, *J. Phys. Chem. B*, **108**, 8614 (2004).  
d) S.-H. Wen, A. Li, J. Song, W.-Q. Deng, K.-L. Han and W. A. Goddard III, *J. Phys. Chem. B*, **113**, 8813 (2009).  
e) I. Yavuz, B. N. Martin, J. Park and K. N. Houk, *J. Am. Chem. Soc.*, **137**, 2856 (2015).
  - 23) a) J. E. Northrup, *Appl. Phys. Lett.*, **99**, 062111 (2011).  
b) H. Kobayashi, N. Kobayashi, S. Hosoi, N. Koshitani, D. Murakami, R. Shirasawa, Y. Kudo, D. Hobara, Y. Tokita and M. Itabashi, *J. Chem. Phys.*, **139**, 014707 (2013).
  - 24) a) S. Datta, W. Tian, S. Hong, R. Reifengerger, J. I. Henderson and C. P. Kubiak, *Phys. Rev. Lett.*, **79**, 2530 (1997).  
b) W. Tian, S. Datta, S. Hong, R. Reifengerger, J. I. Henderson and C. P. Kubiak, *J. Chem. Phys.*, **109**, 2874 (1998).  
c) M. Brandbyge, J.-L. Mozos, P. Ordejón, J. Taylor and K. Stokbro, *Phys. Rev. B*, **65**, 165401 (2002).  
d) Z. Li, M. Smeu, T.-H. Park, J. Rawson, Y. Xing, M. J. Therien, M. A. Ratner and E. Borguet, *Nano Lett.*, **14**, 5493 (2014).  
e) K. H. Khoo, Y. Chen, S. Li and S. Y. Quek, *Phys. Chem. Chem. Phys.*, **17**, 77 (2015).
  - 25) P. Prins, F. C. Grozema and L. D. A. Siebbeles, *J. Phys. Chem. B*, **110**, 14659 (2006).
  - 26) Sumitomo Chemical Co., Ltd., JP 2012-241038 A.
  - 27) a) T. Fujiwara, J. Locklin and Z. Bao, *Appl. Phys. Lett.*, **90**, 232108 (2007).  
b) D. A. Lyashenko, A. A. Zakhidov, V. A. Pozdin and G. G. Malliaras, *Org. Elec.*, **11**, 1507 (2010).  
c) I. Osaka, T. Abe, S. Shinamura and K. Takimiya, *J. Am. Chem. Soc.*, **133**, 6852 (2011).  
d) B. H. Hamadani, D. J. Gundlach, I. McCulloch and M. Heeney, *Appl. Phys. Lett.*, **91**, 243512 (2007).  
e) H. N. Tsao, D. Cho, J. W. Andreasen, A. Rouhanipour, D. W. Breiby, W. Pisula and K. Müllen, *Adv. Mater.*, **21**, 209 (2009).
  - 28) E. Runge and E. K. U. Gross, *Phys. Rev. Lett.* **52**, 997 (1984).
  - 29) Y. Zempo, N. Akino, M. Ishida, M. Ishitobi and Y. Kurita, *J. Phys.: Condens. Matter*, **20**, 064231 (2008).
  - 30) K. Yabana and G. F. Bertsch, *Phys. Rev. B*, **54**, 4484 (1996).
  - 31) J. Chelikowsky, N. Troullier, K. Wu and Y. Saad, *Phys. Rev. B*, **50**, 11355 (1994).

- 32) M. Ishida, N. Akino, Y. Zempo, S. Urashita, S. Shingu, H. Suno and N. Nishikawa, "The User Report of the Earth Simulator Industrial Application Project", p.29 (2010).
- 33) Y. Zempo, N. Akino, M. Ishida, E. Tomiyama and H. Yamamoto, *J. Phys. Conf. Series*, in submitted.
- 34) M. Ishida, "HPCI User Report for Industrial Use of K computer, 2013", <https://www2.hpci-office.jp/output/hp120028/outcome.pdf> (refer at April 16, 2015).
- 35) S. Kuroda, *OYOBUTSURI*, **76**, 795 (2007).
- 36) H. Masuhara, *Polymers*, **60**, 53 (2011).
- 37) S. Habuchi and M. Vacha, *Polymers*, **60**, 54 (2011).
- 38) J. H. Hildebrand and R. L. Scott, "The Solubility of Non-Electrolytes", Reinhold Publishing Corp. (1949).
- 39) P. A. Small, *J. Appl. Chem*, **3**, 77 (1953).
- 40) R. F. Fedors, *Polym. Eng. Sci.*, **14**, 147&472 (1974).
- 41) D.W. Van Krevelen and P. J. Hoftyzer, "Properties of Polymers, Their Estimation and Correlation with Chemical Structure", Elsevier Publishing Company, Amsterdam (1972).
- 42) C. M. Hansen, "Hansen Solubility Parameters", CRS Press (2000).
- 43) K. L. Hoy, *J. Paint. Technol.*, **42**, 76 (1970).
- 44) J. Brandrup, E. H. Immergut and E. A. Grulke, "Polymer Handbook 4<sup>th</sup>", Wiley (2003).
- 45) H. Ebata, T. Izawa, E. Miyazaki, K. Takimiya, M. Ikeda, H. Kuwabara and T. Yui, "*J. Am. Chem. Soc.*", **129**, 15732 (2007).
- 46) T. Okamoto, C. Mitsui, M. Yamaguchi, K. Nakahara, J. Soeda, Y. Hirose, K. Miwa, H. Sato, A. Yamano, T. Matsushita, T. Uemura and J. Takeya, *Adv. Mater.*, **25**, 6392 (2013).
- 47) Sumitomo Chemical Co., Ltd., JP 2009-267372A (2009), JP 5480510 B2 (2014).
- 48) the WHITE HOUSE PRESIDENT BARACK OBAMA, "Materials Genome Initiative for Global Competitiveness, June 2011", [https://www.whitehouse.gov/sites/default/files/microsites/ostp/materials\\_genome\\_initiative-final.pdf](https://www.whitehouse.gov/sites/default/files/microsites/ostp/materials_genome_initiative-final.pdf) (Ref. 2015/4/16).
- 49) Center for Research and Development Strategy, Japan Science and Technology Agency, "STRATEGIC PROPOSAL; Materials Informatics Materials Design by Digital Data Driven Method." (2013), <http://www.jst.go.jp/crds/pdf/2013/SP/CRDS-FY2013-SP-01.pdf> (refer at April 16, 2015).

## PROFILE

*Masaya ISHIDA*

Sumitomo Chemical Co., Ltd.  
Advanced Materials Research Laboratory  
Senior Research Associate

*Akiko NAKAZONO*

Sumitomo Chemical Co., Ltd.  
Advanced Materials Research Laboratory  
Senior Research Specialist

*Yasuyuki KURITA*

Sumitomo Chemical Co., Ltd.  
Advanced Materials Research Laboratory  
Senior Research Specialist,  
Doctor of Agriculture