ORIGINAL PAPER

Optimizing a coarse-grained space for approximate normal-mode vibrations of molecular heterodimers

Makoto Isogai[1] · **Masataka Seshimo**[1] · **Hirohiko Houjou**[1]

Abstract

We applied the method of coarse-graining the intermolecular vibrations to molecular heterodimers assembled by double hydrogen bonding. This method is based on principal component analysis, by which the original atomic displacement vectors are projected onto a lower-dimensional space spanned by a basis set of translations, librations, and intramolecular vibrations of the constituent molecules. Compared with homodimers, the following points are particularly noted: (1) alignment of the constituent molecules in a non-symmetric atomic arrangement of the whole system and (2) the scheme of reordering the bases to construct an optimal coarse-grained space. We tested three schemes for reordering the intramolecular vibration vectors to determine that the best one is equivalent to size reduction based on the singular value decomposition. The coarse-graining analysis affords three parameters, Φ_{intra}, Φ_{inter}, and Φ_{app}, which are relevant to the mechanical nature of the molecular assembly. The Φ_{intra} values account for the internal stiffness of molecules, while the Φ_{inter} values are true stiffness constants of the intermolecular force and show a good correlation with the association energies of the dimers. The Φ_{app} values are the apparent intermolecular stiffness smaller than Φ_{inter}, as a result of compensation for neglecting intramolecular vibrations. All these values are consistent with each other under the coupled oscillator model, showing that the present coarse-graining analysis is valid for heterodimers as well as homodimers.

Keywords Molecular vibration · Molecular assembly · Supramolecular synthon · Low-frequency mode

Introduction

Material properties are governed by both molecular structure and the packing of molecular species in the solid state; hence, the design, synthesis, and modification of novel materials with unique physicochemical properties are a significant pursuit for crystal engineering in many fields [1–3]. Indeed, crystal polymorphs, which possess identical chemical compositions but exist in different molecular arrangements and/or conformations, display distinct physical and chemical properties, including energetic properties and bioavailability [4–7]. Therefore, polymorphism has been an important consideration in many fields, including pharmaceuticals [8], food science [9], dyes/pigments [10], and organic electronics [11]. Although it is not always possible to observe a sufficient

number of polymorphs and pseudopolymorphs [12–15], researchers have employed sophisticated techniques to this problem, including cocrystallization [16–20] and salification [21–23]. Furthermore, in recent years, the elastic or plastic behavior of molecular crystals, some of which may involve local changes in molecular arrangement, has attracted increasing attention [24]. For these reasons, analyzing supramolecular synthons or heterojunctions in molecular assemblies, in terms of elastic body mechanics, is of considerable importance [25]. The nature of intermolecular forces is typically reflected in the vibrational modes of molecular assemblies, especially in the low-frequency (LF) region. Recently, several spectroscopic methods, including terahertz absorption [26–30], Raman scattering [31–34], and resonant two-photon ionization [35–37], have been used to directly observe LF-mode vibrations. These methods cover the detection of hydrogen bonding, van der Waals interactions, overall molecular distortion, and libration in systems such as hydrated sugars, nucleobase pairs, and crystal polymorphs of some drugs.

Currently, the number of reports on normal-mode analysis of large molecular assemblies such as crystals is increasing

✉ Hirohiko Houjou
   houjou@iis.u-tokyo.ac.jp

1 Institute of Industrial Science, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, Japan

[38–41]. Several high-level quantum chemical calculations with explicit full-atom representation have proven to be quite successful [42–46]. However, strict atomistic models often face problems related to the computational cost of the associated calculations. Unlike the intramolecular vibrations observed by infrared spectroscopy, no appropriate simple models are available for intermolecular vibrations. What makes it difficult to reduce the computational cost is the non-negligible coupling among the intermolecular and intramolecular vibrations in LF-mode vibrations. In an attempt to reduce the computational cost while maintaining high accuracy, we have been developing a method of coarse-graining atomic displacement vectors for normal vibrational modes [47–50]. From the coarse-grained displacement vector and frequency for each vibrational mode, we can evaluate the stiffness constant of intermolecular hydrogen bonds. To date, we have demonstrated that this size-reduction method is appropriate for reproducing homodimers that are assembled through double hydrogen bonding. For an assembly composed of rigid molecules, in which only small coupling among the intra- and intermolecular vibrations is expected, the displacement vectors can be sufficiently represented in a coarse-grained space spanned by twelve bases corresponding to translational and rotational motions of the constituent molecules [47]. For an assembly with an LF-mode intramolecular vibration, such as methyl libration, the inclusion of certain vibration modes in the coarse-grained space drastically improves the representation of the original vibration motions [48].

For molecular homodimers that consist of two identical molecules, maintaining an appropriate symmetry (centric, two-fold, mirror), the construction of the coarse-grained space is rather straightforward: one can select the intramolecular vibrations (alternately from two units in the assembly) in increasing order of frequency until the dimension reaches the number of the LF-mode vibrations of the original system. However, the situation is not that simple for heterodimers that consist of two different molecules or two identical molecules in a non-symmetric orientation. It is necessary to overcome this inevitable problem to apply our idea of coarse-graining to molecular crystals and other assemblies. In this paper, we describe our attempt to analyze the vibrations of 21 homo- and heterodimers that are composed of six different molecules (Fig. 1). The results are examined using a fidelity index that we have proposed to evaluate the quality of the coarse-grained

space [50]. In addition, based on the coupled oscillator model, we demonstrate that the intermolecular vibrations can be understood in terms of rigid-body mechanics, but far beyond the conventional pseudodiatomic model.

Theory

For the Hessian analysis of molecular vibration, the displacement vectors $\mathbf{C}$ are obtained by diagonalizing the mass-weighted Hessian matrix ($\mathbf{M}^{-1/2}\mathbf{K}\mathbf{M}^{-1/2}$) with the eigenvalue matrix $\mathbf{\Omega}$ that contains the corresponding frequency ω.

$$\left(\mathbf{M}^{-1/2}\mathbf{K}\mathbf{M}^{-1/2}\right)\left(\mathbf{M}^{1/2}\mathbf{C}\right) = \left(\mathbf{M}^{1/2}\mathbf{C}\right)\mathbf{\Omega}^2 \tag{1}$$

The mass-weighted displacement (MWD) vectors $\mathbf{W}$ are prepared by multiplying $\mathbf{M}^{1/2}$ with $\mathbf{C}$ and normalized with matrix $\mathbf{L}^{-1}$, where $\mathbf{L}^2$ is the modal mass matrix [51]. The hat (^) on $\mathbf{C}$, $\mathbf{W}$, etc., denotes that these matrixes are normalized at least with respect to each column.

$$\begin{cases} \widehat{\mathbf{W}} = \mathbf{M}^{1/2}\widehat{\mathbf{C}}\mathbf{L}^{-1} \\ \mathbf{L}^2 = \widehat{\mathbf{C}}^{\mathrm{T}}\mathbf{M}\widehat{\mathbf{C}} \end{cases} \tag{2}$$

The dimensions of the mass ($\mathbf{M}$) and stiffness ($\mathbf{K}$) matrixes are $3N \times 3N$, where N is the number of constituent atoms. Size reduction based on the idea of principal component analysis (a priori Karhunen-Loève analysis) leads to the following formulation [47].

$$\left(\mathbf{\Gamma}^{1/2}\mathbf{\Phi}\mathbf{\Gamma}^{1/2}\right)\left(\mathbf{\Gamma}^{1/2}\mathbf{\Xi}\right) = \left(\mathbf{\Gamma}^{1/2}\mathbf{\Xi}\right)\mathbf{\Omega}^2 \tag{3}$$

The matrices $\mathbf{\Gamma}^{-1}$ and $\mathbf{\Phi}$, which were named after the *GF* method [52], contain coarse-grained inertial loads and force constants, respectively. These matrices are given by basis transformation from the Cartesian to an internal coordinate system, using a transformation matrix (coarse-graining matrix) $\mathbf{B}$ ($3N \times n$), which is composed of a reduced number of atomic displacement vectors ($n \le 3N$) as basic motions [49]. The coarse-grained displacement $\widehat{\mathbf{\Xi}}$ is constructed using $\widehat{\mathbf{C}}$ ($3N \times n$), which is a partial matrix of the original displacement vectors, in case $n < 3N$.

$$\mathbf{\Gamma}^{-1} \equiv \mathbf{B}^{\mathrm{T}}\mathbf{M}\mathbf{B} \qquad (n \times n) \tag{4}$$

$$\mathbf{\Phi} \equiv \mathbf{B}^{\mathrm{T}}\mathbf{K}\mathbf{B} \qquad (n \times n) \tag{5}$$
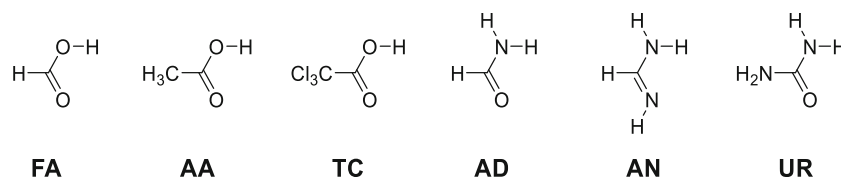

Fig. 1 Molecular structures of monomers studied. Formic acid (FA), acetic acid (AA), trichloroacetic acid (TC), formamide (AD), formamidine (AN), and urea (UR) are abbreviated as indicated

$$\hat{\widehat{\Xi}} \equiv \mathbf{\Gamma} \mathbf{B}^{\mathrm{T}} \mathbf{M} \hat{\mathbf{C}} \quad (n \times n) \tag{6}$$

Again, the coarse-grained MWD vectors (columns in $\mathbf{U}$) are prepared by multiplying $\mathbf{\Gamma}^{1/2}$ with Ξ and normalized with matrix $\mathbf{\Lambda}^{-1}$, where $\mathbf{\Lambda}^2$ is the modal mass matrix under the reduced dimensions of motions.

$$\begin{cases} \hat{\mathbf{U}} \equiv \mathbf{\Gamma}^{-1/2} \hat{\widehat{\Xi}} \mathbf{\Lambda}^{-1} & (n \times n) \\ \mathbf{\Lambda}^2 \equiv \hat{\widehat{\Xi}}^{\mathrm{T}} \mathbf{\Gamma}^{-1} \hat{\widehat{\Xi}} & (n \times n) \end{cases} \tag{7}$$

If the set of frequencies for normal mode vibrations are determined either computationally or experimentally, the force constants can be obtained under the reduced dimension of the selected motions (a partial matrix of $\mathbf{\Omega}^2$ is used if $n < 3N$). This set of force constants, combined with the appropriate information concerning inertial loads, such as molecular weight or moment of inertia, can reproduce the eigenfrequencies in the LF region with comparable accuracy to that derived from the original calculations or measurements.

$$\mathbf{\Phi} = \mathbf{\Gamma}^{-1/2} \hat{\mathbf{U}} \mathbf{\Omega}^2 \hat{\mathbf{U}}^{\mathrm{T}} \mathbf{\Gamma}^{-1/2} \qquad (n \times n) \tag{8}$$

It is more convenient to convert $\mathbf{\Phi}$ into $\widetilde{\mathbf{\Phi}}$ by applying weighting factors $\mathbf{N}^{1/2}$, so that we can compare the stiffness constants of different molecules with various molecular weights using the same standard. The matrix $\mathbf{N}^{1/2}$ is defined so that $\mathbf{N}^{1/2} \mathbf{\Gamma}^{-1} \mathbf{N}^{1/2}$ gives the molecular weight (M_{m}) and the tensor of inertia ($\mathbf{I}_{\mathrm{m}}$) of the m-th (m = I, II, ..., X) molecular unit [47].

$$\begin{cases} \widetilde{\mathbf{\Phi}} = \mathbf{N}^{1/2} \mathbf{\Phi} \mathbf{N}^{1/2} \\ \mathbf{N}^{1/2} \mathbf{\Gamma}^{-1} \mathbf{N}^{1/2} = \begin{pmatrix} M_{\mathrm{I}} \mathrm{E} & O & & & \\ O & \mathrm{I}_{\mathrm{I}} & \ddots & & \\ & & & M_{\mathrm{x}} \mathrm{E} & O \\ & & & O & \mathrm{I}_{\mathrm{x}} \end{pmatrix} \end{cases} \tag{9}$$

For the sake of simplicity, $\widetilde{\mathbf{\Phi}}$ is denoted $\mathbf{\Phi}$ in the rest of this paper.

For a given molecular assembly, its LF-mode vibrations can be approximately represented as a linear combination of six basic motions, namely, three translations ($\mathbf{T}_{\mathrm{m}x}$, $\mathbf{T}_{\mathrm{m}y}$, $\mathbf{T}_{\mathrm{m}z}$) and three librations ($\mathbf{R}_{\mathrm{m}x}$, $\mathbf{R}_{\mathrm{m}y}$, $\mathbf{R}_{\mathrm{m}z}$), for the m-th constituent molecule (hereafter called "unit m"), if they are substantially regarded as rigid bodies. For a molecular dimer (m = I or II), for example, the twelve basic motions in total are contained in the matrix $\mathbf{B}^{\circ}$, each column of which is implicitly normalized but not necessarily orthogonalized.

$$\begin{cases} \mathbf{B}^{\circ} = \begin{pmatrix} \mathbf{B}_{\mathrm{I}}^{\circ} & O \\ O & \mathbf{B}_{\mathrm{II}}^{\circ} \end{pmatrix} \\ \mathbf{B}_{\mathrm{m}}^{\circ} = \begin{pmatrix} \mathbf{T}_{\mathrm{m}x} & \mathbf{T}_{\mathrm{m}y} & \mathbf{T}_{\mathrm{m}z} & \mathbf{R}_{\mathrm{m}x} & \mathbf{R}_{\mathrm{m}y} & \mathbf{R}_{\mathrm{m}z} \end{pmatrix}, \; \mathrm{m} = \mathrm{I} \; \mathrm{or} \; \mathrm{II} \end{cases} \tag{10}$$

When the molecules have non-negligible flexibility, matrix $\mathbf{B}$ should additionally contain some displacement vectors $\mathbf{b}_{\mathrm{I},1}$, $\mathbf{b}_{\mathrm{I},2}$, $\mathbf{b}_{\mathrm{I},3}$, ... $\mathbf{b}_{\mathrm{I},k\mathrm{I}}$ and $\mathbf{b}_{\mathrm{II},1}$, $\mathbf{b}_{\mathrm{II},2}$, $\mathbf{b}_{\mathrm{II},3}$, ... $\mathbf{b}_{\mathrm{II},k\mathrm{II}}$, the intramolecular vibration modes of units I and II, respectively [48]. When all the intramolecular vibrations are included in the basis set (i.e., $k_{\mathrm{I}} = 3 N_{\mathrm{I}}$-6; $k_{\mathrm{II}} = 3 N_{\mathrm{II}}$-6), the eigenvectors are virtually identical to those obtained from the original Hessian matrix represented in the Cartesian coordinate system. However, when the number of bases is restricted to a certain number, the order of incorporating bases should be carefully selected to optimize the coarse-grained space. In other words, for the column numbers $\{j \mid 1 < j < 3(N_{\mathrm{I}} + N_{\mathrm{II}})\}$ of matrix $\mathbf{B}$, it is necessary to determine a new series of column vectors using permutation mapping σ (Eq. 12).

$$\widetilde{\mathbf{B} \sim} = \begin{pmatrix} \mathbf{B}^{\circ} & \mathbf{b}_1 \cdots \mathbf{b}_{k_{\mathrm{I}}+k_{\mathrm{II}}} \end{pmatrix}, \; \mathbf{b}_j$$
$$= \begin{pmatrix} \mathbf{b}_{\mathrm{I},\sigma(j)} \\ O \end{pmatrix} \mathrm{or} \begin{pmatrix} O \\ \mathbf{b}_{\mathrm{II},\sigma(j)} \end{pmatrix} \tag{11}$$

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & j & \cdots & 3(N_{\mathrm{I}} + N_{\mathrm{II}}) \\ \sigma(1) & \sigma(2) & & \sigma(j) & & \sigma(3(N_{\mathrm{I}} + N_{\mathrm{II}})) \end{pmatrix} \tag{12}$$

The squared elements of Ξ represent contributions from the basic motions and individual intramolecular vibrations to the normal mode vibration of the entire molecular system. To maintain Ξ as a regular matrix, the row dimension of $\mathbf{B}$ should coincide with that of $\mathbf{C}$. Therefore, for a given truncated $\mathbf{C}$ that contains displacement vectors of selected vibration modes, it is necessary to retain the suitable components of the basis transformation matrix, so that every displacement vector is satisfactorily represented as a linear combination of the column vectors of $\widetilde{\mathbf{B}}$. The tilde ($\sim$) on $\mathbf{B}$, $\mathbf{C}$, etc. denotes that they are partial matrixes of the original ones.

$$\widetilde{\mathbf{C}} \simeq \widetilde{\mathbf{B}} \hat{\Xi} \quad (3N \times n) \tag{13}$$

The construction of the $\widetilde{\mathbf{B}}$ matrix requires selection of the $\mathbf{b}_{\mathrm{m},k}$ vector, which is accomplished by extracting eigenvectors for the normal mode vibrations of an isolated single molecule as a model of the unit m (hereafter called "model m"). However, the structure of the model may change slightly after forming an assembly, especially when hydrogen bonding is involved. To remove the arbitrariness in choosing the coordination axes of the monomer, we determined the orientation of the models so that their principal axes of inertia coincide with those of the corresponding units arranged in the assembly (Fig. 2). Such an alignment is more important in handling heterodimers than in homodimers to minimize the error that may arise from the decreased symmetry. In addition, the definition of the coordinate system of the assembly is important because the coarse-grained force constants such as $\Phi_{\mathrm{T}x,\mathrm{T}x}$ correspond to the directions in a given coordinate. In the present

Fig. 2 Definition of terms used in this work (e.g., TC + AD dimer). Normal mode analysis is performed for an assembly that consists of two molecular units. Each model of the units is calculated independently, after which its orientation is adjusted to the corresponding unit in the assembly so that their principal axes (arrows in red, green, and blue) coincide
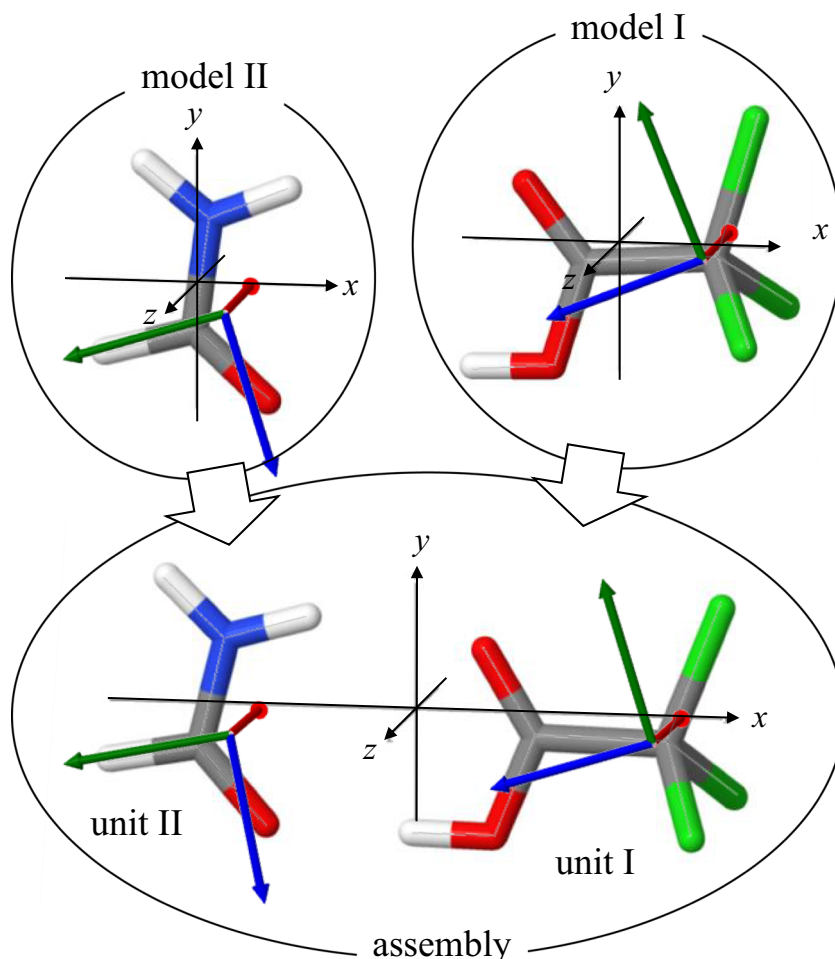
study, we defined the xy plane as the least-square plane of the four atoms participating in the double hydrogen bond and the x-axis as the average of two hydrogen-bonding vectors projected onto the xy plane.

To obtain an appropriate permutation mapping σ, we examined three ways of augmenting the matrix **B**, where the twelve basic motions are preferentially arranged in the original order (Fig. 3). First, similar to a previous study regarding homodimers [50], we incorporated the vectors in increasing order of the frequency of intramolecular vibration (mapping #1). Because the dimensions of the intramolecular vibration differ between the two units, the vectors cannot be alternately acquired from the two models. Second, we noted the elements of Ξ as a measure of the contributions of the bases to reproduced atomic displacement vectors in **C** (Eq. 13) (mapping #2). Third, we noted the elements of **U** as a measure of the contributions of the MWD bases to the reproduced atomic MWD vectors in **W** (mapping #3). This permutation reorders j such that the sum of squared $(\sigma(j), k)$ components of **U** ($1 < k < j$) is the maximum of $\{U_{\sigma(j)\,k}{}^{2} \mid \sigma(j) < k < 3(N_I + N_{II})\}$.

When we note the selected atomic displacement vectors contained in **C**, the MWD vectors are calculated using Eq. 14.

$$\begin{cases} \widetilde{\mathbf{W}} \equiv \mathbf{M}^{1/2}\widetilde{\mathbf{C}}\widetilde{\mathbf{L}}^{-1} \\ \widetilde{\mathbf{L}}^{2} \equiv \widetilde{\mathbf{C}}^{\mathrm{T}}\mathbf{M}\widetilde{\mathbf{C}} \end{cases} \tag{14}$$

Substituting **C** in Eq. 14 with that in Eq. 13, we can obtain the MWD vectors $(\overline{\mathbf{W}})$ reproduced in the coarse-grained space.

$$\begin{aligned} \overline{\mathbf{W}} &\equiv \mathbf{M}^{1/2}\widetilde{\mathbf{B}}\widehat{\Xi}\Lambda^{-1} \\ &= \mathbf{M}^{1/2}\widetilde{\mathbf{B}}\Gamma^{1/2}\widehat{\mathbf{U}} \end{aligned} \tag{15}$$

The quality of the coarse-grained space can be evaluated by comparing the MWD vectors on the original atomistic Cartesian coordinate bases with those reproduced on the reduced bases. We previously proposed the severest criterion F_4, which is the determinant of the correlation matrix **R**. The matrix **R** can be further rewritten as the product of $\widetilde{\mathbf{L}}^{-1}$ and Λ, and the diagonal elements are the ratio of the reproduced modal masses to the original ones.

$$\mathbf{R} \equiv \widetilde{\mathbf{W}}^{\mathrm{T}}\mathbf{W} = \widetilde{\mathbf{L}}^{-1}\Lambda \tag{16}$$

Fig. 3 Flowchart for constructing the coarse-grained space. The cards titled mappings #1, #2, and #3 can replace each other to compare the validity of those methods for selecting the bases
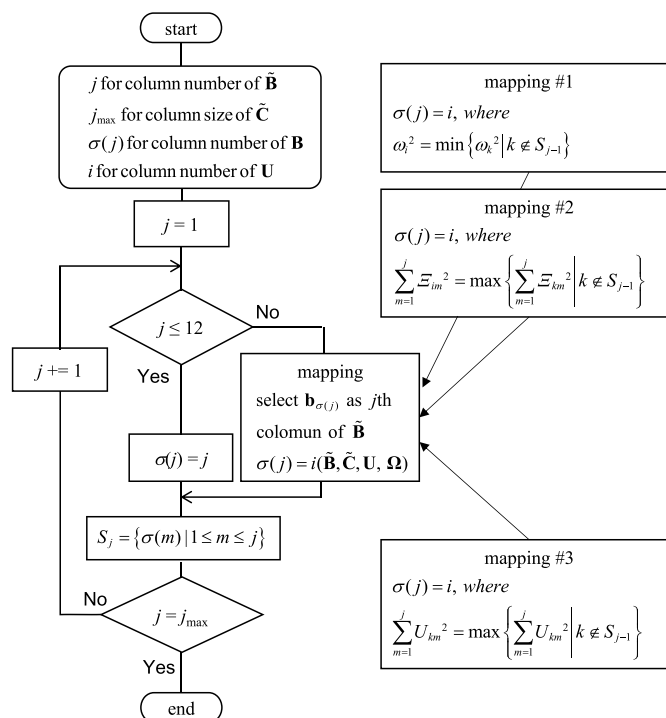
$$F_4 = \det(\mathbf{R}) \tag{17}$$

If the matrix $\overline{\mathbf{W}}$ is identical to $\widetilde{\mathbf{W}}$, the matrix $\mathbf{R}$ is equal to the unit matrix $\mathbf{E}$, meaning that F_4 is maximized to unity when the coarse-grained space is complete to describe the normal-mode vibrations of the assembly. Therefore, $\widetilde{\mathbf{W}}$ is properly approximated by $\overline{\mathbf{W}}$, which is further rewritten from Eq. 15 into its singular value decomposition (SVD), where $\widetilde{\mathbf{H}}$ is a partial matrix containing MWD bases modified by Löwdin symmetry orthogonalization, $\mathbf{J}$ is an orthonormal matrix, and $\rho^{1/2}$ is a set of singular values.

$$\begin{cases} \widetilde{\mathbf{W}} \simeq \mathbf{W} \equiv \widetilde{\mathbf{H}} \rho^{1/2} \mathbf{J}^{\mathrm{T}} \\ \widetilde{\mathbf{H}} \equiv \mathbf{M}^{1/2} \widetilde{\mathbf{B}} \mathbf{\Gamma}^{1/2} \\ \mathbf{J}^{\mathrm{T}} \equiv \rho^{-1/2} \widehat{\mathbf{U}} \end{cases} \tag{18}$$

By definition of SVD theory, the square of singular values is obtained by diagonalizing $\overline{\mathbf{W}} \overline{\mathbf{W}}^{\mathrm{T}}$ (Eq. 19). Consequently, $\widetilde{\mathbf{H}}$ is a partial matrix of $\widehat{\mathbf{H}}$, an orthonormal matrix that diagonalizes $\overline{\mathbf{W}} \overline{\mathbf{W}}^{\mathrm{T}}$, of which the eigenvalues are the diagonal elements of $\mathbf{U} \mathbf{U}^{\mathrm{T}}$, namely, unity.

$$\rho = \widehat{\mathbf{H}}^{\mathrm{T}} \left(\overline{\mathbf{W}} \overline{\mathbf{W}}^{\mathrm{T}} \right) \widehat{\mathbf{H}} = \widehat{\mathbf{U}} \widehat{\mathbf{U}}^{\mathrm{T}} \tag{19}$$

In a coarse-grained space, some of the eigenvalues of $\overline{\mathbf{W}} \overline{\mathbf{W}}^{\mathrm{T}}$ become substantially zero (rank deficient), meaning that we can pick up some non-zero singular values (ρ_i) that can properly approximate $\widetilde{\mathbf{W}}$ based on Eq. 18. The singular value is given by Eq. 20, which is actually an indicator maximized when selecting the intramolecular vibration vector in mapping #3.

$$\rho_i \simeq \left(\widetilde{\mathbf{U}} \widetilde{\mathbf{U}}^{\mathrm{T}} \right)_{ii} = \sum_m^i U_{im}^2 \tag{20}$$

The above formulation supports the use of mapping #3 in reordering matrix $\mathbf{B}$. Mapping #2 also appears rational to some extent, but the non-orthonormality of the $\mathbf{\Xi}$ matrix shows that the diagonal elements of $\mathbf{\Xi}^{\mathrm{T}} \mathbf{\Xi}$ cannot be singular values of any form of SVD.

## Computational details

According to the formulation described in the "Theory" section, we modified our in-house program to calculate the $\mathbf{\Phi}$ matrix, so that it is applicable to heterodimers. The molecules shown in Fig. 1 were used to construct 21 combinations, including 6 homodimers and 15 heterodimers (Fig. 4). The geometry of these dimers was optimized using the Hartree-Fock method with the 6-311G** basis set and subsequently subjected to normal vibration analysis at the same level of calculation. The energies of the monomers and dimers were used to calculate the dimerization energy (E_{assoc}) according to Eq. 21,

$$E_{\mathrm{assoc}} = (E_{\mathrm{model\ I}} + E_{\mathrm{model\ II}}) - E_{\mathrm{assembly}} \tag{21}$$

where $E_{\mathrm{model\ I}}$ and $E_{\mathrm{model\ II}}$ are the energies of models I and II, and E_{assembly} is that of the dimer. The basis set superpositional error was corrected by the counterpoise method. These

Fig. 4 Molecular structures of homo- and heterodimers studied
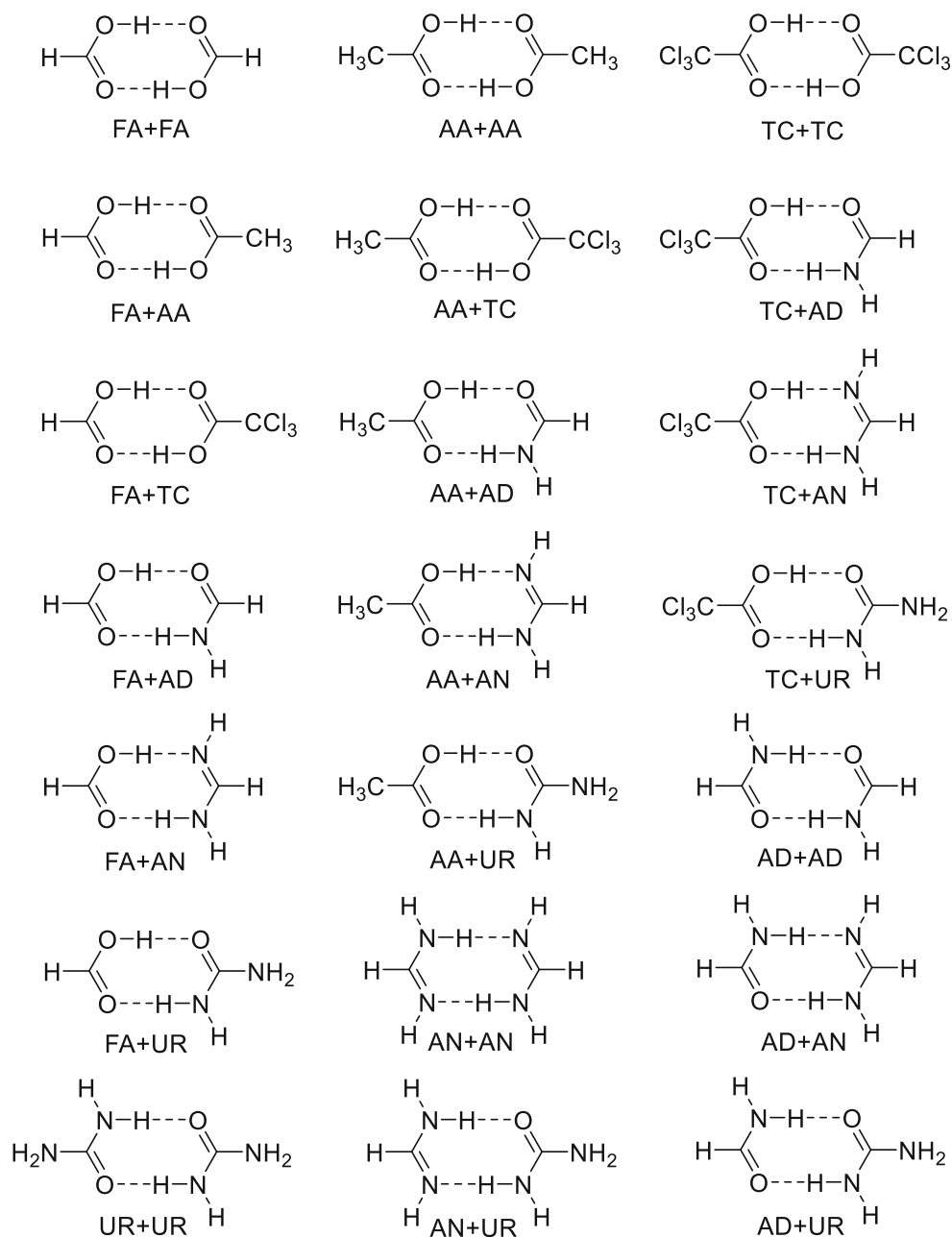
molecular orbital calculations were performed in the GAUSSIAN 09 W programs [53]. Visualization of the molecules was performed with Jmol ver.14.4.0 [54].

Results and discussion

As a representative heterodimer, the results of TC + AD are described here. Figure 5 plots the fidelity index (F_4 in Eq. 17) as a function of the dimension of the coarse-grained space that was constructed according to methods #1–3 in the "Theory" section. In Fig. 5a, where the bases were arranged in increasing frequency order, the F_4 value swings between high and low levels at random, indicating that the bases were not properly selected, especially for the space of less than 25 dimensions. In Fig. 5b, where the bases were arranged in decreasing $(\mathbf{\Xi}^{\mathrm{T}}\mathbf{\Xi})_{ii}$ order, the F_4 value seems much improved, maintaining a sufficiently high level, although there are moderate depressions at 20, 24, 27, and 32 dimensions. As shown in Fig. 5c, where the bases were arranged in decreasing $(\mathbf{U}^{\mathrm{T}}\mathbf{U})_{ii}$ order, the F_4 value maintains a high level as well, and the depression at 24–27 dimensions was significantly improved. As mentioned in the "Theory" section, the sorting scheme of mapping #3 seems quite reasonable, so that we can properly construct a coarse-grained space for a given dimension, although there are some occasional collapses. Such collapses imply a serious
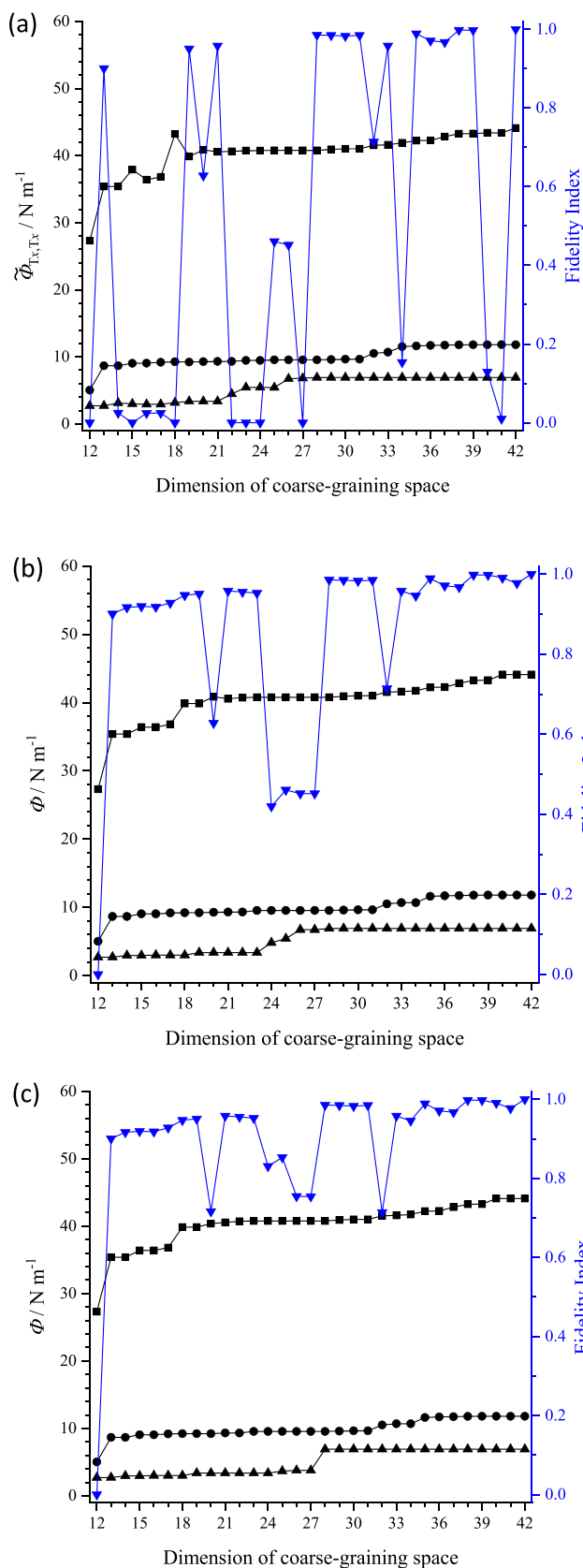
unsatisfactory basis set, namely, a lack of a certain vibrational motion crucial to the normal mode vibrations, which is partly due to the restriction of Ξ and **U** to square matrixes. For all reordering methods, the fidelity index is extremely low when the dimension is 12, which can be explained by the lack of C–C bond libration (#1 vibration mode of TC model).

As demonstrated previously, we are interested in intermolecular stiffness with respect to a double hydrogen bond direction, namely, the diagonal element of Φ with respect to T_x, T_y, and T_z motions. Figure 5 overlays the elements $\Phi_{Tx,Tx}$, $\Phi_{Ty,Ty}$, and $\Phi_{Tz,Tz}$ as a function of the column dimension of **B** [55]. The line graph shows a stepwise increment with the increasing number of bases employed. This discontinuity indicates that the translational motions of the two units are coupled with a limited number of intramolecular vibrations: in a poor basis set, an apparent stiffness with respect to this motion is to some extent underestimated to compensate for the neglected intramolecular flexibility [49]. The step height between the neighboring plots is a measure of coupling with the added vibration mode. Similar to the case already reported for homodimers, the continuity of Φ elements is not greatly disturbed, even if the fidelity index sharply decreases, except for the 12-dimensional space, which lacks an essential basis responsible for the molecular flexibility. Thus, we can define the minimum coarse-graining dimension to be 13 for the TC + AD case.

Similar results were obtained for the other homo- and heterodimers, suggesting that reordering mapping #3 was reliable for constructing the coarse-grained space. The results of the $\Phi_{Tx,Tx}$ plots against the dimension of the space, together with the fidelity index F_4, are shown in Fig. S1. For all dimers, the trend of the $\Phi_{Tx,Tx}$ plot was similar to that of TC + AD, and the F_4 value remained at a moderately high level in many dimensions. For the dimers that contain either AA or TC, there was an LF vibration mode that strongly contributed from intramolecular C–C libration mode; hence, the minimum coarse-grained space comprised 13 dimensions. For AA+AA, AA+TC, and TC + TC dimers, the minimum dimensionality was 14. For the other dimers, coarse-graining in the 12-dimensional space afforded a sufficiently high value for the fidelity index.

Figure 6a displays a heatmap representation of the full matrix **U** for a dimer composed of TC and AD before reordering. The column index i designates a consecutive number of translations ($1 \leq i \leq 3$), librations ($4 \leq i \leq 6$), and normal mode vibrations ($7 \leq i \leq 42$) of the TC + AD dimer, while the row index j designates a consecutive number of translations and librations, as well as normal mode vibrations of TC ($1 \leq j \leq 6$,
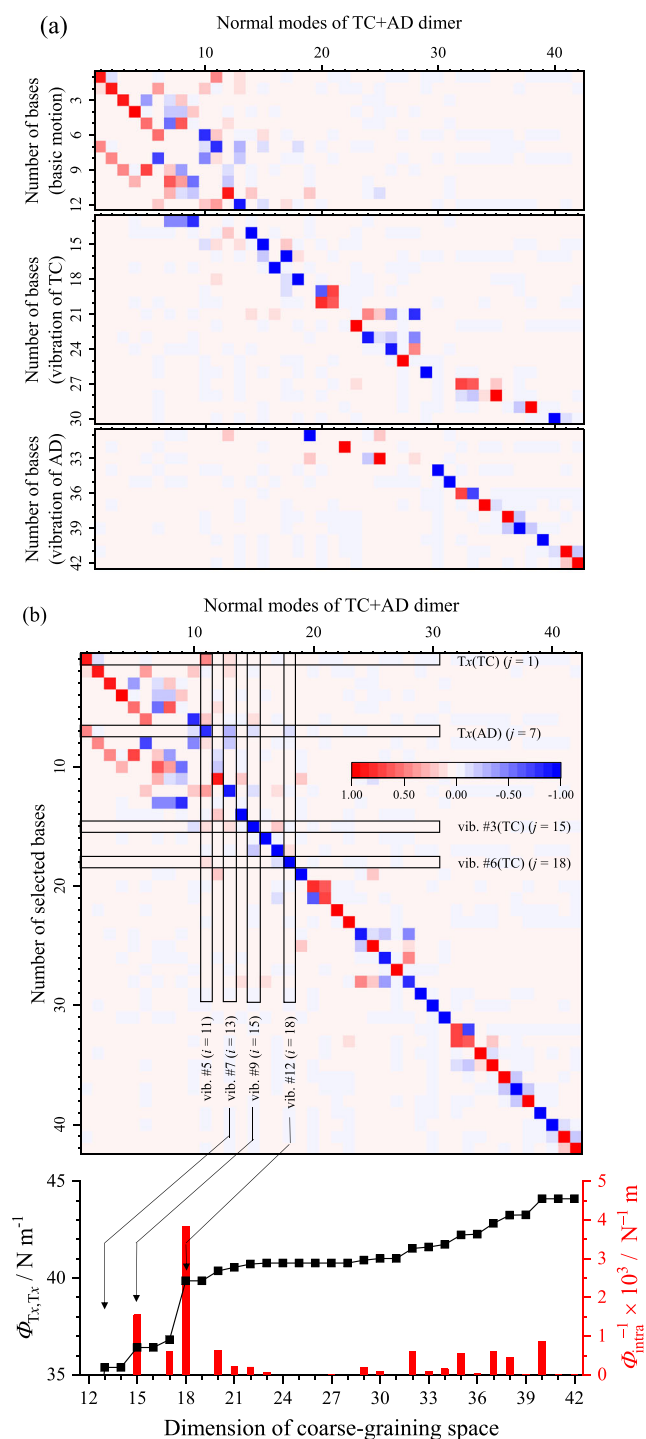
(a)

(b)

Fig. 6 Heatmap of the matrix **U** for TC + AD dimer. **a** Bases arranged by frequency for each model; **b** bases arranged by selected order using mapping #3. Enlarged versions of original figures are given in Figs. S2 and S3
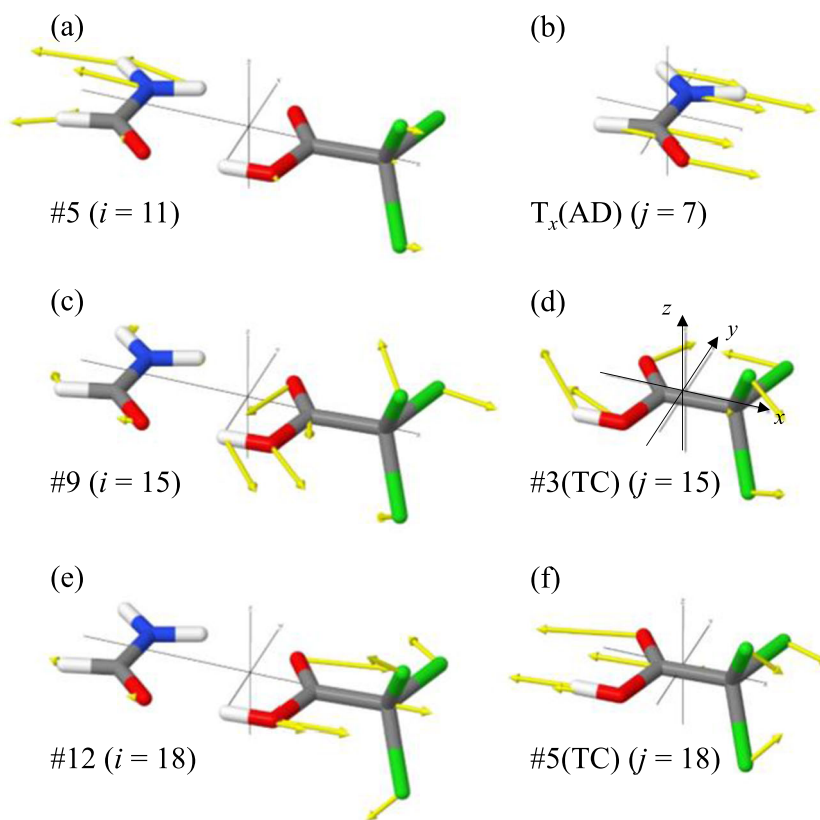
selection scheme, mapping #3 (Fig. 6b). For example, the #5 vibrational mode of the dimer ($i = 11$) is predominantly attributed to T_x motions of TC ($j = 1$) and AD ($j = 7$) with an opposite sign, together with an additional contribution from the R_z motion of AD ($j = 12$). This can be visually understood from the molecular model (Fig. 7a), which shows that the #5 vibration is attributed to the stretching mode between TC and AD, although the center of mass is quite biased toward the TC side; hence, it is well approximated with the T_x motion of the AD model (Fig. 7b). In other words, regarding the TC + AD dimer, the #5 vibration directly reflects information about the intermolecular force of their double hydrogen bonding.

Figure 6b shows that the T_x motions of TC and AD also contribute to the #13, #15 (Fig. 7c), and #18 (Fig. 7e) vibrations of the dimer, and the main components of those vibrations are the R_z motion of AD ($j = 12$), #3 mode of TC ($j = 15$) (Fig. 7d), and #6 mode of TC ($j = 18$) (Fig. 7f), respectively. Consequently, these intramolecular vibrations are moderately coupled with the T_x motions of the constituent units through hydrogen bonding. As described earlier, this type of coupling is responsible for the stepwise behavior of the Φ plot. This is easily understood if one imagines that the assembly is approximated as a series connection of springs, where the apparent force constant is underestimated because of the neglect of the coupled springs. After dissolving all of the contributions from coupled springs, we can obtain the true force constant of the spring of interest. Regarding the intermolecular hydrogen bonding interactions discussed in this study, we denote the apparent stiffness in the minimum coarse-grained space (mostly, the fidelity index >0.8) as Φ_{app} and the true stiffness in the full-dimensional space as Φ_{inter} [48]. Table 1 summarizes the values of Φ_{app} and Φ_{inter} for all molecular dimers studied.

Concerning the physical meaning of the step height of the Φ graphs, we previously proposed an analysis of intermolecular vibrations based on a coupled oscillator model that contains a series connection of three springs. The formulation of Φ_{intra} (Eq. 22) was initially derived from the rigorous treatment of a symmetric four-body model and later proved to approximately hold for an asymmetric model. According to Eq. 22, we can estimate the intramolecular stiffness Φ_{intra} averaged over TC and AD. For example, this value is calculated to be 115 N m^{-1} using Eq. 22, with values of $\Phi_{app} = 35.4$ N m^{-1} and $\Phi_{inter} = 44.1$ N m^{-1}. Table 2 summarizes the average intramolecular stiffness for all the dimers studied. The diagonal items in this table are the values from homodimers, namely, intramolecular stiffness intrinsic to respective molecules, and some of them are refinements of our previous report [48, 56]. TC exhibits an extremely low value (86.2 N m^{-1}), probably because of its flexible C–Cl bond as compared with a C-H bond. Conversely, AD gave the highest value (370 N m^{-1}) among the six molecules, reflecting its rigid

$13 \leq j \leq 30$) and AD ($7 \leq j \leq 12$, $31 \leq j \leq 42$). We can see from this figure that the modes of unit motion largely contribute to the mode of vibration of the assembly, while it becomes clearer when the rows are reordered according to the basis set

Fig. 7 Schematic representation of the molecular motion for **a** #5, **c** #9, and **e** #12 vibration modes of TC + AD dimer; **b** T_x motion of AD; **d** #3, and **f** #6 vibration modes of TC

(a) #5 ($i = 11$)

(b) T_x(AD) ($j = 7$)

(c) #9 ($i = 15$)

(d) #3(TC) ($j = 15$)

(e) #12 ($i = 18$)

(f) #5(TC) ($j = 18$)

skeleton supported by extended π-conjugation.

$$\Phi_{intra} = \frac{\Phi_{inter}^2 - \left(\Phi_{inter} - \Phi_{app}\right)^2}{2\left(\Phi_{inter} - \Phi_{app}\right)} \tag{22}$$

The idea of the averaged intramolecular stiffness is applied to the individual one for each contribution of the intramolecular vibration mode taken into account in the n-th step. In Eq. 23, $\widetilde{\Phi}_{app,n}$ and $\widetilde{\Phi}_{intra,n}$ are the apparent stiffness and newly contributed intramolecular stiffness, respectively, in the dimension of n.

$$\Phi_{intra,n} = \frac{\Phi_{app,n}^2 - \left(\Phi_{app,n} - \Phi_{app,n-1}\right)^2}{2\left(\Phi_{app,n} - \Phi_{app,n-1}\right)} \tag{23}$$

Here, we define Φ_{ser}^{-1} as the sum of $\Phi_{intra,n}^{-1}$ over n from the minimum to the full dimensions of the coarse-grained space. This corresponds to the synthetic compliance of a series connection of springs.

$$\Phi_{ser}^{-1} = \sum_n^{3N} \Phi_{intra,n}^{-1} \tag{24}$$

Interestingly, the plot of Φ_{ser} against Φ_{intra} shows excellent linearity over all 21 dimers, with an incline of nearly 1

Table 1 Values of force constants (in N m^{-1}) for Φ_{app} (left) and Φ_{inter} (right)

	FA	AA	TC	AD	AN	UR
FA	41.2/44.8	41.5/45.8	36.5/44.7	38.0/40.7	39.5/44.5	38.7/43.1
AA		41.7/46.7	37.2/46.5	37.3/40.5	38.3/43.4	38.0/42.5
TC			32.2/42.9	35.4/44.1	37.8/51.9	36.6/47.7
AD				31.2/32.7	29.4/31.4	31.2/33.2
AN					27.1/29.3	29.1/31.4
UR						31.1/33.6

Table 2 Intramolecular stiffness of molecular dimers derived from Eq. 22

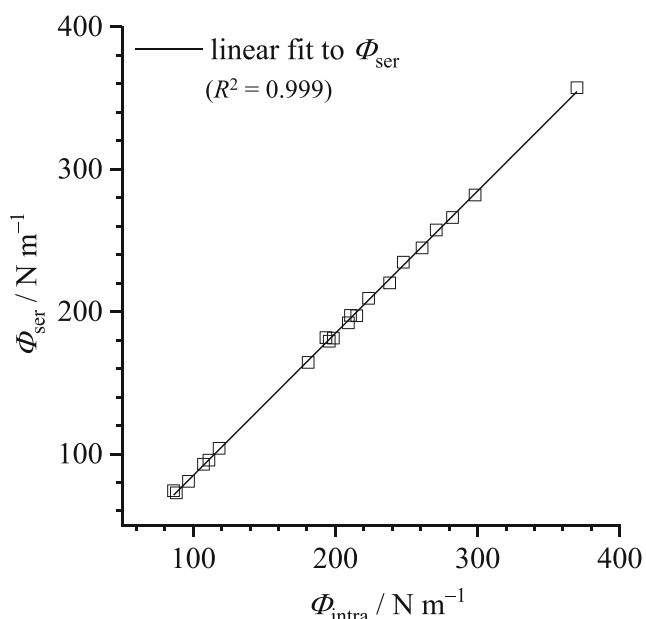|	FA	AA	TC	AD	AN	UR
FA	282	238	118	298	196	209
AA		215	111	261	181	199
TC			86.2	107	88.2	96.8
AD				370	248	271
AN					193	211
UR						223

Fig. 8 Correlation between Φ_{ser} (synthetic stiffness based on a series connection model of springs) and Φ_{intra} (averaged intramolecular stiffness)

(incline = 0.999; intercept = −15.0 N m^{-1}; R^2 = 0.999) (Fig. 8). This result suggests that the internal stiffness of the molecule is well explained by a series connection of springs, for each of which the force constant is $\Phi_{intra,n}$: for a coarse-grained space with a small dimension, most springs are "locked" to behave as rigid bodies and the apparent stiffness of the entire coupled oscillator is somewhat smaller than the true Φ_{inter}. As the dimension increases, the rigid bodies are "unlocked" to behave as springs and the apparent stiffness approaches the true value of Φ_{inter}.
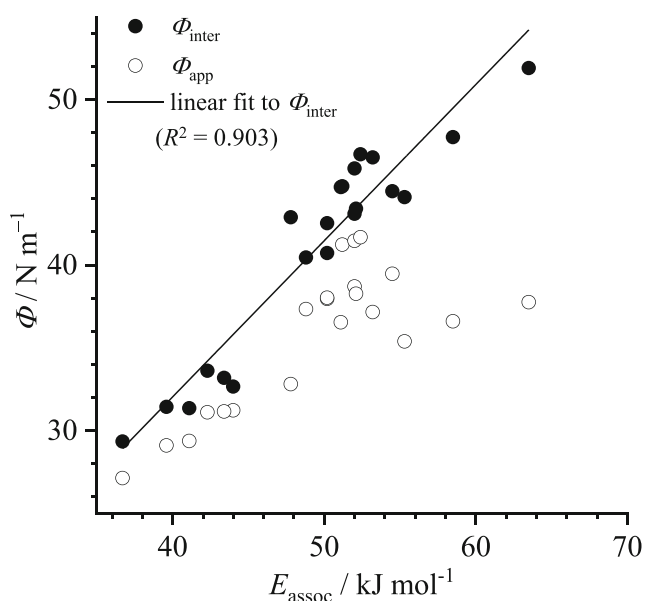

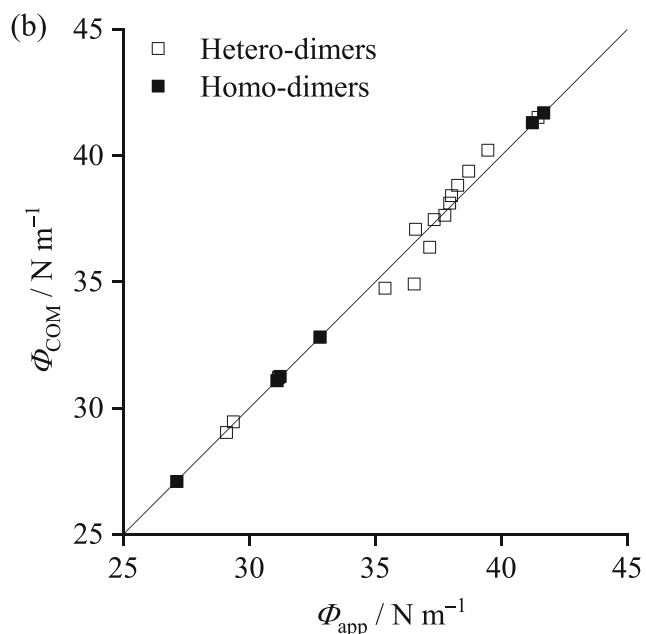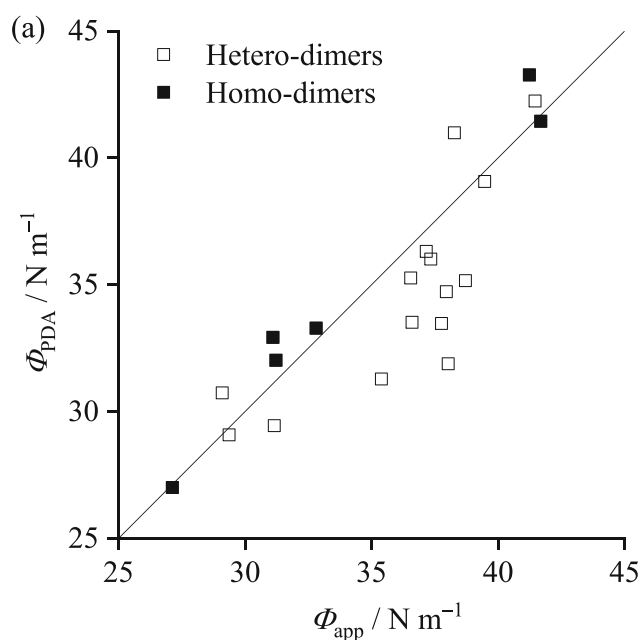Fig. 9 Correlations among Φ_{inter}, Φ_{app}, and E_{assoc}

Fig. 10 a Comparison of Φ_{PDA} vs. Φ_{app}; b comparison of Φ_{COM} vs. Φ_{app}

The Φ_{inter} values reflect the intrinsic nature of intermolecular interactions, namely, hydrogen bonding interactions in the present case. Therefore, we can expect some relationship between the Φ_{inter} and the association energy of the dimer, E_{assoc}. The smallest and largest association energies among those studied were 36.7 (AN+AN) and 63.5 (TC + AN) kJ mol^{-1}, respectively, in which dimers had the smallest and largest Φ_{inter} values of 29.3 and 51.9 N m^{-1}. Figure 9 shows that both values demonstrate a moderate positive correlation, with an incline of 9.4×10^{-4} N m^{-1}/J mol^{-1} and an intercept of

-5.7 N m^{-1} ($R^2 = 0.903$). It seems natural but not quite obvious that the more the assembly is energetically favored, the higher the intermolecular stiffness is. In other words, E_{assoc} and Φ_{inter} are independent parameters that determine the depth and curvature of the potential curve. The unit of the incline can be converted to mol m^{-2}, implying that the inverse square root of this value indicates a distance somewhat characteristic to these series of double hydrogen-bonded dimers. Based on the harmonic oscillator model, we can propose a simple relationship (Eq. 25) between energy E and distance D. A characteristic distance D_0 can then be defined as the half-width of the hypothetical quadratic potential curve at $E = 0$, which is a measure of the positional tolerance of the constituent molecules. The linear relationship in Fig. 9 affords an average D_0 value of 1.9 Å for the series of double hydrogen-bonded dimers in this study. This index might change depending on the nature of intermolecular interactions, including π-stacking, van der Waals forces, and hydrogen bonding.

$$\begin{cases} E = \dfrac{1}{2} N_A \Phi_{inter} D^2 - E_{assoc} \\ D_0 = \sqrt{\dfrac{2 E_{assoc}}{N_A \Phi_{inter}}} \end{cases} \tag{25}$$

For Φ_{app}, conversely, there was an appreciable deviation from the least-square line (line not shown). In view of the nature of the Φ_{app}, which is the apparent stiffness that has been modified by coupling with intramolecular vibrations, it seems natural that it shows no apparent correlation with Φ_{inter}. The discrepancy between Φ_{app} and Φ_{inter} is attributed to the following reasons: (i) the stretching mode vibration of a molecular dimer is not always pure antisymmetric translational motion, but rather a mixture of shearing, libration, and so on; (ii) neglect of the intrinsic flexibility of the constituent molecules causes serious errors in calculating the force constant. For reason (i), contributions other than antisymmetric translation became appreciably large when the symmetry axes of the monomers did not coincide with the direction of hydrogen bonding. Even though a moderately symmetric AA+ AA dimer is employed, the stretching mode has a nearly 20% contribution from $R_z + R_z$ motion, and the situation seems similar for all the dimers studied in this paper. The effect arising from reasons (i) and (ii) will be highlighted by a comparison of the Φ_{app} with the force constant Φ_{PDA}

derived from pseudodiatomic approximation (PDA), in which each constituent molecule is regarded as a frozen rigid body with a given molecular weight [57]. The force constant Φ_{PDA} can be calculated with the angular frequency $\omega_{stretch}$ for stretch mode vibration and the reciprocal masses Γ_I and Γ_{II}, which are the inverse of the molar masses of M_I and M_{II}, respectively. Because of its simplified derivation, it is often pointed out that there is only a poor correlation between the force constants based on PDA and the association energy [58]. Figure 10(a) shows a plot of the Φ_{PDA} values against the Φ_{app}, demonstrating sizable deviation from the diagonal line, although there is a moderate correlation with the homodimer data.

$$\Phi_{PDA} = \frac{\omega_{stretch}^2}{\Gamma_I + \Gamma_{II}} \tag{26}$$

Next, we attempt to obtain an explicit representation of the coupling of vibration with respect to the inter- and intramolecular stiffness. To this end, we have previously employed a four-body coupled oscillator model (COM) composed of four masses, A, B, C, D, and springs between them (Fig. S4). One monomer is approximated with a weight M_I equally distributed on A and B, and a spring with a force constant of $\Phi_{intra,I}$. Another monomer with a weight of M_{II} is represented by a similar oscillator composed of C and D connected with a spring of $\Phi_{intra,II}$. The two oscillators are connected with a spring with a force constant of Φ_{inter} ($<<\Phi_{intra,I}$, $\Phi_{intra,II}$) between B and C, and the four bodies are constrained to one-dimensional movement. Diagonalization of the mass-weighted Hessian matrix of this model is performed, where the reciprocal masses Γ_I and Γ_{II} again give us a set of squared angular frequencies as eigenvalues. The lowest non-zero solution, ω_{LF}^2, corresponds to a vibration mode primarily attributed to the stretching of B–C. Meanwhile, high-frequency solutions $\omega_{HF(ap)}^2$ and $\omega_{HF(ip)}^2$ correspond to the anti-phase and in-phase vibrations, respectively, localized at the A-B and C-D pairs. Eq. 28 allows us to calculate an apparent force constant Φ_{COM} for the spring between B and C. This procedure assumes consolidation of the oscillators A–B and C–D pairs, but moves far beyond the simple pseudodiatomic approximation, in that we can properly take into account the intramolecular stiffness of the constituent molecules.

$$\begin{pmatrix} \Gamma_I \Phi_{intra,I} & -\Gamma_I \Phi_{intra,I} & 0 & 0 \\ -\Gamma_I \Phi_{intra,I} & \Gamma_I(\Phi_{intra,I} + \Phi_{inter}) & -\sqrt{\Gamma_I \Gamma_{II}} \Phi_{inter} & 0 \\ 0 & -\sqrt{\Gamma_I \Gamma_{II}} \Phi_{inter} & \Gamma_{II}(\Phi_{intra,II} + \Phi_{inter}) & -\Gamma_{II} \Phi_{intra,II} \\ 0 & 0 & -\Gamma_{II} \Phi_{intra,II} & \Gamma_{II} \Phi_{intra,II} \end{pmatrix} U = U \begin{pmatrix} 0 & & & O \\ & \omega_{LF}^2 & & \\ & & \omega_{HF(ap)}^2 & \\ O & & & \omega_{HF(ip)}^2 \end{pmatrix} \tag{27}$$

$$\Phi_{COM} = \frac{\omega_{LF}^2}{\Gamma_I + \Gamma_{II}} \tag{28}$$

Because it is difficult for the eigenvalue equation, Eq. 27, to derive an analytical solution in the case of a heterodimer, we numerically calculated the Φ_{COM} values for all 21 dimers. Figure 10(b) shows an excellent correlation between Φ_{COM} and Φ_{app}, meaning that Φ_{app} can be explained by the lowest-frequency mode of the coupled oscillator model. In a coarse-grained space with the minimum dimension, Φ_{app} corresponds to the stiffness of the intermolecular force, but is to some extent underestimated to compensate for the neglected intra-molecular flexibility. As the number of bases increases, the contributions from the monomers are gradually decoupled. Thus, the force constant approaches Φ_{inter}, which is the true stiffness of the central spring.

Conclusions In this study, we developed a method of coarse-graining the intermolecular vibrations of molecular heterodimers assembled by double hydrogen bonding. In contrast to our previous study on homodimers, the coarse-grained space for heterodimers needs to be constructed with special care so that all normal-mode vibrations can be reproduced with a fidelity that is as high as possible at every dimension adopted. To this end, we tested three schemes of reordering the intra-molecular vibration vectors to determine the best method to note the contributions of such vectors to the mass-weighted displacement in the coarse-grained space. This method is shown to be equivalent to the size-reduction of the MWD matrix based on singular value decomposition. Using this method, we successfully obtained the apparent stiffness constants at each dimension, resulting in a monotonously increasing stepwise graph. The stepwise behavior is rationalized in terms of the mechanics of a series connection of springs, which accounts for the internal flexibility of a molecule. The true stiffness constants of the intermolecular force were obtained at the high-dimension limit of the coarse-grained space. Interestingly, the true stiffness constants show a good correlation with the association energies of the dimers, based on which we have proposed a new parameter (a characteristic distance D_0) to evaluate the nature of the interaction. The set of stiffness constants, Φ_{intra} and Φ_{inter}, obtained for each dimer are consistent enough to reproduce the Φ_{app} based on the coupled oscillator model, showing that the model is valid, far beyond the pseudodiatomic model, for heterodimers as well as homodimers. The present work provides a firm foundation for our coarse-graining theory with application to the mechanical properties of molecular crystals and other hetero-geneous molecular assemblies.

Declarations

References

1. Desiraju GR (2013) Crystal engineering: from molecule to crystal. J Am Chem Soc 135:9952–9967. https://doi.org/10.1021/ja403264c

2. Braga D, Grepioni F, Maini L (2010) The growing world of crystal forms. Chem Commun (Camb) 46:6232–6242. https://doi.org/10.1039/c0cc01195a

3. Erk P, Hengelsberg H, Haddow MF, van Gelder R (2004) The innovative momentum of crystal engineering. CrystEngComm 6: 474. https://doi.org/10.1039/b409282a

4. Yu L (2010) Polymorphism in molecular solids: an extraordinary system of red, orange, and yellow crystals. Acc Chem Res 43: 1257–1266. https://doi.org/10.1021/ar100040r

5. Reddy JP, Swain D, Pedireddi VR (2014) Polymorphism and phase transformation behavior of solid forms of 4-amino-3,5-dinitrobenzamide. Cryst Growth Des 14:5064–5071. https://doi.org/10.1021/cg500673a

6. Zimmermann A, Frøstrup B, Bond AD (2012) Polymorphs of pridopidine hydrochloride. Cryst Growth Des 12:2961–2968. https://doi.org/10.1021/cg300185n

7. Zhang Q, Mei X (2016) Two new polymorphs of huperzine A obtained from different dehydration processes of one monohydrate. Cryst Growth Des 16:3535–3542. https://doi.org/10.1021/acs.cgd.6b00493

8. Docherty R, Pencheva K, Abramov YA (2015) Low solubility in drug development: de-convoluting the relative importance of sol-vation and crystal packing. J Pharm Pharmacol 67:847–856. https://doi.org/10.1111/jphp.12393

9. Ramel PR, Marangoni AG (2017) Insights into the mechanism of the formation of the most stable crystal polymorph of milk fat in model protein matrices. J Dairy Sci 100:6930–6937. https://doi.org/10.3168/jds.2017-12758

10. Paulus EF, Leusen FJJ, Schmidt MU (2009) Crystal structures of quinacridones. CrystEngComm 9:131–143. https://doi.org/10.1039/b613059c

11. Chung H, Diao Y (2016) Polymorphism as an emerging design strategy for high performance organic electronics. J Mater Chem C 4:3915–3933. https://doi.org/10.1039/c5tc04390e

12. Arod F, Gardon M, Pattison P, Chapuis G (2005) The ??2-poly-morph of salicylideneaniline. Acta Crystallogr C Cryst Struct Commun 61:317–320. https://doi.org/10.1107/S010827010500911X

13. Kawato T, Kanatomi H, Amimoto K, Koyama H (1999) Enclathrated solvent effects on the photochromic process of N-salicylideneaniline in the crystal state. Chem Lett 28:47–48

14. Taneda M, Amimoto K, Koyama H, Kawato T (2004) Photochromism of polymorphic 4 , 4J-methylenebis- (N-salicylidene-2 , 6-diisopropylaniline) crystals. Org Biomol Chem 2:499–504

15. Safin DA, Robeyns K, Babashkina MG et al (2016) Polymorphism driven optical properties of an anil dye. CrystEngComm 18:7249–7259. https://doi.org/10.1039/C6CE00266H

16. Carletta A, Buol X, Leyssens T et al (2016) Polymorphic and isomorphic cocrystals of a N-salicylidene-3-aminopyridine with dicarboxylic acids: tuning of solid-state photo- and thermochromism. J Phys Chem C 120:10001–10008. https://doi.org/10.1021/acs.jpcc.6b02734

17. Carletta A, Spinelli F, d'Agostino S et al (2017) Halogen-bond effects on the thermo- and photochromic behaviour of anil-based molecular co-crystals. Chem Eur J 23:5317–5329. https://doi.org/10.1002/chem.201605953

18. Hutchins KM, Dutta S, Loren BP, Macgillivray LR (2014) Co-crystals of a salicylideneaniline: photochromism involving planar dihedral angles. Chem Mater 26:3042–3044. https://doi.org/10.1021/cm500823t

19. Mercier GM, Robeyns K, Leyssens T (2016) Altering the photochromic properties of N-salicylideneanilines using a co-crystal engineering approach. Cryst Growth Des 16:3198–3205. https://doi.org/10.1021/acs.cgd.6b00108

20. Haneda T, Kawano M, Kojima T, Fujita M (2007) Thermo-to-photo-switching of the chromic behavior of salicylideneanilines by inclusion in a porous coordination network. Angew Chem Int Ed 46:6643–6645. https://doi.org/10.1002/anie.200700999

21. Jacquemin PL, Robeyns K, Devillers M, Garcia Y (2015) Photochromism emergence in N-salicylidene p-aminobenzenesulfonate diallylammonium salts. Chemistry 21:6832–6845. https://doi.org/10.1002/chem.201406573

22. Carletta A, Colaco M, Mouchet SR, Plas A, Tumanov N (2018) Tetraphenylborate anion induces photochromism in N-salicylideneamino-1-alkylpyridinium derivatives through formation of tetra-aryl boxes. J Phys Chem C 122:10999–11007. https://doi.org/10.1021/acs.jpcc.8b02329

23. Johmoto K, Sekine A, Uekusa H (2012) Photochromism control of salicylideneaniline derivatives by acid-base co-crystallization. Cryst Growth Des 12:4779–4786. https://doi.org/10.1021/cg300454q

24. Saha S, Mishra MK, Reddy CM, Desiraju GR (2018) From molecules to interactions to crystal engineering: mechanical properties of organic solids. Acc Chem Res 51:2957–2967. https://doi.org/10.1021/acs.accounts.8b00425

25. Saha S, Desiraju GR (2018) Acid···amide supramolecular synthon in cocrystals: from spectroscopic detection to property engineering. J Am Chem Soc 140:6361–6373. https://doi.org/10.1021/jacs.8b02435

26. Plusquellic DF, Siegrist K, Heilweil EJ, Esenturk O (2007) Applications of terahertz spectroscopy in biosystems. ChemPhysChem 8:2412–2431. https://doi.org/10.1002/cphc.200700332

27. Heyden M, Bründermann E, Heugen U et al (2008) Long-range influence of carbohydrates on the solvation dynamics of water—answers from terahertz absorption measurements and molecular modeling simulations. J Am Chem Soc 130:5773–5779. https://doi.org/10.1021/ja0781083

28. Williams MRC, True AB, Izmaylov AF et al (2011) Terahertz spectroscopy of enantiopure and racemic polycrystalline valine. Phys Chem Chem Phys 13:11719–11730. https://doi.org/10.1039/c1cp20594c

29. Taday PF, Bradley IV, Arnone DD, Pepper M (2003) Using terahertz pulse spectroscopy to study the crystalline structure of a drug: a case study of the polymorphs of ranitidine hydrochloride. J Pharm Sci 92:831–838. https://doi.org/10.1002/jps.10358

30. Zhang F, Kambara O, Tominaga K et al (2014) Analysis of vibrational spectra of solid-state adenine and adenosine in the terahertz region. RSC Adv 4:269. https://doi.org/10.1039/c3ra44285c

31. Giraud G, Wynne K (2003) A comparison of the low-frequency vibrational spectra of liquids obtained through infrared and Raman spectroscopies. J Chem Phys 119:11753–11764. https://doi.org/10.1063/1.1623747

32. Guido R, Valle D, Venuti E, et al (2004) Intramolecular and low-frequency intermolecular vibrations of pentacene polymorphs as a function of temperature. J Phys Chem B 108:1822–1826. https://doi.org/10.1021/jp0354550

33. Nesbitt DJ, Suhm MA (2010) Chemical dynamics of large amplitude motion. Phys Chem Chem Phys 12:8151. https://doi.org/10.1039/c0cp90051f

34. Xue Z, Suhm MA (2009) Probing the stiffness of the simplest double hydrogen bond : the symmetric hydrogen bond modes of jet-cooled formic acid dimer. J Chem Phys 131:054301. https://doi.org/10.1063/1.3191728

35. Muller A, Talbot F, Leutwyler S (2000) Intermolecular vibrations of jet-cooled (2-pyridone)(2): a model for the uracil dimer. J Chem Phys 112:3717–3725

36. Müller A, Talbot F, Leutwyler S (2001) Intermolecular vibrations of the jet-cooled 2-pyridone·2-hydroxypyridine mixed dimer, a model for tautomeric nucleic acid base pairs. J Chem Phys 115:5192–5202. https://doi.org/10.1063/1.1394942

37. Müller A, Talbot F, Leutwyler S (2002) S1/S2 exciton splitting in the (2-pyridone)2 dimer. J Chem Phys 116:2836–2847. https://doi.org/10.1063/1.1434987

38. Day GM, Zeitler JA, Jones W et al (2006) Understanding the influence of polymorphism on phonon spectra : lattice dynamics calculations and terahertz spectroscopy of carbamazepine. J Phys Chem B 110:447– 456. https://doi.org/10.1021/jp055439y

39. Williams MRC, Aschaffenburg DJ, Ofori-Okai BK, Schmuttenmaer C a. (2013) Intermolecular vibrations in hydrophobic amino acid crystals: experiments and calculations. J Phys Chem B 117:10444–10461. https://doi.org/10.1021/jp406730a

40. Zhang F, Hayashi M, Wang HW et al (2014) Terahertz spectroscopy and solid-state density functional theory calculation of anthracene: effect of dispersion force on the vibrational modes. J Chem Phys 140. https://doi.org/10.1063/1.4873421

41. Zhang F, Wang H-W, Tominaga K, Hayashi M (2015) Intramolecular vibrations in low-frequency normal modes of amino acids: L-alanine in the neat solid state. J Phys Chem A 119:3008–3022. https://doi.org/10.1021/jp512164y

42. Zhang F, Wang HW, Tominaga K et al (2019) High-resolution THz spectroscopy and solid-state density functional theory calculations of polycyclic aromatic hydrocarbons. J Infrared Millimeter Terahertz Waves. https://doi.org/10.1007/s10762-019-00621-0

43. Saito S, Inerbaev TM, Mizuseki H et al (2006) Terahertz vibrational modes of crystalline salicylic acid by numerical model using periodic density functional theory. Jpn J Appl Phys 45:4170–4175. https://doi.org/10.1143/JJAP.45.4170

44. King MD, Korter TM (2011) Application of London-type dispersion corrections in solid-state density functional theory for predicting the temperature-dependence of crystal structures and terahertz spectra. Cryst Growth Des 11:2006–2010. https://doi.org/10.1021/cg200211x

45. King MD, Korter TM (2011) Noncovalent interactions between modified cytosine and guanine DNA base pair mimics investigated by terahertz spectroscopy and solid-state density functional theory. J Phys Chem A 115:14391–14396. https://doi.org/10.1021/jp208883t

46. Jepsen PU, Clark SJ (2007) Precise ab-initio prediction of terahertz vibrational modes in crystalline systems. Chem Phys Lett 442:275–280. https://doi.org/10.1016/j.cplett.2007.05.112

47. Houjou H (2009) Coarse graining of intermolecular vibrations by a Karhunen-Loeve transformation of atomic displacement vectors. J Chem Theor Comput 5:1814–1821

48. Houjou H (2011) Evaluation of coupling terms between intra- and intermolecular vibrations in coarse-grained normal-mode analysis: does a stronger acid make a stiffer hydrogen bond. J Chem Phys 135:154111. https://doi.org/10.1063/1.3652102

49. Houjou H (2017) Modelling intra- and intermolecular vibrations under the harmonic oscillator approximation: from symmetry-adapted to coarse-grained coordinate approaches. J Math Chem 55:532–551. https://doi.org/10.1007/s10910-016-0692-x

50. Isogai M, Houjou H (2018) Indices to evaluate the reliability of coarse-grained representations of mixed inter / intramolecular vibrations. J Mol Model 24:221. https://doi.org/10.1007/s00894-018-3757-x

51. Ochterski JW (1999) Vibrational analysis in Gaussian. White Papers and Technical Notes, see http://www.gaussian.com/g_whitepap/vib.htm. Accessed 31 Mar 2008

52. Wilson Jr EB, Decius JC, Cross PC (1980) Molecular vibrations. Dover, New York

53. Frisch MJ, Trucks GW, Schlegel HB et al (2013) Gaussian 09w, Revison D. Gaussian, Inc, Wallingford

54. Jmol: an open-source Java viewer for chemical structures in 3D, see http://www.jmol.org/. Accessed 24 April 2021

55. Note that the accurate description of $\Phi_{Tx,Tx}, \ldots$ is $\widetilde{\Phi}_{Tmx,Tmx}, \ldots$ (m = I or II), where $\widetilde{\Phi}_{TIx,TIx}$ and $\widetilde{\Phi}_{TIIx,TIIx}$ have exactly the same absolute value but with opposite signs. This is due to the constraint that the force constant $\widetilde{\Phi}_{TIx,TIx} + \widetilde{\Phi}_{TIIx,TIIx}$ for the pure translation should be zero. The force constant for the x-stretch mode vibration is calculated as $\widetilde{\Phi}_{TIx,TIx} - \widetilde{\Phi}_{TIIx,TIIx} = 2\widetilde{\Phi}_{TIx,TIx}$. Similarly, $2\widetilde{\Phi}_{TIy,TIy}$ and $2\widetilde{\Phi}_{TIz,TIz}$ give the force constants for y-shear and z-shear mode vibrations, respectively

56. Isogai M, Houjou H (2016) Quantification of inter/intramolecular stiffness by coarse-graining intermolecular vibrations of Homo/hetero-dimers. J Comput Chem Jpn 15:60–62. https://doi.org/10.2477/jccj.2016-0039

57. Legon AC, Millen DJ (1986) Gas-phase spectroscopy and the properties of hydrogen-bonded dimers: HCN•••HF as the spectroscopic prototype. Chem Rev 86:635–657

58. Houjou H, Koga R (2008) Explicit representation of anisotropic force constants for simulating intermolecular vibrations of multiply hydrogen-bonded systems. J Phys Chem A 112:11256–11262. https://doi.org/10.1021/jp8057614CCC