NMscatt: a program for calculating inelastic scattering from large biomolecular systems using classical force-field simulations ☆

Franci Merzel [a,*], Fabien Fontaine-Vive [b,c], Mark R. Johnson [b,*]

[a] *National Institute of Chemistry, Hajdrihova 19, 1000 Ljubljana, Slovenia*
[b] *Institute Laue Langevin, BP 156, 38042 Grenoble cedex 9, France*
[c] *Radiation, Reactors and Radionuclides Department, Faculty of Applied Sciences, Delft University of Technology, Mekelweg 15, 2629 JB Delft, The Netherlands*

Abstract

Computational tools for normal mode analysis, which are widely used in physics and materials science problems, are designed here in a single package called NMscatt (Normal Modes & scattering) that allows arbitrarily large systems to be handled. The package allows inelastic neutron and X-ray scattering observables to be calculated, allowing comparison with experimental data produced at large scale facilities. Various simplification schemes are presented for analyzing displacement vectors, which are otherwise too complicated to understand in very large systems.

Program summary

Title of program: NMscatt
Catalogue identifier: ADZA_v1_0
Program summary URL: http://cpc.cs.qub.ac.uk/summaries/ADZA_v1_0.html
Program obtainable from: CPC Program Library, Queen's University of Belfast, N. Ireland
Licensing provisions: no
No. of lines in distributed program, including test data, etc.: 573 535
No. of bytes in distributed program, including test data, etc.: 4 516 496
Distribution format: tar.gz
Programming language: FORTRAN 77
Computer: x86 PC
Operating system: GNU/Linux, UNIX
RAM: Depends on the system size to be simulated
Word size: 32 or 64 bits
Classification: 16.3
External routines: LAPACK
Nature of problem: Normal mode analysis, phonons calculation, derivation of incoherent and coherent inelastic scattering spectra.
Solution method: Full diagonalization (producing eigen-vectors and eigen-values) of dynamical matrix which is obtained from potential energy function derivation using finite difference method.
Running time: About 7 hours per one k-point evaluation in sampling all modes dispersion curves for a system containing 3550 atoms in the unit cell on AMD Athlon 64 X2 Dual Core Processor 4200+.

---

1. Introduction

At large scale facilities for neutron and X-ray scattering, large quantities of experimental data are produced. For complex, nanoscale systems, understanding this data requires computer models. In the case of inelastic scattering, molecular dynamics (MD) simulations [1] are widely used to equilibrate structures and explore dynamics as a function of temperature and other experimental parameters. However MD only gives a partial description of vibrational modes through the partial density of states and when knowledge about specific vibrational modes is required, normal mode analysis (NMA) [2] has to be performed. For physics and materials science problems, NMA gives a description of the lattice dynamics via the dispersion (k-vector dependence) of the mode frequencies [3]. For small systems (< 200 atoms in the simulation box) very accurate results can be obtained using density functional theory (DFT) methods to determine interatomic force constants [4,5]. By combining DFT and software that constructs and diagonalizes the dynamical matrix and calculates the experimental observables, experimentalists now have sophisticated tools to analyse their data. The PHONON [6] package is one of the best examples.

Phonon codes are traditionally limited to small systems for a number of reasons. For example, for small unit cells, the reciprocal lattice is big and stronger effects of dispersion are expected. If DFT methods are used to determine force constants then these methods are themselves restricted to a few hundred atoms. However the development of nanoscale structures of (partially) crystalline materials stimulates a need for phonon codes to be extended to much larger systems. Parameterized force fields [7] can be used to determine the inter-atomic force constants and, as will be seen in the example presented here, strong dispersion effects are observed in spite of the small reciprocal lattice.

In biomolecular systems, the need for NMA has long been recognized and codes like CHARMM [8] allow the gamma point normal modes to be calculated for moderately big systems. In addition, the neutron scattering quantities can be directly calculated from the simulations using the time-correlation function formalism [9], as implemented in the nMOLDYN program [10]. A combination of neutron scattering experiments and atomic detail computer simulations has proven to be a powerful technique for studying internal molecular vibrations [11–13]. In this approach one can validate the applied numerical models, i.e. force field parameterizations, depending on the agreement between the experimental and calculated spectra.

In this paper we present a software package that extends the functionality of codes like PHONON [6] and Climax [14] to arbitrarily large systems and extends the gamma point only analysis already available for larger systems to include k-vector dependence. The software reads a Hessian matrix of force con-

stants, constructs and diagonalizes the dynamical matrix for any k-vector and calculates neutron and X-ray scattering observables. The computational bottleneck remains the diagonalization of correspondingly large dynamical matrices and we comment on approximations that have to be used when the Brillouin zone cannot be sampled at a large number of points. In large systems, atomistic detail in the displacement vectors can be difficult to interpret due to the large number of degrees of freedom and we present two methods for simplifying this information. The first entails summing displacement vectors over atoms in user-defined beads, while the second involves a reduction of the degrees of freedom in the dynamical matrix by summing over force constants, which has the advantage of reducing the number of modes to be examined.

2. Theoretical background

The standard approach, also called a direct method [15], to the lattice vibration problem of crystals is based on the explicit knowledge of the interaction between all atom-pairs in the system. Subsequently, one deduces the corresponding force constants, and constructs and diagonalizes the dynamical matrix for any k-vector in order to obtain the frequencies of the normal modes. A reasonable atomic detail description of interactions within large biomolecular systems are provided using empirical force fields.

In the following we will briefly summarize the aspects of the classical theory of lattice vibrations [16] and proceed to the description of the explicit phonon calculations.

The individual atomic positions in the crystal can be assigned as

$$\vec{R}_{n\mu}(t) = \vec{R}_n + \vec{r}_\mu + \vec{u}_{n\mu}(t), \tag{1}$$

where $\vec{R}_n$ is unit cell lattice vector and $\vec{u}_{n\mu}(t)$ is displacement of atom μ from its equilibrium positions $\vec{r}_\mu$. Within the harmonic approximation we concentrate on expansion of the small differences of potential energy V due to the small changes in atom positions:

$$V(\mathbf{u}) \approx V_0 + \sum_{n\mu\alpha} \frac{\partial V}{\partial u_{n\mu\alpha}} u_{n\mu\alpha}$$
$$+ \frac{1}{2} \sum_{n\mu\alpha, m\nu\beta} u_{n\mu\alpha} \mathbf{D}_{n\mu\alpha, m\nu\beta} u_{m\nu\beta} + \cdots, \tag{2}$$

where the second derivative defines the force constant between the atoms μ and ν:

$$\mathbf{D}_{n\mu\alpha, m\nu\beta} = \frac{\partial^2 V}{\partial u_{n\mu\alpha} \partial u_{m\nu\beta}} \quad (\alpha, \beta = x, y, z). \tag{3}$$

As each unit cell is identical to every other unit cell in the crystal, the displacement pattern of a normal mode has to be identical to that in any other cell to within a phase difference $\vec{k}(\vec{R}_n - \vec{R}_m)$. The representation of the atom displacement is

Fig. 1. A 2D periodic system representing the effect of the minimum image convention while constructing dynamical matrix. If a distance between atom ν (striped circle) and μ (full black circle) is larger than the cut-off distance (dashed circle around the atom μ), and if the force constant $\mathbf{H}_{\mu\nu}$ turns out to be nonzero, it means that the value of $\mathbf{H}_{\mu\nu}$ has to be related to the "image" atom ν in the image cell m (thatched circle), with the position vector $\vec{r}_\nu + \vec{l}_x$ ($\vec{l}_x = \vec{R}_m - \vec{R}_n$), instead of to the atom $\vec{r}_\nu$ in the primary cell n. The corresponding lattice translation vector $\vec{l}_x$ is thus determined for the element $\mathbf{H}_{\mu\nu}$.

chosen to be a plain wave ansatz of the form:

$$\vec{u}_{n\mu}(\vec{k}, t) = \frac{u_0}{\sqrt{M_\mu}} \vec{e}_{\mu\vec{k}} \exp\bigl(i[\vec{k}\vec{R}_n - \omega_{\vec{k}}t]\bigr), \tag{4}$$

where $\vec{e}_{\mu\vec{k}}$ is the polarization vector and M_μ is the mass of the atom μ. We omit writing Cartesian component subscripts. Solving the equation of motion with ansatz (4) is equivalent to the eigen-value problem

$$\omega_{\vec{k}j}^2 \vec{e}_{\mu\vec{k}j} = \sum_\nu \mathcal{D}_{\mu\nu}(\vec{k})\vec{e}_{\nu\vec{k}j}, \tag{5}$$

where

$$\mathcal{D}_{\mu\nu}(\vec{k}) = \sum_m \frac{1}{\sqrt{M_\mu M_\nu}} \mathbf{D}_{n\mu,m\nu} \exp\bigl[i\vec{k}(\vec{R}_m - \vec{R}_n)\bigr] \tag{6}$$

is the so-called dynamical matrix.

The form of the dynamical matrix (6) requires the atom pairs for which one atom belongs to a different unit cell, i.e. $(m \neq n)$, to be identified. These terms contribute to the so-called Bloch-factor $\exp[i\vec{k}(\vec{R}_m - \vec{R}_n)]$ and make the dynamical matrix complex. But in case of applying periodic boundary conditions (PBC) as implemented in computer simulation programs, the potential energy of a crystal is given as an explicit function of only the atom positions in the primary unit cell. As a consequence, we obtain the second derivative matrix $\mathbf{H}$ in which the contributions from the inter-cell atomic pairs are mapped and added to the corresponding image atom pairs in the primary unit cell:

$$\left[\sum_m \mathbf{D}_{n\mu,m\nu}\right] \rightarrow \mathbf{H}_{\mu,\nu} = \frac{\partial^2 V^{\mathrm{PBC}}}{\partial u_\mu \partial u_\nu}. \tag{7}$$

One can directly obtain $\mathbf{D}_{n\mu\alpha,m\nu\beta}$ by increasing the size of the unit cell by one or more layers of periodically arranged image cells and calculate force constants in the extended super-cell. However, this approach is unfavorable when dealing with very large systems.

A similar approach is to decompose the potential energy V^{PBC} in Eq. (7) into individual contributions from the image cells, $V^{\mathrm{PBC}} = V_0 + \sum_m V_m$ and evaluate the second derivative matrix for each term upon the same minimized structure.

The situation is simpler if the interaction is truncated at some cutoff distance r_{cut} so that the "minimum image convention" (MIC) is obeyed. The MIC states that each atom interacts at most with one image of every other atom in the system (which is repeated to fully enclose the primary unit cell with the periodic boundary conditions). This has the effect of limiting the interaction cutoff, for example, to no more than half the length of the minimum side when simulating the orthorhombic cell, $r_{\mathrm{cut}} < \min\{a/2, b/2, c/2\}$. It should be noted that the size of nanoscale crystals usually far exceeds the spatial range of forces between atoms ($\sim <12$ Å) allowing physically reasonable cut-off radii to be introduced.

According to the MIC we can conclude that there is always only one translation $\vec{t}$ per atom pair (ν, μ) giving rise to the minimum distance $|(\vec{r}_\nu + \vec{t}) - \vec{r}_\mu|$:

$$\exists s_\alpha = \{-1, 0, 1\}_{\alpha=1,2,3}: \vec{t} = s_1\vec{l}_1 + s_2\vec{l}_2 + s_3\vec{l}_3;$$
$$\text{so that:} \quad \bigl|(\vec{r}_\nu + \vec{t}) - \vec{r}_\mu\bigr| = \min, \tag{8}$$

where $\vec{l}_\alpha$ are lattice translation vectors, cf. Fig. 1. In the case of a nonvanishing $\mathbf{H}_{\mu,\nu}$ we get the following expression for a dynamical matrix element

$$\mathcal{D}_{\mu\nu}(\vec{k}) = \frac{1}{\sqrt{M_\mu M_\nu}} \begin{cases} \mathbf{H}_{\mu,\nu}, & \text{if } |\vec{r}_\mu - \vec{r}_\nu| < r_{\mathrm{cut}}, \\ \mathbf{H}_{\mu,\nu} \exp(i\vec{k}\vec{t}), & \text{if } |\vec{r}_\mu - \vec{r}_\nu| > r_{\mathrm{cut}}. \end{cases} \tag{9}$$

According to Eq. (5) the diagonalization of matrix $\mathcal{D}(\vec{k})$ yields the phonon frequencies $\omega_{\vec{k}j}$ and corresponding polarization vectors $\vec{e}_{\mu\vec{k}j}$ for a given phonon wave vector $\vec{k}$. A complete solution leads to the phonon dispersion relations. The subscript j denotes a branch in the phonon dispersion. In a crystal of N atoms, there are $3N$ branches.

2.1. Inelastic scattering

The dynamic structure factor $S(\vec{q}, \omega)$ contains information about the structure and dynamics of the sample. It can be split into a coherent part arising from the cross-correlations of atomic motions and an incoherent part describing self-correlations of single atom motions. According to the standard theory [17,18], which is based on the harmonic approximation, we obtain the following expressions for the coherent and incoherent dynamical structure factors:

$$
\begin{aligned}
S(\vec{q}, \omega)_{\text{coh}} = \sum_{\vec{G}} \sum_{\vec{k}, j} \frac{\hbar}{2\omega_{\vec{k}j}} \\
\times \left| \sum_{\mu} \sigma_{\mu}^{\text{coh}} \frac{\vec{q} \cdot \vec{e}_{\mu\vec{k}j}}{\sqrt{M_{\mu}}} \exp\left(-W_{\mu}(\vec{q}) + i\vec{q}\vec{r}_{\mu}\right) \right|^2 \\
\times \left(n(\omega_{\vec{k}j}) + 1\right)\delta(\omega - \omega_{\vec{k}j})\delta(\vec{q} + \vec{k} - \vec{G}),
\end{aligned} \quad (10)
$$

and

$$
\begin{aligned}
S(\vec{q}, \omega)_{\text{inc}} = \sum_{\mu} \sigma_{\mu}^{\text{inc}} \sum_{\vec{k}, j} \frac{\hbar}{2M_{\mu}\omega_{\vec{k}j}} \left|\vec{q} \cdot \vec{e}_{\mu\vec{k}j}\right|^2 \left(n(\omega_{\vec{k}j}) + 1\right) \\
\times \exp\left[-2W_{\mu}(\vec{q})\right]\delta(\omega - \omega_{\vec{k}j}),
\end{aligned} \quad (11)
$$

where $\vec{q}$ is the scattering vector, σ_{μ} is the corresponding atomic scattering length, $(n(\omega) + 1)$ refers to the phonon creation process (absorption spectrum) and $n(\omega)$ is the mean number of phonons of frequency ω at temperature T according to the Bose–Einstein statistics

$$
\left(n(\omega) + 1\right) = \frac{\exp(\hbar\omega/k_B T)}{\exp(\hbar\omega/k_B T) - 1} \quad (12)
$$

The factor $\exp[-2W_{\mu}(\vec{q})]$ is called the Debye–Waller factor:

$$
W_{\mu}(\vec{q}) = \frac{\vec{q} \cdot \mathbf{B}(\mu) \cdot \vec{q}}{2}, \qquad B_{\alpha\beta}(\mu) = \langle u_{\mu\alpha} u_{\mu\beta}\rangle_T, \quad (13)
$$

where $\mathbf{B}(\mu)$ is a 3×3 symmetric tensor representing the thermodynamic mean square displacement of an atom μ, which can be expressed by the partial atomic phonon density of states $g_{\alpha\beta,\mu}(\omega)$:

$$
B_{\alpha\beta}(\mu) = \frac{3N\hbar}{2M_{\mu}} \int_0^{\infty} \frac{d\omega}{\omega} g_{\alpha\beta,\mu}(\omega) \coth\left(\frac{\hbar\omega}{2k_B T}\right). \quad (14)
$$

The partial atomic density of states is a weighted distribution of normal modes

$$
g_{\alpha\beta,\mu}(\omega) = \frac{1}{3Nn} \sum_{\vec{k}} \sum_{j=1}^{3N} e_{\alpha\mu\vec{k}j} e_{\beta\mu\vec{k}j}^* \delta(\omega - \omega_{\vec{k}j}), \quad (15)
$$

where n is the number of sampling $\vec{k}$-points in the first Brillouin zone.

The evaluation of the Debye–Waller B factors using Eq. (14) requires extra attention due to the "zero-phonon" term resulting from the singularity caused by phonon (acoustic) branches where $\omega(\vec{q}, j) = 0$. We separate the contributions to $B_{\alpha,\beta}(\mu)$ due to optic and acoustic modes as $B_{\alpha,\beta}(\mu) = B'_{\alpha,\beta}(\mu) +$

$\widetilde{B}_{\alpha,\beta}(\mu)$, where $B'_{\alpha,\beta}(\mu)$ is defined as in Eq. (14), only that the acoustic modes ($j = 1, 3$) are excluded from calculation of the partial atomic density of states when using Eq. (15). The contribution of the 3 acoustic modes described by $\widetilde{B}_{\alpha,\beta}(\mu)$ is treated separately by using the Debye approximation. This means that density of states due to acoustic modes is approximated as $g_{ac}(\omega) \propto \omega^2$, providing the normalization $\int_0^{\omega_{\text{max}}} g_{ac}(\omega)\, d\omega = 3$, where ω_{max} is the maximum frequency up to which the acoustic dispersion curve is linear. We obtain the following expression for B factor due to acoustic modes:

$$
\begin{aligned}
\widetilde{B}_{\alpha\beta}(\mu) &= \sum_{j=1}^{3} \frac{3\hbar}{M_{\mu}\omega_{j\,\text{max}}^3} \\
&\quad \times \int_0^{\omega_{j\,\text{max}}} \omega e_{\alpha\mu\vec{k}j} e_{\beta\mu\vec{k}j}^* \left(\frac{1}{\exp(\hbar\omega/k_B T) - 1} + \frac{1}{2}\right) d\omega \\
&= \sum_{j=1}^{3} \frac{3\hbar e_{\alpha\mu\vec{k}j} e_{\beta\mu\vec{k}j}^*}{M_{\mu}\omega_{j\,\text{max}}} \\
&\quad \times \left[\left(\frac{k_B T}{\hbar\omega_{j\,\text{max}}}\right)^2 \int_0^{x_{j\,\text{max}}} \frac{x}{\exp(x) - 1}\, dx + \frac{1}{4}\right]. \quad (16)
\end{aligned}
$$

A new variable $x_{j\,\text{max}} = \hbar\omega_{j\,\text{max}}/k_B T$ was introduced in the last equation.

3. Analyzing the displacement vectors

For systems containing thousands of atoms (N) the displacement vectors obtained by diagonalizing the dynamical matrix can be difficult to understand, especially for low frequency modes which involve the displacement of many (or all) atoms. One simple solution to this problem is to sum over the displacements of atoms within beads, which represent logical coarse grains of the system, for example base molecules in the DNA example below. This treatment allows different bead definitions to be applied to the calculated displacement vectors but has the disadvantage of not reducing the number of displacement vectors from $3N$.

A related approach could be used to reduce the atomic level Hessian matrix to lower dimension by mapping the inter-atomic force constants on to inter-bead force constants [19]. For N' beads, the resulting dynamical matrices have dimension $3N'$ and therefore result in $3N'$ displacement vectors for any k-vector. Apart from losing the information about internal vibrations within beads (which are usually not of interest when low-frequency motions are to be investigated), such simplification also causes a loss in information about the rotational degrees of freedom of the beads (rigid bodies), which leads to erroneous dispersion properties of low-frequency modes.

4. Example

To verify the implemented formalism we have simulated a B-form DNA molecule (right-handed, 10 base-pairs per turn, pitch 33.6 Å) using CHARMM [8]. The model was constructed
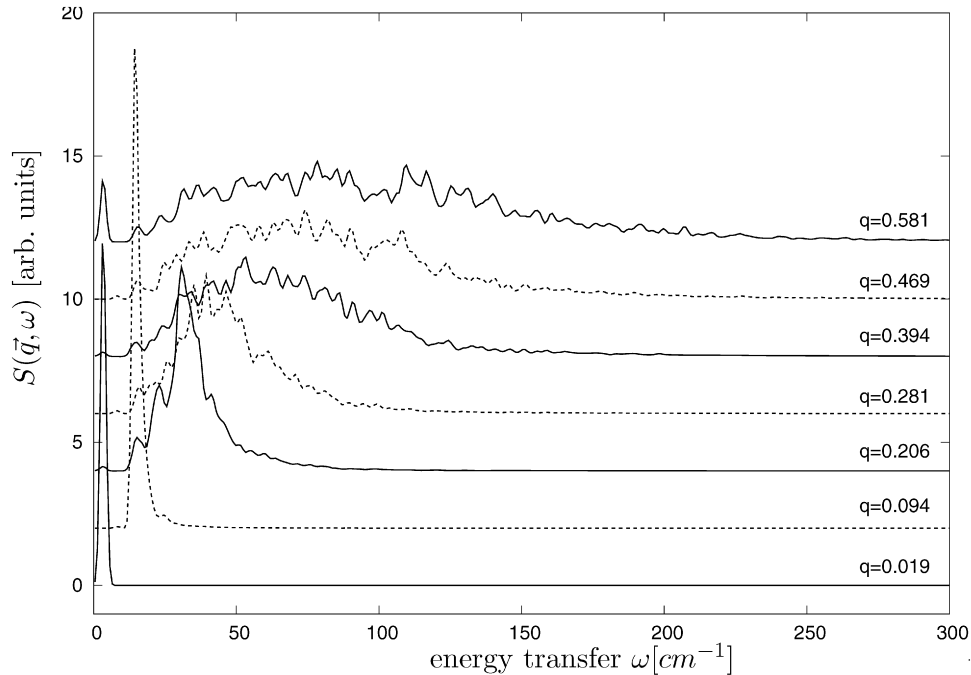
Fig. 2. Coherent dynamical structure factor of B-DNA. Scattering vector $\vec{q}$ is varied along the helical axis $(0, 0, 1)$, $|\vec{q}| = 0.019 \ldots 0.581$ Å^{-1}, leading to a shift of $S(\vec{q}, \omega)$ intensity.

manually choosing an arbitrary base-pair sequence and following the structural rules specific for the B-form DNA in the presence of Li$^+$ ions [20]. The full crystal environment was generated using periodic boundary conditions for an orthorhombic unit cell containing one helix of DNA. The dimension of the unit cell is $a = 32.2$ Å, $b = 31.8$ Å and $c = 33.5$ Å, with c parallel to the helical axis. The starting configuration was obtained by minimization of the potential energy of the crystal structure obtained as the time average over a 1 ns MD simulation at 100 K. The Hessian matrix of force constants was generated by displacing each atom in turn from equilibrium and calculating the forces induced on all other atoms. Diagonalization of the resulting dynamical matrices was performed using the routine zheev from the LAPACK library [21].

As we use Debye approximation for evaluation of the acoustic-modes contribution to the Debye–Waller factors, we show first how sensitive the results are to this approximation. According to Eq. (16), the range of approximation is determined by the normalized frequencies $x_{j\,\text{max}}$ of each of the three acoustic modes corresponding to the cutoff momentum k_D which defines the linear regime of the acoustic dispersion curves. By varying the value of k_D we can observe the effect of the approximation. Results for isotropic atomic B-factors $\bar{B}(\mu) = 1/3(B_{11}(\mu) + B_{22}(\mu) + B_{33}(\mu))$ are given for two selected atoms, one heavy (phosphorus) and one light (hydrogen) atom, in Table 1 and clearly demonstrate modest variation in the values of the B-factors proving the application of the approximation to be reasonable.

The typical coherent spectrum of DNA, Fig. 2, obtained from Eq. (10), shows a low frequency spectrum in which the centre of gravity moves along the frequency dimension upon varying momentum transfer $\vec{q}$. Fitting the spectral profile with

Table 1
Isotropic B-factors for two selected atoms, phosphorous and hydrogen as a function of κ defining the cutoff momentum $k_D = \kappa(2\pi/c)$, which determines the range of validity of the Debye approximation

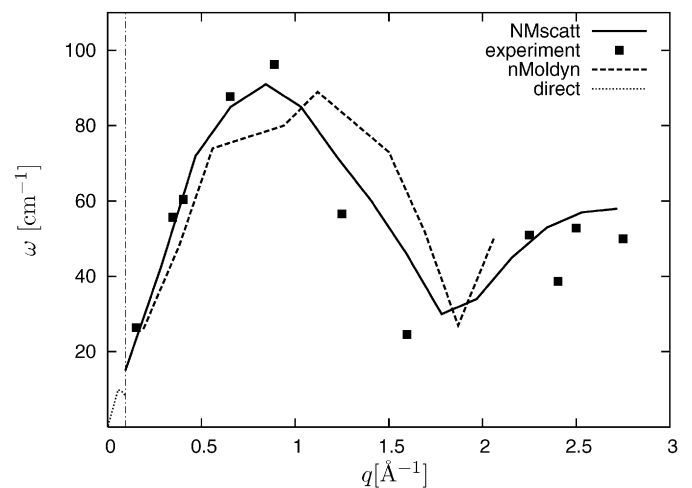| κ	$\bar{B}(P)$ [Å^2]	$\bar{B}(H)$ [Å^2]
0.15	0.00256	0.01702
0.20	0.00246	0.01691
0.25	0.00239	0.01684
0.30	0.00236	0.01681

Fig. 3. Dispersion curves obtained by Gaussian fitting of the moving Brillouin peak shown in $S(\vec{q}, \omega)$ profiles on Fig. 2 for inelastic X-ray scattering using NMscatt (solid curve), and nMoldyn (dashed curve). Measured values from IXS experiment [22] are shown by black squares.

a Gaussian as a function of wave-vector gives the dispersion curve shown in Fig. 3, which compares well with the recent

(a) B-DNA: acoustic mode
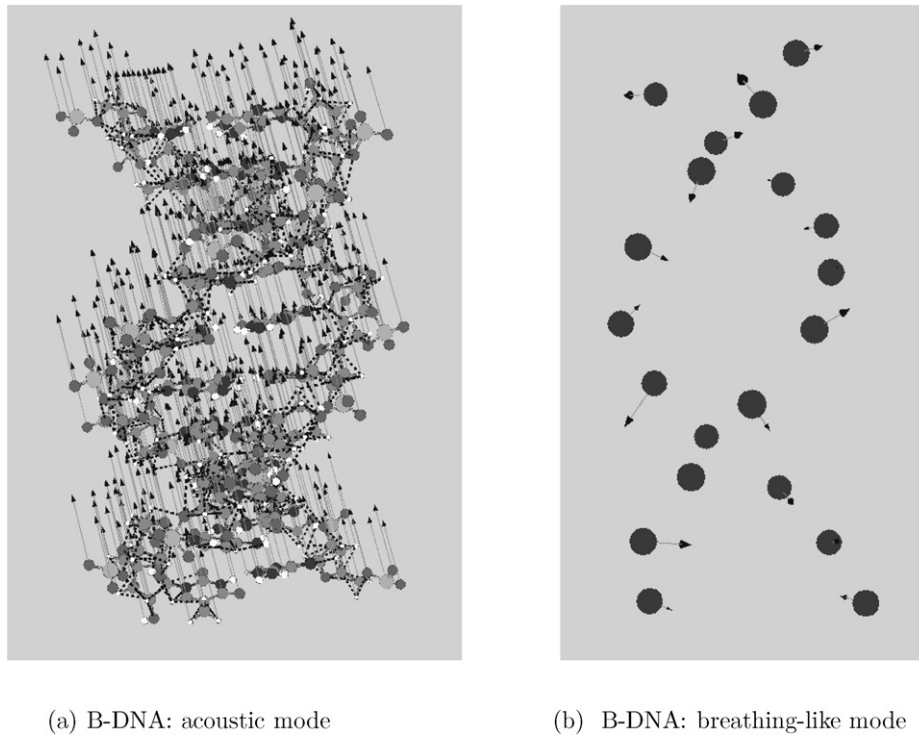

(b) B-DNA: breathing-like mode

Fig. 4. Full-atom representation (a), and bead representation of each nucleotide (b) of the selected vibrational mode.

experimental results [22]. Individual dispersion curves are not shown since there are 1000 in the first 100 cm^{-1} and these have not been observed experimentally. The acoustic phonons reach a maximum of 10 cm^{-1} at the Brillouin zone boundary of the helix 0.09 Å^{-1} and give rise to the sharp peak at 0.019 and 0.581 Å^{-1} with spectral intensity following $S(q)$ (see [22]). At frequencies above 10 cm^{-1}, the q-dependence of weakly dispersed optic modes gives the apparent dispersion behavior seen in Figs. 2 and 3. The minimum in the apparent dispersion curve at 1.8 Å^{-1} corresponds to the extended Brillouin zone of the base-pair separation. Fig. 3 also shows the result of the equivalent analysis of a 300K MD simulation on the same model of B-DNA using nMoldyn, which calculates $S(q, \omega)$ via the van Hove correlation functions. The similarity of the two calculated curves validates the new phonon calculations. The differences may contain important information about the structural fluctuations of DNA at 300K and will be discussed elsewhere.

In order to gain insight into the nature of low frequency dynamics we can analyse the displacement vectors at an atomic level (see Fig. 4(a) for an acoustic mode). By summing over the displacement vectors in terms of beads, where base molecules, sugar molecules and phosphate groups are treated as single units, a simplified picture of the normal modes is obtained. Fig. 4(b) shows a high frequency mode which has a pronounced contribution from the phosphate groups.

5. Conclusion

The new, user-friendly computational package NMscatt presented here enable an efficient atomic detail analysis of different types of inelastic scattering applied to arbitrarily large nanoscale systems. The ability to perform molecular dynamics and phonon calculations on large nano- and bio-molecular materials means that one can efficiently pursue the investigation of some poorly understood structural and dynamical features of these systems.

Appendix A. Program package description

There are four main modules in the NMscatt program package *phonon, coh, incoh* and *bead*, and the overall NMscatt structure is given in Fig. 5. Below are described the corresponding modules.

- *phonon* (running script: phonon *n f*1 *f*2 *f*3 > out) Providing the full Hessian matrix for a given energy-minimized atomic structure within the specified crystallographic unit cell this module constructs dynamical matrix and calculates its eigenvalues and eigenvectors at given wave vector $\vec{k} = f1 * r(\vec{a}) + f2 * r(\vec{b}) + f3 * r(\vec{c})$, where $r(\vec{a})$ means the reciprocal vector of the Bravais lattice vector $\vec{a}$. At input this module requires to specify the Bravais
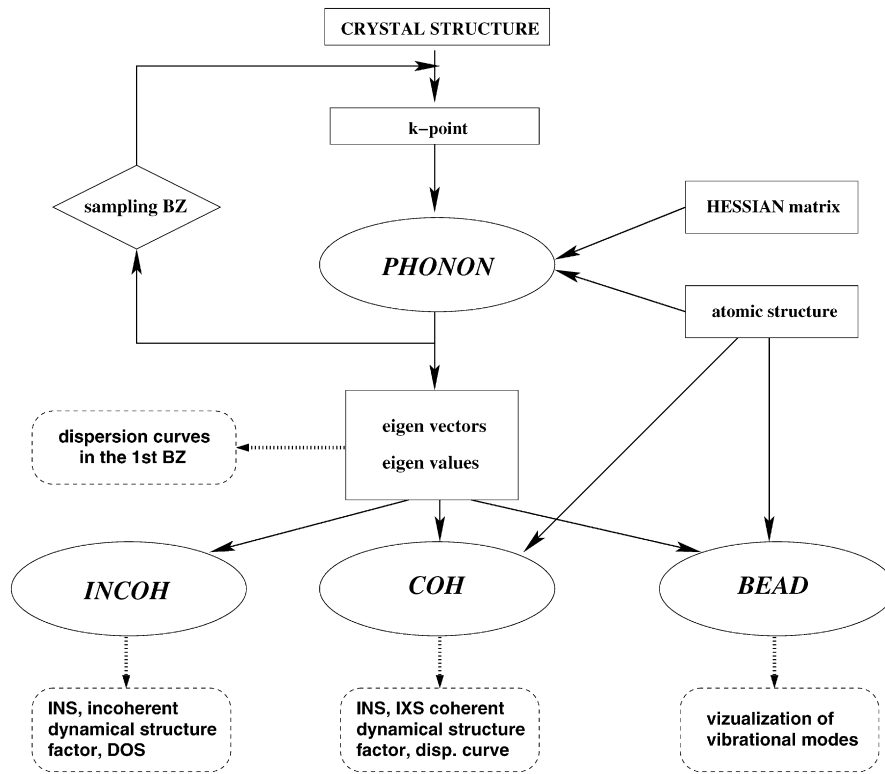
Fig. 5. Flow-chart of the `NMscatt` program package.

lattice vectors that were beforehand used in the molecular mechanics/dynamics simulation package to satisfy the periodic boundary conditions while generating the minimized structure. The cut-off radius must be given at which the long range interactions are truncated while calculating the Hessian matrix in the simulation. At output two separate files `eig_val_n` and `eig_vec_n` are written containing eigenvalues and eigenvectors, respectively, where n assigns a consecutive number of the sampling k-point in the first Brillouin zone. These files serve as an input for other modules of `NMscatt`. The choice $n = 0$ is assumed to be reserved for the Γ-point, $\vec{k} = (0, 0, 0)$. The lowest n should sample the vicinity of the Γ-point. The k-points are to be specified in the fractional coordinates $(f1, f2, f3)$ of the reciprocal lattice vectors.

The Hessian matrix `hessian` should be provided either in the binary format or (default) ASCII format (no special format required). Only the upper triangle of the matrix has to be given. First line of the Hessian matrix file contains a dimension of the matrix, which is equal to $3N_{at}$, where N_{at} is the number of atoms. Each subsequent line contains two integers and real number, defining the matrix element. Force constants in the Hessian matrix must be given in kcal/mol/Å^2 units. The atomic coordinates should be provided in the standard CHARMM coordinate format.

- *incoh* (running script: `incoh` n_{max} nqp n_a `> out`)
  This module allows to calculate atomic Debye–Waller factors representing the mean square displacements and dynamical structure factor of incoherent one-phonon neutron scattering on mono-crystals and from orientationally av-

eraged powder. In the latter case we need to specify an absolute value of momentum transfer q, range of the k-point index n: 0–n_{max} for picking up the corresponding k-point eigenvalues- and eigenvectors-files generated by *phonon*, number of random orientations nqp of q vector to provide spherical averaging and absolute temperature. Also, one has to define the k-point index n_a, for which the k-point range 0–n_a corresponds to the linear regime of the acoustic-mode-dispersion curves. This is needed for the proper derivation of Debye–Waller factors using Debye approximation. As a result, the function $S(q, \omega)$ is given in the file `Sqw_inch`. The convoluted spectra with the resolution function are given in the file `Sqw_inch_smth`. Optionally, one can also obtain density of states (DOS) in the file `DOS.dat`. The wavenumber cm^{-1} units are used for energy transfer and Å^{-1} units for momentum transfer.

- *coh* (running script: `coh` n_{max} n_c st n_a `> out`)
  This module calculates dynamical structure factor of coherent one-phonon neutron or X-ray scattering on monocrystals. In addition to the input data required by *incoh* (except for the spherical averaging), we need to define for module *coh* also the type of scattering: neutron/X-ray $st = n/x$, the number of higher Brillouin zones included for sampling momentum transfer vector, and the range of the k-point index 0–n_c evaluated by module *phonon* and assigning the k-points in the first Brillouin zone, which lie along the selected direction of the momentum transfer vector $\vec{q}$. As a result, the function $S(q, \omega)$ is given in the file `Sqw_coh_m`, where m stays for the index of the

Brillouin zone. The wavenumber cm^{-1} units are used for energy transfer and $Å^{-1}$ units for momentum transfer.

- *bead* (running script: `bead` m > out)
  This module is used to enable visualization of 5 selected consecutive vibrational modes starting with m obtained by running *phonon* for the Γ-point, such that atomic displacement vectors are projected on to the beads, which are defined as the center of mass of larger atomic groups of the system, for example, residues. The output files: `atm.xyz` (all atom structure) and `bead.xyz` (bead model structure), are readable by the program `xmakemol` [23] which enables direct visualization of the vibrational modes by plotting displacement vectors.

A.1. Compiling

The Makefile is provided to compile the program package. It is important to note that 64-bits processors are prerequisite for applying the `NMscatt` to analysis of large systems (containing more than 2000 atoms). One needs to install the LA-PACK library on the computer beforehand, prior to `NMscatt`, because of the internal call of the `zheev` routine used for diagonalization of dynamical matrix in the `phonon` module. The Fortran compiling switches `g77 -mcmodel=medium -funroll-all-loops -O3` are recommended when installed on 64-bits processors running Linux.

A.2. Benchmark results

Benchmark results presented above were obtained on AMD Athlon 64 X2 Dual Core Processor 2.2 GHz running Linux for B-form DNA simulated with the CHARMM program.

Program module specific parameters used in the source code are given bellow.

phonon

```
mxat
      maximum number of atoms
      (physical dimension of matrices)
ch
      suffix for the eigen-value/vector
      output files
fct(1),fct(2),fct(3)
      fractional coordinates of k-point
      in the 1st Brillouin zone
cutoff
      cutoff radius
brva(1),brva(2),brva(3)
      Bravais lattice vector a
brvb(1),brvb(2),brvb(3)
      Bravais lattice vector b
brvc(1),brvc(2),brvc(3)
      Bravais lattice vector c
lf_binr
      logical flag for the format
      of the hessian matrix
```

```
      .true. : binary,.false. : ASCII
name
      name of the hessian matrix file
name_crd
      name of the atom-coordinates file
      (standard CHARMM cor file)
deut
      logical flag for deuteration,
      .true.: deuterated
```

incoh

```
mxat
      maximum number of atoms
      (physical dimension of matrices)
nfr
      max k-point suffix n
      according to files
      eig_val/vec_n, n = 0..nfr 0:
      Gamma-point
nqp
      number of q vector orientations
      for powder averaging
ixf
      n corresponding to k-point
      defining the linear regime
      (0-n) of acoustic branch
lf_dos
      logical flag for calculation
      of DOS density-of-states
frqmx1
      freq. range for S(q,w): 0-frqmx1
temp0
      temperature [K]
deut
      logical flag for deuteration,
      .true.: deuterated
name_crd
      name of the atom-coordinates file
      (standard CHARMM cor file)
nip
      Gaussian integration points

c instrument geometry
enf
      fixed neutron energy of the
      scattered beam in [meV]
alp
      fixed angle between incident and
      scattered beam
```

coh

```
mxat
      maximum number of atoms
      (physical dimension of matrices)
```

```
nfr
      max k-point suffix n
      according to files
      eig_val/vec_n, n = 0..nfr 0:
      Gamma-point
      for evaluation of DW factors
nkp
      k-point suffix indicating q sampling
      range n=0..nkp, which fixes
      a q-direction;
      each k_n must lie along the chosen
      q direction
sct
      scattering type (x-ray/neutrons)
ixf
      n corresponding to k-point defining
      the linear regime (0-n)
      of acoustic branch
brva(1),brva(2),brva(3)
      Bravais lattice vector a
brvb(1),brvb(2),brvb(3)
      Bravais lattice vector b
brvc(1),brvc(2),brvc(3)
      Bravais lattice vector c
frqmx1
      freq. range for S(q,w): 0-frqmx1
temp0
      temperature [K]
deut
      logical flag for deuteration,
      .true.: deuterated
ggg0(1),ggg0(2),ggg0(3)
      reciprocal lattice vector G
      assigning a direction for sampling
      the momentum transfer vector q
nbz1, nbz2
      nbz1..nbz2: Brillouin zones defining
      the sampling interval
      of the momentum transfer vector q
name_crd
      name of the atom-coordinates file
      (standard CHARMM cor file)
nip
      Gaussian integration points
```

bead

```
mxat
      maximum number of atoms
      (physical dimension of matrices)
ism
      5 sequential normal modes
      to display:
      starting with ism:
```

```
      (ism, ism+1, ism+2...)
ch
      suffix of the eigen-vector
      file used for
      visualization of displacements
      (Gamma-point)
nat1
      number of protein atoms
nres
      number of residues in the
      protein part
scl0
      scaling factor for displacement
      vector
name_crd
      name of the atom-coordinates file
      (standard CHARMM cor file)
```

References

[1] A.R. Leach, Molecular Modelling: Principles and Applications, Pearson Education, Harlow, 2001.
[2] B.R. Brooks, D. Janežič, M. Karplus, J. Comp. Chem. 16 (1995) 1522.
[3] W. Frank, C. Elsässer, M. Fähnle, Phys. Rev. Lett. 74 (1995) 1791.
[4] K. Parlinski, G. Chapuis, J. Chem. Phys. 110 (1999) 6406.
[5] M.R. Johnson, K. Parlinski, I. Natkaniec, B.S. Hudson, Chem. Phys. 291 (2003) 53.
[6] K. Parlinski, PHONON Manual, Version 4.22, http://wolf.ifj.edu.pl/phonon/.
[7] A.D.J. MacKerrel, D. Bashford, M. Bellott, et al., J. Phys. Chem. B 102 (1998) 3586.
[8] B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, M. Karplus, J. Comp. Chem. 4 (1983) 187.
[9] L. van Hove, Phys. Rev. 95 (1954) 249.
[10] G.R. Kneller, V. Keiner, M. Kneller, M. Schiller, Comput. Phys. Comm. 91 (1995) 191.
[11] J.C. Smith, Q. Rev. Biophys. 24 (1991) 227.
[12] G.R. Kneller, W. Doster, M. Settles, S. Cusack, J.C. Smith, J. Chem. Phys. 97 (1992) 8864.
[13] A.M. Micu, D. Durand, M. Quilichini, M.J. Field, J.C. Smith, J. Phys. Chem. 99 (1995) 5645.
[14] G.J. Kearley, Nucl. Inst. Meth. Phys. Res. A 354 (1995) 53.
[15] K. Parlinski, Z.Q. Li, Y. Kawazoe, Phys. Rev. Lett. 78 (1997) 4063.
[16] B. Donovan, J.F. Angress, Lattice Vibrations, Chapman and Hall Ltd., London, 1971.
[17] M. Bee, Quasielastic Neutron Scattering: Principles and Applications in Solid State Chemistry, Biology and Materials Science, Hilger, Bristol, 1988.
[18] S. Lovesey, Theory of Neutron Scattering from Condensed Matter, International Series of Monographs on Physics, vol. 72, Oxford Science, Oxford, 1984.
[19] G. Venkataraman, V.C. Sahni, Rev. Mod. Phys. 42 (1970) 409.
[20] C.R. Cantor, P.R. Schimmel, Biophysical Chemistry, Part I, The Conformation of Biological Macromolecules, W.H. Freeman and Company, San Francisco, 1980.
[21] http://www.netlib.org/lapack/.
[22] Y. Liu, S.H. Chen, D. Berti, P. Baglioni, A. Alatas, H. Sinn, E. Alp, A. Said, J. Chem. Phys. 123 (2005) 214909.
[23] http://www.nongnu.org/xmakemol/.