Check for updates

Probing the environmental sensitivity of thymine C=O vibrations through infrared spectra simulations

Giovanni Parolin, [a] Cedrix J. Dongmo Foumthuim, [a] Stefano Corni[ab] and Laura Zanetti-Polzi [*b]

Infrared spectroscopy is widely used to probe the structural organization of biologically relevant molecules, including peptides, proteins, and nucleic acids. The latter show significant structural diversity, and specific infrared bands provide insights into their conformational ensembles. Among DNA/RNA infrared bands, the C=O stretching modes are especially useful, as they are sensitive to the distinct structural arrangements within nucleic acids. However, the relationship between different spectral lineshapes and specific structural features is often non-trivial, especially in highly flexible systems such as single-stranded DNA. In this work, we propose a hybrid quantum-classical computational approach based on the perturbed matrix method for calculating infrared bands in nucleic acids. This approach, previously applied to calculate C=O stretching modes in peptides and proteins, is applied here for the first time to DNA. Specifically, using molecular dynamics simulations combined with density functional theory (B3LYP) calculations, we calculate the spectrum arising from the two C=O stretching modes of the thymine base, both in water and deuterated water, with a specific focus on the sensitivity of the spectral lineshape to the systems' conformational ensemble. We compute the spectra for 1-methylthymine, thymidine 5'-monophosphate, and a single-stranded oligomer composed of ten thymine bases, and critically compare them to their corresponding experimental signals. Our results indicate that the difference in the relative intensity of the two bands experimentally observed between the spectrum of a single solvated base and that of the oligomer, also captured in the calculated spectra, results from stacking and hydrogen bonding interactions among the bases.

1 Introduction

Structural polymorphism is one of the distinguishing features of DNA that can be found in a wide variety of secondary structures depending on the sequence and environmental conditions. In addition to well-known DNA duplexes (which include B-, A-, and Z-DNA), other non-canonical structures can be formed through non-Watson–Crick hydrogen bonds between nucleobases, such as triplexes, G-quadruplexes, intercalated motives or Holliday junctions.[1,2] All these arrangements were shown to be involved in various fundamental biological processes.[3] In addition, structural flexibility plays a relevant role in DNA-based nanotechnologies, which are attracting a lot of attention. Natural and/or artificial DNA can be used to design molecular constructs for various technological applications[4] such as molecular machine engineering,[5]

synthetic tissue growth,[6] and nanoscopic bioimaging.[7] Therefore, a deep knowledge of the structural and dynamic properties of DNA is essential both for understanding DNA functioning, also in interaction with other systems (*e.g.* drugs or proteins), and for the design of DNA nanoconstructs.

Given its importance, the investigation of the conformational ensemble of the nucleic acid has been addressed with several experimental approaches.[8] Among other techniques, infrared (IR) spectroscopy has proven to be a suitable tool for gaining insights into the structure and dynamics of DNA because of its high molecular specificity and applicability to a wide variety of samples under native conditions.[9] Most IR experiments targeting nucleic acids focus on the 1500–1800 cm^{-1} frequency region of in-plane base vibrations, where strongly IR-active modes (double bonds stretching, C=O and C=C, and NH$_2$/NH bending) give rise to absorption bands that are highly sensitive to the various conformations of DNA.[10] In particular, the carbonyl stretching signal, which is present in all bases except adenine, displays intensity changes and frequency shifts as a function of the surrounding environment and was

a Department of Chemical Sciences, University of Padova, Padova 35131, Italy
b Center S3, CNR Institute of Nanoscience, via Campi 213/A, 41125 Modena, Italy.
E-mail: laura.zanettipolzi@nano.cnr.it

This journal is © the Owner Societies 2025

Phys. Chem. Chem. Phys., 2025, **27**, 19797–19808 | **19797**

A

B

C$_2$=O
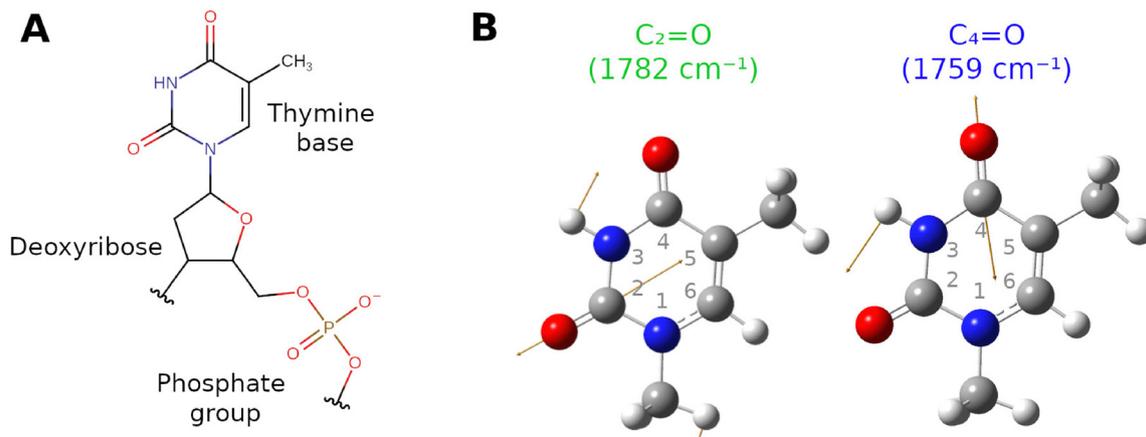(1782 cm^{-1})

C$_4$=O
(1759 cm^{-1})

Fig. 1 (A) Structure of the thymine nucleotide linked to the DNA chain *via* phosphodiester bonds. (B) Structure and numbering convention for 1-methylthymine (1MT); C$_2$=O (left) and C$_4$=O (right) stretching mode eigenvectors and corresponding frequencies as obtained from gas-phase calculations. An analogous picture for 1MT–D is reported in Fig. S1 in the SI.

successfully used to discriminate among different structural motives.[10,11] On the other hand, the presence of in-plane vibrations in all nucleobases leads to the overlap of several bands in the same frequency region, resulting in broad spectra of difficult interpretation. The delocalisation of vibrational modes due to interactions among chromophores further complicates the assignment of specific spectral features to specific molecular arrangements.[12] Nevertheless, linear spectroscopy is a routinely used and quick technique that does not require any specialised experimental setup and, therefore, the understanding of the link between DNA conformations and spectral signatures in linear spectra remains a relevant point.

Computational modelling of IR spectra has the potential to complement experimental data, providing a molecular-level interpretation of specific spectral features, in principle also disentangling the contribution of the environment from that arising from mode coupling effects. In fact, a number of works focused on the calculation of IR spectra arising from DNA in-plane base vibrations, exploring the effect of variations in the local electric field resulting from, *e.g.* base pairing or hydrogen bonding with water. For example, the Hessian matrix reconstruction method, initially developed and applied for the calculation of amide I bands in polypeptides,[13] was extended to investigate base pair vibrational modes and to simulate the IR spectra of various DNA oligomers.[12,14–16] The so-called frequency maps, *i.e.* semiempirical models developed to reconstruct vibrational spectra for specific IR modes on the basis of *ab initio* calculations and experimental data, which were initially developed for peptides,[17,18] have been recently extended to describe the C=O stretching in nucleobases.[19–21] The calculation of non-linear spectra arising from base pair vibrational modes was also addressed.[22–26]

In this work, we propose the use of a computational method based on molecular dynamics (MD) simulations and the perturbed matrix method (PMM), the MD-PMM approach,[27] to compute the IR absorption bands of DNA nucleobases. This approach has already been successfully applied to the calculation of IR spectra in polypeptides, both in the amide I

region[28–30] and for specific side chain bands.[31,32] The MD-PMM approach is a mixed quantum-classical methodology based on the approximation of dividing the system into a region that is treated at the quantum mechanical (QM) level (in the case of IR spectrum calculations, the vibrating chromophore is treated at the QM level) and a complementary region (the environment) that is treated at a classical level and perturbs electrostatically the quantum region. The main strength of the approach is that it allows us to calculate the IR absorption signal arising from a large number of chromophore + environment conformations (*i.e.*, all conformations obtained from a classical MD simulation). In addition, it does not require the use of empirical or adjustable parameters.

Using the MD-PMM approach, we compute here the two IR C=O stretching bands in the thymine base. More specifically, we address the calculation of the bands arising from C$_2$=O and C$_4$=O stretching in 1-methylthymine (1MT, see Fig. 1), thymidine 5′-monophosphate (TMP), and in a 10-base single-stranded oligonucleotide of thymine (dT10) both in water and in D$_2$O. Thymine absorption bands are known to display frequency shifts and/or intensity changes depending on environmental conditions (*e.g.*, single *vs.* double stranded, base pairing, and base stacking)[10] and therefore are a useful marker for DNA conformational variations. In addition, thymine is the only nucleobase featuring two C=O stretching modes, being therefore a good test case for the investigation of possible coupling effects among different vibrational modes. Our calculated spectra show an overall good agreement with the corresponding experimental ones, giving credit to the proposed approach and allowing us to relate subtle spectroscopic changes to modifications in the environment surrounding the chromophore.

2 Materials and methods

2.1 The MD-PMM approach

The MD-PMM methodology has been frequently employed to reconstruct infrared spectra of polypeptides, and thus the

19798 | *Phys. Chem. Chem. Phys.*, 2025, **27**, 19797–19808

This journal is © the Owner Societies 2025

theoretical foundation and computational procedure to obtain vibrational spectra were thoroughly detailed in previous studies.[27,29,33] In this section, we provide a brief overview of this theoretical–computational approach.

In the MD-PMM methodology, a portion of the system, the quantum centre (QC), is treated at the quantum level, and the remainder of the system is described at a classical atomistic level. Unlike other hybrid quantum-classical methods, in the MD-PMM method the entire system, including the QC, is sampled through classical MD simulations. First, the (unperturbed) quantum properties of the isolated QC are calculated quantum mechanically in the gas phase. Next, the electrostatic effect of the environment on the QC electronic states is included using a perturbation operator $\hat{V}$. For each (entire system) configuration generated by the all-atom classical MD simulation, the environment's instantaneous atomic arrangement provides a distinct electrostatic effect. This allows us to obtain the perturbed properties of the QC for a large number of configurations of the environment taking into account the effect of the fluctuating perturbing environment. Specifically, the electronic Hamiltonian operator $\hat{H}$ of the QC embedded in the perturbing environment of a configuration k is expressed as follows:

$$\hat{H}_k = \hat{H}_0 + \hat{V}_k \tag{1}$$

where $\hat{H}_0$ is the QC unperturbed electronic Hamiltonian and the elements $V_{i,j}$ of the perturbation operator between the electronic states i and j are expressed as follows:

$$V_{i,j} = \delta_{i,j}\sum_N \mathscr{V}(\mathbf{R}_N)q_{N,i} - \left(\boldsymbol{E}(\mathbf{r}_0) \cdot \boldsymbol{\mu}_{i,j}\right)\left(1 - \delta_{i,j}\right) \tag{2}$$

with N running over all QC atoms. $\boldsymbol{R}_N$ is the nucleus position of the Nth atom of the QC, $q_{N,i}$ is its atomic charge in the state i, $\mathscr{V}$ is the electrostatic potential exerted by the perturbing environment, $\boldsymbol{E}$ is the perturbing electric field ($\boldsymbol{E} = -\partial\mathscr{V}/\partial\mathbf{r}$) evaluated at the centre of mass $\mathbf{r}_0$ of the QC, $\boldsymbol{\mu}_{i,j}$ is the QC transition dipole moment and $\delta_{i,j}$ is the Kronecker delta.

In this framework, the computational workflow to calculate IR spectra is as follows.

(1) The mass-weighted Hessian eigenvectors are calculated quantum mechanically on the optimised geometry of the QC. In the present case, 1-methylthymine (1MT) is selected as the QC (see Section 2.2). Then, we select the modes of interest (in the present case, the $C_2{=}O$ and $C_4{=}O$ stretching modes of 1MT).

(2) Along the eigenvectors corresponding to the modes of interest, a number of QC configurations are generated. QM calculations are performed on each of these configurations (i.e., points along the mode of interest), providing the unperturbed properties of the QC to be used within the MD-PMM approach (see Section 2.2).

(3) A classical MD simulation of the whole system is performed. In the present case, we perform MDs of 1MT, TMP, and dT10 in water (see Section 2.3).

(4) For each frame of the MD simulations, the perturbed ground state energy of the QC is calculated by constructing and diagonalizing the perturbed Hamiltonian matrix (eqn (1)). This is done for all the configurations generated at point 2 and for both the $C_2{=}O$ and $C_4{=}O$ stretching modes. Note that for calculating the spectrum of dT10, this procedure is repeated for every base of the oligomer. To obtain the spectrum of the first base of the oligomer, the first base is the QC, while the DNA backbone, the remaining 9 bases, and the solvent/counterions are the perturbing environment. To obtain the spectrum of the second base of the oligomer, the second base is the QC, while the DNA backbone, the remaining 9 bases, and the solvent/counterions are the perturbing environment (and so on). In this way, we include the electrostatic effects of the environment, accounting for specific interactions of the chromophore such as hydrogen bonding with the solvent and/or neighboring bases.

(5) The perturbed energy curve obtained at each MD frame (and, for dT10, for each base) is modelled by a Morse potential providing the perturbed vibrational frequency for each MD frame. Note that this fitting procedure also allows for estimation of the anharmonic contribution to the perturbed frequency. A histogram of these frequencies provides the vibrational band of the mode of interest for the QC embedded in its fluctuating perturbing environment, naturally incorporating inhomogeneous broadening.

For dT10, the interaction of multiple vibrational centres (i.e., mode coupling) has to be included in the calculation. This is done by constructing, at each MD frame, an excitonic coupling matrix that includes the coupling among modes at the transition dipole coupling level of approximation, i.e., by expanding the chromophore–chromophore interaction operator up to the dipolar terms. In the present case, coupling effects are considered among the $C_2{=}O$ stretching modes of the 10 bases, among the $C_4{=}O$ stretching modes of the 10 bases, and between the $C_2{=}O$ stretching mode of each residue i and the $C_4{=}O$ stretching mode of each residue j, with $i \neq j$. The excitonic coupling Hamiltonian matrix is given by:[33]

$$\tilde{H} = \tilde{I}\mathscr{U}_{vb}^0 + \Delta\tilde{H} \tag{3}$$

where $\mathscr{U}_{vb}^0$ is the vibrational ground state energy of the interacting chromophores and $\Delta\tilde{H}$ is the excitation matrix, whose diagonal elements are:

$$[\Delta\tilde{H}]_{k,k} = h\nu_k \tag{4}$$

and whose non-zero off-diagonal elements are given by the chromophore–chromophore interaction operator (at the dipolar approximation):

$$\hat{V}_{k,k'(k \neq k')} = \frac{\hat{\boldsymbol{\mu}}_k \cdot \hat{\boldsymbol{\mu}}_{k'}}{R_{k,k'}^3} - 3\frac{\left(\hat{\boldsymbol{\mu}}_k \cdot \mathbf{R}_{k,k'}\right)\left(\hat{\boldsymbol{\mu}}_{k'} \cdot \mathbf{R}_{k,k'}\right)}{R_{k,k'}^5} \tag{5}$$

In the above equations, ν_k is the kth chromophore excitation frequency, $\hat{\boldsymbol{\mu}}_k$ is the kth chromophore dipole operator and $\mathbf{R}_{k,k'}$ is the k' to k chromophore displacement vector. By diagonalizing the above matrix and using the transition dipole for the $0 \rightarrow i$ excitonic transition ($\boldsymbol{\mu}_{0,i}$) obtained via the excitonic

This journal is © the Owner Societies 2025

Phys. Chem. Chem. Phys., 2025, **27**, 19797–19808 | **19799**

eigenvectors, we reconstruct the spectral signal of the excitonic system by summing the absorbance due to each $0 \rightarrow i$ transition:

$$\varepsilon(\nu) = \sum_i \frac{\left|\boldsymbol{\mu}_{0,i}\right|^2 \rho_i(\nu)h\nu}{6\varepsilon_0 c\hbar^2} \tag{6}$$

where ρ_i is the probability density in ν frequency space for the ith excitation and ε_0 is the vacuum dielectric constant.

2.2 QM calculations

The quantum centre to be used in the framework of the MD-PMM approach is, in the present case, 1-methylthymine (1MT). QM calculations are carried out on isolated 1MT and 1MT-D (*i.e.*, 1MT deuterated at N_3, see Fig. 1) at the density functional theory (DFT)[34,35] level with the 6-31+G(d) basis set[36,37] in conjunction with the B3LYP functional[38] and Grimme's dispersion correction with Becke–Johnson damping,[39] *i.e.* at the same level of theory previously employed for the calculation of the infrared spectra of polypeptides with the MD-PMM approach.[29,31,32]

The mass-weighted Hessian matrix is calculated on the optimised geometry of 1MT(-D) at the B3LYP+D3BJ/6-31+G(d) level of theory and subsequently diagonalized to obtain unperturbed eigenvectors and related eigenvalues. The eigenvectors corresponding to the $C_2{=}O$ and $C_4{=}O$ stretching (see Fig. 1B and Fig. S1 in the SI) are then used to generate a number of configurations for each mode: along both the $C_2{=}O$ and $C_4{=}O$ stretching eigenvector 25 points (configurations) are produced using a step of 0.02 a.u. and spanning an energy range of ≈ 50 kJ mol^{-1}. For each point, calculations on the ground state and the lowest three excited states (at the time-dependent DFT level[40]) are performed using the same functional and basis sets, providing the unperturbed properties that will be used in MD-PMM calculations. All QM calculations are performed using the Gaussian16 package.[41] The output files from the calculations performed on both 1MT and 1MT-D are provided in the SI.

2.3 MD simulations

Atomistic MD simulations for 1MT, TMP and the single-stranded dT10 oligonucleotide are performed with the CHARMM36 force field,[42–44] combined with the modified version of the TIP3P water model.[45,46] Force field parameters for 1MT are obtained from the CHARMM general force field (CGenFF).[47] The starting coordinates of TMP are prepared using the Avogadro tool (version 1.2.0).[48] Force field parameters are manually tuned from the default CHARMM36 TMP parameters for the corresponding ribonucleotide. Specifically, atoms O2$'$ and H2$'$ are deleted and replaced by a single hydrogen atom, and their corresponding charges are summed up to the remaining atoms, thereby preserving the overall initial charge of -2. The initial coordinates of dT10 are obtained by generating a double helical structure of B-DNA and then removing the complementary adenine chain. Two and nine Na$^+$ counterions are added in the simulation box of TMP

and dT10, respectively, to neutralise the negative charges of the backbone, while the C$_3'$ and C$_5'$ ends in dT10 are capped with an OH group. Simulations are performed in the isothermal–isobaric ensemble (NPT), using a velocity-rescale thermostat[49] and a Parrinello–Rahman barostat:[50] the reference temperature and pressure are set at 298.15 K and 1 bar, respectively. After energy minimisation, thermalisation (200 ps) and equilibration under NPT conditions (500 ps), trajectories are propagated for 50 ns (1MT), 100 ns (TMP), and 200 ns (dT10) with a time step of 2 fs. Coordinates are saved every 2 ps for the entire system, providing a statistically meaningful sampling of solute and solvent configurations. Following the common prescriptions for the selected force field, all bonds involving H atoms are constrained using the LINCS algorithm.[51] Moreover, a 1.2 nm cutoff is chosen for both the electrostatic and van der Waals interactions. In the latter case, a force-switching function is applied between 1.0 nm and the cutoff, and no long-range correction is considered. On the other hand, long range electrostatic is taken into account by means of the particle-mesh Ewald scheme (PME).[52,53] The size of the simulation boxes (41 nm^3, 43 nm^3, and 264 nm^3 for 1MT, TMP, and dT10, respectively) allows inclusion of a number of solvent molecules sufficient to avoid any unphysical interaction between periodic replicas in adjacent cells, as well as to provide a converged electrostatic perturbation on the QC. All MD simulations and preliminary analyses of the trajectories are performed with the Gromacs suite of packages.[54] Initial structures, input parameters, and topology files for all investigated systems are provided in the SI.

3 Results and discussion

In the spectral region between 1500 and 1800 cm^{-1}, the base in the thymine nucleotide (Fig. 1A) displays three main bands: two $C{=}O$ stretching bands corresponding to the $C_2{=}O$ and $C_4{=}O$ stretching vibrations and the ring vibration associated with the $C_5{=}C_6$ stretching.[58] Here, we focus on the two $C{=}O$ stretching modes and use 1-methylthymine (1MT) as the model system to properly account for the bond between the base and the sugar-phosphate backbone in extended DNA structures. Previous works in fact showed that 1MT is a suitable choice to model the vibrational properties of thymine in more complex structures.[26] For each mode, we compute the gas-phase (unperturbed) stretching frequency and the corresponding eigenvector at the harmonic approximation.[55] The unperturbed stretching frequencies for 1MT are 1782 cm^{-1} and 1759 cm^{-1} for the $C_2{=}O$ and $C_4{=}O$ stretching modes, respectively. These values can be compared to the experimental $C{=}O$ stretching frequencies of isolated 1MT in argon and nitrogen matrices at 10 K, which are 1734 cm^{-1} and 1705 cm^{-1} for the $C_2{=}O$ and $C_4{=}O$ stretching modes, respectively (Table 1). Discrepancies of this order of magnitude (≈ 50 cm^{-1}) are common in this kind of calculations and are due to both intrinsic inaccuracies of the QM calculations and the use of the harmonic approximation.[31,33] As a matter of fact, the experimental $C{=}O$ stretching frequencies of isolated 1MT in argon and

19800 | *Phys. Chem. Chem. Phys.*, 2025, **27**, 19797–19808

This journal is © the Owner Societies 2025

PCCP

Paper

Table 1 Experimental and calculated frequencies in cm^{-1} for 1MT and TMP

		Calc. gas phase	Exp. Ar/N$_2$	Rigid shift	Calc. in solvent	Calc. in solvent shifted	Exp. in solvent	Calc. shift upon solvation	Exp. shift upon solvation
1MT[a]	C$_2$=O	1782	1734[b]	48	1738	1690	1695[d]	44	39
	C$_4$=O	1759	1705[b]	54	1722	1668	1661[d]	37	44
TMP[a]	C$_2$=O	1778	1736[c]	42	1737	1695	1690[e]	41	46
	C$_4$=O	1746	1694[c]	52	1734	1682	1663[e]	12	31

[a] Gas phase calculations are performed on 1MT and 1MT-D for the calculation of the spectrum of 1MT and TMP, respectively. [b] From the study of Morzyk-Ociepa *et al.*[55] [c] Estimated from the study of Ivanov *et al.*,[56] see the main text. [d] From the study of Fick *et al.*[57] [e] From the study of Peng *et al.*[58]

nitrogen matrices also include anharmonic effects. By fitting the unperturbed energy curve to a Morse potential, we estimate the anharmonic contribution to the unperturbed frequency to be 12 cm^{-1} and 13 cm^{-1} for the C$_2$=O and C$_4$=O stretching modes, respectively. Furthermore, by fitting at each MD frame the perturbed energy curve to a Morse potential, the anharmonic contribution to the spectra calculated in water (see below) can also be obtained. This calculation shows that, for both C=O stretching modes, the anharmonic contribution to the perturbed frequencies is essentially constant along the trajectory and unaltered with respect to the gas-phase (see the SI, Fig. S4A). Therefore, we recover the above-mentioned discrepancy between the calculated gas phase frequency and the corresponding experimental frequency in Ar/N$_2$ matrices by applying a rigid frequency shift that corrects the unperturbed frequencies for both anharmonicity and inaccuracies of QM calculations. Thus, each perturbed frequency obtained at the harmonic level is downshifted by 48 cm^{-1} and 54 cm^{-1} for the C$_2$=O and C$_4$=O stretching modes, respectively, as summarised in Table 1.

Following the procedure outlined in Section 2.1, we then compute the C$_2$=O and C$_4$=O bands for 1MT in water. We use to this aim all the environment configurations obtained from the classical MD simulation of 1MT in explicit water coupled to the QM calculations of 1MT (see Section 2). The calculated spectra are shown in Fig. 2A and compared to the experimental spectrum of 1MT in water.[57] Since for a single 1MT in water no mode coupling effects have been included, the bands reported in Fig. 2A correspond to the distribution of the perturbed frequencies, weighted by the corresponding transition dipoles, obtained at each MD frame for each mode. It can be observed that the experimental frequency downshift due to solvation is well captured by the calculated bands (see also Table 1). More specifically, the frequency peaks of both calculated bands are in agreement with the corresponding experimental peaks within 10 cm^{-1} (1690 cm^{-1} calc. *vs.* 1695 cm^{-1} exp. for the C$_2$=O band; 1668 cm^{-1} calc. *vs.* 1661 cm^{-1} exp. for the C$_4$=O band). However, the relative intensity of the two experimental peaks is not exactly reproduced in the calculated bands (*i.e.*, the intensity of the C$_2$=O band is slightly underestimated). Note that in the experimental spectrum a low-frequency shoulder is also present, arising from the C$_5$=C$_6$ stretching band that is not included in our calculations.

Due to overlap with the water bending mode (≈ 1640 cm^{-1}), C=O stretching bands in water result from the subtraction of
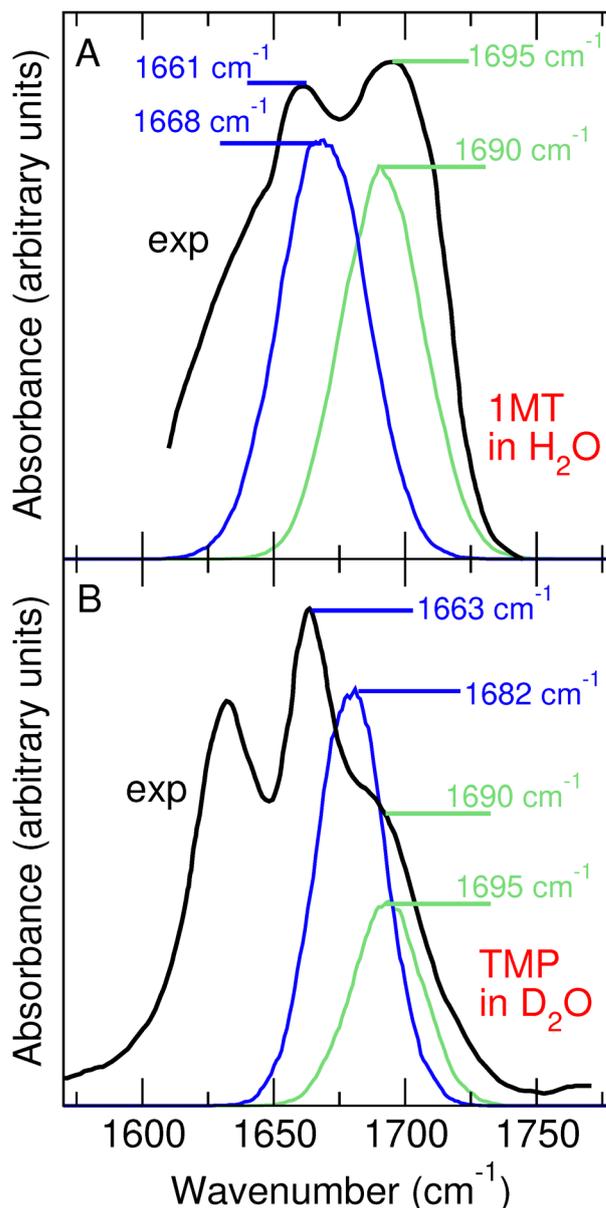

Fig. 2 Calculated and experimental IR signals for 1MT in water (A) and TMP in D$_2$O (B). The calculated C$_2$=O and C$_4$=O bands are reported in green and blue, respectively. The experimental spectra of 1MT in water[57] and TMP in D$_2$O[58] are reported in black.

the water absorption band from the solution spectra, with possible inaccuracies in both the peak position and the

This journal is © the Owner Societies 2025

Phys. Chem. Chem. Phys., 2025, **27**, 19797–19808 | **19801**

intensity. Therefore, experiments are often performed in D_2O to avoid the interference of H_2O bending modes. We thus perform the calculation of the $C_2{=}O$ and $C_4{=}O$ bands in thymidine 5'-monophosphate (TMP, *i.e.*, the thymine base bound to D-ribose-5-phosphate) in D_2O. We use for this purpose the QM calculations on 1MT-D (see Section 2.2) and the MD simulation of TMP in water. In fact, the explicit inclusion of deuterated water and deuteration of the TMP molecule in the MD simulation would not affect the equilibrium properties of the system and therefore would not affect the calculated spectrum. QM calculations on 1MT-D show that the (gas-phase) unperturbed stretching frequencies are 1778 cm^{-1} for the $C_2{=}O$ stretching mode and 1746 cm^{-1} for the $C_4{=}O$ stretching mode (Fig. S1 in the SI). The same calculations successfully predict a third normal mode at 1702 cm^{-1}, whose larger amplitudes are on the $C_5{=}C_6$ stretching of the pyrimidine ring, in agreement with previous experimental observations and assignments[55,57,58] (see also Fig. S2 and S3 in the SI). Although it has not been explicitly modeled in this work, this mode is responsible for the lowest energy band in the reference experimental spectra shown in the following.

The calculated $C_2{=}O$ and $C_4{=}O$ gas-phase frequencies should be compared with the experimental values for isolated 1MT-D. However, we could not find any experimental estimate of these frequencies. We could instead find an experimental estimate of the C=O stretching frequencies in an Ar matrix for thymidine and thymidine deuterated at N_3.[56] Therefore, we apply the frequency shift upon deuteration experimentally observed in thymidine to the experimental frequencies in the argon and nitrogen matrices of 1MT to obtain the reference frequencies for 1MT-D. Using this approach, we downshift each perturbed frequency obtained at the harmonic level by 42 cm^{-1} and 52 cm^{-1} for the $C_2{=}O$ and $C_4{=}O$ stretching modes, respectively (Table 1). Fitting the unperturbed energy curve to a Morse potential, we estimate the anharmonic contribution to the shift between the reference and calculated vacuum frequencies of 1MT-D to be 11 cm^{-1} and 7 cm^{-1} for the $C_2{=}O$ and $C_4{=}O$ stretching modes, respectively, and verified that, also for 1MT-D, the anharmonic contribution to the perturbed frequencies is essentially constant along the trajectory and unaltered with respect to the gas-phase (see the SI, Fig. S4B).

The calculated bands are reported in Fig. 2B and compared to the total experimental spectrum of TMP in D_2O.[58] Unlike the water case, and in accordance with the experiments, the two bands in D_2O have a different intensity ($I(C_2{=}O) \approx 0.6 \cdot I(C_4{=}O)$). This feature depends on the difference in the gas phase transition dipole moments associated with the vibrational eigenvectors of the $C_2{=}O$ stretching mode in 1MT and 1MT-D, and was previously observed in similar gas phase calculations.[14] It can be yet observed that while the peak position of the $C_2{=}O$ band is well reproduced by the calculation, that of the $C_4{=}O$ band is overestimated with respect to the corresponding experimental peak by 19 cm^{-1} (note that in the experimental spectrum the lowest frequency peak arises from the $C_5{=}C_6$ stretching mode that is not included in our calculations). To understand the origin of the underestimated

shift upon solvation of the $C_4{=}O$ band in D_2O, we compute the spectrum of 1MT-D in D_2O. Although we do not have an experimental counterpart for the spectrum of 1MT-D in D_2O, this test allows us to understand if the origin of the above discrepancy can be ascribed to the presence of the D-ribose-5-phosphate group. As shown in Table S1 and in Fig. S5 in the SI, the $C_2{=}O$ and $C_4{=}O$ bands calculated for 1MT-D in D_2O are very similar to those calculated for TMP, suggesting that the presence of the D-ribose-5-phosphate group does not affect the spectra. In addition, to compute the bands for 1MT in water and 1MT-D in D_2O, we use the same simulation and, therefore, the difference between the two $C_4{=}O$ bands cannot be ascribed to the classical MD. The other possible sources of underestimation of the $C_4{=}O$ band in D_2O are the methodology used for the calculation and/or the QM calculations. In fact, it has been highlighted that in our approach the presence of water molecules hydrogen-bonded to the thymine base is only included as a perturbative electrostatic effect, thus neglecting the effects due to the quantum delocalisation of protons. Nonetheless, this approximation is expected to affect the $C_4{=}O$ band calculated in water more than that in D_2O and thus does not explain the underestimation of the shift in D_2O. We therefore focus on the possible effects of the QM calculation, and specifically on the Hessian eigenvectors defining the $C_2{=}O$ and $C_4{=}O$ stretching modes in 1MT and 1MT-D.

In fact, assuming that off-diagonal couplings with excited states are negligible compared to the diagonal correction (*i.e.* if we neglect the second term in eqn (2)) in determining the perturbed ground state energy profile along a given normal mode Q, the corresponding frequency can be obtained as follows:

$$\nu \approx \frac{\partial^2}{\partial Q^2}\left(H^0 + \sum_N \mathscr{V}(\mathbf{R}_N)q_{N,0}\right)$$
$$= \nu^0 + \sum_N \frac{\partial^2(\mathscr{V}(\mathbf{R}_N)q_{N,0})}{\partial Q^2} \tag{7}$$

where $q_{N,0}$ is the partial charge on the Nth atom in the electronic ground state. The correction to the unperturbed frequency ν^0 can be further expanded as follows:

$$\frac{\partial^2(\mathscr{V}(\mathbf{R}_N)q_{N,0})}{\partial Q^2} = \frac{\partial^2 \mathscr{V}(\mathbf{R}_N)}{\partial Q^2}q_{N,0} + \mathscr{V}(\mathbf{R}_N)\frac{\partial^2 q_{N,0}}{\partial Q^2}$$
$$+ 2\frac{\partial q_{N,0}}{\partial Q}\frac{\partial \mathscr{V}(\mathbf{R}_N)}{\partial Q} \tag{8}$$

Remarkably, as the definition of a vibrational eigenvector dictates the molecular deformation along the associated mode, it affects both the calculation of atomic charges and the (electrostatic) perturbation acting on them: in the latter case, the dependence is explicit, while in the former it is implicit – through the parametric dependence of the electronic density on nuclear geometry. Indeed, by substituting the derivative of the potential as follows:

$$\frac{\partial \mathscr{V}(\mathbf{R}_N)}{\partial Q} = \sum_a \frac{\partial \mathscr{V}(\mathbf{R}_N)}{\partial \mathbf{R}_N^a}\frac{\partial \mathbf{R}_N^a}{\partial Q} = -\sum_a \mathbf{E}_N^a \frac{\partial \mathbf{R}_N^a}{\partial Q} \tag{9}$$

19802 | *Phys. Chem. Chem. Phys.*, 2025, **27**, 19797–19808

This journal is © the Owner Societies 2025

where $\mathbf{R}_N^a$ is the Cartesian displacement of the Nth atom of the QC along the direction a, $\mathbf{E}_N^a$ is the Cartesian component a of the electric field acting on the Nth QC atom, and $\partial \mathbf{R}_N^a/\partial Q$ is the Cartesian displacement along a of the same atom for mode Q; it becomes evident that all terms in eqn (8) encode the sensitivity of either (atomic) charges or positions to the underlying displacement along the selected normal mode. We thus refer to an "electrostatic" effect and a "geometric" effect, respectively, associated with mode definition in the MD-PMM calculation.

By dissecting the total shift upon solvation into these two contributions (see Table S2 in the SI), we observe that the downshift due to the geometric effect is similar for $C_2{=}O$, both in 1MT and 1MT-D, as well as for $C_4{=}O$ in 1MT, while it is significantly reduced for $C_4{=}O$ in 1MT-D. As shown in Fig. 1, the eigenvectors associated with the $C_2{=}O$ and $C_4{=}O$ bands in 1MT are very similar, with a major component associated with the C=O stretching and a minor but significant component associated with the NH bending. Coherently with deuteration at N_3, in 1MT-D, the NH bending component is almost absent and a minor component on the other C=O group appears (Fig. S1 in the SI). Yet, the projections on the two C=O groups are in the same direction in the $C_2{=}O$ mode and in opposite directions in the $C_4{=}O$ mode, determining a different geometric effect. Regarding the electrostatic effect, the charge variation along the mode for the groups interacting with water (*i.e.*, C=O and N–H) is in line with the observed downshifts (Table S2 in the SI), with $C_4{=}O$ in 1MT-D featuring a significantly lower charge variation along the mode (Fig. S6 in the SI). The above results suggest that the QM calculations performed here do not provide an optimal description of the eigenvector associated with the $C_4{=}O$ band in 1MT-D. We checked that the definition of both $C_2{=}O$ and $C_4{=}O$ vibrational modes shows negligible dependence on the basis set (see Tables S3 and S4 in the SI). A thorough investigation of the effect of different exchange and correlation functionals, as well as of moving to other electronic structure methods, will be pursued in future work. For the scopes of the present work, we apply an additional rigid downshift of 19 cm^{-1} to the $C_4{=}O$ band in D_2O, obtaining a good reproduction of the experimental lineshape (see Fig. S7 in the SI). In fact, our main aim is to reproduce the spectral changes due to the surrounding environment (*e.g.*, structural arrangements, hydrogen bonding patterns, and interaction with water).

Then, we test the capability of our calculations to reproduce the difference that is experimentally observed between the single molecule case and more extended DNA structures. We use this aim a single-stranded 10-base thymine oligonucleotide (dT10, 5′-TTTTTTTTTT-3′) for which the experimental spectrum is available.[59] The comparison between the experimental spectra of TMP and dT10 in D_2O (Fig. 3A) shows that no appreciable difference in the frequency position of the two bands is present. However, the relative intensity of the two bands is different in the two cases: in the spectrum of dT10 the highest frequency shoulder has a higher intensity with respect to the TMP case. To calculate the vibrational bands arising from both $C_2{=}O$ and $C_4{=}O$ stretching modes in dT10 in D_2O, we
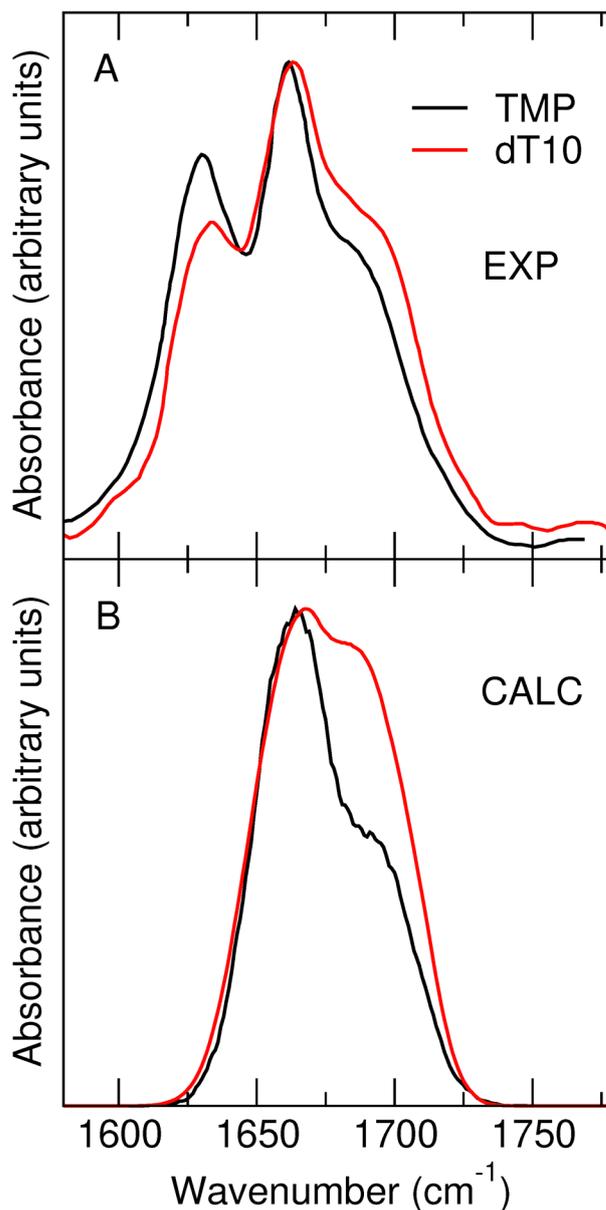

Fig. 3 Experimental (A) and calculated (B) IR signals for TMP (black)[58] and dT10[59] (red) in D_2O.

use the QM calculations on 1MT-D and compute the perturbed frequencies for both vibrational modes for each of the 10 bases of the oligomer along the MD simulation of dT10. To obtain the final spectrum, we also include intra- and inter-mode coupling effects within the transition dipole coupling approximation (see Section 2.1). The starting structure of the MD simulation of dT10 is an extended unrealistic structure (see Section 2.3); therefore, we exclude from the trajectory the first tens of nanoseconds in which the conformational ensemble can be biased by the initial configuration. For the calculation of the dT10 spectrum, we thus use the portion of trajectory from 60 to 200 ns. The calculated spectrum of dT10 is shown in Fig. 3B compared to the calculated spectrum of TMP and the experimental spectra of TMP and dT10 (Fig. 3A). It can be observed

from the figure that the intensity increase in the high-frequency shoulder that is experimentally observed going from TMP to dT10 is present also in the calculated spectrum, although somewhat overestimated. It has also been noted that the experimental spectral lineshape, characterised by a major peak and a higher frequency shoulder, is reproduced in the calculated spectra only if the excitonic contribution is included in the calculation and considering both intra- and inter-mode couplings. In fact, as shown in Fig. 4, if no excitonic coupling or only intra-mode excitonic coupling is included, the total calculated dT10 spectrum is a broad band in which the two experimental features cannot be distinguished. This observation suggests (i) that the relative orientation of the bases in the oligomer plays a relevant role in the high-frequency shoulder increase and (ii) that the $C_2{=}O$ and $C_4{=}O$ bands are strongly coupled and therefore the high-frequency shoulder does not only arise from the $C_2{=}O$ stretching mode.

To understand the origin of the high-frequency shoulder increase in terms of the structural arrangement of bases in the oligomer, we investigate the conformational ensemble of dT10 in our MD in terms of base stacking. More specifically, we investigate the stacking fraction for each base and the time evolution of the stacking fraction of the whole oligomer along the trajectory (Fig. 5). To this end, we consider two bases stacked in a specific MD frame if the distance between the centre of mass of the two bases is less than or equal to 4 Å and if the angle between the two normals to the base planes is less
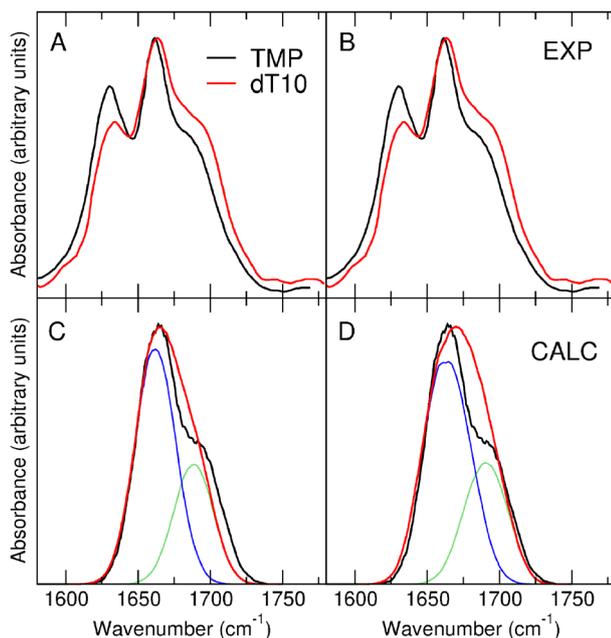

Fig. 4 (A) and (B) Experimental IR signals for TMP[58] (black) and dT10[59] (red) in D_2O. (C) and (D) Calculated IR signals for TMP (black) and dT10 (red) in D_2O. For dT10, the calculated $C_2{=}O$ and $C_4{=}O$ bands are reported in green and blue, respectively. The total calculated spectrum arising from the sum of the $C_2{=}O$ and $C_4{=}O$ bands is reported in red. In (C), the two calculated bands are obtained without the inclusion of the excitonic coupling; in (D), only the intramode excitonic coupling is included.
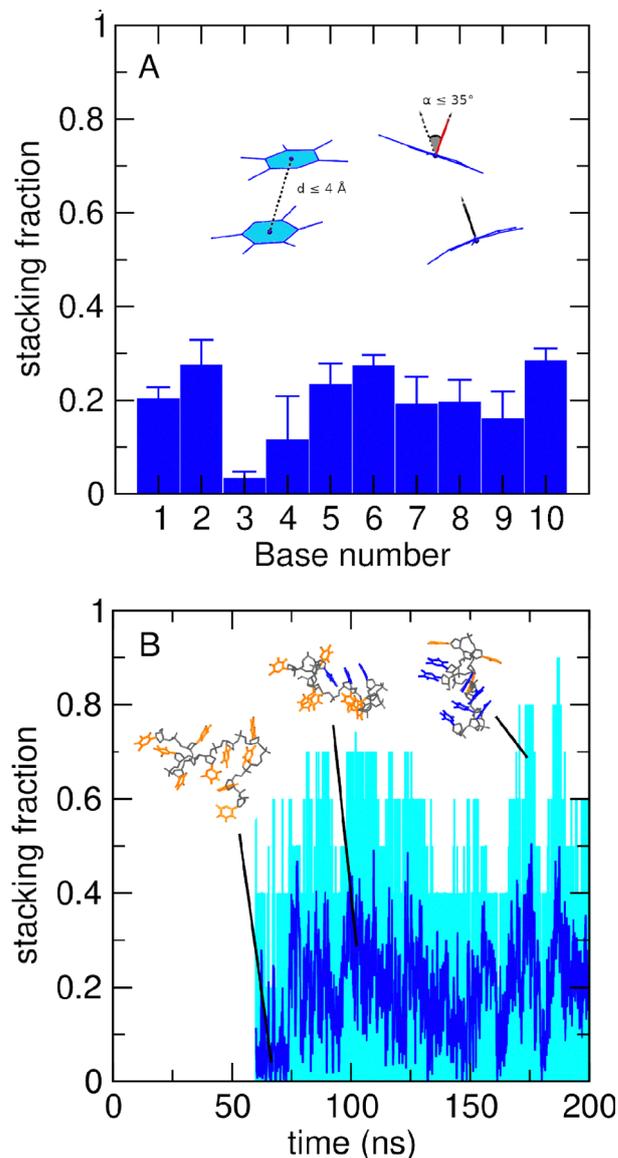

Fig. 5 (A) Stacking fraction of each base along the dT10 MD trajectory. In the inset, the geometrical criteria defining two bases as stacked are sketched. (B) Time evolution of the total stacking fraction along the MD simulation of the dT10 oligomer. In cyan, the stacking fraction for each MD frame is reported, in blue fast fluctuations are filtered with a rolling average. In the inset, representative snapshots corresponding to a zero, 0.3 and 0.7 stacking fraction are reported (stacked bases are highlighted in blue and solvent exposed bases are highlighted in orange).

than or equal to 35°. Using these criteria, we obtain that the average stacking fraction of the simulated dT10 oligomer is 0.20 ± 0.08. This observation is consistent with a previous experimental estimate obtained from small-angle X-ray scattering measurements: the stacking fraction in a single-stranded oligomer composed of 30 thymine bases was found to be 0.16 ± 0.12.[60] As illustrated in Fig. 5A, the 0.2 stacking fraction is sampled almost uniformly by all oligomer bases, with the exception of base 3, which exhibits a significantly lower stacking fraction. Fig. 5B also shows that, while the average fraction

19804 | *Phys. Chem. Chem. Phys.*, 2025, **27**, 19797–19808

This journal is © the Owner Societies 2025

is essentially constant along the trajectory, there are relevant fluctuations: there are portions of the trajectory in which a stacking fraction of around 0.7 is reached. We therefore select along our MD a trajectory interval featuring an approximately null stacking fraction and a trajectory interval featuring an above average (≈ 0.4) stacking fraction. We then compute the spectra arising from the $C_2=O$ and $C_4=O$ stretching modes on these trajectory subparts and compare them to the total spectrum. It can be clearly observed from Fig. 6 that the relative intensity of the low- and high-frequency bands is sensitive to the stacking fraction. More specifically, at low stacking fractions, the high-frequency band can barely be observed as a shoulder, while at high stacking fractions it becomes a separate peak (see also the inset of Fig. 6). These data suggest that the increase in intensity of the high-frequency shoulder in dT10 with respect to the TMP observed in both the experimental and the calculated spectra can be attributed to stacking interactions in the dT10 oligomer.

We then investigate the dT10 structural ensemble also in terms of hydrogen bonds among bases. The plots in Fig. 7 show that the spectral lineshape, and in particular the relative intensity between the low-frequency peak and the high-frequency shoulder, is also sensitive to the absence or presence of hydrogen bonds. Throughout the trajectory, the bases are rarely hydrogen-bonded (as shown in Fig. 7C, no hydrogen bond is present in 80% of the cases) providing the spectrum shown in Fig. 2B and recalled in Fig. 7A. Calculating the dT10 spectrum on two independent portions of the trajectory featuring a higher fraction of conformations in which a hydrogen
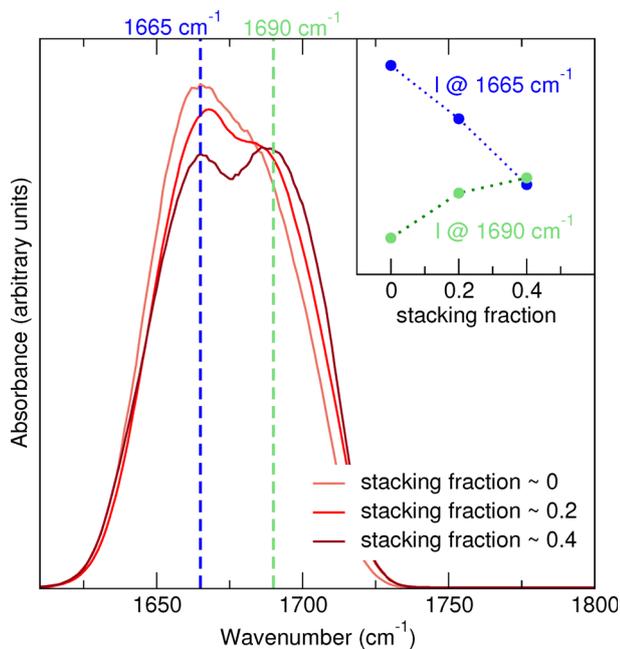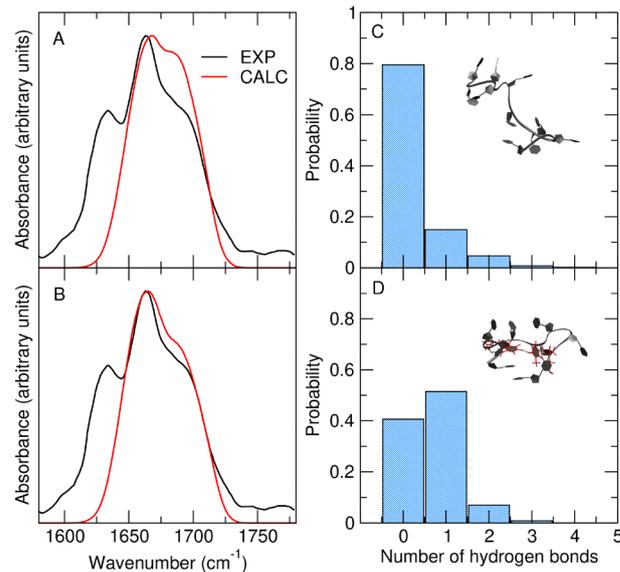

Fig. 7 (A) and (B) Experimental (black) and calculated (red) spectra of dT10 considering the whole trajectory (A) or a portion of the trajectory along which a persistent hydrogen bond between two bases is present (B). (C) and (D) Histogram of the number of hydrogen bonds among bases along the whole trajectory (C) and along the portion of trajectory used to calculate the spectrum shown in (B) (D). The insets of panels (C) and (D) show two representative snapshots with no hydrogen bonds (C) and one hydrogen bond (D) between two bases (highlighted in red).

bond is present between two bases (Fig. 7D), results in a different relative intensity of the two bands. More specifically, the overestimation of the high-frequency shoulder intensity is recovered (Fig. 7C). This result suggests that the experimental signal corresponds to a ≈ 0.5 fraction of hydrogen-bonded bases, also suggesting that this fraction is underestimated by the force field. We note that the presence of hydrogen-bonded bases in dT10 implies the coexistence of extended and coil-like conformations, already observed in single-stranded DNA structures.[61] This observation is in line with the heterogeneous conformational ensemble reconstructed from small angle X-ray scattering experiments on dT30,[60] in which both extended and more collapsed structures are present. It should however be considered that the different length of the oligomers (30 vs. 10 bases) might affect the overall conformational ensemble.

We then move to investigating the coupling effects among the vibrational modes in dT10 to verify that, as suggested by the comparison between the spectra computed by including only intra- or both intra- and inter-mode couplings (*vide supra*), the $C_2=O$ and $C_4=O$ bands are strongly coupled. We therefore analyse the single excitonic state spectra contributing to the total spectrum. As shown in Fig. 8A, the spectra arising from the excitonic states from 1 to 11 (with the excitonic states ordered from the lowest to the highest energy/wavenumber) roughly correspond to the low-frequency band. On the other hand, the spectra arising from the excitonic states from 12 to 20 roughly correspond to the high-frequency band. The diagonalisation of the excitonic Hamiltonian matrix (eqn (3)) provides at each MD frame the eigenvectors representing each excitonic
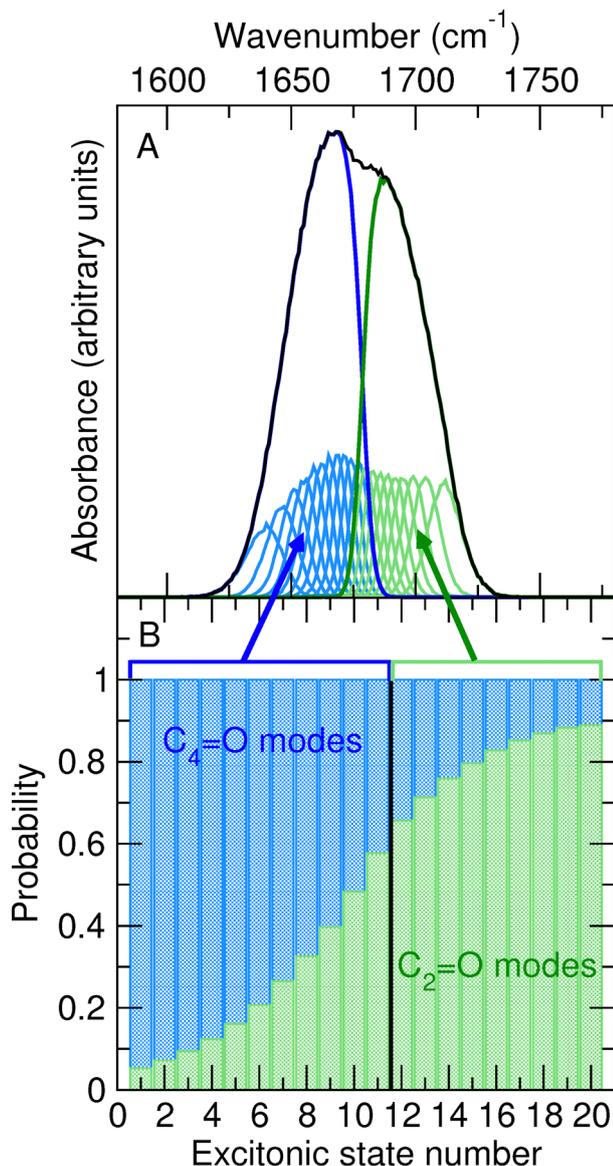

Fig. 6 Calculated spectrum of dT10 under different stacking fraction conditions: light red, zero stacking fraction; red, ≈ 0.2 stacking fraction; dark red ≈ 0.4 stacking fraction. In the inset, the intensity at the frequency corresponding to the low (1665 cm^{-1}, blue) and high (1690 cm^{-1}, green) frequency peaks is reported for the three stacking fractions.

This journal is © the Owner Societies 2025

Phys. Chem. Chem. Phys., 2025, **27**, 19797–19808 | **19805**

Fig. 8 (A) Calculated spectrum of dT10 in D$_2$O (black). The spectra corresponding to the 20 excitonic states, arising from the inter- and intra-mode coupling of the 10 C$_2$=O and 10 C$_4$=O modes, are also reported: the spectra of the excitonic states from 1 to 11 are colored in light blue and those of the excitonic states from 12 to 20 are in green. The spectrum arising from the sum of single excitonic states 1–11 is shown in blue and that arising from the sum of single excitonic states 12–20 is shown in green. (B) For each excitonic state, the probability of originating from C$_2$=O modes (green) or from C$_4$=O modes (blue) is reported.

state in the basis of the 10 C$_2$=O and 10 C$_4$=O modes. By averaging the square modulus of each component of each of the 20 excitonic eigenvectors, we can obtain the probability of the ith excitonic state to excite the vibration of a single specific chromophore. In other words, we can estimate the delocalisation of each excitonic state on the 10 C$_2$=O and C$_4$=O modes arising from the 10 bases in dT10. In Fig. 8B, we report this probability considering together all C$_2$=O modes (green) or all C$_4$=O modes (blue). As anticipated based on the frequency

position of the C$_2$=O and C$_4$=O bands when considered as isolated (see Fig. 2 and 4), the excitonic states at the lowest frequencies mostly arise from C$_4$=O modes and those at the highest frequency mostly arise from C$_2$=O modes. Nonetheless, the contribution of both C$_2$=O and C$_4$=O modes is present in all excitonic states and is particularly relevant for the excitonic states at intermediate frequencies. More specifically, it can be observed that the spectra of the three highest energy excitonic states, which contribute the most to the splitting of the high-frequency shoulder from the low-frequency peak, feature an average C$_4$=O component of ≈12%. Therefore, there is a non-negligible contribution of C$_4$=O modes in the increase of the high-frequency shoulder in dT10 with respect to TMP.

Finally, we also observe that the inverse participation ratio,

$$1 \bigg/ \sum_i |c_i|^4,$$

providing an estimate of the effective delocalisation of the excitonic wavefunction on the interacting states, is equal to 2.02 ± 0.18 (mean and standard deviation over all the excitonic states). This value suggests that excitonic coupling mostly involves pairs of thymine residues, enabling the hybridization of their respective vibrational modes. As coupling is largely favored by reducing the distance between the interacting bases, both paired and stacked arrangements represent favorable, though not exclusive, conformations for achieving efficient excitonic hybridization. Importantly, the energy and brightness of the resulting eigenstates (*i.e.*, their contribution to the overall spectrum) depend on the details of the interaction geometries. For instance, the set of excitonic states originating from the subpopulation of π-stacked thymine pairs can be classified as H-like excitons,[62] characterized by a blue-shifted bright transition: these insights into the underlying excitonic character can complement the previous findings that stacking is the key ingredient underlying the higher frequency shoulder of the absorption profile (see Fig. 6).

4 Conclusions

We present the calculation of infrared spectra arising from the two C=O stretching modes in the thymine base (C$_2$=O and C$_4$=O) using a theoretical–computational approach that combines classical molecular dynamics simulations, quantum mechanical calculations on the isolated base, and the perturbed matrix method. First, we calculate the spectra of 1-methylthymine in water and thymidine 5′-monophosphate in deuterated water, showing that the proposed approach effectively captures most of the frequency shift due to solvation effects. However, we also note that the C$_4$=O stretching frequency is overestimated by the calculation by 19 cm^{-1} in deuterated water. The investigation of this disagreement suggests that the overestimated frequency originates from the definition of the eigenvector corresponding to the C$_4$=O mode in the QM calculations of 1MT-D.

We then compute the spectrum of a single-stranded oligomer composed of 10 thymine bases, showing that the

19806 | *Phys. Chem. Chem. Phys.*, 2025, **27**, 19797–19808

This journal is © the Owner Societies 2025

experimentally observed difference in the relative intensity of the two bands between the spectrum of the solvated single TMP nucleotide and that of the oligomer is reproduced by the calculated spectra. We also show that the high-frequency band intensity increase can be ascribed to excitonic effects due to stacked configurations, including contributions of both the C_2=O and C_4=O modes that result to be strongly coupled. Our results also suggest that the conformational ensemble of the single-stranded dT10 oligomer is characterised by a modest degree of stacking and hydrogen bonding interactions among the bases, which nevertheless influence the spectral lineshape.

Author contributions

SC: conceptualisation, writing – review and editing, supervision, funding acquisition, and project administration. CJDF: software, formal analysis, methodology, visualisation, writing – original draft, and writing – review and editing. GP: investigation, formal analysis, methodology, software, visualisation, writing – original draft, and writing – review and editing. LZP: conceptualisation, methodology, formal analysis, investigation, writing – original draft, writing – review and editing, visualisation, supervision, funding acquisition, and project administration.

Conflicts of interest

There are no conflicts to declare.

Data availability

The data supporting this article have been included as part of the SI. Supplementary information available: additional figures and tables; output files from the quantum mechanical calculations performed on both 1MT and 1MT-D; initial structures, input parameters and topology files for all molecular dynamics simulations. See DOI: https://doi.org/10.1039/d5cp02037a

Acknowledgements

Notes and references

1 D. Varshney, J. Spiegel, K. Zyner, D. Tannahill and S. Balasubramanian, *Nat. Rev. Mol. Cell Biol.*, 2020, **21**, 459–474.
2 M. Dalla Pozza, A. Abdullrahman, C. J. Cardin, G. Gasser and J. P. Hall, *Chem. Sci.*, 2022, **13**, 10193–10215.
3 M. Kaushik, S. Kaushik, K. Roy, A. Singh, S. Mahendru, M. Kumar, S. Chaudhary, S. Ahmed and S. Kukreti, *Biochem. Biophys. Rep.*, 2016, **5**, 388–395.
4 N. C. Seeman and H. F. Sleiman, *Nat. Rev. Mater.*, 2017, **3**, 1–23.
5 H. Ramezani and H. Dietz, *Nat. Rev. Genet.*, 2020, **21**, 5–26.
6 N. Arulkumaran, M. Singer, S. Howorka and J. R. Burns, *Nat. Commun.*, 2023, **14**, 1314.
7 F. Li, J. Li, B. Dong, F. Wang, C. Fan and X. Zuo, *Chem. Soc. Rev.*, 2021, **50**, 5650–5667.
8 L. Salmon, S. Yang and H. M. Al-Hashimi, *Annu. Rev. Phys. Chem.*, 2014, **65**, 293–316.
9 K. B. Beć, J. Grabska and C. W. Huck, *Anal. Chim. Acta*, 2020, **1133**, 150–177.
10 M. Banyay, M. Sarkar and A. Gräslund, *Biophys. Chem.*, 2003, **104**, 477–488.
11 J. Liquier and E. Taillandier, *Infrared Spectroscopy of Biomolecules*, Wiley-Liss, New York, 1996.
12 C. Lee, K.-H. Park and M. Cho, *J. Chem. Phys.*, 2006, **125**, 114508.
13 S. Ham, S. Cha, J.-H. Choi and M. Cho, *J. Chem. Phys.*, 2003, **119**, 1451–1461.
14 C. Lee and M. Cho, *J. Chem. Phys.*, 2006, **125**, 114509.
15 C. Lee, K.-H. Park, J.-A. Kim, S. Hahn and M. Cho, *J. Chem. Phys.*, 2006, **125**, 114510.
16 C. Lee and M. Cho, *J. Chem. Phys.*, 2007, **126**, 145102.
17 S. Ham and M. Cho, *J. Chem. Phys.*, 2003, **118**, 6915–6922.
18 J. R. Schmidt, S. A. Corcelli and J. L. Skinner, *J. Chem. Phys.*, 2004, **121**, 8887–8896.
19 Y. Jiang and L. Wang, *J. Phys. Chem. B*, 2019, **123**, 5791–5804.
20 Y. Jiang and L. Wang, *J. Chem. Phys.*, 2020, **152**, 084114.
21 W. Meng, H.-C. Peng, Y. Liu, A. Stelling and L. Wang, *J. Phys. Chem. B*, 2023, **127**, 2351–2361.
22 A. T. Krummel, P. Mukherjee and M. T. Zanni, *J. Phys. Chem. B*, 2003, **107**, 9165–9169.
23 A. T. Krummel and M. T. Zanni, *J. Phys. Chem. B*, 2006, **110**, 13991–14000.
24 C. Greve, N. K. Preketes, R. Costard, B. Koeppe, H. Fidder, E. T. Nibbering, F. Temps, S. Mukamel and T. Elsaesser, *J. Phys. Chem. A*, 2012, **116**, 7636–7644.
25 C. Greve, N. K. Preketes, H. Fidder, R. Costard, B. Koeppe, I. A. Heisler, S. Mukamel, F. Temps, E. T. Nibbering and T. Elsaesser, *J. Phys. Chem. A*, 2013, **117**, 594–606.
26 J.-J. Ho, D. R. Skoff, A. Ghosh and M. T. Zanni, *J. Phys. Chem. B*, 2015, **119**, 10586–10596.
27 L. Zanetti-Polzi, S. Del Galdo, I. Daidone, M. D'Abramo, V. Barone, M. Aschi and A. Amadei, *Phys. Chem. Chem. Phys.*, 2018, **20**, 24369–24378.
28 L. Zanetti Polzi, A. Amadei, M. Aschi and I. Daidone, *J. Am. Chem. Soc.*, 2011, **133**, 11414–11417.
29 L. Zanetti-Polzi, M. Aschi, A. Amadei and I. Daidone, *J. Phys. Chem. B*, 2013, **117**, 12383–12390.
30 C. M. Davis, L. Zanetti-Polzi, M. Gruebele, A. Amadei, R. B. Dyer and I. Daidone, *Chem. Sci.*, 2018, **9**, 9002–9011.
31 S. M. V. Pinto, N. Tasinato, V. Barone, A. Amadei, L. Zanetti-Polzi and I. Daidone, *Phys. Chem. Chem. Phys.*, 2020, **22**, 3008–3016.
32 S. Pinto, N. Tasinato, V. Barone, L. Zanetti-Polzi and I. Daidone, *J. Chem. Phys.*, 2021, **154**, 084105.

This journal is © the Owner Societies 2025

Phys. Chem. Chem. Phys., 2025, **27**, 19797–19808 | 19807

33 A. Amadei, I. Daidone, L. Zanetti-Polzi and M. Aschi, *Theor. Chem. Acc.*, 2011, **129**, 31–43.

34 P. Hohenberg and W. Kohn, *Phys. Rev.*, 1964, **136**, B864.

35 W. Kohn and L. J. Sham, *Phys. Rev.*, 1965, **140**, A1133.

36 A. Petersson, A. Bennett, T. G. Tensfeldt, M. A. Al-Laham, W. A. Shirley and J. Mantzaris, *J. Chem. Phys.*, 1988, **89**, 2193–2218.

37 G. Petersson and M. A. Al-Laham, *J. Chem. Phys.*, 1991, **94**, 6081–6090.

38 C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B:Condens. Matter Mater. Phys.*, 1988, **37**, 785.

39 S. Grimme, S. Ehrlich and L. Goerigk, *J. Comput. Chem.*, 2011, **32**, 1456–1465.

40 M. E. Casida and M. Huix-Rotllant, *Annu. Rev. Phys. Chem.*, 2012, **63**, 287–323.

41 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman and D. J. Fox, *Gaussian~16 Revision C.01*, Gaussian Inc., Wallingford CT, 2016.

42 B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan and M. Karplus, *J. Comput. Chem.*, 1983, **4**, 187–217.

43 A. D. MacKerell, N. Banavali and N. Foloppe, *Biopolymers*, 2000, **56**, 257–265.

44 K. Hart, N. Foloppe, C. M. Baker, E. J. Denning, L. Nilsson and A. D. J. MacKerell, *J. Chem. Theory Comput.*, 2012, **8**, 348–362.

45 W. E. Reiher III, PhD thesis, Department of Chemistry, Harvard University, 1985.

46 E. Neria, S. Fischer and M. Karplus, *J. Chem. Phys.*, 1996, **105**, 1902–1921.

47 K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes and I. Vorobyov, *et al.*, *J. Comput. Chem.*, 2010, **31**, 671–690.

48 M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeersch, E. Zurek and G. R. Hutchison, *J. Cheminf.*, 2012, **4**, 1–17.

49 G. Bussi, D. Donadio and M. Parrinello, *J. Chem. Phys.*, 2007, **126**, 014101.

50 M. Parrinello and A. Rahman, *Phys. Rev. Lett.*, 1980, **45**, 1196–1199.

51 B. Hess, H. Bekker, H. J. C. Berendsen and J. G. E. M. Fraaije, *J. Comput. Chem.*, 1997, **18**, 1463–1472.

52 T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089–10092.

53 U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. G. Pedersen, *J. Chem. Phys.*, 1995, **103**, 8577–8593.

54 D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark and H. J. C. Berendsen, *J. Comput. Chem.*, 2005, **26**, 1701–1718.

55 B. Morzyk-Ociepa, M. J. Nowak and D. Michalska, *Spectrochim. Acta, Part A*, 2004, **60**, 2113–2123.

56 A. Y. Ivanov, S. Stepanian, V. Karachevtsev and L. Adamowicz, *Low Temp. Phys.*, 2019, **45**, 1008–1017.

57 R. J. Fick, A. Y. Liu, F. Nussbaumer, C. Kreutz, A. Rangadurai, Y. Xu, R. D. Sommer, H. Shi, S. Scheiner and A. L. Stelling, *J. Phys. Chem. B*, 2021, **125**, 7613–7627.

58 C. S. Peng, K. C. Jones and A. Tokmakoff, *J. Am. Chem. Soc.*, 2011, **133**, 15650–15660.

59 K.-I. Miyamoto, K. Onodera, R.-T. Yamaguchi, K.-I. Ishibashi, Y. Kimura and M. Niwano, *Chem. Phys. Lett.*, 2007, **436**, 233–238.

60 A. Plumridge, S. P. Meisburger, K. Andresen and L. Pollack, *Nucleic Acids Res.*, 2017, **45**, 3932–3943.

61 K. Chakraborty, S. Mantha and S. Bandyopadhyay, *J. Chem. Phys.*, 2013, **139**, 075103.

62 N. J. Hestand and F. C. Spano, *Chem. Rev.*, 2018, **118**, 7069–7163.

19808 | *Phys. Chem. Chem. Phys.*, 2025, **27**, 19797–19808

This journal is © the Owner Societies 2025