Martini on the Rocks: Can a Coarse-Grained Force Field Model Crystals?

A. Najla Hosseini[,†] and David van der Spoel[*,,‡]

E-mail: David.vanderSpoel@icm.uu.se

Abstract

Computational chemistry is an important tool in numerous scientific disciplines, including drug discovery, chemical reaction design, materials science, and structural biology. Coarse-grained models offer a simplified representation of molecular systems, that in theory enables efficient simulations of large-scale system. Over the last decade, there has been a considerable increase in the adoption of coarse-grained models for simulations of biomolecular systems. Therefore, critical and independent evaluation of such models is warranted. Here, the properties of crystals of amyloid peptides as well as organic molecules are evaluated using the Martini coarse-grained force field. We investigate whether this force field can accurately maintain the crystal structure of amyloid peptides and organic compounds and predict melting temperatures. The peptide crystals change shape in the simulations, in most cases drastically so. Radial distribution functions show that the distance between backbone beads representing intermolecular hydrogen bonds in β-sheets increases by about 1Å in the simulation, breaking the crystals. In addition, the melting points of organic compounds are much lower in the Martini force field than in either an all-atom force field or experiment. Radial distribution functions for pyridine and phenol at 5 K show that the crystals transition into a glassy state when using Martini. Our results suggest that the Martini 3 model lacks the necessary level of specific interactions to accurately simulate peptide

crystals or organic crystals without imposing artificial restraints. We speculate that the problems are exacerbated by the use of the 12-6 Lennard-Jones potential and suggest that a different, softer, potential could prove advantageous in harnessing this model for crystal simulations.

Introduction

The structure of a protein is defined, in principle, by inter-residue hydrogen bonds providing specificity in combination with well-packed side chains of aliphatic or aromatic character, providing thermodynamic stability through the hydrophobic effect.[1] Indeed, it has been posed that the structure of a protein is to a large extent governed by its interaction with the solvent, water.[2] Amyloid peptide fibrils are an intriguing example of these structural features, containing β-sheets in one plane and tightly packed side chains perpendicular to it.[3–5] Amyloid peptide fibrils such as yeast prion protein Sup35, insulin, Alzheimer's amyloid-β, τ, and amylin, have pairs of tightly bound β-sheets known as "steric zipper" structures.[3,4,6,7] These structures run parallel to the fibril axis and play a crucial role in amyloid aggregations.[8] The stability of these peptide structures is influenced by factors such as hydrogen bonds along the fibril axis, Van der Waals interactions, electrostatic interactions, the hydrophobic effect, and π-π stacking between side chains.[9–11]

The elucidation of the molecular structure of amyloid fibrils has been a challenge due to the inherent difficulty in generating well-diffracting crystals.[12] Early structural investigations of amyloid fibrils hence focused on short polypeptides, such as GNNQQNY and KLVFFAE, as they can be assembled *in vitro*, resulting in well-ordered fibers that are suitable for analysis using techniques such as X-ray diffraction, electron microscopy, and solid-state NMR.[13] Over the years, fibril structures of amyloid proteins have been proposed based on solid-state NMR[14] and cryo-electron microscopy experiments,[15] with X-ray diffraction analysis used for shorter amyloid peptides.[3] Michaels and colleagues conducted a comprehensive investigation on the dynamics of oligomeric

species during the aggregation of the amyloid-β42 peptide, using an approach combining theory, experiment and simulation. Their findings demonstrate that, although mature amyloid fibrils stem from oligomers, the majority of amyloid-β42 oligomers dissociate into their monomeric forms rather than undergoing fibril formation and only a small subset of the oligomers undergo a transition to form fibrillar structures.[16] To complement experimental investigations, computer simulations can offer valuable insights by providing molecular models of the biological process of amyloid-aggregation. For instance, Ganguly *et al.* investigated the aggregation of Tau fragments containing the VQIVYK and VQIINK segments and their mixture, using experimental techniques and replica exchange molecular dynamics simulations. Their findings indicate that the VQIVYKPVDLSK fragment has a higher propensity for aggregation than GKVQIINKKLDL, and they suggest that heterodimer interactions may be involved in initiating Tau aggregation.[17] In another study combining experiments with simulations, Chen *et al.* determined that the aggregation-prone VQIVYK peptide, with its upstream sequence, forms metastable compact structures that influence its propensity for aggregation.[18] Nguyen and co-workers reviewed additional amyloid simulation studies .[19]

The last two decades have seen the introduction and wide-spread adoption of coarse-grained (CG) force fields for molecular simulations.[20–23] By grouping atoms together, the total number of particles is reduced and the computational cost of simulations is reduced drastically. If, for instance, four atoms are modeled as one particle the cost is reduced roughly by a factor of 4^2. This promised the possibility to study larger systems and/or use longer simulation times. Grouping of atoms can be done selectively, for instance just for the solvent immersing a molecule of interest,[24] or throughout the entire system. In this manner, degrees of freedom can be averaged out consecutively until a desired "resolution" is obtained.[25] Although the word "resolution" is used ubiquitously in the coarse-grained modeling field,[26] it is written in quotation marks here, to avoid confusion with the resolution obtained in experimental structural biology, which is a property of the data, rather than a model parameter.

In virtually all CG models, the energy function is replaced by a free energy function, which means force field parameters become temperature dependent (although this is not a large problem in all cases [23,27]) and time becomes ill-defined since the "forces" derived from such a potential function include the derivative of entropy with respect to the particle positions.[28,29] These models have been optimized to reproduce free energies starting from atomistic models, however, the reduced atomic detail leads to specific interactions being approximated. This strongly suggests energy barriers will be lowered and kinetics overestimated. Some recent reviews describe the state of the art in CG simulations.[20,23,30] Indeed, seeing that the potential energy surface actually describes free energies, it is unclear what ensemble is produced by CG "simulations", but it is likely different from the ensemble of the atomistic simulations that the coarse grained models are based on.[31]

Seeing that projects are on-going targeting modelling of entire cells or large virus particles, such as SARS-CoV-2, using CG force fields,[32–34] it is worthwhile to consider what predictive power such models could contribute with. Attempts to model dense protein solutions or even the *Escherichia coli* cytoplasm with all-atom (AA) force fields were sobering in that they demonstrated the complexity of such undertakings and shortcomings on the part of the physical models .[35–37] It is therefore questionable whether studies of large complex biological systems would fare better using less-detailed models. Indeed, a recent study of mechanoporation of biological membranes shows that the Martini 3 force field yields a reasonable structural model, but that under pressure pores are formed faster than with a corresponding AA force field.[38] Moreover, Martini 3 has also been found to have other issues, such as underestimation of the radius of gyration of intrinsically disordered proteins by ≈30% and the overestimation of protein–protein interactions when compared to small-angle X-ray scattering (SAXS) data,[39] incorrect prediction of coiled-coil dimer structures, short transmembrane peptides that were not stable inside a membrane[40] as well as failure to insert transmembrane helix dimer proteins into dodecylphosphocholine micelles.[41] It has been shown as well that Martini fails to accurately capture the enthalpy-entropy decomposition for pair correlations in

bilayers and, in addition that it leads to unusual helix-helix attractions in bilayers and exhibits unphysical fluctuations at intermediate length scales for lipid bilayers.[42]

Clearly, the development of force fields necessitates independent assessment of such models. The study of crystal lattices has been a key component in the history of molecular dynamics and such studies have proven useful in investigating the quality of underlying force fields, establishing correlations between simulated ensembles and experimental structure factors .[43,44] Hence, in this work, we investigate whether the Martini 3 model can reproduce properties of crystals consisting of amyloid peptides and organic compounds that we have studied previously using AA models.[45,46] We present CG simulations of twelve amyloid peptide crystals at cryo temperature and the temperature used for growing the crystals and evaluate the stability of the crystal. To do so, we have performed crystal peptide simulations considering the "bare" Martini3 (M3) force field, a variant including side chain restraints (M3′)[47] and a variant including both side chain restraints and restraints on the intramolecular secondary structure (M3″).[48] Two different water models were considered: regular W for both M3 and M3′, and tiny water (TW) for the M3″ force field model. In addition, we determined the melting points of organic crystals, providing a quantitative measure of the accuracy of the Martini 3 force field in comparison to both experimental data and AA models.

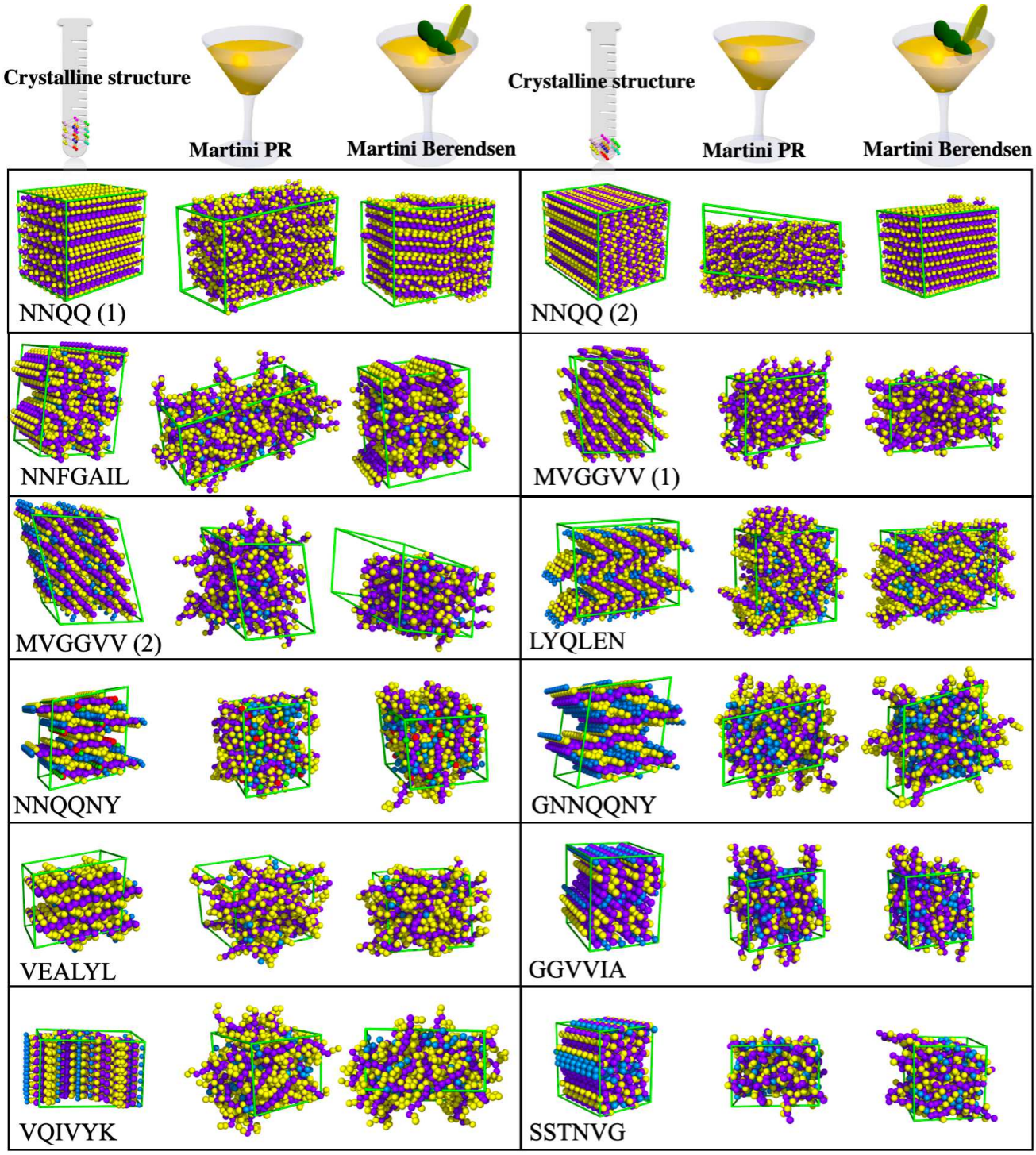

Figure 1: Crystals of twelve amyloid peptides before and after a 200 ns simulation at room temperature with a modified Martini 3 (M3′) force fields using tiny water beads using either the Parrinello-Rahman (PR) or the Berendsen (B) barostat.[49] Color coding is: purple for backbone, yellow for side chain beads, crystal water molecules cyan. Zinc and acetate beads associated with the NNQQNY peptide are colored red and green respectively. For NNQQ as well as MVGGCC two different crystal structures were used as a starting point for the simulations, see ref.[45] for details.

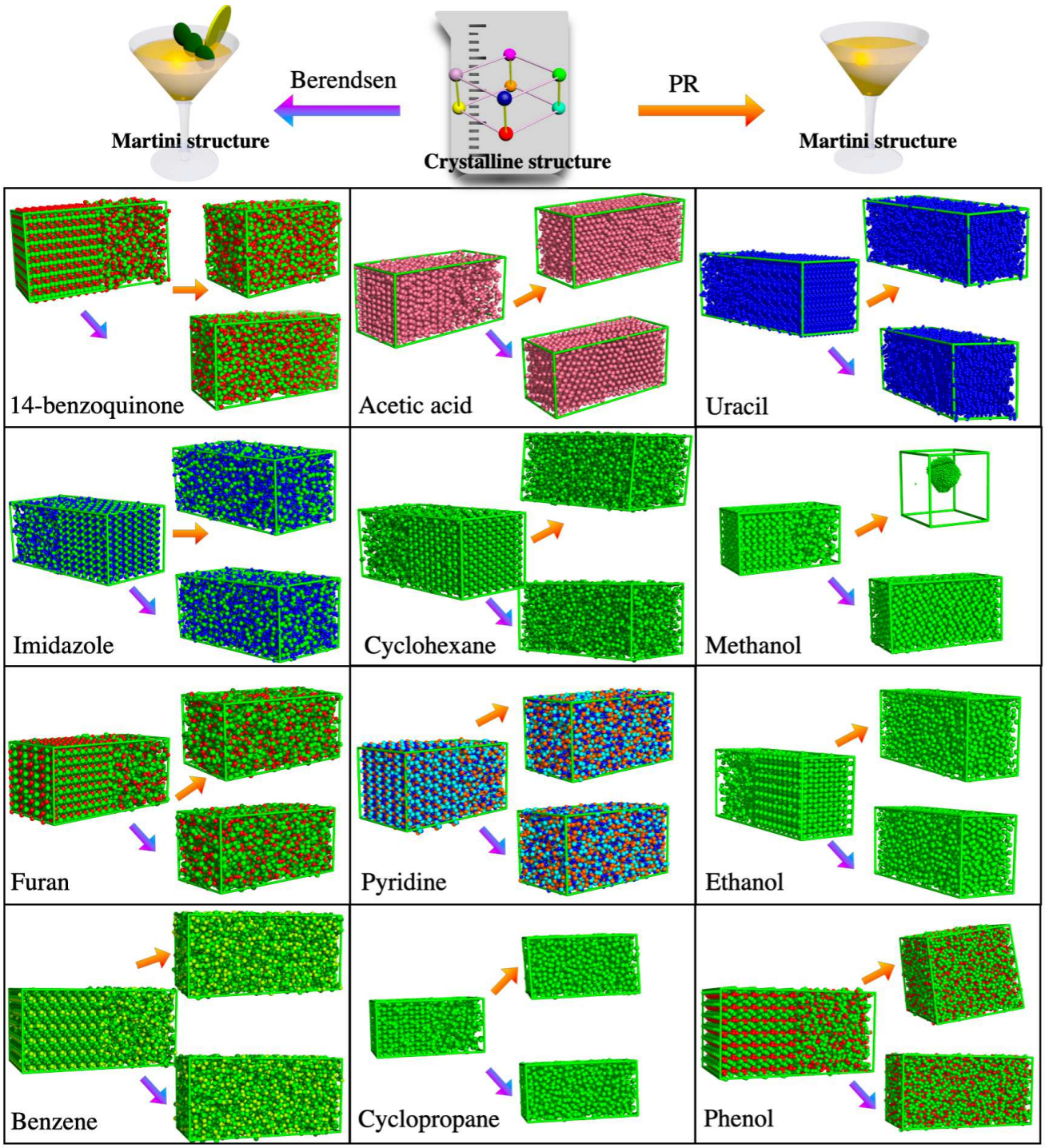

Figure 2: Solid-liquid coexistence simulation systems of twelve organic crystals using the Martini 3 force field. Structures before and after simulation at 180 K, using either the Parrinello-Rahman[50] (PR, orange arrow) or the Berendsen[49] (purple arrow) barostat.

Table 1: Mean absolute deviation from lattice parameters for the AA force fields, AMBER19SB, CHARMM36m, and OPLS-AA/M (simulations from ref.[45]) and Martini 3 force fields (this work) as a function of temperature (CT: Cryo temperature, RT: Room temperature, see methods). Barostat used is indicated as PR (Parrinello-Rahman) or B (Berendsen). M3, M3′, M3″ indicate original Martini3, Martini3 with side chain corrections, and Martini3 with side chain and secondary structure corrections.

Property	T(K)	AMBER B	CHARMM B	OPLS B	M3 PR	M3 B	M3′ PR	M3′ B	M3″ PR	M3″ B
Box edge (%)	CT	0.2	0.3	0.1	15.5	3.7	12.3	3.7	8.5	3.3
	RT	0.4	0.1	0.4	22.4	2.4	20.6	3.2	16.4	2.8
Angle (°)	CT	0.2	0.2	0.1	2.9	0.2	2.1	0.9	3.2	1.0
	RT	0.3	0.2	0.3	8.6	0.8	5.0	0.7	3.8	0.8
Volume (%)	CT	0.3	1.0	0.3	17.8	17.0	12.8	8.9	9.7	7.9
	RT	0.9	0.1	0.5	13.1	8.4	12.9	6.8	7.3	6.1

Results

The kinetic stability of crystals of twelve peptides in CG simulations using three Martini3 variants (M3, M3′, M3″) was evaluated visually (Fig 1) and by evaluating lattice parameters (Figs. S1-S12). Table 1 shows the average deviation from the experimental crystals of the box edge, angle, and volume for the peptides at temperatures corresponding to crystal growing and data collection (cryo temperature). In the case of the M3 model (Figs. S1-S4), a comparison of the lattice size of peptides during the simulations shows better stability of GNNQQNY and NNQQNY using Berendsen barostat than Parrinello-Rahman. Some peptides, like GNNQQNY and NNQQNY keep the cell edges stable over the simulation time after an initial change using the Berendsen barostat. The β and γ unit cell angles for LYQLEN and VQIVYK are both unstable at 310 K and 291 K respectively, using the Parrinello-Rahman barostat. These peptides display rapid and large changes for all cell edges. However, using the Brendsen barostat they show better stability using this model. The cell shapes for all peptides using Berendsen, except GNNQQNY and NNQQNY, change rapidly at room temperature as evidenced from both the length of the supercell edges and the lattice angles. Indeed, MVGGVV (1) shows unstable α, β, and γ angles at 298 K.

M3′ (Figs. S5-S8) and M3″ (Figs. S9-S12) demonstrate a more stable trend, even in

the presence of a significant deviation from the experimental structure, when compared to the M3 model.

Our simulations indicate that the M3″ using the Berendsen thermostat yields the most stable simulations, however, a deviation of 6-8% in the volume of the crystal with respect to the native structure is observed (Table 1). Individual box edges may change by up to 100% in the simulation. An example of such an issue is NNFGAIL at 293 K, demonstrating very rapid change using both Parrinello-Rahman (Fig. S9) and Berendsen (Fig. S11) barostats. In the Martini 3 paper[26] it is suggested the particle mesh Ewald (PME[51]) method may be needed to handle long-range electrostatic interactions for some systems. However, when comparing lattice sizes and angles at both 293 K and 100 K we find that this algorithm does not lead to more stable results either (Figs. S13-S16).

To obtain a more quantitative description of the transformation of peptides from a crystalline to a liquid state in simulations using Martini 3, we calculated the radial distribution function (RDF) for the NNQQ1 peptide, which does not contain any water molecules in its crystal. In the case of the M3 model, there is no difference between Berendsen and Parrinello-Rahman barostats, and the crystal transitions to a liquid state, eventually melting. A comparison between the RDF plots of Berendsen and Parrinello-Rahman shows better stability with Berendsen, considering the M3′ and M3″ force fields at cryo temperature. However, the shift in the calculated RDF and the decrease in RDF height in the simulations, compared to the experimental RDF at both cryo and room temperatures confirm the melting of the peptide and transformation to a liquid state using both Parrinello-Rahman and Berendsen barostats for all three Martini3 variants (Figs. S19-S22).

Compared to our previous work using three atomistic force fields, AMBER19SB, CHARMM36m, and OPLS-AA/M,[45] a much larger deviation of lattice parameters is observed for virtually all peptides using both the Parrinello-Rahman (which is the recommended barostat in Martini 3) and the Berendsen barostat (which is more resistant to fluctuations). The MVGGVV (1) crystal is destabilized and deformed in all

M3, M3′, and M3″ models in a similar manner as AMBER19SB, CHARMM36m, and OPLS-AA/M force fields at room temperature, but to a larger degree. Even at cryo temperature, very rapid changes in angles and edges occur for this peptide. In all cases, there is a large difference between results from CG and AA simulations. Most of the peptide crystal supercells in the Martini simulations are unstable, in contrast to the AA. Table 1 shows that the deviation from the crystal lattice parameters is more than an order of magnitude larger using the Martini 3 force field than in the AA force fields. The comparisons between Parrinello-Rahman and Berendsen barostats suggests better stability is obtained with the use of the Berendsen barostat in Martini simulations. However, it should be noted that even with this choice and addition of intramolecular restraints (M′, M″), stability issues still persist (Figs. S3-S12).

We now turn our attention to molecular crystals. Previous simulation work using an all-atom force field has shown that it is difficult to model crystals of organic molecules (Fig 2), since these often have weak interactions only.[46] It is well established that detailed force field models are needed to estimate melting points[52] and for organic compounds, a root mean square deviation of about 40 K was found comparing AA simulations to experimental numbers.[46] Melting points were computed using the solid-liquid coexistence method,[53] and determined from the diffusion constant as a function of temperature (Table 2, and Figs. S23- S46). There is no or very poor correlation between experimental melting points and those obtained from Martini3 simulations (Fig. S49). By analysing the temperature at which the diffusion constant is zero, even glassy states may be counted as solid. To investigate this more in detail, we calculated radial distribution function for pyridine and phenol at 5 K (Figs. S47,S48). Even at this low temperature, it can be seen that the crystals turn into a glassy state using Martini. Table 2 shows that the Martini 3 model even with this generous definition of the solid state yields melting points that are considerably lower than experimental data or those from AA simulations. The root mean square deviation from experiment is almost five times higher for the CG than the AA model. Upon more detailed analysis, it becomes evident that some of the one-bead compounds persist in the crystalline

form, although not necessarily matching the atomistic crystal structures. Meanwhile, most of the larger compounds undergo melting (Fig. 2, Table 2).

Table 2: Small molecules simulated in the crystalline state. Number of Martini particles (#P) per molecule and number of molecules (#M) in the crystals.[54] Experimental[55] and simulated melting points. Melting temperatures (T_{melt}) for the generalized Amber FF (GAFF[56]) from Schmidt *et al.*,[46] Martini from this work. The T_{melt} is defined here as the lowest temperature for which the diffusion constant deviates from zero (Figs. S23-S46). The pressure scaling algorithm used is indicated for the results from simulations. The last row gives the root mean square deviation (RMSD) in Kelvin from the experiment for both models. Barostat used is indicated as PR (Parrinello-Rahman) or B (Berendsen).

Compound	#P	#M	T_{melt} (K)			
			Exper.	GAFF	Martini	
				B	PR	B
1,4-benzoquinone	4	1152	403	388	140	140
acetic acid	1	1536	290	294	181	181
benzene	3	1536	279	250	5	80
cyclohexane	2	1260	280	302	100	120
cyclopropane	1	840	175	118	120	130
ethanol	1	1600	159	189	150	170
furan	3	1024	188	130	50	80
imidazole	3	1680	364	278	60	100
methanol	1	1024	176	186	0	240
phenol	3	1680	314	279	20	100
pyridine	3	1440	232	246	5	80
uracil	5	1400	611	748	100	200
RMSD				55	247	198

Discussion

Careful benchmarking of methods used in AA force field simulations has led to a reasonable understanding of the merits of the available models. Quantitative evaluation of accuracy is easiest for small (organic) compounds for which standard methods for force field parameterization are readily available[56–60] and for which a large body of data is available from experiments,[61,62] or from high quality quantum chemistry.[63] Examples of such benchmarks are available for the gas phase,[46,64,65] the liquid phase[66–70] as well as the solid phase.[46,70] These studies have established the root mean square errors for

predictions from atomistic force fields to be around 7 kJ/mol for enthalpies of vaporization,[66,67] around 3% for liquid densities,[66,67] 8 kJ/mol for enthalpies of sublimation, 5% for solid densities and 40 K for the melting point.[46] We recently extended this benchmarking work to a study of peptide crystals.[45] In that study, we performed simulations of twelve peptide crystals using three modern atomistic force fields and found that some of the crystals deformed during long simulations (Table 1). From this, it was concluded that simulations of organic crystals and peptides are challenging for atomistic force fields. Nevertheless, it may be possible to study amyloid peptide crystals and fibrils with suitable adaptations and refinement of atomistic force fields.[71] Time scales for fibril formation form another hurdle on the road. Sarthak and colleagues attempted to explore the effects of point mutations on protein condensates, which have implications in neurodegenerative disorders, but even using the most powerful Anton-2 computer,[72] their simulations were unable to reach the time scales of required for fibril formation *in vitro*.[73]

Buell emphasizes that the thermodynamic aspects of fibril formation are an important topic for future amyloid research.[7] Since aggregation of proteins and large peptides is slow, shorter peptides can still be useful to gain deeper understanding of the aggregation process. For instance, the potential of mean force calculations of crystal formation could help to shed light on amyloid thermodynamics,[71] and such calculations should be tractable for all-atom models. Teijlingen *et al.* have investigated the self-assembly of short peptides using different Martini force fields (3/2.1/2.1P/2.2/2.2P).[74] They highlighted the challenge in comparing results between Martini 2.1 and Martini 3 due to differences in bead types (side chain and backbone) and different Lennard-Jones terms for the same beads. One issue they reported is the overstabilization of "π-stacking" effect in Martini 2.1, yielding an energy minimum of -21.0 kJ/mol, compared to high-level quantum chemistry calculations in the gas phase that give -7.5 to -11.7 kJ/mol. On the other hand, Martini 3 produces a more accurate "π-stacking" energy of -12.7 kJ/mol. However, they obtained a stack of nanodiscs instead of the expected tubular structure, highlighting issues with correct packing of compounds.[74]

One of the key motivations for simulating crystal structures is to evaluate and test force field parameters. Indeed, crystals have long served as a crucial testing ground for development of simulation models.[43] The main point of the present work is to evaluate the packing properties in a crystalline system. To do so, simulations of peptides were performed using three Martini3 variants (see method for details). A comparison between the results obtained with the two barostats indicates that kinetic stability is higher when using the Berendsen barostat for both peptides and organic crystals. Despite this choice and the added intramolecular restraints in M′ and M″, stability issues persist. Since peptides are short, the additional stabilization due to the elastic bond corrections are more limited than for example for a protein. The corrections used span over the whole length of the peptide, which means we enforce the β-strand to remain in the same conformation. In combination with the side-chain restraints the peptides becomes almost entirely rigid. Nevertheless, the peptide crystals deform significantly (Figs. 1, S5-S8, Table 1) and molecular crystals are unstable as well (Fig. 2 Table 2). It can be seen from radial distribution functions that the packing interactions are not sufficiently strong for peptide backbones (Figs. S19-S22) and not for organic compounds either (Figs. S47,S48) to maintain stable crystals in CG simulations using Martini 3. Our findings, derived from an exploration of three force field variants, coupled with the implementation of different barostats and treatment of long-range electrostatic, underscore the limitations of the force field in maintaining accurate packing.

It is self-evident that large and complex simulation systems need a long time to equilibrate. For instance, for the small satellite tobacco necrosis virus it was found that one μs of atomistic simulation was not enough for the virus capsid to relax.[75] It therefore seems unfounded to draw conclusions on the "fast dynamics" of components of the much larger SARS-CoV-2 virus system based on a short 500 ns Martini 3 simulation,[34] not in the least because increased kinetic rates are a "feature" of coarse-grained models in general.[31] Both the mentioned dynamics of membrane-embedded proteins and the diverse binding preferences of certain lipid types to specific sites of the membrane-embedded proteins under these conditions could potentially be attributed to model

artifacts.

It is well-established that the Lennard-Jones 12-6 potential[76] has a repulsion that is too steep for atomistic simulations.[77–86] When averaging over atomic interactions, CG beads necessarily have to become larger than atoms to maintain correct densities. It would seem logical to make the beads softer to allow for more flexible interactions and it is curious that the Martini developers continued using the Lennard-Jones 12-6 potential for their models. This potential is the reason that most crystals crack at once in Martini simulations (Figs. 1,2). Moreover, through the elimination of detailed atomistic interactions that give rise to friction to motion, CG models become less viscous, making interpreting of CG dynamics challenging.[28,87] To improve Martini to be able to model crystals, the explicit introduction of electrostatic interactions could be considered.[88] Including electrostatic interactions allows additional control over the properties of molecular systems. For instance, by artificially scaling the charges of the components in a coarse-grained model of an ionic liquid (IL), Saeielli and Wang were able to completely change the phase behavior of the IL.[89] Alternatively, directional potentials could help to stabilize hydrogen bonded structures like amyloid peptides.[90,91] Within the crystal structure prediction community, work is on-going to include anisotropic atoms in order to predict the relative energies of crystal polymorphs,[92] strongly suggesting that models with simplified descriptions of the physics[44] simply lack the detail needed to model peptide or organic crystals. Indeed, Strödel argues, that more, not less detail will be needed to address the biophysics of amyloid formation.[71]

Methods

Peptides

The initial structures of amyloid peptides in CG representation for unit cells were generated using the Martinize2 Python script.[48] Subsequently, we constructed supercells using the genconf tool in GROMACS, following a similar procedure as described in

our prior work. Files are available for inspection on our GitHub repository.[54] Since the Martinize script does not generate water molecules, we used the existing atomistic supercell structures (oxygen atoms of water molecules) and incorporated them into the supercells. Then, we considered two available water models in the Martini3 force field: regular (four water molecules modeled as a single Lennard-Jones site) and tiny beads (two water molecules as a single Lennard-Jones site) for mapping water molecules. After creating supercells, we removed three waters from every closest four waters (averaging the positions of identical neighbor water molecules interacting with the peptide) for the regular model and one water from every two closest waters for the tiny model. Topologies for peptides were generated using the Martinize2 Python script for unit cells. We performed simulations with three different models: the original Martini3 with regular water beads (M3), simulations with side chain fixation with regular water beads (M3′), and simulations with restrained side chains and simultaneous enforcing of the secondary structure for the peptides with tiny water beads (M3″).[26,48]

To conduct CG simulations using GROMACS,[93] the systems were minimized with the steepest descent algorithm, followed by a constant pressure equilibration (NpT) and a slow pressure coupling time of 1000 ps. Position restraints with a force constant of 1000 kJ/mol nm^2 were applied to the backbone of peptides during equilibration and released for production. The temperatures were coupled to a V-rescale thermostat,[94] which was set at crystal growing temperature (see ref.[45] for details) and cryo temperatures (100K) respectively, with a coupling constant of 0.1 ps. The pressure was coupled to an anisotropic Parrinello-Rahman barostat[95] with a coupling constant of 1000 ps, and the reference pressure set to 1 bar. We performed all simulations using the Berendsen barostat[49] with a 1000 ps coupling constant as well, as it dampens fluctuations by design.

Long-range electrostatic interactions were treated using a reaction-field[96] with a relative dielectric constant of 15[26] or using PME.[51] A neutral N-terminus was used for the positively charged peptide VQIVYK, using the Martinize2 script. A neutral C-terminus was used for peptides with negatively charged LYQLEN and VEALYL side

chains. NNQQNY contains Zinc and acetate, hence, we used a neutral N-terminus for this peptide. A Zinc ion is not available in the Martini force field, so we used Calcium beads, but changed the mass to 65.38. To update the neighbor list, the Verlet neighbor search algorithm was used.[97] Production trajectories were generated for 200 ns with cut-offs of 1.1 nm for Lennard Jones potential, no corrections were made for long-range Van der Waals interactions.

Organic compounds

A database of small molecule topology files for Martini 3 is provided by Alessandri *et al.*[98] There are twelve shared compounds between the published Martini compounds and the dataset published by Schmidt *et al.* (Table 2). Initial crystal structures were produced from the all-atom crystals that are available from github.[54] In brief, relevant atoms in the AA structure were mapped to the corresponding CG particles, and new PDB files generated using the GROMACS editconf tool.[93] These were then subject to energy minimization in order to obtain correct CG geometries. It has been suggested that the Martini 3.0 force field, in contrast to previous versions, can reproduce temperature dependent properties of compounds[27] and therefore we used the same temperature series as used by Schmidt *et al.*[46] Temperatures were maintained using the canonical rescaling algorithm[94] with a coupling time τ_T of 0.1 ps. Since all organic compounds considered in this work are neutral there were no Coulomb interactions.

Then, 200 ns constant pressure simulations were performed using Parrinello-Rahman and Berendsen barostats[49,95] with the coupling time τ_P of 100 ps. To compute the melting temperature, we used the liquid/solid direct coexistence approach by conducting a series of NpT simulations at different temperatures as initially suggested in ref.[53] For all the simulations of organic compounds, the diffusion constant was computed for the last 5 ns of the simulations from the mean square displacement. Based on the temperature dependence of the diffusion constant (Figs. S23-S46) the melting temperature was estimated (Table 2). It should be noted that this is rather generous definition since the structures change from the crystal state even at low temperature (see e.g. Figs. S47

and S48).

Analysis of data and availability

Analysis of simulations was done with the GROMACS software suite.[93] Angles were calculated using Python (NumPy, and Pandas).[99,100] Molecular images were produced using the PyMOL software.[101] Matplotlib was used for generating all plots.[102] The scripts are available from the github repository.[54]

Supplementary information

Supplementary figures of diffusion coefficients, lattice parameters, and RDFs are available.

Acknowledgments

This research was supported financially by the project *AI4Research* at Uppsala University, Sweden and by the Swedish Research Council (grant 2020-05059). Funding from eSSENCE - The e-Science Collaboration (Uppsala-Lund-Umeå, Sweden) is gratefully acknowledged. The computations were enabled by resources provided by the National Academic Infrastructure for Supercomputing in Sweden (NAISS) at the national supercomputing center, Linköping, partially funded by the Swedish Research Council through grant agreement no. 2022-06725.

References

(1) Schulz, G. E.; Schirmer, R. H. *Principles of protein structure*; Springer-Verlag: Berlin, Heidelberg, New York, 1998.

(2) Bellissent-Funel, M.-C.; Hassanali, A.; Havenith, M.; Henchman, R.; Pohl, P.;

Sterpone, F.; van der Spoel, D.; Xu, Y.; Garcia, A. E. Water Determines the Structure and Dynamics of Proteins. *Chem. Rev.* **2016**, *116*, 7673–7697.

(3) Nelson, R.; Sawaya, M. R.; Balbirnie, M.; Madsen, A. Ø.; Riekel, C.; Grothe, R.; Eisenberg, D. Structure of the cross-β spine of amyloid-like fibrils. *Nature* **2005**, *435*, 773–778.

(4) Ke, P. C.; Zhou, R.; Serpell, L. C.; Riek, R.; Knowles, T. P.; Lashuel, H. A.; Gazit, E.; Hamley, I. W.; Davis, T. P.; Fändrich, M., et al. Half a century of amyloids: past, present and future. *Chem. Soc. Rev.* **2020**, *49*, 5473–5509.

(5) Chiti, F.; Dobson, C. M. Protein misfolding, amyloid formation, and human disease: a summary of progress over the last decade. *Annu. Rev. Biochem.* **2017**, *86*, 27–68.

(6) Colletier, J.-P.; Laganowsky, A.; Landau, M.; Zhao, M.; Soriaga, A. B.; Goldschmidt, L.; Flot, D.; Cascio, D.; Sawaya, M. R.; Eisenberg, D. Molecular basis for amyloid-β polymorphism. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 16938–16943.

(7) Buell, A. K. Stability matters, too–the thermodynamics of amyloid fibril formation. *Chem. Sci.* **2022**, *13*, 10177–10192.

(8) Sawaya, M. R.; Sambashivan, S.; Nelson, R.; Ivanova, M. I.; Sievers, S. A.; Apostol, M. I.; Thompson, M. J.; Balbirnie, M.; Wiltzius, J. J.; McFarlane, H. T., et al. Atomic structures of amyloid cross-β spines reveal varied steric zippers. *Nature* **2007**, *447*, 453–457.

(9) Seidler, P. M.; Boyer, D. R.; Murray, K. A.; Yang, T. P.; Bentzel, M.; Sawaya, M. R.; Rosenberg, G.; Cascio, D.; Williams, C. K.; Newell, K. L., et al. Structure-based inhibitors halt prion-like seeding by Alzheimer's disease–and tauopathy–derived brain tissue samples. *J. Biol. Chem.* **2019**, *294*, 16451–16464.

(10) Van Gils, J. H. M.; Van Dijk, E.; Peduzzo, A.; Hofmann, A.; Vettore, N.; Schützmann, M. P.; Groth, G.; Mouhib, H.; Otzen, D. E.; Buell, A. K., et al. The hydrophobic effect characterises the thermodynamic signature of amyloid fibril growth. *PLoS Comput. Biol.* **2020**, *16*, e1007767.

(11) Vettore, N.; Buell, A. K. Thermodynamics of amyloid fibril formation from chemical depolymerization. *Phys. Chem. Chem. Phys.* **2019**, *21*, 26184–26194.

(12) Makin, O. S.; Atkins, E.; Sikorski, P.; Johansson, J.; Serpell, L. C. Molecular basis for amyloid fibril formation and stability. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 315–320.

(13) Diaz-Avalos, R.; Long, C.; Fontano, E.; Balbirnie, M.; Grothe, R.; Eisenberg, D.; Caspar, D. L. Cross-beta order and diversity in nanocrystals of an amyloid-forming peptide. *J. Mol. Biol.* **2003**, *330*, 1165–1175.

(14) Wasmer, C.; Lange, A.; van Melckebeke, H.; Siemer, A. B.; Riel, R.; Meier, B. H. Amyloid Fibrils of the HET-s(218–289) Prion Form a β Solenoid with a Triangular Hydrophobic Core. *Science* **2008**, *319*, 1523–1526.

(15) Fitzpatrick, A. W.; Falcon, B.; He, S.; Murzin, A. G.; Murshudov, G.; Garringer, H. J.; Crowther, R. A.; Ghetti, B.; Goedert, M.; Scheres, S. H. Cryo-EM structures of tau filaments from Alzheimer's disease. *Nature* **2017**, *547*, 185–190.

(16) Michaels, T. C.; Šarić, A.; Curk, S.; Bernfur, K.; Arosio, P.; Meisl, G.; Dear, A. J.; Cohen, S. I.; Dobson, C. M.; Vendruscolo, M., et al. Dynamics of oligomer populations formed during the aggregation of Alzheimer's Aβ42 peptide. *Nat. Chem.* **2020**, *12*, 445–451.

(17) Ganguly, P.; Do, T. D.; Larini, L.; LaPointe, N. E.; Sercel, A. J.; Shade, M. F.; Feinstein, S. C.; Bowers, M. T.; Shea, J.-E. Tau assembly: the dominant role of PHF6 (VQIVYK) in microtubule binding region repeat R3. *J. Phys. Chem. B* **2015**, *119*, 4582–4593.

(18) Chen, D.; Drombosky, K. W.; Hou, Z.; Sari, L.; Kashmer, O. M.; Ryder, B. D.; Perez, V. A.; Woodard, D. R.; Lin, M. M.; Diamond, M. I., et al. Tau local structure shields an amyloid-forming motif and controls aggregation propensity. *Nat. Commun.* **2019**, *10*, 2493.

(19) Nguyen, P. H. et al. Amyloid Oligomers: A Joint Experimental/Computational Perspective on Alzheimer's Disease, Parkinson's Disease, Type II Diabetes, and Amyotrophic Lateral Sclerosis. *Chem. Rev.* **2021**, *121*, 2545–2647.

(20) Kmiecik, S.; Gront, D.; Kolinski, M.; Wieteska, L.; Dawid, A. E.; Kolinski, A. Coarse-grained protein models and their applications. *Chem. Rev.* **2016**, *116*, 7898–7936.

(21) Pezeshkian, W.; König, M.; Wassenaar, T. A.; Marrink, S. J. Backmapping triangulated surfaces to coarse-grained membrane models. *Nat. Commun.* **2020**, *11*, 2296.

(22) Tejedor, A. R.; Sanchez-Burgos, I.; Estevez-Espinosa, M.; Garaizar, A.; Collepardo-Guevara, R.; Ramirez, J.; Espinosa, J. R. Protein structural transitions critically transform the network connectivity and viscoelasticity of RNA-binding protein condensates but RNA can prevent it. *Nat. Commun.* **2022**, *13*, 5717.

(23) Jin, J.; Pak, A. J.; Durumeric, A. E.; Loose, T. D.; Voth, G. A. Bottom-up coarse-graining: Principles and perspectives. *J. Chem. Theory Comput.* **2022**, *18*, 5759–5791.

(24) van der Spoel, D.; Zhang, J.; Zhang, H. Quantitative predictions from molecular simulations using explicit or implicit interactions. *Wiley Interdiscip. Rev.-Comput. Mol. Sci.* **2022**, *12*, e1560.

(25) Berendsen, H. J. C. *Simulating the physical world*; Cambridge University Press: Cambridge, 2007.

(26) Souza, P. C. T. et al. Martini 3: a general purpose force field for coarse-grained molecular dynamics. *Nature Meth.* **2021**, *18*, 382–388.

(27) Molza, A.-E.; Gao, P.; Jakpou, J.; Nicolas, J.; Tsapis, N.; Ha-Duong, T. Simulations of the Upper Critical Solution Temperature Behavior of Poly(ornithine-co-citrulline)s Using Martini-Based Coarse-Grained Force Fields. *J. Chem. Theory Comput.* **2021**, *17*, 4499–4511.

(28) de Oliveira dos Santos Soares, R.; Bortot, L. O.; van der Spoel, D.; Caliri, A. Membrane vesiculation induced by proteins of the dengue virus envelope studied by molecular dynamics simulations. *J. Phys. Condens. Matt.* **2017**, *29*, 504002.

(29) Lu, L.; Voth, G. A. The multiscale coarse-graining method. VII. Free energy decomposition of coarse-grained effective potentials. *J. Chem. Phys.* **2011**, *134*.

(30) Pak, A. J.; Voth, G. A. Advances in coarse-grained modeling of macromolecular complexes. *Curr. Opin. Struct. Biol.* **2018**, *52*, 119–126.

(31) Brini, E.; Algaer, E. A.; Ganguly, P.; Li, C.; Rodríguez-Ropero, F.; van der Vegt, N. F. A. Systematic coarse-graining methods for soft matter simulations — a review. *Soft Matter* **2013**, *9*, 2108–2119.

(32) Stevens, J. A.; Grünewald, F.; van Tilburg, P. A. M.; König, M.; Gilbert, B. R.; Brier, T. A.; Thornburg, Z. R.; Luthey-Schulten, Z.; Marrink, S. J. Molecular dynamics simulation of an entire cell. *Frontiers Chem.* **2023**, *11*.

(33) Marrink, S. J.; Monticelli, L.; Melo, M. N.; Alessandri, R.; Tieleman, D. P.; Souza, P. C. T. Two decades of Martini: Better beads, broader scope. *WIREs Comp. Molec. Sci.* **2023**, *13*, e1620.

(34) Wang, D.; Li, J.; Wang, L.; Cao, Y.; Kang, B.; Meng, X.; Li, S.; Song, C. Toward atomistic models of intact severe acute respiratory syndrome coronavirus 2 via Martini coarse-grained molecular dynamics simulations. *Quant Biol* **2023**,

(35) Feig, M.; Yu, I.; Wang, P.-h.; Nawrocki, G.; Sugita, Y. Crowding in Cellular Environments at an Atomistic Level from Computer Simulations. *J. Phys. Chem. B.* **2017**, *121*, 8009–8025.

(36) Bashardanesh, Z.; Elf, J.; Zhang, H.; van der Spoel, D. Rotational and translational diffusion of proteins under crowding conditions. *ACS Omega* **2019**, *4*, 20654–20664.

(37) Bortot, L. O.; Bashardanesh, Z.; van der Spoel, D. Making Soup: Preparing and Validating Molecular Simulations of the Bacterial Cytoplasm. *J. Chem. Inf. Model.* **2019**, *60*, 322–331.

(38) Vo, A. T. N.; Murphy, M. A.; Phan, P. K.; Prabhu, R. K.; Stone, T. W. Effect of Force Field Resolution on Membrane Mechanical Response and Mechanoporation Damage under Deformation Simulations. *Molec. Biotechnol.* **2023**, *0*, 0, doi.org/10.1007/s12033-023-00726-x.

(39) Thomasen, F. E.; Pesce, F.; Roesgaard, M. A.; Tesei, G.; Lindorff-Larsen, K. Improving Martini 3 for disordered and multidomain proteins. *J. Chem. Theory Comput.* **2022**, *18*, 2033–2041.

(40) Spinti, J. K.; Nunes, F. N.; Melo, M. N. Room for improvement in the initial Martini 3 parameterization of peptide interactions. *Chem. Phys. Lett.* **2023**, *819*, 140436.

(41) Claveras Cabezudo, A.; Athanasiou, C.; Tsengenes, A.; Wade, R. C. Scaling Protein–Water Interactions in the Martini 3 Coarse-Grained Force Field to Simulate Transmembrane Helix Dimers in Different Lipid Environments. *J. Chem. Theory Comput.* **2023**,

(42) Jarin, Z.; Newhouse, J.; Voth, G. A. Coarse-grained force fields from the perspective of statistical mechanics: Better understanding of the origins of a MARTINI hangover. *J. Chem. Theory Comput.* **2021**, *17*, 1170–1180.

(43) Dauber-Osguthorpe, P.; Hagler, A. T. Biomolecular force fields: where have we been, where are we now, where do we need to go and how do we get there? *J. Comput. Aid. Mol. Des.* **2019**, *33*, 133–203.

(44) Hagler, A. Force field development phase II: Relaxation of physics-based criteria... or inclusion of more rigorous physics into the representation of molecular energetics. *J. Comput. Aided Mol. Des.* **2019**, *33*, 205–264.

(45) Hosseini, A. N.; van der Spoel, D. Simulations of Amyloid-forming Peptides in the Crystal State. *Prot. J.* **2023**, *42*, 192–204.

(46) Schmidt, L.; van der Spoel, D.; Walz, M.-M. Probing phase transitions in organic crystals using atomistic MD simulations. *ACS Phys Chem Au* **2023**, *3*, 84–93.

(47) Periole, X.; Cavalli, M.; Marrink, S.-J.; Ceruso, M. A. Combining an elastic network with a coarse-grained molecular force field: structure, dynamics, and intermolecular recognition. *J. Chem. Theory Comput.* **2009**, *5*, 2531–2543.

(48) Kroon, P. C.; Grunewald, F.; Barnoud, J.; van Tilburg, M.; Souza, P. C. T.; Wassenaar, T. A.; Marrink, Martinize2 and Vermouth: Unified Framework for Topology Generation. **2023**,

(49) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.

(50) Parrinello, M.; Rahman, A. Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J. Appl. Phys.* **1981**, *52*, 7182–7190.

(51) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An N-log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.

(52) Walz, M.-M.; van der Spoel, D. Systematically improved melting point prediction: a detailed physical simulation model is required. *Chem. Comm.* **2019**, *55*, 12044–12047.

(53) Ladd, A.; Woodcock, L. Triple-point coexistence properties of the Lennard-Jones system. *Chem. Phys. Lett.* **1977**, *51*, 155–159.

(54) Schmidt, L.; A. Najla Hosseini,; Walz, M.-M.; ; van der Spoel, D. Molecular Dynamics Benchmark. https://github.com/dspoel/MDBenchmark, 2023; Date accessed: 2023-11-06.

(55) Lemmon, E. W.; McLinden, M. O.; Friend, D. G. In *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*; Mallard, W. G., Linstrom, P. J., Eds.; National Institute of Standards and Technology: Gaithersburg MD, 20899 (http://webbook.nist.gov), 1998.

(56) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General AMBER Force Field. *J. Comput. Chem.* **2004**, *25*, 1157–1174.

(57) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell Jr., A. D. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **2010**, *31*, 671–690.

(58) Jo, S.; Kim, T.; Iyer, V. G.; Im, W. CHARMM-GUI: a web-based graphical user interface for CHARMM. *J. Comput. Chem.* **2008**, *29*, 1859–1865.

(59) Lee, J.; Hitzenberger, M.; Rieger, M.; Kern, N. R.; Zacharias, M.; Im, W. CHARMM-GUI supports the Amber force fields. *J. Chem. Phys.* **2020**, *153*, 035103.

(60) Yabe, M.; Mori, K.; Ueda, K.; Takeda, M. Development of PolyParGen software to facilitate the determination of molecular dynamics simulation parameters for polymers. *J. Computer Chem., Japan-Int. Ed.* **2019**, *5*, 2018–0034.

(61) Lide, D. R. *"CRC Handbook of Chemistry and Physics 90th edition"*; CRC Press: Cleveland, Ohio, 2009.

(62) Rowley, R. L.; Wilding, W. V.; Oscarson, J. L.; Yang, Y.; Giles, N. F. *"Data Compilation of Pure Chemical Properties (Design Institute for Physical Properties"*; American Institute for Chemical Engineering: New York, 2012.

(63) Kříž, K.; Schmidt, L.; Andersson, A. T.; Walz, M.-M.; van der Spoel, D. An Imbalance in the Force: The Need for Standardised Benchmarks for Molecular Simulation. *J. Chem. Inf. Model.* **2023**, *63*, 412–431.

(64) van der Spoel, D.; Ghahremanpour, M. M.; Lemkul, J. Small Molecule Thermochemistry: A Tool For Empirical Force Field Development. *J. Phys. Chem. A* **2018**, *122*, 8982–8988.

(65) Lim, V. T.; Hahn, D. F.; Tresadern, G.; Bayly, C. I.; Mobley, D. L. Benchmark assessment of molecular geometries and energies from small molecule force fields. *F1000Res.* **2020**, *9*, 1390.

(66) Caleman, C.; van Maaren, P. J.; Hong, M.; Hub, J. S.; Costa, L. T.; van der Spoel, D. Force Field Benchmark of Organic Liquids: Density, Enthalpy of Vaporization, Heat Capacities, Surface Tension, Compressibility, Expansion Coefficient and Dielectric Constant. *J. Chem. Theory Comput.* **2012**, *8*, 61–74.

(67) Fischer, N. M.; van Maaren, P. J.; Ditz, J. C.; Yildirim, A.; van der Spoel, D. Properties of liquids in Molecular Dynamics Simulations with explicit long-range Lennard-Jones interactions. *J. Chem. Theory Comput.* **2015**, *11*, 2938–2944.

(68) Zhang, J.; Tuguldur, B.; van der Spoel, D. Force field benchmark II: Gibbs energy of solvation of organic molecules in organic liquids. *J. Chem. Inf. Model.* **2015**, *55*, 1192–1201.

(69) Zhang, J.; Tuguldur, B.; van der Spoel, D. Correction to Force field benchmark

II: Gibbs energy of solvation of organic molecules in organic liquids. *J. Chem. Inf. Model.* **2016**, *56*, 819–820.

(70) Senftle, T. P.; Hong, S.; Islam, M. M.; Kylasa, S. B.; Zheng, Y.; Shin, Y. K.; Junkermeier, C.; Engel-Herbert, R.; Janik, M. J.; Aktulga, H. M., et al. The ReaxFF reactive force-field: development, applications and future directions. *Npj Comput. Mater.* **2016**, *2*, 1–14.

(71) Strödel, B. Amyloid aggregation simulations: challenges, advances and perspectives. *Curr. Opin. Struct. Biol.* **2021**, *67*, 145–152.

(72) Shaw, D. E.; Grossman, J.; Bank, J. A.; Batson, B.; Butts, J. A.; Chao, J. C.; Deneroff, M. M.; Dror, R. O.; Even, A.; Fenton, C. H., et al. Anton 2: raising the bar for performance and programmability in a special-purpose molecular dynamics supercomputer. SC'14: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis. 2014; pp 41–53.

(73) Sarthak, K.; Winogradoff, D.; Ge, Y.; Myong, S.; Aksimentiev, A. Benchmarking Molecular Dynamics Force Fields for All-Atom Simulations of Biological Condensates. *J. Chem. Theory Comput.* **2023**, *19*, 3721–3740.

(74) van Teijlingen, A.; Smith, M. C.; Tuttle, T. Short Peptide Self-Assembly in the Martini Coarse-Grain Force Field Family. *Acc. Chem. Res.* **2023**, *56*, 644–654.

(75) Larsson, D. S. D.; Liljas, L.; van der Spoel, D. Virus capsid dissolution studied by microsecond molecular dynamics simulations. *PLoS Comput. Biol.* **2012**, *8*, e1002502.

(76) Jones, J. E. On the determination of molecular fields. -II. From the equation of state of a gas. *Proc. Royal Soc. Lond. Ser. A* **1924**, *106*, 463–477.

(77) Buckingham, R. A. The Classical Equation of State of Gaseous Helium, Neon and Argon. *Proc. R. Soc. London Ser. A* **1938**, *168*, 264–283.

(78) Wu, G.; Sadus, R. Molecular simulation of the high-pressure phase equilibria of binary atomic fluid mixtures using the exponential-6 intermolecular potential. *Fluid Phase Equil.* **2000**, *170*, 269–284.

(79) Pirani, F.; Brizi, S.; Roncaratti, L. F.; Casavecchia, P.; Cappelletti, D.; Vecchiocattivi, F. Beyond the Lennard-Jones model: a simple and accurate potential function probed by high resolution scattering data useful for molecular dynamics simulations. *Phys. Chem. Chem. Phys.* **2008**, *10*, 5489–5503.

(80) Wensink, E. J. W.; Hoffmann, A. C.; van Maaren, P. J.; van der Spoel, D. Dynamic Properties of Water/Alcohol Mixtures Studied by Computer Simulation. *J. Chem. Phys.* **2003**, *119*, 7308–7317.

(81) Wang, L.-P.; Chen, J.; Voorhis, T. V. Systematic Parametrization of Polarizable Force Fields from Quantum Chemistry Data. *J. Chem. Theory Comput.* **2013**, *9*, 452–460.

(82) Walz, M. M.; Ghahremanpour, M. M.; van Maaren, P. J.; van der Spoel, D. Phase-transferable force field for alkali halides. *J. Chem. Theory Comput.* **2018**, *14*, 5933–5948.

(83) Burrows, S. A.; Korotkin, I.; Smoukov, S. K.; Boek, E.; Karabasov, S. Benchmarking of Molecular Dynamics Force Fields for Solid-Liquid and Solid-Solid Phase Transitions in Alkanes. *J. Phys. Chem. B.* **2021**, *125*, 5145–5159.

(84) Bernhardt, M. P.; Nagata, Y.; van der Vegt, N. F. A. Where Lennard-Jones Potentials Fail: Iterative Optimization of Ion-Water Pair Potentials Based on Ab Initio Molecular Dynamics Data. *J. Phys. Chem. Lett.* **2022**, *13*, 3712–3717.

(85) Kříž, K.; van Maaren, P. J.; van der Spoel, D. Submitted to J. Chem. Theory Comput.

(86) Horton, J. T.; Boothroyd, S.; Behara, P. K.; Mobley, D. L.; Cole, D. J. A

transferable double exponential potential for condensed phase simulations of small molecules. *Digital Discovery* **2023**, *2*, 1178–1187.

(87) Noid, W. G. Perspective: Coarse-grained models for biomolecular systems. *J. Chem. Phys.* **2013**, *139*.

(88) Heinemann, T.; Palczynski, K.; Dzubiella, J.; Klapp, S. H. L. Coarse-grained electrostatic interactions of coronene: Towards the crystalline phase. *J. Chem. Phys.* **2015**, *143*, 174110.

(89) Saielli, G.; Wang, Y. Role of the electrostatic interactions in the stabilization of ionic liquid crystals: Insights from coarse-grained MD simulations of an imidazolium model. *J. Phys. Chem. B* **2016**, *120*, 9152–9160.

(90) Yu, Z.; Lau, D. Development of a coarse-grained α-chitin model on the basis of Martini forcefield. *J. Mol. Model* **2015**, *21*, 128.

(91) Wolpert, E. H.; Jelfs, K. E. Coarse-grained modelling to predict the packing of porous organic cages. *Chem. Sci.* **2022**, *13*, 13588–13599.

(92) Aina, A. A.; Misquitta, A. J.; Price, S. L. A non-empirical intermolecular forcefield for trinitrobenzene and its application in crystal structure prediction. *J. Chem. Phys.* **2021**, *154*, 094123.

(93) Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; Hess, B.; Lindahl, E. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **2013**, *29*, 845–854.

(94) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007**, *126*, 014101.

(95) Parrinello, M.; Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **1981**, *52*, 7182–7190.

(96) Tironi, I. G.; Sperb, R.; Smith, P. E.; van Gunsteren, W. F. A generalized reaction field method for molecular dynamics simulations. *J. Chem. Phys.* **1995**, *102*, 5451–5459.

(97) Páll, S.; Hess, B. A flexible algorithm for calculating pair interactions on SIMD architectures. *Comput. Phys. Commun.* **2013**, *184*, 2641–2650.

(98) Alessandri, R.; Barnoud, J.; Gertsen, A. S.; Patmanidis, I.; de Vries, A. H.; Souza, P. C. T.; Marrink, S. J. Martini 3 Coarse-Grained Force Field: Small Molecules. *Adv. Theor. Simul.* **2022**, *5*, 2100391.

(99) Van Der Walt, S.; Colbert, S. C.; Varoquaux, G. The NumPy array: a structure for efficient numerical computation. *Comput. Sci. Eng.* **2011**, *13*, 22–30.

(100) McKinney, W. *Python for data analysis: Data wrangling with Pandas, NumPy, and IPython*; " O'Reilly Media, Inc.", 2012.

(101) DeLano, W. L., et al. Pymol: An open-source molecular graphics tool. *CCP4 Newsl. Protein Crystallogr* **2002**, *40*, 82–92.

(102) Hunter, J. D. Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* **2007**, *9*, 90–95.

TOC Graphic

Martini on the Rocks: Can a Coarse-Grained Force Field Model Crystals?

A. Najla Hosseini[,†] and David van der Spoel[*,,‡]

E-mail: David.vanderSpoel@icm.uu.se

List of Figures

S1 Deviation of lattice size of the supercell in %, over the NpT and production runs for all amyloid peptides using Parrinello-Rahman barostat (M3). S5

S2 Deviation of angles of the supercell from experimental crystal structure over the NpT and production runs for all peptides using Parrinello-Rahman barostat (M3). S6

S3 Deviation of lattice size of the supercell in %, over the NpT and production runs for all amyloid peptides using Berendsen barostat (M3). S7

S4 Deviation of angles of the supercell from experimental crystal structure over the NpT and production runs for all peptides using Berendsen barostat (M3). S8

S5 Deviation of lattice size of the supercell in %, over the NpT and production runs for all amyloid peptides considering modifications (M3′) using Parrinello-Rahman barostat. S9

S6 Deviation of angles of the supercell from experimental crystal structure over the NpT and production runs for all peptides considering modifications (M3′) using Parrinello-Rahman barostat. S10

S7 Deviation of lattice size of the supercell in %, over the NpT and production runs for all amyloid peptides considering modifications (M3′) using Berendsen barostat. S11

S8 Deviation of angles of the supercell from experimental crystal structure over the NpT and production runs for all peptides considering modifications (M3′) using Berendsen barostat. S12

S9 Deviation of lattice size of the supercell in %, over the NpT and production runs for all amyloid peptides considering modifications (M3″) using Parrinello-Rahman barostat. S13

S10 Deviation of angles of the supercell from experimental crystal structure over the NpT and production runs for all peptides considering modifications (M3″) using Parrinello-Rahman barostat. S14

S11 Deviation of lattice size of the supercell in %, over the NpT and production runs for all amyloid peptides considering modifications (M3″) using Berendsen barostat. S15

S12 Deviation of angles of the supercell from experimental crystal structure over the NpT and production runs for all peptides considering modifications (M3″) using Berendsen barostat. S16

S13 Deviation of lattice size of the NNFGAIL in %, over the NpT and production runs using reaction-field and particle mesh Ewald (PME) for the (M3″) force field using Berendsen barostat at 293 K. S17

S14 Deviation of lattice size of the NNFGAIL in %, over the NpT and production runs using reaction-field and PME for the (M3″) force field using Berendsen barostat at 100 K. S17

S15 Deviation of angles of the NNFGAIL from experimental crystal structure over the NpT and production runs using reaction-field and PME considering (M3″) force field model with Berendsen barostat at 293 K. S18

S16 Deviation of angles of the NNFGAIL from experimental crystal structure over the NpT and production runs using reaction-field and PME considering (M3″) force field model with Berendsen barostat at 100 K. S18
S17 Radial distribution function (RDF) of the backbone of NNQQ1 at cryo and room temperatures for both experiment and simulation (M3) using the Parrinello barostat. S19
S18 Radial distribution function (RDF) of the backbone of NNQQ1 at cryo and room temperatures for both experiment and simulation (M3) using the Berendsen barostat. S19
S19 Radial distribution function (RDF) of the backbone of NNQQ1 at cryo and room temperatures for both experiment and simulation using the Parrinello barostat (M3′). S20
S20 Radial distribution function (RDF) of the backbone of NNQQ1 at cryo and room temperatures for both experiment and simulation using the Berendsen barostat (M3′). S20
S21 Radial distribution function (RDF) of the backbone of NNQQ1 at cryo and room temperatures for both experiment and simulation using the Parrinello barostat (M3″). S21
S22 Radial distribution function (RDF) of the backbone of NNQQ1 at cryo and room temperatures for both experiment and simulation using the Berendsen barostat (M3″). S21
S23 Diffusion as a function of temperature using Parrinello-Rahman barostat. . . S22
S24 Diffusion as a function of temperature using Berendsen barostat. S22
S25 Diffusion as a function of temperature using Parrinello-Rahman barostat. . . S23
S26 Diffusion as a function of temperature using Berendsen barostat. S23
S27 Diffusion as a function of temperature. S24
S28 Diffusion as a function of temperature. S24

S29 Diffusion as a function of temperature. S25
S30 Diffusion as a function of temperature. S25
S31 Diffusion as a function of temperature. S26
S32 Diffusion as a function of temperature. S26
S33 Diffusion as a function of temperature. S27
S34 Diffusion as a function of temperature. S27
S35 Diffusion as a function of temperature. S28
S36 Diffusion as a function of temperature. S28
S37 Diffusion as a function of temperature. S29
S38 Diffusion as a function of temperature. S29
S39 Diffusion as a function of temperature. S29
S40 Diffusion as a function of temperature. S30
S41 Diffusion as a function of temperature. S30
S42 Diffusion as a function of temperature. S31
S43 Diffusion as a function of temperature. S31
S44 Diffusion as a function of temperature. S31
S45 Diffusion as a function of temperature. S32
S46 Diffusion as a function of temperature. S32
S47 Radial distribution function (RDF) of pyridine at 5 K for both experiment and simulation using the Berendsen barostat. S33
S48 Radial distribution function (RDF) of phenol at 5 K for both experiment and simulation using the Berendsen barostat. S33
S49 Correlation between experimental melting temperature and simulated melting temperatures for 12 organic compounds (see table 2 in main text). S34

Supplementary figures

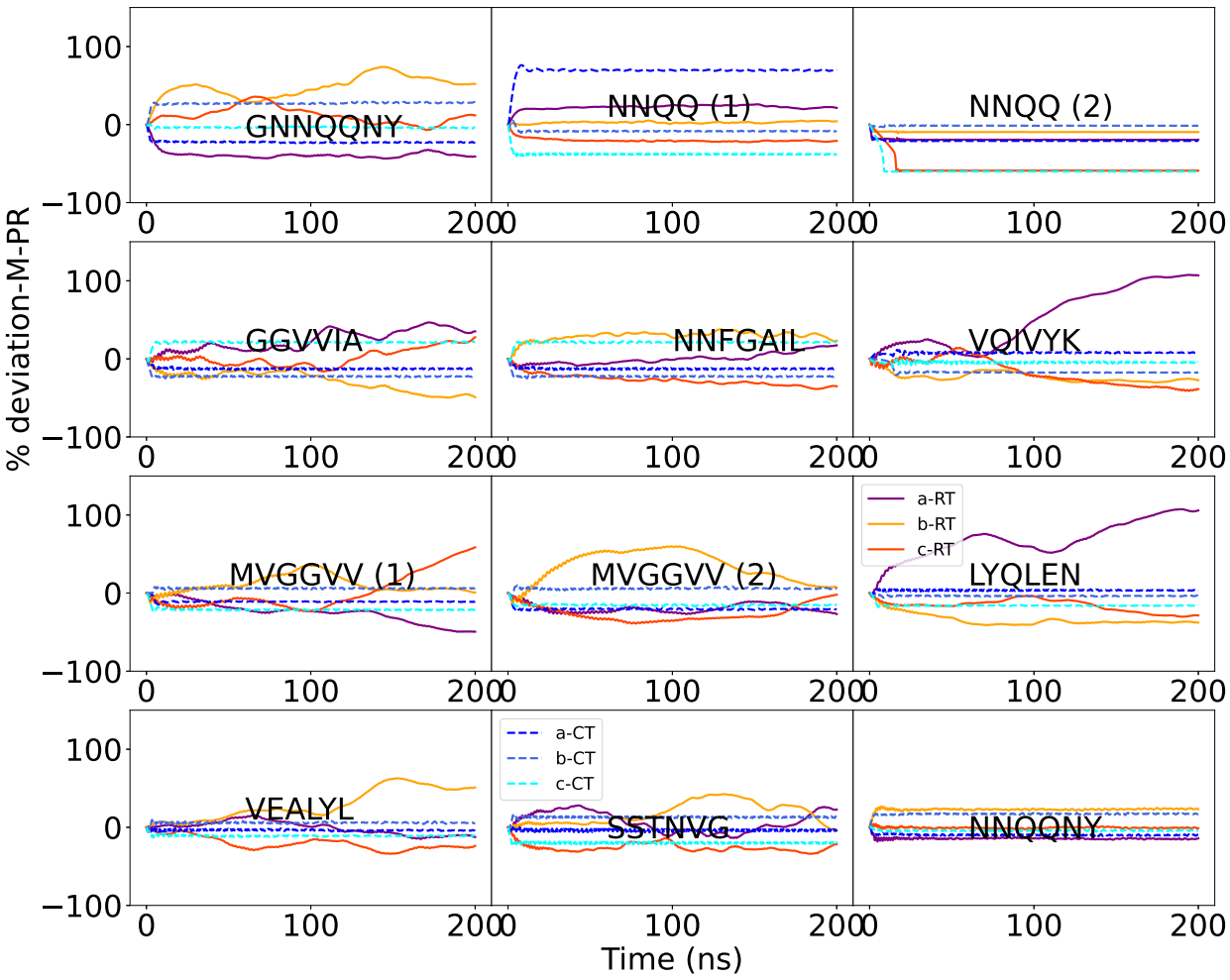

Figure S1: Deviation of lattice size of the supercell in %, over the NpT and production runs for all amyloid peptides using Parrinello-Rahman barostat (M3).

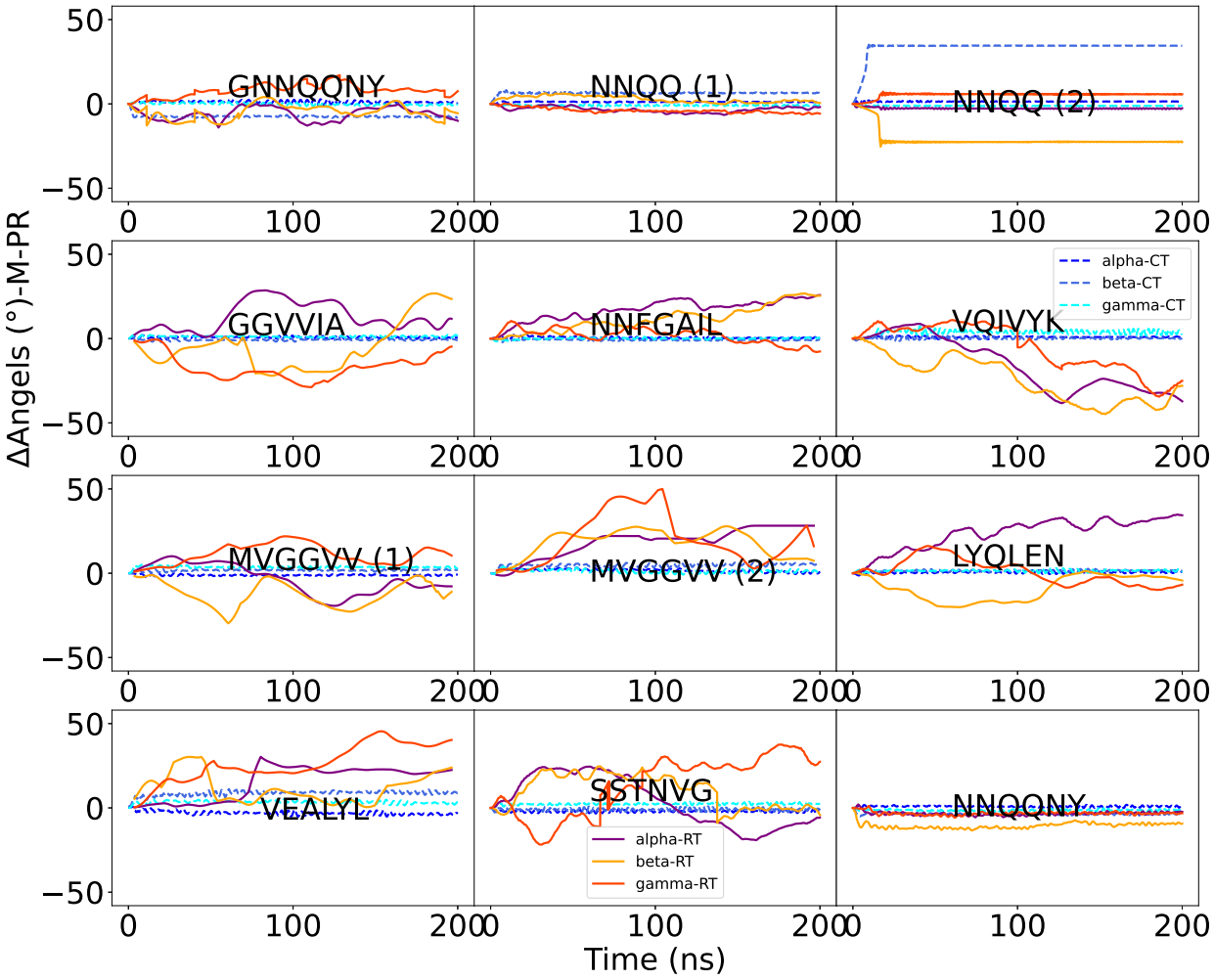

Figure S2: Deviation of angles of the supercell from experimental crystal structure over the NpT and production runs for all peptides using Parrinello-Rahman barostat (M3).

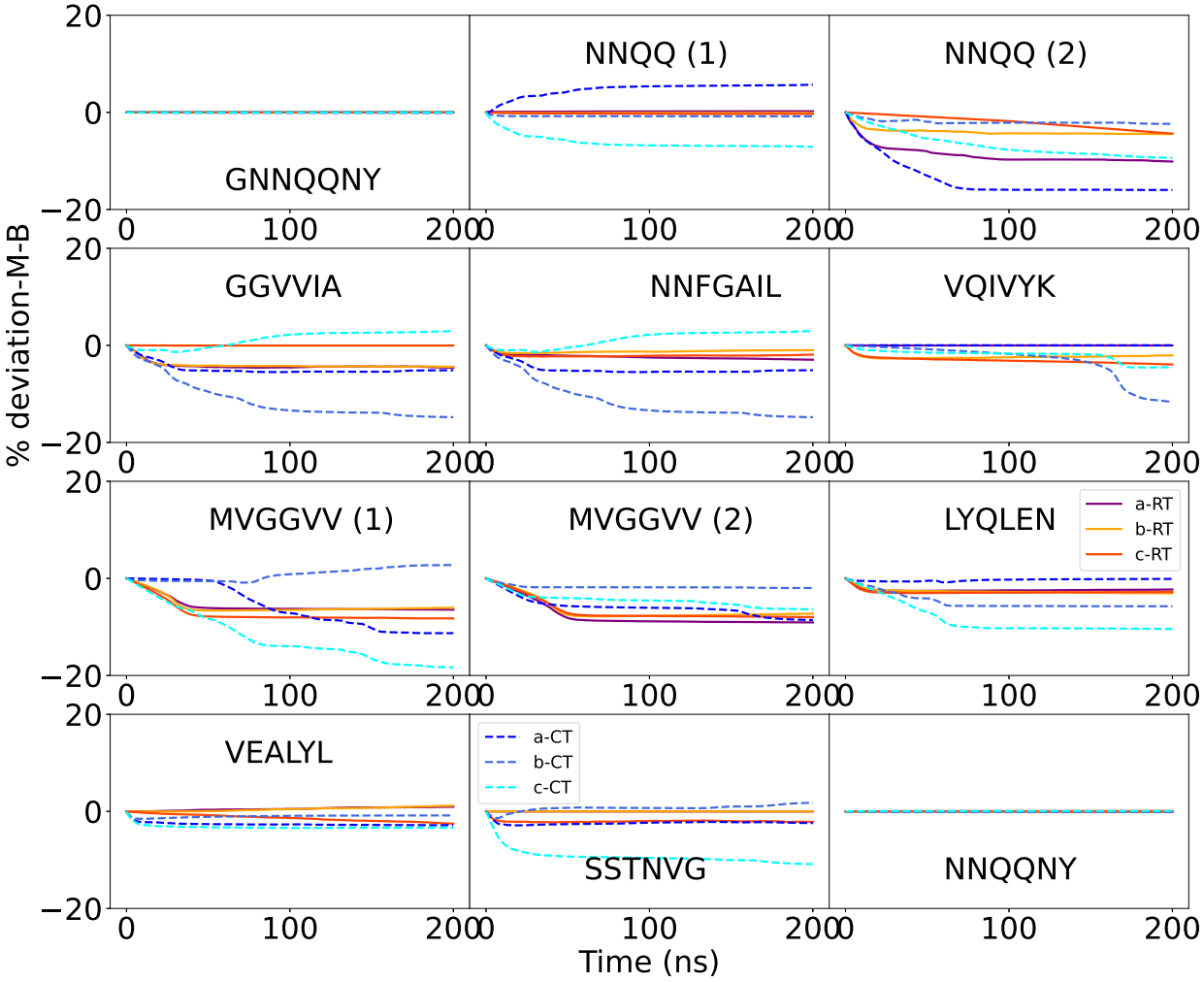

Figure S3: Deviation of lattice size of the supercell in %, over the NpT and production runs for all amyloid peptides using Berendsen barostat (M3).

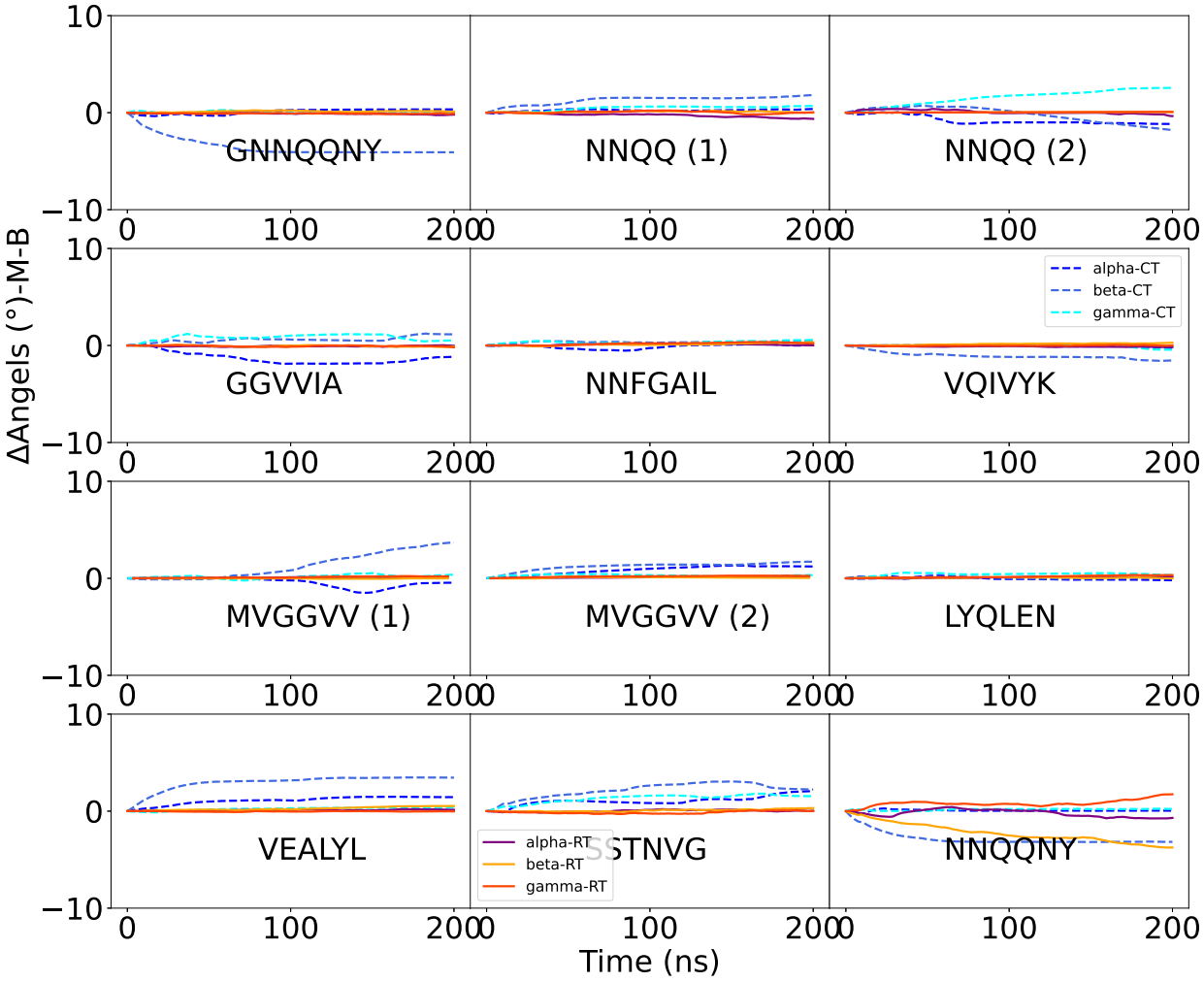

Figure S4: Deviation of angles of the supercell from experimental crystal structure over the NpT and production runs for all peptides using Berendsen barostat (M3).

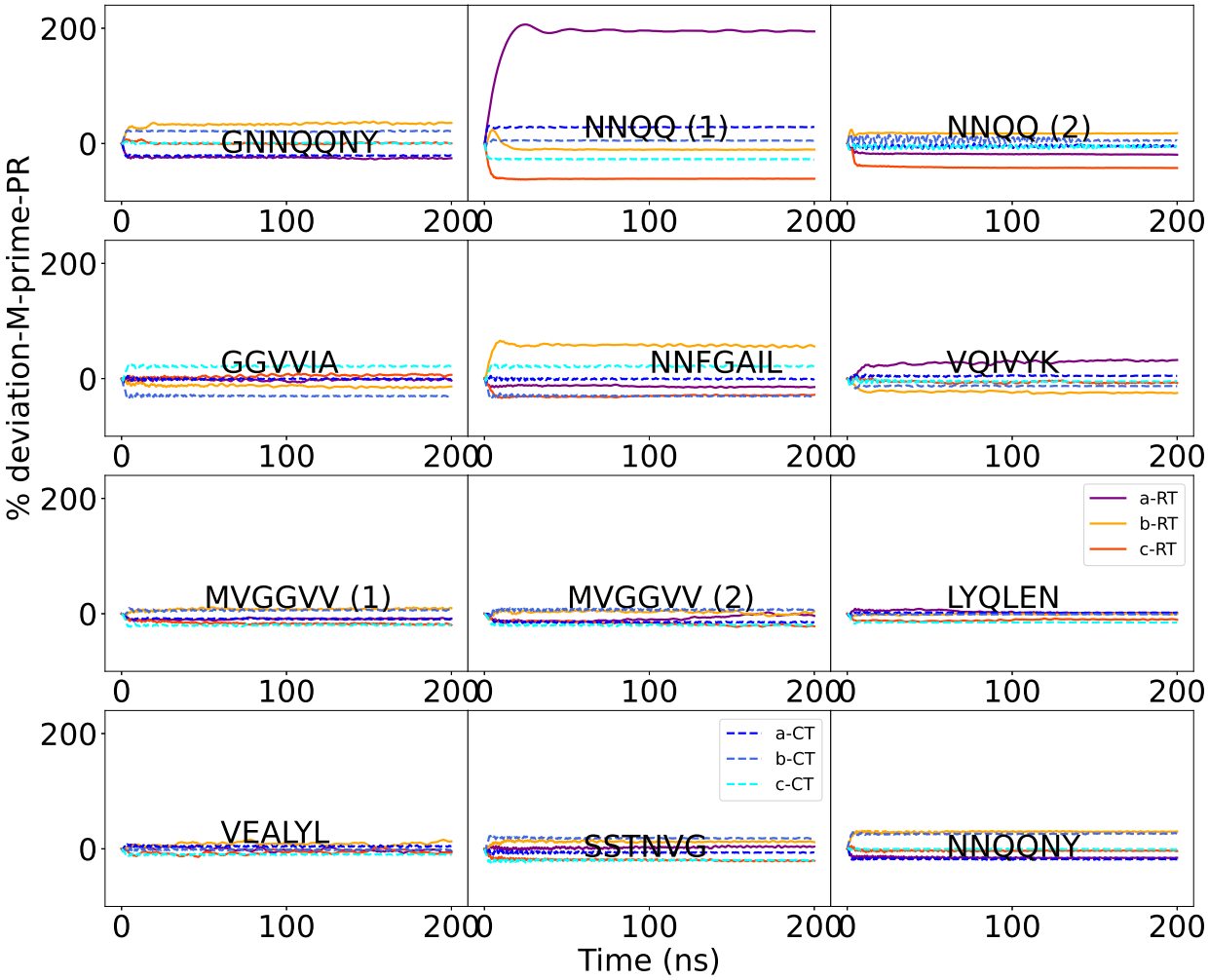

Figure S5: Deviation of lattice size of the supercell in %, over the NpT and production runs for all amyloid peptides considering modifications (M3′) using Parrinello-Rahman barostat.

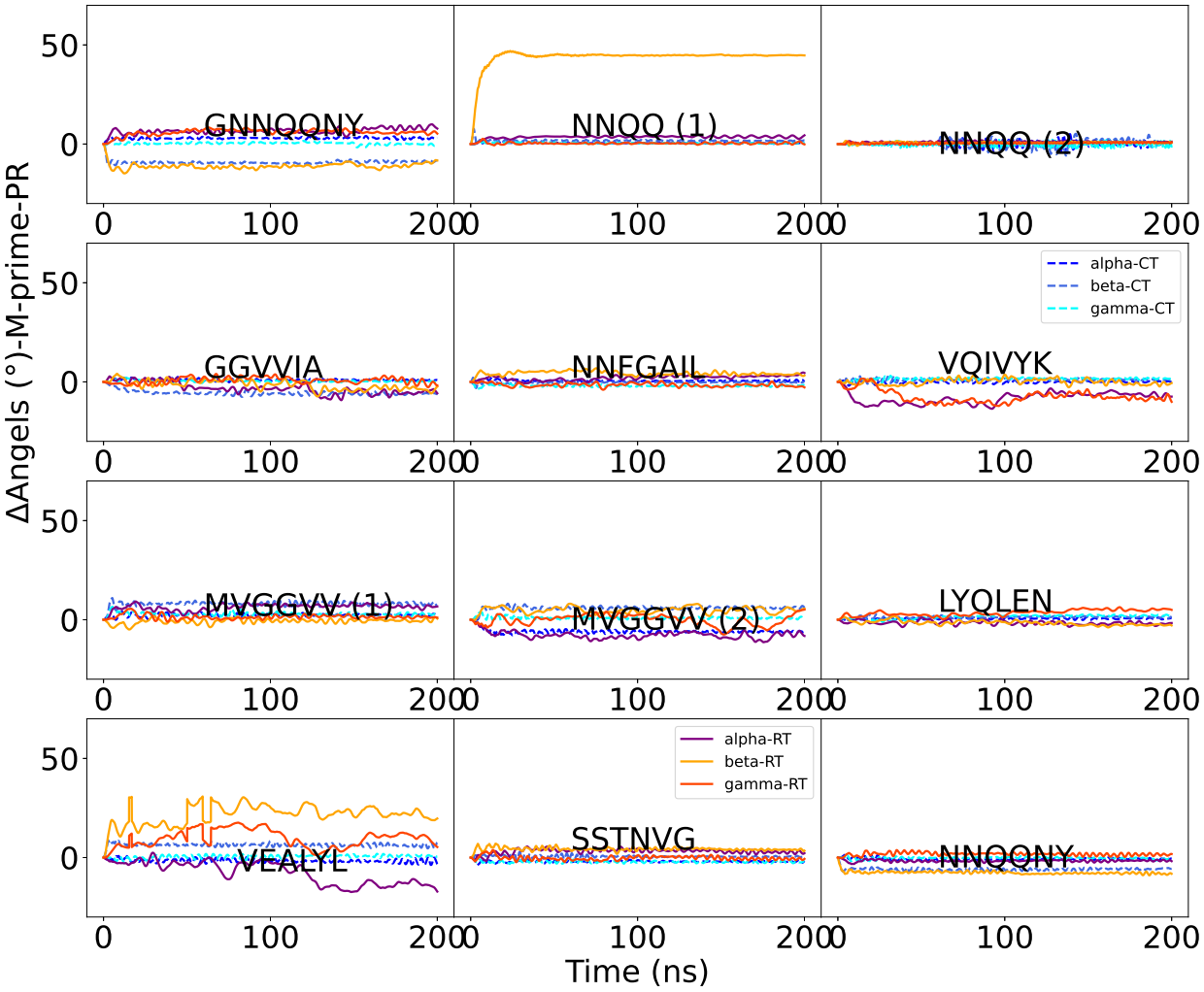

Figure S6: Deviation of angles of the supercell from experimental crystal structure over the NpT and production runs for all peptides considering modifications (M3′) using Parrinello-Rahman barostat.

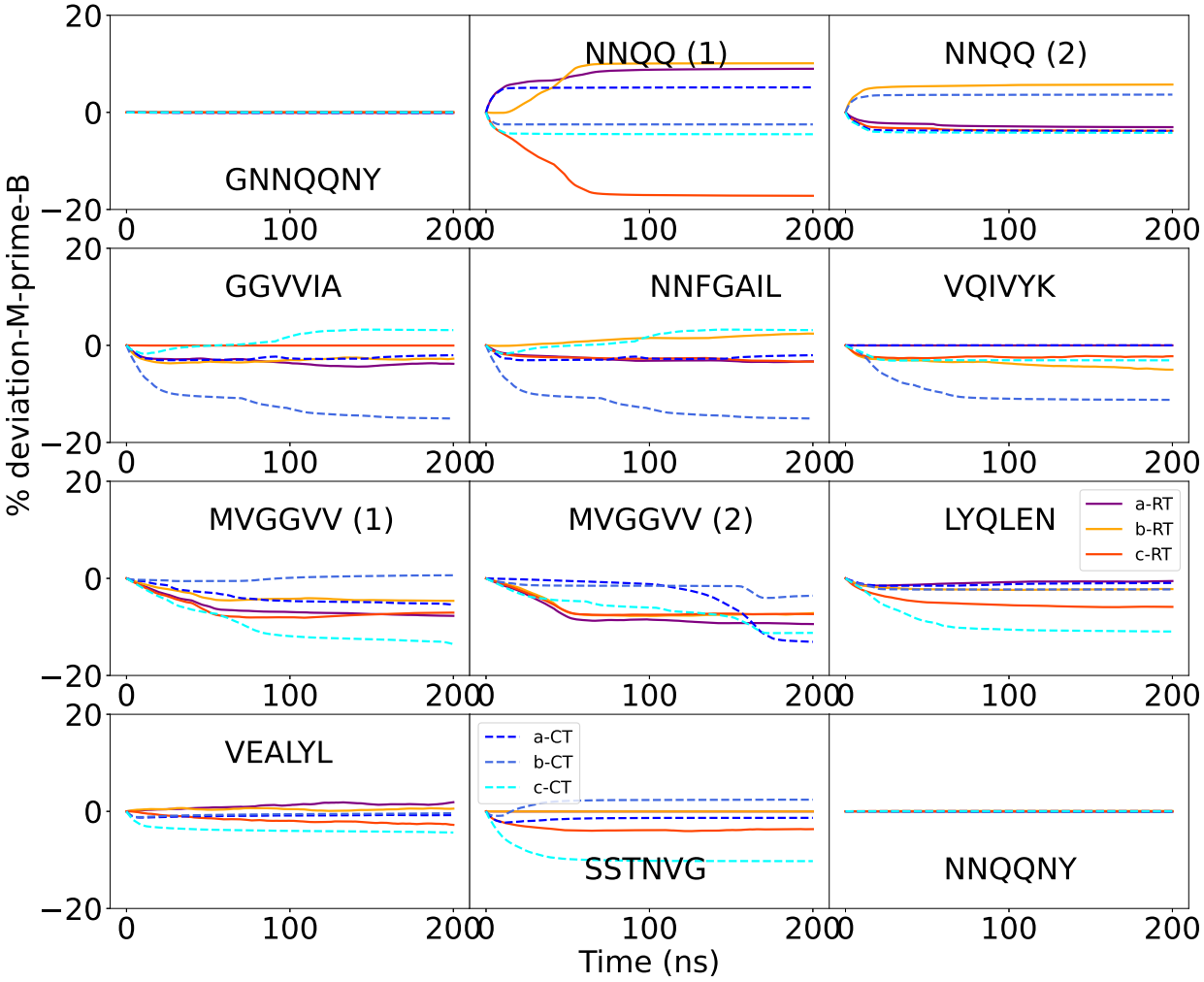

Figure S7: Deviation of lattice size of the supercell in %, over the NpT and production runs for all amyloid peptides considering modifications (M3′) using Berendsen barostat.

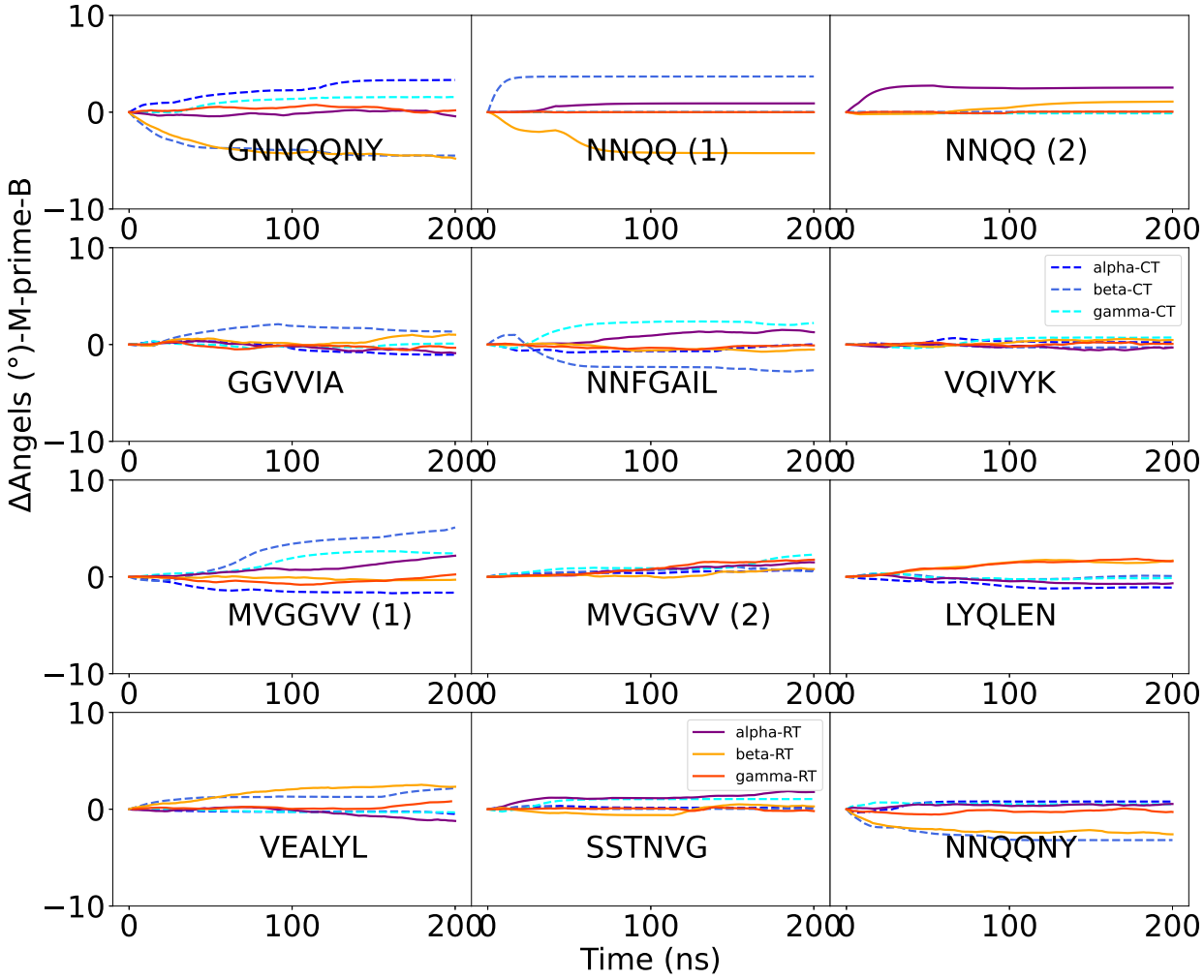

Figure S8: Deviation of angles of the supercell from experimental crystal structure over the NpT and production runs for all peptides considering modifications (M3′) using Berendsen barostat.

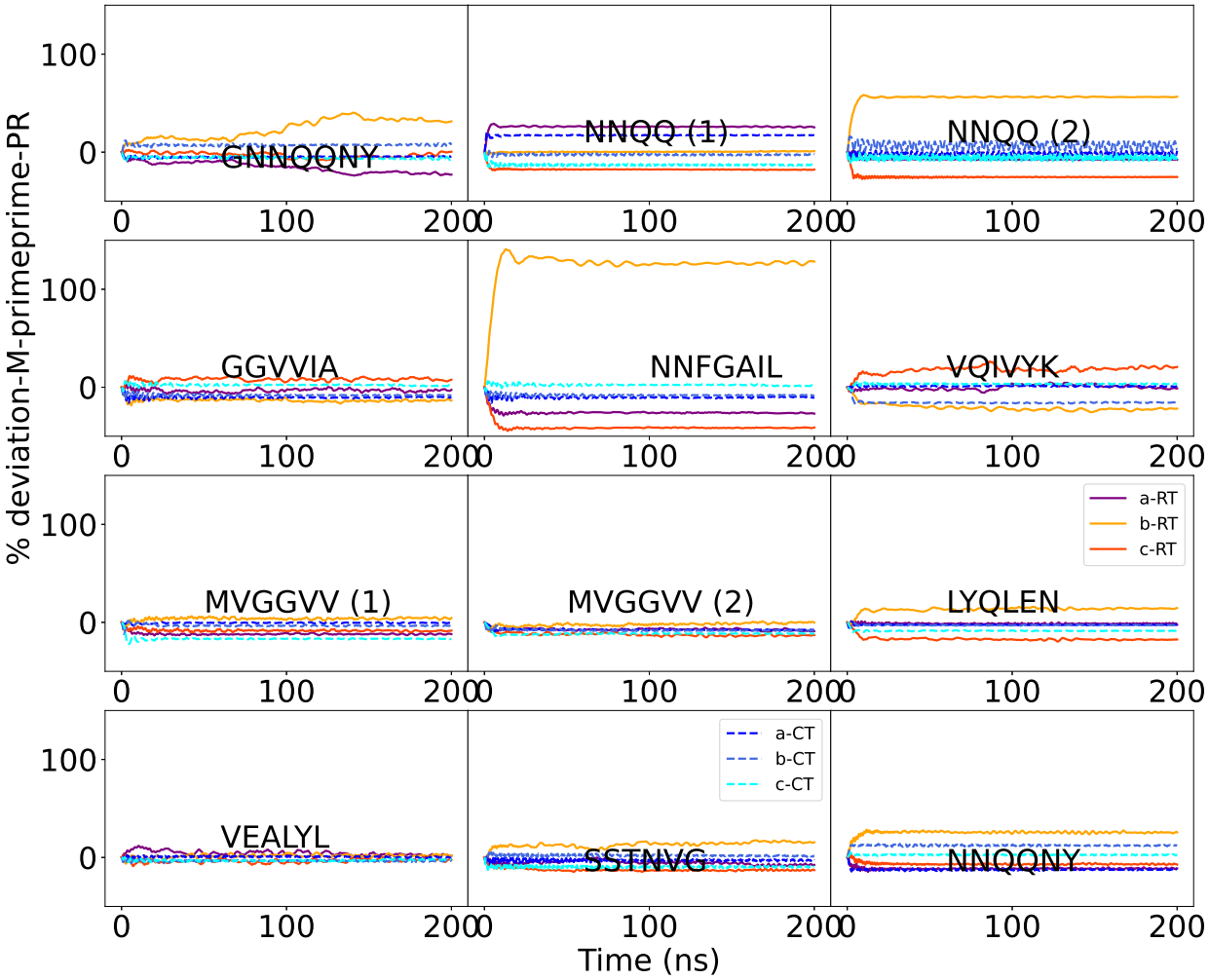

Figure S9: Deviation of lattice size of the supercell in %, over the NpT and production runs for all amyloid peptides considering modifications (M3″) using Parrinello-Rahman barostat.

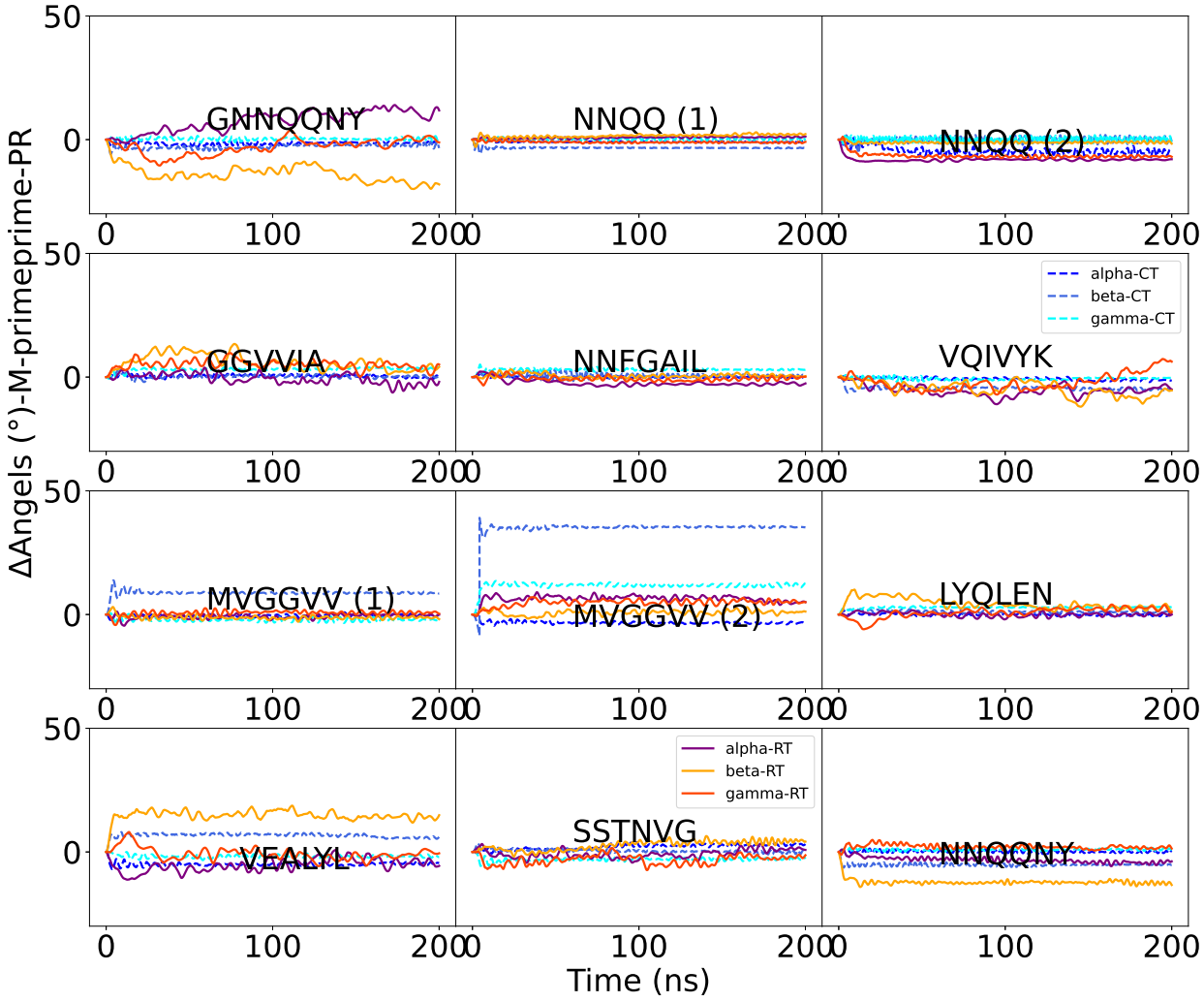

Figure S10: Deviation of angles of the supercell from experimental crystal structure over the NpT and production runs for all peptides considering modifications (M3″) using Parrinello-Rahman barostat.

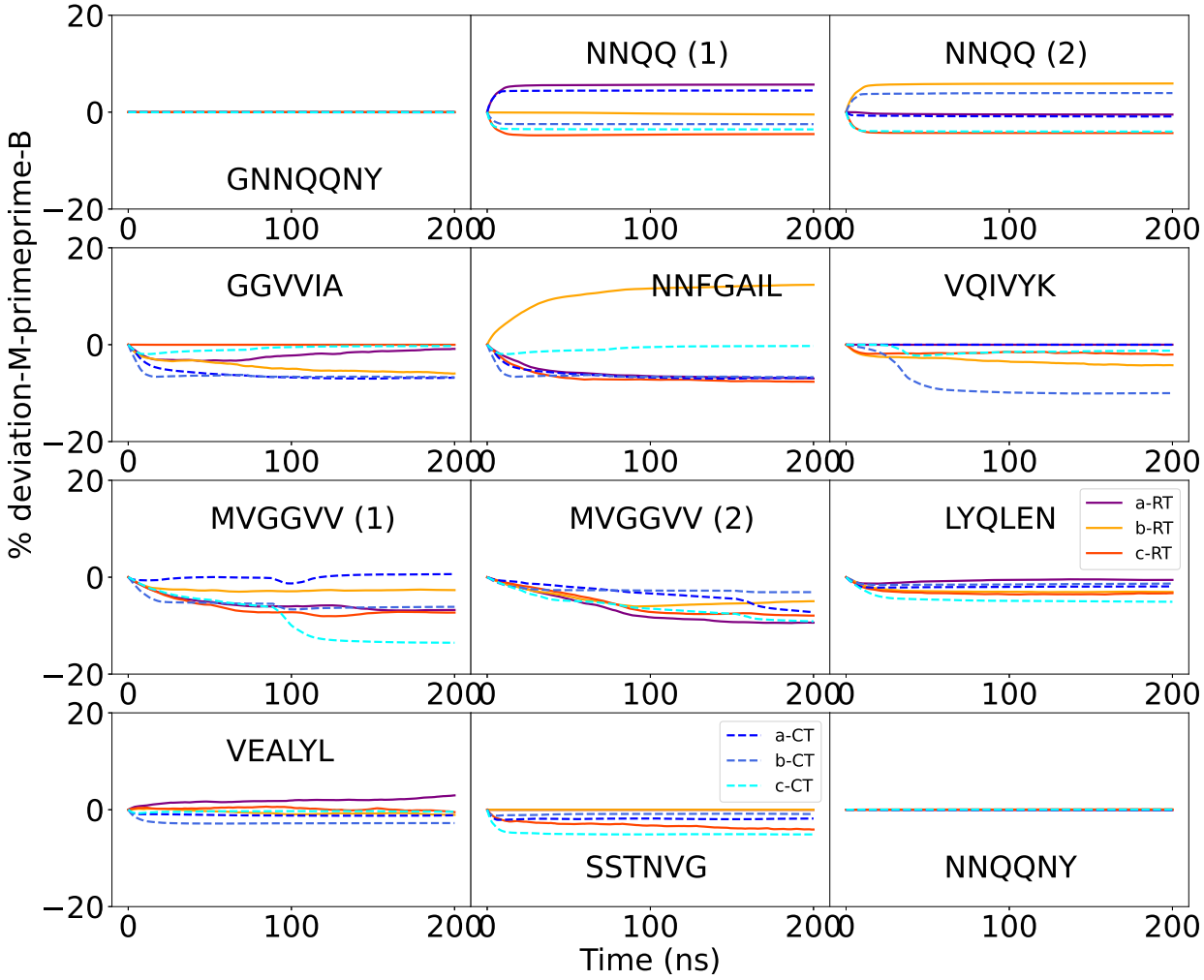

Figure S11: Deviation of lattice size of the supercell in %, over the NpT and production runs for all amyloid peptides considering modifications (M3″) using Berendsen barostat.

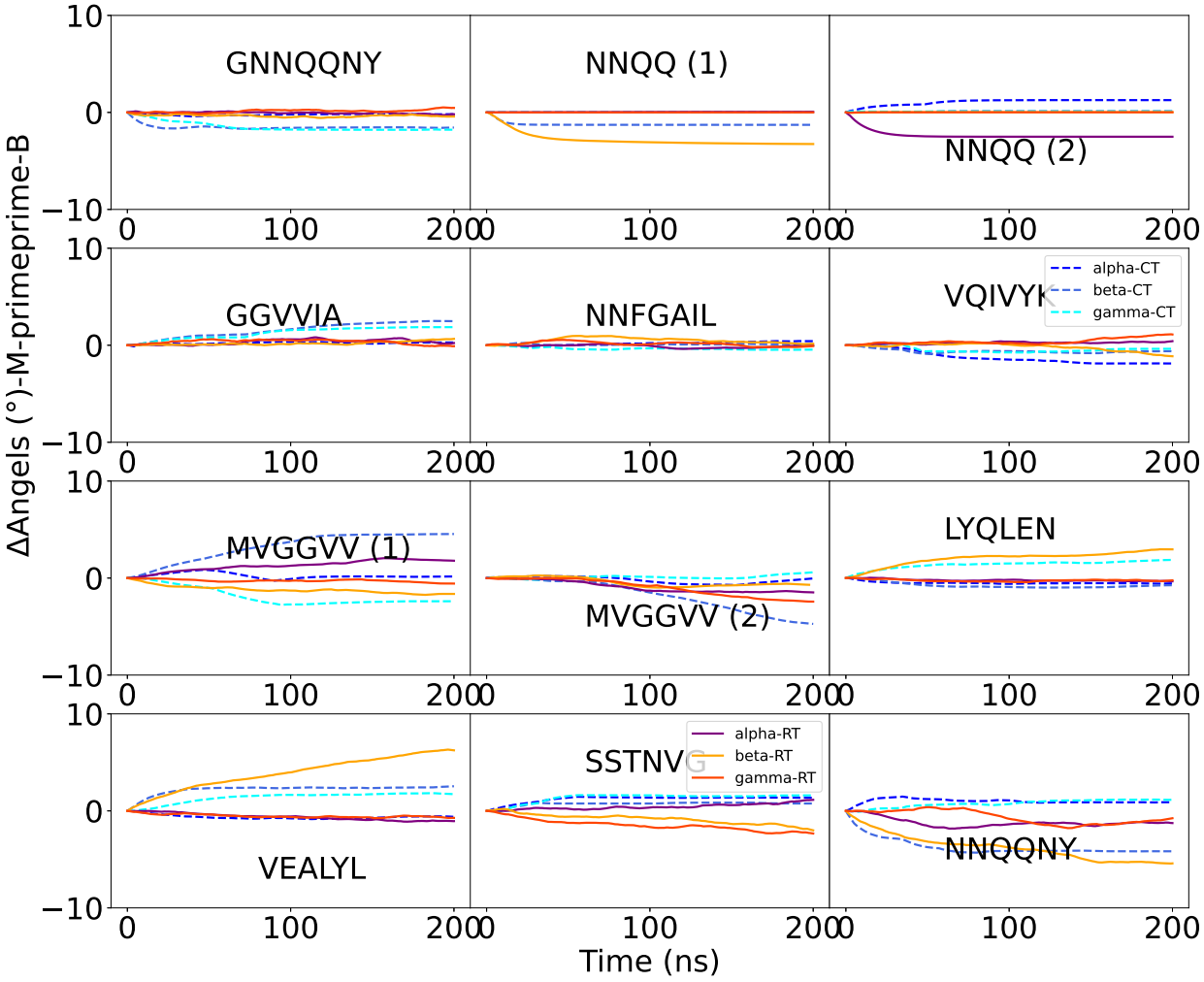

Figure S12: Deviation of angles of the supercell from experimental crystal structure over the NpT and production runs for all peptides considering modifications (M3″) using Berendsen barostat.

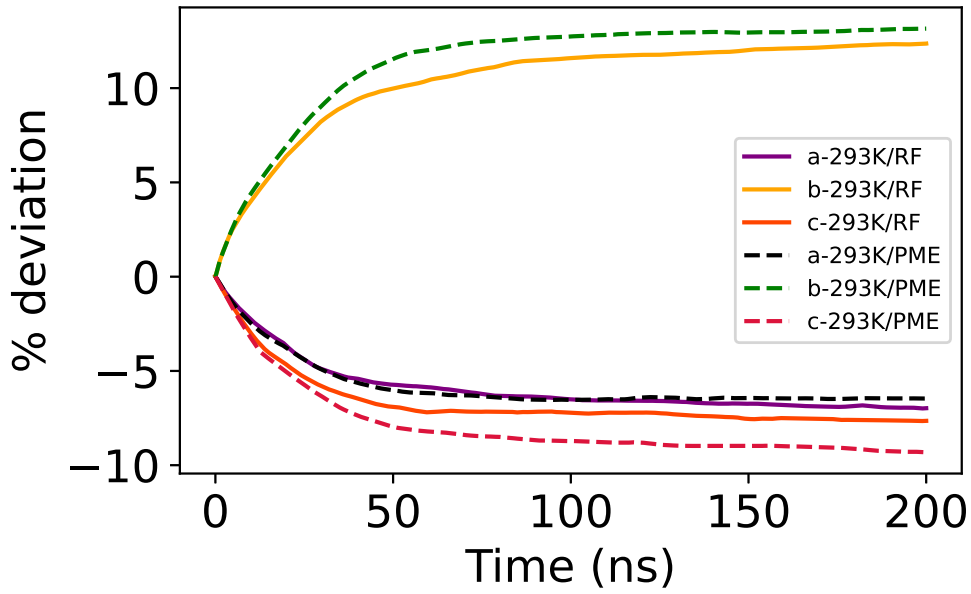

Figure S13: Deviation of lattice size of the NNFGAIL in %, over the NpT and production runs using reaction-field and particle mesh Ewald (PME) for the (M3″) force field using Berendsen barostat at 293 K.

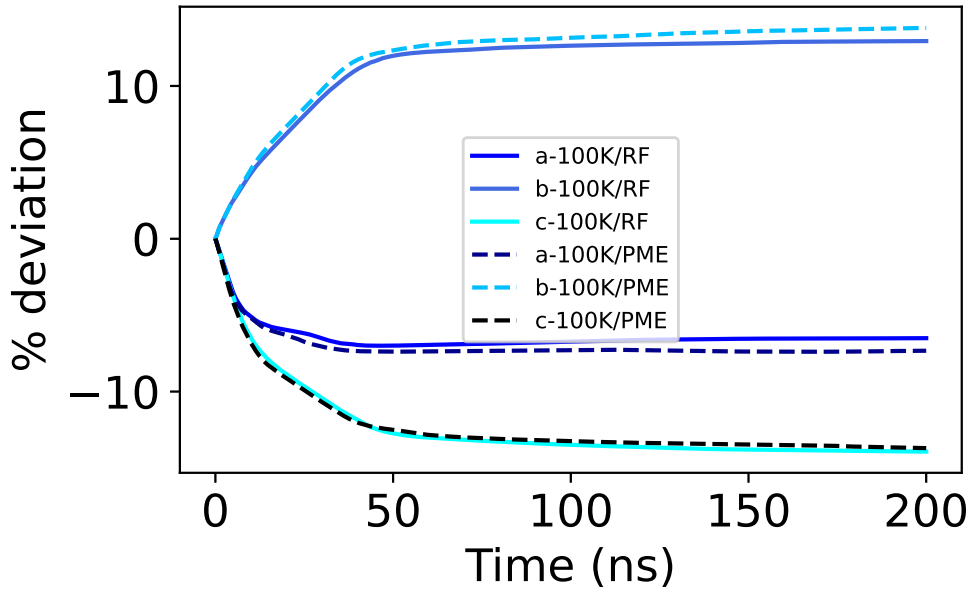

Figure S14: Deviation of lattice size of the NNFGAIL in %, over the NpT and production runs using reaction-field and PME for the (M3″) force field using Berendsen barostat at 100 K.

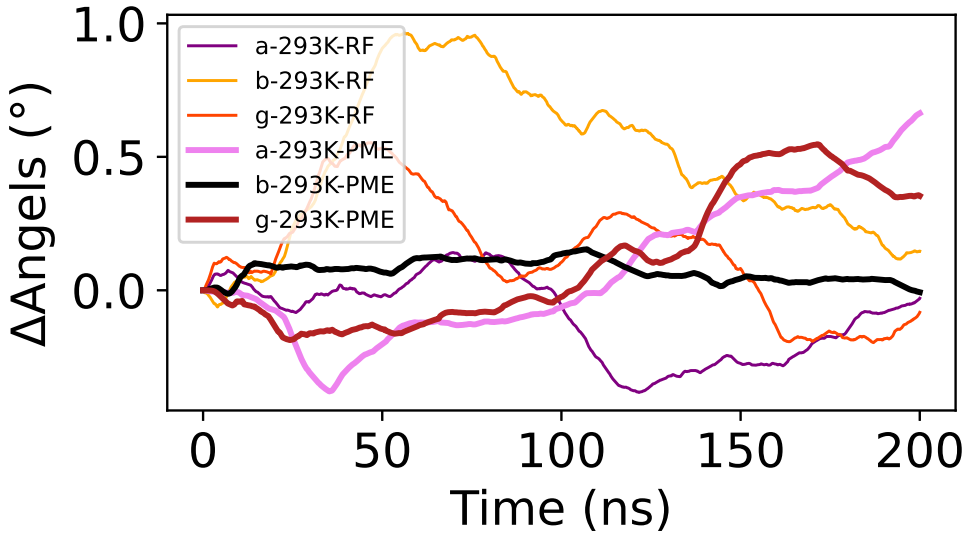

Figure S15: Deviation of angles of the NNFGAIL from experimental crystal structure over the NpT and production runs using reaction-field and PME considering (M3″) force field model with Berendsen barostat at 293 K.

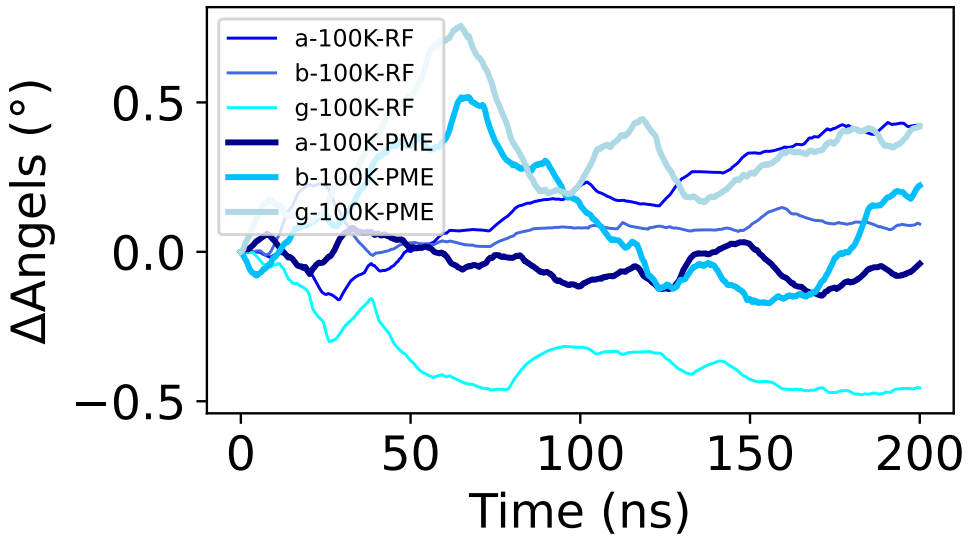

Figure S16: Deviation of angles of the NNFGAIL from experimental crystal structure over the NpT and production runs using reaction-field and PME considering (M3″) force field model with Berendsen barostat at 100 K.

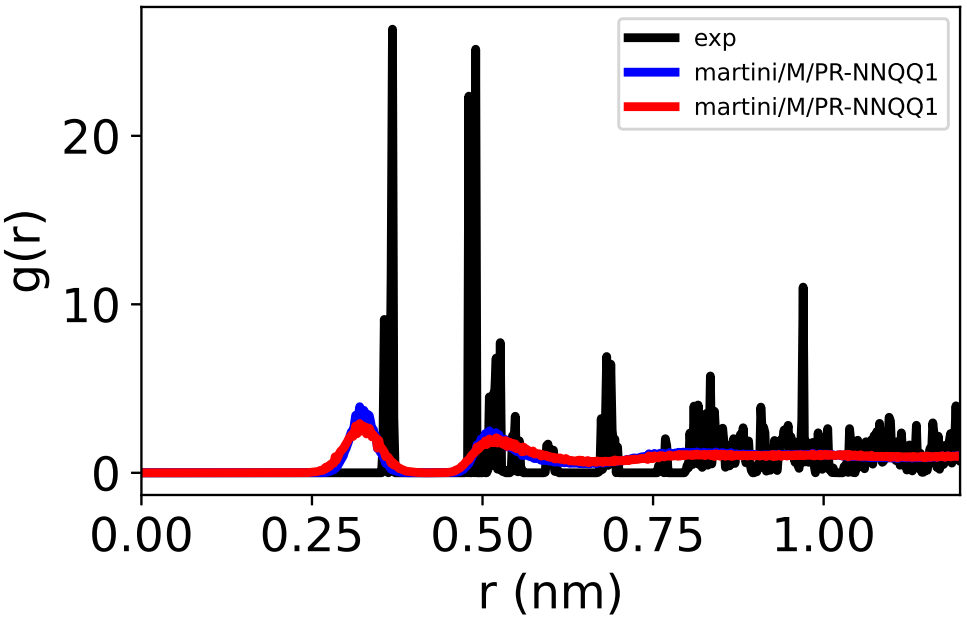

Figure S17: Radial distribution function (RDF) of the backbone of NNQQ1 at cryo and room temperatures for both experiment and simulation (M3) using the Parrinello barostat.

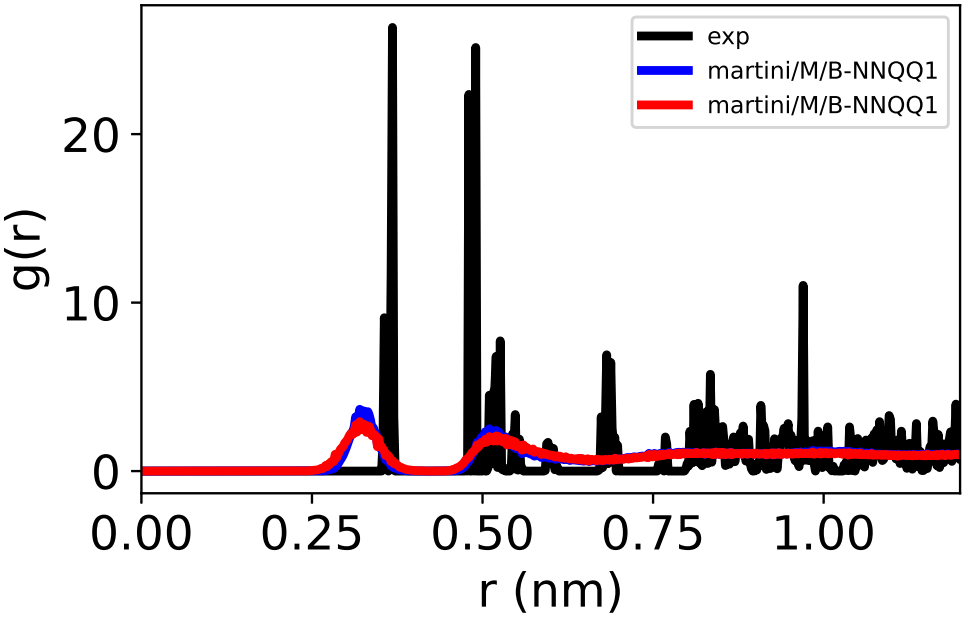

Figure S18: Radial distribution function (RDF) of the backbone of NNQQ1 at cryo and room temperatures for both experiment and simulation (M3) using the Berendsen barostat.

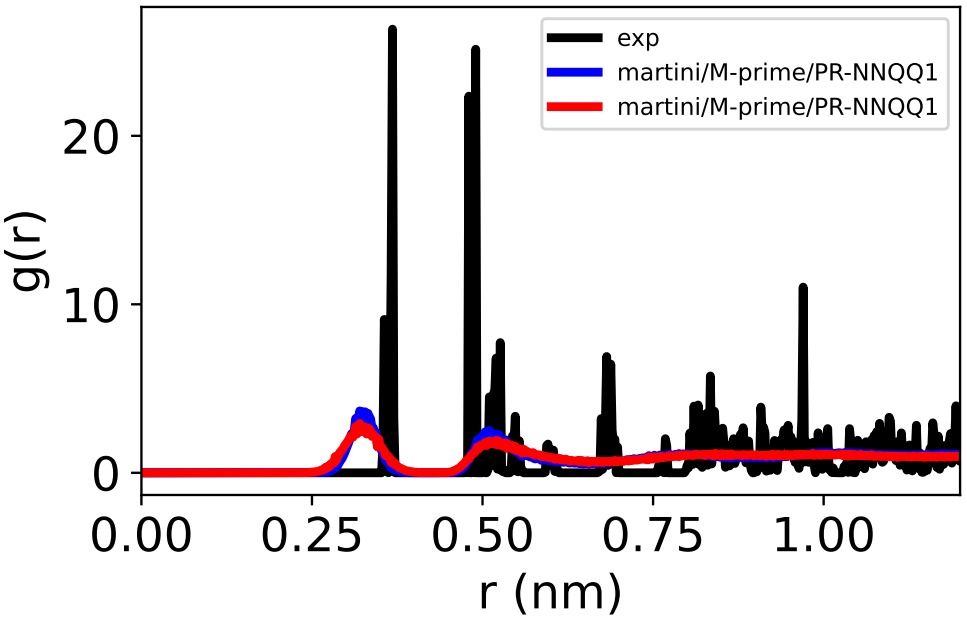

Figure S19: Radial distribution function (RDF) of the backbone of NNQQ1 at cryo and room temperatures for both experiment and simulation using the Parrinello barostat (M3′).

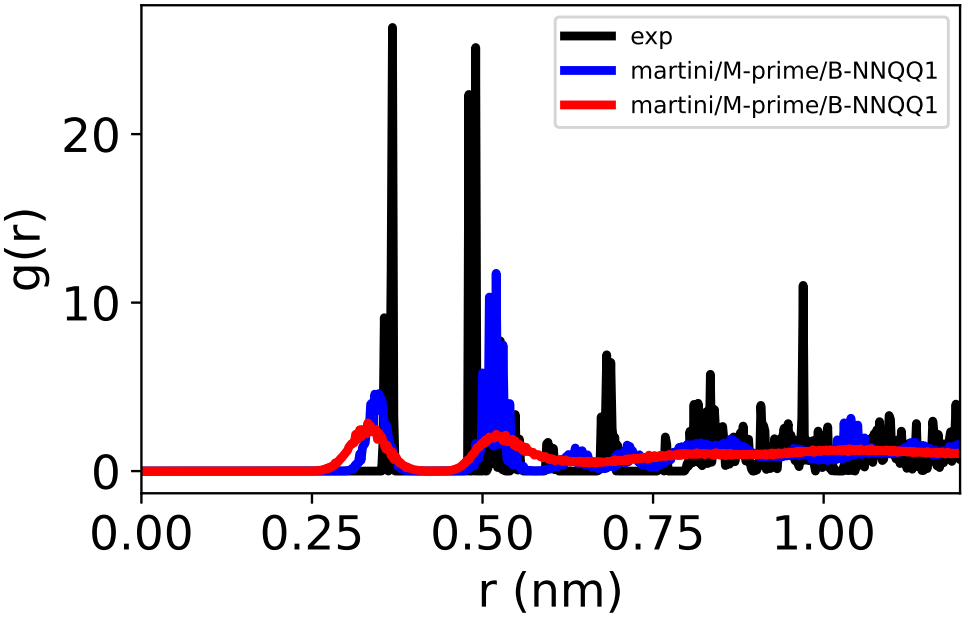

Figure S20: Radial distribution function (RDF) of the backbone of NNQQ1 at cryo and room temperatures for both experiment and simulation using the Berendsen barostat (M3′).

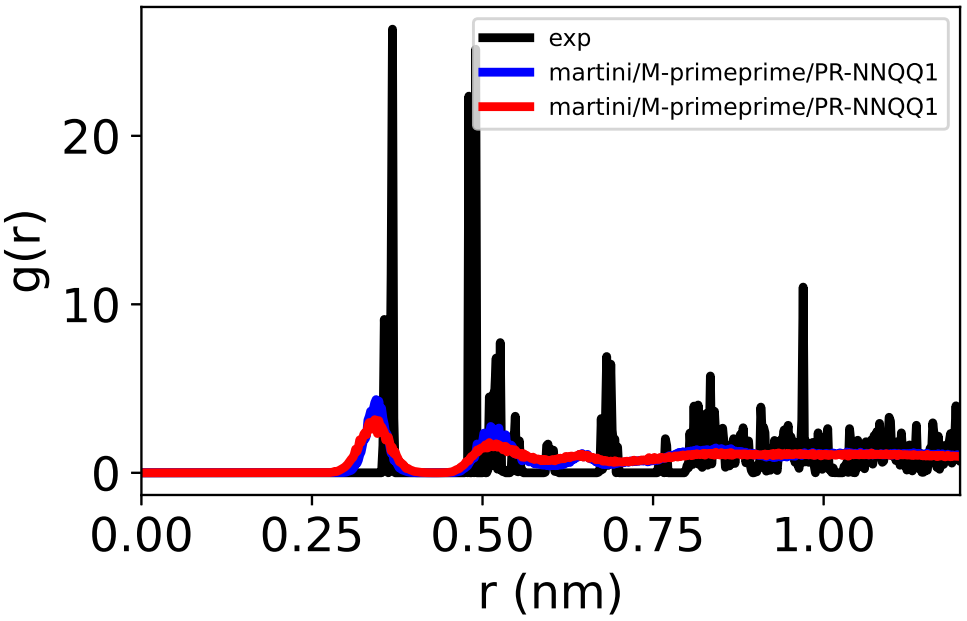

Figure S21: Radial distribution function (RDF) of the backbone of NNQQ1 at cryo and room temperatures for both experiment and simulation using the Parrinello barostat (M3″).

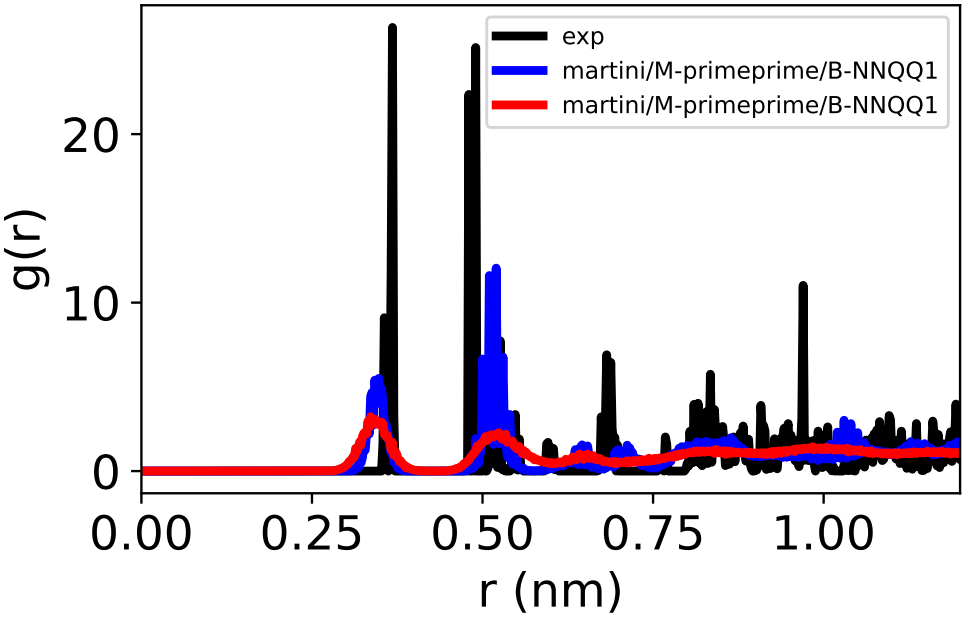

Figure S22: Radial distribution function (RDF) of the backbone of NNQQ1 at cryo and room temperatures for both experiment and simulation using the Berendsen barostat (M3″).

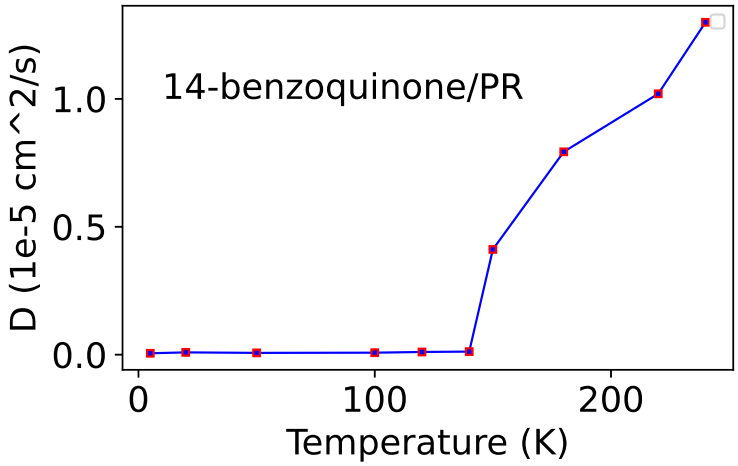

Figure S23: Diffusion as a function of temperature using Parrinello-Rahman barostat.

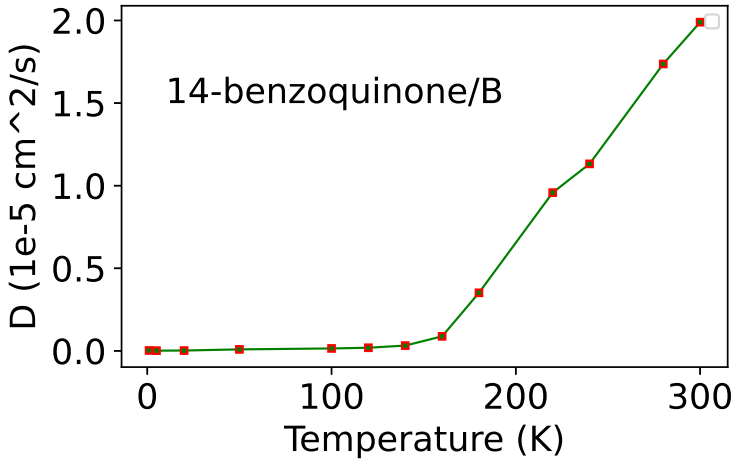

Figure S24: Diffusion as a function of temperature using Berendsen barostat.

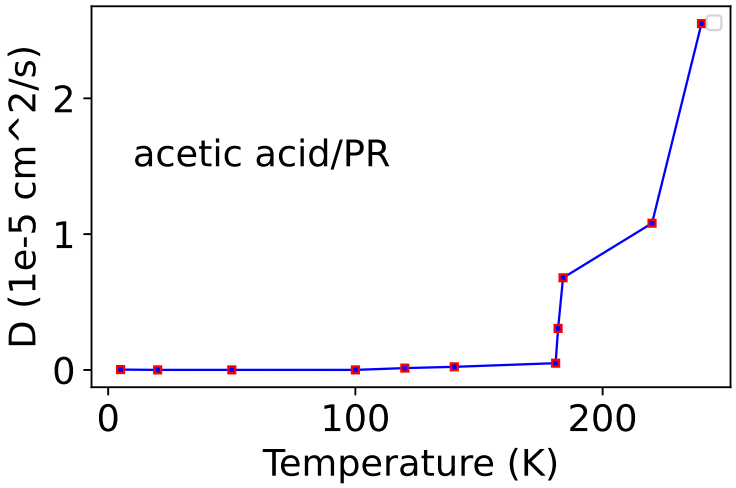

Figure S25: Diffusion as a function of temperature using Parrinello-Rahman barostat.

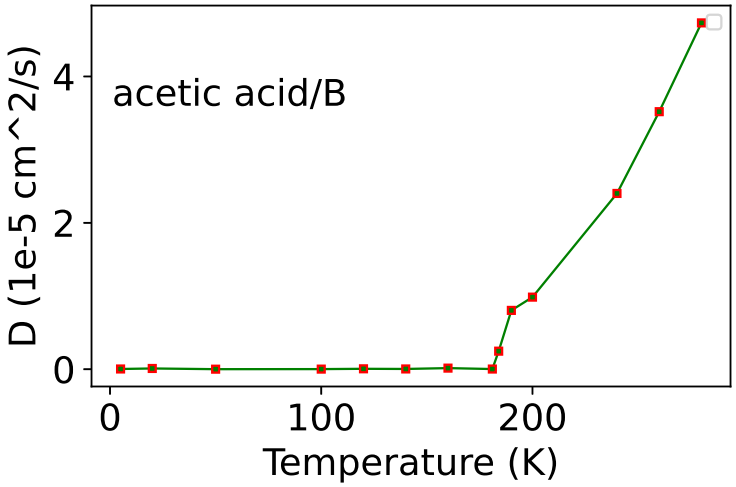

Figure S26: Diffusion as a function of temperature using Berendsen barostat.

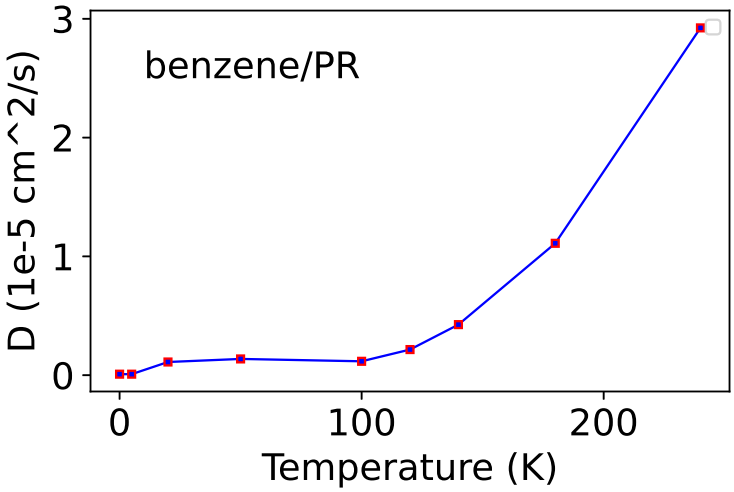

Figure S27: Diffusion as a function of temperature.

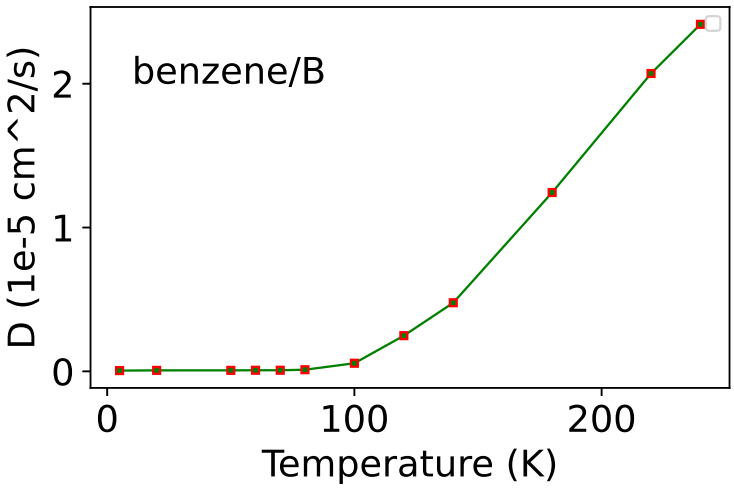

Figure S28: Diffusion as a function of temperature.

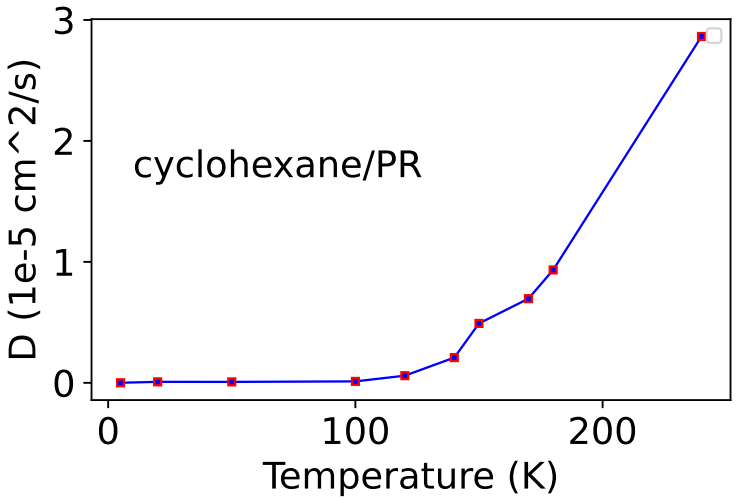

Figure S29: Diffusion as a function of temperature.

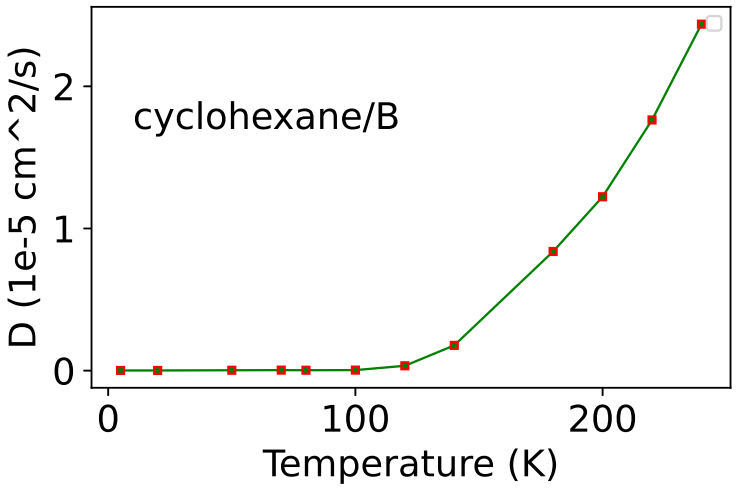

Figure S30: Diffusion as a function of temperature.

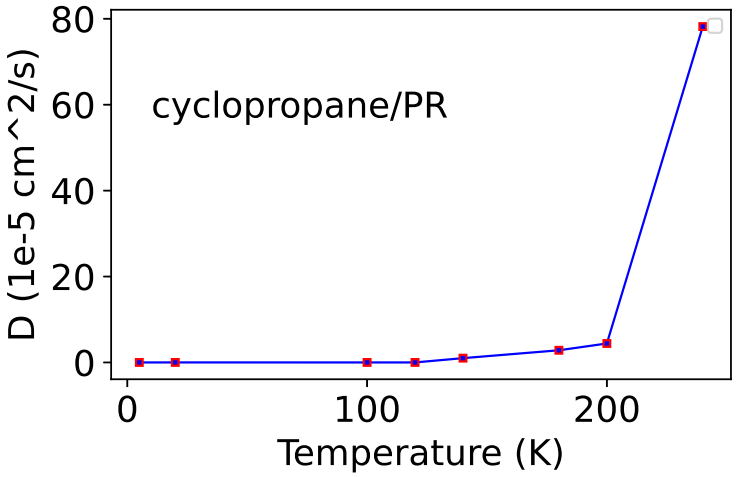

Figure S31: Diffusion as a function of temperature.

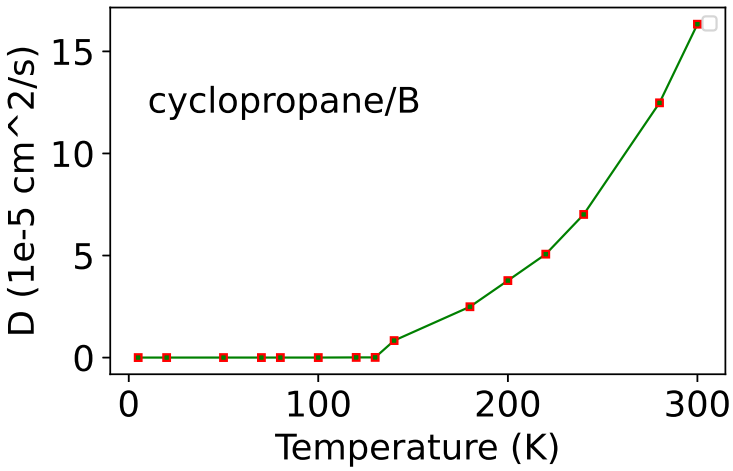

Figure S32: Diffusion as a function of temperature.

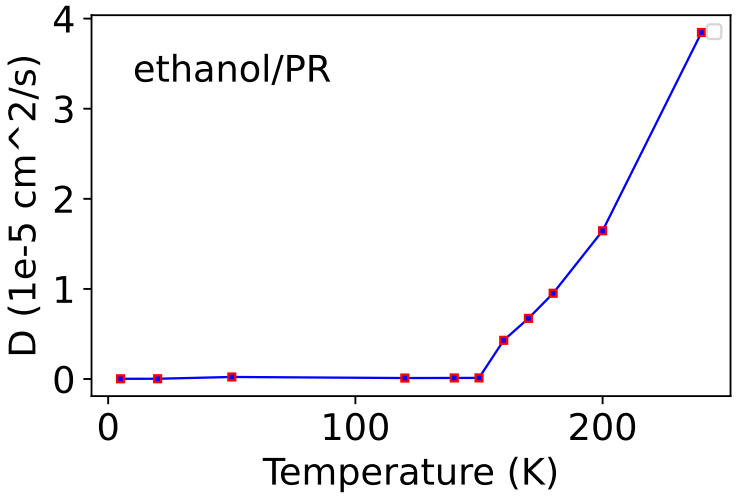

Figure S33: Diffusion as a function of temperature.

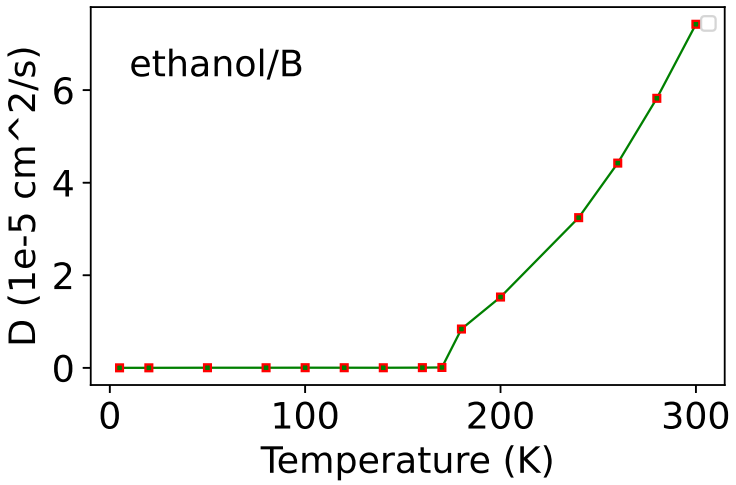

Figure S34: Diffusion as a function of temperature.

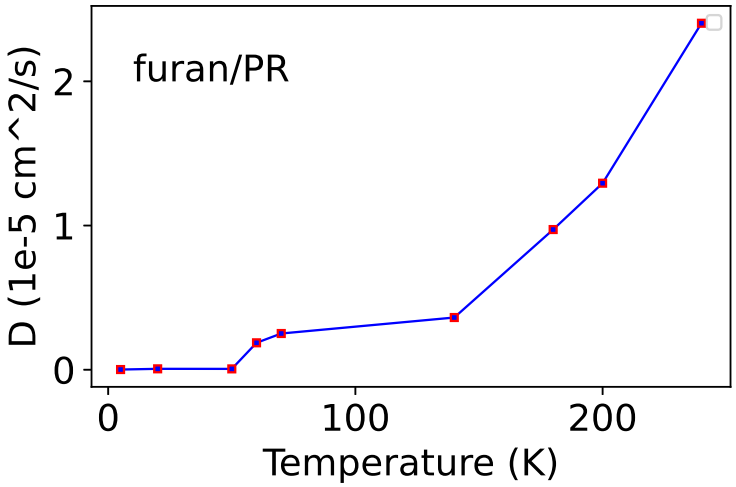

Figure S35: Diffusion as a function of temperature.

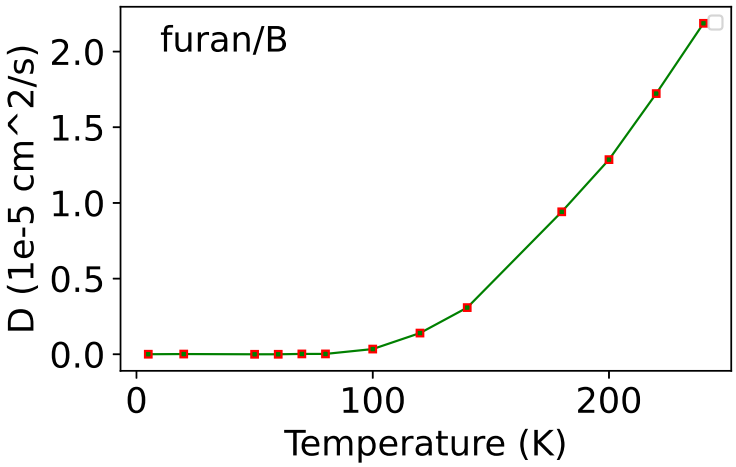

Figure S36: Diffusion as a function of temperature.

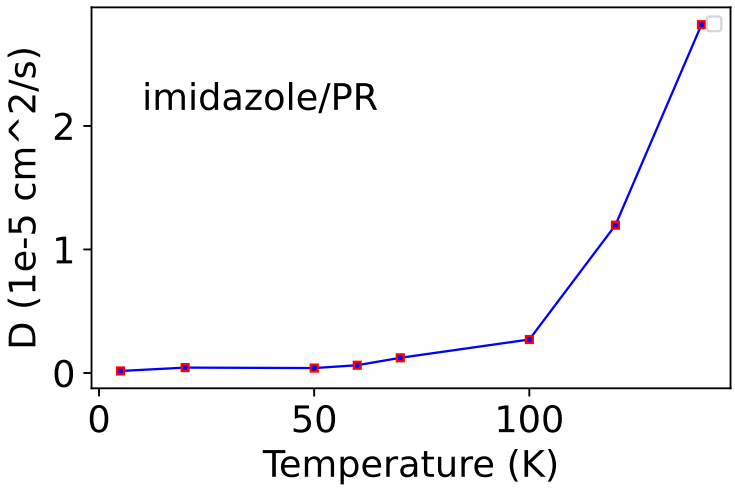

Figure S37: Diffusion as a function of temperature.

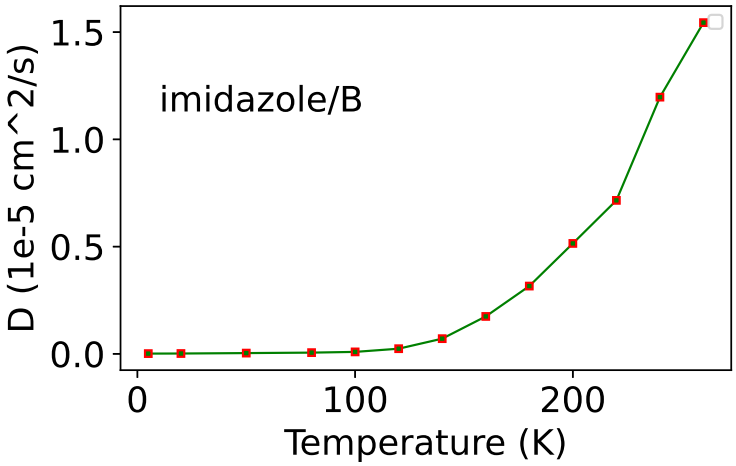

Figure S38: Diffusion as a function of temperature.

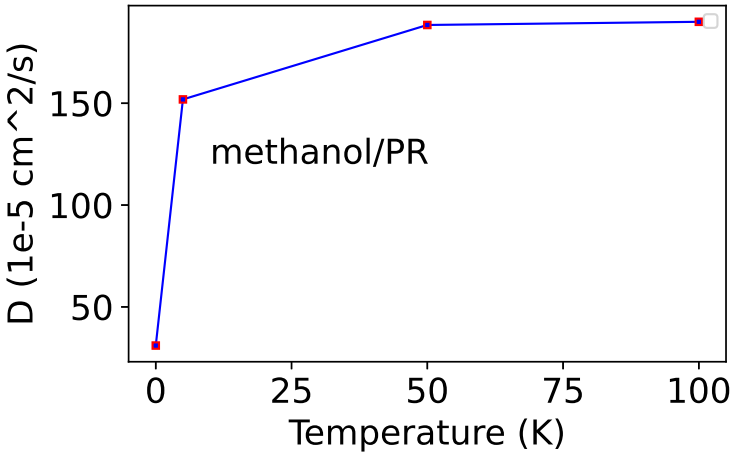

Figure S39: Diffusion as a function of temperature.

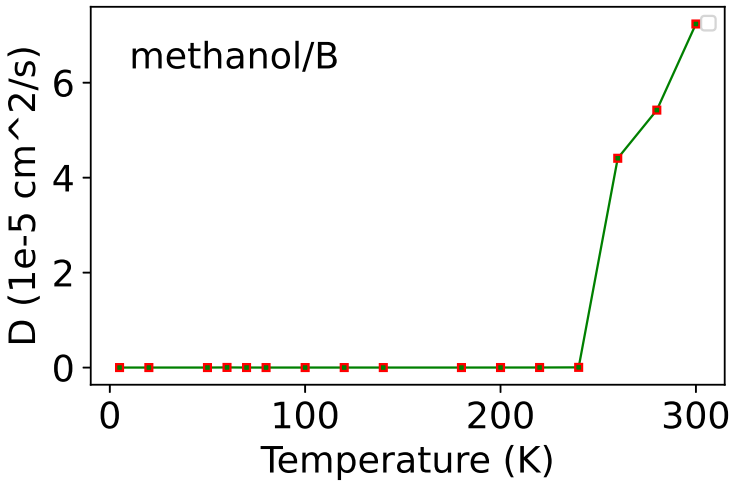

Figure S40: Diffusion as a function of temperature.

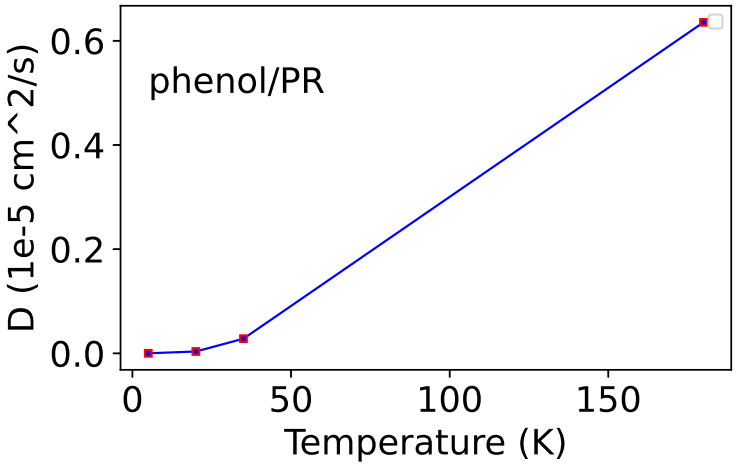

Figure S41: Diffusion as a function of temperature.

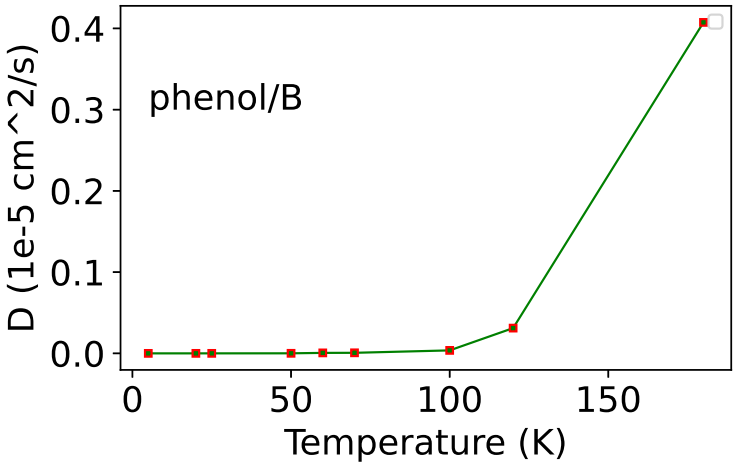

Figure S42: Diffusion as a function of temperature.

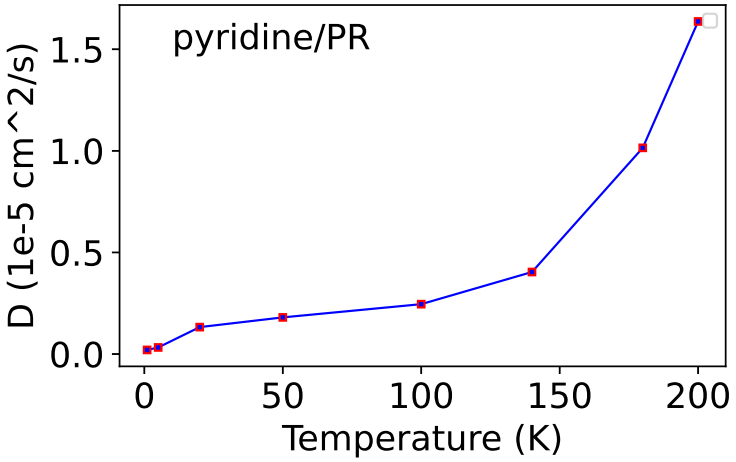

Figure S43: Diffusion as a function of temperature.

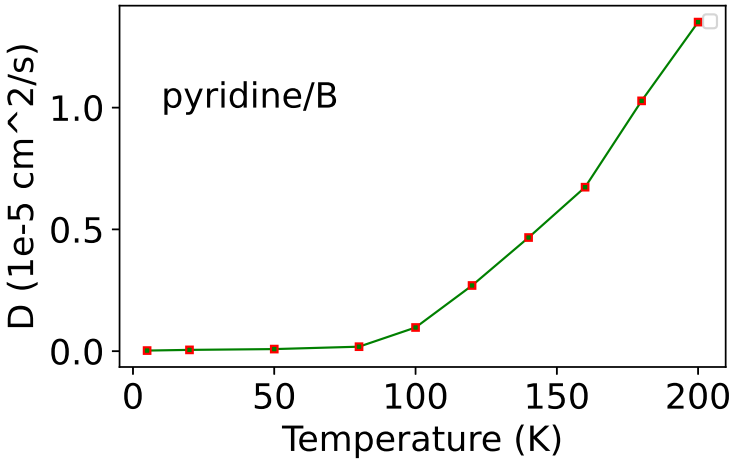

Figure S44: Diffusion as a function of temperature.

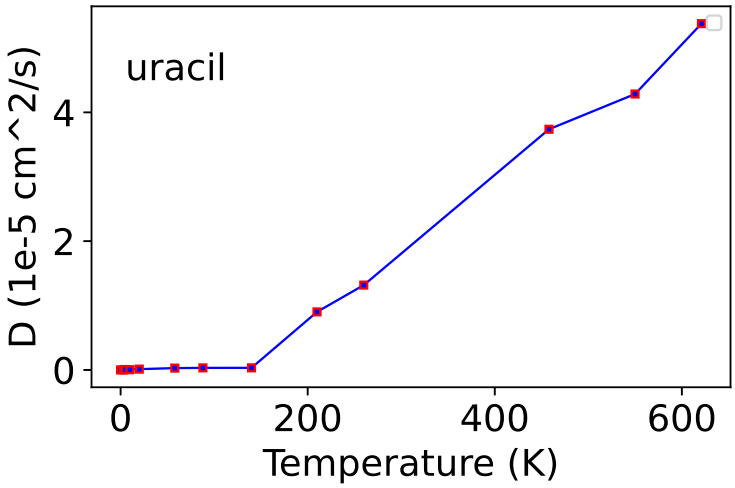

Figure S45: Diffusion as a function of temperature.

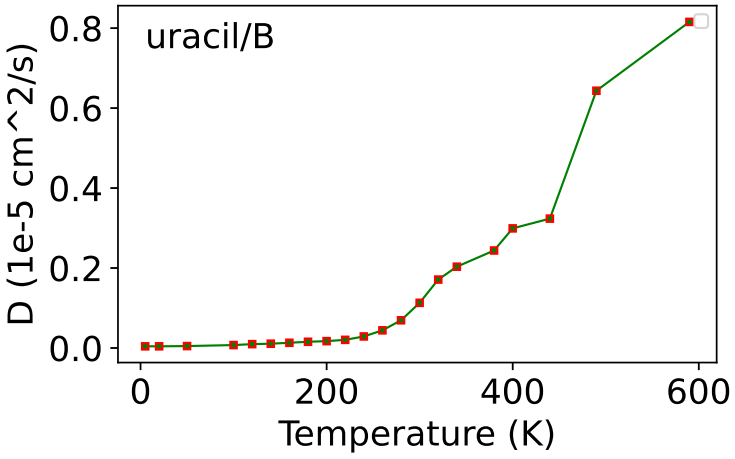

Figure S46: Diffusion as a function of temperature.

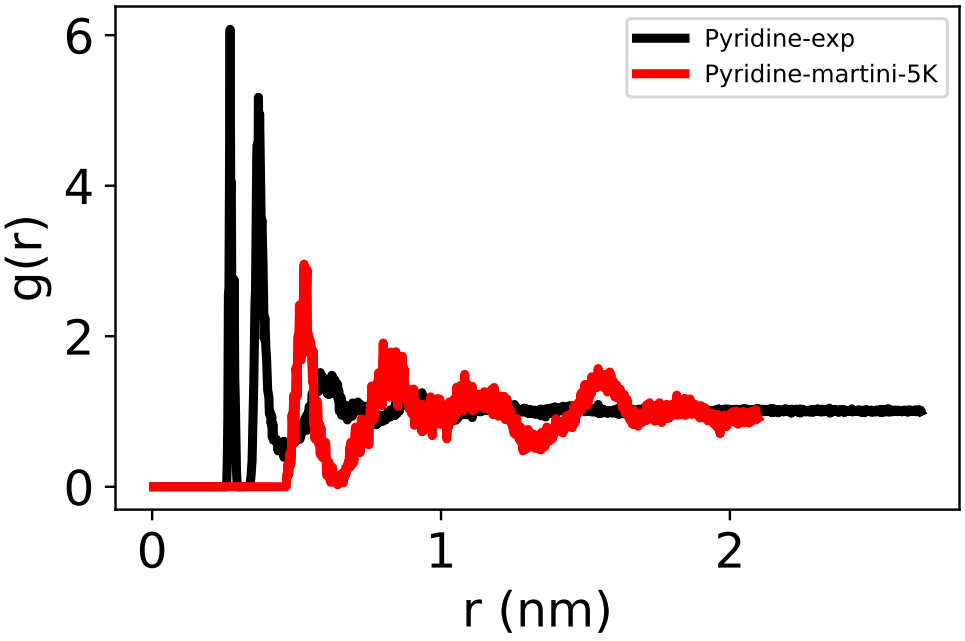

Figure S47: Radial distribution function (RDF) of pyridine at 5 K for both experiment and simulation using the Berendsen barostat.

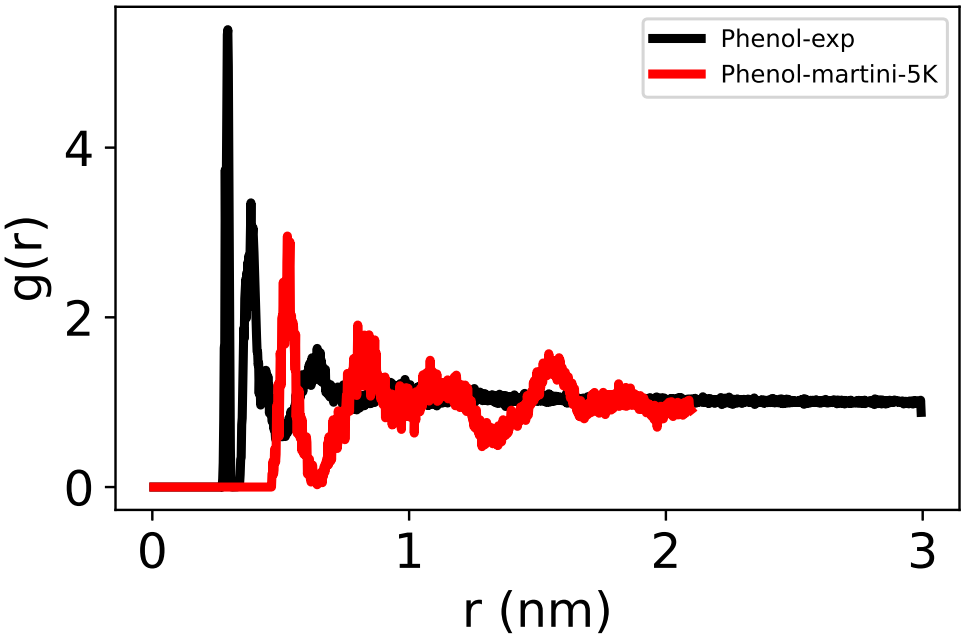

Figure S48: Radial distribution function (RDF) of phenol at 5 K for both experiment and simulation using the Berendsen barostat.

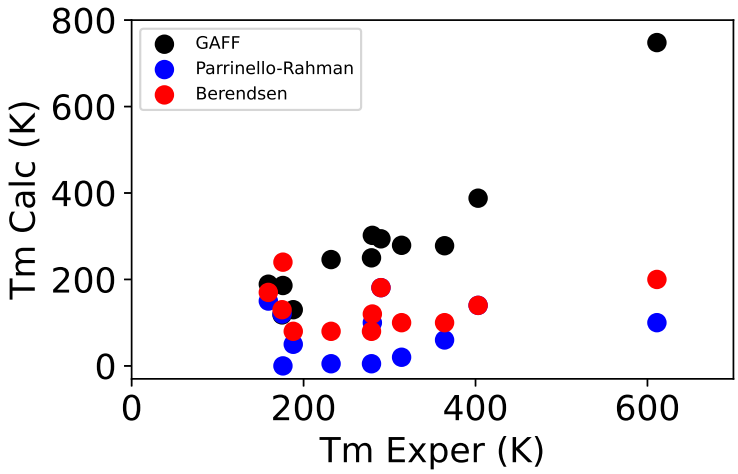

Figure S49: Correlation between experimental melting temperature and simulated melting temperatures for 12 organic compounds (see table 2 in main text).