The liquid structure of tetrachloroethene: Molecular Dynamics simulations and Reverse Monte Carlo modeling with interatomic potentials

Orsolya Gereben and László Pusztai

Institute for Solid State Physics and Optics, Wigner Research Centre for Physics, Hungarian Academy of Sciences, P.O.Box 49, H-1525 Budapest, Hungary

ogereben@freemail.hu, *pusztai.laszlo@wigner.mta.hu*

Corresponding author: László Pusztai

Abstract

The liquid structure of tetrachloroethene has been investigated on the basis of measured neutron and X-ray scattering structure factors, applying molecular dynamics simulations and reverse Monte Carlo (RMC) modeling with flexible molecules and interatomic potentials. As no complete all-atom force field parameter set could be found for this planar molecule, the closest matching OPLS-AA intra-molecular parameter set was improved by equilibrium bond length and angle parameters coming from electron diffraction experiments [Karle, I. L.; Karle, J. *J. Chem. Phys*. **1952,** *20* 63]. In addition, four different intra-molecular charge distribution sets were tried, so in total, eight different molecular dynamics simulations were performed. The best parameter set was selected by calculating the mean square difference between the calculated total structure

factors and the corresponding experimental data. The best parameter set proved to be the one that uses the electron diffraction based intra-molecular parameters and the charges q_C=0.1 and q_{Cl} =–0.05. The structure was further successfully refined by applying RMC computer modeling with flexible molecules that were kept together by interatomic potentials. Correlation functions concerning the orientation of molecular axes and planes were also determined. They reveal that the molecules closest to each other exclusively prefer the parallel orientation of both the molecular axes and planes. Molecules forming the first maximum of the center-center distribution have a preference for <30 º and >60 º axis orientation and >60 º molecular plane arrangement. A second coordination sphere at ~11 Å and a very small third one at ~16 Å can be found as well, without preference for any axis or plane orientation.

Keywords: liquid structure; molecular dynamics; Reverse Monte Carlo modeling

PACS: 61.20.-p; 61.20.Ja; 61.25.Em

I. INTRODUCTION

Tetrachloroethylene (systematic name tetrachloroethene, further referred to as C_2Cl_4) is a colorless liquid with sweet odor and it is an excellent organic solvent. It is widely used as cleaning fluid, appears in paint strippers and degrease metal parts in the metalworking industries[1], therefore it is produced in large quantities. Due to the commercial production, it is a common groundwater contaminant[2], and can be found in abundance at hazardous waste disposal sites. Breathing in larger quantities of the substance causes vomiting, dizziness and eventually loosing consciousness[3]. The possibility to use zero-valent iron as a reducing agent to remove tetrachloroethylene from drinking water was investigated[4].

The gas phase structure was determined by electron diffraction by Karle et al.[5]. Concerning the liquid, neutron diffraction data-based structure analysis, using the Ornstein-Zernike integral equation in an attempt to resolve the partial radial distribution functions (*prdf* or $g_{ij}(r)$), was conducted by Alvarez et al.[6].

Regardless of its many uses and widespread occurrence, our interest in this molecule is due to its simple planar structure. The liquid structure, which may be viewed as a prototype for disordered assemblies of planar molecules, has not yet been considered in any detail, and seems to be and ideal candidate to test our new simulation technique for the first time for planar molecules. These are the main reasons why the present extensive investigation was initiated.

The structure of small organic liquid molecules can be investigated by different experimental techniques (for example electron[7], neutron[8] and X-ray diffraction[9],

infrared[10], Raman[11] and NMR[12] spectroscopy) or using theoretical approaches like ab initio calculations[13], density functional theory[14] or integral equations[15]. Different simulation methods are available, as well, that apply interaction potential functions: molecular dynamics (MD)[16] and standard (Metropolis) Monte Carlo[17] simulations are the best known examples.

As the interpretation of experimental structural data, the liquid structure factor obtained from diffraction measurements, is not entirely straightforward, additional efforts are required for extracting structural information from them. In the 1970-s Bertagnolli and coworkers investigated simple molecular liquids, especially chloroform[18,19], by neutron diffraction using isotopic substitution and by X-ray diffraction[20]. They attempted to extract information on partial radial distribution functions (prdf) by separating the radial distribution function, obtained by Fourier-transformation of the coherent differential cross section, into intra and intermolecular parts, wherever it seemed feasible. Assuming structural models based on the solid structure was also attempted, for instance in the case of acetic acid[21]. Comparing the calculated and experimental functions resulted in agreements that cannot be considered as 'good', according to present standards.

Later, simulation techniques using experimental data during the modeling process were developed; applications of Monte Carlo based methods for organic liquids like the empirical potential structure refinement (EPSR)[22] and Reverse Monte Carlo (RMC)[23,24] simulation may be mentioned. Inverse methods with the aim of extracting only intramolecular parameters by Bayesian fitting[25] have also appeared.

Reverse Monte Carlo structural modeling[26,27,28] is widely used for building 3D models of liquids[29,30] as well as of amorphous materials like semiconductors[31,32] and metallic

glasses[33]. In the traditional RMC modeling no interatomic potentials (except the hard sphere potential) are used: the calculation is driven by the requirement that a given (set of) experimental diffraction data have to be approached as closely as possible. The lack of potential functions can be considered both as a fault and an advantage[27] of this method, as the choice of the potential is obviously a bias. Still, including interatomic potentials into the RMC method has been desired for a long time: such combination could provide direct information concerning the extent that a given potential parameter set is compatible with the experimental structure data.

Therefore lately a new molecular version of the RMC code, RMC_POT, was developed. It operates with flexible molecules that are defined and kept together by means of bonded and non-bonded interatomic potentials. The RMC_POT algorithm has already been successfully applied for revealing the structure of molecular liquids of sulfur-containing small organic materials[34,35]. The approach that is realized by the RMC_POT computer code will be referred to as 'FMP-RMC' from this point on (FMP: 'Flexible Molecules using interatomic Potentials'; the computer code RMC_POT is freely available[36]).

Although in principle, the RMC_POT software was developed for molecules of arbitrary shape, the most challenging of all shapes, the (reasonably flexible) planar geometry has not yet been considered. Therefore liquid C_2Cl_4 was chosen as the most suitable prototype for liquids of planar molecules: it (1) is easily available, (2) is a liquid under ambient conditions, (3) contains a C=C double bond (with sp^2 carbons), which is an essential building block in organic chemistry, and (4) is widely used in everyday life.

In this work the results of a series of molecular dynamics and FMP-RMC simulations, based on measured X-ray and neutron diffraction data, are reported for liquid tetrachloroethene under ambient conditions. The structure of the liquid is characterized by distance dependent correlations between molecular axes and planes; the computer software that calculates these functions has been developed in conjunction with the present study.

II. SIMULATION DETAILS

In both the MD and RMC calculations, 2197 molecules were put in cubic simulation cells, with an atomic number density of $3.5 \cdot 10^{-2}$ $Å^{-3}$; this setup corresponds to the (experimentally determined) bulk density of 1.61 g/cm^3.

A. Molecular dynamics simulations

Molecular dynamics simulations have been performed with the GROMACS[8] simulation package (version 4.0)[37] at T=293 K in the NVT ensemble. GROMACS uses pairwise interactions representing dispersion and repulsion effects (such as the well-known Lennard-Jones potential) and Coulomb terms for describing electrostatic interactions. The 1-2 (i.e., within a molecule, bonded) interactions are handled as harmonic bond stretching, the 1-3 (i.e., non-bonded) interactions as harmonic angle bending, whereas the 1-4 ('torsion', within a molecule) interactions are approximated by dihedral potentials (see the corresponding 'Supplementary Information').

The temperature was kept constant using the Nose-Hover thermostat[38] with $\tau_T = 0.5$ ps. Preliminary tests with time steps of 1 fs and 2 fs were conducted; as there were no

differences regarding the outcome of interest here of the corresponding simulations, the structures discussed below resulted from simulations with a time step of 2 fs. The simulation length was 2000 ps in each case. Equilibrium could be reached under 100 ps and particle configurations in the 'production'-phase (between 500 and 2000 ps) were collected 20 ps apart.

For the MD calculations all-atom force field parameters were needed in order to make the calculation of the *prdf*-s, the total scattering X-ray [$F^X(Q)$] and neutron [$F^N(Q)$] structure factors (*tssf*), and distributions of bond lengths, bond angles and dihedral angles possible. No all-atom force field containing a complete set of parameters for C_2Cl_4 could be found in the literature; even in the study of Alvarez[6] the parameters for CCl_4[39] and CCl_2F_2[40], both with carbon atoms in the sp^3 hybrid state, were used with adjusted charges. For this reason, we were forced to engage in some 'potential developer' type of activity. The OPLS-AA[41] force field contains parameters for chloroalkenes (molecules with, for instance, ClHC=C- fragments), which seem applicable for our case, as the carbon atoms are in the appropriate hybrid state, sp^2. This is why the OPLS-AA force field chloroalkene Lennard-Jones (LJ) parameters of σ_C=3.55 and σ_{Cl}=3.4 Å and ε_C=3.17984·10^{-1} and ε_{Cl}=1.2552 kJ/mol (for the carbon and chlorine atoms, respectively) were taken for the present study. Partial charges also had to be adjusted, in order to make the molecule neutral: four different charge sets were tested in order to find the best parameter set that would produce results closest to the experimental structure factors (see TABLE I). In simulation '*C'* the original partial charge of +0.005 for the chloroalkene carbon atom was applied, and the partial charge of chlorine was adjusted. Analogously, in simulation '*Cl'* the original partial charge of –0.12 for the chloroalkene chlorine atom

was applied, and the partial charge of the carbon was adjusted. In order to explore a wider range, a charge set with smaller charges (‘*Sm’*) and one between *C* and *Cl* called ‘*MID’* were applied as well (see TABLE I). There was no need to use a more polar model than the *Cl* set (see results below).

TABLE I: Partial charges applied in the various MD simulations of liquid C_2Cl_4. The original charges of the force field are highlighted by bold characters.

	C charge	Cl charge
Sm	0.002	-0.001
C	**0.005**	-0.0025
MID	0.1	-0.05
Cl	0.24	**-0.12**

Concerning the intra-molecular parameters, like bond lengths and angles, two different sets were applied (see TABLE II). The first set contained the OPLS-AA parameter values (simulations ‘*ORI’*). The other set operates with the intra-molecular parameters of gas phase electron diffraction (ED)[5] data (simulation sets ‘*ED’*), still using the force constants of the OPLS-AA force field. The molecules were kept planar in both cases by using improper dihedrals[42], according to the convention of the OPLS-AA force field. (For more details concerning intramolecular parameters, see the correponding ‘Supplementary Information’.)

TABLE II: Intra-molecular equilibrium force field parameters used in the molecular dynamics simulations.

	ORI[a]	*ED*[b]
C-C (Å)	1.34	1.3
C-Cl (Å)	1.725	1.72
Cl-C-Cl (º)	*117.0*[c]	113.5
Cl-C-C (º)	121.5	123.25

[a] Ref. [37].

[b] Ref. [5].

[c] Calculated from Cl-C-C *ORI*.

All the MD results reported here have been averaged over 76 time frames; this amount proved to be more than sufficient for obtaining good statistics. Note that apart from being 'stand-alone' results, the final MD configurations provided excellent starting points for the subsequent RMC studies.

B. Reverse Monte Carlo modeling

In order to improve the agreement between experimental and calculated structural data, FMP-RMC calculations were performed, starting from the final particle configurations of the MD simulations *ORI_C* and *ED_MID* (which have turned out to be the most favorable parameter sets, see below).

As details of the RMC algorithm are described elsewhere[26,27,28], here only the basic principle is given. During a conventional RMC simulation one or more atoms are moved randomly in the simulation box. If the squared differences between the experimental data set(s) and the calculated one(s) decrease due to the move then the move is accepted; if

not, then the move may still be accepted with a probability of $\exp\left\{-\left(\chi^2_{new}-\chi^2_{old}\right)/2\right\}$ (the 'χ' values here represent the sums of the squared differences).

When using the flexible molecule 'FMP' approach the total potential energy, V_{pot}, consisting of bonded (bond stretching, angle bending and dihedral) and non-bonded (Lennard-Jones and Coulomb) terms, is also calculated in every step. From the total potential energy values $\chi^2_{pot}=V_{pot}/\sigma_{pot}$ is calculated. (The potential energies are calculated the same way as in the GROMACS[34,35,37] package.) This step precedes the 'normal', data set based χ^2 calculation and acceptance/rejection (for details of the scheme, see, e.g., Refs.[26,28,34]). If χ^2_{pot} decreased then the move is accepted based on the potential energy; if it increased then it is only accepted with a probability $\exp\left\{-\left(\chi^2_{pot,new}-\chi^2_{pot,old}\right)/2\right\}$. If the move based on the potential energy is accepted, only then the calculation and the acceptance/rejection process based on the data sets and constraints are carried out.

Diffraction data have been the most frequently used input for RMC modeling. For the present investigations on liquid tetrachloroethylene, neutron diffraction data have been obtained[43] on the PSD 2-axis diffractometer[44] installed at the Budapest Research Reactor (Budapest, Hungary). Because of its apparent simplicity, liquid C_2Cl_4 has routinely been investigated at various synchrotron X-ray scattering centers, e.g. as a standard material for comparing capabilities of various instruments. In this work the X-ray diffraction data[45] measured on the high energy liquid and amorphous X-ray diffraction beamline BL04B2[46], installed in the SPring-8 synchrotron facility (Hyogo, Japan), are made use of.

Both the neutron and X-ray total scattering structure factors ($F^N(Q)$ and $F^X(Q)$) have been used here as input experimental data sets; in some cases, the three *prdf*-s from *ORI_C* and *ED_MID* MD simulations were also used during FMP_RMC (similarly as in,

for instance, Refs. [29,30,35]). The molecules were kept together by the flexible molecule approach[34,35]. Non-bonding potentials (Lennard-Jones and Coulomb terms) have also been calculated and minimized during the process. The simulation names indicate the used charge set and input data sets; details of FMP-RMC simulations are summarized in TABLE III.

TABLE III. FMP-RMC calculations performed in this work. Data sets marked with '+' used as input and the applied charge set is given for the different FMP-RMC simulations. (The '3 *prdf*-s' are: C-C, C-Cl and Cl-Cl.) The simulation names reflect the data sets applied and the charge set, as well.

	$F^N(Q)$	$F^X(Q)$	*3 ORI_C prdf*-s	*3 ED_MID prdf*-s	Charge set
ORI_fq_sq	+	+	-	-	*C*
ORI_fq_sq_gr	+	+	+	-	*C*
ED_fq_sq	+	+	-	-	*MID*
ED_fq_sq_gr	+	+	-	+	*MID*

In order to account for possibly remaining small systematic experimental uncertainties, a full quadratic refinement during the χ^2 calculation for the $F^N(Q)$ and $F^X(Q)$ data sets was performed in the final stages of all RMC calculations, according to the following formula:

$$\chi^2 = \frac{\sum_{j=1}^{Np_i} (aX_j^E(Q_j) + b + cQ_j + dQ_j^2 - X_j^C(Q_j))^2}{\sigma^2} \qquad (1)$$

a, b, c, and d are renormalization constants; $X(Q)$ stands for $F^X(Q)$ or $F^N(Q)$; the parameters a, b, c and d were determined so that they yield the minimum of χ^2; Q is the modulus of the momentum transfer vector; $F^E(Q)_j$ and $F^C(Q)_j$ are the jth data points of the

experimental and calculated structure factor, respectively; σ is the weighting parameter of the data set in question.

The χ^2 in comparison to the MD average *prdf*–s and cosine distributions of bond angles (with σ=1) were also calculated for each FMP-RMC simulation, as a quantitative measure of the differences between MD and RMC results.

III. RESULTS AND DISCUSSION

A. Molecular dynamics simulations

A.1. General results – molecular structure and 2-body correlations

All the eight MD calculations produced very similar total energies (see TABLE IV), the calculations with the *ED* parameter set yielding slightly lower energies than the simulations using the *ORI* parameter set.

TABLE IV. The average total potential energy and the normalized mean squared difference, χ^2, between the MD average and quadratically refined experimental data sets (calculated with σ=1). The simulation giving the best agreement with experimental data sets for both parameter sets are highlighted with bold characters. The standard deviation for the energy is also given.

	Total E_{pot} (kJ/mol)	$\chi^2_{F^N(Q)}$	$\chi^2_{F^X(Q)}$	$\chi^2_{F^N(Q)+F^X(Q)}$
ORI_Sm	-23.62±0.13	$2.09\cdot10^{-2}$	$3.19\cdot10^{-1}$	$3.40\cdot10^{-1}$
ORI_C	-23.62±0.15	$\mathbf{1.99\cdot10^{-2}}$	$\mathbf{3.10\cdot10^{-1}}$	$\mathbf{3.30\cdot10^{-1}}$
ORI_MID	-23.56±0.17	$2.29\cdot10^{-2}$	$3.17\cdot10^{-1}$	$3.40\cdot10^{-1}$
ORI_Cl	-23.87±0.14	$7.57\cdot10^{-2}$	$4.51\cdot10^{-1}$	$5.27\cdot10^{-1}$
ED_Sm	-24.09±0.16	$2.60\cdot10^{-2}$	$1.84\cdot10^{-1}$	$2.10\cdot10^{-1}$
ED_C	-24.07±0.15	$2.58\cdot10^{-2}$	$1.83\cdot10^{-1}$	$2.09\cdot10^{-1}$
ED_MID	-23.99±0.15	$\mathbf{2.76\cdot10^{-2}}$	$\mathbf{1.77\cdot10^{-1}}$	$\mathbf{2.05\cdot10^{-1}}$
ED_Cl	-24.40±0.14	$7.95\cdot10^{-2}$	$2.93\cdot10^{-1}$	$3.72\cdot10^{-1}$

The average X-ray $F^X(Q)$ and neutron scattering $F^N(Q)$ structure factors were calculated in each case, as well, and their squared differences from the refined

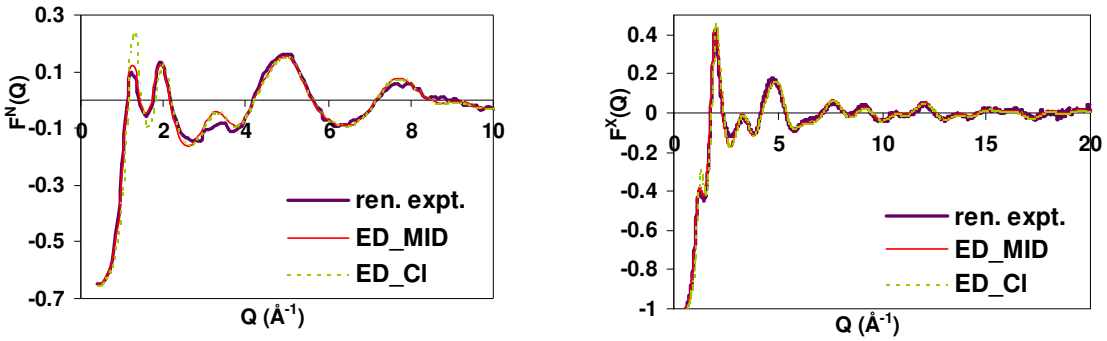
experimental total scattering structure factors (with σ=1) were determined according to equation (1), see TABLE IV. The *ORI_C* simulation seems to be the best out of the *ORI* series, producing the lowest total $\chi^2_{F^N(Q)}$+ $\chi^2_{F^X(Q)}$. The best from the *ED* series and the overall best, from the structural point of view, is the *ED_MID* model. Thus for reproducing the structure of liquid tetrachloroethene, the potential parameters devised in this work appear to be the most successful ones.

Regarding both the neutron and X-ray structure factors, the three potential models with small charges ('*Sm*', '*C*'. '*MID*') produce very similar results but the structure resulting from the '*Cl*' charge set markedly differs from them. The interested reader can find detailed analyses of results of all the MD simulation in the corresponding 'Supplementary Information'; below only comparison between the 'overall best' *ED_MID* and the markedly different *ED_Cl* simulations will be given.

The equilibrium average geometry bond lengths for the *ED_MID* model did not change in comparison with the intramolecular equilibrium distance parameters of the potentials, and only negligible changes could be detected for the equilibrium angles (from 113.5º to ~113.3º for Cl-C-Cl and from 123.25º to ~123º for Cl-C-C).

Considering the neutron diffraction data first, there are noticeable differences in the region Q=1.2-1.8 Å^{-1} [FIG. 1(a))] between model *ED_MID* and the most polar model, *ED_Cl*. The first maximum is much sharper for model *Cl* and consequently, $\chi^2_{F^N(Q)}$ is about three times larger than that for model *ED_MID* (TABLE IV, FIG. 1(a)). Both $F^N(Q)$-s coming from the MD simulations deviate, to some extent, from the experimental data set around the third maximum. On the whole, however, it is fair to say that the

performance of MD matches our most optimistic expectations – and this statement is valid for X-ray diffraction, as well (see below).



(a) (b)

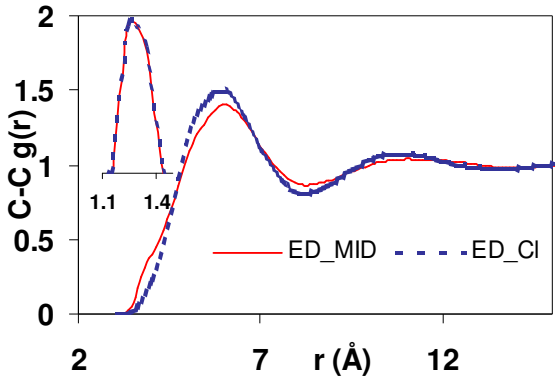
FIG. 1. (a) The renormalized experimental neutron and the *ED_MID* and *ED_Cl* MD neutron weighted total scattering structure factors. (b) The same for the X-ray data sets.

Concerning the X-ray total scattering data, results from the *ED_ MID* charge set also differ from those obtained for *ED_Cl*, mostly around the first (pre-)peak, as it is seen in FIG. 1(b). The *ED_Cl* simulations have higher χ^2 values, similarly to the $F^N(Q)$ data sets, although to a much lesser extent; the origin of the deviation is mainly the difference at the pre-peak and the region around the first minimum (see FIG. 1(b)). For the X-ray data, however, the *ED_MID* model produced considerably lower $\chi^2_{F^X(Q)}$ values, thus making also the sum $\chi^2_{F^N(Q)} + \chi^2_{F^X(Q)}$ lowest for this simulation.

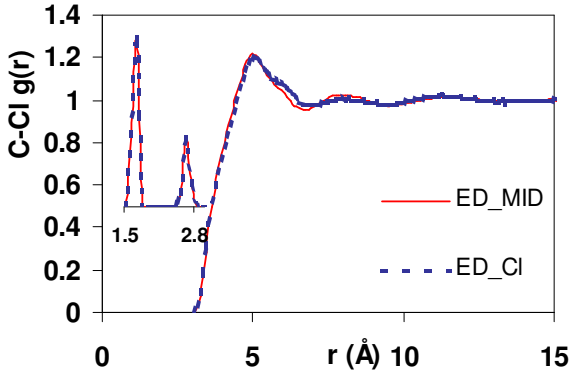
Now let us scrutinize the partial radial distribution functions, in an attempt of identifying the origin of differences between the *ED_MID* and *ED_Cl* models. As it is clear from FIG. 2(a)-(c), only the C-C and C-Cl partials differ notably, over the mid- and longer distance ranges of the intermolecular correlation range. These deviations in *r*-space are manifest around the first maximum of the C-C and C-Cl partial structure

factors, at Q=~1.3 Å^{-1}, as shown in FIG. 2(d). These are the obvious causes of the differences around the first maxima of the neutron and X-ray weighted total scattering structure factors. The normalized coefficients of the partial structure factors for neutron scattering are 0.0663, 0.382 and 0.551 for the C-C, C-Cl and Cl-Cl partials, respectively, and 0.0205, 0.245 and 0.734 at the first maximum, Q=~1.3 Å^{-1}, for X-ray diffraction. Thus it is most probably the C-Cl partial that is responsible for the differences between *tssf*-s (the C-C partial has far lower weight). The second maximum (at Q=~2 Å^{-1}) of the *tssf*-s contains contributions mainly from the Cl-Cl partial; this is why there is no difference here. In conclusion, the difference in terms of the charge distribution between the less polar *ED_MID* model (q_C=0.002-0.1) and the most polar *Cl* model (q_C=0.24) did not cause differences in terms of the intramolecular *prdf*-s, but considerably affected the intermolecular structure.

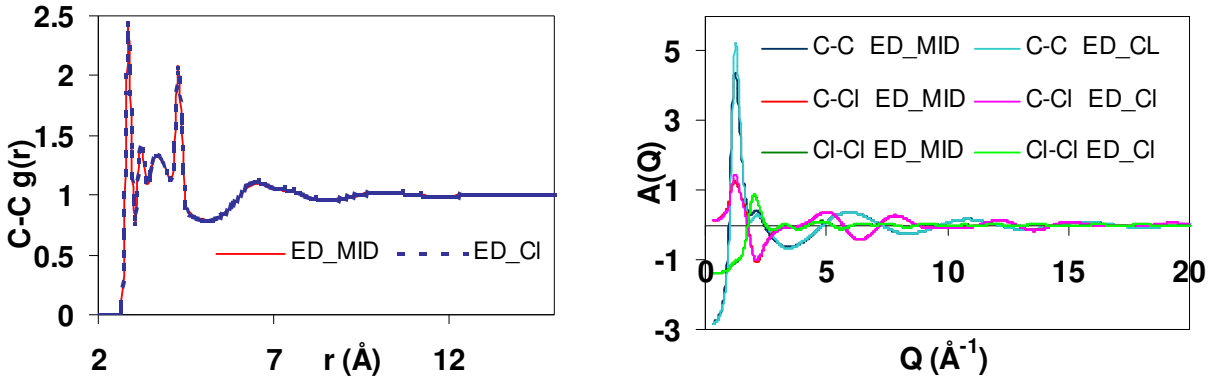
On the other hand, it has to be mentioned, that the difference between the *ED* and *ORI* parameter sets is solely manifesting in the intramolecular part, which is not surprising since the differences between the two sets lie in the bond length and angle parameters (for details, see the corresponding 'Supplementary information').



(a)



(b)



(c) (d)

FIG. 2. (a) C-C, (b) C-Cl, and (c) Cl-Cl partial *rdf*-s for the *ED_MID* and *ED_Cl* MD simulations. (d) Partial structure factors for the *ED_MID* and *ED_Cl* simulations. In parts (a) and (b) the main part of the figure shows the middle and long range distance ranges of the *prdf*, while the inset displays the short range parts.

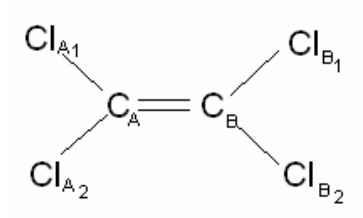


FIG. 3. The labeling of atoms in the C_2Cl_4 molecule.

The planarity of the molecules was examined by calculating the normal vectors of planes defined by atoms (Cl_{A1}, Cl_{A2}, C_B) and (Cl_{B1}, Cl_{B2}, C_A) for each molecule and the angle between the two normal vectors was then determined (see the labeling in FIG. 3). The mean angle for the *ORI* and *ED* series was 6.1±3.2º and 6.5±3.4º, respectively. The average out-of-plane displacement, defined as the absolute value of atom C_A from plane (Cl_{A1}, Cl_{A2}, C_B) and atom C_B from plane (Cl_{B1}, Cl_{B2}, C_A), was 0.052±0.001 Å for the *ORI* and 0.057±0.002 Å for the *ED* series, so the molecules remained 'sufficiently flat' during the simulations.

A.2. **Orientational correlations**

First the distribution of the centers of the molecules (determined as the average position of the six atoms making up the molecule) was calculated for each MD simulation. Here only results for *ED_MID* and *ED_Cl* MD simulations will be discussed; results for the other models can be found in the corresponding 'Supplementary Information'. There is a visible difference between the two curves over the mid- and longer distance ranges, see FIG. 4. The most interesting feature is the presence of a shoulder around 3.9-4.1 Å for the *MID* charge sets; this shoulder does not appear for the *Cl* charge sets (see the inset of FIG. 4). For the less polar model some of the molecules tend to aggregate close to each other, thus creating this shoulder. As the polarity increases, the shoulder disappears for the *Cl* charge set model. Taking into account the fact that the worst agreement between MD and experiment was consistently found for the '*Cl*' charge set, it may be concluded that experimental data require the presence of this shoulder on the center-center distribution function. The orientation of the molecular pairs forming the shoulder will be investigated below.

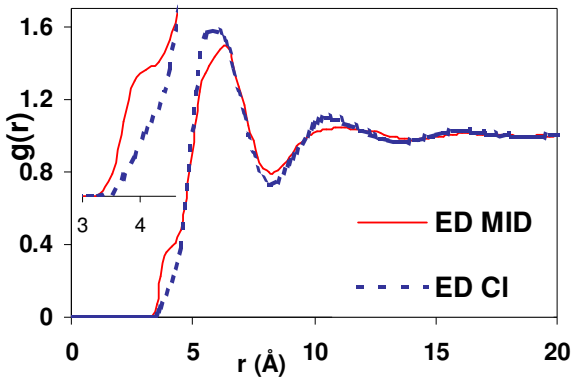


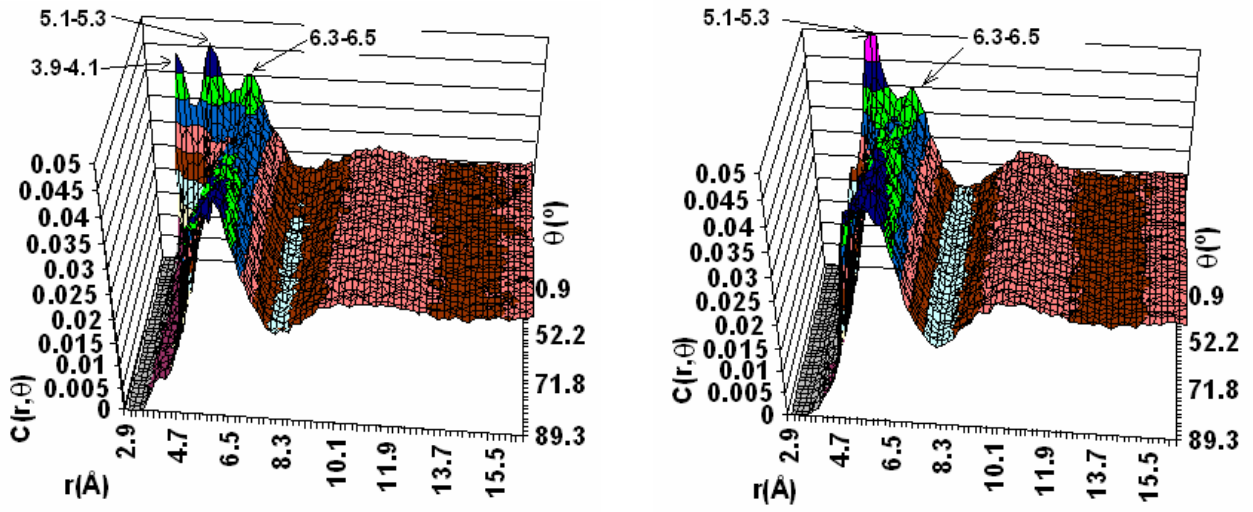
FIG. 4. Distribution of the centers of molecules. The inset (enlarged) shows the region of the shoulder discussed in the text.

The molecular axis defined by the C_A-C_B bond was determined for each molecule and the distribution $C(r,\theta)$ of the molecular pairs depending on the angle confined by the molecular axes and on the distance between the molecular centers have been calculated. The relationship between $C(r,\theta)$ and the center of molecules distribution, $g(r)$, is the following: $g(r) = \sum_{\theta=0}^{90} C(r,\theta)$.

The axis distribution for the *ED_MID* distribution is displayed in FIG. 5(a). The region around the first peak of the center of mass distribution, up to 8 Å, has a complex structure; beyond this distance, all angles occur with the same probability for any distance. Two smaller maxima, around 11 and 16 Å, can also be found; there are no other long-range correlations between molecular axes. Considering the structure of the first peak, the shoulder (shown in FIG. 4.) that appears for the less polar models solely comes from molecular pairs with near parallel axes, as indicated by the '3.9-4.1' (Å) label in FIG. 5(a). The figure also reveals that two maxima, labeled '5.1-5.3' and '6.3-6.5' (Å), at around zero degree, merge to form the asymmetric first maximum shown in FIG. 4. The ridge around 6.3 Å, representing molecular pairs with 0-90 ° angles between molecular axes, is also contained in the very broad first peak of the center-center distribution function.

In case of the *ED_Cl* parameter set the symmetric first maximum of the center of molecules distribution is caused by a higher peak at an angle of ~0 ° and distances of 5.1-5.3 Å, and by the ridge belonging to angles from 0 to 90 ° at around 6.3-6.5 Å (see FIG. 5(b) for *ED_Cl*). That is, the molecules closest to each other even in this case prefer the parallel arrangement of the molecular axes, they are just situated slightly further apart

from each other. Comparing the ridges making up the majority of the first coordination sphere of the molecules for the six less polar models, there is a preference for larger than 60º angles and to a slightly lesser extent, for molecular axis orientations of angles smaller than 30º. The mid-angle range is only slightly less populated. For the *Cl* model, a similar observation can be made but the effect is less visible.



(a) (b)

FIG. 5. Angular distributions depending on the angle between molecular axes and on the distance between molecular centers for (a) the *ED_MID* and (b) the *ED_Cl* models. The numbers in the pictures are in Å, indicating the position of the maxima. The coloring of the distribution is as follows:0-0.005: gray, 0.005-0.01: magenta, 0.01-0.015 pale yellow, 0.015-0.02 turquoise, 0.02-0.025 brow, 0.025-0.03 brick-red, 0.03-0.035 middle blue, 0.035-0.04 green, 0.04-0.045 dark blue and 0.045-0.05 pink.

The molecular plane for each molecule was determined by a least square fit to all the six atoms of the molecule. The average out of plane distance for the atoms from the molecular plane is 0.056±0.042 Å for the *ORI* and 0.058±0.044 Å for the *ED* series. Then the distribution $C(r,\varphi)$, depending on the angle confined by the molecular planes and on the distance between molecular centers, has been calculated. [The relationship between $C(r,\varphi)$ and the center of molecules distribution, *g(r)*, is similar to the case of the axis distribution: $g(r)=\sum_{\varphi=0}^{90} C(r,\varphi)$].

Similarly to the axis distributions, the plane distributions of the *ED_MID* and *ED_Cl* simulations are remarkably different, as it can be seen in FIG. 6(a) and (b). The peak structure is less complicated than for the axis distributions: there is only one maximum, similarly located for the two charge sets. For *ED_MID* the maximum peak height is at *r*=3.9 Å and for *ED_Cl* it is at r=4.1 Å, putting them into the 'shoulder region' (which region is clearly displayed only by the less polar models) of the center of molecules distribution. The peak heights, on the other hand, are very different: the *MID* charge set brings about more than twice the peak height of the *Cl* charge set. In both cases the peak is around the angle 0º, indicating parallel orientation of the molecular planes. The small peak amplitude for the *Cl* charge set models is not sufficient for a shoulder to grow on the main peak of the center of molecules distribution. Looking at the ridge corresponding to the first peak of the center of molecules distribution around 6.5 Å, we can see in both cases that there is a preference for larger angles between the molecular planes, ~90 º; in other words, perpendicular arrangement appears to be the most frequent ones at these distances.

There is no maximum around 5.1 Å as there was in case of the axis distributions; if the cross section of $C(5.1, \varphi)$ is examined closely then a slight preference for smaller than 20º angles can be observed. For larger center of molecule distances, there is no significant preference for any angles between the planes.

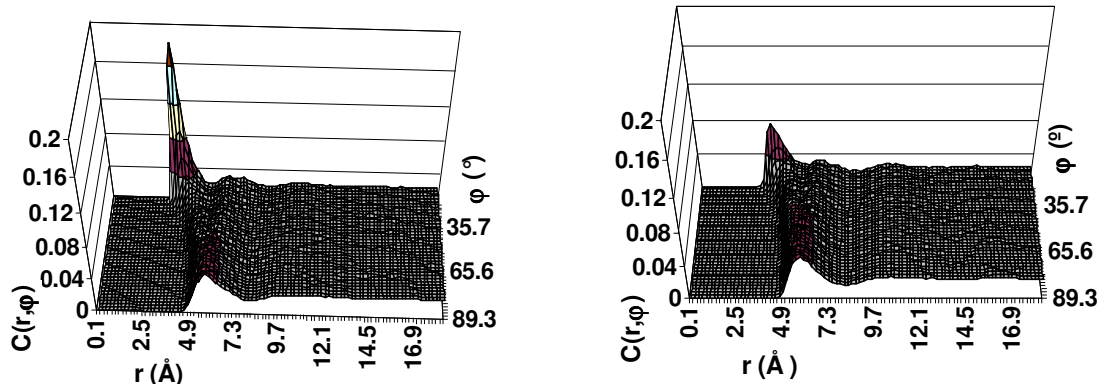


FIG. 6. The molecular plane distribution, depending on the angle confined by the molecular planes and on the distance between molecular centers for (a) the *ED_MID* and (b) the *ED_Cl* models. The coloring of the distribution is as follows:0-0.04: gray, 0.04-0.08: magenta, 0.08-0.12 pale yellow, 0.12-0.16 turquoise, 0.16-0.2 brown.

Summarizing the findings from molecular dynamics simulations concerning the intermolecular structure of liquid C_2Cl_4 and concentrating mainly on the *ED_MID* model (overall best agreement with experimental data), well-defined first and second, and a negligible third coordination shells are seen in the centers of molecules distributions. The first coordination shell itself has a complex structure: a shoulder appears that represents molecules closest to each other, around 3.9-4.1 Å, with a distinctive preference for parallel molecular axes and planes. Similar preference for parallel molecular axes and planes in case of trans-1,2-dichloro-ethylene was found in the work of Rovira-Esteva[47]. Then follows another layer around 5-1-5.3 Å, belonging to the first portion of the first peak of the center of molecules distributions, where there is a preference for parallel axes, and a weak preference for <20 ° angles between the molecular planes. This region is followed by the majority of the molecules belonging to the first maximum of the center-center distribution, where there is a slight preference for angles 0-30 and 60-90º for the

molecular axes, and larger than 60º angles for the molecular planes. For larger distances, no preference either for the molecular axes or the plane orientations can be found.

B. Reverse Monte Carlo modeling with the 'Flexible Molecules via Potentials' approach (FMP-RMC)

The initial particle configurations for the Reverse Monte Carlo refinements, by which we wished to explore whether (the already rather good) structural data provided by MD can be brought closer to diffraction results, were selected based on the MD data.

According to TABLE IV, the most successful MD simulations from a structural point of view are *ORI_C* from the *ORI*-series and *ED_MID* from the E-series (also the overall best). Therefore FMP-RMC simulations starting from both the *ED_MID* and the *ORI_C* MD simulations were performed.

B.1. General results – molecular structure and 2-body correlations

Similarly to our previous works[34,35], first only the scattering data were used as data set constraint in the FMP-RMC simulations, producing models *ORI_fq_sq* and *ED_fq_sq* (see TABLE III). The χ^2 values for the experimental data sets were calculated according to equation (1), the χ^2 values for the *prdf*-s and the cosine distributions of bond angles (in comparison with the MD average curves) have also been determined, in order to be able to analyze differences between resulting particle configurations; the χ^2 values are provided by TABLE V. In the table, the χ^2 values for the final configurations of the MD simulations, used as the starting configuration of the FMP_RMC calculations, are also shown for comparison.

TABLE V. The total χ^2 and its components (calculated with σ=1 for the $F^N(Q)$, $F^X(Q)$ experimental and for $g(r)$ 'quasi-experimental' (i.e., MD) data sets and for the cosine distributions of bond angles) characterizing the FMP-RMC calculations and the final MD configurations. The total potential energy is also given. The overall best FMP-RMC model is highlighted by bold characters.

	χ^2 (σ=1)					Total pot. E
	$F^N(Q)$	$F^X(Q)$	Total $g(r)$	Total *cos*	Total	(kJ/mol)
ORI_C	0.02	0.34	0.15	0.89	1.4	-23.5
ORI_fq_sq	0.01	0.08	13.9	155.5	169.5	-29.4
ORI_fq_sq_gr	0.01	0.15	2.1	25.7	28.0	-27.7
ED_*MID*	0.03	0.18	0.2	1.9	2.3	-24.0
ED_fq_sq	0.01	0.09	221.2	15.5	236.8	-26.6
ED_fq_sq_gr	**0.01**	**0.1**	**0.96**	**6.1**	**7.2**	**-24.0**

As it can be seen from the decreasing $\chi^2_{F^N(Q)}$ and $\chi^2_{F^X(Q)}$ values the agreement between experimental data and the model configurations has improved, the RMC 'refinement' was thus justified (cf. FIG. 7).

We now compare the MD and FMP-RMC structures in detail. We would like to establish (a) what really has changed during the RMC refinement; and (b) which is the structure that is in best possible agreement with diffraction data and at the same time, the closest to the MD final particle arrangement (i.e., a structure that is more ordered than the ones obtainable by standard RMC modeling)? In order to address these issues, FMP-RMC runs were performed that attempted to match *prdf*-s from MD simulations together with the two sets of experimental data. The first thing to notice is that if only the *tssf*-s are modeled (*ORI_fq_sq* and *ED_fq_sq*) the total χ^2 is considerably increased. This is caused mainly by distortions of the cosine distributions for the *ORI_C* and by distortions of the

prdf–s for the *ED_MID* series. 'Distortion' mainly means the heightening/sharpening of some of the peaks; the general shape of the distributions does not change (not shown).

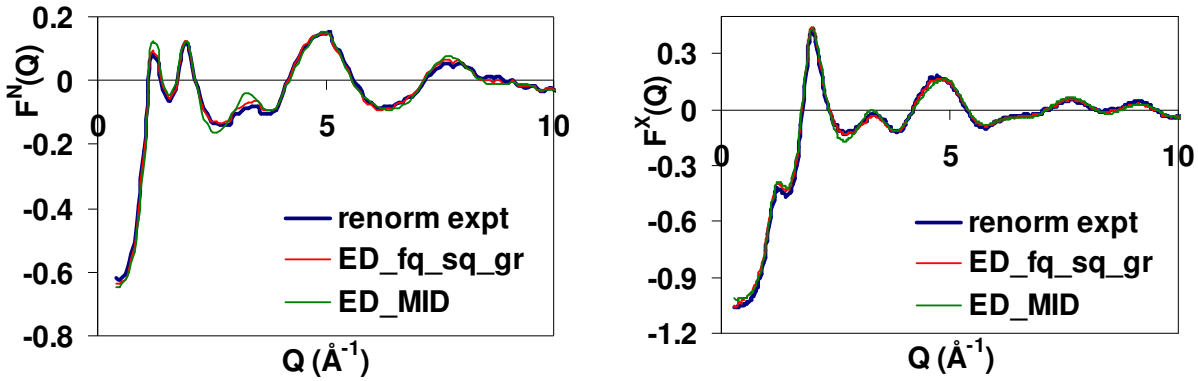
Next, two further FMP-RMC calculations, *ORI_fq_sq_gr* and *ED_fq_sq_gr* (see TABLE III) have been performed in order to see whether it is possible to create a configuration with good agreement with the experimental data and fully compatible with the MD *prdf*-s. In both the '*ORI*' and the '*ED*' cases agreement with MD *prdf*-s and cosine distributions has improved considerably (in comparison with the '*_fq_sq*' runs), although for the calculation *ORI_fq_sq_gr* the agreement with the X-ray data has deteriorated noticeably. Considering all these observations, we think that the calculation *ED_fq_sq_gr* produces the 'best' result (i.e., most ordered while fitting *tssf*-s prefectly), judging also by the low total χ^2 value (TABLE V).

It is of some interest to look at calculations with somewhat higher $\chi^2_{g(r)}$ value, corresponding to the *ORI_fq_sq_gr* model: the origin is mainly the Cl-Cl partial *rdf*, as it is evidenced by FIG. 8(a). While in case of the *ED_fq_sq_gr* model there is only a slight change in terms of the peak heights, for the *ORI_fq_sq_gr* model there are visible distortions of the peak shape.

Now let us turn our attention to the cosine distributions of bond angles. It has to be noted that no explicit cosine distribution constraint was used during the FMP-RMC calculations, only the harmonic angle bending potential kept the distributions in their desired shape. TABLE V shows total χ^2_{cos}: we note that for the *ORI_fq_sq_gr* simulation the Cl-C-C angle was the major contributor, with $\chi^2_{Cl\text{-}C\text{-}C}$ =16.6, while in case of the *ED_fq_sq_gr* the majority of the difference came from the Cl-C-Cl angles ($\chi^2_{Cl\text{-}C\text{-}Cl}$ =5.1). These distributions, in comparison with the respective MD average distributions, are

displayed in FIG. 8(b); in the case of *ED_fq_sq_gr* the peak kept its Gaussian shape, while in case of *ORI_fq_sq_gr* even the shape of the Cl-C-C peak was distorted while it moved to larger angles.

The total potential energy (see TABLE V) decreased slightly for the *ORI_fq_sq_gr* model, from –23.5 to –27.7 kJ/mol, while it remained unchanged for the *ED_fq_sq_gr* model, at -24.0 kJ/mol.



(a) (b)

FIG. 7. Calculated and experimental (a) neutron and (b) X-ray total scattering structure factors for the best FMP-RMC simulation, *ED_fq_sq_gr*, as compared to the *ED_MID* MD averages.

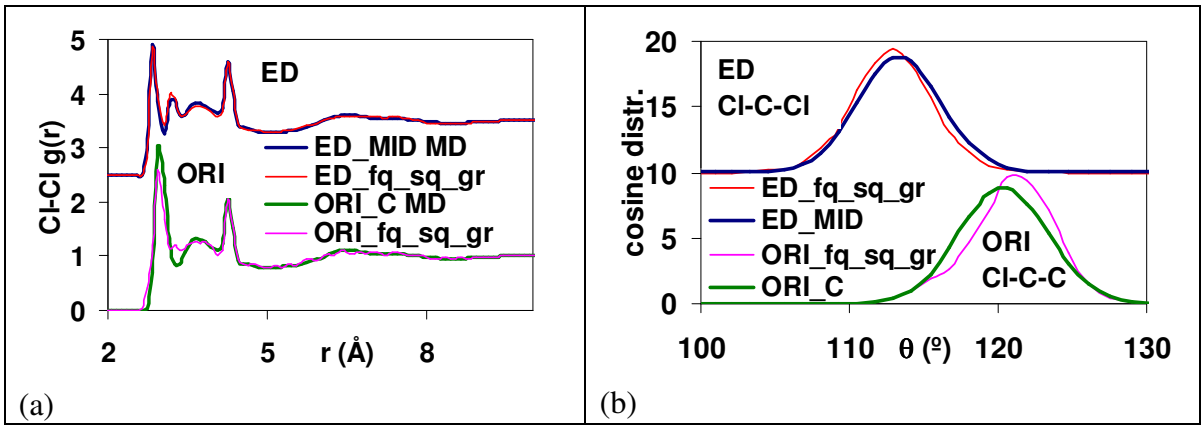


FIG. 8. (a) The Cl-Cl *prdf*-s for the *ED_fq_sq_gr* model, as compared to the *ED_MID* MD average (shifted along the *y*-axis by 2.5) and for the *ORI_fq_sq_gr* model, as compared to the *ORI_C* MD average. (b) Upper part: Cl-C-Cl cosine distribution of bond angles for the *ED_fq_sq_gr* simulation, in comparison with the *ED_MID* MD average (shifted along the *y*-axis by 10); lower part: Cl-C-C cosine distribution of bond angles for the *ORI_fq_sq_gr* simulation, as compared to the *ORI_C* MD average. (Note that θ is shown on the *x*-axis, instead of $\cos(\theta)$.)

The planarity of molecules was examined for the best FMP-RMC calculation *ED_fq_sq_g(r)*, as previously, by calculating the normal vectors of the planes defined by the atoms (Cl_{A1}, Cl_{A2}, C_B) and (Cl_{B1}, Cl_{B2}, C_A), and the angle between the two normal vectors. The mean angle for the *ED_fq_sq_gr* series decreased from the 6.5±3.4º of the *ED_MID* MD simulation to 5.2±2.9º, thus making the molecules more planar-like. The average out-of-plane displacement of the absolute value of atom C_A from plane (Cl_{A1}, Cl_{A2}, C_B) and of atom C_B from plane (Cl_{B1}, Cl_{B2}, C_A) also decreased from 0.057±0.002 Å of the *ED* MD series to 0.049±0.001 Å. This shows that the FMP-RMC algorithm is capable of keeping the molecules planar during RMC modeling.

B.2. **Orientational correlations**

The center of molecules distribution was calculated as previously for the MD configurations, see FIG. 9. The most important observation is that the characteristic shoulder on the low-*r* side of the first maximum still exists.

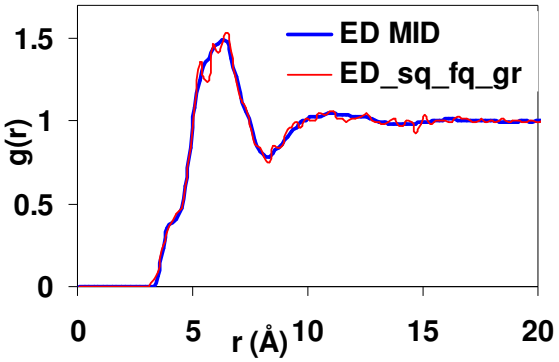


FIG. 9. The center of mass distribution function for the *ED_MID* MD average and the *ED_fq_sq_gr* FMP-RMC calculation.

The molecular axis for each molecule was determined and $C(r,\theta)$, the distribution of angles between pairwise corresponding molecular axes, was calculated as a function of the distance between molecular centers. The basic features are very similar to what were

found for the corresponding *ED_MID* MD average; however, as the distribution in this case is calculated only for the final configuration of the RMC calculation, comparison was made also to the final *ED_MID* MD configuration (that served as a starting configuration for the RMC refinement). Only the cross section at θ~0 ° is displayed in FIG. 10; the first peak, belonging to the shoulder region of the center of molecules distribution, has moved to a slightly higher distance of 4.3 Å, and an extra maximum ('pre-peak') at 4.7 Å appeared, preceding the peak at 5.1 Å.

The ridge region, around 6-3-6.6 Å, remained unchanged, showing a slight preference for the <30 ° or >60 ° angles; there are no orientational preferences at larger distances. All these features point to the fact that the molecules closest to each other prefer the parallel arrangement of their molecular axes.

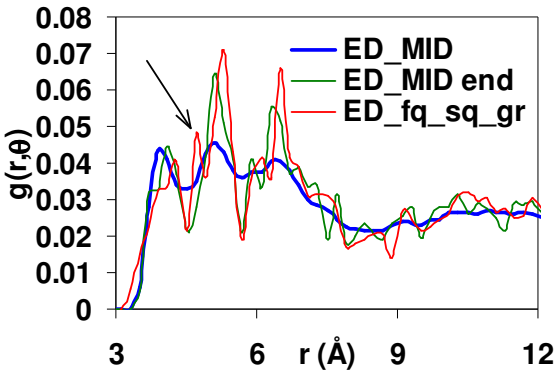


FIG. 10. Cross section of the $C(r,\theta)$ distribution for the *ED_MID* MD simulations averaged over 76 configurations (blue), for its last configuration (olive) and for the final configuration of *ED_MID_fq_sq_gr FMP-RMC* simulation (red) at θ~0 °. The arrow indicates the new maximum at 4.7 Å.

The best-fitting planes to all six atoms of the molecules were determined as previously. The average out-of-plane displacement was found to be 0.045±0.035 Å, smaller than that for the *ED* MD series (see III.A.2.). Considering the orientations of the

molecular planes, the correlation function $C(r,\varphi)$, calculated as described before, was determined. The overall shape of the distribution has not changed compared to the starting MD simulation *ED_MID*: only the position of the peak has moved towards shorter distances, from 3.9 to 3.7 Å, still being in the shoulder region of the center of molecules distribution. The orientation of the planes remained parallel (ϕ~0 º).

IV. SUMMARY AND CONCLUSIONS

The liquid structure of tetrachloroethene was investigated in a series of MD simulations, using various interaction potential parameter sets and charge distributions (see TABLE I), since no entirely appropriate parameterization is available in the literature. We note that potential parameters for the *ED_MID* simulation have partly been introduced by the present work. The efficiency of the simulations was judged by the agreement of the calculated average MD structure factors with the experimental neutron and X-ray total scattering structure factors.

The parameter set producing the overall closest match with the experimental structure factors was *ED_MID*; from the *ORI* parameter set series the *ORI_C* simulation proved to be the most successful one.

FMP-RMC simulations were started from the end configurations of the *ED_MID* and *ORI*_C MD runs, which were already quite close to experiment. The molecular structure (and also energies, *prdf*-s and cosine distributions) stayed close to the MD averages. This was the first time that the recently developed FMP-RMC method, where the molecules are kept together by bonded and non-bonded potentials, was applied to a planar molecule. The planarity of the molecules was assessed in different ways, described in sections III.A.1 and III.A.2. It was found that the planarity of the molecules has improved during

the FMP-RMC simulations. This clearly indicates that the FMP-RMC approach can be successfully applied for planar structures.

For the *ORI_fq_sq_gr* calculation, improving the agreement with the *tssf*-s while maintaining the MD average *prdf-s* and cosine distribution of bond angles was not entirely possible. On the other hand, the FMP-RMC calculation started from the *ED_MID* simulation and using experimental *tssf*-s and the *ED_MID* MD average *prdf*-s resulted in a favorable structure. In the *ED_fq_sq_gr* model differences between experiment and model decreased, while *prdf*-s and cosine distributions of bond angles have hardly changed. This indicates that the *ED_MID* parameter set provides a proper description for the liquid structure of C_2Cl_4.

Orientational correlations, describing preferred relative arrangements of molecules as a function of the distance between their centers, have also been calculated for the *ED_MID* MD simulation and for the *ED_fq_sq_gr* FMP-RMC calculation. The molecules closest to each other, at ~4.3 Å, form a shoulder on the low-*r* side of the center of molecules distribution. These nearest neighbors exclusively prefer a parallel arrangement of their axes and molecular planes. The first maximum of the center-center distribution appears around 6.5 Å. All possible arrangement of molecular axes and molecular planes are found in this region; there is, however, a preference for <30 º and >60 º angles for the molecular axis orientation, and a slight preference for the >60 º molecular plane angle.

Acknowledgment

Financial support for this work was provided by the Hungarian Basic Research Fund (OTKA) via grant No. 083529.

See supplementary material at [***URL***] for more details on the MD simulation and the on the behavior of the different charge set.

References

[1] http://www.hmlindia.com/Specifications/Perchloroethylene.pdf

[2] http://water.epa.gov/drink/contaminants/index.cfm.

[3] http://ephtracking.cdc.gov/showPerchloroetheneHealth.action

[4] D-H. Lim, C. M. Lastoskie, A. Soon, U. Becker Envir Sci. Technol. **43**, 1192, (2009).

[5] I. L. Karle, J. Karle, J. Chem. Phys. **20**, 63, (1952).

[6] M. Alvarez, F. J. Bermejo, P. Chieux, E. Enciso, J. Alonso and N. García, J Phys.: Condensed Matter **1,** 8595, (1989).

[7] L. S. Bartell, E. A. Roth, C. D. Hollowell, Kozo Kuchitsu, J. E. Young J. Chem. Phys. **42**, 2683 (1965).

[8] A. K. Adya, L. Bianchi, C. J. Wormald, J. Chem. Phys. **112**, 4231, (2000).

[9] T. Takamuku, K. Saisho, S. Nozawa, T. Yamaguchi J. Mol. Liq. **119**, 133, (2005).

[10] K. M. Murdoch, T. D. Ferris, J. C. Wright, and Thomas C. Farrar J. Chem. Phys. **116**, 5717 (2002).

[11] L. Nemes, S. Suzuki J. Raman Spectrosc. **2**, 193, (1974).

[12] S. Kawanishi, T. Sasuga, and M. Takehisa, J. Phys. Soc. Jpn. **50**, 3080 (1981).

[13] W. J. Hehre J. Chem. Phys. **63,** 4750 (1975).

[14] M. J. G. Peach, F. de Proft, D. J. Tozer, J. Phys. Chem. Lett. **1**, 2826 (2010).

[15] L. J. Lowden, P. Chandler, J. Chem. Phys. **61**, 5228, (1974).

[16] M. Rovira-Esteva, N. Arul Murugan, L. C. Pardo, S. Busch, J. Ll. Tamarit, G. J. Cuello, F. J. Bermejo J. Chem. Phys 136, 124514 (2012).

[17] M. Tomsic, A. Jamnik, G. Fritz-Popovski, O. Glatter, L. Vlcek, J. Phys. Chem . B **111**, 1738, (2007).

[18] H. Bertagnolli, D. O. Leicht, M. D. Zeidler, Mol. Phys. **35**, 199, (1978).

[19] H. Bertagnolli, D. O. Leicht, M. D. Zeidler, P. Chieux, $CD^{35}Cl_3$ and $CD^{37}Cl_3$, Mol. Phys. **36,** 1769, (1978).

[20] H. Bertagnolli, D. O. Leicht, M. D. Zeidler, Mol. Phys. **35**, 193, (1978).

[21] H. Bertagnolli Chem. Phys. Letters **93**, 287 (1982).

[22] A. K. Soper, J. Mol. Liquids 78, 179, (1998).

[23] L. C. Pardo, J. L. Tamarit, N. Veglio, F. J. Bermejo, G. J. Cuello Phys. Rev. B **76**, 134203, (2007).

[24] I. Waldner, A. Bassen, H. Bertagnolli, K. Tödheide, G. Strauß, A. K. Soper, J. Chem. Phys. **107**, 10667, (1997).

[25] M. Rovira-Esteva, N. A. Murugan, L. C. Pardo, S. Busch, J. Ll. Tamarit, Sz. Pothoczki, G. J. Cuello, F. J. Bermejo, Phys. Rev. B **84**, 064202 (2011).

[26] R. L. McGreevy and L. Pusztai, Molec. Simul. **1**, 359, (1988).

[27] R. L. McGreevy, J. Phys.: Cond. Matter **13**, R877, (2001).

[28] G. Evrard and L. Pusztai, J. Phys: Condens. Matter **17**, S1,(2005).

[29] A. Vrhovšek, O. Gereben, A. Jamnik and L. Pusztai, J. Chem Phys. B **115**, 13473, (2011).

[30] V. Mile, O. Gereben, S. Kohara and L. Pusztai, J. Mol. Liq. **157**, 36, (2010).

[31] T. Fukunaga, K. Itoh, T. Otomo, K. Mori, M. Sugiyama, H. Kato, M. Hasagawa, A. Hirata, Y. Hirotsu and K. Aoki, Physica B: Condensed Matter **385**, 259, (2006), DOI:10.1016.

[32] O. Gereben and L. Pusztai, Phys. Rev. B **50**, 14136, (1994).

[33] O. Gereben and L. Pusztai. Mat. Science and Eng. **A178/A180**, 433, (1994).

[34] O. Gereben and L. Pusztai, J. of Comp. Chem. **33**, 2285 (2012).

[35] O. Gereben and L. Pusztai, J. Phys.: Condens. Matter, **25**, (2013) in press.

[36] http://www.szfki.hu/~nphys/rmc++/opening.html.

[37] http://www.gromacs.org ; D. van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark and H. J. C. Berendsen, J. Comp. Chem. **26**, 1701, (2005).

[38] S. Nose, Mol. Phys., **52**, 255, (1984); W. G. Hoover, Phys. Rev. A **31**, 1695, (1985).

[39] F. Serrano, A. Bańon and J. Santamaría, J. Chem. Phys Lett. **107**, 475, (1984).

[40] R. D. Mountain and G. Morrison, *Mol. Phys.* **64**, 91, (1988).

[41] W. L. Jorgensen, D. S. Maxwell and Tirado-Rives, J. Am. Chem. Soc. **118**, 11225, (1984).

[42] GROMACS manual; http://www.gromacs.org/@api/deki/files/190/=manual-4.5.6.pdf

[43] P. Jóvári, Gy. Mészáros, L. Pusztai and E. Sváb, Physica B **278**, 491, (2000).

[44] E. Sváb, Gy. Mészáros and F. Deák, Mat. Sci. Forum, **247**, 228, (1996).

[45] L. Pusztai, S. Kohara, unpublished results (2006).

[46] S. Kohara, M. Itou, K. Suzuya, Y. Inamura, Y. Sakurai, Y. Ohishi and M. Takata, J. Phys.: Condens. Matter **19**, 506101, (2007).

[47] M. Rovira-Esteva, A. Murugan, L. C. Pardo, S. Busch, M. D. Ruiz-Martín, M.-S. Appavou, J. Ll. Tamarit, C. Smuda, T. Unruh, F. J. Bermejo, G. J. Cuello, S. J. Rzoska, Phys. Rev. B **81**, 092202, (2010).