Chinese Journal of Chemical Engineering 36 (2021) 138–145

Contents lists available at ScienceDirect

Chinese Journal of Chemical Engineering

journal homepage: www.elsevier.com/locate/CJChE

Full Length Article

Testing and validation of a self-diffusion coefficient model based on molecular dynamics simulations

Xia Chen [1], Yan Wang [2], Lianying Wu [1,*], Weitao Zhang [1], Yangdong Hu [1]

[1] College of Chemistry and Chemical Engineering, Ocean University of China, Qingdao 266100, China
[2] College of Chemistry and Chemical Engineering, Qingdao University, Qingdao 266071, China

ARTICLE INFO

Article history:
Received 22 October 2020
Received in revised form 2 April 2021
Accepted 7 April 2021
Available online 26 June 2021

Keywords:
Diffusion
Molecular simulation
Parameter identification
Characteristic length
Diffusion velocity

ABSTRACT

In our previous work, we endowed a new physical meaning of self-diffusion coefficient in Fick's law, which proposed that the diffusion coefficient can be described as the product of the characteristic length and the diffusion velocity. To testify this simple theory, in this work, we further investigated the underlying mechanism of the characteristic length and the diffusion velocity at the molecular level. After a complete dynamic run, the statistical average diffusion velocity and the characteristic length of molecules can be obtained by scripts, and subsequently the diffusion coefficient was determined by our proposed theory. The diffusion processes in 35 systems with a wide range of pressure and concentration variations were simulated using this model. From the simulated results, diffusion coefficients from our new model matched well with the experimental results from literatures. The total average relative deviation of predicted values with respect to the experimental results is 8.18%, indicating that the novel model is objective and rational. Compared with the traditional MSD-*t* model, this novel diffusion coefficient model provides more reliable results, and the theory is simple and straightforward in concept. Additionally, the effect of gas pressure and liquid concentration on the diffusion behavior were discussed, and the microscopic diffusion mechanism was elucidated through the distribution of diffusion velocity and the characteristic length analysis. Moreover, we suggested new distribution functions, providing more reliable data theoretical foundations for the future research about the diffusion coefficient.
© 2021 The Chemical Industry and Engineering Society of China, and Chemical Industry Press Co., Ltd. All rights reserved.

1. Introduction

Diffusion is the tendency of particles to spread into an available area. It plays an important role in every natural or industrial process involving mass transfer. Therefore, the determination of diffusion coefficient is of great importance for the calculation of mass transfer processes. There are two kinds of classical diffusion theories used to define diffusion coefficient: Maxwell-Stefan's (MS) equations and Fick's law [1–4]. The diffusion coefficient of the former has a definite physical meaning, which represents the friction factor between two components [5]. The latter is commonly used in mass transfer process with a simple expression, but the physical meaning for its diffusion coefficient is not clear [6]. While, in our previous work [7], we endowed a novel physical meaning of the diffusion-coefficient in Fick's law, in which the diffusion coefficient can be described as the product of the characteristic length and the diffusion velocity.

In many systems, diffusion coefficients are usually obtained through experimental approaches [8,9], such as conduct metric method, Taylor dispersion technique and digital holography method, *etc.* For example, Rodrigo *et al.* [10] acquired the binary diffusion coefficients of L-histidine methyl ester dihydrochloride solutions by using the Taylor dispersion technique. Zhou *et al.* [11] obtained the diffusion coefficient of the ethylene glycol–water system by employing digital holography method. Mattisson *et al.* [12] measured the diffusion coefficients for lysozyme in gels and liquids by using holographic laser interferometry. Jamnongwong *et al.* [13] studied the oxygen diffusion coefficients in clean water containing salt, glucose or surfactant based on measurements of volumetric mass transfer coefficients. Wang *et al.* [14] determined the gas diffusion coefficient in porous shale. Albeit these experimental methods are direct and easy to operate, the measurement of diffusion coefficients in certain systems is tedious, costineffective and even impractical.

To this end, computer simulation method has been introduced to obtain the diffusion coefficients. Nowadays, molecular dynamics (MD) simulation has become an effective tool to calculate the dif-

* Corresponding author.
E-mail address: wulianying@ouc.edu.cn (L. Wu).

https://doi.org/10.1016/j.cjche.2021.04.036
1004-9541/© 2021 The Chemical Industry and Engineering Society of China, and Chemical Industry Press Co., Ltd. All rights reserved.

fusion coefficients [15–20]. Various studies on the diffusion properties research have been carried out through MD simulation. For example, Yang *et al.* [21] investigated the diffusion of methanol/water mixture through MFI-type zeolite (HZSM-5 and silicalite-1) membranes by employing MD simulation. Hu *et al.* [22] and Zhang *et al.* [23] computed the diffusion of corrosive particle *via* MD simulation. Ghaffari *et al.* [24] explored the self-diffusion coefficients in NaCl aqueous solutions by using MD simulation. In recent study, Zhao *et al.* [25] calculated the transport diffusion coefficients of CO_2 and CH_4 in coal *via* the molecular simulation. Tsimpanogiannis *et al.* [26] presented a detailed discussion about self-diffusion coefficient of water by classical molecular simulation. Moultos *et al.* [27–32] reported a series of papers regrading to the systematic errors in the calculation of self-diffusion caused by the long-range interactions and the imposed periodic boundary conditions.

On the other hand, to study the underlying mechanism of the microscopic phenomena and to predict the diffusion coefficient, various physical and mathematical models have been proposed [33,34]. For instance, Kamgar *et al.* [35] suggested a simple model only with the molecular mobility factor and the thermodynamic correction factor. Considering an analogy between nonideal concentrated solutions and solutions near the consolute point, Cussler *et al.* [36] suggested to mimic the diffusion through the movement of entire molecules clusters.

Although the novel diffusion coefficient model was given in our previous work, we did not give the values of characteristic length and diffusion velocity. In this work, we further investigated the underlying mechanism of our previously proposed diffusion coefficient model at the molecular level. Through MD simulation, the statistical average diffusion velocity and the characteristic length were obtained, and subsequently the diffusion coefficient was given by our proposed model. Based simulated results of 35 systems, this novel diffusion coefficient model was proved to be objective and rational. Moreover, the effect of gas pressure and liquid concentration on the diffusion behavior were discussed, and the microscopic diffusion mechanism was elucidated through the distribution of diffusion velocity and the characteristic length analysis.

2. MD Simulation

2.1. Theory of the diffusion coefficient model

In our previous work, we developed a new diffusion coefficient model [7] by analyzing the unit of Fick's Law diffusion coefficient through the dimension analysis method. The proposed model contains two parts: the characteristic length item and the diffusion velocity item. The clear physical meaning of Fick's Law diffusion coefficient can thus be described as the product of the characteristic length and the diffusion velocity, as follows:

$$D_i = V_i \times L_i \tag{1}$$

where V_i denotes the diffusion velocity; L_i denotes the characteristic length.

V_i is the velocity of molecular diffusion, which can be considered as the statistical average velocity of the molecular motion. L_i is characteristic length, which is the statistical average of the diffusion distance. And the diffusion distance is the distance which traveled by a moving molecule without changing direction. This is similar to the meaning of the free path of gas molecules. To testify this simple theory, we calculated the characteristic length and the diffusion velocity *via* MD simulation.

2.2. Simulation method

In this study, all simulations were performed by means of COMPASS force field in Materials studio software. Each diffusion simulation was repeated more than 10 times to obtain good statistic. The initial configurations were built individually using the Amorphous Cell module. The appropriate length of the simulation box was given in the Tables 1–3 according to the concentration and density. The simulations were performed in a cubic box with periodic boundary conditions and different directions. The Nose thermostat and the Berendsen barostat was used to control the system temperature and pressure, respectively. Ewald method was used to simulate electrostatic and van der Waals interactions with a cut-off value of 15.5 Å (1 Å = 0.1 nm). For the liquid system, the configurations were firstly subjected to geometry optimization, and then run in NPT (1 ns), NVT (1 ns) and NVE (1 ns) ensemble. A small integration time step of 0.1 fs was selected to obtain more statistic samples. The trajectories of liquid molecules were saved every 10 steps for further analysis. For the gas system, the configurations which were built based on the gas density were firstly subjected to geometry optimization, and then run in NVT (8 ns) and NVE (8 ns) ensemble. For the gas system, the integration time step was 1 fs. The trajectories of gas molecules were sampled every 250 steps for further analysis.

MD results can provide the microscopic details of molecular diffusion to reveal the underlying mechanism of diffusion, such details cannot be directly obtained through experimentation. The diffusion velocity is actual the average velocity of the molecular motion, which can be considered as the statistic average velocity form the available statistic samples. After a complete dynamic run, the instantaneous velocity of each molecule per femtosecond directly obtained by scripts. Based on the available statistic samples, the statistical average diffusion velocity was given. Then V_i (the average velocity of the molecular motion) gained. The characteristic length denotes the statistical average of the diffusion distance which traveled by a moving molecule without changing direction. After a complete dynamic run, the position coordinates of each molecule per femtosecond (such as the path in Fig. 1) obtained. The deviation angle of ever frame can be calculated by the arccosine function. For example, as shown in the partial enlarged drawing of the Fig. 1b, the deviation angle of frame A (*i.e.*∠BAC) calculated by the arccosine function (Eq. (2)). We defined that the moving molecule change its direction when the moving direction deviating from a straight line exceeds 2 degrees. Thus, the distance between two adjacent frames that deviation angle exceeded 2 degrees (*e.g.* ∠BAC > 2°) is the diffusion distance. Then L_i (the statistical average of the diffusion distance) obtained.

$$\angle \mathrm{BAC} = \arccos(\cos\angle \mathrm{BAC}) = \arccos\left(\frac{|\mathrm{AC}|^2 + |\mathrm{AB}|^2 - |\mathrm{BC}|^2}{2|\mathrm{AC}| \cdot |\mathrm{AB}|}\right) \tag{2}$$

Through statistical analysis of the displacement and velocity data, the average diffusion distance and velocity can be obtained. And subsequently the diffusion coefficient was determined by our proposed model (Eq. (1)).

For comparison, we also calculated the diffusion coefficient by mean square displacement (MSD) approach. The MSD ($\langle r_{i^{(t)}} - r_{i^{(0)}} \rangle^2$) is the square of particle displacement during a time interval. Based on the well-known Einstein equation (Eq. (3)), the diffusion coefficient can be obtained.

$$D = \frac{1}{6N_\mathrm{a}} \lim_{t\to\infty} \frac{\mathrm{d}}{\mathrm{d}t} \sum_{i=1}^{N_\mathrm{a}} \langle r_{i^{(t)}} - r_{i^{(0)}} \rangle^2 \tag{3}$$

where N_a is the number of diffusion molecules, $r_{i^{(t)}}$ is the position vector of the molecule at time t, and $r_{i^{(0)}}$ is the initial position.

Table 1
The simulated results of various gas and organic vapor systems at 300 K and 1.01 MPa

Group	M/ g·mol^{-1}	Density/ g·L^{-1}	Cell size×10^{10}/ m	V_{cal}×10^{-2}/ m·s^{-1}	L_{cal}×10^{9}/ m	$D_{cal\text{-}LV}$×10^{5}/ m^2·s^{-1}	D_{exp}×10^{5}/ m^2·s^{-1}	RD$_{-1}$/ %	$D_{cal\text{-}MED}$×10^{5}/ m^2·s^{-1}	RD$_{-2}$/ %
N_2	28	1.13	345.27	6.98	34.83	2.43	2.12 [37]	14.65	2.32	9.27
CO	28	1.13	345.26	6.79	33.09	2.25			2.41	
O_2	32	1.29	345.21	6.10	33.97	2.07	2.23 [37]	7.11	2.53	13.35
CO_2	44	1.78	345.24	6.52	17.00	1.11	1.13 [37]	1.89	1.36	20.79
SO_2	64	2.58	345.32	5.60	11.60	0.65			0.75	
CH_4	16	0.65	345.53	8.21	31.62	2.59	2.40 [38]	8.12	3.22	34.03
C_2H_2	26	1.05	345.38	7.58	24.74	1.88	2.01 [38]	6.69	1.79	11.07
C_2H_4	28	1.13	345.44	7.47	20.97	1.57	1.68 [38]	6.70	1.74	3.71
H_2S	34	1.37	345.47	4.77	20.55	0.98			1.49	
C_3H_6	42	1.70	345.44	7.32	13.21	0.97			0.77	
C_6H_{12}	84	3.39	345.44	7.28	6.60	0.48			0.27	
C_6H_6	78	3.15	345.38	7.09	9.57	0.68			0.40	
Average								7.53		15.37

Table 2
The simulated results of the gas systems at 300 K for different pressures

Group	p/ kPa	Density/ g·L^{-1}	Cell Size×10^{10} / m	$D_{cal\text{-}MSD}$×10^{5}/ m^2·s^{-1}	V_{cal}×10^{-2} / m·s^{-1}	$L_{cal\text{-}LV}$×10^{9} / m	D_{cal}×10^{5} / m^2·s^{-1}
CO_2	25.33	0.44	548.03	4.80	6.48	82.88	5.37
	50.66	0.89	434.98	2.66	7.34	44.79	3.29
	101.33	1.78	345.21	1.49	6.52	17.00	1.11
	121.59	2.22	320.49	1.03	7.51	11.48	0.86
	151.99	2.66	301.60	0.87	7.98	9.83	0.78
N_2	25.33	0.28	548.09	9.04	6.80	132.45	9.00
	50.66	0.57	435.02	4.89	6.80	65.49	4.45
	101.33	1.13	345.27	2.32	6.98	34.83	2.43
	121.59	1.41	320.52	1.81	6.59	28.08	1.85
	151.99	1.70	301.62	1.49	6.72	22.03	1.48
O_2	25.33	0.32	547.99	9.00	7.16	123.07	8.82
	50.66	0.65	434.94	4.72	6.32	69.66	4.40
	101.33	1.29	345.21	2.53	6.10	33.97	2.07
	121.59	1.61	320.46	1.86	6.21	27.96	1.74
	151.99	1.94	301.57	1.55	6.00	24.63	1.48

Table 3
The simulated results of different liquid systems at 300 K and 1.01 MPa

Group	c/ mol·L^{-1}	Density ×10^{-3}/ g·L^{-1}	Cell Size ×10^{10} / m	D_{YH}×10^{10} / m^2·s^{-1}	V_{cal}×10^{-2}/ m·s^{-1}	L_{cal}×10^{12}/ m	$D_{cal\text{-}LV}$×10^{10}/ m^2·s^{-1}	D_{exp} [39]×10^{10}/ m^2·s^{-1}	RD$_{-1}$/ %	$D_{cal\text{-}MSD}$×10^{10} / m^2·s^{-1}	RD$_{-2}$/ %
Alanine	0.59	1.03	24.31	0.29	7.12	12.53	9.21	8.68	5.73	16.40	47.07
	1.77	1.05	25.00	0.14	6.92	11.06	7.79	7.79	0.06	9.46	17.61
	3.01	1.02	25.22	0.34	6.88	10.19	7.36	7.20	2.22	9.10	20.91
	3.93	1.02	26.04	0.75	6.85	9.76	7.44	6.63	10.93	6.61	0.25
	5.42	1.03	27.15	0.49	7.09	8.43	6.47	5.60	13.40	5.87	4.55
Glycine	0.5	1.01	23.86	0.59	6.28	16.63	11.03	9.73	11.79	14.94	34.88
Threonine	0.5	1.02	23.98	2.15	7.61	9.54	9.41	7.15	24.00	12.97	44.86
Urea	0.5	1.02	23.73	0.30	6.5	21	13.94	13.43	3.65	25.86	48.06
L-Arginine	0.5	1.04	24.00	0.23	7.4	8.34	6.40	6.04	5.63	5.17	16.83
Leucine	0.5	1.02	23.99	0.53	6.49	13.42	9.24			8.01	
Glutamate	0.5	1.02	24.09	0.84	7.56	13.7	11.19			8.25	
L-Valine	0.5	1.03	23.89	0.80	7.63	14.95	12.21			6.70	
Average									8.60		26.11

3. Results and Discussion

3.1. Computation of MSD

The presence of numerous statistic samples indicates high linearity of the MSD-t curve, resulting in accurate calculation of the diffusion coefficient. Abnormal diffusion, which usually occurs at beginning and ending part of the MSD, is nonlinearly related to time. These irregular data should not be used to calculate the diffusion coefficient. The fitting was conducted only in progresses into normal diffusion. Therefore, the determining time interval ($t_{min} \leq t \leq t_{max}$) is necessary for diffusion coefficient calculation. The diffusion of CO_2 was taken as an example to demonstrate the determination of t_{min} and $t_{max.}$ The MSD–time curve of CO_2 is presented in Fig. 1(a).

Double logarithmic coordinate plots of MSD are shown in Fig. 2(b). The MSD-t curve is straight line when the slope of lg (MSD)-lg(t) curve equal to 1. To observe the fluctuation of the slope of lg(MSD)-lg(t), the derivate of the lg(MSD)-lg(t) is depicted, as shown in Fig. 2(a). This work treats the discrete data by using a centered difference formula to approximate the derivative. The derivative function applied to discrete data points can therefore be written as follows:

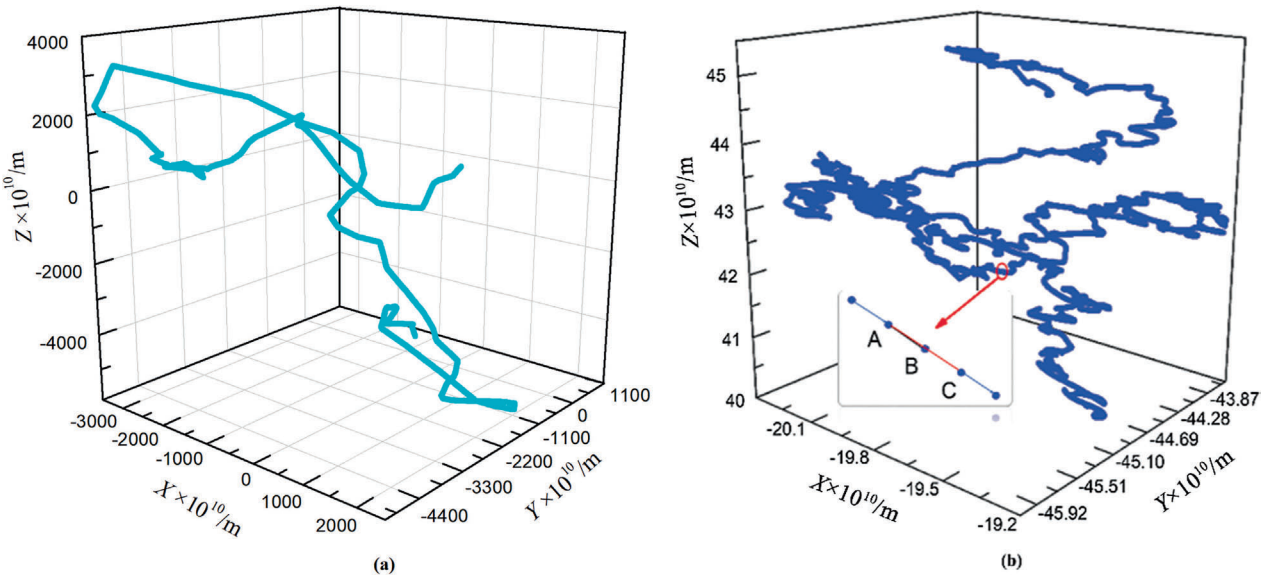

Fig. 1. (a) The diffusion pathway of N_2; (b) The diffusion pathway of Alanine (0.59 mol·L^{-1}).

$$f'(x_i) = \frac{1}{2}\left(\frac{y_{i+1}-y_i}{x_{i+1}-x_i} + \frac{y_i - y_{i-1}}{x_i - x_{i-1}}\right) \tag{4}$$

The MSD slope is considered as constant when the fluctuation of the slope of lg(MSD)-lg(t) is ranging from 0.8 to 1.2. Thus, t_{min} and t_{max} can be determined, as shown in Fig. 2(a). Consequently, the diffusion coefficients can be calculated from this time interval. The simulated results are listed in Tables 1–3 accordingly.

3.2. Results of the diffusion coefficient, characteristic length, and diffusion velocity

After a complete dynamic run, the diffusion velocity and path of each frame for every molecule obtained by scripts. Based on the available statistic samples, the statistical average diffusion velocity (namely the diffusion velocity) was given. The characteristic length got from the diffusion path. As illustrated in Fig. 1, it is obvious that the travel distance before changing direction is much shorter in liquid systems than in gas systems, which means that a moving molecule change direction more easily in the liquid than in the gas. The simulated results of V_i, and L_i for the diffusion of gases, organic vapor, and the liquid system are listed in Tables 1–3 accordingly.

After the characteristic length and the diffusion velocity gained. The diffusion coefficient was determined by our proposed model (Eq. (1)). For liquid system, the self-diffusion coefficients were corrected for finite-size effects by using the Yeh-Hummer correction (see Eqs. (5)–(7)) [30]. And for the gas system, we used the simulation box with large size (1000 molecules) to avoid the error caused by system size.

$$D_{cal-MED} = D'_{cal-MED} + D_{YH} \tag{5}$$

$$D_{cal-LV} = D'_{cal-LV} + D_{YH} \tag{6}$$

$$D_{YH} = \frac{k_B T \xi}{6\pi\eta L} \tag{7}$$

where D_{YH} is the YH correction diffusivity. $D'_{cal\text{-}MSD}$ and $D'_{cal\text{-}LV}$ are the calculated diffusion coefficient from the MSD-t curve and our new model respectively. $D_{cal\text{-}MSD}$ and $D_{cal\text{-}LV}$ are the corrected $D'_{cal\text{-}MSD}$ and $D'_{cal\text{-}LV}$. k_B is the Boltzmann constant, T is the temperature, η is the shear viscosity, L is the length of the cubic

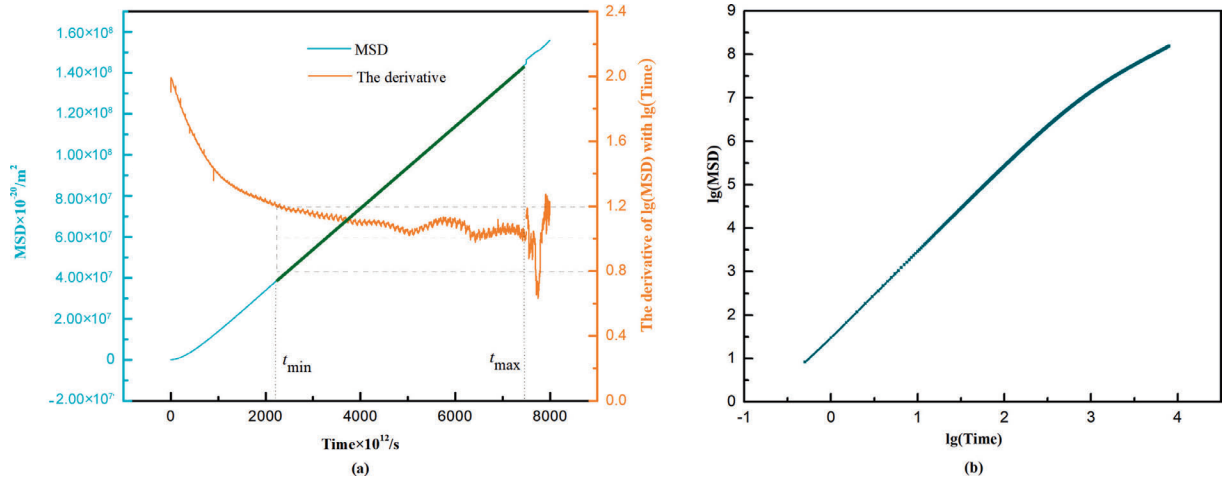

Fig. 2. (a) The MSD-t curve of CO_2; and the derivate of the lg(MSD), t_{min} and t_{max} are marked; (b) The lg(MSD)-lg(t) curve.

simulation box, ξ is a constant which depends on the shape of the simulation box (for a cubic simulation box, ξ = 2.837297).

The simulated results of $D_{\text{cal-LV}}$ and $D_{\text{cal-MSD}}$ for the diffusion of gases, organic vapor, and the liquid system are listed in Tables 1–3 accordingly. The calculation of relative deviations (RD, %) and the mean relative deviations (ARD, %) are based on the following equations:

$$\mathrm{RD}_{-1} = 100\% \cdot \frac{\left|D_{\exp(i)} - D_{\text{cal-LV}(i)}\right|}{D_{\exp(i)}}; \mathrm{ARD}_{-1} = \frac{1}{N_p}\sum_{i=1}^{N_p}\langle \mathrm{RD}_{-1}(i)\rangle \quad (8)$$

$$\mathrm{RD}_{-2} = 100\% \cdot \frac{\left|D_{\exp(i)} - D_{\text{cal-MSD}(i)}\right|}{D_{\exp(i)}}; \mathrm{ARD}_{-2} = \frac{1}{N_p}\sum_{i=1}^{N_p}\langle \mathrm{RD}_{-2}(i)\rangle \quad (9)$$

3.3. Gas system analysis

The simulated results for the diffusion of various gas and organic vapor systems are listed in Tables 1 and 2. From Tables 1 and 2, the calculated diffusion coefficients by our new model are in best agreement with those experimental results from literatures. The minimum value of RD_{-1} is 1.89% and the ARD_{-1} is 7.53%, indicating that the new model proposed previously are objective and rational. Besides, the diffusion coefficients were calculated after the high linearity of the MSD-*t* curve was determined, as listed in Tables 1 and 2. Compared with the traditional MSD-*t* model, this novel diffusion coefficient model provides more reliable results, and the theory is simple and straightforward in concept.

Besides, in this study, the effects of pressure were evaluated by varying pressures of CO_2, N_2 and O_2 in MD simulations. As presented in Table 2, the value of diffusion coefficients and the characteristic length decreased dramatically with increasing the gas pressure, while the diffusion velocity appeared less influenced by the pressure. This is quite understandable, since the intermolecular distance will reduce when the pressure is rising, the characteristic length will decrease. Meanwhile, the molecular velocity is little changed. As a result, the value of diffusion coefficient will reduce.

Moreover, the diffusion coefficients follow the order of CO_2 < O_2 < N_2 at identical pressures, which was probably caused by the increase of the density. This is consistent with the results in Table 1, the smaller density is, the greater the corresponding diffusion coefficient and characteristic length are. At the same temperature and pressure, collisions may not be easy to happen when the gas molecular density is smaller, thus the characteristic length is greater, and then the corresponding diffusion coefficient is larger.

3.4. Liquid system analysis

As shown in Table 3, the calculated diffusion coefficients are perfectly matched with the experimental results. The minimum value of RD_{-1} is 0.06% and the ARD_{-1} is 8.6%. This further evidence that our new model proposed previously accurately described the diffusion coefficient.

As shown in Table 3, we tested Alanine diffusion stimulation at different concentrations. It was found that the increase of the Alanine concentration in the aqueous solution caused decreases in both the diffusion coefficient and the characteristic length, but slight changes in the diffusion velocity. The trend of the characteristic length is caused by the reduction of distance between alanine molecules. However, the diffusion velocity was more sensitive to temperature changes.

The other notable rule is that the characteristic length and the diffusion velocity in liquid system are smaller than that in gas system. Those results about characteristic length verified above leads to a conclusion: when the intermolecular distance is reduced, the diffusion coefficients will be decreased. Generally, there are three different ways of molecule movement: translation, rotation, and vibration. In gas system, particles are far apart, with large intermolecular space. In this case, the intermolecular force of particles is relatively small. Translational motion makes a great contribution to this gas molecular motion. As a result, the direction of gas molecular displacement is not easy to change. Whereas in liquid system, the interaction force among particles is relatively large. In this case, molecular rotations and vibrations should be considered. The direction of molecular displacement is much easier to change in liquid system than in gas system, leading to smaller diffusion distance in liquid system.

3.5. Statistical analysis

According to a large number of statistical data from the trajectories, the distribution of diffusion velocity and the characteristic length can be analyzed through the probability density function.

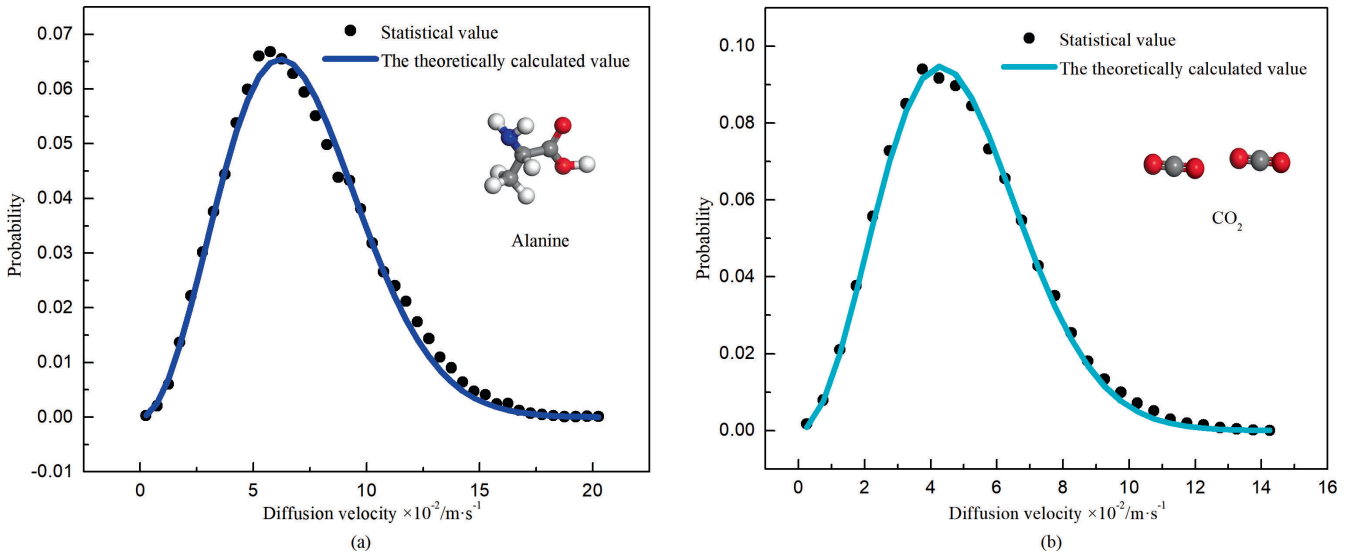

Fig. 3. (a) The diffusion velocity distribution of Alanine (0.59 mol·L^{-1}); (b) The diffusion velocity distribution of CO_2.

Table 4
The parameters of the distribution function for diffusion velocity

Group	p_1	p_2	p_3	$V_{cal}\times10^{-2}$/ m·s^{-1}	Pearson's r	Group	p_1	p_2	p_3	$V_{cal}\times10^{-2}$/ m·s^{-1}	Pearson's r
N_2	0.18	0.13	7.12	6.98	0.99	Alanine	0.13	0.05	6.59	7.12	0.99
O_2	0.16	0.13	5.85	6.10	0.99	Glycine	0.11	0.07	6	6.28	0.99
CO_2	0.07	0.07	6.27	6.52	1.00	Threonine	0.12	0.05	7.11	7.61	0.99
CH_4	0.06	0.05	7.79	8.21	0.99	Urea	0.14	0.06	6.21	6.5	0.99
C_2H_2	0.09	0.10	7.40	7.58	1.00	L-Arginine	0.09	0.05	7.01	7.4	0.99
H_2S	0.10	0.12	4.33	4.77	0.99	Leucine	0.15	0.07	6.16	6.49	0.99
C_3H_6	0.11	0.15	7.20	7.32	1.00	Glutamate	0.13	0.05	7.12	7.56	0.99
C_6H_6	0.12	0.17	6,98	7.09	1.00	L-Valine	0.12	0.05	7.18	7.63	0.99
Average					1.00						0.99

3.5.1. For diffusion velocity

By statistically analyzing of 8 liquid systems and 8 gas systems, we found that the probability density dot for the diffusion velocity firstly increased and then decreased. This pattern is similar to the normal distribution. Hence, we developed a new distribution function for describing the probability distribution of the diffusion velocity based on the normal distribution. The probability distribution function was shown in Eq. (10). As an example, the distribution of diffusion velocity of Alanine (0.59 mol·L^{-1}) and CO_2 are shown in Fig. 3, respectively. From Fig. 3, it illustrated that the probability density dot of diffusion velocity obeys our new distribution function very well.

$$f(V_i) = p_1 \cdot \exp\left(-p_2(V_i - p_3)^2\right) \tag{10}$$

where V_i is the velocity; p_1, p_2, p_3 are parameters.

After calculation, the values of the parameters and the Pearson correlation coefficient (Pearson's r) are listed in Table 4. The results show that the new distribution function in this work are satisfactory for the distribution of diffusion velocity. The minimum value of Pearson's r is 0.99 and the mean value is 0.99. It is worth noting that the parameter p_3 is approximate to the average diffusion velocity. This indicates that the physical meaning of p_3 is related to the average diffusion velocity.

The calculation of the average relative deviation (Pearson's r) was based on the following equation:

$$\text{Pearson}'\text{s } r = \frac{\text{cov}(y, \hat{y})}{\sigma_y \sigma_{\hat{y}}} \tag{11}$$

$\text{cov}(y, \hat{y}\cdot)$ is the covariance; σ_y is the standard deviation of X; $\sigma_{\hat{y}\cdot}$ is the standard deviation of Y.

3.5.2. For characteristic length

Based on the statistically analyzing of 8 liquid systems and 8 gas systems, we found that the tendency of the probability density dot for the characteristic length increase at first, and then decrease, and then go to be horizontal with the characteristic length increasing. Take the characteristic length distribution of Alanine (0.59 mol·L^{-1}) and CO_2 as an example, as shown in Fig. 4.

Molecular motion is a random event, and the molecular diffusion process can be considered into two random events: One is the small amplitude diffusion, and the other is the saltatory

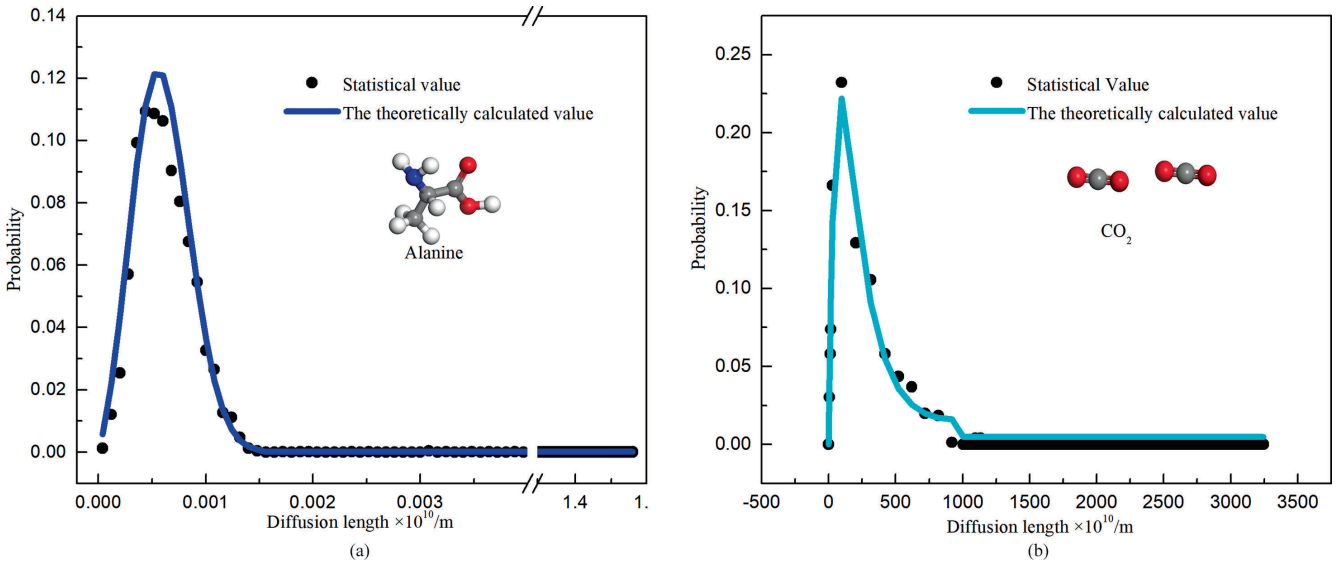

Fig. 4. (a) The characteristic length distribution of Alanine (0.59 mol·L^{-1}); (b) The characteristic length distribution of CO_2.

Table 5
The parameters of the distribution function for diffusion distance

Group	a_1	$a_2\times10^{6}$	$a_3\times10^{-3}$	Pearson's r	Group	$a_1\times10^{-3}$	$a_2\times10^{-8}$	$a_3\times10^{4}$	Pearson's r
N_2	0.50	2.01	−1.62	0.98	Alanine	0.26	0.05	3.8	0.99
O_2	1.02	3.59	−0.67	0.98	Glycine	1.62	0.82	1.25	1
CO_2	0.50	5.02	−0.95	0.99	Threonine	0.22	0.03	4.31	0.99
CH_4	0.50	2.90	−1.32	0.96	Urea	2.11	1.22	1.46	1
C_2H_2	0.50	2.92	−1.34	0.99	L-Arginine	0.19	0.04	5.4	1
H_2S	0.50	3.43	−1.21	0,99	Leucine	2.35	0.87	1.08	0.98
C_3H_6	0.20	5.26	−0.84	0.97	Glutamate	0.98	0.39	1.39	1
C_6H_6	0.02	26.70	−0.18	0.99	L-Valine	1.07	0.32	1.42	0.99
Average				0.98					0.99

diffusion. The former one is a high probability event. This pattern is similar to the normal distribution. And the later one is a small probability event, the regularity of distribution is similar to the uniform distribution. Therefore, we developed a new distribution function for describing the probability distribution of the characteristic length. This new distribution function contains two parts: the former part obeys a new distribution (see Eq.(12)). While the later part obeys a uniform distribution (see Eq.(13)). After calculation, the values of the parameters and the Pearson's r are presented in Table 5. It was showed that the new distribution function could describe the distribution of characteristic length well. The minimum value of Pearson's r is 0.96 and the mean value is 0.99.

$$f(L_i) = a_1 \cdot x \cdot \exp\left(-a_2 \cdot (L_i - a_3)^2\right) \quad 0 < L_i < a \tag{12}$$

$$f(L_i) = \frac{1}{b-a} \quad a < L_i < b \tag{13}$$

where L_i is the diffusion distance; $[a, b]$ is the value range of L_i; a_1, a_2, a_3 is parameters.

4. Conclusions

Based on the molecular dynamics simulations, we further investigated the underlying mechanism of the novel diffusion coefficient model at the molecular level. In this work, molecular dynamics simulations were performed for diffusion process in 35 systems (including 12 liquid systems as well as 23 gas and organic vapor systems), to investigate our novel diffusion coefficient model. With the dynamic simulation, the statistical average diffusion velocity and the characteristic length can be obtained from the trajectory file, and subsequently the diffusion coefficient defined by proposed model was determined. From the simulated results, diffusion coefficients from our new model matched perfectly with those experimental results from literatures. The ARD_{-1} of the simulations results was 7.53% and 8.60% in gas and liquid system, respectively, indicating that the novel model proposed previously are objective and rational. Compared with the traditional MSD-t model, this novel diffusion coefficient model provides more reliable results, and the theory is simple and straightforward in concept. Moreover, the effect of gas pressure and liquid concentration on the diffusion behavior were discussed, and the microscopic diffusion mechanism was elucidated through the distribution of diffusion velocity and the characteristic length analysis. Meanwhile, new distribution functions were suggested to describe the distribution of the diffusion velocity and the characteristic length, which provided theoretical foundations for the future research about the diffusion coefficient.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (21776264 and 21376231). My sincere appreciation goes to Prof. Hu and Prof. Wu who provided scientific guidance for me.

Nomenclature

a_1, a_2, a_3	represent the parameters of the distribution function
c	concentration of component I, $mol\cdot L^{-1}$
$\mathrm{cov}(y,\hat{y})$	the covariance
$D_{\text{cal-MSD}}$	calculated diffusion coefficient from the MSD-t curve, $m^2\cdot s^{-1}$
$D_{\text{cal-LV}}$	calculated diffusion coefficient from our new model, $m^2\cdot s^{-1}$
D_{exp}	experimental diffusion coefficient, $m^2\cdot s^{-1}$
K_B	Boltzmann constant
L_{cal}, L_i	characteristic length, m
L	length of the cubic simulation box
M	molar mass, $g\cdot mol^{-1}$
N_a	number of diffusion molecules
N_p	number of samples
p	pressure, kPa
p_1, p_2, p_3	represent the parameters of the distribution function
T	temperature, K
t	Time, s
$r_{i^{(t)}}$	position vector of the molecule at time t
$r_{i^{(0)}}$	initial position.
V_{cal}, V_i	diffusion velocity, $m\cdot s^{-1}$
η	shear viscosity
ξ	a constant which depends on the shape of the simulation box
σ	represent the standard deviation
σ_y	standard deviation of X
$\sigma_{\hat{y}}$	standard deviation of Y

References

[1] Y. Sun, K. Yang, Analysis of mass transport models based on Maxwell-Stefan theory and Fick's law for protein uptake to porous anion exchanger, *Sep. Purif. Technol.* 60 (2) (2008) 180–189.

[2] X. Liu, T.J.H. Vlugt, A. Bardow, Maxwell-Stefan diffusivities in liquid mixtures: Using molecular dynamics for testing model predictions, *Fluid Phase Equilib.* 301 (1) (2011) 110–117.

[3] K.S. Gandhi, Use of Fick's law and Maxwell-Stefan equations in computation of multicomponent diffusion, *AIChE J.* 58 (11) (2012) 3601–3605.

[4] K. Staszak, K. Prochaska, Estimation of diffusion coefficients based on adsorption measurements in model extraction systems, *Chem. Eng. Technol.* 28 (9) (2005) 985–990.

[5] Z.A. Makrodimitri, D.J. Unruh, I.G. Economou, Molecular simulation of diffusion of hydrogen, carbon monoxide, and water in heavy *n*-alkanes, *J. Phys. Chem. B* 115 (6) (2011) 1429–1439.

[6] O. Annunziata, D.G. Miller, J.G. Albright, Quaternary diffusion coefficients for the sucrose-NaCl-KCl-water system at 25 ℃, *J. Mol. Liq.* 156 (1) (2010) 33–37.

[7] X. Chen, Y. Wang, L.Y. Wu, W.T. Zhang, Y.D. Hu, A new model in correlating and calculating the diffusion-coefficient of electrolyte aqueous solutions, *Fluid Phase Equilib.* 485 (2019) 120–125.

[8] H.S. Harned, R.M. Hudson, The differential diffusion coefficient of potassium nitrate in dilute aqueous solutions at 25°, *J. Am. Chem. Soc.* 73 (2) (1951) 652–654.

[9] W.S. Price, H. Ide, Y. Arata, Self-diffusion of supercooled water to 238 K using PGSE NMR diffusion measurements, *J. Phys. Chem. A* 103 (4) (1999) 448–450.

[10] M.M. Rodrigo, A.J.M. Valente, M.C.F. Barros, L.M.P. Verissimo, M.L. Ramos, L.L.G. Justino, H.D. Burrows, A.C.F. Ribeiro, M.A. Esteso, Binary diffusion coefficients of l-histidine methyl ester dihydrochloride in aqueous solutions, *J. Chem. Thermodyn.* 89 (2015) 240–244.

[11] M. Zhou, Y.H. Zhang, X.G. Yuan, Diffusivity measurement of liquid solutions by real-time digital holography, *Chem. Ind. Eng.* 31 (6) (2014) 24–28.

[12] C. Mattisson, P. Roger, B. Jönsson, A. Axelsson, G. Zacchi, Diffusion of lysozyme in gels and liquids. A general approach for the determination of diffusion coefficients using holographic laser interferometry, *J. Chromatogr. B Biomed. Sci. Appl.* 743 (1–2) (2000) 151–167.

[13] M. Jamnongwong, K. Loubiere, N. Dietrich, G. Hébrard, Experimental study of oxygen diffusion coefficients in clean water containing salt, glucose or surfactant: Consequences on the liquid-side mass transfer coefficients, *Chem. Eng. J.* 165 (3) (2010) 758–768.

[14] J.J. Wang, Q.W. Yuan, M.Z. Dong, J.C. Cai, L. Yu, Experimental investigation of gas mass transport and diffusion coefficients in porous media with nanopores, *Int. J. Heat Mass Transf.* 115 (2017) 566–579.

[15] G. Pranami, M.H. Lamm, Estimating error in diffusion coefficients derived from molecular dynamics simulations, *J. Chem. Theory Comput.* 11 (10) (2015) 4586–4592.

[16] K.C. Ho, D. Hamelberg, Oscillatory diffusion and second-order cyclostationarity in alanine tripeptide from molecular dynamics simulation, *J. Chem. Theory Comput.* 12 (1) (2016) 372–382.

[17] O.A. Moultos, I.N. Tsimpanogiannis, A.Z. Panagiotopoulos, I.G. Economou, Self-diffusion coefficients of the binary (H_2O + CO_2) mixture at high temperatures and pressures, *J. Chem. Thermodyn.* 93 (2016) 424–429.

[18] C.J. Zhang, X.N. Yang, Molecular dynamics simulation of ethanol/water mixtures for structure and diffusion properties, *Fluid Phase Equilib.* 231 (1) (2005) 1–10.

[19] X. Zhao, H. Jin, Investigation of hydrogen diffusion in supercritical water: a molecular dynamics simulation study, *Int. J. Heat Mass Transf.* 133 (2019) 718–728.

[20] H.X. Hu, L. Du, Y.F. Xing, X.C. Li, Detailed study on self- and multicomponent diffusion of CO_2-CH_4 gas mixture in coal by molecular simulation, *Fuel* 187 (2017) 220–228.

[21] J.Z. Yang, Y. Chen, A.M. Zhu, Q.L. Liu, J.Y. Wu, Analyzing diffusion behaviors of methanol/water through MFI membranes by molecular simulation, *J. Membr. Sci.* 318 (1–2) (2008) 327–333.

[22] S.Q. Hu, A.L. Guo, Y.G. Yan, X.L. Jia, Y.F. Geng, W.Y. Guo, Computer simulation of diffusion of corrosive particle in corrosion inhibitor membrane, *Comput. Theor. Chem.* 964 (1–3) (2011) 176–181.

[23] J. Zhang, W.Z. Yu, L.J. Yu, Y.G. Yan, G.M. Qiao, S.Q. Hu, Y. Ti, Molecular dynamics simulation of corrosive particle diffusion in benzimidazole inhibitor films, *Corros. Sci.* 53 (4) (2011) 1331–1336.

[24] A. Ghaffari, A. Rahbar-Kelishami, MD simulation and evaluation of the self-diffusion coefficients in aqueous NaCl solutions at different temperatures and concentrations, *J. Mol. Liq.* 187 (2013) 238–245.

[25] Y.L. Zhao, Y.H. Feng, X.X. Zhang, Molecular simulation of CO_2/CH_4 self- and transport diffusion coefficients in coal, *Fuel* 165 (2016) 19–27.

[26] I.N. Tsimpanogiannis, O.A. Moultos, L.F.M. Franco, M.B.D.M. Spera, M. Erdős, I. G. Economou, Self-diffusion coefficient of bulk and confined water: A critical review of classical molecular simulation studies, *Mol. Simul.* 45 (4–5) (2019) 425–453.

[27] A.T. Celebi, S.H. Jamali, A. Bardow, T.J.H. Vlugt, O.A. Moultos, Finite-size effects of diffusion coefficients computed from molecular dynamics: A review of what we have learned so far, *Mol. Simul.* 2 (2020) 1–15.

[28] S.H. Jamali, A. Bardow, T.J.H. Vlugt, O.A. Moultos, Generalized form for finite-size corrections in mutual diffusion coefficients of multicomponent mixtures obtained from equilibrium molecular dynamics simulation, *J. Chem. Theory Comput.* 16 (6) (2020) 3799–3806.

[29] S.H. Jamali, L. Wolff, T.M. Becker, A. Bardow, T.J.H. Vlugt, O.A. Moultos, Finite-size effects of binary mutual diffusion coefficients from molecular dynamics, *J. Chem. Theory Comput.* 14 (5) (2018) 2667–2677.

[30] I.C. Yeh, G. Hummer, System-size dependence of diffusion coefficients and viscosities from molecular dynamics simulations with periodic boundary conditions, *J. Phys. Chem. B* 108 (40) (2004) 15873–15879.

[31] O.A. Moultos, Y. Zhang, I.N. Tsimpanogiannis, I.G. Economou, E.J. Maginn, System-size corrections for self-diffusion coefficients calculated from molecular dynamics simulations: The case of CO_2, *n*-alkanes, and poly (ethylene glycol) dimethyl ethers, *J. Chem. Phys.* 145 (7) (2016) 074109.

[32] A.T. Celebi, T.J.H. Vlugt, O.A. Moultos, Structural, thermodynamic, and transport properties of aqueous reline and ethaline solutions from molecular dynamics simulations, *J. Phys. Chem. B* 123 (51) (2019) 11014–11025.

[33] A. Moridnejad, T.C. Preston, Models of isotopic water diffusion in spherical aerosol particles, *J. Phys. Chem. A* 120 (49) (2016) 9759–9766.

[34] C.P. Calderon, Estimation and inference of diffusion coefficients in complex biomolecular environments, *J. Chem. Theory Comput.* 7 (2) (2011) 280–290.

[35] A. Kamgar, N. Hamedi, S. Mohsenpour, M.R. Rahimpour, Investigation of using different thermodynamic models on prediction ability of mutual diffusion coefficient model, *J. Mol. Liq.* 243 (2017) 781–789.

[36] E.L. Cussler, Cluster diffusion in liquids, *AIChE J.* 26 (1) (1980) 43–51.

[37] E.B. Winn, The temperature dependence of the self-diffusion coefficients of argon, neon, nitrogen, oxygen, carbon dioxide, and methane, *Phys. Rev.* 80 (6) (1950) 1024.

[38] X. Yun, D. Kou, Y. Li, H. Li, A study of flow perturbation gas chromatography the measurements of the diffusion coefficients, *Acta Scientiarum Naturalium Universitatis Nankaiensis* 28 (1995) 12–17.

[39] Y.G. Ma, C.Y. Zhu, P.S. Ma, K.T. Yu, Studies on the diffusion coefficients of amino acids in aqueous solutions, *J. Chem. Eng. Data* 50 (4) (2005) 1192–1196.