Korean J. Chem. Eng., 39(6), 1651-1657 (2022)
DOI: 10.1007/s11814-022-1096-1

pISSN: 0256-1115
eISSN: 1975-7220

In silico prediction and analysis of dielectric constant of ionic liquids

Chul-Woong Cho*,,† and Yeoung-Sang Yun***,†**

*Department of Integrative Food, Bioscience and Biotechnology, Chonnam National University, Yongbong-ro 77, Buk-gu, Gwangju 61186, Korea
**Department of Bioenergy Science and Technology, Chonnam National University, Gwangju 61186, Korea
***School of Chemical Engineering, Jeonbuk National University, 567 Beakje-dearo, Deokjin-gu, Jeonju, Jeonbuk 54896, Korea
(*Received 20 October 2021 • Revised 24 February 2022 • Accepted 24 February 2022*)

Abstract−Ionic liquids (ILs) are a class of chemicals comprising cations and anions whose properties can be controlled by modifying their chemical structure, which enables a wide range of applications. Among the attractive properties of ILs, dielectric permittivity provides important information related to material solvation and capacitor characteristics. Because there are several ILs and a need to understand the structural effect on their properties, prediction model(s) should be developed. For this, we employed the linear free-energy relationship (LFER) equation to predict the dielectric constant of ILs. In the modeling, we used *in silico* calculated molecular descriptors because the empirically LFER estimated descriptors were limited. The results revealed that the developed model could predict the dielectric constant with an R^2 of 0.882. From the developed model, it was observed that the dielectric constant was more affected by the structure of cations compared to that of anions. In addition, the H-bonding acidity of the cation and basicity of the anion contributed to the dielectric property of ILs, and the dipolarity/polarizability of cations and anions was also important in the prediction. The predictive model is expected to be useful for designing IL structures considering the dielectric constant.

Keywords: Dielectric Permittivity, Quantitative Structure-activity Relationship, Linear Free Energy Relationship, *In Silico* Calculated Molecular Descriptors, Hydrogen-bonding Effect

INTRODUCTION

Ionic liquids (ILs) are a class of chemical species composed of cation(s) and anion(s), which usually have strong ionic bond(s). As they have negligible volatility and flammability because of their ionic bond(s), they can be used as an alternative to conventional volatile organic solvents. Therefore, ILs are considered as greener solvents. The application of ILs in industries includes separation, lubrication, pretreatment of biomass, electric devices, and battery production [1,2]. For the appropriate application and design of ILs, their fundamental properties should be investigated.

Among the properties of ILs, the dielectric constant or relative permittivity (ε), which is the ratio of permittivity of a material to that of the free space, is very important. It is closely related to the polarity of the molecules [3], which is highly correlated with its ability to dissolve substances. A solvent with a high ε has a good ability to dissolve salts [4]. In addition, this property provides information to the design of capacitors and circuits. Molecules with low ε can reduce electric power loss in high-frequency and high-power applications, whereas those with high ε are recommended to be used as small-sized capacitances [5]. Thus, ε is fundamental information that should be provided. However, the number of ILs was theoretically estimated to be 10^{18} [6]; thus, considerable experimental performance with labor and material consumption is required, and a more efficient estimation method should be developed.

Theoretical modeling can be applied as a solution because it can help predict or understand properties. There are several types of modeling methods, such as machine learning, deep learning methods, principle component analysis, and quantitative structure-activity relationship (QSAR). Among them, machine learning and deep learning methods are currently highly used in big data treatment, and they can achieve high predictability. However, they are not effective to understand the chemical meanings related to how the molecular structure affects the molecular properties. As another option, the group contribution modeling method can be applied to predict properties by considering the chemical structure as a numerical value. Although this method can lead to high prediction and understanding of the chemical meaning, it does not sufficiently reflect the electron movement in a molecule by inductive effect. As a complementary method, the QSAR concept can be applied to the modeling of a prediction model [7] and for understanding the chemical aspects, and has been applied in several case studies [8-11].

When applying QSAR modeling, parameter selection is the most important, and the selected parameters should be easy. In addition, *in silico* methods are beneficial for the efficient application of these parameters. From this viewpoint, linear free energy relationship (LFER) modeling by Abraham et al. [12-14] is considered as a suitable method because the model is comprised of well-defined structural information as descriptors, and can be used to develop a highly accurate prediction model [10,11,15,16]. However, it is limited to approximately 8000 chemicals available [17]. To solve this problem, our group previously developed a computational method

†To whom correspondence should be addressed.
E-mail: choicejoe@jnu.ac.kr, ysyun@jbnu.ac.kr

to calculate the descriptors [18], which enables the exploration of a wide range of chemicals, even for theoretically existing substances.

Prediction studies on the dielectric constants of ILs have been performed. Zhou et al. [19] used a group contribution method to predict the dielectric constant of ILs. The model had good predictability with average absolute relative deviation of 7.41-37.47%, but the structural effect on the property was not sufficiently discussed. Rybinska-Fryca [20] developed models using *in silico* calculated two-dimensional (2D) or three-dimensional (3D) related parameters. The developed models achieved high predictability, with R^2 values of 0.87 and 0.90 when applying 2D and 3D descriptors, respectively. Eiden et al. [21] also developed a prediction model using *in silico* parameters calculated by RI-BP86/def-TZVP/COSMO [21]. The model is comprised of simple parameters, that is, molecular volume and Gibbs energy, and can predict the temperature-dependent ε value. However, more diverse prediction models are needed for cross-validation and to improve understanding to the effect of molecular structures on the property.

Therefore, the purpose of this study was to predict the dielectric constants of ILs using the LFER concept with more collected datasets and *in silico* calculated descriptors and to understand which molecular factors contribute to dielectric constants. First, we obtained sub-parameters to calculate LFER descriptors [18], which were used to correlate with the log dielectric constant of ILs. Finally, based on the theoretically developed model, the most important molecular factors for the properties were identified and discussed.

MATERIALS AND METHODS

1. Nomenclature of ILs

Cations: (2-hydroxyethyl)ammonium [Nhhh2OH]; 1-(2-hydroxyethyl)-3-methylimidazolium [IM12OH]; 1,3-dimethylimidazolium [IM11]; 1-ethyl-2-hydroxyethylpyridinum [Py2-2OH]; 1-butyl-1-methylpyrrolidinium [Pyr14]; 1-butyl-2,3,4,5-tetramethylimidazolium [IM14-2Me-4Me-5Me]; 1-butyl-2,3-dimethylimidazolium [IM11-2Me]; 1-butyl-3-methylimidazolium [IM14]; 1-butylpyridinium [Py4]; 1-butyl-3-methylpyridinium [Py4-3Me]; 1-butyl-4-methylpyridinium [Py4-4Me]; 1-ethyl-1-methylpyrrolidinium [Pyr12]; 1-pentyl-1-methylpyrrolidinium [Pyr15]; 1-methyl-1-propylpyrrolidinium [Pyr13]; 1-ethyl-2,3-dimethylimidazolium [IM12-2Me]; 1-ethyl-3-methylimidazolium [IM12]; 1-hexyl-3-methylimidazolium [IM16]; 1-methyl-3-octylimidazolium [IM18]; 1-methyl-3-propylimidazolium [IM13]; 1-methylpyrrolidinium [Pyr1h]; 1-pentyl-3-methylimidazolium [IM15]; butylammonium [Nhhh5]; ethylammonium [Nhhh2]; N-butyl-N,N,N-trimethylammonium [N2224]; triethylsulfonium [S222]; tris(2-hydroxyethyl)ammonium [Nh(2OH)(2OH)(2OH)].

Anions: bis(trifluoromethylsulfonyl)amide [$(CF_3SO_2)_2N$]; dimethyl phosphate [$(CH_3)_2PO_2$]; diethyl phosphate [$(CH_3CH_2)_2PO_2$]; acetate [1COO]; lactate [Lac]; ethyl sulfate [$2OSO_3$]; methyl sulfate [$1OSO_3$]; butyl sulfate [$4OSO_3$]; tetrafluoroborate [BF_4]; trifluoromethanesulfonate [CF_3SO_3]; formate [HCOO]; hydrogen sulfate [HSO_4]; nitrate [NO_3]; hexafluorophosphate [PF_6]; thiocyanate [SCN]; chloride Cl; bromide Br; dicyanamide [$N(CN)_2$]

2. Dataset

The dielectric constant (ε) values of ILs were obtained from several studies [3,22-30]. Some ILs have several ε values. In these cases, the average value was used as the dependent variable. The dataset had a total of 60 points. In addition, ε of ILs is temperature-dependent; therefore, in this study, only the values measured at 298.15 K were collected and used.

3. Amendment of Original LFER Model

The version of the original LFER model, which can cover ionic and nonionic chemicals, is as follows:

$$\log \varepsilon = e\cdot E + s\cdot S + a\cdot A + b\cdot B + v\cdot V + j^-\cdot J^- + j^+\cdot J^+ + c \quad (1)$$

where log ε is the dependent variable and is the dielectric constant in log units. Capital letters represent the inherent molecular properties of a substance [31,32], as follows: E is the excess molar refraction ($cm^3\ mol^{-1}/100$); S is the dipolarity/polarizability; A and B are the H bonding acidity and basicity, respectively; V is the molecular volume ($cm^3\ mol^{-1}/100$); J^+ and J^- are Coulombic interactions of cations and anions, respectively; and small letters and italics (e, s, a, b, v, j^-, and j^+) represent system-dependent parameters that can be determined by multiple linear regressions.

Because ILs are comprised of two parts, that is, cations and anions, Eq. (1) may not be used to predict salt-type chemicals. To address this limitation, Eq. (1) can be split into cationic and anionic parts, as follows:

$$\begin{aligned}\log \varepsilon\,[\text{Farad/meter}] = {} & e_c\cdot E_c^o + s_c\cdot S_c^o + a_c\cdot A_c^o + b_c\cdot B_c^o + v_c\cdot V_c + j^+\cdot J^{+,o} \\ & + e_a\cdot E_a^o + s_a\cdot S_a^o + a_a\cdot A_a^o + b_a\cdot B_a^o + v_a\cdot V_a^o + j^-\cdot J^{-,o} + c,\end{aligned} \quad (2)$$

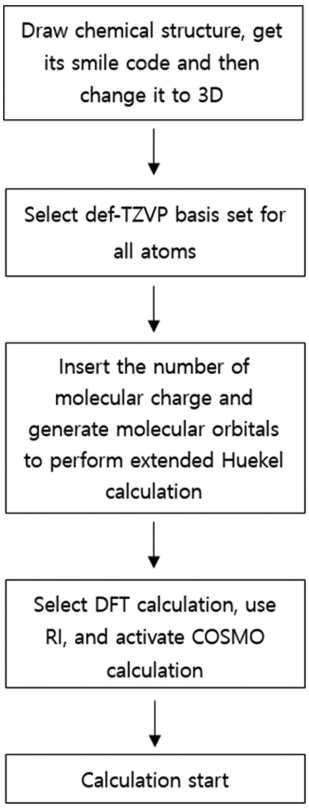
where, the subscripts 'c' and 'a' mean cation and anion, respectively, and the superscript 'o' means that the descriptor was the calculated value.

4. Computational Details

The descriptors were calculated based on density functional theory (DFT) [33] and conductor-like screening model (COSMO) in TmoleX software (COSMOlogic, 1989-2020). Briefly, we first draw the molecular structure as an input and transformed it into 3D information. Then, the structures were optimized using DFT [33] and refined using the def-TZVP basis set [34] to calculate the structures with the lowest energy levels. The calculation flowchart is given in Scheme 1 and the actual figure, describing the calculation steps, is given in supplementary information (Scheme S1). After the calculations, TmoleX generated a cosmo. file for the structure, and the file was sent to be read in the COSMOtherm software [35]. COSMOtherm provided sub-parameters describing surface-charge related terms, that is, sigma moments, H-bonding moments, and dielectric energy values. The meanings of the sub-parameters are given in Table 1. In addition, Open Babel [36] was applied to obtain the refractivity and polar surface area value of a molecule or atom. Moreover, the number of H atoms connected to O or N atoms and the number of rings in the molecule were counted. Previously, using the calculated and counted sub-parameters, Cho et al. (18) presented several equations for each LFER descriptor, that is, Eqs. (3)-(8).

$$V^O\ [(cm^3mol^{-1})/100] = 0.639\cdot V_{cosmo}\ [nm^3] - 0.0046 \quad (3)$$

$$\begin{aligned}E^O\ [cm^3mol^{-1}/10] = {} & 0.341\cdot N_{Ring} + 0.007\cdot PSA \\ & + 0.057\cdot MR - 1.762\cdot V_{cosmo}/100 - 0.113\cdot E_{vdw} \\ & + 0.275\cdot MW + 0.135\cdot \sigma_1 + 0.015\cdot \sigma_4 - 0.037\end{aligned} \quad (4)$$



Scheme 1. Flowchart of the computational procedure.

Table 1. Sub-parameters used and corresponding meanings

Calculated sub-parameter	Meaning
PSA	Polar surface area
V_{cosmo}	Molecular volume calculated by COSMO
E_{vdw}	van der Waals force
MR	Molar refraction
σ_x	Sigma moments meaning surface area
HBD_x	H-bond donor moments
HBA_x	H-bond acceptor moments
N_{Ring}	Number of rings
N_{OH}	Number of H attached to O
N_{HN}	Number of H attached to N

$$A^O\ [\text{dimensionless}]=171\cdot HBD_1^2-0.047\cdot HBD_3^2+0.032\cdot HBD_4^2+73.511\cdot HBD_1+0.654\cdot HBD_2+0.208\cdot HBD_3+0.203\cdot N_{OH}+0.080\cdot N_{HN}-0.019\cdot E_{vdw}-0.132 \quad (5)$$

$$B^O\ [\text{dimensionless}]=0.391\cdot\sigma_2+1.00\cdot\sigma_3+0.421\cdot\sigma_4-0.117\cdot\sigma_5-0.055\cdot\sigma_6+0.112\cdot\sigma_1^2-0.149\cdot\sigma_2^2-0.070\cdot HBA_2/V_{cosmo}+0.074\cdot HBA_3/V_{cosmo}+0.032 \quad (6)$$

$$J^{+,O}\ [\text{dimensionless}]=-0.124-0.106\cdot E_{vdw}+0.421\cdot\sigma_3+0.292\cdot\sigma_4+64.928\cdot HBD_1+0.661\cdot HBD_2-0.049\cdot HBA_2/V_C-0.092\cdot\sigma_6^2/100 \quad (7)$$

$$J^{-,O}\ [\text{dimensionless}]=1.331+4.712\cdot\sigma_2-2.770\cdot\sigma_3-0.832\cdot\sigma_2^2+0.300\cdot\sigma_3^2-0.012\cdot\sigma_4^2-0.155\cdot HBA_2+0.238\cdot HBA_3-0.292\cdot N_{OH}+0.183\cdot N_{Ring} \quad (8)$$

$$S^O\ [\text{dimensionless}]=-1.441\cdot\sigma_1+0.206\cdot\sigma_2^2-0.009\cdot\sigma_4^2-0.122\cdot HBA_4+0.511\cdot\text{Calc. E}+1.524\cdot\text{Calc. B}+0.856\cdot\text{Calc. }J^+ +3.308\cdot(\sigma_1\cdot\text{Calc. }J^-)/\text{Calc. B}-0.099 \quad (9)$$

5. Statistical Analysis

A statistical analysis to determine the system coefficients in Eq. (3) was performed using SPSS software (12.0 K Windows version).

RESULTS AND DISCUSSION

The number of data obtained was 60 and ranged from 0.85 to 1.97 log units, which is rather narrow. The IL cations studied were imidazolium, pyridinium, ammonium, and pyrrolidinium, with different alkyl chain lengths, substituents, and/or functional groups. The tested anions were of 18 types. The collected data were not normally distributed, as if the estimated p-value in the Kolmogorov-Smirnov and Shapiro-Wilk tests were zero. In fact, the data were biased within 1.0-1.2 log units of dielectric constant. Nevertheless, a part of the data was evenly distributed in the remainder of the range, that is, 1.2-2.1.

In modeling, the obtained dataset was not divided into a training set and a test set because of its small size, which can cause less comprehensive structural viewpoints and biased predictability. Therefore, in this study, all data values were used as the modeling set. To develop a prediction model for the dielectric constant based on Eq. (2), the LFER descriptors calculated using Eqs. (3)-(9) were inserted into the datasheet in SPSS, and multiple linear regression was performed. Note that we considered structures of cations and anions separately, which is useful for understanding the effect of cations or anions on the properties and for studying a wide range of ILs. Indeed, using the descriptors of 26 cations and 18 anions, about 460 combinations of ILs can be studied. The studied molecular structures are given in Table S1 in supplementary information. Then, descriptors that contributed more to the dielectric constant of ILs were selected. When selecting a descriptor, the estimated p-value was considered as a criterion. If it was higher than 0.05, meaning that the descriptor was out of 95% of the confidential level, it was excluded. If it was lower, it was considered at an acceptable confidence level. In modeling, several descriptors with p-values higher than 0.05 (i.e., E_c^o, B_c^o, $J^{+,o}$, E_a^o, A_a^o, and V_a^o) were excluded. After exclusion, the following equation was developed:

$$\text{Log }\varepsilon=-0.093(0.019)\cdot S_c^o+0.292(0.025)\cdot A_c^o+0.065(0.012)\cdot S_a^o+0.127(0.023)\cdot B_a^o-0.229(0.041)\cdot J^{-,o}+1.10(0.082) \quad (10)$$

R^2=0.882, R_{adv}^2=0.871, SE=0.084 log units, F=79.1 N=59

where, R^2 of Eq. (9) was 0.882 and the standard error was 0.084 log units, which indicates high predictability. The adjusted R^2 (R_{adv}^2) value revealed that the descriptor selection was reasonable, as the value was close to the R^2 value (0.871 vs. 0.882). The fitting is shown in Fig. 1(a), and the calculated log ε is given in Table 2. The estimated variance inflation factor values of the descriptors used ranged from 1.35 to 3.27, which is lower than 10 according to the criterion, indicating multicollinearity among the descriptors, but it was a slight multicollinearity. By applying Eq. (3) for ε prediction, there was one outlier, [IM12] [$(CH_3CH_2)_2PO_2$]. Eq. (10) slightly overestimated the ε value, providing a result of 1.53 log units, whereas

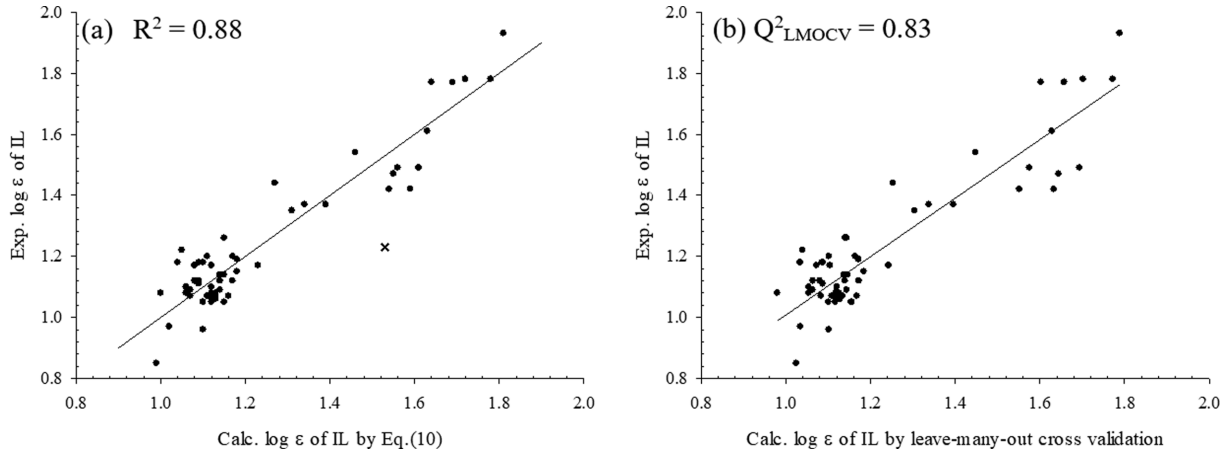


Fig. 1. Correlation between experimental and calculated log dielectric constant (ε) (a) by Eq. (10) and (b) by leave-many-out cross validation. Here, x means an outlier i.e., [IM12] [$(CH_3CH_2)_2PO_2$].

Table 2. Descriptors of ILs calculated by Eqs. (3)-(9), and measured (at 298.15 K) and predicted ε values

Cation	Anion	S^o_c	A^o_c	S^o_a	B^o_a	$J^{-,o}$	Exp. ε	Average of ε	log ε	Calc. log ε by Eq. (10)
[Nhhh2OH]	[Lac]	1.05	2.14	2.53	2.95	1.53	85.6 [22]	85.6	1.93	1.81
[Nhhh2OH]	[1COO]	1.05	2.14	1.51	2.82	1.91	58.3 [22]	58.3	1.77	1.64
[Nhhh2OH]	[HCOO]	1.05	2.14	1.53	2.58	1.57	61 [22], 57.3 [3]	59.2	1.77	1.69
[Nhhh2OH]	[NO_3]	1.05	2.14	2.95	1.63	1.31	60.9 [22]	60.9	1.78	1.72
[IM12OH]	[BF_4]	3.22	1.49	4.06	1.06	1.26	23.3 [22]	23.3	1.37	1.34
[Pyr14]	[$(CF_3SO_2)_2N$]	2.61	0.60	5.24	1.38	1.86	7.0 [30], 14.7 [22], 11.9 [24], 11.7 [23]	11.3	1.05	1.12
[Pyr14]	[$N(CN)_2$]	2.61	0.60	3.47	1.84	1.50	18 [22]	18.0	1.26	1.15
[IM11]	[1OSO_3]	2.18	0.62	3.42	2.20	1.83	17.2 [30]	17.2	1.07	1.16
[IM11-2Me]	[$(CF_3SO_2)_2N$]	2.66	0.44	5.24	1.38	1.86	11.6, 14 [22], 11.45 [25]	12.4	1.09	1.07
[IM11-2Me]	[PF_6]	2.66	0.44	7.43	1.24	2.60	9.4 [22]	9.4	0.97	1.02
[IM11-2Me]	[BF_4]	2.66	0.44	4.06	1.06	1.26	13.3 [22]	13.3	1.12	1.09
[IM14-2Me-4Me-5Me]	[$(CF_3SO_2)_2N$]	2.43	0.26	5.24	1.38	1.86	15 [22]	15.0	1.18	1.04
[IM14-2Me-4Me-5Me]	[BF_4]	2.43	0.26	4.06	1.06	1.26	12 [22]	12.0	1.08	1.06
[IM14]	Cl	2.53	0.59	2.20	2.31	2.05	15 [30]	15.0	1.08	1.00
[IM14]	[$(CF_3SO_2)_2N$]	2.53	0.59	5.24	1.38	1.86	11.6 [24], 14 [22], 11.52 [25], 13.7 [26]	12.7	1.10	1.12
[IM14]	[1OSO_3]	2.53	0.59	3.42	2.20	1.83	14.8 [30]	14.8	1.17	1.12
[IM14]	[$N(CN)_2$]	2.53	0.59	3.47	1.84	1.50	11.3 [22]	11.3	1.05	1.15
[IM14]	[PF_6]	2.53	0.59	7.43	1.24	2.60	14 [30], 14 [22], 11.4 [27], 16.1 [26], 14.1 [26], 11.4 [24]	13.2	1.12	1.08
[IM14]	[BF_4]	2.53	0.59	4.06	1.06	1.26	11.7 [24], 12.9 [30], 13.9 [22], 11.7 [27], 14.6 [26], 14.35 [28], 14.1 [26]	13.3	1.12	1.14
[IM14]	[SCN]	2.53	0.59	3.66	1.52	1.42	13.7 [22]	13.7	1.14	1.14
[IM14]	[CF_3SO_3]	2.53	0.59	3.47	1.70	1.33	13.5 [30], 12.9 [22], 13.2 [24]	13.2	1.12	1.17
[Py4-3Me]	[BF_4]	2.68	0.49	4.06	1.06	1.26	15 [30]	15.0	1.18	1.10
[Py4-4Me]	[BF_4]	2.65	0.45	4.06	1.06	1.26	15.3 [30]	15.3	1.18	1.09

Table 2. Continued

Cation	Anion	S^o_c	A^o_c	S^o_a	B^o_a	$J^{,o}$	Exp. ε	Average of ε	log ε	Calc. log ε by Eq. (10)
[Py4]	$[(CF_3SO_2)_2N]$	2.70	0.66	5.24	1.38	1.86	9.4 [30], 15.3 [22], 11.5 [24], 11.3 [23]	11.9	1.07	1.13
[Py2-2OH]	$[2OSO_3]$	2.35	1.19	3.73	2.39	1.38	35 [22]	35.0	1.54	1.46
[Pyr12]	$[N(CN)_2]$	2.07	0.55	3.47	1.84	1.50	14 [26]	14.0	1.15	1.18
[IM12-2Me]	$[(CF_3SO_2)_2N]$	2.37	0.42	5.24	1.38	1.86	12.8 [22]	12.8	1.11	1.09
[IM12-2Me]	$[PF_6]$	2.35	0.55	7.43	1.24	2.6	14.7 [30]	14.7	1.17	1.08
[IM12]	$[(CF_3SO_2)_2N]$	2.35	0.55	5.24	1.38	1.86	11.5 [30], 12 [22], 12.25 [25], 12.3 [24], 12.3 [26]	12.1	1.08	1.13
[IM12]	$[4OSO_3]$	2.35	0.55	4.98	2.72	1.40	17.5 [24], 30 [22]	23.8	1.37	1.39
[IM12]	$[NO_3]$	2.35	0.55	3.47	1.84	1.50	11 [26], 11.1 [29]	11.1	1.05	1.15
[IM12]	$[(CH_3CH_2)_2PO_2]$	2.35	0.55	6.77	3.90	1.93	16.9 [22]	16.9	1.23	1.53
[IM12]	$[2OSO_3]$	2.35	0.55	3.73	2.39	1.38	13.5 [30], 29.7 [24], 35 [22], 35.2 [37]	27.9	1.44	1.27
[IM12]	$[HSO_4]$	2.35	0.55	4.41	1.91	1.82	18.4 [22]	18.4	1.26	1.15
[IM12]	[1COO]	2.35	0.55	2.53	2.95	1.53	14.9 [37]	14.9	1.17	1.23
[IM12]	$[BF_4]$	2.35	0.55	4.06	1.06	1.26	14.8 [30], 12.9 [24], 12.8 [27], 14.5 [26], 13.6 [26], 14.7 [29]	13.9	1.14	1.15
[IM12]	$[CF_3SO_3]$	2.35	0.55	3.47	1.70	1.33	15.8 [30], 16.5 [22], 15.2 [27], 15.1 [24]	15.7	1.19	1.18
[Nh(2OH)(2OH)(2OH)]	[Lac]	3.34	2.77	2.53	2.95	1.53	59.7 [22]	59.7	1.78	1.78
[IM16]	Br	2.69	0.62	2.90	1.93	1.78	16.6 [22]	16.6	1.22	1.05
[IM16]	$[(CF_3SO_2)_2N]$	2.69	0.62	5.24	1.38	1.86	7.0 [30], 12.7 [26], 16 [26]	11.9	1.08	1.12
[IM16]	$[PF_6]$	2.69	0.62	7.43	1.24	2.60	11.1 [30], 15.5 [26], 8.9 [27]	11.8	1.07	1.07
[IM16]	$[BF_4]$	2.69	0.62	4.06	1.06	1.26	11.3 [30], 12 [26], 15.9 [26], 10.1 [38]	12.3	1.09	1.14
[IM18]	$[(CF_3SO_2)_2N]$	2.84	0.64	5.24	1.38	1.86	6.5 [30], 16.8 [26]	11.7	1.07	1.11
[IM18]	$[PF_6]$	2.84	0.64	7.43	1.24	2.60	9.7 [30], 15.2 [26]	12.5	1.10	1.06
[IM18]	$[BF_4]$	2.84	0.64	4.06	1.06	1.26	7.5 [30], 15.5 [26]	11.5	1.06	1.13
[IM18]	Cl	2.84	0.64	2.20	2.31	2.05	7.0 [30]	7.0	0.85	0.99
[IM13]	$[(CF_3SO_2)_2N]$	2.45	0.58	5.24	1.38	1.86	10.6 [30], 13.3 [22], 11.8 [24], 11.8 [25]	11.9	1.07	1.13
[Pyr1h]	[HCOO]	1.90	1.10	1.53	2.58	1.57	23.1 [22], 22 [3]	22.6	1.35	1.31
[Pyr13]	$[(CF_3SO_2)_2N]$	2.72	0.58	5.24	1.38	1.86	9.1 [30]	09.1	0.96	1.10
[Pyr15]	$[(CF_3SO_2)_2N]$	2.70	0.58	5.24	1.38	1.86	11.1 [24], 12.5 [22]	11.8	1.07	1.11
[IM15]	$[(CF_3SO_2)_2N]$	2.61	0.60	5.24	1.38	1.86	7.9 [30], 11.4 [24], 15 [22], 11.45 [25]	11.4	1.06	1.12
[IM11]	$[(CH_3)_2PO_2]$	2.18	0.62	6.51	3.73	1.85	29.6 [22]	29.6	1.47	1.55
[Nhhh5]	[HCOO]	0.69	1.49	1.53	2.58	1.57	23 [22], 29.2 [3]	26.1	1.42	1.54
[Nhhh2]	[HCOO]	0.45	1.49	1.53	2.58	1.57	31.5 [22], 30.3 [3]	30.9	1.49	1.56
[Nhhh2]	$[NO_3]$	0.45	1.49	2.95	1.63	1.31	26.2 [24], 26.3 [22], 26.2 [3]	26.3	1.42	1.59
[Nhhh1]	[HCOO]	0.31	1.68	1.53	2.58	1.57	41 [22], 40.3 [3]	40.7	1.61	1.63
[N2224]	$[(CF_3SO_2)_2N]$	2.51	0.73	5.24	1.38	1.86	15.7 [22]	15.7	1.20	1.17
[N2225]	$[(CF_3SO_2)_2N]$	2.27	0.43	5.24	1.38	1.86	10.2 [24], 12.5 [22]	11.4	1.05	1.10
[S222]	$[(CF_3SO_2)_2N]$	2.5	0.52	5.24	1.38	1.86	15.8 [22]	15.8	1.20	1.11
[Nh(2OH)(2OH)(2OH)]	[1COO]	3.34	2.77	1.51	2.82	1.91	31 [22]	31.0	1.49	1.61

the experimental value was 1.23 log units. Also, the ε values between 1.0 and 1.2, forming a cluster in the correlation (Fig. 1), were not clearly predictable. This might be due to limitations of the selected parameters. Thus, for more accurate prediction, further studies are needed.

To validate Eq. (10), a leave-many-out cross-validation study was performed. In the validation study, we randomly mixed the dataset, excluded three to four data points, and then developed each model using the remaining data points and the same descriptors selected for Eq. (10). Subsequently, each developed model was applied to predict the log ε value of the excluded ILs, and the calculated values were compared with experimental values. The results revealed that the R^2 value of the leave-many-out cross validation (Q^2_{LMOCV}) was 0.84 (Fig. 1(b)), which meets the criterion of Q^2_{LMOCV} higher than 0.5.

Based on the developed model, the contributions of the molecular properties to the dielectric constant were analyzed. First, to verify the more significant moiety between the cation and anion, the selected terms for each part were used to predict the value. When applying a combination of the cationic descriptors selected in Eq. (10), that is, S^o_c and A^o_c, the two terms had a higher correlation than the combination of anionic descriptors selected in Eq. (10), that is, S^o_a, B^o_a, and $I^{-,o}$. The R^2 value of the former was 0.776, whereas that of the latter was 0.471. Therefore, using the given data, it was estimated that the dielectric permittivity of ILs was mainly contributed by the cation. Second, the most important descriptor was identified. Among the descriptors used in Eq. (10), the H-bonding acidity term of the cation (A^o_c) had the highest correlation with log ε (R^2=0.713), implying that an increase in the H-bonding acidity of the cation leads to a higher log ε value as the sign of the system parameter a_c is positive. Indeed, in the list of ILs, protic ammonium-based cations have higher values because they have higher H-bonding acidity values. Then, the addition of B^o_a to A^o_c led to an increased R^2 value of 0.759. As estimated, the system parameter of b_a also had a positive sign, which means that the H-bonding basicity of the cation affected log ε. Further addition of the S^o_c term helped to increase the predictability to an R^2 of 0.802. The addition of $I^{-,o}$ and S^o_a to the combination of the three terms, that is, A^o_c, B^o_a, and S^o_c, could help to further increase predictability. Note that the dipolarity/polarizability of the cation and anion, that is, S^o_c and S^o_a, negatively affected log ε, as a single correlation of each term with log ε had a negative sign. However, the sign of S^o_a was positive. This may be because the complementary role of the descriptor caused the system parameters to be auto-scaled by multiple linear regression. Overall, the terms related to H-bonding led to a high contribution, followed by the dipolarity/polarizability effect of the ILs. This is reasonable as it has been reported that the number of H-bonding groups is highly correlated with log ε of IL [39] and the H percentage of the molecule and the presence of C-O groups could be used to predict log ε [20].

This study presents an LFER-based prediction model in an *in silico* environment and finds that only five descriptors need to be used to correlate with log ε values of ILs, which helps to understand it more easily on a molecular basis. Its predictability is similar to that based on 2D or 3D parameters by Rybinska-Fryca et al. [20]. As Eq. (10) was developed using a larger dataset than in a previous study [20,21], it is expected that Eq. (10) can be applied to diverse IL structures. However, further experimental and validation studies of the model of Eq. (10) should be performed because the dataset used was limited to only 26 cations and 18 anions. Moreover, it is advisable to develop a temperature-dependent predictive model because ILs can be used in processes where the temperature can vary.

CONCLUSIONS

To investigate the structural effect of ILs on the dielectric constant and derive a more efficient method to estimate it, we developed an LFER model using *in silico* calculated descriptors. The developed model has reasonable predictability, with an R^2 of 0.882 and a standard error of 0.084 log units. The results show that the cationic part has a higher contribution to the dielectric constant than the anionic part. Among the described molecular moments in the developed model, the H-bonding acidity of the cation and H-bonding basicity of the anion considerably increased the dielectric constant of ILs, followed by the dipolarity/polarizability moments of the IL cations and anions. In addition, ionic interactions by anions slightly contributed to the dielectric constant. Therefore, to design ILs, H-bonding and dipolarity/polarizability moments should be considered. These results and discussion are expected to assist in the development of a fast screening method to estimate the dielectric constant of ILs and for efficiently designing the molecular structure of ILs.

ACKNOWLEDGEMENT

This research was supported by the Korean government through the NRF (2017R1A6A3A04003316, 2020R1A2C3009769).

SUPPORTING INFORMATION

Additional information as noted in the text. This information is available via the Internet at http://www.springer.com/chemistry/journal/11814.

REFERENCES

1. A. J. Greer, J. Jacquemin and C. Hardacre, *Molecules*, **25**, 5270 (2020).
2. N. V. Plechkova and K. R. Seddon, *Chem. Soc. Rev.*, **37**, 123 (2008).
3. M. M. Huang and H. Weingartner, *Chemphyschem*, **9**, 2172 (2008).
4. S. C. Moldoveanu and V. David, in *Essentials in modern HPLC separations*, S. C. Moldoveanu and V. David Eds., Elsevier Science Publishing Co Inc, MA (2013).
5. L. W. McKeen (Ed.) *Film properties of plastics and elastomers (Third edition)*, William Andrew Publishing, Boston, MA (2012).
6. J. S. Holbrey and R. Kenneth, *Clean Products and Processes*, **1**, 223 (1999).
7. P. C. S. Costa, J. S. Evangelista, I. Leal and P. Miranda, *Mathematics*, **9**, 3110 (2021).
8. C. W. Cho, U. Preiss, C. Jungnickel, S. Stolte, J. Arning, J. Ranke, A. Klamt, I. Krossing and J. Thöming, *J. Phys. Chem. B*, **115**, 6040 (2011).

9. C. W. Cho, S. Stolte and Y. S. Yun, *Sci. Total Environ.*, **633**, 920 (2018).
10. C. W. Cho and Y. S. Yun, *Environ. Pollut.*, **255**, 113185 (2019).
11. C. W. Cho, Y. F. Zhao, J. W. Choi, J. A. Kim, J. K. Bediako, S. Lin, M. H. Song and Y. S. Yun, *Environ. Res.*, **192**, 110271 (2021).
12. M. H. Abraham and W. E. Acree, *J. Org. Chem.*, **75**, 3021 (2010).
13. M. H. Abraham and W. E. Acree, *Phys. Chem. Chem. Phys.*, **12**, 13182 (2010).
14. M. H. Abraham and W. E. Acree, *J. Chromatogr. A*, **1430**, 2 (2016).
15. C. W. Cho, J. S. Park, S. Stolte and Y. S. Yun, *J. Hazard. Mater.*, **311**, 168 (2016).
16. C. W. Cho, Y. Zhao and Y. S. Yun, *Water Res.*, **151**, 288 (2019).
17. S. Endo, T. N. Brown, N. Watanabe, N. Ulrich, G. Bronner, M. Abraham and K.-U. Goss, Leipzig, Germany, Helmholtz Centre for Environmental Research-UFZ (2015).
18. C. W. Cho, S. Stolte, Y. S. Yun, I. Krossing and J. Thöming, *RSC Adv.*, **5**, 80634 (2015).
19. Y. Zhou, Z. Lin, K. J. Wu, G. H. Xu and C. H. He, *Chin. J. Chem. Eng.*, **22**, 79 (2014).
20. A. Rybinska-Fryca, A. Sosnowska and T. Puzyn, *J. Mol. Liq.*, **260**, 57 (2018).
21. P. Eiden, S. Bulut, T. Köchner, C. Friedrich, T. Schubert and I. Krossing, *J. Phys. Chem. B*, **115**, 300 (2011).
22. M. M. Huang, Y. P. Jiang, P. Sasisanker, G. W. Driver and H. Weingartner, *J. Chem. Eng. Data*, **56**, 1494 (2011).
23. H. Weingartner, P. Sasisanker, C. Daguenet, P. J. Dyson, I. Krossing, J. M. Slattery and T. Schubert, *J. Phys. Chem. B*, **111**, 4775 (2007).
24. H. Weingartner, *Z. Phys. Chemie-Int. J. Res. Phys. Chem. Chem. Phys.*, **220**, 1395 (2006).
25. K. Nakamura and T. Shikata, *Chemphyschem*, **11**, 285 (2010).
26. J. Hunger, A. Stoppa, S. Schrodle, G. Hefter and R. Buchner, *Chemphyschem*, **10**, 723 (2009).
27. C. Wakai, A. Oleinikova, M. Ott and H. Weingartner, *J. Phys. Chem. B*, **109**, 17028 (2005).
28. J. Hunger, A. Stoppa, R. Buchner and G. Hefter, *J. Phys. Chem. B*, **112**, 12913 (2008).
29. A. Stoppa, R. Buchner and G. Hefter, *J. Mol. Liq.*, **153**, 46 (2010).
30. T. Singh and A. Kumar, *J. Phys. Chem. B*, **112**, 12968 (2008).
31. Y. H. Zhao and M. H. Abraham, *J. Org. Chem.*, **70**, 2633 (2005).
32. M. H. Abraham and W. E. Acree, *J. Org. Chem.*, **75**, 1006 (2010).
33. R. G. Y. Parr and Y. Weitao, *Density-functional theory of atoms and molecules*, OUP, Oxford (1989).
34. A. Schäfer, C. Huber and R. Ahlrichs, *Chem. Phys.*, **100**, 5829 (1994).
35. F. Eckert, *COSMOtherm reference manual, version C3.0, Release 15.01.*, Leverkusen, Germany (1999-2014).
36. N. M. O'Boyle, M. Banck, C. A. James, C. Morley, T. Vandermeersch and G. R. Hutchison, *J. Cheminformatics*, **3**, 33 (2011).
37. J. Hunger, A. Stoppa, R. Buchner and G. Hefter, *J. Phys. Chem. B*, **113**, 9527(2009).
38. A. Stoppa, J. Hunger, R. Buchner, G. Hefter, A. Thoman and H. Helm, *J. Phys. Chem. B*, **112**, 4854 (2008).
39. B. L. Shi, *J. Mol. Liq.*, **299**, 112216 (2020).

Supporting Information

In silico prediction and analysis of dielectric constant of ionic liquids

Chul-Woong Cho[*,**,†] **and Yeoung-Sang Yun**[***,†]

*Department of Integrative Food, Bioscience and Biotechnology, Chonnam National University, Yongbong-ro 77, Buk-gu, Gwangju 61186, Korea
**Department of Bioenergy Science and Technology, Chonnam National University, Gwangju 61186, Korea
***School of Chemical Engineering, Jeonbuk National University, 567 Beakje-dearo, Deokjin-gu, Jeonju, Jeonbuk 54896, Korea
(*Received 20 October 2021 • Revised 24 February 2022 • Accepted 24 February 2022*)

Table S1. The studied molecular structures of ionic liquids

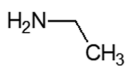
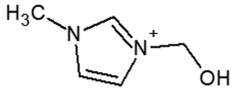
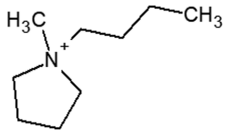
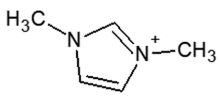
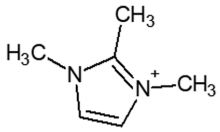
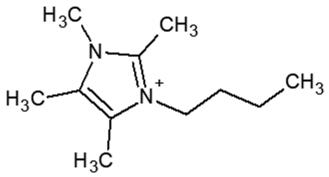
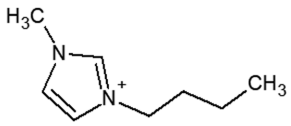
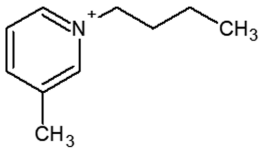
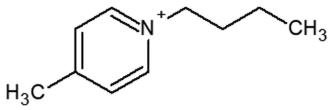
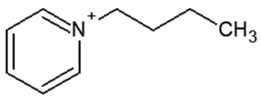
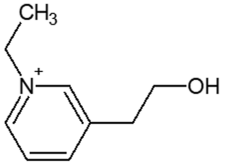
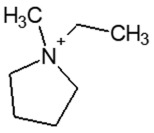
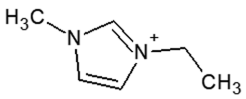
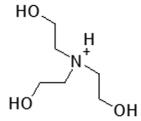
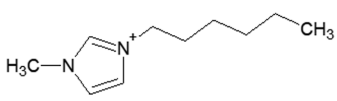
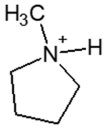
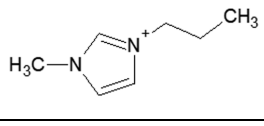
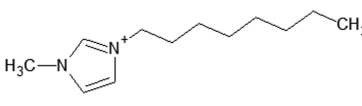
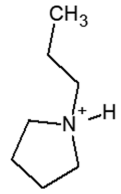
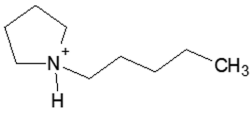
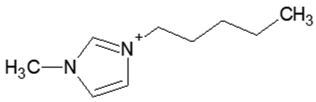
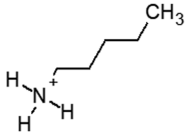
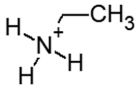
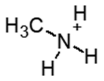
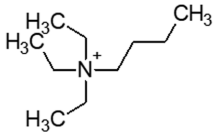
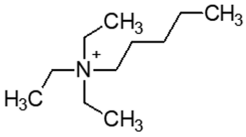
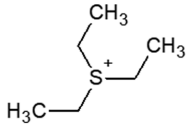
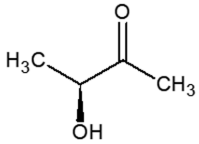
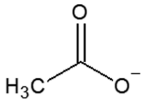
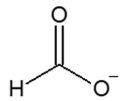
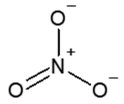
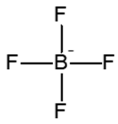
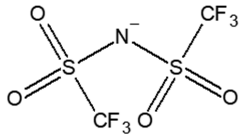
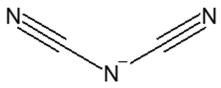
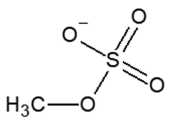
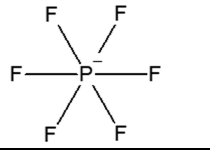
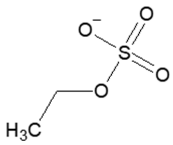
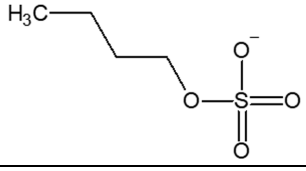
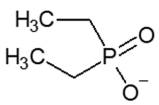
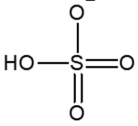
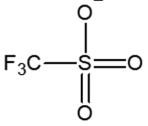
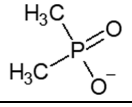
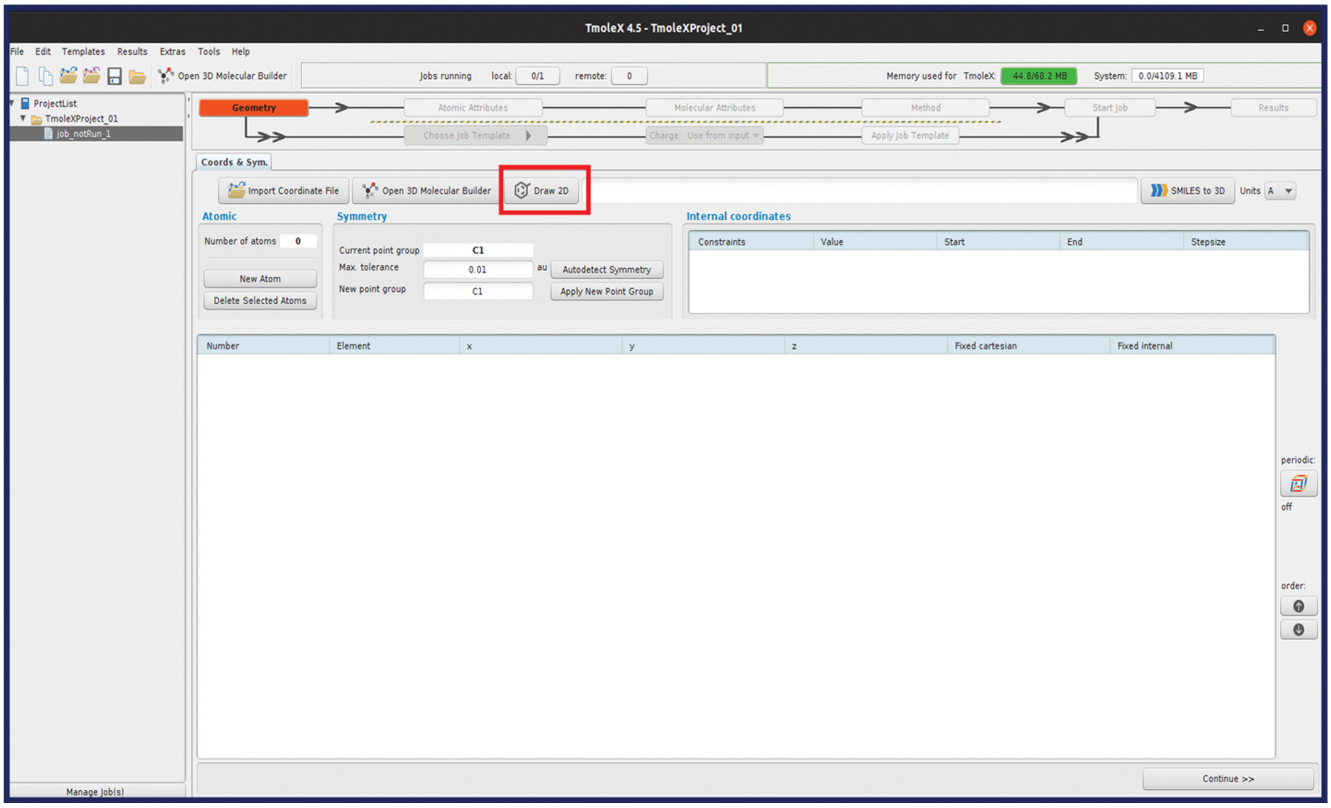
		
[Nhhh2OH]	[IM12OH]	[Pyr14]
		
[IM11]	[IM11-2Me]	[IM14-2Me-4Me-5Me]
		
[IM14]	[Py4-3Me]	[Py4-4Me]
		
[Py4]	[Py2-2OH]	[Pyr12]
		
[IM12]	[Nh(2OH)(2OH)(2OH)]	[IM16]
		
[Pyr1h]	[IM13]	[IM18]

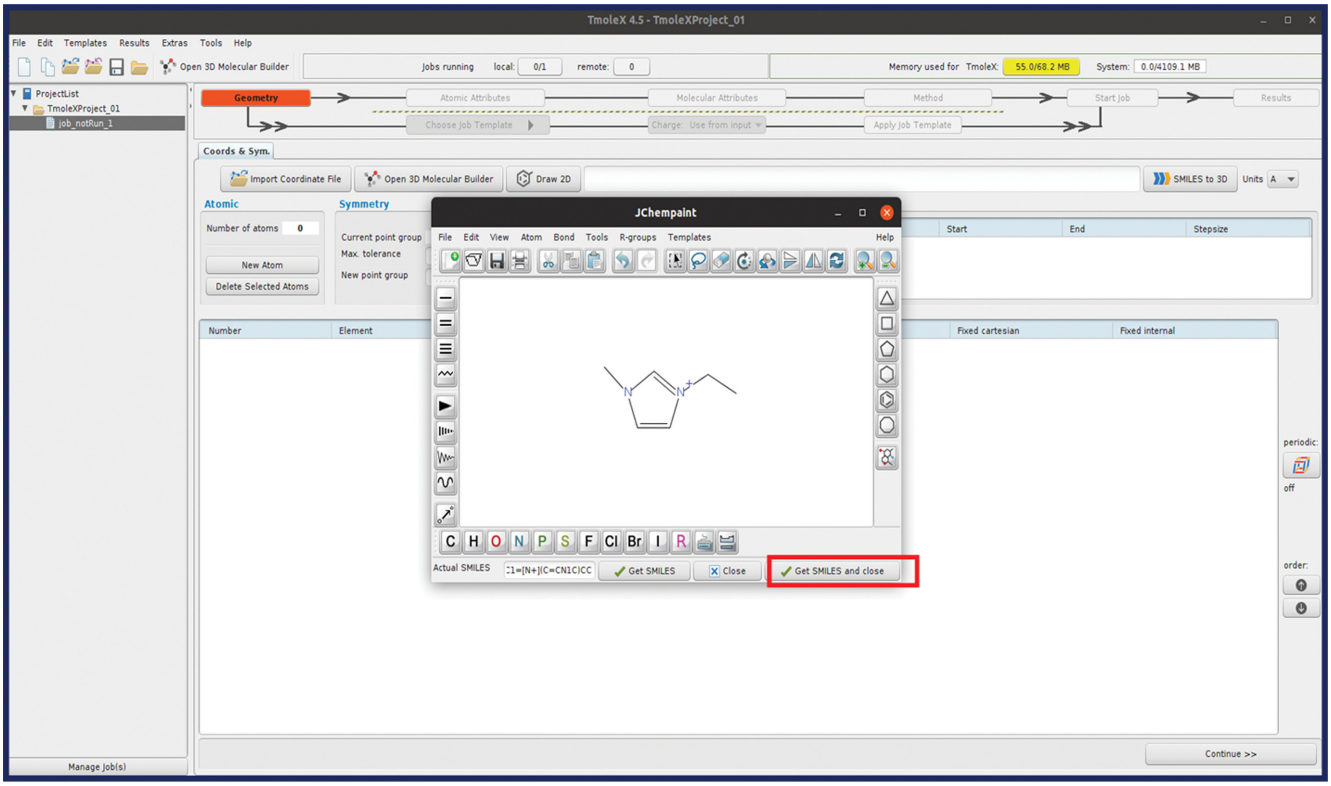
Table S1. Continued

		
[Pyr13]	[Pyr15]	[IM15]
		
[Nhhh5]	[Nhhh2]	[Nhhh1]
		
[N2224]	[N2225]	[S222]
		
[Lac]	[1COO]	[HCOO]
		
[NO_3]	[BF_4]	[$(CF_3SO_2)_2N$]
		
[$N(CN)_2$]	[$1OSO_3$]	[PF_6]
		
[SCN]	[$2OSO_3$]	[$4OSO_3$]
		
[$(CH_3CH_2)_2PO_2$]	[HSO_4]	[CF_3SO_3]
		
[$(CH_3)_2PO_2$]		

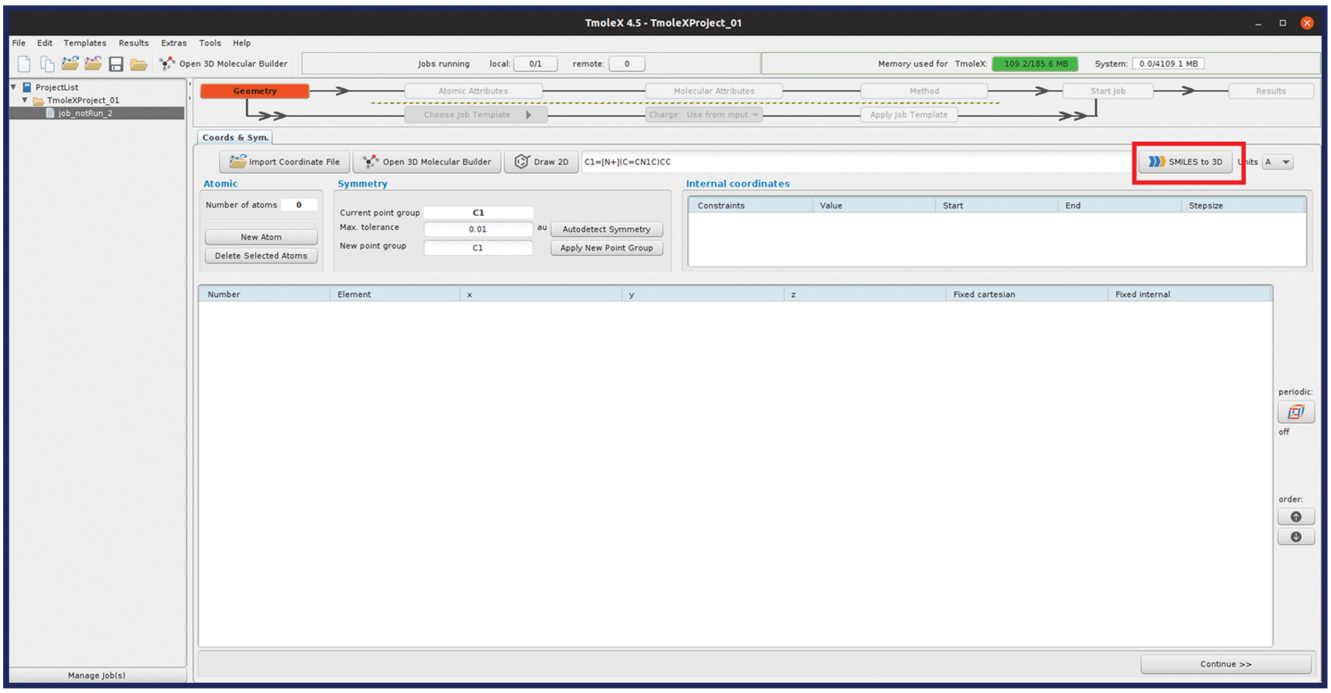
Scheme S1. Calculation steps of ionic liquid cation or anion.
Step 1. Click 'Draw 2D' in red and draw chemical structure.



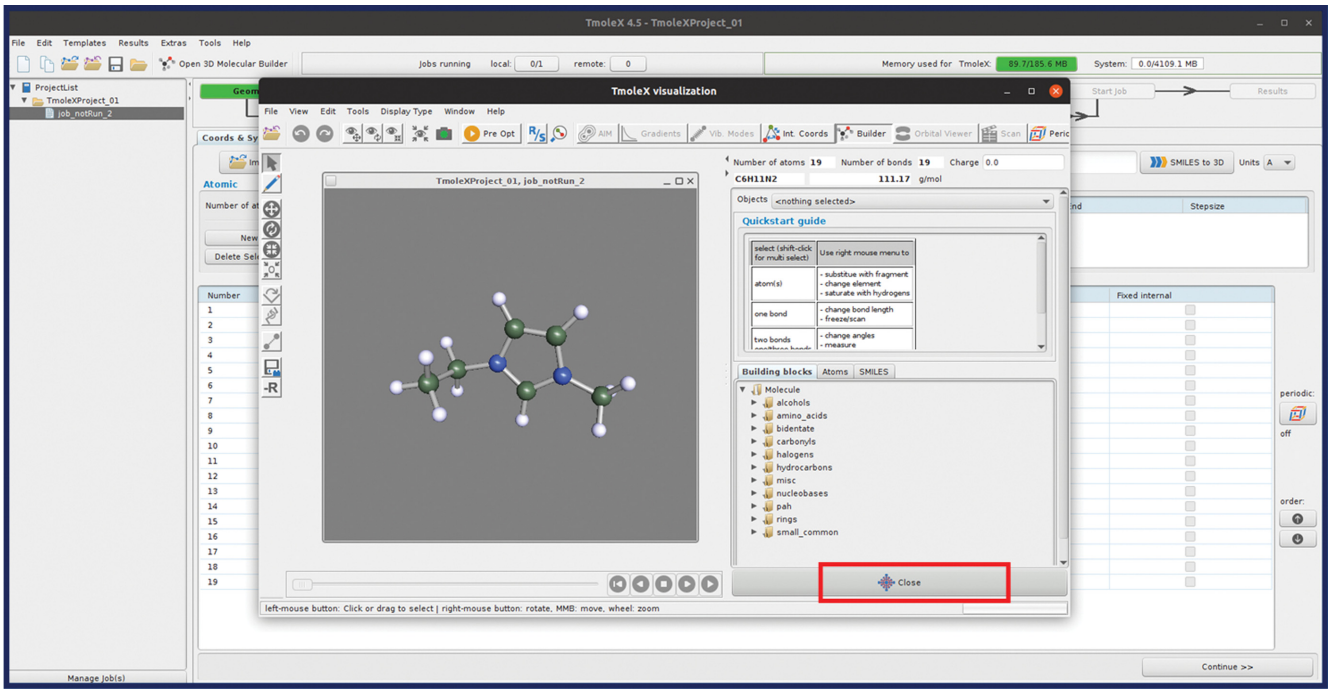
Step 2. After drawing chemical structure in the Jchempaint, click 'Get SMILES and close' in red.



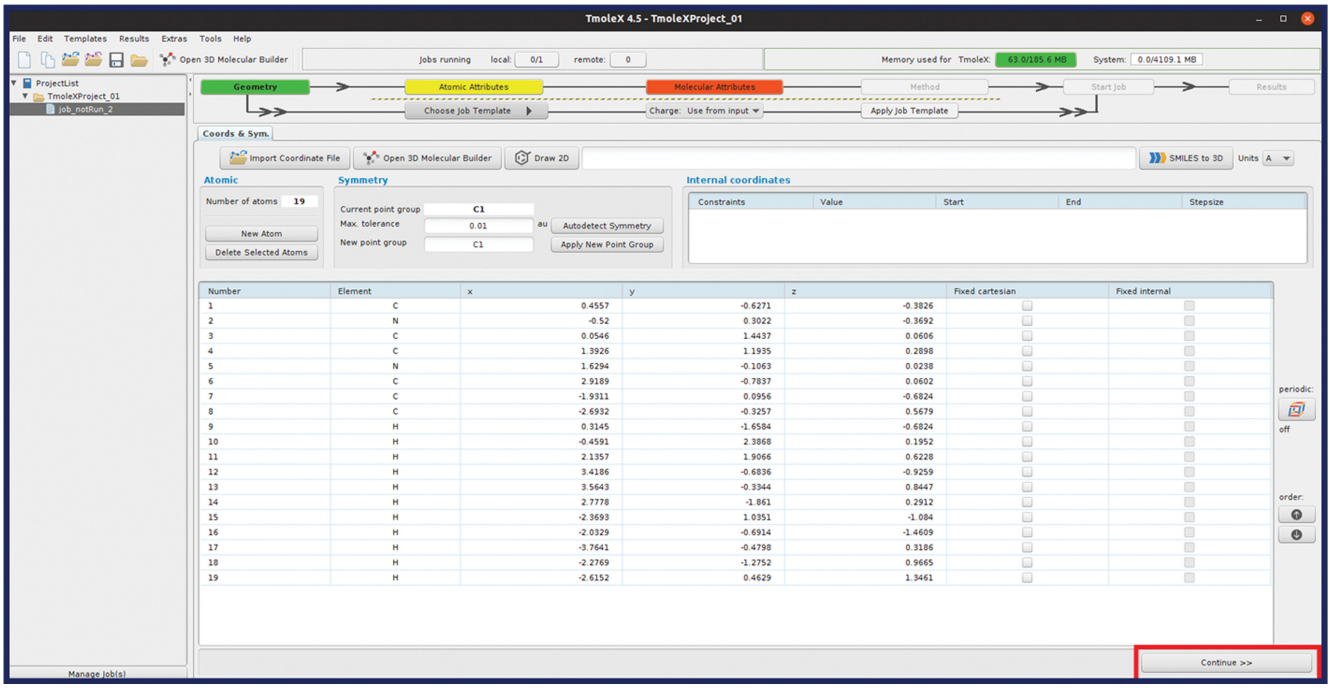
Step 3. After getting the smile code, click 'Smiles to 3D' to change the smile code to 3D structure.



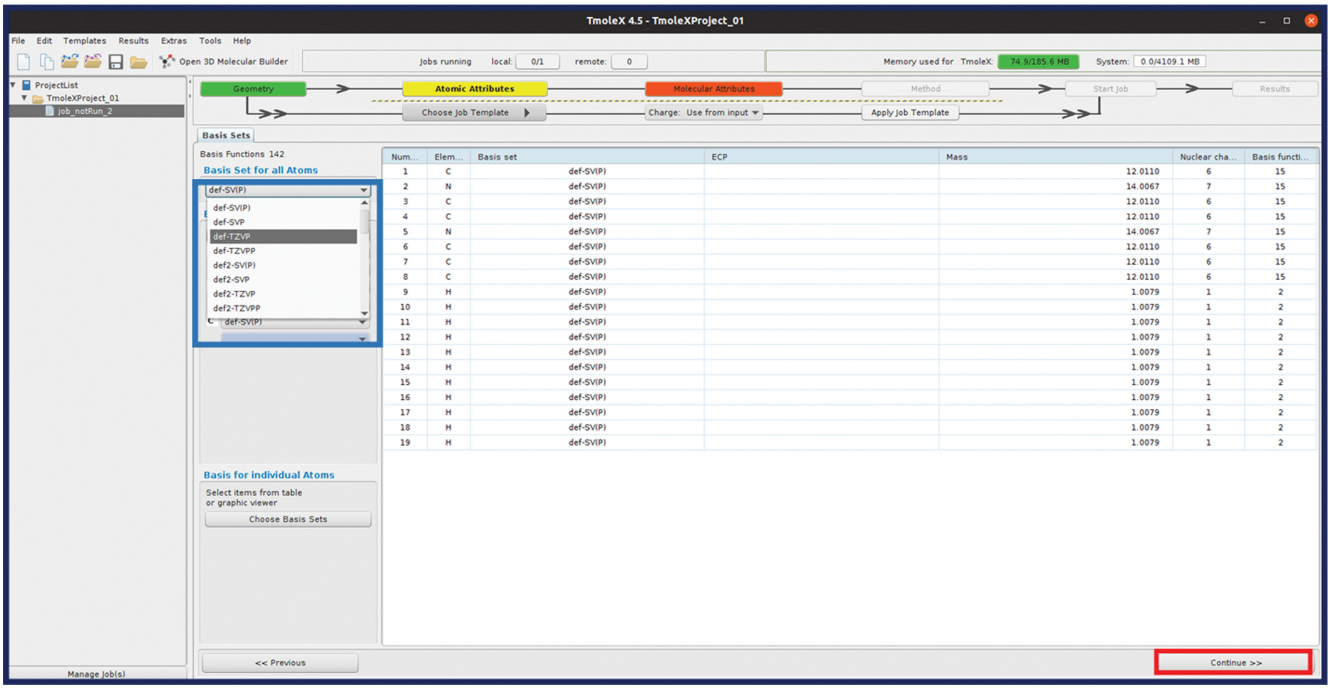
Step 4. After confirming the chemical structure, click 'close' in red.



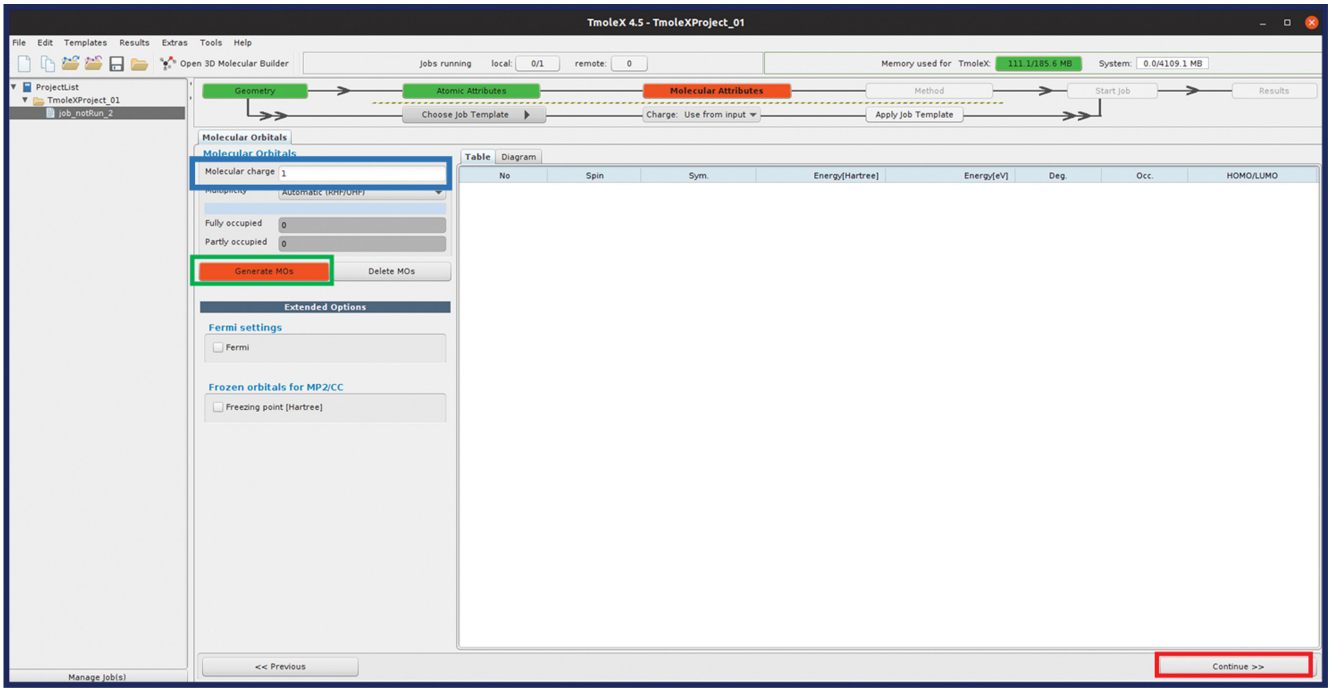
Step 5. Click the 'Continue' in red.



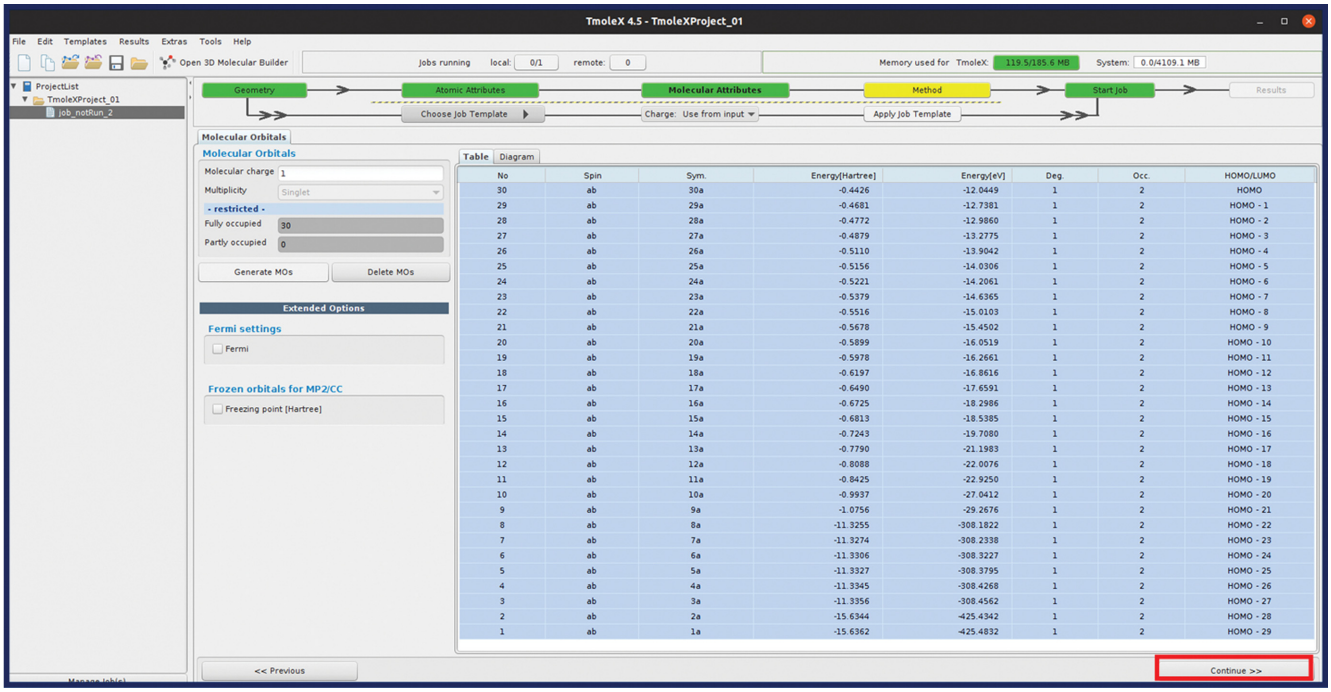
Step 6. Select 'def-TZVP' in blue box and then click the 'continue' in red.



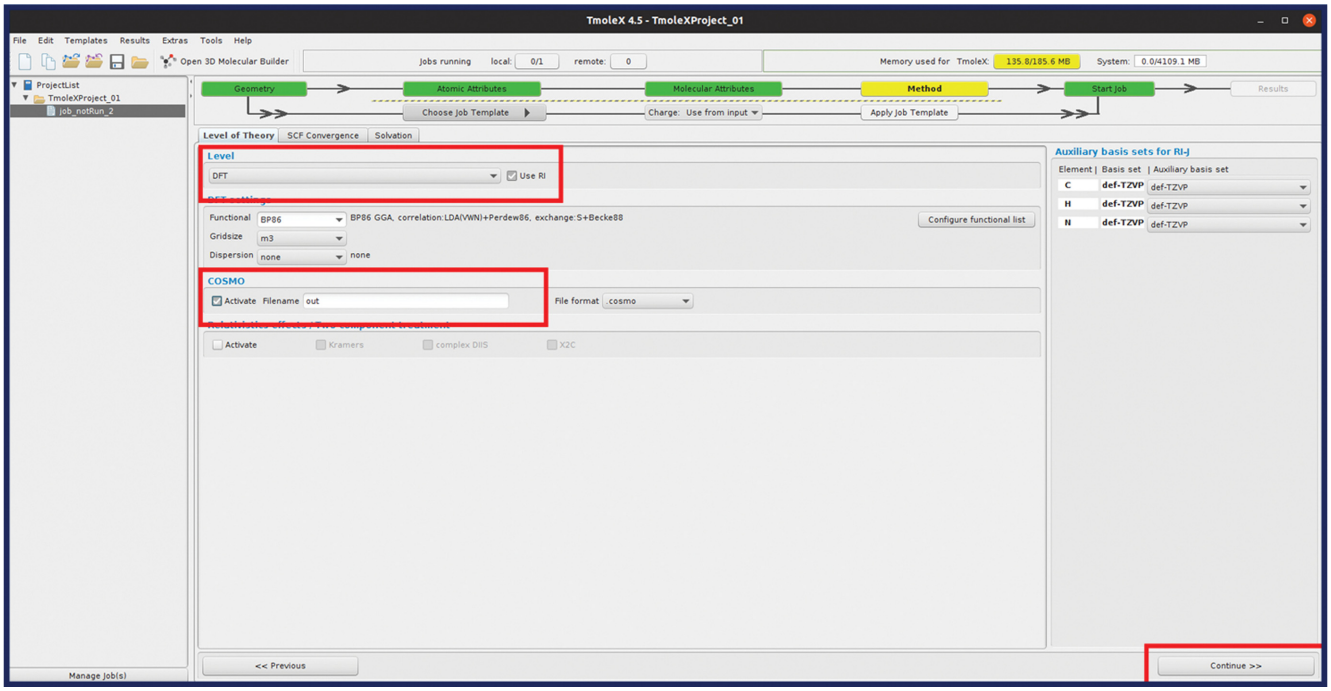
Step 7. add the number of molecular charge in blue box. e.g. monovalent cation=1, monovalent anion=−1, and neutral compound=0. Then, click 'Generate MOs' in green square and click 'Continue' in red square.



Step 8. Click 'Continue' in red square.



Step 9. Select 'DFT', 'USE RI' in red and 'Activate' for COSMO file. Then, click the continue.



Step 10. Click 'Run(local)' in red.