Quantitative Structure and Glass Transition Temperature Relationships Analysis of Polyethylene Analogues

Iqmal Tahir*[,1], Karna Wijaya[1], Bambang Setiaji[1], Yuniawan Hidayat[1], Bernd M. Rode[2]

[1]*Austrian-Indonesian Centre for Computational Chemistry, Chemistry Department, Gadjah Mada University, Jogjakarta 55281, Indonesia.*
[2]*Institute for General, Inorganic and Theoretical Chemistry, Innsbruck University, Innrain 52a, A-6020 Innsbruck, Austria*

Abstract

Quantitative relationships between structure and glass transition temperature T_g and of polyethylene analogues has been studied. The study was done by using molecular modeling of polymer assumed in trimeric compound in their indiotactic form and the calculation was performed by semiempirical AM1 method. The physicochemical properties of molecule was focused on 11 descriptors i.e. atomic net charges of carbon atom as the head and tail of the polymer chain (qC_1 and qC_2), polarizability (α), moment dipole (μ), refractivity index, partition coefficient of n-octanol-water (log P), molecular weight (MW), volume van der Waals (V_{VDW}), molecular surface area, Parachor index and solubility in the water (log SW). Correlation analysis of T_g polymers to those predictors was based on statistical technique of multiple linear regression. The QSPR model resulted is relatively good in terms of accuracy of calculate T_g values of polymer. However the QSPR model is still limited by the validity of the experimental data that were used to derive the regression coefficients of the QSPR equation.

Keywords: *QSPR, glass transition temperature, polymer, polyethylene*

Introduction

In the last several years, there has been increasing interest in organic polymers research. Polymer design and developing new polymer is a subject that was done by many scientists [1]. One tool that was important to study and design new polymer is chemometrics tool namely Quantitative Structure-Property Relationships (QSPR). Actually this subject is developed from medicinal chemistry technique was known as Quantitative Structure-Activity Relationships (QSAR) [2].

In QSPR analysis, there were many properties that could be studied i.e. physical, electrical, optical, and electrochemical properties [3]. An important property of the polymer is glass transition temperature (T_g). This temperature also known as the glass temperature or the glass – rubber transition temperature, is one of the most important properties of amorphous polymers. There were many examples of this polymer i.e. polyethylene, polyacrylate, polystyrene, or polyvinyl [4]. The T_g is difficult to determine experimentally and recently this property have been known empirically that T_g is deeply related with its molecular structure.

Over the past 15 years, there have been numerous attempts to predict T_g's for polymers by various approaches. Wiff et al [5] have been studied of several olygomer prediction using chemical structure as the predictors. Weyland et al [6] used a group additive property (GAP) theory for predicting T_g. The other is Bicerano [7] who used a similar approach by calculating structural parameters and adding increments based on the monomer structure. Most of publication exposed structural parameter to predict T_g value. Caro et al [8] have been published a paper reporting of an energy, volume and mass (EVM) model to predict the value of T_g. In this paper, QSPR is analyzing involved four physico-chemical descriptor variables i.e. van der Waals energy, internal energy, volume and mass, to successfully predict the T_g of aliphatic acrylate and methacrylate polymer.

In the vicinity of T_g, a polymer experiences a sudden increase in the rate of molecular motions and, as a result, undergoes a series of conformational transformations [4]. The torsional oscillation and/or rotation about most of the backbones are activated, which causes a sharp increase and it is effected to T_g value. On the other hand, interaction of the molecule i.e. bonding and non-bonding interaction would give also into T_g value. Representation of the last

phenomena could be study by using electronic parameters i.e. atomic net charge of the atom-atom in polymer molecule, dipole moment and other electronic parameter. Electronic descriptor has been done well to predict biological activity in QSAR analysis of antimalarial compounds [9-11], phenyl ethyl amina [12], benzaldehyde [13], benzoic acid [14] and sunscreen compounds [15].

Utility of electronic structure for polymer have been reported for poly-(arylene ether) using descriptor carried out from molecular orbital calculation of CNDO [16]. In this paper, we used electronic parameter carried out using semimempirical AM1 method as the descriptor to predict T_g of the polyethylene analogues. AM1 methods is based on semiempirical MNDO technique of and its applicability for repeated structure characteristic has been studied successfully [17]. Comparation of three semiempirical techniques i.e. MNDO, AM1 and PM3 with *ab initio* technique 6-31G calculation showed that AM1 is always in qualitative agreement with the *ab initio* results. From these reason we try to use AM1 method applied to calculate electronic structure of the three repeated structure of polyethylene.

Experimental

Preparation of the descriptor

We used 15 polymer compounds of poly ethylene analogues and its value of glass transition temperature (T_g) were listed in Table 1. The program Hyperchem (Hypercube, Inc) was employed to generate models of the polymer. All structures in Table 1 were fully minimized using conjugate gradients Polak-Ribiere algorithm and carried out using molecular orbital calculation AM1. Each polymer was assumed in trimer compound in their indiotactic form. The descriptors that are compiled from this result were atomic net charges of atom C in head position (q_{C1}) and atom C in tail position (q_{C2}), dipole moment (μ) , van der Waals volume (V_{VDW}), and molecular surface area (A).

Table 1 Type of polymer and T_g value of polyethylene homologues

No	Type of polymer	T_g, (^oC)
1	Poly(ethylene)	195
2	Poly(propylene)	233
3	Poly(1,1-dimethylethylene)	199
4	Poly(ethylethylene)	228
5	Poly(1-pentene)	220
6	Poly(4-metil-1-pentene)	302
7	Poly(butylethylene)	220
8	Poly(3,3,3-trifluoropropilene)	300
9	Poly(cyclopenthylethylene)	348
10	Poly(3-cyclopenthyl-1-propilene)	333
11	Poly(cyclohexylethylene)	363
12	Poly(chlorotrifluoroethylene)	373
13	Poly(3-phenyl-1-propylene)	333
14	Poly(3-cyclohexyl-1-propylene)	348
15	Poly(heptafluoropropylethylene)	331

Table 2 Description of the predictors using in QSPR analysis

No	Symbol	Predictor	Unit	Assumption of the model polymer	Calculation method, Software
1	q C_1	Atomic net charge of C_{head}	Coloumb	Optimized trimeric	AM1 semiempirical, Hyperchem
2	q C_2	Atomic net charge of C_{tail}	Coloumb	Optimized trimeric	AM1 semiempirical, Hyperchem
3	α	Polarizability	Å	Monomeric	QSAR Properties, Hyperchem
4	μ	Dipole moment	Debye	Optimized trimeric	AM1 semiempirical, Hyperchem
5	R_D	Refractivity index	-	Monomeric	QSAR Properties, Hyperchem
6	log P	Partition coefficient in n-octanol / water	-	Monomeric	QSAR Properties, Hyperchem
7	MW	Molecular weight	a.m.u	Monomeric	QSAR Properties, Hyperchem
8	V_{VDW}	Van der Waals volume	$Å^3$	Optimized trimeric	AM1 semiempirical, Hyperchem
9	A	Molecular surface area	$Å^2$	Optimized trimeric	AM1 semiempirical, Hyperchem
10	Π	Parachor index	-	Monomeric	Toolkit for Estimating Physichochemical Properties
11	log Sw	Solubility in water	g / 100g water	Monomeric	Toolkit for Estimating Physichochemical Properties

For non repeated-structure we perform monomer form using Hyperchem package program and Toolkit for Estimation of Physicochemical Properties of Organic Compounds (TEPPOC). There were four descriptors compiled from QSAR Properties under Hyperchem i.e. polarizability, refractivity index, partition coefficient in n-octanol/water and molecular weight. Then from the TEPPOC, we have two descriptors i.e. parachor index and solubility in water. Detail of these descriptors was given in Table 2.

Multilinear Regression Analysis

For QSPR evaluation we use logarithmic value of glass transition temperature in reason to get small linear scale relatively to dimension value of the descriptors. If we use T_g value, the highest value is for poly(chlorotrifluoroethylene) that is 373 ^oC and the least is for poly(ethylene) that is 195 ^oC. The correlation models between descriptors and logarithmic value of T_g were evaluated by a stepwise multilinear regression analysis software SPSS for Windows using backward regression method according to the following equation:

$$\log (T_g) = \sum P_{(qi)}q_{(i)} + P_{(\alpha)}\alpha + P_{(\mu)}\mu + P_{(RD}R_D + P_{(\log P}\log P) + P_{(MW}MW + P_{V\,VDW}V_{VDW} + P_{(A)}A + P_{(\Pi)}\Pi + P_{(\log SW)}(\log SW) + D \quad (1)$$

where P is a fitting regression coefficient for corresponding variables and D is a constant.

Evaluating the best QSPR model is looking for the model include number of variables as small as possible relatively, checking the highest of correlation coefficient, minimum standard of error and significantly by F parameter (ratio of F_{calc}/F_{table} is greater than 1 and maximum relatively). We also evaluate the value of PRESS (Predictive residual sum of square) that is given by the following equation :

$$\text{PRESS} = \sum_{i=1}^{N}(y_{pred,i} - y_i)^2 \quad (2)$$

Results and Discussion

Polymer modeling for QSAR purposing is impossible if we use directly to entire molecule because all the polymers possessed high molecular weight. As we know, if the molecular weight is high enough, the terminal groups hold only a very small proportion in a polymer and its effect on the T_g can be ignored [18].

Table 3 Recapitulation of descriptors as the independent variables used for QSAR analysis of polymer

No	Name of polymer	Log $T_{g\ eksp}$	Descriptors										
			q C_1	qC_2	α	μ	R_D	log P	MW	V_{VDW}	A	Π	Log Sw
1	Poly(ethylene)	2.290	-0.157	-0.157	21.937	0.000	9.202	0.793	28.054	44.902	40.213	99.000	-0.620
2	Poly(propylene)	2.367	-0.103	-0.150	37.078	0.230	13.751	1.123	42.081	66.261	57.085	139.000	-0.420
3	Poly(1,1-dimethylethylene)	2.299	-0.038	-0.140	51.828	0.335	18.226	1.556	56.108	87.185	71.614	179.000	-0.260
4	Poly(ethylethylene)	2.358	-0.101	-0.151	52.226	0.233	18.352	1.519	56.108	88.140	77.816	179.000	-0.860
5	Poly(1-pentene)	2.342	-0.097	-0.149	66.269	0.249	22.953	1.916	70.134	108.987	95.115	219.000	-1.340
6	Poly(4-metil-1-pentene)	2.480	-0.103	-0.151	81.190	0.223	27.501	2.246	84.161	130.984	118.786	259.000	-1.630
7	Poly(butylethylene)	2.342	-0.107	-0.152	81.920	0.251	27.554	2.312	84.161	129.719	116.192	259.000	-1.830
8	Poly(3,3,3-trifluoropropilene)	2.477	-0.125	-0.171	39.091	2.772	14.609	1.462	96.052	74.279	70.336	156.400	-1.590
9	Poly(cyclopenthylethylene)	2.542	-0.093	-0.153	93.714	0.202	30.300	2.139	96.172	139.121	115.091	268.900	-1.960
10	Poly(3-cyclopenthyl-1-propilene)	2.522	-0.098	-0.152	109.367	0.222	34.901	2.535	110.199	162.049	137.748	308.600	-2.470
11	Poly(cyclohexylethylene)	2.560	-0.096	-0.159	110.763	0.234	34.901	2.535	110.199	159.973	129.846	308.600	-2.350
12	Poly(chlorotrifluoroethylene)	2.572	0.234	0.051	32.982	0.678	14.255	2.012	116.470	171.136	67.624	153.700	-0.200
13	Poly(3-phenyl-1-propylene)	2.522	-0.095	-0.156	106.155	0.184	38.447	2.737	118.178	156.765	136.369	314.400	-2.540
14	Poly(3-cyclohexyl-1-propylene)	2.542	-0.103	-0.152	124.637	0.218	39.502	2.931	124.226	182.733	157.415	348.600	-2.870
15	Poly(heptafluoropropylethylene)	2.520	-0.120	-0.175	80.006	3.175	24.151	3.036	196.068	128.928	125.212	259.600	-4.640

Table 4 Five selected models and their QSPR equation's parameters for the correlation between predictors and T_g

Model	D	qC_1	qC_2	α	μ	R_D	log P	MW	V_{VDW}	A	Π	Log Sw
Model 1	0.072	-1.435	-10.887	-0.020	-0.230	-0.004	-0.543	0.017	0.020	0.001	---	0.346
Model 2	0.036	-1.795	-10.749	-0.019	-0.226	-0.004	-0.488	0.017	0.020	---	---	0.366
Model 3	0.049	-1.156	-11.319	-0.021	-0.215	---	-0.512	0.015	0.020	---	---	0.307
Model 4	2.349	-3.387	4.056	0.005	---	0.003	-0.249	0.010	---	---	---	0.257
Model 5	2.345	-3.590	4.237	0.006	---	---	-0.236	0.010	---	---	---	0.274

Table 5 Five models and their statistical parameter for the correlation between molecular properties and T_g

Model	Descriptors included in the model	r^2	SE	F_{calc}	Ratio of F_{calc}/F_{table}	PRESS
Model 1	qC_1, qC_2, α, μ, log P, MW, R_D, V_{VDW}, A, log Sw	0.994	0.015	69.489	11.659	0.065
Model 2	qC_1, qC_2, α, μ, log P, MW, R_D, V_{VDW}, log Sw	0.994	0.014	85.373	17.898	0.001
Model 3	qC_1, qC_2, α, μ, log P, MW, V_{VDW}, log Sw	0.992	0.014	95.995	23.131	0.001
Model 4	qC_1, qC_2, α, log P, MW, V_{VDW}, log Sw	0.944	0.034	16.953	4.473	0.009
Model 5	qC_1, qC_2, α, log P, MW, log Sw	0.943	0.032	22.228	6.209	0.643

Here we only focus on their simplest form of the polymer data we used from Katritzky et al.´s work [1]. The molecular model consisting of repeating units, end capped by hydrogen, were chosen as small but it is representative structures to calculate the descriptors. The final molecular descriptors were focused and calculated only for the middle repeating unit of the trimeric structure. However by this procedure, the influence from adjacent repeating units was also taken into account. The terminal groups were not considered separately because all polymer chains were assumed to be terminated by a hydrogen atom. According to this approach, this descriptors are independent of the polymer molecular weight.

Other reason is, in the most cases, T_g remains almost constant with molecular weight above 25.000 amu and all the polymers in the set considered here have molecular weights of > 50.000. As the T_g is itself an intensive property, it was divided by the molecular weight of the repeating unit (MW) and T_g/MW was used an extensive property for the correlation analysis [1]. Table 3 gives a recapitulation of the dependent variable (log T_g) and all of the independent variables (descriptors). Next paragraph we make notes of some interesting descriptors for polymer using in QSPR analysis.

For the value of atomic net charges, all of the compounds have negative values except compound no 12, poly(chlorotrifluoroethylene). All of hydrogen atoms in the vinyl skeleton were changed with one chlorine atom and three fluorine atoms. It would change the electron density in the carbon atoms of the skeleton. The atomic net charges of the carbon atom are positive for this compound. Dipole moment of the compounds are vary positive value except for compound no 1, poly(ethylene). This compound has high level symmetry and all of the chemical bond relate to carbon atom are filled with hydrogen atom.

For multilinear regression we performed using SPSS package program to study the linear correlation between log Tg and all of the descriptors. With stepwise regression method we can have five models (Table 4 and Table 5) that have good correlation in between dependent variable and independent variables.

Evaluation of the best QSPR model

Table 4 reflects mathematical equation of QSPR model using regression parameter related with the descriptors and Table 5 lists statistical parameters that we used for evaluating the best QSPR model.

We can see directly to Table 5 to evaluate which QSPR model was chosen. Based on statistical parameter that we compiled and PRESS value we can evaluate :

a. The determination coefficient value (r^2) of all five models is high relatively for QSAR/QSPR analysis.
b. The value of standard error (SE) is small for all five models.
c. Ratio of F_{calc}/F_{table} is greater than 1 for all five models but the highest is for model 3. We could not compare the value of F without concerning the number of degrees of freedom. The best way is comparing of the ratio of F_{calc}/F_{table}, because we have make relatively for each other model.
d. The PRESS values of all five models are relatively minima but the smallest are for model 2 and 3. PRESS value means of accuracy of the model to predict a theoretical T_g value.

Based on the reason above, we can chose model 3 as the best model for this analysis. The best QSPR equation based on model no 3 using parameters in table 4 could be present by following equation:

$$log\ (T_g) = -0.049 - 1.156.qC_1 - 11.319.qC_2 - 0.021.\alpha - 0.215.\mu - 0.512.logP + 0.015.MW + 0.020.V_{VDW} + 0.307.logS_w$$

In Table 6, we can see t test of the model 3. The most significant descriptors are molecular weight, van der Waals volume and solubility in water. Other descriptors improve indirectly for glass transition temperature.

Table 6 Descriptors involved in model 3 and t test of the descriptors

Descriptor	P	SD	t	significancy
Intercept (D)	0.049	0.428	0.115	0.912
Atomic net charge C head	-1.156	1.087	-1.064	0.328
Atomic net charge C tail	-11.319	3.290	-3.441	0.014
Polarizability	-0.021	0.005	-3.967	0.007
Dipole moment	-0.215	0.035	-6.147	0.001
Log P	-0.512	0.066	-7.746	0.000
Molecular weight	0.015	0.002	7.460	0.000
Van der Waals volume	0.020	0.004	4.823	0.003
Solubility	0.307	0.071	4.355	0.005

In general, the model is covered descriptors of:

a. Molecular weight

This descriptor is classified as the constitutional descriptor that reflects solely the molecular composition of the compound without incorporating any information regarding its geometry or electronic structure. Linear correlation between molecular weight and Tg is clearly founded in this QSPR analysis.

b. Van der Waals volume

This descriptor reflects the geometrical structure or steric effect.

c. Log P and log SW

Lipophilicity of the compounds is represented by descriptors of log P. Log P that defining partition coefficient in n-octanol (non polar solvent) and water (polar solvent) is common used for QSAR of drug. It is reflected how the compound to solve more into non polar or polar solvent. Although log P descriptors are successfully for QSAR analysis, we think it also improves to change the value of T_g in QSPR analysis, because repeated-structure polymer would be changed depends with its lipophilicity. Polymer with less log P value would prefer in high polar system, then if the polymer solves in the water, it would be homogenously distributed in the form of more syndiotactic. In the data, the value of log P that less than 1 is only for poly(ethylene). Others polymer have log P greater than 1 and this phenomenon reflects of al polymer in the water would form of many different stereoregularity. Log P calculation in QSAR Properties under Hyperchem is determined by using atomic approached. Based on this calculation, we know that alkyl branch in a certain polymer would give arising value of log P because hydrogen atom in alkyl improved a positive value of log P. It was clearly if compare the log P value of poly(ethylene) < poly(propylene) < poly(ethylethylene) < poly(buthylethylene). See the value in Table 3. This stereoregularity is possibly would affect to T_g properties. Halogen substitution would give different effect, for F substitution would give positive effect of log P but for Cl, Br and I substitutions would give negative effect and arising follows the atomic number.

d. Atomic net charges and dipole moment

This electronic descriptors reflect the characteristics of the charge distribution of the molecule affected by the difference of electronegativity of two bonded-atom. For example, if we compare the T_gs of poly(1,1-dimethylethylene) and poly(ethylethylene), the high Tg of two polymers is improved by the value of atomic net charges in C head and C tail. The difference of the value for two polymers is because electron distribution ability relatively by different substituent. For poly(ethylethylene), C head atom is substituted with ethyl and hydrogen, but for poly(1,1-dimethylethylene) C head atom is substituted with two methyls. With two methyls in these position, then electron distribution would be easily and homogenous relatively than one ethylene and one hydrogen. The value of partial net charges in poly(1,1-dimethylethylene) would be more negative, consequence of this phenomenon improves for Tg that less than poly(ethylethylene).

If a polymer contains electronegative atoms then inductive effects would be perform because electron effects. Dipole moment would be improved especially by non bonding interaction and therefore mobility of the polymer backbone will be weaker. So the value of glass transition temperature would be increase. We can check in the model, regression coefficient for dipole moment is -0.215 and for polarizability is -0.021. These two descriptors reflect also about the bond length of C head and C tail. The value of Tg would be improved by these parameters.

Reliability of QSPR model

To test the reliability of model no 3, we further finished the multilinear regressions and predicted the T_gs value for the set data. The predicted T_g values are very close to the calculated T_g values see Figure 1. We can see in the figure 1 that deviation of the prediction value is relatively the same with experimentally value. It could be conclude that the QSPR model here is good to illustrate a linear correlation between Tg and the descriptors.

Conclusion

There were good correlation between electronic structure and other molecular structure descriptors with T_gs temperature of a variety of polyethylene analogues. The electronic descriptor related with interaction intermolecular effect especially intermolecular electrostatic interaction and the molecular descriptor related with physical phenomena i.e. steric effect.

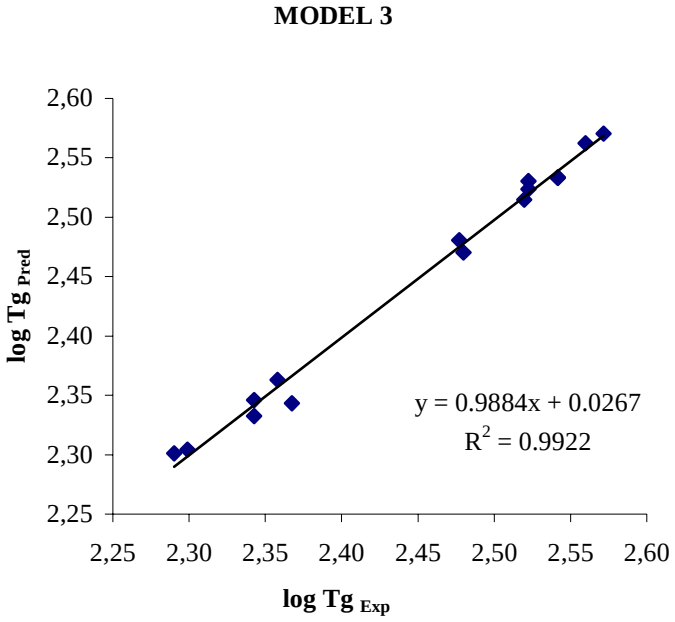

Figure 1 Plot of log Tg polymer prediction versus Tg experimental using model 3

The best model QSPR equation is presented :
$log\ (T_g) = -0.049 - 1.156.qC_1 - 11.319.qC_2 - 0.021.\alpha - 0.215.\mu - 0.512.logP + 0.015.MM + 0.020.V_{VDW} + 0.307.logS_w$

Actually we can make a new model of other polyethylene analogues and from the model we can collect the predictor. Finally we can have theoretical T_g of the polymer. This result encourages the further application of QSPR methods to other classes of polymer, such as polyacrylate, polystyrene and others.

References

[1] A. R. Katritzky, P. Rachwal, K. W. Law, M. Karelson, and V. S. Lobanov, *J. Chem. Inf. Comput. Sci.,* 36 (1996), 880-884.
[2] H. Kubinyi, H., *QSAR : Hansch Analysis and Related Approaches*, VCH Verlagsgesellschaft, Weinheim, 1993, p 14
[3] J.H. Burroughes, D. D. C. Bradley, A. R. Brown, R. N. Marks, K. Mackay, R. H. Friends, P. L. Burns, and A. B. Holmes, *Nature*, 347 (1990), 539
[4] A. R. Katritzky, S. Sild, V. S. Lobanov and M. Karelson, *J. Chem. Inf. Comput. Sci.*, 38 (1998), 300-304.
[5] D. R. Wiff, M. S. Altieri, and I. J. Goldfarb, *J. Polym. Sci*., 23 (1985), 1165-1176
[6] H. G. Weyland P. J. Hoftyzer and D.W. Van Krevelen, *Polymer*, 11 (1970), 79-87
[7] J. Bicerano, Prediction of Polymer Properties, 2nd ed, Marcel Dekker Inc., New York.
[8] A.M. Carro, B. Campisi, P. Camelio, and R. Phan-Tan-Luu, Chemo. Intell. Lab. System, 62 (2002) 79-88
[9] S.U Kokpol, S.V. Hannongbua,., N. Thongrit, and B. M. Rode, *Anal. Sci*., 4 (1988) 565-568.
[10] B. M. Rode, M. G. Schwendinger, S.U. Kokpol, S.V. Hannongbua, and S. Polman, *Monatshefte fur Chemie*, 120 (1989) 913-921.
[11] S. Polman, S. Kokpol, S. Hannongbua and B.M. Rode, *Anal. Sci.*, 5 (1989) 641-644.
[12] I. Tahir, Setiaji, B. and Alim, S.A., Berkala Ilmiah MIPA, X, 1 (2000), 19-37.
[13] I., Tahir, B. Setiaji, and S.A., Alim, Berkala Ilmiah MIPA, XI, 1 (2001), 1-29
[14] S.A. Alim, I. Tahir, and M. F., Pradipta, 2001, Jurnal Nasional Kimia Fisik, III, 2 (2001), 23-26.
[15] I. Tahir, A. Siswandari, B., Setiaji, B., and T.D., Wahyuningsih, *Jurnal Nasional Kimia Fisik*, III, 3 (2001), 73-77
[16] I. Hamerton, B.J. Howlin and V. Larwood, *J. Molec. Graph,* 13 (1995), 14-17
[17] L.Y.A. Davila and M.J. Caldas, *J. Comput. Chem.,* 23 (2002), 12, 1135-1142
[18] C. Cao and Y. Bin, *J. Chem. Inf. Comput. Sci.,* 43 (2003), 643-650