Butwal Campus Journal, Vol. 5, No. 1: 143-159, July 2022
Research Management Cell, Butwal Multiple Campus, Tribhuvan University, Nepal
DOI: https://doi.org/10.3126/bcj.v5i1.50199

COMPUTATIONAL ESTIMATION OF FREE ENERGY USING FORTRAN CODE

Jhulan Powrel
Associate Professor
Butwal Multiple Campus, Department of Physics, Tribhuvan University, Nepal

Article History: *Received 14 May 2022; Reviewed 09 June 2022; Revised 15 July 2022; Accepted 25 July 2022*

ABSTRACT

Theoretical physics based on several computational software. The code for theoretical calculations of free energy is essential in many of the software. One of the codes frequently used in molecular dynamics is FORTRAN (Formula Transformation) code. The free energy for thermal and conformational are found to be $0.2\ \mathrm{KJ/mol}$ and $0.2\ \mathrm{KJ/mol}$ respectively. The average of the total free energy of the theoretical system is found to be $0.40\ \mathrm{KJ/mol}$. The free energy of a system of theoretical molecule of butane is $(-2.75 \pm 0.16)\ \mathrm{KJ/mol}$. The ranges of theoretical and experimental values of free energy are found to be in considerable range. So, the FORTRAN code can also be used in estimating the free energy of the system of molecules.

Keywords: Computational, FORTRAN code, Probability density, Conformational energy, Thermodynamics energy

INTRODUCTION

The free energy is the energy associated with a chemical reaction that can be used to do work. The free energy of a system is the sum of its enthalpy (H) and the product of the temperature (Kelvin) and the entropy (S) of the system i.e., $\mathrm{G = H - TS}$. It is also the change in the enthalpy (H) of the system minus the product of the temperature (Kelvin) and the change in the entropy (S) of the system:

$$\Delta G = \Delta H - T\Delta S$$

The change in free energy that occurs when a compound is formed form its elements in their most thermodynamically stable states at standard-state conditions. In other words, it is the difference between the free energy of a substance and the free energies of its elements in their most thermodynamically stable states at standard-state conditions. The standard-state free energy of reaction can be calculated from the standard-state free energies of formation as well. It is the sum of the free energies of formation of the products minus the sum of the free energies of formation of the reactants. The free energy is the thermodynamic free energy and is the useful work that may be extracted from a closed system. It is the function of energy of system (enthalpy), entropy and temperature of the system (Kittel & McEuen, 1976). It may be zero / positive or negative. Free energy calculation results reproduced for different model of water in experimental work are in well agreement with computational simulation. Free energies for different models are within 1.6 kJ mol^{-1} and entropies are within 3.6 $\mathrm{JK}^{-1}\ \mathrm{mol}^{-1}$. Approximately two-thirds of the entropy comes from translation, a third from rotation, and 5% from conformation. Intermolecular quantum effects for the classical oscillator lying within 0.5 to 0.7 KJmol^{-1}, which is higher than that in the quantum case. The method of thermodynamic integration, direct particle insertion, SMD and grand canonical Monte-Carlo are the different way for calculation of free energy. Free-energy calculations have a fundamental role in the understanding of many natural phenomena ranging from protein folding up to polymorphic transitions in solids. To this end, molecular dynamics (MD) has been extensively used together with a series of algorithms aimed at extending its capabilities far beyond those allowed by straightforward MD. Among some of the most popular enhanced sampling techniques are umbrella sampling (Stradner et al., 2007), Jarzynski equation based methods and umbrella sampling (Jarzynski, 1997). So umbrella sampling is introduced here.

Statistics of free energy

In the case of N particle system in 3D space the system is defined by the Hamiltonian which is the function of position and momentum given by,

$$H(p,q) = \sum_{1}^{N} \frac{P_i^2}{2m_i} + V(r_1, r_2, r_3, \dots\dots, r_N)$$

Thus,

$$F(N,V,T) = -K_B T * \ln \left[\frac{1}{h^{3N}} \iint \exp\left[-\frac{H(p,q)}{K_B T}\right] dpdq\right.$$

Here V-the volume of the system, T-absolute temperature of the system, K_B-Boltzmann constant and h-Planck's constant. As free energy is the change in the energy of the two corresponding state of a closed system hence it is expressed as,

$$\Delta F = F_B - F_A = -K_B T \ln \frac{N_B}{N_A}$$

Thermodynamic integration (TI)

As the state is changed the correspondingly interaction parameter is also changed in between $\lambda = 0$ and $\lambda = 1$, so the Hamiltonian of the system can be written as (DeVries & Hamill, 1995),

$$H \equiv H(p,q,\lambda)$$

Now for state A,

$$H_A(p,q) = H(p,q,\lambda_A),$$

and for state B,

$$H_B(p,q) = H(p,q,\lambda_B)$$

i.e. the Hamiltonian is function of λ, thus the equation for F gives,

$$\frac{dF(\lambda)}{d\lambda} = \left\langle \frac{\partial H(\lambda)}{\partial \lambda} \right\rangle_{\lambda}$$

So the free emery difference of state A and state B of anmolecular system is ,

or, $F_{\lambda_B} - F_{\lambda_A} = \int_{\lambda_A}^{\lambda^B} \left\langle \frac{\partial H(\lambda)}{\partial \lambda} \right\rangle_{\lambda} d\lambda,$

which is the thermodynamic integration (TI) formula. It gives the ensemble average, $\left\langle \frac{\partial H(\lambda)}{\partial \lambda} \right\rangle_{\lambda}$ from the simulation at a series of $\lambda's$ between $\lambda_A = 0$ to $\lambda_B = 1$ (after integration). Now,

$$\text{or, } F_{\lambda_B} - F_{\lambda_A} = \sum_{1}^{\infty} \frac{dF(\lambda)}{d\lambda} (\lambda_B - \lambda_A)^n / n!$$

Umbrella sampling

The method of umbrella sampling is introduced by Torrie and Valleau in 1974 – 1977. In this method blue moon ensemble approach along with some restrained factor of biasing potential instead of reaction coordinate is used. The biasing potential is the harmonic one called umbrella potential as,

$$W(f_1(r_1, r_2, \dots, r_N), S) = \frac{1}{2} K(f_1(r_1, r_2, \dots, r_N) - s)^2.$$

A total potential of $U(r) + W(f_1, \mathrm{s})$ is incorporated in molecular dynamics simulation (Tuckerman, 2001). The U's are choosen at various intermediate points $s^{(1)}, \dots s^{(n)}$ in between initial $s^{(i)}$ and final $s^{(f)}$. Each of the reaction coordinate $\mathrm{q} = f_1(r)$ lies in between two end point values. Thus the new probability distribution function $\mathrm{P}(\mathrm{s}, s^k)$ gives the profile of the true free energy throughout the entire range of the reaction coordinate. An unbiasing techniques of weighted histogram analysis method (Yang & MacKerell Jr, 2015) is implemented in finding the biased probability distribution from each of the molecular dynamics,

$$P(s, s^k) = \exp\,[\beta A_k] \int d^N r e^{-\beta\ (r)} e^{-\beta} W(f_1, s^k) \delta(f_1(r) - q),$$

Where, A_k the biased free energy and

$$e^{-\beta A_k} = \int d^N r e^{-\beta U(r)} e^{-\beta} W(f_1, s^k) = e^{-\beta A_0} \left\langle e^{-\beta W(f_1(r), s^k)} \right\rangle$$

This average is obtain by considering unbiased potential $\mathrm{U(r)}$ and

$$e^{-\beta A_0} = \int d^N r e^{-\beta U(r)}$$

which is the configuration partition function for unbiased case and $\exp\,(\beta A_k)$ is the normalization constant. Now the corresponding unbiased probability distribution function is,

$$P'(q) = e^{-\beta}(A_k - A_0) e^{\beta} W(q, s^k) P(q, s_k)$$

Hence the full probability distribution function is obtained in terms of other two is,

$$P''(q) = \Sigma_1^n C_k(q) P_k(q) = {}_1^n C_k(q) \{ e^{-\beta} (A_k - A_0) e^{\beta} W(q, s^k) P(q, s_k) \}$$

Here $C_k(q)$ is optimised for accurate P″(q) and is $\Sigma C_k(q) = 1$. In oredr to find the coefficient $C_k(q)$ using WHAM (Yang & MacKerell Jr, 2015), let us write $H_k(q)$, biased histogram of each MD, then

$$P(q, s^k) \approx \frac{1}{n_k \Delta q},$$

Where, Δq the bin width and n_k the number of configuration sampled for k simulation. The statistical error for kth window is then,

$$\sigma_k^2 = \epsilon_k(q) H_k(q) / (n_k \Delta q)$$

This epsilon measure the deviation of numerically sampled and actual distribution in the kth window. Now the corresponding unbiased factor is,

$$\sigma'^2_k = \left\{ e^{-2\beta} (A_k - A_0) e^{2\beta W(q,s^k)} \sigma_k^2 \right\}$$

To minimize the total error let us introduce a multiplier λ as, in the error function,

$$\Sigma^2 = \sum_{k=1}^{n} C_k^2(q) \left\{ e^{-2\beta} (A_k - A_0) e^{2\beta W(q,s^k)} \frac{\epsilon_k(q) H_k(q)}{n_k \Delta_k} \right. - \lambda (\Sigma_{k=1}^n C_k(q) - 1)$$

Taking derivation and then on solving in respect of λ

$$C_k(q) = \frac{\lambda n_k \Delta q}{2 \epsilon_k H_k(q) e^{-2\beta} (A_k - A_0) e^{2\beta W(q,s^k)}}$$

Now above equation gives,

$$\lambda \sum_{k=1}^{n} \frac{n_k \Delta q}{2 \epsilon_k H_k(q) e^{-2\beta} (A_k - A_0) e^{2\beta \quad (q,s^k)}} = 1$$

$$\text{or, } \lambda = \frac{1}{\Sigma_1^n \; n_k \Delta q \left[2 \epsilon_k H_k(q) e^{-2\beta} (A_k - A_0) e^{2\beta W(q,s^k)} \right]}$$

Now from the above equation, we write,

$$C_k(q) = \frac{n_k/\epsilon_k H_k(q) e^{-2\beta}(A_k - A_0) e^{2\beta W(q,s^k)}}{\sum_{j=1}^{n} n_j\left[\epsilon_j(q) H_j(q) e^{-2\beta}\left(A_j - A_0\right) e^{2\beta \quad (q,s^j)}\right]}$$

Here we assume that the biased histogram $H_k(q)$ in each umbrella is well estimated and the error function ϵ_k(q) is equal in all n umbrella windows. Thus,

$$H_k(q) \propto e^{\beta}(A_k - A_0) e^{\beta W(q,s^k) P''(q)}$$

Hence the constant can be rearranged as,

$$C_k(q) = \frac{n_k e^{\beta}\left(A_k e^{-\beta W(q,s^{(k)})}\right)}{\sum_{j=1}^{n} n_j e^{-(q,s^{(j)}}}$$

Therefore the distribution is

$$P''(q) = \frac{\sum_{k=1}^{n} n_k P'_k(q)}{\sum_{k=1}^{n} n_k e^{-\beta(A_k - A_0)} e^{-\beta W\left(q,s^{(k)}\right)}}$$

$$\text{or, } e^{-\beta(A_k - A_0)} = P''(q) e^{-\beta W\left(q,s^{(k)}\right)}$$

$$\text{or, } P'' \approx \frac{1}{\sqrt{2\pi\sigma_k^2}} * e^{-(q-q_k)^2}/2\sigma_k^2$$

Where q_k and σ_k^2 are obtaine by MD simulation.

Computation by FORTRAN

Computational simulations are first and time saving technique in modern theoretical physics. Different multi-step calculations are tried to solve by writing computational programming code. Such a basic programming for calculating free energy is FORTRAN program (Raijlich et al., 2001). In the first part of the module, it needs the FORTRAN programming language. To write and run programs in FORTRAN one requires to use a simple text editor to write the code and a compiler to convert it into an executable so that the computer can run. With proper FORTRAN development environment, one can write FORTRAN code, compile it, and run it in a operating system like Linux. Several basic

FORTRAN program (Moses, 1988) like f77, f95 and f90 are available which included as a sample file that calculates the average of the numbers in a file called numbers.txt. Below the code is a line-by-line annotation. Note that in Fortran any line beginning with the character ! is a "comment," which is not translated into the executable program by the compiler. These comments are simply to help the human reader of the code understand it better. All of the code include comments to help or make them more human readable. Linux (Ubuntu) operating system with suitable PC with FORTAN program is used (Yatani et al., 2009) for theoretical estimation of free energy in FORTRAN90 code.

THEORY AND METHODOLOGY

Computational free energy estimation

The conventional method of numerical solution for free energy is non-Boltzmann sampling method (Allen & Tildesley, n.d.) considering the configurationally partition function as,

$$Q_{NVT}^{ew} = \frac{1}{\langle \exp\,(\beta\Psi)\rangle_{NVT}}$$

As per metropolis Monte Carlo (MC) the potential is negative or small positive. So, umbrella sampling is chosen for estimating the free energy. For any two fluid A and B let us write the potentials $\Psi(\mathrm{r})$ and $\Psi_0(r)$ respectively then the free energy of reference fluid can be written as (Callister, 2007),

$$A - A_0 = -k_B\,\mathrm{Tln}\left(\frac{Q}{Q_0}\right) = -k_B\mathrm{T}\,\ln\,(\langle(\exp\,(-\beta\Delta\Psi)\rangle_0\,)$$

The free energy depends on different factors and hence can be calculated as, thermodynamically free energy, van-der-Waal free energy, electrostatics free energy, L-J free energy, proper and improper dihedral free energy and solvation free energy. Among them here is the consideration of the dependency of the free energy only with the thermal (β) and density function (conformational) $\rho_0(\Psi)$. Now,

$$\frac{Q}{Q_0} = \int_{-\propto}^{\propto} d(\Delta\Psi)\exp\,(-\beta\Delta\Psi)\rho_0(\Delta\Psi)$$

Where , $\rho_0(\Delta\Psi)$ is the density representing term of configuration r as,

$\Psi(r) = \Psi_0(r) + \Delta\Psi$.

This ρ_0 contains Boltzmann function $\exp(-\beta\Psi_0)$ Thus the the variation in $\exp(-\beta\Delta\Psi)$ is Boltzmann distribution function (Sevcik, 2017) which favors the configuration with large negative values of$\Delta\Psi$. Torrie and Valleau suggested the general density function as,

$$\rho_w(r) = W(r)\exp(-\beta\Psi_0(r))/\int drW(r)\exp(-\beta\Psi_0(r))$$

Where $W(r) = W(\Delta\Psi(r)$. For this numerical calculation we are considering weighted function for representing the thermal dependency with function $W(-\beta\Delta\Psi)$ which is expressed as,

$$\Delta A = y = W(-\beta\Delta\Psi)$$

Again the density functional dependency is,

$$\Delta A = \rho_0(\Delta\Psi)$$

It is difficult to get weighted function for density dependencyρ_0, even though one can use exponential function as free energy calculation purpose. As it decreases with the rise in temperature, so an exponential function can be used for free energy calculation,

$$\Delta A \propto \exp\big(\rho_0(\Delta\Psi)\big)$$
$$Or, \Delta A = C\exp(\rho_0(\Delta\Psi))$$

Thus the free energy difference is the product of these two functions as given by equation (Allen & Tildesley, n.d.),

$$\Delta A = -K_bT\ln(\langle\exp(-\beta\Delta\Psi \quad _0)\ \rangle\ )$$

FORTRAN code programming

For writing the FORTRAN code of the numerical solution of the free energy let us write the equations,

$y = C\exp(-\beta\Delta\Psi)$.

It is also expressed as,

$$y = a0 * \exp\left((\Delta\Psi/(a3 * (x - a1))/a2\right)$$

And for density functional dependency,

$$y = \Delta A = C \exp\left[\rho_0(\Delta\Psi)\right]$$

It is also expressed as,

$$y = a0\exp\left(-(x - a1)^* * 2/a2\right)$$

Here, a0, a1, a2 and a3 are arbitrary constants.

Code I

The code for conformational contribution of free energy is,

!Conformational free energy is expressed as function of temperature.

!$\mathrm{P}(\mathrm{E}_k) = \mathrm{C}\exp\left\{\frac{-(E_k - \langle E_k\rangle)^2}{3NK_B^2T_0^2}\right\}$

!or , $y = a0 * \exp\left(-(x - a1)^2/a2\right)$

!with, $a0 = 1$; Initia lvalue

!$a1 = 315K$; (AverageTemperature),

!$a2 = \sigma^2 * 128$ K

```
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!programexpo
implicitnone
integer : : i
integer : : x
real : : y
real :: a0, a1, a2, a3
open (20, file = "conformational −FE13. dat")
```

```
a0 = 10E − 14

a1 = 0

a2 = 1

a3 = 0.00138657

!change a3 from 0.00118657 to 0.00198657

do x = 250,360

!temperaturerangemayvarieasrequiredhereis jixedwithin 250 to 360

y = a0 ∗ exp ((15/(a3 ∗ (x − a1)))/a2)

write (∗,∗)x, y

write (20,∗)x, y

!

enddo

endprogram

!graph1isforT315K

!graph 2 for 310 K

!andgraph3 for 320k

!changeonly05, 1015 − −45

!changedata
```

Code II

The code for thermal contribution of free energy is,

```
!Thermal contribution of Free energy is expressed as function of temperature.
```

!made by jhulanpowrelcdptu

! $\mathrm{P}(\mathrm{E}_k) = \mathrm{Cexp}\left\{\frac{-(E_k - \langle E_k \rangle)^2}{3NK_B^2T_0^2}\right.$

!or, $\mathrm{y} = \mathrm{a0} * \exp\left(-(\mathrm{x} - \mathrm{a1})^2/a2\right)$

! with, $\mathrm{a0} = 1$; Initialvalue

! $a1 = 315K$; (AverageTemperature)

! $a2 = \sigma^{2*}128$ K. !!programexpo

implicitnone

integer : : i

integer : : x

real :: y

real : : $a0, a1, a2$

open $(20$, file $=$ "$FE -$ Therm $_{9}28.dat$")

$a0 = 1$

$a1 = 360$

!changea 1 from 250 to360buthereis fixedvalueat $360a2 = 928$

!changea2from $0.28t0928$

$dox = 0{,}450$

!xisalso fixedat 0,450 onecanvariewithdifferentrangeoftemperarture

$y = a0 * \exp\left(-(x - a1) ** 2/a2\right)$

write $(*,*)x, y$

```
write (20,*)x, y

enddo

endprogram

!graph1isfor 250K

!graph 2 for 255k

!andgraph3 for 260 ketc
```

RESULTS AND ANALYSIS

The free energy of a theoretical system with considerable range of parameter is estimated computationally using FORTRAN code. For this purpose we have estimated the conformational as well as thermal free energy separately and then added them to find the resultant free energy of the system.

Plots of conformational and thermal free energy

By considering the parameter of a3 ranging from 0.28 to 928 with interval of 100 the temperature versus free energy graph are plotted for both conformational and thermal contribution of energy.

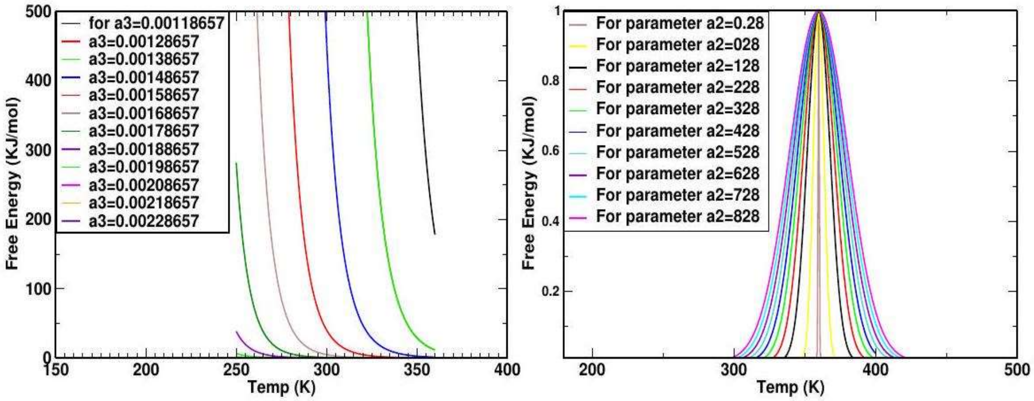


Figure 1: Free energy due to a) conformational and b) due to thermal contribution.

The Figure 1 a) is graph of conformational free energy and figure 1 b) is of thermal contribution of free energy. The different plots are indication of the corresponding free energy. For our convenient we have chosen few of them for estimation of the free energy.

Plots of resultant of free energy

The theoretical plots of conformational and thermal free energy are plotted by Allen and Tildiesely (Figure 2 a)).

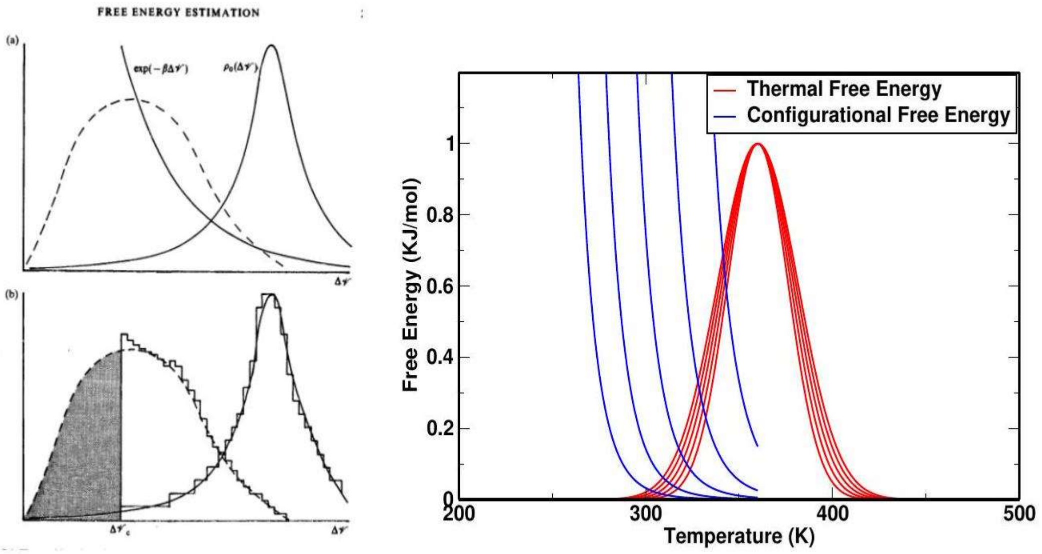


Figure 2: Free energy graph of a) theoretical value b) computational

The resultant plot is shown with dotted line in Figure 2 a) (Allen & Tildesley, n.d.). Selected plots of thermal and conformational free energy are again combines to get the resultant of their free energy as in Figure 2 b).

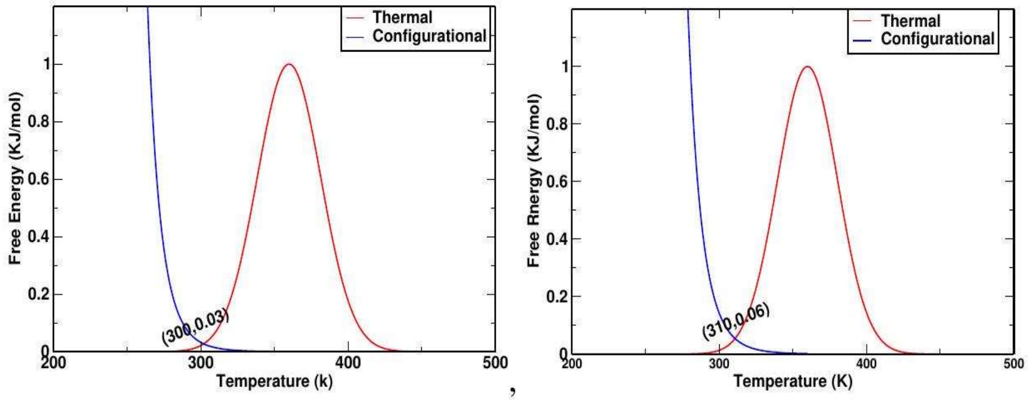
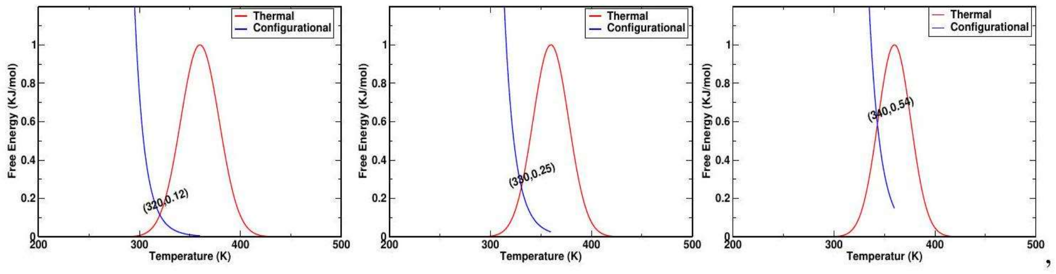


Figure 3: Free energy graph of thermal and conformational contribution at a) 300 K and b) 310 K of temperature.

Individual graphs of thermal and conformational free energy are observed to find their contribution to the total free energy and are tabulated in Table 1. Figure 3a) and 3 b) are the free energy graph giving the energy at temperature of 300 K and 310 K.



a) b) c)

Figure 4: Free energy graph of thermal and conformational contribution at a) 320 K and b) 330 K and at c) 340 K of temperatures

Figure 4 a), 4 b) and 4 c) are the individual graph of thermal and conformational contributions of free energy at temperature of 320 K, 330 K and 340 K respectively.

Estimation of Free Energy

Table 1: Estimated values of free energy ($\mathrm{KJ/mol}$) due to conformational and thermal.

From	Temperature (K)	Conf.	Thermal	Total free energy(KJ/mol)
Computation	300	0.03	0.03	0.06
	310	0.06	0.06	0.12
	320	0.12	0.12	0.24
	330	0.25	0.25	0.50
	340	0.54	0.54	1.08
Experimental(n-Butane)	350	–	–	9.03
Theoretical(n-Butane)	350	–	–	-2.75 ± 0.16

The average of the estimated free energy is calculated as $= 0.4\ \mathrm{kJ/mol}$ computationaly. The free energy calculation has also been performed by computational simulation and is agreed with previously performed work (Torrie & Valleau, 1977). The Gibb's free energy for butane has also been estimated from simulation as, $= 2.75 \pm 0.16\ \mathrm{kJ/mol}$ (Torrie & Valleau, 1977). And the free energy of alkane obtained from simulation is, $(2.75 \pm 0.16)\ \mathrm{kJ/mol}$.

CONCLUSION

FORTRAN code can be used for solving the numerical problem like free energy estimation in molecular dynamics. From graph the average of the computational free energy of a theoretical system is estimated as $0.4\mathrm{KJ/mol}$, which is closed and within the range of theoreticaly calculated (Table 1) and experientally (Table 1) observed free energy.

DATA AVAILABILITY STATEMENT

All data that support the findings of this study are included within the article (and any supplementary files).

AUTHORS CONTRIBUTION

JP carried out all the works and idea of study. Authors read and approved the final manuscript. All data and paper are available with the correspondent author.

REFERENCES

Allen, M., & Tildesley, D. (1987). *J. computer simulations of liquids 1987*. Oxford University Press, New York.

Callister, W. D. (2007*). Materials science and engineering: An introduction. John Wiley & Sons.* Inc.: Utah.

DeVries, P. L., & Hamill, P. (1995*). A first course in computational physics. American Association of Physics Teachers*. USA.

Jarzynski, C. (1997, Apr). *Nonequilibrium equality for free energy differences. Phys. Rev. Lett., 78, 2690-2693.* Retrieved from https://link.aps.org/doi/10.1103/PhysRevLett. 78.2690 doi: 10.1103 /PhysRevLett. 78.2690.

Kittel, C., & McEuen, P. (1976). *Introduction to solid state physics, vol 8 wiley.* New york.

Moses, G. A. (1988*). Engineering applications software development using fortran 77. John Wiley & Sons, Inc.*

Raijlich, V., Wilde, N., Buckellew, M., & Page, H. (2001*).* Software cultures and evolution. *Computer*, 34(9), 24-28.

Sevcik, C. (2017*).* Caveat on the Boltzmann distribution function use in biology. *Progress in Biophysics and Molecular Biology*, 127, 33-42.

Stradner, A., Foffi, G., Dorsaz, N., Thurston, G., & Schurtenberger, P. (2007*).* New insight into cataract formation: Enhanced stability through mutual attraction. *Physical Review Letters*, 99(19), 198103.

Torrie, G. M., & Valleau, J. P. (1977). Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *Journal of Computational Physics,* 23(2), 187 − 199.

Yang, M., & MacKerell Jr, A. D. (2015). Conformational sampling of oligosaccharides using Hamiltonian replica exchange with two-dimensional dihedral biasing potentials and the weighted histogram analysis method (wham). *Journal of Chemical Theory and Computation,* 11(2), 788 − 799.

Yatani, K., Chung, E., Jensen, C., & Truong, K. N. (2009). Understanding how and why open source contributors use diagrams in the development of Ubuntu. *In Proceedings of the Sigchi Conference on Human Factors in Computing Systems* (995-1004).