Evaluating the Reliability of MM-PB/GB-SA Method for the Protein-Ligand Binding Free Energies Using Penicillopepsin-Inhibitor ligands

Twana Salih*
Department of Pharmacognosy & Pharmaceutical Chemistry, College of Pharmacy, University of Sulaimani, Iraq.

Article Info:

Received Jun 2022
Accepted Oct 2022
Corresponding Author email:
twana.salih@univsul.edu.iq
orcid: https://orcid.org/0000-0002-7684-8601

DOI:

Abstract:

An accurate prediction of the ligand-receptor binding free energies (ΔG) is a critical step in the early stages of rational drug design. The Molecular Mechanics-Generalized Born Surface Area (MM-GBSA) method is a popular approach to estimate ΔG. However, correlations between the predicted and the experimental ΔG are variable. The goal of this study is to investigate various approaches to optimize accuracy of the MM-GBSA method. A molecular dynamic (MD) simulations protocol was applied using penicillopepsin receptor against its inhibitor ligands, repeated 50 times for each complex system. After that, ΔG of the five inhibitors were predicted using MM-GBSA method. Moreover, a diverse ΔG values were calculated from the replicate MD simulations of each system. The results were showed correlations not only between the predicted and the experimental binding affinities of the systems but also between the predicted values and root-mean-square deviation. In addition, statistical analysis was evaluated the sample size.

Key words: Predicted binding free energy, Molecular Mechanics-Generalized Born Surface Area method, Molecular dynamic simulations, Correlation coefficient, root-mean-square deviation.

تقييم موثوقية طريقة الميكانيكا الجزيئية – ولادة مساحة سطحية معممة للطاقات الحرة الملزمة للبروتين باستخدام روابط مثبطات بنسلوببسين

توانا صالح*
قسم العقاقير والكيمياء الصيدلانية، كلية الصيدلة، جامعة السليمانية، العراق

الخلاصة:

يعد التقدير الدقيق للطاقات الحرة المرتبطة بين مستقبلات و الربيطة (ΔG) خطوة حاسمة في المراحل الأولى من تصميم الدواء. تعد طريقة الميكانيكا الجزيئية - ولادة مساحة سطحية معممة (MM-GBSA) طريقة شائعة لتقدير ΔG. ومع ذلك ، فإن الارتباطات بين المتوقع و ΔG التجريبية متغيرة. الهدف من هذه الدراسة هو التحقيق في الأساليب المختلفة لتحسين دقة طريقة MM-GBSA. تم تطبيق بروتوكول المحاكاة الديناميكي الجزيئي (MD) باستخدام مستقبلات البنسيلوببسين ضد الروابط المثبطة ، وتكررت 50 مرة لكل نظام معقد. بعد ذلك ، تم التقدير للطاقات الحرة المرتبطة لجزيئات المثبطات الخمسة باستخدام طريقة MM-GBSA. علاوة على ذلك ، تم حساب قيم ΔG متنوعة من محاكاة MD المكررة لكل نظام. أظهرت النتائج ارتباطات ليس فقط بين تقاربات الربط المتوقعة والتجريبية للأنظمة ولكن أيضًا بين القيم المتوقعة وانحراف الجذر التربيعي. بالإضافة إلى ذلك ، تم تقييم التحليل الإحصائي لحجم العينة.

الكلمات المفتاحية: الطاقة الحرة الملزمة المتوقعة ، طريقة الميكانيكا الجزيئية - ولادة مساحة سطحية معممة ، المحاكاة الديناميكية الجزيئية ، معامل الارتباط ، انحراف الجذر التربيعي.

Introduction

Calculating free energies is a vital approach in various subjects of computational drug discovery, for instance, the processes of drug design through predicting binding affinities of the lead molecules with a specific receptor [(1)]. Several computational methods are existed to calculate binding free energies, starting from rapid, but not very accurate approaches, such as Molecular Mechanics-Poisson Boltzmann Surface Area/Generalized Born Surface Area (MM-PBSA/GBSA) calculations [(2)] and linear interaction energy analysis (LIE) [(3)], to the significantly accurate but slow methods like umbrella sampling (US) [4], thermodynamic integration (TI) [5], and free-energy perturbation (FEP) [6]. The above methods analyze trajectories provided through Monte Carlo (MC) or Molecular Dynamics (MD) simulations to calculate binding free energies.

The MM-PBSA/GBSA approach is a popular method to calculate the free energy of the solvated molecules either from different conformations or between interacted and non-interacted states. This technique is widely applied during the drug discovery projects to predict relative values of free energy, as it possesses an acceptable accuracy compared to the experimental results, is computationally affordable, and could apply to various systems [7-9]. Nevertheless, several factors can dramatically affect the accuracy of this approach, such as selecting forcefields and partial charges, inadequate sampling, and trapping the complex systems in local minima for a long-time during MD simulation [8, 10].

This equation is used to calculate the receptor-ligand (ΔG_{bind}) binding free energy from complex energy (ΔG_{com}), energy of receptor (ΔG_{rec}), and energy of ligand (ΔG_{lig}) using MM-PBSA/GBSA approach:

$$\Delta G_{bind}=\Delta G_{com}-\Delta G_{rec}-\Delta G_{lig} \quad (1)$$

The free energy of each ΔG from the eq. 1 is predicted by eq. 2

$$\Delta G=\Delta H-T\Delta S \quad (2)$$

ΔH denotes enthalpy change and -TΔS is the conformational entropy when T denotes the absolute temperature and S indicates the molecule entropy. Eq. 3 is explaining ΔH.

$$\Delta H=\Delta E_{MM}+\Delta G_{solv} \quad (3)$$

When ΔE_{MM} signifies the molecular mechanic's energy of the molecule and ΔG_{solv} is the free energy of desolvation. Both eq. 4 and eq. 5 are clarified ΔE_{MM} and ΔG_{solv} separately.

$$\Delta E_{MM} = \Delta E_{internal} + \Delta E_{electrostatic} + \Delta E_{vdw} \quad (4)$$

$\Delta E_{internal}$ is all internal energies (bond, angle, and dihedral energies), $\Delta E_{electrostatic}$ is electrostatic energy of interactions, and ΔE_{vdw} is Van der Waals energy of interactions [8].

$$\Delta G_{solv}=\Delta G_{PB/GB}+\Delta G_{SA} \quad (5)$$

ΔG_{solv} can be defined as a sum of $\Delta G_{PB/GB}$ (polar or electrostatic contribution of solvation energy) and ΔG_{SA} (nonpolar or nonelectrostatic solvation energy). The electrostatic or polar components are calculated through implicit solvation models, such as Generalized Born (GB) or Poisson-Boltzmann (PB) models. Using explicit solvent rather than continuum solvent for the conformers may lead to an increase in the accuracy of the results [(11)]. For calculating long-range electrostatic interactions, a particle-mesh Ewald procedure is applied [12]. Computing energy components during MM-PBSA/GBSA methods are through conformational snapshots generated from MD simulations [13]. Calculations of energy are achieved through the comparison between the average of conformational ensembles and the reference structure (crystal structure). Therefore, a conformational change of the system components decides on the final energy. In this study, entropy was neglected since the

relative predicted binding free energy of the compounds is required, and ignoring the entropy terms could reduce the probability of error and noise [14, 15].
According to the literature, the accuracy of *in silico* results is related to different factors. First, are the characteristics of the protein-ligand interactions and their structural similarities. Ligands with similar chemical structures could be more correlated than the diverse chemical structures. In addition, charged residues significantly affect the resulted binding free energies. Moreover, parameters applying to the description of the molecular system, like selecting the most appropriate force field have a remarkable effect on the results [16, 17]. Finally, multiple independent media or short MD simulations replicas could increase the correlations between the estimated and experimental results [18,19]. The penicillopepsin enzyme is produced by fungus *Penicillium janthinellum* as an extracellular acid protease enzyme [20]. This molecular modeling study aims to predict the binding free energy of the 5 penicillopepsin-inhibitor ligand complex systems and to explore different approaches to optimize the accuracy of the calculated binding affinity values. Computing energy components of the systems is through using the MM-GBSA method, as it's entirely efficient computationally, even more than MM-PBSA [21]. In addition, the MM-GBSA approach owns a better performance compared to MM-PBSA to predict the relative binding free energy [17]. However, theoretically, the MM-PBSA method is supposed to show higher accuracy to estimate the absolute binding free energy [18]. The MM-GBSA approach is based on the MD simulations data, which are used to calculate the binding free energy as an average over the individual snapshots of the protein-ligand complex [22, 23].

Materials and Methods

MM-GBSA approach was achieved to predict $\Delta G_{binding}$-GB values for the 5 Protein-ligand complexes. The starting coordinates of the complexes were derived from the crystal structures of penicillopepsin-ligand inhibitors; PDB codes are 1APT [24], 1APU, 1APV [25], 2WEA, and 2WEC [26]. The resolution of the first 3 complexes was 1.8 Å, 2WEA is 1.25 Å, and the last complex was 1.5 Å. The inhibitory ligands of 1APT, 1APU, 1APV, 2WEA, and 2WEC were PI1(N-(1-ethoxy-1,3-dihydroxynonan-4-yl)-3-methyl-2-[3-methyl-2-(3-methylbutana-mido) butanamido]butanamide),PI2(N-(1-ethoxy-1,3-dihydroxy-6-methylheptan-4-yl)-3-methyl-2-[3-methyl-2-(3-methylbutanamido) butanamido]butanamide)PI3 (2,2-difluoro-3,3-dihydroxy -N,6-dimethyl-4-{3-methyl-2-[3-methyl-2-(3-methyl-butanamido)butanamido] butanamido} heptanamide), PI4 ([3,6-dioxo-5-(propan-2-yl)-4,7-diazatricyclo[8.6.2.0^{13,17}]octadeca-1 (16), 10(18),11, 13(17),14-pentaen-8-yl][(1-methoxy-1-oxo-3-phenylpropan-2-yl) oxy] phosphinic acid), and PI5 ([(1-methoxy-1-oxo-3-phenylpropan-2-yl)oxy]({3-methyl-2-[2-(naphthalen-1-yl)acetamido]butanamido}methyl)phosphi nic acid), respectively (Figure.1)

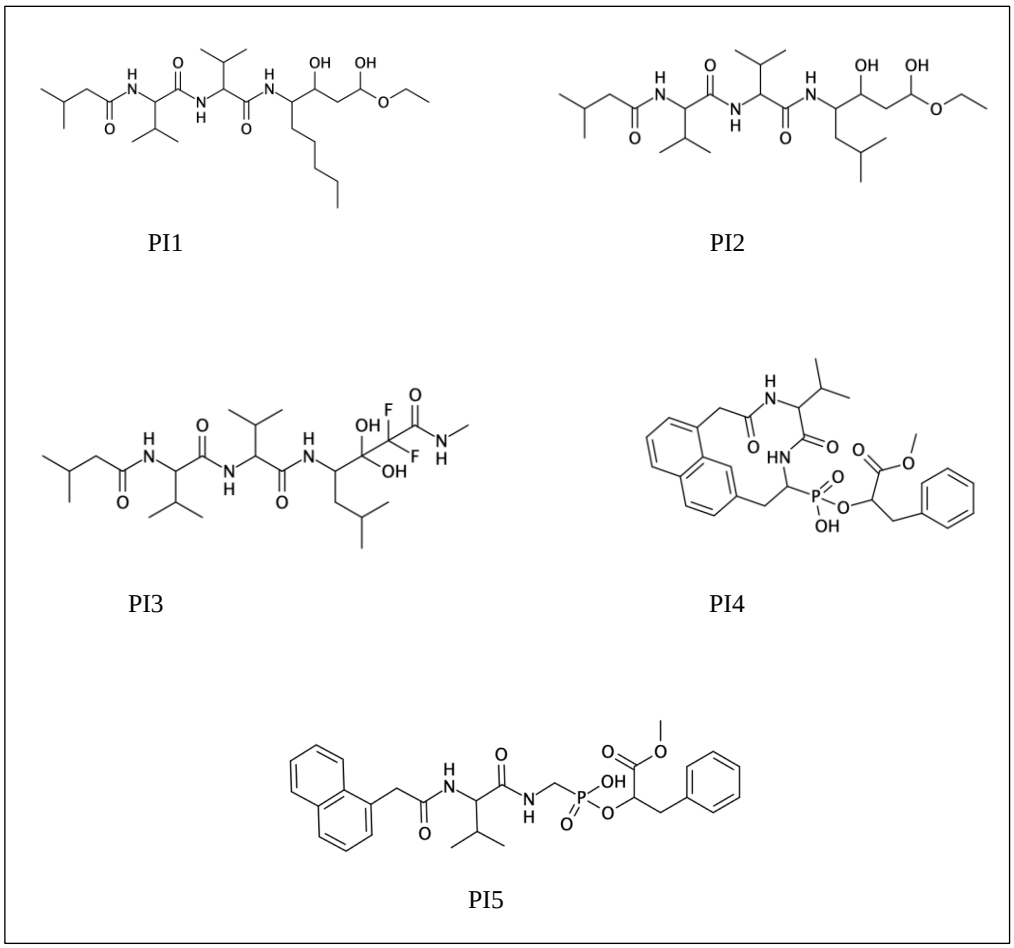

Figure (1): The 2-dimensional structures of the ligand inhibitors of penicillopepsin.

Molecular visualization and manipulation were performed using the Visual Molecular Dynamics (VMD) package [27]. All the MD simulations were executed using AMBER suite version 12 and *tleap* to parameterize the complexes. The force field ff99SB was used in this work [28]. Counterions of Na^+ and Cl^- were added to neutralize the systems, which construct the largest negative or positive Coulombic potential grid around the complexes. Each complex system was immersed in an octahedral box of TIP3P water molecules, which extended 10 Å outside of the complex in all dimensions [29].

MD simulations protocol

Before MD simulations, two minimization steps were applied; restrained and the whole system minimization was achieved to relax the complexes. The restrained minimization procedure was 2000 cycles for both the steepest descent (500 cycles) and the conjugate gradient minimization (1500 cycles). The restrained atoms were the residues of the complex (Protein-Peptide), while all the water and counterions were free of movement. On the other hand, the whole system minimization was 1000 cycles of steepest descent and 2000 cycles of conjugate gradient minimization (3000 cycles).

The next step was producing the MD simulations through applying two consequent steps (restrained MD simulations and the whole system MD simulations). Simulations were run by the pmemd module in Amber 12. After that, the SHAKE algorithm was applied in both MD simulations to restrain bond lengths including hydrogen atoms for MD trajectories [30]. A constant volume periodic boundary and a 10 Å nonbonded cutoff were applied. The simulations time step

was set on 2 femtoseconds (fs). The temperature regulation was through Langevin dynamics applying 1.0 ps^{-1} as a frequency of collision [31]. Then, a constant volume periodic boundary was used to produce the MD simulations. In addition, the system was heated steadily from 0 to 300 Kelvin (K) over 200 picoseconds (ps). Moreover, the Boltzmann distribution was generated random initial velocities.

The last MD simulations were equilibrating the whole system. During this running, the temperature was kept at 300 K. A constant pressure periodic boundary of 1 atm was used. To keep constant pressure dynamics with 2 ps pressure relaxation time, Isotropic position scaling was applied. Also, velocities, coordinates, and box information were read from a formatted coordinate file. Lastly, every 10000 steps or every 20 ps, the snapshots were investigated. All the simulations were achieved using ARCHER (http://www.archer.ac.uk).

.

MM-PB/GB-SA calculations

MM-PB/GB-SA calculations were achieved by applying python scripts (MMPBSA.py)[32]. The MM-PBSA/GBSA energies were calculated from a production run of 10 nanoseconds (ns) at 300 °K, snapshots were taken every 20 picoseconds. The long-range electrostatic interactions was treated by using Particle Mesh Ewald (PME) [33]. Infinite cut-off was executed to removing the water molecules. For running the MD simulations in explicit water, each complex system were created four topology files (non-solvated complex, solvated complex, ligand, and the receptor file) [34]. Subsequently, the generalized Born method (igb =5) and 0.1 M of the salt concentration were executed to strip water and counterions using the MM-GBSA method [35, 36]. Analyzing data and graphs were implemented applying GraphPad Prism Version 9.3.1 (GraphPad Software Inc., San Diego, CA; www.graphpad.com).

Results and Discussion

Regarding the molecular modeling study, single MD simulations for each of the protein-ligand complex were not sufficient, as after repeating MD simulations of the same complex, yielding results that could differ significantly in the calculated MM-GBSA binding affinity. All the replicates originated from the same reference structure and protocols. Although, the predicted energy of each replicate showed a significant difference (Figure 2). It is fairly clear that two main reasons that single MD simulations generate an incorrect result are due to the insufficient sampling of the conformational space and trapping the complex systems in local minima [37]. This suggests that multiple MD simulations could overwhelm such an obstacle. We present here the results of an in-depth analysis of the issue of convergence in such a system. Therefore, data from 50 replicate, 10 ns simulations of the penicillopepsin-ligand inhibitor systems were exposed to statistical analysis. The following table (Table 1) shows MM-GBSA calculations for 5 penicillopepsin-ligand complexes and for each complex, a calculation of 50 replicates was undertaken.

Table 1. Binding free energy values of the penicillopepsin-ligand inhibitors; the experimental results [38]; the predicted average, minimum, and maximum results using the MM/GBSA method.

PDB	Penicillopepsin inhibitor	Experimental ΔG (kcal/mol)	Calculated ΔG average of 50 rep.	Minimum ΔG value (kcal/mol)	Maximum ΔG value (kcal/mol)
1APT	Pepstatin analogue (PI1)	-12.83	-42.47	-14.39	-80.22
1APU	Pepstatin analogues (PI2)	-10.51	-26.35	-9.59	-53.43
1APV	iva-val-val(H)Dfo-n-methylamide (PI3)	-12.27	-28.18	-16.28	-57.97
2WEA	PP6 (PI4)	-8.37	-12.79	-4.41	-25.85
2WEC	PP5 (PI5)	-6.8	-17.32	-7.97	-30.36

Statistical Analysis of the Replicate Values – how many do we need to run?

The proper number of independent simulations should be run to improve the accuracy [39]. Deciding on the sample size is a big concern, as too large samples are expensive and time-consuming, while too small samples could not be accurate. Checking the sample distributions was essential to ensure the expectations of a parametric test are met before use. The histogram of binding energies predicted from the 50 reproduce MD simulations of the receptor-inhibitor systems was incompatible with a standard bell shape (Figure2); nevertheless, it could be acceptable if they draw from a fundamental normal distribution property. Various normality tests were implemented, such as the Shapiro-Wilk test, the Kolmogorov-Smirnov, D'Agostino & Pearson omnibus test. This was one of the indicators to decide on the number of replicates. Additionally, it was valid to apply some standard statistical approaches to expect the number of replicates that need to be run in the chosen confidence limits. An approach to identify the sample size of the complex systems was through equation 6.

$$\mathbf{n=[Z_{\alpha/2}\delta/E]^2} \quad (6)$$

n is a minimum replicate number, **E** is a margin of error, $\mathbf{Z_{\alpha/2}}$ is a critical value, which is 1.96 in the 95% confidence calculations, an **δ** is a standard deviation of population (SD) [40].

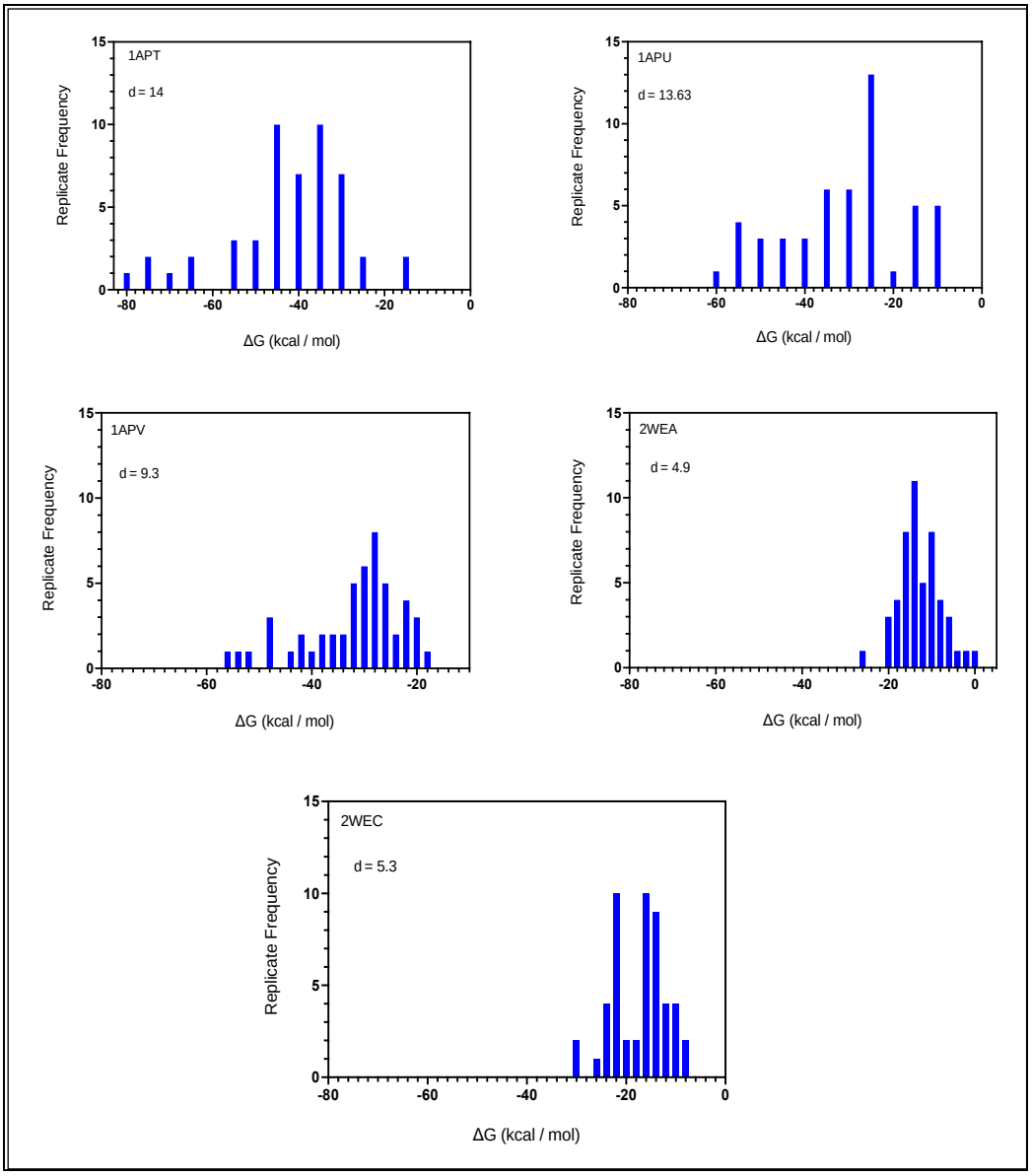

Figure (2): The frequency of ΔG distribution of the penicillopepsin-ligand replicates.

As elucidated in Table 2, the replicate numbers were operated by two factors. First, the various **E** could decide on the **N** value. For example, the minimum sample size of the 1APT complex system was around 753 replicates if 1 kcal/mol was selected as **E**, 188 replicates if 2 kcal/mol was selected, 84 replicates when 3 kcal/mol was chosen, and it was 47 replicates for 4 kcal/mol. The second factor was dissimilar complex systems; the **N** values were different according to the complex systems because each of the ligand-receptor complex system possessed a unique SD, which owns a direct impact on the **N** value. The values of SD were related to the scattering of ΔG results of the complex system replicates. Thus, increasing SD led to an increase in the number of replicates to be running and vice versa. For instance, the maximum number of replicates was 753 for 1APT when **E** was 1 kcal.mol^{-1}. Therefore, when all the systems were set to 1 kcal/mol as **E**, we should run around 750 replicates for all of them to be consistent, which was significantly expensive computationally. Consequently, 4 kcal/mol could be applied as **E**, as the highest number of replicates was the 1APT complex (47 replicates); accordingly, the average of 50

replicates could be acceptable for all the protein-ligand complex systems to organize them consistently.

Table (2): The calculated sample size and SD of the 5 complex systems with various margins of error.

PDB	SD	N value (E=1)	N value (E=2)	N value (E=3)	N value (E=4)
1APT	14	753	188	84	47
1APU	13.6	714	178	79	45
1APV	9.3	332	83	37	21
2WEA	4.9	92	23	10	6
2WEC	5.3	108	27	12	7

The next investigation was the analysis of the relative correlation between the predicted ΔG values with the experimental results. The maximum, minimum, and average ΔG results of the 50 copies in each ligand-receptor system were compared to the experimental results. Coefficient of determination (R^2) was used to measure the strength of correlation. The maximum estimated ΔG of each complex system against the experimental values was shown the highest correlation (R^2 = 0.84). Moreover, the graph of the correlation between the average predicted ΔG of each system versus the experimental results exhibited R^2 = 0.75. On the contrary, the lowest R^2 value (0.71) was exposed from the minimum calculated ΔG values of the replicates. However, all the results indicated a strong correlation between the predicted and the experimental results [41] (Figure 3).

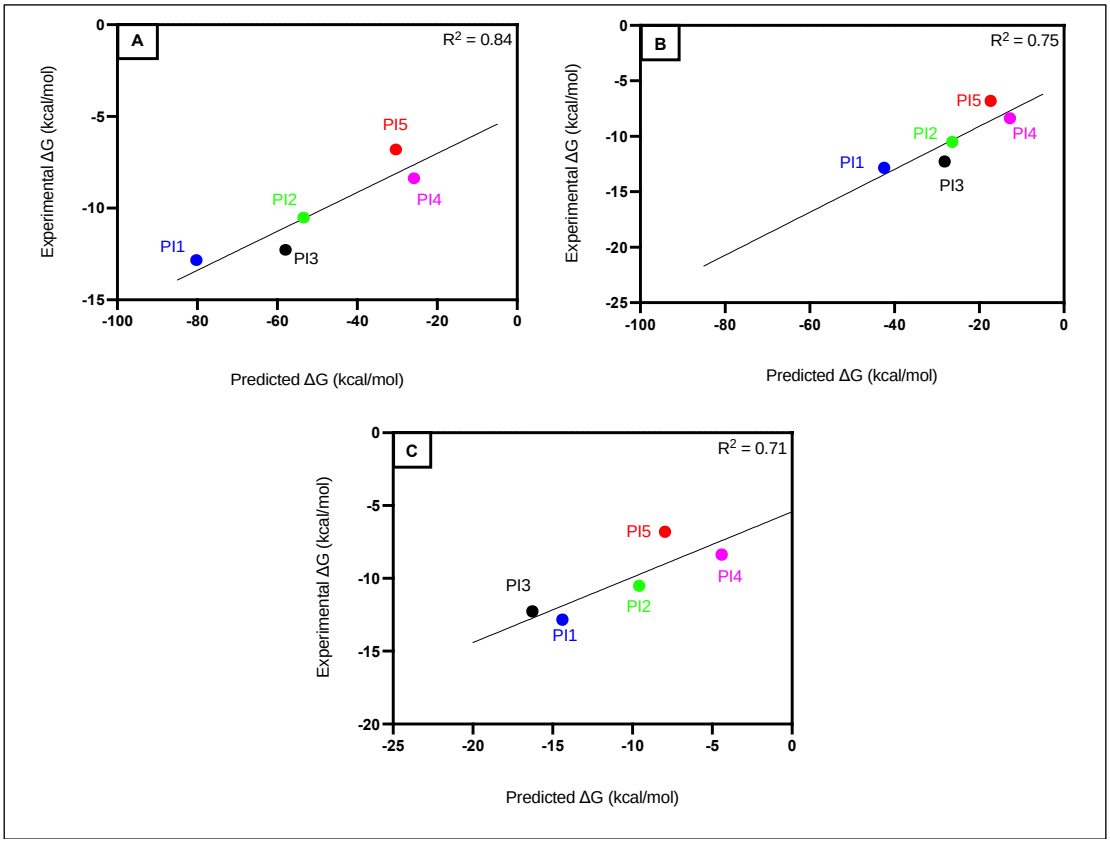

Figure (3): The correlation coefficient between the experimental results and binding free energy values was calculated through the MM-GBSA method for each complex system. A) Correlation between the maximum value of the 50 replicates and the experimental results. B) Correlation between the experimental values and the average of 50 replicates for each complex system. C) Correlation between the experimental ΔG values and the minimum ΔG value of each complex system from the 50 replicate values.

Correlation between binding free energy and RMSD values

Root-mean-square deviation (RMSD) is the system of measurement applies in drug design projects to realize the conformation of diverse models or to compare the divergence of the estimated conformations from the starting structure conformation. According to RMSD technique, the best value is the lowest value, which is the closest conformation from the native structure and the least deviation [42]. This tool is used during MD analysis to observe the equilibration of the systems, check the variations between structural conformations, estimate the amount of sampling of various conformations, and predict the quality of MD simulation [43, 44]. In this study, RMSD was calculated for the 50 replicates of the 1APT, 1APU, and 1APV complex systems (Figure 4). The computed RMSD covered all the snapshots, which calculated over the period of the MD simulations based on the crystal structure. Then, RMSD plots displayed changes in the structural conformations gradually compared with the first snapshot.

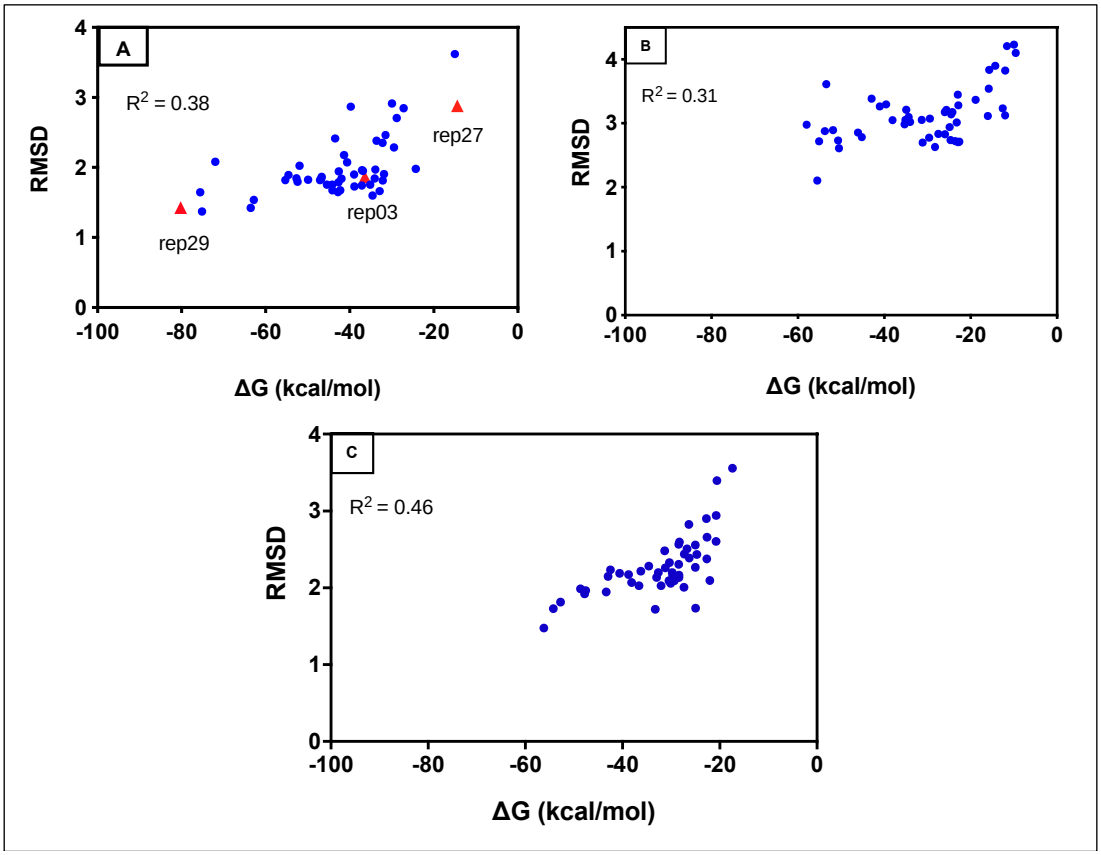

Figure (4): The correlation coefficient (R2) between the predicted free energy and RMSD mean value for each replicate of the 1APT, 1APU, and 1APV complex systems. A) Correlation between the 1APT system average RMSD and ΔG replicate values. B) Correlation between the predicted binding energy of the 1APU system and the average RMSD. C) Correlation between the ΔG value and RMSD of the 1APV complex system.

According to Hou et al. 2011 [45], the success of estimated binding free energy for the various systems is possible to measure through RMSD, as decreasing the RMSD value may lead to an increase ΔG value. This hypothesis means the closest conformation from the reference structure is the closest energy from the global minima. To investigate this hypothesis, the correlation between the predicted binding energies and the average of the RMSD value for each replicate of 1APT, 1APU, and 1APV systems was calculated. As recognized in Figure 4, the R^2 value for the

1APT, 1APU, and 1APV complex systems were 0.38, 0.31, and 0.46, respectively, which showed a low effect size (weak correlation) for the 3 tested systems. However, as shown in Figure 3, the maximum predicted binding free energy of each system was denoted the strongest correlation with the experimental values (R^2 = 0.84) and such predicted ΔG results possessed the lowest RMSD value (Figure 4).

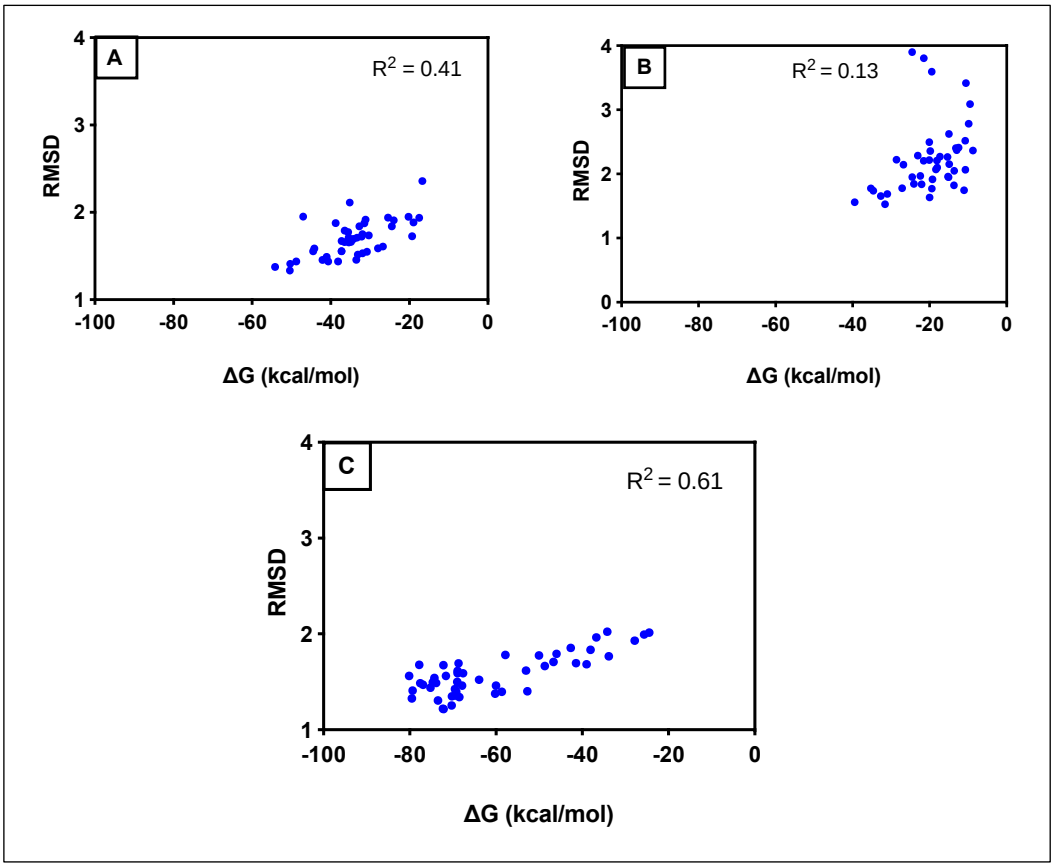

Figure (5): The correlation coefficient between the calculated ΔG values and the RMSD mean value of the replicate 3, 27, and 29 of the 1APT complex system. A) replicate 29; B) replicate 3; C) replicate 27

To further investigate the predicted binding free energies, 3 values were taken from the 50 replicates of 1APT results, which were the highest (replicate 29), middle (replicate 3), and the lowest ΔG values (replicate 27). After that, 50 MD simulations were performed for each of such replicates. The results displayed that replicate 29 possessed a weak correlation (R2 = 0.41), while replicate 3 was considered non-correlation (R2 = 0.13). Nevertheless, replicated 27 indicated a moderate correlation (R2 = 0.61) (Figure 5).

Conclusion

In the current study, the capability and accuracy of the MMGBSA methodology were investigated by predicting binding affinities between penicillopepsin and the ligand inhibitors PI1, PI2, PI3, PI4, and PI5. Each system was individually tested, and 50 replicas of the MD simulation were produced for each complex system. Then, MM-GBSA was applied to estimate the ΔG of individual replica. After calculating ΔG of all complex systems replicas, statistical analysis could identify the number of replicas of each complex system, based on the SD and margin of error. The next outcome was shown that

the correlation coefficient is strong between the predicted ΔG values of the highest, lowest, and average of 50 replicas and the experimental results, specifically, the most profound correlation between the predicted and the experimental values were denoted when the maximum predicted ΔG results of each complex system selected (R2 = 0.84). Finally, it is worth stating that the highest calculated ΔG results possessed the lowest RMSD, accordingly, the further the individual simulation drifts away from the starting structure; the poorer is the calculated ΔG.

Acknowledgments

This work was financially supported by the KRG and the Ministry of the Higher Education/University of Sulaimani. I gratefully acknowledge support from Professor Charles Laughton and the Centre for Biomolecular Sciences, the University of Nottingham, UK.

Conflicts of Interest

The author declares no conflict of interest.

References

1- Huang K, Luo S, Cong Y, Zhong S, Zhang JZH, Duan L. An accurate free energy estimator: based on MM/PBSA combined with interaction entropy for protein-ligand binding affinity. Nanoscale. 2020;12(19) :10737-10750.
2- Weng G, Wang E, Chen F, Sun H, Wang Z, Hou T. Assessing the performance of MM/PBSA and MM/GBSA methods. 9. Prediction reliability of binding affinities and binding poses for protein-peptide complexes. Physical chemistry chemical physics: PCCP. 2019;21(19):10135-10145.
3- He X, Man VH, Ji B, Xie XQ, Wang J. Calculate protein-ligand binding affinities with the extended linear interaction energy method: application on the Cathepsin S set in the D3R Grand Challenge 3. J Comput Aided Mol Des. 2019;33(1):105-117.
4- Ngo ST. Estimating the ligand-binding affinity via λ-dependent umbrella sampling simulations. J Comput Chem. 2021;42(2):117-123.
5- Zou J, Tian C, Simmerling C. Blinded prediction of protein-ligand binding affinity using Amber thermodynamic integration for the 2018 D3R grand challenge 4. J Comput Aided Mol Des. 2019;33(12):1021-1029.
6- Wang L, Chambers J, Abel R. Protein–ligand binding free energy calculations with FEP+. Biomolecular Simulations: Springer; 2019. p. 201-232.
7- Genheden S, Ryde U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. Expert opinion on drug discovery. 2015;10(5):449-461.
8- Wang E, Sun H, Wang J, Wang Z, Liu H, Zhang JZH, et al. End-Point Binding Free Energy Calculation with MM/PBSA and MM/GBSA: Strategies and Applications in Drug Design. Chemical reviews. 2019;119(16):9478-9508.
9- Kollman PA, Massova I, Reyes C, Kuhn B, Huo S, Chong L, et al. Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models. Accounts of Chemical Research. 2000; 33:889-897.
10- Sun H, Li Y, Tian S, Xu L, Hou T. Assessing the performance of MM/PBSA and MM/GBSA methods. 4. Accuracies of MM/PBSA and MM/GBSA methodologies evaluated by various simulation protocols using PDBbind data set. Physical chemistry chemical physics: PCCP. 2014;16(31):16719-16729.
11- Wang C, Greene D, Xiao L, Qi R, Luo R. Recent Developments and Applications of the MMPBSA

Method. Front Mol Biosci. 2017; 4:87.
12- Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A smooth particle mesh Ewald method. The Journal of Chemical Physics. 1995;103(19):8577.
13- Wu X, Xu LY, Li EM, Dong G. Application of molecular dynamics simulation in biomedicine. Chemical biology & drug design. 2022.
14- Geng C, Xue LC, Roel-Touris J, Bonvin AM. Finding the ΔΔG spot: Are predictors of binding affinity changes upon mutations in protein–protein interactions ready for it? Wiley Interdisciplinary Reviews: Computational Molecular Science. 2019;9(5): e1410.
15- Mishra SK, Koča J. Assessing the performance of MM/PBSA, MM/GBSA, and QM–MM/GBSA approaches on protein/carbohydrate complexes: Effect of implicit solvent models, QM methods, and entropic contributions. The Journal of Physical Chemistry B. 2018;122(34):8113-8121.
16- Poli G, Granchi C, Rizzolio F, Tuccinardi T. Application of MM-PBSA Methods in Virtual Screening. Molecules. 2020;25(8).
17- Salo-Ahen OMH, Alanko I, Bhadane R, Bonvin AMJJ, Honorato RV, Hossain S, et al. Molecular Dynamics Simulations in Drug Discovery and Pharmaceutical Development. Processes. 2020;9(1).
18- Tuccinardi T. What is the current value of MM/PBSA and MM/GBSA methods in drug discovery? Expert opinion on drug discovery. 2021;16(11):1233-1237.
19- Wan S, Bhati AP, Zasada SJ, Coveney PV. Rapid, accurate, precise and reproducible ligand-protein binding free energy prediction. Interface Focus. 2020;10(6):20200007.
20- Hofmann T. [35] Penicillopepsin. Methods in Enzymology. 45: Academic Press; 1976. p. 434-452. DOI: https://doi.org/10.1016/S0076-6879(76)45038-3.
21- Chen F, Sun H, Wang J, Zhu F, Liu H, Wang Z, et al. Assessing the performance of MM/PBSA and MM/GBSA methods. 8. Predicting binding free energies and poses of protein-RNA complexes. RNA. 2018;24(9):1183-1194.
22- Patel H, Kukol A. Integrating molecular modelling methods to advance influenza A virus drug discovery. Drug discovery today. 2021;26(2):503-510.
23- Liu X, Shi D, Zhou S, Liu H, Liu H, Yao X. Molecular dynamics simulations and novel drug discovery. Expert opinion on drug discovery. 2018;13(1):23-37.
24- James M, Sielecki A, Moult J, Hruby V, Rich D, editors. Crystallographic analysis of a pepstatin analogue binding to the aspartyl proteinase penicillopepsin at 1.8 Angstroms resolution. Proceedings of the of the Eighth American Peptide Symposium; 1983.
25- James MN, Sielecki AR, Hayakawa K, Gelb MH. Crystallographic analysis of transition state mimics binding to penicillopepsin: difluorostatine-and difluorostatone-containing peptides. Biochemistry. 1992;31(15):3872-3886.
26- Ding J, Fraser ME, Meyer JH, Bartlett PA, James MN. Macrocyclic inhibitors of penicillopepsin. 2. X-ray crystallographic analyses of penicillopepsin complexed with a P3− P1 macrocyclic peptidyl inhibitor and with its two acyclic analogues. Journal of the American Chemical Society. 1998;120(19):4610-4621.
27- Humphrey W, Dalke A, Schulten K. VMD: Visual molecular dynamics. Journal of Molecular Graphics & Modelling. 1996;14(1):33-38.
28- Wickstrom L, Okur A, Simmerling C. Evaluating the Performance of the

ff99SB Force Field Based on NMR Scalar Coupling Data. Biophys J. 2009;97(3):853-856.
29- Aqvist J. Ion-Water Interaction Potentials Derived from Free Energy Perturbation Simulations. J Phys Chem. 1990; 94:8021-8024.
30- Ryckaert JP, Ciccotti G, Berendsen HJC. NUMERICAL-INTEGRATION OF CARTESIAN EQUATIONS OF MOTION OF A SYSTEM WITH CONSTRAINTS - MOLECULAR-DYNAMICS OF N-ALKANES. J Comput Phys. 1977;23(3):327-341.
31- Wu X, Brooks BR. Self-guided Langevin dynamics simulation method. Chemical Physics Letters. 2003;381(3-4):512-518.
32- Miller BR, McGee TD, Swails JM, Homeyer N, Gohlke H, Roitberg AE. MMPBSA.py: An Efficient Program for End-State Free Energy Calculations. Journal of Chemical Theory and Computation. 2012;8(9):3314-3321.
33- Darden T, York D, Pedersen L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. The Journal of Chemical Physics. 1993;98(12):10089.
34- Case DA, Babin V, Berryman JT, Betz RM, Cai Q, Cerutti DS, et al. AMBER 14. University of California, San Francisco. 2014.
35- Tian S, Zeng J, Liu X, Chen J, Zhang JZH, Zhu T. Understanding the selectivity of inhibitors toward PI4KIIIalpha and PI4KIIIbeta based molecular modeling. Physical chemistry chemical physics: PCCP. 2019;21(39):22103-22112.
36- Xu Z, Peng C, Shi Y, Zhu Z, Mu K, Wang X, et al. Nelfinavir was predicted to be a potential inhibitor of 2019-nCov main protease by an integrative approach combining homology modelling, molecular docking and binding free energy calculation. BioRxiv. 2020.
37- Knapp B, Ospina L, Deane CM. Avoiding False Positive Conclusions in Molecular Simulation: The Importance of Replicas. J Chem Theory Comput. 2018;14(12):6127-6138.
38- Hou TJ, Wang JM, Li YY, Wang W. Assessing the Performance of the MM/PBSA and MM/GBSA Methods. 1. The Accuracy of Binding Free Energy Calculations Based on Molecular Dynamics Simulations. Journal of Chemical Information and Modeling. 2011;51(1):69-82.
39- Salih T. Inhibiting protein–protein interactions in telomeres as an approach to cancer chemotherapy: University of Nottingham; 2016.
40- Peat JK, Barton B, ebrary I. Medical statistics: a guide to data analysis and critical appraisal. 1st ed. Malden, Mass.: Blackwell Pub.; 2005. xii, 324 p. p.
41- Moore DS, Notz W, Fligner MA. The Basic Practice of Statistics. 2013. W.H. Freeman and Company; [15-24]. Available from: https://books.google.iq/books?id=aw61ygAACAAJ.
42- Ramírez D, Caballero J. Is It Reliable to Take the Molecular Docking Top Scoring Position as the Best Solution without Considering Available Structural Data? Molecules (Basel, Switzerland). 2018;23(5):1038.
43- E Lohning A, M Levonis S, Williams-Noonan B, S Schweiker S. A practical guide to molecular docking and homology modelling for medicinal chemists. Current topics in medicinal chemistry. 2017;17(18):2023-2040.
44- Perricone U, Gulotta MR, Lombino J, Parrino B, Cascioferro S, Diana P, et al. An overview of recent molecular dynamics applications as medicinal chemistry tools for the undruggable site challenge. MedChemComm. 2018;9(6):920-936.
45- Hou TJ, Wang JM, Li YY, Wang W. Assessing the Performance of the

Molecular Mechanics/Poisson Boltzmann Surface Area and Molecular Mechanics/Generalized Born Surface Area Methods. II. The Accuracy of Ranking Poses Generated from Docking. J Comput Chem. 2011;32(5):866-877.