Sensitivity in predicted relative binding free energies from incremental ligand changes within a model binding site

Nathan Lim*

Department of Pharmaceutical Sciences

E-mail: limn1@uci.edu

Abstract

Despite innovations in sampling techniques for molecular dynamics (MD), reliable prediction of protein-ligand binding free energies from MD remains a challenging problem, even in well studied model binding sites like the apolar cavity of T4 Lysozyme L99A.[1] In this study, we model recent experimental results that show the progressive opening of the binding pocket in response to a series of homologous ligands.[2] Even while using enhanced sampling techniques, we demonstrate that the predicted relative binding free energies (RBFE) are still highly sensitive to the initial protein conformational state. Particularly, we highlight the importance of sufficient sampling of protein conformational changes and possible techniques for addressing the issue.

*To whom correspondence should be addressed

1 Introduction

Medicinal chemistry programs typically focus on changes in ligand binding affinity from incremental changes to the ligand. Focus on how the protein adapts to the changes in the ligand is generally neglected. T4 L99A is well studied experimentally and computationally. It is frequently used as a model binding site in free energy prediction studies. In this study, 8 congeneric ligands were investigated, where addition of a single methyl group was used to lengthen the ligand. Through determination of protein-ligand bound x-ray crystal structures it was revealed T4 Lysozyme adopts 3 discrete conformations in response the series of growing ligands. Consideration of the protein adaptations into discrete conformations may be an important aspect in inhibitor design.

2 Results

Dependence on protein starting conformation

Using the default FEP/REST methodology,[3] we find calculated free energies significantly depend on the protein starting conformation, especially for large perturbations (i.e. opening the cavity from the closed state). To illustrate this, we begin our molecular dynamics simulations from either the protein closed or open conformation and perform alchemical transformations to ligands that occupy the opposite protein conformational state. In this study, root-mean-square-deviation (RMSD) of the backbone atoms in the F-helix is used to determine the conformational state of the protein over the course of the trajectory. Here, we demonstrate the standard 5ns simulation time and default REST region selection are insufficient for adequate sampling of the motion in the F-helix and does not eliminate the dependence on the initial protein state.

An examination of the largest alchemical transformation, benzene to n-hexylbenzene, highlights the sampling problems faced when using the standard FEP/REST protocol.

From experimental data of ligand occupancies (Table 1), we expect in our simulations of n-hexylbenzene to see the protein primarily in the open state over the closed state. Instead, we find the protein remains trapped in its initial conformational state whether we start from closed (Fig 1) or open (Fig 2) over the course of the 5ns trajectory. From the protein closed trajectory, the protein only begins to enter the intermediate state around 3ns but never enters the open conformation. As the protein strains to accommodate n-hexylbenzene and enter its preferred open state, protein-ligand strain results and is reflected in the positivity of $\Delta\Delta G_{calc}$(+4.13 kcal/mol) On the other hand, in the protein open trajectory, the protein already begins in its preferred state for n-hexylbenzene and stays only in this open state. Expectedly, the $\Delta\Delta G_{calc}$ comes out negative(-0.61 kcal/mol) as there is no occurrence of large protein-ligand strain in order to open the cavity. Ultimately, we arrive at two very different relative free energies, where the discrepancy is a whopping +4.74 kcal/mol, for the same transformation of benzene to n-hexylbenzene[1]. Collectively, when we view the discrepancy of all calculations involving closed-open transformations we find the root-mean-square-error (RMSE) to be +4 kcal/mol (Table 6). Clearly, despite the use of FEP/REST, we are unable to adequately sample all the relevant states within the standard 5ns time frame, resulting in such large differences in $\Delta\Delta G_{calc}$.

In the case of more moderate alchemical transformations, such as cases that involve the set of closed ligands (i.e. benzene to n-propylbenzene) to intermediate ligands (i.e. sec-/n-butylbenzene), we find that the calculated free energies still have a dependence on the initial protein conformation using the default protocol, albeit much smaller. For the set of transformations to the intermediate state, the discrepancy RMSE in $\Delta\Delta G_{calc}$ for protein closed versus open simulations is +0.60 kcal/mol (Table 3). Although, when we compare $\Delta\Delta G_{calc}$ against $\Delta\Delta G_{exp}$ for transformations involving n-butylbenzene, we find that simulations starting from the protein closed conformation are further from converging

[1]It should be noted that the binding affinities of n-pentyl/n-hexylbenzene to T4-L99A are not known and were inaccessible in experimental studies due to solubility limits.[2] For cases involving these ligands, we only focus on the convergence of the calculated free energies between simulations starting from protein closed or open.

to $\Delta\Delta G_{exp}$ than when starting from the protein open conformation. From the protein closed trajectories, the RMSE from $\Delta\Delta G_{exp}$ is +1.40 kcal/mol (Table 9) while the protein open trajectories have a RMSE of +0.70 kcal/mol (Table 10). Based from experimental evidence (Table 1), we should expect to see some sampling (30%) of the open conformation for the n-butylbenzene ligand. Evidently, the $\Delta\Delta G_{calc}$ dependence on protein starting conformation is again from remaining trapped throughout the trajectory, resulting in inadequate sampling in the protein closed simulations (Fig. 3) versus the protein open simulations (Fig. 4). Despite performing much smaller alchemical transformations, we still encounter sampling problems, evident when making comparisons of the calculated free energies to experimental.

REST improvements (pREST)

Primarily, we encounter major sampling problems when we begin our simulations from the protein closed state and attempt to open the helix. In order to facilitate protein motion, we included 3 key residues spanning the F-helix region into the REST region, which we will denote simulations using this with pREST. By expanding the REST region, we are able to drive the F-helix out its initial state trap faster by locally heating up key regions and thereby reduce our problem in inadequate sampling.

To demonstrate the REST improvement over the default protocol, we return to the case of benzene to n-hexylbenzene. Here, we show the facilitation of the helix motion by first referring to Fig 1 which shows that there is no sampling of the open state for the default protocol. On the other hand with the pREST, we see a few open state points around 3ns and even a single open point before closing again after our initial timestep (Fig 5. Alternatively, we can further illustrate the enhancement of protein transitions by viewing all replicas collectively. In Fig 6 and Fig 7, we perform the same RMSD analysis but instead represent each time point as a colored bar and no longer plot the raw RMSD. Visually, it is easy to see that there are far less transitions in default simulations (Fig 6) as opposed to pREST simulations (Fig 7).

Collectively, for all our closed/open and closed/intermediate transformations using pREST, we find only some minor improvements in the discrepancy RMSE and even cases where we perform worse. For closed/open transformations (Table 7), the discrepancy RMSE improves to +2.78 kcal/mol (previously +4 kcal/mol) but gets regresses slightly to +0.78 kcal/mol (previously +0.60 kcal/mol) for closed/intermediate cases (Table 4). In general, simulations starting from the closed state had $\Delta\Delta G_{calc}$ values that moved towards favorability (i.e. more negative $\Delta\Delta G_{calc}$) and while protein open simulations $\Delta\Delta G_{calc}$ values tended towards unfavorability (i.e. more positive $\Delta\Delta G_{calc}$). This is indicative of the fact that pREST is indeed improving sampling, but it is evident that our $\Delta\Delta G_{calc}$ are far from convergence given the discrepancy RMSE is still large, especially for closed/open transformations.

Longer pREST simulations

Although we see improvements in sampling with pREST, the standard timeframe of 5ns clearly is not long enough completely capture the transition of closed to open in the helix. Here, we simulate 55ns for closed/open transformations and take the final 15ns of the simulation while for closed/intermediate we simulate 25ns and take the last 10ns of the simulation, discarding the initial time as additional equilibration time. In simply running longer, we allow our simulations to perform more exchanges across replicas and thereby allow for better sampling of all conformational states.

Returning to our most extreme transformation, benzene to n-hexylbenzene, we have shown pREST alone does not allow for adequate sampling of the open state (Fig 5). Now, when we run much longer we see far more frequent transitions between all the protein conformational states in the final 10ns window (Fig 8). In viewing all the replicas (Fig 9), we illustrate the dramatic increase in protein conformational transitions in stark contrast of our 5ns simulations (Fig 7).

By simulating longer with pREST we dramatically increase our sampling and almost entirely eliminate the dependence on the initial protein conformational state. For the set

of closed/open transformations the discrepancy RMSE dramatically falls to +0.57 kcal/mol (Table 8). In the set of closed/intermediate transformations, only a few cases required additional simulation time, after doing so we obtain an overall RMSE of +0.42 kcal/mol (Table 5). There still remains some discrepancy between $\Delta\Delta G_{calc}$ from protein open or closed simulations, but it now falls within a much more reasonable range of less than +1kcal/mol.

3 Discussion

Sensitivity to initial protein conformational state

Here, we found that relative free energy calculations can suffer from substantial convergence problems resulting from relatively modest protein conformational changes. These problems have profound implications for the accuracy of computed relative free energies in these cases. Particularly, we find that calculated relative free energies depend on the initial protein conformational state by up to 4 kcal/mol. Even with simulations of up to 55ns per lambda window, far more than is typically used in the 5ns standard protocol for these calculations, we do not completely converge our calculated relative free energies.

By looking at cases that involve a conformation change in the protein, we can clearly see the large dependence of the final predicted ddG on the initial protein conformational state. These cases are when the alchemical transformation involves changing ligands that primarily occupy the closed state into ligands that occupy the intermediate or open states (Table **??**). Despite use of FEP/REST, free energy predictions when starting from the protein closed or open conformational state, were unable to converge to the same predicted ddG. Visualization of these simulations (ex. closed ligands to open ligands) show the protein remains trapped in its initial state. Free energy predictions from simulations starting from the protein closed state give a positive ddG, a result from strain in the protein as the ligand begins to grow in the binding cavity.

Without prior knowledge of preferred protein conformational states on ligand binding, we

can arrive at very different binding affinity predictions. If we only had the crystal structure of the closed protein-ligand complexes, we would incorrectly arrive at the conclusion that larger ligands such as n-hexylbenzene are much worse than smaller ones. On the other hand, the opposite would be concluded in that larger ligands are better binders, if only the open protein-ligand complexes were available. Fortunately, in the case of T4-L99A, knowledge of protein-ligand conformational states is known from the x-ray crystallography studies. [2]

This knowledge of various protein-ligand states is not always the case, especially in early drug discovery phases, where sometimes only the apo crystal structure of a potential therapeutic target exists. Dangerously, a medicinal chemist can dock a library of ligands and run free energy calculations to predict and rank binding affinities. Without knowing the protein can undergo various changes to accommodate different or larger ligands, the chemist would discard ligands with "low" binding affinities. Here in T4-L99A, a relatively small (1-3Å) and localized motion in a helix is shown to drastically impact binding affinity predictions. This demonstrates that special attention and care should be exercised in binding affinity predictions where regions of flexibility surround the binding site.

Inclusion of protein atoms into the REST region (pREST)

In order to get the protein out of its trapped initial state, we included residues from the F-helix into the REST region. Inclusion of residues into the REST region was expected to allow for faster transitions between states by effectively heating up key regions in the F-helix. Indeed in most cases, we found that this improved the time and frequency to transition between conformational states. But in cases of large alchemical transformations, this was still inadequate to entirely eliminate the dependence on the starting protein conformational state, even with simulations of up to 55ns.

Despite adjustments to FEP/REST and longer simulations where we only take the last 30-40% of snapshots, we do not entirely eliminate the dependence on the starting protein conformational state. It is much easier to do so by performing smaller perturbations and/or

running longer simulations. But this is may not always be feasible in industrial medicinal chemistry programs, where as ligands grow more drug-like they typically increase in size and complexity. As a result of increased complexity, FEP calculations commonly require large insertions and deletions of atoms. In turn, this would require much longer simulation times or additional computational power, which are not desirable factors when working with large ligand libraries.

4 Conclusions

Overall, we have shown that identification of discrete protein conformational states is an important factor to consider in inhibitor design. In this study, we demonstrated that there can be a large dependence of predicted binding free energies on the initial protein conformational state. Even while using enhanced sampling techniques and longer simulations than typical, we were unable to completely converge all of our calculations. Close attention should be exercised when performing alchemical transformations that involve large perturbations and/or result in a large conformational change in the protein structure. More importantly, as FEP calculations become a more common part of the computer-aided drug design process, we illustrate that it still requires some level of expertise. Although FEP calculations have shown tremendous recent successes,[4] we are still bounded by problems of adequate sampling and computational power.

5 Methods

5.1 FEP Protocols

Two FEP protocols developed from Schrodinger were used in this project: FEP+ and LigandFEP. FEP+ is a fully automated workflow that will plan the perturbation pathways based of the LOMAP mapping algorithm which uses the maximum common substructure

(MCS) between any pair of compounds. LigandFEP, aimed for academics, is limited in the sense that the user must plan each perturbation instead. LigandFEP does not use the MCS but rather plans the transformation by minimizing the core RMSD. Both protocols still use the default relaxation protocol and FEP/REST methodology and were run on GPUs with Desmond. FEP/REST simulations were performed for 5-55ns depending on the ligand transformation involved. In both protocols, forcefield parameters OPLS3 and OPLS2005 were used.

INSERT FEP MAPPER IMAGE HERE

Trials including protein atoms in the REST region will be referred to with 'pREST' Considering the F-helix ranged from residues 107-115, residues selected to include into the REST region were Glu108, Val111, and Gly113 as sort of a start, mid and end point. Based on the crystal structures and molecular dynamics (MD) simulations and since the F-helix spans residues 107 to 115, we selected residues Glu108, Val111, and Gly113. At the start of the helix sits Glu108 which appears to serve as a hinge point for the transition between conformational states. Next, Val111 appears in the middle of the helix and was observed to undergo the largest motion during transitions. Towards the end was Gly113 which was observed to undergo some minor motions as well.

5.2 Protein/Ligand Preparation

Protein structures were taken from PDBS: 4W52,4W53,4W54,4W55,4W56,4W57,4W58,4W59 corresponding to bound structures of benzene, toluene, ethylbenzene, n-propylbenzene, sec-butylbenzene, n-butylbenzene, n-pentylbenzene, and n-hexylbenzene, respectively. Each simulation will start from either the protein closed state (PDB:4W52) or the open state (PDF: 4W59). When using the FEP+ protocol, ligand crystal structure positions were used as the starting position of the simulation. When using the LigandFEP protocol, two options

would occur:
(1) If the simulation starts from the protein closed state, the benzene crystal position was used as a reference.
(2) The corresponding ligand in the transformation was built by duplicating benzene in place and adding a methyl from the Build/Fragments Toolbar.

For ligands with a longer tail, the fragments would be added in the same direction as the crystal structure, but not overlaid or docked via Glide. If the simulation starts from the protein open state, the n-hexylbenzene crystal position was used as reference. The corresponding ligand in the transformation was built by duplicating n-hexylbenzene in place and deleting a methyl group. All proteins were prepared and aligned in Maestro using the 'Protein Preparation Wizard' tool and the following settings enabled:

- Preprocess: Assign bond orders, Add hydrogens, Create zero-order bonds to metals, Create disulfide bonds, Cap termini, Delete waters beyond 5Åfrom het groups
- Refine: Sample water orientations, Use PROPKA pH: 7.0, Remove waters with less than 3 H-bonds to non-waters, and restrained minimization.

6 Potential Journals

- Journal of Chemical Theory and Computation (H-index: 94)
- Journal of Computational Chemistry (H-index: 139)
- Journal of Physical Chemistry B. (H-index: 295)
- Journal of American Chemical Society (H-index: 412)

7 Objectives

- Assess convergence of relative free energy calculations involving modest amounts of protein conformational change
- Assess the implication of protein sampling challenges for accuracy of relative free energy calculations
- Highlight how even modest ligand changes can induce conformational changes in "rigid" proteins which pose significant sampling challenges
- Test how well we can model and capture protein conformational change in the T4 Lysozyme L99A apolar cavity
- Accurately predict relative binding free energies while sampling protein conformational changes
- Compare performance of FEP plus[4] and LigandFEP.
- Compare performance of OPLS2005 and OPLS3 forcefield parameters.
- Highlight improvement to FEP/REST[5] method by inclusion of protein atoms.

Acknowledgement

- David L. Mobley
- Robert Abel
- Lingle Wang
- Dima Lupyan
- Joseph Goose

Supporting Information Available

8 Experimental

8.1 Discrete Conformations and the Ligands

T4 L99A contains an engineered apolar cavity which is our binding site of interest. The 8 congeneric ligands are all apolar and begins with a simple benzene ring. Subsequent ligands are simply addition of a methyl group to generate a growing tail up until n-hexylbenzene. In response to the growing ligand, the crystal structures show the protein will adopt into 3 conformations aptly named: closed, intermediate, and open. Primarily, the motion of the protein occurs in the F-helix (residues 107-115), which serves as a sort of gating mechanism into the apolar cavity. As the ligand tail expands, the F-helix transitions from closed to open, exposing the cavity to the bulk solvent. From observed electron densities of the F-helix in[2]-Fig2, the ligands occupy each of the conformations given in Table: 1

From a protein conformation clustering analysis, shown,[2] Val111 occupies 3 distinct states iin accordance to the 3 protein conformations. Visualization of side-chain Val111 from the x-stal structures, shows the backbone alpha-carbon moving 1.25Åwhen transitioning from closed to intermediate, 3.25Åintermediate to open, and 3.50Åclosed to open. From our MD

simulations, it is primarily the repulsive interactions between the ligands and Val111 that drive the F-helix to the open state.

8.2 Ligand Binding Affinities

Table 2 Details of the ligand binding affinities and how they were obtained can be found in the paper[2]

Section "Energy of Ligand Binding and Conformational Strain"

Ligands n-pentylbenzene and n-hexylbenzene affinities are inaccessible due to solubility limits. Generally, as the ligands grow from benzene to n-butylbenzene, the affinity rises linearly.

9 FEP+ vs. LigandFEP

- Comparisons between the two protocols will be used to show that, although limited, LigandFEP does not result in a difference in performance of accurate predicted relative binding free energies.
- Comparisons between FFs will show that the new and improved OPLS3 parameters are give better RBFE predictions for larger transformations.
- ***INSERT TABLE COMPARING OPLS2005/OPLS3 BETWEEN TWO PROTOCOLS***

  Both are within the same MUE/RMSE range
- Highlight inconsistency in predicted RBFE when the protein starting conformation is varied.
- Predicted ddGs are in the wrong direction frequently when starting from protein closed. If we assume smaller to larger ligands should yield favorable (-) ddGs.

10 OPLS2005 vs. OPLS3

11 Case studies

11.1 Small Ligands

- Small transformations and ligands generally occupy protein closed state.
- Equal performance with either FF.
- Highlight good case: Toluene to Ethylbenzene with OPLS3
  - Small perturbation: Adding one carbon
  - Both ligands occupy the closed state with some intermediate
  - Protein starting conformation does not result in a large discrepancy in predicted ddGs.
  - ***INSERT RMSD GRAPH***
  - RMSD graph over the lambda=0 and lambda=1 corresponding to the toluene and ethylbenzene end states shows at each time point which conformation (reference to crystal structure) the simulation has the lowest RMSD to.
    Purple = Closed, Teal = Intermediate, and Green = Open.
  - Highlight good sampling/number of transitions between either state when starting from closed.
  - Highlight points 0-1ns are still stuck in open conformation. 240ps relaxation protocol wasn't sufficient to discard few open points.
    Show that this does not have a large impact in the final ddG from sliding time.
- Highlight not-so-good case: Toluene to n-propylbenzene
  - Perturbation involves adding two carbons

  - Discrepancy from protein starting conformation increases.
  - Highlight more open points are present (0-2ns) in open runs.

- Growing ligands require more time for helix to relax out of open state.
- Inclusion of protein residues in REST region should allow for faster transition between states which will give us a lower discrepancy between open/closed runs.
- Apply pREST to previous case and show faster relaxation out of open conformation resulting in a lower discrepancy.
- ***INSERT COMPARISON BETWEEN NORMAL AND PREST RUNS WITH EXPERIMENTAL ddG ***
- Show pREST increases agreement with experiment and lowers error from protein starting conformation.

11.2 Intermediate Ligands

Insert all data

11.3 Open Ligands

Simulations starting from closed give a predicted ddG of +2.74 kcal/mol and from open +1.37 kcal/mol. Upon closer inspection of the closed simulations corresponding to the final state of n-hexylbenzene, we find that the protein does not remain trapped in its initial state. Instead, the region of the helix around Gly113 briefly opens to relieve the strain but quickly closes back. Hence, we still see some strain energy as the protein fails to completely stabilize into the open conformation. Viewing the open simulations, reveals that the protein is no longer trapped only in the open state and makes transitions into the intermediate and closed states. In making these transitions, the protein also experiences strain as the tail pushes the

F-helix back into open from the other conformational states.

INSERT RMSD GRAPH SHOWING TRANSITIONS BETWEEN STATES FOR BENZENE TO NHEXYL

By comparing these graphs we show that inclusion of protein residues in the REST region does allow for more/faster transitions between states

***Compare pREST runs with default REST in order to show that pREST allows for the protein to get of being trapped in its initial state.

Here we have shown pREST allows for more motion in the helix in comparison to the default REST protocol. Although there is a reduction in the discrepancy, it is still fairly large (+1.37 kcal/mol). Thus, we cannot say these are converged, where this poor convergence results from inadequate sampling of all the protein conformational states. Such is the case of closed pREST simulations of n-hexylbenzene, where we only see a partial opening of the helix. Similarly, in open pREST simulations with n-hexylbenzene, the helix does not enter the closed conformation. Where as it is expected to at least partially occupy the closed state, according to the loop/ligand occupancy table. It seems likely that the overall motion of the helix is not entirely accessible in the range of up to 50ns (longest we have simulated).

INSERT DATA FROM 55ns pREST simulations

Table showing all closed to open cases

Here we can show that the discrepancy does not entirely go away, but gradually gets smaller with smaller perturbations

Show that in 50ns the protein is only barely able to stabilize in the open state and vice versa

This material is available free of charge via the Internet at `http://pubs.acs.org/`.

12 Figures

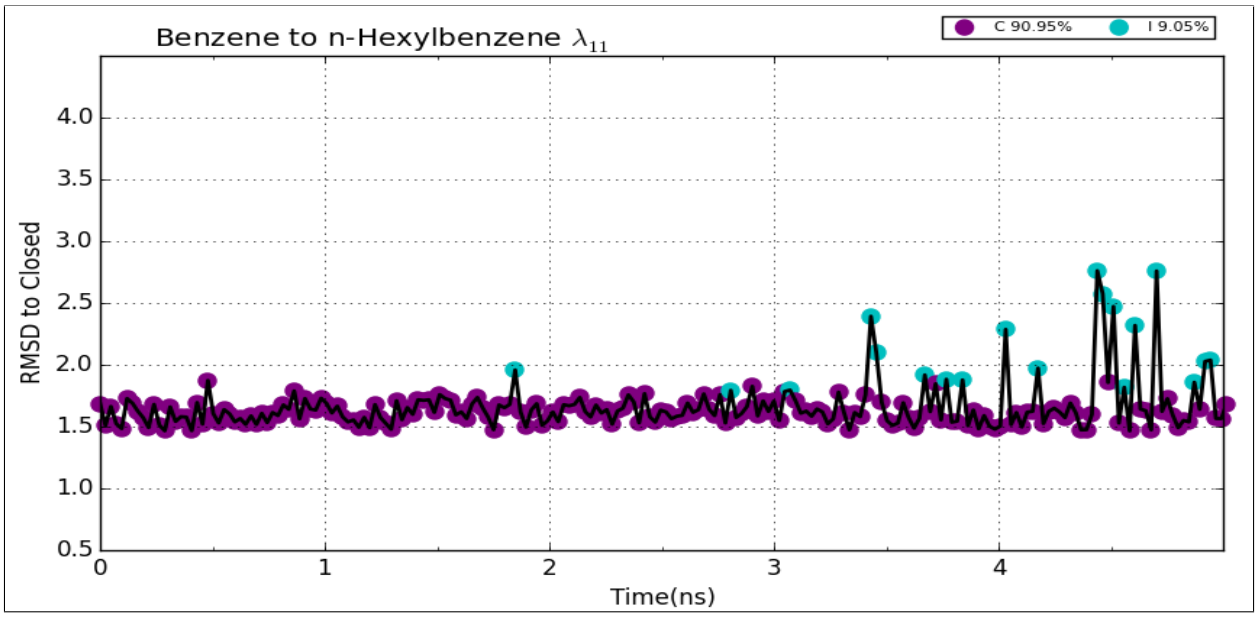

Figure 1: Closed - Benzene to n-Hexylbenzene 0-5ns RMSD Replica11

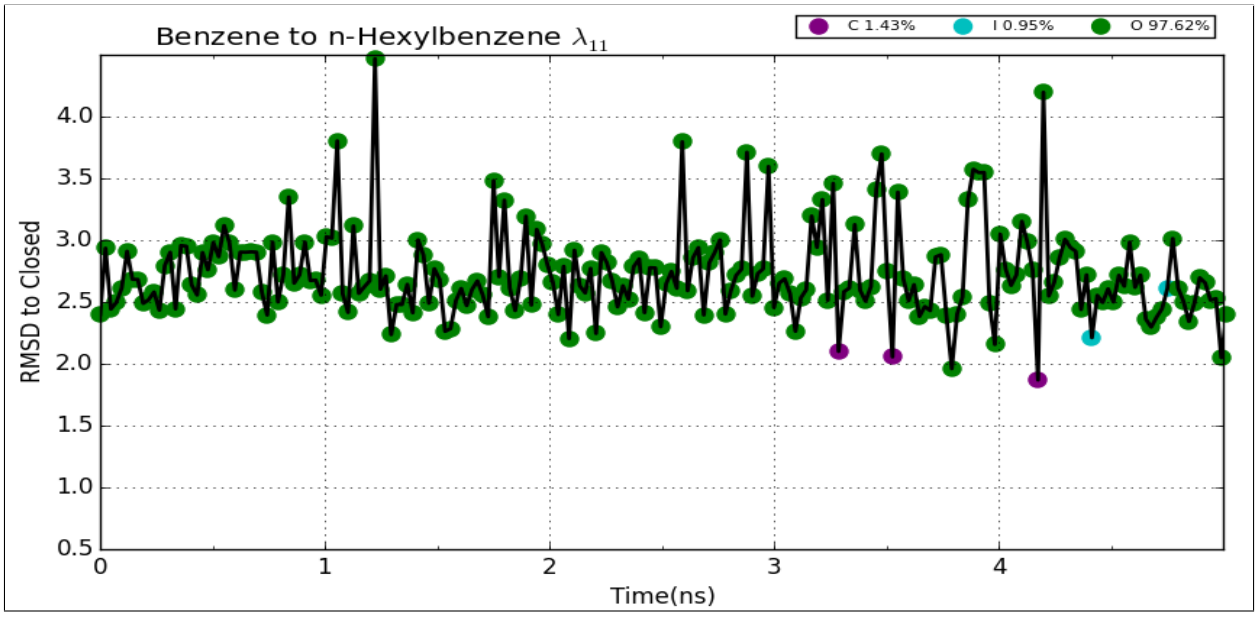

Figure 2: Open - Benzene to n-Hexylbenzene 0-5ns RMSD Replica11

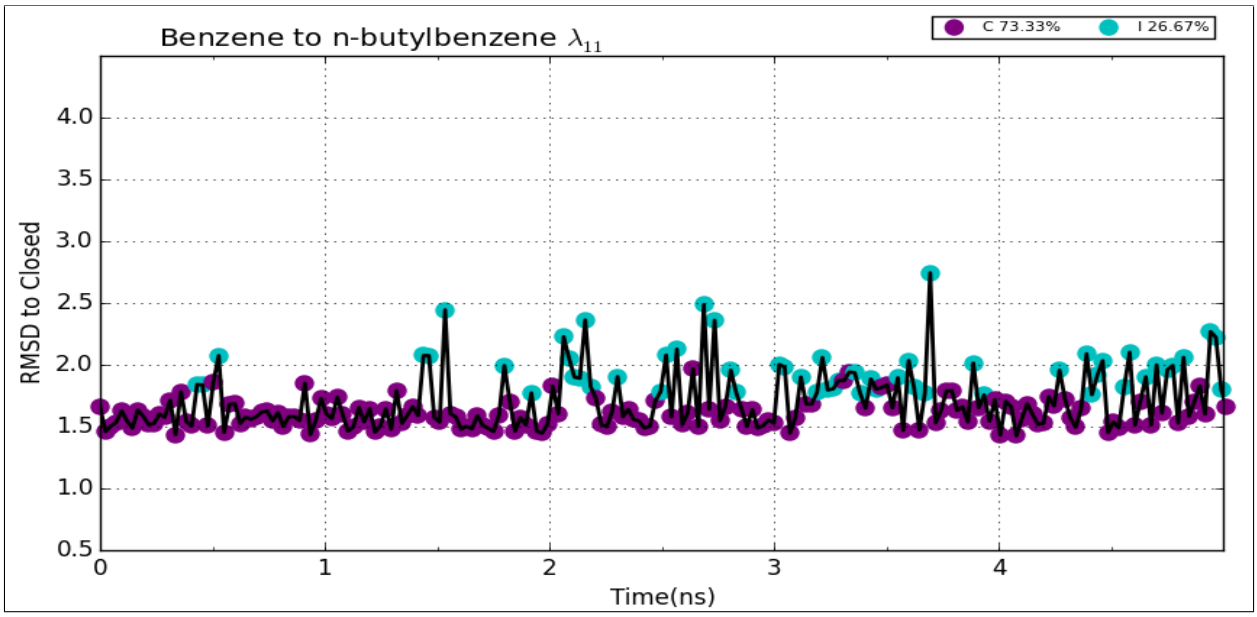

Figure 3: Closed - Benzene to n-butylbenzene 0-5ns RMSD Replica11

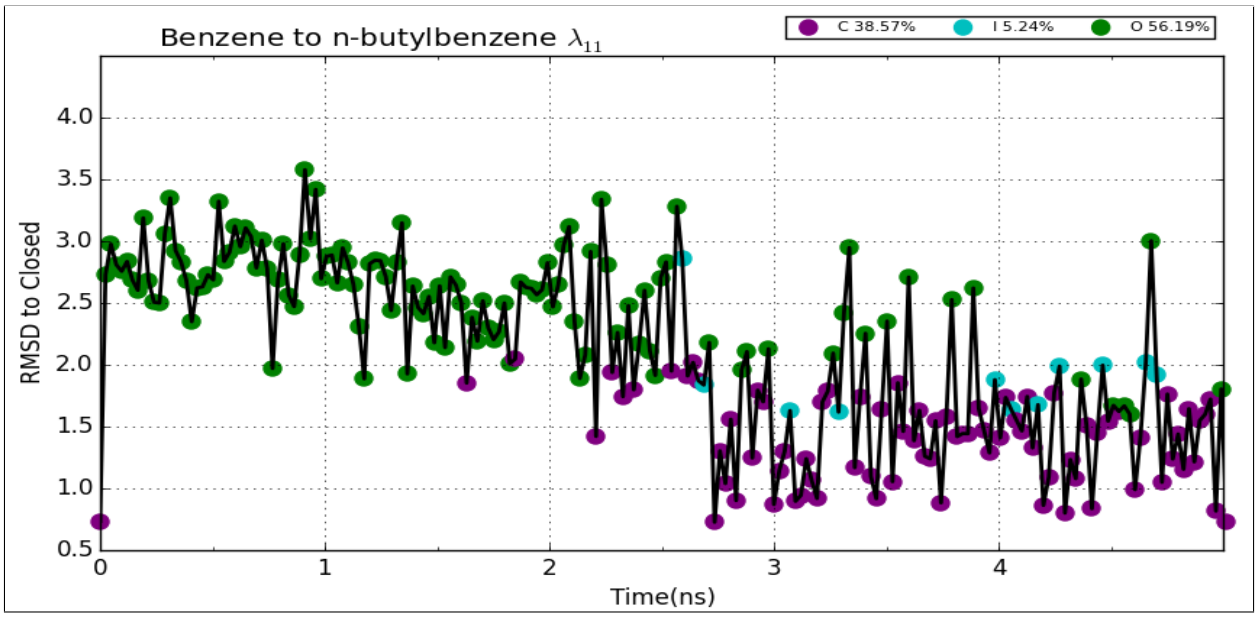

Figure 4: Open - Benzene to n-butylbenzene 0-5ns RMSD Replica11

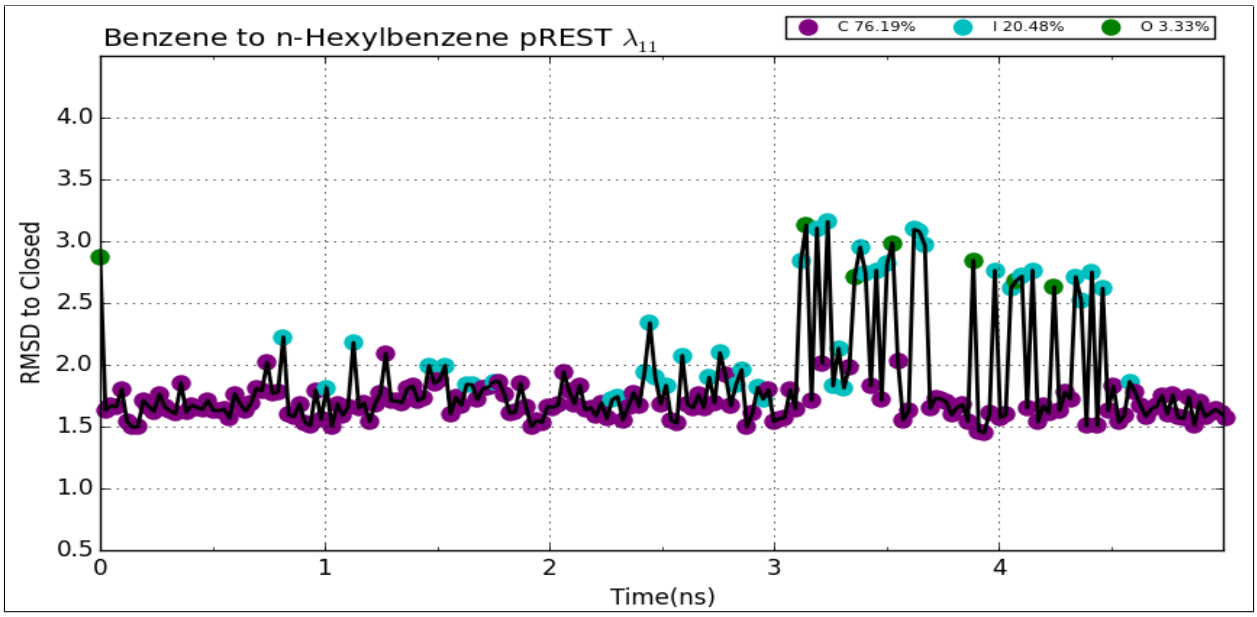

Figure 5: Closed - Benzene to n-Hexylbenzene 0-5ns RMSD Replica11 pREST

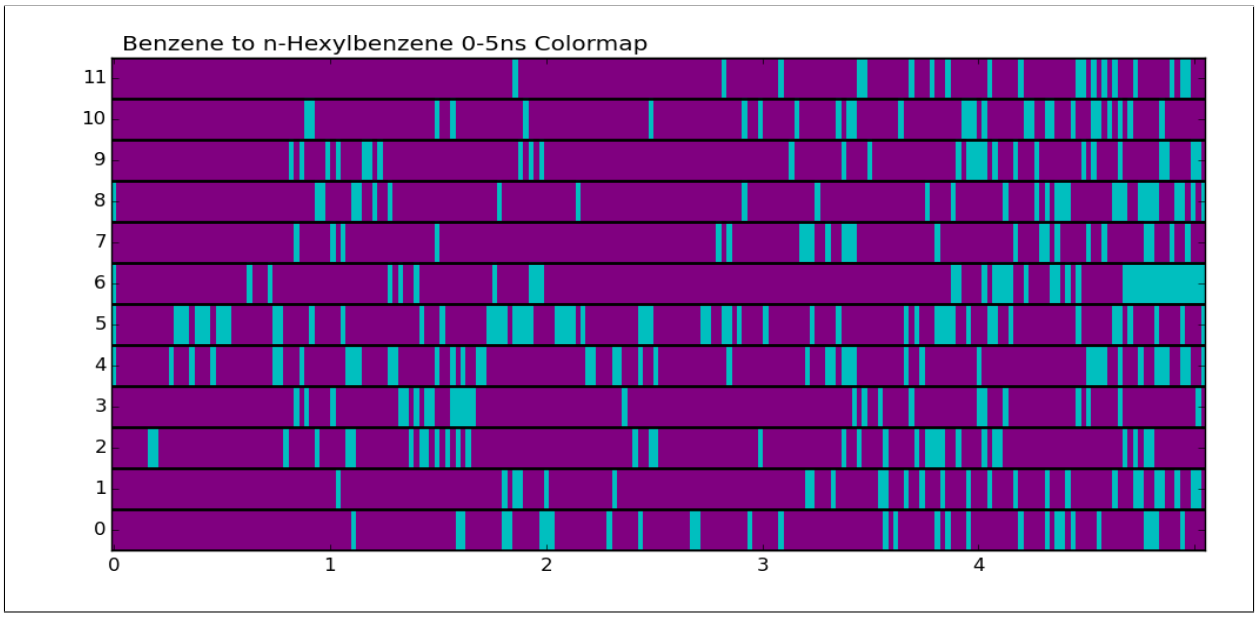

Figure 6: Closed - Benzene to n-Hexylbenzene 0-5ns Colormap

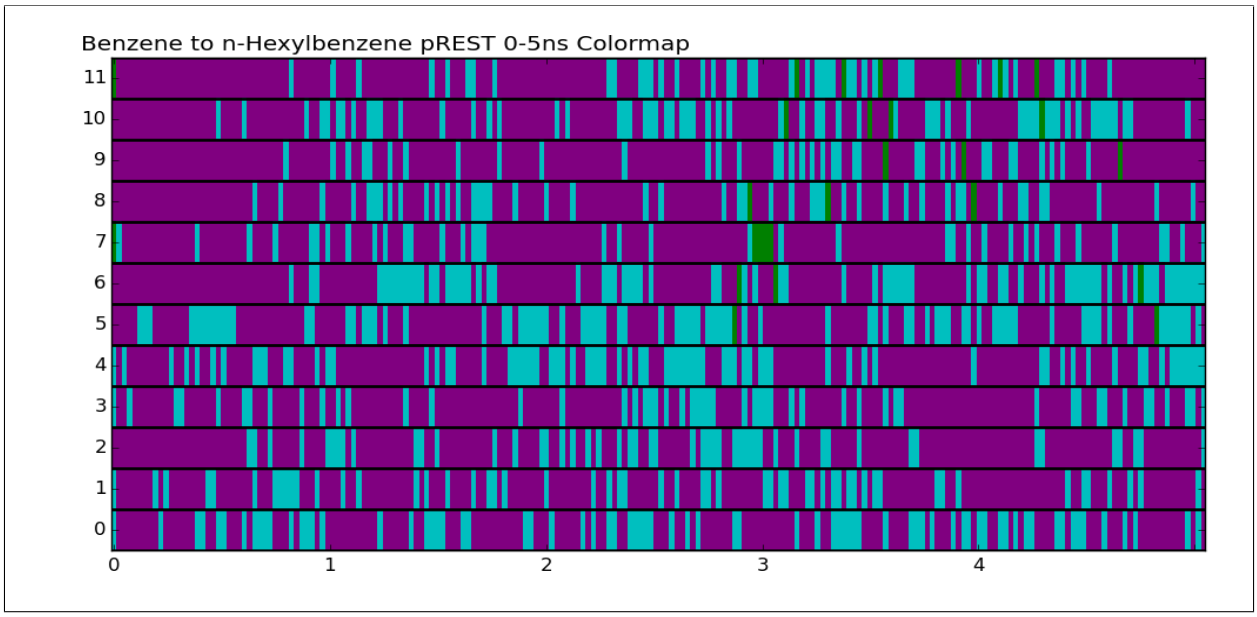

Figure 7: Closed - Benzene to n-Hexylbenzene 0-5ns Colormap pREST

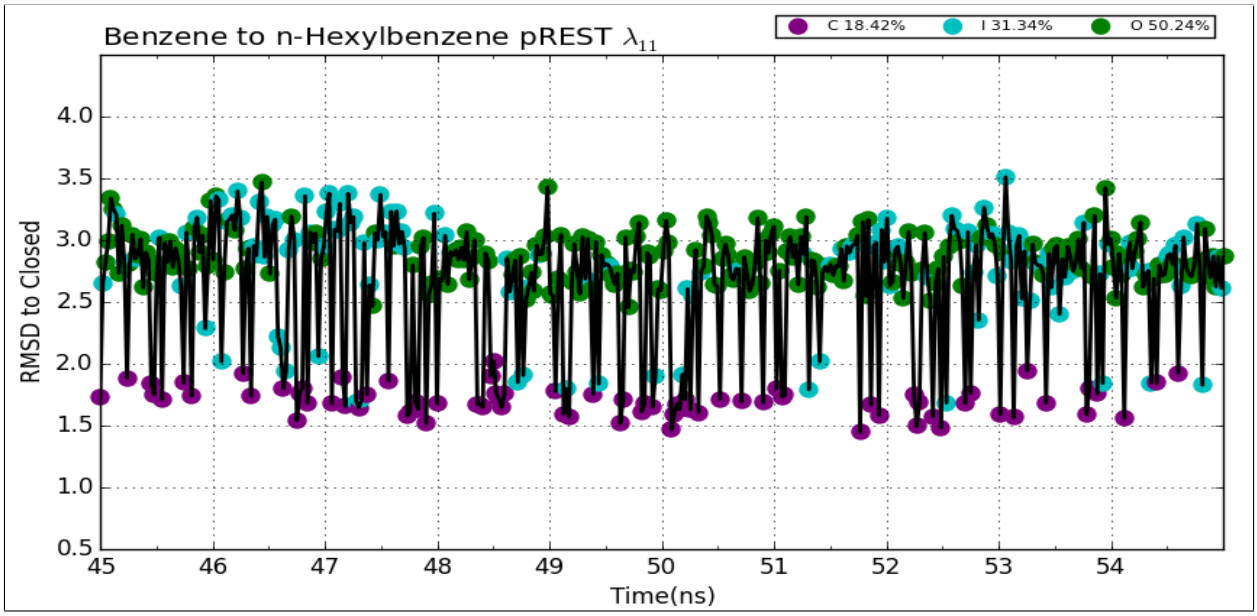

Figure 8: Closed - Benzene to n-Hexylbenzene 45-55ns RMSD Replica11 pREST

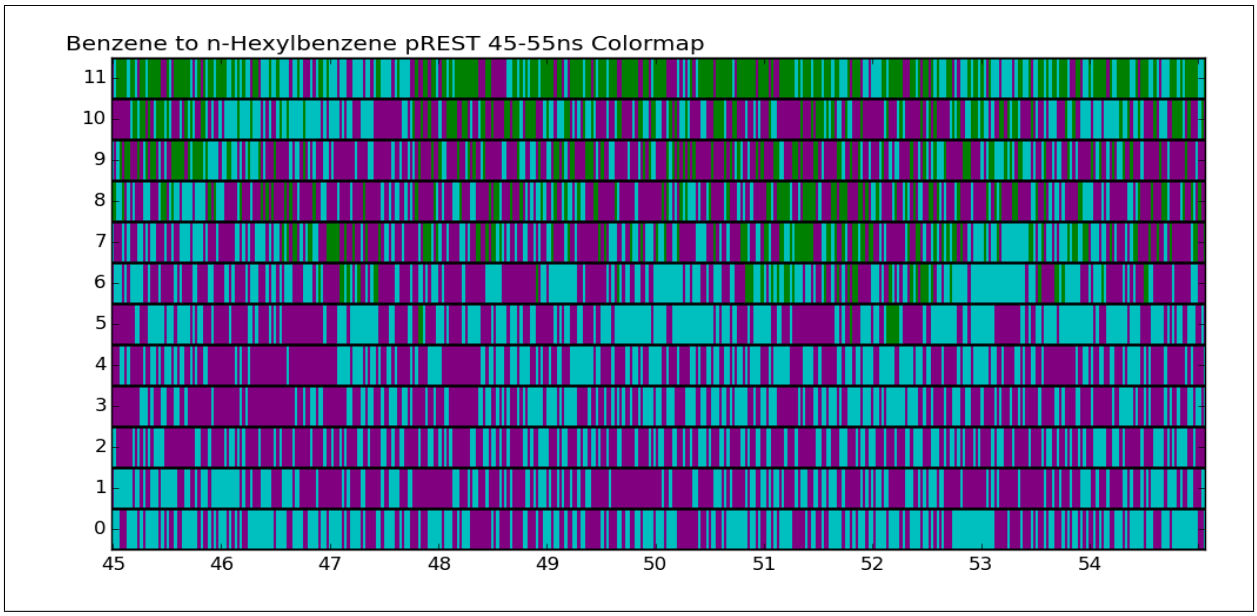

Figure 9: Closed - Benzene to n-Hexylbenzene 45-55ns Colormap pREST

13 Tables

Table 1: Loop Occupancies[2]

Ligand	C	I	O
benzene	0.9	-	-
toluene	0.8	0.2	-
ethylbenzene	0.5	0.5	-
n-propylbenzene	0.6	0.4	-
sec-butylbenzene	0.4	0.6	-
n-butylbenzene	0.1	0.6	0.3
n-pentylbenzene	0.3	-	0.7
n-hexylbenzene	0.3	-	0.7

Table 2: Ligand Binding Affinities

PDB	**Ligand**	**ΔG_{exp}**	**Error**
4W52	benzene	-5.19	0.16
4W53	toluene	-5.52	0.04
4W54	ethylbenzene	-5.76	0.07
4W55	n-propylbenzene	-6.55	0.02
4W56	sec-butylbenzene	N/A	-
4W57	n-butylbenezene	-6.70	0.02
4W58	n-pentylbenzene	N/A	-
4W59	n-hexylbenzene	N/A	-

Table 3: Closed-Intermediate Transformations

Ligand 1	**Ligand 2**	**Closed**	*Err*	**Open**	*Err*	**C-O Diff**
benzene	n-butylbenzene	0.58	*0.07*	-0.59	*0.09*	1.17
toluene	n-butylbenzene	-0.28	*0.06*	-1.27	*0.09*	0.99
ethylbenzene	n-butylbenzene	0.24	*0.07*	-0.23	*0.07*	0.47
n-propylbenzene	n-butylbenzene	0.99	*0.06*	0.63	*0.04*	0.36
benzene	sec-butylbenzene	2.36	*0.09*	2.14	*0.11*	0.22
toluene	sec-butylbenzene	1.47	*0.07*	1.14	*0.09*	0.33
ethylbenzene	sec-butylbenzene	1.90	*0.08*	1.77	*0.07*	0.13
n-propylbenzene	sec-butylbenzene	2.86	*0.06*	2.67	*0.05*	0.19

Table 4: Closed-Intermediate Transformations pREST

Ligand 1	Ligand 2	Closed	*Err*	Open	*Err*	C-O Diff
benzene	n-butylbenzene	-0.10	*0.11*	-0.72	*0.12*	0.62
toluene	n-butylbenzene	0.90	*0.09*	-0.36	*0.09*	1.26
ethylbenzene	n-butylbenzene	-0.20	*0.09*	-0.43	*0.08*	0.23
n-propylbenzene	n-butylbenzene	1.00	*0.07*	0.49	*0.06*	0.51
benzene	sec-butylbenzene	0.45	*0.08*	1.34	*0.12*	0.89
toluene	sec-butylbenzene	0.60	*0.12*	0.60	*0.10*	0.0
ethylbenzene	sec-butylbenzene	1.09	*0.10*	1.69	*0.09*	0.60
n-propylbenzene	sec-butylbenzene	1.88	*0.07*	3.05	*0.08*	1.17

Table 5: Closed-Intermediate Transformations pREST 15-25ns

Ligand 1	Ligand 2	Closed	*Err*	Open	*Err*	C-O Diff
benzene	n-butylbenzene	-1.15	*0.09*	-0.72	*0.12*	0.43
toluene	n-butylbenzene	-0.52	*0.09*	-0.36	*0.09*	0.16
ethylbenzene	n-butylbenzene	-0.20	*0.09*	-0.43	*0.08*	0.23
n-propylbenzene	n-butylbenzene	0.53	*0.07*	0.49	*0.06*	0.04
benzene	sec-butylbenzene	0.45	*0.08*	1.34	*0.12*	0.89
toluene	sec-butylbenzene	0.60	*0.12*	0.60	*0.10*	0.0
ethylbenzene	sec-butylbenzene	1.09	*0.10*	1.69	*0.09*	0.60
n-propylbenzene	sec-butylbenzene	1.88	*0.07*	1.84	*0.06*	0.04

Table 6: Closed-Open Transformations

Ligand 1	Ligand 2	Closed	*Err*	Open	*Err*	C-O Diff
benzene	n-pentylbenzene	2.36	*0.12*	-1.33	*0.11*	3.69
toluene	n-pentylbenzene	1.77	*0.09*	0.34	*0.10*	1.43
ethylbenzene	n-pentylbenzene	2.45	*0.08*	0.46	*0.09*	1.99
n-propylbenzene	n-pentylbenzene	3.46	*0.08*	-0.22	*0.08*	3.68
benzene	n-hexylbenzene	4.13	*0.16*	-0.61	*0.15*	4.74
toluene	n-hexylbenzene	2.90	*0.14*	-1.63	*0.08*	4.53
ethylbenzene	n-hexylbenzene	3.63	*0.11*	-0.76	*0.09*	4.39
n-propylbenzene	n-hexylbenzene	5.85	*0.10*	0.13	*0.06*	5.72

[a] Some text; [b] Some more text.

Table 7: Closed-Open Transformations pREST

Ligand 1	Ligand 2	Closed	*Err*	Open	*Err*	C-O Diff
benzene	n-pentylbenzene	1.45	*0.13*	0.15	*0.10*	1.30
toluene	n-pentylbenzene	1.40	*0.13*	0.82	*0.11*	0.58
ethylbenzene	n-pentylbenzene	2.89	*0.10*	1.32	*0.10*	1.57
n-propylbenzene	n-pentylbenzene	4.40	*0.12*	1.06	*0.09*	3.34
benzene	n-hexylbenzene	2.74	*0.19*	1.37	*0.13*	1.37
toluene	n-hexylbenzene	3.21	*0.15*	-1.08	*0.09*	4.29
ethylbenzene	n-hexylbenzene	3.39	*0.11*	-0.14	*0.10*	3.53
n-propylbenzene	n-hexylbenzene	4.93	*0.12*	1.28	*0.10*	3.65

[a] Some text; [b] Some more text.

Table 8: Closed-Open Transformations pREST 40-55ns

Ligand 1	Ligand 2	Closed	*Err*	Open	*Err*	C-O Diff
benzene	n-pentylbenzene	1.86	*0.06*	1.50	*0.06*	0.36
toluene	n-pentylbenzene	1.03	*0.06*	0.71	*0.06*	0.32
ethylbenzene	n-pentylbenzene	1.69	*0.06*	1.60	*0.06*	0.09
n-propylbenzene	n-pentylbenzene	3.43	*0.04*	2.44	*0.04*	0.99
benzene	n-hexylbenzene	2.14	*0.08*	1.41	*0.07*	0.73
toluene	n-hexylbenzene	0.33	*0.08*	1.16	*0.06*	0.84
ethylbenzene	n-hexylbenzene	1.97	*0.07*	2.39	*0.06*	0.42
n-propylbenzene	n-hexylbenzene	3.49	*0.06*	3.44	*0.05*	0.05

[a] Some text; [b] Some more text.

Table 9: Closed to n-butylbenzene Transformations (Closed Protein)

Ligand 1	Ligand 2	Closed	*Err*	C-Exp Err
benzene	n-butylbenzene	0.58	*0.07*	2.09
toluene	n-butylbenzene	-0.28	*0.06*	0.90
ethylbenzene	n-butylbenzene	0.24	*0.07*	1.18
n-propylbenzene	n-butylbenzene	0.99	*0.06*	1.14

Table 10: Closed to n-butylbenzene Transformations (Open Protein)

Ligand 1	Ligand 2	Open	*Err*	O-Exp Err
benzene	n-butylbenzene	-0.59	*0.09*	0.92
toluene	n-butylbenzene	-1.27	*0.09*	0.09
ethylbenzene	n-butylbenzene	-0.23	*0.07*	0.71
n-propylbenzene	n-butylbenzene	0.63	*0.04*	0.78

References

(1) Boyce, S. E.; Mobley, D. L.; Rocklin, G. J.; Graves, A. P.; Dill, K. A.; Shoichet, B. K. *Journal of Molecular Biology* **2009**, *394*, 747 – 763.

(2) Merski, M.; Fischer, M.; Balius, T. E.; Eidam, O.; Shoichet, B. K. *Proceedings of the National Academy of Sciences* **2015**, *112*, 5039–5044.

(3) Wang, L.; Deng, Y.; Knight, J. L.; Wu, Y.; Kim, B.; Sherman, W.; Shelley, J. C.; Lin, T.; Abel, R. *Journal of Chemical Theory and Computation* **2013**, *9*, 1282–1293, PMID: 26588769.

(4) Wang, L.; Wu, Y.; Deng, Y.; Kim, B.; Pierce, L.; Krilov, G.; Lupyan, D.; Robinson, S.; Dahlgren, M. K.; Greenwood, J.; Romero, D. L.; Masse, C.; Knight, J. L.; Steinbrecher, T.; Beuming, T.; Damm, W.; Harder, E.; Sherman, W.; Brewer, M.; Wester, R.; Murcko, M.; Frye, L.; Farid, R.; Lin, T.; Mobley, D. L.; Jorgensen, W. L.; Berne, B. J.; Friesner, R. A.; Abel, R. *Journal of the American Chemical Society* **2015**, *137*, 2695–2703, PMID: 25625324.

(5) Wang, L.; Friesner, R. A.; Berne, B. J. *The Journal of Physical Chemistry B* **2011**, *115*, 9431–9438, PMID: 21714551.