PLoS one

Auto-FACE: An NMR Based Binding Site Mapping Program for Fast Chemical Exchange Protein-Ligand Systems

Janarthanan Krishnamoorthy[1], Victor C. K. Yu[2], Yu-Keung Mok[1]*

1 Department of Biological Sciences, National University of Singapore, Singapore, Singapore, 2 Department of Pharmacy, National University of Singapore, Singapore, Singapore

Abstract

Background: Nuclear Magnetic Resonance (NMR) spectroscopy offers a variety of experiments to study protein-ligand interactions at atomic resolution. Among these experiments, ^{15}N Heteronuclear Single Quantum Correlation (HSQC) experiment is simple, less time consuming and highly informative in mapping the binding site of the ligand. The interpretation of ^{15}N HSQC becomes ambiguous when the chemical shift perturbations are caused by non-specific interactions like allosteric changes and local structural rearrangement. Under such cases, detailed chemical exchange analysis based on chemical shift perturbation will assist in locating the binding site accurately.

Methodology/Principal Findings: We have automated the mapping of binding sites for fast chemical exchange systems using information obtained from ^{15}N HSQC spectra of protein serially titrated with ligand of increasing concentrations. The automated program Auto-FACE (Auto-FAst Chemical Exchange analyzer) determines the parameters, e.g. rate of change of perturbation, binding equilibrium constant and magnitude of chemical shift perturbation to map the binding site residues. Interestingly, the rate of change of perturbation at lower ligand concentration is highly sensitive in differentiating the binding site residues from the non-binding site residues. To validate this program, the interaction between the protein hBcl$_{XL}$ and the ligand BH3I-1 was studied. Residues in the hydrophobic BH3 binding groove of hBcl$_{XL}$ were easily identified to be crucial for interaction with BH3I-1 from other residues that also exhibited perturbation. The geometrically averaged equilibrium constant (3.0×10^4) calculated for the residues present at the identified binding site is consistent with the values obtained by other techniques like isothermal calorimetry and fluorescence polarization assays (12.8×10^4). Adjacent to the primary site, an additional binding site was identified which had an affinity of 3.8 times weaker than the former one. Further NMR based model fitting for individual residues suggest single site model for residues present at these binding sites and two site model for residues present between these sites. This implies that chemical shift perturbation can represent the local binding event much more accurately than the global binding event.

Conclusion/Significance: Detail NMR chemical shift perturbation analysis enabled binding site residues to be distinguished from non-binding site residues for accurate mapping of interaction site in complex fast exchange system between small molecule and protein. The methodology is automated and implemented in a program called "Auto-FACE", which also allowed quantitative information of each interaction site and elucidation of binding mechanism.

Introduction

Basic research on protein-ligand and protein-protein interaction has contributed a lot to the success of structure-aided drug design and development [1]. A myriad of techniques are available to study such interactions, among which NMR spectroscopy has been unique in giving dynamic details at atomic resolution [2–4]. The chemical shift, a fundamental property of nucleus, gets perturbed when an adjacent nucleus comes in close proximity to it. Such perturbation can be explained with the help of phenomena like "chemical exchange" and "relaxation" [5,6]. Extensive theories are available to explain chemical exchange and

relaxation, based on which, many of the complicated NMR experiments have been successfully established [7–9]. Chemical exchange by definition is the switching of nuclei from one environment to another. For instance, addition of ligand or change in pH and temperature would result in chemical exchange [5]. On the other hand, relaxation is a process by which the excited nucleus return to its ground equilibrium [10,11]. The inherent nature of the nucleus and its surrounding influence the relaxation process.

Both chemical exchange and relaxation modulate the basic line shape characteristics of NMR like the offset or analogously Larmor frequency; the line width at half maximum; and the phase

and intensity of peak [5,12]. For a two state system, where nucleus A is chemically exchanging with nucleus B,

$$A \underset{k_{-1}}{\overset{k_{+1}}{\rightleftarrows}} B$$
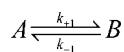
Assume λ_1 and λ_2 represents the Larmor frequency of A and B, M_A and M_B are the respective magnetization. By default, M_A will give rise to a peak at λ_1, but because of chemical exchange with B, it will also give rise to a peak at λ_2. Conversely, M_B will give rise a peak at λ_2 and because of chemical exchange with A, it will also give rise to a peak at λ_1 [12]. The analytical expression for M_A and M_B can be obtained by solving the classical Bloch-McConnell equations [13–15].

To study the chemically exchanging species individually, an easier approach would be to look at the components at λ_1 and λ_2 rather than signals M_A and M_B [12,14,16]. Both M_A and M_B contributes to the component peaks at λ_1 and λ_2. Addition of the λ_1 components from M_A and M_B and the λ_2 components from M_A and M_B, would give a spectrum that can be easily analyzed as λ_1 and λ_2 peaks since these components correlate directly with its population A, B (Figure 1A). Moreover, the rate of chemical exchange is also important as it influences all the above mentioned peak characteristics significantly. Based on the rate (k), the chemical exchange phenomenon can be classified into fast, intermediate and slow exchange regimes. By definition, fast exchange requires $|\lambda_1 - \lambda_2| < k$, whereas for slow exchange, $|\lambda_1 - \lambda_2| > k$. In interme-diate exchange, the difference in Larmor frequency of the exchanging species equals to the exchange rate i.e. $|\lambda_1 - \lambda_2| = k$ [17]. Experimentally, fast exchange systems will show a single peak with the components of A and B appearing at a population weighted frequency λ_{avg}, where λ_{avg} is in between λ_1 and λ_2. In intermediate exchange, a single peak will appear as seen with fast exchange, but the phases of the contributing components A and B are highly distorted and gives rise to a very broad peak. Sometimes, it may even disappear amidst noise peaks due to poor signal to noise ratio. In slow exchange, two individual peaks appear at λ_1 and λ_2 corresponding to the components A and B, the area of which are population weighted. To summarize, chemical shift, phase and peak intensity are population weighted for fast, intermediate and slow exchange systems, respectively (Figure 2) [17].

The fast exchange protein-ligand systems show a characteristic 'peak walking' pattern in spectra on gradual addition of ligand. This variation in chemical shift due to increasing ligand concentrations can be explained analytically by linear combina-tion of population weighted individual chemical shifts [18,19]. For example, in a two state system comprised of free $[P]$ and bound protein $[PL]$, the averaged chemical shift δ is given as,

$$\delta = \delta_{[P]}f_{[P]} + \delta_{[PL]}f_{[PL]}$$

where $f_{[P]} = \dfrac{[P]}{[P_T]}$, the mole fraction of $[P]$, $f_{[PL]} = \dfrac{[PL]}{[P_T]}$, the mole fraction of $[PL]$ and $[P_T] = [P] + [PL]$, the total protein concen-tration. $\delta_{[sub]}$ refers to the chemical shift corresponding to the
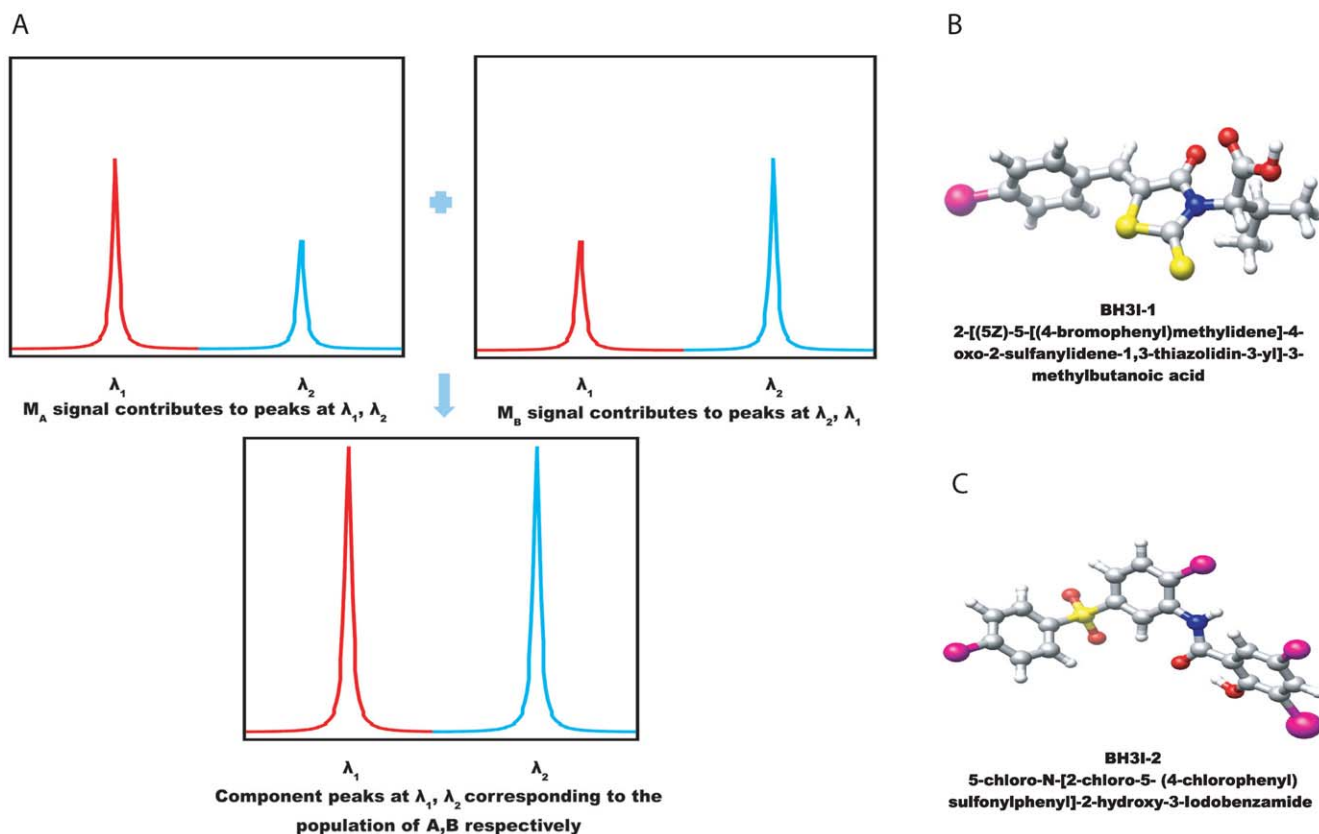
Figure 1. Component signals of population A and B and structural comparison of BH3I-1 and its analogue BH3I-2. (A) M_A and M_B both contributes to the component peaks at λ_1 and λ_2 which are directly correlated with its respective population A and B. (B) & (C) Structural comparison of BH3I-1 and its analogue BH3I-2.
doi:10.1371/journal.pone.0008943.g001
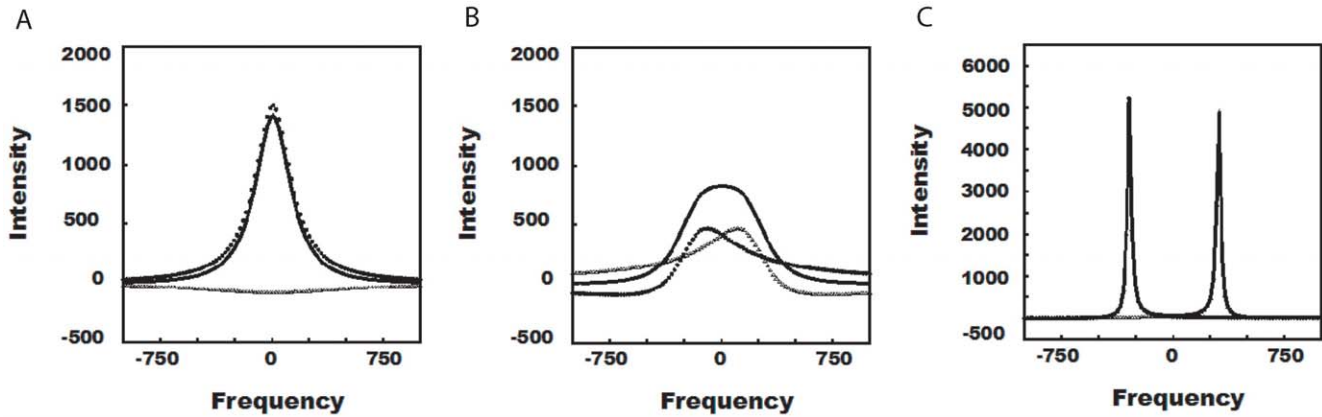
Figure 2. Simulation of fast, intermediate and slow exchange regimes for two site chemical exchange using Mexico 3.1 [53]. The offset (ω_1, ω_2) are set at ± 300Hz for site A and B. The T_2 relaxation rates are 1Hz each. Assuming forward and reverse rates to be same, the chemical exchange rates are set at 2400 sec^{-1}, 1200 sec^{-1} and 100 sec^{-1}, for fast, intermediate and slow exchange systems, respectively. In all cases, the population of A:B is fixed at 1:1. ($\square\square\square$: Component of site A, ■ ■ ■: Component of site B, – : Sum of both components (A+B)).
doi:10.1371/journal.pone.0008943.g002

subscripted free or bound form. Though weakly interacting ligands with complex mechanisms can be studied in detail by making such fast exchange approximations, we were interested in finding out which of the NMR derived parameters correlates well with the binding process rather than non-specific allosteric structural changes [19]. Here, we show that detailed analysis of chemical shift perturbation for complex fast exchange systems enable us to obtain parameters like the rate of change of perturbation, binding equilibrium constant and magnitude of chemical shift perturbation, which can be collectively used to distinguish the binding site residues from the bulk of residues.

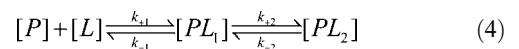## Results and Discussion

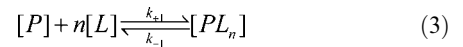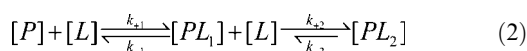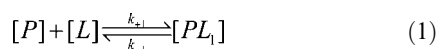### Mechanisms of protein-ligand interaction

The mechanism of interaction can be as simple as a single site binding or much complex sequential binding. Despite, the nature of mechanism, if the ligand interacts weakly with protein exhibiting shorter residence time, fast exchange approximations can be made and explicit analytical expressions can be derived for δ relating $[L]$ [18–20]. In fast exchange approximation, two assumptions are empirically made,

1. The overall chemical shift is the sum of population weighted individual chemical shifts
2. All the exchanging species are in equilibrium

Mechanisms explaining different physical situations are considered; for example,

1. Single site binding (1)
2. Sequential two site binding (2)
3. Simultaneous 'n' site binding (3)
4. Single site binding with allosteric contribution (4)

which are illustrated as,

$$[P]+[L]\underset{k_{-1}}{\overset{k_{+1}}{\rightleftharpoons}}[PL_1] \qquad (1)$$

$$[P]+[L]\underset{k_{-1}}{\overset{k_{+1}}{\rightleftharpoons}}[PL_1]+[L]\underset{k_{-2}}{\overset{k_{+2}}{\rightleftharpoons}}[PL_2] \qquad (2)$$

$$[P]+n[L]\underset{k_{-1}}{\overset{k_{+1}}{\rightleftharpoons}}[PL_n] \qquad (3)$$

$$[P]+[L]\underset{k_{-1}}{\overset{k_{+1}}{\rightleftharpoons}}[PL_1]\underset{k_{-2}}{\overset{k_{+2}}{\rightleftharpoons}}[PL_2] \qquad (4)$$

where $[P]$ and $[L]$ denotes free protein and ligand species. $[PL_1]$, $[PL_2]$ and $[PL_n]$ are the ligand bound protein forms. Assuming fast exchange approximation, the expressions for δ can be written as

$$\delta = \delta_{[P]}f_{[P]} + \delta_{[PL_1]}f_{[PL_1]} \qquad \text{[model (1)]}$$

$$\delta = \delta_{[P]}f_{[P]} + \delta_{[PL_1]}f_{[PL_1]} + \delta_{[PL_2]}f_{[PL_2]} \qquad \text{[model (2)]}$$

$$\delta = \delta_{[P]}f_{[P]} + \delta_{[PL_n]}f_{[PL_n]} \qquad \text{[model (3)]}$$

$$\delta = \delta_{[P]}f_{[P]} + \delta_{[PL_1]}f_{[PL_1]} + \delta_{[PL_2]}f_{[PL_2]} \qquad \text{[model (4)]}$$

where δ represents the averaged or overall chemical shift and $f_{[sub]}$ and $\delta_{[sub]}$ are the mole fraction and chemical shifts for the subscripted species, respectively. Assuming equilibrium, mole fraction can be explicitly written in terms of $[L]$ as follows,

$$\delta = \frac{\delta_{[P]} + \delta_{[PL]}K_A[L]}{1 + K_A[L]} \qquad \text{[model (1)]} \qquad (5)$$

$$\delta = \frac{\delta_{[P]} + \delta_{[PL_1]}K_A[L] + \delta_{[PL_2]}K_AK_B[L]^2}{1 + K_A[L] + K_AK_B[L]^2} \qquad \text{[model (2)]} \qquad (6)$$

$$\delta = \frac{\delta_{[P]} + \delta_{[PL_1]}K_A[L]^n}{1 + K_A[L]^n} \qquad \text{[model (3)]} \qquad (7)$$

$$\delta = \frac{\delta_{[P]} + \delta_{[PL_1]}K_A[L] + \delta_{[PL_2]}K_AK_B[L]}{1 + K_A[L] + K_AK_B[L]} \qquad \text{[model (4)]} \qquad (8)$$

Correction for free ligand concentration

In the above equations the free ligand concentration $[L]$ appears rather than total ligand concentration $[L_T]$. The determination of $[L]$ from $[L_T]$ is mechanism dependent and can be obtained by making use of ligand mass balance. The polynomials used for correction are

$$K_A[L]^2 + (1 + [P_T]K_A - K_A[L_T])[L] - [L_T] = 0 \quad \text{[model (1)]} \quad (9)$$

$$K_A K_B[L]^3 + (K_A + [P_T]K_A K_B - [L_T]K_A K_B)[L]^2$$

$$+ (1 + [P_T]K_A - [L_T]K_A)[L] - [L_T] = 0 \quad \text{[model (2)]} \quad (10)$$

$$K_1[L]^{n+1} + ([P_T] - [L_T]K_1)[L]^n + [L] - [L_T] = 0 \quad \text{[model (3)]} (11)$$

Physically, only one value is possible for $[L]$, so the choice of right root is judiciously made by considering that

1. The root must be real and positive.
2. Its value cannot exceed $[L_T]$.

If many roots meet the above criteria, then the one that is closer to $[L_T]$ is chosen.

Automation using genetic algorithm

A common feature that is seen from the simple model fitting to the complex structure calculation is the optimization of the desired parameters using experimental data as constraints. For any such problem, proper definition of the target or the objective function is critical for steering the optimization towards global minimum. Different protocols are available to perform optimization, e.g. simulated annealing (SA), genetic algorithm (GA), simplex and Levenberg Marquardt algorithm (LVM), etc. To accelerate the convergence step for finding solutions, sometimes the gradient calculations are incorporated along with objective function, e.g. LVM. Instead, methods like SA and GA relies on random sampling of the entire parameter space to obtain the best combination. Here, GA has been implemented to optimize and determine the parameters for different fast chemical exchange models [21].

For a serial titration experiment with n different ligand concentrations, n chemical shift values will be obtained for each residue. The objective function for fitting to an appropriate model is defined as,

$$\text{Obj} = \text{Minimize}\left[\sum_{n'=1}^{n}\left(1 - \frac{\delta[n']_{calc}}{\delta[n']_{exp}}\right)^2\right] \quad (12)$$

Here $\delta[n']_{exp}$ and $\delta[n']_{calc}$ are the experimental and calculated chemical shift values. For the calculation of $\delta[n']_{calc}$, we can consider the model (6), having five parameters namely $\delta_{[P]}$, $\delta_{[PL_1]}$, $\delta_{[PL_2]}$, K_A and K_B to be optimized. Initially, random values for each parameter within the specified lower and upper bound values will be generated. These limits are automatically specified from the experimental data. With the generated parameters, the free ligand concentration $[L]$ will be calculated from $[L_T]$ using the equation (10). From the calculated $[L]$ and parameters, $\delta[n']_{calc}$ will be evaluated for each titrated ligand concentration. The objective

function is finally calculated from $\delta[n']_{exp}$ and $\delta[n']_{calc}$. This process is iterated several times till a convergent minimum value is obtained for objective function. Successful achievement of the global minimum depends primarily on setting the correct lower and upper limits for the parameters. The genetic algorithm based parameters like cross-over rate, mutation rate and number of generation also influence the quality of the fitting. Jack-knife algorithm has been incorporated for determination of standard error for parameters. The whole analysis is automated through an in-house written c program 'Auto-FACE' (Auto-FAst Chemical Exchange analyzer). 'Auto-FACE' is highly interactive, user friendly and portable to different platforms.

Mapping the binding site of BH3I-1 onto hBcl$_{XL}$

hBcl$_{XL}$ is a key member of the anti-apoptotic Bcl2 family of proteins [22,23]. It is up regulated in different types of cancers and confers cancer cells its resistance to normal apoptotic signal [24]. Targeting and inhibiting hBcl$_{XL}$ is one of the therapeutic strategies in treating recalcitrant cancer [25]. BH3I-1 on the other hand is a small ligand (400.31 Da) that has been identified to bind to the BH3 binding groove of hBcl$_{XL}$ (Figure 1B). Similar to BH3I-1, the structural analogue BH3I-2 can also displace Bak peptide from the hydrophobic groove of hBcl$_{XL}$ (Figure 1C). The results of the fluorescence polarization assays (FPA) suggest that the weakly interacting BH3I-1 (7.8 µM KDa) can displace the strongly bound Bak peptide (16 residues; $K_d = 0.34$ µM). The mechanism could be more complex than a simple competitive displacement [26]. Previous studies carried out with BH3I-2 (an analogue of BH3I-1) and hBcl$_{XL}$ generated a differential pattern in HSQC perturbation for a single substitution of Br group to H [27]. Residues like N136, G138, I140, A142, F146, G147, G148 and R91 were differentially perturbed and were identified to be the binding site residues [27]. In the current analysis, we have used hBcl$_{XL}$ and BH3I-1 as a standard system to validate our automated analysis program.

Results of ITC titration

To confirm the interaction of BH3I-1 with hBcl$_{XL}$, ITC titration was performed. The isothermal binding curve fitted well to the three sites sequential binding model with good statistics for parameters (Figure 3 and Table 1). A closer look at the equilibrium constants for all three processes revealed that the last event could merely be a non-specific allosteric change rather than an actual binding process. This is evident from its lower K_{eq} value (10^3) and much higher ΔS value (34 $cals$). A recent comparative work on thermodynamics of protein-ligand interaction shows that ΔH is more correlated with the binding process than ΔS [28,29]. Considering the possibility that the third process might not be significant, the global interaction mechanism could be primarily dictated by the first two enthalpy dominant processes.

NMR titration

BH3I-1 was serially titrated into hBcl$_{XL}$ at increasing ligand concentrations and ^{15}N HSQC spectra were recorded. On overlaying the spectra, more than half the peaks exhibited 'peak walking' pattern characteristic of fast exchange (Figure 4A). Compared to the rest of the residues, stronger perturbations were observed for residues like F146, G148, G94 and G196 (Figure 4B). Structurally, F146 and G148 are 10 Å away from the latter residues G94 and G196. We proceeded with the detailed analysis on chemical shift by fitting the data against binding models like single site, two site sequential, multiple sites simultaneous and single site with allosteric contribution models to interpret the binding
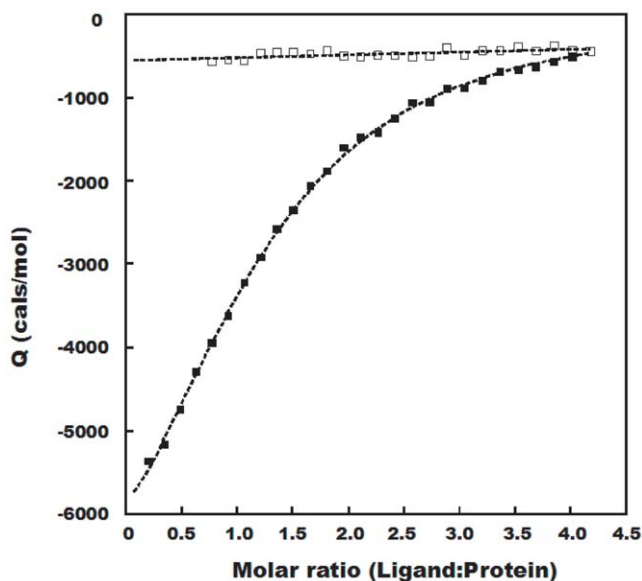
Figure 3. Isothermal binding curve for BH3I-1 titrated into hBcl$_{XL}$.
(□) : Blank experiment where 1 mM of BH3I-1 was titrated into 20 mM phosphate buffer. (■) : 1mM of BH3I-1 was titrated into 25 μM hBcl$_{XL}$. In all buffer solutions, concentration of DMSO was adjusted to 2.5%.
doi:10.1371/journal.pone.0008943.g003

mechanism. Almost all residues fitted well to the single site model and a few remaining ones were represented better with two site sequential model. F-test and Akaike's criteria were used to choose the best simpler model statistically (Figure 5 and Table 2) [30–33].

Binding site analysis using NMR based parameters

The mapping of binding site was carried out using the following parameters,

i)   Binding equilibrium constant
ii) Initial rate of perturbation
iii) Magnitude of the perturbation

Among these parameters, the last two can be either calculated from model equations and fitted parameters or obtained directly from experimental data. For further analysis, a detailed consideration on the fundamental differences between H and N chemical shift is required for correct interpretation of data. The chemical shift calculated from protein structures and quantum mechanical treatments by semi-empirical and *ab initio* methods shows that several factors contribute to the chemical shift value in an additive manner [34,35]. For H resonances, primary contribution comes

Table 1. Thermodynamic parameters obtained from ITC experiment by fitting the data to sequential three site binding model.

Three sequential binding model $\chi^2 = 559.097$		
K$_{Equilibrium}$	Δ H (cals)	Δ S (cals)
$4.869 \times 10^5 \pm 4.412 \times 10^4$	$-6.81 \times 10^3 \pm 0.71 \times 10^2$	3.182
$9.322 \times 10^4 \pm 6.632 \times 10^3$	$-2.387 \times 10^3 \pm 0.88 \times 10^2$	14.73
$1.584 \times 10^3 \pm 1.117 \times 10^2$	$-1.464 \times 10^4 \pm 0.72 \times 10^3$	34.46

doi:10.1371/journal.pone.0008943.t001

from ring-current effect, magnetic-anisotropic effect, electric field effect, and the length and orientation of hydrogen bond [36]. Whereas N resonances are strongly influenced by the side chain conformation of the preceding residue ($i-1$). Hence, backbone torsion angles ($\psi-1$, ϕ) and side chain chi angle ($\chi-1$) are the major contributing components [37]. In a perturbation setting like protein-ligand interaction, H resonances can be interpreted unambiguously as ring-current effect of the ligand itself contributes directly to H shift. But for N shifts, complication arises due to the convoluted contribution from ligand and structural changes. Our present analysis considers both H and N shifts with an underlying assumption that allosteric structural changes are minimum at lower ligand concentrations and the major contribution comes from the direct interaction of ligand with protein. Taking the first derivative of the equation (5) with respect to $[L]$ relates $\frac{d\delta}{dt} \propto \frac{1}{[L]^2}$, which implies that at lower ligand concentration the rate of change of δ will be larger. But at higher concentration of $[L]$, the slope decreases parabolically. Thus the more sensitive information content is encapsulated in the initial perturbation data rather than at later stages of titration. The initial perturbation data at lower ligand concentrations also circumvents non-specific interactions and allosteric structural changes that are more likely to occur at higher ligand concentrations. For example, a recently proposed mechanism for cyclic AMP receptor protein (CRP) and cAMP association suggests that two independent binding processes preceds a subsequent three step conformational changes. In this case, if more emphasis is given to the data content at initial stages, where binding process dominates, the effect of non-specific perturbations caused by conformational changes can be eliminated [38].
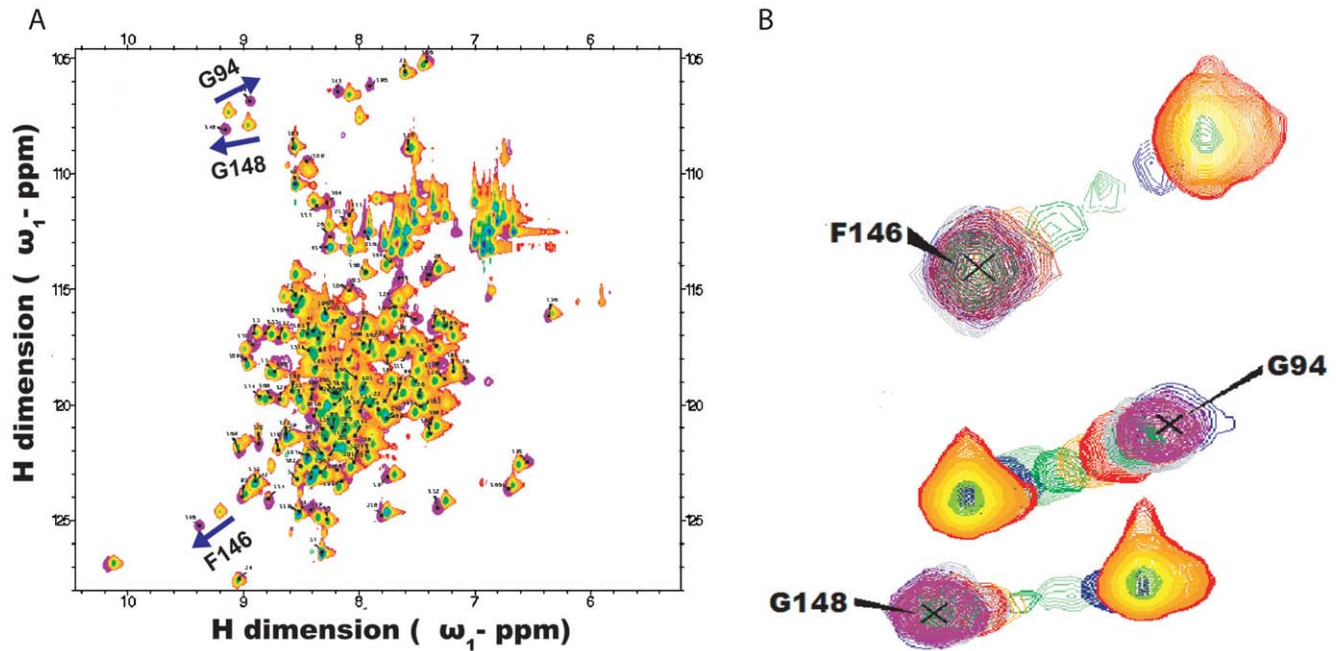
A 3D graphical plot of the listed parameters greatly assists in identifying the binding site residues (Figure 6). The initial perturbation rate, as explained above, is more sensitive in distinguishing the critical binding site residues from the bulk residues (6). On the other hand, binding equilibrium constant and magnitude of perturbation are also correlated with the binding process but influenced by non-specific interactions as well. Hence, these parameters are used in later stages only to refine the residues selected based on initial rate of perturbation. Appropriate threshold levels are set for each parameter statistically or manually. For initial rate of perturbation, $|0.01|$ and $|0.5|$ ppm/mM corresponding to 1.0 σ value was set for H and N resonances, respectively. Only perturbations greater than $|0.01|$ and $|0.1|$ ppm for H and N resonances were considered. Threshold for equilibrium constants was based on median analysis. The values falling within 0.15 and 0.7 quartiles were selected for both H and N resonances.

Residues like G94, E96, Q111, L112, V126, E129, F143, F146, G147, G148, V192 and G196 from the H plot and residues like L90, L99, Q111 and I114 from the N plot were mapped onto the structure of hBcl$_{XL}$ (Figure 7A, Figure S1 & Table 2). Two distinct regions that are adjacent to each other but separated by a minimum distance of 10 $\mathring{A}$ were identified. The first site (A) is located at the edge of the extended hydrophobic BH3 groove near the 'C' terminal region. Residues like G94, E96, L99, V192 and G196 that constitute this site are part of the BH3 domain. The second site (B) is located at the middle of the highly conserved but less exposed hydrophobic groove. Residues like Q111, L112, V126, E129, F143, F146, G147 and G148 that spans the BH3 binding groove are proximally distributed within BH1 and BH2 domain. As mentioned above, the perturbation at saturation limit may or may not be directly related to the binding process. This is evident from residues like F27 and K157 that are not at the binding site, as confirmed by the slope values of 0.035 and 0.014 ppm/mM, but have high perturbation values of 0.211 and 0.346 ppm. This implies that

Figure 4. Generalized ^{15}N HSQC perturbation observed for all residues. (A) Overlaid perturbation spectrum for all residues and (B) Selected residues with significant "peak walking" chemical shift perturbation. Reddish-orange contour represents protein alone spectrum and blue contour represents the spectrum of protein with maximum titrated ligand concentration. The overlaid spectra of gradually titrated ligand concentrations are shown in blue, magenta, green, orange, red, grey and pink contours ranging from 0.133 mM to 1.177 mM of BH3I-1. F146, G148 reaches saturation at the protein to ligand ratio of 1:1, whereas saturation could not be reached for G94.
doi:10.1371/journal.pone.0008943.g004

mapping binding site using perturbation alone could be misleading in complex protein-ligand interactions.

J surface mapping

To localize the binding site, we have also performed J-surface mapping using the same perturbation data. In principle, the ring current effect of the aromatic ligand causes strong perturbation of amide protons present adjacent to it [39,40]. The electron density map calculated for the ligand from the sign and magnitude of perturbation could locate the position of the ligand at the binding pocket. Since BH3I-1 contains an aromatic ring, J surface map could be calculated at all titrated ligand concentrations (Figure 7B & C). At lower ligand concentrations, the J-surface map is localized near the central helix α5, where residues like L90, G94, D95, F97 and V141 are located (site A). But at higher ligand concentrations, the J-surface mapping converged to a region where residues like F143, F146 and G147 are located (site B). The latter site is completely buried and inaccessible to the ligand in the closed conformation of hBcl$_{XL}$.

Binding mechanism

In order to get a quantitative sense of the interaction, the equilibrium constants for the two distinctly mapped regions were geometrically averaged from the individual residues flanking these sites. The equilibrium constants averaged to 2.970×10^4 and 0.775×10^4 for site B and A, respectively. The affinity of site B is 3.8 times stronger than site A. When the results of J surface mapping are also considered, we propose that site A is a weaker site where BH3I-1 makes its first contact with the protein. Because of its dynamical nature [41], this interaction consequently lead to the exposure of the hydrophobic groove for the more critical interaction of BH3I-1 with site B to occur. The consistency in the site predicted by our chemical shift analysis, J surface mapping and the stoichiometry suggested by

ITC all points to the possibility of a complex sequential binding mechanism. This also explains why a small ligand with weak affinity like BH3I-1 can displace the Bak peptide that binds strongly to hBcl$_{XL}$ [26]. Further, more mutation studies with L130A, R139A and R100E suggests that these residues are crucial for BH3I-1 interaction and notably, the first two residues are present at site B and the last one near site A [26].

NMR model fitting suggests single site model to be appropriate and good enough for residues present at site A and site B, this is in contrast to the two site model as suggested by the global interaction mechanism. The inconsistency can be explained by making a valid assumption that chemical shifts are highly dependent on local environment and its perturbation also reflects the same. In this regard, the residues located at site A and B fit well to single site models, but the residues in between these sites, influenced by both the binding processes, would require a two site binding model to explain its behavior. From our analysis, one such residue G148, was found to be represented best with two site model (Figure 5). (Though the model selection is performed based on the values of F-test and Akaike's criteria as mentioned in Table 2, a closer look at the fitted graph suggests that the model 2 agrees well with the experimental data with better Chi-square value (10^{-5} compared to 10^{-4}). Hence we choose model 2 for explaining the behavior of residue G148). Thus NMR titration data, unlike ITC titration data, pictures the local binding mechanism much more accurately than the global binding mechanism.

Docking results

Docking performed with perturbation differences between BH3I-1 and BH3I-2 as constraints resulted in the model as shown in Figure 8A [27]. In our case, initial blind docking resulted in majority of the ligand conformations (80%) docked to site B. The BH3I-1 oriented itself with its phenyl ring buried deeply inside the hydrophobic pocket of site B, making close contacts with L130,
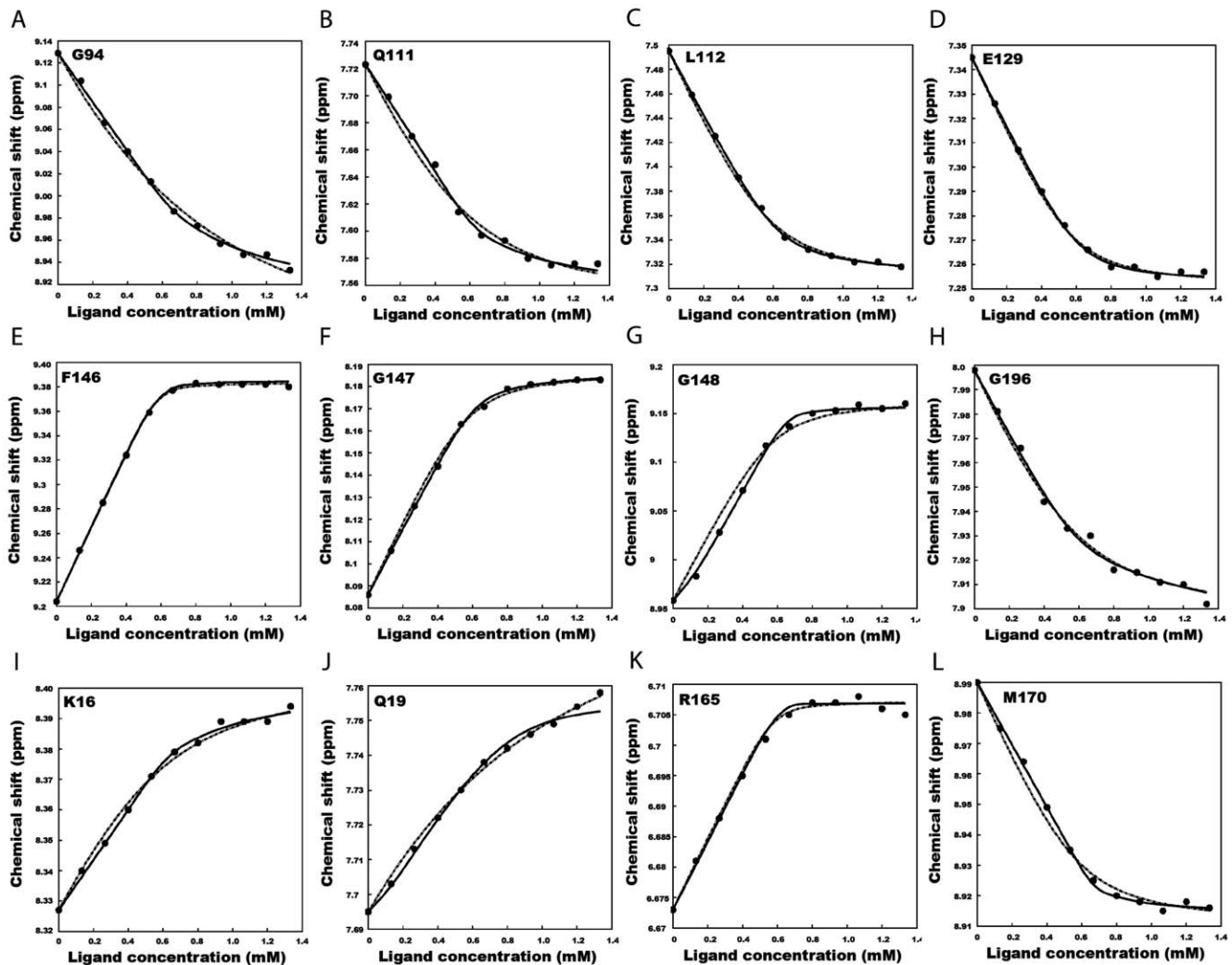
Figure 5. Comparison of single and double site binding models for different residues. Comparison of two different models for residues present at the binding site (A–H) and the non-binding site (I–L). ●●● : Experimental data, ···· : Single site model, – : Two site sequential model.
doi:10.1371/journal.pone.0008943.g005

F146 and A149 (Figure 8B). As blind docking resulted only limited conformations of BH3I-1 at site A, a constrained docking was performed for site A, with the docking grid confined to NMR perturbed residues at this site. In this docked conformation, the phenyl ring of BH3I-1 is partially exposed in a shallow groove, which suggest a weaker interaction for this site.

Conclusions

The chemical shift perturbation contains not only the qualitative details but also the quantitative information on the local environment of an atom, which can be reliably obtained if detailed model based analysis is carried out. With detailed analysis of chemical shift perturbation of the protein-ligand system of hBcl$_{XL}$ and BH3I-1, we have arrived at a conclusion that NMR data, unlike ITC data, contains interaction details at local level rather than at global level. This paves a way to study interactions of each individual atom quantitatively. Further, the initial perturbation data contains more information on binding process compared to data obtained at later stages of titration. By following the dynamic aspect of perturbation, i.e. the rate of change in perturbation at lower ligand concentrations, rather than the overall magnitude of chemical shift perturbation, we

can distinguish the binding site residues from the allosterically perturbed residues. The approach that has been adopted and implemented in 'Auto-FACE' is suitable for simple to complex protein-ligand interactions, particularly mechanisms that involve allosteric structural changes in addition to binding process. 'Auto-FACE' is more useful in distinguishing the binding site residues from the large number of perturbed residues, which resulted because of combined binding and allosteric effects. If only a few residues are perturbed, 'Auto-FACE' would not be required as the perturbed residues must be coming from the binding site residues. Additionally, when the stoichiometry of protein to ligand is more than 1:1, analysis has to take into account of the sequential or simultaneous nature of interaction in addition to correction for free ligand concetration. In such cases, 'Auto-FACE' would be much useful in analyzing the data automatically with minimal user input.

Materials and Methods

Protein expression

The DNA sequence of human hBcl$_{XL}$ starting from residues M1 to M218, with a flexible loop region R45 to A84 being deleted,

Table 2. Parameters determined by fitting of chemical shifts to equations (6) & (7) for hBcl$_{XL}$ and BH3I-1 system.

Residue. no.	δ_0 (ppm)	δ_{max} (ppm)	Equilibrium constant (K_A or K_B)	Slope (ppm/mM)	$|\Delta CS|_{Calc}$ (ppm)	F - statistics (Standard: 5.049)	Akaike's criteria (AC)
Binding site residues							
'H' resonances							
G148	8.958	9.165	$3.1290 \pm 0.2716 \times 10^4$	0.240	0.207	1 (1.311)	1 (−11.89/−9.32)
	8.958	8.959	$0.7709 \pm 0.0012 \times 10^4$				
		9.157	$25.6178 \pm 0.0139 \times 10^4$				(model 2)
F146	9.204	9.383	$24.1123 \pm 4.0886 \times 10^4$	0.224	0.180	1 (1.871)	1 (−12.25/−9.27)
L112	7.495	7.306	$1.7979 \pm 0.1785 \times 10^4$	0.204	0.189	1 (2.716)	1 (−17.05/−12.69)
G94	9.129	8.828	$0.2056 \pm 0.0097 \times 10^4$	0.195	0.300	1 (2.329)	1 (−12.28/−8.78)
V192	7.870	7.658	$0.4296 \pm 0.0344 \times 10^4$	0.180	0.211	1 (0.539)	1 (−5.86/−3.66)
Q111	7.723	7.542	$0.6775 \pm 0.0577 \times 10^4$	0.163	0.181	1 (3.304)	1 (−10.24/−8.98)
E96	8.293	8.444	$4.4784 \pm 1.6428 \times 10^4$	0.162	0.151	2 (118.500)	2 (−8.12/−8.22)
V126	7.172	7.074	$10.5379 \pm 2.6728 \times 10^4$	0.120	0.098	1 (0.239)	1 (−10.46/−8.53)
F143	8.111	8.217	$2.1494 \pm 0.3548 \times 10^4$	0.119	0.106	1 (2.837)	1 (−13.67/−8.81)
G147	8.086	8.817	$3.1401 \pm 0.6247 \times 10^4$	0.116	0.101	1 (2.429)	1 (−16.71/−13.05
E129	7.345	7.250	$2.7540 \pm 0.5867 \times 10^4$	0.108	0.095	1 (1.039)	1 (−12.55/−10.85)
G196	7.998	7.895	$0.9103 \pm 0.1505 \times 10^4$	0.102	0.103	1 (0.536)	1 (−10.96/−9.13)
'N' resonances							
I114	122.620	124.677	$0.3363 \pm 0.0024 \times 10^4$	1.98	2.06	1 (0.303)	1 (−18.59/−14.89)
Q111	118.894	117.731	$3.2953 \pm 0.0555 \times 10^4$	1.3	1.16	1 (2.554)	1 (−10.24/−8.98)
L99	117.247	116.476	$545.5641 \pm 0.7442 \times 10^4$				
		116.173	$0.3046 \pm 0.0006 \times 10^4$	1.11	1.08	2 (143.581)	2 (−7.99/−9.88)
L90	120.121	119.009	$0.3163 \pm 0.0043 \times 10^4$	0.81	1.11	1 (1.447)	1 (−11.17/−10.78)
F144	120.848	121.301	$10.9532 \pm 0.6110 \times 10^4$	0.53	0.45	1 (0.781)	1 (−13.91/−12.79)
Non−binding site residues							
'H' resonances							
M170	8.990	8.931	$487.0460 \pm 21.5370 \times 10^4$				
		8.914	$0.7460 \pm 0.9310 \times 10^4$	0.088	0.080	2 (927.198)	2 (−14.63/−18.37)
K16	8.327	8.405	$0.6013 \pm 0.1229 \times 10^4$	0.069	0.078	1 (0.827)	1 (−13.15/−10.83)
N136	6.327	6.376	$0.3091 \pm 0.0909 \times 10^4$	0.039	0.050	1 (0.618)	1 (−15.28/−13.04)
F27	7.317	7.527	$0.0221 \pm 0.0003 \times 10^4$	0.035	0.211	1 (2.229)	1 (−11.65/−8.29)
T41	8.249	8.344	$0.0549 \pm 0.0102 \times 10^4$	0.031	0.096	1 (2.771)	1 (−14.60/−10.02)
S2	8.374	8.394	$3.4162 \pm 0.0358 \times 10^4$	0.024	0.021	1 (1.463)	1 (−12.94/−11.33)
V86	7.528	7.502	$0.3209 \pm 0.1847 \times 10^4$	0.020	0.025	1 (2.766)	1 (−15.02/−10.47)
K157	7.696	7.349	$0.0015 \pm 0.0001 \times 10^4$	0.014	0.346		(model 1)
	7.696	7.695	$446.112 \pm 0.7390 \times 10^4$				
		7.569	$0.0097 \pm 0.0001 \times 10^4$			2 (221.858)	2 (−11.91/−14.23)

Most of the residues fitted well to single site binding model with a few residues fitted better to two site sequential model. Initial perturbation rate (slope), binding affinity and magnitude of perturbation ($|\Delta CS|_{Calc}$) are listed below. For model selection, F-test and Akaike's criteria based analysis were performed. $F_{critical}$ value for F-test is chosen based on the degree of freedom (the number of parameters present in the model, i.e. 3 and 5 for single and two site, respectively) and level of significance ($\alpha = 0.05$). The $F_{Standard}$ or $F_{critical}$ was found to be 5.049. $F_{calc} > F_{critical}$ suggests simpler model (i.e. single site) cannot explain the experimental data satisfactorily. The null hypothesis is rejected and the complex two site binding model is chosen. Akaike's information criteria was calculated for both models, the model with lower AC value represents the data better. The non-binding site residues are represented by the residues that are $>10\text{\AA}$ away from the conventional BH3 binding groove residues.
doi:10.1371/journal.pone.0008943.t002

was subcloned into a modified pET-32a (Novagen) vector which lacks *S*-tag and thioredoxin genes. The plasmid was transformed into *E. coli* BL21(DE3) strain and the His tagged protein was expressed at 37°C. IPTG was added to a final concentration of 0.4 mM when the optical density of cells reached 0.6 (measured at 600 nm). The culture was allowed to grow at the same temperature for another 8 hours before the cells were harvested. The bacterial culture was centrifuged at $6,891 \times g$ and the pellet was collected and sonicated. The suspension was clarified by centrifugation at $26,581 \times g$ at 4°C. The supernatant was taken and passed through Ni–NTA agarose column (Qiagen) and washed thoroughly with wash buffer (20 mM of Tris, pH 7.9
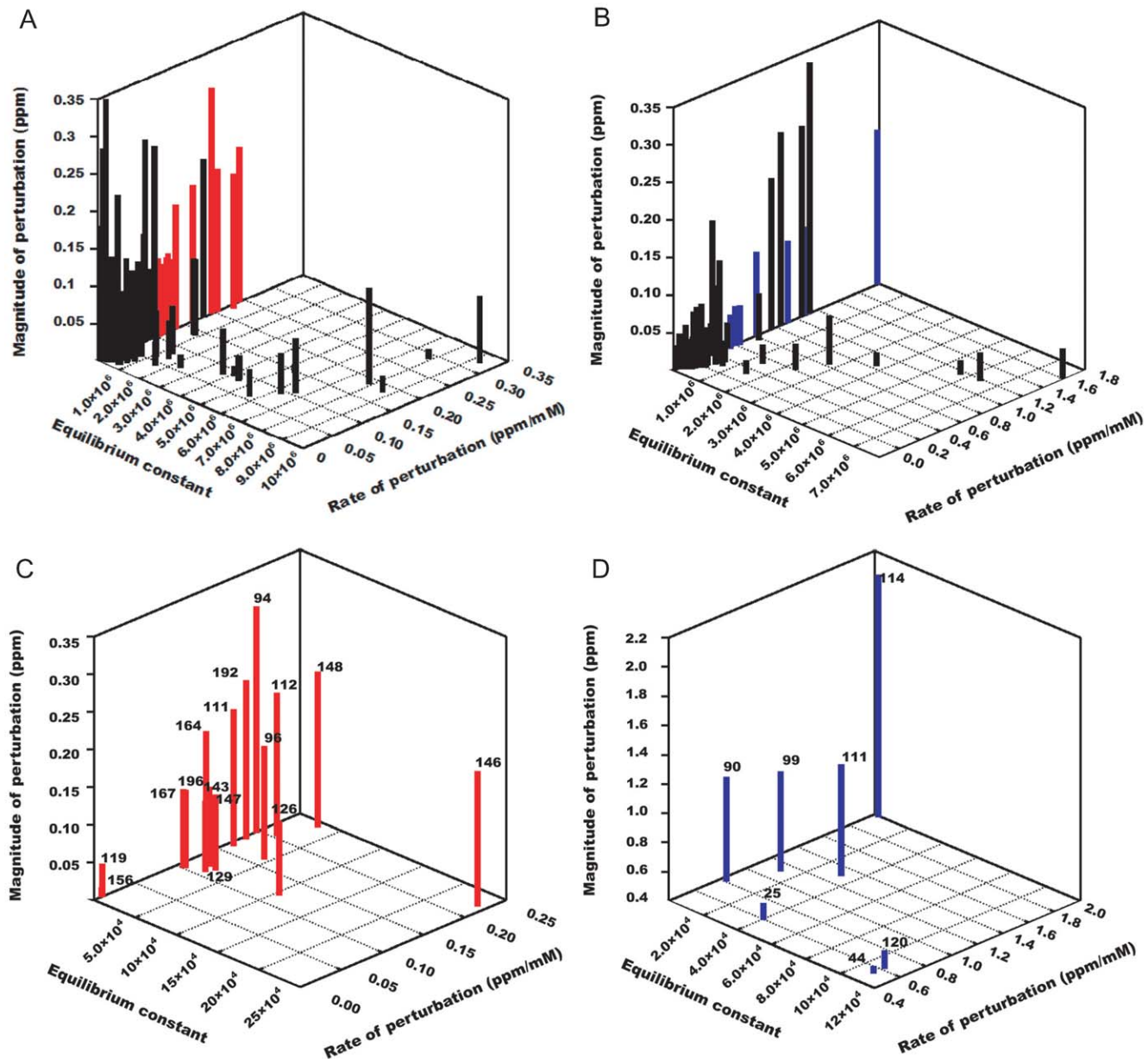
Figure 6. '3D' plot to differentiate the binding site residues from bulk residues. (A) and (B) are plots for H and N resonances, with no threshold set for slope and magnitude of perturbation. (C) and (D) are plots for H and N resonances, with threshold set at $\sigma = 1.00$ which corresponds to 0.01 and 0.5 ppm/mM for slope values of H and N residues and to $|0.01|$ and $|0.1|$ ppm for magnitude of perturbation of H and N residues. For both plots, equilibrium constants falling within 0.15 to 0.7 percentile were used.
doi:10.1371/journal.pone.0008943.g006

containing 30 mM of imidazole and 0.5M sodium chloride) before eluted with wash buffer containing 0.5 M of imidazole. The eluate was dialyzed against 50 mM Tris pH 7.9 overnight at 4°C. The dialysed protein was concentrated to 4 mL. Thrombin and calcium chloride were added to a final concentrations of 3 units/mg of protein and 3 mM, respectively, to cleave the His tag. After digestion, $hBcl_{XL}$ was purified further on a superdex 75 prep grade column (GE Healthcare) using 50 mM Tris pH 7.9 buffer containing 0.5 M sodium chloride and with a flow rate of 1ml/ min. Finally, the purified fractions containing $hBcl_{XL}$ were pooled together and dialyzed against 20 mM phosphate buffer at pH 7.0. NMR sample was prepared by concentrating the above sample to 0.6 mM using centrifugal concentrator with a membrane cutoff of

5 kDa (Viva-spin 20, Sartorius). For preparation of ^{15}N labeled sample, the protein was expressed in M9 minimal media containing ^{15}N ammonium chloride as the sole nitrogen source, while LB medium was used for preparing the unlabeled samples.

ITC titration

4 mL of 25 μM of $hBcl_{XL}$ and 0.8 mL of 1 mM BH3I-1 were prepared in 20 mM of phosphate buffer pH 7.0 containing 2.5% DMSO and degassed under vacuum for 20 minutes. In the reference cell, 20 mM of phosphate buffer at pH 7.0 and containing 2.5% DMSO was used. 0.3 mL of BH3I-1 was titrated into 1.2 mL of $hBcl_{XL}$ at 25°C over 28 injections of 10 μL each.
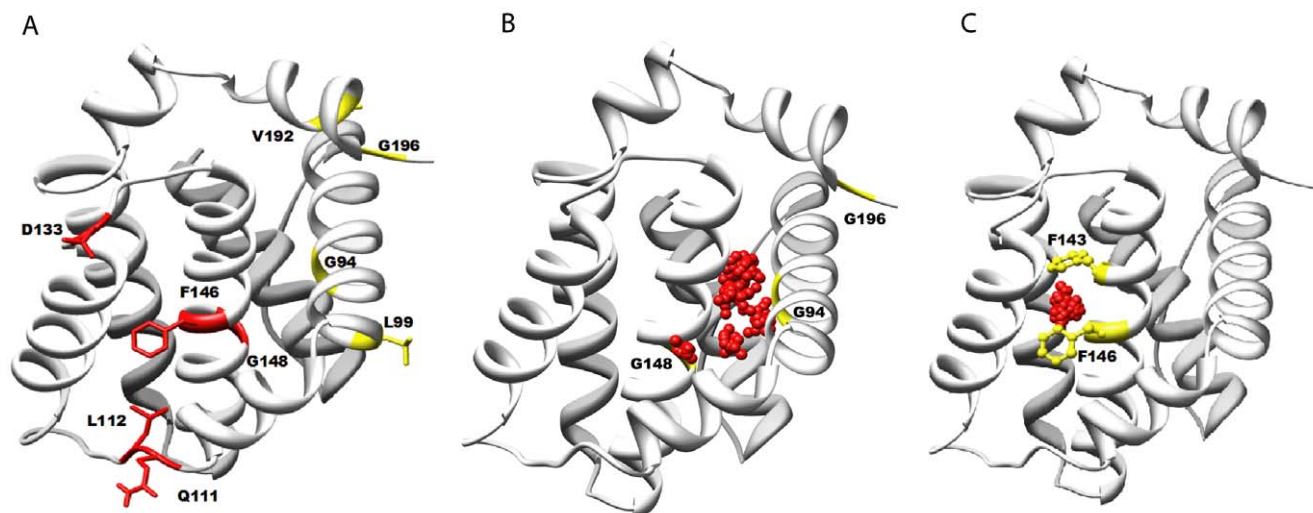
Figure 7. Mapping of the unique residues identified from '3D' plot onto the structure of hBcl$_{XL}$ and comparison with J-surface mapping. (A) Two distinct regions are shown which are colored differently (red, yellow); (B) & (C) are the J-surface mapping of BH3I-1 at lower (P:L::1:0.229) and higher (P:L::1:0.918) ligand concentrations, respectively. Each red dot represents the possible location of the centroid of the aromatic ring of BH3I-1. The collection of dots suggests that the aromatic ring could be anywhere in that mapped region. The initial map appears diffused covering G94, G196, G148 residues but slowly converges near F143 and F146 as the concentration of ligand increases. J-surface map were calculated using JSURF program considering perturbations $> |0.05|$ ppm. Other parameters like *ms* (standard deviation for data spread), *np* (number of random points to fill the sphere) and *ar* (an offset in $\mathring{A}$ added to radius of sphere) were set at 3, 2000 and 1, respectively. All the figures were made using the software Chimera [54].
doi:10.1371/journal.pone.0008943.g007

Blank experiment was performed by titrating BH3I-1 into sample cell containing 1.2 mL of buffer alone. Buffer alone was titrated into protein sample to confirm that the heat of protein dilution was negligible. The isothermal chromatogram was integrated and analyzed using the commercial software Origin 5.0.

^{15}N HSQC titration

20 μL of 40 mM of BH3I-1 in D$_6$ DMSO was titrated serially into 550 μL of 0.58 mM ^{15}N labeled hBcl$_{XL}$. The ^{15}N HSQC spectra were recorded at 25°C for different protein to ligand ratios of 1:0.23, 1:0.46, 1:0.69, 1:0.92, 1:1.15, 1:1.14, 1:1.61, 1:1.82, 1:2.07 and 1:2.30. The data was acquired with a resolution of 2048×128 points in the direct and indirect dimensions. Eight scans were accumulated for each titration. The obtained spectra were processed with NMRPipe 9 [42,43]. using the following parameters. Solvent and polynomial baseline corrections were done with an auto flag. The data was padded with zeros to twice its size in both dimensions to increase the digital resolution of peaks. Apodization using phase shifted sine bell function ($\theta = 90°$) of order one was performed for the acquired dimension and of order two for the indirect dimension. Linear prediction was done for the indirect dimension before apodization. The phase corrected spectrum was assigned using Sparky 3.114 and resonance lists were generated for all spectra [44].

J-Surface mapping

J-Surface mapping requires ^{15}N HSQC titration data and PDB coordinates of the protein. "jsurf" module written by McCoy and G. Moyna was integrated with an in house written program to automate and analyze all the serially titrated data. The coordinates of all the amide protons were sorted from the PDB file, and the chemical shift perturbation, $\Delta CS = CS_{protein} - CS_{protein+ligand}$, for the corresponding protons were determined from the sparky assignment files. Electron density map was calculated from the magnitude and direction ($\pm$) of perturbation values. The region showing higher 'j' density was identified to be the binding site for ligand.

Molecular docking

Automated docking was performed using Autodocksuite-4.0.1 [45]. The coordinates of complexed (1BXL and 2YXJ) and free (1LXL) hBcl$_{XL}$ were obtained from the protein database [46]. Structures of (R, S) BH3I-1 were generated in SYBYL-7.0 and atom types were assigned with considerations for stereo-specificity. Prior to docking, protons and charges were added to protein and ligand structures using MGLTools-1.5.2 [47]. For BH3I-1, the number of rotatable bonds were set to 4 and docking was performed with Lamarckian-Genetic algorithm. The variable resolution was set at 250 (population size) and energy evaluation was performed for 25×10^5 conformations per run. 100 such runs were performed. Ligand conformations within 1 RMSD difference were clustered together. Unlike blind docking, where the docking grid covered the whole protein, constrained docking was performed with the grid confined to NMR perturbed residues.

Automated data analysis

The resonance list file generated by 'Sparky' is used as an input to our in house written 'c' program (Auto-FACE). The software and its manual are freely available on request. Curve fitting for different models were performed for individual residues and the parameters with its standard error were written in separate files. Using binding affinity, initial rate of perturbation and magnitude of perturbation, binding site analysis was performed and '3D' plots were generated for H and N resonances. The quality of the plot depends on the threshold set for each of these parameters. Except the affinity constant, which is obtained only by model fitting, the other parameters can either be calculated or obtained from experimental data.

The number of binding constants depends on the models used, e.g. equation (10) has two equilibrium constants, K_A and K_B. For analysis, either an individual binding constant (K_A or K_B) or geometrically averaged value (K_A and K_B) can be used. If the data is poor, the model fitting may fail for some residues and result in excessively high or low values for equilibrium constant. Such
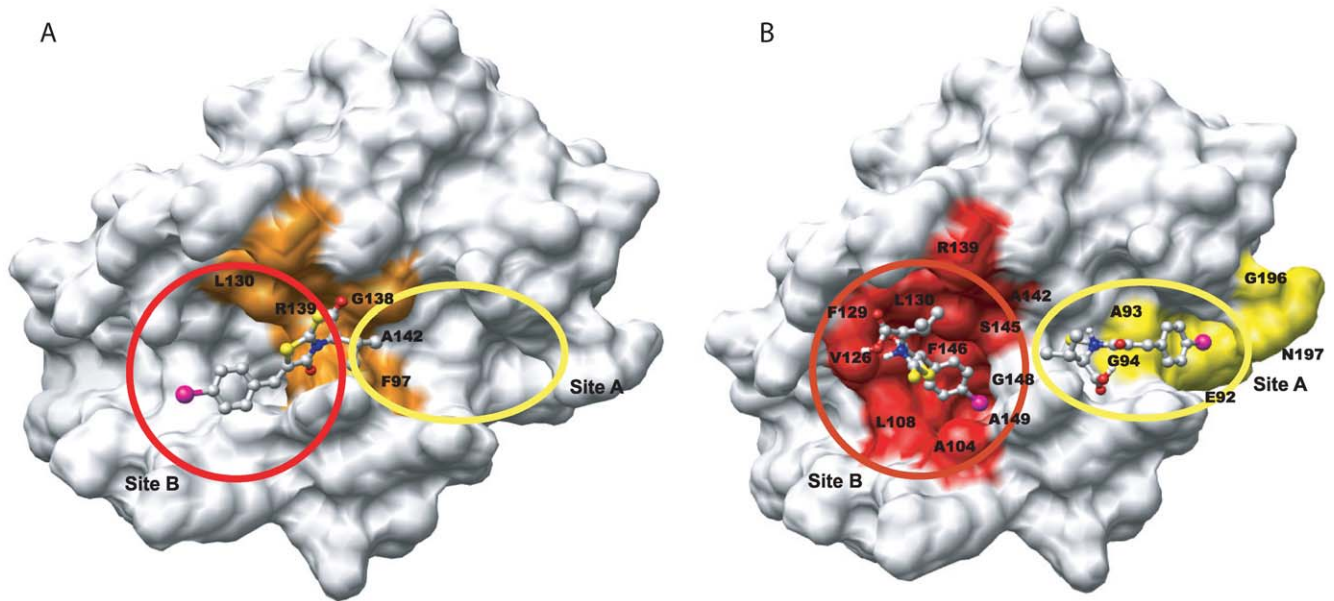
Figure 8. Comparison of the previous and current docked models of BH3I-1 on to hBcl$_{XL}$. (A) and (B) compares the published and current docked models of BH3I-1 on hBcl$_{XL}$, respectively. In the published model, the stoichiometry was constrained to a single site, so the ligand preferred the site in between the two adjacent pockets. The key residues that interact with BH3I-1 within 5 $\mathring{A}$ are highlighted in orange. In the current model, two BH3I-1 molecules bind adjacent to each other with distinctive affinities. Site A and B are circled and highlighted in yellow and red color, respectively.
doi:10.1371/journal.pone.0008943.g008

artifacts can be eliminated by median analysis. The user can specify the upper and lower quartile values for residue selection.

The initial rate of perturbation is calculated using the following expression,

$$\overline{S} = \text{mean}(S_1, S_2, \dots, S_n) \quad \text{where } S_{n'} = \left| \frac{\delta_{n'+1} - \delta_{n'}}{[L_{n'+1}] - [L_{n'}]} \right| \quad (13)$$

$L_{n'}$ and $L_{n'+1}$ are the n'^{th} and its subsequent higher ligand concentration; $\delta_{n'}$ and $\delta_{n'+1}$ are the corresponding chemical shift values. The average of the rates of first few data points well below the half saturation limit was used for binding site analysis. Statistically, the slope values exhibited a normal distribution. In this regard, most of the bulk residues would have their slope value centered near the mean ($\overline{x}$) and the binding site residues having large slopes would be present outside ± 1 or 2 σ (standard deviation). The residues were selected depending on stringency of σ and the threshold.

The magnitude of perturbation is the absolute difference between the chemical shift of free protein and ligand complexed protein. User can define threshold in terms of ppm for 'H' and 'N' resonances. Final '3D' plots would be generated using the software 'gnuplot' [48]. Interested users can download the 'Auto-FACE' program along with its manual and source code from http://www. dbs.nus.edu.sg/staff/henry.htm.

Deriving complex models

The derivation of two site sequential binding is explained below.

$$[P] + [L] \underset{k_{-1}}{\overset{k_{+1}}{\rightleftharpoons}} [PL_1] + [L] \underset{k_{-2}}{\overset{k_{+2}}{\rightleftharpoons}} [PL_2]$$
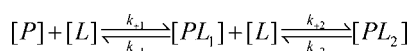
In this mechanism, the protein exists as $[P]$, $[PL_1]$ and $[PL_2]$ in solution. The averaged chemical shift is

$$\delta = \delta_{[P]} f_{[P]} + \delta_{[PL_1]} f_{[PL_1]} + \delta_{[PL_2]} f_{[PL_2]} \quad (14)$$

δ and f refers to the chemical shift and mole fraction of the appropriately subscripted molecular species i.e. free or bound form. The mole fractions $f_{[P]}, f_{[PL_1]}$ and $f_{[PL_2]}$ can be expressed in terms of ligand concentration assuming equilibrium for the system. Here, a more general approach of framing differential equations for each exchanging species is adopted.

$$\frac{d[P]}{dt} = -k_{+1}[P][L] + k_{-1}[PL_1] \quad (15)$$

$$\frac{d[PL_1]}{dt} = +k_{+1}[P][L] - k_{-1}[PL_1] + k_{-2}[PL_2] - k_{+2}[PL_1][L] \quad (16)$$

$$\frac{d[PL_2]}{dt} = +k_{+2}[PL_1][L] - k_{-2}[PL_2] \quad (17)$$

The terms on R.H.S are constituted by multiplying the rate constant with its corresponding reactant. The sign indicates whether a particular rate increases (+) or decreases (−) the concentration of the considered species. At equilibrium, the above equations are equated to zero as concentration of $[P]$, $[PL_1]$ and $[PL_2]$ will not vary with respect to time.

$$\frac{d[P]}{dt} = 0 \quad \text{gives} \quad k_{+1}[P][L] = k_{-1}[PL_1]$$

$$\frac{d[PL_1]}{dt} = 0 \quad \text{gives} \quad (k_{-1} + k_{+2}[L])[PL_1] = k_{+1}[P][L] + k_{-2}[PL_2]$$

$$\frac{d[PL_2]}{dt} = 0 \quad \text{gives} \quad k_{-2}[PL_2] = k_{+2}[PL_1][L]$$

In fact, the above relations can also be obtained from conventional equilibrium assumption. But when time dependent analysis is required, e.g. non-steady state systems, the above simultaneous differential equations have to be solved analytically to obtain the mole fractions. The resulting expression for chemical shift would then depend not only on ligand concentration but also on time [49–52]. On rearranging the above equations, $[PL_1]$ and $[PL_2]$ can be expressed in terms of $[P]$ as follows,

$$[PL_1] = K_A[P][L] \qquad (18)$$

$$[PL_2] = K_A K_B[P][L]^2 \qquad (19)$$

where $K_A = \dfrac{k_{+1}}{k_{-1}}$ and $K_B = \dfrac{k_{+2}}{k_{-2}}$. Since total protein $[P_T]$ is equal to the sum of free as well as bound forms,

$$[P_T] = [P] + [PL_1] + [PL_2] \qquad (20)$$

Substituting (18), (19) into (20) gives expression for $[P]$, $[PL_1]$ and $[PL_2]$ in terms of $[P_T]$ and $[L]$.

$$[P] = \frac{[P_T]}{1 + K_A[L] + K_A K_B[L]^2} \qquad (21)$$

$$[PL_1] = \frac{[P_T]K_A[L]}{1 + K_A[L] + K_A K_B[L]^2} \qquad (22)$$

$$[PL_2] = \frac{[P_T]K_A K_B[L]^2}{1 + K_A[L] + K_A K_B[L]^2} \qquad (23)$$

Substituting $[P]$, $[PL_1]$ and $[PL_2]$ back into (14) yields,

$$\delta = \frac{\delta_{[P]} + \delta_{[PL_1]} K_A[L] + \delta_{[PL_2]} K_A K_B[L]^2}{1 + K_A[L] + K_A K_B[L]^2} \qquad (24)$$

Calculating $[L]$ from $[L_T]$

The total ligand concentration $[L_T]$ is equal to the sum of free and complexed forms of ligand. For two site sequential binding, the ligand can exist in three states,

$$L_T = [L] + [PL_1] + [PL_2]$$

$[PL_1]$ and $[PL_2]$ can be written in terms of [L] as explained by equations (22) and (23). Therefore,

$$L_T = [L] + P_T\left[\frac{K_A[L]}{1 + K_A + K_A K_B[L]^2}\right] + P_T\left[\frac{K_A K_B[L]^2}{1 + K_A + K_A K_B[L]^2}\right]$$

On rearranging, the polynomial equation that has to be solved is obtained.

$$K_A K_B[L]^3 + (K_A + P_T K_A K_B - L_T K_A K_B)[L]^2$$

$$+ (1 + P_T K_A - L_T K_A)[L] - L_T = 0$$

Supporting Information

Figure S1 Location of binding site residues of BH3I-1 in the primary sequence and 3D structure of hBcl$_{XL}$. The binding site residues are interspersed among the BH3 (red), BH1 (green) and BH2 (cyan) domains in the primary sequence and are highlighted with yellow and red color for site A and B respectively in both (A) primary sequence and (B) structure of hBcl$_{XL}$.
Found at: doi:10.1371/journal.pone.0008943.s001 (9.08 MB EPS)

Acknowledgments

Author Contributions

Conceived and designed the experiments: YKM. Performed the experiments: JK. Analyzed the data: JK. Contributed reagents/materials/analysis tools: VCKY YKM. Wrote the paper: JK.

References

1. Van Dongen M, Weigelt J, Uppenberg J, Schultz J, Wikstrom M (2002) Structure-based screening and design in drug discovery. Drug Discov Today 7: 471–478.
2. Carlomagno T (2005) Ligand-target interactions: what can we learn from NMR? Annu Rev Biophys Biomol Struct 34: 245–266.
3. Takeuchi K, Wagner G (2006) NMR studies of protein interactions. Curr Opin Struct Biol 16: 109–117.
4. Roberts G (2000) Applications of NMR in drug discovery. Drug Discov Today 5: 230–240.
5. Bain AD (2003) Chemical exchange in NMR. Prog Nucl Magn Reson Spectrosc 43: 63–103.
6. Palmer A III (2004) NMR characterization of the dynamics of biomacromolecules. Chem Rev 104: 3623–3640.
7. Jayalakshmi V, Krishna NR (2002) Complete relaxation and conformational exchange matrix analysis (CORCEMA) of intermolecular saturation transfer effects in reversibly forming ligand–receptor complexes. J Magn Reson 155: 106–118.
8. Jayalakshmi V, Rama Krishna N (2004) CORCEMA refinement of the bound ligand conformation within the protein binding pocket in reversibly forming weak complexes using STD-NMR intensities. J Magn Reson 168: 36–45.
9. Moseley H, Lee W, Arrowsmith C, Krishna N (1997) Quantitative determination of conformational, dynamic, and kinetic parameters of a ligand-protein/DNA complex from a complete relaxation and conformational exchange matrix analysis of intermolecular transferred NOESY. Biochemistry-US 36: 5293–5299.
10. Keeler J (2005) Understanding NMR spectroscopy Wiley West Sussex. 459 p.
11. Cavanagh J (1996) Protein NMR spectroscopy: principles and practice Academic Press. 885 p.
12. Bain AD, Duns GJ (1996) A unified approach to dynamic NMR based on a physical interpretation of the transition probability. Can J Chem 74: 819–824.
13. Gutowsky H, Holm C (1956) Rate processes and nuclear magnetic resonance spectra. II. Hindered internal rotation of amides. J Chem Phys 25: 1228–1234.
14. Binsch G (1969) Unified theory of exchange effects on nuclear magnetic resonance line shapes. J Am Chem Soc 91: 1304–1309.
15. McConnell HM (1958) Reaction rates by nuclear magnetic resonance. J Chem Phys 28: 430–431.
16. Johnson CS (1965) Chemical rate processes and magnetic resonance. Adv Magn Reson 1: 33–102.
17. Bain AD (1998) Blurring the distinction between slow and intermediate chemical exchange. Biochem Cell Biol 76: 171–176.

18. Davies D, Eaton R, Baranovsky S, Veselkov A (2000) NMR investigation of the complexation of daunomycin with deoxytetranucleotides of different base sequence in aqueous solution. J Biomol Struct Dyn 17: 887.

19. Davies D, Veselkov A (1996) Structural and thermodynamical analysis of molecular complexation by 1H NMR spectroscopy. Intercalation of ethidium bromide with the isomeric deoxytetranucleoside triphosphates 5-d (GpCpGpC) and 5-d (CpGpCpG) in aqueous solution. J Chem Soc Faraday T 92: 3545–3557.

20. Veselkov A, Evstigneev M, Rozvadovskaya A, Hernandez Santiago A, Zubchenok O, et al. (2004) 1H NMR structural and thermodynamical analysis of the hetero-association of daunomycin and novatrone in aqueous solution. J Mol Struct 701: 31–37.

21. Deb K, Pratap A, Agarwal S, Meyarivan T (2002) A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE T Evolut Comput 6: 182–197.

22. Vaux D, Strasser A (1996) The molecular biology of apoptosis. Proc Natl Acad Sci U S A 93: 2239–2244.

23. Adams J, Cory S (2007) Bcl-2 regulated apoptosis: mechanism and therapeutic potential. Curr opin Immunol 19: 488.

24. Berghella AM, Pellegrini P, Contasta I, Beato TD, Adorno D (1998) Bcl-2 and drugs used in the treatment of cancer: new strategies of biotherapy which should not be underestimated. Cancer Biother Radio 13: 225–236.

25. Huang Z (2000) Bcl-2 family proteins as targets for anticancer drug design. Oncogene 19: 6627–6631.

26. Zhang YH, Bhunia A, Wan KF, Lee MC, Chan SL, et al. (2006) Chelerythrine and sanguinarine dock at distinct sites on hBcl$_{XL}$ that are not the classic BH3 binding cleft. J Mol Biol 364: 536–549.

27. Lugovskoy AA, Degterev AI, Fahmy AF, Zhou P, Gross JD, et al. (2002) A novel approach for characterizing protein ligand complexes: molecular basis for specificity of small-molecule Bcl-2 inhibitors. J Am Chem Soc 124: 1234–1240.

28. Tjelvar SG, Olsson, Mark A, Williams, William R, et al. (2008) The thermodynamics of protein-ligand interaction and solvation: insights for ligand design. J Mol Biol 384: 1002–1017.

29. Harding S, Chowdhry B (2001) Protein-ligand interactions: hydrodynamics and calorimetry. , USA: Oxford University Press. 360 p.

30. Mandel A, Akke M, Palmer A III (1995) Backbone dynamics of *Escherichia coli* ribonuclease HI: correlations with structure and function in an active enzyme. J Mol Biol 246: 144–163.

31. Kovrigin E, Loria J (2006) Characterization of the transition state of functional enzyme dynamics. J Am Chem Soc 128: 7724–7725.

32. Beach H, Cole R, Gill M, Loria J (2005) Conservation of mus-ms enzyme motions in the apo- and substrate-mimicked state. J Am Chem Soc 127: 9167–9176.

33. Kovrigin E, Kempf J, Grey M, Loria J (2006) Faithful estimation of dynamics parameters from CPMG relaxation dispersion measurements. J Magn Reson 180: 93–104.

34. Osapay K, Case DA (1991) A new analysis of proton chemical shifts in proteins. J Am Chem Soc 113: 9436–9444.

35. Neal S, Nip AM, Zhang H, Wishart DS (2003) Rapid and accurate calculation of protein 1H, 13C and 15N chemical shifts. J Biomol NMR 26: 215–240.

36. Sharma Y, Kwon OY, Brooks B, Tjandra N (2002) An *ab initio* study of amide proton shift tensor dependence on local protein structure. J Am Chem Soc 124: 327–335.

37. Le H, Oldfield E (1996) *Ab initio* studies of amide - N chemical shifts in dipeptides: applications to protein NMR spectroscopy. J Phys Chem 100: 16423–16428.

38. Gorecki A, Kkepys B, Bonarek P, Wasylewski Z (2009) Kinetic studies of cAMP-induced propagation of the allosteric signal in the cAMP receptor protein from *Escherichia coli* with the use of site-directed mutagenesis. Int J Biol Macromol 44: 262–270.

39. McCoy MA, Wyss DF (2002) Spatial localization of ligand binding sites from electron current density surfaces calculated from NMR chemical shift perturbations. J Am Chem Soc 124: 11758–11763.

40. McCoy M, Wyss D (2002) Structures of protein-protein complexes are docked using only NMR restraints from residual dipolar coupling and chemical shift perturbations. J Am Chem Soc 124: 2104–2105.

41. Lama D, Ramasubbu, Sankararamakrishnan (2008) Anti-apoptotic hBcl$_{XL}$ protein in complex with BH3 peptides of pro-apoptotic Bak, Bad, and Bim proteins: comparative molecular dynamics simulations. Proteins 73: 492–514.

42. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, et al. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6: 277–293.

43. NMRpipe 9. http://spin.niddk.nih.gov/NMRPipe.

44. Goddard TD, Kneller DG. SPARKY 3.1.1. http://www.cgl.ucsf.edu/home/sparky/distrib-3.106.

45. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, et al. (1998) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. J Comput Chem 19: 1639–1662.

46. PDB. http://www.rcsb.org/pdb.

47. Dallakyan S, Omelchenko A, Sanner M, Karnati S. MGLTools 1. 5. 2. http://mgltools.scripps.edu.

48. Merritt E. gnuplot 4. 2. 4. http://www.gnuplot.info/download.html.

49. King EL, Altman (1956) A schematic method of deriving the rate laws for enzyme-catalyzed reactions. J Phys Chem 60: 1375–1378.

50. Bowden AEC (1977) An automatic method for deriving steady-state rate equations. J Biochem 165: 55–59.

51. Huang CY (1979) Derivation and initial velocity and isotope exchange rate equations. Method Enzymol 63: 54–84.

52. Segel HL (1975) Enzyme kinetics. New York: John Wiley and Sons press. 992 p.

53. Bain AD (2002) MEXICO 3.1 : The McMaster program for exchange lineshape calculations. http://www.chemistry.mcmaster.ca/bain/exchange.html.

54. Ferrin T. Chimera 1. 3. 2577. http://www.cgl.ucsf.edu/chimera/download.html.