DOI: 10.1002/cmdc.200500013

Effects of Conformational Dynamics on Predicted Protein Druggability

Scott P. Brown and Philip J. Hajduk*[a]

Drug discovery is a costly and time-consuming process that is critically dependent on the identification of disease-modifying proteins that can be targeted with high affinity and specificity by small-molecule drugs.[1] Given the explosion of potential new targets derived from genome research, there is currently a great need for rapid and reliable computational tools that can be used in the process of selecting new targets to prioritize those that have the highest potential for clinical impact. We have recently shown that the propensity of a protein target to bind small, druglike molecules (its "druggability") can be quantitatively correlated to descriptors of the binding region on the protein surface with reasonable correlation to experimentally known propensities.[2] For a given protein binding site, its druggability index, I_D, provides a measure of the potential to associate with small-molecule candidate drugs. Typical values of I_D lie in the range of -3.0 to 0.0. Binding pockets with $I_D > -1.0$ are considered to have a high potential for druggability, whereas pockets with $I_D < -1.5$ are considered to have a low potential for small-molecule druggability.

Whereas the above druggability analysis allowed identification of the known druggable binding sites on the majority of proteins examined in the investigation, only static representations of the protein structures were used. That is, the influence of protein conformational dynamics on predicted druggability was not assessed. Conformational fluctuations of protein binding pockets due to inherent thermal motions have the potential to influence predictions of protein druggability in a nontrivial way. A recent example of the effects of dynamics in altering pocket shape is the case of HIV integrase,[3] in which a new binding space is observed during the course of an analysis of a molecular dynamics (MD) trajectory. Additionally, other realizations based largely upon consideration of the effects of receptor-site dynamics and inherent flexibility have recently been made on HIV integrase.[4,5] Many proteins, especially those involved in protein–protein interactions, have flexible interfaces that can adopt several conformations. A case in point is Bcl-xL, for which large differences are observed between the unbound and peptide-bound complexes.[6,7] This is illustrated in Figure 1, which shows the unbound structure of Bcl-xL with a pocket that cannot accommodate the Bak peptide. In fact, the predicted I_D value for the bound complex is -0.5 (high druggability), whereas the predicted value for the apo structure is -2.0 (low druggability).

[a] *S. P. Brown, P. J. Hajduk*
Global Pharmaceutical Research and Development
Abbott Laboratories
100 Abbott Park Road, R46Y, AP10-LL
Abbott Park, IL 60064 (USA)
Fax: (+1) 847-938-2478
E-mail: philip.hajduk@abbott.com

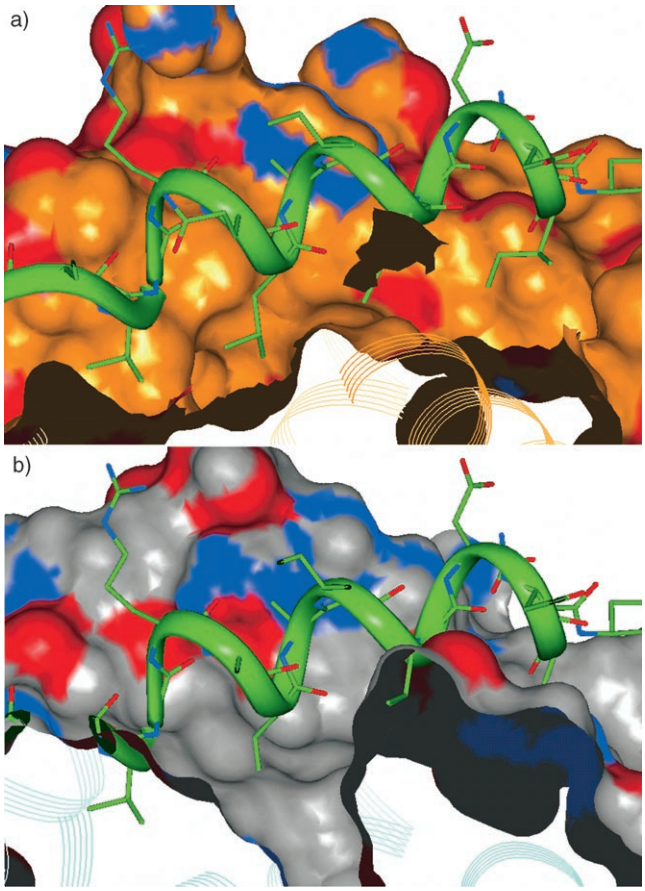

Figure 1. Surface representation of Bcl-xL for a) Bak-peptide-bound (orange surface) and b) unbound (gray surface) conformations. The Bak peptide (green ribbon) is shown in its experimental position in a) and is superimposed onto the structure shown in b). The Bak-peptide-bound complex has a predicted druggability index of -0.5 (high druggability), whereas the unbound structure has a value of -2.0 (low druggability).

To investigate the influence of a protein's conformational flexibility on its druggability assessment, we report herein MD simulations on three proteins: FKBP, the PH domain of Akt (Akt-PH), and apo Bcl-xL. By using structures taken from the MD trajectories at regular intervals, we calculated I_D values over the course of the simulation, and investigated how inherent receptor flexibility and the accompanying dynamic fluctuations influence predicted pocket druggability.

The values of I_D are calculated from the MD trajectory at intervals of 1 ps during the MD runs. This allows inspection of both the inherent dynamic fluctuations in I_D values as well as their convergence to an average value over the course of the dynamics trajectory. We take the average I_D value to indicate the overall druggability of a protein receptor site. Notably, this average is an arithmetic mean over snapshots from the MD trajectory, as our MD runs are performed under constant-energy conditions.

FKBP has a high degree of druggability as determined from both experiment ($I_D = -0.03$) and predictions that are based on the analysis of its static structure ($I_D = -0.24$).[2] Note that the experimental measure of I_D is obtained by taking the base-ten logarithm of the observed hit rate, which is expressed as the

percentage of compounds (from a large compound library) that are observed by NMR spectroscopy to bind to the region of interest on the protein target.

From MD we obtained an average druggability for FKBP of −0.6, which also places FKBP in the highly druggable category. FKBP has relatively low average magnitude fluctuations in I_D ($\langle \delta I_D^2 \rangle = 0.2$), and, as shown in Figure 2a, nearly always presents a highly druggable pocket ($I_D > -1.0$ for 90% of the conformations).

Akt-PH, in contrast, has a low degree of druggability from both experiment ($I_D = -1.91$)[2] and prediction from a static structure ($I_D = -1.98$). From our MD simulation, we obtained an average druggability index of −2.1. Figure 2b shows that there are a number of conformational fluctuations of the binding site in Akt-PH for which its calculated druggability index lies within the druggable regime ($I_D > -1.0$ for 11% of the conformations). Despite the relatively large magnitude of the fluctuations of I_D ($\langle \delta I_D^2 \rangle = 0.6$), the predicted druggability of Akt-PH remains low, which is in agreement with experimental hit rates. Thus transient pocket openings are insufficient to shift the overall druggability of Akt-PH into even the moderately druggable regime.

This situation is quite different for Bcl-xL. As mentioned above (and as shown in Figure 1), the free and Bak-bound static structures of Bcl-xL yield significantly different values for predicted druggability. Bcl-xL is experimentally known to be druggable,[8] yet the predicted druggability based solely on analysis of the static structure of the apo form gives an I_D value of −2.0 (very low druggability). In this case, protein flexibility plays a substantial role. As can be observed in Figure 2d, the average druggability index for Bcl-xL increases substantially as the MD simulation progresses from an initial starting value of $\langle I_D(0) \rangle \approx -2.0$ to its final value of $\langle I_D(100) \rangle \approx -1.3$. Furthermore, as observed in Figure 2c, the binding pocket on Bcl-xL explores a wide range of I_D values, persistently exposing a surface region that is predicted to be highly druggable ($I_D > -1.0$ for 28% of the conformations). The average magnitude of fluctuations in Bcl-xL is $\langle \delta I_D^2 \rangle = 0.6$, which is equivalent to that observed for Akt-PH. Thus, for Bcl-xL, the inclusion of dynamic fluctuations is critical for an appropriate computational assessment of protein druggability.

It is clear from the instantaneous values of I_D during the MD runs (Figure 2a–c) that the three proteins (FKBP, Akt-PH, and Bcl-xL) all have binding pockets that vary significantly in their properties. All of the binding pockets appear to experience roughly the same rate of thermal fluctuation of I_D about its average value. Yet apparently, it is the lifetimes of these transient fluctuations in conjunction with their magnitudes, which

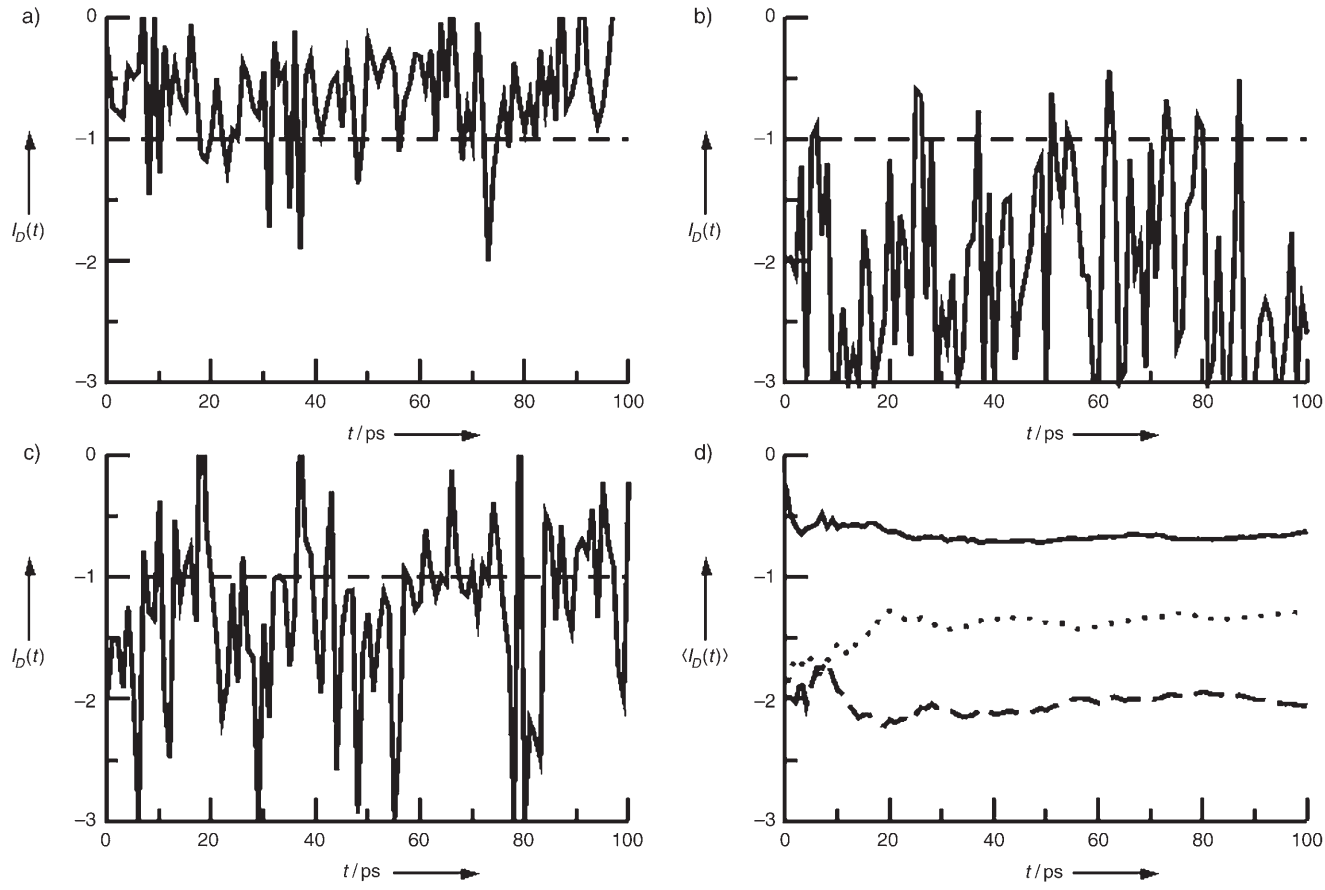

Figure 2. Calculated indices of druggability I_D from MD simulations. Instantaneous values of I_D for the protein snapshots along the MD trajectories taken at time t for a) FKBP, b) Akt-PH, and c) Bcl-xL. The horizontal dashed lines in a)–c) are merely guides for the eye to help delineate the cut-off boundaries in I_D (see text for details). d) The average index of druggability $\langle I_D(t) \rangle$ versus time t for FKBP (——), Akt-PH (- - - -), and Bcl-xL (••••). $\langle I_D(t) \rangle$ appears roughly converged after 100 ps with the following final averages: $\langle I_D \rangle \approx -0.6$ for FKBP, $\langle I_D \rangle \approx -2.1$ for Akt-PH, and $\langle I_D \rangle \approx -1.3$ for Bcl-xL.

serve to differentiate the predicted druggabilities for the three proteins.

For FKBP and Akt-PH, the presence of transient "openings" and "closings" of the binding pocket is insufficient to produce significant change in predicted druggability in comparison with static-structure calculations. By contrast, binding-pocket dynamics of Bcl-xL persistently explore longer-lived conformations with I_D values in the druggable regime.

These results illustrate the importance of considering a conformational ensemble for investigations of the druggability of a particular binding site, as the inspection of a single static-structure representation may produce erroneous results. By including dynamics into the assessment of druggability, one incorporates the inherent thermal fluctuations in binding-pocket shape that can affect small-molecule binding. As evidenced herein for Bcl-xL, the predicted druggability can be sharply dependent on the subtleties of conformational flexibility.

Computational Methods

The initial structures for FKBP and Akt-PH were obtained from the Protein Data Bank, with PDB IDs 1FKJ and 1H10, respectively. In the case of Bcl-xL, an in-house crystallographic structure was used.[6] This Bcl-xL structure was missing a loop segment between residues Phe 27 and Ile 36. The missing segment was replaced with an eight-residue polyglycine loop attached on each end to residues Phe 27 and Ile 36. All three proteins were initially stripped of all nonpeptide atoms to leave only the remaining protein atoms.

Molecular dynamics: All MD simulations were preformed by using the SANDER program from AMBER[9] (version 8). As we were not interested in specific solvent interactions, we used the Generalized Born model of Onufriev et al.[10] to capture solvent effects. The parm99 force field[11] was used for simulations of all proteins. Hydrogen bonds were constrained with SHAKE, and a time step of 2 fs was used. A distance cut-off of 18.0 Å was used in all runs.
All initial structures were minimized by performing 5000 steps of steepest decent minimization followed by 5000 steps of conjugate gradient minimization. The minimized structures were then equilibrated with gradual heating from 0 to 300 K over 20 ps. The final structures from equilibration were then used to start constant-energy MD production runs over trajectories of 100 ps. Structures for subsequent analysis were saved every 1 ps during the production runs.

Calculation of druggability. The snapshot structures saved during the production MD runs were used as input in calculating instantaneous I_D values. This was done by using a previously reported method.[2] Briefly, the predicted index of druggability was calculated from a linear combination of terms, with each term including both a linear and logarithmic function of an active-site descriptor,

$$I_D = \sum_{i=1}^{N} [a_i X_i + b_i \log(X_i)]$$

in which N is the number active-site descriptors, X_i is the i^{th} descriptor, and a_i and b_i are coefficients (obtained from fit to data) for the linear and logarithmic terms of the i^{th} parameter, respectively. For a description of the specific terms and values for the coefficients used in the equation, see Table 2 in the report by Hajduk, et al.[2]

Acknowledgments

The authors thank Dr. Steve Muchmore for his assistance in the loop construction on Bcl-xL.

Keywords: conformation analysis · drug delivery · druggability · molecular dynamics

[1] A. L. Hopkins, C. R. Groom, *Nat. Rev. Drug Discovery* **2002**, *1*, 727 – 730.
[2] P. J. Hajduk, J. R. Huth, S. W. Fesik, *J. Med. Chem.* **2005**, *48*, 2518 – 2525.
[3] J. Schames, R. Henchman, J. Siegel, C. Sotriffer, H. Ni, J. McCammon, *J. Med. Chem.* **2004**, *47*, 1879 – 1881.
[4] J. Deng, K. Lee, T. Sanchez, M. Cui, N. Neamati, J. Briggs, *J. Med. Chem.* **2005**, *48*, 1496 – 1505.
[5] R. Dayam, N. Neamati, *Bioorg. Med. Chem.* **2004**, *12*, 6371 – 6381.
[6] S. W. Muchmore, M. Sattler, H. Liang, R. P. Meadows, J. E. Harlan, H. S. Yoon, D. Nettesheim, B. S. Chang, C. B. Thompson, S. L. Wong, S. L. Ng, S. W. Fesik, *Nature* **1996**, *381*, 335 – 341.
[7] M. Sattler, H. Liang, D. Nettesheim, R. P. Meadows, J. E. Harlan, M. Eberstadt, H. S. Yoon, S. B. Shuker, B. S. Chang, A. J. Minn, C. B. Thompson, S. W. Fesik, *Science* **1997**, *275*, 983 – 986.
[8] T. Oltersdorf, S. W. Elmore, A. R. Shoemaker, R. C. Armstrong, D. J. Augeri, B. A. Belli, M. Bruncko, T. L. Deckwerth, J. Dinges, P. J. Hajduk, M. K. Joseph, S. Kitada, S. J. Korsmeyer, A. R. Kunzer, A. Letai, C. Li, M. J. Mitten, D. G. Nettesheim, S. Ng, P. M. Nimmer, J. M. O'Connor, A. Oleksijew, A. M. Petros, J. C. Reed, W. Shen, S. K. Tahir, C. B. Thompson, K. J. Tomaselli, B. Wang, M. D. Wendt, H. Zhang, S. W. Fesik, S. H. Rosenberg, *Nature* **2005**, *435*, 677 – 681.
[9] D. A. Pearlman, D. A. Case, J. W. Caldwell, W. S. Ross, I. T. E. Cheatham, S. DeBolt, D. Ferguson, G. Seibel, P. Kollman, *Comput. Phys. Commun.* **1995**, *91*, 1 – 41.
[10] A. Onufriev, D. Bashford, D. A. Case, *Proteins Struct. Funct. Genet.* **2004**, *55*, 383 – 394.
[11] J. Wang, P. Cieplak, P. A. Kollman, *J. Comput. Chem.* **2000**, *21*, 1049 – 1074.

Received: July 19, 2005

