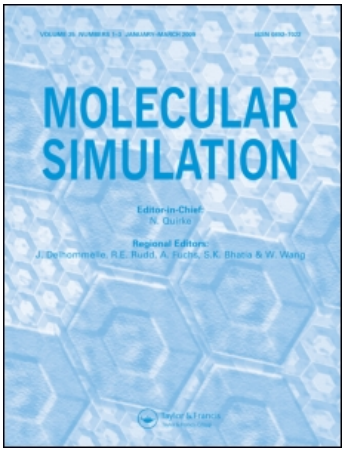

Molecular Simulation
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713644482

FlexE ensemble docking approach to virtual screening for CDK2 inhibitors

J. Kim[a]; J. G. Park[a]; Y. Chong[a]
[a] Division of Bioscience and Biotechnology, Institute of Biomedical Science and Technology, Konkuk University, Seoul, South Korea

To cite this Article Kim, J. , Park, J. G. and Chong, Y.(2007) 'FlexE ensemble docking approach to virtual screening for CDK2 inhibitors', Molecular Simulation, 33: 8, 667 — 676
To link to this Article: DOI: 10.1080/08927020701297401
URL: http://dx.doi.org/10.1080/08927020701297401

Molecular Simulation, Vol. 33, No. 8, July 2007, 667–676

Taylor & Francis
Taylor & Francis Group

FlexE ensemble docking approach to virtual screening for CDK2 inhibitors

J. KIM, J. G. PARK and Y. CHONG*

Division of Bioscience and Biotechnology, Institute of Biomedical Science and Technology, Konkuk University, 1 Hwayang-Dong, Gwangjin-Gu, Seoul 143-701, South Korea

(Received December 2006; in final form February 2007)

In spite of a proven potential and effectiveness of FlexE in docking flexible ligands into an ensemble of protein structures, FlexE has rarely been successful in virtual screening situations. In this study, we constructed cyclin-dependent kinase 2 (CDK2) ensemble structures which have exactly the same backbone conformations as 1AQ1 but differ only at the side chain torsion angles of the key amino acid (Lys33, Phe80, Lys89 and Asp145) residues: the torsion angles observed in the 17 CDK2 crystal structures were adapted to represent conformational flexibility. FlexE ensemble docking protocol then completely samples the full conformational fields generated by combination of torsions of the four amino acids in the ATP binding site. Virtual screening for CDK2 inhibitors by using the FlexE ensemble docking of a database composed of 48,703 inactives and 82 actives showed significant enrichment factor (EF = 18.5), and successfully identified 71 actives among the top 132 ligands (53.8%) ranked by total energy scores. Moreover, total energy scoring followed by visual inspection filtered-off non-specific binders among the highly-ranked ligands to increase the ratio of actives-to-inactives to 71:13 at the top 5% of the virtual screening solutions.

Keywords: CDK2; Docking; FlexE; Virtual screening

1. Introduction

Cyclin-dependent kinase 2 (CDK2) is an enzyme involved in the regulation of the cell cycle, which undergoes a series of conformational changes upon binding to cyclin and phosphorylation to yield a fully active complex [1]. ATP and inhibitor binding at the ATP-binding site also induce further conformational changes [2]. Thus, more than 60 CDK2 crystal structures are publicly available at present, which have subtle but significant structural differences [2]. In a crystal structure, however, a protein conformation is optimally adapted for interaction with one specific ligand, which makes it difficult to choose a basal protein structure for use in high-throughput docking of thousands of ligands of diverse structures. One way around this problem is to incorporate protein flexibility into the docking program, and the program FlexE [3] tackles receptor flexibility using a combinatorial approach, which generates new, recombined protein conformations based on a set of experimentally determined structures of a target protein. Thus, FlexE approach is based on a united protein description generated from the superimposed structures of the ensemble. For varying parts of the protein, discrete alternative conformations are explicitly taken into account, which can be combinatorially joined to create new valid protein structures. FlexE can then dock flexible ligands into an ensemble of protein structures. However, in spite of the successful representation of the protein flexibility when docking a small set of ligands, FlexE has rarely been investigated for screening large databases. To our knowledge, only three studies have been directed for evaluation of the FlexE in virtual screening situation [4–6]. The first FlexE ensemble docking to virtual screening for CDK2 inhibitors was reported by Steffen *et al.* [4]. During the enrichment study, 6011 ligands were docked using the FlexE/ScreenScore docking scheme, and the authors found that enrichments in different ensembles could not even reach that yielded by FlexX and stated that the enrichment declined sharply with an increasing number of structures in the ensemble. Also, two independent studies recently reported [5,6] concluded that the flexibility solution provided by FlexE does not seem to reflect true flexibility, but FlexE could only be a useful tool for merging different crystal structures and

*Corresponding author. Email: chongy@konkuk.ac.kr

Molecular Simulation
ISSN 0892-7022 print/ISSN 1029-0435 online © 2007 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/08927020701297401

docking into them simultaneously, which saves significant computational time but does not necessarily improve the enrichment. It was also noted that the best result of FlexE was the reproduction of the enrichment given by FlexX, and this was limited to the four-membered ensemble, which is coincident with the results of Steffen *et al* [4]. We have also experienced the same problem in which the FlexE failed to run on the protein ensemble with more than four to five CDK2 structures. However, we found that FlexE run could be successfully performed (data not shown) with certain combinations of CDK2 structures, which commonly exclude three PDB structures (1FIN, 1FVV and 1H1P) among 17 CDK2 structures used for FlexE run. In order to understand the structural characteristics of these three PDB structures, a CDK2 structure complexed with an inhibitor staurosporine (PDB code: 1AQ1) was chosen as a reference and 17 CDK2 structures were aligned by homology. Interestingly, the three PDB structures (1FIN, 1FVV and 1H1P) showed significant conformational differences (root−mean−square−deviations, RMSD $\geqq$ 2.0 Å) compared with the reference structure (1AQ1), while the others matched well (RMSD $<$ 1.0 Å) (table 1). Based on these findings, it was assumed that, in general, FlexE does not tolerate flexibility in a wide range of the conformational field which may include the polypeptide backbone. Thus, for successful virtual screening by FlexE ensemble docking, it is necessary to generate a united protein structure which represents different protein conformations with a combination of minimal structural variations.

Recently, Thomas *et al.* showed that there are some CDK2 structures that are clearly better than others for docking study, and the main determinants of this are the volume of the binding site into which the ligands are docked [7]. Moreover, by comparison of 20 CDK2 structures, they concluded that the more the side chains of Lys33, Phe80, Lys89 and Asp145 protruded into the ATP binding site, the smaller the number of correctly docked ligands tended to be. Taken together, even though the relationship was not clearly notified by the authors, it is obvious that the volume of the ligand binding site would be determined by conformations of the four amino acid residues (Lys33, Phe80, Lys89 and Asp145). Superposition of 17 CDK2 structures (figure 1) clearly shows that they share almost the same backbone conformation but differ in the side chain conformation of the four amino acid residues (Lys33, Phe80, Lys89 and Asp145).

Table 1. RMSD of 17 crystal structures (align structures by homology).

PDB	*RMSD (Å)*	*PDB*	*RMSD (Å)*	*PDB*	*RMSD (Å)*
1AQ1	0.0000[†]	1FIN	**2.2482**	1H1P	**2.0551**
1B38	0.7297	1FVT	0.7180	1HCK	0.7375
1DI8	0.8783	1FVV	**2.0957**	1HCL	0.7014
1DM2	0.7396	1G5S	0.7835	1JVP	0.8081
1E1V	0.8778	1GZ8	0.7490	1OIT	0.7370
1E1X	0.8929	1H00	0.7141		

[†] CDK2 complexed with an inhibitor staurosporine (1AQ1) was chosen as a reference structure for comparison.

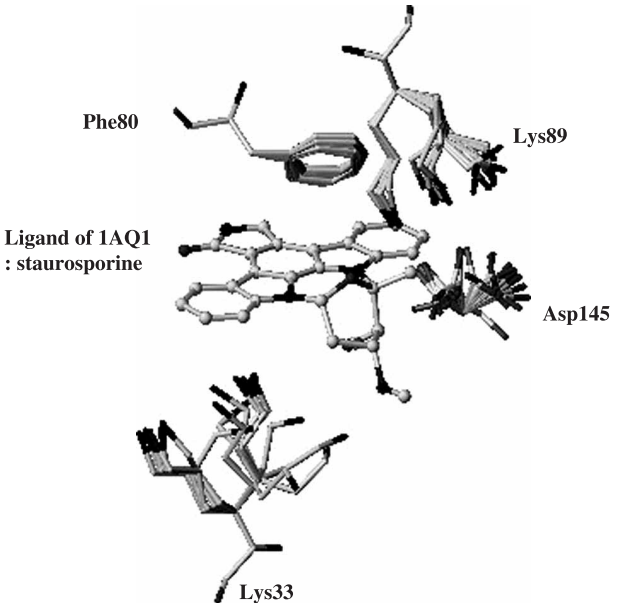

Figure 1. Superposition of 17 CDK2 structures. Fine gray lines indicate superimposed backbone Cα's, and the four residues (Lys33, Phe80, Lys89 and Asp145) are shown as capped sticks. A ball and stick model at the center is a ligand of the PDB ID 1AQ1.

Thus, in this study, we focused on the conformational flexibilities of the four key amino acid residues (Lys33, Phe80, Lys89 and Asp145) of CDK2 and, with the minimized conformational flexibilities in the protein ensemble, we tested FlexE ensemble docking to virtual screening for CDK2 inhibitors.

2. Materials and methods

2.1 Protein structures

The following CDK2 protein structures, taken from the Protein Data Bank [8], were used in this work: 1AQ1 [9], 1B38 [10], 1DI8 [11], 1DM2 [12], 1E1V [13], 1E1X [13], 1FIN [14], 1FVT [15], 1FVV [15], 1G5S [16], 1GZ8 [17], 1H00 [18], 1HIP [19], 1HCK [20], 1HCL [21], 1JVP [22] and 1OIT [23]. These 17 structures were chosen because of their diversity in terms of resolution [1.30 Å (1GZ8) ∼ 2.80 Å (1FVV)], size of the bound ligand [ATP (1FIN, 1HCK) ∼ staurosporine (1AQ1)], the presence or absence of other proteins (1FIN, 1FVV, 1H1P: complexed with cyclin A), phosphorylation status (1H1P: phosphorylated at Tyr160), and ATP binding site conformation [status of the loop around the active site (T-loop): ordered (1B38, 1E1V, 1E1X, 1FIN, 1FVV, 1GZ8, 1H1P, 1HCK and 1HCL) or disordered (1AQ1, 1DI8, 1DM2, 1FVT, 1G5S, 1H00, 1JVP and 1OIT)]. The side chains of lysine and arginine residues were protonated and the carboxylate groups of aspartic and glutamic acid were ionized. Water molecules contained in the PDB file were removed. In order to define the active sites of the proteins, all 17 enzyme structures were superimposed together with their ligand structure using the "Align

Structures by Homology" module of Sybyl 7.2 (Tripos Inc., St Louis, MO, USA). All atoms located less than 10 Å apart from any ligand atom were considered in ensembles. In addition, the complete amino acid was selected if at least one of its atoms is picked. The bound inhibitor was not included in the docking run. 1AQ1 was used as a reference structure for the united protein preparation.

2.2 Ligand structures

First, the ligand coordinates of the non-hydrogen atoms were extracted from the original PDB. They were used as reference for the calculation of the RMSD values later on. The ligand input files were obtained by defining correct atom types (including hybridization states) and correct bond types, adding hydrogen atoms, assigning formal charges to each atom, and finally energy-minimizing the reference structure (using the standard Tripos force field with Gästeiger–Hückel charges until the energy gradient converged to below 0.05 kcal/mol). The energy minimization guarantees a low-energy conformation with suitable bond length and angles. This new geometry and the fact that the minimized structures is not translated according to its original crystal structure guarantees that there is no implicit docking information about the protein–ligand complex of the PDB structure in the ligand input file. In general, all carboxylic acid groups are ionized while all amino groups but no amide groups are protonated.

2.3 Virtual screening and preparation of the ligand database

LeadQuest is a screening library commercialized by Tripos, Inc. Our screening library involves a subset of LeadQuest compound libraries (LQDB) as inactive molecules. The LQDB has molecular weight less than 550, ClogP less than 6, no more than one chiral center, no non-drug-like fragments (a toxicity filter, based on a growing knowledge base), and at least one chromophore. Our active set was compiled by the diverse selection of 35 ligands extracted from the crystal structures of CDK2 (1AQ1, 1CKP, 1D18, 1DM2, 1E1V, 1E1X, 1E9H, 1FVT, 1FVV, 1G5S, 1GZ8, 1H00, 1H01, 1H06, 1H07, 1H08, 1H0V, 1H0W, 1H1Q, 1H1R, 1H1S, 1JSV, 1JVP, 1KE5, 1KE6, 1KE7, 1KE8, 1OGU, 1OIQ, 1OIR, 1OIT, 1P2A, 1PF8, 1PXK and 1PXL; figure 2) and 47 oxindole analogues [24] (figure 3). Thus, the input library comprised 48,703 inactive and 82 active compounds, and the final screening library was comprised of both active and inactive sets, which were stored as a Sybyl SLN list and converted to Sybyl mol2 format by means of Concord™.

2.4 Docking protocol

Virtual screening experiments were performed using FlexE. Standard parameters were used as implemented in the SYBYL 7.2 package. The FlexE approach is based on the united protein description which handles the similarities and differences of the protein structures of the ensemble. Formal charges and particle concept options were always checked. Thirty poses were saved in mol2 files for further analysis. All stored poses were rescored using the CScore™ module of SYBYL 7.2 comprising five different scoring functions including Dock [25], Chem [26], FlexX [27], PMF [28] and Gold [29]. In order to eliminate conformations with impossible torsion energy values, we introduced torsion energy constraints by setting up the maximal torsion energy term to default value (20 kJ/mol).

3. Results and discussion

3.1 Developing a virtual screening protocol. Comparison of X-ray structures

Comparison of the X-ray structures has revealed that overall folds of both apo enzyme and its ligand-bound forms are very similar, which is shown by low RMSD values of Cα's in table 1. However, even though CDK2 structures share almost the same backbone conformations, superposition of the 17 CDK2 crystal structures (figure 1) revealed that they show major differences in the side chain conformations at the four residues (Lys33, Phe80, Lys89 and Asp145), which are responsible for the different volume and thereby different three dimensional structure of the ATP-binding site of CDK2. The conformations of the four residues are summarized in figure 4. Even though Asp145 is distributed in a relatively wide conformational field compared with other amino acids, the four key residues (Lys33, Phe80, Lys89 and Asp145) show conserved backbone conformations (phi and psi) throughout 17 crystal structures (figure 4). The Ramachandran plot for the four amino acid residues clearly show that the conformations of these residues lie in the confined energetically favored regions (figure 5) without significant conformational differences. Therefore, the contribution of the flexibility of the protein backbone, if any, to the overall flexibility of CDK2 can be assumed to be marginal.

However, side chain conformations of the four residues in the 17 crystal structures studied were not conserved at all (figure 6), which confirms that the conformational flexibilities of the four amino acids are focused on their side chains.

As these residues are directly involved in the formation of the ATP binding site, it is obvious that the conformational differences would exert significant impact on the overall structure of the ATP binding site and thereby different ligand specificity. Taken together, we assumed that the conformational differences in various CDK2 X-ray structures result from different side chain conformations of the key four amino acids (Lys33, Phe80, Lys89 and Asp145) in the ATP binding site and, therefore, protein ensembles could be constructed by a simple

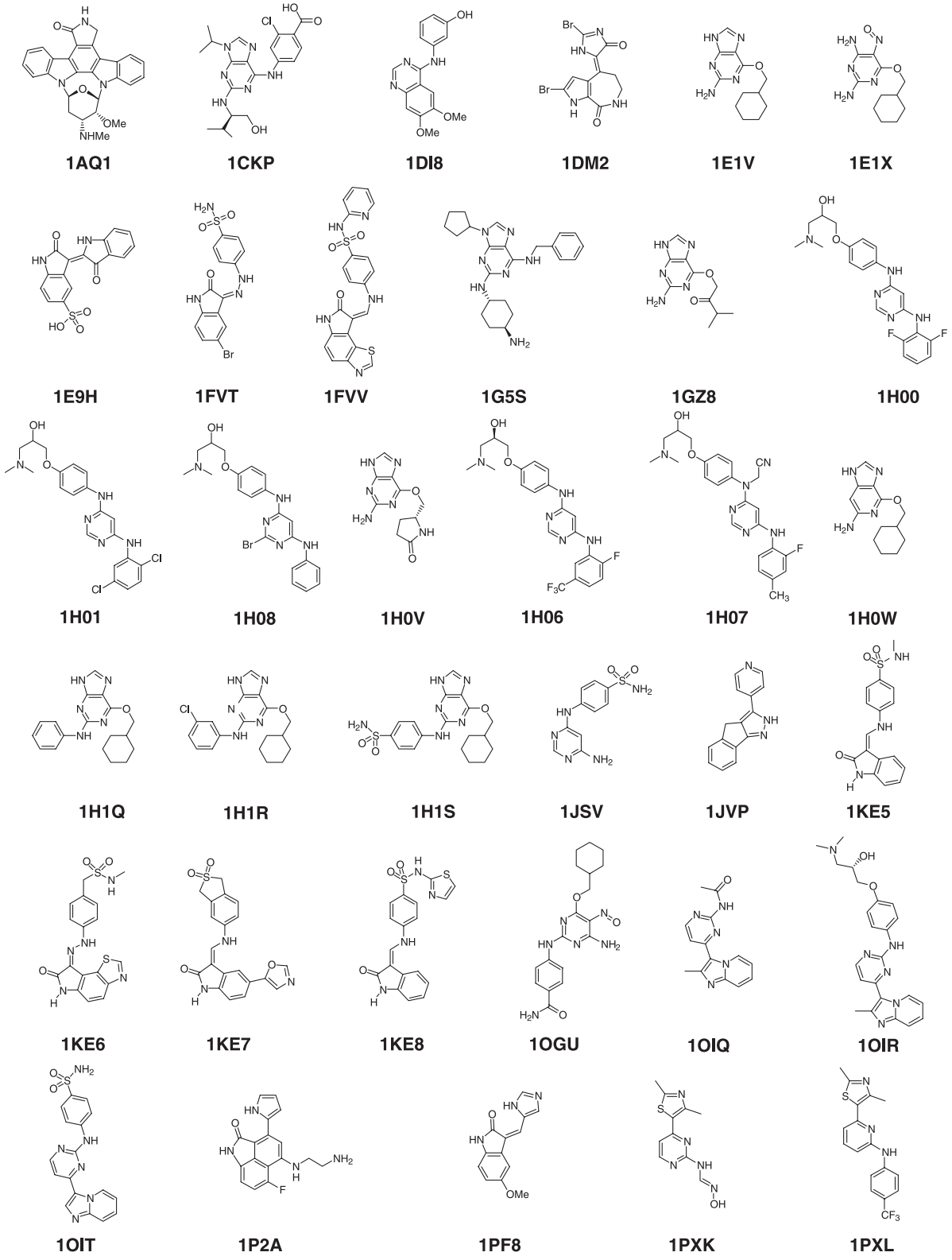

Figure 2. Structures of the 35 ligands extracted from the CDK2 crystal structures.

combination of the side chain conformations of the four amino acid residues on a consensus backbone structure. Then, the generated protein ensembles would be compatible with FlexE by virtue of the minimized Cα RMSD endowed by sharing exactly the same backbone structures. Nevertheless, the key structural features which represent the conformational flexibilities of the ATP binding site are reserved in the protein ensemble by

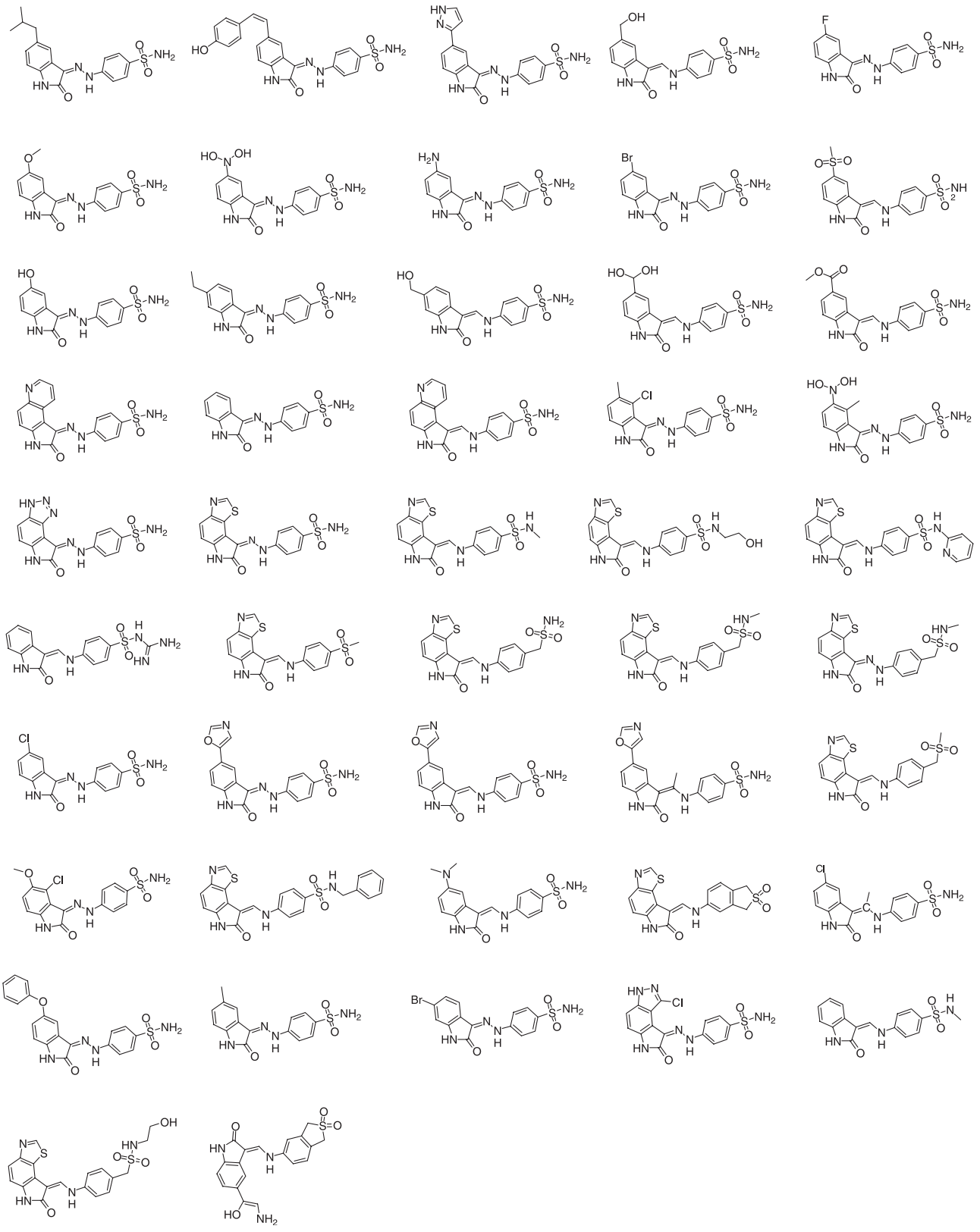

Figure 3. Structures of the 47 oxindole analogues used as active molecules.

combination of the torsions of the four amino acid residues found in the 17 crystal structures of CDK2.

A CDK2 crystal structure with staurosporine bound at the ATP binding site (PDB ID, 1AQ1 [9]) was used as our platform structure for this purpose because 1AQ1 has a wide open ATP binding site from which various combinations of side chain torsions of the four amino acids without invoking steric hindrance with nearby

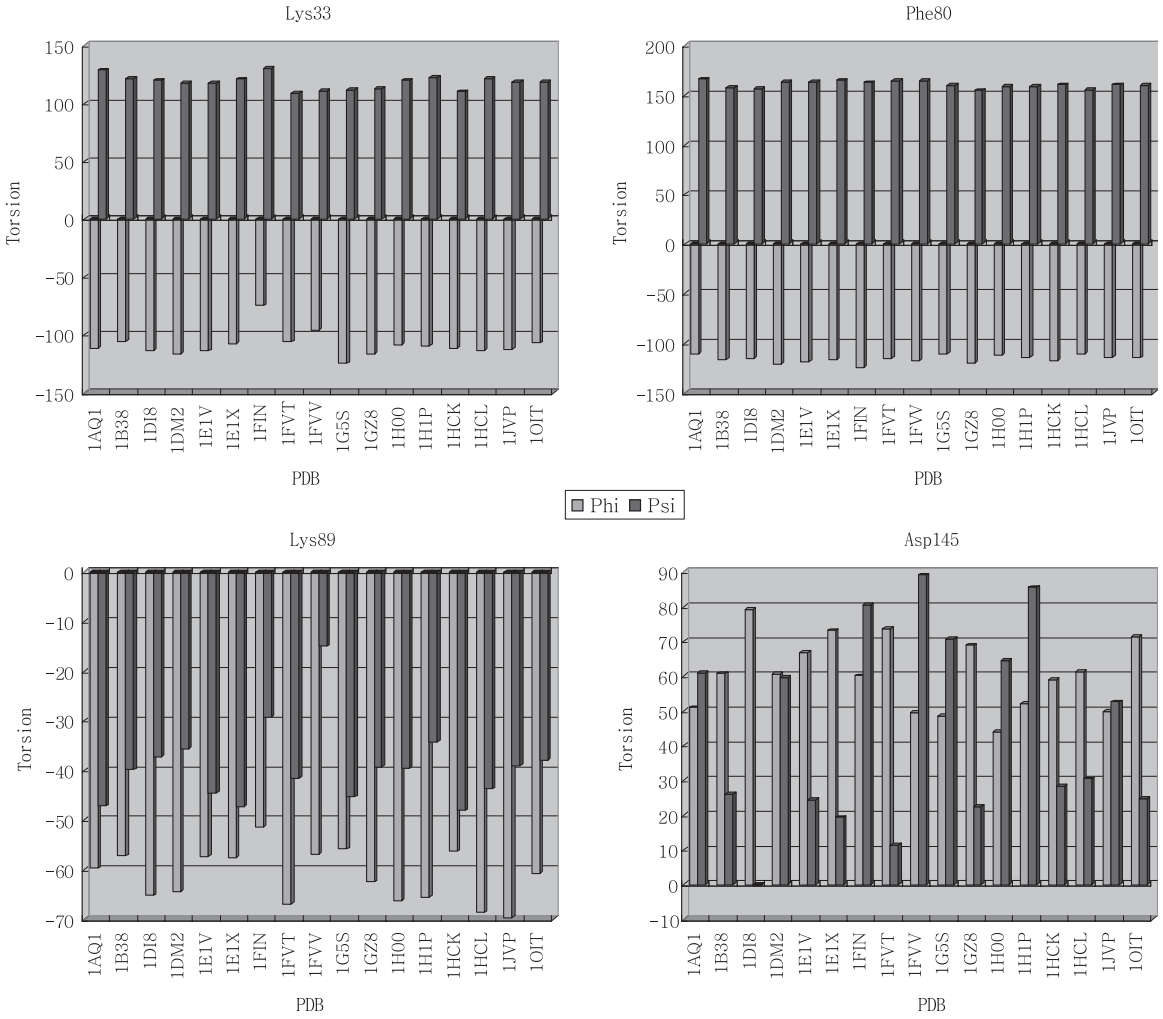

Figure 4. Backbone torsion angles of the four residues (Lys33, Phe80, Lys89 and Asp145).

residues can be feasible. Based on this platform structure, 17 ensemble structures were manually constructed by changing the side chain torsion angles of the four amino acids according to the torsion angles observed in the 17 crystal structures (figures 4 and 6).

3.2 Virtual screening protocol

FlexE algorithm then samples discrete receptor conformations in a combinatorial fashion while incrementally building ligands into the binding site. Thus, the full conformational fields which are generated by combination of torsions of the four amino acids in the ATP binding site can be covered by this method with minimum computational efforts.

3.3 Enrichment studies

To show the effectiveness of our screening protocol, we calculated enrichment factors (EF). The LQDB database containing 48,703 unique molecules was used as a background of drug-like decoys for enrichment calculations, and the screening library also contained 82 active compounds (35 ligands extracted from the crystal structures of CDK2 and 47 oxindole analogues). We assumed that molecules annotated as inhibitors of a given enzyme in the LQDB are true positives and the remaining molecules are true negatives (neither of these assumptions is likely to be entirely correct). The quality of enrichment is measured as the proportion of true binders found in selected subsets from the docking (or rescoring) calculations compared with the proportion expected from random selection. The EF assesses the quality of the rankings [24]:

$$\mathrm{EF}(\%) = \frac{(N_{\mathrm{active}(\%)}/N_{(\%)})}{(N_{\mathrm{acitive}}/N_{\mathrm{all}})}$$

where EF (%) is given as the percentage of the ranked database, $N_{\mathrm{active}(\%)}$ is the number of active compounds in a selected subset of the ranked database, $N_{(\%)}$ is the number of active molecules and the number of compounds in the screening database.

Thirty poses generated for all of the ligands from FlexE docking run were saved, and the saved poses were then

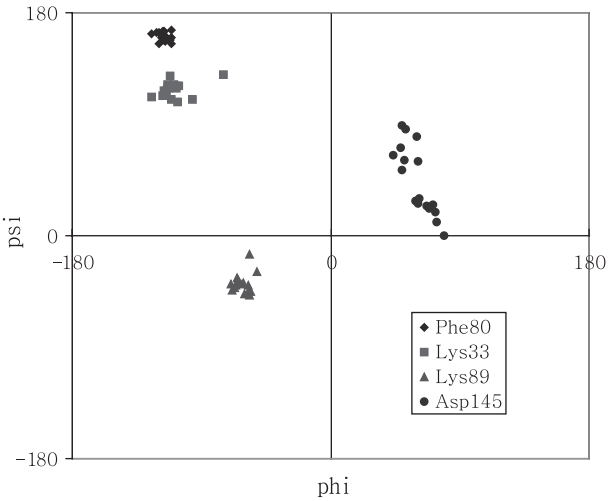

Figure 5. Ramachandran plot for the four residues (Lys33, Phe80, Lys89 and Asp145).

scored using the CScore™ module of SYBYL 7.2 comprising five different scoring functions including Dock [25], Chem [26], FlexX [27], PMF [28] and Gold [29]. Then the ligands with the lowest scores were ranked with the help of the total energy scores. Compared with the previous report (EF = 2.8–5.8) [4], our novel virtual screening protocol for CDK2 inhibitors using the FlexE ensemble docking provided significant improvement in the EF (EF = 18.5). More interestingly, 25 out of 27 top ranked ligands (top 1% of the ranked database) proved to be actives. At the top 3 and 5% of the ranked database, 55.7% (44 out of 79) and 53.8% (71 out of 132) of the actives were identified among the docked ligands, respectively (table 2).

With this result in hand, we performed visual inspection of 2644 ligands in the ranked database in order to improve our novel virtual screening protocol. In 1AQ1, staurosporine forms main polar interactions with Glu81 and Leu83 of the enzyme (figure 7).

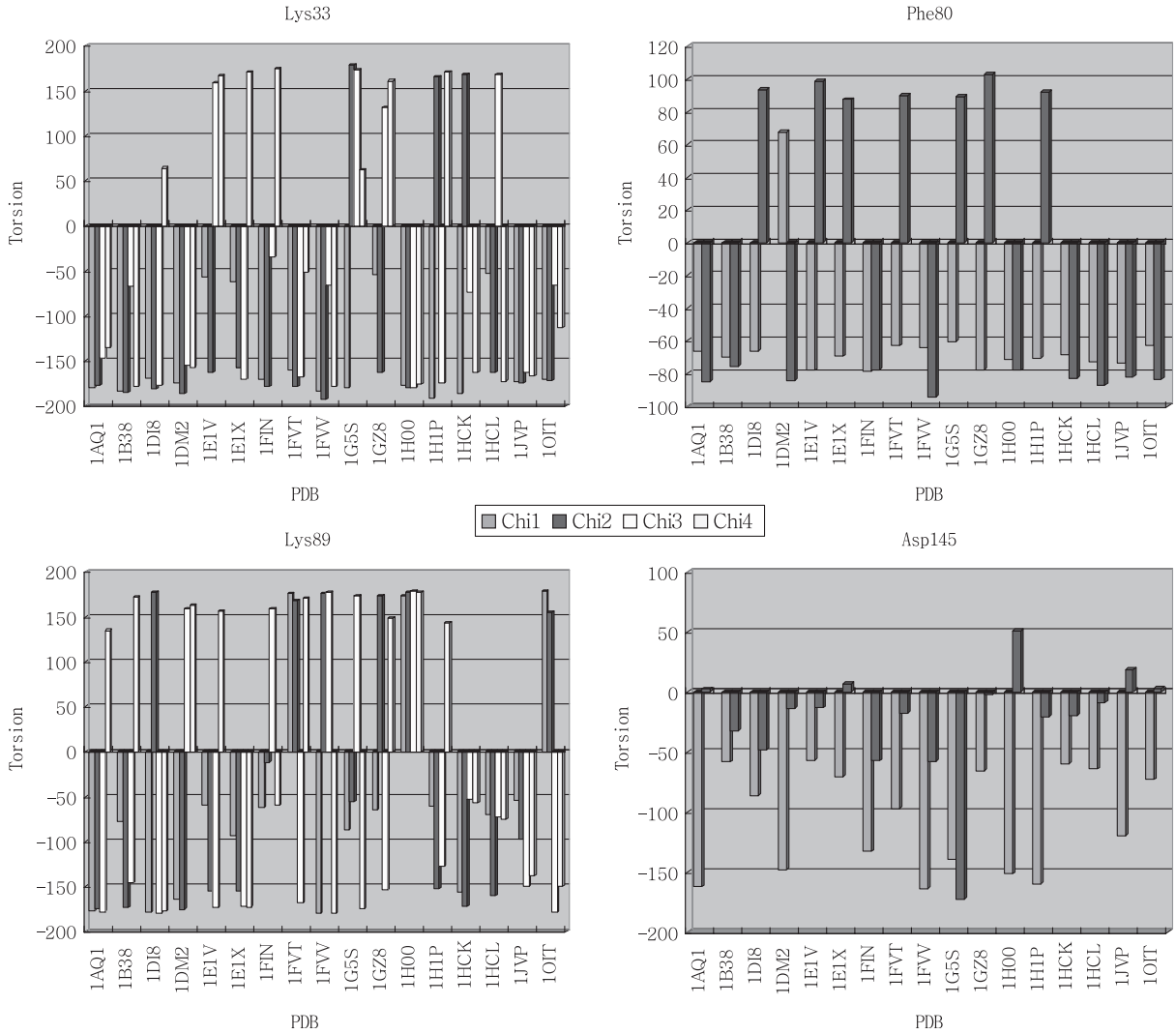

Figure 6. Side chain torsion angles of the four residues (Lys33, Phe80, Lys89 and Asp145).

Table 2. Number of the actives found at the top 1, 3 and 5% of the ranked database.

	Scoring function[†]			*Visual inspection filter*[‡]		
	No. of the actives	*No. of the inactives*	*Percentage of the actives*[¶]	*No. of the actives*	*No. of the inactives*[§]	*Percentage of the actives*
1%	25	2	92.6	25	2	92.6
3%	44	35	55.7	44	11	80.0
5%	71	61	53.8	71	13	84.5

Total number of compounds in the screening DB: 48,785. Total number of the actives in the screening DB: 82. Total number of the docked ligands in the ranked DB: 2644. EF: 18.5.[†] Number of the docked ligands determined only by the final total energy scores. [‡] Number of the docked ligands determined after filtering-off non-specific binders by visual inspection. [¶] (No. of the actives)/(No. of the actives + No. of the inactives) × 100. [§] Inactive ligands without specific hydrogen bonding interactions were filtered-off.

Structural comparison of 17 CDK2 complexes strengthens the importance of similar specific interactions. In this regard, it should be noted that CDK2 crystal structures identified to date shows that ligands binding at the ATP binding site should form a hydrogen bond with the backbone of Leu83 [7]. Even though additional hydrogen bonds with Lys33, Lys89 or Gln131 take part in fine-tuning the binding affinities of some ligands, the key interaction is hydrogen bonding between ligand and backbone CO as well as NH of Leu83.

Thus, ligands docked into the ATP binding site without specific hydrogen bonding interaction with backbone CO and NH of Leu83 can be considered as nonspecific binders, which were filtered-off by visual inspection. This process of filtering nonspecific binders reduced the number of the inactives without affecting the number of the actives to give remarkably increased ratio of actives-to-inactives (table 2). It is noteworthy that all of the active compounds observed in the high rank showed hydrogen bonding interaction with the enzyme backbone. In conjunction with this result, our virtual screening protocol was also examined in terms of the ability to reproduce X-ray ligand positions in a near-native binding geometry (within 2.0 Å heavy atom RMSD). As 17 CDK2 structures were used for construction of the protein ensembles, ligands which were extracted from those 17 X-ray structures were investigated. Interestingly, 13 out of 17 ligands (76.5%) were successfully docked at the top 5% of the ranked database, and it is noteworthy that the docking poses of the docked ligands were well reproduced in the RMSD range of 0.02–0.47 Å with near-native binding geometries (table 3).

4. Conclusions

There have been several efforts to tackle the problem of docking into conformationally flexible enzymes but, due to the complexities of the specially designed approaches (docking multiple ligands to multiple enzyme structures)

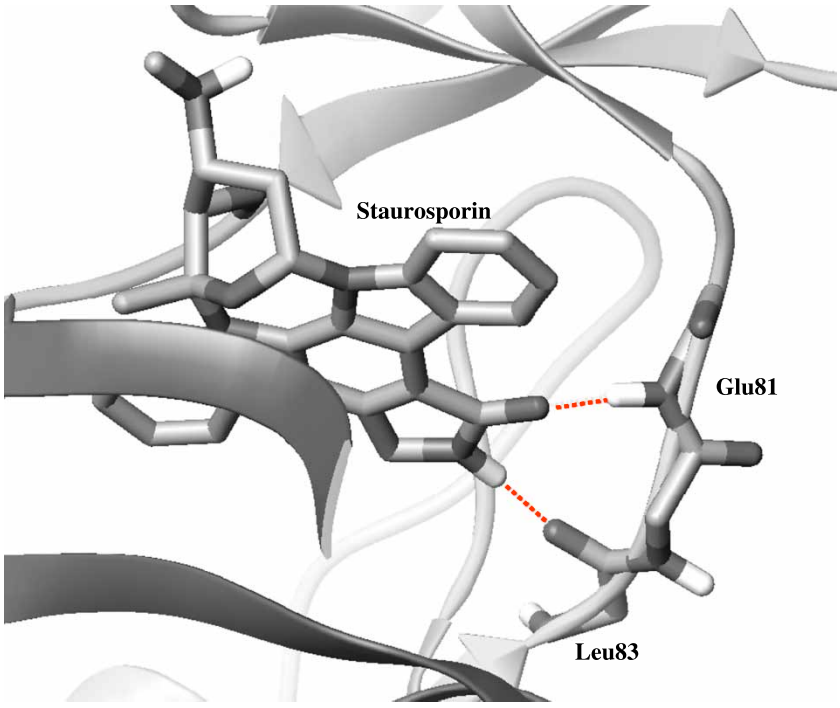

Figure 7. CDK2 complexed with staurosporine (PDB code 1AQ1). Staurosporine is hydrogen bonded to backbone NH and CO of Glu81 and Leu83, respectively.

Table 3. RMSD of the docked ligands from the X-ray ligand positions.

PDB	*RMSD (Å)*	*PDB*	*RMSD (Å)*	*PDB*	*RMSD (Å)*
1AQ1	0.18	1KE6	0.12	1P2A	0.12
1E9H	0.10	1KE7	0.19	1B38	–[†]
1FVV	0.15	1OGU	0.02	1HCK	–[†]
1H1R	0.10	1O1R	0.03		
1H1S	0.10	1OIT	0.47		

[†] This structure contains an ATP as a ligand bound at the ATP binding site. The crystal structure also includes a Mg ion chelated by ATP but, in the docking study, Mg ion could not be considered. As a result, RMSD was not estimated.

[7], there is an ongoing need for simple, intuitive, straightforward and easily reproducible method which can be generally applied. In this study, by assuming the major conformational difference among the numerous CDK2 crystal structures lies at the side chains of four amino acid residues (Lys33, Phe80, Lys89 and Asp145) at the ATP-binding site, we constructed 17 ensemble structures by changing side chain torsion angles of the four key amino acids on the platform structure (1AQ1) according to the torsion angles observed in the 17 crystal structures. FlexE algorithm then efficiently samples the full conformational fields generated by combination of torsions of the four amino acids in the ATP binding site. Our novel virtual screening protocol showed significant EF (EF = 18.5) in virtual screening of a database composed of 48,703 inactives and 82 actives, and successfully identified 25 actives among the top 27 ligands ranked by total energy scores. Moreover, total energy scoring followed by visual inspection filtered-off nonspecific binders among the highly-ranked ligands to increase the ratio of actives-to-inactives.

In summary, combination of the key structural features which confer conformational flexibilities to CDK2 instead of sampling whole enzyme structures in virtual screening for CDK2 inhibitors by using the FlexE ensemble docking successfully identified the actives from the database of inactives.

Acknowledgements

This work was supported by grant KRF-2006-33-C00176 from the Korea Research Foundation, Republic of Korea (MOEHRD, Basic Research Promotion Fund) and by grants from Biogreen 21 (Korea Ministry of Agriculture and Forestry), and the second Brain Korea 21 (Korea Ministry of Education). Jinyoung Kim is supported by the second Brain Korea 21.

References

[1] M. Huse, J. Kuriyan. The conformational plasticity of protein kinases. *Cell*, **109**, 275 (2002).

[2] S.Y. Wu, I. McNae, G. Kontopidis, S.J. McClue, C. McInnes, K.J. Stewart, S. Wang, D.I. Zheleva, H. Marriage, D.P. Lane, P. Taylor, P.M. Fischer, M.D. Walkinshaw. Discovery of a novel family of CDK inhibitors with the program LIDAEUS: structural basis for ligand-induced disordering of the activation loop. *Structure*, **11**, 399 (2003).

[3] H. Claußen, C. Buning, M. Rarey, T. Lengauer. FLEXE: efficient molecular docking considering protein structure variations. *J. Mol. Biol.*, **308**, 377 (2001).

[4] A. Steffen, J. Günther, H. Briem. Evaluation of the applicability of the FlexE ensemble docking approach to virtual screening for CDK2 inhibitors. paper presented at the Darmstädter Molecular Modelling Workshop, Erlangen, Germany, 19th April (2004).

[5] T. Polgár, G.M. Keserü. Ensemble docking into flexible active sites. Critical evaluation of FlexE against JNK-3 and β-secretase. *J. Chem. Inf. Model.*, **46**, 1795 (2006).

[6] T. Polgár, A. Baki, G.I. Szendrei, G.M. Keserü. Comparative virtual and experimental high-throughput screening for glycogen synthase kinase-3β inhibitors. *J. Med. Chem.*, **48**, 7946 (2005).

[7] M.P. Thomas, C. McInnes, P.M. Fischer. Protein structures in virtual screening: a case study with CDK2. *J. Med. Chem.*, **49**, 92 (2006).

[8] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne. The Protein Data Bank. *Nucleic Acids Res.*, **28**, 235 (2000).

[9] A.M. Lawrie, M.E. Noble, P. Tunnah, N.R. Brown, L.N. Johnson, J.A. Endicott. Protein kinase inhibition by staurosporine revealed in details of the molecular interaction with CDK2. *Nat. Struct. Biol.*, **4**, 796 (1997).

[10] N.R. Brown, M.E. Noble, A.M. Lawrie, M.C. Morris, P. Tunnah, G. Divita, L.N. Johnson, J.A. Endicott. Effects of phosphorylation of threonine 160 on cyclin-dependent kinase 2 structure and activity. *J. Biol. Chem.*, **274**, 8746 (1999).

[11] L. Shewchuk, A. Hassell, B. Wisely, W. Rocque, W. Holmes, J. Veal, L.F. Kuyper. Binding mode of the 4-anilinoquinazoline class of protein kinase inhibitor: X-ray crystallographic studies of 4-anilinoquinazolines bound to cyclin-dependent kinase 2 and p38 kinase. *J. Med. Chem.*, **43**, 133 (2000).

[12] L. Meijer, A.M. Thunnissen, A.W. White, M. Garnier, M. Nikolic, L.H. Tsai, J. Walter, K.E. Cleverley, P.C. Salinas, Y.Z. Wu, J. Biernat, E.M. Mandelkow, S.H. Kim, G.R. Pettit. Inhibition of cyclin-dependent kinases, GSK-3β and CK1 by hymenialdisine, a marine sponge constituent. *Chem. Biol.*, **7**, 51 (2000).

[13] C.E. Arris, F.T. Boyle, A.H. Calvert, N.J. Curtin, J.A. Endicott, E.F. Garman, A.E. Gibson, B.T. Golding, S. Grant, R.J. Griffin, P. Jewsbury, L.N. Johnson, A.M. Lawrie, D.R. Newell, M.E. Noble, E.A. Sausville, R. Schultz, W. Yu. Identification of novel purine and pyrimidine cyclin-dependent kinase inhibitors with distinct molecular interactions and tumor cell growth inhibition profiles. *J. Med. Chem.*, **43**, 2797 (2000).

[14] P.D. Jeffrey, A.A. Russo, K. Polyak, E. Gibbs, J. Hurwitz, J. Massague, N.P. Pavletich. Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex. *Nature*, **376**, 313 (1995).

[15] S.T. Davis, B.G. Benson, H.N. Bramson, D.E. Chapman, S.H. Dickerson, K.M. Dold, D.J. Eberwein, M. Edelstein, S.V. Frye, R.T. Gampe Jr., R.J. Griffin, P.A. Harris, A.M. Hassell, W.D. Holmes, R.N. Hunter, V.B. Knick, K. Lackey, B. Lovejoy, M.J. Luzzio, D. Murray, P. Parker, W.J. Rocque, L. Shewchuk, J.M. Veal, D.H. Walker, L.F. Kuyper. Prevention of chemotherapy-induced alopecia in rats by CDK inhibitors. *Science*, **291**, 134 (2001).

[16] M.K. Dreyer, D.R. Borcherding, J.A. Dumont, N.P. Peet, J.T. Tsay, P.S. Wright, A.J. Bitonti, J. Shen, S.H. Kim. Crystal structure of human cyclin-dependent kinase 2 in complex with the adenine-derived inhibitor H717. *J. Med. Chem.*, **44**, 524 (2001).

[17] A.E. Gibson, C.E. Arris, J. Bentley, F.T. Boyle, N.J. Davies, T.G. Curtin, J.A. Endicott, B.T. Golding, S. Grant, R.J. Griffin, P. Jewsbury, L.N. Johnson, V. Mesguiche, D.R. Newell, M.E. Noble, J.A. Tucker, H.J. Whitfield. Probing the ATP ribose-binding domain of cyclin-dependent kinases 1 and 2 with O6-substituted guanine derivatives. *J. Med. Chem.*, **45**, 3381 (2002).

[18] J.F. Beattie, G.A. Breault, R.P.A. Ellston, S. Green, P.J. Jewsbury, C.J. Midgley, R.T. Naven, C.A. Minshull, R.A. Pauptit, J.A. Tucker, J.E. Pease. Cyclin-dependent kinase 4 inhibitors as a treatment for cancer. Part 1: identification and optimisation of substituted 4,6-bis anilino pyrimidines. *Bioorg. Med. Chem. Lett.*, **13**, 2955 (2003).

[19] C.W. Carter Jr., J. Kraut, S.T. Freer, N.H. Xuong, R.A. Alden, R.G. Bartsch. Two angstrom crystal structure of oxidized chromatium high potential iron protein. *J. Biol. Chem.*, **249**, 4212 (1974).

[20] U. Schulze-Gahmen, H.L. De Bondt, S.H. Kim. High-resolution crystal structures of human cyclin-dependent kinase 2 with and

without ATP: Bound waters and natural ligand as guides for inhibitor design. *J. Med. Chem.*, **39**, 4540 (1996).

[21] U. Schulze-Gahmen, J. Brandsen, H.D. Jones, D.O. Morgan, L. Meijer, J. Vesely, S.H. Kim. Multiple modes of ligand recognition: crystal structures of cyclin-dependent protein kinase 2 in complex with ATP and two inhibitors, olomoucine and isopentenyladenine. *Proteins*, **22**, 378 (1995).

[22] P. Furet, T. Meyer, A. Strauss, S. Raccuglia, J.M. Rondeau. Structure-based design and protein X-ray analysis of a protein kinase inhibitor. *Bioorg. Med. Chem. Lett.*, **12**, 221 (2002).

[23] M. Anderson, J. Beattie, G. Breault, J. Breed, K. Byth, J. Culshaw, R. Ellston, S. Green, C. Minshull, R. Norman, R. Pauptit, J. Stanway, A. Thomas, P. Jewsbury. Imidazo[1,2-a]pyridines: a potent and selective class of cyclin-dependent kinase inhibitors identified through structure-based hybridization. *Bioorg. Med. Chem. Lett.*, **13**, 3021 (2003).

[24] H.N. Bramson, J. Corona, S.T. Davis, S.H. Dickerson, M. Edelstein, S.V. Frye, R.T. Gampe Jr., P.A. Harris, A. Hassell, W.D. Holmes, R.N. Hunter, K.E. Lackey, B. Lovejoy, M.J. Luzzio, V. Montana, W.J. Rocque, D. Rusnak, L. Shewchuk, J.M. Veal, D.H. Walker, L.F. Kuyper. Oxindole-based inhibitors of cyclin-dependent kinase 2 (CDK2): design, synthesis, enzymatic activities, and X-ray crystallographic analysis. *J. Med. Chem.*, **44**, 4339 (2001).

[25] (a) C. Meng, B.K. Shoichet, I.D. Kuntz. Automated docking with grid-based energy evaluation. *J. Comp. Chem.*, **13**, 505 (1992) (b) S.J. Weiner, P.A. Kollman, D.T. Nguyen, D.A. Case. An all atom force field for simulations of proteins and nucleic acids. *J. Comp. Chem.* **7**, 230 (1986). (c) R.M.A. Knegtel, D.M. Bayada, R.A. Engh, W. v.d. Saal, V.J.v. Geerestein, P.D.J. Grootenhuis. Comparison of two implementations of the incremental construction algorithm in flexible docking of thrombin inhibitors. *J. Comput. Aided. Mol. Des.* **13**, 167 (1999).

[26] M.D. Eldridge, C.W. Murray, T.R. Auton, G.V. Paolinine, R.P. Mee. Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *J. Comput. Aided. Mol. Des.*, **11**, 425 (1997).

[27] (a) M. Rarey, B. Kramer, T. Lengauer, G. Klebe. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.*, **261**, 470 (1996) (b) M. Rarey, B. Kramer, T. Lengauer. Docking of hydrophobic ligands with interaction-based matching algorithms. *Bioinformatics*, **15**, 243 (1999). (c) H.-J. Böhm. The development of a simple empirical scoring function to estimate the binding constant for a protein–ligand complex of known three-dimensional structure. *J. Comput. Aided. Mol. Des.* **8**, 243 (1994).

[28] (a) I. Muegge, Y.C. Martin. A general and fast scoring function for protein–ligand interactions: a simplified potential approach. *J. Med. Chem.*, **42**, 791 (1999) (b) F.C. Bernstein, T.F. Koetzle, G.J.B. Williams, E.F. Meyer Jr., M.D. Brice, J.R. Rodgers, O. Kennard, T. Shimanouchi, M. Tasumi. The Protein Data Bank. A computer based archival file for macromolecular structures. *Eur. J. Biochem.* **80**, 319 (1977).

[29] (a) G. Jones, P. Willett, R.C. Glen. Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. *J. Mol. Biol.*, **245**, 43 (1995) (b) G. Jones, P. Willett, R.C. Glen, A.R. Leach, R. Taylor. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **267**, 727 (1997).