IKKβ inhibitors identification part II: Ligand and structure-based virtual screening

Shanthi Nagarajan [a,b], Hyunah Choo [a], Yong Seo Cho [a], Kwang-Seok Oh [c], Byung Ho Lee [c], Kye Jung Shin [a], Ae Nim Pae [a,*]

[a] Neuro-Medicine Center, Life/Health Division, Korea Institute of Science and Technology, PO Box 131, Cheongryang, Seoul 130-650, Republic of Korea
[b] School of Science, Korea University of Science and Technology, 52 Eoeun dong, Yuseong-gu, Daejeon 305-333, Republic of Korea
[c] Drug Discovery Division, Korea Research Institute of Chemical Technology, Daejeon, Republic of Korea

ABSTRACT

IκB kinase (IKK) is critical in proinflammatory cytokine-induced IκBα phosphorylation and subsequent activation of the nuclear transcription factor NF-κB complex. The activated NF-κB plays a major role in the pathogenesis of a number of human disorders, such as rheumatic and chronic inflammatory diseases. The inhibition of NF-κB activation by small molecule inhibitors that targets IKKβ may provide a pharmacological basis for interfering with these acute processes. To date, only three inhibitors have passed pre-clinical trials; on the other hand, identifying novel IKKβ inhibitors could evolve as potential candidates to meet the clinical requirements in the future. In the present work, we have employed a virtual screening (VS) method to identify novel compounds. The VS scheme is comprised of pharmacophore filtering and, subsequently, receptor based screening. The VS scheme was applied to the databases of 1.04 million compounds to identify three novel compounds that can inhibit the IKKβ at a micro molar range. Moreover, these compounds can be raised into a potential anti-inflammatory drug candidate after optimizing and passing several phases of clinical trials.

© 2010 Published by Elsevier Ltd.

1. Introduction

The nuclear factor kappa-B (NF-κB) is a transcription factor required for the activation of several genes involved in the regulation of apoptosis and cell proliferation.[1] NF-κB is activated by a wide variety of agents, including phorbol esters, IL-1, TNF-α, lipopolysaccharide (LPS), double-stranded RNA, cAMP, bacteria, and viral trans activators.[1,2] The NF-κB activation requires the removal of IκB from NF-κB by inducible proteolysis, which liberates the transcription factor for migration to the nucleus, where it binds to the κB-regulatory elements and induces transcription.[2] The NF-κB exists in the cytoplasm in an inactive form associated with regulatory proteins called inhibitor of κB (IκB), of which the most important may be IκBα, IκBβ, and IκBε.[3] The IκB kinase (IKK) complex activation initiates the IκB phosphorylation, ubiquitination, and degradation by proteasomes, thereby releasing NF-κB/Rel dimers from the cytoplasmic NF-κB/Rel-IκB complex for nuclear translocation.[4]

The transcription factor NF-κB binds to the κB enhancer elements of the target genes, inducing the transcription of proinflammatory genes.[3] Activation and regulation of NF-κB is tightly controlled by IκB proteins. The understanding of the regulation of Rel/NF-κB activity is of great interest to a wide variety of basic biological and medical fields[5] in order to cure inflammatory related disorders. IKKβ (also known as IKK-2) is critical for NF-κB activation more than IKKα (also known as IKK-1).[6] Activation of IKKβ, rather than IKKα, participates in the primary pathway by which proinflammatory stimuli induce NF-κB function.[7] The classical NF-κB pathway has been closely linked to the orchestration of inflammatory responses and the survival of professional immune cells by coordinated expression of multiple inflammatory and innate immune genes.[8] The NF-κB activation has been shown to play a role in causing a wide range of diseases in humans, such as, cancer, Rhemotoid arthritis, diabetes, etc. However, its predominant role is causing immunosuppression and may result in chronic inflammatory diseases.[8] Birrell et al. demonstrated that the inhibition of the IKKβ activity completely blocks the release of cytokine released from cultured cells.[9] Some of the small molecules like sulfasalazine and its salicylate moiety 5-aminosalicylic acid, aspirin, and leflunomide are known to block the NF-κB action. Sulfasalazine and leflunomide block the nuclear translocation of NF-κB through the inhibition of IκBα degradation.[3] The inhibition of NF-κB activation by antioxidants and specific protease inhibitors may provide a pharmacological basis for interfering with these acute processes.[10]

A recent survey of the Integrity database reports that there are 230 compounds in the IKK inhibitor category in which only three inhibitors passed preclinical studies. EB-1627[11] (Fig. 1a) is a prodrug proceeding in the Phase I/II clinical trials, which is specifically designed as an anticancer agent. IMD-1041 is in Phase I, and, it was confirmed that IMD-1041 had a reliable safety and pharmacokinetic profile for further development. This inhibitor is mainly targeting the IKKβ system to control chronic obstructive pulmonary disease (COPD), one of the easiest of this inhibitor to work with as it can be orally administered.[12] Another inhibitor in Phase I is CHS-828[13–15] (Fig. 1b), a cyanoguanidine-based compound, which has demonstrated potent antitumor activity in preclinical tumor models.[16] In addition to Phase I/II, there are six compounds in preclinical trials and the rest are in a biological testing level. There is no marketed drug that targets IKKβ yet; on the other hand, a few compounds are only in the highest level of clinical trials. Therefore, demand is increasing for the compounds targeting the IKKβ.

Virtual screening is a complementary approach to experimental screening, and when it is coupled with structural biology, it promises to enhance success[17] called structure-based drug designing. The increased robustness of computational algorithms, scoring functions, availability of affordable computational power, and the potential of structural determination of the target molecule has provided new opportunities for the VS and has made it more practical.

Structure-based VS encompasses a variety of sequential computational phases, including target and database preparation, docking and post docking analysis, and prioritization for experimental testing.[17] The ligand-based method is more generalized and efficient when there is a strong correlation between the structural diversity and activity of known compounds. The structure information, activity, and non-arbitrary descriptors can be fed to the chemometric models to develop categorization rules. These pre-trained models can be used to rapidly classify the vast databases quantitatively or qualitatively. However, the ligand-based methods are often prone to misclassify due to the lack of access toward the realistic environment. Nevertheless, combining the ligand- and structure-based approaches under one scheme can overcome the artifact of the computational modeling.

In addition, VS is the rational method in choosing compounds, which are to be tested experimentally from the commercial database. A huge number of academics and researchers have been practicing this method over the years to discover novel scaffolds.[18–21] In the present work, we made an attempt to virtual screen the IKKβ inhibitors from the ChemDiv and Asinex databases. Approximately, 1.04 million compounds were used in the screening process. The databases supplied in an electronic version of the physically synthesized compounds by the suppliers. The VS protocol includes both the ligand- and structure-based approaches. The huge number of compounds are managed by pharmacophore filtering and followed by docking; this approach is helpful in reducing and prioritizing the compounds to be tested experimentally. A common feature pharmacophore model consist of five feature is generated using seven IKKβ inhibitors and a homology model generated as there is no experimentally derived structure for the IKKβ protein. Docking of the known inhibitors is helpful in validating the protocol as well as to calculate the enrichment factor.

From the VS cascade, we selected 60 compounds and those compounds are subjected to experimental screening, among which three compounds find to have IC$_{50}$ from 6.7 to 10.8 μM. In our lab, we have started the process of optimizing these compounds to locate the compounds having higher potency.

2. Results and discussions

2.1. Common feature hypothesis

In total, ten, five-feature hypothesis were generated. Each hypothesis was comprised of five features: one ring aromatic (RA), one hydrophobic (Hy), one hydrogen bond donor (HBD), and two hydrogen bond acceptor (HBA). In a pharmacophore model, the tolerance sphere defines the area in space that should be occupied by a specific type of chemical functionality. Among the ten hypothesis generated, the 8th ranked has shown to have the acceptor–donor feature in proper spacing. Therefore, this pharmacophore can bring out most of the compounds having acceptor and donor features together from the database which would facilitate the hinge region contact. The pharmacophore maps well with the training compounds and the fit value is always over 3.2. The fit value means that a minimum of three out of five features can map well with all training compounds (Table 1, Fig. 2). Hence, we applied the same criteria in database searching; the compounds

Table 1
Pharmacophore fit values corresponding to the training set compounds

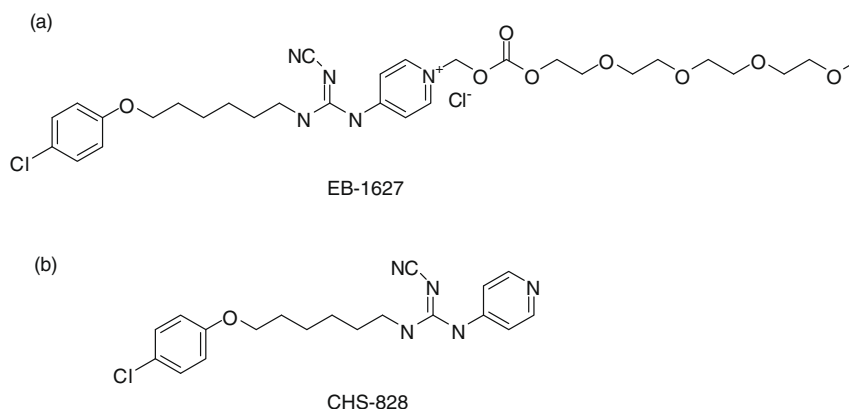| Compounds	Fit value	IC$_{50}$
2-Anilino-4-thiophenyl-pyrimidines derivative	3.75974	25 nM
Thiophenecarboxamides derivative	3.75839	25 nM
SPC-839	3.23021	62 nM
ML120B	3.38284	45 nM
Carboxyl-substituted 2-anilo-pyrimidines	3.46753	40 nM
Aztra Zeneca	5	40 nM
BI	3.3045	<1 μM

(a)

EB-1627

(b)

CHS-828

Figure 1. The compounds in the highest phase of clinical trial.
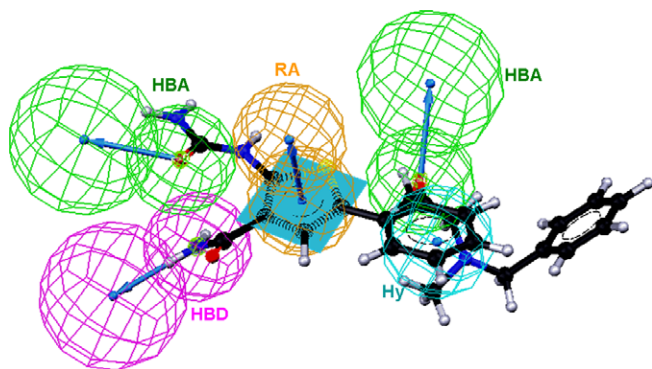
Figure 2. Five-feature pharmacophore mapped with one of the training set compounds (Aztra Zeneca compound), HBA; hydrogen bond acceptor, HBD; hydrogen bond donor, RA; ring aromatic, Hy; hydrophobic.

having a fit value below 3.5 were not allowed to pass the second filter of the VS.

The pharmacophore model was validated by searching with various multi-conformer databases including the IKK database (Table 2). The 8th ranked hypothesis was validated by searching the different databases designed for IKKβ, GSK-3 (a kinase database) and a 5HT$_{2C}$ agonist, which is a GPCR inhibitor category. The pharmacophore could generate 80% of the hits from the IKKβ database, 63% and 0.006% from GSK-3 and 5HT$_{2C}$ agonist databases, respectively. The pharmacophore is specific to the IKKβ compounds; however, a number of hits reported from the GSK database were not much less: this is mainly due to the commonalities among the kinase inhibitors. Most of the kinase compounds have similar chemical features, such as a hinge region interacting atoms. At the same time, in 5HT$_{2C}$ database number of compounds screened was negligible since the pharmacophore model is specific to IKKβ inhibitors. The validation insures the reliability of the pharmacophore hypothesis in VS of the commercial databases. However, introducing an additional filter to reduce the hits and prioritize the compounds to be screened in vitro is necessary as the pharmacophore searching yields too many compounds in the specified fit value criteria.

2.2. Homology model

The upgraded IKKβ model (Fig. 3) also exhibits well-defined β strands at the N-lobe and α helices in the C-lobe. In the same way, the G-loop forms a portion of an ATP binding pocket and the hinge region connects the N and C-terminal domains. Phosphoserine 177 and 181[22] are responsible for the IKKβ activation; these two serine residues placed in the middle portion of the A-loop, which is well exposed to cell sap. The homology model generated previously[23] had a root mean square deviation (RMSD) of 6.397 Å with the upgraded model as the model upgraded by new templates. We observed major differences at the ATP binding pocket, while the secondary elements of the C- and N-terminal domains remain unchanged. The hinge region residues were shifted down;

Table 2
Pharmacophore model validation result

Database searched	% of hit
IKKβ	80
GSK	63
5HT$_{2C}$ -Agonist	0.006

The pharmacophore is specific to the IKKβ database comparing to the GSK and 5HT$_{2C}$ database.
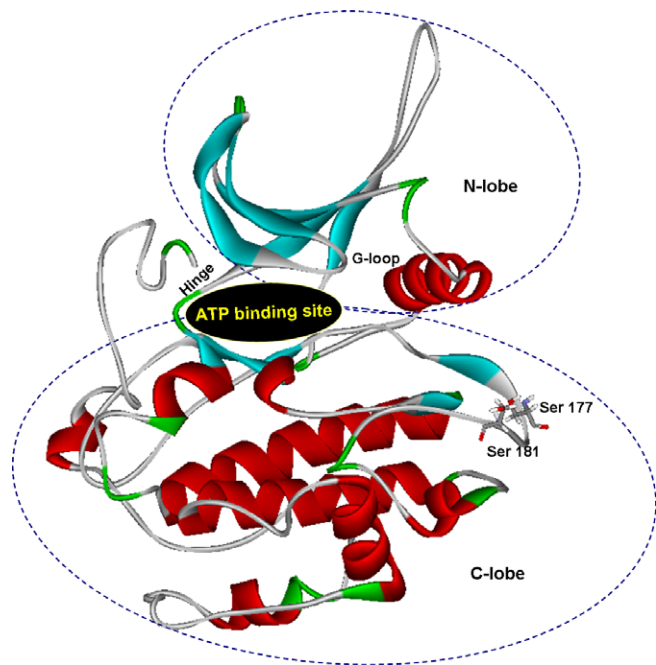

Figure 3. Architecture of IKKβ model shows a well-defined N- and C-lobe; phosphoserine 177 and 181 shown at the A-loop.

Glu172, considered to be important for the selectivity, was not seen in the ATP binding pocket in this model, and instead the residue was placed in the activation loop. Same way, Tyr169 pulled away from the ATP binding pocket and placed at the activation loop. The upgraded model exhibits Try169 is in π–π stacking with the imidazole ring of His143. The Phe26 of G-loop faces outwards from the ATP binding site, hence significantly reducing the hydrophobic interaction. Moreover, the present architecture of IKKβ represents the larger ATP binding pocket.

Among the three models generated, we chose the best model based on the following criteria for the structure-based VS. The DOPE score suggests model 2 as a most stable model, whereas the ProCheck result favors the model 1 (Table 3). On the other hand, docking studies conducted with three different models suggest the model 2 as a suitable conformation for VS, since more number of active compounds docked successfully. Among the 176 IKKβ inhibitors docked, 168 compounds were docked without any constraints, whereas, models 1 and 3 interact only with 73 and 3 inhibitors, respectively. However, the docking does not mean the ranking of the compounds; despite, this analysis accounts for the number of active compounds docked at ATP binding pocket. In the model 2, a few compounds are failed to dock at ATP binding pocket. Similarly, we docked the same 176 compounds with the Cys99 donor essential, Glu97 acceptor and Cys99 acceptor as optional constraints. A larger number of compounds docked with model 2 than the other models (Table 3). Moreover, the ATP molecule was stabilized properly at the binding site of model 2 when the molecule is docked with constraints; the adenine portion of the ATP makes a hinge region interaction. Table 3 summarizes the docking results of the IKKβ inhibitors and ATP molecule. Docking of an ATP molecule reveals that the binding site has a hydrophobic core to facilitate the adenine portion binding and the phosphate region of the ATP molecule binds to the solvent accessible area (Fig. 4).

2.3. Impact of ligand protonation

Table 4 represents the Goodness of Hits (GH) calculated for the protonated and normal compounds docking using FlexX; addition-

Table 3
Comparisons of three models; the second and third columns represent the DOPE score and ProCheck results, respectively

Model	DOPE score	ProCheck—Ramachandran plot		No. of compounds docked	Constraints ADA[c]	Constraints ADA[d]	Docking with ATP	
		a	b				No cons.	Cons
1	−26471.88	98.7	1.2	176	72	11	At binding site	No
2	−26952.14	97.6	2.4	176	168	144	At binding site	Properly placed in hinge ATP binding pocket.
3	−25929.33	97.6	2.4	176	3	9	Allosteric	No

The other columns summarize the docking analysis of known inhibitors and binding modes of an ATP molecule.
[a] Allowed region (%).
[b] Disallowed region (%).
[c] All the constraints set to be optional.
[d] Donor essential.
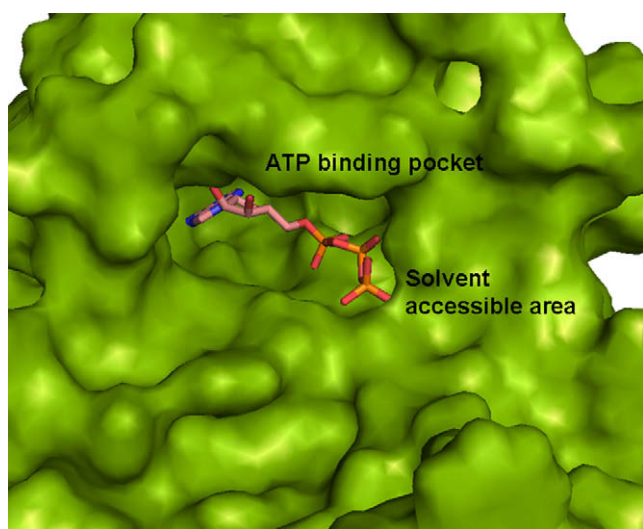
Figure 4. ATP molecule bound IKKβ homology model; the adenine portion forms interaction with the hinge region and the phosphate region extends to the solvent accessible area.

Table 4
Consolidation of the goodness of hit (GH) calculations based on the docking hits retrieved from protonated compounds, normal compounds and theoretical ideals

S. No.	Parameter	Protonated state	Normal compounds	Ideal
1	Total molecules in database (D)	540	540	540
2	Total number of actives in database (A)	40	40	40
3	Total hits (Ht)	43	46	40
4	Active hits (Ha)	29	30	40
5	% Yield of actives [(Ha/Ht) × 100]	67.441	65.217	100
6	% Ratio of actives [(Ha/A) × 100]	72.5	75	100
7	Enrichment factor (E) [($Ha \times D$)/($Ht \times A$)]	9.10	8.80	13.5
8	False negatives [$A - Ha$]	11	10	0
9	False positives [$Ht - Ha$]	14	16	0
10	Goodness of hit score	0.67	0.65	1

ally, an ideal GH score is provided to compare the results. There was no drastic change between protonated and normal library docking; however, protonated docking shows a slightly higher score, that is a 0.02 difference than normal docking. The percentage of true positive yielded and enrichment factor calculated is relatively higher in protonated docking. The overall hits returned by the protonated library docking were lesser, whilst composed of more active compounds. Therefore, we can state that protonating

the database before docking certainly improves the true positives rates in structure-based screening of IKKβ inhibitors.

2.4. Database screening

As a result of the pharmacophore screening of three databases containing ~1.04 million compounds yields 15,229 compounds. These compounds are retained based on the fit value threshold of ⩾3.5. To further filter the pharmacophore hits, we have performed two different docking screenings by means of running heavy and light docking methods in parallel. Finally, 352 compounds were retained from the heavy constrained docking and 4193 compounds from the light constrained docking (Table 5). The compounds were sorted according to F-score; the heavy constrained docking score began from −33.233 to 2.018, whereas light constrained docking scores range from −36.953 to 5.477. The top scoring 100 compounds were selected from each docking category, and then some compounds are eliminated based on the visual inspection of the interactions. At last only 60 compounds that are 0.0058% of the database compounds are subjected to in vitro screening and for this purpose, we used IKKβ-TR-FRET reactions to measure the inhibition rate of the compounds (Fig. 5).

2.5. Novel scaffold

Of the 60 compounds screened in the IKKβ enzyme inhibition assay, three compounds, VH01, VH02 and VH03 showed 84.8%, 76.4% and 53.1% of inhibition, respectively, at 10 μM (Table 6, Fig. 6). VH01 ((E)-4-(4-((5-(1,3-dioxoisoindolin-5-yl)furan-2-yl)methylene)-3-methyl-5-oxo-4,5-dihydro-1H-pyrazol-1-yl)benzenesulfonamide) had the highest inhibition rate of 84.8% and an IC$_{50}$ value of 6.7 μM. VH02 ((E)-N-(5-(2-(2-(4-hydroxy-3-methoxy-benzylidene)hydrazinyl)-2-oxoethyl)-1,3,4-thiadiazol-2-yl)-4-nitrobenzamide) was able to inhibit 76.4% and had an IC$_{50}$ value of 6.98 μM. Whereas the compound VH03 ((E)-N-(5-(2-(2-(4-hydroxy-3-methoxybenzylidene)hydrazinyl)-2-oxoethyl)-1,3,4-thiadiazol-2-yl)-4-methoxybenzamide) had a less inhibitory effect of 53.1% and an IC$_{50}$ value of 10.8 μM. To monitor the assay system, we used a Bayer's compound (Bayer-5a) as a positive control, the IC$_{50}$ value measured to this compound by Murata et al.[24] is 25 nM. Whereas the TR-FRET analysis reports the compound's IC$_{50}$ value of 0.174 μM, the screening system we used reports that the Bayer compound has 6.96 folds lesser potency, owing to the different assay system. Based on the IC$_{50}$ value measured to a positive control, we expected the hit compounds to inhibit more potently in the recombinant human IKKβ inhibition assay. It is worth noting that all the active compounds reported were screened from the light constrained docking, none of the heavy constrained docking were shown to retrieve a potential inhibitor. Clearly the light constraint is capable to overcome the artifact of the heavy

Table 5
Virtual screening summary of each database

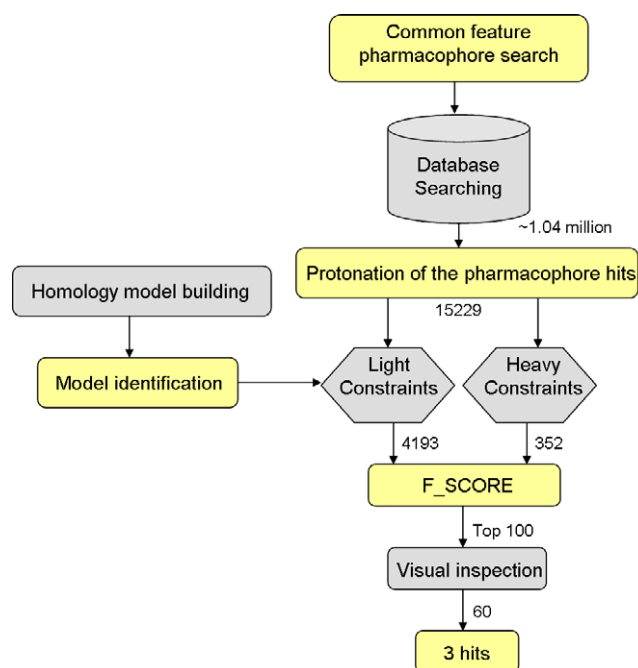| Database	Size in million	Pharmacophore hits	Protonation	Docking hits		Biological screening hits
				Heavy	Light	
ChemDiv	0.7	7000	7101	46	1248	2
Asinex—Gold	0.22	4655	4728	86	1831	1
Asinex—Platinum	0.12	3574	3672	220	1114	0

Figure 5. Workflow of the virtual screening scheme.

constraint. Among the three databases screened, the Asinex platinum databases did not yield any active compounds, whereas VH01 and VH03 from the ChemDiv database and VH02 from Asinex Gold. Although the VH02 and VH03 do not belong to the same database, the VS protocol is efficient in screening similar compounds from different databases enriched with a large number of compounds. The three hit compounds were highly diverse among the known IKKβ inhibitors, similarity analysis revealed that none of the hit compounds had a tanimoto similarity index greater than 0.34. Hence, it is clear that the hit compounds are novel, diverse, and opens a new door to IKKβ drug discovery research.

2.6. Interaction of the hit compounds

All hit compounds have been identified from light constrained docking; hence these compounds are capable of forming two hinge region hydrogen bonds (Fig. 7). Nitrogen (NH) in isoindo-line-dione of the VH01 forms a hydrogen bond with Glu97 carbonyl oxygen and isoindoline oxygen acts as a hydrogen bonding acceptor with the Cys99 donor (backbone NH). In case of VH02, the methoxyphenol forms a hinge region interaction, the hydroxyl portion donates an electron to Glu97, and the methoxy oxygen accepts an electron from Cys99 to form a hydrogen bonding. VH03 also forms a similar manner of binding, while the hydroxyl portion makes two hydrogen bonds with Glu97 and Cys99, despite the fact that no conspicuous hydrophobic interaction was observed. In addition to the hinge region interaction, there are several other hydrogen bonds formed between the lead compounds and receptor.

VH01: (Fig. 7a) pyrazolon nitrogen forms a hydrogen bond with the Lys106 side-chain amino group; sulfonamide oxygen accepts an electron from the Arg105 side-chain guanidinium group. The amino group is a hydrogen bond donor to the Lys106 carbonyl of the amide portion.

VH02: (Fig. 7b) amide nitrogen of this compound is a hydrogen bond donor to the Lys106 backbone carbonyl, another carbonyl group of the amide portion accepts an electron to form a hydrogen bond with Lys44 side-chain amino group.

VH03: (Fig. 7c) shares similarities with the VH02, but has a nitrobenzene group in the place of anisole. The nitro group oxygen interacts with the guanidinium group of Arg105 and the amide nitrogen forms a hydrogen bond with the carbonyl oxygen of Gly168. Amide group nitrogen forms an interaction with the side-chain carboxylic oxygen of Asp103.

3. Conclusion

Virtual screening is a complement method to a high throughput screening, a computation method to select the promising compounds from large databases and focused libraries to in vitro screening. The VS scheme comprised of ligand- and structure-based approach led the way to discover novel inhibitors. The effectiveness of the VS protocol has been explained by the pharmacophore constraints. To the best of our knowledge, this is the first VS approach of IKKβ inhibitors that identifies three novel compounds and becomes a basis for diverse and novel scaffolds. The hit compounds VH01, VH02, and VH03 had IC$_{50}$ values of 6.7, 6.98 and 10.8 μM, respectively. Among the 60 compounds tested three compounds are potentially inhibiting the IKKβ enzyme, the hit compounds ratio (1:20) meticulously validates and supports the effectiveness of the VS protocol. The hit compounds identified were novel among the known IKKβ inhibitors and further optimization of these lead compounds could result in potential non-steroidal anti-inflammatory drugs (NAIDs).
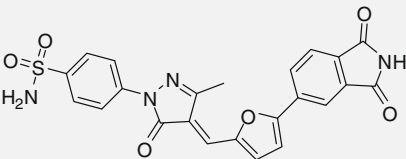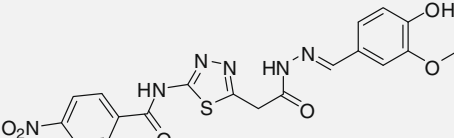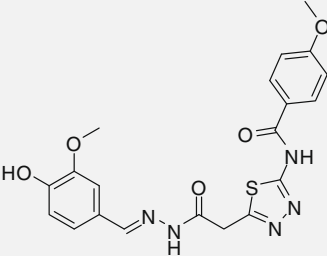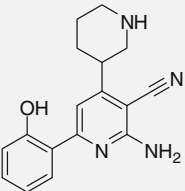
4. Materials and methods

4.1. Pharmacophore generation

The most critical task in the drug discovery process is developing an appropriate model to predict the activity of given molecules. However, most of the QSAR models have lapses that can lead to erroneous predictions. In such a scenario, the application of a common feature pharmacophore would be of great help in retaining the features required for the activity as it is a non-QSAR method. Despite the advantage, it can also report a large number of false positives; hence in a VS scheme a secondary filter is necessary to refine the hits.

A total of seven compounds[25–29] were used in a common feature model generation (Fig. 8). The main reasons to choose these compounds in pharmacophore modeling are sensitivity and selectivity, the compounds are tend to inhibit IKKβ at nano molar concentration, and having several fold selectivity over IKKα. The Catalyst 4.11(Accelrys, Inc., San Diego) module was used for pharmacophore modeling. Possible conformers were generated to all

Table 6
Virtual screening hit compounds and its corresponding docking score IC$_{50}$ values

| Code	Compound	F-Score	% of inhibition at 10 μM	IC$_{50}$ in μM
VH01	 (E)-4-(4-((5-(1,3-Dioxoisoindolin-5-yl)furan-2-yl)methylene)-3-methyl-5-oxo-4,5-dihydro-1H-pyrazol-1-yl)benzenesulfonamide	−27.248	84.8	6.7
VH02	 (E)-N-(5-(2-(2-(4-Hydroxy-3-methoxybenzylidene)hydrazinyl)-2-oxoethyl)-1,3,4-thiadiazol-2-yl)-4-nitrobenzamide	−27.830	76.4	6.98
VH03	 (E)-N-(5-(2-(2-(4-Hydroxy-3-methoxybenzylidene)hydrazinyl)-2-oxoethyl)-1,3,4-thiadiazol-2-yl)-4-methoxybenzamide	−22.863	53.1	10.8
Bayer-5a		ND	ND	0.174

ND—Not determined.
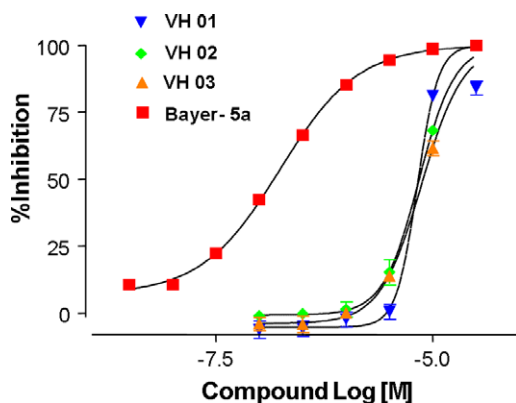
Figure 6. Dose-dependent inhibition measured by TR-FRET analysis of IKKβ phosphorylation. The IC$_{50}$ values reported for virtual screening hits and Bayer-5a as a positive control.

the compounds by best conformer generation method with a minimum energy range of 20 kcal/mol. Desired feature definition is an important stage in the pharmacophore model generation. All the training set compounds consist of at least one hydrogen bond acceptor (HBA), a hydrogen bond donor (HBD), and a ring aromatic (RA) feature. In addition, some of the compounds have a hydrophobic (Hy) feature. Based on the training set observations, we decided to define the important four features. As we know, the kinase compounds tend to make a hinge contact.[30–32] Hinge loop is a small segment that connects the N- and C-terminal lobes of any kinase, in the case of IKKβ Glu97 and Cys99 forms the acceptor–donor–acceptor interaction; hence it is apparent that the IKKβ inhibitor must have the feature to facilitate the hinge interaction. On that basis, we insisted at least one donor and acceptor feature should be retained in the pharmacophore model. As a consequence of various trials conducted to optimize the pharmacophore generation, we identified the best pharmacophore when the importance was given to ML120B and Aztra Zeneca by setting the principal values of 2 and 1 to other training sets. The value 2 insures that all of the chemical features in the compound are considered for building hypothesis space. As per our features setting, the model can have a maximum of ten features and should have at least four features. We modified the spacing to 100 picometer in order to achieve the acceptor–donor features together, and the rest of the parameters set to default, a total of ten pharmacophores were generated.

4.2. Homology model building

We explained the usefulness of the IKKβ homology model in VS of our previous work.[23] In the current VS scheme, we updated the homology model using the same procedure with three new tem-
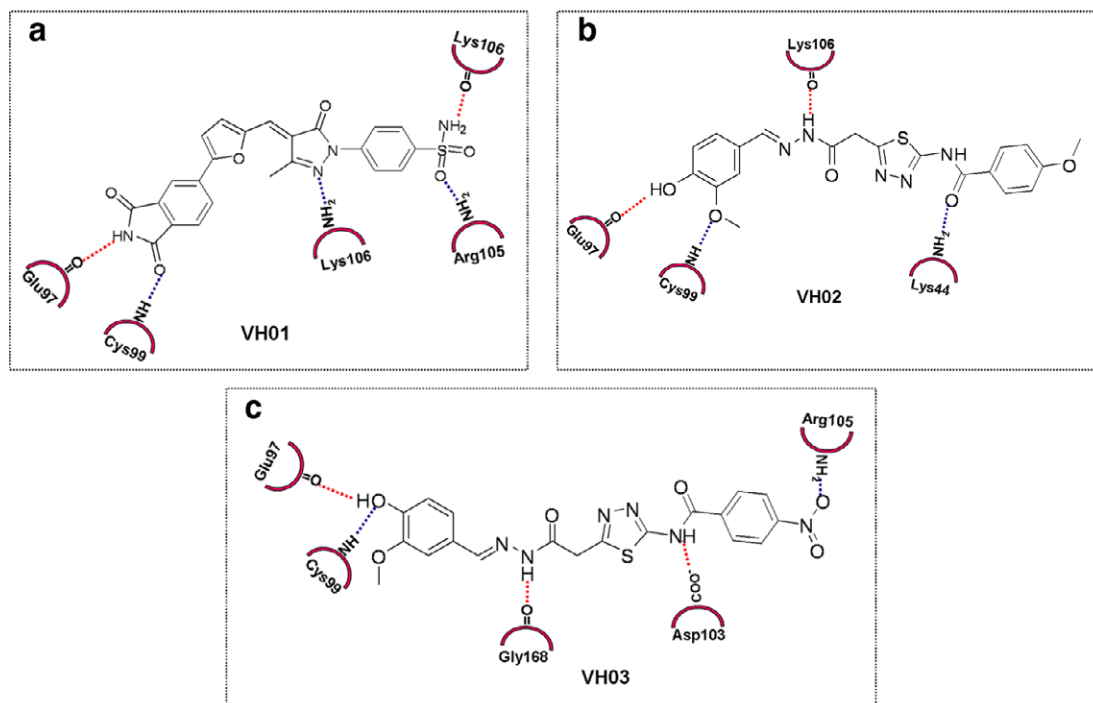
Figure 7. 2D representation of hydrogen bonding interaction of hit compounds. The red dotted line represents the hydrogen bond accepting nature of receptor and the blue dotted line represents the hydrogen bond donating nature.
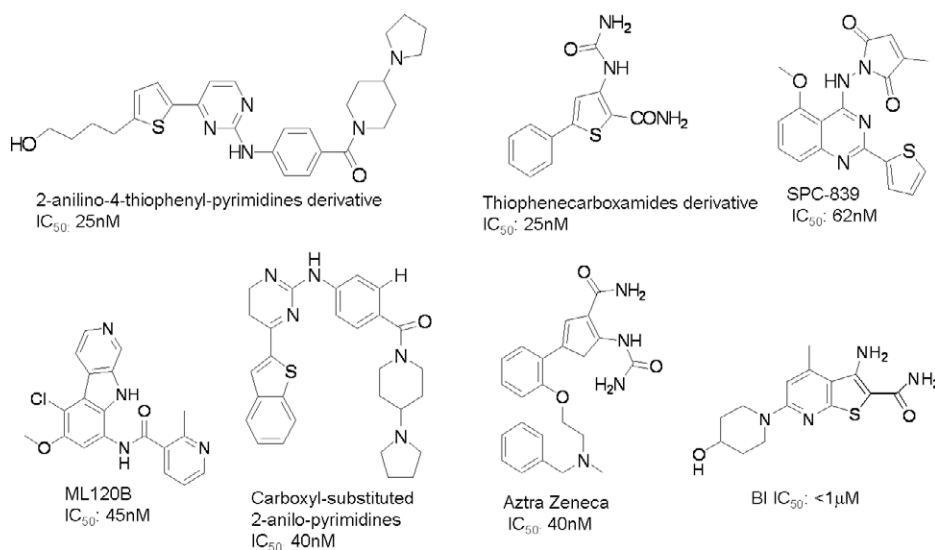

Figure 8. The seven compounds used as training sets to model a common feature.

plates (Table 7). Employing more than one template increase the backbone flexibility and thus expanding the accessible conformational space.[33] The automated sequence alignment and analysis of the template and target is carried out using the Clustal X 1.81[34] and the GeneDoc 2.6[35] programs. The alignment depicts the G-loop and the hinge regions have a high sequence similarity, whereas the activation loop (A-loop) has lesser similarity (Fig. 9). However, there is no appropriate crystal structure available to model the A-loop. Besides, A-loop does not interact with the ATP-competitive inhibitors and modeling the accurate loop conformation is a difficult task in the homology modeling approach as the loop region is highly flexible.

Table 7
Templates used to model an IKKβ protein

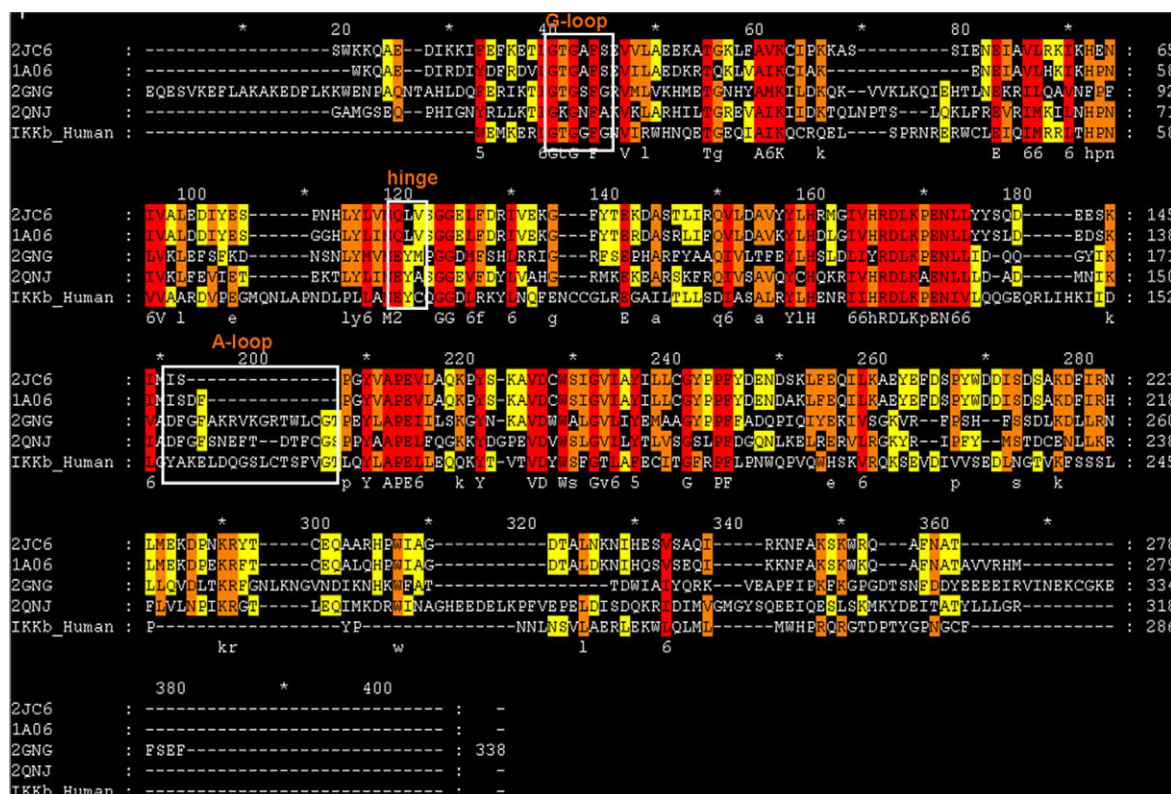| PDB	Identity (%)	Protein function	Resolution in Å
2JC6	31	Crystal structure of human calmodulin-dependent protein kinase 1D	2.30
1A06	29	Calmodulin-dependent protein kinase from rat	2.50
2QNJ	32	Kinase and ubiquitin-associated domains of MARK3/Par-1	2.70
2GNG	29	Protein kinase A—fivefold mutant model of Rho-kinase	1.87

Figure 9. The sequence alignment held between IKKβ protein sequence and the template sequence. G-loop, Hinge and A-loop regions are indicated in the box.

We used Discovery Studio 2.0 (http://accelrys.com/) program to homology model an IKKβ protein, which uses a built-in Modeler[36] package. The four different kinase template sequences[37–40] were aligned together with the IKKβ sequence and this alignment was supplied along with the 3D coordinates of templates as an input to the program. The Modeler program implements the comparative protein structure modeling by satisfying the spatial restraints.[36,41,42] A 3D model was obtained by optimization of a molecular probability density function (PDF). The molecular PDF for comparative modeling is optimized with the variable target function procedure in cartesian space that employs methods of conjugate gradients and molecular dynamics with simulated annealing.[36] To assess the quality of the minimized models, a Profunc procheck[43–45] analysis was employed. Prior to docking, the Aspartic acid, glutamic acid, lysine, and arginine residues were protonated according to the physiological pH (7.4) level.

4.3. Homology model selection

In order to choose the suitable model to structure-based virtual screening, we have conducted various analyses to rank order the best model. To quickly access the models the discrete optimized protein energy (DOPE) score and ProCheck Ramachandran plot has been observed. Furthermore, to analyze the detailed quality of the models we have carried out exhaustive docking studies using 176 known IKKβ inhibitors and ATP molecule, the docking simulation performed using FlexX. The inhibitors and ATP molecule has been docked in two trails with and without the hydrogen bond constraints (Table 5).

4.4. Database preparation

We used two databases, ChemDiv and Asinex (gold and platinum), to VS. These databases had approximately 1.04 million com-

pounds. They are fragmented and drug-like building blocks which have several points of diversity. The compounds are supplied with validated quality, pre-filtered by drug-likeness properties, no duplicate records and the compounds can be supplied in a short period of time. The pharmacophore screening requires a multi-conformer database; therefore, we generated a minimum of 300 conformers to every compound. The database contains conformations of separate isomers generated within the relative energy threshold of 20.0 kcal/mol. The Discovery Studio 2.1 is used to generate a database in a catalyst format; this program uses the CatDB module of Catalyst. The database automatically indexed with substructure, pharmacophore features, and shape information. As the pharmacophore searching reported the hits in 3D space, there was no need to convert the structure into 3D coordinates prior to docking. However, all the compounds low energy conformers were generated and chirality was set right. In addition, every compound was ionized at physiological pH 7.4 and tautomers were enumerated in the library. The LigPrep[46] tool of Maestro 8.0 was used to prepare the compounds. Polgar et al.[47] demonstrated the VS by FlexX-Pharm provided the best results and showed dependency on ligand protonation.

4.5. Docking procedure

The second filter used in the VS scheme was a structure-based approach, Flex-X program was employed to dock the compounds. Before docking, hydrogen atoms were added to the protein and minimized using the Steepest Descent algorithm for about 500 steps. The amino acids Phe26, Val29, Ala42, Lys44, Met65, Val74, Ala76, Glu97, Tyr98, Cys99, Lys106, Val152, Gly168 and the surrounding residues within the distance range of 6.5 Å were defined as active sites. Flex-X[48] uses an incremental construction algorithm to place flexible ligands into a fully specified active site; its empirical-based scoring function estimates the binding free energy

based on the physicochemical properties. The FlexX-Pharm[49] module was used to define the constraints and to direct the FlexX docking of several compounds into the specified active site simultaneously. FlexX-Pharm insures that an interaction is formed between the specified interacting group in the active site and the ligand in a valid docking solution. There are many research groups who successfully employed the constraints in structure-based VS, which increases the enrichment factor[19,47,50,51] of active compounds. As we know, most of the ATP-competitive inhibitors of kinase make two or three hinge[52–54] region hydrogen bonding interactions. Hence, we applied the hydrogen bonding ability of a molecule as constraints to filter out the compounds that can possibly make hinge contacts. In the docking simulation two different sets of constraints were applied, namely 'Heavy' and 'Light' (Fig. 10). The Heavy constraint is strict in choosing the compounds. According to this specification, compounds forming three hydrogen bonds with the hinge region acceptor (Glu97)–donor (Cys99)–acceptor (Cys99) are only reported as hits. Whereas, in the Light constraint the middle donor interaction is essential and at least one acceptor hydrogen bonding interaction is essential. The resultants of the pharmacophore search were submitted to the docking protocol as a library in structure data format (SDF). A maximum of 30 conformers were retained to each compound passing the constraints criteria. In our previous work,[23] we demonstrated that the f-scoring function was capable of discriminate the IKKβ inhibitors from the decoys; henceforth we applied the f-scoring function in the current VS as well.

4.6. Goodness of hits calculation

We measured the efficiency of the protonated docking by calculating GH.[55] The GH is a measure that accounts for the efficiency of the model in VS. It can account for both the % of yield (the fraction of active structures hit) and % of actives that are retrieved from the database. In this experiment, we used the same set of compounds that were used to calculate the enrichment factor (EF)[23] in our previous study. A small library, consisting of 40 IKKβ inhibitors is used as actives and 500 compounds selected from the ChemDiv database were used as decoys and were used in the GH calculation. The library was protonated using the LigPrep[46] module and only the best possible protonation state was retained in the library. Two different docking calculations were performed simultaneously with the protonated and the normal library. In addition, an ideal GH score was calculated theoretically for the highest possible

goodness. The ideal GH score was calculated by assuming the model reports only true positives as hits.

4.7. Virtual screening

A five-feature pharmacophore model developed, and used to search the ChemDiv and the Asinex subsets database such as Gold and Platinum. The databases have ~1.04 million compounds (Fig.10); a pharmacophore search results were sorted by fit value. Fit value is criteria that measure how well a compound that can fit into the pharmacophore. The Fast method was used in the pharmacophore search, which performs a rigid fit of the conformers against the pharmacophore. In each database, compounds having a pharmacophore fit value $\geqslant 3.5$ were considered for docking based screening. Prior to performing docking simulations, we prepared the library as results of pronation, tautomer enumeration and energy minimization. The prepared library was used in the heavy as well as light docking. We further limited the compounds to be tested experimentally based on the visual inspection and commercial availability.

4.8. In vitro analysis: IKKβ enzyme inhibition assay

IKKβ-TR-FRET reactions for the search of the IKKβ inhibitors were carried out based upon the suggestions of the IMAP-TR-FRET system (MDS Analytical Technologies, Sunnyvale, Calif., USA). IKKβ kinase reactions were performed in a reaction buffer (10 mM Tris–HCl, pH 7.2, 10 mM $MgCl_2$, 0.05% NaN_3), containing 1 mM DTT and 0.01% Tween-20 (Sigma–Aldrich Co., St. Louis, Mo., USA) to help stabilize the enzyme. The reactions were done at room temperature for 2 h in white standard 384 plates (3572, Corning Life Sciences, Lowell, Mass., USA), using 0.5 μg/ml IKKβ (Millipore Co., Billerica, Mass., USA), 1 μM $I_KB\alpha$-derived substrate (5FAM-GRHDSGLDSMK-NH2; R7574, MDS Analytical Technologies), and 3 μM ATP (Sigma–Aldrich Co.) unless otherwise noted. The total reaction volumes were 20 μl and 10 μM and compounds were pre-incubated with IKKβ enzyme for 10 min before the substrate and ATP were added. For the TR-FRET reaction, 60 μl of the detection mixture (1:600 dilution of IMAP binding reagent and 1:400 dilution of Terbium donor supplied by MDS Analytical Technologies) was added 15 h before reading the plate. The energy transfer signal was measured in a multilabel counter with a TR-FRET option (Victor II, PerkinElmer Oy, Turku, Finland). The counter setting was 340 nm excitation, 100 μs delay, and dual-emission collection for 200 μs at 495 and 520 nm. The energy transfer signal data were used to calculate the percentage inhibition and IC_{50} values.

References and notes

1. Kishore, N.; Sommers, C.; Mathialagan, S.; Guzova, J.; Yao, M.; Hauser, S.; Huynh, K.; Bonar, S.; Mielke, C.; Albee, L.; Weier, R.; Graneto, M.; Hanau, C.; Perry, T.; Tripp, C. S. *J. Biol. Chem.* **2003**, *278*, 32861.
2. Jobin, C.; Sartor, R. B. *Am. J. Phys. Cell Phys.* **2000**, *278*, C451.
3. Tak, P. P.; Firestein, G. S. *J. Clin. Invest.* **2001**, *107*, 7.
4. Tak, P. P.; Gerlag, D. M.; Aupperle, K. R.; Aupperle, K. R.; van de Geest, D. A.; Overbeek, M.; Bennett, B. L.; Boyle, D. L.; Manning, A. M.; Firestein, G. S. *Arthritis Rheum.* **2001**, *44*, 1897.
5. Sizemore, N.; Agarwal, A.; Das, K.; Lerner, N.; Sulak, M.; Rani, S.; Ransohoff, R.; Shultz, D.; Stark, G. R. *PNAS* **2004**, *101*, 7994.
6. Ying, z.; Ryan, M. J.; Webb, R. C.; Hall, J. E. *FASEB J.* **2007**, *21*, LB115.
7. Delhase, M.; Hayakawa, M.; Chen, Y.; Karin, M. *Science* **1999**, *284*, 309.
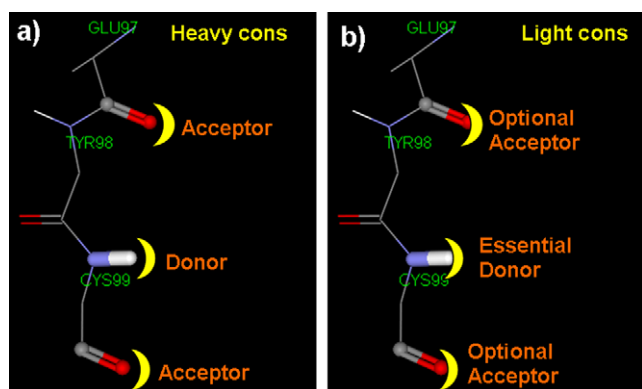8. Uwe, S. *Biochem. Pharmacol.* **2008**, *75*, 1567.

Figure 10. Pictorial representation of constraints. (a) 'Heavy' constraint—all three hinge hydrogen bonds are required for successful docking. (b) 'Light' constraint—middle donor is essential and flanking hydrogen bonds are optional, yet according to the light constraint definition anyone of the acceptor interactions is a 'must' along with the donor interaction.

9. Birrell, M. A.; Hardaker, E.; Wong, S.; McCluskie, K.; Catley, M.; De Alba, J.; Newton, R.; Haj-Yahia, S.; Pun, K. T.; Watts, C. J.; Shaw, R. J.; Savage, T. J.; Belvisi, M. G. *Am. J. Res. Criti. Care Med.* **2005**, *172*, 962.
10. Baeuerle, P. A.; Henkel, T. *Annu. Rev. Immunol.* **1994**, *12*, 141.
11. Binderup, E.; Bjorkling, F.; Hjarnaa, P. V.; Latini, S.; Baltzer, B.; Carlsen, M.; Binderup, L. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 2491.
12. http://www.immd.co.jp/en/news/news20060824.html. 2009.
13. Hjarnaa, P. J. V.; Jonsson, E.; Latini, S.; Dhar, S.; Larsson, R.; Bramm, E.; Skov, T.; Binderup, L. *Cancer Res.* **1999**, *59*, 5751.
14. Lovborg, H.; Wojciechowski, J.; Larsson, R.; Wesierska-Gadek, J. *Cancer Res.* **2002**, *62*, 4206.
15. Olsen, L. S.; Hjarnaa, P. J. V.; Latini, S.; Holm, P. K.; Larsson, R.; Bramm, E.; Binderup, L.; Madsen, M. W. *Int. J. Cancer* **2004**, *111*, 198.
16. Hovstadius, P.; Larsson, R.; Jonsson, E.; Skov, T.; Kissmeyer, A. M.; Krasilnikoff, K.; Bergh, J.; Karlsson, M. O.; Lonnebo, A.; Ahlgren, J. *Clin. Cancer Res.* **2002**, *8*, 2843.
17. Lyne, P. D. *Drug Discovery Today* **2002**, *7*, 1047.
18. Kovac, A.; Konc, J.; Vehar, B.; Bostock, J. M.; Chopra, I.; Janezic, D.; Gobec, S. *J. Med. Chem.* **2008**, *51*, 7442.
19. Polgar, T.; Keseru, G. M. *J. Med. Chem.* **2005**, *48*, 3749.
20. Cho, Y.; Ioerger, T. R.; Sacchettini, J. C. *J. Med. Chem.* **2008**, *51*, 5984.
21. Kiss, R.; Kiss, B.; Konczol, A.; Szalai, F.; Jelinek, I.; Laszlo, V.; Noszal, B.; Falus, A.; Keseru, G. M. *J. Med. Chem.* **2008**, *51*, 3145.
22. Mercurio, F.; Zhu, H. Y.; Murray, B. W.; Shevchenko, A.; Bennett, B. L.; Li, J. W.; Young, D. B.; Barbosa, M.; Mann, M. *Science* **1997**, *278*, 860.
23. Nagarajan, S.; Doddareddy, M. R.; Choo, H.; Cho, Y. S.; Oh, K. S.; Lee, B. H.; Pae, A. N. *Bioorg. Med. Chem.* **2009**, *17*, 2759.
24. Murata, T.; Shimada, M.; Kadono, H.; Sakakibara, S.; Yoshino, T.; Masuda, T.; Shimazaki, M.; Shintani, T.; Fuchikami, K.; Bacon, K. B.; Ziegelbauer, K. B.; Lowinger, T. B. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 4013.
25. Baxter, A.; Brough, S.; Cooper, A.; Floettmann, E.; Foster, S.; Harding, C.; Kettle, J.; McInally, T.; Martin, C.; Mobbs, M.; Needham, M.; Newham, P.; Paine, S.; St-Gallay, S.; Salter, S.; Unitt, J.; Xue, Y. F. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 2817.
26. Coish, P. D. G.; Wickens, P. L.; Lowinger, T. B. *Expert Opin. Ther. Patents* **2006**, *16*, 1.
27. Karin, M.; Yamamoto, Y.; Wang, Q. M. *Nat. Rev. Drug Disc.* **2004**, *3*, 17.
28. Strnad, J.; Burke, J. R. *Trends Pharmacol. Sci.* **2007**, *28*, 142.
29. Waelchli, R.; Bollbuck, B.; Bruns, C.; Buhl, T.; Eder, J.; Felfel, R.; Hersperger, R.; Janser, P.; Revesz, L.; Zerwes, H. G.; Schlapbach, A. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 108.
30. Breitenlechner, C.; Gassel, M.; Hidaka, H.; Kinzel, V.; Huber, R.; Engh, R. A.; Bossemeyer, D. *Structure* **2003**, *11*, 1595.
31. Liu, Y.; Gray, N. S. *Nat. Chem. Biol.* **2006**, *2*, 358.
32. Zhang, J. M.; Yang, P. L.; Gray, N. S. *Nat. Rev. Cancer* **2009**, *9*, 28.
33. Kneissl, B.; Leonhardt, B.; Hildebrandt, A.; Tautermann, C. S. *J. Med. Chem.* **2009**, *52*, 3166.
34. Thompson, J. D.; Gibson, T. J.; Plewniak, F.; Jeanmougin, F.; Higgins, D. G. *Nucleic Acids Res.* **1997**, *25*, 4876.
35. Nicholas, H. B., Jr.; Graves, S. B. *J. Mol. Biol.* **1983**, *171*, 111.
36. Sali, A.; Blundell, T. L. *J. Mol. Biol.* **1993**, *234*, 779.
37. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235.
38. Bonn, S.; Herrero, S.; Breitenlechner, C. B.; Erlbruch, A.; Lehmann, W.; Engh, R. A.; Gassel, M.; Bossemeyer, D. *J. Biol. Chem.* **2006**, *281*, 24818.
39. Goldberg, J.; Nairn, A. C.; Kuriyan, J. *Cell* **1996**, *84*, 875.
40. Murphy, J. M.; Korzhnev, D. M.; Ceccarelli, D. F.; Briant, D. J.; Zarrine-Afsar, A.; Sicheri, F.; Kay, L. E.; Pawson, T. *PNAS* **2007**, *104*, 14336.
41. Fiser, A.; Do, R. K. G.; Sali, A. *Protein Sci.* **2000**, *9*, 1753.
42. Marti-Renom, M. A.; Stuart, A. C.; Fiser, A.; Sanchez, R.; Melo, F.; Sali, A. *Ann. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 291.
43. Laskowski, R. A.; Macarthur, M. W.; Moss, D. S.; Thornton, J. M. *J. Appl. Cryst.* **1993**, *26*, 283.
44. Laskowski, R. A.; Watson, J. D.; Thornton, J. M. *Nucleic Acids Res.* **2005**, *33*, W89.
45. Morris, A. L.; Macarthur, M. W.; Hutchinson, E. G.; Thornton, J. M. *Protein Struct. Funct. Genet.* **1992**, *12*, 345.
46. LigPrep. http://www.schrodinger.com/ProductDescription.php?mID=6&sID=704/09.
47. Polgar, T.; Magyar, C.; Simon, I.; Keseru, G. M. *J. Chem. Inf. Model.* **2007**, *47*, 2366.
48. Tripos Inc., S. L., MO 63144, USA.
49. Hindle, S. A.; Rarey, M.; Buning, C.; Lengauer, T. *J. Comput. Aided Mol. Des.* **2002**, *16*, 129.
50. Polgar, T.; Baki, A.; Szendrei, G. I.; Keseru, G. M. *J. Med. Chem.* **2005**, *48*, 7946.
51. Virag, I.; Polgar, T.; Keseru, G. M. *J. Mol. Struct. (Theochem.)* **2005**, *725*, 239.
52. Aronov, A. M.; McClain, B.; Moody, C. S.; Murcko, M. A. *J. Med. Chem.* **2008**, *51*, 1214.
53. Ghose, A. K.; Herbertz, T.; Pippin, D. A.; Salvino, J. M.; Mallamo, J. P. *J. Med. Chem.* **2008**, *51*, 5149.
54. Liao, J. J. L. *J. Med. Chem.* **2007**, *50*, 409.
55. Jones, G.; Willet, P. In *Phamacophore Perception, Development, and Use in Drug Design*; Osman, F. G. D., Ed.; International University Line: La Jolla, 2000; pp 13–43.