Journal of Molecular Graphics and Modelling 29 (2010) 72–81

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

IKK-β inhibitors: An analysis of drug–receptor interaction by using Molecular Docking and Pharmacophore 3D-QSAR approaches

Antonino Lauria[a,*], Mario Ippolito[a], Marco Fazzari[a,b,c], Marco Tutone[a], Francesco Di Blasi[c], Francesco Mingoia[b], Anna Maria Almerico[a]

[a] Dipartimento Farmacochimico, Tossicologico e Biologico, Università di Palermo, Via Archirafi 32, 90123 Palermo, Italy
[b] Istituto per lo Studio dei Materiali Nanostrutturati, CNR, Via Ugo La Malfa 153, 90146 Palermo, Italy
[c] Istituto di Biomedicina Molecolare, CNR, Via Ugo La Malfa 153, 90146 Palermo, Italy

ARTICLE INFO

Article history:
Received 31 December 2009
Received in revised form 28 April 2010
Accepted 30 April 2010
Available online 10 May 2010

Keywords:
NF-kB
IKK-β
Binding Database
Homology modeling
Molecular Docking
Pharmacophore

ABSTRACT

The IKK kinases family represents a thrilling area of research because of its importance in regulating the activity of NF-kB transcription factors. The discovery of the central role played by IKK-β in the activation of transcription in response to apoptotic or inflammatory stimuli allowed to considerate its modulation as a promising tool for the treatment of chronic inflammation and cancer. To date, several IKK-β inhibitors have been discovered and tested. In this work, an analysis of the interactions between different classes of inhibitors and their biological target was performed, through the application of Molecular Docking and Pharmacophore/3D-QSAR approaches to a set of 141 inhibitors included in the Binding Database. In order to overcome the difficulty due to the lack of crystallographic data for IKK-β, a homology model of this protein has been built and validated. The results allowed to study in depth the structural bases for the interaction of each family of inhibitors and provided clues for further modifications, with the aim of improving the activity and selectivity of designed drugs targeting this enzyme.

1. Introduction

The Nuclear Factor-kB (NF-kB) proteins are a group of highly related dimeric transcription factors which regulate the immune and inflammatory response, as well as the apoptotic pathways [1–3]. The association with inhibitory proteins, belonging to the IkB family, retains the NF-kB dimer in the cytoplasm and prevents its activation by masking a nuclear localization signal (NLS) [4]. In response to inflammatory, physical and chemical stresses, IkB is phosphorylated by the IKK kinase and addressed to the proteasomic degradation, while the NF-kB dimer migrates into the nucleus and binds to specific DNA sequences, regulating thus the transcription of target genes [5].

The IKK family of kinases represent a thrilling area of research because of its importance in regulating the activity of NF-kB transcription factors. The IKK complex is constituted by three subunits [6–8]: the two catalytic subunits IKK-α and IKK-β, and the regulatory subunit IKK-γ (also termed as NEMO). The discovery of the central role played by IKK-β in the activation of transcription in response to apoptotic or inflammatory *stimuli* allowed to consider its inhibition as a promising tool for the treatment of chronic inflammation and cancer. In fact, it has been widely proven [9,10] that a variety of solid and haematological tumours increases the activation of NF-kB pathway suppressing their susceptibility to drug-induced apoptosis. Moreover, a catalytically inactive mutant of IKK-β was demonstrated to inhibit the transcription by TNFα, IL1-β, LPS and anti-CD3/antiCD28 stimulation [11]. Thus, the inhibition of IKK-β results in a decrease of the transcriptional activity, leading to an enhanced susceptibility to the induction of apoptosis.

The discovery of small molecules acting as selective IKK-β inhibitors (Fig. 1) has been one of the main goals in the research of new anti-inflammatory and anti-cancer treatments. The use of High Throughput Virtual Screening (HTVS) [12] led to the identification of SC-514 (**1**), the first potent and orally active compound. SAR modifications on this compound allowed the design of 2-(aminocarbonyl)amino-5-acetylenyl-3-thiophenecarboxamides (**2**) [13], a novel class of inhibitors with an improved activity and selectivity. More recently [14], the IKK-β inhibitors providing a 2-amino-3,5-diarylbenzamide scaffold (**3**) were discovered starting from structure–activity studies on a series of IKK-ε inhibitors, part of an PMA-inducible IkB kinase complex with a structural homology to IKK-α of 30% [15].

Further SAR studies allowed the identification of BMS-345541 (**4**) as a highly selective and potent inhibitor; starting from this lead compound, amino derivatives of benzimidazoquinoxaline, benzoimidazoquinoline and benzopyrazoloquinazoline were syn-

* Corresponding author. Tel.: +39 0916161606; fax: +39 0916169999.
E-mail address: lauria@unipa.it (A. Lauria).

1093-3263/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2010.04.008

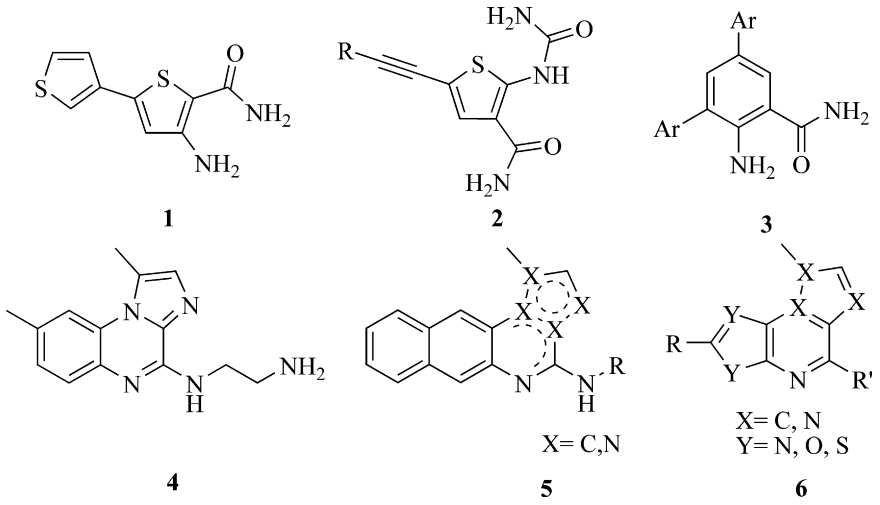

Fig. 1. IKK-β inhibitors.

thesized (**5**) [16], showing an overall improved activity when compared to the lead compound. However, due to toxicity and solubility problems, novel tricycle inhibitors of IKK-β have been designed and synthesized (**6**) [17], with the aim of improving the activity and selectivity against IKK-β and of minimizing side-effects.In order to investigate the binding modes of each class of IKK-β inhibitors, we performed Molecular Modeling studies on the 141 inhibitors as collected in the Binding Database (http://www.bindingdb.org) [18]. It consists of a public accessible database in which several activity data of small heterocycle molecules, interacting with many biological targets, are continuously updated. By using the frozen conformations obtained from Molecular Docking experiments, it was also possible to develop several Pharmacophore Hypotheses for each structural class of compounds. These models can be helpful to predict the activity pattern and to guide the design and synthesis of more effective IKK-β inhibitors.

In recent works quite similar approaches were used involving CoMFA methodology [19], pharmacofore modeling [20], homology modeling [21,22] and Molecular Docking on investigated structures [21–23]. Herein, as improvement in the molecular modeling approach, we proposed the use of a flexible docking protocol to refine the protein structure obtained by homology modeling. Using this method, four different protein–ligand complexes were obtained, one for each structural class of inhibitors investigated; further, these models (validated through the examination of Ramachandran plots) were used to rigidly dock the inhibitors belonging to each class. Finally, the 3D-QSAR models were obtained starting from the frozen conformations obtained in previous stages.

2. Materials and methods

In all molecular modeling facilities below reported (software or web resources), unless otherwise specified, the default parameters were used.

2.1. Homology modeling of IKK-β

Due to the unavailability of a crystal structure of IKK-β, it was necessary to develop a homology model of the protein in order to carry out further experiments. The sequence of Human NF-kB inhibitor kinase beta was retrieved from Swiss-Prot Database (http://www.expasy.ch/sprot/) [24] (Accession Code: O14920) and submitted to the MODWEB server (http://modbase.compbio.ucsf.edu/ModWeb20-html/modweb.html) [25]. This is a comparative modeling web server which accepts one or more sequences in FASTA format and calculates the models using an automatic software pipeline for comparative modeling based on various MODELLER modules [26,27]. Sequence–structure matches are established using a variety of fold-assignment methods, including sequence–sequence, profile–sequence, and profile–profile alignments; a representative model for each alignment is then chosen through a ranking process based on the atomic distance-dependent statistical potential DOPE [28]. Finally, the fold of each model is evaluated using a multiple scoring algorithm that includes parameters such as the length of the modelled sequence, the sequence identity implied by the sequence–structure alignment, the fraction of gaps in the alignment, the compactness of the model and various statistical potential *Z*-scores [28–30]. Since only models that are assessed to have the correct fold were included in the final model sets, this approach enables a through exploration of fold assignments, sequence–structure alignments and conformations, in order to find the model with the best evaluation score. The chosen model was optimized by adding hydrogen atoms and charges, and by minimizing the structure using Schrödinger Prime Refinement Tool [31].

2.2. Induced fit docking

The structure of IKK-β obtained by homology modeling was still not suitable to carry out Molecular Docking experiments since its binding pocket was in a “closed” rigid conformation. In fact, the problem of efficiently handling the receptor flexibility is still considered one of the hardest challenges to overcome; even though the target macromolecule can be considered as a rigid body, it has been shown [32] that large localized rearrangements of the protein surface frequently occur, especially in the ligand binding domains and in large flexible amino acid chains.

Instead of forcing ligands into a closed active site using constraints or manual placing [12], we used a mixed Molecular Docking/Dynamics protocol, called Induced Fit Docking (IFD) [33], developed by Schrödinger Inc. (http://www.schrodinger.com/). This approach combines, in an iterative fashion, the ligand docking techniques with those for modeling receptor conformational changes. The Glide docking program [34] is used for ligand flexibility, while the refinement module in Prime program [31] is used to account for receptor flexibility: the side chain degrees of freedom are mainly sampled, while minor backbone movements are allowed through minimization. The strategy is to dock first ligands

into a rigid receptor using a softened energy function such that steric clashes do not prevent at least one pose from assuming a conformation close to the correct one (ligand sampling step). Further, the receptor degrees of freedom are sampled, and a global ligand/receptor energy minimization is performed for many ligand poses which attempts to identify low free-energy conformations of the whole complex (protein sampling step). A second round of ligand docking is then performed on the refined protein structures, using a hard potential function to sample ligand conformational space within the refined protein environment (ligand resampling step). Finally a composite score function is applied to rank the complexes, accounting for the receptor/ligand interaction energy as well as strain and solvation energies (scoring step). The composite score, used for final ranking of compounds, is given by GlideScore + 0.05 PrimeEnergy. The validity of the whole process was previously tested, and was able to successfully reproduce ligand conformations observed by X-ray crystallography [35].

The goal of this process was to obtain a ligand/receptor model for each class of IKK-β inhibitors that could be used as a starting point to dock the whole class. Among the 141 IKK-β inhibitors included in the Binding Database, we selected four derivatives, one for each structural class, which showed the highest activity data expressed as IC_{50}. These models were also useful to topologically evaluate the active site in terms of binding pockets and interactions within the ligand.

2.3. Molecular Docking

The ligand/receptor models obtained with the IFD protocol were used as a starting point to dock the remaining derivatives belonging to each structural class using the Glide software [34]. This program uses a hierarchical series of filters to search for possible locations of the ligand in the active-site region of the receptor, while the shape and properties of the receptor are represented on a grid by several sets of fields that progressively provide more accurate scoring of the ligand poses. An extensive conformational search is performed using a heuristic screen that rapidly eliminates unsuitable conformations (e.g. long-range internal H-bonds or sterical clashes). The entire amount of poses generated is then hierarchically classified, refined and further minimized into the active site grid before being finally scored using the proprietary *GlideScore* function, defined as:

$$GScore = 0.065 \times vdW + 0.130 \times Coul + Lipo + Hbond + Metal + BuryP + RotB + Site$$

where *vdW* is the van der Waals energy term, *Coul* is the Coulomb energy, *Lipo* is a Lipophilic contact term which rewards favourable hydrophobic interactions, *Hbond* is a H-bonding term, *Metal* is a metal-binding term (where applicable), *BuryP* is a penalty term applied to buried polar groups, *RotB* is a penalty for freezing rotatable bonds and *Site* is a term used to describe favourable polar interactions in the active site.

Our experiments were carried out using the default parameters and the following protocol: initially, the Standard Precision (SP) level of accuracy was used for the generation and scoring of 10 poses for each ligand. The top-scored conformation was further re-docked using the Extra-Precision (XP) algorithm. Finally the best poses were refined using the "Refine" option (this performs an optimization of the ligand structure in the field of the receptor) at the XP level of accuracy. The Extra-Precision mode combines a powerful sampling protocol with the use of a custom scoring function designed to identify ligand poses that would be expected to have unfavourable energies. The presumption is that only active compounds will have available poses that avoid these penalties and receive favourable scores for appropriate hydrophobic contact between the protein and the ligand. The XP sampling method is based on an anchor and refined growth strategy: anchor fragments of the docked ligand (typically rings) are chosen from the set of SP poses and the molecule is re-grown bond by bond from these anchor positions; a complete minimization and XP scoring cycle is carried out on the large ensemble of poses generated by growing method. The XP scoring functions integrates the basic GlideScore functions with additional terms: to model solvation an explicit water term is used, parameterized on several known protein–ligand complexes; similarly, additional parameterized rewards are calculated for buried hydrophobic pockets occupied by hydrophobic ligand groups, including the detection of π-cation and π–π stacking interactions.

2.4. Phamacophore modeling and 3D-QSAR

The top-scored conformations of the ligands, obtained with the previously described docking protocol, were used for the generation of Pharmacophore Hypotheses for each class of derivatives and further for the development of 3D-QSAR models. The software Phase [36] was employed to carry out the calculations. The conformations were first imported in the program together with their activity data as available in the Binding Database, expressed in terms of pIC_{50}. The second step was to use a set of available pharmacophore features (H-bond donor, H-bond acceptor, Hydrophobic group, Aromatic Ring, Positively and Negatively charged group) to create several sets of pharmacophore sites (site points) for all the ligands. Then, common pharmacophores were identified using a tree-based partitioning technique that groups together similar pharmacophores according to the distances between pairs of sites in the pharmacophore (termed *intersite distances*). Each *k*-point pharmacophore is represented by a vector of *n* distances, where $n = k(k-1)/2$. Each intersite distance *d* is filtered through a binary decision tree. The algorithm projects a map of each pharmacophore, according to its intersite distances, into a box of finite size. All pharmacophores that are mapped into the same box are considered to be similar enough to facilitate identification of a common pharmacophore. Thus if each of the minimum required number of ligands contributes at least one pharmacophore belonging to a particular box, then that box represents a common pharmacophore. Such boxes are said to "survive" the partitioning procedure, while all others are eliminated. By means this automatic iterative procedure it is possible to fully explore the entire number of pharmacophore hypotheses available for each ligand set and choose the real representative ones, discarding redundant or unreliable models.

The use of Phase software also allows to build 3D-QSAR models for a set of ligands that are aligned to a selection of hypotheses; the resulting models are 3D-QSAR models, in which chemical features of ligand structures are mapped to a cubic 3D grid. The ligands are first aligned to the set of pharmacophore features in the selected hypotheses using a standard least-squares procedure; a rectangular grid is defined to encompass the space occupied by the aligned ligands. This grid divides the occupied space into *N* uniformly sized cubes, typically 1 Å on each side. The independent variables in the regression are given by the binary-valued occupancies ("bits") of the cubes (by structural components), while the dependent variables are the activities. Because the number of bits is typically much larger than the number of training set molecules, the system is said to be highly underdetermined; thus, the regression is performed by a partial least squares (PLS) method [37], in which a series of models is constructed with an increasing number of PLS factors. The accuracy of the models increases with increasing number of PLS factors until over-fitting starts to occur.

In examining each pharmacophore model several parameters were monitored: *Survival*, which represents a weighted combination of the vector, site, volume scores; *Post-Hoc*: the result of rescoring after minimization; *Site*: a parameter measuring how

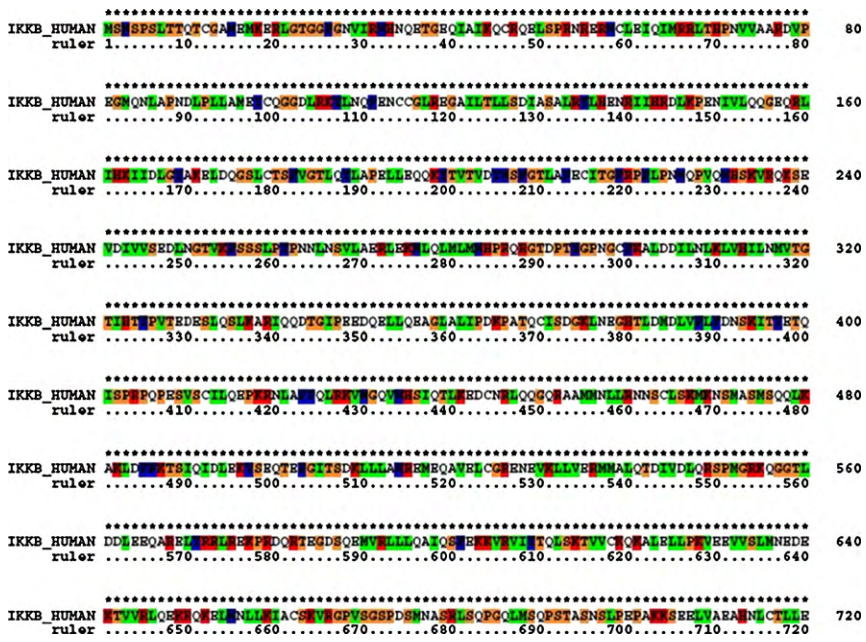

Fig. 2. Human IKK-β sequence (Swiss-Prot Database Accession Code: O14920).

closely the site points are superimposed in an alignment to the pharmacophore of the structures that contribute to this hypothesis; *Vector*: a parameter which measures how well the vectors for acceptors, donors, and aromatic rings are aligned in the structures that contribute to this hypothesis; *Volume*: how much the volumes of the contributing structures overlap when aligned on the pharmacophore. Further, each 3D-QSAR model was characterized by the following parameters: *Factors*: number of PSL factors of the model; *SD*: standard deviation; r^2: correlation coefficient; q^2: cross-validated r^2; *RMSE*: root-mean square error; *F*: Fisher test.

3. Results and discussion

One of the aim of this work was the use of Molecular Modeling techniques to investigate the binding mode of some IKK-β inhibitors, and to provide topological studies on the active site of this enzyme. Unfortunately, no crystallographic structure of the IkB kinase is available in the databases; thus it was necessary to use Homology Modeling Techniques to build a useful model for further investigations.

The sequence of Human IKK-β was retrieved from SwissProt Sequence Database [34] (Fig. 2) and submitted to the MODWEB server [25], a comparative modeling web server based on MODELLER pipelines.

The calculation produced 68 models of IKK-β relative to various portions of the submitted sequences. According to previous studies [12,14,38,39] those models representing portions of the enzyme far from the active site (the ATP binding pocket comprised between residues 44, 90, and 145) were discarded. The resulting models were further discriminated by examining the percentage of identity with the template sequence and the provided ModelScore value. In fact, this parameter, based on potential energy parameters [30], gives evidence about the reliability of the model (at least 95% confidence for values higher than 0.7). This evaluation process allowed

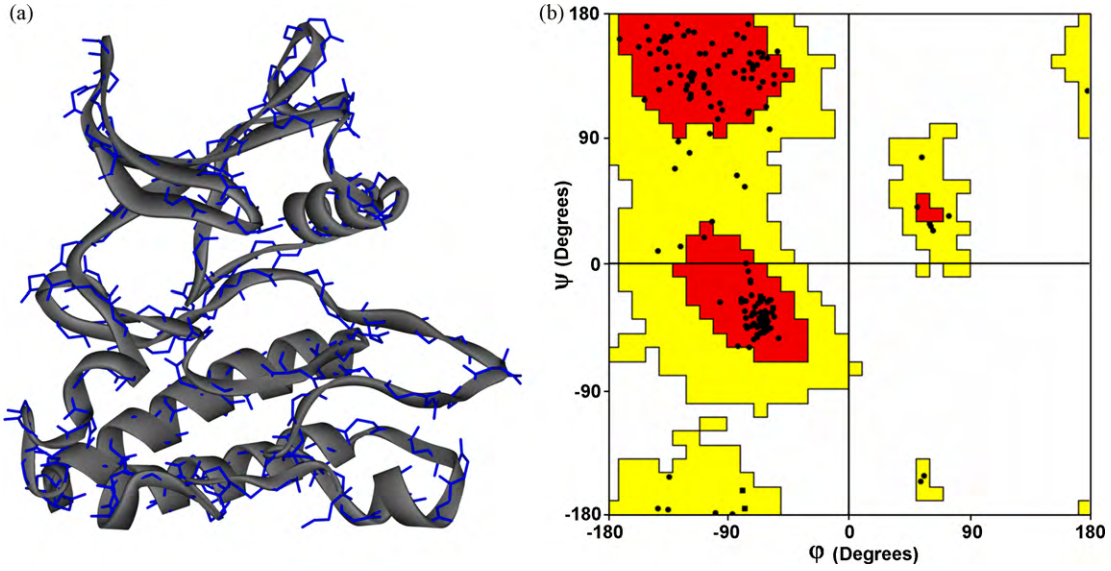

Fig. 3. (a) The homology model of IKK-β from PAK-1 Kinase 3FXZ; (b) Ramachandran plot for the model.

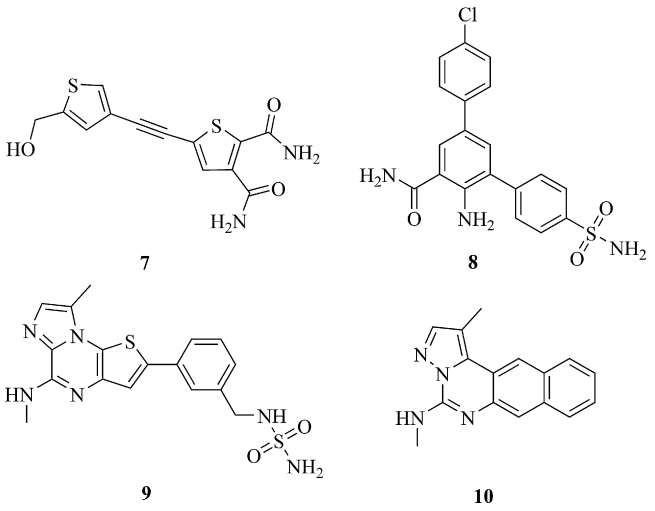

Fig. 4. IKK-β inhibitors used for the generation of ligand/receptor models through IFT.

us to choose, as a reliable model for further analyses, the structure built starting from PAK-1 kinase crystal structure, which showed a good degree of identity in its sequence (36%) (PDB code: 3FXZ [40]) (Fig. 3a).

The resulting protein is characterized by the highest value of homology score (ModelScore = 1) that ensures the good quality of the model. Moreover it reasonably reproduces the catalytic domain of IKK-β, since a total amount of 212 residues are represented (those residues between Glu19 and Gln230). An additional optimization step was performed using the Refinement module of Prime software [31]; moreover, the examination of Ramachandran Plot (Fig. 3b) provided further evidence of the model strength, since all residues fall in allowed regions of the plot.

However, this model was not suitable to carry out Molecular Docking experiments, since the active site is in a closed conformation, which does not allow the insertion of a ligand structure. In order to overcome this obstacle, we used a mixed Molecular Docking/Dynamics approach (IFD). The structures of 141 IKK-β inhibitors were retrieved from the Binding Database [18], a public web-accessible database of binding affinities. These inhibitors can be grouped according to their core structure in four different classes: (a) thiophenecarboxamides (**2**); (b) diarylbenzamides (**3**); (c) tetracycle derivatives (**5**); (d) tricycle derivatives (**6**) (Fig. 1). For each of these classes the most active derivatives **7–10** were selected (Fig. 4) and docked into the IKK-β ATP binding site through IFD in order to validate and refine the macromolecule structure obtained through homology modelling.

By means this protocol it was possible to obtain four Ligand/Receptor models, one for each structural class of compounds (Fig. 5). Moreover, the application of IFD allowed to obtain for each investigated derivative Docking Score values (Table 1) which are in agreement with its activity profile. This trend can be appreciated by examining the plot of the Mean-Weighted Docking Score (WDS, defined as the IFD Score for each compound normalized through dividing it by the mean of scores) values against pIC_{50} (Fig. 6) which confirms that these values are highly correlated with a $r^2 = 0.898$.

The binding models for each structural class were further used as a starting point for the docking of all the derivatives belonging

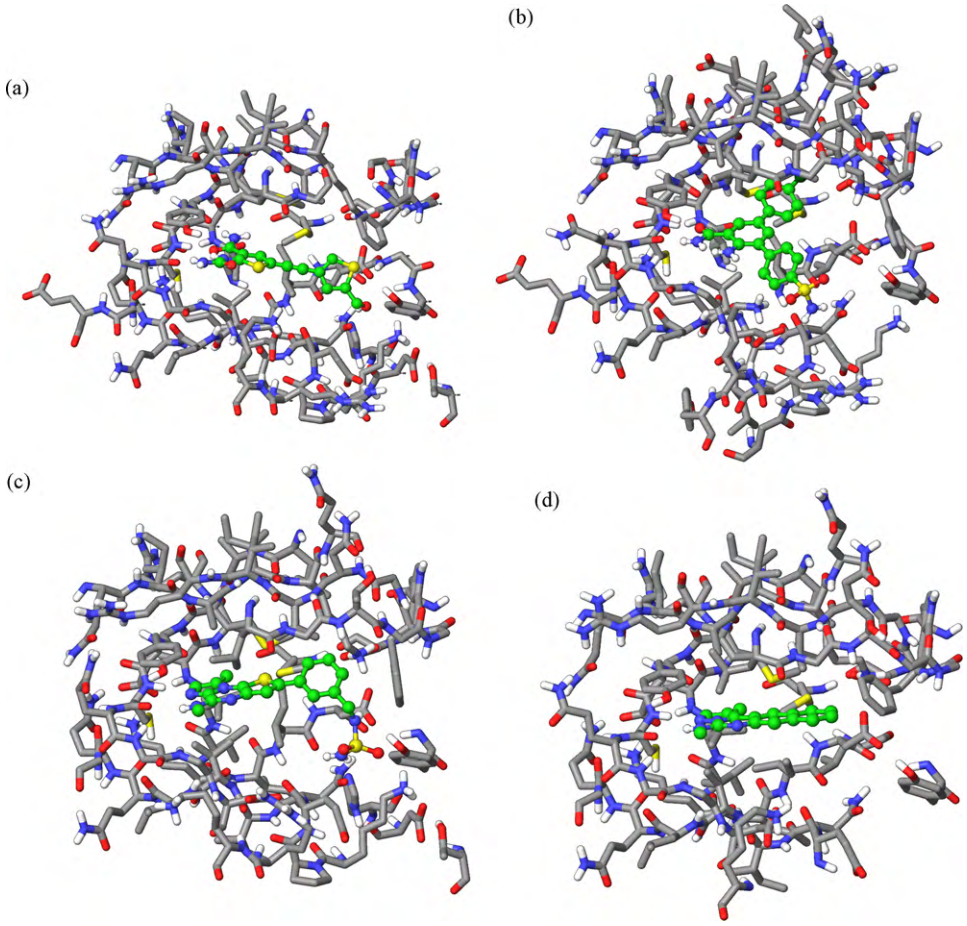

Fig. 5. Models of the ligand/receptor complexes obtained by IFD (see Fig. 4).

Table 1
Activity values and Docking Scores for the derivatives investigated by IFD.

Entry	Class	Binding DB ID	IC_{50} (μM)	pIC_{50}	Docking Score	WDS
7	Thiophenes	26078	0.195	0.71	−6.837	0.88
8	Diarylbenzamides	27467	0.100	1.00	−7.130	0.91
9	Tricycles	25960	0.003	2.52	−9.442	1.21
10	Tetracycles	25924	0.011	1.96	−7.803	1.00

to the corresponding family. With the aim of obtaining a higher degree of reliability, we used a docking protocol involving two different degrees of precision, as available in Glide software. First, the structures were docked by using the SP algorithm; further, the top-scored poses were re-docked and refined with the XP algorithm.

From these experiments it was possible to analyze in deep both the interactions typical of each class of compound and the topological features of the IKK-β binding pocket (Fig. 7).

This pocket is constituted by a central hydrophobic pavement delimited by three different hydrophobic pockets. The first and the most extended one includes residues Ala95, Met96, Val79, Glu61, and Lys44, involved in accommodating large aromatic and heteroaromatic rings, variously substituted. This interaction can be completed by the possibility of additional hydrogen bonds between ring substituents and the side chain of Lys44 and Glu61, respectively. The second hydrophobic pocket is comprised between residues Phe26, Val29, Ile165, and Ile168, and usually interacts with alkyl side chains or simple phenyl rings; without providing any additional H-bond. Finally the third pocket is the most exposed to the solvent side of the protein. It extends between residues Leu21, Arg105, Lys106, and Val152. This pocket is also able to provide additional hydrogen interactions with amino terminal of Lys106 or guanidine moiety of Arg105. According to the model proposed by Traxler and Furet [41] the presence of a loop (corresponding to the Traxler's "Adenine Binding Region") constituted by Glu97, Tyr98, Cys99, Glu100, and Gly101, is remarkable, since it allows the formation of hydrogen bonds in several positions. These interactions stabilize the inhibitor pose and contribute to disrupt the catalytic activity of the enzyme.

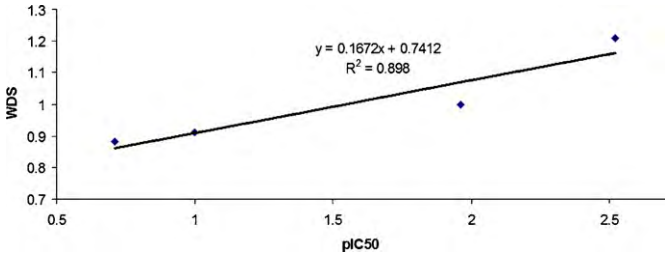

Fig. 6. Plot of the Mean-Weighted Docking Score (WDS) versus pIC_{50} for the derivatives investigated by IFD.

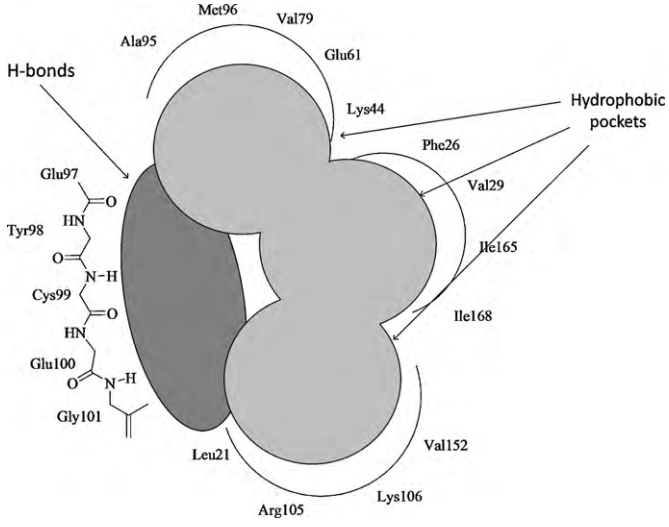

Fig. 7. IKK-β binding pocket and interactions involved in ligand binding as emerged by Molecular Docking experiments.

The analysis of the binding modes adopted by each class of ligands further supports these observations. Thiophenecarboxamide derivatives [13] (Fig. 8) fit very well in the binding pocket; the thiophene moiety represents the central scaffold interacting with the active site pavement.

From the central ring, 5-acetilenyl substituents projects into the whole binding pocket and terminates with an aromatic ring which further stabilizes the interaction. Strong hydrogen bonds are present between the amide substituents on the positions 2 and 3 of the thiophene ring and backbone atoms of Gly97 and Cys99; additional bonds occur with side chains of Asn150 and Arg185.

The diarylbenzamide derivatives [14] (Fig. 9) provided some remarkable differences. In particular the most deeply buried hydrophobic pocket (Ala95, Met96, Val79, Glu61) is occupied by aryl or haloaryl substituents on the position 5 of the benzamide ring.

On the contrary the substituents on the position 3 insert into the solvent-exposed cavity and form additional hydrogen bonds with Lys106 and Asp103. As previously observed, it is relevant the contribution of the H-bonds with the amidic and amino group on the central phenyl ring interacting with Cys99 and Gly102 backbone atoms respectively.

Although the class constituted by tetracycle derivatives [16] (Fig. 10) contains the smallest number of structures (only **7**), our docking results allowed to predict a binding mode similar to that observed for the thiophenecarboxamide class. In particular the central scaffold (benzoquinoline, benzoquinoxaline or benzoquinazoline ring) points into the hydrophobic pocket formed by residues Phe26, Val29, Ile165, and Ile168. Favourable hydrogen bonds are formed between the ligand imidazole ring and Cys99 backbone

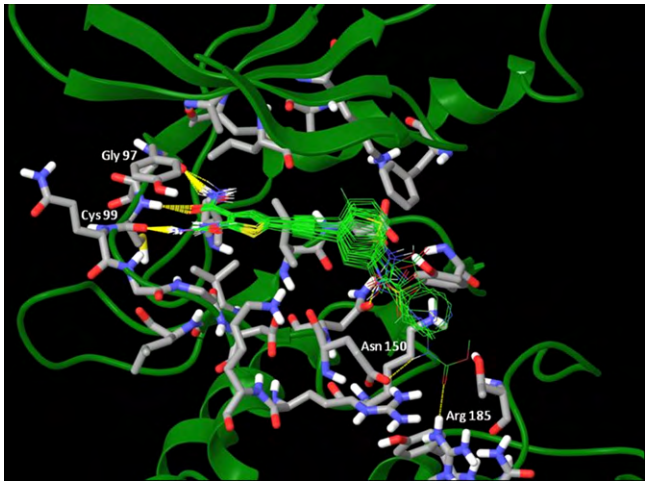

Fig. 8. Binding modes of the thiophenecarboxamide class of compounds.

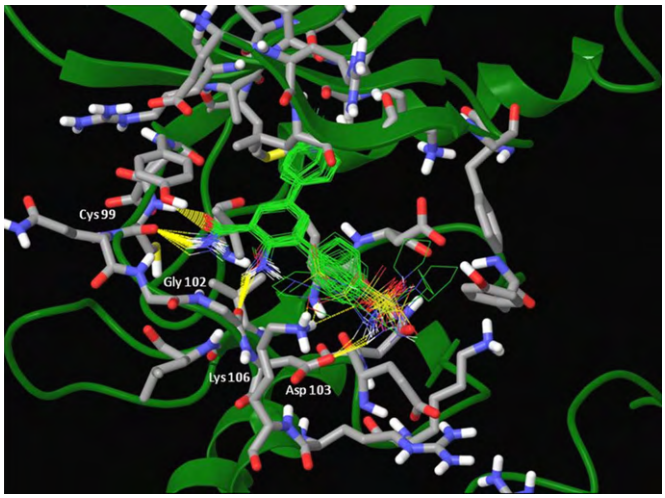

Fig. 9. Binding modes of the diarylbenzamide class of compounds.

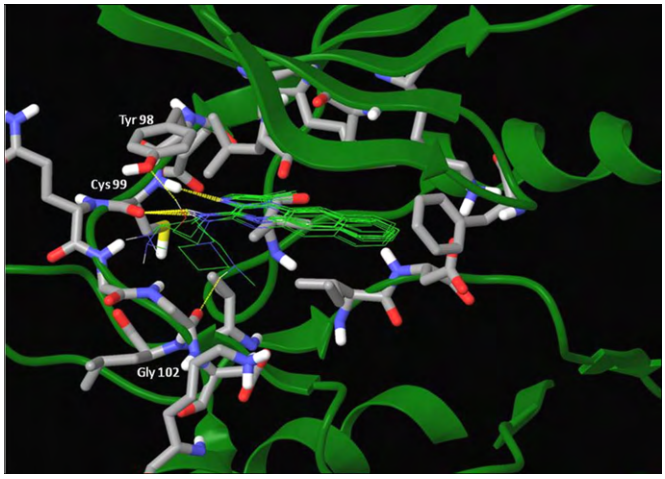

Fig. 10. Binding modes of the tetracycle class of compounds.

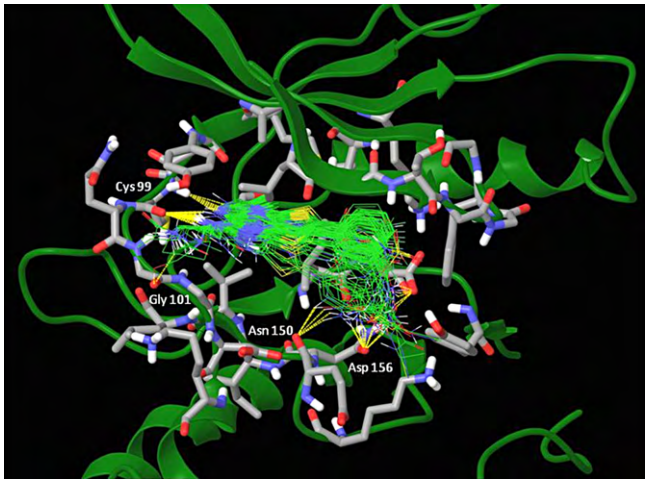

Fig. 11. Binding modes of the tricycle class of compounds.

nitrogen, between the ligand secondary amino group and Cys99 backbone oxygen. Additional H-bonds are observed with Gly102 backbone oxygen and Tyr98 hydroxyl group.

The last examined structural class was constituted by tricycle derivatives [17]. Their binding mode (Fig. 11) involves the possibility of projecting one of the phenyl substituents toward the solvent-exposed hydrophobic cavities, and eventually forming hydrogen bonds with Asn150 amide group and Asp156 backbone oxygen. The methyl substitutents on the tricycle scaffold directly point to the most buried hydrophobic cavity (formed by residues Ala95, Met96, Val79, Glu61, and Lys44). According to the trend already observed for tetracycle derivatives (Fig. 10) strong hydrogen bonds are formed with Cys99 backbone. Moreover, the substituents containing an amino moiety are able to form additional bonds with Gly101 backbone oxygen.

These interesting results suggest that all investigated inhibitors share a common pattern of interactions when approaching the IKK-β active site. The broadest differences can be found in the residues involved in the formation of hydrogen bonds with each class of ligands and in the hydrophobic cavity occupancies. In particular whereas all classes of inhibitors provide hydrogen bonds between Cys99 backbone oxygen and nitrogen, the diarylbenzamide class provides additional strong bonds with Gly102 due to the 2-amino group present in all the investigated structures, while the thiophenecarboxamide class forms an additional bond with Gly97 backbone oxygen.

Starting from the frozen ligand conformations obtained by Molecular Docking using the Extra-Precision algorithm of Glide [34], pharmacophore and 3D-QSAR models were built in order to obtain a more quantitative analysis of the interactions between IKK-β and its inhibitors. The software Phase was used to perform our experiments. According to the protocol previously followed for Molecular Docking experiments, the inhibitors were grouped following their structural class. In this case, however, the tricycle and tetracycle class of compounds were merged together. This choice can be easily explained considering that: (a) the two classes share the same binding mode and strength of interactions with key residues such as Cys99; (b) these inhibitors are derived by a common lead compound [16,17] (BMS-345541, Fig. 1 (**4**)); (c) the number of tetracycle derivatives is too small to carry out and test a statistically valid Pharmacophore/3D-QSAR model (a model based on too few structures easily leads to over-fitting problems). In addition, each class of compounds was divided in a training set and a test set by an automatic procedure which randomly chooses the 20% of whole structures to constitute the test set. After this first evaluation steps, the following classes of compounds were used to build Pharmacophore/3D-QSAR models: (a) thiophenecarboxamides (22 derivatives with a pIC50 in the range 0.71 and −1.301 divided into a 17 structures training set and 5 structures test set); (b) diarylbenzamides (33 derivatives with a pIC50 in the range 1.0 and −1.20 divided into a 27 structures training set and 6 structures test set); (c) tri/tetracycles compounds (86 derivatives with a pIC50 in the range 2.52 and −1.398 divided into a 70 structures training set and 16 structures test set).

For the thiophenecarboxamide class of compounds a four-points pharmacophore was obtained (Fig. 12a and Table 2). This hypothesis is based on two centres containing aromatic rings, one of which corresponds to the thiophene ring, placed at 6.80 Å distance each other. A H-bond donor centre and a H-bond acceptor centre (placed at 3.56 and 4.57 Å from the thiophene ring and at 2.50 Å each other) are also present and correspond to the amide moiety on the thiophene ring. The high value of Post-Hoc score obtained for this hypothesis (3.61) indicates that the model is highly reliable. Moreover, the Site and Vector Score values (0.90 and 0.99 respectively) confirm that the site points are well superimposed to the ligands and the H-bond centres are well aligned to the corresponding moieties.

The 3D-QSAR model obtained for these derivatives (Fig. 12b and Table 2) is highly reliable, with a $r^2 = 0.98$ and $q^2 = 0.71$. The analysis of the 3D areas related to the distribution of hydrophobic properties suggests that of hydrophobic substituents on the aromatic rings

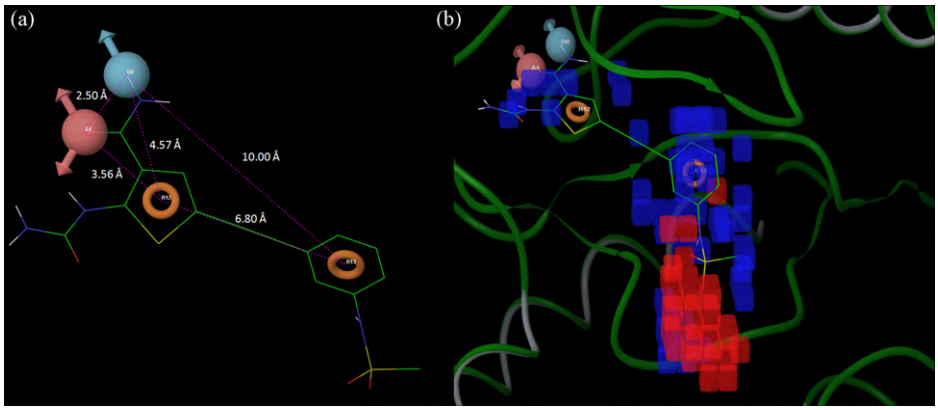

Fig. 12. (a) Pharmacophore model from the thiophenecarboxamide class of compounds; (b) 3D-QSAR areas relative to the hydrophobic properties of the ligands. Blue: favourable interactions; Red: unfavourable interactions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

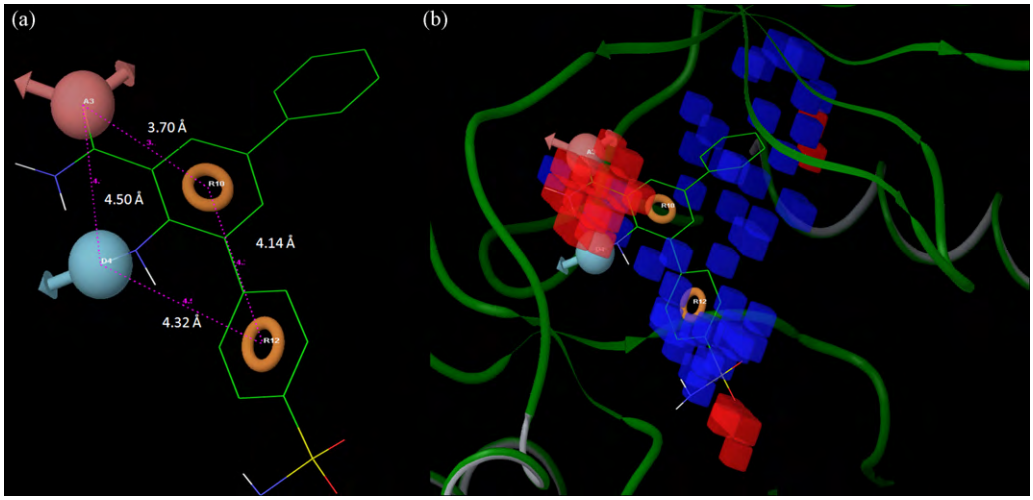

Fig. 13. Pharmacophore hypothesis and 3D-QSAR model for diarylbenzamide class of compounds (see Fig. 12 for further explanations).

are well tolerated, as these groups fit well into the protein cavities. However, further substitutions to the phenyl ring are not suitable to improve the interaction with the enzyme, because these groups directly interact with Asn150 and Arg185 (see Molecular Docking experiments).

The diarylbenzamide class of compounds also provided a four-points pharmacophore (Fig. 13a and Table 2). Two aromatic rings are present, located at 4.14 Å distance each other; additional features are a H-bond donor centre (placed at 4.32 Å from the external ring) and a H-bond acceptor centre located at 4.50 Å from the H-bond donor and at 3.70 Å from the central ring. A comparison with the pharmacophore obtained for the thiophenecarboxamide derivatives reveals that there is high correspondence between the two models. In particular, it is possible a superimposition between the two aromatic ring centres and between the two H-bond acceptor areas, whereas the positions of the external rings and the H-bond donor areas reflect the different pattern of interactions with the active site and the different occupancies of the binding pockets.

Moreover, the analysis of the 3D-QSAR model ($r^2 = 0.96$ and $q^2 = 0.82$) gives a further evidence of this behaviour. According to this model (Fig. 13b) the favoured hydrophobic interactions are those who fill the pockets. Particularly the cavities between Ala95, Met96, Val79, Glu61 and between Leu21, Arg105, Lys106, and Val152 (see Figs. 7 and 10); the hydrophobic substitution on the 5-phenyl ring appear to be highly favoured, since these groups inserts in the most buried cavity, while the increase of hydrophobicity on the 3-phenyl ring lead to the disruption of a hydrogen bond network involving the solvent-exposed cavity of the enzyme.

The last pharmacophore model was obtained for the class constituted by tri- and tetra-cycle derivatives (Fig. 14a and Table 2). As in the case of previous models, this pharmacophore hypothesis contains four centres, with two aromatic rings placed at very small distance (only 2.19 Å), due to the presence of a very compact

Table 2
Pharmacophore and 3D-QSAR models parameters (see Appendix BMaterials and Methods and Supplementary Material for further information).

Class	Survival	Post-Hoc	Site	Vector	Volume	Factor	SD	r^2	q^2	RMSE	F
Thiophenes	6.41	3.61	0.90	0.99	0.72	3	0.07	0.98	0.71	0.17	298.1
Diarylbenzamides	7.90	3.34	0.86	0.87	0.62	4	0.10	0.96	0.82	0.32	386.4
Tri-tetracycles	1864.83	3.648	0.96	0.99	0.70	3	0.36	0.69	0.58	0.39	48.9

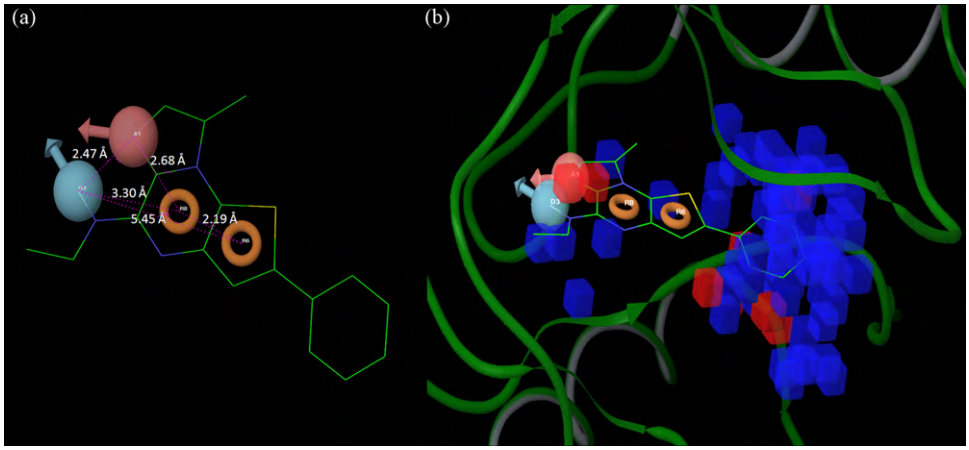

Fig. 14. Pharmacophore hypothesis and 3D-QSAR model for the tri- and tetracycle class of compounds (see Fig. 12 for further explanations).

ring system. The remaining centres follow the pattern of hydrogen bonds detected in previous models, with H-bond acceptor and donor areas placed at 2.47 Å each other. Nevertheless these centres are differently oriented and reflect the specific pattern of interaction of this class of compounds with the H-bond loop (see Fig. 7).

The obtained 3D-QSAR model has a lower degree of predictive capability if compared to previous results ($r^2 = 0.69$ and $q^2 = 0.58$). This behaviour can be partially explained by considering the difficulties to obtain a model valid for such a complex pattern of substitutions and different polycyclic ring systems. However, the analysis of 3D-QSAR areas according to the distribution of hydrophobicity features leads to the same conclusions as those drawn for the previous models. As seen in the thiophenecarboxamide model, hydrophobic moieties on the phenyl group seem to be highly favoured, as this group is kept buried into the cavity formed by residues Phe26, Val29, Ile165, and Ile168, as well as the elongation of the aliphatic chain on the secondary amine, which points on the external binding pocket. By contrast, the increase of the steric hindrance in the H-bond loop region lead to a decrease in overall activity.

A comparative analysis of the results obtained for pharmacophore/3D-QSAR models confirms the central role played by the presence of hydrogen bonds between the inhibitor and the loop constituted by residues Glu97, Tyr98, Cys99, Glu100, Gly101, and Gly102. In particular, all the obtained pharmacophores contain a H-bond acceptor area in correspondence of Cys99 bound to an aromatic ring centre placed at distance between 2.68 and 3.70 Å. Additional interactions are given by the presence of a H-bond donor area placed near Gly97 or Gly102. The selectivity for one of these residues depends on the degree of flexibility and opening of the structural scaffold. Moreover, the presence of a central skeleton constituted by an aryl or heteroaryl ring is a common feature for all the investigated ligands which emphasizes the need of simulating the purine ring of ATP, the natural substrate of IKK-β.

Due to the high hydrophobic character of the binding pocket, it is no surprising that hydrophobic substitutions tend to enhance the fitting of each ligand in the active site. This can be explained considering that the increase of the surface complementarity forces the inhibitor to maintain a conformation which points the moieties involved in the formation of H-bonds toward the catalytic residues of IKK-β. The presence of additional H-bonds with the solvent-exposed pocket of the receptor (Asp103, Lys106 and Asn150) further stabilizes the interaction.

4. Conclusions

IKK-β plays a crucial role in the control of NK-kB activity, and the design of new inhibitors providing a higher activity and selectivity represent one of the most ambitious challenges of these years. Although several classes of inhibitors have been discovered and tested for their ability to inhibit the enzyme, no crystallographic structure of a IKK-β/inhibitor complex is available to date.

In this work a deep analysis of the interactions between the four classes of inhibitors included in the Binding Database and IKK-β was carried out, by using a computational approach involving Homology Modelling, Molecular Docking, and Pharmacophore/3D-QSAR analysis.

In order to overcome the obstacle given by the unavailability of crystallographic data, a model of the drug–receptor complex has been built by Homology Modelling and validated. Furthermore, the 141 inhibitor structures, as collected in the Binding Database, were docked in the active site by means a Docking algorithm involving two level of accuracy. The frozen conformations obtained were used in combination with the activity data to carry out and test Pharmacophore/3D-QSAR models, the analysis of which allowed to get insight deeper on the study the nature of the interaction with their target.

In consideration of the obtained results, the molecular mechanism involved in the inhibition of IKK-β is similar to that observed in the case of other kinase proteins. It consist in the formation of a ligand–receptor complex able to mimic ATP, the natural substrate of such class of proteins. Moreover, these studies allowed to characterize topologically the binding site and revealed the presence of three broad hydrophobic pockets able to accept a considerable steric hindrance. On the opposite side of the cavity a loop constituted by residues Glu97, Tyr98, Cys99, Glu100, Gly101, and Gly102 represents a region involved in the formation of several hydrogen bonds with each ligand.

The Molecular Docking approach allowed to verify that all investigated compounds provided a similar binding mode, with differences in the hydrophobic cavity occupied by the side chain group. In particular, as suggested by the Pharmacophore models, the presence of a central aromatic ring or ring system is a common feature shared by all classes of inhibitors; starting from this central scaffold the substituent groups project towards the different pockets and interact with different residues in the H-bond loop, according to their degree of flexibility. In agreement to the obtained 3D-QSARs, the increase of steric hindrance of the side chains and

the introduction of moieties able to give or accept H-bonds is a convenient strategy to improve the strength of the interaction and, hopefully, the activity of newly lead compounds.

Acknowledgment

This work was supported in part by CNR fund, RSTL project PM ID 624.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmgm.2010.04.008.

References

[1] P.A. Bauerle, T. Henkel, Function and activation of NF-kβ in the immune system, Annu. Rev. Immunol. 12 (1994) 141–179.
[2] G. Ghosh, M.J. May, E.B. Kopp, NF-kβ and Rel proteins: evolutionarily conserved mediators of immune responses, Annu. Rev. Immunol. 16 (1998) 225–260.
[3] M. Karin, Y. Yamamoto, Q. May Wang, The IKK NF-B system: a treasure trove for drug development, Nat. Rev. Drug Discov. 3 (2004) 17–26.
[4] A.A. Beg, S.M. Ruben, R.I. Scheinmann, S. Haskill, C.A. Rosen, A. Baldwin, I kappa B interacts with the nuclear localization sequences of the subunits of NF-kappa B: a mechanism for cytoplasmic retention, Genes Dev. 6 (1992) 1899–1913.
[5] M. Karin, Y. Ben-Neriah, Phosphorylation meets ubiquitination: the control of NF-kB activity, Annu. Rev. Immunol. 18 (2000) 621–663.
[6] Y. Hu, V. Baud, T. Oga, K.I. Kim, K. Yoshida, M. Karin, IKKα controls formation of the epidermis independently of NF-kβ, Nature 410 (2001) 710–714.
[7] Z.W. Li, W. Chu, Y. Hu, M. Delhase, T. Deerink, M. Ellisman, R. Johnson, M. Karin, The IKKβ subunit of IB kinase (IKK) is essential for nuclear factor B activation and prevention of apoptosis, J. Exp. Med. 189 (1999) 1839–1845.
[8] C. Makris, V.L. Godfrey, G. Krähn-Senftleben, T. Takahashi, J.L. Roberts, T. Schwarz, L. Feng, R. Johnson, M. Karin, Female mice heterozygous for IKK gamma/NEMO deficiencies develop a dermatopathy similar to the human X-linked disorder incontinentia pigmenti, Mol. Cell 5 (2000) 969–979.
[9] B. Haefner, NF-κB: arresting a major culprit in cancer, Drug Discov. Today 7 (2002) 653–663.
[10] M. Karin, A. Lin, NF-κB at the crossroads of life and death, Nat. Immunol. 3 (2002) 221–227.
[11] M. Delhase, M. Hayakawa, Y. Chen, M. Karin, Positive and negative regulation of IB kinase activity through IKK subunit phosphorylation, Science 284 (1999) 309–313.
[12] A. Baxter, S. Brough, A. Cooper, E. Floettmann, S. Foster, H. Carding, J. Kettle, C. Martin, M. Mobbs, M. Needham, P. Newham, S. Paine, S. St-Gallay, S. Salter, J. Unitt, J. Xue, Hit-to-lead studies: the discovery of potent, orally active, thiophenecarboxamide IKK-2 inhibitors, Bioorg. Med. Chem. Lett. 14 (2004) 2817–2822.
[13] D. Bonafoux, S. Bonar, L. Christine, M. Clare, A. Donnelly, J. Guzova, N. Kishore, P. Lennon, L. Aibby, S. Mathialagan, W. McGhee, S. Rouw, C. Sommers, M. Tollefson, C. Tripp, R. Weier, R. Wolfson, Y. Min, Inhibition of IKK-2 by 2-[(aminocarbonyl)amino]-5-acetylenyl-3-thiophenecarboxamides, Bioorg. Med. Chem. Lett. 15 (2005) 2870–2875.
[14] J.A. Christopher, B.G. Avitabile, P. Bamborough, A.C. Champigny, G.J. Cutler, S.L. Dyos, K.G. Grace, J.K. Kerns, J.D. Kitson, G.W. Mellor, J.V. Morey, M.A. Morse, C.F. O'Malley, C.B. Patel, N. Probst, W. Rumsey, C.A. Smith, M.J. Wilson, The discovery of 2-amino-3,5-diarylbenzamide inhibitors of IKK-α and IKK-β kinases, Bioorg. Med. Chem. Lett. 17 (2007) 3972–3977.
[15] R.T. Peters, S.M. Liao, T. Maniatis, IKKε is part of a novel PMA- inducible IκB kinase complex, Mol. Cell 5 (2000) 513–522.
[16] F. Beaulieu, C. Ouellet, E.H. Ruediger, M. Belema, Y. Qiu, X. Yang, B. Janville, J.R. Burke, K. Gregor, J.F. McMaster, A. Marte, K.W. McIntyre, M.A. Pattoli, F.C. Zusi, D. Vyas, Synthesis and biological evaluation of 4-amino derivatives of benzimidazoquinoxaline, benzimidazoquinoline and benzopyrazoloquinazoline as potent IKK inhibitors, Bioorg. Med. Chem. Lett. 17 (2007) 1233–1237.
[17] J. Kemptson, S.H. Spergel, J. Guo, C. Quesnele, P. Gill, D. Belanger, A.J. Dyckman, T. Li, S.H. Watterson, C.M. Langevine, J. Das, R.V. Moquin, J.A. Furch, A. Marinier, M. Dodier, A. Martel, D. Nirschl, K. Van Kirk, J.R. Burke, M.A. Pattoli, K. Gillooly, K.W. McIntyre, L. Chen, Z. Yang, P.H. Marathe, D. Wang-Iverson, J.H. Dodd, M. McKinnon, B.J.C. arrish, W.J. Pitts, Novel tricyclic inhibitors of IkappaB kinase, J. Med. Chem. 52 (2009) 1994–2005.
[18] T. Liu, Y. Lin, X. Wen, R.N. Jorissen, M.K. Gilson, D.B. Binding, A web-accessible database of experimentally determined protein–ligand binding affinities, Nucleic Acids Res. 35 (2007) D198–D201.
[19] W. Long, P. Liu, Q. Li, Y. Xu, J. Gao, 3D-QSAR studies on a class of IKK-2 inhibitors with GALAHAD used to develop molecular alignment models, QSAR Comb. Sci. 27 (2008) 1113–1119.
[20] J.A. Christopher, P. Bamborough, C. Alder, A. Campbell, G.J. Cutler, K. Down, A.M. Hamadi, A.M. Jolly, J.K. Kerns, F.S. Lucas, G.W. Mellor, D.D. Miller, M.A. Morse, K.D. Pancholi, W. Rumsey, Y.E. Solanke, R. Williamson, Discovery of 6-Aryl-7-alkoxyisoquinoline inhibitors of IκB Kinase-β (IKK-β), J. Med. Chem. 52 (2009) 3098–3102.
[21] S. Nagarajan, M. Doddareddy, H. Choo, Y. Cho, K. Oh, B. Lee, A. Pae, IKKβ inhibitors identification. Part I: homology model assisted structure based virtual screening, Bioorg. Med. Chem. 17 (2009) 2759–2766.
[22] C.M. Avila, N.C. Romeiro, C.M. Sant'Anna, E.J. Barreiro, C.A. Fraga, Structural insights into IKKbeta inhibition by natural products staurosporine and quercetin, Bioorg. Med. Chem. Lett. 19 (2009) 6907–6910.
[23] H. Sugiyama, M. Yoshida, K. Mori, T. Kawamoto, S. Sogabe, T. Takagi, H. Oki, T. Tanaka, H. Kimura, Y. Ikeura, Synthesis and structure–activity relationship studies of benzothieno[3,2-b]furan derivatives as a novel class of IKKβ inhibitors, Chem. Pharm. Bull. 55 (2007) 613–624.
[24] B.E. Suzek, H. Huang, P. McGarvey, R. Mazumder, C.H. Wu, UniRef: comprehensive and non-redundant UniProt reference clusters, Bioinformatics 23 (2007) 1282–1288.
[25] U. Pieper, N. Eswar, B.M. Webb, D. Eramian, L. Kelly, D.T. Barkan, H. Carter, P. Mankoo, R. Karchin, M.A. Marti-Renom, F.P. Davis, A. Sali, MODBASE, a database of annotated comparative protein structure models and associated resources, Nucleic Acids Res. 37 (2009) D347–D354.
[26] M.A. Marti-Renom, A. Stuart, A. Fiser, R. Sanchez, F. Melo, A. Sali, Comparative protein structure modeling of genes and genomes, Annu. Rev. Biophys. Biomol. Struct. 29 (2000) 291–325.
[27] N. Eswar, B. John, N. Mirkovic, A. Fiser, V.A. Ilyn, U. Pieper, A.C. Stuart, M.A. Marti-Renom, M.S. Madhusudhan, B. Yerkovich, A. Sali, Tools for comparative protein structure modeling and analysis, Nucleic Acids Res. 31 (2003) 3375–3380.
[28] M.Y. Shen, A. Sali, Statistical potential for assessment and prediction of protein structures, Protein Sci. 15 (2006) 2507–2524.
[29] D. Eramian, M.Y. Shen, D. Devos, F. Melo, A. Sali, M.A. Marti-Renom, A composite score for predicting errors in protein structure models, Protein Sci. 15 (2006) 1653–1666.
[30] F. Melo, R. Sanchez, A. Sali, Protein structure modeling for structural genomics, Protein Sci. 11 (2002) 430–448.
[31] Prime, version 2.1, Schrödinger, LLC, New York, NY, 2009.
[32] J. Cherfils, J. Janin, Protein docking algorithms: simulating molecular recognition, Curr. Opin. Struct. Biol. 3 (1993) 265–269.
[33] W. Scherman, T. Day, M.P. Jacobson, R.A. Friesner, R. Farid, Novel procedure for modeling ligand/receptor induced fit effects, J. Med. Chem. 49 (2006) 534–553.
[34] Glide, version 5.5, Schrödinger, LLC, New York, NY, 2009.
[35] A. Lauria, M. Ippolito, A.M. Almerico, Inside the Hsp90 inhibitors binding mode through induced fit docking, J. Mol. Graph. Model. 27 (2009) 712–722.
[36] (a) S.L. Dixon, A.M. Smondyrev, E.H. Knoll, S.N. Rao, D.E. Shaw, R.A. Friesner, PHASE: a new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening. 1. Methodology and preliminary results, J. Comput. Aid. Mol. Des. 20 (2006) 647–671;
(b) Phase, version 3.1, Schrödinger, LLC, New York, NY, 2009.
[37] R. Todeschini, Introduction to Chemometrics, EdiSes, Naples, 1998.
[38] W. Long, P. Liu, Q. Li, W. Xu, J. Gao, 3D-QSAR studies on a class of IKK-2 inhibitors with GALAHAD used to develop molecular alignment models, QSAR Comb. Chem. 27 (2008) 1113–1119.
[39] R. Buijsman, Structural aspects of kinases and their inhibitors, in: Chemogenomics and Drug Discovery: A Medicinal Chemistry Perspective, Wiley–VCH, Weinheim, 2004.
[40] J. Maksimoska, L. Feng, K. Harms, C. Yi, J. Kissil, R. Marmorstein, E. Meggers, Targeting large kinase active site with rigid, bulky octahedral ruthenium complexes, J. Am. Chem. Soc. 130 (2008) 15764–15765.
[41] P. Traxler, P. Furet, Strategies toward the design of novel and selective protein tyrosine kinase inhibitors, Pharmacol. Ther. 82 (1999) 195–206.