Rational design of new protein kinases inhibitors of pharmaceutical interest

----

Bortolato, Andrea

UNIVERSITÉ DE GENÈVE

Section des Sciences Pharmaceutiques

UNIVERSITÁ DEGLI STUDI DI PADOVA

Department of Pharmaceutical Sciences

FACULTÉ DES SCIENCES

Professeur Leonardo Scapozza

FACOLTÁ DI FARMACIA

Professeur Stefano Moro

Rational Design of New Protein Kinases Inhibitors of Pharmaceutical Interest

THÈSE
en co-tutelle

présentée à la Faculté des sciences de l'Université de Genève
pour obtenir le grade de Docteur ès sciences, mention sciences pharmaceutiques

par

Andrea BORTOLATO

de

Spinea (Italie)

Thèse N° 3962

PADOVA

Legatoria Artigiana

2008

Doctorat ès sciences
mention sciences pharmaceutiques

Thèse en co-tutelle avec **L'Universitá degli Studi di Padova, Italia**

Thèse de *Monsieur Andrea BORTOLATO*
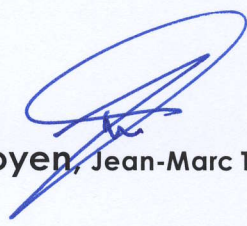
intitulée :

"Rational Design of New Protein Kinases Inhibitors of Pharmaceutical Interest"

La Faculté des sciences, sur le préavis de Messieurs L. SCAPOZZA, professeur ordinaire et codirecteur de thèse (Groupe de biochimie pharmaceutique, Laboratoire de chimie thérapeutique), S. MORO, professeur et codirecteur de thèse (University of Padova, Molecular modeling section, Department of pharmaceutical sciences, Padova, Italy), F. GERVASIO, docteur (Eidgenössische Technische Hochschule Zurich, Department of Chemistry and Applied Biosciences, Lugano, Switzerland) et de Madame M. ROBERTI, professeure (University of Bologna, Department of Pharmaceutical Sciences, Bologna, Italy), autorise l'impression de la présente thèse, sans exprimer d'opinion sur les propositions qui y sont énoncées.

Genève, le 27 mars 2008

Thèse - 3962 -

Le Doyen, Jean-Marc TRISCONE

N.B.- La thèse doit porter la déclaration précédente et remplir les conditions énumérées dans les "Informations relatives aux thèses de doctorat à l'Université de Genève".

Nombre d'exemplaires à livrer par colis séparé à la Faculté : - 7 -

"Some scientists work so hard
there is no time left for serious thinking."
Francis Crick

Dedicated to Francesca

Contents

List of Figures

Abbreviations

PK	Protein kinase
A-loop	Activation loop
P-loop	nucleotide-binding loop
SH1	Src homology domain 1, usually the kinase domain
SH2	Src homology domain 2
SH3	Src homology domain 3
CK2	Casein kinasi 2
CML	Chronic myeloid leukaemia
Ph	Philadelphia chromosome
Bcr	Breakpoint cluster region
Abl	Abelson proto-oncogene protein
Bcr-Abl	Bcr-Abl fusion protein
v-Src	PTK encoded by the chicken Rous sarcoma virus
c-Src	Cellular counterpart of c-Src in mammals
v-Kit	Viral oncogene first identified in the Hardy-Zuckerman IV feline sarcoma virus
Kit	Cellular counterpart of v-Kit, stem-cell factor/stem-cell factor receptor
GIST	Gastrointestinal stromal tumor
PDGFR	Platelet-derived growth factor receptor
PDGF	Platelet-derived growth factor
STI-571	Signal Transduction Inhibitor-571, known also as CGP-57148, imatinib, Gleevec (in the United States) and Glivec (in Europe)
HES	Idiopathic hypereosinophilic syndrome
CEL	Chronic eosinophilic leukemia
LIE	Linear Interaction Energy Method
MD	Molecular Dynamics
MM-PBSA	Molecular Mechanics Poisson Boltzmann Surface Area
MM-GBSA	Molecular Mechanics Generalized Born Surface Area

MM-PB/GBSA	Molecular Mechanics Poisson Boltzmann and Generalized Born Surface Area
MM-PBSA-C	Molecular Mechanics Poisson Boltzmann Surface Area Colony Energy
MM-GBSA-C	Molecular Mechanics Generalized Born Surface Area Colony Energy
MM-PB/GBSA-C	Molecular Mechanics Poisson Boltzmann and Generalized Born Surface Area Colony Energy

Riassunto

Le cellule rispondono a stimoli esterni grazie a diverse cascate del segnale attivate da recettori transmembrana come i recettori accoppiati alle proteine G e recettori tirosin chinasici. Tali vie del segnale non trasmettono semplicemente il segnale ma lo processano, decodificano ed integrano in un complesso sistema finemente regolato che determina un preciso comportamento cellulare. In tale contesto le protein chinasi sono enzimi chiave per la vita di ogni cellula, grazie al controllo che compiono sulle cascate di trasduzione del segnale. Una perturbazione della loro attività influenza pesantemente questo importante equilibrio causando diverse patologie, incluso cancro, diabete ed infiammazione. Tali enzimi esplicano il loro controllo sui substrati fosforilando alcuni specifici residui di Serina, Treonina o Tirosina, utilizzando una molecola di ATP come donatore di gruppi fosfato. In tale studio sono state prese in considerazione principalmente due proteine chinasi, CK2 e Abl, la cui disfunzione è stata dimostrata essere implicata con lo sviluppo e la progressione di diverse forme tumorali. La prima, nota anche come Casein Chinasi 2, è una proteina pleiotropica, costitutivamente attiva, implicata in innumerevoli processi fisiologici e sovra-espressa in varie linee cellulari cancerose. La protein Tirosin chinasi Abelson (Abl) invece è la causa molecolare insieme agli enzimi PDGFR e Kit della Leucemia Mieloide cronica e del tumore Stromale Gastrointestinale.

Le protein chinasi costituiscono un famiglia di bersagli molecolari per una nuova generazione di farmaci antitumorali in grado di distruggere i meccanismi di base necessari al tumore per la sua sopravvivenza e sviluppo. In questo studio si è cercato di raggiungere tale scopo utilizzando un approccio razionale di sviluppo di nuove piccole molecole organiche a potenziale effetto antitumorale grazie alla loro attività inibitoria verso tali chinasi. Le metodiche impiegate sfruttano le più avanzate tecniche computazionali di drug design basate sulla struttura tridimensionale del bersaglio terapeutico e sono state impiegate solamente dopo un'intensa fase di validazione interna basata sui dati sperimentali disponibili in letteratura e derivati da importanti collaborazioni con diversi

gruppi di ricerca a livello nazionale ed internazionale.

La morfologia del target proteico in studio può derivare dalla spettrometria NMR, di diffrazione a raggi X o da modelli computazionali derivati da templati di struttura nota. Se non è disponibile la conformazione del complesso proteina-inibitore può essere valutata attraverso uno studio di docking molecolare che mira alla predizione della conformazione energeticamente più favorevole del ligando nella tasca enzimatica. Nel caso in cui ci siano sufficienti dati sperimentali disponibili, le prestazioni di tale approccio possono essere migliorate sfruttando i risultati di uno studio qualitativo di relazione struttura-attività (SAR) di un set di inibitori noti che può essere utilizzato come base per lo sviluppo di un modello farmacoforico. Il farmacoforo costituisce la collezione qualitativa delle caratteristiche chimico fisiche chiave per l'interazione dell'inibitore con il sito attivo e può essere sfruttato per guidare la ricerca conformazionale durante il docking o per selezionare le migliori conformazioni ottenute. Lo step successivo, la stima dell'energia libera di legame del ligando alla proteina bersaglio, può derivare o da una funzione matematica (scoring function) collegata ai vari contributi energetici in cui può essere scomposta l'interazione, spesso ottimizzata su base empirica o da un consenso di più scoring function. Approcci più dispendiosi computazionalmente, ma più rigorosi da un punto di vista chimico-fisico sono Linear Interaction Energy Method (LIE) e Molecular Mechanics Poisson Boltzmann Surface Area (MM-PBSA). Il primo crea un modello energetico attraverso una correlazione lineare di tre termini energetici (un contributo di van der Waals, una parte elettrostatica e il contributo non polare) ad un set di dati sperimentali, valutando la differenza di energia tra lo stato libero in soluzione dell'inibitore e lo stato legato al bersaglio. Nell'approccio MM-PBSA invece non vi è alcuna componente empirica e l'energia è valutata su una serie di conformazioni derivate da uno studio di dinamica molecolare del complesso.

Utilizzando tali metodologie computazionali è stata effettuato un completo studio della tasca recettoriale di CK2, creando diverse ipotesi farmacoforiche basate su dati strutturali cristallografici e affiancandole a studi che miravano alla comprensione dell'elettrostatica del sito legante l'ATP. Tali approcci oltre a valutare qualitativamente le interazioni cruciali inibitore-proteina hanno evidenziato l'importanza di alcune molecole d'acqua per un legame efficace alla tasca enzimatica[A]. Tali conoscenze sono state il punto di partenza per lo studio quantitativo LIE di derivati benzo-imidazolici[B,C] e cumarinici[D] che hanno evidenziato il contributo chiave dell'effetto idrofobico per un elevato potere inibitorio verso CK2 e di diverse interazioni di van der Waals con le varie catene laterali apolari che caratterizzano il binding pocket di tale chinasi. È stato utiliz-

zato invece l'approccio MM-PBSA per l'analisi delle differenze di inibizioni, con importanti implicazioni farmacologiche per la selettività, del farmaco Gleevec (STI-571) e del suo stretto analogo CGP-582 verso la protein tirosin chinasi Abl. Tale metodica è stata combinata con successo con l'approccio Colony Energy per migliorare la robustezza dei risultati diminuendo la dipendenza della stima dell'energia di legame da piccole variazioni conformazionali.

La conoscenza delle caratteristiche chiave per l'inibizione è il primo passo del processo di drug discovery basato su informazioni biostrutturali ed il protocollo integrato computazionale utilizzato in tale studio ha reso possibile comprendere le peculiarità chimico fisiche dei bersagli terapeutici e dei ligandi consentendo lo sviluppo di nuovi inibitori a potenziali attività antitumorale.

A R. Battistutta, M. Mazzorana, L. Cendron, **A. Bortolato**, S. Sarno, Z. Kazimierczuk, G. Zanotti, S. Moro and L. A. Pinna. Role of Water Molecules in Affecting Potency and Selectivity of Protein Kinase CK2 Inhibitors. *Chembiochem*, 2007, **8**,1804-1809.

B **A. Bortolato** and S. Moro. In Silico Binding Free Energy Predictability by Using the Linear Interaction Energy (LIE) Method: Bromobenzimidazole CK2 Inhibitors as a Case Study. *J. Chem. Inf. Model.* 2007, **47**, 572-82.

C **A. Bortolato**, E. Fioravanzo, M. Mabilia and S. Moro. Linear Interaction Energy (LIE) as interesting tool to predict ligand-receptor binding free energy. *Il Farmaco*, October 2007.

D A. Chilin, R. Battistutta, **A. Bortolato**, G. Cozza, S. Zanatta, G. Poletto, G Zagotto, E. Uriarte, A. Guiotto, L. A. Pinna, F. Meggio and S. Moro. Coumarin as attractive casein kinase 2 (CK2) inhibitor scaffold: an integrate approach to elucidated the putative binding motif and explain structure-activity relationships. *J. Med. Chem.*, Accepted 2008.

Summary

Cells respond to externals cues thanks to several different signals cascade activated by transmembrane receptors as G coupled protein receptors and tyrosin kinase receptors. Signals are not simply transducted, but also amplified and propagated using a precise tuned elaborating system resulting in a well defined cell behaviour. In this contest, protein kinases are key enzymes for cell life, thanks to their signal transduction pathways control. A perturbation of their activity can influence in an important way this essential equilibrium, resulting in several pathologies as cancer, diabetes and inflammation. These enzymes act controlling substrates activity phosphoryling some specific residues of Serine, Threonine and Tyrosin, using an ATP molecule as phosphate donor. In this study have been studied principally two protein kinases, CK2 and Abl, implicated with the development and progression of several tumors. The first, know also as Casein Kinase 2, is a pleiotropic, constitutively active protein demonstrated to be implicated in many physiological process and over-expressed in several cancer cell lines. On the other hand the protein tyrosin kinase Abelson (Abl), together with PDGFR and Kit enzymes, is the molecular cause of chronic myelogenous Leukaemia and gastrointestinal stromal tumor.

Protein kinases are an interesting family of molecular targets for a new generation of antitumors drugs aiming the molecular mechanisms exploited by cancer to survive and develop. This study represents an attempt to reach this result by means of a rational approach to design new small organic molecules with a potential antitumoral activity based on their inhibition of these protein kinases. Methodologies used rely on the most advanced drug design computational techniques based on the threedimensional structure of the therapeutic target and they have been used only after a intense internal validation phase based on the experimental data available in literature and on important collaborations with several research groups at national and international level.

Protein target morphologies can derive from NMR spectroscopy, X-ray diffraction spectroscopy and computational model derived from templates of known

structure. If the structure of the inhibitor-protein complex is not know, it can be evaluated using a molecular docking study, methodology that aims to the prediction of the ligand most energetically favorable conformation in the enzymatic pocket. In the case that the amount of structural data known are not enough, the performance of this approach can be improved exploiting results of a structure-activity relationshionship study of a known inhibitors set, that can be used to develop a pharmacophore model. The pharmacophore is the qualitative collection of the key chemical-physical features for the inhibitor interaction with the active site and can be used to drive the conformational search during the docking or to select the best conformations obtained. The next step, the free energy of binding estimation, can result or from a mathematical function (scoring function) derived from the interaction decomposition in different energy contributions, often optimized on empiric basis, or from a consensus of different scoring functions. More computational demanding approaches, but more rigorous from a chemical-physical point of view, are the Linear Interaction Energy method (LIE) and the Molecular Mechanics Poisson Boltzmann Surface Area (MM-PBSA). The first creates an energy model using a linear correlation of three energy terms (a van der Waals contribution, an electrostatic part and a not-polar contribution) to an experimental data set, evaluating the free energy difference from the inhibitor free in solution state and the target bounded state. On the contrary in the MM-PBSA approach is not present any empiric component and the energy is evaluated using a molecular mechanics force field and an implicit solvent on a collection of conformations derived from a molecular dynamic study

Using these computational methodologies a complete study of CK2 ATP-binding pocket has been carried out, investigating its electrostatic features and creating several pharmacophoric hypothesis based on crystallographic structural data. These approaches helped us to evaluate from a qualitative point of view the interactions crucials for the inhibitor-protein binding and to highlight the importance for the binding mode in the enzymatic pocket of some waters molecules[A]. These informations have been the starting point for a quantitative LIE study of benzoimidazolic[B,C] and coumarins derivatives[D]. They have point out the key contribution of the hydrophobic effect to achieve a powerful inhibitor effect of CK2 and of van der Waals interactions with several apolar side chains characterizing the binding site of this kinase. On other hand the MM-PBSA approach has been used to analyze the different inhibition power, with important pharmacological implications on the selectivity, of the drug known as Gleevec (STI-571) and its similar analogue CGP-582 versus the protein kinase Abl. This

method has been combined with success with the Colony Energy approach to improve the results robustness, diminishing the dependence of the free energy of binding estimation from small conformational changes.

The key features knowledge for the inhibition is the first step for the drug discovery process based on biostructural information and the integrate computational protocol used in this study had allowed the comprehension of the chemical-physical peculiarities of therapeutical targets as well as ligands, allowing the development of new inhibitors with potential antitumoral activity.

A R. Battistutta, M. Mazzorana, L. Cendron, **A. Bortolato**, S. Sarno, Z. Kazimierczuk, G. Zanotti, S. Moro and L. A. Pinna. Role of Water Molecules in Affecting Potency and Selectivity of Protein Kinase CK2 Inhibitors. *Chembiochem*, 2007, **8**,1804-1809.

B **A. Bortolato** and S. Moro. In Silico Binding Free Energy Predictability by Using the Linear Interaction Energy (LIE) Method: Bromobenzimidazole CK2 Inhibitors as a Case Study. *J. Chem. Inf. Model.* 2007, **47**, 572-82.

C **A. Bortolato**, E. Fioravanzo, M. Mabilia and S. Moro. Linear Interaction Energy (LIE) as interesting tool to predict ligand-receptor binding free energy. *Il Farmaco*, October 2007.

D A. Chilin, R. Battistutta, **A. Bortolato**, G. Cozza, S. Zanatta, G. Poletto, G Zagotto, E. Uriarte, A. Guiotto, L. A. Pinna, F. Meggio and S. Moro. Coumarin as attractive casein kinase 2 (CK2) inhibitor scaffold: an integrate approach to elucidated the putative binding motif and explain structure-activity relationships. *J. Med. Chem.*, Accepted 2008.

Résumé

Pour coordonner leurs activités, les cellules ont besoin de recevoir et de transmettre divers signaux. Souvent les réponses cellulaires aux médiateurs de la signalisation impliquent des séquences de réactions chimiques. Ces cascades de signalisation sont activées par des récepteurs transmembranaires comme p.ex. les récepteurs couplés à la protéine G ou les récepteurs protéine-kinases. Une telle organisation de la transmission permet l'amplification, la modulation ainsi que la propagation strictement contrôlées du stimulus et garantit le bon fonctionnement cellulaire. En tant que régulateurs de la signalisation cellulaire, les protéine-kinases sont des enzymes essentielles au maintien de l'homéostase. Dans des conditions physiologiques, elles assurent la croissance, la différentiation et la communication cellulaire. Dans des conditions pathologiques par contre, leur dérégulation a principalement été corrélée à l'apparition de tumeurs, de diabète et d'inflammation. Les protéine-kinases sont des enzymes qui catalysent la réaction de phosphorylation des protéines au niveau de résidus tyrosine *(tyrosine kinases)* ou de résidus sérine et thréonine *(sérine/thréonine kinases)* présents dans des séquences spécifiques en utilisant ATP en tant que donneur du groupement phosphate. La présente thèse étudie principalement deux protéine-kinases; CK2 et Abl, impliquées dans le développement et la progression de plusieurs cancers. La première, anciennement appelée Caséine Kinase 2, est pleiotropique et présente chez tous les eucaryotes. Son extrme conservation au cours de l'évolution, et le fait qu'elle est constitutivement active sont des propriétés indicatives d'une fonction critique de cette enzyme dans la cellule. CK2 est surexprimée dans plusieurs lignées de cellules tumorales. La tyrosine kinase d'Abelson (Abl) avec PDGFR et Kit représente, à son tour, l'origine moléculaire de la leucémie myélode chronique et de la tumeur stromale gastro-intestinale.

Les protéine-kinases sont des cibles moléculaires d'un grand intért pour la recherche des traitements anticancéreux de nouvelle génération. La conception rationnelle de nouveaux médicaments vise à interférer avec les mécanismes moléculaires que la maladie utilise pour survivre et pour se développer. Dans

le cadre de cette approche pharmacologique, ce travail décrit le développement rationnel d'inhibiteurs de la CK2 et de l'Abl protéine-kinases en tant qu'anti tumoraux potentiels. Les méthodologies utilisées sont basées sur les techniques computationnelles les plus avancées et considèrent la cible thérapeutique dans sa structure tridimensionnelle. Toute approche est utilisée uniquement après l'établissement d'un protocole de validation interne et chaque validation repose sur les données publiées dans la littérature et les importantes collaborations scientifiques d'ordre national ou international.

Les structures tridimensionnelles des protéines cibles analysées ici ont été définies par cristallographie aux rayons X, par spectroscopie RMN ou par modélisation par homologie. Dans le cas o la structure du complexe protéine-inhibiteur n'était pas connue, celle-ci a été définie à l'aide des études de docking moléculaire. Le docking moléculaire in silico évalue l'ensemble des interactions intervenant lors de la formation de complexes entre le ligand (petite molécule) et la cible biologique d'intért thérapeutique, généralement protéique. Le logiciel de docking a donc été utilisé pour la prédiction de la conformation la plus énergétiquement favorable du ligand dans sa poche enzymatique. Dans les cas o la quantité de données structurelles pour le ligand n'a pas été suffisante, la performance du docking a été améliorée en y incorporant les résultats obtenus par une étude structure-activité (SAR). L'étude SAR a été basée sur un set d'inhibiteurs connus et utilisés pour créer un modèle pharmacophore. Ce dernier représente l'ensemble qualitatif des caractéristiques physico-chimiques assurant l'interaction inhibiteur-protéine selon un arrangement spatial adéquat et peut tre utilisé pour la sélection conformationnelle lors du docking ou pour désigner les meilleures résultats du docking. La deuxième étape est l'estimation de l'énergie libre de liaison, qui peut tre obtenue à partir d'une équation mathématique empirique nommée ≪scoring function≫ qui inclut toutes les contributions énergétiques participant à l'énergie libre de liaison ou de la combinaison du résultat de plusieurs ≪ scoring functions ≫ qui a comme but de voir le consensus entre les différents résultats (Consensus Scoring). Pour estimer l'énergie libre de liaison il y a d'autres approches, plus coteux en terme de temps machine, mais aussi plus rigoureux du point de vue physico-chimique, comme ≪Linear Interaction Energy Method (LIE)≫ et ≪Molecular Mechanics Poisson Boltzmann Surface Area (MM-PBSA)≫. Le premier crée un modèle énergétique à partir d'une corrélation linéaire entre trois termes énergétiques (les contributions de van der Waals, électrostatique et des parties non polaire) faite sur un set de données expérimentales, qui évalue la différence d'énergie entre l'état de l'inhibiteur libre en solution et en complexe avec la cible thérapeutique. L'approche MM-PBSA

ne contient aucun composant empirique et l'énergie libre de liaison est estimée en partant d'une série de conformations obtenues par des études de dynamique moléculaire sur le complexe.

Les méthodes computationnelles décrites plus haut ont permis la description intégrale du site de fixation d'ATP de la CK2 protéine kinase. La caractérisation a été basée sur les études des propriétés électrostatiques de ce domaine protéique ainsi que sur l'établissement de plusieurs pharmacophores développés à partir des structures cristallographiques de la protéine. Une telle démarche a également démontré de faon qualitative les interactions essentielles à la liaison ≪inhibiteur-protéine kinase≫ et l'importance de certaines molécules d'eau de la poche enzymatique pour le mode de liaison[A]. Ces informations ont été approfondies par une étude quantitative LIE sur des dérivés de *benzoimidazole*[B,C] et de *coumarine*[D]. Les études quantitatives ont démontré l'importance du caractère hydrophobe de l'inhibiteur pour son activité sur CK2. La force d'inhibition a été corrélée également à des liaisons de type van der Waals que l'inhibiteur établit avec plusieurs chanes latérales apolaires au sein du site de liaison de la kinase. Pour étudier la spécificité de liaison, l'approche MM-PBSA a été adoptée afin d'analyser les différences entre les forces d'inhibition de Gleevec (STI-571) et de CGP-582 sur Abl protéine-kinase. Cette dernière méthode, combinée à Colony Energy, s'avère solide et peu dépendante des variations de l'énergie libre lors de petits changements conformationnels.

La connaissance des caractéristiques biologiques et structurelles essentielles à l'inhibition est le premier pas vers une nouvelle recherche rationnelle de médicaments. Le protocole intégral, que représente cette thèse, permet la compréhension des particularités physico-chimiques de certaines cibles thérapeutiques et de leurs ligands dans le but de faciliter le développement de nouveaux inhibiteurs des protéine-kinases en tant que traitement contre le cancer.

A R. Battistutta, M. Mazzorana, L. Cendron, **A. Bortolato**, S. Sarno, Z. Kazimierczuk, G. Zanotti, S. Moro and L. A. Pinna. Role of Water Molecules in Affecting Potency and Selectivity of Protein Kinase CK2 Inhibitors. *Chembiochem*, 2007, **8**,1804-1809.

B **A. Bortolato** and S. Moro. In Silico Binding Free Energy Predictability by Using the Linear Interaction Energy (LIE) Method: Bromobenzimidazole CK2 Inhibitors as a Case Study. *J. Chem. Inf. Model.* 2007, **47**, 572-82.

C **A. Bortolato**, E. Fioravanzo, M. Mabilia and S. Moro. Linear Interaction Energy (LIE) as interesting tool to predict ligand-receptor binding free energy. *Il Farmaco*, October 2007.

D A. Chilin, R. Battistutta, **A. Bortolato**, G. Cozza, S. Zanatta, G. Poletto, G Zagotto, E. Uriarte, A. Guiotto, L. A. Pinna, F. Meggio and S. Moro. Coumarin as attractive casein kinase 2 (CK2) inhibitor scaffold: an integrate approach to elucidated the putative binding motif and explain structure-activity relationships. *J. Med. Chem.*, Accepted 2008.

Chapter 1

Introduction

1.1 Protein Kinases

1.1.1 The Kinome

The kinome, defined as the protein kinases (PKs) complement of the human genome, encompasses more than 500 enzymes. They are grouped into a number of subsets based primarily on sequence and structural similarities [1]. The phylogenetic tree shown in Fig. 1.1 [2] highlights the structural relationships of the different protein kinases. PKs catalyze the transfer of the ATP terminal phosphoryl group to the peptide substrate:

$$\text{MgATP}^{-1} + \text{Protein-OH} \longrightarrow \text{Protein-OPO}_3^{2-} + \text{MgADP} + \text{H}^+$$

On the contrary the phosphatases remove the phosphate group from the substrate allowing the reversibility of the phosphorylation cycle [3]. These enzymes together exert a thigh and reversible control on proteins phosphorylation. Both these enzymatic categories can be subdivided on tyrosine- or serine/threonine-specific, based on their catalytic specificity. In addition some of them, like Casein Kinase 2 (CK2), possess dual specificity for both tyrosine and serine/threonine.

1.1.2 Kinases Control of Life

PKs are versatile and sophisticated proteins able to strictly control the complexity of cell signalling networks. Thanks to an interplay of activations and inhibitions they can produce a wide range of responses, tuning a precise cell behavior. The signal is transduced, amplified, elaborated and propagated using
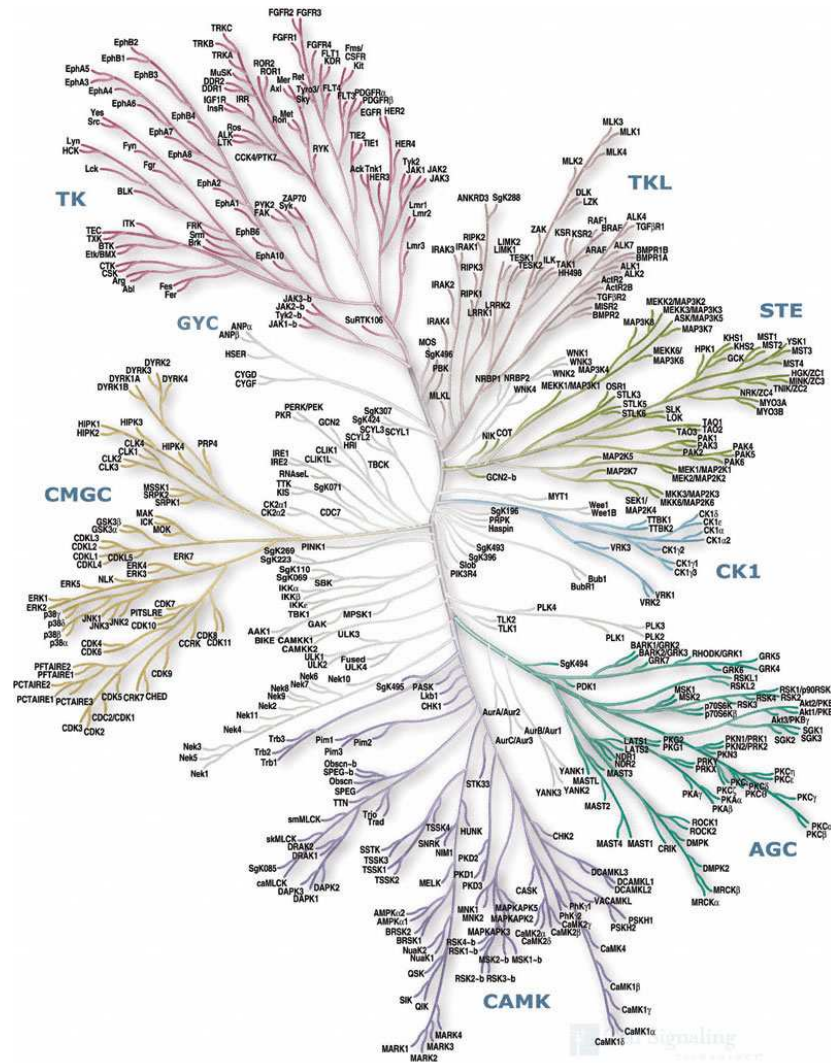
Figure 1.1: Phylogram of the human protein kinase superfamily inferred from amino acid sequences of the kinase domain.The protein kinase subsets have been defined as follows: AGC containing PKA, PKG, PKC families; CAMK Calcium/calmodulin-dependent protein kinase; CK1 Casein kinase 1; CMGC containing CDK, MAPK, GSK3, CLK families; STE Homologs of yeast Sterile 7, Sterile 11, Sterile 20 kinases; TK Tyrosine kinase; TKL Tyrosine kinaselike. Other kinases are shown in the center of the tree, colored gray. The relationships shown on the tree can be used to predict protein substrates and biological function for many of the over 100 uncharacterized kinases presented here. The protein kinases studied in this research work are CK2 (dual specificity Kinase) and Abl (belonging to the Tyrosin Kinase family) [2].

a well defined myriad of interconnections linking different signalling pathways characterized by redundancy and spatial-temporal modulation. In such a way cells can answer robustly to unwanted perturbations and at the same time can respond specifically and sensitively to relevant inputs [4].

With such a critical role in signal transduction, the tight regulation of kinase activity is crucial and linked to several cellular process like cell metabolism, transcription, cytoskeletal rearrangement and cell movement, cell cycle progression and cell differentiation, apoptosis, physiological responses, homeostasis and function of the immune system. A perturbation of kinases activity from activating and inactivating mutations results in a wide number of disease, including cancer, diabetes and inflammation [5].

1.1.3 Structure and Regulation

All protein kinases possess the catalytic domain, called also kinase domain, of about 300 residues, with common structural features. Its two-subdomains architecture consists of an N-terminal lobe, with five-antiparallel-stranded β-sheets and one prominent α-helix (αC helix), and a C-terminal lobe, with mainly α-helixes. The flexibility of this catalytic domain allows a complex auto-regulatory mechanism with important conformational changes. ATP binds in the cleft between the two lobes and the part of the substrate with the tyrosine residue interacts with amino acids in the C-terminal lobe (Fig. 1.2). Some regions of the active site result determinant even for the activation of the protein: with their conformational changes they contribute to open and close the catalytic site to the ATP and the substrate. These regions are the nucleotide-binding loop (P-loop) that is a conserved loop forming the roof of the active site, the α-C helix and the activation loop (A-loop), a 20-25 residue segment which serves as platform for substrate binding. The DFG motif is constitute by the first three amino acids of the A-loop. These residues highly conserved are an aspartate, essential for the enzymatic activity, a phenylalanine and a glycine. All these regions adopt distinct conformations that are associated with the active and inhibited state of the kinase domain. In many kinases, the αC-helix can rotate and change its position, thereby changing the orientation of key catalytic residues. *Trans*-phosphorylation of one or more tyrosines in the activation loop leads to a repositioning of the A-loop and a concomitant increase in catalytic efficiency. In the unphosphorylated state, the A-loop acts as a pseudo-substrate inhibitor. The catalytic activity of a PTK that has not been activated, for example by ligand-mediated phosphorylation, is defined as basal-level activity. This state is maintained through these various auto-regulatory mechanisms.
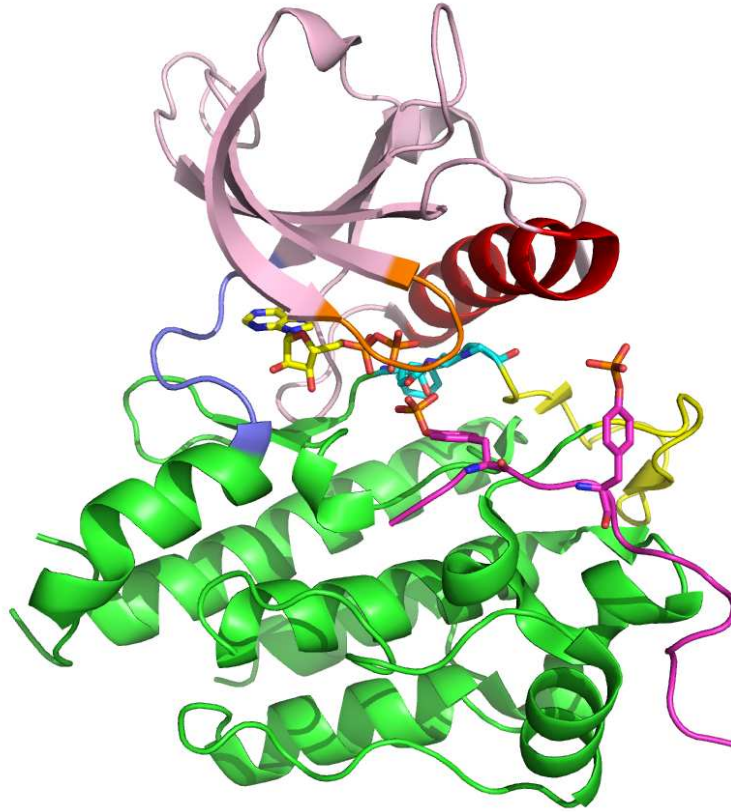
Figure 1.2: Protein kinase domain architecture. This is the ribbon diagram of the x-ray diffraction crystal structure of c-Kit PTK (PDB entry 1PKG). The N-lobe is shown in pink, the C-lobe in green, they are linked by the hinge region in violet, the C-helix in red, the P-loop in orange, the A-loop in yellow. In the middle between the two lobes there is the ADP (carbons in yellow). The substrate, with phosphorylated tyrosines, is colored in magenta and the DFG motif in cyan.

The high level of complexity of cell signal pathways ruled by protein phosphorylation is possible from a molecular point of view thanks to precise protein-protein interactions [6] (Fig. 1.3). Different levels of organization are controlled by specific chemical-physical interactions:

1. Substrate target site specificity: kinase active sites might only accommodate certain amino acid sequences. These preferences, however, especially for tyrosine kinase are not very stringent and often the preferred sequences are not very different. In general this is not enough to explain the in vivo specificity observed in kinase signaling pathways.

2. Modular protein-protein recognition domains, often fused to the catalytic domain. These modular protein domains are present in a variety of other proteins involved in cell signaling, like phospholipases and noncatalytic adaptor proteins. A common example of such domains, that play a important function in signal transduction are the SH2 and SH3 domains, that mediate specific protein-protein interactions [7]. SH2 domains recognize short peptide motifs containing phosphotyrosine, and SH3 domains bind to proline-rich sequences. Both can act as adapters to promote complex formation, or modulate catalytic activity [8].

3. Allosteric site: interactions involving binding surfaces of the catalytic domain but different from the catalytic active site.

While the catalytic site must satisfy precise stereochemical properties to keep the phosphorylation capability, the other regulation mechanisms can freely vary and evolve to achieve the desired protein recognition needed for the complex networks regulation. The final pourpose is simply separate protein association from the catalytic function, influencing kinase activity through allosteric mechanisms.

Interestingly in general allosteric sites are present in serine/threonine kinases and phosphatases, while on the contrary the majority of them does not contain any recognition or targeting domains, including protein-protein, protein-lipid and transmembrane motifs. On other hand, the structure regulation of tyrosine kinases and phosphatases is based on one or more recognizable modular targeting elements outside of the catalytic domain [9]. This observation is in agreement with the theory that the signaling system based on serine/threonine phosphorylation evolved before the tyrosin phosphorylation, based instead on recombination with other modular interaction domains

Moreover another level of organization that governs specificity in vivo is based on the protein control in space and time: protein subcellular distribution, multimolecular organization and the timing of their activation ensure that they will only encounter certain potential substrates [10].

Physiological regulation of PTKs is the key to understand mechanisms causing their oncogenic activation [11, 12].

1.1.4 Difficulties of Pharmaceutical Research

Today more than ever before, it is available a tremendous technical development on all fields of life sciences, from the characterization of molecular mechanisms underlying cellular function to resources to test and evaluate pharmaceutical
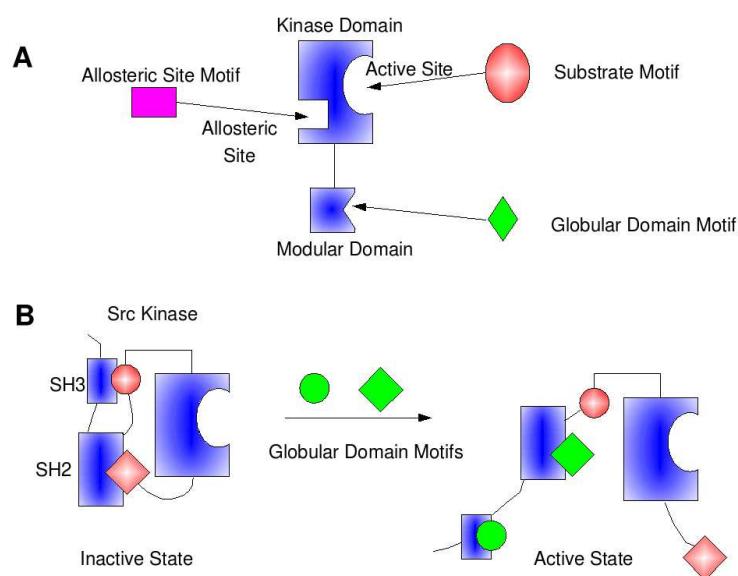
Figure 1.3: Protein kinase activity control systems. On the upper panel three kinases and phosphatases protein recognition site are shown: appropriate proteins pattern can bind the allosteric site or modular domain to regulate enzyme activity. On the lower panel the Src control system operated by SH2 and SH3 domains is shown as example.

targets and chemical compounds and from cell and tissue analysis to molecular diagnostics [13, 14]. With such impressive expertise we are all waiting for a major revolution in the treatment of human disease. However recent reports show 30-year decline in pharmaceutical research and development productivity: the ever-increasing spending on pharmaceutical research did not result in an improving of the rate of new drug approvals [15, 16]. In the period 1993-2004 the industry research and development investment increased of 147%, while the drugs approvals have risen only 38% [17, 18]. If we can state that drug discovery is difficult in general, for oncology despite the huge public and private investment the success rate is three time lower than for cardiovascular disease as shown in Fig. 1.4 [19, 20].

1.1.5   Cancer Drug Discovery and Protein Kinases

Nowadays almost 200 drugs, with a large variety of mechanisms of action, have been approved for therapeutic use and hundreds are in development:
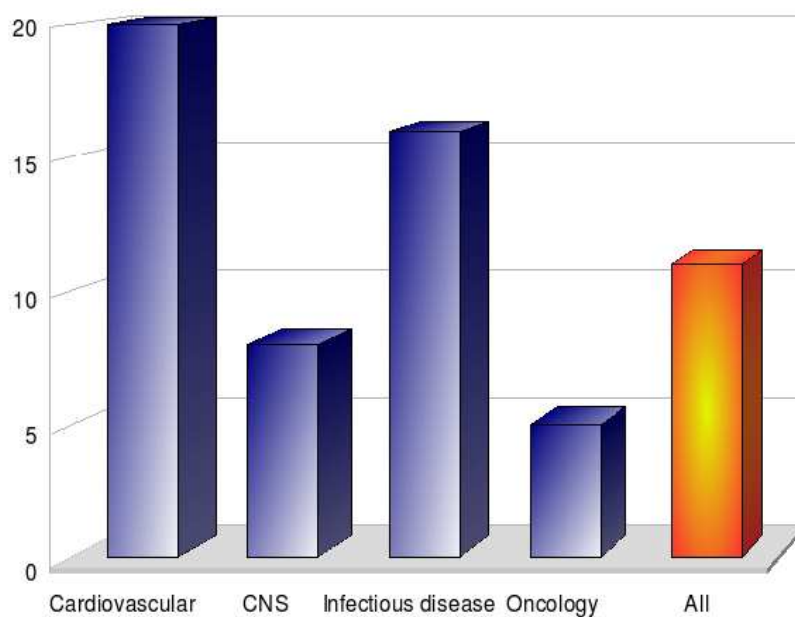
Figure 1.4: Percentage of success rates from first-in-human to registration for ten large pharmaceutical companies (AstraZeneca, Bristol-Myers Squibb, Eli Lilly, F. Hoffman-LaRoche, GlaxoWellcome, Johnson & Johnson, Novartis, Pfizer, Pharmacia, Schering-Plough and SmithKline Beechamin)in the United States and Europe for the period 1991-2000 are shown [19].

Mechanism of action	Target example (drug)
Activation of apoptosis	BCL2
Interference with signalling	ABL (Gleevec)
Metastasis inhibition	Cathepsin K
Induction if senescence	Telomerase
Angiogenesis interference	VEGF (Avastin)
Antibody-directed cytotoxicity	CD20 (Rituxan)
Interference with cell division	Microtubules (Taxol)
DNA-modifying agents	DNA (Cisplatin)
Reduction of essential metabolite	Thymidylate synthase (5-FU)
Inhibition of acceleration of protein degradation	Proteasome (Velcade)
DNA methylation	HDAC interactions
Interference with cellular stress buffering	ATPase/chaperone superfamily
Anti-hormonal activity	Aromatase (Letrazole)

Even if the discrimination cannot be always clear, therapeutics targets for cancer treatment can be distint between essential for the life of at lest one cell type, or not-essential. The latter represent the large part of the proteome. Targeting the firsts often results on narrow therapeutic windows, while on the contrary the other should cause less toxic side effects.

Kinase inhibitors represent a new paradigm in anticancer therapy: small-molecule drugs inhibiting elements in key signalling pathways. This approach can be considered different from traditional cytotoxic tumor therapy since enzymes targeted might not be essential in normal adult cells. For this reason pharmaceutical companies started extensive research projects on these targets and now there are more than 30 kinase inhibitors in clinical trials [21]. In this contest Imatinib, the first kinase inhibitor entered in commerce for the treatment of Chronic Myeloid Leukemia (CML), is a prototype for these new cancer drugs thanks to its action on Abelson kinase (Abl), the molecular basis of the pathology. The low toxic profile of Imatinib, a precious feature for an anticancer drug, is linked to the fact that most probably animals have a limited requirement for Abl activity, as demonstrated by the mouse knockout phenotype. On other hand targeting kinases belonging to pathways crucial for survival and proliferation might behave like cytotoxic drug with narrow therapeutic windows [22].

An additional problem is linked to capability of cancer cells to acquire resistance during the progress of the therapy. For this reasons clinicians suggest that a possible solution can be related to a multi-targeting drug [23] and in particular many clinical-stage kinase inhibitors are fairly non-selective [24], however the risk can be a less predictable effect and toxicity. However Imatinib, thanks to its activity on three protein kinase, Abl, PDGFR and c-Kit, demonstrated how a multi-targeting action can improve the final effect and amplify the therapeutic range of action from CML to gastrointestinal stromal tumor (GIST) and the hypereosinophilia [25, 26]

1.2 The Many Faces of CK2, Casein Kinase 2

1.2.1 A Cell Destiny Protagonist

CK2 represents one of most exceptional protein kinase. With its physiological constitutive activity [27] and its ubiquitous presence in eukaryotic organisms this enzyme has been demonstrated to be able to phosphorylate more than 300 substrates, and probably this number is going to grow again [28]. It is a messenger-independent protein serine/threonine kinase [29], but experimental data supports its dual-specificity, acting also on tyrosines residues [30]. It was

one of the firsts PKs discovered, as enzyme able to phosphorylate in vitro the casein, from this Casein Kinase II [31], even if now it seems very unlikely that in vivo this could happen.

In humans CK2 exists in a tetrameric form composed by two catalytic units, CK2α with three possible isoforms, and two regulatory units, CK2β [32]. In the majority of the case the catalytic unit has demonstrated to phosphorylate a particular consensus sequence different from other PK known. This is composed by 4 aminoacids: Ser-X-X-Acidic, where the acidic residues can be Glu, Asp, pSer or pTyr [33]. On the other hand the β units are able to stabilize the tetrameric complex and at the same time to enhance and to modulate the activity thanks to a crucial role in substrates recruitment [34,35]. Therefore the answer to the physiological constitutively activity paradox of CK2 seems to rely on this regulation unit, but also has been hypothesized a regulatory high order interactions between CK2 tetramers.

The sophisticated charm of CK2 regulation continues in cells. This protein kinase is something different, the first rule is not followed: phosphorylation is not essential for activation [36,37]. There is not a clear and strong regulation system, but maybe a collection of mechanisms that participate to some degree to it, and scientists struggle with not always consisten studies and hypothesis. The human capability to understand is unfortunately linked to the possibility to simplify and CK2 with its 300 substrates and even more interacting proteins, does not seem to like this. In the arena there are regulation interactions with proteins like FGF-1 [38], tubulin [39] and Hsp90 [40], or positive charged compounds like polyamines [41]] and possible assembly regulation systems. A fascinating theory comes from the analogy between CK2β and cyclins [42], the essential regulatory units of cyclin-dependent kinases. Interesting, in agreement with this, the level of CK2 presence in cells corresponds to their proliferation [43].

CK2 multitude of proteins targets and its presence in a variety of cellular compartments is linked to its large number of functions [44, 45]. It is essential for viability [46] and it regulatory control of several protein kinase pathways that have been demonstrated to be linked to transformation and cancer [47]. In general it seems to have an anti-apoptotic activity supported by its modulation of proteins susceptibility to caspases, it protects them from caspase-mediated degradation [48]. Moreover experimental data support CK2 ability to protect cells from drug-induced apoptosis. On other hand CK2 essential role for cell cycle, in particular for the progression from G_1/S and G_2/M transition [46], is the consequence of its regulation of proteins important for cell cycle, like p34^{cdc2} [49], cdc34 [50] and topoisomerase II [51].

1.2.2 CK2 Key Role in Cancer Development

Several experimental data support the CK2 importance for cancer transformation and development: CK2 has been found to be overexpressed in tumors in the head and neck [52], in prostate [53], in the kidney [54], in mammary gland [55] and lung [56]. The final effect on CK2 seems linked to an action on both oncogenes and tumor suppressor proteins. It promotes a number of proto-oncogenic products stimulating cell proliferation and differentiation and inhibiting apoptosis. Several pathways are influenced by CK2, thanks to its control of substates like β-catenin [57], NF-κB [58], Max [59], c-Jun [60], c-Myc [61] and c-Myb [62]. Among the tumor suppressor proteins phosphorylated by CK2 there is PTEN resulting in its inactivation. This protein is involved in the regulation of cell survival through its action on the phosphatidylinositol-3 kinase pathway [63]. Another example of CK2 substrate is p53, implicated in the regulation of differentiation, cell cycle and apoptosis [64, 65] and p38 MAP kinase [66]. Also PML, a growth suppression protein, is influenced by CK2: its phosphorylation at Ser517 induces its degradation in proteasome [67].

1.3 CK2: From Structural Informations to Inhibitors Design

1.3.1 Inhibitors Scaffold Census

The CK2 importance in cancer progression and development as well as the observation that many viruses are exploited in host cell CK2 for the phosphorylation of proteins essential to their life cycle, indicates this enzyme as a potential target for antineoplastic and antiviral drugs [68]. As soon as the pathological role of CK2 was demonstrated the research of potent and selective CK2 inhibitors started and today are know several ATP-competitive small molecules inhibitors with different scaffold and pharmacokinetics.

Benzimidazole Derivatives

The first class of CK2 inhibitors derived from the molecule 5,6-dichloro-1-(b-D-ribofuranosyl)benzimidazole (DRB) [69] a nucleoside analogue. The starting activity of 23 μM was improved removing the sugar moiety [70] and then exploring the potentiality of the benzoimidazole scaffold decorability. In such a way several potent and selective inhibitors were discovered [71,72] and their pharmacodynamic properties explored from a structural point of view by means of X-ray

diffraction spectrometry [73–75]. Among them TBB (4,5,6,7-tetrabromobenzotriazole), selective for CK2 on a panel of 30 protein kinase [76], was demonstrated to have a Ki of 0.4 μM, further improved of one folder in the analogue K25 (2-dimethylamino-4,5,6,7-tetrabromo-1H-benzimidazole) [77] (Fig. 1.5). A more recent series of derivatives are instead based on the isoindoldione scaffold, the best inhibitor of the series, TID46 (2-(4,5,6,7-tetraiodo-1,3-dioxo-1,3-dihydro-2H-isoindol-2-yl)propanoic acid), reaches a IC$_{50}$ of 0.15 μM [78](Fig. 1.6).
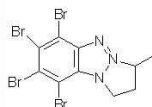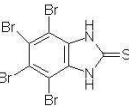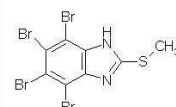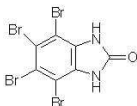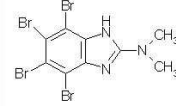
PDB code	Inhibitor co-crystallized		PDB code	Inhibitor co-crystallized	
2OXY		**TBI, K17, TBBz** Ki=0.30μM IC$_{50}$=0.70μM	1ZOH		**K44** Ki=0.10μM
2OXX		**K22** Ki=0.20μM IC$_{50}$=0.91μM	1ZOG		**K37** Ki=0.07μM
2OXD		**K32** Ki=0.15μM	1ZOE		**K25, DMAT** Ki=0.04μM IC$_{50}$=0.14μM
1J91		**TBB** Ki=0.40μM IC$_{50}$=0.90μM			

Figure 1.5: Inhibitors with benzimidazole scaffold co-crystallized in complex with CK2.

Carboxyl Acid Derivatives

A structurally evolution of the poli-bromurated benzoimidazole scaffold can be considered the cinnamic acid derivatives, and in particular the most active, TBCA the tetrabromo analogue, showed a Ki of 0.1 μM [79](Fig. 1.6). This molecule possess an carboxyl acid moiety presents also in another scaffold rapresented by the indoloquinazoline IQA (5-oxo-5,6- dihydroindolo-(1,2-a)quinazolin-7-yl acetic acid, also know as CGP029482) from Novartis(Fig. 1.7). Even if it showed a good activity (IC$_{50}$= 0.39 [80]), this scaffold was not further developed cause instability problems in water where its lactam ring is subjected to slow hydrolysis. An interesting aspect common to IQA and one benzimidazole derivative, K37 (4,5,6,7,-tetrabromo-2-(methylsulfanyl)-1H-benzimidazole), is
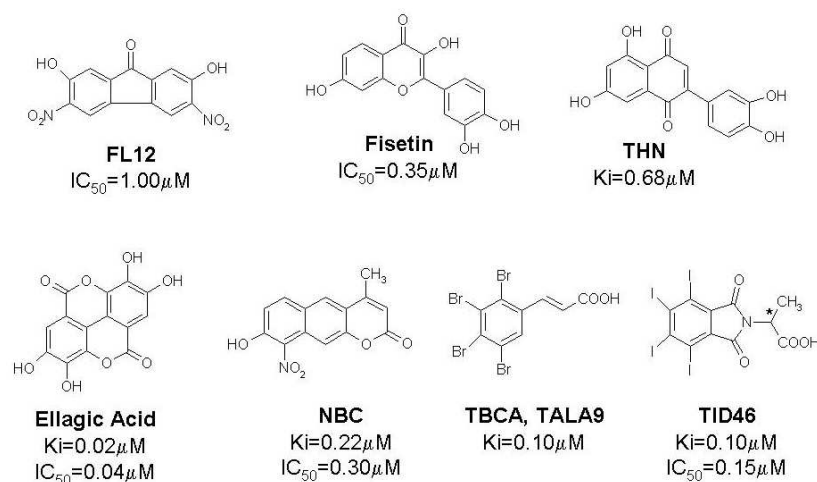
Figure 1.6: Remarkable CK2 inhibitors without any structurally information freely available.

the fact that both show two possible orientations in the crystal structure, thing quite peculiar of CK2, linked to the possibility of comparable potential energy wells in the binding energy surface separated by low, easily to overcome energy barriers.

Figure 1.7: Inhibitors with antraquinone scaffold co-crystallized in complex with CK2. The coumarin DBC and the quinazolin derivative IQA is also shown.

Antraquinone, Xantenone and Fluorenone

A third important class of inhibitors is based on the antraquinone and xantenone scaffold. Four X-ray crystal structures [81, 82] demonstrated how this versatile structure, if adequately functionalized could be compatible with polar interactions with both the hinge region or the opposite region near the DFG motif Aspartic (Asp175). These molecular scaffolds were initially chosen for the know potential activity as antiviral, antimicrobial or anticancer drug (19), even if there is the risk of an action as DNA-intercalators with expectable cytotoxic effects. However the good results obtained for emodin (Ki=1.5μM), a natural compound presents in the Rheum palmatum used for a long time in the Orient to cure inflammatory and neoplastic diseases, encouraged a intensive structure-based inhibitors screening resulting in compound submicromolar. Among them 1,8-dihydroxy-4-nitroanthraquinone (MNA), 1,8-dihydroxy-4-nitro-xanthen-9-one (MNX), and DAA (1,4-diamino-5,8-dihydroxy-anthraquinone) [82] (Fig. 1.7). Other scaffold similar to them are fluorenone based, like FL12 (2,7-dihydroxy-3,6-dinitro-fluoren-9-one) with an IC$_{50}$ of 1μM and THN, a tetra-hydroxy-benzonaphtone with an IC$_{50}$ of 0.68 μM, also inspired to the flavonoids inhibitors [83, 84](Fig. 1.6).

Natural Derivatives: Flavonoids, Ellagic Acid and Coumarins

Three other compounds families active versus CK2 are natural products or their derivatives. The first class are flavonoids derivatives, a group of substances present in fruit, vegetables, grains, bark, roots, stems, flowers, tea, and wine with variable phenolic structures [85]. More than 4000 varieties of flavonoids have been identified and several studies demonstrated their importance for scavenging of oxygen-derived free radicals, their activity as antiinflammatory, antiallergic, antiviral, and anticarcinogenic compounds [86]. Moreover they participate, together with other natural compounds to the French paradox [87], the low cardiovascular mortality rate observed in Mediterranean populations in association with red wine consumption and a high saturated fat intake. Several flavonoids demonstrated to have an inhibitory activity on CK2 [83]: Morin (2',3,4',5,7-pentahydroxy-flavone) and Chrysin (5,7-dihydroxy-flavone) showed a low micromolar IC$_{50}$, while Myrecitin (3,3',4',5,5',7-hexahydroxy-flavone), Quercetin (3,3',4',5,7-penta-hydroxy-flavone), Apigenin (4',5,7-tri-hydroxy-flavone) were sub-micromolar and Fisetin (3,3',4',7-tetrahydroxy-flavone) was the stronger flavonoid with an IC$_{50}$ estimated of 0.35 μM (Fig. 1.6).

Another poliphenolic compound Ellagic Acid (Fig. 1.6), a naturally occurring tannic acid derivative, has demonstrated to be a very potent (Ki=20 nM) and

quite specific CK2 inhibitor, capable to simultaneously bind the hinge region and the phosphate-binding region of the ATP-binding cleft [88].

The last family of inhibitors derivatives of natural products are coumarins. As flavonoids, coumarins are largely distributed in plants, like Apiaceae and Rutaceae. It is possible to find reference to these family already in persian books of 980 A. C [89]. They are considered fitolexine as plants products able to give a defence against other organisms or after wounds. Noteworthy is the fact that the dietary exposure to benzopyrones derivatives, as coumarins, is estimated of 1 g/day in western countries [90]. For such reason in literature are present several coumarins compounds with antitumoral activity [91], for example against prostate cancer [92], melanoma [93] and metastatic kidney carcinoma [94]. Moreover several coumarins anticoagulant have shown the capability to low metastasis in animals models. Generally the hydroxy-derivatives were particularly promising [93, 94] as anticancer molecules and in particular for CK2 inhibition the 8-hydroxy-4-methyl-9-nitrobenzo[g]chromen-2-one (NBC, Fig. 1.6) and the 3,8-dibromo-7- hydroxy-4-methylchromen-2-one (DBC), showed a costant of inhibition of respectively 0.22 and 0.06 μM [84]. The structure-activity of this scaffold was then further investigated and the X-ray crystal structure of the CK2-DBC complex solved [95](Fig. 1.7).

Pyrazolo-triazine Scaffold

The last and more recent family is based on the pyrazolo triazine scaffold. These derivatives, developed by Polaris Pharmaceuticals Inc., are the most potent CK2 inhibitors known at present. The better of the series reaches a Ki of 0.35 nM, even if selectivity data are not available [96]. The scaffold was included in a marocyclic ring that improved the cell activity of the molecule [97]. Moreover seven X-ray crystal structures of these derivatives helped the lead optimization process and have highlighted the capability of the better inhibitors to interact with both the hinge region of CK2 and the opposite part of the pocket near, the phosphate binding cleft, as hypothesized for ellagic acid(Fig. 1.8).

1.3.2   Computational Techniques to Target CK2

CK2 rational inhibitors design was based principally on a structure based approach thanks to a multidisciplinary work connecting chemical synthesis, biological assays, structural information. Different molecular modelling techniques have been used to find and to analyze, propose and corroborate hypothesis about essential inhibitors features for CK2 binding. The principal approach used was molecular docking, a computational technique that try to find the
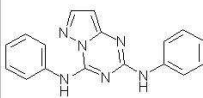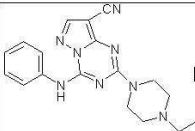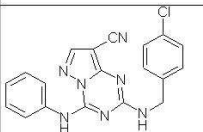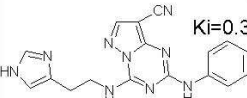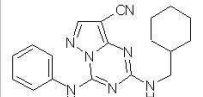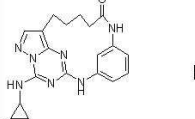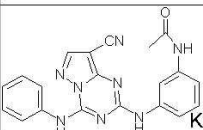
Figure 1.8: Inhibitors with pyrazolo-triazine scaffold co-crystallized in complex with CK2.

best, the most probable conformations of the inhibitor inside the protein active pocket and at the same time to evaluate and estimate as precise as possible the different chemical-physical interactions with the target that are linked to the final activity [98].

At the moment 20 X-ray diffraction crystal structures of CK2 kinase domain in complex with different ATP-competitive inhibitors, are publicly available on the Protein Data Bank (http://www.rcsb.org/pdb/home/home.do) [99]. Always these structural informations were used as starting point and exploited to validate computational approaches. For example for docking we studied the capability of the different programs and of possible combination of search algorithms and scoring function to reproduce the inhibitor binding pose present in the X-ray crystal. A consensus scoring approach [100] for example has allowed us to identify the ellagic acid inhibitor [88].

Moreover inhibitors binding interactions presents in crystal structures were analyzed and utilized to create CK2 optimized scoring functions for docking studies [Moro et al. not published data] and on other hand they were the starting point for Linear Interaction Energy (LIE) method [101]. Thanks to the LIE approach is possible to build up a binding free energy prediction model to rationalize from a quantitative point of view the structure-activity relationship
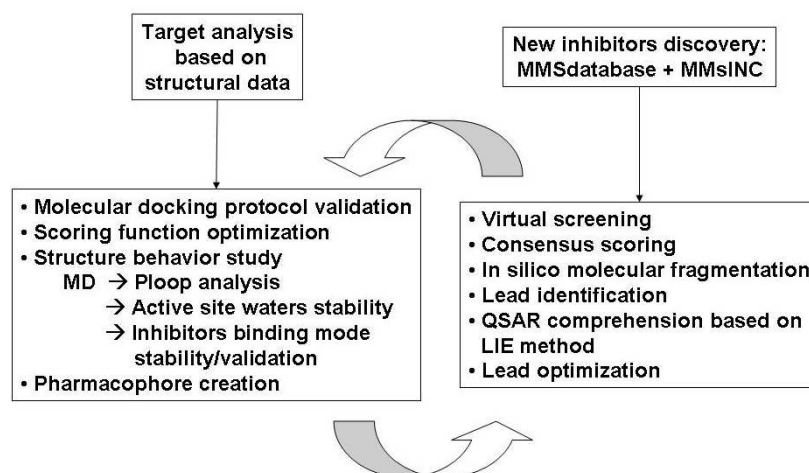
Figure 1.9: Molecular modelings technique used in the rational CK2 inhibitors design. The approach aims to continuous improvement of the protocol based on the informations that is possible to gather from a precise analysis of the structural informations available. Virtula screening has been carried out on an in-house database (MMSdatabase) and MMsINC database (http://mms.dsfarm.unipd.it/MMcINC.html). This last one is a free web-oriented database of commercially-available compounds for virtual screening and chemoinformatic applications. MMsINC contains over 4 million non-redundant chemical compounds in 3D formats. It is provided by the Molecular Modeling Section in the Department of Pharmaceutical Sciences at the University of Padova, (Italy) in collaboration with the Software Support Services & Development Laboratory (S3D) at the Center for Advanced Studies, Research and Development (CRS4) in Sardinia.

of a particular scaffold.

The structure stability and in a particular way the movement of CK2's P-loop and the position in the active site of crucial waters molecules conserved in the crystal structures available were investigated and evaluated by means of molecular dynamics approaches [Moro et al. not published data].

The common idea was to identify and understand the principal interactions important for the final inhibitor activity. A useful representation of this qualitative rationalization could be the pharmacophore, defined by IUPAC as an ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interactions with a specific biological target and to trigger (or block) its biological response. The approach was to distinguish the essential features from the subsidiary properties by means of in silico Molecular Fragmentation. After that the final pharmacophore has been individuated this could be

adapted to one particular scaffold and implemented to a quantitative level using the LIE approach as shown for coumarins [95](Fig. 1.9)

In Silico Molecular Fragmentation and Scaffold Hybridization

A possible approach to fully understand structure recognition is based on in silico molecular fragmentation (Fig. 1.10). In theory is possible to evaluate the free energy of binding as an additive collection of distinct moieties able to create specific chemical-physical interactions. Starting from a lead compound is possible to progressively simplify the scaffold to reach the minimum pharmacophore in such a way that is possible to understand the distinct contribution. The chemical space is sampled in the opposite common direction from a lead to a low complexity fragment. This is based on the fact that a recent analysis [102] indicated that a sufficiently large fragment screen provides excellent coverage of clinically useful chemical space. On other hand once that the minimum scaffold is reached is possible to take a different direction in the chemical space moving toward a new scaffold. Moreover is possible to combine these fragment on the basis of a docking study of their binding mode, or using an approach known as scaffold hybridization where the known positions of two ligands are used to recombine fragments from each to generate a novel ligand. This new molecule will generally be a hybrid of the two scaffolds or a transfer of a substituent from one scaffold to the other [103].

A Common Pharmacophore for CK2 Inhibition

CK2 activity and regulation originality goes on affecting drug design. Again the basic rule is not followed: hydrogen bonding with the hinge region is not needed for binding. The common pharmacophore for CK2 activity is composed by an important hydrophobic/aromatic area in the center of the pocket (Fig. 1.11). This part is essential and present in all submicromolar inhibitor, and it is understandable on the basis of the many apolar bulky aminoacids characterizing the CK2 ATP-binding pocket such as Ile66, Ile174, Met163. On other hand possible hydrogen bonding with the hinge region (backbone of Glu114 and Val117) could improve the activity, even if are not essentials, as demonstrated by the crystal structures of K25 or K44 [74]. On the opposite part of the protein, near the DFG motif, has been individuated a region with the shape of an half-moon with a positive electrostatic potential equal to 1.5 kcal/mol. In this region in the CK2 apo form are present three very conserved waters molecules [75] that can be substituted by sufficient polar group of the inhibitors(Fig. 1.12). In such a way the inhibitor can adapt its chemical-physical properties exploiting these

in silico Molecular Fragmentation Approach
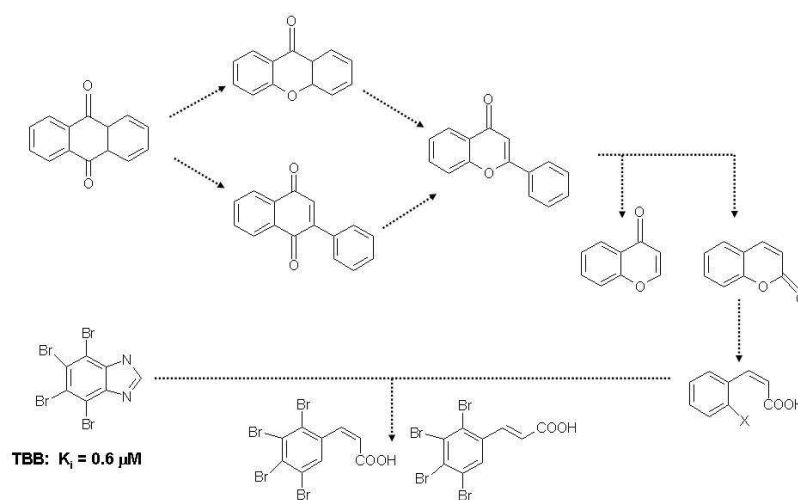
Figure 1.10: Scaffold relationship and evolution in the in silico molecular fragmentation approach used for the rational CK2 inhibitors design.

waters and are possible direct or water-mediated interactions between inhibitors and residues like Asp175, Lys68, Glu81 and Trp186.

Linear Interaction Energy Method

All the contributions and the crucial interactions rapresented with the pharmacophore can be evaluated with the LIE approach. All the contributions in this way are compared and a final weight of the hydrophobic, van der Waals and electrostatic contribution to the free energy of binding is estimated. In our approach the final importance of the energy terms is not an absolute value but a relative weight. This because the limits of the theoric framework is too strong and empiric correction are needed to correlate the computational evaluation to the experimental data. For such reason the pharmacophore is adapted to the single scaffold in study and used to select a training set coherent with the binding mode hypothesized. Anyway the final quantitative results are in agreement with the intrinsic characteristics of CK2 active pocket and with the inhibitors peculiarities.
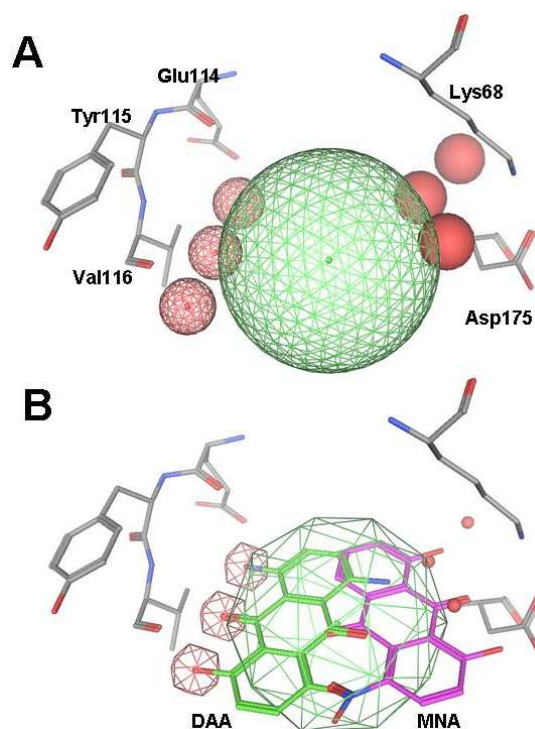
Figure 1.11: CK2 inhibitors pharmacophore. A) Hydrophobic/aromatic region is shown as a green sphere, hydrogen bonding capable group are shown in red and conserved waters molecules in CPK representation. B) Superimposition of the pharmacophore to the two inhibitors DAA and MNA. MNA can successfully substitute two waters molecules with polar hydrogen bonding capable moieties, on the contrary DAA can interact with the hinge region and both share similar hydrophobic interactions with CK2 ATP-binding pocket.

1.4 Imatinib

1.4.1 The First Kinase Inhibitor Drug

Imatinib (4-(4-Methyl-piperazin-1-ylmethyl)-N-[4-methyl-3-(4-pyridin-3-yl-pyrimidin-2-ylamino)-phenyl]-benzamide) is the first drug targeting a protein kinase entered in clinical use. It is also called STI-571 (Signal Transduction Inhibitor-571), CGP 57148, Gleevec (in the United States), and Glivec (in Europe). It has been developed thanks to a time-consuming process of testing a large number of compounds for inhibition of protein kinase C [104, 105]. The starting phenylaminopyrimidine scaffold was modified to enhance selectivity, solubility and bioavailability. The final molecule, Imatinib (Fig 1.13), was selective for two
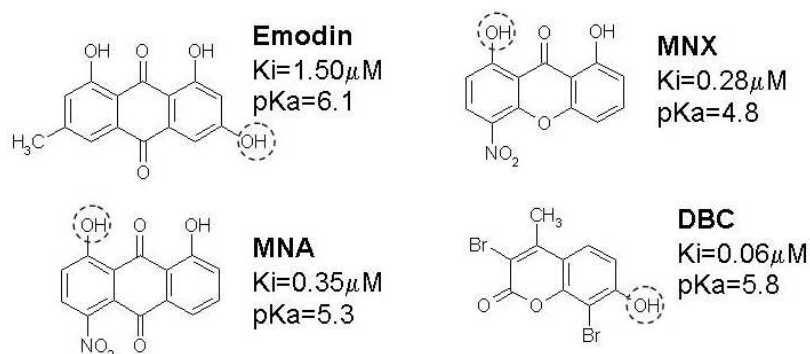
Figure 1.12: pKa of some CK2 inhibitors. The acidity of groups inside the circle is important for the final binding mode, allowing the substitution of active site waters. The final activity is the results of this capability and hydrophobic and van der Waals interactions with the several apolar aminoacid side chains of the pocket.

protein kinase receptors, the platelet-derived growth factor receptor (PDGFR) and the stem-cell factor/mast-cell factor receptor (Kit), and a non-receptor TK, the Abelson protein kinase (Abl).
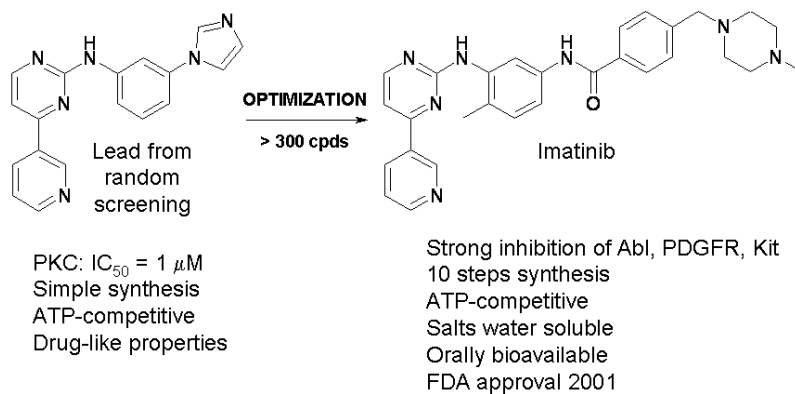
Figure 1.13: Imatinib development.

X-ray crystallography data [106,107] and kinetics studies [108] highlight the ATP-competitive inhibition of Imatinib on Abl. It binds the cleft between the two lobes and penetrates further than ATP in the hydrophobic core of the kinase. Moreover it forces the activation loop into an inactive conformation, in such a way that it mimics the binding of protein substrates like that of the inactive insulin-receptor protein-tyrosine kinase [109]. In this conformation the

Phe401 instead of the catalytic aspartate residue of DFG motif, is flipped and points toward the ATPbinding site.

The Imatinib binding mode to c-Kit is similar, even if the inactive conformation of the protein is not induced by the inhibitor as in Abl. It replaces the normal molecular mechanism that has evolved to keep c-Kit in its off state partially modifying it: it induces a rearrangement of the DFG motif caused by a steric clash with the Phe810 side chain [110].

The introduction of Imatinib was the first answer of medicinal chemistry efforts to the pressing need for the development of novel agents to treat cancer. With its remarkable clinical success it demonstrated the efficacy of this therapeutical approach. Now thanks to a better understanding of structural biology and new computational approaches we can be optimistic regarding the possibility of innovative kinase inhibitors for the treatment of tumors.

1.4.2 Therapeutic Applications

The first therapeutic application of Imatinib was chronic myeloid leukaemia (CML), responsible of about the 20% of cases of adults leukemia in Western populations. This pathology is caused by the reciprocal translocation between chromosomes 9 and 22 resulting in the Philadelphia chromosome encoding for the chimeric protein Bcr-Abl (Fig 1.14-A). The consequence is a unregulated proliferation of myeloid cells in the bone marrow. Additionally this oncogenic protein is present in the 25% of patients with B-cell acute lymphoblastic leukemia [111, 112]. Even if Imatinib induces a complete cytogenetic response in more than 80% of patients with CML in chronic phase, the patients majority still harbour molecular residual disease and some develop resistance associated with Abl mutations [113].

Once that the mechanism of action of Imatinib and the low toxicity, evaluated in clinical trials, were demonstrated other several studies started to understand if different clinical application of the inhibitor were possible. In particular all pathologies linked to a constitutively activation of Abl, Kit and PDGFR were analyzed. In particular either KIT (75-80%) or PDGFR (5-10%) possess activating mutations in gastrointestinal stromal tumours (GISTs). These rapresent the most common mesenchymal neoplasm of the gastrointestinal tract (Fig 1.14-B) and are highly resistant to conventional chemotherapy and radiotherapy [116, 117]. Imatinib efficacy on these pathologies results in a recent approval in the USA and EU for the treatment of unresectable and metastatic GISTs. However, even in this case, even if Imatinib has greatly improved the quality of life and survival of patients with advanced gastrointestinal stromal tu-
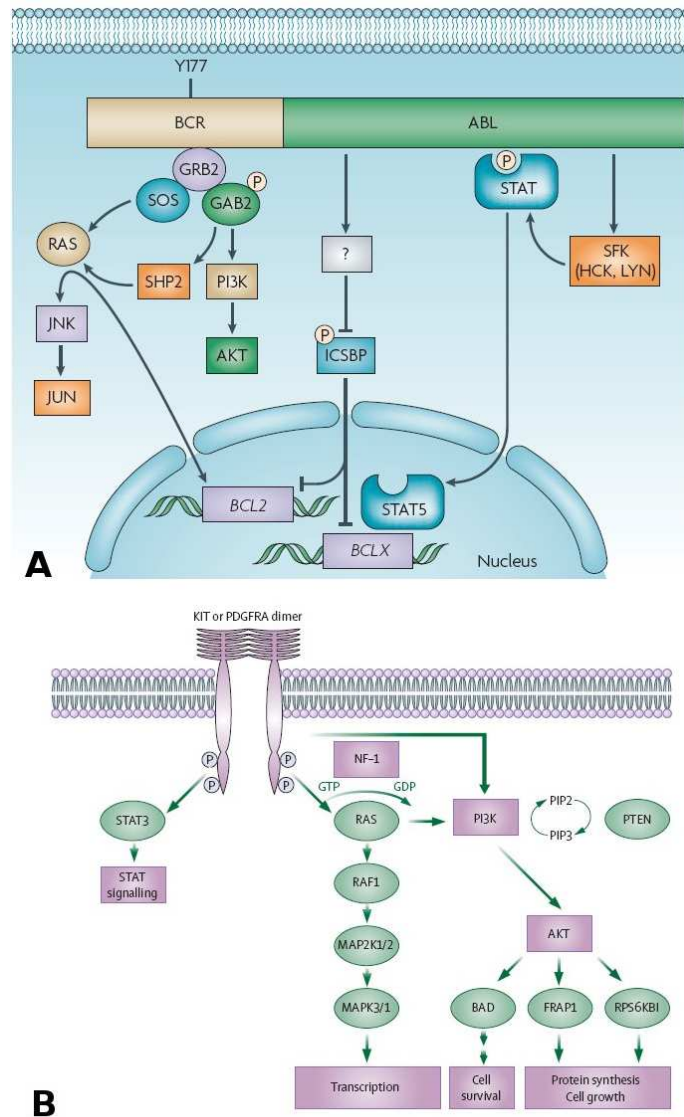
Figure 1.14: Bcr-Abl, Kit and PDGFR cancer signalling. A- Some of the more important oncogenic signalling of Bcr-Abl taken from [114]. Bcr-Abl constitutively activation leads to an increase expression of BCL2 and BCLX and inhibition if ICSBP transcription causing cell proliferation and survival through activation of RAS, SHP2 and PI3K-AKT pathways. B- Predominant oncogenic pathways in GISTs [115]. Mutations on Kit and PDGFR lead to ligand-independent activation finally linked to cell proliferation and inhibition of apoptosis principally thanks to the PI3K and STAT3 pathways.

mours, most patients are not cured cause the development of Imatinib resistance due to the emergence or expansion of individual clones [118].

Finally recent reports show a beneficial effect of Imatinib versus idiopathic hypereosinophilic syndrome (HES) and chronic eosinophilic leukemia (CEL). They comprise a spectrum of indolent to aggressive diseases characterized by unexplained, persistent hypereosinophilia. The complete hematologic remissions observed is rapid and present at lower Imatinib doses than used in (CML) [119]. The action is probably due to the inhibition of FIP1L1-PDGFRA fusion protein, a PK derived from an interstitial deletion, that fuses PDGFR gene to an uncharacterized human gene FIP1-like-1 (FIP1L1). Interesting, approximately 40% of responding patients lack the FIP1L1-PDGFRA fusion, suggesting an activity versus other PKs [120].

1.4.3 CGP-582: Improving Imatinib Selectivity

CGP-582 is a Imatinib analogue designed in the group of Prof. Scapozza. The only difference from the precursor is the substitution of the benzene ring near the piperazin with a pyridin ring. This small modification results in an activity loss of about 2 fold versus Abl in the enzymatic assay, enough to achieve a selectivity versus PDGFR and Kit on a cellular level (Fig 1.15).
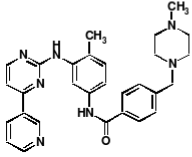
	Abl (IC$_{50}$ μM)		
	Kinase Assay	Proliferation	
STI-571	0.037 ± 0.010	0.400 ± 0.070	Selective Inhibition: Abl, PDGFR, Kit
CGP-582	2.200 ± 0.800	>10	PDGFR, Kit

Figure 1.15: CGP-582

This result is particular interesting from a medicinal chemistry point of view

addressing the common issue of specificity in drug design. Usually specific inhibitors of protein kinases take advantage of limited sequence variation surrounding the ATPbinding site. However in this case is possible to speculate, on the basis of the structural informations available on Imatinib and Abl, that CGP-582 does not lose any direct interaction with the Abl active pocket.

For this reason the first question that is addressed in this study is which are the chemical-physical basis of this difference of activity and on other hand try to find out a computational technique with the sensibility to answer to this very difficult issue. In theory then, this in silico approach would be able to help the computational chemist to design a selective inhibitor, modulating its activity without modifying clearly important and crucial interactions with the target.

Chapter 2

LIE Method Applications to Study CK2 Inhibitors

LIE methodology was developed by Åqvist as a plausible compromise between accuracy and computational speed in determining binding free energy ($\Delta\mathrm{G}_{bind}$) values [121, 122]. Different authors have already described the applicability and the accuracy of this approach in which free energy predictions have been estimated with an error on the order of 1 kcal/mol [123–125]. Application of LIE methodology in docking/scoring applications have been also reported [126, 127]. An important key-point in using LIE methodology is the accessibility of a reasonable pose of the ligand inside the binding cavity and, if available, it is usually derived from the crystal structure of ligand-protein complex. As alternative, a reasonable ligand starting pose can be obtained by molecular docking simulations, but in this case a special attention must be aware into the selection of the most representative binding conformation [128].

2.1 Linear Interaction Energy Method (LIE)

The LIE method is based on the assumption that the inhibitor free energy of binding to a macromolecule is linearly correlated to several energy terms that can be calculated using a molecular mechanic force field [121]. The $\Delta\mathrm{G}_{bind}$ is derived using the following formula:

$$\Delta G_{bind} = \alpha(< U_{vdw}^b > - < U_{vdw}^f >) + \beta(< U_{elec}^b > - < U_{elec}^f >) + \\ + \gamma(< U_{cav}^b > - < U_{cav}^f >) \quad (2.1)$$

Two different simulations must be performed: one describes the molecule free in solution (U^f) and the other considers the ligand-protein complex (U^b). This is necessary to approximate the binding event as replacement of the interaction of the inhibitor with water molecules, with the interaction in a mixed aqueous-protein environment. The energy values collected estimate the van der Waals interaction (U_{vdw}), the electrostatic contributions (U_{elec}) and the cavity parameter (U_{cav}). This last one, introduced by Carlson and Jorgensen [129], is linked to the energy penalty for forming a solute cavity. Indeed, this term is proportional to the solvent accessible area and, in addition, it contains a component corresponding to the van der Waals solute-solvent interactions, not zero even for buried atoms up to a certain distance from the solute-exposed surface [130]. The brackets indicate that the ensemble average of the energy terms is taken in account during the simulation. Using a training set of molecules with known activity, a semi-empirical energy model is built by fitting the three different parameter coefficients (α, β and γ) to the free energy of binding. This empirical element can be useful to compensate part of the limits of the method, due, for example, to the force field or to the difficult to evaluate the entropic cost in the intra-molecular changes of receptors and ligands during the complex formation. In this study the solvent has been treated using the surface-generalized Born (SGB) continuum solvation model, approach that has been validate in several studies [125, 131, 132] and reduces of ten times the computational effort. In this method the electrostatic energy is due to a Coulomb (U_{coul}) term and the SGB-solvent reaction energy (U_{rxnf}):

$$U_{elec} = U_{coul} + 2U_{rxnf} \tag{2.2}$$

Molecular dynamic sampling or Monte Carlo simulation can be used to calculate the energy terms, however if the initial docked position of the ligands is known a simple minimization protocol can provide comparable results with small degradation of accuracy. A common approach consists in using the crystal structure of an available protein-inhibitor complex as starting pose for the fast minimization protocol of other molecules with comparable scaffold. LIE method seems to not permit the parameter transferability [125] even if the enzymes belong to the same class, nevertheless more investigations are needed.

2.2   Benzimidazole Scaffold

The availability of several crystal structures of different TBB-analogues in concert with a large number of activity data for this class of CK2 inhibitors, sug-

gested us the possibility to explore the structure-activity relationship (SAR) of this scaffold also from a quantitative point of view using the linear interaction energy (LIE) method. This approach allows us to build up a binding free energy prediction model of this TBB-like class of CK2 inhibitors.

In this study, we have developed a binding free energy prediction model based on all known TBB-like inhibitors by using LIE method in tandem with the Surface Generalized Born (SGB) approximation for the treatment of the electrostatic solvatation energy [131]. The CK2-inhibition activity of 45 TBB-like together with the availability of four different crystal structures has been provided a robust base to build up and evaluate the LIE method performances in both rationalization of the observed binding motifs and in the prediction of the corresponding binding free energies [71, 72, 84].

2.2.1 Methods

Prepartion of CK2

Coordinates of the human CK2 were downloaded from the RSCB Protein Data Bank [99] (PDB code 1JWH [133]). Only the A chain was used and the other three monomers were removed together with all water molecules, ions and the ligand AMP. In the protein were included two water molecules present in the crystal structure of TBB in complex with CK2 and considerate important for the binding of inhibitors. TBB (PDB entry 1J91 [73]) and these two water molecules (Water 1, Water 2) were superimposed to the kinase domain. Using the Protein Preparation [134] utility (Schrödinger Inc.), the residues that were beyond 8-12 Å from TBB were neutralized, possible conflicts in the hydrogen bonding were corrected and the two alternate tautomers of histidine side chain were considered, especially if near the binding pocket. Using the OPLS force field the hydrogenated CK2-TBB complex was energy minimized until the average RMS deviation of the non-hydrogen atoms reached 0.3 Å. This protocol was exploited for both the two tautomeric states of TBB, that influences the orientation of the hydrogens of water molecules.

Preparation of Inhibitors

All inhibitors were built using LigPrep [135] utility in Maestro (Schrödinger Inc.), and all species existing at pH= 7±1, included tautomers and enantiomers were generated. Using MMFFs force field in vacuum a short conformational search was performed to relax the structure. As starting pose of the LIE calculation the molecules were superimposed to TBB, K37A or K37B, since all these

ligands have similar scaffold is likely that they share a common binding mode. For every inhibitor the choice of its initial position, enantiomer or ionic form was made in agree with chemical-physical considerations on the LIE simulation final pose and with the statistical validation of the LIE final model.

LIE Calculation

The LIE energy model was created using the Liaison package [134] implemented by Schrdinger Inc. All the calculations were carried out using OPLS2003 force field and the implicit solvent was treated using the surface generalized Born (SGB) model. A Truncated Newton minimization was performed for free ligands with a residue-based cutoff distance of 15 Å, a 0.01 root mean square (rms) gradient for convergence and a maximum of 500 steps. For the complex CK2-inhibitor has been used the same protocol, but with a rms gradient for convergence of 0.05 and all the protein residue beyond 10 Å from TBB have been frozen. The energy terms calculated with this protocol were used to build the binding free energy models.

2.2.2   CK2 Crystal Structure

The crystal structures of the Zea mais CK2 in complex with tetrabromo-benzo-diazole derivates were solved for TBB [73] (PDB code 1J91), K25 (PDB code 1ZOE), K44 (PDB code 1ZOH) and K37 (PDB code 1ZOG) [74]. The last one presents two different ligand poses called K37A and K37B (as shown in Figure 2.1). The inhibitor activity assays have been performed using rat CK2 that share a perfectly identical kinase domain with the human one, and presents an identity of 76.1% with Zea mais CK2. Indeed, in a range of 6 Å (using TBB as a reference ligand) of the ATP-binding site, only four amino acids are not identical between human and maize CK2: Leu45Val and Val66Ile (positioned on the P-loop), Ile95Val, and His115Tyr (located in the hinge region). However, only Val66Ile position is closed to the ligand-binding region. Following this structure similarity, in this computational study we selected the crystal of human CK2 in complex with an ATP analogue (PDB entry 1JWH) [133] as reference structure.

Another important issue considered in the present study was the induced fit by ligand to the residue His160; according to the available crystal structures His160 can assume two different conformers depending on the nature of the inhibitor bounded. According to the B-factors, both conformers should be plausible. In the present study, we have approximated two distinct binding modes: the first one similar to TBB (TBB-like binding mode) with histidine in a close position and a second one (K-like binding mode) with histidine in a open

conformation. In particular, the crystal structures of the maize CK2, His160 is its open-status leaving the active site completely accessible to K-like analogues.
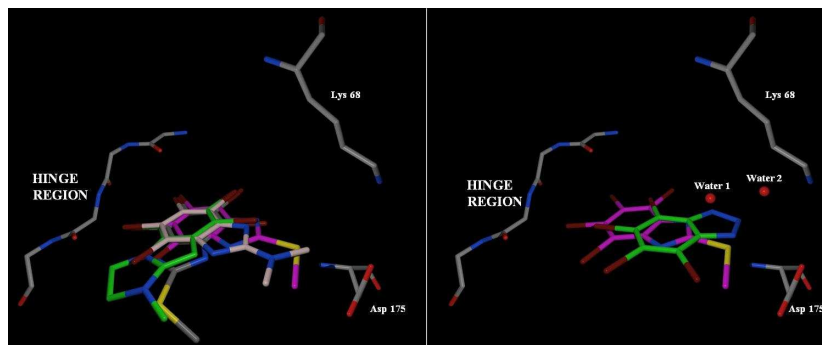
Figure 2.1: CK2 ATP binding site from X-ray diffraction crystal structure (PDB entry 1JWH) [133]. The residue Lys68 and the DFG motif amino acid Asp175 are shown, together with the Hinge region (the side chains are hidden). On the left, superimposition of K25 (white, PDB entry 1ZOE), K44 (green, PDB entry 1ZOH), K37 A and B (respectively grey and magenta, PDB entry 1ZOG) [74] point out that the hydrophobic interactions are almost the same. The molecules rotate around the benzo moiety and three out of the four bromides are in a conserved position. On the right, TBB (green, PDB entry 1J91) [73] is superimposed to K37 in conformation B (magenta, PDB entry 1ZOG) [74]. For this computational study was used CK2 from the PDB 1JWH with the two water molecules from PDB 1J91 (shown on the right). While the second water molecule (Water 2) is preserved in all these crystals, the first one (Water 1) can be substituted by a chloride or sodium ion and its location can change in a way that seems directly correlated to the structure and final inhibitor pose. For the energy model CK2scoreA the starting LIE position of each inhibitor, without crystal structure available, was obtained by analogy to TBB and K37A, whereas CK2scoreB was built on the binding mode of TBB and K37B.

2.2.3   Water Molecules and Inhibitor Binding Modes

Pospisil et al. have recently demonstrated the critical role of explicit water molecules in structure-based drug design applications [136]. In this LIE study, we decided to include two water molecules in the active site of the human CK2, derived from the crystal structure of the protein in complex with TBB (Water 1 and Water 2, PDB code 1J91). Water 2 is structurally conserved in all available crystal structures while Water 1 can be replaced by ions such as chloride (K25; PDB code:1ZOE) or sodium (K44; PDB code:1ZOH). Theoretically, each TBB-like inhibitors might have the possibility of interact through hydrogen bonding
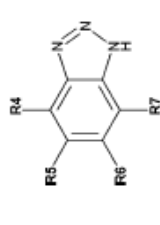
with Water 1 (Fig. 2.1). Moreover, we have considered all possible tautomeric forms of the imidazole moiety to correctly evaluate which tautomer might preferably bound the water molecule. Interestingly, the minimization protocol applied to K37B-CK2 complex was able to reproduce the crystallographic pose of Water 1 found in the corresponding crystal structure (H_2O1227; PDB code 1ZOG). In this LIE study, for all other inhibitors not yet crystallized in complex with CK2, we have considerate three possible starting poses in the minimization protocol: TBB-like, K37A-like, or K37B-like binding mode (Fig. 2.1).

2.2.4   Training Set
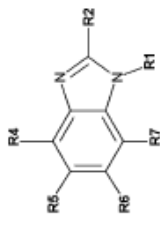
As summarized in Figure 2.2, all benzoimidazole derivates (25 compounds) with a known inhibitory constant (Ki) value were used as training set to build up a LIE model. From a structural point of view, the main difference among them is the nature of the substituent at position 2.

As already anticipated, the possible different starting pose of ligands (TBB-like, K37A-like or K37B-like), the possible tautomeric forms and, in some case, the different ionic state have been taken into account during the construction of all LIE models.

Two different criteria were applied to evaluate which combinations of effects better correlated with experimental Ki values. The first one is based on the quality of the final pose resulted after the minimization protocol. In particular, since the conformational search with this method is extremely limited, it is important that each starting pose is reasonable close to a local energy minimum. An appropriate visual inspection in tandem with a structural and a potential energy analysis have been performed to properly evaluate the quality of each final pose. The second criterion is based on the statistical validation and cross-validation of the LIE energy model. Following the abovementioned strategy, we have developed a stepwise workflow as shown in Figure 2.3.

inhibitor	R1	R2	R4	R5	R6	R7	K_i (μM)	ΔG_{exp} (kcal/mol)
1 TBB			-Br	-Br	-Br	-Br	0.60	-8.54
2 diclo-2azaB			-H	-Cl	-Cl	-H	10.00	-6.86
3 tetraclo-2azaB			-Cl	-Cl	-Cl	-Cl	5.00	-7.28
4 K25	-H	-N(CH2)2	-Br	-Br	-Br	-Br	0.04	-10.16
5 K37	-H	-SCH3	-Br	-Br	-Br	-Br	0.07	-9.82
6 DiClBz	-H	-H	-H	-Cl	-Cl	-H	35.00	-6.12
7 TCBz	-H	-H	-Cl	-Cl	-Cl	-Cl	21.00	-6.42
8 TBI	-H	-Br	-Br	-Br	-Br	-Br	0.70	-8.45
9	-H	-NH2	-Br	-Br	-Br	-Br	0.23	-9.11
10	-H	-NHCH3	-Br	-Br	-Br	-Br	0.09	-9.67
11	-H	-NH(CH2)2OH	-Br	-Br	-Br	-Br	0.09	-9.67
12	-H	-NHCH(CH3)2	-Br	-Br	-Br	-Br	0.13	-9.45
13	-H	-NH(CH2)2OCH3	-Br	-Br	-Br	-Br	0.06	-9.91
14	-H	--Br	-Br	-Br	-Br	-Br	0.14	-9.41
15	-CH3	--Br	-Br	-Br	-Br	-Br	0.54	-8.60
16	-CH2CH=CH2	--Br	-Br	-Br	-Br	-Br	0.63	-8.51
17	-CH3	-OH (=O)^a	-Br	-Br	-Br	-Br	0.20	-9.20
18	-CH3	-N(CH3)2	-Br	-Br	-Br	-Br	0.20	-9.20
19	-CH3	-NHCH(CH3)2	-Br	-Br	-Br	-Br	0.19	-9.23
20	-CH3	-SCH3	-Br	-Br	-Br	-Br	0.36	-8.84
21	-H	-SCH2CH(OH)CH2OH	-Br	-Br	-Br	-Br	0.14	-9.41
22	-H	-OH (=O)^a	-Br	-Br	-Br	-Br	0.15	-9.38
23	-H	-NH(CH2)2N(CH3)2	-Br	-Br	-Br	-Br	0.16	-9.33
24	-CH2CONH2	--Br	-Br	-Br	-Br	-Br	1.90	-7.85
25	-H	-SCH2COOH	-Br	-Br	-Br	-Br	0.18	-9.26

[a] Both keto- and enol-tautomers were investigated in two different simulations. The keto-tautomer was finally chosen to build both CK2scoreA and CK2scoreB models.

Figure 2.2: CK2 inhibitors used as training set to create CK2scoreA and CK2scoreB.
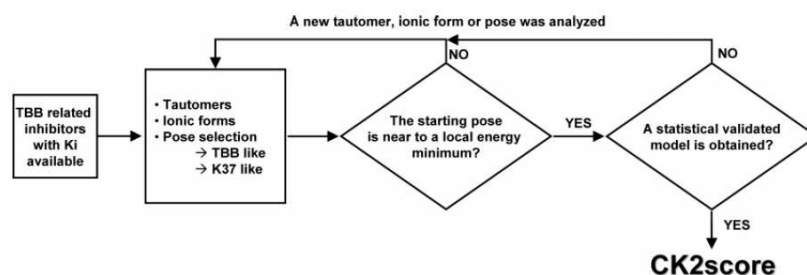
Figure 2.3: Workflow used to select the inhibitor starting pose and, tautomeric and ionic form. An initial model was built using a small set of molecules composed by ligands crystallized in complex with CK2 and their most structurally related compounds. Starting from the more reasonable pose, tautomeric forms and ionic state hypothesis, a first model was built. Its validation was both chemical-physical, checking if the energy minimization protocol found a near local energy minimum and a plausible pose (in a way similar to the docking visual inspection), and statistical, using the correlation factor and cross validated q^2. After that the model was progressively implemented with the remaining molecules, with special attention to their tautomeric and ionic hypothesis, their starting and final pose and the energy terms calculated during the minimization (van der Waals, Electrostatic, Reaction Field and Cavity energy terms). With this approach if some inhibitors were not linearly correlated to the others, they were considerate outlier and excluded from the final model, denominated CK2score. This deep molecular interaction study is useful to try to understand the meaning of these outliers.

A first LIE model was created using a small subset of the entire training set characterized by the highest structural similarity to TBB, K25, and K37. The starting poses for TBB, K25 and K37 were simply collected from the corresponding crystal structures, whereas for all analogues we have used TBB and K37 poses as a template for their structural superimpositions into the CK2 binding cavity. In this way, possible errors due to erroneous starting poses were minimized. After both chemical-physical and the statistical validations, this first model has been progressively implemented with including one by one all the remaining structures until the creation of the final LIE model. Since the crystal structure of the CK2-K37 complex highlights two possible binding modes (A and B) of K37 inhibitor, we investigated and compared both starting poses as possible templates for the superimposition of all other TBB-like derivatives. Consequently, two different energy models were ultimately obtained: "CK2scoreA" created by using as starting poses those from TBB and K37A, and "CK2scoreB" based on TBB and K37B conformations. This stepwise protocol is very useful to

create simple "binding mode rules" able to generate robust binding free energy models and helpful to understand which chemical features can relevantly determine the final binding pose (Fig. 2.4). For example, using this approach, we were able to select the carbonyl form instead of the corresponding enol-tautomer for both **17** and **22** derivatives comparing the quality of the corresponding pKi prediction with respect to the experimental data (see Table 2.5). Analogously, we found that K25 is probably unprotonated in its bounded state, on the contrary of what Pagano et al previously suggested [72].
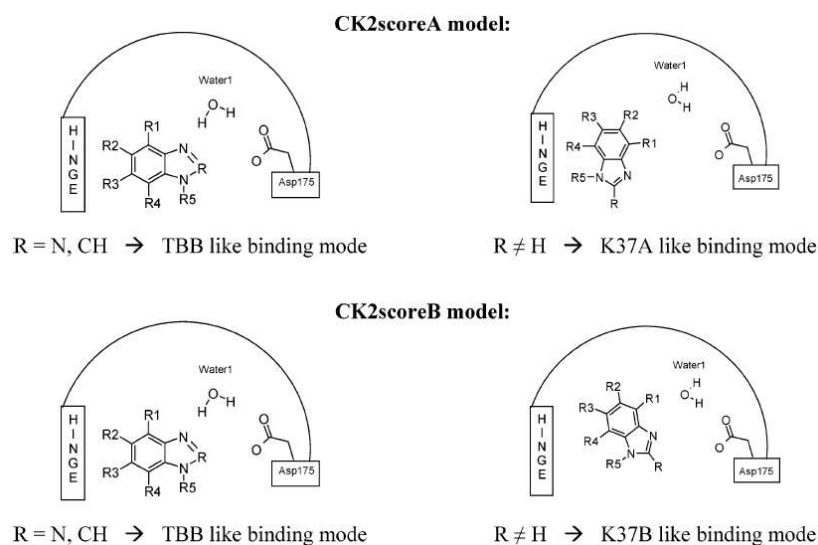
Figure 2.4: Binding modes rules derived from the best models CK2scoreA and CK2scoreB. The starting pose was chosen according to the moiety in position 2. Both the tautomeric forms were investigated during the model creation. The final choice shown is the one resulting in the energy model with the best r^2 and q_{cv}^2. For CK2scoreA the possible binding modes are TBB-like and K37A-like, while for CK2scoreB are TBB-like and K37B-like.

The calculated energy terms, collected in 2.6, underline that the most crucial binding interactions in this TBB-like class of inhibitors are due to hydrophobic and van der Waals contributions and, in particular, emphasize the importance played by the solvent accessible area, which is linked to the energy penalty for forming a solute cavity (U_{cav}). Moreover, the analysis of the binding energy terms indicate that the presence of a poly-bromide moiety guarantees the necessary driving force for the ligand binding into the CK2 active site. An additional hydrogen bonding with water (Water 1), or further van der Waals interactions of the substituent at position 2, provide the final binding mode of each ligands.

Both CK2scoreA and CK2scoreB models, indicate molecules **23**, **24** and **25** as possible outliers. In fact, the predicted energy values are considerably different to the experimental data. A possible explanation might be found considering some chemical peculiarities of these three TBB-like derivatives. Considering CK2scoreB model, ligand **23** contains a protonable tertiary amine in the side chain at 2-position that, apparently, does not interact with any CK2 residues in the bounded state. Therefore, the equilibrium is moved to the direction of the free molecule in solution rather than the bound state and an evaluation with CK2score gives a sub-estimated free energy of about 10 kcal/mol. Considering derivative **25** in its K37B-like pose, the carboxyl group interacts so strongly with the hydroxyl moiety of Ser51, located in the P-loop domain, that LIE model overestimated the free binding energy of 7 kcal/mol. The well known mobility of this region of the protein kinase does not allow us to consistently speculate on this tight interaction, even if in principle achievable. The modeling study of inhibitor **24** presents the same type of misinterpretation as argued for **25**: in this case, the amide group in the side chain at 2-position can strongly interact with the backbone of His160. As already anticipated, the conformational mobility of this specific residue, again, does not allow us to appropriately quantify this interaction overestimating the corresponding binding free energy of about 2.5 kcal/mol. Similarly, also CK2scoreA model overestimates both **24** and **25** inhibitors due to the orientation of the side chain at 2-position towards a region of the binding cavity more exposed to the solvent and consequently less easier to precisely quantify the solvation energy term of these poses . For these reasons, we have considered molecules **23**, **24** and **25** as outliers and, consequently, they were not included in the training set used to build up the final energy models. Moreover, also K44 has been not included into the training set, but incorporated to the test set, because its binding mode is more rotated towards the hinge region (Fig. 2.1) respect all other TBB-like inhibitors.

Both final LIE energy models, CK2scoreA and CK2scoreB, show acceptable statistics even if CK2scoreA model presents slightly better performance, as summarized in Figure 2.7. Moreover, in Table 2.5 are collected the experimental and the predicted binding free energies (ΔG_{bind}) for both models. CK2scoreA energy terms resulted in: α=0.173±0.003, β=0.081±0.005 and γ=-0.772±0.027. It is very clear that for this class of CK2 inhibitors, the energy penalty coming from the formation of a solute cavity (γ coefficient), seems to play a crucial role in determining ligand binding. Probably this is one of the most important driving forces in determining the ligand movement from outside to the binding cavity, where stabilizing van der Waals interactions are suddenly established

(α coefficient). The complementary between the tetrabromo-benzo moiety and the protein active site is almost perfect, and its importance is highlighted by the significant loss of activity caused by the substitution of one of these bromide with chloride atoms. The final minimized posed shows that, in several molecules, the substituent in 2 can contribute with additional van der Waals interactions. Similar consideration can be made for CK2scoreB, where energy terms resulted to be: α=0.191±0.003, β=0.108±0.007 and γ=-0.460±0.025. In this second model the cavity contribution is still the crucial contributor to binding free energy even if less than for CK2scoreA. This result is in agreement with the consideration that electrostatic interactions with Water 1 or with Asp175 are not accessible for the K37A-like binding mode, while are often established in the K37B-like. Moreover, CK2score energy terms analysis is also compatible with the alogen-driven binding hypothesis proposed by Battistutta et al. [74].

inhibitor	ΔG_{exp} (kcal/mol)	$\Delta G_{\text{CK2scoreA}}$ (kcal/mol)	$\Delta G_{\text{CV\_A}}$ (kcal/mol)	$\Delta G_{\text{CK2scoreB}}$ (kcal/mol)	$\Delta G_{\text{CV\_B}}$ (kcal/mol)
1	−8.54	−8.93	−9.07	−8.61	−8.62
2	−6.86	−5.93	−5.79	−5.85	−5.70
3	−7.28	−6.96	−6.94	−6.82	−6.80
4	−10.16	−9.93	−9.90	−9.14	−9.04
5	−9.82	−9.03	−8.95	−8.83	−8.71
6	−6.12	−4.91	−4.80	−5.04	−4.90
7	−6.42	−6.44	−6.44	−6.64	−6.69
8	−8.45	−8.07	−8.00	−7.97	−7.89
9	−9.11	−8.68	−8.65	−8.38	−8.23
10	−9.67	−7.99	−7.76	−8.59	−8.52
11	−9.67	−9.29	−9.24	−9.12	−9.07
12	−9.45	−9.72	−9.74	−9.54	−9.55
13	−9.91	−9.93	−9.93	−10.08	−10.10
14	−9.41	−9.91	−9.99	−10.52	−10.63
15	−8.60	−8.66	−8.68	−8.61	−8.61
16	−8.51	−9.11	−9.31	−9.31	−9.50
17	−9.20	−9.67	−9.75	−8.90	−8.86
18	−9.20	−10.07	−10.14	−9.82	−9.87
19	−9.23	−9.73	−9.77	−10.39	−10.48
20	−8.84	−9.02	−9.06	−9.28	−9.36
21	−9.41	−10.22	−10.52	−10.27	−10.46
22	−9.38	−9.53	−9.57	−9.84	−10.14

Figure 2.5: Comparison of experimental (ΔG_{exp}, kcal/mol), predicted ($\Delta G_{CK2scoreA}$ and $\Delta G_{CK2scoreB}$, kcal/mol) and jackknife cross validated (ΔG_{CV_A} and ΔG_{CV_B}, kcal/mol) free energies of binding between the two energy models CK2scoreA and CK2scoreB for the training set.

inhibitor		$U^f_{coul}{}^b$ (kcal/mol)	$U^f_{vdw}{}^b$ (kcal/mol)	U^f_{nonf} (kcal/mol)	U^f_{cav} (kcal/mol)	U^b_{coul} (kcal/mol)	U^b_{vdw} (kcal/mol)	U^b_{nonf} (kcal/mol)	U^b_{cav} (kcal/mol)
1	CK2scoreA	0	0	−7.3369	−3.9708	−25.9540	−28.2600	−1.1124	−1.6834
	CK2scoreB	0	0	−7.3369	−3.9708	−25.9540	−28.2600	−1.1124	−1.6834
2	CK2scoreA	0	0	−8.0632	−1.9537	−26.5010	−18.6490	−0.5601	−0.3302
	CK2scoreB	0	0	−8.0632	−1.9537	−26.5010	−18.6490	−0.5601	−0.3302
3	CK2scoreA	0	0	−8.2269	−2.6004	−20.3440	−25.7910	−1.5122	−0.1051
	CK2scoreB	0	0	−8.2269	−2.6004	−20.3440	−25.7910	−1.5122	−0.1051
4	CK2scoreA	0	0	−7.8606	−4.5924	−8.8304	−36.4620	−3.6214	0.0487
	CK2scoreB	0	0	−7.8606	−4.5924	−8.8304	−36.4620	−3.6214	0.0487
5	CK2scoreA	0	0	−5.8040	−3.2395	0.6488	−39.5260	−3.8976	0.0560
	CK2scoreB	0	0	−8.2309	−3.5706	−5.4229	−37.4250	−5.5969	0.0276
6	CK2scoreA	0	0	−10.6060	−1.1149	−27.1810	−17.3850	−1.9863	0.2891
	CK2scoreB	0	0	−10.6060	−1.1149	−27.1810	−17.3850	−1.9863	0.2891
7	CK2scoreA	0	0	−9.9687	−1.8555	−27.0050	−23.9830	−2.3543	−0.1418
	CK2scoreB	0	0	−9.9687	−1.8555	−27.0050	−23.9830	−2.3543	−0.1418
8	CK2scoreA	0	0	−7.6129	−3.0885	−24.4980	−27.5350	−1.6406	−0.1457
	CK2scoreB	0	0	−7.6129	−3.0885	−24.4980	−27.5350	−1.6406	−0.1457
9	CK2scoreA	0	0	−6.0590	−3.8315	−1.3455	−38.3150	−4.0858	−0.6597
	CK2scoreB	0	0	−6.1283	−4.1155	−1.8397	−37.0610	−3.3133	−0.4056
10	CK2scoreA	0	0	−11.0160	−4.9734	0.0713	−36.2320	−6.5187	−6.5187
	CK2scoreB	0	0	−11.1650	−5.0441	−14.3230	−33.1000	−6.3554	−1.2314
11	CK2scoreA	0	0	−8.4315	−5.2075	0.7733	−37.7250	−5.2796	−0.9014
	CK2scoreB	0	0	−8.7779	−5.2547	−12.2000	−34.1390	−4.1418	−0.2877
12	CK2scoreA	0	0	−13.6680	−5.1610	−12.2250	−37.5540	−7.3914	−0.9671
	CK2scoreB	0	0	−13.6070	−5.1384	−16.6110	−37.1250	−5.9417	−0.1172
13	CK2scoreA	0	0	−7.7826	−5.0738	0.2795	−40.2860	−4.9879	−0.6444
	CK2scoreB	0	0	−7.8203	−5.0567	−11.9860	−37.8760	−3.4786	0.3434
14	CK2scoreA	0	0	−9.6616	−5.3114	−0.0291	−40.6240	−5.3420	−0.6836
	CK2scoreB	0	0	−9.4952	−5.3267	−15.2920	−39.2910	−4.1977	0.1101

15	CK2scoreA	0	−2.6495	−2.2237	−1.2393	−39.8380	−1.8082	0.0983
	CK2scoreB	0	−2.8130	−2.5318	−2.3673	−38.0970	−1.6814	0.3324
16	CK2scoreA	0	−2.6057	−2.0137	−1.9885	−42.2550	−1.8014	2.6127
	CK2scoreB	0	−3.0168	−2.3688	−3.8413	−40.8160	−1.6899	0.6311
17	CK2scoreA	0	−10.3210	−6.0790	−6.1568	−36.6910	−4.7716	−1.2720
	CK2scoreB	0	−10.9300	−6.1630	−12.9160	−33.890	−4.7848	−1.0334
18	CK2scoreA	0	−4.5568	−3.2973	0.2302	−41.5190	−2.7188	0.8368
	CK2scoreB	0	−4.5959	−3.1823	−7.2072	−39.3360	−2.4997	1.1179
19	CK2scoreA	0	−4.7107	−3.5682	−1.1517	−40.8110	−3.4040	0.0289
	CK2scoreB	0	−5.1187	−3.6615	−13.6370	−38.6240	−2.337	0.9863
20	CK2scoreA	0	−4.2830	−1.8848	−1.5686	−41.3510	−2.9071	0.6429
	CK2scoreB	0	−4.7006	−2.2029	−7.4554	−39.8620	−2.1566	0.8545
21	CK2scoreA	0	−19.9740	−3.5731	−29.9950	−38.1490	−13.4830	−6.9798
	CK2scoreB	0	−21.0080	−3.8815	−41.3180	−38.9110	−6.4401	−0.5889
22	CK2scoreA	0	−11.5860	−6.6001	−5.4465	−36.0510	−5.1827	−1.5775
	CK2scoreB	0	−12.1450	−6.9093	−27.1050	−31.8910	−1.8854	−0.3021
23	CK2scoreA	0	−73.5960	−3.5507	54.8200	−42.1260	−3.7417	0.2921
	CK2scoreB	0	−74.3630	−3.3062	38.8130	−41.9110	−39.8020	−0.3901
24	CK2scoreA	0	−8.6384	−5.1098	−24.1580	−34.1820	−3.6828	−0.4103
	CK2scoreB	0	−8.9871	−5.3484	−8.9025	−41.0740	−4.5262	0.1529
25	CK2scoreA	0	−76.8190	−4.2250	−117.1700	−36.8700	−68.7050	−0.6296
	CK2scoreB	0	−77.3850	−4.5163	−122.9700	−40.4070	−47.5200	−0.1659

Figure 2.6: Comparison of the Ensemble Average LIE Energy Terms for the Surface Generalized Born Continuum Solvent Model for the Training Set Free State (f) and Bound State (b) between CK2scoreA and CK2scoreB. U_{coul}, U_{vdw}, U_{rxnf} and U_{cav} are the Coulomb, van der Waals, Reaction Field and Cavity energy terms. [b]The absence of explicit interaction between the ligands and solvent in the free state results in a van der Waals and Coulomb energy equal to zero.
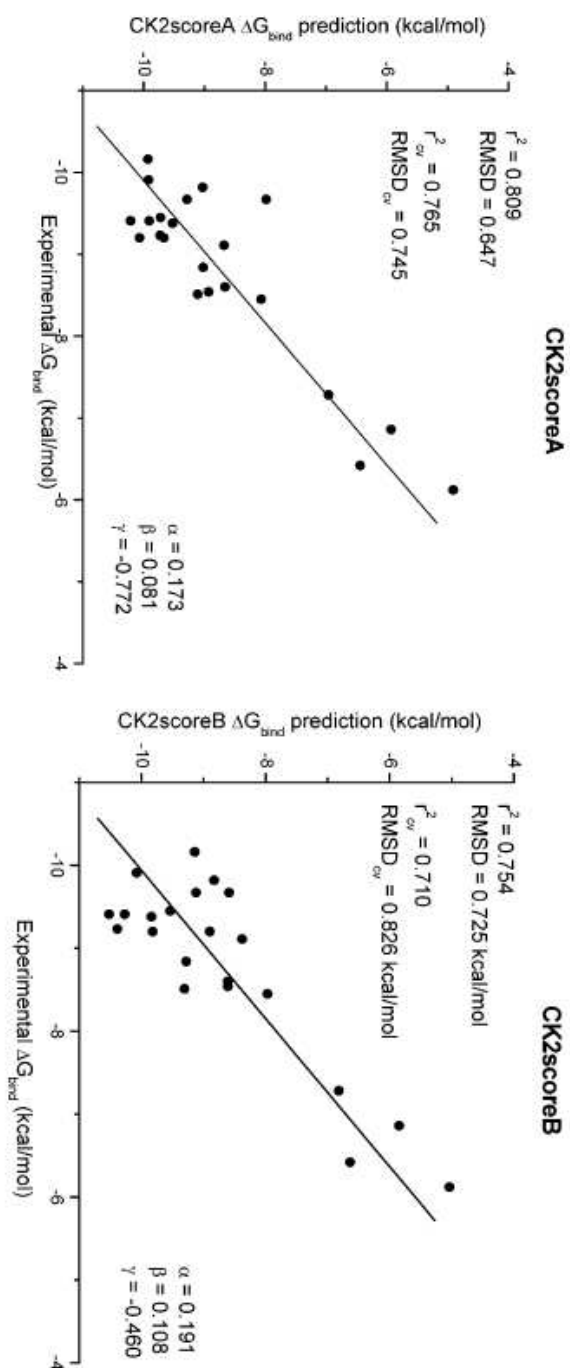
Figure 2.7: Free energy of binding estimated by CK2scoreA and CK2scoreB of each inhibitor used in the training set versus the experimentally measured data. Correlation coefficient (r^2) and cross-validated coefficient (r_{cv}^2) are shown. For the whole set RMSD and cross validate RMSD ($RMSD_{cv}$) were calculated. α, β and γ are the resulting coefficients of the LIE equation for CK2score.

2.2.5 Test Set

CK2score was validated with an external test set of 20 inhibitors with known IC_{50} (Fig. 2.8). In this case we have used IC_{50} instead of Ki considering that it has been precedently demonstrated the correlation between IC_{50} and Ki in the CK2 inhibitory activity assay [72]. For all test set inhibitors, starting binding poses and tautomeric forms have been chosen following the same criteria developed and validated during the LIE model generation.

From a statistical point of view, the predictability of CK2scoreA model is significantly better than CK2scoreB (q^2=0.676 and q^2=0.457, respectively) as summarized in Chart 2. We have identified four possible outliers (**41**, **42**, **43**, and **44**) in generating CK2scoreA model, and only three outliers (**42**, **43**, and **44**) in generation CK2scoreB model (Fig. 2.10, 2.9).

Interestingly both the energy models identified three commons outliers (**43**, **44** and **42**) indicating that this specific behavior is pose-independent. Derivative **43**, overestimated by about 2.5 kcal/mol by both models, contains a nitro-benzo substituent at 2-position not present in any ligands used in the training set. In the final poses, this side chain creates several favorable van der Waals interactions, that result in an overestimation of its binding affinity. This overestimation can be explained considering the loss of rotational degrees of freedom during the complex formation resulting in an important reduction of entropy. This relevant entropic factor is probably diluted into the different energy terms of the LIE model parameters altering the prediction of its real binding free energy. For the second outlier (derivative **44**) in its K37B-like binding mode, a stabilizing hydrogen bonding interaction is formed between the ester moiety of the side chain at 2-position and the exposed hydroxyl group of Ser51. We already described a very similar phenomena considering derivative **25** of the training set. Nevertheless, in this case CK2scoreB overestimates of only about 2.5 kcal/mol instead of the 7 kcal/mol of molecule **25**. This is in agree with the weaker interaction formed by the ester group respect the one formed by the carboxyl moiety of inhibitor **25**.

Moreover, we have used CK2scoreB model to understand what determines a two fold difference of activity transforming the amino group present in **19** in an oxygen as present in molecule **41**. A comparative semi-empirical conformational analysis was performed using the force field MMFF94x, a stochastic conformational search in vacuum was carried out with MOE package [137] to detect starting conformations near the energy minimum. From these initial geometries, we have calculated the heat of formation (ΔH_f) of the equilibrium geometry at ground state with the Spartan program [138]. The semi-empirical

calculations have been performed with AM1 and PM3 with the aqueous solvation energy using the SM 5.4 procedure of Cramer, Truhlar and co-workers [139].

This quantum-mechanic conformational analysis with the aqueous solvation energy, points out that the amino moiety provides two different conformations corresponding to comparable energy minimums. The first conformer, in the final energy minimized docked position, creates two hydrogen bonds: the first with the water molecule (Water 1) and the second with the catalytic aspartic acid Asp175 (Figure 2.11-A). Whereas in the second conformer, the amino group is rotated of about 180 degrees and the steric hindrance of the isopropyl moiety does not allow any hydrogen bonding interactions (Figure 2.11-B). On the contrary, for molecule **41** only this last conformer exists, owing to the repulsion of the oxygen and imidazole nitrogen lone pairs (Figure 2.11-C). Through the semi-empiric AM1 calculation, the rotational isomer of molecule **41**, corresponding to the other minimum of molecule **19** of the training set, is estimated to have an energy of 8.9 kcal/mol higher than the absolute minimum. The conformational analysis of derivative **42** provides similar results to derivative **41**. However, using the correct starting poses resulting from the quantum-mechanical conformational study, CK2score overestimates molecule **41**, and **42**. Concluding, the missing of entropic contributions, erroneous estimations of the inhibitor internal energy resulting from the binding conformation, and dynamic interactions with different conformational states of the protein are only some of the possible limits of the LIE approach in predicting binding free energies.

Another interesting consideration can be made considering the capacity of CK2scoreB model to accurately predict the energy of binding of K44, characterized by a binding mode of the bromo-benzoimidazole scaffold rotated around the benzo moiety towards the hinge region, a position similar to K37A. This observation can be useful to understand the differences and the robustness of the two alternative LIE models.

inhibitor	R1	R2	R4	R5	R6	R7	IC$_{50}$ (μM)	ΔG_{exp} (kcal/mol)
26	-H	-CF$_3$	-H	-Br	-Br	-H	28.00	−6.25
27	-H	-CF$_3$	-Br	-H	-Br	-H	40.00	−6.04
28	-H	-CF$_3$	-Br	-Br	-Br	-H	1.20	−8.13
29	-H	-CF$_3$	-Br	-Br	-Br	-Br	0.60	−8.54
30	-CH$_3$	-CF$_3$	-Br	-Br	-Br	-Br	1.71	−7.92
31	-CH$_2$CH$_3$	-CF$_3$	-Br	-Br	-Br	-Br	6.12	−7.16
32	-H	-CF$_2$CF$_3$	-Br	-Br	-Br	-Br	0.40	−8.78
33	-H	-(CF$_2$)$_2$CF$_3$	-Br	-Br	-Br	-Br	1.28	−8.09
34	-H	-(CF$_2$)$_3$CF$_3$	-Br	-Br	-Br	-Br	1.48	−8.00
35	-H	-Cl	-Br	-Br	-Br	-Br	0.49	−8.66
36	-H	-CF$_3$	-Br	-Cl	-Br	-Br	0.39	−8.80
37	-H	-CF$_3$	-Cl	-Br	-Cl	-Br	1.25	−8.10
38	-H	-CF$_3$	-Br	-Cl	-Cl	-Br	1.96	−7.83
39	-H	-SH	-Br	-Br	-Br	-Br	0.91	−8.29
40	-H	-CF$_3$	-H	-H	-Br	-H	40.00	−6.04
41	-CH$_3$	-OCH(CH$_3$)$_2$	-Br	-Br	-Br	-Br	20.39	−6.44
42	-CH$_3$	-OCH$_2$CH$_3$	-Br	-Br	-Br	-Br	40.00	−6.04
43	-H	-S—C$_6$H$_4$—NO$_2$	-Br	-Br	-Br	-Br	3.48	−7.49
44	-H	-SCH$_2$COOCH$_2$CH$_3$	-Br	-Br	-Br	-Br	1.55	−7.97
45 K44	(structure)						0.10^a	−9.61

^{a} K_i was used instead of IC$_{50}$.

Figure 2.8: CK2 Inhibitors Used as Test Set for CK2scoreA and CK2scoreB.
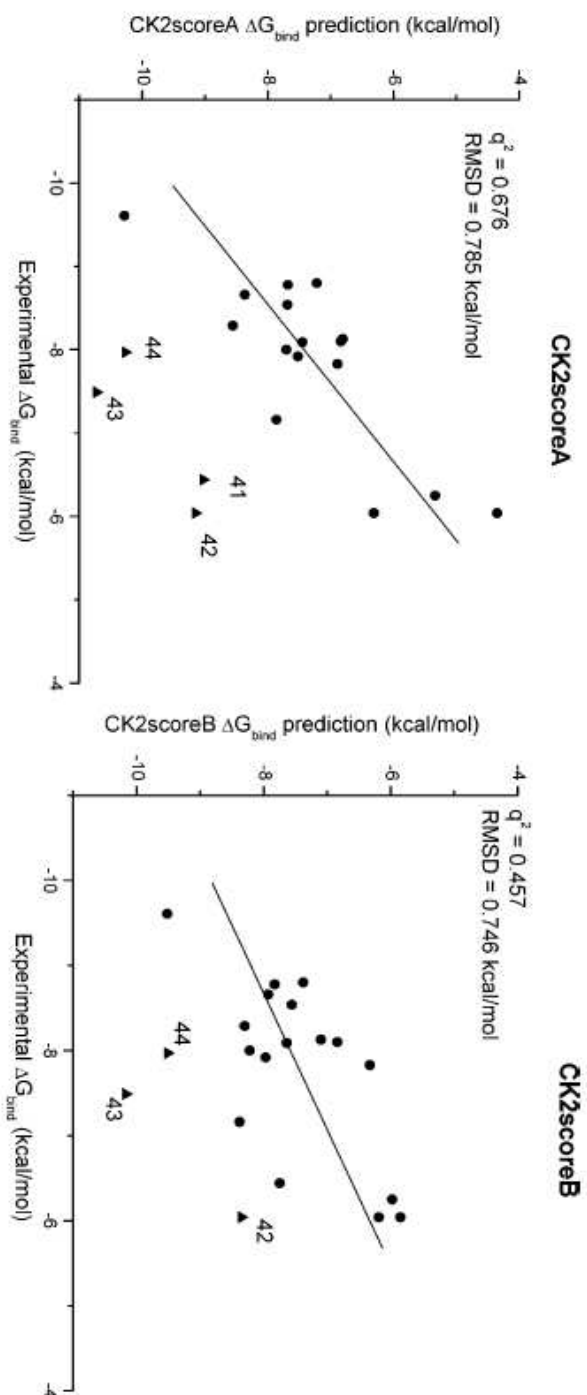
Figure 2.9: Comparison of the Predicted versus Experimental ΔG_{bind} of the Molecules Used as Test Set for CK2scoreA and CK2scoreB. RMSD and q^2 are referred to linear fit of the points indicated with circles. The points represented by triangles are outliers, their number is referred to Fig. 2.8.

inhibitor	ΔG_{exp} (kcal/mol)	$\Delta G_{\text{CK2scoreA}}$ (kcal/mol)	$\Delta G_{\text{CK2scoreB}}$ (kcal/mol)
26	−6.25	−5.34	−5.98
27	−6.04	−6.31	−6.19
28	−8.13	−6.81	−7.10
29	−8.54	−7.69	−7.56
30	−7.92	−7.52	−7.97
31	−7.16	−7.86	−8.38
32	−8.78	−7.68	−7.83
33	−8.09	−7.45	−7.64
34	−8.00	−7.71	−8.22
35	−8.66	−8.38	−7.93
36	−8.80	−7.22	−7.38
37	−8.10	−6.84	−6.84
38	−7.83	−6.89	−6.33
39	−8.29	−8.56	−8.30
40	−6.04	−4.35	−5.85
41	−6.44	−9.03	−7.75
42	−6.04	−9.15	−8.35
43	−7.49	−10.72	−10.17
44	−7.97	−10.27	−9.51
45	−9.61	−10.28	−9.52

Figure 2.10: Test Set Observed (ΔG_{exp}, kcal/mol) and Calculated ($\Delta G_{CK2scoreA}$ and $\Delta G_{CK2scoreB}$, kcal/mol) Free Energies of Binding for CK2scoreA and CK2scoreB.

2.2.6 General Considerations

The proposed stepwise protocol in combination with the linear interaction method provides a powerful and fast computational technique to accurately predict the binding free energies of this potent class of CK2 inhibitors.. The estimated energy terms (α, β and γ), the calculated interaction energies, the starting binding poses and the corresponding final minimized poses, together with a robust statistical validation can be utilized to appropriately describe at molecular level the possible differences in the CK2-inhibitor recognition processes. In this scenario, the presence of outliers can be useful to understand the limits of the model and the applicability of the theoretical framework. Moreover, it also offers a chemical insight into the key binding interactions, such as the possible role of water molecules, and rationalize the different binding modes shown by the available crystal structures. This important knowledge can be the key for a possible development of new bromo-benzimidazole derivatives. For example, could be interesting to understand if all four bromide moieties are equally important for the ligand binding or if one of these could be substituted to obtain hydrogen bonding with the hinge region. To achieve a dual polar interaction at the two opposite site of the ATP-binding pocket was in fact demonstrated by the new inhibitor ellagic acid to improve greatly the inhibitor activity [88].

Different inhibitors binding modes could be investigated accurately using the LIE approach shown in this work. Even if crystal structures of their complex with the protein are not available, the starting poses used for the LIE study can be obtained using common docking programs. The purpose could be to try to map the key parts of the ligands and how they can effect binding modes or the coefficients of the LIE equation.
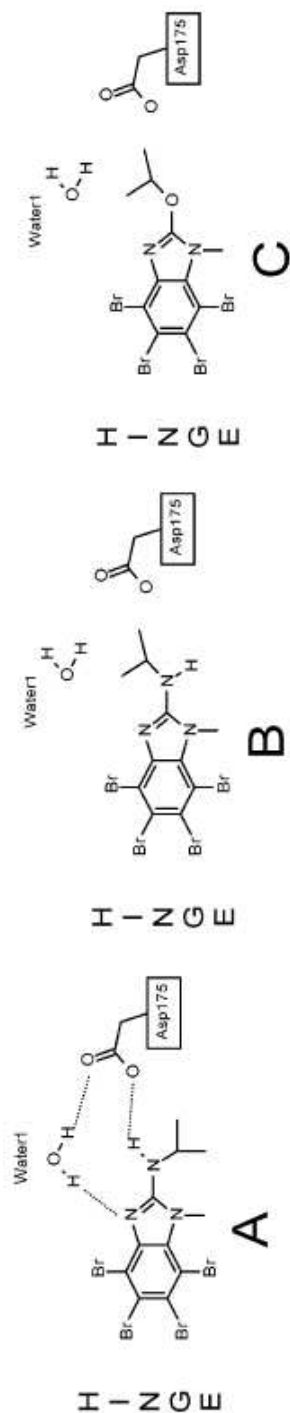
Figure 2.11: Comparison between the possible binding mode of molecule **19** of the training set (Ki=0.19 μM) and molecule **41** of the test set (IC$_{50}$=20.39 μM). Quantum mechanical study with the aqueous solvation energy of ligand **19** point out two different absolute energy minimum conformation. (A) In the first one, finally used for the LIE study, the inhibitor creates two hydrogen bonding, one with a Water 1 and the other with the catalytic aspartic acid. (B) In the second there is no electrostatic interaction with the protein. (C) For molecule **41**, the only energy minimum present, corresponds to this last one of ligand **19** and no hydrogen bonding is created with CK2.

2.3   Activity and Binding Mode

This LIE study can help us to extract general rules to understand what controls the activity of these inhibitors: as we discuss before electrostatic interactions seem to be not crucial for the free energy of binding. However it is necessary to interpretate carefully the meaning of α, β and γ, the three energy coefficients of the LIE model. First of all there is an empirical part included in them that in theory could be not strictly linked to the energy term: it can be for example linked to entropy contribution or internal energy factors. Another observation is that they rapresent from a statistical point of view the importance of the energy term, but this is a relative weight among the series in study.

2.3.1   Electrostatic Contribution

If we move from the activity point of view to the binding mode issue we can see that electrostatic is crucial to control the orientation and position of inhibitors. The crystal structure of the CK2 apo form shows 3 waters molecules near the catalytic center of the enzyme (Fig. 2.12) where a Poisson-Boltzmann study identifies an iso-surface of about 1.5 kcal/mol positive electrostatic potential. The calculation was performed using the default protocol of the Electrostatic Feature Map implemented in the Molecular Operating Environment (MOE) suite [137]. The space-dependent electric potential equations were determined by solution of non-linear Poisson-Boltzmann equation (PBE) using finite difference multi-grid technique [140].

   This is particularly interesting if we consider that several crystal structures of CK2 inhibitors complexes show one or two of these waters can be substituted by particular functional group of the ligands. These conserved molecules are capable to adapt the receptor to the inhibitor or from another point of view the inhibitor can adapt itself to the protein exploiting these waters.

2.3.2   Van der Waals and Hydrophobic Contribution

On other hand the final binding mode is the result of important electrostatic but also hydrophobic and van der Waals interactions. In Figure 2.13 the comparison of the crystal structure of the apo form of CK2 and CDK2 is shown to highlight that also CDK2 posses 3 waters near the DFG motif, but the 3rd is farer from the positive isosurface. This is due to a different aminoacid: the Ile174 in CK2 is substituted by a less bulky Ala that does not block the water in the isosurface. This difference together with another mutation always of an Isoleucine in Alanine results in the shift of TBB more toward the hinge region,
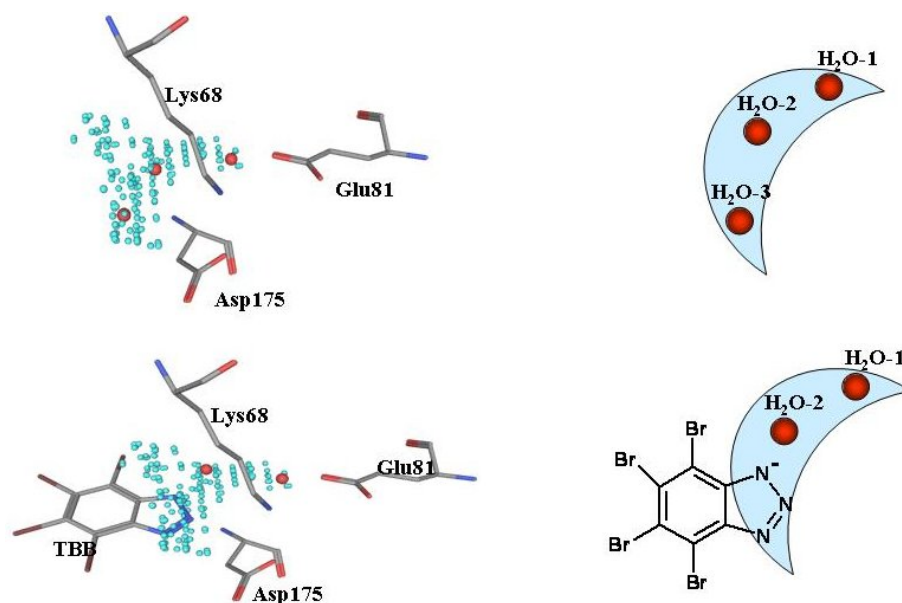
Figure 2.12: CK2 ATP-binding site. The residue Lys68 and the catalytic amino acid Asp175 are shown, together with Glu81 from the αC-helix. On the top, the three Apo-CK2 water molecules (W1, W2 and W3) nearer to the DFG conserved amino acids, are located inside an iso-surface with 1.5-2.0 kcal/mol positive electrostatic potential represented by small red spheres. On the bottom a comparison with the crystal structure of the CK2-TBB complex highlights how the inhibitor interactions with the active pocket involve the substitution of W3 with the ligand anionic nitrogen, while the other two waters keep an almost identical position.

resulting in a lost of activity of about 30 times. This hypothesis is confirmed by the IC_{50} of TBB on a mutant form of CK2 were both the Ile are substituted to Ala. In this way the ATP binding site of CK2 becomes more similar to the CDK2's one and TBB loses its inhibition power.

The hydrophobic character of CK2 ATP binding pocket is a peculiarity of this enzyme and a key feature for inhibitors binding. The more the small organic molecule is able to cover the apolar side chains of the aminoacids of the cleft the more the ligand will be active. In Figure number 2.14 is shown the linear correlation between predicted activity using the LIE model and the protein surface area in contact with the inhibitor. In this way is possible to understand how the increment of the number of bromide moieties in the benzotriazole scaffold results in an improvement of the complementarity of the molecule to the protein active site. Moreover thanks to the LIE function is possible to estimate the relative importance of the bromide in the different position to the final bind-
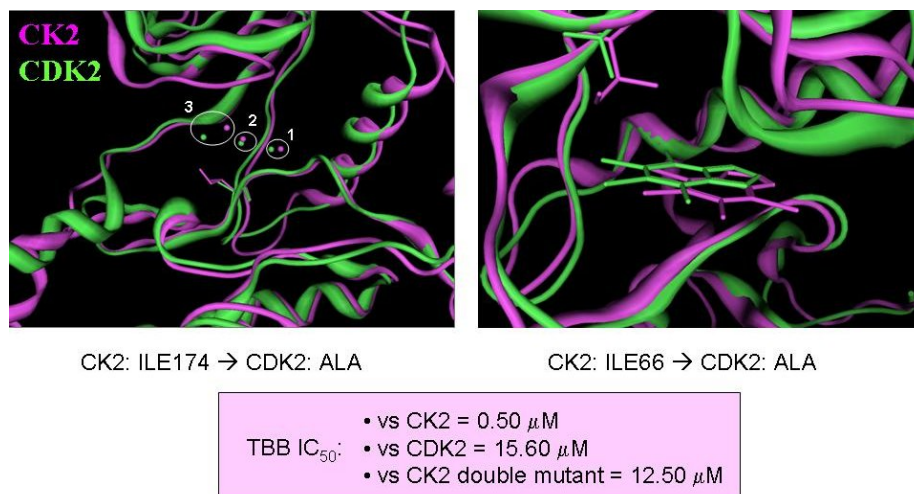
Figure 2.13: On the left superimposition of the apo form of CK2 (magenta) and CDK2 (green). The conserved waters are inside circles and Ile174 of CK2 and the corresponding Ala of CDK2 are shown in sticks. On the right Ile66 (CK2) and the corresponding Ala of CDK2 are shown in stick as well as TBB from crystal structures (PDBcode 1J91 and 1P5E).

ing free energy (Fig. 2.15). In particular the one in position 1, deep inside the pocket, is the most important, followed by the one in position 4, 2 and 3, more solvent exposed. This can be an example of how the LIE model could be used for lead optimization studies: in this case a possible evolution of TBB could be a molecule with a different group in position 3, ables to create an hydrogen bonding with the hinge region
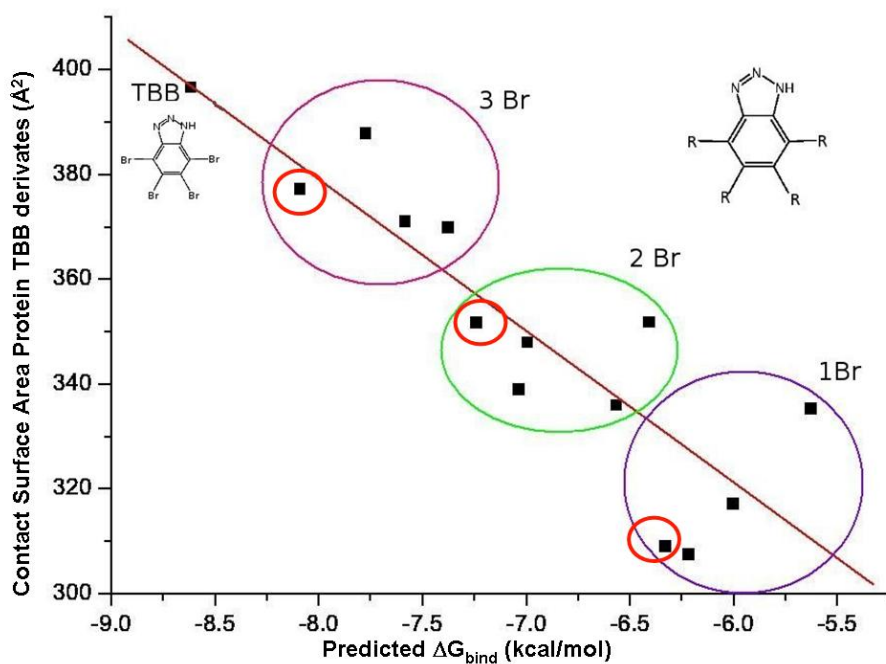
Figure 2.14: Linear correlation between predicted activity using the LIE model and the protein surface area in contact with the inhibitor. All possible combinations were tested for every number of bromide atoms. The red circle indicates the most active structural isomer for every series.
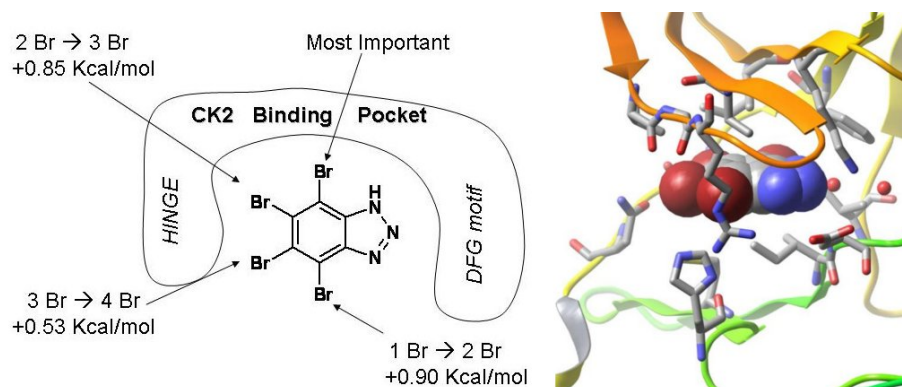
Figure 2.15: Importance of the bromide positions for the final TBB activity. On the left it is shown the increment of activity from 1 to 4 bromides, starting from the most important position to the less crucial. On the right TBB in CPK inside the CK2 ATP binding pocket.

2.4   Coumarin Scaffold

To further validate the hypothesys developed for TBB derivatives and to understand if they are applicable to other CK2 inhibitors, we focused our attention to the coumarin derivative DBC. Interestingly, dietary exposure to coumarins and benzopyrones in general, is quite significant, as these compounds are found in vegetables, fruit, seeds, nuts, coffee, tea and wine. Coumarin is a natural substance that has shown anti-tumour activity in vivo, with the effect believed to be due to its metabolites (e.g. 7-hydroxycoumarin) [93].

We studied an in-house focused library of coumarin derivatives with known CK2 activity (more than 60, Fig. 2.16) with the aim of elucidated the putative binding motif and explain structure-activity relationships. In this study, the X-ray diffraction crystal structure of CK2 in complex with DBC has been exploited as starting point for a linear interaction energy study to rationalize the different free energy of binding and the key interactions of all coumarin derivatives. This computational approach is an efficient tool to quantitatively explore scaffold decorability, and to evaluate regions of the protein active site crucial for ligand binding, as already proved for a class of CK2 inhibitors such as bromo-benzoimidazole [101].

2.4.1   Methods

Protein and Inhibitors Preparation

Using the software package Molecular Operating Environment (MOE 2006.08) [137], all the ligands and the X-ray diffraction crystal structure of CK2 in complex with the ATP competitive inhibitor DBC were prepared for the linear interaction energy analysis. In the protein active site was included for the molecular modeling study the water W1. This molecule indeed is well conserved in almost all CK2-inhibitors complexes crystal structures and seems crucial for their binding. Hydrogen were added and energy minimized using the AMBER99 force field [141] until the energy gradient reached 0.05 kcal/mol. This protocol was exploited for both the neutral and anionic states of DBC to achieve two possible orientations of the water as donor or acceptor of hydrogen bonding. Furthermore to fully understand the importance of acidity of the 7-hydroxy coumarins derivatives, has been investigated the hypothesis of the ionic couple formation in the active site created by the anionic DBC form and the hydronium ion. All inhibitors were built with MOE and energy minimized using MMFF94x force field [129] until a 0.01 energy gradient was attained. Their ionic states were evaluated using ACDLabs suite [73] and as starting pose of the LIE calculation

the molecules were superimposed to DBC, since all these ligands have similar scaffold is likely that they share a common binding mode.

Linear Interaction Energy Model

A linear interaction energy (LIE) approach has been used to evaluate the binding free energy of this class of CK2 inhibitors following a similar computational approached used for TBB derivatives. The LIE energy model was created using the default options of the MOE-LIE suite [142]. Briefly, using the MMFF94x force field and the surface generalized Born (SGB) implicit solvent [130], a Truncated Newton minimization was performed for ligands in complex with CK2 and free in solution with a residue-based cutoff distance of 15 Å, a 0.5 root mean square (rms) gradient for convergence and a maximum of 500 steps. The energy terms calculated with this protocol were used to build the binding free energy model.

It was possible to include a total of 16 coumarins, shown in Figure 2.17, in the training set to generate the quantitative LIE model (Fig. 2.21). Most probably this is linked to the fact that some particular chemical features, present only in this subset of molecules, are the key to achieve a DBC-like binding mode, that was used as starting point for the LIE calculation. First of all, an hydroxyl group capable to strongly interact with the conserved water W1 (see Figure 2.19) must be present at the 7 position of the coumarin scaffold. This is the only crucial electrostatic interaction shown by this class of inhibitors, and it seems to be strongly modulated in its intensity by the presence of an electron-withdrawing substituent at the 6 position of the coumarin scaffold. According to the Hammett theory, the presence of an electron-withdrawing substituent in orto position of phenol OH increases the acidic behavior of the phenolic group. To simulate this highly polarized hydrogen bonding, we created two possible working hypothesis:

1. In the first case the hydroxyl group in 7 becomes anionic before forming the complex with CK2, and it is considered with a formal net charge equal to -1;

2. In the second assumption, phenol OH participates at the active site in its neutral form, and the hydrogen bonding with water W1 is so highly polarized that can be more suitable modeled as an ionic couple between an hydronium ion (water W1 is transformed into H_3O^+) and the anionic form of the phenol OH.

Both these possibilities are in accordance with the fact that only ligands with a
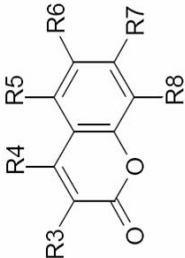
low pKa, estimated using the ACDLabs [143] package, can be studied using the LIE approach considering a DBC-like binding mode. Both hypothesis resulted in two high-quality and roughly comparable energy models, the best one presented here is derived by in situ ionic couple theory. Finally, almost all possess hydrophobic groups at the 8 position of the coumarin moiety, an hydrogen atom in 5, and a not-hydrophobic group at the 6 position.

The calculated energy terms (Fig. 2.17) result in a final model with $\alpha = 0.029$, $\beta = -0.003$ and $\gamma = 0.016$. The obtained correlation factor is $r^2 = 0.771$ with an RMSD of 0.303 kcal/mol and the cross-validated q^2 is slightly lower ($q^2 = 0.633$, cross-validated RMSD = 0.491 kcal/mol). Free energies of binding, calculated from experimental IC_{50}, and predicted ΔG_{bind} are shown in Figure 2.18, and finally plotted in Figure 2.19.

2.4.2 Qualitative Structure-Activity Relationship (SAR)

Several important structural features of coumarin derivatives can be individuated from a first qualitative analysis of their activity on CK2. Primarily, the hydroxyl group at the 7 position of the coumarin moiety can be considered an essential feature even if not sufficient to achieve an IC_{50} in a sub-micromolar range. Indeed, to increase the inhibitory potency an electron withdrawing substituent should be simultaneously present at the position 8. For instance, in derivative **5** (IC_{50}=0.74) a simple substitution of the bromide atom with a methoxy group (derivative 50) results in a complete loss of activity. This important effect can be rationalized according to the increase of the 7-hydroxyl group's acidity as confirmed by our in silico predictions, where to achieve a sub-micromolar activity is necessary that the pKa of the 7-hydroxyl group is lower than 7. This hypothesis is also supported by the experimental evidence that the substitution on DBC of the 7-hydroxyl group with the 7-amine group results again in a complete loss of activity (see derivative **62**).

Hydrophobic groups at positions 3 and 4, such as bromide or methyl, improve the activity while substitutions at 5 position result in a loss of inhibition efficiency. Halogen or nitro substituents at position 6 can be compatible with an activity around 1-4 μM according to the other decorations of coumarin moiety.

Name	R3	R4	R5	R6	R7	R8	IC$_{50}$ (µM)	ΔG$_{exp}$ (kcal/mol)	Predicted pK$_a$[a]	Ionic State
1 DBC	-Br	-CH$_3$	-H	-H	-OH	-Br	0.10 (K$_i$=0.06 µM)	-9.61	5.88	-1
2	-Br	-CH$_3$	-H	-H	-OH	-I	0.28	-9.00	5.96	-1
3	-Br	-CH$_3$	-H	-H	-OH	-Cl	0.32	-8.92	5.94	-1
4	-Br	-CH$_3$	-H	-Br	-OH	-H	0.66	-8.48	5.88	-1
5	-Br	-CH$_3$	-H	-H	-OH	-Br	0.74	-8.42	6.48	-1
6	-H	-CH$_3$	-H	-H	-OH	-I	0.80	-8.37	6.56	-1
7	-Br	-CH$_3$	-H	-Br	-OH	-Br	1.90	-7.85	4.35	-1
8	-Br	-CH$_3$	-H	-H	-OH	-CN	1.90	-7.85	4.62	-1
9	-Br	-CH$_3$	-H	-OH	-OH	-Br	2.00	-7.82	5.72	-1
10	-H	-CH$_3$	-H	-H	-OH	-Cl	2.20	-7.77	6.54	-1
11	-Br	-CH$_3$	-H	-H	-OCOCH$_3$	-Br	2.50	-7.69	/	0
12	-Br	-CH$_3$	-H	-H	-OH	-CH=NOH	2.50	-7.69	6.53	-1
13	-H	-CH$_3$	-H	-I	-OH	-I	2.66	-7.65	5.12	-1
14	-Br	-CH$_3$	-H	-H	-OH	-COH	3.30	-7.52	5.62	-1
15	-Br	-CH$_3$	-H	-NO$_2$	-OH	-H	3.30	-7.52	4.20	-1
16	-Br	-CH$_3$	-H	-Br	-OH	-OH	3.60	-7.47	5.72	-1
17	-H	-CH$_3$	-H	-H	-OH	-NO$_2$	4.00	-7.41	2.95	-1

No.	R1	R2	R3	R4	R5	R6	IC50			
18	-H	-CH₂Cl	-H	-H	-OH	-OH	4.00	-7.41	7.44	0
19	-Br	-CH₃	-H	-Br	-OH	-NO₂	4.00	-7.41	0.59	-1
20	-Br	-CH₃	-H	-H	-OH	-NO₂	4.00	-7.41	2.35	-1
21	-Br	-CH₃	-H	-NO₂	-OH	-Br	4.00	-7.41	2.42	-1
22	-Br	-CH₃	-H	-Br	-OH	-Cl	4.03	-7.41	4.41	-1
23	-H	-CH₃	-H	-H	-OH	-CN	4.20	-7.38	5.22	-1
24	-Br	-CH₃	-H	-H	-OH	-OH	10.50	-6.83	9.99	0
25	-H	-H	-H	-H	-OH	-Cl	10.70	-6.82	6.43	-1
26	-Br	-H	-H	-Br	-OH	-Br	15.21	-6.61	4.24	-1
27	-H	-CH₃	-OH	-H	-OH	-C=NOH	20.00	-6.45	7.13	0
28	-Br	-CH₃	-H	-Br	-OH	-Br	21.50	-6.41	4.14	-2
29	-H	-CH₃	-H	-H	-OH	(structure)	22.70	-6.37	5.36	-1
30	-CH₃	-CH₃	-H	-H	-OH	-OCH₃	27.00	-6.27	8.13	0
31	-Br	-OH	-H	-NO₂	-OH	-Br	28.00	-6.25	6.49	-2
32	-H	-CH₃	-H	-NO₂	-OH	-CH₃	30.00	-6.21	5.19	-1
33	-H	-CH₃	-H	-NO₂	-OH	-H	30.00	-6.21	4.80	-1
34	-Br	-CH₃	-OH	-H	-OH	-H	31.00	-6.19	7.25	0
35	-H	-H	-OH	-H	-OH	-H	34.96	-6.12	8.61	0
36	-H	-CH₃	-H	-NO₂	-OH	-CH₃	34.96	-6.12	8.72	0
37	-Br	-H	-H	-H	-OH	-H	35.00	-6.12	4.48	-1
38	-Br	-OH	-H	-H	-OH	-H	39.00	-6.05	7.4	0
39	-COOEt	-H	-Br	-H	-H	-H	117.5	-5.39	/	0
40	-COO'	-H	-Br	-H	-H	-H	147.00	-5.26	/	-1
41	-H	-CH₃	-H	-H	-OH	-H	>100.00	>-5.50	8.00	0
42	-H	-H	-H	-H	-OH	-H	>100.00	>-5.50	7.89	0
43	-H	-H	-Br	-H	-OH	(structure)	>100.00	>-5.50	3.78	-1
44	-H	-H	-Br	(structure)	-H	-H	>100.00	>5.50	/	0
45	-H	-H	-H	(structure)	-H	-Br	>100.00	>-5.50	/	0
46	-NO₂	-CH₃	-H	-NO₂	-OH	-H	>100.00	>-5.50	4.36	-1
47	-H	-CH₃	-H	-NO₂	-OH	-OH	>100.00	>-5.50	4.61	-1
48	-NO₂	-CH₃	-H	-H	-H	-H	>100.00	>-5.50	/	-1
49	-NO₂	-CH₃	-H	-NO₂	-OH	-CH₃	>40	>-6.00	4.77	-1

						IC$_{50}$	pKa	pKa^a	Net Charge
50	-H	-CH$_3$	-H	-H	-OH	-OCH$_3$	>40	>6.00	8.02	0
51	-H	-CH$_2$Cl	-H	-H	-OH	-OCH$_3$	>40	>6.00	7.74	0
52	-CH$_3$	-H	-OCH$_3$	-H	-OH	-H	>40	>6.00	7.64	0
53	-H	-CH$_2$OH	-H	-H	-OH	-OH	>40	>6.00	10.48	0
54	-NO$_2$	-CH$_3$	-H	-NO$_2$	-OH	-NO$_2$	>100.00	>5.50	0.71	-1
55	-H	-CH$_3$	-H	-NO$_2$	-OCOCH$_3$	-H	>100.00	>5.50	/	0
56	-H	-CH3	-H	-H	-OCOCH$_3$	-NO$_2$	>100.00	>5.50	/	0
57	-H	-H	-H	[structure]	-OH	-OCH$_3$	>100.00	>5.50	6.64	-1
58	-H	-H	-H	[structure]	-OH	-OCH$_3$	>100.00	>5.50	5.54	-1
59	-H	-CH$_2$OH	-H	-H	-OH	-H	>40	>6.00	7.89	0
60	-H	-CH$_2$Cl	-H	-H	-OH	-H	>40	>6.00	7.72	0
61	-NH$_2$	-H	-H	-H	-OH	-OCH$_3$	>40	>6.00	8.16	0
62	-Br	-CH$_3$	-H	-H	-NH$_2$	-Br	>40	>6.00	/	0
63	-Br	-CH$_3$	-H	-Br	-NH$_2$	-Br	>40	>6.00	/	0
64	-H	-CH$_3$	-H	-H	-OH	-COH	>40	>6.00	6.22	-1
65	-Br	-OH	-H	-Br	-OH	-Br	>40	>6.00	4.49	-2
66	-Br	-CH$_3$	-H	-Br	-OCH$_3$	-H	>40	>6.00	/	0
67	-Br	-CH$_3$	-H	-H	-OCH$_3$	-Br	>40	>6.00	/	0

Figure 2.16: Inhibition of CK2 by Coumarins Analogues. The IC$_{50}$ activity data represent the means of three independent experiments with SEM never exceeding 15%. The in silico Prediction of R7 Substituent pKa and the Net Charge of the Most Populated Ionic State at Physiological pH is Shown. ^{a}The error estimated on the predicted pKa of the phenolic OH at position 7 is ±0.02.

Name	U^{b-f}_{elect} kcal/mol	U^{b-f}_{vdw} kcal/mol	$HASA^{b-f}$ kcal/mol	ASA^{b-f} kcal/mol	$ASA+^{b-f}$ kcal/mol	$ASA-^{b-f}$ kcal/mol	$CASA+^{b-f}$ kcal/mol	$CASA-^{b-f}$ kcal/mol	VSA^{b-f} kcal/mol	$PASA^{b-f}$ kcal/mol
1 DBC	-197.7653	1.5702	-448.1037	-557.2585	-100.0504	-288.9721	-74.2579	868.0718	2806.5571	-109.1548
2	-193.7874	4.4171	-451.7560	-564.7581	-100.9600	-294.8176	-76.9676	-187.3542	2543.4438	-113.0021
3	-215.5250	1.5011	-430.4421	-540.7503	-100.0653	-273.8589	-59.2773	2885.0308	1206.3640	-110.3082
4	-220.7834	18.9818	-447.4083	-557.1554	-89.2716	-303.8774	-103.1900	4220.9419	-1929.1779	-109.7472
5	-167.8966	2.2074	-397.6865	-512.5418	-113.6143	-233.5788	-64.8164	4264.6226	2366.2527	-114.8553
6	-169.0103	5.0361	-398.7990	-516.8976	-113.9363	-237.4862	-66.6977	3336.0249	2459.3225	-118.0987
8	-219.6383	-0.4217	-374.5649	-553.6124	-126.7920	-260.2887	-60.4464	4194.0859	2781.6345	-179.0476
9	-231.1221	4.3391	-414.0914	-570.3610	-109.6718	-292.0350	-89.8760	2781.8528	3626.5566	-156.2696
10	-182.6900	3.4665	-379.5748	-494.7814	-113.7104	-217.3335	-50.6239	2417.8528	962.4318	-115.2067
12	-229.2684	13.3805	-381.2005	-574.2347	-160.1861	-248.5138	-88.1841	6509.5933	-1039.3051	-193.0342
14	-212.7549	2.0311	-388.4405	-543.1951	-145.6445	-230.2745	-61.6509	7457.0674	3919.6196	-154.7546
17	-192.0742	3.4807	-373.1952	-560.5029	-101.5250	-291.5494	-66.2913	5628.3013	5357.7446	-187.3077
20	-197.6469	6.5256	-319.2815	-513.6910	-112.6414	-237.0944	-56.6301	9666.1660	4908.8560	-194.4096
23	-218.6269	3.7350	-320.2022	-503.8472	-139.2490	-199.9986	-52.0343	8096.2407	3728.5420	-183.6451
25	-191.6230	3.6139	-356.0089	-470.6675	-122.9562	-235.5503	-46.7456	6553.1392	3481.0146	-114.6586
27	-202.7459	12.6623	-332.0560	-538.6724	-173.7345	-192.0848	-78.5291	10961.5870	-1002.8069	-196.6165

Figure 2.17: Ensemble Average LIE Energy Terms for the Inhibitors Used to Build the Energy Model. U^{b-f} elect, U^{b-f} vdw, $HASA^{b-f}$, ASA^{b-f}, $ASA+^{b-f}$, $ASA-^{b-f}$, $CASA+^{b-f}$, $CASA-^{b-f}$, VSA^{b-f}, $PASA^{b-f}$ are respectively Electrostatic, van der Waals, total hydrophobic surface area, water accessible surface area, positive accessible surface area, negative accessible surface area, charge-weighted positive surface area, charge-weighted negative surface area, van der Waals surface area and total polar surface area energy terms. All the energy terms are the result of the subtraction from the bound state energy estimation of the free state energy estimation. Together with U^{b-f} elect and U^{b-f} vdw, $HASA^{b-f}$ surface has been chosen to built the final energy model.

2.4.3 SAR Considering Structural Information

The crystal structure of DBC in complex with CK2 nicely explains the pharmacophore derived by SAR considerations, as summarized in Figure 2.20.

In fact, hydrophobic groups in position 3, 4 and 8 positions can interact through van der Waals interactions with the CK2 active site while the importance of the 7-hydroxyl group is explained by its key interactions with Lys68 and water W1, a water molecule very well conserved in all CK2-inhibitor complexes available. These represent the only electrostatic interactions exploited by DBC in the binding to CK2, and most probably all other coumarin derivatives must possess this ionizable hydroxyl group to achieve a similar binding mode.

However, considering the strong hydrophobic nature of CK2 ATP-binding pocket, the crucial role of both hydrophobic interactions and desolvation effect in ligand binding is doubtless. It is most likely that when the hydrophobic contribution to the binding becomes sufficiently strong it is possible to have active compounds that could present a binding mode different from DBC as seen for the dibromobenzo imidazole and triazole derivatives. In fact, crystal structure analysis and molecular modeling studies have demonstrated that these scaffolds can bind the CK2 active pocket into two different orientations. When the inhibitor can interact through a polarized hydrogen bonding with the water molecule, like Tetrabromobenzotriazole (TBB) derivatives, the position of the planar aromatic scaffold is similar to the one presented by the DBC-CK2 crystal structure; on the contrary, the ligand is shifted and rotated nearer the hinge region, like already observed in K44, K37, K25 or K17 crystal structure. In other words, even if DBC and TBB present different chemical structures, their ATP-binding pocket recognition are comparable, and water molecules seem to adapt the protein to the ligand peculiarities.

2.4.4 Quantitative Free Energy of Binding Model

Starting from these qualitative considerations, we analyzed the structure-activity relationship of coumarin derivatives from a more quantitative point of view. Using the Linear Interaction Energy method, it has been possible to build up a quantitative model using 16 coumarin derivatives and including in this training set all sub-micromolar inhibitors (Figure 2.20). The final model is statistically acceptable. This computational technique allows us to evaluate the principal contributions for inhibitors activity. The strongest effect is resulting from van der Waals interactions (α=0.029), in agreement with the steric complementarily between CK2 active pocket and almost all coumarins derivatives studied. This

Name	ΔG_{exp} (kcal/mol)	ΔG_{pred} (kcal/mol)	ΔG_{cv} (kcal/mol)
1 DBC	-9.61	-9.04	-8.87
2	-9.00	-9.03	-9.04
3	-8.92	-8.69	-8.63
4	-8.48	-8.44	-8.37
5	-8.42	-8.30	-8.25
6	-8.37	-8.23	-8.18
8	-7.85	-7.82	-7.81
9	-7.82	-8.29	-8.45
10	-7.77	-7.92	-7.94
12	-7.69	-7.49	-7.41
14	-7.52	-8.00	-8.07
17	-7.41	-7.78	-7.82
20	-7.41	-6.79	-6.59
23	-7.38	-6.81	-6.57
25	-6.82	-7.50	-7.59
27	-6.45	-6.80	-6.97

Figure 2.18: Comparison of Experimental (ΔG_{exp}, kcal/mol), Predicted (ΔG_{pred} kcal/mol) and Cross Validated (ΔG_{CV} kcal/mol) Free Energies of Binding for the 16 inhibitors included in the model.

is followed in absolute importance by the hydrophobic effect (γ=0.016) constituting the most variable part among the test set. Indeed this aspect is the most important to discriminate the relative activity of coumarins studied, as shown by γ normalized value of 0.854. On the contrary the very low absolute weight of electrostatic contribution (β=-0.003) is linked to the fact that the strong interaction with water molecule W1 is common in all inhibitors. For this reason even if this interaction is crucial for the inhibitor orientation in the active pocket, it is not important to discriminate the activity difference inside the series studied.

In particular, non polar contribution to the free energy of solvation in a LIE study is represented by the difference of the solvent accessible surface area between the ligand in complex and free in solution (ASA^{b-f}). Since this contribution seems crucial for the binding of CK2 inhibitor, we tried to include it
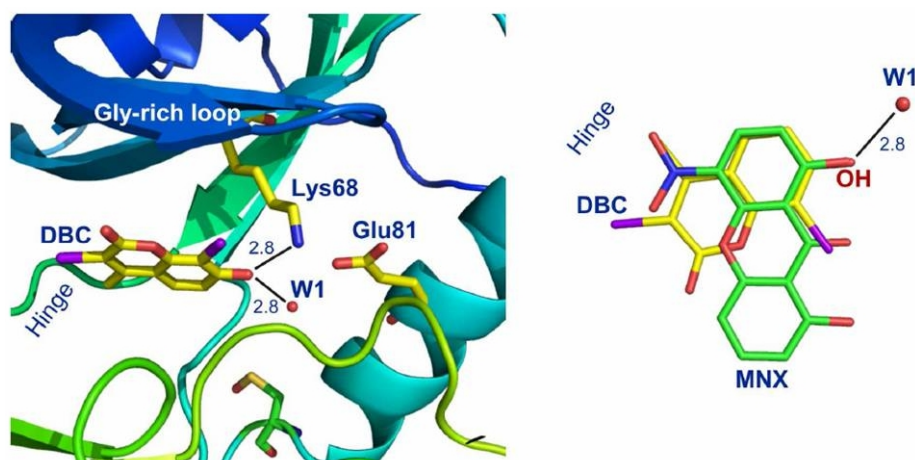
Figure 2.19: Left panel: DBC bound to the CK2 ATP-binding site. The hydroxyl function is involved in two hydrogen bonds (at a distance of 2.8 Å), with Lys68 and the conserved water molecule W1. Right panel: superimposition of the two inhibitors DBC (yellow carbon atoms), and MNX (green carbon atoms). Bromine atoms are shown in magenta. Note the identical position of the OH function of DBC and MNX. As reference, water molecule W1 and the zone of the hinge region are indicated.

in our LIE model also using alternative types of solvent accessible surface:such as the positive and negative accessible surface area (respectively $\text{ASA}+^{b-f}$ and ASA-^{b-f}), the charge-weighted positive and negative surface area (respectively $\text{CASA}+^{b-f}$ and CASA-^{b-f}), the van der Waals surface area (VSA^{b-f}), the total polar surface area (PASA^{b-f}) and the total hydrophobic surface area (HASA^{b-f}) as collected in Figure 2.17. Interestingly, only using the total hydrophobic surface area (HASA^{b-f}), it is possible to obtain an acceptable LIE model (Fig. 2.22), pointing out the important role of the hydrophobic contribution in the final free energy of binding. This is a clear indication of the major role played by the several non-polar aminoacids side chains that characterize the CK2 active pocket.

The final energy model derived was able to predict almost all inhibitors with an hydrophobic group at the 8 position and with the pKa value of the 7-hydroxyl group lower than 7 units. A possible reason can be related to the peculiarity of CK2 ATP binding site to allow different binding modes, even among similar inhibitors, as demonstrated by several structural data available. Indeed these chemical-physical requisites of coumarins seem to discriminate between a DBC-like binding mode, used as starting point for the LIE study, and other possible ligand orientation.
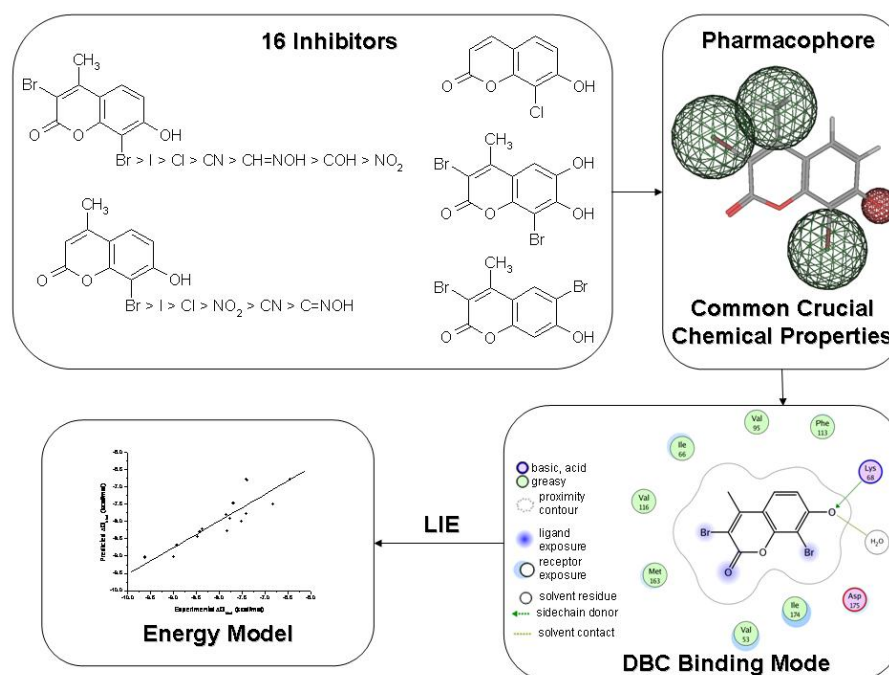
Figure 2.20: The 16 inhibitors used to build the LIE energy model using the DBC binding mode as starting position for the computational study are shown in the upper-left. The more the scaffold satisfies the base pharmacophore shown on the upper-right (the green spheres represent hydrophobic molecular regions, the small red region stand for a strongly polarized hydrogen bonding donor) higher the probably it will possess the chemical features to achieve a DBC similar binding mode to the CK2 ATP binding pocket.

2.4.5 General Considerations

CK2 inhibitors design strategy that started more than 15 years ago with the firsts ribofuranosyl-benzimidazole derivatives has been continuously evolving with a progressive development of many sub-micromolar inhibitors, constantly improving the structural knowledge of their activity. Results presented here can be considered a new important tessera in the always more complete CK2 mosaic. The multidisciplinary approach used in this study connect in a virtuous circle chemical synthesis, biological assays, structural information and different molecular modelling techniques to rational design new CK2 inhibitors. More than 60 coumarins derivatives synthesized and tested represent a strong base to analyze, propose and corroborate hypothesis about essential inhibitors features for CK2 binding. In this particular contest, the crystallization of DBC-CK2 complex

represent a very important improvement in mapping the most crucial chemical features into CK2 recognition process. Fully understanding of principal binding features is the first step in the inhibitors design process based on biostructural information. Finally, starting from this new crystallographic information, a Linear Interaction Energy method has been used as an efficient computational approach to understand the importance of different energy contributions to the final free energy of binding.

Figure 2.21: Free energy of binding estimated by the LIE model of each inhibitor used versus the experimentally measured data. On the right the correlation coefficient (r^2) and cross-validated coefficient (q^2) are shown. For the whole set RMSD and cross validate RMSD ($RMSD_{cv}$) were calculated. α, β and γ are the resulting coefficients of the LIE equation for CK2score.
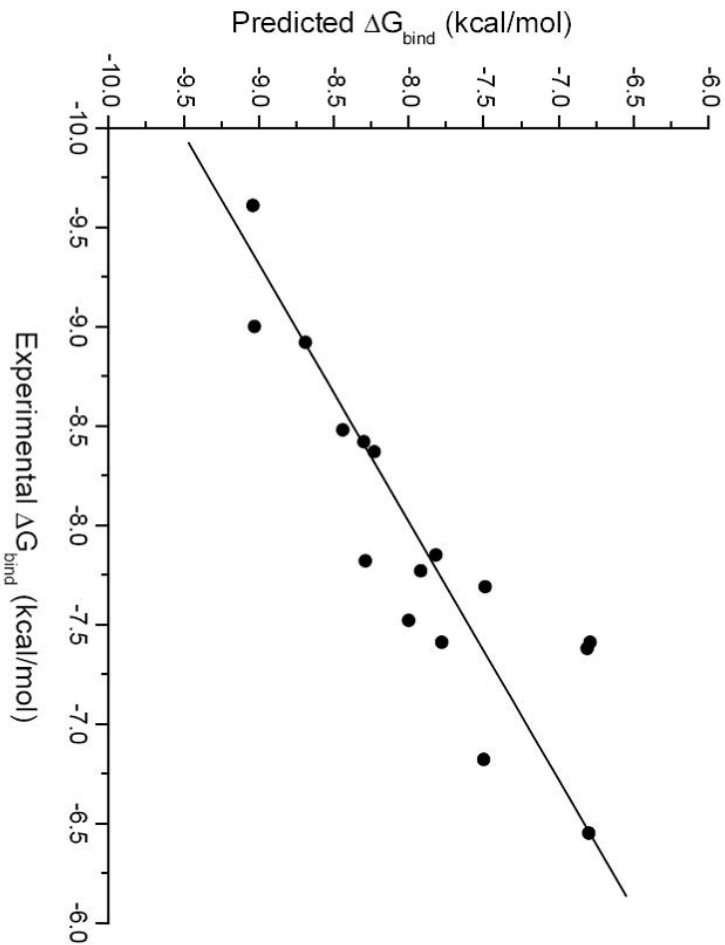
Estimated Linear Energy Model	
r^2	0.771
RMSD	0.303
q^2_{cv}	0.633
$RMSD_{cv}$	0.491
α (van der Waals energy term)	0.029
β (electrostatic energy term)	-0.003
γ (HASA, hydrophobic surface area)	0.016

Estimated Normalized Energy Model	
α / ΔG_{exp} Standard Deviation	0.184
β / ΔG_{exp}Standard Deviation	-0.081
γ / ΔG_{exp}Standard Deviation	0.854

Relative Importance of Descriptors	
α (van der Waals energy term)	0.216
β (electrostatic energy term)	0.095
γ (HASA, hydrophobic surface area)	1.000

ASA

$r^2 = 0.34$ $r^2_{CV} = 0.05$
RMSD = 0.36   RMSD$_{CV}$ = 0.89

ASA-

$r^2 = 0.49$ $r^2_{CV} = 0.24$
RMSD = 0.57   RMSD$_{CV}$ = 0.72

ASA+

$r^2 = 0.45$ $r^2_{CV} = 0.25$
RMSD = 0.59   RMSD$_{CV}$ = 0.72

Predicted ΔG_{bind} (kcal/mol)

Experimental ΔG_{bind} (kcal/mol)

Figure 2.22: LIE models dependence from the way the hydrophobic contribution to the free energy of binding is described. ASA^{b-f}, $ASA\text{-}^{b-f}$, $ASA\text{+}^{b-f}$, $CASA\text{-}^{b-f}$, $CASA\text{+}^{b-f}$, VSA^{b-f}, $PASA^{b-f}$ and $HASA^{b-f}$ are respectively total water accessible surface area, negative accessible surface area, positive accessible surface area, charge-weighted negative surface area, charge-weighted positive surface area, van der Waals surface area, polar surface area and total hydrophobic surface area. All the corresponding LIE model are shown in the charts resulting from the plot of experimental vs predicted free energy of bindings for the 16 inhibitors. Correlation and prediction factors are shown, as well as the cross-validated ones.

2.5   Conclusion

LIE method has for sure some important limits, but if we have enough structural and activity data it is a useful approach to create and energy function optimized for the system in study, and since it is the modelist to build the model on a training set, he will know very well its limits, its potentiality, its applicability and how to improve it.

When you have a good knowledge of the rules behind the binding mode and an optimized energy function your can have more chances to have success in the design of new inhibitors, as well as the possibility to understand errors and improve the system comprehension. One strategy can be to integrate crystal structures informations, docking studies, LIE models and crucial interactions insight to go from the target peculiarities to efficient new molecules. For example using a scaffold hybridization technique it is possible to start from 2 known inhibitors to try to create a new molecule that will posses the best characteristics of both (Fig. 2.23).
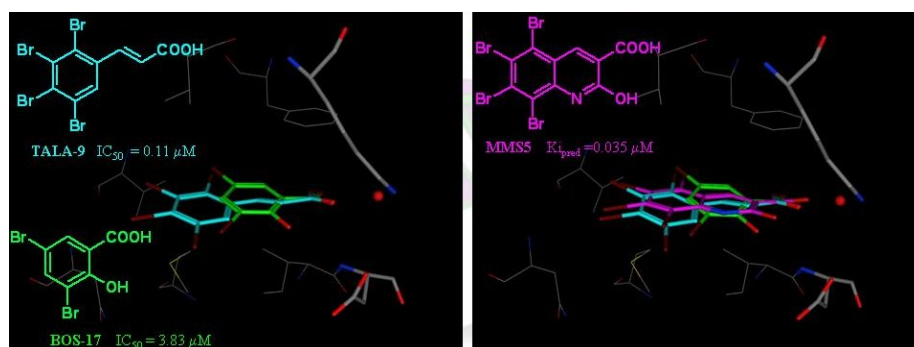
Figure 2.23: New inhibitors design using a scaffold hybridization approach based on crystal structures informations, docking studies, MD, LIE models and pharmacophoric analysis. In this example the rational behind the design of compound MMS5 is shown. Starting from two inhibitors with known activity (TALA-9, IC_{50}=0.11μM and BOS-17, IC_{50}=3.83μM), their hypothetical binding modes are predicted using different docking studies and pharmacophore approaches and validated by means of molecular dynamics. These structural informations are used to design a new molecule, MMS5, based on a scaffold hybridization idea. MMS5 binding mode is predicted and its stability validated and finally the free energy of binding estimated using the LIE energy model.

Chapter 3

Improving the Selectivity of Imatinib

To understand the molecular reasons of the activity loss of CGP-582 for Abl, the difference of free energy of binding between STI-571 and CGP-582 was evaluated using the molecular mechanics Poisson-Bolzmann surface area (MM-PBSA) and molecular mechanics Generalized Born surface area (MM-GBSA) computational techniques. These methods, involving both force field and solvation terms important for the binding, have been employed successfully in several biophysical studies to predict free energy changes [144].

3.1 Imatinib and CGP-582

The complexity of the multidimensional energy surface of a macromolecule does not allow a complete exploration of the conformational space of the system, too demanding from a computational point of view. To solve this issue we start the molecular dynamics (MD) study from a crystallographic conformation, assuming to be near a energy minimum, that we can explore exhaustively.

Since the only difference between CGP-582 and imatinib is the presence of the nitrogen group in benzene ring near the methyl piperazine, a common binding mode of the two inhibitors was assumed. This was considered similar to the one shown by STI-571 crystal structure in complex with Abl. However two distinct possible orientations of the pyridine ring that identifies CGP-582 are allowed, CGP-582 A and B (Figure 3.1), therefore MD and MM-PBSA studies were performed for both these two alternatives and for STI-571.
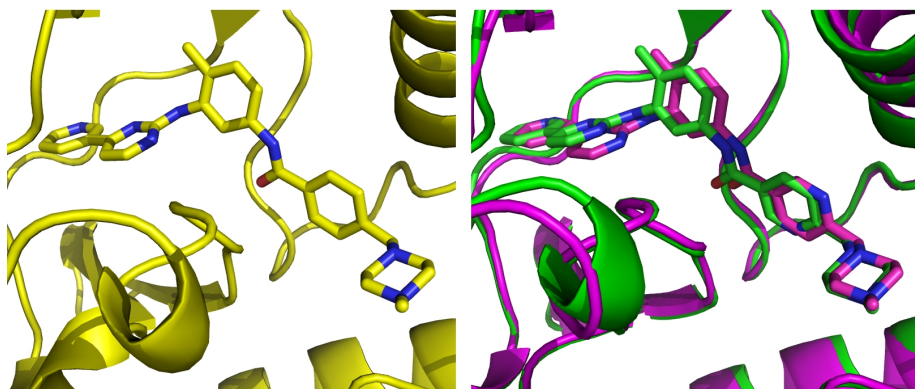
Figure 3.1: Ligands placements as average structure of 1 ns MD. On the left, the STI-571 Abl kinase domain complex is shown and on the right, the two different binding modes studied of CGP-582: in magenta CGP-582 binding mode A superimposed to CGP-582 binding mode B (green).

3.2   MM-PB/GBSA

3.2.1   Methods

Molecular Dynamics Simulation Procedure

The structure of Abl kinase domain in complex with STI-571 (PDB code 1IEP [145]) has been downloaded from the Brookhaven Protein Data Bank [99], while the structure of CGP-582 was created modifying Imatinib using Sybyl 7.0 [146]. After the optimization at the HF/6-31G* level using Gaussian 98 package [147], the ligands RESP partial charges [148] were calculated. All the molecular dynamics simulation were performed using the Amber 9 [149] suite. The force field used for the protein was ff03 [150], gaff [151] for the ligands and TIP3P [152] model for the waters. Each complex Abl-inhibitor was neutralized by adding an appropriate number of sodium counter-ions and solvated in a rectangular box of pre-equilibrated waters molecules with a minimum solute-wall distance of 10 Å. First, the protein-ligand was frozen and solvent waters molecules with counter-ions were allowed to move during a 3000-steps energy minimization. After full relaxation of the entire system, the temperature was increased from 10 to 300 K in 35 ps Langevin dynamic at constant volume, followed by 200 ps Langevin dynamics at constant temperature and pressure. Finally, the production molecular dynamics (MD) simulation was performed for 1 ns with a periodic boundary condition in the NPT ensemble at T=300 K with Berendsen temperature coupling and a constant pressure (P=1 atm) with isotropic

molecule-based scaling. The SHAKE algorithm was used to constrain bonds involving hydrogen while the particle mesh Ewald (PME) method was applied to treat long-range electrostatic interactions with a residue-based cutoff of 10 Å.

Binding Free Energy Calculation

The binding free energies were calculated using the MM-PB/GBSA free energy calculation method as detailed in ref. 1. In this computational technique, the free energy of the inhibitor binding, ΔG_{bind}, is obtained from the difference between the free energy of the receptor-ligand complex (G_{cpx}), and the unbound receptor (G_{rec}) and ligand (G_{lig}) according to Equation 3.1.

$$\Delta G_{bind} = G_{cpx} - (G_{rec} + G_{lig}) \tag{3.1}$$

The binding free energy (ΔG_{bind}) was evaluated in 1000 molecular dynamics snapshots taken at 10 ps intervals, as a sum of changes in the energy of three different contributions:

1. a force field term (E_{FF}) for bond, angle, torsional, van der Waals, and electrostatic potential energies;

2. a polar solvation free energy part (ΔG_{PB}), calculated by the finite-difference solution to the Poisson-Boltzmann equation as implemented in the program pbsa or using the Generalized Born implicit solvent;

3. a non-polar contribution to the solvation free energy (ΔG_{NP})

$$\Delta G_{bind} = E_{FF} + \Delta G_{PB} + \Delta G_{NP} \tag{3.2}$$

$$\Delta G_{NP} = \alpha SASA + \beta \tag{3.3}$$

SASA represents the solvent accessible surface area of the solute, while α and β are two empirical parameters, 0.92 kcal/mol and 0.00542 kcal/Å^2 respectively if the MM-PBSA is used and α=0.0072 and β=0 if the Generalized Born implicit solvent is present. The molecular similarity of STI-571 and CGP-582, together with the fact that the difference of activity was investigated, allows to not include in this study the entropy contribution to the binding free energy.

3.2.2   Results

Chemical-physical analysis of molecular dynamics simulation showed that the protocol chosen was enough to reach a equilibrate state for STI-571 and both the binding mode of CGP-582 studied. As shown in Figure 3.3, energy, temperature, pression, volume, density were stable in the period used for the MM-PBSA calculation. For every snapshots the kinetic energy plus the potential energy gives the total energy, the temperature fluctuation during the Langevian dynamics are slightly stronger than when the Berendsen thermostat is used as expected. Correctly after the first part at constant volume the system reach an equilibrium with a density near 1 g/cm^3, the density of pure water, and the volume changes during time are the specular image of density fluctuation. Pression variation are stronger, but the mean value is comparable to the atmospheric pression. The system, and in particular the ligand, RMSD reaches soon a low fluctuation. The inhibitors gaff parameterization was satisfactory and ligands movement during the MD were plausible.

However the MM-PB/GBSA energy estimation range during the 1 ns MD production phase was of about 25 kcal/mol as show in Figure 3.4. This variation does not allow an accurate prediction of the free energy of binding difference between the two inhibitors in study. This high window of energy fluctuation derives from both the force field energy term and from the solvation energy estimation and it is linked to several possible reasons:

- First of all snapshots derive from an explicit solvent MD, while for the binding energy estimation we use implicit solvents. For this reason the conformations obtained could be not optimal. Anyway recent studies showed that a snapshot energy minimization in implicit solvent does not always result in an improvement of the correspondence of the prediction and the experimental free energy of binding [153].

- One principal issue is the strong energy dependence of the solvation energy from small conformational changes of inhibitors and of the ATP binding pocket.

- Another major problematic is the correct evaluation of the protein dielectric constant used in the implicit solvent parameterization.

- Difficulties of a correct evaluation of the electrostatic energy resulting from charged ligands (like both STI-571 and CGP-582) can also contribute to errors in the final energy estimation.

This fluctuation of the energy estimation creates a problem also in the snapshots selection for the ΔG_{bind} evaluation. The different approaches generally utilized are the use of all the 1000 snapshots, or a snapshots selection from all the trajectory or from the last part of the MD. These possibilities in this case would result in important difference in the final results as summarized in Figure 3.2.

snapshots	STI-571	CGP-582 A	CGP-582 B
0-100	-66.40	-59.20	-69.02
100-200	-67.61	-59.24	-64.67
200-300	-68,94	-61.05	-67.16
300-400	-68.14	-64.68	-62.31
400-500	-58.47	-62.39	-61.39
500-600	-62.32	-67.12	-65.60
600-700	-63.33	-65.22	-64.90
700-800	-65.41	-65.62	-66.23
800-900	-61.65	-64.02	-67.89
900-100	-64.41	-67.20	-66.65
0-1000	-64.67	-63.58	-65.56

Figure 3.2: Binding energy dependence from snapshot selection. On the first column the snapshot interval is shown and on the others the prediction energy in kcal/mol is indicated.
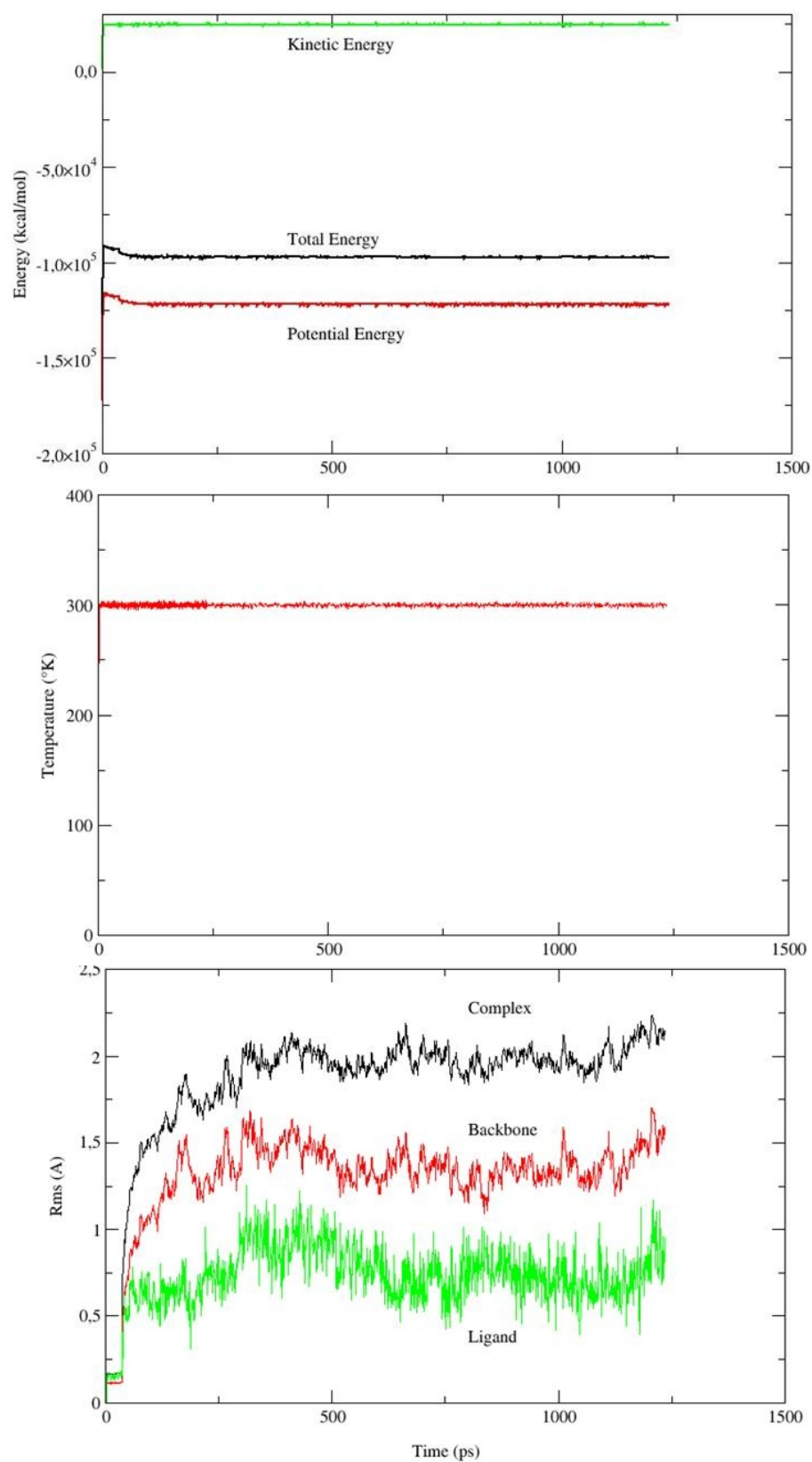
Figure 3.3: STI-571 MD equilibration.
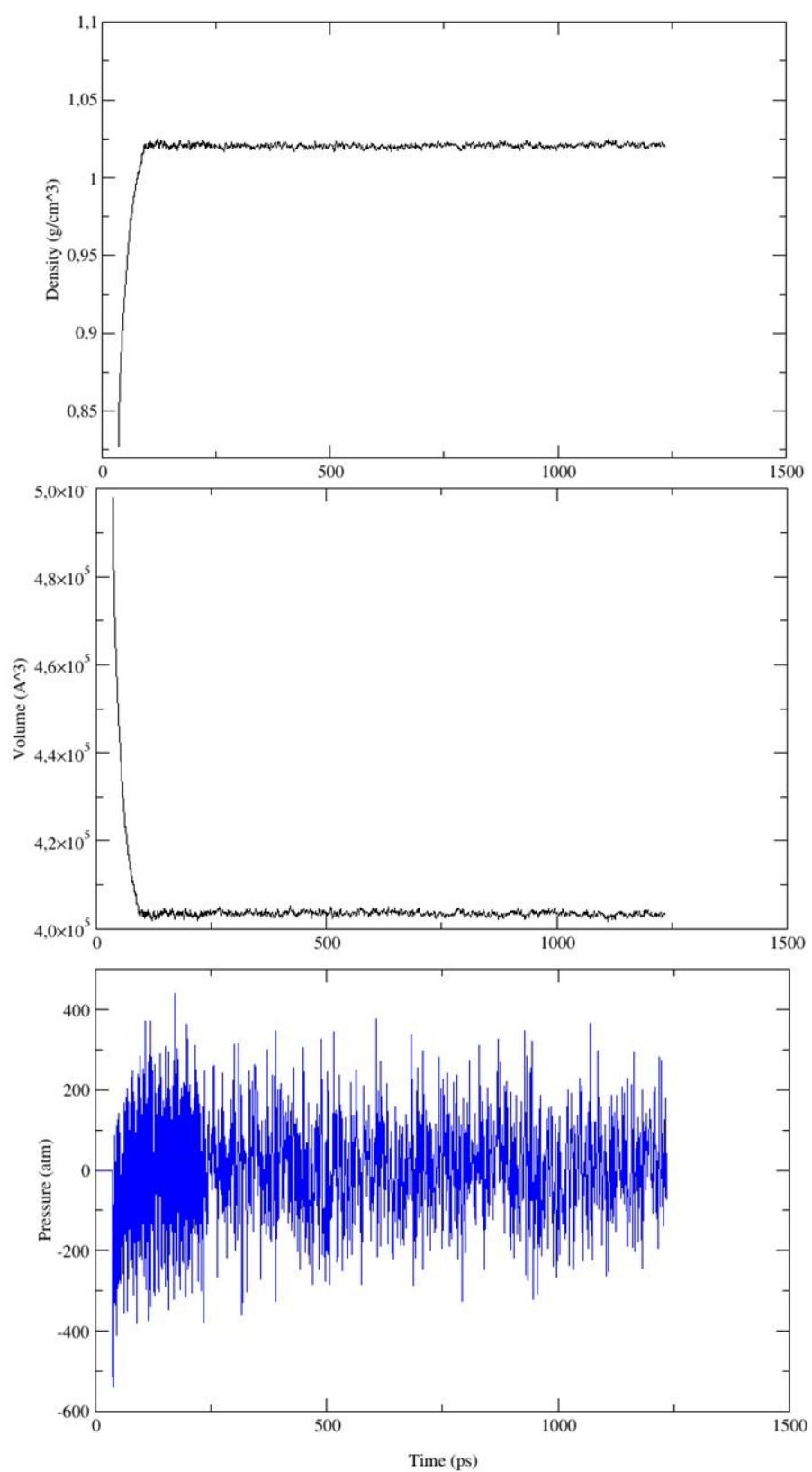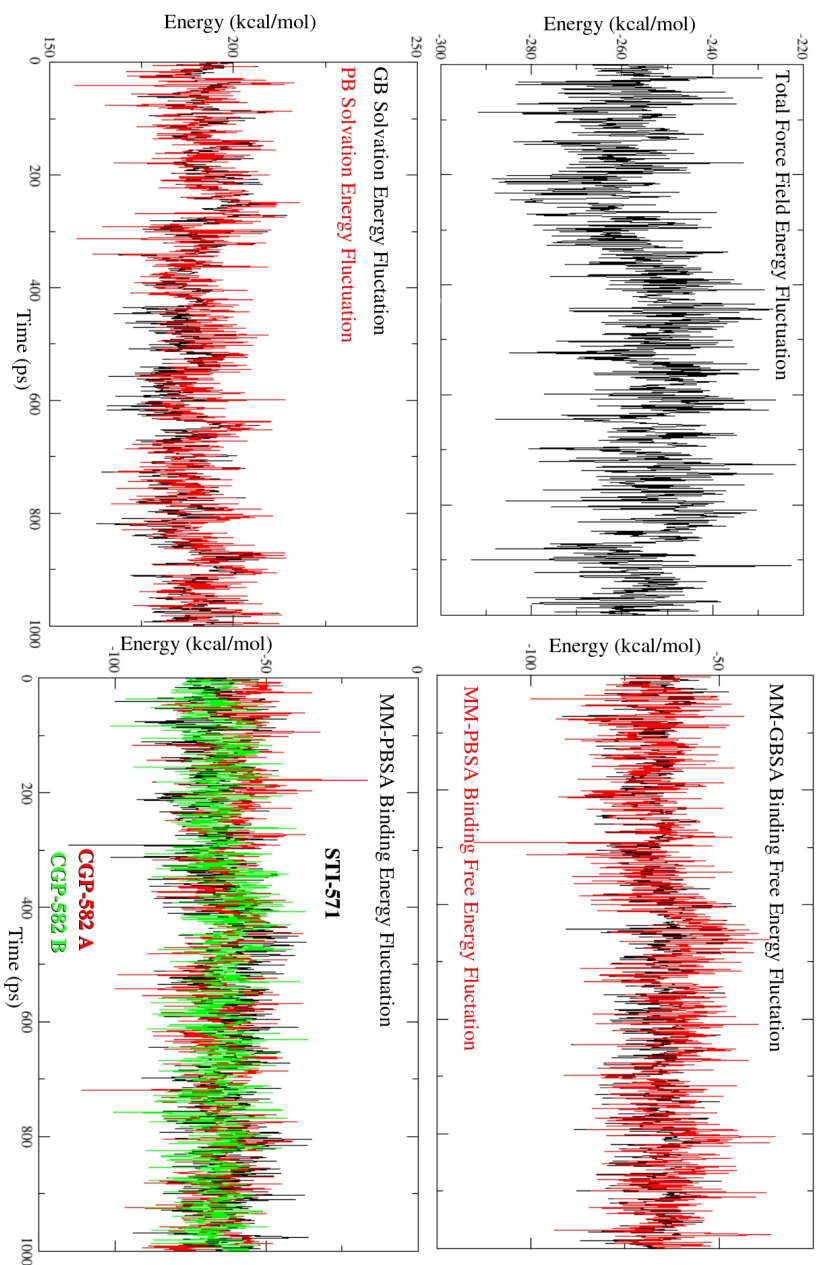
Figure 3.4: MM-PB/GBSA study energy fluctuation. During the equilibrated MD there is a energy fluctuation of about 25 kcal/mol in the total force field energy, in solvation energy and binding energy for both the Poisson-Boltzmann and Generalized Born method. Such fluctuation does not allow to distinguish the experimental activity difference of about 2.4 kcal/mol between STI-571 and CGP-582.

3.3 Colony Energy Approach

We tried to increase the robustness of the MM-PB/GBSA method using the recently proposed method by Honig and co-workers [154] based on the Colony Energy approach. This technique tries to reduce the possible large changes in various energy terms e.g., covalent, van der Waals, and electrostatic terms resulting from small changes in atomic coordinates (e.g., of sidechain atoms). It has been developed to evaluate conformational free energies and tested to predict loop conformation [155] and in this study it is for the first time used to predict and compare ligands free energy of binding.

3.3.1 Methods

In the Colony Energy approach each conformation is considered a sample of a colony of similar conformations. For this reason the collections of snapshots deriving from the MD study is analyzed, and all the poses are compared to find the colony they belong to. It is assumed that the conformational space is sampled in a statistically significant way and the following equation is applied to evaluate the free energy corresponding to the colony:

$$\Delta G_i^C = -kTln\frac{\sum_j exp\left(\frac{-W(r_1^j,\cdots,r_n^j)}{kT} - \frac{rmsd_{ij}^3}{t^3}\right)}{\sum_j exp\left(\frac{-W(r_1^j,\cdots,r_n^j)}{kT}\right)} \qquad (3.4)$$

i is index of the reference structure and j the index of each other sampled conformation, W rapresents the free energy of binding estimated using the MM-PBSA or MM-GBSA approach. The t value is a threshold RMSD: if the RMSD between the reference conformation and the sampled conformation is under this value, they are considered part of the same colony.

t is an important value discriminating related and unrelated poses and it effectively smoothes out the force-field potential, lowering high-energy conformations that are close in structure to low-energy conformations. In this way the strong dependence of the estimated ΔG on small coordinate variations is sensibily reduced [155]. It broads the energy surface when many nearby state are detected, but on the other hand high conformations energy are not discarted.

The final Colony Energy calculated can be considered a weighted mean of the energy of all conformations that take in account the pose similarity (indicated as RMSD) to the other colony members.

3.3.2   Results

The threshold parameter t is the RMSD value below which two poses are considerated members of the same colony. From a practical point of view the higher is t the stronger will be the smoothing effect, if it is too high all conformations are considered of the same colony and the energy of all is flattened at the level of the best one (lowest free energy of binding) and the plot will display a single horizontal line. It is possible to see this effect in Figure 3.5. In this study we decide to evaluate the RMSD of the ligand and the aminoacids in a radius of 5Å from it and we tested two t values: 0.3 and 0.5Å. These are in the range of the active site fluctuations during the MD.

The Colony Energy changed the energy fluctuation window, removing the noise. The standard deviation of the ΔG_{bind} improved from about 7 kcal/mol for MM-PB/GBSA to 4 and 2 kcal/mol for MM-PB/GBSA-C with a t of 0.3 or 0.5 Å respectively (Fig. 3.6). This lower error lets us to be more confident about the energy evaluation.

The error of the prediction remains anyway a complex issue. From a statistical point of view the error depends on the number of the snapshots used for the energy evaluation and can be calculated in the following way:

$$\sigma_{final} = \frac{\sigma}{\sqrt{N}} \tag{3.5}$$

Where N is the snapshots number. In our case using 1000 conformations per ligand the error results very low. However in the scientific community the general applicability of this statistical approach to this particular case is not completely accepted: it can be plausible to evaluate the statistical error of the ensemble, but not the ones linked to the force field, and to the continuum solvent model itself.

The Colony Energy approach does not only reduces the standard deviation of the free energy of binding prediction, but also change the resulting relative power of inhibitors in study. As we can see from Figure 3.7, the MM-PB/GBSA binding energy for CGP-582 in the binding mode B was evaluated stronger than STI-571 in opposition with the experimental data. On the contrary after the Colony Energy approach Imatinib resulted always the more powerful and the experimental free energy of binding difference is interesting always between the one calculated for the binding mode A and B. The Colony Energy approach changes in this way the wrong MM-PB/GBSA prediction in a correct and precise evaluation of the ΔG_{bind}.
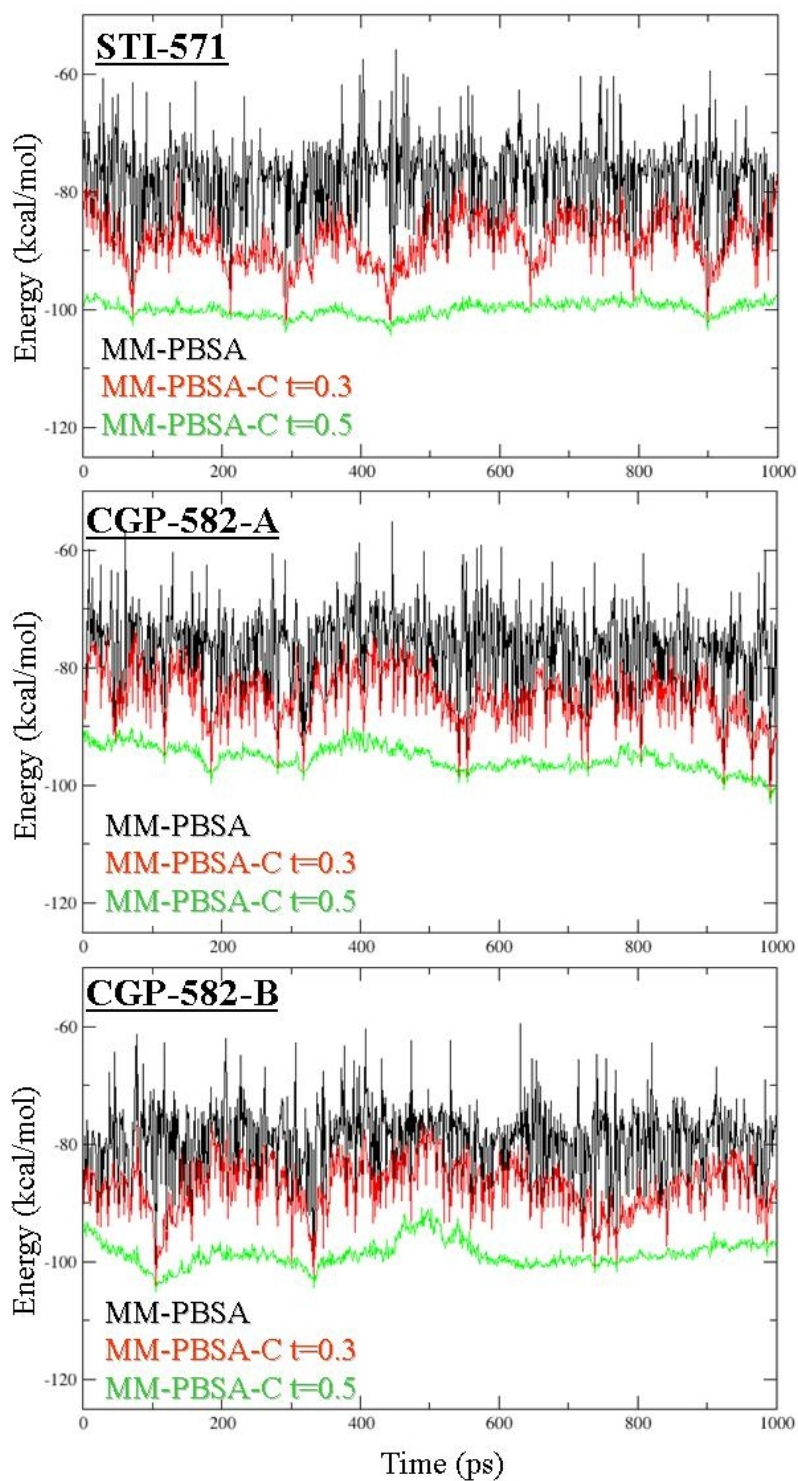
Figure 3.5: Smoothing effect on the MM-PBSA binding energy of the Colony Energy approach for the three inhibitors depending on the value of t indicated in Å.
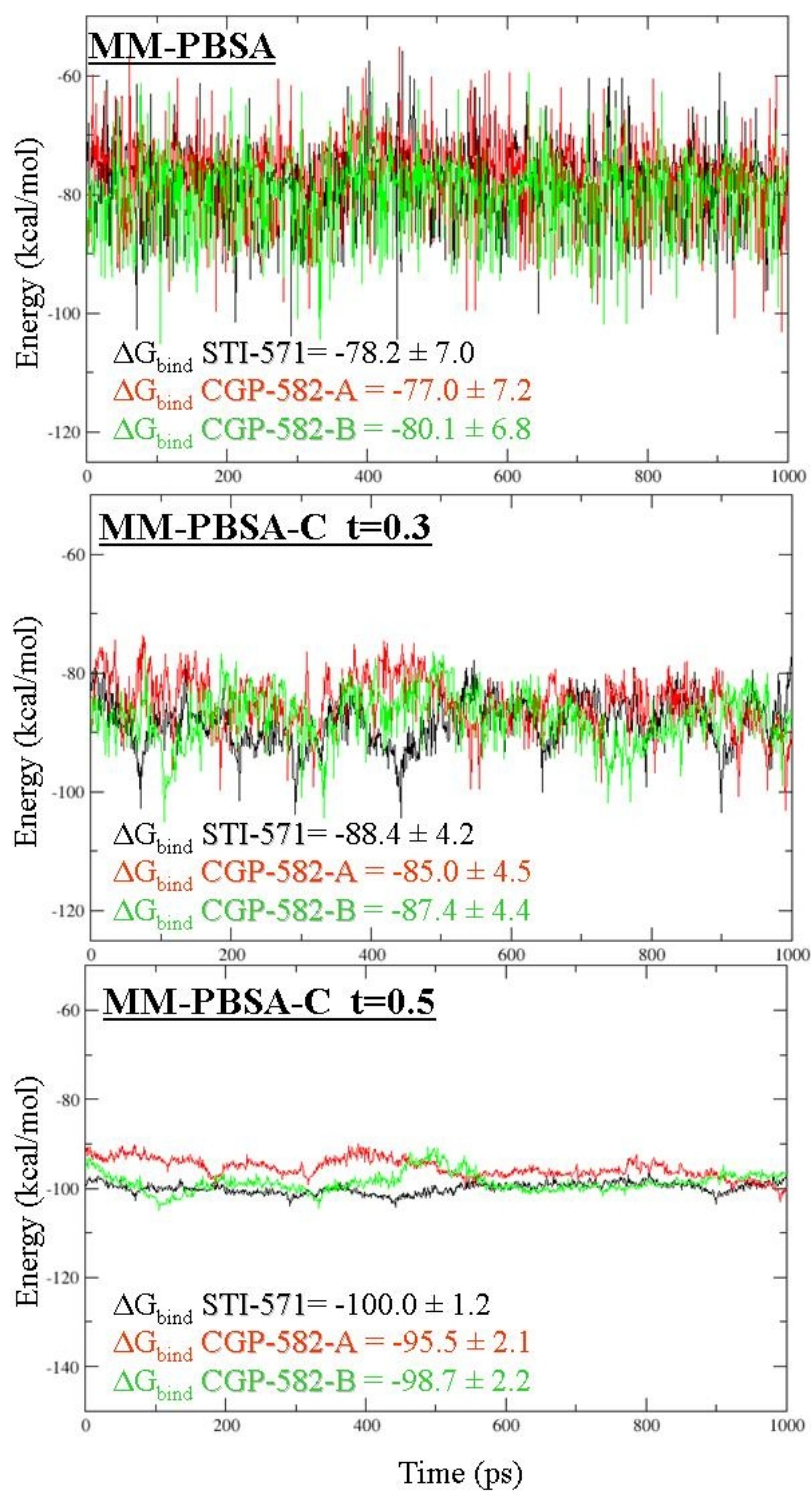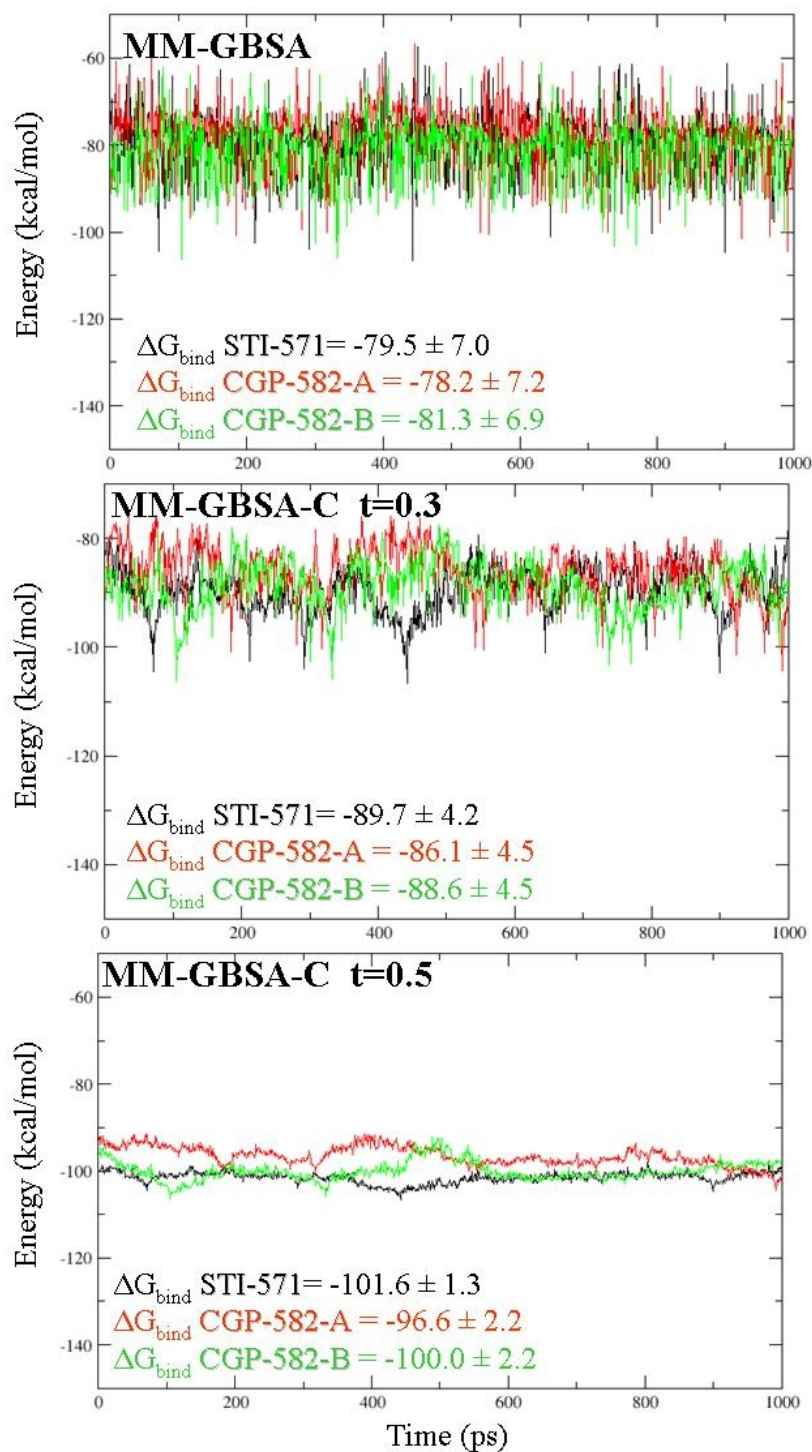
Figure 3.6:   Energy fluctuation differences between MM-PB/GBSA and MM-PB/GBSA-C for STI-571 and the two CGP-582 binding modes. In the Colony Energy approach two different possible t value have been used, t=0.3 and t=0.5 Å. The binding free energy results using all the 1000 snapshots and their standard deviation is indicated.

ΔΔG$_{bind}$(kcal/mol) STI-571-CGP-582		
	CGP-582 binding mode A	CGP-582 binding mode B
MM-PBSA	-1.2	+1.9
MM-PBSA-C t=0.3	-3.4	-1.0
MM-PBSA-C t=0.5	-4.5	-1.3
MM-GBSA	-1.3	+1.8
MM-GBSA-C t=0.3	-3.6	-1.1
MM-GBSA-C t=0.5	-5.0	-1.6
Experimental	-2.4	

Figure 3.7: Final MM-PB/GBSA and MM-PB/GBSA-C ΔΔG$_{bind}$.

3.4 Conclusions

The difficulties to understand the loss of activities of CGP-582 for Abl are based on the absence of a evident strong unfavorable interaction in the binding pocket to justify one fold activity difference from STI-571. In this case where docking studies and qualitative analysis cannot help, the MM-PB/GBSA technique in concert with a Colony Energy post-processing approach has demonstrated to be an efficient tool to understand the chemical insight of the inhibitor binding. The complex formation in this case is the result of several multiple small energies contributes that to be fully understood must be evaluated from the point of view of the equilibrium and the events dynamics. In this contest the Colony Energy theory helps to evaluate correctly, removing the basal noise, smoothing the energy factors on the basis of a conformational similarity coherence.

Bibliography

[1] G. Manning, D. B. Whyte, R. Martinez, T. Hunter, and S. Sudarsaxnam. The protein kinase complement of human genome. *Science*, **298**:1912–1934, 2002.

[2] D. R. Robinson, Yi-Mi Wu, and Su-Fang Lin. The protein tyrosine kinase family of the human genome. *Oncogene*, **19**:5548–5557, 2000.

[3] R. Roskoski Jr. STI-571: an anticancer protein-tyrosine kinase inhibitor. *Biochemical and Biophysical Research Communications*, **309**:709–717, 2003.

[4] R. P. Araujo, L. A. Liotta, and E. F.Petricoin. Proteins, drug targets and the mechanisms they control: the simple truth about complex networks. *Nature Reviews Drug Discovery*, **6**:871–880, 2007.

[5] C. Ortutay, J. Väliaho, K. Stenberg, and M. Vihinen. KinMutBase: a registry of disease-causing mutations in protein kinase domains. *Human Mutation*, **25**:435–442, 2005.

[6] R. P. Bhattacharyya, A. Reményi, B. J. Yeh, and W. A. Lim. Domains, motifs, and scaffolds: the role of modular interactions in the evolution and wiring of cell signaling circuits. *Annual Review Biochemistry*, **75**:655–80, 2006.

[7] W. T. Miller. Determinants of substrate recognition in nonreceptor tyrosine kinases. *Accounts of Chemical Research*, **36**:393–400, 2003.

[8] T. Pawson and P. Nash. Protein-protein interactions define specificity in signal transduction. *Genes & Development*, **14**:1027–1047, 2000.

[9] R. M. Biondi and A. R. Nebreda. Signalling specificity of Ser/Thr protein kinases through docking-site-mediated interactions. *Biochemical Journal*, **372**:1–13, 2003.

[10] T. Pawson and J. D. Scott. Signaling through scaffold, anchoring, and adaptor proteins. *Science*, **278**:2075–2080, 1997.

[11] S. R. Hubbard. Protein tyrosine kinase: autoregulation and small-molecule inhibition. *Current Opinion in Structural Biology*, **12**:735–741, 2002.

[12] S. R. Hubbard. Juxtamembrane autoinhibition in receptor tyrosine kinase. *Nature Reviews Molecular Cell Biology.*, **5**:464–470, 2004.

[13] S. F. Kingsmore. Multiplexed protein measurement: technologies and applications of protein and antibody arrays. *Nature Review Drug Discovery*, **5**:310–320, 2006.

[14] K. Lundstrom. Latest development in drug discovery on G protein-coupled receptors. *Current Protein and Peptide Science*, **7**:465–470, 2006.

[15] N. Steeghs, J. W. Nortier, and H. Gelderblom. Small molecule tyrosine kinase inhibitors in the treatment of solid tumors: an update of recent developments. *Annals of Surgical Oncology*, **14**:942–953, 2007.

[16] B. Turk. Targeting proteases: successes, failures and future prospects. *Nature Review Drug Discovery*, **5**:785–799, 2006.

[17] J. Couzin. Cancer drugs. smart weapons prove tough to design. *Science*, **298**:522–525, 2002.

[18] T. J. Lynch, D. W. Bell, R. Sordella, S. Gurubhagavatula, R. A. Okimoto, B. W. Brannigan, P. L Harris, S. M Haserlat, J. G. Supko, F. G. Haluska, D. N. Louis, D. C. Christiani, J. Settleman, and D. A. Haber. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to Gefitinib. *New England Journal*, **350**:2129–2139, 2004.

[19] I. Kola and J. Landis. Can the pharmaceutical industry reduce attrition rates? *Nature Review Drug Discovery*, **3**:711–715, 2004.

[20] T. G. Roberts Jr., B. H. Goulart, L. Squitieri, S. C. Stallings, E. F. Halpern, B. A. Chabner, G. S. Gazelle, S. N. Finkelstein, and J. W. Clark. Trends in the risks and benefits to patients with cancer participating in phase 1 clinical trials. *The Journal of American Medical Association*, **292**:2130–2140, 2004.

[21] J. Dancey and E. A. Sausville. Issues and progress with protein kinase inhibitors for cancer treatment. *Nature Review Drug Discovery*, **2**:296–313, 2003.

[22] T. Lahaye, B. Riehm, U. Berger, P. Paschka, M. C. Müller, S. Kreil, K. Merx, U. Schwindel, C. Schoch, R. Hehlmann, and A. Hochhaus. Response and resistance in 300 patients with BCR-ABL-positive leukemias treated with Imatinib in a single center: a 4.5-year follow-up. *Cancer*, **103**:1659–69, 2005.

[23] H. Daub, K. Specht, and A. Ullrich. Strategies to overcome resistance to targeted protein kinase inhibitors. *Nature Review Drug Discovery*, **3**(12):1001–10, 2004.

[24] J. Litz, G. S. Warshamana-Greene, G. Sulanke, K. E. Lipson, and G. W. Krystal. The multi-targeted kinase inhibitor SU5416 inhibits small cell lung cancer growth and angiogenesis, in part by blocking Kit-mediated VEGF expression. *Lung Cancer*, **46**:283–91, 2004.

[25] M. S. Saund, G. D. Demetri, and S. W. Ashley. Gastrointestinal stromal tumors (GISTs). *Current Opinion Gastroenterology*, **20**:89–94, 2004.

[26] A. M. S. Müller, U. M. Martens, S. C. Hofmann, L. Bruckner-Tuderman, R. Mertelsmann, and M. Lübbert. Imatinib mesylate as a novel treatment option for hypereosinophilic syndrome: two case reports and a comprehensive review of the literature. *Annual Hematology*, **85**:1–16, 2006.

[27] M. E. K. Olsten and D. W. Litchfield. Order or chaos? an evaluation of the regulation of protein kinase CK2. *Biochemistry and Cell Biology*, **82**:681–693, 2004.

[28] D. W. Litchfield. Protein kinase CK2: structure, regulation and role in cellular decisions of life and death. *Biochemical Journal*, **369**:1–15, 2003.

[29] S. K. Hanks and T. Hunter. Protein kinases 6. the eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB Journal*, **9**:576–596, 1995.

[30] L. K. Wilson, N. Dhillon, J. Thorner, and G. S. Martin. Casein kinase II catalyzes tyrosine phosphorylation of the yeast nucleolar immunophilin Fpr3. *Journal of Biological Chemistry*, **272**:12961–12967, 1997.

[31] M. Lasa, O. Marin, and L. A. Pinna. Rat liver Golgi apparatus contains a protein kinase similar to casein kinase of lactating mammary gland. *European Journal of Biochemistry*, **243**:719–725, 1997.

[32] R. D. Gietz, K. C. Graham, and D. W. Litchfield. Interactions between the subunits of casein kinase II. *Journal Biological Chemistry*, **270**:13017–13021, 1995.

[33] L. A. Pinna and F. Meggio. Protein kinase CK2 ("casein kinase-2") and its implication in cell division and proliferation. *Progress in Cell Cycle Research*, **3**:77–97, 1997.

[34] K. C. Graham and D. W. Litchfield. The regulatory beta subunit of protein kinase CK2 mediates formation of tetrameric CK2 complexes. *Journal Biological Chemistry*, **275**:5003–5010, 2000.

[35] F. Meggio, B. Boldyreff, O. Marin, L. A. Pinna, and O. G. Issinger. Role of the beta subunit of casein kinase-2 on the stability and specificity of the recombinant reconstituted holoenzyme. *European Journal of Biochemistry*, **204**:293–297, 1992.

[36] P. Agostinis, J. Goris, L. A. Pinna, and W. Merlevede. Regulation of casein kinase 2 by phosphorylation/dephosphorylation. *Biochemical Journal*, **248**:785–789, 1987.

[37] D. W. Litchfield, F. J. Lozeman, M. F. Cicirelli, M. Harrylock, L. H. Ericsson, C. J. Piening, and E. G. Krebs. Phosphorylation of the beta subunit of casein kinase II in human A431 cells. identification of the autophosphorylation site and a site phosphorylated by p34cdc2. *Journal of Biological Chemistry*, **266**:20380–20389, 1991.

[38] C. S. Skjerpen, T. Nilsen, J. Wesche, and S. Olsnes. Binding of FGF-1 variants to protein kinase CK2 correlates with mitogenicity. *EMBO Journal*, **21**:4058–4069, 2002.

[39] M. Faust, N. Schuster, and M. Montenarh. Specific binding of protein kinase CK2 catalytic subunits to tubulin. *FEBS Letters*, **462**:51–56, 1999.

[40] Y. Miyata and I. Yahara. Interaction between casein kinase II and the 90-kDa stress protein, HSP90. *Biochemistry*, **34**:8123–8129, 1995.

[41] P. T.Tuazon and J. A. Traugh. Casein kinase I and II–multipotential serine protein kinases: structure, function, and regulation. *Advances in Second Messenger and Phosphoprotein Research*, **23**:123–164, 1991.

[42] N. P. Pavletich. Mechanism of cyclin-dependent kinase regulation: structure of Cdks, their cyclin activators, and Cip and INK4 inhibitors. *Journal of Molecular Biology*, **287**:821–828, 1999.

[43] U. Münstermann, G. Fritz, G. Seitz, Y. P. Lu, H. R. Schneider, and O. G. Issinger. Casein kinase II is elevated in solid human tumours and rapidly proliferating non-neoplastic tissue. *European Journal of Biochemistry*, **189**:251–257, 1990.

[44] L. A. Pinna. Casein kinase 2: an "eminence grise" in cellular regulation? *Biochimica et Biophysica Acta*, **1054**:267–284, 1990.

[45] M. Faust and M. Montenarh. Subcellular localization of protein kinase CK2. a key to its function? *Cell Tissue Research*, **301**:329–340, 2000.

[46] C. V. Glover. On the physiological role of casein kinase II in Saccharomyces cerevisiae. *Progress in Nucleic Acid Research and Molecular Biology*, **59**:95–133, 1998.

[47] J. S. Duncan and D. W. Litchfield. Too much of a good thing: The role of protein kinase CK2 in tumorigenesis and prospects for therapeutic inhibition of CK2. *Biochimica et Biophysica Acta*, 2007.

[48] S. Desagher, A. Osen-Sand, S. Montessuit, E. Magnenat, F. Vilbois, A. Hochmann, L. Journot, B. Antonsson, and J. C. Martinou. Phosphorylation of bid by casein kinases I and II regulates its cleavage by caspase 8. *Molecular Cell*, **8**:601–611, 2001.

[49] G. L. Russo, M. T. Vandenberg, I. J. Yu, Y. S. Bae, B. R. Franza, and D. R. Marshak. Casein kinase II phosphorylates p34cdc2 kinase in G1 phase of the HeLa cell division cycle. *Journal of Biological Chemistry*, **267**:20317–20325, 1992.

[50] K. Block, T. G. Boyer, and P. R. Yew. Phosphorylation of the human ubiquitin-conjugating enzyme, CDC34, by casein kinase 2. *Journal of Biological Chemistry*, **276**:41049–41058, 2001.

[51] J. R. Daum and G. J. Gorbsky. Casein kinase II catalyzes a mitotic phosphorylation on threonine 1342 of human DNA topoisomerase IIalpha, which is recognized by the 3F3/2 phosphoepitope antibody. *Journal Biological Chemistry*, **273**:30622–30629, 1998.

[52] R. A. Faust, S. Tawfic, A. T. Davis, L. A. Bubash, and K. Ahmed. Antisense oligonucleotides against protein kinase CK2-alpha inhibit growth

of squamous cell carcinoma of the head and neck in vitro. *Head Neck*, **22**:341–346, 2000.

[53] S. Yenice, A. T. Davis, S. A. Goueli, A. Akdas, C. Limas, and K. Ahmed. Nuclear casein kinase 2 (CK-2) activity in human normal, benign hyperplastic, and cancerous prostate. *Prostate*, **24**:11–16, 1994.

[54] G. Stalter, S. Siemer, E. Becht, M. Ziegler, K. Remberger, and O. G. Issinger. Asymmetric expression of protein kinase CK2 subunits in human kidney tumors. *Biochemical and Biophysical Research Communications*, **202**:141–147, 1994.

[55] E. Landesman-Bollag, R. Romieu-Mourez, D. H. Song, G. E. Sonenshein, R. D. Cardiff, and D. C. Seldin. Protein kinase CK2 in mammary gland tumorigenesis. *Oncogene*, **20**:3247–3257, 2001.

[56] M. Daya-Makin, J. S. Sanghera, T. L. Mogentale, M. Lipp, J. Parchomchuk, J. C. Hogg, and S. L.Pelech. Activation of a tumor-associated protein kinase (p40TAK) and casein kinase 2 in human squamous cell carcinomas and adenocarcinomas of the lung. *Cancer Research*, **54**:2262–2268, 1994.

[57] D. C. Seldin, E. Landesman-Bollag, M. Farago, N. Currier, D. Lou, and I. Dominguez. CK2 as a positive regulator of Wnt signalling and tumourigenesis. *Molecular Cell Biochemistry*, **274**:63–67, 2005.

[58] S. F. Eddy, S. Guo, E. G. Demicco, R. Romieu-Mourez, E. Landesman-Bollag, D. C. Seldin, and G. E. Sonenshein. Inducible IkappaB kinase/IkappaB kinase epsilon expression is induced by CK2 and promotes aberrant nuclear factor-kappaB activation in breast cancer cells. *Cancer Research*, **65**:11375–11383, 2005.

[59] K. Busset, M. Henriksson, J. M. Luscher-Firlaff, and B. Luscher. Identification of casein kinase II phosphorylation sites in max: effectson DNA-binding kinetics of max homo and Myc/Max. *Oncogene*, **8**:3211–3220, 1992.

[60] A. Lin, J. Frost, T. Deng, T. Smeal, N. al Alawi, U. Kikkawa, T. Hunter, D. Brenner, and M. Karin. Casein kinase II is a negative regulator of c-Jun DNA binding and AP-1 activity. *Cell*, **70**:777–789, 1992.

[61] P. Channavajhala and D. C. Seldin. Functional interaction of protein kinase CK2 and c-Myc in lymphomagenesis. *Oncogene*, **21**:5280–5288, 2002.

[62] M. Oelgeschläger, J. Krieg, J. M. Lüscher-Firzlaff, and B. Lüscher. Casein kinase II phosphorylation site mutations in c-Myb affect DNA binding and transcriptional cooperativity with NF-M. *Molecular Cell Biology*, **15**:5966–5974, 1995.

[63] J. Torres and R. Pulido. The tumor suppressor PTEN is phosphorylated by the protein kinase CK2 at its C terminus. implications for PTEN stability to proteasome-mediated degradation. *Journal Biological Chemistry*, **276**:993–998, 2001.

[64] A. M. Oliveira, J. S. Ross, and J. A. Fletcher. Tumor suppressor genes in breast cancer: the gatekeepers and the caretakers. *American Journal of Clincal Pathology*, **124 Suppl**:S16–S28, 2005.

[65] M. Kapoor and G. Lozano. Functional activation of p53 via phosphorylation following DNA damage by UV but not gamma radiation. *Proceedings of the National Academy of Sciences of the United States of America*, **95**:2834–2837, 1998.

[66] M. Sayed, S. O. Kim, B. S. Salh, O. G. Issinger, and S. L. Pelech. Stress-induced activation of protein kinase CK2 by direct interaction with p38 mitogen-activated protein kinase. *Journal of Biological Chemistry*, **275**:16569–16573, 2000.

[67] P. P. Scaglioni, T. M. Yung, L. F. Cai, H. Erdjument-Bromage, A. J. Kaufman, B. Singh, J. Teruya-Feldstein, P. Tempst, and P. P. Pandolfi. A CK2-dependent mechanism for degradation of the PML tumor suppressor. *Cell*, **126**:269–283, 2006.

[68] L. A. Pinna and F. Meggio. Protein kinase CK2 ("casein kinase-2") and its implication in cell division and proliferation. *Progress in Cell Cycle Research*, **3**:77–97, 1997.

[69] R. Zandomeni, M. C. Zandomeni, D. Shugar, and R. Weinmann. Casein kinase type II is involved in the inhibition by 5,6-dichloro-1-beta-D-ribofuranosylbenzimidazole of specific RNA polymerase II transcription. *Journal of Biological Chemistry*, **261**:3414–3419, 1986.

[70] F. Meggio, D. Shugar, and L. A . Pinna. Ribofuranosyl-benzimidazole derivatives as inhibitors of casein kinase-2 and casein kinase-1. *European Journal of Biochemistry*, **187**:89–94, 1990.

[71] M. Andrzejewska, M. A. Pagano, F. Meggio, A. M. Brunati, and Z. Kazimierczuk. Polyhalogenobenzimidazoles: synthesis and their inhibitory activity against casein kinases. *Bioorganic and Medicinal Chemistry*, **11**:3997–4002, 2003.

[72] M. A. Pagano, M. Andrzejewska, M. Ruzzene, S. Sarno, L. Cesaro, J. Bain, M. Elliott, F. Meggio, Z. Kazimierczuk, and L. A. Pinna. Optimization of protein kinase CK2 inhibitors derived from 4,5,6,7-tetrabromobenzimidazole. *Journal of Medicinal Chemistry*, **47**:6239–6247, 2004.

[73] R. Battistutta, E. Moliner, S. Sarno, G. Zanotti, and L. A. Pinna. Structural features underlying selective inhibition of protein kinase CK2 by ATP site-directed tetrabromo-2-benzotriazole. *Protein Science*, **10**:2200–2206, 2001.

[74] Roberto Battistutta, Marco Mazzorana, Stefania Sarno, Zygmunt Kazimierczuk, Giuseppe Zanotti, and Lorenzo A Pinna. Inspecting the structure-activity relationship of protein kinase CK2 inhibitors derived from tetrabromo-benzimidazole. *Chemistry and Biology*, **12**:1211–1219, 2005.

[75] R. Battistutta, M. Mazzorana, L. Cendron, A. Bortolato, S. Sarno, Z. Kazimierczuk, G. Zanotti, S. Moro, and L. A. Pinna. The atp-binding site of protein kinase CK2 holds a positive electrostatic area and conserved water molecules. *Chembiochem*, **8**:1804–1809, 2007.

[76] S. Sarno, H. Reddy, F. Meggio, M. Ruzzene, S. P. Davies, A. Donella-Deana, D. Shugar, and L. A. Pinna. Selectivity of 4,5,6,7-tetrabromobenzotriazole, an ATP site-directed inhibitor of protein kinase CK2 ('casein kinase-2'). *FEBS Letters*, **496**:44–48, 2001.

[77] M. A. Pagano, F. Meggio, M. Ruzzene, M. Andrzejewska, Z. Kazimierczuk, and L. A. Pinna. 2-Dimethylamino-4,5,6,7-tetrabromo-1H-benzimidazole: a novel powerful and selective inhibitor of protein kinase CK2. *Biochemical and Biophysical Research Communications*, **321**:1040–1044, 2004.

[78] A. G. Golub, O. Y. Yakovenko, A. O. Prykhod'ko, S. S. Lukashov, V. G. Bdzhola, and S. M. Yarmoluk. Evaluation of 4,5,6,7-tetrahalogeno-1h-isoindole-1,3(2H)-diones as inhibitors of human protein kinase CK2. *Biochim Biophys Acta*, 2007.

[79] M. A. Pagano, G. Poletto, G. Di Maira, G. Cozza, M. Ruzzene, S. Sarno, J. Bain, M. Elliott, S. Moro, G. Zagotto, F. Meggio, and L. A. Pinna. Tetrabromocinnamic acid (TBCA) and related compounds represent a new class of specific protein kinase CK2 inhibitors. *Chembiochem*, **8**:129–139, 2007.

[80] S. Sarno, E. Moliner, M. Ruzzene, M. A. Pagano, R. Battistutta, J. Bain, D. Fabbro, J. Schoepfer, M. Elliott, P.Furet, F. Meggio, G. Zanotti, and L. A. Pinna. Biochemical and three-dimensional-structural study of the specific inhibition of protein kinase CK2 by [5-oxo-5,6-dihydroindolo-(1,2-a)quinazolin-7-yl]acetic acid (IQA). *Biochemical Journal*, **374**:639–646.

[81] R. Battistutta, S. Sarno, E. Moliner, E. Papinutto, G. Zanotti, and L. A. Pinna. The replacement of ATP by the competitive inhibitor emodin induces conformational modifications in the catalytic site of protein kinase CK2. *Journal of Biological Chemistry*, **275**:29618–29622, 2000.

[82] E. Moliner, S. Moro, S. Sarno, G. Zagotto, G. Zanotti, L. A. Pinna, and R. Battistutta. Inhibition of protein kinase CK2 by anthraquinone-related compounds. a structural insight. *Journal of Biological Chemistry*, **278**:1831–1836, 2003.

[83] S. Sarno, S. Moro, F. Meggio, G. Zagotto, D. Dal Ben, P. Ghisellini, R. Battistutta, G. Zanotti, and L. A. Pinna. Toward the rational design of protein kinase casein kinase-2 inhibitors. *Pharmacology and Therapeutics*, **93**:159–168, 2002.

[84] F. Meggio, M. A. Pagano, S. Moro, G. Zagotto, M. Ruzzene, S. Sarno, G. Cozza, J. Bain, M. Elliott, A. D. Deana, A. M. Brunati, and L. A. Pinna. Inhibition of protein kinase CK2 by condensed polyphenolic derivatives. an in vitro and in vivo study. *Biochemistry*, **43**:12931–12936, 2004.

[85] R. J. Nijveldt, E. van Nood, D. E. van Hoorn, P. G. Boelens, K. van Norren, and P. A. van Leeuwen. Flavonoids: a review of probable mechanisms of action and potential applications. *American Journal of Clinical Nutrition*, **74**:418–425, 2001.

[86] E. Middleton. Effect of plant flavonoids on immune and inflammatory cell function. *Advances in Experimental Medicine and Biology*, **439**:175–182, 1998.

[87] J. V. Formica and W. Regelson. Review of the biology of Quercetin and related bioflavonoids. *Food and Chemical Toxicology*, **33**:1061–1080, 1995.

[88] G. Cozza, P. Bonvini, E. Zorzi, G. Poletto, M. A. Pagano, S. Sarno, A. Donella-Deana, G. Zagotto, A. Rosolen, L. A. Pinna, F. Meggio, and S. Moro. Identification of ellagic acid as potent inhibitor of protein kinase CK2: a successful example of a virtual screening application. *Journal of Medicinal Chemistry*, **49**:2363–2366, 2006.

[89] H. Madari and R. S. Jacobs. An analysis of cytotoxic botanical formulations used in the traditional medicine of ancient Persia as abortifacients. *Journal of Natural Products*, **67**:1204–1210, 2004.

[90] F. Borges, F. Roleira, N. Milhazes, L. Santana, and E. Uriarte. Simple coumarins and analogues in medicinal chemistry: occurrence, synthesis and biological activity. *Current Medicinal Chemistry*, **12**:887–916, 2005.

[91] A. Lacy and R. O'Kennedy. Studies on coumarins and coumarin-related compounds to determine their therapeutic role in the treatment of cancer. *Current Pharmaceutical Design*, **10**:3797–811, 2004.

[92] D. Yim, R. P. Singh, C. Agarwal, S. Lee, H. Chi, and R. Agarwal. A novel anticancer agent, decursin, induces G1 arrest and apoptosis in human prostate carcinoma cells. *Cancer Research*, **65**:1035–44, 2005.

[93] M. Kawase, H. Sakagami, K. Hashimoto, S. Tani, H. Hauer, and S. S. Chatterjee. Structure-cytotoxic activity relationships of simple hydroxylated coumarins. *Anticancer Research*, **23**:3243–3246, 2003.

[94] G. Finn, B. Creaven, and D. Egan. Modulation of mitogen-activated protein kinases by 6-nitro-7-hydroxycoumarin mediates apoptosis in renal carcinoma cells. *European Journal of Pharmacology*, **481**:159–167, 2003.

[95] A. Bortolato, G. Cozza, S. Zanatta, G. Poletto, G Zagotto, E. Uriarte, A. Guiotto, L. A. Pinna, F. Meggio, and S. Moro. Coumarin as attractive casein kinase 2 (CK2) inhibitor scaffold: an integrate approach to elucidated the putative binding motif and explain structure-activity relationships. *Journal of Medicinal Chemistry*, **Accepted**, 2008.

[96] Z. Nie, C. Perretta, P. Erickson, S. Margosiak, R. Almassy, J. Lu, A. Averill, K. M. Yager, and S. Chu. Structure-based design, synthesis, and study of pyrazolo[1,5-a][1,3,5]triazine derivatives as potent inhibitors of protein kinase CK2. *Bioorganic and Medicinal Chemistry Letters*, **17**:4191–4195, 2007.

[97] Z. Nie, C. Perretta, P Erickson, S. Margosiak, J. Lu, A. Averill, R. Almassy, and S. Chu. Structure-based design and synthesis of novel macrocyclic pyrazolo[1,5-a] [1,3,5]triazine compounds as potent inhibitors of protein kinase CK2 and their anticancer activities. *Bioorganic and Medicinal Chemistry Letters*, 2007.

[98] D. B. Kitchen, H. Decornez, J. R. Furr, and J. Bajorath. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nature Review Drug Discovery*, **3**:935–949, 2004.

[99] H. M. Berman, T. Battistuz, T. N. Bhat, W. F. Bluhm, P. E. Bourne, K. Burkhardt, Z. Feng, G. L. Gilliland, L. Iype, S. Jain, P. Fagan, J. Marvin, D. Padilla, V. Ravichandran, B. Schneider, N. Thanki, H. Weissig, J. D. Westbrook, and C. Zardecki. The Protein Data Bank. *Acta Crystallographica. Section D, Biological Crystallography*, **58**:899–907, 2002.

[100] M. A. Miteva, W. H. Lee, M. O. Montes, and B. O. Villoutreix. Fast structure-based virtual ligand screening combining FRED, DOCK, and Surflex. *Journal of Medicinal Chemistry*, **48**:6012–22, 2005.

[101] A Bortolato and S. Moro. In silico binding free energy predictability by using the linear interaction energy (LIE) method: bromobenzimidazole CK2 inhibitors as a case study. *Journal of Chemical Information and Modeling*, **47**:572–582, 2007.

[102] P. J. Hajduk, J. R. Huth, and S. W. Fesik. Druggability indices for protein targets derived from NMR-based screening data. *Journal of Medicinal Chemistry*, **48**:2518–2525, 2005.

[103] A. C. Pierce, G. Rao, and G. W. Bemis. BREED: Generating novel inhibitors through hybridization of known ligands. application to CDK2, p38, and HIV protease. *Journal of Medicinal Chemistry*, **47**:2768–2775, 2004.

[104] B. J. Druker and N. B. Lydon. Lessons learned from the development of an abl tyrosine kinase inhibitor for chronic myelogenous leukemia. *The Journal of Clinical Investigation*, **105**:3–7, 2000.

[105] R. Capdeville, E. Buchdunger, J. Zimmermann, and A. Matter. Glivec (STI571, Imatinib), a rationally developed, targeted anticancer drug. *Nature Review Drug Discovery*, **1**:493–502, 2002.

[106] T. Schindler, W. Bornmann, P. Pellicena, W. T. Miller, B. Clarkson, and J. Kuriyan. Structural mechanism for STI-571 inhibition of abelson tyrosine kinase. *Science*, **289**:1938–1942, 2000.

[107] B. Nagar, W. G. Bornmann, P. Pellicena, T. Schindler, D. R. Veach, W. T. Miller, B. Clarkson, and J. Kuriyan. Crystal structures of the kinase domain of c-Abl in complex with the small molecule inhibitors PD173955 and imatinib (STI-571). *Cancerer Research*, **62**:4236–4243, 2002.

[108] S. W. Cowan-Jacob, V. Guez, G. Fendrich, J. D. Griffin, D. Fabbro, P. Furet, J. Liebetanz, J. Mestan, and P. W. Manley. Imatinib (STI571) resistance in chronic myelogenous leukemia: molecular basis of the underlying mechanisms and potential strategies for treatment. *Mini Review in Medicinal Chemistry*, **4**:285–299, 2004.

[109] S. R. Hubbard and J. H. Till. Protein tyrosine kinase structure and function. *Annual Review in Biochemistry*, **69**:373–398, 2000.

[110] C. D. Mol, D. R. Dougan, T. R. Schneider, R. J. Skene, M. L. Kraus, D. N. Scheibe, G. P. Snell, H. Zou, B. Sang, and K. P. Wilson. Structural basis for the autoinhibition and STI-571 inhibition of c-Kit tyrosine kinase. *Journal of Biological Chemistry*, **279**:31655–31663, 2004.

[111] E. Shtivelman, B. Lifshitz, R. P. Gale, and E. Canaani. Fused transcript of abl and bcr genes in chronic myelogenous leukaemia. *Nature*, **315**:550–554, 1985.

[112] Y. Ben-Neriah, G. Q. Daley, A. M. Mes-Masson, O. N. Witte, and D. Baltimore. The chronic myelogenous leukemia-specific P210 protein is the product of the bcr/abl hybrid gene. *Science*, **233**:212–214, 1986.

[113] B. J. Druker, F. Guilhot, S. G. O'Brien, I. Gathmann, H. Kantarjian, N. Gattermann, M. W. N. Deininger, R. T. Silver, J. M. Goldman, R. M. Stone, F. Cervantes, A. Hochhaus, B. L. Powell, J. L. Gabrilove, P. Rousselot, J. Reiffers, J. J. Cornelissen, T. Hughes, H. Agis, T. Fischer, G. Verhoef, J. Shepherd, G. Saglio, A. Gratwohl, J. L. Nielsen, J. P. Radich, B. Simonsson, K. Taylor, M. Baccarani, C. So, L. Letvak, and R. A. Larson. Five-year follow-up of patients receiving imatinib for chronic myeloid leukemia. *New England Journal of Medicine*, **355**:2408–2417, 2006.

[114] A. Quintás-Cardama, H. Kantarjian, and J. Cortes. Flying under the radar: the new wave of BCR-ABL inhibitors. *Nature Review Drug Discovery*, **6**:834–848, 2007.

[115] B.P. Rubin, M. C. Heinrich, and C. L. Corless. Gastrointestinal stromal tumour. *Lancet*, **369**:1731–1741, 2007.

[116] E. Wardelmann, I. Losen, V. Hans, I. Neidt, N. Speidel, E. Bierhoff, T. Heinicke, T. Pietsch, R. Büttner, and S. Merkelbach-Bruse. Deletion of Trp-557 and Lys-558 in the juxtamembrane domain of the c-kit protooncogene is associated with metastatic behavior of gastrointestinal stromal tumors. *International Journal of Cancer*, **106**:887–895, 2003.

[117] M. C. Heinrich, C. L. Corless, A. Duensing, L. McGreevey, C. Chen, N. Joseph, S. Singer, D. J. Griffith, A. Haley, A. Town, G. D. Demetri, C. D. M. Fletcher, and J. A. Fletcher. PDGFRA activating mutations in gastrointestinal stromal tumors. *Science*, **299**:708–710, 2003.

[118] M. C. Heinrich, C. L. Corless, G. D. Demetri, C. D. Blanke, M. von Mehren, H. Joensuu, L. S. McGreevey, C. Chen, A. D. Van den Abbeele, B. J. Druker, B. Kiese, B. Eisenberg, P. J. Roberts, S. Singer, C. D. M. Fletcher, S. Silberman, S Dimitrijevic, and J. A. Fletcher. Kinase mutations and imatinib response in patients with metastatic gastrointestinal stromal tumor. *Journal Clinical Oncology*, **21**:4342–4349, 2003.

[119] J. Gotlib, J. Cools, J. M. Malone, S. L. Schrier, D. G. Gilliland, and S. E. Coutré. The FIP1L1-PDGFRalpha fusion tyrosine kinase in hypereosinophilic syndrome and chronic eosinophilic leukemia: implications for diagnosis, classification, and management. *Blood*, **103**:2879–2891, 2004.

[120] J. Score, C. Curtis, K. Waghorn, M. Stalder, M. Jotterand, F. H. Grand, and N. C. P. Cross. Identification of a novel imatinib responsive KIF5B-PDGFRA fusion gene following screening for PDGFRA overexpression in patients with hypereosinophilia. *Leukemia*, **20**:827–32, 2006.

[121] J. Aqvist, C. Medina, and J. E. Samuelsson. A new method for predicting binding affinity in computer-aided drug design. *Protein Engineering*, **7**:385–391, 1994.

[122] T. Hansson and J. Aqvist. Estimation of binding free energies for HIV proteinase inhibitors by molecular dynamics simulations. *Protein Engineering*, **8**:1137–1144, 1995.

[123] B. O. Brandsdal, J. Aqvist, and A. O. Smalås. Computational analysis of binding of P1 variants to trypsin. *Protein Science*, **10**:1584–1595, 2001.

[124] M. Graffner-Nordberg, K. Kolmodin, J. Aqvist, S. F. Queener, and A. Hallberg. Design, synthesis, computational prediction, and biological evaluation of ester soft drugs as inhibitors of dihydrofolate reductase from Pneumocystis carinii. *Journal of Medicinal Chemistry*, **44**:2391–2402, 2001.

[125] D. Huang and A. Caflisch. Efficient evaluation of binding free energy using continuum electrostatics solvation. *Journal of Medicinal Chemistry*, **47**:5791–5797, 2004.

[126] K. B. Ljungberg, J. Marelius, D. Musil, P. Svensson, B. Norden, and J. Aqvist. Computational modelling of inhibitor binding to human thrombin. *European Journal of Pharmaceutical Science*, **12**:441–446, 2001.

[127] J. Wang, R. Dixon, and P. A. Kollman. Ranking ligand binding affinities with avidin: a molecular dynamics-based interaction energy study. *Proteins*, **34**:69–81, 1999.

[128] F. Osterberg and J. Aqvist. Exploring blocker binding to a homology model of the open hERG K+ channel using docking and molecular dynamics methods. *FEBS Letters*, **579**:2939–2944, 2005.

[129] A. Carlson and W. L. Jorgensen. An extended linear response method for determining free energies of hydration heather. *Journal of Physical Chemistry*, **99**:10667–10673, 1995.

[130] E. Gallicchio, L. Y. Zhang, and R. M. Levy. The SGB/NP hydration free energy model based on the surface generalized born solvent reaction field and novel nonpolar hydration free energy estimators. *Journal of Computational Chemistry*, **23**:517–529, 2002.

[131] B. A. Tounge and C. H. Reynolds. Calculation of the binding affinity of beta-secretase inhibitors using the linear interaction energy method. *Journal of Medicinal Chemistry*, **46**:2074–2082, 2003.

[132] P. Singh, A. M. Mhaka, S. B. Christensen, J. J. Gray, S. R. Denmeade, and J. T. Isaacs. Applying linear interaction energy method for rational design of noncompetitive allosteric inhibitors of the sarco- and endoplasmic reticulum calcium-ATPase. *Journal of Medicinal Chemistry*, **48**:3005–3014, 2005.

[133] K. Niefind, B. Guerra, I. Ermakowa, and O. G. Issinger. Crystal structure of human protein kinase CK2: insights into basic properties of the CK2 holoenzyme. *EMBO Journal*, **20**:5320–5331, 2001.

[134] Schrödinger LLC. Liaison, version 3.5. *New York, NY*, 2005.

[135] Schrödinger LLC. Ligprep, version 1.6. *New York, NY*, 2005.

[136] P. Pospisil, T. Kuoni, L. Scapozza, and G. Folkers. Methodology and problems of protein-ligand docking: case study of dihydroorotate dehydrogenase, thymidine kinase, and phosphodiesterase 4. *Journal of Receptor Signal Transduction Research*, **22**:141–154, 2002.

[137] Chemical Computing Group Inc. Molecular Operating Enviroment. *1255 University Street, Suite 1600, Montreal, Quebec, Canada H3B 3X3*, 2005.

[138] Wavefunctio Inc. Spartan'02. *Irvine, CA*, 2002.

[139] C. C. Chambers, G. D. Hawkins, C. J. Cramer, and D. G. Truhlar. Model for aqueous solvation based on class IV atomic charges and first solvation shell effects. *Journal of Physical Chemistry*, **100**:16385–16398, 1996.

[140] J. A. Grant, B. T. Pickup, and A. Nicholls. A smooth permittivity function for Poisson-Boltzmann solvation methods. *Journal of Computational Chemistry*, **2001**:608–640, 2001.

[141] W. D. C. P. Cornell, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman. A second generation force field for the simulation of proteins, nucleic acids and organic molecules. *Journal of Americal Chemistry Society*, **117**:5179–5196, 1995.

[142] S. Sarno, P. Vaglio, F. Meggio, O. G. Issinger, and L. A. Pinna. Protein kinase CK2 mutants defective in substrate recognition. purification and kinetic analysis. *Journal of Biological Chemistry*, **271**:10595–10601, 1996.

[143] Advanced Chemistry Development Inc. ACD/pka, version 10.0. *Toronto ON, Canada.*, 2006.

[144] P. A. Kollman, Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D. A. Case, and T. A. Cheatham III. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Accounts of Chemical Research*, **33**:889–897, 2000.

[145] B. Nagar, W. G. Bornmann, P. Pellicena, T. Schindler, D. R. Veach, W. T. Miller, B. Clarkson, and J. Kuriyan. Crystal structures of the kinase domain of c-Abl in complex with the small molecule inhibitors PD173955 and imatinib (STI-571). *Cancer Research*, **62**:4236–4243, 2002.

[146] Tripos Inc. Sybyl 7.0. *1699 South Hanley Rd., St. Louis, Missouri, 63144, USA*, 2006.

[147] M. J. Frisch et al. Gaussian98. *Gaussian Inc.*, 1998.

[148] C. I. Bayly, P. Cieplak, W. D. Cornell, and P. A. Kollman. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *Journal of Physical Chemistry*, **97**:10269–10280, 1993.

[149] D. A. Case, T. A. Darden, T. E. Cheatham III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, K. M. Merz, D. A. Pearlman, M. Crowley, R. C. Walker, W. Zhang, B. Wang, S. Hayik, A. Roitberg, G. Seabra, K. F. Wong, F. Paesani, X. Wu, S. Brozell, V. Tsui, H. Gohlke, L. Yang, C. Tan, J. Mongan, V. Hornak, G. Cui, P. Beroza, D. H. Mathews, C. Schafmeister, W. S. Ross, and P. A. Kollman. Amber 9, University of California, San Francisco. 2006.

[150] Y. Duan, C. Wu, S. Chowdhury, M. C. Lee, G. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, J. Caldwell, J. Wang, and P. Kollman. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *Journal of Computational Chemistry*, **24**:1999–2012, 2003.

[151] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case. Development and testing of a general amber force field. *Journal of Computational Chemistry*, **25**:1157–1174, 2004.

[152] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. Comparison of simple potential functions for simulating liquid water. *Journal of Chemical Physic*, **79**:926–935, 1983.

[153] A. M. Ferrari, G. Degliesposti, M. Sgobba, and G. Rastelli. Validation of an automated procedure for the prediction of relative free energies of binding on a set of aldose reductase inhibitors. *Bioorganic and Medicinal Chemistry*, **15**:7865–77, 2007.

[154] Z. Xiang, C. S. Soto, and B. Honig. Evaluating conformational free energies: the colony energy and its application to the problem of loop prediction. *The Proceedings of the National Academy of Sciences of the United States of America*, **99**:7432–7437, 2002.

[155] F. Fogolari and S. C. E. Tosatto. Application of MM/PBSA colony free energy to loop decoy discrimination: toward correlation between energy and root mean square deviation. *Protein Science*, **14**:889–901, 2005.

Acknowledgement

I would like to express my gratitude to all those who gave me the possibility to complete this thesis. Many are the people helped me so much to reach this crucial result and I am afraid to do not remember them all.

During these years I pass thought different labs, different ways and mentalities to do science, I had the opportunity to meet really extraordinary scientists and wonderful friends. The Molecular Modeling Section of the University of Padova, where always I feel at home: with Stefano, more a friend than a supervisor, Giorgio, Serena, Erika, Lisa, Magda, Francesca, Mattia, StefanoS, Marcof, Massimo, FrancescaF, Giuseppe, Santi, Nicola, Elena, Chiara, MarcoM, Federico.

The Scapozza Lab at the University of Geneva, my second home: Leonardo that always helped me, Loris, Andrea, Remo, Anja, Shaheen, Ralitza, Sabine, Yvonne, LeonardoL, Carole.

The Filizola Lab at NYC, my new sweet home: Marta that gave me a great opportunity, MartaM, Juan Carlos, Debashree.

A special thanks to Barbara, Giuseppe, Prof. Fogolari, Prof. Battistutta, Prof. Scorrano, Prof. Mehler, Christian, Samuele.

And for sure I cannot forget the most important people of my life: Francesca, my one and only, my parents that always support me in every decision and my brothers Paolo and Stefano.