



# Efficace tecnica per progettare molecole

*Per accelerare e migliorare il processo di Drug Design, negli ultimi decenni si è diffuso l'utilizzo di modelli computazionali. In tale ambito rivestono grande importanza metodi per la predizione dell'energia di interazione tra il legante e la sua proteina target. Tra quelli sviluppati recentemente c'è il Linear Interaction Energy (LIE), applicato nello studio in esame per lo sviluppo razionale di un modello energetico volto alla comprensione strutturale di una famiglia di inibitori della protein chinasi CK2*

**I**l grande interesse dell'industria farmaceutica per i processi di fosforilazione cellulare è strettamente legato all'importanza di tale reazione per il controllo della funzione proteica e di conseguenza dell'intera vita cellulare. Il fosfoproteoma, definito come la totalità di proteine fosforilate all'interno della cellula, è controllato e mantenuto in un preciso equilibrio dinamico dalle proteine chinasi e fosfatasi. Le proteine ad azione chinasi catalizzano il trasferimento del gruppo fosforilico gamma di una molecola d'ATP o GTP al peptide substrato, mentre le proteine ad azione fosfatasi sono responsabili della rimozione di tale gruppo rendendo il processo reversibile [1]. In tale contesto l'attività del chinoma, cioè la totalità cellulare delle protein chinasi, deve essere finemente regolata affinché non si incorra in gravi patologie come il cancro, diabete e infiammazione, che rendono tale famiglia di enzimi il secondo più grande gruppo di bersagli a potenziale interesse farmaceutico dell'era post-genomica [2].

La progettazione di molecole o Drug Design è un processo iterativo che inizia quando s'identifica un composto che mostra un interessante profilo biologico e termina quando sia il profilo d'attività che la sintesi della nuova entità chimica sono stati ottimizzati. Negli ultimi decenni si è diffuso l'utilizzo di modelli computazionali per accelerare e migliorare questo processo. Nell'ambito di questo approccio grande importanza hanno i metodi per la predizione dell'energia di interazione tra una piccola molecola, il legante, e la sua proteina target. Punto di par-

tenza per questi studi sono le strutture tridimensionali delle proteine determinate sperimentalmente tramite tecniche NMR o di spettroscopia mediante raggi X o generate attraverso metodi computazionali quali la modellazione per omologia o il protein folding. La struttura del complesso legante-proteina è generata tramite tecniche di docking, con le quali è possibile analizzare velocemente diverse orientazioni e conformazioni (pose) dei leganti all'interno del sito di binding di una proteina per identificare la posa più favorevole. Esistono poi diversi approcci per stimare l'energia di binding a partire dalla struttura 3D del complesso legante-proteina. Tra i metodi sviluppati recentemente c'è il metodo Linear Interaction Energy (metodo LIE). Il metodo LIE si basa sull'assunto che l'energia libera di interazione di un legante a una macromolecola sia linearmente correlata con diversi termini energetici che possono essere calcolati utilizzando un campo di forza di meccanica molecolare. Il valore dell'energia di interazione ( $\Delta G_{bind}$ ) viene stimato utilizzando la seguente formula [3]:

$$\Delta G_{bind} = \alpha (\langle U_{vdw}^b \rangle - \langle U_{vdw}^f \rangle) + \beta (\langle U_{elec}^b \rangle - \langle U_{elec}^f \rangle) + \gamma (\langle U_{cav}^b \rangle - \langle U_{cav}^f \rangle)$$

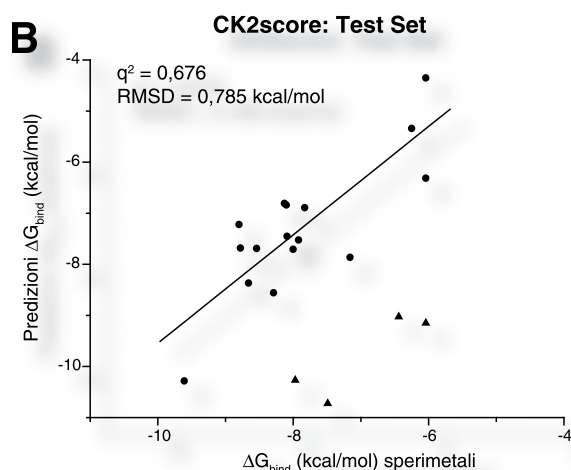
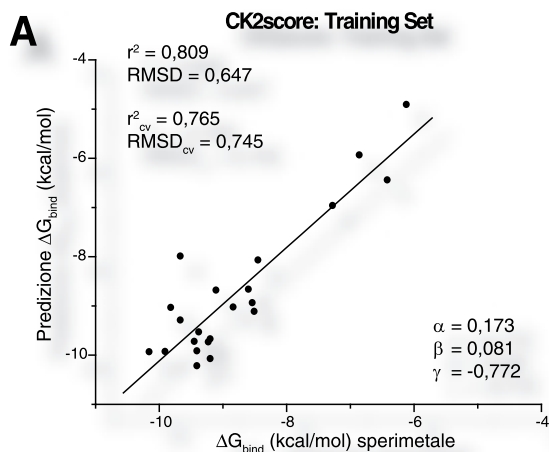
dove i termini  $U^b$  e  $U^f$  descrivono rispettivamente lo stato legato e in soluzione delle molecole coinvolte mentre i termini  $U_{vdw}$ ,  $U_{elec}$  e  $U_{cav}$  sono i valori di energia che stimano rispettivamente i contributi dovuti a interazioni di van der Waals, interazioni elettrostatiche e il parametro di cavità. Questo ultimo parametro è stato introdotto per modellare le interazioni con il solvente. Utilizzando un training set di molecole ad attività

nota si determina il valore dei coefficienti ( $\alpha$ ,  $\beta$  e  $\gamma$ ) arrivando ad ottenere una equazione per la stima dell'energia di interazione.

In tale studio il metodo LIE viene applicato per lo sviluppo razionale di un modello energetico volto alla comprensione strutturale di una famiglia di inibitori della protein chinasi CK2 [4]. Tale enzima, noto anche come Casein chinasi 2, fu una delle prime proteine ad azione chinasi scoperte e controlla una svariata serie di eventi cellulari come l'espressione genica, la regolazione del ciclo cellulare, la sintesi di RNA e proteine [5]. Da un punto di vista patologico diversi dati sperimentali hanno evidenziato sia una sua eccessiva espressione in diverse forme tumorali, sia uno sfruttamento di CK2 da parte di virus che la utilizzano per la fosforilazione di proteine essenziali al loro ciclo vitale, dati che supportano il fatto che tale enzima possa costituire un potenziale bersaglio per farmaci antitumorali e antivirali [6]. La disponibilità di svariati dati di inibizione dell'attività di CK2 ad opera di una serie di derivati benzo di- e tri-azolici e di diverse strutture tridimensionali, ottenute grazie alla spettrometria di diffrazione a raggi X, di complessi ligando-proteina [7,8] costituiscono una base solida per l'utilizzo efficiente del metodo Linear Interaction Energy per creare un modello di predizione dell'energia libera di legame per tale classe di inibitori.

## Materiali e Metodi

Le coordinate di CK2 umana sono state scaricate dal sito internet del RSCB Protein Data Bank (codice PDB 1J91) e sono state incluse per lo



**A)** È indicato il coefficiente di correlazione ( $r^2$ ) e cross-validato ( $r_{cv}^2$ ). Per l'intero test set è stato calcolato l'RMSD e l'RMSD cross validato ( $RMSD_{cv}$ ).  $\alpha$ ,  $\beta$  e  $\gamma$  sono i coefficienti risultanti per l'equazione LIE finale. **B)** RMSD e  $q^2$  riferiti alla retta d'interpolazione corrispondenti alle molecole indicate dai cerchi. I punti indicati con triangoli rappresentano gli outliers.

**Grafico 1.** Grafico della stima dell'energia libera di legame, per il training set (A) e test set (B) data dal modello CK2score per ogni inibitore utilizzato, contro il valore misurato sperimentalmente

studio computazionale 2 molecole d'acqua presenti nel sito attivo e ritenute importanti per il legame dei derivati benzodiazolici. Il sistema è stato preparato per il calcolo LIE attraverso il modulo Protein Preparation [9], mentre la conformazione degli inibitori disegnati è stata ottimizzata utilizzando il programma LigPrep [9], che ha consentito la creazione delle possibili forme tautomeriche e valutazione dei loro vari stati ionici a  $pH=7\pm 1$ . Infine il modello energetico LIE è stato creato grazie al programma Liaison [9] utilizzando il campo di forza OPLS2003 e un modello di solvente implicito (surface generalized Born).

## Risultati e Discussione

Una totale di 25 derivati benzoimidazolici, di cui era nota la costante d'inibizione ( $K_i$ ), sono stati utilizzati come training set per creare il modello energetico basato sul metodo LIE (Figura 1). In tale approccio risulta cruciale la scelta corretta della posizione dell'inibitore nella tasca recettoriale, dello stato ionico dei ligandi e di possibili diverse forme tautomeriche. Tutti questi aspetti sono stati risolti in modo progressivo in conformità a considerazioni chimico-fisiche e statistiche (Figura 2).

Si è deciso infatti di costruire un primo modello con le molecole di cui era disponibile la struttura cristallografica del complesso inibitore proteina, in modo tale da possedere un dato sperimentale che potesse indicare un binding mode ragionevolmente simile con il minimo d'energia di legame tra la proteina e l'inibitore (Figura 3). Per avere un'altra conferma di tale ipotesi di lavoro e allo stesso tempo validare l'efficienza del campo di forza utilizzato, si è inoltre verificato che la posizione della struttura dopo un protocollo di minimizzazione dell'energia potenziale fosse coerente con il dato cristallografico di

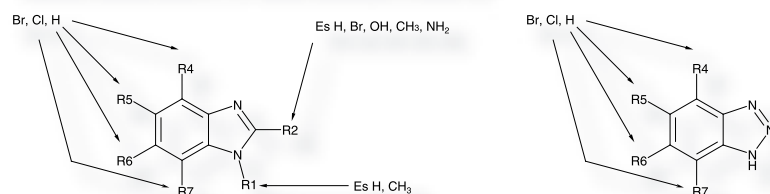
partenza. Sono poi stati considerati i più probabili stati ionici del ligando, oltre che forme tautomeriche, utilizzando come criteri ultimi di scelta 1) l'ottenimento di una pose finale simile a quella di partenza; 2) un'analisi delle interazioni principali di tipo elettrostatico e van der Waals; 3) valutando su base statistica la bontà del modello ottenuto, verificabile attraverso la coerenza delle predizioni dell'energia libera di legame con i dati d'attività sperimentale. Tale primo modello parziale è stato quindi progressivamente implementato utilizzando le altre molecole che presentavano maggiori problematiche di tipo chimico-fisico. Le varie scelte in tal caso sono state basate, oltre che sui principi esposti precedentemente, anche coerentemente ai risultati ottenuti per le molecole di cui era disponibile un dato cristallografico. In particolare è stato possibile evidenziare che per

ottenere un buon modello LIE era necessario considerare un binding mode simile a quello mostrato dall'inibitore TBB o da K37, a seconda del sostituito in posizione 2 (Figura 4). Tali regole di scelta per la pose iniziale sono risultate essenziali per lo studio a causa dell'influenza cruciale che causano alla metodica computazionale LIE. In tale logica è interessante notare come la struttura cristallografica di CK2 in complesso con il ligando K37 mostri due diversi possibili modalità d'interazione con il recettore, uno dei quali stimato essere presente nel 60% (K37A) delle unità cristallografiche, l'altro nel 40% (K37B).

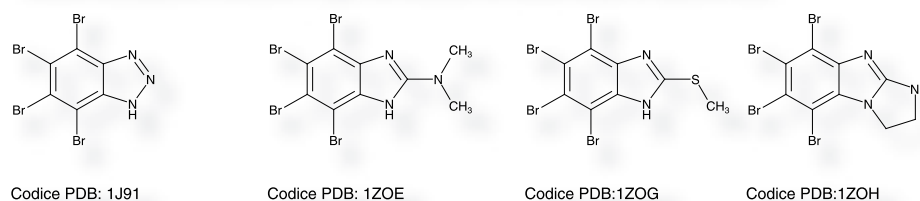
Per investigare maggiormente tale aspetto e approfondire l'ipotesi di binding mode multipli per alcuni derivati abbiamo provato a sfruttare come modello per la pose iniziale entrambe le conformazioni presenti nella struttura cristal-

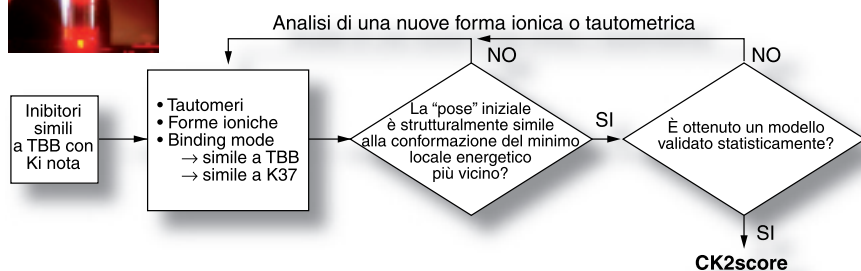
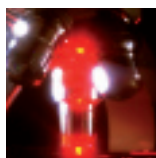
**Figura 1 -** In alto sono rappresentati degli esempi di strutture di derivati benzoimidazolici e triazolici utilizzati nello studio, fra questi gli inibitori indicati in basso sono quelli di cui è pubblicamente disponibile nel Protein Data Bank la struttura cristallografica del loro complesso con CK2

Esempi di derivati benzo bi- e tri- azolici utilizzati per lo studio LIE:



Strutture cristallografiche di diffrazione ai raggi X dei complessi CK2-inibitore utilizzati nel seguente studio:





**Figura 2 - Diagramma di flusso che rappresenta lo schema logico per la scelta del binding mode, dello stato ionico e della forma tautomeric per la costruzione del modello energetico LIE denominato CK2score. Tale approccio è basato su un'implementazione successiva delle diverse molecole partendo da quelle di cui era disponibile il dato cristallografico fino a quelle più problematiche per possibili diversi stati ionici, forme tautomeriche o diverso binding mode. Le scelte sono basate su considerazioni energetiche conformazionali e statistiche, inerenti al modello energetico**

lografica di K37. Si è evidenziato che, per alcuni inibitori, entrambe le conformazioni erano compatibili con un modello LIE anche se il migliore in coefficiente di correlazione era quello basato su K37A in accordo con il dato sperimentale cristallografico che mostrava questa come la modalità di legame più frequente.

I coefficienti dei termini energetici della equazione LIE ( $\alpha = 0,173 \pm 0,003$ ,  $\beta = 0,081 \pm 0,005$  and  $\gamma = -0,772 \pm 0,027$ ) del modello finale denominato CK2score sottolineano che la principale interazione di tali inibitori è dovuta ad effetti idrofobici ( $U_{cav}$ ), seguiti dal contributo di van der Waals. Probabilmente la presenza dei sostituenti poli-bromurati guida il legame al sito attivo di CK2, che viene stabilizzato inoltre in alcuni ligandi da un ponte idrogeno con una molecola d'acqua ed in altri da ulteriori intera-

zioni di van der Waals costituite dal sostituente in posizione 2. Da un punto di vista statistico CK2score presenta un alto coefficiente di correlazione ( $r^2 = 0,809$ ) e un errore sulla stima delle predizioni di circa 0,6 kcal/mole, comparabile con l'incertezza della stima sperimentale delle Ki (Grafico 1-A).

CK2score è stato quindi validato con un test set esterno di 20 ligandi di cui era nota l'attività inibitoria come IC50.

Per tali molecole il binding mode iniziale, la forma tautomeria e lo stato ionico sono stati scelti sulla base degli stessi criteri sviluppati e validati durante la generazione del modello LIE. Da un punto di vista statistico il coefficiente di predizione è risultato buono ( $q^2 = 0,676$ ) dopo l'individuazione ed esclusione di quattro outliers (Grafico 1-B).

### Conclusioni

Il protocollo di progressivo sviluppo del modello basato sul metodo Linear Interaction Energy costituisce una tecnica computazionale veloce ed efficace per predire accuratamente l'energia libera di legame di tale potente classe d'inibitori di CK2.

I valori energetici stimati ( $\alpha$ ,  $\beta$  e  $\gamma$ ), l'energie d'interazione calcolate, la scelta accurata del binding mode e l'analisi attenta della pose finale in concerto con una validazione statistica robusta possono essere utilizzati per descrivere in modo appropriato le possibili differenze nel processo di riconoscimento di CK2 da parte dell'inibitore.

In tale contesto, l'individuazione degli outliers può risultare utile per capire i limiti del modello e l'applicabilità dell'approccio. Inoltre offre una

**Figura 3 - Sito attivo di CK2 umana derivato dalla struttura cristallografica (codice PDB 1JWH). È mostrata la Lisina 68 e l'aminoacido catalitico Asp175, insieme alla Hinge Region (le catene laterali sono nascoste). A sinistra la sovrapposizione di K25 (carboni colorati in bianco), K44 (in verde), i due binding mode di K37 (K37A in grigio e K37B in magenta), mostra la similarità nella posizione dei gruppi idrofobici. A destra TBB (in verde) è sovrapposto a K37B (magenta), sono indicate inoltre due molecole d'acqua che partecipano attivamente al legame di alcuni derivati. Per la costruzione di CK2score la posizione iniziale di ogni inibitore di cui non era disponibile la struttura cristallina è stato ottenuto per analogia a TBB o K37A**



Divisione Pharma

Progettiamo e costruiamo cappe a flusso laminare DOWN CROSS e moduli a flusso laminare LAF.



all@termodesign.it



termoindustriale

1964

TERMOINDUSTRIALE S.p.A.

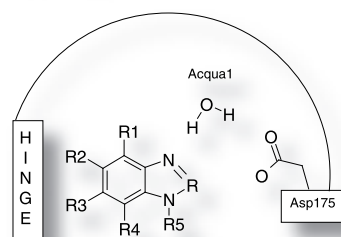
12069 - Santa Vittoria d'Alba (CN)  
Strada Statale 231, 80/D

Tel. 0172.479355 - Fax 0172.479242

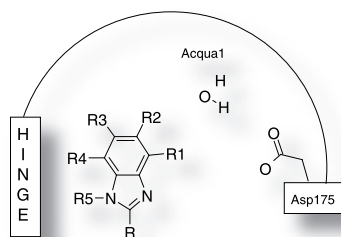
e-mail: pharma@termoindustriale.it

http: www.termoindustriale.it

CK2score:



R = N, CH → binding mode simile a TBB



R ≠ H → binding mode simile a K37

comprensione chimica delle interazioni chiave per il legame, analizza i possibili ruoli delle molecole d'acqua e razionalizza possibili differenti modalità di legare la proteina come mostrato dalle diverse strutture cristallografiche disponibili. Tali conoscenze possono essere la chiave per lo sviluppo di possibili nuovi potenti derivati bromo-benzimidazolici.

#### Afferenze degli autori

B. Bortolato e S. Moro, Molecular Modeling Section, Department of Pharmaceutical Sciences, University of Padova  
E. Fioravanzo e M. Mabilia, 2S.IN - Soluzioni Informatiche S.a.s.

#### Ringraziamenti

Il lavoro di modellistica molecolare coordinato da S.M. è stato condotto grazie al supporto del Ministero Italiano per l'Università e la Ricerca (MIUR), Roma, Italia e l'Università degli studi di Padova, Padova, Italia. S.M. è molto grato a Stuart Murdock e Schrödinger Inc. per l'aiuto tecnico e scientifico. Si ringrazia Chemical Computing Group per la lunga e fruttuosa collaborazione.

#### Referenze

- [1] Cohen, P. The origins of protein phosphorylation. *Nat. Cell Biol.* **2002**, 4, E127-E130.
- [2] Cohen, P. Protein kinases- the major drug targets of the twenty-first century? *Nat. Rev.* **2002**, 1, 309-315.

**Figura 4 - Regole derivate dal modello energetico CK2score: il binding mode iniziale è stato scelto a seconda del sostituito in posizione 2, che costituisce la principale differenza presente tra i vari derivati**

- [3] Åquist, J.; Medina, C.; Samuelsson, J.-E. A new method for predicting binding affinity in computer-aided drug design. *Protein Eng.* **1994**, 7, 385-391.
- [4] Pinna, L. A. A historical view of protein kinase CK2. *Cell. Mol. Biol. Res.* **1994**, 40, 383-390.
- [5] Meggio, F.; Pinna, L. A. One-thousand-and-one substrates of protein kinase CK2? *FASEB J.* **2003**, 17, 349-368.
- [6] Pinna, L. A.; Meggio, F. Protein kinase CK2 ("casein kinase-2") and its implication in cell division and proliferation. *Prog. Cell Cycle Res.* **1997**, 3, 77-97.
- [7] Battistutta, R.; De molier, E.; Sarno, S.; Zanotti, G.; Pinna, L. A. Structural features underlying selective inhibition of protein kinase CK2 by ATP site-directed tetrabromo-2-benzotriazole. *Protein Sci.* **2001**, 10, 2200-2206. w
- [8] Sarno, S.; Moro, S.; Meggio, F.; Zagotto, G.; Dal Ben, D.; Ghisellini, P.; Battistutta, R.; Zanotti, G.; Pinna, L. A. Toward the rational design of protein kinase casein kinase-2 inhibitors. *Pharmacol. Ther.* **2002**, 93, 159-168. w
- [9] Schrödinger, LLC, New York, NY, **2005**.

**MEISSNER**  
FILTRATION PRODUCTS, INC.

Elementi filtranti per  
l'industria farmaceutica  
e per bio-processo

Excellence in  
filtration  
solutions

**ASCO**  
FILTRI

Distributore per l'Italia  
Asco Filtri s.r.l.  
20059 Rozzano (MI) - ITALY  
Via Silvio Pellico, 5/1  
Tel. +39 02 89057 1  
Fax +39 02 89267201  
e-mail: asco@ascofiltri.com  
internet: www.ascofiltri.com

A UNIT OF  
**Finder**