

INTRODUZIONE

La chimica farmaceutica è la disciplina che studia gli aspetti chimici e strutturali che determinano o influenzano l'attività dei farmaci. Queste sostanze, per lo più di natura organica, sono ottenute per via sintetica o mediante estrazione da materiali naturali, e vengono usate nella pratica medica per prevenire o curare le malattie. I principali obiettivi della chimica farmaceutica sono la sintesi e lo studio delle relazioni struttura - attività dei farmaci. Per molti anni, la scoperta di nuovi farmaci è avvenuta in modo casuale attraverso la modificazione di un *lead compound*, senza spiegare perché la variazione strutturale di una molecola *lead* provocasse dei mutamenti nell'attività biologica. Appare evidente che a determinare l'attività biologica siano le proprietà intrinseche (chimico-fisiche) delle molecole. Pertanto conoscendo tali proprietà, compresa l'attività biologica, è possibile derivare un modello, il quale metta in relazione i vari fattori o variabili (proprietà). Questi modelli descrivono le relazioni quantitative tra le proprietà intrinseche delle molecole e le attività biologiche e sono, infatti, detti QSAR (*Quantitative Structure Activity Relationship*).

Per modello si intende una funzione matematica (ma può essere anche grafica) che è in grado di spiegare un evento. I modelli, per loro natura, sono necessariamente delle rappresentazioni approssimate della realtà. La loro accuratezza, ossia la capacità dei modelli di descrivere, spiegare e quindi di prevedere un evento, dipende dalla conoscenza dei fattori (variabili) che lo influenzano. Poiché tale conoscenza, specie se l'evento è complesso come una interazione farmaco-recettore, è incompleta, il modello che ne deriva è una rappresentazione approssimata della realtà. I modelli QSAR, tuttavia, possono essere usati per spiegare l'effetto biologico delle sostanze studiate, e quindi come utile punto di partenza per progettare nuovi farmaci, che presentino il quadro di attività biologica voluto.

Le QSAR classiche usano descrivere le molecole attraverso proprietà chimico-fisiche (descrittori). I modelli così ottenuti presentano un buon carattere descrittivo, ossia spiegano perché una molecola è attiva o inattiva, ma possiedono uno scarso potere predittivo, cioè non sono in grado di prevedere l'attività di nuove molecole. Per permettere una progettazione di farmaci più efficace è, quindi, necessario conoscere la struttura tridimensionale della conformazione bioattiva del composto *lead*, che molto spesso è il ligando naturale. In questo modo non si sfruttano solo le proprietà chimico-fisiche della molecola, ma anche quelle conformazionali. Sono stati, così, sviluppati una serie di approcci per le QSAR tridimensionali che sfruttano la struttura 3D dei ligandi. I modelli QSAR 3D pur presentando un potere predittivo maggiore rispetto alle QSAR classiche, hanno lo stesso una validità locale; poiché non è possibile progettare molecole molto diverse da quelle usate per la creazione del modello.

Un miglioramento in termini di predittività di questi prototipi, può essere ottenuto attraverso una conoscenza più dettagliata delle interazioni ligando-proteina (*target*). Tale conoscenza deriva dalla risoluzione della struttura 3D della proteina recettoriale o meglio del complesso ligando-*target*. Tali strutture si possono ottenere attraverso raggi X e NMR (quando è possibile isolare la proteina recettoriale), oppure, (nel caso in cui non sia possibile isolarla) dedurla dalla sequenza primaria per omologia con altre proteine, che svolgono funzioni analoghe e di cui sia nota la struttura 3D. Altri metodi per migliorare la QSAR sono le *Scoring Functions*, che oltre a dare un responso sull'attività, cercano di trovare una relazione generale applicabile ad altri composti.

Al momento, sono disponibili diversi programmi che assistono il chimico farmaceutico nella progettazione di nuove strutture molecolari capaci di interagire con i più svariati recettori di cui è conosciuta la struttura tridimensionale (*structure-based drug design*).

Tra le tecniche introdotte, quelle di maggiore innovazione sono la chimica combinatoriale e la modellazione molecolare al computer. Mentre il primo

approccio permette la rapida evoluzione di composti capostipiti o *lead compounds* mediante la sintesi di numerosi derivati in un tempo relativamente breve; la chimica computazionale si inserisce nella chimica farmaceutica sia ottimizzando i *lead compounds*, sia sviluppando nuove strutture avvalendosi di basi puramente teoriche, supportate da dati sperimentali come la cristallografia ai raggi X e le prove di “binding” enzimatico. I campi di applicazione della chimica computazionale vanno, quindi, dallo studio puramente strutturale, all’applicazione della statistica su serie di composti modellati, per generare modelli matematici capaci di predire le caratteristiche chimico-fisiche e biologiche di nuove molecole non ancora sintetizzate.

Una comune procedura di *structure-based drug design* comprende i seguenti stadi:

1. ricerca delle posizioni chiave all’interno del sito attivo, le quali possono interagire con differenti gruppi funzionali;
2. connessione dei gruppi funzionali secondo le loro posizioni tridimensionali così da formare le nuove molecole;
3. predizione dell’attività-affinità delle molecole progettate per il determinato recettore;
4. sintesi delle molecole che in fase di progettazione sono risultate le migliori su base di interazione farmaco-recettore.

Un metodo per predire l’attività di un nuovo ligando si basa sul calcolo diretto dei parametri termodinamici, che prendono parte alla formazione del complesso ligando-recettore.

In particolare, per predire l’analogia farmaco–recettore, si usano dei metodi computazionali, che sfruttano 2 fasi consecutive:

- 1 *Lead generation*: caratterizzata da *scoring function* veloci ed abbastanza accurate; queste vengono usate soprattutto per identificare le strutture più attive in una serie di composti
- 2 *Lead optimization*: richiede *force field* accurati ed analizza solo le strutture più attive, senza curarsi del tempo di calcolo.

I *force field* utilizzati in questi tipi di elaborazione, possono essere raggruppati in 3 diverse classi:

- 1 *All atom force field* (come AMBER, MM3, CHARMM, ecc.), i quali sfruttano il metodo dell'energia libera perturbazionale (FEP), che attraverso la dinamica molecolare calcola le energie di legame relative ai ligandi di una proteina. Il metodo è accurato, ma computazionalmente costoso e limitato a ligandi a struttura simile. Questo metodo è comunque molto usato nelle *lead optimization*.
- 2 *Empirical approaches* (come SCORE 1, LUDI, VALIDATE, ecc.), nei quali l'energia di legame è approssimata da una equazione *master*, che è la somma di più funzioni corrispondenti ai contributi di ΔS e ΔU . I metodi di *scoring functions* più comuni sono basati su questa equazione *master*.
- 3 *Knowledge-based methods* (come SMOG, PMF, BLEEP, ecc.), sfruttano la meccanica statistica. L'energia libera di legame è la somma delle energie libere (o dei potenziali delle forze medie) dei contatti interatomici tra ligando e sito attivo della proteina; i quali vengono calcolati dalla loro frequenza nei *database* delle strutture. Questi metodi sono veloci come le *scoring function*, e possono essere, perciò, usati per le *lead generation*.

In questa tesi saranno prese in considerazione diverse tecniche di *scoring functions*, che, secondo le più recenti pubblicazioni, sono le metodologie che meglio permettono la creazione di modelli altamente predittivi , ed in particolare verrà sviluppata ed applicata su appositi *test sets* il metodo VALIDATE.