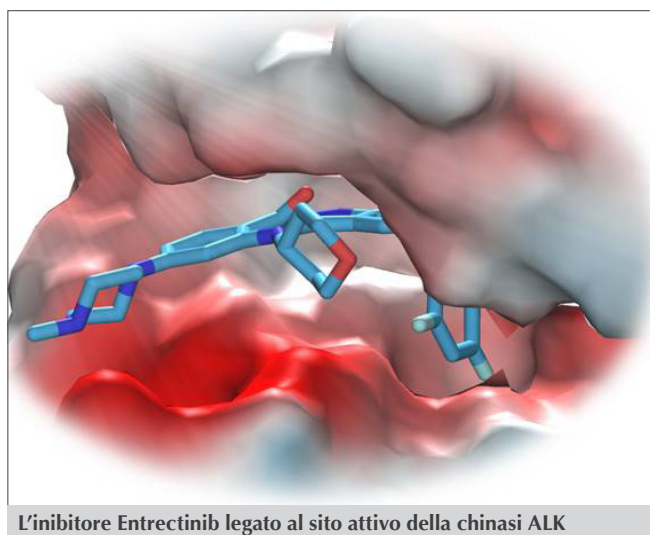




GIULIA CANEVARI, MARINA FASOLINI, ELENA CASALE  
DIPARTIMENTO DI CHEMICAL CORE TECHNOLOGIES  
NERVIANO MEDICAL SCIENCES SRL, NERVIANO (MI)  
ELENA.CASALE@NERVIANOMS.COM

## IL FARMACO VISTO AI RAGGI X

*Non solo chimica e biologia, anche la cristallografia ha un ruolo chiave nella progettazione di nuovi farmaci. La conoscenza dettagliata delle interazioni tra la proteina bersaglio e il suo legante facilita il drug design, aumentando la probabilità di successo in molte fasi del processo di drug discovery.*



L'inibitore Entrectinib legato al sito attivo della chinasi ALK

A più di mezzo secolo dalle sue prime applicazioni in campo biologico, la cristallografia a raggi X rimane la tecnica d'elezione per ottenere un modello a risoluzione atomica della struttura tridimensionale di macromolecole d'interesse biologico, siano esse acidi nucleici, proteine, complessi multiproteici o virus. Le informazioni strutturali ottenute con questa metodologia, non solo hanno aiutato a rivelare e comprendere la complessità dei sistemi biologici, ma hanno anche contribuito alla scoperta e allo sviluppo di numerosi farmaci, con l'avvento del cosiddetto Structure-Based Drug Design (SBDD). Avere a disposizione un modello dettagliato della macromolecola bersaglio (*drug target*, nella maggior parte dei casi una proteina) in forma libera o in complesso con piccole molecole (ad esempio cofattori, substrati o inibitori) permette di capire dove e come i leganti interagiscano con il target, fornendo prezio-

se indicazioni per la progettazione di un farmaco. Le prime strutture cristallografiche di proteine risalgono agli anni Cinquanta e Sessanta, con i pionieristici lavori di Kendrew (struttura della mioglobina, 1958) [1] e Perutz (struttura dell'emoglobina, 1960) [2]. Tuttavia, la mancanza di opportuni supporti sperimentali e computazionali ha inizialmente rallentato lo sviluppo applicativo della tecnica, il cui utilizzo nella ricerca farmaceutica risale a poco più di una trentina di anni fa. I primi successi dello SBDD risalgono alla fine degli anni Ottanta, con lo sviluppo di inibitori delle proteasi dell'HIV [3]; pochi anni dopo nascono compagnie biotech come Vertex, Syrrx e Astex che puntano principalmente su questo approccio per l'identificazione di nuovi farmaci. Oggi la cristallografia di proteine è comunemente utilizzata nell'industria farmaceutica e si può considerare parte integrante del processo di sviluppo di un farmaco (*drug discovery*).

### Legenda

ALK	Anaplastic Lymphoma Kinase
cryo-EM	Crio-Microscopia Elettronica
FBDD	Fragment Based Drug Design
HTS	High-Throughput Screening
MAD	Multiple Anomalous Dispersion
MIR	Multiple Isomorphous Replacement
MR	Molecular Replacement
NMR	Risonanza Magnetica Nucleare
PDB	Protein Data Bank
SAD	Single Anomalous Diffraction
SAR	Relazione Attività-Struttura
SBDD	Structure-Based Drug Design
SPR	Risonanza Plasmonica Superficiale



## La determinazione della struttura cristallografica

Per determinare la struttura è necessario innanzitutto ottenere i cristalli della proteina di interesse; questo è purtroppo ancora oggi il “collo di bottiglia” di tutto il processo (Fig. 1). Nella fase di cristallizzazione, si devono identificare le condizioni chimico-fisiche che riducono la solubilità della proteina e ne favoriscono il passaggio dalla soluzione allo stato solido “cristallino”. Il processo è strettamente empirico e consiste nel testare (per esempio mediante la tecnica della diffusione di vapore) numerose condizioni in cui vengono variati parametri come precipitante, tampone, pH e temperatura. Il successo è legato alla qualità della macromolecola da cristallizzare, che deve essere pura ed omogenea (ovvero avere per esempio lo stesso stato di fosforilazione, oligomerizzazione etc.). Non sempre è possibile ottenere cristalli della proteina intera (full length) perché troppo flessibile ed è quindi necessario lavorare con forme troncate o mutate (costrutti) o con singoli domini. Le variabili coinvolte sono numerose e la cristallizzazione può richiedere tempi molto lunghi; i progressi tecnologici degli ultimi decenni hanno in parte velocizzato il processo aumentandone le probabilità di successo. Lo sviluppo della tecnologia del DNA ricombinante ha, per esempio, reso disponibili metodi e strumentazioni che permettono il clonaggio, l'espressione e la purificazione di diversi costrutti, che possono essere testati in parallelo. In fase di cristallizzazione, sono oggi disponibili kit commerciali di soluzioni precipitanti (sparse matrix screens) e vengono impiegati sistemi automatizzati che consentono di disegnare e preparare centinaia di prove di cristallizzazione utilizzando quantità minime di campione.

Nel processo SBDD è importante ottenere cristalli della proteina legata con più ligandi. I complessi possono essere preparati per cocrystallizzazione (aggiungendo l'inibitore alla proteina in soluzione e cristallizzando il complesso) o per soaking (immergendo i cristalli della proteina nativa in una “soluzione madre” contenente il ligando). Per entrambi i metodi è comunque cruciale identificare una condizione di cristallizzazione robusta e facilmente riproducibile, in quanto spesso è necessario determinare decine di strutture.

Ottenuti i cristalli, la fase successiva è quella della raccolta dei dati di diffrazione. A questo scopo, i cristalli sono congelati in azoto liquido per ridurre il danneggiamento causato dalla radiazione e ottenere quindi dati a più alta risoluzione.

La raccolta dei dati è oggi effettuata principalmente utilizzando la radiazione di sincrotrone, che ha il vantaggio di essere molto intensa (migliaia di volte in



Fig. 1 - Le fasi della determinazione della struttura cristallografica in un processo di SBDD

più rispetto alle sorgenti di laboratorio) e modulabile (ovvero è possibile scegliere la lunghezza d'onda del fascio, cosa utile per esempio per gli esperimenti di dispersione anomala MAD, Multiple Anomalous Dispersion). I sincrotroni di ultima generazione producono un fascio di raggi X estremamente brillante e focalizzato, che permette l'analisi di cristalli anche molto piccoli (poche decine di micron di lato) e sono equipaggiati con sistemi automatizzati per il montaggio dei cristalli e con rivelatori molto veloci, che permettono la raccolta di un set di dati in pochi minuti.

Le immagini di diffrazione sono costituite da una serie di riflessi, per ciascuno dei quali è possibile misurare l'intensità  $I(hkl)$ , che è proporzionale al quadrato del modulo del fattore di struttura  $F(hkl)$ . Quest'ultimo è una quantità complessa (dotata cioè di modulo e fase) da cui è possibile risalire alla densità elettronica del cristallo mediante anti-trasformata di Fourier. Per eseguire quest'operazione, però, non è sufficiente conoscere il modulo di  $F(hkl)$ , ma è necessario anche ricavarne la fase, incognita. Per la soluzione del problema della fase sono stati sviluppati diversi approcci. Se si ha a disposizione la struttura di una proteina simile a quella di interesse, si può utilizzare il metodo del Molecular Replacement (MR): mediante l'uso di funzioni di rotazione e traslazione, il modello viene posizionato all'interno della cella elementare del cristallo e poi utilizzato per calcolare le fasi e la mappa di densità

elettronica. In assenza di un modello di partenza, si utilizzano metodi che sfruttano lo scattering anomalo osservato per alcuni elementi chimici; ad esempio, il Multiple Isomorphous Replacement (MIR) richiede di raccogliere dati dal cristallo nativo e in complesso con più atomi pesanti. Il Single Anomalous Diffraction (SAD) sfrutta il segnale dello zolfo, mentre il MAD prevede la raccolta di più dataset a diverse lunghezze d'onda e può richiedere l'inserimento di seleniomietionine all'interno della proteina.

Una volta ottenuta la mappa di densità elettronica, si costruisce un modello iniziale che è poi migliorato mediante cicli iterativi d'affinamento e ricostruzione per migliorare l'accordo tra il modello e i dati sperimentali.

Per l'analisi dei dati cristallografici sono oggi disponibili software molto efficienti, che hanno permesso una notevole automatizzazione del processo.

Testimone della straordinaria evoluzione della tecnica nel tempo è la "Protein Data Bank" (PDB, la banca dati che raccoglie le coordinate e i dati di diffrazione di tutte le strutture macromolecolari pubblicate), che, alla fine del 2016, contava più di 112.000 strutture depositate, di cui circa l'80% nell'ultimo decennio.

Lo sviluppo tecnologico ha rappresentato un enorme vantaggio anche per la drug discovery; oggi è possibile acquisire i dati di diffrazione e risolvere la struttura in poche ore, direttamente dal computer del proprio ufficio, collegandosi in remoto ad un sincrotrone. Il chimico farmaceutico può quindi apprendere, entro pochi giorni dalla sua sintesi, come un inibitore si lega al target, consentendone l'ottimizzazione per cicli di sintesi e feedback strutturistico, quasi "in tempo reale".

## Il contributo della cristallografia di proteine nel processo di drug discovery

La scoperta di un farmaco è un percorso lungo e complesso che nasce dalla sinergia di diverse competenze e tecnologie. Tutto comincia con l'identificazione del bersaglio biologico (target identification). Mediante saggi biochimici *in vitro*, si testano poi diverse librerie di composti chimici (High-Throughput Screening, HTS) e si identificano i primi "hits", ovvero composti dotati di una minima attività biologica nei confronti del bersaglio (hit identification). La fase successiva consiste nell'ottimizzazione chimica degli "hits" per ottenere il "lead", ovvero una molecola, all'interno di una classe chimica profilata per relazione tra struttura e attività (SAR), con una buona attività biologica e un profilo di selettività e tossicità accettabili (lead identification). Livelli successivi di ottimizzazione chimica,

finalizzati a migliorare le caratteristiche farmacodinamiche, metaboliche e tossicologiche trasformano poi il lead in quello che sarà il potenziale farmaco (lead optimization).

La cristallografia può contribuire a tutte queste fasi, ma senza dubbio il suo impatto è maggiore negli stadi iniziali. Idealmente sarebbe desiderabile ottenere i cristalli della proteina target il più presto possibile, addirittura in fase di validazione del bersaglio molecolare. La struttura tridimensionale può rivelare, infatti, se il target è "druggable" ovvero se esistono delle cavità (sito attivo, tasche allosteriche, interfacce proteina-proteina) che sono accessibili e adatte a legare un inibitore, o se vi siano elementi strutturali che permettano di ottenere selettività. Questo aspetto non è da trascurare, soprattutto quando il bersaglio appartiene a una famiglia di proteine strutturalmente simili, come per esempio le chinasi.

Conoscendo la struttura, si possono identificare inibitori utilizzando metodi computazionali: disegnare molecole che si legano in modo simile a leganti naturali o comunque noti (rational drug design) oppure effettuare uno screening virtuale di librerie chimiche (virtual screening) e, in base a calcoli modellistici, selezionare un numero ristretto di molecole a cui dare priorità di valutazione in test biologici.

Al termine della campagna HTS, ottenere la struttura della proteina bersaglio legata ai diversi hits identificati può facilitare le scelte dei chimici farmaceutici. L'impatto è tanto maggiore quante più strutture cristallografiche si hanno a disposizione e quanto più strutturalmente diversi fra loro sono gli hits. Le informazioni che si possono ottenere dalla comparazione di più strutture cristallografiche sono molteplici. È possibile caratterizzare in maniera dettagliata il sito di legame rivelando gli amminoacidi chiave per l'interazione con la molecola bioattiva e quelli potenzialmente importanti per modulare affinità e selettività. A volte l'interazione con un legante può indurre dei cambiamenti conformazionali della proteina non facilmente prevedibili e che possono offrire nuove opportunità. Un esempio sono gli inibitori chinasi di tipo II, molecole che si legano alla proteina che si trova in una conformazione inattiva, definita "DFG out" [4]. Interagendo con zone diverse della proteina questi composti possono avere potenze e profili di selettività differenti rispetto agli inibitori che occupano solo il sito attivo (inibitori di tipo I). L'imatinib (Glivec), primo farmaco inibitore di chinasi messo sul mercato per la cura della leucemia mieloide cronica è appunto un inibitore di tipo II [5].

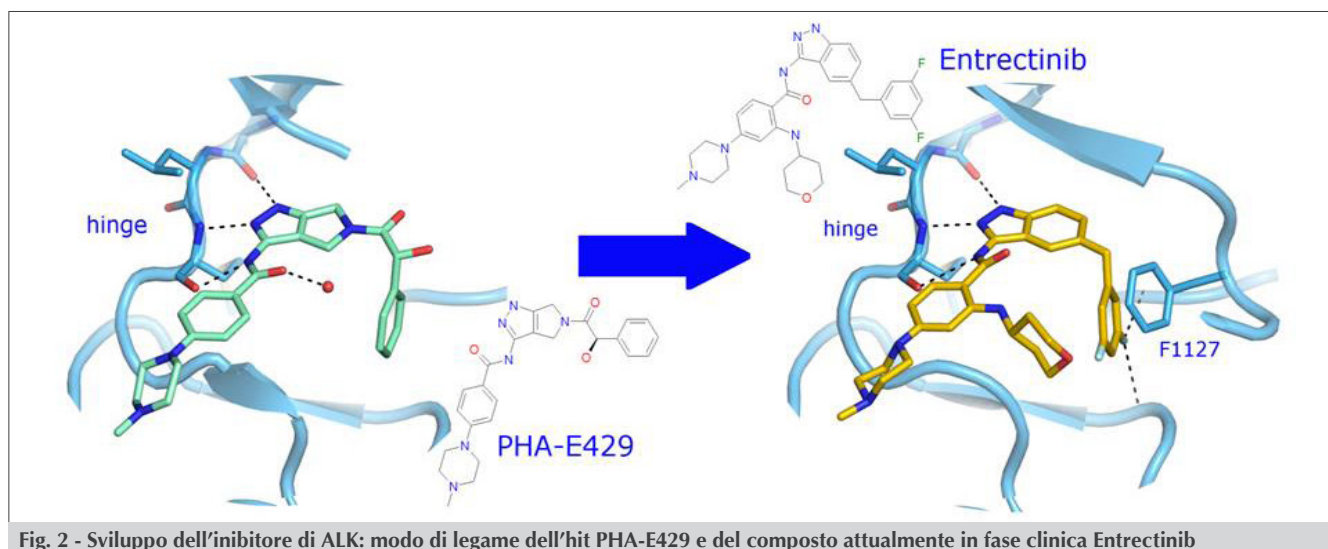


Fig. 2 - Sviluppo dell'inibitore di ALK: modo di legame dell'hit PHA-E429 e del composto attualmente in fase clinica Entrectinib

Da un punto di vista più strettamente chimico, la disponibilità di dati strutturali facilita la comprensione della relazione tra struttura e attività (SAR) in quanto permette di identificare i gruppi funzionali dell'inibitore (donatori o accettori di legame idrogeno, superfici lipofile, sostituenti aromatici) responsabili dell'attività biologica. Queste informazioni, unite a valutazioni di fattibilità chimica, contribuiscono alla selezione degli hit che inizieranno il percorso di ottimizzazione per divenire il futuro farmaco.

Il supporto strutturale continua nelle fasi di hit e lead optimization in modo interattivo e a stretto contatto con il chimico medicinale. Le nuove molecole sintetizzate sono cristallizzate e, basandosi sui dati SAR e SBDD, si identificano i gruppi che possono essere modificati per migliorare inizialmente le caratteristiche di potenza e selettività o, in fase finale, quelle farmacocinetiche o metaboliche.

È opportuno sottolineare che la caratterizzazione strutturale del sito di legame permette anche la progettazione di inibitori di seconda generazione in grado di legare forme mutate del target resistenti (poco affini) al farmaco.

In un processo SBDD è importante ottenere cristalli che diffrangano ad una buona risoluzione ( $\leq 2,0$ - $2,5$  Å). Avendo a disposizione un modello atomico ad alta risoluzione non solo si determinano in modo accurato le interazioni tra legante e residui amminoacidici (distanze di legami idrogeno o interazioni Van der Waals) ma è anche possibile visualizzare la presenza di molecole d'acqua strutturali, legate alla proteina. Queste possono mediare le interazioni tra la molecola bioattiva e il target, quindi contribuire alla sua affinità. L'opportuna modifica dell'inibitore con gruppi funzionali che mimano le interazioni delle molecole di acqua può portare un miglioramento della sua potenza.

Questo tipo di analisi è stato utile per l'identificazione di inibitori della proteina chinasi ALK (Anaplastic Lymphoma Kinase), progetto sviluppato in Nerviano Medical Sciences (NMS) che riportiamo come esempio di approccio SBDD [6]. Mutazioni attivanti o traslocazioni di ALK sono responsabili dell'insorgenza di diversi tipi di tumore, come ad esempio il linfoma a grandi cellule anaplastico e il tumore del polmone non a cellule piccole (NSCLC). Oggi ALK è un target validato, per il quale sono stati sviluppati diversi inibitori, alcuni di questi (crizotinib e alectinib) già approvati dalle autorità regolatorie statunitensi (FDA). Quando il progetto ALK ha avuto inizio in NMS non erano noti né inibitori di questa proteina né la sua struttura cristallografica. Una campagna di screening della collezione di molecole sintetiche proprietarie ha permesso di identificare alcuni hit, tra cui i più promettenti sono stati i composti 3-ammino indazoli. Parallelamente sono iniziate le prove per ottenere i cristalli del dominio chinasi di ALK. L'identificazione delle condizioni di cristallizzazione non è stata facile; dopo numerosi tentativi, utilizzando diversi costrutti, la strategia vincente è risultata essere quella di co-purificare la proteina in presenza dell'inibitore d'interesse. Questo approccio ci ha permesso di ottenere la struttura ad alta risoluzione di ALK in complesso sia con inibitori della classe chimica degli indazoli sia con altri hit emersi dallo screening. Tutti i composti si legano al sito attivo della chinasi in modo competitivo rispetto al legante naturale (ATP), formano due o tre legami a idrogeno chiave con la regione della "hinge region" (una breve sequenza di amminoacidi che funge da "cardine" tra i due lobi delle chinasi) e si estendono all'interno della tasca, dando luogo ad un network di interazioni, polari e non polari, con i residui del sito attivo. In parti-

colare, la struttura di ALK in complesso con l'inibitore PHA-E429 (Fig. 2) ha rivelato la presenza di una molecola d'acqua nella regione normalmente occupata dalla parte zuccherina dell'ATP, legata mediante un ponte a idrogeno con il composto. Questo dettaglio ha suggerito che l'introduzione di un opportuno sostituito sull'anello, in posizione 2', potesse spiazzare la molecola d'acqua instaurando interazioni favorevoli. Utilizzando queste informazioni, facendo leva sulla disponibilità di librerie combinatoriali di analoghi rappresentate nella collezione dello screening e sui protocolli di sintesi modulare di queste librerie, con un approccio combinato di chimica medicinale e docking *in silico*, è stato possibile consolidare e affinare una SAR che ha poi portato in breve tempo all'identificazione della molecola lead, Entrectinib (Fig. 2). La successiva determinazione della struttura cristallografica ha confermato le assunzioni fatte in fase di ottimizzazione, mostrando, inoltre, una particolare conformazione "chiusa" del "glycine-rich loop" (una regione flessibile che funge da "tetto" del sito attivo delle chinasi), mai osservata in precedenza, in cui la fenilalanina 1127 forma un'interazione di stacking favorevole con il composto. Il cammino di Entrectinib è proseguito con successo, e la molecola si trova ora in fase di sperimentazione clinica.

## La cristallografia e il Fragment Based Drug Design

Lo screening di frammenti (Fragment Based Drug Design, FBDD) è un metodo, complementare al tradizionale screening HTS, utile per la selezione e lo sviluppo degli hits.

Il FBDD si basa sull'identificazione di frammenti, ovvero molecole molto piccole (150-300 Dalton) in grado di legarsi, seppur debolmente, alla proteina target riconoscendo siti specifici nonostante la loro semplicità e limitatezza di punti di interazione. La determinazione della struttura cristallografica di complessi con frammenti di debole affinità è cruciale per confermare e localizzare il riconoscimento specifico e facilitare un disegno razionale di espansione sintetica. Gli hits, essendo strutturalmente semplici, sono facilmente modificabili da un punto di vista chimico e rappresentano un ottimo punto di partenza per l'ottenimento di inibitori strutturalmente nuovi. Questo approccio si è rivelato molto potente [7] e ha portato negli ultimi anni allo sviluppo di due farmaci attualmente in commercio: Vemurafenib, un inibitore della forma mutata della chinasi BRAF, utilizzato nella cura del melanoma [8], e Venetoclax, farmaco per la leucemia linfatica cronica, che agisce inibendo la proteina Bcl-2 [9].

La cristallografia ha un ruolo chiave in questo processo e i suoi contributi sono molteplici. I raggi X possono essere utilizzati come metodo iniziale di screening, mediante il soaking di cristalli della proteina target con miscele di frammenti, oppure nella riconferma degli hits identificati con altre tecniche (Risonanza Magnetica Nucleare (NMR) o la Risonanza Plasmonica Superficiale (SPR)) per escludere effetti artefactuali, frequenti nella misurazione di deboli attività.

Il ruolo della cristallografia è fondamentale soprattutto nella fase di ottimizzazione degli hits; conoscere come uno o più frammenti si legano nel sito attivo della proteina permette di utilizzare tecniche come il "fragment growing" (crescita del frammento mediante l'introduzione di gruppi funzionali che aumentano le interazioni con la proteina) o il "fragment linking" (ottimizzazione basata sull'unione di due frammenti che occupano zone prossimali diverse del sito attivo) che possono portare, attraverso un processo iterativo, all'ottenimento di un inibitore con l'attività biologica desiderata.

L'approccio FBDD è stato applicato con successo presso il Nerviano Medical Sciences per l'identificazione di inibitori di HSP90 [10]. La Heat Shock Protein 90 (Hsp90) è una chaperonina che regola la stabilità conformazionale e l'attività di molte proteine oncogeniche, ed è quindi un target oncologico interessante. Il progetto è iniziato in modo "classico", conducendo una campagna HTS, che però ha portato all'identificazione di molecole poco adatte ad essere ottimizzate. Si è perciò deciso di testare, mediante un saggio NMR, la collezione proprietaria di frammenti (1200 composti). Lo screening ha permesso di identificare un derivato di tipo resorcinolico con una debole attività enzimatica (Fig. 3). In questa fase non ha importanza la potenza, ma è rilevante l'efficienza di legame relativa al peso molecolare (ligand efficiency).

Per facilitare il programma di sintesi chimica è stata determinata la struttura cristallografica di questa molecola in complesso con il dominio catalitico di HSP90. Il gruppo resorcinolico occupa il sito dell'ATP, formando un network d'interazioni con la proteina mediate da diverse molecole d'acqua, mentre l'anello dicloro fenilico s'inserisce in una piccola tasca di natura idrofobica che non esiste nella proteina nella forma libera ma è indotta dalla presenza del composto. La combinazione di informazioni strutturali, modeling e attività di sintesi ha permesso di modificare il frammento, introducendo sostituenti che si estendono nel sito dell'ATP (fragment growing) e di identificare un composto potente, con buone caratteristiche di farmacocinetica, adatto ad un successivo sviluppo clinico (Fig. 3).

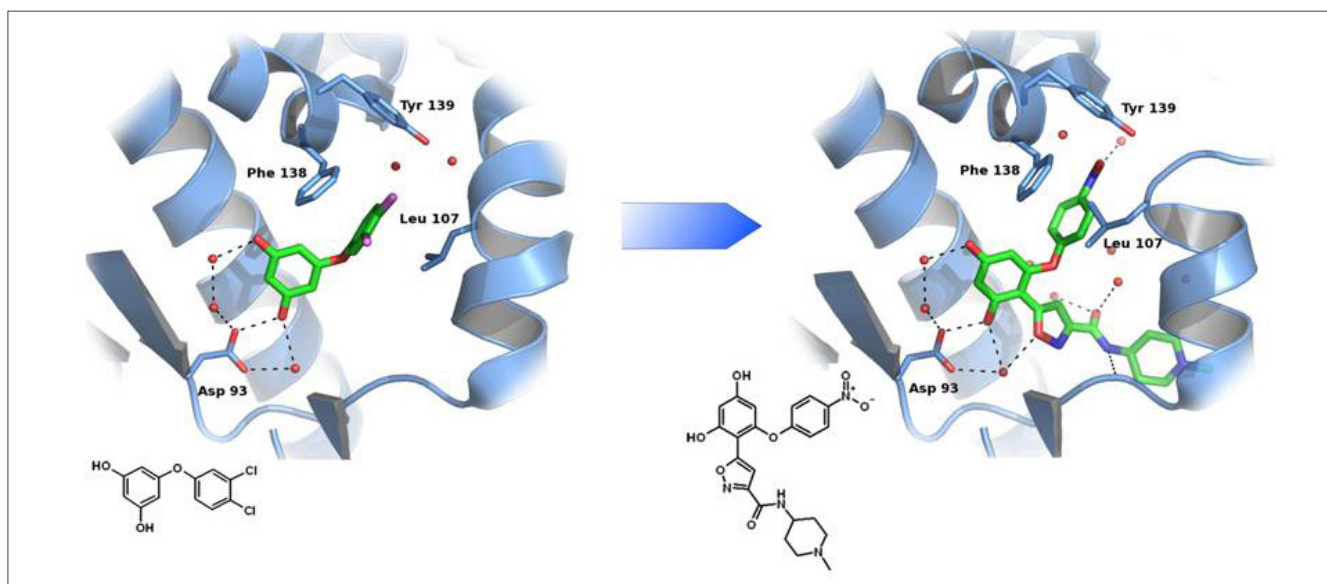


Fig. 3 - Sviluppo FBDD di inibitori di HSP90: modo di legame del frammento iniziale e della molecola identificata alla fine del processo di ottimizzazione

### Le sfide della cristallografia

La cristallografia rimane indubbiamente la metodologia principale per la determinazione di strutture di macromolecole biologiche. Tuttavia, la tecnica presenta alcune fondamentali limitazioni. L'ostacolo principale è rappresentato dalla necessità di ottenere cristalli della proteina bersaglio. Se da un lato vi sono stati enormi progressi nelle tecniche di cristallizzazione, d'altro canto l'interesse della ricerca si è progressivamente spostato verso target sempre più ambiziosi, come i grandi complessi proteici e le proteine di membrana, la cui cristallizzazione è ben lontana dal diventare una routine. Non va inoltre dimenticato che il modello ottenuto mediante raggi X è un'immagine statica, che non dà informazioni sulla dinamica della proteina e sulla sua flessibilità, aspetti che in alcuni casi sono determinanti per la comprensione del meccanismo d'azione del farmaco. L'ambizione di conoscere a fondo le proprietà di dinamica molecolare di un target è tendenzialmente in aumento per poter meglio valutare le probabilità e gli effetti di cambi conformazionali della proteina. Una tecnica complementare alla cristallografia, da questo punto di vista, è senza dubbio l'NMR che permette di ricostruire la struttura delle diverse varianti conformazionali presenti nel campione in esame. Per lo studio di proteine o complessi di grandi dimensioni, invece, la crio-microscopia elettronica (cryo-EM) per singola particella sta emergendo come una tecnica di indagine molto potente. Inizialmente la cryo-EM era in grado di fornire modelli a bassa risoluzione, ma, recentemente, sono stati pubblicati esempi di strutture di complessi proteina-inibitore a risoluzione atomica

dove è possibile caratterizzare il modo di legame del legante [11]. Appare pertanto evidente che lo studio strutturale di target sempre più complessi non potrà prescindere dall'uso complementare e sinergico di più tecniche biofisiche.

### BIBLIOGRAFIA

- [1] J.C. Kendrew *et al.*, *Nature*, 1958, **181**, 662.
- [2] M.F. Perutz *et al.*, *Nature*, 1960 **185**, 416.
- [3] N. Roberts *et al.*, *Science*, 1990, **249**, 527.
- [4] D.K. Treiber *et al.*, *Nat. Chem. Biol.*, 2015, **11**, 818.
- [5] T. Schindler *et al.*, *Science*, 2000, **289**, 1938.
- [6] a) R.T. Bossi *et al.*, *Biochemistry*, 2010, **49**, 6813; b) M. Menichincheri *et al.*, *J. Med. Chem.*, 2016, **59**, 3392.
- [7] C.N. Johnson *et al.*, *J. Med. Chem.*, 2017, **60**, 89.
- [8] J. Tsai *et al.*, *PNAS*, 2007, **105**, 3041.
- [9] A.J. Souers *et al.*, *Nature Medicine*, 2013, **19**, 202.
- [10] M.G. Brasca *et al.*, *Bioorg. Med. Chem.*, 2013, **21**, 7047.
- [11] A. Merk *et al.*, *Cell*, 2016, **165**, 1698.

### X-Rays Shed Light on the Drug

Macromolecular X-ray crystallography is an important and powerful technique in drug discovery. The availability of high resolution structures of the target protein in complex with small molecules of interest, in fact, can boost a more focused chemical optimization of drug candidates.