

# DIPARTIMENTO DI FARMACIA



DOTTORATO DI RICERCA IN SCIENZA DEL FARMACO

XXVI CICLO

***Sviluppo di nuovi approcci computazionali per la progettazione di potenziali agenti antitumorali attraverso la modulazione delle interazioni proteina-proteina.***

COORDINATORE

CH. MA PROF.SSA MARIA VALERIA D'AURIA

CANDIDATO

DOTT. SSA ILARIA MORONESE

TUTOR

CH. MA PROF. CATERINA FATTORUSSO

## SOMMARIO

1. INTRODUZIONE .....	5
1.1 <i>Importanza delle interazioni proteina-proteina nei processi cellulari</i> .....	5
1.2 BIBLIOGRAFIA .....	9
2. Progettazione di nuovi inibitori della fostaditil inositolo-3-chinasi (PI3K) quali potenziali agenti antitumorali .....	10
2.1 INTRODUZIONE .....	10
2.1a <i>Caratteristiche generali e strutturali della PI3K</i> .....	10
2.1b <i>Attivazione della PI3K</i> .....	13
2.1c <i>La via di segnalazione PI3K: un potenziale bersaglio per lo sviluppo di più efficaci terapie antitumorali</i> .....	16
2.1d <i>Principali classi di inibitori della PI3K</i> .....	19
2.1e <i>Descrizione del sito di legame dell'ATP</i> .....	20
2.2 SCOPO E RAZIONALE DEL LAVORO .....	20
2.3 PARTE SPERIMENTALE .....	22
2.3a <i>Valutazione delle possibili Forme Tautomeriche e Calcolo del pKa degli inibitori</i> .....	23
2.3b <i>Analisi Conformazionale degli inibitori cristallizzati e del lead</i> .....	23
2.3c <i>Analisi Strutturale e Bioinformatica</i> .....	24
2.3d <i>Procedura di Docking</i> .....	25
2.3e <i>Pharmacophore Fitting del lead sull'inibitore AS-605240</i> .....	27
2.3f <i>Studi di Spettroscopia di Risonanza Magnetica Nucleare</i> .....	28
2.3g <i>Analisi Conformazionale dei nuovi composti progettati</i> .....	28
2.3h <i>Pharmacophore Fitting degli inibitori progettati su AS-605240 ed AS-604850</i> .....	28
2.3i <i>Costruzione del modello della PI3K<math>\alpha</math></i> .....	30
2.3l <i>Procedura di Docking degli inibitori progettati</i> .....	32
2.3m <i>Progettazione dei Peptidi</i> .....	32
2.4 RISULTATI E DISCUSSIONE .....	34
2.4a <i>Analisi Strutturale degli enzimi PI3K di classe Ia</i> .....	34
2.4a.1 <i>Analisi dei Domini e dei Motivi Funzionali</i> .....	37
2.4a.2 <i>Analisi del Sito di Legame dell'ATP e delle sue interazioni con i ligandi co-cristallizzati</i> .....	47
2.4a.3 <i>Analisi dell'accomodamento degli inibitori AS-604850 ed AS605240 nel sito ATPasico</i> .....	52
2.4b <i>Analisi delle Relazioni Struttura-Attività (SARs) degli inibitori noti</i> .....	56
2.4c <i>Analisi Conformazionale dei composti di riferimento e dei nuovi lead</i> .....	58
2.4d <i>Identificazione dei Requisiti Farmacoforici necessari per l'Inibizione del sito ATPasico della PI3K</i> .....	62
2.4d.1 <i>Analisi della Struttura a raggi X del Wortmannin legato al Sito ATPasico della PI3K:</i> .....	66
2.4d.2 <i>Sovrapposizione dei conformeri selezionati al wortmaninn in complesso con la PI3K</i> .....	68
2.4e <i>Analisi degli Studi di Docking sul lead</i> .....	72
2.4f <i>Ottimizzazione del lead: progettazione di nuovi inibitori potenti e specifici</i> .....	77

2.4g Studi Computazionali ed Analisi delle Relazioni Struttura-Attività (SARs) dei nuovi derivati progettati .....	80
2.4g.1 Analisi degli Studi di Docking sugli inibitori risultati più attivi .....	89
2.4h Progettazione di peptidi in grado di influenzare le interazioni con i partner proteici della p110 $\alpha$ .....	93
2.4h.1 Peptidi Progettati per inibire la Forma Non Mutata della PI3K $\alpha$ .....	94
2.4h.2 Peptidi Progettati per inibire la Forma Mutata (E542K-E545K) della PI3K $\alpha$ .....	96
2.5 CONCLUSIONI .....	98
2.6 BIBLIOGRAFIA .....	99
3. Progettazione di nuovi inibitori delle Deacetilasi Istoniche (HDAC) .....	101
3.1 INTRODUZIONE .....	101
3.1a Struttura della Cromatina e Codice Istonico .....	101
3.1b Struttura e Funzione degli HDAC .....	104
3.1c HDAC Zn <sup>2+</sup> -Dipendenti .....	107
3.1d Inibitori degli HDAC di Classe I e II .....	109
3.1e Potenziali Bersagli Terapeutici degli HDACi .....	110
3.1f Approcci Terapeutici .....	112
3.2 SCOPO E RAZIONALE DEL LAVORO .....	114
3.3 PARTE SPERIMENTALE .....	116
3.3.1a Analisi Conformazionale dei nuovi inibitori caratterizzati da un nuovo gruppo coordinante lo zinco (ZBG) .....	116
3.3.1b Analisi conformazionale dei nuovi inibitori caratterizzati da un nucleo benzodiazepinico (BDZ) .....	117
3.3.2 Calcoli DFT .....	119
3.3.3 Analisi Strutturale e Bioinformatica .....	119
3.3.4 Studi di Pharmacophore Fitting .....	120
3.3.5a Procedura di Docking dei nuovi inibitori caratterizzati da un nuovo gruppo coordinante lo zinco (ZBG) .....	121
3.3.5b Procedura di Docking dei nuovi inibitori caratterizzati da un nucleo benzodiazepinico .....	124
3.4.1 RISULTATI E DISCUSSIONE .....	125
3.4.1a Analisi Strutturale e Bioinformatica .....	125
3.4.1b Basi Razionali della Progettazione .....	131
3.4.1c Razionalizzazione delle Relazioni Struttura-Attività (SARs) dei nuovi inibitori progettati .....	136
3.4.1d Analisi dei Calcoli DFT .....	138
3.4.1e Analisi degli Studi di Docking .....	145
3.4.2 CONCLUSIONI .....	151
3.5 Progettazione di nuovi inibitori delle Deacetilasi Istoniche (HDAC) caratterizzati da un nucleo benzodiazepinico .....	152
3.5.1 RISULTATI E DISCUSSIONE .....	152
3.5.1a Implementazione del Modello Farmacoforico precedentemente progettato .....	152
3.5.1b Analisi del farmacoforo di ligandi noti in complesso con l'enzima. ....	154

3.5.1c <i>Analisi delle Relazioni Struttura-Attività (SARs) dei nuovi derivati benzodiazepinici</i> .....	157
3.5.1d <i>Studi di Pharmacophore Fitting</i> .....	168
3.5.1e <i>Analisi degli Studi di Docking</i> .....	170
3.5.2 CONCLUSIONI E PROSPETTIVE FUTURE .....	179
3.6 <i>Progettazione di nuovi inibitori delle Deacetilasi Istoniche (HDAC) caratterizzati dalla presenza di sostituenti in grado di interagire estesamente con la superficie dell'enzima</i> .....	180
3.6.1 RISULTATI E DISCUSSIONE .....	180
3.7 <i>Bibliografia</i> .....	187
4.1 INTRODUZIONE .....	192
4.2 SCOPO DEL LAVORO .....	195
4.3 PARTE SPERIMENTALE .....	196
4.3a <i>Analisi Conformazionale</i> .....	196
4.3b <i>Studi di Docking</i> .....	196
4.4 RISULTATI E DISCUSSIONE .....	198
4.4a <i>Analisi Conformazionale</i> .....	198
4.4b <i>Risultati degli Studi di Docking</i> .....	200
4.5 CONCLUSIONI .....	202
4.6 BIBLIOGRAFIA .....	203

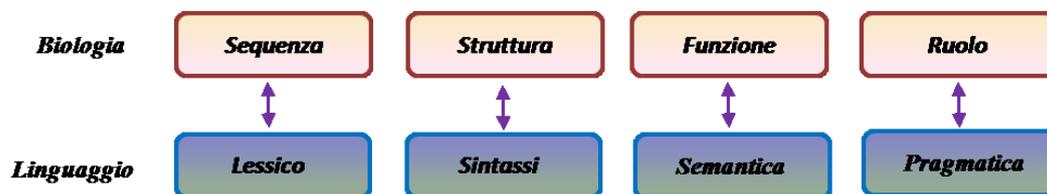
# ***1. Introduzione***

## ***1.1 Importanza delle interazioni proteina-proteina nei processi cellulari***

Molti dei più importanti processi cellulari come, ad esempio, la replicazione del DNA, avvengono in “macchine molecolari” costituite da un gran numero di componenti proteici, strutturalmente e funzionalmente organizzati attraverso interazioni proteina-proteina. Le interazioni proteiche sono state studiate dal punto di vista della biochimica, bioinformatica, dinamica molecolare e trasduzione del segnale. Le *interazioni proteina-proteina* (PPI) sono il core dei sistemi di interazione di qualsiasi cellula vivente. Tali interazioni sono importanti per la maggior parte delle funzioni che svolgono processi biologici: quali la trasmissione dei segnali provenienti dall'esterno, l'adesione cellulare, lo sviluppo e la comunicazione tra le cellule. Le informazioni su queste interazioni migliorano la comprensione sulle malattie e possono fornire le basi di un nuovo approccio terapeutico. Le PPI avvengono quando due o più proteine si legano spesso per svolgere la loro funzione biologica. Specifiche proteine svolgono un ruolo cruciale per mediare queste interazioni e reclutare recettori, enzimi, adattatori in determinati compartimenti cellulari. Negli eucarioti, l'attivazione della cascata di trasduzione del segnale richiede interazioni multiple al fine di costituire complessi proteina-proteina e proteina-lipide in specifici siti all'interno della cellula; tutti gli stimoli esterni e i segnali che vengono percepiti dalle cellule sono successivamente propagati all'interno attraverso interazioni proteiche con una forte gerarchia spaziale e temporale che permette alla cellula stessa di comportarsi e rispondere in maniera adeguata rispetto all'ambiente circostante riscontrando sempre un equilibrio tra ambiente esterno e nucleo. Lo studio della propagazione intracellulare dei segnali è attualmente arrivato a descrivere in maniera minuziosa gli eventi biochimici iniziali, l'attivazione delle chinasi dei recettori, i segnali dovuti a fosforilazioni di serine, treonine e tirosine, la partecipazione di adattatori che riconoscono questi segnali e partecipano addensando proteine nel punto richiesto. Le PPI rappresentano quindi un aspetto fondamentale delle funzioni proteiche, le funzioni di una proteina infatti non emergono dalla semplice sequenza aminoacidica ma si possono definire attraverso l'analisi dei suoi *partners* funzionali<sup>1</sup>. Le funzioni cellulari sono basate proprio su transitori o permanenti legami fisici tra due o più proteine, al fine di realizzare i rispettivi compiti; le caratteristiche generali, la specificità di tali interazioni di legame e la relazione tra la conformazione proteica e il legame stesso rivestono un'importanza cruciale. In particolare, la creazione di una mappa d'interazioni proteina-proteina di una cellula può essere di gran

valore per la comprensione della biologia della cellula stessa. L'approccio allo studio del genoma attraverso gli strumenti della proteomica ha portato all'accumulo di una grande mole di informazioni riguardo la concentrazione e distribuzione delle molecole biologiche e le loro relative interazioni. Lo studio sistematico delle interazioni proteina-proteina è uno strumento altamente efficace per l'annotazione funzionale dei geni e delle proteine da essi codificate e per la caratterizzazione degli interattomi nella ricerca post-genomica. L'accezione 'interattoma' comprende la totalità delle interazioni molecolari che hanno luogo in un organismo o, più in dettaglio, che permettono lo svolgersi di una cascata molecolare regolativa, metabolica, ecc. o, ancora, che fanno parte del meccanismo d'azione di una proteina (es. un recettore, un enzima, un fattore di trascrizione, ecc.). Il numero sempre crescente di strutture tridimensionali risolte a livello atomico, ha permesso di allargare lo studio delle strutture proteiche e delle loro interazioni su larga scala e l'identificazione sperimentale di numerose coppie di proteine interagenti<sup>2</sup>. Tuttavia, il meccanismo d'interazione e i residui coinvolti non sono ancora, in generale, noti e la maggior parte delle interazioni deve ancora essere strutturalmente caratterizzata. Per arrivare ad una conoscenza più dettagliata dei meccanismi alla base dei processi biologici è quindi di fondamentale importanza lo studio della struttura proteica e il riconoscimento delle sue superfici d'interazione, in quest'ottica, si rivela molto utile lo sviluppo di metodi computazionali per la predizione dei siti di interazione<sup>3</sup>. La predizione delle PPI presuppone una dettagliata analisi delle sequenze e soprattutto delle strutture ed ha due problemi principali: il numero d'interazioni possibili è estremamente elevato e i modi d'interazione possibili sono molteplici (omocomplessi, eterocomplessi, complessi obbligati o complessi transienti). La maggior parte di queste interazioni scaturiscono dal fatto che le proteine sono strutturate in modo modulare per mezzo di una specifica combinazione di *domini funzionali*, tali domini guidano i segnali dei polipeptidi verso specifici complessi multi-proteici reclutandoli e limitandoli in un'appropriata localizzazione subcellulare (sito specifica)<sup>4</sup>. Tipicamente, i domini d'interazione proteina-proteina il cui ripiegamento (*folding*) è altamente conservato indipendentemente dalla proteina stessa, di 35-150 amminoacidi, possono essere espressi da soli rispetto alle proteine di cui fanno parte; mantenendo l'abilità intrinseca di legare i loro partners fisiologici. Tutte le interazioni proteina-proteina coinvolgono la capacità di un dominio foldato d'interazione di riconoscere un breve motivo peptidico per mezzo di amminoacidi altamente conservati. In particolare, mentre *i motivi* sono combinazioni regolari di strutture secondarie, o meglio sono una combinazione specifica di strutture secondarie con una topologia particolare organizzati in una struttura tridimensionale caratteristica, (es. coiled

– coil, elica – *loop* – elica, dita di zinco, ecc...) i *domini strutturali e funzionali* sono moduli di strutture terziarie caratterizzati da aspetti interessanti come l'abbondanza insolita di un dato aminoacido (un dominio ricco di prolina, un dominio acido, un dominio ricco di glicina). È divenuto ormai di uso comune parlare di un vero e proprio linguaggio delle proteine per indicare una chiara e precisa comunicazione tra esse tanto da affermare che il linguaggio umano può essere considerato come una trasformazione del linguaggio delle cellule, se tale teoria è perseguibile, è esplicitativo paragonare la sintassi con la semantica e il rapporto struttura–funzione (*Figura 1*), così come è interessante che la complessità della linguistica computazionale abbia sviluppato tecniche di calcolo usate non solo per interpretare i dati biologici ma anche per studiare ed investigare la lingua e la grammatica più a fondo<sup>5</sup>.



*Figura 1: Rappresentazione schematica e chiarificatrice della notevole somiglianza tra il linguaggio delle proteine (riga superiore) e quello umano (riga inferiore): importanza del valore semantico delle informazioni che sono depositate in una frase o in una proteina in quanto definisce un ruolo biologico o pragmatico in un contesto molto più ampio (Figura riprodotta da Nature 7, 68-73 (2006)).*

La comprensione della struttura e della funzione delle proteine, unita alla possibilità di influenzarne l'interazione a livello molecolare, rappresenta una grande sfida per la comunità scientifica, motivata, tra l'altro, anche dalle applicazioni nella progettazione di farmaci e nelle nano biotecnologie. In conclusione, le PPI sono di importanza centrale per ogni processo attivo in una cellula vivente, le informazioni che scaturiscono dallo studio di queste interazioni possono migliorare la comprensione di un gran numero di processi fisiologici attivati ad esempio in maniera aberrante e possono quindi fornire la base per nuovi metodi terapeutici.

Sulla base di queste considerazioni il mio progetto di dottorato ha riguardato lo sviluppo di nuovi approcci computazionali per la progettazione di potenziali agenti antitumorali che vadano ad inibire il sito catalitico estendendosi anche sulla superficie della proteina impedendo il legame con proteine effettrici; e/o stabilizzare la conformazione inattiva dell'enzima fungendo da ligandi-bi-adesivi modulando, quindi, in entrambi i casi le interazioni proteina-proteina .

Mi sono occupata della cascata di attivazione *fostaditil inositolo-3-chinasi* (PI3K)-Akt e degli enzimi *istone de-acetilasi* (HDAC). In particolare, la cascata di attivazione PI3K-Akt è

coinvolta nella regolazione di un elevato numero di processi cellulari quali: trascrizione, migrazione, angiogenesi, crescita cellulare, apoptosi e metabolismo del glucosio. La maggior parte delle cellule tumorali presentano mutazioni a livello di questo pathway che consentono una continua attivazione degli enzimi e quindi la continua proliferazione delle cellule mutate. Invece, gli HDAC che sono una famiglia di metalloproteasi, deacetilano gli istoni nucleosomali giocando un ruolo importante nelle funzioni del genoma come l'assemblaggio della cromatina, la riparazione e la ricombinazione del DNA. Numerose altre proteine non istoniche si sono rivelate come il substrato degli HDAC compresi fattori di trascrizione, proteine citoscheletriche ed alcuni importanti fattori nucleari. Questo espande e complica le funzioni degli HDAC e indica che la regolazione della loro attività può essere utile in molte terapie come cancro, infezioni virali, malattie neurodegenerative, infiammazioni, etc. Di seguito verrà riportato, in dettaglio, il lavoro svolto su ciascun progetto.

## ***1.2 Bibliografia***

- (1) Ofra Y, Rost B “*Analysing six types of protein-protein interfaces*” J Mol Biol 325, 377-387(2003)
- (2) Chothia C, Janin J “*Principles of protein-protein recognition*” Nature, 256, 705-8 (1975).
- (3) Jones S, Thornton JM “*Principles of protein-protein interactions*” Proc Natl Acad Sci U S A, 93, 13-20 (1996)
- (4) Cohen G. B. et al. “*Modular binding domains in signal transduction proteins*” Cell 80, 237–248 Review (1995)
- (5) Baker, M. C. “*The atoms of language*” Basic books, New York (2001)

## ***2. Progettazione di nuovi inibitori della fosfatidil inositolo-3-chinasi (PI3K) quali potenziali agenti antitumorali***

### ***2.1 Introduzione***

#### ***2.1a Caratteristiche generali e strutturali della PI3K***

Le fosfatidil inositolo-3-chinasi (PI3Ks) sono chinasi lipidiche ubiquitarie, ampiamente espresse nell'organismo, che agiscono sia come trasduttori del segnale in differenti vie cellulari a seguito della stimolazione dei recettori presenti sulla superficie cellulare, sia nei processi di smistamento delle proteine. Tutte le PI3Ks sono enzimi a specificità duale; possiedono infatti un'attività lipidico-chinasica necessaria per la fosforilazione in posizione 3' dell'anello inositolico del fosfatidil inositolo-4,5-bifosfato (PI(4,5)P<sub>2</sub>), ed un'attività proteino-chinasica<sup>1</sup>.

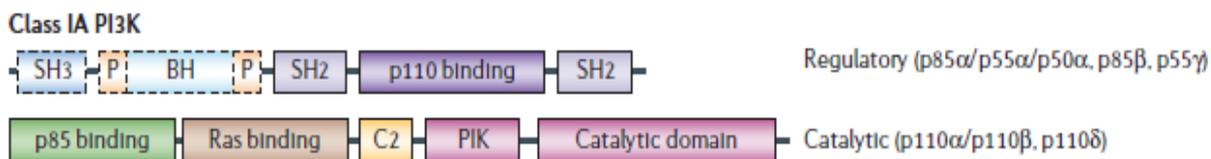
La famiglia delle PI3K comprende otto membri divisi in tre classi sulla base dell'omologia di sequenza, della specificità di substrato e sulla base dei domini proteici che la costituiscono; nelle cellule umane in particolare sono espressi tre gruppi rappresentativi, di questi la classe I è quella meglio caratterizzata e comprende una serie di elementi chiave per l'integrazione di diversi segnali cellulari. Le PI3K di classe I risultano essere molecole-chiave della segnalazione intracellulare, che integrano una grande varietà di segnali attivati da alcuni fattori di crescita polipeptidici, i quali, legandosi a recettori specifici di membrana, ovvero i recettori tirosin chinasi, ne guidano l'attivazione mediante treonin-chinasi, proteine G eterotrimeriche o Ras.

La classe I è caratterizzata da quattro membri, identificati e sub-classificati sulla base del loro meccanismo d'azione, quali: enzimi eterodimerici regolati da recettori, che in vivo preferenzialmente fosforilano il fosfatidil inositolo 4,5-trifosfato (PI(4,5)P<sub>2</sub>) trasformandolo in fosfatidil inositolo 3,4,5-trifosfato (PIP<sub>3</sub>) e sono costituiti da una subunità catalitica chiamata p110 e da una subunità con funzione di regolatore/adattatore chiamata p85 o p55, la quale serve ad ancorare la PI3K alla membrana plasmatica mediante l'interazione dei domini SH-2 con i residui fosfotirosinici dei recettori attivati<sup>2</sup>.

La classe I viene ulteriormente suddivisa in due sottoclassi, Ia e Ib.

Le PI3K di classe Ia vengono attivate da recettori ad attività tirosina-chinasica (RTK) e sono composte da una subunità catalitica del peso di 110 kDa (p110 $\alpha$ ), e da una subunità

regolatoria del peso di 85KDa (p85 $\alpha$ ). La subunità p110 è codificata da tre differenti geni (p110 $\alpha$ , p110 $\beta$ , p110 $\delta$ ) e contiene diversi domini tra cui: l'ABD che è localizzato vicino ad una delle estremità del complesso, contatta l'iSH2 nella regione N-Terminale, il dominio C2 (proteinkinase-C homology-2), che è stato proposto legare la membrana cellulare, interagisce con l'iSH2 nella regione C-Terminale (*Figura 1*).



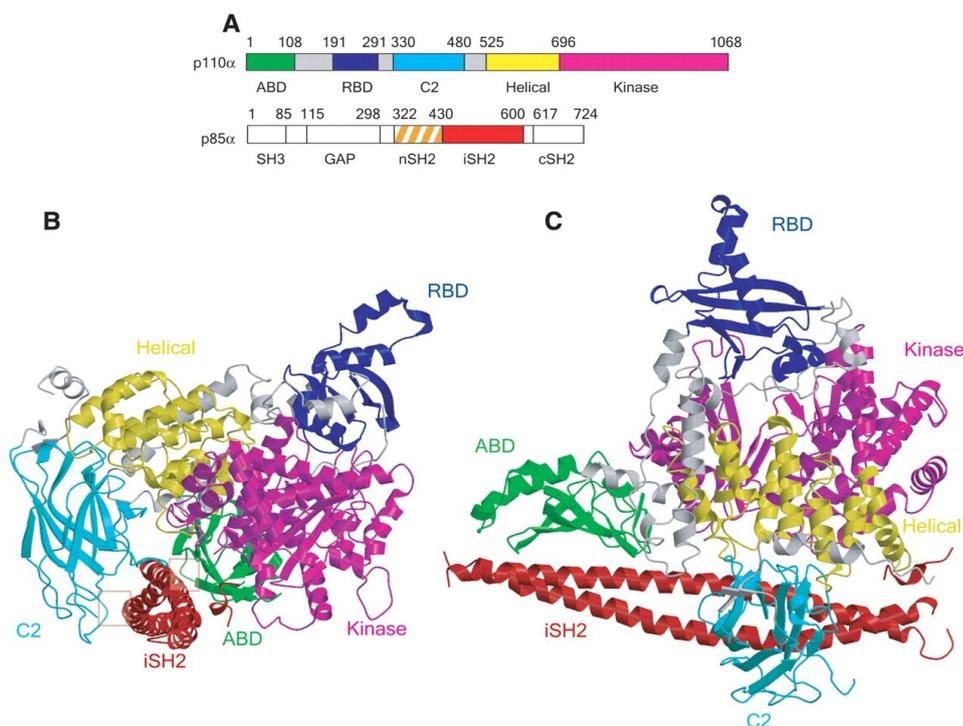
**Figura 1: Domini costituenti le subunità, regolatoria e catalitica, delle PI3K di classe Ia**

Il dominio ad elica e quello chinasi formano il grosso del centro del complesso ma non si trovano in diretto contatto con l'iSH2 della p85. Il dominio di legame a Ras (RBD), contiene il dominio chinasi ed è situato dal lato opposto del dominio ad elica. L'RBD così come l'ABD sono entrambi piccoli domini globulari di circa 100 residui di lunghezza, un lungo linker contenente quattro eliche connette questi due domini. Il dominio ad elica con funzione ancora non del tutto chiara, si ripiega come una struttura completamente ad  $\alpha$ -elica; nel complesso p110 $\alpha$ /niSH2 il ruolo principale di tale dominio è quello di fare da ponte tra il dominio C2 e il dominio chinasi che contiene il sito catalitico dell'enzima, gli unici due domini con il quale è in diretto contatto. Quest'ultimo dominio è in particolare una struttura costituita da due lobi: N-terminale e C-terminale, a sua volta formati da 2 $\alpha$ / $\beta$  sub-domini separati da una fenditura; questa è un'architettura tipica delle protein-chinasi lipidiche. La subunità regolatoria invece non ha attività lipidico-chinasi ma contiene dei domini che fungono da modulatori e che possono essere separati funzionalmente e spazialmente dal resto della proteina. In particolare la subunità p85 contiene un dominio N-terminale SH3 che media specifiche interazioni proteina-proteina ed è seguito da un dominio ricco di proline che è un potenziale partner di interazione del dominio SH3; un dominio BH (breakpoint cluster dominio regione omologa) responsabile dell'attività di GAP (famiglia di proteine regolatorie i cui membri possono legare e attivare proteine G stimolando la loro attività GTPasica) che interagisce con Cdc42, Rac1 (piccole proteine G) e riconosce una sequenza simile a quella riconosciuta dalle proteine 14\_3\_3; un secondo dominio ricco di proline anch'esso potenziale partner del dominio SH3 e due domini SH2 coinvolti nell'attivazione/inattivazione della subunità catalitica e legati a sua volta ad un dominio coiled-coil (iSH2 o inter-SH2-domain) (*Figura 2*). Il dominio iSH2 lega la regione N-terminale della subunità catalitica e contiene il residuo Ser608<sup>3</sup>, che può essere fosforilato dalla subunità catalitica p110 $\alpha$

riducendo l'attività lipidico-chinasica. Questo è un meccanismo potenzialmente importante per la regolazione dell'attività della PI3K, in quanto è stato provato che induce un decremento della regolazione dell'attività lipidico-chinasica in cellula.

Ci sono otto isoforme di p85 $\alpha$  generate da fenomeni di splicing alternativo a partire dai prodotti di tre geni, tutte contengono i due domini SH2 (Src omologhi) che interagiscono con fosfotirosine sul RTK attivato; provocando il reclutamento di enzimi presso la membrana plasmatica e l'attivazione dell'attività catalitica della PI3K.

Nel suo stato basale, la subunità p110 $\alpha$  è vincolata ed inibita dalla subunità adattatrice p85 $\alpha$ ; quando sono presenti appropriati stimoli cellulari, i domini nSH2 e cSH2 di p85 $\alpha$  legano fosfotirosine di recettori attivati e il legame a queste tirosine fosforilate attiva la subunità p110 $\alpha$ . Diversi studi sul complesso p110 $\alpha$  /p85 $\alpha$  hanno mostrato che la p85 $\alpha$  inibisce la p110 $\alpha$  e che invece l'attivazione dei fosfopeptidi contribuisce all'attivazione del dimero p110 $\alpha$ /p85 $\alpha$  riflettendo così una disinibizione di p110 $\alpha$  da parte di p85 $\alpha$ . Il dominio iSH2 della subunità regolatoria p85 $\alpha$  è sufficiente per legare la subunità p110 $\alpha$  ma non per inibirla; l'inibizione richiede la presenza del dominio nSH2 legato a quello iSH2, ciò suggerisce che l'nSH2 è un dominio coinvolto direttamente nella regolazione di p110 $\alpha$  <sup>4</sup>.



**Figura 2: Schema dell'organizzazione dei domini della subunità catalitica e della subunità regolatoria di PI3K di classe Ia.** In figura A vengono messi in evidenza con colori differenti i domini della subunità catalitica p110 $\alpha$  e quelli della subunità regolatoria p85 $\alpha$  di PI3K e vengono inoltre mostrate in figura B e C due diverse disposizioni degli stessi domini.

Le PI3K di classe Ib, invece, non posseggono la subunità regolatoria p85 necessaria per il legame con il RTK, per cui sembrano essere attivate esclusivamente da GPCRs. Similmente alla classe Ia, si tratta di eterodimeri costituiti da una subunità regolatoria p101 e una subunità catalitica p110 $\gamma$ <sup>5</sup>. La subunità p110 $\gamma$  è espressa nei leucociti ed è attivata dal legame con la subunità G $\beta\gamma$  delle proteine G, interazione mediata dalla subunità regolatoria p101.

La classe II delle PI3K comprende tre membri (PI3KC2  $\alpha$ ,  $\beta$ ,  $\gamma$ ) ed è caratterizzata da un dominio C2 carbossi terminale che lega fosfolipidi e risultano essere una famiglia di enzimi che utilizza PI e PI(4)P come substrato. Gli enzimi di classe II sono legati principalmente alla membrana e attivati da recettori di membrana inclusi RTK e integrine.

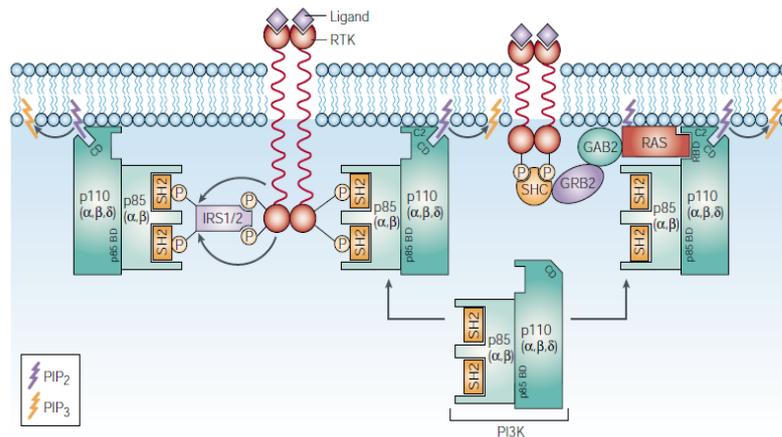
La classe III comprende una chinasi detta VPS34p responsabile della produzione di PIP3 e coinvolta nel traffico di proteine attraverso i lisosomi, sono enzimi eterodimerici costituiti da una subunità catalitica (Vps34, 100 kDa) e una subunità adattatrice (p150)<sup>6</sup>.

La mia attenzione è stata posta solo sulla classe Ia che è quella meglio caratterizzata e che rappresenta un importante bersaglio farmacologico in quanto principalmente mutata e responsabile della formazione di cellule tumorali.

### ***2.1b Attivazione della PI3K***

La maggior parte delle conoscenze riguardo il meccanismo di trasduzione del segnale della PI3K deriva dagli studi condotti sulla classe Ia, che è quella che è stata meglio caratterizzata, e comprende una serie di elementi chiave per l'integrazione di diversi segnali cellulari.

Il meccanismo che porta all'attivazione delle PI3K di classe Ia prevede che in condizioni di riposo il complesso p85 $\alpha$ -p110 $\alpha$  sia presente in forma inattiva all'interno del citoplasma e che, in seguito a stimolazione dei RTK da parte di specifici ligandi, i recettori dimerizzino e si fosforilino reciprocamente sui residui di tirosina presenti a livello delle loro porzioni citoplasmatiche. La dimerizzazione del recettore e la sua attivazione per fosforilazione permettono, quindi, il reclutamento alla membrana plasmatica del complesso p85 $\alpha$ -p110 $\alpha$  della PI3K. Il legame della subunità regolatoria p85 $\alpha$  al recettore determina a sua volta variazioni conformazionali della PI3K che attivano la subunità catalitica oltre a permetterle di raggiungere il suo substrato lipidico. (*Figura 3*)



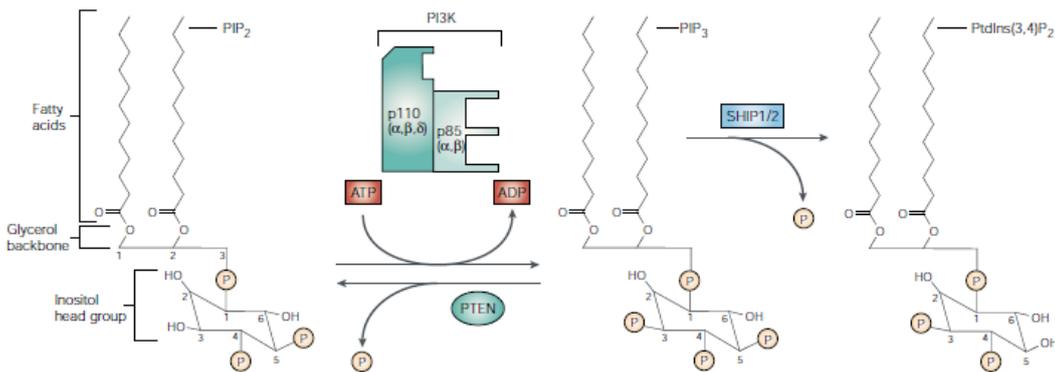
**Figura 3: Modello di attivazione delle PI3K di classe Ia.**

L'auto-fosforilazione del RTK, attivato dall'interazione con uno specifico ligando, determina il reclutamento dell'eterodimero inattivo p85-p110 da parte del recettore, ciò avviene attraverso l'interazione dei domini SH2 della p85 con residui di fosfotirosina della sequenza consenso pYXXM presente sul RTK attivato o, in alcuni casi, su molecole "adattatrici" come IRS1/IRS2. Queste ultime vengono fosforilate dal recettore attivato, generando così i siti di ancoraggio per i domini SH2 della p85.

L'interazione SH2-fosfotirosina porta la PI3K in prossimità del suo substrato sulla membrana plasmatica e determina una riduzione dell'azione inibitoria che la p85 esercita sulla p110 allo stato basale; la p110 quindi, a questo punto, è libera di convertire il PIP2 in PIP3.

In alternativa, il legame della PI3K a RAS-GTP può determinare il reclutamento dell'enzima sulla membrana e l'attivazione del suo dominio catalitico; ciò si realizza mediante il reclutamento di proteine adattatrici, quali SHC, GRB2 e GAB2 da parte del RTK attivato.

La conseguenza principale dell'attivazione della PI3K è la genesi del PIP3, a partire dal PIP2, nella regione intracellulare della membrana plasmatica. Il PIP3 è un potente secondo messaggero di natura lipidica, responsabile dell'attivazione di numerose molecole a valle attraverso il legame con i loro domini PH (domini pleckstrin omologhi). (Figura 4)

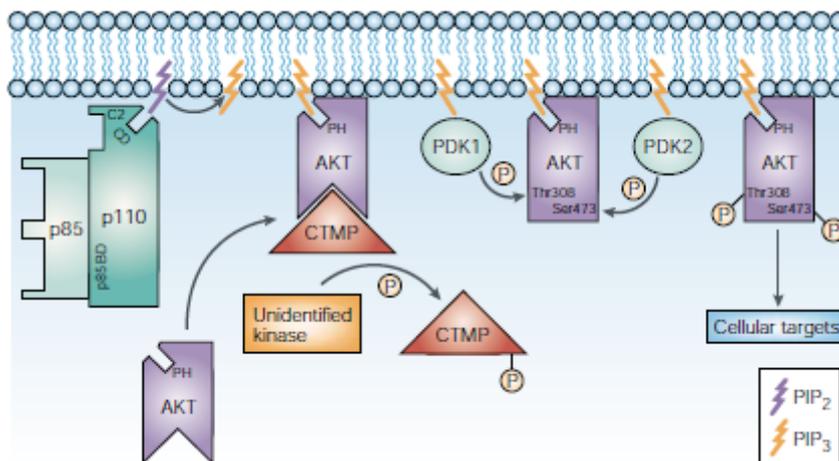


**Figura 4: Ciclo PtdIns(4,5)P2-PtdIns(3,4,5)P3.**

Il fosfatidilinositol-fosfato è classificato come un glicerol-fosfolipide, costituito da uno scheletro di glicerolo, due code apolari di acidi grassi e una testa polare di inositolo fosforilato.

La PI3K può fosforilare il PIP2 nella posizione D3 per formare il secondo messaggero PIP3; tale fosforilazione è necessaria per il legame al dominio PH della chinasi AKT (non in figura). La defosforilazione del PIP3 per rigenerare il PIP2 è operata, invece, dalla fosfatasi PTEN. Inoltre il PIP3 può essere anche defosforilato nella posizione D5 dalle proteine SHIP1 e SHIP2 per generare PtdIns (3,4) P2, un altro potenziale secondo messaggero.

Una di queste è la serin/treonin chinasi AKT (nota anche come PKB) che rappresenta il principale target del PIP3. (Figura 5)



**Figura 5: Attivazione e regolazione dell'attività dell'AKT.**

L'attivazione dell'AKT ha inizio con la traslocazione della proteina dal citoplasma alla membrana, che si verifica dopo la stimolazione cellulare e la genesi del PIP3. Il reclutamento dell'AKT sulla membrana plasmatica si realizza attraverso una specifica interazione tra il suo dominio PH e il PIP3. A livello della membrana, l'associazione con la proteina CTMP impedisce la fosforilazione dell'AKT e, conseguentemente, la sua completa attivazione. La fosforilazione di CTMP da parte di una chinasi non ancora identificata, porterebbe al distacco della proteina dall'AKT, permettendo dunque a quest'ultima di poter essere fosforilata dalle chinasi PDK1 e PDK2 sui residui di Thr308 e Ser473, rispettivamente; la fosforilazione di questi due residui porta a una completa attivazione della chinasi.

Il legame del PIP3 all'AKT porta al reclutamento di quest'ultima a livello della membrana e la sua conseguente fosforilazione da parte del complesso Rictor-mTOR (bersaglio della rapamicina nei mammiferi) e da parte di PDK1 (3-chinasi fosfoinositide dipendente); ciò porta alla piena attivazione della AKT che, a sua volta, attraverso la fosforilazione di numerose proteine target regola una vasta gamma di processi cellulari<sup>5</sup>.

I livelli di PIP3 sono appena rilevabili nelle cellule di mammifero in condizioni di crescita non stimolata e sono rigorosamente controllati, sia dall'azione della PI3K, sia dall'azione di una serie di fosfatasi, tra cui PTEN, SHIP1 e SHIP2. In particolare, quella che sembra essere più chiaramente coinvolta nel processo di oncogenesi è *PTEN* (fosfatasi e tensina omologa chiamata anche MMAC1), una fosfatasi lipidica che catalizza la conversione del PIP3 in PIP2 inibendo così l'attivazione dell'AKT ad opera della PI3K e agendo, quindi, come regolatore negativo di questa via di segnalazione. Evidenze sperimentali dimostrano che p110 $\alpha$  può essere attivata anche in maniera indiretta attraverso l'intervento di molecole adattatrici, tra cui GRB2 (fattore di crescita del recettore legato alla proteina 2) o di Ras. La proteina GRB2 è in grado di legare le tirosine fosforilate di RTK e reclutare a sua volta la famiglia di molecole adattatrici GAB (proteine leganti GRB2 associati) che legano anche la p85 $\alpha$ <sup>7</sup>. La proteina Ras, invece, è una switch protein di circa 189 amminoacidi con

attività GTPasica, è in particolare una proteina G monomericata formata da  $\alpha$ -eliche e  $\beta$ -sheets unite tra loro tramite regioni di legame (*loops*) e dotata di due regioni switch I (32-40) e switch II (60-67) che cambiano la loro conformazione a seconda se siano legate a GDP o GTP; Ras si alterna infatti tra uno stato "acceso" in cui è legata al GTP ad uno "spento" in cui è legata a GDP. Ras in complesso con GTP si lega alla conformazione cataliticamente attiva di p110 attraverso le regioni dello switch I e switch II, per cui l'attivazione della subunità catalitica p110 in seguito a stimolazione di RTK e l'attivazione di Ras sono sinergistiche. Quest'ultima è accelerata dal fattore di scambio dei nucleotidi guaninici (GEF), il quale si lega a Ras permettendo il distacco del GDP. Poiché la concentrazione cellulare di GTP è nettamente superiore a quella del GDP, il GTP si lega liberamente alle proteine Ras lasciate "vuote" permettendo il rilascio di GEF. La sua disattivazione è invece regolata dalla proteina GAP che ha attività GTPasica<sup>8</sup>.

### ***2.1c La via di segnalazione PI3K: un potenziale bersaglio per lo sviluppo di più efficaci terapie antitumorali***

Il processo di cancerogenesi è un evento complesso e multifasico in cui sono coinvolte diverse vie di trasmissione di segnale connesse tra di loro, le quali regolano la sopravvivenza e la crescita cellulare. Componenti che contribuiscono al processo di segnalazione cellulare che stimola la crescita, morte e divisione cellulare svolgono un ruolo di rilevante importanza medica nel tentativo di identificare gli squilibri molecolari che causano cancro e che alterano la morte cellulare programmata<sup>9</sup>.

Le nuove conoscenze sui meccanismi coinvolti nel processo di trasformazione e progressione tumorale, ed il riconoscimento delle proteine coinvolte nella regolazione di questi processi, ha aperto una nuova era nella formulazione e nella valutazione clinica di nuovi farmaci, le cosiddette terapie mirate. Le proteine che regolano la proliferazione, la differenziazione, l'apoptosi e l'invasività cellulare sono alla base della trasformazione neoplastica e sono il bersaglio di questo nuovo approccio terapeutico. Le nuove linee di ricerca farmacologica si sono quindi rivolte all'identificazione di agenti in grado di interferire in maniera selettiva contro bersagli molecolari specifici (terapia mirata), al fine di aumentare la selettività del bersaglio e di ridurre gli effetti collaterali sistemici.

In questo scenario, l'attivazione non controllata della via di trasduzione del segnale PI3K porta ad un alterato controllo della duplicazione cellulare e dell'apoptosi, con conseguente proliferazione incontrollata da parte delle cellule tumorali<sup>10</sup>.

Questo pathway gioca un ruolo di fondamentale importanza non solo nello sviluppo tumorale ma anche nella potenziale risposta al trattamento del cancro. Le più comuni alterazioni a valle della via di segnalazione PI3K comprendono amplificazioni/mutazioni attivanti, incrementi dell'espressione dei geni che codificano per i recettori tirosin-chinasici, aumento dell'espressione dei geni che codificano per i fattori di crescita.

La maggiore alterazione descritta riguarda l'amplificazione del gene per PI3K; il gene PI3KCA, che codifica per la subunità catalitica p110 $\alpha$  ed è localizzato sul cromosoma 3q26, una regione che è frequentemente amplificata in diversi tumori umani, in particolare mutazioni abituali si hanno nel colon umano, glioblastoma, polmone e mammella. Le più frequenti mutazioni di p110 $\alpha$  sono E542K, E545K e H1047R. Tali mutazioni incrementano l'attività enzimatica, stimolano il segnale di Akt e stimolano fattori di crescita causando tumori in vivo.

L'attività indotta dalle mutazioni E542K e E545K, è altamente dipendente dal legame Ras-GTP, mentre il guadagno di funzione enzimatica indotta da H1047R è forse dipendente dal cambiamento allosterico mediato da p85 e questo cambiamento mima il legame Ras-GTP, rendendo l'attività catalitica di questo mutante indipendente dall'interazione con Ras-GTP.

Studi biochimici e strutturali supportano la teoria che il dominio d'elica mutato E542K e E545K, impedisca un'interazione inibitoria tra il dominio SH2 N-terminale di p85 e il dominio ad elica di p110 $\alpha$ . D'altra parte, nel mutante H1047R, la regolazione del pathway Akt appare essere sufficiente per l'induzione della trasformazione oncogenica da parte di PI3K.

Per esaminare gli effetti del legame di p85 nell'attività del mutante, sono stati eliminati i legami di p110 $\alpha$  a p85 mediante delezione di 72 aminoacidi N-terminali da p110 $\alpha$ .

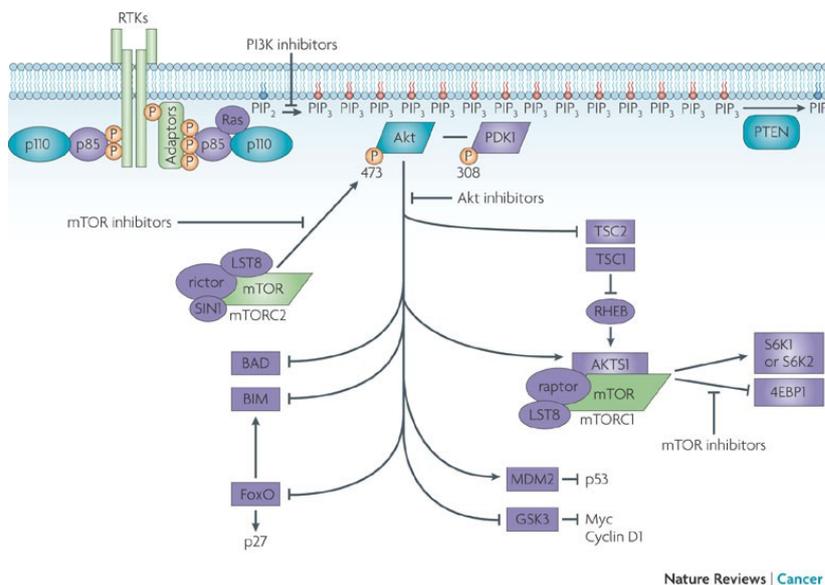
L'eliminazione del sito di legame a p85 ha un debole effetto sulla trasformazione dell'attività del dominio ad elica mutato, ma è completamente abolita la trasformazione indotta dalla mutazione del dominio chinasico H1047R. Tali risultati suggeriscono che un processo non ancora identificato dipendente da p85, gioca un ruolo critico nell'oncogenicità di H1047R<sup>11</sup>.

La presenza delle mutazioni E542K, E545K e H1047R costituiscono una delle principali cause di insorgenza di tumori (*Tabella 1*) e sono correlati alla resistenza che si instaura nei confronti di diversi farmaci anti-tumorali.

**Tabella 1:** Frequenza delle mutazioni a carico del pathway PI3K nel cancro

Mutazioni genetiche	Tipi di cancro	Percentuale di frequenza
<b>PI3KCA (p110<math>\alpha</math>)</b>		
<b>Mutazioni</b>	Seno	26% (176/684)
	Colon	26% (88/337)
	Glioma	8% (14/182)
	Epatocellulare	36% (26/73)
	Ovarico	10% (35/365)
	Polmone	2% (4/253)
	Gastrico	7% (24/338)
<b>Amplificazioni</b>	Testa e Collo	42% (54/128)
	Tiroide	9% (12/128)
	Polmone:	
	Cellule squamose	66% (46/70)
	Adenocarcinoma	5% (4/86)
	Seno	9% (9/92)
	Gastrico	36% (20/55)
	Adenocarcinoma esofageo	6% (5/87)
	Cervicale	69% (11/16)

Attualmente, molti composti, hanno come bersaglio l'asse di sopravvivenza di PI3K/AKT/mTOR; in particolare sono state sviluppate molecole che vanno ad inibire PI3K, AKT ed mTOR sia allosteriche (rapamicina e derivati) che del sito catalitico (Figura 6)<sup>12</sup>.



**Figura 6:** Rappresentazione schematica dell'inibizione dell'asse di sopravvivenza PI3K/Akt/mTOR: in figura vengono messi in evidenza le possibili e alternative vie di inibizione a carico delle varie fasi e dei vari componenti coinvolti nel pathway PI3K/Akt/mTOR.

In conclusione, essendo la funzione di PI3K quella più comunemente alterata nei tumori umani e nella chemio-resistenza, la sua inibizione risulta essere un bersaglio attraente per lo sviluppo di nuovi farmaci nella terapia del cancro.

## ***2.1d Principali classi di inibitori della PI3K***

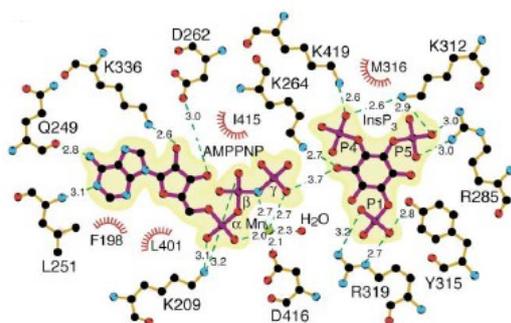
Tra gli inibitori della PI3K sono stati identificati due composti principali, LY294002 e wortmannin, comunemente usati per inibire la proliferazione e la crescita delle cellule tumorali e sensibilizzare le cellule maligne al trattamento con farmaci chemioterapici e alla radioterapia<sup>6</sup>. L'inibizione si realizza attraverso l'interazione a livello del sito di legame per l'ATP posto nella subunità catalitica dell'enzima; LY294002 è un derivato morfolinico della Quercetina che si è dimostrato essere un potente inibitore reversibile della PI3K. Studi condotti sia in vitro che in vivo, ad esempio, hanno evidenziato la sua capacità di invertire la resistenza al Paclitaxel nel carcinoma ovarico; tuttavia, la scarsa biodisponibilità ha precluso la possibilità di un ulteriore sviluppo di questo inibitore. È stato sviluppato, tuttavia, un profarmaco, SF1126, che permette di aumentare la quantità di LY294002 in grado di raggiungere le cellule bersaglio; studi di fase I hanno mostrato che è ben tollerato e ha una buona attività antitumorale.

Il wortmannin, invece, è un metabolita micotico che inibisce irreversibilmente PI3K; è un composto instabile, meno attivo in vivo che in vitro ed attualmente non trova uso clinico. Al fine di superare i suoi problemi farmacologici, sono stati sviluppati dei derivati, tra i quali PX-866 (derivato più potente del wortmannin attualmente disponibile) risultato essere ben tollerato nel trattamento di tumori solidi negli studi di fase I, ed attualmente oggetto di studi clinici in combinazione con chemioterapici citotossici.

## ***2.1e Descrizione del sito di legame dell'ATP***

La tasca di legame dell'ATP a cui si legano gli inibitori sopra descritti, è situata tra i lobi N e C terminale, l'adenina attraverso legami idrogeno lega i residui di Gln249 e Leu251 (Figura 7). Le tasche idrofobiche dell'adenina e del ribosio sono formate dai residui Leu251 e Leu252 insieme con residui dei lobi N e C (Leu401 e Ile415 parte altamente conservata del motivo GSSLL<sup>401</sup> e <sup>415</sup>IDFG, rispettivamente). Gli ossigeni del ribosio dell'ATP sono legati con un legame idrogeno alle catene di Lys336 e Asp262.

L'Asp262 nel motivo DxK è altamente conservato dai membri della famiglia delle PI3K. I residui coordinanti i fosfati dell'ATP sono conservati all'interno della famiglia delle PI3K. La Lys209 del  $\beta$ 2N coordina il gruppo fosfato in  $\alpha$  e  $\beta$ . È altamente conservato l'Asp416 nelle inositol-polifosfato chinasi (IPKs) coordinati l' $\alpha$  e il  $\gamma$  fosfato con il cofattore  $Mn^{2+}$ . Questo aspartato sembra posizionare il  $\gamma$  fosfato per l'attacco nucleofilo<sup>13</sup>.



**Figura 7:** Sito di legame dell'ATP nella PI3K.

## ***2.2 Scopo e rationale del lavoro***

L'obiettivo del progetto di ricerca sulla PI3K, svolto in collaborazione con l'Industria Farmaceutica Sigma-Tau (Sigma-Tau, S.p.A., Pomezia, Italia), è stato quello di individuare l'approccio farmacologico più appropriato per influenzare il segnale promosso dalla cascata di fosforilazioni iniziata dall'enzima PI3K, responsabile dei meccanismi di proliferazione cellulare. È stata inizialmente valutata la possibilità di progettare composti "duali" in grado di colpire più componenti della cascata di fosforilazione quali PI3K ed mTOR. Da un'attenta ricerca bibliografica e dall'analisi biostrutturale e bioinformatica di tutte le strutture sperimentalmente determinate attualmente disponibili nella Brookhaven Protein Data Bank (PDB; [www.rcsb.org](http://www.rcsb.org)), è risultato che le sole mutazioni della subunità p110 $\alpha$  sono sufficienti a causare l'insorgenza e lo sviluppo di tumori nell'uomo, nonché sono coinvolte nei fenomeni di chemioresistenza, e che ciò è causato da un'interferenza nel legame della subunità p110 $\alpha$  con i suoi partner proteici. Sulla base di questi studi, quindi, è stato successivamente abbandonato l'obiettivo di progettare inibitori "duali" che vadano a colpire gli enzimi PI3K ed mTOR ed è stata posta particolare attenzione sulla subunità p110 $\alpha$  di PI3K (ovvero sulla classe Ia) con l'intento di inibire la sua interazione con alcuni partner proteici cruciali a valle (Ras) della cascata di segnali cellulari di cui è promotrice.

Di conseguenza l'obiettivo principale è stato quello: i) di inibire il sito catalitico e contemporaneamente inibire il legame con Ras considerando in particolare la forma mutata H1047R di PI3K, ii) stabilizzare l'interazione della subunità catalitica p110 $\alpha$  con il dominio N-SH2 della subunità regolatoria p85 $\alpha$ , con conseguente inibizione del legame fra la PI3K ed il recettore tirosin-chinasico, tenendo in considerazione il meccanismo d'azione dei mutanti E542K ed E545K.

La progettazione di nuovi composti in grado di inibire il sito catalitico della PI3K è stata effettuata selezionando come scheletro molecolare di partenza (*lead*) il composto ST0358AA1 proveniente da un database di inibitori prodotti dalla Sigma-Tau e che studi farmacologici hanno mostrato inibire scarsamente la PI3K (a concentrazioni >10  $\mu$ M).

I composti progettati, sono stati sintetizzati dal gruppo del Prof. Maurizio Taddei dell'Università di Siena e testati dall'Industria Farmaceutica Sigma-Tau, gli studi NMR sono stati condotti dal gruppo di ricerca del Prof. Orazio Tagliatela Scafati dell'Università degli Studi di Napoli "Federico II". I test farmacologici sono stati effettuati dal dott. Pisano presso i laboratori della Sigma-Tau di Pomezia.

La progettazione, invece, di composti in grado di stabilizzare l'interazione della subunità catalitica p110 $\alpha$  con il dominio N-SH2 della subunità regolatoria p85 $\alpha$ , è stata effettuata a

partire da studi approfonditi sui domini/motivi coinvolti nelle interazioni delle due subunità prese in considerazione.

## ***2.3 Parte Sperimentale***

Tutti gli studi di modellistica molecolare sono stati eseguiti su Silicon Graphics SGI Origin 200 8XR12000 ed E4 Server Twin 2 x Dual Xeon—5520, equipaggiati con due nodi; ogni nodo: 2 x Intel® Xeon® QuadCore E5520—2,26Ghz, 36 GB RAM.

### ***2.3a Valutazione delle possibili Forme Tautomeriche e Calcolo del pKa degli inibitori***

I valori di pKa di tutti i composti presi in esame sono stati calcolati utilizzando il software ACD/pKa versione 12.00 (*Advanced Chemistry Development Inc., Toronto, Canada*) considerando tutte le possibili forme tautomeriche. Di conseguenza, le percentuali delle forme neutre/protonate sono state determinate considerando il valore di pH citoplasmatico (7.2) usando l'equazione di Henderson-Hasselbach. In base ai risultati ottenuti, ciascun tautomero di ogni composto, nel suo stato di protonazione a pH fisiologico, è stato poi sottoposto ad un'approfondita analisi conformazionale.

### ***2.3b Analisi Conformazionale degli inibitori cristallizzati e del lead***

Le molecole sono state costruite utilizzando il modulo *Builder di Insight 2005* tenendo conto della forma ionica prevalente di ogni tautomero al valore di pH considerato. I potenziali atomici e le cariche parziali sono state assegnate utilizzando il campo di forza *CFF91*<sup>14</sup>.

Lo spazio conformazionale di questi composti è stato campionato attraverso 200 cicli di *Simulated Annealing* ( $\epsilon=1$ ) seguendo il protocollo di seguito riportato. Una temperatura iniziale di 1000 K è stata applicata al sistema per 2000 fs (time step= 3.0 fs) al fine di superare tutte le barriere torsionali; successivamente la temperatura è stata ridotta a 300 K in 1000 fs (time step= 1.0 fs) con un decremento di 0.5 K/fs. Le strutture risultanti sono state sottoposte a minimizzazione energetica utilizzando il modulo *Discover\_3 di Insight 2005* (campo di forza CFF91;  $\epsilon= 1$ ). La minimizzazione energetica è stata eseguita usando come algoritmo di minimizzazione *Conjugate Gradient*<sup>15</sup>; il criterio di convergenza è stato stabilito in modo che il massimo valore di RMS sulla derivata prima sia inferiore a 0.001 kcal/Å.

I conformeri ottenuti sono stati classificati secondo i valori di energia conformazionale, gli angoli torsionali, la presenza di legami idrogeno intra-molecolari e le distanze interatomiche facendo uso di una macro in linguaggio bcl ideata dal gruppo di ricerca presso il quale ho svolto il mio lavoro di ricerca.

Al fine di analizzare in maniera dettagliata le proprietà elettroniche, tutti i conformeri, ottenuti dai calcoli di meccanica e dinamica molecolare, sono stati sottoposti ad una piena ottimizzazione geometrica usando il metodo quanto-meccanico PM6<sup>16</sup> e di Mopac 2009<sup>17</sup>. Al fine di raggiungere un gradiente geometrico (GNorm) fissato al valore di 0.01 è stato utilizzato l'algoritmo di minimizzazione EF (Eigenvector Following routine)<sup>18</sup>. Per ottenere una completa ottimizzazione della geometria di partenza abbiamo aumentato i criteri per la terminazione di tutte le ottimizzazioni di un fattore di 100, usando la keyword *PRECISE*.

### ***2.3c Analisi Strutturale e Bioinformatica***

Le strutture cristallografiche delle PI3K sono state scaricate dalla *Protein Data Bank* (PDB, <http://www.rcsb.org/pdb/>). Gli idrogeni sono stati aggiunti ad esse considerando un valore di pH di 7.2.

Le strutture prese in considerazione sono state le seguenti:

*PI3K subunità alfa* PDB IDs: 3HHM (mutante), 3HIZ (mutante), 2ENQ, 2RD0, 2REA, 2RED, 2V1Y, 2IWL, 2IUG, 2IUH, 2IUI, 1H90, 1FU5, 1FU6, 1PHT, 1PKS, 1PKT, 2AR5, 2B3R, 1QAD, 1BF1, 1BFJ, 1PIC.

*PI3K subunità gamma* PDB IDs: 3IBE, 3ENE, 3DPD, 3DBS, 2V4L, 3CSF, 3CST, 2CHW, 2CHZ, 2CHX, 2A4Z, 2A5U, 1HE8, 1E7U (porcina), 1E7V(porcina), 1E8W(porcina), 1E8Y, 1E8Z(porcina), 1E90(porcina), 1E8X(porcina), 3LJ3, 3L13, 3L16, 3L17, 3L08, 3L54, 3APC, 3APD, 3APF, 3QJZ, 3QKO

*mTOR* PDB IDs: 2NPU, 2GAQ, 3FAP, 4FAP, 1AUE, 1W1N

Tutte le strutture cristallografiche sono state analizzate utilizzando i moduli *Biopolymer e Homology di Insight2005*.

Inoltre sono stati effettuati studi di previsione di strutture secondarie sulla subunità di PI3K ed mTOR utilizzando il server *Structure Prediction e Sequence Analysis (PredictProtein)* (<http://www.predictprotein.org/>). La correttezza delle previsioni è stata poi verificata confrontando i dati della predizione con le strutture cristallografiche.

Sono stati effettuati allineamenti di sequenze utilizzando il server BLAST (<http://www.ebi.ac.uk/tools/blast/>) e il server PROMALS 3D (<http://prodata.swmed.edu/promals3d/promals3d.php>).

I motivi funzionali e i domini strutturali coinvolti nelle interazioni proteina-proteina sono stati identificati utilizzando il server *Eukaryotic Linear Motif* (<http://elm.eu.org/>) che consente di

identificare piccoli siti funzionali su proteine eucariotiche, così come quelli soggetti a modifiche post-traduzionali o coinvolti in interazioni proteina-proteina.

Dalla *Protein Data Bank* (PDB, <http://www.rcsb.org/pdb/>) sono state scaricate le strutture determinate sperimentalmente del complesso dei seguenti domini di interazione proteina-proteina complessati con i loro motivi partner: SH2 STAT5 (PDB ID: 1Y1U), SH2 SRC (PDB ID: 2EU0), SH2 GRB2 (PDB ID: 1QG1), SH3 (PDB ID: 1SEM), 14-3-3 (PDB ID: 2B05), FHA2 (PDB ID: 1G6G), PDZ\_3 (PDB ID: 2KOH), TRAF2\_1 (PDB ID: 1QSC), BRCT\_BRCA1 (PDB ID: 1T29), WW (PDB ID: 1F8A), IQ (PDB ID: 1M45).

Al fine di identificare i domini strutturali di interazione proteina-proteina all'interno delle strutture tridimensionali di PI3K (p110 $\alpha$ /p85 $\alpha$ ) è stata effettuata una sovrapposizione strutturale dei domini di interazione proteina - proteina noti con analoghe strutture secondarie identificate all'interno delle proteine in esame.

### ***2.3d Procedura di Docking***

La struttura utilizzata per i calcoli di *docking* è stata quella del mutante H1047R di PI3K (PDB ID: 3HHM). I potenziali atomici sono stati assegnati usando il campo di forza CFF91<sup>14</sup>. Sebbene nel successivo protocollo di docking flessibile l'intero sistema venga comunque perturbato dal metodo *Monte Carlo* e dal *Simulated Annealing*, tuttavia la procedura di docking dinamico richiede una struttura iniziale ragionevole. Il ligando è stato collocato nel sito catalitico prendendo come riferimento la struttura cristallografica del complesso PI3K/wortmannin sempre sulla base dei risultati precedentemente ottenuti.

Il docking flessibile è stato eseguito utilizzando il modulo *Affinity* di Insight 2005 e scegliendo la procedura di *SA\_Docking* (*Simulated Annealing Docking*).

Sono stati effettuati numerosi calcoli di docking al fine di: i) valutare le variazioni conformazionali della subunità catalitica p110 $\alpha$  in presenza o meno della subunità regolatoria p85, ii) testare l'utilizzo di vari criteri di selezione (*Energy Test*; *Metropolis Test*) delle possibili pose molecolari, iii) variare il metodo di calcolo (*Cell Multipole*; *Group Based*; *Quartic\_No\_Coulomb*) delle interazioni di non legame.

La porzione della proteina lasciata completamente libera di muoversi durante i calcoli eseguiti (*binding site*), è stata ottenuta dall'unione di tutti gli amminoacidi aventi almeno un atomo entro un raggio di 10Å da un qualsiasi atomo del ligando con quelli aventi almeno un atomo entro un raggio di 10 Å dal dominio di legame di Ras (*RBD domain*).

Le molecole d'acqua presenti nel sito catalitico sono state lasciate libere di muoversi, applicando una forza di 10 kcal/mol/Å.

Dal complesso iniziale sono state generate casualmente 20 nuove strutture utilizzando la procedura combinata Monte Carlo/minimizzazione con un criterio di tolleranza energetico di  $10^6$  Kcal/mol per assicurare un'ampia variabilità delle strutture da minimizzare (2500 iterazioni;  $\epsilon=1$ ). Durante questa fase i ligandi sono stati mossi attraverso una combinazione casuale di movimenti traslazionali e rotazionali, e, variando il valore degli angoli torsionali (opzione *Flexible\_Ligand*, considerando tutti i legami ruotabili), al fine di campionare sia lo spazio conformazionale del ligando che la sua orientazione rispetto al sito catalitico ( $MxRChange = 3 \text{ \AA}$ ;  $MxAngChange = 180^\circ$ ). Al fine di evitare forti divergenze nelle energie di van der Waals (vdW) e Coulomb, sono stati scalati i termini relativi alle forze di vdW e di Coulomb di un fattore pari a 0.1.

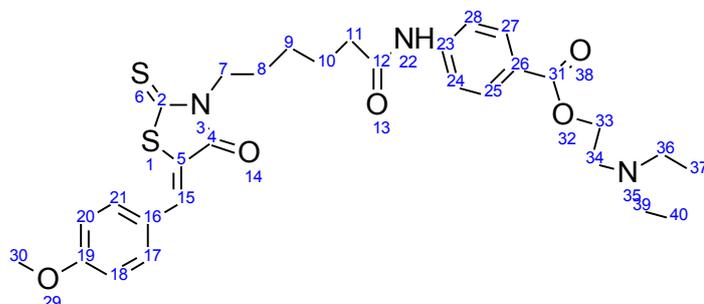
Dopo la fase di minimizzazione (*Conjugate Gradient*; 2500 iterazioni;  $\epsilon=1$ ), sono stati applicati un criterio di controllo di similitudine delle strutture (RMS tolerance=0.3 Kcal/Å) e due criteri di energia (Energy test; T=50 K, Metropolis test; T=310 K) per selezionare 20 soluzioni accettabili.

I complessi selezionati sono stati prima minimizzati e poi sottoposti a 50 cicli di *Simulated Annealing*, ciascuno di 100 fs. Durante la durata di un ciclo di *Simulated Annealing* la temperatura del sistema è stata diminuita linearmente da 500K a 300K, contemporaneamente le forze di vdW e di Coulomb sono state riportate dal 10% del loro valore al 100%.

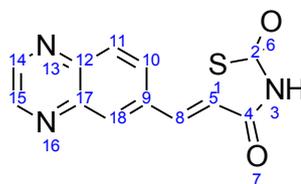
Alla fine dei calcoli di dinamica molecolare ciascun complesso ottenuto è stato sottoposto ad una minimizzazione finale di  $10^4$  cicli.

### 2.3e Pharmacophore Fitting del lead sull'inibitore AS-605240

Sono state avanzate due diverse ipotesi di sovrapposizione realizzate utilizzando il comando *Transform Superimpose*; in particolare è stato sovrapposto l'anello 1,3-tiazolidin-4-one del *lead* (Figura 8) sull'anello 1,3-tiazolidin-2,4-dione dell'inibitore AS-605240 (Figura 9).



**Figura 8:** struttura bidimensionale del lead con numerazione IUPAC



**Figura 9:** struttura bidimensionale dell'inibitore AS-605240 con numerazione IUPAC

Le sovrapposizioni sono state effettuate nel modo seguente:

- I tipo di sovrapposizione: S1 (ST0358AA1) su N3 (AS-605240), C2 su C2, N3 su S1, C4 su C5; e C5 sul C4;
- II tipo di sovrapposizione: C2 (ST0358AA1) su N3 (AS-605240), S1 su C2, C5 su S1, C4 su C5, N3 sul C4.

Questi due tipi di sovrapposizione hanno dato origine a due differenti modalità di legame possibili nel sito catalitico, caratterizzate da un diverso accomodamento delle porzioni R ed R'.

### ***2.3f Studi di Spettroscopia di Risonanza Magnetica Nucleare***

Gli studi di spettroscopia di Risonanza Magnetica Nucleare (NMR), sono stati effettuati dal prof. Orazio Tagliatela Scafati del Dipartimento di Farmacia dell'Università Federico II di Napoli, per determinare la configurazione del doppio legame dei carboni in  $\alpha$  e  $\beta$  al carbonile dell'anello 1,3-tiazolidin-4-one. Tali studi sono stati condotti in DMSO ed hanno confermato la configurazione *cis* del doppio legame considerato. Sulla base di questi risultati, sono stati eseguiti calcoli di dinamica e meccanica molecolare e calcoli quanto-meccanici (PM7) fissando il doppio legame in configurazione *cis*.

### ***2.3g Analisi Conformazionale dei nuovi composti progettati***

Per l'analisi conformazionale degli inibitori progettati è stata eseguita la stessa procedura descritta nel paragrafo 2.3b. utilizzando il metodo quanto-meccanico PM7<sup>16</sup>. Sulla base degli studi NMR su è stata applicata una restrizione al del doppio legame dei carboni in  $\alpha$  e  $\beta$  al carbonile dell'anello 1,3-tiazolidin-4-one. I calcoli sono stati condotti utilizzando sia la costante dielettrica dell'acqua ( $\epsilon=80.0*r$ ) che quella del vuoto ( $\epsilon=1.0*r$ ). I conformeri ottenuti sono stati quindi classificati secondo:

- i) i valori di energia conformazionale ( $\Delta E < 5$  Kcal);
- ii) le distanze interatomiche fra i gruppi funzionali;
- iii) il valore degli angoli torsionali;
- iv) la presenza di legami idrogeno intramolecolari.

### ***2.3h Pharmacophore Fitting degli inibitori progettati su AS-605240 ed AS-604850***

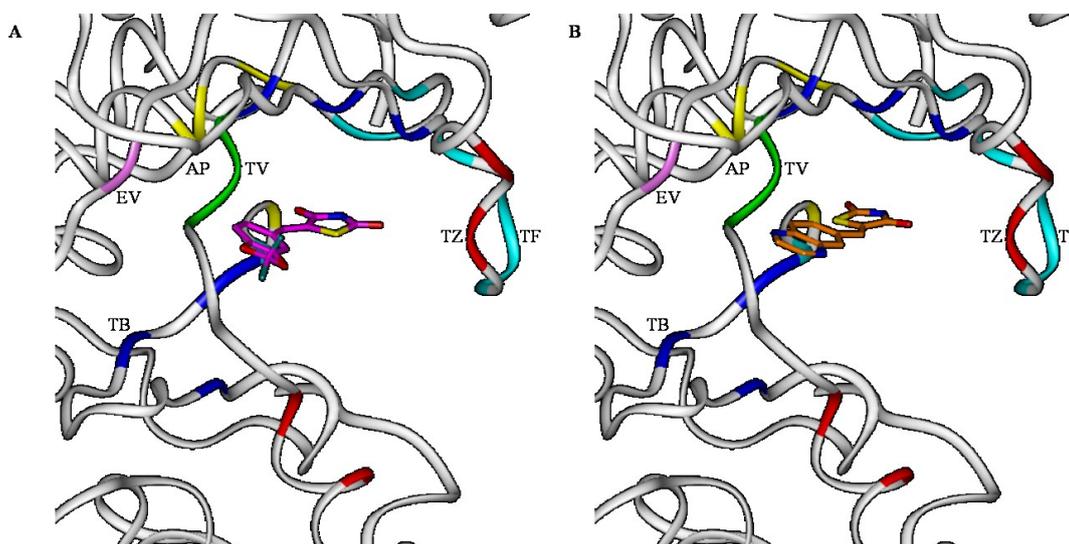
Partendo dai risultati dell'analisi delle proprietà conformazionali ed elettroniche degli inibitori progettati; sono stati presi in considerazione, per ciascun ligando, i seguenti conformeri:

- a) il minimo globale ottenuto dal calcolo condotto in acqua ( $\epsilon = 80.0*r$ );
- b) il minimo globale ottenuto dal calcolo condotto nel vuoto ( $\epsilon = 1.0*r$ );
- c) il conformero a più bassa energia che meglio riesce a posizionarsi nel sito catalitico.

Tali conformeri, al fine di valutarne l'accomodamento nel sito catalitico delle PI3K $\alpha/\gamma$ , sono stati sovrapposti a due inibitori noti AS-605240 (PDB ID: 2A5U) e AS-604850 (PDB ID:

2A4Z) a livello delle porzioni strutturali simili: anello 1,3-tiazolidin-4-one (per i composti in esame) ; anello 1,3-tiazolidin-2,4-dione (per gli inibitori AS-605240 e AS-604850).

Le sovrapposizioni sono state realizzate utilizzando il comando *Transform Superimpose*. I nuovi composti sono stati sovrapposti una volta all'inibitore AS-604850 (*Figura 10A*) una volta all'inibitore AS-605240 (*Figura 10B*) nel modo riportato nella tabella 2 seguente:



**Figura 10:** Le figure 10A ed 10B mostrano l'accomodamento, rispettivamente, dell'inibitore AS-604850 (magenta) e dell'inibitore AS-605240 (arancione) nel sito catalitico della PI3K. Sono indicate: la tasca della valina (TV), in verde; la tasca dove si posiziona lo zucchero dell'ATP (TZ,) in rosso; la tasca dove si posiziona la base dell'ATP (TB) in blu; la tasca dove si posiziona il fosfato dell'ATP (TF) in ciano; una prima regione dove si puo' accomodare un eventuale volume molecolare aggiuntivo, rispetto a quello dell'ATP, (EV), in rosa; una seconda regione dove si puo' accomodare un eventuale volume molecolare aggiuntivo, rispetto a quello dell'ATP, detta tasca di affinità (AP) in giallo. Gli eteroatomi dei ligandi sono colorati sulla base dei tipi atomici (N: blu,O: rosso e S: giallo).

**Tabella 2:** Atomi considerati nella sovrapposizione.

<b>AS-605240/AS(604850)</b>	<b>N3</b>	<b>C2</b>	<b>S1</b>	<b>C5</b>	<b>C4</b>
<b>nuovo inibitore progettato</b>	<b>N3</b>	<b>C2</b>	<b>S1</b>	<b>C5</b>	<b>C4</b>

Effettuando tale sovrapposizione sono state ottenute due differenti modalità di legame possibili nel sito catalitico da noi denominate modalità di legame I e modalità di legame II:

- modalità di legame I: per la sovrapposizione è stato considerato l'inibitore AS-604850 (*Figura 10A*)
- modalità di legame II: per la sovrapposizione è stato considerato l'inibitore AS-605240 (*Figura 10B*)

Per ciascun conformero sono state valutate le :

- a) compenetrazioni steriche con gli aminoacidi presenti nel sito attraverso il calcolo e l'analisi delle superfici di vdW, utilizzando il comando *Molecule Surface*;
- b) repulsioni/attrazioni elettroniche con gli aminoacidi presenti nel sito, attraverso la visualizzazione ed analisi delle cariche parziali ottenute dai calcoli quanto-meccanici.

### 2.3i Costruzione del modello della PI3K $\alpha$

Dal momento che nelle strutture cristallografiche della PI3K (PDB IDs: 3HHM, 3HIZ) mancavano le regioni di quattro regioni di legame flessibili per la p110 $\alpha$  e tre regioni di legame flessibili della p85, questi sono stati modellati utilizzando la procedura standard di costruzione di modelli prevista dal modulo *Homology* di *Insight 2005* (Accelrys, San Diego). Tutti i calcoli successivi sono stati eseguiti utilizzando il campo di forza CFF91<sup>14</sup>. per parametrizzare i residui aminoacidici della proteina.

Sono state assegnate le coordinate della SCRs con il comando *Sequence/AssignCoords* nel seguente modo (Tabella 3) :

**Tabella 3:** assegnazione coordinate delle regioni di legame mancanti nella PI3K $\alpha$

<b><i>SUBUNITA'</i></b>	<b><i>REGIONE DI LEGAME</i></b>	<b><i>SEQUENZE</i></b>	<b><i>PDB ID</i></b>
p110 $\alpha$	Regione di legame 1	P5-R309	3HHM
		E321-V409	3HHM
	Regione di legame 2	K410-G414	3HIZ
		L422-N515	3HHM
Regione di legame 3	L422-N515	3HHM	
	R524-H940	3HHM	
Regione di legame 4	R524-H940	3HHM	
	P553-I1062	3HHM	
p85 $\alpha$	Regione di legame 5	W1-S78	3HIZ
		N79-Q105	3HHM
	Regione di legame 6	N79-Q105	3HHM
		V109-D113	3HHM
	Regione di legame 7	V109-D113	3HHM
		A117-K120	3HHM
	Regione di legame 8	A117-K120	3HHM
		E124-Q252	3HHM

Le porzioni N-terminale MPPRP (1-5) e C-terminale KQHALN (1063-1068) sono state aggiunte con il comando *END-REPAIR*; è stata unmergiata (eliminata) l'acqua di 3HHM ed è stata mergiata (unita) nel modello e sono stati assegnati i potenziali con il campo di forza CFF91 carica totale -2 (gli aminoacidi carichi positivamente sono 179 mentre gli aminoacidi carichi negativamente 181). Le porzioni aggiunte al modello sono state ottimizzate fissando tutto il resto della proteina.

I parametri utilizzati sono stati i seguenti: *Step*: 10.000; *Steepest descent*: fino a convergenza 10; *Conjugate gradient*: fino a 1.

A partire dal risultato precedente abbiamo eseguito un'ottimizzazione fissando tutta la proteina ad eccezione delle regioni di legame aggiunti N-terminali e C-terminali.

### ***Costruzione modello (Tabella4)***

**Tabella 4:** Regioni di legame PI3K $\alpha$  modellate

<b><i>SUBUNITA'</i></b>	<b><i>LOOP</i></b>	<b><i>AMINOACIDI</i></b>	<b><i>SEED</i></b>	<b><i>INTERNAL-EXTERNAL OVERLAP</i></b>	<b><i>LOOP SCELTO</i></b>
p110 $\alpha$	<i>Regione di legame 1</i>	309-321	102861	0.5	<i>Regione di legame 5</i>
	<i>Regione di legame 2</i>	414-422	539973	0.5	<i>Regione di legame 4</i>
	<i>Regione di legame 3</i>	515-524	609972	0.5	<i>Regione di legame 1</i>
	<i>Regione di legame 4</i>	940-953	28361	0.5	<i>Regione di legame 7</i>
p85 $\alpha$	<i>Regione di legame 5</i>	105-109	1564353	0.5	<i>Regione di legame 4</i>
	<i>Regione di legame 6</i>	113-117	1245654	0.5	<i>Regione di legame 10</i>
	<i>Regione di legame 7</i>	120-124	1683749	0.5	<i>Regione di legame 10</i>

Queste regioni di legame sono state generate utilizzando il *Modulo Homology* e il comando *loops-generate*.

## ***2.3l Procedura di Docking degli inibitori progettati***

Al fine di confermare gli studi di *pharmacophore fitting* nella PI3K degli inibitori presi in considerazione; sono stati eseguiti studi di *docking* dinamico sui complessi p110 $\alpha$ /p85 $\alpha$  usando la procedura descritta nel paragrafo 2.3e in questo caso però per il calcolo delle interazioni di non legame è stato scelto solo il metodo: *Cell\_Multipole*<sup>19</sup> e solo il criterio energetico (*Metropolis* test; T=310 K) per selezionare le 20 soluzioni accettabili.

Come conformazioni iniziali per i ligandi sono stati selezionati, in accordo con gli studi precedentemente descritti, i minimi energetici derivanti dalla ricerca conformazionale dei ligandi ST0388AA1, ST0394AA1 scelti sulla base della loro elevata capacità inibitoria sulla PI3K $\alpha$ ; mentre per quanto riguarda la proteina è stata considerata quella del modello costruito che presenta i) la mutazione H1047R; ii) non presenta regioni di legame mancanti nella subunità catalitica p110 $\alpha$ , iii) presenta gran parte della subunità regolatoria p85.

I ligandi sono stati collocati nel sito catalitico prendendo come riferimento le due modalità di legame precedentemente descritte (vedi paragrafo 2.3i)

Le molecole di acqua presenti nelle strutture cristallografiche sono state mantenute, ad eccezione di quelle stericamente sovrapposte al ligando.

Per valutare nel dettaglio il ruolo giocato dalle molecole d'acqua presenti nel sito catalitico sono state eseguite due tipologie di *docking*: una dove le molecole d'acqua sono state lasciate libere di muoversi, l'altra dove sono state vincolate applicando una forza di 10 kcal/mol/Å; inoltre è stata applicata una restrizione del doppio legame dei carboni in  $\alpha$  e  $\beta$  al carbonile dell'anello 1,3-tiazolidin-4-one che studi NMR hanno mostrato essere sempre in *cis*.

## ***2.3m Progettazione dei Peptidi***

I peptidi in esame sono stati costruiti utilizzando il modulo *Biopolymer di Insight 2005*. I potenziali atomici e le cariche parziali sono state assegnate utilizzando il campo di forza CFF91<sup>14</sup>.

Il protocollo di progettazione dei peptidi ha previsto:

a) *Analisi strutturale e bioinformatica*: analisi dei motivi di legame dei domini SH2 per individuare la sequenza canonica riconosciuta ed analisi strutturale della porzione di p110 $\alpha$  che contatta l'N-SH2 (PDB ID: 3HHM) (modulo *Homology; Insight 2005*).

c) *Studi di dinamica molecolare*: lo spazio conformazionale dei peptidi è stato campionato attraverso 500 cicli di *Simulated Annealing* ( $\epsilon=80^*r$ ) seguendo il protocollo di seguito

riportato. Una temperatura iniziale di 1000 K è stata applicata al sistema per 2000 fs (time step= 1.0 fs) al fine di superare tutte le barriere torsionali; successivamente la temperatura è stata ridotta a 300 K in 1000 fs (time step= 1.0 fs) con un decremento di 0.5 K/fs.

d) *Studi di meccanica molecolare*: le strutture risultanti sono state sottoposte a minimizzazione energetica utilizzando il modulo *Discover\_3* di *Insight 2005* (campo di forza CFF91;  $\epsilon = 80 \cdot r$ ). La minimizzazione energetica è stata eseguita usando come algoritmo di minimizzazione *Conjugate Gradient*<sup>15</sup>; il criterio di convergenza è stato stabilito in modo che il massimo valore di RMS sulla derivata prima sia inferiore a 0.001 kcal/Å.

## ***2.4 Risultati e Discussione***

### ***2.4a Analisi Strutturale degli enzimi PI3K di classe Ia***

Le PI3K di classe Ia sono formate da due subunità: la p110 e la p85. La p110 $\alpha$  è costituita da un dominio N-terminale che lega la p85 (p85BD), un dominio che lega Ras (RBD), un dominio protein chinasi di tipo C2, un dominio ad elica e un dominio chinasi C-terminale.

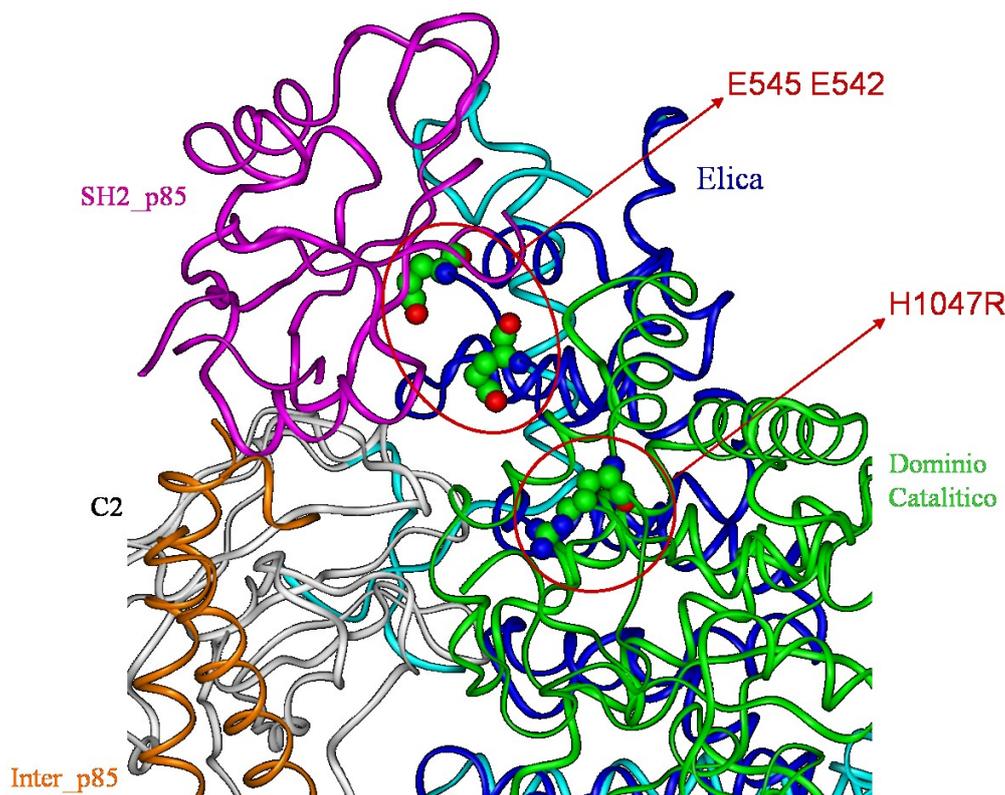
La p110 $\alpha$  è costitutivamente associata con una subunità regolatoria di cui p85 $\alpha$  è quella più studiata.

Le due subunità esistono nelle cellule come complessi eterodimerici. Nelle cellule quiescenti, p85 $\alpha$  stabilizza p110 $\alpha$  e inattiva l'attività di PI3K. Su stimolazione di fattori di crescita, i recettori tirosin chinasi (RTKs) subiscono un'autofosforilazione, creando un sito di legame per proteine contenenti un dominio proteico di tipo Src2 (SH2). I domini SH2 di p85 si legano al motivo YxxM fosforilato nei RTKs. Questo legame attiva la p110 $\alpha$  e permette il suo reclutamento a livello della membrana plasmatica. La diretta interazione tra Ras-GTP e p110 $\alpha$  attraverso l'RBD aumenta l'attività catalitica di p110 $\alpha$ .

Le più frequenti mutazioni di p110 $\alpha$  sono E452K, E454K e H1047R. Tali mutazioni incrementano l'attività enzimatica, stimolano il segnale di Akt (protein chinasi B) e stimolano fattori di crescita causando tumori in vivo.

L'attività indotta dalle mutazioni E542K e E545K, è altamente dipendente dal legame Ras-GTP, mentre il guadagno di funzione enzimatica indotta da H1047R è forse dipendente dal cambiamento allosterico mediato da p85 $\alpha$ , e questo cambiamento mima il legame Ras-GTP, rendendo l'attività catalitica di questo mutante indipendente dall'interazione con Ras-GTP.

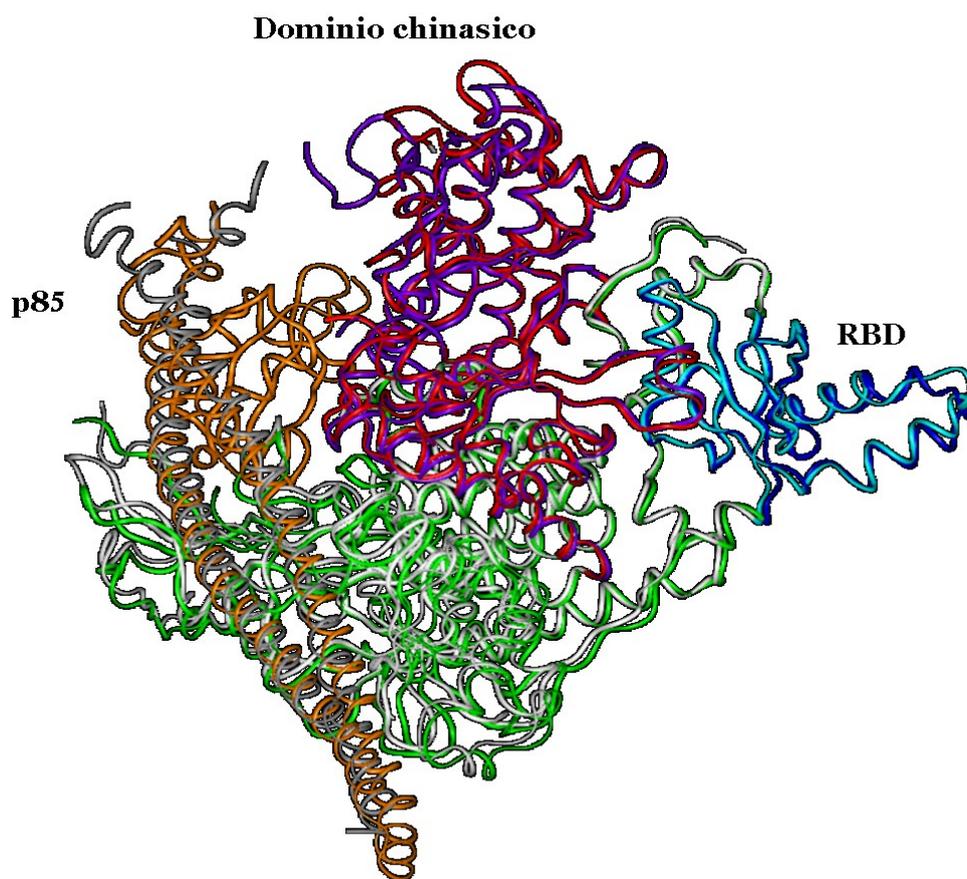
Studi biochimici e strutturali supportano la teoria che il dominio d'elica mutato E542K, E545K (*Figura 11*), rivela un'interazione inibitoria tra il dominio SH2 N-terminale di p85 $\alpha$  e il dominio d'elica p110 $\alpha$ .



**Figura 11:** In figura sono riportate le principali mutazioni che riguardano la subunità p110 $\alpha$  e sono indicati i siti in cui esse sono presenti: in verde il dominio catalitico in cui è presente la mutazione H1047R, in blu il dominio ad elica in cui sono presenti le mutazioni E542K, E545K, in magenta l'N-SH2 della p85, il bianco il dominio C2 e in arancio la regione centrale di connessione tra i domini (inter-region) della p85.

Nel mutante H1047R, la regolazione della cascata enzimatica di Akt appare essere sufficiente per l'induzione della trasformazione oncogenica da parte di PI3K. Alcuni risultati suggeriscono che un processo sconosciuto dipendente da p85 $\alpha$  gioca un ruolo critico nell'oncogenicità di H1047R<sup>11</sup>. Infatti, per esaminare gli effetti del legame di p85 $\alpha$  nell'attività del mutante, sono stati eliminati i legami di p110 $\alpha$  a p85 $\alpha$  mediante delezione di 72 aminoacidi N-terminali da p110 $\alpha$ . L'eliminazione del sito di legame a p85 $\alpha$  ha un minore effetto sulla trasformazione dell'attività del dominio ad elica mutato, ma è completamente abolita la trasformazione indotta dalla mutazione del dominio chinasi H1047R.

I risultati di tutti gli studi effettuati hanno comunque indicato che il mutante H1047R di p110 $\alpha$  costituisce una delle principali cause di insorgenza di tumori ed è inoltre sovra-espresso nelle cellule tumorali stesse. D'altra parte i risultati della nostra analisi strutturale hanno evidenziato che in tale mutante l'interazione con la subunità regolatoria p85 $\alpha$  risulta essere stabilizzata da un cambiamento conformazionale della subunità catalitica p110 $\alpha$  che coinvolge anche la regione comprendente sia il sito di legame con Ras-GTP che il sito ATPasico (Figura 12).



**Figura 12:** in figura è riportata la sovrapposizione strutturale della PI3K  $\alpha$  nella forma mutata H1047R (PDB ID:3HHM) e non mutata (PDB ID:2RD0). In particolare sono evidenziati il dominio di legame di Ras (RBD) in blu della forma mutata e ciano della forma non mutata; il dominio chinasi in viola della forma mutata e in rosso della forma non mutata; la p85 in arancione della forma mutata e in grigio della forma non mutata e le restanti parti della proteina in bianco della forma mutata e in verde della forma non mutata.

In conclusione, dagli studi effettuati mediante un'attenta ricerca bibliografica, e mediante l'analisi delle strutture sperimentalmente determinate attualmente disponibili nella *Brookhaven Protein Data Bank* ([www.rcsb.org](http://www.rcsb.org)), è emerso che; le mutazioni a carico della subunità p110 $\alpha$  risultano essere la principale causa dello sviluppo di tumori nell'uomo e che il loro effetto è legato ad una variazione conformazionale dell'enzima che ne influenza l'interazione con cruciali partner proteici quali la subunità regolatoria p85. Si è deciso quindi di adottare un approccio basato sulla sola inibizione della PI3K e di effettuare la progettazione dei nuovi inibitori considerando la struttura cristallografica del complesso p110 $\alpha$  H1047R e p85 $\alpha$  (PDB ID: 3HHM) bersagliando in maniera specifica la conformazione dell'enzima stabilizzata da tale mutazione. Pertanto i miei studi si sono focalizzati sulla progettazione di i) inibitori selettivi della subunità p110 $\alpha$  di PI3K in grado di legare il sito ATPasico ed influenzare il legame con partner proteici quali Ras, ii) peptidi in grado di stabilizzare l'interazione della p110 $\alpha$  con la subunità p85 in modo da impedire l'attivazione catalitica

dell'enzima. A tal fine ho proceduto con i differenti tipi di analisi riportati nei successivi paragrafi.

### ***2.4a.1 Analisi dei Domini e dei Motivi Funzionali***

È stata verificata nelle strutture cristallografiche della PI3K (PDB ID: 3HHM) la presenza dei seguenti domini proteici: SH2\_STAT5 (PDB ID: 1Y1U), SH2\_SRC (PDB ID: 2EU0), SH2\_GRB2 (PDB ID: 1QG1), SH3\_3 (PDB ID: 1SEM), 14-3-3\_3 (PDB ID: 2B05), FHA2 (PDB ID: 1G6G), PDZ (PDB ID: 2KOH), TRAF2\_1 (PDB ID: 1QSC), BRCT\_BRCA1 (PDB ID: 1T29), WW (PDB ID: 1F8A) e IQ (PDB ID: 1M45). Per determinare l'eventuale presenza di domini all'interno della PI3K, è stata ricavata la successione delle strutture secondarie sia della proteina in esame che di tutti i domini noti, utilizzando il server *Structure Prediction and Sequence Analysis (Predict Protein)* (<http://www.predictprotein.org/>) ed effettuata un'attenta analisi bioinformatica dei risultati ottenuti. Nel caso in cui nelle strutture della PI3K sia stata riscontrata la medesima successione delle strutture secondarie dei domini di interazione proteina-proteina considerati, il dominio di interesse è stato sovrapposto alla corrispondente sub-struttura di PI3K e sono stati considerati gli amminoacidi e le molecole d'acqua aventi almeno un atomo entro un raggio di 4Å da un qualsiasi atomo del motivo ad esso complementare.

Inoltre, al fine di identificare, all'interno del dominio di volta in volta in esame, gli specifici amminoacidi direttamente coinvolti nell'interazione proteina-proteina, è stato utilizzato il server *Conseq* (<http://conseq.tau.ac.il/>) e sono stati classificati i residui aminoacidici fondamentali sia da un punto di vista funzionale che strutturale.

I risultati ottenuti hanno evidenziato la presenza nella struttura di PI3K dei seguenti domini di interazione proteina-proteina: SH2Grb2, SH3, FHA1, PDZ e 14\_3\_3. (*Figura 13*: è riportato come esempio il risultato ottenuto per il dominio SH2Grb2):

#### *Dominio SH2Grb2 (PDB: 1QG)*

I domini SH2 di tipo Grb2 sono moduli ripiegati di circa 100 amminoacidi contenenti foglietti  $\beta$  centrali antiparalleli circondati da due  $\alpha$  eliche. Sono in grado di legare motivi peptidici specifici che contengono tirosine fosforilate (pY), la presenza di residui conservati contribuisce sia al mantenimento del core idrofobico che al riconoscimento della pY, mentre più residui variabili contribuiscono allo specifico riconoscimento dei residui all'estremità C-terminale. Nella PI3K esso è fondamentale per l'interazione e la successiva attivazione dell'enzima da parte del recettore tirosin-chinasico che, attivato, presenta una sequenza contenente una tirosina fosforilata.

Sequenza delle strutture secondarie: N-terminale (Gly)- foglietti  $\beta$  -  $\alpha$  elica - foglietti  $\beta$  antiparalleli - foglietti  $\beta$  - foglietti  $\beta$  antiparalleli - foglietti  $\beta$  -  $\alpha$  elica - foglietti  $\beta$  - C terminale (Thr)

Aminoacidi funzionali conservati del sito di interazione proteina-proteina

R12  $\rightarrow$  G,R situati nella elica  $\alpha$ 1

R 31  $\rightarrow$  R situati su foglietto  $\beta$ 1

S 33  $\rightarrow$  A,C,R,S situati sulla regione di legame tra i due foglietti  $\beta$ 1,2

V 55  $\rightarrow$  I,V situati dopo il foglietto  $\beta$ 3 del core

Un'analisi più approfondita del dominio di interazione proteina-proteina SH2Grb2 in complesso con il motivo da esso riconosciuto (PDB ID: 1QG) ha rivelato che:

Tre Ser (33, 35, 41) e due Arg (12, 31) del dominio interagiscono con il fosfato dell'amminoacido fosforilato presente sul motivo complementare.

L'His52 del dominio dà un legame con la Ser precedente alla Tyr fosforilata del motivo

I due Asp del motivo riconosciuto sono importanti nell'orientare la parte negativa con quella positiva delle Arg presenti sul dominio.

Nel motivo la Val16 si inserisce nella tasca idrofobica formata dalla Phe53 e dal Trp66 del dominio.

La Asn17 del motivo dà un legame idrogeno con il carbonile dello scheletro aminoacidico della Lys54 del dominio. L'Asn17 forma un legame idrogeno intra-molecolare con l'NH dello scheletro peptidico della Val18 presente sul motivo stesso.

Paragone dell'SH2 individuato nella struttura della subunità regolatoria p85 (PDB ID: 3HMM)  $\alpha$  con il dominio SH2-Grb2 (PDB ID: 1QG)

R, G 12 $\rightarrow$ (19)  $\rightarrow$ R 31  $\rightarrow$ (2)  $\rightarrow$ S, A, C, R 33  $\rightarrow$  (22)  $\rightarrow$ V, I 55

Il dominio SH2 di Grb2 è presente nella struttura cristallografica PDB ID: 3HMM ed è situato nella subunità regolatoria p85 $\alpha$  conservando i seguenti residui identificati come importanti dal punto di vista funzionale per l'interazione proteina-proteina.

R 523  $\rightarrow$  R 542  $\rightarrow$  R 544  $\rightarrow$  I 566

Tali aminoacidi sono situati su di una struttura ad  $\alpha$ -elica.

### Paragone delle strutture secondarie

Aminoacidi utilizzati per la sovrapposizione di tali strutture:

**Tabella 5:** Sovrapposizione dei foglietti  $\beta$ .

<b>p85</b>	<b>SH2-GRB2</b>
378 – 384	50 – 56
373 – 371	45 – 43
354 - 357	27 - 30

**Tabella 6:** Sovrapposizione delle  $\alpha$  eliche.

<b>p85</b>	<b>SH2-GRB2</b>
406 – 403	75 – 78
345 - 342	14 - 17

**Tabella 7:** Principali amminoacidi conservati/omologhi all'interno delle strutture secondarie.

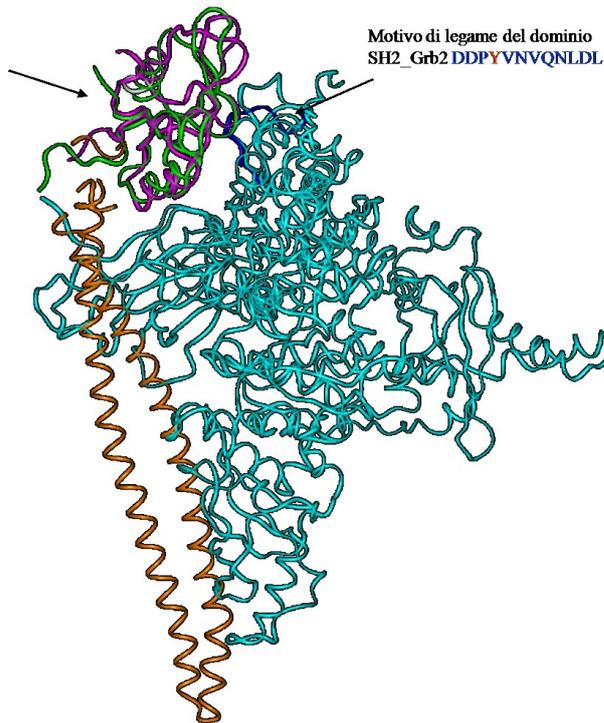
<b>p85</b>	<b>Grb2</b>
Thr369	Ser41
Thr362	Ser35
Ala360	Ser33
Arg340	Arg12
Arg358	Arg31
Leu380	His52

### Motivi che legano il dominio di interazione proteina-proteina SH2Grb2:

Il motivo “complementare” al dominio di interazione proteina-proteina SH2Grb2 (*PDB ID: 1QG*) si sovrappone ad una sequenza della subunità catalitica p110 $\alpha$  (*Figura 13*). È interessante notare come un residuo di Asp della p110 $\alpha$  si sovrappone al fosfato della Tyr fosforilata del motivo riconosciuto dal dominio SH2Grb2, “sostituendosi” ad esso nell'interazione. Inoltre all'interno del dominio SH2 identificato nella subunità p85 $\alpha$  è presente anche il motivo che lega tale dominio, avente sequenza YRNE (*Figura 14*).

Motivo di legame (bianco) del dominio SH2\_Grb2 in p85 alfa con sequenza YRNE che si allinea con HRST del dominio SH2\_Grb2 (PDB:1QG1).

Motivo di legame del dominio SH2\_Grb2 DDPYVNVQNLDL

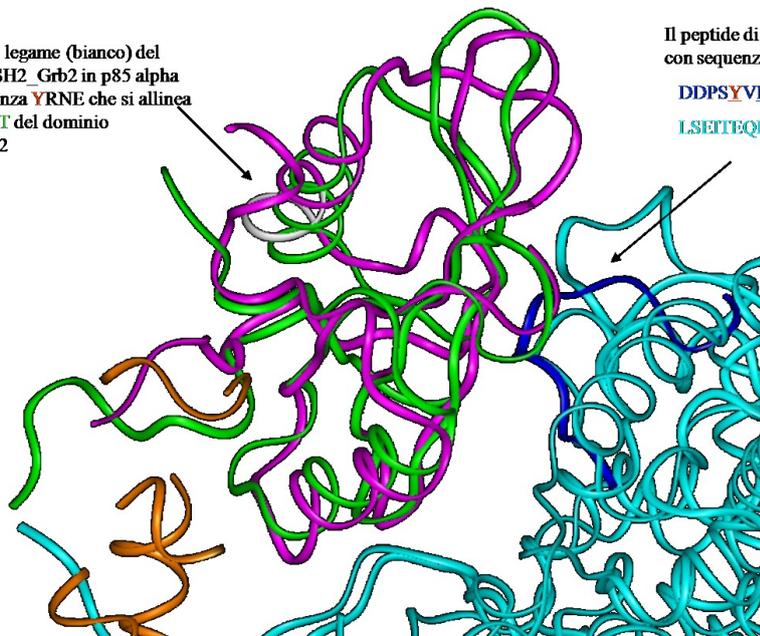


**Figura13:** Sovrapposizione del dominio SH2\_GRB2 (PDB ID: 1QG1; verde), con il dominio SH2\_Grb2 individuato in p85 alfa (PDB ID: 3HHM, magenta): in figura viene mostrato in arancione la p85 alfa, in ciano la subunità p110 alfa, in magenta il dominio SH2\_Grb2 di p85, in verde l'SH2\_Grb2, in blu il peptide di legame del dominio SH2\_Grb2 e in bianco il motivo di legame del dominio SH2\_Grb2 di p85 alfa.

Motivo di legame (bianco) del dominio SH2\_Grb2 in p85 alfa con sequenza YRNE che si allinea con HRST del dominio SH2\_Grb2

Il peptide di legame del dominio SH2\_Grb2 con sequenza

DDPSYVNVQNLDL si allinea con  
LSEITEQELDFLW della p110 alfa



**Figura 14:** Particolare della sovrapposizione del dominio SH2\_GRB2 (PDB ID: 1QG1) con il dominio SH2\_Grb2 individuato in p85 alfa (PDB ID: 3HHM): in figura viene mostrato in arancione la p85 $\alpha$ , in magenta il dominio SH2\_Grb2 di p85, in verde l'SH2\_Grb2, in blu il peptide di legame del dominio SH2\_Grb2 e in bianco il motivo di legame del dominio SH2\_Grb2 di p85 $\alpha$ .

### Dominio SH3 di p85 (PDB ID: 3I5R)

I domini SH3 (Src homology) sono presenti in molte proteine coinvolte nella trasduzione dei segnali intracellulari e legano motivi peptidici ricchi di prolina. Così come i domini SH2, anch'essi sono importanti “adesivi” molecolari, infatti sono responsabili della formazione di complessi macroproteici e quindi essenziali nel mediare specifiche interazioni proteina-proteina.

Confrontando le strutture secondarie presenti nella struttura cristallografica del dominio SH3 di p85 con quella predetta dal server *Predict Protein* risulta che una zona predetta ad  $\alpha$  elica è in realtà presente nella struttura cristallografica come foglietto  $\beta$ .

Struttura secondaria: N-terminale (Ser2) – 2 foglietti  $\beta$  – 2  $\alpha$  eliche – 3 foglietti  $\beta$  – C terminale (Ile82)

La struttura secondaria rispetta la classica struttura del dominio SH3 in cui ci sono 5  $\beta$  *strand* antiparalleli che formano due foglietti  $\beta$  perpendicolari.

Il motivo di legame del dominio ha una sequenza del tipo KRPLPPLPS.

Confrontando la sequenza del peptide con la sequenza consenso di p85 ottenuta dal software *ELM*, essa risulta rispettare la successione n(RKY)..P..P che è quella del motivo che lega il dominio SH3\_1.

Il dominio SH3\_1 è anch'esso presente nella subunità p85 $\alpha$  in quanto la sua sequenza si allinea con quella della p85 $\alpha$ .

Motivi: considerando i motivi che legano i domini SH3\_1 sulla p85 $\alpha$  ed allineando quest'ultima con la struttura del dominio (PDB ID: 3I5R), è possibile notare che sono presenti solo due aminoacidi nella p85 $\alpha$  di quelli individuati dalla sequenza consenso, in quanto il resto della sequenza non è stata cristallizzata. È possibile comunque dedurre che, nella PI3K i motivi di legame del dominio SH3\_1 sono in prossimità del dominio stesso.

Effettuando un allineamento tra la struttura del dominio (PDB ID: 3I5R) e quella della PI3K (PDB ID: 3HHM), è possibile notare che le due strutture non presentano un buon allineamento; inoltre, allineando la sequenza della p85 $\alpha$  con la struttura della PI3K (PDB ID: 3HHM), si evidenzia che i motivi considerati sulla p85 $\alpha$  non sono contenuti nella struttura cristallografica PDB ID: 3HHM, di conseguenza non è possibile definire con precisione dove essi siano situati.

Allineamento sequenze:

p85 $\alpha$ –PDB ID: 3I5R (SH3\_1 di p85) → Ser2 – Ile82

P85 $\alpha$ –PDB ID: 3HHM (p85 + p110 $\alpha$ ) → Met326 – Lys363 ; Gly366 – Gln433; His450 – Gln579

Dall'analisi effettuata è stato possibile dedurre che il dominio SH3 considerato appartiene ai domini SH3 di classe I in quanto riconoscono entrambi la stessa sequenza.

#### Domini 14\_3\_3 PDB ID: 2B05

I domini 14\_3\_3 sono presenti altamente conservati in una famiglia ubiquitaria di proteine di peso molecolare di circa 30 kDa, coinvolte in un'ampia varietà di processi cellulari (apoptosi, risposta allo stress, ciclo cellulare), che attraverso il dominio 14\_3\_3 mediano la trasduzione dei segnali cellulari mediante il legame a proteine contenenti residui fosfoserinici (pS).

Struttura secondaria: N-terminale (Met3) – 9  $\alpha$  eliche – C terminale (Ser232)

Aminoacidi coinvolti nell'interazione dominio-motivo:

-Ser505 del motivo interagisce con Arg57 del dominio;

-Asp129 e Ile504 del motivo interagiscono con la Leu226 del dominio.

Analizzando la struttura secondaria della PI3K (PDB ID: 3HHM), sono state ritrovate 9  $\alpha$  eliche (546 – 693) consecutive che però confrontate con la struttura terziaria del dominio (PDB ID: 2B05) non risultano essere corrispondenti; di conseguenza il dominio 14\_3\_3 non è contenuto né in p110 $\alpha$  né in p85.

Motivi: i motivi di legame RTQSSS (151 – 156) e RSPSIP (228 – 233) del dominio 14\_3\_3 di p85 non sono contenuti nella struttura cristallografica di PI3K (PDB ID: 3HHM).

Il motivo di legame RLCLSIC (401 – 407) è presente nel foglietto  $\beta$  di p110 $\alpha$ , il motivo KTVSQR (678 – 683) è presente in un' $\alpha$  elica di p110 $\alpha$ . I motivi RNSHTIM (852 – 858) e KAISTR (532 – 537) sono contenuti in un  $\alpha$  elica di p110 $\alpha$ .

Residui conservati del dominio 14\_3\_3:

Y20 (9 S) – D22 (7) – V30 (9 S) – L37 (9 S) – V52 (9 S) – K69 (9 F) – L101 (8) – K125 (9 F) – E136 (9 F) – L175 (9 S)

Residui corrispondenti nei raggi X della PDB ID: 3HHM (porzione 609-790):

Y606- D608- V616– L623 – V638 – K655 – L687 – K711 – E722- L76

#### Dominio PDZ\_3 PDB: 2KOH

I domini PDZ sono piccoli domini globulari, costituiti da  $\alpha$  eliche e foglietti  $\beta$ , presenti in diverse proteine dove svolgono funzioni importanti nel riconoscimento e nell'assemblaggio di complessi proteici.

Struttura secondaria: N-terminale (Gly579) – 3 foglietti  $\beta$  –  $\alpha$  elica – foglietto  $\beta$  –  $\alpha$  elica – foglietto  $\beta$  – C terminale (Glu689).

Considerando la struttura secondaria del dominio PDZ\_3 (PDB ID: 2KOH) e la predizione di struttura secondaria di p110 $\alpha$  e p85 $\alpha$  ottenuta con il server *Predict Protein*, non è stata

ritrovata alcuna successione che rispecchi quella della struttura secondaria del dominio. Di conseguenza il dominio PDZ\_3 non è presente né il p110α né in p85α.

Motivi che legano il dominio PDZ\_3 identificati in p110α → VECL (22 – 25); EEKI (109 – 112); FDMV (128 – 131); KEAV (148 – 151); LEDL (452 – 455); EDLL (453 – 456); MDAL (745 – 748); SELL (790 – 793).

Motivi che legano il dominio PDZ\_3 identificati in p85α → EEDI (19 – 22); PEEI (50 – 53); LEMI (174 – 177); EDNI (439 – 442); IEAV (442 – 445); NEWL (595 – 598); DEDL (613 – 616); NDSL (707 – 710), mentre i motivi presenti nella struttura cristallografica 3HHM\_SDPL (393 – 396); NESL (410 – 413); QDQV (433 – 436); YDRL (463 – 466); NETI (488 – 491); KEYI (506 – 509); PDLI (568 – 571). Tali motivi non sono tuttavia presenti nella struttura cristallografica di p85α.

La sequenza di riconoscimento del dominio PDZ comprende una porzione alifatica che si inserisce nella tasca alifatica del dominio. La zona polare del motivo ha residui negativi che interagiscono con residui positivi del dominio.

*Aminoacidi coinvolti (KOH):* 682 (R) 683 (R) 686 (R)

*Corrispondenti aminoacidi di PI3K (PDB ID: 3HHM):* 51(K) 54 (R) 55(K)

Zona idrofobica: Val, Leu, Ile

Aminoacidi coinvolti nell'interazione dominio-motivo: L (601) L (603) V (605) V (642) V (662) L (678)

Aminoacidi omologhi nel dominio PDZ\_3 (PDB ID:2KOH) e in PI3K (PDB ID: 3HHM):

PDZ\_3: I619 V621 I624

PI3K: V101 I102 V105

Sebbene da un lato non sia corretto affermare che i domini ritrovati in base alla somiglianza di struttura secondaria siano sicuramente presenti in PI3K in quanto non si sovrappongono ad essa perfettamente, dall'altro non si può escludere che un cambiamento conformazionale della proteina possa condurre ad una conformazione tale da determinare l'esatta sovrapposizione domini/proteina e confermare quindi la loro presenza nella proteina stessa.

Inoltre, la non perfetta conservazione dei residui strutturali/funzionali identificati da *ConSeq* potrebbe implicare l'interazione con motivi di riconoscimento non ancora identificati.

Sulla base del concetto sopra espresso, è stata effettuata un'analisi dei domini noti che caratterizzano le diverse isoforme di p110 per mettere in evidenza eventuali porzioni che cambiano conformazione in seguito al legame di Ras al dominio RBD.

Si è proceduto effettuando un'analisi bioinformatica delle strutture cristallografiche (non mutate e mutate) di PI3K da sola o in complesso con Ras. Le strutture considerate sono identificate con i seguenti codici PDB IDs: 3HHM, 2RD0, 3HIZ (p110 $\alpha$ ), 1HE8,1E8X (p110 $\gamma$ ). Sono stati identificati tutti i domini funzionali che caratterizzano le due diverse subunità catalitiche di PI3K (p110 $\alpha$  e p110 $\gamma$ ). I risultati sono riportati nella Tabella 8 seguente.

**Tabella 8:** domini funzionali e sequenze delle subunità p110 $\alpha$  e p110 $\gamma$ .

<b>p110<math>\alpha</math></b>		<b>p110<math>\gamma</math></b>	
<b>Dominio funzionale</b>	<b>Sequenza</b>	<b>Dominio funzionale</b>	<b>Sequenza</b>
<i>ABD</i>	1-108	<i>ABD</i>	Non presente
<i>RBD</i>	191-291	<i>RBD</i>	221-311
<i>C2</i>	330-480	<i>C2</i>	357-529
<i>ELICA</i>	525-696	<i>ELICA</i>	548-725
<i>DOMINIO CHINASICO</i>	696-1068	<i>DOMINIO CHINASICO</i>	726-1092
<i>REGIONE DI LEGAME DI ATTIVAZIONE</i>	940,953	<i>REGIONE DI LEGAME DI ATTIVAZIONE</i>	967,981
<i>DFG</i>	933-935	<i>DFG</i>	964-966
<i>LOBO N TERMINALE</i>	696-851	<i>LOBO N TERMINALE</i>	725-882
<i>LOBO C TERMINALE</i>	853-933	<i>LOBO C TERMINALE</i>	884-965

E' seguita un'approfondita analisi delle strutture in oggetto al fine di valutare i cambiamenti conformazionali nella struttura terziaria, dovuti alla diversa strutturazione/interazione dei domini funzionali identificati. A tal fine sono state effettuate tre diverse sovrapposizioni strutturali: i) sovrapposizione del dominio di legame con Ras, ii) sovrapposizione del lobo N-terminale, iii) sovrapposizione del dominio ad elica. I risultati ottenuti sono stati i medesimi indipendentemente dal tipo di sovrapposizione effettuata. Di seguito è quindi riportato un esempio del tipo di analisi eseguita, in particolare prendendo in considerazione la sovrapposizione effettuata utilizzando il dominio di legame di Ras.

Confronto tra le diverse strutture cristallografiche analizzate:

PDB ID: 1HE8-1E8X:

Presentano la medesima strutturazione, ma in PDB ID: 1E8X manca la porzione 254-267.

In *Predict Protein* la porzione 253-260 non è predetta avere alcun tipo di struttura, mentre la porzione 260-269 è altamente predetta avere la struttura di *loop*.

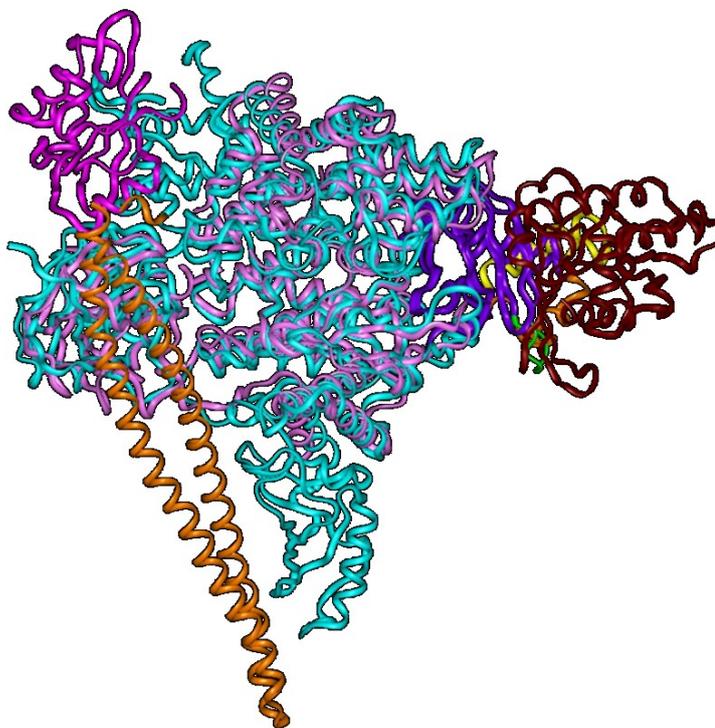
PDB ID: 3HHM-1HE8:

La PDB ID: 3HHM rispetto alla PDB ID: 1HE8 presenta una regione di legame di maggiore lunghezza che va dall'amminoacido 197 a 204. Un'ulteriore variazione si ha nella zona 253-269 di PDB ID: 1HE8 che in PDB ID: 3HHM corrisponde alla zona 229-249. Tali porzioni si strutturano in due regioni di legame e sono disposti in modo tale da essere uno l'immagine speculare dell'altro (Figure 15 e 16)

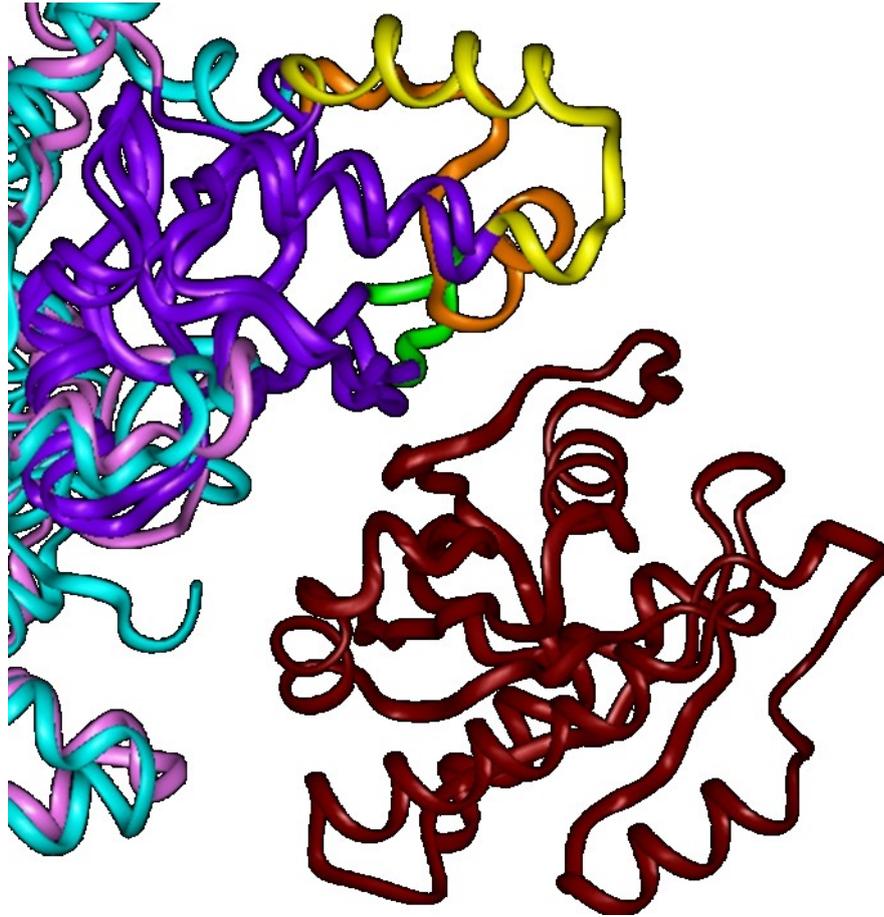
*Risultati del server Predict Protein:*

PDB ID: 1HE8: 253-260 è una porzione non predetta, mentre la sequenza che va da 260 a 269 è altamente predetta avere la struttura di *loop*.

PDB ID: 3HHM: 229-249 è una zona non predetta, eccetto per le brevi sequenze 233-235, 245-247 altamente predette a *loop*.



**Figura15: Sovrapposizione PI3K $\alpha$  (PDB ID:3HHM) con PI3K $\gamma$  (PDB ID:1HE8) e visualizzazione delle variazioni nel dominio che lega Ras (RBD):** in figura viene mostrata la subunità p110 $\alpha$  in ciano, quella p110 $\gamma$  in rosa, in arancione l'inter region della p85, in magenta l'SH2 della p85, in viola l'RBD sia di alfa che di gamma, in verde la regione di legame in più della p110 alfa, in giallo e in arancione le variazioni corrispondenti delle due subunità, in particolare in giallo la porzione di alfa e in arancione quella di gamma, in porpora viene indicata la proteina Ras.



*Figura16: particolare della sovrapposizione PI3K $\alpha$  (PDB ID:3HHM) con PI3K $\gamma$  (PDB ID:1HE8) e visualizzazione delle variazioni nel dominio che lega Ras (RBD): in figura viene mostrato in viola l'RBD sia di alfa che di gamma, in verde la regione di legame in più della p110 alfa, in giallo e in arancione le variazioni corrispondenti delle due subunità, in particolare in giallo la porzione di alfa e in arancione quella di gamma, in porpora viene indicata la proteina Ras.*

PDB ID: 3HHM-3HIZ:

Tra le due strutture si hanno delle variazioni nelle zone che vanno da 198-202, 229-238 in entrambi i casi la struttura di 3HIZ si restringe di più rispetto a quella di 3HHM.

PDB ID: 2RD0-3HHM:

Si verificano dei cambiamenti conformazionali nelle porzioni 198-202, 229-238.

*Risultati del server Predict Protein:*

- 198-202 è una zona altamente predetta avente struttura di *loop* in entrambe le proteine
- 229-238: zona non predetta, eccetto per la breve sequenza 233-235 altamente predetta avente struttura di *loop* in entrambe le proteine.

#### PDB ID: 2RD0- 3HIZ:

La porzione che va da 228-243 non è diversa conformazionalmente ma risulta essere sfalsata nella sua struttura terziaria, la porzione che cambia conformazione è, invece, quella che va da 198-202.

#### PDB ID: 2RD0-1HE8:

Le variazioni conformazionali si trovano nella regione 228-247 in 2RD0 che in 1HE8 corrisponde alla regione 252-267. Le due regioni nelle rispettive strutture sono disposte in direzione opposta l'una rispetto all'altra. Inoltre varia la strutturazione della regione di legame 196-205 della struttura cristallografica PDB ID:2RD0 rispetto al corrispondente struttura cristallografica PDB ID: 1HE8 (226-229).

Per implementare l'analisi precedentemente descritta e allo scopo di progettare nuovi inibitori del sito catalitico della PI3K è stata effettuata un'analisi dettagliata di tutte le strutture a raggi X relative alle diverse isoforme della PI3K. In particolare, è stata presa come riferimento la struttura cristallografica della PI3K $\gamma$  contenente l'ATP (PDB ID: 1E8X) e sono state sovrapposte ad essa sia le altre strutture a raggi X della PI3K che le strutture di altre chinasi lipidiche contenenti anche il substrato. Con questa metodologia è stato possibile caratterizzare il sito catalitico della PI3K individuando le tasche occupate dalle diverse porzioni di ciascun inibitore considerato.

Di seguito viene riportata l'analisi dettagliata del complesso PI3K $\gamma$  /ATP e schematicamente la suddivisione in tasche del dominio catalitico della PI3K con gli aminoacidi che lo caratterizzano e le corrispondenze nelle diverse isoforme.

#### ***2.4a.2 Analisi del Sito di Legame dell'ATP e delle sue interazioni con i ligandi co-cristallizzati***

L'ATP si inserisce nel dominio chinasi della PI3K: l'adenina si posiziona a livello dell'*hinge region* (regione a cerniera) (881-885; (numerazione della PI3K $\gamma$ )) contattando sia il lobo N-terminale (725-882) che il lobo C-terminale (884-965); i gruppi idrossilici dello zucchero sono rivolti verso il lobo C-terminale (884-965); i fosfati interagiscono con il G-loop (803-810) e con il motivo DFG (964-966). Il motivo DFG è una sequenza di tre residui Asp-Phe-Gly ed è localizzato nel sito catalitico immediatamente prima della regione di legame di attivazione. Nella sua conformazione attiva il motivo DFG adotta una conformazione detta DFG-in, in cui l'acido aspartico catalitico punta nella tasca di legame dell'ATP e la Phe punta lontano dall'ATP. Nelle strutture cristallografiche di alcune chinasi

il motivo DFG può adottare una differente conformazione in cui l'Asp e la Phe del DFG cambiano la posizione e puntano nella direzione opposta rispetto alla conformazione DFG-in, determinando così un movimento della Phe verso l'interno che ostruisce parzialmente il sito di legame dell'ATP. Questo diverso arrangiamento conformazionale crea una nuova tasca di legame idrofobica adiacente al sito di legame dell'ATP. La risultante conformazione DFG-out, risulta nell'inattivazione dell'attività enzimatica della PI3K. Di conseguenza il blocco conformazionale del motivo DFG nella conformazione inattiva (DFG-out) rappresenta un approccio molto attraente per la progettazione di nuovi inibitori allosterici delle chinasi.

I residui coordinanti i fosfati dell'ATP sono assolutamente conservati nei membri della famiglia delle inositolo 3-fosfato 3-chinasi (IP3-3K). Di seguito viene riportata in dettaglio l'analisi delle interazioni dell'ATP nelle diverse tasche evidenziate.

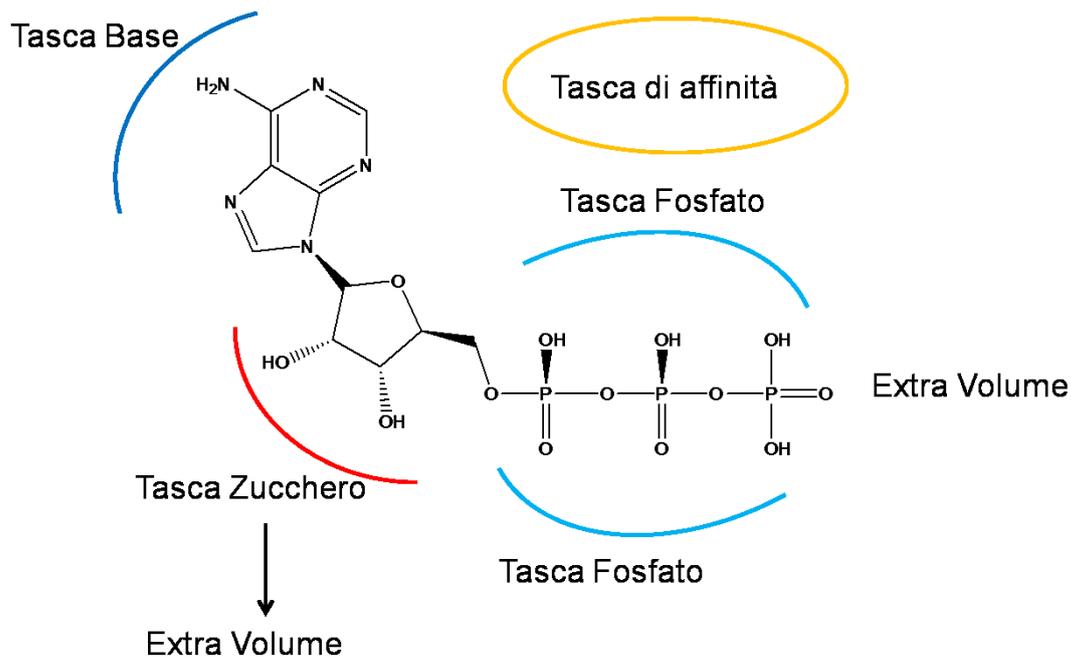
Nella struttura della p110 $\gamma$  (PDB ID: 1E8X) sono presenti le seguenti interazioni:

*Tasca base:* l'adenina si posiziona tra diversi aminoacidi alifatici che sono: Ile881 (regione a cerniera); Ile831 e Ile879 (lobo N-terminale); Met953 e Ile963 (lobo C-terminale); Pro810 (*G-loop*). Inoltre dà un'interazione a T sia con il Trp812 (*G-loop*) che con la Tyr867 (lobo N-terminale) e forma un legame idrogeno con l'NH dello scheletro peptidico della Val882 e con due molecole di acqua (A2227H-A2207H). Quest'ultime interagiscono a loro volta con l'Asp841 (Lobo N-terminale).

*Tasca zucchero:* la Thr887 interagisce con la Lys890 ed entrambe danno interazioni con i due gruppi idrossilici dello zucchero (lobo C-terminale). L'ossigeno del ribosio dà un'interazione con lo zolfo della Met804 (*G-loop*).

*Tasca fosfati:* i gruppi fosfato legano due ioni Lutenio (che in questa struttura sostituiscono il magnesio) di cui uno è legato al  $\beta$  e  $\gamma$  fosfato e l'altro si trova tra l' $\alpha$  e il  $\gamma$  fosfato. L' $\alpha$  fosfato dà un'interazione con la Lys833 che a sua volta interagisce con l'Asp836 (lobo N-terminale). Il  $\beta$  fosfato interagisce, invece, sia con la Lys833 (lobo N-terminale) che con la Ser806 (*G-loop*). Infine, il  $\gamma$  fosfato punta all'esterno del sito verso la tasca di legame dell'inositolo ed è posizionato tra l'Asn951 (lobo C-terminale) e Lys807, Lys808 e Lys809 (*G-loop*) ed interagisce attraverso la molecola d'acqua A2193H con l'Asp836. Lo ione Lutenio che lega il fosfato  $\gamma$  interagisce con l'Asn951 (lobo C-terminale) mentre il secondo ione Lutenio legato al  $\beta$  e  $\gamma$  fosfato, viene coordinato dall'Asp964 che a sua volta interagisce con la Lys833 (lobo N-terminale). Il Lutenio si lega anche all'Asp836 (lobo N-terminale).

La rappresentazione schematica delle tasche occupate dall'ATP è riportata nella figura successiva (*Figura 17*)



**Figura 17:** Visione schematica del sito di legame dell'ATP nella PI3K $\gamma$

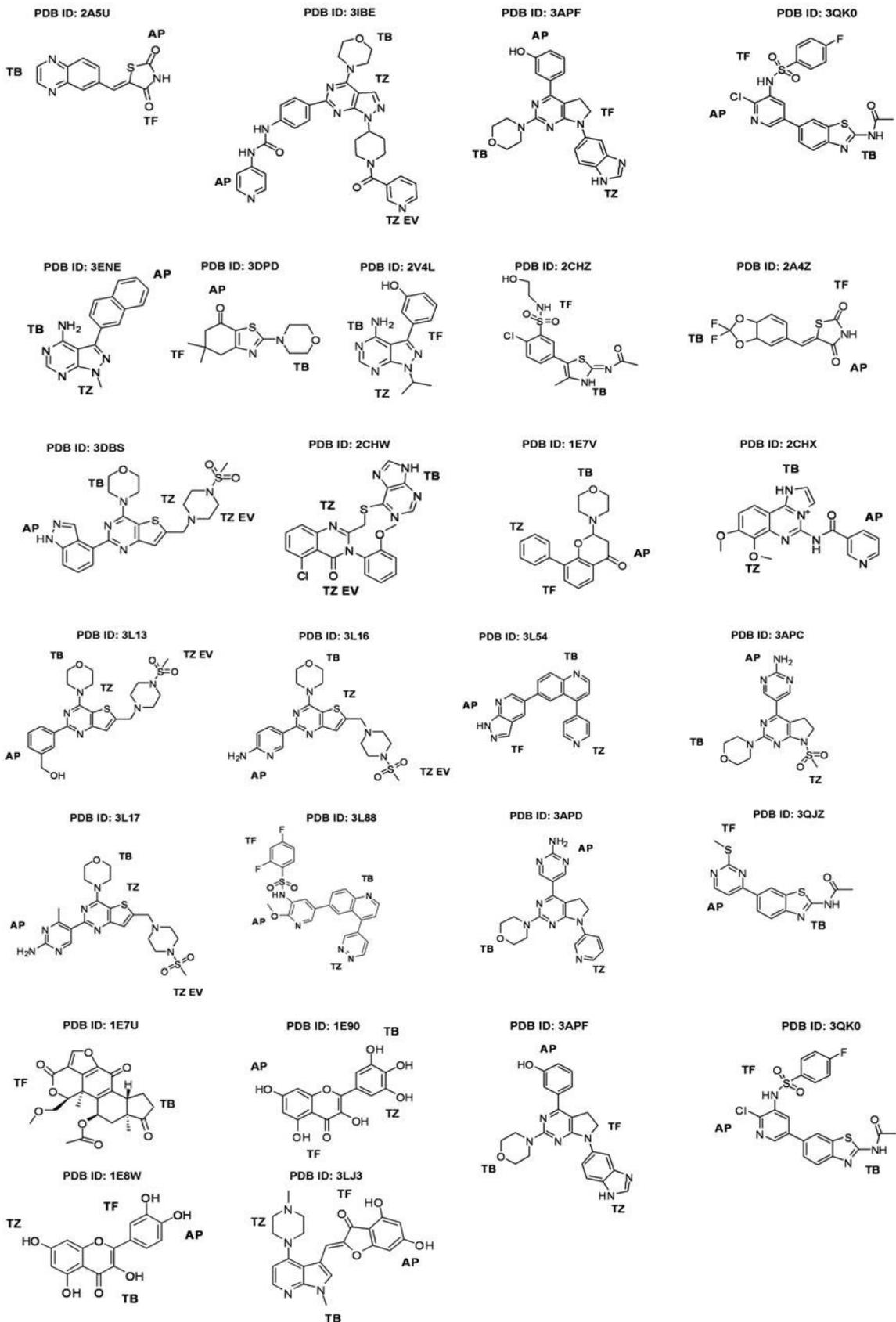
Di seguito è riportata una tabella (Tabella 9) che schematizza tutte le tasche d'interazione individuate nel dominio catalitico della PI3K con gli aminoacidi che lo caratterizzano e le corrispondenze nelle diverse isoforme.

**Tabella 9:** schema della suddivisione in tasche del sito catalitico e corrispondenze aminoacidi che nelle varie isoforme.

<b>TASCA BASE (TB)</b>		
<b>Gamma</b>	<b>Alfa</b>	<b>Delta</b>
Met953	Met922	Met900
Ile881	Val850	Val827
Ile831	Ile800	Ile877
Ile879	Ile848	Ile825
Pro810	Pro778	Pr0758
Ile963	Ile932	Ile910
Trp812	Trp780	Trp760
Phe961	Phe930	Phe908
<b>SOTTOTASCA DELLA VALINA (TASCA BASE)</b>		
<b>Gamma</b>	<b>Alfa</b>	<b>Delta</b>
Glu880	Glu849	Glu826
Ile881	Val850	Val827
Val882	Val851	Val828
<b>EXTRA VOLUME OCCUPATO DAI LIGANDI NEL SITO CATALITICO (EV)</b>		
<b>Gamma</b>	<b>Alfa</b>	<b>Delta</b>
Tyr867	Tyr836	Tyr813
Asp841	Asp810	Asp787

<b>TASCA DI AFFINITA' "AFFINITY POCKET" (AP)</b>		
<b>Gamma</b>	<b>Alfa</b>	<b>Delta</b>
Gly966	Gly935	Gly914
Gln840	Gln809	Gln886
Asp964	Asp933	Asp911
Lys833	Lys802	Lys779
Leu838	Leu807	Leu784
Asp841	Asp810	Asp787
Asp836	Asp805	Asp783
<b>ZONA TRA TASCA DI AFFINITA' E TASCA FOSFATO</b>		
<b>Gamma</b>	<b>Alfa</b>	<b>Delta</b>
Lys833	Lys802	Lys779
Asp964	Asp933	Asp911
<b>TASCA FOSFATO (TF)</b>		
<b>Gamma</b>	<b>Alfa</b>	<b>Delta</b>
( $\alpha$ - $\beta$ fosfato)		
Lys833	Lys802	Lys779
Asp964	Asp933	Asp911
Pro810	Pro778	Pro758
Asp836	Asp805	Asp783
Lys808	Lys776	Met756
B- $\gamma$ fosfato		
Lys807	Ala775	Lys755
Ser806	Ser774	Ser754
<b>TASCA ZUCCHERO (TZ)</b>		
<b>Gamma</b>	<b>Alfa</b>	<b>Delta</b>
Lys802	Gln859	Thr750
Thr887	Thr856	Thr833
Lys890	Arg770	Arg836
Met804	Met772	Met752
<b>TASCA INOSITOLO</b>		
<b>Gamma</b>	<b>Alfa</b>	<b>Delta</b>
Tyr757	Thr727	Thr706
Lys807	Ala775	Lys755
Lys808	Lys776	Met756
Gly835	Gly804	Gly781
Asp836	Asp805	Asp782
Leu1090	Thr1061	Arg1037
Asp964	Asp933	Asp911

Sulla base dell'analisi precedentemente descritta, è stata effettuata una rappresentazione schematica delle modalità di interazione di tutti gli inibitori cristallizzati nella PI3K. In particolare, sono state evidenziate le diverse tasche da loro occupate (*Figura 18*), utilizzando gli acronimi mostrati nella Tabella 9 precedente:



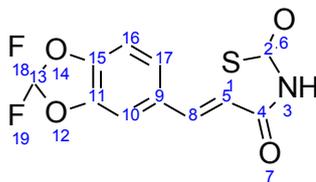
**Figura 18:** in figura è mostrato l'accomodamento delle diverse porzioni degli inibitori cristallizzati nel sito catalitico della PI3K utilizzando gli acronimi della Tabella 9.

Da tale analisi è emerso che sub-strutture uguali (per esempio l'anello pirimidinico delle strutture cristallografiche PDB ID:3IBE e 2V4L in *Figura 18*) ma connessi in maniera differente ad altri gruppi possono occupare tasche diverse oppure, lo stesso gruppo (per esempio l'anello 1,3-tiazolidin-2,4-dione delle strutture cristallografiche PDB ID: 2A5U, 2A4Z in *Figura 18*) occupa la stessa tasca ma si orienta in maniera opposta.

Sulla base di studi precedentemente effettuati che hanno permesso di suddividere il sito catalitico della PI3K in diverse tasche di legame, è stato possibile effettuare anche un'analisi dettagliata dell'accomodamento degli inibitori AS-605240 (PDB ID:2A5U) e AS-604850 (PDB ID:2A4Z) considerati perché presentano una porzione simile al composto di riferimento ST0358AA1.

### ***2.4a.3 Analisi dell'accomodamento degli inibitori AS-604850 ed AS605240 nel sito ATPasico***

L'inibitore AS-604850 (*Figura 19*) (PDB ID : 2A4Z) è anch'esso, come AS-605240 (PDB ID: 2A5U), un inibitore maggiormente selettivo verso la PI3K $\gamma$  ed analogamente è in grado di riprodurre alcune delle principali interazioni dell'ATP nel sito catalitico dell'enzima PI3K (sia nell' isoforma  $\alpha$  che nella  $\gamma$ ).



***Figura 19: struttura bidimensionale dell'inibitore AS-604850 con numerazione IUPAC.***

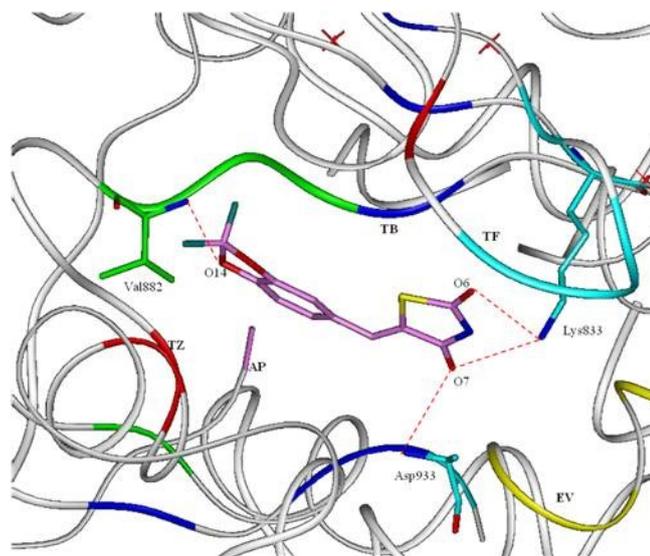
L'inibitore AS-604850 si inserisce nel dominio chinasi della PI3K $\gamma$  come descritto di seguito (*Figura 20*):

- L'anello 1,3-tiazolidin-2,4-dione si inserisce nella TF con il C2O6 rivolto verso l'AP: gli ossigeni O6 ed O7 possono dare entrambi un legame idrogeno con l'NH della catena laterale della Lys833(logo N), mimando l'interazione dell'O del  $\beta$ -fosfato dell'ATP con la stessa Lys833; l'NH dà interazione con la Lys833 ed Asp964.

Tale anello instaura interazioni idrofobiche di vdW con i seguenti residui aminoacidici: Asp964, Ile963, Tyr867, Pro810, Lys833, Ile879, Ile963.

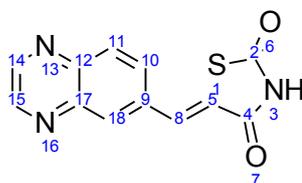
- L'anello 1,3-benzodiossolo si posiziona nella tasca base in maniera non del tutto planare : l'O14 dà un legame idrogeno con l'NH dello scheletro peptidico della Val882; il Fluoro dà un' interazione con la Ser854 (non appartiene ad alcuna tasca).

L'anello 1,3-benzodiossolico dà interazioni idrofobiche con : Ile831, Ile879 , Ile881 ,Val882, Met953, Ile963, mimando la base dell'ATP. La modalità di interazione con l'enzima da parte dell'inibitore AS-604850 verrà da qui in poi indicato con il termine di modalità di legame I.



**Figura 20:** Accomodamento dell'inibitore AS-604850 (rosa) nel sito catalitico della PI3K. Sono riportati gli aminoacidi chiave di interazione del sito con l'inibitore. Sono indicate: la tasca della valina (TV,) in verde; la tasca dove si posiziona lo zucchero dell'ATP (TZ,) in rosso; la tasca dove si posiziona la base dell'ATP (TB) in blu; la tasca dove si posiziona il fosfato dell'ATP (TF) in ciano; una prima regione dove si può accomodare un eventuale volume molecolare aggiuntivo, rispetto a quello dell'ATP, (EV), in rosa; una seconda regione dove si può accomodare un eventuale volume molecolare aggiuntivo, rispetto a quello dell'ATP, detta tasca di affinità (AP) in giallo. Gli eteroatomi del ligando sono colorati sulla base dei tipi atomici (N: blu, O: rosso, F: verde scuro e S: giallo)

Per quanto riguarda l'altro inibitore, AS-605240 (Figura 21) (PDB ID: 2A5U), esso è maggiormente selettivo verso la PI3K $\gamma$  rispetto alle altre isoforme.

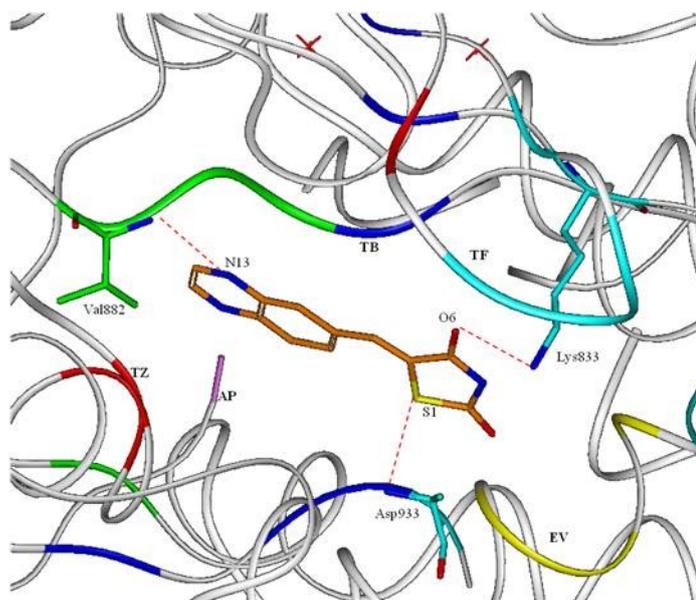


**Figura 21:** struttura bidimensionale dell'inibitore AS-605240 con numerazione IUPAC

Gli studi condotti hanno mostrato che questo composto ha un accomodamento nel sito catalitico in grado di riprodurre alcune delle principali interazioni dell'ATP in diverse tasche enzimatiche (sia nella isoforma  $\alpha$  che nella  $\gamma$ ).

L'inibitore AS-605240 si inserisce nel dominio chinasi della PI3K $\gamma$  come descritto di seguito (Figura 22):

- L'anello 1,3-tiazolidin-2,4-dione si accomoda nella tasca fosfato (TF), con una funzione carbonilica direzionata verso la tasca di affinità (AP) ; l'O6 dà un legame a idrogeno con l'NH della catena laterale della Lys833 (lobo N), mimando l'interazione dell'O del  $\beta$ -fosfato dell'ATP con la stessa Lys833; l' S1 interagisce con l'NH dello scheletro peptidico dell'Asp964 (lobo C), mimando l'interazione del Lutenio legato a  $\beta$  e  $\gamma$ -fosfato con lo stesso aminoacido;
- l'anello 1,3-tiazolidin-2,4-dione dà interazioni di vdW con i residui di Tyr867 (lobo N), Asp964 (lobo C), Ile963 (lobo C), Ile831, Ile879, Met953.
- l'anello chinossalinico si accomoda in tasca base (TB); l'N13 dà un ponte a idrogeno con l'NH dello scheletro peptidico della Val882 . Tale anello dà interazioni di vdW con i residui di Ile831 (lobo N), Ile881 (regione a cerniera), Val882 (regione a cerniera), Ile879 (lobo N), Ile963 (lobo C), Met953, mimando la base dell'ATP; instaura, inoltre, interazione di tipo  $\pi$  con i residui di Tyr867 e Phe961. La modalità di interazione con l'enzima da parte dell'inibitore AS-605240 verrà da qui in poi indicato con il termine di modalità di legame II.



**Figura 22:** Accomodamento dell'inibitore AS-605240 (arancione) nel sito catalitico della PI3K. Alcune interazioni chiave con gli aminoacidi presenti nel sito sono mostrate. Sono indicate: la tasca della valina (TV,) in verde; la tasca dove si posiziona lo zucchero dell'ATP (TZ,) in rosso; la tasca dove si posiziona la base dell'ATP (TB) in blu; la tasca dove si posiziona il fosfato dell'ATP (TF) in ciano; una prima regione dove si può accomodare un eventuale volume molecolare aggiuntivo, rispetto a quello dell'ATP, (EV), in rosa; una seconda regione dove si può accomodare un eventuale volume molecolare aggiuntivo, rispetto a quello dell'ATP, detta tasca di affinità (AP) in giallo. Gli eteroatomi del ligando sono colorati sulla base dei tipi atomici (N: blu, O: rosso e S: giallo).

L'analisi delle strutture cristallografiche condotta sui composti AS-605240 e AS-604850 mostra come l'inibizione enzimatica si realizza attraverso l'instaurarsi di due principali interazioni nel sito di legame per l'ATP posto nella subunità catalitica dell'enzima:

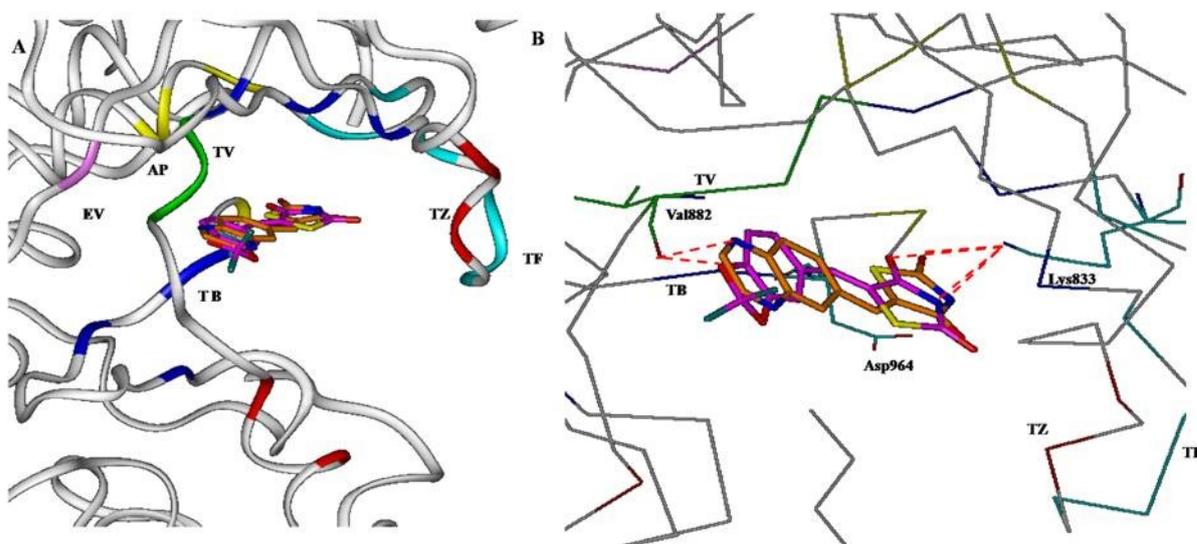
a) un ponte ad idrogeno tra l'N dell'anello 1,3-tiazolidin-2,4-dione e l'NH della catena laterale della Lys833;

b) un secondo ponte ad idrogeno tra l'N dell'anello chinossalinico o l'ossigeno dell'anello 1,3-benzodiossolico e l'NH dello scheletro peptidico della Val882.

Le differenze nell'accomodamento nel sito catalitico dei due inibitori sono, dunque, relative al fatto che l'anello 1,3-tiazolidin-2,4-dione si posiziona per entrambi i ligandi in TF, ma in maniera differente. Infatti, l'inibitore AS-604850 si posiziona nel sito in modo che l'O7 punta verso l'AP mentre l'inibitore AS-605240 si accomoda in modo tale che è l'O6 a puntare verso l'AP. Questi ossigeni danno le stesse interazioni, con la differenza che gli ossigeni dell'inibitore AS-604850 possono dare entrambi interazione con la Lys833 mentre nel caso dell'inibitore AS-605240 solo l'ossigeno O6 dà la stessa interazione.

L'S1 dell' AS-605240 dà un legame a idrogeno con l'NH dello scheletro peptidico dell'Asp964, interazione che non si osserva per AS-604850.

L'anello 1,3 benzodiossolico non si accomoda in modo planare all'interno della TB dell'enzima, a differenza dell'anello chinossalinico (*Figura 23*)

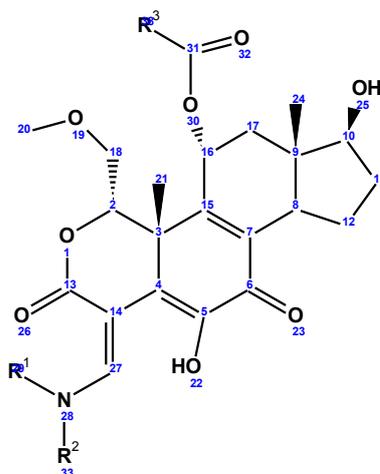


**Figura 23:** La figura A mostra le differenze nell'accomodamento dei due inibitori nel sito. In particolare, in arancio è mostrato l'inibitore AS-605240; mentre in magenta AS-604850. La figura B mette in evidenza le interazioni che i due binding consentono di avere con gli aminoacidi del sito catalitico. Sono indicate: la tasca della valina (TV,) in verde; la tasca dove si posiziona lo zucchero dell'ATP (TZ,) in rosso; la tasca dove si posiziona la base dell'ATP (TB) in blu; la tasca dove si posiziona il fosfato dell'ATP (TF) in ciano; una prima regione dove si puo' accomodare un eventuale volume molecolare aggiuntivo, rispetto a quello dell'ATP, (EV), in rosa; una seconda regione dove si puo' accomodare un eventuale volume molecolare aggiuntivo, rispetto a quello dell'ATP, detta tasca di affinità (AP) in giallo. Gli eteroatomi dei ligandi sono colorati sulla base dei tipi atomici (N: blu, O: rosso, F: verde scuro e S: giallo)

## 2.4b Analisi delle Relazioni Struttura-Attività (SARs) degli inibitori noti

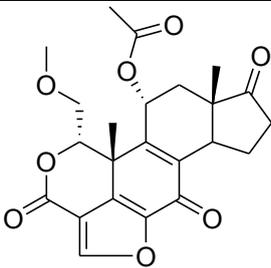
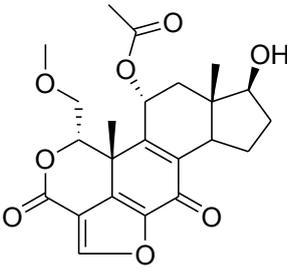
Allo scopo di progettare nuovi inibitori specifici per il sito catalitico di PI3K è stata effettuata, per cominciare, un'analisi delle relazioni struttura-attività (SARs) di alcuni inibitori noti. Nelle Tabelle 10 e 11 sono riportati i risultati dell'analisi SARs dei derivati del wortmannin. Tali dati sono stati combinati con gli studi computazionali eseguiti sia sugli inibitori noti che su quelli forniti dall'Industria Farmaceutica Sigma-Tau, riportati nella Tabella 12.

**Tabella 10:** Struttura e attività inibitoria dei derivati del wortmannin



Composti	R1	R2	R3	PI3K <i>IC</i> <sub>50</sub> μM	mTOR <i>IC</i> <sub>50</sub> μM
5	-CH <sub>2</sub> CH <sub>3</sub>	-CH <sub>2</sub> CH <sub>3</sub>	-COCH <sub>3</sub>	0.382	14
6	-CH <sub>2</sub> CH <sub>3</sub>	-CH <sub>2</sub> CH <sub>3</sub>	H	7.850	>20
7	-CH <sub>2</sub> CH <sub>3</sub>	H	-COCH <sub>3</sub>	>10	>20
8	-C(CH <sub>3</sub> ) <sub>3</sub>	-(CH <sub>2</sub> ) <sub>2</sub> OH	-COCH <sub>3</sub>	1.110	28.60
9	-C(CH <sub>3</sub> ) <sub>3</sub>	-CH <sub>3</sub>	-COCH <sub>3</sub>	0.440	6.50
10	-C(CH <sub>3</sub> ) <sub>3</sub>	H	-COCH <sub>3</sub>	>10	>20
11	-(CH <sub>2</sub> ) <sub>3</sub> CH <sub>3</sub>	-CH <sub>3</sub>	-COCH <sub>3</sub>	0.635	8.70
12	-(CH <sub>2</sub> ) <sub>4</sub> -	-	-COCH <sub>3</sub>	0.496	>20
13	-(CH <sub>2</sub> ) <sub>3</sub> N(CH <sub>3</sub> ) <sub>2</sub>	-CH <sub>3</sub>	-COCH <sub>3</sub>	0.006	2.92
14	-(CH <sub>2</sub> ) <sub>3</sub> N(CH <sub>3</sub> ) <sub>2</sub>	H	-COCH <sub>3</sub>	>10	>20
15	-(CH <sub>2</sub> ) <sub>3</sub> N(CH <sub>3</sub> ) <sub>2</sub>	-CH <sub>3</sub>	H	0.638	>20
16	-(CH <sub>2</sub> ) <sub>2</sub> N(CH <sub>3</sub> ) <sub>2</sub>	-C(CH <sub>3</sub> ) <sub>3</sub>	-COCH <sub>3</sub>	0.009	0.82
17	-(CH <sub>2</sub> ) <sub>2</sub> N(CH <sub>3</sub> ) <sub>2</sub>	-CH <sub>2</sub> Ph	-COCH <sub>3</sub>	0.066	5.60
4	-CH <sub>2</sub> CCH	-CH <sub>2</sub> CCH	-COCH <sub>3</sub>	0.088	3.10

**Tabella 11:** Struttura ed attività inibitoria del wortmannin (1) e del 17-idrossi wortmannin (2)

Composti	Struttura	PI3K <i>IC</i> <sub>50</sub> $\mu$ M	mTOR <i>IC</i> <sub>50</sub> $\mu$ M
1		0.001	ND
2		0.003	0.19

Dall'analisi delle relazioni struttura attività del wortmannin e dei suoi derivati emerge che le sostituzioni effettuate sui sostituenti  $R_1$ ,  $R_2$  ed  $R_3$  possono essere sfruttate per approfondire le basi molecolari della selettività di inibizione fra PI3K ed mTOR:

-  $R_3$ - Una variazione del sostituito in tale posizione (-COCH<sub>3</sub>/H) porta ad una perdita di attività di un ordine di grandezza su PI3K, ed a una perdita totale di attività su mTOR. Altre interessanti considerazioni possono essere fatte per quanto riguarda i sostituenti sull'azoto enamminico,  $R_1$  ed  $R_2$ ;

-  $R_2$ - In particolare, quando  $R_2$ =H il composto perde di attività su entrambi gli enzimi, mentre quando  $R_2$  è sostituito con una catena alchilica, l'attività su mTOR varia più sensibilmente di quella su PI3K. Infatti, quando  $R_2$ =CH<sub>2</sub>-CH<sub>2</sub>-OH il composto è meno attivo, così come quando l'azoto viene chiuso in un ciclo pirrolidinico con  $R_1$  ed  $R_2$ . Pertanto, l'assenza di un idrogeno in posizione  $R_2$  è un requisito cruciale per l'attività su entrambi gli enzimi, mentre quando  $R_2$  è sostituito con un alchile, la natura del sostituito è in grado di modulare l'attività su mTOR. La situazione cambia quando si introducono sostituenti aromatici, infatti, se  $R_2$ =CH<sub>2</sub>-Phe il composto perde di un ordine di grandezza di attività su PI3K mentre non viene condizionata l'attività su mTOR.

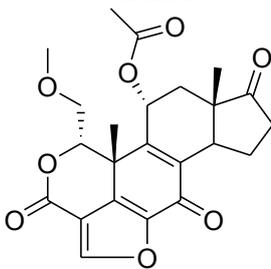
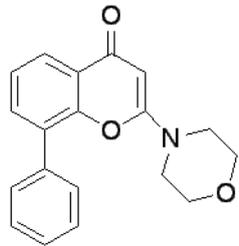
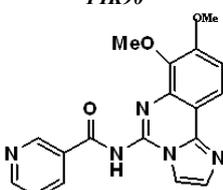
-  $R_1$ - Per quanto riguarda l'influenza del sostituito su  $R_1$ , in questo caso è evidente la tolleranza di entrambi gli enzimi verso la variazione dell'ingombro sterico dei sostituenti in tale posizione. In particolare mTOR guadagna di un ordine di grandezza in attività quando

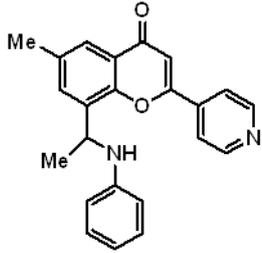
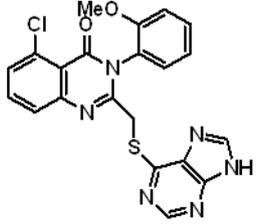
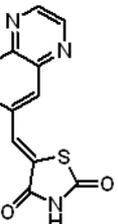
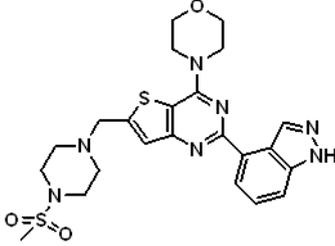
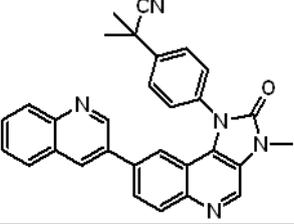
$R_1 = -(CH_2)_3CH_3$ . Quando la posizione  $R_1$  viene sostituita con una catena propilamminica, caricata positivamente a pH fisiologico, il risultato è un aumento di un ordine di grandezza della capacità inibitoria verso mTOR e di due ordini di grandezza di quella verso PI3K. Incrociando questi risultati con i dati relativi al 17-idrossi-wortmannin, emerge che la maggiore attività di quest'ultimo rispetto ai derivati N-sostituiti sembra attribuibile alla mancanza di ingombro sterico da un lato del doppio legame C14-C27 e/o alla presenza di un legame idrogeno intramolecolare.

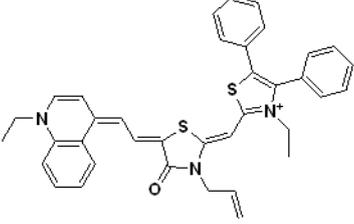
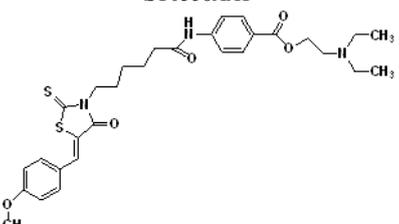
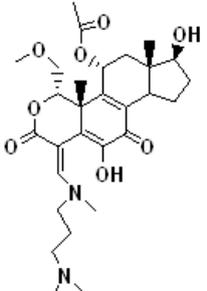
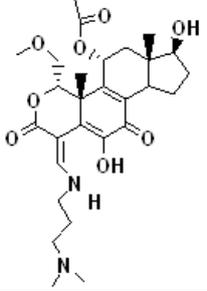
### 2.4c Analisi Conformazionale dei composti di riferimento e dei nuovi lead

E' stata effettuata un'estesa analisi conformazionale sulle strutture degli inibitori forniti dalla Sigma-Tau quali ST0358AA1, ST7318ES1 e dei seguenti composti di riferimento: PIK39, LY294002, TGX286, PIK90, AS605240, wortmannin, GDC0941, NVP\_BEZ235, PIK75. In tabella 12 sono riportate le strutture e i valori di inibizione, verso la subunità p110 di PI3K ed mTOR dei composti considerati.

**Tabella 12:** Struttura ed attività biologica del set di molecole in esame

NOME MOLECOLA	PI3K $IC_{50}$ $\mu M$				mTOR $IC_{50}$ $\mu M$	
	p110 $\alpha$	p110 $\beta$	p110 $\gamma$	p110 $\delta$	TORC1	TORC2
<p><i>Wortmannin</i></p> 	0.001	0.014	0.005	0.009	ND	ND
<p><i>LY294002</i></p> 	0.630	0.340	1.601	6.000	ND	ND
<p><i>PIK90</i></p> 	0.011	0.350	0.018	0.058	1.05	ND

<p><i>TGX286</i></p> 	4.500	0.120	10.0	1	>100	ND
<p><i>PIK75</i></p> 	0.003	0.850	0.040	ND	ND	ND
<p><i>PIK39</i></p> 	200	11	17	0.180	>100	>100
<p><i>AS605240</i></p> 	0.006	0.027	0.008	0.300	ND	ND
<p><i>GDC0941</i></p> 	0.003	0.033	0.075	0.003	ND	ND
<p><i>NVP-BEZ235</i></p> 	0.004	0.075	0.005	0.070	ND	ND

<p><b>ST7318SE1</b></p> 	>10	>10	>10	>10	>10	10
<p><b>ST0358AA1</b></p> 	>10	>10	>10	>10	>10	>10
<p><b>WORT13</b></p> 	0.0064	0.013	0.008	0.011	2.92	2.92
<p><b>WORT14</b></p> 	>10	>10	>10	>10	>10	>10

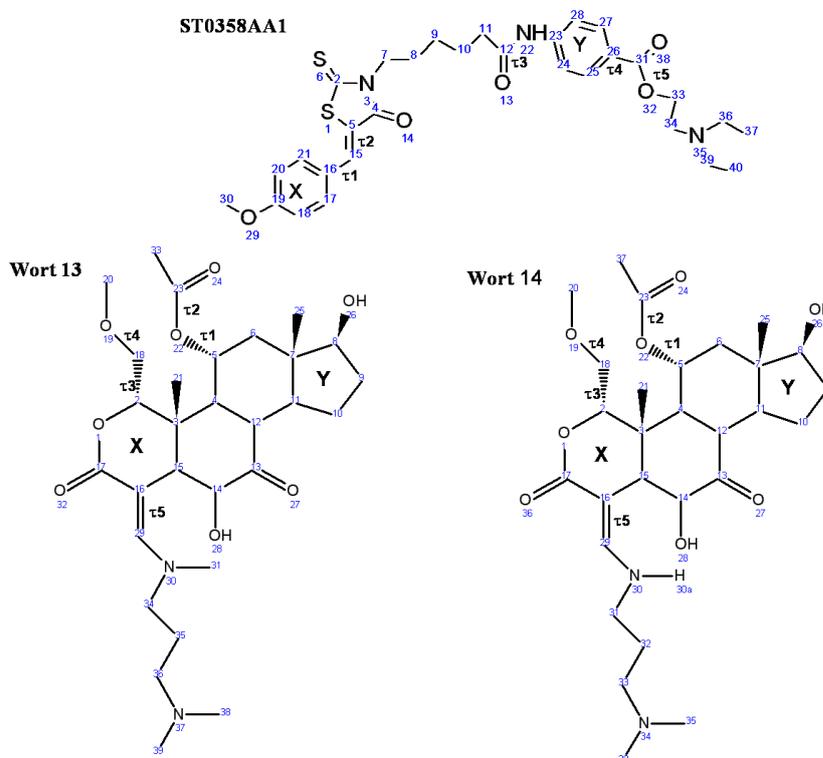
Come riportato in tabella 12 il composto ST7318SE1 ha una discreta attività su mTOR e una citotossicità inferiore a 20 $\mu$ M, è un prodotto non attivo su PI3K; mentre il composto ST0358AA1 non è attivo su mTOR ed è attivo su PI3K, ha una citotossicità inferiore a 20 $\mu$ M ed uno scheletro molecolare dotato di migliori proprietà farmacocinetiche, ed è pertanto stato selezionato come scheletro molecolari di partenza (*lead*) per effettuare il processo di ottimizzazione del profilo farmacologico.

È stata effettuata un'approfondita analisi conformazionale su tutti i composti in esame come descritto nella parte sperimentale. I conformeri ottenuti sono stati classificati secondo criteri geometrici ed energetici. Di seguito sono riportati i risultati dell'analisi conformazionale di tre composti : Wort13, Wort14 ed ST0358AA1, sui quali è stata posta una particolare attenzione

e che hanno permesso poi di progettare le modifiche da apportare al *lead* ST0358AA1. Per quanto riguarda il composto di riferimento Wort14 i calcoli di pKa effettuati con il software ACDLab (vedere parte sperimentale) hanno indicato la presenza di due diverse forme ioniche in equilibrio fra loro: una forma protonata (da noi chiamata Wort14) predetta essere presente al 52% ed una forma zwitterionica deprotonata sull'ossigeno ossidrilico (da noi chiamata Wort14B) predetta essere presente al 48%. Di conseguenza, tutti i calcoli sono stati condotti su entrambe le forme ioniche. In accordo con i loro valori calcolati di pKa, sia il Wort13 che l' ST0358AA1 sono stati, invece, considerati solo nella forma protonata.

Tutti i 200 conformeri ottenuti per ciascun composto sono stati, quindi, classificati, applicando i seguenti criteri:

- i. I valori di energia conformazionale.
- ii. La distanza interatomica tra i due sistemi aromatici X e Y, e l'atomo di azoto, denominato N<sub>35</sub> in ST0358AA1; N<sub>37</sub> in Wort13 ed N<sub>34</sub> in WORT14, come indicato nella successiva Figura 24.
- iii. I legami idrogeno intramolecolari.
- iv. Valori degli angoli torsionali indicati nella successiva Figura 24 .



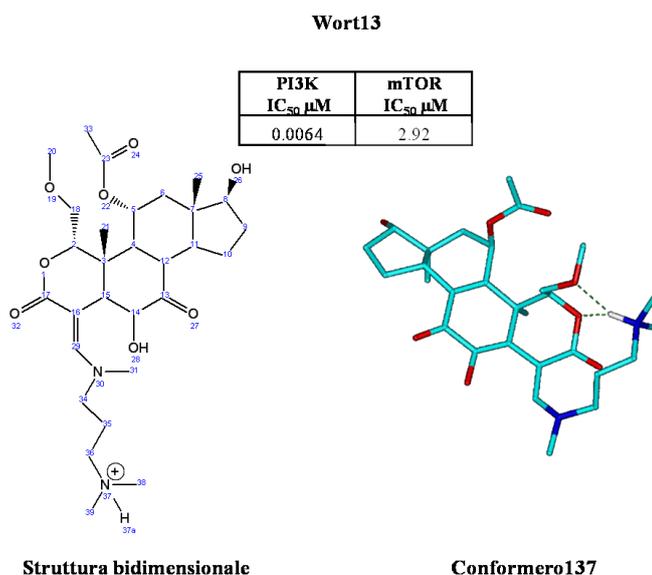
**Figura 24:** in figura sono indicati: i sistemi aromatici (X,Y) l'atomo di azoto N (N35 di ST0358AA1; N37 del Wort13 ed N34 del Wort14) per il calcolo delle distanze interatomiche, e gli angoli torsionali  $\tau_1$  (C6C5O22C23),  $\tau_2$  (C5O22C23C37)  $\tau_3$  (C3C2C18O19)  $\tau_4$  (C2C18O19C20)  $\tau_5$  (C15C16C29N30) dei composti Wort 13, Wort14;  $\tau_1$  (C21C16C15C5),  $\tau_2$  (C15C5S1C2)  $\tau_3$  (O13C12N22H23)  $\tau_4$  (C25C26C31O38)  $\tau_5$  (C26C31O32C33) del composto ST0358AA1, considerati per la classificazione dei conformeri ottenuti dalla ricerca conformazionale.

In accordo con l'analisi SARs riportata nel paragrafo precedente i risultati ottenuti hanno indicato che la differente preferenza conformazionale della catena basica dei composti di riferimento è relativa alla presenza di legami idrogeno intra-molecolari fortemente polarizzati.

## 2.4d Identificazione dei Requisiti Farmacoforici necessari per l'Inibizione del sito ATPasico della PI3K

In seguito all'analisi conformazionale dei derivati del wortmannin e del *lead*, sono stati selezionati i conformeri a più bassa energia. In particolare, la selezione è stata effettuata confrontando le distanze farmacoforiche, riportate nelle tabelle di seguito, dei diversi derivati del wortmannin con quelle del nostro *lead* ST0358AA1. Successivamente è stata effettuata una sovrapposizione di tali conformeri nella struttura a raggi X del wortmannin stesso legato al sito ATPasico della PI3K. In tal modo, è stato possibile selezionare la possibile conformazione bioattiva sia dei derivati del wortmannin che di ST0358AA1. Sono poi stati valutati gli amminoacidi presenti nel sito catalitico di mTOR attraverso l'allineamento delle sequenze dei due enzimi (PROMALS 3D <http://prodata.swmed.edu/promals3d/promals3d.php>). Di seguito sono riportati i risultati più rappresentativi ottenuti.

Analisi delle distanze farmacoforiche del Wort 13 (*Figura 25*)

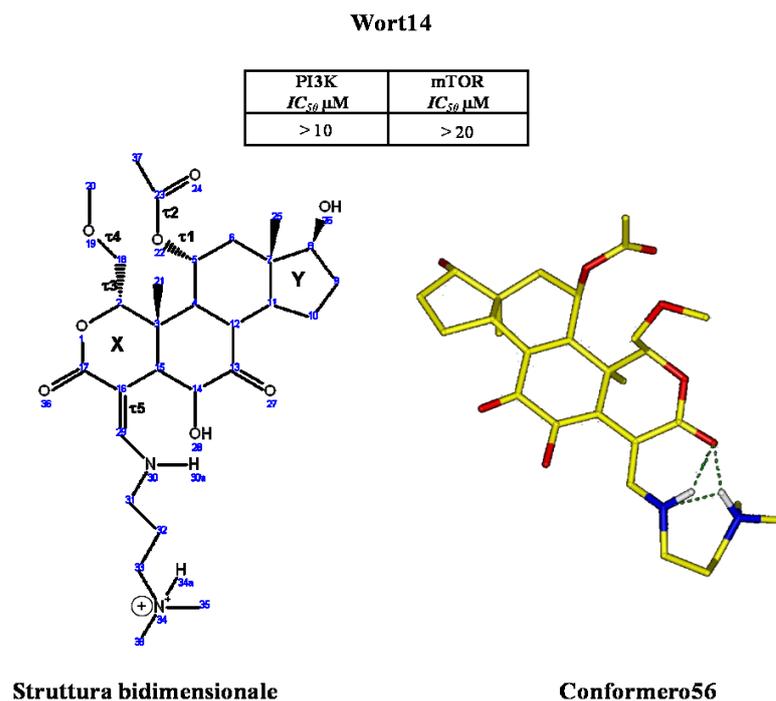


**Figura 25:** in figura sono rappresentati la struttura bidimensionale con numerazione IUPAC; l'immagine tridimensionale della conformazione bioattiva (conformero 137) del Wort 13 (celeste) e le attività del composto su PI3Kα ed mTOR. Gli idrogeni sono stati omessi per motivi di chiarezza, ad eccezione di quello legato all'azoto amminico. Gli atomi del ligando sono colorati sulla base dei tipi atomici (H: bianco, N: blu e O: rosso).

**Tabella 13:** Analisi delle distanze farmacoforiche dei minimi ottenuti dalla ricerca conformazionale

Conformero	$\Delta E_{GM}$	X-Y	Y-N37	X-N37
137	0	6.73	3.97	10.31
29	3.76	6.69	4.91	11.14
14	4.96	6.69	4.20	9.70

Analisi delle distanze farmacoforiche del Wort 14 (*Figura 26*)



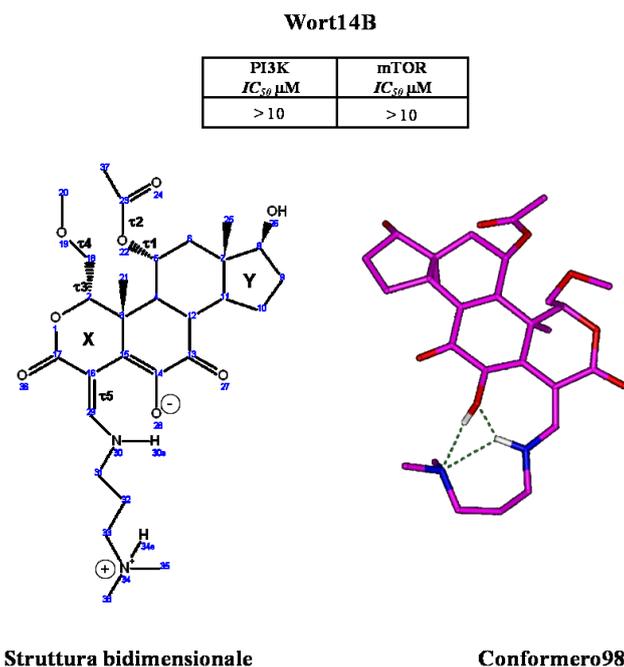
**Figura26:** in figura sono rappresentati la struttura bidimensionale con numerazione IUPAC; l'immagine tridimensionale della conformazione bioattiva (conformero 56) del Wort 14 (giallo) e le attività del composto su PI3K $\alpha$  ed mTOR. Gli idrogeni sono stati omessi per motivi di chiarezza, ad eccezione di quelli legati agli azoti amminico ed enamminico. Gli atomi del ligando sono colorati sulla base dei tipi atomici (H: bianco, N: blu e O: rosso).

Forma protonata presente al 52 % in accordo con il calcolo del pKa effettuato con ACDLab (vedere parte sperimentale).

**Tabella 14:** Analisi delle distanze farmacoforiche dei minimi ottenuti dalla ricerca conformazionale.

Conformero	$\Delta E_{GM}$	X-Y	Y-N34	X-N34
56	0	6.72	4.95	11.08
6	1.09	6.69	4.28	9.79
99	2.23	6.69	4.78	10.97
180	3.56	6.71	5.18	11.84

Analisi delle distanze farmacoforiche del Wort 14B (Figura 27).



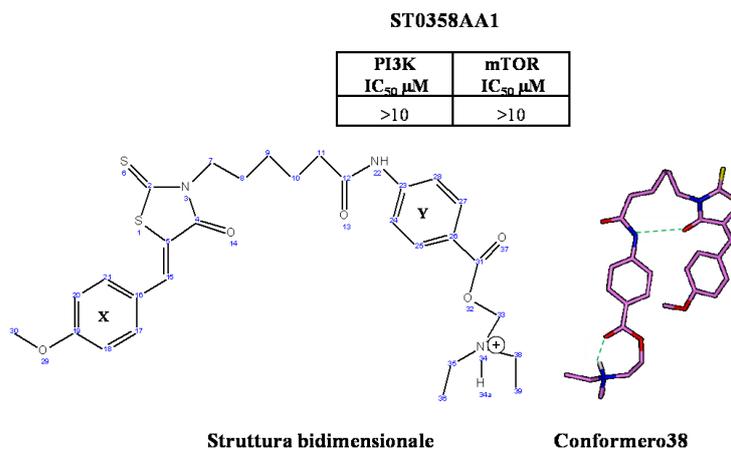
**Figura 27:** in figura sono rappresentati la struttura bidimensionale con numerazione IUPAC; l'immagine tridimensionale della conformazione bioattiva (conformero 98) del Wort 14B (magenta) e le attività del composto su PI3K $\alpha$  ed mTOR. Gli idrogeni sono stati omessi per motivi di chiarezza, ad eccezione di quelli legati agli azoti amminico ed enamminico. Gli atomi del ligando sono colorati sulla base dei tipi atomici (H: bianco, N: blu e O: rosso).

Forma zwitterionica presente al 48 % in accordo con il calcolo del pKa effettuato con ACDLab.

**Tabella 15:** Analisi delle distanze farmacoforiche dei minimi ottenuti dalla ricerca conformazionale.

Conformero	$\Delta E_{GM}$	X-Y	Y-N2	X-N2
<b>98</b>	0	6.71	6.03	7.55
<b>118</b>	2.57	6.71	5.15	9.11
<b>68</b>	3.06	6.70	5.10	9.35
<b>33</b>	4.08	6.69	4.60	8.94

Analisi delle distanze farmacoforiche del *lead* ST0358AA1 (Figura 28)



**Figura 28:** in figura sono rappresentati la struttura bidimensionale con numerazione IUPAC; l'immagine tridimensionale della conformazione bioattiva (conformero 38) di ST0358AA1 (rosa) e le attività del composto su PI3K $\alpha$  ed mTOR. Gli idrogeni sono stati omessi per motivi di chiarezza, ad eccezione di quello legato all'azoto amminico. Gli atomi del ligando sono colorati sulla base dei tipi atomici (H: bianco, N: blu e O: rosso).

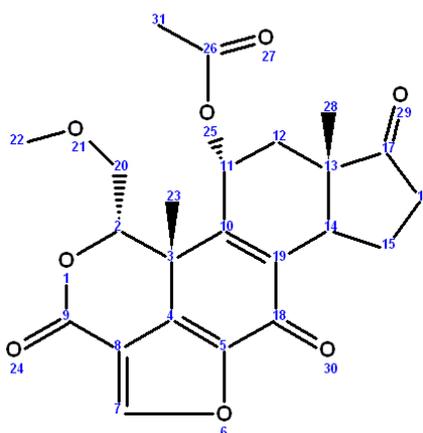
**Tabella 16:** Analisi delle distanze farmacoforiche della conformazione bioattiva del composto di riferimento.

Conformero	$\Delta E_{GM}$	X-Y	Y-N2	X-N2
<b>38</b>	3.95	4.96	6.10	9.11

### 2.4d.1 Analisi della Struttura a raggi X del Wortmannin legato al Sito ATPasico della PI3K:

#### Wortmannin

PI3K <i>IC</i> <sub>50</sub> $\mu$ M	mTOR <i>IC</i> <sub>50</sub> $\mu$ M
0.001	ND



**Figura29:** in figura è rappresentata la struttura bidimensionale con numerazione IUPAC del wortmannin e le attività del composto su PI3K $\alpha$  ed mTOR

#### Interazioni idrofobiche

**PI3K:** le Leu848, Leu800, Leu932 Met772, Pro778, Met922, Val850 e Val851 determinano la forma della tasca idrofobica dove si accomoda l'inibitore, a ciò partecipano, dando anche interazioni a trasferimento di carica, la Phe930 e la Tyr836.

**mTOR:** la Met772 è un'Ile, l'Ile800 è una Leu, la Val850 è un Trp, la Phe961 è una Leu, la Phe930 diventa una Leu. La Pro778, la Tyr836, l'Ile848, la Met922, l'Ile932 e la Val851 sono conservati.

#### Legami idrogeno

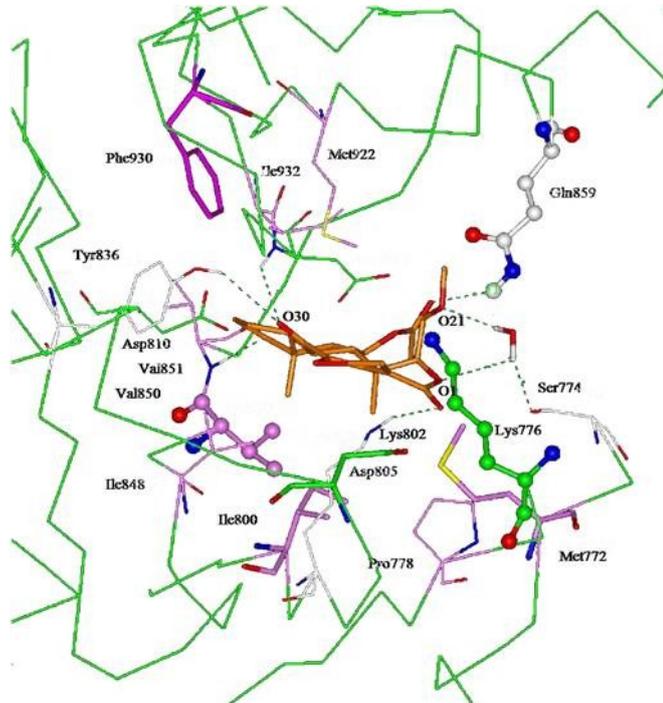
**PI3K:** gli amminoacidi idrofili che circondano l'inibitore sono:, i) la Ser774 forma un legame idrogeno con l'ossigeno etero in catena laterale (O21) e con l'ossigeno etero del lattone (O1) attraverso una molecola d'acqua, e, infine, ii) l'Asp933 (legato alla Lys776), iii) l'OH della Tyr836 dà un legame idrogeno con il carbonile C-18 dell'inibitore, iv) L'NH dello scheletro peptidico della Val851 dà un legame idrogeno con l'ossigeno carbonilico in C-17, v) L'NH dello scheletro peptidico dell'Asp933 dà un legame idrogeno con l'ossigeno del carbonile C-18, vi) la Gln859 dà interazione di legame H con l'acetile su R<sub>3</sub> (O27).

**mTOR:** La Ser774, l'Asp933 e l'Asp810 sono conservati. La Lys776 diventa Gln, l'Asp805 è un Glu e la Gln859 è un'Ala in mTOR.

## Interazioni con amminoacidi carichi

PI3K: la Lys802, da' un legame idrogeno con l'ossigeno carbonilico del lattone dell'inibitore (O24).

mTOR: La Lys802 è conservata.

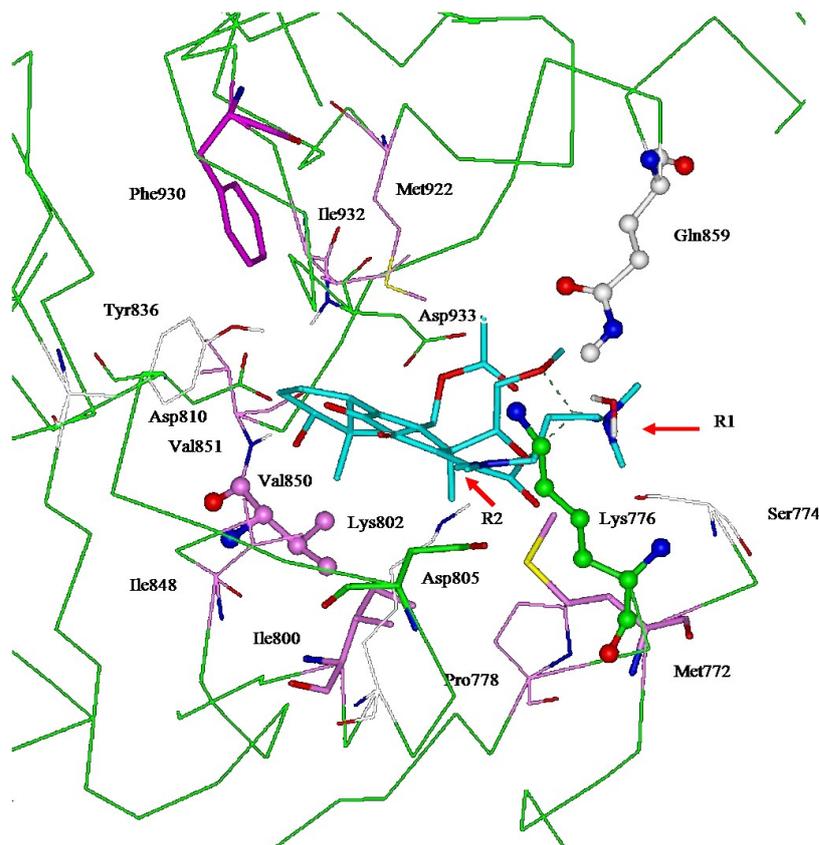


**Figura 30:** in figura sono mostrate le interazioni del wortmannin (arancione) nel sito ATPasico della PI3K. Gli amminoacidi coinvolti in interazioni idrofobiche sono colorati in rosa, quelli coinvolti in trasferimento di carica in magenta, quelli coinvolti in legami idrogeno in bianco. Gli amminoacidi riportati in lines sono conservati in mTOR, quelli riportati in stick sono sostituiti con amminoacidi omologhi in mTOR, quelli riportati in ball & stick non sono conservati in mTOR. In figura sono indicati i nomi degli atomi che sono coinvolti nelle interazioni con gli amminoacidi del sito catalitico. Gli ossigeni del ligando sono colorati in rosso.

## 2.4d.2 Sovrapposizione dei conformeri selezionati al wortmaninn in complesso con la PI3K

### Wort13

In tutti i conformeri selezionati tramite l'analisi delle distanze farmacoforiche l'azoto enamminico si trova sempre dalla parte opposta all'ossigeno ossidrilico e, proietta di conseguenza, i sostituenti R1 ed R2 (Figura 31).

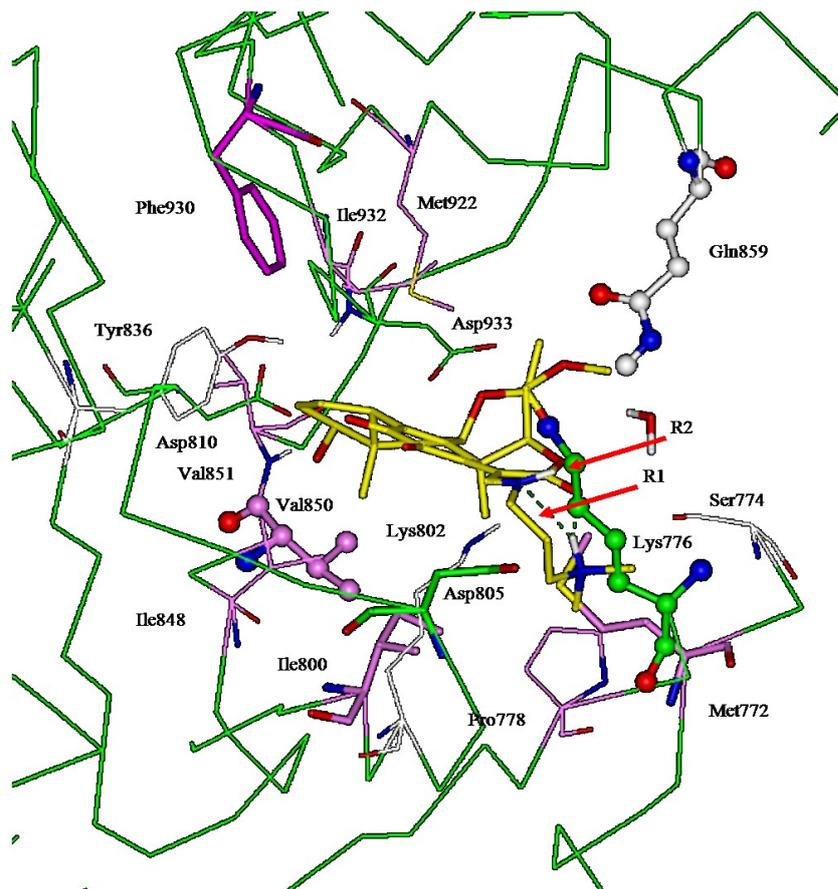


**Figura 31:** in figura sono mostrate le interazioni del derivato Wort13 (celeste) nel sito ATPasico della PI3K. Gli amminoacidi coinvolti in interazioni idrofobiche sono colorati in rosa, quelli coinvolti in trasferimento di carica in magenta, quelli coinvolti in legami idrogeno in bianco. Gli amminoacidi riportati in lines sono conservati in mTOR, quelli riportati in stick sono sostituiti con amminoacidi omologhi in mTOR, quelli riportati in ball & stick non sono conservati in Mtor. La differente posizione dei sostituenti R1 ed R2 è evidenziata con le frecce rosse. Gli eteroatomi del ligando sono colorati sulla base dei tipi atomici (N: blu e O: rosso).

Ipotesi di legame: La catena basica su R<sub>1</sub> spiazza la molecola d'acqua legata alla Ser774, che funge da ponte nell'interazione fra quest'ultima e gli ossigeni O1 e O21 dell'inibitore stesso. L'azoto protonato rimpiazza tutte le interazioni instaurate dalla molecola d'acqua. Bilancio: perdita di due legami idrogeno e guadagno di due legami idrogeno assistiti da carica positiva. Attività biologica paragonabile a quella del wortmannin (Figura 29). Il conformero selezionato rappresenta il minimo globale di energia.

## Wort14

*Ipotesi di legame*: la catena laterale basica è bloccata da un network di legami idrogeno che comprende l'azoto protonato, l'NH enamminico ed il carbonile del lattone, spiazzando sia la Lys802 che la Lys776 ed instaurando un conflitto sterico con lo scheletro peptidico della Lys776 e della Pro778. Composto inattivo (Figura 26). Il conformero selezionato rappresenta il minimo globale di energia (Figura 32).

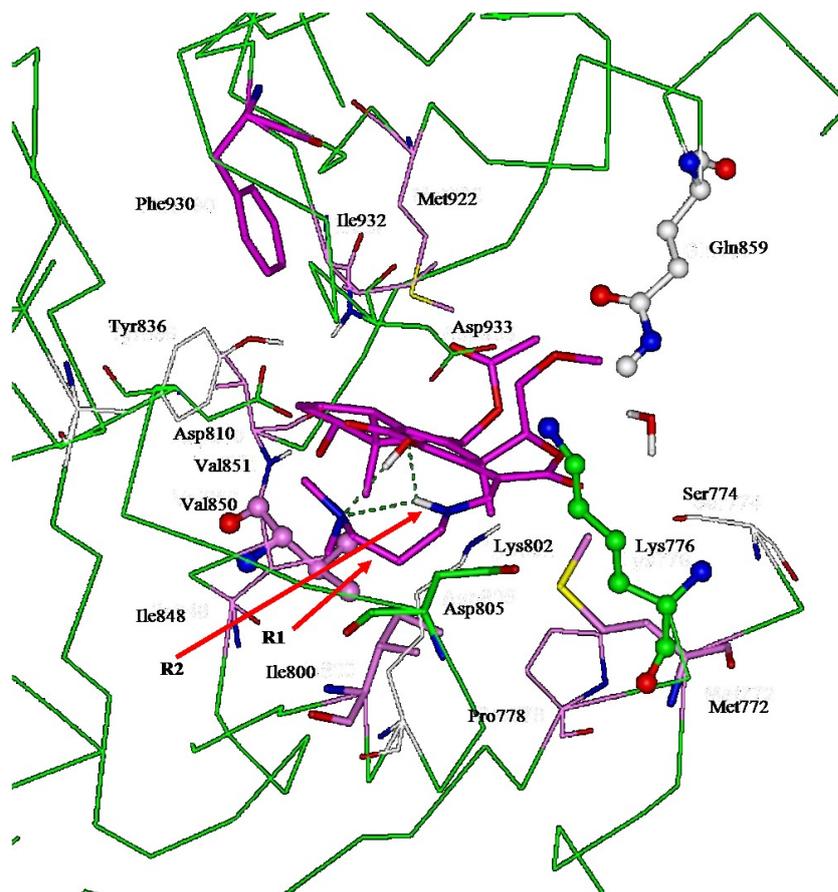


**Figura 32:** in figura sono mostrate le interazioni del derivato Wort14 nel sito ATPasico della PI3K. Gli amminoacidi coinvolti in interazioni idrofobiche sono colorati in rosa, quelli coinvolti in trasferimento di carica in magenta, quelli coinvolti in legami idrogeno in bianco. Gli amminoacidi riportati in lines sono conservati in mTOR, quelli riportati in sticks sono sostituiti con amminoacidi omologhi in mTOR, quelli riportati in ball & stick non sono conservati in mTOR. La differente posizione dei sostituenti R1 ed R2 è evidenziata con le frecce rosse. Gli atomi del ligando sono colorati sulla base dei tipi atomici (H: bianco, N: blu e O: rosso).

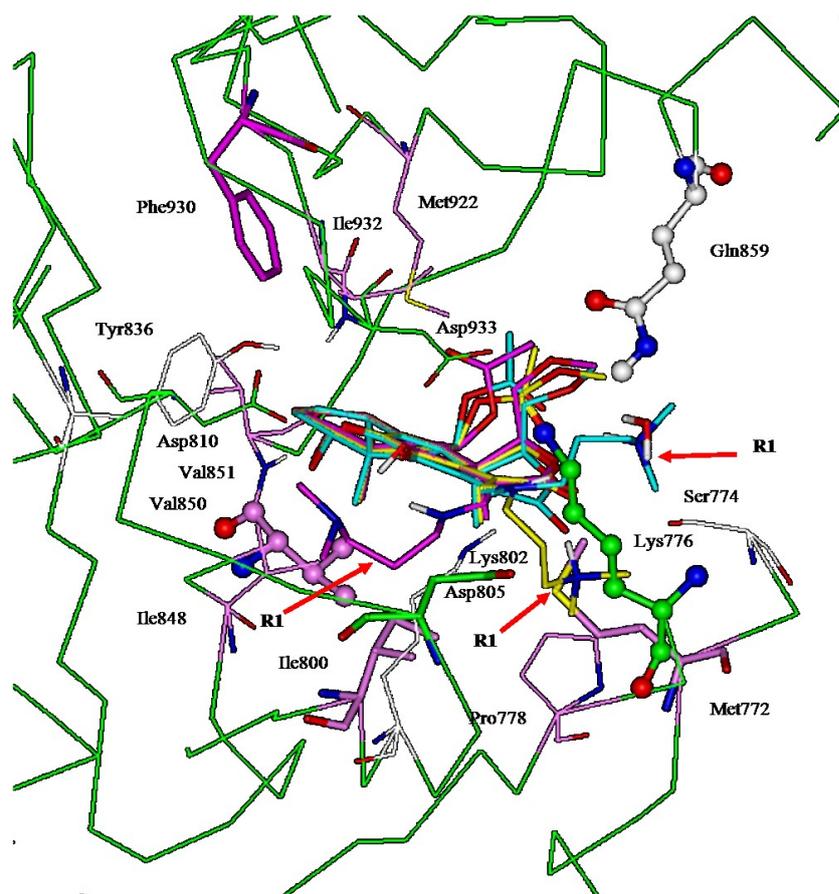
## Wort 14B

Nella struttura zwitterionica del composto 14 si nota nei calcoli quanto-meccanici PM6 lo slittamento dell'idrogeno dall'azoto basico (ammina terziaria su R2) all'ossigeno anionico con la formazione di un network di legami idrogeno fra l'NH enamminico, la catena basica e l'ossidrilica. Ciò porta inevitabilmente ad un differente posizionamento della catena laterale basica.

*Ipotesi di legame*: la catena basica a causa delle interazioni sopra descritte si trova in una posizione tale da spiazzare la Lys802 ed instaurare un conflitto sterico con la Ile848 . La Lys802 costituisce l'interazione più forte che il wortmannin instaurava con la proteina. Composto inattivo (Figura 27). Il conformero selezionato rappresenta il minimo globale di energia (Figura 33).



**Figura 33:** in figura sono mostrate le interazioni del derivato Wort 14B nel sito ATPasico della PI3K. Gli amminoacidi coinvolti in interazioni idrofobiche sono colorati in rosa, quelli coinvolti in trasferimento di carica in magenta, quelli coinvolti in legami idrogeno in bianco. Gli amminoacidi riportati in lines sono conservati in mTOR, quelli riportati in sticks sono sostituiti con amminoacidi omologhi in mTOR, quelli riportati in ball & stick non sono conservati in mTOR. La differente posizione dei sostituenti R1 ed R2 è evidenziata con le frecce rosse. Gli atomi del ligando sono colorati sulla base dei tipi atomici ( H: bianco, N: blu e O: rosso).

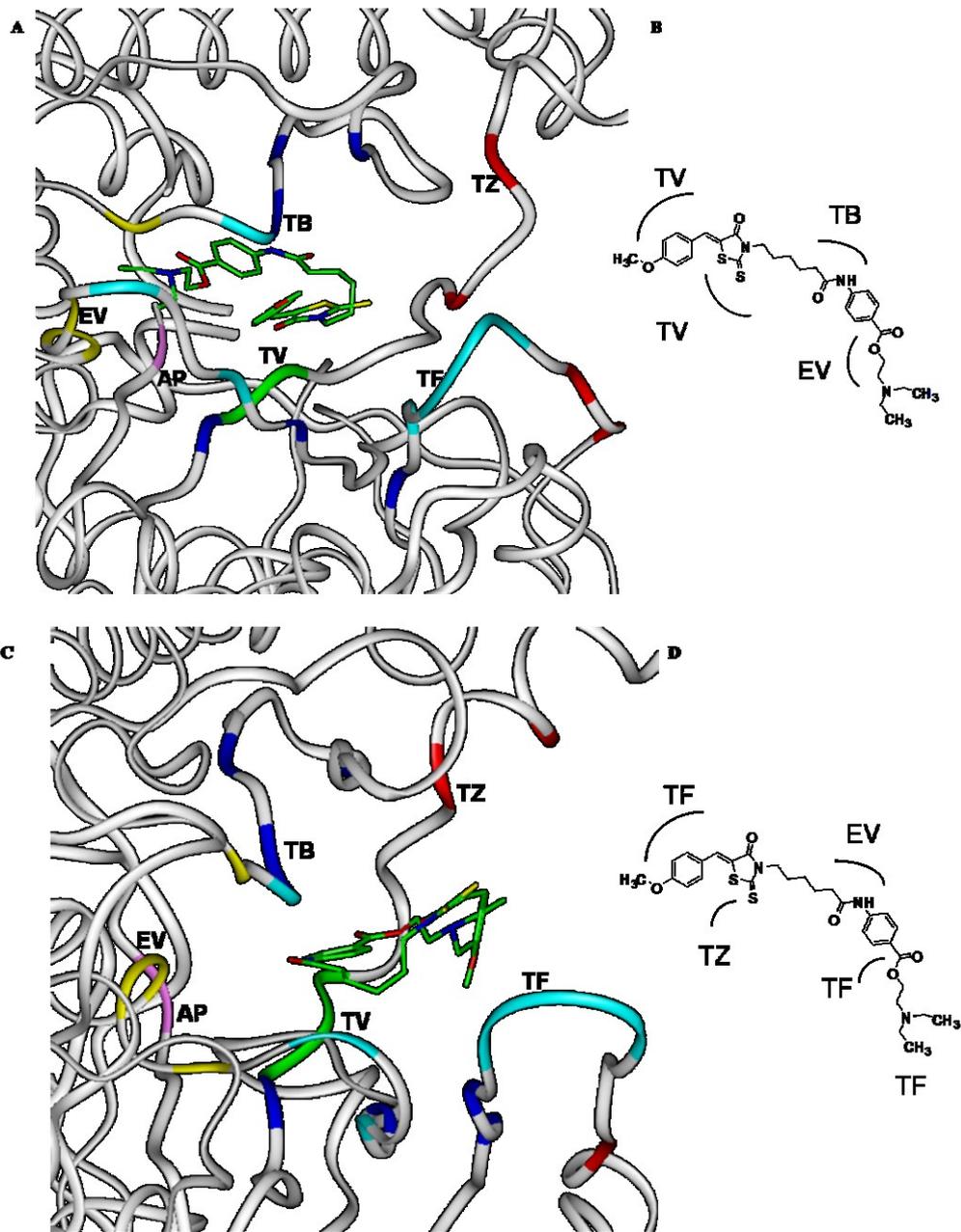


**Figura 34:** in figura è mostrata la sovrapposizione di Wort13 (ciano), Wort14 (giallo), Wort14B (magenta). La differente posizione della catena basica R1 è evidenziata con le frecce rosse. Gli amminoacidi coinvolti in interazioni idrofobiche sono colorati in rosa, quelli coinvolti in trasferimento di carica in magenta, quelli coinvolti in legami idrogeno in bianco. Gli amminoacidi riportati in lines sono conservati in mTOR, quelli riportati in sticks sono sostituiti con amminoacidi omologhi in mTOR, quelli riportati in ball & stick non sono conservati in mTOR. . Gli atomi dei ligandi sono colorati sulla base dei tipi atomici (H: bianco, N: blu e O: rosso).

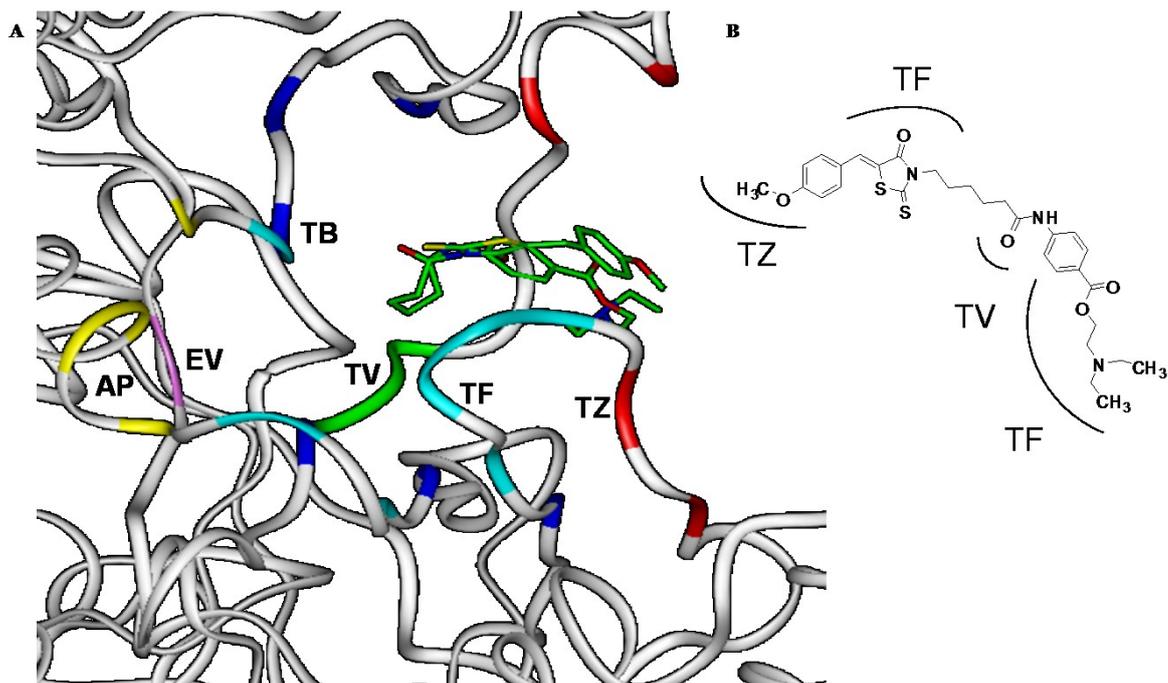
I risultati ottenuti dalle nostre analisi hanno permesso, quindi, di ipotizzare il farmacoforo per l'interazione con la PI3K del nostro *lead* ST0358AA1. In particolare, tramite l'analisi effettuata sui derivati del wortmannin, sono state individuate posizioni "attive" ed "inattive" dell'azoto amminico protonato a pH fisiologico, consentendo di definire il possibile posizionamento relativo delle porzioni farmacoforiche X, Y ed N35 di ST0358AA1 (Figura 24) all'interno del sito ATPasico della PI3K.

## ***2.4e Analisi degli Studi di Docking sul lead***

Per il calcolo di *docking*, è stata selezionata la conformazione bioattiva del composto di riferimento ottenuta dai calcoli precedentemente effettuati e in accordo con le distanze farmacoforiche degli inibitori noti. Tutti i complessi ottenuti dalla procedura di *docking* sono stati classificati per il valore della loro energia conformazionale e per quello dell'energia di interazione proteina-ligando. L'energia d'interazione di ciascun complesso è stata valutata calcolando: i) l'energia totale d'interazione tra il ligando e la PI3K, utilizzando il comando *Evaluate* nel modulo *Docking* di *Insight 2005* (energie di vdW ed elettrostatiche; no CUT\_OFF); ii) l'energia d'interazione (ovvero di non legame) tra il ligando e la PI3K utilizzando il modulo *Discover\_3* di *Insight 2005* (energie di vdW ed elettrostatiche; no CUT\_OFF). I complessi con il migliore compromesso tra l'energia conformazionale del ligando e l'energia di interazione con la proteina sono stati selezionati. I risultati sono riportati nelle *Figure 35 e 36* seguenti.



**Figura 35:** in figura sono riportate le possibili modalità di legame fra ST0358AA1 (riportato in stick) e la p110 $\alpha$  (mostrata in nastro). Nella figura A è mostrato il risultato Docking 1 (no p85); nella figura B è rappresentata l'immagine bidimensionale del ligando e la sua modalità di accomodamento nel sito catalitico. Nella Figura C viene mostrato il risultato del Docking 2 (no p85); nella figura D è rappresentata l'immagine bidimensionale del ligando e la sua modalità di accomodamento nel sito catalitico. Sono indicate: la tasca della valina (TV,) in verde; la tasca dove si posiziona lo zucchero dell'ATP (TZ,) in rosso; la tasca dove si posiziona la base dell'ATP (TB) in blu.; la tasca dove si posiziona il fosfato dell'ATP (TF) in ciano; una prima regione dove si puo' accomodare un eventuale volume molecolare aggiuntivo, rispetto a quello dell'ATP, (EV), in rosa; una seconda regione dove si puo' accomodare un eventuale volume molecolare aggiuntivo, rispetto a quello dell'ATP, detta tasca di affinità (AP) in giallo. Gli eteroatomi dei ligandi sono colorati sulla base dei tipi atomici (S: giallo, N: blu e O: rosso).

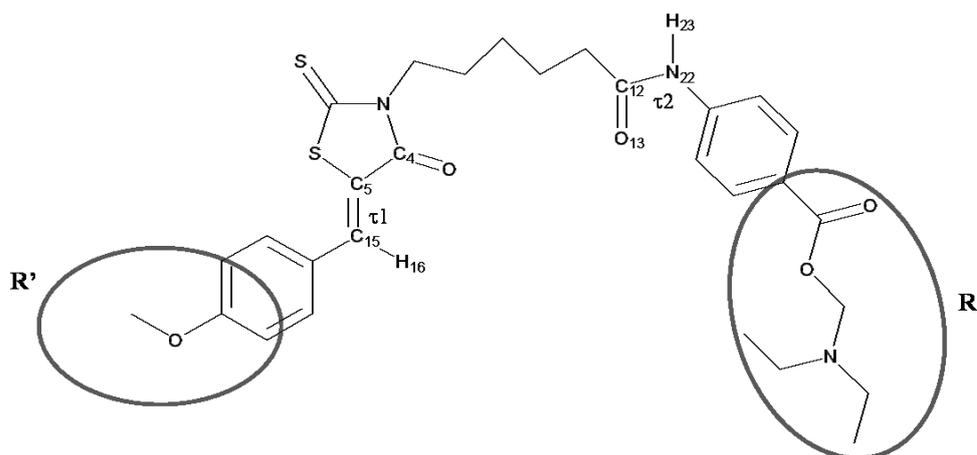


**Figura 36:** nella figura A è mostrato il risultato del Docking fra ST0358AA1 (riportato in stick) e la p110 $\alpha$  in complesso con la p85 (mostrata in nastro); nella figura B è rappresentata l'immagine bidimensionale del ligando e la sua modalità di accomodamento nel sito catalitico. Sono indicate: la tasca della valina (TV,) in verde; la tasca dove si posiziona lo zucchero dell'ATP (TZ,) in rosso; la tasca dove si posiziona la base dell'ATP (TB) in blu; la tasca dove si posiziona il fosfato dell'ATP (TF) in ciano; una prima regione dove si puo' accomodare un eventuale volume molecolare aggiuntivo, rispetto a quello dell'ATP, (EV), in rosa; una seconda regione dove si puo' accomodare un eventuale volume molecolare aggiuntivo, rispetto a quello dell'ATP, detta tasca di affinità (AP) in giallo. Gli eteroatomi del ligando sono colorati sulla base dei tipi atomici (S: giallo, N: blu e O: rosso).

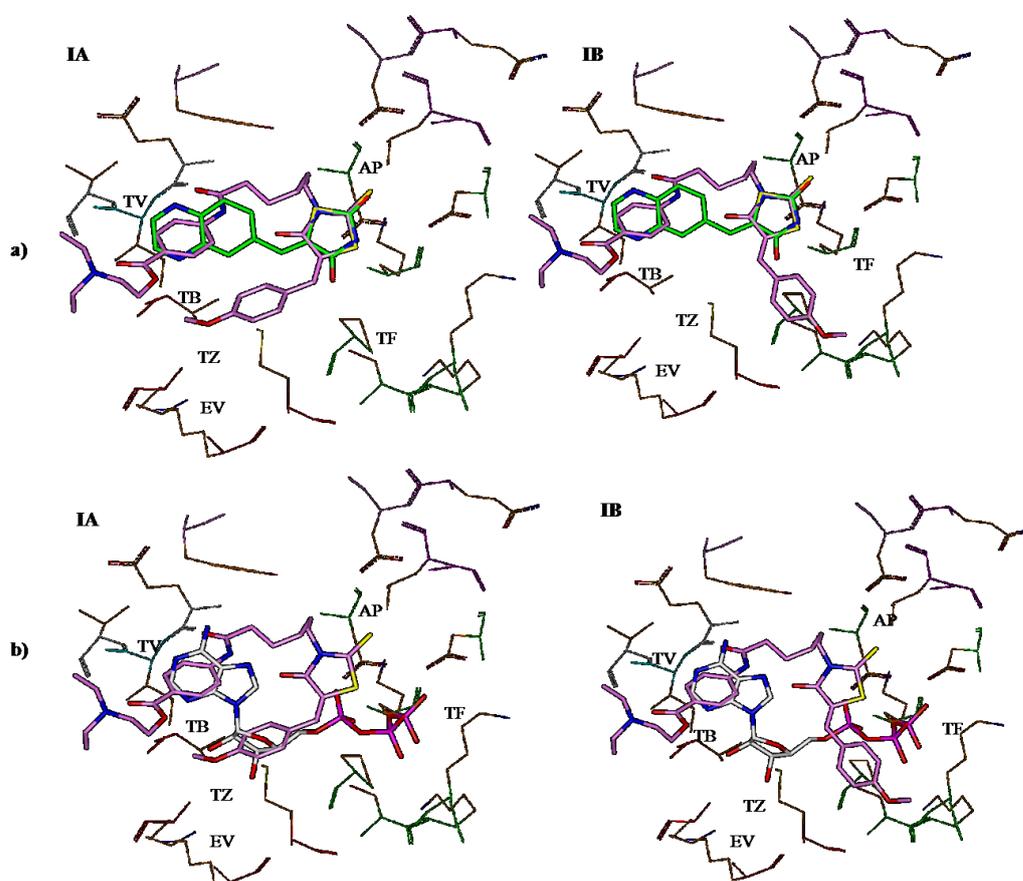
Come mostrato nelle immagini sono risultati possibili diverse modalità di legame con la proteina, data l'alta flessibilità della struttura del ligando e dell'enzima stesso, quest'ultima considerata in assenza (Figura 34) ed in presenza (Figura 35) della subunità regolatoria p85.

E' stato ipotizzato che ciascuna di queste porzioni molecolari del ligando debbano accomodarsi in uno specifico sotto-sito del sito catalitico di PI3K; inoltre, la conformazione bioattiva potrebbe essere stabilizzata da almeno due legami idrogeno intramolecolari: quello fra l'ossigeno carbonilico dell'anello 1,3-tiazolidin-4-one e l'NH del legame ammidico e quello fra il carbonile estereo e l'ammina terziaria (protonabile a pH fisiologico).

Al fine di valutare tutte le possibili ipotesi di interazione del *lead* con la PI3K, oltre a quelle ottenute dagli studi di *docking*, si è sovrapposto il composto nella sua conformazione bioattiva all'inibitore AS-605240 a livello delle porzioni strutturali simili: sono stati effettuati due diversi tipi di sovrapposizione (Figure 37 e 38) ipotizzando, in tal modo, due differenti modalità di legame nel sito catalitico a seconda della zona di accomodamento delle porzioni R ed R'.



### Sovrapposizione di tipo I



**Figura 37: Sovrapposizione di tipo I**

Nell'immagine **a)** è mostrata la sovrapposizione di I tipo del lead ST0358AA1 sull'inibitore AS-605240, realizzata a livello della porzione strutturale simile: anello 1,3-tiazolidin-4-one (ST0358AA1) – anello 1,3-tiazolidin-2,4-dione (AS-605240).

L'immagine **b)** mostra l'accomodamento che il I tipo di sovrapposizione ha determinato per le porzioni R ed R' dell'inibitore nelle diverse tasche enzimatiche di PI3K, individuate considerando le porzioni dell'ATP: la porzione R va ad accomodarsi nella tasca dove si posiziona la base dell'ATP (TB), mentre la porzione R' va ad accomodarsi nella tasca dove si posiziona lo zucchero dell'ATP (TZ,) se il torsionale  $\tau_1$  è cis (IA) oppure nella tasca dove si posiziona il fosfato dell'ATP (TF) se il torsionale  $\tau_1$  è trans (IB).

La denominazione delle sotto-aree del sito catalitico si riferisce a quella riportata in Figura 17 ed utilizzando gli acronimi della tabella 9. Gli eteroatomi dei ligandi sono colorati sulla base dei tipi atomici (S: giallo, N: blu e O: rosso).

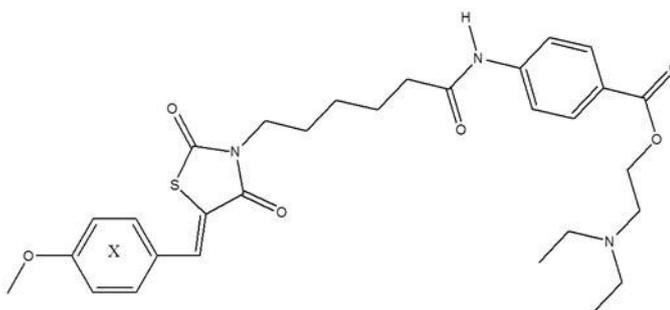


## 2.4f Ottimizzazione del lead: progettazione di nuovi inibitori potenti e specifici.

Sulla base dei risultati ottenuti nelle precedenti fasi della ricerca si è proceduto alla progettazione di modifiche mirate da effettuare sul composto di riferimento ST0358AA1, al fine di ampliare le relazioni struttura-attività del nuovo *lead* e di ottenere composti con un migliore profilo di inibizione enzimatica e una potenziata citotossicità su linee cellulari tumorali. Le modifiche molecolari sono state guidate da calcoli di *docking* dinamico nel sito dell'enzima e dagli studi di *pharmacophore fitting* del composto ST0358AA1

Sono stati variati, dunque, i seguenti punti della molecola:

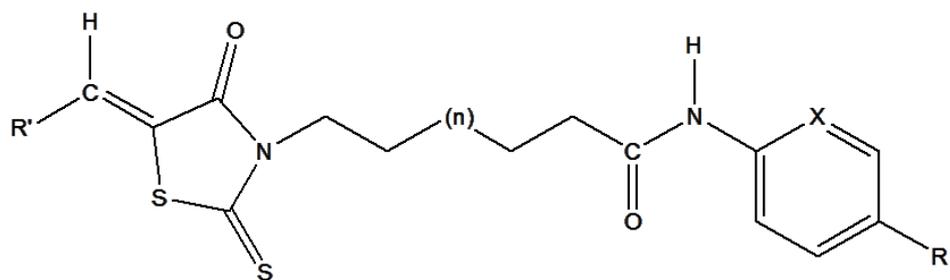
- i) i sostituenti sull'anello aromatico indicato con X in *Figura 39*
- ii) la lunghezza del linker metilenico
- iii) il gruppo estereo
- iv) la catena contenente l'azoto protonabile



**Figura 39:** rappresentazione bidimensionale del lead

Sulla base dei risultati ottenuti nelle precedenti fasi della ricerca si è proceduto, quindi, alla progettazione di nuove molecole riportate nella tabella seguente (*Tabella 17*).

**Tabella 17:** set di ligandi progettati



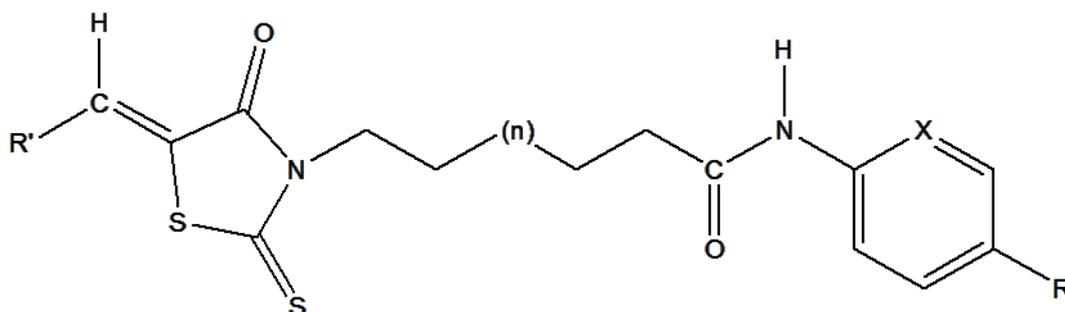
ST	R'	n	R	X
0360		0		CH
0386		1		CH
0387		1		CH
0359		1		CH
0415		1		CH
0388		1		CH
0362		1		CH
0390		1		CH
0389		1		CH
0391		1		CH
0392		1		CH
0393		1		CH
0364		1		CH

0416		1		CH
0363		1		CH
0354		1		CH
0355		1		CH
0356		1		CH
0357		1		CH
0361		1		CH
0394		1	-	N

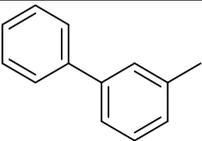
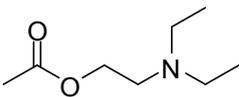
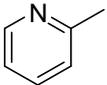
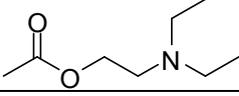
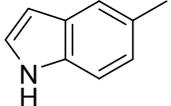
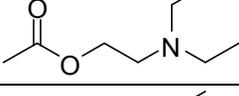
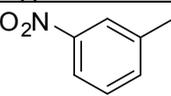
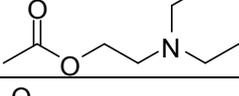
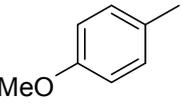
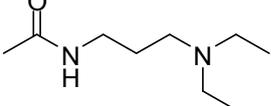
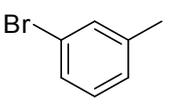
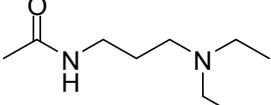
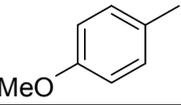
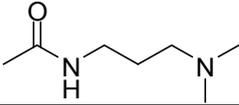
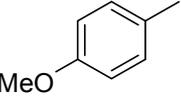
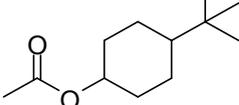
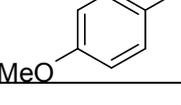
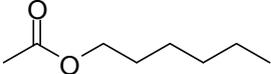
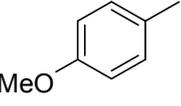
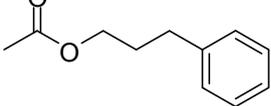
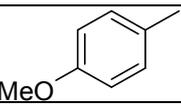
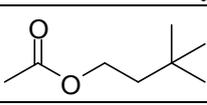
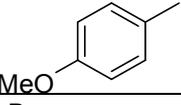
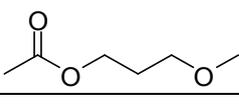
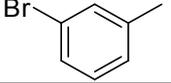
## 2.4g Studi Computazionali ed Analisi delle Relazioni Struttura-Attività (SARs) dei nuovi derivati progettati

Il profilo di selettività e potenza del *lead* e dei nuovi analoghi progettati è mostrato nella seguente tabella 18.

**Tabella 18:** Attività sulle diverse isoforme della PI3K e dati di citotossicità del *lead* e dei nuovi analoghi progettati sulle cellule tumorali renali A498 e di osteosarcoma U20S



ST	R'	n	R	X	PI3K IC50 (μM)				Citotossicità IC50 (μM)	
					α	β	δ	γ	A498 Renali	U20S Osteosarcoma
0358 <i>lead</i>		1		CH	>10	7.1	>10	>10	NT	NT
0360		0		CH	>10	1.10	>10	>10	NT	NT
0386		1		CH	0.80	>10	>10	>10	5.7	3.0
0387		1		CH	>10	>10	>10	>10	NT	NT
0359		1		CH	0.31	>10	>10	>10	9.6	6.0
0415		1		CH	>10	3	>10	>10	NT	NT
0388		1		CH	0.001	>10	>10	>10	3.6	1.9
0362		1		CH	0.09	>10	>10	>10	<10	<10
0390		1		CH	0.16	>10	>10	>10	1.8	2.5

0389		1		CH	>10	>10	>10	>10	NT	NT
0391		1		CH	>10	>10	>10	>10	NT	NT
0392		1		CH	>10	>10	>10	>10	NT	NT
0393		1		CH	>10	6.2	>10	>10	NT	NT
0364		1		CH	0.13	>10	>10	>10	13.1	9
0416		1		CH	>10	>10	>10	>10	NT	NT
0363		1		CH	>10	>10	>10	>10	NT	NT
0354		1		CH	>10	>10	>10	>10	20	15.6
0355		1		CH	>10	>10	>10	>10	>20	>20
0356		1		CH	>10	>10	>10	>10	>20	>20
0357		1		CH	5.6	>10	>10	>10	>20	>20
0361		1		CH	>10	>10	>10	>10	>20	>20
0394		1	-	N	0.001	>10	>10	>10	7.1	7.5

NT = non testato.

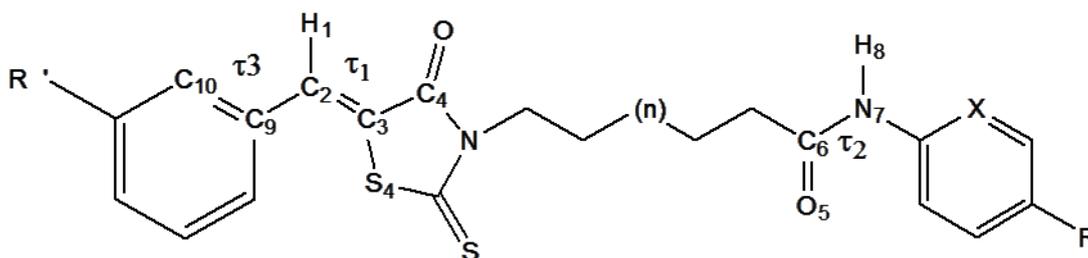
Per l'analisi delle SARs degli inibitori progettati è stata effettuata un'analisi conformazionale dei ligandi nella quale è stata fatta una valutazione statistica dei diversi parametri considerati, al fine di effettuare un'analisi approfondita delle relazioni struttura-attività.

In particolare, sono stati considerati:

- conformazione dell'ammide ( $\tau_2$  O5C6N7H8);
- posizionamento del sostituente del gruppo R' ( $\tau_3$  C10C9C2H1) mediante la misura degli angoli torsionali (parametro non presente quando R' è in posizione para);
- legami idrogeno intramolecolari;
- la modalità di legame (in tabella denominata modalità di legame I e modalità di legame II) che ciascun minimo conformazionale ha riprodotto negli studi di *pharmacophore fitting* effettuati sovrapponendo i minimi risultanti dalla ricerca conformazionale all'inibitore AS-604850 o all'inibitore AS-605240 come descritto nel paragrafo 2.3i.

I risultati dell'analisi condotta sono riportati nella seguente Tabella 19

**Tabella 19:** risultati della ricerca conformazionale



Nome molecola	Conformazione Ammide	Modalità di legame I (%)	Modalità di legame II (%)	Posizionamento R' ( $\tau_3$ )	Legami H intramolecolari
ST0359		47	53	A <sup>a</sup> = 50 % B <sup>b</sup> = 50%	19,4 % NH ammide- O tiazolo 100 % N protonato- estere (CO)
	TRANS 97%	47	50		
	CIS 3 %	-	3		
ST0388		56	44	-	6,8 % NH ammide- O tiazolo 100 % N protonato- estere (CO)
	TRANS 100%	56	44		
ST0394		49,5	50,5	A <sup>a</sup> = 49,5% B <sup>b</sup> = 50,5%	26,5 % NH ammide- O tiazolo 85 % N piridina- H ammide
	TRANS 83%	41,5	41,5		
	CIS 17%	8	9		

<b>ST0415</b>		52	48	A <sup>a</sup> = 59% B <sup>b</sup> = 41%	13% NH ammide- O tiazolo 100% N protonato- CO ammido
	TRANS 100%	52	48		
<b>ST0354</b>		52	48	-	36,4 % NH ammide- O tiazolo
	TRANS 83%	44	39		
	CIS 17%	8	9		
<b>ST0355</b>		47	53	-	34 % NH ammide- O tiazolo
	TRANS 92%	43	49		
	CIS 8%	4	4		
<b>ST0356</b>		47,5	52,5	-	34,3 % NH ammide- O tiazolo
	TRANS 88%	44,5	43,5		
	CIS 12,%	3	9		
<b>ST0357</b>		48	52	-	27,7 % NH ammide- O tiazolo
	TRANS 80%	39	41		
	CIS 20%	9	11		
<b>ST0358</b>		46	54	-	16 % NH ammido- O tiazolo 94 % N protonato- estere (CO)
	TRANS 97%	46	51		
	CIS 3%	-	3		
<b>ST0361</b>		54	46	-	37,5 % NH ammide- O tiazolo
	TRANS 97%	52	45		
	CIS 3%	2	1		
<b>ST0362</b>		50	50	-	31,5 % NH ammide- O tiazolo 92,6 % N protonato- estere (CO)
	TRANS 100%	50	50		

<b>ST0360</b>		52	48	-	39,6 % NH ammide- O tiazolo  98 % N protonato- estere (CO)
	TRANS 96%	48	48		
	CIS 4%	4 (D3)	-		
<b>ST0363</b>		30,4	69,6	-	17,4 % NH ammide- O tiazolo  100 % N protonato- ammide R (CO)
	TRANS 100%	30,4	69,6		
	CIS 5%	-	5		
<b>ST0364</b>		45	55	-	30 % NH ammide- O tiazolo  100 % N protonato- ammide R (CO)
	TRANS 100 %	45	55		
	CIS 0%	-	-		
<b>ST0386</b>		75,5	24,5	A <sup>a</sup> = 49% B <sup>b</sup> = 51%	34,7 % NH ammide- O tiazolo  98 % N protonato- estere (CO)
	TRANS 95,9%	71,4	24,5		
	CIS 4,1%	4,1	-		
<b>ST0387</b>		43,15	56,85	A <sup>a</sup> = 41% B <sup>b</sup> = 59%	21,6% NH ammide- O tiazolo  100% N protonato- estere (CO)
	TRANS 80,4%	37,25	43,15		
	CIS 19,6%	5,9	13,7		
<b>ST0416</b>		39	61	A <sup>a</sup> = 39 % B <sup>b</sup> = 61%	10 % NH ammide- O tiazolo  100% N protonato- estere (CO)
	TRANS 100%	39	61		
<b>ST0390</b>		38,7	61,3	A <sup>a</sup> = 48,4% B <sup>b</sup> = 51,6 %	29 % NH ammide- O tiazolo  90% N protonato- estere (CO)
	TRANS 100%	38,7	61,3		

<b>ST0389</b>		40	60	A <sup>a</sup> = 60% B <sup>b</sup> = 40 %	25 % NH ammide- O tiazolo
	TRANS 100%	40	60		100% N protonato- estere ( CO)
<b>ST0391</b>		52	48	-	12 % NH ammide- O tiazolo
	TRANS 96%	48	44		96 % N protonato- estere ( CO)
	CIS 4 %	4	4		
<b>ST0393</b>		45	55	A <sup>a</sup> = 70% B <sup>b</sup> = 30 %	50 % NH ammide- O tiazolo
	TRANS 95 %	45	50		100 % N protonato- estere ( CO)
	CIS 5%	-	5		

<sup>a</sup>  $\tau_3 = -40, -130$ ; <sup>b</sup>  $\tau_3 = +40, +130$ .

L'analisi conformazionale ha messo in evidenza che:

- l'ammide tende ad essere principalmente in *trans*;
- i sostituenti in meta ed orto del gruppo R' si orientano con uguale probabilità in entrambe le posizioni possibili;
- è sempre presente il legame ad idrogeno tra l'estere o ammidine di R con l'N protonato.

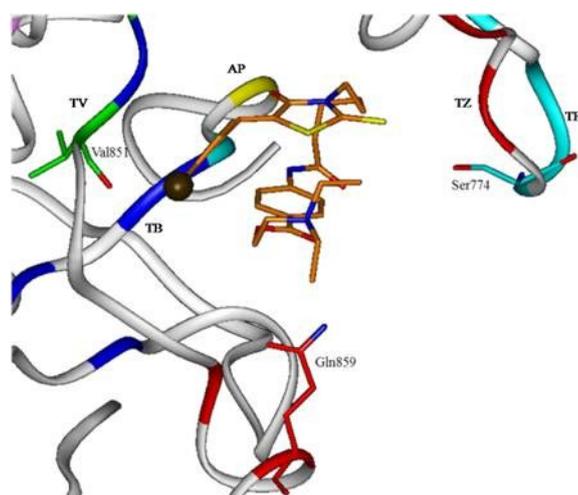
Sulla base dei risultati ottenuti, è stato possibile razionalizzare le relazioni struttura-attività:

### ***1) Modifiche su R'.***

La sostituzione del fenile in R' con un indolo (ST0392AA1, *Tabella 18*) o con una 2-piridina (ST0391AA1, *Tabella 18*) fa cadere l'attività. Sia la 2-piridina che l'anello indolico sono mimetici dell'ATP, in particolare, in grado di riprodurre le interazioni della base azotata con l'enzima. Di conseguenza, i risultati ottenuti indicano che tale porzione della molecola non è coinvolta in interazioni con la tasca del sito catalitico preposta al legame della base dell'ATP (TB, *Figura 17*).

A parità di R e di lunghezza del linker (n = 1), la sostituzione in posizione para (ST0358AA1, ST0388AA1, ST0362AA1, *Tabella 18*) evidenzia il seguente trend di attività: Br>OBn>>>OMe. La maggior attività è data dal Br in quanto meno ingombrante rispetto

all'O-CH<sub>3</sub> e all'OBn, suggerendo che riesce meglio ad accomodarsi nella tasca della base (Figura 40), e ad orientarsi verso la Val851 (aminoacido importante nelle interazioni della TB), instaurando interazioni ottimali per l'accomodamento di tale porzione. La sostituzione in posizione meta indica che sia il Br (ST0359AA1, Tabella 18) che l'OBn (ST0390AA1, Tabella 18) portano a composti attivi e di potenza comparabile; la sostituzione con un gruppo OMe (ST0386AA1, Tabella 18) modula negativamente la potenza, pur mentendo una inibizione a concentrazioni sub-micromolari. Al contrario, la presenza di un sostituito fenilico (ST0389AA1, Tabella 18) o di un gruppo nitro (ST0393AA1, Tabella 18) causa la perdita di attività inibitoria. Infine, la sostituzione con un OMe in posizione orto porta anch'essa ad un composto inattivo (ST0387AA1, Tabella 18).

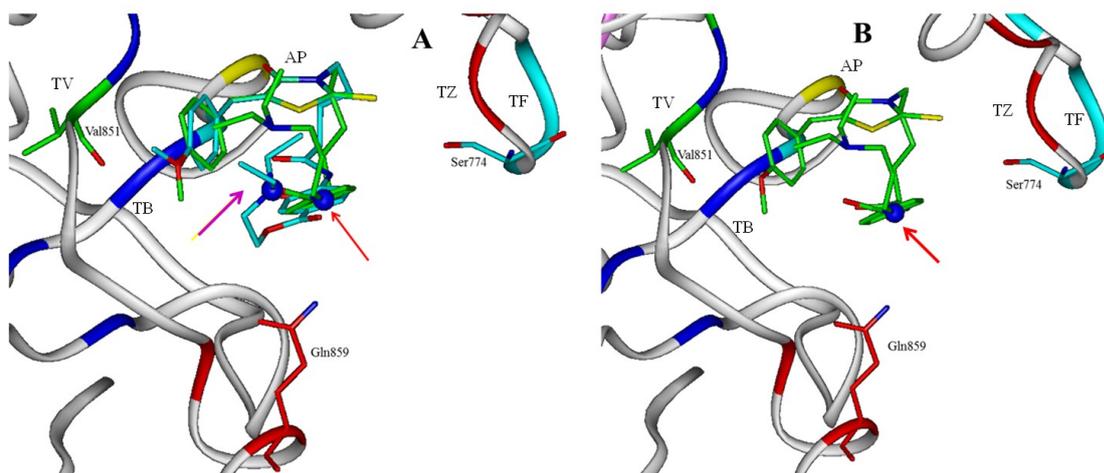


**Figura 40:** in figura è mostrato l'accomodamento del Br nella tasca dove si posiziona la base dell'ATP (TB) rivolto verso la Val851. Sono indicate: la tasca della valina (TV,) in verde; la tasca dove si posiziona lo zucchero dell'ATP (TZ,) in rosso; la tasca dove si posiziona la base dell'ATP (TB) in blu,; la tasca dove si posiziona il fosfato dell'ATP (TF) in ciano; una prima regione dove si puo' accomodare un eventuale volume molecolare aggiuntivo, rispetto a quello dell'ATP, (EV), in rosa; una seconda regione dove si puo' accomodare un eventuale volume molecolare aggiuntivo, rispetto a quello dell'ATP, detta tasca di affinità (AP) in giallo. . Gli eteroatomi del ligando sono colorati sulla base dei tipi atomici (S: giallo, N: blu ; O: rosso e C: arancione).

## II) Modifiche su R.

Per quanto riguarda la presenza di un gruppo ammidico (donatore/accettore di legami H) o estereo (donatore di legami H), quest'ultimo risulta dare una maggiore attività (ST0359AA1 vs ST0415AA1, Tabella 18). Come evidenziato dai risultati ottenuti dai calcoli quantomeccanici, la parziale carica negativa dell'estere è delocalizzata tra i due ossigeni, mentre nell'ammide è l'ossigeno carbonilico ad avere la carica negativa più elevata. Ciò fa sì che in presenza dell'estere la molecola ha più libertà conformazionale rispetto a quando c'è

l'ammide, infatti, l'interazione tra azoto protonato e gruppo carbonilico dell'ammide è più forte che con quello dell'estere. In particolare, la presenza dell'estere consente all'azoto protonato di posizionarsi nella tasca delimitata dalla Ser774 e dalla Gln859; invece, la presenza dell'ammide tende a sfavorire il posizionamento dell'azoto protonato in questa tasca. E' importante sottolineare che, in perfetto accordo con l'attività inibitoria misurata, quando è presente l'ammide ma il linker è più lungo di un metilene (ST0364AA1 vs ST0360AA1, *Tabella 18*) ciò consente all'azoto ammidico di posizionarsi comunque tra la Ser774 e la Gln859 (*Figura 41*). Queste rappresentano interazioni fondamentali per il legame con l'enzima, individuate dalle analisi riportate nei precedenti 2.4d che mostrano come l'azoto protonato del derivato Wort13 del wortmannin tenda a posizionarsi al posto di una molecola d'acqua che fa da ponte per l'interazione con la Ser774 e gli ossigeni dell'inibitore stesso, nella tasca delimitata dalla Ser774 e la Gln859 (*Figura 31*). L'azoto protonato, quindi, rimpiazza tutte le interazioni instaurate dalla molecola d'acqua. Un caso particolare è rappresentato dal composto ST0394AA1 (*Tabella 18*), nel quale il sostituente è assente ed il fenile in "coda" è rimpiazzato con una 2-piridina (X = N, *Tabella 5*). Anche in questo caso la 2-piridina è stata utilizzata in qualità di analogo della base azotata dell'ATP, producendo uno dei composti più attivi della serie. La presenza dell'ammide (a causa del legame idrogeno intra-molecolare) tende invece a sfavorire l'interazione dell'azoto protonato con questa tasca. E' importante sottolineare che, se viene allungata la catena alchilica tra l'ammide e l'azoto protonato aggiungendo un ulteriore CH<sub>2</sub>, l'azoto protonato può posizionarsi correttamente per interagire con la suddetta tasca, infatti, il composto ST0364AA1 (*Tabella 18*), presenta un'ottima attività inibitoria. A conferma di ciò è stato osservato che tutte le molecole che non presentano su R un azoto protonabile non sono dotate di attività.



**Figura 41:** in figura è evidenziato come l'azoto protonato (freccia magenta) del linker della ST0360AA1 (molecola in ciano) si posizioni nella stessa regione dell'azoto amidico (freccia rossa) del gruppo R della ST0364AA1 (molecola verde). La figura 41B mostra come la presenza di un  $\text{CH}_2$  in più nel gruppo R tra la funzione amminica e la funzione esterea favorisce il posizionamento dell'azoto amminico (freccia rossa) di R nella tasca d'entrata tra la Ser774 e la Gln859. Sono indicate: la tasca della valina (TV,) in verde; la tasca dove si posiziona lo zucchero dell'ATP (TZ,) in rosso; la tasca dove si posiziona la base dell'ATP (TB) in blu; la tasca dove si posiziona il fosfato dell'ATP (TF) in ciano; una prima regione dove si può accomodare un eventuale volume molecolare aggiuntivo, rispetto a quello dell'ATP, (EV), in rosa; una seconda regione dove si può accomodare un eventuale volume molecolare aggiuntivo, rispetto a quello dell'ATP, detta tasca di affinità (AP) in giallo. . Gli eteroatomi del ligando sono colorati sulla base dei tipi atomici (S: giallo, N: blu e O: rosso).

### III) Linker

L'effetto dell'accorciamento del linker, ottenuto modificando il valore di n, non può essere verificato nella presente serie di composti in quanto ST0358AA1 ed ST0360AA1 (Tabella 18), sulla base della analisi SAR su riportata, sono entrambi inattivi a causa della sostituzione su R'.

### IV) Conformazione bioattiva

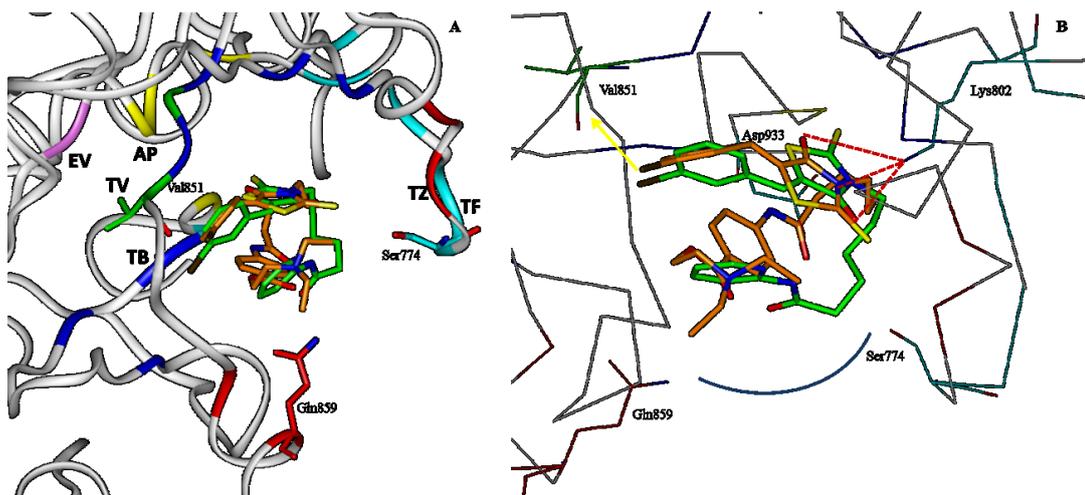
La conformazione che, secondo gli studi di *docking* e di *pharmacophore fitting*, determina il miglior accomodamento dei nuovi derivati nel sito catalitico è rappresentata dalla conformazione di minimo energetico globale per le molecole risultate attive, mentre è una conformazione a più alta energia per quelle risultate non attive.

L'analisi condotta ha mostrato che il set di ligandi progettati riesce a riprodurre in maniera ottimale la modalità di legame I eccetto per la molecola ST0394AA1 (Tabella 18), in cui non è presente la catena R ed il fenile che lega R è in questo caso una piridina, che tende a preferire la modalità di legame II.

Sulla base di queste considerazioni le molecole che maggiormente rispettano i requisiti farmacoforici sopra elencati sono la ST0388AA1, ST0394AA1 (Tabella 18).

In particolare, la ST0388AA1 (Tabella 18) si accomoda con la modalità di legame I nel sito catalitico posizionando il fenile con il Br in para nella tasca dove si posiziona la base

dell'ATP (TB) direzionandosi verso la Val851, il fenile con il *linker* che presenta l'azoto protonato nella tasca dove si posiziona lo zucchero dell'ATP (TZ); nel caso della ST0394AA1 (Tabella 18) si accomoda in modo simile a ST0388AA1 (Tabella 18) ma utilizzando la modalità di legame II; ossia il fenile con il Br in meta si accomoda nella tasca dove si posiziona la base dell'ATP (TB) rivolto verso la Val851, la piridina nella tasca dove si posiziona lo zucchero dell'ATP (TZ); rivolgendo l'azoto verso la Ser774 e Gln859 (Figura 42).



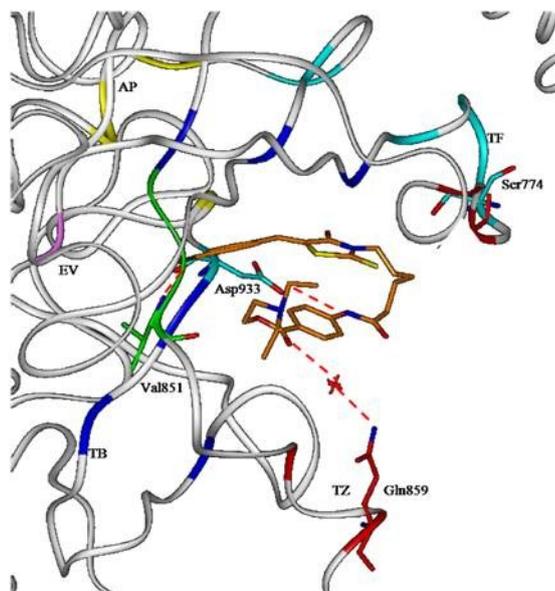
**Figura 42:** in figura è mostrato l'accomodamento del ligando ST0388AA1 in arancio (modalità di legame I) e ST0394AA1 in verde (modalità di legame II). Pur adottando due differenti modalità di legame, le molecole posizionano nella stesse direzione le porzioni fondamentali per le interazioni con gli aminoacidi del sito catalitico. La figura B mostra lo stesso accomodamento ma in modo frontale evidenziando le interazioni con gli aminoacidi del sito catalitico da parte di ciascun anello 1,3-tiazolidin-4-one. Sono indicate: la tasca della valina (TV,) in verde; la tasca dove si posiziona lo zucchero dell'ATP (TZ,) in rosso; la tasca dove si posiziona la base dell'ATP (TB) in blu; la tasca dove si posiziona il fosfato dell'ATP (TF) in ciano; una prima regione dove si può accomodare un eventuale volume molecolare aggiuntivo, rispetto a quello dell'ATP, (EV), in rosa; una seconda regione dove si può accomodare un eventuale volume molecolare aggiuntivo, rispetto a quello dell'ATP, detta tasca di affinità (AP) in giallo. Gli eteroatomi del ligando sono colorati sulla base dei tipi atomici (S: giallo, N: blu e O: rosso).

Le due molecole, quindi, possono accomodarsi nel sito catalitico mediante due diverse modalità di legame, ma essere ugualmente attive, in quanto riescono a posizionare nella stessa direzione i punti farmacoforici importanti per le interazioni con gli aminoacidi del sito.

### 2.4g.1 Analisi degli Studi di Docking sugli inibitori risultati più attivi

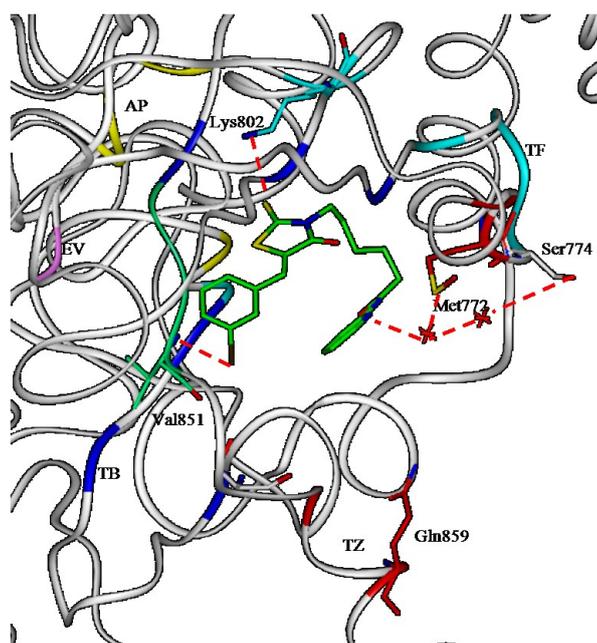
Gli studi di docking sono stati effettuati sui ligandi ST0388AA1 ed ST0394AA1 in quanto presentano la maggiore selettività (attività nano molare) per l'isoforma  $\alpha$  della PI3K. Il punto di partenza di ciascun inibitore è stato effettuato tenendo in considerazione gli studi *pharmacophore fitting* i cui risultati sono riportati nella tabella 19.

I risultati ottenuti hanno confermato l'ipotesi di legame precedentemente descritta: in particolare, la ST0388AA1 si accomoda nel sito catalitico adottando la modalità di legame I: posizionando il fenile con il Br in para nella tasca dove si posiziona la base dell'ATP (TB) e dà un legame idrogeno con l'NH dello scheletro peptidico della Val851, l'NH dell'ammide interagisce con il CO della catena laterale dell'Asp933, il CO dell'estere mediante una molecola d'acqua interagisce con l'NH della catena laterale della Gln859 e l'N protonato si inserisce nella tasca tra Ser774 e Gln859 (Figura 43)



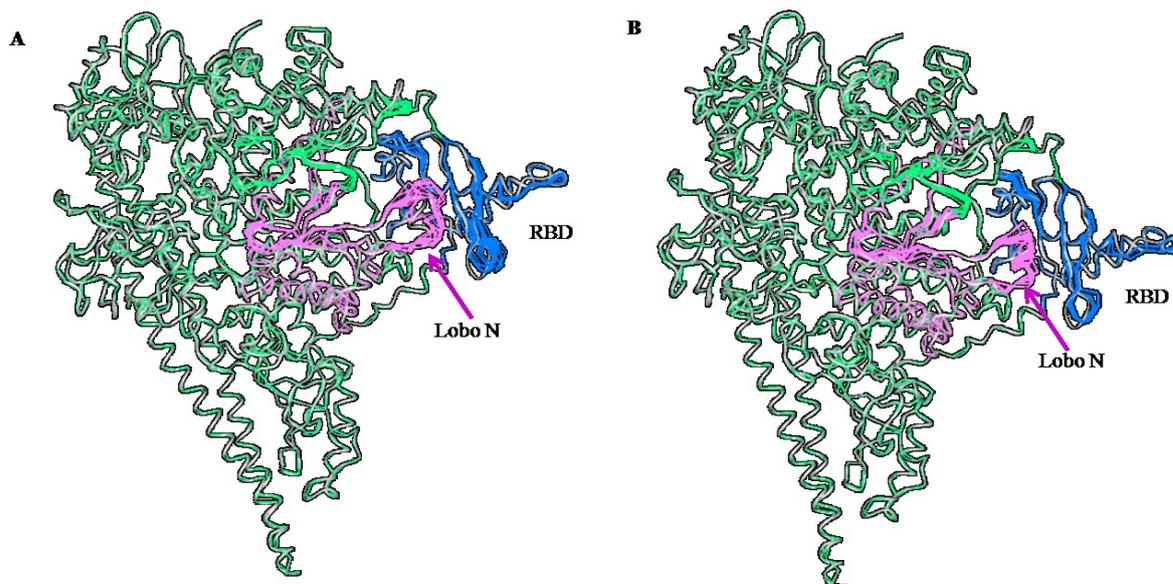
**Figura 43:** in figura è mostrato il risultato del docking dell'inibitore ST0388AA1 nel sito catalitico della PI3K $\alpha$ . Sono indicate: la tasca della valina (TV,) in verde; la tasca dove si posiziona lo zucchero dell'ATP (TZ,) in rosso; la tasca dove si posiziona la base dell'ATP (TB) in blu; la tasca dove si posiziona il fosfato dell'ATP (TF) in ciano; una prima regione dove si puo' accomodare un eventuale volume molecolare aggiuntivo, rispetto a quello dell'ATP, (EV), in rosa; una seconda regione dove si puo' accomodare un eventuale volume molecolare aggiuntivo, rispetto a quello dell'ATP, detta tasca di affinità (AP) in giallo. Gli eteroatomi del ligando sono colorati sulla base dei tipi atomici (S: giallo, N: blu ; O: rosso e C:arancione).

La ST0394AA1 si accomoda in modo simile a ST0388AA1 ma utilizzando la modalità di legame II; ossia il fenile con il Br in meta si accomoda nella tasca dove si posiziona la base dell'ATP (TB) ed interagisce con l'NH dello scheletro peptidico della Val851, il CS dell'anello 1,3-tiazolidin-4-one interagisce con l'NH della catena laterale della Lys802 e l'N piridinico mediante due molecole d'acqua interagisce con il CO della catena laterale della Ser773, mentre mediante una molecola d'acqua interagisce con lo zolfo della catena laterale della Met772 e si inserisce nella tasca delimitata dalla Ser774 e dalla Gln859 (Figura 44).



**Figura 44:** in figura è mostrato il risultato del docking dell'inibitore ST0394AA1 nel sito catalitico della PI3K $\alpha$ . Sono indicate: la tasca della valina (TV,) in verde; la tasca dove si posiziona lo zucchero dell'ATP (TZ,) in rosso; la tasca dove si posiziona la base dell'ATP (TB) in blu; la tasca dove si posiziona il fosfato dell'ATP (TF) in ciano; una prima regione dove si puo' accomodare un eventuale volume molecolare aggiuntivo, rispetto a quello dell'ATP, (EV), in rosa; una seconda regione dove si puo' accomodare un eventuale volume molecolare aggiuntivo, rispetto a quello dell'ATP, detta tasca di affinità (AP) in giallo. Gli eteroatomi del ligando sono colorati sulla base dei tipi atomici (S: giallo, N: blu ; O: rosso; C:verde; Br:marrone).

Il simile accomodamento degli inibitori determina dei cambiamenti conformazionali determinati a livello della porzione di interazione con Ras; in quanto l'interazione dell'anello con il Br con l'NH dello scheletro peptidico della Val851 favorisce un cambiamento conformazionale del Lobo N che a sua volta determina una variazione conformazionale del dominio di legame di Ras (RBD) al quale è connesso, facendo ipotizzare che l'interazione dei ligandi con la proteina le faccia assumere uno stadio conformazione non compatibile con la sua interazione con Ras (Figura 45)

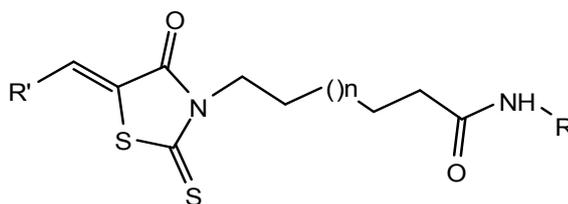


**Figura 45:** la figura 45A mostra il cambiamento conformazionale del ras binding domani (RBD) e del Lobo N indicati in figura ottenuti dai calcoli di docking dell'inibitore ST0388AA1; la figura 45B mostra il cambiamento conformazionale del RBD e del Lobo N indicati in figura ottenuti dai calcoli di docking dell'inibitore ST0394AA1. In particolare in bianco è riportata la conformazione di partenza della proteina; in blu il cambiamento conformazionale del RBD, in rosa il cambiamento conformazionale del Lobo N

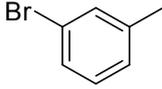
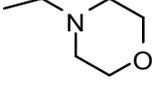
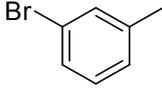
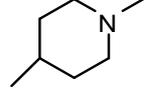
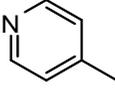
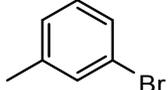
In accordo con gli studi effettuati, le SARs hanno confermato che, per quanto riguarda questa classe di inibitori della PI3K, giocano un ruolo chiave sia il sostituente R che R', in quanto piccole variazioni di tali sostituenti sono fondamentali per l'attività inibitoria. Sulla base dei risultati raggiunti, è stato possibile suggerire nuove modifiche strutturali che hanno portato alla progettazione di un nuovo set di ligandi.

Sono di seguito riportate le nuove strutture proposte (Tabella 20):

**Tabella 20:** nuovi inibitori progettati



ST	R'	n	R
0446		0	
0449		0	

0450		0	
0451		0	
0452		0	

### ***2.4h Progettazione di peptidi in grado di influenzare le interazioni con i partner proteici della p110 $\alpha$***

Per stabilizzare, invece, l'interazione inibitoria del dominio N-SH2 sulla p110 $\alpha$  sia per la forma mutata della PI3K (E542K, E545K) che per la forma non mutata si è proceduto alla progettazione di peptidi bioadesivi.

A tale scopo, è stata eseguita una dettagliata analisi bioinformatica e strutturale del dominio N-SH2 della subunità regolatoria p85 e della porzione di p110 $\alpha$  che contatta tale dominio.

Questi studi riportati in dettaglio nel paragrafo 2.4a.1, hanno permesso di individuare le porzioni aminoacidiche coinvolte in questa interazione inibitoria rispettivamente <sup>539</sup>PLSEITE<sup>545</sup> per p110 $\alpha$  e gli amminoacidi R340, T369, K379, L381 per il dominio N-SH2.

Prendendo in considerazione le due sequenze aminoacidiche individuate, sono stati progettati diversi peptidi caratterizzati da un linker costituito da un residuo di Pro e un residuo di Gly/Asp in modo da stabilizzare una conformazione strutturalmente chiusa dove i due motivi proteici sono disposti in modo tale da poter interagire separatamente con i loro partner.

Di seguito sono riportati i peptidi progettati per inibire la forma non mutata (Tabella 21) e a seguire (Tabella 24) quelli progettati per inibire la forma mutata della PI3K. Per validare l'ipotesi di progettazione tutti i peptidi sono stati sottoposti a studi computazionali. In particolare, è stato investigato lo spazio conformazionale dei nuovi peptidi attraverso la procedura di *Simulated Annealing* seguita da una minimizzazione energetica a livello di meccanica molecolare (MM).

I risultati ottenuti da tali studi hanno confermato che i peptidi progettati sono in grado di assumere una conformazione "bioadesiva".

### 2.4h.1 Peptidi Progettati per inibire la Forma Non Mutata della PI3K $\alpha$

Di seguito sono riportati i peptidi progettati per inibire la forma non mutata della PI3K (Tabella 21)

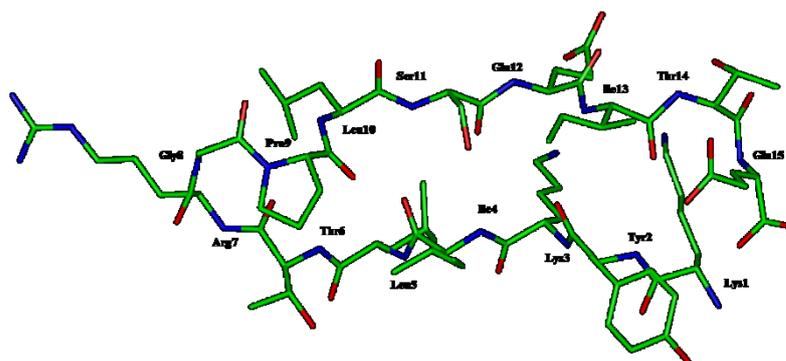
**Tabella 21:** peptidi progettati per la forma non mutata

Peptide	Sequenza
1	H-Lys-Tyr-Lys-Ile-Leu-Thr-Arg-Gly-Pro-Leu-Ser-Glu-Ile-Thr-Glu-OH
2	H-Lys-Ile-Leu-Thr-Arg-Gly-Pro-Leu-Ser-Glu-Ile-Thr-OH
3	H-Lys-Tyr-Lys-Ile-Leu-Thr-Arg-Asp-Pro-Leu-Ser-Glu-Ile-Thr-Glu-OH
4	H-Lys-Ile-Leu-Thr-Arg-Asp-Pro-Leu-Ser-Glu-Ile-Thr-OH
5	H-Ser-Glu-Ile-Thr-Glu-Asn-Gly-Pro-Lys-Ile-Leu-Thr-Arg-OH
6	H-Ser-Glu-Ile-Thr-Glu-Gly-Pro-Lys-Ile-Leu-Thr-Arg-OH

Sulla base degli studi effettuati il peptide che meglio conservava una conformazione ottimale per l'interazione tra il dominio N-SH2 della p85  $\alpha$  e il corrispondente motivo sulla p110 $\alpha$  è il peptide 1 di seguito riportato (Figura 46)

#### PEPTIDE 1

	Leu10	Ser11	Glu12	Ile13	Thr14	Glu15	
Pro9							
Gly8	Arg7	Thr6	Leu5	Ile4	Lys3	Tyr2	Lys1



**Figura 46:** la figura mostra la sequenza e la struttura tridimensionale del peptide selezionato per l'interazione tra il dominio N-SH2 della p85  $\alpha$  e il corrispondente motivo sulla p110 $\alpha$ . Gli eteroatomi del peptide sono colorati sulla base dei tipi atomici (N: blu ; O: rosso e C: verde).

Sono stati effettuati studi di *pharmacophore fitting* per validare la capacità di accomodamento del peptide tra le due porzioni sopra descritte: in particolare nella figura seguente viene mostrato in magenta il dominio N-SH2 della p85 $\alpha$  al quale è stato sovrapposto nella zona di

interazione con il motivo della p110  $\alpha$  la porzione del peptide corrispondente in bianco; stesso discorso è stato fatto per il motivo della p110 $\alpha$  in ciano al quale è stato sovrapposto la porzione del peptide corrispondente in giallo (*Figura 47*)

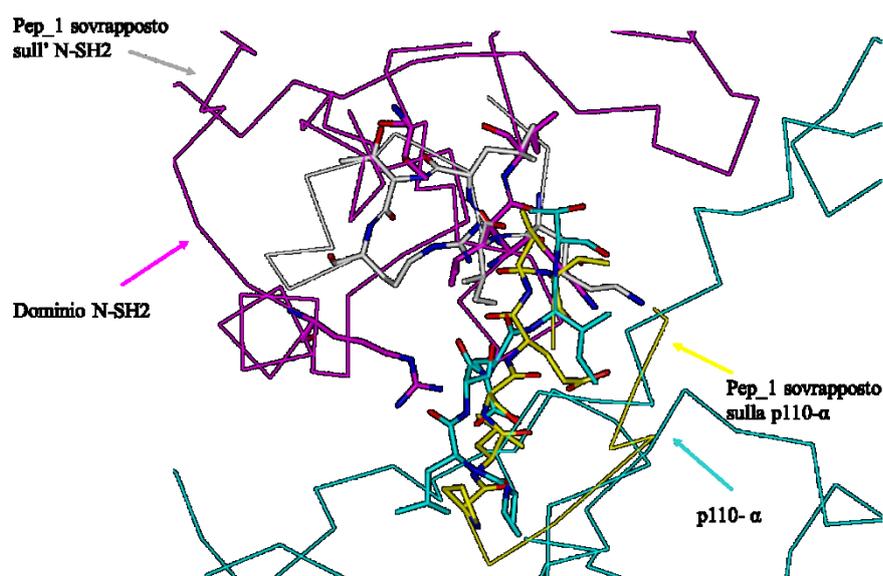
Di seguito sono riportate due tabelle (Tabella 22-23) in cui vengono mostrati gli aminoacidi considerati nella sovrapposizione:

**Tabella 22:** Sovrapposizione del peptide sull'N-SH2 di p85 $\alpha$

<b>N-SH2</b>	R340	T369	L380	I348	K379
<b>PEPTIDE</b>	R7	T6	L5	I4	K3

**Tabella 23:** Sovrapposizione del peptide sul motivo della p110 $\alpha$

<b>Motivo p110<math>\alpha</math></b>	P539	L540	S541	E542	I543	T544
<b>PEPTIDE</b>	P9	L10	S11	E12	I13	T14



**Figura 47:** nella figura è mostrata la sovrapposizione del peptide 1 (*Pep\_1*) alla PI3K $\alpha$ ; in magenta il dominio N-SH2 della p85 $\alpha$ ; in bianco sequenza del peptide 1 che si sovrappone all'N-SH2; in ciano motivo della p110 $\alpha$ , in giallo sequenza del peptide 1 che si sovrappone al motivo della p110 $\alpha$ . Gli eteroatomi del peptide sono colorati sulla base dei tipi atomici (N: blu ; O: rosso ).

## 2.4h.2 Peptidi Progettati per inibire la Forma Mutata (E542K-E545K) della PI3K $\alpha$

Dal momento che i residui aminoacidici della p110 $\alpha$  Glu542 e Glu545 che interagiscono con i residui Arg340 e Lys379 del dominio N-SH2 sono mutati in lisina, sono stati progettati nuovi peptidi con le stesse sequenze aminoacidiche di quelli precedenti variate solo nei residui di arginina e lisina sostituiti con due residui di glutammato.

Di seguito sono riportate le sequenze peptidiche progettate (Tabella 24):

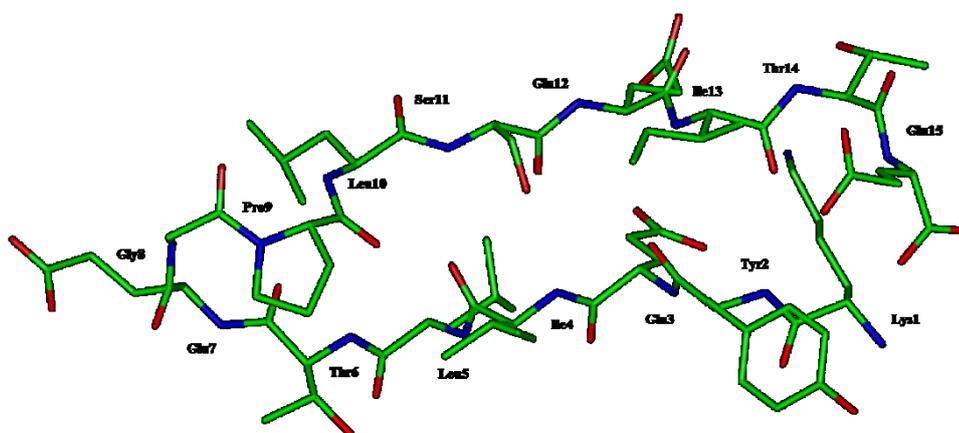
**Tabella 24:** peptidi progettati per la forma mutata

Peptide	Sequenza
7	H-Lys-Tyr-Glu-Ile-Leu-Thr-Glu-Gly-Pro-Leu-Ser-Glu-Ile-Thr-Glu-OH
8	H-Glu-Ile-Leu-Thr-Glu-Gly-Pro-Leu-Ser-Glu-Ile-Thr-OH
9	H-Lys-Tyr-Glu-Ile-Leu-Thr-Glu-Asp-Pro-Leu-Ser-Glu-Ile-Thr-Glu-OH
10	H-Glu-Ile-Leu-Thr-Glu-Asp-Pro-Leu-Ser-Glu-Ile-Thr-OH
11	H-Ser-Glu-Ile-Thr-Glu-Asn-Gly-Pro-Glu-Ile-Leu-Thr-Glu-OH
12	H-Ser-Glu-Ile-Thr-Glu-Gly-Pro-Glu-Ile-Leu-Thr-Glu-OH

Il peptide selezionato per la forma mutata della PI3K $\alpha$  è il peptide\_7 (Figura 48)

### PEPTIDE 7

	Leu10	Ser11	Glu12	Ile13	Thr14	Glu15	
Pro9							
Gly8	Glu7	Thr6	Leu5	Ile4	Glu3	Tyr2	Lys1



**Figura 48:** la figura mostra la sequenza e la struttura tridimensionale del peptide 7 selezionato per l'interazione tra il dominio N-SH2 della p85  $\alpha$  e il corrispondente motivo sulla p110 $\alpha$  con mutazione E542K-E545K. Gli eteroatomi del peptide sono colorati sulla base dei tipi atomici (N: blu; O: rosso e C: verde).

Tutti i nuovi inibitori della PI3K peptidici saranno sintetizzati dal prof. Maurizio Taddei del Dipartimento Farmaco Chimico Tecnologico dell'Università di Siena.

## ***2.5 Conclusioni***

In conclusione, gli studi condotti, hanno riguardato lo sviluppo di un'approccio computazionale per la progettazione di nuovi potenziali agenti antitumorali. I risultati ottenuti hanno portato allo sviluppo di un modello farmacoforico che ha permesso di ottenere inibitori (ST0388AA1, ST0394AA1) ad elevata attività ed altamente selettivi per la PI3K $\alpha$  partendo da un *lead* scarsamente attivo (ST0358AA1). I composti si sono rivelati citotossici in saggi preliminari riguardanti due linee cellulari tumorali. Inoltre, è stata effettuata la progettazione di peptidi potenzialmente in grado di stabilizzare il legame della p110 $\alpha$  mutata con la p85, al fine di combattere i fenomeni di tumorigenesi associati alle mutazioni E542K-E545K della PI3K $\alpha$ .

## 2.6 Bibliografia

1. Edward H. Walker et.al. “*Structural insights into phosphoinositide 3\_kinase catalysis and signalling*” Nature, 1999
2. Chuan-Hsiang Huang et. al. “*The structure of a Human p110 $\alpha$ /p85  $\alpha$  complex Elucidates the effects of oncogenic PI3K mutations*” Science, 2007
3. Tcherkezian et.al. “*Current knowledge of the large RhoGAP family of proteins*” BiolCell, 2007, 2, 67-86
4. Yu J, Zhang Y et. al. “*Regulation of the p85/p110 PI3K: stabilization and inhibition of the p110 $\alpha$  catalytic subunit by p85regulatory subunit*” Molecular Cell Biology, 1998, 18, 1379-1387
5. Jeffrey A. Engelman, et.al.” *The evolution of phosphatidylinositol 3-kinases as regulators of growth and metabolism*” Nature Reviews Genetics, 2006, 7, 606-619
6. Bing-Hua Jiang et. al. “*PI3K/PTEN signaling in tumorigenesis and angiogenesis*” Biochimica et Biophysica Acta, 2008, 1784, 150–158
7. Jimenez C et. al. “*The p85 regulatory subunit controls sequential activation of PI3K by Tyr kinase and Ras*” The Journal of Biological Chemistry, 2002, 277, 41556-41562
8. Pacold, M.E. et. al. “*Crystal structure and functional analysis of Ras binding to its effector phosphoinositide 3-kinase gamma*” Cell, 2000, 103, 931
9. Pawson T et. al. “*Protein–protein interactions define specificity in signal transduction*” Genes & Development, 2000, 14, 1027-1047
10. Chang F. Lee J. et al. “*Involvement of PI3K/Akt pathway in cell cycle progression, apoptosis and neoplastic transformation: a target for cancer chemotherapy*” Leukemia, 2003, 17, 590-603
11. Li Zhao and Peter K. Vogt. “*Helical domain and kinase domain mutations in p110 $\alpha$  of phosphatidylinositol 3-kinase induce gain of function by different mechanisms*” The National Academy of Science, 2008
12. M. Osaki et. al. “*PI3K-Akt pathway its function and alterations in human cancer*” Apoptosis, 2004, 9, 667-676
13. González B et.al. “*Structure of a human inositol 1,4,5-trisphosphate 3-kinase: substrate binding reveals why it is not a phosphoinositide 3-kinase*” Molecular Cell, 2004, 15, 689-701.

14. Maple J. R. et al. "*Derivation of class II force fields. VIII. Derivation of a general quantum mechanical force field for organic compounds*" Journal of Computational Chemistry, 2001, 22, 1782 – 1800
15. Fletcher, R. "*Unconstrained Optimization. In Practical Methods of Optimization John Wiley & Sons*" New York 1980
16. MOPAC2007, James J. P. Stewart, Stewart Computational Chemistry, Colorado Springs, CO, USA, <http://OpenMOPAC.net>, 2007
17. Stewart J. J. P. Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements, Journal of Molecular Modeling, 2007, 13, 1173-1213
18. J. Baker." *An Algorithm for the Location of Transition States.*" Journal of Computational Chemistry, 1986, 7, 385-395
19. Ding, H. Q. et.al. "*III Atomic level simulations on a million particles: The cell multipole method for Coulomb and London non-bond interactions*" The Journal of Chemical Physics, 1992, 97, 4309-4315

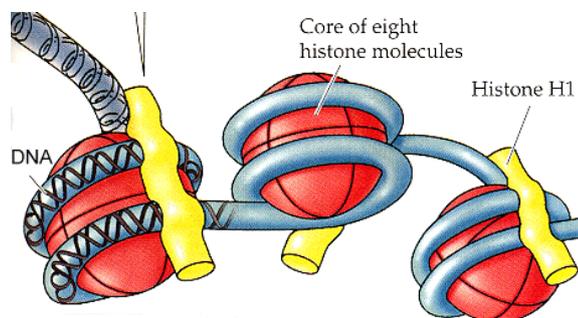
### ***3. Progettazione di nuovi inibitori delle Deacetilasi Istoniche (HDAC)***

#### ***3.1 Introduzione***

##### ***3.1a Struttura della Cromatina e Codice Istonico***

Nel nucleo delle cellule eucariotiche il genoma è impaccato nella cromatina (dal greco *Khroma*, colorato), un nome dato da Fleming basato sulla proprietà di questa sostanza di colorarsi intensamente con i coloranti basici<sup>1</sup>.

La cromatina è un complesso proteine-DNA altamente organizzato. L'unità fondamentale della cromatina è il nucleosoma (*Figura 1*), costituito da un ottamero formato da quattro nuclei istonici (un tetramero H3/H4 e due dimeri H2A e H2B), attorno al quale il DNA si avvolge in modo sinistrorso per 1,65 giri corrispondenti a 146 coppie di basi<sup>2</sup>. Vi è inoltre un istone H1 con la funzione di *linker*, che permette l'assemblaggio della cromatina ad un livello di organizzazione successivo e più complesso.



**Figura 1: Struttura del nucleosoma.** Il DNA avvolge il nucleosoma, formato dalle otto molecole istoniche. In rosso è rappresentato il core nucleosomico, in blu il DNA e in giallo l'istone H1.

Esistono infatti diversi livelli di organizzazione della cromatina. Il primo livello è lo stadio detto "filo a collana di perle" per il suo aspetto, costituito da una fibra di 11 nm di diametro. In tale fibra il DNA è avvolto attorno ai nucleosomi, intervallati l'uno dall'altro da un tratto di DNA *linker* di lunghezza variabile, da poche coppie a circa 80 nucleotidi, senza ulteriori ripiegamenti. La formazione dei nucleosomi converte una molecola di DNA in un filo di cromatina lungo circa un terzo della lunghezza iniziale. Questo stadio non necessita della presenza dell'istone H1.

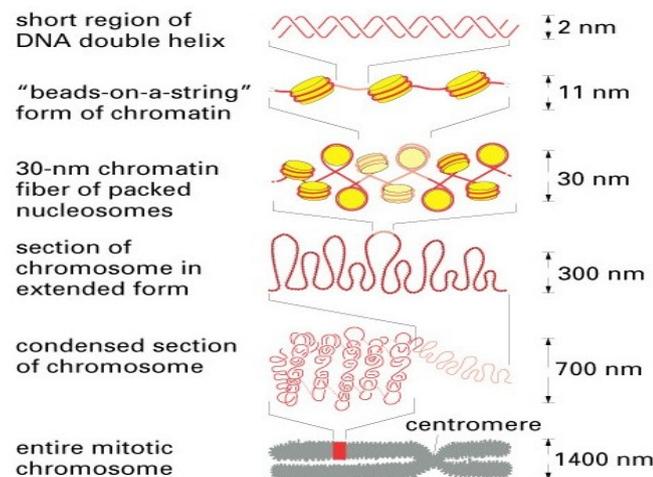
Il secondo livello è la fibra da 30 nm di diametro; in esso la cromatina assume un aspetto sinusoidale grazie alle interazioni che gli istoni H1 formano fra di loro: è lo stadio in cui si

trova la cromatina attiva in interfase, periodo compreso fra due divisioni cellulari, quando cioè la cromatina viene trascritta. La fibra si avvolge su se stessa, con una periodicità di 6 nucleosomi per ogni giro.

Il terzo livello è quello della “fibra ad ansa” da 300 nm di diametro, in questo stadio la cromatina si ripiega ulteriormente su se stessa grazie anche all'aiuto di altre proteine.

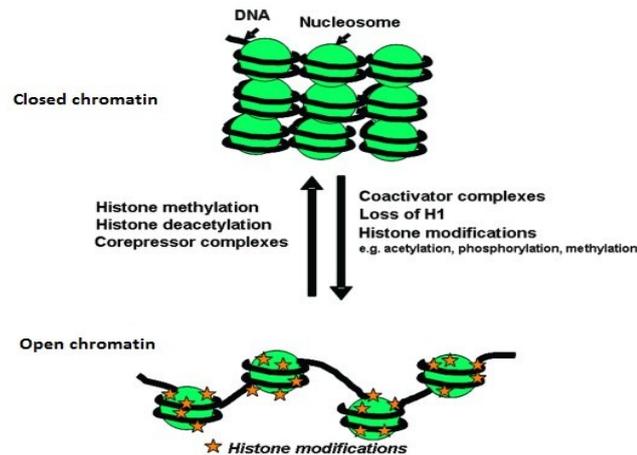
Il quarto livello è caratterizzato dalla formazione dei cromatidi: la cromatina si superavvolge in una fibra da 700 nm di diametro.

Il quinto ed ultimo livello vede la formazione dei cromosomi metafasici; è il livello di condensazione massimo che costituisce la fibra da 1400 nm di diametro (Figura 2).



**Figura 2: livelli di impacchettamento della cromatina.** Dall'alto: breve tratto di DNA; filo a collana di perle, di 11 nm; fibra da 30 nm, in cui gli istoni H1 interagiscono tra loro; fibra ad ansa, da 300 nm; formazione dei cromatidi, fibra da 700 nm; formazione dei cromosomi metafasici, fibra da 1400 nm

Per essere accessibile al gran numero di complessi responsabili della trascrizione e della replicazione il DNA non si deve trovare avvolto agli istoni. Per questo motivo il nucleosoma esiste in una conformazione condensata, in cui la trascrizione è repressa, o in una conformazione aperta, che è trascrizionalmente attiva. Le proteine regolatorie causano cambiamenti nella conformazione del nucleosoma attraverso modifiche chimiche degli istoni. La combinazione di tali trasformazioni, dunque, definirebbe i diversi stati della cromatina e fa ipotizzare la presenza di un vero e proprio codice istonico<sup>3</sup> (Figura 3).



**Figura 3: stati di condensazione della cromatina.** La perdita dell'istone H1, complessi coattivi e modificazioni istoniche quali l'acetilazione, la fosforilazione e la metilazione, comportano l'apertura della cromatina. Al contrario la metilazione e la deacetilazione degli istoni e complessi corepressori determinano la chiusura della cromatina.

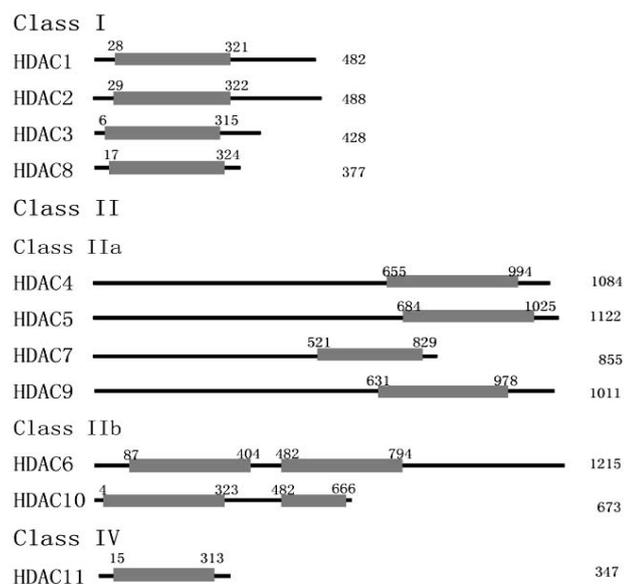
Queste modifiche includono fosforilazione, metilazione, ribosilazione, acetilazione e deacetilazione. Molte riguardano i residui di lisina vicini alla porzione N-terminale delle proteine istoniche. Questi residui di lisina sono carichi positivamente a pH fisiologico. La carica positiva dei residui di lisina interagisce con l'ossigeno anionico dei gruppi fosfato del backbone del DNA, formando interazioni ioniche forti che mantengono il DNA avvolto attorno agli istoni. L'enzima istone-acetiltransferasi è responsabile del trasferimento di un gruppo acetile sulla porzione amminica del residuo di lisina, che neutralizza la carica positiva e permette al DNA di svolgersi dall'istone. In questa conformazione il DNA è trascrizionalmente attivo. L'enzima istone-deacetilasi catalizza la reazione inversa, ossia la rimozione del gruppo acetile dal residuo di lisina, instaurando nuovamente l'interazione ionica, favorendo il ritorno del DNA alla conformazione condensata, trascrizionalmente inattiva. In breve, l'istone deacetilasi (HDAC) promuove la repressione del gene, mentre l'acetiltransferasi (HAT) ne promuove la trascrizione. La funzione degli HDAC verrà trattata in modo più specifico nei capitoli successivi<sup>4</sup>.

### ***3.1b Struttura e Funzione degli HDAC***

Gli HDAC appartengono a una famiglia di metalloproteasi che rimuovono il gruppo acetile dalla porzione  $\epsilon$ -amminica dei residui di lisina presenti negli istoni nucleosomali. L'idrolisi del gruppo acetile comporta l'aumento della densità della carica positiva sulla porzione N-terminale degli istoni nucleosomali, rafforzando così l'interazione con la carica negativa del DNA e bloccando l'accesso ai fattori di trascrizione. L'azione degli HDAC gioca un ruolo determinante anche in altri meccanismi, quali il rimodellamento della cromatina, la ricombinazione e la riparazione del DNA. Recentemente una gran quantità di proteine non istoniche sono state riconosciute come substrato degli HDAC, compresi fattori di trascrizione, proteine citoscheletriche, chaperoni molecolari. Questo ovviamente estende e complica la funzione degli HDAC e ci indica come regolare l'attività di questi enzimi potrebbe essere di aiuto ad esempio nello sviluppo di terapie anti-cancro, nella cura delle infezioni virali, delle malattie neurodegenerative e delle infiammazioni.

Finora, nell'uomo sono stati identificati diciotto tipi di HDAC, che possono essere suddivisi in quattro classi in base alle loro analogie con gli HDAC del lievito *Saccharomyces cerevisiae*, alla loro localizzazione subcellulare, alla loro specificità tissutale e alla loro attività enzimatica.

Gli HDAC di classe I (HDAC 1, 2, 3 e 8), omologhi della proteina RPD3 del lievito, sono prevalentemente nucleari, ubiquitariamente espressi e funzionano come corepressori della trascrizione. La Classe II degli HDAC comprende gli HDAC 4, 5, 6, 7, 9 e 10 ed è ulteriormente suddivisa in due sottoclassi, la IIa (HDAC 4, 5, 7 e 9), in cui è presente un solo dominio deacetilasi, e la IIb (HDAC 6 e 10), in cui ne sono presenti due (*Figura 4*). Questi enzimi sono omologhi della proteina Hda1 del lievito e fanno da spola tra il nucleo ed il citoplasma in risposta a determinati segnali cellulari; sono selettivamente espressi nei vari tessuti e questo suggerisce che abbiano ruoli diversi nella differenziazione cellulare e nei processi di sviluppo. Gli HDAC della classe IIa si trovano prevalentemente nel cuore, nella muscolatura scheletrica e nel cervello; quelli della classe IIb hanno una maggiore localizzazione citoplasmatica e si trovano principalmente nella milza, nel rene e nel fegato<sup>2</sup>. Gli enzimi appartenenti alla classe I sono costituiti da 350-500 amminoacidi, mentre gli enzimi di classe II da circa 1000 amminoacidi. La disparità di grandezza tra le due classi è dovuta alla lunga porzione N-terminale che caratterizza la classe II. Entrambe sono però accomunate da una regione di omologia di sequenza di 320 residui amminoacidici, che corrisponde alla sezione dell'enzima che forma la tasca catalitica<sup>5</sup>.



**Figura 4: classificazione degli HDAC.** I numeri sulla destra indicano da quanti amminoacidi è composta ogni isoforma. La barra grigia rappresenta il sito attivo

La tasca catalitica altamente conservata rappresenta una sfida per lo sviluppo di inibitori enzimatici selettivi. Tuttavia, alcune differenze strutturali sono state trovate e vengono utilizzate per generare inibitori selettivi per le varie isoforme delle diverse classi<sup>6</sup>.

L'HDAC11 è l'isoforma di più recente scoperta, ed è classificata come una deacetilasi di classe IV. Presenta caratteristiche simili agli HDAC di classe I e II, ma non tali da includerlo in esse. Ciò che accomuna la classe I, II e IV è che gli enzimi che ne fanno parte sono tutti  $Zn^{2+}$  dipendenti. Gli HDAC di classe III sono invece definiti Sirtuine o Sir2 (Sirt 1-7). L'acronimo sta per *Silent Information Regulator 2*, un'enzima istone deacetilasi  $NAD^+$  dipendente, che utilizza il  $NAD^+$  per la rimozione dei gruppi acetili dai residui di lisina acetilati. Questi enzimi si distinguono quindi dagli altri HDAC per differenti meccanismo d'azione e struttura<sup>7</sup>. Fatta eccezione per l'HDAC8, questa categoria enzimatica non si presenta mai in singoli monomeri, piuttosto si ritrova in complessi multiproteici in cui differenti classi di HDAC sono spesso associate sia con specifici coregolatori sia con altri enzimi che modificano la cromatina<sup>8</sup>.

L'attività biologica degli HDAC può essere separata in due aree d'azione distinte: quella enzimatica, caratterizzata dall'abilità di deacetilare gli istoni o altri substrati proteici non istonici, e l'attività funzionale, che consiste nella capacità di regolare la trascrizione e altri processi biologici; non in tutte le circostanze la funzione biologica degli HDAC è strettamente dipendente dall'attività enzimatica<sup>9</sup>.

L'esatto ruolo biologico degli HDAC nelle cellule non è stato ancora del tutto compreso e molti gruppi di ricerca sono attivi in questo campo. La classificazione delle distinte funzioni

dei vari membri degli HDAC è stata messa in evidenza usando varie tecniche di indagine biologica, quali inibizione dell'RNA ed esperimenti sui “topi *knockout*”, oppure farmacologica utilizzando piccole molecole inibitrici. La sperimentazione sui “topi *knockout*” è il tipo di indagine che ha fornito maggiori informazioni. Un topo *knockout* è un topo geneticamente modificato, nel quale uno o più geni sono stati “spenti” attraverso l'aggiunta nel genoma di un particolare gene detto *knockout*. Tali studi rappresentano importanti modelli per analizzare il ruolo di geni sequenziati ma dalle funzioni ignote; la comparsa di eventuali anomalie nelle normali funzionalità dell'animale fornisce importanti informazioni circa la funzione svolta dal gene silenziato e dalle molecole derivanti dalla sua traduzione.

In questi studi si è notato che la distruzione di entrambi gli alleli del gene per gli HDAC1 causa difetti di proliferazione e ritardo nello sviluppo, con conseguente morte embrionale. Contemporaneamente è indotta l'espressione di HDAC 2 e 3, ma la presenza di questi enzimi non riesce a compensare l'assenza dell'HDAC1. Questo suggerisce una unicità di funzione degli enzimi. I topi in cui, invece, viene silenziato il gene per gli HDAC2 sopravvivono fino al periodo prenatale per poi soccombere per l'insorgenza di difetti cardiaci, e questo ci suggerisce la correlazione tra questi enzimi e la regolazione delle funzioni cardiache. È stato dimostrato invece che gli HDAC3 sono regolatori critici della mitosi, indipendenti dai processi trascrizionali. L'espressione degli HDAC8, infine, è stata trovata nelle cellule muscolari lisce, dove questi enzimi sono associati all'actina citoscheletrica e regolano la capacità contrattile.

Come già accennato precedentemente, a differenza della classe I, la classe IIa mostra un'espressione tessuto-specifica ed è caratterizzata da una coda N-terminale molto lunga che conferisce a questa classe la capacità di rispondere bene a un'ampia varietà di segnali di trasduzione. Altri topi *knockout* si sono ottenuti per gli HDAC 4, 5 e 9. I topi privi dell'HDAC4 mostrano un fenotipo caratterizzato da condrociti ipertropici, responsabili di una formazione ectopica dell'osso. È stato dimostrato che nei normali neuroni questi enzimi sono mediatori della morte neuronale. Gli HDAC 5 e 9 sono invece coinvolti nello sviluppo cardiaco perché i topi *knockout* per questi enzimi mostrano ipertrofia cardiaca.

L'HDAC7, infine, risulta essere un sistema timo-specifico e inibisce l'espressione del gene Nur77 coinvolto nell'apoptosi e nella regolazione negativa dello sviluppo dei timociti<sup>10</sup>. I sottotipi 6 e 10 della classe IIb presentano due domini catalitici, ma si pensa che il secondo non sia funzionante nell'HDAC10. L'HDAC6 è l'unica deacetilasi conosciuta ad agire sulla tubulina. La deacetilazione della tubulina è necessaria per la formazione degli aggresomi, organelli necessari alla rimozione di alcune proteine citoplasmatiche tossiche. Gli HDAC6

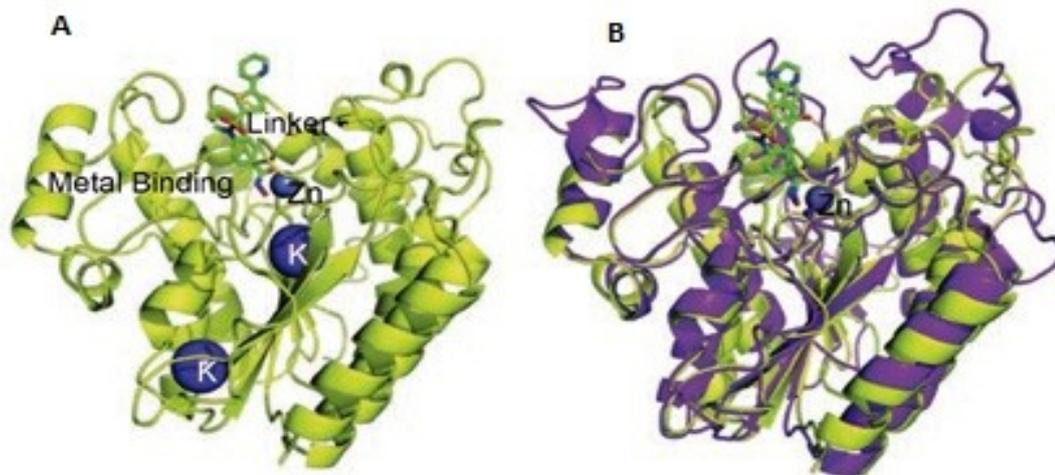
citoplasmatici hanno anche il compito di prevenire associazioni non corrette tra catene polipeptidiche non ripiegate, sia in condizioni fisiologiche che in condizioni di stress. L'HDAC10 agisce sia nel nucleo che nel citoplasma, dove svolge attività deacetilasi. Infine per l'HDAC11, l'ultimo ad essere individuato, non è ancora stata chiarita del tutto la funzione svolta<sup>9</sup>.

**Tabella 1:** funzioni biologiche degli HDAC di classe I e II

Isoforme	Funzioni biologiche	Modelli di riferimento
Classe I		
HDAC1	proliferazione, regolazione genetica, espressione dell'apoptosi	topo knockout, siRNA nelle cellule tumorali
HDAC2	proliferazione, modificazioni cardiache	topo knockout, siRNA nelle cellule tumorali
HDAC3	proliferazione mitotica, regolazione dell'espressione di INF	siRNA nelle cellule tumorali
HDAC8	capacità contrattile, regolazione della proliferazione, attività telomerasica	siRNA nelle cellule muscolari lisce umane, siRNA nelle cellule tumorali
Classe IIa		
HDAC4	ipertrofia dei condrociti, mediatori della morte neuronale, repressione di PLZF_RAR, stabilità di HIF_1	topo knockout, siRNA nei neuroni, siRNA nella leucemia promielocitica
HDAC5	soppressione degli stress cardiaci	topo knockout
HDAC7	regolazione dell'apoptosi	siRNA nei timociti
HDAC9	sviluppo cardiaco	topo knockout
Classe IIb		
HDAC6	acetilazione della tubulina, acetilazione di hsp90	distruzione del gene nelle cellule staminali, siRNA nella telomerasi, siRNA nelle cellule umane del rene

### 3.1c HDAC Zn<sup>2+</sup>-Dipendenti

Le prime informazioni sulla struttura degli HDAC Zn-dipendenti derivano dagli studi condotti sulla struttura cristallina dell'istone-deacetilasi proteina simile (HDLP), un enzima simile agli HDAC presente nel batterio *Aquifex aelicus*, complessato con gli inibitori SAHA e tricostatina A (TSA)<sup>11</sup>. Un altro importante passo avanti nella definizione della struttura degli HDAC è stato fatto nel 2004, da due gruppi di studio che hanno pubblicato, indipendentemente l'uno dall'altro, ma raggiungendo risultati simili, la struttura dell'HDAC8 umano, cristallizzato con cinque differenti inibitori (*Figura 5*)<sup>12</sup>.



**Figura 5: strutture cristallografiche dell'HDAC8.** a) struttura cristallografica dell'HDAC8 in complesso con TSA. b) sovrapposizione dell'HDAC8, in giallo, e dell'HDLP, in magenta, in complesso con TSA.

La struttura generale di questo enzima comprende un singolo dominio compatto  $\alpha/\beta$  composto da otto foglietti  $\beta$  centrali, situati tra le  $\alpha$  eliche undici o quattordici. Approssimativamente metà degli amminoacidi sono contenuti negli elementi della struttura secondaria, mentre l'altra metà forma delle regioni di legame che congiungono diversi elementi di questa struttura. Dallo studio della struttura cristallina dell'HDAC8 e dalla sua comparazione con l'HDLP, è emerso che, nonostante le grosse differenze tra i vari tipi di HDAC, esiste in tutti un dominio catalitico estremamente conservato che probabilmente è alla base del loro meccanismo d'azione. Le maggiori differenze risiedono invece proprio nelle regioni di legame. In particolar modo la regione di legame L1 sembra essere importante nella specificità di substrato in quanto delinea una larga porzione su una parete del sito attivo e si estende fino alla superficie della proteina. Questo comporta un aumento della larghezza del sito attivo vicino alla sua apertura, rispetto agli HDLP (simili tra l'altro agli HDAC 1, 2 e 3), rendendo l'HDAC8 capace di idrolizzare molti più substrati rispetto agli altri HDAC della classe I. Queste diversità giocano un ruolo cruciale nella selettività degli HDAC e nelle loro interazioni con le altre proteine. Entrambi i gruppi di ricerca hanno trovato due cationi monovalenti nella struttura cristallina dell'HDAC8, uno vicino al sito attivo e l'altro distante dal sito.

Il sito attivo dell'enzima consiste in un tunnel lipofilo lungo circa 12 Å, in fondo al quale si trova il sito catalitico. Gli amminoacidi che formano la parete del tunnel, la Phe152, Phe208, His180, Gly151, Met274 e la Tyr306, sono conservati negli HDAC di classe I eccetto la Met274, che è una leucina negli altri membri della famiglia. Particolarmente interessanti sono

le due catene laterali delle fenilalanine, che sono parallele tra loro e formano una sorta di alloggiamento, sito di interazione di alcuni HDACi.

L'architettura del sito attivo è conservata tra l'HDAC8 e l'HDLP, e sulla base dell'omologia di sequenza è molto simile negli altri HDAC Zn-dipendenti. Alla fine del canale è presente lo ione  $Zn^{2+}$ , il centro del sito catalitico; lo  $Zn^{2+}$  è penta-coordinato con tre siti occupati da Asp178, His180 e Asp267 e con altri due siti occupati dagli ossigeni della porzione acetilica e dell'acqua, presente normalmente nel substrato.

Recentemente è stata studiata la struttura di un batterio omologo all'HDAC6 della classe IIb. Questo presenta le caratteristiche degli HDAC di classe I e contiene lo  $Zn^{2+}$  nel sito catalitico. La principale differenza strutturale con gli HDAC di classe I consiste nelle regioni di legame situate in prossimità dell'apertura del sito attivo, confermando l'importanza selettiva di questa regione<sup>10</sup>

### ***3.1d Inibitori degli HDAC di Classe I e II***

Gli enzimi istone-deacetilasi possono essere inibiti da piccole molecole. L'inibitore si lega strettamente al sito attivo e probabilmente lega il cofattore zinco, così che la catalisi non può avvenire.

La maggior parte degli inibitori HDAC, compresi quelli nella fase clinica, sono reversibili, sebbene si conoscano anche inibitori irreversibili<sup>13</sup>.

Gli HDACi degli enzimi appartenenti alle classi I e II sono caratterizzati da elementi strutturali comuni che permettono l'interazione con le differenti zone del tunnel catalitico dell'enzima. Queste caratteristiche sono riassunte in un modello farmacoforico comune a tutti i composti di questa classe. Il gruppo funzionale basilare di un HDACi è lo *zinc binding group* (ZBG), che chela lo ione zinco e lega il sito attivo attraverso diversi legami a idrogeno. Lo ZBG è tipicamente un tiolo, un acido idrossammico, un acido carbossilico, un chetone o delle aniline sostituite.

Gli inibitori irreversibili hanno invece un gruppo elettrofilo, solitamente un cheto-epossido, nella regione ZBG, che potrebbe legare covalentemente un residuo nucleofilo, come un'istidina, un aspartato, o una tirosina nel sito attivo<sup>14</sup>. Il *linker* è generalmente una catena idrofobica, omologa alla catena laterale idrofobica della lisina. Il linker congiunge lo ZBG e il dominio capace di interagire con il margine superiore del tunnel catalitico dell'enzima, detto anche CAP. Il CAP è legato al linker attraverso un'unità di connessione, solitamente un gruppo ammidico. Questa porzione spesso è costituita da sistemi anellari aromatici ed

eteroaromatici. Sono stati costruiti inibitori in cui i CAP sono costituiti da diversi gruppi, e alcuni di questi mostrano una maggiore potenza e selettività<sup>15</sup>. Il farmacoforo per costruire un HDACi è quindi costituito da uno ZBG, un linker, il CAP e la porzione che congiunge il linker al CAP. Tutti i domini funzionali negli inibitori dell'istone deacetilasi possono essere ottimizzati per aumentare la potenza dell' inibitore e per generare selettività verso le isoforme.

### ***3.1e Potenziali Bersagli Terapeutici degli HDACi***

Ad oggi la principale applicazione farmacologica per gli inibitori HDAC è il trattamento del cancro.

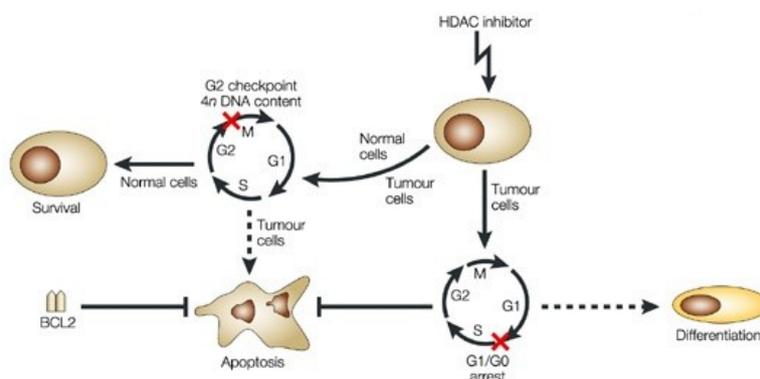
I ricercatori focalizzano la propria attenzione sulla progettazione di farmaci capaci di agire su specifiche alterazioni molecolari trovate solo nelle cellule trasformate. Il fine della ricerca è associare ogni tumore all'espressione di un determinato gene, definendo così le alterazioni responsabili di ogni tipo di cancro. I processi molecolari che portano all' attivazione o alla repressione della trascrizione sono considerati come possibili bersagli per le terapie anticancro.

Gli HDAC sono stati accomunati al tumore per la prima volta durante gli studi sulla leucemia promielocitica acuta<sup>16</sup>. Da allora, si è dimostrato che la silenziamento o l'inibizione degli HDAC ha effetto sul ciclo cellulare, sulla crescita cellulare, sulla decondensazione della cromatina, sulla differenziazione cellulare, l'apoptosi e l'angiogenesi in diversi tipi di cellule cancerogene. È stato ricercato il ruolo dell'HDAC1 nelle cellule tumorali usando RNAi mediato da proteine *knockdown* e si è dimostrato il coinvolgimento di questa isoforma nell'inibizione della crescita e nell'apoptosi. Per quanto riguarda l'HDAC2 si è riscontrato un aumento della proliferazione di questo enzima nelle cellule gastriche affette da cancro nell'uomo. Un effetto antiproliferativo è presente nelle cellule cancerogene trattate con siRNA per HDAC3. L'inibizione dell'HDAC4 nelle cellule renali umane affette da carcinoma dimostra come questo enzima, assieme all'HDAC6, sia un partner della subunità R del fattore ipossia inducibile (HIF-1R), probabilmente regolando il livello di acetilazione della proteina e quindi la sua stabilità. Nella leucemia promielocitica è provato che l'HDAC4 reprime la fusione della proteina PLZF-RAR $\alpha$ , responsabile di una repressione trascrizionale.

Gli HDACi potrebbero quindi indurre l'espressione dei geni. Nella realtà, tuttavia, numerosi geni della trascrizione subiscono una *up-regulation* mentre altri una *down-regulation* influenzando così molti altri meccanismi d'azione e coinvolgendo diverse proteine accessorie. Inoltre è stato dimostrato che l'attività antitumorale degli HDACi è collegata con l'acetilazione

delle proteine non istoniche, il che complica ulteriormente il meccanismo d'azione. La razionalizzazione di queste scoperte è fondamentale per poter effettuare studi clinici efficaci e sviluppare nuove terapie.

Un aumento dell'acetilazione degli istoni durante la fase S (di sintesi del DNA) e la fase G2 (la fase pre-mitotica) del ciclo cellulare può portare a un controllo in G2 che causa l'arresto delle cellule nella suddetta fase. La perdita del controllo G2 è un evento frequente nelle cellule cancerogene e può comportare, in parte, un aumento della sensibilità di queste cellule rispetto a quelle sane agli effetti proapoptotici degli HDACi <sup>17</sup>(Figura 6). Al di là del cancro, l'attività degli inibitori HDAC può essere di interesse in numerose altre aree terapeutiche, come l'immunomodulazione. Gli HDACi hanno effetti sull'acetilazione dei fattori chiave nella regolazione delle funzioni delle cellule immunologiche. Il risultato è la soppressione delle citochine proinfiammatorie come TNFR e IL-1. Un'altra possibile applicazione è nell'apoplezia: gli HDACi incrementano l'espressione di proteine neuroprotettive, come Hsp70 e Bcl2, presenti nel cervello. Questa azione sulle proteine neuroprotettive può essere una possibile modalità di trattamento di pazienti colpiti da ischemia.



**Figura 6: azione degli inibitori HDAC sul ciclo cellulare.** Le cellule sane sopravvivono a un arresto in G2 indotto dagli HDACi, quelle tumorali vanno

Dai vari studi sugli HDAC, anche utilizzando tecniche di DNA ricombinante, è emerso, inoltre, che HDACi strutturalmente differenti possono avere determinate attività su diverse classi di HDAC o persino su specifiche isoforme<sup>10</sup>.

### ***3.1f Approcci Terapeutici***

Sempre più numerosi sono i gruppi di ricerca impegnati ad indagare, tramite approcci di tipo biologico e chimico, il ruolo chiave delle le interazioni fra le singole proteine nei sistemi biologici complessi. Diverse classi di composti a basso peso molecolare hanno mostrato un'interessante capacità di modulare funzioni biologiche mediate da enzimi facenti parte di macro-complessi proteici, generando quindi degli effetti farmacologici di grande interesse. Essi a ben ragione rappresentano quindi degli strumenti utili per chiarire il meccanismo catalitico di molti enzimi coinvolti nella regolazione epigenetica della trascrizione.

L'elevato livello di informazioni attualmente accumulate nel campo dell'epigenetica, insieme alla necessità di scoprire nuove e più efficaci molecole utilizzabili in terapia, spiega il grande interesse suscitato in quest'area ed in particolare verso la possibilità di ottenere un'inibizione degli enzimi responsabili delle modifiche chimiche alla base della regolazione della trascrizione genica, attraverso l'utilizzo di composti aventi caratteristiche chimico-fisiche adatte al loro potenziale utilizzo, quali farmaci (*drugable compounds*).

Infatti, c'è una crescente evidenza che l'espressione genica, governata dalla regolazione epigenetica, è cruciale nella fase iniziale di malattie quali il cancro. L'insorgenza e la progressione di questa e di numerose altre patologie è caratterizzata da processi multistadio che implicano alterazioni cromosomiali che portano ad una perdita della funzione di alcuni geni o l'iperattivazione di altri. Le alterazioni epigenetiche alterano i livelli dell'espressione genica senza modificare strutturalmente il DNA, perciò provocando malattie il cui decorso, contrariamente a quanto avviene per le mutazioni di DNA, può essere potenzialmente invertito attraverso l'inibizione di enzimi responsabili delle modificazioni epigenetiche.

Nell'ambito della regolazione epigenetica, gli enzimi HDAC svolgono molte funzioni fisiologiche e, sebbene la loro azione sia ben lontana dall'essere completamente chiarita, tuttavia la relazione tra le anomalie della loro funzionalità e l'origine di diverse patologie ha motivato il crescente interesse verso questi enzimi e di conseguenza anche verso i loro inibitori. Teoricamente bloccando l'attività degli HDAC, attraverso un inibitore, si favorisce la decondensazione della cromatina con un incremento generale della trascrizione genica. Gli inibitori HDAC inducono, inoltre, un aumento dell'acetilazione degli istoni. Nella realtà tuttavia, molti geni della trascrizione subiscono una up-regulation mentre altri una down-regulation influenzando così molti altri meccanismi d'azione e coinvolgendo diverse proteine accessorie. Inoltre è stato dimostrato che l'attività antitumorale degli inibitori HDAC è anche collegata con l'acetilazione delle proteine non istoniche, questo complica ulteriormente il

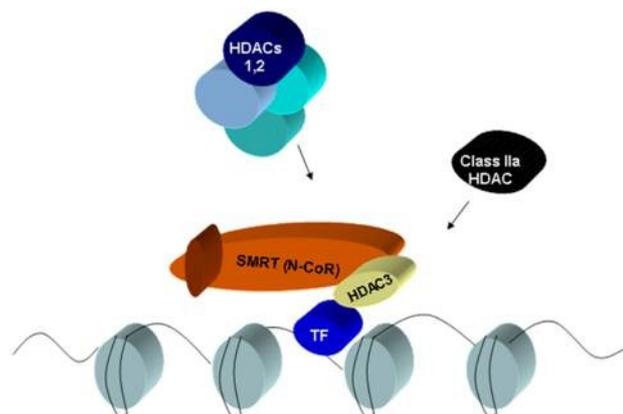
meccanismo d'azione. La razionalizzazione di queste scoperte è fondamentale per poter effettuare studi clinici efficaci e sviluppare nuove terapie. Oggetto del mio lavoro di dottorato è stato, quindi, lo studio della struttura molecolare degli HDAC umani e dei loro inibitori al fine di chiarirne la potenzialità quale bersaglio terapeutico e di indirizzare lo sviluppo di nuovi potenziali farmaci.

### 3.2 Scopo e Razionale del lavoro

La ricerca riguardante gli HDAC è stata svolta in collaborazione con l'Industria Farmaceutica Sigma-Tau (Sigma-Tau, S.p.A., Pomezia, Italia). In particolare, il mio obiettivo è stato quello di progettare inibitori in grado di: i) intrappolare l'enzima dal punto di vista conformazionale, ii) impedirgli di legare i cofattori necessari al legame con le proteine istoniche, iii) favorire l'omodimerizzazione dell'enzima attraverso la contemporanea occupazione dei due rispettivi siti catalitici. A tal fine, sono stati progettate nuove strutture caratterizzate da: I) diversi gruppi in grado di coordinare lo zinco presente nel sito catalitico (*zinc binding groups*, ZBGs), II) diversi sostituenti in grado di proiettarsi all'esterno del sito catalitico verso la superficie dell'enzima.

Infatti, l'attività catalitica delle diverse isoforme di HDAC è associata alla formazione di macrocomplessi molecolari, la cui composizione varia a seconda della sottofamiglia a cui appartengono. Gli HDAC di classe I formano macrocomplessi con almeno quattro tipi di corepressori genici. In particolare, gli HDAC3 sono reclutati unicamente dal complesso SMRT/NCoR<sup>18</sup>, mentre HDAC1 e HDAC2 sono attivati attraverso il reclutamento di alcuni corepressori come Sin3A<sup>19</sup>, CoREST<sup>20</sup>, and NuRD<sup>21</sup>.

L'HDAC4 e l'HDAC7 (classe IIa), invece, si associano direttamente al macrocomplesso HDAC3- SMRT/NCoR che, a sua volta, si lega al fattore trascrizionale e va a deacetilare le proteine istoniche<sup>22</sup>. (Figura 7)



**Figura 7:** La classe I HDACs forma macrocomplessi con almeno quattro dei maggiori corepressori. Gli HDAC3 sono reclutati unicamente dal complesso SMRT/NCoR, mentre HDAC1 e HDAC2 sono attivati attraverso il reclutamento di alcuni corepressori come Sin3A, CoREST, and NuRD. Gli HDAC di classe II, invece, si associa direttamente al macrocomplesso HDAC3- SMRT/NCoR che a sua volta si lega al fattore trascrizionale e va a deacetilare le proteine istoniche.

Infine, in collaborazione con l'Industria Farmaceutica Indena (Indena, S.p.A., Milano) ho investigato le basi molecolari dell'interazione fra la curcumina e gli HDAC. Infatti, l'attività

inibitoria della curcumina è stata misurata essere micro molare nei confronti di estratti cellulari contenenti HDAC di classe I ma non è mai stata misurata sulle diverse isoforme isolate. Tutti gli studi effettuati riguardanti la progettazione dei differenti tipi di inibitori, verranno descritti in dettaglio seguendo l'ordine sopra riportato.

### ***3.3 Parte Sperimentale***

Tutti gli studi di modellistica molecolare sono stati eseguiti su Silicon Graphics SGI Origin 200 8XR12000 ed E4 Server Twin 2 x Dual Xeon—5520, equipaggiati con due nodi; ogni nodo: 2 x Intel® Xeon® QuadCore E5520—2,26Ghz, 36 GB RAM.

#### ***3.3.1a Analisi Conformazionale dei nuovi inibitori caratterizzati da un nuovo gruppo coordinante lo zinco (ZBG)***

Le forme tautomeriche e i valori di pKa dei nuovi composti progettati, sono stati calcolati utilizzando il software ACD/Percepta, versione 14.0.0 (Advanced Chemistry Development Inc., Toronto, Canada)<sup>23</sup>.

Di conseguenza, le percentuali delle forme neutre/protonate sono state determinate considerando il valore di pH citoplasmatico (7.2) usando l'equazione di Henderson-Hasselbach. Le molecole progettate sono state costruite nella forma ionica prevalente per ogni tautomero considerando valori di pH 7.2 utilizzando il modulo *Builder di Insight 2005* (Accelrys Software Inc., San Diego). I potenziali atomici e le cariche parziali sono state assegnate utilizzando il campo di forza CVFF<sup>24</sup>. In particolare per quanto riguarda lo ZBG le molecole prese in considerazione per il calcolo del pKa sono state: ST7727, ST7729, ST7732, ST7991

Lo spazio conformazionale di questi composti è stato campionato attraverso 200 cicli di *Simulated Annealing* ( $\epsilon=1$ ) eseguito mediante la minimizzazione energetica (MM). Una temperatura iniziale di 1000 K è stata applicata al sistema per 2000 fs (time step= 3.0 fs) al fine di superare tutte le barriere torsionali; successivamente la temperatura è stata ridotta a 300 K in 1000 fs (time step= 1.0 fs) con un decremento di 0.5 K/fs. Le strutture risultanti sono state sottoposte a minimizzazione energetica utilizzando il modulo Discover\_3 di Insight 2005 (campo di forza CVFF;  $\epsilon= 1$ ). La minimizzazione energetica è stata eseguita usando come algoritmo di minimizzazione *Conjugate Gradient*<sup>25</sup>; il criterio di convergenza è stato stabilito in modo che il massimo valore di RMS sulla derivata prima sia inferiore a 0.001 kcal/Å. Al fine di analizzare in maniera dettagliata le proprietà elettroniche, tutti i conformeri, ottenuti dai calcoli di meccanica e dinamica molecolare, sono stati sottoposti ad una piena ottimizzazione geometrica (*full geometry optimization*) usando il metodo quanto-meccanico PM6<sup>26</sup> di Mopac 2009<sup>27</sup>.

Al fine di raggiungere un gradiente geometrico (GNorm) fissato al valore di 0.01 è stato utilizzato l'algoritmo di minimizzazione EF<sup>28</sup> (*Eigenvector Following routine*). Per ottenere

una completa ottimizzazione della geometria di partenza abbiamo aumentato i criteri per la terminazione di tutte le ottimizzazioni di un fattore di 100, usando la *keyword* *PRECISE*.

I conformeri ottenuti sono stati quindi classificati per quanto riguarda lo ZBG secondo:

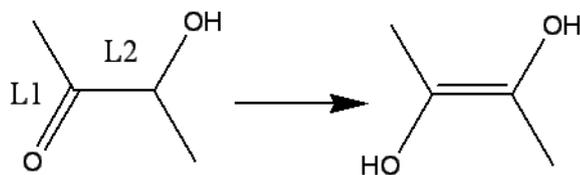
- i) i valori di energia conformazionale ( $\Delta E < 5$  Kcal);
- ii) la distanza interatomica tra il carbonio chetonico o il carbonio imminico dello ZBG e il centroide del CAP (distanze:  $< 5 \text{ \AA}$ ,  $5-7.5 \text{ \AA}$ ,  $7.5-10 \text{ \AA}$ ,  $> 10 \text{ \AA}$ );
- iii) angoli torsionali dello ZBG ( chetoesteri, chetoammidi, immine estere e immine ammidi)
- iv) legami idrogeno intramolecolari.

Poichè tutti gli inibitori cristallizzati negli enzimi HDAC mostrano una conformazione estesa della catena alchilica (*linker*, *Figura 8 e 9*), al fine di analizzare sia le preferenze conformazionali che la capacità dei nuovi composti di coordinare lo zinco, gli stereoisomeri energeticamente più stabili di ogni tautomero ST7727, ST7729, ST7732, ST7991, aventi una conformazione estesa della catena alchilica (distanza ZBG-CAP of  $> 7.5 \text{ \AA}$ ) sono stati selezionati per i calcoli DFT (effettuati dalla Dott.ssa Francesca Rondinelli).

### ***3.3.1b Analisi conformazionale dei nuovi inibitori caratterizzati da un nucleo benzodiazepinico (BDZ)***

I valori di pKa dei nuovi composti progettati RR89, EC332, EC226, EC229, RR67, RR96, RR6, RR84, RR94, EC233, RR62, RR11, EC190, RR1\_127, RR1\_125, RR1\_116, RR90, sono stati calcolati utilizzando il *software* ACD/pKa DB versione 12.00 (Advanced Chemistry Development Inc., Toronto, Canada) considerando tutte le possibili forme tautomeriche. Di conseguenza, le percentuali delle forme neutre/protonate sono state determinate considerando il valore di pH citoplasmatico (7.4) usando l'equazione di Henderson-Hasselbach. La solubilità e il LogD delle diverse forme tautomeriche è stata valutata a pH 7.4 considerando il valore in mol/l. Dato che ogni molecola presenta tre forme tautomeriche è stata condotta una ricerca bibliografica al fine di stabilire quale forma tautomerica utilizzare nei nostri calcoli mediante il server *SciFinder* (<https://scifinder.cas.org>). Inoltre sono state analizzate le strutture cristallografiche di tutti gli  $\alpha$ -idrossichetoni liberi (codici CDS: BUSFEA, COSRUY, DATSIB10, BUVFAA, CABBIQ, CERXOM, CEZTAD, CIJBAY, CIWBAM, CIJHAE, EBIFUQ, CUTGAZ, DISZIO, EYAWOQ), e di quelli in complesso con i metalli (codici CDS: AJARAE, BAQPAK, CEMRER, DIBZIX, DIBZOD), depositate nella *Cambridge Structural Database* (CSD) utilizzando il *software* *Conquest* 1.12 della CSDB (Cambridge Structural Database System).

Unendo tutti i risultati ottenuti è stato possibile ipotizzare che la tautomeria sia favorita dal legame a idrogeno intramolecolare e dalla planarità del sistema. In questo caso si verifica un allungamento del legame L1 ed un accorciamento del legame L2 nelle strutture cristallografiche esaminate (Figura 8).

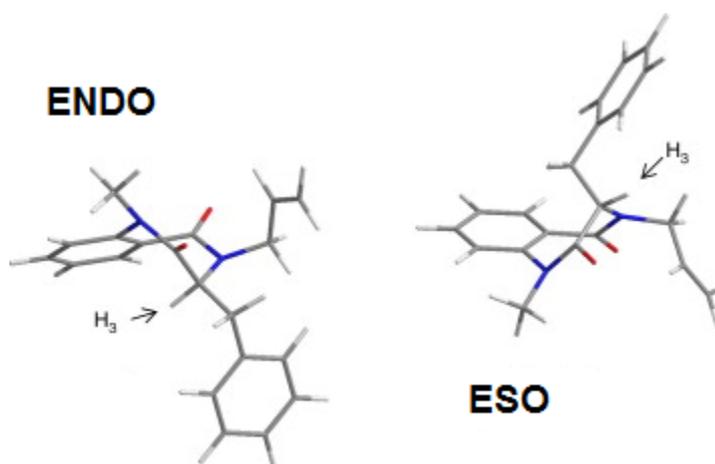


**Figura 8:** rappresentazione schematica della tautomeria del gruppo  $\alpha$ -idrossichetone.

Tale situazione viene indotta dalla presenza di basi o per effetto della temperatura<sup>29</sup>.

Pertanto i ligandi sono stati considerati nella forma idrossi-chetonica. Contemporaneamente è stato possibile stabilire la geometria di coordinazione con lo zinco, da utilizzare nei successivi calcoli di *docking*.

La ricerca bibliografica ha permesso di definire la conformazione del nucleo benzodiazepinico. Quest'ultimo è caratterizzato da sistema triciclico abbastanza rigido e adottano una forma ad ali di farfalla<sup>30</sup>. Lo spettro <sup>1</sup>H NMR di tutti i composti ha dimostrato chiaramente la presenza di due conformeri, in un rapporto 60:40, non separabili cristallograficamente. Ciò suggerisce che c'è un'inversione dell'anello troppo rapida perchè sia possibile una separazione a temperatura ambiente, ma lenta abbastanza da permettere la scoperta dei due conformeri attraverso l'analisi NMR. Attraverso l'esecuzione dei calcoli ab-initio con il modello chimico standard B3-LYP/6-31G<sup>31</sup>, sono state trovate due conformazioni, una *endo* e una *exo* (Figura 9), con una piccola differenza di energia, a seconda di come l'H3 si posiziona nell'anello a sette termini.



**Figura 9:** conformazione delle benzodiazepine. A destra è riportata la conformazione *exo*, mentre a sinistra quella *endo*.

Le molecole progettate sono state costruite secondo la procedura descritta nel paragrafo 3.3.1a.

### 3.3.2 Calcoli DFT

Calcoli a livello della Teoria del Funzionale della Densità. L'ottimizzazione a livello B3LYP/6-311++G\*\*/LanL2DZ<sup>32-34</sup> è stata effettuata su conformeri PM6 nella loro forma di base e connessa allo Zn<sup>2+</sup>. Il set di funzioni di base 6-311++G\*\* è stato scelto per gli atomi di C, H, N ed O in modo da trattare correttamente gli orbitali molecolari dei composti anionici mediante le funzioni diffuse. Lo pseudopotenziale e il set di funzioni di base LanL2DZ sono stati usati per il metallo. Tutti i minimi sono stati ottenuti senza imporre alcuna costrizione geometrica e l'analisi vibrazionale è stata realizzata allo stesso livello di teoria dell'ottimizzazione. L'analisi delle caratteristiche di legame di tutti i composti è stata realizzata mediante calcoli NBO<sup>35</sup> (Natural Bond Orbital). L'affinità del metallo allo ZBG (Zinc Binding Group) è stata stimata attraverso la valutazione dell'energia di legame (BE), secondo la seguente equazione:

$$BE = E_{\text{ione metallico}} + E_{\text{inibitore}} - E_{\text{complesso metallico}}$$

nella quale lo ione metallico Zn<sup>2+</sup> è stato considerato nel suo stato fondamentale. Le energie elettroniche del complesso metallico e dell'inibitore sono comprensive della correzione ZPE. La distribuzione delle cariche è stata elucidata mediante calcoli APT<sup>36</sup> (Atomic Polar Tensor), generalmente preferiti alle analisi di popolazione naturale e all'approccio di Mulliken per lo studio di composti organometallici<sup>37</sup>. Gaussian 09 è stato usato per realizzare tutti i calcoli DFT<sup>38</sup>.

### 3.3.3 Analisi Strutturale e Bioinformatica

Le sequenze di tutti gli istoni deacetilasi (HDACs) sono state scaricate da *UniProtKB/Swiss-Protein Data Bank* (<http://www.uniprot.org>; codici HDAC1: Q13547; HDAC2: Q92769; HDAC3: O15379; HDAC4: P56524; HDAC5: Q9UQL6; HDAC6: Q9UBN7; HDAC7: Q8WUI4; HDAC8: Q9BY41; HDAC9: Q9UKV0) e l'allineamento è stato eseguito utilizzando il server *PROMALS3D* (<http://prodata.swmed.edu/promals3d/promals3d.php>).<sup>39</sup> Tutte le strutture sperimentalmente determinate degli HDAC sono state scaricate dalla banca dati *Protein Data Bank* (PDB, <http://www.rcsb.org/pdb/>). HDAC1 (PDB ID: 4BKX), HDAC2

(PDB ID: 3MAX), HDAC3 (PDB ID: 4A69), HDAC4 (PDB IDs: 2VQW, 2VQV, 2VQJ, 2VQM, 2VQO, 2VQQ, 4CBT, 4CBY), HDAC7 (PDB IDs: 3C0Y, 3C0Z, 3C10), and HDAC8 (PDB IDs: 3EW8, 3EWF, 3EZP, 3EZT, 3F06, 3F07, 3F0R, 2V5W, 2V5X, 1W22, 1VKG, 1T69, 1T67, 1T64, 3SFF, 3SFH, 3MZ3, 3MZ4, 3MZ6, 3MZ7, 3RQD), Gli idrogeni sono stati aggiunti a tutte le strutture cristallografiche considerando un pH di 7.2. Tutte le strutture sperimentalmente determinate degli HDAC sono state analizzate utilizzando il modulo *Homology di Insight 2005* (Accelrys, San Diego, CA). In particolare, le strutture cristallografiche sono state sovrapposte per gli aminoacidi del sito catalitico, le sequenze delle strutture sperimentalmente determinate sono state estratte dalle strutture cristallografiche e poi, allineate utilizzando l'algoritmo *Multiple\_Alignment (Homology module, Accelrys, San Diego)*. Le strutture secondarie delle PDB sono state analizzate utilizzando il comando *Secondary Classification di ProStat*, nel modulo *Homology di Insight 2005*. Al fine di determinare il modello farmacoforico degli inibitori HDAC, è stata analizzata la modalità di legame di tutti gli inibitori co-cristallizzati con gli HDAC considerando gli aminoacidi e le molecole d'acqua aventi almeno un atomo entro un raggio di 5Å da un qualsiasi atomo del ligando utilizzando i moduli *Biopolymer e Homology di Insight 2005* (Accelrys Software Inc., San Diego, CA). Nell'analisi è stata valutata la coordinazione dello ZBG, i torsionali del linker, l'accomodamento, le interazioni e le posizioni dei gruppi carbonilici del CAP.

### ***3.3.4 Studi di Pharmacophore Fitting***

Le strutture cristallografiche degli HDAC sono state scaricate dalla *Protein Data Bank* (PDB; <http://www.rcsb.org/pdb>) e sono state analizzate come descritto nel paragrafo 3.3.3.

Il modello farmacoforico è stato suddiviso in tre parti: il CAP; il *linker*; lo ZBG e il sostituente ingombrante nella tasca idrofobica. Per ogni struttura cristallografica sono stati esaminati l'accomodamento e le interazioni del CAP, i torsionali della catena alchilica, e la coordinazione dello ZBG. Per lo studio delle interazioni del CAP è fondamentale osservare il posizionamento dei gruppi carbonilici, in quanto determinano l'accomodamento dei diversi sostituenti nelle tasche di superficie dell'enzima. Per questo motivo sono state selezionate sette strutture cristallografiche che presentano dieci diverse posizioni del gruppo carbonilico dell'inibitore. In particolare, sono state considerate le seguenti strutture cristallografiche: PDB IDs: 3EWF, 2V5X, 1W22, 1VKG, 1T67 appartenenti all'HDAC8, la PDB ID: 2VQM della famiglia dell'HDAC4, e la PDB ID: 3C10 appartenente all'HDAC7. Per valutare inoltre i differenti accomodamenti degli inibitori, nei canali dell'enzima, sono state selezionate altre

due strutture cristallografiche dell'HDAC8: la PDB 2V5X, in cui il ligando occupa le tasche 1 e 3, e la PDB 3F07 che presenta occupazione della tasca 2. L'analisi della posizione dello ZBG nel sito catalitico è stata effettuata così come descritto nel paragrafo 3.4.1a.

Partendo dai risultati dell'analisi delle proprietà conformazionali ed elettroniche i conformeri sono stati sovrapposti alla struttura cristallografica PDB 2V5X, isoforma appartenente alla famiglia degli HDAC8 di classe I, al fine di valutare l'accomodamento nel sito catalitico degli HDAC. Tramite il comando *Transform Superimpose* i composti sono stati sovrapposti a livello delle porzioni strutturali simili: i coordinanti dello  $Zn^{2+}$ ; il carbonio carbonilico; il primo metilene del *linker*; il carbonio chirale che congiunge il *linker* al CAP.

### ***3.3.5a Procedura di Docking dei nuovi inibitori caratterizzati da un nuovo gruppo coordinante lo zinco (ZBG)***

Gli studi di docking sono stati condotti su ST7727, ST7728, ST7729, ST7730, ST7732, ST7733, ST7990, ST7991 in complesso con l'HDAC2 (PDB ID: 3MAX) ed HDAC3 (PDB ID: 4A69). Per raggiungere questo obiettivo, è stata utilizzata una metodologia di *docking* (Affinity, SA\_Docking; Insight2005, Accelrys, San Diego, CA) che considera tutto il sistema flessibile (i.e., ligando, proteina e molecole d'acqua). I potenziali atomici sono stati assegnati utilizzando il campo di forza ESFF<sup>40</sup>, che contiene i parametri dello zinco; mentre le cariche parziali sono state assegnate utilizzando il campo di forza CVFF<sup>24</sup>.

Nel protocollo di dinamica del docking, tutto il sistema è stato perturbato mediante le procedure *Monte Carlo* e *Simulated Annealing*, tuttavia la procedura di *docking* dinamico richiede una struttura di partenza ragionevole. Al fine di definire la conformazione di partenza dello ZBG, sono stati considerati tutti i conformeri di ST7727, ST7729, ST7732 e ST7991 in complesso con il catione  $Zn^{2+}$  derivanti dai calcoli di DFT, sono stati inseriti nel sito attivo delle strutture sperimentalmente determinate degli HDAC umani e sovrapposti all'inibitore dell'HDAC4 (PDB ID:2VQV) considerando gli atomi coordinanti lo zinco e l'atomo di zinco stesso. Le conformazioni dello ZBG che non mostravano ingombri sterici con gli aminoacidi del sito catalitico, sono state selezionate come possibili conformazioni di partenza dello ZBG per i calcoli di docking. Di conseguenza, i conformeri ottenuti dai calcoli PM6 di ST7727, ST7728, ST7729, ST7730, ST7732, , ST7733, ST7990 e ST7991 caratterizzati dalla conformazione così selezionata dello ZBG e da una conformazione estesa del linker (distanza ZBG-CAP > 7.5 Å), sono stati sostituiti nel sito attivo degli enzimi sulla base dell'orientazione degli inibitori co-cristallizzati.

Le molecole d'acqua sono state mantenute in tutte le strutture cristallografiche considerate con l'eccezione di quelle che si sovrapponevano ai ligandi.

I calcoli di *docking* flessibili sono stati effettuati utilizzando il modulo *Affinity di Insight 2005*, mediante la procedura *SA\_Docking*<sup>41</sup>, e utilizzando il metodo *Cell\_Multipole* per le interazioni di non legame<sup>42</sup>.

Il dominio di legame è stato definito combinando un *subset* attorno al ligando dal raggio di 10 Å che considera tutti gli aminoacidi della proteina e le molecole d'acqua comprese in tale raggio, più altre due regioni di legame (Asp23-Arg39; Ser83-Phe108; numerazione dell' HDAC2 umana, Asp12-Arg28, Arg72-Phe97 numerazione dell' HDAC3 umana). Quest'area di legame è stata ampliata al fine di consentire il cambiamento conformazionale della superficie della proteina indotto dal CAP. Tutti gli atomi inclusi nel dominio di legame sono liberi di muoversi durante il corso dei calcoli di *docking*, al fine di evitare risultati non realistici, agli ioni zinco e potassio è stata applicata una costante di forza che li vincolasse alla loro posizione iniziale di 100 (Kcal/mol)/ Å; inoltre una restrizione di 2.5 Å è stata applicata tra gli atomi degli aminoacidi di coordinazione e gli ioni utilizzando una forza di 100 (Kcal/mol)/ Å. Al fine di ottenere risultati più rappresentativi, la procedura di *docking* è stata incrementata applicando per ogni complesso di partenza o una costante di forza che li vincolasse alla loro posizione iniziale di 30 Kcal/Å, oppure una restrizione di 2.3 Å tra gli atomi coordinanti del composto e lo zinco catalitico, utilizzando una costante di forza di 100 (Kcal/mol)/ Å.

Sono state generate casualmente 20 strutture utilizzando la procedura combinata Monte Carlo/minimizzazione con un criterio di tolleranza energetico di 10<sup>6</sup> Kcal/mol per assicurare un'ampia variabilità delle strutture da minimizzare (2500 iterazioni; ε=1). Durante questa fase i ligandi sono stati mossi attraverso una combinazione casuale di movimenti traslazionali e rotazionali e variando il valore degli angoli torsionali (opzione *Flexible\_Ligand*), considerando tutti i legami ruotabili, al fine di campionare sia lo spazio conformazionale del ligando che la sua orientazione rispetto allo zinco (MxRChange = 3 Å; MxAngChange = 180°). Al fine di evitare forti divergenze nelle energie di van der Waals (vdW) e Coulomb, sono stati scalati i termini relativi alle forze di vdW e di Coulomb di un fattore pari a 0.1.

Dopo la fase di minimizzazione (*Conjugate Gradient*; 2500 interazioni; ε=1), sono stati applicati sia un criterio di controllo di similitudine delle strutture (RMS tolerance = 0.3 Kcal/Å), sia un criterio di energia (Energy test; T = 50 K), per selezionare 20 soluzioni accettabili.

I complessi selezionati sono stati prima minimizzati e poi sottoposti a 50 cicli di *Simulated Annealing*, ciascuno di 100 fs. Inizialmente è stata applicata una temperatura di 500 K in modo tale da superare la barriera torsionale, e consentire al ligando e alla proteina di non essere conformazionalmente costretti. Nel corso di un ciclo di *Simulated Annealing*, la temperatura del sistema viene diminuita linearmente da 500K a 300K in 5 ps (velocità di integrazione Verlet; time step di 1,0 fs), e contemporaneamente le forze di vdW e di Coulomb sono state riportate dal 10 al 100% del loro valore. In questa procedura di *Simulated Annealing*, quindi, la temperatura viene variata in intervalli di tempo, passando dalla temperatura iniziale a quella finale. Queste variazioni di temperatura sono da ricondurre a modificazioni dell'energia cinetica della molecola e delle velocità degli atomi. Nella prima fase, le velocità iniziali sono state generate casualmente dalla distribuzione di Boltzman, in base alla temperatura utilizzata, mentre nelle fasi successive sono state generate attraverso la funzione *Dynamics Restart Data*. Alla fine dei calcoli di dinamica molecolare ciascun complesso ottenuto è stato sottoposto ad una minimizzazione finale di  $10^4$  cicli.

Dopo questa procedura i risultati dei *docking* sono stati classificati mediante la loro energia conformazionale. Il risultante complesso di *docking* è stato selezionato eliminando i risultati che presentassero una conformazione ripiegata del *linker* dell'inibitore o una conformazione *cis* dell'ammide e prendendo in considerazione le precedenti analisi riportate sullo ZBG e l'orientazione degli inibitori cristallizzati.

I risultati dei calcoli di *docking* sono stati analizzati considerando: i) l'energia di interazione ligando-enzima di ciascun complesso che è stata valutata calcolando l'energia totale di interazione tra ligando e il sito di legame (tutti i residui entro i 10 Å da ogni atomo del ligando), utilizzando il comando *Evaluate* nel modulo *Docking di Insight2005* (energia di vdW ed elettrostatiche; CUT\_OFF=10); ii) la differenza di energia tra la conformazione del ligando legato alla proteina (1SCF, PM6, MOPAC2012) e la conformazione di minimo globale (PM6, MOPAC 2012); iii) le distanze tra i coordinanti e lo zinco catalitico; iv) la geometria di coordinazione dello ZBG; v) la planarità del metile ed etile rispetto allo ZBG, considerando sia i risultati ottenuti dai calcoli di *docking* che quelli dei calcoli DFT.

### ***3.3.5b Procedura di Docking dei nuovi inibitori caratterizzati da un nucleo benzodiazepinico***

Sono stati eseguiti studi di *docking* sulle seguenti isoforme HDAC: HDAC2 (PDB ID: 3MAX), HDAC3 (PDB ID: 4A69), HDAC4 (PDB ID: 2VQJ), HDAC7 (PDB ID: 3COY) e HDAC8 (PDB ID: 3F07), utilizzando la metodologia di *docking* descritta nel paragrafo 3.3.5a; in questo caso però i ligandi sono stati collocati nel sito catalitico prendendo come riferimento la struttura cristallografica del complesso substrato/HDAC8 (PDB ID: 2V5W). L'inserimento del ligando nel sito catalitico degli HDAC è stato effettuato, previa rimozione degli inibitori con il comando *Unmerge*, utilizzando il comando *Merge* nel modulo *Biopolymer* di *Insight2005* (Accelrys, San Diego). Le molecole di acqua presenti nelle strutture cristallografiche sono state mantenute, eccetto quelle stericamente sovrapposte al ligando. Nel caso dell'HDAC8, dato che nella 3F07 erano presenti solo cinque molecole d'acqua, sono state introdotte quelle della struttura cristallografica PDB 3SFF (PDB tra gli HDAC8 con migliore risoluzione), con il comando *Merge* del modulo *Biopolymer* di *Insight2005* (Accelrys, San Diego).

L'area del sito di legame è stata definita combinando un *subset* attorno al ligando, caratterizzato da: i) tutti i residui e le molecole d'acqua che presentano almeno un atomo che ricade in un raggio di 10 Å da un qualsiasi atomo del ligando stesso, ii) due regioni di legame adiacenti al sito catalitico. Quest' ultime sono costituite dalle seguenti sequenze amminoacidiche: nell'HDAC2 23-39 e 83-108; nell'HDAC3 12-28 e 72-97; nell'HDAC4 83-118 e 17-37; nell'HDAC7 527-547 e 592-631; nell'HDAC8 23-37 e 81-105. Quest'area è stata inclusa per permettere al dominio dello ZBG eventuali cambiamenti conformazionali.

I risultati ottenuti sono stati selezionati sulla base della loro energia conformazionale, dei valori degli angoli torsionali del *linker* del ligando, della conformazione del CAP e della geometria di coordinazione con lo zinco. Quella che presentava il migliore compromesso tra questi parametri è stata selezionata come la più probabile ipotesi di modalità di legame. Sulla base dei risultati del *pharmacophore fitting* sono stati scartati quei complessi che presentavano: i) una posizione dello ZBG incompatibile con il legame allo zinco catalitico, ii) il secondo e quarto torsionale in cis del *linker* e iii) una conformazione non distesa del *linker*.

## ***3.4 Progettazione di inibitori caratterizzati da un nuovo gruppo coordinante lo zinco(ZBG)***

### ***3.4.1 Risultati e Discussione***

#### ***3.4.1a Analisi Strutturale e Bioinformatica***

Lo studio del dominio catalitico degli HDAC appartenenti alla classe I e II è stato effettuato mediante l'analisi strutturale e bioinformatica delle strutture HDAC sperimentalmente determinate (in complesso, o meno, con degli inibitori) presenti nella banca dati *Brookhaven Protein Data Bank*.

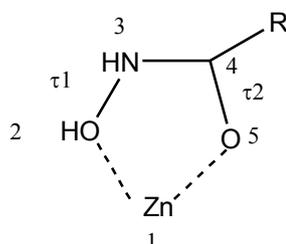
I risultati ottenuti hanno permesso di identificare diverse regioni strutturali, caratterizzate dalla presenza di specifici residui amminoacidici, responsabili i) del funzionamento catalitico dell'enzima e ii) del riconoscimento dei *partners* proteici facenti parte del macro-complesso che regola di fatto la trascrizione in vivo. Inoltre, sono state analizzate e classificate le diverse posizioni occupate dai coordinanti dello zinco degli inibitori co-cristallizzati e le loro geometrie di coordinazione.

Particolare attenzione è stata posta: i) sul canale d'uscita dal quale si ipotizza che fuoriesca il gruppo acetato, ovvero uno dei prodotti della reazione di catalisi (HDAC2: Tyr29, Met35, Phe114, Leu144; HDAC3: Tyr18, Met24, Phe103, Leu133), questa porzione è fondamentale per l'accomodamento del gruppo metile dei nostri composti; ii) sugli aminoacidi coinvolti nell'interazione con i cofattori responsabili della formazione dei macrocomplessi e dell'attivazione enzimatica,

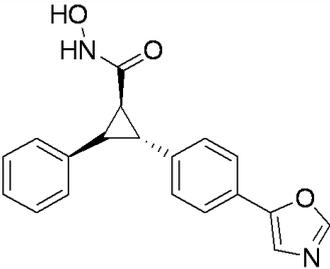
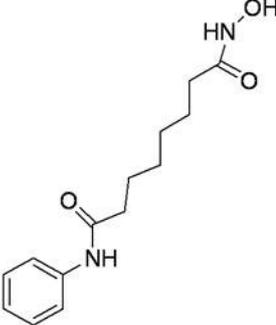
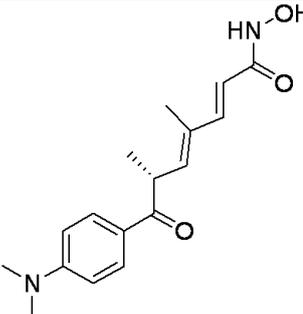
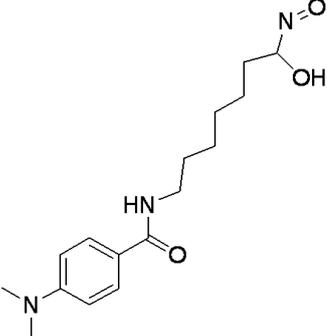
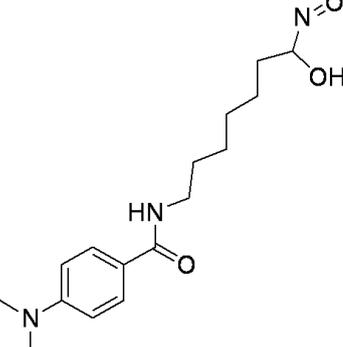
Per quanto riguarda l'HDAC3 l'interazione con il cofattore NCoR è data dagli aminoacidi: Val47, Lys29, Asp12, Asn15, Glu330, Glu327 (HDAC2: Ile58, Arg60, Asp23, Asn26, Glu340, Asp337); mentre quella con l'IP4 è data dagli aminoacidi Gly21, Arg265, Lys25, His17, Arg301 (HDAC2: Gly32, Arg275, Lys36, Tyr28, Arg311).

In primo luogo, quindi, è stata valutata la modalità con cui i coordinanti degli inibitori co-cristallizzati coordinano lo zinco catalitico e sono stati calcolati i valori degli angoli torsionali di seguito riportati (Tabella 2).

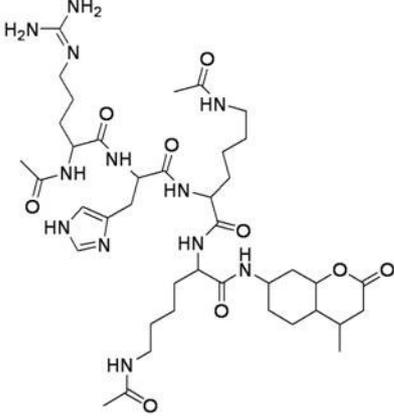
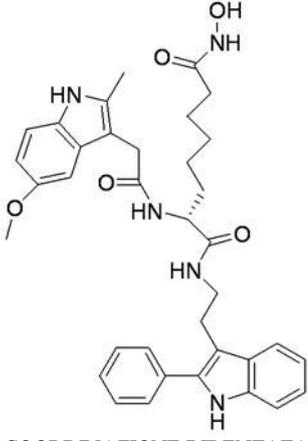
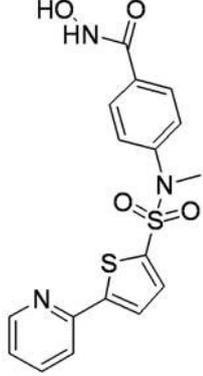
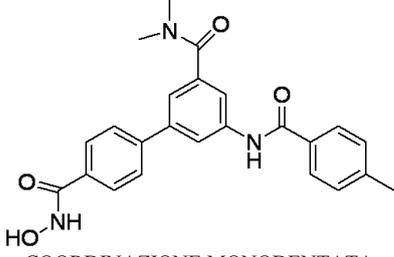
**Tabella 2:** valori degli angoli torsionali ( $\tau_1$  (1-2-3-4);  $\tau_2$  (1-5-4-3)) delle strutture cristallografiche

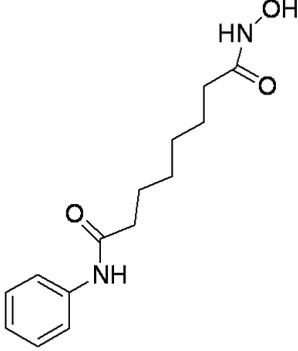
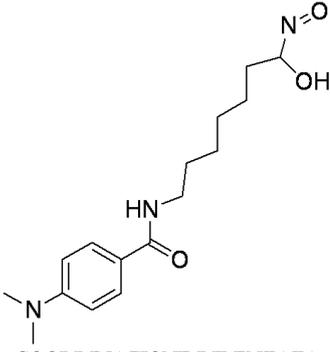
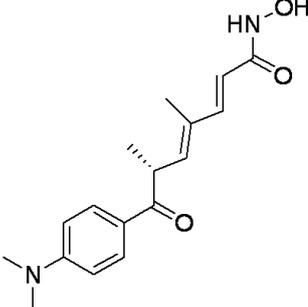
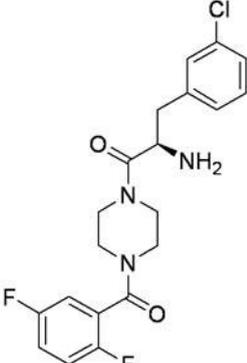
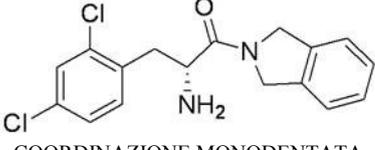
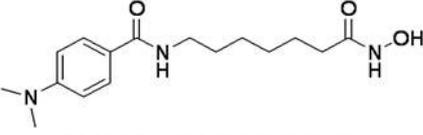


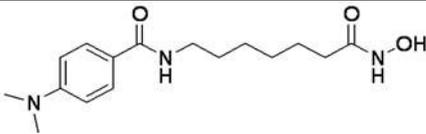
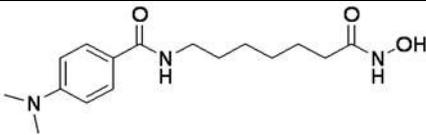
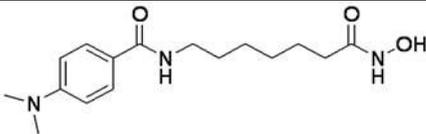
PROTEINA	LIGANDO	$\tau_1$	$\tau_2$
<b>HDAC2</b> <b>PDB ID = 3MAX</b> CONF. CHIUSA	 COORDINAZIONE BIDENTATA	-77.57	-56.67
<b>HDAC 4</b> <b>PDB ID: 2VQM</b> CONF. APERTA	 COORDINAZIONE MONODENTATA	H2O	-76.80
<b>HDAC 4</b> <b>PDB ID: 2VQV_MA</b> (MUTANTE H976Y) CONF. APERTA	 COORDINAZIONE MONODENTATA	39.21	-63.30
<b>HDAC4</b> <b>PDB ID: 4CBT_MA</b> CONF. CHIUSA		32.0	-26.99

	COORDINAZIONE BIDENTATA		
<p><b>HDAC4</b>  <b>PDB ID: 4CBY_MA</b>            CONF. CHIUSA</p>	 <p>COORDINAZIONE BIDENTATA</p>	19.54	-26.44
<p><b>HDAC 7</b>  <b>PDB ID: 3C0Z_MA</b>            CONF. CHIUSA</p>	 <p>COORDINAZIONE MONODENTATA</p>	-69.38	-43.58
<p><b>HDAC 7</b>  <b>PDB ID: 3C10_MA</b>            CONF. CHIUSA</p>	 <p>COORDINAZIONE MONODENTATA</p>	-47.15	-45.58
<p><b>HDAC8</b>  <b>PDB ID: 3EW8</b>            (MUTANTE D101L)            CONF. CHIUSA</p>	 <p>COORDINAZIONE MONODENTATA</p>	49.58	-56.86
<p><b>HDAC8</b>  <b>PDB ID: 3EZF_MA</b>            (MUTANTE D101N)            CONF. CHIUSA</p>		21.70	-23.18

	COORDINAZIONE BIDENTATA		
<p><b>HDAC8</b>  <b>PDB ID: 3EZT_MA</b>  (MUTANTE D101E)  CONF. CHIUSA</p>	<p>COORDINAZIONE BIDENTATA</p>	21.91	-22.86
<p><b>HDAC8</b>  <b>PDB ID: 3F06_MA</b>  (MUTANTE D101A)  CONF. CHIUSA</p>	<p>COORDINAZIONE BIDENTATA</p>	14.28	16.26
<p><b>HDAC8</b>  <b>PDB ID: 3F07_MA</b>  CONF. CHIUSA</p>	<p>COORDINAZIONE BIDENTATA</p>	53.69	-60.54
<p><b>HDAC8</b>  <b>PDB ID: 3F0R_MA</b>  CONF. CHIUSA</p>	<p>COORDINAZIONE BIDENTATA</p>	42.76	-43.25

<p><b>HDAC8</b>  <b>PDB ID: 2V5W_MA</b>  CONF. CHIUSA</p>	 <p>COORDINAZIONE MONODENTATA</p>	<p>H20</p>	<p>-84.70</p>
<p><b>HDAC8</b>  <b>PDB ID: 2V5X_MA</b>  CONF. CHIUSA</p>	 <p>COORDINAZIONE BIDENTATA</p>	<p>29.17</p>	<p>-27.89</p>
<p><b>HDAC8</b>  <b>PDB ID: 1W22_MA</b>  CONF. CHIUSA</p>	 <p>COORDINAZIONE BIDENTATA</p>	<p>38.37</p>	<p>-44.73</p>
<p><b>HDAC8</b>  <b>PDB ID: 1VKG_MA</b>  CONF. APERTA</p>	 <p>COORDINAZIONE MONODENTATA</p>	<p>-13.49</p>	<p>-37.05</p>

<p><b>HDAC8</b> <b>PDB ID:1T69</b> CONF. CHIUSA</p>	 <p>COORDINAZIONE BIDENTATA</p>	-36.68	-9.42
<p><b>HDAC8</b> <b>PDB ID:1T67</b> CONF. CHIUSA</p>	 <p>COORDINAZIONE BIDENTATA</p>	8.41	-21.50
<p><b>HDAC8</b> <b>PDB ID:1T64_MA</b> CONF. CHIUSA</p>	 <p>COORDINAZIONE BIDENTATA</p>	5.05	-38.09
<p><b>HDAC8</b> <b>PDB ID:3SFF_MA</b> CONF. CHIUSA</p>	 <p>COORDINAZIONE MONODENTATA</p>	38.13	-47.37
<p><b>HDAC8</b> <b>PDB ID:3SFH_MA</b> CONF. CHIUSA</p>	 <p>COORDINAZIONE MONODENTATA</p>	33.02	-54.45
<p><b>HDAC8</b> <b>PDB ID: 3MZ3_MA</b> CONF. CHIUSA</p>	 <p>COORDINAZIONE BIDENTATA</p>	18.07	-15.97

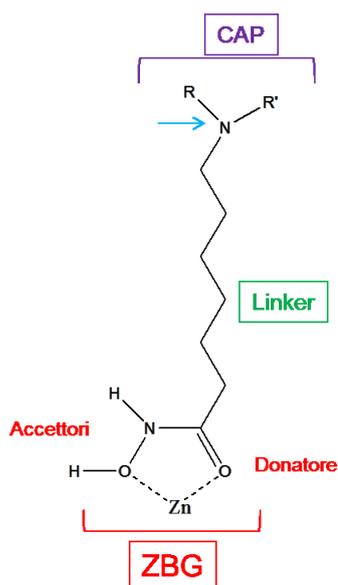
<p><b>HDAC8</b>  <b>PDB ID: 3MZ4_MA</b>            CONF. CHIUSA</p>	 <p>COORDINAZIONE BIDENTATA</p>	33.57	-23.25
<p><b>HDAC8</b>  <b>PDB ID: 3MZ6</b>            (MUTANTE D101L)            CONF. CHIUSA</p>	 <p>COORDINAZIONE BIDENTATA</p>	44.92	-27.52
<p><b>HDAC8</b>  <b>PDB ID: 3MZ7</b>            (MUTANTE D101L)            CONF. CHIUSA</p>	 <p>COORDINAZIONE BIDENTATA</p>	36.71	-21.47

### 3.4.1b Basi Razionali della Progettazione

Sulla base dei risultati dell'analisi strutturale e bioinformatica sono stati progettati nuovi inibitori che, pur contenendo le caratteristiche strutturali fondamentali degli inibitori HDAC noti (*Figura 10*), presentassero uno ZBG differente.

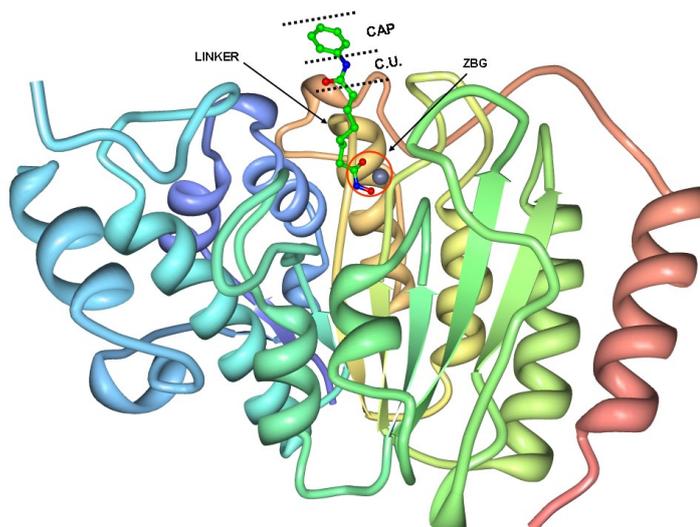
Le caratteristiche strutturali fondamentali degli inibitori finora sviluppati sono:

- uno ZBG che coordini lo zinco e formi legami ad idrogeno con gli amminoacidi circostanti;
- una catena alchilica (*linker*), con una lunghezza che va dai 6.6 ai 7.3 Å, che congiunge lo ZBG al CAP
- un gruppo donatore/accettore di legami ad idrogeno chiamato *connection unit* (CU), interposto tra il CAP ed il *linker*;
- un CAP che interagisca con la superficie della proteina,



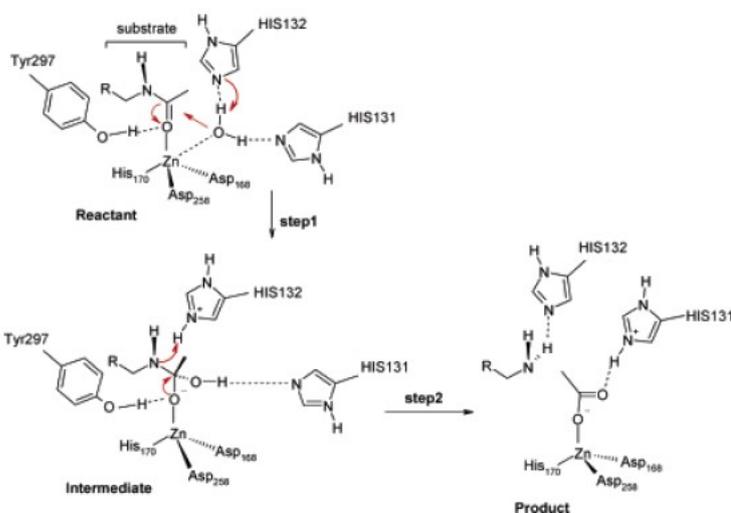
**Figura 10** : Rappresentazione schematica delle porzioni strutturali degli inibitori HDAC

L'acido idrossamico rappresenta lo ZBG maggiormente presente negli inibitori HDAC sviluppati sinora ed è l'unico ZBG presente in quelli utilizzati in terapia. Tale gruppo funzionale, tuttavia, mostra diversi inconvenienti dal punto di vista farmacocinetico, che ne determinano una breve emivita in vivo. È stata, quindi, valutata la capacità di diversi gruppi chelanti lo zinco di fungere da ZBG di un inibitore ed inserirsi all'interno del sito catalitico di un enzima HDAC. Modificando lo ZBG si determina un cambiamento sia nella modalità di interazione con il metallo che nella rete di legami idrogeno che lo ZBG stesso instaura con gli amminoacidi presenti nel sito catalitico. Ciò, a sua volta, determina: i) un diverso accomodamento dei sostituenti eventualmente presenti all'estremità dello ZBG e ii) una differente proiezione della catena idrofobica, del CU e del CAP verso la superficie dell'enzima (*Figura 11*).



**Figura 11:** Struttura cristallografica dell'HDAC8 umano in complesso con il SAHA (DOI:10.2210/pdb1t69/pdb; RCSB PDB Protein Workshop 3.9). La proteina e l'inibitore sono mostrati, rispettivamente, in ball & stick e sticks. Lo zinc binding group (ZBG), il linker, la connecting unit (C.U.) ed il CAP sono indicati. Gli eteroatomi del ligando sono colorati sulla base dei tipi atomici (N: blu ; O: rosso e C:verde).

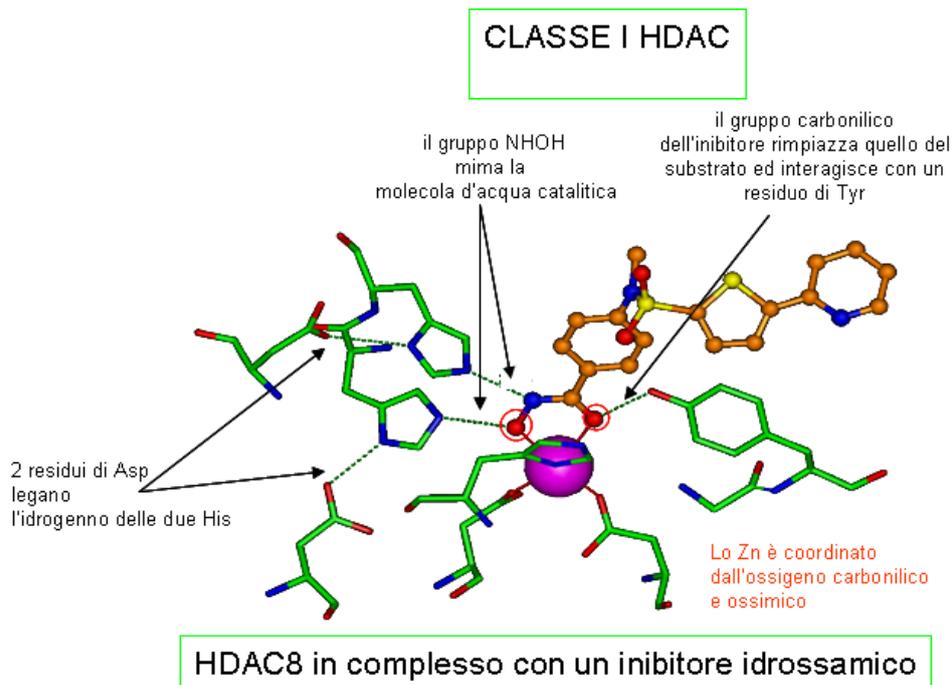
Al fine di progettare nuovi ZBG è stato considerato anche il meccanismo di idrolisi enzimatica ipotizzato per la deacetilazione della lisina da parte degli HDAC (Figura 12)



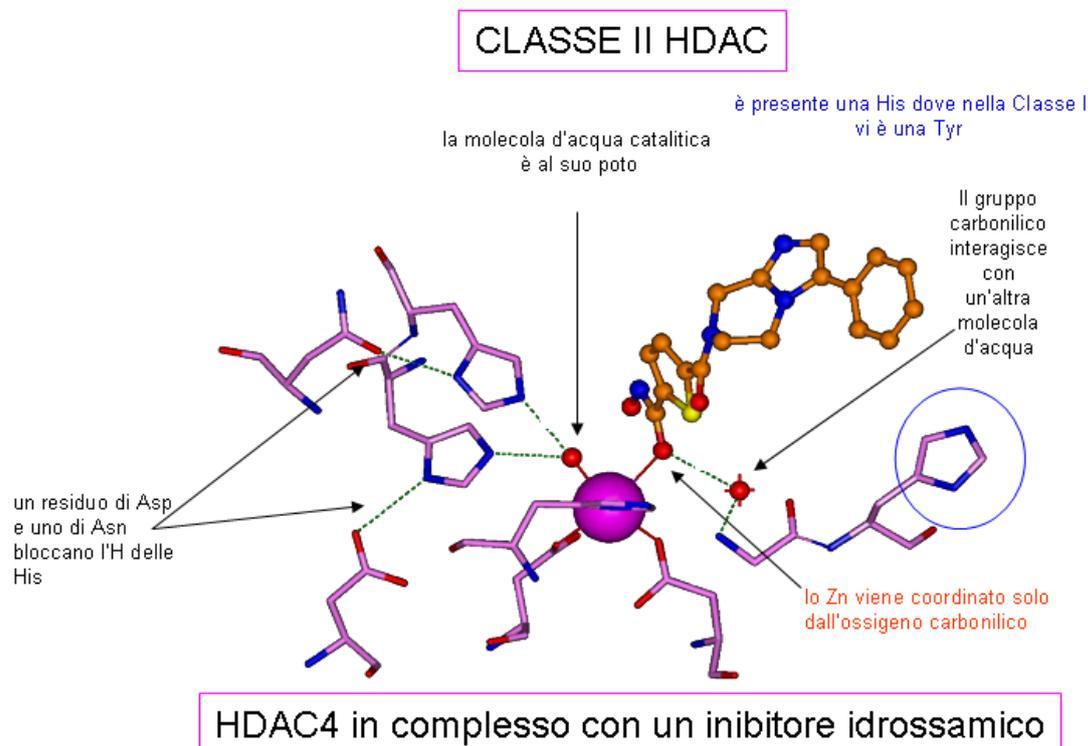
Zhang et al JACS 2006, 128, 4530-4531.

**Figura 12:** Meccanismo ipotizzato di idrolisi della lisina acetilata da parte degli enzimi HDAC di classe I.

Su queste basi, l'analisi bioinformatica e strutturale (vedere parte sperimentale) ha consentito di sfruttare i dati attualmente disponibili sulle interazioni instaurate dagli inibitori HDAC, sia nella classe I (Figura 13) che nella classe II (Figura 14).



**Figura 13:** Rappresentazione delle principali interazioni fra il sito catalitico dell'HDAC8 e lo ZBG di un inibitore idrossamico. Gli eteroatomi del ligando sono colorati sulla base dei tipi atomici ( N: blu ; O: rosso; S: giallo e C: arancione).

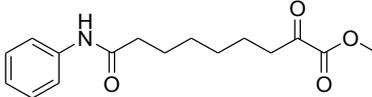
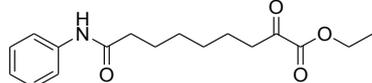
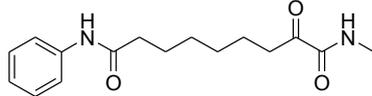
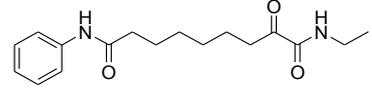
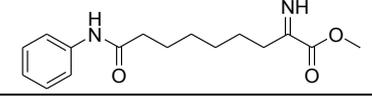
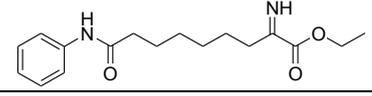
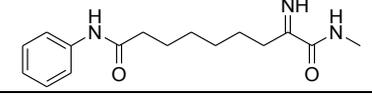
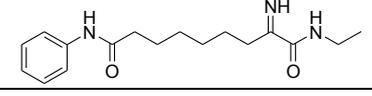


**Figura14:** Rappresentazione delle principali interazioni fra il sito catalitico dell'HDAC4 e lo ZBG di un inibitore idrossamico. Gli eteroatomi del ligando sono colorati sulla base dei tipi atomici ( N: blu ; O: rosso; S: giallo e C: arancione).

La progettazione è stata effettuata seguendo il concetto di “intrappolare” l’enzima in una specifica conformazione, in modo da inibire sia l’attività catalitica che le interazioni proteina-

proteina responsabili della formazione del macro-complesso proteico funzionale contenente gli HDAC. Inoltre, i nuovi inibitori (*Tabella 3*) sono stati progettati con l'intento di legare selettivamente l'isoforma HDAC3, il principale componente del macro-complesso proteico responsabile della repressione genica. Pertanto, oltre a modificare lo ZBG, è stato aggiunto un ulteriore sostituente potenzialmente in grado di occupare la tasca di uscita dell'acetato (paragrafo 3.4a), la quale appare particolarmente accessibile nell'isoforma HDAC3.

**Tabella 3** : strutture dei nuovi inibitori progettati.

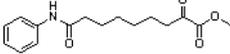
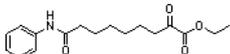
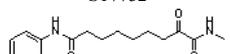
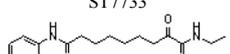
Nome strutture	Strutture
ST7991	
ST7990	
ST7732	
ST7733	
ST7727	
ST7728	
ST7729	
ST7730	

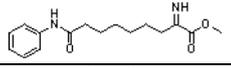
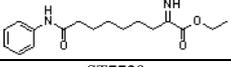
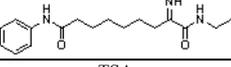
### 3.4.1c Razionalizzazione delle Relazioni Struttura-Attività (SARs) dei nuovi inibitori progettati

Il set di ligandi progettati; è stato sintetizzato dall'Industria Farmaceutica Sigma-Tau (Sigma-Tau, S.p.A., Pomezia, Italia) in collaborazione con l'Università degli Studi di Siena (Prof. Maurizio Taddei) e l'Università degli Studi di Salerno (Dott.ssa Manuela Rodriguez) ed è stato testato su diverse isoforme di HDAC isolate dal gruppo dal Dott. Ralph Mazitschek dell'Harvard Medical School e dalla Reaction Biology Corporation, mentre i saggi cellulari sono stati eseguiti dal gruppo del Prof. Antonello Petrella, della Dott.ssa Valentina Bizzarro del Dipartimento di Farmacia dell'Università degli Studi di Salerno.

L'attività dei nostri composti sulle diverse isoforme HDAC (Tabella 4) è stata testata utilizzando diversi substrati. In particolare, i dati riportati in tabella sono stati evidenziati a seconda del tipo di substrato utilizzato: i) in giallo vi sono i risultati ottenuti utilizzando il peptide fluorogenico corrispondente alla porzione 379-382 di p53 RHKK(Ac); ii) in rosso vi sono i dati per i quali è stato utilizzato il substrato specifico per la classe IIa Boc-Lys (trifluoroacetil)-AMC; iii) in magenta vi sono i risultati ottenuti utilizzando il substrato per l'HDAC8, il peptide fluorogenico corrispondente alla porzione 379-382 di p53 RHK(Ac)K(Ac). Tali saggi sono stati condotti dalla *Reaction Biology*. I valori in grassetto rappresentano i risultati ottenuti dal Dott. Ralph Mazitschek dell'Harvard Medical School, che ha utilizzato differenti substrati rispetto a quelli della *Reaction Biology*; il substrato tripeptidico acetilato per gli HDAC 1, 2, 3, e 6; ed il substrato tripeptidico trifluoroacetil-lisina per gli HDAC 4, 5, 7, 8 e 9. La sigla NA (non attiva), riportata in tabella 4, indica l'assenza di un'attività inibitoria che sia compatibile con una curva di IC<sub>50</sub>.

**Tabella 4:** valori di inibizione sulle diverse isoforme HDAC dei nuovi composti progettati

Structure	Class I Kd (µM)				Class IIa Kd (µM)				Class IIb Kd (µM)
	HDAC 1	HDAC 2	HDAC 3	HDAC 8	HDAC 4	HDAC 5	HDAC 7	HDAC 9	HDAC 6
 ST7991	NA	NA	0.15	NA	NA	NA	6.42	NA	9.85
 ST7990	0.06 0.07	0.06 1.45	0.03 0.04	NA NA	NA NA	NA 0.81	NA 4.82	NA NA	NA 16.50
 ST7732	0.44	0.86	0.30	80.65	1.89	0.57	2.10	3.54	NA
 ST7733	0.06 0.30	0.07 0.57	0.03 0.14	NA 25.46	NA 0.78	NA 0.45	NA 0.43	NA 3.22	NA NA

 ST7727	7.48	NA	1.23	NA	15.57	4.79	13.41	22.24	NA
 ST7728	0.32	1.50	0.16	NA	9.05	1.87	7.80	20.42	NA
 ST7729	5.29	19.77	3.66	NA	23.99	6.46	4.53	NA	NA
 ST7730	NA	NA	70.81	NA	25.22	NA	NA	NA	NA
TSA	0.001	0.002	0.003	0.11	0.05	0.003	0.03	0.01	0.0002

La sigla NA sta per non attiva

Va sottolineato che da un'attenta analisi dei risultati sperimentali forniti dalla ditta Reaction Biology, è emerso che i dati di attività relativi alla classe IIa sono inaffidabili in quanto il substrato utilizzato in questi saggi è sempre quello della classe I, e, pertanto, l'attività rilevata è associata non all'inibizione della classe IIa ma all'inibizione dell'HDAC3, con il quale questa classe forma complessi multimerici in grado di regolare l'attività catalitica, e che è sempre presente, anche nei saggi su riportati. Ciò è confermato dal fatto che, sulla base dei saggi effettuati dal Dott. Ralph Mazitschek dell'Harvard Medical School, che ha utilizzato il substrato specifico per gli enzimi di classe IIa, tutti i composti sono risultati inattivi su tale classe di HDAC e, quindi, essi sono risultati selettivi per la classe I. Infine, va specificato che nel caso dell'HDAC3 non è possibile purificare questa isoforma separando il co-fattore NCoR (regolatore della sua attività catalitica) che sarà quindi sempre presente in tutti i saggi riportati in Tabella 4. Tali considerazioni sono valide per tutti i dati di inibizione delle diverse isoforme di HDAC riportati nella presente tesi di dottorato.

I risultati ottenuti hanno, quindi, confermato la selettività prevista verso gli HDAC di classe I e la possibilità di bloccare l'enzima in una specifica conformazione. Infatti, la disparità dei dati ottenuti dai saggi condotti dalla *Reaction Biology* e quelli dell'Harvard Medical School è correlata ai differenti substrati utilizzati e viene giustificata dalla possibilità degli inibitori di legarsi a stati conformazionalmente diversi, determinando lo spostamento dell'equilibrio verso una forma piuttosto che un'altra.

Sono state considerate tutte le possibili forme tautomeriche delle strutture riportate in tabella 4 e sono stati valutati i rispettivi valori di pKa. È stata quindi eseguita un'accurata analisi conformazionale di tutte le forme considerate comprendente calcoli di dinamica, meccanica molecolare e calcoli di quanto-meccanica (PM6 e DFT). Tutti i conformeri ottenuti dai calcoli quanto-meccanici PM6 sono stati classificati in base a: i) valori di energia conformazionale ( $\Delta E < 5$  Kcal), ii) distanze interatomiche tra il carbonio carbonilico (di cheto-esteri, cheto-ammidi, ossime) o il carbonio imminico (di immino-esteri, immino-ammidi) dello ZBG ed il

centroide del CAP (intervalli di distanze utilizzati per la classificazione : <5 Å; [5-7.5 Å]; [7.5-10 Å]; >10 Å), iii) angoli torsionali del carbonio carbonilico ( di cheto-esteri e cheto-ammidi) e del carbonio imminico ( di immino-esteri, immino-ammidi), iv) legami idrogeno intramolecolari. Per eseguire i calcoli DFT in complesso con lo zinco, sono stati selezionati i conformeri con una conformazione distesa della catena alchilica (distanza ZBG-CAP >7.5 Å) e, fra questi, sono stati considerati i conformeri a più bassa energia dei diversi tautomeri possibili dello ZBG.

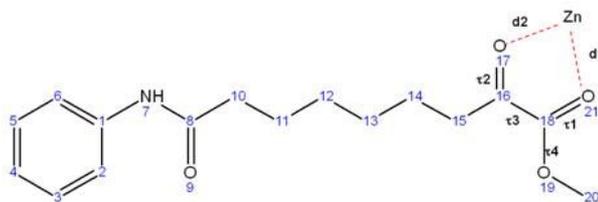
In questo modo sono stati selezionati i conformeri di partenza dei ligandi da utilizzare nei successivi calcoli di *docking* in complesso con le strutture delle isoforme cristallizzate.

### 3.4.1d Analisi dei Calcoli DFT

Per ciascun ligando sono state valutate le distanze dei coordinanti dallo zinco, gli angoli diedri che determinano la geometria di coordinazione dei coordinanti rispetto al metallo e la planarità del metile rispetto all'anello dello ZBG. Di seguito vengono riportati i risultati dell'analisi per ciascun ligando considerato:

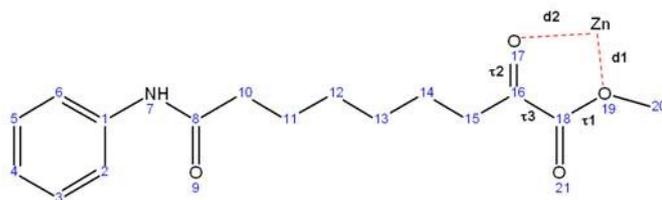
Chetoesteri: sono caratterizzati da due possibili conformazioni dell'angolo diedro  $\tau_3$  (O17C16C18O21, *Tabella 5*, 6); una conformazione *sin-periplanare* ( $\tau_3 \sim 0^\circ$ ) o una conformazione *anti-periplanare* ( $\tau_3 \sim 180^\circ$ ) degli atomi coordinanti lo zinco. Come mostrato nelle tabelle 5 e 6, i valori degli altri angoli diedri presi in considerazione indicano la perfetta co-planarità dell'anello con il metallo. Infine, anche il gruppo metile è co-planare al sistema ZBG-zinco, e ciò avviene per entrambi i conformeri.

**Tabella 5:** sono riportate le distanze dei coordinanti dallo zinco e i valori degli angoli diedri di ST7991.



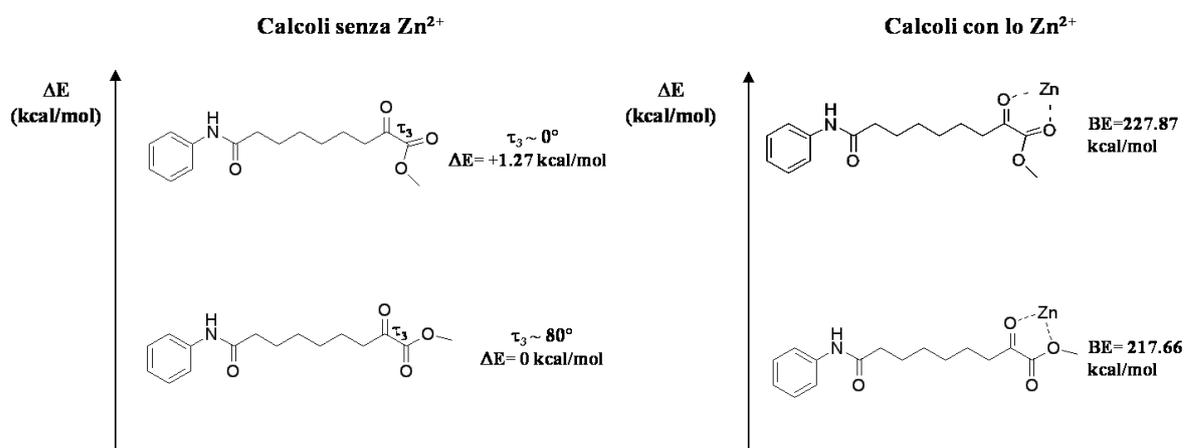
Conformazione ZBG	d1 (ZnO21)	d2 (ZnO17)	$\tau_1$ (ZnO21C18C16)	$\tau_2$ (ZnO17C17C18)	$\tau_3$ (O17C16C18O21)	$\tau_4$ (C16C18O19C20)
Sin-periplanare	2,15	2,18	0.56	-0.02	-0.37	-179.27

**Tabella 6:** sono riportate le distanze dei coordinanti dallo zinco e i valori degli angoli diedri di ST7991.



Conformazione ZBG	d1 (ZnO19)	d2 (ZnO17)	$\tau_1$ (ZnO19C18C16)	$\tau_2$ (ZnO17C16C18)	$\tau_3$ (O17C16C18O19)	$\tau_1$ (ZnO19C18C16)
Anti-periplanare	2.27	2.12	-0.38	-0.05	179.85	179.38

Entrambe le possibili conformazioni di  $\tau_3$  sono state considerate come conformazioni di partenza per i calcoli di *docking*. Infatti, la conformazione *sin-periplanare* e la conformazione *anti-periplanare* presentano una differenza di energia di sole 1.27 Kcal/mol. Anche nella scelta del risultato del *docking* sono stati considerati entrambi i conformeri, pur tenendo presente che quello *sin-periplanare* presenta una energia di legame nei confronti dello zinco (BE) più elevata rispetto a quello *anti-periplanare* (Figura 15)



**Figura 15:** nella colonna “calcoli senza  $Zn^{2+}$ ” sono mostrate le due conformazioni possibili dell’ ST7991 in assenza di zinco e riportati i relativi  $\Delta E$ . Nella colonna “calcoli con lo  $Zn^{2+}$ ” sono riportati i valori di BE di entrambe le conformazioni.

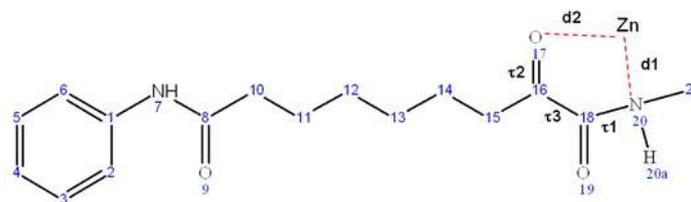
Chetoammidi: sono possibili due forme tautomeriche dello ZBG denominate tautomero A (Figura 16) e tautomero B (Figura 17), ciascuna forma è caratterizzata da due conformazioni possibili dell’angolo diedro  $\tau_3$  (O17C16C18O19, Tabella 7, 8) ovvero la conformazione *sin-periplanare* ( $\tau_3 \sim 0^\circ$ ) e la conformazione *anti-periplanare* ( $\tau_3 \sim 180^\circ$ ) degli atomi coordinanti lo zinco.

In particolare:

**Tautomero A:** Come mostrato nelle tabelle 7 e 8, quando  $\tau_3 \sim 180$  i valori degli altri angoli diedri presi in considerazione indicano la non planarità del sistema di coordinazione, a causa della presenza fra i coordinanti dell'atomo di azoto N20, caratterizzato da una ibridazione di tipo  $sp^3$ . Al contrario, quando la coordinazione avviene attraverso i due ossigeni ( $\tau_3 \sim 0^\circ$ ) la planarità con il metallo è garantita dalla ibridazione  $sp^2$  di entrambi i coordinanti e dalla conseguente delocalizzazione elettronica. Infine, il gruppo metile risulta non co-planare rispetto al sistema ZBG-zinco per la conformazione *anti-periplanare* del diedro  $\tau_3$ , mentre è co-planare al sistema ZBG-zinco per la conformazione *sin-periplanare*.

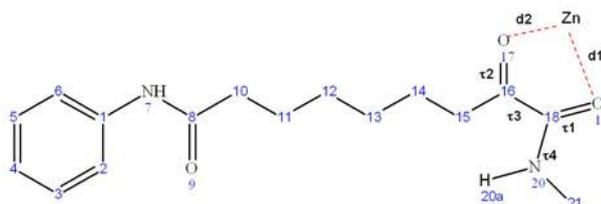
Solo la conformazione *anti-periplanare* di  $\tau_3$  è stata considerata come conformazioni di partenza per i calcoli di *docking* in quanto, in assenza di  $Zn^{2+}$ , è l'unica conformazione possibile dello ZBG. Nella scelta del risultato del *docking* sono state, invece, prese in considerazione sia le soluzioni con la conformazione *sin-periplanare* che *anti-periplanare* di  $\tau_3$ , anche se per la prima non è possibile calcolare la BE con il metallo poiché non è possibile calcolare l'energia del corrispondente conformero in assenza di zinco (Figura 16).

**Tabella 7:** sono riportate le distanze dei coordinanti dallo zinco e i valori degli angoli diedri di ST7732.



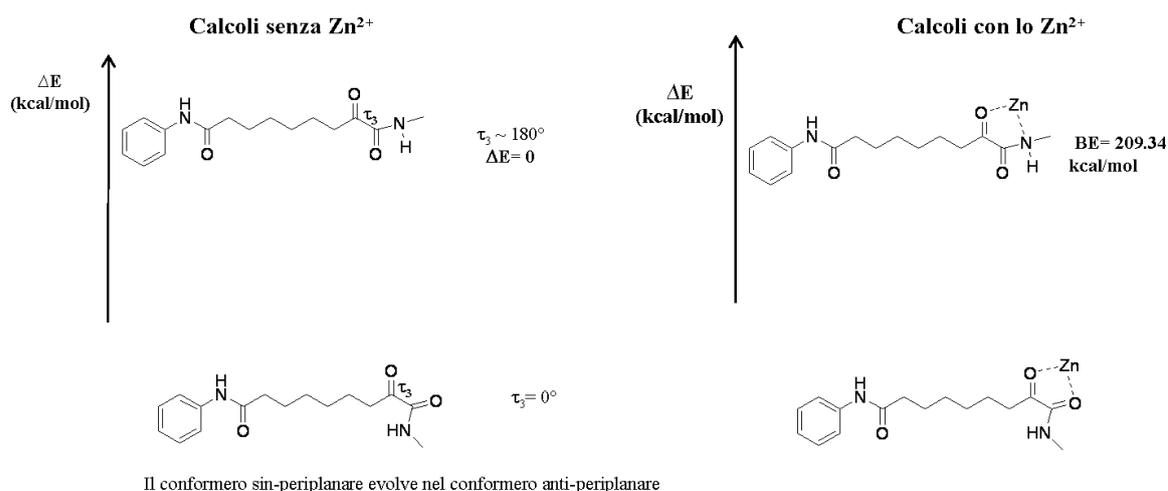
Conformazione ZBG	d1 (ZnN20)	d2 (ZnO17)	$\tau_1$ (Zn N20C18C16)	$\tau_2$ (ZnO17C16C18)	$\tau_3$ (O17C16C18O19)	$\tau_1$ (C16C18N20C21)
Tautomero A Anti-periplanare	2.34	2.16	40.15	-5.09	142.79	173.8

**Tabella 8:** sono riportate le distanze dei coordinanti dallo zinco e i valori degli angoli diedri di ST7732.



Conformazione ZBG	d1 (ZnO19)	d2 (ZnO17)	$\tau_1$ (Zn O19C18C16)	$\tau_2$ (ZnO17C16C18)	$\tau_3$ (O17C16C18O19)	$\tau_4$ (C16C18N20C21)
Tautomero A Sin-periplanare	2.06	2.18	0.64	0.098	0.495	-179.7

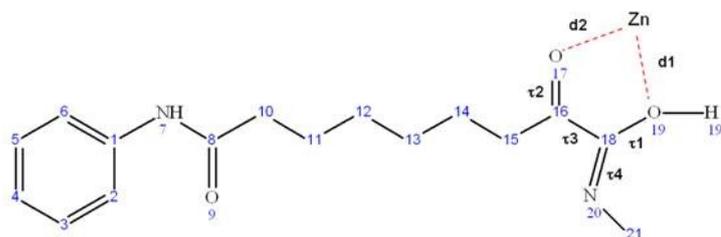
## Tautomero A



**Figura 16:** nella colonna “calcoli senza  $Zn^{2+}$ ” sono mostrate le conformazioni possibili di ST7732 Tautomero A in assenza di zinco e riportati i relativi  $\Delta E$ . Nella colonna “calcoli con lo  $Zn^{2+}$ ” sono riportati i valori di BE della conformazione antiperiplanare ma non della conformazione sinperiplanare in quanto non è calcolabile.

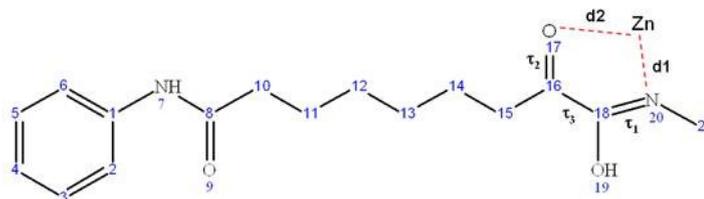
**Tautomero B:** Come mostrato nelle tabelle 9 e 10, i valori degli altri angoli diedri presi in considerazione indicano la perfetta co-planarità dei coordinanti con il metallo, anche il gruppo metile è co-planare al sistema ZBG-zinco, e ciò avviene per entrambi i conformeri. Solo la conformazione *sin-periplanare* di  $\tau_3$  è stata considerata come conformazione di partenza per i calcoli di *docking* in quanto è quella di gran lunga favorita in assenza di zinco (Figura 17). Inoltre, nella scelta del risultato del *docking* è stata presa in considerazione solo la conformazione dello ZBG con  $\tau_3 \sim 0^\circ$  poiché essa presenta una energia di legame nei confronti dello zinco (BE) molto più elevata rispetto alla conformazione con  $\tau_3 \sim 180^\circ$  (Figura 17).

**Tabella 9:** sono riportate le distanze dei coordinanti dallo zinco e i valori degli angoli diedri di ST7732.



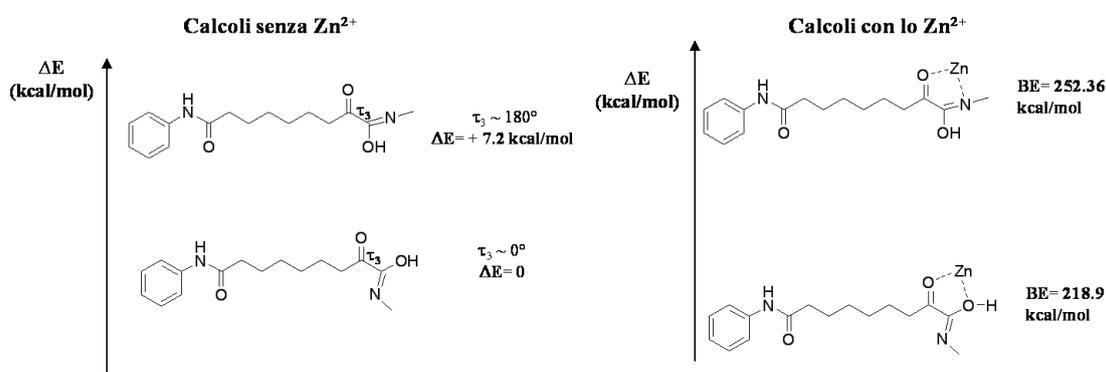
Conformazione ZBG	d1 (ZnO19)	d2 (ZnO17)	$\tau_1$ (ZnO19C18C16)	$\tau_2$ (ZnO17C16C18)	$\tau_3$ (O17C16C18O19)	$\tau_4$ (C16C18N20C21)
Tautomero B Sin-periplanare	2,23	2,09	1.09	0.23	-0.89	-179.8

**Tabella 10:** sono riportate le distanze dei coordinanti dallo zinco e i valori degli angoli diedri di ST7732.



Conformazione ZBG	d1 (ZnN20)	d2 (ZnO17)	$\tau_1$ (ZnN20C18C16)	$\tau_2$ (ZnO17C16C18)	$\tau_3$ (O17C16C18O19)	$\tau_1$ (C16C18N20C21)
Tautomero B Anti-periplanare	2.15	2.14	0.23	-0.24	179.41	179.43

### Tautomero B

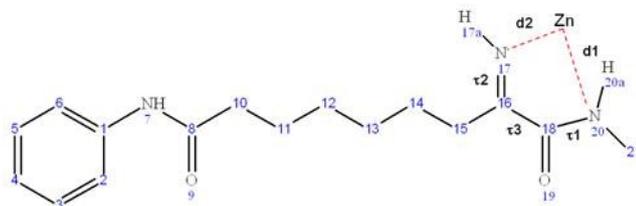


**Figura 17:** nella colonna “calcoli senza  $Zn^{2+}$ ” sono mostrate le due conformazioni possibili di ST7732 tautomero B in assenza di zinco e riportati i relativi  $\Delta E$ . Nella colonna “calcoli con lo  $Zn^{2+}$ ” sono riportati i valori di BE di entrambe le conformazioni del tautomero B.

Immino ammidi: sono possibili due conformazioni del diedro  $\tau_3$  (N17C16C18O19, Tabella 11, 12), ovvero la conformazione *sin-periplanare* ( $\tau_3 \sim 0^\circ$ ) e la conformazione *anti-periplanare* ( $\tau_3 \sim 180^\circ$ ) degli atomi coordinanti lo zinco.

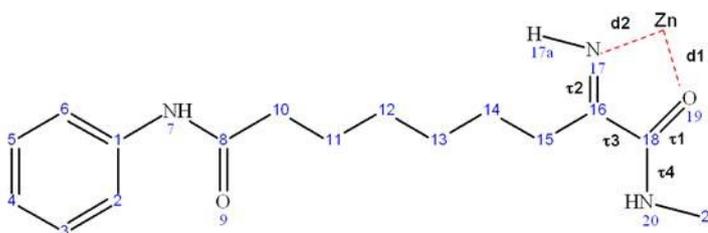
Quando  $\tau_3 \sim 180^\circ$ , come mostrato nelle tabelle 11 e 12, i valori degli altri angoli diedri presi in considerazione indicano la non planarità del sistema di coordinazione, a causa della presenza fra i coordinanti dell'atomo di azoto N20, caratterizzato da una ibridazione di tipo  $sp^3$ . Al contrario, quando la coordinazione avviene con  $\tau_3 \sim 0^\circ$  la planarità con il metallo è consentita dalla presenza come coordinanti dell'ossigeno e dell'azoto N17, entrambi ibridati  $sp^2$ . Di conseguenza la BE di coordinazione con lo zinco è molto maggiore nel caso della conformazione *sin-periplanare* (Figura 18). In assenza del metallo, è favorita la conformazione *anti-periplanare* di  $\tau_3$  di 3.8 kcal/mol rispetto a quella *sin-periplanare* (Figura 18).

**Tabella 11:** sono riportate le distanze dei coordinanti dallo zinco e i valori degli angoli diedri di ST7729.

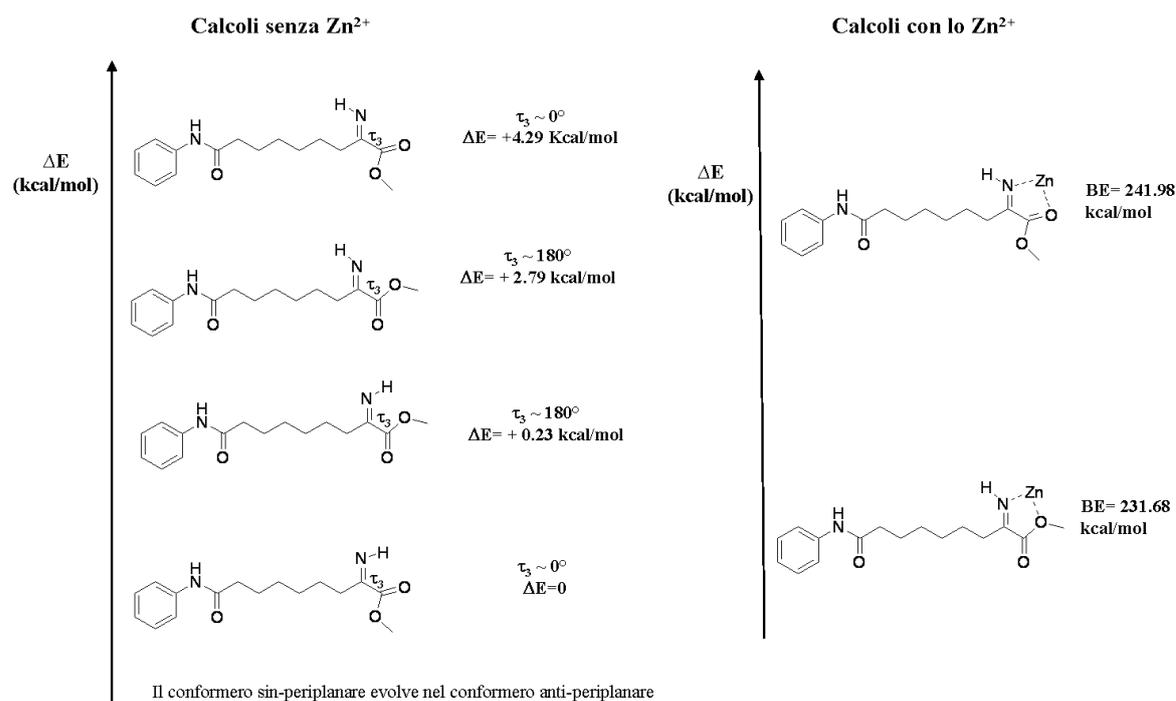


Conformazione ZBG	d1 (ZnN20)	d2 (ZnN17)	$\tau$ 1 (ZnN20C18C16)	$\tau$ 2 (ZnN17C16C18)	$\tau$ 3 (N17C16C18O19)	$\tau$ 1 (C16C18N20C21)
Anti-periplanare	2.30	2.13	42.55	-2.84	141.18	175.87

**Tabella 12:** sono riportate le distanze dei coordinanti dallo zinco e i valori degli angoli diedri di ST7729.



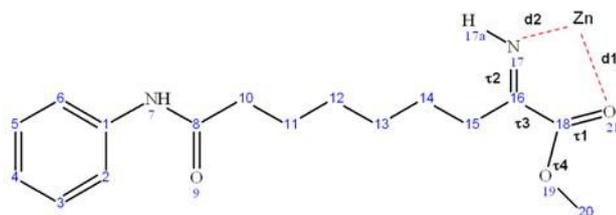
Conformazione ZBG	d1 (ZnO19)	d2 (ZnN17)	$\tau$ 1 (ZnO19C18C16)	$\tau$ 2 (ZnN17C16C18)	$\tau$ 3 (N17C16C18O19)	$\tau$ 4 (C16C18N20C21)
Sin-periplanare	2.06	2.14	1.46	1.45	-2.00	-179.32



**Figura 18:** nella colonna “calcoli senza Zn<sup>2+</sup>” sono mostrate le conformazioni possibili di ST7729 in assenza di zinco e riportati i relativi  $\Delta E$  qualora calcolabili. Nella colonna “calcoli con lo Zn<sup>2+</sup>” sono riportati i valori di BE delle conformazioni qualora calcolabili.

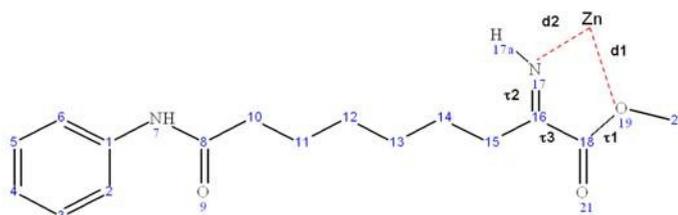
**Immino esteri:** sono possibili due conformazioni dell’angolo diedro  $\tau_3$  (N17C16C18O2, Tabelle 13, 14), una conformazione *sin-periplanare* ( $\tau_3 \sim 0^\circ$ ) o una conformazione *anti-periplanare* ( $\tau_3 \sim 180^\circ$ ) degli atomi coordinanti lo zinco. Come mostrato nelle tabelle 13 e 14, i valori degli altri angoli diedri presi in considerazione indicano la perfetta co-planarità dei coordinanti con il metallo. Infine, anche il gruppo metile è co-planare al sistema ZBG-zinco, e ciò avviene per entrambi i conformeri. Anche in questo caso la coordinazione energeticamente favorita è quella con  $\tau_3 \sim 0^\circ$ . In assenza di zinco entrambi i conformeri possono essere ugualmente presenti ( $\Delta E = 0.23$  Kcal/mol) (Figura 19).

**Tabella 13:** sono riportate le distanze dei coordinanti dallo zinco e i valori degli angoli diedri di ST7727.

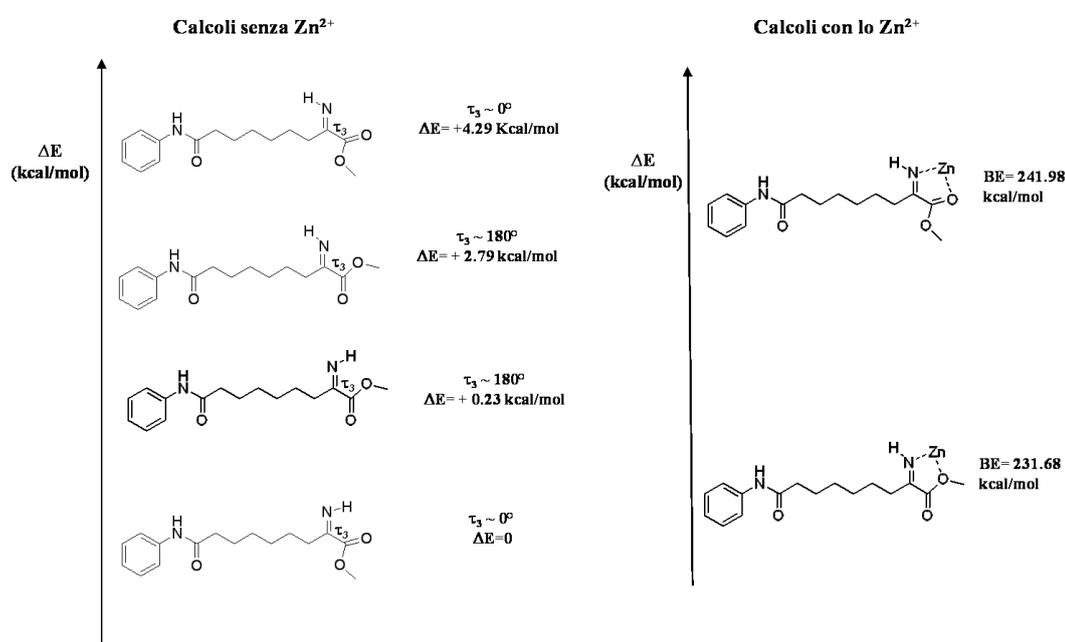


Conformazione ZBG	d1 (ZnO21)	d2 (ZnN17)	$\tau_1$ (ZnO21C18C16)	$\tau_2$ (ZnN17C16C18)	$\tau_3$ (N17C16C18O21)	$\tau_4$ (C16C18O19C20)
Sin-periplanare	2.13	2.14	1.00	1.72	-1.89	-179.10

**Tabella 14:** sono riportate le distanze dei coordinanti dallo zinco e i valori degli angoli diedri di ST7727.



Conformazione ZBG	d1 (ZnO19)	d2 (ZnN17)	$\tau_1$ (ZnO19C18C16)	$\tau_2$ (ZnN17C16C18)	$\tau_3$ (N17C16C18O21)	$\tau_1$ (C16C18O19C20)
Anti-periplanare	2,21	2,11	1.69	1.03	177.60	179.49



**Figura 19:** nella colonna "calcoli con lo Zn<sup>2+</sup>" sono mostrate le conformazioni possibili di ST7727 in assenza di zinco e riportati i relativi  $\Delta E$ . Nella colonna "calcoli con lo Zn<sup>2+</sup>" sono riportati i valori di BE delle conformazioni che possono coordinare lo zinco.

### 3.4.1e Analisi degli Studi di Docking

Le strutture ottenute dai calcoli di *docking* sono state analizzate sulla base dei risultati ottenuti dall'analisi strutturale dei complessi determinati sperimentalmente e delle relazioni struttura-attività (SARs) degli inibitori precedentemente progettati. In particolare, i risultati dei calcoli di *docking* sono stati classificati considerando: i) l'energia di interazione ligando-enzima di ciascun complesso; ii) la differenza di energia tra la conformazione del ligando legato alla proteina e la conformazione di minimo globale; iii) le distanze tra i coordinanti e lo zinco catalitico; iv) la geometria di coordinazione dello ZBG; v) la planarità del metile ed etile

rispetto allo ZBG, considerando sia i risultati ottenuti dai calcoli di *docking* che quelli dei calcoli DFT.

I requisiti strutturali fondamentali per l'interazione con l'enzima sono risultati essere quindi:

i) l'accomodamento del sostituito ingombrante nel canale di uscita dell'acetato, ii) la formazione di una rete di legami idrogeno con i residui presenti nel sito catalitico, iii) la conformazione favorevole dello ZBG per l'interazione con lo zinco ricavata dai valori di BE dei calcoli DFT.

Di seguito (*Tabella 15*) sono riportati i risultati dei docking ottenuti per l'HDAC2.

**Tabella 15.** Complessi ottenuti dai calcoli DFT e dagli studi di *docking*; distanze (d1,d2), angoli torsionali ( $\tau_1$ - $\tau_4$ ) e modalità di legame riprodotta dei gruppi chelanti lo zinco; energia di legame (BE) dei complessi ligando-proteina.

Molecola	d1	d2	$\tau_1$	$\tau_2$	$\tau_3$	$\tau_4$	Modalità di legame Raggi X	BE
ST7991 <sup>a</sup> Sin-periplanare	2.15 Zn-O21	2.18 ZnO17	0.56 ZnO21C18C16	-0.02 ZnO17C17C18	-0.37 O17C16C18O21	-179.27 C16C18O19C20		
ST7991 <sup>a</sup> Anti-periplanare	2.27 ZnO19	2.12 ZnO17	-0.38 ZnO19C18C16	-0.05 ZnO17C16C18	179.85 O17C16C18O19	179.38 ZnO19C18C16		
ST7990 <sup>b</sup>	2.62	2.32	41.33	-45.81	-2.40	-179.50	2VQV	-56.66
ST799 <sup>b</sup>	2.63	2.30	-12.53	16.70	-179.49	-178.21	2VQV	-44.11
ST7732 <sup>a</sup> Tautomero A Anti-periplanare	2.34 ZnN20	2.16 ZnO17	40.15 ZnN20C18C16	-5.09 ZnO17C16C18	142.79 O17C16C18O19	173.8 C16C18N20C21		
ST7732 <sup>a</sup> Tautomero A Sin-periplanare	2.06 ZnO19	2.18 ZnO17	0.64 ZnO19C18C16	0.098 ZnO17C16C18	0.495 O17C16C18O19	-179.7 C16C18N20C21		
ST7732 <sup>a</sup> Tautomero B Sin-periplanare	2.23 ZnO19	2.09 ZnO17	1.09 ZnO19C18C16	0.23 ZnO17C16C18	-0.89 O17C16C18O19	-179.8 C16C18N20C21		
ST7732 <sup>a</sup> Tautomero B Anti-periplanare	2.15 ZnN20	2.14 ZnO17	0.23 ZnN20C18C16	-0.24 ZnO17C16C18	179.41 O17C16C18O19	179.43 C16C18N20C21		
ST7733 <sup>b</sup> Tautomero A	2.94	2.29	39.41	-35.43	164.15	177.40	2VQV	-48.71
ST7733 <sup>b</sup> Tautomero B	2.58	2.27	35.54	-44.34	0.96	178.48	2VQV	-48.91
ST7732 <sup>b</sup> Tautomero A	3.07	2.29	40.37	-38.47	163.10	179.84	2VQV	-51.39

ST7732 <sup>b</sup> Tautomero B	2.51	2.27	27.68	-41.31	6.49	-178.92	2VQV	-48.10
ST7729 <sup>a</sup> Anti-periplanare	2.30 ZnN20	2.13 ZnN17	42.55 ZnN20C18C16	-2.84 ZnN17C16C18	141.18 N17C16C18O19	175.87 C16C18N20C21		
ST7729 <sup>a</sup> Sin-periplanare	2.06 ZnO19	2.14 ZnN17	1.46 ZnO19C18C16	1.45 ZnN17C16C18	-2.00 N17C16C18O19	-179.32 C16C18N20C21		
ST7730 <sup>b</sup>	2.82	2.40	50.17	-22.31	151.38	178.66	2VQV	-21.56
ST7729 <sup>b</sup>	3.02	2.43	37.85	-12.62	154.36	179.60	2VQV	-22.49
ST7727 <sup>a</sup> Sinperiplanare	2.13 ZnO21	2.14 ZnN17	1.00 ZnO21C18C16	1.72 ZnN17C16C18	-1.89 N17C16C18O21	-179.10 C16C18O19C20		
ST7727 <sup>a</sup> Antiperiplanare	2.21 ZnO19	2.11 ZnN17	1.69 ZnO19C18C16	1.03 ZnN17C16C18	177.60 N17C16C18O21	179.49 C16C18O19C20		
ST7728 <sup>b</sup>	2.43	2.44	13.08	-21.38	6.30	177.66	2VQV	-51.98
ST7727 <sup>b</sup>	2.47	2.44	9.74	-18.02	5.95	177.85	2VQV	-49.54

<sup>a</sup> Complessi ottenuti dai calcoli DFT; <sup>b</sup> complessi ottenuti dagli studi di *docking*.

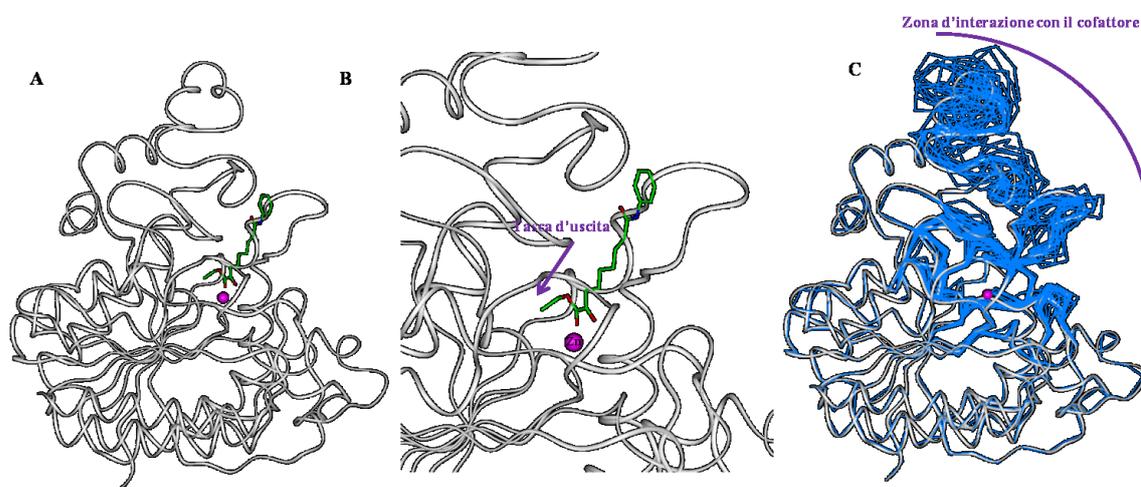
La ST7990 presenta rispetto alla ST7991 un simile valore di DE del dal minimo globale (i.e., energia conformazionale) ma un maggiore guadagno energetico nell'interazione con l'enzima. Ciò si esplica con l'occupazione della tasca di uscita dei prodotti di catalisi, in grado di bloccare l'enzima in una conformazione non compatibile con il legame con i cofattori necessari per la catalisi. Inoltre, la soluzione del *docking* della ST7990 presenta una coordinazione dello zinco con lo ZBG nella conformazione *cis*, in accordo con i dati ottenuti dai calcoli DFT. Al contrario, la ST7991 assume una conformazione *trans*, caratterizzata da una minore tendenza a coordinare lo zinco; motivo per cui quest'ultima risulta essere inattiva. Inoltre, essendo la tasca di uscita dei prodotti della catalisi (acetato) nell'HDAC2 più larga di quella presente nell'HDAC3, è necessario un sostituito più ingombrante del metile per avere una variazione conformazionale di L1 tale da non permettere all'enzima di interagire con i suoi cofattori (*Figura 20*).

Per le molecole ST7733 ed ST7732 sono stati considerati sia il tautomero A che il tautomero B, entrambi caratterizzati da un buon valore di BE con la proteina e da un'energia conformazionale paragonabile a quella degli altri ligandi. E' interessante sottolineare che durante i calcoli di *simulated annealing* (vedere parte sperimentale) avviene la rotazione dello ZBG sia per il tautomero A che per quello B, indicando una vasta gamma di possibili modalità di legame che inducono altrettanti variazioni conformazionali nella proteina, in accordo con la buona attività inibitoria riportata per queste molecole.

La ST7730 e la ST7729 sono entrambe risultate inattive ed, infatti, presentano un valore di BE con l'enzima di molto inferiore a quello degli altri ligandi. Inoltre, coordinano il metallo nella conformazione con  $\tau_3 \sim 180^\circ$  che non è quella energeticamente favorita.

La ST7728, ancora attiva sugli HDAC2 rispetto alla ST7727 completamente inattiva, evidenzia che la presenza dell'etile sullo ZBG è fondamentale per determinare sia un guadagno energetico nell'interazione ligando-proteina che una variazione conformazionale tale da destabilizzare la regione di legame L1 di interazione con il cofattore, similmente a quanto osservato per i derivati ST990 ed ST991.

Di seguito (Figura 20) viene riportato il risultato dei *docking* della molecola più attiva ST7990 (Tabella 4).



**Figura 20:** in figura è mostrato l'accomodamento di ST7990 nel sito catalitico dell'HDAC2, la figura 18B è un particolare del risultato del *docking* che mette in evidenza l'accomodamento dello ZBG nel sito catalitico con occupazione della tasca d'uscita e la coordinazione dello zinco catalitico. La figura 18C mostra invece il movimento della proteina determinato dall'accomodamento del ligando nel sito catalitico e che quindi va a perturbare la zona di interazione con il cofattore. Gli eteroatomi del ligando sono colorati sulla base dei tipi atomici (N: blu ; O: rosso e C:verde).

Di seguito (Tabella 16) vengono riportati i risultati dei *docking* ottenuti per l'HDAC3

**Tabella 16.** Complessi ottenuti dai calcoli DFT e dagli studi di *docking*; distanze (d1,d2), angoli torsionali ( $\tau_1$ - $\tau_4$ ) e modalità di legame riprodotta dei gruppi chelanti lo zinco; energia di legame (BE) dei complessi ligando-proteina.

Molecola	d1	d2	$\tau_1$	$\tau_2$	$\tau_3$	$\tau_4$	Modalità di legame Raggi X	BE
ST7991 <sup>a</sup> Sin-periplanare	2,15 ZnO21	2,18 ZnO17	0,56 ZnO21C18C16	-0,02 ZnO17C17C18	-0,37 O17C16C18O21	-179,27 C16C18O19C20		
ST7991 <sup>a</sup> Anti-periplanare	2,27 ZnO19	2,12 ZnO17	-0,38 ZnO19C18C16	-0,05 ZnO17C16C18	179,85 O17C16C18O19	179,38 ZnO19C18C16		

ST7990 <sup>b</sup>	2.54	2.27	45.15	-47.60	-3.94	-178.74	2VQM	-77.55
ST7991 <sup>b</sup>	2.41	2.30	6.54	-14.38	5.23	177.60	2VQV	-68.18
ST7991 <sup>b</sup>	3.08	2.30	34.20	-45.06	-6.37	-177.43	2VQM	-66.56
ST7732 <sup>a</sup> Anti-periplanare	2.34 ZnN20	2.16 ZnO17	40.15 ZnN20C18C16	-5.09 ZnO17C16C18	142.79 O17C16C18O19	173.8 C16C18N20C21		
ST7732 <sup>a</sup> Sin-periplanare	2.06 ZnO19	2.18 ZnO17	0.64 ZnO19C18C16	0.098 ZnO17C16C18	0.495 O17C16C18O19	-179.7 C16C18N20C21		
ST7732 <sup>a</sup> Sin-periplanare	2.23 ZnO19	2.09 ZnO17	1.09 ZnO19C18C16	0.23 ZnO17C16C18	-0.89 O17C16C18O19	-179.8 C16C18N20C21		
ST7732 <sup>a</sup> Anti-periplanare	2.15 ZnN20	2.14 ZnO17	0.23 ZnN20C18C16	-0.24 ZnO17C16C18	179.41 O17C16C18O19	179.43 C16C18N20C21		
ST7733 <sup>b</sup> Tautomero A	2.86	2.38	41.86	-24.71	159.17	177.61	2VQV	-33.47
ST7733 <sup>b</sup> Tautomero B	2.45	2.27	-33.32	-8.48	33.12	177.09	2VQV OH su N 1VKG	-56.31
ST7732 <sup>b</sup> Tautomero A	2.94	2.38	36.78	-30.84	166.50	179.46	2VQV	-34.20
ST7732 <sup>b</sup> Tautomero B	2.50	2.27	-0.70	-15.09	10.52	177.63	2VQV	-58.16
ST7729 <sup>a</sup> Anti-periplanare	2.30 ZnN20	2.13 ZnN17	42.55 ZnN20C18C16	-2.84 ZnN17C16C18	141.18 N17C16C18O19	175.87 C16C18N20C21		
ST7729 <sup>a</sup> Sin-periplanare	2.06 ZnO19	2.14 ZnN17	1.46 ZnO19C18C16	1.45 ZnN17C16C18	-2.00 N17C16C18O19	-179.32 C16C18N20C21		
ST7730 <sup>b</sup>	2.85	2.50	49.66	-21.16	152.80	178.86	2VQV	26.71
ST7729 <sup>b</sup>	3.00	2.49	37.54	-25.07	163.74	-179.85	2VQV	-22.50
ST7727 <sup>a</sup> Sin-periplanare	2.13 ZnO21	2.14 ZnN17	1.00 ZnO21C18C16	1.72 ZnN17C16C18	-1.89 N17C16C18O21	-179.10 C16C18O19C20		
ST7727 <sup>a</sup> Anti-periplanare	2.21 ZnO19	2.11 ZnN17	1.69 ZnO19C18C16	1.03 ZnN17C16C18	177.60 N17C16C18O21	179.49 C16C18O19C20		
ST7728 <sup>b</sup>	2.51	2.54	-39.69	-4.27	34.07	176.96	2VQV CO su N 1VKG	-26.87
ST7727 <sup>b</sup>	2.53	2.53	5.11	-14.26	6.62	176.76	2VQV	-67.47

<sup>a</sup> Complessi ottenuti dai calcoli DFT <sup>b</sup> complessi ottenuti dagli studi di docking

ST7991 presenta una minore energia di interazione con la proteina rispetto a ST7990, in accordo con i dati di inibizione enzimatica.

Come avveniva per l'HDAC2, anche in questo caso ST7732 e ST7733 presentano diverse modalità di interazione con la proteina, tutte caratterizzate da una buona energia di legame e

una significativa variazione conformazionale della zona L1. Ciò è in accordo con la loro buona attività inibitoria.

Vale la pena sottolineare che i composti ST7991 ed ST7733, i più attivi nei confronti dell'HDAC3, sono anche quelli che presentano una maggiore variazione conformazionale della regione L1, coinvolta nel legame con il cofattore NCoR (anch'esso presente nei saggi di inibizione enzimatica).

I derivati ST7729 e ST7730 presentano un'energia di legame con la proteina di gran lunga inferiore rispetto agli altri ligandi, giustificando la perdita di attività inibitoria di questi composti.

La ST7728 è di un ordine di grandezza più attiva della ST7727 tuttavia è caratterizzata da un guadagno energetico molto minore in seguito all'interazione con l'enzima. Tuttavia, l'accomodamento del sostituente etile nella tasca di uscita determina una maggiore variazione conformazionale delle regioni di legame L1 e L2 che va a destabilizzare maggiormente l'interazione con il cofattore.

Di seguito (*Figura 21*) viene riportato il risultato dei *docking* della molecola più attiva ST7990 (*Tabella 4*).



**Figura 21:** in figura è mostrato l'accomodamento dell'ST7990 nel sito catalitico dell'HDAC3, la figura 19B è un particolare del risultato del docking che mette in evidenza l'accomodamento dello ZBG nel sito catalitico con occupazione della tasca d'uscita e la coordinazione dello zinco catalitico. La figura 19C mostra invece il movimento della proteina determinato dall'accomodamento del ligando nel sito catalitico e che quindi va a perturbare la zona di interazione con NCoR. Gli eteroatomi del ligando sono colorati sulla base dei tipi atomici (N: blu ; O: rosso; C:verde).

I composti riportati in Tabella 4 sono tutti attualmente in esame presso i laboratori del dott. Petrella che sta effettuando i saggi di citotossicità su diverse linee cellulari tumorali e non. I dati preliminari ottenuti indicano, per i composti risultati attivi quali inibitori degli HDAC di classe I, un'attività a concentrazioni basso micromolari ( $\sim 2 \mu\text{M}$  paragonabili a quella del SAHA) su cellule di linfoma di tipo Jurkat e U937, paragonabile a quella del SAHA, tuttavia, rispetto al SAHA presentano una maggiore tossicità selettiva nei confronti di altre linee cellulari, sia tumorali che non.

### 3.4.2 Conclusioni

I composti progettati hanno confermato l'ipotizzata selettività verso gli HDAC di classe I e la possibilità di bloccare l'enzima in una specifica conformazione. Infatti, la disparità dei dati ottenuti dai saggi condotti dalla *Reaction Biology* e quelli dell'Harvard Medical School è correlata ai differenti substrati utilizzati e viene giustificata dalla possibilità degli inibitori (come dei diversi substrati peptidici utilizzati) di legarsi a stati conformazionalmente diversi dell'enzima.

I risultati di *docking* hanno mostrato che negli HDAC di classe I i diversi ZBG progettati determinano una diversa coordinazione del metallo, un diverso accomodamento del gruppo metile/etile e diverse interazioni con gli amminoacidi del sito catalitico; ciò determina un differente accomodamento del CAP ed una diversa variazione conformazionale del sito di legame con i co-fattori coinvolti nella loro attività catalitica. In particolare, l'accomodamento dei sostituenti metile/etile nella tasca d'uscita dei prodotti di catalisi determina un cambiamento conformazionale della regione L2, connessa direttamente alla regione di legame dei cofattori L1. Combinando, quindi, i risultati ottenuti i) dall'analisi conformazionale dei ligandi progettati, ii) dall'analisi strutturale degli inibitori cristallizzati e iii) dagli studi di *docking*, con le attività sulle diverse isoforme HDAC degli inibitori progettati, è stato dunque possibile formulare un'ipotesi di legame con l'enzima per i nostri ligandi che ne consentirà la modifica razionale per lo sviluppo di nuovi potenziali agenti terapeutici.

I composti progettati sono attualmente oggetto di investigazione farmacologica da parte del gruppo di ricerca del dott. Petrella del Dipartimento di Farmacia dell'Università degli Studi di Salerno.

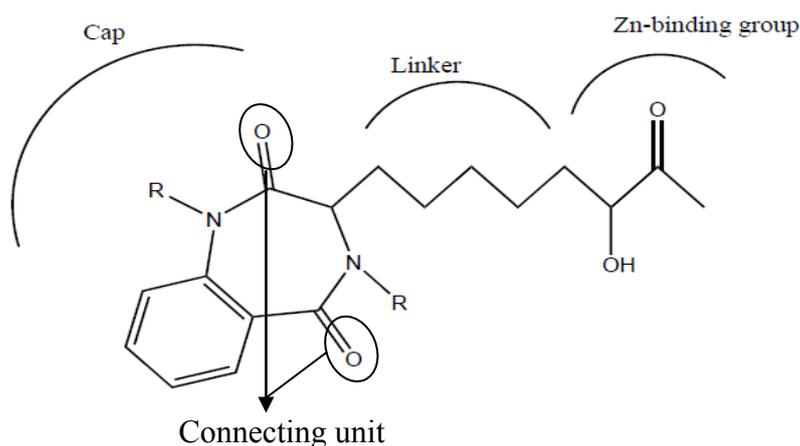
## 3.5 Progettazione di nuovi inibitori delle Deacetilasi Istoniche (HDAC) caratterizzati da un nucleo benzodiazepinico

### 3.5.1 Risultati e Discussione

#### 3.5.1a Implementazione del Modello Farmacoforico precedentemente progettato

Al fine di raggiungere gli obiettivi descritti nel paragrafo 3.2, sono stati progettati anche composti aventi un nucleo benzodiazepinico, considerato uno scheletro versatile, dimostratosi efficace nell'agire sugli HDAC e trattare diverse forme di leucemia<sup>1</sup>. Sulle nuove molecole progettate sono stati svolti: i) studi di relazione struttura attività; ii) un'analisi delle proprietà conformazionali e delle modalità di interazione dei composti con gli enzimi. Prima di eseguire queste ricerche è stata condotta un'analisi bioinformatica sulle strutture cristallografiche per determinare i punti farmacoforici degli inibitori a livello del CAP. In particolare, gli studi sono stati svolti partendo dal farmacoforo degli inibitori HDAC presentato nel paragrafo 3.4.1b (Figure 10,11).

Su queste basi è stata formulata un'ipotesi di farmacoforo per lo scheletro molecolare a struttura benzodiazepinica dei nuovi ligandi (Figura 1):



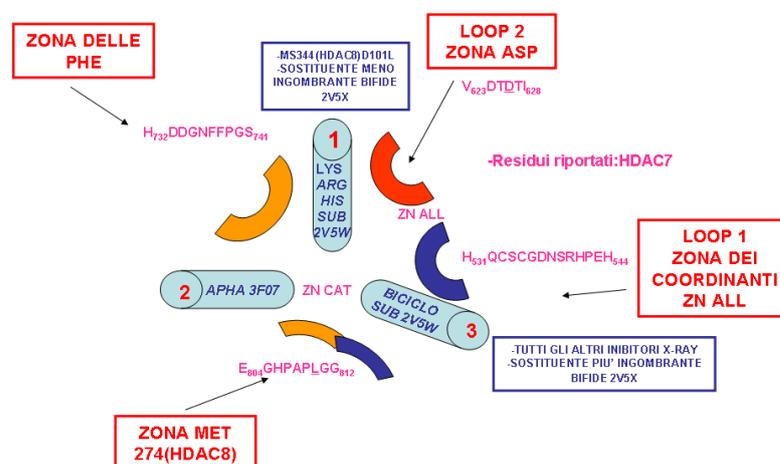
**Figura 1:** struttura generale di un HDACi a struttura benzodiazepinica. La molecola è caratterizzata dalla presenza di uno ZBG, un linker e un CAP.

Le molecole sintetizzate sono state sottoposte ad analisi conformazionale ed il loro posizionamento nel sito attivo delle varie isoforme HDAC è stato eseguito prendendo in

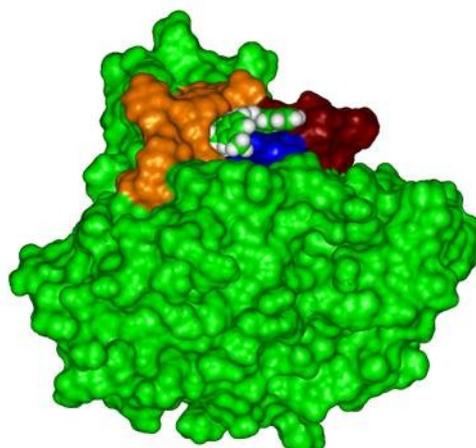
considerazione il farmacoforo dei ligandi cristallizzati in complesso con le varie isoforme degli enzimi.

Il farmacoforo dei nostri ligandi è stato quindi comparato e sovrapposto con quello risultante da tale analisi (*pharmacophore fitting*). Contemporaneamente sono stati eseguiti calcoli di *docking* flessibile dei nostri ligandi in complesso con le strutture delle isoforme cristallizzate. Al fine di razionalizzare al meglio i risultati ottenuti, è stato anche utilizzato uno schema, che suddivide il sito di legame delle diverse isoforme di HDAC in canali e relative tasche di superficie (*surface pockets*), di seguito riportato e sviluppato nel laboratorio presso il quale ho svolto la mia attività di ricerca.

**Schema 1:** canali e tasche di un HDAC



In accordo con tale schema, la superficie molecolare delle strutture degli enzimi è stata suddivisa in diverse regioni, come riportato nella figura 2 seguente:



**Figura 2:** superficie di Connolly dell'HDAC in complesso con un inibitore. La zona in arancio rappresenta il canale 1, la zona in blu il canale 2, quella in rosso il canale 3.

Sulla base dei risultati ottenuti è stato dunque possibile formulare un'ipotesi di legame con l'enzima per i nostri ligandi che ne consentirà la modifica razionale per lo sviluppo di nuovi potenziali agenti terapeutici.

### ***3.5.1b Analisi del farmacoforo di ligandi noti in complesso con l'enzima.***

L'analisi bioinformatica condotta ha consentito di definire il posizionamento del CAP dell'inibitore nelle diverse isoforme di HDAC umani presenti nella banca dati *Protein Data Bank*. Sono state quindi individuate una serie di strutture cristallografiche in cui il gruppo carbonilico dell'unità di connessione (CU) dei ligandi (*Figura farmacoforo*) va ad occupare posizioni diverse. Le strutture cristallografiche selezionate sono state suddivise in due gruppi in base al tipo di composto legato:

- substrato: struttura cristallografica PDB ID: 3EWF (HDAC8);
- inibitori: strutture cristallografiche PDB IDs: 2V5X, 1T67, 1W22, 1VKG (HDAC8), dalla PDB ID: 3C10 (HDAC7) e dalla PDB ID: 2VQM (HDAC4).

La struttura cristallografica PDB ID: 3EWF (mutante H143A) presenta il carbonile del CAP del substrato, orientato verso il canale 1, che interagisce attraverso una molecola d'acqua con l'NH dello scheletro peptidico della Phe208 (loop delle Phe). Il carbonile del substrato, orientato invece verso il canale 3, interagisce attraverso una molecola d'acqua con l'His180 catalitica. Le due molecole d'acqua coinvolte in questi legami interagiscono tra di loro. Inoltre il gruppo carbonilico orientato verso il canale 3 potrebbe anche interagire con un'altra molecola d'acqua, a sua volta coinvolta in una rete di legami ad idrogeno con altre molecole d'acqua, che interagiscono con il gruppo carbonilico dello scheletro peptidico della Pro273 e Met274.

Nella struttura cristallografica PDB ID: 2V5X (che presenta un ligando con il CAP ramificato in due diverse porzioni molecolari) il carbonile del CAP dell'inibitore è orientato verso il canale 1 e interagisce attraverso una molecola d'acqua con lo scheletro peptidico della regione di legame delle fenilalanina (l'NH dello scheletro peptidico della Phe208). Questo carbonile ha interazioni con questa stessa regione di legame anche attraverso una molecola d'acqua, che interagisce con il carbonile dello scheletro peptidico della Gly206. Il carbonile del CAP del substrato, orientato verso il canale 3, attraverso una molecola d'acqua dà legami ad idrogeno

con l'His180 catalitica. Le due molecole d'acqua coinvolte in questi legami interagiscono tra di loro.

Nella struttura cristallografica PDB ID: 1T67 (MS344) il carbonile del CAP ha interazione con l'His180 catalitica attraverso una molecola d'acqua.

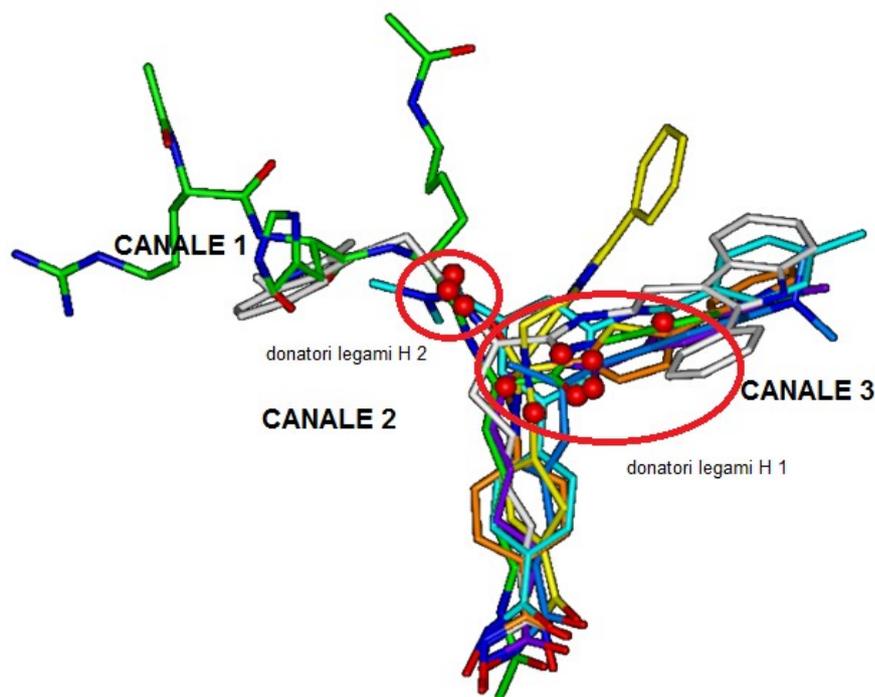
Nella struttura cristallografica PDB ID: 1W22, l'ossigeno del gruppo SO del CAP dà interazione con l'His180 mediante una molecola d'acqua. L'istidina a sua volta interagisce con la molecola d'acqua che è legata all'NH dello scheletro peptidico della Phe208 (regione di legame della fenilalanina).

Nella struttura cristallografica PDB ID: 1VKG (CRA-19156) il carbonile del CAP dell'inibitore, orientato verso il canale 3, interagisce attraverso una rete di molecole d'acqua sia con l'His180 che con il gruppo carbonilico dello scheletro peptidico della Met274. Il carbonile del CAP, orientato verso il canale 1, attraverso una molecola d'acqua vede il gruppo carbonilico dello scheletro peptidico della Gly206 (regione di legame della Phe208).

Nella struttura cristallografica PDB ID: 3C10 (TSA) il carbonile del CAP del TSA interagisce, attraverso una molecola d'acqua, sia con il gruppo carbonilico dello scheletro peptidico della Pro809, sia con l'His709 catalitica. Il gruppo carbonilico della Pro809 occupa la posizione della molecola d'acqua che permette al substrato di interagire con il gruppo carbonilico dello scheletro peptidico della Pro273 e della Met274.

Nella struttura cristallografica PDB ID: 2VQM (inibitore acido idrossammico) il carbonile del CAP dell'inibitore, attraverso una molecola d'acqua vede l'His198 catalitica e il gruppo carbonilico dello scheletro peptidico della Pro298. Inoltre attraverso due molecole d'acqua si ha un'interazione con la regione di legame delle fenilalanine (NH dello scheletro peptidico della Phe227).

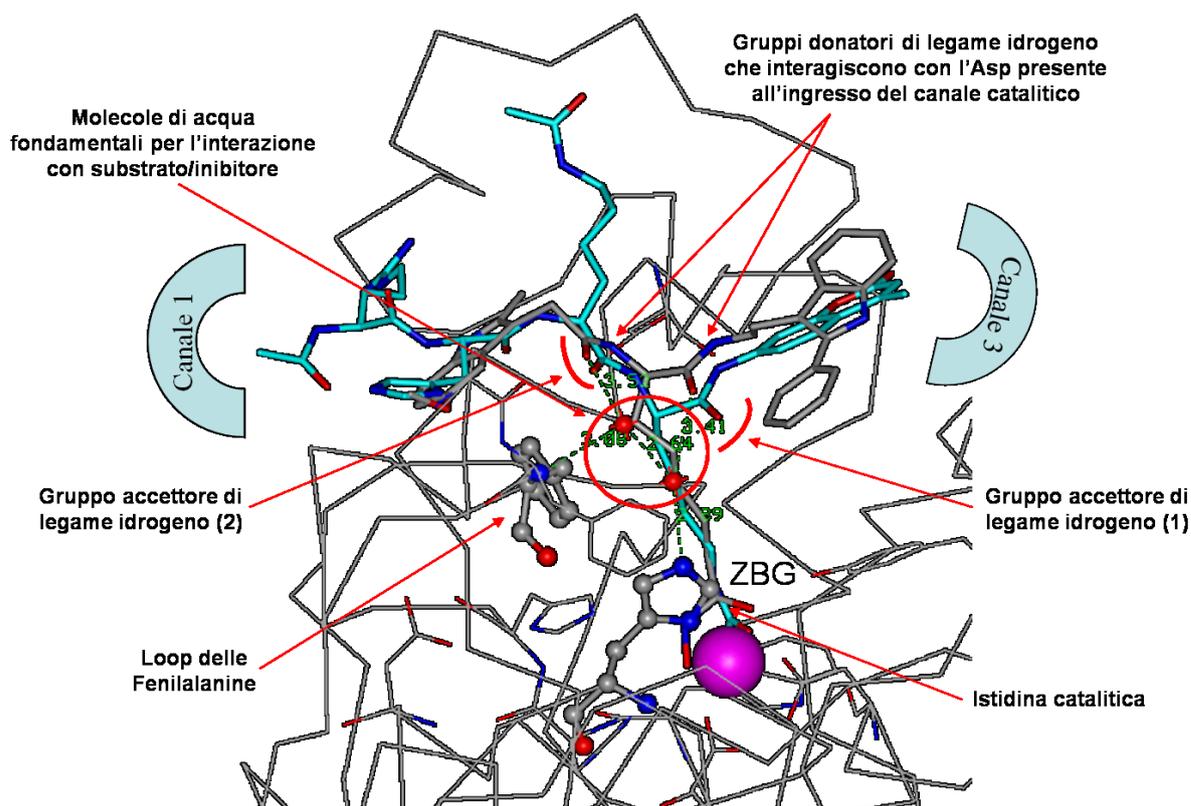
Le differenti posizioni del carbonile del CAP dei vari ligandi sono raffigurate nell'immagine seguente (*Figura 3*).



**Figura 3 : posizioni del carbonile del CAP degli inibitori cristallizzati.** In figura sono riportate tutte le PDB utilizzate con le rispettive posizioni dei gruppi carbonilici, evidenziati con un pallino rosso. In verde la 3EWF (HDAC8/Substrato); in bianco la 2V5X (HDAC8/acido idrossammico); in arancione la 1W22 (HDAC8/acido idrossammico); in ciano la 1VKG (HDAC8/acido idrossammico); in viola la 1T67(HDAC8/MS-344); in blu scuro la 3C10 (HDAC7/TSA); in giallo la 2VQM (HDAC8/acido idrossammico). Gli eteroatomi del ligando sono colorati sulla base dei tipi atomici (N: blu ; O: rosso).

L'analisi bioinformatica degli inibitori cristallizzati in complesso con gli HDAC, ha permesso di individuare ulteriori punti farmacoforici rilevanti per l'inibizione dell'enzima (Figura 4) e, di conseguenza, per la progettazione di una nuova classe di inibitori quali:

- conformazione del CAP ad *ali di farfalla*, o *flip a V*, tale da occupare almeno una delle due posizioni dei gruppi carbonilici individuati;
- accomodamento dei vari sostituenti che caratterizzano il CAP in canali (ovvero tasche di superficie dell'enzima) specifici, in particolare, il sostituito ingombrante deve posizionarsi nel canale 3, l'anello benzodiazepinico *benzo-fuso* nel canale 1 e il sostituito meno ingombrante nel canale 2.



**Figura 4:** la figura mostra il modello farmacoforico 3D di un inibitore HDAC evidenziando la posizione dei sostituenti nei due canali, la posizione dei gruppi donatori e accettori di legame idrogeno del CAP, il linker disteso, e la posizione dello ZBG. In questa figura è mostrata la sovrapposizione del substrato (celeste; PDB ID: 2V5W) sull'inibitore della struttura PDB 2V5X (grigio). Gli eteroatomi del ligando sono colorati sulla base dei tipi atomici (N: blu ; O: rosso).

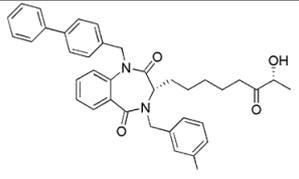
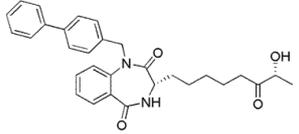
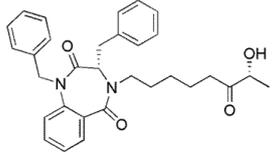
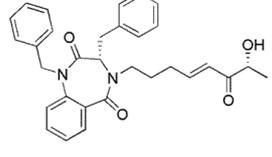
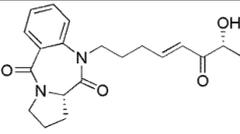
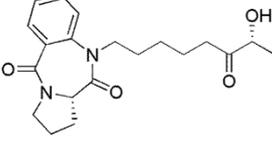
### 3.5.1c Analisi delle Relazioni Struttura-Attività (SARs) dei nuovi derivati benzodiazepinici

Parallelamente agli studi bioinformatici, i composti progettati sono stati sottoposti a saggi di citotossicità su linee cellulari di leucemia promieloide U937 e le molecole sono state testate in DMSO, in tempi diversi. Di seguito sono riportati (Tabella 1) gli inibitori considerati con le corrispondenti attività  $IC_{50}$  in DMSO solvente a 24h, 48h e 72h.

**Tabella 1:** attività  $IC_{50}$  degli inibitori nel solvente DMSO a 24,48 e 72h.

Composto	Struttura	Citotossicità U937 48h ( $\mu$ M)	$IC_{50}$ ( $\mu$ M) valori in DMSO solvente		
			24h	48h	72h
RR_89		14	NA	13	72

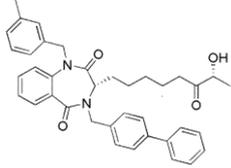
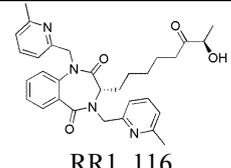
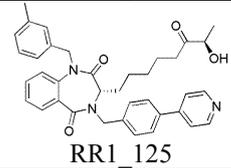
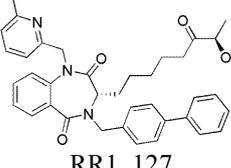
RR_67		0.04	NA	85	NT
RR_6		NT	NT	NT	NT
RR_84		75	NA	76	NT
RR_62		90	NA	74	NT
RR_90		LS <sup>c</sup>	NT	NT	NT
RR1_125		NT	NT	NA	NA
EC_332		35	NA	24	NT
EC_233		80	NA	89	V
RR1_116		NT	NT	NA	NA
RR1_127		NT	NT	114	4.8

<b>RR_94</b>		LS	NA	81	NT
<b>RR_96</b>		65	NA	52	NT
<b>EC_226</b>		NT	NA	53	NT
<b>EC_229</b>		36	NA	57	NT
<b>RR_11</b>		NT	NT	NT	NT
<b>EC_190</b>		NT	NT	NT	NT

NA = non attiva; NT= non testata, LS= scarsa solubilità

Putroppo, dai saggi effettuati è emerso un problema di solubilità dei nuovi composti, probabilmente dovuto alla loro capacità di formare aggregati molecolari. Di conseguenza sono stati condotti ulteriori studi per valutare l'attività delle molecole selezionate in diversi solventi ( *Tabella 2*).

**Tabella2:** attività IC<sub>50</sub> degli inibitori in DMSO a 24, 48 e 72h; in MetOh a 48 e 72h; in PBS1x a 48 e 72h; in PBS/MetOh a 48 e 72h

Struttura	IC <sub>50</sub> μM DMSO			IC <sub>50</sub> μM MetOH		IC <sub>50</sub> μM PBS1x		IC <sub>50</sub> μM PBS/MetOH	
	24h	48h	72h	48h	72h	48h	72h	48h	72h
 <p>RR 89</p>	NA	13	72	53	36	NT	NT	NT	NT
 <p>RR1 116</p>	NA	NA	NA	NA	30	NA	NA	>10 <sup>-3</sup>	77
 <p>RR1 125</p>	NA	NA	NA	NA	22	NA	NA	>10 <sup>-3</sup>	31
 <p>RR1 127</p>	NT	114	4.8	45	1.2	>100	>100	>10 <sup>-3</sup>	28

NA=non attiva; NT = non testata

Contemporaneamente, le stesse molecole, sono state sottoposte a saggi di inibizione sulle diverse isoforme umane di HDAC, utilizzando differenti substrati. (Tabella 3) Purtroppo, anche in questo caso, la scarsa solubilità dei composti ha fortemente limitato la corretta valutazione della loro attività biologica.

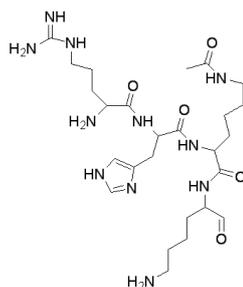
**Tabella 3:** attività Kd degli inibitori nelle diverse isoforme enzimatiche

Composti	Classe I HDAC Kd $\mu$ M				Classe IIa HDAC Kd $\mu$ M				Classe IIb HDAC Kd $\mu$ M		Classe IV HDAC Kd $\mu$ M
	1	2	3	8	4	5	7	9	6	10	11
RR_89	39.68	16.29	27.35	NA	NA	NA	NA	NA	96.36	ND	ND
	NA	NA	NA	82.3	NA	NA	NA	NA	93.7	NA	NA
RR1_125	12.447	6.064	9.785	12.407	47.430	12.187	24.797	NA	1.164	NA	NA
RR1_127	NA	9.774	25.622	85.185	NA	NA	NA	NA	4.989	NA	NA
RR1_116	NA	12.015	30.128	67.685	NA	NA	NA	NA	7.494	NA	NA
TSA	0.0017	0.007	0.011	0.373	3.150	0.830	2.884	4.454	0.001	0.001	0.003

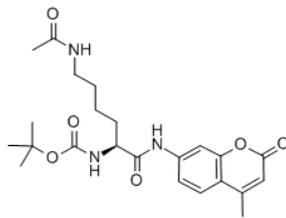
NA= non attiva; ND= non determinata

Le attività riportate nella successiva *Tabella 3* sono state evidenziate a seconda del tipo di substrato utilizzato nei saggi condotti dalla *Reaction Biology*. In particolare, le attività in giallo riguardano i risultati ottenuti utilizzando il substrato generale, il peptide fluorogenico del residuo 379-382 di p53 RHKK(Ac) (*Figura 5*); per quelle evidenziate in rosso è stato utilizzato il substrato specifico per la classe IIa, il substrato fluorogenico per HDAC di classe IIa Boc-Lys (trifluoroacetil)-AMC (*Figura 6*); per le attività in magenta è stato utilizzato il substrato per l'HDAC8, il peptide fluorogenico del residuo 379-382 di p53 RHK(Ac)K(Ac) (*Figura 7*).

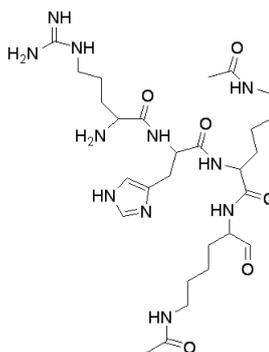
I valori in grassetto rappresentano i risultati ottenuti dal Dott. Ralph Mazitschek dell'Harvard Medical School, che ha utilizzato differenti substrati: il substrato tripeptidico acetilato per gli HDAC1,2,3,e 6 (*Figura 8*); il substrato tripeptidico trifluoro-acetil- lisinico per gli HDAC 4,5,7,8 e 9 (*Figura 9*). Di seguito sono riportate le strutture di tutti i substrati con cui sono state saggiate le attività delle molecole progettate.



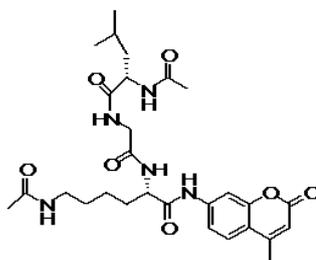
**Figura 5:** è rappresentata la struttura del substrato generale Arg-His-Lys-Lys(Ac)



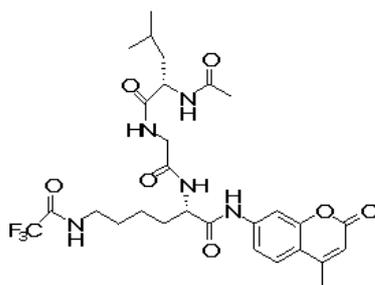
**Figura 6:** è rappresentata la struttura del substrato fluorogenico per la classe Hdac IIa, Boc-Lys (trifluoroacetil)-AMC



**Figura 7:** è rappresentata la struttura del substrato per l'Hdac8, peptide fluoro genico, Arg-His-Lys(Ac)-Lys(Ac)



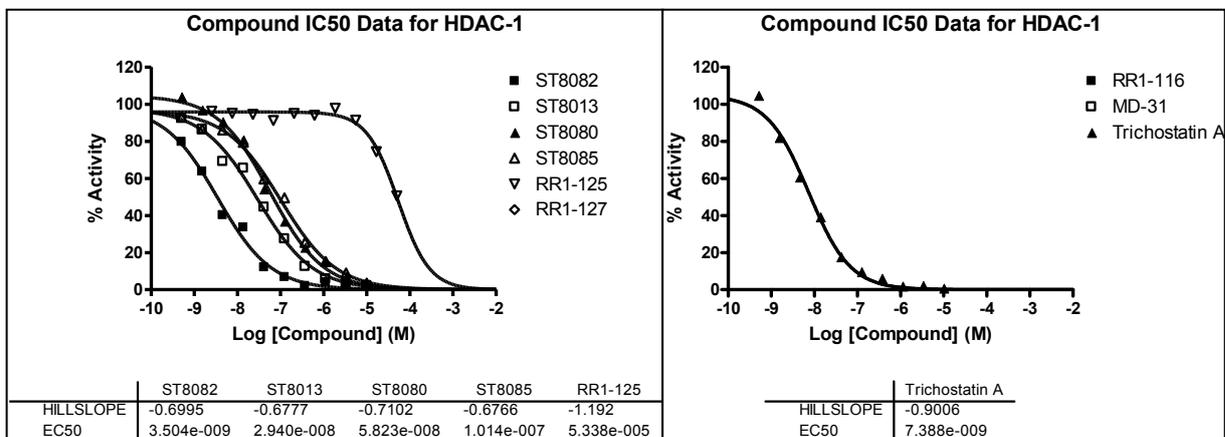
**Figura 8:** è rappresentata la struttura del substrato tripeptidico acetilato utilizzato per gli HDAC 1, 2, 3, e 6



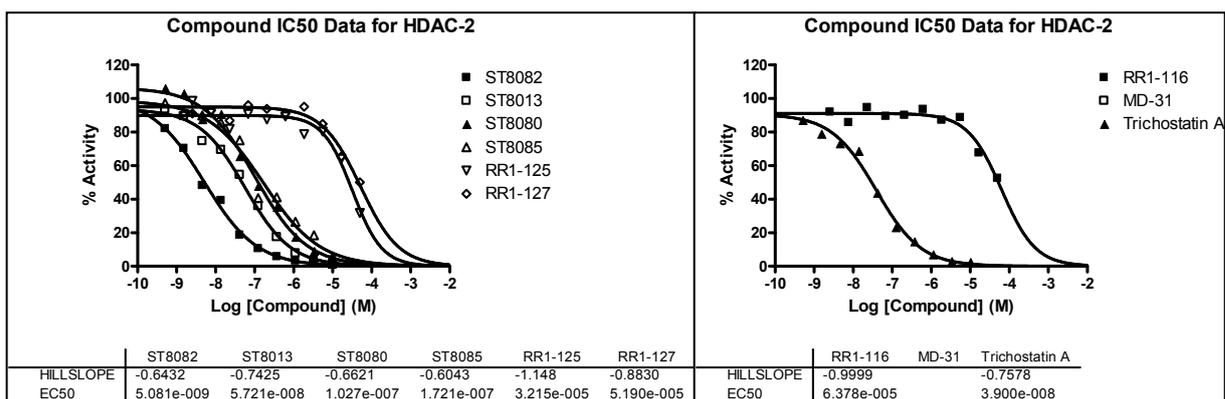
**Figura 9:** è rappresentata la struttura del substrato tripeptidico trifluoroacetillisinico utilizzato per gli HDAC 4, 5, 7, 8 e 9

Le attività di ciascuna molecola sono state ricavate dai grafici seguenti.

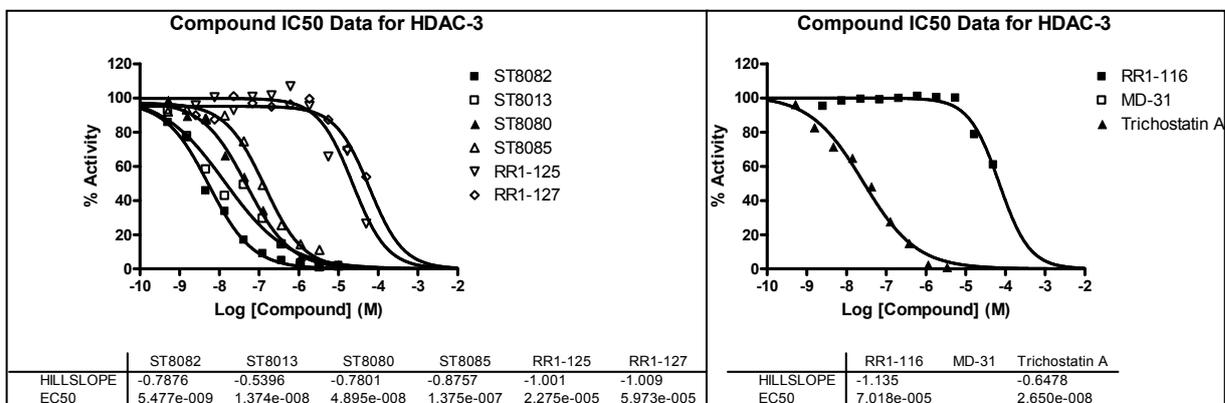
**Grafico HDAC1:** sull'asse delle ascisse è riportato il Log della concentrazione del composto espressa in moli e sull'asse delle ordinate la percentuale di attività.



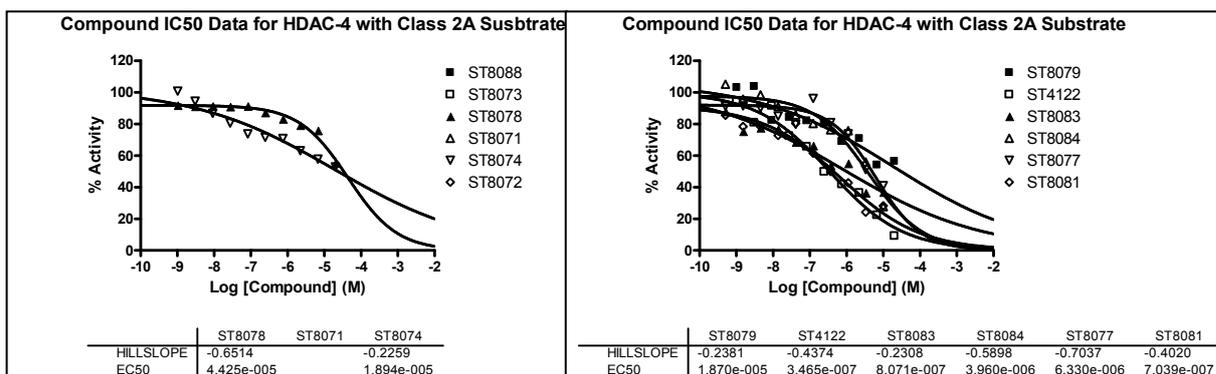
**Grafico HDAC2:** sull'asse delle ascisse è riportato il Log della concentrazione del composto espressa in moli e sull'asse delle ordinate la percentuale di attività.



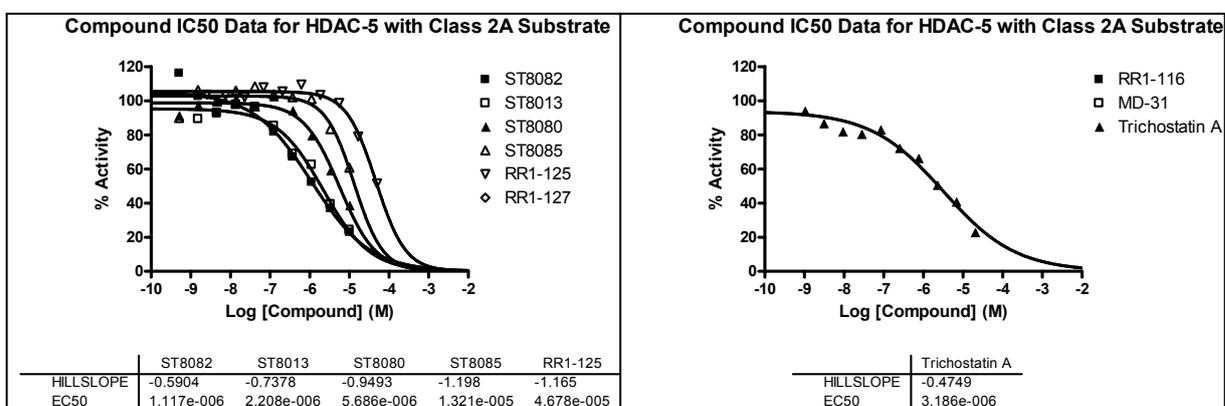
**Grafico HDAC3:** sull'asse delle ascisse è riportato il Log della concentrazione del composto espressa in moli e sull'asse delle ordinate la percentuale di attività.



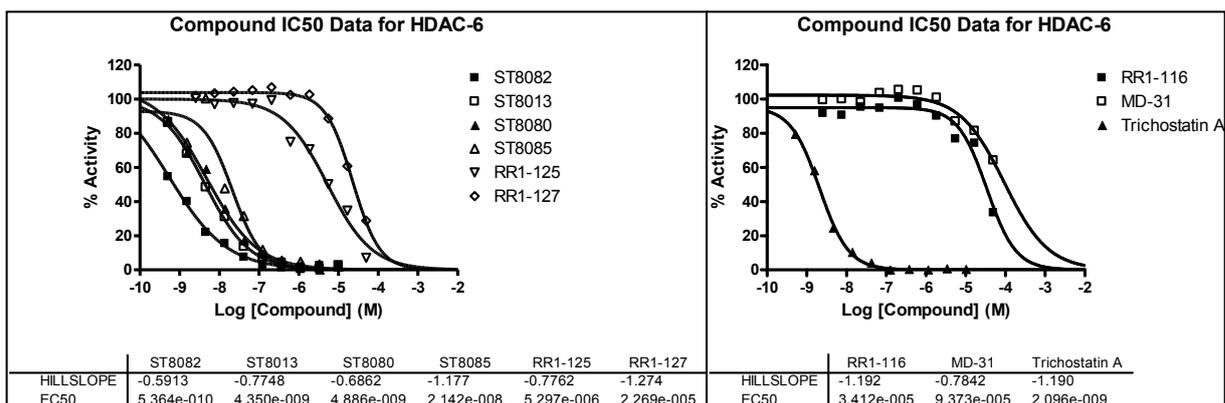
**Grafico HDAC4:** sull'asse delle ascisse è riportato il Log della concentrazione del composto espressa in moli e sull'asse delle ordinate la percentuale di attività.



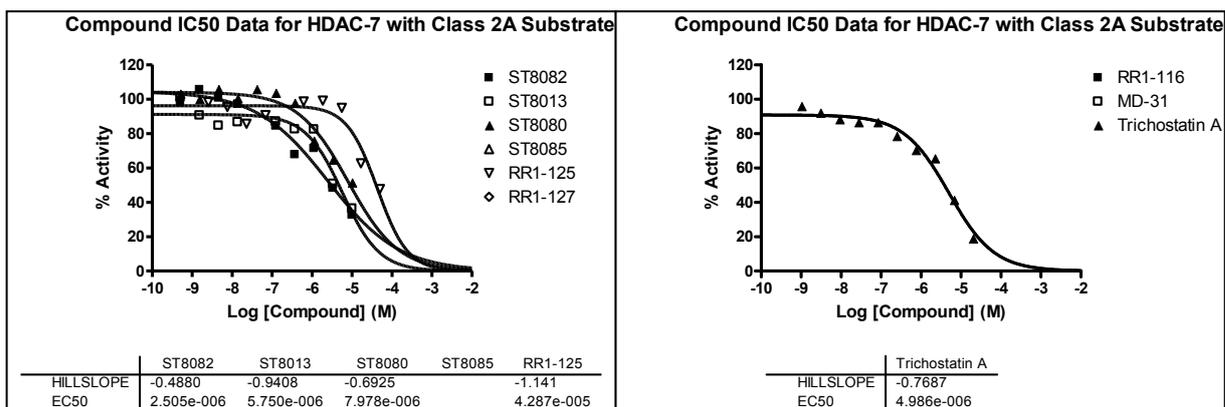
**Grafico HDAC5:** sull'asse delle ascisse è riportato il Log della concentrazione del composto espressa in moli e sull'asse delle ordinate la percentuale di attività.



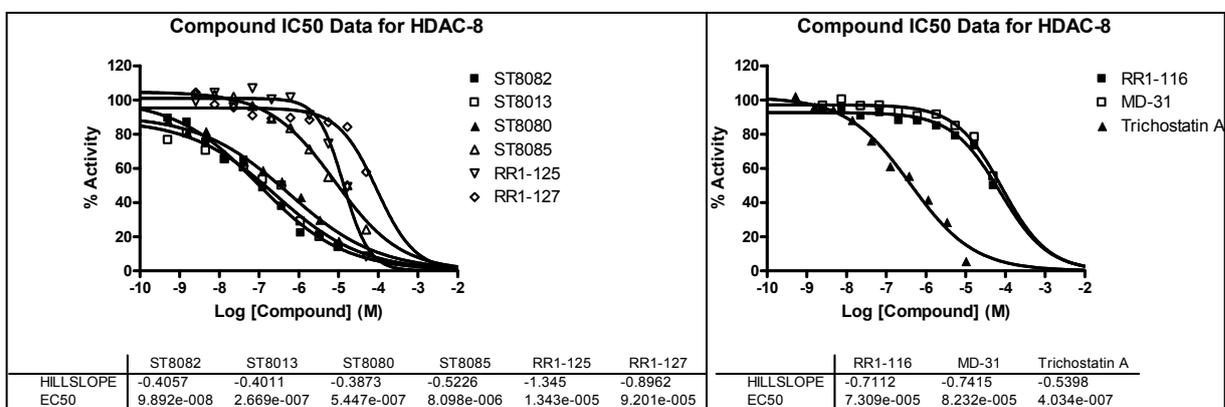
**Grafico HDAC6:** sull'asse delle ascisse è riportato il Log della concentrazione del composto espressa in moli e sull'asse delle ordinate la percentuale di attività.



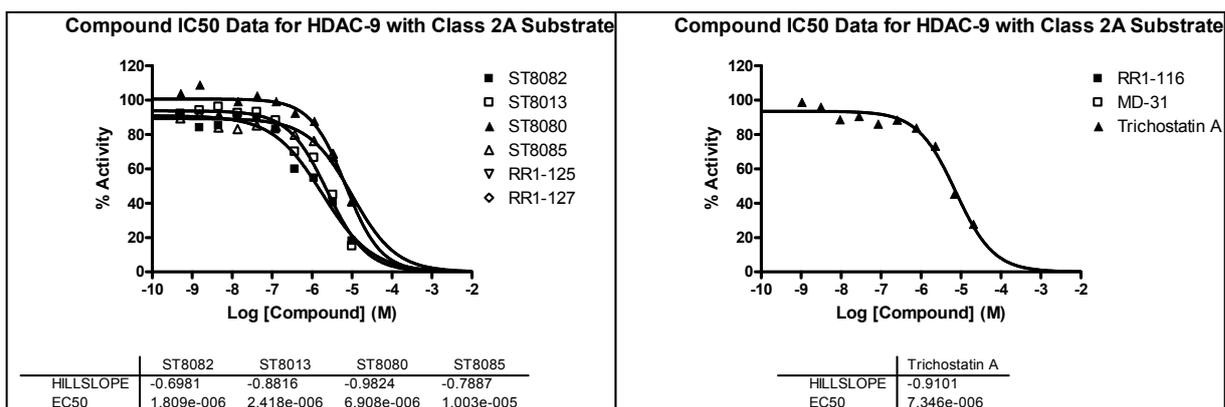
**Grafico HDAC7:** sull'asse delle ascisse è riportato il Log della concentrazione del composto espressa in moli e sull'asse delle ordinate la percentuale di attività.



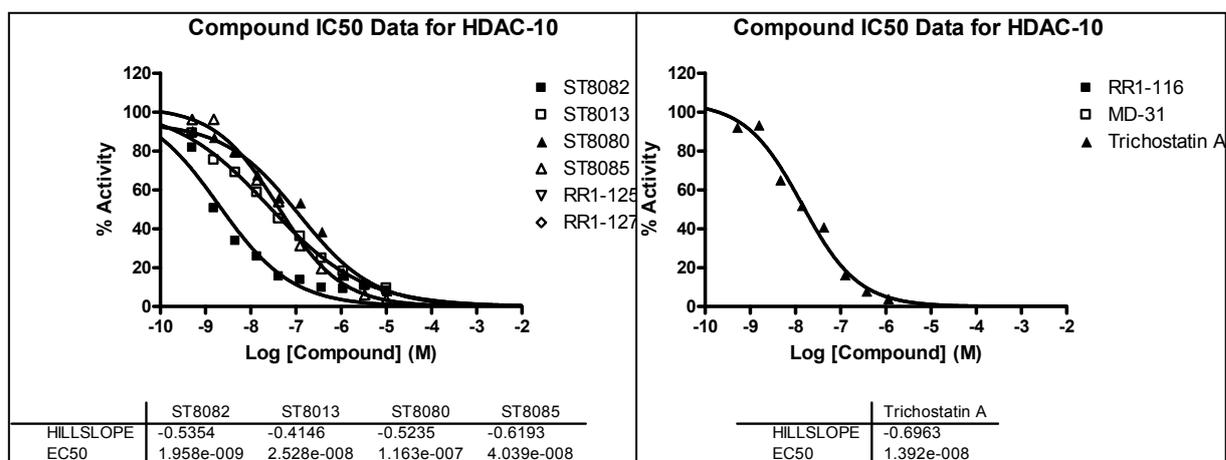
**Grafico HDAC8:** sull'asse delle ascisse è riportato il Log della concentrazione del composto espressa in moli e sull'asse delle ordinate la percentuale di attività.



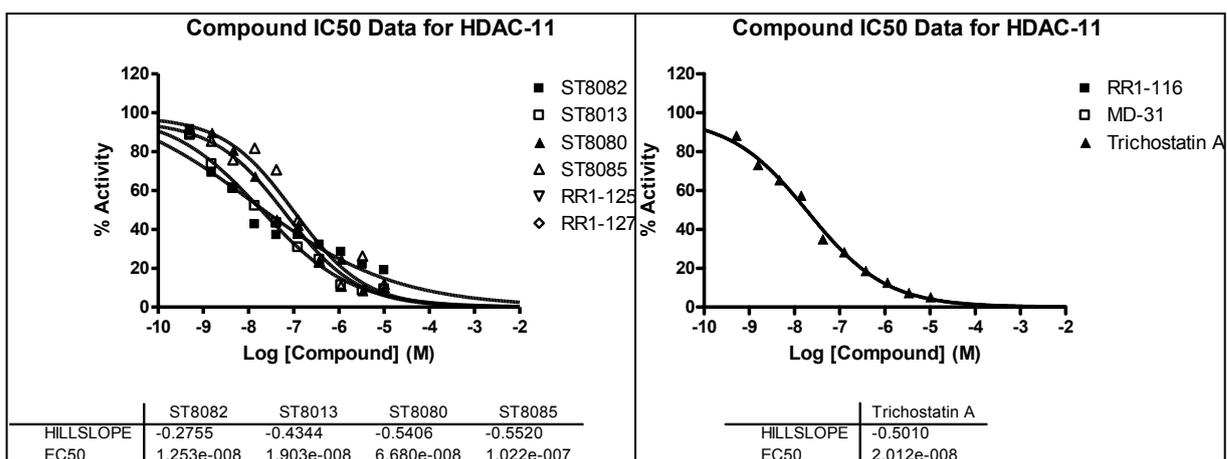
**Grafico HDAC9:** sull'asse delle ascisse è riportato il Log della concentrazione del composto espressa in moli e sull'asse delle ordinate la percentuale di attività.



**Grafico HDAC10:** sull'asse delle ascisse è riportato il Log della concentrazione del composto espressa in moli e sull'asse delle ordinate la percentuale di attività.



**Grafico HDAC11:** sull'asse delle ascisse è riportato il Log della concentrazione del composto espressa in moli e sull'asse delle ordinate la percentuale di attività.



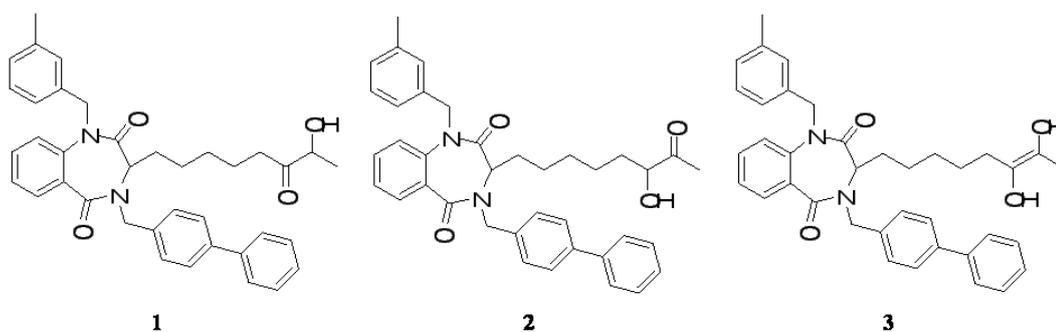
I composti sono stati testati in 10 dosi IC<sub>50</sub> con tre diluizioni seriali a partire da 50 µM.

Il composto di controllo, la Tricostatina A (TSA), è stato testato in 10 dosi IC<sub>50</sub> con tre diluizioni seriali a partire da 10 µM, mentre per la classe IIa è stato testato in 10 dosi IC<sub>50</sub> con tre diluizioni seriali a partire da 20 µM. Durante l'esecuzione del saggio, i composti non hanno mostrato alcun fondo fluorescente che potesse interferire con i saggi. La disparità dei risultati ottenuti dai saggi condotti dalla *Reaction Biology* e quelli dell'Harvard Medical School è correlata ai differenti substrati utilizzati. Il substrato inoltre può legarsi a stati conformazionalmente diversi. Lo spostamento dell'equilibrio verso una forma piuttosto che un'altra giustifica le differenti attività.

È plausibile affermare che queste attività siano sottostimate, a causa della scarsa solubilità delle molecole. Inoltre il fatto che i saggi effettuati dall'Harvard Medical School per uno

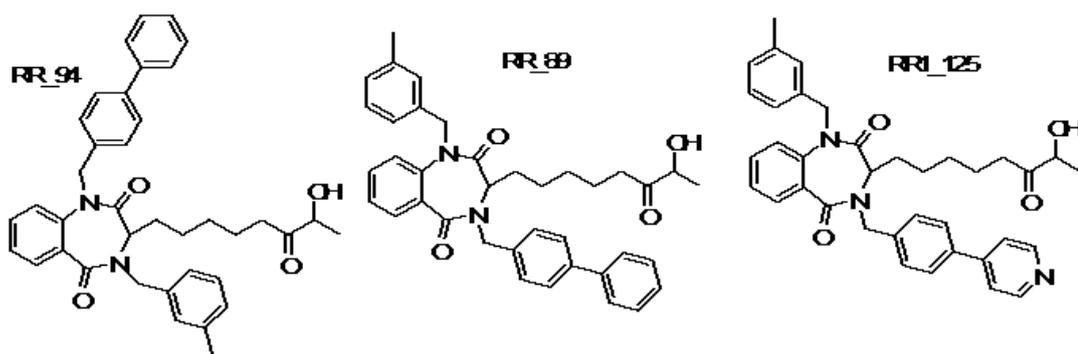
stesso composto determinino due differenti valori di attività, all'interno della stessa isoforma enzimatica, lascia pensare che la metodologia utilizzata, per superare i problemi della formazione di aggregati e aumentare la solubilità, non sia sempre la stessa. Ciò porta ovviamente a risultati differenti.

Dall'analisi di questi studi condotti sull'attività e sulla solubilità di tutte le strutture nei diversi substrati e solventi, è emersa una linearità di risultati tra la RR\_89 e la RR1\_125, in quanto quest'ultima è risultata sempre più attiva, e ha raggiunto una  $IC_{50}$  di circa  $10\mu M$  nelle uniche isoforme verso le quali la RR\_89 è attiva. Questo dato potrebbe implicare il fatto che una migliore attività possa essere legata, in questo caso, semplicemente a una migliore solubilità. A conferma di ciò la RR1\_127, più solubile della RR\_89, ma leggermente meno della RR1\_125, ha un'attività intermedia. L'inserimento di un azoto nel CAP porterebbe quindi a una migliore attività in generale. Al contrario la RR1\_116, che presenta due azoti aromatici, pur essendo più solubile, ha mostrato un profilo farmacologico simile a quello di RR1\_127. Questo porterebbe a pensare che il metile in meta dà problemi di ingombro sterico, o che è necessaria la presenza di un secondo fenile in para per ottenere un incremento di attività. È inoltre interessante notare come lo “scambio” dei due sostituenti aromatici sui due azoti dello scheletro benzodiazepinico di RR\_89 porti ad un composto, RR\_94, completamente inattivo. L'analisi del pKa (vedi parte sperimentale), indica la presenza di tre forme tautomeriche possibili (Figura 10).



**Figura 10:** la figura mostra le tre forme tautomeriche dell'RR\_89. Dalla sinistra il tautomero 1, il 2 e il 3.

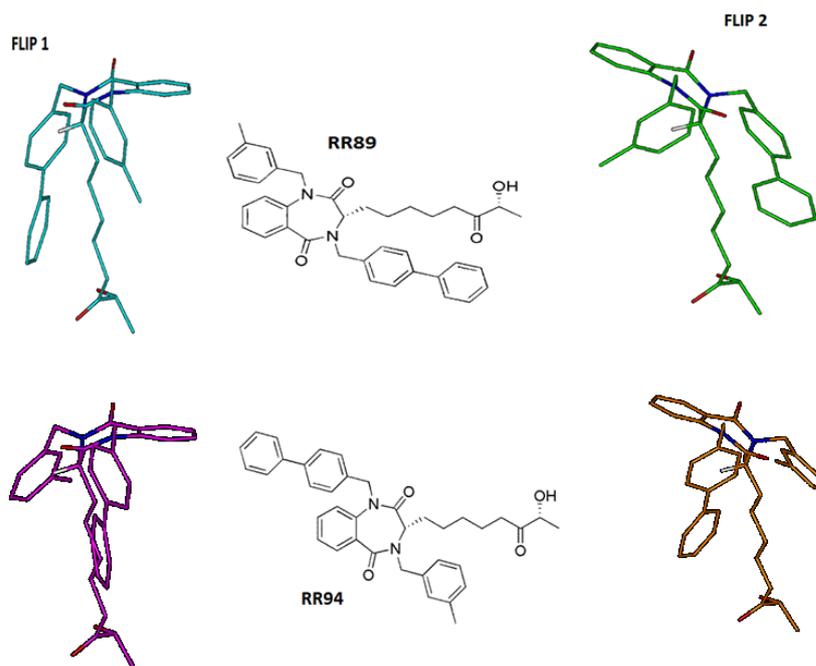
I risultati dell'analisi di strutture cristallografiche note, così come quelli del *pharmacophore fitting*, convergono nell'indicare che il tautomero da considerare nei miei studi è il tautomero 1. In seguito all'analisi dei ligandi e agli studi descritti precedentemente, il mio interesse si è rivolto su tre strutture in particolare, ovvero tre 1,4 benzodiazepine: l'RR\_89 e l'RR\_94, che presentano gli stessi sostituenti sul CAP, ma disposti in posizioni opposte tra loro, e la RR1\_125, che differisce dalla RR\_89 per la presenza di un azoto in para sul sostituente più ingombrante del CAP (Figura 11).



**Figura 11:** strutture delle benzodiazepine RR\_94, RR\_89 e RRI\_125.

### 3.5.1d Studi di Pharmacophore Fitting

Lo studio di tutte le possibili geometrie che le tre strutture possono assumere e di tutte le possibili posizioni spaziali che i sostituenti del CAP occupano all'interno della tasca catalitica, è stato indirizzato all'individuazione della conformazione responsabile del legame con l'enzima (conformazione bioattiva). Come detto nei paragrafi precedenti, sono stati considerati i due possibili *flip* del sistema benzodiazepinico. Infatti, è stato possibile costruire due *flip* uguali per gli inibitori. Quest'ultimi infatti presentano due possibili geometrie: un *flip* 1 che presenta una disposizione pseudo-assiale e un *flip* 2 caratterizzato da una disposizione pseudo-equatoriale. Un'analisi delle energie che caratterizzano le due conformazioni porta alla luce che il *flip* 1, quello con disposizione pseudo-assiale, ha una minore energia ed è quindi più stabile (*Figura 12*).



**Figura 12:** strutture e flip. Al centro sono mostrate le strutture bidimensionali delle benzodiazepine RR\_89 e RR\_94; ai lati i due flip che queste molecole possono assumere. Gli eteroatomi del ligando sono colorati sulla base dei tipi atomici (N: blu ; O: rosso e H:bianco).

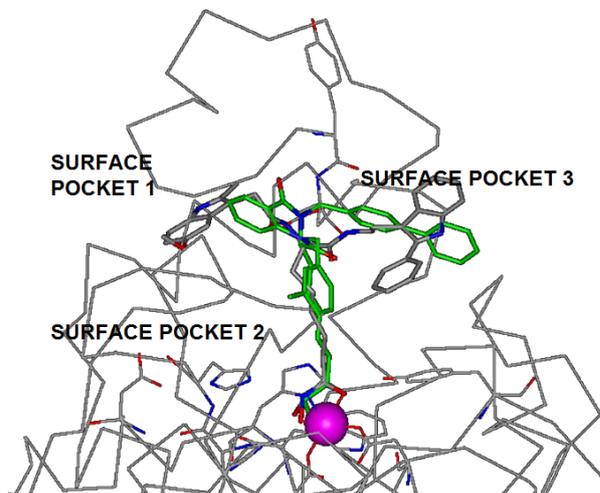
A questo punto le benzodiazepine sono state poste all' interno dell' enzima per studiarne l'accomodamento. L' isoforma utilizzata è la PDB ID: 2V5X, appartenente alla famiglia degli HDAC 8 di classe I.

La RR\_89, e di conseguenza la RR1\_125, con *flip* pseudo-equatoriale (*flip* 2), rispettano perfettamente il modello farmacoforico: gli ossigeni carbonilici del CAP interagiscono con le molecole d'acqua presenti all' interno del sito; l'anello benzodiazepinico è diretto nel canale 1; il sostituito più ingombrante va nel canale 3, quello con dimensione ridotta nel canale 2; lo ZBG riproduce le posizioni occupate dai coordinanti della PDB in esame.

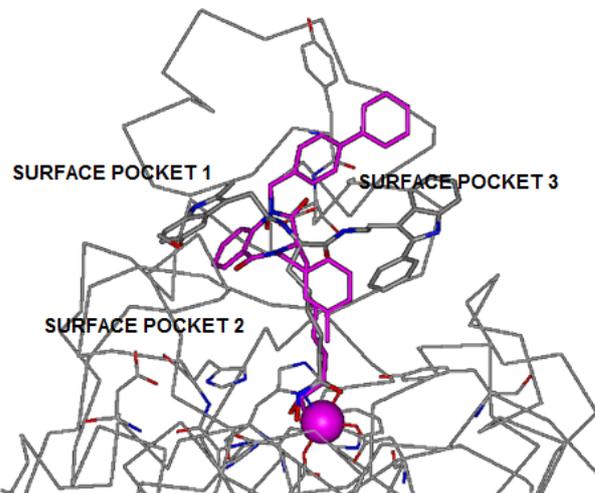
Diversa è invece la situazione per quanto riguarda il composto RR\_94. Il *flip* 2, che nella RR\_89 e nella RR1\_125 è caratterizzato da un buon accomodamento, presenta il sostituito più ingombrante diretto nel canale 2. Questa situazione è sfavorevole, in quanto i punti farmacoforici non sono rispettati e si perde l'interazione con l'enzima. Il *flip* 1, che si presenta perfettamente speculare al *flip*2 della RR\_89 e della RR1\_125, presenta la possibilità di sovrapporre i sostituito allo stesso modo. L'anello benzodiazepinico però, in questa conformazione, determina ingombri sterici perché si schiaccia verso il basso nella proteina e il sostituito più ingombrante risente dell'ingombro sterico del gruppo carbonilico della catena

laterale dell'Asp101, presente all'interno del canale. L'unico accomodamento che sembra essere ottimale è quello del sostituente meno ingombrante.

Quest'analisi spaziale conferma quindi il dato della minore attività dell'RR\_94 rispetto all'RR\_89 (Figura 13 e 14).



**Figura 13:** flip 2 dell'RR\_89 nell'HDAC8



**Figura 14:** flip 1 dell'RR\_94 nell'HDAC8

Grazie a questi risultati è auspicabile la sintesi di nuovi inibitori selettivi di HDAC, combinando i diversi sostituenti, le diverse posizioni e finalizzando gli studi a un miglioramento della potenza e dell'affinità di questi farmaci per le varie isoforme enzimatiche.

### ***3.5.1e Analisi degli Studi di Docking***

Gli studi di *docking* sono stati condotti utilizzando come punto di partenza una disposizione dell'inibitore, nel sito catalitico, tale da rispettare tutti i punti farmacoforici, importanti per una buona interazione tra ligando e proteina. In particolare è stata utilizzata una coordinazione dello zinco come quella dell'inibitore dell'HDAC8 (PDB ID: 2V5X). Prima di arrivare a questo risultato sono stati fatti vari tentativi per ottenere una geometria ottimale dei coordinanti. Quest'ultimi ad esempio sono stati ruotati di 180°, ottenendo però compenetrazioni steriche con la proteina. Inoltre, il secondo e quarto torsionale del *linker* sono stati mantenuti in trans, il resto del *linker* è stato lasciato in una conformazione distesa, il CAP è stato posizionato in modo tale da occupare con i gruppi carbonilici, una delle posizioni occupata dal carbonile di ligandi noti e sono stati disposti: l'anello benzodiazepinico nel canale 1, il sostituente ingombrante nel canale 3 e quello meno ingombrante nel canale 2. I

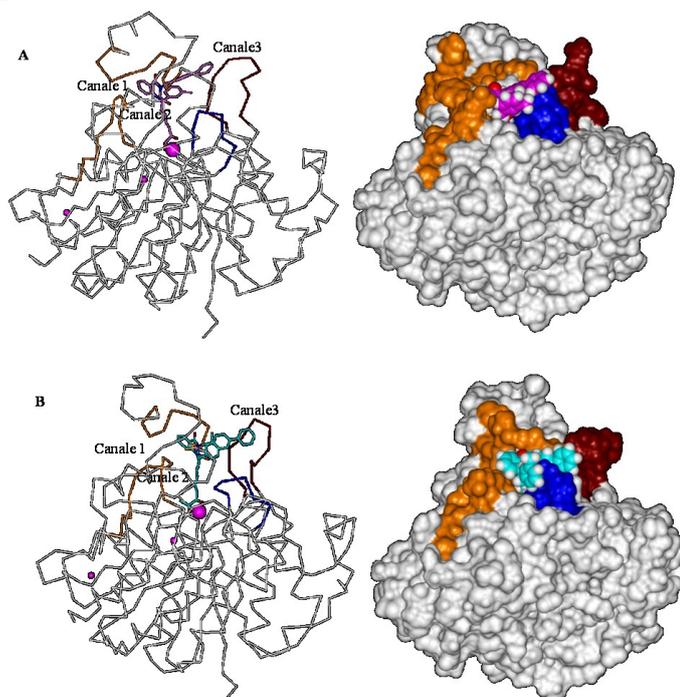
risultati dei *docking* sono stati selezionati utilizzando lo stesso criterio adottato per il punto di partenza.

Sono di seguito riportati i risultati degli studi di docking del composto RR\_89 effettuati sulle diverse isoforme HDAC.

## HDAC 2

**Tabella 4:** sono riportati i criteri di selezione dei *docking*; in particolare sono stati valutati: gli angoli diedri ( $\tau_1, \tau_2$ ), le distanze dei coordinanti (d1,d2) e l'angolo dello ZBG rispetto allo zinco, l'accomodamento dei sostituenti nei canali, le posizioni dei coordinanti e le posizioni dei gruppi carbonilici del CAP.

Frame docking	$\tau_1$	$\tau_2$	d1	d2	Angolo	Canali sostituenti			Posizioni coordinanti	Posizioni CO
						Anello benzodiazepinico	R	R'		
CAR	-45.04	-9.95	2.24	1.67	91.36	1	2	3	2V5X	2V5X
FR2	39.97	-5.82	2.55	2.41	68.27	1	2	3	2V5X	2VQM
FR17	28.65	16.38	2.52	2.41	67.36	1	2	3	2V5X	2V5X

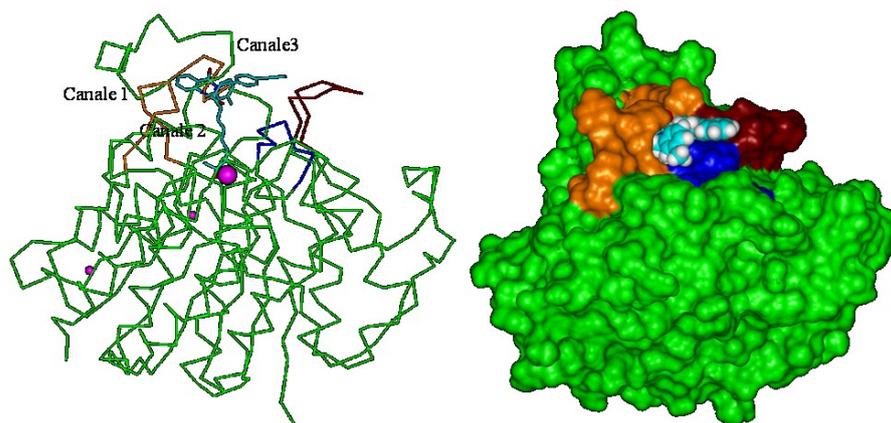


**Figura 15:** RR\_89 nell'HDAC2. A) nella figura A è mostrato il risultato 2(magenta) della RR\_89 nell'HDAC2; a destra lo stesso risultato è rappresentato con la proteina in superfici di connoly e il ligando in CPK. B) nella figura B è mostrato il risultato 17(ciano) della RR\_89 nell'HDAC2, a destra lo stesso risultato è rappresentato con la proteina in superfici di connoly e il ligando in CPK. In arancione è mostrato il canale 1, in blu ed arancio il canale 2 e il blu e rosso il canale 3. Gli eteroatomi dei ligandi sono colorati sulla base dei tipi atomici (N: blu e O: rosso).

## HDAC3

**Tabella 5:** sono riportati i criteri di selezione dei *docking*; in particolare sono stati valutati: gli angoli diedri ( $\tau_1, \tau_2$ ), le distanze dei coordinanti (d1,d2) e l'angolo dello ZBG rispetto allo zinco, l'accomodamento dei sostituenti nei canali, le posizioni dei coordinanti e le posizioni dei gruppi carbonilici del CAP.

Frame docking	$\tau_1$	$\tau_2$	d1	d2	Angolo	Canali sostituenti			Posizioni coordinanti	Posizioni CO
						Anello benzodiazepinico	R	R'		
CAR	-45.73	-12.72	2.18	1.85	88.80	1	2	3	2V5X	2V5X
FR9	-25.95	9.46	2.48	2.46	67.83	1	2	3	2V5X	2V5X

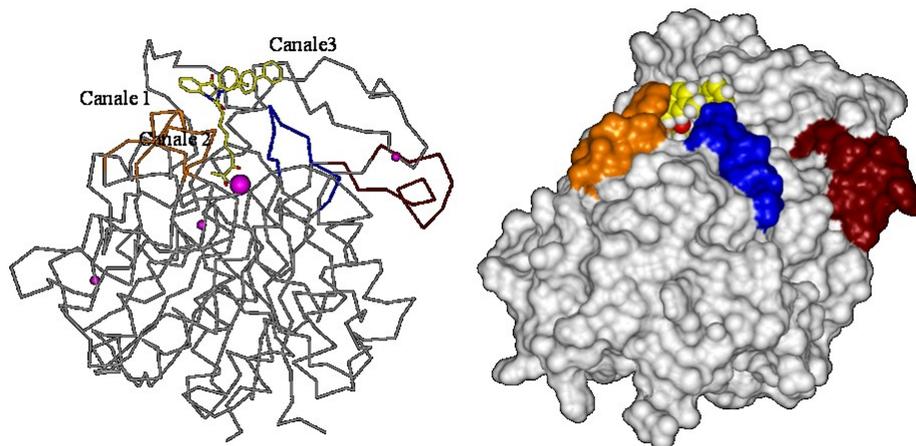


**Figura 16:** RR\_89 nell'HDAC3. Nella figura è mostrato il risultato 9 della RR\_89 nell'HDAC3; a destra lo stesso risultato è rappresentato con la proteina in superfici di connoly e il ligando in CPK. In arancione è mostrato il canale 1, in blu ed arancio il canale 2 e il blu e rosso il canale 3. Gli eteroatomi del ligando sono colorati sulla base dei tipi atomici (N: blu; C: celeste e O: rosso).

## HDAC4

**Tabella 6:** sono riportati i criteri di selezione dei *docking*; in particolare sono stati valutati: gli angoli diedri ( $\tau_1, \tau_2$ ), le distanze dei coordinanti (d1,d2) e l'angolo dello ZBG rispetto allo zinco, l'accomodamento dei sostituenti nei canali, le posizioni dei coordinanti e le posizioni dei gruppi carbonilici del CAP.

Frame docking	$\tau_1$	$\tau_2$	D1	D2	Angolo	Canali sostituenti			Posizioni coordinanti	Posizioni CO
						Anello benzodiazepinico	R	R'		
CAR	-52.81	-9.91	2.40	2.30	73.79	1	2	3	2V5X	2V5X
FR12	-12.74	20.20	2.39	2.31	67.91	1	2	3	1W22	-

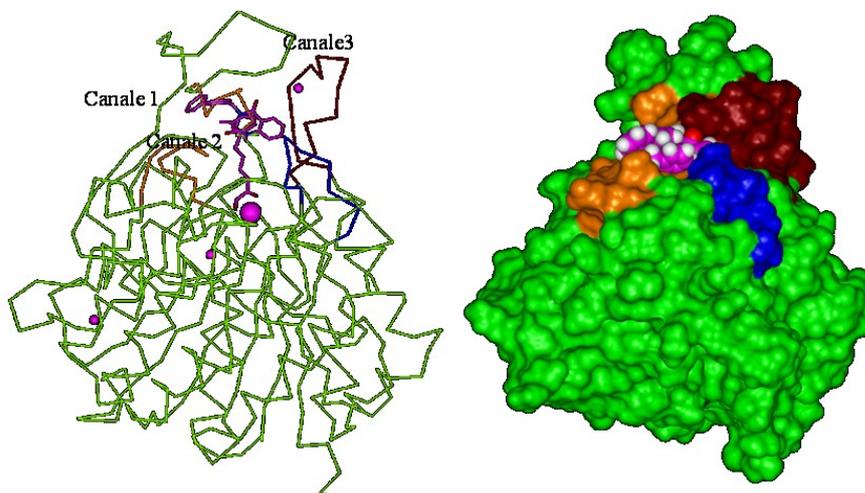


**Figura 17:** RR\_89 nell'HDAC4. Nella figura è mostrato il risultato 12 della RR\_89 nell'HDAC4; a destra lo stesso risultato è rappresentato con la proteina in superfici di connoly e il ligando in CPK. In arancione è mostrato il canale 1, in blu ed arancio il canale 2 e il blu e rosso il canale 3. Gli eteroatomi del ligando sono colorati sulla base dei tipi atomici (N: blu; C: giallo e O: rosso).

### HDAC7: nessun risultato ottimale

**Tabella 7:** sono riportati i criteri di selezione dei *docking*; in particolare sono stati valutati: gli angoli diedri ( $\tau_1, \tau_2$ ), le distanze dei coordinanti (d1, d2) e l'angolo dello ZBG rispetto allo zinco, l'accomodamento dei sostituenti nei canali, le posizioni dei coordinanti e le posizioni dei gruppi carbonilici del CAP.

Frame docking	$\tau_1$	$\tau_2$	d1	d2	Angolo	Canali sostituenti			Posizioni coordinanti	Posizioni CO
						Anello benzodiazepinico	R	R'		
CAR	-21.83	-50.21	2.68	1.85	74.57	1	2	3	2V5X	2V5X
Fr10	-21.91	-4.32	2.59	2.42	66.85	3	2	1	1W22	-

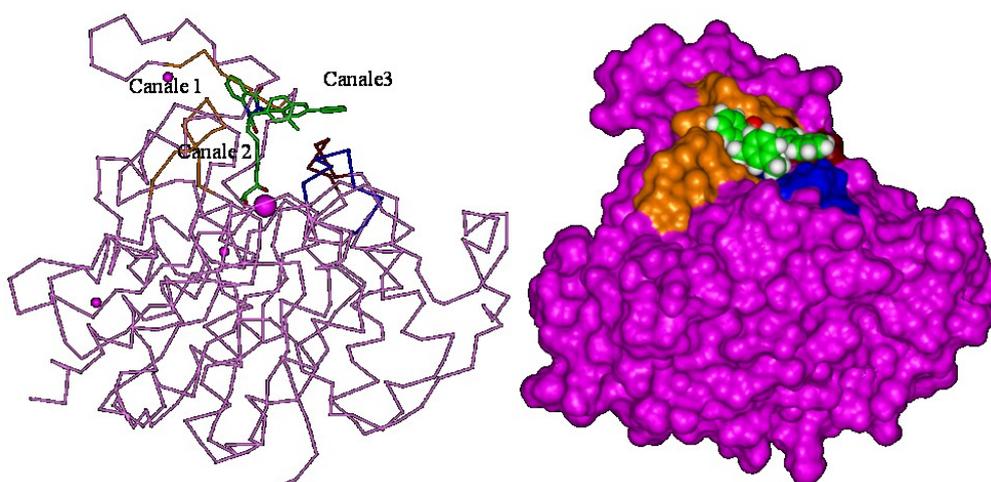


**Figura 18:** RR\_89 nell'HDAC7. Nella figura è mostrato il risultato 10 della RR\_89 nell'HDAC7; a destra lo stesso risultato è rappresentato con la proteina in superfici di connoly e il ligando in CPK. In arancione è mostrato il canale 1, in blu ed arancio il canale 2 e il blu e rosso il canale 3. Gli eteroatomi del ligando sono colorati sulla base dei tipi atomici (N: blu; C: magenta e O: rosso).

## HDAC8

**Tabella 8:** sono riportati i criteri di selezione dei *docking*; in particolare sono stati valutati: gli angoli diedri ( $\tau_1, \tau_2$ ), le distanze dei coordinanti (d1,d2) e l'angolo dello ZBG rispetto allo zinco, l'accomodamento dei sostituenti nei canali, le posizioni dei coordinanti e le posizioni dei gruppi carbonilici del CAP.

Frame docking	$\tau_1$	$\tau_2$	D1	D2	Angolo	Canali sostituenti			Posizioni coordinanti	Posizioni CO
						Anello benzodiazepinico	R	R'		
CAR	-43.13	-17.33	2.15	1.90	88.12	1	2	3	2V5X	2V5X
FR1	0.95	-32.79	2.43	2.42	67.20	1	2	3	1W22	-



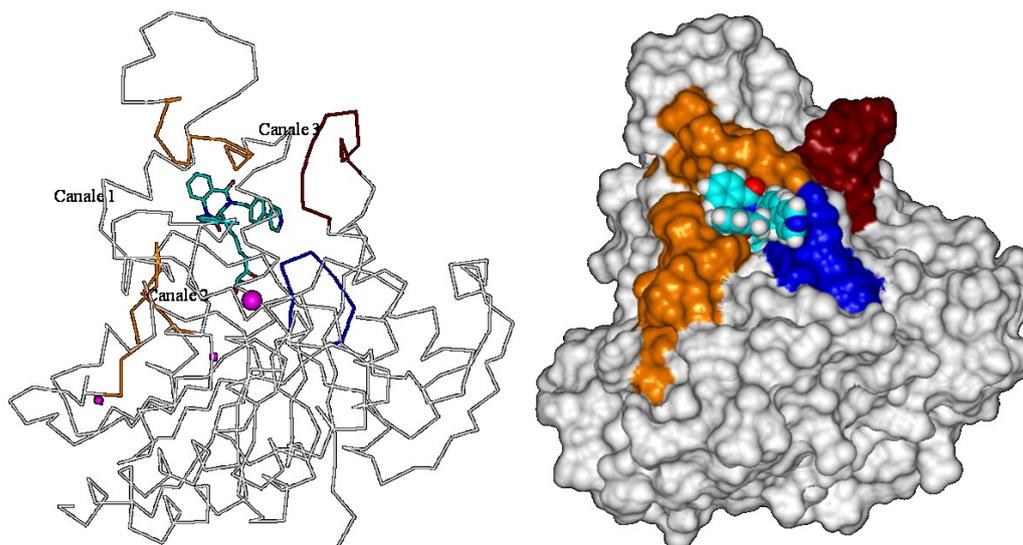
**Figura 19: RR\_89 nell'HDAC8.** Nella figura è mostrato il risultato 1 della RR\_89 nell'HDAC8; a destra lo stesso risultato è rappresentato con la proteina in superfici di connoly e il ligando in CPK. In arancione è mostrato il canale 1, in blu ed arancio il canale 2 e il blu e rosso il canale 3. Gli eteroatomi del ligando sono colorati sulla base dei tipi atomici (N: blu; C: verde e O: rosso).

Sono di seguito riportati i risultati degli studi di docking del composto RR1\_125 effettuati sulle diverse isoforme HDAC.

## HDAC2

**Tabella 9:** sono riportati i criteri di selezione dei *docking*; in particolare sono stati valutati: gli angoli diedri ( $\tau_1, \tau_2$ ), le distanze dei coordinanti (d1,d2) e l'angolo dello ZBG rispetto allo zinco, l'accomodamento dei sostituenti nei canali, le posizioni dei coordinanti e le posizioni dei gruppi carbonilici del CAP.

Frame docking	$\tau_1$	$\tau_2$	D1	D2	Angolo	Canali sostituenti			Posizioni coordinanti	Posizioni CO
						Anello benzodiazepinico	R	R'		
CAR	-26.07	16.24	2.65	2.37	65.88	1	2	3	2V5X	2V5X
FR15	-24.00	-14.99	2.63	2.51	64.26	1	2	3	2V5X	1W22

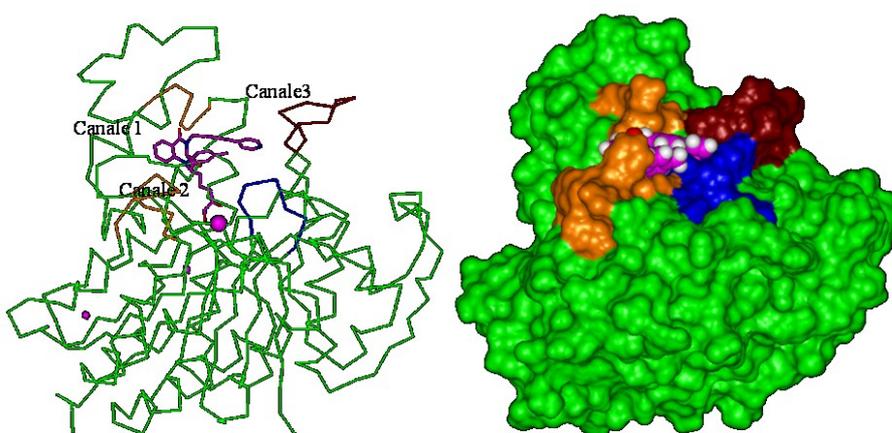


**Figura 20:** RR1\_125 nell'HDAC2. Nella figura è mostrato il risultato 15 della RR1\_125 nell'HDAC2; a destra lo stesso risultato è rappresentato con la proteina in superfici di connoly e il ligando in CPK. In arancione è mostrato il canale 1, in blu ed arancio il canale 2 e il blu e rosso il canale 3. Gli eteroatomi del ligando sono colorati sulla base dei tipi atomici ( N: blu; C:celestese e O: rosso).

### HDAC3

**Tabella 10:** sono riportati i criteri di selezione dei *docking*; in particolare sono stati valutati: gli angoli diedri ( $\tau_1, \tau_2$ ), le distanze dei coordinanti (d1,d2) e l'angolo dello ZBG rispetto allo zinco, l'accomodamento dei sostituenti nei canali, le posizioni dei coordinanti e le posizioni dei gruppi carbonilici del CAP.

Frame docking	$\tau_1$	$\tau_2$	D1	D2	Angolo	Canali sostituenti			Posizioni coordinanti	Posizioni CO
						Anello benzodiazepinico	R	R'		
CAR	-25.99	9.44	2.48	2.47	67.69	1	2	3	2V5X	2V5X
FR8	-34.74	0.31	2.55	2.53	65.19	1	2	3	2V5X	1W22

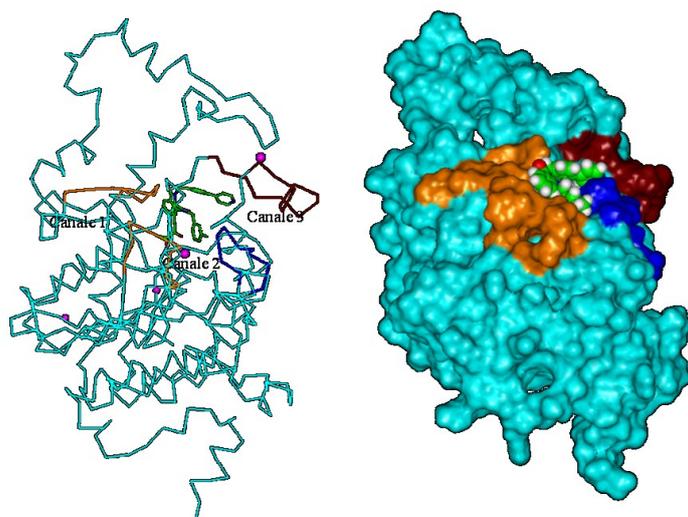


**Figura 21:** RR1\_125 nell'HDAC3. Nella figura è mostrato il risultato 8 della RR1\_125 nell'HDAC3; a destra lo stesso risultato è rappresentato con la proteina in superfici di connoly e il ligando in CPK. In arancione è mostrato il canale 1, in blu ed arancio il canale 2 e il blu e rosso il canale 3. Gli eteroatomi del ligando sono colorati sulla base dei tipi atomici ( N: blu; C:magenta e O: rosso).

## HDAC4

**Tabella 11:** sono riportati i criteri di selezione dei *docking*; in particolare sono stati valutati: gli angoli diedri ( $\tau_1, \tau_2$ ), le distanze dei coordinanti (d1,d2) e l'angolo dello ZBG rispetto allo zinco, l'accomodamento dei sostituenti nei canali, le posizioni dei coordinanti e le posizioni dei gruppi carbonilici del CAP.

Frame docking	$\tau_1$	$\tau_2$	D1	D2	Angolo	Canali sostituenti			Posizioni coordinanti	Posizioni CO
						Anello benzodiazepinico	R	R'		
CAR	-41.28	-21.45	2.36	2.07	78.95	1	2	3	2V5X	2V5X
FRI	-45.86	2.52	2.79	2.90	58.22	1	2	3	2V5X	2V5X

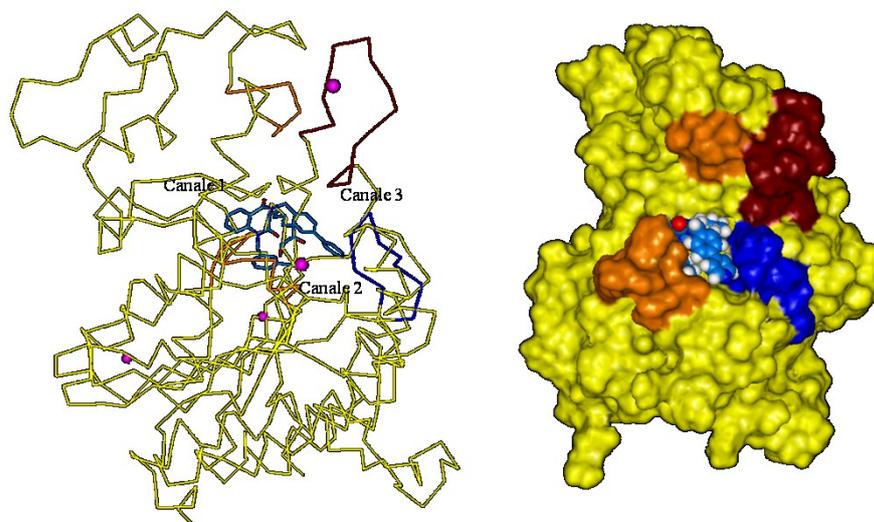


**Figura 22:** RRI\_125 nell'HDAC4. Nella figura è mostrato il risultato 1 della RRI\_125 nell'HDAC4; a destra lo stesso risultato è rappresentato con la proteina in superfici di conolly e il ligando in CPK. In arancione è mostrato il canale 1, in blu ed arancio il canale 2 e il blu e rosso il canale 3. Gli eteroatomi del ligando sono colorati sulla base dei tipi atomici (N: blu; C: verde e O: rosso).

## HDAC7:

**Tabella 12:** sono riportati i criteri di selezione dei *docking*; in particolare sono stati valutati: gli angoli diedri ( $\tau_1, \tau_2$ ), le distanze dei coordinanti (d1,d2) e l'angolo dello ZBG rispetto allo zinco, l'accomodamento dei sostituenti nei canali, le posizioni dei coordinanti e le posizioni dei gruppi carbonilici del CAP.

Frame docking	$\tau_1$	$\tau_2$	D1	D2	Angolo	Canali sostituenti			Posizioni coordinanti	Posizioni CO
						Anello benzodiazepinico	R	R'		
CAR	-33.44	-29.76	2.37	1.79	84.15	1	2	3	2V5X	2V5X
Fr9	-52.07	-35.17	2.57	2.49	70.31	1	2	3	2V5X	-

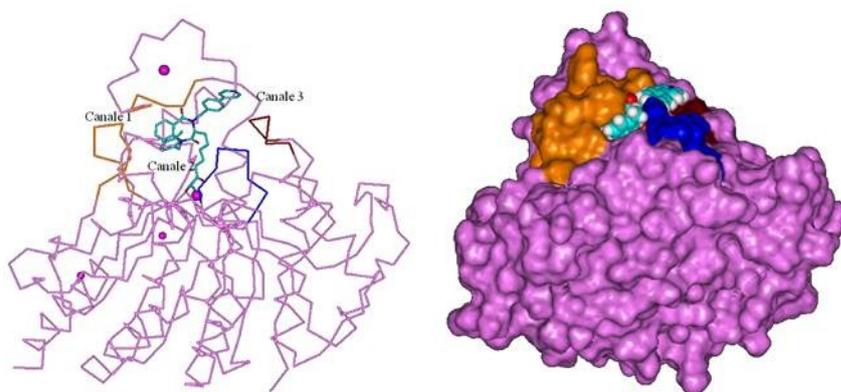


**Figura 23:** *RRI\_125 nell'HDAC7.* Nella figura è mostrato il risultato 9 della *RRI\_125 nell'HDAC7*; a destra lo stesso risultato è rappresentato con la proteina in superfici di conolly e il ligando in CPK. In arancione è mostrato il canale 1, in blu ed arancio il canale 2 e il blu e rosso il canale 3. Gli eteroatomi del ligando sono colorati sulla base dei tipi atomici (N: blu; C: blu e O: rosso).

#### HDAC8:

**Tabella 13:** sono riportati i criteri di selezione dei *docking*; in particolare sono stati valutati: gli angoli diedri ( $\tau_1, \tau_2$ ), le distanze dei coordinanti (d1,d2) e l'angolo dello ZBG rispetto allo zinco, l'accomodamento dei sostituenti nei canali, le posizione dei coordinanti e le posizione dei gruppi carbonilici del CAP.

Frame docking	$\tau_1$	$\tau_2$	D1	D2	Angolo	Canali sostituenti			Posizioni coordinanti	Posizioni CO
						Anello benzodiazepinico	R	R'		
CAR	-31.22	-32.00	2.13	2.00	86.40	1	2	3	2V5X	2V5X
Fr1	-44.62	-35.68	2.47	2.44	70.94	1	2	-	2V5X	-



**Figura 24:** *RRI\_125 nell'HDAC8.* Nella figura è mostrato il risultato 1 della *RRI\_125 nell'HDAC8*; a destra lo stesso risultato è rappresentato con la proteina in superfici di conolly e il ligando in CPK. In arancione è mostrato il canale 1, in blu ed arancio il canale 2 e il blu e rosso il canale 3. Gli eteroatomi del ligando sono colorati sulla base dei tipi atomici (N: blu; C: celeste e O: rosso).

I risultati del *docking* effettuato per la RR\_89 e per la RR1\_125, che differisce dalla prima per la presenza di un azoto in para nel bifenile, sono in perfetto accordo con i risultati ottenuti dal *pharmacophore fitting*: l'anello benzodiazepinico del CAP va nel canale 1, il sostituito R nel 2 e il sostituito R' nel 3 quando le molecole risultano avere una buona o discreta attività sull'enzima. Dal momento che gli accomodamenti sono gli stessi per entrambi i composti, è possibile affermare che le differenti attività biologiche osservate possano essere giustificate da solubilità diverse.

### ***3.5.2 Conclusioni e prospettive future***

Lo studio dei nuovi composti a struttura benzodiazepinica ha sottolineato l'importanza della conformazione ad ali di farfalla del sistema triciclico per consentire la corretta disposizione del gruppo carbonilico del CAP e dei diversi sostituenti del CAP, affinché si accomodino all'interno di specifiche tasche di superficie dell'enzima. L'esito dei saggi farmacologici sia in cellula che sugli enzimi isolati è stato inficiato dalla scarsa solubilità dei composti, probabilmente legata alla loro tendenza a formare aggregati. Lo studio computazionale sarà in grado di guidare successive modifiche strutturali da effettuare per ottenere nuovi composti più solubili che verranno sintetizzati e sottoposti agli opportuni saggi farmacologici sull'enzima isolato ed in cellula. L'obiettivo futuro è, quindi, quello di introdurre sostituenti protonabili a pH fisiologico e che siano anche in grado di ottenere una migliore interazione con l'enzima.

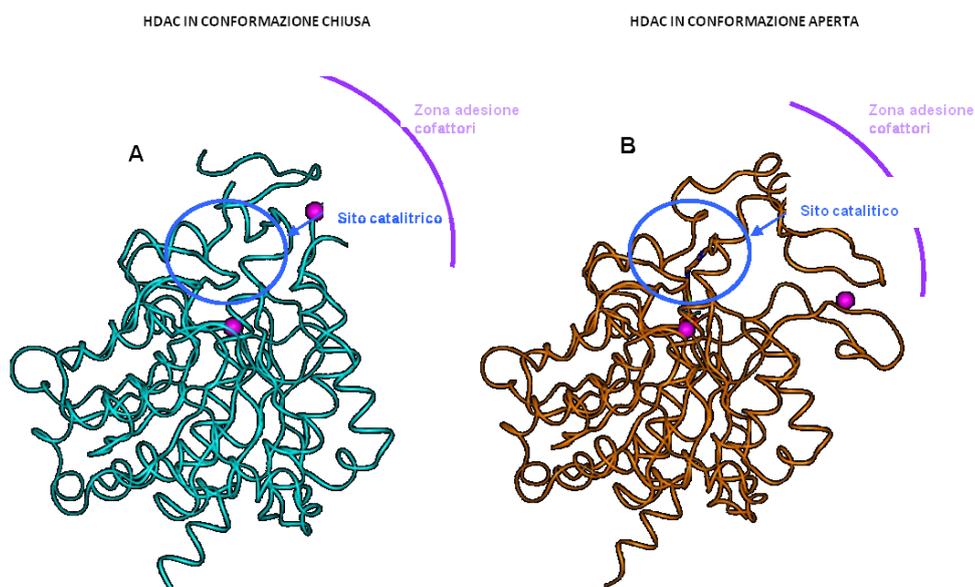
### ***3.6 Progettazione di nuovi inibitori delle Deacetilasi Istoniche (HDAC) caratterizzati dalla presenza di sostituenti in grado di interagire estesamente con la superficie dell'enzima***

#### ***3.6.1 Risultati e Discussione***

Al fine di raggiungere gli obiettivi descritti nel paragrafo 3.2, la continuazione del progetto HDAC ha portato allo sviluppo di una nuova serie di inibitori, implementando il modello farmacoforico di partenza (descritto nel paragrafo 3.4b), attraverso una funzionalizzazione del CAP che consentisse di occupare estesamente la zona di superficie dell'enzima e di stabilire interazioni con essa, in modo da i) intrappolare l'enzima in una conformazione "non prevista" interagendo contemporaneamente con il sito catalitico e con le regioni di modulazione allosterica L2 ed L1, o, alternativamente, ii) indurre l'omodimerizzazione dell'enzima attraverso la simultanea occupazione dei due siti catalitici.

Sono state prese in considerazione le varie conformazioni degli HDAC, sia quelle ottenute dalle strutture a raggi X che quelle ottenute dai nostri calcoli di dinamica molecolare, e sono stati ipotizzati diversi meccanismi di modulazione allosterica, considerando anche i risultati ottenuti da studi di mutagenesi riportati in letteratura.

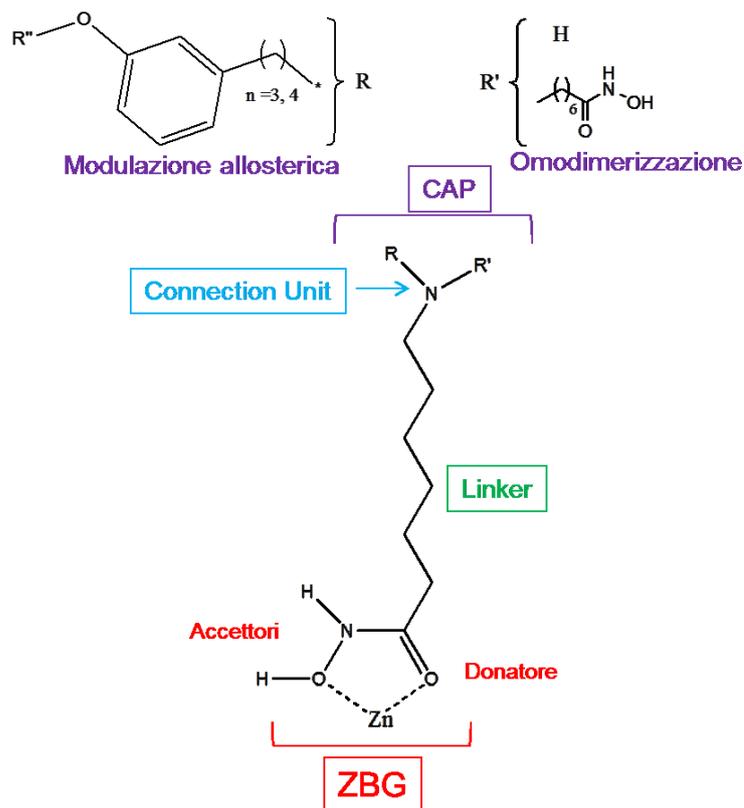
L'immagine seguente (*Figura 1*) mostra come la modulazione allosterica sia importante nel determinare una diversa conformazione dell'enzima e nell'inibire non solo l'attività catalitica ma anche la formazione di macrocomplessi proteici con i cofattori necessari per il riconoscimento delle proteine istoniche.



**Figura 1:** la figura 1A mostra la struttura cristallografica dell' HDAC nella conformazione chiusa mentre la figura 1B mostra la struttura cristallografica dell' HDAC nella conformazione aperta. In entrambi i casi viene messo in evidenza il sito catalitico e la porzione dell'enzima coinvolta nell'interazione con i cofattori .

I risultati ottenuti hanno evidenziato l'esistenza di una sorta di comunicazione conformazionale fra il sito catalitico e la superficie dell'enzima responsabile della stabilizzazione dei diversi stadi conformazionali dell'enzima e, quindi, della modulazione sia dell'attività catalitica che dell'abilità di legare altre proteine. In particolare, gli studi effettuati hanno permesso di individuare gli aminoacidi chiave coinvolti nella regolazione allosterica dell'attività catalitica sia per gli HDAC di classe I (1,2,3 ed 8) che di classe IIa (4,7). Negli HDAC di classe IIa e nell'HDAC 8 sono coinvolti gli aminoacidi impegnati nella coordinazione con uno Zn presente sulla superficie dell'enzima. Il posizionamento e la modalità di coordinazione dello Zn variano fra la classe IIa (4 e 7) e l'HDAC8, che è di classe I. Negli HDAC 1,2 e 3, di classe I, invece, la modulazione allosterica svolta dal legame con lo Zn è sostituita dal legame con i cofattori e quindi sono coinvolti gli aminoacidi responsabili di questa interazione, in particolare per quanto riguarda l'HDAC3 corrispondono a quelli coinvolti nell'interazione con NCoR.

Partendo da questi risultati è stato implementato il modello farmacoforico sviluppato sulla base dell'esperienza acquisita durante gli studi precedentemente effettuati sugli inibitori HDAC, e i cui risultati sono stati precedentemente riportati. Il nuovo farmacoforo prodotto è schematizzato nell'immagine seguente (Figura 2).



**Figura 2:** rappresentazione schematica del nuovo modello farmacoforico degli HDACi

Ai punti farmacoforici precedentemente individuati, ossia i) ZBG, che coordina lo ione zinco e interagisce con gli amminoacidi presenti nel sito catalitico attraverso diversi legami a idrogeno; ii) il *linker*, rappresentato generalmente una catena idrofobica, omologa alla catena alchilica della lisina, che attraversa il canale idrofobico dal sito catalitico alla superficie dell'enzima; iii) la *connection unit*, capace di interagire con un Asp conservato in tutti gli HDAC e situato all'ingresso del canale idrofobico, è stato aggiunto un CAP, ovvero la porzione del ligando che si proietta al di fuori del canale verso la superficie dell'enzima, razionalmente progettato al fine di interagire i) con gli amminoacidi responsabili della modulazione allosterica dell'attività enzimatica, in particolare dell'HDAC3, ii) favorire l'omodimerizzazione degli HDAC riuscendo a coordinare contemporaneamente lo Zn catalitico di due monomeri di HDAC, in entrambi i casi portando al blocco dell'enzima in una conformazione tale da non poter più svolgere le proprie funzioni.

Su queste basi è stata effettuata la progettazione mirata di un ristrettissimo set di inibitori. Il CAP dei nuovi inibitori progettati è stato funzionalizzato con una carica positiva; un gruppo guanidinico ad opportuna distanza dall'unità di connessione (CU) in modo da interagire con due residui di Aspartato esclusivamente conservati negli HDAC di classe I e che normalmente interagiscono con la carica positiva di un residuo di lisina del substrato e sono intimamente

coinvolti nella regolazione allosterica esistente fra la superficie dell'enzima ed il sito catalitico.

I nuovi derivati, sintetizzati dal Prof. Maurizio Taddei (Università di Siena), sono state testate dal gruppo di ricerca del Dr. Ralph Mazitschek dell'Harvard Medical School, i substrati utilizzati sono quelli riportati nel capitolo 3.4.1. Le strutture e le relative attività sulle diverse isoforme di HDAC sono di seguito riportate.

**Tabella 1:** valori di inibizione sulle diverse isoforme HDAC dei nuovi composti progettati anche qui solo quelli riportati nella relazione.

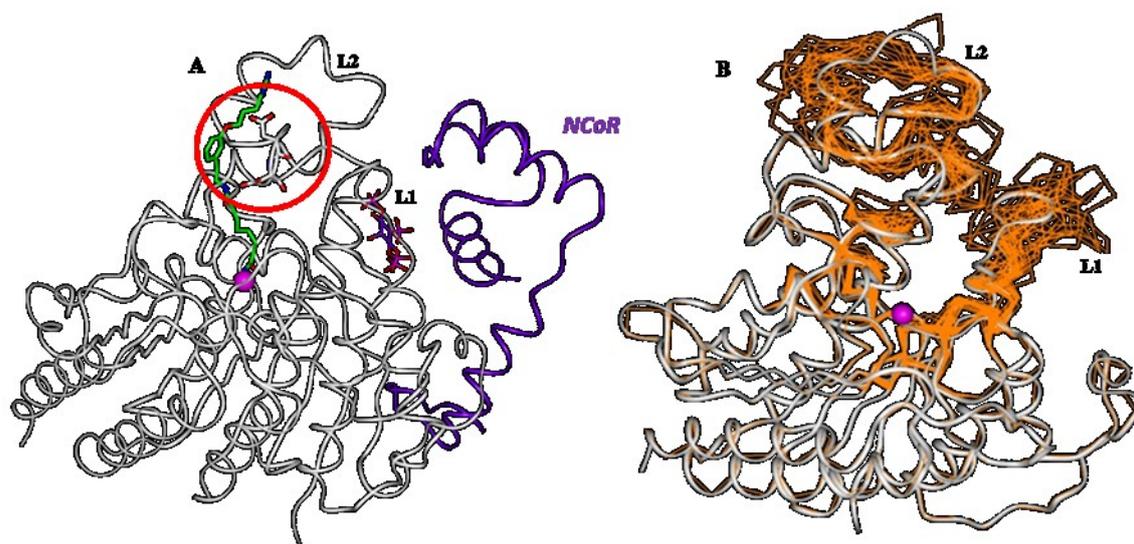
COMPOSTO	Classe I Kd ( $\mu$ M)				Classe IIa Kd ( $\mu$ M)			
	HDAC 1	HDAC 2	HDAC 3	HDAC 8	HDAC 4	HDAC 5	HDAC 7	HDAC 9
<b>ST7785</b> 	0.023	0.022	0.032	1.38	NA	NA	NA	NA
<b>ST7884</b> 	0.027	0.031	0.048	1.973	NA	NA	NA	NA
<b>ST8020CL1</b> 	0.0009	0.0004	0.0019	0.469	NA	98.55	NA	NA
<b>SAHA</b> 	0.014	0.045	0.107	0.177	0.083	0.039	1.433	0.012
<b>TSA</b> 	0.005	0.006	0.007	0.192	13.085	0.733	2.270	13.210

NA sta per non attiva

Tutte le molecole sintetizzate hanno dimostrato un'elevata attività inibitoria sugli HDAC 1-3. In particolare, gli inibitori ST7785 ed ST7884 mostrano avere una Kd 3 volte superiore al SAHA che è il composto attualmente utilizzato nella terapia antitumorale; mentre il ligando ST8020 presenta una Kd sull'HDAC3 56 volte superiore a quella del SAHA e 4 volte

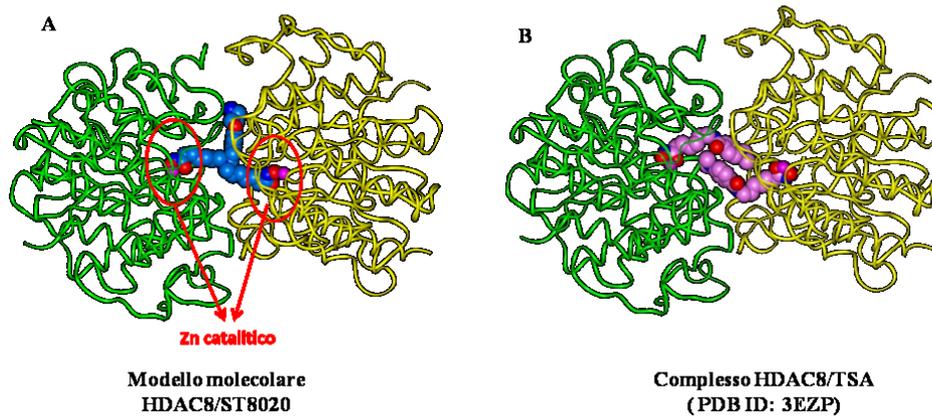
superiore a quella del TSA, uno degli inibitori HDAC più potenti attualmente in circolazione. Inoltre, i nuovi composti progettati, contrariamente ai composti di riferimento SAHA e TSA, mostrano una spiccata selettività verso gli HDAC di classe I.

Gli studi di *pharmacophore fitting* mostrano che i nuovi composti progettati riescono a stabilire con la loro carica positiva delle interazioni con due cariche negative (Glu99, Asp100 dell' HDAC1; Glu101, Asp100 dell' HDAC2; Asp93, Asp94 dell' HDAC3) esclusivamente conservati negli HDAC di classe I; nella figura 3A seguente viene mostrato come l'inibitore ST7785 riesce a interagire mediante il CAP con l'Asp determinando un cambiamento conformazionale della regione di legame L2 che è direttamente connessa alla regione di legame L1 e che come mostrato nella figura 3B studi di *docking* dinamico precedentemente effettuati sugli altri inibitori, confermano che tale cambiamento conformazionale va ad interferire nella regione di legame dei cofattori.



**Figura 3:** la figura 3A mostra l'interazione del ligando ST7785 con l'Asp93 dell'HDAC3; la figura 3B mostra il cambiamento conformazionale della regione di legame L1 ed L2 ottenuto dagli studi di *docking*. Gli eteroatomi del ligando sono colorati sulla base dei tipi atomici (N: blu; C: verde e O: rosso).

Per l'inibitore ST8020 il cui CAP è stato, invece, funzionalizzato aggiungendo due punti di aggancio per la coordinazione dello zinco catalitico, è possibile ipotizzare un processo di omodimerizzazione delle proteine così come avviene per il complesso HDAC8/TSA sperimentalmente determinato. (Figura 4)



**Figura 4:** la figura 4A mostra il modello molecolare HDAC8/in complesso con l'inibitore ST8020; la figura 4B rappresenta il complesso HDAC8/TSA della struttura cristallografica PDB ID:3EZF

### ***3.6.2 Conclusioni***

In conclusione, nel suo complesso il progetto HDAC svolto in collaborazione con la Sigma-Tau, attraverso un processo iterativo che ha previsto l'analisi del meccanismo d'azione a livello molecolare e la progressiva ottimizzazione della struttura di nuovi inibitori sulla base del modello farmacoforico e delle SAR ottenute, ha portato all'ottenimento di inibitori HDAC caratterizzati da un elevato profilo di selettività e potenza.

### 3.7 Bibliografia

1. W. Flemming. "*Cell substance, Nucleus and Cell Division*", 1882
2. Yingjie Zhang, et.al. "*The Structure and Function of Histone Deacetylases: The Target for Anti-cancer Therapy*" *Current Medicinal Chemistry* 2008, 15, 2840-2849.
3. B.M. Turner "*Cellular memory and the histone code*" *Cell*, 2002, 111, 285–291.
4. Kyle V. Butler and Alan P. Kozikowski; "*Chemical Origin of Isoform Selectivity in Histone Deacetylase Inhibitors*", *Current Pharmaceutical Design*, 2008, 14, 000-000.
5. Hildmann C, et al. "*Substrate and inhibitor specificity of class 1 and class 2 histone deacetylases*" *Journal of biotechnology* 2006, 124, 258-70.
6. . Wang DF, et.al. "*Toward selective histone deacetylase inhibitor design: homology modeling, docking studies, and molecular dynamics simulations of human class I histone deacetylases*" *Journal of medicinal chemistry* 2005, 48, 6936-4.
7. Rodriguez M, et.al. "*Chemistry and biology of chromatin remodelling agents: state of art and future perspectives of HDAC inhibitors*" *Current Medicinal Chemistry* , 2006, 13, 1119-39.
8. Yang XJ, et.al. "*Collaborative spirit of histone deacetylases in regulating chromatin structure and gene expression*" *Current Opinion in Genetics & Development*, 2003, 13:143-153.
9. Paola Gallinari, et.al "*HDACs, histone deacetylation and gene transcription: from molecular biology to cancer therapeutics*" *Cell Research*, 2007, 17:195-211.
10. Marielle Paris, et.al. "*Histone Deacetylase Inhibitors: From Bench to Clinic*" *Journal of Medicinal Chemistry*, 2008, 51, 6.
11. Finnin, M. S. et.al "*Structure of a histone deacetylase homologue bound to TSA and SAHA inhibitors*" *Nature*, 1999, 401, 188–193
12. a) Vannini, A. et.al. "*Crystal structure of a eukaryotic zinc-dependent histone deacetylase, human HDAC8, complexed with a hydroxamic acid inhibitor*" *PNAS*, 2004, 101, 15064–15069.
- b) Somoza, J. R.; et.al. "*Structural snapshots of human HDAC8 provide insight into class I histone deacetylases*" *Structure*, 2004, 12,1325–1334.
13. Suzuki T, et.al. "*Non-hydroxamate histone deacetylase inhibitors*" *Current Medicinal Chemistry*, 2005, 12: 2867-80.

14. Suzuki T, et al. “*Novel histone deacetylase inhibitors: design, synthesis, enzyme inhibition, and binding mode study of SAHA-based non-hydroxamates*” *Bioorganic e Medicinal Chemistry letters*, 2003, 13, 4321-6.
15. Wang D, et.al. “*On the function of the 14 A long internal cavity of histone deacetylase-like protein: implications for the design of histone deacetylase inhibitors*” *Journal of Medicinal Chemistry*, 2004, 47, 3409-17.
16. Warrel, R. P.; et.al. “*Therapeutic targeting of transcription in acute promyelocitic leukemia by use of an inhibitor of histone deacetylases*” *Journal of the National Cancer Institute*, 1998, 90, 1621–1625.
17. Milos Dokmanovic et.al. “*Prospects: Histone Deacetylase Inhibitors*” *Journal of Cellular Biochemistry*, 96, 293–304 2005.
18. a) Guenther, M.G., et.al “*A core SMRT corepressor complex containing HDAC3 and TBL1, a WD40-repeat protein linked to deafness*”. *Genes & Development* , 2000, 14, 1048–1057
- b) Wen, Y.D., et.al. “*The histone deacetylase-3 complex contains nuclear receptor corepressors*”. *PNAS*, 2000, 97, 7202–7207.
19. Laherty, C.D., et.al. “*Histone eacetylases associated with the mSin3 corepressor mediate mad transcriptional repression*”. *Cell*, 1997, 89, 349–356.
20. Humphrey, G.W., et.al. “*Table histone deacetylase complexes distinguished by the presence of SANT domain proteins CoREST/kiaa0071 and Mta-L1*”. *The Journal of Biological Chemistry*, 2001, 276, 6817–6824.
- 21 a) Xue, Y., et.al. “*NURD, a novel complex with both ATP-dependent chromatin-remodeling and histone deacetylase activities*” *Molecular Cell*, 1998, 2, 851–861.
- b) Zhang, Y., et.al. “*Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation*” *Genes & Development*, 1999, 13, 1924–1935.
22. Christopher J. Millard, et.al.”*Class I HDACs Share a Common Mechanism of regulation by Inositol Phosphates*” *Cell* 2013, 51, 57–67
23. ACD/Percepta, version 14.0.0, Advanced Chemistry Development, Inc., Toronto, ON, Canada, www.acdlabs.com, 2013.
24. Dauber-Osguthorpe, P.; et.al. “*Structure and energetics of ligand binding to proteins: E. coli dihydrofolate reductase-trimethoprim, a drug-receptor system*” *Proteins*, 1988, 4, 31-47.
25. Fletcher, R. “*Unconstrained Optimization. In Pratical Methods of Optimization John Wiley & Sons*” New York 1980.

26. Stewart J. J. P., “*Optimization of Parameters for Semiempirical Methods V: Modification of NDDO Approximations and Application to 70 Elements*” *Journal Molecular Modeling* 2007, 13, 1173-1213
27. MOPAC2012, James J. P. Stewart, Stewart Computational Chemistry, Colorado Springs, CO, USA, [HTTP://OpenMOPAC.net](http://OpenMOPAC.net) (2012)
28. J. Baker. “*An Algorithm for the Location of Transition States*” *Journal of Computational Chemistry*, 1986, 7, 385-395
29. a) H. Brunner, F. Stöhr *Eur. The Journal of Organic Chemistry*, 2000, 2777-2786. b) B. A. Marshall and William A. Waters. *Organic Compounds by Cupric Salts. Part I.* 1980 c) R.M. L. Wolfrom, W. Lee Lewis *J. Am. Chem. Soc.* 1928
30. McDowell, R. S.; Blackburn, B. K.; Gadek, T. R.; McGee, L. R.; Rawson, T.; Reynolds, M. E.; Robarge, K. D.; Somers, T. C.; Thorsett, E. D.; Tischler, M.; Webb, R. R., II; Venuti, M. C. *Journal of the American Chemical Society*, 1994, 116, 5077
31. a) Theoretical calculations at the density functional B3-LYP/6-31G(d) level were performed using the Gaussian 98 package 22b; b) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*; Gaussian, Inc.: Wallingford, CT, 2009.
32. (a) Becke, A. D. “*Density-functional thermochemistry. III. The role of exact exchange*”. *The Journal of Chemical Physics*, 1993, 98, 5648–5652; (b) Lee, C.; et.al. “*Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density*” *Physical Review*, 1988, 37, 785–789.
33. Frisch, M. J.; et.al. “*Self-consistent molecular orbital methods. 25. Supplementary functions for Gaussian basis sets*” *The Journal of Chemical Physics*, 1984, 80, 3265–3269.

34. (a) Hay, P. J.; et.al. “*Ab-initio effective core potentials for molecular calculations. Potentials for the transition metal atoms Sc to Hg*” The Journal of Chemical Physics, 1985, 82, 270–283. (b) Wadt, W. R.; et.al. “*Ab- initio effective core potentials for molecular calculations. Potentials for main group elements Na to Bi*” The Journal of Chemical Physics, 1985, 82, 284–298 (c) Hay, P. J.; et.al. “*Ab-initio effective core potentials for molecular calculations. Potentials for K to Au including the outermost core orbitals*” The Journal of Chemical Physics, 1985, 82, 299–310.
35. (a) NBO, version 3.1: Glendening, E. D.; Reed, A. E.; Carpenter, J. E.; Weinhold F. (b) Foster, J. P.; et.al. “*Natural hybrid orbitals*”. Journal of the American Chemical Society 1980, 102, 7211–7218. (c) Reed, E.; et.al. “*Natural bond orbital analysis of near-Hartree-Fock water dimer*” The Journal of Chemical Physics, 1983, 78, 4066–4073. (d) Reed, A. E.; et.al. “*Natural population analysis*” The Journal of Chemical Physics, 1985, 83, 735–746. (e) Reed, A. E.; et.al. “*Natural localized molecular orbitals*” The Journal of Chemical Physics, 1985, 83, 1736–1740. (f) Carpenter, J. E. “*Extension of Lewis Structure Concepts to Open-Shell and Excited-State Molecular Species*” Ph.D. Thesis, University of Wisconsin, Madison, WI, 1987. (g) Carpenter, J. E.; Weinhold, F. “*Analysis of the geometry of the hydroxymethyl radical by the different hybrids for different spins natural bond orbital procedure*” Journal of Molecular Structure, THEOCHEM, 1988, 46, 41–62. (h) Reed, A. E.; et.al. “*Intermolecular interactions from a natural bond orbital, donor-acceptor viewpoint*” Chemical Reviews, 1988, 88, 899–926. (i) Weinhold, F.; et.al. “*In The Structure of Small Molecules and Ions*” Naaman, R., Vager, Z., Eds.; Plenum: New York, 1988; p 227.
36. Cioslowski, J. “*A new population analysis based on atomic polar tensors*” Journal of the American Chemical Society, 1989, 11, 8333–8336
37. Irshaidat, T. “*A unique and novel cyclopropylmethyl cation intermediate: a DFT study*” Tetrahedron Letters, 2008, 49, 5894–5898. (b) Naoki, K. J. “*Density functional theory study of electroreductive hydro-coupling of  $\alpha,\beta$ -unsaturated carbonyl compounds*” The Journal of Organic Chemistry, 2006, 71, 9203–9207. (c) Faza, O. N.; et.al. “*Mechanism of the gold(I) catalyzed Rautenstraucht rearrangement: a center-to-helix-to-center chirality transfer*” Journal of the American Chemical Society, 2006, 128, 2434–2437. (d) Lev, D. A.; et.al. “*Reversal of reactivity in diene-complexed o-quinone methide complexes: insights and explanations from ab initio density functional theory calculations*” Organometallics, 2005, 24, 4240. (e) Sertchook, R.; et.al. “*Rozen’s epoxidation reagent,  $CH_3CN \cdot HOF$ : a theoretical study of its structure, vibrational spectroscopy, and reaction mechanism*” The Journal of Physical Chemistry A, 2006, 110, 8275–8281. (f) Iron, M. A.; et.al “*A computational foray*

*into the formation and reactivity of metallabenzenes*” Journal of the American Chemical Society, 2004, 126, 11699–11710.

38. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian 09; Gaussian, Inc.: Wallingford, CT, 2009.

39. Pei, J.; Kim, B. H.; Grishin, N. V. “*PROMALS3D: a tool for multiple sequence and structure alignment*” Nucleic Acids Research, 2008, 36, 2295–2300

40. Shi S, et.al. “*An extensible and systematic force field, ESFF, for molecular modeling of organic, inorganic, and organometallic systems*” Journal of Computational Chemistry, 2003, 24,1059-1076

41. Senderowitz, H.; et.al. “*A smart Monte Carlo technique for free energy simulations of multiconformational molecules Direct calculations of the conformational populations of organic molecules*” Journal of the American Chemical Society, 1995, 117, 8211–8219

42. Ding, H. Q.; et.al. “*A III Atomic level simulations on a million particles: the cell multipole method for Coulomb and London non-bond interactions*” The Journal of Chemical Physics, 1992, 97, 4309–4315

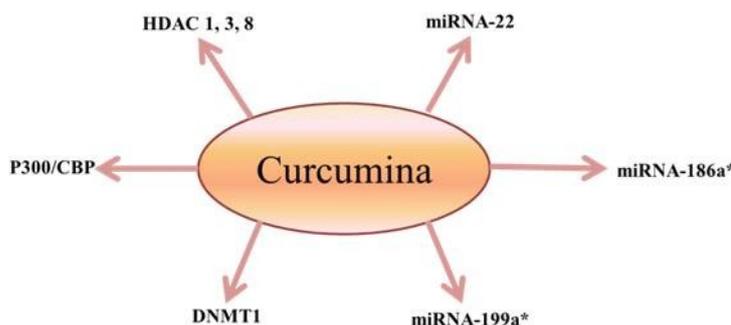
43. Guandalini, A. “*Design, synthesis and preliminary biological evaluation of new hydroxamate histone deacetylase inhibitors as potential antileukemic agents*” Bioorganic & Medicinal Chemistry Letters 2008, 18, 5071–5074.

# ***4. Investigazione delle Basi Molecolari della Curcumina***

## ***4.1 Introduzione***

Alla fine del mio dottorato di ricerca ho cominciato ad occuparmi di un ulteriore progetto in collaborazione con l'industria farmaceutica Indena S.p.A. (Milano). Il composto dell'Indena denominato *Meriva* è costituito da estratti di *Curcuma longa* il cui componente principale è rappresentato dalla Curcumina, i miei studi hanno riguardato, quindi, l'investigazione delle basi molecolari dell'attività biologica della curcumina. Tale progetto proseguirà anche dopo la conclusione del mio dottorato di ricerca grazie ad una borsa di studio finanziata dall'Indena.

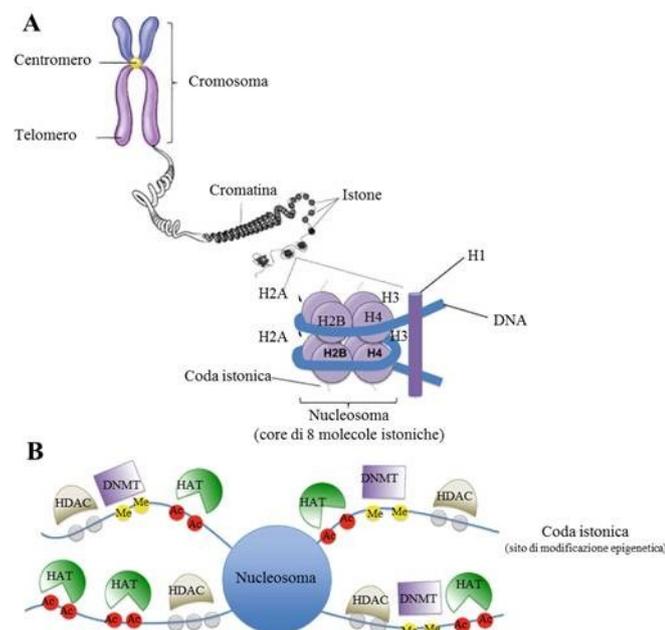
La Curcumina (diferuloilmetano), componente principale della *Curcuma longa*, è una molecola altamente pleiotropica, ovvero lega numerosi bersagli (*target*) molecolari<sup>1</sup>, quali molecole infiammatorie, proteine coinvolte nella sopravvivenza cellulare, protein chinasi, protein reduttasi, istoni acetil transferasi, istoni deacetilasi, proteosoma, HIV1 integrasi, DNA metiltransferasi 1, proteine carrier e ioni metallici. Essa va ad interferire sui meccanismi di controllo epigenetico della cellula, ovvero sui cambiamenti dell'espressione dei geni contenuti nella sequenza del DNA (*Figura 1*).



***Figura 1*** : rappresentazione schematica dei fattori epigenetici regolati dalla curcumina<sup>2</sup>

I meccanismi epigenetici includono cambiamenti nella metilazione del DNA, modificazioni istoniche e alterazioni dell'espressione dei microRNA (miRNA). Modifiche sulla struttura della cromatina influenzano l'espressione genica mediante l'inattivazione dei geni, che avviene quando la cromatina è chiusa (eterocromatina) o mediante l'attivazione di geni quando la cromatina è aperta (eucromatina). Il nucleosoma che è fondamentale una

ripetizione di unità della cromatina è composto da DNA raggruppato attorno all'ottamero istonico, formato da quattro istoni *partners*, un tetramero H3-H4, due dimeri H2A-H2B. Ogni core nucleosomale successivo è separato da un DNA *linker* associato con una singola molecola di istone H1 (*Figura 2A*). Le modificazioni della cromatina di solito avvengono all'amminoacido N-terminale dell'istone e facilitano od ostacolano l'associazione delle proteine che riparano il DNA e dei fattori trascrizionali con la cromatina. Questi core istonici subiscono una vasta gamma di modificazioni post-traduzionali, compresa l'acetilazione controllata da un istone acetil transferase HAT, l'espressione del gene, deacetilazione controllata dagli HDAC , inattivazione genica, metilazione, fosforilazione, ubiquitinazione, ribosilazione dell'ADP e possibili biotinilazione<sup>2</sup> (*Figura 2B*)

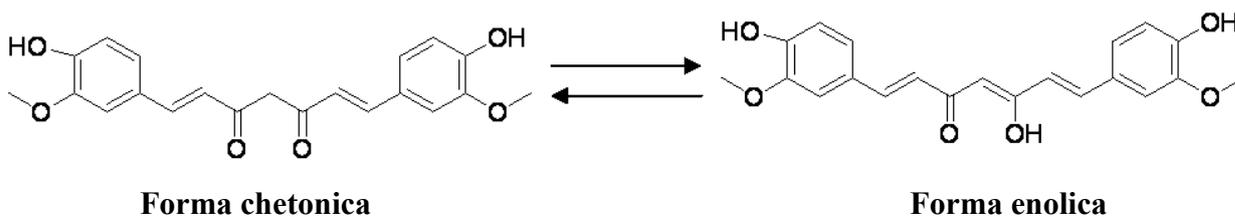


**Figura 2 :** *A*) Sono rappresentate le unità fondamentali di ripetizione della cromatina nel nucleosoma. Un singolo core nucleosomale è composto da un frammento di DNA avvolto attorno all'ottamero istonico, formato da un H4-H4 tetramero e due H2A-H2B dimero. Ogni successivo core nucleosomale è separato da un linker di DNA associato con una singola molecola di istone H1. *B*) Le modificazioni della cromatina solitamente avvengono all'amminoacido N-terminale degli istoni. Queste code istoniche sono il sito di modificazioni post-traduzionali, includendo l'acetilazione controllata dagli istoni acetil transferasi (HAT), deacetilazione controllata dalle deacetilasi istoniche (HDACs), e la metilazione controllata dal DNA metiltransferasi (DNMTs)<sup>2</sup>.

Inoltre la curcumina può avere effetti sull'espressione dei miRNA (*Figura 2*). Espressioni aberranti dei miRNA possono avvenire attraverso numerosi meccanismi, compresi anomalie genomiche, regolazione trascrizionale, processamento del miRNA. Quest'ultimi sono piccole molecole, endogene, singolo filamento di RNA di 19-25 nucleotidi lunghi che regolano l'espressione genica, per esempio si legano in modo imperfetto alla regione 3' non tradotta dell' mRNA bersaglio, portando alla repressione traduzionale, o al clivaggio dell'mRNA bersaglio a causa dell'imperfetta complementarietà tra miRNA ed mRNA<sup>2</sup>.

La curcumina, in conclusione, nei saggi cellulari ed in vivo, mostra avere un ampio spettro di effetti farmacologici come l'attività anti-proliferativa, antiinfiammatoria, antiossidante e pro-apoptotica ed interagisce con una varietà di proteine modificando la loro espressione ed attività, inibendo in ultima analisi la proliferazione cellulare, invasione, angiogenesi e la metastasi di differenti tipi di cancro.

Dal punto di vista chimico, la curcumina è un bis- $\beta$ -dichetone  $\alpha,\beta$ -insaturo ed esiste in equilibrio con il suo tautomero enolico. (Figura 3)



**Figura 3:** sono mostrate le due forme tautomeriche della curcumina, rispettivamente la forma chetonica e la forma enolica in equilibrio tra loro.

Si presenta come una polvere cristallina arancio-giallo insolubile in acqua ed etere ma solubile in etanolo, dimetil-sulfossido e acetone. Ad un pH tra 1-7, le molecole di curcumina sono in uno stato neutro con un colore giallo e hanno una bassa solubilità in acqua, mentre ad un pH >7.5 questo colore muta in rosso. L'eptadienone tra i due anelli metossi-fenolici contiene un atomo di carbonio altamente reattivo. La curcumina possiede tre protoni che sono ionizzabili in acqua, il protone enolico con un pKa di circa 7.8 e due protoni dei fenoli con un pKa di circa 9.0 (in una miscela di solventi di alcool e acqua).<sup>3</sup>

Nell'intervallo di pH 3-7, la curcumina agisce come un potente donatore di atomi di H, mentre sopra un pH circa di 8, predomina la forma enolica della catena eptadienonica e il composto agisce come un donatore di elettroni similmente a molti fenoli antiossidanti<sup>3</sup>.

## ***4.2 Scopo del lavoro***

Lo scopo del progetto di ricerca svolto in collaborazione con l'Industria Farmaceutica Indena (Indena, S.p.A., Milano, Italia) e con l'Università degli Studi di Salerno, è stato quello di investigare le basi molecolari dell'interazione fra la curcumina e gli HDAC. L'attività inibitoria della curcumina sugli HDAC è stata misurata essere micro molare nei confronti di estratti cellulari contenenti HDAC di classe I ma non è mai stata misurata sulle diverse isoforme isolate.

## 4.3 Parte Sperimentale

Tutti gli studi di modellistica molecolare sono stati eseguiti su Silicon Graphics SGI Origin 200 8XR12000 ed E4 Server Twin 2 x Dual Xeon—5520, equipaggiati con due nodi; ogni nodo: 2 x Intel® Xeon® QuadCore E5520—2,26Ghz, 36 GB RAM.

### 4.3a Analisi Conformazionale

La metodologia del calcolo delle forme tautomeriche, dei valori di pKa e la procedura ed i parametri utilizzati per l'analisi conformazionale dei composti in esame, sono riportati nel paragrafo 3.3a.

I conformeri ottenuti dal metodo di calcolo quanto-meccanico PM7 sono stati classificati mediante: i) energia conformazionale ( $\Delta E < 5$  Kcal), (ii) legami idrogeno intramolecolari, (iii) angoli torsionali; iv) distanze interatomiche dei centroidi dei due anelli aromatici (Pseudo\_Atom Define command, Biopolymer Module, Insight 2005).

Inoltre le strutture cristallografiche della curcumina, libera (codici CSDs: AXOGIE, AXOGOK, BINMEQ, BINMEQ01, BINMEQ02, BINMEQ03, BINMEQ04, BINMEQ05, BINMEQ06, BINMEQ07, BINMEQ08, FAPLIR) o in complesso con i metalli (codice CSDs: RESXUI, RAWDOK), sono state selezionate e scaricate dal *Cambridge Structural Database* (CSD) utilizzando il CSDS (Cambridge Structural Database System) software *Conquest 1.15*; e sono state analizzate utilizzando il modulo *Biopolymer di Insight 2005* (Accelrys, San Diego, CA).

### 4.3b Studi di Docking

Gli studi di *docking* sono stati condotti sulla curcumina considerando tutte le sue forme tautomeriche, in complesso con le istoni deacetilasi umane HDAC2 (PDB ID: 3MAX) ed HDAC3 (PDB ID: 4A69). La procedura ed i parametri utilizzati per gli studi di *docking*, sono stati gli stessi di quelli descritti nel paragrafo 3.3d.

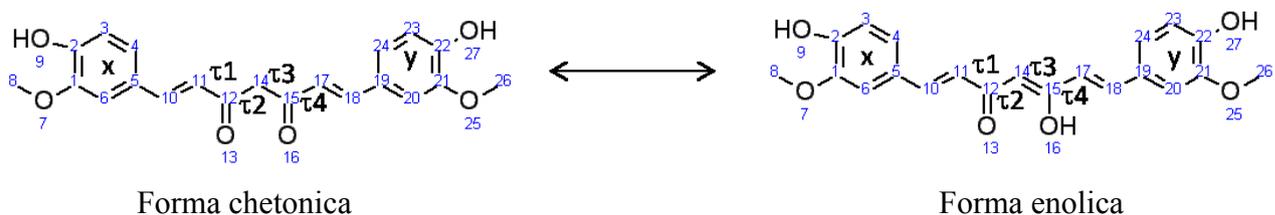
Sono stati considerati diversi gruppi coordinanti lo zinco nel sito catalitico, in particolare: i) i due gruppi carbonilici della forma chetonica, ii) il gruppo idrossile e carbonile della forma enolica; iii) i sostituenti 2-metossi 3-idrossi dell'anello fenilico. Come struttura di riferimento per il posizionamento dei conformeri della curcumina nel sito catalitico, è stata utilizzata la struttura a raggi X dell'HDAC2 (PDB ID: 3MAX). I ligandi sono stati inseriti nel sito catalitico tenendo in considerazione il modello farmacoforico sviluppato degli inibitori HDAC (i-e. ZBG, linker e CAP vedi paragrafo 3.4.1b, *Figure 10 e 11*).

I complessi risultanti dai calcoli di *docking* sono stati classificati e selezionati sulla base della loro energia conformazionale e della loro abilità di riprodurre le modalità di legame riportate sperimentalmente per gli inibitori HDAC.

## 4.4 Risultati e Discussione

### 4.4a Analisi Conformazionale

L'investigazione delle proprietà conformazionali delle forme tautomeriche della curcumina ha evidenziato che la curcumina nella forma chetonica si trova sempre in una conformazione ripiegata, mentre nella forma enolica si trova sempre in una conformazione distesa. Quest'ultima, a sua volta, nella metà dei casi presenta una conformazione denominata Conf 1 del torsionale  $\tau 1$  e nell'altra metà dei casi una conformazione denominata Conf 2 (Tabella 1). In particolare, per la classificazione dei minimi energetici ottenuti dalla ricerca conformazionale sono stati considerati gli angoli diedri e le distanze interatomiche dei centroidi degli anelli aromatici (X-Y), come mostrato nell'immagine seguente (Figura 4)



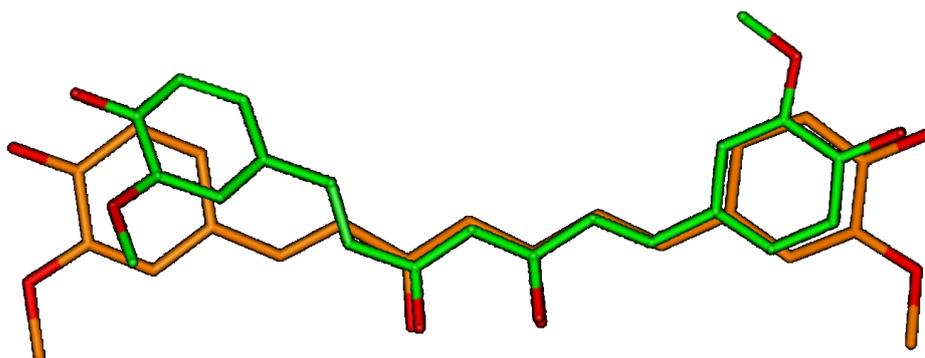
**Figura 4:** sono rappresentate le forme tautomeriche della curcumina con numerazione IUPAC ed indicati gli angoli diedri  $\tau 1$  ( $C_{10}C_{11}C_{12}C_{14}$ ),  $\tau 2$  ( $C_{11}C_{12}C_{14}C_{15}$ ),  $\tau 3$  ( $C_{12}C_{14}C_{15}C_{17}$ ),  $\tau 4$  ( $C_{14}C_{15}C_{17}C_{18}$ ) e i centroidi degli anelli (X-Y) considerati per la classificazione in famiglia.

**Tabella 1:** sono riportati in tabella i valori di  $\Delta E$  dei minimi considerati, le distanze tra i centroidi degli anelli aromatici (X-Y), i valori degli angoli diedri ( $\tau 1, \tau 2, \tau 3, \tau 4$ ), e le percentuali delle conformazioni ritrovate.

Forma tautomerica	$\Delta E$	X-Y	$\tau 1^a$	$\tau 2^b$	$\tau 3^c$	$\tau 4^d$	% conformazioni
Chetonica	0	4.98	-153.51	32.16	32.09	-53.48	100
Enolica	0	12.45	178.53	-179.78	-179.66	-178.71	50
Enolica	0.728	11.63	42.73	179.36	-179.70	-178.00	50

<sup>a</sup>  $\tau 1$  ( $C_{10}C_{11}C_{12}C_{14}$ ), <sup>b</sup>  $\tau 2$  ( $C_{11}C_{12}C_{14}C_{15}$ ), <sup>c</sup>  $\tau 3$  ( $C_{12}C_{14}C_{15}C_{17}$ ), <sup>d</sup>  $\tau 4$  ( $C_{14}C_{15}C_{17}C_{18}$ )

La ricerca effettuata nella *Cambridge Structural Database* sulla curcumina ha evidenziato, in accordo con la ricerca conformazionale sopra descritta, la presenza di due diverse conformazioni della curcumina per la forma enolica, corrispondenti alla Conf1 e 2 di  $\tau 1$  da noi individuate (Conf 1- Codici CSDs: AXOGIE, AXOGOK, BINMEQ 06, BINMEQ 07, BINMEQ 08; Conf 2- Codici CSDs: BINMEQ, BINMEQ01, BINMEQ 02, BINMEQ 03, BINMEQ 04, BINMEQ 05) (Figura 5) .



**Figura 5:** sono rappresentate le due possibili conformazioni della curcumina in forma enolica ritrovate nelle strutture a raggi X e risultate dall'analisi conformazionale.

Inoltre, è stata trovata solo una struttura della curcumina in complesso con un metallo ed, in particolare, con lo zinco (codice CSD: RAWDOK). Per verificare l'esatta natura dei legami coniugati presenti nella struttura, sono state calcolate tutte le distanze di legame e sono state paragonate alle seguenti distanze di riferimento: C-C (1.5Å); C=C (1.3Å); C-O (1.43-1.47Å), C=O (1.17-1.26Å). Di seguito sono riportate le distanze considerate per ciascuna struttura presa in esame (Tabella 2).

**Tabella 2:** sono riportati i valori delle distanze dei legami che caratterizzano i composti considerati.

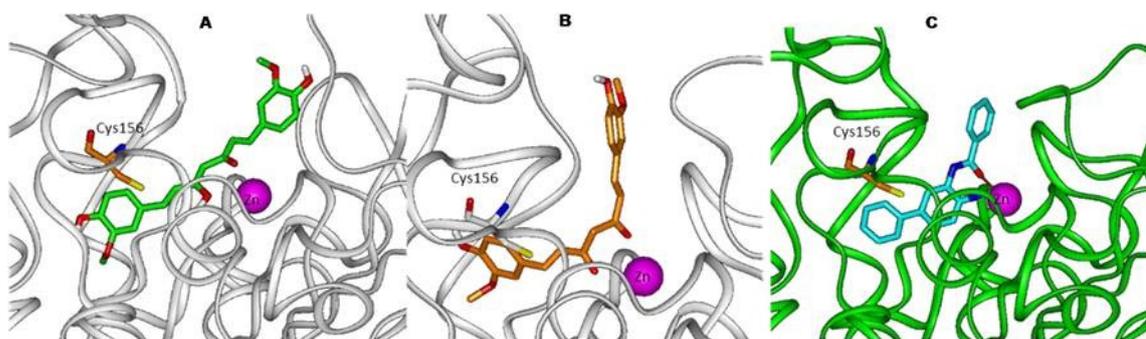
Codici CSD	C12-O13	C15-O16	C5-C10	C10-C11	C11-C12	C12-C14	C14-C15	C15-C17	C17-C18	C18-C19
AXOGIE	1.29	1.33	1.49	1.26	1.48	1.41	1.35	1.47	1.28	1.53
AXOGOK	1.29	1.33	1.46	1.34	1.46	1.42	1.38	1.44	1.34	1.46
BINMEQ06	1.27	1.33	1.45	1.33	1.47	1.41	1.37	1.44	1.33	1.46
BINMEQ07	1.28	1.33	1.45	1.32	1.45	1.43	1.36	1.47	1.33	1.46
BINMEQ08	1.28	1.32	1.46	1.37	1.47	1.42	1.38	1.45	1.34	1.46
BINMEQ	1.31	1.31	1.47	1.34	1.46	1.39	1.40	1.45	1.37	1.46
BINMEQ01	1.29	1.30	1.48	1.32	1.46	1.40	1.39	1.45	1.33	1.45
BINMEQ02	1.28	1.30	1.49	1.30	1.46	1.40	1.39	1.45	1.32	1.47
BINMEQ03	1.29	1.30	1.47	1.31	1.46	1.39	1.39	1.45	1.33	1.46
BINMEQ04	1.30	1.30	1.46	1.34	1.46	1.40	1.41	1.46	1.33	1.47
BINMEQ05	1.30	1.30	1.45	1.34	1.46	1.40	1.40	1.46	1.33	1.46
RAWDOK	1.28	1.28	1.45	1.34	1.47	1.41	1.38	1.48	1.33	1.46

#### ***4.4b Risultati degli Studi di Docking***

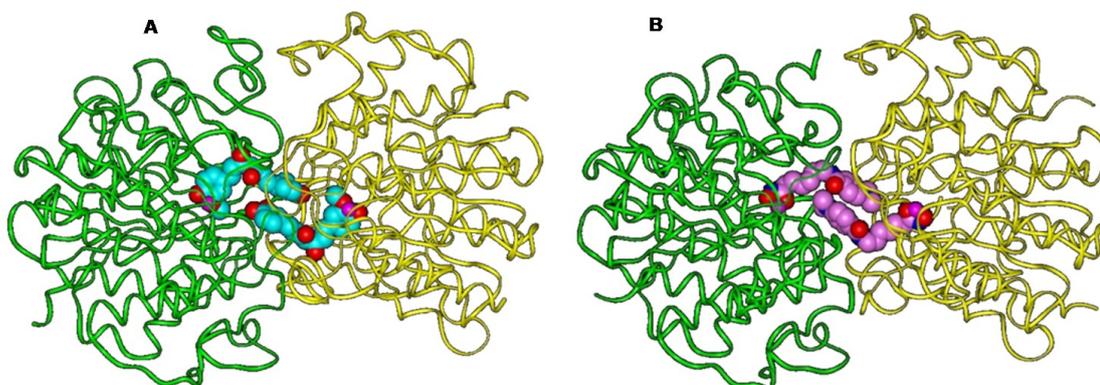
Sulla base degli studi effettuati, i calcoli di *docking* sono stati condotti sia sulla forma chetonica che sulla forma enolica della curcumina; in quanto seppur dalla ricerca nella *Cambridge Structural Database System* sia stata trovata solo la forma enolica, non si può escludere che tale forma possa esistere, seppure in minima parte, in soluzione, ed essere in grado di legare il sito attivo dell'enzima, venendo così progressivamente sottratta all'equilibrio. Inoltre, la curcumina potrebbe entrare nel sito attivo dell'enzima nella forma enolica e poi coordinare il metallo cedendo l'idrogeno dell'enolo ad un amminoacido circostante. Sono state ipotizzate due modalità di legame da noi chiamate modalità di legame I e modalità di legame II a seconda della diversa modalità di coordinazione dello zinco. In particolare, nella modalità di legame I lo zinco catalitico viene coordinato o dai due gruppi carbonilici del gruppo  $\beta$ -dichetone della curcumina nella forma chetonica o dal gruppo idrossile e dal gruppo carbonile dell'enolo della curcumina nella forma enolica; nella modalità di legame II, invece, lo zinco catalitico viene coordinato dai sostituenti 2-metossi-3-idrossi dell'anello fenilico della curcumina nella forma enolica.

I punti di partenza sono stati effettuati prendendo i minimi energetici dei conformeri ottenuti dalla ricerca conformazionale; visto che la forma chetonica presenta sempre una conformazione ripiegata, mentre quella enolica sempre una conformazione distesa come descritto nel paragrafo 6.4b, la modalità di legame II è stato possibile effettuarla solo per la forma enolica in quanto la conformazione ripiegata della forma chetonica non consente il corretto posizionamento e la corretta coordinazione dello zinco.

I risultati di *docking* hanno mostrato come la modalità di legame I è possibile sia per la forma chetonica che per la forma enolica, in quanto in entrambi i casi si va a mimare l'accomodamento dell'inibitore a struttura benzamidica co-cristallizzato nell'HDAC2 PDB ID: 3MAX (*Figura 6*). Il risultato degli studi di *docking* relativo alla modalità di legame II fa ipotizzare, invece, l'induzione di una possibile omodimerizzazione degli HDAC inibiti, così come osservato nel caso del complesso, sperimentalmente determinato, HDAC8/TSA (*Figura 7*)



**Figura 6:** rappresentazione dei risultati dei docking con la modalità di legame II della forma chetonica ed enolica nell'HDAC2 paragonati all'accomodamento della benzammide nell'HDAC2: nella figura A è rappresentato il risultato del docking della forma enolica nell'HDAC2; nella figura B è mostrato il risultato del docking della forma chetonica nell'HDAC2; nella figura C è mostrata la benzammide in complesso con l'HDAC2



**Figura 7:** rappresentazione della modalità di legame II della forma enolica nell'HDAC2 paragonata al dimero TSA/HDAC8: in particolare la figura A mostra l'ipotesi di dimerizzazione della forma enolica che coordina il metallo con la modalità di legame II; nella figura B, invece, viene mostrata la struttura cristallografica del dimero TSA/HDAC8

## ***4.5 Conclusioni***

Gli studi effettuati, anche se preliminari, hanno evidenziato la complessità del comportamento chimico della curcumina, probabilmente alla base del suo meccanismo d'azione ancor oggi non del tutto compreso; ma hanno anche permesso di definire due diverse modalità di interazione con gli enzimi HDAC, capaci di influenzarne sia l'attività catalitica che l'interazione con i cofattori. I saggi enzimatici sulle diverse isoforme degli HDAC saranno effettuati dal gruppo di ricerca del Dr. Ralph Mazitschek dell'Harvard Medical School mentre studi sulle interazioni molecolari saranno condotti dal gruppo di ricerca Dr. Cristiano Ferlini Danbury Hospital; i saggi farmacologici, invece, saranno eseguiti dalla Dr. Antonella Riva dell'Indena al fine di validare le ipotesi di interazione con gli enzimi HDAC di classe I; inoltre vista la molteplicità dei bersagli biologici della curcumina, ci si propone di investigare i suoi effetti anche su altri bersagli molecolari.

## **4.6 Bibliografia**

1. Gupta et. al. “*Natural Product Reports, Multitargeting by curcumin as revealed by molecular interaction studies*” *Natural Product Reports*, 2011, 28, 1937
2. Simone Reuter et.al. “*Epigenetic changes induced by curcumin and other natural compounds*” *Genes & Nutrition* , 2011, 6, 93–108
3. Alok Vyasa et.al. “*Perspectives on New Synthetic Curcumin Analogs and their Potential Anticancer Properties*” *Current Pharmaceutical Design*, 2013 , 19, 2047–2069.