

UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI SCIENZE DEL FARMACO

**CORSO DI LAUREA MAGISTRALE IN
CHIMICA E TECNOLOGIA FARMACEUTICHE**

TESI DI LAUREA

*Protocollo computazionale per l'identificazione di potenziali
inibitori della Casein Chinasi 1 di Plasmodium falciparum*

Relatore: Prof. Stefano Moro

Correlatore: Dott.ssa Veronica Salmaso

Laureanda: Anna Brocchetti

Anno Accademico: 2023/2024

Sommario

1. Abstract	1
2. La malaria	5
2.1 Cenni storici	5
2.2 La malaria oggi.....	7
2.3 Eziopatogenesi della malaria.....	9
2.4 I <i>Plasmodi</i> responsabili della malaria umana.....	10
2.5 Manifestazioni cliniche della malaria.....	12
2.6 Profilassi e terapia antimalarica	13
2.7 Resistenza ai farmaci antimalarici.....	18
2.8 Vaccini	19
3. I chinomi di Homo sapiens e del Plasmodium	20
3.1 PfCK1.....	21
3.2 Casein chinasi 1.....	21
3.3 CK1 δ umana.....	21
4. Scopo della tesi	24
5. Hardware e Software	27
6. Virtual Screening	29
6.1 <i>Homology modeling</i> ⁴⁶	31
6.2 Operazioni preliminari al <i>virtual screening</i>	34
6.3 Meccanica molecolare.....	36
6.4 <i>Docking</i>	38
6.5 Dinamica molecolare.....	41
6.6 Modello Farmacoforico.....	42
6.7 PLIF (<i>Protein Ligand Interaction Fingerprint</i>)	43
6.8 <i>Solvent exposure</i>	43
7. Generazione del modello di omologia	48
7.1 Validazione del modello.....	50
7.2 Interazione con la Leu85 del <i>backbone</i>	57
8. Scelta della libreria di ligandi	62

9. Docking	63
9.1 Filtraggio delle <i>poses</i>	63
9.2 Generazione del farmacoforo	63
9.3 Filtro per farmacoforo	64
9.4 Filtro per PLIF (<i>Protein Ligand Interaction Fingerprint</i>).....	65
9.5 Ulteriori filtri.....	66
10. Caratterizzazione dei composti selezionati	67
11. Indagine sulla selettività dei composti selezionati nei confronti di CK1δ umana	88
12. Conclusione	93

Indice delle figure

Figura 1: C. L. A. Laveran.....	5
Figura 2: La mappa mostra il tasso di incidenza previsto per tutte le età (casi clinici per 1.000 abitanti all'anno) della malaria da <i>Plasmodium falciparum</i> nel 2020 ⁴	7
Figura 3: il ciclo replicativo del <i>Plasmodium falciparum</i> all'interno dell'organismo umano.	10
Figura 4: Le due immagini raffigurano due vetrini di sangue. Il vetrino a sinistra è infettato da <i>Plasmodium falciparum</i> , mentre quello a destra da <i>Plasmodium vivax</i> . È possibile notare le differenze morfologiche tra le due specie più diffuse di <i>Plasmodium</i>	11
Figura 5: Nell'immagine sono mostrati i principali segni e sintomi della malaria.....	12
Figura 6: struttura della cloroquina.	13
Figura 7: strutture dei più noti farmaci chinolinici e focus sulla somiglianza con la chinina. (Da Foye's – Principi di Chimica Farmaceutica, ed. Piccin).....	14
Figura 8: struttura della sulfadossina.	15
Figura 9: struttura della pirimetamina.	15
Figura 10: struttura dell'atovaquone.	16
Figura 11: struttura del proguanil.....	16
Figura 12: l' <i>Artemisia annua</i>	17
Figura 13: struttura dell'artemisinina e dei suoi derivati. Da ¹³	17
Figura 14: Stato globale di resistenza a cloroquina e sulfadossina/pirimetamina, i due farmaci antimalarici più ampiamente utilizzati. I dati provengono dall'OMS.....	18
Figura 15: Albero filogenetico del chinoma di <i>Plasmodium falciparum</i> . In blu sono evidenziate le chinasi del parassita discusse come bersaglio di scoperte farmacologiche in corso. ¹⁷	20
Figura 16: Figura 16. Struttura tridimensionale di CK1 delta umana in forma apo, da PDB (5IH4).....	22
Figura 17: organizzazione informatica interna al laboratorio MMS.....	27
Figura 18: timeline del percorso di scoperta e sviluppo di un farmaco. ⁴⁵	29
Figura 19: schema della serie di processi da svolgere al fine di generare un modello di omologia ⁴⁶	33
Figura 20: esemplificazione grafica del docking molecolare proteina-ligando.....	38
Figura 21: Allineamento delle sequenze amminoacidiche di CK1 (6RCG) e PfCK1.....	50
Figura 22: legenda delle interazioni 2D tra proteina e ligando.....	52
Figura 23: grafico dell'andamento del RMSD in funzione del tempo.	56
Figura 24: rappresentazione grafica degli atomi coinvolti nel legame a idrogeno bidentato con il residuo Leu85.....	57
Figura 25: Struttura tridimensionale di CK1delta umana (PDB: 6RCG).....	58
Figura 26: Struttura tridimensionale del modello di omologia di PfCK1 generato con MOE.....	58

<i>Figura 27: Struttura tridimensionale del modello di omologia di PfCK1 dopo il rilassamento in dinamica.</i>	59
<i>Figura 28: Sovrapposizione tra il modello di PfCK1 pre (in viola) e post (in rosso) dinamica.</i>	59
<i>Figura 29: Sovrapposizione tra la struttura cristallografica di 6RCG (in verde), e il modello di PfCK1 pre (in viola) e post (in rosso) dinamica.</i>	60
<i>Figura 30: Tabella con i valori di RMSD tra le strutture 6RCG modello pre-MD e post-MD.</i>	60
<i>Figura 31: Rappresentazione grafica del modello farmacoforico.</i>	64
<i>Figura 32: pose selezionata e rappresentazione 2D delle interazioni con il target di MolPort-002-054-502.</i>	67
<i>Figura 33: pose selezionata e rappresentazione 2D delle interazioni con il target di MolPort-001-803-629.</i>	69
<i>Figura 34: pose selezionata e rappresentazione 2D delle interazioni con il target di MolPort-002-527-585.</i>	71
<i>Figura 35: pose selezionata e rappresentazione 2D delle interazioni con il target di MolPort-001-987-036.</i>	73
<i>Figura 36: pose selezionata e rappresentazione 2D delle interazioni con il target di MolPort-002-510-879.</i>	75
<i>Figura 37: pose selezionata e rappresentazione 2D delle interazioni con il target di MolPort-002-535-751.</i>	77
<i>Figura 38: pose selezionata e rappresentazione 2D delle interazioni con il target di MolPort-002-529-576.</i>	79
<i>Figura 39: pose selezionata e rappresentazione 2D delle interazioni con il target di MolPort-007-825-628.</i>	80
<i>Figura 40: pose selezionata e rappresentazione 2D delle interazioni con il target di MolPort-035-700-089.</i>	82
<i>Figura 41: pose selezionata e rappresentazione 2D delle interazioni con il target di MolPort-029-884-875.</i>	84
<i>Figura 42: pose selezionata e rappresentazione 2D delle interazioni con il target di MolPort-005-911-320.</i>	86
<i>Figura 43: allineamento tra le sequenze amminoacidiche di CK1δ umana (PDB_{ID}: 6RCG) e PfCK1. In verde sono evidenziati i residui del pocket che vengono conservati tra i due omologi, mentre in rosso i residui mutati.</i>	89
<i>Figura 44: sovrapposizione tra la struttura cristallografica di 6RCG (in verde) e il modello di omologia di PfCK1 (in rosso). Vengono mostrate le catene laterali degli amminoacidi mutati tra le due proteine.</i>	90
<i>Figura 45: A sinistra, viene mostrata la pose del composto MolPort-029-884-875 nella tasca di binding di PfCK1 mutata, a destra le interazioni con il ligando in 2D.</i>	91

Figura 46: a sinistra, la pose del composto MolPort -005-911-320 nella tasca di binding di PfCK1 mutata, a destra vengono mostrate le interazioni con il ligando in 2D. 92

Indice delle tabelle

<i>Tabella 1: Classificazione di farmaci antimalarici, divisi per categoria.</i>	<i>13</i>
<i>Tabella 2: Software utilizzati nello svolgimento del lavoro di tesi.</i>	<i>28</i>
<i>Tabella 3: Risultato della ricerca della sequenza amminoacidica ottenuta tramite Protein BLAST.</i>	<i>49</i>
<i>Tabella 4: Nomenclatura, Kd e peso molecolare dei due ligandi veri positivi per PfCK1.</i>	<i>51</i>
<i>Tabella 5: Confronto tra le pose dei due ligandi veri positivi sul modello di PfCK1 pre-dinamica.</i>	<i>52</i>
<i>Tabella 6: Rappresentazione 3D e 2D della pose di docking ottenuta post rilassamento in dinamica molecolare con il ligando ChEMBL4461850.</i>	<i>54</i>
<i>Tabella 7: confronto tra la pose pre e post-MD del ligando ChEMBL4591417.</i>	<i>55</i>
<i>Tabella 8: confronto tra la pose pre e post-MD del ligando ChEMBL4461850.</i>	<i>55</i>
<i>Tabella 9: Nomenclatura e proprietà chimiche di Molport-002-054-502.</i>	<i>67</i>
<i>Tabella 10: nomenclatura e proprietà chimiche di Molport-001-803-629.</i>	<i>69</i>
<i>Tabella 11: nomenclatura e proprietà chimiche di Molport-002-057-585.</i>	<i>71</i>
<i>Tabella 12: nomenclatura e proprietà chimiche di Molport-001-987-036.</i>	<i>73</i>
<i>Tabella 13: nomenclatura e proprietà chimiche di Molport-002-510-879.</i>	<i>75</i>
<i>Tabella 14: nomenclatura e proprietà chimiche di Molport-002-535-751.</i>	<i>77</i>
<i>Tabella 15: nomenclatura e proprietà chimiche di Molport-002-529-576.</i>	<i>79</i>
<i>Tabella 16: nomenclatura e proprietà chimiche di Molport-007-825-628.</i>	<i>80</i>
<i>Tabella 17: nomenclatura e proprietà chimiche di Molport-035-700-089.</i>	<i>82</i>
<i>Tabella 18: nomenclatura e proprietà chimiche di Molport-029-884-875.</i>	<i>84</i>
<i>Tabella 19: nomenclatura e proprietà chimiche di Molport-005-911-320.</i>	<i>86</i>

Indice delle equazioni

<i>Equazione 1. Equazione di potenziale.....</i>	<i>36</i>
<i>Equazione 2. Potenziale di stretching.....</i>	<i>37</i>
<i>Equazione 3. Potenziale di bending.....</i>	<i>37</i>
<i>Equazione 4. Potenziale dei diedri.....</i>	<i>37</i>
<i>Equazione 5. Potenziale elettrostatico.....</i>	<i>37</i>
<i>Equazione 6. Potenziale di van der Waals.....</i>	<i>37</i>
<i>Equazione 7: formula per il calcolo della costante di dissociazione K_d.....</i>	<i>51</i>
<i>Equazione 8: Equazione del RMSD.....</i>	<i>56</i>

1. Abstract

La lotta contro la malaria ha attraversato secoli di progressi scientifici, dalla sua prima descrizione nei testi di Ippocrate fino alle moderne terapie antimalariche e ai recenti sviluppi nel campo dei vaccini. Sebbene siano stati fatti notevoli progressi, la malaria persiste come una delle principali emergenze sanitarie globali, soprattutto nelle regioni tropicali e subtropicali dell’Africa, dell’Asia e del Sud America.

La malaria viene trasmessa all’uomo attraverso la puntura di una zanzara del genere *Anopheles*. L’agente eziologico della malattia è il *Plasmodium*, un protozoo appartenente al phylum degli *Apicomplexa*. Questo studio si concentra su *Plasmodium falciparum*, la specie di *Plasmodium* più diffusa e che causa la sintomatologia più grave. Inoltre, la comparsa di ceppi resistenti ai farmaci antimalarici più utilizzati rappresenta una minaccia crescente e sottolinea la necessità di approcci terapeutici innovativi.

La recente approvazione da parte dell’OMS del vaccino RTS,S/AS01 segna un traguardo significativo nella prevenzione della malaria. Tuttavia, l’utilizzo sinergico di vaccini, farmaci antimalarici e misure di controllo delle zanzare è fondamentale per affrontare in modo completo ed efficiente la malattia.

La malaria continua, dunque, a richiedere un impegno globale e collaborativo per raggiungere gli obiettivi di controllo e eradicazione. In questo contesto si inserisce la ricerca farmacologica, che si sta avvalendo sempre più di approcci computazionali avanzati per accelerare il processo di scoperta e sviluppo di nuovi farmaci. Questo studio si concentra sull’applicazione di due metodologie chiave: il *docking* molecolare e la dinamica molecolare.

Il *docking* molecolare è uno strumento utilizzato per simulare e studiare a livello strutturale l’interazione tra molecole e i loro potenziali *target* biologici. Viene sfruttato il suo ruolo nel predire interazioni intermolecolari e valutare la geometria dei complessi proteina-ligando.

Le simulazioni di dinamica molecolare possono rivelare informazioni dettagliate sulla flessibilità dei sistemi studiati e sulle loro variazioni conformazionali, contribuendo così alla comprensione dell’evoluzione tempo-dipendente del sistema molecolare in esame.

Questo studio presenta un protocollo di *virtual screening* sviluppato per selezionare composti naturali, o loro derivati, con potenziale attività inibitoria nei confronti di Casein Chinasi 1, una proteina appartenente al chinoma di *Plasmodium falciparum* coinvolta nella replicazione del parassita.

La nostra analisi ha coinvolto una vasta libreria di più di 113mila composti naturali e loro derivati, con l'obiettivo di identificare molecole di interesse per lo sviluppo farmaceutico.

Il protocollo applicato si basa su una combinazione di approcci computazionali, consistenti nella generazione di un modello strutturale e nella sua validazione, nel *virtual screening* tramite *docking* molecolare della libreria di composti, e nella applicazione di una serie di filtri, che hanno permesso di ridurre razionalmente la numerosità della libreria, consentendo di prioritizzare i candidati più promettenti.

Il nostro protocollo ha permesso di identificare 11 composti con potenziale attività inibitoria nei confronti di PfCK1. Questi candidati sono emersi come risultato della valutazione di proprietà chimiche e interazionali con il bersaglio biologico di interesse.

La nostra ricerca è dunque incentrata sullo sviluppo di un protocollo capace di selezionare un insieme di composti naturali con potenziale applicazione farmaceutica nell'ambito della malaria, come potenziali inibitori di Casein Chinasi 1 di *Plasmodium falciparum*. La priorità assegnata a questi 11 candidati offre una base per ulteriori studi sperimentali.

Parte prima:

INTRODUZIONE

2. La malaria

2.1 Cenni storici

La malaria nell'antichità

Nell'antichità, le prime testimonianze scritte della malattia che oggi identifichiamo con il termine "malaria" si ritrovano nei testi dell'illustre medico greco Ippocrate, nato nel 460 a.C circa e morto non prima del 377 a.C. Nei suoi scritti delle Epidemie, datati intorno al IV secolo a.C., vengono descritte dettagliatamente le febbri intermittenti caratteristiche della malattia, la terzana e la quartana¹.

Ippocrate documenta gli episodi di febbre terzana benigna e quartana maligna, che in tempi successivi sono stati associati all'infezione rispettivamente da *Plasmodium vivax* e da *Plasmodium malariae*. Inoltre, il medico greco nota la comparsa di splenomegalia come sintomo correlato, particolarmente evidente tra coloro che risiedevano in regioni paludose.

Per quanto concerne l'Italia, secoli dopo, il medico romano Galeno (130 d.C. - 216 d.C.) avanzò l'ipotesi che la malattia potesse essere contratta attraverso l'inalazione delle esalazioni provenienti da zone caratterizzate da acque stagnanti e paludose. Da questa teoria deriva l'etimologia della parola "malaria", con "mal'aria" indicante il presunto veicolo di contagio.

La scoperta del Plasmodium

I primi studi microbiologici sulla malaria si svolsero nel XIX secolo, in un contesto intriso di interesse scientifico, alimentato dallo sviluppo degli imperi coloniali europei e dalla necessità di difendere coloni e comunità dall'insidiosa infezione malarica.

Fu il chirurgo dell'esercito francese Charles Louis Alphonse Laveran (Figura 1), nel 1878, a compiere una pionieristica osservazione attraverso la microscopia ottica: la visualizzazione dei gametociti di *Plasmodium* nel sangue dei suoi pazienti. Il 6 novembre 1880, presso l'ospedale di Constantine in Algeria, Laveran registrò la prima osservazione del *Plasmodium vivax*, visualizzando i gametociti in un giovane soldato di 24 anni².

Questo evento si rivelò fondamentale nel dimostrare che l'"impaludismo", termine utilizzato all'epoca, aveva origine parassitaria anziché batterica.

Per l'importanza delle sue scoperte, Laveran venne insignito del premio Nobel per la medicina nel 1907.

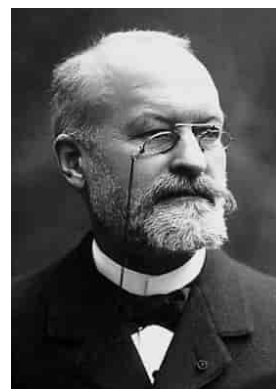


Figura 1: C. L. A. Laveran.

La successiva rivelazione del ruolo cruciale delle zanzare come vettori esclusivi della malaria fu guidata dall'intuizione del medico inglese Patrick Manson. Basandosi sulla sua precedente scoperta riguardante la trasmissione di altri parassiti del sangue tramite zanzare, Manson formulò l'ipotesi che anche i parassiti della malaria potessero essere veicolati da questi insetti, stabilendo così un legame tra la malattia e le regioni paludose sovrappopolate da zanzare.

Nel 1897, in India, il chirurgo dell'esercito britannico Ronald Ross dimostrò che il *Plasmodium relictum*, agente eziologico della malaria negli uccelli, veniva trasmesso dalle zanzare *Culex*, confermando, due anni dopo, che i parassiti della malaria umana potevano essere veicolati dalle zanzare *Anopheles*.

Tuttavia, è doveroso menzionare che già nel 1898, i malariologi italiani Giovanni Battista Grassi, Amico Bignami e Giuseppe Bastianelli avevano fornito la prova formale della trasmissione della malaria da uomo a uomo attraverso le zanzare *Anopheles*.

Nei due anni successivi, gli stessi ricercatori italiani compresero in dettaglio il ciclo completo di *P. vivax*, *P. falciparum* e *P. malariae* all'interno della zanzara. Ulteriori progressi furono compiuti negli anni successivi, culminando nella scoperta nel 1947 di schizonti pre-eritrocitari di *Plasmodium cynomolgi* nel fegato di una scimmia *Rhesus* da parte di Henry Shortt e Cyril Garnham. Infine, nel 1982, Wojciech Krotoski dimostrò la presenza di stadi dormienti di *P. vivax* nel fegato, spiegando così le ricadute della malaria³.

In conclusione, oltre un secolo di impegno da parte di centinaia di scienziati è stato necessario per chiarire e comprendere appieno il ciclo completo dei parassiti della malaria umana.

La malaria in Italia

Si stima che in Italia tre specie di *Plasmodium* (*P. vivax*, *P. falciparum* e *P. malariae*) fossero endemiche dal II secolo a.C., soprattutto nelle regioni paludose del Veneto, dell'Italia centro-meridionale e delle isole.

Dal 1928, la lotta contro la malaria assunse un ruolo centrale nella politica fascista italiana. La strategia adottata prevedeva l'impiego del chinino in combinazione con tre pratiche di bonifica: idraulica, mediante il drenaggio delle paludi e il controllo delle fonti d'acqua; agraria, con il ripopolamento dell'area attraverso incentivi agricoli; igienica, promuovendo condizioni abitative sicure con murature imbiancate. Particolare attenzione fu rivolta all'Agro Pontino, a sud di Roma, come area su cui concentrare gli sforzi.

Tuttavia, con l'ingresso dell'Italia nella Seconda Guerra Mondiale nel 1940, la lotta antimalarica subì un rallentamento, portando a un aumento dei casi fino al 1945. Il piano sanitario che instradò successivamente l'eliminazione della malaria fu implementato grazie all'*United Nations Relief and*

Rehabilitation Administration (UNRRA) e alla *Rockefeller Foundation*. Inoltre, con l'introduzione di nuovi pesticidi come il DDT (para-diclorodifeniltricloroetano), fu possibile effettuare campagne intensive su tutto il territorio italiano.

Nel 1962, si registrò l'ultimo caso a Palermo, mentre nel 1965 si verificò l'ultimo episodio di trasmissione di *P. malariae* da parte di una zanzara locale.

Cinquantaquattro anni fa, il 17 novembre del 1970, l'Organizzazione Mondiale della Sanità (OMS) certificò ufficialmente il debellamento della malaria in Italia.

2.2 La malaria oggi

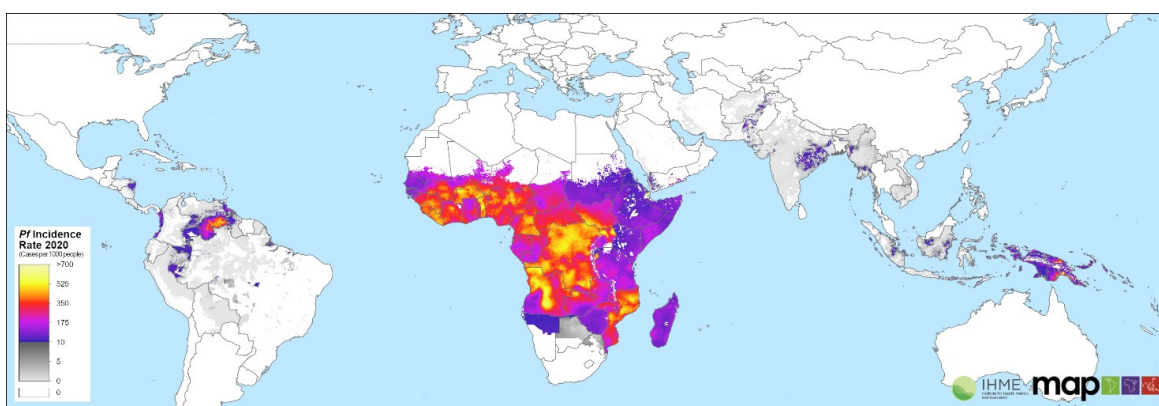


Figura 2: La mappa mostra il tasso di incidenza previsto per tutte le età (casi clinici per 1.000 abitanti all'anno) della malaria da *Plasmodium falciparum* nel 2020⁴.

Nonostante i progressi compiuti nel controllo e nell'eliminazione della malattia nel corso degli anni, la malaria persiste come una delle principali emergenze sanitarie mondiali, affliggendo in particolare le regioni tropicali e subtropicali del pianeta, come l'Africa Sub-Sahariana, l'Asia e l'America del Sud. Insieme alla tubercolosi e all'AIDS, la malaria continua a mietere milioni di vite, con un impatto devastante sulla salute globale.

Le statistiche evidenziano che circa l'80% dei casi e il 90% delle morti sono concentrati tra le donne in gravidanza e i bambini al di sotto dei 5 anni. Ogni anno, il World Malaria Report dell'OMS offre una panoramica dettagliata delle tendenze globali nella lotta contro la malattia, tenendo conto dei progressi nelle diverse fasi di intervento, come la prevenzione, la diagnosi, il trattamento, l'eliminazione e la sorveglianza.

Il rapporto del 2022, basato su informazioni provenienti da 84 Paesi endemici, rivela che nel 2021 si sono registrati 247 milioni di casi di malaria, con un incremento rispetto all'anno precedente. Tra il 2019 e il 2021, si stima che 13,4 milioni di casi aggiuntivi siano attribuibili all'interruzione dei servizi essenziali durante la pandemia di COVID-19.

Nonostante i progressi registrati tra il 2000 e il 2019, con una costante riduzione dei decessi globali per malaria, il 2020 ha visto un aumento del 10% rispetto all'anno precedente, con una stima di 625.000 morti. Nel 2021, sebbene si sia assistito a una lieve diminuzione, i decessi per malaria si sono attestati 619.000, sottolineando l'urgente necessità di rafforzare gli sforzi nel controllo e nella gestione di questa emergenza sanitaria persistente⁵.

2.3 Eziopatogenesi della malaria

L'agente eziologico della malaria è il *Plasmodium*, un protista appartenente al phylum *Apicomplexa*, caratterizzato dalla presenza di un organello specializzato nell'invasione, chiamato "complesso apicale".

Il ciclo vitale di questo protista si svolge all'interno di tessuti differenti di due ospiti distinti: un ospite intermedio (un organismo vertebrato) e un ospite definitivo (una zanzara del genere *Anopheles*). Quando una zanzara infetta punge un ospite intermedio, gli sporozoiti vengono iniettati nel derma e, attraverso la circolazione sanguigna, raggiungono il fegato, infettando gli epatociti e dando inizio al "ciclo schizogonico eso-eritrocitario".

Durante questa fase epatica, gli sporozoiti si sviluppano prima in trofozoita epatico e successivamente, attraverso divisioni nucleari, in schizonte, liberando centinaia di merozoiti. Uno o più merozoiti invadono un singolo eritrocita, avviando il "ciclo intra-eritrocitario". Il parassita attraversa gli stadi di trofozoita giovane, trofozoita maturo e schizonte, dando origine a 32 merozoiti. Questi provocano la lisi dei globuli rossi, causando l'infiammazione e i sintomi iniziali della malaria, come la febbre.

In risposta a stimoli o stress cellulari, alcuni schizonti si evolvono in forme pre-sessuate, chiamate gametociti femminili e maschili. Durante il pasto di sangue della zanzara, i gametociti vengono prelevati e, nel lume intestinale dell'insetto, generano un gamete femminile e otto gameti maschili mobili. Questi si uniscono per formare uno zigote, chiamato oocinete, unica cellula diploide del ciclo vitale del parassita. L'oocinete è in grado di penetrare la parete intestinale e di raggiungere la lamina basale, dove forma un'oocisti, che dà inizio ad una divisione meiotica e ad una serie di divisioni mitotiche da cui si sviluppano numerosi sporozoiti. Una volta maturi, gli sporozoiti rompono l'oocisti e si riversano nell'emocele dell'insetto, dove raggiungono le ghiandole salivari. Da qui restano in attesa di essere inoculati nel nuovo ospite durante il successivo pasto di sangue, completando così il complesso ciclo vitale del *Plasmodium*⁶.

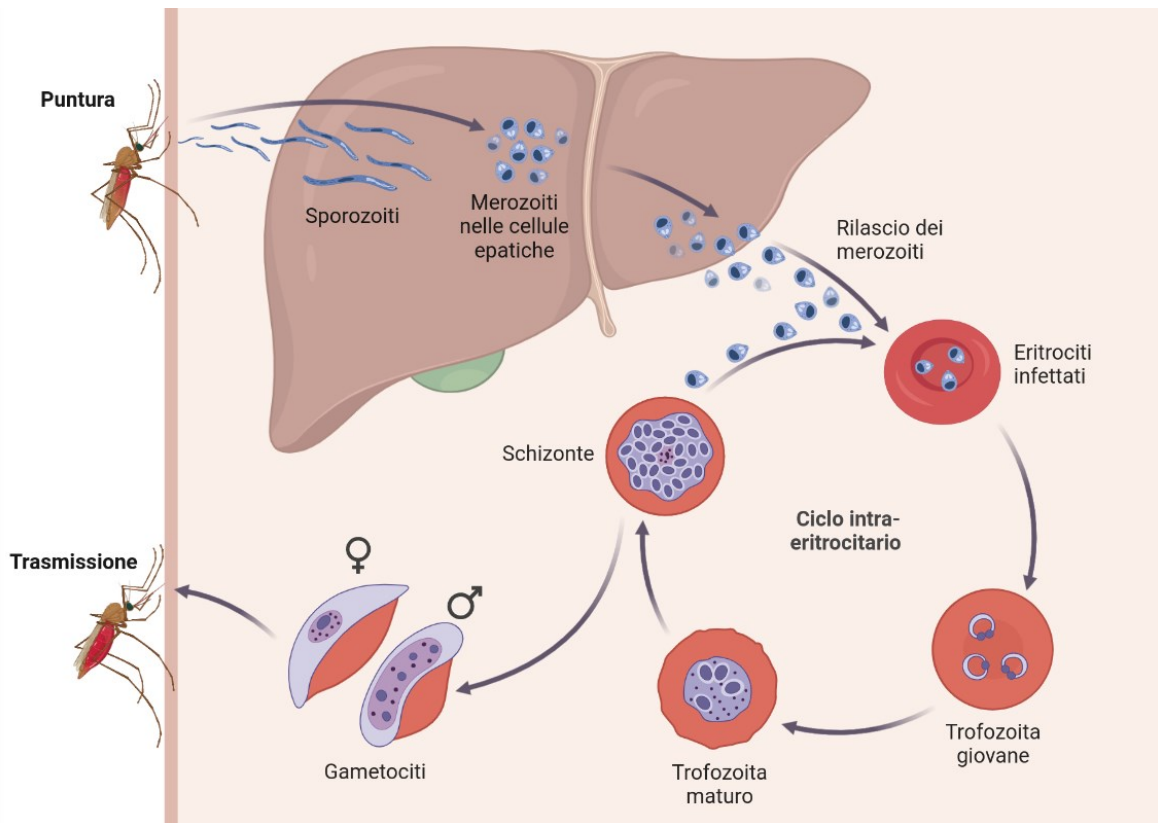


Figura 3: il ciclo replicativo del *Plasmodium falciparum* all'interno dell'organismo umano.

2.4 I *Plasmodi* responsabili della malaria umana

Le specie di *Plasmodium* responsabili della malaria nell'essere umano sono:

1. ***Plasmodium falciparum***, dalla caratteristica forma a falce, è il responsabile della cosiddetta “terzana maligna”. Ha un'incubazione di 1-3 settimane ed effettua un ciclo di replicazione nel sangue umano ogni 48 ore. Endemico nelle zone dell'Africa sub-Sahariana, Asia e America Latina, Papua-Nuova Guinea e in alcune isole del Pacifico orientale. È la specie che provoca la sintomatologia più grave.
2. ***Plasmodium vivax***, con periodo di incubazione di 1-4 settimane, è responsabile della “terzana benigna”. È diffuso prevalentemente in Africa orientale, Asia e America Latina.
3. ***Plasmodium ovale***, con periodo di incubazione di 9-18 giorni, molto comune principalmente in Africa occidentale. Negli individui che vengono infettati da questa specie sono comuni delle ricadute causate dalla capacità del parassita di rimanere in forma silente (ipnozoita) nel tessuto epatico per lunghi periodi di tempo.
4. ***Plasmodium malariae***, responsabile della “quartana”, che ha un'incubazione di 2-4 settimane e il cui ciclo asessuato si completa ogni 72 ore. È diffuso in aree tropicali e subtropicali in modo non uniforme e con una frequenza piuttosto bassa

5. *Plasmodium knowlesi*, morfologicamente somigliante a *P. malariae*, presente in varie aree del Sud-est asiatico e responsabile della malaria di alcune scimmie. A causa dei numerosi focolai umani identificati in alcune regioni del Sud-est asiatico, dal 2008 viene considerato a tutti gli effetti una delle specie di plasmodio responsabile di malaria nell'uomo.

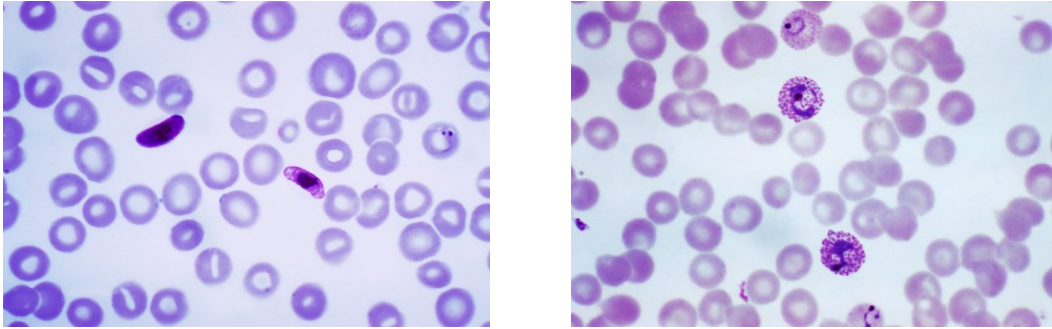


Figura 4: Le due immagini raffigurano due vetrini di sangue. Il vetrino a sinistra è infettato da *Plasmodium falciparum*, mentre quello a destra da *Plasmodium vivax*. È possibile notare le differenze morfologiche tra le due specie più diffuse di *Plasmodium*.

2.5 Manifestazioni cliniche della malaria

I segni iniziali della malaria si manifestano con la lisi dei globuli rossi infetti, un evento cruciale al termine della fase intraeritrocitaria. Tale processo induce un'inflammatione sistemica, culminando con la comparsa della febbre, primario sintomo della malattia. L'immagine sottostante mostra i sintomi successivi che coinvolgono vari organi del corpo.

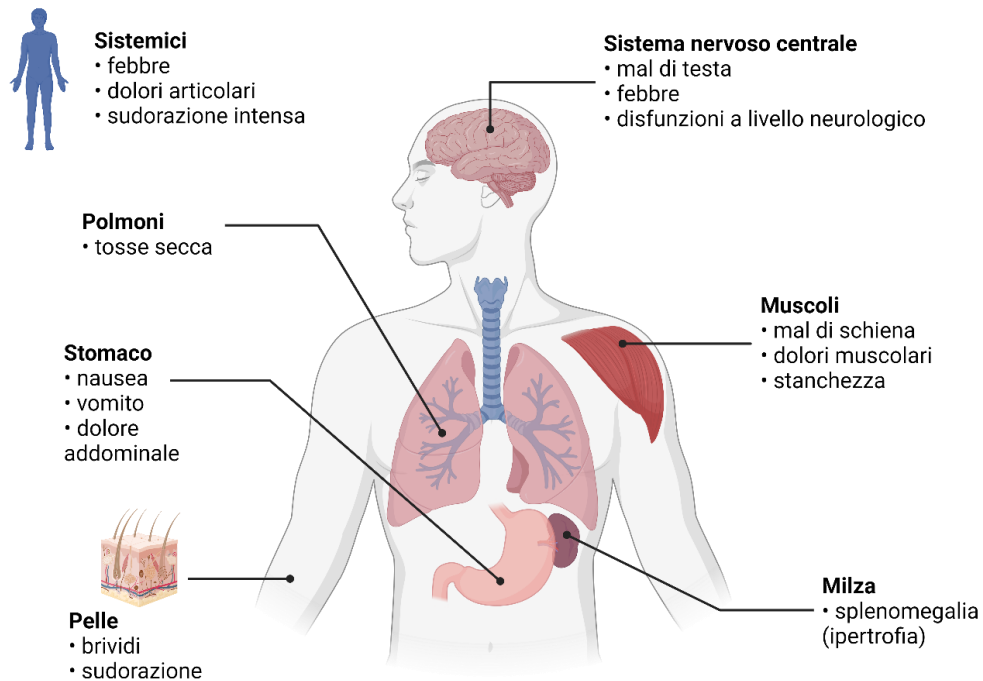


Figura 5: Nell'immagine sono mostrati i principali segni e sintomi della malaria.

Oltre agli effetti della patologia riportati precedentemente, è stato dimostrato come la malaria cerebrale sia una delle cause maggiori di disfunzioni a livello neurologico nei bambini africani⁷, in quanto non solo comporta gravi implicazioni durante il suo corso, ma ha dimostrato di persistere nelle conseguenze anche dopo la guarigione⁷.

2.6 Profilassi e terapia antimalarica

La terapia della malaria si avvale di una vasta gamma di farmaci appartenenti ad almeno cinque classi distinte, caratterizzate da differenze a livello di principio attivo, meccanismo d'azione e target molecolare. (Tabella 1).

CHINOLINICI	ANTIFOLATI	ATOVAQUONE	ARTEMISININE	ANTIBIOTICI
Chinina	Pirimetamina	Atovaquone	Artemisina	Tetraciclina
Chinidina	Sulfadossina		Artemeter	Tossiciclina
Cloroquina	Proguanil		Arteeter	Clindamicina
Alofantrina			Artesunato	
Meflochina			Diidroartemisina	
Lumefantrina				

Tabella 1: Classificazione di farmaci antimalarici, divisi per categoria.

Chinolinici

La cloroquina è il farmaco di elezione per fronteggiare la malaria causata da *P. vivax*, *P. ovale* e *P. malariae*. Non può essere utilizzato nel caso di *P. falciparum*, a causa della marcata resistenza al farmaco sviluppata nelle regioni endemiche del Sud America e in gran parte dell'Asia, Oceania e Africa occidentale.

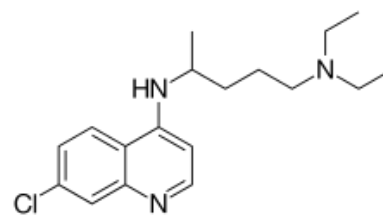


Figura 6: struttura della cloroquina.

La chinina è un alcaloide naturale originariamente derivato dalla corteccia della *Cinchona*, utilizzato come rimedio tradizionale in Sud America per le febbri intermittenzi. Si tratta del più antico farmaco impiegato per curare la malaria fin dal 1820, ma conserva la sua rilevanza nei confronti dei ceppi di *Plasmodium falciparum* resistenti alla cloroquina. Si conferma essenziale nella gestione delle manifestazioni gravi e complesse della malattia, grazie alla sua buona biodisponibilità e alla possibilità di essere somministrato sia per via parenterale che endovenosa.

La chinidina è lo stereoisomero destrogiro della chinina, utilizzata anch'essa come antimalarico. Tra le due l'efficacia terapeutica è simile, con l'inconveniente nella prima di una tossicità più marcata.

Altri farmaci appartenenti a questa categoria sono: Meflochina, Alofantrina, Lumefantrina, Pamachina e Primachina.

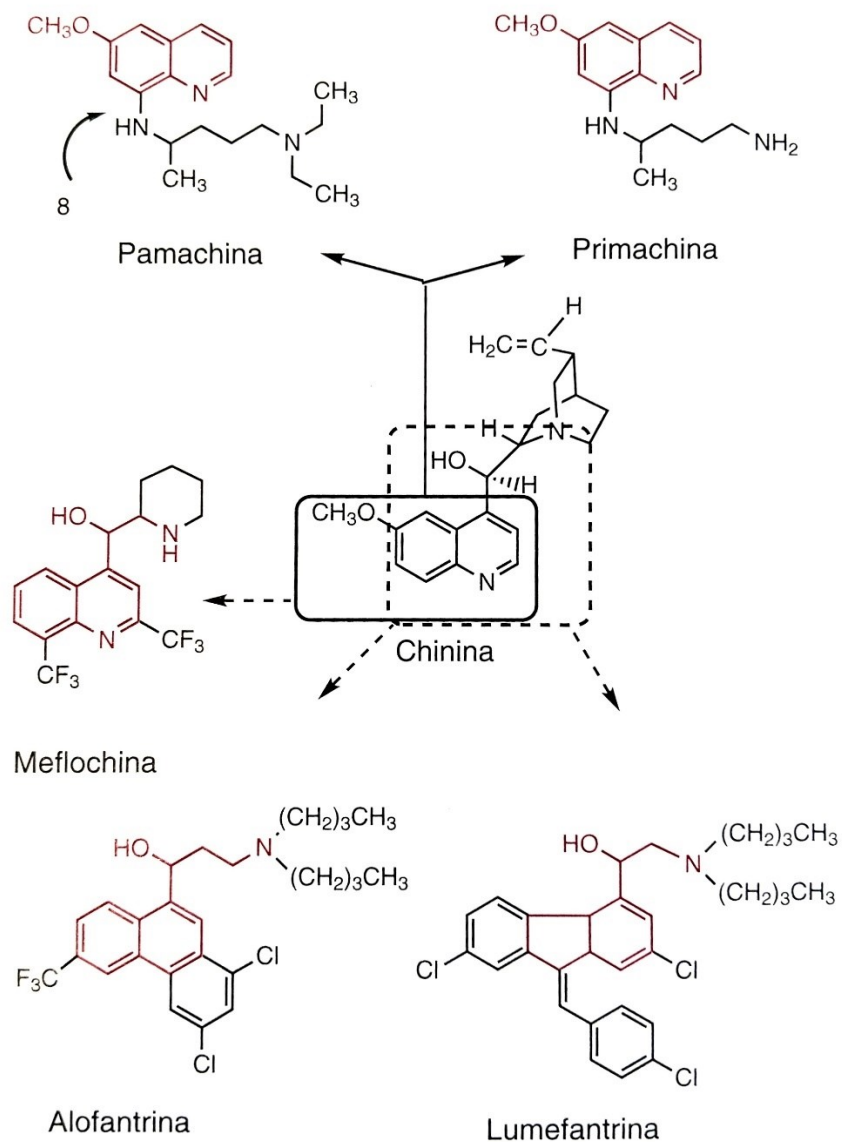


Figura 7: strutture dei più noti farmaci chinolinici e focus sulla somiglianza con la chinina. (Da Foye's – Principi di Chimica Farmaceutica, ed. Piccin).

Antifolati

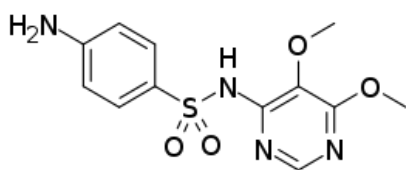


Figura 8: struttura della sulfadossina.

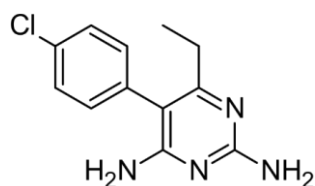


Figura 9: struttura della pirimetamina.

Il farmaco principale appartenente alla classe degli antifolati è il *Fansidar*. Esso è composto dalla sulfadossina, che interagisce con l'enzima diidropteroato sintasi (DHPS), e dalla pirimetamina, un inibitore della diidrofolato reduttasi (DHFR).

L'interazione sinergica di sulfadossina e pirimetamina è cruciale per il meccanismo d'azione del *Fansidar*. La sulfadossina ostacola l'attività dell'enzima DHPS, interferendo con la sintesi del tetraidrofolato, mentre la pirimetamina agisce come inibitore della DHFR, bloccando la conversione del diidrofolato in tetraidrofolato, una molecola necessaria per la sintesi di aminoacidi e acidi nucleici. Questa doppia azione compromette la replicazione del DNA parassitario, contribuendo così ad una buona efficacia antimalarica.

Particolarmente rilevante è il ruolo del *Fansidar* nel contesto africano, dove ha assunto la posizione di trattamento di prima linea per la malaria da *P. falciparum* resistente alla cloroquina, sia a livello preventivo che a livello di trattamento effettivo della patologia.

La resistenza al *Fansidar* è però un problema rilevante, in quanto in alcune aree del mondo, tra cui l'Africa Orientale⁸⁹, l'Asia¹⁰ e l'America Latina¹¹, questo farmaco risulta meno efficace contro l'infezione da *Plasmodium falciparum*.

Atovaquone

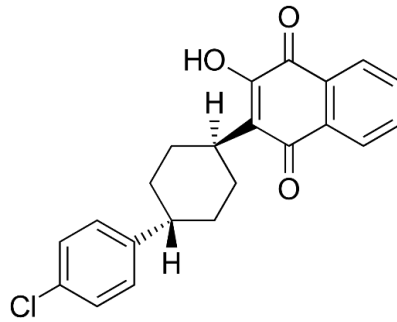


Figura 10: struttura dell'atovaquone.

L'atovaquone si distingue per l'inibizione della Citocromo c reduttasi, un complesso enzimatico vitale nella catena di trasporto degli elettroni all'interno dei mitocondri.

La sua capacità di agire come inibitore si deve alla sua notevole somiglianza strutturale con il suo substrato naturale, il coenzima Q, anche detto Ubichinone.

Per mitigare il rischio di sviluppare resistenze, l'atovaquone viene combinato all'antifolato proguanil, dando vita a un farmaco denominato *Malarone*. È importante notare che l'utilizzo di questa combinazione è circoscritto da una limitazione finanziaria, a causa dell'elevato costo del atovaquone¹².

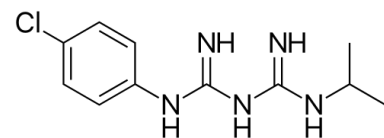


Figura 11: struttura del proguanil.

Artemisinine

L'artemisinina e i suoi derivati sono stati isolati dall'*Artemisia annua*, ed erano usati in Cina già dal 168 a.C.

Per migliorarne l'emivita ed aumentarne la selettività sono state introdotte delle modifiche nella sua struttura originaria. La prima modifica coinvolse il gruppo chetonico in posizione 10, ottenendo così la diidroartemisinina.



Figura 12: l'*Artemisia annua*

Successivamente, ulteriori modifiche della molecola portarono ad ottenere composti tra cui l'artesunato, estere dell'acido succinico, somministrabile per via orale e per via endovenosa; e l'artemether, estere lipofilo, che può essere assunto per via orale.

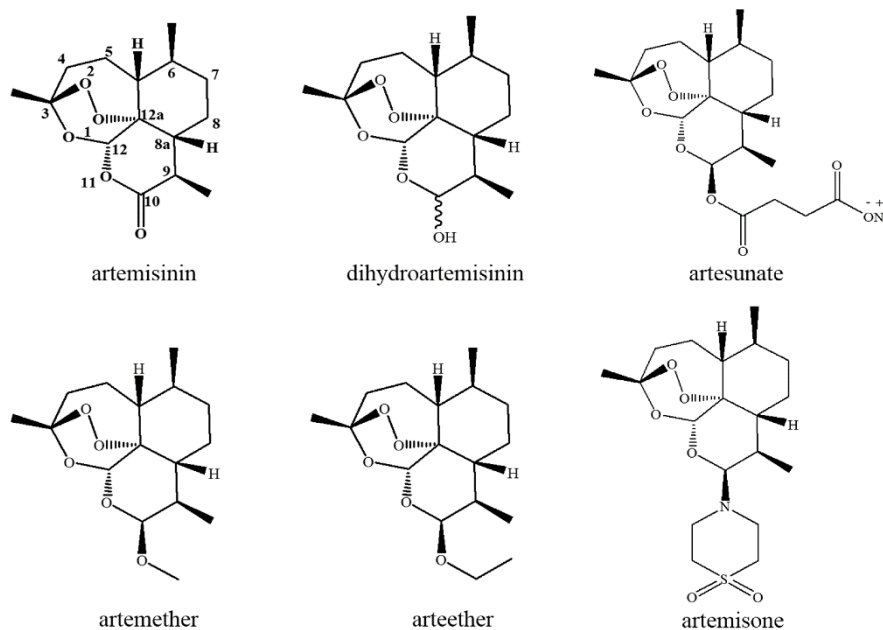


Figura 13: struttura dell'artemisinina e dei suoi derivati. Da ¹³

La struttura tetraciclica dell'artemisinina e dei suoi derivati è nota per svolgere un ruolo essenziale nel meccanismo d'azione antimalarico. L'interazione dell'endoperossido con l'eme o il Fe (II) libero porta alla formazione di radicali liberi intermedi. L'importanza di questi radicali liberi non è ancora del tutto chiarita, ma è opinione diffusa che provochino la morte del parassita¹⁴.

Le terapie combinate a base di artemisinina (ACTs – Artemisin-based Combination Therapy), vengono utilizzate come trattamento di prima linea per casi di malaria da *P. falciparum* multifarmaco-resistente. Tra le combinazioni maggiormente utilizzate sono presenti: artesunato-fosmidomicina, amodiachina-artesunato, cloroquina-artemisinina e artesunato-sulfadossina-pirimetamina. Attualmente le ACTs costituiscono la scelta predominante nelle zone endemiche¹⁵.

Tuttavia, studi recenti hanno evidenziato uno sviluppo di resistenza, anche nei confronti di questo approccio terapeutico, soprattutto nel Sud-Est Asiatico^{16,17}.

In risposta a questa crescente minaccia l'OMS ha avviato il *Global Plan for Artemisinin Resistance Containment* (GPARC). Questa iniziativa mira a frenare la diffusione di parassiti resistenti all'artemisinina, perseguendo l'eradicazione dei ceppi di *P. falciparum* resistenti in specifiche aree.

Questo intervento radicale ha un'importanza cruciale dal momento che ancora oggi non sono disponibili altri farmaci antimalarici che offrono la stessa efficacia e la tollerabilità dei derivati dell'artemisinina e non esistono alternative per l'immediato futuro.

2.7 Resistenza ai farmaci antimalarici

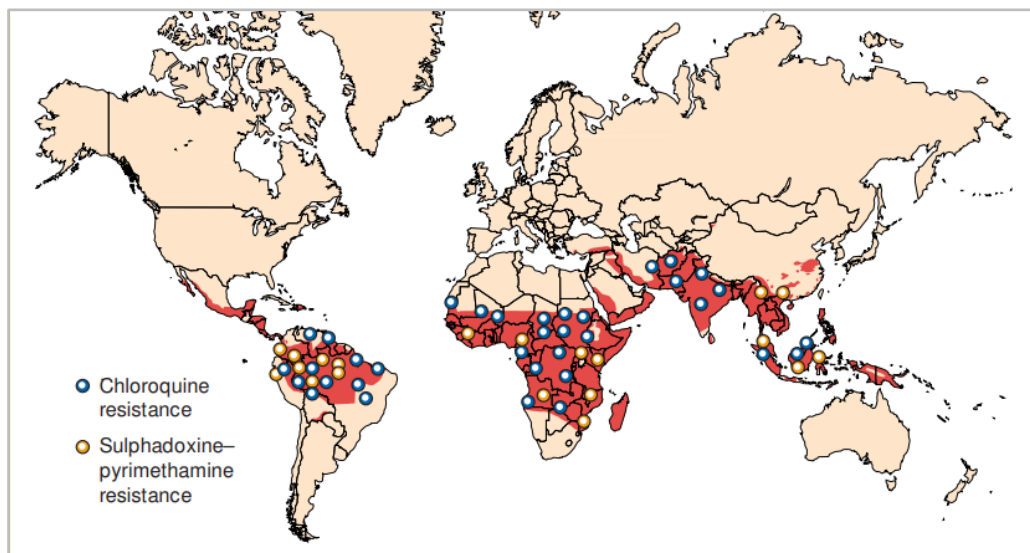


Figura 14: Stato globale di resistenza a cloroquina e sulfadossina/pirimetamina, i due farmaci antimalarici più ampiamente utilizzati. I dati provengono dall'OMS

La farmacoresistenza di *Plasmodium falciparum* è un importante ostacolo al controllo, al trattamento ed all'eliminazione della malaria. La recente comparsa nel Sud-Est Asiatico della resistenza all'artemisinina e ai suoi derivati ha sollevato preoccupazioni per la salute globale in quanto questa classe di farmaci risulta essere attualmente la migliore frontiera terapeutica.

La resistenza ai farmaci antimalarici è stata segnalata per la prima volta negli anni '50 tra la Cambogia e la Thailandia, e si è successivamente diffusa anche in altre aree endemiche.

La resistenza si è evoluta nel corso degli anni, contribuendo all'aumento della morbilità e della mortalità, soprattutto infantile.

La farmacoresistenza può svilupparsi in diversi momenti: durante la fase sessuata del parassita nell'ospite definitivo (la zanzara), durante lo stadio epatico pre-eritrocitario e durante lo stadio asessuato eritrocitario nell'ospite intermedio (l'uomo).

I meccanismi di resistenza ai farmaci antimalarici richiamano quelli degli antibiotici o chemioterapici, e sono causati da diversi fattori, tra cui la somministrazione inadeguata, la scarsa qualità e l'uso estensivo dei farmaci, che favoriscono la comparsa di ceppi farmaco-resistenti.

2.8 Vaccini

Il 6 ottobre 2021 l'OMS ha dato il via libera definitivo all'utilizzo del vaccino RTS,S/AS01 (*Mosquirix*, prodotto da GlaxoSmithKline) nei bambini di età inferiore ai 5 anni che vivono in aree a endemia moderata e elevata¹⁸.

Questo vaccino ricombinante, composto dalla proteina circumsporozoita del *Plasmodium falciparum* e combinato con antigeni di superficie dell'epatite B, viene somministrato per via intramuscolare in tre dosi intervallate da un mese l'una, con un richiamo a 18 mesi dall'ultima somministrazione.

Uno studio del 2016 nella regione africana del Sahel su oltre 6.800 bambini di età compresa tra 5 e 17 mesi ha portato a risultati molto positivi. L'uso di vaccini, chemioprolifassi e zanzariere ha infatti ridotto il rischio di ospedalizzazione e morte per malaria di oltre il 60%¹⁹.

Se implementata su larga scala, la vaccinazione potrebbe prevenire milioni di casi e salvare migliaia di vite all'anno. Tuttavia, l'uso delle zanzariere impregnate di repellente continua ad essere essenziale. Numerose campagne di sensibilizzazione, promosse da ONG, progetti locali e governi africani, mirano a distribuire zanzariere e a educare sull'utilizzo corretto, consolidando così una difesa multifattoriale contro la malaria.

3. I chinomi di *Homo sapiens* e del *Plasmodium*

Nel 2022 è stato pubblicato uno studio comparativo dei chinomi di *Homo sapiens*, *Plasmodium falciparum* e *Plasmodium vivax*²⁰.

È emerso che esistono differenze minime tra i chinomi delle due specie di *Plasmodium*, e che entrambi mostrano una divergenza significativa rispetto al chinoma umano. In particolare, il gruppo delle tirosin-chinasi (TK) è assente nei parassiti della malaria, mentre nel chinoma umano comprende circa 47 proteine. Inoltre, *Plasmodium* possiede anche famiglie come le protein chinasi calcio-dipendenti (CDPK) e le FIKK (che prendono il nome da un peptide condiviso Phe-Ile-Lys-Lys) che sono assenti invece nel chinoma umano. I chinomi dei parassiti includono numerose proteine appartenenti anche alle chinasi eucariotiche, ma in molti casi questi enzimi sono dissimili dai loro omologhi umani, il che suggerisce la possibilità di ottenere un'inibizione selettiva²¹.

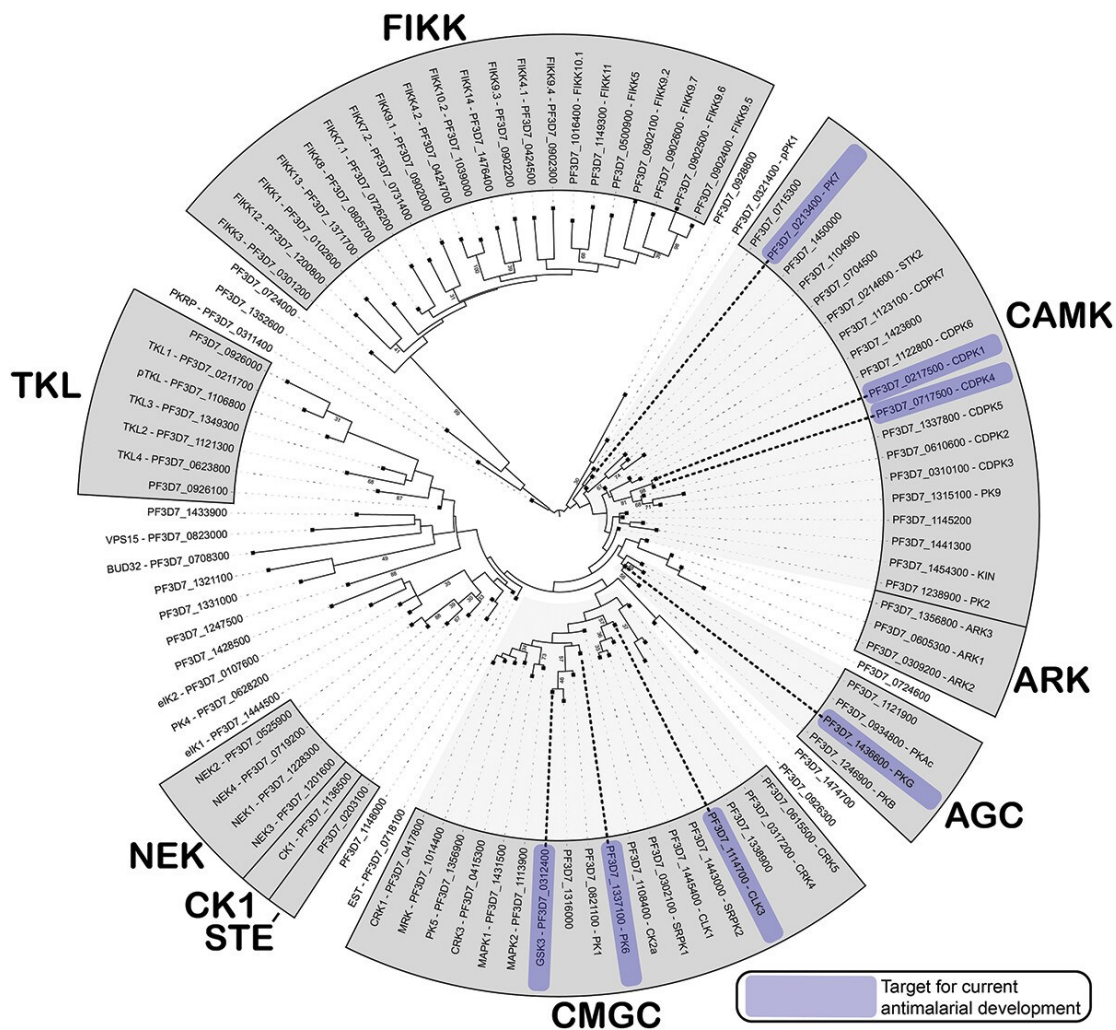


Figura 15: Albero filogenetico del chinoma di *Plasmodium falciparum*. In blu sono evidenziate le chinasi del parassita discusse come bersaglio di scoperte farmacologiche in corso.¹⁷

3.1 PfCK1

Una delle chinasi importanti per la sopravvivenza del *Plasmodium* è Casein chinasi 1 (PFCK1).

In particolare, PfCK1 sembra essere indispensabile per la proliferazione del parassita all'interno degli eritrociti. Le ricerche indicano come i parassiti intracellulari come Plasmodium possano alterare le vie di segnalazione delle cellule ospiti, esportando la loro CK1 nella cellula colpita dell'ospite^{22,23}.

Plasmodium falciparum codifica una singola isoforma di CK1²⁴, che viene espressa in tutte le fasi del ciclo vitale del parassita. Essa condivide un'identità di sequenza pari al 69% con la CK1 δ umana. Come altre CK1, PfCK1 interagisce con diversi ligandi, suggerendo il suo potenziale coinvolgimento nella regolazione di varie vie, tra cui quelle legate alla trascrizione, alla traduzione e alla traslocazione post-traduzionale delle proteine in specifiche zone della cellula.

Non sono disponibili in letteratura molte informazioni su questa proteina, però è noto che è un enzima espresso in tutti gli eucarioti.

3.2 Casein chinasi 1

La Proteina Casein chinasi 1 (CK1) si può trovare in tutti i tessuti e compartimenti cellulari e all'interno delle cellule si localizza principalmente nel nucleo e nel citoplasma, ma anche ancorata stabilmente alla membrana cellulare o al citoscheletro²⁵.

Si tratta di una proteina appartenente alla famiglia delle Ser-Thr Chinasi, enzimi in grado di catalizzare reazioni di fosforilazione che coniugano covalentemente il gruppo fosfato in posizione γ dell'ATP a un idrossile di una Serina o di una Treonina.

Nei mammiferi sono presenti sette diverse isoforme della Casein chinasi 1: α , β , γ_1 , γ_2 , γ_3 , δ e ϵ .

Le varie isoforme si differenziano tra loro per la diversa lunghezza degli estremi C-terminale e N-terminale, ma presentano un'identità di sequenza molto elevata.

In particolare, le isoforme δ e ϵ presentano una similarità di sequenza del 97,5% nel dominio chinamico, per questo motivo talvolta vengono accorpate in un'unica sottofamiglia δ/ϵ ²⁶.

3.3CK1 δ umana

La Casein chinasi 1 δ (CK1 δ) è un enzima presente nel chinoma di *Homo sapiens*. Omologhi di CK1 δ sono stati isolati da diversi organismi come lieviti, piante, alghe o protozoi.

È costituita da una porzione N-terminale (composta principalmente da filamenti organizzati in struttura secondaria β -sheet) e da una C-terminale (costituita prevalentemente da strutture secondarie ad α -elica e *loop*), collegate tra loro da una regione cerniera (*hinge region*) altamente flessibile.

Quest'ultima accoglie nel suo intorno il sito catalitico della chinasi, dove si legano i diversi substrati e l'ATP per le reazioni di fosforilazione in cui è coinvolta la proteina. La sua flessibilità porta ad una certa variabilità conformazionale del sito attivo, causata dalla convergenza o divergenza delle regioni N-terminale e C-terminale.

In Figura 16 è raffigurata come esempio una delle strutture tridimensionali di CK1 δ umana risolte tramite cristallografia a raggi X depositate nel *Protein Data Bank* (PDB), un archivio di strutture tridimensionali sperimentali di macromolecole (PDB_{ID}: 5IH4).

Si possono vedere le regioni caratterizzate da strutture ad α -elica colorate in rosso, i β -sheet in giallo, i *turn* (regioni di raccordo) in blu, e i *loop* (regioni flessibili che non hanno una struttura terziaria definita) in bianco.

Intorno al *backbone* della proteina, in grigio, è stata costruita la superficie molecolare di Connolly, che mostra la superficie di van der Waals della biomolecola²⁷.

L'immagine è stata generata con l'utilizzo del *software* MOE (*Molecular Operating Environment*) utilizzando la struttura proveniente dal *Protein Data Bank*.

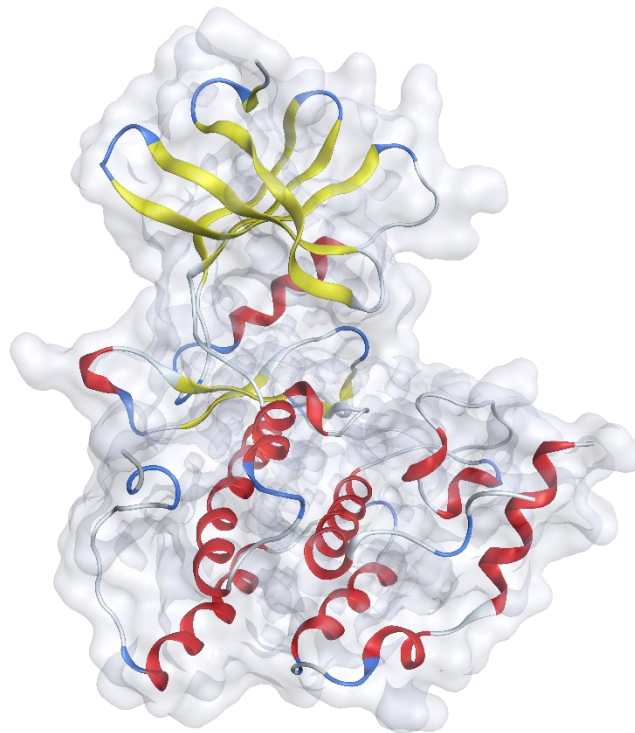


Figura 16: Figura 16. Struttura tridimensionale di CK1 delta umana in forma apo, da PDB (5IH4)

Funzioni di CK1δ

CK1δ svolge un ruolo significativo in diversi pathway cellulari, come la progressione del ciclo cellulare, la stabilità genomica, la mitosi e la meiosi²⁸⁻³¹.

A livello del ciclo cellulare, la sua importanza risiede nella regolazione della polimerizzazione dei microtubuli, alla stabilità dell'apparato del fuso e dei centrosomi durante la mitosi mediante la fosforilazione diretta delle α -, β - e γ -tubuline³²⁻³⁴.

Di conseguenza la deregolazione di CK1δ contribuisce alla perdita del fine controllo del ciclo cellulare, portando alla tumorigenesi³⁵.

Infine, CK1δ ha anche un ruolo nella modulazione dei ritmi circadiani. Essi sono controllati da un ciclo di feedback trascrizionale negativo, nel quale le proteine PERIOD (PER) svolgono un ruolo cruciale.

È stato dimostrato che CK1δ è in grado di regolare l'abbondanza temporale e l'attività delle proteine PER attraverso meccanismi di degradazione mediata dalla fosforilazione³⁶.

Rilevanza clinica

Dal punto di vista della carcinogenesi CK1δ, sembra essere coinvolta in diversi modi; in alcune forme tumorali, l'mRNA che codifica per CK1δ risulta sovraespresso.

In parallelo, alcuni studi hanno evidenziato livelli ridotti di espressione dell'mRNA di CK1δ in altri contesti tumorali. Queste variazioni di espressione suggeriscono un coinvolgimento complesso della proteina CK1δ in diverse neoplasie³⁵.

Inoltre, l'espressione anomala di CK1δ nel tessuto cerebrale è stata associata a diverse malattie neurodegenerative: tra tutte, la malattia di Alzheimer, il morbo di Parkinson, la sclerosi laterale amiotrofica (SLA) e la paralisi sopranucleare progressiva e la sindrome familiare della fase avanzata del sonno (conosciuta anche come Advanced Sleep Phase Syndrome o FASPS)³⁷.

CK1δ infine può influenzare la disfunzione metabolica, in particolare nei soggetti obesi, migliorando la tolleranza al glucosio, riducendo l'espressione genica della gluconeogenesi e la secrezione di glucosio e aumentando l'assorbimento di glucosio al basale e stimolato dall'insulina. Inoltre, CK1δ svolge un ruolo nella modulazione della formazione della forma biologicamente attiva a più alto peso molecolare dell'adiponectina. Di conseguenza a questo livello la perdita della sua regolazione è molto probabilmente coinvolta nello sviluppo di disordini metabolici^{38,39}.

4. Scopo della tesi

L'obiettivo principale di questa tesi è affrontare la crescente problematica della resistenza di *Plasmodium falciparum* alle terapie antimalariche maggiormente utilizzate tramite un approccio di Computer-Aided Drug Design (CADD).

La necessità di trovare nuovi approcci terapeutici può beneficiare dell'identificazione di nuovi target molecolari che possano fungere da potenziali bersagli per lo sviluppo di farmaci. Campagne di screening focalizzate su bersagli alternativi consentono infatti di esplorare lo spazio chimico e di identificare nuovi composti potenzialmente in grado di modulare *pathway* molecolari che eludano i meccanismi di resistenza già in essere.

Nel contesto dello screening di grandi librerie di composti, l'approccio computazionale si distingue per il suo vantaggio in termini di tempi e costi rispetto a *high-throughput screening* sperimentali. Tecniche come il *docking* e la dinamica molecolare fungono da supporto a test *in vitro*: infatti queste tecniche computazionali consentono di valutare *in silico* le capacità interattive di grandi librerie di composti con il target molecolare di interesse, prioritizzando un sottogruppo di molecole promettenti da sottoporre in modo focalizzato ad uno screening *in vitro*, con un vantaggio economico.

In questo contesto, si è scelto di focalizzare l'attenzione sulla proteina PfCK1 come possibile *target* di interesse a fini antimalarici, infatti CK1 è coinvolta in vari processi vitali del parassita *Plasmodium falciparum*, e potrebbe quindi essere un *target* di interesse a scopo terapeutico⁴⁰.

Questo lavoro di tesi ha l'obiettivo di sottoporre il target PfCK1 ad una campagna di *virtual screening* utilizzando una libreria di composti naturali e loro analoghi contenente più di centotredicimila molecole, mirando ad identificare potenziali inibitori dell'enzima la cui attività possa essere successivamente verificata con saggi sperimentali.

Questo progetto nasce dalla collaborazione con il gruppo di ricerca del Professor Giampiero Spalluto e della Professoressa Stephanie Federico dell'Università di Trieste che si occuperà di testare l'attività *in vitro* dei composti identificati in questo lavoro su CK1 tramite saggi di inibizione enzimatica, e dalla collaborazione con il gruppo di ricerca della Professoressa Maria Laura Idda dell'Università di Sassari, che si occuperà di testare i composti attraverso saggi di vitalità di *Plasmodium falciparum*.

Parte seconda:

MATERIALI E METODI

5. Hardware e Software

Il lavoro di tesi è stato svolto presso il laboratorio di modellistica molecolare MMS (*Molecular Modeling Section*) del Dipartimento di Scienze del Farmaco dell'Università degli Studi di Padova, supervisionato dal Professor Stefano Moro.

L'infrastruttura informatica del laboratorio, rappresentata nell'immagine sottostante, comprende:

- 10 *user workstation*
- 4 nodi di calcolo CPU (per un totale di 384 *core*)
- 9 nodi di calcolo GPU (30 GPU NVIDIA, serie *Titan*, GTX e RTX)
- 1 *web server*
- 3 nodi di *storage* da 250 TB di capacità

L'architettura indicata utilizza il sistema operativo Linux Ubuntu aggiornato alla versione 16.04 LTS e alla versione 20.04 LTS.

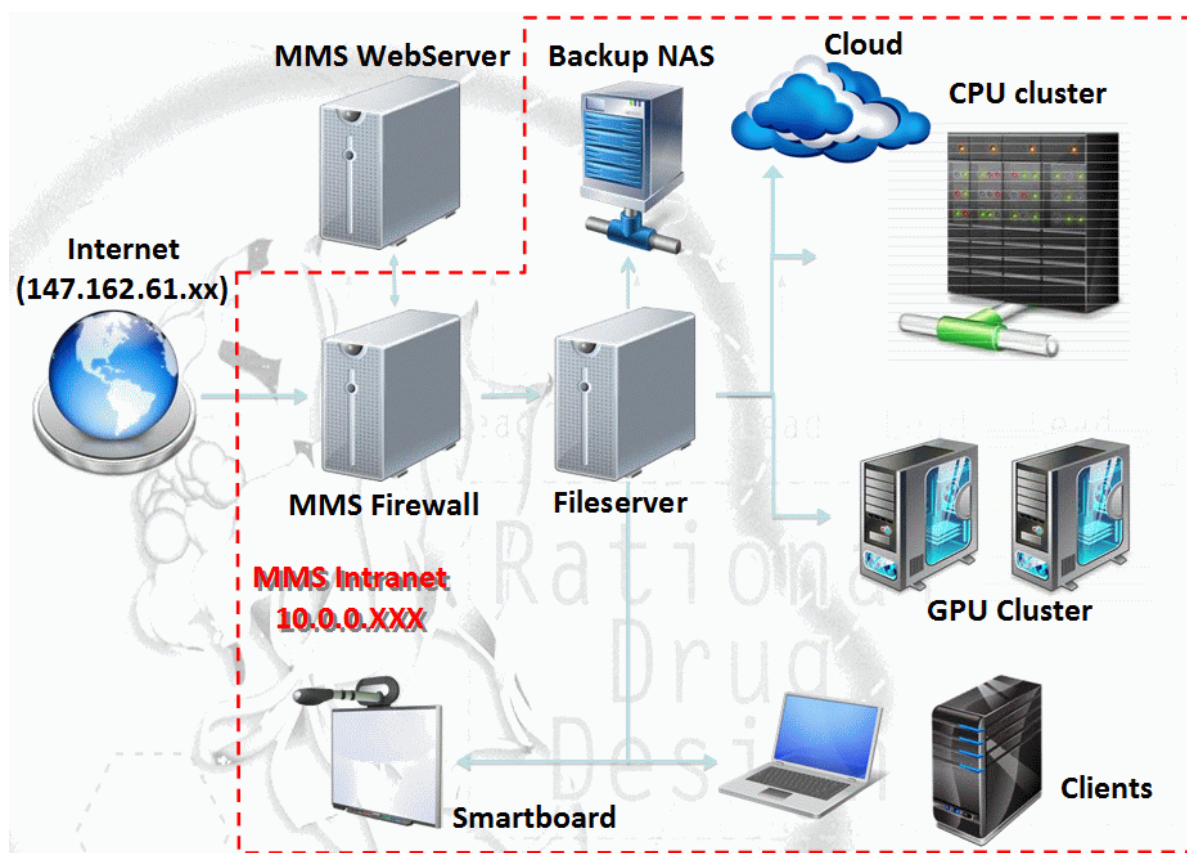


Figura 17: organizzazione informatica interna al laboratorio MMS

Per lo svolgimento dell'attività di tesi, oltre a una serie di *script* SVL (*Scientific Vector Language*), BASH (*Bourne Agains Shell*), e Python, sono stati utilizzati i software in Tabella 2.

SOFTWARE	UTILIZZO
MOE 2022 (Molecular Operating Environment)²⁷	Preparazione strutture proteiche e ligandi
MOE Structure-Based Homology Modeling	Generazione di un modello di omologia structure-based
PLANTS v1.2⁴¹	Docking molecolare
AmberTools14⁴²	Preparazione sistema per la dinamica molecolare
ACEMD 3⁴³	Motore di dinamica molecolare
VMD (Visual Molecular Dynamics)⁴⁴	Gestione e visualizzazione delle traiettorie di dinamica molecolare

Tabella 2: Software utilizzati nello svolgimento del lavoro di tesi.

6. Virtual Screening

Nel processo di ricerca e sviluppo di un farmaco, soprattutto nelle fasi iniziali, l'operazione di *screening* assume un ruolo di primaria importanza. Lo scopo principale è di esaminare un'ampia varietà di piccole molecole organiche sullo stesso bersaglio, al fine di individuare uno o più *hit compounds* che manifestino un'interazione promettente con il *target* di interesse.

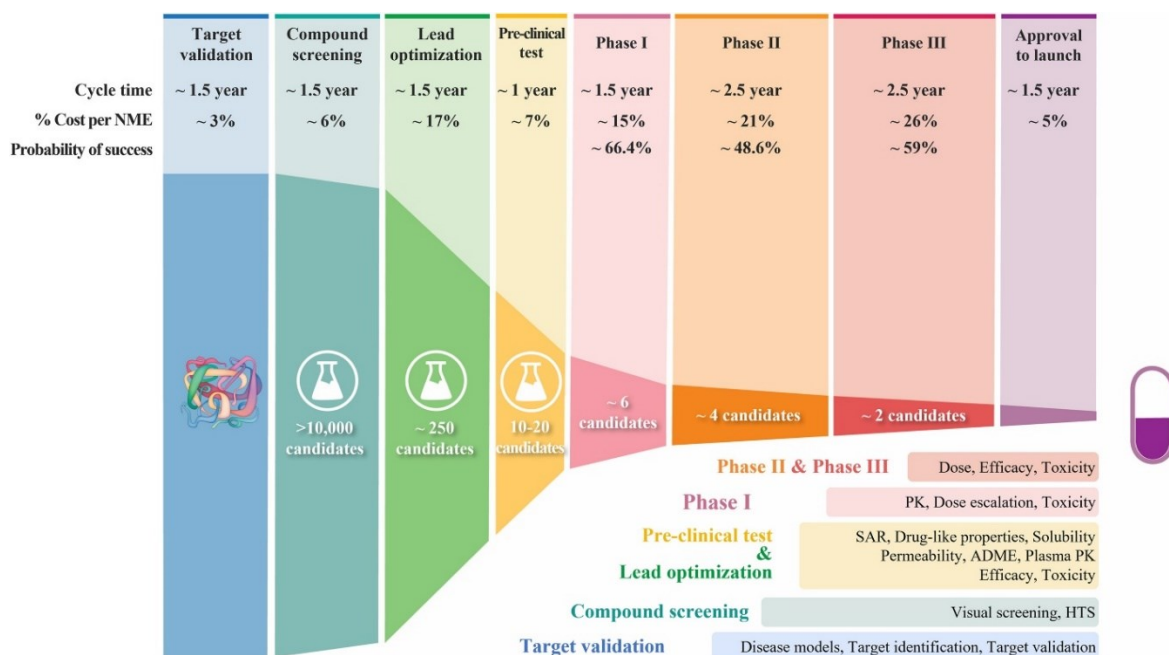


Figura 18: timeline del percorso di scoperta e sviluppo di un farmaco.⁴⁵

In questo contesto, una delle tecniche più utilizzate è quella dei saggi HTS (*High Throughput Screening*), ottimizzati per testare un grande numero di molecole in tempi relativamente ridotti. Nonostante tali test costituiscano un approccio indubbiamente valido, il loro principale svantaggio deriva dalla vastità dello spazio chimico disponibile. Infatti, la quantità di composti esaminabili tramite questi saggi rappresenta solo una frazione minima di tutti i composti disponibili. Allo stesso tempo, un altro significativo limite è legato alla disponibilità sul mercato, ai costi per l'acquisto delle molecole da testare e per lo svolgimento dei test, ma anche al tempo necessario al loro svolgimento che diventerebbe insostenibilmente grande. Infine, un ultimo inconveniente è rappresentato dall'enorme spreco di materiale che ne risulterebbe, al netto di una probabilità di successo molto bassa.

Per ovviare a questi problemi, si può anteporre alla fase di *screening* sperimentale una fase precoce di test *in silico*. L'idea di fondo è di fare una pre-selezione al fine di trovare, all'interno di un *pool* molto ampio (nell'ordine delle centinaia di migliaia, anche milioni) di possibili candidati, quella frazione che presenta caratteristiche tali da far presupporre una possibile attività nei confronti del *target* di interesse. Questo permette quindi di esplorare uno spazio chimico molto più ampio, con un

risparmio notevole sia in termini economici che temporali. È importante però ricordare che le tecniche *in silico* costituiscono un processo di preselezione, e non sono sostitutive dei saggi sperimentali. Il valore del *virtual screening* si manifesta pienamente quando viene supportato e integrato con successivi test *in vitro*.

Negli ultimi anni, grazie alla vasta disponibilità di strutture sperimentalmente risolte di biomolecole di interesse terapeutico e ai progressi negli algoritmi per la predizione della struttura terziaria delle proteine, la maggior parte dei *virtual screening* si basa sul *docking*. Questa tecnica consente di predire o quantomeno ipotizzare quale sia la migliore modalità di interazione dei ligandi nel sito attivo di un *target* definito. L'ampio utilizzo di questo approccio è dovuto al fatto che, al giorno d'oggi, il *docking* rappresenta il miglior compromesso tra la precisione nella previsione dell'interazione farmaco-recettore e la rapidità di esecuzione del calcolo.

Non è possibile definire un protocollo di *docking* valido per tutti i casi di studio; esso viene infatti costruito su misura per il *target* e i ligandi di nostro interesse.

Per valutare l'attendibilità del protocollo definito si possono prendere più strade, in base ai dati che si hanno a disposizione. La più semplice è quella che prevede di calibrare il protocollo in modo che esso sia in grado di selezionare, da una libreria di composti fornita, quelli che in base a dati in letteratura su test sperimentali abbiano un'efficacia di interazione verificata nei confronti di quel ligando. Una seconda via prevede che il protocollo venga calibrato effettuando un test con delle molecole sperimentalmente attive di cui esiste la struttura cristallografica con quel *target*. In questo caso un risultato può definirsi ottimale se le *binding mode* (modalità di interazione) risultanti dal *docking* riproducono quelle determinate sperimentalmente.

La similarità tra due *binding mode*, quella sperimentalmente risolta e quella predetta, può essere determinata in base al valore di RMSD (*Root Mean Square Deviation*). Esso quantifica la deviazione quadratica media delle posizioni degli atomi nei due insiemi di coordinate, e fornisce un'indicazione della discrepanza della localizzazione spaziale tra le specie confrontate. Valori bassi di RMSD indicano una buona convergenza tra le coordinate delle due strutture, mentre valori elevati indicano una maggiore divergenza. Conseguentemente, più piccolo sarà il valore di RMSD, più accurato sarà il protocollo nel fornire valori vicini al dato sperimentale, e quindi tanto più sarà adatto a selezionare (all'interno della libreria da testare) i composti più affini al *target* d'interesse.

Un ulteriore aspetto da considerare è correlato alla potenza dell'*hardware* a disposizione. La quantità di composti che può essere processata tramite un *virtual screening* è proporzionale alla potenza di calcolo disponibile. In presenza di risorse computazionali limitate, il numero di composti esplorabili tramite questa tecnica è limitato.

Per superare questo problema, un approccio possibile consiste nell'effettuare una *clusterizzazione* del *database* di partenza. Tale procedura prevede il raggruppamento di composti aventi uno *scaffold* chimico in comune, con la selezione di un rappresentante per ciascun gruppo. Questa strategia comporta una significativa riduzione delle dimensioni della libreria iniziale, garantendo al contempo un'efficiente esplorazione dello spazio chimico.

6.1 Homology modeling⁴⁶

L'assenza di una struttura sperimentalmente risolta limita fortemente un approccio *structure-based*. Per ovviare a questa mancanza, si può generare un modello della proteina di interesse per omologia, partendo dalla sua sequenza primaria. Tale approccio si fonda su due assunzioni fondamentali: la prima ipotizza che la struttura tridimensionale (struttura terziaria) della proteina sia determinata dalla sua sequenza amminoacidica, mentre la seconda si basa sul presupposto che sequenze simili assumano una disposizione conservata nello spazio tridimensionale.

Il *software* utilizzato per la generazione di questo modello è il *tool* di MOE *Structure-based Homology model*.

La generazione del modello si articola in diversi step:

1. **Identificazione di un *template*.** Il primo passo fondamentale nella creazione di un modello di omologia consiste nella selezione di un *template*. A partire dalla sequenza amminoacidica della proteina di interesse, si utilizzano alcuni web-server come Protein BLAST al fine di identificare quali proteine, all'interno di un database definito (es. PDB), presentano la massima percentuale di identità di sequenza in comune con quella di interesse. Per un risultato attendibile, è bene prendere in considerazione strutture con una significativa identità di sequenza, che non dovrebbe mai scendere sotto il 25%.
2. **Allineamento di sequenza.** Dopo aver definito il *template*, la sequenza *target* deve essere allineata con quella del *template*. Questo step è fondamentale, pertanto un allineamento accurato riveste un'importanza cruciale per la validità del modello.
3. **Creazione del modello.** Le tecniche utilizzate per la generazione di modelli tradizionali si suddividono in diverse categorie:
 - a. **Rigid-body modeling:** questa tecnica scompone la proteina d'interesse in un insieme di parti, ciascuna rappresentante una porzione conservata. Le varie parti vengono poi assemblate, ottenendo la struttura finale;
 - b. **Segment matching:** in questa tecnica, la sequenza amminoacidica della proteina viene suddivisa in segmenti, e ogni segmento viene allineato separatamente con la

struttura di riferimento omologa corrispondente. Il modello completo viene poi assemblato fondendo i vari segmenti allineati;

- c. ***Spatial restraint***: qui vengono applicate delle restrizioni spaziali durante il processo di modellazione, basate sulla struttura nota della proteina template. Le restrizioni spaziali guidano la creazione del modello garantendo che le distanze e gli angoli rispettino la conformazione 3D desiderata;
- d. ***Artificial evolution***: La tecnica di evoluzione artificiale utilizza il metodo di assemblaggio a corpo rigido (*rigid-body modeling*) e introduce mutazioni, fino a quando la sequenza del *template* coincide con la sequenza *target*.

4. **Modellamento dei *loop***. I *loop* hanno un ruolo cruciale nell'organizzazione 3D delle proteine, tuttavia, avendo una elevata flessibilità, hanno una minore conservazione strutturale. La modellazione dei *loop* è dunque più complessa a causa della maggiore variabilità rispetto agli elementi di struttura secondaria come α -eliche o β -sheet. Per il modellamento dei *loop* si utilizzano due metodi principali: la ricerca in banche dati e la ricerca conformazionale. La ricerca su *database* identifica i segmenti critici da strutture note, mentre la ricerca conformazionale utilizza una *scoring function*.
5. ***Side chain modeling***. Il modellamento delle catene laterali viene realizzato facendo coincidere le coordinate del *backbone* con quelle del *template*. L'accuratezza della previsione della posizione delle catene laterali è maggiore ad alti livelli di identità di sequenza. Il processo si basa sulla selezione di rotameri, conformazioni a bassa energia delle catene laterali, e sulla base di funzioni energetiche. L'accuratezza è generalmente elevata per i residui idrofobici collocati nel *core*, e inferiore per i residui esposti alla superficie.
6. **Ottimizzazione del modello**. L'ottimizzazione del modello di solito viene effettuata a partire da una minimizzazione dell'energia, per la quale vengono utilizzati i campi di forza della meccanica molecolare.
7. **Validazione del modello**. L'accuratezza del modello costruito è determinante per la sua applicazione in vari settori, per questo motivo è necessario che venga validato. Viene effettuata un'analisi della stereochimica, valutando parametri quali la lunghezza di legame, l'angolo di torsione e l'angolo di rotazione. Il diagramma di Ramachandran è un potente determinante della qualità della struttura proteica, infatti i residui con problemi di stereochimica non rientrano nelle regioni accettabili del diagramma di Ramachandran.

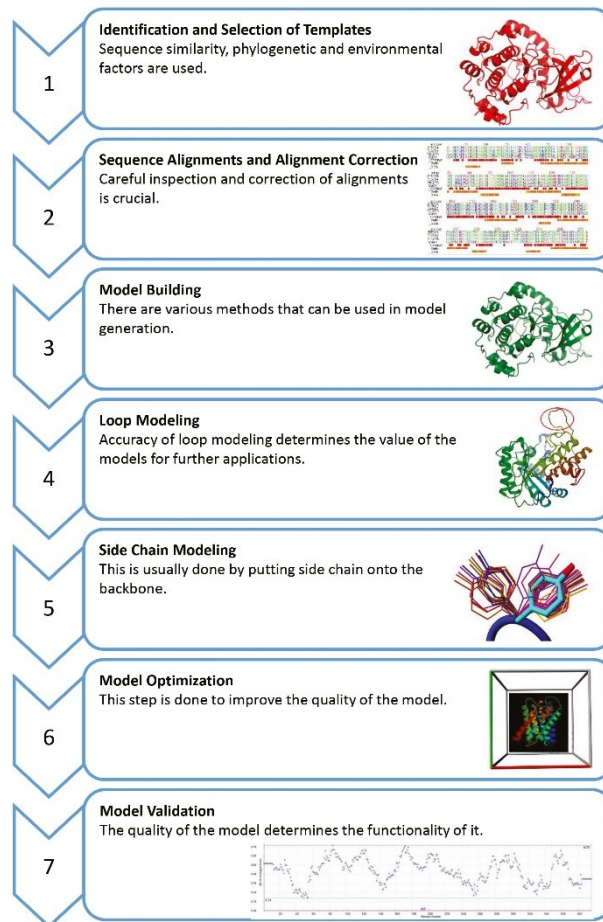


Figura 19: schema della serie di processi da svolgere al fine di generare un modello di omologia ⁴⁶

6.2 Operazioni preliminari al *virtual screening*

Prima di iniziare il *virtual screening*, è necessario svolgere alcune operazioni preliminari, essenziali per ottenere risultati affidabili. Esse consistono nella preparazione del *target* e dei ligandi che saranno utilizzati come *input* per il *docking*.

Per la selezione dei ligandi, è possibile adottare due strategie. La prima prevede di recuperare librerie di ligandi da *database* specializzati, come ad esempio ZINC. Alternativamente, è possibile creare una libreria di composti personalizzata e prepararla mediante l'utilizzo di specifici *software*.

Preparazione dei ligandi

Per la preparazione della struttura dei ligandi, si applica un protocollo diviso in vari punti:

- 1. Generazione dello stato di protonazione e tautomerico.** Le molecole organiche possono esplorare equilibri tautomerici, e ciò può impattare radicalmente il profilo di interazione di una molecola con il *target*. Lo stesso vale per gli stati di protonazione dei gruppi funzionali, che possono essere influenzati dal pH dando equilibri acido-base. Nello step preliminare di preparazione di una libreria, vengono perciò generati gli stati protomerici e tautomerici più probabili nel solvente e pH dati.
- 2. Generazione delle coordinate 3D.** Generalmente le librerie virtuali di composti consistono in *database* di molecole espresse in forma monodimensionale, ad esempio tramite stringhe SMILES, o bidimensionale. Prima di sottoporre i composti a simulazioni di *docking* molecolare, è necessario generare una struttura tridimensionale plausibile per ogni molecola, con angoli o lunghezze di legame verosimili. Bisogna inoltre assicurarsi che la tipizzazione di ogni atomo sia compatibile con il *set* di *atom type* riconosciuto dal programma di *docking* scelto. Ciò previene errori nell'interpretazione del profilo di interazione e nella determinazione della struttura del ligando.
- 3. Verifica delle cariche formali e assegnazione delle cariche parziali.** Le cariche formali dipendono dalla struttura del composto, dai gruppi funzionali presenti e dalle condizioni del contorno (es. pH), mentre le cariche parziali descrivono la polarizzazione dei legami e la distribuzione degli elettroni. Possono essere calcolate con metodi quantomeccanici, semi-empirici o basati su campi di forze.

Per questo studio i ligandi sono stati preparati con l'ausilio del software MOE, usando il suo specifico *tool Wash*. Gli atomi di idrogeno mancanti sono stati aggiunti con il *tool Protonate 3D*, impostando il valore di pH a 7.

Preparazione della proteina *target*

Per quanto riguarda la preparazione della struttura della proteina, è essenziale focalizzare l'attenzione sui seguenti aspetti:

- 1. Valutazione della qualità della struttura.** Per quanto concerne le strutture cristallografiche, la qualità complessiva della struttura può essere valutata mediante i parametri dell'*R-factor* e della risoluzione. Un valore ridotto dell'*R-factor* e un aumento della risoluzione indicano un miglioramento della qualità generale. Per un'analisi più dettagliata delle singole porzioni proteiche, si considera il parametro del *B-factor*: valori più bassi indicano una minore incertezza nella posizione atomica. Un altro parametro che viene valutato è l'*occupancy*, variabile tra 0 e 1, che rappresenta la probabilità che un gruppo funzionale si trovi effettivamente nella posizione indicata nel cristallo, risultando cruciale per i residui amminoacidici con equilibri conformazionali nelle catene laterali degli amminoacidi. Altrettanto cruciale è assicurarsi che le regioni di interesse, come il sito attivo, siano risolte in modo accurato. Nel caso in cui, come in questo studio, la struttura sperimentale del ligando d'interesse non sia disponibile, si ricorre a un modello di omologia come visto in precedenza, a condizione di avere a disposizione una struttura sperimentale di una proteina ad elevata percentuale di identità di sequenza da utilizzare come *template*, che andrà comunque preliminarmente preparata come riportato in seguito.
- 2. Preparazione e pulizia della struttura.** Tutte le molecole estranee alla proteina e ai ligandi, presenti nel cristallo a causa di fattori legati al protocollo sperimentale di cristallizzazione, come ioni o molecole di solvente/tampone, devono essere eliminate. Alcune sezioni della struttura proteica potrebbero risultare mancanti a causa di zone di densità elettronica difficili da risolvere, o potrebbero essere presenti conformazioni alternative dei residui di Asn, Gln e His. Questi problemi possono essere risolti attraverso l'uso di *tool* dedicati, come *Structure Preparation* di MOE.
- 3. Aggiunta degli idrogeni e minimizzazione della loro energia.** Le strutture cristallografiche, per limiti intrinseci della tecnica, non includono gli atomi di idrogeno ed è necessario aggiungerli tramite appositi *tool* informatici. Poiché la posizione degli atomi di idrogeno viene introdotta in modo arbitrario, è essenziale minimizzare le loro coordinate dal punto di vista energetico. Per questo studio sono stati impiegati i software di MOE *Protonate3D* ed *Energy Minimize*, utilizzando il campo di forze AMBER 10:EHT.

6.3 Meccanica molecolare

Nell'analisi dei modelli molecolari a livello computazionale, si adoperano due approcci principali per calcolare l'energia del sistema: i metodi quantomeccanici e quelli basati sulla meccanica molecolare.

Nei metodi quantomeccanici viene studiato il comportamento degli elettroni, tenendo conto delle particelle subatomiche. Questo comporta calcoli teorici complessi, che richiedono una elevata capacità di calcolo e tempi di esecuzione prolungati, non applicabili a macromolecole.

Nei metodi basati sulla meccanica molecolare, si adotta la meccanica classica (meccanica Newtoniana), in cui gli atomi vengono semplificati ad entità puntiformi con massa e carica specifiche. Ciascun atomo presenta una posizione spaziale definita, codificata in coordinate tridimensionali, e le interazioni tra entità vicine contribuiscono a formare il potenziale energetico del sistema.

L'energia potenziale di un sistema viene descritta da un insieme di equazioni parametriche chiamato campo di forze (o *force field*). Ne esistono vari tipi, ciascuno dei quali è ottimizzato per descrivere una specifica tipologia di sistema. Tra i vari *force field* utilizzati si menzionano CHARMM, utile a descrivere sistemi di membrana; AMBER, ideale quando vengono descritte proteine solubili; e MMFF che viene usato per descrivere piccole molecole organiche.

Ogni campo di forza considera due componenti nel calcolo dell'energia potenziale complessiva di un sistema: il contributo di legame e quello di non legame.

$$E = E_{bonded} + E_{non-bonded}$$

$$E_{bonded} = E_{stretching} + E_{bending} + E_{dihedral}$$

$$E_{non-bonded} = E_{electrostatic} + E_{vdW}$$

Equazione 1. Equazione di potenziale.

Contributi di legame (E_{bonded})

- 1. Bond Stretching (stiramento):** è definito come l'energia del legame tra due atomi, ed è funzione della variazione della distanza tra i centri di massa delle due entità puntiformi coinvolte nel legame covalente rispetto al valore all'equilibrio. Il potenziale di *stretching* viene descritto tramite la legge di Hooke (equazione della forza elastica). K_b indica la costante di forza, mentre $(r - r_0)$ indica la variazione di lunghezza del legame rispetto alla lunghezza all'equilibrio.

$$E_{stretching} = \sum K_b (r - r_0)^2$$

Equazione 2. Potenziale di stretching.

- 2. Bond Bending (piegamento):** è definito come l'energia associata alla variazione dell'angolo di legame tra due legami adiacenti. Anche in questo caso, il potenziale di *bending* viene definito dalla legge di Hooke, dove K_θ indica la costante di deformazione angolare, mentre $(\theta - \theta_0)$ indica la variazione di ampiezza dell'angolo rispetto all'equilibrio.

$$E_{bending} = \sum K_\theta (\theta - \theta_0)^2$$

Equazione 3. Potenziale di bending.

- 3. Dihedral torsion (torsione dei diedri):** considerando un sistema composto da 4 atomi interconnessi da 3 legami, di cui il centrale è libero di ruotare, l'angolo diedro è rappresentato dall'angolo che si crea tra i due legami esterni. Il potenziale torsionale viene calcolato tramite la seguente equazione trigonometrica, dove V_n indica la costante di forza torsionale, n rappresenta la molteplicità, ϕ corrisponde all'angolo diedro e γ determina il fattore di fase.

$$E_{dihedrals} = \sum \frac{V_n}{2} [1 + \cos(n\phi - \gamma)]$$

Equazione 4. Potenziale dei diedri.

Contributi di non legame

- 1. Potenziale elettrostatico:** deriva dalle interazioni tra le cariche parziali degli atomi che compongono il sistema, e viene calcolato tramite la legge di Coulomb. In questa equazione q_i e q_j indicano le cariche parziali dei due atomi, r_{ij} indica la loro distanza, mentre ϵ_0 rappresenta la costante dielettrica nel mezzo.

$$E_{electrostatics} = \sum_i \sum_j \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}}$$

Equazione 5. Potenziale elettrostatico.

- 2. Potenziale di van der Waals:** viene calcolato tramite la funzione di Lennard-Jones, che tiene conto delle forze *short-range* di natura repulsiva (repulsioni steriche) e *long-range* di natura attrattiva (legata alla formazione di dipoli istantanei).

$$E_{vdW} = \sum_i \sum_j \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6}$$

Equazione 6. Potenziale di van der Waals.

6.4 Docking

Il docking molecolare costituisce un approccio computazionale finalizzato all'analisi e alla previsione della modalità di legame di un ligando all'interno di una binding pocket di un target molecolare, tramite la simulazione della conformazione che il ligando può assumere e delle interazioni che possono crearsi tra il ligando e il target molecolare.

Per effettuare un *docking* è necessario innanzitutto disporre della struttura tridimensionale della proteina *target*, la quale viene ottenuta mediante tecniche quali cristallografia, Risonanza Magnetica Nucleare (NMR) o Microscopia Crioelettronica (Cryo-EM). In situazioni in cui la struttura non è risolta sperimentalmente, è possibile ricorrere alla costruzione di modelli di omologia o, alternativamente, alla predizione de novo, consentendo così di esplorare bersagli proteici di cui si conosce solamente la sequenza amminoacidica.

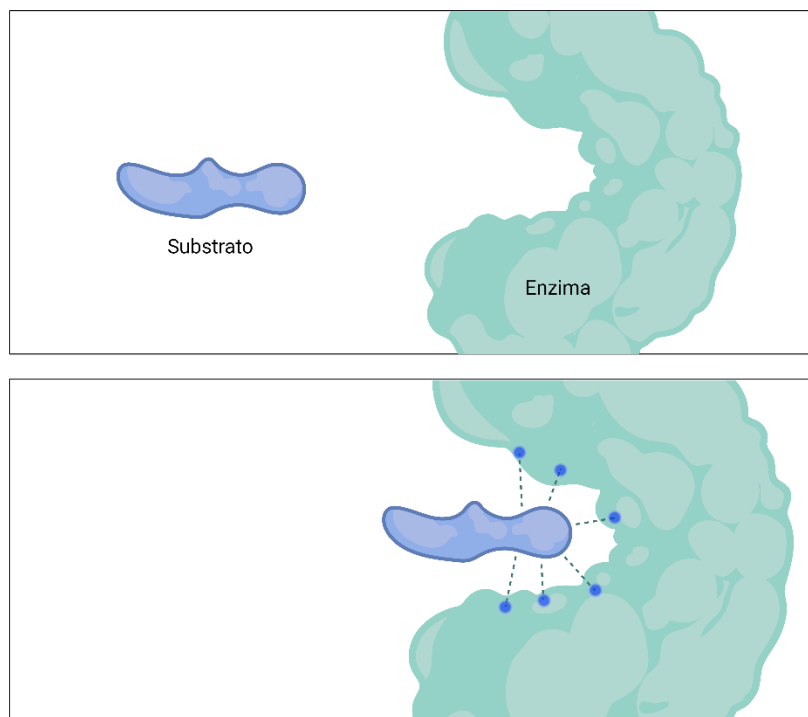


Figura 20: esemplificazione grafica del docking molecolare proteina-ligando.

Data una libreria di composti e una proteina *target* di cui viene definito il sito attivo, il *docking* è una tecnica computazionale che permette di generare un numero definito di *poses* dei composti di interesse all'interno del *binding site*. All'interno della struttura di interesse viene quindi definito un volume da esplorare (solitamente tramite centro e raggio di una sfera o alternativamente un cubo) che contenga il sito attivo. Ogni *pose* di docking rappresenta una possibile modalità di interazione ligando-proteina nel panorama delle differenti conformazioni e orientazioni del ligando. Ogni protocollo di docking è formato da due componenti: un algoritmo di ricerca conformazionale

(responsabile dell'esplorazione dei possibili gradi di libertà del ligando all'interno della *search box*) e una *scoring function* (responsabile della valutazione della qualità del riconoscimento ligando-recettore). L'algoritmo di ricerca tiene conto dei gradi di libertà traslazionali, rotazionali e conformazionali del ligando e, in funzione del tipo di docking che si vuole effettuare questi parametri possono essere modificati al fine di incrementare la velocità di esecuzione (se, ad esempio, il numero di composti da testare è molto elevato) oppure per massimizzare l'accuratezza del risultato.

Algoritmi di ricerca conformazionale

Gli algoritmi di ricerca si suddividono in:

1. **Docking rigido:** rappresenta il primo tipo di protocollo di docking sviluppato. I componenti del sistema, ossia gli atomi dei ligandi e della proteina, vengono trattati come entità rigide, e vengono considerati esclusivamente sei gradi di libertà roto-traslazionali. Questa forma di approssimazione attualmente trova applicazione principalmente nel docking proteina-proteina. In generale, tali approcci modellano il sito di legame e il ligando assimilandoli a "hot-point" la cui sovrapposizione comporta il *binding*.
2. **Docking semiflessibile:** è stato sviluppato al fine di superare le limitazioni del docking rigido e viene largamente utilizzato nell'ambito dei virtual screening. Oltre ai sei gradi di libertà roto-traslazionale, in questo protocollo viene considerata anche la libertà conformazionale del ligando data dall'esplorazione dei suoi diedri rotabili. La proteina invece viene mantenuta rigida. Questa semplificazione implica che l'attendibilità dei risultati sia strettamente legata alla struttura della proteina utilizzata, costituendo la principale limitazione di questo protocollo.
3. **Docking flessibile:** questo protocollo è il più completo e avanzato, in quanto considera sia la flessibilità della proteina che del ligando. Permette di descrivere eventi di *binding* di tipo *induced-fit* e selezione conformazionale. Tuttavia, questa maggiore precisione è raggiunta a discapito della velocità di esecuzione dello screening. Per questo motivo il docking flessibile non è molto utilizzato e vengono preferite tecniche più avanzate come la dinamica molecolare.

Scoring functions

Gli algoritmi di scoring si suddividono in:

1. **Scoring function basate sul force field:** Il campo di forze rappresenta un elemento cruciale nella meccanica molecolare, esso approssima l'energia potenziale di un sistema a una combinazione di interazioni intra e intermolecolari. Per quanto riguarda la componente

intermolecolare, le *scoring function* in questo contesto sono prevalentemente composte da due elementi principali: il potenziale elettrostatico descritto attraverso la funzione di Coulomb e l'energia di van der Waals, delineata dal potenziale di Lennard-Jones. Un esempio di *scoring function* basata sul campo di forze è GoldScore, utilizzata nel protocollo di docking GOLD.

- 2. *Scoring function empiriche:*** esse rappresentano un buon compromesso tra la precisione del punteggio assegnato e la rapidità di valutazione delle *poses* molecolari. Comprendono la somma di diversi termini energetici determinati empiricamente, tra cui quelli relativi alle interazioni elettrostatiche, di van der Waals, di desolvatazione, entropiche, idrofobiche e i legami a idrogeno. Esempi di funzioni di *scoring* empiriche includono ChemScore, impiegata nel protocollo di docking GOLD, e ChemPLP utilizzata nel protocollo PLANTS.
- 3. *Scoring function knowledge-based:*** si basano sull'analisi statistica dei contatti intermolecolari ravvicinati all'interno di vasti database di strutture 3D, come il Protein Data Bank. Questo metodo si basa sull'assunzione che interazioni intermolecolari a corto raggio tra alcuni tipi di atomi o gruppi funzionali che si verificano più spesso rispetto ad altre siano particolarmente favorevoli per il *binding*. La *scoring function* mappa le interazioni tra coppie di atomi ligando-proteina e le confronta con il database, restituendo uno *score* per le *pose* ottenute. Sono meno impiegate rispetto ad altri approcci. Un esempio di *scoring function knowledge-based* è rappresentato da GOLD-ASP.
- 4. *Scoring function basate sul machine learning:*** a differenza delle *scoring function* descritte in precedenza, quelle basate sul *machine learning* non presentano una forma funzionale predeterminata, ma essa viene dedotta direttamente dai dati. Questo tipo di funzioni di *scoring* hanno dimostrato di riuscire a superare quelle classiche nella previsione dell'affinità di legame per vari complessi proteina-ligando.

PLANTS-ChemPLP^{41,47}

PLANTS (*Protein-Ligand Ant System*) è un protocollo di *docking* molecolare basato su un algoritmo ispirato al comportamento sociale delle formiche. Nell'ambiente naturale, le comunità di formiche lasciano una traccia di feromoni che svolge la funzione di guidare gli spostamenti delle altre formiche, orientandole verso il percorso ottimale per raggiungere il cibo.

Gli algoritmi ACO (*Ant Colony Optimization*) mimano questo comportamento caratteristico delle formiche e vengono utilizzati in svariati ambiti per individuare il percorso più breve che unisce un numero definito di nodi.

PLANTS esplora lo spazio conformazionale del ligando, considerando diverse conformazioni e orientazioni della molecola per trovare quelle che ottimizzano l'interazione con la proteina target.

Ogni grado di libertà è associato ad un “feromone”, e le “formiche virtuali” esplorano gradi di libertà, contribuendo al deposito di feromoni nel caso di identificazione di conformazioni fruttuose. L’esplorazione da parte di ulteriori formiche è poi direzionata sulla base dei valori dei “feromoni”.

PLANTS utilizza un protocollo di ricerca stocastico, dove il ligando viene trattato come una entità flessibile. Il ligando avrà dunque $6+r_1$ gradi di libertà (traslazionali, rotazionali e torsionali); per quanto riguarda la proteina è ammessa la ricerca conformazionale solo di idrogeni legati a eteroatomi. Le *docking poses* generate da PLANTS tramite *local search and refinement* vengono ordinate secondo una *scoring function* ascendente (più negativo è lo *score*, migliore sarà la *binding mode* ottenuta). Il numero di *pose* generate per il ligando viene scelta dall’utente (10 in questo lavoro), e per evitare *poses* ridondanti e poco eterogenee, vengono salvate conformazioni che massimizzano la loro diversità, con un RMSD (*Root Mean Square Deviation*, che quantifica la deviazione quadratica media delle posizioni degli atomi) sopra un valore soglia (2Å in questo lavoro).

Le *scoring function* disponibili con PLANTS sono PLANTSPLP e PLANTSHEMPLP. Per questo lavoro è stata impiegata la seconda, che è in realtà una derivazione della prima. Mentre PLANTSPLP valuta la complementarità sterica tra proteina e ligando, CHEMPLP considera anche le interazioni chimiche, grazie all’inclusione di termini derivanti da ChemScore⁴⁷.

La funzione di *scoring* CHEMPLP valuta l’accuratezza delle diverse *poses* di *docking* generate dall’algoritmo, tenendo conto di diversi contributi energetici, inclusi legami idrogeno, interazioni elettriche, forze di van der Waals e altri termini energetici. Questa valutazione dettagliata mira a fornire una stima accurata dell’energia complessiva dell’interazione proteina-ligando.

L’applicazione di PLANTS per la predizione di complessi ligando-proteina, lo studio di modalità di interazione e la guida nella progettazione di nuovi composti, si è rivelata di estrema efficacia. Tuttavia, per le motivazioni sopra descritte, può richiedere risorse significative da un punto di vista computazionale.

6.5 Dinamica molecolare

La dinamica molecolare (MD) è una tecnica computazionale finalizzata alla simulazione dell’evoluzione temporale di sistemi molecolari, spesso di rilevanza biologica. Per generare una traiettoria di dinamica molecolare, è essenziale calcolare le posizioni (coordinate) degli atomi nel corso del tempo. A tal fine, si applicano le leggi del moto di Newton per prevedere la posizione spaziale di ciascun atomo in funzione del tempo.

Queste simulazioni si rivelano strumenti estremamente utili per diverse ragioni. Ad esempio, consentono l’osservazione della posizione e del movimento dei sistemi a livello atomico, un’informazione difficile da ottenere con tecniche sperimentali.

La dinamica molecolare si configura nel contesto della meccanica molecolare descritta nel paragrafo 6.3. Ogni atomo viene approssimato ad una massa puntiforme associata ad una carica parziale, e i legami vengono approssimati da una molla. L'energia potenziale del sistema viene calcolata sulla base del campo di forze, che comprende termini energetici riguardanti le interazioni intermolecolari e intramolecolari, come descritto in precedenza. L'evoluzione temporale del sistema viene descritta attraverso l'integrazione delle equazioni newtoniane del moto, utilizzando una discretizzazione del tempo in *timestep* nell'ordine del femtosecondo. Grazie a questa semplificazione la scala dei tempi richiesti per le simulazioni di questi sistemi si accorcia notevolmente e con essa le risorse necessarie da un punto di vista computazionale.

Il punto di partenza di una simulazione è una struttura tridimensionale ottenuta sperimentalmente oppure generata tramite *homology modeling* e *de novo prediction*. A seguito della minimizzazione delle strutture, viene assegnata una velocità randomica ad ogni atomo, tale da ottenere complessivamente un'energia cinetica media che corrisponda a una determinata temperatura, usualmente 310K. Successivamente, tramite l'integrazione delle leggi newtoniane del moto ad intervalli definiti *timestep*, vengono calcolate le posizioni atomiche e le nuove velocità di ciascun atomo in maniera sequenziale a seguito dell'accelerazione dovuta alle forze descritte dal campo di forze. Ciò viene fatto in maniera sequenziale e i *frame* vengono salvati con una frequenza definita fino a creare una traiettoria.

La dinamica classica, anche grazie al progresso tecnologico degli acceleratori GPU, ha permesso di raggiungere tempi di simulazione computazionali di nanosecondi/microsecondi a seconda della potenza di calcolo e del tempo reale a disposizione. Tuttavia, eventi di *binding* e *unbinding* possono essere difficili da osservare poiché richiedono tempi di simulazione lunghi. Le tecniche di dinamica di *enhanced sampling* sono state sviluppate per velocizzare l'osservazione di questi fenomeni.

6.6 Modello Farmacoforico

Un farmacoforo è un modello molecolare che rappresenta le caratteristiche strutturali e chimiche di una molecola che sono responsabili dell'interazione con uno specifico bersaglio biologico e di conseguenza cruciali per l'attività biologica della molecola.

Il modello farmacoforico non rappresenta una molecola reale, ma è un'approssimazione delle diverse caratteristiche chimiche che accomunano diversi ligandi attivi nella capacità di riconoscere il *target*. In particolare, un modello farmacoforico riassume gli elementi interattivi condivisi tra molecole attive e loro distanza relativa.

Una volta creato un modello farmacoforico in cui si evidenziano le caratteristiche comuni ai vari ligandi attivi, è possibile analizzare una libreria di composti ad attività ignota per valutare la

somiglianza con l'ipotesi farmacoforica. Questo può portare alla scoperta di nuovi possibili ligandi per il target in esame.

A livello pratico, ligandi ad attività nota vengono "tradotti" in una serie di *features farmacoforiche* come la presenza di: accettori di legami idrogeno, donatori di legami idrogeno, anelli aromatici, gruppi idrofobici ecc. Ogni singola *feature* è solitamente rappresentata da una sfera con un determinato raggio che fissa la tolleranza dalla posizione esatta. Le *features* possono essere verificate singolarmente o come combinazioni logiche di "AND" ("e"), "OR" ("o") e "NOT" ("non"), indicanti la compresenza necessaria/possibile/impedita di diverse *features*.

Tra i software utilizzabili per la generazione di farmacofori si ricordano MOE (utilizzato in questo lavoro di tesi), OpenBabel e Glide (Schrödinger).

6.7 PLIF (*Protein Ligand Interaction Fingerprint*)

Il *tool* per la generazione delle PLIF (*Protein Ligand Interaction Fingerprints*) di MOE è un metodo per computare le interazioni che si formano all'interfaccia proteina-ligando. Questo può essere applicato a database di complessi ottenuti tramite *docking* molecolare, infatti l'uso delle *fingerprints* per rappresentare le interazioni proteina-ligando è un modo efficace di filtrare database numerosi sulla base della presenza o meno di interazioni con residui chiave del target. Le interazioni, come legami a idrogeno, interazioni ioniche e contatti tra superfici, sono classificate per ogni ligando.

Ci sono 9 tipi di interazioni che possono coinvolgere un residuo: legami a idrogeno (donatori e accettori) che coinvolgono la catena laterale, legami a idrogeno (donatori e accettori) che coinvolgono il *backbone*, legami a idrogeno (donatori e accettori) che coinvolgono i solventi, interazioni ioniche, interazioni di legame con ioni metallici e interazioni π .

È possibile introdurre dei *cutoff* al fine di mostrare solo le interazioni più deboli o più forti per ogni tipologia di interazione che si desidera investigare.

6.8 Solvent exposure

In MOE (*Molecular Operating Environment*), l'esposizione al solvente si riferisce alla misura in cui la superficie di un ligando è esposta al solvente, e si esprime come una percentuale dell'area superficiale del ligando.

Parte terza:

RISULTATI E DISCUSSIONE

Questo progetto di tesi ha portato all'identificazione di 11 potenziali inibitori della Casein chinasi 1 di *Plasmodium falciparum* (PfCK1), attraverso la definizione di uno specifico protocollo di *virtual screening* basato sul *docking*, applicato ad un *database* di composti naturali e analoghi.

La crescente preoccupazione per la resistenza di *Plasmodium falciparum* alle principali terapie in uso ci ha portati a voler indagare un *target* che finora non è stato oggetto di ricerche approfondite nel campo del *drug discovery*, ma che da studi in letteratura sembrerebbe essere di interesse terapeutico, la proteinchinasi CK1 di *Plasmodium falciparum*.

7. Generazione del modello di omologia

Dal momento che per svolgere un *virtual screening* basato sul *docking* è fondamentale la conoscenza della struttura tridimensionale della proteina *target* e data l'assenza, ad oggi, di una struttura risolta sperimentalmente della PfCK1, è stato necessario costruirne un modello.

Il primo passo dell'indagine è stato l'identificazione di un *template* finalizzato alla costruzione di un modello per omologia. Questo viene fatto attraverso il confronto della sequenza amminoacidica della proteina desiderata (dalla struttura ignota) con le sequenze di proteine la cui struttura tridimensionale è nota, alla ricerca di sequenze che massimizzino l'identità di sequenza con il target. La sequenza di PfCK1 è stata confrontata con la sequenza di tutte le proteine umane per le quali è disponibile una struttura tridimensionale nel *Protein Data Bank* (PDB). Questo approccio è stato adottato per ottenere la proteina umana più simile a PfCK1, con una percentuale di identità di sequenza sufficientemente elevata da poter essere utilizzata come *template* affidabile per la generazione del modello di PfCK1. Per effettuare tale confronto, è stata utilizzata la *suite Protein BLAST*, fornita dal *National Institute of Health* (NIH)⁴⁸. Attraverso questa piattaforma, è possibile inserire la sequenza amminoacidica di interesse (nel nostro caso, la sequenza di PfCK1) e richiedere un confronto con un *database* predefinito. Nella nostra indagine, abbiamo limitato la ricerca alle proteine appartenenti al proteoma di *Homo sapiens* che presentano una struttura risolta nel PDB.

La sequenza amminoacidica di Casein chinasi 1 di *Plasmodium falciparum* è stata scaricata da PlasmoDB. La ricerca ha dato il risultato visibile in Tabella 3.

Vengono di seguito descritti i diversi punteggi che vengono calcolati da *Protein BLAST* per valutare la similarità tra proteine:

1. **Max[imum] Score (Punteggio Massimo):** è il punteggio di allineamento più alto calcolato dalla somma tra ricompense per gli amminoacidi corrispondenti e penalità per le non corrispondenze e gli spazi vuoti;
2. **Tot[al] Score (Punteggio Totale):** è la somma dei punteggi di allineamento di tutti i segmenti provenienti dalla stessa sequenza di riferimento;
3. **Query Cover[age] (Copertura della Query):** indica la percentuale della lunghezza della sequenza di ricerca che è inclusa nei segmenti allineati;
4. **E[xpect] Value (Valore di Aspettativa):** rappresenta il numero di allineamenti, con il punteggio calcolato o migliore, dovuti al caso. **Il valore di aspettativa è la metrica di ordinamento predefinita;** per allineamenti significativi, il valore di aspettativa dovrebbe essere molto vicino a zero;

5. Ident[jity] (Identità): Rappresenta la percentuale più alta di identità per un insieme di segmenti allineati alla stessa sequenza di riferimento;

Dall'output emerge che la proteina umana più simile a PfCK1 è Casein chinasi 1 nell'isoforma delta. In particolare, tra le varie strutture cristallografiche di questa proteina presenti sul PDB, la migliore risulta essere quella con il codice PDB 6RCG, con una percentuale di identità di sequenza del 69,07%. Essa è dunque la struttura scelta come *template* per costruire il modello di omologia.

	Description	Max Score	Tot Score	Query Cover	E value	Per. Ident.	Acc Len	Accession
1	Chain A, Casein kinase I isoform delta [Homo sapiens]	452	452	90%	1e-161	69.07%	296	6RCG_A
2	Chain A, Casein kinase I isoform delta [Homo sapiens]	451	451	90%	2e-161	69.07%	296	3UYS_A
3	Chain A, Casein kinase I isoform delta [Homo sapiens]	451	451	90%	2e-161	69.07%	294	5IH4_A
4	Chain A, Casein kinase I isoform delta [Homo sapiens]	452	452	90%	2e-161	68.60%	301	8D7M_A
5	Chain A, Casein kinase I isoform delta [Homo sapiens]	452	452	90%	3e-161	68.60%	317	4KBC_A
6	Chain A, Casein kinase I isoform delta [Homo sapiens]	451	451	90%	3e-161	69.07%	296	7P7F_A
7	Chain A, Casein kinase I isoform delta [Homo sapiens]	452	452	90%	3e-161	69.07%	318	4TN6_A
8	Chain A, Casein kinase I isoform delta [Homo sapiens]	451	451	90%	4e-161	69.07%	295	4TW9_A
9	Chain A, Casein kinase I isoform delta [Homo sapiens]	451	451	90%	4e-161	69.07%	314	5MQV_A
10	Chain A, Casein kinase I isoform epsilon [Homo sapiens]	451	451	90%	5e-161	69.07%	296	4HNI_A
11	Chain A, Casein kinase I isoform delta [Homo sapiens]	450	450	90%	7e-161	68.73%	294	6PXP_A
12	Chain A, Casein kinase I isoform delta [Homo sapiens]	451	451	90%	7e-161	69.07%	314	5OKT_A
13	Chain A, Casein kinase I isoform delta [Homo sapiens]	452	452	89%	5e-160	68.97%	415	6PXN_A
14	Chain A, Casein kinase I isoform delta [Homo sapiens]	447	447	89%	2e-159	68.62%	331	4KB8_A
15	Chain A, Casein kinase I isoform delta [Homo sapiens]	427	427	99%	2e-151	62.35%	337	7WTT_A

Tabella 3: Risultato della ricerca della sequenza amminoacidica ottenuta tramite Protein BLAST.

Chain A, Casein kinase I isoform delta [Homo sapiens]

Sequence ID: [6RCG_A](#) Length: 296 Number of Matches: 1

[See 24 more title\(s\)](#) [See all Identical Proteins\(IPG\)](#)

Range 1: 3 to 293 [GenPept](#) [Graphics](#)

[▼ Next Match](#) [▲ Previous Match](#)

Score	Expect	Method	Identities	Positives	Gaps
452 bits(1162)	1e-161	Compositional matrix adjust.	201/291(69%)	251/291(86%)	0/291(0%)
Query 1	MEIRVANKYALGKKLGSFSGFDIYVAKDIVTMEEFVAVKLESTRSKHPQLLYESKLYKILG				60
Sbjct 3	ME+RV N+Y LG+K+GSGSFGDIY+ DI EE A+KLE ++KHPQL ESK+YK++ MELRVGNRYRLGRKIGSFGDIYLTGTDIAAGEEVAIKLECVKTKHPQLHIESKIYKMMQ				62
Query 61	GGIGVPKVVYWGIEGDFTIMVLDLLGPSLEDLFTLCNRKFSKTKVLTADQMLNRIEYVH				120
Sbjct 63	GG+G+P + W G EGD+ +MV++LLGPSLEDLF C+RKFSKTKVL+ ADQM++RIEY+H GGVGIPTRIRWCGAEGDYNVMVMELLGPSLEDLNFCSRKFSKTKVLLADQMISRIEYIH				122
Query 121	SKNFIHRDIKPDNFLIGRGKVTLIHIIDFGLAKKYRDSRSHTHIPYKEGKNLTGTARYA				180
Sbjct 123	SKNFIHRD+KPDNFL+G GKK L++IIDFGLAKKYRD+R+H HIPY+E KNLGTGTARYA SKNFIHRDVKPDNFLMGLGKKNLVYIIDFGLAKKYRDARTHQHIPYRENKLTGTARYA				182
Query 181	SINTHLGIEQSRDDIEALGYVLMYFLRGLPWQGLKAISKKDYKIMEKKISTSVEVL				240
Sbjct 183	SINTHLGIEQSRDD+E+LGYVLMYF GSLPWQGLKA +K+ KY++I EKK+ST +EVL SINTHLGIEQSRDDLES LGYVLMYFNLSLPWQGLKAATKRQYERISEKKMSTPIEVL				242
Query 241	CRNASFEFVTYLNYCRSLRFEDRDPDYTYLRRLLKDLFIREGFTYDFLFDWT			291	
Sbjct 243	C+ EF TYLN+CRSLRF+D+PDY+YLR+L ++LF R+GF+YD++FDW CKGYPSEFATYLNFCRSLRFDDKPDYSYLRQLFRNLFHRQGFSDYDVFDDWN			293	

Figura 21: Allineamento delle sequenze amminoacidiche di CK1 (6RCG) e PfCK1

Attraverso l'uso del tool apposito di MOE *Homology Model*, fornendo la struttura tridimensionale di 6RCG come *template* e la sequenza amminoacidica di PfCK1 come *query*, è stato calcolato l'allineamento tra le due sequenze e generato il modello tridimensionale di PfCK1.

Il modello è stato preparato con i tool Protein Preparation e Protonate 3D per l'assegnazione della conformazione dello stato tautomerico di istidina, dello stato di protonazione e conformazione di asparagina e glutammina.

7.1 Validazione del modello

Il modello generato necessita di essere validato ed ottimizzato prima di procedere con le simulazioni di *virtual screening*. In questo caso, l'approccio è stato quello di identificare ligandi di PfCK1 (veri positivi, con dati di *binding* sperimentali) riportati in letteratura e verificare la loro capacità di interagire con la proteina *in silico*: in altri termini, è stata testata la compatibilità del modello con una modalità di *binding* ragionevole per ligandi veri positivi di PfCK1.

Due molecole, secondo uno studio⁴⁹, hanno dimostrato di avere capacità di legare PfCK1. Il dato di *binding* disponibile per PfCK1 è la Kd apparente, ossia la costante di dissociazione del complesso proteina-ligando ottenuta tramite saggi di competizione.

Si ricorda che il valore di Kd è inversamente proporzionale alla stabilità del complesso, secondo l'Equazione 7.

$$K_d = \frac{[\text{prodotti}]}{[\text{reagenti}]} = \frac{[\text{proteina}] * [\text{ligando}]}{[\text{complesso}]}$$

Equazione 7: formula per il calcolo della costante di dissociazione K_d .

Di conseguenza la concentrazione della proteina e del ligando dissociati sono direttamente proporzionali alla K_d , mentre la concentrazione del complesso è inversamente proporzionale alla K_d .

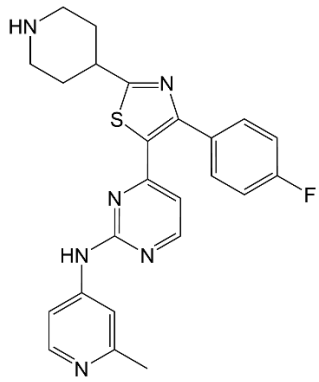
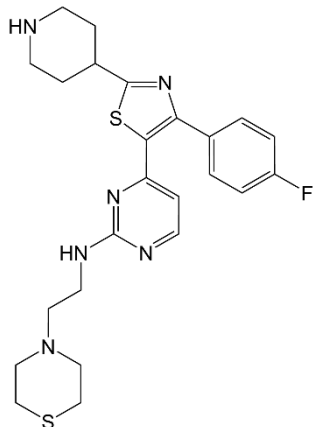
ChEMBL ID	Structure	Kd	MW
CHEMBL4591417		150.0 nM	446.56 Da
CHEMBL4461850		1500.0 nM	484.67 Da

Tabella 4: Nomenclatura, K_d e peso molecolare dei due ligandi veri positivi per PfCK1.

È stato eseguito un *docking* di queste due molecole sul modello precedentemente creato di PfCK1. Il centro utilizzato per definire la posizione della *search box* per il *docking* è stato il centro di massa del ligando cristallografico di 6RCG (che si ricorda essere la struttura di CK1 δ umana usata come *template*), assumendo che le molecole siano ligandi competitivi rispetto all'ATP.

Per le immagini che raffigurano le interazioni proteina-ligando in 2D, la legenda adottata è visibile in Figura 22: legenda delle interazioni 2D tra proteina e ligando. Figura 22.

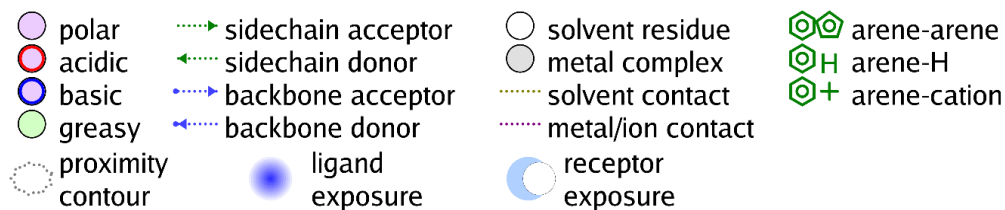


Figura 22: legenda delle interazioni 2D tra proteina e ligando.

In particolare, per quanto concerne i legami a idrogeno, si è deciso di mostrare solo quelli con una energia superiore a 1.5 kcal/mol.

La *pose* migliore per ognuno dei due composti, secondo la *scoring function* CHEMPLP, è quella riportata in Tabella 5.

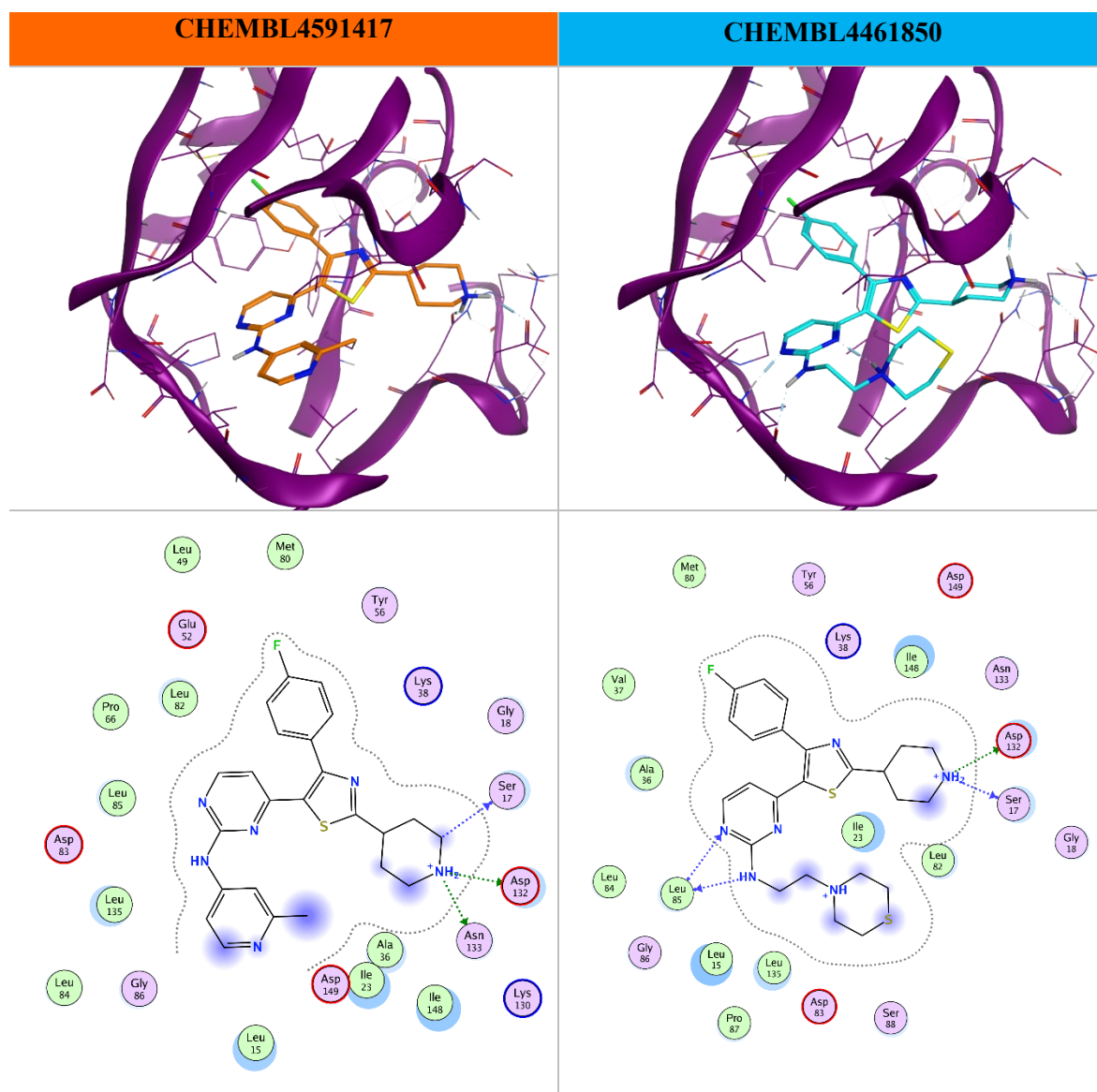


Tabella 5: Confronto tra le pose dei due ligandi veri positivi sul modello di PfCK1 pre-dinamica.

Nelle *binding mode* generate e selezionate per i due composti veri positivi, lo *scaffold* 2-aminopirimidinico viene posizionato in prossimità del residuo Leucina 85 (Leu85). Da un'analisi che ha coinvolto le strutture cristallografiche di vari complessi di CK1 δ umana legata a vari inibitori nella *ATP-binding pocket*, si è osservato un coinvolgimento proprio del residuo Leu85 (conservato tra CK1 δ umana e PfCK1) nell'interazione con la maggior parte dei ligandi. In particolare, si è visto che le interazioni più conservate sono due legami a idrogeno con Leu85, in particolare un primo legame a idrogeno in cui i ligandi si comportano da donatori nei confronti dell'ossigeno carbonilico del *backbone* di Leu85, e un secondo in cui si comportano da accettori verso l'azoto della porzione amminica dell'amminoacido. Leu85 è un residuo che si colloca nella *hinge region*, la regione cerniera flessibile nelle cui prossimità si colloca il sito attivo della chinasi.

Tornando ai risultati di docking dei due ligandi di PfCK1, si osserva che il legame a idrogeno bidentato è presente soltanto nella molecola CHEMBL4461850, colorata in azzurro in Tabella 5. Tuttavia, il nucleo 2-aminopirimidinico della molecola CHEMBL4591417 non è distante dal residuo Leu85.

È stata perciò effettuata una ottimizzazione del modello al fine di migliorare il pattern di interazione dei veri positivi con l'enzima, in modo conforme alla modalità di legame tipica degli inibitori delle Casein chinasi.

Per ottimizzare il modello, si è deciso quindi di svolgere una simulazione di dinamica molecolare. È stato utilizzato ai fini dell'ottimizzazione unicamente il complesso formato dal modello proteico con il vero positivo a più alto peso molecolare e con maggiore ingombro sterico, ossia CHEMBL4461850. La simulazione di dinamica è stata ottenuta tramite l'impiego del motore di dinamica ACEMD.

Il complesso è stato inizialmente inserito in un box d'acqua prismatico a base rettangolare, con un *padding* di 15 Å. Sono stati poi aggiunti ioni monovalenti per neutralizzare il sistema, portandolo alla concentrazione 0.154 M. Infine, è stata eseguita una minimizzazione a 500 step per rimuovere eventuali *clash* e due fasi di equilibratura. Nella prima fase di equilibratura (0.1 ns), condotta in condizioni di volume e temperatura costante (310K), solamente acqua e ioni sono liberi di muoversi; nella seconda fase (0.5 ns), condotta a temperatura e pressione costanti (310K e 1 atm), acqua, ioni e catene laterali della proteina sono liberi di muoversi.

Infine, il complesso è stato sottoposto a 25 ns di dinamica classica senza *constraint*, a volume e temperatura costanti (310 K). Al termine della simulazione, è stato estratto l'ultimo *frame* della traiettoria, che è riportato in Tabella 6.

CHEMBL4461850

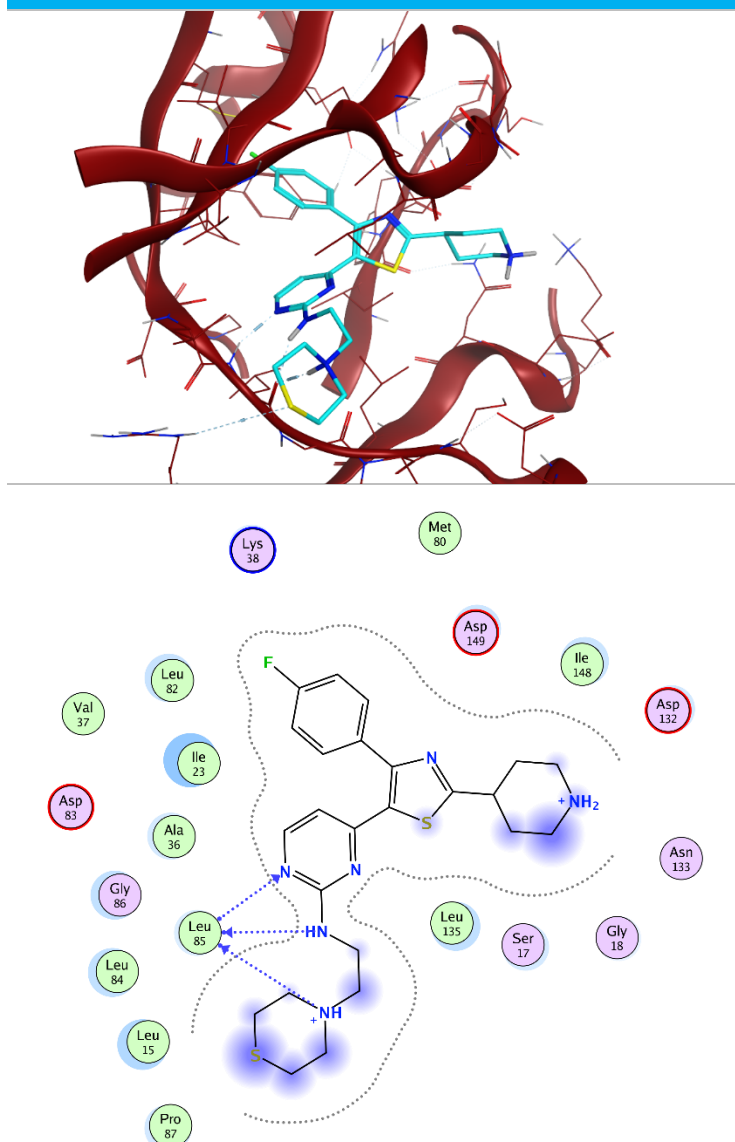


Tabella 6: Rappresentazione 3D e 2D della pose di docking ottenuta post rilassamento in dinamica molecolare con il ligando CHEMBL4461850.

Il modello rifinito tramite dinamica è stato validato effettuando un ulteriore *docking* dei due ligandi veri positivi. Dall'ultimo *frame* della dinamica sono state perciò estratte le coordinate della proteina, e utilizzate in una simulazione di *docking*. In questo caso le coordinate del centro di massa considerato per definire la posizione della *search box* per il *docking* sono state modificate, utilizzando quelle del ligando CHEMBL4461850 a seguito del rilassamento in dinamica, ed è stato impostato un raggio di 10.5 Å.

Nelle tabelle successive vengono riassunte le *poses* di *docking* delle due molecole appena discusse, sul modello pre- e post-ottimizzazione tramite dinamica molecolare, con un focus sul sito attivo.

Nella colonna di sinistra, dove la proteina è colorata in viola, vengono riportate le *poses* delle due molecole prima della simulazione di dinamica, mentre nella colonna di destra, con la proteina in colore rosso, le *poses* dopo la simulazione di dinamica.

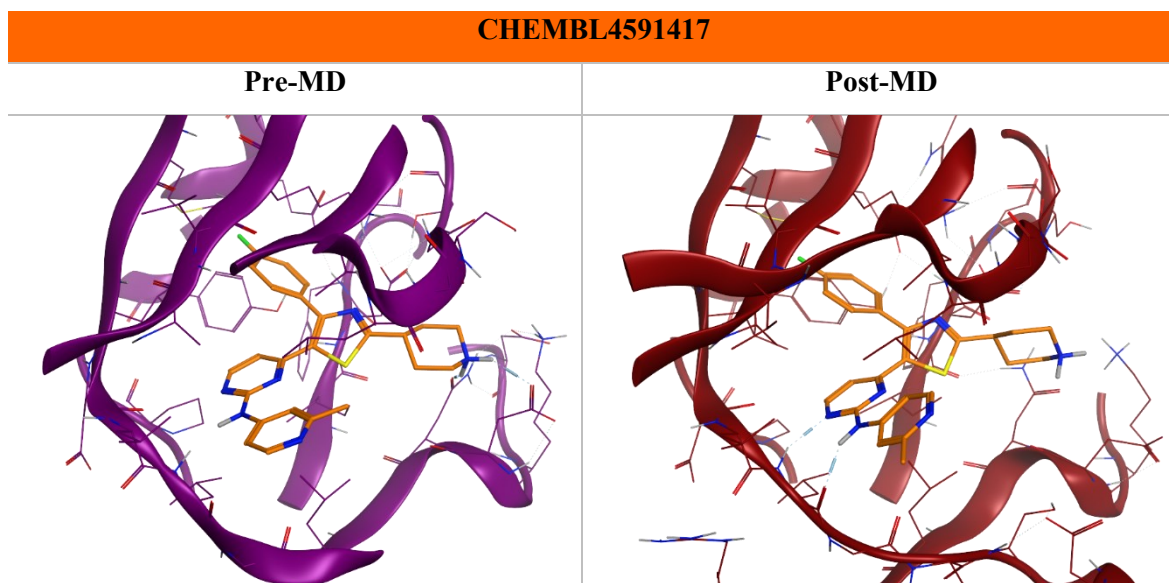


Tabella 7: confronto tra la pose pre e post-MD del ligando CHEMBL4591417.

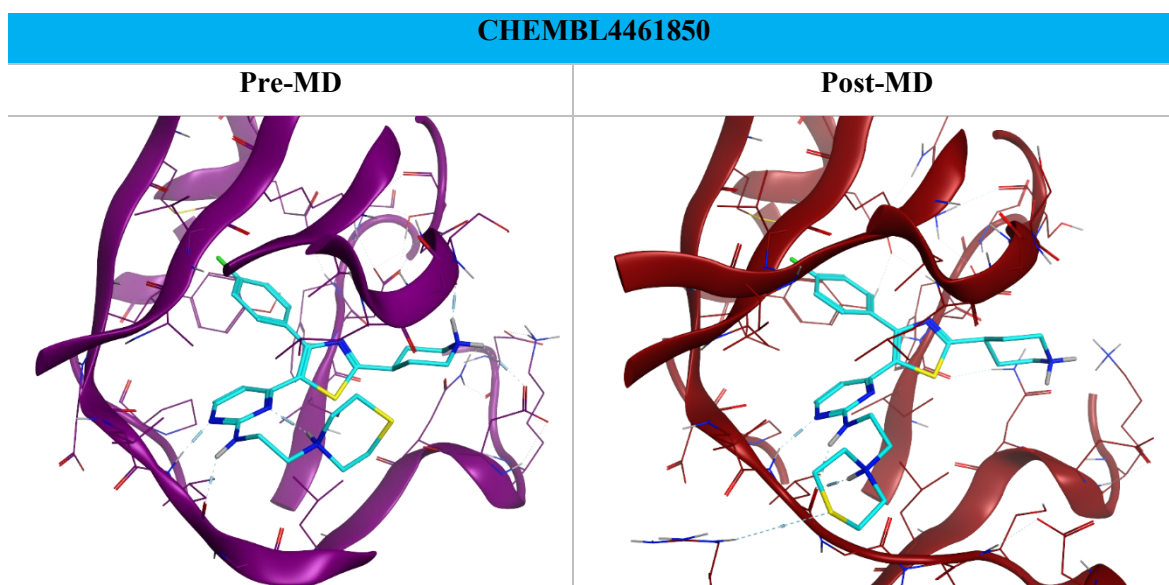


Tabella 8: confronto tra la pose pre e post-MD del ligando CHEMBL4461850.

Si può notare che entrambi i ligandi, dopo la simulazione di dinamica molecolare, presentano il *pattern* di interazioni ideale che coinvolge la Leu85 nella regione *hinge*, formando con essa due legami a idrogeno. Prima del rilassamento in dinamica del complesso ligando-proteina solo il ligando meno attivo (e più ingombrato) mostrava questo *pattern* d'interazione, mentre il ligando più attivo (e meno ingombrato) formava solo un legame a idrogeno anziché due.

Per validare ulteriormente la bontà del modello, è stata indagata la sua stabilità in dinamica tramite il calcolo del RMSD ad ogni *frame*. L'RMSD è indice della deviazione nel tempo delle coordinate

del sistema rispetto ad un riferimento. In particolare, per tralasciare i moti di rotazione e traslazione della proteina ed evidenziare soltanto le variazioni conformazionali, la proteina durante il corso dell'analisi del RMSD è stata sovrapposta al *frame* iniziale.

Viene riportata in seguito la formula per il calcolo del RMSD. Essa viene applicata per ogni *frame* della simulazione, dove X, Y e Z rappresentano le coordinate atomiche del primo *frame* usato come riferimento (r), e del *frame* corrente (f).

$$RMSD(r, f) = \sqrt{\frac{\sum_{i=1}^{N_{atomi}} ((X_{ri} - X_{fi})^2 + (Y_{ri} - Y_{fi})^2 + (Z_{ri} - Z_{fi})^2)}{N_{atomi}}}$$

Equazione 8: Equazione del RMSD.

Il grafico mostra in blu le fluttuazioni del valore di RMSD del *backbone* dell'intera proteina, mentre in giallo quelle del *backbone* dei residui del *binding site*, definito tramite l'utilizzo del *tool* di MOE *Near Residues*, calcolato sul ligando CHEMBL4461850 a seguito del *docking*. Solitamente il sito di legame è più stabile dell'intera proteina, poiché quest'ultima comprende molte regioni ricche di strutture ad alta flessibilità come i *loop*, i quali contribuiscono ad aumentare la variazione di RMSD e il suo valore medio. La proteina e in particolar modo il *binding site* hanno globalmente dei valori bassi di RMSD. Il valore medio è 1.46 Å per l'intera proteina e 0.78 Å per il *binding site*, comparabili con la risoluzione della struttura cristallografiche 6RCG usata come *template*, e questo è indice di una significativa stabilità del modello durante la traiettoria.

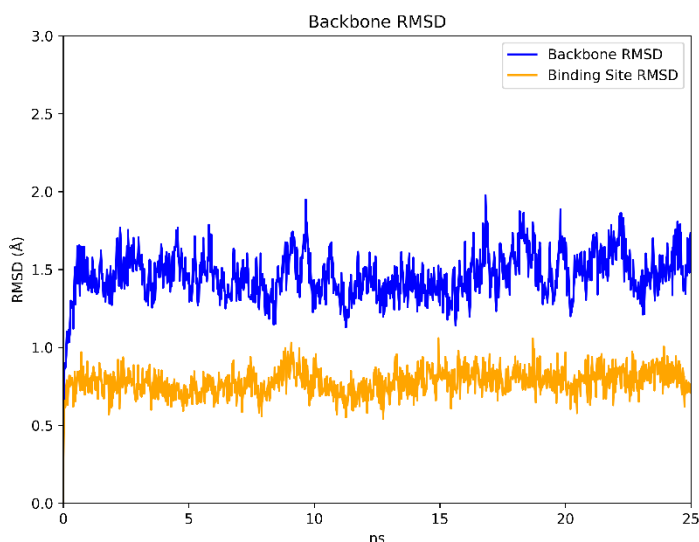


Figura 23: grafico dell'andamento del RMSD in funzione del tempo.

Il modello proteico ottimizzato dalla simulazione di dinamica, dopo queste validazioni, viene ritenuto idoneo ad essere utilizzato come *input* per il *virtual screening* di una libreria di composti. Il centro

di massa che viene utilizzato per il *virtual screening* sarà quello del ligando vero positivo CHEMBL4461850 a seguito del rilassamento in dinamica. Il raggio è definito a 10.5 Å.

7.2 Interazione con la Leu85 del *backbone*

Nella struttura della chinasi si possono distinguere due porzioni ben differenziate, un lobo N-terminale più piccolo, maggiormente composto da β -foglietti, ed un lobo C-terminale più grande, nel quale si ha prevalenza di strutture a α -elica e la presenza di *loops*.

I due domini strutturali sono collegati da una zona di raccordo detta *hinge region*.

L'interazione ligando-proteina comune tra tutti i ligandi di CK1 considerati, che viene mantenuta anche nel *docking* dei composti veri positivi di PfCK1 dopo il rilassamento in dinamica della struttura, è il legame a idrogeno bidentato con la Leu85, situata proprio nella regione *hinge*.

Nell'immagine sottostante viene mostrata una *pose* di *docking* evidenziando il caratteristico legame a idrogeno bidentato tra il modello finale (dopo l'ottimizzazione ottenuta con il rilassamento in dinamica) di PfCK1 e una delle due molecole ad attività nota. Viene rappresentato in giallo il residuo Leu85. Si può notare che la porzione dell'amminoacido che funge da accettore di legame a idrogeno è l'ossigeno carbonilico del *backbone*, mentre il donatore è l'idrogeno legato all'azoto della porzione amminica del *backbone*.

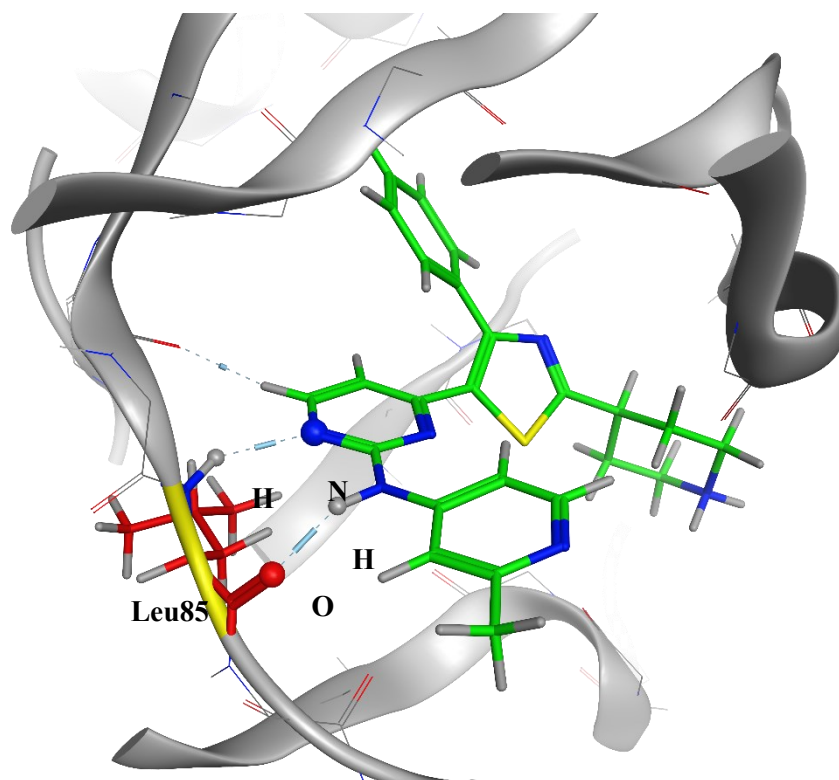


Figura 24: rappresentazione grafica degli atomi coinvolti nel legame a idrogeno bidentato con il residuo Leu85.

Il residuo Leu85 è conservato sia in CK1delta che in PfCK1.

Nelle pagine seguenti, sono mostrate le strutture coinvolte nelle varie fasi della generazione del modello (da Figura 25 a Figura 29).



Figura 25: Struttura tridimensionale di CK1delta umana (PDB: 6RCG).

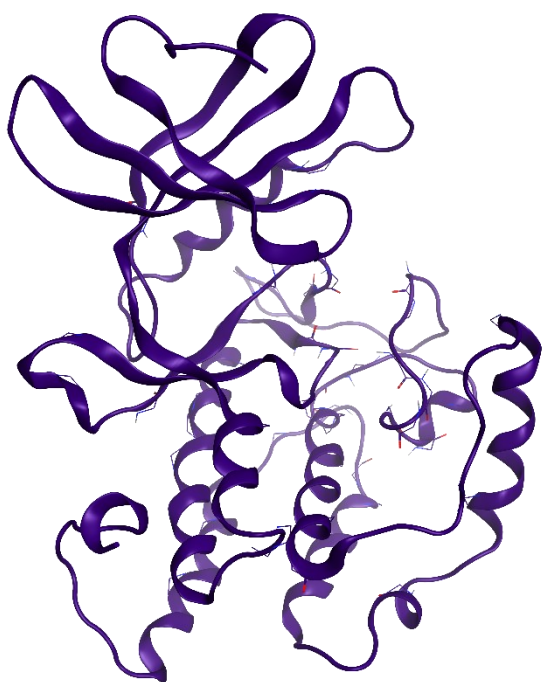


Figura 26: Struttura tridimensionale del modello di omologia di PfCK1 generato con MOE.

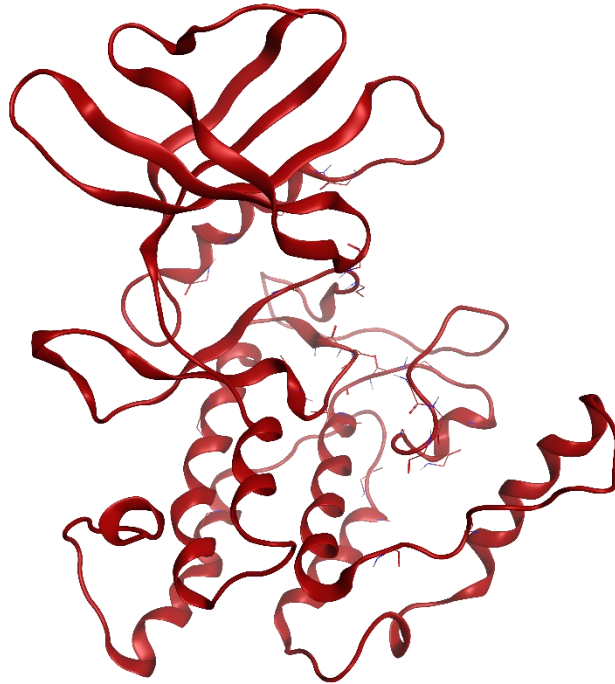


Figura 27: Struttura tridimensionale del modello di omologia di PfCK1 dopo il rilassamento in dinamica.

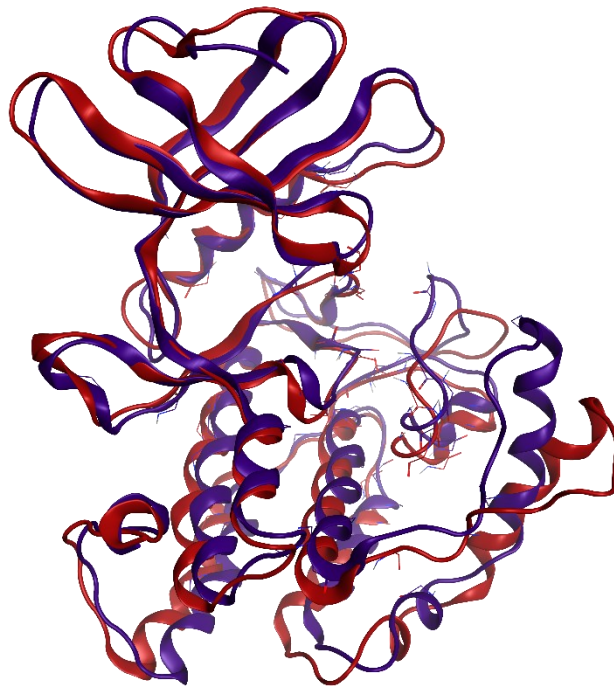


Figura 28: Sovrapposizione tra il modello di PfCK1 pre (in viola) e post (in rosso) dinamica

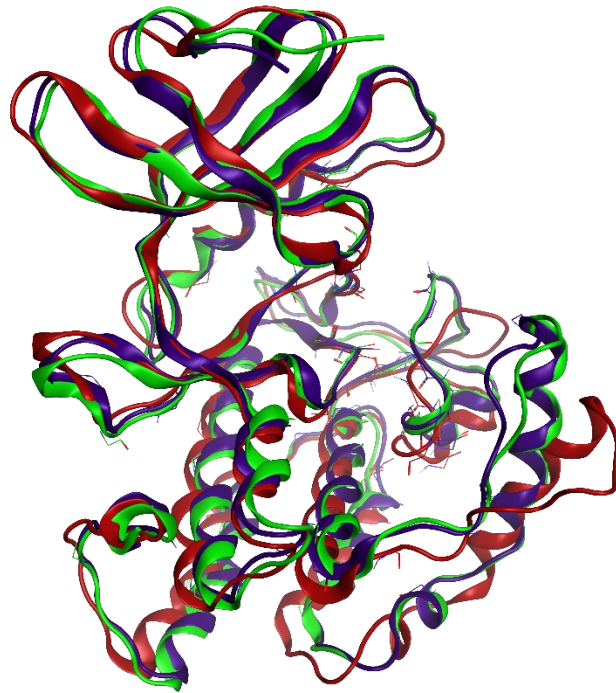


Figura 29: Sovrapposizione tra la struttura cristallografica di 6RCG (in verde), e il modello di PfCK1 pre (in viola) e post (in rosso) dinamica.

In Figura 30 viene mostrata la tabella contenente i valori di RMSD (*Root Mean Square Deviation*, Deviazione Quadratica Media) relativi all'ultima immagine, dove si può vedere la sovrapposizione tra la struttura cristallografica di CK1 δ umana e del modello di PfCK1 pre e post-dinamica. Il valore è ottenuto post allineamento sul *binding site*.

RMSD = 0.589 Å

	1	2	3	
				4.0
				3.5
				3.0
1	0.00	0.63	0.47	2.5
2	0.63	0.00	0.65	2.0
3	0.47	0.65	0.00	1.5
				1.0
				0.5
				0.0

Figura 30: Tabella con i valori di RMSD tra le strutture 6RCG modello pre-MD e post-MD.

L'RMSD è utilizzato come misura quantitativa della somiglianza tra le strutture proteiche.

Con il numero **1** è indicato il modello pre-dinamica di PfCK1, con il numero **2** il modello ottimizzato post-dinamica e con il numero **3** la CK1 δ umana (PDB_{ID} = 6RCG).

Si può notare quanto queste strutture siano sovrapponibili tra loro, in accordo con la percentuale di omologia tra la proteina 6RCG e la sequenza amminoacidica di PfCK1, che si attesta al 69%.

8. Scelta della libreria di ligandi

Partendo dal presupposto che i farmaci storicamente più impiegati per il trattamento della malaria sono derivati da sostanze naturali, come il chinino o l'artemisinina, si è voluto conservare tale approccio in questo studio, orientato all'individuazione di potenziali inibitori della Casein chinasi di *Plasmodium falciparum*. A tal fine, la libreria oggetto della nostra analisi è costituita da un *database* di composti naturali e analoghi reperibile sul sito MolPort e aggiornata al 6 dicembre 2022.

MolPort è una piattaforma *online* per la vendita di reagenti chimici e molecole. Offre agli utenti la possibilità di cercare, confrontare e acquistare una vasta gamma di molecole da diversi fornitori. Inoltre, fornisce informazioni dettagliate su ciascuna molecola, inclusi dati chimici, disponibilità e prezzi.

Questa libreria contiene 113687 composti. Di ciascuno di essi viene fornito solo un codice identificativo e la stringa SMILE corrispondente.

Di conseguenza è stato necessario preparare la libreria, al fine di convertire gli SMILES in coordinate 3D. Per fare questo è stato usato il tool *Wash* di MOE.

9. Docking

Il *docking* è stato svolto tramite PLANTS e la *scoring function* ChemPLP, impostando la generazione di 10 *poses* per ogni molecola. Le coordinate del centro di massa valutate per definire la posizione della *search box* per il *docking* sono quelle del ligando CHEMBL4461850 a seguito del rilassamento in dinamica. Il raggio di ricerca utilizzato per l'algoritmo di *docking* è di 10.5Å.

Da un numero iniziale di composti contenuti nella libreria pari a 113687, è stata effettuata una simulazione di *docking* con PLANTS generando 10 *poses* per ogni molecola. I composti analizzati sono stati 112770, per un totale di 1127700 *docking poses*.

9.1 Filtraggio delle *poses*

Vari filtri sono stati applicati alle *pose* generate dall'algoritmo di *docking*, al fine di scremare il numero iniziale di composti e consentire la selezione di un numero ragionevole da sottoporre a futuri saggi sperimentali.

9.2 Generazione del farmacoforo

Per la creazione del modello farmacoforico, sono state utilizzate le informazioni strutturali relative agli inibitori noti di CK1δ. Questi dati sono stati estratti dalle strutture cristallografiche dei complessi ligando-proteina depositati nel Protein Data Bank (PDB). Le strutture così ottenute sono state successivamente sottoposte all'analisi del *tool* di modellazione farmacoforica *Pharmacophore* di MOE. Le caratteristiche di interazione che hanno in comune almeno il 50% dei ligandi sono state individuate e quindi utilizzate per generare il modello farmacoforico finale.

Come evidenziato in Figura 31 il modello finale è composto da tre *features* farmacoforiche distinte rappresentate graficamente come sfere. Queste includono un accettore e un donatore di ponte idrogeno (rappresentati dalla sfera rosa indicata nell'immagine con il numero 1), entrambi nei confronti del residuo Leu85, e un anello aromatico nelle vicinanze della regione di *hinge* (rappresentato dalla sfera rosa indicata con il numero 2). Le tre sfere più piccole verdi sono posizionate per indicare il collocamento spaziale dell'anello aromatico.

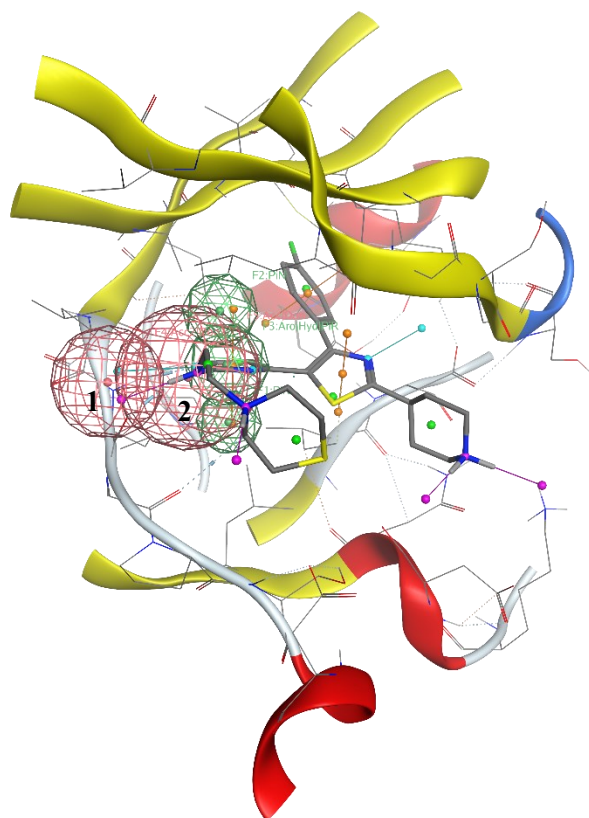


Figura 31: Rappresentazione grafica del modello farmacoforo.

I complessi proteina-ligando di CK1 δ disponibili nel *Protein Data Bank* che sono stati considerati per la generazione del modello farmacoforo sono quelli identificati dai seguenti codici PDB: 3UYT, 3UZP, 4HGT, 4HNF, 4KB8, 4KBA, 4KBC, 4KBK, 4TN6, 4TW9, 4TWC, 5IH5, 5IH6, 5MQV, 5OKT, 5W4W, 6F1W, 6F26, 6GZM, 6HMP, 6HMR, 6RCG, 6RCH, 6RU6, 6RU7, 6RU8, 7NZY, 7P7F, 7P7G, 7P7H.

I 30 complessi cristallografici sono stati scaricati e preparati per le successive analisi attraverso la *suite* MOE. Per la preparazione delle strutture è stato usato il *tool Structure Preparation*, tramite *Protonate 3D* sono stati definiti gli stati di protonazione più probabili a pH 7. Successivamente, gli idrogeni aggiunti sono stati sottoposti a una minimizzazione energetica usando il *force field* AMBER10: EHT. Dopo la fase di preparazione, i complessi proteina-ligando sono stati allineati e sovrapposti tramite MOE, al fine di uniformare le coordinate del sito di legame tra le diverse strutture cristallografiche. Questi complessi sono stati poi utilizzati per la generazione del modello farmacoforo.

9.3 Filtro per farmacoforo

Il primo metodo utilizzato per filtrare le pose ottenute dal *virtual screening* è stato basato sul modello farmacoforo citato nel paragrafo precedente. Tale tecnica permette di isolare, in un'ampia libreria

di molecole di cui è stato fatto un docking, solo quelle in grado di rispettare le *features farmacoforiche* ideali, ossia solo quei composti che hanno una struttura tale da poter formare, se posizionate correttamente dal docking, interazioni con i residui nella tasca di *binding* uguali alle interazioni che fanno i ligandi veri positivi.

È stato utilizzato il tool *Pharmacophore* di MOE, al quale sono state fornite tutte le *poses* derivanti dall'*output* del docking, che ha consentito di selezionare la *pose* migliore (in termini di compatibilità con le sfere farmacoforiche) tra le 10 presenti per ogni molecola.

Il risultato di questa primaria selezione ha permesso di ridurre la libreria iniziale di quasi 107 mila unità, portando all'identificazione di 6749 composti.

9.4 Filtro per PLIF (*Protein Ligand Interaction Fingerprint*)

Il passaggio successivo è stata la generazione delle PLIF (*Protein-Ligand Interaction Fingerprint*) per studiare nel dettaglio le interazioni che i vari composti formano con i residui del *target*. È stato applicato un criterio di selezione per i composti che nella migliore *pose* di *docking* formano un legame bidentato accettore-donatore con il residuo Leu85 presente sulla regione *hinge*.

La formazione di questo legame in particolare, come abbiamo descritto in precedenza, è presente nei ligandi di CK1 δ umana, ed è predetto anche per i due ligandi veri positivi di PfCK1 con il modello. In particolare, per rendere più severo il filtro, sono stati selezionati solo i composti in grado di formare legami a idrogeno a tenacia maggiore con quel residuo. La forza del legame a idrogeno viene stabilita da un *cutoff* regolabile liberamente a priori che in questo caso è stato fissato a 1.5 kcal/mol. La forza di un legame a idrogeno viene infatti espressa in unità di energia come kcal/mol (kilocalorie per mole). Questa è una misura dell'energia necessaria per rompere una mole di legami a idrogeno in una determinata molecola o in una sostanza.

Il risultato del filtro per PLIF ha portato alla selezione di 480 molecole.

Successivamente è stata fatta una minimizzazione delle *poses*, attraverso l'uso del tool *Dock* di MOE, con l'utilizzo della *scoring function* GBVI/WSA dG. Questo processo serve a rilassare le strutture dei composti a partire dalla *docking pose*, consentendo un avvicinamento ulteriore al minimo energetico, e minimizzando *clash* sterici.

A seguito della minimizzazione è stato ripetuto il filtro per farmacoforo, al fine da selezionare le *poses* che soddisfino le *features* descritte in precedenza, ottenendo 227 molecole.

È stato ripetuto anche il filtro per PLIF *strong*, dal quale rimangono 71 molecole.

9.5 Ulteriori filtri

Di queste 71 molecole risultanti, tramite il *tool* per il calcolo dei descrittori molecolari di MOE, sono state computate l'energia di interazione di van der Waals tra molecole e proteina, e la percentuale di superficie esposta al solvente per ogni composto.

Sono stati applicati i seguenti filtri:

1. **Ervdw < 0**, utilizzato per identificare le pose che presentino *clash* eccessivi con la proteina.
2. **Percentuale di ligando esposto al solvente (*lig_exp*) < 15**. Questa soglia è stata tarata sul valore di *lig_exp* dei due veri positivi nelle *poses* ottenute tramite *docking* sul modello proteico rifinito in dinamica. In particolare, CHEMBL4461850 risulta avere un valore di *lig_exp* pari a 13.0629, mentre CHEMBL4591417 di 14.7694. Per questo motivo il *cutoff* è stato regolato a 15.

I risultati ottenuti sono i seguenti:

1. Tutti i ligandi esaminati, come anche i due ligandi veri positivi, presentano un valore di Ervdw minore di 0. Questo significa che non sono presenti *clash* causati dalla sovrapposizione delle sfere di van der Waals.
2. Tra i 71 composti, 43 superano il filtro *lig_exp* < 15.
3. Di questi 43 è stata fatta una valutazione del peso molecolare, decidendo di dare priorità alle molecole mature piuttosto che ai frammenti; quindi, in questa fase sono state escluse molecole con peso molecolare inferiore a 300 Da.
4. Questa selezione ha portato a 11 composti finali.

10. Caratterizzazione dei composti selezionati

1. MolPort-002-054-502 (8-Bromoguanosina)⁵⁰

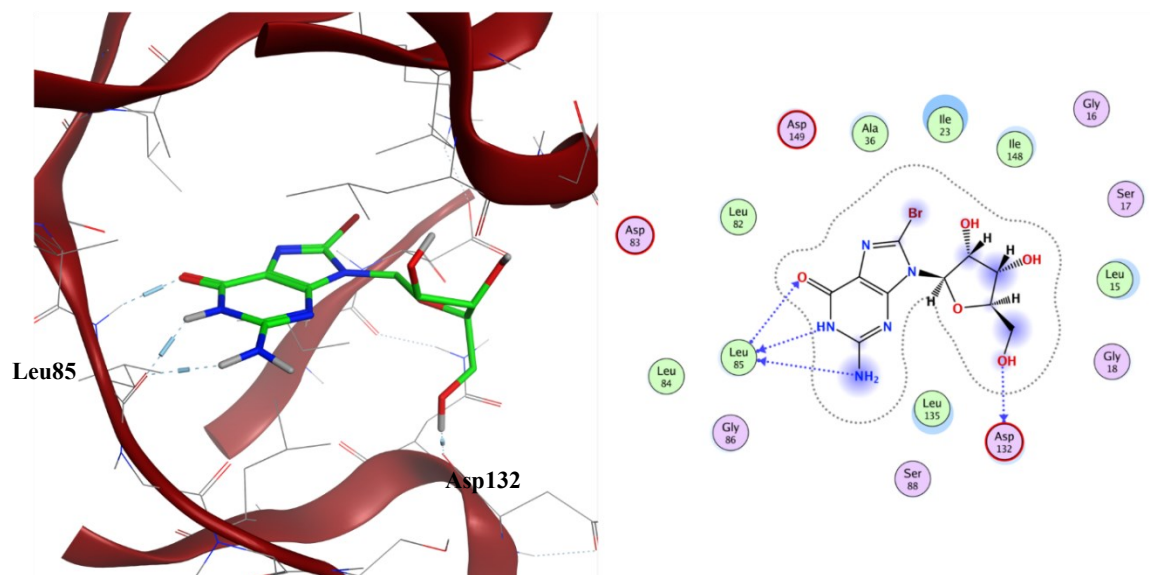


Figura 32: pose selezionata e rappresentazione 2D delle interazioni con il target di MolPort-002-054-502.

Molport ID	Molport-002-054-502
IUPAC Name	2-amino-8-bromo-9-[(2R,3R,4S,5R)-3,4-dihydroxy-5-(hydroxymethyl)oxolan-2-yl]-1H-purin-6-one
Molecular Formula	C ₁₀ H ₁₂ BrN ₅ O ₅
Canonical SMILES	C(C1C(C(C(O1)N2C3=C(C(=O)NC(=N3)N)N=C2Br)O)O)O
Molecular Weight	362.14 g/mol
XLogP3	-0.9
Hydrogen Bond Donor Count	5
Hydrogen Bond Acceptor Count	7
Rotatable Bond Count	2
Undefined Bond Stereocenter Count	0

Tabella 9: Nomenclatura e proprietà chimiche di Molport-002-054-502.

Questa molecola è un nucleoside purinico, una guanosina derivatizzata con un bromo in posizione 8 della guanina.

Nella *pose* selezionata l'ossigeno carbonilico sul *backbone* della Leu85 del *target* funge da accettore di due legami a idrogeno, il primo con uno degli idrogeni della porzione amminica legata al carbonio

in posizione 2 della guanina, e il secondo con l'ammina in posizione 1. Un terzo legame a idrogeno coinvolge l'ammina del *backbone* della Leu85, che si comporta da donatore, e l'ossigeno carbonilico legato al carbonio in posizione 6 della guanina.

Un quarto legame a idrogeno presente nella *pose* selezionata coinvolge il residuo Asp132 e l'ossidrile terminale legato al carbonio in posizione 5' del ribosio.

Questa molecola è commercialmente disponibile presso numerosi *vendors*.

Data la sua somiglianza con l'ATP (scaffold purinico), substrato endogeno delle chinasi, non stupisce il fatto che molecole con uno *scaffold* di questo tipo vengano selezionate come possibili inibitori di PfCK1.

2. MolPort-001-803-629⁵¹

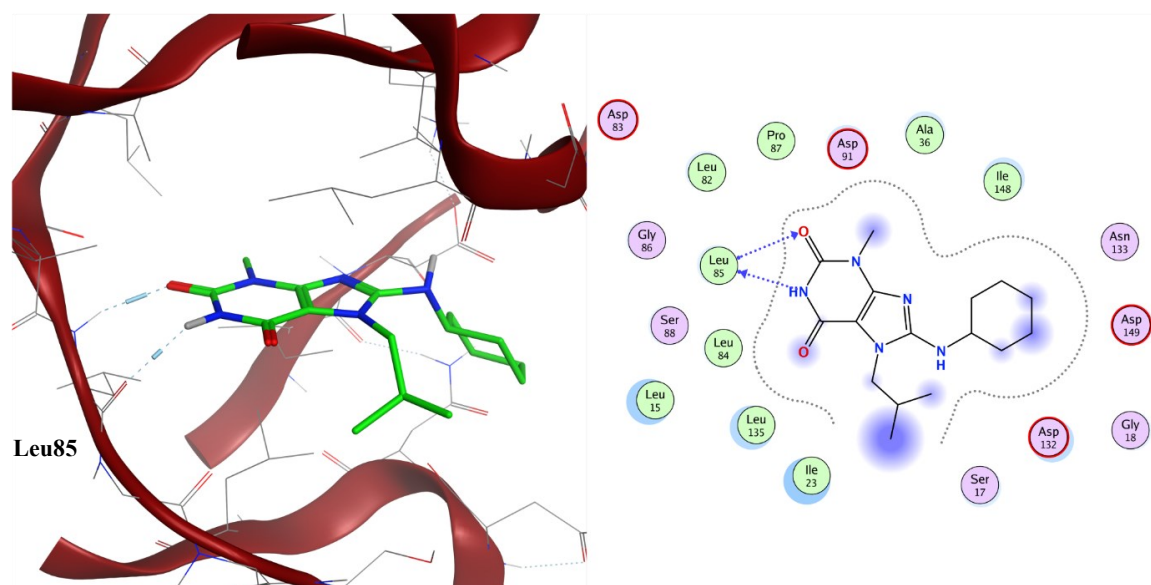


Figura 33: pose selezionata e rappresentazione 2D delle interazioni con il target di MolPort-001-803-629

Molport ID	Molport-001-803-629
IUPAC Name	8-(cyclohexylamino)-3-methyl-7-(2-methylpropyl)purine-2,6-dione
Molecular Formula	C ₁₆ H ₂₅ N ₅ O ₂
Canonical SMILES	CC(C)Cn1c(NC2CCCCC2)nc2n(C)c(=O)[nH]c(=O)c12
Molecular Weight	319.409 g/mol
XLogP3	2.8
Hydrogen Bond Donor Count	2
Hydrogen Bond Acceptor Count	4
Rotatable Bond Count	4
Undefined Bond Stereocenter Count	0

Tabella 10: nomenclatura e proprietà chimiche di Molport-001-803-629.

Questa molecola presenta uno *scaffold* principale simile a una base azotata purinica, in particolare una guanina la cui ammina esociclica legata al carbonio in posizione 2 è sostituita da un gruppo carbonilico. Si può notare inoltre la presenza di un sostituito isobutilico legato all'azoto in posizione 7, e di un sostituito legato al carbonio 8 formato da un'ammina secondaria a cui si lega un cicloesano.

Nella *pose* selezionata l'ossigeno carbonilico sul *backbone* della Leu85 del *target* si comporta da accettore di un legame a idrogeno, con l'NH in posizione 1 della purina. Il secondo legame a idrogeno

coinvolge l'NH ammidico del *backbone* della Leu85, che si comporta da donatore, e l'ossigeno carbonilico in posizione 2 della purina.

È una molecola disponibile per l'acquisto presso numerosi *vendors*.

Da quanto riportato in letteratura, l'attività di questo composto è stata studiata nei confronti di numerosi *target* umani, tra cui Ciclina A2 e CDK2 (Chinasi Ciclina-dipendente 2), e virali (sul virus dell'afta epizootica o *Foot-and-mouth disease virus*), risultando tuttavia sempre inattiva.

3. MolPort-002-527-585⁵²

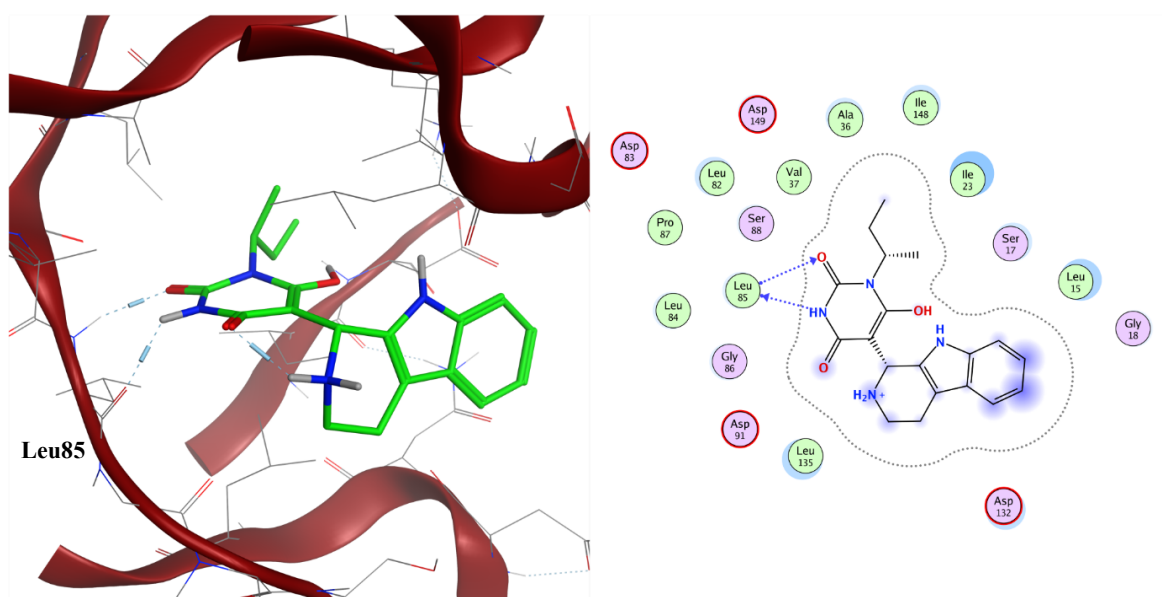


Figura 34: pose selezionata e rappresentazione 2D delle interazioni con il target di MolPort-002-527-585

Molport ID	Molport-002-527-585
IUPAC Name	1-butan-2-yl-6-hydroxy-5-(2,3,4,9-tetrahydro-1H-pyrido[3,4-b]indol-1-yl)pyrimidine-2,4-dione
Molecular Formula	C ₁₉ H ₂₂ N ₄ O ₃
Canonical SMILES	CCC(C)N1C(=C(C(=O)NC1=O)C2C3=C(CCN2)C4=CC=CC=C4N3)O
Molecular Weight	354.41 g/mol
XLogP3	1.7
Hydrogen Bond Donor Count	4
Hydrogen Bond Acceptor Count	4
Rotatable Bond Count	3
Undefined Atom Stereocenter Count	2

Tabella 11: nomenclatura e proprietà chimiche di Molport-002-057-585.

Questa molecola è formata da un tricyclo condensato, composto da un benzene, un pirrolo e una piperidina. A questo eterociclo è legata una struttura pirimidinica simile a una timina sostituita.

Nella *pose* selezionata l'ossigeno carbonilico sul *backbone* della Leu85 del *target* si comporta da accettore di un legame a idrogeno nei confronti del gruppo NH in posizione 3 della pirimidina. Il secondo legame a idrogeno coinvolge l'NH ammidico del *backbone* della Leu85, che si comporta da donatore, e l'ossigeno carbonilico legato al carbonio in posizione 2 della pirimidina.

Si tratta di una molecola disponibile per l'acquisto presso differenti venditori. La sua criticità principale nell'ottica del suo utilizzo in ambito farmacologico e per i futuri test *in vitro* è data dalla presenza di due stereocentri che portano all'esistenza di 4 differenti stereoisomeri. La stereochimica nel database Molport non è assegnata in modo univoco, perciò i vari stereoisomeri sono stati sottoposti a docking e solo uno ha passato i filtri di selezione.

4. MolPort-001-987-036⁵³

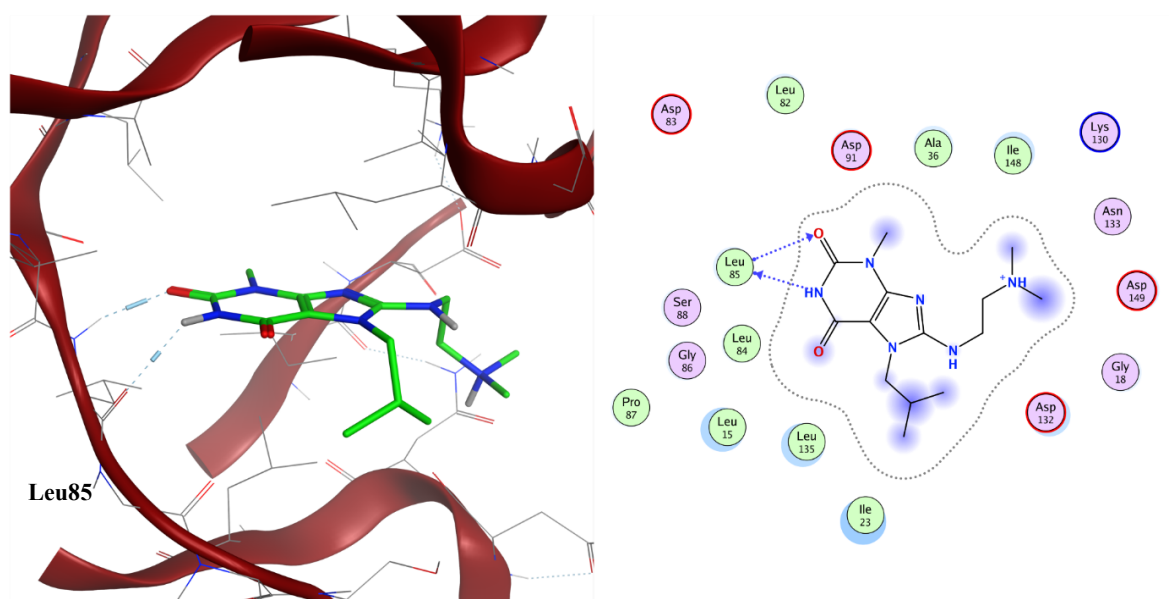


Figura 35: pose selezionata e rappresentazione 2D delle interazioni con il target di MolPort-001-987-036

Molport ID	Molport-001-987-036
IUPAC Name	8-(2-Dimethylamino-ethylamino)-7-isobutyl-3-methyl-3,7-dihydro-purine-2,6-dione
Molecular Formula	C ₁₄ H ₂₄ N ₆ O ₂
Canonical SMILES	CC(C)CN1C2=C(N=C1NCCN(C)C)N(C(=O)NC2=O)C
Molecular Weight	308.38 g/mol
XLogP3	1
Hydrogen Bond Donor Count	2
Hydrogen Bond Acceptor Count	5
Rotatable Bond Count	6
Undefined Bond Stereocenter Count	0

Tabella 12: nomenclatura e proprietà chimiche di Molport-001-987-036.

Questa molecola, molto simile alla numero 2, presenta uno *scaffold* principale purinico, in particolare una guanina la cui ammina legata al carbonio in posizione 2 è stata sostituita da un gruppo carbonilico. Si può notare inoltre la presenza di un sostituente isobutilico legato all'azoto in posizione 7, e di un sostituente legato al carbonio 8 formato da un'ammina secondaria a cui si lega una dimetil propil ammina.

Nella *pose* selezionata l'ossigeno carbonilico sul *backbone* della Leu85 del *target* si comporta da accettore di un legame a idrogeno, con l'idrogeno del NH in posizione 1 della purina. Il secondo

legame a idrogeno coinvolge l'ammina del *backbone* della Leu85, che si comporta da donatore, e l'ossigeno carbonilico legato al carbonio in posizione 2 della purina.

È una molecola disponibile per l'acquisto presso numerosi *vendors*.

Una criticità che potrebbe limitare la sua applicabilità in ambito farmacologico è data dalla presenza di 6 legami rotabili, che causano un'umentata flessibilità molecolare.

5. MolPort-002-510-879⁵⁴

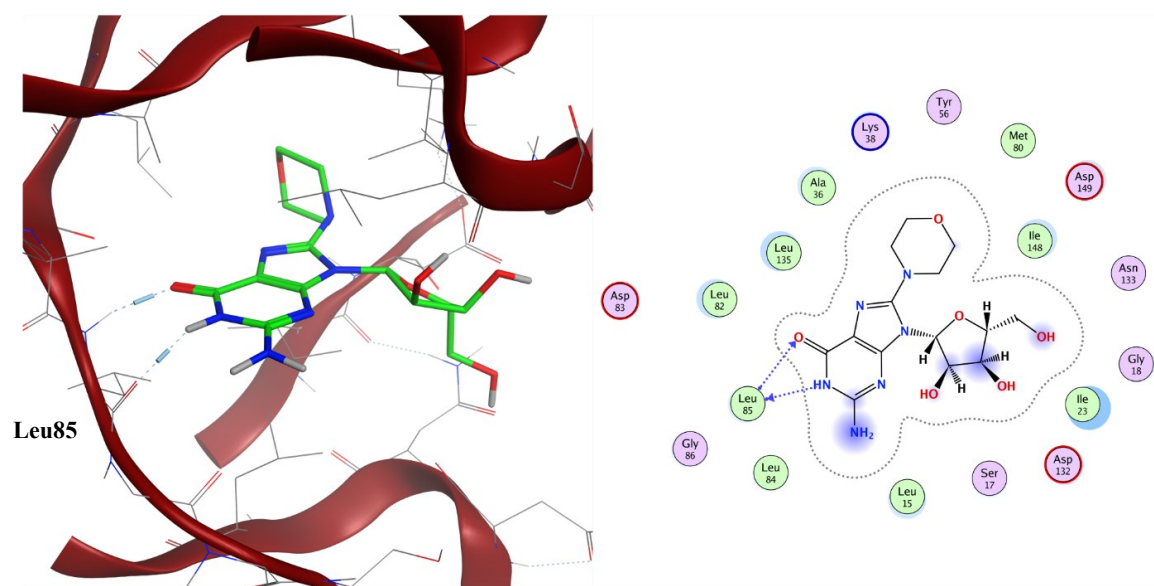


Figura 36: pose selezionata e rappresentazione 2D delle interazioni con il target di MolPort-002-510-879.

Molport ID	Molport-002-510-879
IUPAC Name	2-amino-9-[(2R,3R,4S,5R)-3,4-dihydroxy-5-(hydroxymethyl)oxolan-2-yl]-8-morpholin-4-yl-1H-purin-6-one
Molecular Formula	C ₁₄ H ₂₀ N ₆ O ₆
Canonical SMILES	C1COCCN1C2=NC3=C(N2C4C(C(C(O4)CO)O)O)N=C(NC3=O)N
Molecular Weight	368.35
XLogP3	-1.8
Hydrogen Bond Donor Count	5
Hydrogen Bond Acceptor Count	9
Rotatable Bond Count	3
Undefined Atom Stereocenter Count	0

Tabella 13: nomenclatura e proprietà chimiche di Molport-002-510-879.

Questa molecola presenta uno *scaffold* purinico, una guanosina sostituita in posizione 8 con un gruppo morfolinico.

Nella *pose* selezionata l'ossigeno carbonilico sul *backbone* della Leu85 del *target* funge da accettore di un legame a idrogeno con l'NH in posizione 1. Un secondo legame a idrogeno coinvolge l'NH ammidico del *backbone* della Leu85, che si comporta da donatore, e l'ossigeno carbonilico legato al carbonio in posizione 6 della guanina.

È una molecola disponibile per l'acquisto presso numerosi *vendors*.

Da quanto riportato in letteratura, è stata testata come inibitore della proliferazione su alcune linee cellulari tumorali umane, risultando sempre inattiva.

6. MolPort-002-535-751⁵⁵

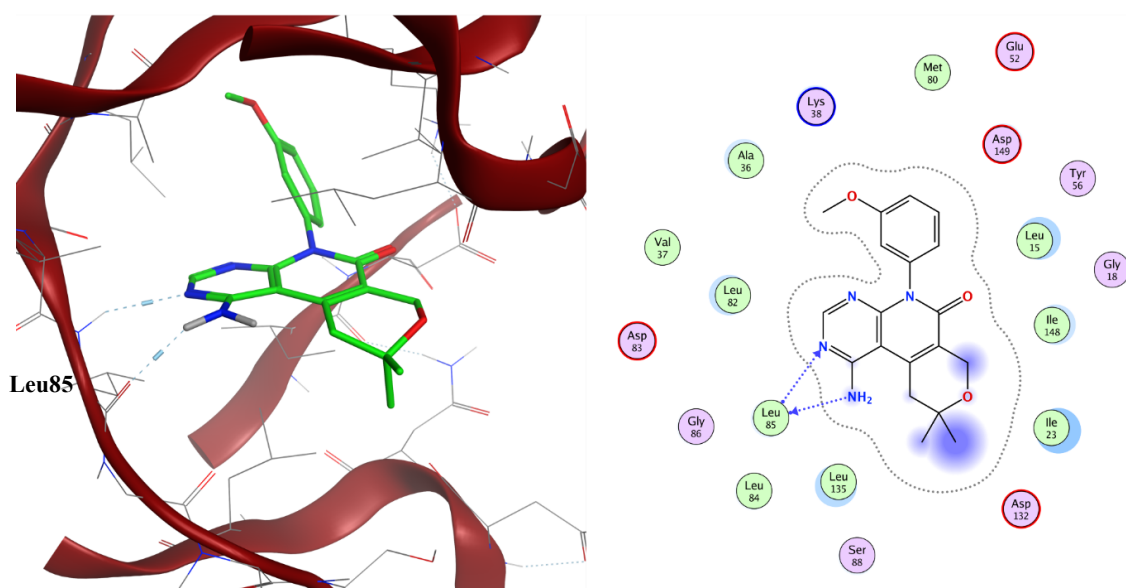


Figura 37: pose selezionata e rappresentazione 2D delle interazioni con il target di MolPort-002-535-751.

Molport ID	Molport-002-535-751
IUPAC Name	3-amino-8-(3-methoxyphenyl)-13,13-dimethyl-12-oxa-4,6,8-triazatricyclo[8.4.0.02,7]tetradeca-1(10),2,4,6-tetraen-9-one
Molecular Formula	C ₁₉ H ₂₀ N ₄ O ₃
Canonical SMILES	CC1(CC2=C(CO1)C(=O)N(C3=NC=NC(=C2)3)N)C4=CC(=CC=C4)OC)C
Molecular Weight	352.394
XLogP3	1.2
Hydrogen Bond Donor Count	1
Hydrogen Bond Acceptor Count	6
Rotatable Bond Count	2
Undefined Bond Stereocenter Count	0

Tabella 14: nomenclatura e proprietà chimiche di Molport-002-535-751

Questa molecola presenta uno *scaffold* pirimido-isochinolinico, con un sostituito fenilico metossilato. Nella *pose* selezionata l'ossigeno carbonilico sul *backbone* della Leu85 del *target* funge da accettore di un legame a idrogeno con l'ammina esociclica. Come si vede in figura, è presente anche un secondo legame a idrogeno in cui l'NH ammidico del *backbone* funge da donatore nei confronti dell'azoto sul ciclo aromatico.

È una molecola disponibile per l'acquisto presso numerosi *vendors*.

Da quanto riportato in letteratura, è emerso un dato di attività di questo composto come inibitore della tirosinchinasi EphB4. Lo studio è stato fatto mediante *docking* di una libreria, seguito da simulazioni di dinamica molecolare a solvente esplicito per la valutazione della modalità di legame. Il dato di attività è stato ottenuto attraverso un saggio in cui viene valutato lo spiazzamento di ATP radiomarcato da parte di questo inibitore. È stato ottenuto un valore di IC_{50} pari a $14.3\mu M$ ⁵⁶.

7. MolPort-002-529-576⁵⁷

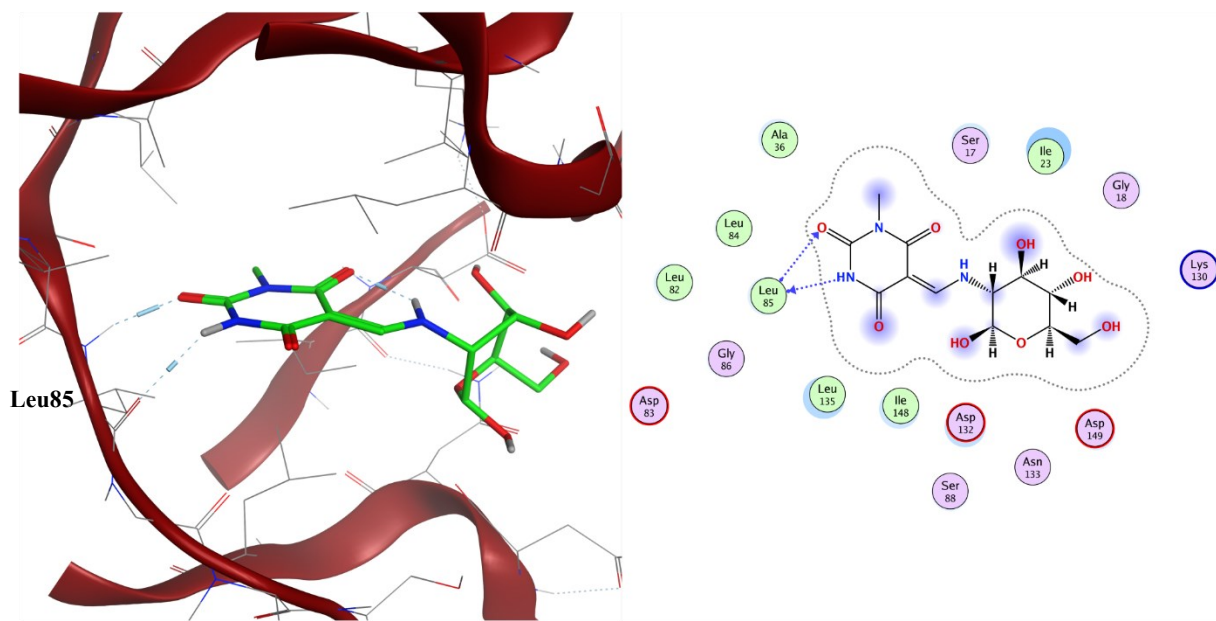


Figura 38: pose selezionata e rappresentazione 2D delle interazioni con il target di MolPort-002-529-576

Molport ID	Molport-002-529-576
IUPAC Name	6-hydroxy-1-methyl-5-[[[(2R,3R,4R,5S,6R)-2,4,5-trihydroxy-6-(hydroxymethyl)oxan-3-yl]iminomethyl]pyrimidine-2,4-dione
Molecular Formula	C ₁₂ H ₁₇ N ₃ O ₈
Canonical SMILES	CN1C(=C(C(=O)NC1=O)C=NC2C(C(C(OC2O)CO)O)O)O
Molecular Weight	331.28 g/mol
XLogP3	-3.3
Hydrogen Bond Donor Count	6
Hydrogen Bond Acceptor Count	9
Rotatable Bond Count	3
Undefined Bond Stereocenter Count	0

Tabella 15: nomenclatura e proprietà chimiche di Molport-002-529-576

Questa molecola presenta uno *scaffold* pirimidinico, legato al carbonio in posizione 2 di un glucosio tramite uno spaziatore amminico. La pirimidina forma il legame bidentato con la Leu85, attraverso un pattern di interazione già visto nella caratterizzazione delle molecole precedenti.

È una molecola disponibile per l'acquisto, ma ad oggi non sono presenti dati di attività derivanti da studi su questa molecola.

8. MolPort-007-825-628⁵⁸

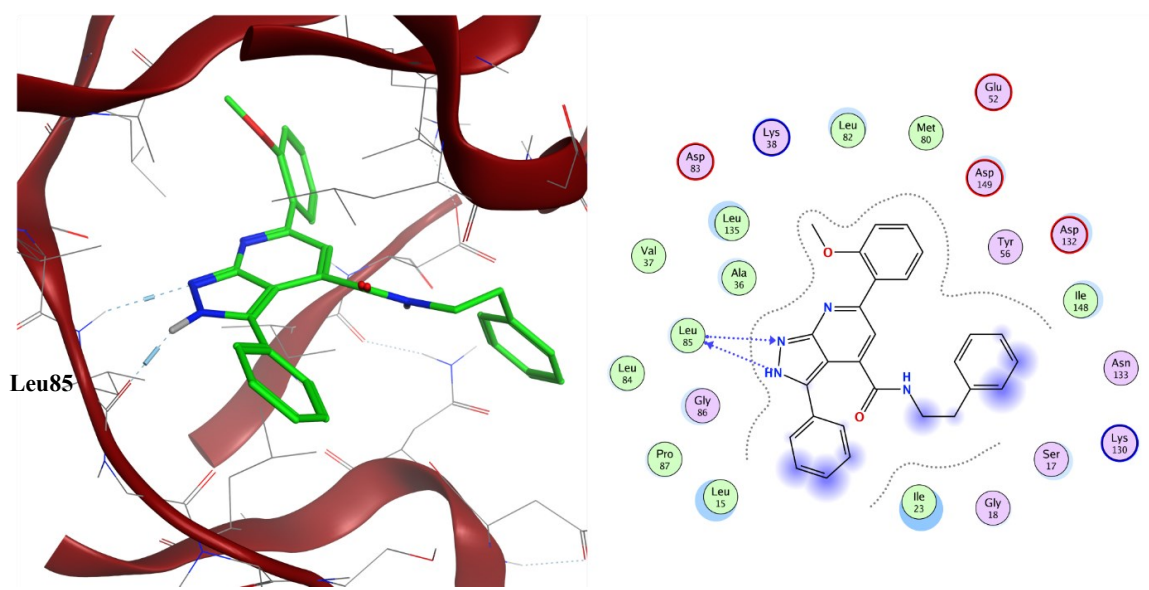


Figura 39: pose selezionata e rappresentazione 2D delle interazioni con il target di MolPort-007-825-628

Molport ID	Molport-007-825-628
IUPAC Name	6-(2-methoxyphenyl)-3-phenyl-N-(2-phenylethyl)-2H-pyrazolo[3,4-b]pyridine-4-carboxamide
Molecular Formula	C ₂₈ H ₂₄ N ₄ O ₂
Canonical SMILES	<chem>COC1=CC=CC=C1C2=NC3=NNC(=C3C(=C2)C(=O)NCCC4=CC=CC=C4)C5=CC=CC=C5</chem>
Molecular Weight	448.5 g/mol
XLogP3	5.3
Hydrogen Bond Donor Count	2
Hydrogen Bond Acceptor Count	4
Rotatable Bond Count	7
Undefined Bond Stereocenter Count	0

Tabella 16: nomenclatura e proprietà chimiche di Molport-007-825-628

Questa molecola è caratterizzata da uno *scaffold* pirazolo-piridinico, che interagisce con la Leu85 mediante la formazione di due legami a idrogeno. Il primo vede il gruppo ossidrilico del *backbone* del residuo amminoacidico fungere da accettore nei confronti dell'idrogeno dell'ammina in posizione 7 della molecola. Il secondo legame a idrogeno coinvolge l'azoto in posizione 8, che funge da accettore verso l'NH ammidico del *backbone* amminoacidico.

Questa molecola è disponibile per l'acquisto e, da quanto riportato in letteratura, è stata studiata la sua attività *in vitro* nei confronti di *Schistosoma mansoni*. Non vengono riportati gli esiti di questo test.

9. MolPort-035-700-089⁵⁹

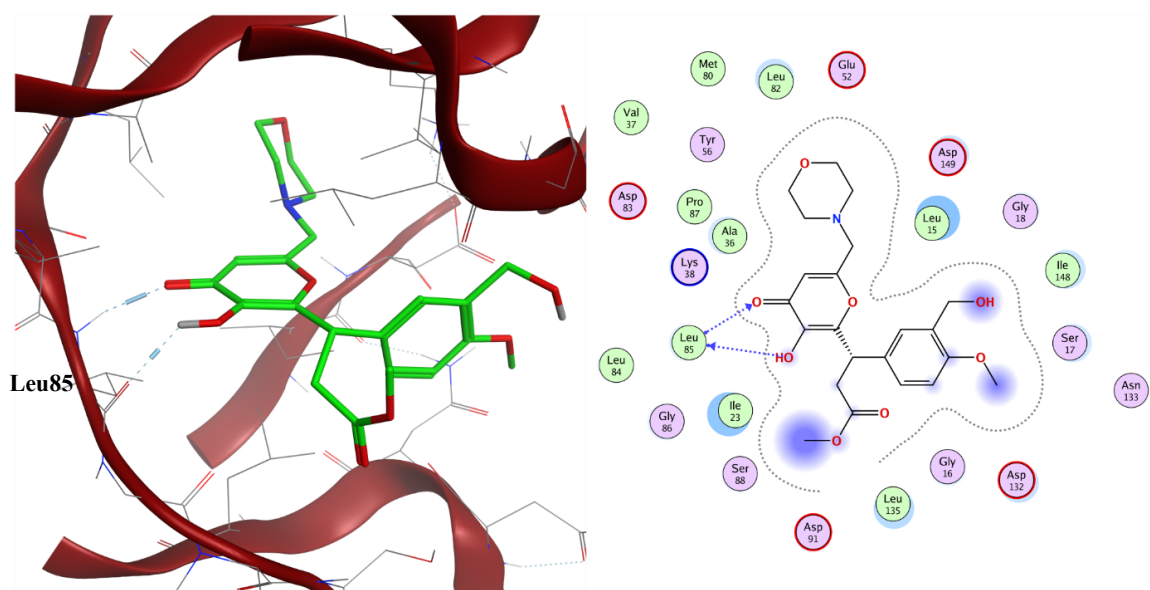


Figura 40: pose selezionata e rappresentazione 2D delle interazioni con il target di MolPort-035-700-089

Molport ID	Molport-035-700-089
IUPAC Name	Methyl3-[3-(hydroxymethyl)-4-methoxyphenyl]-3-[3-hydroxy-6-(morpholin-4-ylmethyl)-4-oxopyran-2-yl]propanoate
Molecular Formula	C ₂₂ H ₂₇ N ₃ O ₈
Canonical SMILES	<chem>COC1=C(C=C(C=C1)C(CC(=O)OC)C2=C(C(=O)C=C(O2)CN3CCOCC3)O)CO</chem>
Molecular Weight	433.5 g/mol
XLogP3	0.1
Hydrogen Bond Donor Count	2
Hydrogen Bond Acceptor Count	9
Rotatable Bond Count	9
Undefined Bond Stereocenter Count	1

Tabella 17: nomenclatura e proprietà chimiche di Molport-035-700-089

Questa molecola presenta una struttura molto complessa, con due anelli aromatici: un benzene e un pirano con un carbonile in posizione 4, variamente sostituiti con ossidrili, un metossile e un gruppo estereo. È presente inoltre un anello morfolinico.

Le interazioni che il composto forma con la Leu85 della proteina coinvolgono la porzione piranica della molecola. In particolare, l'ossigeno carbonilico del *backbone* di Leu85 funge da accettore di legame a idrogeno nei confronti dell'ossidrile in posizione 3 del piranone, mentre l'NH ammidico

del *backbone* di Leu85 funge da donatore di ponte a idrogeno nei confronti dell'ossigeno carbonilico in posizione 4 del piranone.

Questa molecola presenta tre criticità che potrebbero limitare il suo utilizzo in ambito farmaceutico: la prima è la presenza di uno stereocentro (non unicamente assegnato in Molport) che porta all'esistenza di 2 differenti stereoisomeri, la seconda consiste nel numero di legami rotabili (9) che la rende particolarmente flessibile, e la terza vede la presenza di un estere facilmente idrolizzabile. Una parte dei legami rotabili e la funzione esterea si trovano però alle estremità della molecola, in regioni che non vengono coinvolte nel legame con il *target*.

È una molecola disponibile per l'acquisto e ad oggi non sono presenti dati sulla sua attività in letteratura.

10. MolPort-029-884-875⁶⁰

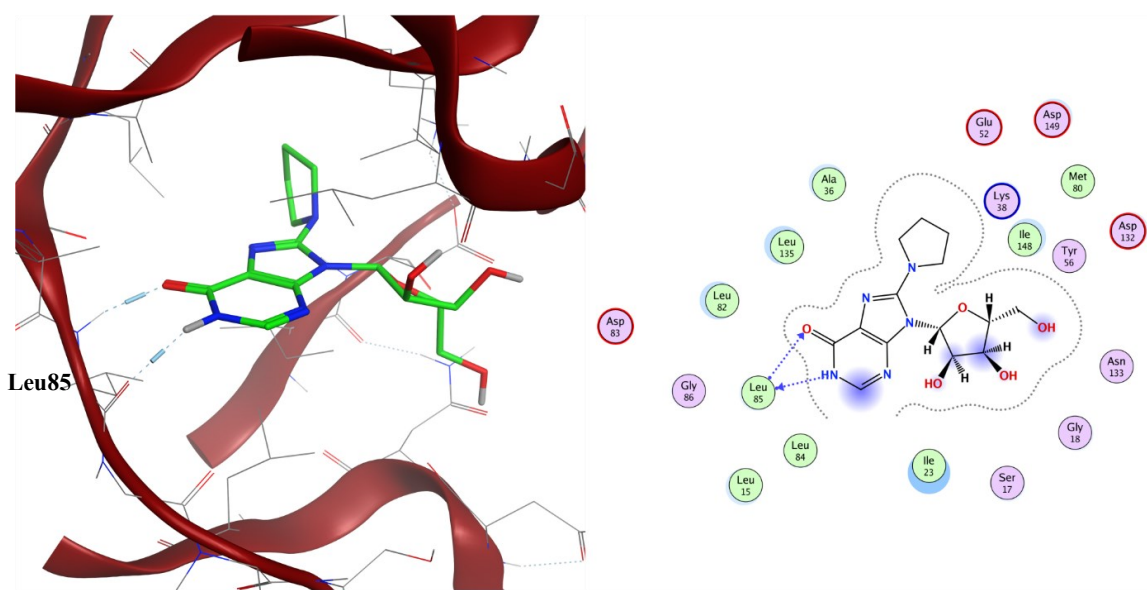


Figura 41: pose selezionata e rappresentazione 2D delle interazioni con il target di MolPort-029-884-875

Molport ID	MolPort-029-884-875
IUPAC Name	9-[(2R,3R,4S,5R)-3,4-dihydroxy-5-(hydroxymethyl)oxolan-2-yl]-8-pyrrolidin-1-yl-1H-purin-6-one
Molecular Formula	C ₁₄ H ₁₉ N ₅ O ₅
Canonical SMILES	C1CCN(C1)C2=NC3=C(N2C4C(C(C(O4)CO)O)O)N=CNC3=O
Molecular Weight	337.33 g/mol
XLogP3	-1.2
Hydrogen Bond Donor Count	4
Hydrogen Bond Acceptor Count	8
Rotatable Bond Count	3
Undefined Bond Stereocenter Count	0

Tabella 18: nomenclatura e proprietà chimiche di Molport-029-884-875

Questa molecola presenta uno *scaffold* purinico con un carbonile in posizione 6, a cui sono legati una pirrolidina in posizione 8 e un ribosio in posizione 9.

Nella *pose* selezionata l'ossigeno carbonilico sul *backbone* della Leu85 del *target* funge da accettore di legame a idrogeno con l'NH in posizione 1, e l'NH ammidico del *backbone* della Leu85 si comporta da donatore di legame a idrogeno verso l'ossigeno carbonilico legato al carbonio in posizione 6 della purina.

È una molecola disponibile per l'acquisto e ad oggi non sono presenti dati sulla sua attività in letteratura.

11. MolPort-005-911-320⁶¹

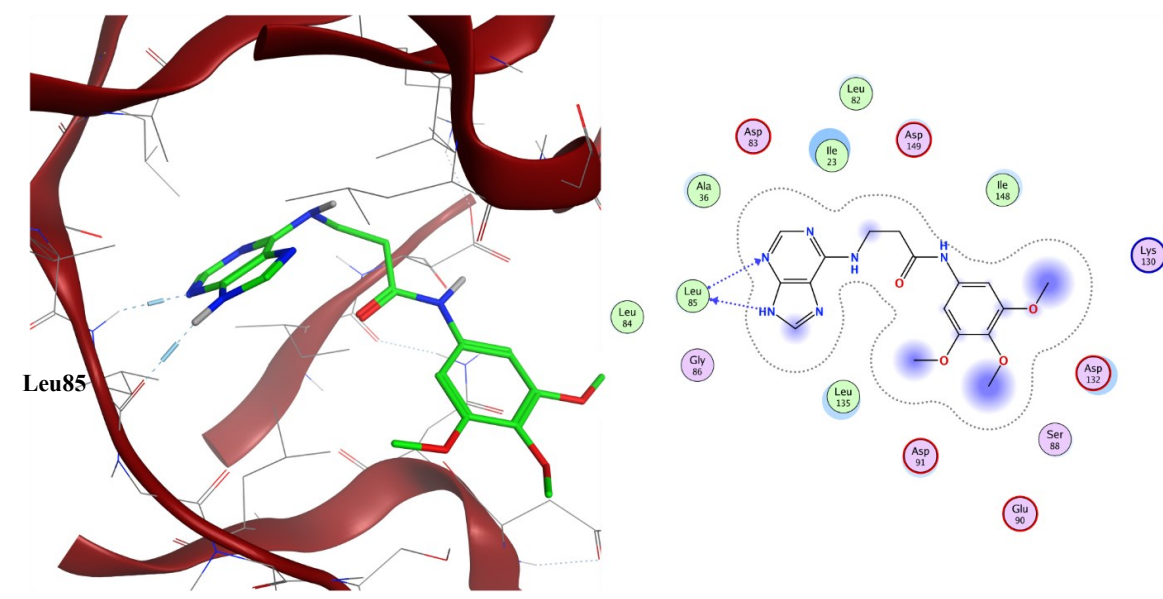


Figura 42: pose selezionata e rappresentazione 2D delle interazioni con il target di MolPort-005-911-320

Molport ID	MolPort-005-911-320
IUPAC Name	3-(7H-purin-6-ylamino)-N-(3,4,5-trimethoxyphenyl)propanamide
Molecular Formula	C ₁₇ H ₂₀ N ₆ O ₄
Canonical SMILES	<chem>COC1=CC(=CC(=C1OC)OC)NC(=O)CCNC2=NC=NC3=C2NC=N3</chem>
Molecular Weight	372.4 g/mol
XLogP3	1.7
Hydrogen Bond Donor Count	3
Hydrogen Bond Acceptor Count	8
Rotatable Bond Count	8
Undefined Bond Stereocenter Count	0

Tabella 19: nomenclatura e proprietà chimiche di Molport-005-911-320

Questo composto presenta uno *scaffold* purinico sostituito in posizione 1 con una catena idrocarburica che contiene un'ammina, un'ammide e termina con un benzene trisostituito da tre gruppi ossimetilici. La purina interagisce con il *target* mediante la formazione di due legami a idrogeno. Il primo coinvolge l'ossigeno carbonilico del *backbone* di Leu85 (che funge da accettore) e l'NH in posizione 9 della purina. Il secondo coinvolge l'NH ammidico del *backbone* di Leu85 che funge da donatore nei confronti dell'azoto in posizione 3 della purina.

Una limitazione dell'applicazione farmacologica di questa molecola è data dal grande numero di legami rotabili (8) che, come discusso in precedenza, conferiscono un'eccessiva flessibilità alla molecola.

È una molecola disponibile per l'acquisto e ad oggi non sono presenti dati sulla sua attività in letteratura.

11. Indagine sulla selettività dei composti selezionati nei confronti di CK1 δ umana

È stata svolta infine un'analisi al fine di determinare se i composti selezionati potessero mostrare un profilo di selettività *in silico* nei confronti di PfCK1 rispetto a CK1 δ umana.

Residui mutati nel sito attivo

In Figura 43 sono mostrate le sequenze amminoacidiche della CK1 δ umana (indicata nell'immagine come 6RCG.A), e della PfCK1 (indicata come last_frame).

In verde e in rosso sono selezionati i residui che fanno parte della tasca di legame delle due proteine.

In particolare, in verde sono evidenziati i residui del *pocket* che vengono conservati tra i due omologhi, mentre in rosso i residui mutati. Come si può notare gli amminoacidi mutati facenti parte della regione deputata al legame sono cinque, i seguenti:

1. Arg13 in CK1 δ viene sostituita da Lys13 in PfCK1.
2. Ile15 in CK1 δ viene sostituita da Leu15 in PfCK1.
3. Leu25 in CK1 δ viene sostituita da Val25 in PfCK1.
4. Met82 in CK1 δ viene sostituita da Leu82 in PfCK1. Questa mutazione ha una rilevanza particolare in quanto l'amminoacido mutato è il residuo *gatekeeper*.

Il *gatekeeper* è il residuo che ha il compito di regolare l'accesso delle molecole all'interno della tasca. Questo residuo può avere un'influenza significativa sulla forma tridimensionale del *binding site*, modulando la sua accessibilità.

5. Glu83 in CK1 δ viene sostituita da Asp83 in PfCK1.



Figura 43: allineamento tra le sequenze amminoacidiche di CK1δ umana (PDBID: 6RCG) e PfCK1. In verde sono evidenziati i residui del pocket che vengono conservati tra i due omologi, mentre in rosso i residui mutati.

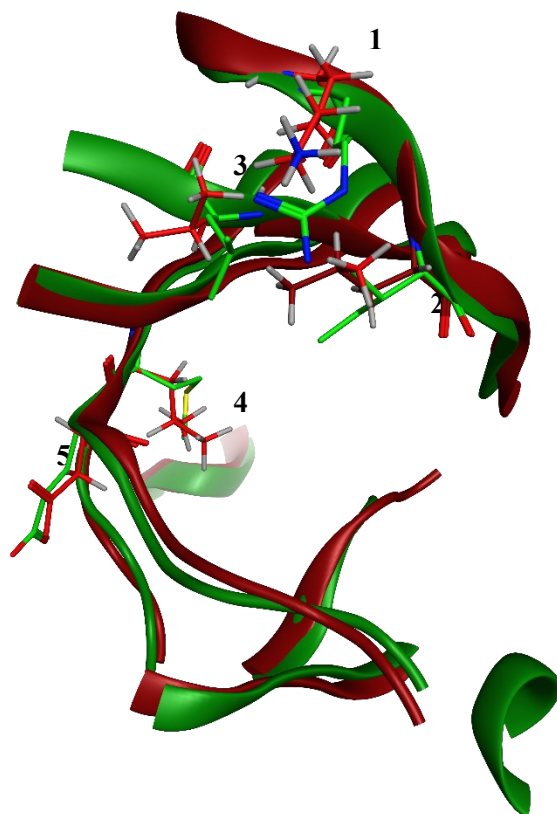


Figura 44: sovrapposizione tra la struttura cristallografica di 6RCG (in verde) e il modello di omologia di PfCK1 (in rosso). Vengono mostrate le catene laterali degli amminoacidi mutati tra le due proteine.

La struttura del *target* umano considerato per effettuare l'indagine sulla selettività dei ligandi selezionati è la proteina con PDB_{ID} 6RCG, la stessa utilizzata per generare il modello di omologia, al fine da minimizzare le differenze dipendenti da strutture cristallografiche alternative.

Per portare avanti questa caratterizzazione è stato svolto per prima cosa un *docking* delle 11 molecole selezionate e descritte nel capitolo precedente sulla struttura di CK1 δ umana. Il centro della sfera relativa alla ricerca conformazionale è stato definito come il centro di massa del ligando cristallografico della struttura 6RCG.

Dopo aver ottenuto l'*output* della procedura di *docking*, è stato verificato se fosse possibile ottenere *pose* che garantissero il pattern di interazione discusso in precedenza anche in CK1 δ umana. Le pose sono state perciò sottoposte al filtro farmacoforico, e solamente 8 molecole su 11 lo hanno superato.

Analizzando le interazioni che ognuno di questi ligandi è in grado di formare con la proteina, attraverso la generazione delle PLIF (Protein Ligand Interaction Fingerprint), risulta che nessuna molecola forma il legame a idrogeno bidentato con la Leu85, di conseguenza nessuna delle molecole viene selezionata.

Il mancato ottenimento di pose significative nel caso di CK1 δ umana potrebbe suggerire una selettività dei composti identificati, ma data l'elevata identità di sequenza tra la proteina umana CK1 δ

(in particolare il cristallo con PDB_{ID} 6RCG) e l'omologa PfCK1, in particolare nella loro tasca di legame, una selettività sarebbe inaspettata.

Per evitare che ci siano differenze dovute al posizionamento da parte dell'algoritmo di ricerca conformazionale stocastica del docking o al fatto che il modello di PfCK1 sia stato rilassato in dinamica, mentre il cristallo di CK1 δ no, si è indagato l'impatto delle diverse mutazioni puntiformi che differenziano i *binding site* delle due proteine sull'ottenimento di *binding mode* promettenti su quei composti.

Sono state quindi introdotte le 5 mutazioni puntiformi degli amminoacidi del *pocket* che differenziano il *sito di legame* della proteina PfCK1 da quello della proteina omologa umana, mostrate in *Figura 44*. Vengono minimizzate le catene laterali di questi 5 amminoacidi utilizzando Amber14EHT. Si assume quindi che le differenze siano solo in quei 5 residui, mentre tutto il resto rimanga invariato.

Facendo questo, si ottiene una proteina con struttura tridimensionale identica al modello del PfCK1, in cui l'unica differenza è introdotta dalla mutazione dei 5 residui.

A questo punto, le *poses* degli 11 composti selezionati dal *virtual screening* fatto sul modello di PfCK1 vengono minimizzate *in place* con il *tool Dock* di MOE, attraverso l'utilizzo della *scoring function* GBVI/WSA dG. Questo *step* permette di osservare se le stesse *poses* possano avere un accomodamento nella tasca in cui l'unica variazione è data dalle mutazioni che discriminano CK1 δ da PfCK1.

Viene ripetuto a questo punto il filtro tramite farmacoforo e PLIF (*Protein Ligand Interaction Fingerprint*). Le molecole che superano questi filtri sono 9 su 11.

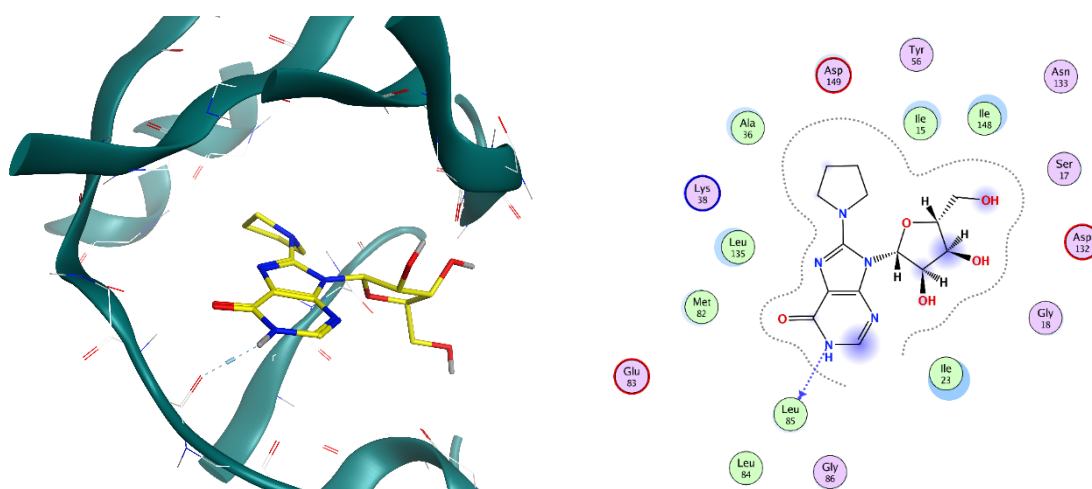


Figura 45: A sinistra, viene mostrata la pose del composto MolPort-029-884-875 nella tasca di binding di PfCK1 mutata, a destra le interazioni con il ligando in 2D.

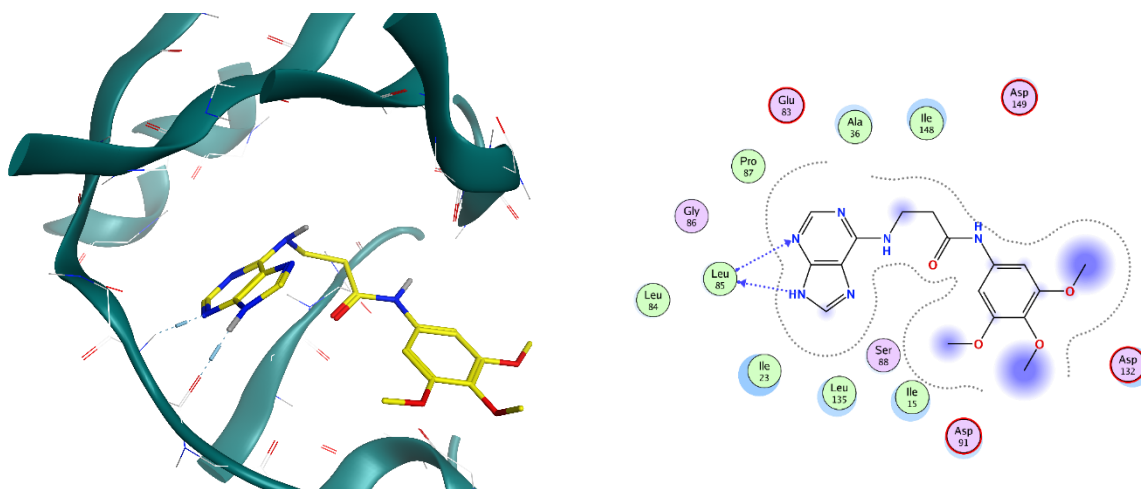


Figura 46: a sinistra, la pose del composto MolPort -005-911-320 nella tasca di binding di PfCK1 mutata, a destra vengono mostrate le interazioni con il ligando in 2D.

Le molecole che non passano i filtri per farmacoforo e per PLIF nel modello mutato sono la decima e l'undicesima molecola dell'elenco dei composti selezionati, **MolPort-029-884-875** e **MolPort-005-911-320**. La causa del loro fallimento risiede nella formazione di un solo legame a idrogeno nei confronti della Leu85 della proteina per la prima molecola (mancata formazione del legame a idrogeno bidentato), e nell'orientazione non perfettamente allineata al modello farmacoforico per la seconda.

Per quanto la geometria di queste due *poses* non sia ottimale, il *pattern* di interazione non è così distante da quello ideale, e quindi è difficile sostenere la tesi che ci sia una selettività di questi composti nei confronti di PfCK1 rispetto all'omologo umano.

12. Conclusione

Questo studio si proponeva di elaborare un protocollo volto a identificare, all'interno di un database di composti naturali, quelli capaci di stabilire interazioni promettenti con la proteina Casein chinasi 1 di *Plasmodium falciparum* (PfCK1), come potenziali inibitori dell'enzima.

La mancanza di una struttura sperimentalmente risolta della proteina di interesse è stata superata mediante la creazione di un modello di omologia, che è stato generato utilizzando la proteina CK1 δ umana come *template*, data la sua elevata identità di sequenza con PfCK1 (69%). Il modello è stato ottimizzato e validato grazie a simulazioni di *docking* e *dinamica molecolare*, utilizzando dei ligandi veri positivi.

Il modello ottimizzato è stato utilizzato per effettuare un *virtual screening* di una libreria di composti naturali e analoghi commercialmente disponibili, tramite *docking* molecolare. L'obiettivo era quello di prioritizzare *in silico* un gruppo di composti presenti nella libreria, selezionando quelli che mostravano un pattern d'interazione promettente con la proteina. Sono stati in questo modo identificati e scelti 11 composti. Essi mostrano un *pattern* di interazioni *in silico* analogo a quello che accomuna e contraddistingue i ligandi di CK1 δ umana, lo stesso *pattern* che viene predetto *in silico* anche per i ligandi veri positivi di PfCK1. Una indagine *in silico* sulla selettività di questi composti verso PfCK1 rispetto a CK1 δ umana non è conclusiva, ma non ha dato risultati incoraggianti, come era prevedibile data l'elevata identità di sequenza delle due proteine. Tuttavia, l'indagine del profilo di selettività è prematura in questa fase del progetto, che mira all'identificazione di *hit compounds* da validare sperimentalmente. Successive fasi di ottimizzazione potranno eventualmente essere volte al miglioramento della selettività, come di altre proprietà farmacodinamiche e farmacocinetiche, delle molecole qualora risultassero essere *hit compounds* sperimentalmente validati.

Gli 11 composti selezionati verranno acquistati dall'Università di Sassari per essere sottoposti a test *in vitro* al fine di valutarne l'attività su *Plasmodium falciparum* tramite test sperimentali.

Bibliografia

1. Pappas, G., Kiriaze, I. J. & Falagas, M. E. Insights into infectious disease in the era of Hippocrates. *International Journal of Infectious Diseases* **12**, 347–350 (2008).
2. Boualam, M. A., Pradines, B., Drancourt, M. & Barbieri, R. Malaria in Europe: A Historical Perspective. *Front Med (Lausanne)* **8**, 691095 (2021).
3. Cox, F. E. History of the discovery of the malaria parasites and their vectors. *Parasit Vectors* **3**, 1–9 (2010).
4. Malaria Atlas Project | Home. <https://malariaatlas.org/>.
5. World malaria report 2022. <https://www.who.int/publications/i/item/9789240064898>.
6. Bannister, L. H., Hopkins, J. M., Fowler, R. E., Krishna, S. & Mitchell, G. H. A Brief Illustrated Guide to the Ultrastructure of Plasmodium falciparum Asexual Blood Stages. *Parasitology Today* **16**, 427–433 (2000).
7. Nankabirwa, J. *et al.* Malaria in school-age children in Africa: an increasingly important challenge. *Tropical Medicine & International Health* **19**, 1294–1309 (2014).
8. Hess, U., Timmermans, P. M. & Jones, M. Combined Chloroquine/Fansidar®-Resistant Falciparum Malaria Appears in East Africa. *Am J Trop Med Hyg* **32**, 217–220 (1983).
9. Eriksen, J. *et al.* Antimalarial resistance and DHFR/DHPS genotypes of Plasmodium falciparum three years after introduction of sulfadoxine-pyrimethamine and amodiaquine in rural Tanzania. *Trans R Soc Trop Med Hyg* **102**, 137–142 (2008).
10. Hurwitz, E. S., Johnson, D. & Campbell, C. C. RESISTANCE OF PLASMODIUM FALCIPARUM MALARIA TO SULFADOXINE-PYRIMETHAMINE ('FANSIDAR') IN A REFUGEE CAMP IN THAILAND. *The Lancet* **317**, 1068–1070 (1981).
11. Oostburg, B. F. J. & Lozefzoon, L. M. E. Fansidar resistant Plasmodium falciparum infection in Surinam. *Trop Geogr Med* **35**, 243–247 (1983).
12. Hyde, J. E. Mechanisms of resistance of Plasmodium falciparum to antimalarial drugs. *Microbes Infect* **4**, 165–174 (2002).
13. Li, Z., Li, Q., Wu, J., Wang, M. & Yu, J. Artemisinin and Its Derivatives as a Repurposing Anticancer Agent: What Else Do We Need to Do? *Molecules* 2016, Vol. 21, Page 1331 **21**, 1331 (2016).

14. Jahan, M. *et al.* Structure–Activity Relationships of the Antimalarial Agent Artemisinin 10. Synthesis and Antimalarial Activity of Enantiomers of rac-5 β -Hydroxy-d-Secoartemisinin and Analogs: Implications Regarding the Mechanism of Action. *Molecules* 2021, Vol. 26, Page 4163 **26**, 4163 (2021).
15. Ridley, R. G. Medical need, scientific opportunity and the drive for antimalarial drugs. *Nature* 2002 415:6872 **415**, 686–693 (2002).
16. Noedl, H. *et al.* Evidence of Artemisinin-Resistant Malaria in Western Cambodia. *New England Journal of Medicine* **359**, 2619–2620 (2008).
17. Dondorp, A. M. *et al.* Artemisinin Resistance in Plasmodium falciparum Malaria . *New England Journal of Medicine* **361**, 455–467 (2009).
18. Williams, K. L. *et al.* A candidate antibody drug for prevention of malaria. *Nature Medicine* 2024 30:1 **30**, 117–129 (2024).
19. Dicko, A. *et al.* Seasonal vaccination with RTS,S/AS01E vaccine with or without seasonal malaria chemoprevention in children up to the age of 5 years in Burkina Faso and Mali: a double-blind, randomised, controlled, phase 3 trial. *Lancet Infect Dis* **24**, 75–86 (2024).
20. Adderley, J. & Doerig, C. Comparative analysis of the kinomes of Plasmodium falciparum, Plasmodium vivax and their host Homo sapiens. *BMC Genomics* **23**, 1–15 (2022).
21. Ong, H. W., Adderley, J., Tobin, A. B., Drewry, D. H. & Doerig, C. Parasite and host kinases as targets for antimalarials. *Expert Opin Ther Targets* **27**, 151–169 (2023).
22. Batty, M. B., Schittenhelm, R. B., Dorin-Semblat, D., Doerig, C. & Garcia-Bustos, J. F. Interaction of Plasmodium falciparum casein kinase 1 with components of host cell protein trafficking machinery. *IUBMB Life* **72**, 1243–1249 (2020).
23. Barik, S., Taylor, R. E. & Chakrabarti, D. Identification, Cloning, and Mutational Analysis of the Casein Kinase 1 cDNA of the Malaria Parasite, Plasmodium falciparum: STAGE-SPECIFIC EXPRESSION OF THE GENE *. *Journal of Biological Chemistry* **272**, 26132–26138 (1997).
24. Dorin-Semblat, D. *et al.* Malaria Parasite-Infected Erythrocytes Secrete PfCK1, the Plasmodium Homologue of the Pleiotropic Protein Kinase Casein Kinase 1. *PLoS One* **10**, e0139591 (2015).
25. Gross, S. D. & Anderson, R. A. Casein Kinase I: Spatial Organization and Positioning of a Multifunctional Protein Kinase Family. *Cell Signal* **10**, 699–711 (1998).

26. Longenecker, K. L., Roach, P. J. & Hurley, T. D. Three-dimensional Structure of Mammalian Casein Kinase I: Molecular Basis for Phosphate Recognition. *J Mol Biol* **257**, 618–631 (1996).
27. Molecular Operating Environment (MOE) | MOEsaic | PSILO. <https://www.chemcomp.com/Products.htm>.
28. Greer, Y. E., Gao, B., Yang, Y., Nussenzweig, A. & Rubin, J. S. Lack of Casein Kinase 1 Delta Promotes Genomic Instability - The Accumulation of DNA Damage and Down-Regulation of Checkpoint Kinase 1. *PLoS One* **12**, e0170903 (2017).
29. Greer, Y. E. *et al.* Casein kinase 1 δ functions at the centrosome and Golgi to promote ciliogenesis. *Mol Biol Cell* **25**, 1629–1640 (2014).
30. Johnson, A. E., Chen, J. S. & Gould, K. L. CK1 is required for a mitotic checkpoint that delays cytokinesis. *Current Biology* **23**, 1920–1926 (2013).
31. Phadnis, N. *et al.* Casein Kinase 1 and Phosphorylation of Cohesin Subunit Rec11 (SA3) Promote Meiotic Recombination through Linear Element Formation. *PLoS Genet* **11**, e1005225 (2015).
32. Behrend, L. *et al.* Interaction of casein kinase 1 delta (CK1 δ) with post-Golgi structures, microtubules and the spindle apparatus. *Eur J Cell Biol* **79**, 240–251 (2000).
33. Stöter, M. *et al.* Inhibition of casein kinase I delta alters mitotic spindle formation and induces apoptosis in trophoblast cells. *Oncogene 2005 24:54* **24**, 7964–7975 (2005).
34. Wolff, S. *et al.* Interaction of casein kinase 1 delta (CK1 δ) with the light chain LC2 of microtubule associated protein 1A (MAP1A). *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1745**, 196–206 (2005).
35. Schitteck, B. & Sinnberg, T. Biological functions of casein kinase 1 isoforms and putative roles in tumorigenesis. *Molecular Cancer 2014 13:1* **13**, 1–14 (2014).
36. Lee, H., Chen, R., Lee, Y., Yoo, S. & Lee, C. Essential roles of CKI δ and CKI ϵ in the mammalian circadian clock. *Proc Natl Acad Sci U S A* **106**, 21359–21364 (2009).
37. Xu, Y. *et al.* Functional consequences of a CKI δ mutation causing familial advanced sleep phase syndrome. *Nature 2005 434:7033* **434**, 640–644 (2005).
38. Cunningham, P. S. *et al.* Targeting of the circadian clock via CK1 δ/ϵ to improve glucose homeostasis in obesity. *Scientific Reports 2016 6:1* **6**, 1–11 (2016).

39. Xu, P. *et al.* Gene expression levels of Casein kinase 1 (CK1) isoforms are correlated to adiponectin levels in adipose tissue of morbid obese patients and site-specific phosphorylation mediated by CK1 influences multimerization of adiponectin. *Mol Cell Endocrinol* **406**, 87–101 (2015).
40. Borba, J. V. V. B. *et al.* Update and elucidation of Plasmodium kinomes: Prioritization of kinases as potential drug targets for malaria. *Comput Struct Biotechnol J* **20**, 3708–3717 (2022).
41. Korb, O., Stützle, T. & Exner, T. E. An ant colony optimization approach to flexible protein–ligand docking. *Swarm Intelligence 2007 1:2* **1**, 115–134 (2007).
42. Case, D. A. *et al.* The Amber biomolecular simulation programs. *J Comput Chem* **26**, 1668–1688 (2005).
43. Harvey, M. J., Giupponi, G. & De Fabritiis, G. ACEMD: Accelerating biomolecular dynamics in the microsecond time scale. *J Chem Theory Comput* **5**, 1632–1639 (2009).
44. Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *J Mol Graph* **14**, 33–38 (1996).
45. Sun, D., Gao, W., Hu, H. & Zhou, S. Why 90% of clinical drug development fails and how to improve it? *Acta Pharm Sin B* **12**, 3049–3062 (2022).
46. Muhammed, M. T. & Aki-Yalcin, E. Homology modeling in drug discovery: Overview, current applications, and future perspectives. *Chem Biol Drug Des* **93**, 12–20 (2019).
47. Korb, O., Stützle, T. & Exner, T. E. Empirical scoring functions for advanced Protein-Ligand docking with PLANTS. *J Chem Inf Model* **49**, 84–96 (2009).
48. Protein BLAST: search protein databases using a protein query. https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome.
49. Matralis, A. N. *et al.* Development of Chemical Entities Endowed with Potent Fast-Killing Properties against Plasmodium falciparum Malaria Parasites. *J Med Chem* **62**, 9217–9235 (2019).
50. 8-Bromoguanosine | C₁₀H₁₂BrN₅O₅ | CID 135465599 - PubChem. <https://pubchem.ncbi.nlm.nih.gov/compound/135465599>.
51. 8-(cyclohexylamino)-3-methyl-7-(2-methylpropyl)-2,3,6,7-tetrahydro-1H-purine-2,6-dione | C₁₆H₂₅N₅O₂ | CID 719388 - PubChem. <https://pubchem.ncbi.nlm.nih.gov/compound/719388>.

52. 1-(sec-butyl)-6-hydroxy-5-(2,3,4,9-tetrahydro-1H-pyrido[3,4-b]indol-1-yl)pyrimidine-2,4(1H,3H)-dione | C₁₉H₂₂N₄O₃ | CID 4833775 - PubChem. <https://pubchem.ncbi.nlm.nih.gov/compound/4833775>.
53. 8-(2-Dimethylamino-ethylamino)-7-isobutyl-3-methyl-3,7-dihydro-purine-2,6-dione | C₁₄H₂₄N₆O₂ | CID 651390 - PubChem. <https://pubchem.ncbi.nlm.nih.gov/compound/651390>.
54. 2-Amino-9-((2R,3R,4S,5R)-3,4-dihydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-8-morpholino-1H-purin-6(9H)-one | C₁₄H₂₀N₆O₆ | CID 135921573 - PubChem. <https://pubchem.ncbi.nlm.nih.gov/compound/135921573>.
55. 3-Amino-8-(3-methoxyphenyl)-13,13-dimethyl-12-oxa-4,6,8-triazatricyclo[8.4.0.0^{2,7}]tetradeca-1(10),2,4,6-tetraen-9-one | C₁₉H₂₀N₄O₃ | CID 16407651 - PubChem. <https://pubchem.ncbi.nlm.nih.gov/compound/16407651>.
56. Zhao, H., Dong, J., Lafleur, K., Nevado, C. & Caflisch, A. Discovery of a novel chemotype of tyrosine kinase inhibitors by fragment-based docking and molecular dynamics. *ACS Med Chem Lett* **3**, 834–838 (2012).
57. 6-hydroxy-1-methyl-5-[[2R,3R,4R,5S,6R)-2,4,5-trihydroxy-6-(hydroxymethyl)oxan-3-yl]iminomethyl]pyrimidine-2,4-dione | C₁₂H₁₇N₃O₈ | CID 135549403 - PubChem. <https://pubchem.ncbi.nlm.nih.gov/compound/135549403>.
58. 6-(2-methoxyphenyl)-3-phenyl-N-(2-phenylethyl)-2H-pyrazolo[3,4-b]pyridine-4-carboxamide | C₂₈H₂₄N₄O₂ | CID 20939597 - PubChem. <https://pubchem.ncbi.nlm.nih.gov/compound/20939597>.
59. methyl 3-(3-hydroxy-6-(morpholinomethyl)-4-oxo-4H-pyran-2-yl)-3-(3-(hydroxymethyl)-4-methoxyphenyl)propanoate | C₂₂H₂₇N₃O₈ | CID 91636750 - PubChem. <https://pubchem.ncbi.nlm.nih.gov/compound/91636750>.
60. 9-[(2R,3R,4S,5R)-3,4-dihydroxy-5-(hydroxymethyl)oxolan-2-yl]-8-pyrrolidin-1-yl-1H-purin-6-one | C₁₄H₁₉N₅O₅ | CID 136234725 - PubChem. <https://pubchem.ncbi.nlm.nih.gov/compound/136234725>.
61. 3-(7H-purin-6-ylamino)-N-(3,4,5-trimethoxyphenyl)propanamide | C₁₇H₂₀N₆O₄ | CID 42507449 - PubChem. <https://pubchem.ncbi.nlm.nih.gov/compound/42507449>.

Ringraziamenti

Vorrei ringraziare per primo il professor Stefano Moro, che mi ha accolta letteralmente a braccia aperte nel suo laboratorio, per avermi sempre supportata e aiutata, e per avermi permesso di portare avanti un progetto di ricerca così stimolante.

Grazie agli MMSini, che in quest'ultimo anno sono stati per me una seconda famiglia. In particolare grazie a Matteo per la sua infinita disponibilità, chiarezza e pazienza. La tua presenza e il tuo aiuto sono stati fondamentali per riuscire a concretizzare tutto questo. Grazie per tutte quelle volte in cui mi hai spiegato e rispiegato le stesse cose. Il tuo sostegno non è mai stata una cosa scontata, per questo motivo meriti un ringraziamento speciale. Grazie a Veronica per aver provato ad insegnarmi *l'ordine e la disciplina*, per la tua presenza e supporto durante il mio percorso di ricerca. La tua competenza ha arricchito la mia esperienza, fornendomi una guida preziosa.

Grazie ad Andrea per avermi sostenuta nei momenti in cui ero giù, per avermi motivata ogni volta che ero convinta di non potercela fare, per avermi confortata e rassicurata quando qualcosa non andava come avrei voluto, e per aver festeggiato con me ogni mio successo. Grazie perché vedi sempre il bicchiere mezzo pieno, e per avermi ricordato, ogni volta che le mie insicurezze prendevano il sopravvento, che il traguardo era davvero dietro l'angolo. In questi mesi mi hai conosciuta sia nel bene che nel male, e sei sempre stato al mio fianco nel modo più dolce e sincero. La tua presenza rende ogni giornata più bella.

E per concludere il capitolo MMS, grazie a Silvia, Davide Ross, Gianluca, Chiara e Davide A.I. per aver condiviso con me (tante) gioie e (pochi) dolori, per tutto il sostegno che non mi avete mai fatto mancare. Grazie per le indimenticabili giteMMS, per i pomeriggi gin tonic e soppresata, e in generale per tutti i momenti di pausa spensierati trascorsi insieme. Siete il miglior *gruppo di ricerca* che potessi trovare. Grazie per avermi fatta sentire a casa, il periodo in MMS è stato il più bello di tutti gli anni di università.

Ringrazio i miei genitori che mi hanno sempre supportata, sia materialmente che emotivamente, durante tutto il mio percorso. Il vostro costante sostegno e la vostra immensa disponibilità sono stati fondamentali. Grazie per avermi compresa e per non avermi mai fatta sentire sotto pressione.

Ringrazio Gloria, Marica e Elisa, che mi hanno accolta e fatta sentire a casa dai primi giorni in cui sono arrivata qui a Padova. In particolare ringrazio Gloria per essere stata una coinquilina meravigliosa e per aver condiviso con me un anno di sopravvivenza in quel covo di criminali e novax chiamato Camplus, che però porto nel cuore.

Ringrazio Viola per tutte le serate spensierate che abbiamo passato su Discord e Sergio per i meme, le serate di giochi e il tempo trascorso insieme.

Ringrazio Anna, per apprezzarmi e sostenermi per quella che sono, per i succosissimi *drama* che anni fa erano all'ordine del giorno, e anche se ora sono un po' diminuiti forse è meglio così. Grazie per esserci sempre.

Ringrazio tutte le persone che, pur talvolta essendo fisicamente distanti, durante questi ultimi anni mi hanno fatta stare bene e alle quali (anche se probabilmente non l'ho mai detto) voglio tanto bene.

E per concludere, ringrazio me stessa. Sono stati anni complicati, ricchi di momenti in cui tutto mi sembrava solo un brutto incubo dal quale speravo solo di risvegliarmi presto. Ringrazio me stessa per non aver mollato, per essere cresciuta in questi anni, per aver finalmente capito quanto sia importante stare bene e quanto la vita sia meravigliosa.

Anna