

Modelli biofisici in silico per la predizione di inibitori enzimatici in modelli batterici

Matteo Calcagnile and Pietro Alifano

*University of Salento, Lecce, Department of Biological and Environmental Sciences
and Technologies (DiSTeBA)*

Lecce, 73100

e-mail address: matteo.calcagnile@unisalento.it

Abstract

The development of new drugs is a hard and expensive goal for industry and scientists. The computational approach can drive this process reducing time and costs. This work shows pipeline and software to model the interaction between a target enzyme and a small molecule (inhibitor). Usually, allosteric compounds can interact with the protein, inducing conformational rearrangement that affects the functionality of the enzyme. Here the 3D model of one enzyme, the isocitrate lyase (ICL), was modeled starting from a sequence of ICL of *Streptomyces coelicolor* and, using online tools, allosteric pockets were identified. Finally, the potential inhibitors were screened using a docking method. ICL is an important factor of virulence and persistence in *Mycobacterium tuberculosis*, and one commercially available drug was identified: rapamycin, an mTOR inhibitor used in prevention of transplant rejection.

Riassunto

Gli inibitori enzimatici sono una classe di farmaci utilizzati per interferire a livello molecolare con la funzionalità di proteine “target”. Spesso questi composti si comportano come modulatori allosterici, ovvero legano l’enzima in posizione diversa dal sito attivo inducendo dei cambiamenti conformazionali che disattivano funzionalmente l’enzima. Oggi, grazie a modelli computazionali, è possibile ottenere modelli in 3D di enzimi a struttura sconosciuta, identificando il sito attivo e le tasche allosteriche coinvolte in questo tipo di regolazione. Inoltre, grazie al *docking* è possibile stimare quanto un inibitore è affine ad una determinata tasca. Sfruttando quest’approccio abbiamo creato un modello biofisico di alcuni enzimi (isocitrato liasi, ICL) e identificato delle molecole potenzialmente utilizzabili come inibitori. In particolare, questo approccio è interessante poiché l’ICL è un fattore di virulenza e persistenza batterico. Fra gli inibitori identificati spicca il caso della rapamicina, molecola già disponibile in commercio utilizzata come immunosoppressore e come anti-neoplastico e che potenzialmente potrebbe essere utilizzata anche contro la tubercolosi.

Keywords: inibitori enzimatici, streptomiceti, isocitrato liasi, *Mycobacterium tuberculosis*, rapamicina

1. Introduzione

Gli inibitori enzimatici sono piccole molecole che possono interagire con un enzima modulandone la funzione catalitica. Ci sono vari livelli di regolazione che modulano le reazioni enzimatiche nei sistemi biologici. Alcuni tipi di regolazione avvengono prima che l'enzima sia stato sintetizzato, viceversa esistono meccanismi di regolazione che agiscono direttamente modulando l'attività di un enzima. Quest'ultimo tipo di controllo è formato essenzialmente da un sistema macromolecolare costituito da un interattore proteico (enzima) e una piccola molecola che, legandosi all'enzima ne modula la funzionalità. Un sistema di questo tipo può oggi essere modellizzato a livello teorico grazie delle leggi della termodinamica e *in silico*, ovvero utilizzando approcci di calcolo e simulazione computazionale.

Inizialmente bisogna considerare che gli enzimi sono costituiti da lunghe catene amminoacidiche che si ripiegano su sé stesse per formare da prima le strutture secondarie (eliche, foglietti beta) e conseguentemente strutture globulari (terziarie). Infine, più strutture globulari possono interagire fra loro formando strutture quaternarie: in questo caso si definiscono omo-oligomeri gli agglomerati che sono formati da sub-unità identiche fra loro o etero-oligomeri quelli in cui le sub-unità siano differenti proteine. Tali oligomeri possono essere formati da 2 fino a 8 sub-unità. Questo tipo di interazioni, chiamate *foldings*, dipendono essenzialmente dalla diminuzione dell'energia libera che possiede il sistema, che secondo uno dei principi fondamentali della termodinamica deve assumere il valore massimo (minimo di energia del sistema) [1]. Normalmente all'interno della struttura globulare di un enzima possono essere individuate varie tasche, cavità interne che formano siti di legame per altre macromolecole. Il substrato di un enzima lega proprio una di queste tasche, definita sito attivo dove viene convertito in prodotto [1].

A livello termodinamico le reazioni possono avvenire spontaneamente quando il contenuto energetico del substrato è superiore a quello del prodotto e, quindi, quando la reazione è termodinamicamente favorita. Viceversa, se il substrato ha un contenuto in termini di energia maggiore del prodotto occorre una reazione accoppiata che fornisca l'energia mancante, come ad esempio la conversione dell'ATP in ADP. In entrambe le ipotesi per far avvenire la reazione deve essere superata una barriera di energia (figura 1.1 A). Questa barriera di energia rappresenta lo stato di transizione, ovvero una molecola instabile che esiste solo per un tempo limitato e che possiede un'energia superiore sia al substrato che al prodotto [1].

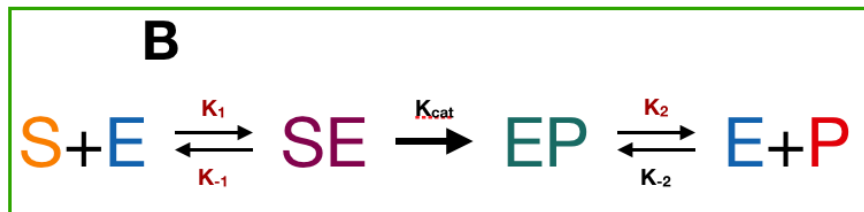
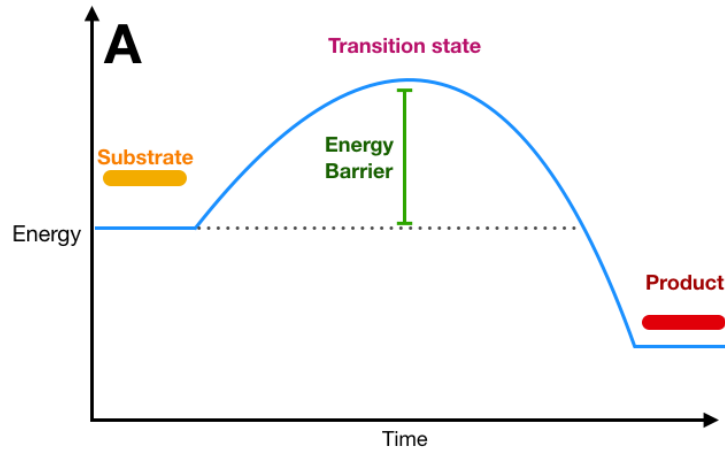


Figura 1.1: A- grafico rappresentante il tempo nelle ordinate e il contenuto energetico nelle ascisse. In verde è rappresentata l'energia necessaria al superamento della barriera energetica. B- Schema generale di una reazione con le relative costanti K evidenziate. S= substrato, E=enzima, SE=complesso enzima substrato, EP= complesso enzima prodotto; P= prodotto.

Generalmente, le reazioni enzimatiche possono essere descritte da varie fasi consecutive (figura 1.1 B): innanzitutto substrato ed enzima devono incontrarsi e formare un complesso enzima-substrato (SE). Naturalmente la formazione del complesso avviene con una precisa cinetica rappresentata da K_1 , mentre K_{-1} rappresenta il processo inverso (il complesso che si dissocia). Formazione e rilascio del complesso sono processi che avvengono spontaneamente in soluzione. Nel passaggio seguente il complesso enzima-substrato si trasforma nel risultante complesso enzima-prodotto (EP). Tale processo può essere espresso tramite un'ulteriore costante K_{cat} . Quest'ultimo, analogamente alla formazione del complesso SE ma in maniera contraria deve dissociarsi per liberare il prodotto (K_2), mentre è possibile che il complesso EP si riformi (K_{-2}). L'affinità di un substrato, di un prodotto o di un qualsiasi altro modulatore per un enzima può essere stimata grazie a *software*, calcolando la variazione dell'energia libera di Gibbs (ΔG) mentre non è possibile calcolare la K_{cat} con metodi computazionali.

La funzione dell'enzima consiste proprio nell'abbassare l'energia d'attivazione associata allo stato di transizione, ovvero l'energia necessaria a superare la barriera. Le equazioni cinetiche che descrivono questo processo sono le equazioni di Michaelis-Menten 1.1 e 1.2, in cui: V è la velocità della reazione; V_{max} è la sua velocità massima, $[S]$ è la concentrazione di prodotto; K_m è la costante di Michaelis-Menten [1].

$$V = \frac{V_{max}[S]}{K_m + [S]} \quad (1.1)$$

$$K_m = \frac{k_{-1} + K_2}{K_{-2}} \quad (1.2)$$

Gli inibitori possono interferire con questo meccanismo di differenti maniere: i) legando il sito attivo, competendo, quindi, con il substrato per l'accesso a questo sito attivo; ii) legando un ulteriore sito, chiamato allosterico causando una perturbazione nella conformazione terziaria e quaternaria della proteina e modulando l'attività enzimatica o iii) un meccanismo che comprende entrambi i casi. I modulatori allosterici possono essere negativi (inibitori) o positivi. L'equazione di Michaelis-Menten può essere modificata considerando gli inibitori competitivi (1.3 che causano un cambiamento nella costante K_m) e allosterici (1.4) [1].

$$V = \frac{V_{max}[S]}{\alpha K_m + [S]} \quad (1.3)$$

$$V = \frac{V_{max}[S]}{K_m + \alpha [S]} \quad (1.4)$$

In questo sistema, utilizzando metodi computazionali è possibile calcolare: a) il *folding* di un enzima a struttura sconosciuta; b) L'affinità (ΔG) del sito attivo di un enzima con substrato e/o prodotto; c) l'affinità (ΔG) di eventuali modulatori in tasche allosteriche dell'enzima.

Lo scopo di questo studio *in silico* è individuare con metodi computazionali possibili modulatori allosterici per l'enzima isocitrato liasi (ICL o AceA) codificato dal gene *aceA*. Questo enzima fa parte di un bypass attivo a livello del ciclo degli acidi tricarbossilici (TCA o ciclo di Krebs) unitamente ad un altro enzima, chiamato malato sintasi (AceB) e codificato dal gene *aceB* (figura 1.2). Queste due reazioni permettono di saltare alcune tappe del ciclo TCA, che normalmente permette di convertire il carbonio dalla forma ridotta a quella ossidata immagazzinando potenziale riducente che verrà successivamente convertito in ATP dalla catena di trasporto degli elettroni. Durante questo il carbonio viene espulso come CO_2 , mentre se è attivo il bypass, il flusso è veicolato verso la biosintesi, non permettendo l'accumulo di potenziale riducente sotto forma di NADH (ottenuto per riduzione del NAD^+). Tale bypass è chiamato anche ciclo del gliossilato poiché l'ICL produce questa molecola che successivamente viene convertita in malato a sua volta intermedio del ciclo TCA. Di fatto l'ICL rappresenta un interruttore molecolare che regola il bilancio fra carbonio ridotto da cui estrarre energia e carbonio ridotto veicolato verso la biosintesi.

Alcuni studi sperimentali sottolineano l'importanza di questo enzima nei batteri: in *Escherichia coli* l'espressione di questo enzima è indotta da lipidi tramite il fattore CRP e il cAMP [2-3-4] mentre in *Coynebacterium glutamicum* l'espressione è indotta dall'acetato [5] e, infine, in *Mycobacterium tuberculosis* questo enzima è definito come un fattore di persistenza e virulenza, in quanto la sua espressione permette al batterio di sopravvivere ai trattamenti e causa un aumento di virulenza del batterio [6].

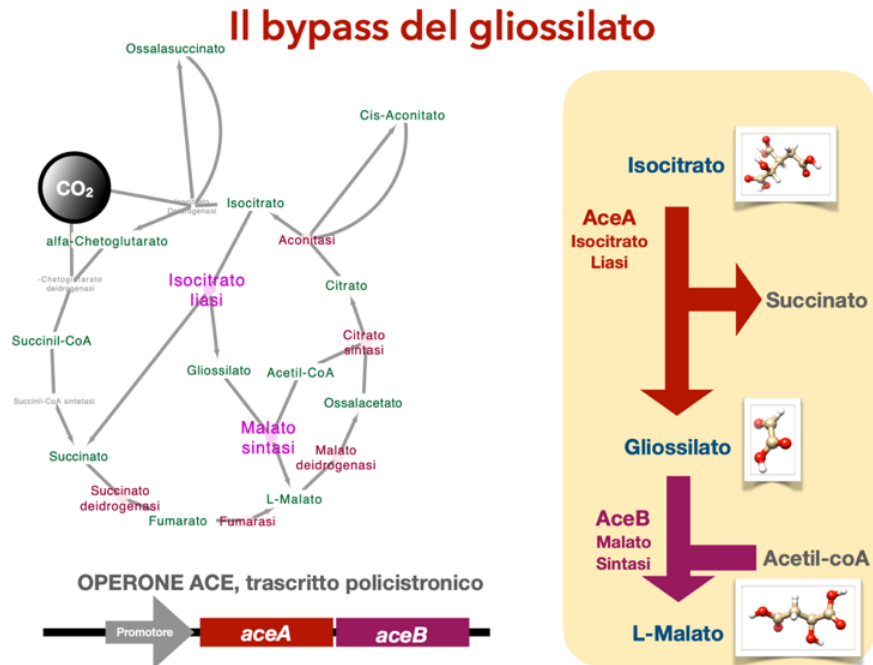


Figure 1.2: Bypass del gliossilato: a sinistra considerando il ciclo TCA, a destra evidenziando substrati e i prodotti degli enzimi AceA e AceB.

2. Materiali e Metodi

Molecular modeling. Per modellare la struttura dell'enzima sono stati utilizzati diversi *software*. Innanzitutto, partendo dalla sequenza amminoacidica dell'enzima, deve essere prodotto un modello 3D della proteina in questione. Molti modelli sono già stati dedotti da studi cristallografici e sono riportati in *database online* come il Protein Data Bank (PDB) [7]. Tuttavia, in questo studio abbiamo deciso di concentrare le nostre attenzioni sugli enzimi degli attinomiceti, fra i quali troviamo il *M. tuberculosis* e gli streptomiceti, questi ultimi conosciuti per il loro potenziale biosintetico (molti antibiotici sono prodotti da fermentazioni utilizzando batteri di questo genere). Nei database online non sono presenti modelli di ICL ricavati da streptomiceti, quindi utilizzando la sequenza primaria dell'ICL di *Streptomyces coelicolor* come input, e il *server online* di modellistica molecolare I-Tasser [7] è stato ottenuto un primo modello di enzima. Questo modello è stato successivamente sottomesso al server online Galaxy-Homomer [8] per ricavare una struttura omo-oligomerica quaternaria.

Predizione dei siti allosterici, dei modulatori e docking. Al fine di predire quali siano i siti allosterici e gli eventuali modulatori è stato utilizzato un insieme di algoritmi eseguiti su un *server online*: AlloFider [9-10]. Questo *tool* da prima individua i potenziali siti allosterici e successivamente permette di individuare potenziali modulatori, utilizzando un database scelto dall'utente, in questo caso il database scelto è ZincAC [11] che ha il vantaggio di riportare composti disponibili in commercio. Una volta individuati i potenziali modulatori, 300 di queste molecole che presentavano uno *score* alto (>0,8 su un massimo

di 1), ovvero che più probabilmente modulano l'attività enzimatica, sono state scaricate dai database online (PubChem e ChemSpider) [12-13] ed utilizzate per uno *screening in silico* utilizzando il *software stand-alone* AutoDock Vina [14], che stima la variazione di energia libera associata ad ogni potenziale sito allosterico, per tutti i 300 composti identificati. I risultati sono stati analizzati con PAST [15] al fine di ricavare grafici in multivariata mentre le immagini sono state prodotte utilizzando i *software* Avogadro [16] e Chimera [17]. Infine, per gli allineamenti delle sequenze primarie è stato usato il software LALIGN [18].

3. Risultati

Tasche di legame per modulatori allosterici. Come accennato in materiali e metodi, sottomentendo la sequenza dell'enzima ICL di *S. coelicolor* ad I-Tasser è possibile ottenere: una struttura tridimensionale della proteina e alcune informazioni riguardo al sito attivo che lega il substrato. Utilizzando il *software* Chimera è stata visualizzata la struttura della proteina e gli amminoacidi indicati come facenti parte del sito attivo sono stati segnati in rosso (figura 3.1).

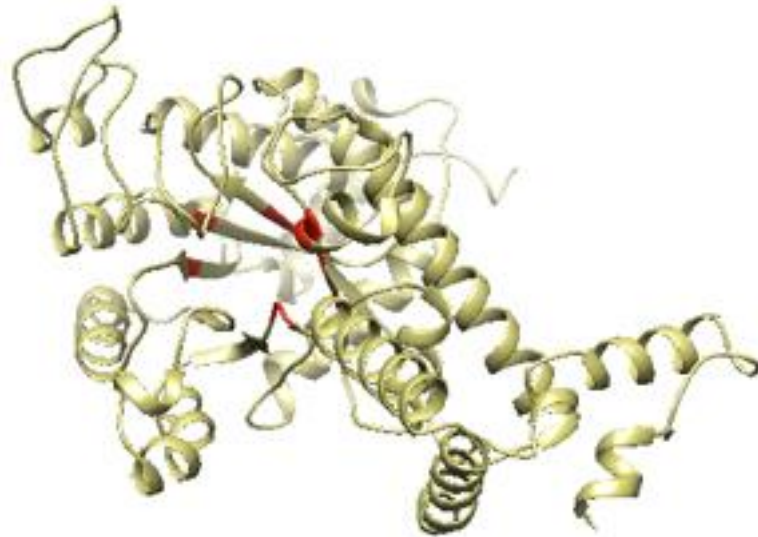


Figure 3.1: modello dell'ICL di *S. coelicolor* ricavato utilizzando il server online I-Tasser. In rosso sono evidenziati gli amminoacidi facenti parte del sito attivo che lega il substrato.

Successivamente il *software* Galaxy-Homomer ha previsto che questo enzima nella sua conformazione oligomerica sia formato da 4 subunità identiche ripetute (figura 3.2, gli amminoacidi del sito attivo sono indicati nuovamente in rosso).

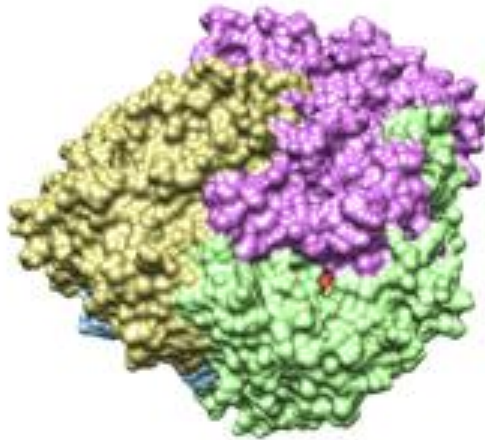


Figure 3.2: modello tetra-oligomero dell'ICL di *S. coelicolor* ottenuto da Homomer, sottomettendo il modello precedentemente ottenuto (mostrato in figura 3.1). Il sito attivo è segnato in rosso.

Sottomettendo la struttura tridimensionale ottenuta tramite I-Tasser in AlloFinder sono stati identificate le potenziali tasche allosteriche. La prima parte dell'algoritmo utilizzato da AlloFinder (denominata AlloSite) ha predetto 10 tasche allosteriche (numerate da 0 a 9) ma soltanto 3 avevano uno score superiore a 0,8 (il massimo è 1). Queste tasche erano la 0 la 3 e la 6 e sono evidenziate nella figura 3.3.

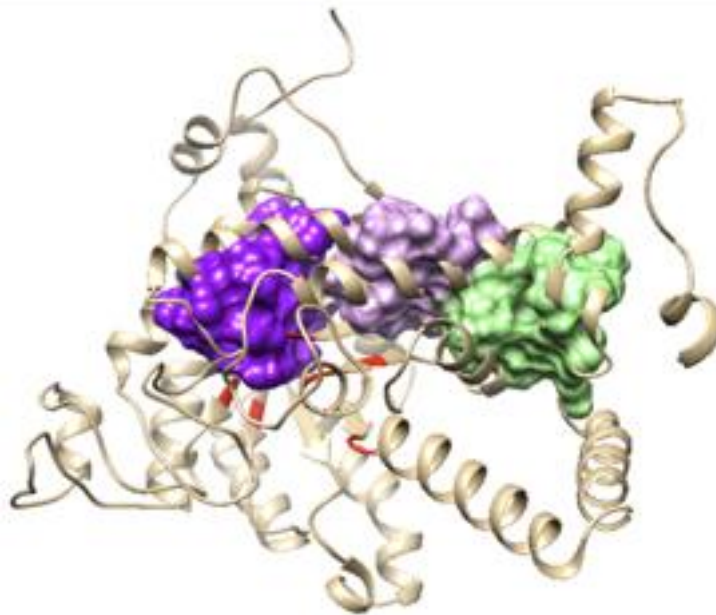


Figure 3.3: Modello di ICL di *S. coelicolor* con evidenziate le tre tasche allosteriche: tasca 0 viola scuro, tasca 3 viola chiaro, tasca 6 verde.

Una volta ottenute le probabili tasche allosteriche predette da AlloSite, l'algoritmo di AlloFinder prosegue calcolando i composti che potenzialmente hanno maggiore affinità per queste tasche individuate. Come detto precedentemente, il database da cui provengono tali composti è ZincAC che ha il vantaggio di contenere soltanto molecole disponibili in commercio. Per ogni sito sono predetti vari modulatori da cui sono stati selezionati i 100 composti con lo score più alto per ogni tasca, selezionando in tutto 300 composti.

Docking. Ottenute le predizioni delle tasche e i 300 composti, questi ultimi sono stati scaricati in file formato mol dai database online SpiderChem e PubChem. Per ogni tasca sono stati utilizzati come ligandi tutti i 300 composti, eseguendo 900 simulazioni di docking in totale. Questi dati sono stati analizzati tramite il software PAST per produrre un'analisi multivariata (non riportata). Tale analisi è opportuna per poter confrontare i risultati di ogni ligando per tutte e tre le tasche. I ligandi con score maggiore sono poi stati usati come input per altre simulazioni di docking utilizzando, tuttavia, un altro enzima target, l'ICL del *M. tuberculosis*. In questo caso la struttura tridimensionale dell'enzima è stata ottenuta con la cristallografia ed è disponibile sui *database online*. La tabella 3.1 riporta i migliori ligandi selezionati grazie a questo *screening in silico*.

Tabella 3.1 – Ligandi selezionati dopo lo screening in silico dei composti ottenuti in Allofinder (sia contro le tre tasche predette sul modello di ICL di *S. coelicolor* sia sull'unica tasca predetta per *M. tuberculosis*).

Composti	<i>S. coelicolor</i>				<i>Mycobacterium</i>
	0 (Kj/mol)	3 (Kj/mol)	6 (Kj/mol)	Media (Kj/mol)	Kj/mol
ZINC000010109500	-8,0	-9,6	-8,6	-8,7	-8,2
ZINC000011868123	-8,3	-8,8	-8,8	-8,6	-8,2
ZINC000010301115	-8,4	-8,6	-8,6	-8,5	-7,9
Rapamycin	-4,5	-6,6	-7,6	-6,23	-5,9
c-diGMP	-6,6	-8,8	-9,4	-8,27	-6,8
NADPH	-6,6	-6,5	-7,5	-6,87	-6,5
UDP-glucose	-6,4	-6,8	-7,1	-6,77	-5,7
NAD+	-6,5	-6,7	-7	-6,73	-6,0
GTP	-6	-7	-6,8	-6,60	-6,3
NADH	-6,4	-6,1	-7,1	-6,53	-6,2
GDP	-6,3	-6,3	-6,7	-6,43	-7,1

Confronto fra ICL del micobatterio e degli streptomiceti. Al fine di poter confrontare le strutture dell'ICL in *S. coelicolor* e in *M. tuberculosis* è stato eseguito un allineamento delle strutture primarie utilizzando l'algoritmo LALIG. Tale allineamento dimostra come: i) le due sequenze coincidono per lunghezza e nessuna delle due presenta gap rispetto all'altra; ii) hanno una coerenza del 88% (nel 88% dei casi gli amminoacidi se pur diversi hanno le stesse caratteristiche chimico-fisiche); iii) sono identiche per il 78% dei residui amminoacidici coinvolti. La figura 3.4 riporta il confronto fra l'unica tasca allosterica identificata utilizzando sempre Allofinder in *M. tuberculosis* e la corrispondente tasca nell'enzima modellato a partire dalla sequenza di *S. coelicolor*. Da questa figura si può notare come le tasche siano sovrapponibili.

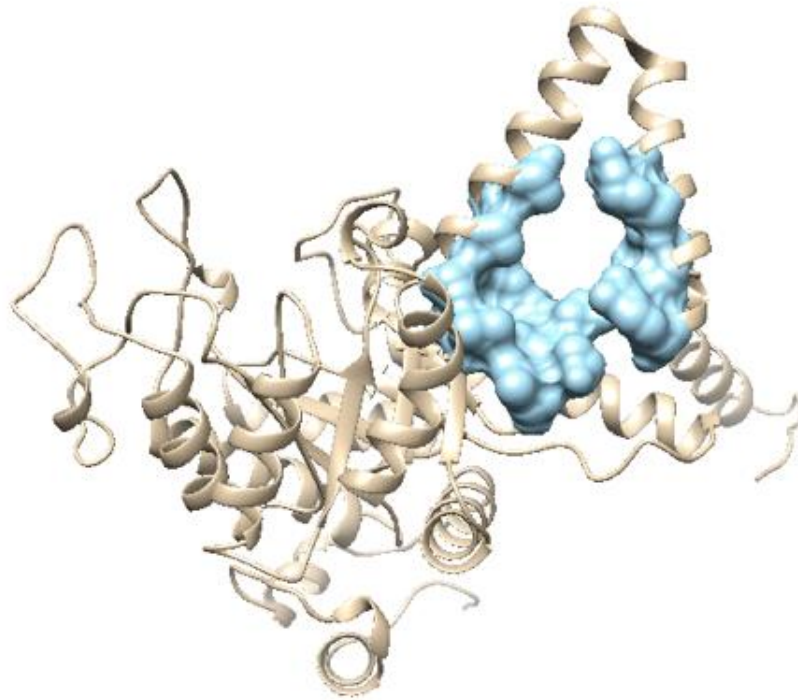


Figure 3.3: Modello di ICL di *M. tuberculosis* scaricato dal database PDB. In celeste è evidenziata la stessa tasca allosterica predetta per *S. coelicolor*. Dalla figura è evidente come questo sito allosterico sia omologo nei due enzimi.

4. Conclusioni

I composti identificati con i metodi precedentemente descritti comprendono sia composti utilizzati comunemente in terapie farmacologiche (in particolare la rapamicina) sia composti che possiamo definire endogeni, ovvero presenti comunemente nei sistemi biologici. Fra i composti endogeni troviamo molte molecole che agiscono come intermediari nel trasporto dell'energia sotto forma sia di potenziale riducente (NAD nelle sue forme ossidate e ridotte), sia sotto forma di legami ad alta energia (fosfodiesteri, come il GTP). Questo indica come la regolazione del bypass del gliossilato sia strettamente associata allo stato energetico del microrganismo: grandi contenuti di NADH, ad esempio,

indicano un buon approvvigionamento. Come detto precedentemente l'attivazione del bypass sposta il flusso di carbonio dalla produzione di energia alla biosintesi, quindi tali modulatori potrebbero regolare l'attivazione dell'ICL solo quanto vi è abbastanza energia evitando, in questo modo, che il microrganismo non sopravviva.

Caso interessante è quello della rapamicina, farmaco prodotto da *Streptomyces hygroscopicus* ed utilizzato nelle terapie verso tumori e malattie autoimmuni. Tale farmaco, secondo i dati prodotti dai software, potrebbe modulare l'attività dell'ICL anche nel *M. tuberculosis*, dove funziona come fattore di virulenza e persistenza. Tale farmaco, già approvato per l'uso nell'uomo potrebbe quindi essere usato anche per le malattie infettive, riducendo di molto i costi legati allo sviluppo di nuovi farmaci. Anche se tali approcci *in silico* non sono sostitutivi delle sperimentazioni, possono indirizzare i ricercatori verso composti farmacologici più certi per un target conosciuto e ridurre il costo di nuove sperimentazioni, utilizzando composti già usati per l'uomo.

Referenze

- [1] A.L. Lehninger, M.M Cox, D.L Nelson. (2008) Lehninger principles of biochemistry /New York: W.H. Freeman.
- [2] M.H. Saier. (1996). Cyclic AMP-independent catabolite repression in bacteria. FEMS Microbiol. Lett. 138, 97/103.
- [3] M.H. Saier. (1998). Multiple mechanisms controlling carbon metabolism in bacteria. Biotechnol. Bioeng. 58, 170/174.
- [4] S.R. Maloy, M. Bohlander, W.D. Nunn. (1980). Elevated levels of glyoxylate shunt enzymes in Escherichia coli strains constitutive for fatty acid degradation. J Bacteriol. 1980 Aug;143(2):720-5.
- [5] R. Gerstmeir, V.F. Wendisch, S. Schnicke, H. Ruan, M. Farwick, D. Reinscheid, B.J. Eikmanns. Acetate metabolism and its regulation in Corynebacterium glutamicum. J Biotechnol. 2003 Sep 4;104(1-3):99-122.
- [6] J.D. McKinney, K. Höner zu Bentrup, E.J. Muñoz-Elías, A. Miczak, B. Chen, W.T. Chan, D. Swenson, J.C. Sacchettini, W.R. Jacobs, D.G. Russell. Persistence of Mycobacterium tuberculosis in macrophages and mice requires the glyoxylate shunt enzyme isocitrate lyase. Nature. 2000 Aug 17;406(6797):735-8.
- [7] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne. (2000) The Protein Data Bank Nucleic Acids Research, 28: 235-242.
- [8] M. Baek, T. Park, L. Heo, C. Park, S. Seok. GalaxyHomomer: a web server for protein homo-oligomer structure prediction from a monomer sequence or structure. Nucleic Acids Res. 2017 Jul 3;45(W1):W320-W324. doi: 10.1093/nar/gkx246. PMID: 28387820; PMCID: PMC5570155.
- [9] Wenkang Huang, Shaoyong Lu, Zhimin Huang, Xinyi Liu, Linkai Mou, Yu Luo, Yanlong Zhao, Yaqin Liu, Zhongjie Chen, Tingjun Hou, Jian Zhang, Allosite: a method for predicting allosteric sites, Bioinformatics, Volume 29, Issue 18, 15 September 2013, Pages 2357–2359, <https://doi.org/10.1093/bioinformatics/btt399>

- [10] Huang, M.; Song, K.; Liu, X.; Lu, S.; Shen, Q.; Wang, R.; Gao, J.; Hong, Y.; Li, Q.; Ni, D.; Xu, J.; Chen, G.; Zhang, J. AlloFinder: A Strategy for Allosteric Modulator Discovery and Allosterome Analyses. *Nucleic Acids Res.* 2018, 46, W451–W458.
- [11] Irwin JJ, Shoichet BK. ZINC--a free database of commercially available compounds for virtual screening. *J Chem Inf Model.* 2005;45(1):177–182. doi:10.1021/ci049714+
- [12] Kim S, Chen J, Cheng T, Gindulyte A, He J, He S, Li Q, Shoemaker BA, Thiessen PA, Yu B, Zaslavsky L, Zhang J, Bolton EE. PubChem 2019 update: improved access to chemical data. *Nucleic Acids Res.* 2019 Jan 8; 47(D1):D1102-1109. doi:10.1093/nar/gky1033.
- [13] Harry E. Pence and Antony Williams. ChemSpider: An Online Chemical Information Resource. *Journal of Chemical Education* 2010 87 (11), 1123-1124
DOI: 10.1021/ed100697w
- [14] O. Trott, A. J. Olson, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading, *Journal of Computational Chemistry* 31 (2010) 455-461
- [15] Hammer, Ø., Harper, D.A.T., Ryan, P.D. 2001. PAST: Paleontological statistics software package for education and data analysis. *Palaeontologia Electronica* 4(1): 9pp
- [16] Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem.* 2004 Oct;25(13):1605-12.
- [17] Marcus D Hanwell, Donald E Curtis, David C Lonie, Tim Vandermeersch, Eva Zurek and Geoffrey R Hutchison. Avogadro: An advanced semantic chemical editor, visualization, and analysis platform. *Journal of Cheminformatics* **2012**, 4:17.
- [18] X. Huang and W. Miller (1991) *Adv. Appl. Math.* 12:373-381