

Un approccio computazionale per lo studio di nodi metabolici coinvolti con la sintesi di spiramicina I in *Streptomyces ambofaciens*

Calcagnile M¹, Damiano F¹, Talà A¹, Alifano P¹.

¹ Dipartimento di Scienze e Tecnologie Biologiche ed Ambientali, Università del Salento, Lecce.
matteo.calcagnile@gmail.com

Abstract

Gli streptomiceti rappresentano il genere di organismi maggiormente utilizzati per la produzione industriale di metaboliti secondari. In particolare *S. ambofaciens* è il ceppo produttore della spiramicina (antibiotico). Una delle sfide scientifiche più importanti in questo campo è la capacità di modulare il metabolismo veicolando il flusso di carbonio verso la sintesi della molecola di interesse. In questo studio, utilizzando metodi computazionali, è stata indagata la funzione di una proteina sconosciuta (pirina) che regola la tappa iniziale della sintesi della spiramicina. In particolare uno dei target della pirina è una deidrogenasi che è correlata con la β -ossidazione e che controlla in maniera indiretta la formazione di etilmalonil-CoA importante precursore dell'antibiotico in questione.

Introduzione

Gli streptomiceti sono batteri gram-positivi che sono conosciuti soprattutto in virtù del grande numero di metaboliti secondari che producono, fra cui alcuni utilizzati come antibiotici. La crescita degli streptomiceti è miceliale e la formazione del micelio aereo, che avviene dopo 18 ore dalla germinazione delle spore, corrisponde con l'attivazione del metabolismo secondario. Il genoma di questi batteri è generalmente di grandi dimensioni e rispecchia un alto potenziale metabolico per cui una singola specie di streptomicete produce un grande numero di metaboliti secondari. *Streptomyces ambofaciens* [1], scoperto a Parigi durante gli anni '50, è il ceppo produttore della spiramicina, utilizzata ad esempio per curare la toxoplasmosi in gravidanza [2]. In realtà il batterio selvatico produce tre forme di spiramicina di cui solo la

I è prodotta dai ceppi industriali e disponibile in commercio (figura 1). Altri antibiotici prodotti dallo stesso streptomicete sono, ad esempio, la congocidina e le antimicine. Data la complessità del metabolismo di questi batteri, è importante capire in profondità i meccanismi molecolari che permettono l'attivazione del metabolismo secondario che corrisponde anche ad alcuni cambiamenti fenotipici, ad esempio la pigmentazione.

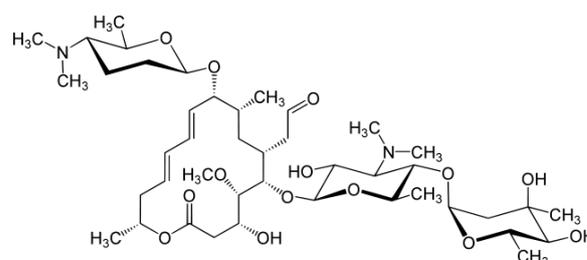


Figura 1: Struttura della spiramicina: si può notare lo scheletro macrolattone centrale e gli zuccheri aggiunti con la reazione di decorazione. La sintesi dello scheletro centrale avviene grazie alle PKS che utilizzano alcuni precursori. La

sintesi dei precursori può essere considerata la tappa primaria per la formazione dell'antibiotico.

Alcuni geni sono conosciuti come regolatori pleiotropici grazie alla capacità che hanno di controllare numerosi caratteri. Ad esempio i geni *bld* in *S. coelicolor* se inattivati provocano un abbassamento nella sintesi dei metaboliti secondari e inibiscono la crescita del micelio aereo [3]. Altri geni regolano in maniera specifica la produzione di un particolare metabolita. Ad esempio la produzione di spiramicina è regolata a livello trascrizionale grazie a due geni *srm22* e *srm40*, il primo regola la trascrizione del secondo che a sua volta attiva la trascrizione di tutti gli altri geni del cluster, compresi i complessi enzimatici (PKS di tipo I) [4]. Questi ultimi assemblano l'anello macrolattone a 16 atomi di carbonio che rappresenta lo scheletro dell'antibiotico a partire dai precursori la cui sintesi è la fase iniziale del processo di formazione dello stesso antibiotico. Infine sono aggiunti gli zuccheri nella fase finale detta di decorazione [5]. Il precursore fondamentale per la spiramicina è l'etilmalonil-CoA prodotto di una reazione il cui substrato di partenza è rappresentato dal crotonil-CoA. La sintesi del precursore è finemente regolata ed è di fondamentale importanza capire come il flusso di carbonio è canalizzato verso la produzione di spiramicina. Evidenze sperimentali suggerivano che una proteina appartenente alla famiglia delle pirine potesse regolare la biosintesi di spiramicina grazie al controllo di enzimi implicati nei nodi metabolici, come ad esempio la sintesi dell'etilmalonil-CoA. Le pirine sono proteine altamente conservate presenti in eucarioti e procarioti formate da domini tipici della super famiglia delle cupine che solitamente legano un atomo metallico. Fra le pirine conosciute alcune hanno funzione enzimatica (come la quercetina diossigenasi di *Escherichia coli* che lega cadmio [6]) mentre nell'uomo la pirina forma un'oligomero con Bcl-3 e NF- κ B responsabile della regolazione trascrizionale di alcuni geni pro-apoptotici [7]. In *Serratia marcescens* la

pirina controlla l'attività dell'enzima piruvato deidrogenasi che rappresenta uno snodo metabolico fondamentale che separa il ciclo dell'acido citrico dalla glicolisi [8]. In *S. ambofaciens* esistono quattro differenti geni *pirin-like* a funzione sconosciuta. In particolare l'inattivazione di una delle quattro pirine (chiamata PirA) provoca un crollo della sintesi di spiramicina. La sequenza codificante per PirA si trova in un locus nei pressi di un enzima annotato come un'acil-CoA deidrogenasi (ACAD). Esistono differenti forme di ACAD che si differiscono l'una dall'altra rispetto alla lunghezza dei substrati che utilizzano: possono utilizzare catene lunghe o molto lunghe (LCAD e VLCAD), catene medie (MCAD) oppure catene corte (SCAD). Nell'uomo il fattore HIF-1 si comporta da regolatore redox che, in condizioni d'ipossia, inibisce l'attività di enzimi come LCAD e MCAD, regolando in negativo l'apporto di substrati (acidi grassi) per la β -ossidazione che a sua volta è implicata nella formazione di acetil-CoA utile come metabolita primario correlato con il ciclo dell'acido citrico (TCA). Lo stesso fattore HIF-1 regola anche alcuni nodi metabolici coinvolti nella glicolisi che a sua volta produce acetil-CoA [9]. Negli streptomiceti è ipotizzabile che PirA svolga un ruolo di regolatore diretto per l'acil-CoA deidrogenasi, identificata come una ML-CAD e di seguito denominata BCAD, controllando la produzione di etilmalonil-CoA e quindi la biosintesi della spiramicina. Per indagare la funzione di questa proteina è stato utilizzato un approccio computazionale basato su programmi di 3D modeling e docking molecolare. In particolare questo tipo di tecnologie informatiche utilizzano diversi algoritmi per selezionare ed elaborare dati presenti in un database oppure calcolano i valori che un dato parametro può assumere secondo una funzione. In particolare questo tipo di software e/o server sono spesso freeware, semplici da utilizzare e molto più veloci rispetto alle tecniche sperimentali. Tuttavia i risultati prodotti, denominati

appunto predizioni, sono abbastanza incerti. I risultati devono essere quindi comprovati o tramite l'utilizzo di software differenti o con esperimenti in vitro o in vivo. Grazie a questi approcci è stato possibile ottenere informazioni sull'interazione specifica fra BCAD e PirA. In particolare in quest'articolo sarà approfondito l'aspetto computazionale soffermandosi sugli algoritmi utilizzati e ipotizzando un meccanismo molecolare e biofisico.

Materiali e metodi

Partendo dalla sequenza primaria della proteina attraverso i software riportati in tabella 1 si sono ottenuti dei dati (output) che sono stati confrontati fra loro. Inoltre è necessario rilevare come alcuni output ottenuti siano stati utilizzati come input per programmi di calcolo utilizzati in seguito.

Nome	Output ottenuto	Sito
I-Tasser	Modello tridimensionale in PDB	http://zhanglab.cmb.med.umich.edu/I-TASSER/
I-Tasser - Coach	Software interno di I-Tasser - previsione del ligando	
I-Tasser - Binding Site	Software interno di I-Tasser - previsione AA di binding	
I-Tasser - Enzyme	Software interno di I-Tasser - previsione funzione enzima	
Rampage	Angoli diedri degli AA e qualità del modello	http://mordred.bioc.cam.ac.uk/~rapper/rampage.php
Gemini	omodimerizzazione	http://galaxy.seoklab.org/cgi-bin/submit.cgi?type=GEMINI
Blast (delta-blast)	Allineamento pir/genoma di S. Ambobaciens	http://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch&BLAST_SPEC=MicrobialGenomes
Conserved Domain	Domini conservati	https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi
Grammx	Docking	http://vakser.compbio.ku.edu/resources/gramm/grammx/
PatchDock	Docking	http://bioinfo3d.cs.tau.ac.il/PatchDock/
FireDock	raffinamento output dei software di docking Grammx e PatchDock	http://bmm.crick.ac.uk/~SwarmDock/submit.cgi
SwarmDock	Docking	http://bioinfo3d.cs.tau.ac.il/FireDock/index.html
Psiver	interfacce conservate, possibili interfacce di interazione	http://tardis.nbio.go.jp/PSIVER/theory.html
3D Ligand Site	Previsione sito di binding	http://www.sbg.bio.ic.ac.uk/3dligandsite/
Allopred	Previsione sito di binding	http://www.sbg.bio.ic.ac.uk/allopred/home
Chimera	Visualizzazione PDB e produzione delle immagini	http://www.rbvi.ucsf.edu/chimera

Generazione e valutazione dei modelli.

Tramite il server I-Tasser [10] sono stati generati differenti modelli 3D per entrambe le proteine, di default i modelli possono essere massimo 5. Questo server si serve del *threading* utilizzando una libreria per generare *template* per la proteina input. Inoltre fornisce informazioni aggiuntive come i ligandi (substrati, cofattori o interattori proteici), i residui attivi e di binding e l'eventuale funzione enzimatica. I modelli ottenuti sono valutati autonomamente dalla suite di I-Tasser attraverso uno score (I-Score). Un'altra valutazione è stata ottenuta attraverso Rampage che genera un grafico di Ramachandran [11]. In questo caso l'algoritmo calcola gli angoli diedri per ogni residuo amminoacidico classificandone i valori come favoriti, permessi o anomali. Naturalmente minore è la percentuale di residui con valori anomali più è alta la qualità del modello. I modelli ottenuti sono sottomessi a Galaxy Gemini [14] che prevede se più unità identiche fra loro si assemblano per formare l'enzima funzionale.

Studio dei residui di binding per cofattori e substrati e dei residui attivi dell'enzima BCAD.

Al fine di prevedere gli AA (amminoacidi) di legame dell'enzima BCAD per il substrato e/o i cofattori, sono stati utilizzati i software 3DLig, Conserved Domain e Binding Site. Il primo di questi grazie a Phyre2 elabora un modello 3D che è automaticamente sottomesso su 3DLig, che prevede i siti di legame per il FAD [12]. Conserved Domain è un software che si attiva di default con Delta Blast e riporta il sito di legame per il FAD, il sito di legame per il substrato e il sito catalitico [13]. Altri AA di legame con il FAD sono previsti da Binding Site (I-Tasser) [10].

Docking.

I residui che potenzialmente interagiscono fra PirA e proteina target sono stati computati tramite docking (Gramm-X, PatchDock, FireDock e SwarmDock) e visualizzati tramite Chimera. Il docking è effettuato fra il modello migliore di PirA e il

modello migliore della BCAD. I software di docking scelti usano algoritmi differenti. Gramm-X è basato sul rigid-body docking ed utilizza una curva di potenziale intermolecolare di Lennard-Jonnes [15], mentre PatchDock usa un algoritmo di riconoscimento di forma basato sulla corrispondenza geometrica di superfici proteiche concave, convesse o piatte [16]. Queste metodiche di computazione presentano dei limiti e possono essere perfezionate utilizzando algoritmi di docking flessibile, come quello di FireDock [17]. Quest'ultimo software è usato per perfezionare gli output di Gramm-X e PatchDock e fornisce una variazione dell'energia libera dopo il legame fra le due proteine input. Anche SwarmDock fornisce questo risultato e usa un algoritmo di docking flessibile [18]. Un altro modo di perfezionare i risultati del docking è identificare gli amminoacidi di contatto e/o legame e valutare la probabilità che questi ultimi siano accessibili e possano far parte d'interfacce di legame. PSiver predice le interfacce partendo dalla sola sequenza primaria e può essere utilizzato per filtrare i risultati di docking [19].

Costruzione delle immagini e interpretazione dei risultati.

Confrontando tutti i residui identificati con i metodi precedenti, si è proceduto in questa maniera al fine di estrapolare i risultati e riportarli in un'unica figura:

1. i residui di legame per substrato e cofattori sono stati verificati con la sequenza ottenuta da PSiver, eliminando residui che non ricadono in regioni accessibili;
2. Lo stesso procedimento è stato usato con i residui identificati dai risultati del docking analizzati con Find Clashes/Contacts e FindHbond di Chimera[20]. Queste ultime due funzioni sono state utilizzate anche per identificare gli AA di legame nell'omopolimero.
3. Tutti i residui sono stati riportati in un'unica immagine in 3D.

Risultati

Inizialmente le sequenze della proteina PirA e della BCAD sono state sottomesse sulla versione on-line di I-Tasser, e come descritto in materiali e metodi, i modelli ottenuti sono stati validati con Rampage. Il C-score di I-Tasser varia da -5 a 2, con risultati maggiori per modelli di maggiore qualità. Il C-score per BCAD è 0,06 mentre per la pirina è 0,22. Il programma Rampage, invece produce dei grafici di Ramachandran, che mettono in relazione gli angoli diedri(Ψ e Φ) degli amminoacidi (AA) nel modello in 3D prodotto. Naturalmente, le combinazioni dei due angoli teoricamente possibili, considerando l'ingombro sterico degli AA, possono essere impossibili in realtà. Il grafico è diviso in zone delimitate da contorno che individuano regioni in cui i valori di angoli sono favoriti, regioni in cui sono permessi e regioni in cui i valori non sono permessi. Per la deidrogenasi il numero di residui in regioni favorite è il 79,9%, in regioni permesse 12,8% e in regioni anomale 7,3%. Per la PirA la percentuale di residui in regioni favorite è 82,4%, in regioni permesse 10,8% e in regioni anomale 6,8%. I modelli generati da Galaxy Gemini sono visualizzati con chimera (figura 2), che attraverso il tool FindHbond, permette di ricavare i legami idrogeno.

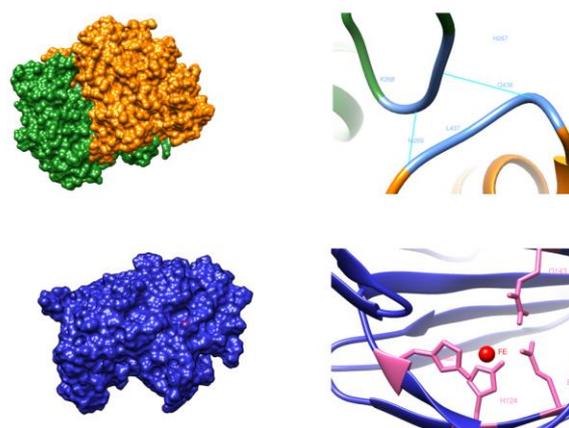


Figura 2: In alto a sinistra modello dimerico della BCAD (catene A e B), i due monomeri identici sono colorati in arancione e verde. In alto a destra dettaglio di due dei quattro legami idrogeno presenti nella forma dimerica. Gli altri due cadono sugli stessi residui a catene invertite. In basso a sinistra modello monometrico della pirina, che contiene un atomo di ferro (in dettaglio in basso a sinistra).

I legami sono in tutto 4, la catena A (BCAD-A residui LYS 268, HIS 267), forma con la catena B (BCAD-B residui GLU439, PHE 436) due legami idrogeno. Altri due legami sono formati dai residui invertiti (catena A GLU439, PHE 436 e catenaB LYS 268, HIS 267). I legandi delle due proteine sono previsti dal software Coach, contenuto in I-Tasser. La pirina presenta come ligandi un cofattore che è il ferro. Questo indica che questa proteina possa avere funzione di sensore redox capace di inibire in maniera specifica l'attività di altre proteine. I due ligandi proteici identificati, invece, sono due cupine. Anche se di basso score (più è alto migliore è il legame), considerando che le due specie sono molto differenti dagli streptomiceti, potrebbe esserci un'interazione fra cupine dello streptomicete e pirina. La BCAD lega come cofattore il FAD. Tale molecola è un mediatore redox utilizzato dalla proteina per ridurre o ossidare il substrato. Infine utilizzando differenti software, sono stati individuati i residui dell'enzima BCAD che potenzialmente legano il substrato. In particolare i software utilizzati sono 3DLig/Phyre2, Conserved Domain, Ligand Binding Site di I-Tasser. Una volta analizzati i siti di legame per substrato e cofattori della BCAD, bisogna capire se questa proteina può interagire con PirA. Per ottenere questo risultato si utilizzano tre software di docking (Gramm-X, PatchDock e SwarmDock) e un ulteriore software (FireDock) è utilizzato per rifinire i risultati di Gramm-X e PatchDock. Questi programmi utilizzano approcci e algoritmi differenti, tuttavia il risultato è sempre una predizione della capacità delle proteine di legarsi. Una volta ottenuti i complessi attraverso Chimera (funzione Find Clashes/Contacts) è possibile individuare i residui che formano contatti fra le due proteine. Mentre SwarmDock fornisce direttamente il valore dell'abbassamento dell'energia di legame (-34,34 KJ/mol), per gli altri due software il valore è generato da FireDock (-70,34 KJ/mol per Gramm-X e -5,22

K J/mol per PatchDock). L'ultimo programma online usato è PSiver. Quest'ultimo permette di ricavare i residui che con maggior facilità sono accessibili rispetto alla struttura tridimensionale della proteina. Di fatto è un metodo alternativo al docking e all'analisi del sito di binding per substrato e cofattore, per verificare se i contatti previsti ricadono in regioni accessibili della proteina. Il risultato migliore generato con Gramm-X è riportato in figura 3.

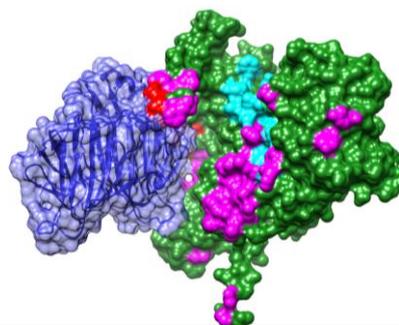


Figura 3: Interazione migliore prevista (Gramm-X) fra PirA e BCAD. In questo caso l'enzima è rappresentato come uno solo dei monomeri. La pirina attraverso una particolare regione si inserisce in prossimità del canale formato da alcune strutture ad alfa elica, giungendo in prossimità del sito attivo (regione color ciano). I residui di contatto fra PirA (blu) e BCAD catena A (verde) sono colorati in rosso, mentre in viola sono evidenziati i residui di contatto fra monomero A e B.

Discussione

L'obiettivo di questo lavoro è comprendere il meccanismo biofisico e molecolare grazie al quale la proteina PirA potrebbe regolare l'attività dell'enzima BCAD. Per analizzare in maniera soddisfacente i risultati computazionali ottenuti bisogna tener presente i limiti e i vantaggi di ogni algoritmo utilizzato. Ad esempio i modelli generati da I-Tasser sono valutati autonomamente dallo stesso ma anche grazie a RAMPAGE. In questo modo è stato selezionato il modello migliore in accordo con i risultati. Sicuramente la proteina BCAD funzionale è costituita da due monomeri di BCAD che probabilmente fanno parte di un complesso che contiene catene polipeptidiche differenti. Nella struttura proteica del dimero di BCAD si possono notare due eliche parallele che formano un'apertura nella struttura globulare della proteina e vari foglietti beta.

Tale apertura è un canale d'ingresso che conduce al sito attivo e un'interfaccia di interazione con altre proteine. Ogni monomero presenta un altro canale, quindi il dimero che ha una forma allungata contiene in tutto 3 canali, uno all'interfaccia fra i monomeri al centro e due alle estremità. Lo studio funzionale dell'enzima rivela come sicuramente questa proteina lega il FAD utile come mediatore redox e acidi grassi attivati dal legame tioestere con il coenzima A. Quest'enzima è fondamentale poiché è coinvolto nelle reazioni del catabolismo degli acidi grassi, ovvero la beta ossidazione. In questo caso la via metabolica porta alla formazione di acetil-CoA, intermedio del metabolismo primario, e potrebbe indirettamente controllare il nodo metabolico responsabile della sintesi di etilmalonil-CoA. PirA contiene un atomo di ferro e, dato che la reazione catalizzata dalla BCAD è un'ossidazione, possiamo presupporre che possa avere funzione di sensore redox che traduce il segnale a più livelli, controllando in particolare l'attività della BCAD e quindi la sintesi dei precursori. Per verificare quest'ipotesi bisogna analizzare i risultati del Docking. In questo caso bisogna evidenziare come i risultati ottenuti dai differenti software siano molto differenti, ciò può essere spiegato dai differenti algoritmi. PatchDock-FireDock restituiscono il risultato più basso probabilmente poiché le simulazioni sono state effettuate monomero di BCAD contro monomero PirA utilizzando un algoritmo basato sulla corrispondenza geometrica. Questo approccio considera la proteina come un solido trascurando le modifiche conformazionali che possono insorgere in seguito ad il legame con altre proteine, ligandi e substrati. SwarmDock restituisce un risultato intermedio mentre il migliore è stato registrato con GrammX-FireDock. In quest'ultimo caso l'interazione della pirina dovrebbe avvenire a cavallo delle eliche in modo tale che PirA penetri nell'enzima avvicinandosi al sito attivo. Attraverso questo meccanismo può essere

spiegato come lo stato redox influisca su PirA che a sua volta modula l'attività della BCAD. Anche secondo gli altri software, benché i risultati siano diversi, i residui di contatto fra PirA e BCAD cadono sempre in prossimità dei siti di binding e del sito attivo dell'enzima. I-TASSER prevede come la prima legni due proteine la Faseolina di *Phaseolus vulgaris* e BacB di *Bacillus subtilis* implicata nella sintesi di antibiotici. Questo dato suggerisce che la superfamiglia delle proteine definite cupine correlate con le pirine possano formare un network di interazioni molecolari capace di controllare nodi metabolici correlati con il metabolismo secondario e risposte a segnali di specie vegetali. Fra le cupine presenti nel genoma di *S. ambifaciens* troviamo fattori trascrizionali AraC-like ed Xre-like entrambi identificati già in *S. coelicolor* in cui se inattivati si registra un calo della sintesi di antibiotici. Altre cupine possono legare ormoni vegetali come le auxine oppure avere attività enzimatica di quercetina diossigenasi. Un dominio cupinico è presente anche in una fosfomannosio isomerasi. Enzimi come quest'ultimo sono stati identificati come importanti per la sintesi di antibiotici..

Conclusioni

Lo studio computazionale presentato evidenzia come l'interazione fra PirA e BCAD non solo possa essere funzionale, ma anche capace di regolare la sintesi di spiramicina controllando la produzione dei precursori degli antibiotici. In questo meccanismo sicuramente è implicato lo stato redox della cellula in cui probabilmente la pirina funge da sensore e modulatore. Comprendere il ruolo delle pirine in questione ma anche degli altri tre geni in *S. ambifaciens* è fondamentale sia per i risvolti applicativi che potrebbero esserci per quanto riguarda la produzione di molecole di interesse farmaco-industriale ma anche per avere ulteriori dati riguardo all'evoluzione molecolare della stessa proteina.

References

- [1] Pinnert-Sindico. Une nouvelle espèce de *Streptomyces* productrice d'antibiotiques: *Streptomyces ambofaciens*. n. sp. caractères culturaux. *Ann Inst Pasteur Paris*, (1954);
- [2] Stray-P B. Toxoplasmosis in pregnancy. *Baillieres Clin Obstet Gynaecol.*, (1993);
- [3] S. Hackl and A. Bechthold. The Gene *bldA*, a Regulator of Morphological Differentiation and Antibiotic Production in *Streptomyces*, *Arch. Pharm.*, (2015);
- [4] Ben Shen. Biosynthesis of Aromatic Polyketides, *Topics in Current Chemistry*, (2000);
- [5] S. Donadio, S., G. Marino. *Biotechnologie microbiche*, Casa Editrice Ambrosiana, (2008);
- [6] Adams MA1, Jia Z. Structural and biochemical evidence for an enzymatic quinone redox cycle in *Escherichia coli*: identification of a novel quinol monooxygenase, *J Biol Chem.* (2005);
- [7] Dechend R1, Hirano F, Lehmann K, Heissmeyer V, Ansieau S, Wulczyn FG, Scheiderei C, Leutz A. The Bcl-3 oncoprotein acts as a bridging factor between NF-kappaB/Rel and nuclear co-regulators. *Oncogene.* (1999);
- [8] Soo PC1, Horng YT, Lai MJ, Wei JR, Hsieh SC, Chang YL, Tsai YH, Lai HC. Pirin regulates pyruvate catabolism by interacting with the pyruvate dehydrogenase E1 subunit and modulating pyruvate dehydrogenase activity. *J Bacteriol.* (2007);
- [9] De Huang, Tingting Li, Xinghua Li, Long Zhang, Linchong Sun, Xiaoping He, Xiuying Zhong, Dongya Jia, Libing Song, Gregg L. Semenza, Ping Gao, and Huafeng Zhang. HIF-1-Mediated Suppression of Acyl-CoA Dehydrogenases and Fatty Acid Oxidation Is Critical for Cancer Progression, *Cell Reports*, (2014);
- [10] J Yang, R Yan, A Roy, D Xu, J Poisson, Y Zhang. The I-TASSER Suite: Protein structure and function prediction. *Nature Methods*, (2015);
- [11] Lovell, I.W. Davis, W.B. Arendall III, P.I.W. de Bakker, J.M. Word, M.G. Prisant, J.S. Richardson & D.C. Richardson, Structure validation by C α geometry: ϕ/ψ and C β deviation. *Proteins: Structure, Function & Genetics*, (2002);
- [12] Wass MN, Kelley LA, Sternberg MJ. 3DLigandSite: predicting ligand-binding sites using similar structures. *NAR* 38 Suppl, (2010);
- [13] Grzegorz M Boratyn, Alejandro A Schäffer, Richa Agarwala, Stephen F Altschul, David J Lipman, and Thomas L Madden. Domain enhanced lookup time accelerated BLAST. *Biol Direct.* (2012);
- [14] Hasup Lee, Hahnbeom Parky, Junsu Koz and Chaok Seok. GalaxyGemini: a web server for protein homo-oligomer structure prediction based on similarity, *Bioinformatics, Oxford Journals*, (2013).
- [15] Andrey Tovchigrechko and Ilya A. Vakser. GRAMM-X public web server for protein-protein docking, *Nucleic Acids Research*, (2006);
- [16] Dina Schneidman-Duhovny, Yuval Inbar, Ruth Nussinov, and Haim J. Wolfson. PatchDock and SymmDock: servers for rigid and symmetric docking, *Nucleic Acids Research*, (2005);
- [17] Efrat Mashlach, Dina Schneidman-Duhovny, Nelly Andrusier, Ruth Nussinov and Haim J. Wolfson, "FireDock: a web server for fast interaction refinement in molecular docking", *Nucleic Acids Research*, (2008);
- [18] Mieczyslaw Torchala, Iain H. Moal, Raphael A. G. Chaleil, Juan Fernandez-Recio and Paul A. Bates. SwarmDock: a server for flexible protein-protein docking, *Bioinformatics, Oxford Journals*, (2013);
- [19] Murakami, Y. and Mizuguchi, K. Applying the Naive Bayes classifier with kernel density estimation to the prediction of protein-protein interaction sites, *Bioinformatics*, (2010);
- [20] 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).