



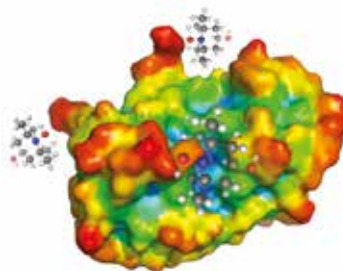
# UNA PICCOLA SONDA PARAMAGNETICA PER IDENTIFICARE SITI ATTIVI IN PROTEINE (A CACCIA DI *HOT SPOT* PROTEICI)

STUDI NMR DI **PROTEINE** IN SOLVENTI RESI PARAMAGNETICI DALLA PRESENZA DI RADICALI LIBERI STABILI E NEUTRI, COME IL **TEMPOLO**, PERMETTONO DI ACQUISIRE INFORMAZIONI SULLA DISTRIBUZIONE DEGLI **HOT SPOT** SULLA SUPERFICIE PROTEICA, FORNENDO UN FILTRO SPERIMENTALE A STUDI DI **SCREENING VIRTUALE** DI FARMACI

Che la Vita sia, se osservata alla risoluzione dell'atomo, nient'altro che un intricato dialogo tra molecole grandi e piccole, è da molto tempo del tutto noto. Ciò nonostante, meno noti sono i modi con i quali tale dialogo avviene, specialmente a media e lunga distanza. D'altra parte, in una situazione di affollamento di molecole, quale quello esistente all'interno delle cellule, tali meccanismi devono essere sufficientemente specifici per minimizzare il ruolo della casualità come fattore determinante per gli incontri intermolecolari che sono alle base dei processi biologici.

La possibilità, quindi, di interferire con tale dialogo molecolare, alterando la rete delle interazioni chimiche all'interno del vitale miscuglio molecolare caratterizzante ciascun organismo, costituisce una fase fondamentale dalla quale fare emergere strategie farmacologiche innovative.

Nella cosiddetta attuale "era post-genomica" della ricerca, la nostra capacità di caratterizzare, anche strutturalmente, i singoli componenti del citato miscuglio molecolare è aumentata in modo enorme, suggerendo ottimistiche previsioni per significativi progressi nella cura di molte patologie, dal cancro alle infezioni batteriche e virali. Infatti, abbiamo a



disposizione un esteso repertorio strutturale per indagare quale siano le caratteristiche molecolari che stabilizzano l'associazione tra biopolimeri o tra questi e le piccole molecole organiche, PMO, tramite il pubblico accesso alla *Protein Data Bank* [1]. Così, oltre 100.000 strutture di proteine libere o legate ad altre specie molecolari sono adesso disponibili, costituendo una solida base per fare predizioni sui possibili siti di interazione delle superfici proteiche [2]. A questo proposito è da notare che la distribuzione delle regioni della superficie proteica che possono essere rilevanti per l'interazione intermolecolare è stata a lungo considerata a partire da strutture proteiche statiche. Solo nell'ultimo decennio la presenza sulla superficie proteica di siti transienti in grado di legare PMO, è stata messa

in evidenza principalmente per mezzo di studi NMR in presenza di sonde paramagnetiche legate covalentemente alla proteina [3] o presenti come co-soluto [4].

La marcata variabilità degli incrementi paramagnetici al rilassamento, PRE, osservati in presenza della sonda paramagnetica come co-soluto nei sistemi fino ad oggi studiati, suggerisce come l'interazione tra lo spin elettronico ed i circostanti spin nucleari possa seguire un meccanismo a "sfera interna", pur in presenza di deboli interazioni proteina-sonda. In questo caso, i PRE misurati possono essere analizzati mediante le equazioni di Solomon e Bloembergen [5].

Come mostrato in Eq. (1), le velocità di rilassamento osservate in presenza della sonda paramagnetica, libera o legata,  $(1/T_1)_{para}$  sono date dalla somma di due contributi  $(1/T_1)_{el}$  e  $(1/T_1)_{nuc}$ , rispettivamente derivanti dall'interazione scalare e dipolare tra gli spin elettronici e nucleari e dall'interazione dipolare tra gli spin nucleari che determina il rilassamento nel sistema diamagnetico.

Se le condizioni  $\omega_{I_{TC}} \ll 1$ ,  $\omega_{S_{TC}} \gg \omega_{I_{TC}}$ , e  $J(\omega_S \pm \omega_I) \approx J(\omega_S)$  sono soddisfatte, il contributo paramagnetico alle velocità di rilassamento nucleare è dato dalle Eq 2, 3, secondo

Nell'esempio riportato in Fig. 1, si possono facilmente distinguere le differenti intensità della perturbazione causata dalla presenza del TEMPOLO sui segnali  $^1\text{H}$ - $^{15}\text{N}$  HSQC di una soluzione acquosa di una chemochina, CXCL12, lunga 70 residui e ampiamente strutturata.

È da sottolineare che tra tutte le sonde paramagnetiche che sono state usate per ottenere informazioni sulla caratterizzazione strutturale della superficie delle proteine, il TEMPOLO, ovvero il 2,2,6,6-tetrametil-4-piperidino-*N*-ossile, ha rappresentato una pietra miliare, principalmente per merito della sua elevata stabilità chimica, dell'alta solubilità in acqua ed in solventi meno polari e dell'assenza di cariche elettriche che potrebbero indurre interazioni intermolecolari preferenziali. Dall'utilizzo iniziale del TEMPOLO per discriminare tra nuclei proteici interni od esposti [6], progressivamente abbiamo notato che la sonda paramagnetica dava perturbazioni paramagnetiche principalmente dipendenti da fattori dinamici, quali l'effettiva accessibilità di limitate regioni della superficie proteica nei suoi confronti [7]. Infatti, dall'analisi combinata dei profili delle perturbazioni paramagnetiche indotte dalla presenza di TEMPOLO e, mediante simulazioni di *Molecular Dynamics* (MD), dei tempi di vita dei legami ad idrogeno intramolecolari coinvolgenti i gruppi ammidici del *backbone* delle proteine studiate, emerge quanto sia determinante per l'entità dei PRE osservati l'interazione tra il gruppo *N*-ossile del TEMPOLO e gli idrogeni ammidici. Così, tali idrogeni, che risultano poco impegnati in legami ad idrogeno intramolecolari ed in posizioni esposte al solvente sulla superficie proteica, risultano i maggiormente influenzati dal TEMPOLO, confermando come la formazione di legami ad idrogeno intermolecolari costituisca uno dei principali meccanismi di approccio tra la sonda paramagnetica e la proteina. In Fig. 2 sono mostrate le caratteristiche steriche del legame ad idrogeno che si instaura tra TEMPOLO e scheletro covalente della proteina, a sostegno di quanto sia forte l'interazione dipolare a breve distanza tra lo spin dell'elettrone spaiato, residente al 61% sull'azoto e al 39% sull'ossigeno del gruppo *N*-ossile [8], e lo spin dell'idrogeno ammidico della proteina. A questo proposito è da notare che, in sistemi nei quali un analogo del TEMPOLO è legato covalentemente alla proteina, sono rilevabili PRE a distanze tra gli spin I ed S fino a 20 Å.

$$\left(\frac{1}{T_j}\right)_{\text{parata}} = \left(\frac{1}{T_j}\right)_{\text{di}} + \left(\frac{1}{T_j}\right)_{\text{sc}}, \quad j = 1, 2 \quad (1)$$

$$\left(\frac{1}{T_j}\right)_{\text{di}} = \frac{2}{15} \left(\frac{\mu_0}{4\pi}\right)^2 \frac{\gamma_I^2 \gamma_S^2 \pi^2 S(S+1)}{r^6} \chi \left( \frac{3\tau_{lc}}{1 + \omega_I^2 \tau_{lc}^2} + \frac{7\tau_{lc}}{1 + \omega_S^2 \tau_{lc}^2} \right) + \frac{S(S+1)}{3} \left(\frac{\alpha}{\hbar}\right)^2 \chi \left( \frac{2\tau_{lc}}{1 + \omega_I^2 \tau_{lc}^2} \right) \quad (2)$$

$$\left(\frac{1}{T_j}\right)_{\text{sc}} = \frac{1}{15} \left(\frac{\mu_0}{4\pi}\right)^2 \frac{\gamma_I^2 \gamma_S^2 \pi^2 S(S+1)}{r^6} \chi \left( 4\tau_{lc} + \frac{3\tau_{lc}}{1 + \omega_I^2 \tau_{lc}^2} + \frac{13\tau_{lc}}{1 + \omega_S^2 \tau_{lc}^2} \right) + \frac{S(S+1)}{3} \left(\frac{\alpha}{\hbar}\right)^2 \chi \left( \tau_{lc} + \frac{\tau_{lc}}{1 + \omega_I^2 \tau_{lc}^2} \right) \quad (3)$$

$$\frac{1}{\tau_{jc}} = \frac{1}{\tau_M} + \frac{1}{\tau_R} + \frac{1}{T_{jS}}, \quad j = 1, 2 \quad (4)$$

$$\frac{1}{\tau_{je}} = \frac{1}{\tau_M} + \frac{1}{T_{jS}}, \quad j = 1, 2 \quad (5)$$

le condizioni riportate nelle Eq. 4 e 5, in cui  $\tau_{jc}$  rappresenta il tempo effettivo di correlazione,  $\tau_{jS}$  è il tempo di rilassamento elettronico longitudinale o trasversale,  $\tau_R$  è il tempo di correlazione rotazionale del complesso,  $\tau_M$  è il tempo di vita dell'addotto intermolecolare,  $\tau_{je}$  è il tempo di correlazione scalare,  $r$  è la distanza tra gli spin elettronici, I, e nucleari, S;  $\omega_I$  e  $\omega_S$  sono le frequenze di Larmor per gli spin I ed S;  $\gamma_I$  e  $\gamma_S$  sono rispettivamente il rapporto giromagnetico del nucleo osservato e dello spin elettronico,  $\mu_0$  è la costante di permeabilità magnetica,  $a$  è la costante di accoppiamento scalare dell'interazione di Fermi e  $\chi$  è la frazione

di popolazione, per ciascun spin nucleare, che si trova in stretta vicinanza dello spin elettronico. Dalle equazioni sopra riportate risulta evidente la forte dipendenza esistente tra PRE misurati e distanza tra gli spin I e S.

Un'accurata valutazione quantitativa dei PRE potrebbe essere fatta sulla base di misure di velocità di rilassamento, ma una stima di quali siano i nuclei di una proteina maggiormente esposti a frequenti incontri ravvicinati con la sonda paramagnetica può essere fatta più velocemente, misurando le variazioni delle intensità relative dei segnali NMR negli spettri ottenuti in assenza ed in presenza della sonda.

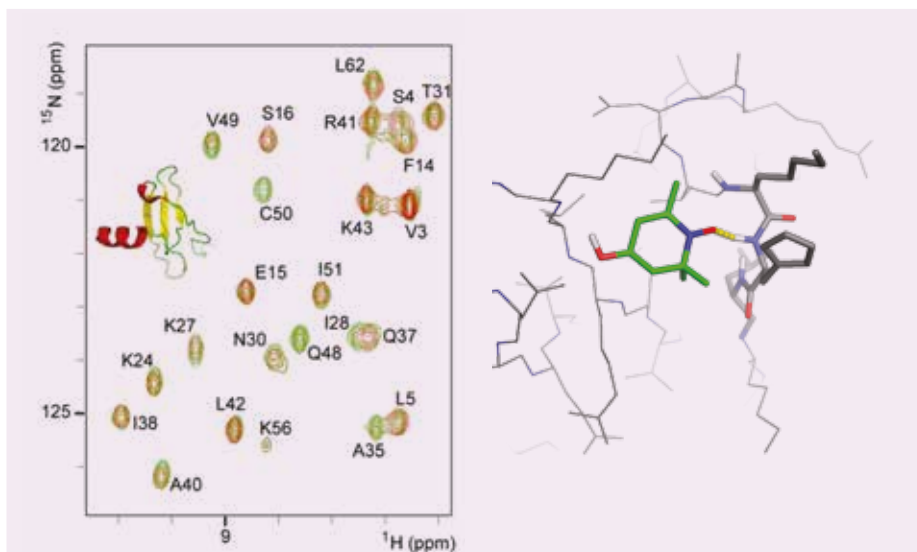


Fig. 1 Sovrapposizione degli spettri  $^1\text{H}$ - $^{15}\text{N}$  HSQC ottenuti per una soluzione acquosa 0,5 mM diamagnetica (picchi verdi) e paramagnetica (picchi rossi) di CXCL12, una chemochina la cui struttura è riportata nella rappresentazione a nastro; le etichette sui segnali denotano la posizione nella sequenza di ciascun gruppo ammidico del *backbone* proteico

Fig. 2 Il legame ad idrogeno che si instaura tra il gruppo *N*-ossile del TEMPOLO, evidenziato dalla colorazione verde dei carboni, ed un generico gruppo ammidico del *backbone* proteico determina una breve distanza tra i nuclei osservati I e lo spin spaiato S

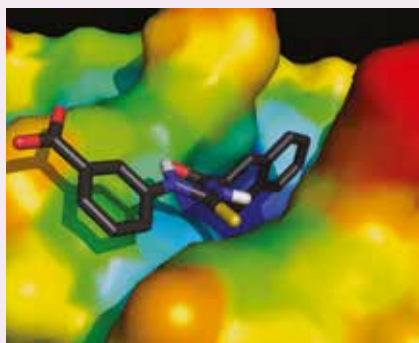


Fig. 3

Il posizionamento della PMO con PubChem CID 6945315 in una tasca transiente individuata sulla superficie di CXCL12 in accordo con le perturbazioni paramagnetiche indotte dal TEMPOLO

L'intensità della perturbazione paramagnetica è di fondamentale importanza, perché suggerisce come possano essere evidenziate anche interazioni transienti all'interno della dinamica di interazione TEMPOLO-proteina [9]. La possibilità di osservare le PRE direttamente sui nuclei che sono coinvolti nell'interazione dipolare I-S mediante l'analisi delle attenuazioni indotte dalla presenza del TEMPOLO sull'intensità dei segnali ammidici di spettri NMR di tipo  $^1\text{H}$ - $^{15}\text{N}$  HSQC, risulta essere un modo molto sensibile per valutare l'accessibilità alla sonda paramagnetica da parte di specifiche regioni, anche transienti, della superficie proteica. Inoltre, il fatto che il TEMPOLO abbia tutte le caratteristiche per essere utilizzato nella ricerca per frammenti di siti di legame sulle superfici proteiche (vedi PubChem CID 137994), soddisfacendo alla cosiddetta "regola del tre" [10], suggerisce che questa sonda paramagnetica possa essere utilizzata come un potente strumento per individuare i punti di massima accessibilità delle superfici proteiche, generalmente note come *hot spot*. Così, questa informazione, sperimentalmente derivata, può costituire un filtro razionale per validare interazioni tra proteine e PMO predette da tecniche di *screening* virtuale. La possibilità di disporre di una tecnica sperimentale per identificare siti transienti di interazione tra proteine e potenziali frammenti di molecole farmacologicamente attive apre nuove strade per la progettazione razionale di nuovi farmaci, particolarmente nel settore degli antitumorali che agiscono interferendo con interazioni proteina-proteina che sono alla base dello sviluppo e proliferazione di neoplasie. Infatti, in molti casi è stato osservato

che i siti di interazione di una PMO con la proteina bersaglio hanno la caratteristica di non essere presenti sulla superficie statica della proteina sia nella sua forma libera, sia in quella legata ad altri partner proteici [11].

In una nostra recente pubblicazione [4] abbiamo messo in evidenza quanto l'uso combinato dell'analisi dei profili paramagnetici indotti dal TEMPOLO e delle traiettorie di MD abbia permesso di trovare la base strutturale per l'affinità esistente tra una PMO (PubChem CID: 6945315) ed una proteina ritenuta responsabile della proliferazione di processi metastatici, la CXCL12 [12]. In Fig. 3 è mostrato come una PMO (PubChem CID 6945315) si posizioni sulla superficie proteica in una tasca transiente che non è identificabile in nessuna delle 19 strutture ad oggi disponibili per CXCL12 nella Protein Data Bank, ma che corrisponde ad una zona in cui l'attenuazione paramagnetica è particolarmente forte. L'interazione tra TEMPOLO e gruppi ammidici del *backbone* proteico esposti al solvente è sufficientemente debole da non causare variazioni nello spostamento chimico anche dei segnali maggiormente interessati dalla perturbazione paramagnetica, ma, al tempo stesso, permette una accurata mappatura della distribuzione degli *hot spot* sia stabili sia transienti sulla superficie delle proteine. Anche nel caso di corte sequenze di RNA, il TEMPOLO è stato utilizzato con successo per evidenziare le regioni di mancato accoppiamento A-U e C-G delle basi nucleiche via legame ad idrogeno. In questi casi, infatti, i donatori di legame ad idrogeno presenti sulle basi azotate rimangono disponibili all'interazione con il gruppo *N*-ossile della sonda paramagnetica. In questo modo, ad esempio, è stata evidenziata una regione funzionalmente rilevante del dominio *intramolecular stem-loop* dello RNA nucleare U6 [13]. Una limitazione nell'uso di questa sonda paramagnetica, come pure dei suoi analoghi nitrossidici, è rappresentato dalla sua attività redox. Nei sistemi proteici l'eventuale presenza di tioli liberi ed accessibili sulla superficie molecolare, come pure quella di gruppi prostetici, può causare l'ossidazione o riduzione dello *N*-ossile dando luogo a specie diamagnetiche quali, rispettivamente, il catione *N*-ossoammonio o l'anione aminossile, oltre ad alterazioni strutturali del biopolimero in esame. Fatta salva questa limitazione, l'uso della sonda paramagnetica TEMPOLO, fornendo informazioni su caratteristiche dinamiche quali l'accessibilità delle superfici proteiche, espande il già esteso repertorio delle conoscenze

strutturali in una direzione che potrà fornire nuovi indizi per la progettazione di molecole che possano interferire con specifiche interazioni proteina-proteina o proteina-PMO.

## BIBLIOGRAFIA

- [1] F.C. Bernstein *et al.*, *J. Mol. Biol.*, 1977, **112**, 535.
- [2] Y. Yuan *et al.*, *Curr. Pharm. Des.*, 2013, **19**, 2326.
- [3] J. Iwahara, G.M. Clore, *Nature*, 2006, **440**, 1227.
- [4] A. Bernini *et al.*, *Biochim. Biophys. Acta.*, 2014, **1844**, 561.
- [5] I. Solomon, N. Bloembergen, *J. Chem. Phys.*, 1956, **25**, 261.
- [6] G. Esposito *et al.*, *J. Mol. Biol.*, 1992, **224**, 659.
- [7] A. Bernini *et al.*, *Progr. NMR Spectr.*, 2009, **54**, 278.
- [8] D. Bordeaux *et al.*, *Z. Naturforsch.*, 1993, **48a**, 117.
- [9] N.J. Anthis, G.M. Clore, *Q. Rev. Biophys.*, 2015, epub head of print DOI: 10.1017/S0033583514000122
- [10] J. Zuegg, M.A. Cooper, *Curr. Top. Med. Chem.*, 2012, **12**, 1500.
- [11] J.A. Wells, C.L. McClendon, *Nature*, 2007, **450**, 1001.
- [12] [http://en.wikipedia.org/wiki/Stromal\\_cell-derived\\_factor\\_1](http://en.wikipedia.org/wiki/Stromal_cell-derived_factor_1)
- [13] V. Venditti *et al.*, *Nucleic Acids Res.*, 2008, **36**, e20.

### Paramagnetic Fragment Based Search for Protein Active Sites (Hunting for Protein Hot Spots)

Solvent spin-labeling with stable and neutral paramagnets, such as TEMPOL, allows a unique NMR strategy to localize protein surface hot spots. This information is of primary relevance as an experimental filter to fragment-based virtual screening in drug discovery.

NERI NICCOLAI

DIPARTIMENTO DI BIOTECNOLOGIE, CHIMICA  
E FARMACIA  
UNIVERSITÀ DEGLI STUDI DI SIENA

NERI.NICCOLAI@UNISI.IT