



UNIVERSITÀ  
DEGLI STUDI  
FIRENZE

Scuola di  
Scienze Matematiche  
Fisiche e Naturali

Corso di Laurea in Chimica  
Curriculum Scienze Chimiche

# **Metodi di denoising innovativi applicati a studi di interazione silice-lisozima**

## **Innovative denoising methods applied to silica- lysozyme interaction studies**

**Relatore**

Enrico Ravera

**Correlatore**

Leonardo Tenori, Lucia Gigli

**Candidato**

Francesco Bruno

# Indice

<b>1</b>	<b>Introduzione</b>	<b>1</b>
1.1	Perché ssNMR di silice e silicati . . . . .	1
1.2	Tecniche per migliorare il segnale . . . . .	4
1.3	Metodi post-acquisizione . . . . .	7
1.4	Lisozima e silice: un rapporto complicato . . . . .	8
<b>2</b>	<b>Materiali e metodi</b>	<b>11</b>
2.1	Preparazione dei campioni . . . . .	11
2.2	Strumentazione . . . . .	12
2.3	Processing degli spettri . . . . .	12
<b>3</b>	<b>Il problema scientifico</b>	<b>14</b>
<b>4</b>	<b>Risultati e discussione</b>	<b>15</b>
4.1	Multivariate Curve Resolution . . . . .	15
4.2	Cadzow . . . . .	17
4.3	Il denoising combinato . . . . .	18
4.4	Discussione sui risultati . . . . .	19
<b>5</b>	<b>Analisi del materiale</b>	<b>27</b>
5.1	La prospettiva del lisozima . . . . .	27
5.2	La prospettiva della silice . . . . .	28
5.3	Il composito . . . . .	29
<b>6</b>	<b>Conclusioni</b>	<b>34</b>
	<b>Bibliografia</b>	<b>35</b>

# 1. Introduzione

## 1.1 Perché ssNMR di silice e silicati

Le tecniche spettroscopiche sono gli strumenti più utilizzati per la caratterizzazione di nuovi materiali. In particolare, la spettroscopia di risonanza magnetica nucleare di stato solido (*ssNMR*) é senza dubbio una delle più utili, dal momento che fornisce informazioni sulla natura e sul comportamento chimico dell'intorno di singoli atomi: essa osserva come la struttura elettronica cambia l'energia delle transizioni di ciascuno spin nucleare.

Sebbene abbia intrinsecamente una risoluzione elevatissima, il problema fondamentale della spettroscopia NMR é la bassa sensibilità, dovuta principalmente a due fattori: la scarsa abbondanza naturale dei nuclei osservati e la piccola differenza di popolazione tra i livelli di spin nucleare.

La sensibilità di una tecnica spettroscopica dipende dalla differenza in energia tra i livelli energetici:

$$P_i = \frac{e^{-E_i/k_B T}}{\sum_j e^{-E_j/k_B T}} \quad (1.1)$$

Nell'NMR la differenza in energia dei livelli di spin dipende dall'intensità del campo magnetico applicato ( $B_0$ ) e dal rapporto giromagnetico ( $\gamma$ ) del nucleo in esame:

$$\Delta E = \hbar \gamma B_0 \quad (1.2)$$

Inserendo nelle equazioni precedenti il valore del rapporto giromagnetico del silicio e moltiplicando per un campo magnetico di intensità ragionevole, si osserva che a temperatura ambiente viene rilevato circa un nucleo attivo su diecimila. Per aumentare il segnale di spin nucleare si può agire sull'intensità del campo magnetico e sulla temperatura del campione.

Altri fattori che influenzano la sensibilità dell'NMR sono il volume del campione, l'arricchimento isotopico dei nuclei attivi, la temperatura dell'elettronica dello strumento e la qualità delle bobine per l'emissione di radiofrequenze. In queste condizioni l'intensità dei segnali risulta bassa, dunque gli spettri presentano un basso rapporto segnale-rumore, senza contare che solo il 4% degli atomi di silicio é magneticamente attivo.

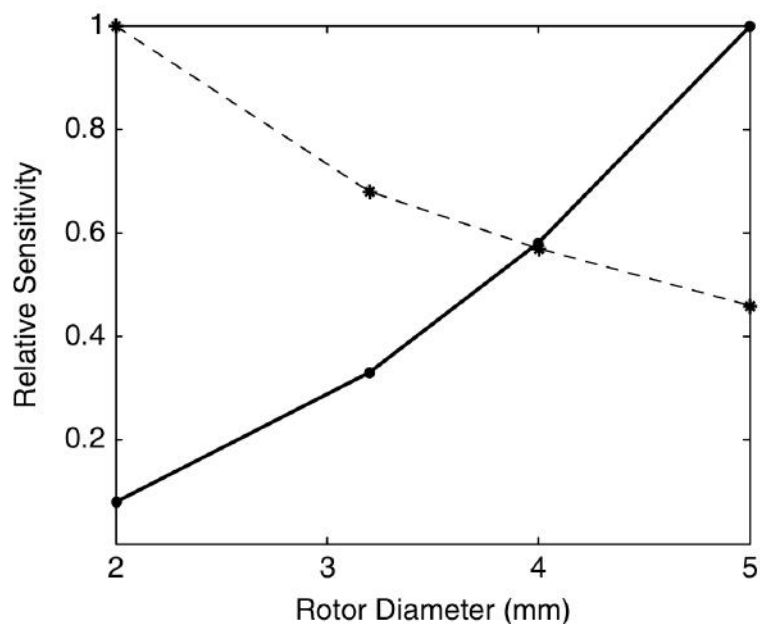
Inoltre, gli spin nucleari sono poco accoppiati con l'ambiente circostante e, piuttosto spesso, gli esperimenti richiedono lunghi tempi di acquisizione anche a causa della scarsa efficienza dei meccanismi di rilassamento.

Negli esperimenti di stato solido si deve anche fare i conti con un consistente allargamento delle linee spettrali, dovuto principalmente alle interazioni che sussistono tra i nuclei nel sistema in esame, come l'anisotropia del chemical shift (*Chemical Shift Anisotropy, CSA*) e l'accoppiamento dipolare (tabella 1.1).

**Tabella 1.1:** Dipendenza spaziale delle interazioni nucleari e loro media rotazionale sotto MAS. Adattata da [4].

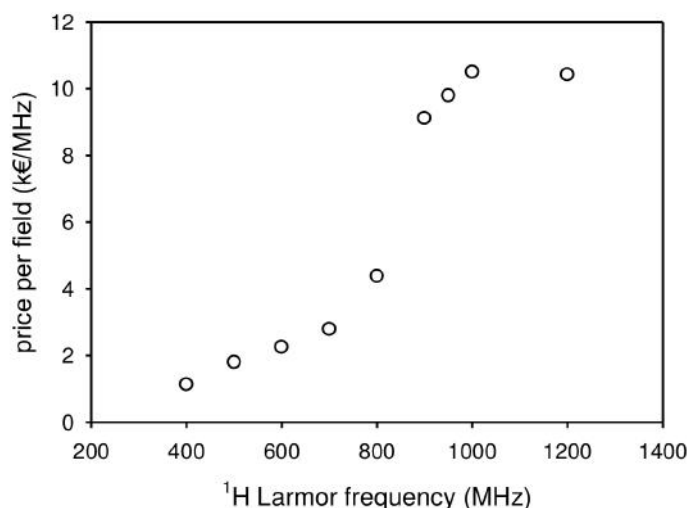
Interazione	Dipendenza dall'orientazione	Si media a
Chemical shielding	$3 \cos^2 \theta - 1$	Chemical shift
Shielding dipolare	$3 \cos^2 \theta - 1$	Shift di pseudocontatto
Dipolo-dipolo eteronucleare	$3 \cos^2 \theta - 1$	0
Quadrupolo nucleare (1° ordine)	$3 \cos^2 \theta - 1$	0
Quadrupolo nucleare (2° ordine)	$35 \cos^4 \theta - 30 \cos^2 \theta + 3$	Non mediato dal MAS ma scalato a $-7/18$ del valore statico. Il valore isotropo si chiama shift quadrupolare
Shielding di contatto	$3 \cos^2 \theta - 1$ , anisotropo solo quando causato da spin-polarizzazione	Shift di contatto

Le interazioni nucleari sono, in generale, anisotrope, ovvero dipendono dalla orientazione della molecola rispetto al campo magnetico. In soluzione, queste interazioni sono mediate dai moti molecolari veloci, perciò possono contribuire solamente al rilassamento (allargamento omogeneo). In uno spettro di polvere i segnali restano invece distribuiti in range di migliaia di Hz (allargamento inhomogeneo). In uno stesso campione possono inoltre essere presenti ambienti leggermente diversi, che provocano un ulteriore allargamento gaussiano dei segnali detto allargamento eterogeneo. Per ridurre l'allargamento inhomogeneo si ricorre allo spinning all'angolo magico (*Magic Angle Spinning, MAS*) e all'impiego di sequenze di disaccoppiamento eteronucleare ad alte prestazioni. Queste tecniche richiedono una rotazione meccanica del campione molto veloce (dell'ordine della frequenza dell'interazione che si intende eliminare), che riduce la quantità di campione utilizzabile. Quindi, l'intensità dei segnali può essere ancora non sufficiente per la caratterizzazione di un materiale (figura 1.1).



**Figura 1.1:** Confronto della sensibilità dei probe per diverse dimensioni del rotore impiegato: la sensibilità relativa dell'intero rotore (linea continua) e sensibilità relativa per unità di volume (linea tratteggiata). [15]

Aumentare il campo magnetico applicato è attualmente possibile fino a valori di 28 T. Il costo di uno strumento cresce con il campo magnetico (un magnete da 16.4 T costa circa 2 milioni di euro, uno da 20 T circa 8 milioni di euro e uno da 28 T circa 13 milioni di euro; grafico in figura 1.2), rendendo poco accessibile la spettroscopia di risonanza magnetica.



**Figura 1.2:** Andamento del prezzo di acquisto di strumenti NMR commerciali con magneti superconduttori. [2]

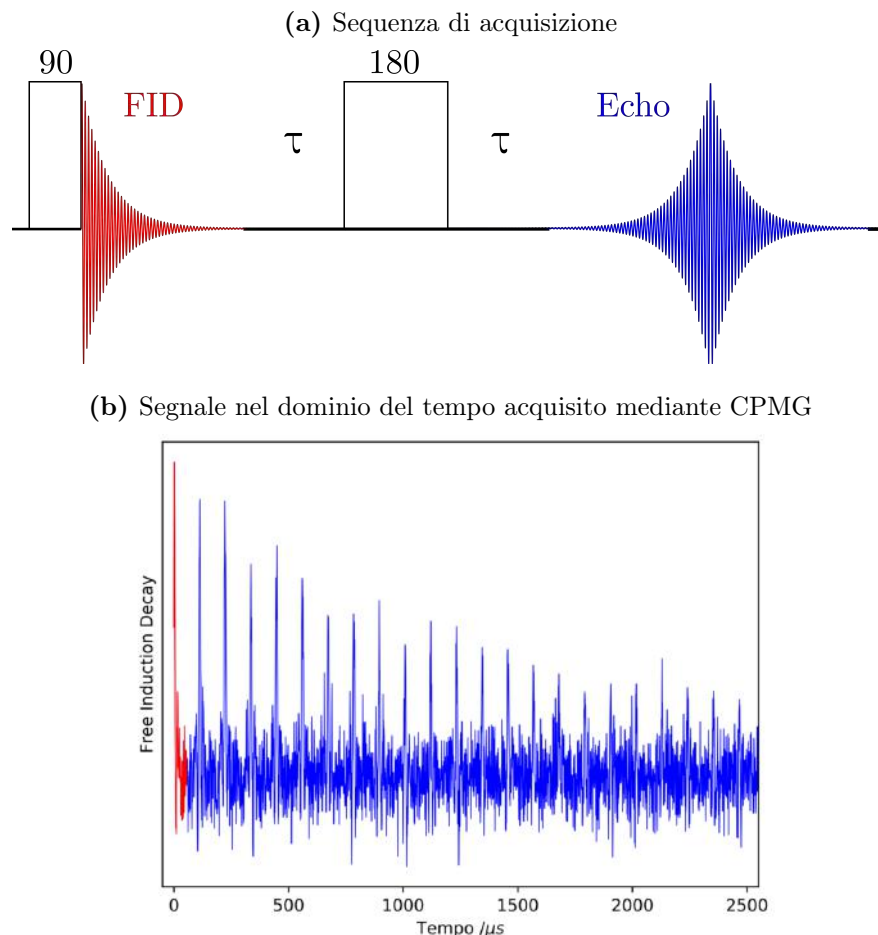
Anche ridurre la temperatura del campione è problematico: servono attrezzature dedicate e una grande quantità di gas criogenici liquefatti. Si può stimare

che, in un giorno di misure a 80 K, si usino circa 200 litri di azoto liquido, con una spesa di circa 800 €. Un giorno di misure a 5 K porta al consumo di una quantità ancora maggiore di elio liquido, con un costo stimato di 5000 € al giorno.

## 1.2 Tecniche per migliorare il segnale

### 1.2.1 CPMG

Quando l'allargamento é eterogeneo ma il  $T_2$  é lungo é possibile incrementare sensibilmente il numero di scansioni, mantenendo il tempo di acquisizione ragionevole, sfruttando l'esperimento Carr-Purcell-Meiboom-Gill (CPMG). [23] Attraverso questa sequenza di impulsi gli spettri vengono registrati in forma di treno di echi (*echo-train*) impiegando il *pulse program* riportato in figura 1.3a, che consiste in un primo impulso a  $90^\circ$ , seguito da una serie di impulsi a  $180^\circ$  con lo scopo di generare un echo della magnetizzazione per ogni impulso. La FID così acquisita appare come in figura 1.3b. Gli echi acquisiti possono essere scomposti in "singole fid" e sommati, come se si trattasse appunto di transienti differenti.



**Figura 1.3:** L'esperimento CPMG.

L'intensità dell'echo dipende dalla magnetizzazione presente sul piano  $xy$ , che a sua volta é influenzata dal rilassamento longitudinale. Infatti, questo espe-

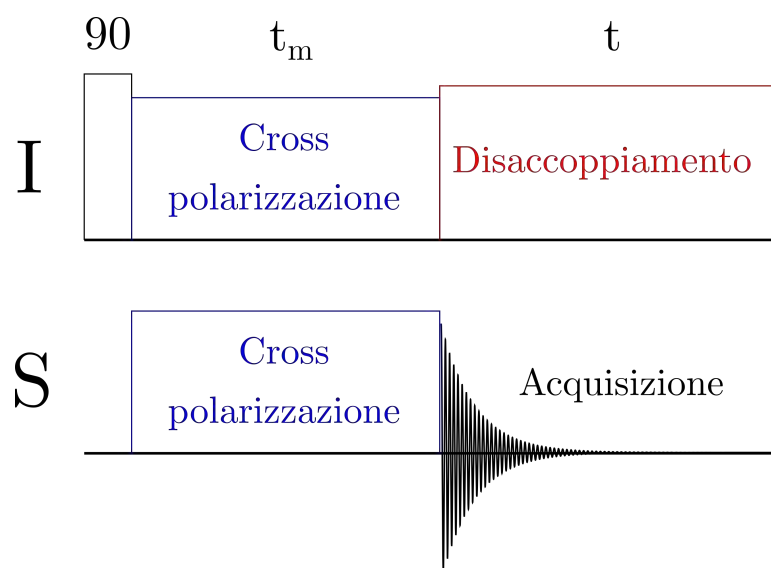
rimento fornisce una stima del tempo di rilassamento  $T_2$ , misurato fittando il decadimento degli echi in funzione del tempo.

Per ottenere uno spettro a partire da una FID acquisita mediante CPMG é necessario un processing ulteriore: sommando gli echi tra loro si riesce a incrementare sensibilmente l'intensità dei segnali abbattendo il rumore di fondo. Lo spettro che si ottiene in questo modo appare come se fosse acquisito con un numero di scansioni molto superiore, con un conseguente aumento di risoluzione, tuttavia si perdono le informazioni sul rilassamento.

### 1.2.2 Cross-polarizzazione

Una tecnica sperimentale più accessibile per incrementare la rilevabilità dei segnali é la *cross-polarizzazione*. [16][13] Si tratta di un elemento base negli esperimenti ssNMR di nuclei poco sensibili, che si basa sul trasferimento di magnetizzazione da parte di nuclei più sensibili presenti in concentrazione più elevata, al fine di incrementare il segnale acquisito. Il trasferimento di magnetizzazione avviene mediante l'accoppiamento dipolare tra i nuclei poco sensibili ( $S$ ) e i nuclei più sensibili ( $I$ ). Un ulteriore requisito per poter effettuare questo esperimento é che l'accoppiamento omonucleare  $S$ - $S$  sia trascurabile rispetto all'accoppiamento  $I$ - $I$ . Il sistema può essere descritto in termini di termodinamica di spin, secondo la quale i nuclei  $I$  ed  $S$  rappresentano due riserve termiche a temperatura diversa. Entrambe sono a contatto con l'intorno, che si trova a temperatura ambiente. Stabilendo un contatto termico tra le due riserve, si ottiene trasferimento di energia da quella più calda ( $S$ ) a quella più fredda ( $I$ ), risultando in una diminuzione della temperatura di spin di  $S$ . In termini di livelli energetici di spin, diminuire la temperatura comporta un aumento della differenza di popolazione tra lo stato fondamentale e lo stato eccitato, perciò si osserva un aumento dell'intensità del segnale associato alla transizione.

La sequenza di acquisizione dell'esperimento é riportata in figura 1.4.



**Figura 1.4:** Pulse program per un esperimento standard di cross-polarizzazione. [16].

Quando il sistema si trova all'equilibrio, lo splitting tra i livelli di spin è determinato dall'intensità del campo magnetico statico  $B_0$ : essendo i due nuclei diversi, anche la differenza energetica tra gli stati  $I$  ed  $S$  è diversa. Durante il tempo  $t_m$  dell'esperimento, viene applicato un campo magnetico  $B_1$  come spin-lock ad entrambi i nuclei, in modo tale che questi non sentano più l'influenza di  $B_0$ . Modulando l'intensità dello spin-lock è possibile far sì che la differenza energetica tra i livelli di ciascun nucleo sia uguale (condizione di Hartmann-Hahn): in queste condizioni, il trasferimento di magnetizzazione non richiede dispendio di energia.

L'aumento dell'intensità del segnale di  $S$  è proporzionale al rapporto  $\gamma_I/\gamma_S$ : per questo motivo il massimo guadagno in sensibilità si ha quando il nucleo  $I$  è  $^1\text{H}$ . Inoltre si ha un accorciamento dei tempi di riciclo della magnetizzazione, dato che il tempo di rilassamento trasversale  $T_1$  è guidato dal rilassamento di  $I$ .

Il tempo di contatto ottimale dipende dalla velocità di trasferimento di magnetizzazione, proporzionale alla forza dell'accoppiamento eteronucleare, e dal tempo di rilassamento  $T_{1\rho}$ , il quale risulta accelerato dal moto delle molecole nelle scale dei tempi della frequenza di Larmor.

Una sequenza di disaccoppiamento omonucleare deve essere applicata sullo spin  $I$  al fine di aumentare la risoluzione durante il tempo di acquisizione.

Data l'elevata potenza applicata durante quest'ultima fase, i tempi di acquisizione devono essere il più corti possibile per evitare l'usura del probe e il riscaldamento del campione. I tempi di riciclo tra due acquisizioni successive sono tipicamente dell'ordine dei 3s, ma sono comunque limitati dal rilassamento trasversale.

### 1.2.3 PMLG

Come descritto nel paragrafo precedente, per aumentare la risoluzione del segnale è necessario applicare una sequenza di disaccoppiamento omonucleare allo spin  $I$ . In questo lavoro è stata impiegata la versione più evoluta del semplice esperimento *Lee-Goldburg* (LG), ovvero il *Phase-Modulated Lee-Goldburg* (PMLG). [21].

L'esperimento LG si basa sull'invio di un impulso off-resonance, continuo, di frequenza  $\omega_1/\sqrt{2}$ , dove  $\omega_1$  è la frequenza dell'impulso. La magnetizzazione in questo modo precede in un sistema di riferimento disallineato con quello di laboratorio, facendo sì che la media dell'accoppiamento dipolare al prim'ordine si annulli. L'evoluzione successiva è il *Frequency Switch Lee-Goldburg* (FSLG), nel quale la frequenza dell'impulso viene alternata tra  $\pm\omega_1/\sqrt{2}$ . Ad ogni switch di frequenza si introduce anche uno shift di fase di  $\pi$ . Per annullare l'effetto dell'accoppiamento dipolare, ogni inversione deve avvenire dopo un tempo uguale a una precessione di  $2\pi$  della magnetizzazione di  $I$  nel sistema di riferimento ruotato. Dato che l'FSLG consiste in una sequenza molto semplice, è stato modificato per renderlo più performante, giungendo allo sviluppo del PMLG. La sequenza del PMLG consiste in una serie di  $m$  impulsi, ciascuno di fase diversa, che possono essere ripetuti  $n$  volte. L'incremento di fase è inversamente proporzionale al numero di impulsi inviati. Per mantenere il profilo di radiofrequenze inviate il più regolare possibile bisogna far sì che il numero di impulsi sia elevato, riducendone la durata, che di conseguenza deve essere regolata in base all'elettronica dello spettrometro per evitare la formazione di artefatti. Rispetto all'FSLG, questo

metodo è più versatile in quanto consente di modulare intensità e forma degli impulsi RF.

### 1.3 Metodi post-acquisizione

In accordo con quanto esposto nelle sezioni precedenti, la spettroscopia ssNMR presenta non poche complicazioni, tuttavia permette di ottenere informazioni non deducibili attraverso l'impiego di altre tecniche. Numerosi studi sono orientati al *denoising*, l'insieme dei metodi di processing che consentono l'estrazione del segnale da spettri rumorosi "a costo zero".

Un buon metodo di denoising riduce il livello di rumore senza alterare la forma o l'intensità dei segnali, intervenendo solamente sull'incremento del rapporto segnale-rumore. Durante tutte le fasi del processing il rumore deve rimanere bianco, ovvero la sua distribuzione deve mantenere un andamento gaussiano.

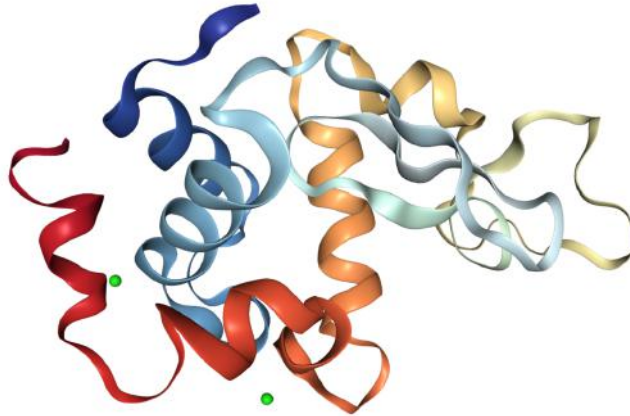
L'applicazione di una procedura di denoising efficace consente di acquisire una maggior quantità di informazioni dagli spettri di nuclei poco sensibili.

In letteratura esistono numerose opzioni per ridurre il rumore, ad esempio: wavelet transform [1], Savitzky-Golay, [22] random QR denoising, [6] Singular Spectrum Analysis, [8] e il filtro di Cadzow, [10, 12] ciascuno con i propri pro e contro.

## 1.4 Liozima e silice: un rapporto complicato

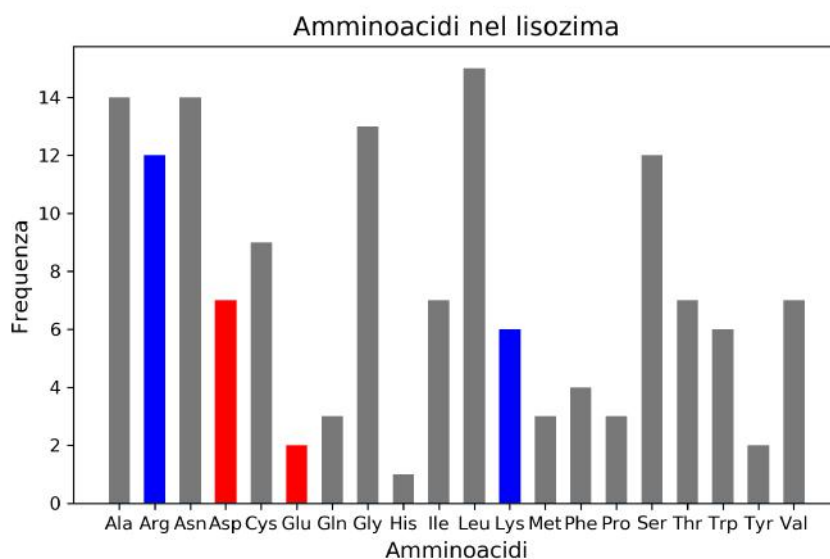
### 1.4.1 Il liozima

Il liozima (figura 1.5) è una piccola proteina globulare, del peso di 14.4 kDa, presente nella maggior parte degli organismi animali. Si trova in concentrazione rilevante in secrezioni umane, quali lacrime e saliva, e nell'albume d'uovo. Ha funzione antibatterica.



**Figura 1.5:** Schematizzazione della struttura del liozima.

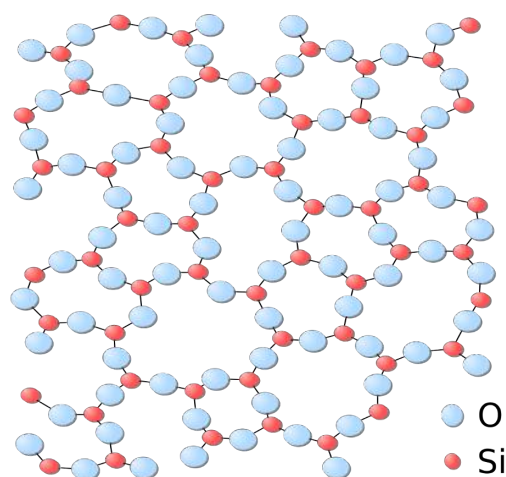
La sua struttura primaria consiste di 149 amminoacidi, organizzati in due domini e distribuiti come riportato nell'istogramma di figura 1.6. Da questo si può notare che l'80% degli amminoacidi presenta catena laterale a carica nulla, quindi la proteina è principalmente neutra. Tuttavia, ciò che la caratterizza è la prevalenza di catene laterali cariche positivamente, che conferiscono alla superficie esterna una spiccata carica positiva. Su questa peculiarità si basa parte della funzione antibatterica del liozima: attraverso interazioni elettrostatiche è in grado di favorire la fagocitosi dei batteri riducendone la carica negativa superficiale.



**Figura 1.6:** Istogramma rappresentativo della distribuzione dei vari amminoacidi nel lisozima.

### 1.4.2 L'interazione con la silice

Il diossido di silicio  $\text{SiO}_2$ , conosciuto in nomenclatura tradizionale come *silice*, è un solido polimerico. In natura si trova sotto forma di polvere cristallina bianca in cui ciascun atomo di silicio coordina tre atomi di ossigeno. È il principale costituente dei minerali silicati, tra i quali il quarzo, e di molte rocce sedimentarie.



**Figura 1.7:** Struttura schematizzata della silice cristallina.

La silice è un materiale chimicamente inerte, molto duro (7 nella scala di Mohs), e inadatto alla conduzione di calore. Queste proprietà la rendono ideale per diverse applicazioni, ad esempio può essere impiegata come abrasivo o isolante termico, ed è il materiale d'elezione per la costruzione di colonne impaccate per cromatografia analitica. Inoltre, la sua struttura la rende particolarmente adatta all'adsorbimento di piccole molecole, fungendo da supporto per la sintesi di catalizzatori eterogenei.

Nella silice il silicio é coordinato in piú modi, e si identificano siti con caratteristiche diverse: disilanoli  $\text{SiO}_2(\text{OH})_2^{2-}$  (Q<sup>2</sup>), monosilanoli  $\text{SiO}_3(\text{OH})^{3-}$  (Q<sup>3</sup>) e siti completamente condensati  $\text{SiO}_4^{4-}$  (Q<sup>4</sup>), che generano negli spettri NMR di <sup>29</sup>Si segnali caratteristici rispettivamente a -90, -100 e -110 ppm. I precursori per la formazione della silice in acqua sono specie chimiche di stechiometria  $[\text{SiO}_4\text{H}_x]^{(4-x)-}$ , alle quali ci si riferisce con il termine generico di *acido silicico*. In una soluzione neutra, l'acido silicico é presente prevalentemente in forma indissociata e in forma monodeprotonata ( $pK_a = 6.9$ ). L'acido silicico tende spontaneamente a polimerizzare, ma si pensa che la presenza di lisozima nell'ambiente avvicini le particelle mediante attrazioni elettrostatiche, facilitando la formazione di un network molto condensato, che in soluzione appare come una dispersione colloidale lattiginosa.

Nonostante la sintesi di silice "bioispirata" sia conosciuta da diverso tempo, mancano una caratterizzazione approfondita del materiale formato e, soprattutto, uno studio focalizzato sul destino della proteina dopo la formazione del network. A tale proposito, questo lavoro di tesi é finalizzato alla caratterizzazione, mediante spettroscopia NMR di stato solido, delle interazioni che sussistono tra silice e lisozima in questo tipo di composti, e alla formulazione di modelli compatibili con i dati sperimentali. Si propone inoltre un metodo di denoising innovativo che consenta di ricavare informazioni piú accurate dagli spettri registrati.

## 2. Materiali e metodi

### 2.1 Preparazione dei campioni

Come precursori per la preparazione del campione sono stati impiegati lisozima da albume d'uovo e tetrametil ortosilicato (*TetraMethyl OrthoSilicate*, *TMOS*), acquistati da Sigma-Aldrich e utilizzati senza purificazione ulteriore.

Il lisozima è stato disciolto in buffer TRIS a  $pH = 7.5$  (soluzione A; concentrazione  $22 \text{ mg/cm}^3$ ). Il TMOS è stato idrolizzato con una soluzione  $1 \text{ mmol/dm}^3$  di HCl al fine di preparare uno stock di acido silicico  $1 \text{ mol/dm}^3$ , successivamente diluito 1:10 con lo stesso buffer dopo cinque minuti di agitazione a temperatura ambiente (soluzione B; concentrazione  $0.1 \text{ mol/dm}^3$ ).

Le soluzioni A e B sono state mescolate per ottenere il precipitato, che è stato poi centrifugato e lavato con acqua deionizzata fino a scomparsa di perdite di lisozima dal composto.

Per investigare il processo di folding della proteina, il precipitato è stato lavato con una soluzione di cloruro di guanidinio (GdnHCl)  $6 \text{ mol/dm}^3$  e ditioneitrato (DTT)  $50 \text{ mmol/dm}^3$  in buffer TRIS  $20 \text{ mmol/dm}^3$  a  $pH = 7.4$  e poi lavato nuovamente con acqua deionizzata, centrifugando a ogni step a  $10\,000 \text{ rpm}$  (raggio della centrifuga  $98 \text{ mm}$ ) per due minuti. È stato necessario allungare i tempi di centrifugazione a causa dell'apparente minor densità del precipitato.

#### 2.1.1 Campioni ssNMR

I campioni sono stati inseriti in un rotore di ossido di zirconio da  $3.2 \text{ mm}$ . Gli spettri sono stati registrati alla temperatura di  $265 \text{ K}$  con frequenza del MAS di  $64 \text{ kHz}$ .

Allo scopo di rivelare il più possibile sugli aspetti molecolari dell'interazione, sono stati preparati e analizzati cinque campioni:

1. lisozima nel composto liofilizzato;
2. campione 1 reidratato con  $\text{H}_2\text{O}$ ;
3. campione 1 reidratato con  $\text{D}_2\text{O}$ ;
4. campione 1 lavato con GdnHCl come agente denaturante;
5. campione 4 lavato con  $\text{H}_2\text{O}$ .

## 2.2 Strumentazione

Lo spettro  $^1\text{H}$  monodimensionale del lisozima è stato acquisito ad uno spettrometro Bruker Avance III operante ad un campo magnetico di 22 T, corrispondente alla frequenza di Larmor per il protone di 950 MHz. Un tubo di vetro da 5 mm è stato riempito con una soluzione 10 mg/cm<sup>3</sup> di lisozima in buffer fosfato 20 mmol/dm<sup>3</sup> a  $pH = 7.2$ , in miscela H<sub>2</sub>O/D<sub>2</sub>O 90:10 oppure 98% D<sub>2</sub>O.

Gli spettri ssNMR del resto dei campioni sono stati registrati ad uno spettrometro Bruker Avance II operante ad un campo magnetico di 16.4 T, corrispondente alle frequenze di Larmor di 700 MHz, 176 MHz e 139 MHz rispettivamente per  $^1\text{H}$ ,  $^{13}\text{C}$  e  $^{29}\text{Si}$ . Lo strumento è dotato di un probe 3.2 BVT MAS in doppia risonanza, sintonizzato alternativamente sui due eteronuclei. Il campo dello spettrometro è riferito al segnale del  $^{13}\text{CH}_2$  dell'adamantano a 38.48 ppm come riferimento secondario. L'ottimizzazione dei parametri di acquisizione ha definito impulsi a 90 di 2.4  $\mu\text{s}$  per  $^1\text{H}$ , 3.5  $\mu\text{s}$  per  $^{13}\text{C}$  e 4.7  $\mu\text{s}$  per  $^{29}\text{Si}$ .

## 2.3 Processing degli spettri

Tutti gli spettri sono stati processati impiegando la classe *NMR PIPE* contenuta nella libreria *NMRglue*, scritta in *python*.

Gli spettri monodimensionali sono stati processati con i seguenti parametri:

**Tabella 2.1:** Parametri di processing per spettri 1D  $^{29}\text{Si}$ .

Nucleo	$^{29}\text{Si}$
Funzione di apodizzazione	Gaussiana
LB	-20 Hz
GB	0.2
Zero-filling	4096 pt

**Tabella 2.2:** Parametri di processing per spettri 1D  $^{13}\text{C}$ .

Nucleo	$^{13}\text{C}$
Funzione di apodizzazione	Gaussiana
LB	-20 Hz
GB	0.2
Zero-filling	4096 pt

Gli spettri bidimensionali sono stati processati con i seguenti parametri:

**Tabella 2.3:** Parametri di processing per spettri 2D  $^1\text{H}$ - $^{29}\text{Si}$ .

Dimensione	F2	F1
Nucleo	$^{29}\text{Si}$	$^1\text{H}$
TD	57	60
Funzione di apodizzazione	Sine bell	Gaussiana
	Potenza: 2	LB = -20 Hz
	Offset: $0.35\pi$	GB = 0.1
Zero-filling	4096	512

**Tabella 2.4:** Parametri di processing per spettri 2D  $^1\text{H}$ - $^{13}\text{C}$ .

Dimensione	F2	F1
Nucleo	$^{13}\text{C}$	$^1\text{H}$
TD	1700	60
Funzione di apodizzazione	Gaussiana	Gaussiana
LB	-100 Hz	-100 Hz
GB	0.1	0.1
Zero-filling	2048	256

### 3. Il problema scientifico

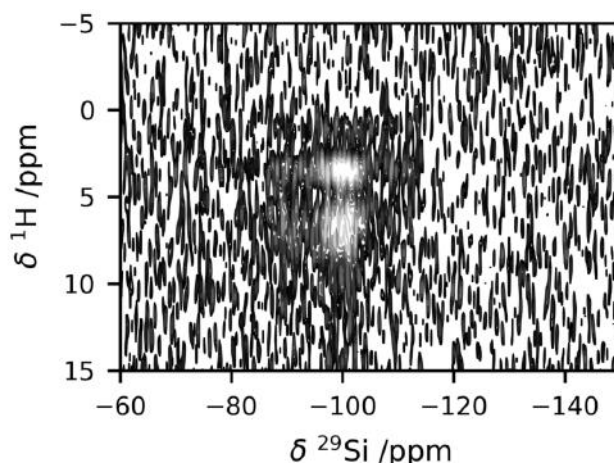
La preparazione di silice bioispirata ha suscitato un notevole interesse in quanto permette di generare una vasta gamma di materiali compositi, la cui struttura e reattività é modificabile in base alla biomolecola scelta come "stampo", mantenendo i costi di produzione incredibilmente bassi.

In questo caso la formazione del frame di silice é catalizzata dall'elevata carica superficiale del lisozima, tuttavia le interazioni che guidano l'aggregazione di silice e lisozima non sono ancora conosciute, come anche le modalit  con cui la proteina interagisce con il materiale. Caratterizzare la natura di queste forze é fondamentale per poter creare una vasta gamma di materiali ibridi con propriet  definite a priori.

Dato che queste interfacce ibride sono per loro natura non-cristalline, la spettroscopia ssNMR é lo strumento ideale per analizzare la struttura a livello atomico e dare ulteriori informazioni sulle interazioni a corto e medio raggio.

Tuttavia, come ampiamente esposto nelle sezioni precedenti, le limitazioni di questa tecnica sono tali da rendere impossibile l'interpretazione degli spettri (figura 3.1). Anche se fosse possibile acquisire alcune informazioni da spettri a bassa risoluzione, queste potrebbero essere incomplete a causa dei picchi nascosti dal rumore, conducendo nel peggiore dei casi a deduzioni errate sulle caratteristiche del campione analizzato.

Per questo motivo é stato impiegato il metodo di denoising descritto nel dettaglio nella sezione 4.



**Figura 3.1:** Esempio di spettro 2D  $^1\text{H}$ - $^{29}\text{Si}$ . É evidente che la risoluzione é cos  bassa da rendere impossibile l'estrapolazione qualunque informazione da questo tipo di dati.

## 4. Risultati e discussione

La procedura di denoising sviluppata prevede l'applicazione successiva di due metodi: *Multivariate Curve Resolution* e il filtro di *Cadzow*. [5] Sebbene entrambi possiedano limitazioni, l'utilizzo combinato dei due consente di sfruttare al meglio le potenzialità di ciascuna tecnica. Il metodo combinato si presta anche al processing di più spettri concatenati, aumentando il rapporto segnale-rumore e incrementando sensibilmente l'informazione ottenibile dagli spettri più rumorosi.

L'intero algoritmo è stato implementato in `python2.7` sfruttando la classe *NMR PIPE* contenuta nella libreria *NMRGlue*. [9]

### 4.1 Multivariate Curve Resolution

Multivariate Curve Resolution (MCR) è un metodo chemometrico principalmente impiegato per estrarre informazioni sulle componenti pure dai dati registrati su miscele complesse, principalmente spettri UV-vis e cromatogrammi. È stato impiegato successivamente per risolvere dati provenienti da un gran numero di tecniche analitiche; le sue applicazioni attuali includono: monitoraggio di reazioni chimiche, imaging spettroscopico, monitoraggio ambientale e data-processing in scienze omiche come genomica e metabolomica.

In questo lavoro di tesi si propone un utilizzo di MCR per il processing di spettri 2D ssNMR  $^{29}\text{Si}$ - $^1\text{H}$ , diverso da come era stato fatto in precedenza: l'intero algoritmo viene applicato ai dati acquisiti nel dominio del tempo, ovvero alle FID, rendendo possibile l'impiego di processing ulteriori.

#### 4.1.1 L'algoritmo

Si è scelto di presentare la trattazione mantenendo il formalismo chemometrico originale.

MCR decompone la matrice dei dati  $\mathbb{D}$  (di dimensione  $m \times n$ ) in una matrice "concentrazioni"  $\mathbb{C}$  (di dimensione  $m \times k$ ) e una matrice "spettri"  $\mathbb{S}$  (di dimensione  $k \times n$ ), lasciando una matrice residua  $\mathbb{E}$  (di dimensione  $m \times n$ ).

$$\mathbb{D} = \mathbb{C}\mathbb{S}^T + \mathbb{E} \quad (4.1)$$

Intuitivamente in uno spettro NMR 2D non si ha una variazione in concentrazione, bensì una variazione nell'intensità dei segnali dovuta all'evoluzione nella dimensione indiretta. Pertanto la matrice  $\mathbb{C}$  contiene le tracce della dimensione indiretta, mentre  $\mathbb{S}$  contiene le FID della dimensione diretta.

Per generare una guess iniziale delle due matrici  $\mathbb{C}$  ed  $\mathbb{S}$  è stato tradotta in *python* la funzione `pure.m` del pacchetto MATLAB "ALS Routines".[20, 18, 19] Lo



## 4.2 Cadzow

Il metodo di Cadzow é stato proposto con discreti risultati per il denoising di spettri NMR monodimensionali, applicando la procedura sui dati processati. Per lavorare su un dataset bidimensionale, si scompone la matrice dei dati  $\mathbb{D}$  in  $n$  righe  $\mathbf{d}_j$ , ciascuna contenente una FID della dimensione diretta. L'algoritmo viene applicato su ogni transiente indipendentemente dagli altri, quindi viene ricostruita la matrice processata concatenando le  $n$  righe.

### 4.2.1 L'algoritmo

Il transiente  $\mathbf{d}_j$  di lunghezza  $z$  viene trasformato in una matrice di Hankel  $\mathbb{H}_j$  (4.6), ovvero una matrice ad antidiagonali costanti, di dimensione  $h \times (z - h)$ , dove  $h$  é un parametro specificato dall'operatore. In una matrice di questo tipo, le informazioni sul segnale sono mantenute nella prima riga e nell'ultima colonna di  $\mathbb{H}$ .

$$(a \ b \ c \ d \ e \ f \ g \ h \ i \ j) \rightarrow \begin{pmatrix} a & b & c & d & e \\ b & c & d & e & f \\ c & d & e & f & g \\ d & e & f & g & h \\ e & f & g & h & i \\ f & g & h & i & j \end{pmatrix} \quad (4.6)$$

La matrice  $\mathbb{H}$  viene fattorizzata mediante decomposizione a valori singolari (*Singular Value Decomposition, SVD*):

$$\mathbb{H}_j = \mathbb{U}_j \mathbb{S}_j \mathbb{V}_j^* \quad (4.7)$$

dove  $\mathbb{U}$  e  $\mathbb{V}$ , rispettivamente di dimensioni  $h \times h$  e  $(z - h) \times (z - h)$ , sono dette *matrici singolari sinistra e destra*.  $\mathbb{S}$  é una matrice di dimensione  $h \times (z - h)$  i cui elementi  $s_{ij} = s_i \delta_{ij}$  sono ordinati in ordine decrescente e vengono chiamati *valori singolari* della matrice  $\mathbb{H}_j$ . Ovviamente il numero di valori singolari generati dalla SVD dipende dal parametro  $h$ .

L'idea di base del metodo é azzerare un certo numero di valori singolari, ottenendo la matrice  $\mathbb{S}'_j$ :

$$\mathbb{S}_j = \begin{pmatrix} s_{11} & 0 & 0 & 0 & 0 \\ 0 & s_{22} & 0 & 0 & 0 \\ 0 & 0 & s_{33} & 0 & 0 \\ 0 & 0 & 0 & s_{44} & 0 \\ 0 & 0 & 0 & 0 & s_{55} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \rightarrow \mathbb{S}'_j = \begin{pmatrix} s_{11} & 0 & 0 & 0 & 0 \\ 0 & s_{22} & 0 & 0 & 0 \\ 0 & 0 & s_{33} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (4.8)$$

A questo punto si ricostruisce la matrice iniziale impiegando le stesse matrici singolari ma  $\mathbb{S}'_j$  al posto di  $\mathbb{S}_j$ , ottenendo la matrice  $\mathbb{J}_j$ :

$$\mathbb{U}_j \mathbb{S}'_j \mathbb{V}_j^* = \mathbb{J}_j \quad (4.9)$$

Tuttavia  $\mathbb{J}_j$  non é una matrice Hankel-type, pertanto, per fare in modo che lo diventi, si sostituiscono le antidiagonali di  $\mathbb{J}_j$  con la media aritmetica dei valori

su ciascuna antidiagonale, ottenendo la matrice  $\mathbb{H}'_j$ :

$$\mathbb{H}'_j = \begin{pmatrix} a' & b' & c' & d' & e' \\ b' & c' & d' & e' & f' \\ c' & d' & e' & f' & g' \\ d' & e' & f' & g' & h' \\ e' & f' & g' & h' & i' \\ f' & g' & h' & i' & j' \end{pmatrix} \quad (4.10)$$

La procedura descritta viene iterata fino a convergenza di  $\mathbb{S}_j$ . Il criterio di arresto é stato stabilito come descritto nell'equazione 4.11, nella quale il simbolo ad apice indica lo step iterativo considerato.

$$R^k = \sqrt{\sum_i (s_{ii}^k - s_{ii}^{k-1})^2} \quad R^k \leq 1\% \min_i s_{ii}^0 \quad (4.11)$$

Infine, dalla matrice  $\mathbb{H}'_j$  finale si ricava il transiente processato  $\mathbf{d}'_j$  e si passa al transiente successivo.

$$\begin{pmatrix} a' & b' & c' & d' & e' \\ b' & c' & d' & e' & f' \\ c' & d' & e' & f' & g' \\ d' & e' & f' & g' & h' \\ e' & f' & g' & h' & i' \\ f' & g' & h' & i' & j' \end{pmatrix} \rightarrow (a' \ b' \ c' \ d' \ e' \ f' \ g' \ h' \ i' \ j') \quad (4.12)$$

### 4.3 Il denoising combinato

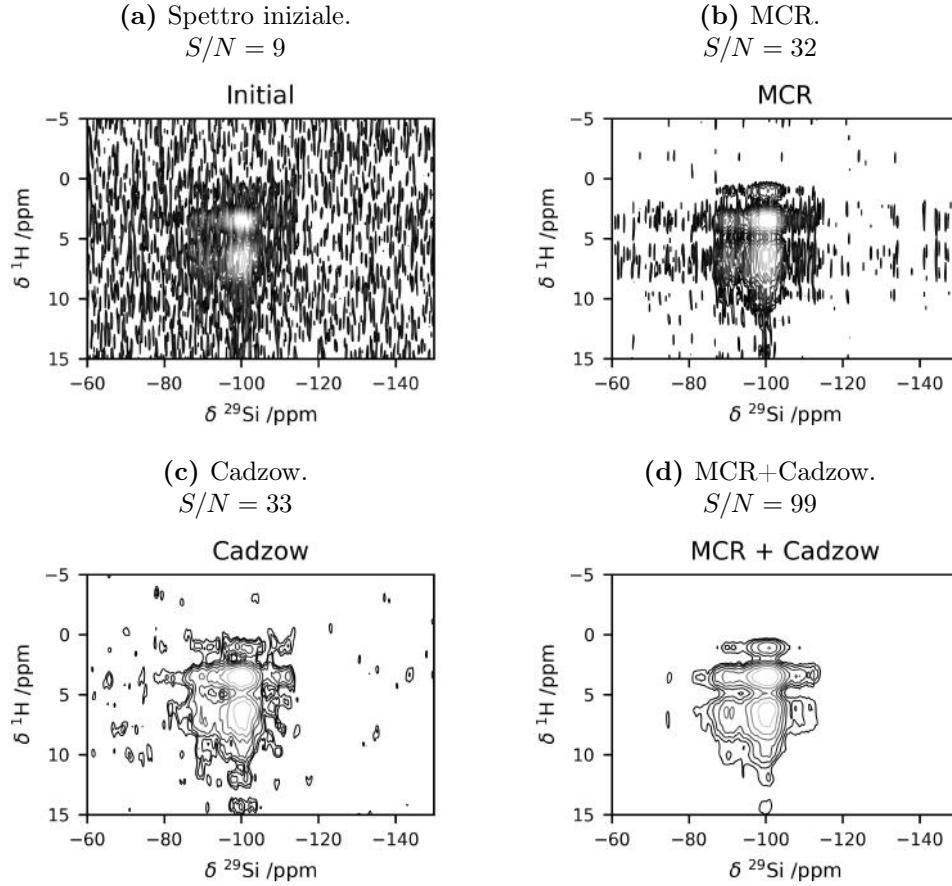
Da varie prove su alcuni dataset sperimentali é risultato evidente che il metodo di Cadzow, da solo, é un metodo piuttosto efficace, in quanto incrementa il rapporto segnale-rumore di circa tre volte. Un risultato pressoché analogo si ottiene applicando singolarmente MCR.

Risultati completamente diversi si ottengono, invece, applicando il metodo di Cadzow sull'output di MCR: in questo caso, l'incremento del rapporto segnale-rumore é circa un fattore dieci. Si suppone che la capacità di MCR di discriminare i segnali dal rumore consenta a Cadzow di eseguire un denoising più mirato.

In figura 4.2 é riportato uno spettro di esempio processato nelle diverse modalità sopra descritte.

La procedura di denoising é fortemente dipendente dai parametri di input, per MCR il numero di componenti pure da estrarre  $k$  e per Cadzow il numero di valori singolari totali  $h$  e quanti di essi porre uguali a zero. Per i dataset processati in questo lavoro sono stati scelti i seguenti parametri:

- $k = 4$ ;
- %rumore = 10%;
- $h = 35$ ;
- numero valori singolari azzerati: 15.



**Figura 4.2:** Applicazione sullo spettro **a** di solo MCR (**b**), del solo metodo di Cadzow (**c**) e del processing combinato MCR+Cadzow (**d**). Come confermato anche dai valori assunti dal rapporto  $S/N$ , il metodo combinato risulta essere molto più performante.

## 4.4 Discussione sui risultati

La procedura di denoising è stata testata prima su una serie di spettri simulati, ai quali è stato aggiunto a posteriori del rumore gaussiano. Dopodiché si è passati ad un set di spettri sperimentali, e su di essi è stato effettuato anche uno studio sulla distribuzione del rumore.

Il rapporto segnale-rumore è stato stimato dividendo l'intensità massima del segnale per la deviazione standard del rumore, calcolata come segue:

$$\text{noise} = \frac{1}{\sqrt{N-1}} \sqrt{\sum_{i=-n}^n y(i)^2 - \frac{1}{N} \left[ \left( \sum_{i=-n}^n y(i) \right)^2 + \frac{3}{N^2-1} \left( \sum_{i=1}^n i(y(i) - y(-i)) \right)^2 \right]} \quad (4.13)$$

dove  $N$  è il numero totale di punti nella regione del rumore,  $n = (N-1)/2$ , e  $y(i)$  è l' $i$ -esimo punto nella regione del rumore. Come regione del rumore rappresentativa è stata selezionata una traccia di  $^{29}\text{Si}$  nella quale non è presente segnale.

### 4.4.1 Test sui dati simulati

É stata generata una serie di tre spettri contenenti fino a nove picchi gaussiani di intensità relativa variabile, nella stessa finestra spettrale dei dataset sperimentali, come illustrato nella tabella 4.1.

F2	F1	Intensità segnale 1	Intensità segnale 2	Intensità segnale 3
-89.8 ppm, 1570 Hz	7.50 ppm, 4650 Hz	2	1	1
-89.8 ppm, 1570 Hz	4.75 ppm, 3142 Hz	1	1	1
-89.8 ppm, 1570 Hz	0.20 ppm, 3142 Hz	0	0	0.5
-99.8 ppm, 1570 Hz	7.50 ppm, 4650 Hz	3	3	3
-99.8 ppm, 1570 Hz	4.75 ppm, 3142 Hz	2	3	3
-99.8 ppm, 1570 Hz	0.20 ppm, 3142 Hz	2	3	1
-109.8 ppm, 1570 Hz	7.50 ppm, 4650 Hz	0	0	1
-109.8 ppm, 1570 Hz	4.75 ppm, 3142 Hz	0.5	2	2
-109.8 ppm, 1570 Hz	0.20 ppm, 3142 Hz	1	2	2

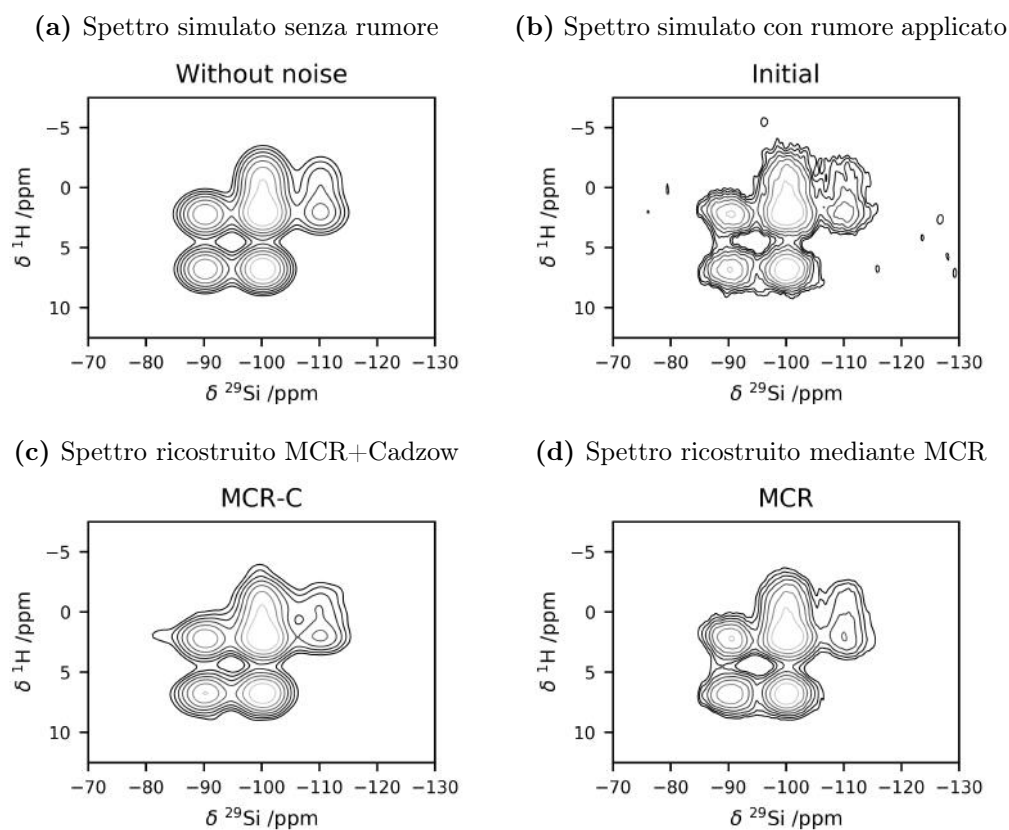
**Tabella 4.1:** Tabella illustrativa di posizione, larghezza di riga e intensità relativa dei segnali nel set di tre spettri simulati.

### MCR singolo

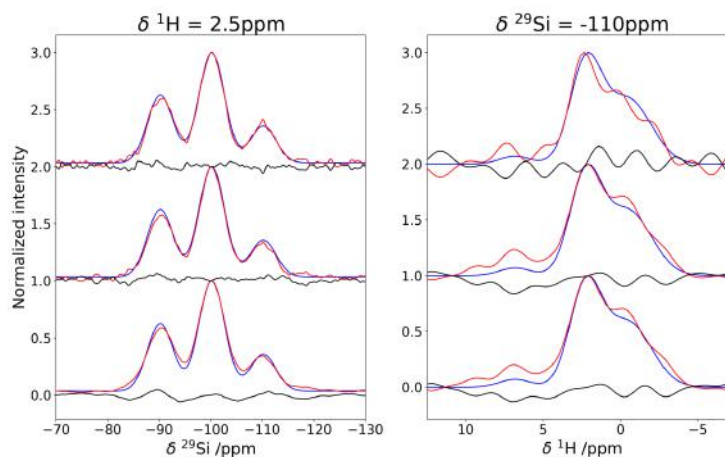
Il metodo é stato testato inizialmente su un singolo spettro, al quale é stato aggiunto rumore gaussiano, ottenendo un rapporto S/N pari a 46. Il denoising incrementa il rapporto S/N a 221 dopo MCR e 237 dopo il denoising di Cadzow. In figura 4.3 é illustrato il contour plot dello spettro nelle diverse fasi.

Per studiare l'efficienza del denoising e il suo impatto su intensità e larghezze di riga dei segnali, sono state selezionate alcune tracce significative dallo spettro con rumore aggiunto e sono state confrontate con le stesse tracce estratte dallo spettro ricostruito mediante il denoising e dello spettro senza rumore, riportato in figura 4.4. Nella stessa figura é riportata anche la differenza tra le tracce, che, se la ricostruzione fosse perfetta, dovrebbe contenere solamente rumore.

Dalla figura si osserva un notevole miglioramento in "pulizia" del segnale é notevole e senza alterazione delle intensità relative, tuttavia la ricostruzione dell'informazione non é perfetta. Si nota infatti un lieve bias della traccia "differenza" nei pannelli inferiori, dovuto alla maggiore enfasi che Cadzow dà ai primi punti del dominio del tempo, imponendo un modesto line broadening.



**Figura 4.3:** Miglioramento in risoluzione di uno spettro simulato nei vari step del denoising.



**Figura 4.4:** Efficacia della ricostruzione per un dataset processato singolarmente. La figura mostra a sinistra una traccia di  $^1\text{H}$  e a destra una traccia di  $^{29}\text{Si}$ . In ogni pannello, la traccia blu corrisponde allo spettro senza rumore, mentre la traccia rossa si riferisce nelle figure in alto allo spettro con rumore aggiunto, nelle figure al centro allo spettro ricostruito con MCR e nelle figure in basso allo spettro ricostruito con MCR+Cadzow. La corrispondente differenza é riportata in nero.

### MCR multiplo

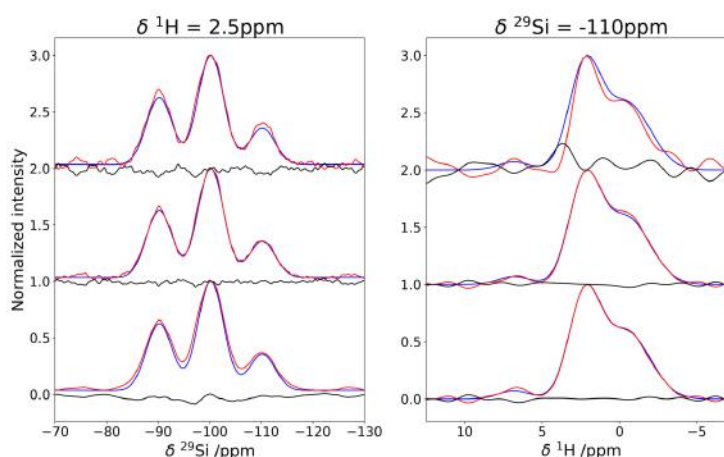
I tre spettri generati come in tabella 4.1 sono stati concatenati e sulla matrice risultante è stato applicato MCR. Il metodo di Cadzow invece è stato applicato sui tre spettri singolarmente.

È interessante osservare che, mentre il rapporto S/N degli spettri B e C migliora meno significativamente quando sono co-processati con A, il rapporto S/N di A incrementa molto di più. Allo stesso tempo, la ricostruzione di tutti e tre gli spettri mostra un'aderenza allo spettro senza rumore molto più grande rispetto a quelli processati individualmente (figura 4.5). È appunto su questo fatto che si può affermare di poter utilizzare anche spettri con rapporto S/N non elevato.

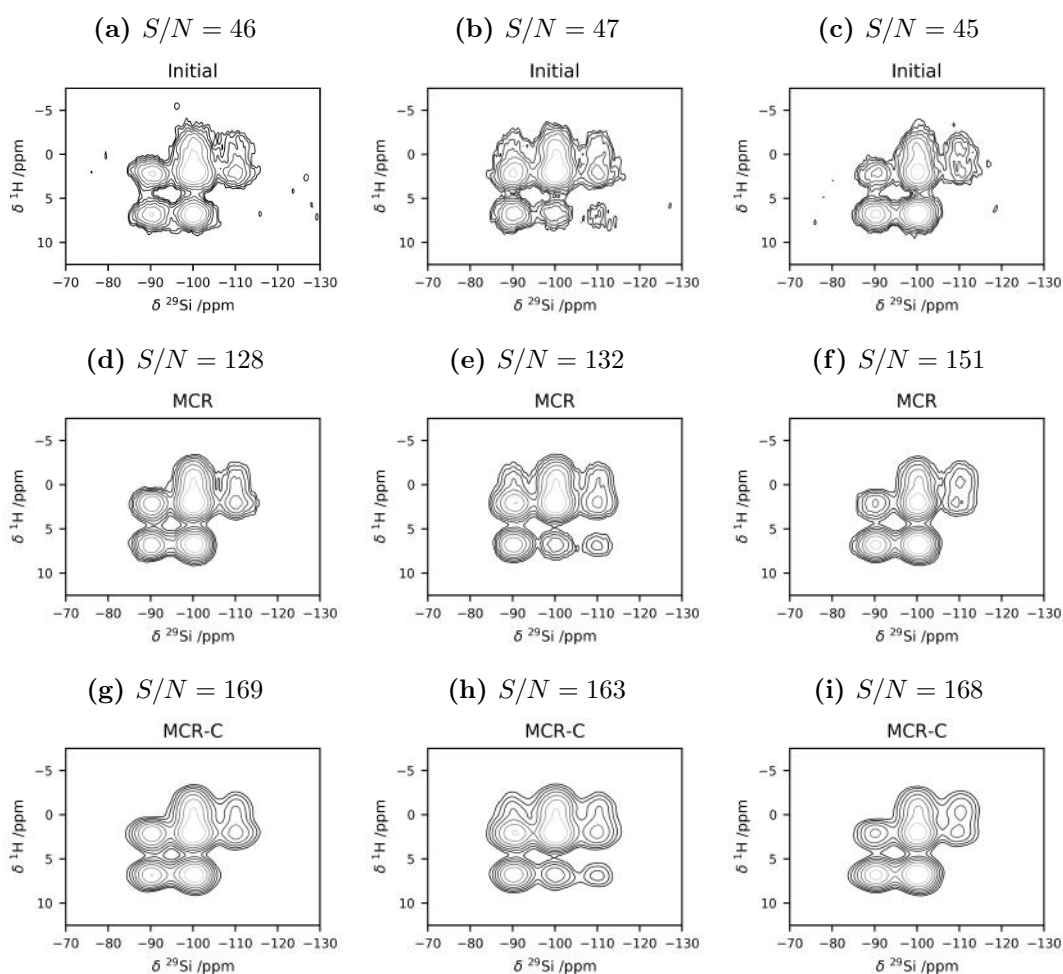
In tabella 4.2 sono riportati a confronto i rapporti S/N di tutti gli spettri processati singolarmente e coprocessati.

**Tabella 4.2:** Tabella illustrativa dei vari rapporti S/N durante gli step del processing. Lo spettro riportato come esempio nelle figure 4.4 e 4.5 è lo spettro **B**.

Step ↓/ Spet- tro →	A		B		C	
S/N Iniziale	45		45		47	
	<b>Sing.</b>	<b>Coproc.</b>	<b>Sing.</b>	<b>Coproc.</b>	<b>Sing.</b>	<b>Coproc.</b>
S/N MCR	106	206	221	150	250	153
S/N MCR+C	130	253	237	191	299	195



**Figura 4.5:** Efficacia della ricostruzione del segnale per lo stesso spettro di figura 4.4 coprocessato con gli altri due. La figura mostra a sinistra una traccia di  $^1\text{H}$  e a destra una traccia di  $^{29}\text{Si}$ . In ogni pannello, la traccia blu corrisponde allo spettro senza rumore, mentre la traccia rossa si riferisce nelle figure in alto allo spettro con rumore aggiunto, nelle figure al centro allo spettro ricostruito con MCR e nelle figure in basso allo spettro ricostruito con MCR+Cadzow. La corrispondente differenza è riportata in nero.

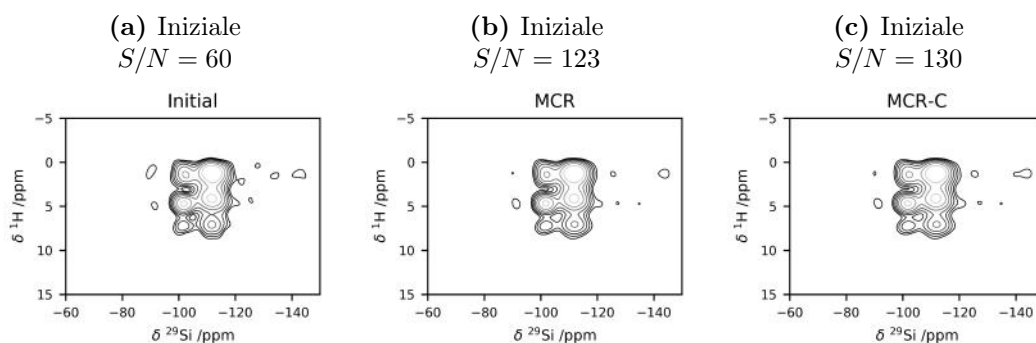


**Figura 4.6:** Risultati del denoising effettuato coprocessando i tre spettri simulati. In ciascuna colonna si può osservare l'evoluzione del processing per ogni dataset: nella prima riga sono riportati gli spettri iniziali (a, b e c), nella seconda riga l'output di MCR (d, e ed f), e nell'ultima riga il risultato finale (g, h e i)

#### 4.4.2 Test su dataset sperimentali

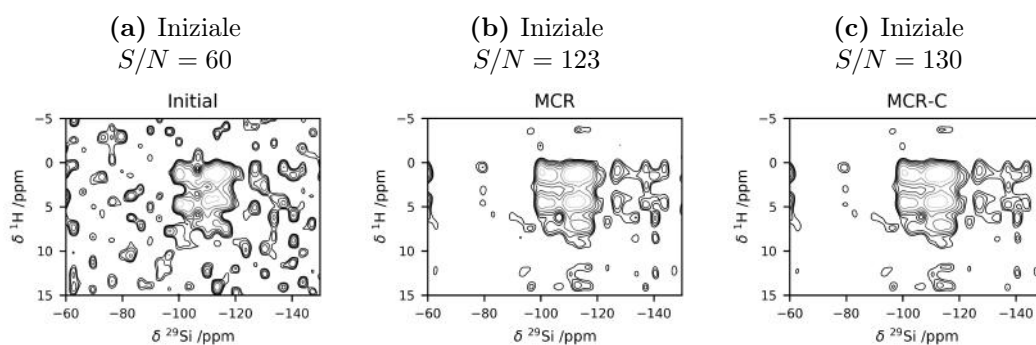
Dato che l'informazione quantitativa degli spettri sembra essere preservata durante le fasi del denoising, è stata applicata la stessa procedura a un dataset sperimentale, acquisito sul composto ibrido silice-lisozima. Lo spettro designato è un bidimensionale  $^1\text{H}$ - $^{29}\text{Si}$  acquisito impiegando la sequenza di disaccoppiamento omonucleare PMLG nella dimensione indiretta e l'esperimento CPMG lungo la dimensione diretta. Tutti gli echi di ogni transiente sono stati sommati tra loro, ottenendo un rapporto S/N iniziale di 60.

L'applicazione di MCR porta il rapporto S/N a 123 e il metodo di Cadzow in successione risulta in un ulteriore incremento fino a 130 (figura 4.7).



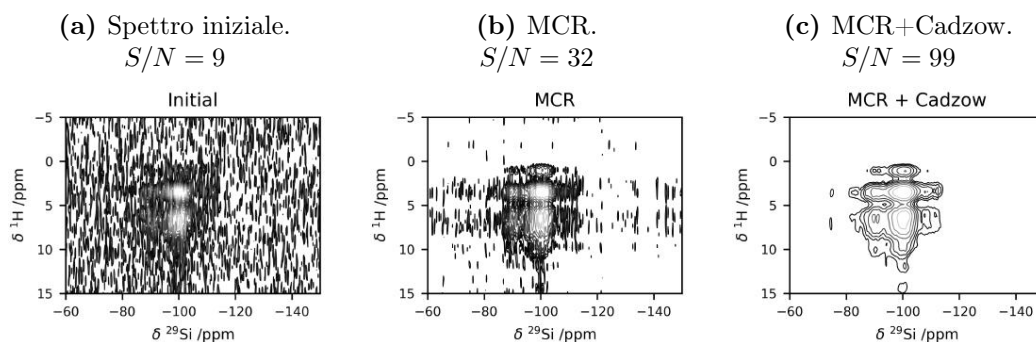
**Figura 4.7:** Risultato del denoising su un singolo spettro sperimentale.

Per effettuare un confronto, dallo stesso spettro é stato selezionato un singolo echo, per il quale il rapporto S/N é di solamente 10. L'applicazione di MCR lo porta a 27 e non incrementa ulteriormente dopo Cadzow. (figura 4.8).



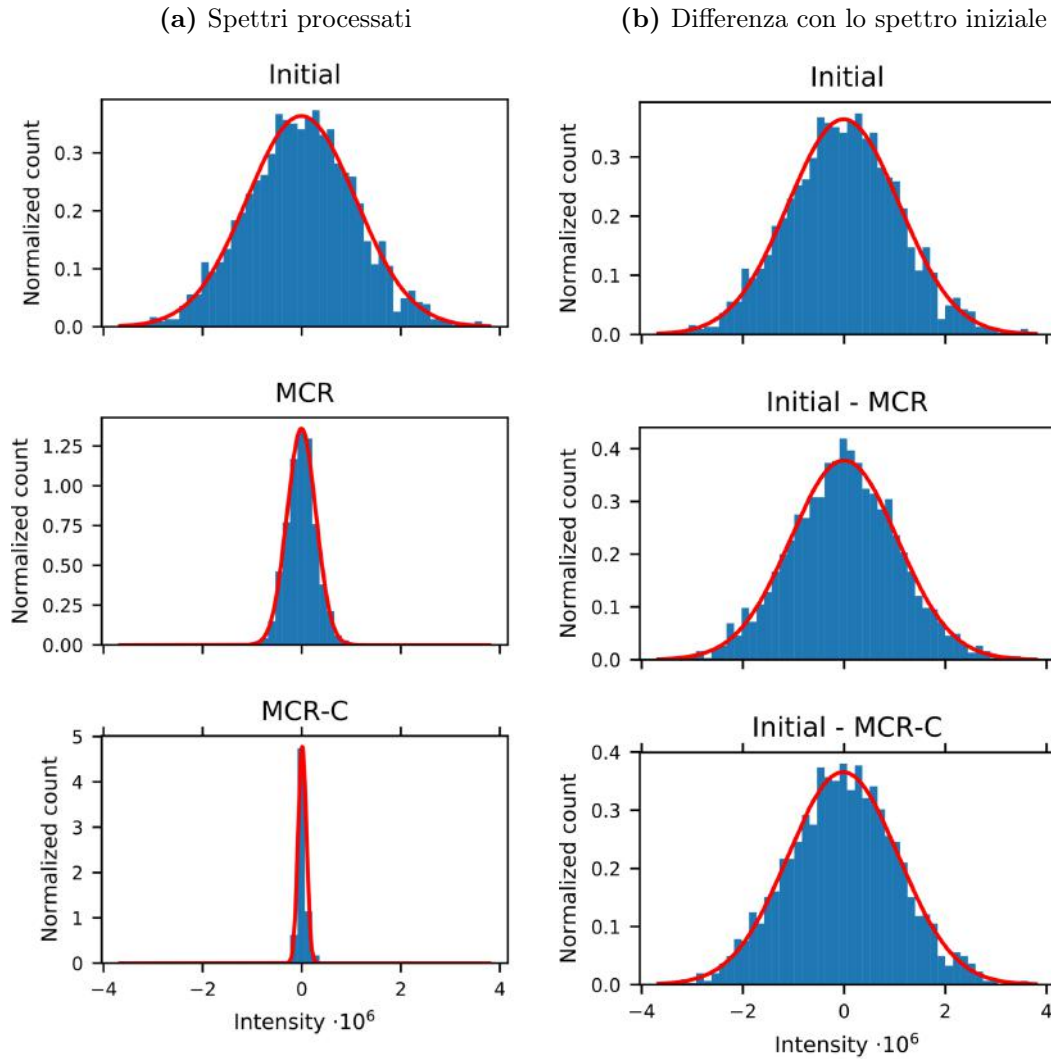
**Figura 4.8:** Risultato del denoising su un singolo spettro sperimentale con rapporto S/N inferiore.

É estremamente importante inoltre valutare le proprietà del rumore, per verificare che non cambi significativamente durante la procedura di denoising. É quindi stato selezionato uno spettro diverso, acquisito sulla dimensione diretta non come echo-train ma con un semplice esperimento di Cross-Polarizzazione, e su di esso é stata eseguita un'analisi sulla distribuzione del rumore. Gli spettri ottenuti dall'applicazione del denoising sono riportati in figura 4.9.



**Figura 4.9:** Applicazione del denoising su uno spettro con basso rapporto S/N per eseguire analisi sulla distribuzione del rumore.

Dai risultati mostrati in figura 4.10a si osserva che il rumore rimane gaussiano nell'intero processo, ma la larghezza della distribuzione decresce significativamente. La figura 4.10b mostra la differenza tra lo spettro iniziale e gli spettri ricostruiti, evidenziando che la forma della gaussiana rimane la stessa, indice del fatto che nessuna componente del segnale non è andata perduta nel rumore. I parametri sono riportati in tabella 4.3.



**Figura 4.10:** Analisi sulla distribuzione del rumore di uno spettro sperimentale preso come esempio.

**Tabella 4.3:** Tabella illustrativa dei parametri ottenuti per le gaussiane in figura 4.10.

Spettro	$\mathbb{D}$	
	$\mu = -0.01 \times 10^6, \sigma = 1.10 \times 10^6$	
Spettro	$\mathbb{C}\mathbb{S}^T$	$\mathbb{E}$
MCR	$\mu = -0.01 \times 10^6, \sigma = 0.29 \times 10^6$	$\mu = -0.002 \times 10^6, \sigma = 1.05 \times 10^6$
MCR+C	$\mu = 0.01 \times 10^6, \sigma = 0.08 \times 10^6$	$\mu = -0.02 \times 10^6, \sigma = 1.09 \times 10^6$

Abbiamo anche applicato il denoising combinato a una serie di spettri bidimensionali  $^1\text{H}$ - $^{29}\text{Si}$  acquisiti con diversi tempi di contatto. Tutti gli echi generati dal CPMG in ogni transiente sono stati sommati. I risultati sono mostrati in figura 4.11 e in tabella 4.4.

**Tabella 4.4:** Tabella illustrativa dei vari rapporti S/N durante gli step del processing.

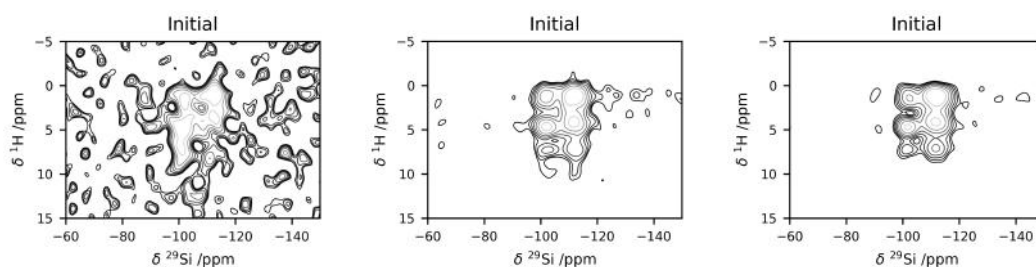
A: tempo di contatto 0.5 ms

B: tempo di contatto 6 ms

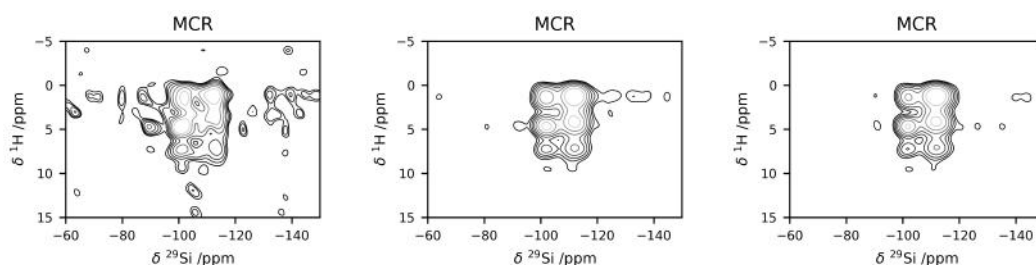
C: tempo di contatto 10 ms

Step ↓/ Spettro →	A	B	C
S/N Iniziale	10	43	60
S/N MCR	26	119	160
S/N MCR+C	28	123	177

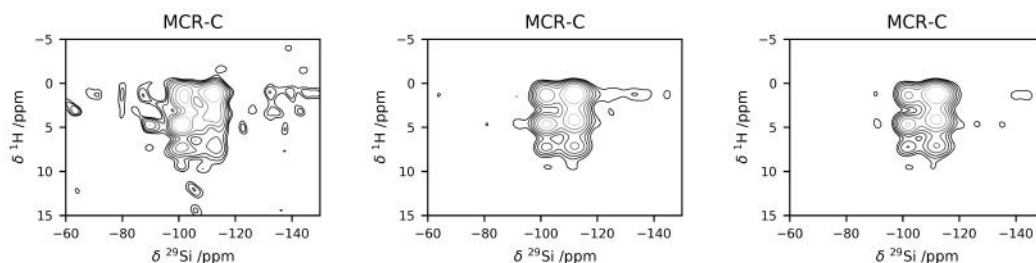
(a) Tempo di contatto: 0.5 ms (b) Tempo di contatto: 6 ms (c) Tempo di contatto: 10 ms



(d) Tempo di contatto: 0.5 ms (e) Tempo di contatto: 6 ms (f) Tempo di contatto: 10 ms



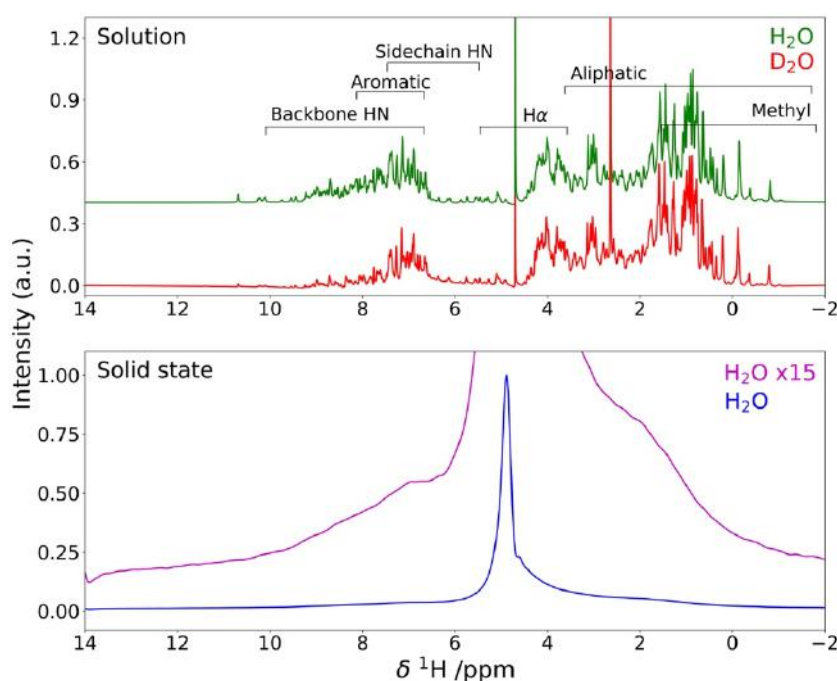
(g) Tempo di contatto: 0.5 ms (h) Tempo di contatto: 6 ms (i) Tempo di contatto: 10 ms



**Figura 4.11:** Risultati del denoising effettuato coprocessando i tre spettri sperimentali. In ciascuna colonna si può osservare l'evoluzione del processing per ogni dataset: nella prima riga sono riportati gli spettri iniziali (a, b e c), nella seconda riga l'output di MCR (d, e ed f), e nell'ultima riga il risultato finale (g, h e i)

## 5. Analisi del materiale

### 5.1 La prospettiva del lisozima



**Figura 5.1:** Nel pannello superiore é riportato lo spettro  $^1\text{H}$ -NMR di una soluzione  $10\text{ mg/cm}^3$  di lisozima in miscela  $\text{H}_2\text{O}/\text{D}_2\text{O}$  90:10 (in verde) e in  $98\%\text{D}_2\text{O}$  (in rosso). Nella figura sono evidenziate le regioni approssimative in cui cadono i segnali dei diversi protoni. Il pannello inferiore invece mostra lo spettro di stato solido dello stesso campione (in blu), con anche un ingrandimento di 15 volte (in viola) per osservare meglio le caratteristiche spettrali attorno al segnale dell'acqua.

Lo spettro  $^1\text{H}$ -NMR del lisozima é quello tipico di una proteina foldata, con segnali distribuiti tra 12 e  $-1$  ppm. Lo spettro acquisito con la soppressione del segnale dell'acqua é riportato in figura 5.1 (pannello superiore, in verde), dove gli assegnamenti collettivi sono basati sulla referenza [17]. In seguito allo scambio con  $\text{D}_2\text{O}$  si osserva un'attenuazione dei segnali delle ammidi (figura 5.1, pannello superiore, in rosso). Nella stessa regione dello spettro di stato solido del composito (figura 5.1, pannello inferiore) si osservano alcuni tratti del segnale.

In figura 5.2 sono riportati gli spettri  $^{13}\text{C}$ -NMR di stato solido del lisozima in campioni diversi.

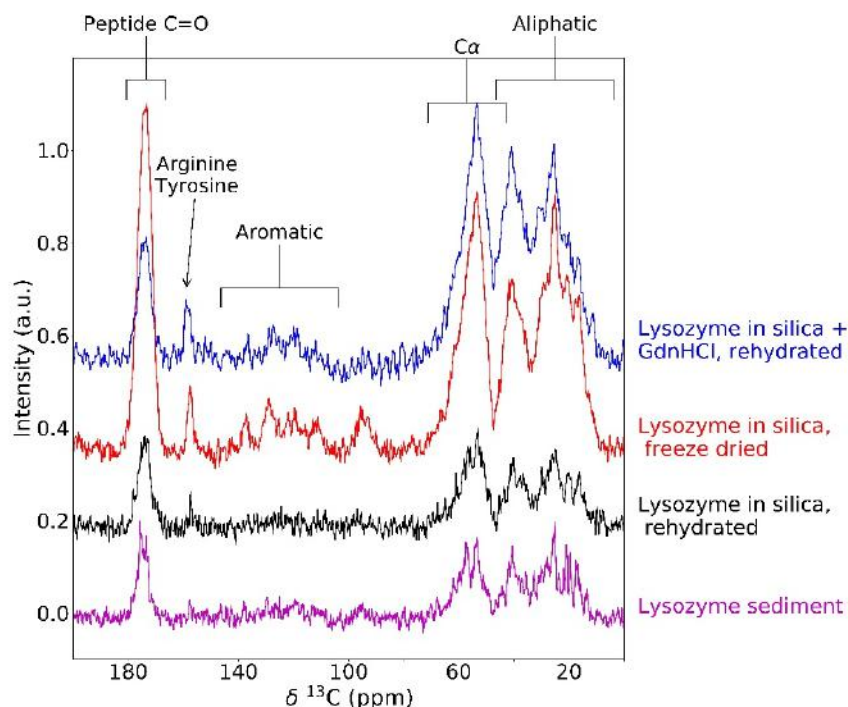


Figura 5.2: Spettri  $^{13}\text{C}$  CP-MAS NMR del lisozima in diverse condizioni.

Come si osserva tipicamente per campioni di proteine liofilizzate, la risoluzione dei segnali della proteina è bassa a causa dell'eterogeneità strutturale. Tuttavia in seguito a reidratazione si osserva un incremento in intensità di alcuni segnali dovuto alle dinamiche locali che l'acqua consente all'interno della proteina. È un'osservazione piuttosto comune in spettroscopia ssNMR di biomolecole.

Il confronto tra lo spettro del composito e del lisozima sedimentato mostra che il largo involuppo di segnali del composito include il range di frequenze osservati per la proteina sedimentata. Anche con uno spettro risultante molto largo, si possono osservare caratteristiche distintive di una proteina foldata, come la distribuzione complessiva dei segnali alifatici tra 65 e 5 ppm, oppure lo shift dei segnali dei metili a meno di 20 ppm quando il core idrofobico della proteina è intatto.

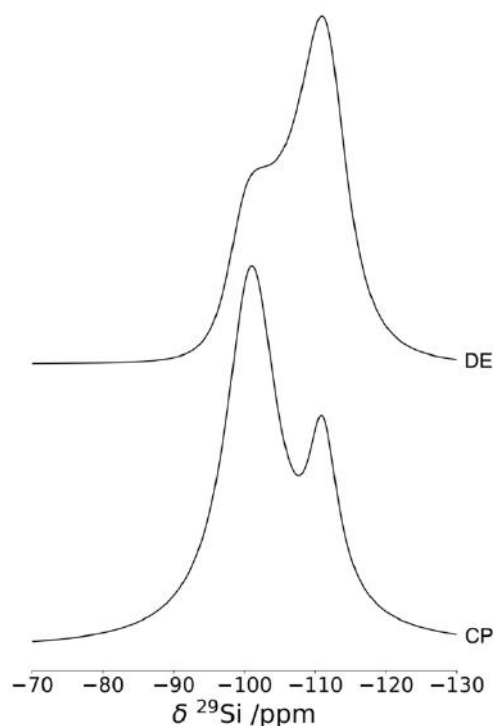
Dato che la distribuzione di chemical shift dei segnali di  $^{13}\text{C}$ -NMR è una sonda per il corretto folding delle proteine si può concludere che il lisozima all'interno del materiale mantiene la sua struttura terziaria.

## 5.2 La prospettiva della silice

La caratterizzazione della matrice di silice è cominciata dal campione liofilizzato. Di solito, gli spettri  $^{29}\text{Si}$  NMR della silice mostrano segnali provenienti dai silossani ( $Q^4$ , circa  $-110$  ppm), monosilanoli ( $Q^3$ , circa  $-100$  ppm) e silanoli geminali ( $Q^2$ , circa  $-90$  ppm).

Il rapporto tra i siti completamente condensati  $Q^4$  e i meno condensati  $Q^3$  è utilizzato come sonda del grado di condensazione della rete di silice: gli spettri di eccitazione diretta di  $^{29}\text{Si}$  (figura 5.3) forniscono un rapporto  $Q^3/Q^4 = 0.51$ ,

indicativo di un network piuttosto condensato, in accordo con altri campioni preparati in condizioni simili. [7, 14]



**Figura 5.3:** Spettri  $^{29}\text{Si}$  NMR registrati in eccitazione diretta (DE) o mediante esperimenti di cross-polarizzazione (CP,  $^1\text{H}$ - $^{29}\text{Si}$ ), che mostrano le strutture  $\text{Q}^3$  e  $\text{Q}^4$ . Gli spettri sono stati processati applicando il denoising di Cadzow.

### 5.3 Il composito

Gli spettri monodimensionali da soli non consentono di estrarre informazioni sulle interazioni tra proteina e silice. Quindi sono stati acquisiti e confrontati spettri di correlazione  $^1\text{H}$ - $^{29}\text{Si}$  e  $^1\text{H}$ - $^{13}\text{C}$ , nei quali l'interfaccia tra silice e proteina é rappresentata dalle proiezioni sulla dimensione di protone.

In figura 5.4 é stato riportato lo spettro della silice. Aumentando il tempo di contatto da 1 a 6 ms si osserva l'effetto della spin diffusione dalla superficie della silice verso i siti piú condensati.

Lo spettro di protone ottenuto proiettando lo spettro bidimensionale lungo la dimensione del silicio mostra quali protoni sono fonte di magnetizzazione: sembra che la cross-polarizzazione provenga da diverse sorgenti.

Tutti gli spettri presentano un segnale molto largo con picchi a circa 3.5 ppm e 7.8 ppm: mentre il segnale a 3.5 ppm é consistente con i segnali di protone osservati soitamente per la silice, la spiegazione piú probabile per quello a 7.8 ppm é la spin diffusione dalla proteina verso la silice. Dopo reidratazione, la sorgente di magnetizzazione principale per tutti i siti della silice é l'acqua.

- (a) Spettro bidimensionale  $^1\text{H}$ - $^{29}\text{Si}$  del campione secco, impostando il tempo di contatto a 6 ms.
- (b) Tracce ottenute sommando le porzioni corrispondenti alle frequenze di  $^{29}\text{Si}$  dei siti  $Q^2$ ,  $Q^3$  e  $Q^4$ .

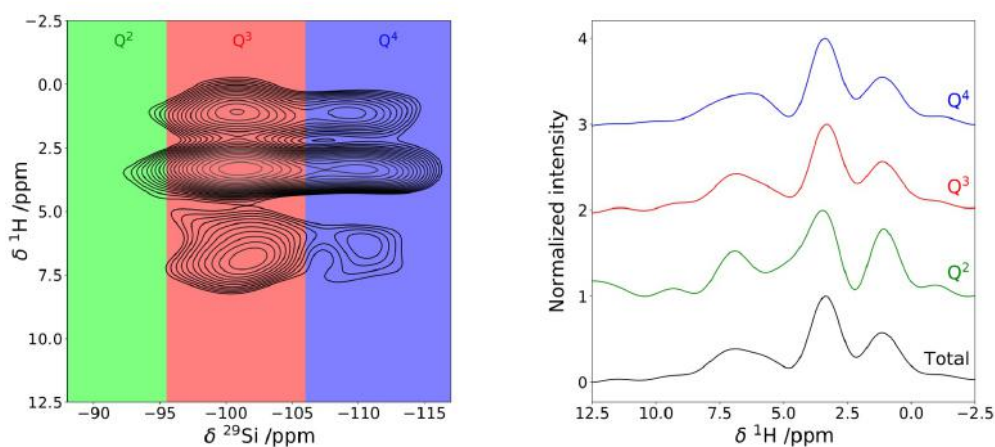
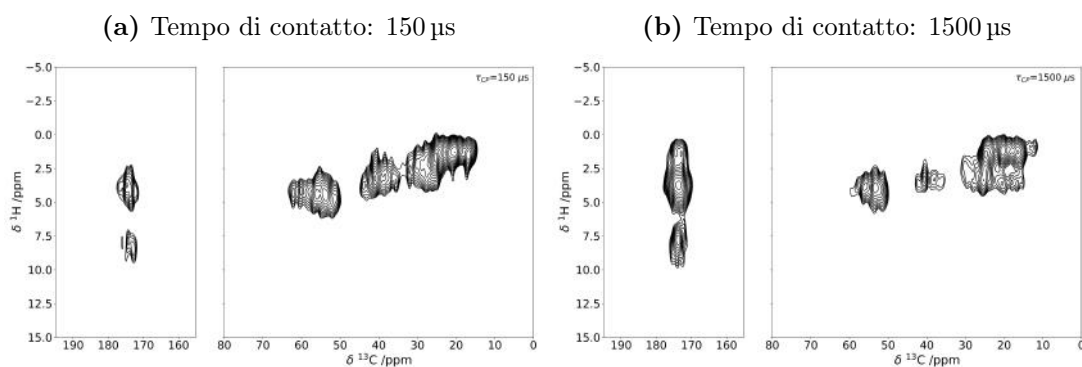


Figura 5.4

Per verificare se i nuclei di carbonio nella proteina e i nuclei di silicio nella matrice di silice vengono polarizzati dalla stessa sorgente di protone, gli esperimenti di correlazione sono stati ripetuti, sempre sul campione liofilizzato, con i nuclei di  $^{13}\text{C}$  come bersaglio. In questo tipo di spettri si osserva il massimo di trasferimento di magnetizzazione a  $150\ \mu\text{s}$  di tempo di mixing per le coppie protone-carbonio legate direttamente (come  $\text{H}_\alpha\text{-C}_\alpha$ ), e a tempi più lunghi per i carbonili polarizzati da  $\text{H}_\alpha$  e protoni ammidici (figura 5.5).

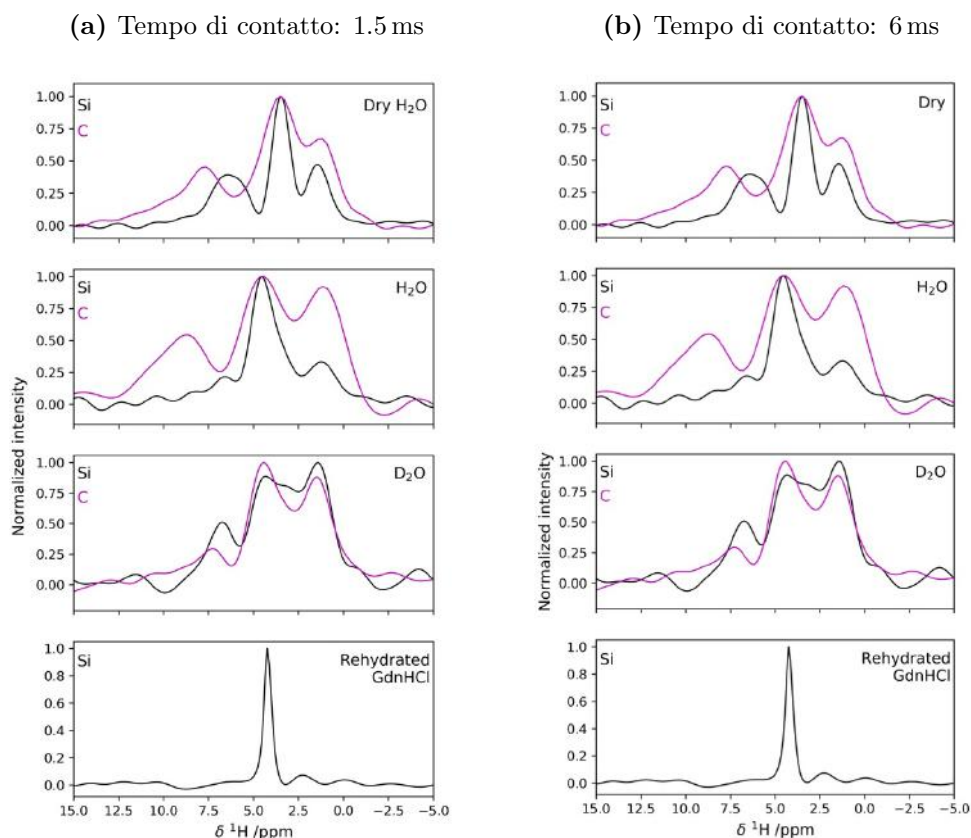


**Figura 5.5:** Spettri bidimensionali  $^1\text{H}$ - $^{13}\text{C}$  acquisiti sul campione liofilizzato a diversi tempi di contatto.

In figura 5.6 le proiezioni ottenute dagli spettri  $^1\text{H}$ - $^{29}\text{Si}$  e  $^1\text{H}$ - $^{13}\text{C}$  sono state confrontate tra di loro e con lo spettro acquisito con il disaccoppiamento omoneucleare. Si osserva che, a tempi di contatto corti, le tracce di protone degli esperimenti di correlazione sono molto diverse da come dovrebbero essere, perché il  $^{29}\text{Si}$  riceve magnetizzazione principalmente dai silanoli e dalle molecole d'acqua superficiali, mentre il  $^{13}\text{C}$  dagli atomi di carbonio della proteina.

I segnali che compaiono attorno a 7 ppm nella traccia di silicio aumentando il tempo di mixing sono di dubbia origine. Un valido assegnamento potrebbe essere

associato ai silanoli protonati, se non fosse che questi scompaiono quando il campione viene trattato con cloruro di guanidinio. Perciò la spiegazione alternativa può essere il trasferimento di magnetizzazione da parte dei residui aromatici, dalle ammidi del backbone, o dalle ammine e dal gruppo guanidinico rispettivamente delle lisine e delle arginine. Reidratando con D<sub>2</sub>O i picchi di protoni corrispondenti alle ammidi sono assenti nelle tracce estratte da entrambi gli spettri di <sup>13</sup>C e <sup>29</sup>Si.



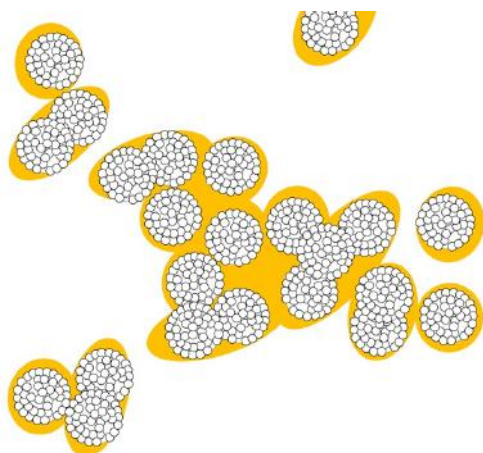
**Figura 5.6:** Tracce di <sup>1</sup>H estratte dagli spettri 2D <sup>1</sup>H-<sup>13</sup>C, in magenta, e <sup>1</sup>H-<sup>29</sup>Si, in nero. Dall'alto al basso: campione liofilizzato, reidratato con H<sub>2</sub>O, reidratato con D<sub>2</sub>O, e reidratato con H<sub>2</sub>O dopo denaturazione della proteina con GdnHCl.

Il campione denaturato e poi reidratato sembra contenere un bassissimo ammontare di proteina, ma l'analisi non è potuta essere portata avanti in modo completo a causa dei lunghissimi tempi di acquisizione degli spettri bidimensionali.

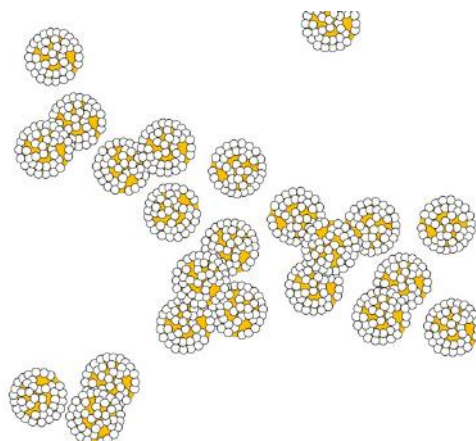
Dall'analisi degli spettri presentati si deduce che il lisozima interagisce con la silice, rimanendo in stretto contatto con la matrice condensata, ma la presenza di queste interazioni non altera abbastanza la struttura della proteina da influenzarne il folding.

Si propongono diversi modelli per spiegare il dato sperimentale, riportati in figura 5.7.

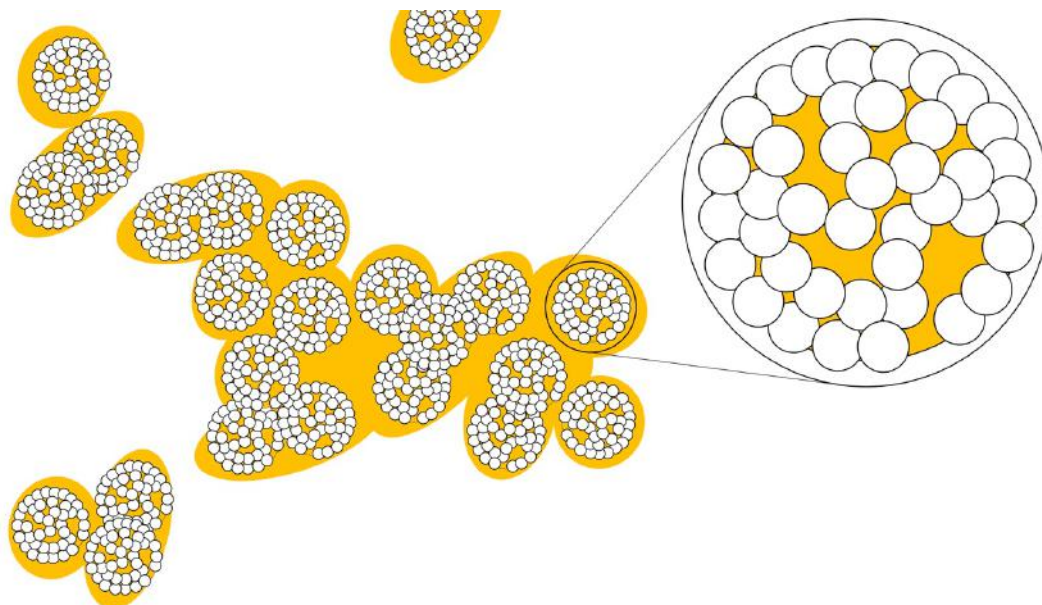
(a) Modello 1: il lisozima é adsorbito sulla superficie della silice.



(b) Modello 2: il lisozima é intrappolato all'interno della silice.



(c) Modello 3: il lisozima é sia all'interno che adsorbito sulla superficie del network di silice.



**Figura 5.7:** Rappresentazione grafica dei modelli proposti per descrivere il sistema. Lisozima: regioni arancioni; silice: sfere bianche.

Le osservazioni derivanti dal test con il cloruro di guanidinio sembrerebbero suggerire un'interazione prevalentemente elettrostatica, che permette il rilascio della proteina all'aumentare della forza ionica. Per verificare questa ipotesi, il campione di composito é stato trattato aggiungendo quattro diversi reagenti:

1. solo acqua;
2. GdnHCl 6 mol/dm<sup>3</sup> e DTT 50 mmol/dm<sup>3</sup>;
3. urea 8 mol/dm<sup>3</sup> e DTT 50 mmol/dm<sup>3</sup>;
4. NaCl 6 mol/dm<sup>3</sup>.

Assumendo che l'interazione sia rappresentata dal modello di figura 5.7c e che sia prevalentemente elettrostatica, sia la condizione 2 che la condizione 4

causerebbero il distacco della proteina dalla silice a causa dell'aumento di forza ionica del sistema.

Invece, si è osservato che il rilascio quantitativo della proteina dal materiale avviene nelle condizioni 2 e 3, suggerendo un'interazione prevalentemente sterica. Il modello che rappresenta meglio questa situazione è quello di figura 5.7b.

## 6. Conclusioni

In questo lavoro abbiamo introdotto l'uso di Multivariate Curve Resolution per migliorare la sensibilità di esperimenti bidimensionali NMR di stato solido, registrati allo scopo di caratterizzare una matrice silicica.

Questo metodo si è rivelato essere estremamente efficiente nell'aumentare il rapporto segnale-rumore, lasciando inalterate sia la forma dei picchi che le loro intensità relative. Inoltre, il suo utilizzo accoppiato con il filtro di Cadzow è in grado di incrementare la potenza di quest'ultimo, riducendo ulteriormente il livello di rumore. Offre inoltre la possibilità di coprocessare spettri dello stesso campione acquisiti in condizioni diverse, consentendo l'utilizzo quantitativo anche degli spettri con segnali meno intensi.

Date le numerose problematiche che intrinsecamente possiede la spettroscopia ssNMR di nuclei poco sensibili, uno strumento come il metodo di denoising proposto in questo lavoro è in grado di ridurre significativamente il tempo di acquisizione dei dati, migliorandone contemporaneamente la qualità.

Una possibile futura applicazione di MCR potrebbe essere l'analisi di spettri NMR di soluzione, caratterizzati da un elevato numero di picchi molto stretti, diversamente dallo stato solido in cui i segnali sono pochi e larghi. A causa della grande differenza tra i due tipi di dataset, ad ora non siamo in grado di garantire la stessa efficacia della procedura anche sugli spettri di soluzione.

Per quanto riguarda l'analisi del composito silice-lisozima, l'applicazione della spettroscopia NMR ha fornito informazioni molto dettagliate sulla natura chimica del materiale, evidenziando la conservazione del fold della proteina all'interno della struttura. Questo fatto fornisce una giustificazione sperimentale al risultato pubblicato nella pubblicazione originale di Luckarift et al. [11]

Tuttavia, gli esperimenti lasciano ancora una parziale ambiguità riguardo alla sistemazione relativa della proteina e della silice nel composito. Pur favorendo il modello della proteina intrappolata dentro la matrice, non possiamo concludere che il ruolo delle interazioni elettrostatiche sia irrilevante. Ulteriori esperimenti mirati alla proteina in soluzione saranno svolti per tentare di chiarire quest'ultimo aspetto.

# Bibliografia

- [1] J. P. Antoine, C. Chauvin e A. Coron. «Coron, Wavelets and related time-frequency techniques in magnetic resonance spectroscopy». In: *NMR in Biomedicine* 14 (2006), pp. 265–270.
- [2] Lucia Banci et al. «Biomolecular NMR at 1.2 GHz». In: *arXiv.org* (2019).
- [3] Adi Ben-Israel e Thomas N.E. Greville. *Generalized inverses: Theory and applications (2nd ed.)* Springer, 2003.
- [4] Ivano Bertini et al. *NMR of Paramagnetic Molecules*. Elsevier, 2016.
- [5] Caroline Brissac, Thérèse E. Malliavin e Marc A. Delsuc. «Use of the Cadzow procedure in 2D NMR for the reduction of noise». In: *Journal of Biomolecular NMR* 6 (1995), pp. 361–365.
- [6] L. Chiron et al. «Efficient denoising algorithms for large experimental datasets and their applications in Fourier transform ion cyclotron resonance mass spectrometry». In: *Proceedings of the National Academy of Sciences* 111 (2014), pp. 1385–1390.
- [7] S.C. Christiansen et al. «Sensitivity considerations in polarization transfer and filtering using dipole-dipole couplings: Implications for biomineral systems». In: *Solid States Nuclear Magnetic Resonance* 29 (1-3) (2006), pp. 170–182.
- [8] R. Ghanati et al. «Filtering and parameter estimation of surface-NMR data using singular spectrum analysis». In: *Journal of Applied Geophysics* 130 (2016), pp. 118–130.
- [9] J. J. Helmus e C. P. Jaroniec. «Nmrplug: an open source Python package for the analysis of multidimensional NMR data». In: *Journal of Biomolecular NMR* 55(4) (2013), pp. 355–367.
- [10] G. Laurent et al. «Denoising applied to spectroscopies – part I: concept and limits». In: *Applied Spectroscopy Reviews* 54 (2019), pp. 602–630.
- [11] H.R. Luckarift et al. «Rapid, Room-Temperature Synthesis of Antibacterial Bionanocomposites of Lysozyme with Amorphous Silica or Titania». In: *Small* 2 (2006), pp. 640–643.
- [12] P. P. Man, C. Bonhomme e F. Babonneau. «Denoising NMR time-domain signal by singular-value decomposition accelerated by graphics processing units». In: *Solid States Nuclear Magnetic Resonance* 61-62 (2014), pp. 28–34.

- [13] David Marks e Shimon Vega. «A Theory for Cross-Polarization NMR of Nonspinning and Spinning Samples». In: *Journal of Magnetic Resonance* 118 (1996), pp. 157–172.
- [14] T. Martelli et al. «Atomic level quality assessment of biosilica encapsulated and autoencapsulated enzymes». In: *Chemistry - A European Journal* 22 (2016), pp. 425–432.
- [15] Seth A. McNeill et al. «Optimizing ssNMR experiments for dilute proteins in heterogeneous mixtures at high magnetic fields». In: *Magnetic Resonance in Chemistry* 45 (2007), S209–S220.
- [16] A Pines, M. G. Gibby e J. S. Waugh. «Protonenhanced NMR of dilute spins in solids». In: *Journal of Chemical Physics* 59 (2) (1973), pp. 569–590.
- [17] C. Redfield e C. M. Dobson. «Sequential  $^1\text{H}$  NMR assignments and secondary structure of hen egg white lysozyme in solution». In: *Biochemistry* 27(1) (1988), pp. 122–136.
- [18] R. Tauler, A Smilde e B. Kowalski. «Selectivity, local rank, three-way data analysis and ambiguity in multivariate curve resolution». In: *Journal of Chemometrics* 9 (1995), pp. 31–58.
- [19] Roma Tauler. «Calculation of maximum and minimum band boundaries of feasible solutions for species profiles obtained by multivariate curve resolution». In: *Journal of Chemometrics* 15 (2001), pp. 627–646.
- [20] Roma Tauler. «Multivariate curve resolution applied to second order data». In: *Chemometrics and Intelligent Laboratory Systems* 30 (1995), pp. 133–146.
- [21] Elena Vinogradov, P.K. Madhu e Shimon Vega. «High-resolution proton solid-state NMR spectroscopy by phase-modulated Lee–Goldburg experiment». In: *Chemical Physics Letters* 314 (1999), pp. 443–450.
- [22] G. Vivò-Truyols e P.J. Schoenmakers. «Automatic Selection of Optimal Savitzky-Golay Smoothing». In: *Analytical Chemistry* 78 (2006), pp. 4598–4608.
- [23] J. W. Wiench, V. S.-Y. Lin e M. Pruski. « $^{29}\text{Si}$  NMR in Solid State with CPMG acquisition under MAS». In: *Journal of Magnetic Resonance* 193 (2) (2008), pp. 233–242.