

8. Sovrapposizione e confronto di strutture proteiche

In questa esercitazione utilizzeremo il programma SwissPdbViewer (versione 4.0.1) per effettuare la sovrapposizione e il confronto di strutture proteiche. Il software è gratuito e può essere scaricato dal sito <http://spdbv.vital-it.ch>. Sul sito sono disponibili le istruzioni per l'installazione, oltre a una breve e semplice guida introduttiva all'utilizzo. Di seguito vedremo solo alcune delle molteplici funzioni che il programma mette a disposizione. Si raccomanda pertanto di leggere la guida utente per avere una panoramica più completa delle sue potenzialità.

- Collegatevi al sito della Protein Data Bank (www.rcsb.org/pdb) per recuperare le coordinate delle proteine 1SBT e 2PRK e memorizzarle sul computer.
- Avviate il programma SwissPdbViewer e caricate i file contenenti le coordinate delle due proteine come mostrato in **Figura 1**.

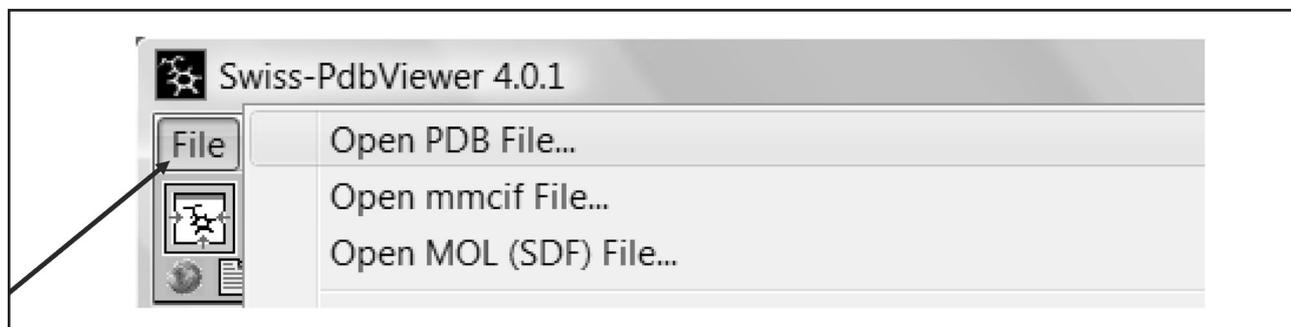


Figura 1 Particolare della barra dei menu del programma SwissPdbViewer da cui è possibile leggere le coordinate di una struttura contenute in un file.

- Visualizzate le tracce dei $C\alpha$ delle due proteine (cioè i $C\alpha$ uniti da un segmento che indica il legame peptidico) e colorate le due strutture con colori diversi in modo da che siano facilmente distinguibili, attraverso i comandi disponibili nel pannello di controllo (**Figura 2**).
- Selezionate la voce *MagicFit* nel menu *Fit*. Vedrete comparire una finestra attraverso cui si possono selezionare le strutture da sovrapporre (in questo caso ce ne sono solo due) e gli atomi di cui tenere conto nel calcolo (in genere si utilizzano i $C\alpha$). Le due tracce vengono sovrapposte. La procedura utilizzata dal programma è la seguente: si allineano le sequenze e si identificano le zone di maggiore somiglianza (evidenziate in rosso nel pannello di controllo). A partire da queste zone il programma esegue la sovrapposizione. Cliccando sulla voce *Calculate RMS* nel menu *Fit* si può visualizzare la RMSD risultante e il numero di atomi equivalenti (**Figura 3**).

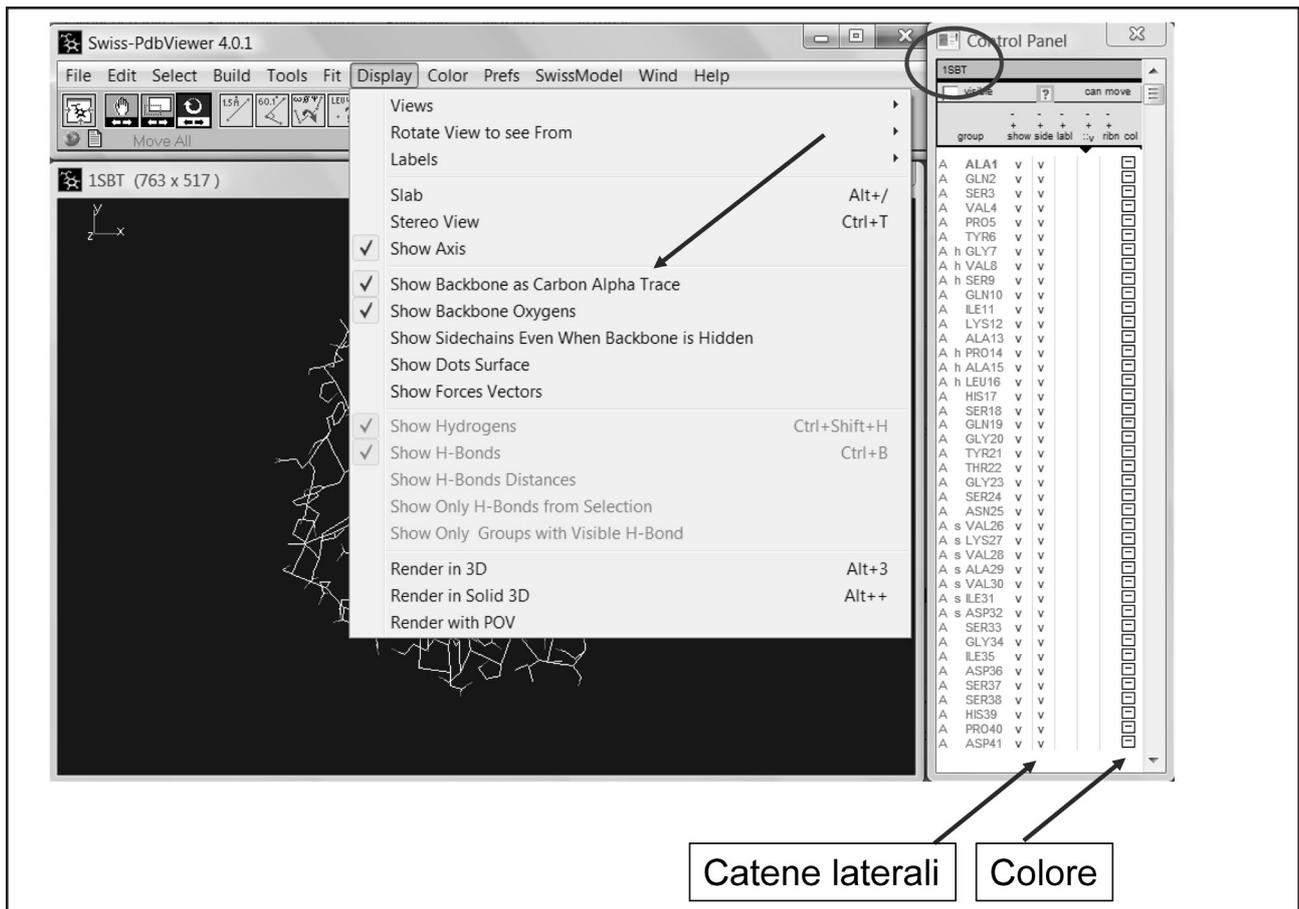


Figura 2 Pannello di controllo (a destra) che può essere visualizzato attraverso il menu *Wind* nella barra superiore. Per visualizzare solo la traccia dei C α , bisogna selezionare i residui nel pannello di controllo; quelli selezionati vengono evidenziati in rosso. Poi si rimuove la visualizzazione delle catene laterali (cliccando con il puntatore sulla colonna *side* nel pannello di controllo). Infine dal menu *Display* si seleziona la voce *Show Backbone as Carbon Alpha Trace*. L'attribuzione del colore avviene in modo analogo. Dopo aver selezionato i residui si clicca su uno dei riquadri indicati dalla freccia "colore"; si apre un pannello attraverso cui si seleziona il colore desiderato. Se nel programma sono presenti più strutture, tutte le operazioni coinvolgono la proteina selezionata nel pannello di controllo e visualizzata nell'ovale della figura.

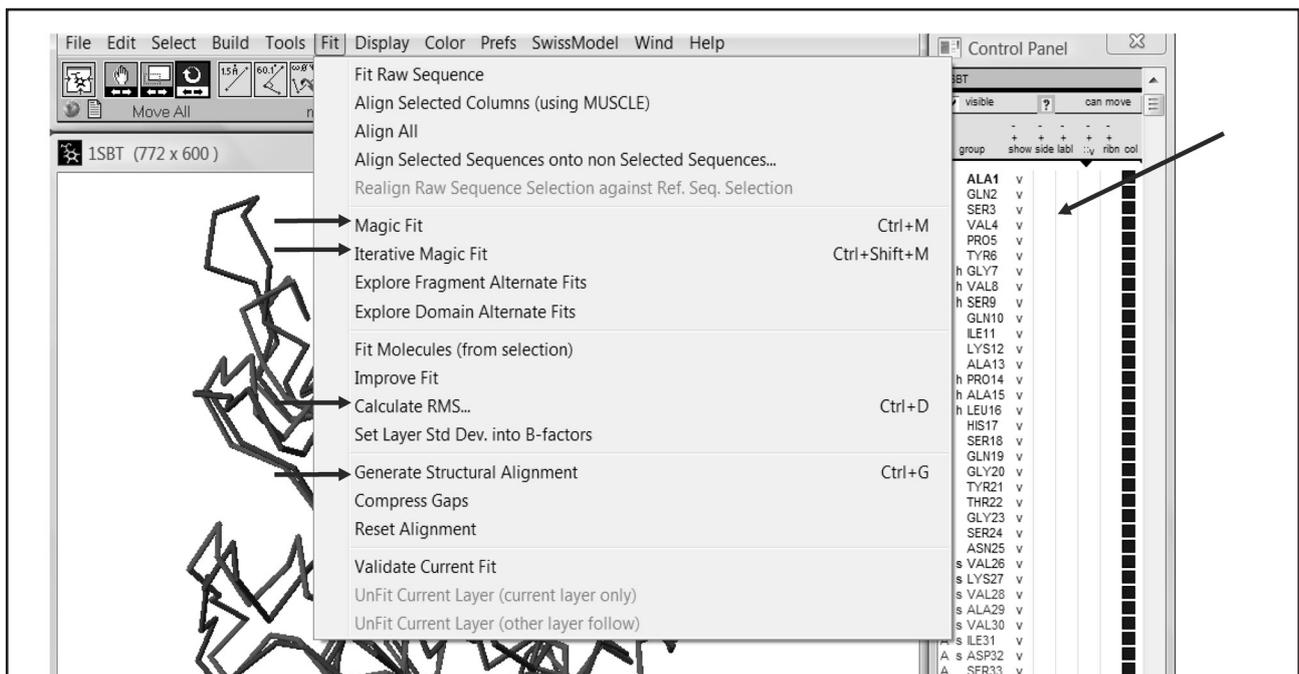


Figura 3 Comandi utilizzati nel calcolo della sovrapposizione di strutture. I residui strutturalmente equivalenti vengono visualizzati in rosso nel pannello di controllo.

Remove Selected Residues nel menu *Build*, rimuovete i residui selezionati. Analogamente, isolate la catena L della struttura 1CPC.

- Sovrapponetevi le due strutture con *Magic Fit*. Calcolate la RMSD e annotate i valori; visualizzate l'allineamento strutturale.
- Migliorate l'allineamento con *Iterative Magic Fit*: il programma interrompe la procedura perché questo calcolo non può essere effettuato con proteine molto diverse, come in questo caso.

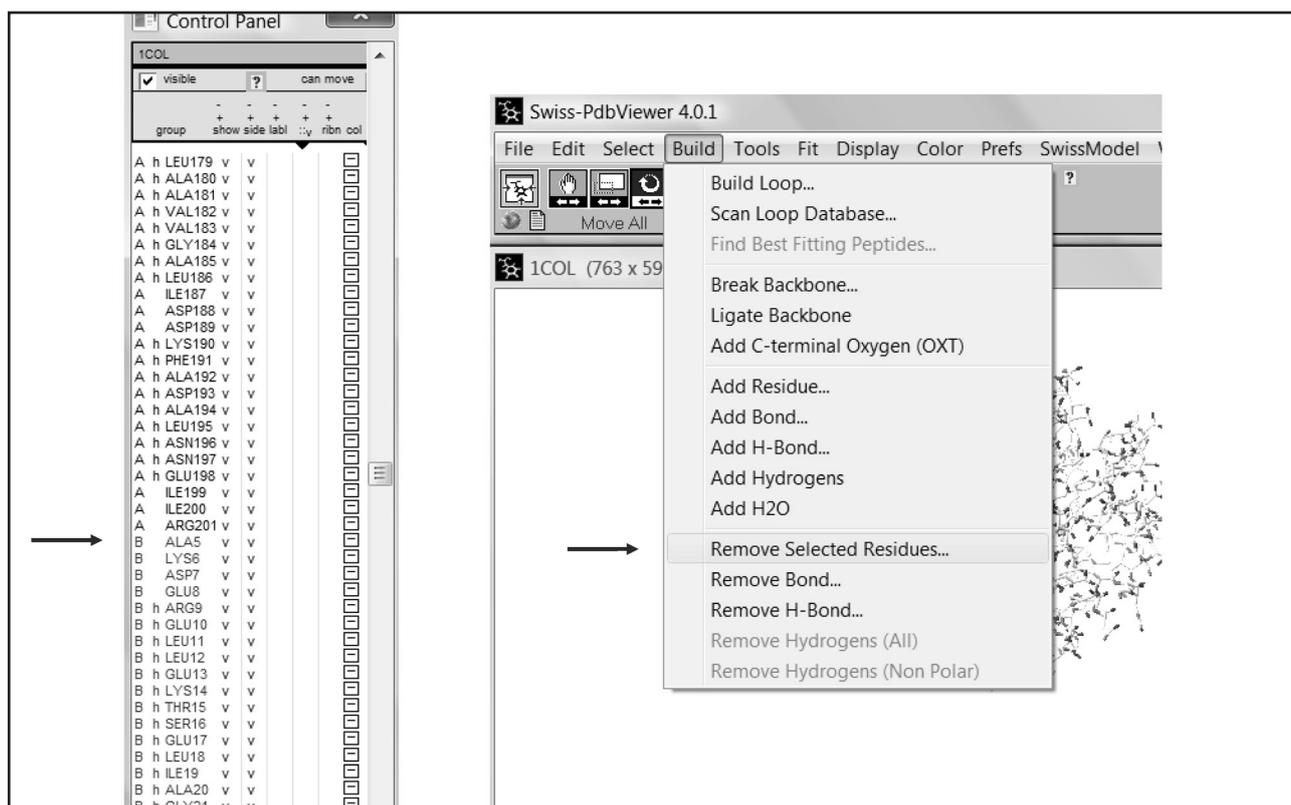


Figura 5 Procedura per rimuovere una catena di una struttura. Si deve cliccare sul carattere della catena da rimuovere (nel caso specifico B). I residui vengono evidenziati in rosso. Nel menu *Build* si seleziona la voce *Remove Selected Residues*.

Se le proteine sono molto diverse, l'uso di algoritmi di sovrapposizione per frammenti può risolvere il problema. Mostriamo ora come si utilizzano i programmi SSAP e Dali a questo scopo.

- Collegatevi al sito www.cathdb.info/cgi-bin/SsapServer.pl e inserite nei campi visualizzati il nome delle strutture da confrontare seguite dall'indicatore di catena: 1colA e 1cpcL (**Figura 6**). Avviate il confronto cliccando sul tasto *Continue*. In pochi secondi il programma calcolerà la sovrapposizione e visualizzerà i risultati. Notate i seguenti parametri: RMSD, numero di carboni equivalenti e percentuale d'identità tra le sequenze allineate. Confrontate il risultato con quello ottenuto con SwissPdbViewer.
- Collegatevi al sito www.ebi.ac.uk/Tools/dalilite/ e inserite negli appositi campi i codici delle strutture e gli identificatori di sequenza (**Figura 8**). Lanciate il programma e analizzate i risultati. Confrontateli con quelli precedenti e notate come algoritmi diversi forniscano risposte parzialmente diverse.

SSAP Server

Pairwise structure comparison

The SSAP server allows users to compare the structures of two proteins full path to a local PDB file (right box) for each structure.

Structure 1

PDB/CATH domain code (e.g. '2bop', '2bopA' or '2bopA00')

→

Structure 2

PDB/CATH domain code (e.g. '1by9', '1by9A' or '1by9A00')

→

Email

> CATHDB_V3_2_0 (change)

Home

SSAP Server - submitted job

Domain1	Length	Domain2	Length	Equiv. Res.	Overlap (%)	Seq. id (%)	Score (0-100)	RMSD
1colA0	197	1cpcl0	172	143	72	9	72.73	9.81

Alignment (readable)

```

1colA0:pdhoo 5      15      25      35
1colA0      |      |      |      |
1colA0:aa   ANDEKELLETSELIAGWQKIQEHLGQRYVATARDYADW
  
```

Figura 6 Modulo di inserimento dei codici pdb nel server SSAP. Nel pannello inferiore è riportata la schermata contenente i risultati.

DaliLite Pairwise comparison of protein structures

DaliLite is a program for pairwise structure comparison. Compare your structure(first structure) to a reference structure(second structure).

	RESULTS	SEARCH TITLE	YOUR EMAIL
	<input type="text" value="interactive"/>	<input type="text" value="Testo bioinformatica"/>	<input type="text" value="Stefano.Pascarella@ur"/>
First Structure	PDB entry code: → <input type="text" value="1col"/>	Chain ID: <input type="text" value="A"/>	or upload a file in PDB format (.pdb,.ent,.dat,.brk) <input type="text"/> <input type="button" value="Sfoggia..."/>
Second Structure	PDB entry code: → <input type="text" value="1cpc"/>	Chain ID: <input type="text" value="L"/>	or upload a file in PDB format (.pdb,.ent,.dat,.brk) <input type="text"/> <input type="button" value="Sfoggia..."/>
<input type="button" value="Run DaliLi"/>		<input type="button" value="Reset"/>	

First Structure & Chain: mol1A

No.	Second Structure & Chain	Z-Score	Aligned Residues	RMSD [Å]	Seq. Identity [%]	Structural Alignment	Superimposed C-alpha Traces	PDB Files: mol2 is rotated / translated
1	mol2L	5.8	114	3.7	12	click here	CA 1.pdb	mol1 original.pdb mol2 1.pdb

Figura 7 Modulo di inserimento dei dati sul server DaliLite sul sito dell'EBI. Nel pannello inferiore sono riportati i risultati dell'elaborazione. Attraverso le voci selezionabili con il cursore è possibile visualizzare l'allineamento strutturale e copiare sul proprio calcolatore le coordinate trasformate, in modo da poter visualizzare la sovrapposizione mediante un programma di grafica molecolare.

Gli algoritmi per il confronto strutturale tra proteine sono utilizzati anche dai sistemi di ricerca di somiglianze strutturali in banche dati. Un esempio è il server DALI.

- Collegatevi al sito http://ekhidna.biocenter.helsinki.fi/dali_server/ e avviate una ricerca nella PDB di strutture simili alla catena A di 1COL (**Figura 8**). I risultati sono riportati sottoforma di elenco di strutture in ordine decrescente di *Z-score*. Nell'elenco sono presenti collegamenti attivi che permettono di visualizzare direttamente la sovrapposizione e l'allineamento tra due o più strutture e trasferire le coordinate trasformate sul proprio computer. Esaminate la lista dei risultati e localizzate la catena L della struttura 1CPC (**Figura 9**); confrontate la RMSD, il numero di atomi strutturalmente equivalenti, la percentuale di identità con i risultati precedenti.

Protein Structure Database Searching by DaliLite v. 3

The Dali server is a network service for comparing protein structures in 3D. You submit the coordinates of a query protein structure and Dali compares them against those in the Protein Data Bank (PDB). You receive an email notification when the search has finished. In favourable cases, comparing 3D structures may reveal biologically interesting similarities that are not detectable by comparing sequences.

Requests can also be submitted by e-mail to *dali-server at helsinki dot fi*. The body of the e-mail message must contain atomic coordinates in PDB format.

If you want to know the structural neighbours of a protein already in the Protein Data Bank (PDB), you can find them in the Dali Database.

If you want to superimpose two particular structures, you can do it in the pairwise DaliLite server.

Upload a structure:

Or enter PDB identifier: **chain:** (optional)

(Keyword search for PDB identifiers)

Job name: (optional)

Enter email address for notification: (recommended)

Figura 8 Modulo di avvio di ricerca con DaliLite sul server DALI.

MOLECULE: COLICIN A;

Select neighbours (check boxes) for viewing as multiple structural alignment or 3D superimposition. The list of neighbours to pairwise structural alignment with the query structure, to pre-computed structural neighbours in the Dali Database, and

Expand gaps

Summary

No:	Chain	Z	rmsd	lali	nres	%id	PDB	Description
<input checked="" type="checkbox"/>	1: 1col-A	42.2	0.0	197	197	100	PDB	MOLECULE: COLICIN A;
<input type="checkbox"/>	2: 1col-B	38.3	0.7	197	197	100	PDB	MOLECULE: COLICIN A;
<input checked="" type="checkbox"/>	3: 3few-X	33.8	1.1	197	431	76	PDB	MOLECULE: COLICIN S4;
<input type="checkbox"/>	4: 1rh1-A	33.5	1.2	196	488	60	PDB	MOLECULE: COLICIN B;
<input type="checkbox"/>	5: 1a87	30.9	1.3	194	297	57	PDB	MOLECULE: COLICIN N;
<input type="checkbox"/>	6: 1a87-A	30.9	1.3	194	297	57	PDB	MOLECULE: COLICIN N;
<input checked="" type="checkbox"/>	7: 1cii	14.9	3.1	161	602	22	PDB	MOLECULE: COLICIN IA;
<input type="checkbox"/>	8: 1cii-A	14.9	3.1	160	602	23	PDB	MOLECULE: COLICIN IA;
<input type="checkbox"/>	9: 2i88-A	13.5	3.0	155	178	25	PDB	MOLECULE: COLICIN-E1;
<input type="checkbox"/>	10: 1jbo-A	7.4	3.0	111	162	10	PDB	MOLECULE: C-PHYCOCYANIN ALPHA CHAIN;

Figura 9 Risultati della ricerca con DALI. Le strutture della banca dati vengono elencate in ordine decrescente di *Z-score*. Oltre a questo parametro sono riportati altri dati riguardanti la qualità dell'allineamento (RMSD, numero di carboni equivalenti, numero di residui totali, percentuale di identità). Attraverso gli appositi riquadri si possono selezionare più strutture, in modo da poter visualizzare l'allineamento strutturale e la sovrapposizione con la struttura sonda.