

## **Assassinio in aeroporto**

Diventa un investigatore e indaga sulle cause della morte di un turista americano a Amsterdam.

### **Il caso**

All'aeroporto di Amsterdam è stato trovato il cadavere di un turista americano. Apparentemente sembra aver avuto delle convulsioni e una emorragia interna. Vicino al cadavere è stata trovata una bottiglietta con dei resti di una bevanda che ha l'aspetto del latte.

Viene predisposta un'autopsia: in laboratorio dal contenuto dello stomaco e dal sangue si individuano alcune proteine sospette. Ti viene chiesto di analizzare e identificare queste proteine per capire se una di queste possa essere stata la causa del decesso del turista.

Utilizza il computer per cercare informazioni sulle proteine. La consultazione di alcune banche dati e l'uso di alcuni strumenti bioinformatici potranno aiutarti in questa ricerca.

### **Identificazione della proteina sospetta**

Ti viene data la sequenza aminoacidica delle 4 proteine (p.sospetta 1 – 4 )

Vedi allegato 1

Attenzione: Le sequenze aminoacidiche sono date con il codice a una lettera . Consulta la tabella per trovare le corrispondenze fra il codice a 1 lettera e quello a 3 lettere.

Allegato 2

Adesso puoi incominciare la tua indagine.

Rispondi a queste 5 domande per ciascuna delle proteine sospette.

1. Di che proteina si tratta?
2. Da che organismo proviene?
3. Quale è la sua funzione?
4. Questa proteina può essere la proteina incriminata? Potrebbe essere la responsabile della morte del turista? Perché sì? Perché no?
5. Questa proteina ha altre caratteristiche, rilevanti per l'indagine?

BLAST è il programma usato per confrontare una data sequenza (nucleotidica o aminoacidica) con tutte le sequenze già presenti in banca dati. Per confrontare una sequenza aminoacidica con le sequenze presenti in Swiss Prot (il maggior database di proteine, contiene tutte le proteine conosciute), useremo il software BLAST nel server MRS (uno strumento di accesso a molte banche dati utilizzato dai bioinformatici).

Per familiarizzare con MRS e BLAST, analizziamo insieme la prima proteina sospetta.

Successivamente, ripeterai la ricerca per le proteine sospette 2, 3 e 4 per scoprire quale delle 4 proteine è “colpevole”.

Nella pagina seguente trovi le istruzioni per rispondere alle 5 domande e per interpretare i risultati ottenuti.

## ESERCIZIO 1

Ti appare questa pagina

## ESERCIZIO 2

Fai click su BLAST nella colonna di sinistra per iniziare il programma. Nel riquadro sotto, è stata inserita la sequenza aminoacidica della proteina sospetta numero 1. La prima riga deve sempre incominciare con `>`, seguito dal nome della proteina (questo formato è chiamato FASTA) . Qui abbiamo scelto la “p. sospetta 1”.

The screenshot shows the BLAST web interface. On the left is a navigation menu with links: Home, Blast, Blast results, Clustal, and Databanks. The main content area has a 'Run Blast' button at the top right. Below it is the 'Protein Query' section with a text area containing a FASTA sequence: `>suspect1` followed by several lines of amino acid sequences. Below the text area is a 'Browse...' button. The 'Protein Databanks' section shows 'Choose the databank to search' set to 'Swiss-Prot'. Below that is a field for 'Optionally enter an MRS query to limit the search space'. The 'Blast Options' section has a checkbox for 'Filter sequence (low complexity)' which is checked, and an 'E-value cutoff' set to '10.0'.

### ESERCIZIO 3

Apri l'allegato 1 in cui c'è la lista delle proteine sospette. Le sequenze aminoacidiche sono scritte nel formato FASTA, quello da usare per la ricerca in MRS.

- \* Copia la sequenza p.sospetta1 nella finestra di BLAST
- \* Verifica che l'opzione Filter sequence (low complexity) sia selezionata
- \* Fai click su Run BLAST (in alto a destra)

Nella schermata successiva, vedi BLAST al lavoro. Se molti utenti lo stanno usando nello stesso momento, la tua ricerca può richiedere un pò di tempo. Ricorda che la banca dati che stai consultando contiene centinaia di migliaia di sequenze proteiche!

Quando la ricerca è completata, compare una tabella dei risultati simile a questa (i numeri possono variare a causa dei continui aggiornamenti di SwissProt ):

Nr	ID	Description	Databank	Hits	Best E-value	Status
1	<a href="#">suspect1</a>		<a href="#">sprot</a>	89	5.36E-113	finished

Nella banca dati SwissProt sono state trovate 89 proteine con una sequenza aminoacidica simile a p.sospetta1. Il valore e-Value dice quanto sia alto il livello di similitudine. Se questo valore è basso ( come in questo caso,  $5.36 \times 10^{-113}$ ), il risultato è affidabile .

*NB: Se vuoi ritornare alla pagina dei risultati di altre proteine sospette lo puoi fare facendo click su BLAST results nella colonna di sinistra.*

### ESERCIZIO 4

Fai Click sull'intera riga (diventa blu) per avere ulteriori dettagli sui risultati . Comparirà questa schermata:

Hits 1-11 of 11 page: 1

Nr	ID	Coverage	Description	Hsps	BitScore	E-value
1	<a href="#">casa1_bovin</a>		Alpha-S1-casein precursor.	1	372	7.06E-103
2	<a href="#">casa1_bubbu</a>		Alpha-S1-casein precursor.	1	352	7.4E-97
3	<a href="#">casa1_sheep</a>		Alpha-S1-casein precursor.	1	328	1.24E-89
4	<a href="#">casa1_caphi</a>		Alpha-S1-casein precursor (Alpha-S1-CN).	1	326	4.97E-89
5	<a href="#">casa1_pig</a>		Alpha-S1-casein precursor.	1	131	2.49E-30
6	<a href="#">casa1_camdr</a>		Alpha-S1-casein precursor.	1	119	1.02E-26
7	<a href="#">casa1_rabit</a>		Alpha-S1-casein precursor (Alpha-casein).	1	85	1.75E-16
8	<a href="#">casa1_cavpo</a>		Alpha-S1-casein precursor (Casein-B).	1	80	5.62E-15
9	<a href="#">casa1_human</a>		Alpha-S1-casein precursor [Contains: Casoxin-D].	1	56	9.42E-8
10	<a href="#">casa1_mouse</a>		Alpha-S1-casein precursor (Alpha-casein).	2	46	9.65E-5
11	<a href="#">sok2_yeast</a>		Protein SOK2.	1	30	6.32E0

In questa schermata, trovi la cosiddetta hitlist, cioè la lista delle proteine simili a p.sospetta1 trovate nella banca dati SwissProt . La prima proteina della lista è la più simile alla sequenza “query” cioè alla proteina p.sospetta1 (di fatto, è identica). Il codice di questa proteina è casein alpha s1, bovina.

Il primo risultato è sempre quello della proteina più simile.

Andando da sinistra a destra, per ciascuno degli hits vengono indicati:

- \* il codice della sequenza proteica nella banca dati. il nome della proteina, con il nome (abbreviato) dell’organismo di provenienza, in questo caso il bue.




- \* Coverage: la linea colorata mostra quale parte della sequenza della proteina “query” (p.sospetta1) corrisponde alla sequenza della banca dati (sfondo grigio). Nel caso del primo risultato della lista, la sequenza “query” e la sequenza casa1\_bovin mostrano perfetta corrispondenza per tutta la lunghezza.

- \* Description: una breve descrizione della sequenza della banca dati che corrisponde alla sequenza “query”.

- \* Hsps, Bitscore, and E-value: valori calcolati da BLAST che danno un’idea del grado di affidabilità del risultato. Il risultato migliore è sempre indicato per primo.

## ESERCIZIO 5

Diamo un'occhiata al primo risultato, casa1\_bovin. Fai Click sulla barra colorata, e successivamente sulla barra a due colori. Osserva i risultati.

Nr	ID	Coverage	Description	Hsps	BitScore	E-value		
1	 casa1_bovin		Alpha-S1-casein precursor.	1	372	7.06E-103		
Hsp Nr	Alignment	Score	BitScore	E-value	Length	Identity	Similarity	Gaps
1		954	372	7.06E-103	199	199 (100%)	199 (100%)	0 (0%)
Q:	1 RPKHPIKHQGLPQEVLENLLRFFVAPFPEVFGKEKVNELSKDIGSESTEDQAMEDIKQM						60	
	RPKHPIKHQGLPQEVLENLLRFFVAPFPEVFGKEKVNELSKDIGSESTEDQAMEDIKQM							
S:	16 RPKHPIKHQGLPQEVLENLLRFFVAPFPEVFGKEKVNELSKDIGSESTEDQAMEDIKQM						75	
Q:	61 EAESISSSEEIVPNSVEQKHIQKEDVPSERYLGYLEQLRLKKYKVPQLEIVPNSAEERL						120	
	EAESISSSEEIVPNSVEQKHIQKEDVPSERYLGYLEQLRLKKYKVPQLEIVPNSAEERL							
S:	76 EAESISSSEEIVPNSVEQKHIQKEDVPSERYLGYLEQLRLKKYKVPQLEIVPNSAEERL						135	
Q:	121 HSMKEGIHAQQKEPMIGVNQELAYFYPELFRQFYQLDAYPSGAWYYVPLGTQYTDAPSFS						180	
	HSMKEGIHAQQKEPMIGVNQELAYFYPELFRQFYQLDAYPSGAWYYVPLGTQYTDAPSFS							
S:	136 HSMKEGIHAQQKEPMIGVNQELAYFYPELFRQFYQLDAYPSGAWYYVPLGTQYTDAPSFS						195	
Q:	181 DIPNPIGSENSEKTTMPLW	199						
	DIPNPIGSENSEKTTMPLW							
S:	196 DIPNPIGSENSEKTTMPLW	214						

E' mostrato l'allineamento delle due sequenze aminoacidiche. Q (query) è la sequenza che tu hai inserito, in questo caso p.sospetta1. S è la sequenza della proteina della banca dati che mostra la maggior somiglianza con p.sospetta1, nel nostro caso casa1\_bovin. Nella riga fra le due sequenze sono indicate le lettere corrispondenti agli aminoacidi identici nelle due sequenze Q e S. Se non tutti gli amino acidi corrispondono, vengono inseriti dei *gap* (spazi vuoti). Ne vedremo più avanti.

Come vedi, la sequenza p.sospetta1 è identica al 100% alla sequenza della caseina alpha-S1- (tutti i 199 amino acidi sono identici). In conclusione, sei riuscito a identificare una delle 4 proteine. Fai Click su casa1\_bovin per avere tutte le informazioni conosciute su questa proteina, per scoprire se essa può essere coinvolta nella morte del turista.

## ESERCIZIO 6

La scheda della caseina bovina casa1\_bovin contiene tutte le informazioni che ti servono per rispondere alle 5 domande.

1. Di che proteina si tratta?
2. Da che organismo proviene?
3. Quale è la sua funzione?

4. Questa proteina può essere la proteina incriminata? Potrebbe essere la responsabile della morte del turista? Perché sì? Perché no?

5. Questa proteina ha altre caratteristiche, rilevanti per l'indagine?

Se non sai dove trovare le informazioni richieste, usa questi suggerimenti:

- \* Che informazioni trovi sotto Protein name?

- \* Che informazioni trovi sotto From?

- \* Che informazioni trovi sotto Keywords? Sono rilevanti per la tua indagine?

- \* Che informazioni trovi sotto Comments? Sono rilevanti per la tua indagine?

Quali sono le tue conclusioni? La proteina p.sospetta1 è coinvolta nella morte del turista? Sì o No?

## ESERCIZIO 7

Adesso ripeti la stessa analisi per le proteine sospette.

Fai una ricerca BLAST con le sequenze delle altre 3 proteine e trascrivi i risultati.

Quali sono le conclusioni finali sull'omicidio? Quale è stata la causa della morte del turista?

Risposta:

Vuoi continuare le indagini.....

Bene, hai identificato la causa della morte della vittima. Alla polizia il compito di completare le indagini. Durante la tua ricerca, hai imparato molte cose su una parte della bioinformatica: l'analisi di sequenze proteiche interrogando banche dati per mezzo di software specifici. Questo ti ha portato a concludere la tua inchiesta.

Se hai ancora tempo e voglia di continuare la tua ricerca bioinformatica, potrai scoprire altre straordinarie possibilità offerte dagli strumenti bioinformatici. In questa parte supplementare, troverai alcuni esercizi per esplorare ulteriormente le sequenze proteiche, gli aminoacidi più importanti, le omologie tra organismi, etc.

## Supplemento di indagine sulla proteina sospetta

In questa parte, non dovrai fare nessuna nuova ricerca BLAST. Ritorna alla tua precedente ricerca MRS session e utilizza i risultati BLAST già ottenuti.

Probabilmente, già sai che proteine di organismi diversi che mostrano alta somiglianza nelle sequenze (dette omologhe) hanno spesso funzioni simili.

Adesso approfondiremo questo aspetto, partendo dalle 4 proteine che hai appena identificato.

### ESERCIZIO 8

Torna ai risultati BLAST nella colonna sinistra della schermata di MRS, e riprendi i risultati della proteina sospetta2.

Il primo risultato è lacb\_bovin. La p.sospetta2 è una  $\beta$ -lactoglobulina bovina. Adesso cerca la  $\beta$ -lactoglobulina della capra. Troverai che le sequenze delle due proteine sono identiche al 96%. Solo pochi amino acidi sono diversi nelle due sequenze.

Trova quanti sono gli amino acidi diversi nelle  $\beta$ -lactoglobuline di capra e bovina.

Quale è la tua risposta?

Come prima, la riga fra le due sequenze contiene tutti gli amino acidi identici nelle sequenze Q e S. In questo caso, alcune posizioni sono vuote o contengono un segno +.

Ad esempio, nella prima riga dell'allineamento si trova una L (leucina) nella sequenza bovina e una I (isoleucina) nella sequenza di capra. I due amino acidi leucina e isoleucina sono molto simili, segno + nella riga centrale. In alcune posizioni, nella riga centrale c'è un *gap* (spazio vuoto), ad esempio subito prima della fine della sequenza. Nella sequenza bovina c'è una E (glutammato) e nella sequenza di capra c'è una G (glicina). La glicina e il glutammato sono amino acidi molto diversi, e questo è indicato dallo spazio vuoto.

Usando questo tipo di rappresentazione quando si confrontano due sequenze, si possono identificare facilmente e rapidamente le regioni in cui le proteine sono uguali o diverse.

## ESERCIZIO 9

Usando la lista dei risultati BLAST, confronta le sequenze della  $\beta$ -lactoglobulina bovina e canina e trova quanti sono gli amino acidi identici.

Quale è la tua risposta?

## ESERCIZIO 10

Guarda i risultati BLAST del confronto delle sequenze della  $\beta$ -lactoglobulina bovina e canina.

```
Bovine: 1 LIVTQTMKGLDIQKVAGTWYSLAMAASDISLLDAQSAPLRVYVEELKPTPEGDLEILLQK 60
++V +TM+ LD+QKVAGTW+S+AMAASDISLLD+++APLRVY++EL+PTP+ +LEI+L+K
Dog: 1 IVVPRTMEDLDLQKVAGTWHSMAMAASDISLLDSETAPLRVYIQELRPTPDNLEIVLRK 60

Bovine: 61 WENGEC AQKKIIAEKTKIPAVFKIDALNENKVLVLDTDYKKYLLFCMENSAEPEQSLACQ 120
WE+G CA++K++AEKT++PA FKI+ + EN++ +LDTDY YL FC N+ P+QSL CQ
Dog: 61 WEDGRCAEQKVLAEKTEVPAEFKINYVEENQIFLLD TDYDNYLFFCEMNADAPQQLMCQ 120

Bovine: 121 CLVRTPEVDDEALEKFDKALKALPMHIRLSFNPTQLEEQCHI 162
CL RT EVD+E +EKF++ALK LP+H++L NPTQ EEQC I
Dog: 121 CLARTLEV DNEVMEKFN RALKTL PVHMQ L LNPTQAEEOCLI 161
```

Tutti gli amino acidi importanti della  $\beta$ -lactoglobulina sono identici nelle due sequenze. Sono stati mantenuti uguali (conservati) durante l'evoluzione. Ad esempio, nella proteina bovina esiste un ponte disolfuro tra Cys66 e Cys160. Queste due cisteine sono evidenziate. Come puoi vedere, due cisteine si trovano esattamente nella stessa posizione nella proteina di cane. I numeri corrispondenti sono 66 e 159 (invece di 66 e 160). In posizione 150 della sequenza della proteina bovina c'è una Ser (S) che manca nella proteina di cane. Questo è il risultato di una delezione, ossia della perdita, durante l'evoluzione, di 3 nucleotidi nel gene del cane. Di conseguenza, la proteina del cane ha un amino acido in meno.

## ESERCIZIO 11

Adesso guardiamo da vicino la sequenza della proteina p.sospetta3, la metalloproteinasi atrolisina del serpente a sonagli, che produce forti emorragie



nelle vittime morse dal serpente. Usando BLAST results, ritorna alla hit list della proteina p.sospetta3.

Guarda i primi risultati della lista. Noterai che tutte le proteine sono metalloproteinasi. In posizione 13 (o nelle vicinanze) , trovi acl3\_agkac. Questa è una metalloproteinasi del veleno del serpente “Cento-passi” (*Deinagkistrodon acutus*). Anche questa proteina provoca emorragie interne. Le due sequenze amino acidiche mostrano 54% di identità.

L'allineamento è il seguente:

```
rattlesnake: 5 PQRYIELVVVADHRVFMKYNSDLNTIRTRVHEIVNFINGFYRSLNIHVSLTDLEIWSNED 64
PQ IEL ++ DH ++ KYNS+ + I T + VN +N Y SLN+ ++L+ +E+WS D
hundred-pace: 3 PQTSIELFLIVDHSYAKYNSNSSKITTTTLKARVNIMNAIYSSLNLVITLSGIEMWSAAD 62

rattlesnake: 65 QINIQSASSDTLNAFAEWRETDLNLRKSHDNAQLLTAIELDEETLGLAPLGTMCDPKLSI 124
I +QS+S +TL FA WRETDL R S+DNAQLLTA + T+GLA L TMC+ K S+
hundred-pace: 63 LITVQSSSRNTLKLFWRETDLKRTSNDNAQLLTATNFNGNTVGLAYLKTMCNSKYSV 122

rattlesnake: 125 GIVQDHSPINLLMGVTMAHELGHNLGMEHDGKDCIRGASLCIMRPGTLKGRSYEFSDDSM 184
G++QDHS I LLM VTMAHELGHNLGM HDG C + CIM P L+ G + FSD S
hundred-pace:123 GLIQDHSAIPLLMVAVTMAHELGHNLGMNHDGAGC--SCATCIMAPVLSSGPAKSFSDCSK 180

rattlesnake: 185 HYYERFLKQYKPOCILN 201
H Y+ FL +KPQC+LN
hundred-pace:181 HDYQSFLTIIHKPOCLLN 197
```

Sebbene differiscano all'incirca per il 50% , entrambe sono metalloproteinasi di serpente.

Dall'allineamento si possono fare delle interessanti osservazioni:

- \* Le due proteine hanno una regione estremamente simile, mentre sono molto diverse nella parte finale (amino acidi 160-185). Gli amino acidi responsabili delle differenze fra le due proteine sono concentrati in questa regione.

- \* L'amino acido glutammato (E evidenziata) è l'amino acido del sito attivo dell'enzima. Questo è l'amino acido più importante di tutta la proteina, perchè è responsabile dell'attività della metalloproteinasi. I due veleni hanno questo amino acido nella stessa posizione, sia nella sequenza che nella struttura tridimensionale.

- \* Ci sono 3 istidine (H evidenziate) che coordinano il legame con un atomo di zinco conservate nelle due sequenze (diremo di più di queste istidine nella parte successiva, sulla struttura 3D del farmaco).

- \* Ci sono dei legami disolfuro in entrambe le proteine; uno di questi è presente nelle due proteine in posizione equivalente: Cys118-Cys198 nel serpente a sonagli,

che corrisponde a quella delle Cys116-Cys194 nel serpente “cento-passi” (in evidenza). Questo ponte disolfuro è stato conservato durante l’evoluzione.

Tuttavia, nel serpente a sonagli c’è un ulteriore ponte disolfuro (Cys158-Cys165), mentre nell’altro serpente ce ne sono due: Cys156-Cys178 e Cys158-Cys161. Trova le posizioni di queste cisteine nell’allineamento e evidenziale.

Dove sono?

Possiamo dedurre che questi ponti disolfuro non sono essenziali per la funzione di queste proteine, ma piuttosto possono spiegarne le differenze!

### **In conclusione**

Se vai un pò avanti nella lista dei risultati di BLAST, troverai anche il veleno del serpente Habu di Okinawa (codice hr2\_trifl). Questa è un’altra metalloproteinasi, al 52% identica all’atrolisina del serpente a sonagli (e al 51% al veleno del serpente “cento-passi”), ma questa non causa emorragie.

Da questa osservazione si capisce come le diverse caratteristiche delle proteine dipendono da pochi amino acidi.

A questo punto, hai imparato molte cose su una parte della bioinformatica: l’analisi di sequenze proteiche interrogando banche dati per mezzo di software specifici.

Ma con la bioinformatica si può fare ben di più che analizzare delle sequenze!

Nella parte successiva, vedremo come analizzare le strutture tridimensionali (3-D) delle proteine e come è possibile disegnare un antidoto contro il veleno atrolisina.

### **Come disegnare un farmaco in 3D**

#### **Introduzione**

L’atrolisina, il veleno del *Crotalus atrox*, meglio conosciuto sotto il nome di serpente a sonagli dal dorso a losanghe è molto potente. Negli Stati Uniti il morso di questo serpente è la causa del maggior numero di morti da morso di serpente riscontrati in un anno. Il veleno è una metalloproteinasi, un enzima che distrugge le pareti dei vasi sanguigni, provocando gravi emorragie interne che portano a morte le vittime. E’ ovvio che è necessario identificare urgentemente un antidoto a questo veleno .



*Crotalus atrox*, serpente a sonagli dal dorso a losanghe

In questa esercitazione bioinformatica, cercheremo di disegnare un antidoto a questo veleno, partendo dalla struttura tridimensionale 3-D dell'atrolisina. La struttura spaziale di una proteina è molto importante per la sua funzione. Capire la struttura 3-D di una proteina è indispensabile per risolvere problemi dovuti alla presenza di proteine alterate (come accade in molte malattie genetiche) o a proteine tossiche con cui il nostro organismo può venire a contatto, come in questo caso.

Per procedere, abbiamo bisogno di avere alcune conoscenze di base sulla struttura delle proteine, sugli aminoacidi e sulle interazioni e i legami fra atomi, come ad esempio i legami idrogeno.

Se ritieni di non avere sufficienti conoscenze di base in questo campo, prima di continuare vai a rivedere questo argomento nella parte teorica.

Iniziamo dalla struttura primaria

Nella prima parte, hai avuto a che fare con la struttura primaria dell'atrolisina, è cioè con la sua sequenza amino acidica. Dalla struttura primaria dipende la struttura terziaria (o tridimensionale 3-D) della proteina, fondamentale per espletare la sua funzione. Una proteina con alterazioni della struttura primaria e di conseguenza anche della struttura terziaria può essere una proteina che ha perso parzialmente o completamente la sua funzione. Tuttavia, piccole modificazioni della struttura primaria possono non avere alcun effetto sulla funzione della proteina o causare solo modeste variazioni della sua attività. In una proteina non tutti gli aminoacidi hanno la stessa importanza. Gli studi sugli effetti delle mutazioni e gli allineamenti multipli sono molto informativi per capire l'importanza dei singoli aminoacidi all'interno di una proteina.

## ESERCIZIO 1

Spiega perchè la mutazione sistematica, uno ad uno, di tutti gli aminoacidi di una proteina può aiutare a capire quali sono gli aminoacidi importanti.

Quale è la tua risposta?

Negli enzimi, gli aminoacidi più importanti si trovano nel sito attivo e cioè nella regione della proteina in cui effettivamente avviene la reazione enzimatica. Prima di andare a vedere quale è il sito attivo del nostro veleno, iniziamo a fare alcuni esercizi per familiarizzare con la visualizzazione delle strutture 3-D.

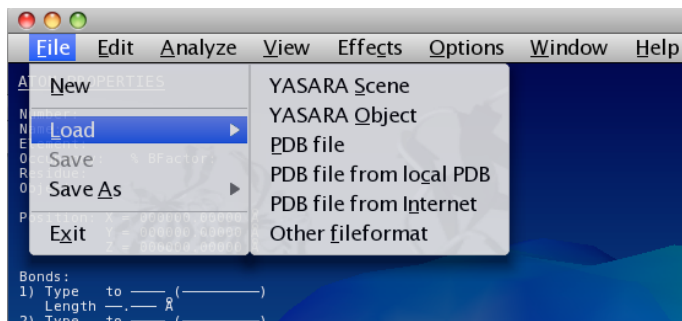
Impariamo a usare Yasara

Yasara è un programma di grafica utile per visualizzare e manipolare I modelli delle proteine. E' stato sviluppato da Elmar Krieger, al CMBI (Centre for Molecular and Biomolecular Informatics) in Olanda. Se non hai installato il software Yasara nel computer, cercalo nella sezione software.

Quando avrai imparato ad usare questo software, potrai andare a vedere la struttura di altre proteine interessanti. Con questo programma puoi creare immagini da inserire in lavori e progetti scientifici. Puoi trovare una descrizione dettagliata nella sezione: More in 3D.

## ESERCIZIO 2

Avvia YASARA. In YASARA, usa il menu in alto a sinistra per caricare una proteina. File > Load > YASARA Scene e poi scegli introduction.sce e fai click su OK.



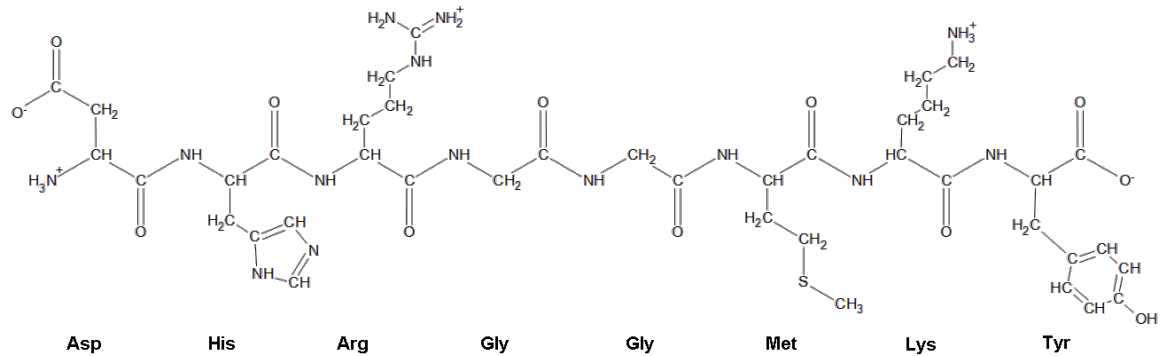
Vedrai la struttura del peptide con la sequenza Asp-His-Arg-Gly-Gly-Met-Lys-Tyr nella rappresentazione detta a “ball and stick”. I singoli atomi sono rappresentati da pallini, e sono connessi fra loro da bastoncini che rappresentano i legami chimici. NB: per semplificare il modello gli atomi di idrogeno non sono rappresentati.

Noterai che il peptide è avvolto in una “nuvola” che rappresenta la superficie di Van der Waals. Questa superficie mostra lo spazio effettivamente occupato dagli atomi: in una proteina non esiste praticamente alcuno spazio vuoto fra un atomo e l’altro.

Muovendo il mouse lungo la parte inferiore dello schermo di YASARA, compare la barra della sequenza. Per mantenere visibile la barra, clicca sul pulsante blu all’estremità sinistra della barra. Cliccando sui singoli aminoacidi compare (si illumina) il carbonio alfa ( $C\alpha$ ). Premendo il tasto Ctrl mentre si clicca un aminoacido, si può far ruotare la proteina e avvicinarla con lo zoom in modo da vedere distintamente l’atomo  $C\alpha$ . Tenendo premuto il tasto del mouse e muovendolo, puoi manipolare la proteina. Prova a fare le azioni seguenti:

- \* Pulsante sinistro: ruotare
- \* Pulsante di centro: trasportare
- \* Pulsante destro: Zoom

Gli atomi sono colorati secondo un codice specifico per ognuno. Cliccando su un atomo, nella parte sinistra della finestra compaiono ulteriori informazioni su di esso.

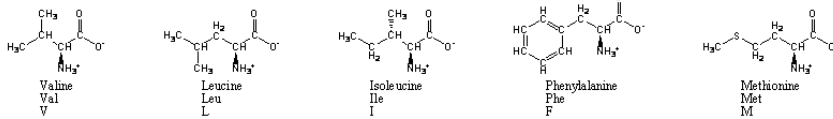


Struttura 2D del peptide

### ESERCIZIO 3

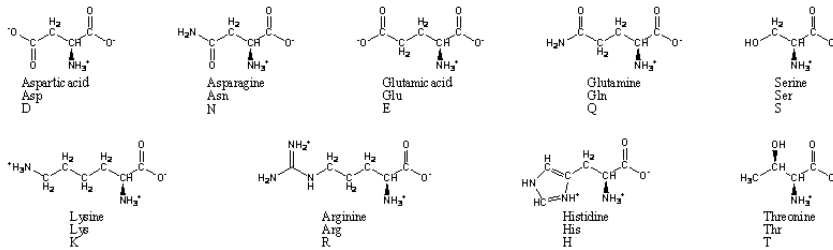
A. Usando questa lista di aminoacidi, identifica quali sono gli atomi presenti nelle proteine.

#### Amino acids with hydrophobic side chains

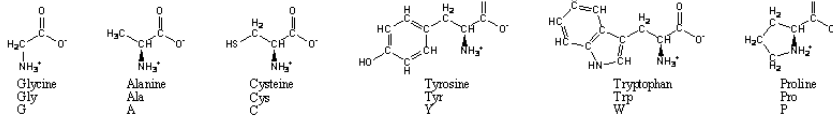


Quale è la tua risposta?

#### Amino acids with hydrophilic side chains



#### Amino acids with intermediate side chains



B. Quale è il codice colore usato in YASARA per i diversi atomi?

Rosso =

Blu =

Verde =

Azzurro =

#### ESERCIZIO 4

In questa rappresentazione, i legami covalenti si vedono come bastoncini corti di un colore specifico. Quale è la differenza fra i legami gialli e quelli bianchi in questa rappresentazione?

Quale è la tua risposta?

Nella struttura 3-D sono mostrati 3 dei 4 diversi tipi di interazioni importanti per l'assunzione della struttura terziaria (*folding*) di una proteina (per ulteriori informazioni, vedi la parte teorica). Questi tre diversi tipi di interazioni sono rappresentati da bastoncini verdi, blu e arancione.

#### ESERCIZIO 5

1. A quale tipo di interazione corrisponde ciascun colore?
2. Localizza queste interazioni nella rappresentazione 2-D del peptide.

### **Il veleno atrolisina**

Adesso hai informazioni sufficienti per analizzare il veleno atrolisina con cui hai fatto conoscenza nelle indagini sul delitto in aeroporto.

Incomincia a caricare la struttura 3-D della molecola.

#### ESERCIZIO 6

Inizia una nuova ricerca.

Con YASARA apri il file veleno.pdb (via File > Load > PDB file (come nell'esercizio 2)).

Apparirà la rappresentazione a pallini del veleno. Si vedono tutti gli atomi, salvo quelli di idrogeno. Questa rappresentazione non è molto chiara, ma fortunatamente esistono delle rappresentazioni alternative. Puoi navigare da una all'altra premendo i tasti da F1 a F8 sulla tastiera:

F1: rappresentazione a pallini (singoli atomi)

F2: rappresentazione a pallini e bastoncini (Ball-and-stick) (singoli atomi + legami)

F3: rappresentazione a bastoncini (gli atomi sono dei bastoncini)

F4: traccia dei C $\alpha$  (sono rappresentati solo i C $\alpha$ , connessi con bastoncini)

F5: traccia dell'ossatura della proteina (senza catene laterali)

F6: rappresentazione a cartoon (elementi della struttura secondaria)

F7: rappresentazione a cartoon alternativa

F8: Aggiunge/Rimuove le catene laterali degli amino acidi (gruppi "R") nelle rappresentazioni a bastoncini e a pallini e bastoncini

## ESERCIZIO 7

Con i tasti F5, F6 e F7 si possono facilmente mettere in evidenza gli elementi della struttura secondaria. Quale è la struttura secondaria delle regioni rosse e blu della proteina?

## ESERCIZIO 8

Trova i seguenti cinque elementi (o strutture):

Un atomo di ossigeno con un doppio legame

Un legame peptidico

La catena laterale di una istidina

Una alfa elica

Un ponte disolfuro

Utilizza per ciascun elemento la rappresentazione appropriata e prendi nota.

Esempio: Un atomo di ossigeno con un doppio legame. Rappresentazione a pallini e bastoncini.

## Il sito attivo

Adesso andiamo a cercare il sito attivo della proteina, e cioè il sito in cui avviene effettivamente la reazione catalizzata dall'enzima. Dato che non disponiamo di

nessuna proteina con mutazioni nel sito attivo, dobbiamo trovare un altro sistema per identificarlo.

Un aspetto caratteristico dei siti attivi degli enzimi è che, nella maggior parte dei casi, si trovano in tasche o cavità sulla superficie della proteina. Quindi, il sistema più sbrigativo per identificare il sito attivo di un enzima è quello di cercare la cavità più ampia presente sulla superficie della proteina.

#### ESERCIZIO 9

Perché il sito attivo si trova in cavità della superficie?

#### ESERCIZIO 10

Cerca il sito attivo facendo ruotare la proteina e avvicinandola con lo zoom. Scegli la rappresentazione che ritieni più appropriata. Non superare il minuto in questa ricerca. Se dopo un minuto non hai ancora trovato niente, non ti preoccupare. Nella pagina seguente useremo un altro sistema per identificare il sito attivo, basandoci sul fatto che molti enzimi utilizzano ioni metallici per svolgere la loro attività.

#### ESERCIZIO 11

Conosci almeno una proteina umana che usa uno ione metallico per la sua attività?

Localizzare gli ioni metallici, è un sistema alternativo per identificare i siti attivi degli enzimi. Attenzione: ci sono molte altre ragioni per cui si può trovare uno ione metallico in una struttura 3-D di una proteina, quindi per identificare il sito attivo non affidarti ciecamente a questo metodo.

#### ESERCIZIO 12

Inizia una nuova ricerca con YASARA cliccando File > New e quindi Yes. Poi carica il file veleno2.pdb via File > Load > PDB file (come hai già fatto in precedenza).

#### ESERCIZIO 13

Uno ione zinco ( $\text{Zn}^{2+}$ ) è molto ben visibile come un pallino viola nella rappresentazione appropriata. Usando questo ione, cerca di trovare il sito attivo.

Suggerimento: è più facile usando la barra della sequenza: alla fine della sequenza si parla di Zn. Cliccando su Zn mentre si tiene premuto il tasto Ctrl, la proteina automaticamente ruota e lo ione diventa visibile.



## Disegna l' antidoto

Adesso che conosciamo il sito attivo, possiamo cercare di disegnare un antidoto. Partiremo dal principio della chiave e della serratura usato dagli enzimi per svolgere la loro azione.

### ESERCIZIO 14

Spiega in poche parole in che cosa consiste il principio della chiave e della serratura, eventualmente aiutandoti con un disegno.

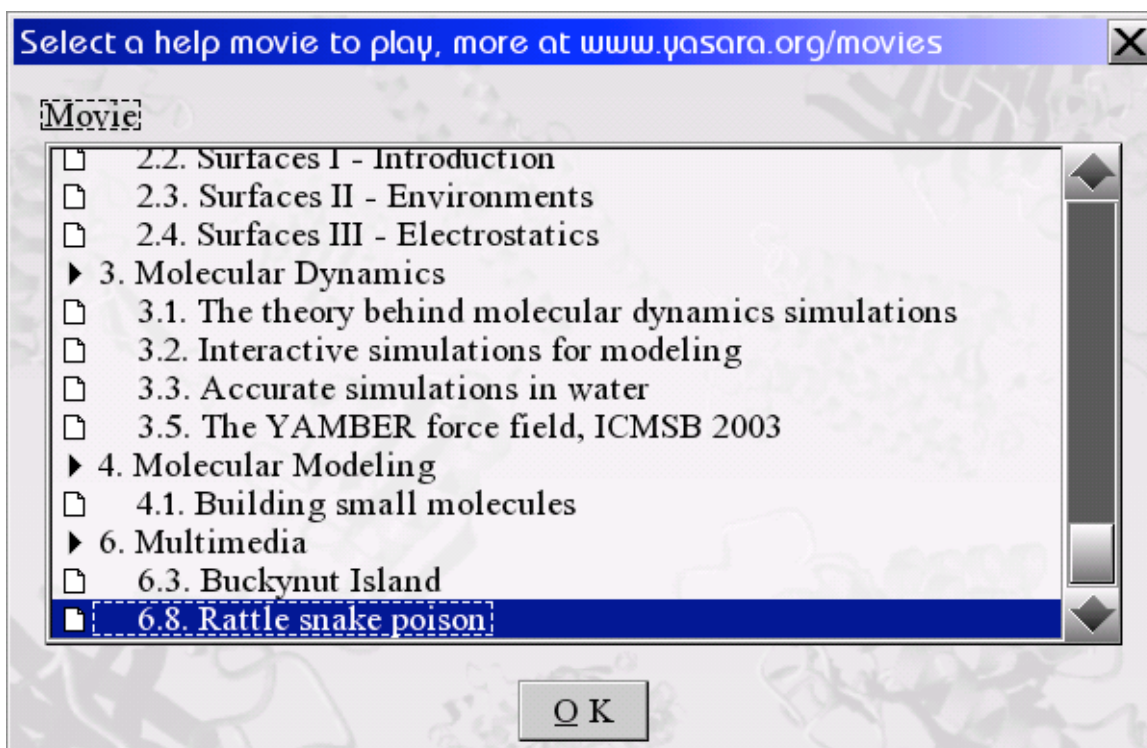
Probabilmente sai (magari per esperienza diretta) che una serratura può bloccarsi se viene occupata da un corpo estraneo. La chiave non può più entrare e la serratura è fuori uso. Il nostro antidoto deve funzionare secondo lo stesso principio.

Dobbiamo quindi cercare una molecola che sia in grado di legarsi in modo irreversibile al veleno e che quindi lo renda inattivo.

Hai già visto che l'avvolgimento di una proteina dipende da specifiche interazioni fra atomi. Anche il legame dell'antidoto si basa su interazioni simili. Dobbiamo quindi cercare una molecola che sia in grado di formare legami idrogeno, interazioni idrofobiche e legami ionici e che, inoltre, si adatti alla cavità del sito attivo.

### ESERCIZIO 15

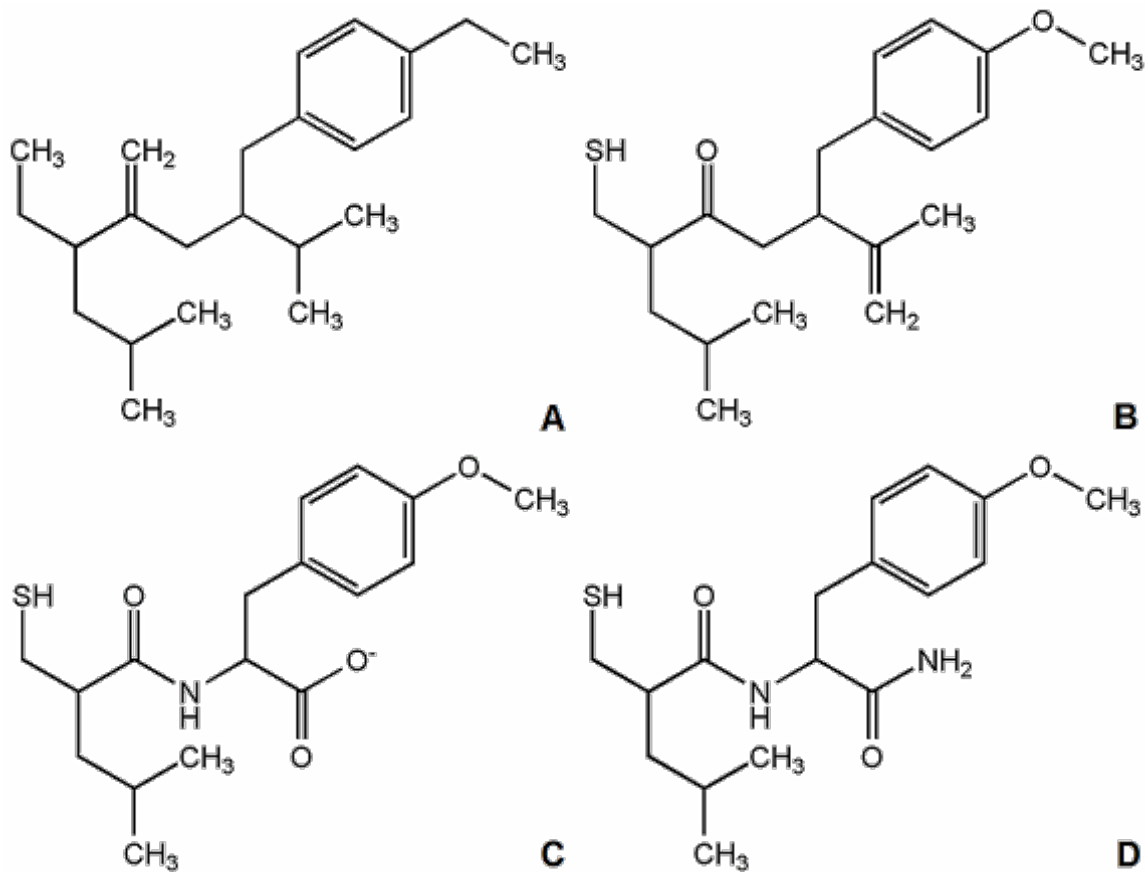
Apri Help > Play help movie > 6.8 Rattle snake poison > OK e guarda il filmato. Nella animazione si vede una piccola molecola che entra nel sito attivo e si lega alla proteina velenosa. Una piccola molecola di questo tipo che alloggia bene in una



cavità di una proteina si chiama ligando. Questa molecola, tuttavia, è composta solo da atomi di carbonio e non è in grado di formare molte interazioni forti con il sito attivo. Nell'ultima parte di questa esercitazione, ti verranno mostrate tre versioni modificate del ligando. A te decidere quale delle tre funzionerà meglio come antidoto. Il ligando migliore è quello che forma il maggior numero di interazioni con la proteina e che quindi dovrebbe essere anche l'antidoto migliore.

### Trova l'antidoto migliore

Sono mostrate 4 molecole diverse. La prima (A) è il ligando visto nell'animazione, le altre tre (B, C e D) rappresentano le versioni 1, 2 e 3 dell'antidoto. Per ciascuna di queste, in YASARA esiste un modello 3-D.



## ESERCIZIO 16

Guardando le diverse strutture degli antidoti, identifica gli atomi che possono essere coinvolti in interazioni con la proteina.

## ESERCIZIO 17

Analizza i diversi modelli dell'antidoto in YASARA.

Per facilitarti il compito, in YASARA puoi trovare delle simulazioni che ti consentiranno di valutare rapidamente le differenze fra i modelli. Apri facendo click su File > Load > YASARA Scene. Scegli ligandB.sce.

Si aprirà una rappresentazione “ ball-and-stick “ del ligando. Il veleno è rappresentato con un cartoon, e gli amino acidi vicini al ligando sono nella rappresentazione con gli stick.

Dopo aver osservato ligandB.sce, passa a ligandC.sce e ligandD.sce.

### **Domanda finale:**

Sapresti ritrovare la posizione dello zinco negli schemi degli antidoti confrontandoli con la ricostruzione in 3D?

Quale dei 3 ligandi ritieni sia l'antidoto migliore? Spiega il perchè.

Verifica la tua risposta osservando aprendo i file delle interazioni, rappresentate come bastoncini arancioni, dell'antidoto con il veleno. Caricali via File > Load > YASARA Scene > intB.sce, intC.sce and intD.sce.

### **Un vero antidoto**

Sei arrivato alla fine, e dovresti aver identificato quale è l'antidoto giusto. Questa molecola viene usata effettivamente come antidoto al veleno del serpente a sonagli.

Questo è un esempio di come la Bioinformatica abbia dato un contributo essenziale per risolvere un problema di biomedicina.