

Modellering av sekvensen til Cape Cobra toksinet

SERGIO MANTZETTI

Som første trinn i serien til opplæring i studier av funksjoner og strukturer til proteiner, iverksettes det her modellering og analyse av en slange-gift som agerer som inhibitor av proteaser. Som vi ser, så skal vi omsider modellere den ukjente strukturen til slange-giften til «Cape Cobra» (Fig 1). Inhibitoren tilhører familien BPTI (bovine pancreatic trypsin inhibitor), en kjent protease-inhibitor familie. Inhibitorers mekanismer er å binde seg til enzymet den er komplementær til, uten å bli konvertert til produkt. Inhibitoren har derfor svært like egenskaper som enzymets naturlige substrat, men gitt visse strukturelle/funksjonelle forskjeller ved det inhiberende setet, greier den å binde seg med høy K_m uten å legge grunn for katalyse.

DET FØRSTE DU gjør er å laste ned programvaren du trenger. Hent ned Swiss PDB Viewer på <http://www.expasy.org/spdbv> og installer den et sted på harddisken din.

1. INNHENTING AV SEKVENNS.

Til å begynne med skal vi hente proteinsekvensen fra PubMed Central (<http://www.ncbi.nlm.nih.gov/entrez>). Søkeord er «snake venom protease inhibitor», og vi velger tredje sekvens med id-navn TINJVC.



Figur 1: Cape Cobra, Naja nivea.

Når du klikker deg inn på siden til Inhibitoren:

- Velg rullemenyen ved siden av Display (der det står GenPept) og klikk FASTA.
- FASTA er formatet på sekvensen, enkel ett-bokstavs format per aminosyre. Klikk så «Display».
- Når du ser sekvensen, klikk så «Send» og vær sikker på at det står «all to file».
- Last ned og lagr sekvensen som «Cape Cobra.txt» i en katalog kalt NBS-modellering et sted på disken din.

2. SØK ETTER TEMPLATER.

Nå skal vi finne et templat-struktur til å modellere vår sekvens på. Til dette bruker vi SWISS Model, <http://www.expasy.org/swissmod>.

- Klikk «Search» på menyen til venstre.
- Lim inn sekvensen til slangetoksinen i vinduet og klikk «Submit Request».
- Listen som dukker opp inneholder de strukturene som ligner mest på vår sekvens

Første strukturen er krystallstrukturen til dendrotoksinet fra Black Mamba. Klikk der det står idtk_, og last så ned struktur-filen på hard-disken din. Klikk også den andre filen under denne, rjc6A. Dette er en annen inhibitor.

Klikk så på «detail» til idtk_. Her ser du at sekvensen til Cape Cobra (Query) mot sekvensen til Black Mamba (Sbjct) inneholder mange fragmenter som er like sekvensen til Cape Cobra, men ikke alle områder er komplementære. Disse «gaps» eller tomrom blir viktige å se på i sammenheng med multippel sekvens oppstillingen. Ved å ha nedlastet tre strukturer istedenfor bare ett, har man mulighet til å se flere mønstre av strukturell evolusjon i forhold til lengde og posisjon av konserverte beta-plater og alfa-helikser, og hvor «coiled-regions» inntreffer.

3. MODELLERING.

I denne prosedyren, skal vi modellere en 3-D modell til sekvensen til protease-inhibitoren til Cape Cobra, basert på 3 lignende krystallstrukturer. Vi åpner SWISS PDB Viewer, og ved å gå på «File» og «Open PDB File» åpner vi i følgende rekkefølge: 1) idtk_ 2) rjc6A. Gå så til «SWISS Modell» øverst til høyre på meny linjen, og gå til «Load Raw sequence to Modell». Velg nå fasta-sekvens filen som du nedlastet fra NCBI hjemmeside i punkt 1.

1. For oversiktens skyld, gå til «Window» så til «Alignment».

2. For å rette opp de to templat-strukturen korrekt til hverandre, gå til «Fit», så til «Iterative Fit», som søker etter beste strukturell oppstilling av templatene. Velg her idtk_ som øverst alternativ (referanse strukturen for templatene) og så rjc6A, klikk så OK.

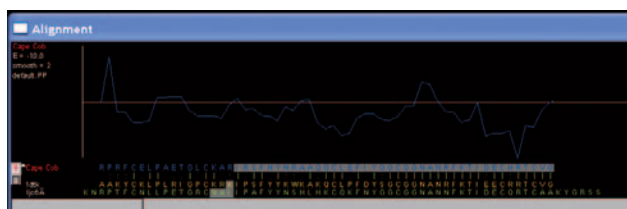
3. Klikk så på Cape Cobra sekvensen i alignment vinduet, så den blir rød (aktivisert), og så gå til «Fit» og «Magic Fit».

4. Nå er sekvensen oppstilt etter idtk_ som hovedreferanse, og rjc6A som sekundær referanse for modelleringen.

4. THREADING (OPPSTILLING AV SEKVENSMOT KRYSTALLSTRUKTUR-TEMPLAT)

«Threading» er en viktig prosedyre som tillater oss å optimalisere oppstillingen av sekvensen mot modellen. Hovedtrekkene i threading er (listet i prioritert rekkefølge):

- bevaring av alfa helikser i henhold til mest



Figur 2: Manuell sekvensalignement.

konserverte mønstre.

- bevaring av beta plater i henhold til mest konserverte mønstre.

- god hydrogenbånd-nettverk mellom hydrofile atomer i kjernen av proteinet. Minst mulig «misfornøyde» hydrofile atomer i kjernen. Fornøyde hydrofile atomer er parret i H-bindinger.

- identifisering og beskyttelse av hydrofobe regioner i loops slik at de ikke er eksponerte mot en imaginær vannfase og vendes inn mot kjernen av proteinet.

- Minimalisert total energi av hele proteinet, slik at den energisk-laveste konfigurasjonen er noenlunde identifisert.

Prioritetene listet her er med andre ord de energetisk optimale konfigurasjonene innvendig i proteinet. Jobber man med multimerer, så søker man også etter en optimal energetisk konfigurasjon mellom dem. Bevaring av lavest energi er fundamental i modellering, fordi energien til et protein/enzym må være lavest mulig slik at stabiliteten til den bevares ved økninger i energier under katalyse og interaksjoner mellom proteiner. Lavest energi prinsippet for strukturer er universelt for modellering og er tross alt hovedmålet med alle strukturell-bioinformatikk metoder.

For at threading'en skal bli mest mulig presis, bør vi vende oss til tanken på å gjøre mest mulig manuelt. I 1999 dukket det dog opp et automatisk threading script tilgjengelig på Internett som er svært effektivt, SAM-T99. SAM-T99 er et automatisk threading script, utviklet ved University of California av Professor Karplus. SAM-T99 er tilgjengelig på: <http://www.so.e.ucsc.edu/research/compbio/HMM-apps/T99-tuneup.html>.

For å forberede sekvensene for threading, klikker du på hver av PDB templatene, idtk og rjc6 i «alignment» vinduet så de blir røde (aktivisert). Så går du til «File» og «Save» og så «Alignment». Lagre så alle sekvens-filene i en gitt mappe.

Vi begynner threading-prosedyren: Gå til SAM-T99 siden.

Sett inn email adressen din. Sett så inn sekvensen til Cape Cobra i øverste vindu, og sekvensene til hhv. rdtk og tjc6 en av gangen i nederste vindu. Klikk «Press here to Submit». Du skal da få tilsendt i emailen dine totalt 2 parvis «alignment-tuneup» mellom Cape Cobra og de to sekvensene hver for seg.

Når du har mottatt alignment tuneup, reproduserer du den i alignment-vinduet i Swiss PDB Viewer. Dette gjøres ved å markere sekvens-områder på Cape Cobra og holde SHIFT-tasten inne. Når du har markert et område du vil flytte langs oppstillings-rekken, bruker du høyre og venstre pil.

Når oppstillingen er gjort korrekt etter SAM-T99 kan du velge å gå inn og sjekke manuelt at det ikke finnes noen abnorme oppstillinger, som for eksempel en Prolin midt i en beta-plate, eller andre urimelige oppstillinger av sekvens til struktur.

Gå til «Swiss Modell» øverst i menyen. Klikk «Submit Modelling request». Du blir da bedt om å lagre prosjektet på harddisken; gjør det og følg videre instruksene som kommer opp.

Om noe er galt med SAM-T99: Dersom noe går galt med SAM-T99, utfører du oppstillingen manuelt, ved å flytte på sekvensene hhv. til hverandres størst likhet, ved å markere et område i alignment vinduet med shift-tasten hold inne, og flytte sekvensene til venstre eller høyre med piltastene. En manuell variant av dette prosjektet kan se slik ut (Fig 2).

Nå har vi kommet dit i prosessen hvor det genereres en PDB-fil for sekvensen til Cape Cobra ved Swiss Model Serveren (www.expasy.org/swissmod.html). Du vil motta den i emailen så fort modelleringen er ferdig (5-10 minutter).

5. OPTIMALISERING AV MODELLEN

For at kriteriene under punkt 4 skal være mest mulig oppfylt, utfører vi en lett reorganisering av atomenes posisjoner i henhold til hverandre. Dette gjøres ved «Energy Minimization», som ut fra en liste av parametere (force field) for atomers interaksjon med hverandre, optimaliserer posisjonene til atomene slik at deres energi er lavest mulig og at de er i «tilfredsstillende» fysisk-kjemiske omgivelser.

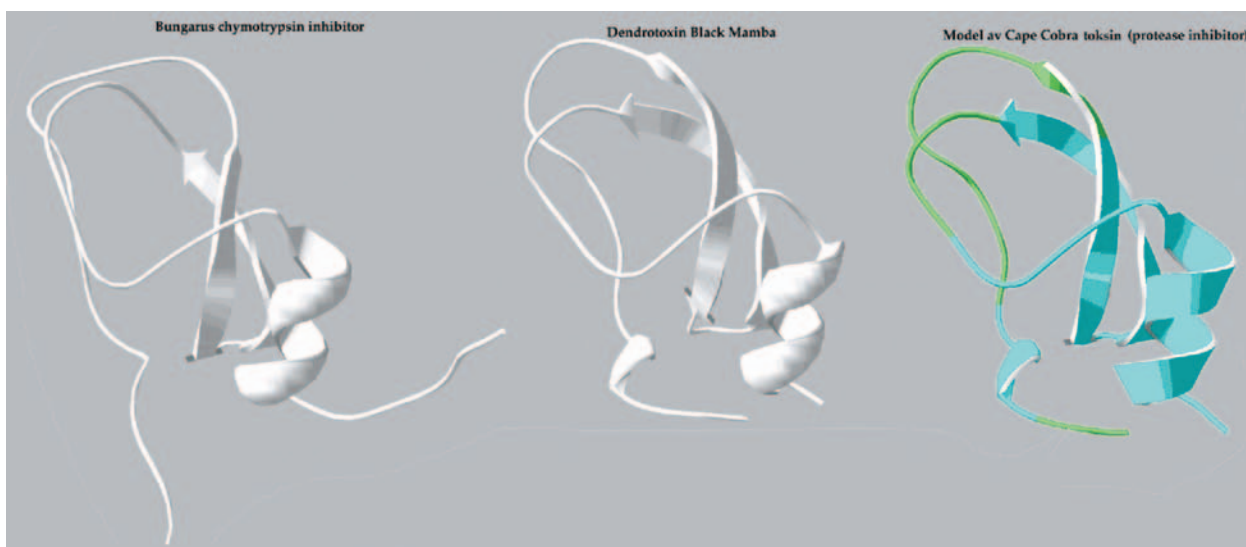
Gå til «Preferences», og så til «Energy Minimization».

Øverst i vinduet som dukker opp setter du inn 100 for «Steepest Descent», 100 for «Conjugate Gradients» og 200 «for Steepest Descent». Trykk «ctrl+n».

En negativ verdi tyder på at atomene er organisert på en favorabel måte, med hensyn til hydrogenbindinger, van der Waal krefter, ladninger og bindinger. En positiv verdi tyder på at noe frastøtes eller at bindinger er unaturlige. Se resultat siden som dukker opp. En verdi på ca. -3500kJ/mol for et slikt protein er god.

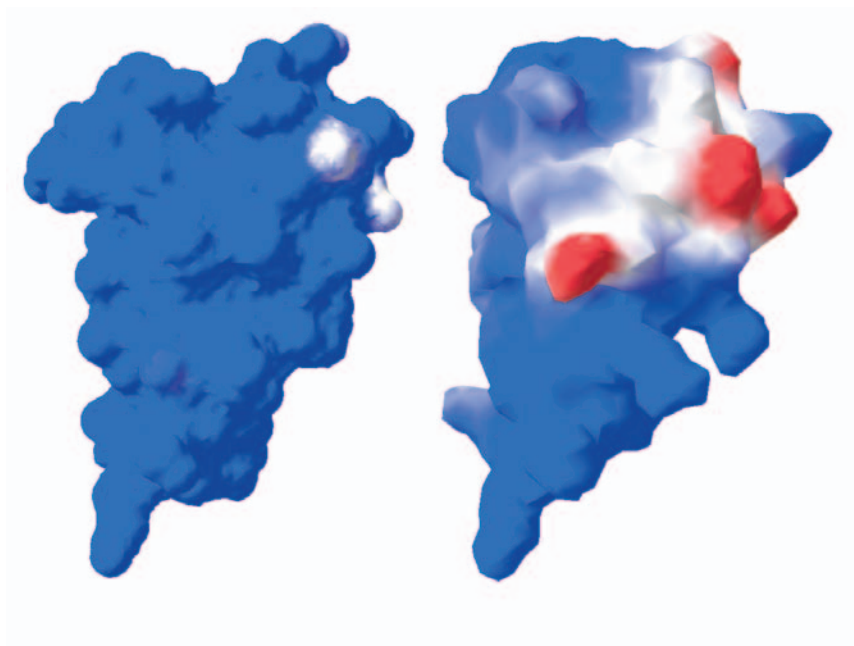
Dette vil ut fra Steepest Descent utføre en «snill» reorganisering av atomenes posisjoner, og så med Conjugate Gradients (CG) bevege seg litt grovere ut av posisjonene for å falle inn i ett energi-minimum (som kan enklere finnes ved

Figur 3: Proteinene har samme fold.



Denrotoksin (Black Mamba)

Cape Cobra toksin modell



Figur 4: Elektrostatisk potensial.

CG, dersom det ikke er altfor langt unna). Når CG har arbeidet i noen sykluser vil Steepest Descent finpusse på det optimale minimum (strukturell organisering av best energisk kompatibilitet) i 200 sykluser.

6. ANALYSE AV MODELLEN

En bilde-analyse (Fig. 3) viser tydelig at folden til proteinene er det samme. Som de fleste protein familier, så vil en likhet fra sekvenslikhet fra 20% og oppover kunne resultere i at tertiær strukturen forblir konserverert. For trypsin protease familiene, hvor strukturellikheten er bevart fra en sekvens-likhet på 20%, er årsaken «hovedankrene» i proteinet, som proliner, glyciner og cysteiner som spiller stor rolle for strukturell konservering. For mer informasjon om disse residiene i BPTI-familien, se relaterte publikasjoner [1, 2].

For mer bildebasert differensiering mellom strukturene kan man generere en elektrostatisk profil. Denne indikerer elektrostatisk karakter på overflate til proteinet. Den elektrostatiske profilen kan genereres ved å gå til «Tools», «Compute Molecular Surface» og «Compute Electrostatic Potential». Så går du til «File Discard» og «Discard Electrostatic Map». Da kommer det opp en slik illustrasjon (Fig. 4), hvor den molekylære overflaten er farget etter positiv ladning (blå), negativ ladning (rød) og

nøytrale områder (hvit).

For de interesserte vil det være viktig å utføre en analyse av sidekjeder til funksjonelle residuer, som har den inhiberende funksjonen. Slik informasjon vil finnes i relevante publikasjoner [2]. Ved å overlappere strukturene og så «highlight» sidekjedene ved Kontrollpanelet (Window, Control Panel) kan man utforske strukturene langt dypere og analysere hvilke typer sidekjeder som spiller rolle i inhiberingsmekanismen.

REFERANSER

1. Hokama, Y., Iwanaga, S., Tatsuki, T. and Suzuki, T. (1976). Snake venom proteinase inhibitors. III. Isolation of five polypeptide inhibitors from the venoms of *Hemachatus haemachatus* (Ringhal's cobra) and *Naja nivea* (Cape cobra) and the complete amino acid sequences of two of them. *J. Biochem.* **79**, p. 559-578.
2. Hanson WM, Beeser SA, Oas TG, Goldenberg DP. (2003). Identification of a residue critical for maintaining the functional conformation of BPTI. *J Mol Biol.* **333**, p. 425-41.



SERGIO MANTZETTI
Molecular Biology
Product Section
E. Pedersen & Sønn