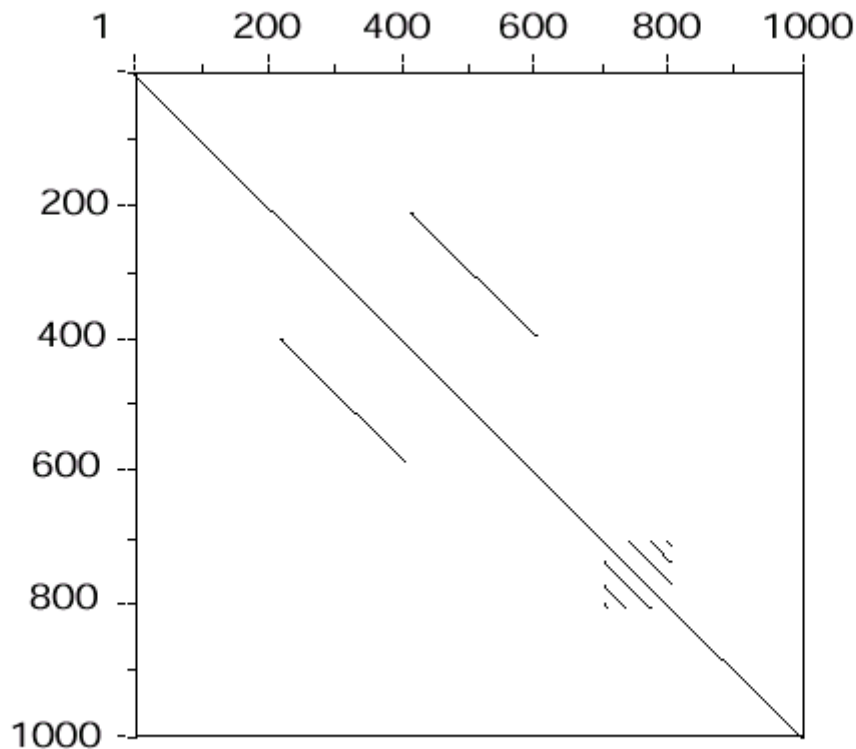


Oppgave 1

Figuren viser en idealisert prikkmatrise (dot matrix) laget ved å sammenligne et 1000 aminosyrer langt protein mot seg selv. Det er satt en prikk i matrisen dersom 5 påfølgende aminosyrer er helt identiske. Gi en tolkning av figuren og skisser hvordan dette proteinet er bygget opp, basert på dette resultatet.



Oppgave 2

ClustalW har blitt kalt et "unabashedly *ad hoc*" program for konstruksjon av flersekvenssammenstillinger. Delvis skyldes nok denne karakteristikken at ClustalW benytter en heuristisk søkemetode.

- Hva menes med en heuristisk søkemetode?
- Hvorfor kan det være fordelaktig å benytte heuristiske metoder for å lage flersekvenssammenstillinger?
- En annen grunn til karakteristikken av ClustalW over kan være at programmet prøver å ta biologiske hensyn under sammenstillingsprosessen, blant annet når det gjelder innføring av gap i sekvensene. Gi et eller flere eksempler på hvordan ClustalW innfører variabel gap-straff (gap penalty) under konstruksjon av sammenstillinger.

Oppgave 3

- Finne den eller de optimale lokale sammenstillingene av sekvensens ACDED og DCEF. Benytt poeng for sammenstilling av to like aminosyrer lik 1, to ulike aminosyrer -0.5 og poeng for gap -0.8 .

- e) Poenggivingen over er svært enkel, men neppe særlig biologisk ”korrekt”. Margareth Dayhoff og medarbeidere lagde i 60- og 70-årene mer anvendbare poengtabeller. Forklar kort hvordan disse tabellene ble fremstilt og hvordan de skiller seg fra poengsystemet over?

Oppgave 4

BLAST finnes i flere utgaver:

blastp
blastn
blastx
tblastn
tblastx
psi-blast

Fortell hva slags type søk de forskjellige variantene utfører og gi eksempler på problemstillinger for de forskjellige variantene vil være spesielt godt egnet.

Oppgave 5

Hva viser et Ramachandran plott?

Hva gjenspeiler B-faktor når man fremviser proteinstruktur?

Hvordan kan man analysere likhet mellom to fastlagte proteinstrukturer?

Oppgave 6

Du har klonet et nytt gen, og finner at den ene halvdel har 70 % sekvenslikhet med et protein med kjent struktur, mens resten ikke ligner på noe som er tidligere beskrevet. Hvordan kan man gå frem for å finne ut hvilken struktur proteinet kan ha?

Beskriv de forskjellige aktuelle metodene og de viktigste usikkerhetsfaktorer.

Oppgave 7

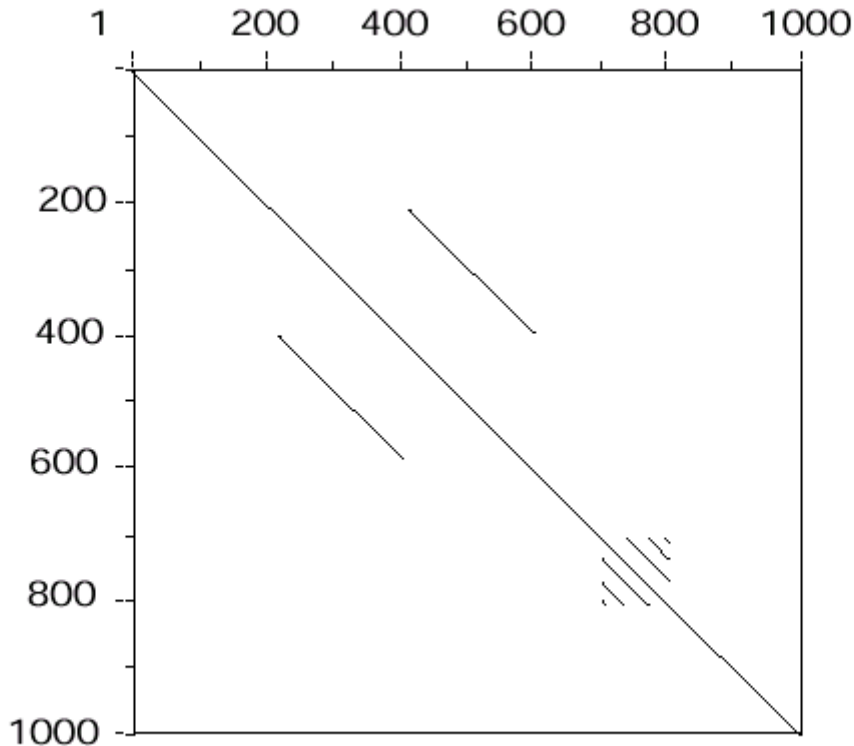
Hva er et proteinmotiv? Beskriv hvordan motiver i proteiner kan representeres og hvordan de kan benyttes for å utlede mulig struktur og/eller funksjon av et protein.

Oppgave 8

Hva er homologi? Beskriv hovedtrinn og fremgangsmåter for synliggjøring av slektskapet mellom homologer.

Exercise 1

The figure shows an idealized dot matrix obtained by comparing a protein with 1000 amino acids with itself. A dot is introduced in the matrix if 5 consecutive amino acids are identical. Interpret the drawing and indicate how this protein is built up, based on this result.



Exercise 2

- ClustalW has been characterized as an “unabashedly *ad hoc*” program for constructing multiple sequence alignments. This characterization may partly be due to the heuristic approach that the program uses.
- What is meant by a heuristic algorithm?
- Why may it be an advantage to use heuristic approaches in the construction of multiple sequence alignments?
- Another reason for the characterization above may be that the program tries to take biological considerations during the assembly process, for instance concerning introduction of gaps into the sequences. Give one or more examples on how ClustalW introduces a variable gap penalty during alignment construction.

Exercise 3

- Find the optimal **local** alignment(s) of the sequences ACDED and DCEF. Use a score of 1 for alignment of two identical residues, a score of -0.5 for alignment of two different residues and a score of -0.8 for aligning a residue with a gap.

- b) The scoring system used above is very simple, but hardly very relevant biologically. In the 60s and 70s Margareth Dayhoff and co-workers developed more interesting scoring matrices. Explain briefly how these tables were constructed and in what ways they differ from the scoring system above.

Exercise 4

BLAST is available in several variants:

blastp
blastn
blastx
tblastn
tblastx
psi-blast

Explain what kind of searches these BLAST versions perform and give examples of problems for which the various versions are especially well suited.

Exercise 5

What does a Ramachandran plot show?

What does the B factor reflect when protein structure is shown?

How can similarity between two determined protein structures be analysed?

Exercise 6

You have cloned a new gene and find that one half has 70 % sequence identity with a protein with known structure while the rest shows no similarity with any described sequences. How can you proceed to infer possible structural elements of the protein?

Exercise 7

What is a protein motif? Describe how protein motifs may be represented and how they can be used to infer the potential structure and/or function of a protein..

Exercise 8

What is homology? Describe the key steps and methods in visualising the relationship between homologues.