

TØ 6 – Nukleinsyrestruktur 1

Table of contents

1	Opgave 1. Kærlighed ved første blik	2
1.1	Forklar specificitet i major groove	2
1.2	Identificer aminosyrer til sekvensspecifik genkendelse	2
1.3	Undersøg groove-mønster for 1BNA	2
2	Opgave 2. Der skal to til tango	4
2.1	Giv eksempel på et basepar-palindrom	4
2.2	Beskriv DNA-bindingsdomæner	4
2.3	Diskuter fordele ved heterodimer-binding	4
2.4	Beskriv ligand-binding regulering af DNA	4
3	Opgave 3. TATA-box	5
3.1	Find TBP-bindingsgroove	5
3.2	Beskriv groove-mønstret ved TBP-bindingssted	5
3.3	Beskriv TBPs sekundære DNA-bindingsstruktur	5
3.4	Undersøg phenylalanin-interkalering	5
4	Opgave 4. Mad-Max	6
4.1	Beskriv dimeriseringsdomænet	6
4.2	Find bindingsgroove for de tre komplekser	6
4.3	Beskriv mønster af H-bindings-donor-acceptorer	6
4.4	Beskriv hydrogenbindinger for aminosyrerne	6
5	Opgave 5. At TALE med DNA	7
5.1	Find TALE-proteinets bindingsgroove	7
5.2	Beskriv TALE-repeat genkendelse af nukleotider	7
5.3	Forslå protein-funktionaliteter for TALE	7
6	Opgave 6. Trp trap træsko	8
6.1	Bestem Trp-repressors bindingsgroove	8
6.2	Undersøg Trp-bindings effekt på repressor	8
6.3	Beskriv groove-mønstret for Trp-repressor	8
6.4	Identificer Arg69's base-interaktion	8
7	Opgave 7. Arc repressor	9
7.1	Lav script til Arc-repressor-analyse	9
7.2	Beskriv DNA-bøjning ved Arc-binding	10
7.3	Beskriv sekundær struktur i DNA-binding	10
7.4	Find DNA-helixparameter for afstanden	10
7.5	Identificer H-bindingspartnere i active site	10

1 Opgave 1. Kærlighed ved første blik

I denne opgave skal I lære hvordan proteiner genkender DNA (eller omvendt) for at kunne indgå i en specifik (“romantisk”) binding til hinanden. Hemmeligheden bag denne genkendelse kan findes i Liljas “Textbook of Structural Biology” afsnit 10.2, hvor den specifikke interaktion mellem aminosyrer og DNA-baser beskrives. Aminosyrerne kan genkende både major og minor groove via den såkaldte “groove code”, der er vist i Liljas figur 10.18. Som det ses i figuren herunder er der eksponerede H-bindings-donorer og acceptorer i både major og minor groove og det er disse, som aminosyrerne kan genkende og binde til:

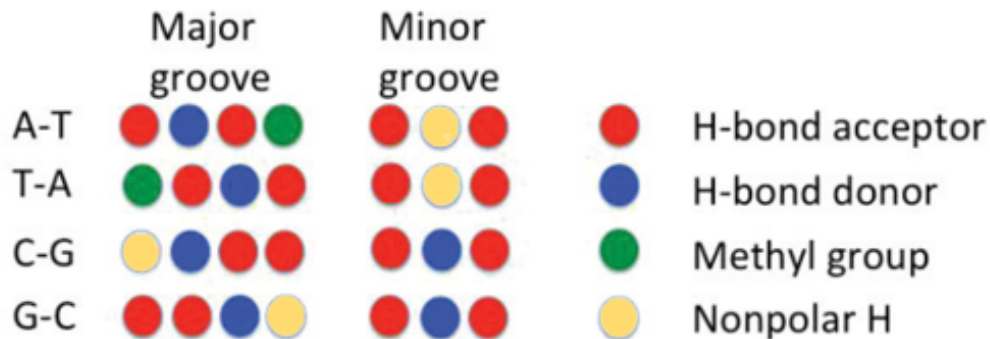


Figure 1: Figur 1

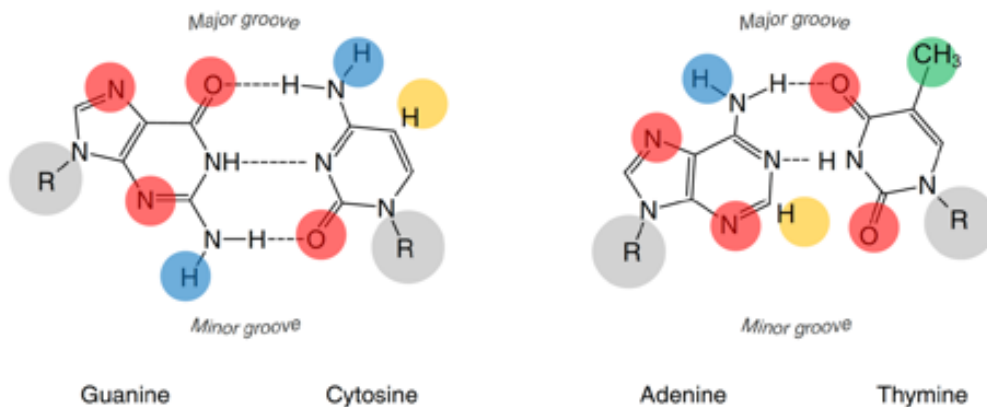


Figure 2: Figur 2

1.1 Forklar specificitet i major groove

Forklar hvorfor der er mere specificitet i binding til major groove end til minor groove.

1.2 Identificer aminosyrer til sekvensspecifik genkendelse

Hvilke aminosyrer vil du tro ofte bruges til sekvensspecifik genkendelse og hvorfor?

1.3 Undersøg groove-mønster for 1BNA

Groove-koden kan være lidt svært at få øje på når man analyserer en DNA dobbelthelix i PyMOL, men med den rette repræsentation bliver den mere tydelig: **DNA-groove.pml** scriptet, som findes under denne uge på Brightspace, viser DNA-dobelthelixa som en kuglemodel (spheres), hvor H-bindingsacceptorer er farvet røde, H-bindingsdonorer blå, metylgrupper grønne og ikke-polære hydrogenatomer gule.

PyMOL info: Sphere-repræsentationen viser Van der Waals-radius af atomer og giver dermed et billede af molekylets realistiske form og fylde. Farvning af atomerne ses tydeligere pga. deres størrelse. Til gengæld er det ikke let at aflæse den kemisk sammensætning (som man ser det med sticks) eller den sekundære struktur (som man ser det med cartoon).

Åben DNA dobbelthelix med PDB-ID: 1BNA i PyMOL og fjern vand-molekyler med PyMOL kommandoen "remove solvent". Brug nu **DNA-groove.pml** script til at ændre på repræsentationen og farvelægge donorer- og acceptorer i grooves. Rotér molekylet, så du ser ned i major groove (så symmetrien bliver tydelig).

Hvad er specielt ved groove-kode-mønstret for denne DNA dobbelthelix?

2 Opgave 2. Der skal to til tango

Det er vigtigt for DNA-bindende proteiner at kunne gribe fat om DNA-dobbelthelixen og interagere med specifikke base-sekvenser. DNA-molekylet bindes ofte af homodimeriske DNA-bindende proteiner i områder med såkaldte "basepar-palindromer".

2.1 Giv eksempel på et basepar-palindrom

Giv et eksempel på et "basepar-palindrom" og forklar hvilke geometriske egenskaber det har og hvorfor homodimeriske proteiner kan lide at binde der.

2.2 Beskriv DNA-bindingsdomæner

Som det ses i Tabel 10.1 i Liljas "*Textbook of Structural Biology*" findes der flere forskellige familier/typer af DNA-bindende proteiner, der via forskellige DNA-bindingsdomæner interagerer med DNA dobbelthelixen.

Beskriv hovedtyperne af DNA-bindingsdomæner og hvordan de interagerer med DNA.

2.3 Diskuter fordele ved heterodimer-binding

DNA-bindende proteiner har en vigtig rolle som gen-specifikke transskriptionsfaktorer, der kan aktivere eller hæmme mRNA syntese. I eukaryoter findes der flere heterodimeriske transkriptionsfaktorer.

Hvad er de mulige fordele ved DNA-binding af en heterodimer?

2.4 Beskriv ligand-binding regulering af DNA

Binding af en ligand til et DNA-bindende protein kan regulere DNA-bindingen.

Hvad er mekanismen hvormed ligand-binding kan påvirke proteinets evne til at binde DNA? Nævn et par eksempler på regulering med denne mekanisme.

3 Opgave 3. TATA-box

Det TATA-box bindende protein (TBP) er involveret i promoter-genkendelse mht. at starte RNA polymerase transskription. TATA-boksen er en 8-basepar del af promoteren, der normalt er lokaliseret 30 nukleotider før transskription start site (TSS). Konsensus-sekvensen er TATA(A/T)A(A/T)(G/A).

Åben `TATA.pml` scriptet i PyMOL. **F1** viser bindingen af TBP langs den heliske akse for DNA-dobbelthelix.

3.1 Find TBP-bindingsgroove

I hvilken groove af DNA-dobbelthelix binder TBP?

PyMOL Hint

Groove kan identificeres ved at zoome ind på basepar og identificere Högsteen- og sukker-side. Vis f.eks. et basepar med kommandoen `show sticks, resi 7+23` og zoom ind for at undersøge dette med kommandoerne `zoom resi 7+23` og juster dybden med musehjulet, så der fokuseres på baseparret. Rotér derefter så du ser baseparret fra siden. Tryk **F2** for at vise den ønskede visning.*

3.2 Beskriv groove-mønsteret ved TBP-bindingssted

Tryk **F1** og brug groove-kode scriptet (`DNA-groove.pml`) til at vise DNA-dobbelthelixa som kuglemodel, hvor H-bindingsacceptorer er farvet røde, H-bindingsdonorer blå, metylgrupper grønne og ikke-polære hydrogenatomer gule. Skjul TBP for bedre at kunne se groove-mønster.

Hvad karakteriserer mønsteret i groove, hvor TBP binder?

3.3 Beskriv TBPs sekundære DNA-bindingsstruktur

Hvad karakteriserer TBPs sekundære struktur, der binder til DNA groove?

3.4 Undersøg phenylalanin-interkalering

Tryk **F3/F4** for at se et vigtigt element i DNA-protein bindingen, hvor to phenylalaniner interkalerer mellem to basepar.

Hvordan er de to basepar påvirket af interaktionen med de to phenylalaniner? Hvilken standardparameter for den relative orientation af baser (Textbook of Structural Biology, figur 5.16) beskriver bedst konformationen af disse to basepar?

4 Opgave 4. Mad-Max

Max-Myc-Mad er et heterodimerisk transkriptionsfaktor system, der binder specifikt til DNA segmentet CACGTG som tre forskellige komplekser: Myc-Max, Max-Max og Mad-Max. Ifølge forfatterne til den oprindelige artikel om Max-Myc-Mad systemet har navngivningen af Mad (det andet protein i Mad-Max heterodimer) ikke noget at gøre med filmen af samme navn (**Youtube link**), men kommer fra forkortelsen af “Max dimerization”. Hent PyMOL-scriptet `Mad-Max.pml` og kørs det i PyMOL. Tryk **F1** for at se Myc-Max komplekset, tryk **F2** for at se Max-Max komplekset, og **F3** for at se Mad-Max komplekset.

4.1 Beskriv dimeriseringsdomænet

Beskriv sammensætningen af dimeriseringsdomænet og hvilken egenskab α -helixerne, der binder DNA, har.

4.2 Find bindingsgroove for de tre komplekser

I hvilken groove binder Myc-Max, Max-Max og Mad-Max?

4.3 Beskriv mønster af H-bindings-donor-acceptorer

Orientér molekylet i PyMOL, så du ser ned langs pseudosymmetriaksen af Myc-Max og gem derefter Myc-Max, så kun DNA-dobbelhelixen er synlig. Brug groove-kode scriptet (`DNA-groove.pml`) til at vise helixgroove-mønsteret.

Hvad er karakteristisk ved mønsteret af H-bindings-donor og -acceptorer?

4.4 Beskriv hydrogenbindinger for aminosyrerne

Tryk **F4** for at se hvordan genkendelses-helixer i Myc-Max interagerer med DNA segmentet.

Tryk **F5** for at se hvordan genkendelses-helixer i Max-Max interagerer med DNA segmentet.

Tryk **F6** for at se hvordan genkendelses-helixer i Mad-Max interagerer med DNA segmentet.

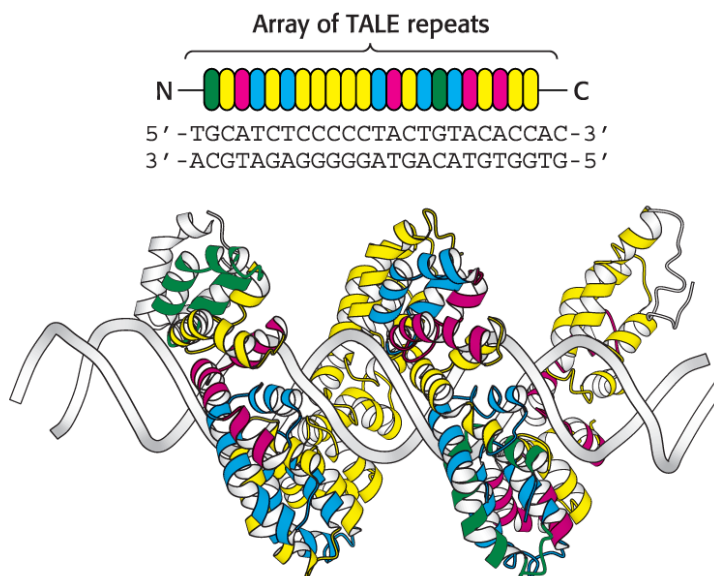
Beskriv hydrogen-bindingerne for de tre aminosyrer (His, Glu, Arg) i genkendelses-helixer i Myc-Max ved at specificere hvilke baser og hvilke atomer på baserne, der indgår i bindingen. Er der andre interaktioner end med baserne? Er interaktionerne symmetriske i Myc-Max, Max-Max og Mad-Max og hvorfor?

PyMOL Hint

Brug menuen “Mouse > Selection mode > Atoms”, så du bedre kan identificere de individuelle atomer, der indgår i hydrogen-bindingerne.*

5 Opgave 5. At TALE med DNA

Transkriptionsfaktor-lignende effector nukleaser (TALENs) binder specifikke DNA-sekvenser vha. en række DNA-bindende domæner kaldet TALE-repeats. TALEN-proteiner kan designes til at genkende og kløve specifikke DNA-sekvenser og kan dermed anvendes til genom-redigering (genome editing). TALE-proteiner er omtalt i Berg Biochemistry, udgave 10, side 292, og opbygningen kan ses i figur 9.37 (se nedenstående figur).



Berg et al., *Biochemistry*, 9e, © 2019 W. H. Freeman and Company

Figure 1: BER_11467_05_F38

5.1 Find TALE-proteinets bindingsgroove

Download og åben scriptet [TALE.pml](#) i PyMOL. Tryk **F1** og **F2** for at se DNA-protein komplekset hhv. fra siden og langs helix akse. Tryk **F3** for at se sphere-repræsentation, hvor man lettere kan se repeat-strukturen. Tryk **F4** for at se en specifik interaktion mellem protein og DNA.

I hvilken groove af DNA-dobbelthelixen binder TALE-proteinet? Forklar hvordan du observerer dette.

5.2 Beskriv TALE-repeat genkendelse af nukleotider

Hvert TALE-repeat indeholder 34 aminosyrer og to alfa-helicer, men kun to aminosyre-rester er ansvarlige for unik genkendelse af et enkelt nukleotid i DNA-dobbelthelixen.

Tryk **F4-F7** for at se hvordan TALE-repeats genkender specifikke nukleotider. Beskriv hvilke aminosyrer, der interagerer med hvilke baser og hvilken type interaktion der er tale om. Hvilken kant af baserne er involveret i interaktionen og hvilken af DNA-strengene genkendes?

5.3 Forslå protein-funktionaliteter for TALE

Man kan designe TALE-proteiner, så de binder til unikke DNA-sekvenser, ved at organisere rækkefølgen af TALE-repeats ihht. DNA-sekvensen.

Forslå hvilke protein-funktionaliteter man kan sætte i enden af et TALE-proteindomæne og beskriv hvilke biologiske effekter det kan have og/eller hvilke teknologiske muligheder det kan give.

6 Opgave 6. Trp trap træsko

I *E. coli* binder Trp-repressor en specifik DNA-sekvens under tilstedeværelse af aminosyren tryptofan (Trp) og hæmmer dermed udtryk af fem gener i *trp*-operon. Trp-repressor danner en homodimer, der binder til DNA-dobbelthelix via to alpha-helicer (kaldet "genkendelseshelicer"). Åben PyMOL-scriptet `Trp.pml` i PyMOL, der laver en strukturel alignment af Trp-repressor (bundet til Trp og DNA) med *apo*-formen af proteinet (ikke bundet til Trp eller DNA). **F1** viser denne alignment.

6.1 Bestem Trp-repressors bindingsgroove

I hvilken groove af DNA binder Trp-repressor med de to genkendelseshelicer? Mål afstanden mellem de to genkendelseshelicer i 1TRO. Hvilken DNA-helix-parameter svarer denne afstand til og hvorfor observeres denne afstand typisk for transkriptionsfaktorer?

6.2 Undersøg Trp-bindings effekt på repressor

Hvilken del af proteinet påvirkes mest af binding til Tryptofan? Mål afstanden mellem de to genkendelseshelicer for *apo*-repressoren og beskriv hvilken effekt ændringen må forventes at have på binding af DNA.

6.3 Beskriv groove-mønstret for Trp-repressor

Brug groove-kode scriptet (`DNA-groove.pml`) til at vise DNA-dobbelthelixen som kuglemodel, hvor H-bindingsacceptorer er farvet røde, H-bindingsdonorer blå, metylgrupper grønne og ikke-polære hydrogenatomer gule. Skjul Trp repressor, så kun DNA dobbelthelix ses.

Hvad er karakteristisk ved mønsteret af H-bindingsdonorer og -acceptorer og hvad fortæller det om bindingen af Trp repressor?

6.4 Identificer Arg69's base-interaktion

Tryk **F2** for kun at se polære kontakter mellem genkendelseshelicer og DNA-dobbelthelix. Hvilken kant af basen interagerer Arginin-69 med? Hvilke typer af interaktioner ser man typisk bidrage til bindingsspecificitet mellem DNA og protein?

7 Opgave 7. Arc repressor

PyMOL-scripting opgave: I denne opgave skal I analysere DNA-protein binding i PyMOL. I vil lære hvordan man farver proteiner efter sekundær struktur og hvordan man finder polære bindinger mellem to selektioner.

Arc-repressor (PDB-ID: 1BDT) er en transskriptionsfaktor fra bakteriofag P22, der hæmmer dets eget gen *arc*. Arc-repressor binder kooperativt som to dimere til et 21-basepar operator site. DNA operator-sekvensen er vist herunder, hvor TAGA-sekvensen er farvet grå. Det brudte palindroms pseudosymmetri-akse går igennem det centrale G-C basepar.

- 5' - ATAGTAGAGTGCTTCTATCAT - 3'
- 3' - TATCATCTCACGAAGATAGTA - 5'

7.1 Lav script til Arc-repressor-analyse

Lav et script, der gør følgende:

- a. Henter Arc-repressoren og viser DNA-protein komplekset med pseudosymmetri-akse langs PyMOLs y-akse. Gem scenen med navnet **F1** (scene F1, store).
- b. Lav en scene kaldet **F2**, der farver sekundær struktur (se nedenstående PyMOL info).
- c. Lav en scene kaldet **F3**, der viser afstanden mellem de to bindingselementer. Hint: Brug distance-funktionen mellem to atom selektioner på aminosyrer i bindingselementerne og brug mode=0, se nedenstående PyMOL info.
- d. Lav en scene kaldet **F4**, der viser DNAs interaktion med proteinet. Vis genkendelses-element (rest 8-14) og DNA (chain E+F) som sticks med element farver. Find nu hydrogenbindinger mellem genkendelses-element (rest 8-14) og DNA. Hint: Brug distance-funktionen mellem genkendelses-element og DNA med mode=2, se nedenstående PyMOL info.

i PyMOL Info: Sæt farve på sekundærstruktur

Når man farver proteiner efter sekundær struktur er der ikke en samlet kommando, der gør det. Man er nødsaget til at farve hver type, altså alphahelicer, betastrands og loops, enkeltvis. Loops indeholder alle sekundære strukturer, der ikke er betastrands eller alphahelicer. I har tidligere lært at property selektoren for sekundær struktur er `ss`. PyMOLs navn for alphahelicer, betastrands og loops er hhv. `h`, `s` og `l+`". Man ville altså farve loops hvide med kommandoen: `color white, ss l+`".

i PyMOL Info: Distance funktionens mode

Distance funktionen kan tage et ekstra argument kaldet mode, som har flere indstillinger, men de vigtigste er `mode=0` og `mode=2`. Med mode kan man bestemme hvilken type interaktioner man vil medtage. `mode=0` er default og er bare afstanden mellem alle punkter i de to selektioner. `mode=2` viser udelukkende de polære bindinger mellem to selektioner. Læs mere her.

Brug nu dit script til at svare på de følgende spørgsmål.

7.2 Beskriv DNA-bøjning ved Arc-binding

Tryk **F1**. Bindingen af Arc-repressor påvirker DNA-dobbelhelixen og fører til en bøjning. Hvilken basestakparameter (Liljas figur 5.16) beskriver bedst bøjning?

7.3 Beskriv sekundær struktur i DNA-binding

Tryk **F2**. Beskriv bindingen af Arc-proteinet til DNA. Hvilken type protein-sekundær struktur binder til hvilken groove på DNA dobbelhelixen? Beskriv hvordan du observerer hvilken groove der bindes.

7.4 Find DNA-helixparameter for afstanden

Tryk **F3**. Hvilken DNA-helix parameter svarer denne afstand til?

7.5 Identificer H-bindingspartnere i active site

Tryk **F4**. Hvilke aminosyrer hydrogenbinder til baserne? Hvilke andre typer af interaktioner bidrager til bindingspecificitet?

