Looking into DNA Recognition: Zinc Finger Binding Specificity

Guillaume Paillard, Cyril Deremble and Richard Lavery

Presented By:
Jamie Duke

Background Information

- Zinc Fingers
  - Nucleic Acid binding domain
  - Classic C2H2 conformation – coordinating a zinc ion
  - Conserved Pattern: x-C-x(1-5)-C-x(12)-H-x(3-6)-H
  - Conserved aromatic ring
  - 24 residue $\beta - \beta - \alpha$ motif
  - Multiple domains used to recognize specific DNA sequences
  - Most commonly studied family is Zif268 with 3 zinc finger domains

Referenced from Pfam Acc. No: PF00096 (http://www.sanger.ac.uk/cgi-bin/Pfam/getacc?PF00096)
Zinc Finger Binding

- Each Finger recognizes 3 – 5 nucleotides
- Recognition occurs in the α-helix of the finger
- Recognition is overlapped by the 3 domains
- DNA binding site can be changed with mutation to the protein

![Diagram of Zinc Finger Binding](Paillard et al. Fig 1A and 1B)

Introduction

- “Here, we hope to delve deeper into the binding mechanisms by breaking recognition down into its components, looking at how individual amino acids contribute to specificity, whether indirect effects contribute and, if so, what is the relative importance of different structural components of DNA deformation.” (Paillard et al)
ADAPT

- Program developed to overcome expensive modeling studies
- Uses a protein – DNA complex with all atoms based off of crystallographic data
  - Incorporates complexation energy using molecular mechanics force field
- Compares all possible DNA sequences for N given nucleotides, and determines the best consensus sequence

ADAPT Overview

- Construction of a protein – DNA Complex
  - Unpaired DNA is removed
  - Incorporation of Lexides
    - Replace nucleotides with a multi-copy base that contains the four standard bases superimposed upon one another
    - Used to create unbound reference DNA structures and allow for easier energy calculations
  - Addition of nucleotide pairs to the end of the sequence
  - Energy Minimization → protein – DNA complex conformations
  - Energy calculation for all nucleotide sequences
  - Selection of optimal sequence
Validation of ADAPT with Zif268

- Consensus sequence recognizes G at positions 1, 3, 6, 7, and 9 and C at position 2
- 677 strong binding sequences were computed
- Computed consensus sequence shows overall agreement within the binding site
- Direct interactions account for 89% of the selectivity

Paillard et al. Fig 1C.
TATA – Zif

- Cα RMSD between TATA – Zif and Zif268 = 1Å
- Bound DNA has B-DNA conformation with enlarged major groove where TATA – Zif is bound
- Mutations occur at positions -1, 1, 2, 3, 4, 5 and 6 of each finger → different DNA selectivity
- More amino acids contact DNA and several residues contact more than one base
- Increased overlap between binding sites of successive fingers
- No experimental consensus available

Paillard et al. Fig 4.
Conclusions

- Certain residues can influence selectivity at more than one site in the target DNA sequence
- Direct interactions cannot support binding specificity alone
  - DNA deformation accounts for ~10% of recognition in WT Zif268 and ~30% of recognition in TATA-Zif
- Prediction of the binding sites for Zinc finger proteins must include all fingers
References
