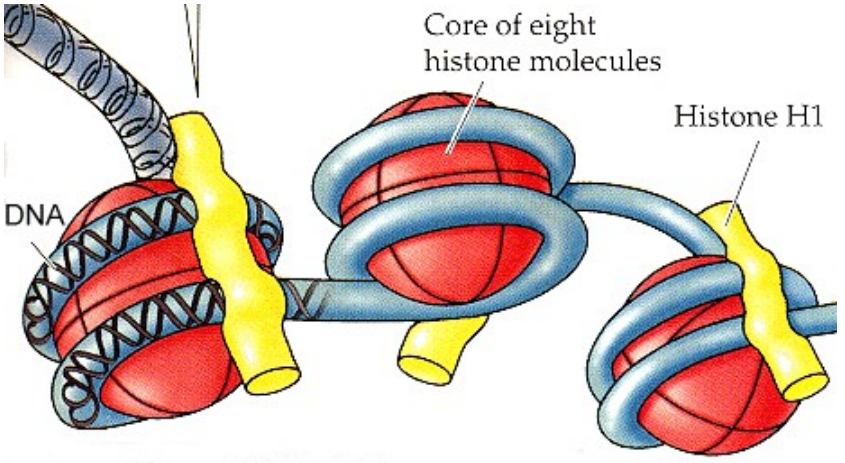
The mechanics behind DNA sequence-dependent properties of the nucleosome

Eugene Y. D. Chua, Dileep Vasudevan, Gabriela E. Davey, Bin Wu and Curt A. Davey
Nucleic Acids Research Advance Access published March 26, 2012

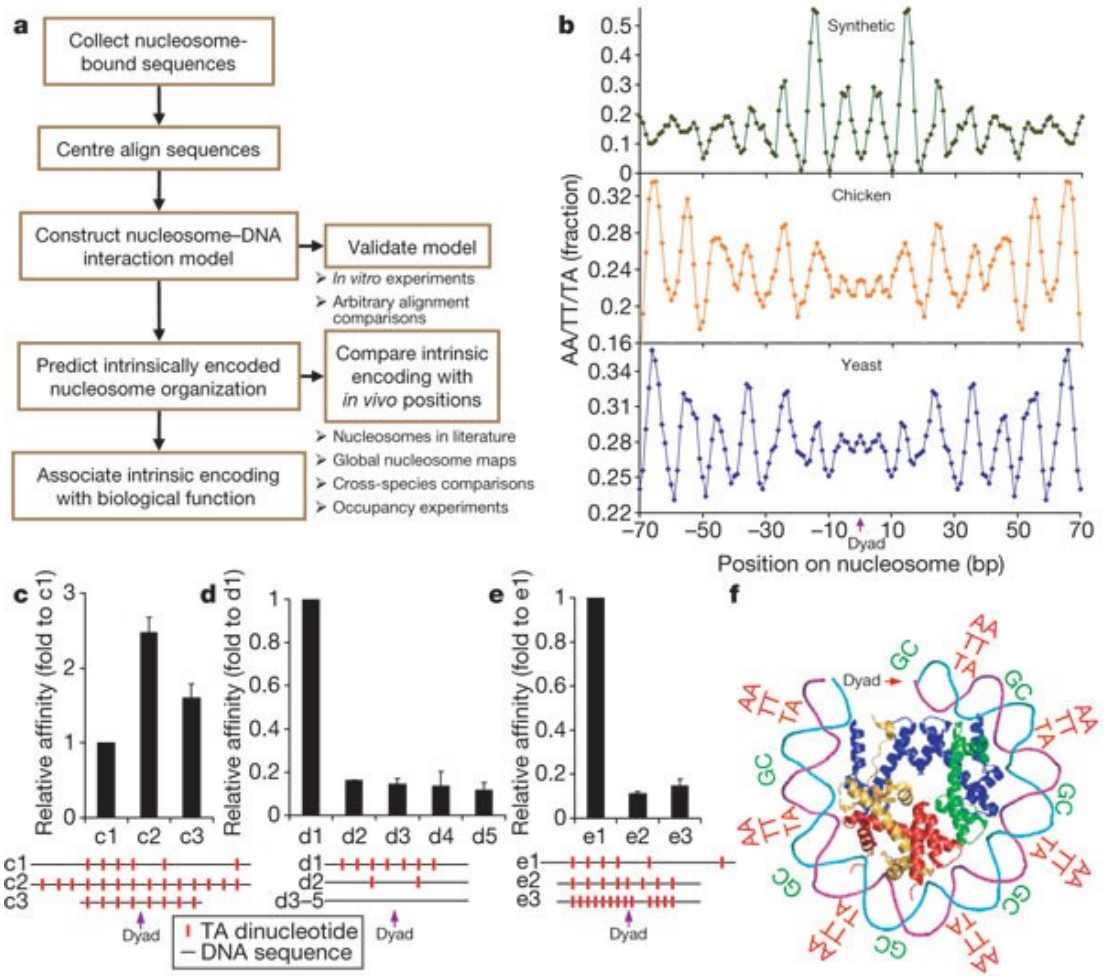
Lauris Kaplinski
Bioinformatics Journal Club
2012/04/16

Nucleosomes

- The DNA in single eucaryotic sell is about 2m long
- Packed to a structure only few uM in diameter
- The primary packing structure is histone octamer
- Histones have to bind with arbitrary double-stranded DNA sequences
- They still have preferences
- This may have effect in gene regulation



Nucleosome consensus



b, Fraction (3-bp moving average) of AA/TT/TA dinucleotides at each position of centre-aligned yeast, chicken or random chemically synthesized17 nucleosome-bound DNA sequences, showing 10-bp periodicity of these dinucleotides.

c–e, In vitro experiments. Positions of the key AA/TT/TA dinucleotides on the tested sequences are indicated

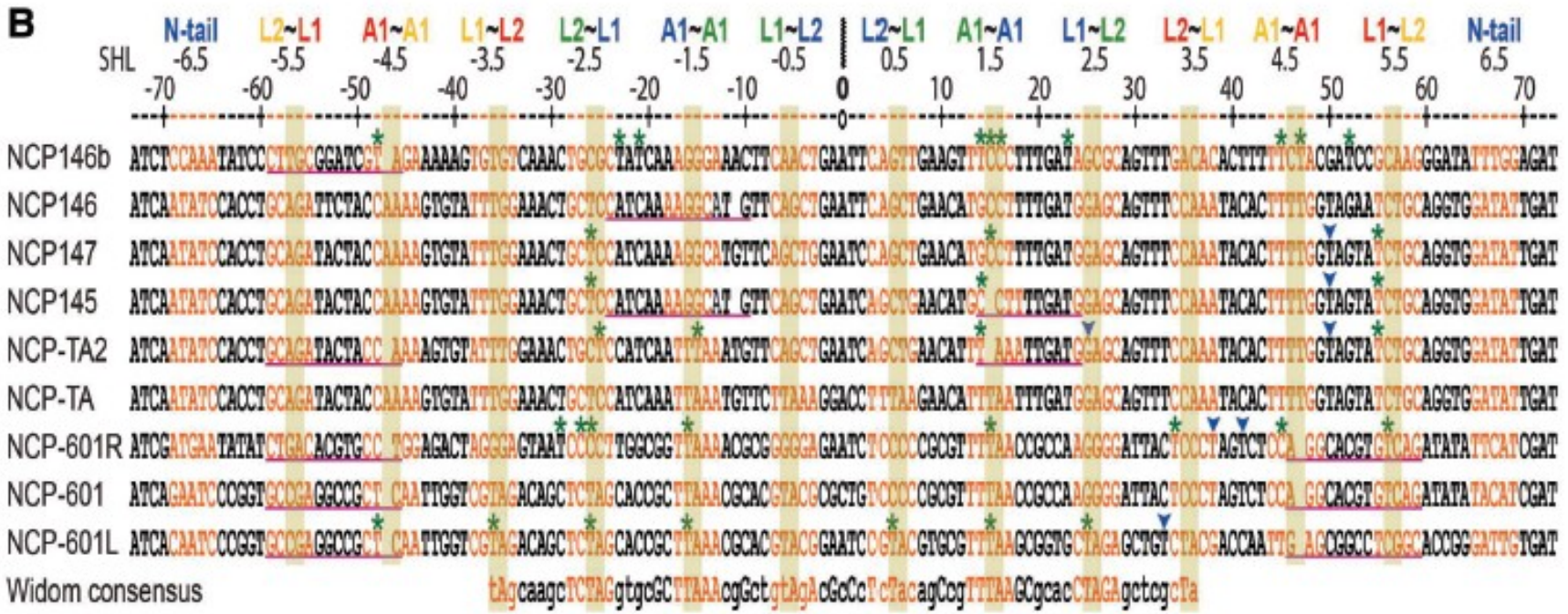
f, Key dinucleotides inferred from the alignments are shown relative to the three-dimensional structure of one-half of the symmetric nucleosome.

Segal et al. 2006

Segal et al 2006

- A genomic code for nucleosome positioning
  Nature 442, 772-778 (17 August 2006)
- 10-bp periodic AA/TT/TA dinucleotides and out of phase periodic GC dinucleotides
- AT dinucleotides are positioned in **bound minor groove**
- AT is the most flexible dinucleotide pair
- Predicted 73bp highest affinity nucleotide sequence binding to H3/H4 tetramer (Widom consensus)
- The methodology did not allow to predict optimal flanking regions
- Synthetic very strongly binding full-length sequence has recently found (Widom 601)

Known histone-bound sequences

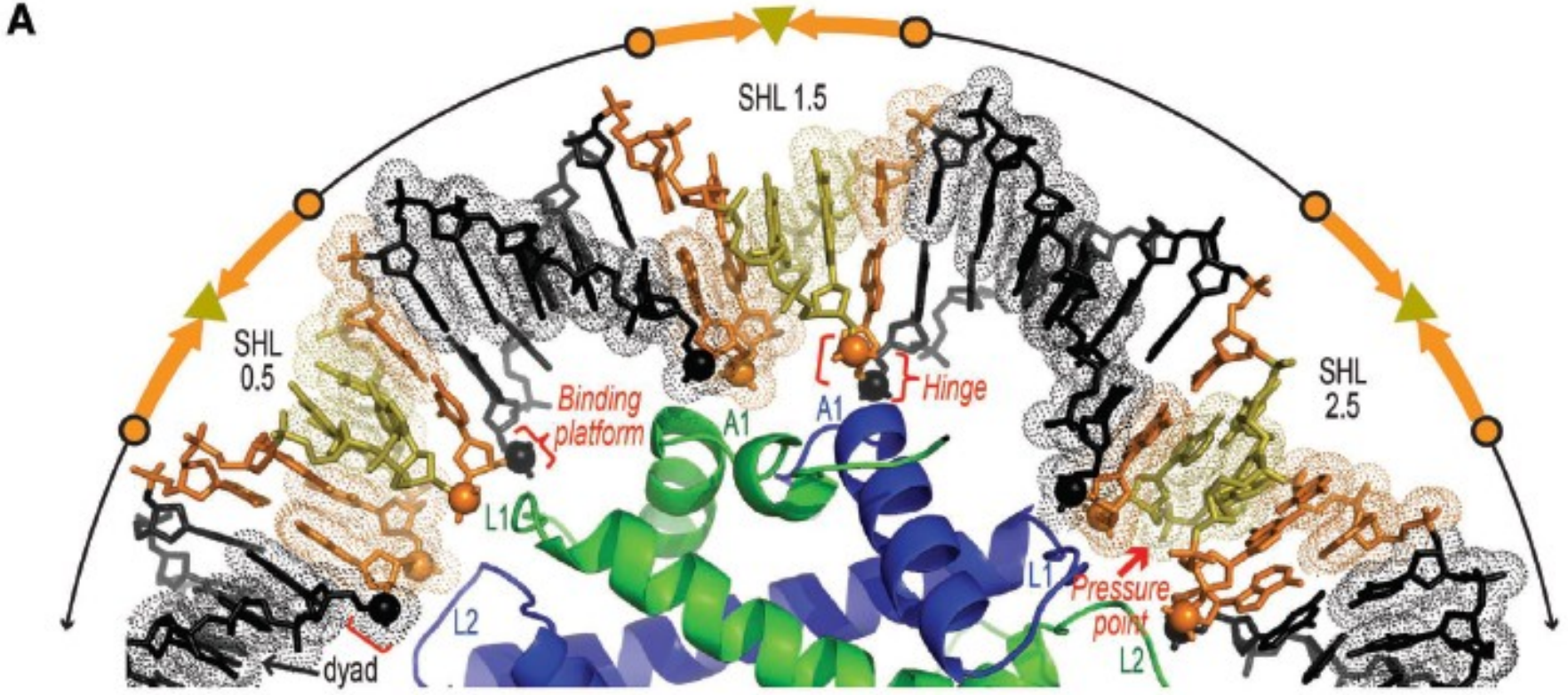


- NCP constructs are arranged in order of increasing salt stability. Severe kinks at locations of DNA stretching around SHL ±2 or ±5 (magenta underlines), associated with a single base pair shift in histone-nucleotide register, are depicted as gaps in the sequence. DNA-permanganate reactivity hotspots in the nucleosomal state from footprinting analysis (six constructs) are indicated with green asterisks. Sites where the nucleosomal DNA shows reduced permanganate reactivity relative to the naked state are indicated with blue arrowheads. Capitalized bases in the Widom consensus sequence represent the most highly conserved nucleotides (17). The histone–DNA register assignments for NCP-601R and the Widom consensus sequence, for which crystal structures are not available, were inferred from the structures of NCP-601 and NCP-601L.

Nucleosome structure

- Crystal structures of several alpha-satellite sequences on nucleosome have been determined
- In total 14 "attach points" with NCP
- Structure of strongest-binding Widom 601 is mostly known
- Precise atom-positioning cannot be determined because it is non-palindromic
- Synthesized separate left- and right-half palindromic sequences 601L and 601R
- Determined crystal structure with 2-3 Å resolution
- Determined salt stability of different structures
- The chemical stability along DNA chain was determined by oxidizing with $KMNO_4$

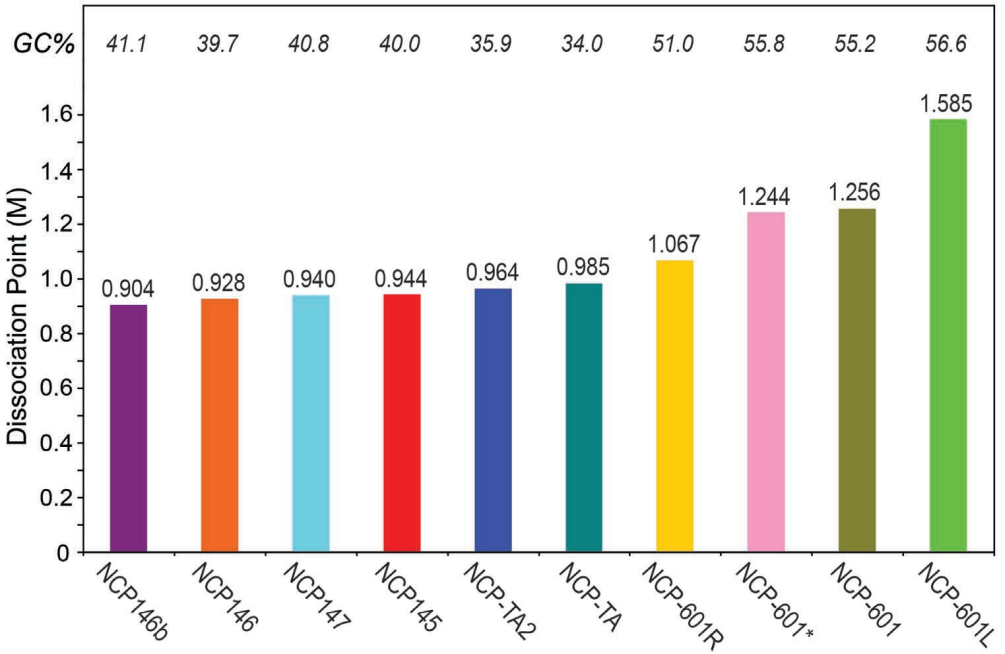
How is DNA arranged in nucleosome?



- Figure 1. Double helix association, conformation and positioning on the histone octamer. (A and B) Minor and major groove-inward-facing regions are orange and black, respectively, with ‘pressure points’ at minor groove-inward centres highlighted gold. Histone proteins are blue, H3, green, H4, yellow, H2A and red, H2B (DNA-binding motifs: L, loop, A, a-helix). (A) Section of the NCP-601L crystal structure with phosphorous atoms of the ‘binding platforms’ shown as spheres. Bound single-strand regions act as a ‘hinge’, allowing conformational variation between different DNA sequences.

Nucleosome structure

- Four phosphates are bound at "pressure points" and act as "hinges" around whose DNA strains can bend
- TTTAA motifs at positions +/- 1.5 (strongest nucleosome binding signals)
- NCP 601L had extraordinary salt stability, 601R much lower

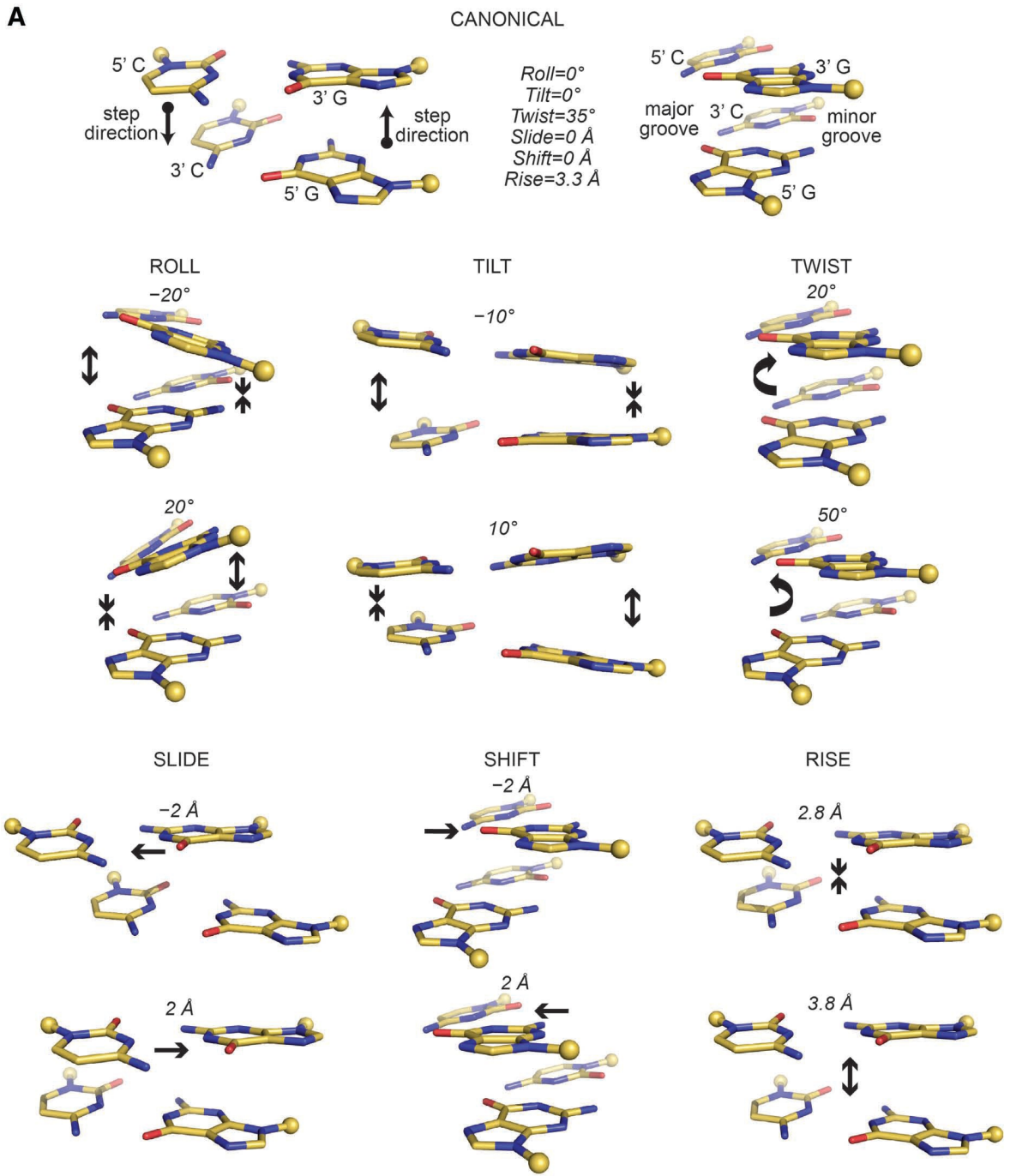


DNA sequence-dependent salt stability of the nucleosome core. Dissociation points correspond to the midpoints of NaCl-induced DNA–histone dissociation measured by tyrosine fluorescence spectroscopy

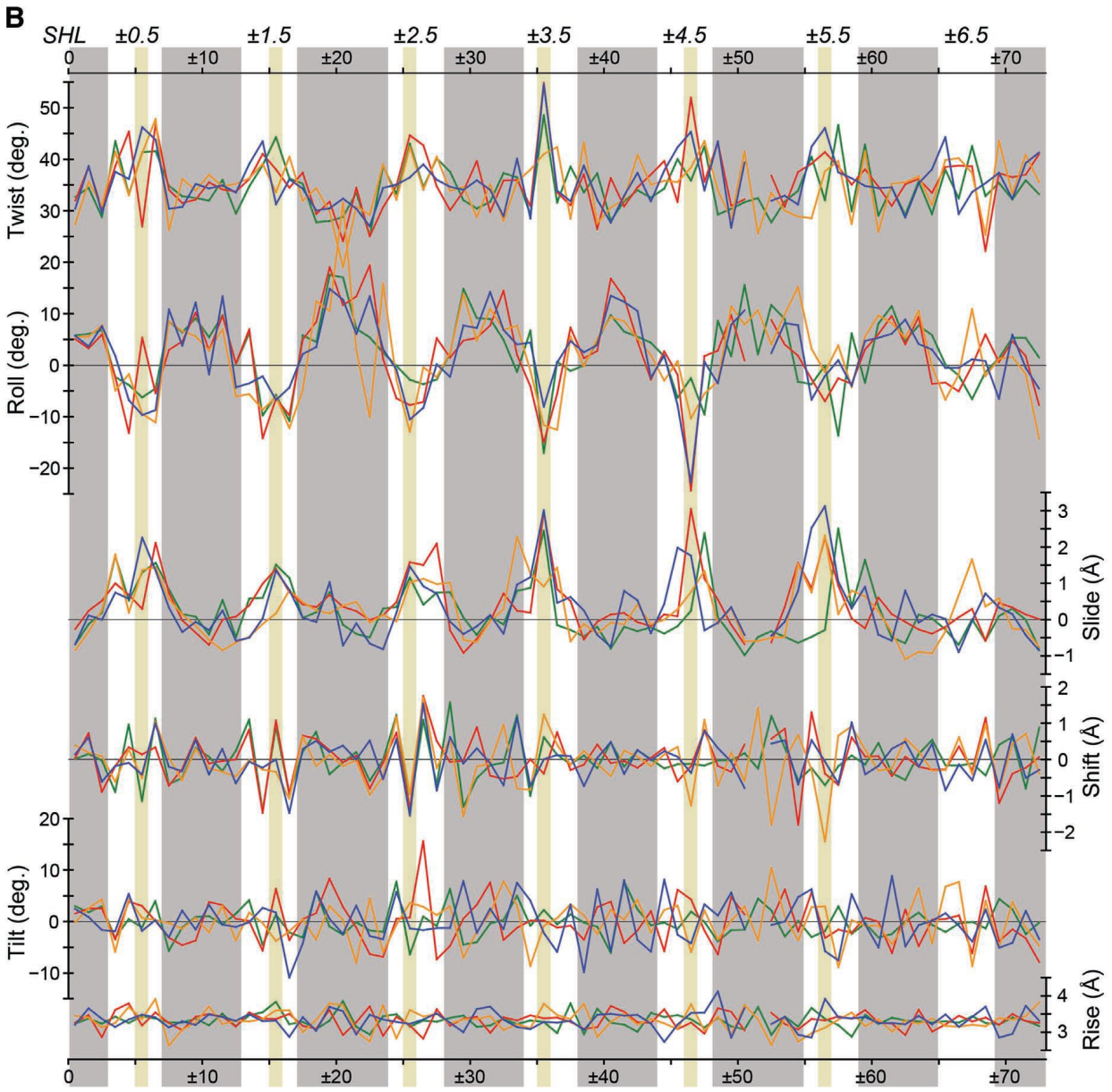
Results

- Crystal structure did not confirm kinked conformation around minor groove TA dinucleotides as expected
- 601R had good stability, but GGG(G) motif in 6 (of 8) central binding points
- Thus it is not clear, what role TA dinucleotide plays
- It is not as simple, as suggested by Widom original article

Dinucleotide step parameters



(A) Illustration of the six degrees of freedom for DNA structure at the base pair step level.



Dinucleotide step parameters

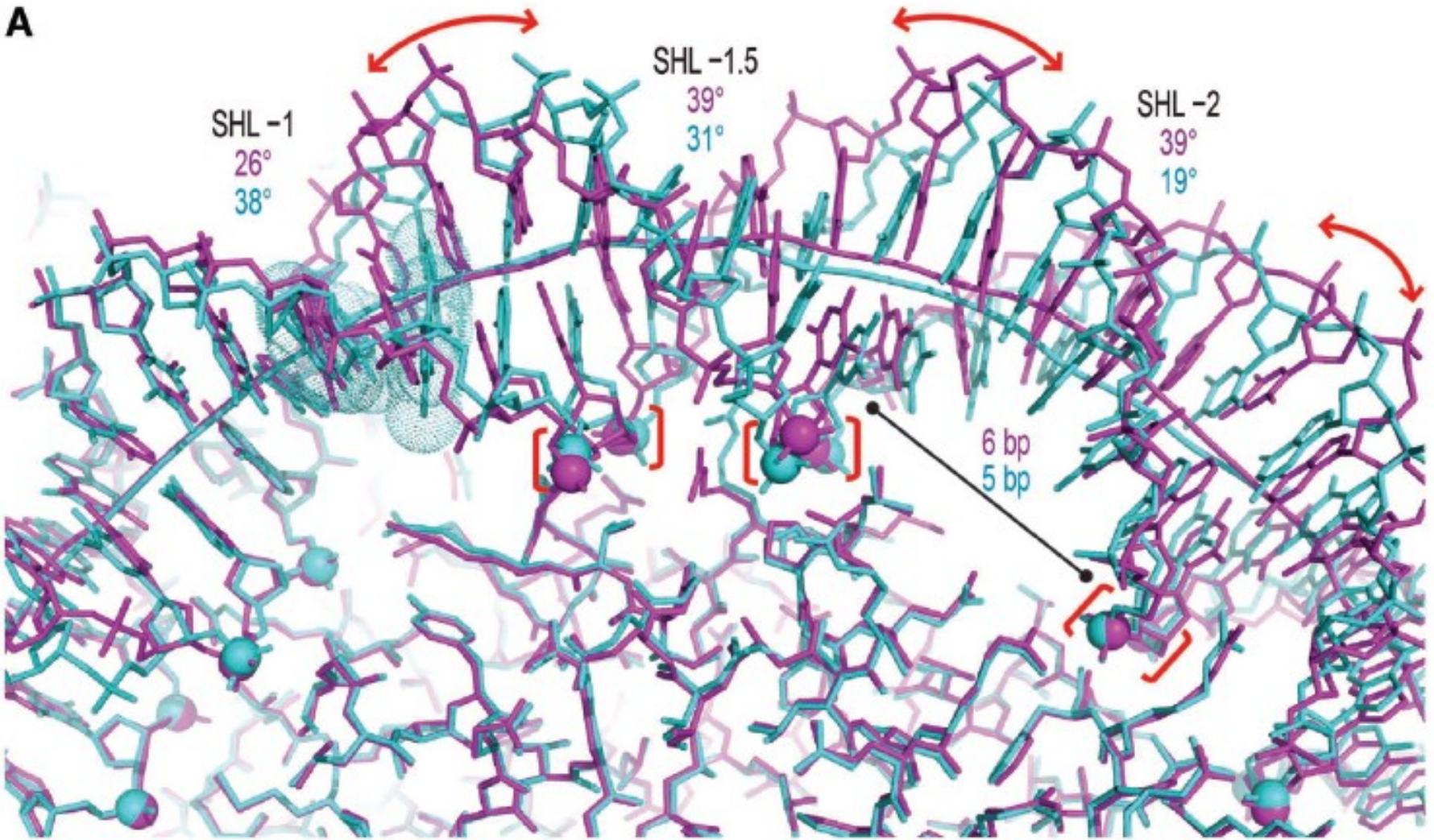
(B) Dinucleotide step values for **NCP-601L** (**blue**) and **NCP147** (**green**) averaged over one particle half and for the two particle halves of **NCP146b** (**red** and **yellow**; NCP146b displays a distinct DNA–histone register in each half).

Dinucleotide steps in major groove-inward sections in addition to the flanking major-to-minor groove-inward interface steps have a grey shaded background. The four dinucleotide steps in each minor groove-inward section have a white background with a gold shading indicating the step located at the pressure point

DNA sequence dependency

- Previous models for the sequence-dependent mechanics of wrapping in the nucleosome have been based on the contribution of specific DNA conformational parameters, foremost base pair step roll and slide
- Should not forget DNA-histone interactions
- The phosphorus atoms of four phosphate groups, designated as the ‘binding platform’, lie roughly in a plane and show the least variation in position between different DNA sequences
- Inward facing minor- and major grooves have to be compressed
- Two sides of binding platform act as "hinges"
- These allow big amount of conformational freedom depending on DNA sequence (145 and 147)

Different structure



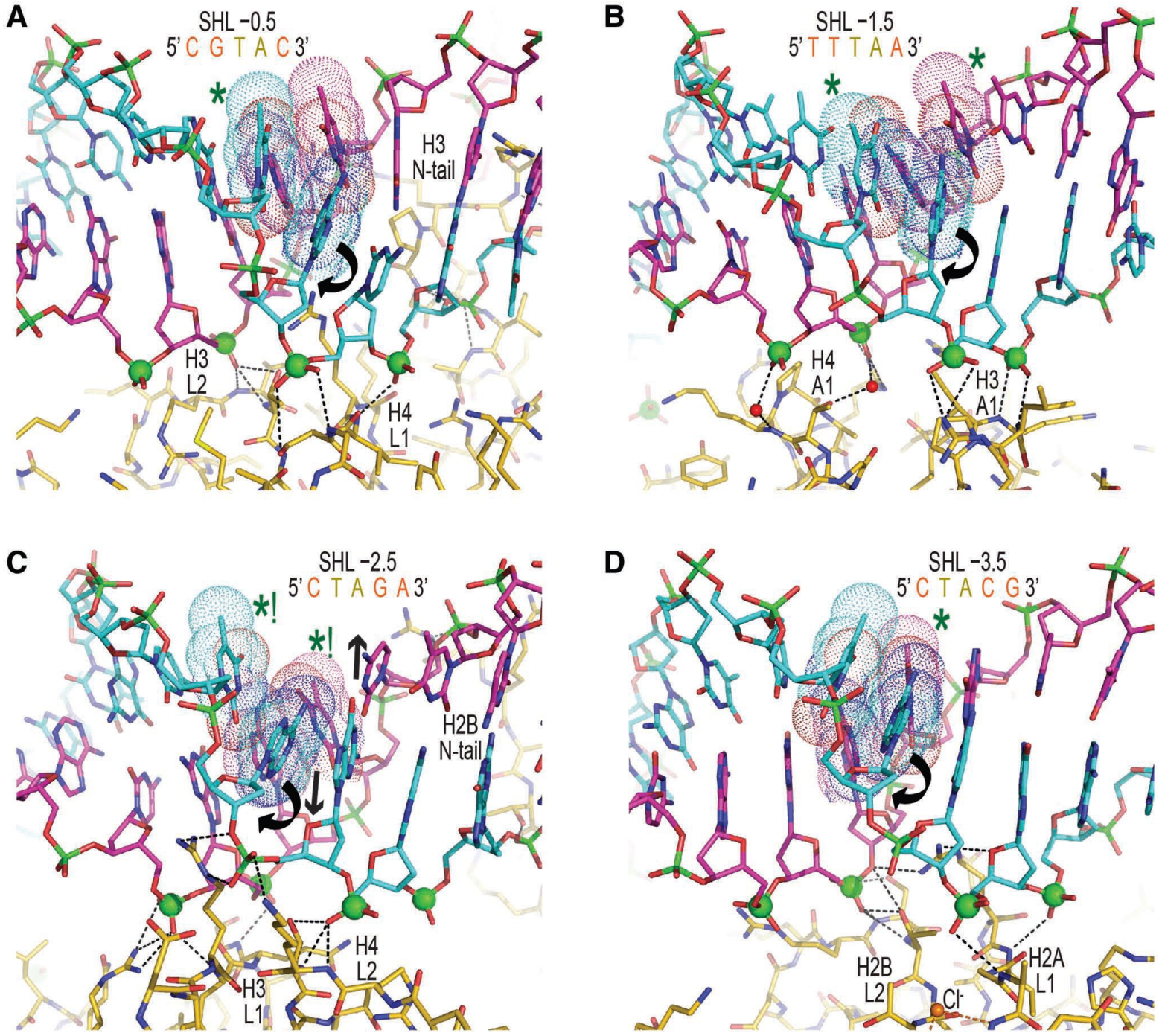
DNA single-strand histone attachment points act as hinges to allow double helix conformational freedom.

H3–H4 tetramer dominates sequence dependency

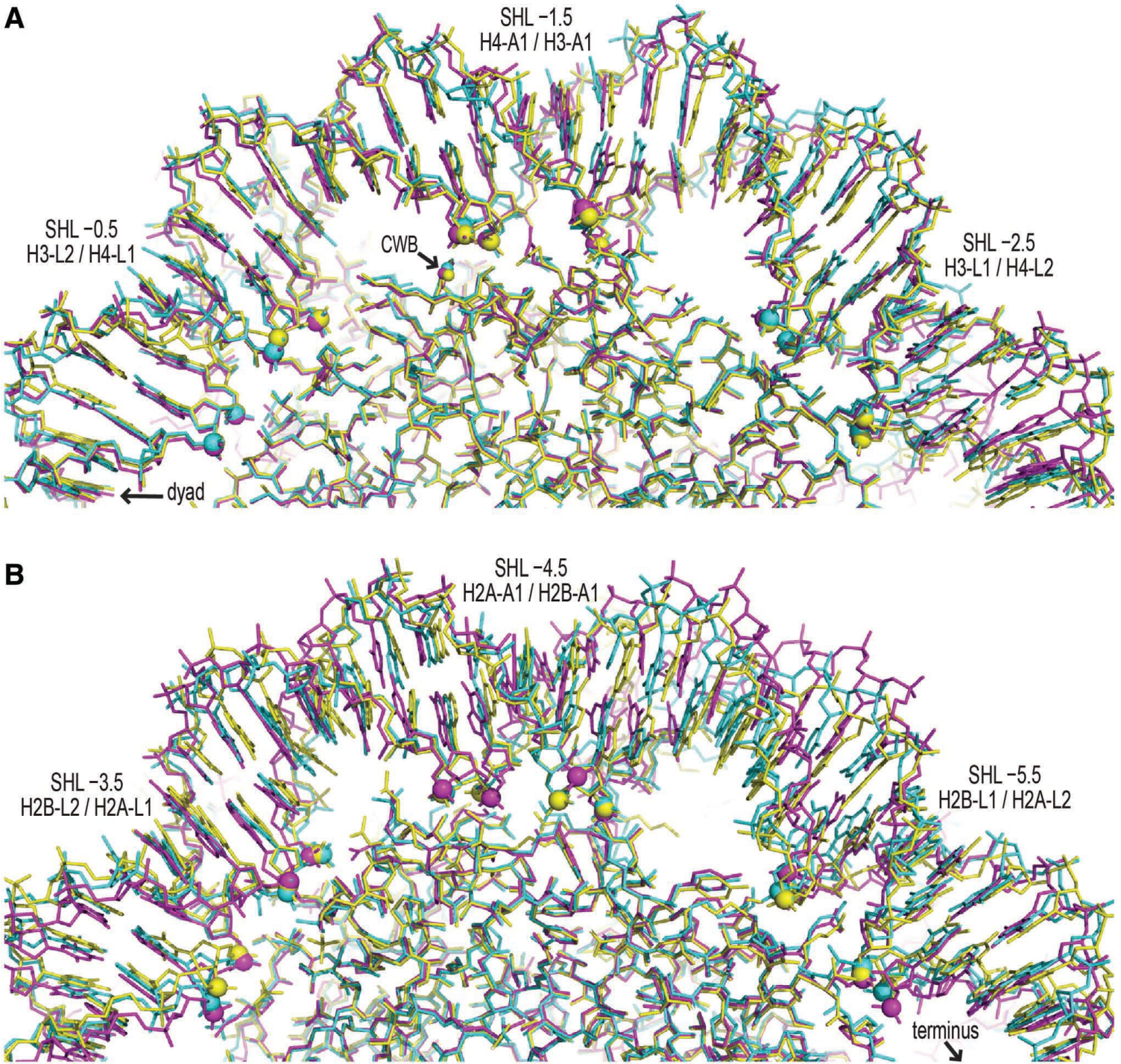
- DNA–histone interaction strength is greater over the H3–H4 tetramer compared to the H2A–H2B dimers
- So there is more conformational freedom at H2 binding sites
- And H3/H4 associated sequence has more influence over stability
- H3–H4 association could impose discrimination between the two distinct forms of minor groove bending – isolated kinks or smooth bending with alternating shift
- Thus, the double helix seems free to adopt one of a subset of favourable conformational modes within the confines of the binding
- There is histone-specific preferences – SHL +/- 2.5 consistently has alternating shift pattern

Minor groove inwards regions

- Dominate positioning and stability
- Major groove inwards regions have more conformational freedom and are less conserved (although generally GC-rich)
- α-satellite sequences had kinks on AT dinucleotides
- 601L had smooth (although higher than normal) bending... but still AT dinucleotides had unusual positioning
- Much reduced roll compared to α-satellite sequences
- AT dinucleotides coincide with the places of maximum reactivity
- Extreme propeller twisting of AT pairs, other shifts close to ideal
- Weakened hydrogen bonds
- TTTAA is naturally compressed. If this motif is introduced to α-satellite sequences, kinking is much reduced
- Most severe kinking happens, if GC is in minor groove



601L sequence binding to histone



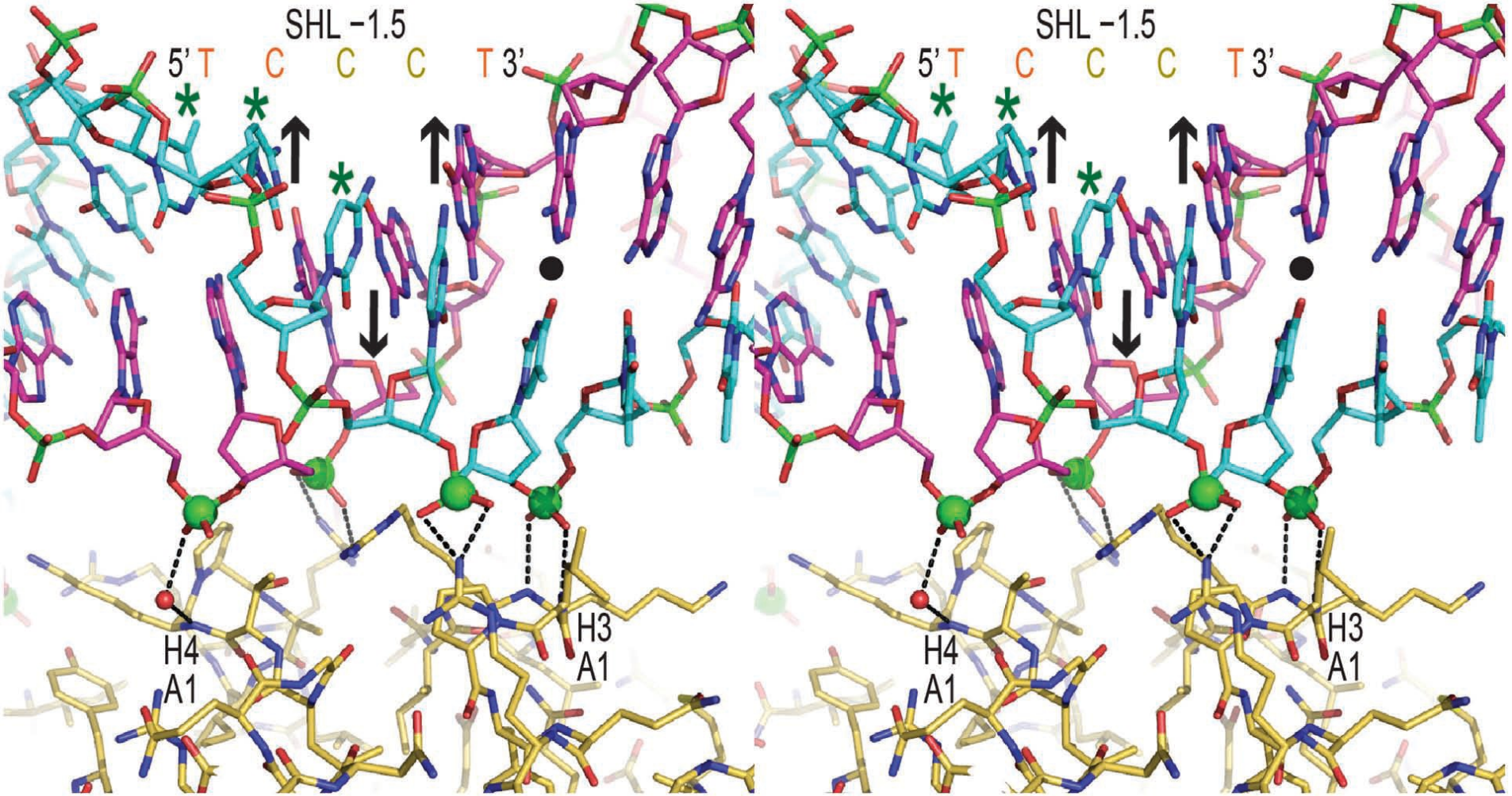
147, 146b and 601L binding comparison

Discussion

- Histone provides rigid scaffold, DNA has to find energetically most favourable conformation
- The ranking of minor-groove inwards binding sites according to contribution to stability is:
  $SHL\pm1.5 > SHL\pm2.5 > SHL\pm0.5 > SHL\pm3.5 > SHL\pm4.5 > SHL\pm5.5$
- The salt stability of various DNA sequences varies roughly twofold
- Permanganate reactivity hotspots coincided with minor-groove inwards AT dinucleotides ($SHL \pm0.5, \pm1.5, \pm2.5, \pm3.5$)
- Positioning of TA steps at these dominant pressure points allows for a minimum distortion energy, because TA is the loosest stacking dinucleotide step in conjunction with the weaker base pairing interaction for AT versus GC
- TA step is free to distort in a great variety of ways

Discussion

- The overall degree of DNA distortion in the nucleosomal state relative to the naked state, estimated by the differential permanganate reactivity, is larger for NCP-601L compared to NCP146b
- The binding energy of NCP146b, on the other hand, is much less
- NCP146b displays reactivity hotspots also at *SHL±2* and *±5* (i.e. major groove inwards sites)
- I.e. Extra energy is needed to bend AT dinucleotides at major groove binding sites compared to GC dinucleotides
- The optimal nucleotide context depends on histone binding site
- GC dinucleotides can take favourable conformations



- GjC-rich elements can undergo energetically favourable distortions at minor groove-inward positions by adopting specialized conformations. Stereo view of the AGGGA (=TCCCT) motif at the SHL1.5 location in NCP146b, which displays smooth bending with pronounced alternating displacement of base pairs into the minor groove (downward-pointing arrow and filled circle indicating displacement away from the viewer) and major groove (upward-pointing arrows) via fluctuating shift

Discussion

- The reactivity of TA at TTTAA is the next strongest in bound state at SHL+-1.5
- The same motif has very low reactivity in native state
- Ideal for groove narrowing that happens at SHL+-1.5
- CTAGA motif has high reactivity in native state (i.e. distorted structure)
- Ideal for alternating shift at SHL+-2.5
- Highest affinity sequences are GC rich (except for atrategic AT dinucleotides at binding sites)
- Second-best motifs are sometimes GC rich:
  - SHL±2.5 AGGGG
  - SHL ±0.5 GGGGA

Conclusions

- Different DNA motifs take different configurations at histone binding sites
- Different configuration changes can result in similar energetically favourable bindings
- Step-flexible and step-rigid dinucleotides can coperatively help binding at nucleosome
- GC steps are generally more rigid for bending but more flexible for shifting
- GA/TC is very rigid, but it has highest shift value (0.28±0.46A for GA) in native state and thus can contribute to binding
- This may be useful for predicting histone binding sites in genome or creating synthetic genomes