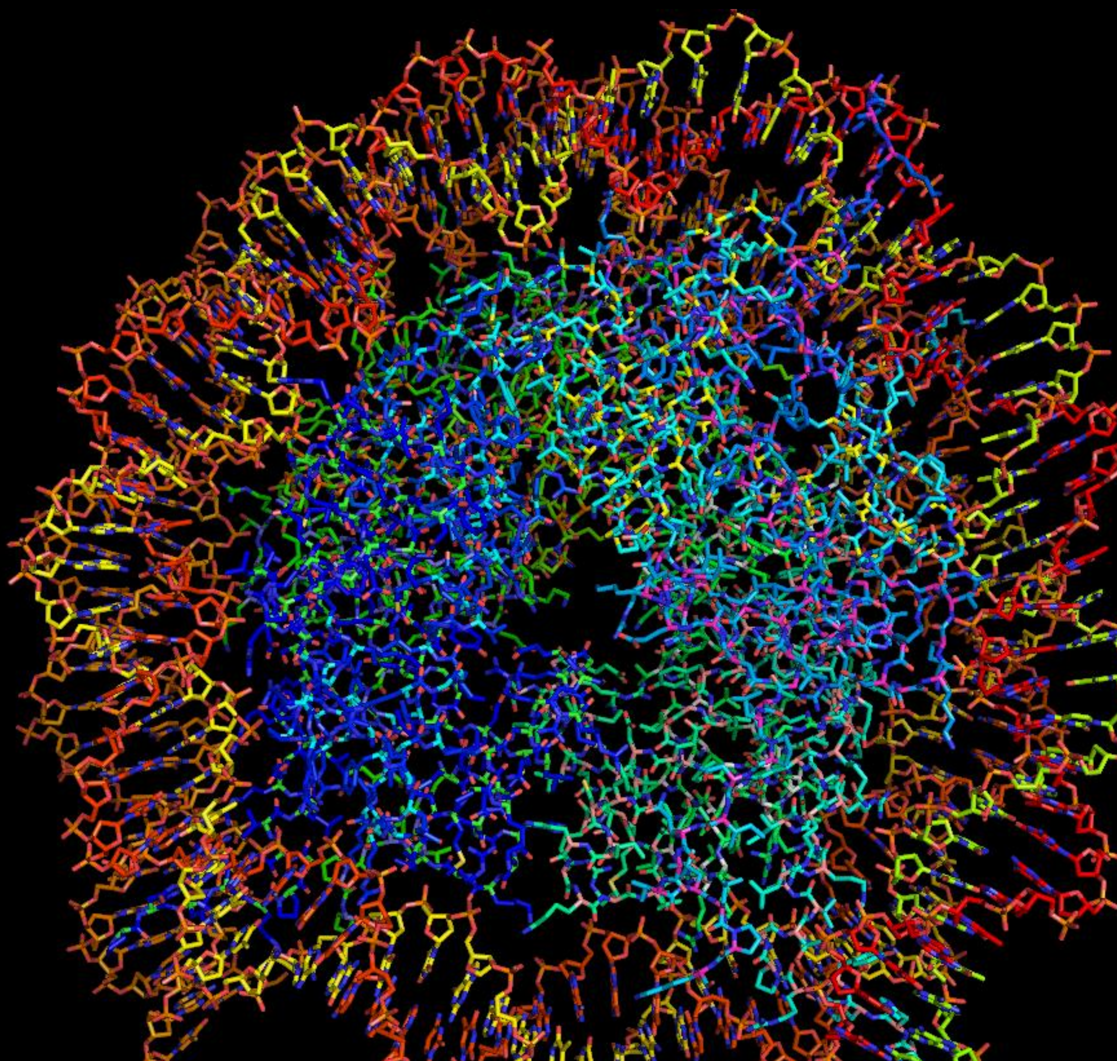
NUCLEOSOME



Victòria Brugada Ramentol
Pedro Fuentes Varela
Structural Biology 2012

Index

Nucleosome

A **Nucleosome** is the basic unit of DNA packaging in eukaryotes, consisting of a segment of DNA wound in sequence around eight histone protein cores.

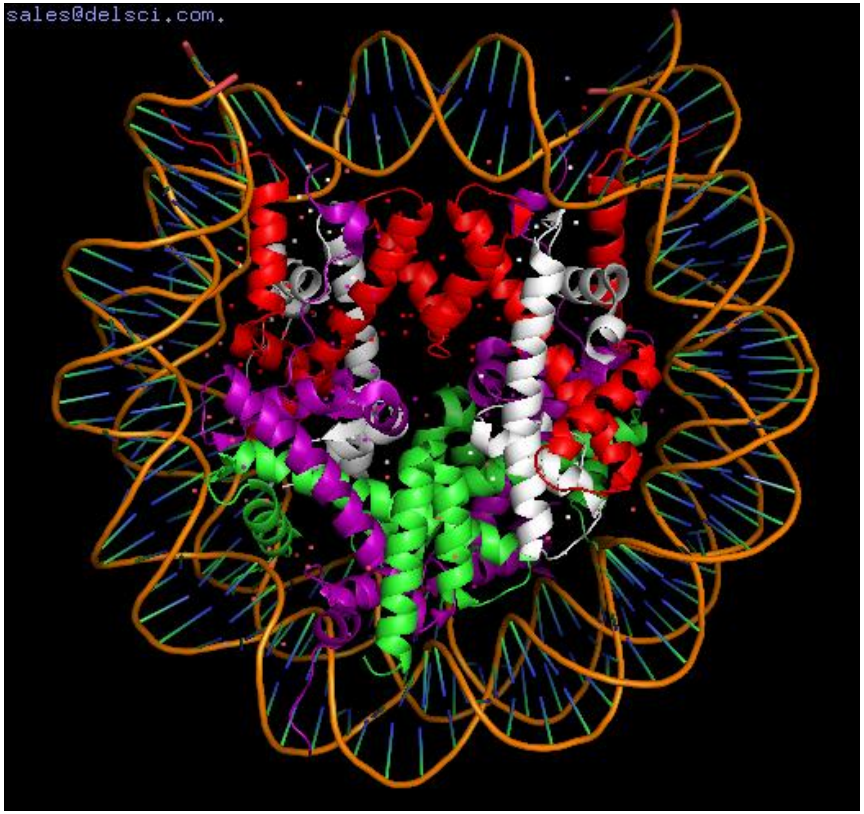


Figure1: nucleosome

DNA compaction

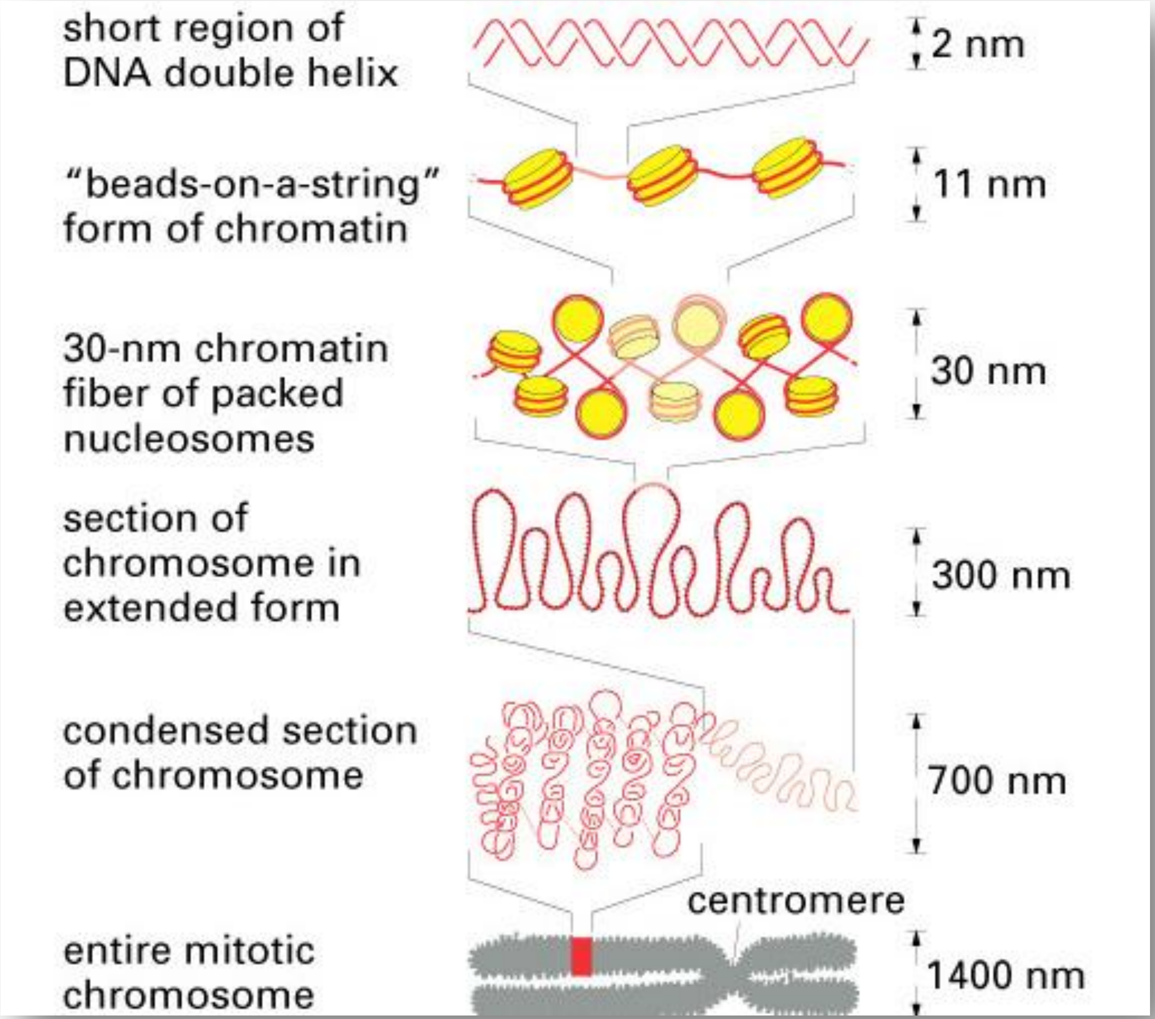


Figura2: levels of DNA condensation in the eukaryotic cell

Nucleosome's discovery

- 1888, Albrecht Kossel, isolated the histones and he identified them as the basic substances that bind to nucleic acids.

- If the naked DNA is partially digested with a nonspecific endonuclease is obtained a range of polynucleotide fragments.

- Within the chromatin, Kornberg found that roughly 200 bp of DNA are wrapped around an octamer of histone proteins.

- At the same time, images were obtained by electron microscope of fibers extended chromatin.

- In the 1970s, Kornberg discovered the nucleosome as the basic protein complex packaging chromosomal DNA in the nucleus of eukaryotic cells.

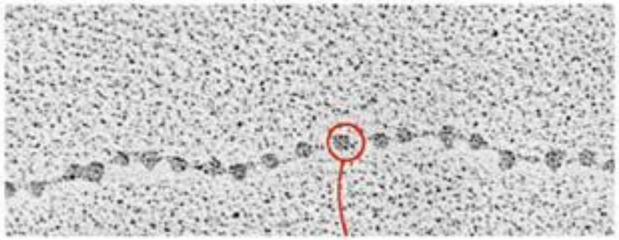


Figure3: Fibers extended chromatin

The highly basic nature of histones, aside from facilitating DNA-histone interactions, contributes to their water solubility.

	Histone3	Histone4	Histone 2A	Histone 2B
Basic residues	24.4%	26.4%	23.3%	24.8%

Normal basic content: 14%

Histones

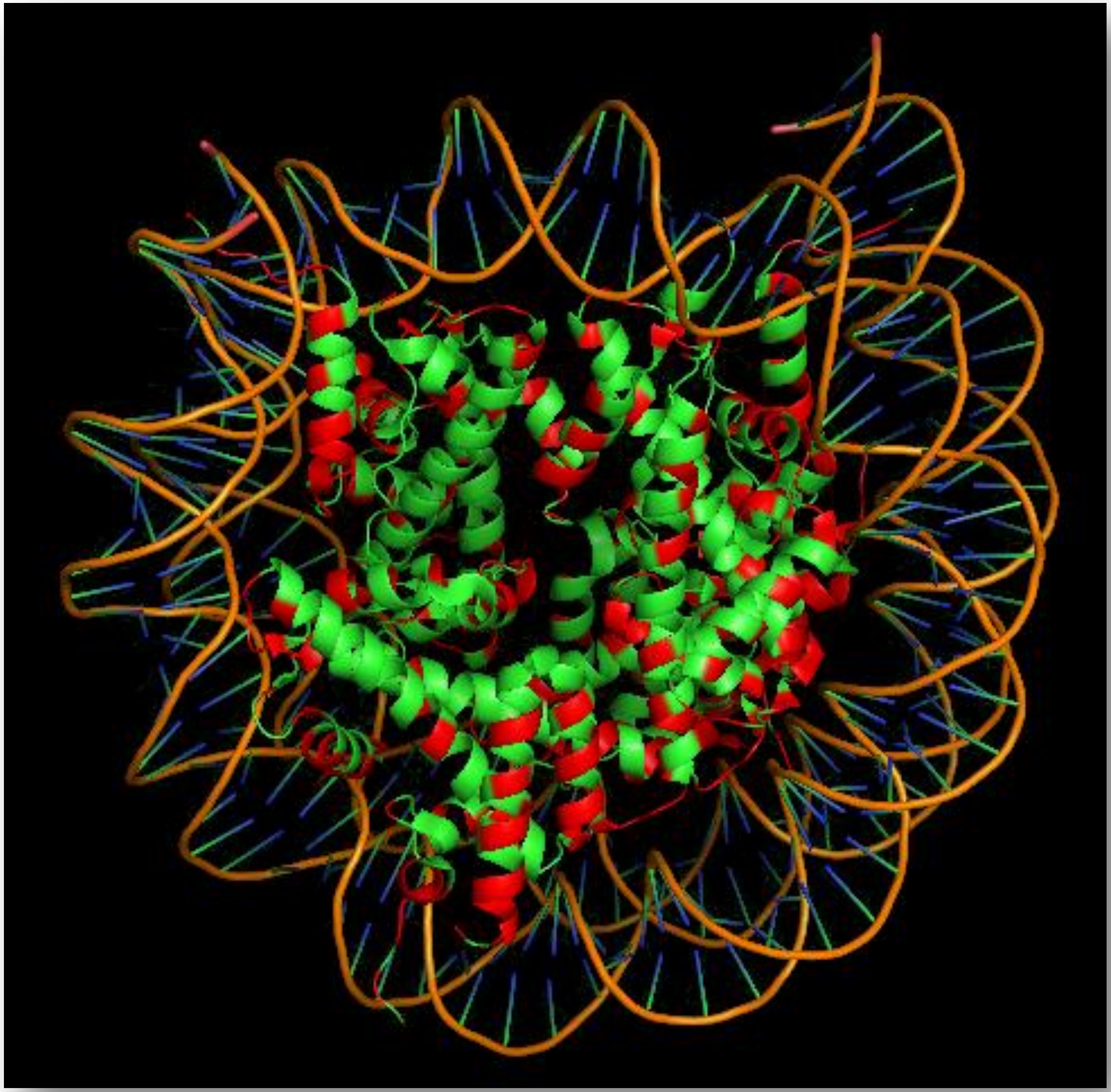


Figure5: histone core: four pais of histones and 147 bp of DNA (PDBid: 3AFA)

Introduction

NCP:
- H2A
- H2B
- H3
- H4
- 147 base pairs of DNA

Linker:
- H1

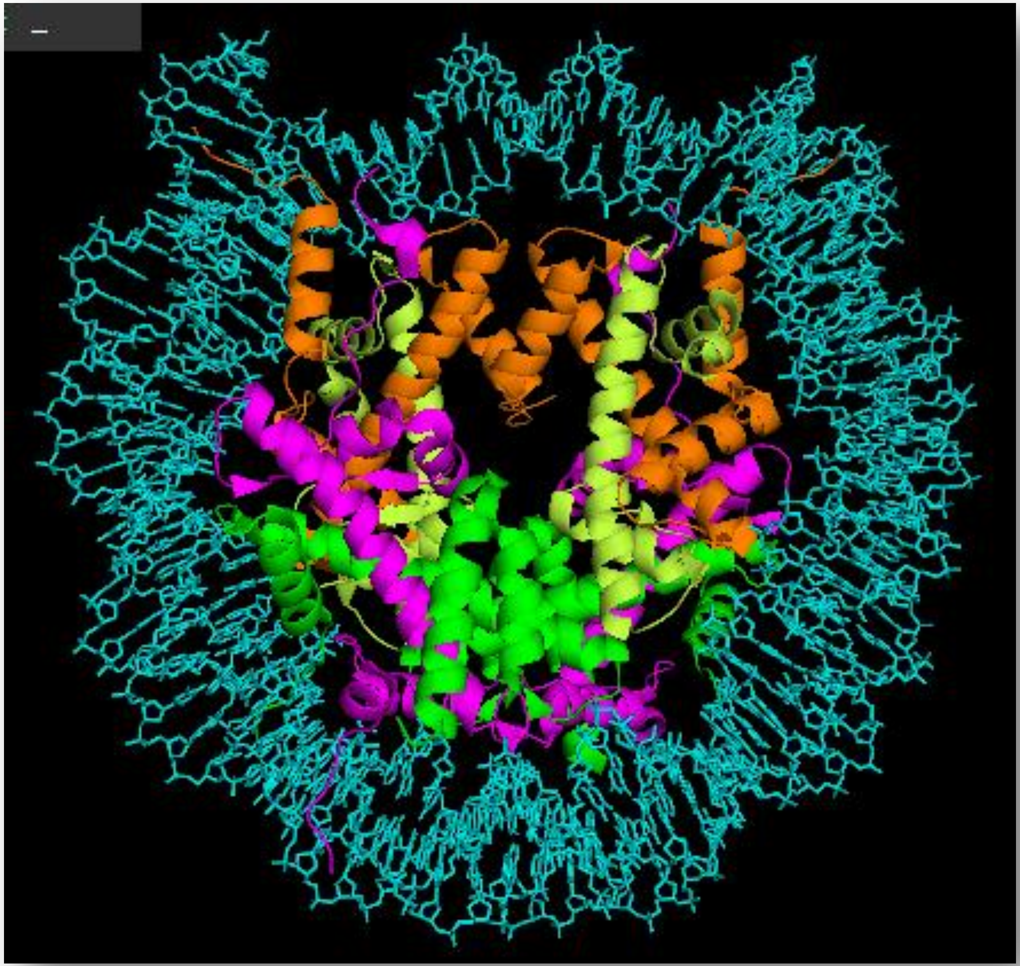


Figura6: histone core: four pais of histones and 147 bp of DNA (PDBid: 3AFA)

Introduction

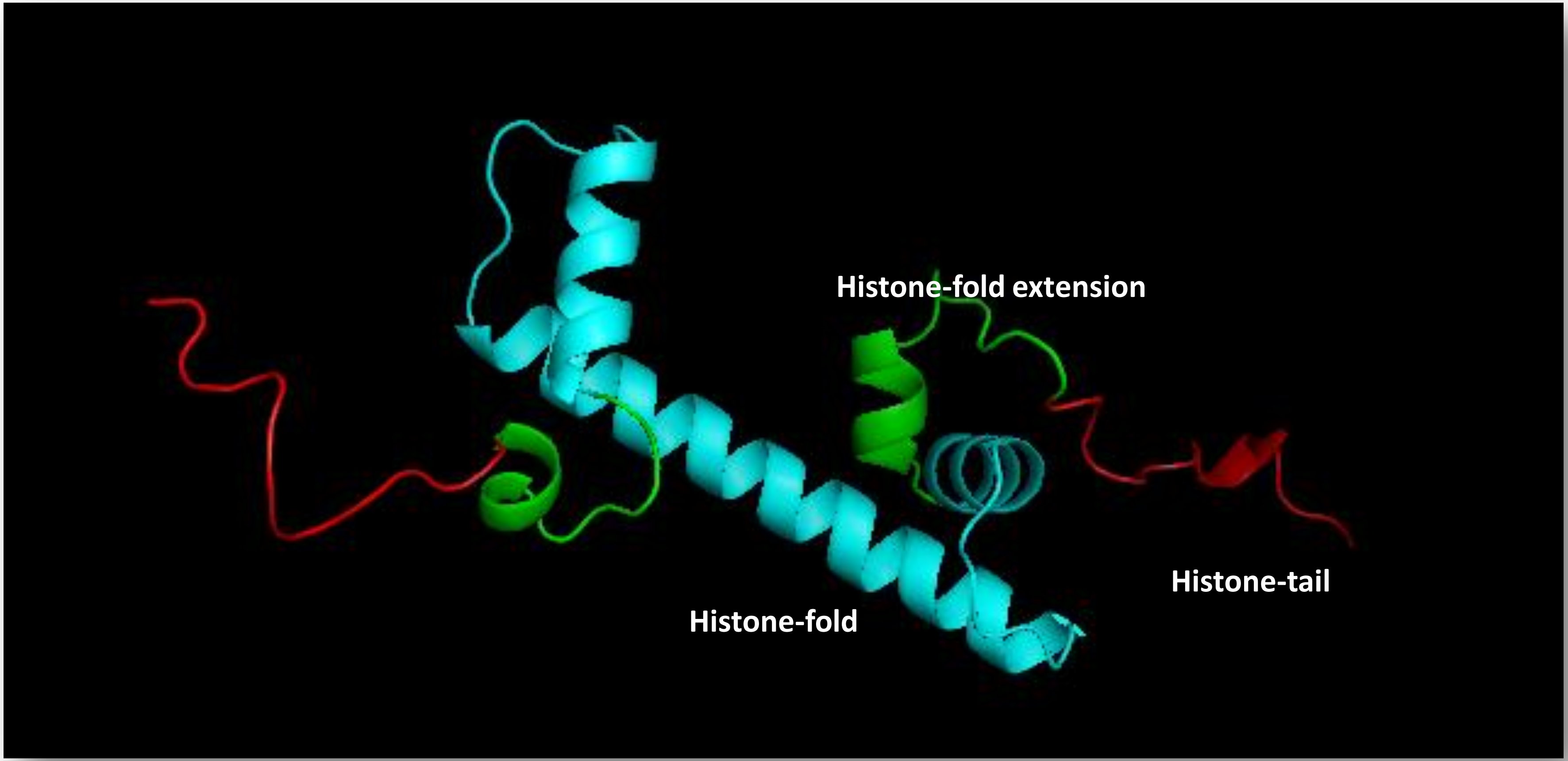


Figura7: Parts of the histone H2A (PDBid:1AOI)

Introduction

SCOP Classification

- **Class:** All alpha proteins
- **Fold:** Histone-fold
- **Superfamily:** Histone-fold
- **Family**: Nucleosome Core histones

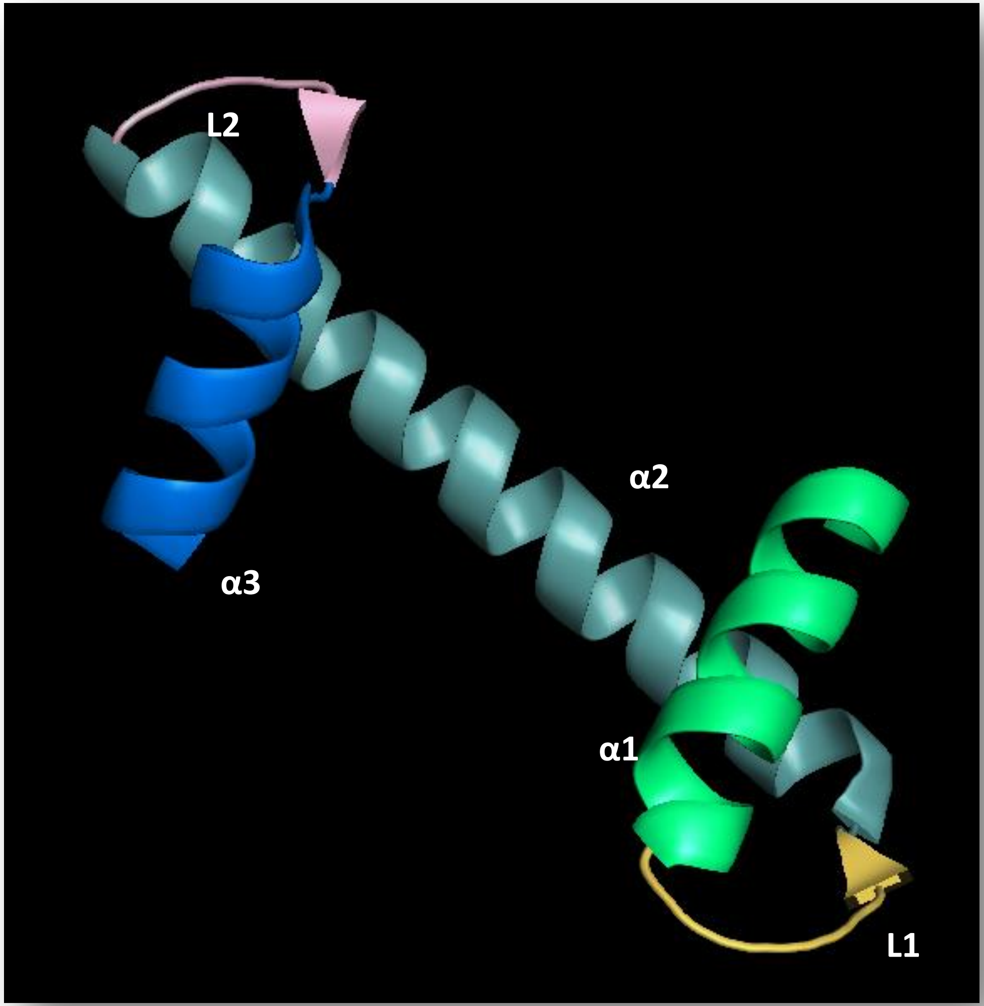


Figura8: Structure form the histone-fold from Histone-3 (PDBid: 3AFA)

Origin of the histone fold

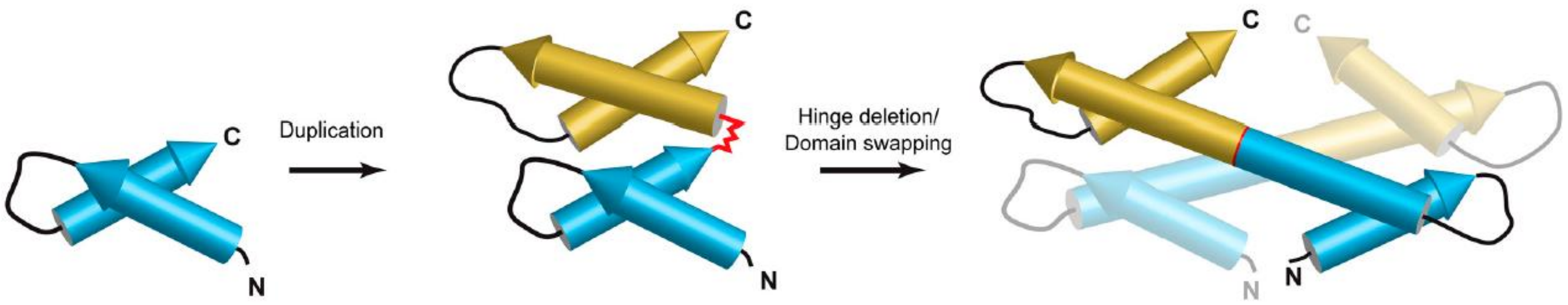
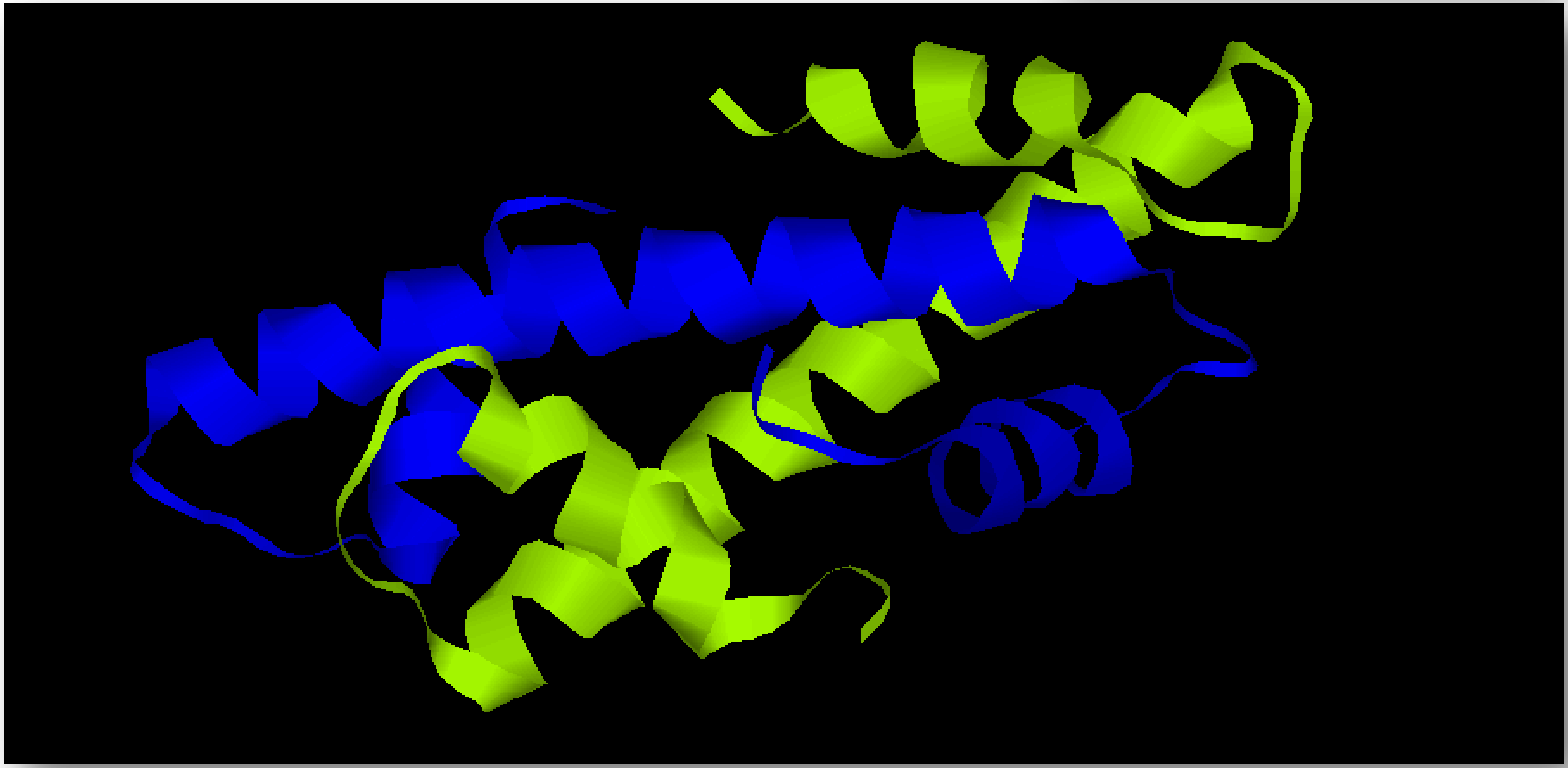


Figure9: Proposed evolutionary trajectory of the histone fold structure.
Hadjithomas M, Moudrianakis E. Experimetal evidence for the role of domain swapping in the evolution of the histone fol. PNAS. 2011; 108 (33): 13462-7.

Experimental data suggests that the histone-fold motif appeared from a duplication and 3-D domain swapping of two HSH units.

Handshake motif

Figure10: *handshake motif.*



Structure conservation

The histone fold is conserved in the eight nucleosomal histones.

RMSD	SCORE
1.53	8.65

Figure11: superimposition of the 8 core histones (PDBid: 3AFA)

Structure conservation

```
CLUSTAL W(1.60) multiple sequence alignment

H2A         -VGRVHRLLRK-G-N-Y-SERVGAGAPVYLAAVLEYLTAEILELAGNAARDNKKTRIIPRHLQLAIRN-
H2A'        -VGRVHRLLRK-G-N-Y-SERVGAGAPVYLAAVLEYLTAEILELAGNAARDNKKTRIIPRHLQLAIRN-
H2B         -----SIYVYKVLKQVHPDTGISSKAMGIMNSFVNDIFERIAGEASRLAHYNKRSTITSREIQTAVRLL
H2B'        -----SIYVYKVLKQVHPDTGISSKAMGIMNSFVNDIFERIAGEASRLAHYNKRSTITSREIQTAVRLL
H3          RKLPFQRLVREIAQDFKTDLRFQSSAVMALQEACEAYLVGLFEDTNLCAIHAKRVTIMPKDIQLARRIR
H3'         RKLPFQRLVREIAQDFKTDLRFQSSAVMALQEACEAYLVGLFEDTNLCAIHAKRVTIMPKDIQLARRIR
H4          -----KPAIRRLARRGGV-KRISGLIYEETRGVLKVFLENVIRDAVTYTEHAKRKTVTAMDVVYALKR-
H4'         -----KPAIRRLARRGGV-KRISGLIYEETRGVLKVFLENVIRDAVTYTEHAKRKTVTAMDVVYALKR-
```

Figure12: Multiple sequence alignment from NCP.

Structure conservation

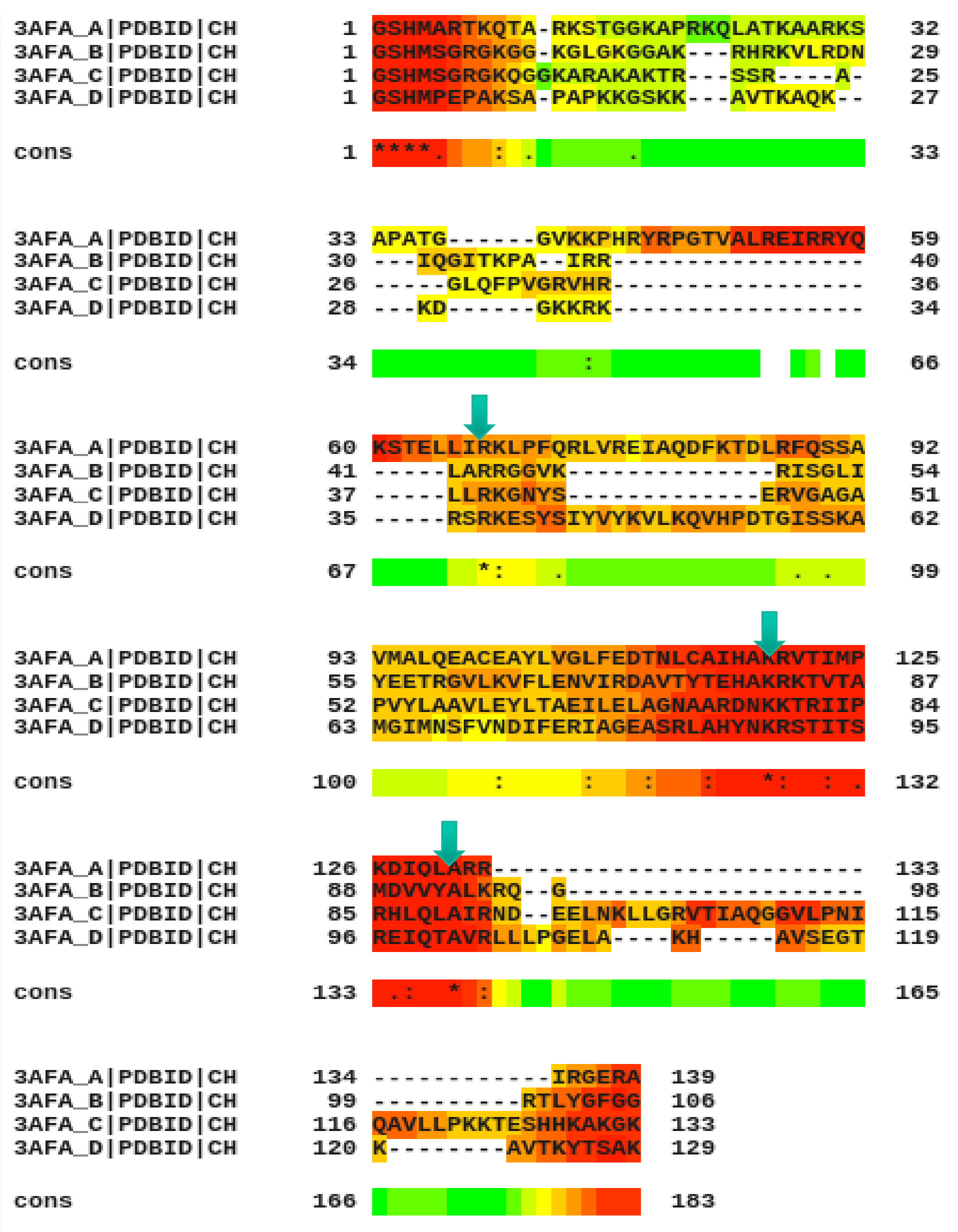
```
STAMP Structural Alignment of Multiple Proteins
 by Robert B. Russell & Geoffrey J. Barton
 Please cite PROTEINS, v14, 309-323, 1992

Running roughfit.

    Sc = STAMP score, RMS = RMS deviation, Align = alignment length
    Len1, Len2 = length of domain, Nfit = residues fitted
    Secs = no. equivalent sec. strucs. Eq = no. equivalent residues
    %I = seq. identity, %S = sec. str. identity
    P(m)  = P value (p=1/10) calculated after Murzin (1993), JMB, 230, 689-694

     No.  Domain1         Domain2         Sc     RMS    Len1 Len2  Align NFit Eq. Secs.   %I    %S    P(m)
Pair   1  H2A             H2B             7.24   2.08     62   65     62   57  55     0  25.45 100.00 0.00058
Pair   2  H2A             H3              7.58   1.60     62   68     66   58  57     0  28.07 100.00 0.00026
Pair   3  H2A             H4              7.13   1.41     62   63     63   50  48     0  12.50 100.00 1.00000
Pair   4  H2B             H3              8.55   1.34     65   68     65   65  64     0  20.31 100.00 0.00609
Pair   5  H2B             H4              8.26   1.38     65   63     64   60  59     0  15.25 100.00 1.00000
Pair   6  H3              H4              7.69   1.39     68   63     65   59  58     0  18.97 100.00 0.03320
Reading in matrix file histonefold.mat...
Doing cluster analysis...
Cluster:  1 (      H2B   &        H3 ) Sc  8.55 RMS    1.30 Len  65 nfit  64
 See file histonefold.1 for the alignment and transformations
Cluster:  2 (       H4   &       H2B        H3 ) Sc  8.79 RMS    1.28 Len  64 nfit  59
 See file histonefold.2 for the alignment and transformations
Cluster:  3 (      H2A   &        H4       H2B        H3 ) Sc  8.67 RMS    1.53 Len  66 nfit  53
 See file histonefold.3 for the alignment and transformations
```

Figure13: Output of the structural alignment performed with STAMP (PDBid: 3AFA)

Pairwise alignment shows that H2A is the one that differs the most, while the other three are very similar.

Structure conservation

However, the sequence conservation is very low.



```
3AFA_A|PDBID|CH     1 GSHMARTKQTA-RKSTGGKAPRKQLATKAARKS  32
3AFA_B|PDBID|CH     1 GSHMSGRGKGG-KGLGKGGAK---RHRKVLRDN  29
3AFA_C|PDBID|CH     1 GSHMSGRGKQGGKARAKAKTR---SSR----A-  25
3AFA_D|PDBID|CH     1 GSHMPEPAKSA-PAPKKGSKK---AVTKAQK--  27

cons                1 ****.   :  .      .                33

3AFA_A|PDBID|CH    33 APATG------GVKKPHRYRPGTVALREIRRYQ  59
3AFA_B|PDBID|CH    30 ---IQGITKPA--IRR-----------------  40
3AFA_C|PDBID|CH    26 -----GLQFPVGRVHR-----------------  36
3AFA_D|PDBID|CH    28 ---KD------GKKRK-----------------  34

cons               34                :                    66

3AFA_A|PDBID|CH    60 KSTELLIRKLPFQRLVREIAQDFKTDLRFQSSA  92
3AFA_B|PDBID|CH    41 -----LARRGGVK--------------RISGLI  54
3AFA_C|PDBID|CH    37 -----LLRKGNYS-------------ERVGAGA  51
3AFA_D|PDBID|CH    35 -----RSRKESYSIYVYKVLKQVHPDTGISSKA  62

cons               67       *:    .                 .  .  99

3AFA_A|PDBID|CH    93 VMALQEACEAYLVGLFEDTNLCAIHAKRVTIMP 125
3AFA_B|PDBID|CH    55 YEETRGVLKVFLENVIRDAVTYTEHAKRKTVTA  87
3AFA_C|PDBID|CH    52 PVYLAAVLEYLTAEILELAGNAARDNKKTRIIP  84
3AFA_D|PDBID|CH    63 MGIMNSFVNDIFERIAGEASRLAHYNKRSTITS  95

cons              100         :     :   :   :    *:   : . 132

3AFA_A|PDBID|CH   126 KDIQLARR------------------------- 133
3AFA_B|PDBID|CH    88 MDVVYALKRQ--G-------------------- 98
3AFA_C|PDBID|CH    85 RHLQLAIRND--EELNKLLGRVTIAQGGVLPNI 115
3AFA_D|PDBID|CH    96 REIQTAVRLLLPGELA----KH-----AVSEGT 119

cons              133  .:   *  :                          165

3AFA_A|PDBID|CH   134 ------------IRGERA 139
3AFA_B|PDBID|CH    99 ----------RTLYGFGG 106
3AFA_C|PDBID|CH   116 QAVLLPKKTESHHKAKGK 133
3AFA_D|PDBID|CH   120 K--------AVTKYTSAK 129

cons              166                    183
```

Figure14: Sequence alingment of the four histones in a NCP using the program T-COFFEE (PDBid: 3AFA).

Structure conservation

However, the sequence conservation is very low.

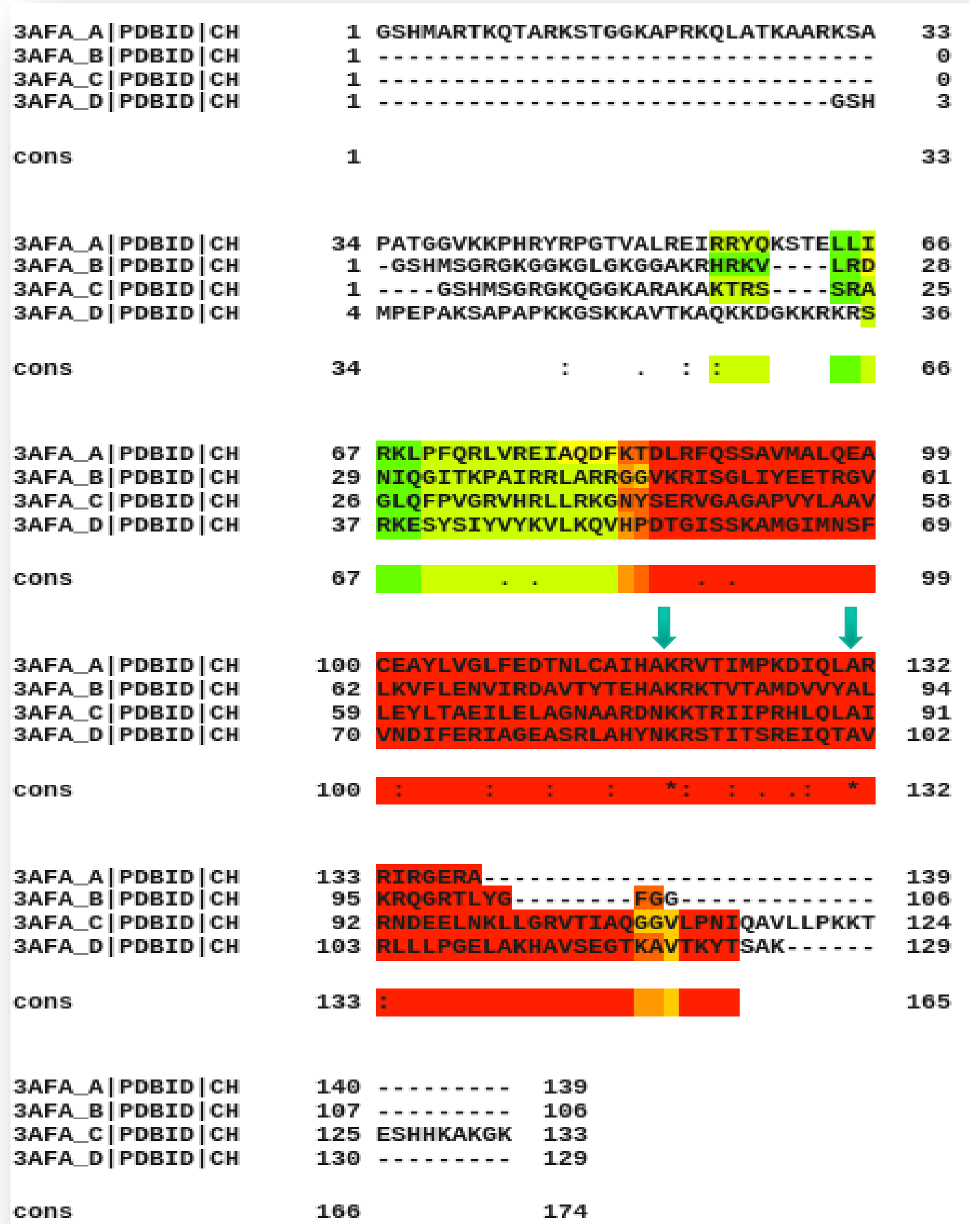


Figure15: Sequence alingment of the four histones in a NCP using the program M-COFFEE(PDBid: 3AFA).

H1 Linker histone

There's no histone-fold structure in H1 Linker histone.

SCOP Classification

- **Class**: All alpha proteins
- **Fold**: DNA/RNA binding 3-helical bundle.
- **Superfamily**: winged helix DNA binding domain.
- **Family**: Linker histone H1/H5

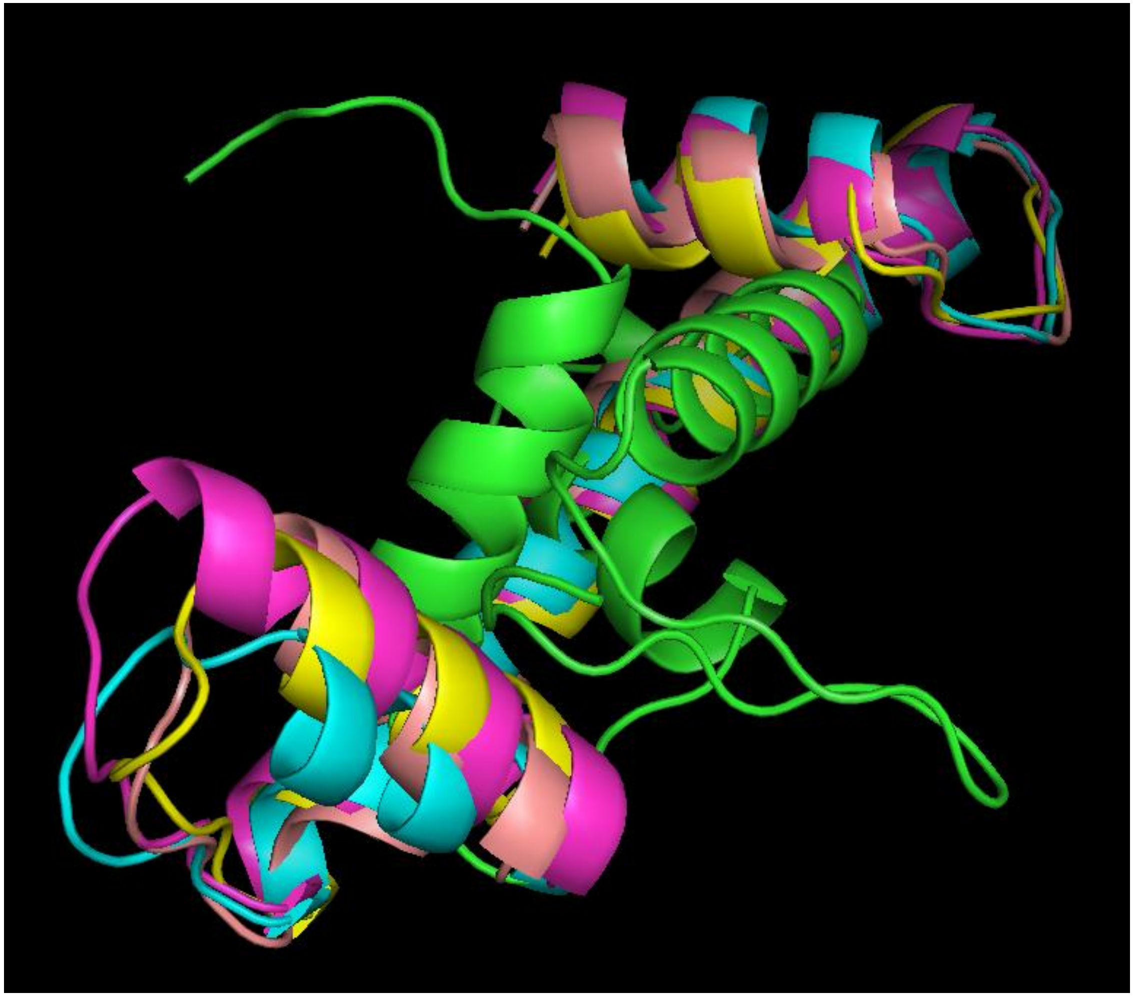


Figura16: Superimposition of the four core histones with the H1 linker.

H1 Linker histone

```
STAMP Structural Alignment of Multiple Proteins
 by Robert B. Russell & Geoffrey J. Barton
 Please cite PROTEINS, v14, 309-323, 1992

Running roughfit.

    Sc = STAMP score, RMS = RMS deviation, Align = alignment length
    Len1, Len2 = length of domain, Nfit = residues fitted
    Secs = no. equivalent sec. strucs. Eq = no. equivalent residues
    %I = seq. identity, %S = sec. str. identity
    P(m)  = P value (p=1/10) calculated after Murzin (1993), JMB, 230, 689-694

     No.  Domain1         Domain2         Sc     RMS    Len1 Len2  Align NFit Eq. Secs.   %I    %S    P(m)
Pair   1  H2A             H2B             7.24   2.08     62   65     62   57  55     0  25.45 100.00 0.00058
Pair   2  H2A             H3              7.58   1.60     62   68     66   58  57     0  28.07 100.00 0.00026
Pair   3  H2A             H4              7.13   1.41     62   63     63   50  48     0  12.50 100.00 1.00000
Pair   4  H2A             H1              1.49   2.38     62   78     67   25  25     0   4.00 100.00 1.00000 LOW SCORE
Pair   5  H2B             H3              8.55   1.34     65   68     65   65  64     0  20.31 100.00 0.00609
Pair   6  H2B             H4              8.26   1.38     65   63     64   60  59     0  15.25 100.00 1.00000
Pair   7  H2B             H1              0.32 100.00     65   78     73    2   0     0   0.00   0.00 1.00000 LOW SCORE
Pair   8  H3              H4              7.69   1.39     68   63     65   59  58     0  18.97 100.00 0.03320
Pair   9  H3              H1              0.36 100.00     68   78     86    2   0     0   0.00   0.00 1.00000 LOW SCORE
Pair  10  H4              H1              0.70   1.55     63   78     84    7   6     0  16.67 100.00 1.00000 LOW SCORE
Reading in matrix file stamp_h1.mat...
Doing cluster analysis...
Cluster:  1 (     H2B  &       H3 ) Sc  8.55 RMS   1.30 Len  65 nfit  64
 See file stamp_h1.1 for the alignment and transformations
Cluster:  2 (      H4  &      H2B       H3 ) Sc  8.79 RMS   1.28 Len  64 nfit  59
 See file stamp_h1.2 for the alignment and transformations
Cluster:  3 (     H2A  &       H4      H2B       H3 ) Sc  8.67 RMS   1.53 Len  66 nfit  53
 See file stamp_h1.3 for the alignment and transformations
Cluster:  4 (      H1  &      H2A       H4      H2B       H3 ) Sc  1.60 RMS   3.47 Len  80 nfit  11 LOW SCORE
 See file stamp_h1.4 for the alignment and transformations
```

Figura17: STAMP from four histone core with histone 1.

H1 Linker histone

```
CLUSTAL W(1.60) multiple sequence alignment

H1         EASSKSYRELIIEGLTALKERKG------------------------SSRP-ALKKFIK
H2A        --------------------VGRVHRLLRK----GNYAERVGAGAPVYL--AAVLEYLT-
H3         --------------------KLPFQRLVREIAQDFKTDLRFQSSAVMAL--QEASEAYL-
H2B        ------------------Y-AIYVYKVLK-Q---VHPDTGISSKAMSIM--NSFVNDVF-
H4         --------------------KPAIRRLAR-R---GGV-KRISGLIYEET--RGVLKVFL-


H1         ENYPIVGSAS-N--------FDL--YFN--NAIKKGVEAGDFEQPKGPAGAVKLAKK
H2A        ---AEILELAGNAARDNKKTRIIPRH-LQLA-VR---------------------
H3         ---VALFEDTNLCAIHAKRVTIMPKD-IQLARRIR--------------------
H2B        ---ERIAGEASRLAHYNKRSTITSRE-IQTAVRLL--------------------
H4         ---ENVIRDAVTYTEHAKRKTVTAMD-VVYALKRQ--------------------
```

Figura18: CLUSTAL. Estructural alignment

Structure conservation

Lys49 and Ala61* are the only two very conserved residues

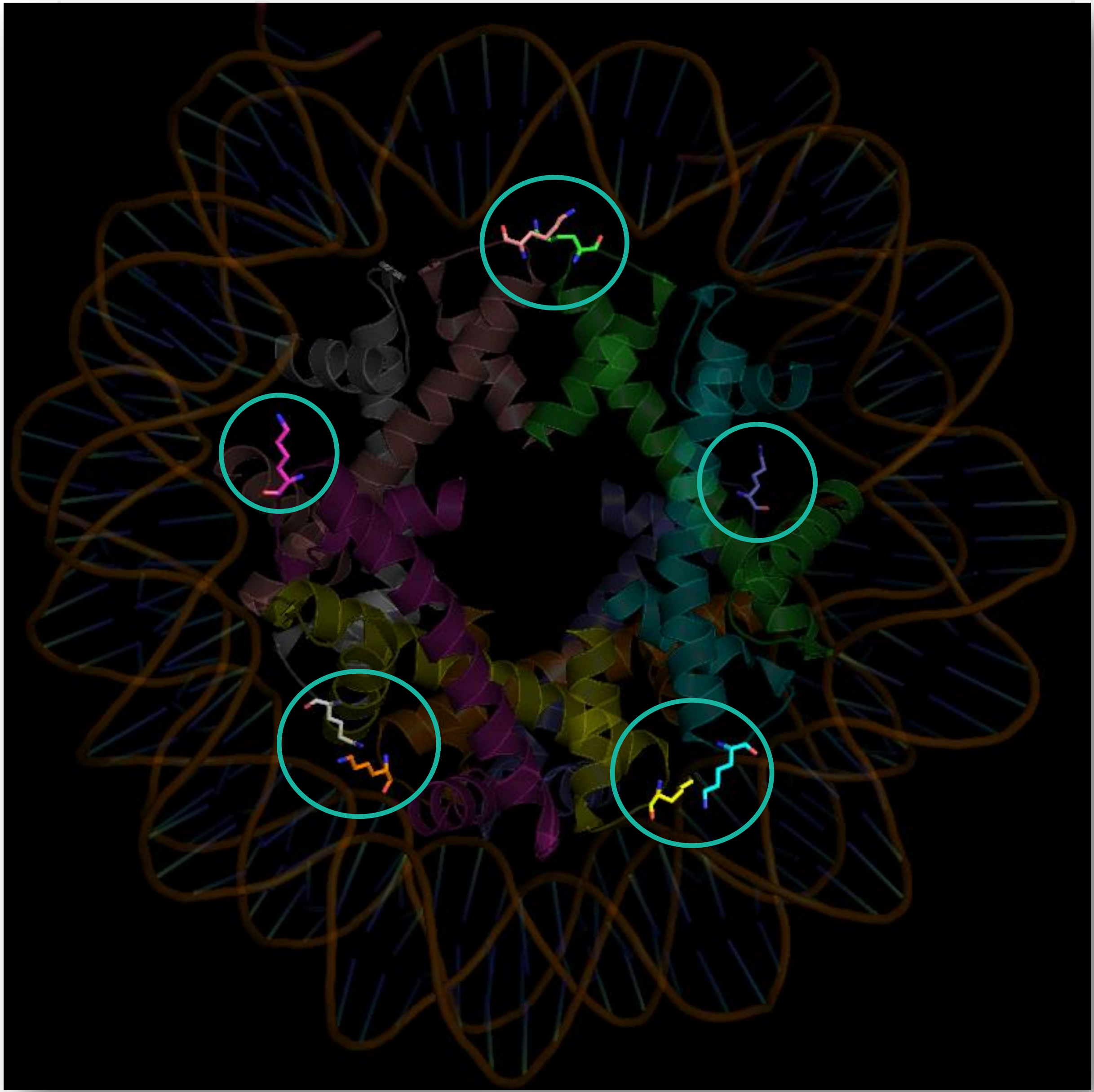


Figure19: Lys49* sequence conservation, shown in the nucleosome structure.

* Being 1 the first residue of each histone-fold sequence

STAMP LYS

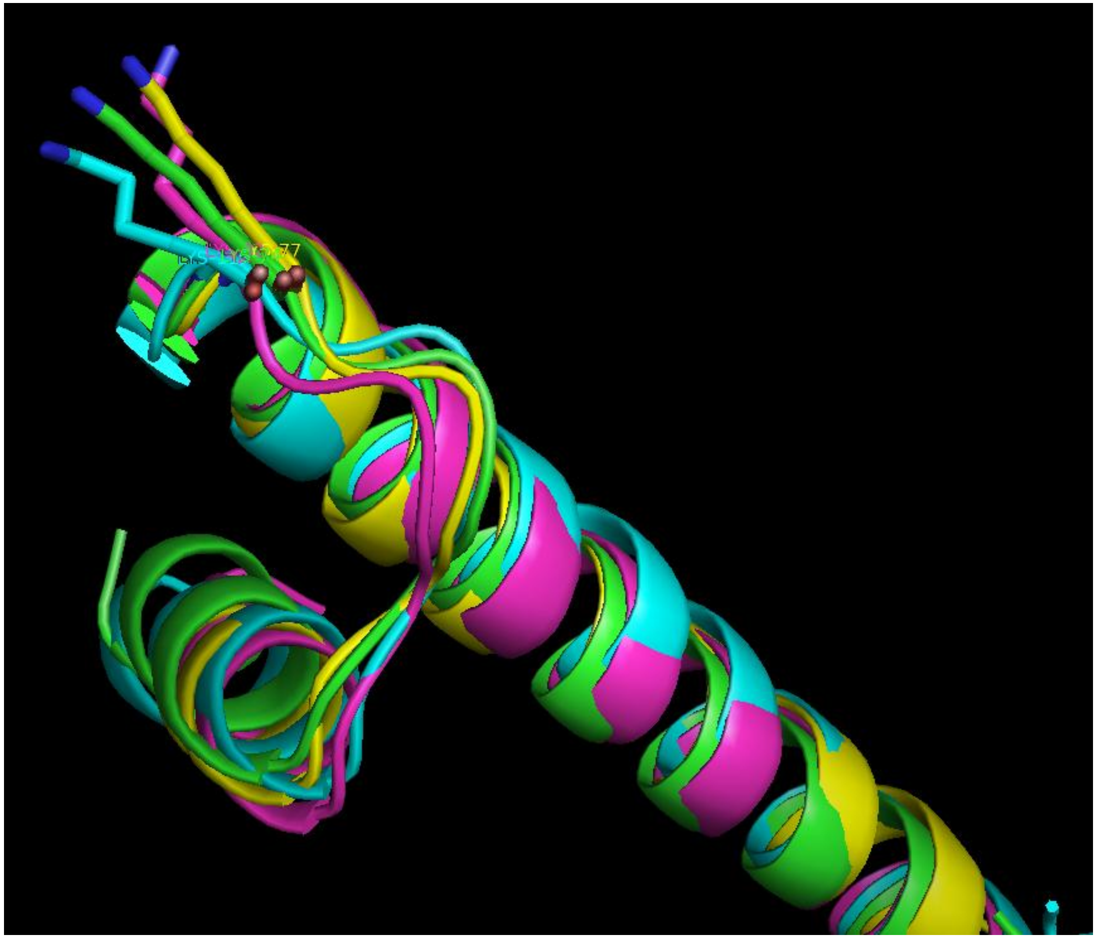


Figura20: STAMP. LYS.

Structure conservation

Lys49 and Ala61* are the only two very conserved residues

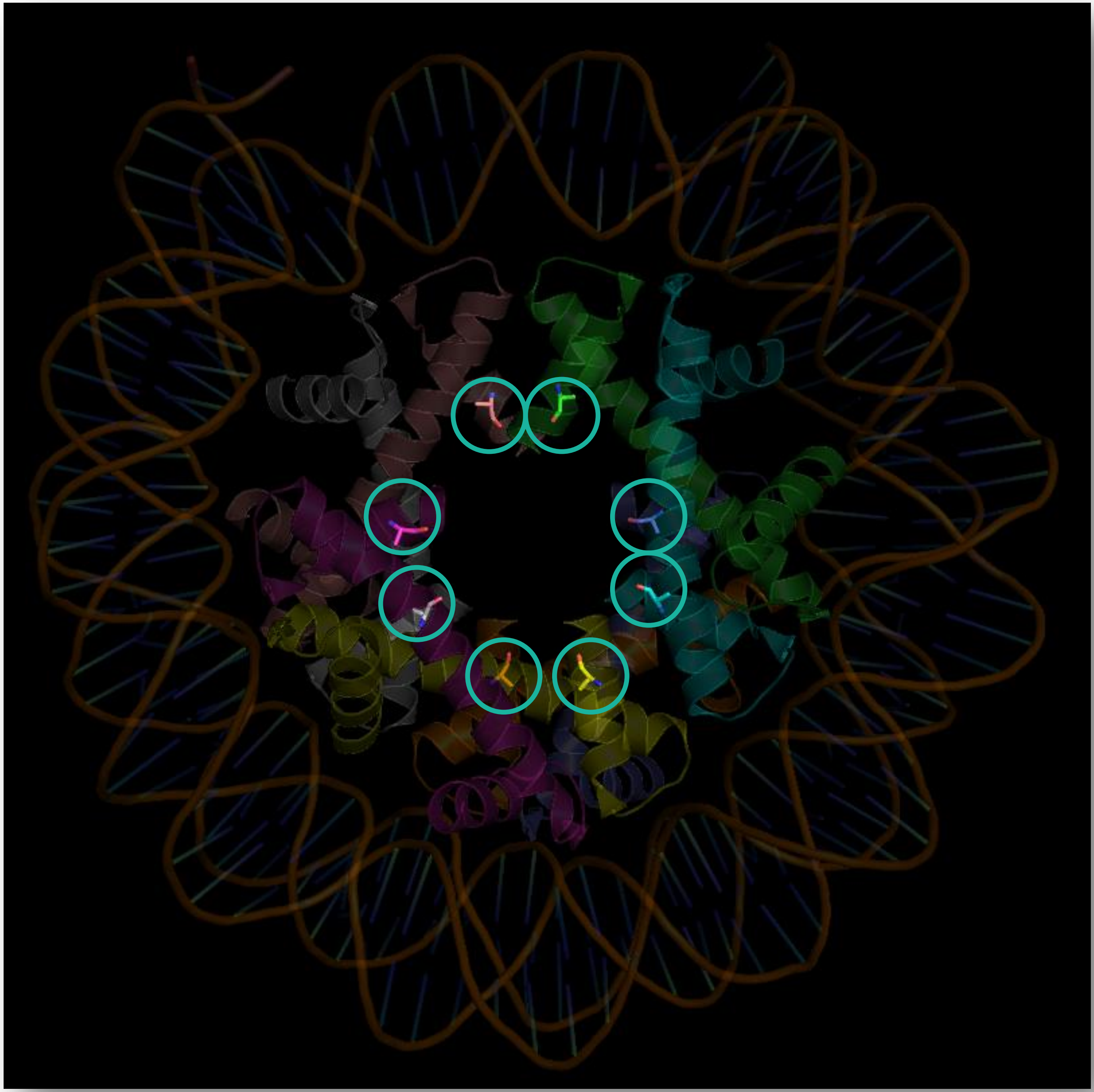


Figure21: Ala61* sequence conservation, shown in the nucleosome structure.

* Being 1 the first residue of each histone-fold sequence

Histone-fold in other proteins

Family:

1. Nucleosome core particles
2. Archeal histone
3. TBP-associated factors
4. Archeal histone-fold protein

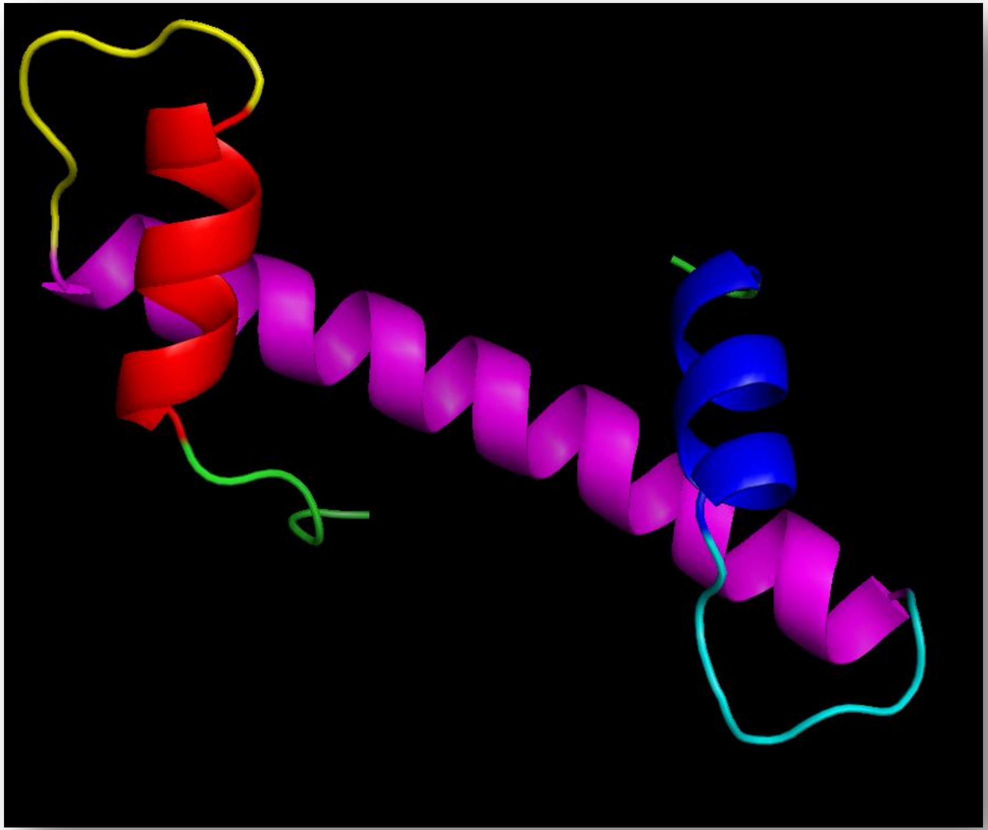


Figure22: Archeal histone A, *Methanothermus fervidus* (PDBid: 1HTA)

Histone-fold in other proteins

Family:

1. Nucleosome core particles
2. Archeal histone
3. TBP-associated factors
4. Archeal histone-fold protein

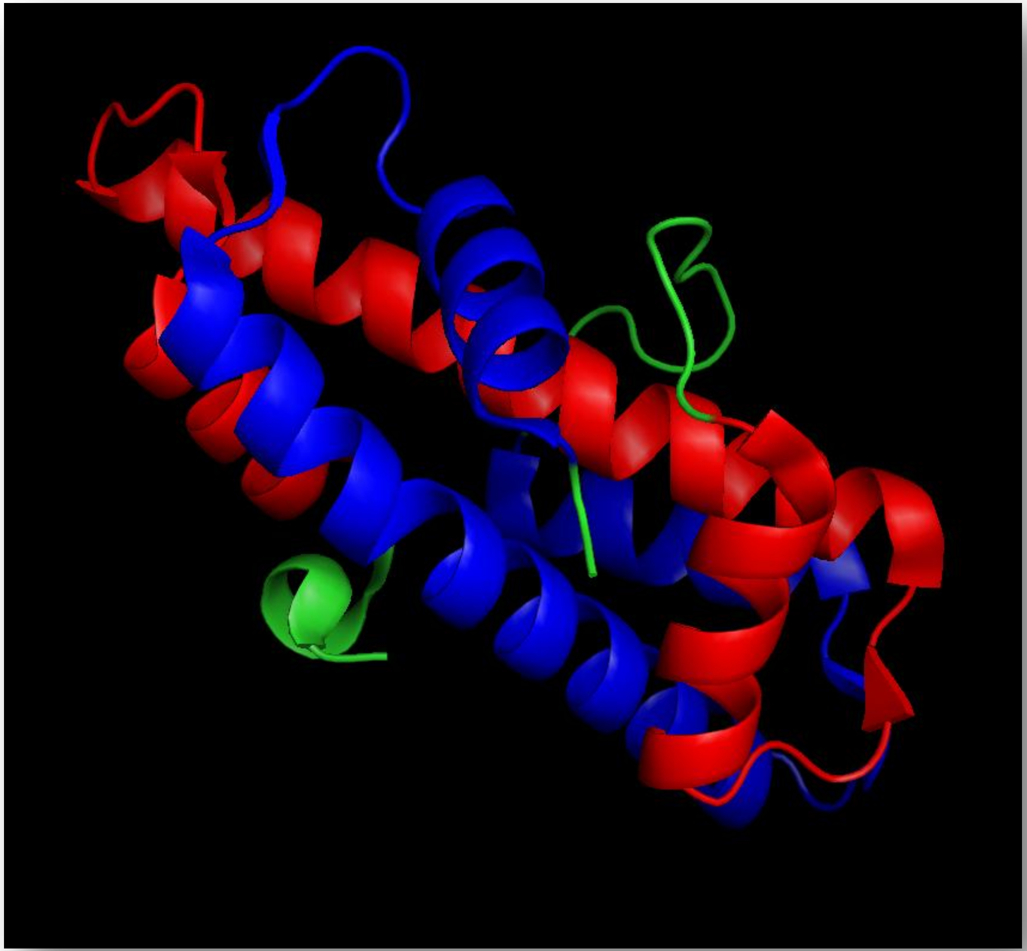


Figure23: Archeal histone-fold protein (PDBid: 1F1E)

Histone-fold in other proteins

Family:

1. Nucleosome core particles
2. Archeal histone
3. TBP-associated factors
4. Archeal histone-fold protein

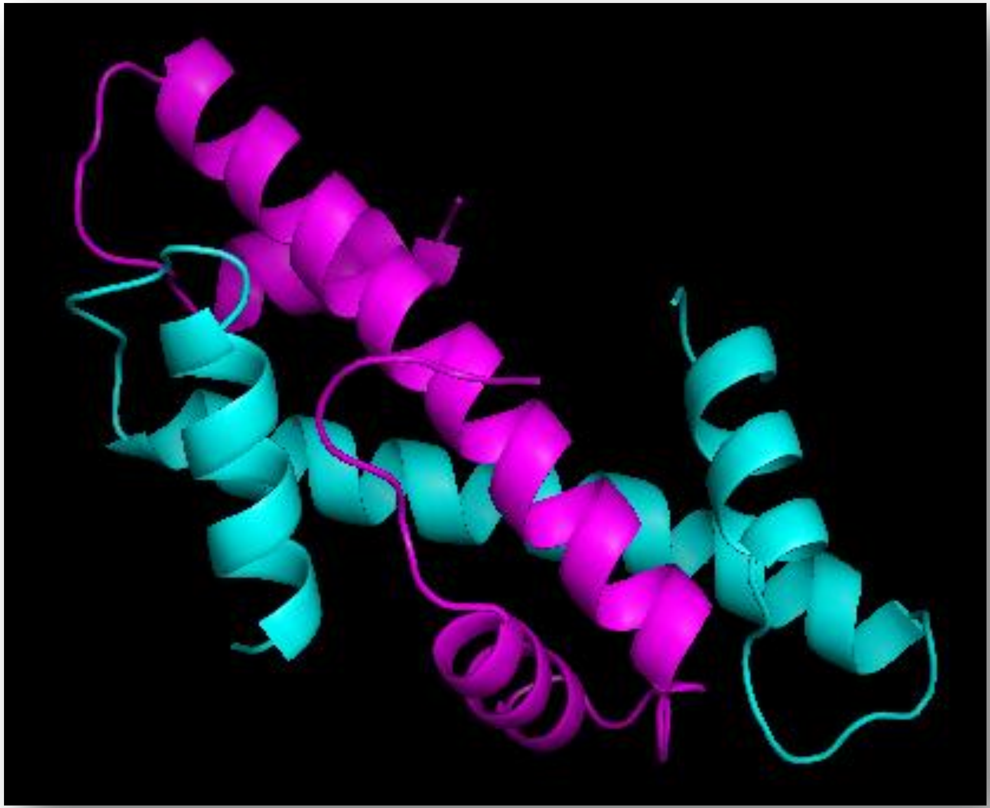


Figure24: TFIID TBP Associated Factor (PDBid: 3TAF)

TBP Associated Factor (TAFs)

TAF are important coactivatios an core promoter recognition factors.

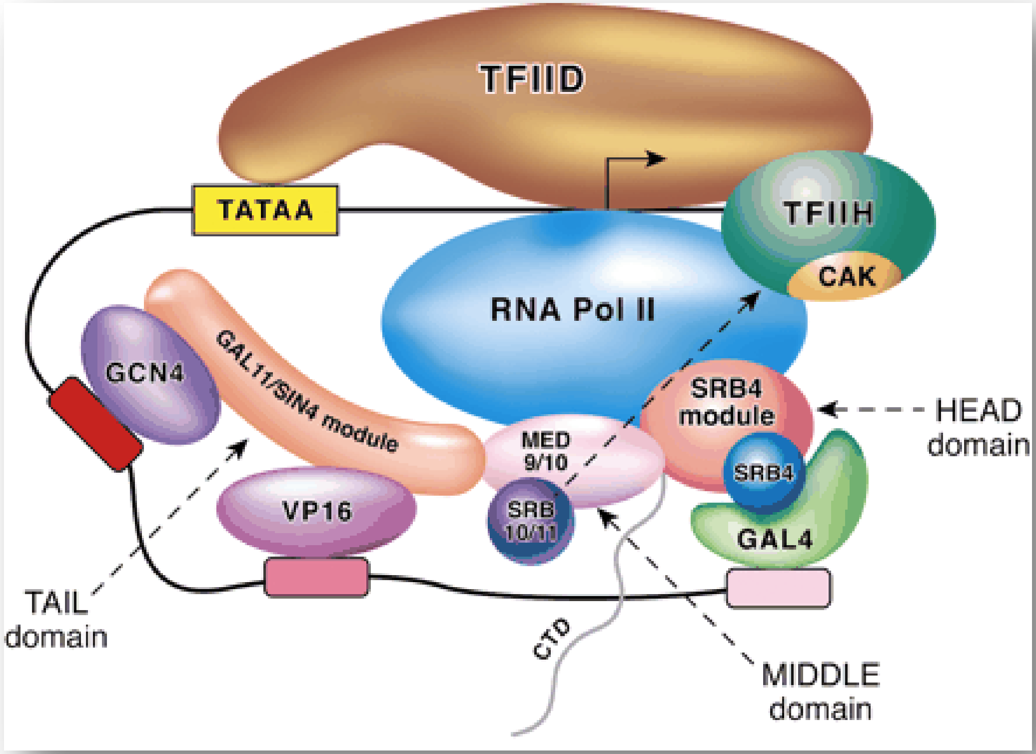


Figure25. Transcription complex.
Lewis B, Reinberg D. *The mediator coactivator complex: functional and physical roles in transcriptional regulation*. Journal of Cell Science. 2003; 116: 3667-75.

TBP Associated Factor (TAFs)

Figure26. STAMP. Four histone core with TAFs

Quaternary structure

- **Protein-protein interaction**
  - Histone dimerization
  - Histone tetramerization
  - Histone octamer
- **Protein-DNA interaction**
  - α1α1 motif
  - L1L2 motif

Protein-protein interaction: Dimerization

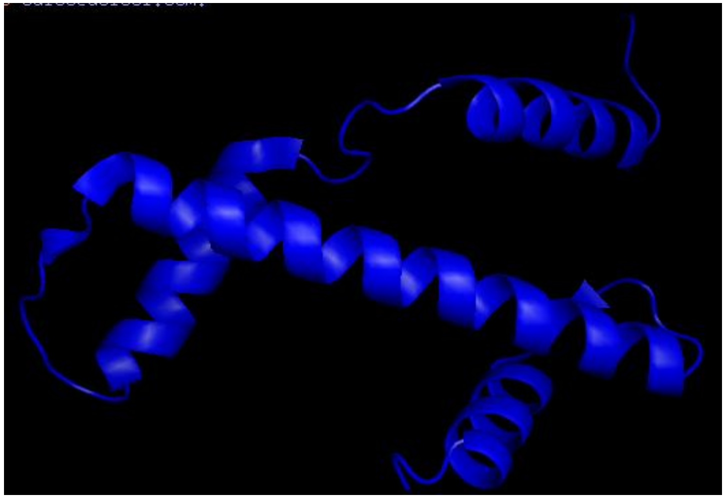


Figure26. Histone 3

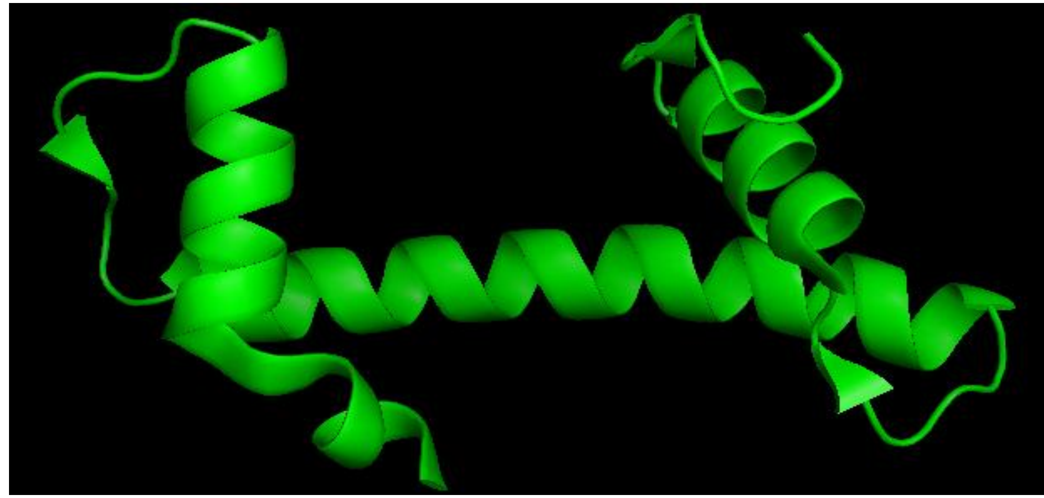
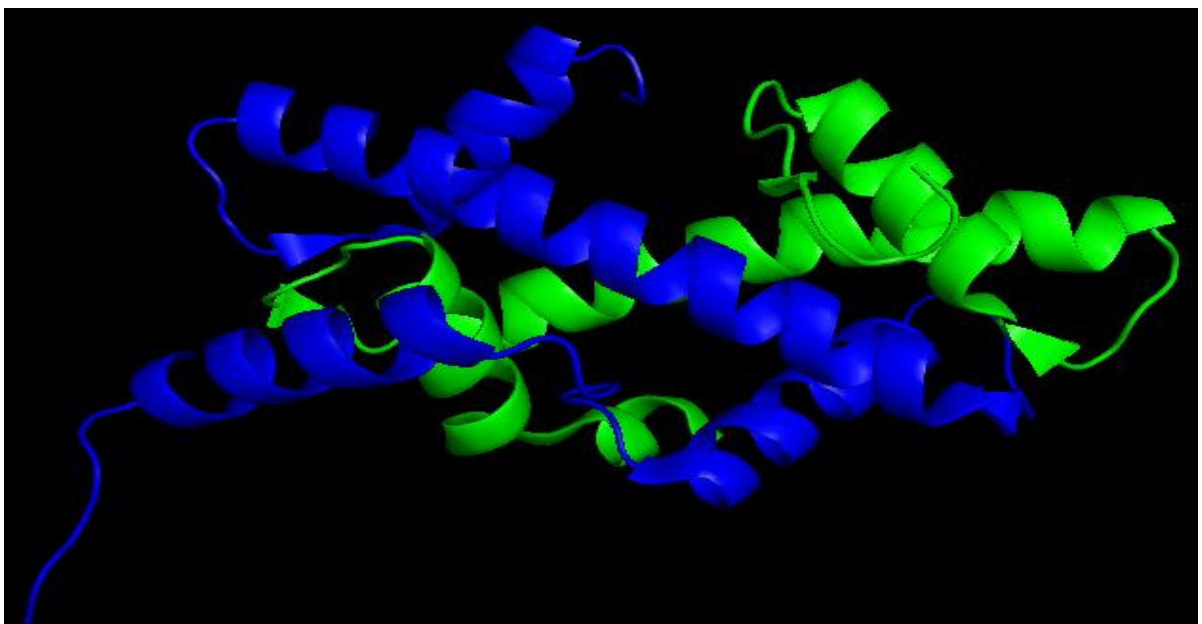


Figure27. Histone 4



H3
H4

Figure28. Dimer Histone 3-Histone 4

Protein-protein interaction: Dimerization

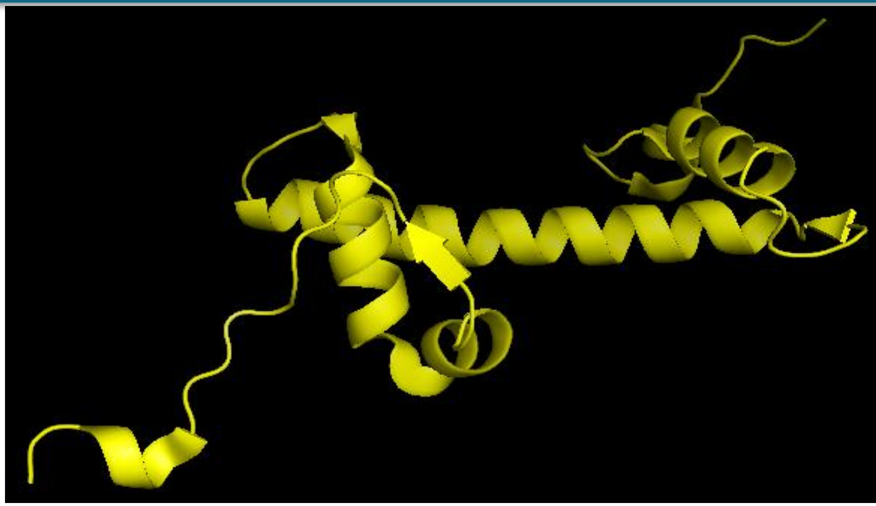


Figure29 Histone H2A

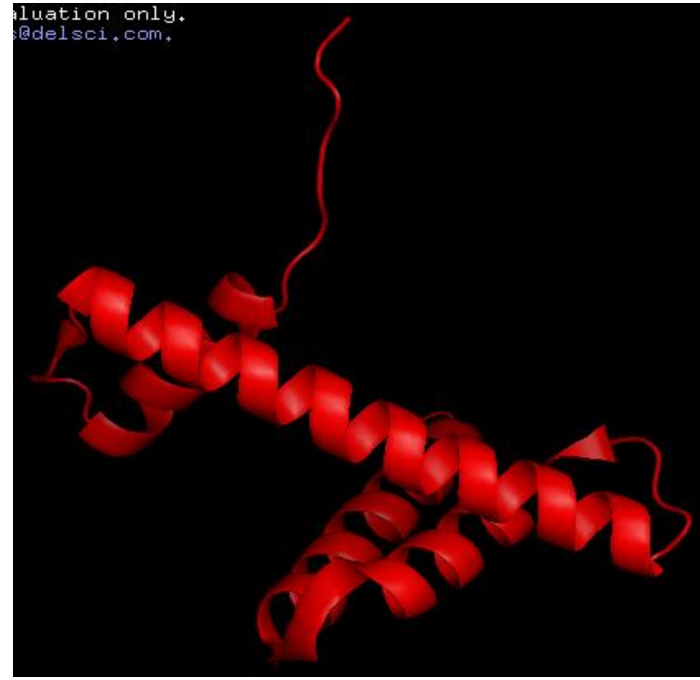
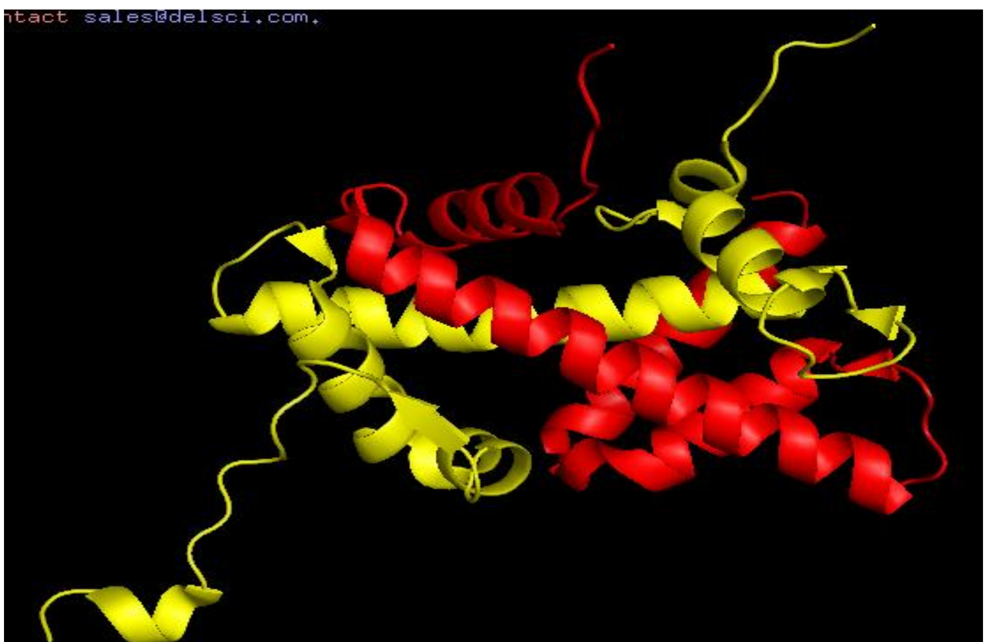


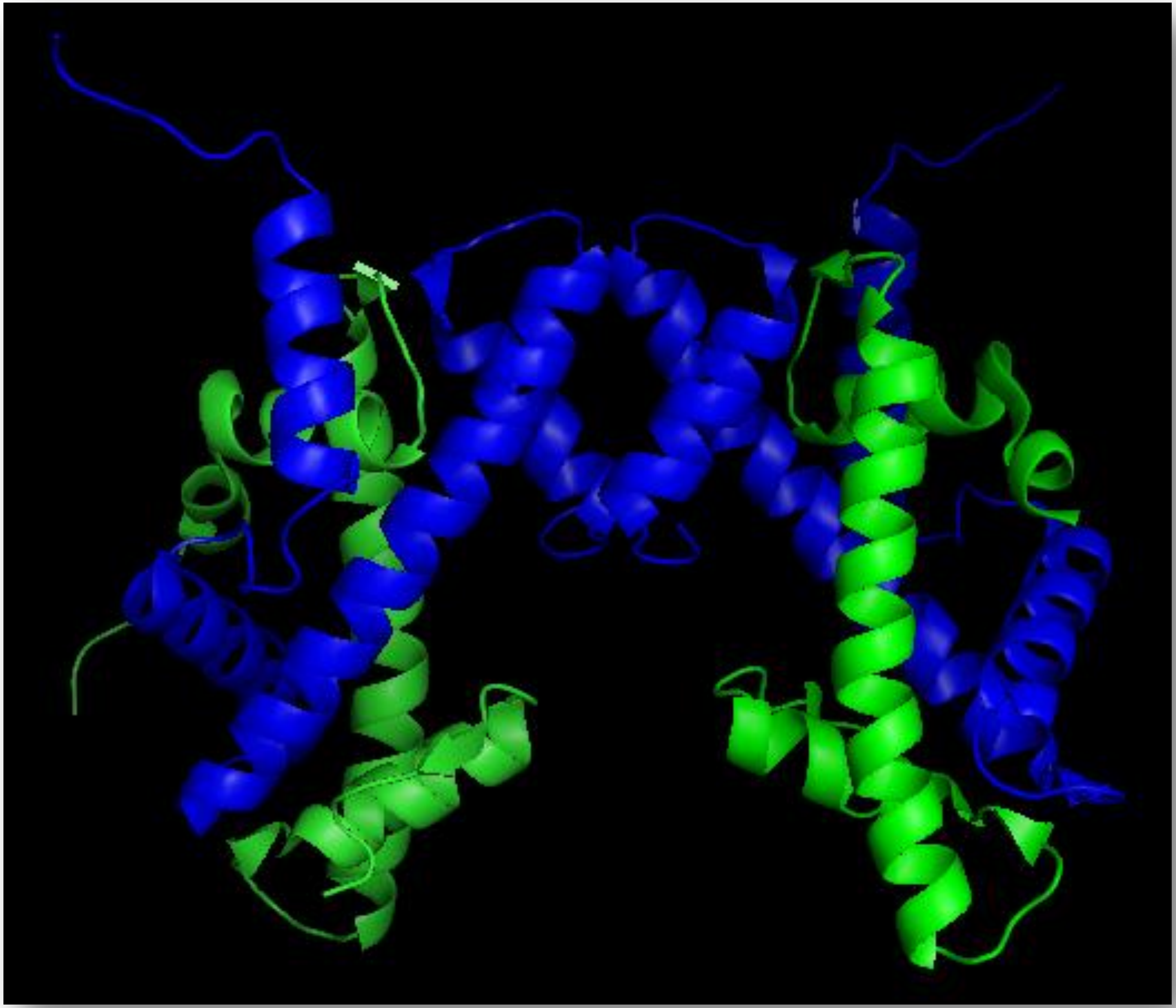
Figure29 Histone H2B



H2A
H2B

Figure31. DImer Histone H2A-Histone H2B

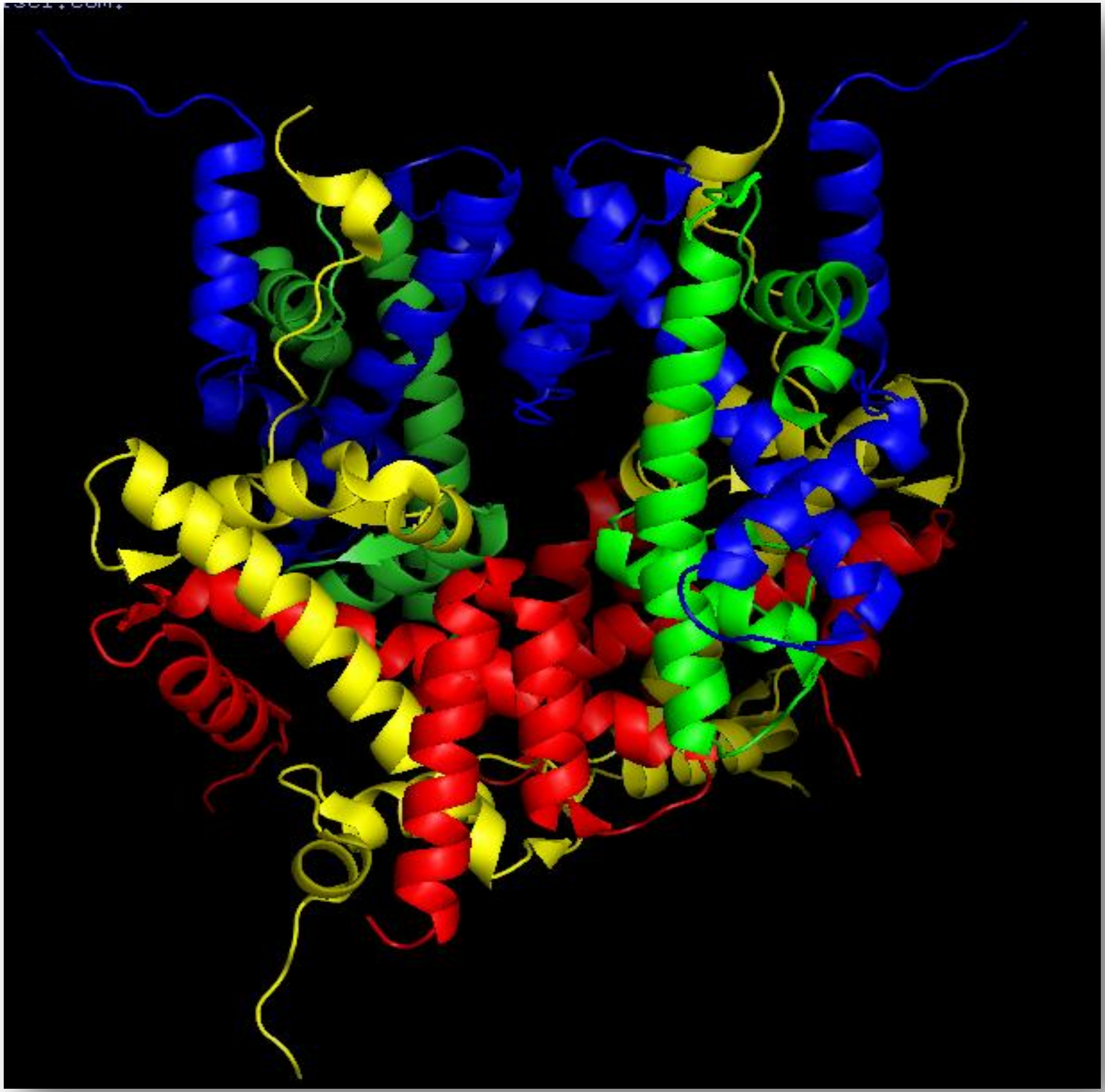
Protein-protein interaction: Tetramer



H3
H4

Figure31. Tetramer

Protein-protein interaction: Histone octamer

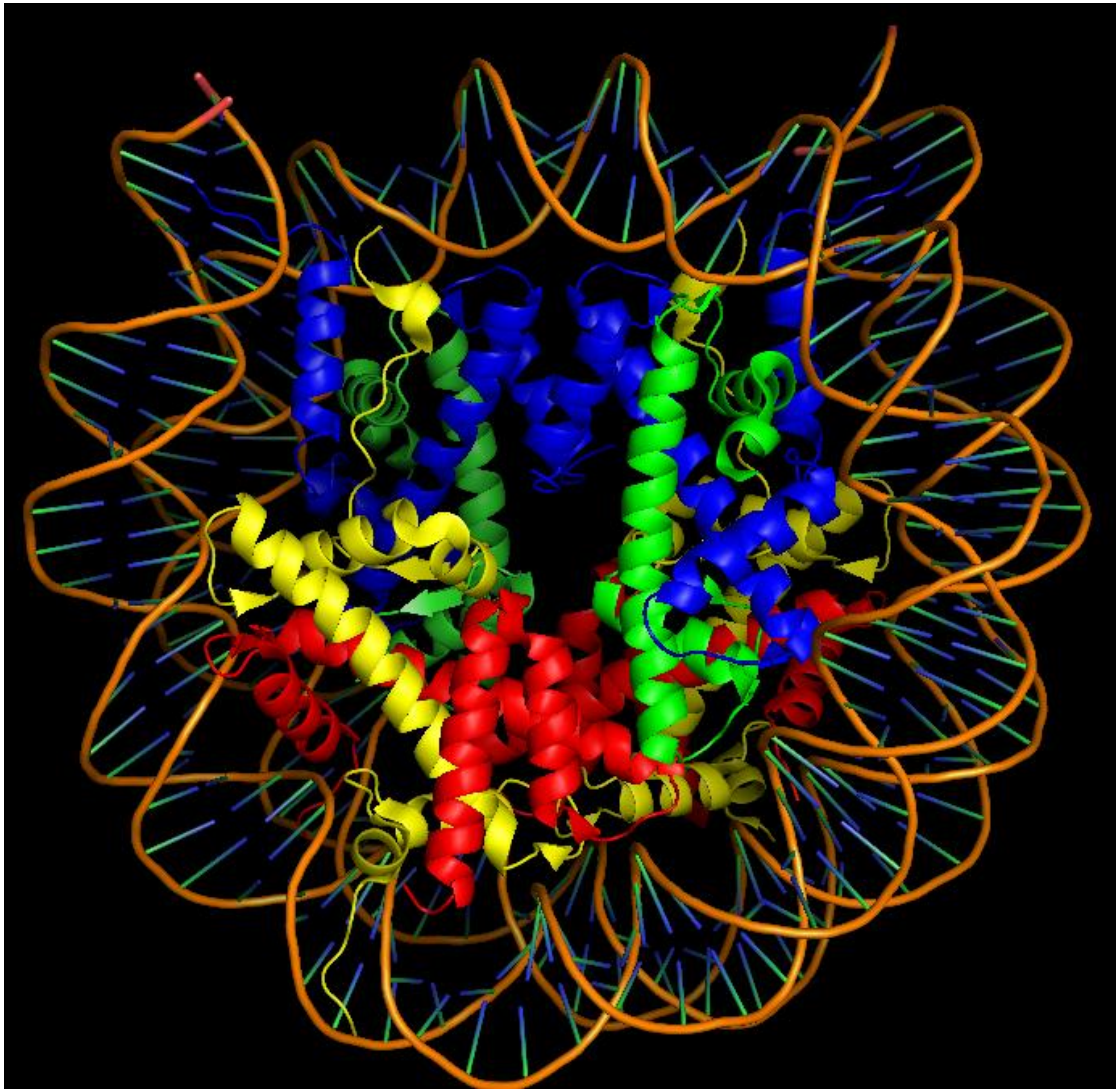


H3
H4

H2A
H2B

Figure32. Octamer

Protein-protein interaction: Nucleosome



DNA

H3
H4

H2A
H2B

Figure33. Nucleosome

Protein-protein interaction: Dimerization

The H3 and H4 histones use L1-L2 motif

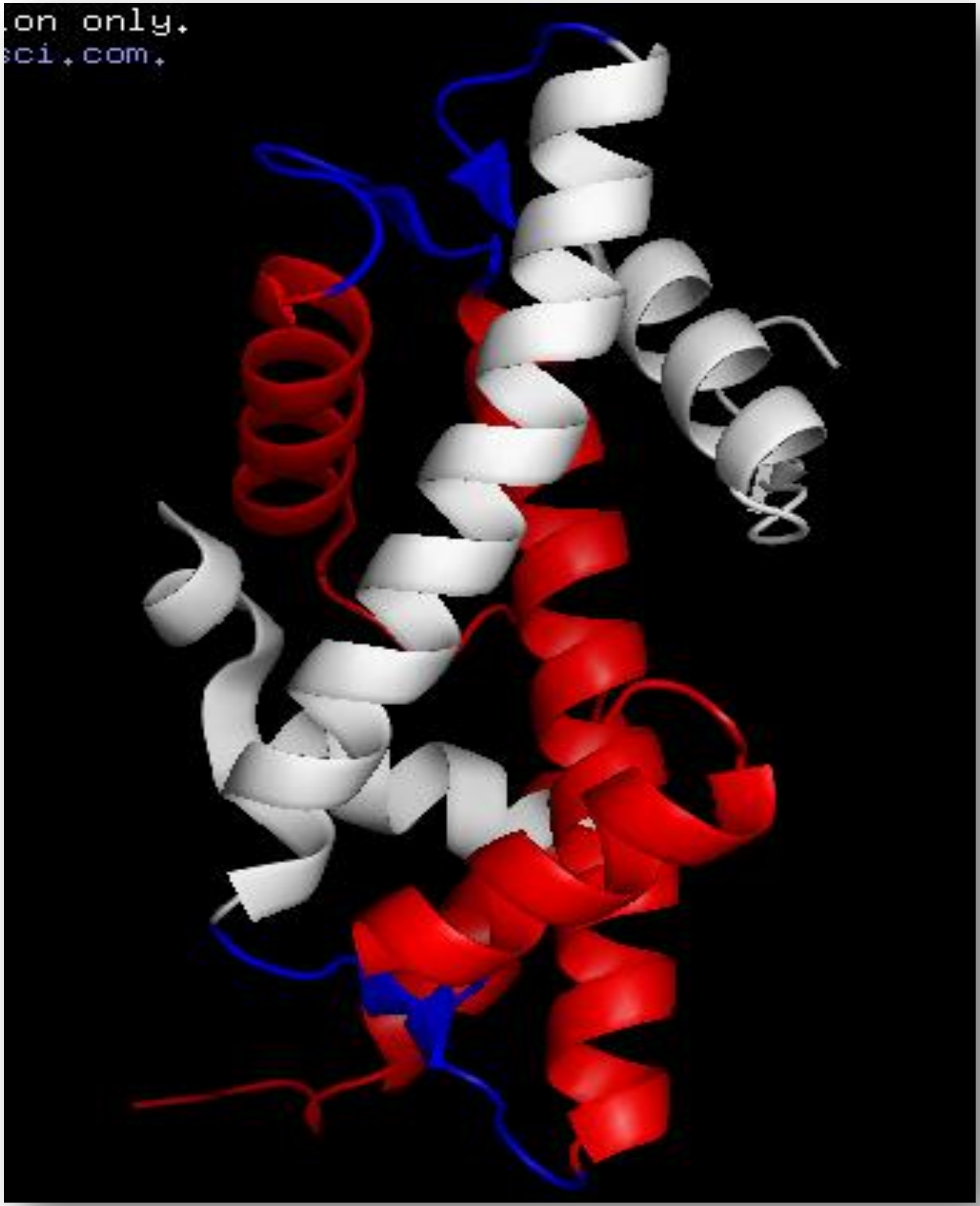


Figure34. Dimer H3-H4

Protein-protein interaction: Dimerization

There are several hydrophobic interaction. H3-H4

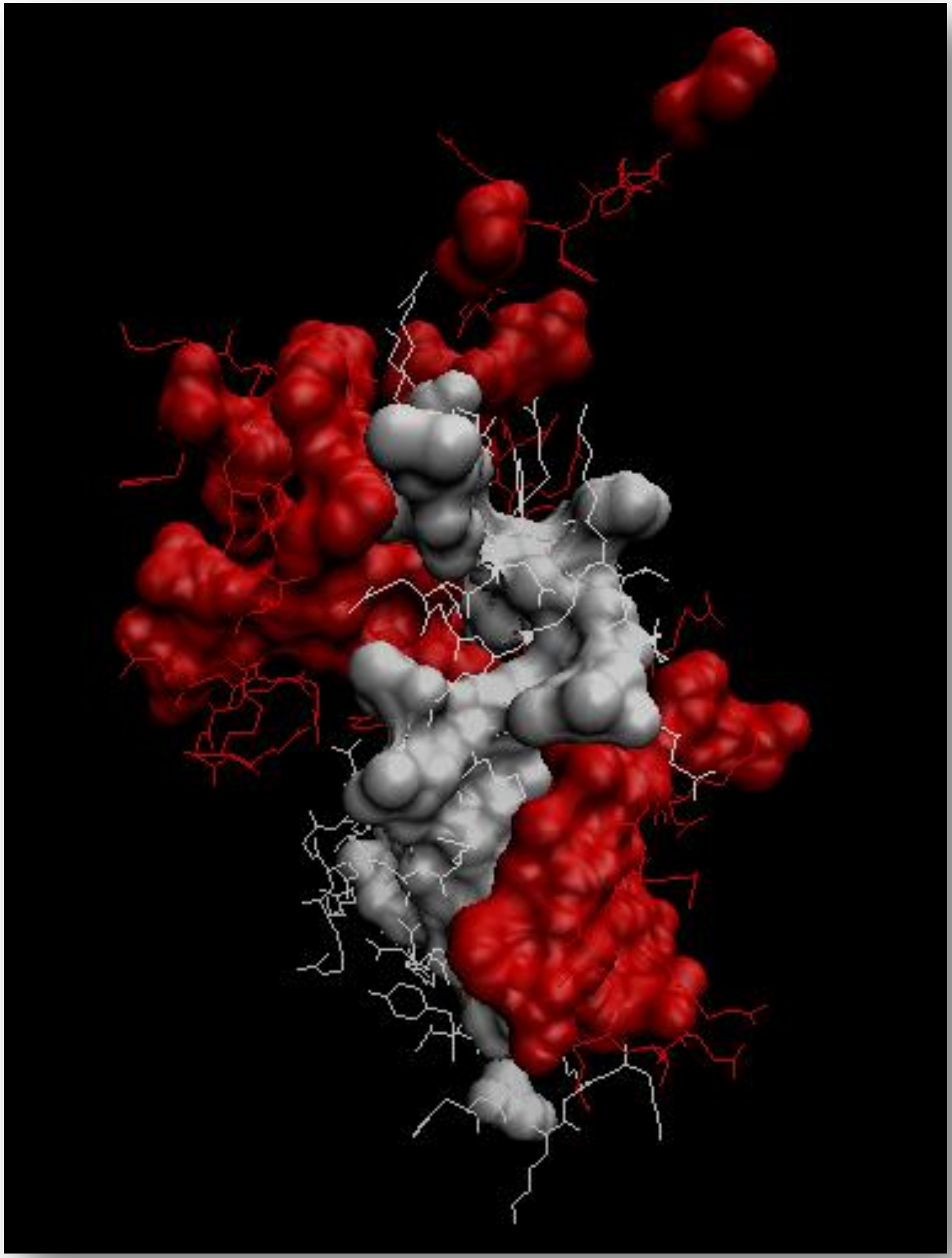


Figure35. Hydrophobic forces in dimer H3-H4

Protein-protein interaction: Dimerization

H4/R45(SC)-**H3/V117(MC)**
H4/R45(SC)-**H3/I119(SC)**

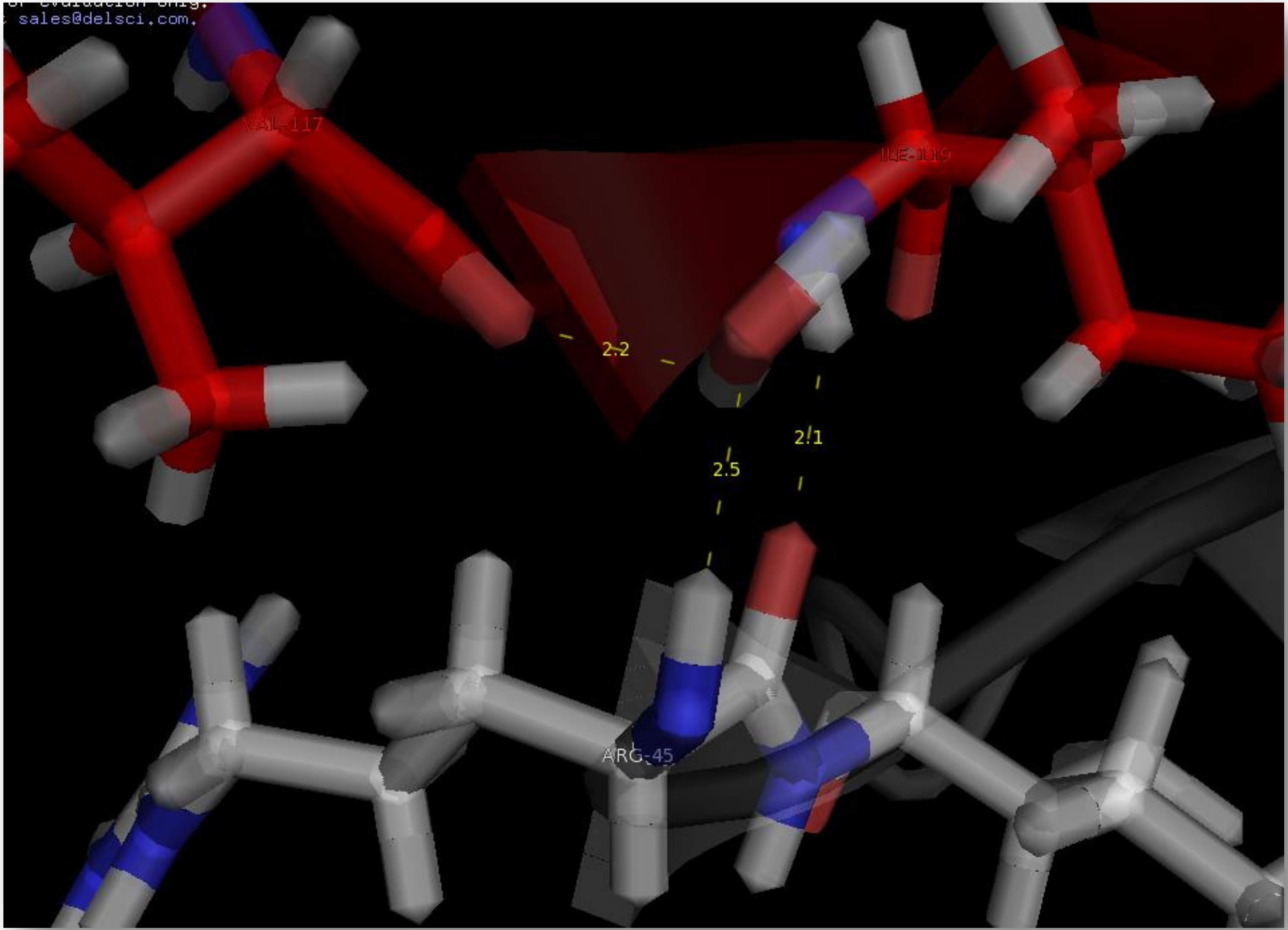


Figure 36. Dimer H3-H4. Interaction in the L1L2 motif.

Protein-protein interaction: Dimerization

H4/K79(MC)-H3/R83(SC)
H4/V81(MC)-H3/R83 (SC)

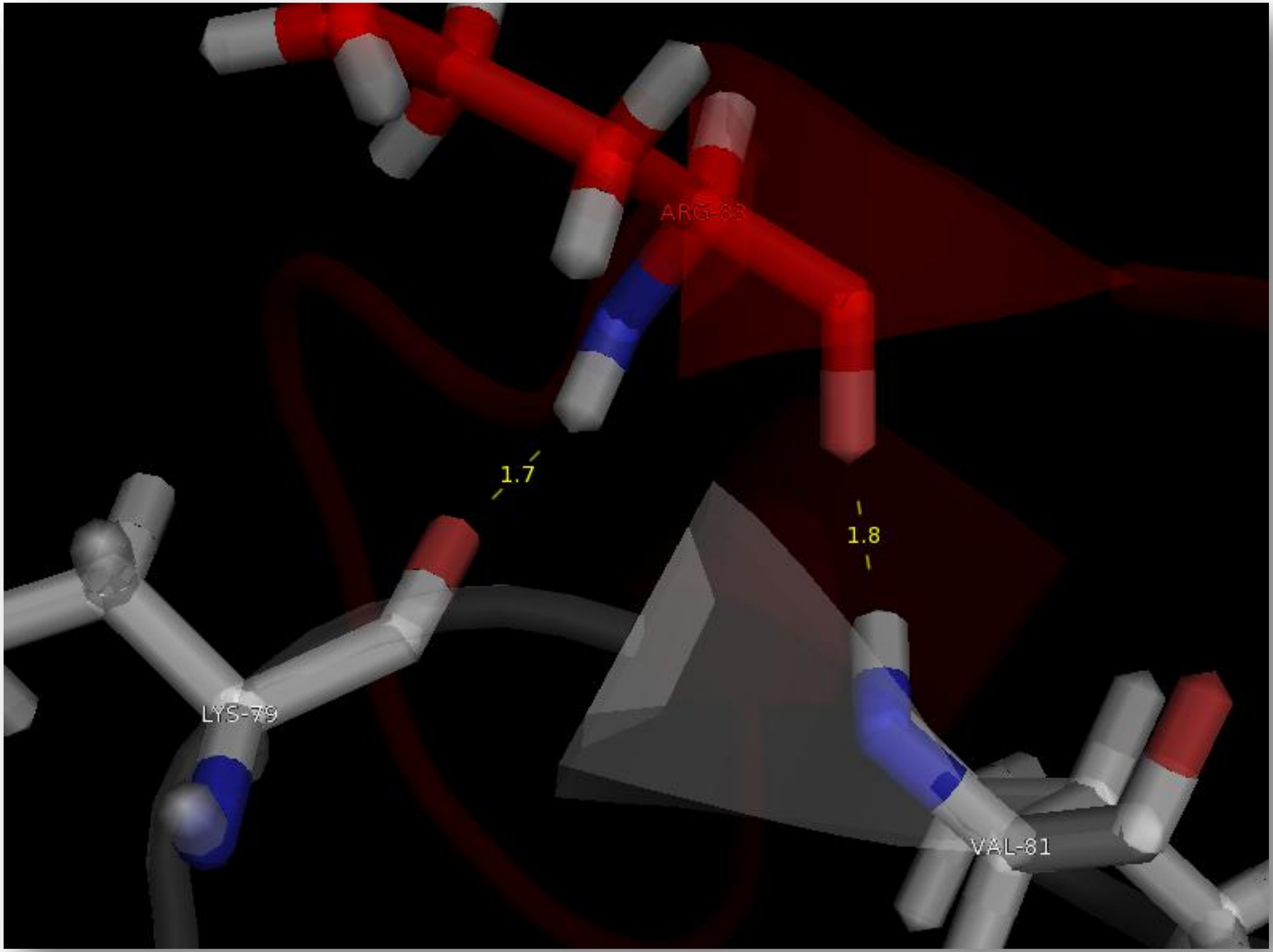


Figure 37. Dimer H3-H4. Interaction in the L1L2 motif.

Protein-protein interaction: Dimerization

The H2A and H2B use the same motif L1L2

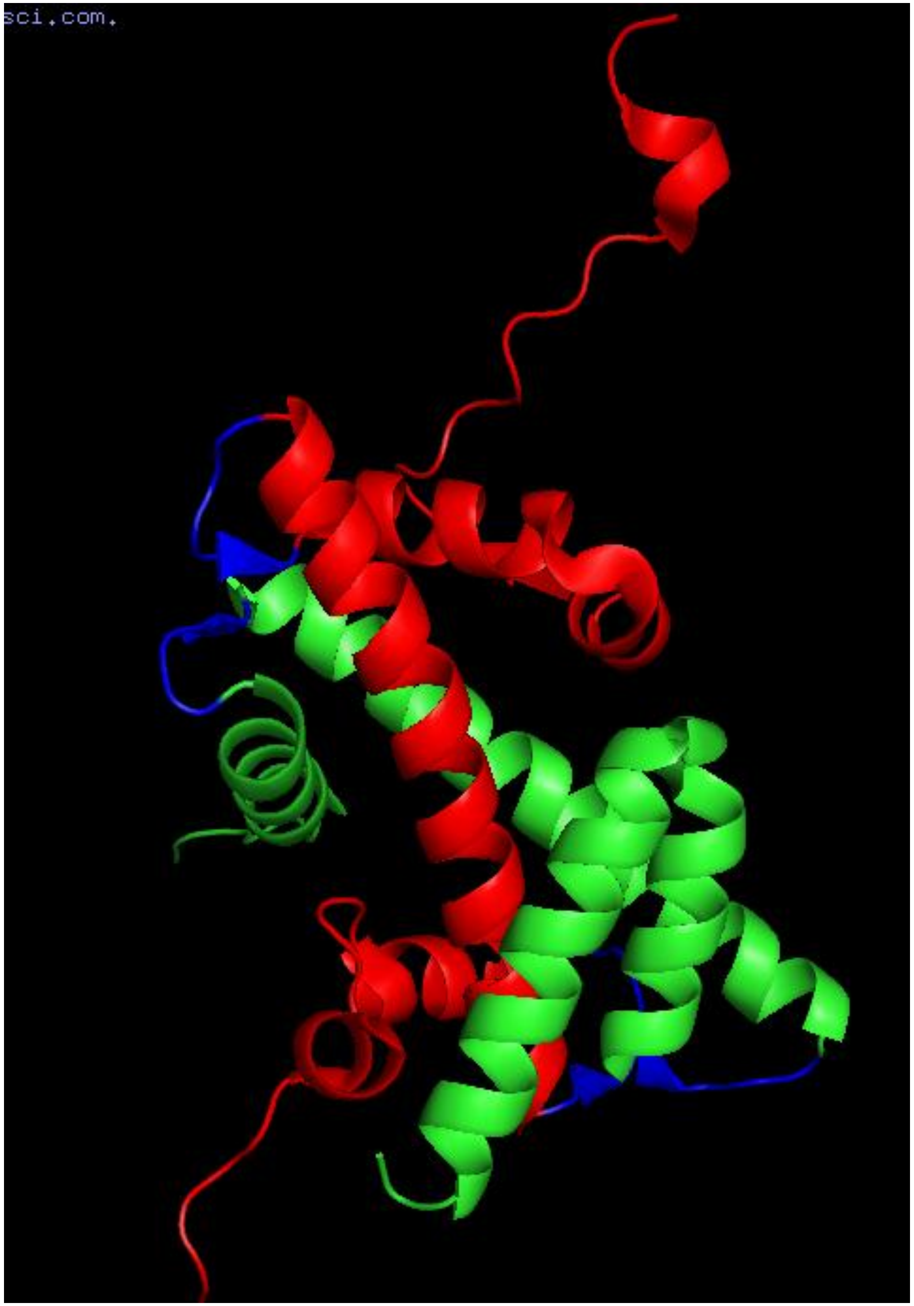


Figure38. Dimer H2A-H2B

Protein-protein interaction: Dimerization

There are several hydrophobic interaction. H2A, H2B

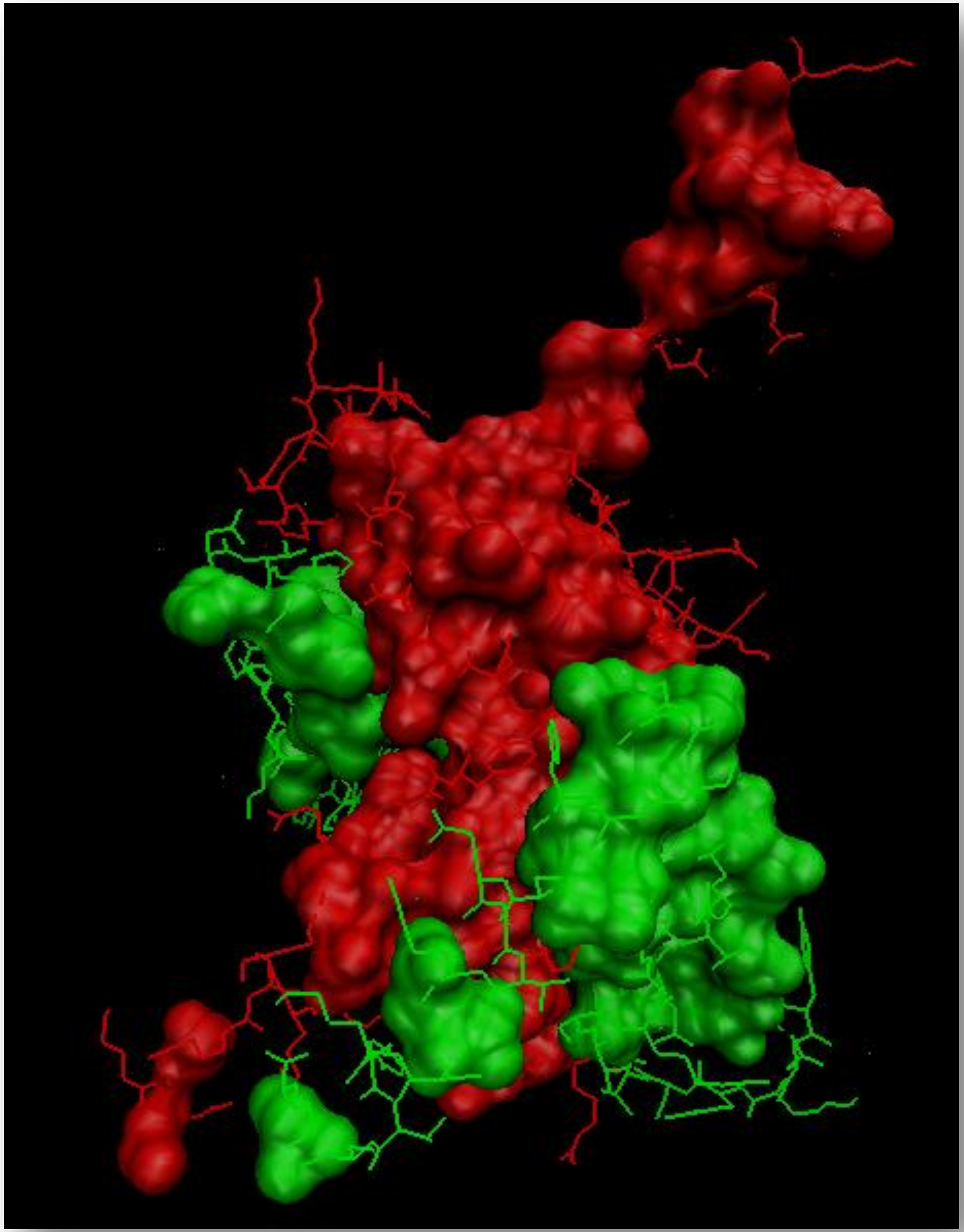


Figure39. Hydrophobic forces in dimer H3-H4

Protein-protein interaction: Dimerization

H2A/R42(SC)-H2B/I86(SC),
H2A/R42(SC)-H2B/T85(MC),
H2A/G44(MC)-H2B/I86(MC)

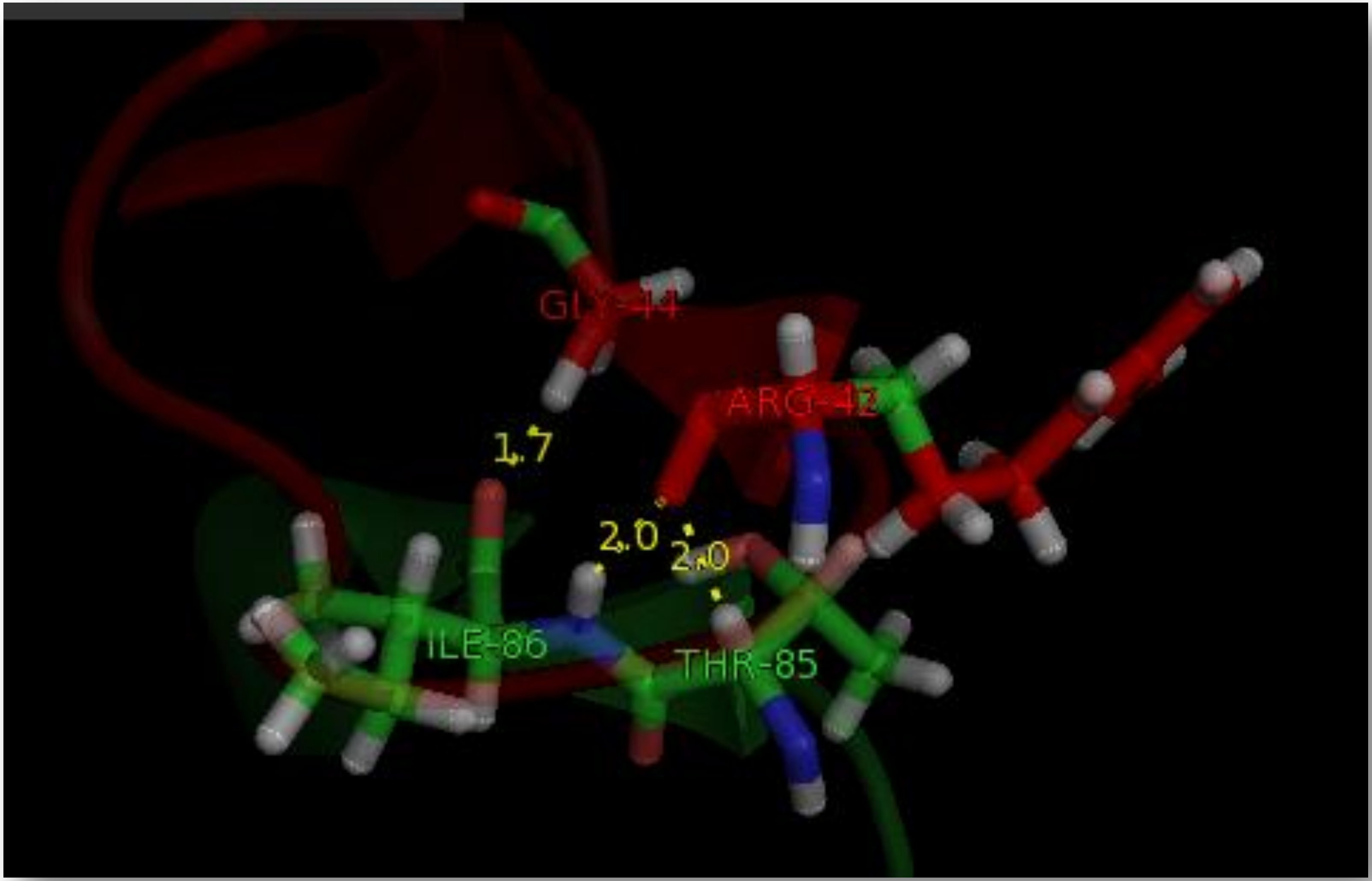
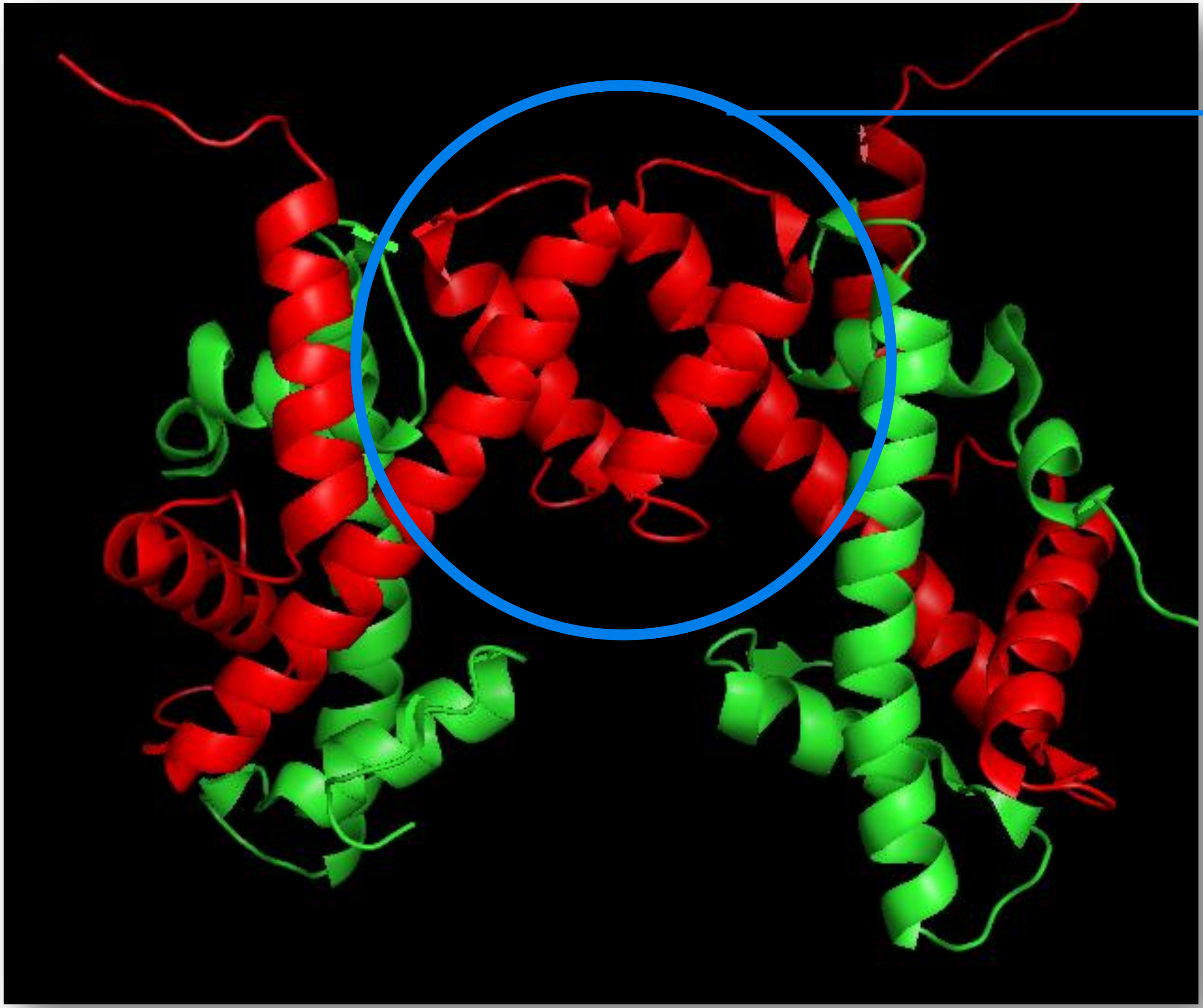


Figure40. Dimer H2A-H2B. Interaction in the L1L2 motif.

Protein-protein interaction: Tetramer

Two H3-H4 dimers are arranged in a tetramer via a four-helix bundle



H3-H3’ 4-HELIX BUNDLE STRUCTURE

Figure41. Tetramer

Protein-protein interaction: Tetramer

Several hydrophobic interactions. H3-H3'

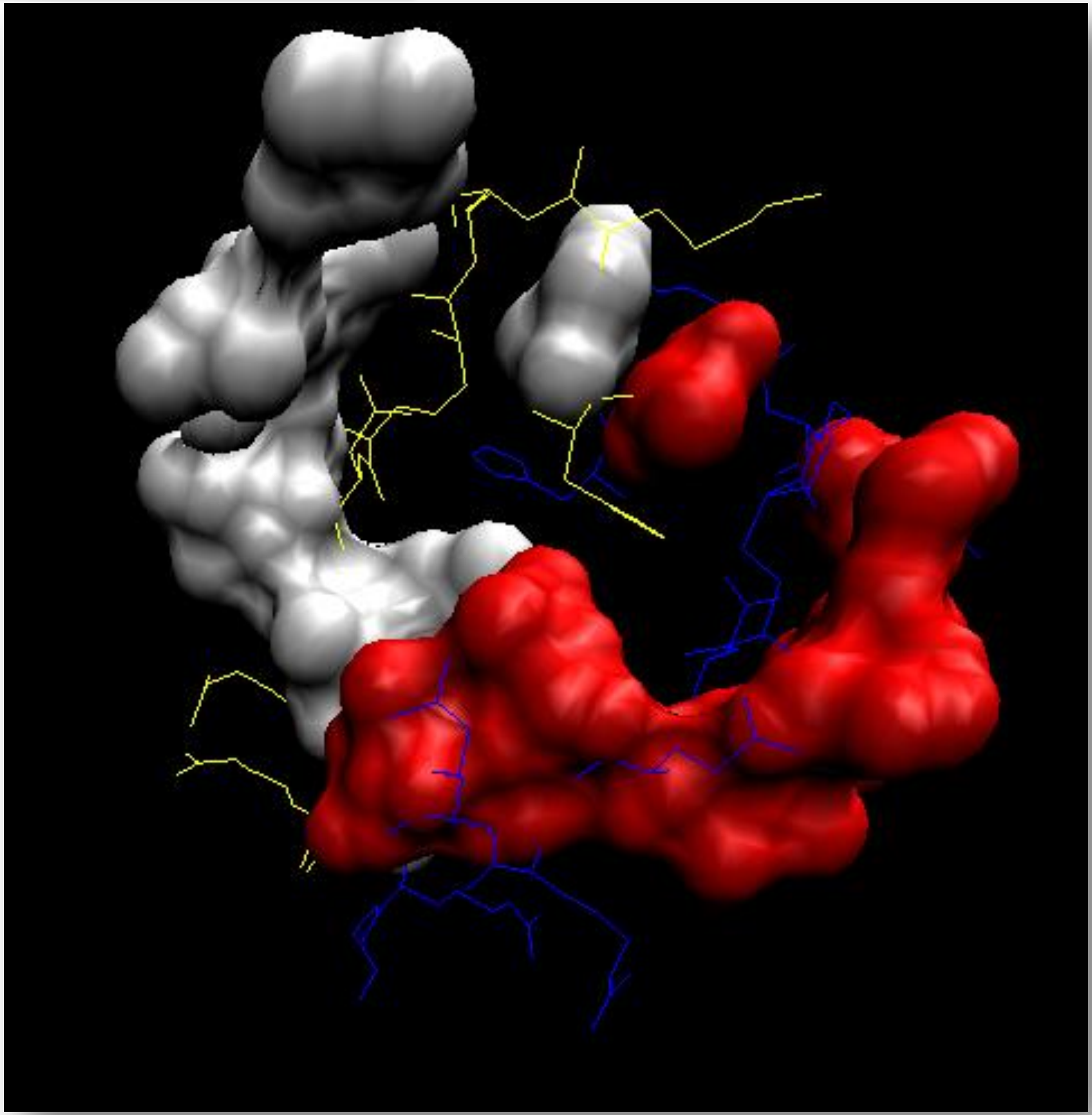


Figure42. Hydrophobic forces in dimer H2B-H2A

Protein-protein interaction: Tetramer

H3/D123(SC)-H3'/H113(SC), H3/H113(SC)-H3'/D123(SC)

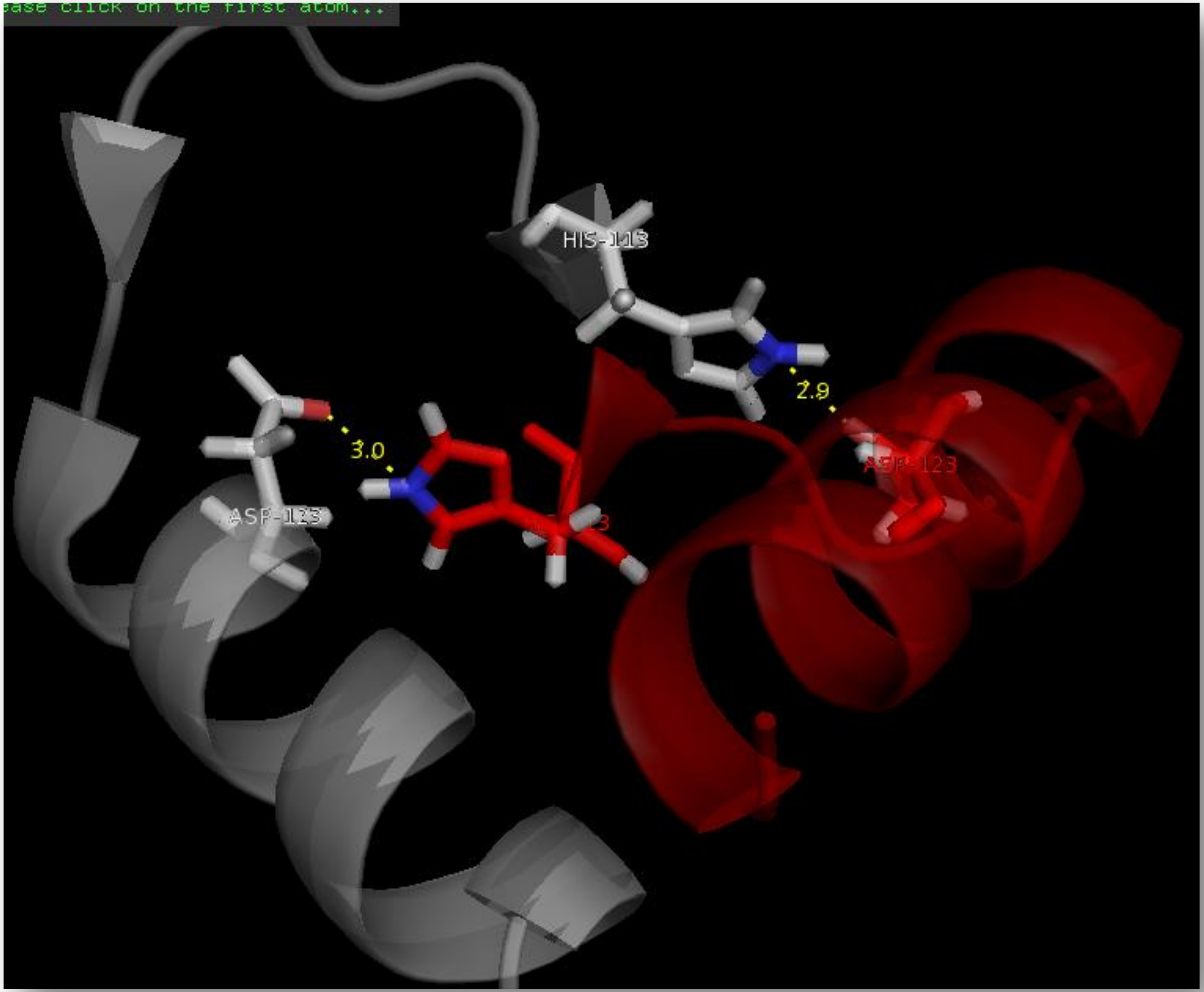


Figure43. Dimer H2A-H2B.

Protein-protein interaction: Histones octamer

•Single H3'-H3 association to form the tetramer and two H2B-H4 associations to produce the octamer.

Figure44. Histone Octamer.

Protein-protein interaction: Histones octamer

H4, H2B

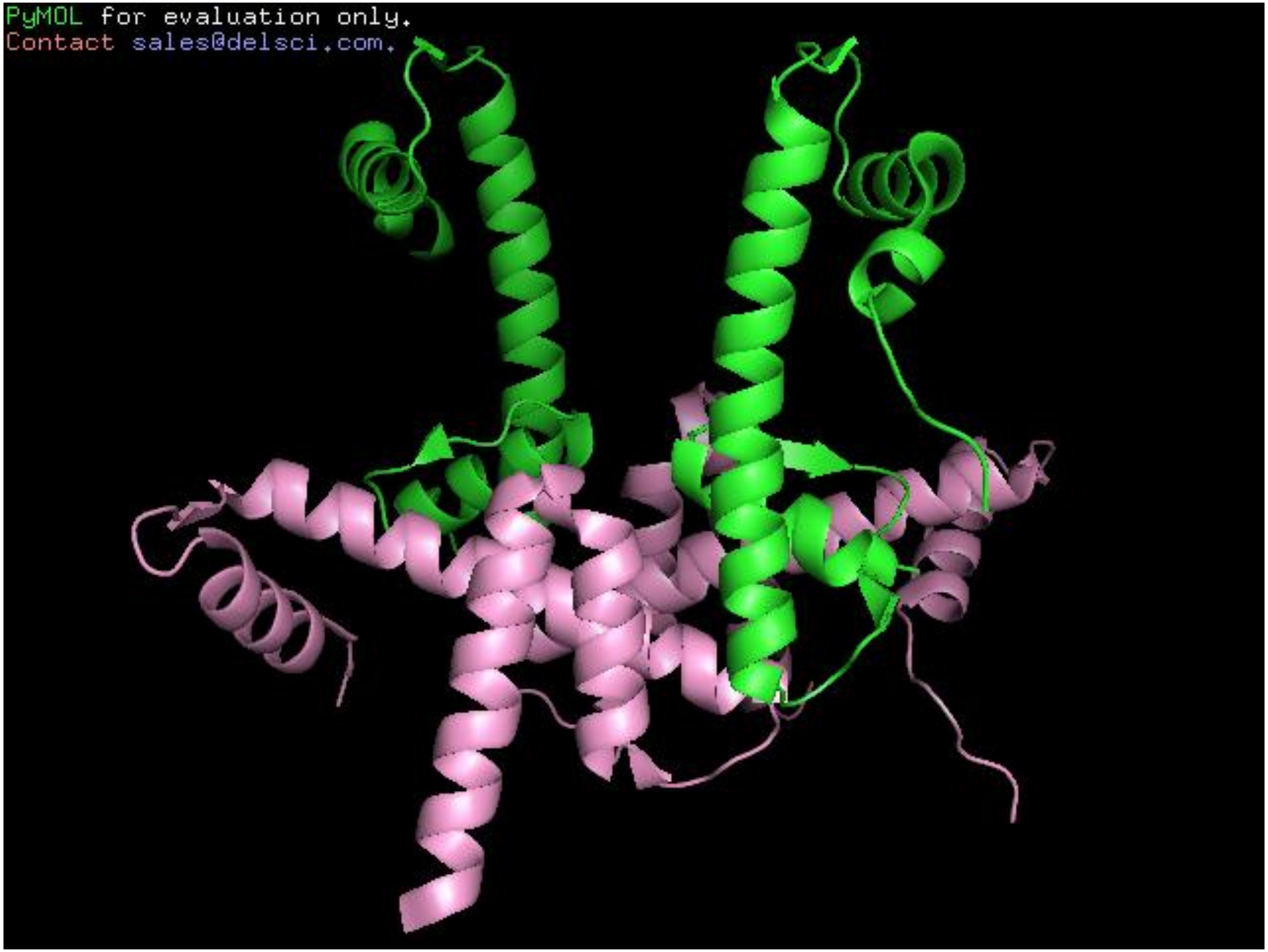


Figure 45 . Histone octamer. H2B-H4

Protein-protein interaction: Histones octamer

Several hydrophobic bonds. H4, H2B

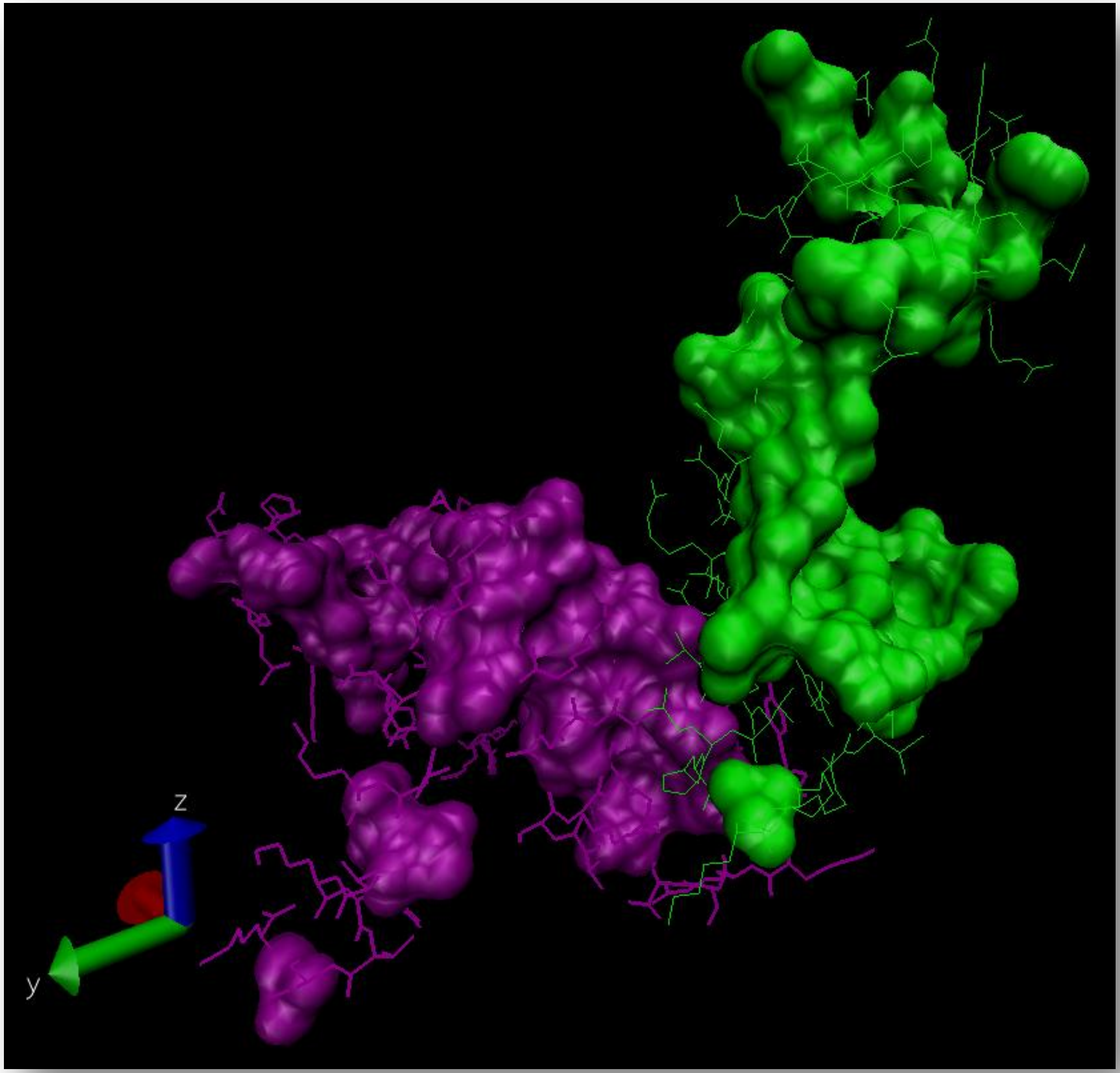


Figure46. Hydrophobic forces in H2B-H4

Protein-protein interaction: Histones octamer

H4/H75(SC)-H2B/E90(SC)

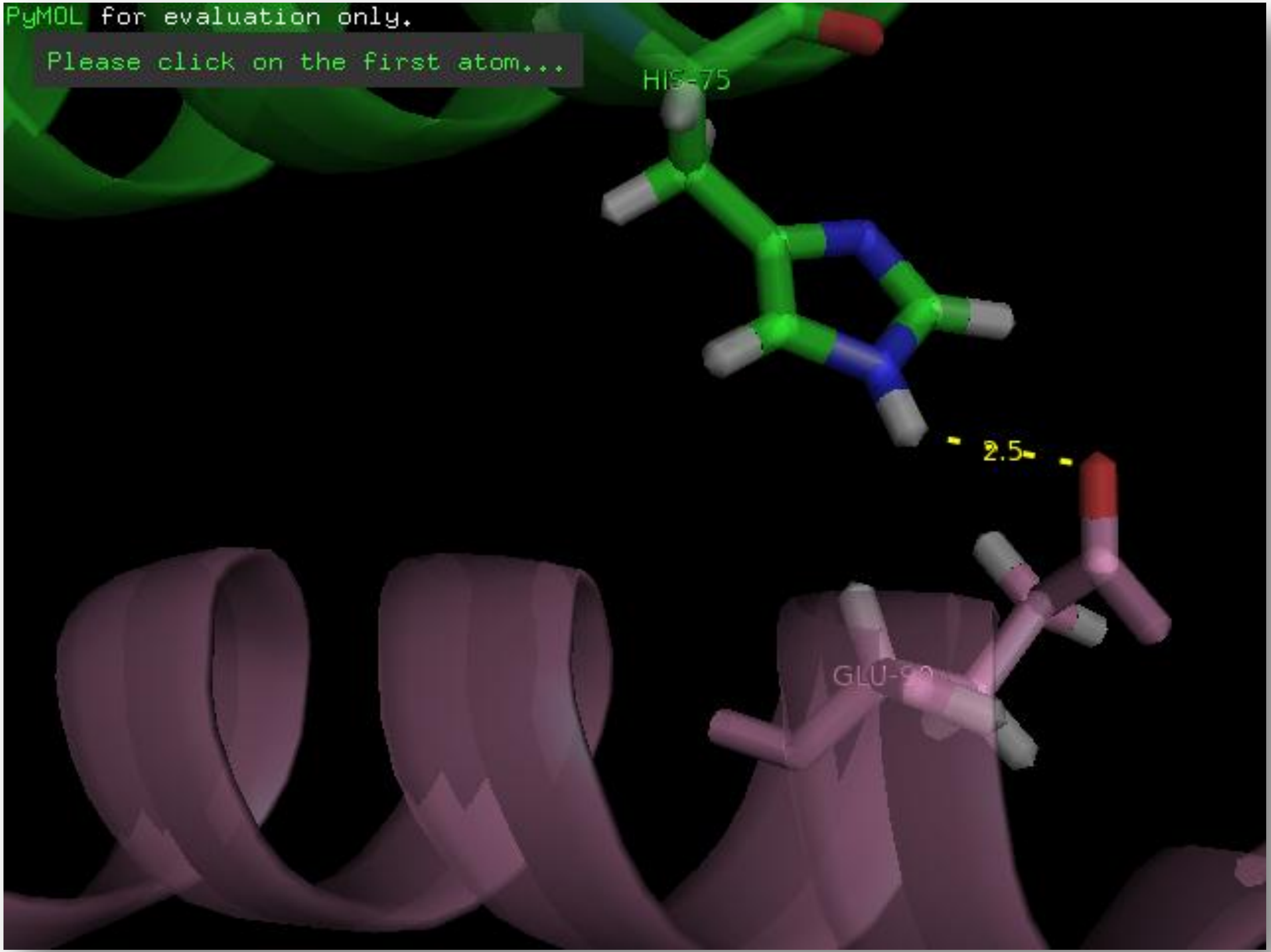


Figure47. H2B-H4. Intermolecular interaction

Protein-protein interaction: Histones octamer

H4/D68(SC)-H2B/L97(SC),
H4/R92(SC)-H2B/E73(SC),
H4/Y72(SC)-H2B/E73(SC)

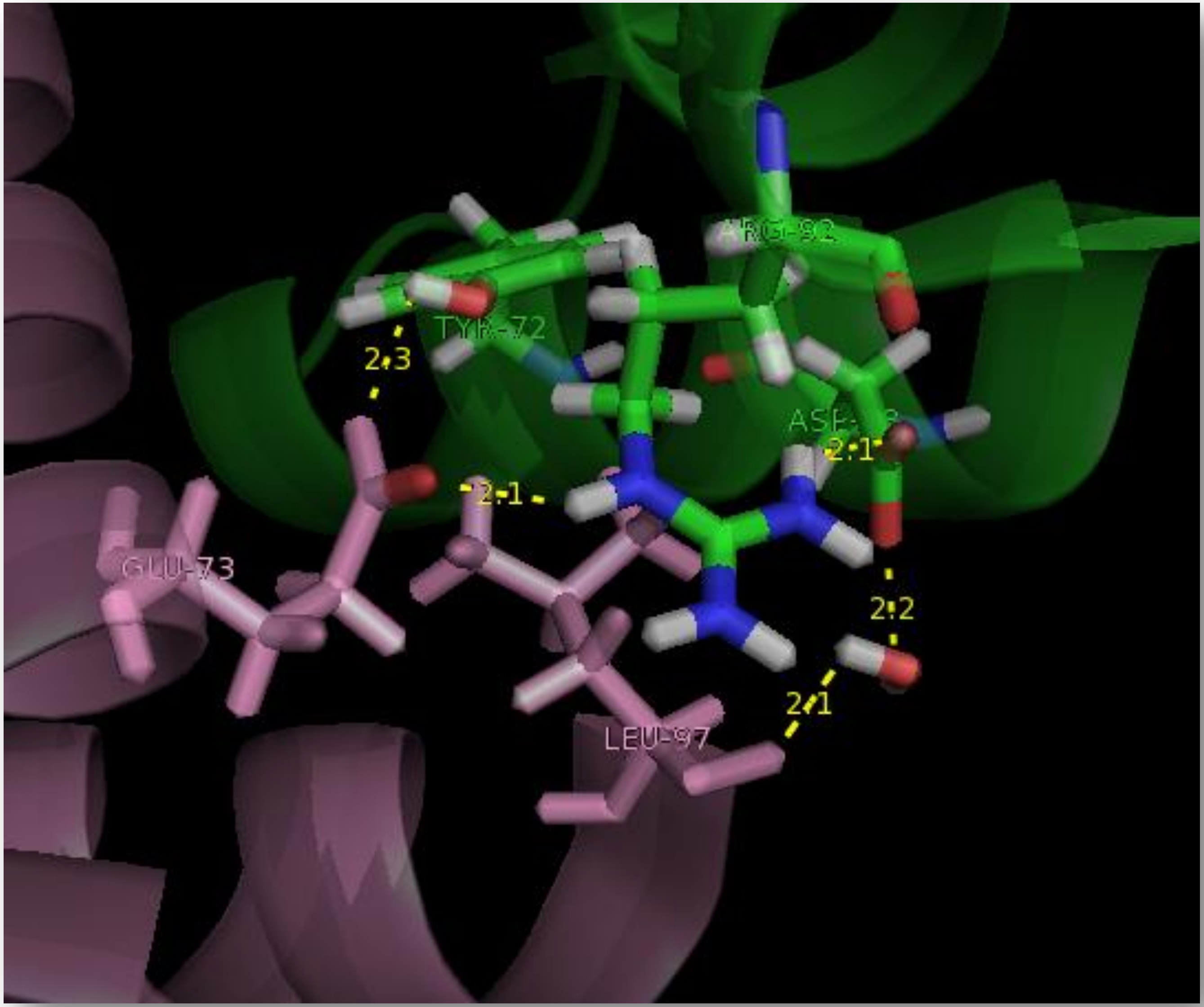


Figure48. H2B-H4. Intermolecular interaction

Protein-protein interaction: Histones octamer

Hydrophobic cluster: H4/Y72, H4/Y88, H2B/H80

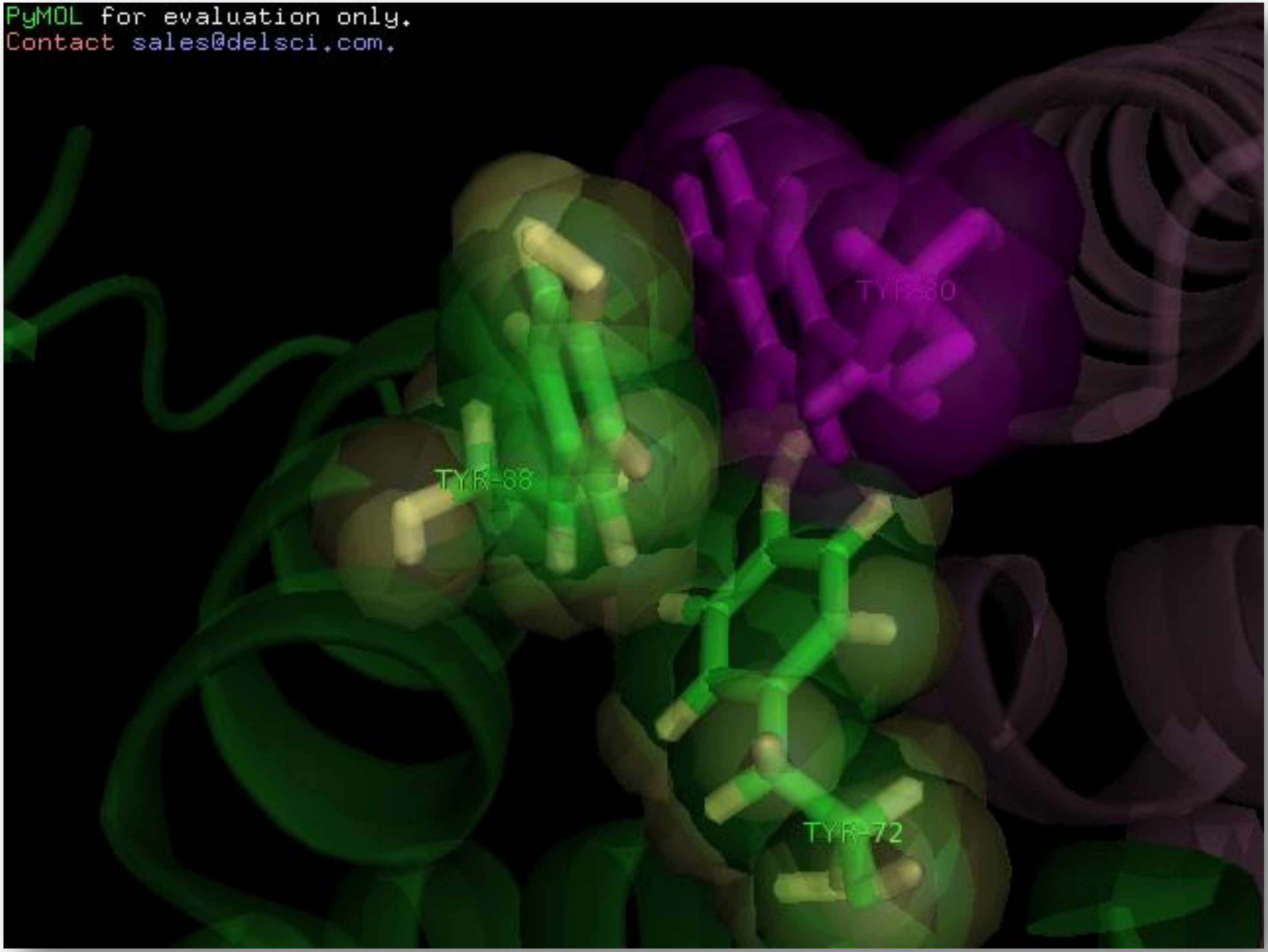


Figure49. Hydrophobic cluster in H2B-H4 interaction.

Protein-DNA interaction.

- The histone-fold DNA-binding sites can be divided into two types:
  - α1α1 motif.
  - L1L2 motif and termini of the α2 helices.
- The binding of the DNA to the nucleosome core particle is not sequence-specific.

Protein-DNA interaction: α1α1 binding site

H3/K64→main chain
H3/R63, H3/R69, H3/R72→side chain

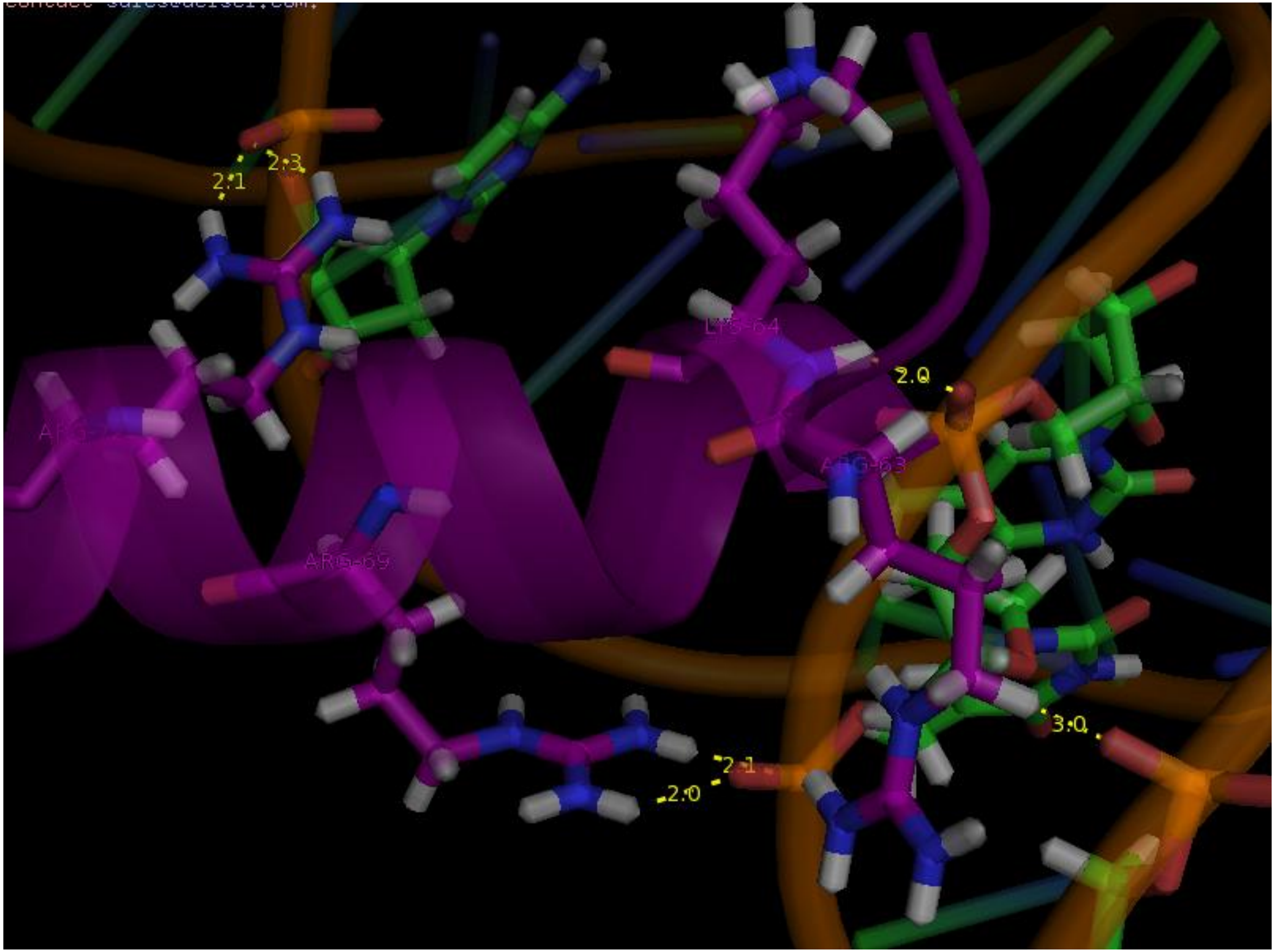


Figure 50. α1α1 binding site. Interaction of histone 3 with DNA.

Protein-DNA interaction: α1α1 binding site

H4/R36, H4/R35→side chain.

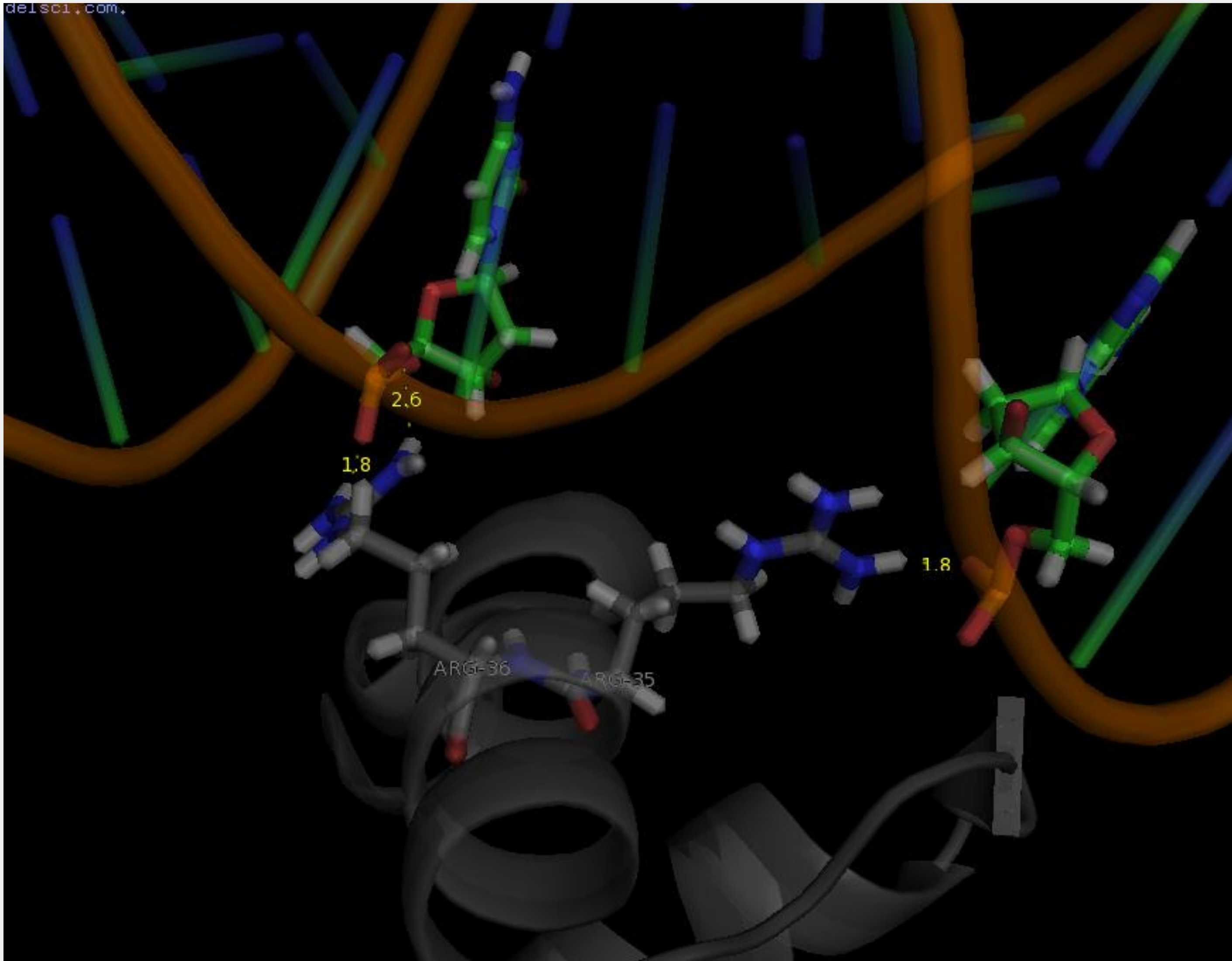


Figure 51. α1α1 binding site. Interaction of histone 4 with DNA.

Protein-DNA interaction: α1α1 binding site

H2B/I36, H2B/S33, H2B/Y37→main chain

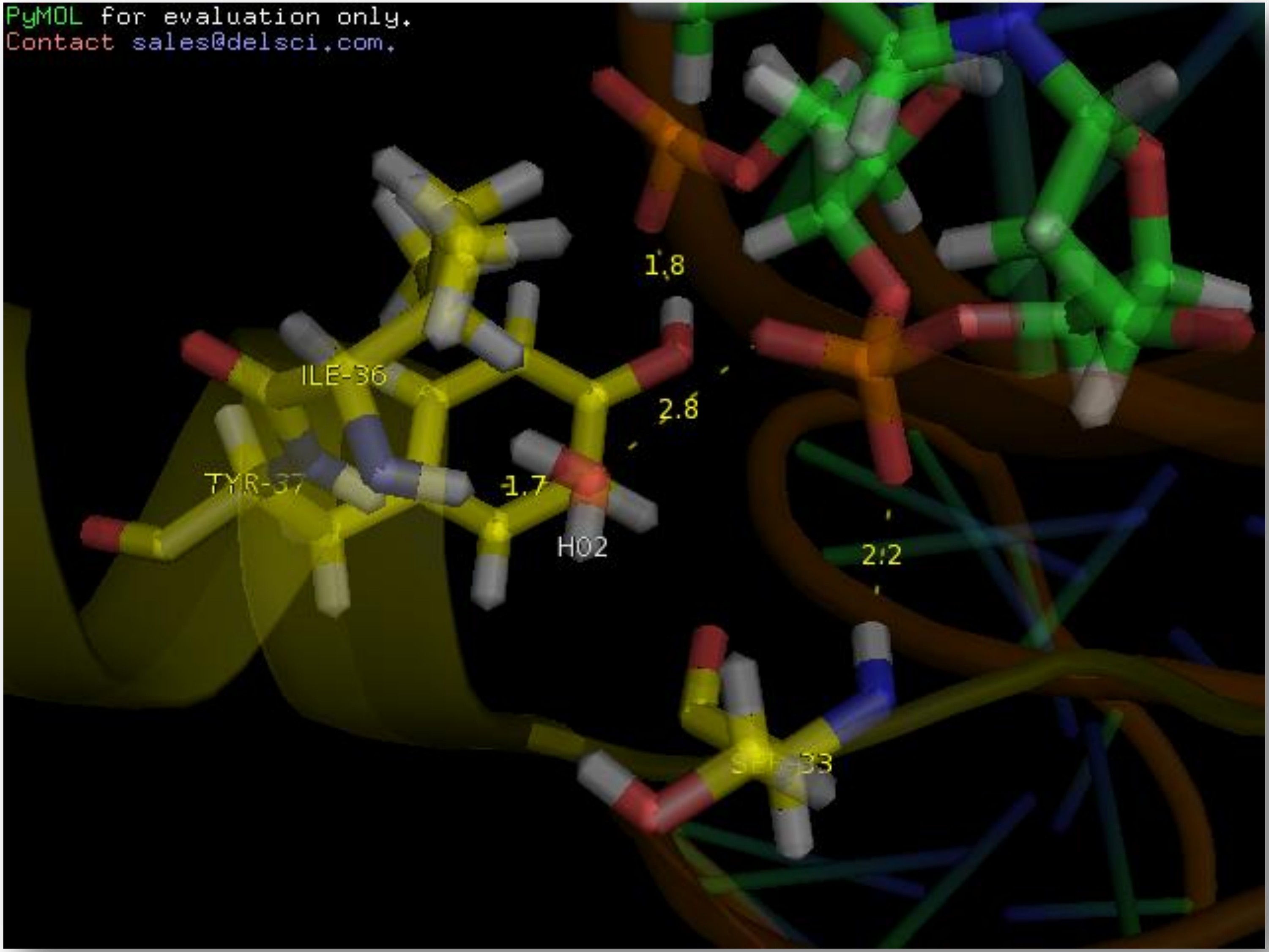


Figure 52. α1α1 binding site. Interaction of histone 2B with DNA.

Protein-DNA interaction: α1α1 binding site

H2A/R29, H2A/R32→side chain

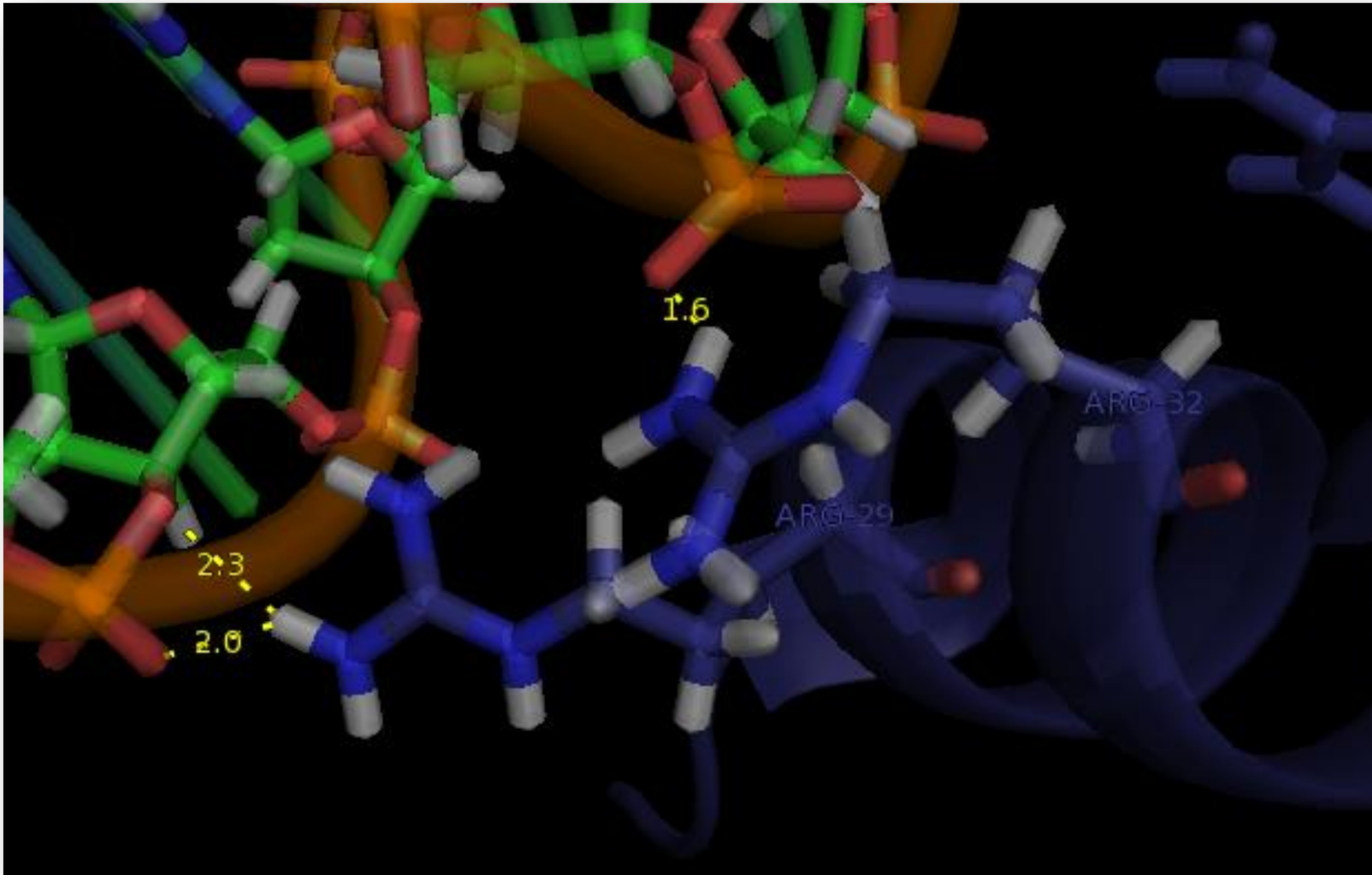


Figure 53. α1α1 binding site. Interaction of histone 2A with DNA.

Protein-DNA interaction: L1L2 binding site

H3/K115, H3/R83→side chain H3/T118, H3/F84→main chain

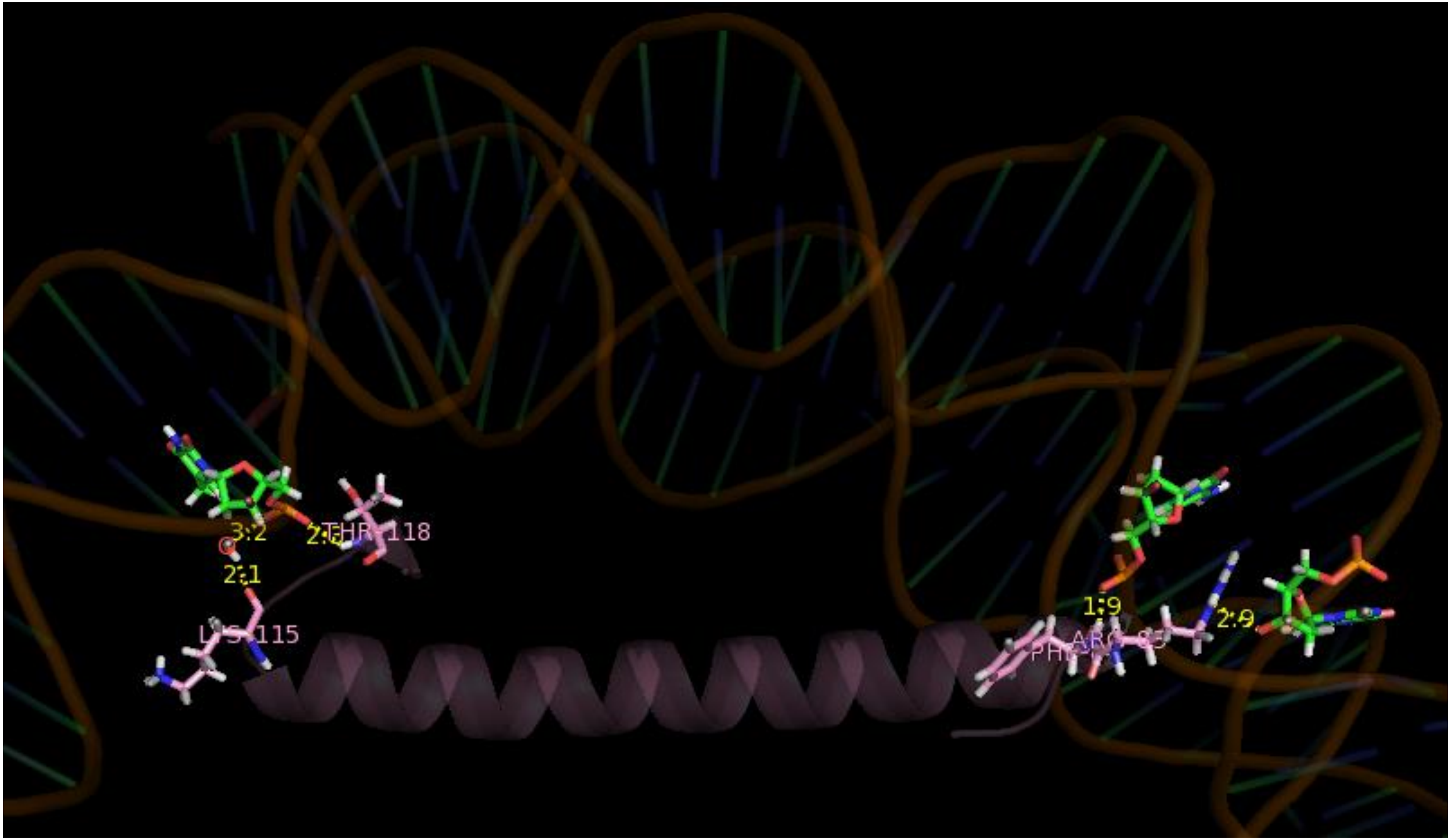


Figure 54. L1L2 binding site. Interaction of histone 3 with DNA.

Protein-DNA interaction: L1L2 binding site

H3/K115→side chain
H3/T118→main chain

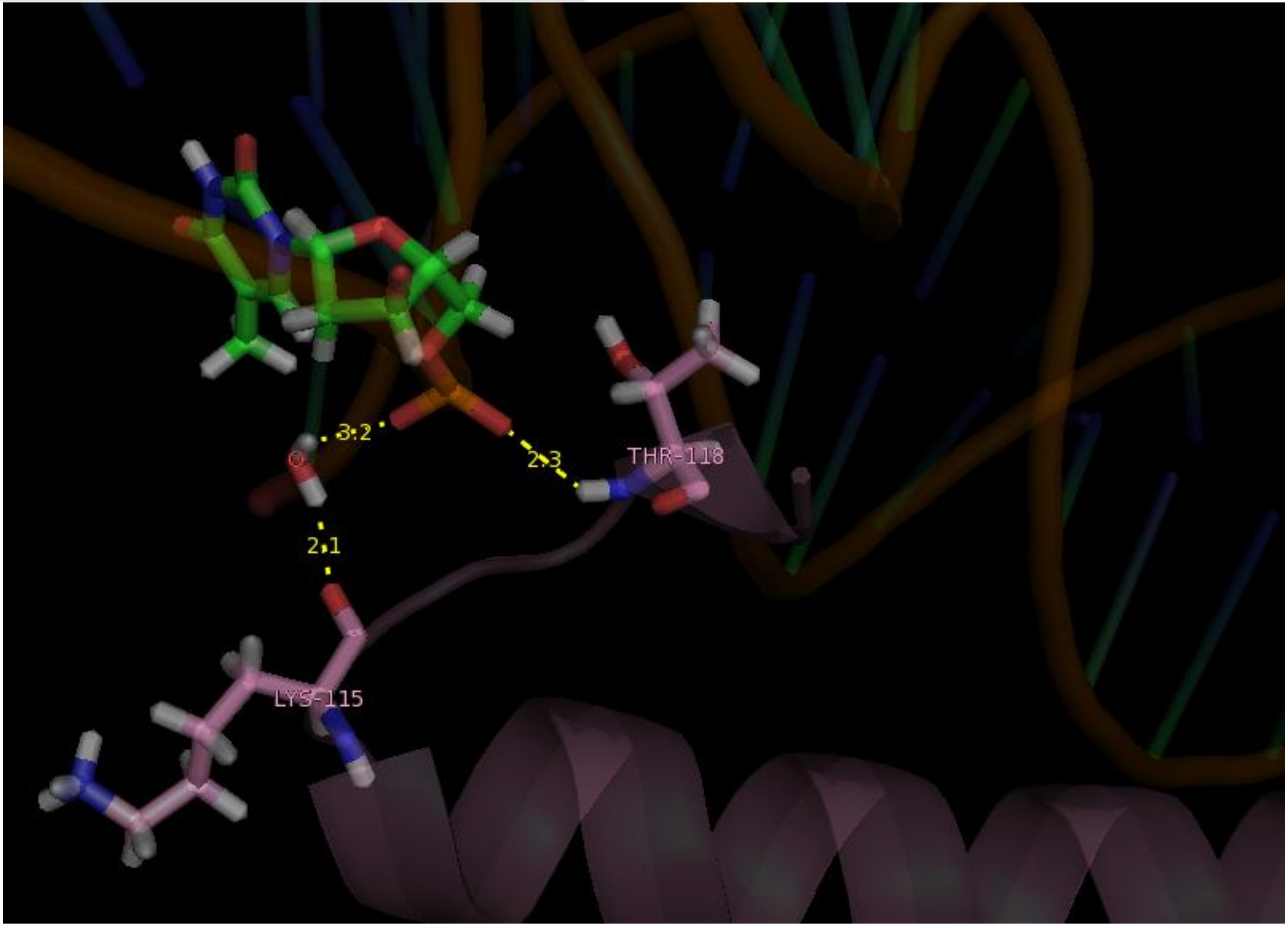


Figure 55. L1L2 binding site. Interaction of histone 3 with DNA.

Protein-DNA interaction: L1L2 binding site

H3/R83→side chain
H3/F84→main chain

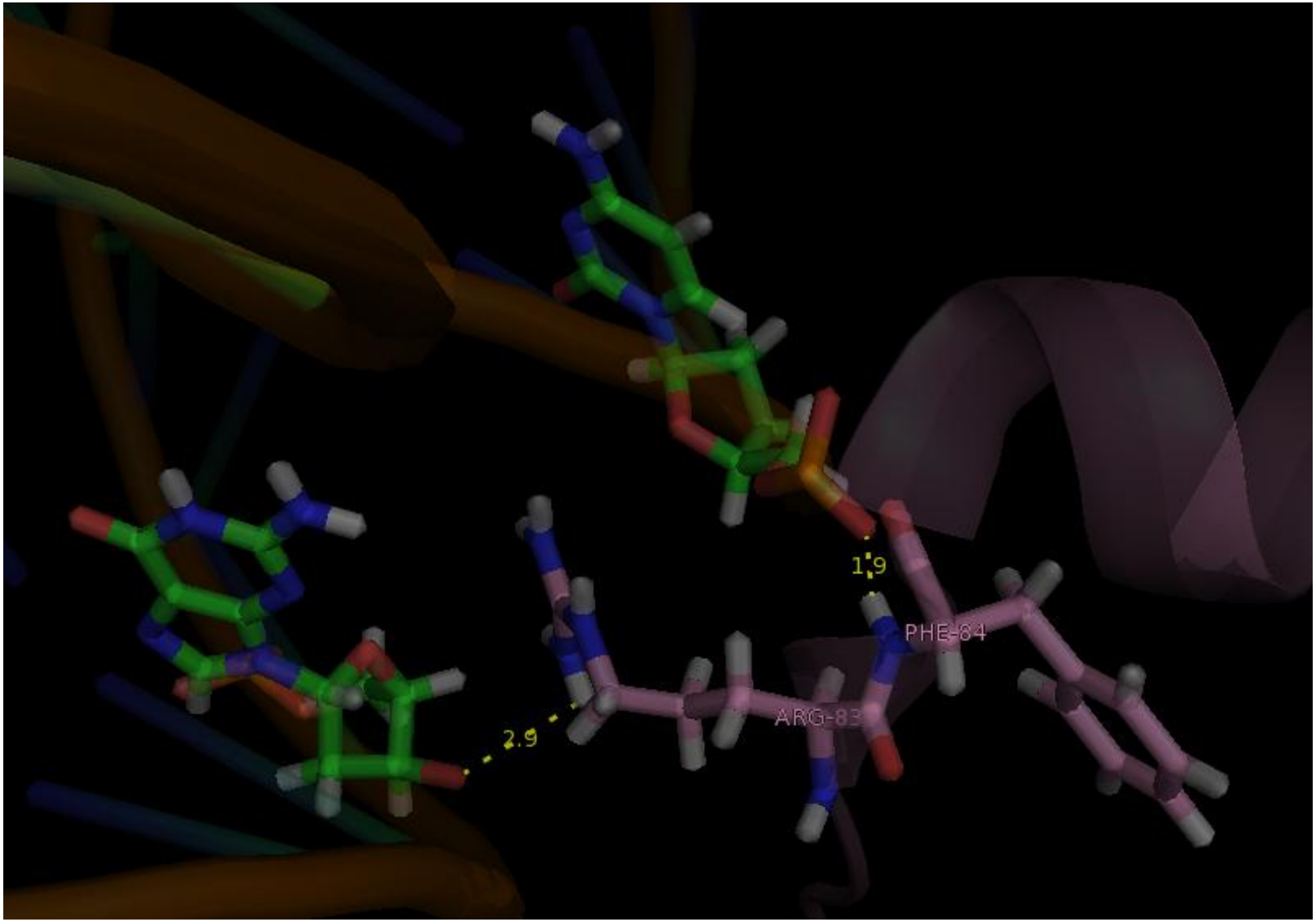


Figure 56. L1L2 binding site. Interaction of histone 3 with DNA.

Protein-DNA interaction: L1L2 binding site

H4/R45, H4/K77→side chain.
H4/I46, H4/G48, H4/K79→main chain.

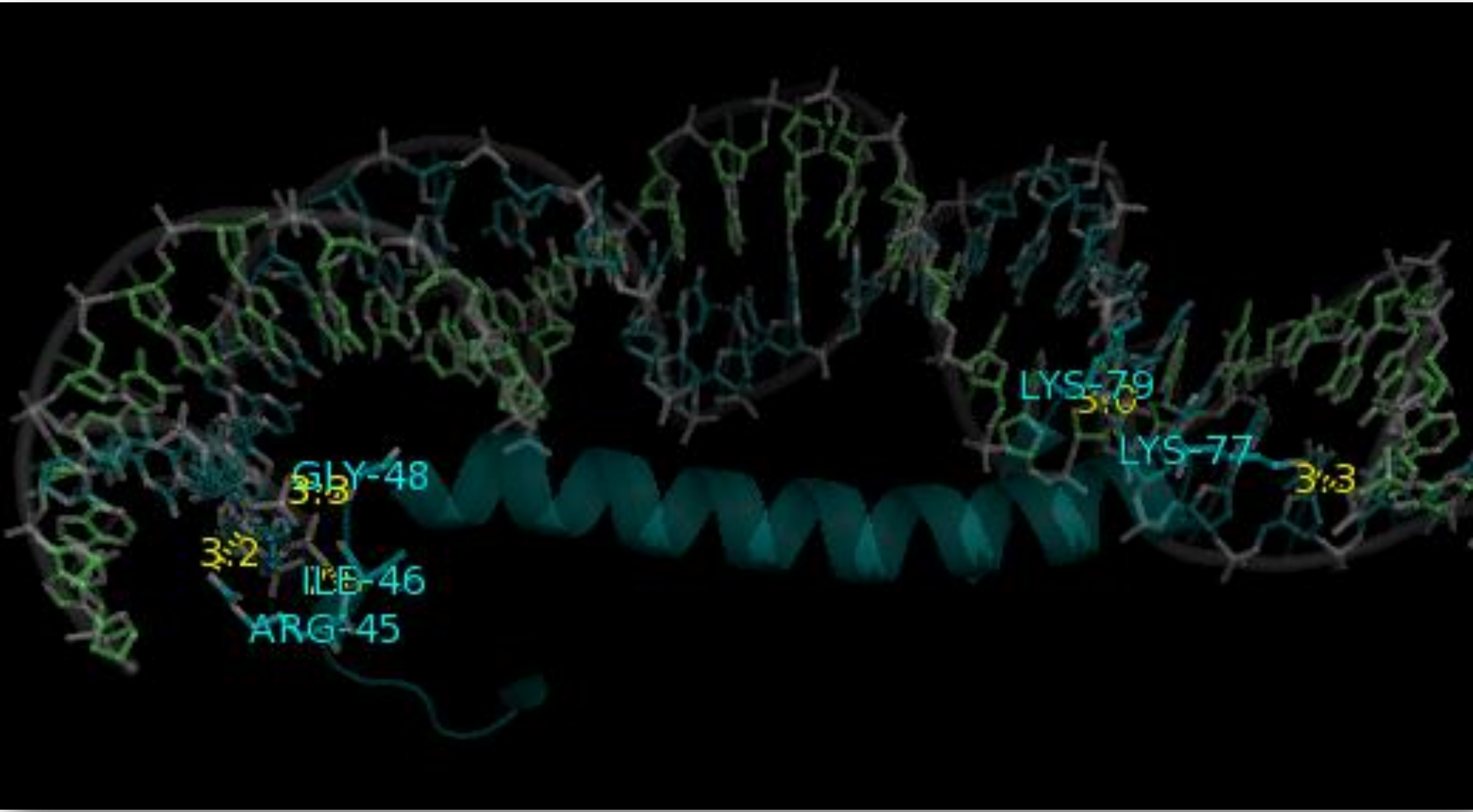


Figure 57. L1L2 binding site. Interaction of histone 4 with DNA.

Protein-DNA interaction: L1L2 binding site

H4/R45→side chain
H4/I46, H4/G48 →main chain.

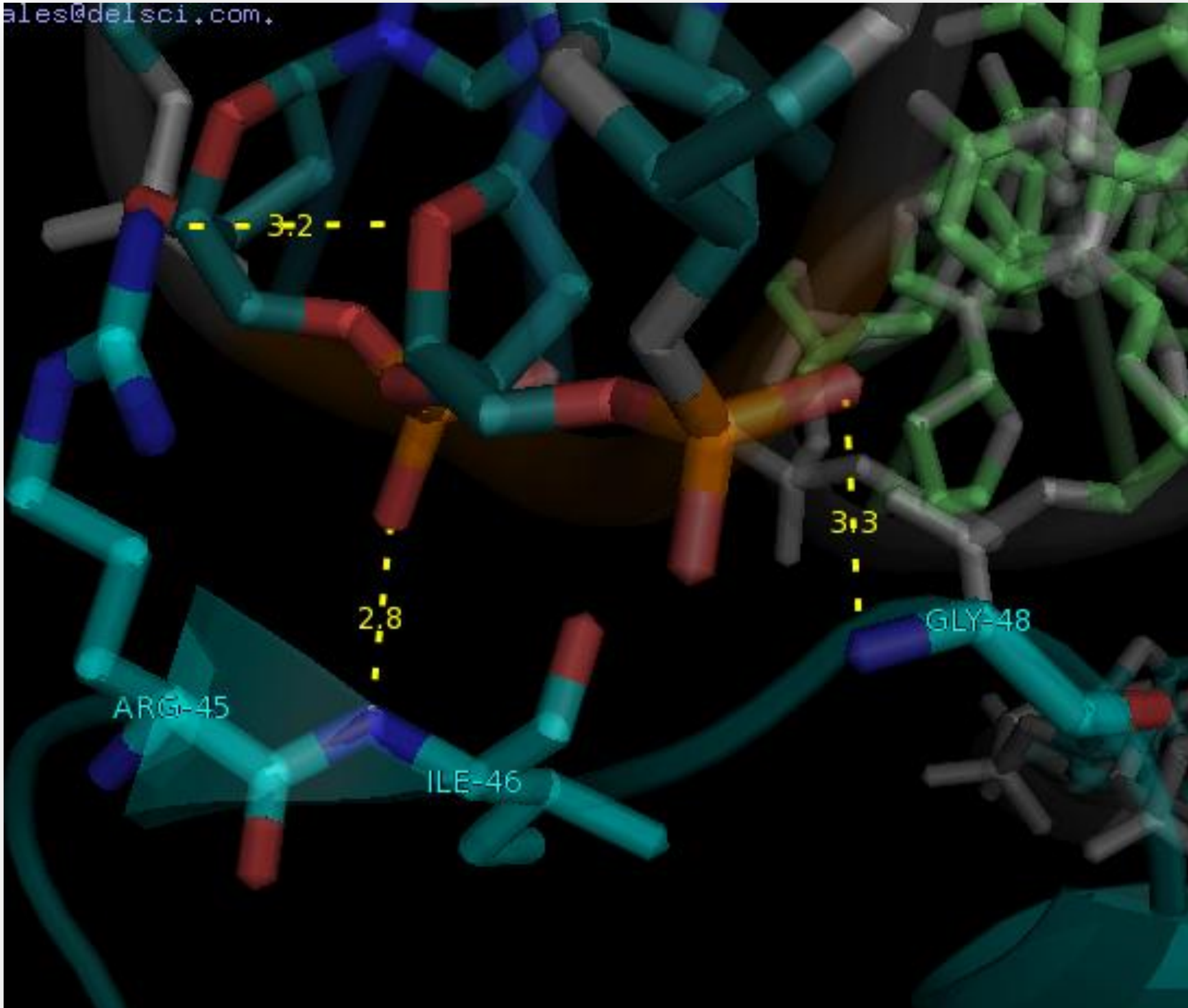


Figure 58. L1L2 binding site. Interaction of histone 4 with DNA.

Protein-DNA interaction: L1L2 binding site

H4/K77→side chain.
H4/K79→main chain

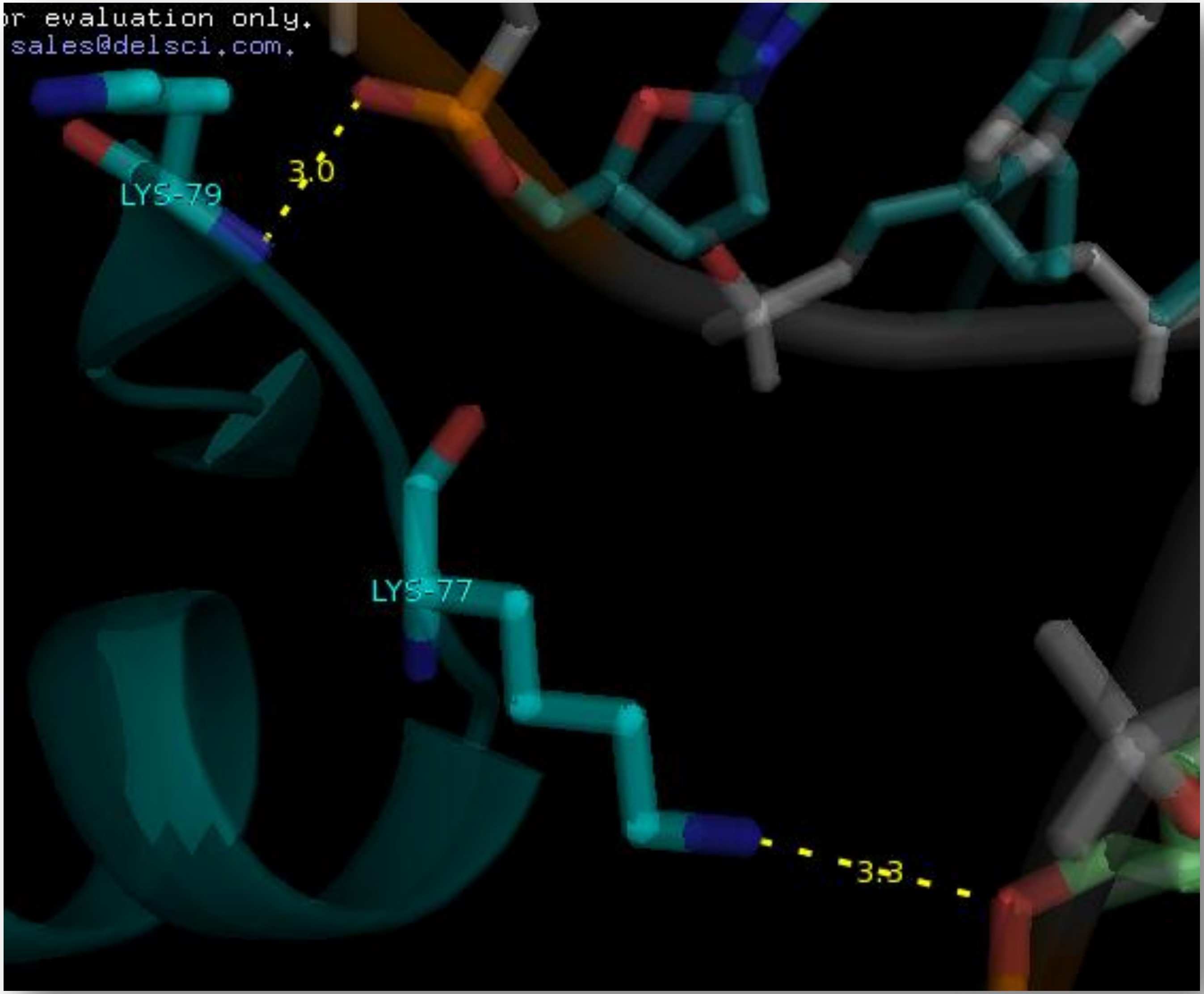


Figure 59. L1L2 binding site. Interaction of histone 4 with DNA.

Protein-DNA interaction: L1L2 binding site

H2A/V43, H2A/A45, H2A/R77→ main chain.
H2A/K74→ side chain.

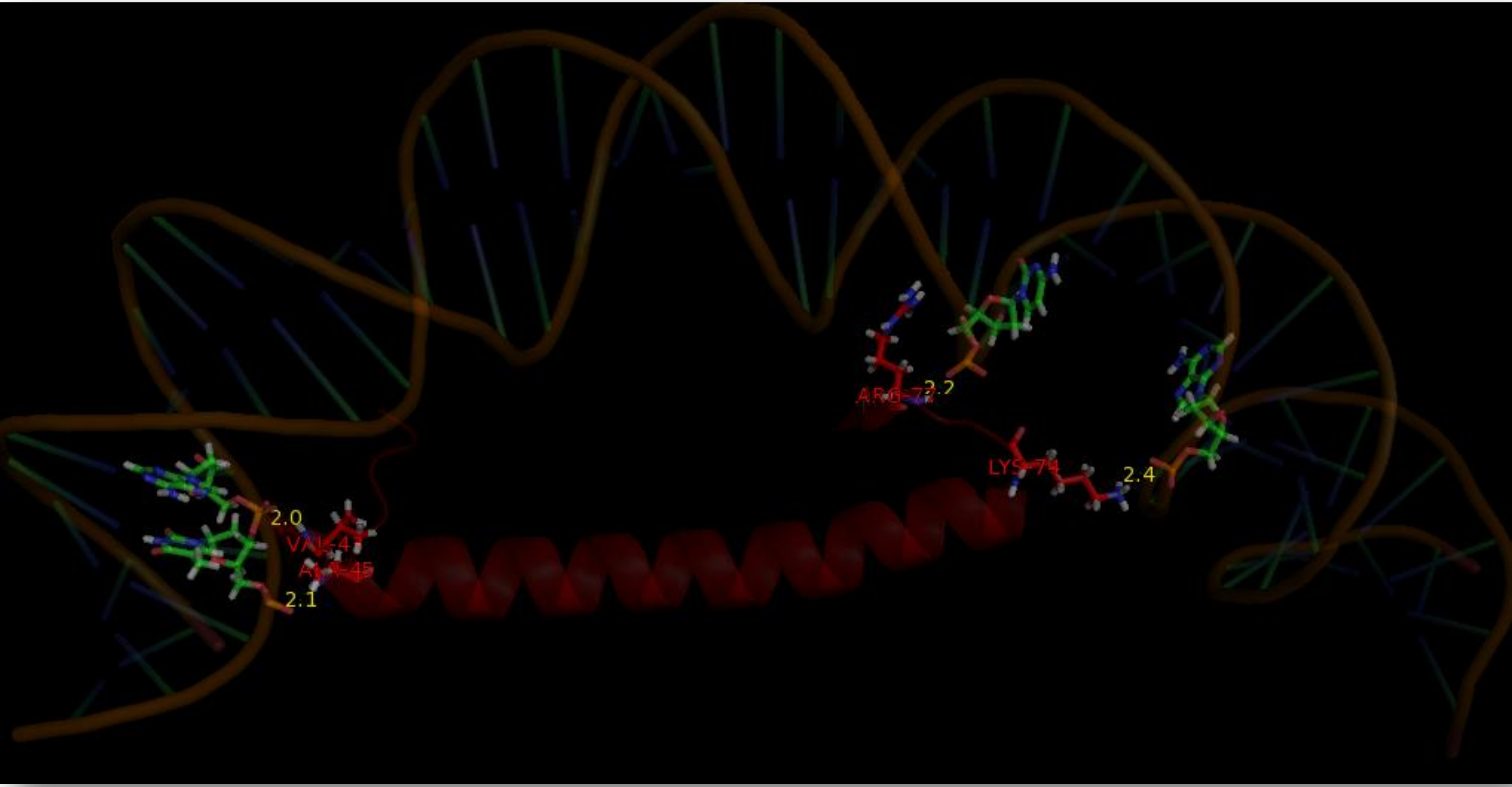


Figure 60. L1L2 binding site. Interaction of histone 2A with DNA.

Protein-DNA interaction: L1L2 binding site

H2A/V43, H2A/A45→ main chain

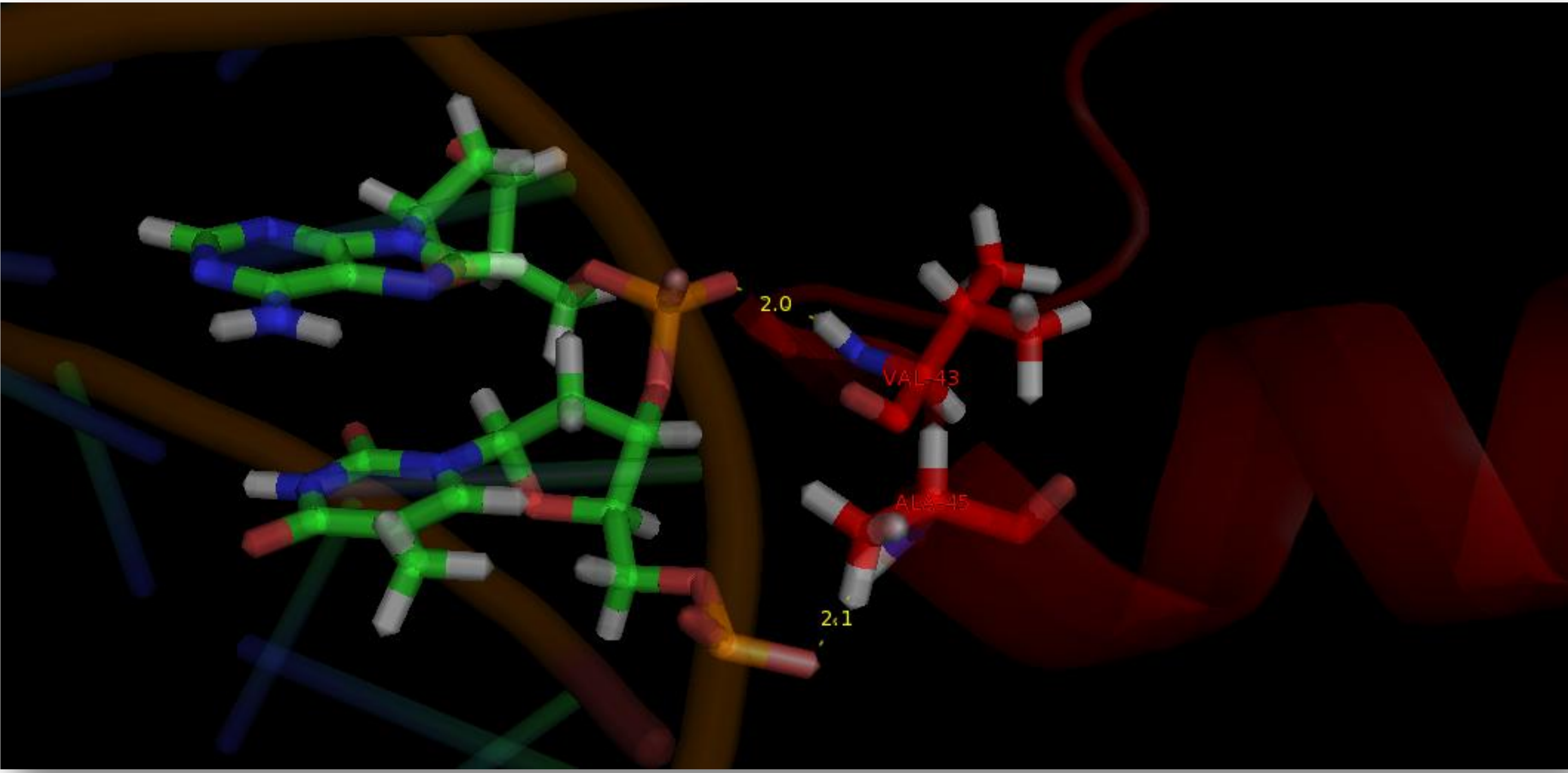


Figure 61. L1L2 binding site. Interaction of histone 2A with DNA.

Protein-DNA interaction: L1L2 binding site

H2A/R77→ main chain
H2A/K74→ side chain

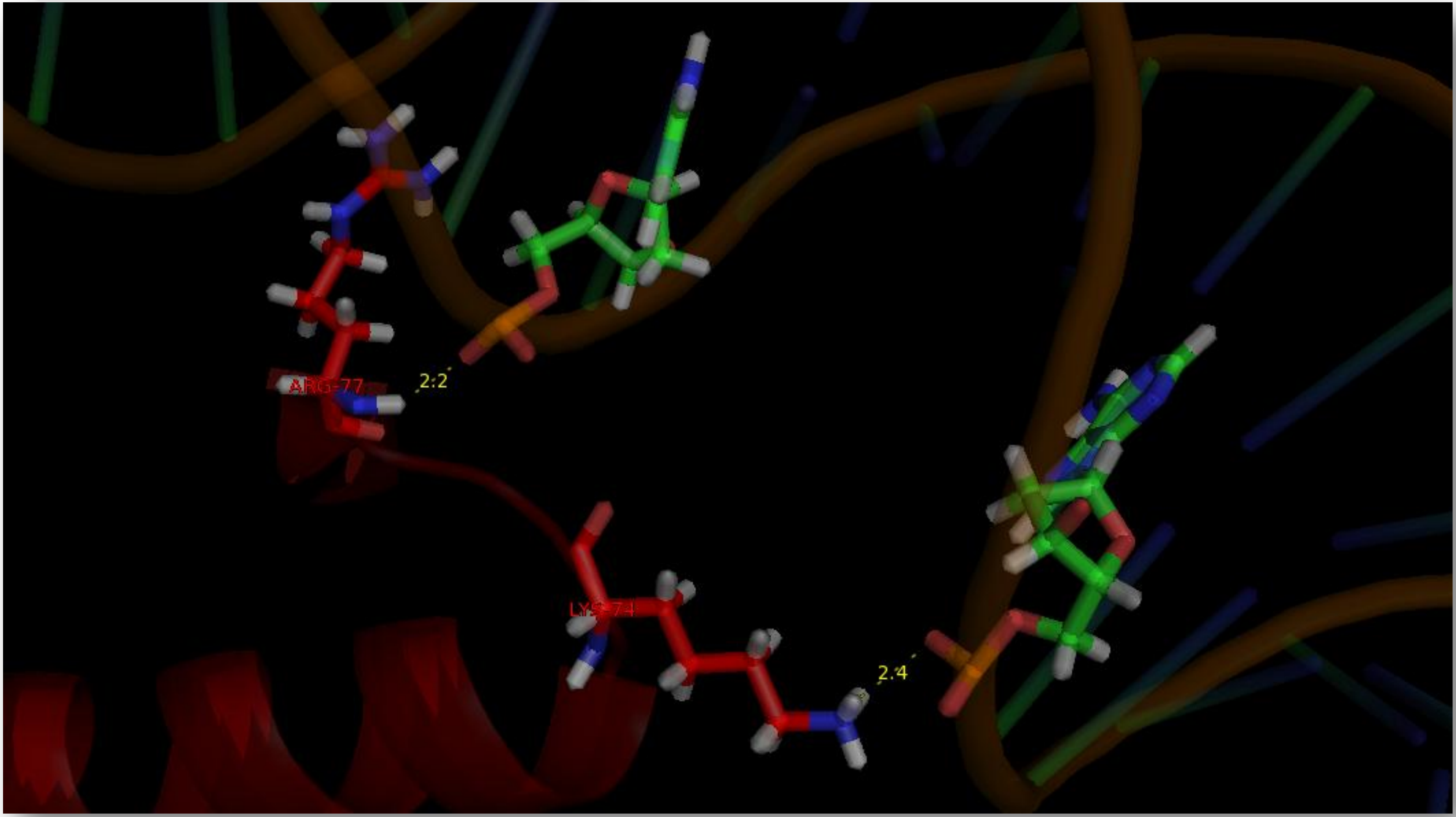


Figure 62. L1L2 binding site. Interaction of histone 2A with DNA.

Protein-DNA interaction: L1L2 binding site

H2B/K82, H2B/R83→ side chain
H2B/T85, H2B/S53→ main chian.

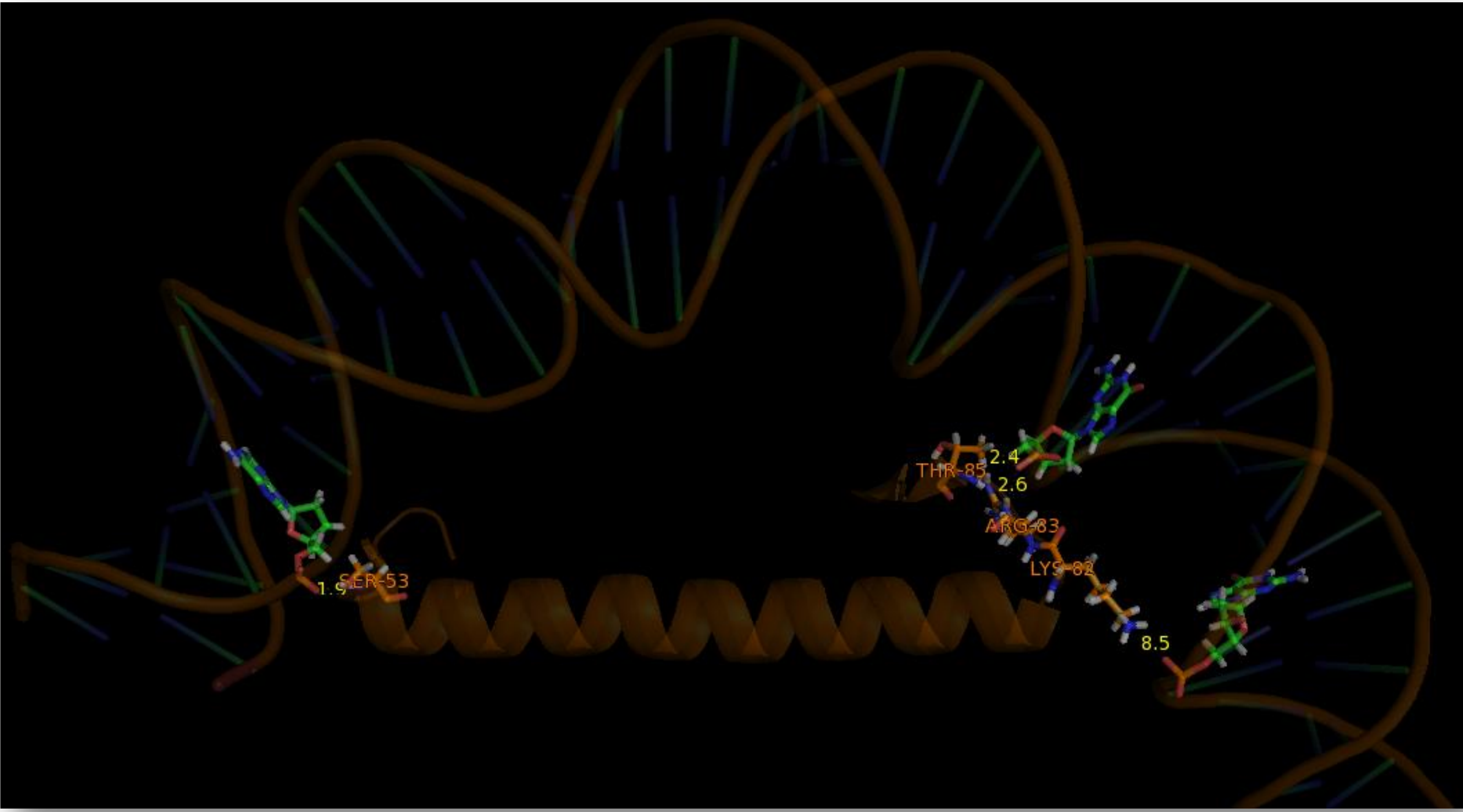


Figure 63. L1L2 binding site. Interaction of histone 2B with DNA.

Protein-DNA interaction: L1L2 binding site

H2B/S53→ main chian

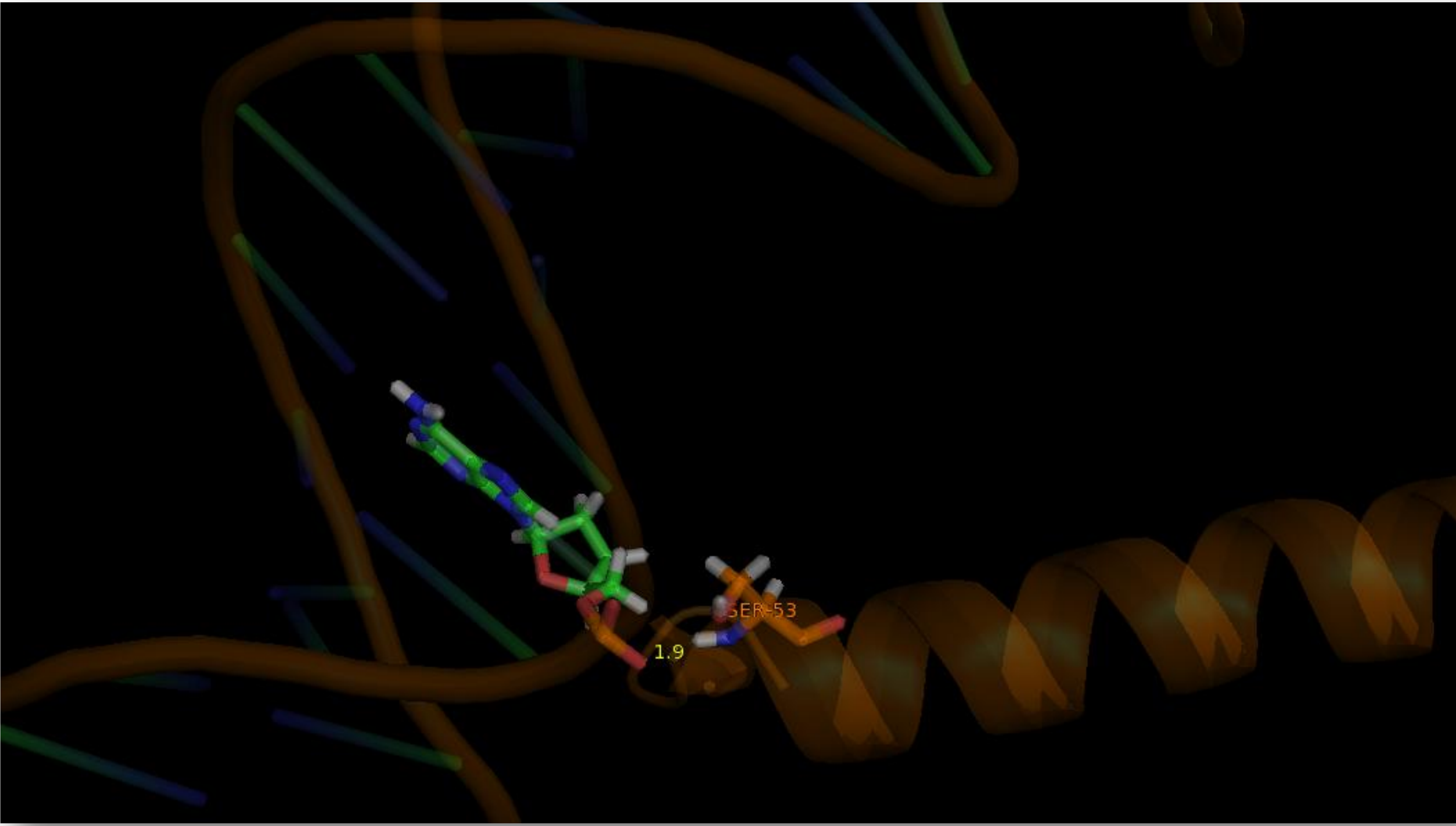


Figure 64 L1L2 binding site. Interaction of histone 2Bwith DNA.

Protein-DNA interaction: α1α1 binding site

H2B/K82, H2B/R83→ side chain
H2B/T85→ main chian

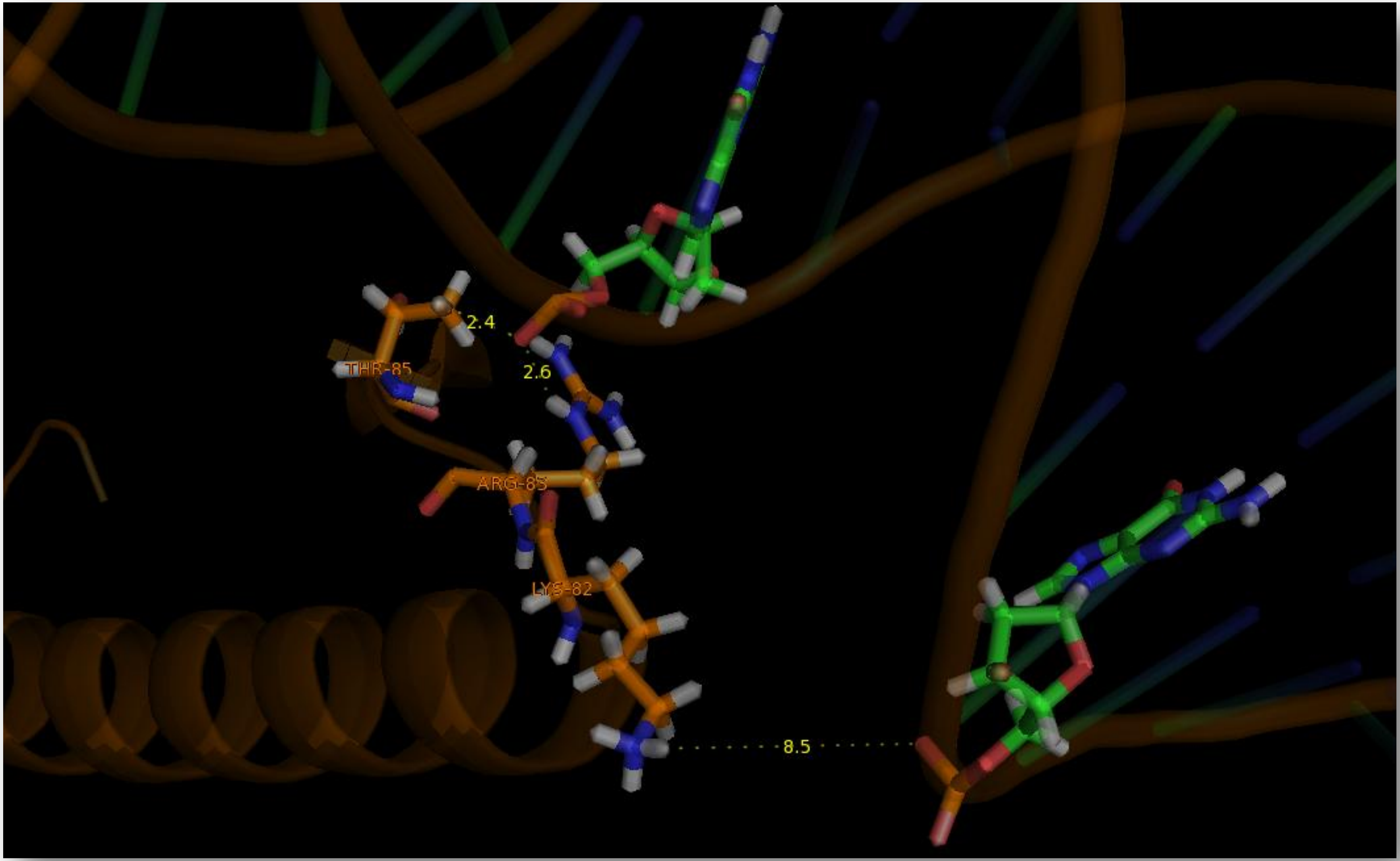


Figure 65. L1L2 binding site. Interaction of histone 2Bwith DNA.

Introduction

- Histones are found in animals, plants and lower eukaryotes.
- They always act as mediators of DNA compactation into chromatin.
- Archaebacteria show two strains of small histone-like proteins and induce super-coils: HMf and HMt.

HMf and HMt

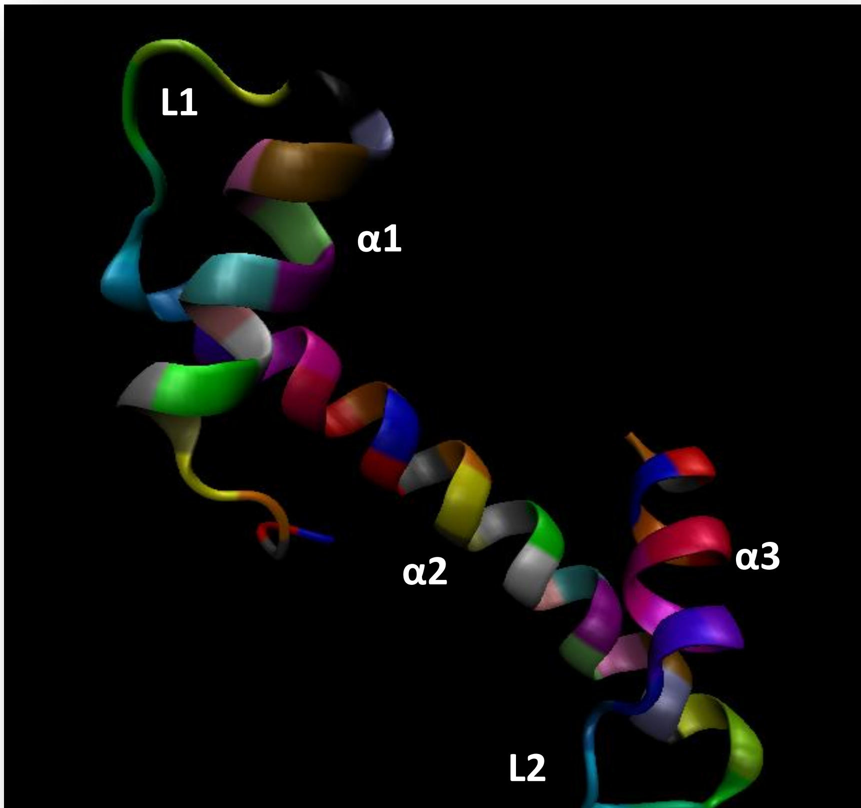


Figure66: structure of the HMfA of *Methanothermus fervidus* (PDBid: 1HTA)

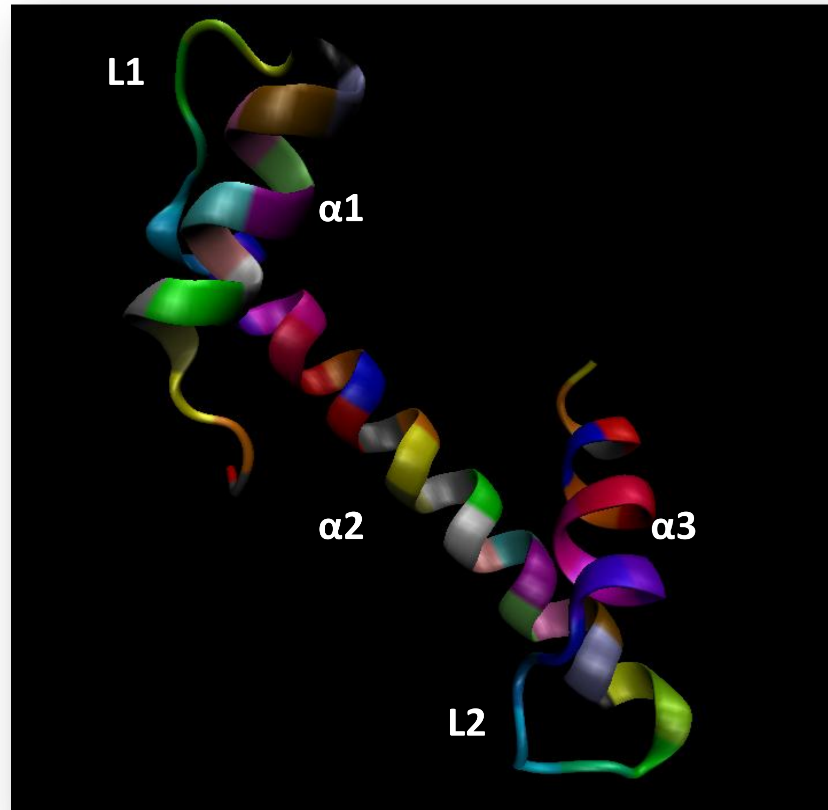


Figure67: structure of the HMfB of *Methanothermus fervidus* (PDBid: 1A7W)

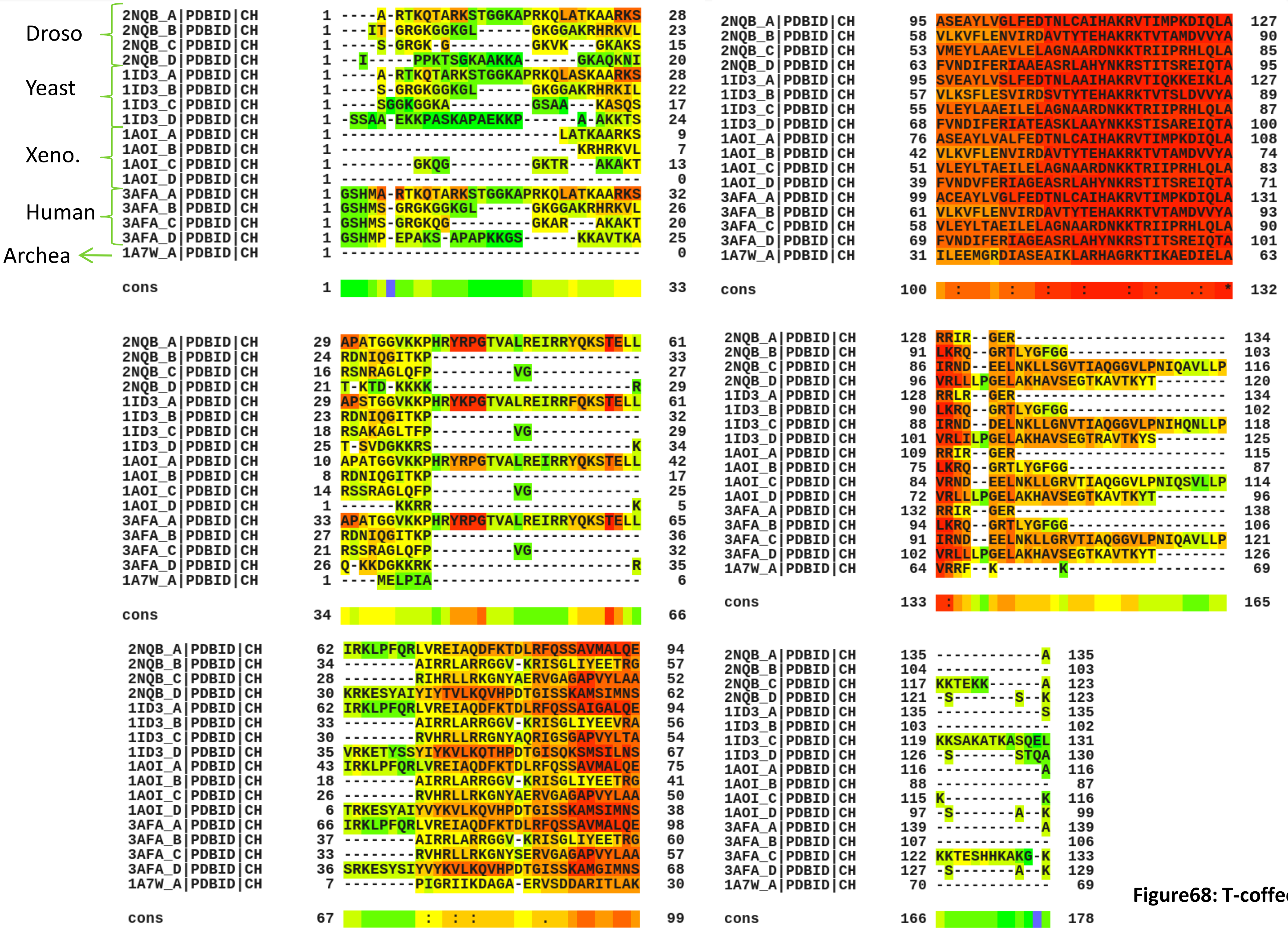


Figure68: T-coffee

H3

	RMSD	SCORE
H2A	1.73	8.89
H2B	1.22	8.90
H3	1.10	8.70
H4	1.31	8.48

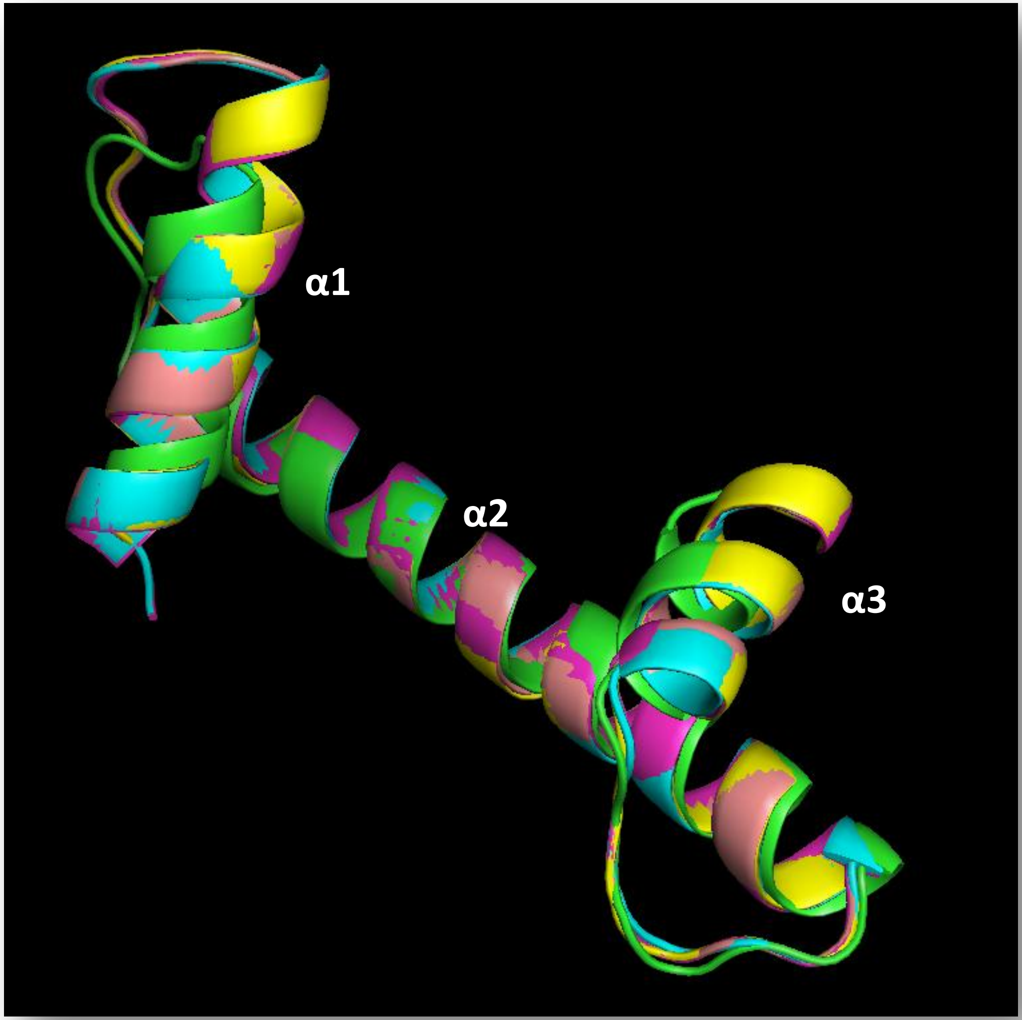


Figure69: Evolutionary structure conservation of H3 + HMfA (PDBid: 1A7W, 1AOI, 2NQB, 3AFA, 1ID3)

H3

```
2NQB_A|PDBID|CH      1 ----ARTKQTARKSTGGKAPRKQLATKAARKSA  29
1AOI_A|PDBID|CH      1 -----------------------LATKAARKSA  10
3AFA_A|PDBID|CH      1 GSHMARTKQTARKSTGGKAPRKQLATKAARKSA  33
1ID3_A|PDBID|CH      1 ----ARTKQTARKSTGGKAPRKQLASKAARKSA  29
1A7W_A|PDBID|CH      1 ---------------------------------   0

cons                 1                                    33

2NQB_A|PDBID|CH     30 PATGGVKKPHRYRPGTVALREIRRYQKSTELLI  62
1AOI_A|PDBID|CH     11 PATGGVKKPHRYRPGTVALREIRRYQKSTELLI  43
3AFA_A|PDBID|CH     34 PATGGVKKPHRYRPGTVALREIRRYQKSTELLI  66
1ID3_A|PDBID|CH     30 PSTGGVKKPHRYKPGTVALREIRRFQKSTELLI  62
1A7W_A|PDBID|CH      1 ----------------------------MELP-   4

cons                34                              **    66

2NQB_A|PDBID|CH     63 RKLPFQRLVREIAQDFKTDLRFQSSAVMALQEA  95
1AOI_A|PDBID|CH     44 RKLPFQRLVREIAQDFKTDLRFQSSAVMALQEA  76
3AFA_A|PDBID|CH     67 RKLPFQRLVREIAQDFKTDLRFQSSAVMALQEA  99
1ID3_A|PDBID|CH     63 RKLPFQRLVREIAQDFKTDLRFQSSAIGALQES  95
1A7W_A|PDBID|CH      5 -IAPIGRIIKDAGAERVSDD-----ARITLAKI  31

cons                67    *: *:::: . :  :*      *  :* :   99

2NQB_A|PDBID|CH     96 SEAYLVGLFEDTNLCAIHAKRVTIMPKDIQLAR 128
1AOI_A|PDBID|CH     77 SEAYLVALFEDTNLCAIHAKRVTIMPKDIQLAR 109
3AFA_A|PDBID|CH    100 CEAYLVGLFEDTNLCAIHAKRVTIMPKDIQLAR 132
1ID3_A|PDBID|CH     96 VEAYLVSLFEDTNLAAIHAKRVTIQKKEIKLAR 128
1A7W_A|PDBID|CH     32 LEEMGRDIASEAIKLARHAGRKTIKAEDIELAV  64

cons               100  *     : .::   * ** * **  ::*:**   132

2NQB_A|PDBID|CH    129 -RIRGERA 135
1AOI_A|PDBID|CH    110 -RIRGERA 116
3AFA_A|PDBID|CH    133 -RIRGERA 139
1ID3_A|PDBID|CH    129 -RLRGERS 135
1A7W_A|PDBID|CH     65 RRF---KK  69

cons               133  *:   :  140
```

Figure70: Sequence alignment for histone H3 of eukaryotes and HMf.

H3

```
CLUSTAL W(1.60) multiple sequence alignment

1A7W        -IAPIGRIIKD----A-GAERVSDDARITLAKILEEM
1ID3        RKLPFQRLVREIAQDFKTDLRFQSSAIGALQESVEAY
3AFA        RKLPFQRLVREIAQDFKTDLRFQSSAVMALQEACEAY
1AOI        -KLPFQRLVREIAQDFKTDLRFQSSAVMALQEASEAY
2NQB        -KLPFQRLVREIAQDFKTDLRFQSSAVMALQEASEAY


1A7W        GRDIASEAIKLARHAGRKTIKAEDIELAVR--
1ID3        LVSLFEDTNLAAIHAKRVTIQKKEIKLARR--
3AFA        LVGLFEDTNLCAIHAKRVTIMPKDIQLARRIR
1AOI        LVALFEDTNLCAIHAKRVTIMPKDIQLARRIR
2NQB        LVGLFEDTNLCAIHAKRVTIMPKDIQLARRIR
```

Figure71: CLUSTAL, estructural alignment.

H4

```
2NQB_B|PDBID|CH      1 ---ITGRGKGGKGLGKGGAKRHRKVLRDNIQGI   30
1ID3_B|PDBID|CH      1 ----SGRGKGGKGLGKGGAKRHRKILRDNIQGI   29
1AOI_B|PDBID|CH      1 -------------------KRHRKVLRDNIQGI   14
3AFA_B|PDBID|CH      1 GSHMSGRGKGGKGLGKGGAKRHRKVLRDNIQGI   33

cons                 1                    *****:********   33


2NQB_B|PDBID|CH     31 TKPAIRRLARRGGVKRISGLIYEETRGVLKVFL   63
1ID3_B|PDBID|CH     30 TKPAIRRLARRGGVKRISGLIYEEVRAVLKSFL   62
1AOI_B|PDBID|CH     15 TKPAIRRLARRGGVKRISGLIYEETRGVLKVFL   47
3AFA_B|PDBID|CH     34 TKPAIRRLARRGGVKRISGLIYEETRGVLKVFL   66

cons                34 ************************.*.*** **   66


2NQB_B|PDBID|CH     64 ENVIRDAVTYTEHAKRKTVTAMDVVYALKRQGR   96
1ID3_B|PDBID|CH     63 ESVIRDSVTYTEHAKRKTVTSLDVVYALKRQGR   95
1AOI_B|PDBID|CH     48 ENVIRDAVTYTEHAKRKTVTAMDVVYALKRQGR   80
3AFA_B|PDBID|CH     67 ENVIRDAVTYTEHAKRKTVTAMDVVYALKRQGR   99

cons                67 *.****:*************::***********   99


2NQB_B|PDBID|CH     97 TLYGFGG  103
1ID3_B|PDBID|CH     96 TLYGFGG  102
1AOI_B|PDBID|CH     81 TLYGFGG   87
3AFA_B|PDBID|CH    100 TLYGFGG  106

cons               100 *******  106
```

Figure72:T-coffee histones 4.

H2A

```
2NQB_C|PDBID|CH      1  ----SGRG-KGGKVKGKAKSRSNRAGLQFPVGR   28
1ID3_C|PDBID|CH      1  S---GGKGGKAGSAAKASQSRSAKAGLTFPVGR   30
1AOI_C|PDBID|CH      1  -----G--KQGGKTRAKAKTRSSRAGLQFPVGR   26
3AFA_C|PDBID|CH      1  GSHMSGRGKQGGKARAKAKTRSSRAGLQFPVGR   33

cons                 1       *   :.*..   :::** :*** *****   33

2NQB_C|PDBID|CH     29  IHRLLRKGNYAERVGAGAPVYLAAVMEYLAAEV   61
1ID3_C|PDBID|CH     31  VHRLLRRGNYAQRIGSGAPVYLTAVLEYLAAEI   63
1AOI_C|PDBID|CH     27  VHRLLRKGNYAERVGAGAPVYLAAVLEYLTAEI   59
3AFA_C|PDBID|CH     34  VHRLLRKGNYSERVGAGAPVYLAAVLEYLTAEI   66

cons                34  :*****:***::*:*:******:**:***:**:   66

2NQB_C|PDBID|CH     62  LELAGNAARDNKKTRIIPRHLQLAIRNDEELNK   94
1ID3_C|PDBID|CH     64  LELAGNAARDNKKTRIIPRHLQLAIRNDDELNK   96
1AOI_C|PDBID|CH     60  LELAGNAARDNKKTRIIPRHLQLAVRNDEELNK   92
3AFA_C|PDBID|CH     67  LELAGNAARDNKKTRIIPRHLQLAIRNDEELNK   99

cons                67  ************************:***:****   99

2NQB_C|PDBID|CH     95  LLSGVTIAQGGVLPNIQAVLLPKKTEKK-----  122
1ID3_C|PDBID|CH     97  LLGNVTIAQGGVLPNIHQNLLPKKSAKATKASQ  129
1AOI_C|PDBID|CH     93  LLGRVTIAQGGVLPNIQSVLLPKK---------  116
3AFA_C|PDBID|CH    100  LLGRVTIAQGGVLPNIQAVLLPKKTESHHKAKG  132

cons               100  **. ************:   *****          132

2NQB_C|PDBID|CH    123  -A   123
1ID3_C|PDBID|CH    130  EL   131
1AOI_C|PDBID|CH    117  --   116
3AFA_C|PDBID|CH    133  -K   133

cons               133       134
```

Figure73:T-coffee histones 2A.

H2B

```
2NQB_D|PDBID|CH      1 ---IP---PKTSGKAAKKAGK-AQKNITKTD-K   25
1ID3_D|PDBID|CH      1 SS-AAEKKPASKAPAEKKPAA--KKTSTSVDGK   30
1AOI_D|PDBID|CH      1 -------------------------------K    1
3AFA_D|PDBID|CH      1 GSHMPE--PAKSAPAPKKGSKKAVTKAQKKDGK   31

cons                 1                                *   33


2NQB_D|PDBID|CH     26 KKKRKRKESYAIYIYTVLKQVHPDTGISSKAMS   58
1ID3_D|PDBID|CH     31 KRSKVRKETYSSYIYKVLKQTHPDTGISQKSMS   63
1AOI_D|PDBID|CH      2 KRRKTRKESYAIYVYKVLKQVHPDTGISSKAMS   34
3AFA_D|PDBID|CH     32 KRKRSRKESYSIYVYKVLKQVHPDTGISSKAMG   64

cons                34 *: : ***:*: *:*.****.*******.*:*.   66


2NQB_D|PDBID|CH     59 IMNSFVNDIFERIAAEASRLAHYNKRSTITSRE   91
1ID3_D|PDBID|CH     64 ILNSFVNDIFERIATEASKLAAYNKKSTISARE   96
1AOI_D|PDBID|CH     35 IMNSFVNDVFERIAGEASRLAHYNKRSTITSRE   67
3AFA_D|PDBID|CH     65 IMNSFVNDIFERIAGEASRLAHYNKRSTITSRE   97

cons                67 *:******:***** ***:** ***:***::**   99


2NQB_D|PDBID|CH     92 IQTAVRLLLPGELAKHAVSEGTKAVTKYTSS--  122
1ID3_D|PDBID|CH     97 IQTAVRLILPGELAKHAVSEGTRAVTKYSSSTQ  129
1AOI_D|PDBID|CH     68 IQTAVRLLLPGELAKHAVSEGTKAVTKYTSA--   98
3AFA_D|PDBID|CH     98 IQTAVRLLLPGELAKHAVSEGTKAVTKYTSA--  128

cons               100 *******:**************:*****:*:    132


2NQB_D|PDBID|CH    123 K  123
1ID3_D|PDBID|CH    130 A  130
1AOI_D|PDBID|CH     99 K   99
3AFA_D|PDBID|CH    129 K  129

cons               133    133
```

Figure73:T-coffee histones 2B.

To sum up…

- The conservation of structure is not sequence conservation dependent.
- The histone fold origin is a duplication form and HSH segment.
- Protein and DNA interactions are very conserved.
- 75% of the histone structure is has an helix conformation.
- Many histone residues are critical in the contacts with other histones and with DNA.
- NCP are very important for the chromosomal molecules.


Questions and comments
ONE MONTH LEFT,
CABO VERDE!!

Bibliography

- Cooper J, Henikoff S. Adaptative evolution of the histone fold domain in centromeric histones. Mol Biol Evol. 2004; 21(9): 1712-1718.
- Luger K. Nucleosomes: Structure and Function. Encyclopedia of life sciences. 2001.
- White C, Suto R, Luger K. Structure of the yeast nucleosme core particle reveals fundamental changes in internucleosome interactions. EMBO. 2001; 20(18):5207-18.
- Alva V, Ammelburg M, Söding J, Lupas A. On the origin of the histone fold. BCM Struc Bio. 2007; 7:17
- Arents G, Moudrianakis E. The histone fold: A ubiquitous architectural motif utilized in DNA compactation and protein dimerization. Proc Natl Acad Sci. 1995; 92: 11170-74.
- Hadjithomas M, Moudrianakis E. Experimental evidence for the role of domain swapping in the evolution of the histone fold. PNAS. 2011; 108(33): 13462-7.
- Zlatanova J. Archeal chromatin: virtual or real?. Proc. Natl. Acad. Sci. USA. 1997; 94: 12251-4.

Bibliography

- Curt AD, Sargent DF, Luger K, Maeder AW, Richmond TJ. Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 A resolution. JMC. 2002; 319, 1097-1113.
- Lugar K, Richmond TJ. DNA binding within the nucleosome core. Current opinion in Structural Biology. 1998; 8:33-40.
- Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ. Crystal structure of the nucleosome core particle at 2.8A resolution. Nature. 1997; 389, 251-260.