Mid-term report:

HPC-LEAP
EUROPEAN JOINT DOCTORATES

HPC-based molecular simulation approaches to investigate Protein-DNA interactions

Wenping Lyu
Forschungszentrum Juelich
April 18, 2017

Supervisors:

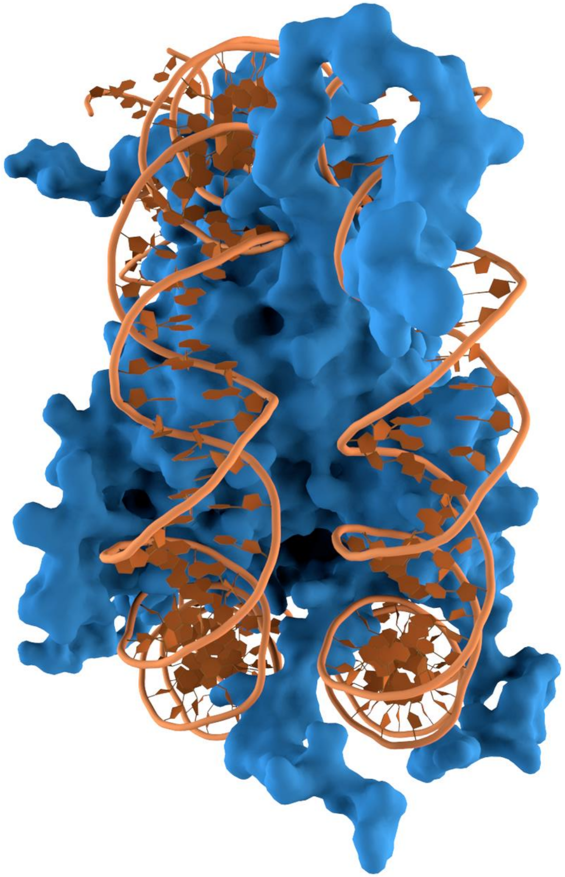
Prof. Dr. Paolo Carloni	RWTH Aachen
Jun-Prof. Dr. Giulia Rossetti	RWTH Aachen
Prof. Dr. George Christophides	The Cyprus Institute

ESR-6: Wenping Lyu

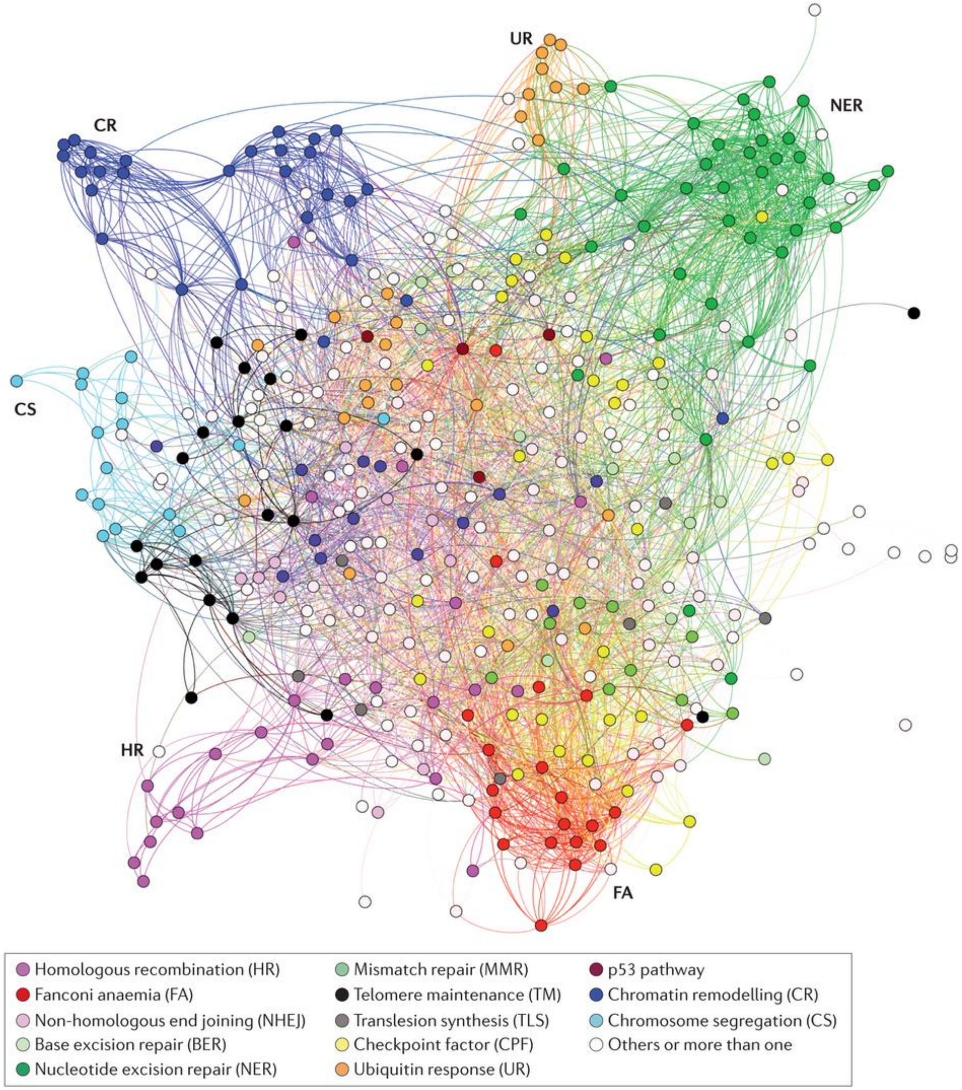
- B.S. in physics (Southwest University)
- M.S. in computational biophysics (Dalian University of Technology)

Background

Proteins are not working alone. --Arthur Lesk



Interaction of DNA (in orange)
with histones (in blue)



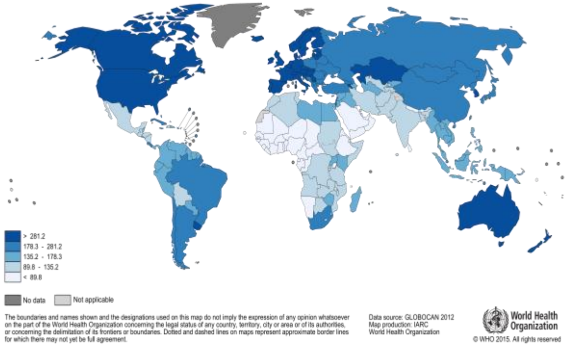
A network view of the DNA damage response.

Nature Reviews | Cancer

Background

Cancer, is an abnormal, uncontrolled multiplication of cells.

Men

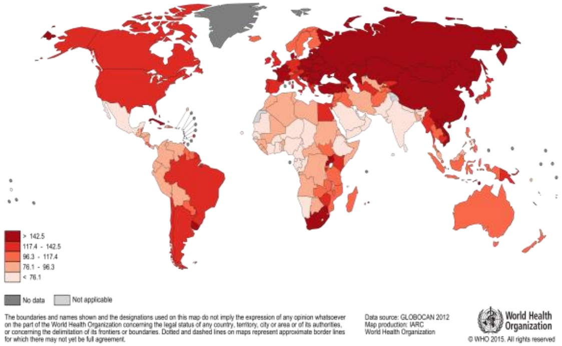


The boundaries and names shown and the designations used on this map do not imply the expression of any opinion whatsoever on the part of the World Health Organization concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Dotted and dashed lines on maps represent approximate border lines for which there may not yet be full agreement.

Data source: GLOBOCAN 2012
Map production: IARC
World Health Organization

© WHO 2015. All rights reserved

Men

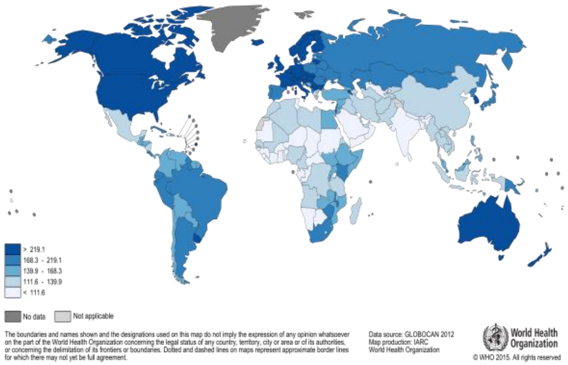


The boundaries and names shown and the designations used on this map do not imply the expression of any opinion whatsoever on the part of the World Health Organization concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Dotted and dashed lines on maps represent approximate border lines for which there may not yet be full agreement.

Data source: GLOBOCAN 2012
Map production: IARC
World Health Organization

© WHO 2015. All rights reserved

Women

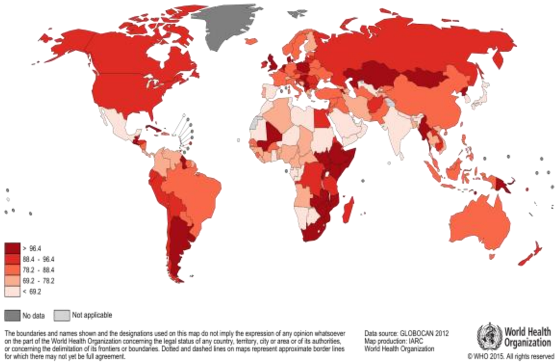


The boundaries and names shown and the designations used on this map do not imply the expression of any opinion whatsoever on the part of the World Health Organization concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Dotted and dashed lines on maps represent approximate border lines for which there may not yet be full agreement.

Data source: GLOBOCAN 2012
Map production: IARC
World Health Organization

© WHO 2015. All rights reserved

Women



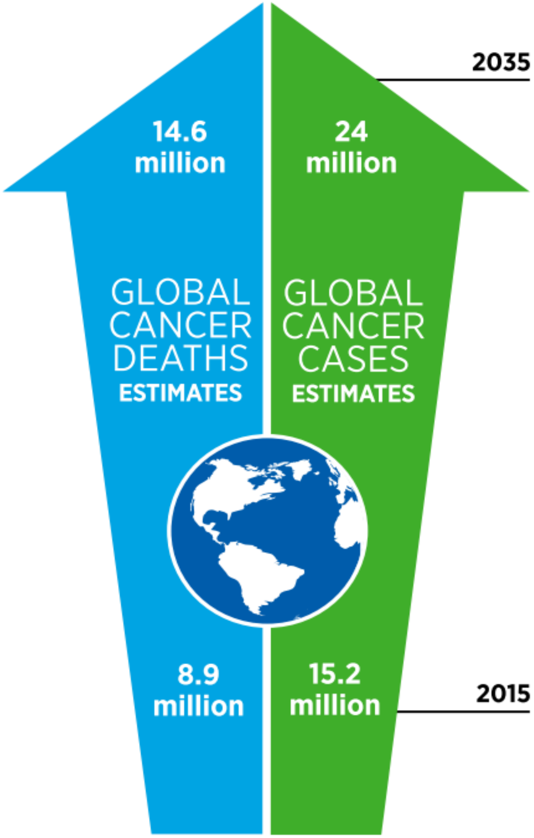
The boundaries and names shown and the designations used on this map do not imply the expression of any opinion whatsoever on the part of the World Health Organization concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Dotted and dashed lines on maps represent approximate border lines for which there may not yet be full agreement.

Data source: GLOBOCAN 2012
Map production: IARC
World Health Organization

© WHO 2015. All rights reserved

Estimated Cancer Incidence (left) and Mortality (right) Worldwide in 2012.

A GLOBAL CHALLENGE !

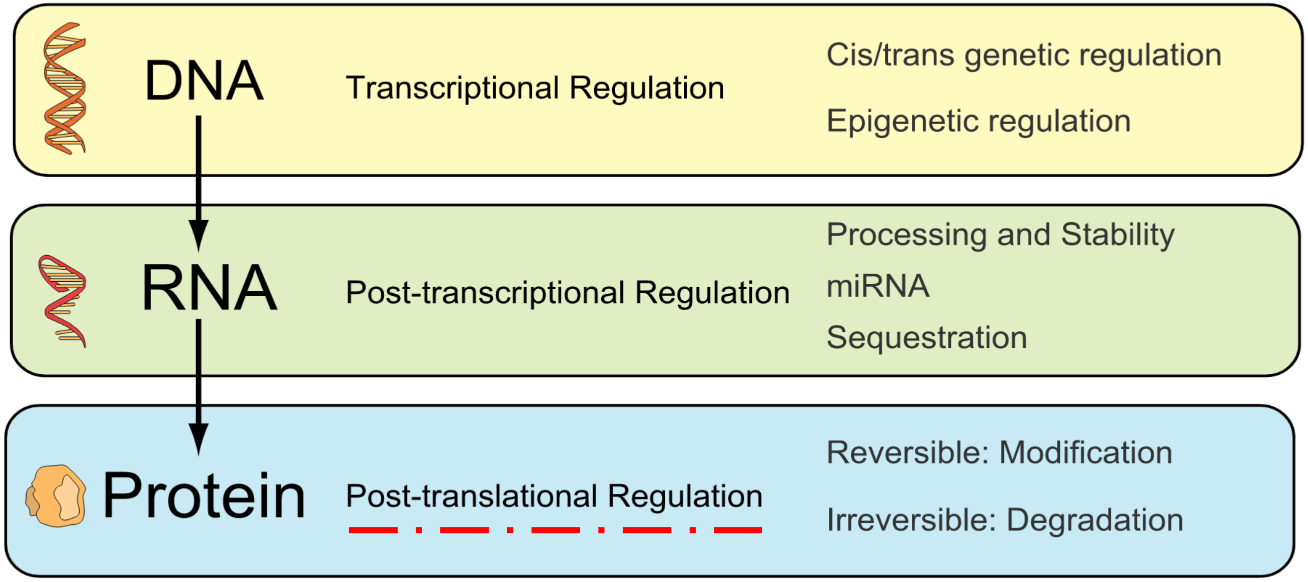


International Agency for Research on Cancer; 2017: http://globocan.iarc.fr

AACR CANCER PROGRESS REPORT 2016

Background

Cell replication starts from gene expression:



DNA
Transcriptional Regulation
Cis/trans genetic regulation
Epigenetic regulation

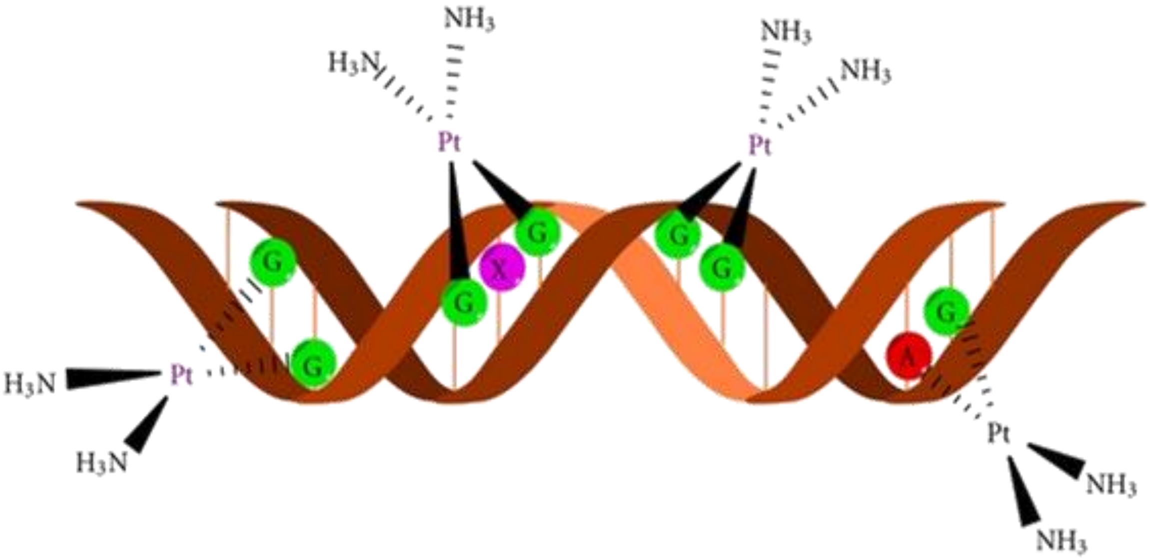
RNA
Post-transcriptional Regulation
Processing and Stability
miRNA
Sequestration

Protein
Post-translational Regulation
Reversible: Modification
Irreversible: Degradation

Regulation of gene expression involves in multiple steps, depends on various cellular and developmental conditions.

Post-translational modification,
nature's escape from genetic imprisonment,
induces additional dynamic information encoding processes.

Background



Cisplatin
inhibits gene replication
and transcription

Apoptotic Signals

Kill Cancer Cells

An powerful anticancer drug

Signal
Attenuation

Drug resistance that ensues is a major limitation of cisplatin.

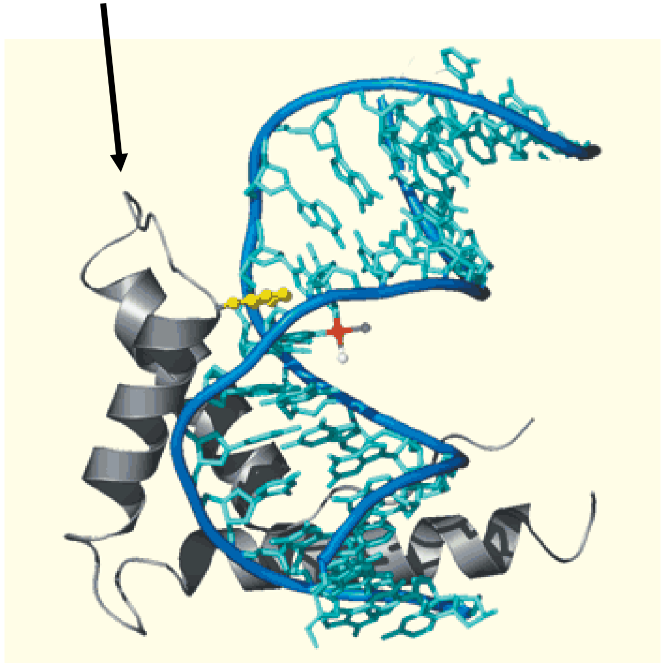
Oncogene (2003) **22,** 7265–7279.

Background

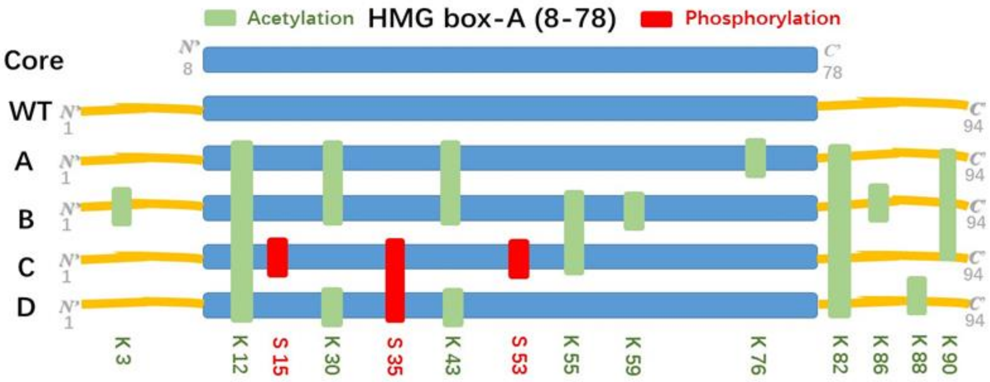
HMGB1 (High mobility group Box 1) can specifically binds about 90% of the cisplatin–DNA adducts formed in vivo.

Mol Med. 2008 Jul-Aug; 14(7-8): 476–484.

“*L-shape*” **HMGB1A**



Science 10 Apr 1992, 256, 5054, 234-237



HMGB1 isoforms bearing multisite PTMs in vivo.

Chem. Sci., 2015,**6**, 2074-2078

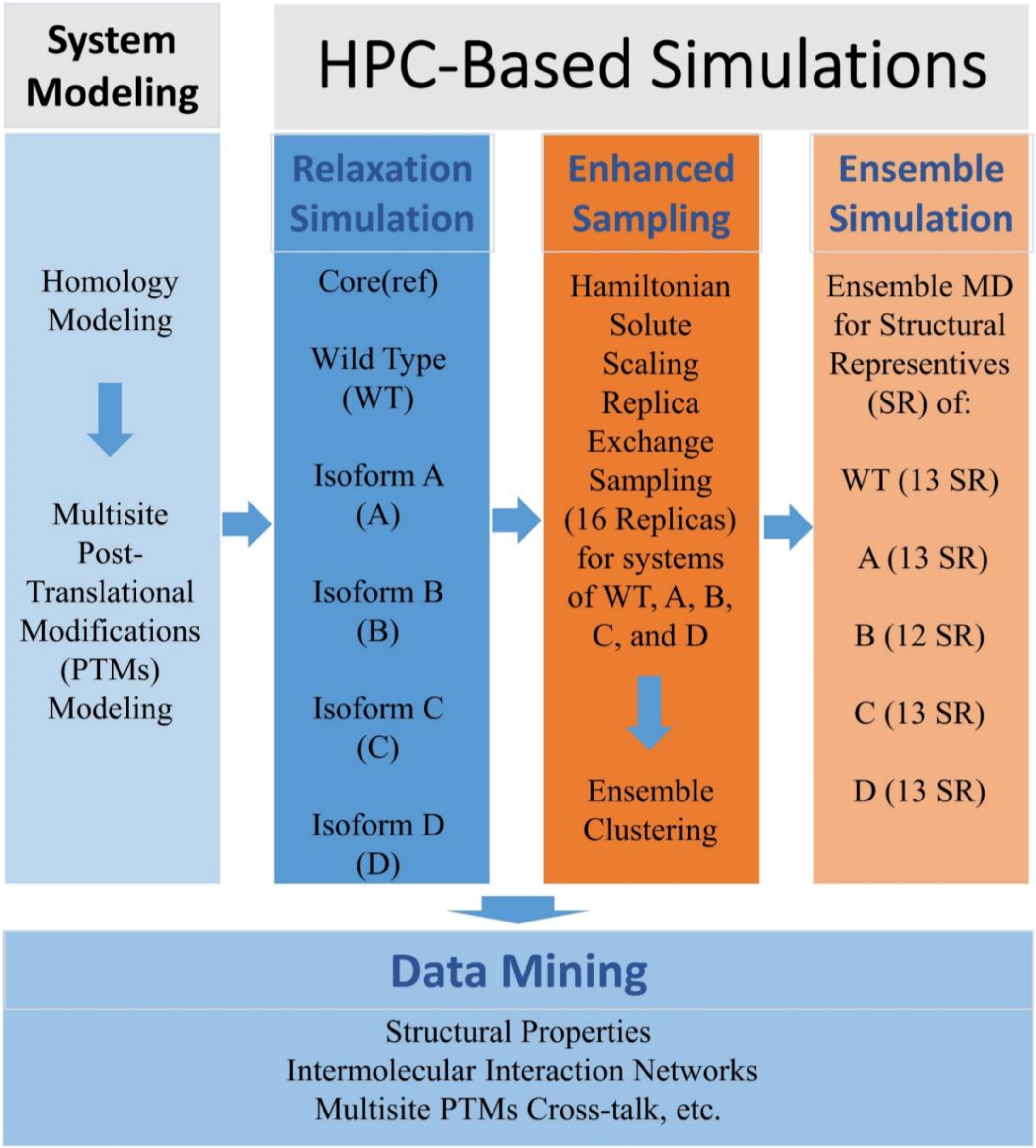
Multisite PTMs

potential ‘mod-forms’ →

Encoding complex biological meanings

No Structural Basis

Protein-Nucleic Acid interactions



System Modeling

Homology Modeling

Multisite Post-Translational Modifications (PTMs) Modeling

HPC-Based Simulations

Relaxation Simulation

Core(ref)

Wild Type (WT)

Isoform A (A)

Isoform B (B)

Isoform C (C)

Isoform D (D)

Enhanced Sampling

Hamiltonian Solute Scaling Replica Exchange Sampling (16 Replicas) for systems of WT, A, B, C, and D

Ensemble Clustering

Ensemble Simulation

Ensemble MD for Structural Representives (SR) of:

WT (13 SR)

A (13 SR)

B (12 SR)

C (13 SR)

D (13 SR)

Data Mining

Structural Properties

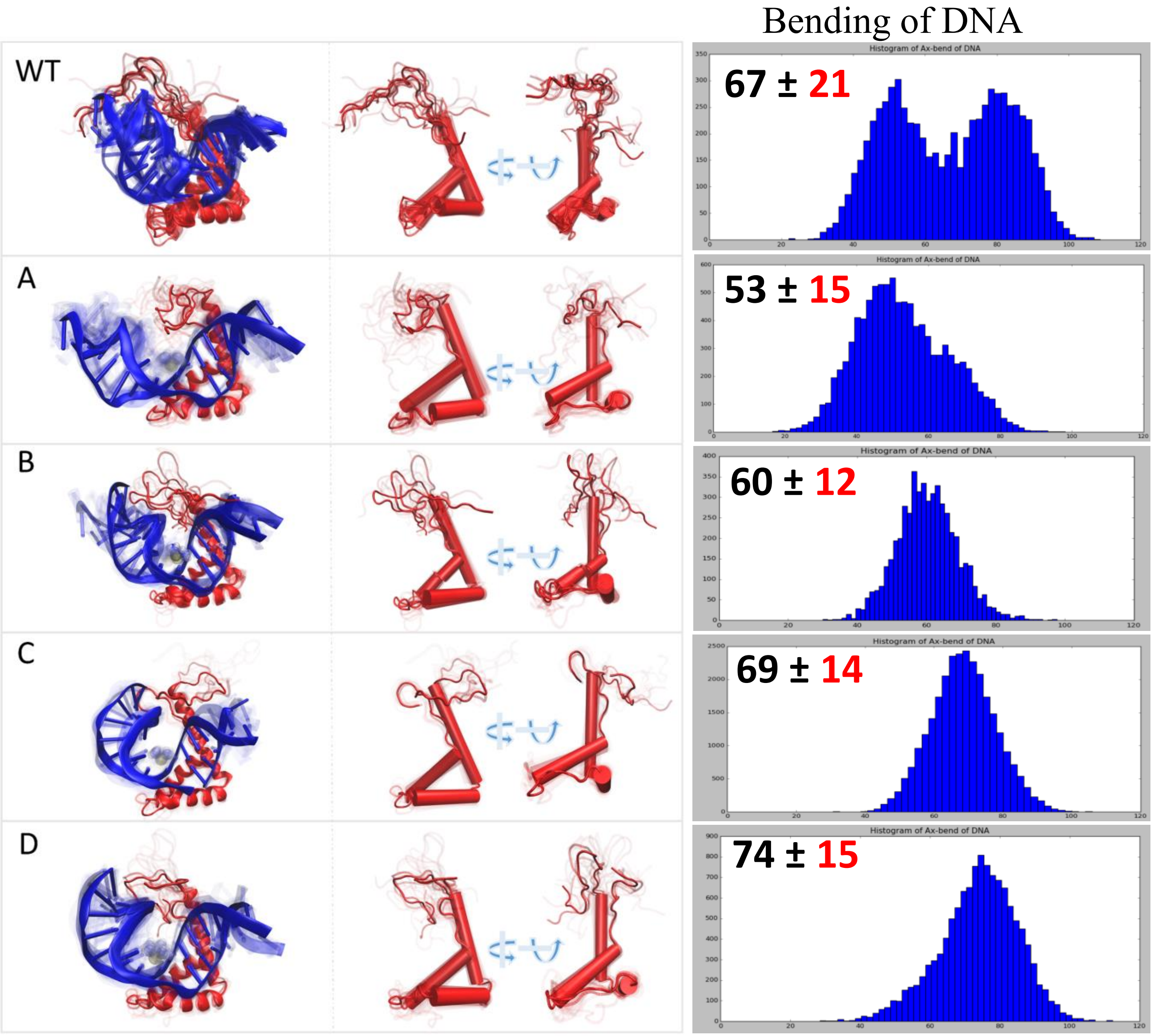
Intermolecular Interaction Networks

Multisite PTMs Cross-talk, etc.

Established HPC-based molecular simulation workflow for protein-DNA interactions investigation.

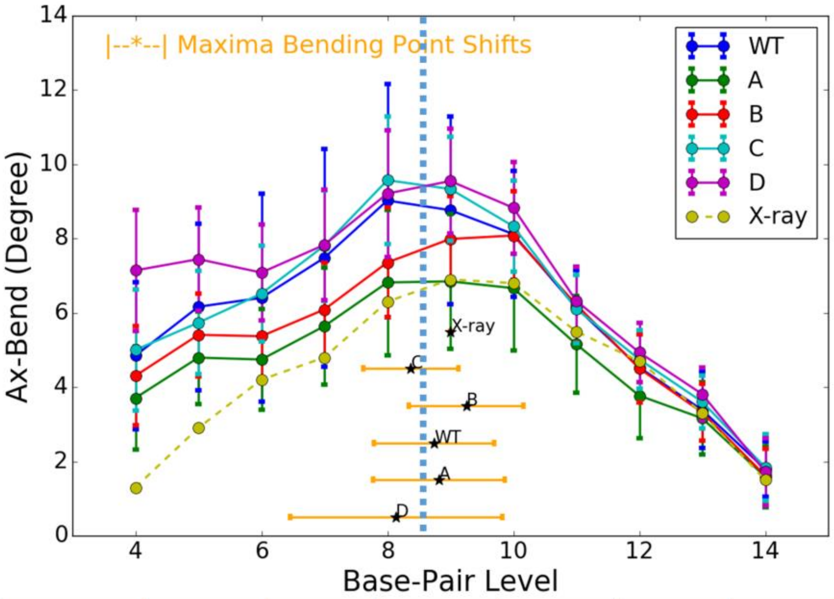
- ✓ *How PTMs regulate protein conformation?*
- ✓ *Will the DNA structure be affected by PTMs?*
- ✓ *What's the difference/ similarity of the DNA-Protein interfaces between wild-type and PTM-modified proteins?*
- ✓ *What's the contributions of these multisite PTMs to the DNA-Protein interactions?*

Protein-Nucleic Acid interactions

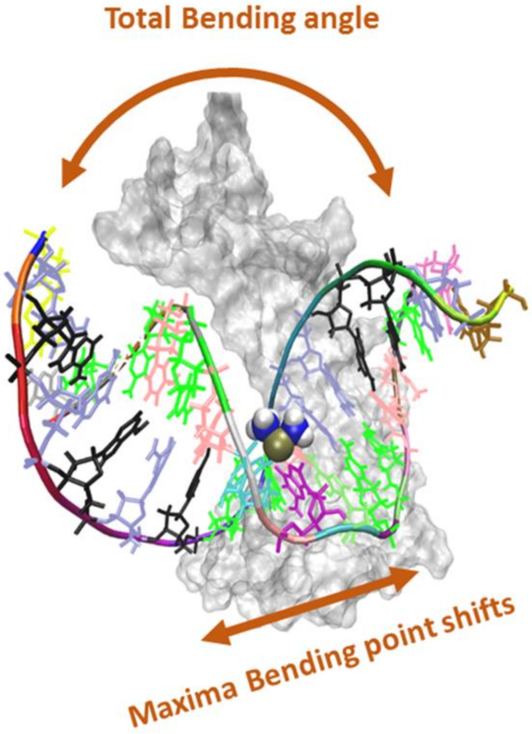


Structure reprensentives of the widetype (WT) and isoforms A , B, C, D proteins using halmitionian replica exchange sampling.

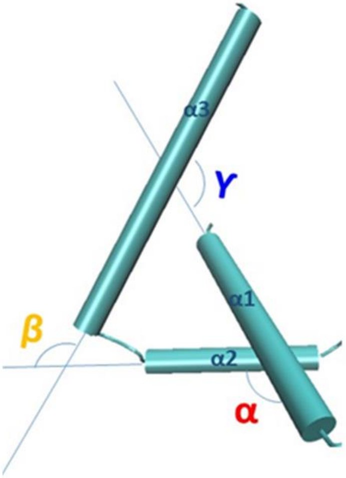
Protein-Nucleic Acid interactions

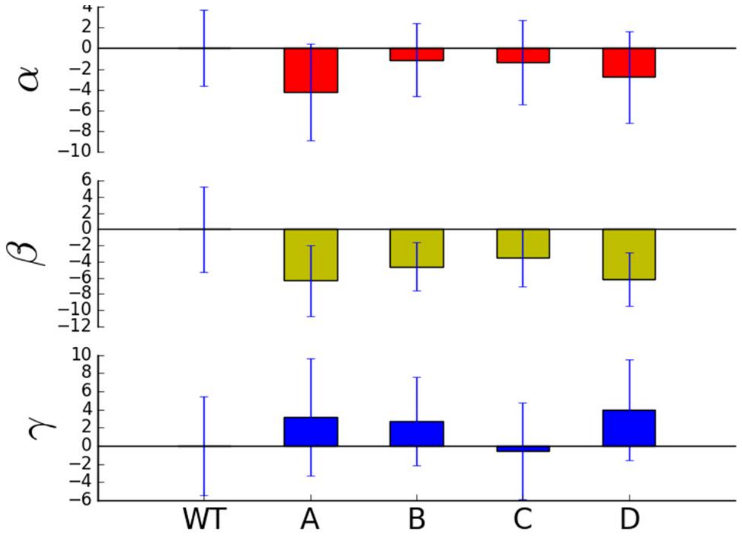


	WT	A	B	C	D	X-ray
Total Bending	67	53	60	69	74	48
Standard Div.	21	15	12	14	15	--

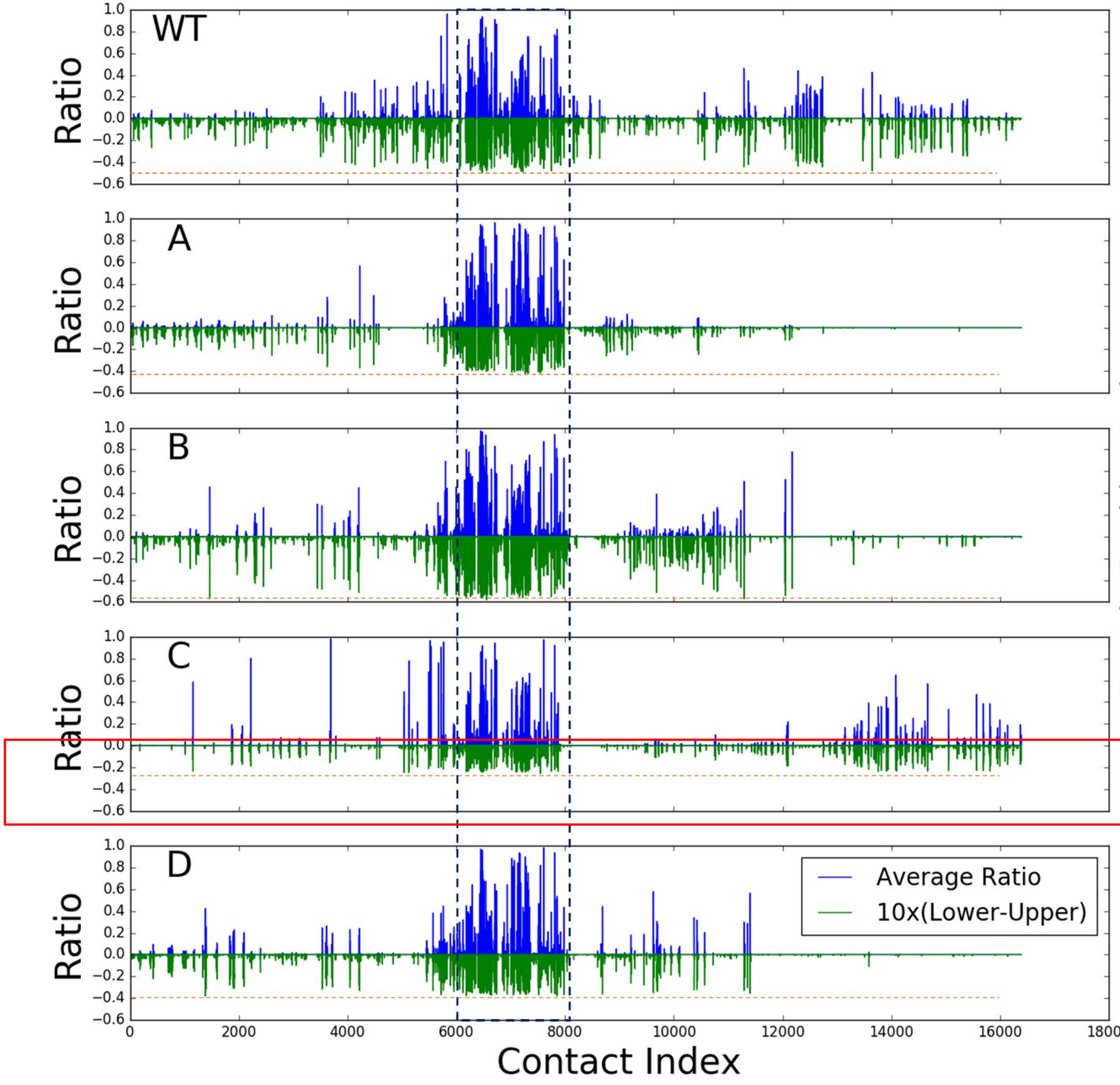


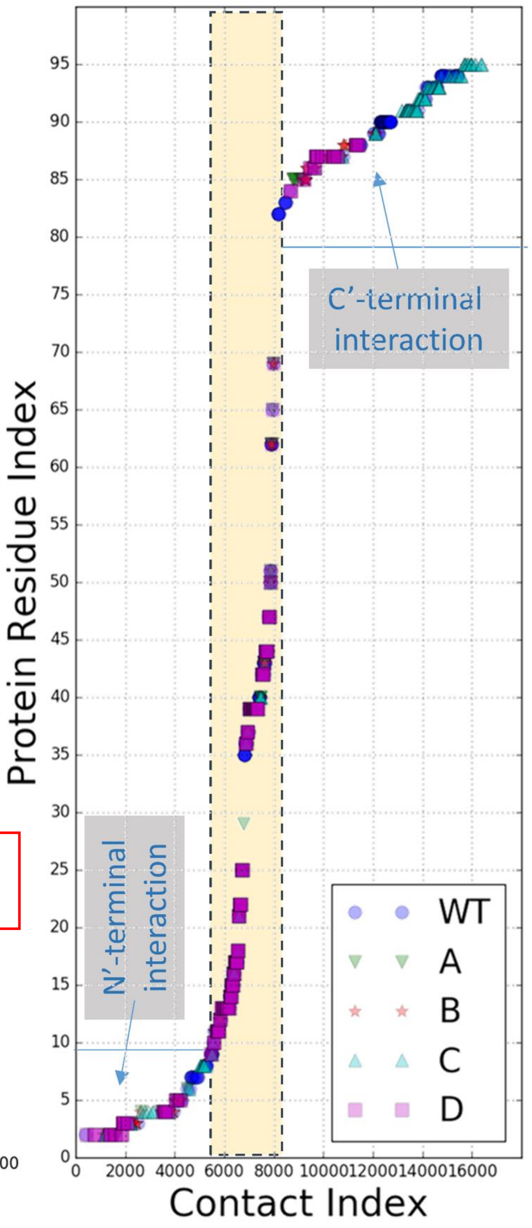
PTMs induced deforming of the L-shape of HMGB1A protein.





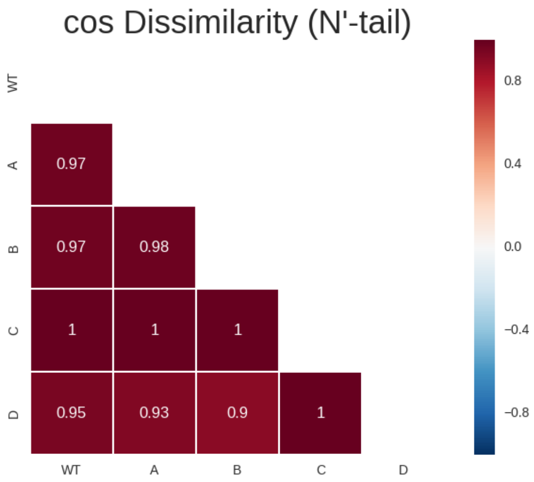
Protein-Nucleic Acid interactions

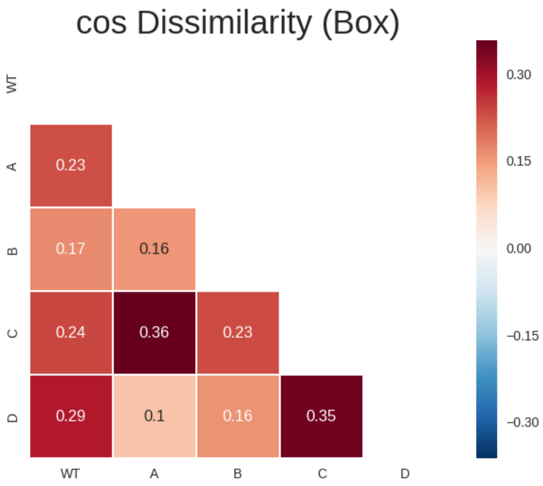


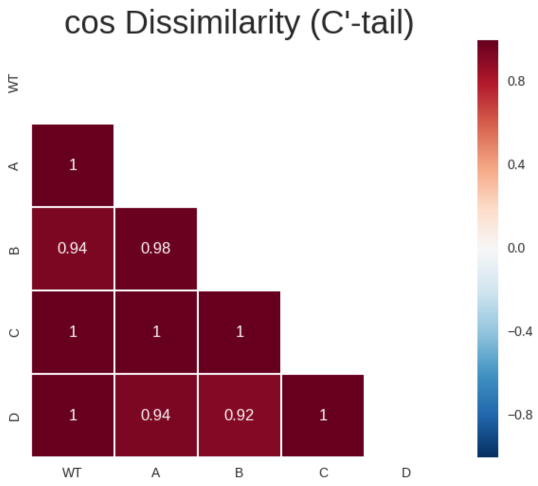


Note: Lower/Upper is the lower/upper bounds of 95% confidence interval

Protein-Nucleic Acid interactions

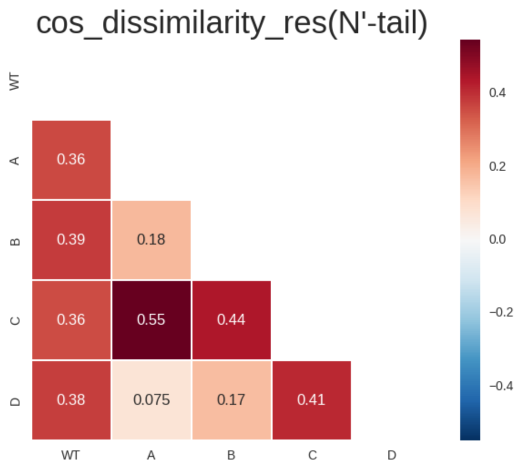

cos Dissimilarity (N'-tail)

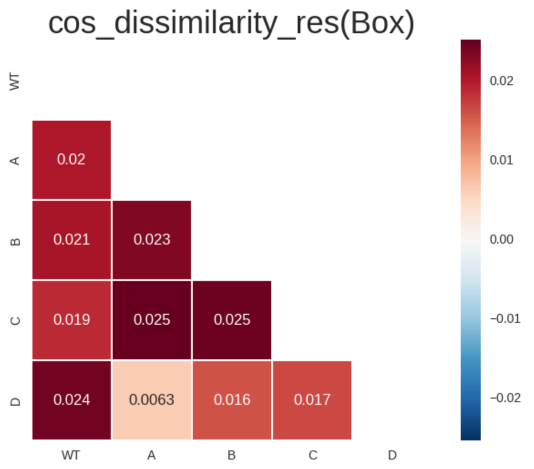

cos Dissimilarity (Box)

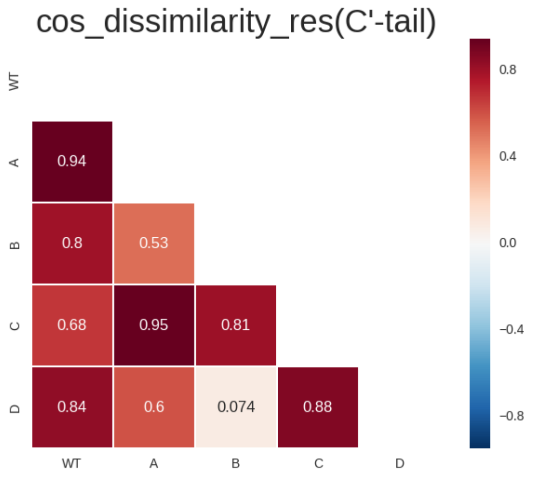

cos Dissimilarity (C'-tail)

Atomic-level

Residual-Level

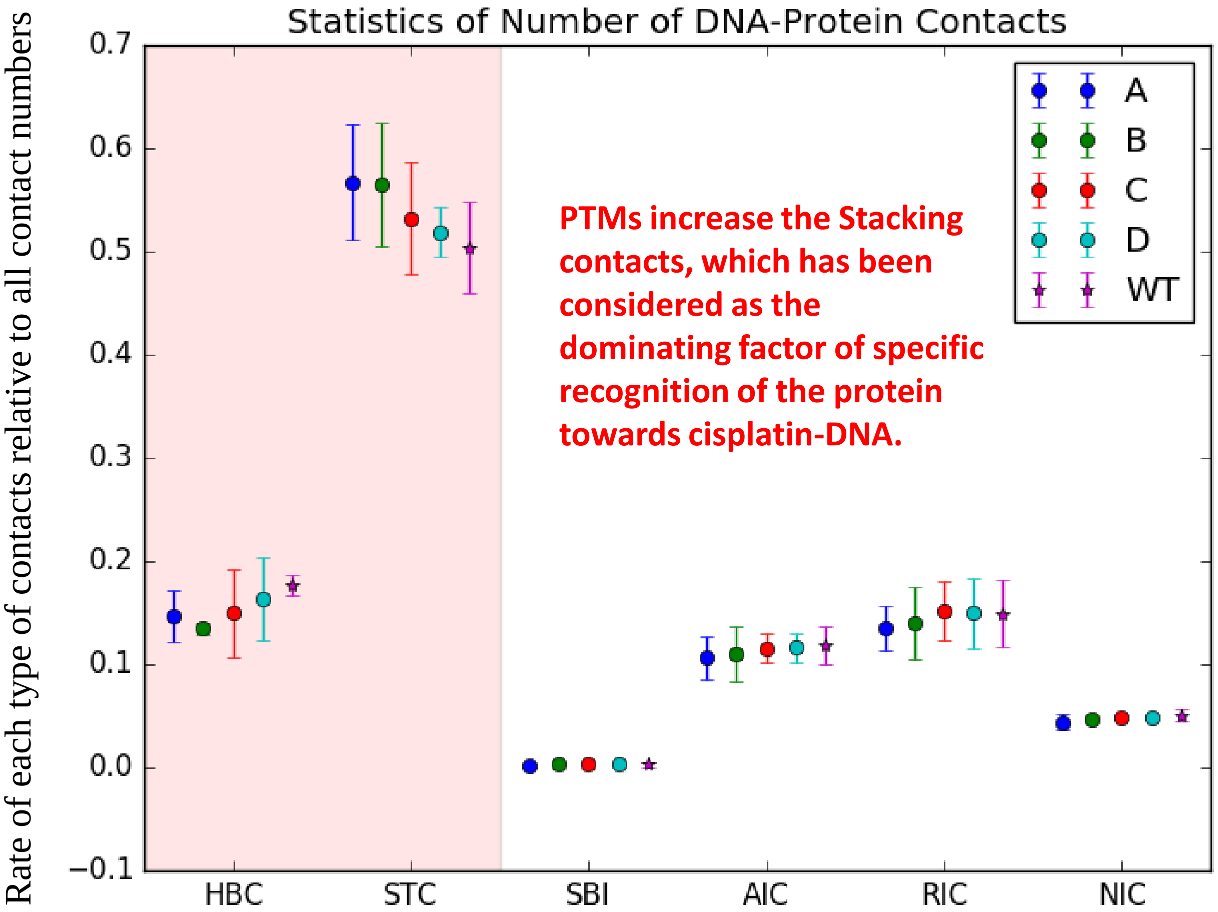

cos_dissimilarity_res(N'-tail)


cos_dissimilarity_res(Box)


cos_dissimilarity_res(C'-tail)

Protein-Nucleic Acid interactions

- ○ Hydrogen bonding contacts (**HBC**)
- ○ Stacking contacts(**STC**)
- ○ Salt bridge contacts (**SBC**)
- ○ Residual attractive contacts (**AIC**)
- ○ Residual repulsive contacts (**RIC**)
- ○ Residual neutral contacts (**NIC**)

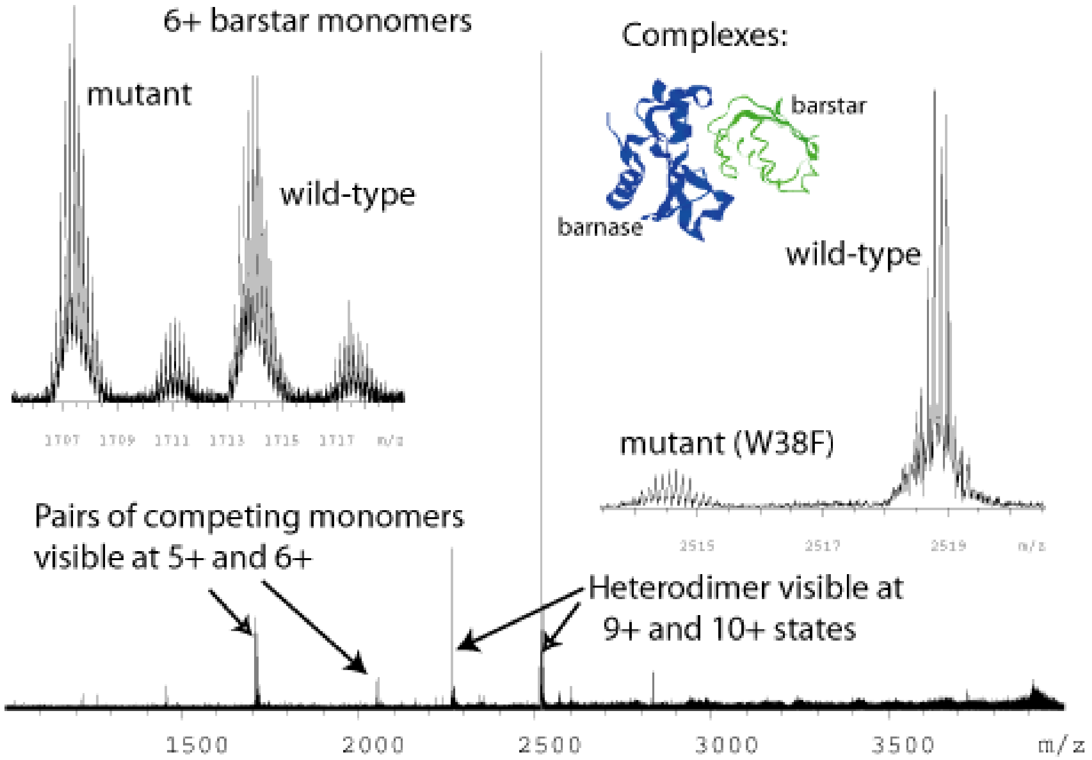


PTMs increase the Stacking contacts, which has been considered as the dominating factor of specific recognition of the protein towards cisplatin-DNA.

Data analysis is in progress, and a manuscript is in preparing.

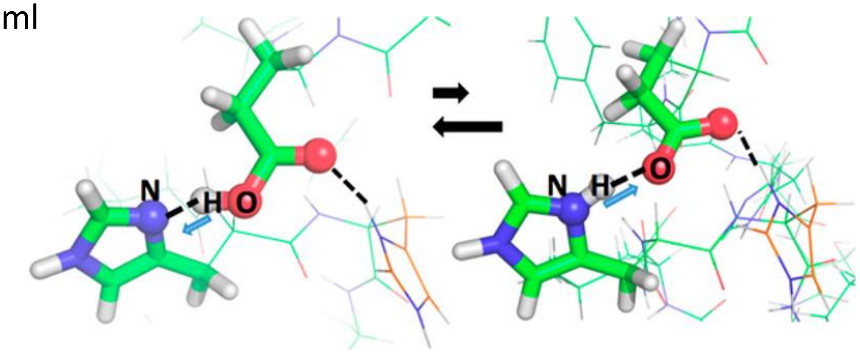
Protein-Protein interactions

Native Mass Spectrometry of Proteins and Protein Complexes



http://www.cchem.berkeley.edu/erwgrp/science.html

- *Do proton migrate from one site to the other?*

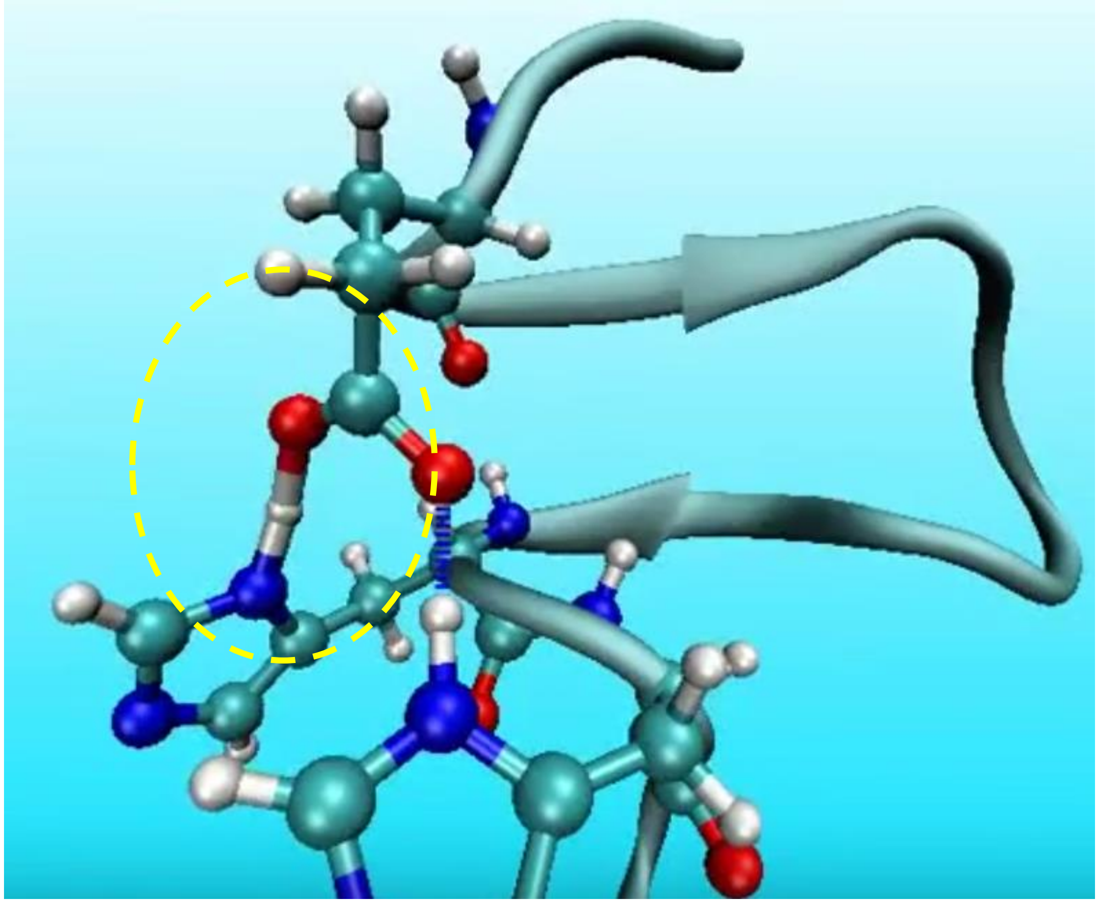


First, QM/MM calculations at B3LYP level were performed on the Aβ(1-16) peptide, to address these fascinating issues.

^{1}DAEFRHDSGYEVHHQK^{16}

Protein-Protein interactions

Intramolecular Proton transfer occurs in ps timescale in MS condition



https://youtu.be/QieEf7DwzxM

Jinyu Li, Wenping Lyu (co-first author) and et. al.
Proton Dynamics in Protein Mass Spectrometry,
J. Phys. Chem. Lett., 2017, 8 (6), pp 1105–1112.

Now, we are studying the intermolecular proton transfer of insulin dimer.

Summary and Perspectives

- HPC-based enhanced sampling: Multisite PTMs on HMGB1A can stabilize the protein-DNA interface in both structure and interaction.

- Hybrid QM/MM calculation: Intramolecular proton transfers has been discovered in the gas-phase (mass spectrometry condition).

- We expect to apply these protocols to investigate intermolecular interactions of protein and DNA in mosquitos regarding malaria.

Acknowledgments

Prof. Dr. Paolo Carloni

Prof. Dr. George Christophides

Jun-Prof. Dr. Giulia Rossetti

Prof. Dr. Giovanni Natile

IAS-5/INM-9 Group members

Financial Supporting:

Horizon 2020 project under the Marie Sklodowska-Curie grant agreement No. 642069.

HPC–LEAP
EUROPEAN JOINT DOCTORATES

Computational and experimental resources:

JURECA at Jülich Supercomputing Centre (JSC)
MareNostrum III at Barcelona Supercomputing Center (BSC)

JÜLICH
FORSCHUNGSZENTRUM

BSC
Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Hercules foundation for funding of the Synapt G2 instrument.

Thank you for your attention!