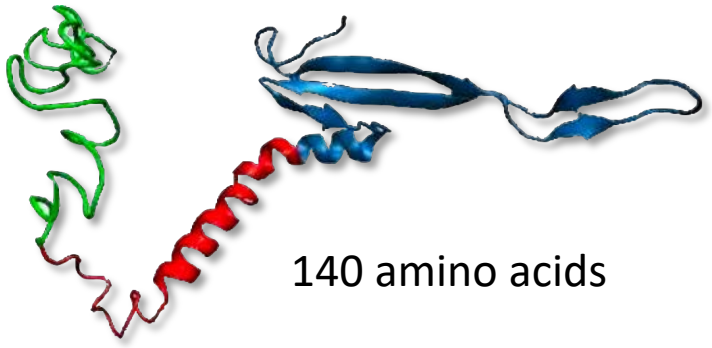
Effective Free-energy landscape of an Intrinsically Disordered Protein: α-synuclein

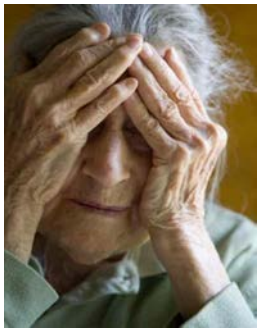
Patrick Senet

psenet@u-bourgogne.fr

α-synuclein



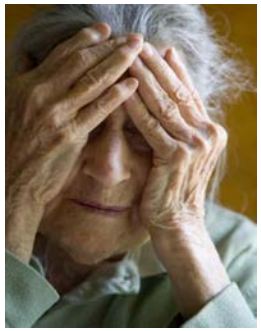
140 amino acids



Parkinson Disease

Lewy body Dementia

RESEARCH NEEDED ON THE SYNUCLEOPATHIES

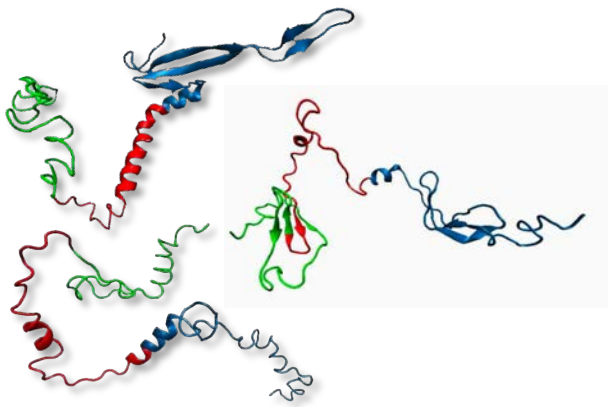


Parkinson Disease
200000 patients (France)
25000 increase/year
Incurable
Lewy Body Disease
Idem

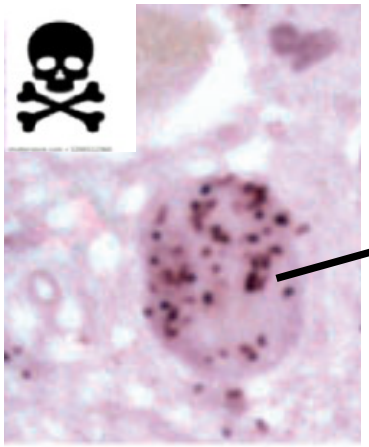
What we know

Related to **α-synuclein**

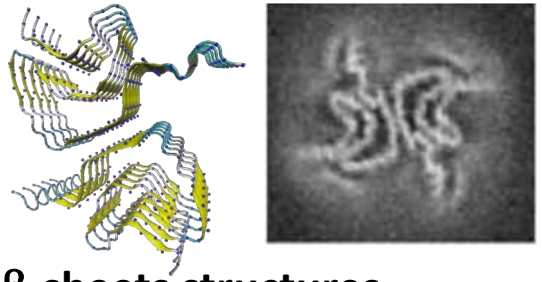
Occurs as a **disordered monomer (IDP) in the brain (normal state)**



What we know



α-synuclein aggregation in neurons (fibrils)



β-sheets structures amyloids

Familial mutations: A30P, E46K, A53T

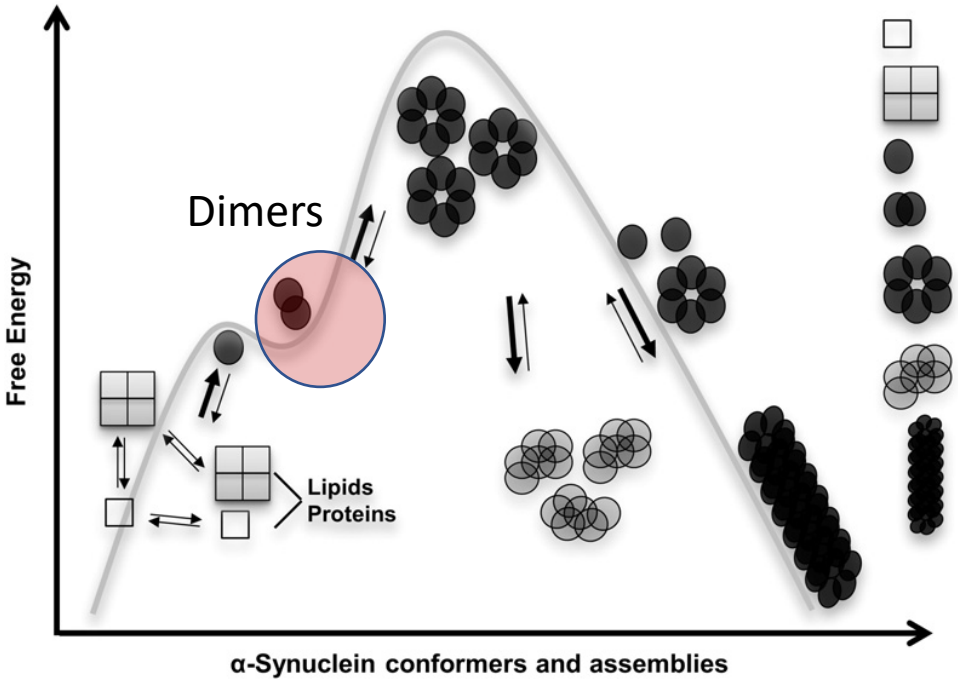
*What we **don't know:***

Why and how α-synuclein aggregates ?
Does aggregate size matters for toxicity?
How mutations modify the conformations?

Still fundamental questions to answer…

SMALL MULTIMERS/OLIGOMERS OF α-SYNUCLEIN PLAY A CRUCIAL ROLE

Simplified model *



Monomers

Tetramer and multimers

Aggregation competent monomers

Dimers

β-Sheet rich soluble oligomers

Non-amyloid aggregates

Typical amyloid fibrils

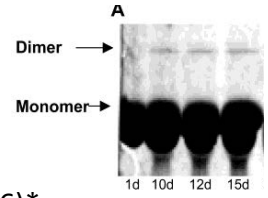
Toxic ? (highly probable)

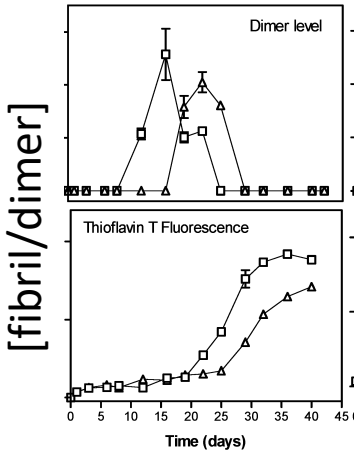
Toxic ? **(unclear)**

Structures

???

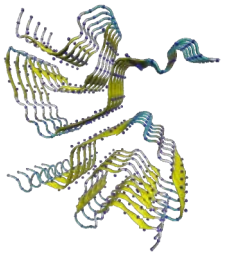
In vitro kinetics





Dimer

Fibrils



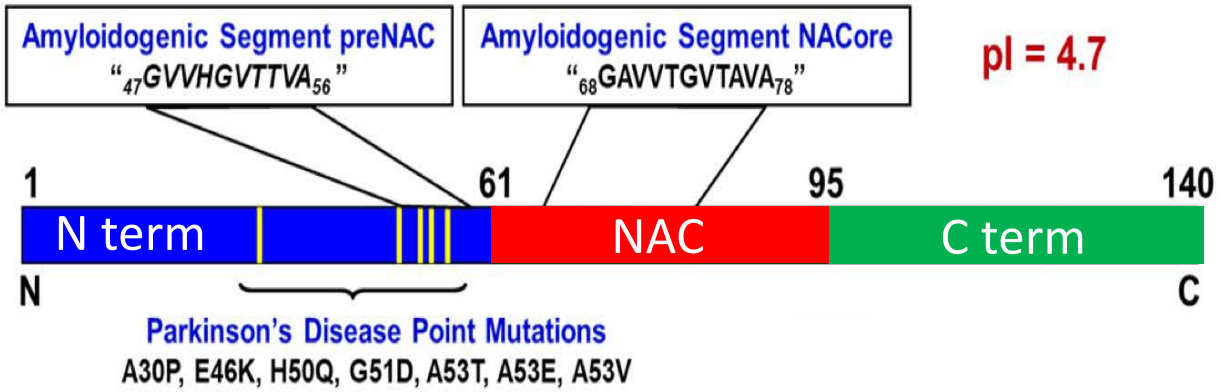
D.E. mor et al., Neurobiology of disease, 88, 66 (2016)*

Roostaee et al. Molecular Neurodegeneration, 8, 5 (2013)

S. Krishnan et al., Biochemistry 42, 829 (2003)

AIMS OF THE PRESENT STUDY

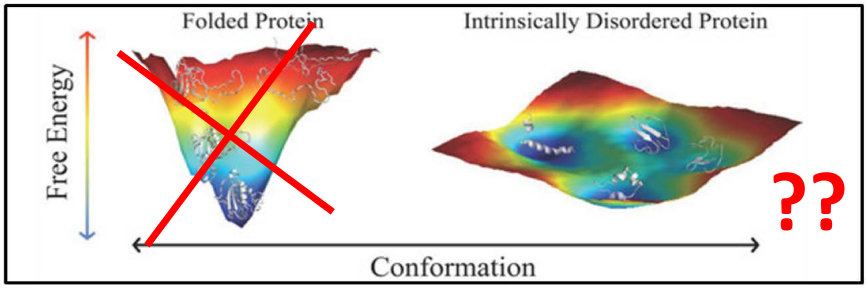
From α-synuclein sequence



5 complete KTKEGV motifs

Lucas HR and Fernandez RD, Neural Regen Res 15(3):407-415 (2020)

To (effective) free-energy landscapes (**FEL**) from molecular dynamics

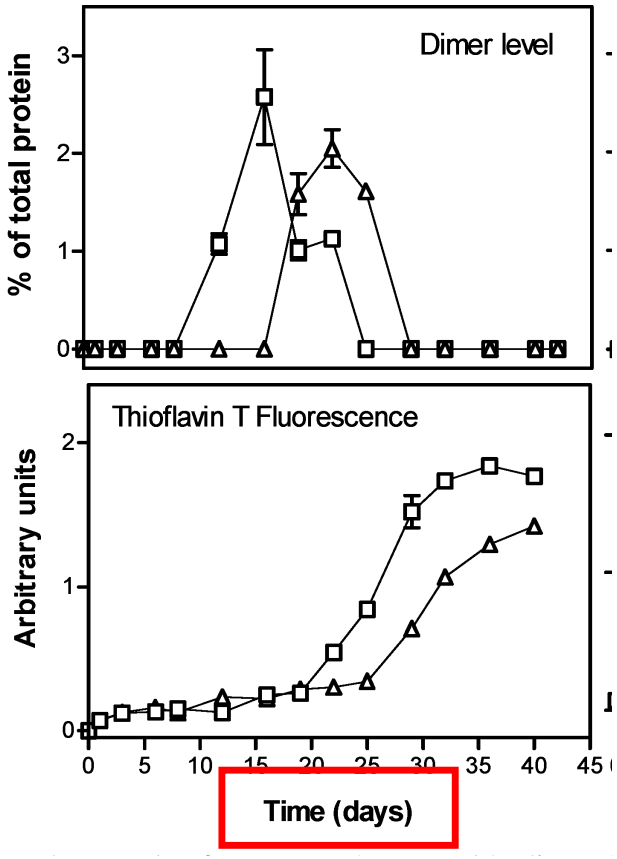


How to quantify/characterize the FEL of the free monomer (IDP) ?

How to quantify/characterize the FEL of the dimers ?

How the mutations influence the conformations and the FELs ?

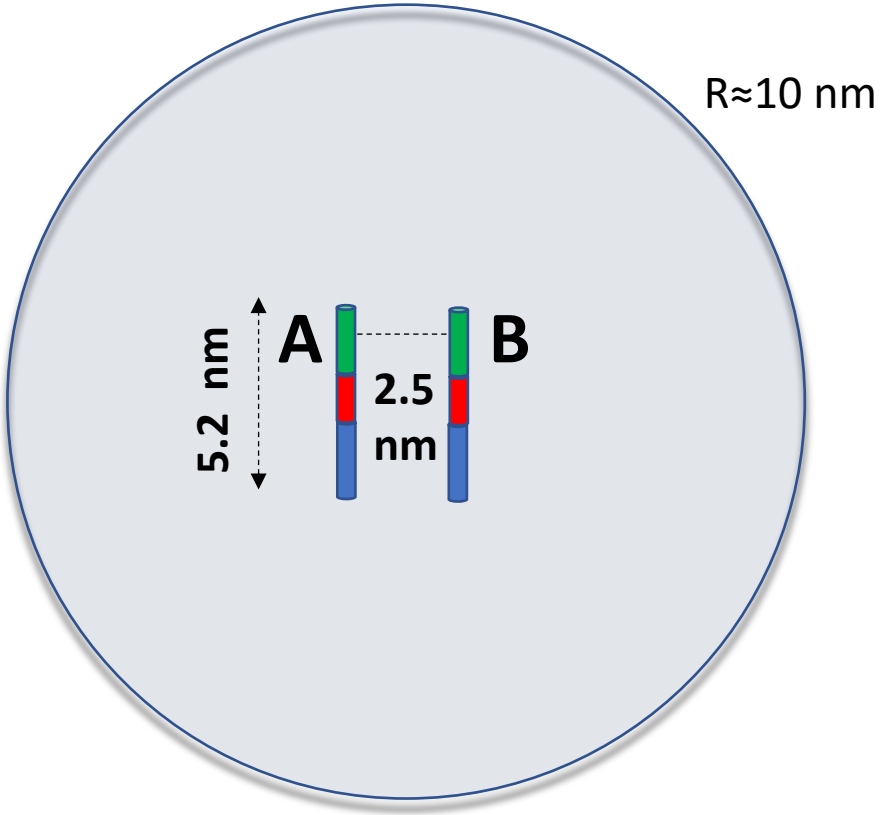
MODEL & METHODOLOGY



IN VITRO EXPERIMENT

[α-synuclein] in vitro ≃ **5 mg/ml**
[α-synuclein] in the brain (estimated) ≃ **0.09 mg/ml**

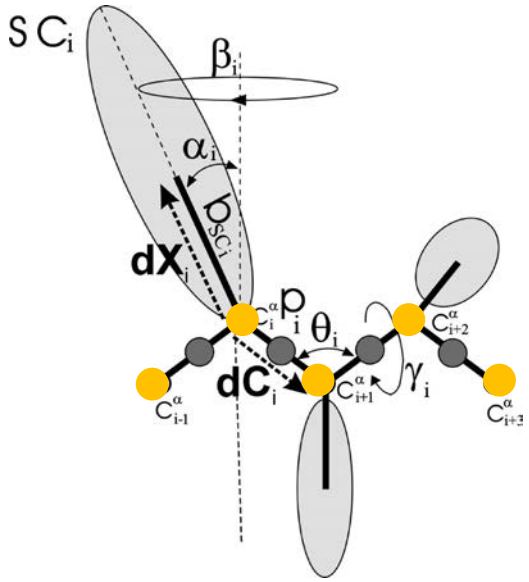
Biased initial conditions of the present simulations for WT & mutants



α-synuclein concentration in present simulations ≃ **14 mg/ml**
Limited diffusion (confinement)

MODEL & METHODOLOGY – UNRES COARSE GRAINED MODEL

Coarse-grained model
United RESidue (UNRES)



Maisuradze GG et al., J. Phys. Chem. A 114, 4471 (2010)
Liwo A et al., J. Chem. Phys. 150, 155104 (2019)

http://www.unres.pl

C^{α} and peptide interaction centers
CG angles α, β, θ, γ

UNRES Force-Field
Implicit solvent
Energy function (PMF)
built from all-atom MD trajectories

Sampling of the conformational space

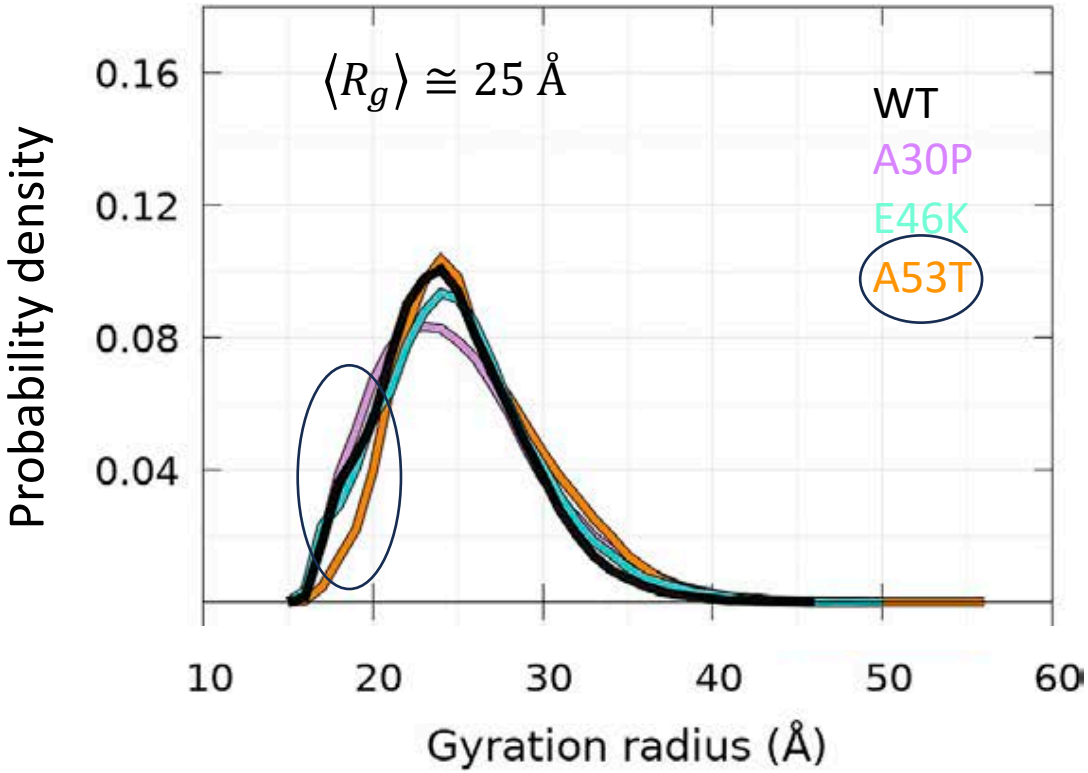
Langevin thermostat (effective time step 4.9 ps)

Replica Exchange Molecular Dynamics
72 trajectories
32 trajectories 300K
5x8 trajectories at 310K, 323K, 337K, 353K, 370K

Applied for WT, A30P, E46K, A53T

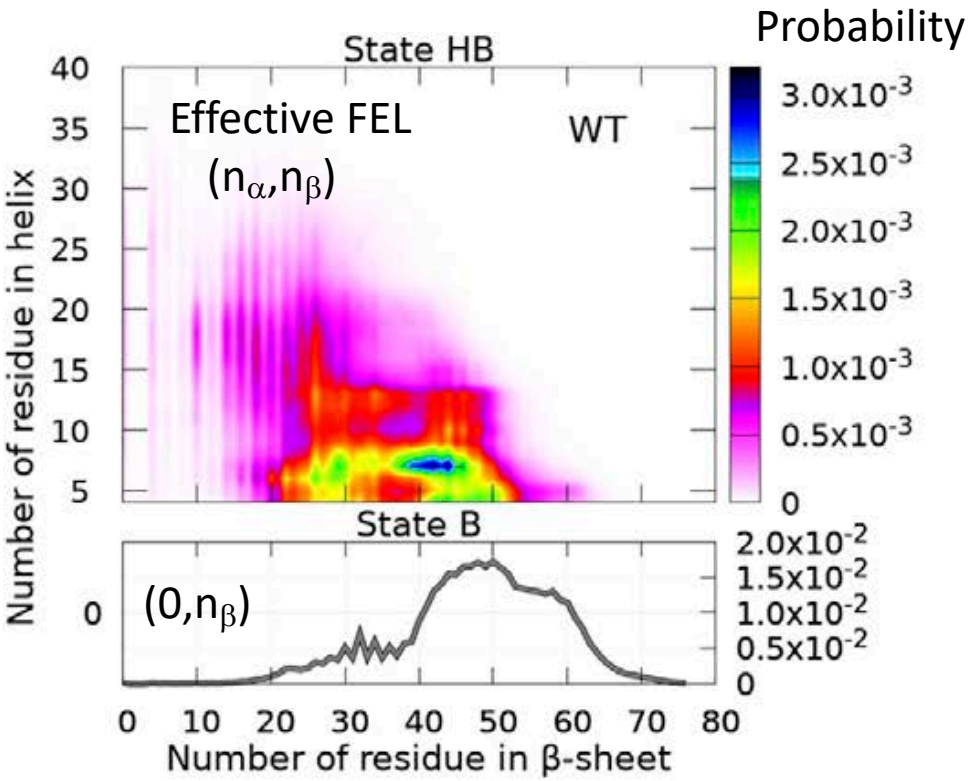
RESULTS – MONOMERS – A GLOBAL STRUCTURAL ORDER PARAMETER: Rg

α-synuclein monomer is an IDP

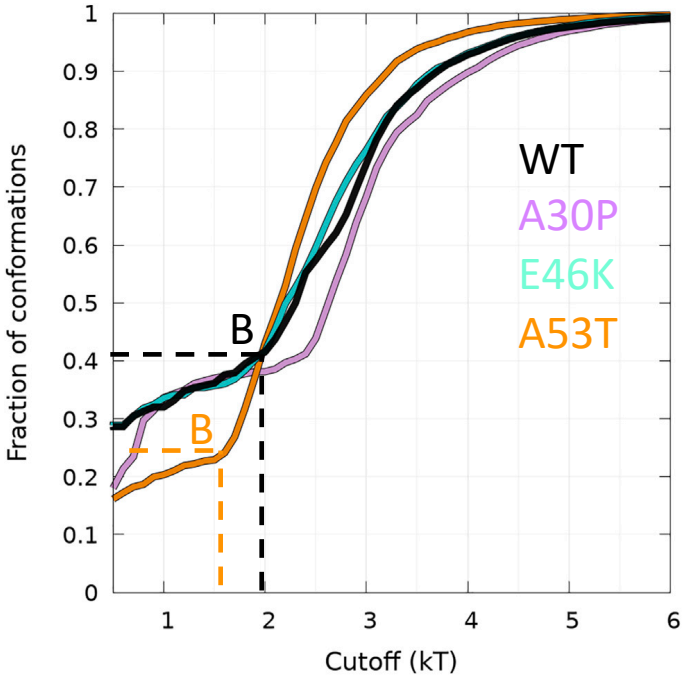


Experimental values (SAXS) depend on buffer solution
*Experimental value extrapolated at infinite dilution for **WT**:* $\langle R_g \rangle \cong 27$Å

RESULTS – MONOMERS – LOCAL STRUCTURE DERIVED ORDER PARAMETERS n_α, n_β

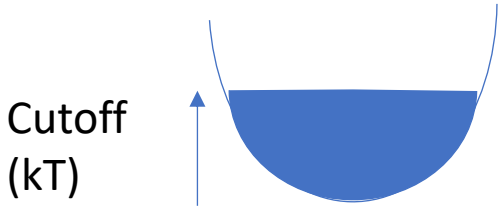


Two main « states » : HB & B revealed by the DOS in the FEL



*Experimental values **WT** extracted from experiments vary a lot*

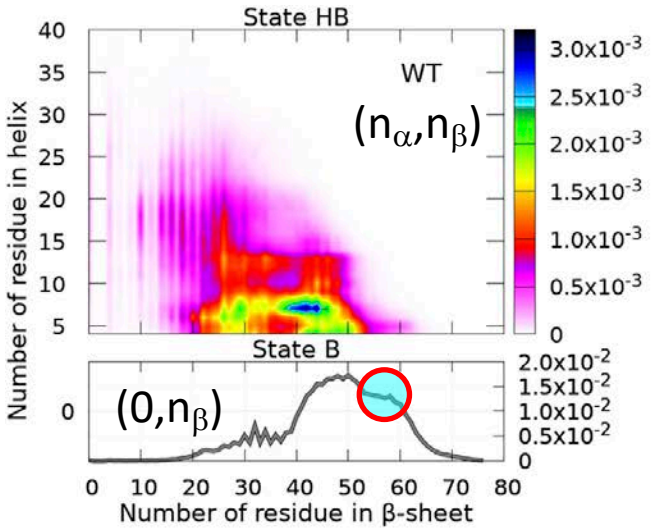
<α-helix> 3-6 residues
<β-sheets> 15-42 residues

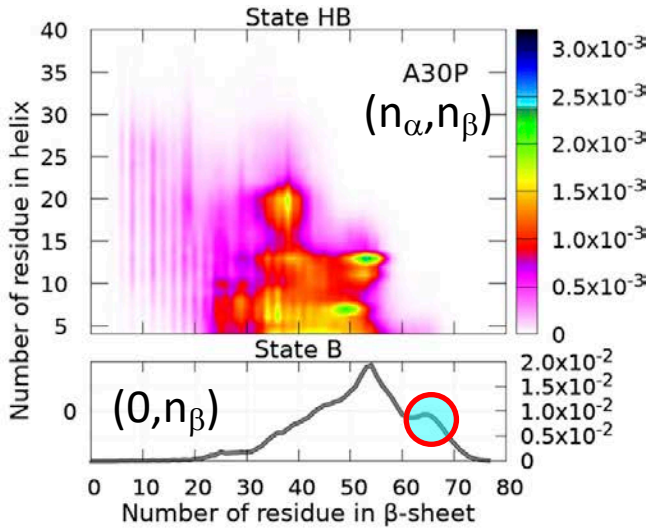


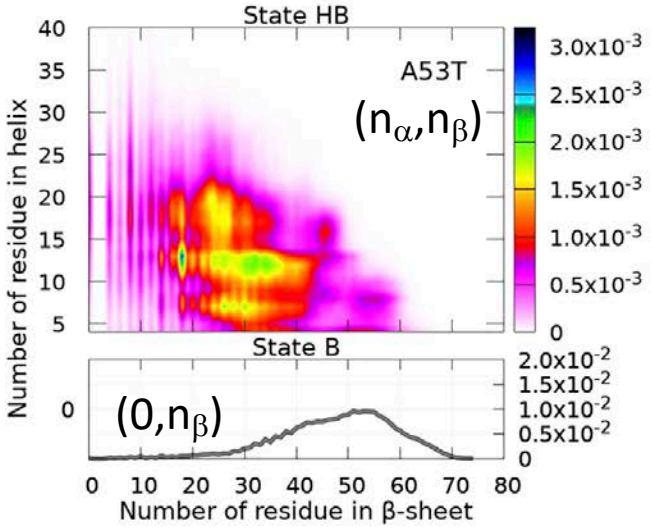
Minimum of (n_α,n_β) FEL

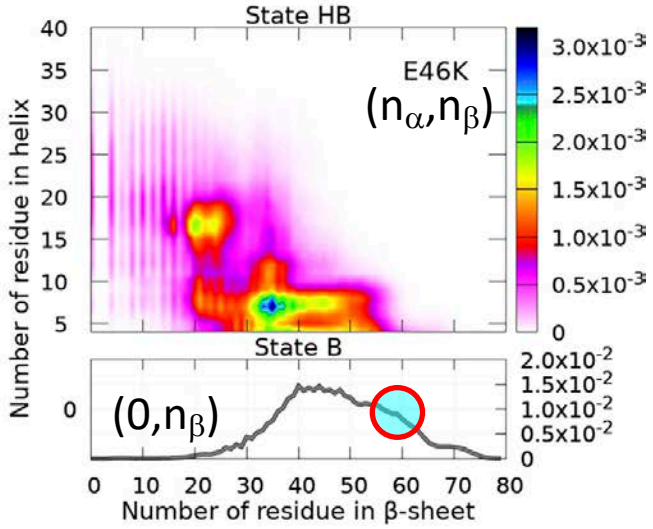
RESULTS – MONOMERS – 1D PROFILE (B state) & 2D LANDSCAPES (HB state)

1 single amino-acid substitution induces subtle differences

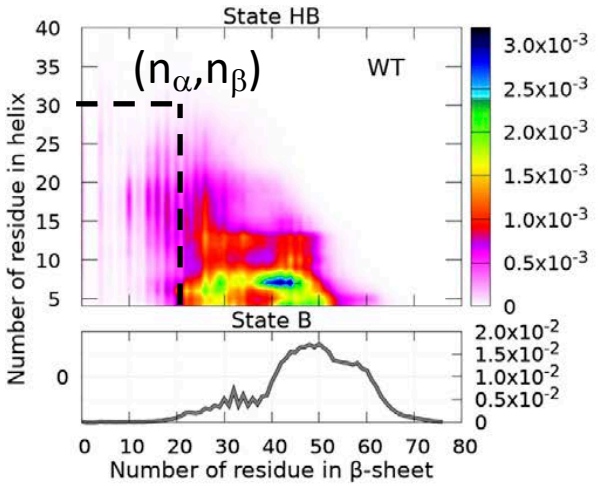






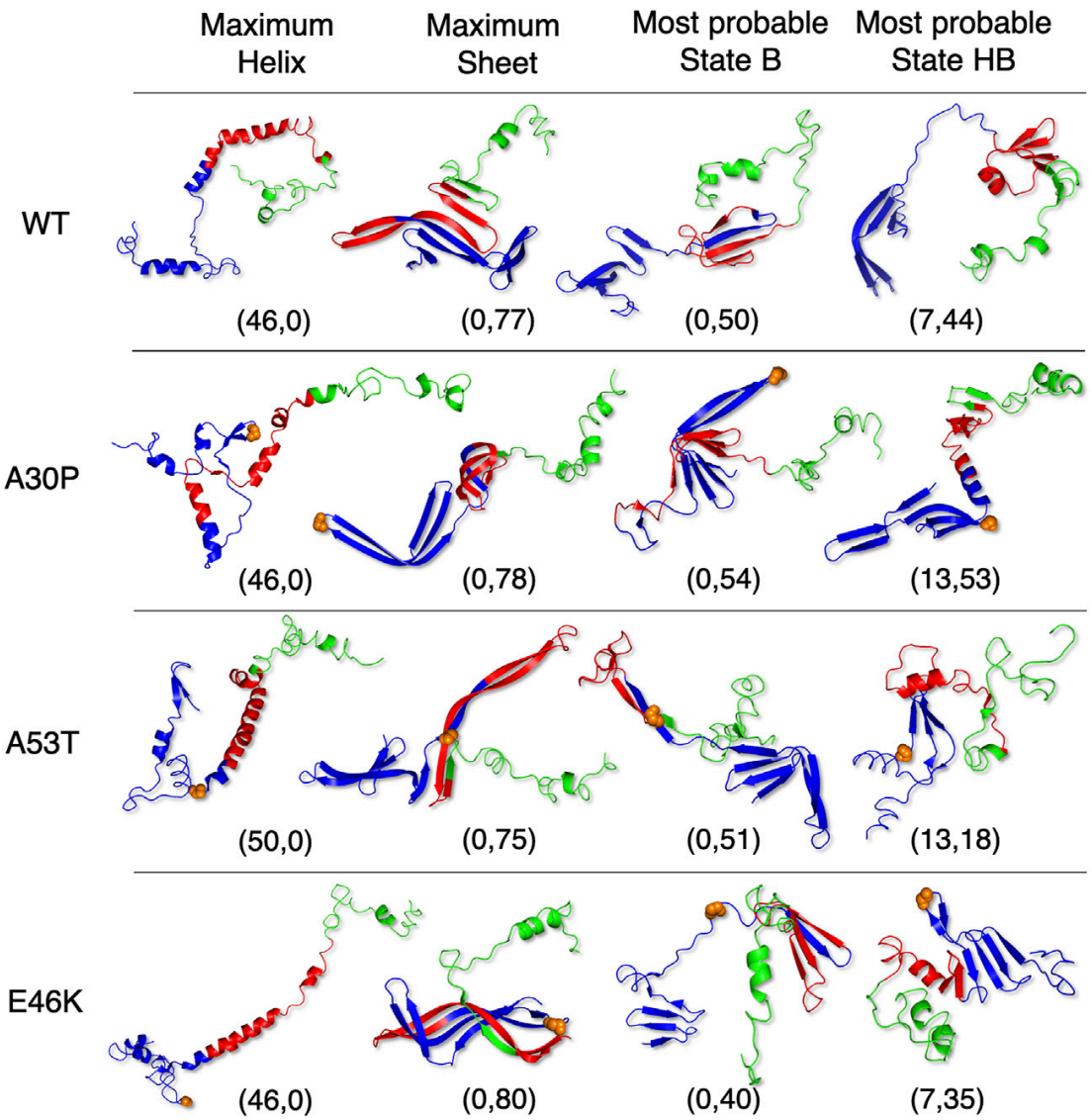


RESULTS – MONOMERS – LABELLING OF ENSEMBLE OF MICRO-STATES

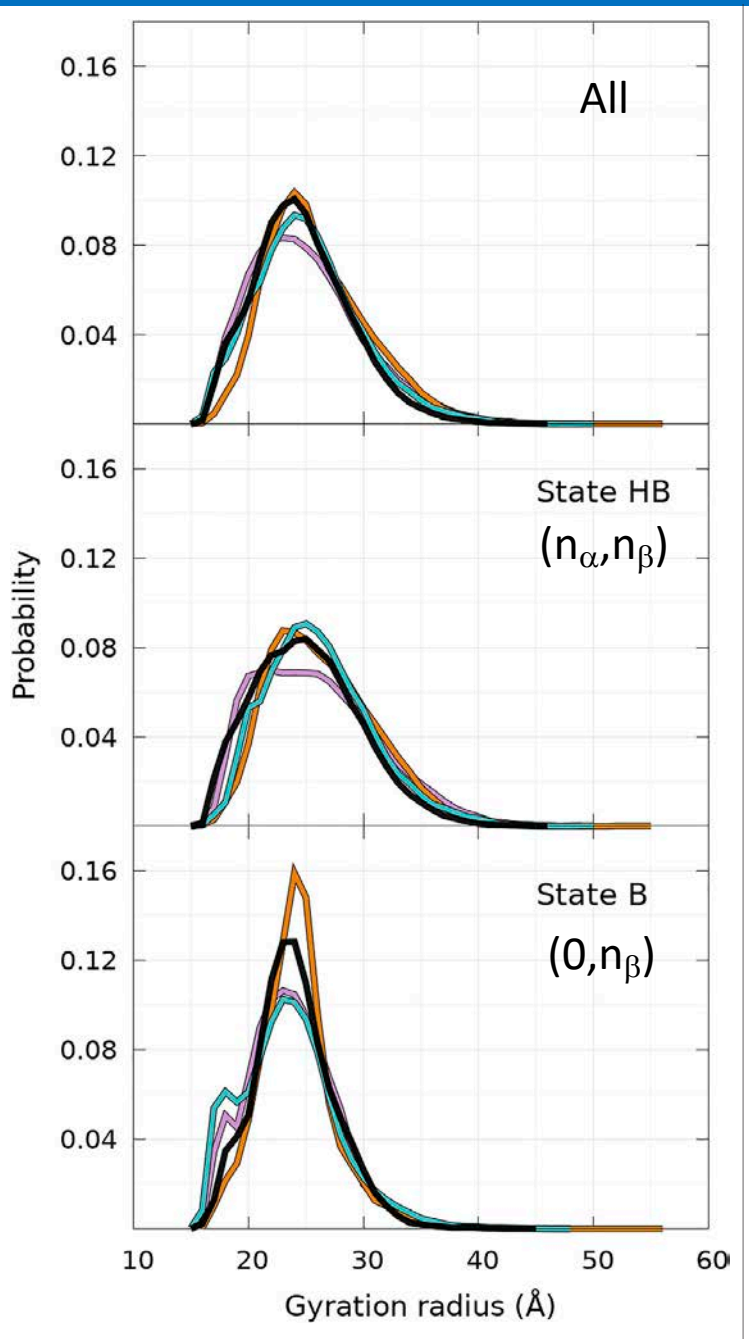


On going (EXP):
Single-molecule
Raman mapping
Amide bands
$\alpha \simeq 1652\ cm^{-1}$
$\beta \simeq 1615\ cm^{-1}$ - $1637\ cm^{-1}$

Illustration of how diverse are the IDP micro-states



RESULTS – MONOMERS – COMBINATION OF ORDER PARAMETERS

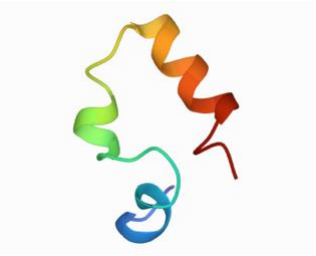


Distinct peak of very compact structures for WT, A30P, E46K

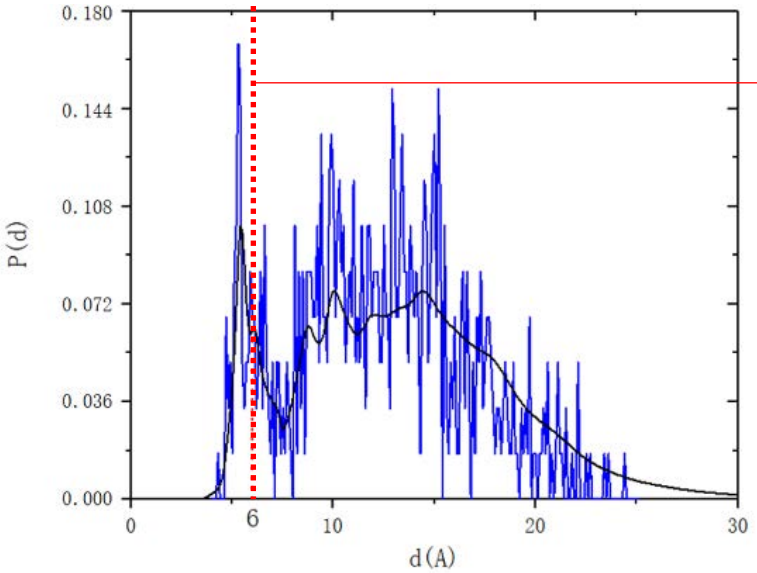
Structural analysis reveals contacts N-term <-> C-term

On going work

1. A protein graph (PG) is built from a 3D structure

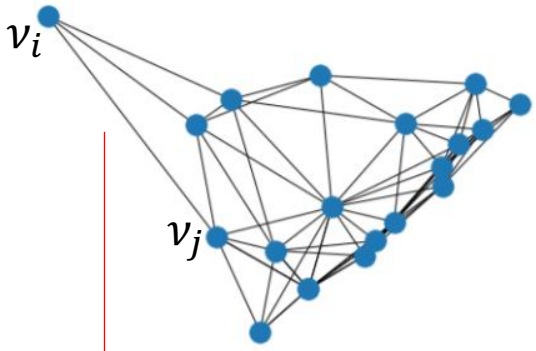


Example
HP-36 protein



Pair distribution function

Link between 2 nodes = contact < 6 Å



Laplacian matrix

$L_{ii} = \deg(v_i)$

$L_{ij} = -1 \; \forall \; i \neq j$ if neighbors and 0 if not

$$L_{ij} = \sum_{\alpha} \lambda_\alpha \, e_\alpha(i) e_\alpha(j)$$

2. Topological descriptors are computed for each graph (=each structure)

Global force constant K

$$\frac{1}{K} = \sum_{\alpha}{}' \frac{1}{\lambda_\alpha}$$

Related to the network criticality in a communication Network (measure of the ribustness of the network)

$$C = \frac{2N}{K}$$

Average shortest path length

$$l = \frac{1}{N(N-1)} \sum_{v_i v_j \neq v_i} d(v_i v_j)$$

Related to the Wiener index w

$$l = \frac{2\,w}{N(N-1)}$$

3. Free-energy landscape (K,l) for α-synuclein monomer

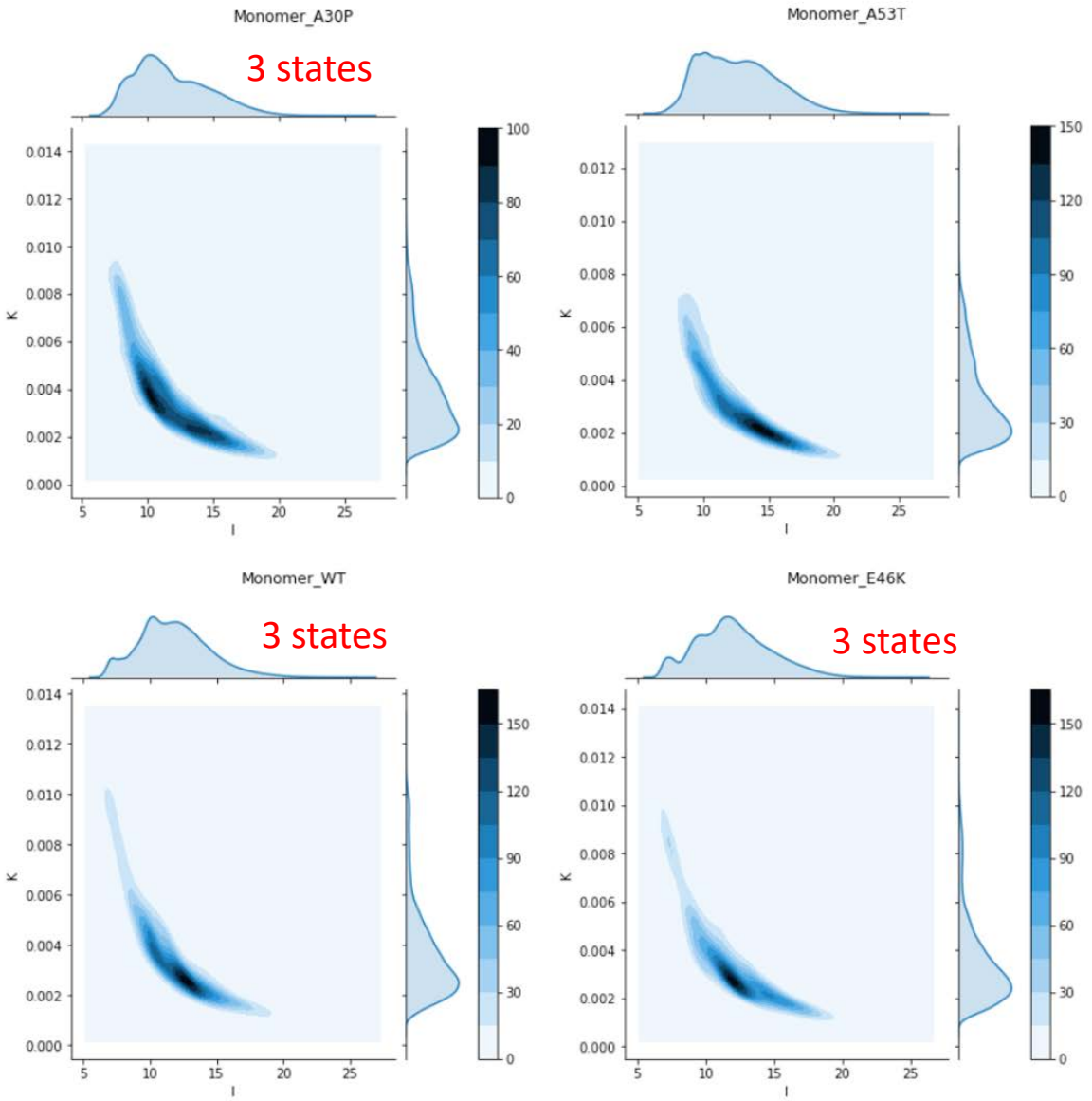
K & l are inversely related !
(universal for proteins)

Analytical result for an unfolded protein

$$\frac{1}{K} = \frac{3}{2} l^2 - l$$

On going work
Local descriptors

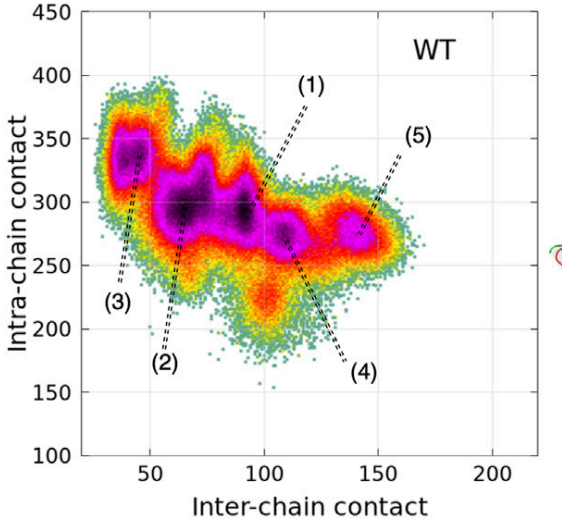
Graph of PG



MAIN MESSAGES FROM THE MONOMER SIMULATIONS

- IDP (R_g)
- Two « states » : B & HB (n_α,n_β)
- Two « states » of Rg in the B state (n_α,n_β + Rg) except A53T
- PDF of the topological descriptor l shows 3 different sub-states except A53T (further analysis with graph is on going)
- The global force constant is inversely related to l (general)
- Complexity of the FELs -> multidimensional approach to pursue Experimental test – single-molecule spectroscopy…

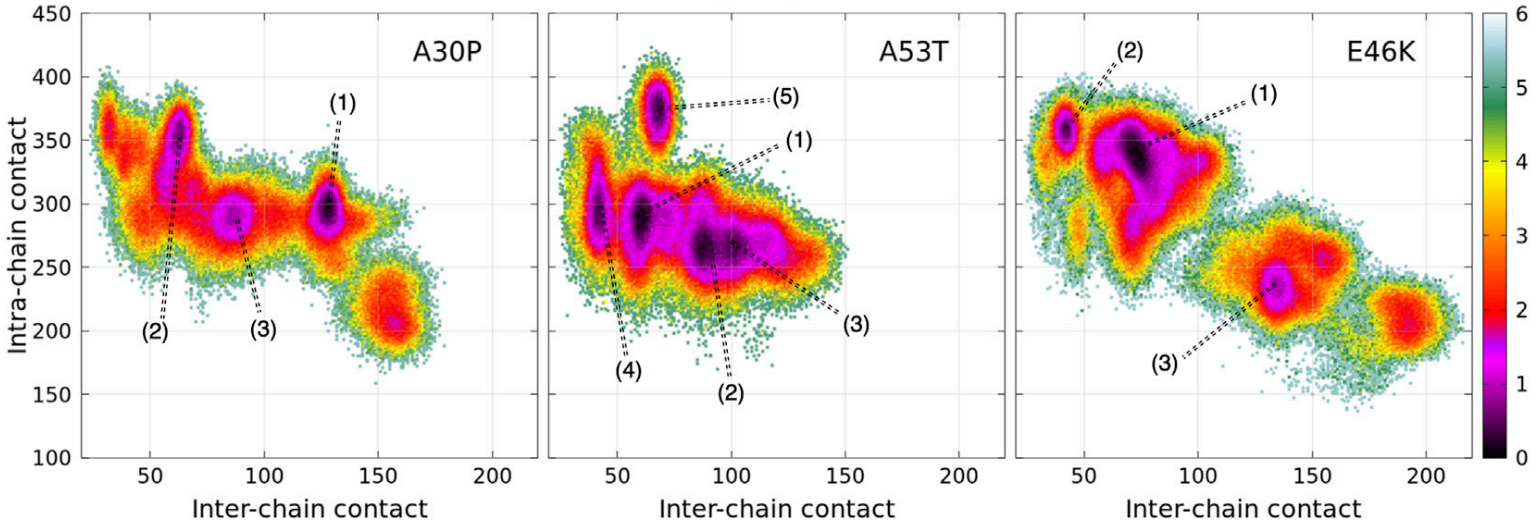
CONTACTS (C^{α}-C^{α} distance < 5 Å)



PMF wells < 1 kT

TABLE 1 | Effective (dimensionless) free-energy difference ($-ln[\frac{P_i}{P_1}]$), where P_1 and P_i are the probabilities of the minimum 1 and the of *i*th minima shown in **Figure 1** for the WT and the variants.

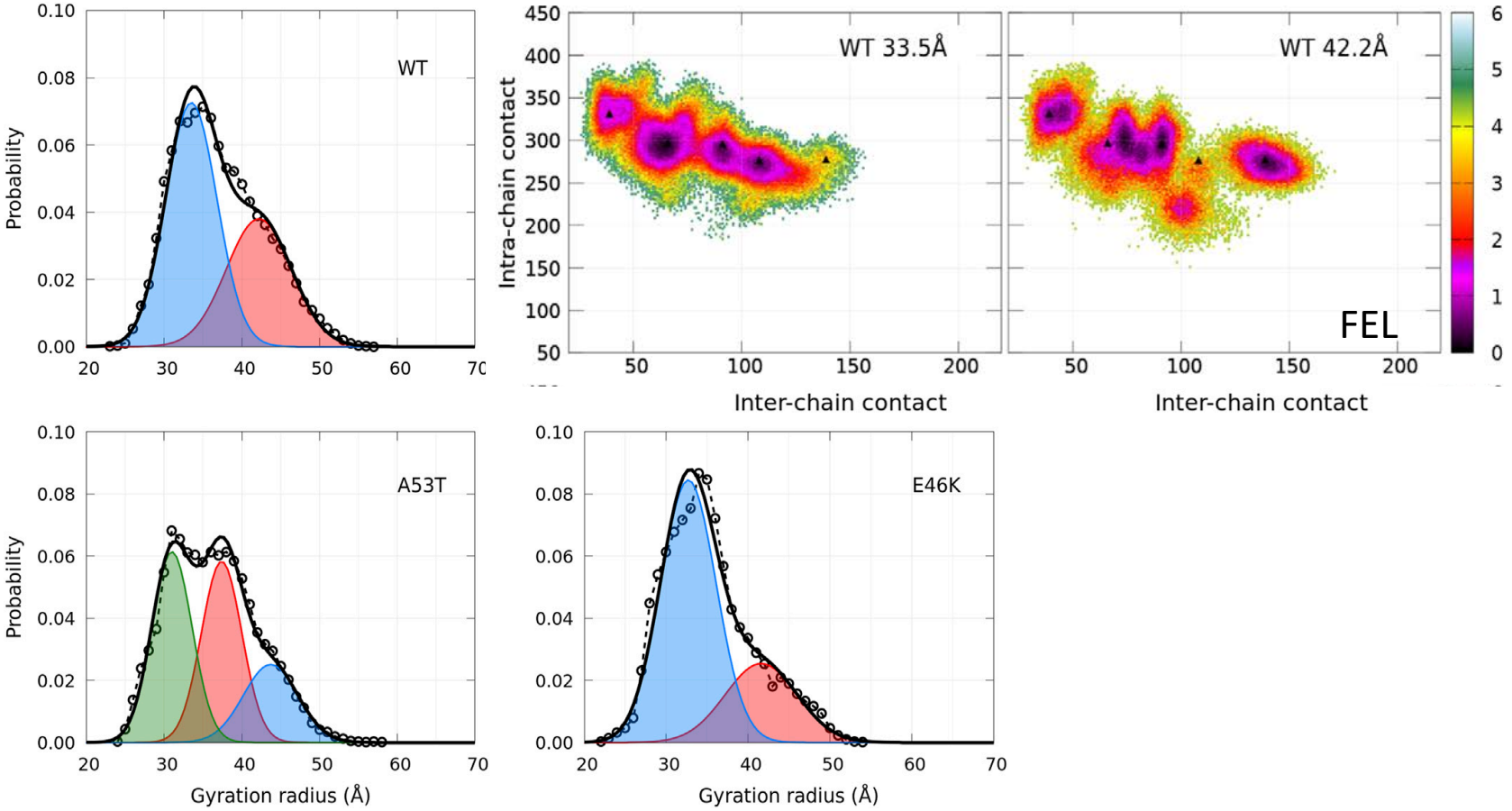
Protein	Min 2	Min 3	Min 4	Min 5
WT	0.08	0.39	0.40	0.84
A30P	0.42	0.54	-	-
A53T	0.04	0.12	0.14	0.24
E46K	0.32	0.72	-	-



RESULTS – PART 1 DIMERS - COMPLEXITY OF THE FEL (II, Rg)

Complexity of the FEL (II) R_g

Hides a large diversity



Multimodal $Maxima\ R_g \approx 3.4 - 4.8\ nm$

$\langle R_g \rangle \approx 3.5 - 3.7\ nm$

No SAXS experimental data

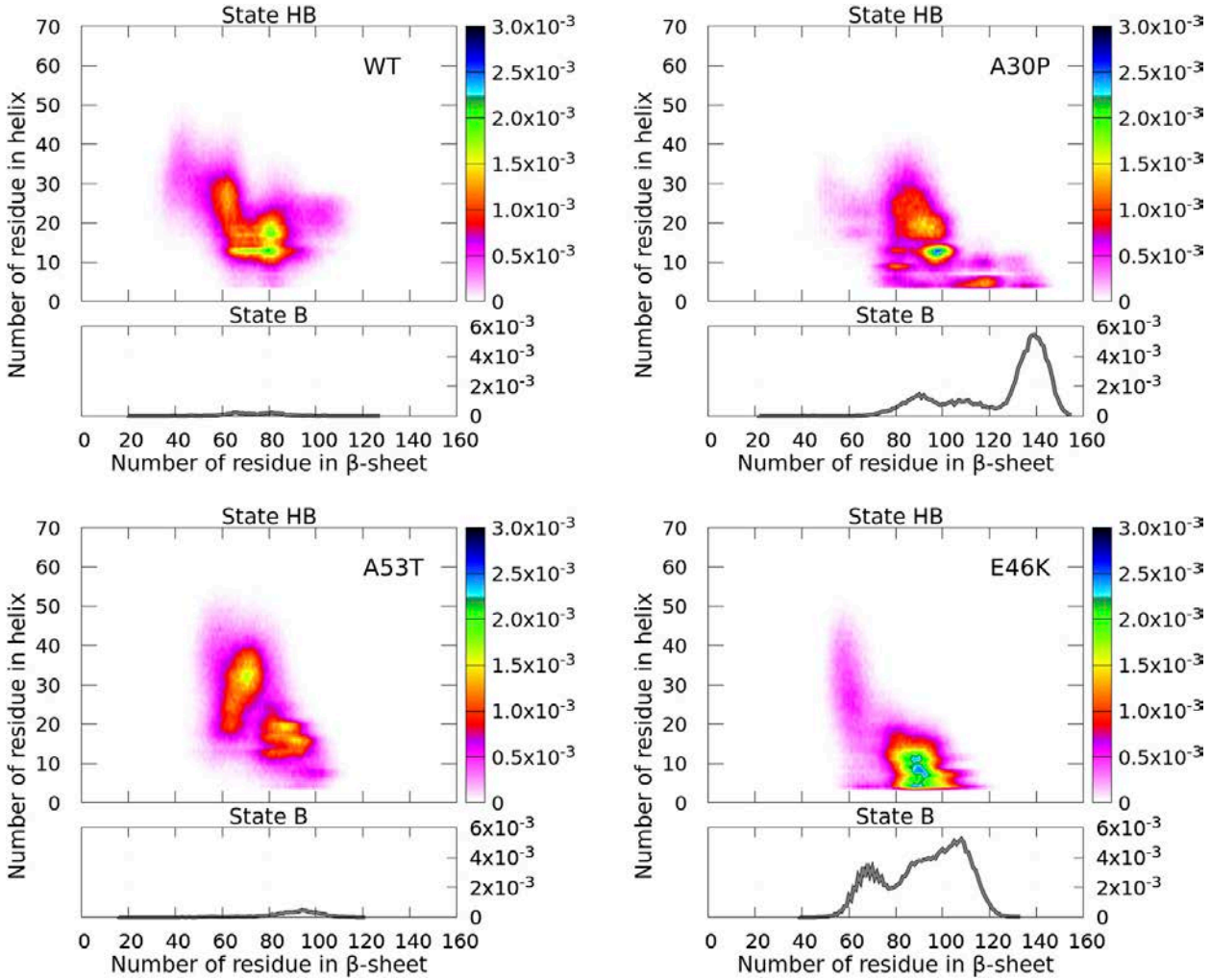
RESULTS – PART 1 DIMERS - COMPLEXITY OF THE FEL (III, n_α,n_β)

Complexity of the FEL (III) (n_α,n_β)

B state of dimers is much less probable

HB states have some correlations with HB state of monomers

Probability

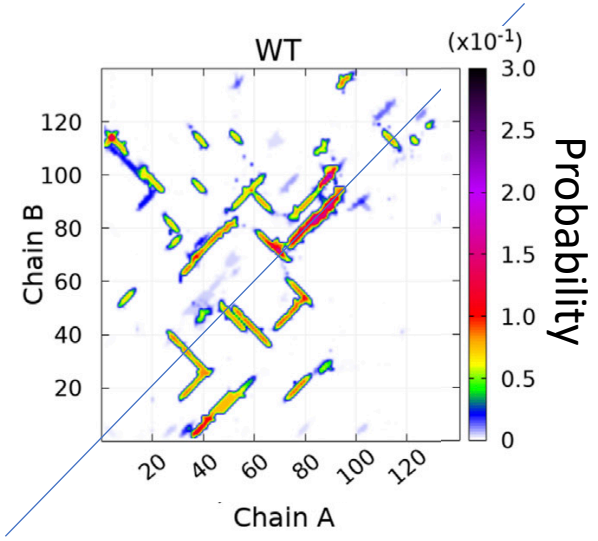


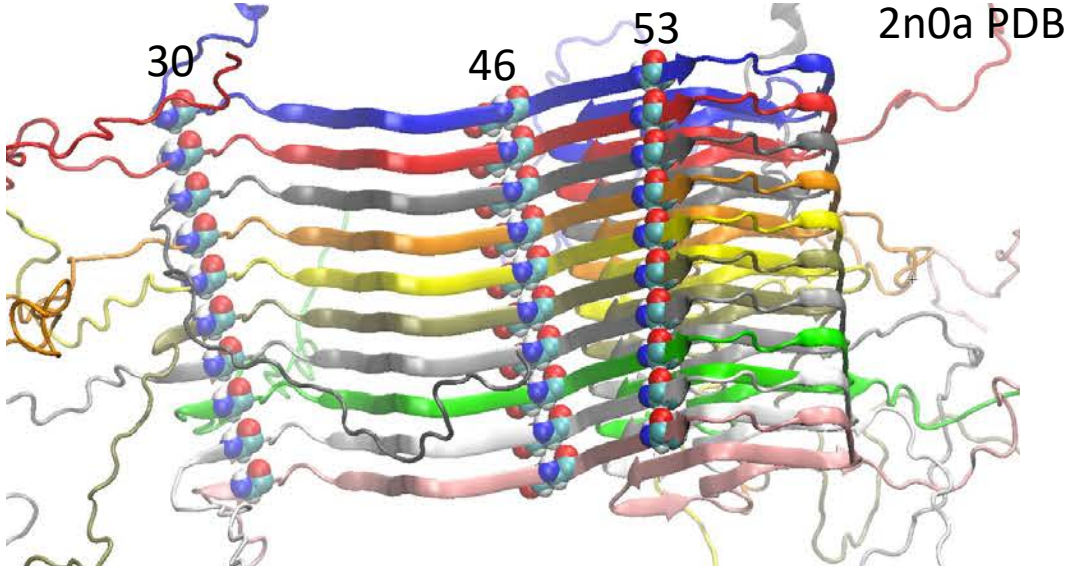
MAIN MESSAGES FROM THE DIMER SIMULATIONS – PART 1

- On the simulation time-scale dimers are disordered structures as IDP
- Gyration radius distribution multimodal (differences WT/mutants)
- B state of dimers correspond to a very small number of configurations
- Complexity of the FELs -> multidimensional approcah to pursue

Among all disordered structures can we detect the nucleation of pre-fibrillar structures ? Possible nucleation centers for the fibrils ?

Maps of intermolecular contacts computed on the whole ensemble of dimers



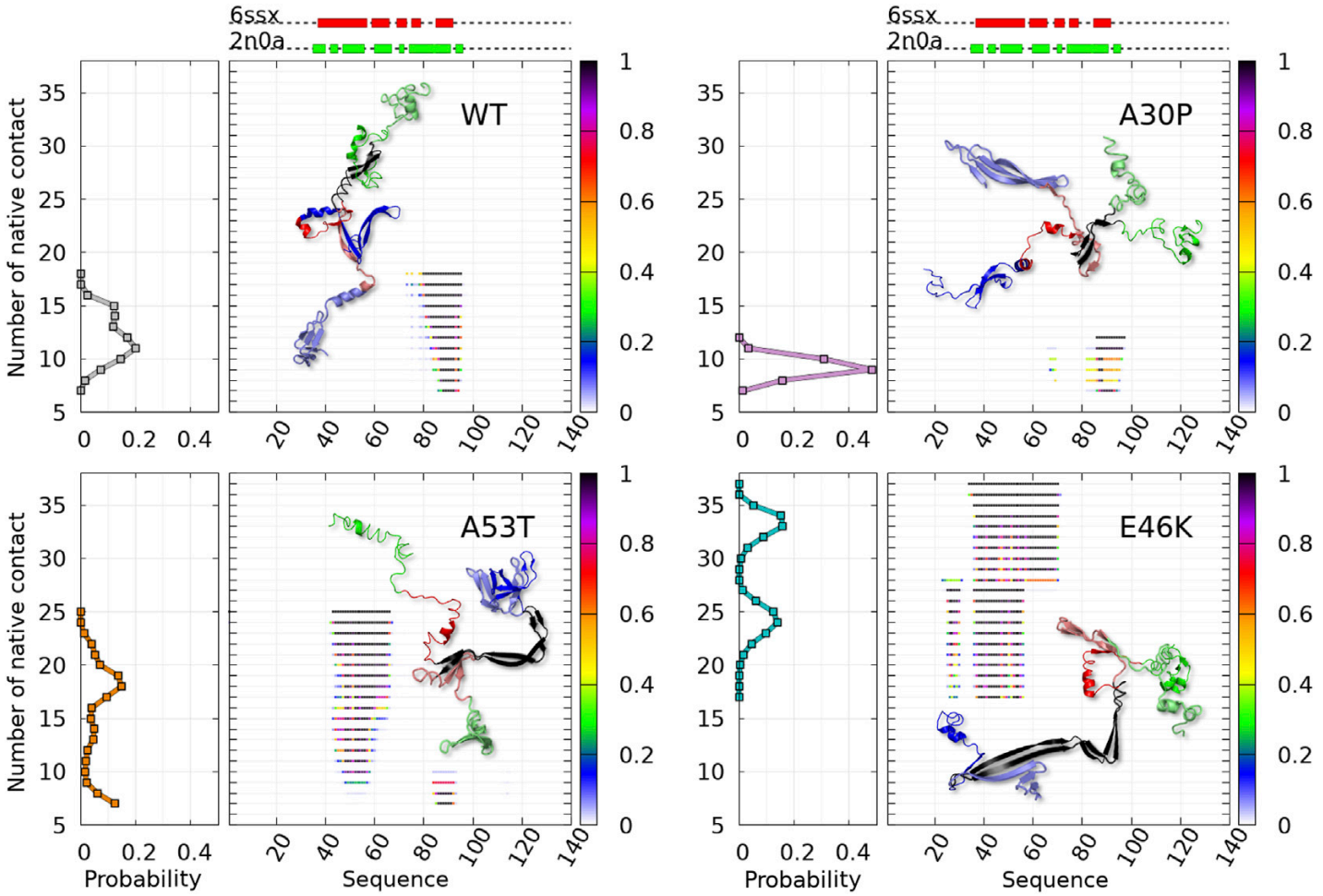


Parallel β-sheets in protofilament of amyloid fibrils

We define **Dfnc structures** = Dimers with fibril native contacts
Structure with at least 5 consecutive native contacts

Fraction of Dfnc in dimers (minority):
WT = 8% < A53T=11% < E46K=14% < A30P=16%

1 single familial mutation changes the % of pre-fibrillar structures
What about the nucleation region within the sequence ?



In addition we identify key amino-acids as for example L38, Y39

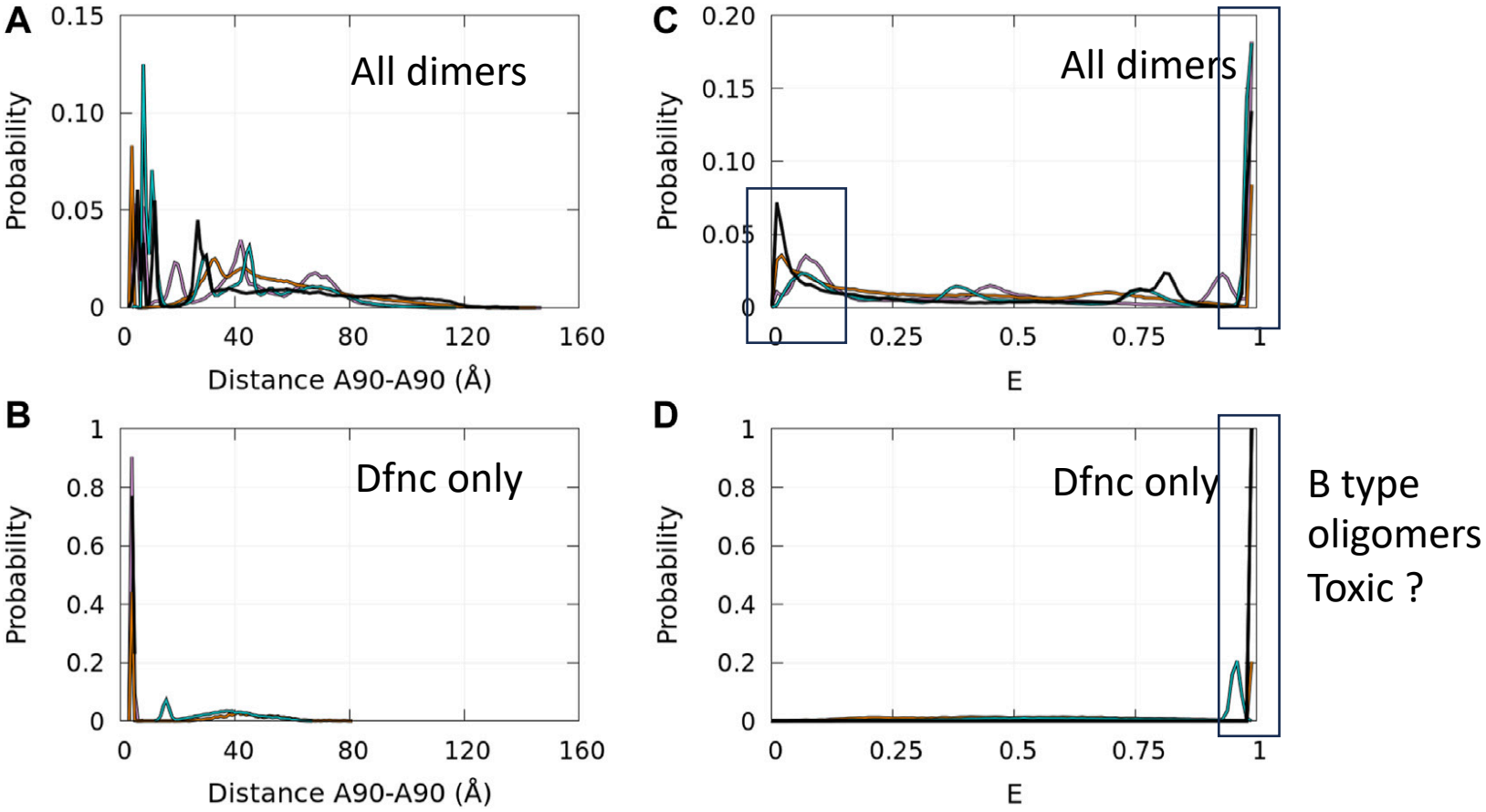
RESULTS – PART 2 DIMERS – COMPARISON WITH FRET DATA

Comparison ***to single-molecule FRET data***

Fluorophores at residue 90

Two types of oligomers : A (low E, not toxic) & **B (high E, toxic**) against cells

Horrocks et al., Anal. Chem. 87, 8818 (2015)



B type oligomers Toxic ?

MAIN MESSAGES FROM THE DIMER SIMULATIONS – PART 2

- **8-16% of prefibrillar structures (Dfnc)**
  WT = 8% < A53T=11% < E46K=14% < A30P=16%
- **Shorter native contact region in A30P & WT but large number of dimers could be interpreted as a slow fibril growth (as observed)**
- **Larger regions of native contacts in E46K and A53T could be interpreted as a faster fibril growth (as observed)**
- **Importance of N-terminal region for dimerisation (E46K, A53T)**
- **Key amino-acids in aggregation (not shown)**
- **Agree with experimental data (AFM, FRET) – predictions for Rg (SAXS)**

Laboratoire Interdisciplinaire Carnot de Bourgogne

Collaborateurs

Adrien GUZZO (ICB, main contributor, PhD thesis 2018-2022)
Steve TYLER (former master student, graph)
Ruoyang GUO (former master student, graph)
Patrice DELARUE (ICB)
Adrien NICOLAI (ICB)
Ana ROJAS (Schrödinger Inc)
Gia MAISURADZE (Cornell University)
Patrick SENET (ICB & Cornell University)

Related publications:

Adrien GUZZO, PhD thesis, Université de Bourgogne, 2022
A. Guzzo et al., Frontiers in Molecular Biosciences, 9, 910104 (2022)
A. Guzzo et al., Frontiers in Molecular Biosciences, 8, 786123 (2021) (included CUTABI)
P. Grassein at al., J. Phys. Chem. B , 124, 4391 (2020)

Laboratoire Interdisciplinaire Carnot de Bourgogne

THANK YOU FOR YOUR ATTENTION

QUESTIONS **?**

RÉGION BOURGOGNE FRANCHE COMTE

l'Europe s'engage en Bourgogne-Franche-Comté

EIPHI-BFC GRADUATE SCHOOL CROSS DISCIPLINARY SCIENCE AND TECHNOLOGY

INVESTIR L'AVENIR

cnrs

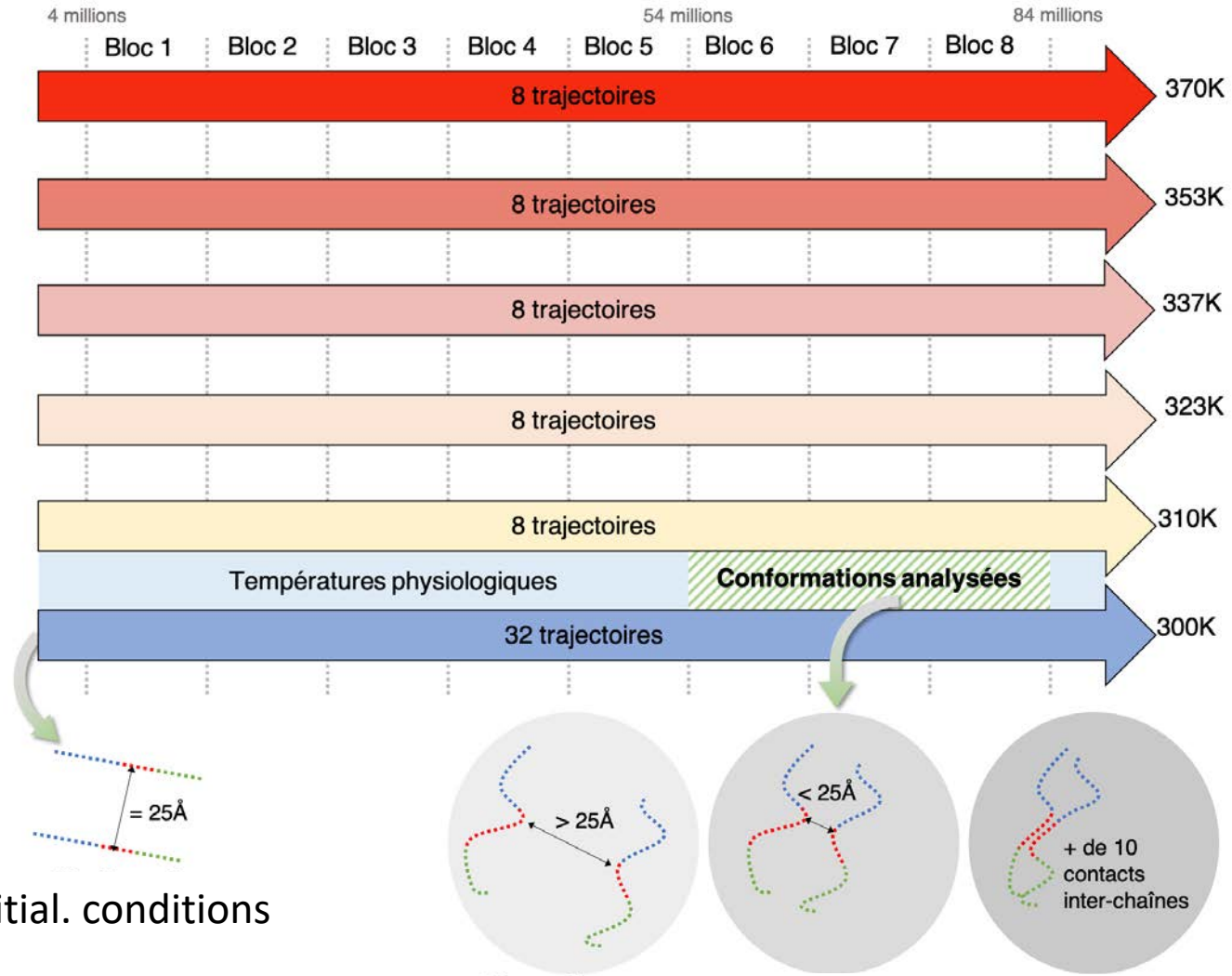
uB UNIVERSITÉ DE BOURGOGNE

utbm université de technologie Belfort-Montbéliard

UBFC UNIVERSITÉ BOURGOGNE FRANCHE-COMTÉ

SUPPORTING INFORMATION

MODEL & METHODOLOGY – REPLICA EXCHANGE MD



Initial. conditions

WT numbers

Monomers (56%)

Transient (13%)

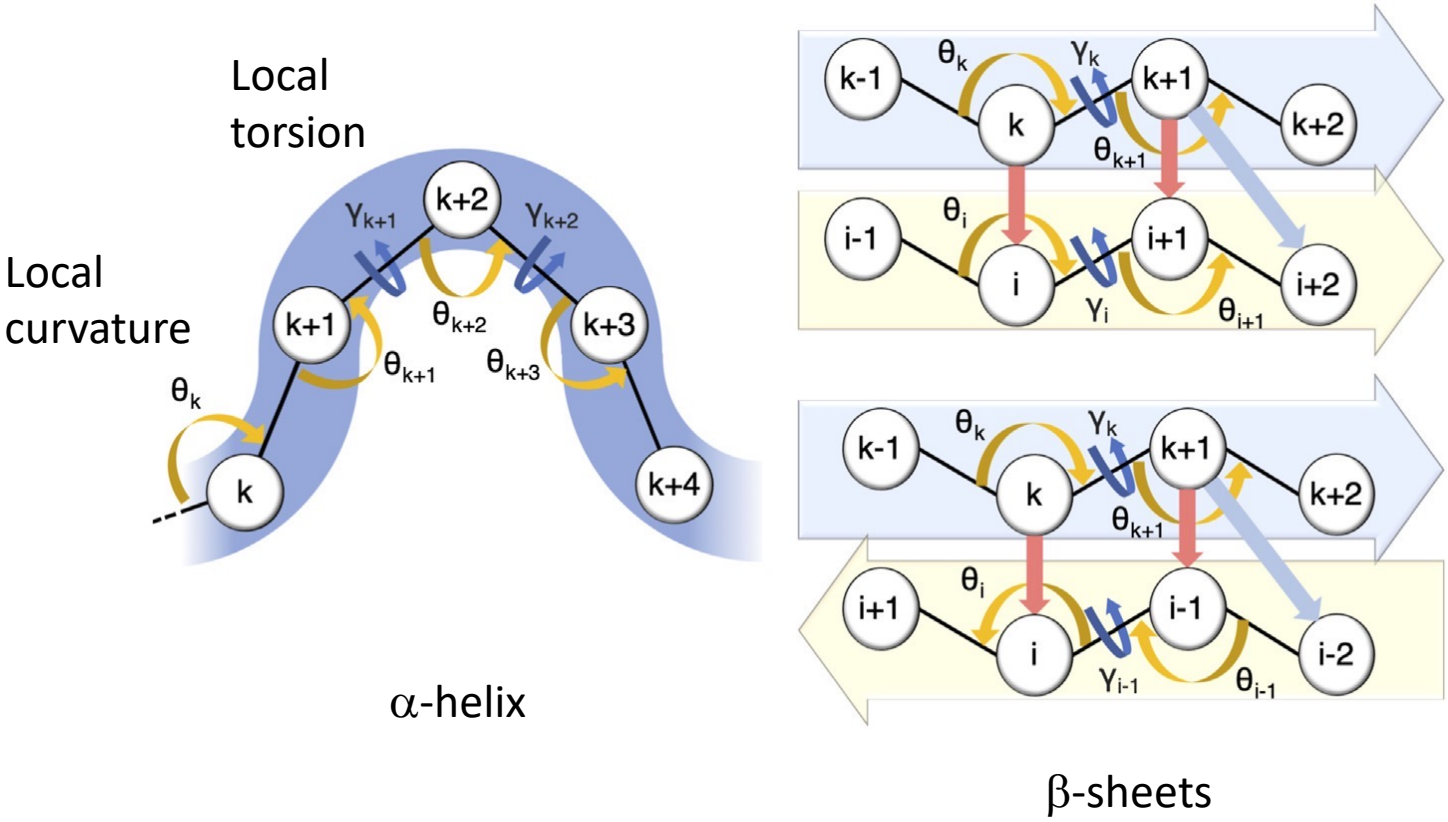
Dimers (31%)

Prefibrillar (2.6%)

Production data = 5.9 milliseconds of effective time scale for each protein at 300-310 K

CUrvature and **T**orsion based of **A**lpha-helix and **B**eta-sheet **I**dentification

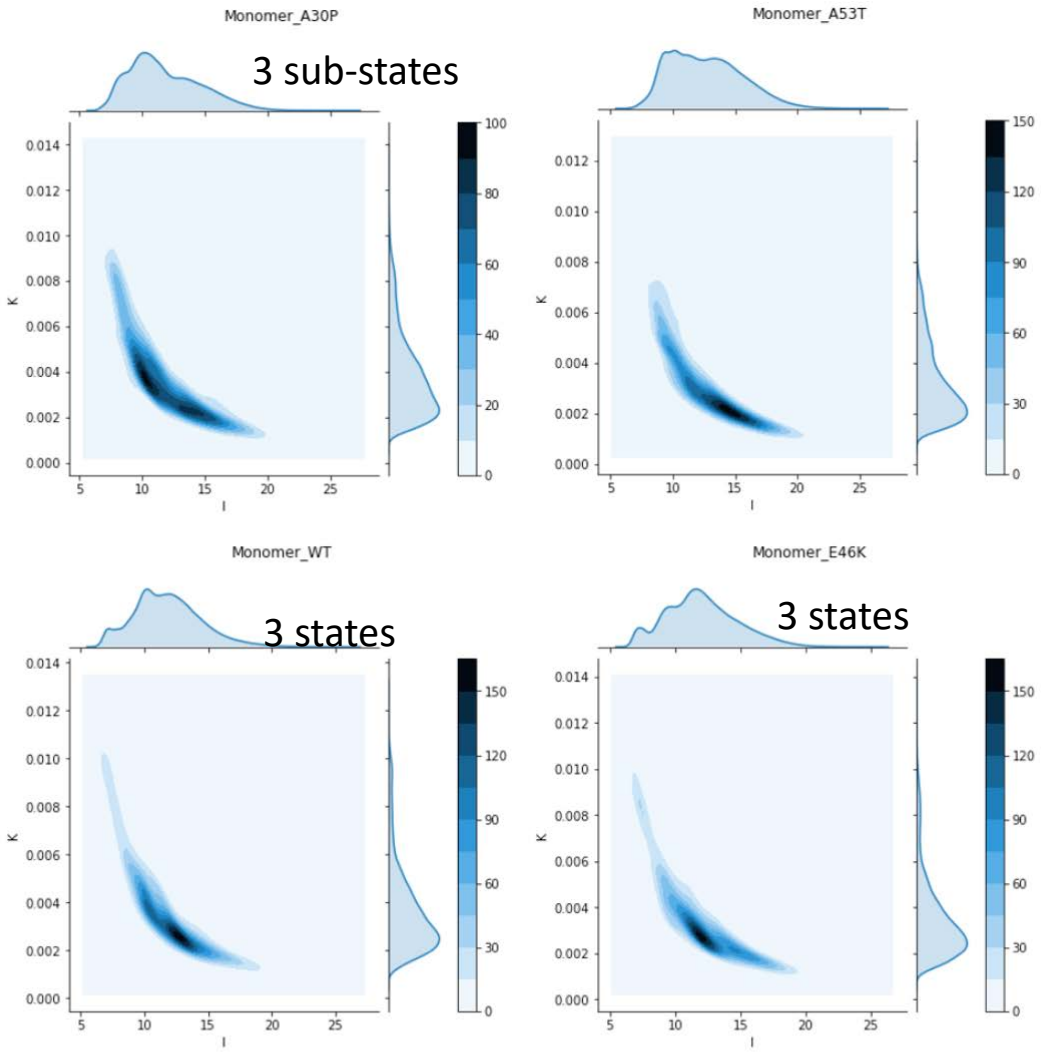
With CUTABI, no need to convert coarse-grained structures in all atom to identify secondary structures (works even better than DSSP in some cases)



For details see A. Guzzo et al., Frontiers in Molecular Biosciences, 8, 786123 (2021)

3. Free-energy landscape (K,l) for α-synuclein monomer

K & l are inversely related !



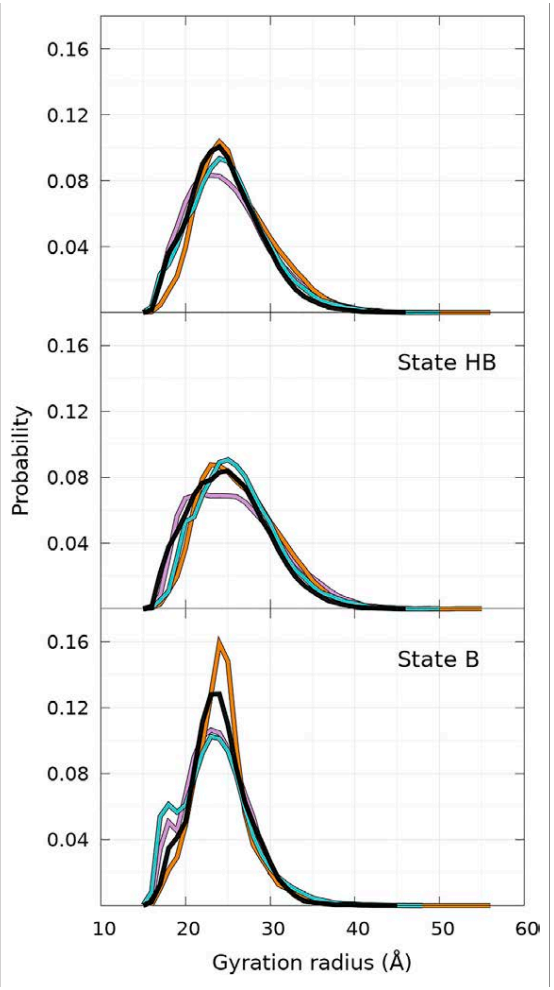
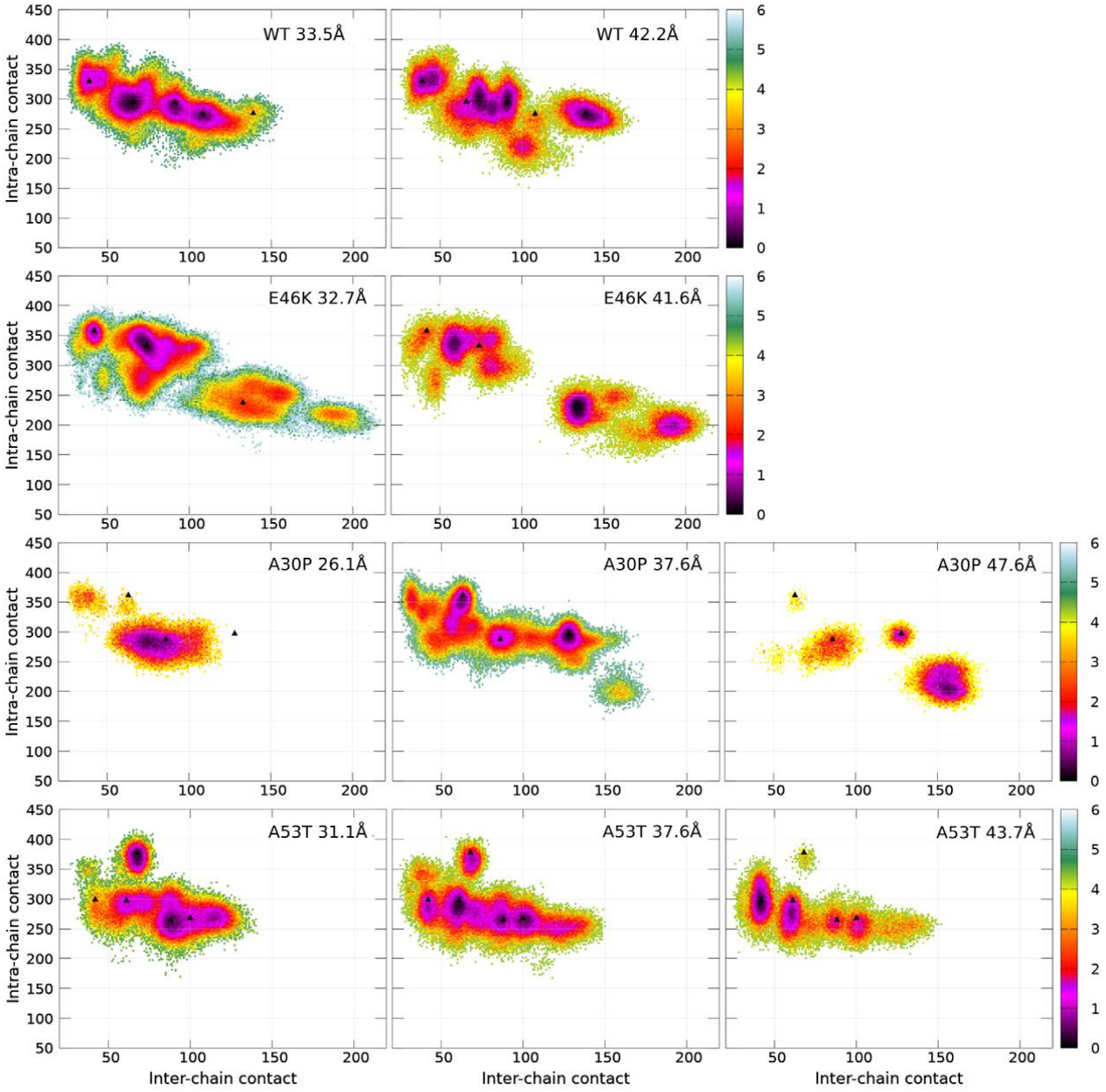
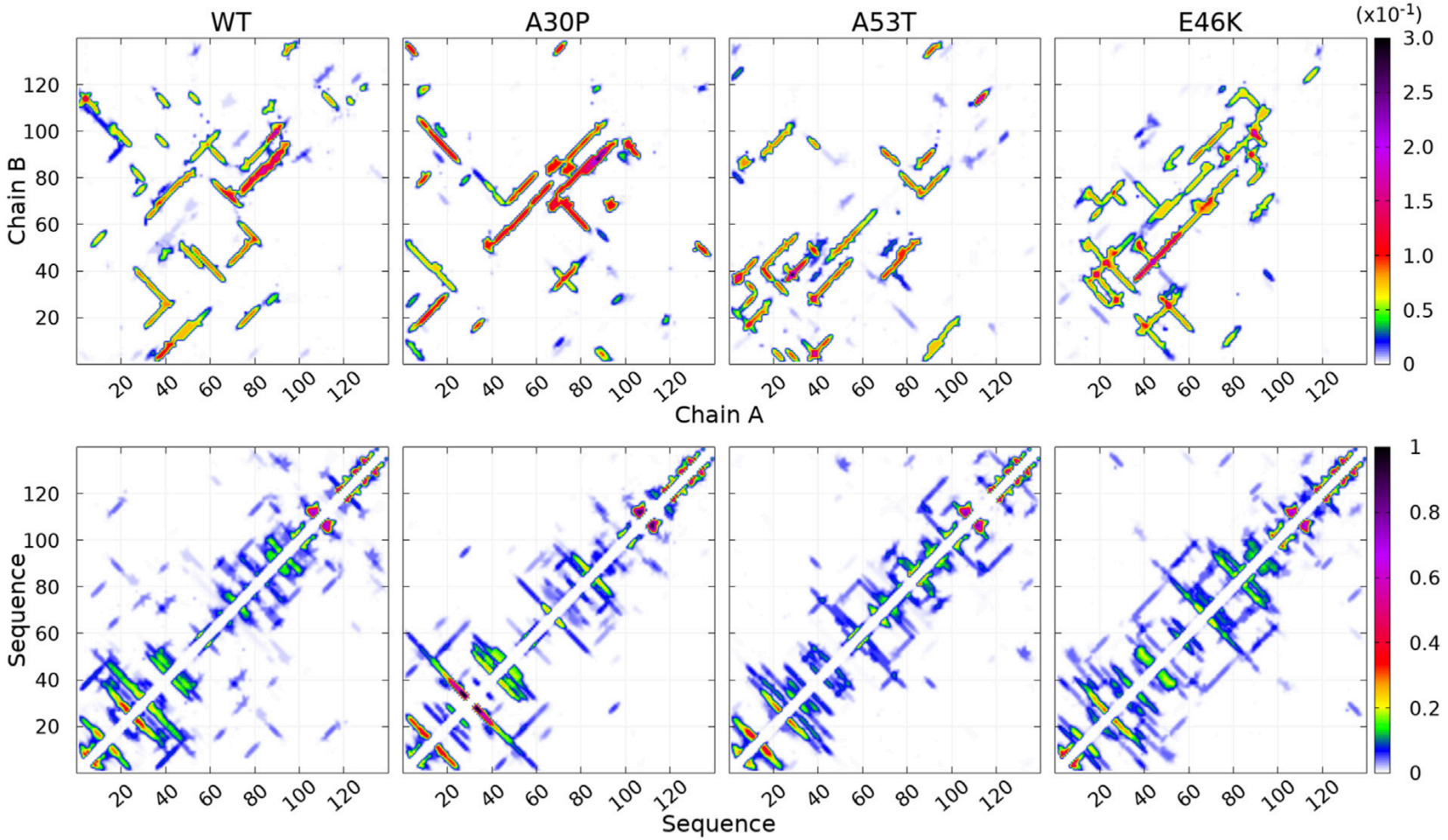


Fig. 20 Two-dimensional PDF of global K and l for alpha-synuclein monomer (WT and A30P, A53T, E46K mutants). The one-dimensional PDF of K and l are shown on the secondary vertical and horizontal axis, respectively.


WT 33.5Å
WT 42.2Å
E46K 32.7Å
E46K 41.6Å
A30P 26.1Å
A30P 37.6Å
A30P 47.6Å
A53T 31.1Å
A53T 37.6Å
A53T 43.7Å
Intra-chain contact
Inter-chain contact
50
100
150
200
250
300
350
400
450
0
1
2
3
4
5
6


WT
A30P
A53T
E46K
(x10-1)
3.0
2.5
2.0
1.5
1.0
0.5
0
Chain B
Chain A
Sequence
Sequence
1
0.8
0.6
0.4
0.2
0
20
40
60
80
100
120

RESULTS - DIMERS

Analyze of the sub-group of Dfnc along the amino-acid sequence

Blue = helix
Green= intra molecular β

Black = inter molecular β
Red = native

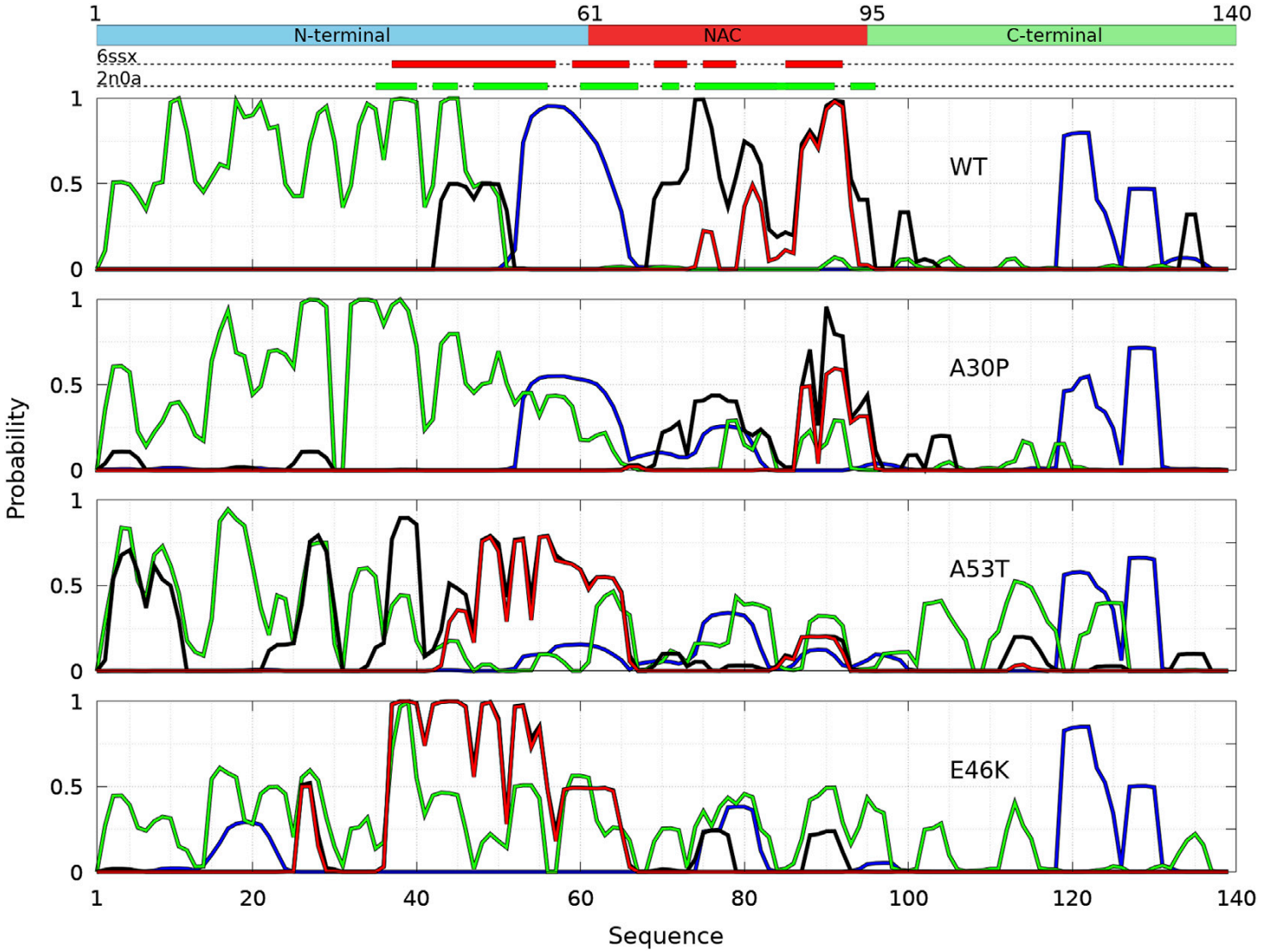


TABLE 4 | Main residues or segments identified by MD as important for the dimerization of α-syn from the maxima of propensity for the mean contact, intermolecular β-sheets, and Nfcs.

Protein	Mean contact	Intermolecular β-sheet (all dimers)	Nfcs in Dfncs
WT	I88	I88	T75, A76, K80, T81,V82, A85, S87-G93
A30P	V70	A90	S87, I88, A90-V95
A53T	L38	L38	T44-N65, S87-T92
E46K	V49	Y39	V26-E28, V37-N65

TABLE 3 | Clustering of the gyration radius probability density using the GMM algorithm. The values in brackets are the corresponding % of the ensemble of the conformations.

Protein	Cluster 1	Cluster 2	Cluster 3	Average value
WT	33.5 Å (60%)	42.2 Å (40%)	-	37.0 Å (100%)
A30P	37.6 Å (73%)	47.6 Å (16%)	26.1 Å (11%)	37.9 Å (100%)
A53T	31.1 Å (41%)	37.6 Å (38%)	43.8 Å (21%)	36.2 Å (100%)
E46K	32.7 Å (72%)	41.7 Å (28%)	-	35.2 Å (100%)

TABLE 4 | Main residues or segments identified by MD as important for the dimerization of α-syn from the maxima of propensity for the mean contact, intermolecular β-sheets, and Nfcs.

Protein	Mean contact	Intermolecular β-sheet (all dimers)	Nfcs in Dfncs
WT	I88	I88	T75, A76, K80, T81,V82, A85, S87-G93
A30P	V70	A90	S87, I88, A90-V95
A53T	L38	L38	T44-N65, S87-T92
E46K	V49	Y39	V26-E28, V37-N65

RESULTS - DIMERS

Comparison ***to AFM rupture-force experimental data***

MD :

WT & A30P form native contacts in shorter localized segments
A53T & E46 K more native contacts in two regions

AFM:

large number of multiple rupture force events observed for A53T and E46K compared to WT and A30P

MD : Stretchable part of the polymers:
WT & A30P = 35,7 nm
A53T = 38 nm & 58 nm
E46K = 58 nm

AFM: *Stretchable part of the polymers:*
34 nm & 44 nm