Electronic Supplementary Material (ESI) for RSC Advances.

Structural transition of ETS1 from auto-inhibited to functional state upon association with $p16^{INK4a}$ native and mutated promoter region

Kannan Muthu, Manivel Panneerselvam, Nishith Saurav Topno and Krishna Ramadas*

Centre for Bioinformatics, Pondicherry University, Puducherry, India-605014

*Corresponding Author

Dr. R. Krishna,

Assistant Professor,

Centre for Bioinformatics,

Pondicherry University

Puducherry-605014

Supplementary methods

T-Pad analysis

Understanding the protein plasticity plays a key role in molecular recognition and numerous cellular processes like metabolism, protein aggregation, gene expression and molecular signaling [1-4]. T-pad analysis is performed on residue-by-residue to identify flexible and rigid sites in proteins. This method also pinpoints backbone transitions concerning two conformations of Ramachandran plot which could subsequently explain the structural adaptation through its hinge point in a molecular mechanism. Specifically, in our case we used T-pad analysis to understand the structural transition of ETS1 protein residues upon binding with native and mutated $p16^{INK4a}$ promoter region. Accordingly, each 50 ns trajectory of native and mutated complexes were prepared by removing the water molecules, ions and un-equilibrated trajectories by using the *trjconv* tool and used for T-pad analysis. Previously, the protein plasticity was described based on the circular spread Ramachandran angles Φ or ψ (CSΦ or CSψ) but the protein backbone conformation was not accounted when using these angular dispersion indices. Hence, Protein Angular Dispersion (PADω), a new quantity was introduced to overcome these complications and by the assessment of both CSω and PADω, protein backbone plasticity can be enumerated clearly. PADω is advantageous over CSω due to the two features: i) the function ω (ω = Φ + ψ) is dependent on both the Ramachandran angles; ii) and it is formulated in the range between 0° and 180° [5].

Supplementary figure legends:

S1: The interaction of ETS1 (green) with A) Native_ $p16^{INK4a}$, B) M1_ $p16^{INK4a}$, C) M2_ $p16^{INK4a}$ and D) M3_ $p16^{INK4a}$. Color codes; ETS1 protein: green, Native_ $p16^{INK4a}$: Cyan, M1_ $p16^{INK4a}$: Cornflower blue, M2_ $p16^{INK4a}$: Purple and M3_ $p16^{INK4a}$: Gray. Here, atoms are shown in heteroatom type.

S2: The major and minor groove width of native and mutated $p16^{INK4a}$ promoters with ETS1 protein.

S3: The structural transition of ETS1 in complex with M2_$p16^{INK4a}$ promoter. A) Describes the four types of mechanism governing folding of HI1 helix and changes in DNA bending. The docked (protein: plum and DNA: tan) and free energy represented structure (protein: green and DNA: purple) superimposed and explains the changes in the protein (orange arrow) and DNA (black arrow). Mechanism I) the hydrophobic interaction formed between P334 and (green spheres) DNA, loss of triangle basic patch interaction (blue sphere) and triplet residues contribution (orange spheres) are shown. Mechanism II) the series of hydrophobic interaction (residues yellow spheres) and hydrogen bond (green spheres) formed between H2 and H1 helix. Mechanism III) defines the hydrophobic interaction between H1 and HI2 helix. Mechanism IV) folding of HI1 helix and consequent basic to acidic patch and hydrophobic interaction were shown. (B) *Roll (ρ)* - *Twist (Ω)* plot, (C) *Slide* (D_y) - *Twist (Ω)* plot and (D) major (straight line) - minor groove (dotted line) width parameters explain the mode of DNA distortion. The reference (B and C) / docked (D) and free energy representative (B, C, D) DNA is shown in black and red, respectively. The shaded band (in C) shows the conformational channel responsible for naked DNA mode I distortion.

S4: The structural transition of ETS1 in complex with M3_$p16^{INK4a}$ promoter. A) Describes the four types of mechanism governing folding of HI1 helix and changes in DNA bending. The docked (protein: plum and DNA: tan) and free energy represented structure (protein: green and DNA: gray) superimposed and explains the changes in the protein (orange arrow) and DNA (black arrow). Mechanism I) the hydrophobic interaction formed between P334 and (green spheres) DNA, loss of triangle basic patch interaction (blue spheres) and triplet residues

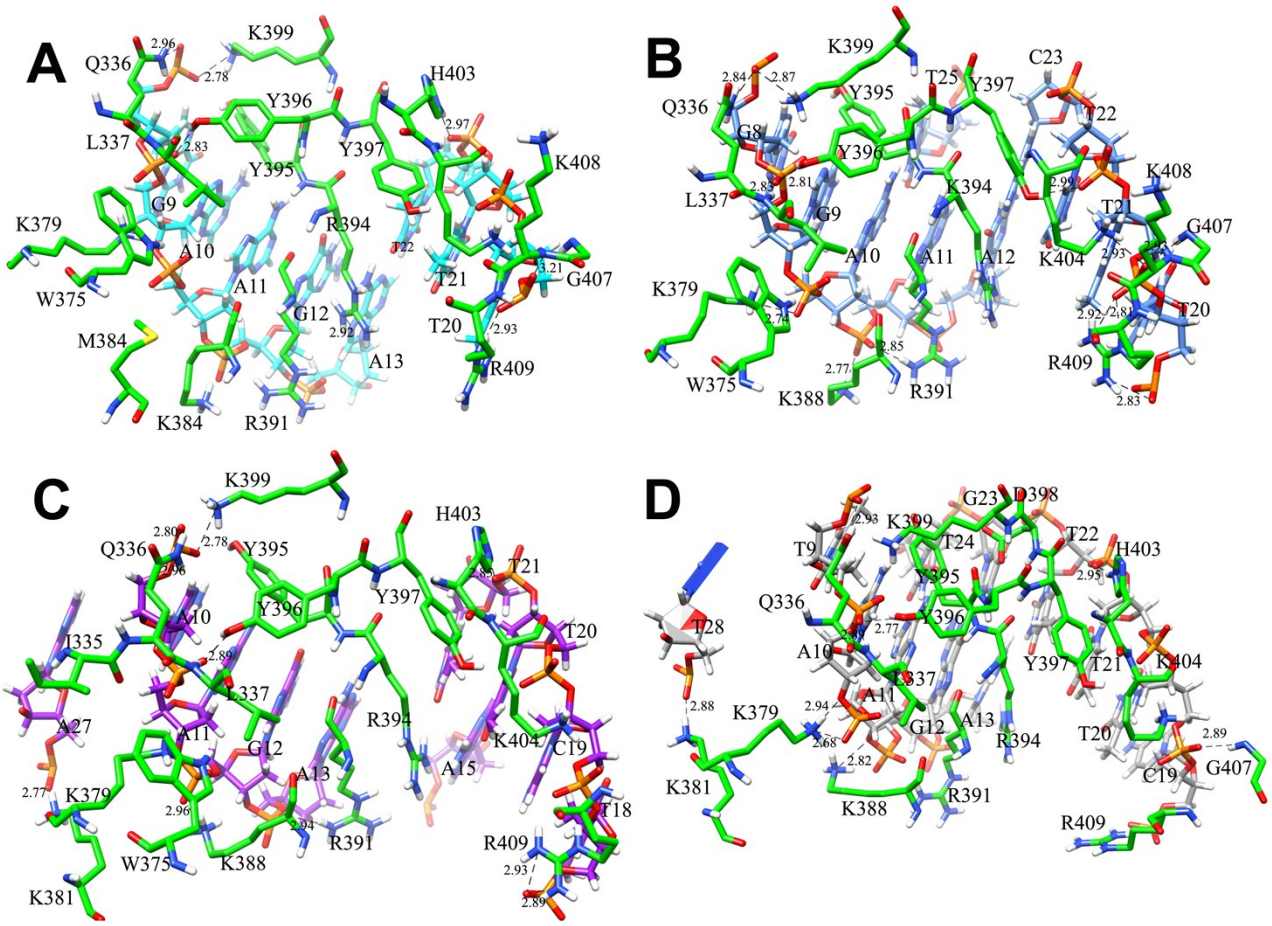
contribution (orange spheres) are shown. Mechanism II) the series of hydrophobic interaction (yellow spheres) and hydrogen bond (green spheres) formed between H2 and H1 helix. Mechanism III) defines the hydrophobic interaction formed between H1 and HI2 helix. Mechanism IV) folding of HI1 helix and consequent basic to acidic patch and hydrophobic interaction formed. (B) *Roll (ρ)* - *Twist (Ω)* plot, (C) *Slide* (D_y) - *Twist (Ω)* plot and (D) major (straight line) - minor groove (dotted line) width parameters explain the mode of DNA distortion. The reference (B and C) / docked (D) and free energy representative (B, C, D) DNA is shown in black and red, respectively. The shaded band (in C) shows the conformational channel responsible for naked DNA mode I distortion.

Supplementary Table legends

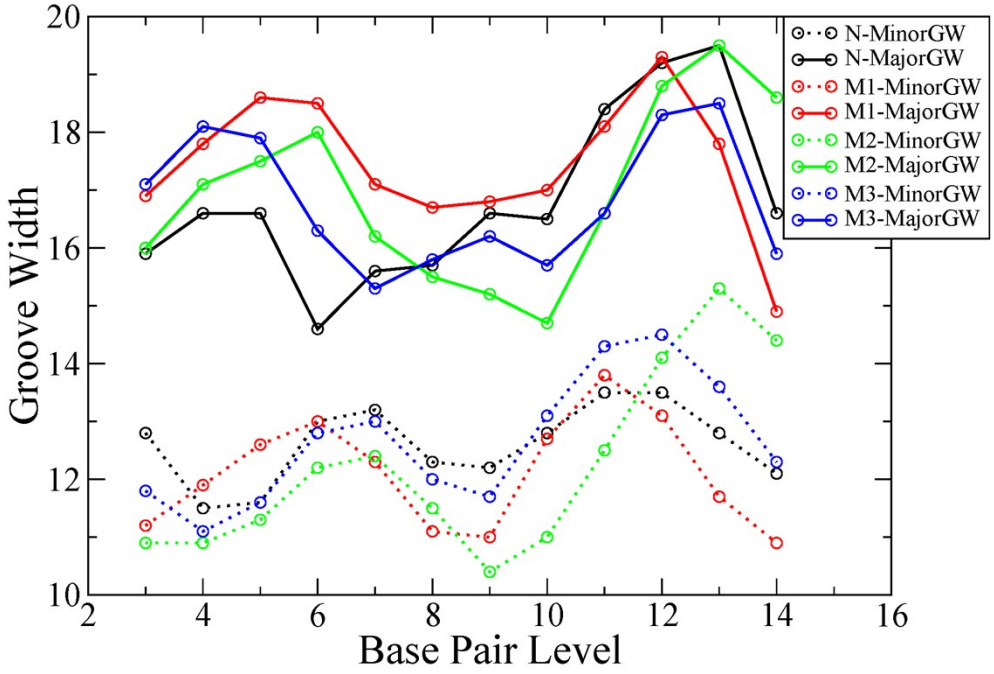
ST1: The parameters, *Roll*, *Slide*, *Twist*, H-twist, minor and major groove width calculated using 3DNA tool for the selected single nucleotide base pair steps in each free energy representative DNA structure are listed. The high *Roll* values of single nucleotide base pair steps highlighted in bold and discussed.

ST2: The distance of helical orientation and number of hydrogen bond maintained between the helices was calculated throughout the simulation and listed here. The ETS domain helical distance was highlighted in bold.

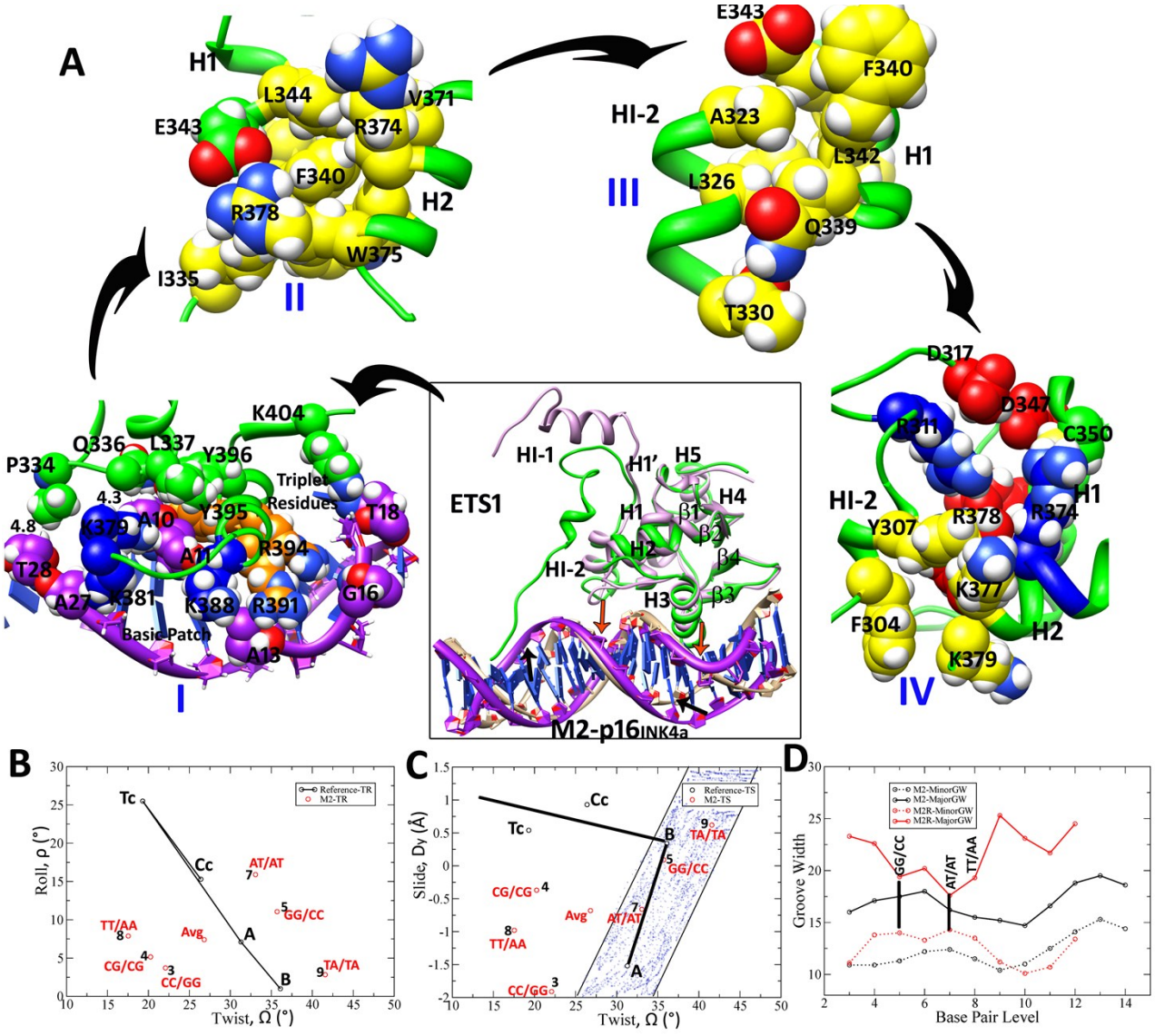
S1:


A
Q336
K399
H403
Y396
L337
Y397
Y395
K408
G9
K379
R394
A10
T21
G407
A11
W375
G12
T20
M384
A13
R409
K384
R391
B
K399
C23
Y395
T25
Y397
Q336
T22
G8
Y396
K408
L337
K394
T21
G9
G407
A10
A11
A12
K404
K379
T20
W375
K388
R391
R409
C
K399
H403
Q336
Y395
T21
Y397
Y396
A10
T20
T335
A337
R394
A27
A11
G12
K404
C19
A13
A15
K379
T18
R391
R409
W375
K388
K381
D
G23
D398
K399
T22
T9
T24
H403
Y395
Q336
Y396
T28
A10
Y397
T21
K404
L337
A11
K379
G12
A13
R394
T20
G407
C19
K381
K388
R391
R409

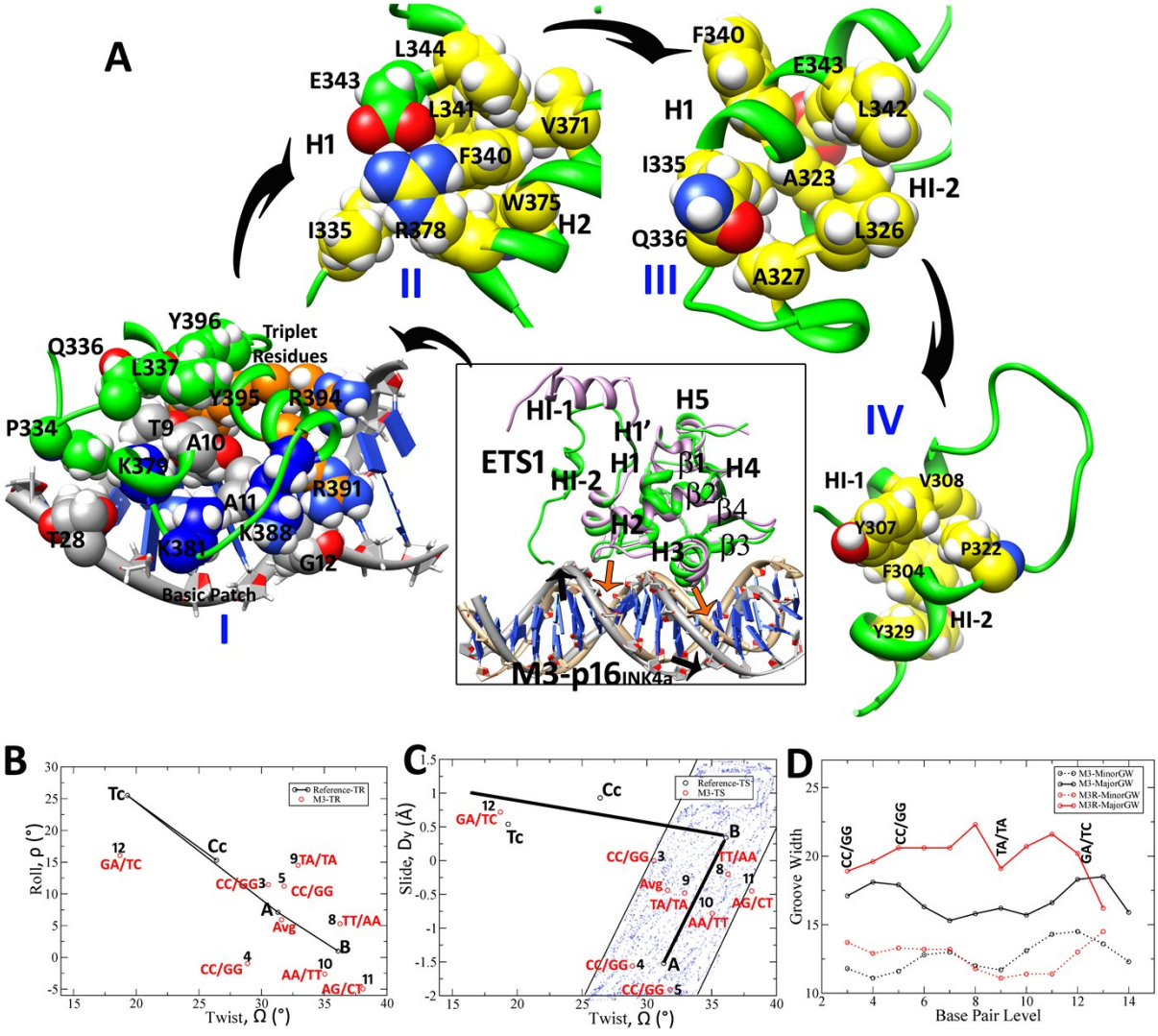
S2:


Groove Width
Base Pair Level
N-MinorGW
N-MajorGW
M1-MinorGW
M1-MajorGW
M2-MinorGW
M2-MajorGW
M3-MinorGW
M3-MajorGW
20
18
16
14
12
10
2
4
6
8
10
12
14
16

S3:


A
H1
L344
V371
E343
R374
F340
H2
R378
W375
I335
II
E343
F340
HI-2
A323
L342
H1
III
L326
Q339
T330
D317
D347
P311
C350
HI-2
H1
Y307
R378
R374
K377
F304
H2
K379
IV
K404
Q336
L337
Y396
P334
Triplet
Residues
4.3
Y395
T18
4.8
K379
A10
A11
R394
T28
A27
K381
K388
R391
G16
Basic Patch
A13
I
HI-1
H1'
H5
ETS1
H4
H1
β1
β2
H2
β4
HI-2
H3
β3
M2-p16INK4a
B
Reference-TR
M2-TR
Tc
Cc
AT/AT
GG/CC
TT/AA
Avg
A
CG/CG
CC/GG
TA/TA
B
Roll, ρ (°)
Twist, Ω (°)
C
Reference-TS
M2-TS
Cc
Tc
B
TA/TA
GG/CC
CG/CG
Avg
AT/AT
TT/AA
A
CC/GG
Slide, Dy (Å)
Twist, Ω (°)
D
M2-MinorGW
M2-MajorGW
M2R-MinorGW
M2R-MajorGW
GG/CC
AT/AT
TT/AA
Groove Width
Base Pair Level

S4:


A
L344
E343
L341
V371
H1
F340
W375
H2
I335
R378
II
F340
E343
H1
L342
I335
A323
HI-2
Q336
L326
A327
III
IV
Y396
Triplet Residues
Q336
L337
Y395
R394
P334
T9
A10
K379
R391
A11
K388
T28
K381
G12
Basic Patch
I
HI-1
H5
H1'
ETS1
H1
β1
H4
HI-2
β2
β4
H2
H3
β3
M3-p16INK4a
HI-1
V308
Y307
P322
F304
Y329
HI-2
B
Reference-TR
M3-TR
Tc
Cc
12 GA/TC
9 TA/TA
CC/GG 3
5 CC/GG
A
Avg
8 TT/AA
B
CC/GG 4
10 AA/TT
AG/CT 11
Roll, ρ (°)
Twist, Ω (°)
C
Reference-TS
M3-TS
Cc
12 GA/TC
Tc
B
CC/GG 3
TT/AA
8
Avg
9
11
TA/TA
10
AG/CT
AA/TT
CC/GG 4
A
CC/GG 5
Slide, Dy (Å)
Twist, Ω (°)
D
M3-MinorGW
M3-MajorGW
M3R-MinorGW
M3R-MajorGW
CC/GG
CC/GG
TA/TA
GA/TC
Groove Width
Base Pair Level

ST1:

Representative structure	Base pair step		Roll ρ	Slide D_y	Twist Ω	H-twist	Minor Width	Major Width	DNA form
N_P16^{INK4a} – ETS1	**4**	**CG/CG**	**12.92**	**-0.35**	**31.78**	**34.33**	**13.6**	**20.7**	**B**
	5	GG/CC	6.24	-2.07	26.96	27.76	11.7	22.5	-
	6	GA/TC	1.19	-1.45	33.49	33.52	10.9	22.6	-
	10	AA/TT	-0.48	-0.46	20.64	20.87	12.2	22.2	-
	11	**AG/CT**	**16.19**	**0.26**	**35.82**	**39.52**	**12.0**	**18.8**	**B**
	12	GA/TC	-2.15	-0.01	40.24	40.30	9.7	16.8	B
	Avg		*2.17*	*-0.30*	*30.58*	*30.57*			
	S.D		*6.43*	*1.12*	*10.89*	*13.57*			
M1_P16^{INK4a} – ETS1	**3**	**CC/GG**	**16.16**	**-1.36**	**34.34**	**38.70**	**9.4**	**21.3**	**B**
	4	CC/GG	-0.94	-0.81	36.19	36.32	9.7	22.8	B
	5	CC/GG	7.42	-0.68	37.44	38.17	12.7	22.9	B
	6	**CA/TG**	**8.80**	**0.66**	**11.92**	**15.00**	**15.6**	**18.6**	**B**
	10	AA/TT	-4.16	-1.22	35.96	36.25	10.3	23.3	B
	11	AG/CT	0.19	-1.18	31.78	32.19	12.5	24.5	-
	12	**GA/TC**	**18.77**	**-1.44**	**31.43**	**33.11**	**14.5**	**22.8**	**-**
	Avg		*5.71*	*-0.43*	*29.95*	*32.00*			
	S.D		*8.46*	*0.97*	*8.20*	*7.25*			
M2_P16^{INK4a} – ETS1	3	CC/GG	3.71	-1.91	22.08	22.87	11.1	23.3	-
	4	CG/CG	5.15	-0.37	20.26	21.61	13.8	22.6	B
	5	**GG/CC**	**11.06**	**0.09**	**35.71**	**37.82**	**14.0**	**19.4**	**B**
	7	**AT/AT**	**15.90**	**-0.66**	**33.07**	**37.31**	**14.3**	**17.6**	**B**
	8	**TT/AA**	**7.88**	**-0.98**	**17.54**	**23.48**	**13.5**	**19.3**	**-**
	9	TA/TA	2.83	0.62	41.58	41.83	11.2	25.3	B
	Avg		*7.40*	*-0.68*	*26.81*	*27.37*			
	S.D		*8.74*	*0.76*	*16.40*	*20.63*			
M3_P16^{INK4a} – ETS1	**3**	**CC/GG**	**11.45**	**0.00**	**30.54**	**32.59**	**13.7**	**18.9**	**B**
	4	CC/GG	-0.98	-1.56	28.88	29.00	12.9	19.6	B
	5	**CC/GG**	**11.20**	**-1.91**	**31.79**	**33.93**	**13.3**	**20.6**	**B**
	8	TT/AA	5.30	-0.20	36.25	36.79	11.8	22.3	B
	9	**TA/TA**	**14.50**	**-0.48**	**32.91**	**36.38**	**11.1**	**19.1**	**B**
	10	AA/TT	-2.63	-0.78	35.04	35.30	11.4	20.7	B
	11	AG/CT	-4.96	-0.45	38.09	38.50	11.4	21.6	B
	12	**GA/TC**	**16.06**	**0.72**	**18.70**	**24.50**	**13.0**	**20.2**	**-**
	Avg		*5.93*	*-0.44*	*31.60*	*33.17*			
	S.D		*6.58*	*0.80*	*5.28*	*4.47*			

ST 2:

Complex-simulation data	N_P16^{INK4a}-ETS1	M1_P16^{INK4a}-ETS1	M2_P16^{INK4a}-ETS1	M3_P16^{INK4a}-ETS1
Distance of helices and sheet (Å)				
H1-βsheet	11.5	11.5	11.7	11.5
H1-H2	11.5	11.25	11.5	11.5
H1-H3	13.5	13.5	13.5	13.5
H2- βsheet	17.5	17.5	17.8	17.5
H2-H3	**12.0**	**12.3**	**12.5**	**12.3**
H3- βsheet	**12.5**	**12.3**	**12.3**	**11.5**
HI2-H4	**13.0**	**12.0**	**11.3**	**12.3**
Hydrogen bond between the helices and sheet (number)				
H1-βsheet	1	1	1	1
H1-H2	3	3	3- 1 maintains & 2 rarely	3
H1-H3	1	1	1 rarely	1
H2-H3	1	0	1 rarely	1 rarely
H3- βsheet	2	1 rarely	2-1 maintain & 1 rarely	3
HI2-H4	1	1	1	1
Number of hydrogen bonds in interface area				
Numbers	10-15	7-12	5-10 for 30,00 & 7-13 remaining	10-17

Supplementary references

1. R. Huber, *Nature*, 1979, 280, 538-539.
2. J. H. Lin, *Current topics in medicinal chemistry*, 2011, 11, 171-178.
3. M. Nocker and P. Cozzini, *Current topics in medicinal chemistry*, 2011, 11, 133-147.
4. K. Teilum, J. G. Olsen and B. B. Kragelund, *Biochim Biophys Acta*, 2011, 1814, 969-976.
5. R. Caliandro, G. Rossetti and P. Carloni, *Journal of Chemical Theory and Computation*, 2012, 8, 4775-4785.