Supplemental Information

Structural basis of diversity of Zinc-Finger-Associated Domains (ZAD) that mainly form homodimers in Drosophila melanogaster

Artem Bonchuk, Konstantin Boyko, Anna Fedotova, Alena Nikolaeva, Sofia Lushchekina, Anastasia Khrustaleva, Vladimir Popov, Pavel Georgiev*

SUPPLEMENTARY METHODS

Molecular dynamics simulations

The lengths of productive MD trajectories were chosen based on RMSD evolution. While 300 ns of MD simulation was enough to achieve a stable RMSD level for the CG2712 homodimer and CG2712–Zw5 heterodimer, for the Zw5 homodimer, RMSD was increasing during first 300 ns; thus, the trajectory was prolonged up to 500 ns (Supplementary Figure S6).

Using 100 ns SMD runs allow the dimer dissociation process to occur closer to equilibrium. During the first 20 ns of pulling, the dimer structure is ruptured, but contact between monomers persists for up to 50-60 ns of simulation, after which the monomers are moving far enough to lose interactions first, followed by electrostatic influence when the distance becomes bigger than the cutoff radius.

SUPPLEMENTARY TABLES

Supplementary Table S1. Primers used for cloning. Restriction enzyme sites are shown in small letters, the corresponding enzymes are noted in brackets. Nucleotide substitutions in mutagenic primers are also shown in small letters.

CG10321 ZAD d: GGgaattcATGGATCGCAAGAAGGTCCTG (*Eco*RI)
CG10321 ZAD r: GGctcgagTACCAGTGCCACGGTAAGACC (*Xho*I)
CG17328 ZAD d: GGgaattcATGGACTATAATATCCACAAAATATGC (*Eco*RI)
CG17328 ZAD r: AActcgagCTGGCGCCAGGATTCCAGAAT (*Xho*I)
CG6654 ZAD d: GGgaattcATGGATGTGGACAAGATCTGTC (*Eco*RI)
CG6654 ZAD r: GGctcgagGCTATCGGGCTCCTCCG (*Xho*I)
CG6808 ZAD d: GGgaattcATGGAAATCAGCATTAAATGGAGC (*Eco*RI)
CG6808 ZAD r: GGctcgagTGCCGATGCGGAGGACAGA (*Xho*I)
wek ZAD d: GGgaattcATGGGAGTTCCCACAAGCGATT (*Eco*RI)
wek ZAD r: GGctcgagGCTATCCTGCGGCTCCGAGT (*Xho*I)
Pita ZAD dir: GGgaattcAACATGGCCGCGAAACTGGA (*Eco*RI)
Pita ZAD rev TTctcgagGGGCGCACGGGGCTTCTCC (*Xho*I)
Grau ZAD dir: GGgaattcAACATGGATATCTGCCGCCTC (*Eco*RI)
Zw5 ZAD dir: GGgaattcAACATGATGAATAGCAAGATCGCC (*Eco*RI)
Zw5d AAActcgagGGAAAGCGGATAAGGACA (*Xho*I)
9793ZADd TTgaattcATGGAAGGAAAAGTGTGTCG (*Eco*RI)
9793ZADr TCctcgagGGGAATGGAGCAACTCCC (*Xho*I)
9797ZADd AAgaattcAAGAGCACCTGCCGCG (*Eco*RI)
9797ZADr CTctcgagATCCCAAGAGTAGCTCAC (*Xho*I)
15073ZADd ATgaattcATGATTTGCCGTTTGTGC (*Eco*RI)
15073ZADr TGctcgagTTTCGGTGAGCAAACGATG (*Xho*I)
2712ZADd TTcaattgATGACGACCATGTCGC (*Mun*I)
2712ZADr TTctcgagTCTGCCGCAGAATTCTGTC (*Xho*I)
17328ZADr AActcgagGATATTGTCGCAGTCGGAG (*Xho*I)
8145cDNAd AAAcaattgATGTTAATAAATGTGTGCCGAG
8145_90r TACActcgagTTTCCTCGCTGGCTCCCAC (*Xho*I)
8159cDNAd ACTgaattcATGCTGCAAAATGTGTGCCG (*Eco*RI)
8159_93r TATCctcgagGACGAACTGAGGAGCGGAATG (*Xho*I)
31457gDNAd AAAgaattcATGAATGTTTGTCGCCTGTGC (*Eco*RI)
31457_90r TAGCctcgagCCTCCTCCACTTCACTCAG (*Xho*I)
7386gDNAd ATTgaattcATGGAGCTCAAGTACTCCGTC (*Eco*RI)
7386_101r TATGctcgagAATCATCAGCTATAGATTCG (*Xho*I)
10269gDNAd AAAgattcATGAACGAGGGCAGGCAG (*Eco*RI)
10269_107r TATActcgagAGTCCTGCACGGCAACGTTG (*Xho*I)
10270gDNAd TTGgaattcATGAATGAGGAGAGCCAGTAC (*Eco*RI)
10270_111r TAGTctcgagCCTCGCCCGGCAGATCC (*Xho*I)
1792_98r TAAActcgagTCTCCACACCAACTGCGAAG (*Xho*I)
11695cDNAd TGTgaattcATGATATGCCGCCTGTGCC (*Eco*RI)
11695_89r TATGctcgagCGTCCTCGGAGAACGGC (*Xho*I)
11696cDNAd TTTgaattcATGATATGCCGACTCTGCC (*Eco*RI)
11696_95r TATGctcgagTCAGTTCGGGCAGCTCCG (*Xho*I)
18764cDNAd ACTgaattcATGGCGCTCCAGTGCCG (*Eco*RI)
18764_96r TATGctcgagATGCCATCTCCTCAACATC (*Xho*I)
11762cDNAd TTAgaattcATGCTGAACATCGTGTGTAGAG (*Eco*RI)
11762_83r TACActcgagAAGAGTCGTCCTCGATGG (*Xho*I)
dv2712d TGAcaattgATGTCCATTTGCGAGGACAC (*Mun*I)
dv2712_95r TTAAGgtcgacGCTTTCTAAGCAGATGGTCG (*Sal*I)
dvZw5d AAgaattcATGATGAGCGGCAGCAAATC (*Eco*RI)
dvZw5_96r TTAAGgtcgacCAATTTCCTCGAGGCCCAG (*Sal*I)
31365d GACgaattcATGGATGCACAGCCG (*Eco*RI)
31365_101r TTAATctcgagAGAAGACGCGGCTTCGC (*Xho*I)
Zw5D57Ed GCCGCAAGAgGGCGCCGCC
Zw5D57Er GGCGGCGCCcTCTTGCGGC
S82MH84Id CAAGCGGAtgGTCatTCGCATGG
S82MH84Ir GCGAatGACcaTCCGCTTGAATTC
2712V16Td GTGCCGCacGTGCCACACAG
2712V16Tr CTGTGTGGCACgtGCGGCACAAG

2712M86Sd CCATCGCtcGGACAGAATTCTGC
2712M86Sr ATTCTGTCCgaGCGATGGACAC
2712 E58Sd CTAGAGGCGtcGCTCTTCCTGC
2712 E58Sr CAGGAAGAGCgaCGCCTCTAGG
Zw5M82Sd CAAGCGCtcGGTCATCGATTCG
Zw5M82Sr GAATCGATGACCgaGCGCTTGAAG
4707d GGTgaattcATGACTGTCTGTGTGCTC (*Eco*RI)
4707_90r GATgtcgacCTCTGTCTGACCGCAAAATTC (*Sal*I)
4282d CAAgaattcATGGATTCCTCTCCTAGCTG (*Eco*RI)
4282_89r TAGgtcgacATCCTCCATGAGCGC (*Sal*I)
10654_29d ACCgaattcATGGATGGATGCGATGG (*Eco*RI)
10654r_135r CGAgtcgacGTCCAAGTCGTGCCAAG (*Sal*I)
17568d ATTgaattcATGACGGACGGCAGC (*Eco*RI)
17568_91r TGTgtcgacTCGCCGCAGGACCTCAAAG (*Sal*I)
10274d ACAgaattcATGCAGGAGGAGGCC (*Eco*RI)
10274_97r TGGgtcgacAAAGACGTTCCTGGCCAC (*Sal*I)

Supplementary Table S2. Data collection, processing and refinement.

Dataset	Remote	Peak
Data collection		
Diffraction source	BL41XU, Spring8	
Wavelength (Å)	0.9	1.28268
Temperature (K)	100	
Detector	PILATUS 6M	
Crystal-to-detector distance (mm)	360	360
Rotation range per image (°)	0.5	0.5
Total rotation range (°)	110	360
Exposure time per image (s)	0.1	0.1
Space group	$P4_12_12$	
a, *b*, *c* (Å)	97.65 97.65 48.48	
α, β, γ (°)	90.00 90.00 90.00	
Resolution range (Å)	97.65-2.0 (2.05-2.0)	97.49-2.43 (2.53-2.43)
Completeness (%)	98.9 (99.7)	99.9 (100.0)
Average redundancy	7.4 (7.8)	21.7 (23.6)
$\langle I/\sigma(I)\rangle$	9.5 (4.4)	28.6 (14.2)
Rmeas (%)	19.5 (80.9)	9.9 (25.5)
$CC_{1/2}$	99.3 (94.7)	99.9 (99.4)
Overall *B* factor from Wilson plot ($Å^2$)	11.7	18.6
Refinement		
R_{fact} *(%)*	16.0	
$R_{free.}$ *(%)*	19.2	
Bonds (Å)	0.02	
Angles (°)	2.27	
Ramachandran plot		
Most favoured (%)	100	
Allowed (%)	0	
PDB entry code	6FP5	

Values for the highest resolution shell are given in parentheses

Supplementary Table S3. Binary interactions between ZADs tested with yeast two-hybrid assay.

Cluster	Protein 1	Protein 2	Y2H interactions	identity positions,%	consensus positions,%
1	CG31457 (1-90)	CG3941 (1-111)	+	39	52
	CG31457 (1-90)	CG31365 (1-103)	-	39	51
	CG17328 (1-109)	CG10321 (1-90)	-	35	51
	CG31365 (1-103)	CG3941 (1-111)	-	35	49
2	CG8145 (1-90)	CG11762 (1-83)	-	40	61
	CG8145 (1-90)	CG8159 (1-93)	+	45	60
	CG1792 (1-98)	CG18764 (1-96)	+	49	57
	CG8159 (1-93)	CG11762 (1-83)	-	35	56
	CG9797 (9-114)	CG8159 (1-93)	+	39	55
	CG9793 (1-124)	CG8159 (1-93)	-	43	53
	CG9793 (1-124)	CG11762 (1-83)	-	35	52
	CG18764 (1-96)	CG9793 (1-124)	-	43	52
	CG18764 (1-96)	CG9797 (9-114)	-	40	52
	CG9793 (1-124)	CG9797 (9-114)	-	40	51
	CG9793 (1-124)	CG8145 (1-90)	-	33	51
	CG9797 (9-114)	CG11762 (1-83)	-	35	49
	CG1792 (1-98)	CG9797 (9-114)	-	36	49
	CG1792 (1-98)	CG9793 (1-124)	-	39	48
	CG9797 (9-114)	CG8145 (1-90)	-	29	48
3	CG10269 (1-107)	CG10270 (1-111)	+	75	86
	CG10274 (1-97)	CG10270 (1-111)	+	65	81
	CG10274 (1-97)	CG10269 (1-107)	+	61	80
	CG7386 (1-101)	CG10269 (1-107)	+	52	70
	CG7386 (1-101)	CG10270 (1-111)	+	49	66
	CG10274 (1-97)	CG7386 (1-101)	-	43	61
	CG10654 (29-135)	CG7386 (1-101)	-	31	46
	CG10654 (29-135)	CG10269 (1-107)	-	31	44
	CG10654 (29-135)	CG10270 (1-111)	-	29	42
	CG10274 (1-97)	CG10654 (29-135)	-	26	40
4	CG11695 (1-89)	CG11696 (1-95)	+	58	65
	CG17568 (1-91)	CG4148 (1-90)	-	46	62
	CG33133 (1-125)	CG11695 (1-89)	+	35	49
	CG15073 (1-95)	CG33133 (1-125)	+	42	48
	CG33133 (1-125)	CG17568 (1-91)	-	24	47
	CG15073 (1-95)	CG11695 (1-89)	-	30	44
	CG15073 (1-95)	CG11696 (1-95)	-	28	42
	CG17568 (1-91)	CG11696 (1-95)	-	26	41
	CG33133 (1-125)	CG11696 (1-95)	-	27	41
	CG33133 (1-125)	CG4148 (1-90)	-	18	37
	CG17568 (1-91)	CG11695 (1-89)	-	25	36

	CG15073 (1-95)	CG4148 (1-90)	-	15	36
	CG11695 (1-89)	CG4148 (1-90)	-	22	32
	CG15073 (1-95)	CG17568 (1-91)	-	18	31
	CG11696 (1-95)	CG4148 (1-90)	-	17	31

Supplementary Table S4. Dimerization interfaces of CG2712 and Grauzone ZADs as revealed by their crystal structures.

	Buried area, Å^2		Residue 1 (monomer A)	Atom	Residue (monomer B)	Atom	Distance, Å
CG2712	1150	**Hydrogen bonds**					
		1	H84	NE2	S50	O	2.81
		2	S50	O	H84	NE2	2.76
		Salt bridges					
		1	E78	OE1	K72	NZ	3.41*
		2	R91	NH2	E58	OE1	2.64
		3	R91	NH1	E58	OE2	2.87
		4	E58	OE1	R91	NH2	2.74
		5	E58	OE2	R91	NH1	2.88
		Hydrophobic interactions (ΔG^{int} = -18.2 kcal/mol)**					
		V16, H18, L46, L52, L61, P62, K72, L73, L75, F79, V83, M86, L90					
Grauzone	1039	**Hydrogen bonds**					
		1	R5	O	Q74	NE2	2.81
		2	L6	O	Q74	NE2	2.70
		3	Q74	NE2	L6	O	2.70
		4	Q74	NE2	R5	O	2.81
		Hydrophobic interactions (ΔG^{int} = -19.6 kcal/mol)**					
		Q71, F63, I48, Y67, F66, F40, C7, E47, L8, F38, W39, Y77, H37, I70, A73, V60, C56, Q59					

* - Side chain of Glu78 has different conformations in each monomer within a dimer, with only one forming salt bridge.

** - estimated by PDBePISA server.

Supplementary Table S5. Residues stabilizing CG2712 ZAD homodimer along MD trajectory.

Polar residues		
Donor	Acceptor	Occupancy
ARG91-Side	***GLU58-Side***	***199.97%***
SER76-Side	SER76-Side	85.51%
SER82-Side	VAL16-Main	67.22%
LYS80-Side	SER50-Main	44.71%
HIS84-Side	*SER50-Main*	*41.11%*
LYS80-Side	CYS49-Main	13.29%
LYS72-Side	***GLU78-Side***	***6.49%***
ARG88-Side	**GLU53-Side**	**3.58%**
Hydrophobic residues		
K72, L75, P60, R88, V83, F79, L90, I45, V16, L73, L61, L52, H18, L46, M86, P62		

Contacts in bold are salt bridges.
Contacts in italic are found in X-ray structure of CG2712 ZAD homodimer and listed in the Table S4 above.

Supplementary Table S6. Residues stabilizing Zw5 ZAD homodimer along MD trajectory.

Polar residues		
Donor	Acceptor	Occupancy
ARG15-Side	**ASP87-Side**	**139.61%**
SER86-Side	THR16-Main	41.32%
ARG91-Side	**GLU54-Side**	**36.49%**
ARG91-Side	**ASP57-Side**	**28.26%**
ARG19-Side	**ASP85-Side**	**25.69%**
SER86-Side	ARG15-Main	15.89%
ARG15-Side	SER86-Side	15.32%
SER86-Side	THR18-Side	14.72%
LYS80-Side	THR49-Main	12.44%
LYS80-Side	CYS48-Main	9.29%
ALA59-Main	LEU90-Main	7.95%
LEU92-Main	ASP57-Side	2.99%
ARG91-Side	PRO55-Main	1.56%
GLN75-Side	GLN75-Side	1.49%
LEU92-Main	ASP57-Main	1.34%
LEU51-Main	GLY50-Main	1.14%
ARG91-Side	SER52-Side	1.13%
ARG19-Side	**ASP85-Main**	**1.13%**
Hydrophobic residues		
F79, L90, L72, A76, L45, L73, L52, M61, V83, P62, A59, L92, P55, I65, M82.		

Contacts in bold are salt bridges.

Supplementary Table S7. Residues stabilizing CG2712-Zw5 ZAD heterodimer along MD trajectory.

Polar residues		
Donor (CG2712)	Acceptor (Zw5)	Occupancy
ARG91-Side	**GLU58-Side**	**99.99%**
ARG91-Side	**GLU54-Side**	**75.70%**
SER82-Side	THR16-Main	37.38%
LYS80-Side	THR49-Main	27.31%
HIS84-Side	THR49-Main	12.85%
LYS80-Side	SER50-Main	7.50%
SER86-Side	ARG15-Main	4.64%
HIS18-Side	SER86-Side	3.58%
SER86-Side	VAL16-Main	3.30%
ARG15-Side	**ASP87-Side**	**4.08%**
ARG91-Side	SER52-Main	2.27%
LYS72-Side	GLN75-Side	1.86%
Hydrophobic residues		
CG2712: V16, L46, L52, F60, L61, P62, L73, L75, F79, V83, L90		
Zw5: L45, L51, M61, P62, L72, L73, A76, F79, M82, V83, I84, L90		

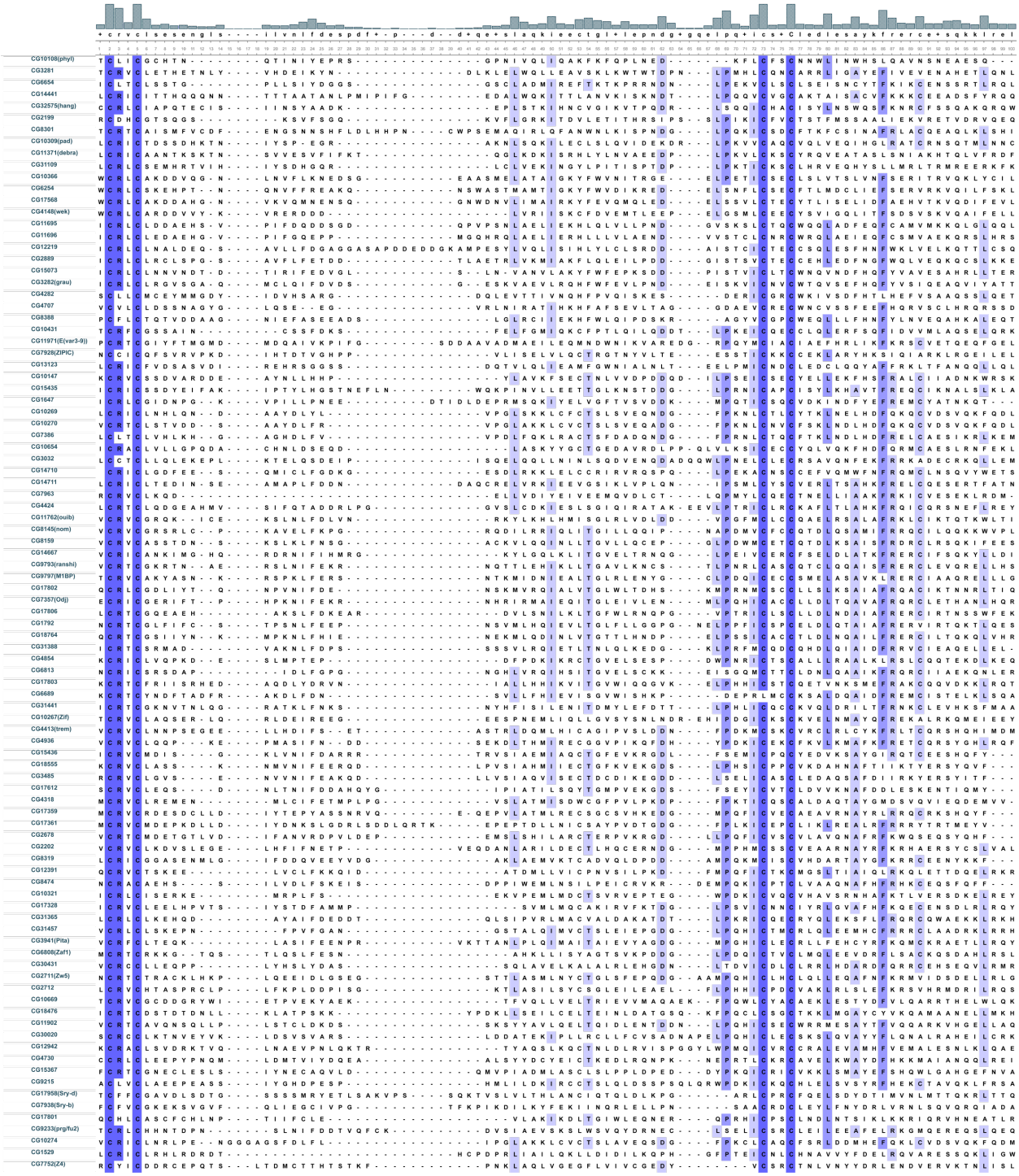
Contacts in bold are salt bridges.

Supplementary Table S8. Residues stabilizing CG2712- dv2712 ZAD heterodimer along MD trajectory.

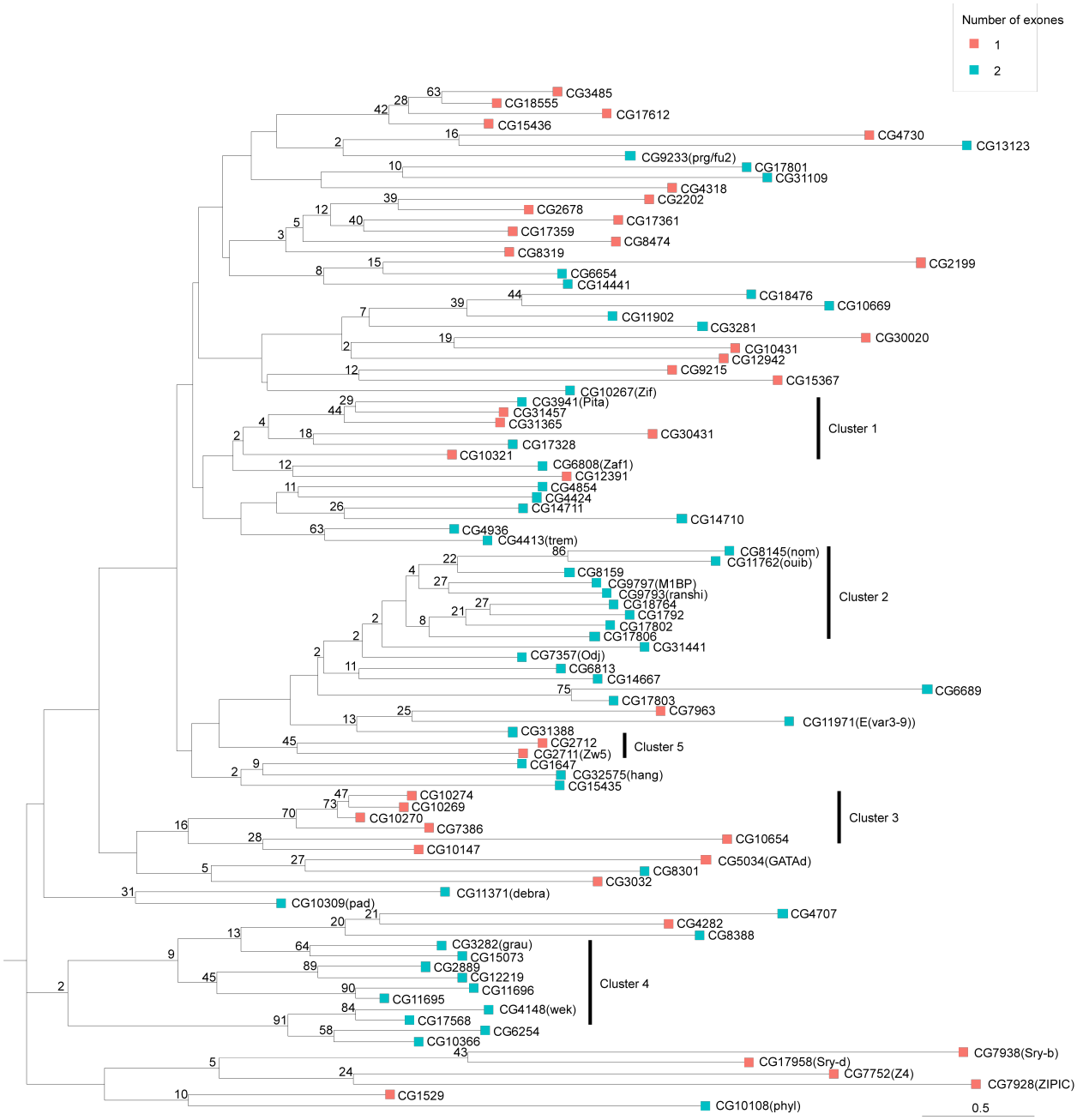
Polar residues		
Donor (dv2712)	Acceptor (CG2712)	Occupancy
ARG93-Side	**GLU58-Side**	**98.91%**
THR88-Side	VAL16-Main	95.40%
ARG85-Side	VAL20-Main	20.45%
LYS80-Side	ASN53-Side	16.51%
LYS82-Side	SER50-Main	15.00%
ARG91-Side	**GLU60-Side**	**13.41%**
ARG85-Side	ASN22-Side	11.30%
HSE84-Side	ALA52-Main	8.06%
LYS80-Side	CYS51-Main	6.19%
LYS80-Side	ALA52-Main	5.44%
ARG15-Side	**ASP89-Side**	**1.81%**
LYS82-Side	CYS49-Main	1.58%
LYS72-Side	**GLU80-Side**	**1.58%**
LYS72-Side	GLN77-Side	1.47%
ARG93-Side	GLU53-Main	1.37%
ARG85-Side	CYM21-Side	1.09%
Hydrophobic residues		
dv2712:V20, L48, A52, L54, F62, L63, L74, L75, V78, F81, A85, L86		
CG2712: V16, L46, L52, L61, P62, L73, L75, F79, V83, M86, I89, L90		

SUPPLEMENTARY FIGURES

Supplementary Figure S1. Multiple sequence alignment of *D. melanogaster* ZAD domains. Amino acids are numbered according to the consensus sequence in the alignment. The sequence logo for the alignment is shown in the Supplementary Figure S5.



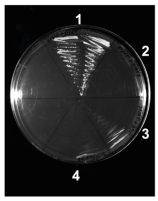
Supplementary Figure S2. Clustering of ZAD-domains. Values at branching points represent bootstrap confidence values. The number of exons encoding ZADs is indicated by color.



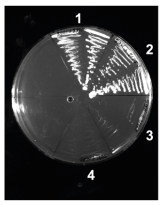
Supplementary Figure S3. Yeast two-hybrid assay results. AD stands for GAL4 Activation Domain, BD – GAL4 DNA-binding Domain. Cluster numbering is according to the Supplementary Table S3.

Cluster 1

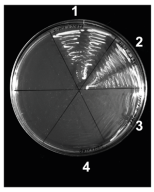
	BD CG31365 (1-103)	
AD CG31365(1-103)	**1**	+
AD CG31457(1-90)	**2**	-
AD CG3941(1-111)	**3**	-
AD	**4**	-



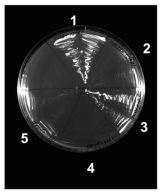
	BD CG3941 (1-111)	
AD CG3941(1-111)	**1**	+
AD CG31457(1-90)	**2**	+
AD CG31365(1-103)	**3**	-
AD	**4**	-



	BD CG31457 (1-90)	
AD CG31457(1-90)	**1**	+
AD CG3941(1-111)	**2**	+
AD CG31365(1-103)	**3**	-
AD	**4**	-

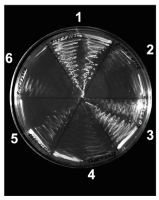


	BD 17328 (1-109)	
AD 17328 (1-109)	**1**	+
AD 10321 (1-90)	**2**	-
AD	**4**	-
	BD 10321 (1-90)	
AD 10321 (1-90)	**3**	+
AD	**5**	-

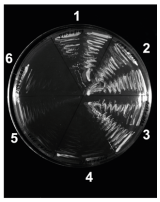


Cluster 2

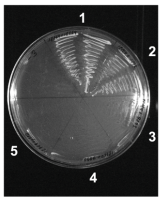
	BD CG8145 (1-91)	
AD CG8145(1-91)	**1**	+
AD CG11762(1-83)	**2**	-
AD CG8159(1-93)	**3**	+
AD CG9793(1-124)	**4**	-
AD CG9797 (9-114)	**5**	-
AD	**6**	-



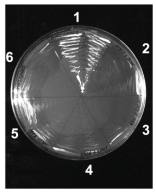
	BD CG8159 (1-93)	
AD CG8159(1-93)	**1**	+
AD CG8145(1-91)	**2**	+
AD CG9797 (9-114)	**3**	+
AD CG11762(1-83)	**4**	-
AD CG9793(1-124)	**5**	-
AD	**6**	-



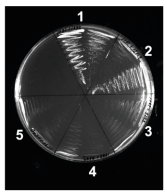
	BD 18764 (1-96)	
AD CG18764(1-96)	**1**	+
AD CG1792(1-98)	**2**	+
AD CG9793(1-124)	**3**	-
AD CG9797 (9-114)	**4**	-
AD	**5**	-



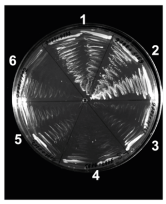
	BD CG11762 (1-83)	
AD CG11762(1-83)	**1**	+
AD CG8145(1-91)	**2**	-
AD CG8159(1-93)	**3**	-
AD CG9793(1-124)	**4**	-
AD CG9797 (9-114)	**5**	-
AD	**6**	-



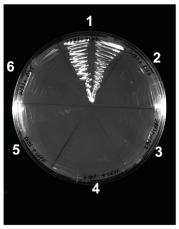
	BD CG1792(1-98)	
AD CG1792(1-98)	**1**	+
AD CG18764(1-96)	**2**	+
AD CG9797 (9-114)	**3**	-
AD CG9793(1-124)	**4**	-
AD	**5**	-



	BD CG9797 (9-114)	
AD CG9797 (9-114)	**1**	+
AD CG8159(1-93)	**2**	+
AD CG18764(1-96)	**3**	-
AD CG9793(1-124)	**4**	-
AD CG11762(1-83)	**5**	-
AD CG1792(1-98)	**6**	-
AD CG8145(1-91)	**7**	-
AD	**8**	-

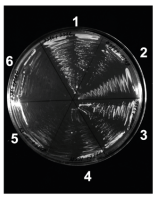


	BD CG9793(1-124)	
AD CG9793(1-124)	**1**	+
AD CG8159(1-93)	**2**	-
AD CG11762(1-83)	**3**	-
AD CG18764(1-96)	**4**	-
AD CG9797 (9-114)	**5**	-
AD CG8145(1-91)	**6**	-
AD CG1792(1-98)	**7**	-
AD	**8**	-

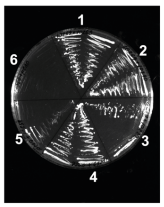


Cluster 3

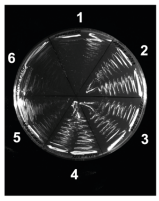
	BD CG7386 (1-101)	
AD CG7386(1-101)	1	+
AD CG10269(1-107)	2	+
AD CG10270(1-111)	3	+
AD CG10274(1-97)	4	-
AD CG10654(29-135)	5	-
AD	6	-



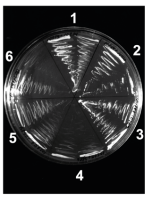
	BD CG10269 (1-107)	
AD CG10269(1-107)	1	+
AD CG10270(1-111)	2	+
AD CG10274(1-97)	3	+
AD CG7386(1-101)	4	+
AD CG10654(29-135)	5	-
AD	6	-



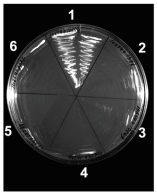
	BD CG10270 (1-111)	
AD CG10270(1-111)	1	+
AD CG10269(1-107)	2	+
AD CG10274(1-97)	3	+
AD CG7386(1-101)	4	+
AD CG10654(29-135)	5	-
AD	6	-



	BD CG10274 (1-97)	
AD CG10274(1-97)	1	+
AD CG10270(1-111)	2	+
AD CG10269(1-107)	3	+
AD CG7386(1-101)	4	-
AD CG10654(29-135)	5	-
AD	6	-

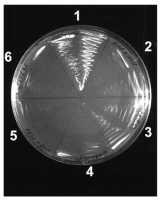


	BD CG10654 (29-135)	
AD CG10654(29-135)	1	+
AD CG7386(1-101)	2	-
AD CG10269(1-107)	3	-
AD CG10270(1-111)	4	-
AD CG10274(1-97)	5	-
AD	6	-

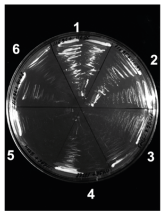


Cluster 4

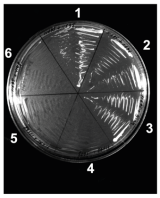
	BD CG4148 (1-90)	
AD CG4148(1-90)	1	+
AD CG17568(1-91)	2	-
AD CG33133(1-125)	3	-
AD CG15073(1-95)	4	-
AD CG11695(1-89)	5	-
AD CG11696(1-95)	6	-
AD	7	-



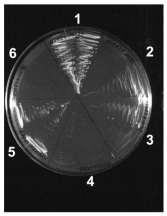
	BD CG11696 (1-95)	
AD CG11696(1-95)	1	+
AD CG11695(1-89)	2	+
AD CG15073(1-95)	3	-
AD CG17568(1-91)	4	-
AD CG33133(1-125)	5	-
AD CG4148(1-90)	6	-
AD	7	-



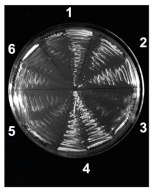
	BD CG11695 (1-90)	
AD CG11695(1-89)	1	+
AD CG17568(1-91)	2	+
AD CG33133(1-125)	3	+
AD CG15073(1-95)	4	-
AD CG11695(1-89)	5	-
AD CG11696(1-95)	6	-
AD	7	-



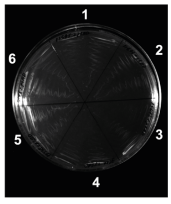
	BD CG15073 (1-95)	
AD CG15073(1-95)	1	+
AD CG11696(1-95)	2	-
AD CG4148(1-90)	3	-
AD CG17568(1-91)	4	-
AD CG33133(1-125)	5	+
AD CG11695(1-89)	6	-
AD	7	-



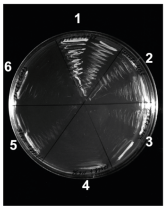
	BD CG33133 (1-125)	
AD CG33133(1-125)	1	+
AD CG11695(1-89)	2	+
AD CG4148(1-90)	3	-
AD CG15073(1-95)	4	+
AD CG11696(1-95)	5	-
AD	6	-



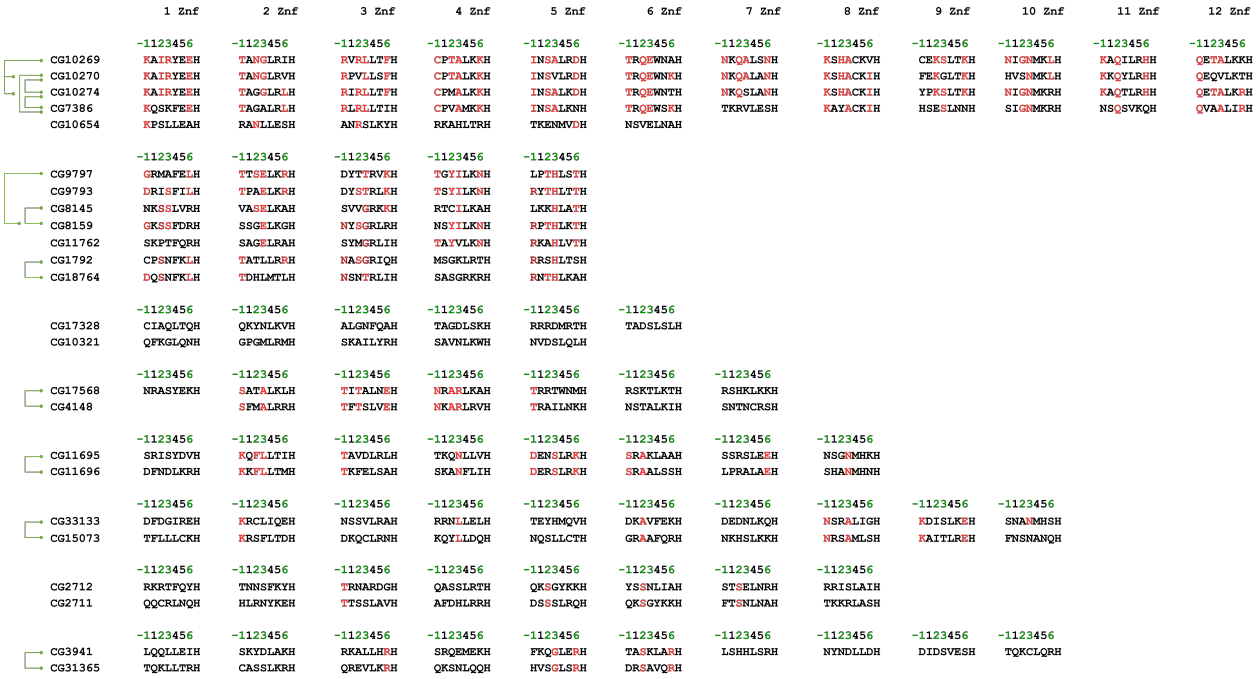
	AD	
AD CG11695(1-89)	1	-
AD CG11696(1-95)	2	-
AD CG15073(1-95)	3	-
AD CG33133(1-125)	4	-
AD CG17568(1-91)	5	-
AD CG4148(1-90)	6	-



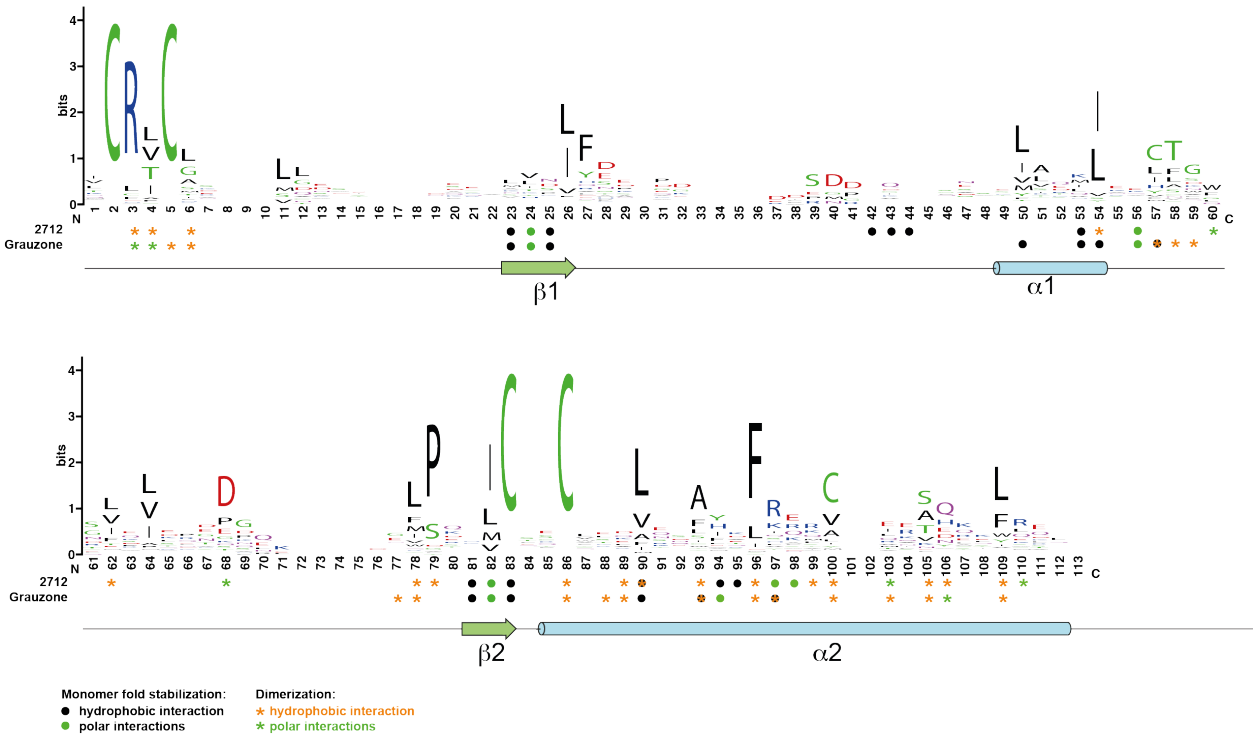
	BD CG17568 (1-91)	
AD CG17568(1-91)	1	+
AD CG4148(1-90)	2	-
AD CG15073(1-95)	3	-
AD CG11696(1-95)	4	-
AD CG11695(1-89)	5	-
AD CG15073(1-95)	6	-



Supplementary Figure S4. Conservation of the residues in DNA-binding positions of zinc-fingers of ZAD-C2H2 proteins from paralogous clusters containing proteins with heterodimerizing ZADs. DNA-binding positions are shown in green. Conserved residue at these positions are shown in red. Green lines indicate heterodimerizing pairs of ZADs.

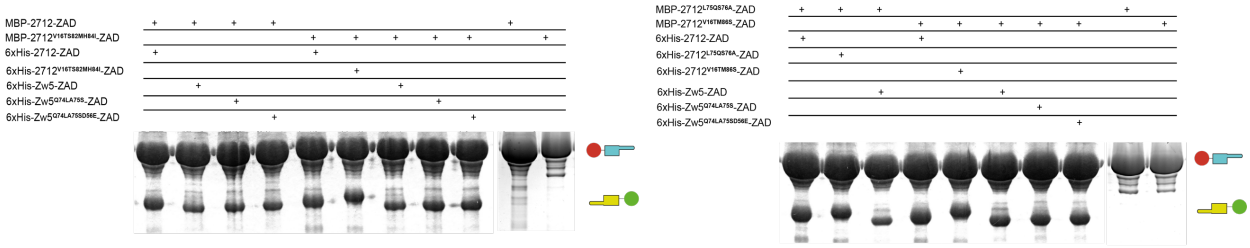


Supplementary Figure S5. Distribution of the residues involved in the monomer folding and dimer formation (according to crystal structures) of Grauzone and CG2712 ZADs shown at the sequence logo for all *Drosophila melanogaster* ZAD based on the multiple sequence alignment shown in the Supplementary Figure S1. Residue numbering is according to the consensus of multiple sequence alignment.

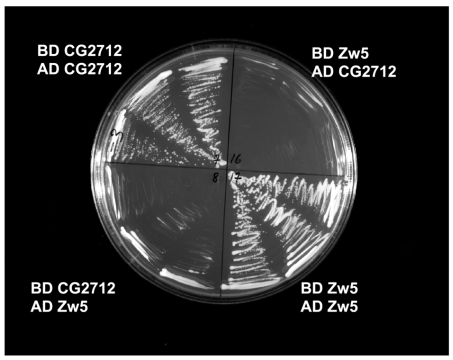


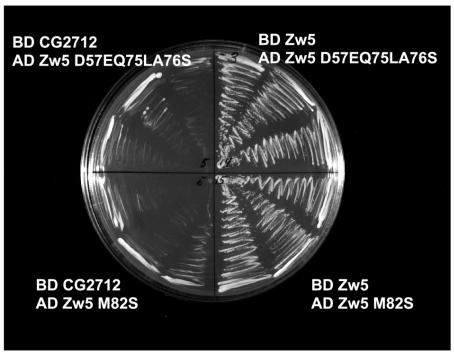
Supplementary Figure S6. (A) Testing of the impact of point mutations on the hetero-dimerization ability of CG2712/Zw5 proteins in MBP pull-down assay after co-expression in bacteria. Pluses show the MBP- and 6xHis-tagged proteins co-expressed in corresponding lane. Small cartoons on the right show the positions of MBP-fused ZADs (red circle, MBP is 46kDa) and Thioredoxin-6xHis-fused ZADs (green circle, Thioredoxin-6xHis is 17kDa). **(B)** Testing of the impact of point mutations on the hetero-dimerization ability of CG2712/Zw5 ZADs in yeast two-hybrid assay.

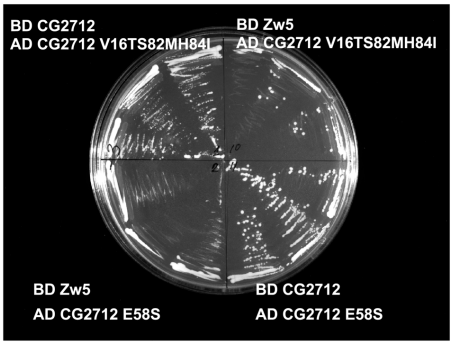
A

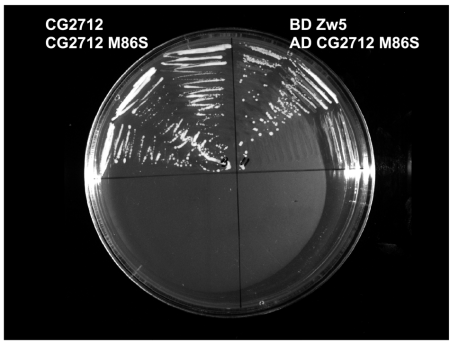


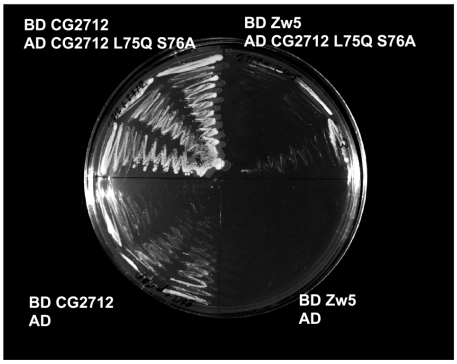
B



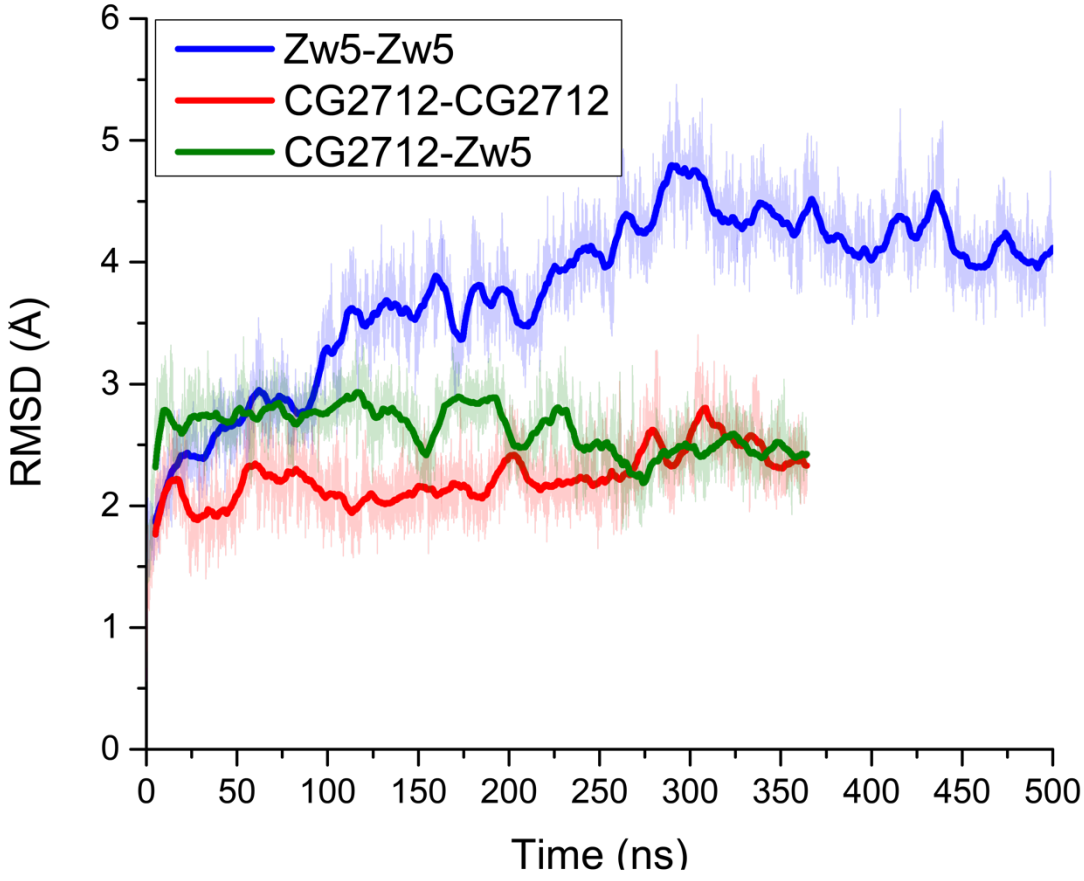




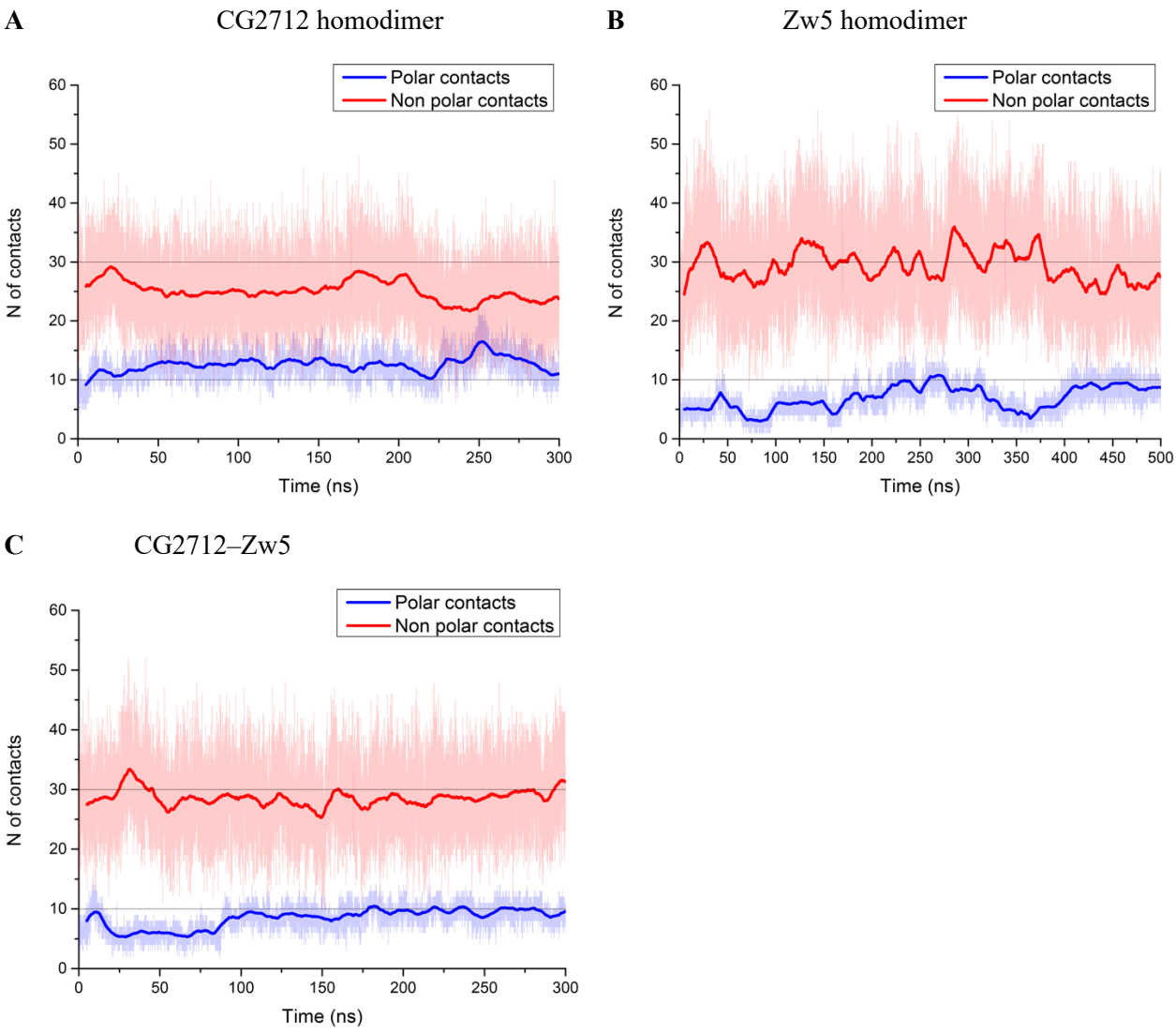




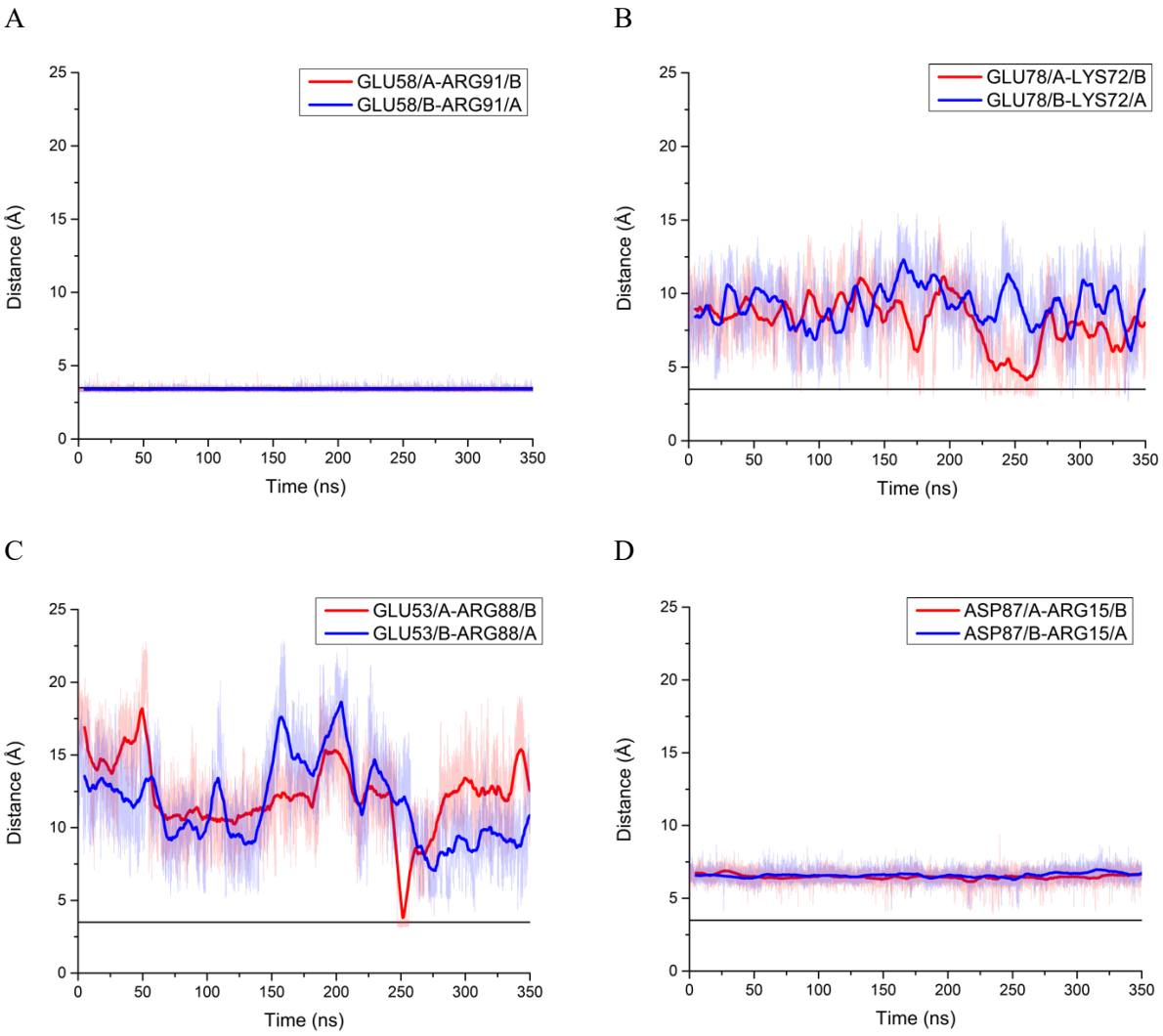
Supplementary Figure S7. Evaluation of RMSD along MD trajectory for corresponding dimers. Bold line shows moving average with 10 ns window.



Supplementary Figure S8. Evolution of polar and hydrophobic interactions along MD trajectory. Bold lines show moving average with 10 ns window.

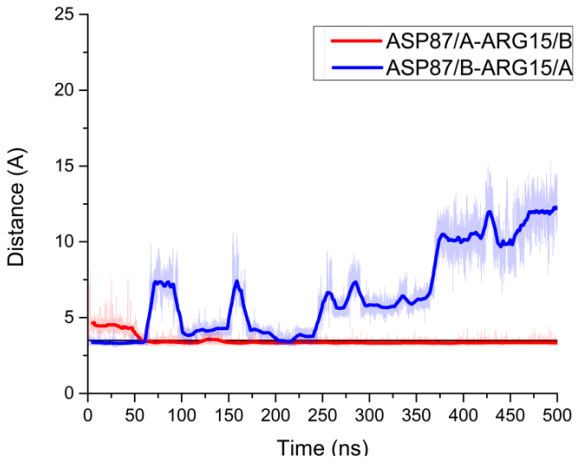


Supplementary Figure S9. Evolution of salt bridges of CG2712 ZAD homodimer along MD trajectory. Bold line shows moving average with 10 ns window.

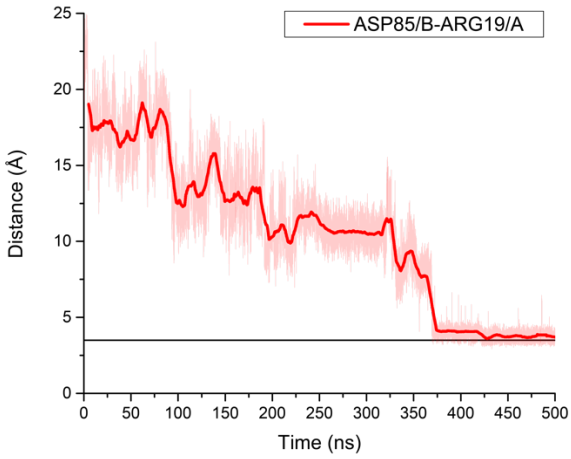


Supplementary Figure S10. Evolution of salt bridges of Zw5 ZAD homodimer along MD trajectory. Bold line shows moving average with 10 ns window.

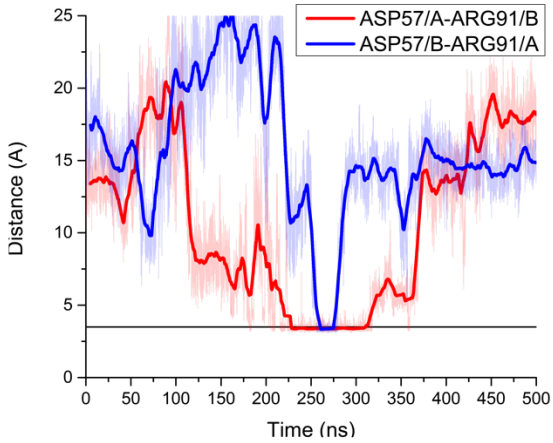
A



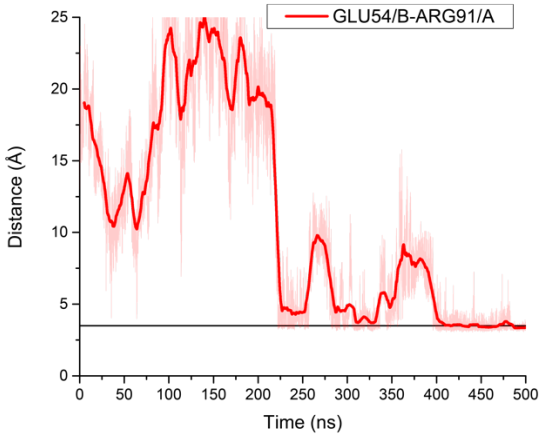
B



C

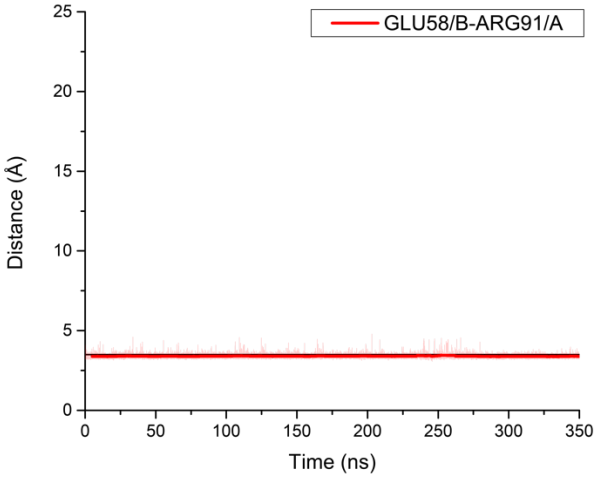


D

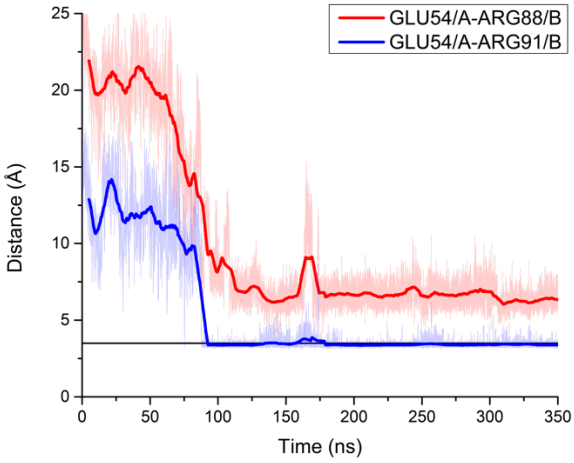


Supplementary Figure S11. Evolution of salt bridges of CG2712-Zw5 ZAD heterodimer along MD trajectory. Bold line shows moving average with 10 ns window.

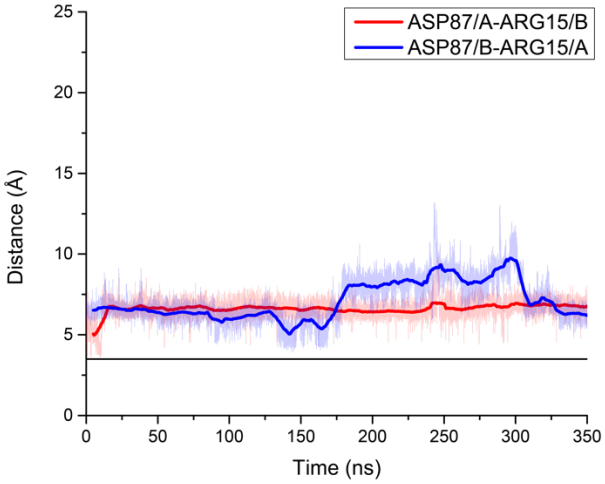
A



B



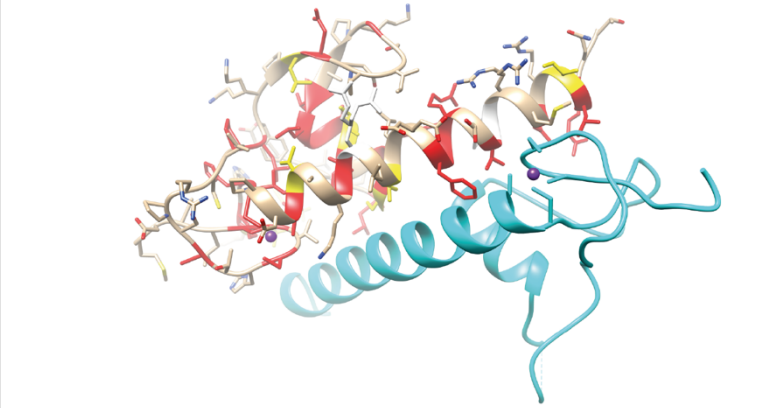
C



Supplementary Figure S12. (A) Crystal structure of CG2712 ZAD-domain colored by homology with Zw5 ZAD (identical residues are shown in red, conserved – in yellow) according to sequence alignment shown above. **(B)** PDBePISA calculation of buried surface area and dimer solvation energy contribution (ΔG) of residues involved in CG2712 ZAD dimerization. The ΔG values (calculated from CG2712 ZAD structure) of residues which are identical/conserved in Zw5/dv2712/dvZw5 ZADs are shown in red/yellow (according to the alignment shown at the Figure 2). The non-conserved residues of CG2712 ZAD contributing to overall ΔG dimer solvation energy are shown in green. Residues which are not involved in dimerization are not shown for clarity. **(C)** PDBePISA analysis of ΔG solvation energy influence of residues involved in Zw5 dimerization (according to the model built with Robetta server), residues identical/conserved in CG2712/dv2712/dvZw5 are shown in red/yellow (according to the alignment shown at the Figure 2). In green are shown residues conserved only in Zw5/dvZw5 contributing to overall ΔG.

A

```
                1         10        20        30        40        50        60        70        80        90
2712  MTTMSPPSPMDALCRVCHTASPRCLPLFKPLDDPISGKPATLASILSYCSGLEILEAE-LFLPHHICPDCVAKLRLSLEFKRSVHRMDRILRQSHADFCR
                1         10        20        30        40        50        60        70        80        90
Zw5   MMNSKIAEVVVLNCRTCTRACKLHKPLQEEIDLGSEGS-TTLASMLNYCTGLSFEPQDGAAMPQHICLHCLQLLEQAFNFKRMVIDSDELLRLGLDEARC
```



B

N	Residue	Buried surface area	ΔiG Solvation energy effect, kcal/mol	zw5	dv2712	dvZw5
14	CYS	0.25	0,00			
15	ARG	42.68	-0,52	-0,52	-0,52	-0,52
16	VAL	74.71	0,50		0,50	
17	CYS	10.93	0,10	0,10	0,10	0,10
18	HIS	42.42	0,20			
46	LEU	17.24	0,28	0,28	0,28	0,28
49	CYS	24.00	-0,15	-0,15	-0,15	-0,15
50	SER	56.74	0,27	0,27	0,27	0,27
51	GLY	3.75	0,06			
52	LEU	82.77	1,32	1,32	1,32	1,32
53	GLU	5.92	-0,03			
54	ILE	3.01	0,05	0,05	0,05	0,05
58	GLU	18.99	-0,26	-0,26	-0,26	-0,26
61	LEU	49.53	0,79	0,79	0,79	0,79
62	PRO	11.39	0,18	0,18	0,18	0,18
65	ILE	0.50	0,01	0,01	0,01	0,01
69	CYS	3.02	-0,01	-0,01	-0,01	-0,01
72	LYS	56.33	-0,17			
73	LEU	9.11	0,13	0,13	0,13	0,13
75	LEU	25.75	0,41			
76	SER	36.69	0,31			
79	PHE	136.30	2,18	2,18	2,18	2,18
80	LYS	33.67	0,05	0,05	0,05	0,05
82	SER	13.56	-0,04			
83	VAL	71.74	1,14	1,14	1,14	1,14
84	HIS	43.62	0,09			
86	MET	71.01	1,51			
87	ASP	36.39	-0,12	-0,12	-0,12	-0,12
90	LEU	95.13	1,31	1,31	1,31	1,31
91	ARG	76.83	-1,18	-1,18	-1,18	-1,18
Overall ΔiG kcal/mol			16,82			

C

N	Residue	Buried surface area	ΔiG Solvation energy effect, kcal/mol	2712	dv2712	dvZw5
14	CYS	0.37	0,00			
15	ARG	36.08	-0,42	-0,42	-0,42	-0,42
16	THR	74.13	0,50			0,5
17	CYS	13.91	0,22	0,22	0,22	0,22
18	THR	37.38	0,60			0,6
45	LEU	10.14	0,16	0,16	0,16	0,16
48	CYS	22.41	-0,21	-0,21	-0,21	-0,21
49	THR	72.18	0,41	0,41	0,41	0,41
50	GLY	1.68	0,03	0,03	0,03	0,03
51	LEU	76.91	1,23	1,23	1,23	1,23
52	SER	11.30	-0,12	-0,12	-0,12	-0,12
53	PHE	8.76	0,14	0,14	0,14	0,14
56	GLN	28.86	-0,36			-0,36
60	ALA	41.30	0,41			0,41
61	MET	50.23	0,92	0,92	0,92	0,92
62	PRO	8.53	0,14	0,14	0,14	0,14
65	ILE	0.34	0,01	0,01	0,01	0,01
68	HIS	0.31	0,01	0,01	0,01	0,01
69	CYS	6.89	0,10	0,1	0,1	0,1
72	LEU	53.53	0,82		0,82	0,82
73	LEU	7.27	0,10	0,1	0,1	0,1
75	GLN	45.16	-0,10		-0,1	-0,1
76	ALA	34.88	0,55			0,55
79	PHE	136.94	2,19	2,19	2,19	2,19
80	LYS	40.97	0,46	0,46	0,46	0,46
82	MET	53.07	1,03			1,03
83	VAL	71.00	1,14	1,14		1,14
84	ILE	23.11	0,37		0,37	0,37
86	SER	43.60	0,33		0,33	0,33
87	ASP	38.05	-0,06	-0,06	-0,06	-0,06
89	LEU	8.03	0,13	0,13	0,13	0,13
90	LEU	92.99	1,32	1,32	1,32	1,32
91	ARG	72.47	-1,15	-1,15	-1,15	-1,15
Overall ΔiG kcal/mol			21,80	13,50	14,06	21,80