PLOS ONE

RESEARCH ARTICLE

The 9aaTAD Transactivation Domains: From Gal4 to p53

Martin Piskacek[1☯‡*], Marek Havelka[1], Martina Rezacova[1], Andrea Knight[2☯‡]

1 Laboratory of Cancer Biology and Genetics, Department of Pathological Physiology, Faculty of Medicine, Masaryk University Brno, Kamenice 5, 625 00, Brno, Czech Republic, **2** Gamma Delta T Cell Laboratory, Department of Pathological Physiology, Faculty of Medicine, Masaryk University Brno, Kamenice 5, 625 00, Brno, Czech Republic

☯ These authors contributed equally to this work.
‡ These authors are joint senior authors on this work.
* Piskacek@med.muni.cz

Citation: Piskacek M, Havelka M, Rezacova M, Knight A (2016) The 9aaTAD Transactivation Domains: From Gal4 to p53. PLoS ONE 11(9): e0162842. doi:10.1371/journal.pone.0162842

Editor: Sumitra Deb, Virginia Commonwealth University, UNITED STATES

Received: July 1, 2016

Accepted: August 29, 2016

Published: September 12, 2016

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported by Ministry of Health of the Czech Republic NT14310 (M.P.) and 15-32935A (A.K.).

Competing Interests: The authors have declared that no competing interests exist.

Abstract

The family of the Nine amino acid Transactivation Domain, 9aaTAD family, comprises currently over 40 members. The 9aaTAD domains are universally recognized by the transcriptional machinery from yeast to man. We had identified the 9aaTAD domains in the p53, Msn2, Pdr1 and B42 activators by our prediction algorithm. In this study, their competence to activate transcription as small peptides was proven. Not surprisingly, we elicited immense 9aaTAD divergence in hundreds of identified orthologs and numerous examples of the 9aaTAD species' convergence. We found unforeseen similarity of the mammalian p53 with yeast Gal4 9aaTAD domains. Furthermore, we identified artificial 9aaTAD domains generated accidentally by others. From an evolutionary perspective, the observed easiness to generate 9aaTAD transactivation domains indicates the natural advantage for spontaneous generation of transcription factors from DNA binding precursors.

Introduction

The transcription factors are versatile regulators of gene expression. Their DNA binding domains, DBD, specifically recognize regulatory elements and their transactivation domains, TAD, mediate activation of transcription. A number of tested TADs is functional in both yeast and mammals e.g. Gal4 and p53 transcription factors [1,2].

The Nine amino acid Transactivation Domain, 9aaTAD, is universally recognized by the transcriptional machinery in eukaryotes. Currently, the 9aaTAD family comprises of over 40 members including Gal4, Oaf1, Pip2, Pdr1, Pdr3, Leu3, Tea1, Pho4, Gln3, Gcn4, Msn2, Msn4, Rtg3, E2A, MLL, p53-TADI, p53-TADII, FOXO3, NF-kB, NFAT, CEBPA/E, ESX, ELF3, ETV1, KLF2/4, EBNA2, VP16, HSF1, HSF2, HsfA, Gli3, Sox18, PIF, Dreb2a, MTF1, OREB1, WRKY45, NS1, MKL1, VP16, EBNA2, KBP220, ECapLL, P201, AH, and B42 transcription factors. We and others have shown the 9aaTAD domains have competence to activate transcription as small peptides [3–17]. We have established the 9aaTAD prediction service online (www.piskacek.org). The 9aaTADs are annotated on protein database UniProt (www.uniprot.org/9aaTAD).

Previously, we predicted two distinct 9aaTAD domains for p53 protein with conserved proximal Leucines [4,16]. Both 9aaTAD domains, called the 9aaTAD-I and the 9aaTAD-II corresponded with the transactivation regions interacting with the KIX domain of CBP [2,18,19]. We have reported that the first transactivation domain of the p53 protein has the highest similarity to the 9aaTAD of the transcription factor E2A (helical structure of the 9aaTAD-I about 12 aa), while the second transactivation domain of p53 has the highest similarity to the 9aaTAD of transcription factor MLL (shorter helical structure of the 9aaTAD-II about 9 aa) [16].

In this study we aimed to determine the evolutional conservation of the 9aaTAD domains and to prove their competence to activate transcription.

Materials and Methods

Constructs

The construct pBTM116-HA (BHA) was generated by Klenow fill-in of oligonucleotides and subcloned in to pBTM116 (B) EcoRI. G1-G45 and H1-H45 were generated by PCR and subcloned in to pBTM116 EcoRI and BamHI sites. All constructs were sequenced by Eurofins Genomics. All construct information, primer sequences and further detailed information are available on request.

Assessment of enzyme activities

β-galactosidase activity was determined in the yeast strain L40 crude extracts using the ONPG substrate [3]. The average value of β-galactosidase activities from at least three experiments is presented as a percentage with standard deviation (means and plusmn; SD; n = 3).

Protein purification

The GST-KIX expression constructs kindly provided by Isabelle Lemasson [20] were then transformed into *Escherichia coli* BL21. Cells were grown in LB medium at 37°C and induced with 1 mM IPTG for 2 h at 25°C. The cells were harvested by centrifugation, suspended in lysis buffer with complete protease inhibitors EDTA-free (Roche 04719948001). Protein extracts were cleared by centrifugation for 10 min at 12,000 g, diluted 10 times with GST buffer (50 mM Tris-HCl pH 6.8, 200 mM NaCl, 5% Glycerol), applied on 20 uL GST beads slurry, incubated for 10 min with gently inverting and washed 4 times. Protein purity was estimated by SDS-PAGE. The yeast strain L40 crude extracts were produced by lysis (lysis buffer Roche 04719948001 with complete protease inhibitors EDTA-free) for 15 min at 25°C and clarified by centrifugation 14.000g for 10 min at 4°C.

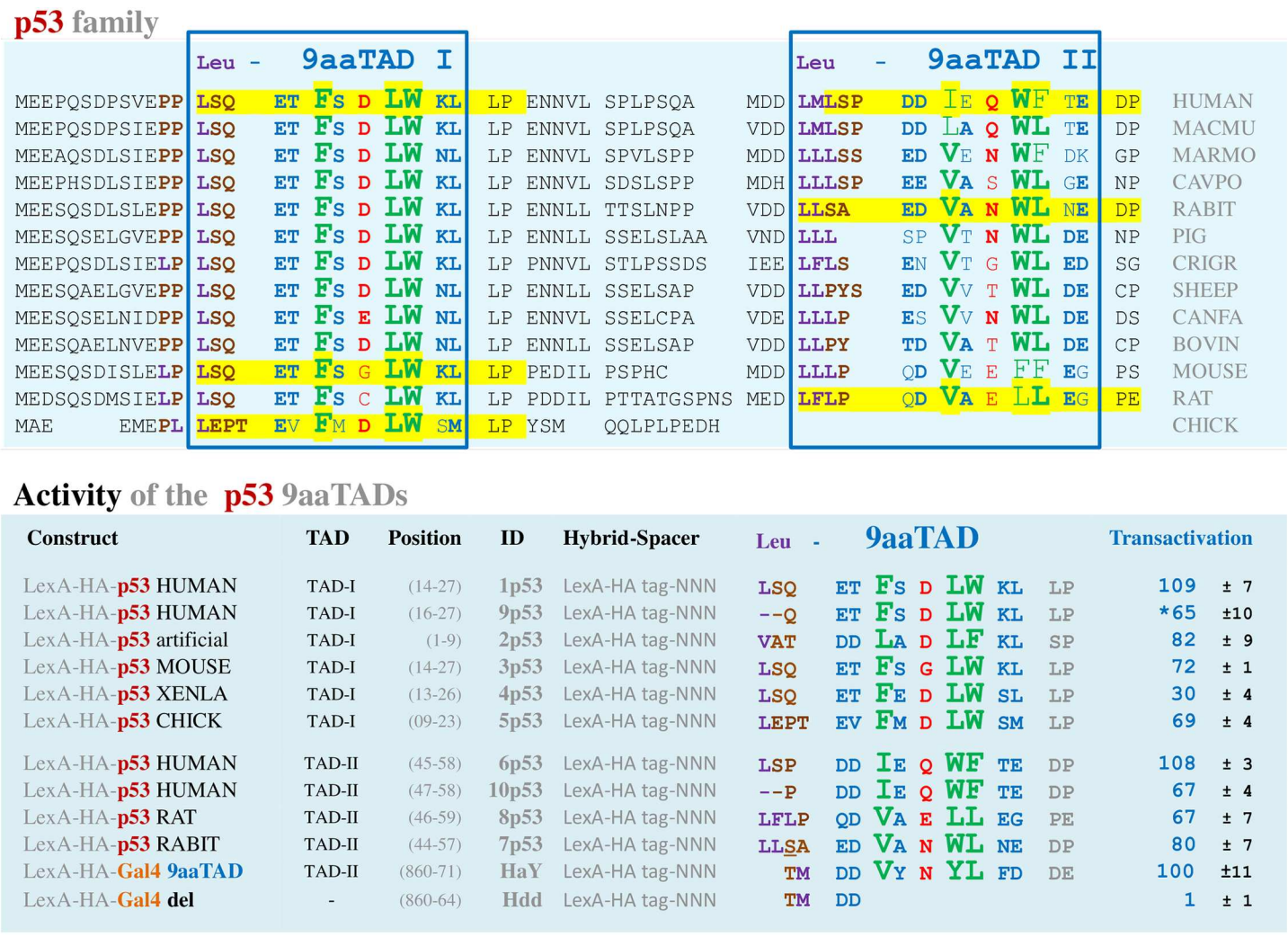
Results

The 9aaTADs in the p53 protein

For both predicted p53 9aaTAD domains, we aimed to prove their competence to activate transcription.

The human p53 9aaTAD constructs were generated with and without conserved 9aaTAD proximal regions including Leucines. We generated p53 9aaTAD constructs from several other species representing variability within the family (mouse, frog and chicken for TAD-I, and rat and rabbit for TAD-II; highlighted in Fig 1).

Importantly, we observed the distinctive similarities of the p53 9aaTAD-II domain with Gal4 (four out of nine amino acids are identical and seven out of nine are similar between rabbit p53 and Gal4 9aaTAD)(Fig 1). The fact that the Gal4 proteins could be found only in lower eukaryotes while the p53 proteins are found only in higher eukaryotes, we argue that their



Activity of the p53 9aaTADs

Construct	TAD	Position	ID	Hybrid-Spacer	Leu - 9aaTAD							Transactivation	
LexA-HA-**p53** HUMAN	TAD-I	(14-27)	1p53	LexA-HA tag-NNN	LSQ	ET	FS	D	LW	KL	LP	109	± 7
LexA-HA-**p53** HUMAN	TAD-I	(16-27)	9p53	LexA-HA tag-NNN	--Q	ET	FS	D	LW	KL	LP	*65	±10
LexA-HA-**p53** artificial	TAD-I	(1-9)	2p53	LexA-HA tag-NNN	VAT	DD	LA	D	LF	KL	SP	82	± 9
LexA-HA-**p53** MOUSE	TAD-I	(14-27)	3p53	LexA-HA tag-NNN	LSQ	ET	FS	G	LW	KL	LP	72	± 1
LexA-HA-**p53** XENLA	TAD-I	(13-26)	4p53	LexA-HA tag-NNN	LSQ	ET	FE	D	LW	SL	LP	30	± 4
LexA-HA-**p53** CHICK	TAD-I	(09-23)	5p53	LexA-HA tag-NNN	LEPT	EV	FM	D	LW	SM	LP	69	± 4
LexA-HA-**p53** HUMAN	TAD-II	(45-58)	6p53	LexA-HA tag-NNN	LSP	DD	IE	Q	WF	TE	DP	108	± 3
LexA-HA-**p53** HUMAN	TAD-II	(47-58)	10p53	LexA-HA tag-NNN	--P	DD	IE	Q	WF	TE	DP	67	± 4
LexA-HA-**p53** RAT	TAD-II	(46-59)	8p53	LexA-HA tag-NNN	LFLP	QD	VA	E	LL	EG	PE	67	± 7
LexA-HA-**p53** RABIT	TAD-II	(44-57)	7p53	LexA-HA tag-NNN	LLSA	ED	VA	N	WL	NE	DP	80	± 7
LexA-HA-**Gal4 9aaTAD**	TAD-II	(860-71)	HaY	LexA-HA tag-NNN	TM	DD	VY	N	YL	FD	DE	100	±11
LexA-HA-**Gal4 del**	-	(860-64)	Hdd	LexA-HA tag-NNN	TM	DD						1	± 1

Fig 1. Predicted 9aaTADs in p53 family. Alignment of predicted 9aaTADs in p53 family retrieved by our 9aaTAD prediction algorithms. Highlighted domains were tested for competence to activate transcription in this study.

doi:10.1371/journal.pone.0162842.g001

9aaTAD domains' similarity reflects rather functional convergence than conservation (Fig 2). As convergence could be seen in nature e.g. tenrec is genetically closer relative to elephant but rather distant to hedgehog, then tenrec and hedgehog are much more similar by appearance, size and way of functional "spiny" protection (Fig 2).

Next, we tested minimal p53 9aaTAD-I domain for interaction with the KIX domain. The LexA hybrid construct with HA-tag and minimal p53 9aaTAD-I domain was expressed in yeast L40 strain. The GST construct with human KIX domain was expressed separately in E. coli strain BL21. Under our experimental conditions, the purified GST-KIX protein was able to specifically pull down p53 9aaTAD-I from whole crude yeast extract (Fig 3).

Finally, we also generated an artificial derivate of the human p53 9aaTAD-I, in which we followed variation found in the p53 TAD-II domain and substituted conserved Tryptophan in TAD-I domain for Phenylalanine conserved in TAD-II domain.

All tested p53 9aaTAD domains have competence to activate transcription (Fig 1). The protein expression of all p53 constructs were confirmed by Western blotting (S1 Fig).

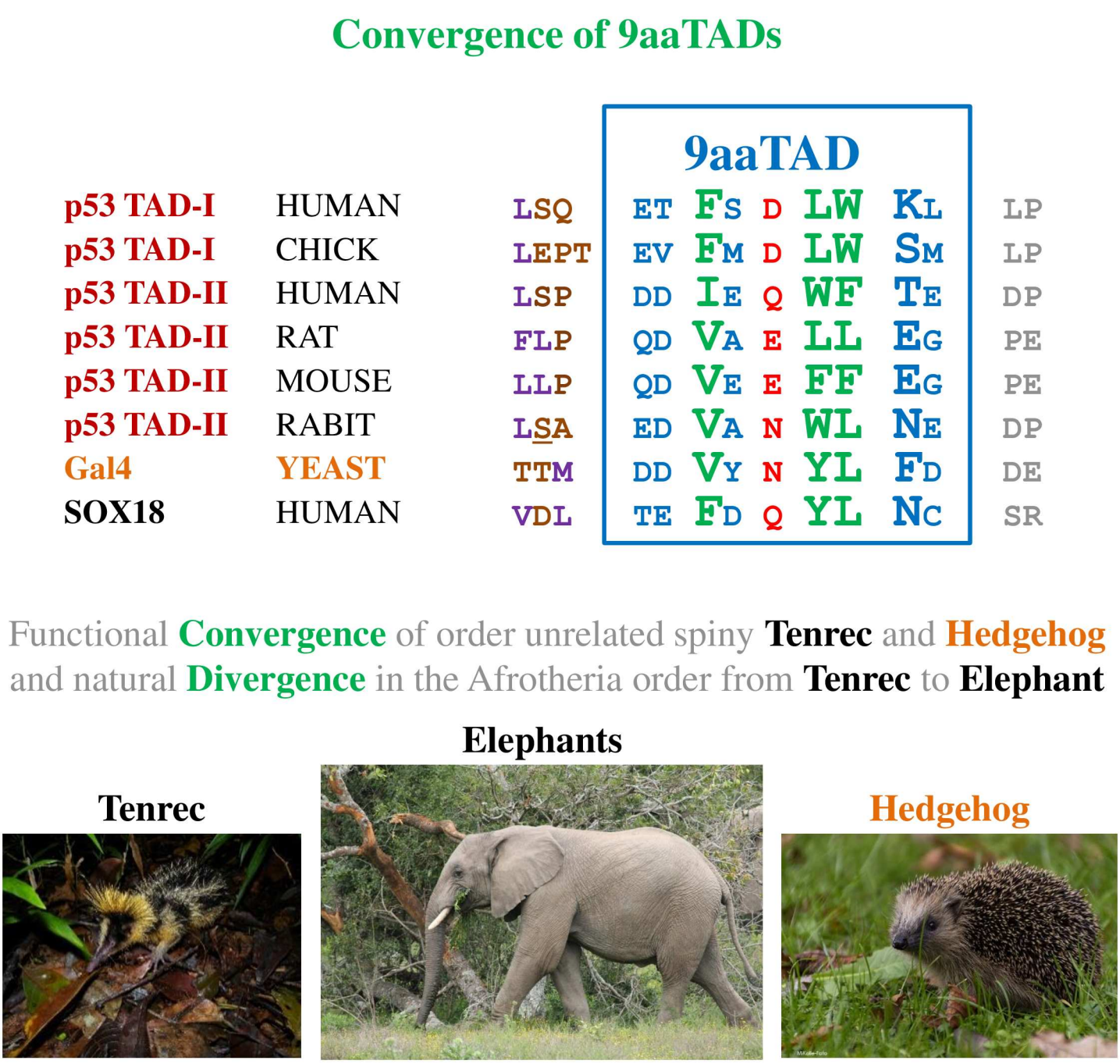


Fig 2. p53 9aaTADs activate transcription as small peptides. The predicted 9aaTADs in p53 from different species were tested for activation of transcription in LexA hybrid constructs. Similarity of p53 with Gal4 and Sox18 are highlighted. The construct 9p53, labelled with asterisk, has lower expression level compared with other constructs (S1 Fig). Animal picture from Flickr: Lowland Streaked Tenrec, Mantadia, Madagascar, Author: Frank Vassen; Elephant, Author: Jon Mountjoy; Igel (Hedgehog), Author: Mi chaela. All pictures have Creative Commons Attribution 2.0 Generic license.

doi:10.1371/journal.pone.0162842.g002

The 9aaTAD domain and MED15

The transcription factors including Gal4, Oaf1 and Pdr1, members of the 9aaTAD family, are known to interact with the MED15 transcriptional mediator. Therefore we focused our attention on other MED15 interacting proteins such as Msn2 and Msn4 transcription factors [21,

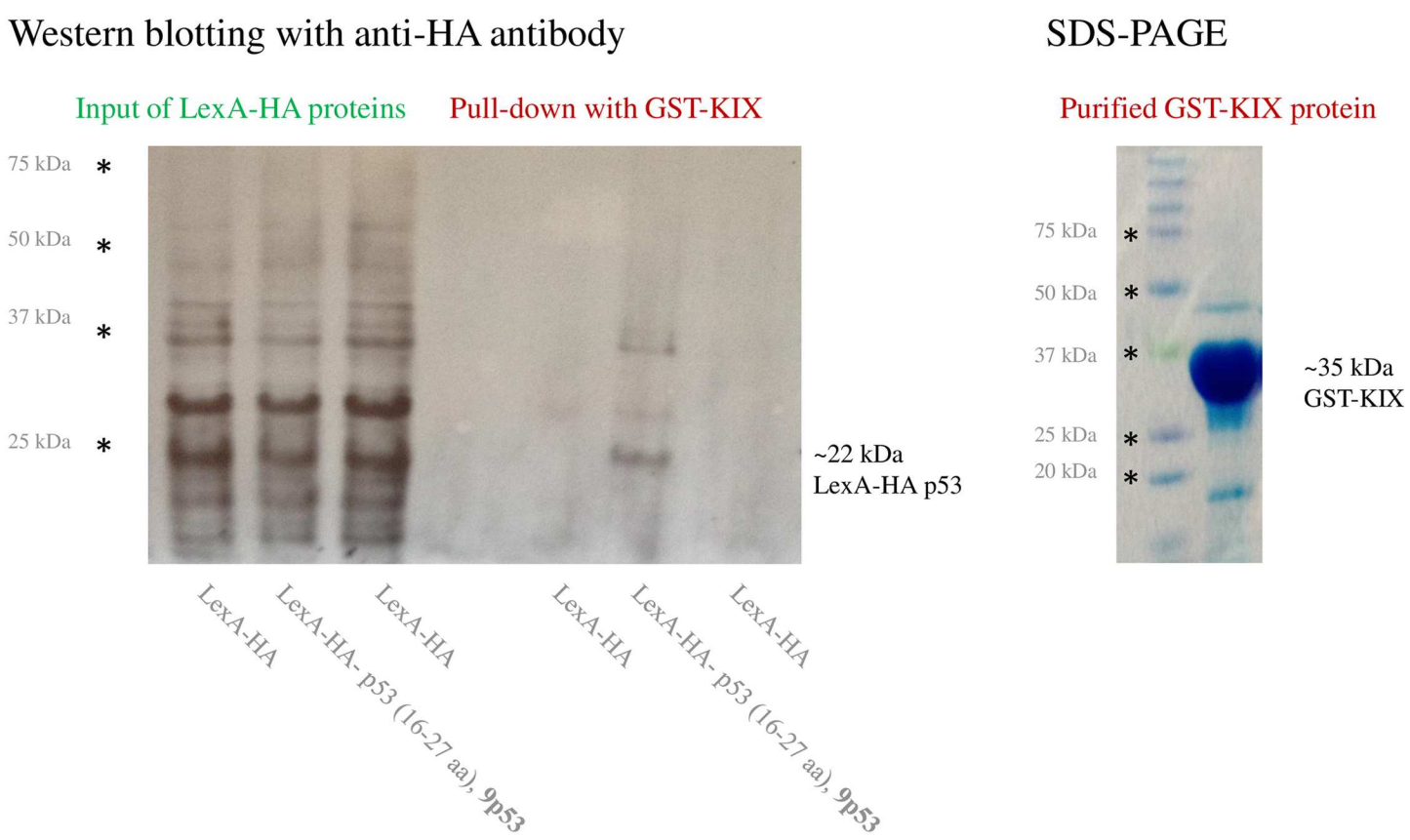


Fig 3. p53 9aaTAD-I is sufficient for interaction with the KIX domain. Interaction of the KIX domain with the p53 9aaTAD-I without its proximal Leucine in pull down experiment was monitored by western blotting. The purity of GST-KIX protein was monitored by SDS-PAGE.

doi:10.1371/journal.pone.0162842.g003

22]. The authors successfully used computational prediction for unstructured regions of the Msn2 protein to localize the transactivation domain.

Initially, we used the ExPASy SIB BLAST to identify the Msn2 and Msn4 orthologs. By using our online 9aaTAD prediction, we have identified the 9aaTAD domains in Msn2, Msn4 and their orthologs. First, we showed that the predicted 9aaTADs were located within the reported transactivation domains (Fig 4). Second, we observed conservation in the 9aaTAD domain in both Msn2 and Gal4 families (Fig 4). Next, we chose representative members of the Msn2 family and tested their predicted 9aaTAD domains with and without 9aaTAD proximal regions for ability to activate transcription.

The resulted activity of all tested Msn2 and Msn4 9aaTAD domains with proximal regions showed up to 30% activity of the Gal4 9aaTAD in LexA hybrid assay, which proved their competence to efficiently activate transcription. The 9aaTAD proximal regions are essential in *S.c.* Msn2 and *S.c.*Msn4 but not in *K.a.*Msn2/4 (Fig 5). Therefore, we assigned both Msn2 and Msn4 transcription factors as proved members of the 9aaTAD family.

Our previously reported 9aaTAD domains for Gal4, Oaf1, Pdr1, Pdr3, Pip2 and other Gal4 paralogs (description and 9aaTAD online annotations, 2006) [4,7] had encouraged further studies with Oaf1 and Pdr1 9aaTAD peptides [23,24]. A novel xenobiotic response domain was reported for Oaf1 and Pdr1 activators and each interacted with the KIX domain of Med15 [23,24]. However, the origin of the Pdr1 peptide (about twelve amino acids) used in the study

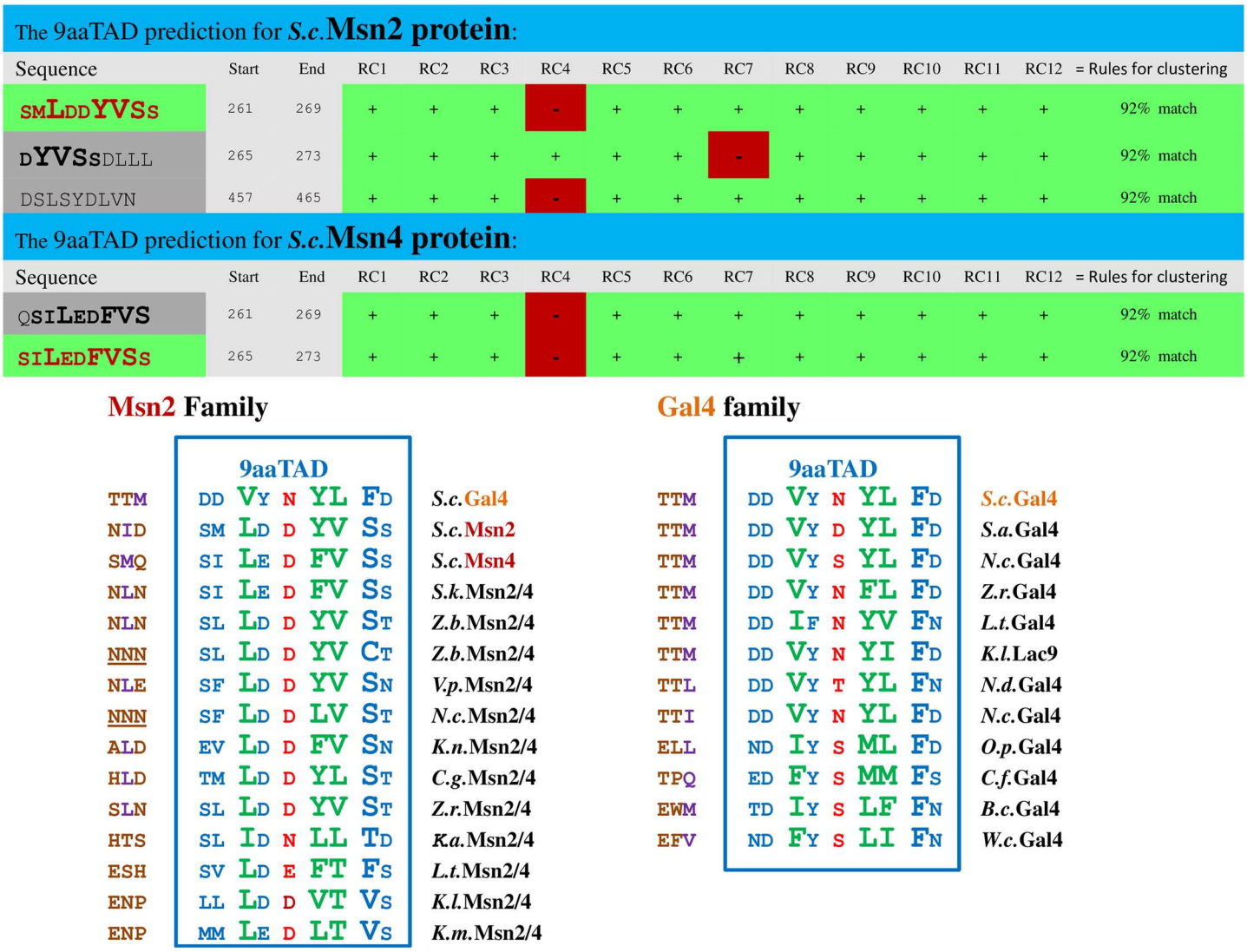


Fig 4. Predicted 9aaTADs in Msn2 family. Prediction result for 9aaTADs in Msn2 and Msn4 activators retrieved by our 9aaTAD prediction algorithms. The conservation and variability of the 9aaTADs in the Msn2 and Gal4 families are shown.

doi:10.1371/journal.pone.0162842.g004

was unknown (the peptide has not been referred to any screen or prediction e.g. for xenobiotic response pattern).

We used the ExPASy SIB BLAST to identify the Pdr1 and Pdr3 orthologs and confirmed the 9aaTAD conservation in the family (Fig 6). Because of the position of predicted Pdr1 9aaTAD domain and the Pdr1 xenobiotic domain were matched, we aimed to prove the corresponding Pdr1 peptide for competence to activate transcription. We have tested the Pdr1 xenobiotic domain (12 aa) including predicted 9aaTAD in LexA hybrid assay for activation of transcription. The resulted activity was comparable with the Gal4 and Oaf1 9aaTADs (Fig 7). Therefore, the Pdr1 xenobiotic domain is a functional transactivation domain conserved in Pdr1 family and is identical to reported Pdr1 9aaTAD domain. Therefore, we also assigned Pdr1 as proved member of the 9aaTAD family.

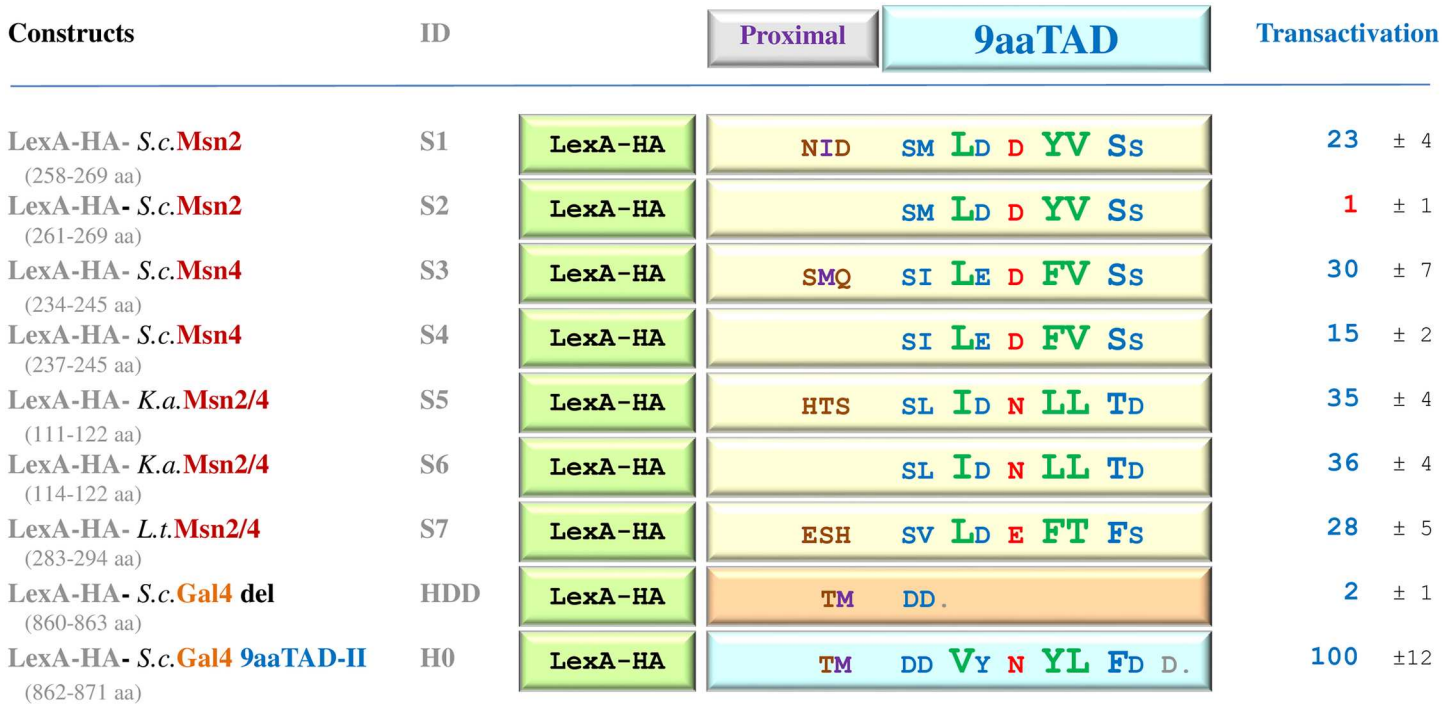


Fig 5. Activity of Msn2 and Msn4 9aaTADs. Msn2 and Msn4 LexA hybrid constructs assayed in L40 strain for transactivation activity.

doi:10.1371/journal.pone.0162842.g005

Spontaneously generated 9aaTADs in Gal4

In recent study (Piskacek et al., 2016), we reported artificial 9aaTAD domains identified by online 9aaTAD prediction (www.piskacek.org) in Gal4 TAD replicas G80BP-A and G80BP-B originally shown in [25]. Similarly, we identified artificial 9aaTAD domain in a strong activator KBP2.20 and in p53 mimetic ECapLL [16][26–29]. Accordingly, we generated an artificial Gcn4 mimetic S11, artificial 9aaTAD domain swapping of the Gal4 residues in to the Gcn4 9aaTAD domain shown in (S4 Fig).

We predicted a half site of the 9aaTAD domain in Gal4 region (92–100 aa), a part of DNA binding domain, DBD, which is not involved in transactivation. This Gal4 region can be fused with other peptides to form strong artificial 9aaTAD domains. We demonstrated the ability of the Gal4DBD to form strong artificial 9aaTAD domain by fusing it with the second half site of the Gal4 9aaTAD domain (construct U39, Fig 8). The fusion construct of Gal4DBD and a half site of the Gal4 9aaTAD domain activated transcription much powerfully than the natural Gal4 9aaTAD domain.

The best examples of accidentally generated artificial 9aaTAD domains involving Gal4DBD domain are the constructs pRJR200 and pRJR213 generated in [30]. In these constructs, the Gal4DBD domains were fused with peptides originated from Gal4 region (840–857 aa), which resulted in formation of strong artificial 9aaTAD called by authors the Gal4 acidic domain. The Gal4DBD fusion constructs accidentally generated artificial 9aaTADs with strong activities (four and six amino acid long peptides of the artificial Gal4 acidic domain) (Fig 8). For a complete set of Gal4 constructs and their natural activity see recent Piskacek et al., 2006.

Similarly, as indicated by alanine scanning for essential amino acids, another fusion protein Gal4DBD-P201 [31] formed artificial 9aaTAD domain shown (S2 Fig). Furthermore, the

The 9aaTAD prediction for *S.c.* **Pdr1 protein**:

Sequence	Start	End	RC1	RC2	RC3	RC4	RC5	RC6	RC7	RC8	RC9	RC10	RC11	RC12	= Rules for clustering
EDLYSILWS	1054	1062	+	+	+	+	+	+	+	+	+	+	+	+	perfect match

Pdr1 family

	9aaTAD	
TTM	DD VY N YL FD	*S.c.*Gal4
GDL	ED LY S IL WS	S.c.Pdr1
GDL	TD LY H TL WN	S.c.Pdr3
TDM	ND LY N VL WS	*N.c.*Pdr1/3
TDL	SD LY N VL WS	*N.d.*Pdr1/3
ADI	DD LY N VL WG	*N.d.*Pdr1/3
GDL	NE LY N SL WG	*C.g.*Pdr1/3
SDM	TD LC N IL WS	*K.a.*Pdr1/3
GDL	ND LC N TL WS	*T.d.*Pdr1/3
GDL	TE LY H TL WN	*S.k.*Pdr1/3
GDL	ND LY S TL WS	*Z.b.*Pdr1/3
TDM	ND LF K TL WN	*K.a.*Pdr1/3

Oaf1 family

	9aaTAD	
GGL	DL FD Y DF LF	S.c.Oaf1
GSL	DF FD Y DL LF	S.c.Pip2
GGL	DL FD Y EF LF	K.a.Oaf1
DNL	DF LD Y DL FF	K.a.Pip2
GNL	DI FN Y DF FF	K.n.Oaf1
GRL	DL FN Y DF LF	T.b.Oaf1
GGL	DL FD Y DF LF	T.p.Oaf1
ATL	GV LD F EF LL	C.g.Oaf1
GGL	DI FD F LF GN	Z.b.Oaf1
GGL	NL FD Y DF LF	Z.r.Oaf1
SSL	GV SE F DL LF	N.c.Oaf1
QSL	GI SD F DG FL	N.d.Oaf1
DDL	DL LD Y EF LF	A.a.Oaf1
DIL	DF FD T FS MN	K.m.Oaf1
HGV	SV LN D YF KQ	K.l.Oaf1
LAG	VS AD F DI VF	K.m.Oaf2
LAG	LS AN F DF VF	K.l.Oaf2
FLG	DI FD R LM GT	W.c.Oaf1
LNG	DF FN S AI IE	W.c.Oaf2

Fig 6. Predicted 9aaTADs in Gal4 family. Prediction results for 9aaTADs in Pdr1 and Pdr3 activators revealed by our 9aaTAD prediction algorithms. The conservation and variability of the 9aaTADs in the Pdr1 and Oaf1 families are shown.

doi:10.1371/journal.pone.0162842.g006

Gal4DBD-P201 construct strongly resembles Gal4DBD+1/2 Gal4TAD construct that we described above (Fig 8).

The diverse synthetic peptide libraries revealed strong artificially transactivation domains; "Activating regions: as many as you like" [32]. Beside the proline and tryptophan repeats, we could identify artificial 9aaTAD domains in the reported activators, e.g. the most potent activator reported called B42 [33]. We tested the B42 9aaTAD domain in LexA hybrid assay for activation of transcription. The resulted activity was comparable with the Gal4 9aaTAD domain (S3 Fig). We obtained similar prediction result (predicted 9aaTAD domain: DTLYLDWLED) for other potent activator B114 that has been reported later (in second series of potent activators) [34].

From the results above, it apparent that spontaneous generation of numerous artificial transactivation domains could be generated by intension [25–29] or by accident [30,32–35]. The artificial activators G80BP-A and G80BP-B (Gal4 9aaTAD mimetics)[25], KBP2.20 (KIX binding peptide, random peptide from screen)[26], ECapLL (p53 derivate)[27–29], S11 (Gcn4 / Gal4 derivate, 9aaTAD domain swapping derivate) (S4 Fig), pRJR200 and pRJR213 (Gal4 acidic domain constructs, artificial 9aaTAD generated accidentally) [30], U39 (Gal4 derivate,

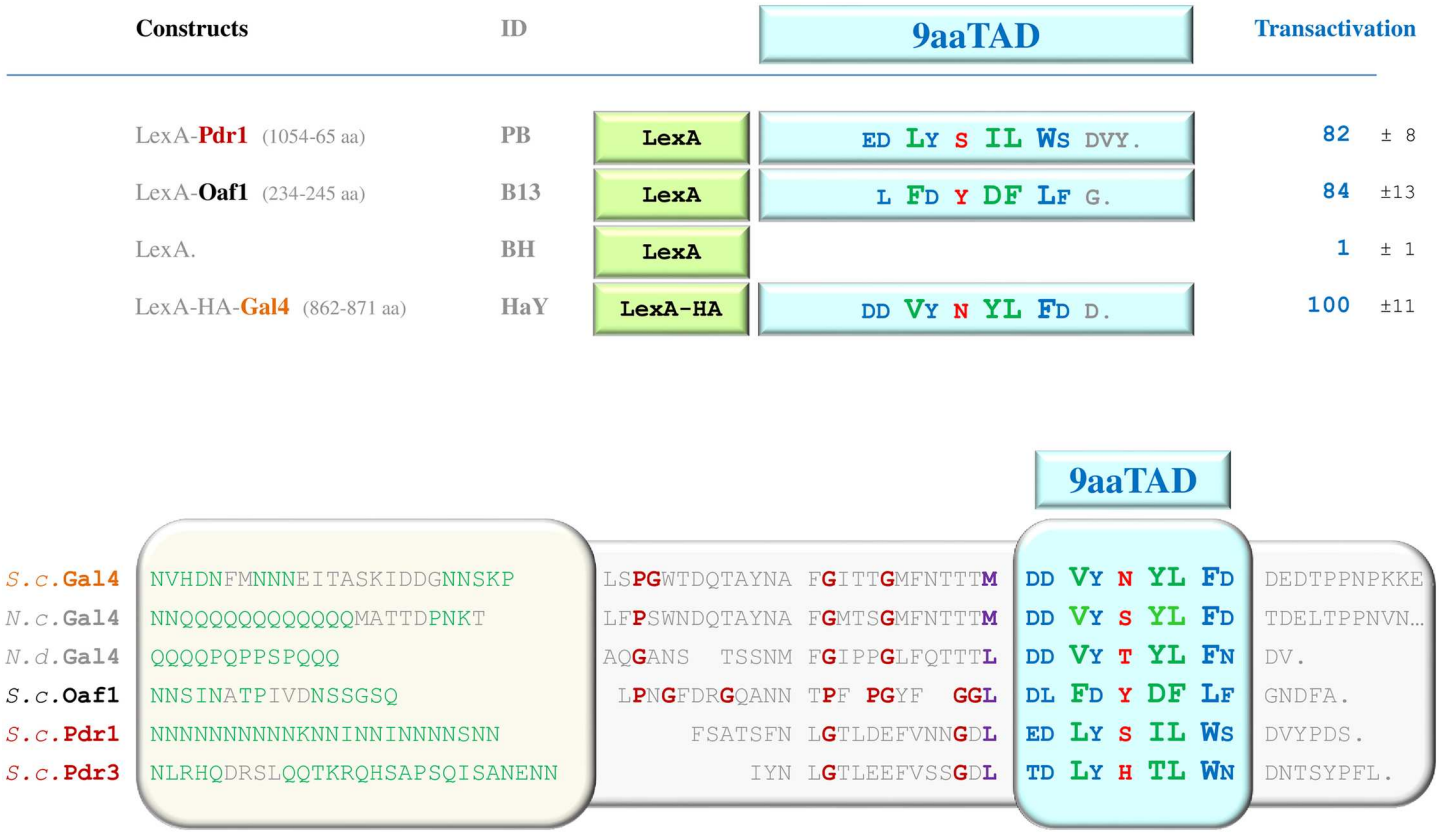


Fig 7. Activity of Pdr1 9aaTAD. Pdr1 LexA hybrid construct assayed in L40 strain for transactivation activity. The similarity with other Gal4 orthologs with adjacent sequences is shown.

doi:10.1371/journal.pone.0162842.g007

artificial 9aaTAD made by fusion, analogy to pRJR200 and pRJR213, generated intentionally) (Fig 8) and B42 (artificial 9aaTAD, random peptide from screening) (S3 Fig) activators were assigned to 9aaTAD family.

Discussion

The Nine amino acid Transactivation Domain, 9aaTAD, is a large family of the transcription activators universally recognized by transcriptional machinery from yeast to man. The 9aaTAD domain is characterized by the disengaged pattern, by amino acid composition and by tandem of hydrophobic clusters (Fig 9). The 9aaTAD domain is well balanced by hydrophilic amino acids, which are usually in proportion of positively and negatively charged. From the structural data for the E2A and MLL in complex with the KIX domain, we observed helix formation for some 9aaTADs, whose length vary from 9 to 12 aa [16]. The online 9aaTAD prediction is available on www.piskacek.org.

Some but not all of the 9aaTAD domains interact with multiple mediators and that with different binding affinity e.g. p53, MLL and E2A (MLL-site and Myb-site of KIX domain, sites of TAZ1 and TAZ2, and IBiD) [23,36–40]. These interactions share some similarity, but also show obvious differences and individuality, e.g. Oaf1 and Pdr1 [23,24]. The 9aaTAD domains may use multiple binding positions and orientations, e.g. p53 and Gcn4 [19,41].

Albeit the 9aaTAD domains have enormous variability and amorphous character, they are universally recognized by transcriptional machinery throughout eukaryotes [1]. The 9aaTAD

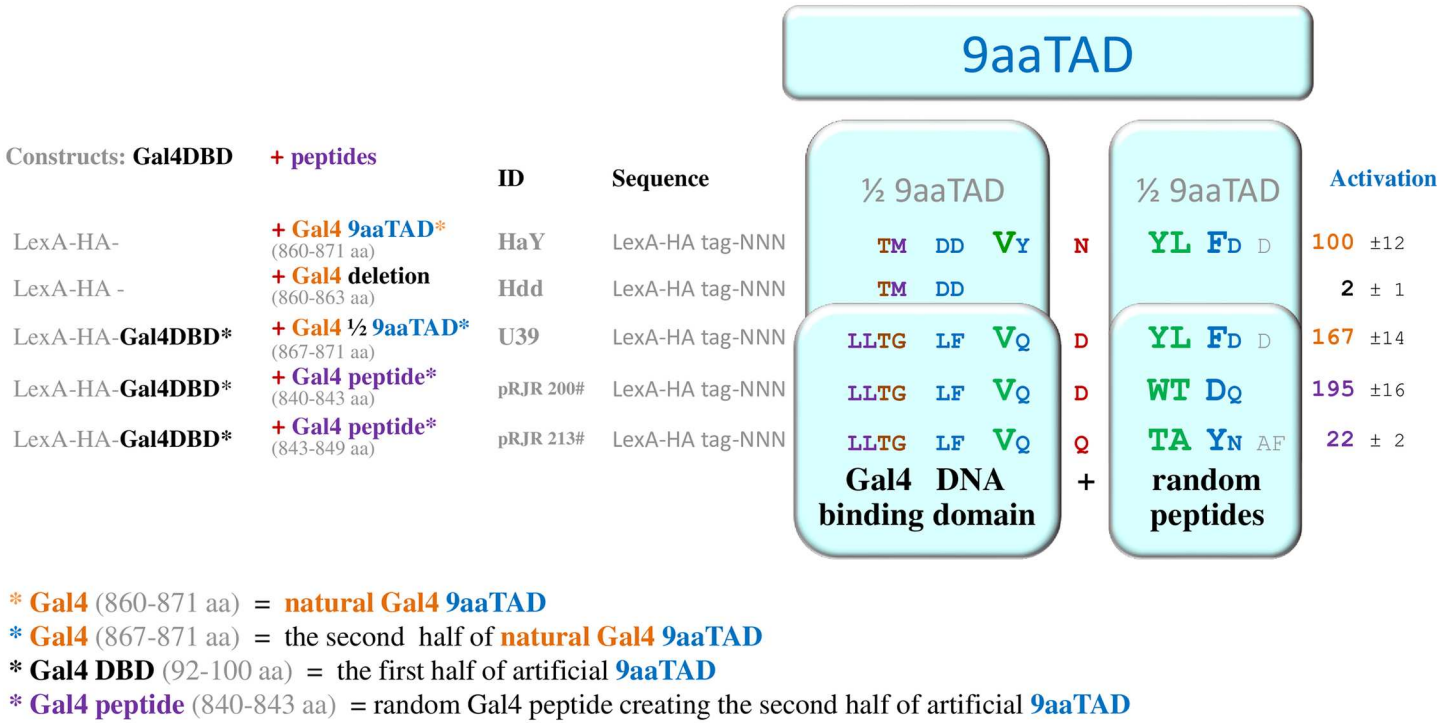


Fig 8. C-terminus of the Gal4DBD domain (92–100 aa) works as a half side of the 9aaTAD domain. A serious concern was found for not real activation function of the Gal4 acidic domain. The artificial 9aaTADs in pRJR200 and pRJR213 constructs were generated accidentally by others and represent so called Gal4 acidic domain. Here we restricted essential part of the Gal4 acidic domain to the recognized functional 9aaTAD region. In this constructs, the functionally unrelated peptides from Gal4 region (840–857 aa) subsidised unintentionally for the second half site of 9aaTAD domain. Artificial 9aaTADs was generated by fusion of the Gal4DBD domain and a half of the 9aaTAD of Gal4. The part of the Gal4DBD domain (92–100 aa) represent first half site of the 9aaTAD domain and was use in constructs to demonstrate capability to generate artificial 9aaTADs by fusion with the second half site of the Gal4 9aaTAD domain.

doi:10.1371/journal.pone.0162842.g008

transcription factors interact with multiple transcriptional mediators [18,19,23,24,26,29,42–54] [55–60]. The conservation of some transcriptional mediators (TAF9 and KIX domain in MED15) might be responsible for the 9aaTAD domain overall occurrence and functionality.

In this study, we showed that both p53 9aaTAD domains activate transcription as small peptides. Nevertheless, we showed that the p53 9aaTAD-I and 9aaTAD-II domains do not have one amino acid identical. We also demonstrated that their shared 9aaTAD motif enables residue swapping. In respect of the shared 9aaTAD motif in p53 protein, the modification of the first p53 9aaTAD-I domain towards the second 9aaTAD motif (in the p53 9aaTAD-II domain) did not result in a lost of the transactivation potential (artificial construct 2p53).

Similar result has been reported for p53 mimetic ECapLL [28]. We recognized analogical modification of the second p53 9aaTAD-II domain towards the first 9aaTAD motif (in the p53 9aaTAD-I domain) [16]. Noteworthy, both p53 9aaTAD-I and 9aaTAD-II domains bind to the same transcriptional mediator subunits of the CBP/p300, what further underpins the shared 9aaTAD motif (MLL-site and Myb-site of KIX domain, sites of TAZ1 and TAZ2, and IBiD)[19,36].

Our results have shown the 9aaTAD convergence in the rabbit p53 and Gal4 proteins, which have 44% overall identity and 78% similarity. Both 9aaTAD domains resembled another 9aaTAD domain, Sox18 [61]. Previously, we observed another 9aaTAD convergence in two unrelated proteins, E2A and MLL (SDLL-D-FS and SDIM-D-FV). Their 9aaTAD domains occupied identical binding site on the KIX domain of the CBP [16]. However, we found only

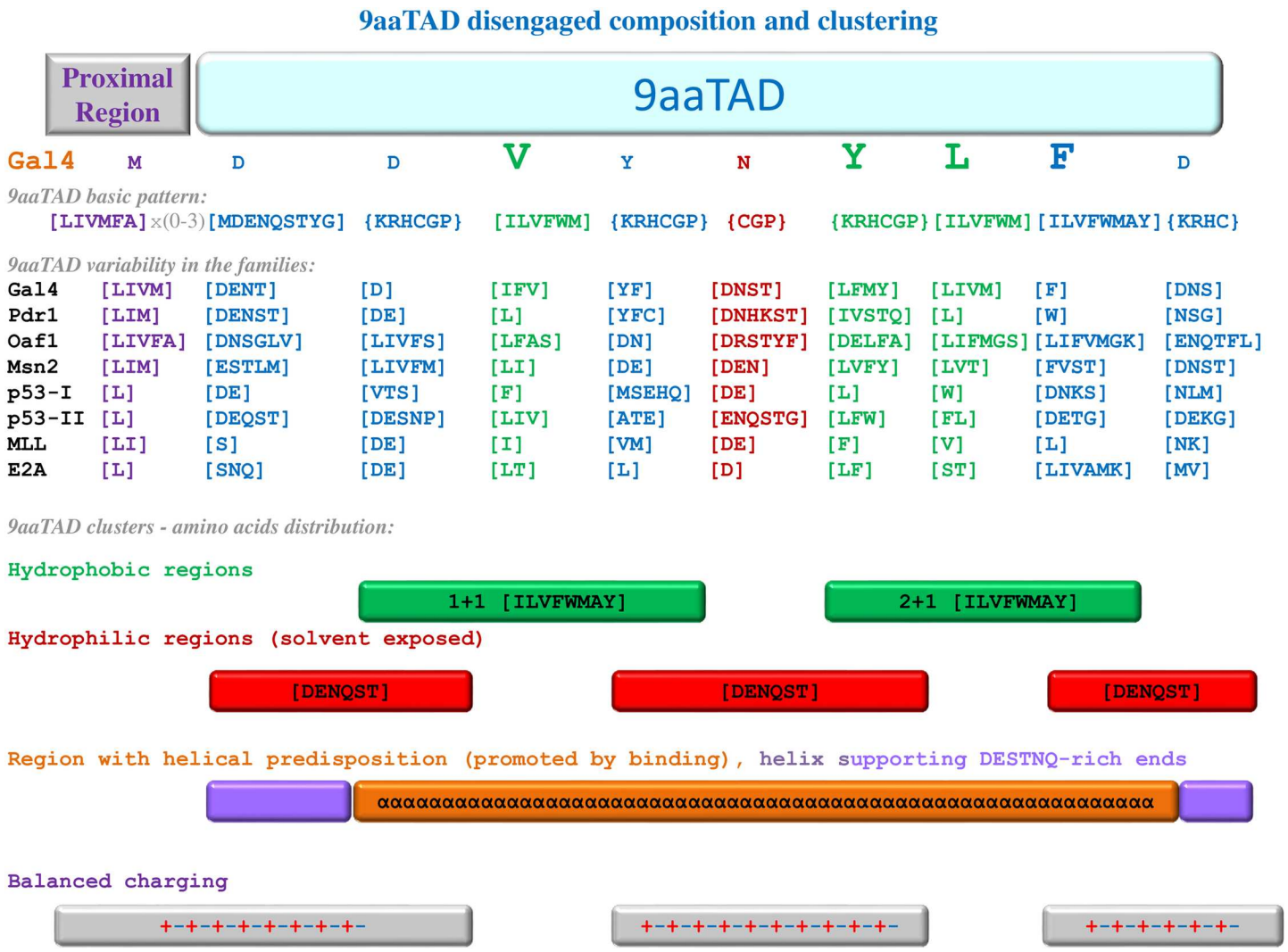


Fig 9. The 9aaTAD domain disengaged composition and clustering.

doi:10.1371/journal.pone.0162842.g009

negligible similarity between rat and human p53 TAD-II, which represented natural 9aaTAD divergence.

The 9aaTAD proximal Leucines are well conserved in the p53 family, but less in other members of the 9aaTAD family, e.g. Msn2 activators. The Leucines are not conserved also in other 9aaTAD proximal regions demonstrating overall natural variability (Isoleucine, Valine, Phenylalanine, Methionine, Threonine and Serine were found in the 9aaTAD proximal regions of MLL, Oaf1, Pdr1, Gcn4 and Gal4). From our results for the Msn2 and p53 activators, we demonstrated that the contribution of the 9aaTAD proximal regions is individually variable and not always essential for transactivation function. The 9aaTAD proximal regions might be part of the 9aaTAD helix e.g. E2A and MLL [16].

The members of the 9aaTAD family share small size, highly variable pattern and hydrophobic and hydrophilic clusters. The 9aaTAD domains could be split in two subdomains and merged with other 9aaTAD subdomains without losing activity. The 9aaTAD domains could be largely and almost freely modified, and new 9aaTAD domain could be easily generated by accident or by intention.

The 9aaTAD family is not exclusive transactivation domain in eukaryotes. There are numerous known transactivation domains unrelated to the 9aaTAD domain with different mode of binding to transcriptional mediators e.g. STAT2 [62] or different amino acid composition e.g. SP1 [63–65].

During the evolution, the fast generation of numerous specific transcription factors has been crucial for tailored regulation of individual genes. The observed easiness of spontaneous generation of artificial 9aaTAD transactivation domains in the labs, "Activating regions: as many as you like" [32], indicates that the 9aaTAD domain represented evolutional advantage for generation of transcription factors from DNA binding precursors.

Supporting Information

S1 Fig. Protein expression. The protein level produced from the constructs 1-10p53, HaA and HaY in L40 strain were monitored by Westernblotting. The proteins comprise LexA a HA tags with a total size of about 21 kDa. The degradation product comprising almost LexA protein (LexA torso) has a total size of about 20kDa.
(TIF)

S2 Fig. Artificial activator P201. We fused the part of the Gal4 DNA binding domain (92–100 aa), DBD, with the second half site of the Gal4 9aaTAD. The Gal4 DNA binding domain region (92–100 aa) substitute for the first half of the 9aaTAD in this and other artificial constructs. The amino acids in fusion region of both Gal4 DNA binding domain and the random peptide are essential for transactivation function. Notice: Gal4 region (1–84 aa) is sufficient for DNA binding. Blue asterisks referred to the results of this study (Fig 8), constructs HaY and U39, black asterisks to the results reported by Lu X et al. 2000 and red asterisks by Lu Z et al. 2005.
(TIF)

S3 Fig. Artificial activator B42. The full sequence of B42 peptide (1–79 aa) and the identified 9aaTAD within are shown. The B42 9aaTAD LexA hybrid construct was assayed in L40 strain for the transactivation activity.
(TIF)

S4 Fig. Putative 9aaTAD motifs in Gcn4. We identified the putative 9aaTADs in both reported Gcn4 transactivation domains (description and 9aaTAD online annotations, 2006), which amino acid variations are very close to mouse and bovine p53 9aaTADs (5 identical and 3 similar amino acids: **K/Q**, **D/E**, **V**, **E**, **S/T**, **F**, **F**, **D**, N/E). Noteworthy, the Gcn4 protein has an unusual Lys in the position 1 of the 9aaTAD-II, which is out of predictive recognition (sequence: KEWTSLFDN). The unusual amino acids in the 9aaTAD domains were found also in other members of the 9aaTAD family e.g. Cysteine and Glycine in rat and mouse p53 9aaTAD-I. We assigned many transactivation domains to the 9aaTAD family, which fit with size, share deliberated 9aaTAD pattern and the clusters of the hydrophobic/hydrophilic amino acids. The amorphous nature of the 9aaTAD domains does not offer any invariant or conserved residues, which let us to generate the absolute reliable pattern for all of them. Therefore i) our prediction is still uncertain, ii) generate many false positives, iii) pattern does not fit for all 9aaTAD variations of the orthologs, and iv) putative 9aaTADs need always to be experimental verified. Nevertheless, there are many examples, where the 9aaTAD prediction works well, e.g. MLL or p53 activators. MLL (Q03164) is 3969 amino acids long protein with only two predicted 9aaTADs, where one of them is confirmed transactivation domain. Over two hundred Gcn4 9aaTAD-I modifications were generated and their competence to activate transcription were assayed by Warfied et al., 2014. Despite of the authors' enormous effort to define the transactivation domain by this approach, they found merely Tryptophan-rich transactivation

domains deprived of acidic residues (AVWWSLFAS, AWWWWAFWS, AFWMWLFAT). We tested the Tryptophan-rich activation domain m120 (AFWMWLFAT) derived from Gcn4 9aaTAD in the standard LexA hybrid assay. The Gcn4 mutant m120 has no activity (>1% ± 1 of the referent Gal4 construct HaY), what indicated serious data inconsistency in the report by Warfied et al., 2014. Therefore we proceeded differently to characterise the Gcn4 TAD. Because of the Gcn4 TAD-I domain fulfils the deliberate criteria for 9aaTAD motif (positive online 9aaTAD prediction, formation of two hydrophobic patches interspersed by hydrophilic residues), we made subdomain swapping between putative Gcn4 9aaTAD and Gal4 9aaTAD and generated a hybrid construct S11 to prove predicted shared motif. The construct S11 has comparable transcriptional activity to the Gal4 9aaTAD in LexA hybrid assay, what proofs the concept for the 9aaTAD motif in Gcn4 TAD-I by the swapping experiment (Gcn4 9aaTAD core spreading from position 3p to 7p was swapped with Gal4 9aaTAD). Noteworthy, the hydrophobic patches in the core of Gcn4 9aaTAD-I (**V_SFF**) and Gcn4 9aaTAD-II (w_**SLF**) have high similarity with the 9aaTAD domain of the *B.a*.Gal4 (I_**SLF**).
(TIF)

Acknowledgments

We thank especially to Alan G. Hinnebusch and Robert Tjian for support of the 9aaTAD project, to Isabelle Lemasson for expression constructs essential for this study.

Author Contributions

Conceptualization: MP.

Investigation: MP MH MR.

Software: MP.

Writing – original draft: MP AK.

References

1. Kakidani H, Ptashne M. GAL4 activates gene expression in mammalian cells. Cell. 1988; 52: 161–167. PMID: 2830021
2. Fields S, Jang SK. Presence of a potent transcription activating sequence in the p53 protein. Science. 1990; 249: 1046–1049. PMID: 2144363
3. Baumgartner U, Hamilton B, Piskacek M, Ruis H, Rottensteiner H. Functional analysis of the Zn(2)Cys (6) transcription factors Oaf1p and Pip2p. Different roles in fatty acid induction of beta-oxidation in Saccharomyces cerevisiae. J Biol Chem. 1999; 274: 22208–22216. PMID: 10428786
4. Piskacek S, Gregor M, Nemethova M, Grabner M, Kovarik P, Piskacek M. Nine-amino-acid transactivation domain: establishment and prediction utilities. Genomics. 2007; 89: 756–768. doi: 10.1016/j.ygeno.2007.02.003 PMID: 17467953
5. Sandholzer J, Hoeth M, Piskacek M, Mayer H, de Martin R. A novel 9-amino-acid transactivation domain in the C-terminal part of Sox18. Biochem Biophys Res Commun. 2007; 360: 370–374. doi: 10.1016/j.bbrc.2007.06.095 PMID: 17603017
6. Lindert U, Cramer M, Meuli M, Georgiev O, Schaffner W. Metal-responsive transcription factor 1 (MTF-1) activity is regulated by a nonconventional nuclear localization signal and a metal-responsive transactivation domain. Mol Cell Biol. 2009; 29: 6283–6293. doi: 10.1128/MCB.00847-09 PMID: 19797083
7. Piskacek M. 9aaTAD Prediction result (2006). Nature Precedings. 2009; doi: 10.1038/npre.2009.3984.1
8. Piskacek M. Common Transactivation Motif 9aaTAD recruits multiple general co-activators TAF9, MED15, CBP and p300. Nature Precedings. 2009; doi: 10.1038/npre.2009.3488.2
9. Piskacek M. 9aaTADs mimic DNA to interact with a pseudo-DNA Binding Domain KIX of Med15 (Molecular Chameleons). Nature Precedings. 2009; doi: 10.1038/npre.2009.3939.1

10. Hong JY, Chae MJ, Lee IS, Lee YN, Nam MH, Kim DY, et al. Phosphorylation-mediated regulation of a rice ABA responsive element binding factor. Phytochemistry. 2011; 72: 27–36. doi: 10.1016/j.phytochem.2010.10.005 PMID: 21055780

11. Shekhawat UKS, Ganapathi TR, Srinivas L. Cloning and characterization of a novel stress-responsive WRKY transcription factor gene (MusaWRKY71) from Musa spp. cv. Karibale Monthan (ABB group) using transformed banana cells. Mol Biol Rep. 2011; 38: 4023–4035. doi: 10.1007/s11033-010-0521-4 PMID: 21110110

12. Lou S, Luo Y, Cheng F, Huang Q, Shen W, Kleiboeker S, et al. Human parvovirus B19 DNA replication induces a DNA damage response that is dispensable for cell cycle arrest at phase G2/M. J Virol. 2012; 86: 10748–10758. doi: 10.1128/JVI.01007-12 PMID: 22837195

13. Matsushita A, Inoue H, Goto S, Nakayama A, Sugano S, Hayashi N, et al. The nuclear ubiquitin proteasome degradation affects WRKY45 function in the rice defense program. Plant J. 2012; doi: 10.1111/tpj.12035

14. Aguilar X, Blomberg J, Brännström K, Olofsson A, Schleucher J, Björklund S. Interaction studies of the human and Arabidopsis thaliana Med25-ACID proteins with the herpes simplex virus VP16- and plant-specific Dreb2a transcription factors. PLoS ONE. 2014; 9: e98575. doi: 10.1371/journal.pone.0098575 PMID: 24874105

15. Scharenberg MA, Pippenger BE, Sack R, Zingg D, Ferralli J, Schenk S, et al. TGF-β-induced differentiation into myofibroblasts involves specific regulation of two MKL1 isoforms. J Cell Sci. 2014; 127: 1079–1091. doi: 10.1242/jcs.142075 PMID: 24424023

16. Piskacek M, Vasku A, Hajek R, Knight A. Shared structural features of the 9aaTAD family in complex with CBP. Mol Biosyst. 2015; 11: 844–851. doi: 10.1039/c4mb00672k PMID: 25564305

17. Qiu Y, Li M, Pasoreck EK, Long L, Shi Y, Galvão RM, et al. HEMERA Couples the Proteolysis and Transcriptional Activity of PHYTOCHROME INTERACTING FACTORs in Arabidopsis Photomorphogenesis. Plant Cell. 2015; doi: 10.1105/tpc.114.136093

18. Goto NK, Zor T, Martinez-Yamout M, Dyson HJ, Wright PE. Cooperativity in transcription factor binding to the coactivator CREB-binding protein (CBP). The mixed lineage leukemia protein (MLL) activation domain binds to an allosteric site on the KIX domain. J Biol Chem. 2002; 277: 43168–43174. doi: 10.1074/jbc.M207660200 PMID: 12205094

19. Lee CW, Arai M, Martinez-Yamout MA, Dyson HJ, Wright PE. Mapping the interactions of the p53 transactivation domain with the KIX domain of CBP. Biochemistry. 2009; 48: 2115–2124. doi: 10.1021/bi802055v PMID: 19220000

20. Cook PR, Polakowski N, Lemasson I. HTLV-1 HBZ protein deregulates interactions between cellular factors and the KIX domain of p300/CBP. J Mol Biol. 2011; 409: 384–398. doi: 10.1016/j.jmb.2011.04.003 PMID: 21497608

21. Lallet S, Garreau H, Garmendia-Torres C, Szestakowska D, Boy-Marcotte E, Quevillon-Chéruel S, et al. Role of Gal11, a component of the RNA polymerase II mediator in stress-induced hyperphosphorylation of Msn2 in Saccharomyces cerevisiae. Mol Microbiol. 2006; 62: 438–452. doi: 10.1111/j.1365-2958.2006.05363.x PMID: 17020582

22. Sadeh A, Baran D, Volokh M, Aharoni A. Conserved motifs in the Msn2-activating domain are important for Msn2-mediated yeast stress response. J Cell Sci. 2012; 125: 3333–3342. doi: 10.1242/jcs.096446 PMID: 22505609

23. Thakur JK, Arthanari H, Yang F, Chau KH, Wagner G, Näär AM. Mediator subunit Gal11p/MED15 is required for fatty acid-dependent gene activation by yeast transcription factor Oaf1p. J Biol Chem. 2009; 284: 4422–4428. doi: 10.1074/jbc.M808263200 PMID: 19056732

24. Thakur JK, Arthanari H, Yang F, Pan S-J, Fan X, Breger J, et al. A nuclear receptor-like pathway regulating multidrug resistance in fungi. Nature. 2008; 452: 604–609. doi: 10.1038/nature06836 PMID: 18385733

25. Han Y, Kodadek T. Peptides selected to bind the Gal80 repressor are potent transcriptional activation domains in yeast. J Biol Chem. 2000; 275: 14979–14984. PMID: 10809742

26. Frangioni JV, LaRiccia LM, Cantley LC, Montminy MR. Minimal activators that bind to the KIX domain of p300/CBP identified by phage display screening. Nat Biotechnol. 2000; 18: 1080–1085. doi: 10.1038/80280 PMID: 11017047

27. Bates CA, Pomerantz WC, Mapp AK. Transcriptional tools: Small molecules for modulating CBP KIX-dependent transcriptional activators. Biopolymers. 2011; 95: 17–23. doi: 10.1002/bip.21548 PMID: 20882601

28. Langlois C, Del Gatto A, Arseneault G, Lafrance-Vanasse J, De Simone M, Morse T, et al. Structure-based design of a potent artificial transactivation domain based on p53. J Am Chem Soc. 2012; 134: 1715–1723. doi: 10.1021/ja208999e PMID: 22191432

29. Pomerantz WC, Wang N, Lipinski AK, Wang R, Cierpicki T, Mapp AK. Profiling the dynamic interfaces of fluorinated transcription complexes for ligand discovery and characterization. ACS Chem Biol. 2012; 7: 1345–1350. doi: 10.1021/cb3002733 PMID: 22725662

30. Wu Y, Reece RJ, Ptashne M. Quantitation of putative activator-target affinities predicts transcriptional activating potentials. EMBO J. 1996; 15: 3951–3963. PMID: 8670900

31. Lu X, Ansari AZ, Ptashne M. An artificial transcriptional activating region with unusual properties. Proc Natl Acad Sci USA. 2000; 97: 1988–1992. doi: 10.1073/pnas.040573197 PMID: 10681438

32. Ptashne M. The chemistry of regulation of genes and other things. J Biol Chem. 2014; 289: 5417–5435. doi: 10.1074/jbc.X114.547323 PMID: 24385432

33. Ma J, Ptashne M. A new class of yeast transcriptional activators. Cell. 1987; 51: 113–119. PMID: 3115591

34. Ruden DM, Ma J, Li Y, Wood K, Ptashne M. Generating yeast transcriptional activators containing no yeast protein sequences. Nature. 1991; 350: 250–252. doi: 10.1038/350250a0 PMID: 2005981

35. Ma J, Ptashne M. Deletion analysis of GAL4 defines two transcriptional activating segments. Cell. 1987; 48: 847–853. PMID: 3028647

36. Teufel DP, Freund SM, Bycroft M, Fersht AR. Four domains of p300 each bind tightly to a sequence spanning both transactivation subdomains of p53. Proc Natl Acad Sci USA. 2007; 104: 7009–7014. doi: 10.1073/pnas.0702010104 PMID: 17438265

37. Gamper AM, Roeder RG. Multivalent Binding of p53 to the STAGA Complex Mediates Coactivator Recruitment after UV Damage. Mol Cell Biol. 2008; 28: 2517–2527. doi: 10.1128/MCB.01461-07 PMID: 18250150

38. Feng H, Jenkins LMM, Durell SR, Hayashi R, Mazur SJ, Cherry S, et al. Structural basis for p300 Taz2-p53 TAD1 binding and modulation by phosphorylation. Structure. 2009; 17: 202–210. doi: 10.1016/j.str.2008.12.009 PMID: 19217391

39. Ferreon JC, Lee CW, Arai M, Martinez-Yamout MA, Dyson HJ, Wright PE. Cooperative regulation of p53 by modulation of ternary complex formation with CBP/p300 and HDM2. Proc Natl Acad Sci USA. 2009; 106: 6591–6596. doi: 10.1073/pnas.0811023106 PMID: 19357310

40. Jenkins LMM, Yamaguchi H, Hayashi R, Cherry S, Tropea JE, Miller M, et al. Two distinct motifs within the p53 transactivation domain bind to the Taz2 domain of p300 and are differentially affected by phosphorylation. Biochemistry. 2009; 48: 1244–1255. doi: 10.1021/bi801716h PMID: 19166313

41. Brzovic PS, Heikaus CC, Kisselev L, Vernon R, Herbig E, Pacheco D, et al. The acidic transcription activator Gcn4 binds the mediator subunit Gal11/Med15 using a simple protein interface forming a fuzzy complex. Mol Cell. 2011; 44: 942–953. doi: 10.1016/j.molcel.2011.11.008 PMID: 22195967

42. Fassler JS, Winston F. The Saccharomyces cerevisiae SPT13/GAL11 gene has both positive and negative regulatory roles in transcription. Mol Cell Biol. 1989; 9: 5602–5609. PMID: 2685570

43. Suzuki Y, Nogi Y, Abe A, Fukasawa T. GAL11 protein, an auxiliary transcription activator for genes encoding galactose-metabolizing enzymes in Saccharomyces cerevisiae. Mol Cell Biol. 1992; 12: 4806. PMID: 1406662

44. Sakurai H, Hiraoka Y, Fukasawa T. Yeast GAL11 protein is a distinctive type transcription factor that enhances basal transcription in vitro. Proc Natl Acad Sci USA. 1993; 90: 8382–8386. PMID: 8378310

45. Lin J, Chen J, Elenbaas B, Levine AJ. Several hydrophobic amino acids in the p53 amino-terminal domain are required for transcriptional activation, binding to mdm-2 and the adenovirus 5 E1B 55-kD protein. Genes Dev. 1994; 8: 1235–1246. PMID: 7926727

46. Jackson BM, Drysdale CM, Natarajan K, Hinnebusch AG. Identification of seven hydrophobic clusters in GCN4 making redundant contributions to transcriptional activation. Mol Cell Biol. 1996; 16: 5557–5571. PMID: 8816468

47. Uesugi M, Nyanguile O, Lu H, Levine AJ, Verdine GL. Induced alpha helix in the VP16 activation domain upon binding to a human TAF. Science. 1997; 277: 1310–1313. PMID: 9271577

48. Ernst P, Wang J, Huang M, Goodman RH, Korsmeyer SJ. MLL and CREB bind cooperatively to the nuclear coactivator CREB-binding protein. Mol Cell Biol. 2001; 21: 2249–2258. doi: 10.1128/MCB.21.7.2249-2258.2001 PMID: 11259575

49. Jeong CJ, Yang SH, Xie Y, Zhang L, Johnston SA, Kodadek T. Evidence that Gal11 protein is a target of the Gal4 activation domain in the mediator. Biochemistry. 2001; 40: 9421–9427. PMID: 11478912

50. De Guzman RN, Goto NK, Dyson HJ, Wright PE. Structural Basis for Cooperative Transcription Factor Binding to the CBP Coactivator. Journal of Molecular Biology. 2006; 355: 1005–1013. doi: 10.1016/j.jmb.2005.09.059 PMID: 16253272

51. Majmudar CY, Wang B, Lum JK, Håkansson K, Mapp AK. A high-resolution interaction map of three transcriptional activation domains with a key coactivator from photo-cross-linking and multiplexed mass

spectrometry. Angew Chem Int Ed Engl. 2009; 48: 7021–7024. doi: 10.1002/anie.200902669 PMID: 19681084

52. Jedidi I, Zhang F, Qiu H, Stahl SJ, Palmer I, Kaufman JD, et al. Activator Gcn4 employs multiple segments of Med15/Gal11, including the KIX domain, to recruit mediator to target genes in vivo. J Biol Chem. 2010; 285: 2438–2455. doi: 10.1074/jbc.M109.071589 PMID: 19940160

53. Arai M, Dyson HJ, Wright PE. Leu628 of the KIX domain of CBP is a key residue for the interaction with the MLL transactivation domain. FEBS Lett. 2010; 584: 4500–4504. doi: 10.1016/j.febslet.2010.10.024 PMID: 20969867

54. Wang N, Lodge JM, Fierke CA, Mapp AK. Dissecting allosteric effects of activator-coactivator complexes using a covalent small molecule ligand. Proc Natl Acad Sci USA. 2014; 111: 12061–12066. doi: 10.1073/pnas.1406033111 PMID: 25049401

55. Chang C, Gonzalez F, Rothermel B, Sun L, Johnston SA, Kodadek T. The Gal4 activation domain binds Sug2 protein, a proteasome component, in vivo and in vitro. J Biol Chem. 2001; 276: 30956–30963. doi: 10.1074/jbc.M102254200 PMID: 11418596

56. Jabbur JR, Tabor AD, Cheng X, Wang H, Uesugi M, Lozano G, et al. Mdm-2 binding and TAF(II)31 recruitment is regulated by hydrogen bond disruption between the p53 residues Thr18 and Asp21. Oncogene. 2002; 21: 7100–7113. doi: 10.1038/sj.onc.1205856 PMID: 12370832

57. Klein J, Nolden M, Sanders SL, Kirchner J, Weil PA, Melcher K. Use of a genetically introduced cross-linker to identify interaction sites of acidic activators within native transcription factor IID and SAGA. J Biol Chem. 2003; 278: 6779–6786. doi: 10.1074/jbc.M212514200 PMID: 12501245

58. Kussie PH, Gorina S, Marechal V, Elenbaas B, Moreau J, Levine AJ, et al. Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain. Science. 1996; 274: 948–953. PMID: 8875929

59. Liu W-L, Coleman RA, Ma E, Grob P, Yang JL, Zhang Y, et al. Structures of three distinct activator-TFIID complexes. Genes Dev. 2009; 23: 1510–1521. doi: 10.1101/gad.1790709 PMID: 19571180

60. Uesugi M, Verdine GL. The alpha-helical FXXPhiPhi motif in p53: TAF interaction and discrimination by MDM2. Proc Natl Acad Sci USA. 1999; 96: 14801–14806. PMID: 10611293

61. Sandholzer J, Hoeth M, Piskacek M, Mayer H, de Martin R. A novel 9-amino-acid transactivation domain in the C-terminal part of Sox18. Biochem Biophys Res Commun. 2007; 360: 370–374. doi: 10.1016/j.bbrc.2007.06.095 PMID: 17603017

62. Wojciak JM, Martinez-Yamout MA, Dyson HJ, Wright PE. Structural basis for recruitment of CBP/p300 coactivators by STAT1 and STAT2 transactivation domains. EMBO J. 2009; 28: 948–958. doi: 10.1038/emboj.2009.30 PMID: 19214187

63. Courey AJ, Holtzman DA, Jackson SP, Tjian R. Synergistic activation by the glutamine-rich domains of human transcription factor Sp1. Cell. 1989; 59: 827–836. PMID: 2512012

64. Courey AJ, Tjian R. Analysis of Sp1 in vivo reveals multiple transcriptional domains, including a novel glutamine-rich activation motif. Cell. 1988; 55: 887–898. PMID: 3142690

65. Pascal E, Tjian R. Different activation domains of Sp1 govern formation of multimers and mediate transcriptional synergism. Genes Dev. 1991; 5: 1646–1656. PMID: 1885006