Polyglutamine tracts as modulators of transcriptional activation from yeast to mammals

Journal Article

Author(s):
Atanesyan, Lilit; Gunther, Viola; Dichtl, Bernhard; Georgiev, Oleg; Schaffner, Walter

Polyglutamine tracts as modulators of transcriptional activation from yeast to mammals

Lilit Atanesyan[1,2], Viola Günther[1], Bernhard Dichtl[1], Oleg Georgiev[1] and Walter Schaffner[1,*]

[1] Institute of Molecular Life Sciences, University of Zürich, CH-8057 Zürich, Switzerland

[2] Life Science Zurich Graduate School, University of Zürich/ETH Zürich, CH-8057 Zürich, Switzerland

*Corresponding author
 e-mail: walter.schaffner@imls.uzh.ch

Abstract

Microsatellite repeats are genetically unstable and subject to expansion and shrinkage. A subset of them, triplet repeats, can occur within the coding region and specify homomeric tracts of amino acids. Polyglutamine (polyQ) tracts are enriched in eukaryotic regulatory proteins, notably transcription factors, and we had shown before that they can contribute to transcriptional activation in mammalian cells. Here we generalize this finding by also including evolutionarily divergent organisms, namely, *Drosophila* and baker's yeast. In all three systems, Gal4-based model transcription factors were more active if they harbored a polyQ tract, and the activity depended on the length of the tract. By contrast, a polyserine tract was inactive. PolyQs acted from either an internal or a C-terminal position, thus ruling out a merely structural 'linker' effect. Finally, a two-hybrid assay in mammalian cells showed that polyQ tracts can interact with each other, supporting the concept that a polyQ-containing transcription factor can recruit other factors with polyQ tracts or glutamine-rich activation domains. The widespread occurrence of polyQ repeats in regulatory proteins suggests a beneficial role; in addition to the contribution to transcriptional activity, their genetic instability might help a species to adapt to changing environmental conditions in a potentially reversible manner.

Keywords: *Drosophila*; evolvability; polyglutamine repeat; *Saccharomyces cerevisiae*; transcription factor; triplet repeat.

Introduction

Polyglutamines (PolyQs) are the most frequent type of homo-amino acid tracts found in the *Drosophila melanogaster* proteome (Bettencourt et al., 2007). At least 30 *Drosophila* proteins possess polyQ tracts of 20 residues or longer, which are conserved in other drosophilid species. In the human proteome, there are several prevalent homo-amino acid tracts, whereby polyQ tracts are the most frequent homo-amino

acid tracts encoded by repeats of a single codon (CAG) (Kozlowski et al., 2010). Screening of the *Saccharomyces cerevisiae* protein database (www.yeastgenome.org) revealed that polyQ tracts are the most common homopolymeric amino acid tracts of eight amino acids length or longer. Functional association analysis showed that the most prevalent triplet repeats found in human ORFs (CGG coding for alanine; CAG, glutamine; CCG, proline; AGG, glutamic acid; ACC, histidine) predominantly occur in genes associated with transcription-related functions (Bettencourt et al., 2007). For example, transcriptional regulators such as Snf5 and Gal11 in *S. cerevisiae*, prospero and dClock in *Drosophila*, and TATA box-binding protein (TBP), forkhead box protein P2 (FoxP2), and the androgen receptor in humans contain polyQ tracts of at least 20 residues. In most of the above-mentioned transcription factors, the role of the polyQ tract was not studied, but in a synthetic transcription factor, polyQs were demonstrated to increase the transactivation potential (Gerber et al., 1994). Furthermore, many regulators of transcription contain glutamine-rich activation domains that are important for protein-protein interactions and gene expression regulation (Courey and Tjian, 1988; Seipel et al., 1992).

PolyQ tracts encoded by CAG repeats are particularly problematic due to their genetic instability caused by out-of-register recombination or DNA strand slippage during replication (Gao et al., 2008), an effect that was also observed with other types of short repeats (e.g., Bois et al., 2001). Nine human neurodegenerative disorders, all of them inherited gain-of-function diseases, are caused by the expansion of CAG repeats. These include Huntington's disease, spinobulbar muscular atrophy (Kennedy disease), several forms of spinocerebellar ataxia (SCA1, SCA2, SCA3, SCA6, SCA7, and SCA17) and dentatorubral pallidoluysian atrophy (Gusella and MacDonald, 2000; Mirkin, 2007; Orr and Zoghbi, 2007). The critical maximal length of polyQ tracts in eukaryotes is usually 35–40 residues. Longer polyQ tracts are prone to self-aggregation into fibrils, wherein the polyQs form stable 'polar zippers' of β-barrels, in agreement with a so-called β-helix nanotube model (Perutz et al., 1994, 2002; Merlino et al., 2006). A main feature of polyQ-associated diseases is intraneural aggregates of proteins with expanded polyQ tracts. Large insoluble aggregates (inclusion bodies) might even have a protective role in affected cells by preventing the alternative formation of many small, more deleterious oligomeric aggregates (Saudou et al., 1998; Sánchez et al., 2003; Arrasate et al., 2004; Miller et al., 2010).

Transcriptional dysregulation has been implicated in several polyQ diseases (Sugars and Rubinsztein, 2003; Riley and Orr, 2006). Expanded polyQ tracts can interact with polyQ-containing or other transcription factors and cofactors

and thus interfere with their normal function. Target proteins can be sequestered by polyQ-protein monomers or recruited into aggregates (Kazantsev et al., 1999; Okazawa, 2003; Helmlinger et al., 2006; Friedman et al., 2007; Huang et al., 2011). Of note, aberrant polyQ interactions were described for CREB-binding protein, p53, Sp1, and TAFII130 and are associated with impaired transcription (Nucifora et al., 2001; Dunah et al., 2002; Bae et al., 2005). Furthermore, expanded trinucleotide repeat sequences within the non-coding region of mRNA can be toxic for the cell and contribute to disease symptoms (Li and Bonini, 2010).

We postulated before that the genetic instability of polyamino acid tracts, notably of polyQs, has a beneficial role in evolution because it allows for the modulation of transcription factor activity (Gerber et al., 1994; see also Künzler et al., 1995). Here we generalize the earlier findings with mammalian cells by including the model systems of *Drosophila* and yeast. We propose that variable repeats located in certain key proteins allow for potentially reversible adaptations to changing environmental conditions. Such a trait would increase the 'evolvability' and thus the fitness of a species.

Results

PolyQs contribute to transcriptional activation in a reporter system driven by a synthetic transcription factor

To test the influence of a polyQ tract on transcriptional activation, we generated chimeric transcription factors in which the DNA-binding domain of yeast Gal4 (Gal4DBD) was linked *via* polyQ tracts of various length to the strong activation domain of the herpes simplex virus protein VP16 (VP16AD[80]) or to the N-terminal half of it (VP16AD[40]) (Figure 1A and C). Gal4-VP16 was used as a test system because of its universal activity in all eukaryotes, implying that it reflects a fundamental mechanism of transcriptional activation and is not, for example, dependent on some cell type-specific coactivator or protein modification. The number of glutamine residues was chosen such that it corresponds to the naturally occurring nonpathogenic size (20Q), close to pathogenic size (40Q), or expanded abnormal length (86Q). When tested in *Drosophila* cell culture, polyQ tracts had a positive effect on transcription efficiency (Figure 2A). Reporter gene transcript levels were highest with 40Q- and 86Q-, and were also elevated with 20Q- containing effector proteins compared with the one without polyQ.

Next we produced transgenic *Drosophila* strains containing the chimeric transcription factors and tested them with a *UAS-GFP* reporter system. We found that polyQs also contributed to transcriptional activation *in vivo*, whereby flies with the 86Q-containing transcription factor showed a stronger GFP signal than the ones with 20Q (Figure 2B). A positive correlation between transactivation potential and the length of the polyQ tract was also observed with a *UAS-lacZ* reporter (Figure 2C).

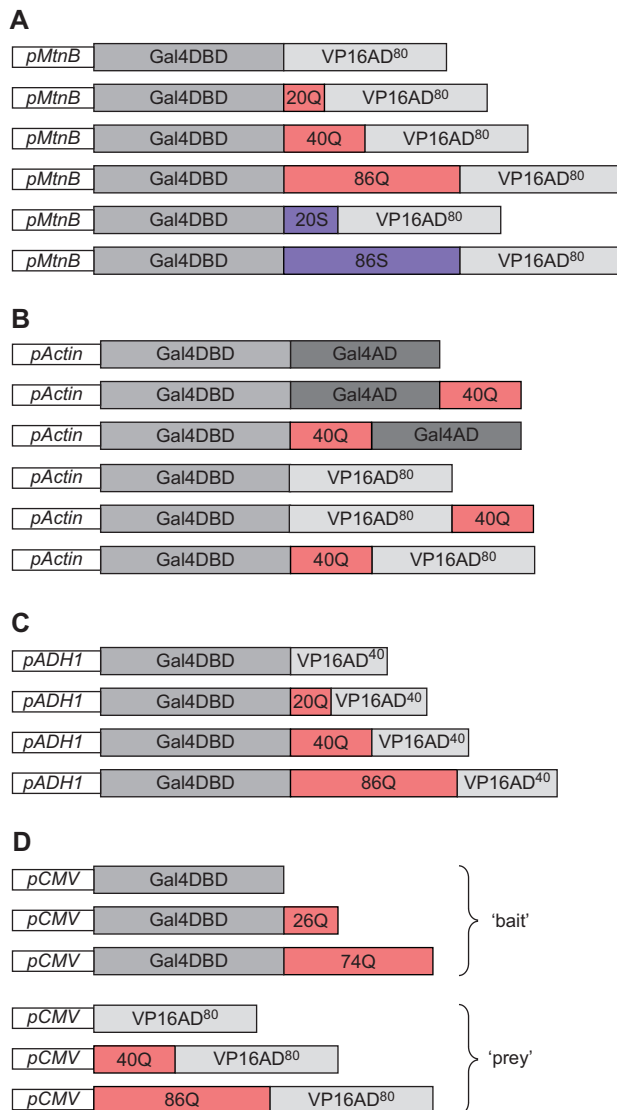Together with polyasparagine, polyproline, and polyglutamic acid, polyQs are often found in proteins that are

Figure 1   Model transcription factors used in this study. Chimeric transcription factors were driven either by (A) the metal-inducible *Drosophila metallothionein B* promoter, (B) the *Drosophila actin 5c* promoter, or (C) the yeast *alcohol dehydrogenase 1* (*ADH1*) promoter. PolyQ stretches encoded by CAG codons are denoted by 20Q, 40Q, and 86Q; 20 and 86 AGC codons encode 20 and 86 serines (20S and 86S, respectively). (D) Fusion proteins used for mammalian two-hybrid studies. The constructs were expressed from the strong enhancer/promoter of human cytomegalovirus.

involved in basal and regulated transcription in baker's yeast. To test the polyQ-containing transcription factors in yeast, we used an *in vivo* plate assay with X-Gal in the medium and the liquid lacZ assay. As in mammalian cells and *Drosophila*, in yeast, the synthetic transcription factors carrying polyQ tracts of various sizes activated the transcription of a *LacZ* reporter gene in a polyQ length-dependent manner (Figure 2D and E). In our yeast studies, a less active subsegment of the VP16 activation domain (VP16AD[40]) was used to avoid selection against the very potent and possibly toxic full-length activation domain (Sadowski et al., 1988) and also to retain the
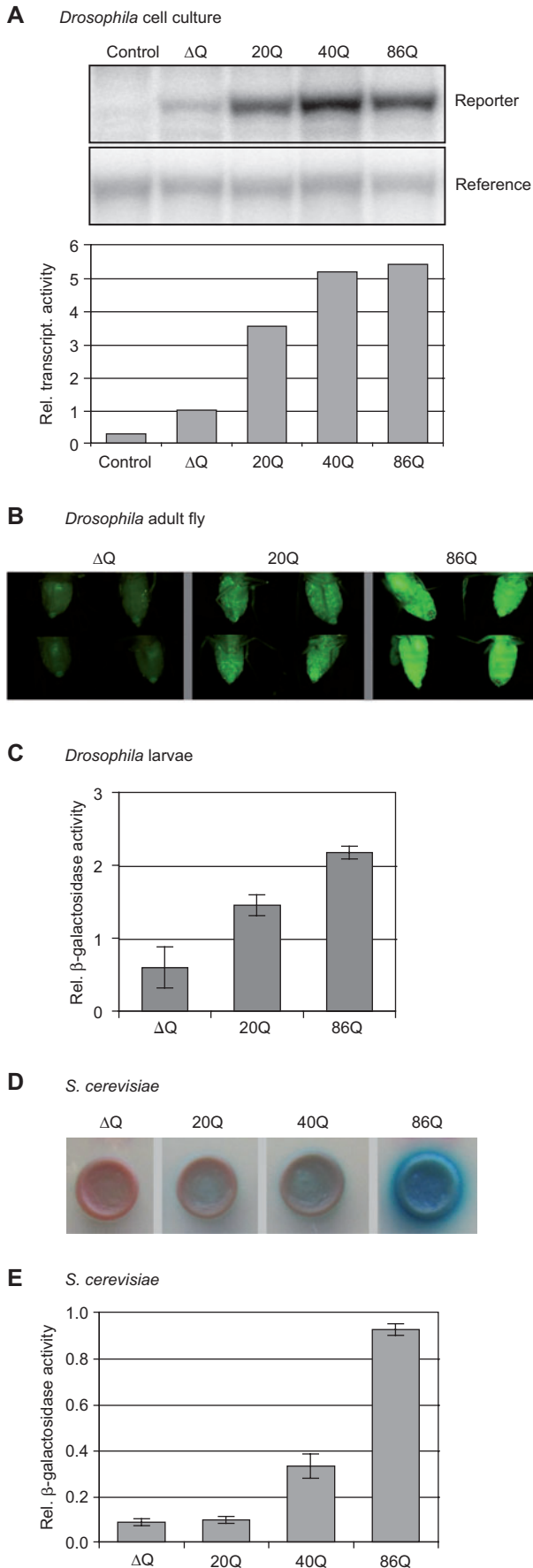
A *Drosophila* cell culture

B *Drosophila* adult fly

C *Drosophila* larvae

D *S. cerevisiae*

E *S. cerevisiae*

Figure 2 PolyQs contribute to transcriptional activation in a length-dependent manner in *Drosophila* cell culture, flies, and yeast. (A) *Drosophila* Kc cells were transfected with 20 µg reporter plasmid, 10 µg effector plasmid (Figure 1A), and 5 µg reference plasmid. The expression of effector genes (*Gal4DBD-nQ-VP16AD[80]*) was induced by supplementing cell growth media with $CuSO_4$ to a final concentration of 100 µM. Forty-eight hours after induction, transcript levels of the *5xGal-OVEC* reporter and *actin-OVEC* reference were determined by the S1 nuclease protection assay. ΔQ=no glutamines. (B) PolyQs boost reporter gene expression in transgenic flies. Freshly enclosed flies of the genotype *y w*; *UAS-GFP/+*; *Gal4DBD-nQ-VP16/+* were chloroform-anesthetized and pictures of four fly abdomens for each genotype were taken with the same exposure time (1200 ms). (C) For the quantitative liquid β-galactosidase assay in *Drosophila*, third instar larvae of the genotype *y w*; *Gal4DBD-nQ-VP16AD[80]/ UAS-LacZ* were collected. β-Galactosidase activity was measured in pools of three larvae. (D) Activity of chimeric transcription factors (Figure 1C) in a yeast *in vivo* plate assay using X-gal in the medium. The increasing blue color intensity of yeast cells indicates increasing activity of the transcription factor that drives the expression of the *LacZ* reporter gene. (E) Quantitative β-galactosidase assay in yeast liquid culture using ONPG as a substrate.

possibility to increase transactivation by introducing a polyQ stretch.

To test whether the increased activity conferred by the polyQ tract was position specific (i.e., confined to the linker region between the DNA-binding and activation domains) or was a more general phenomenon, we placed the polyQ tract at the C-terminus of the chimeric transcription factors, containing the activation domain of either Gal4 or VP16 (Figure 1B). A C-terminally located polyQ tract also boosted transcription, although the effect was not as pronounced as with an internal polyQ 'linker' (Figure 3).

PolyQs, but not polyserines, stimulate transcription

Poly[CAG] can be translated, depending on the reading frame, into three different homopolymeric amino acid stretches, namely, polyglutamine, polyserine, and polyalanine. Homopolymeric serine tracts have been reported to behave in a neutral manner with regard to transcriptional activity, whereas polyalanines act as transcriptional repressors (Lanz et al., 1995). In HEK293T cells, the constructs with 20 or 86 serines between Gal4DBD and VP16AD[80] did not increase transcriptional activity (Figure 4A). A polyserine tract also had no effect in *Drosophila* (Figure 4B). These results show that activation by polyQ tracts located between DNA-binding and activation domain is not merely a linker effect, in line with the finding that a polyQ tract was also active if added at the C-terminus (see above).

Transactivation via PolyQ interactions tested in a mammalian two-hybrid system

It has been reported that polyQ domains of proteins can interact with other polyQ-containing or glutamine-rich transcription factors (Chen et al., 2004). To test for polyQ-polyQ interactions *in vivo*, we generated a mammalian two-hybrid system
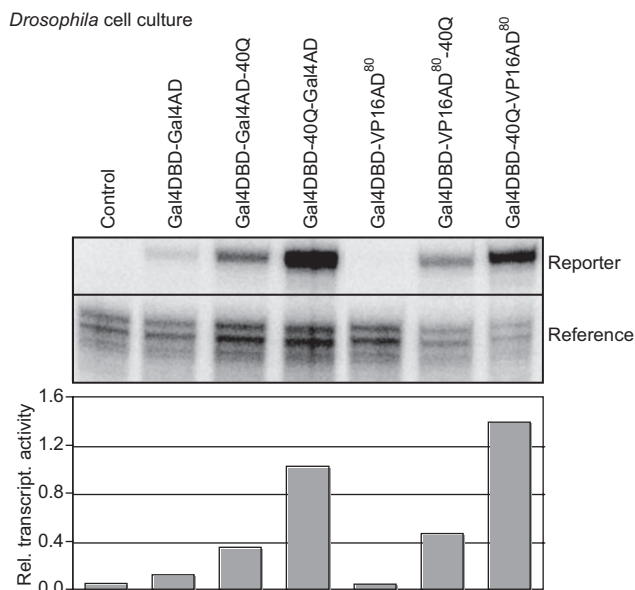
Figure 3 PolyQs at the C-terminus also contribute to transactivation.

Drosophila S2 cells were transfected with 1 µg reporter plasmid, 10 µg effector plasmid (Figure 1A and B), and 2.5 µg reference plasmid. The activity of transcription factors with a polyQ stretch added to the C-terminus (*Gal4DBD-Gal4AD-40Q* or *Gal4DBD-VP16AD[80]-40Q*), with a 'linker' polyQ tract (*Gal4DBD-40Q-Gal4AD* or *Gal4DBD-40Q-VP16AD[80]*), or without polyQ (*Gal4DBD-Gal4AD* or *Gal4DBD-VP16AD[80]*) was studied by measuring the transcript levels of the *5xGal-OVEC* reporter and *actin-OVEC* reference by the S1 assay.

where a polyQ tract fused to the Gal4DBD acts as a 'bait' and a polyQ tract fused to the VP16AD[80] as a 'prey' (Figure 1D). The interaction between the two polyQ-containing proteins indeed induced transcription in HEK293T cells, as shown by the S1-nuclease protection assay, whereby longer polyQs on one or both partners gave stronger signals (Figure 5A). A bandshift assay showed that all of the bait constructs were similarly expressed. However, the interaction between polyQ-containing bait and prey proteins is probably weak, as no supershift with the prey was observed in electrophoretic mobility shift assays (EMSAs) (Figure 5B). We could not observe a two-hybrid interaction in other cell lines and in *Drosophila* (data not shown), likely because HEK293T cells take up and express transfected DNA most efficiently.

Discussion

Despite their genetic instability, especially their tendency to expand and the potential to cause neurodegeneration, polyQs are a widespread feature in eukaryotic proteomes, from yeast to humans, and are enriched in transcriptional regulators. The presence of polyQ tracts in so many proteins, notably in DNA-binding transcription factors and cofactors, suggests a functional role. It was previously shown in mammalian cells that glutamine-rich domains can synergize with acidic activation
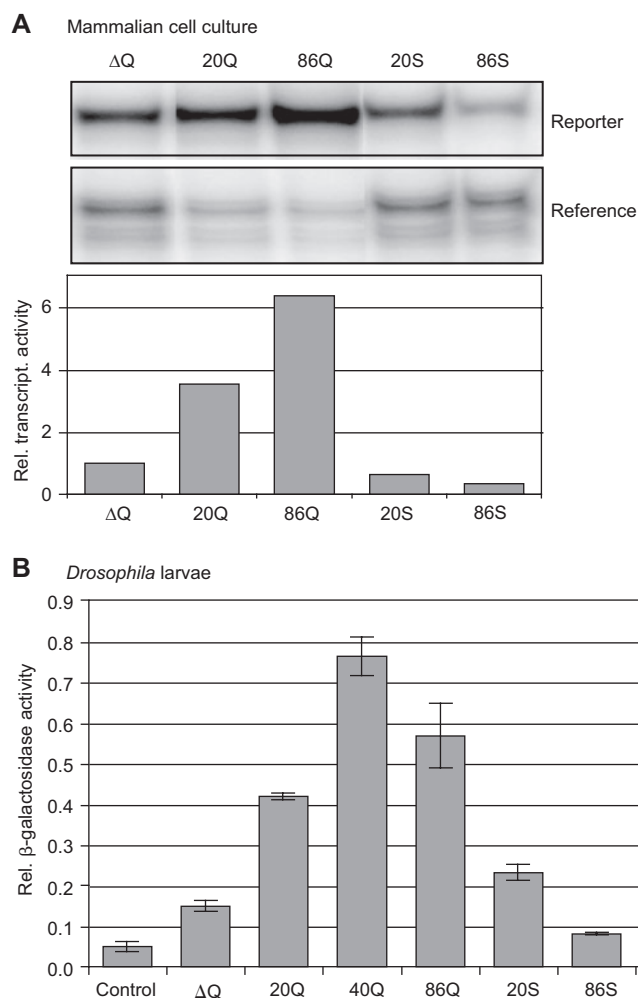

Figure 4 Polyserine tracts do not contribute to transcription activation.

(A) Transcript levels of the *5xGal-OVEC* reporter and *CMV-OVEC* reference in HEK293T cells were determined by the S1 assay. Cells were transfected with 10 µg reporter plasmid, 5 µg of *Gal4DBD-nQ/nS-VP16AD[80]* effector constructs (Figure 1A) and 2 µg reference plasmid. 500 µM $CuSO_4$ was used to induce the *MtnB* promoter for 24 h. (B) PolyQ and polyserine in transgenic *Drosophila* larvae. Third instar larvae of genotype *y w*; *UAS-GFP/+*; *Gal4DBD-nQ/nS-VP16AD[80]/+* were collected. Reporter activity was determined by the liquid LacZ assay.

domains (Seipel et al., 1992). Indeed, introduction of a polyQ stretch into a synthetic transcription factor increased its transcriptional activity. By varying the size of the polyQ stretch within a transcription factor, we could modulate its activity. A similar effect was observed for the androgen receptor; however, there was a negative correlation between polyQ length and transcription factor activity (Chamberlain et al., 1994; Wang et al., 2004). A positive or negative correlation of repeat number and protein activity is not confined to homomeric polyamino acid tracts: for example, cell adhesion in yeast is influenced by the repeat number of a subdomain in the flocculation protein Flo1 (Verstrepen et al., 2005).
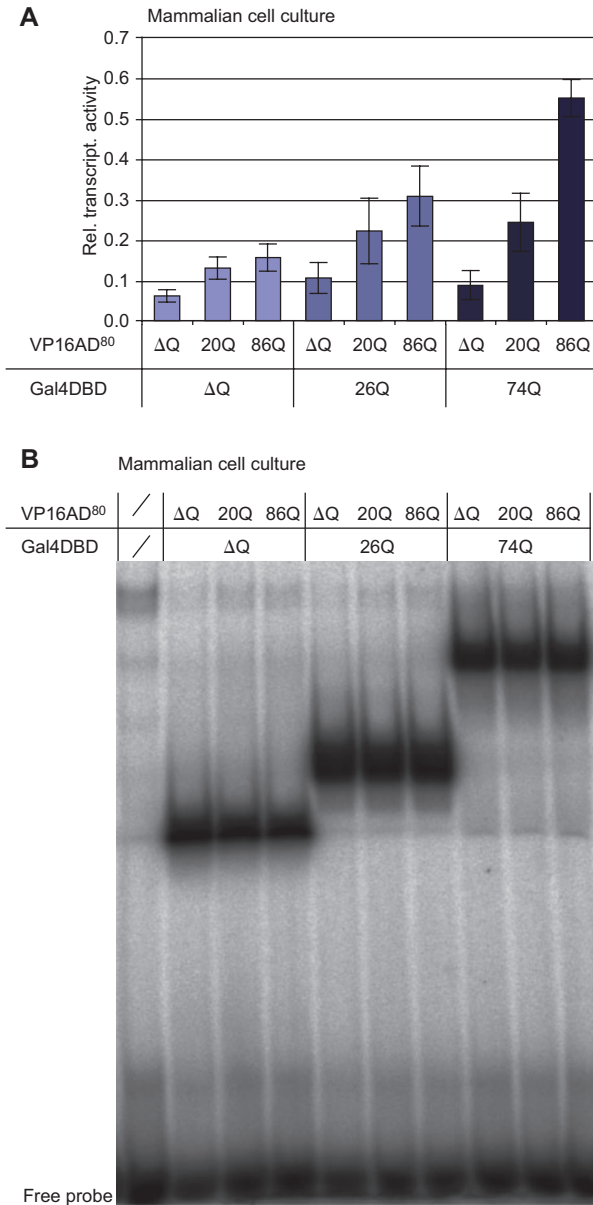
A

B

Figure 5 Analysis of polyQ tract interactions in a mammalian two-hybrid system.
(A) HEK293T cells were transfected with 10 μg *5xGal-OVEC* reporter, 4 μg of each of the two-hybrid effector plasmids (Figure 1D), and 1 μg *CMV-OVEC* reference plasmid. Reporter and reference gene transcripts were quantified by the S1 assay. (B) EMSA of a [^{32}P]-labeled Gal4-binding oligonucleotide with nuclear extracts of HEK293T cells transfected with 10 μg of the 'bait' and 'prey' constructs each, demonstrating that DNA binding is not affected by polyQ length and that the binding of bait (*Gal4DBD-ΔQ/26Q/74Q*) remains unchanged in the presence of prey (*ΔQ/20Q/86Q-VP16AD*). A bait-prey interaction as seen *in vivo* by the two-hybrid assay (A) is not observed under EMSA conditions.

A certain number of repeats is optimal for transcriptional activation *in vivo* in our case, in transgenic *Drosophila* introducing 40Q resulted in the highest transcriptional activity (even stronger than 86Q). The relatively lower activity with

86Q , which was also observed in mammalian cells (Gerber et al., 1994), might be explained by the propensity of oversized polyQ tracts to aggregate and thereby inactivate a fraction of the chimeric transcription factor. Somewhat surprisingly, the synthetic transcription factor with 86Qs was the most potent one in yeast, even though the physiological upper limit for polyQ tracts in yeast, like in *Drosophila* and mammals, is around 40Q.

The type of glutamine-encoding triplet (CAA or CAG) and their degree of intermingling is critical for the genetic stability of repeat size and therefore in the development polyQ-associated neurological disorders. Pure CAG repeats are less stable than repeats encoded by a mix of interspersed CAG and CAA triplets (Gao et al., 2008). An example of the latter is FoxP2, a transcription factor needed for speech and language development in humans, with a long, quite stable polyQ stretch of 40 glutamines (Bruce and Margolis, 2002; Enard et al., 2002; Webb and Zhang, 2005). Most likely, CAA interruptions counteract CAG repeat slippage and expansions in DNA. CAA interruptions also play a role in the mRNA folding of the repeat region in SCA2, SCA1, TBP, FoxP2, and MAML2 transcripts: depending on the number and frequency of CAA interruptions, polyQ transcripts exhibit different secondary structures, such as short or long hairpins or no hairpins at all (Sobczak and Krzyzosiak, 2005).

We have shown that polyQs can increase the activity of a transcription factor in diverse model organisms from yeast to mammals. However, the same effect could, in principle, be achieved without polyQ stretch by simply making a standard activation domain more potent. Thus, the question of the biological role of polyQ tracts remains. We postulated before that genetically unstable polyamino acid tracts have a positive role in evolution: by modulating transcription factor activity, they might result in subtle or overt alterations of an organism's genetic program (Gerber et al., 1994). In recent years, much attention has been given to reversible epigenetic modifications and their role in the adaptation to new ecological niches. Heritable alterations in the length of polyQ tracts, which, in principle, are readily reversible, represent an intermediate state between classical DNA mutations and epigenetic changes *via* chromatin proteins and DNA methylation. Whether environmental conditions can influence the direction of such semistable genetic alterations, as Künzler et al. (1995) speculated, remains to be seen.

Materials and methods

Plasmid constructions

Reporter and reference plasmids for mammalian and *Drosophila* cell culture transcription experiments are based on the OVEC system (Westin et al., 1987). *Gal4DBD-nQ-VP16AD^{80}* plasmids (whereby n indicates the number of glutamine residues), driven by the CMV promoter in the pSCT vector, were described previously (Gerber et al., 1994). For expression vectors encoding Gal4DBD-nQ-VP16AD fusion proteins (Figure 1A) that were inserted into the *Drosophila* genome *via* the φC31 system (Bischof et al., 2007), the inducible 183-bp *metallothionein B* promoter was inserted in between *Sal*I

and *Sac*I sites. Although constructs driven by the strong *actin* promoter could be used in cell culture, we could not obtain transgenic *Drosophila* with these constructs, possibly due to the toxicity of the construct. The 40 carboxyl terminal CAG codons followed by a stop codon were inserted into the *Age*I site of plasmids *Gal4DBD-GalAD* and *Gal4DBD-VP16AD⁸⁰*, in this case, both driven by the *Drosophila actin* promoter (Figure 1B).

For experiments in *S. cerevisiae*, we used the high-copy number pAS2ΔΔ vector, where fusions of Gal4DBD with the N-terminal 40 amino acids of VP16AD (VP16AD⁴⁰) were driven by the yeast *alcohol dehydrogenase 1* (*ADH1*) promoter (Figure 1C). *VP16AD⁴⁰*, encompassing only the N-terminal half of the strong VP16 activation domain, was generated by introducing several stop codons into *VP16AD⁸⁰*.

'Prey' and 'bait' constructs for two-hybrid studies (*Gal4DBD-nQ* and *nQ-VP16*) were generated by removing either the Gal4DBD or the VP16AD moieties from the CMV-driven *Gal4DBD-nQ-VP16* constructs. Untypical polyQ sizes (26 or 74Q) were the result of instability of triplet repeats during the cloning procedure (Figure 1D).

Detailed information about plasmid sequences and cloning strategies is available upon request.

Cell culture and transfections

Culturing and transfection of human embryonic kidney 293T (HEK293T) cells was carried out in Dulbecco's modified Eagle medium (Invitrogen, Basel, Switzerland) supplemented with 10% fetal bovine serum (Biochrom AG, Berlin, Germany), 100 U/ml penicillin, 100 μg/ml streptomycin and 2 mM L-glutamine (Invitrogen, Basel, Switzerland). Reporter and reference genes were transfected into cells together with the expression vectors by the calcium-phosphate coprecipitation method (Graham and van der Eb, 1973; Wigler et al., 1978). Sixteen hours after transfection, cells were washed with TBS and incubated for 24 h in growth medium before harvesting them for RNA extraction or nuclear protein extract preparation. For transfections, a total of 20 μg DNA adjusted with herring sperm DNA was used per 10 cm cell culture dish. *Drosophila* Schneider 2 (S2) and Kc cells were cultured in Schneider's *Drosophila* medium (Invitrogen, Basel, Switzerland) containing 10% heat-inactivated fetal bovine serum (Biochrom AG, Berlin, Germany), 100 U/ml penicillin, and 100 μg/ml streptomycin. As for mammalian cells, the calcium-phosphate coprecipitation method was used for transfections. Cells were harvested for further manipulations 48 or 72 h after transfection.

S1 nuclease protection assay

Exponentially growing cells in a 10-cm culture dish were transfected with 1–5 μg of the particular expression plasmids, 10–20 μg of reporter, and 1–10 μg of reference plasmid, as indicated in the figure legends. The isolation of RNA and the S1 nuclease protection assay were performed as described previously (Weaver and Weissmann, 1979; Westin et al., 1987), with 100–150 μg of total RNA per sample. Signals were visualized using the fluorescent image analyzer FLA-7000 and quantified using the ImageGauge software (Fujifilm Life Science, Bucher Biotec AG, Basel, Switzerland). Reporter signals were normalized to the reference signals.

Electrophoretic mobility shift assay

Mammalian HEK293T cells were transiently transfected with the respective constructs and collected 48 h later. EMSAs were performed as described by Radtke et al. (1993) and Zhang et al.

(2001). Binding reactions were performed by incubating 20 fmol of [γ-³²P]-ATP end-labeled oligonucleotides containing the DNA-binding site for Gal4 with 5 μg of nuclear extracts prepared according to (Schreiber et al., 1993). Oligonucleotides used for EMSA are as follows (complementary base pairs are indicated in capital letters):

5′-CGG ATT AGA ACC CTC CGC Cgc-3′
3′-gta cGC CTA ATC TTG GGA GGC GG-5′

Liquid lacZ assay with *Drosophila* larvae

For measuring β-galactosidase activity, samples consisting of three wandering larvae were lysed by repeated freeze-that cycles in liquid nitrogen. The samples were vortexed with 100 μl Z-buffer (60 mM Na₂HPO₄, 40 mM NaH₂PO₄, 10 mM KCl, 1 mM MgSO₄) for 20 s. 700 μl Z-buffer supplemented with 50 mM β-mercaptoethanol was added to each sample. The reaction was started with 160 μl *ortho*-nitrophenyl-β-galactoside (ONPG) substrate (4 mg/ml). After 5 min at 30°C, 400 μl of 1 M Na₂CO₃ was added and the samples centrifuged for 10 min at 13 000 rpm. The OD of the clear supernatant was measured at 420 nm.

Yeast strains and gene expression

Yeast was grown in a rich YPD medium (2% glucose, 2% bacto-tryptone, 1% yeast extract) or in a synthetic drop-out medium if selection for plasmid-dependent growth was required (2% glucose, 0.17% yeast nitrogen base, 0.5% ammonium sulfate, 1×amino acid mix). The following *S. cerevisiae* strains were used: Y190 (MATa, *gal4-542*, *gal80-538*, *his3*, *trp1-901*, *ade2-101*, *ura3-52*, *leu2-3,112*, URA3::GAL1-LacZ, Lys2::GAL1-HIS3cyhr) and Y187 (*MATα*, *ura3-52*, *his3-200*, *ade2-101*, *trp1-901*, *leu2-3,112*, *gal4Δ*, *met-*, *gal80Δ*, *MEL1*, URA3::GAL1$_{UAS}$-GAL1$_{TATA}$-*lacZ*). Gene expression in yeast was monitored by an *in vivo* plate assay using X-Gal (bromo-chloro-indolyl-galactopyranoside) in the medium and by a quantitative liquid culture assay using ONPG as substrate (Clontech protocols).

Fly food, transgenic fly strains

One liter of fly food is composed of 55 g cornmeal (Maisgriess 54.401.025; Meyerhans Hotz AG, Weinfelden, Switzerland), 100 g yeast (Hefe Schweiz AG, Stettfurt, Switzerland), 75 g sugar (dextrose monohydrate), 8 g agar, and 15 ml antifungal Nipagin [Nipagin 33 g/l and Nipasol 66 g/l (Brenutag Schweizerhall AG, Basel, Switzerland) in 96% ethanol]. For the induction of the *MtnB* promoter, food was supplemented with CuSO₄ to 100 μM. Transgenes encoding chimeric transcription factors were introduced into the *Drosophila* genome at the *86Fb* locus on the third chromosome by the φC31 integrase-mediated transgenesis technique (Bischof et al., 2007). The *UAS-GFP* (second chromosome) and *UAS-lacZ* (third chromosome) transgenes were used as reporters.

Acknowledgments

References

Arrasate, M., Mitra, S., Schweitzer, E., Segal, M., and Finkbeiner, S. (2004). Inclusion body formation reduces levels of mutant huntingtin and the risk of neuronal death. Nature *431*, 805–810.

Bae, B.-I., Xu, H., Igarashi, S., Fujimuro, M., Agrawal, N., Taya, Y., Hayward, S.D., Moran, T.H., Montell, C., Ross, C.A., et al. (2005). p53 mediates cellular dysfunction and behavioral abnormalities in Huntington's disease. Neuron *47*, 29–41.

Bettencourt, B., Hogan, C., and Nimali, M. (2007). Polyglutamine expansion in *Drosophila*: thermal stress and Hsp70 as selective agents. J. Biosci. *32*, 537–547.

Bischof, J., Maeda, R., Hediger, M., Karch, F., and Basler, K. (2007). An optimized transgenesis system for *Drosophila* using germline-specific phiC31 integrases. Proc. Natl. Acad. Sci. USA *104*, 3312–3317.

Bois, P.R., Southgate, L., and Jeffreys, A.J. (2001). Length of uninterrupted repeats determines instability at the unstable mouse expanded simple tandem repeat family MMS10 derived from independent SINE B1 elements. Mamm. Genome *12*, 104–111.

Bruce, H. and Margolis, R. (2002). FOXP2: novel exons, splice variants, and CAG repeat length stability. Hum. Genet. *111*, 136–144.

Chamberlain, N.L., Driver, E.D., and Miesfeld, R.L. (1994). The length and location of CAG trinucleotide repeats in the androgen receptor N-terminal domain affect transactivation function. Nucleic Acids Res. *22*, 3181–3186.

Chen, S., Peng, G.-H., Wang, X., Smith, A.C., Grote, S.K., Sopher, B.L., and La Spada, A.R. (2004). Interference of Crx-dependent transcription by ataxin-7 involves interaction between the glutamine regions and requires the ataxin-7 carboxy-terminal region for nuclear localization. Hum. Mol. Genet. *13*, 53–67.

Courey, A.J. and Tjian, R. (1988). Analysis of Sp1 in vivo reveals multiple transcriptional domains, including a novel glutamine-rich activation motif. Cell *55*, 887–898.

Dunah, A., Jeong, H., Griffin, A., Kim, Y., Standaert, D., Hersch, S., Mouradian, M., Young, A., Tanese, N., and Krainc, D. (2002). Sp1 and TAFII130 transcriptional activity disrupted in early Huntington's disease. Science *296*, 2238–2243.

Enard, W., Przeworski, M., Fisher, S.E., Lai, C.S.L., Wiebe, V., Kitano, T., Monaco, A.P., and Paabo, S. (2002). Molecular evolution of FOXP2, a gene involved in speech and language. Nature *418*, 869–872.

Friedman, M.J., Shah, A.G., Fang, Z.-H., Ward, E.G., Warren, S.T., Li, S., and Li, X.-J. (2007). Polyglutamine domain modulates the TBP-TFIIB interaction: implications for its normal function and neurodegeneration. Nat. Neurosci. *10*, 1519–1528.

Gao, R., Matsuura, T., Coolbaugh, M., Zühlke, C., Nakamura, K., Rasmussen, A., Siciliano, M., Ashizawa, T., and Lin, X. (2008). Instability of expanded CAG/CAA repeats in spinocerebellar ataxia type 17. Eur. J. Hum. Genet. *16*, 215–222.

Gerber, H., Seipel, K., Georgiev, O., Höfferer, M., Hug, M., Rusconi, S., and Schaffner, W. (1994). Transcriptional activation modulated by homopolymeric glutamine and proline stretches. Science *263*, 808–811.

Graham, F.L. and van der Eb, A.J. (1973). A new technique for the assay of infectivity of human adenovirus 5 DNA. Virology *52*, 456–467.

Gusella, J.F. and MacDonald, M.E. (2000). Molecular genetics: unmasking polyglutamine triggers in neurodegenerative disease. Nat. Rev. Neurosci. *1*, 109–115.

Helmlinger, D., Tora, L., and Devys, D. (2006). Transcriptional alterations and chromatin remodeling in polyglutamine diseases. Trends Genet. *22*, 562–570.

Huang, S., Ling, J.J., Yang, S., Li, X.-J., and Li, S. (2011). Neuronal expression of TATA box-binding protein containing expanded polyglutamine in knock-in mice reduces chaperone protein response by impairing the function of nuclear factor-Y transcription factor. Brain *134*, 1943–1958.

Kazantsev, A., Preisinger, E., Dranovsky, A., Goldgaber, D., and Housman, D. (1999). Insoluble detergent-resistant aggregates form between pathological and nonpathological lengths of polyglutamine in mammalian cells. Proc. Natl. Acad. Sci. USA *96*, 11404–11409.

Kozlowski, P., de Mezer, M., and Krzyzosiak, W. (2010). Trinucleotide repeats in human genome and exome. Nucleic Acids Res. *38*, 4027–4039.

Künzler, P., Matsuo, K., and Schaffner, W. (1995). Pathological, physiological, and evolutionary aspects of short unstable DNA repeats in the human genome. Biol. Chem. Hoppe Seyler *376*, 201–211.

Lanz, R.B., Wieland, S., Hug, M., and Rusconi, S. (1995). A transcriptional repressor obtained by alternative translation of a trinucleotide repeat. Nucleic Acids Res. *23*, 138–145.

Li, L.-B. and Bonini, N.M. (2010). Roles of trinucleotide-repeat RNA in neurological disease and degeneration. Trends Neurosci. *33*, 292–298.

Merlino, A., Esposito, L., and Vitagliano, L. (2006). Polyglutamine repeats and beta-helix structure: molecular dynamics study. Proteins *63*, 918–927.

Miller, J., Arrasate, M., Shaby, B., Mitra, S., Masliah, E., and Finkbeiner, S. (2010). Quantitative relationships between huntingtin levels, polyglutamine length, inclusion body formation, and neuronal death provide novel insight into Huntington's disease molecular pathogenesis. J. Neurosci. *30*, 10541–10550.

Mirkin, S. (2007). Expandable DNA repeats and human disease. Nature *447*, 932–940.

Nucifora, F.J., Sasaki, M., Peters, M., Huang, H., Cooper, J., Yamada, M., Takahashi, H., Tsuji, S., Troncoso, J., Dawson, V., et al. (2001). Interference by huntingtin and atrophin-1 with cbp-mediated transcription leading to cellular toxicity. Science *291*, 2423–2428.

Okazawa, H. (2003). Polyglutamine diseases: a transcription disorder? Cell. Mol. Life Sci. *60*, 1427–1439.

Orr, H. and Zoghbi, H. (2007). Trinucleotide repeat disorders. Annu. Rev. Neurosci. *30*, 575–621.

Perutz, M., Johnson, T., Suzuki, M., and Finch, J. (1994). Glutamine repeats as polar zippers: their possible role in inherited neurodegenerative diseases. Proc. Natl. Acad. Sci. USA *91*, 5355–5358.

Perutz, M., Finch, J., Berriman, J., and Lesk, A. (2002). Amyloid fibers are water-filled nanotubes. Proc. Natl. Acad. Sci. USA *99*, 5591–5595.

Radtke, F., Heuchel, R., Georgiev, O., Hergersberg, M., Gariglio, M., Dembic, Z., and Schaffner, W. (1993). Cloned transcription factor MTF-1 activates the mouse metallothionein I promoter. EMBO J. *12*, 1355–1362.

Riley, B. and Orr, H. (2006). Polyglutamine neurodegenerative diseases and regulation of transcription: assembling the puzzle. Genes Dev. *20*, 2183–2192.

Sadowski, I., Ma, J., Triezenberg, S., and Ptashne, M. (1988). GAL4-VP16 is an unusually potent transcriptional activator. Nature *335*, 563–564.

Saudou, F., Finkbeiner, S., Devys, D., and Greenberg, M.E. (1998). Huntingtin acts in the nucleus to induce apoptosis but death does not correlate with the formation of intranuclear inclusions. Cell *95*, 55–66.

Schreiber, E., Tobler, A., Malipiero, U., Schaffner, W., and Fontana, A. (1993). cDNA cloning of human N-Oct3, a nervous-system

specific POU domain transcription factor binding to the octamer DNA motif. Nucleic Acids Res. *21*, 253–258.

Seipel, K., Georgiev, O., and Schaffner, W. (1992). Different activation domains stimulate transcription from remote ('enhancer') and proximal ('promoter') positions. EMBO J. *11*, 4961–4968.

Sobczak, K. and Krzyzosiak, W.J. (2005). CAG repeats containing CAA interruptions form branched hairpin structures in spinocerebellar ataxia type 2 transcripts. J. Biol. Chem. *280*, 3898–3910.

Sugars, K. and Rubinsztein, D. (2003). Transcriptional abnormalities in Huntington disease. Trends Genet. *19*, 233–238.

Sánchez, I., Mahlke, C., and Yuan, J. (2003). Pivotal role of oligomerization in expanded polyglutamine neurodegenerative disorders. Nature *421*, 373–379.

Verstrepen, K.J., Jansen, A., Lewitter, F., and Fink, G.R. (2005). Intragenic tandem repeats generate functional variability. Nat. Genet. *37*, 986–990.

Wang, Q., Udayakumar, T.S., Vasaitis, T.S., Brodie, A.M., and Fondell, J.D. (2004). Mechanistic relationship between androgen receptor polyglutamine tract truncation and androgen-dependent transcriptional hyperactivity in prostate cancer cells. J. Biol. Chem. *279*, 17319–17328.

Weaver, R.F. and Weissmann, C. (1979). Mapping of RNA by a modification of the Berk-Sharp procedure: the 5′ termini of 15 S beta-globin mRNA precursor and mature 10 s beta-globin mRNA have identical map coordinates. Nucleic Acids Res. *7*, 1175–1193.

Webb, D.M. and Zhang, J. (2005). FoxP2 in song-learning birds and vocal-learning mammals. J. Hered. *96*, 212–216.

Westin, G., Gerster, T., Müller, M., Schaffner, G., and Schaffner, W. (1987). OVEC, a versatile system to study transcription in mammalian cells and cell-free extracts. Nucleic Acids Res. *15*, 6787–6798.

Wigler, M., Pellicer, A., Silverstein, S., and Axel, R. (1978). Biochemical transfer of single-copy eucaryotic genes using total cellular DNA as donor. Cell *14*, 725–731.

Zhang, B., Egli, D., Georgiev, O., and Schaffner, W. (2001). The Drosophila homolog of mammalian zinc finger factor MTF-1 activates transcription in response to heavy metals. Mol. Cell. Biol. *21*, 4505–4514.