Supplemental Information and Figures for:

Biophysical and Functional Analyses Suggest That Adenovirus E4-ORF3 Protein Requires Higher-order Multimerization to Function against Promyelocytic Leukemia Protein Nuclear Bodies

Vadim Patsalo, Mark A. Yondola, Bowu Luan, Ilana Shoshani, Caroline Kisker, David F. Green, Daniel P. Raleigh, and Patrick Hearing

Plasmid Constructs and Purification of the E4-ORF3 Protein –The WT Ad5 E4-ORF3 reading frame was cloned by PCR into bacterial expression plasmids pET-15b (Novagen), pET-duet (Novagen), and pProEx-HTb (Invitrogen) while the non-functional E4-ORF3 mutant, L103A, was cloned into pProEx-HTb for protein expression in *E. coli*. These constructs expressed E4-ORF3 with a six-histidine-residue (His_6)[1] tag fused to the N-terminus and differed only in the protease cleavage sites used for removal of the His_6-tag. WT E4-ORF3 protein also was expressed using the IMPACT-CN system (New England Biolabs) where E4-ORF3 was expressed as a bipartite fusion protein containing a self-splicing intein and a chitin binding domain. Recombinant proteins expressed from each of these vectors behaved identically. The WT and L103A E4-ORF3 sequences fused to the His_6 tag and TEV protease cleavage site were excised from the pProEx-HTb vectors and inserted into mammalian expression plasmid pcDNA3 (Invitrogen).

Recombinant E4-ORF3 protein was expressed using the BL21 DE3 Rosetta strain of *E. coli* (Novagen). Cultures were induced with 0.5mM IPTG[2] at an OD_{600} = 0.6-0.8 after cooling to room temperature. Inductions were performed for 20 h at 25 ^{o}C and cell pellets were frozen prior to protein purification. In order to generate approximately 5 mg of pure WT E4-ORF3 protein (5 mg/mL), cell pellets from a 2 liter culture were lysed in 300 mL of lysis buffer (10 mM Tris-HCl, 10 mM βME[3], pH 7.5). E4-ORF3 contains three cysteine residues and reducing agent was included, as indicated, to prevent disulfide bond formation. Sonication was performed at a 50% duty cycle at setting five using a Branson sonicator. Two rounds of sonication lasting two min. each were performed on 40 mL aliquots maintained on ice throughout the process. Protease inhibitors were added to all lysis and resuspension buffers and included: PMSF, benzamidine, pepstatin, leupeptin, and aprotinin. Lysates were clarified by centrifugation at 20,000 x g for 30 min. The soluble fraction was decanted and NaCl was added to a final concentration of 100 mM. The sample was then agitated at 4 ^{o}C for 20 minutes. The E4-ORF3 protein was precipitated by centrifugation at 10,000 x g for 15 min. and the precipitate was resuspended in an equal volume of 10 mM Tris-HCl, 10 mM βME, pH 8.5 by mechanical disruption followed by agitation at 4 ^{o}C for 20 min. The sample was centrifuged at 13,000 x g for 20 min to remove any protein that did not resuspend. The soluble fraction was recirculated over a Ni^{2+}-NTA column overnight, the column was washed with 5 volumes of 10 mM Tris-HCl, 10 mM βME, 20 mM imidazole, pH 8.5. E4-ORF3 protein was eluted with the same buffer containing 200 mM imidazole. Individual fractions were assayed for protein by Bradford assay (Bio-Rad) and sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). E4-ORF3-containing fractions were pooled and concentrated via a Centricon Bio-Max 5 filter (5000 molecular weight cutoff) with a buffer change to the final E4-ORF3 storage buffer (10mM Tris-HCl, 5 mM βME, pH 8.5). Size exclusion chromatography, the ultimate purification step for the E4-ORF3 WT and L103A proteins, was performed using a Superdex 200 16/60 column (GE Healthcare) in

---

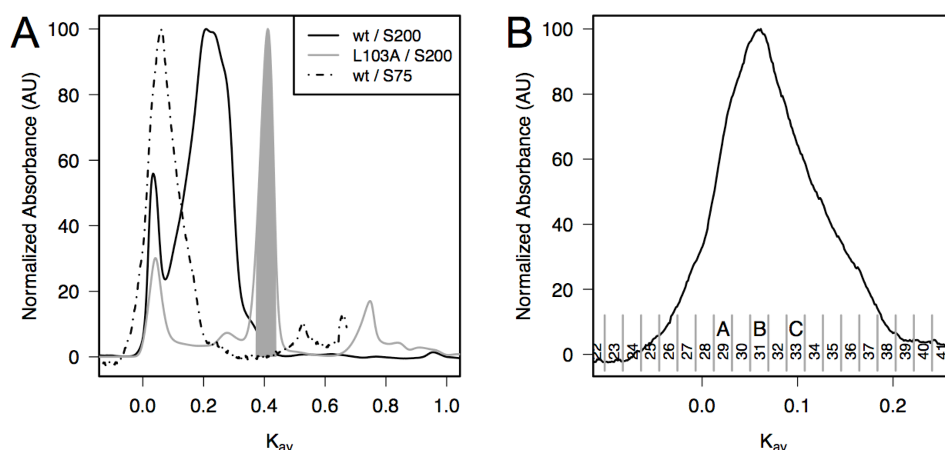
[1] His_6, six-histidine-residue

[2] IPTG, isopropyl-beta-D-thiogalactopyranoside

[3] βME, 2-mercaptoethanol

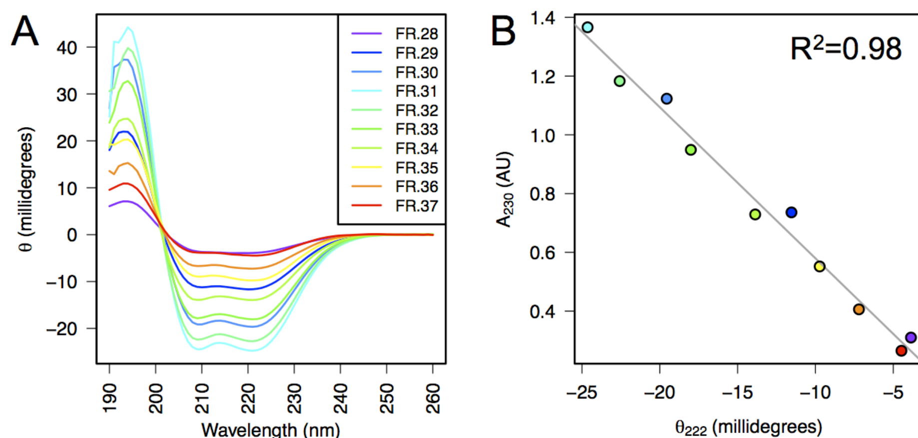
Buffer A (10 mM Tris-HCl, 0.5 mM TCEP[4], pH 8.5). A separate purification of the WT E4-ORF3 protein used a Superdex 75 26/60 column to take advantage of the increased loading capacity. 30 uL aliquots of peak fractions of E4-ORF3 proteins were flash frozen in liquid N_2 and stored at –80 ^oC.

For *E. coli* co-expression studies, the gene for WT E4-ORF3 was cloned into pET-28a, where the protein product lacked a His_6-affinity tag. BL21 (DE3) cells were transformed with pProEx-L103A plasmid (Amp), and chemically-competent cell stocks were made. These cell stocks were then transformed with pET-28a-WT (Kan), and plated on LB plates containing both antibiotics. Protein expression was performed as indicated above, except that both ampicillin and kanamycin were included in the culture medium. The salting out purification step was skipped to avoid coprecipitating the two proteins; we were interested in visualizing protein complexes formed between the two variants during protein expression, rather than those that might arise as a consequence of salt precipitation.

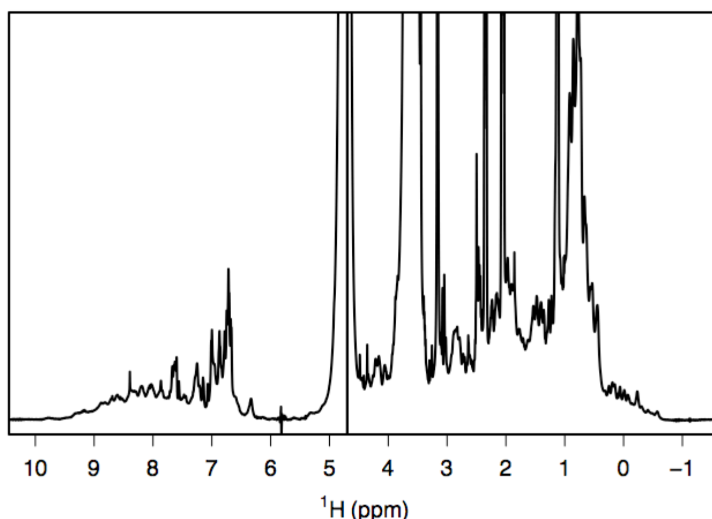


Supplemental Figure 1. Gel filtration profiles of E4-ORF3 WT and L103A proteins. *A*, Size exclusion chromatography was performed with the E4-ORF3 WT and L103A proteins using Superdex 200 and Superdex 75 columns**.** The WT E4-ORF3 protein (solid black line) eluted as a complex mixture. A large fraction of the E4-ORF3 L103A protein (gray line) eluted as a monodisperse peak (shaded) at K_{av} of 0.4 (~75 mL) on a Superdex 200 16/60 column. This peak was collected and characterized for secondary structure content and oligomerization state. The elution times are normalized by column bed and void volumes. K_{av} of 0 is the void volume of the column. *B*, The WT E4-ORF3 protein was fractionated using a Superdex 75 26/60 column (dotted line in panel *A*). Fraction numbers (22–41) and designations (A, B, and C) are indicated. The protein eluted as an asymmetric peak near the column void volume. Superdex 75 is less effective at separating large particles (~200 kDa) than Superdex 200. All fractions were collected and analyzed for secondary structure content using CD (Supplemental Fig. 3*A*).

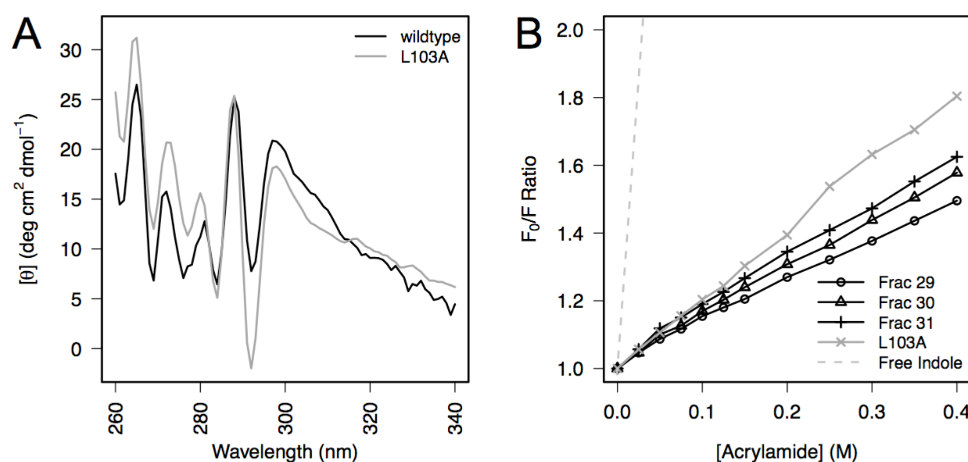
[4] TCEP, Tris(2-carboxyethyl)phosphine)



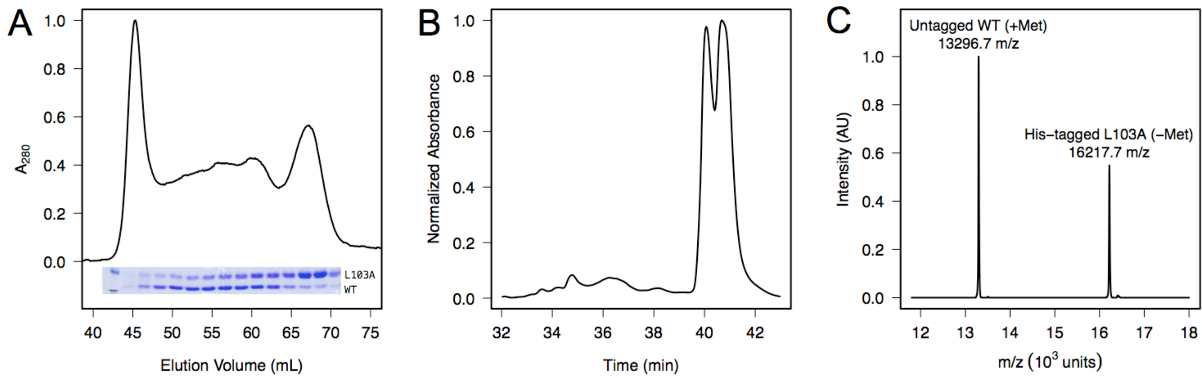
Supplemental Figure 2. CD suggests that species present in different size exclusion chromatography fractions of WT E4-ORF3 contains similar secondary structure. *A*, Far-UV CD spectra of size exclusion fractions of WT E4-ORF3 chromatographed on a Superdex 75 26/20 column (Supplemental Fig. 1*B*). The spectra are given as raw measured ellipticity (are not concentration-normalized). *B*, Helical content (ellipticity at 222 nm) is presented as a function of protein concentration, given by A_{230}. Each point represents a distinct sample for which we measured A_{230} and collected a CD spectrum. The colors used correspond to those in panel *A*. Helical content is directly proportional to protein concentration, indicating that all fractions possess similar secondary structure, and that the secondary structure of recombinant E4-ORF3 is not significantly altered by oligomeric state. A linear least-squares fit is shown as a gray line.



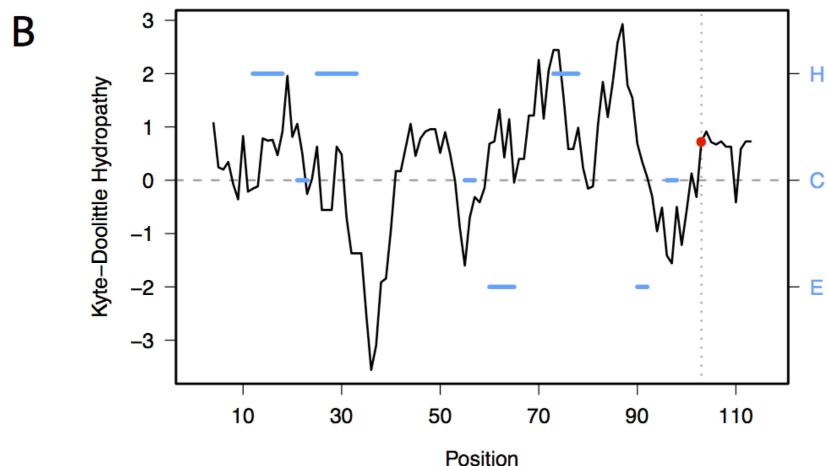
Supplemental Figure 3. One-dimensional NMR spectrum of L103A. Proton chemical shift is shown on the x-axis. The spectrum reveals the presence of ring current-shifted methyl resonances below 0 ppm, indicating close tertiary packing of methyl group-containing sidechains against aromatic groups. In addition, the amide region of the spectrum is well-dispersed. This suggests that the protein interior is tightly packed, and is not a molten globule.



Supplemental Figure 4. Probing tertiary structure in E4-ORF3. *A*, Near-UV CD spectra collected for WT and L103A variants of E4-ORF3. The spectra revealed the presence of narrow bands for both variants, indicating a similar environment for the protein aromatic sidechains. The two variants contained concordant bands arising from dipole orientation and electronic structure of residues which absorb in this region (Trp, Tyr, Phe). The spectral bands seen for WT and L103A were not identical in intensity, but displayed remarkable agreement in their position and linewidth. *B*, Collisional quenching of Trp fluorescence by acrylamide revealed that Trp35, the unique Trp residue in E4-ORF3 is sequestered from solvent, as was suggested by the fluorescence emission experiments. Fluorescence intensity diminishment (F_0/F) is plotted *versus* concentration of acrylamide. Stern-Volmer constants of approximately 1 M^{-1} were derived both several samples of WT E4-ORF3 and for L103A. The dashed line indicates quenching of an indole group, representing maximal solvent accessibility; the Stern-Volmer constant of quenching the free indole is 14 M^{-1}.



Supplemental Figure 5. Recombinant co-expression of WT and L103A variants of E4-ORF3. *A*, A Superdex 200 gel filtration profile of the Ni-affinity purified co-expression is shown. Individual fractions were probed for protein content by SDS-PAGE (overlaid.) Material eluting near the void volume of 45 mL was determined to be contaminating nucleic acids using SDS-PAGE and ethidium bromine fluorescence. Note the presence of both proteins in all examined gel filtration fractions. *B*, Gel filtration fractions from panel *A* were pooled and loaded onto an analytical C_4 HPLC column. The proteins were eluted with a linear gradient of 0–90% isopropanol, 0.1% trifluoroacetic acid. While the peaks were not resolved enough for accurate quantification, we approximate (using HPLC and SDS-PAGE from panel *A*) that a nearly-equal amount of both proteins was present. *C*, MALDI/TOF mass spectrometry was used to confirm the presence and identity of both proteins. For His_6-tagged L103A, the calculated MW is 16221.5 Da (m/z of 16217.7 observed), after removal of N-terminal leading methionine. For untagged WT E4-ORF3, the calculated MW is 13298.4 Da (m/z of 13296.7 observed); for this protein product, the N-terminal methionine was retained.



Supplemental Figure 6. Analysis of E4-ORF3 hydrophobicity. *A*, The amino-acid sequence of WT E4-ORF3 (Uniprot accession code P04489) is provided. *B*, The Kyte-Doolittle hydrophobicity plot of WT E4-ORF3 using a window of 7 residues. Negative and positive scores indicate polar and non-polar regions, respectively. The site of the L103A mutation is highlighted with a red point and a vertical line. S econdary-structure prediction of PSI-PRED is shown in blue lines, and indicate helical, coil, and extended regions (H, C, and E, respectively). Only regions with the prediction is of high confidence (>6) are shown.