Supporting Information

Kang et al. 10.1073/pnas.1211078109

SI Text

SI Materials and Methods. ***Structural bioinformatics procedure for prediction of filament-specific cation binding sites***

I. Selected actin structures.
The following modifications were performed before using these files to predict cation-binding sites: For 1J6Z [Protein Data Bank ID number, (1)], all non-protein atoms (including water, tetramethylrhodamine-5-maleimide, and six Ca^{2+} ions) were stripped from the file except for those belonging to the bound ADP molecule. This allowed for a fair comparison with 3MFP (2), which only contains a bound ADP molecule in addition to the protein atoms. Besides the F and G conformations, the only difference between the actin monomer files submitted for cation-binding site prediction was that 1J6Z contains three fewer residues (Asp-Glu-Asp) at the N terminus than 3MFP. 1J6Z is also missing the last three residues (Lys-Cys-Phe), which were not resolved in the structure. Therefore predicted cation-binding sites near these additional residues in the F-actin monomer (3MFP) were ignored for the purpose of comparing with the G-actin monomer (1J6Z). Neither 1J6Z nor 3MFP contain the N-terminal Met and Cys residues indicated in the representative mammalian skeletal muscle actin sequence from the National Center for Biotechnology Information (NCBI) protein database with the name "ACTS_HUMAN." Both 1J6Z and 3MFP have an identical protein residue-numbering scheme, which ends with 375 (two shorter than ACTS_HUMAN).

II. WebFEATURE (3) Ca^{2+} binding site prediction model: The model was trained using structures deposited in the Protein Data Bank (PDB) that contain bound Ca^{2+} ions (4, 5), and uses a vector of 66 structural characteristics within a 7-Å radius to score putative Ca^{2+} binding sites. We compared the positions of Ca^{2+} ions bound to the G-actin monomer (1J6Z) to the top-predicted Ca^{2+} binding sites. Out of six bound Ca^{2+} cations in 1J6Z, WebFEATURE predicts two of the crystallographic Ca^{2+} sites (to within 2.0 Å, including the nucleotide-associated site), the only two that have two or more acidic residues within 7 Å. The other four all have either one or zero acidic residues within 7 Å, which likely indicates a strong preference for the prediction algorithm to include acidic residues. Also, the primary coordination shell of the other four crystallographic Ca^{2+} is mostly water oxygens (two cases have only one protein atom in the first shell), which likely do not constitute "strongly bound" Ca^{2+}.

III. Detailed subtraction process using VMD and custom Python scripts.

1. The G-actin monomer (1J6Z) was aligned with the F-actin monomer (3MFP), using the "measure fit" command within the VMD Tcl scripting interface. Only backbone atoms for residues 4 to 372 (both 1J6Z and 3MFP numbering) were used to calculate the best-fit rotation/translation matrix. This matrix was then applied to all the G-actin atoms including the predicted cation-binding sites (saved as single pseudo-atom residues named "HIT" in the WebFEATURE output). This resulted in cation-binding site HIT predictions from the G-actin structure that were moved into close proximity with HIT predictions from the F-actin structure, based on structurally equivalent protein residue positions between the G- and F-actin structures.
2. We exhaustively considered each of the HIT predictions for the F-actin monomer and compared its position to all HIT predictions for the G-actin monomer. If any HIT prediction for the G-actin monomer was within 10 Å of any HIT prediction for the F-actin monomer, those F-actin HITs were ignored. This distance criterion is conservative in terms of identifying discrete F-actin-specific binding sites. The results did not change significantly down to a distance of 6 Å (larger numbers of sites with lower summed prediction scores in step IV below), while greater than 10 Å starts to delete F-actin specific predictions that are not near any G-actin predicted sites (or F-actin polymer vs. F-actin monomer below). In the limit of an overly large distance cut-off for comparisons, no F-actin-specific site predictions would survive this step.
3. Any "HIT" predictions that survived step III.2 above constituted potential cation binding sites that were predicted to stabilize the F conformation over the G conformation of an actin monomer.
4. We repeated steps III.1–3 above, comparing the F-actin monomer (3MFP) HIT predictions to the F-actin 5-mer (3MFP "biological assembly"). This was achieved by iteratively aligning the F-actin monomer (including HIT prediction pseudo-atoms) with each subunit of the F-actin 5-mer following the procedure in step III.1 above and saving the new monomer HIT coordinates for each of the five subunit alignment steps. Then step III.2 above was repeated by looping over all of the F-actin 5-mer hit predictions and ignoring any that were found within 10 Å of any F-actin monomer HIT prediction.

IV. Predicting and ranking discrete binding sites: Criteria for grouping individual HITs using custom Python scripts.

1. We considered only those HIT predictions that were not eliminated from the "subtraction" steps described above.
2. We individually grouped HIT predictions for either the F-actin monomer (after subtracting G-actin monomer HITs), or the F-actin polymer (after subtracting F-actin monomer HITs) as follows:

   a. We started with any HIT prediction and formed a new Group "A" (or B, C, etc.)
   b. We joined all HIT predictions within 4.5 Å of the starting HIT prediction into Group A (or B, C, etc.). 4.5 Å was chosen as a conservative estimate of the maximum distance between any two neighboring residues that contribute to a single typical cation binding site, or conversely the minimum distance between neighboring cations occupying two adjacent discrete cation-binding sites.
   c. We joined any HIT prediction within 4.5 Å of any other member of Group A (or B, C, etc.) and repeated this step until no other HIT predictions occured within 4.5 Å of any member of Group A.
   d. We selected a new HIT prediction that was not a member of Group A and repeated steps 2a to 2c above with each new Group (B, C, etc.). We repeated this step until all HIT predictions from step 1 above were assigned to a Group (including Groups containing only 1 HIT prediction).
   e. For each Group, we calculated its total prediction score based on the sum of all individual HIT prediction scores within that Group
   f. For each Group, we calculated its weight-averaged position

as follows:

$$r_{\text{Avg.}} = \frac{\sum_{i=1}^{N} w_i(r_i)}{\sum_{i=1}^{N} w_i} \qquad \text{[S1]}$$

where $r_{\text{Avg.}}$ is the weight-averaged position of the group, r_i is the position vector (x,y,z coordinates) of the ith HIT prediction within the Group, w_i is the WebFEATURE prediction score of the ith HIT within the Group consisting of N WebFEATURE HITs.

Actin residue side-chain pKa predictions. We used the F-actin 5-mer model deposited in the PDB by Namba and co-workers (2) as input for residue side-chain pK_a predictions to determine if specific residues show a shift in going from G- to F-actin in their predicted microscopic pK_a that would put them into a range where changing the solution pH from >7.0 to <7.0 would promote protonation of this specific residue. We chose to use the web-based implementation of the PROPKA software (6).

We found that the predicted pK_a of only one protonatable side-chain (Asp288) changes from near its model value (3.9) to a value in this "interesting range" (7.1) in going from the G-actin structure (1J6Z) to any protomer of the F-actin model (3MFP 5-mer) that has an additional subunit toward the pointed end making longitudinal close-contacts with this residue (three out of the five subunits) while the other two subunits have a predicted pK_a near the model value in the absence of inter-subunit contacts from the polymer lattice.

1. Otterbein LR, Graceffa P, Dominguez R (2001) The crystal structure of uncomplexed actin in the ADP state. *Science* 293:708–711.
2. Fujii T, Iwane AH, Yanagida T, Namba K (2010) Direct visualization of secondary structures of F-actin by electron cryomicroscopy. *Nature* 467:724–728.
3. Liang MP, Banatao DR, Klein TE, Brutlag DL, Altman RB (2003) WebFEATURE: An interactive web tool for identifying and visualizing functional sites on macromolecular structures. *Nucleic Acids Res* 31:3324–3327.
4. Wei L, Altman RB (1998) Recognizing protein binding sites using statistical descriptions of their 3D environments. *Pac Symp Biocomput* 497–508.
5. Liu T, Altman RB (2009) Prediction of calcium-binding sites by combining loop-modeling with machine learning. *BMC Struct Biol* 9:72.
6. Bas DC, Rogers DM, Jensen JH (2008) Very fast prediction and rationalization of pK(a) values for protein-ligand complexes. *Proteins* 73:765–783.

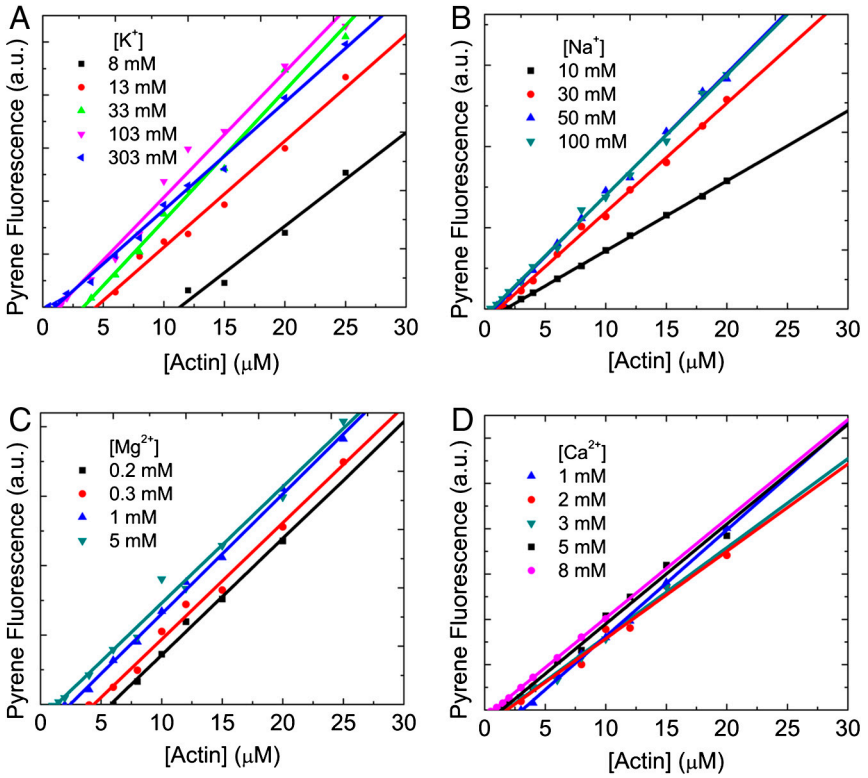

Fig. S1. Salt dependence of ADP-actin polymerization. Rabbit skeletal muscle actin (5% pyrene labeled) polymerized with (*A*) [KCl], (*B*) [NaCl], (*C*) [$MgCl_2$], or (*D*) [$CaCl_2$]. The solid lines represent the best fits of the data to a linear function, yielding C_c from the x-intercept value.

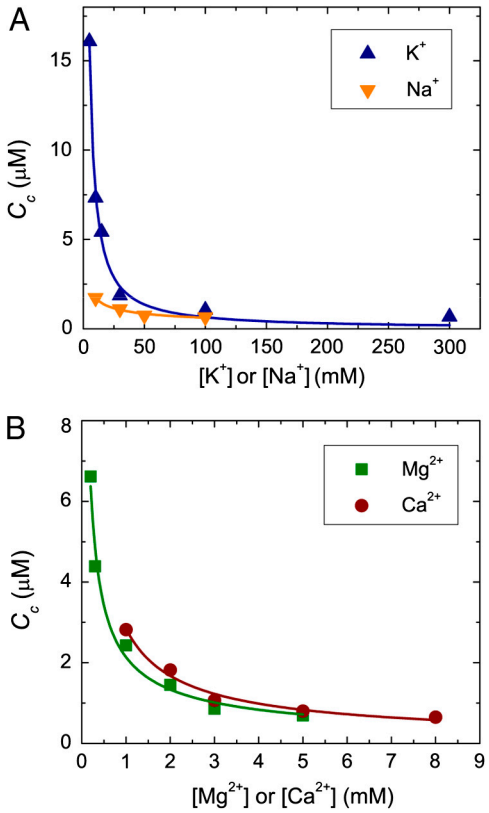

Fig. S2. Cation dependence of the ADP-actin critical concentration (C_c). Rabbit skeletal muscle actin (5% pyrene labeled) polymerized with (*A*) monovalent cations or (*B*) divalent cations. The solid lines through the data are for presentation.

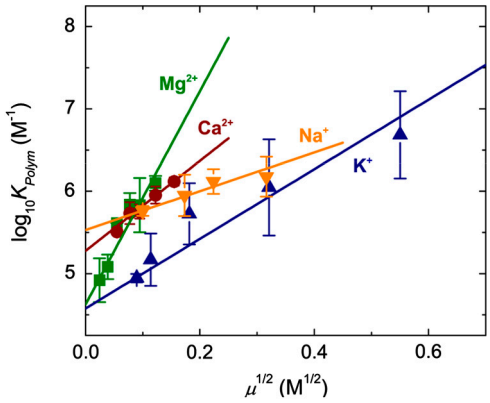

Fig. S3. Ionic strength (μ)-dependence of the ADP-actin K_{polym}. Rabbit skeletal muscle actin (5% pyrene labeled) polymerized with the indicated salts. The solid lines represent best fits of the data to $\log K_{polym} = \log K_0 + m\sqrt{\mu}$ (1).

1 Record MT, Lohman TM, deHaseth PL (1976) Ion effects on ligand-nucleic acid interactions. *J Mol Biol* 107:145–158.

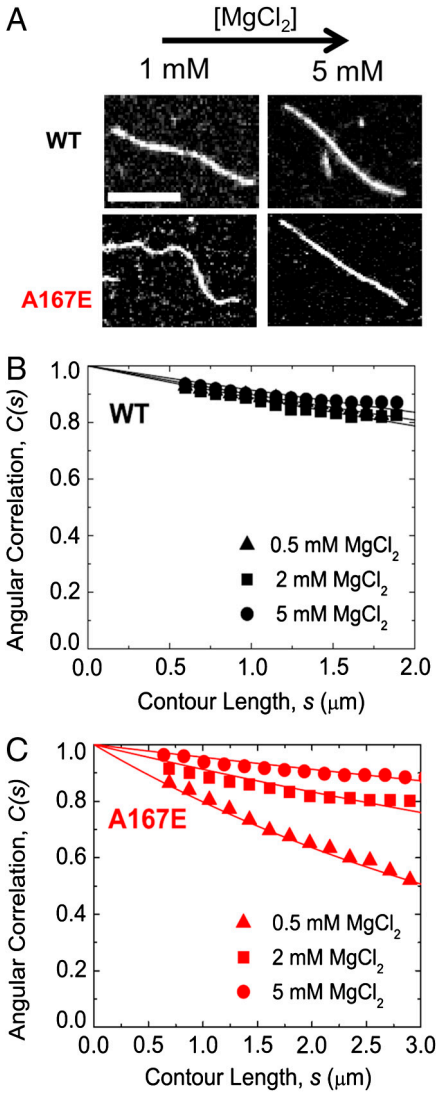

Fig. S4. Cation-dependence of actin filament bending rigidity. (*A*) Representative images of wild type (wt) and A167E mutant yeast actin filaments at $[Mg^{2+}] = 1$ and 5 mM (scale bar = 3 μm). $[Mg^{2+}]$ dependence of cosine angular correlation function $[C(s)]$ of (*B*) wt and (*C*) A167E yeast actin filaments. The solid lines through data represent the best fits to Eq. **4**, yielding filament bending persistence lengths (L_p).

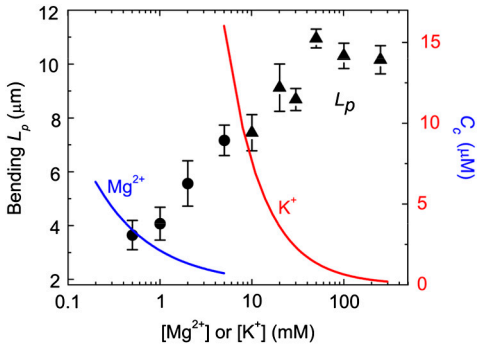

Fig. S5. Cation-dependence of actin filament critical concentration and bending rigidity. Rabbit skeletal muscle actin (5% pyrene or Alexa 488 labeled) polymerized with the indicated salts. The solid lines represent C_c values and the symbols represent L_p values.