Structural studies on Full-length Talin1 reveal a compact auto-inhibited dimer: Implications for talin activation

Benjamin T Goult[1], Xiao-Ping Xu[2], Alexandre R Gingras[1*], Mark Swift[2], Bipin Patel[1], Neil Bate[1], Petra M Kopp[1], Igor L Barsukov, David R Critchley[1], Niels Volkmann[2], Dorit Hanein[2]

SUPPLEMENTARY FIGURES

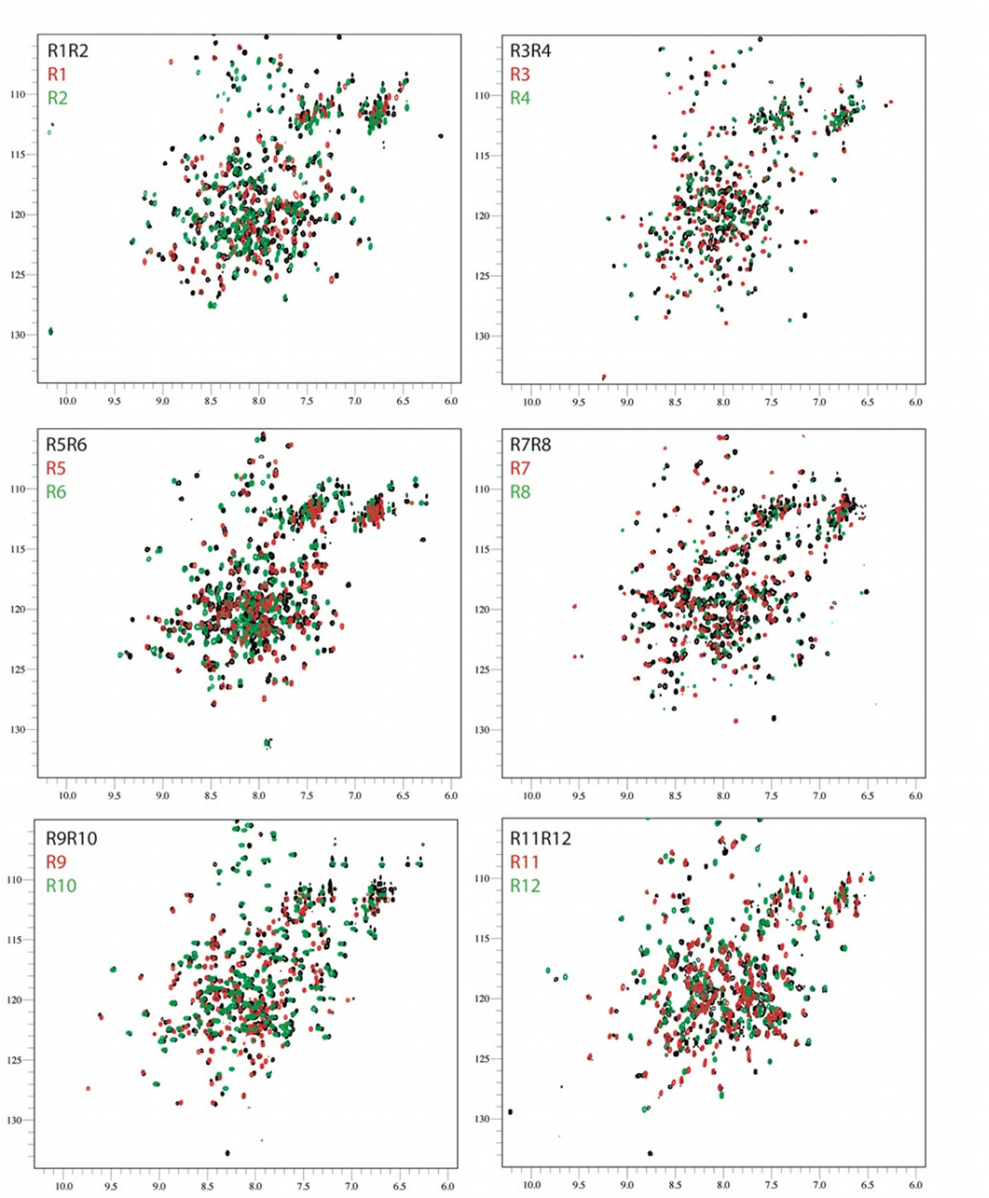

Figure S1. $^{1}H,^{15}N$-HSQC spectra of ^{15}N-labeled talin1 double domains. NMR spectra of R1-R2 (482-911), R3-R4 (913-1206), R5-R6 (1046-1357), R7-R8 (1357-1653), R9-R10 (1655-1973), R11-R12 (1974-2294) (100 µM). The relevant individual domain spectra are overlaid (red and green) on the double domain spectra (black), in general, the spectra of the individual domains overlap well with those of the corresponding double domains, supporting the idea that domains are joined by short flexible linkers. However, large differences are observed in the spectra of R1 and R2 compared with the R1R2 double domain. This is due to the extensive interface between the two domains (Papagrigoriou et al., 2004). The spectrum of R11R12 is markedly broader than those of the other double domains since the two domains are linked by an almost continuous helix (Gingras et al., 2009).

```
CLUSTAL W (1.83) multiple sequence alignment

sp|P26039|TLN1_MOUSE  MVALSLKISIGNVVKTMQFEPSTMVYDACRMIRERIPEALAGPPNDFGLFLSDDDPKKGIWLEAGKALDYYMLRNGDTME 80
sp|P54939|TLN1_CHICK  MVALSLKISIGNVVKTMQFEPSTMVYDACRMIRERVPEAQMGQPNDFGLFLSDEDPKKGIWLEAGKALDYYMLRNGDTME 80
                      ************************************:***  * **********:************************

sp|P26039|TLN1_MOUSE  YRKKQRPLKIRMLDGTVKTIMVDDSKTVTDMLMTICARIGITNHDEYSLVRELMEEKKDEGTGTLRKDKTLLRDEKKMEK 160
sp|P54939|TLN1_CHICK  YKKKQRPLKIRMLDGTVKTVMVDDSKTVTDMLTTICARIGITNYDEYSLVREIMEEKKEEVTGTLKKDKTLLRDEKKMEK 160
                      *:*****************:************ **********:********:*****:* ****:***************

sp|P26039|TLN1_MOUSE  LKQKLHTDDELNWLDHGRTLREQGVEEHETLLLRRKFFYSDQNVDSRDPVQLNLLYVQARDDILNGSHPVSFDKACEFAG 240
sp|P54939|TLN1_CHICK  LKQKLHTDDELNWLDHGRTLREQGIDDNETLLLRRKFFYSDQNVDSRDPVQLNLLYVQARDDILNGSHPVSFDKACEFAG 240
                      ************************:::::************************************************

sp|P26039|TLN1_MOUSE  FQCQIQFGPHNEQKHKAGFLDLKDFLPKEYVKQKGERKIFQAHKNCGQMSEIEAKVRYVKLARSLKTYGVSFFLVKEKMK 320
sp|P54939|TLN1_CHICK  YQCQIQFGPHNEQKHKPGFLELKDFLPKEYIKQKGERKIFMAHKNCGNMSEIEAKVRYVKLARSLKTYGVSFFLVKEKMK 320
                      :***************.***:*********:********* ******:********************************

sp|P26039|TLN1_MOUSE  GKNKLVPRLLGITKECVMRVDEKTKEVIQEWSLTNIKRWAASPKSFTLDFGDYQDGYYSVQTTEGEQIAQLIAGYIDIIL 400
sp|P54939|TLN1_CHICK  GKNKLVPRLLGITKECVMRVDEKTKEVIQEWSLTNIKRWAASPKSFTLDFGDYQDGYYSVQTTEGEQIAQLIAGYIDIIL 400
                      ********************************************************************************

sp|P26039|TLN1_MOUSE  KKKKSKDHFGLEGDEESTMLEDSVSPKKSTVLQQQYNRVGKVEHGSVALPAIMRSGASGPENFQVGSMPPAQQQITSGQM 480
sp|P54939|TLN1_CHICK  KKKKSKDHFGLEGDEESTMLEDSVSPKKSTVLQQQFNRVGKAELGSVALPAIMRTGAGGPENFQVGTMPQAQMQITSGQM 480
                      ************************************:*****.* **********:**.********:** ** *******

sp|P26039|TLN1_MOUSE  HRGHMPPLTSAQQALTGTINSSMQAVQAAQATLDDFETLPPLGQDAASKAWRKNKMDESKHEIHSQVDAITAGTASVVNL 560
sp|P54939|TLN1_CHICK  HRGHMPPLTSAQQALTGTINSSMQAVNAAQATLDDFETLPPLGQDAASKAWRKNKMDESKHEIHSQADAITAGTASVVNL 560
                      **************************:****************************************.*************

sp|P26039|TLN1_MOUSE  TAGDPAETDYTAVGCAVTTISSNLTEMSRGVKLLAALLEDEGGNGRPLLQAAKGLAGAVSELLRSAQPASAEPRQNLLQA 640
sp|P54939|TLN1_CHICK  TAGDPADTDYTAVGCAVTTISSNLTEMSKGVKLLAALMEDEGGNGRQLLQAAKNLASAVSDLLKTAQPASAEPRQNLLQA 640
                      ******:**********************:********:******** ******.**.***:**::***************
                                  R1-R2
sp|P26039|TLN1_MOUSE  AGNVGQASGELLQQIGESDTDPHFQDVLMQLAKAVASAAAALVLKAKSVAQRTEDSGLQTQVIAAATQCALSTSQLVACT 720
sp|P54939|TLN1_CHICK  AGLVGQTSGELLQQIGESDTDPRFQDMLMQLAKAVASAAAALVLKAKNVAQKTEDSALQTQVIAAATQCALSTSQLVACT 720
                      ** ***:***************:***:*********************.***:****.**********************
                                                                                          R2-R3
sp|P26039|TLN1_MOUSE  KVVAPTISSPVCQEQLVEAGRLVAKAVEGCVSASQAATEDGQLLRGVGAAATAVTQALNELLQHVKAHATGAGPAGRYDQ 800
sp|P54939|TLN1_CHICK  KVVAPTISSPVCQEQLIEAGKLVAKSAEGCVEASKAATNDDQLLKQVGVAATAVTQALNDLLQHIKQHATGGQPIGRYDQ 800
                      ****************:***:****:.****.**:***:*.***: **.**********:****:* ****. * *****

sp|P26039|TLN1_MOUSE  ATDTILTVTENIFSSMGDAGEMVRQARILAQATSDLVNAIKADAEGESDLENSRKLLSAAKILADATAKMVEAAKGAAAH 880
sp|P54939|TLN1_CHICK  ATDTILNVTENIFSSMGDAGEMVRQARILAQATSDLVNAIKADAEGETDLENSRKLLSAAKILADATAKMVEAAKGAAAH 880
                      ******.*****************************************:******************************
                                        R3-R4
sp|P26039|TLN1_MOUSE  PDSEEQQQRLREAAEGLRMATNAAAQNAIKKKLVQRLEHAAKQAAASATQTIAAAQHAASAPKASAGPQPLLVQSCKAVA 960
sp|P54939|TLN1_CHICK  PDSEEQQQRLREAAEGLRMATNAAAQNAIKKKLVHKLEHAAKQAAASATQTIAAAQHAAASNKNPAAQQ-QLVQSCKVVA 959
                      **********************************::************************:: * .*. *  ******.**

sp|P26039|TLN1_MOUSE  EQIPLLVQGVRGSQAQPDSPSAQLALIAASQSFLQPGGKMVAAAKASVPTIQDQASAMQLSQCAKNLGTALAELRTAAQK 1040
sp|P54939|TLN1_CHICK  DQIPMLVQGVRGSQSQPDSPSAQLALIAASQNFLQPGGKMVAAAKATVPTITDQASAMQLSQCAKNLAAALAELRTAAQK 1039
                      :***:*********:****************.***************:**** ***************.:************
                       R4-R5
sp|P26039|TLN1_MOUSE  AQEACGPLEMDSALSVVQNLEKDLQEIKAAARDGKLKPLPGETMEKCTQDLGNSTKAVSSAIAKLLGEIAQGNENYAGIA 1120
sp|P54939|TLN1_CHICK  AQEACGPLEIDSALGLVQSLERDLKEAKAAARDGKLKPLPGETMEKCAQDLGNSTKAVTSAIAHLLGEVAQGNENYTGIA 1119
                      *********:****.:**.**:**:* ********************:**********:****:****:*******:***

sp|P26039|TLN1_MOUSE  ARDVAGGLRSLAQAARGVAALTSDPAVQAIVLDTASDVLDKASSLIEEAKKASGHPGDPESQQRLAQVAKAVTQALNRCV 1200
sp|P54939|TLN1_CHICK  AREVAQALRSLSQAARGVAANSSDPQAQNAMLECASDVMDKANNLIEEARKAVAKPGDPDSQQRLVQVAKAVSQALNRCV 1199
                      **:** .****:******** :*** .*  :*: ****:***..*****:** .:****:*****.******:*******
                       R5-R6
sp|P26039|TLN1_MOUSE  SCLPGQRDVDNALRAVGDASKRLLSDSLPPSTGTFQEAQSRLNEAAAGLNQAATELVQASRGTPQDLARASGRFGQDFST 1280
sp|P54939|TLN1_CHICK  NCLPGQRDVDAAIRMVGEASKRLLSDSFPPSNKTFQEAQSQLNRAAAGLNQSANELVQASRGTPQDLAKSSGKFGQDFNE 1279
                      .********* *:* **:**********:***. ********:**.*******:*.***************::**:******.
                                                                                          R6-R7
sp|P26039|TLN1_MOUSE  FLEAGVEMAGQAPSQEDRAQVVSNLKGISMSSSKLLLAAKALSTDPASPNLKSQLAAAARAVTDSINQLITMCTQQAPGQ 1360
sp|P54939|TLN1_CHICK  FLQAGVEMASLSPTKEDQAQVVSNLKSISMSSSKLLLAAKALSADPTSPNLKSQLAAAARAVTDSINQLITMCTQQAPGQ 1359
                      **:******. :*::**:********.****************:**:**********************************
```

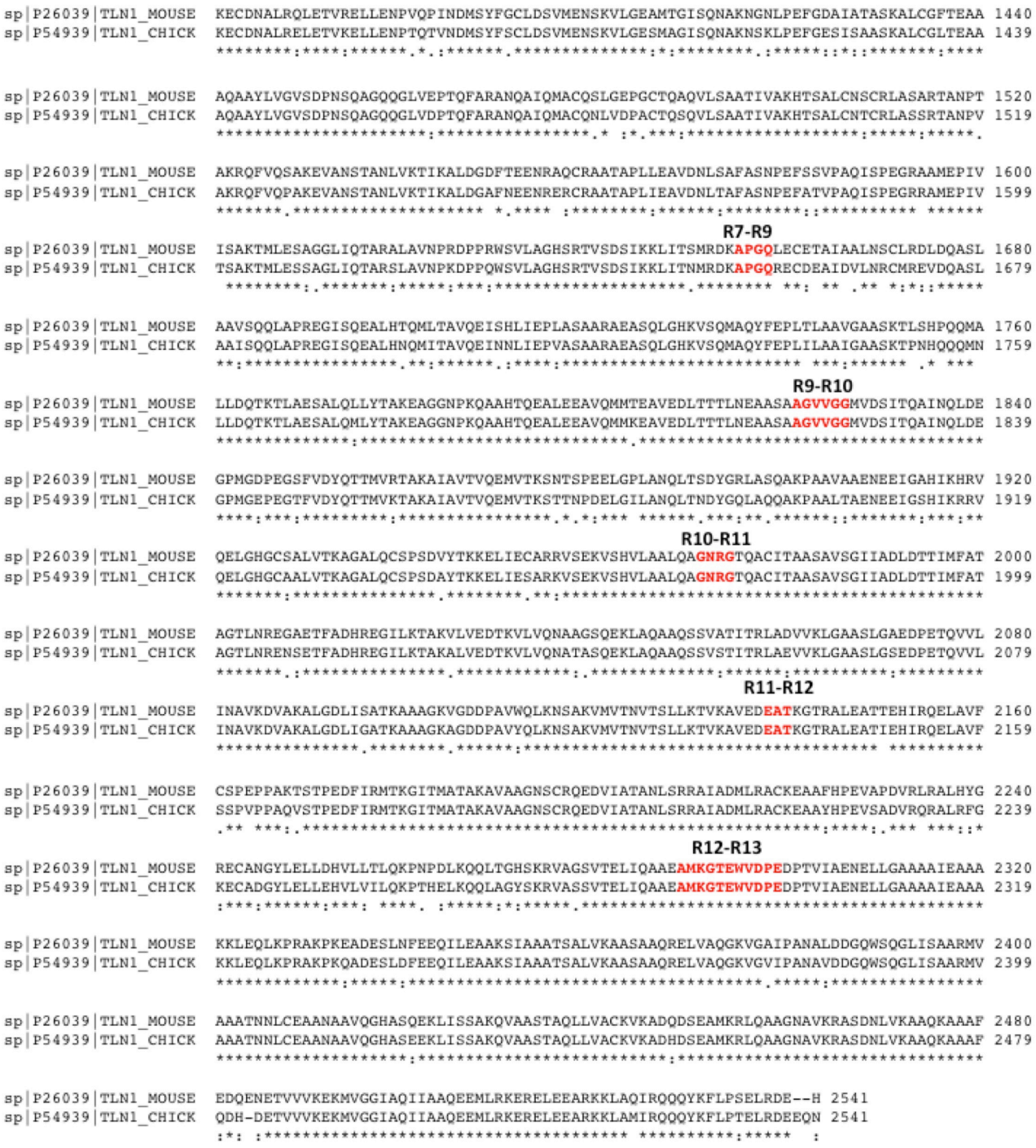

Figure S2. Sequence alignment of mouse and chicken talin 1. The linker regions between the rod domains R1-R13 are marked in red.

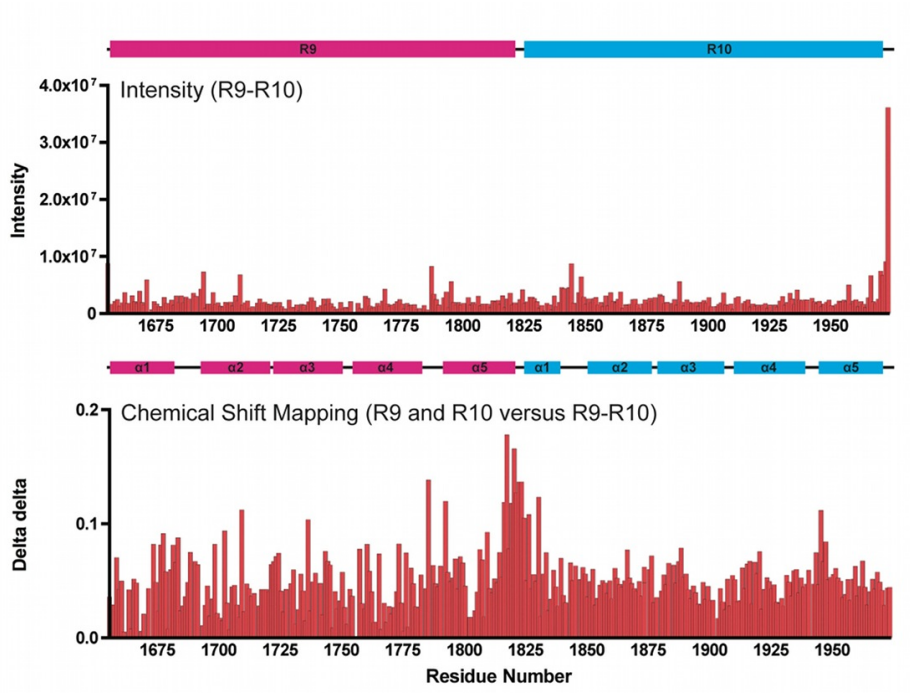

Figure S3. Analysis of the R9-R10 double domain. The R9-R10 double domain is separated by a short hinge. The domain boundaries are shown schematically above the plots, the secondary structure elements are shown schematically between the plots. Top. Plot of $^{1}H,^{15}N$-HSQC peak height as a function of residue number. The peak height is essentially uniform across. Bottom. Chemical shift mapping (CSM) of the individual R9 and R10 domains compared with the R9-R10 double domain. Small chemical shift changes are observed across the whole of the protein with the main changes clustering in the linker region between the two domains.

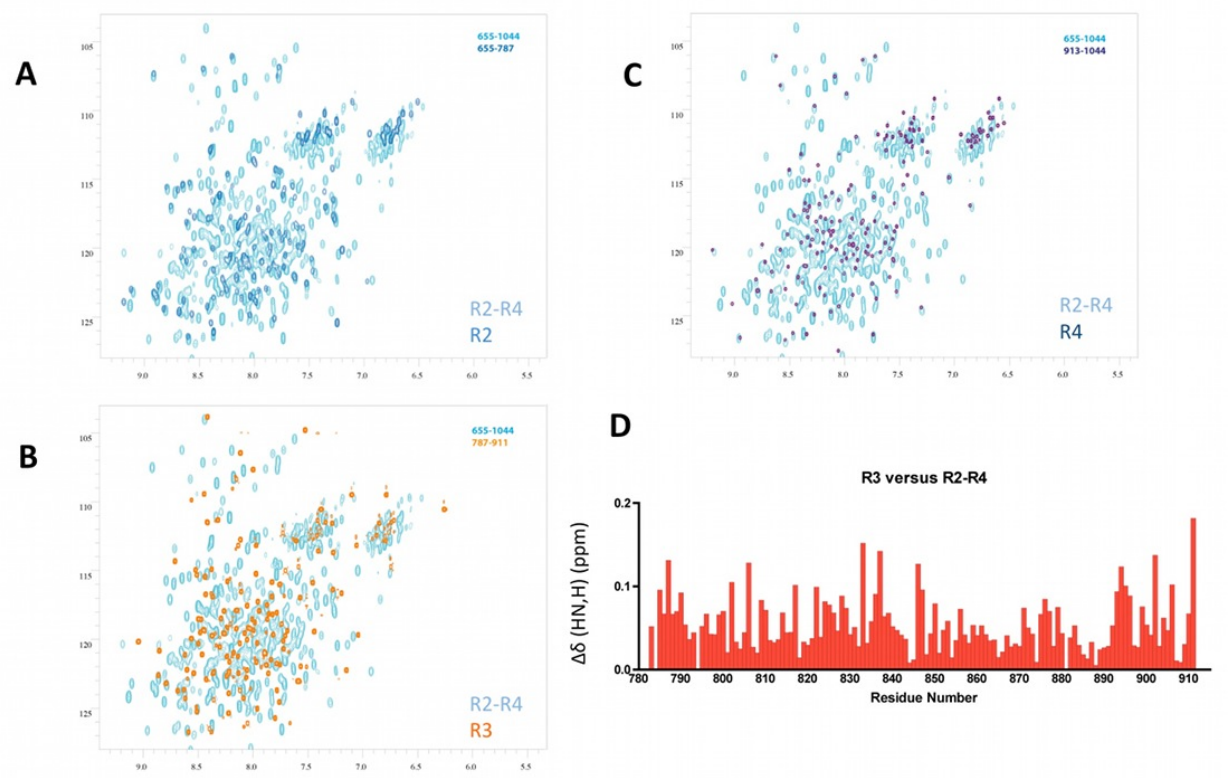

Figure S4. $^{1}H,^{15}N$-HSQC spectra of the three 4-helix bundles in the context of the triple domain R2-R3-R4. The triple domain R2-R3-R4 (sky blue) overlaid with (A) R2 (blue), (B) R3 (orange) and (C) R4 (navy blue). (D) Chemical shift mapping of R3 in the context of R2-R4.

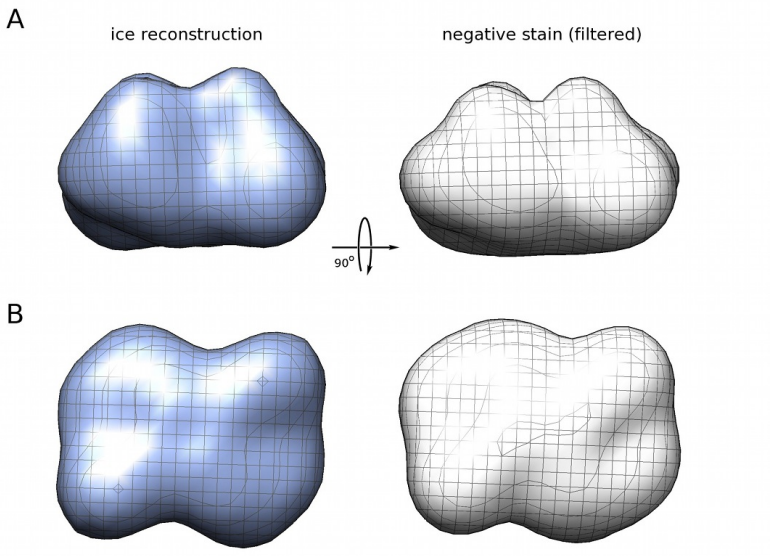

Figure S5. Cryo-EM reconstruction corresponds to negative stain reconstruction but is more limited in resolution. Comparison of the cryo-EM reconstruction at (3.5-nm resolution) (A) with the negative-stain reconstruction filtered to the same resolution (B). The correlation between the two maps is 0.97 ± 0.02 indicating that there is no difference between the two at the resolution of the cryo-EM reconstruction. Standard deviation was calculated from comparisons of reconstructions from different image processing runs.

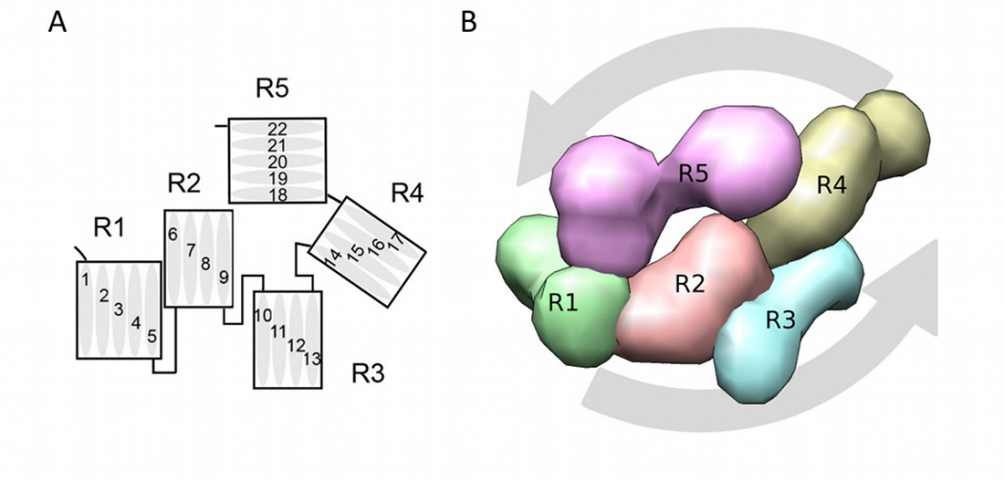

Figure S6. Fitting the N-terminal region of the rod. (A) Schematic representation of the topology and connectivity of the N-terminal region of the rod. (B) The same region shown as resulting from fitting the SAXS envelops of individual rod domains constraint by the inter-domain angles observed by SAXS into the EM reconstruction. The arrows indicate the overall directionality of the chain, which reverses at R4.

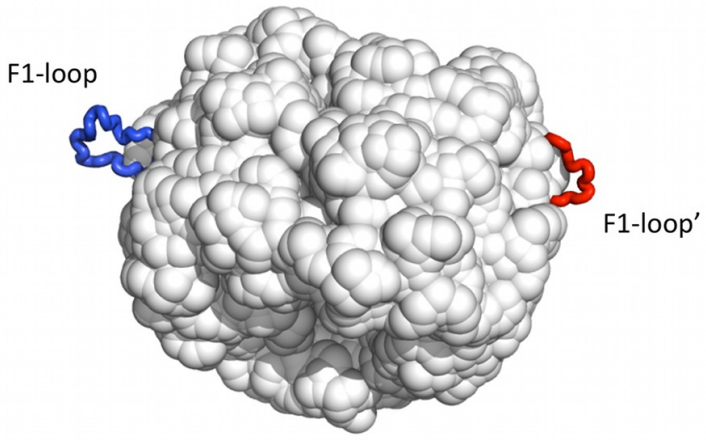

Figure S7. Structural modeling of the F1-loop onto the EM structure of talin. The F1-loops in the two talin subunits are large enough (>30 residues) to protrude from the interior of the molecule.

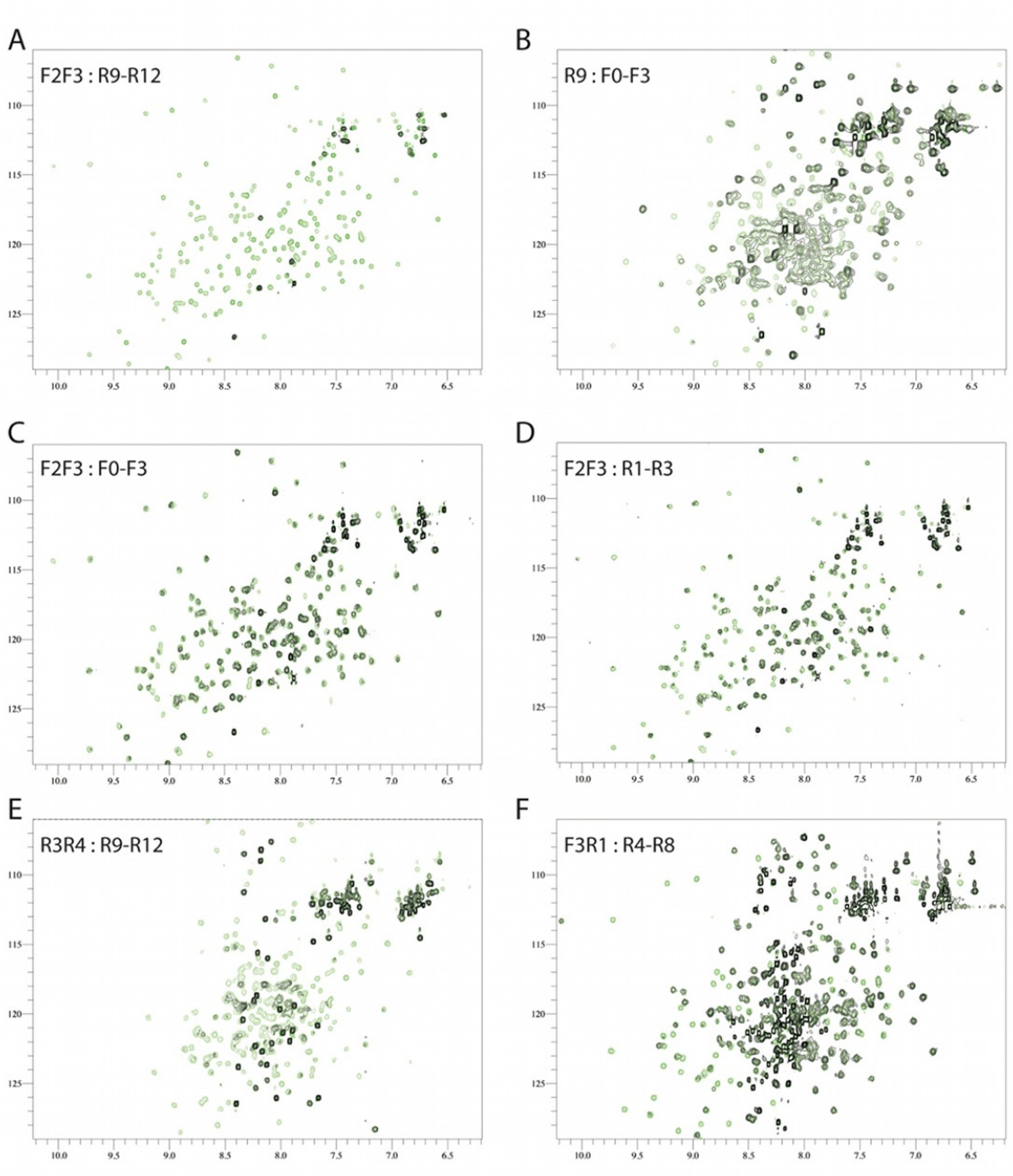

Figure S8. Identification of intra-molecular interactions in talin1 by NMR. $^{1}H,^{15}N$-HSQC spectra of 100 µM ^{15}N-labeled talin1 domain in the absence (green) or presence (black) of a 3-fold molar excess of unlabeled talin fragment. (A) ^{15}N-labelled F2F3 with unlabeled R9-R12; (B) ^{15}N-labelled R9 with unlabeled F0-F3; (C) ^{15}N-labelled F2F3 with unlabeled F0-F3; (D) ^{15}N-labelled F2F3 with unlabeled R1-R3; (E) ^{15}N-labelled R3R4 with unlabeled R9-R12; (F) ^{15}N-labelled F3-R1 with unlabeled R4-R8.

SUPPLEMENTARY REFERENCES

Gingras, A.R., N. Bate, B.T. Goult, L. Hazelwood, I. Canestrelli, J.G. Grossmann, H. Liu, N.S. Putz, G.C. Roberts, N. Volkmann, D. Hanein, I.L. Barsukov, and D.R. Critchley, 2008. The structure of the C-terminal actin-binding domain of talin. EMBO J 27: 458-69.

Gingras, A.R., W.H. Ziegler, A.A. Bobkov, M.G. Joyce, D. Fasci, M. Himmel, S. Rothemund, A. Ritter, J.G. Grossmann, B. Patel, N. Bate, B.T. Goult, J. Emsley, I.L. Barsukov, G.C. Roberts, R.C. Liddington, M.H. Ginsberg, and D.R. Critchley, 2009. Structural determinants of integrin binding to the talin rod. J Biol Chem 284: 8866-76.

Papagrigoriou, E., A.R. Gingras, I.L. Barsukov, N. Bate, I.J. Fillingham, B. Patel, R. Frank, W.H. Ziegler, G.C. Roberts, D.R. Critchley, and J. Emsley, 2004. Activation of a vinculin-binding site in the talin rod involves rearrangement of a five-helix bundle. Embo J 23: 2942-2951.