PNAS
www.pnas.org

Supplementary Information for

Identifying Sequence Perturbations to an Intrinsically Disordered Protein that Determine Its Phase Separation Behavior

Benjamin S. Schuster[1,a,b], Gregory L. Dignon[1,c,d], Wai Shing Tang[e], Fleurie M. Kelley[b], Aishwarya Kanchi Ranganath[b], Craig N. Jahnke[f], Alison G. Simpkins[f], Roshan Mammen Regy[c], Daniel A. Hammer[a,f], Matthew C. Good[a,g], Jeetain Mittal[c,2]

[a] Department of Bioengineering, University of Pennsylvania, Philadelphia, PA, 19104
[b] Department of Chemical and Biochemical Engineering, Rutgers University, Piscataway, NJ 08854
[c] Department of Chemical and Biomolecular Engineering, Lehigh University, Bethlehem, PA 18015
[d] Laufer Center for Physical and Quantitative Biology, Stony Brook University, Stony Brook, NY, 11794
[e] Department of Physics, Brown University, Providence, RI, 02912
[f] Department of Chemical and Biomolecular Engineering, University of Pennsylvania, Philadelphia, PA, 19104
[g] Department of Cell and Developmental Biology, University of Pennsylvania, Philadelphia, PA, 19104
[1] These authors contributed equally to this work
[2] Corresponding author, jeetain@lehigh.edu

This PDF file includes:

Supplementary Text
Supplementary Materials and Methods
Figures S1 to S10
Legends for Movies S1 to S3
SI References

Other supplementary materials for this manuscript include the following:

Movie S1
Movie S2
Movie S3

1. Supplementary Text

1.1 Sequences used in in vitro work (including His tag and XhoI restriction site).

LAF-1 RGG WT [Highlighted residues: 21-30 (red); 82-91 (blue); 101-110 (green)]:
MESNQSNNGG SGNAALNRGG RYVPPHLRGG DGGAAAAASA GGDDRRGGAG GGGYRRGGGN SGGGGGGGYD
RGYNDNRDDR DNRGGSGGYG RDRNYEDRGY NGGGGGGGNR GYNNNRGGGG GGYNRQDRGD GGSSNFSRGG
YNNRDEGSDN RGSGRSYNND RRDNGGDGLE HHHHHH

LAF-1 RGG Δ21-30:
MESNQSNNGG SGNAALNRGG DGGAAAAASA GGDDRRGGAG GGGYRRGGGN SGGGGGGGYD RGYNDNRDDR
DNRGGSGGYG RDRNYEDRGY NGGGGGGGNR GYNNNRGGGG GGYNRQDRGD GGSSNFSRGG YNNRDEGSDN
RGSGRSYNND RRDNGGDGLE HHHHHH

LAF-1 RGG Δ82-91:
MESNQSNNGG SGNAALNRGG RYVPPHLRGG DGGAAAAASA GGDDRRGGAG GGGYRRGGGN SGGGGGGGYD
RGYNDNRDDR DDRNYEDRGY NGGGGGGGNR GYNNNRGGGG GGYNRQDRGD GGSSNFSRGG
YNNRDEGSDN RGSGRSYNND RRDNGGDGLE HHHHHH

LAF-1 RGG Δ101-110:
MESNQSNNGG SGNAALNRGG RYVPPHLRGG DGGAAAAASA GGDDRRGGAG GGGYRRGGGN SGGGGGGGYD
RGYNDNRDDR DNRGGSGGYG RDRNYEDRGY GYNNNRGGGG GGYNRQDRGD GGSSNFSRGG YNNRDEGSDN
RGSGRSYNND RRDNGGDGLE HHHHHH

LAF-1 RGG_{shuf}:
MNNSGDNDRG SGNYGLRNSF GDDGYGDNGN DEGNSGYRNR GLGGDRADEY GNSGGNGDNE AAPNASDRDD
AHYYDSDDYD DGGGGRGSGG AGGGGARGPG SNRAGRYGGG GRRGRGRGNG YNGNRSQRRR GGGRGRGNRG
YRVGNGNGQS GGRNSRGGGG GNGGANYGLE HHHHHH

LAF-1 $RGG_{shuf\text{-}pres}$:
MGGYGYGSSG DGGGDDYGDA **RYVPPHLR**GY GDGAGDDGGD NNDDSDDADR DYNGGLSGGA GGNSGGDGEN
GGDGNGRNNA RSGNNRGGNG NYRYFGANYG AGEGRGRNGQ GGEGSGNNRG GGGRYGRRRR QGSRGGRGSG
GNYGGNSNRS GRAGGRDNNA RNRRRNGSLE HHHHHH

LAF-1 $RGG_{shuf\text{-}control}$:
MSGGARNNRS GSNGGGHFSG GRGGYGYGDG QYRDAGGSAR RDDGNGGGGG RENGDNRYSY QLGNRGDYAN
NGSAGGDGGN GGRDGDRGES RNDNRDRGER GGSRRGNGVN AGGYAGRDGP NGNNGYDNGY GGYYGRNGSR
GADGGRDGNL GNARSNDGGD RRNGPGGSLE HHHHHH

LAF-1 RGG R to K:
MESNQSNNGG SGNAALNKGG KYVPPHLKGG DGGAAAAASA GGDDKKGGAG GGGYKKGGGN SGGGGGGGYD
KGYNDNKDDK DNKGGSGGYG KDKNYEDKGY NGGGGGGGNK GYNNNKGGGG GGYNKQDKGD GGSSNFSKGG
YNNKDEGSDN KGSGKSYNND KKDNGGDGLE HHHHHH

LAF-1 RGG Y to F:
MESNQSNNGG SGNAALNRGG RFVPPHLRGG DGGAAAAASA GGDDRRGGAG GGGFRRGGGN SGGGGGGGFD
RGFNDNRDDR DNRGGSGGFG RDRNWEDRGF NGGGGGGGNR GFNNNRGGGG GGFNRQDRGD GGSSNFSRGG
FNNRDEGSDN RGSGRSFNND RRDNGGDGLE HHHHHH

1.2 LAF-1 homologs used in sequence alignment (accession numbers)

LAF-1, C. elegans (NP_001254859.1):
MESNQSNNGG SGNAALNRGG RYVPPHLRGG DGGAAAAASA GGDDRRGGAG GGGYRRGGGN SGGGGGGGYD
RGYNDNRDDR DNRGGSGGYG RDRNYEDRGY NGGGGGGGNR GYNNNRGGGG GGYNRQDRGD GGSSNFSRGG
YNNRDEGSDN RGSGRSYNND RRDNGGDGQN TRWNNLDAPP SRGTSKWENR GARDERIEQE LFSGQLSGIN

FDKYEEIPVE ATGDDVPQPI SLFSDLSLHE WIEENIKTAG YDRPTPVQKY SIPALQGGRD LMSCAQTGSG
KTAAFLVPLV NAILQDGPDA VHRSVTSSGG RKKQYPSALV LSPTRELSLQ IFNESRKFAY RTPITSALLY
GGRENYKDQI HKLRLGCHIL IATPGRLIDV MDQGLIGMEG CRYLVLDEAD RMLDMGFEPQ IRQIVECNRM
PSKEERITAM FSATFPKEIQ LLAQDFLKEN YVFLAVGRVG STSENIMQKI VWVEEDEKRS YLMDLLDATG
DSSLTLVFVE TKRGASDLAY YLNRQNYEVV TIHGDLKQFE REKHLDLFRT GTAPILVATA VAARGLDIPN
VKHVINYDLP SDVDEYVHRI GRTGRVGNVG LATSFFNDKN RNIARELMDL IVEANQELPD WLEGMSGDMR
SGGGYRGRGG RGNGQRFGGR DHRYQGGSGN GGGGNGGGGG FGGGGQRSGG GGGFQSGGGG GRQQQQQQRA
QPQQDWWS

DDX3X, H. sapiens (NP_001180345.1):
MSHVAVENAL GLDQQFAGLD LNSSDNQSGG STASKGRYIP PHLRNREATK GFYDKDSSGW SSSKDKDAYS
SFGSRSDSRG KSSFFSDRGS GSRGRFDDRG RSDYDGIGSR GDRSGFGKFE RGGNSRWCDK SDEDDWSKPL
PPSERLEQEL FSGGNTGINF EKYDDIPVEA TGNNCPPHIE SFSDVEMGEI IMGNIELTRY TRPTPVQKHA
IPIIKEKRDL MACAQTGSGK TAAFLLPILS QIYSDGPGEA LRAMKENGRY GRRKQYPISL VLAPTRELAV
QIYEEARKFS YRSRVRPCVV YGGADIGQQI RDLERGCHLL VATPGRLVDM MERGKIGLDF CKYLVLDEAD
RMLDMGFEPQ IRRIVEQDTM PPKGVRHTMM FSATFPKEIQ MLARDFLDEY IFLAVGRVGS TSENITQKVV
WVEESDKRSF LLDLLNATGK DSLTLVFVET KKGADSLEDF LYHEGYACTS IHGDRSQRDR EEALHQFRSG
KSPILVATAV AARGLDISNV KHVINFDLPS DIEEYVHRIG RTGRVGNLGL ATSFFNERNI NITKDLLDLL
VEAKQEVPSW LENMAYEHHY KGSSRGRSKS RFSGGFGARD YRQSSGASSS SFSSSRASSS RSGGGGHGSS
RGFGGGGYGG FYNSDGYGGN YNSQGVDWWG N

DEAD box helicase 3b isoform 5X, D. rerio (XP_005168849.1)
MSHVAVENVH GLDQQLAALD LSSADVQGVT GRRYIPPHLR NKEAAKNDAP GGWDNGRSNG FVNGYHDGRD
NRMNGGSSFA GRGPIRSDRG GRGGFRGKST ASYNPIQPMQ SAGFGYDNKE AGGWNVPKDN AYNSFGGRSD
RGKSSFFNDR GSSSRGRYER GGFGGGGNSR WVEECRDEDW SKPLPPNERL EHELFSGSNT GINFEKYDDI
PVEATGHNGP QPIDRFHDLE MGEIIMGNIN LSRYTRPTPV QKHAIPIIKS KRDLMACAQT GSGKTAAFLL
PVLSQIYTDG PGEALQAAKN SAQENGKYGR RKQYPISLVL APTRELALQI YDEARKFSYR SHVRPCVVYG
GADIGQQIRD LERGCHLLVA TPGRLVDMME RGKIGLDYCN YLVLDEADRM LDMGFEPQIR RIVEQDTMPP
KGLRQTMMFS ATFPKEIQIL ARDFLEDYIF LAVGRVGSTS ENITQKVVWV EENDKRSFLL DLLNATGKDS
LTLVFVETKK GADALEDFLY REGYACTSIH GDRSQRDREE ALHQFRSGRC PILVATAVAA RGLDISNVKH
VINFDLPSDI EEYVHRIGRT GRVGNLGLAT SFFNDKNGNI TKDLLDILVE AKQEVPSWLE SLAYEHQHKS
SSRGRSKRFS GGFGARDYRQ NSSSGGGGFG GRGGRSTGGH GGNRGFGGGG FGNFYSSDGY GGNYSQVDWWG
N

DEAD-box helicase 3 X-linked L homeolog, X. laevis (NP_001080283.1)
MSHVAVENVL NLDQQFAGLD LNSADAESGV AGTKGRYIPP HLRNKEASRN DSNWDSGRGG NGYINGMQDD
RDGRMNGYDR GGYGSRGTGR SDRGFYDREN SGWNSGRDKD AYSSFGSRGE RGKGSLFNDK GSGSRRPDES
RPDGFDGVGN RGNNSSFGRF DRGNSRWSDE RNDEDDWSKP LAPNDRVEQE LFSGSNTGIN FEKYDDIPVD
ATGSNCPPHI ECFQDVDMGE IIMGNIQLTR YTRPTPVQKH AIPIIIGKRD LMACAQTGSG KTAAFLLPIL
SQIYADGPGD AMKHLKDNGR YGRRKQFPLS LVLAPTRELA VQIYEEARKF AYRSRVRPCV VYGGADIGQQ
IRDLERGCHL LVATPGRLVD MMERGKIGLD FCKYLVLDEA DRMLDMGFEP QIRRIVEQDT MPPKGVRQTM
MFSATFPKEI QILARDFLDE YIFLAVGRVG STSENITQKV VWVEEMDKRS FLLDLLNATG KDSLTLVFVE
TKKGADALED FLYHEGYACT SIHGDRSQRD REEALHQFRS GKCPILVATA VAARGLDISN VKHVINFDLP
SDIEEYVHRI GRTGRVGNLG LATSFFNEKN INITKDLLDL LVEAKQEVPS WLENMAYEQH HKSSSRGRSK
SRFSGGFGAK DYRQSSSAGS SFGSSRGGRS SGHGGSRAFG GGYGGFYNSD GYGGNYGGSS QVDWWGN

Belle isoform B, D. melanogaster (NP_001262379.1)
MSNAINQNGT GLEQQVAGLD LNGGSADYSG PITSKTSTNS VTGGVYVPPH LRGGGGNNNA ADAESQGQGQ
GQGQGFDSRS GNPRQETRDP QQSRGGGGEY RRGGGGGGRG FNRQSGDYGY GSGGGGRRGG GGRFEDNYNG
GEFDSRRGGD WNRSGGGGGG GRGFGRGPSY RGGGGGSGSN LNEQTAEDGQ AQQQQQPRND RWQEPERPAG
FDGSEGGQSA GGNRSYNNRG ERGGGGYNSR WKEGGGSNVD YTKLGARDER LEVELFGVGN TGINFDKYED
IPVEATGQNV PPNITSFDDV QLTEIIRNNV ALARYDKPTP VQKHAIPIII NGRDLMACAQ TGSGKTAAFL
VPILNQMYEL GHVPPPQSTR QYSRRKQYPL GLVLAPTREL ATQIFEEAKK FAYRSRMRPA VLYGGNNTSE
QMRELDRGCH LIVATPGRLE DMITRGKVGL ENIRFLVLDE ADRMLDMGFE PQIRRIVEQL NMPPTGQRQT
LMFSATFPKQ IQELASDFLS NYIFLAVGRV GSTSENITQT ILWVYEPDKR SYLLDLLSSI RDGPEYTKDS
LTLIFVETKK GADSLEEFLY QCNHPVTSIH GDRTQKEREE ALRCFRSGDC PILVATAVAA RGLDIPHVKH
VINFDLPSDV EEYVHRIGRT GRMGNLGVAT SFFNEKNRNI CSDLLELLIE TKQEIPSFME DMSSDRGHGG

AKRAGRGGGG RYGGGFGSRD YRQSSGGGGG GRSGPPPRSG GSGSGGGGGS YRSNGNSYGK FGGNSGGGGY
YGGGAGGGSY GGSYGGGSAS HSSNAPDWWA Q

DDX3X-like RNA helicase, E. pallida (XP_020899200.1)
MSHVAPGNQQ SLDQRFAGLD LNSGVGNNPD AGHNQRQQRY VPPHLRRNPQ ELFHNDPRNP VNFPSGGAPQ
QFQGGGRDGA FRGMNYGGKY NNFGGGGGYG GGGGGYGGRG GYGGAGYRRG GGGGNWRERG GNNYWGNNSG
YDDRDSYAKT ARPEDWSKLL PKNDRIEREL FGGHNTGINF EKYDDIPVEA TGQDCPQNIE SFTDVDLGEI
LTHNIQLANY SKPTPVQKYA IPIVKHKRDL MACAQTGSGK TAAFLIPILS RIYQEGPPPA PDAKHTSRRR
QYPVCLVLAP TRELAVQIFD EARKFAYCSL VRPCVVYGGA DIGSQLRELD RGCHLLVATP GRLVDMMDRG
RIGLDVIKFL VLDEADRMLD MGFEPQIRRI VDQDTMPKAG DRQTLMFSAT FPKEIQILAR DFLDNYIFLA
VGRVGSTSEN ITQKIVWVDE YDKRSFLLDL LNASGPDALT LVFVETKKGA DSLELFLYKD GYQCTSIHGD
RSQSEREEAL RSFRSGKTPI LVATAVAARG LDINNVRHVI NFDLPSDIEE YVHRIGRTGR VGHTGLATSF
FNEKNKNVAK DLLSLVTETG QEVPSWLESI AYESNQNSKR GPRRYGGFGG SRDYRQQRGN SAQMNQMHGY
GGYGGGGGGY MHYGGYSGGG GGGGSGGRYH GGGGGGGGQD WWN

Hypothetical protein, M. brevicolis (XP_001747837.1)
MSNGANPNGS DLSQHMADLD LTKTKPSGGS RYVPPHLRNR QPSGPAPPSG GRTAAPPVSA PPPSSNGGGR
DFGSSRPPRG SRDGSRDMGG SRPPRDGGRG GSWDVQPRFQ QEDWTRPLKR NERMEEELFG SNHRTGGINF
EKYDDIPVEA SGNNVPAHIS EFATAGLCEL MTGNLELARY TVPTPVQKYS IPIVQAKRDL MACAQTGSGK
TAAFLVPILN RVYETGPVPP PPNARRSQQF PVALILAPTR ELAIQIYGEA QKFSYRSRVR ICCVYGGASP
RDQIQDLRRG CQLLVATPGR LVDFMERGVI GLDSIRFLVL DEADRMLDMG FEPQIRRIVE EDNMPQVGIR
QTLMFSATFP KDIQMLAQDF LDDYVHLSVG RVGSTSENIQ QIVHWIDEAD KRPSLLDLIS AASSEDLFLI
FVETKKAADA LEYYLTMQGR PATSIHGDRT QYEREEALAD FRAGRRPILV ATAVAARGLD IPNVKHVINF
DLPSDIDEYV HRIGRTGRAG HKGTAVSFFN DKNRNVARDL LN

Dbp1p, S. cerevisiae (AJW08300.1)
MADLPQKVSN LSINNKENGG DGGKSSYVPP HLRSRGKPSF ERSTPKQEDK VTGGDFFRRA GRQTGNNGGF
FGFSKERNGG TSANYNRGGS SNYKSSGNRW VNGKHIPGPK NAKLEAELFG VHDDPDYHSS GIKFDNYDDI
PVDASGKDVP EPILDFSSPP LDELLMENIK LASFTKPTPV QKYSIPIVTK GRDLMACAQT GSGKTGGFLF
PLFTELFRSG PSPVPEKAQS FYSRKGYPSA LVLAPTRELA TQIFEEARKF TYRSWVRPCV VYGGAPIGNQ
MREVDRGCDL LVATPGRLND LLERGKVSLA NIKYLVLDEA DRMLDMGFEP QIRHIVEECD MPSVENRQTL
MFSATFPVDI QHLARDFLDN YIFLSVGRVG STSENITQRI LYVDDMDKKS ALLDLLSAEH KGLTLIFVET
KRMADQLTDF LIMQNFKATA IHGDRTQAER ERALSAFKAN VADILVATAV AARGLDIPNV THVINYDLPS
DIDDYVHRIG RTGRAGNTGV ATSFFNSNNQ NIVKGLMEIL NEANQEVPTF LSDLSRQNSR GGRTRGGGGF
FNSRNNGSRD YRKHGGSGSF GSTRPRNTGT SNWGSIGGGF RNDNEKNGYG SSNASWW

Sum3, S. pombe (NP_588033.1)
MSDNVQQQVD SVGSVTEKLQ KTNISRPRKY IPPFARDKPS AGAAPAVGDD ESVSSRGSSR SQTPSEFSSN
YGGRREYNRG GHYGGGEGRQ NNYRGGREGG YSNGGGYRNN RGFGQWRDGQ HVIGARNTLL ERQLFGAVAD
GTKVSTGINF EKYDDIPVEV SGGDIEPVNE FTSPPLNSHL LQNIKLSGYT QPTPVQKNSI PIVTSGRDLM
ACAQTGSGKT AGFLFPILSL AFDKGPAAVP VDQDAGMGYR PRKAYPTTLI LAPTRELVCQ IHEESRKFCY
RSWVRPCAVY GGADIRAQIR QIDQGCDLLS ATPGRLVDLI DRGRISLANI KFLVLDEADR MLDMGFEPQI
RHIVEGADMT SVEERQTLMF SATFPRDIQL LARDFLKDYV FLSVGRVGST SENITQKVVH VEDSEKRSYL
LDILHTLPPE GLTLIFVETK RMADTLTDYL LNSNFPATSI HGDRTQRERE RALELFRSGR TSIMVATAVA
SRGLDIPNVT HVINYDLPTD IDDYVHRIGR TGRAGNTGQA VAFFNRNNKG IAKELIELLQ EANQECPSFL
IAMARESSFG GNGRGGRYSG RGGRGGNAYG ARDFRRPTNS SSGYSSGPSY SGYGGFESRT PHHGNTYNSG
SAQSWW

1.3 Homolog sequence alignment

```
NP_001254859.1      ------------------MESN--QS----NNGGSGNAALNRGGRYVPPHLRGGDGGAAA       36
NP_001262379.1      MSNAINQNGTGLEQQVAGLDLNGGSADYSGPITSKTSTNSVTGGVYVPPHLRGGGGNNNA       60
XP_001747837.1      MSNGANPNGSDLSQHMADLDLTKTKP--------------SGGSRYVPPHLRNRQPSGPA       46
XP_020899200.1      MSHVAPGNQQSLDQRFAGLDLNSGVGNNP------DAGHNQRQQRYVPPHLRRNPQELFH       54
XP_005168849.1      MSHVAVENVHGLDQQLAALDLSSADV--Q------G----VTGRRYIPPHLRNKEAAKN-       47
NP_001180345.1      MSHVAVENALGLDQQFAGLDLNSSDN--Q------SGGSTASKGRYIPPHLRNREA----       48
NP_001080283.1      MSHVAVENVLNLDQQFAGLDLNSADA--E------SG-VAGTKGRYIPPHLRNKEASRN-       50
```

```
AJW08300.1          MAD--------LPQKVSNLS-I---------N-NKENGGDGGKSSYVPPHLRSRGKPSFE        41
NP_588033.1         MSDN-------VQQQVDSVGSV---------TEKLQKTNISRPRKYIPPFARDKPSAGAA        44
                                     :                        *:**. *

NP_001254859.1      AASAGGDDRRGGAGGGGYRRG-----------------------GGNS-----------        61
NP_001262379.1      ADAESQGQGQ--GQGQGFDSRSGNPRQETRDPQQSRGGGGEYRRGGGGGGRGFNRQSGDY        118
XP_001747837.1      ----PPS---GGRTAAPPVS-------------------------------------A        60
XP_020899200.1      NDPRNPVNFPSGGAPQQFQGG-----------------GRDGAFRGMNYGG---------        88
XP_005168849.1      ---DAPGGWDNGR-SNGFVNG-----------------YHDGRDNRMNGGSSF---AGRG        83
NP_001180345.1      ------------------------------------------------------------        48
NP_001080283.1      ---DS--NWDSGRGGNGYING----------------MQDDRDGRMNGYDR--------        80
AJW08300.1          RSTPKQED-----------------KV---------------------------------        51
NP_588033.1         PAVGDDESVSSR----------GSSRSQ---------------------------------        62

NP_001254859.1      ----GGGGGGG----YDRGYNDNRDD-----------------RDNRGGSGGYGRDRNY        95
NP_001262379.1      GYGSGGGGRRGGGGRFEDNY--------NGGEFDSRRGGDWNRSGGGGGGGRGFGRGPSY        170
XP_001747837.1      PPPSSNGG-------------------------------------GRDFGSSRPP        78
XP_020899200.1      ------------------------------------------KYNNFGGGGGY        99
XP_005168849.1      PIRSDRGGRGGFRGKSTASYNPIQPMQSAGFGYDNKEAGGWNVPKD--NAYNSFGGRSD-        140
NP_001180345.1      ----------------------------TKGFYDKDSSGWSSSKDK-DAYSSFGSRSDS        78
NP_001080283.1      ------G---GYGSRG---------TGRSDRGFYDRENSGWNSGRDK-DAYSSFGSRGE-        120
AJW08300.1          ---------------------------TGGDF-----------------FR-RAGRQTG        65
NP_588033.1         ----------------------------TPSEF-----------------SSNYGGRREY        77
                                                                        .

NP_001254859.1      EDRGYNGGGGGGGNRGYNNNRGGGGGG-----YNRQDRGDGGSSNFSRGGYNNRDEGSDN        150
NP_001262379.1      -----RGGG-----GGSGSNLNEQTAEDGQAQQQQQPRNDRWQEPERPAGFDGSEGG-QS        219
XP_001747837.1      -----RGSR-----D-------------------------------------G-----        84
XP_020899200.1      -----GGGG-----GGYG----G-------------------------RGGYGGAG-----        116
XP_005168849.1      -----RGKS-----SFFN----DRGSS-----------SRGRYER---------------        160
NP_001180345.1      -----RGKS-----SFFS----DRGSG-----------SRGRFDDRGRSDYDGIG-----        108
NP_001080283.1      -----RGKG-----SLFN----DKGSG-----------SRRP-DESRPDGFDGVG-----        149
AJW08300.1          NNGGFF---------GFSKERNG-----------------------------GT------        81
NP_588033.1         NRGGHYGGG-----EGRQNNYRG-----------------------------GR------        97

NP_001254859.1      RGSGRSYNNDRRDNGGDGQNTRWNNLDA-------PPSRGTSKWENRGARDERIEQELFS        203
NP_001262379.1      AGGNRSYNN-RGERGGGGYNSRWKE-----------GGGSNVDYTKLGARDERLEVELFG        267
XP_001747837.1      ---SRDMGG-SRPPRDGGRGGSWDV----------QPRFQQEDWTRPLKRNERMEEELFG        130
XP_020899200.1      ---YRRGGG-GGNWRERGGNNYWGNNSGYDDRDSYAKTARPEDWSKLLPKNDRIERELFG        172
XP_005168849.1      ------------GGFGGGGNSRWV-----------EEC-RDEDWSKPLPPNERLEHELFS        196
NP_001180345.1      ---SRGDRS-GFGKFERGGNSRWC------------DKSDEDDWSKPLPPSERLEQELFS        152
NP_001080283.1      ---NRGNNS-SFGRFDRGN-SRWS-----------DERNDEDDWSKPLAPNDRVEQELFS        193
AJW08300.1          ---SANYNR-GGSSNYKSSGNRWVN-----------------GKHIPGPKNAKLEAELFG        120
NP_588033.1         ---EGGYSN-GGGYRNNRGFGQWRD-----------------GQHVIGARNTLLERQLFG        136
                                      *                          .  :* :**.

NP_001254859.1      GQ------LSGINFDKYEEIPVEATGDDVPQPISLFSDLSLHEWIEENIKTAGYDRPTPV        257
NP_001262379.1      VG------NTGINFDKYEDIPVEATGQNVPPNITSFDDVQLTEIIRNNVALARYDKPTPV        321
XP_001747837.1      SNH----RTGGINFEKYDDIPVEASGNNVPAHISEFATAGLCELMTGNLELARYTVPTPV        186
XP_020899200.1      GH------NTGINFEKYDDIPVEATGQDCPQNIESFTDVDLGEILTHNIQLANYSKPTPV        226
XP_005168849.1      GS------NTGINFEKYDDIPVEATGHNGPQPIDRFHDLEMGEIIMGNINLSRYTRPTPV        250
NP_001180345.1      GG------NTGINFEKYDDIPVEATGNNCPPHIESFSDVEMGEIIMGNIELTRYTRPTPV        206
NP_001080283.1      GS------NTGINFEKYDDIPVDATGSNCPPHIECFQDVDMGEIIMGNIQLTRYTRPTPV        247
AJW08300.1          VHDDPDYHSSGIKFDNYDDIPVDASGKDVPEPILDFSSPPLDELLMENIKLASFTKPTPV        180
NP_588033.1         AVADGTKVSTGINFEKYDDIPVEVSGGDI-EPVNEFTSPPLNSHLLQNIKLSGYTQPTPV        195
                              **:*::*::***:.:* :    :  *    : . :  *:  : :  ****

NP_001254859.1      QKYSIPALQGGRDLMSCAQTGSGKTAAFLVPLVNAILQDGPDAVHRSVT------SSGGR        311
NP_001262379.1      QKHAIPIIINGRDLMACAQTGSGKTAAFLVPILNQMYELGHVPPPQST-------RQYSR        374
XP_001747837.1      QKYSIPIVQAKRDLMACAQTGSGKTAAFLVPILNRVYETGPVPPPPNA----------RR        236
XP_020899200.1      QKYAIPIVKHKRDLMACAQTGSGKTAAFLIPILSRIYQEGPPPAPDA--------KHTSR        278
XP_005168849.1      QKHAIPIIKSKRDLMACAQTGSGKTAAFLLPVLSQIYTDGPGEALQAAKNSAQENGKYGR        310
NP_001180345.1      QKHAIPIIKEKRDLMACAQTGSGKTAAFLLPILSQIYSDGPGEALRAM----KENGRYGR        262
NP_001080283.1      QKHAIPIIIGKRDLMACAQTGSGKTAAFLLPILSQIYADGPGDAMKHL----KDNGRYGR        303
AJW08300.1          QKYSIPIVTKGRDLMACAQTGSGKTGGFLFPLFTELFRSGPSPVPEKA-------QSFYS        233
NP_588033.1         QKNSIPIVTSGRDLMACAQTGSGKTAGFLFPILSLAFDKGPAAVPVDQ----DAGMGYRP        251
                    ** :** :   ****:*********..**.*:..     *

NP_001254859.1      KKQYPSALVLSPTRELSLQIFNESRKFAYRTPITSALLYGGRENYKDQIHKLRLGCHILI        371
```

```
NP_001262379.1      RKQYPLGLVLAPTRELATQIFEEAKKFAYRSRMRPAVLYGGNN-TSEQMRELDRGCHLIV       433
XP_001747837.1      SQQFPVALILAPTRELAIQIYGEAQKFSYRSRVRICCVYGGAS-PRDQIQDLRRGCQLLV       295
XP_020899200.1      RRQYPVCLVLAPTRELAVQIFDEARKFAYCSLVRPCVVYGGAD-IGSQLRELDRGCHLLV       337
XP_005168849.1      RKQYPISLVLAPTRELALQIYDEARKFSYRSHVRPCVVYGGAD-IGQQIRDLERGCHLLV       369
NP_001180345.1      RKQYPISLVLAPTRELAVQIYEEARKFSYRSRVRPCVVYGGAD-IGQQIRDLERGCHLLV       321
NP_001080283.1      RKQFPLSLVLAPTRELAVQIYEEARKFAYRSRVRPCVVYGGAD-IGQQIRDLERGCHLLV       362
AJW08300.1          RKGYPSALVLAPTRELATQIFEEARKFTYRSWVRPCVVYGGAP-IGNQMREVDRGCDLLV       292
NP_588033.1         RKAYPTTLILAPTRELVCQIHEESRKFCYRSWVRPCAVYGGAD-IRAQIRQIDQGCDLLS       310
                     : :*  *:*:*****  **. *::** * : :  . :***      *::.:  **.::

NP_001254859.1      ATPGRLIDVMDQGLIGMEGCRYLVLDEADRMLDMGFEPQIRQIVECNRMPSKEERITAMF       431
NP_001262379.1      ATPGRLEDMITRGKVGLENIRFLVLDEADRMLDMGFEPQIRRIVEQLNMPPTGQRQTLMF       493
XP_001747837.1      ATPGRLVDFMERGVIGLDSIRFLVLDEADRMLDMGFEPQIRRIVEEDNMPQVGIRQTLMF       355
XP_020899200.1      ATPGRLVDMMDRGRIGLDVIKFLVLDEADRMLDMGFEPQIRRIVDQDTMPKAGDRQTLMF       397
XP_005168849.1      ATPGRLVDMMERGKIGLDYCNYLVLDEADRMLDMGFEPQIRRIVEQDTMPPKGLRQTMMF       429
NP_001180345.1      ATPGRLVDMMERGKIGLDFCKYLVLDEADRMLDMGFEPQIRRIVEQDTMPPKGVRHTMMF       381
NP_001080283.1      ATPGRLVDMMERGKIGLDFCKYLVLDEADRMLDMGFEPQIRRIVEQDTMPPKGVRQTMMF       422
AJW08300.1          ATPGRLNDLLERGKVSLANIKYLVLDEADRMLDMGFEPQIRHIVEECDMPSVENRQTLMF       352
NP_588033.1         ATPGRLVDLIDRGRISLANIKFLVLDEADRMLDMGFEPQIRHIVEGADMTSVEERQTLMF       370
                    ****** *.: :* :.:   .:*******************:**:   *     * * **

NP_001254859.1      SATFPKEIQLLAQDFLKENYVFLAVGRVGSTSENIMQKIVWVEEDEKRSYLMDLLDAT--       489
NP_001262379.1      SATFPKQIQELASDFLS-NYIFLAVGRVGSTSENITQTILWVYEPDKRSYLLDLLSSIRD       552
XP_001747837.1      SATFPKDIQMLAQDFLD-DYVHLSVGRVGSTSENIQQIVHWIDEADKRPSLLDLISAA--       412
XP_020899200.1      SATFPKEIQILARDFLD-NYIFLAVGRVGSTSENITQKIVWVDEYDKRSFLLDLLNAS--       454
XP_005168849.1      SATFPKEIQILARDFLE-DYIFLAVGRVGSTSENITQKVVWVEENDKRSFLLDLLNAT--       486
NP_001180345.1      SATFPKEIQMLARDFLD-EYIFLAVGRVGSTSENITQKVVWVEESDKRSFLLDLLNAT--       438
NP_001080283.1      SATFPKEIQILARDFLD-EYIFLAVGRVGSTSENITQKVVWVEEMDKRSFLLDLLNAT--       479
AJW08300.1          SATFPVDIQHLARDFLD-NYIFLSVGRVGSTSENITQRILYVDDMDKKSALLDLLSA---       408
NP_588033.1         SATFPRDIQLLARDFLK-DYVFLSVGRVGSTSENITQKVVHVEDSEKRSYLLDILHTL--       427
                    ***** :** ** ***. :*:.*:*********** * :  : : :*:  *:*:: :

NP_001254859.1      ----GDSSLTLVFVETKRGASDLAYYLNRQNYEVVTIHGDLKQFEREKHLDLFRTGTAPI       545
NP_001262379.1      GPEYTKDSLTLIFVETKKGADSLEEFLYQCNHPVTSIHGDRTQKEREEALRCFRSGDCPI       612
XP_001747837.1      ----SSEDLFLIFVETKKAADALEYYLTMQGRPATSIHGDRTQYEREEALADFRAGRRPI       468
XP_020899200.1      ----GPDALTLVFVETKKGADSLELFLYKDGYQCTSIHGDRSQSEREEALRSFRSGKTPI       510
XP_005168849.1      ----GKDSLTLVFVETKKGADALEDFLYREGYACTSIHGDRSQRDREEALHQFRSGRCPI       542
NP_001180345.1      ----GKDSLTLVFVETKKGADSLEDFLYHEGYACTSIHGDRSQRDREEALHQFRSGKSPI       494
NP_001080283.1      ----GKDSLTLVFVETKKGADALEDFLYHEGYACTSIHGDRSQRDREEALHQFRSGKCPI       535
AJW08300.1          ----EHKGLTLIFVETKRMADQLTDFLIMQNFKATAIHGDRTQAERERALSAFKANVADI       464
NP_588033.1         ----PPEGLTLIFVETKRMADTLTDYLLNSNFPATSIHGDRTQRERERALELFRSGRTSI       483
                          . * *:*****: *. *  :*   .   .:**** .* :**. *  *::.   *

NP_001254859.1      LVATAVAARGLDIPNVKHVINYDLPSDVDEYVHRIGRTGRVGNVGLATSFFNDKNRNIAR       605
NP_001262379.1      LVATAVAARGLDIPHVKHVINFDLPSDVEEYVHRIGRTGRMGNLGVATSFFNEKNRNICS       672
XP_001747837.1      LVATAVAARGLDIPNVKHVINFDLPSDIDEYVHRIGRTGRAGHKGTAVSFFNDKNRNVAR       528
XP_020899200.1      LVATAVAARGLDINNVRHVINFDLPSDIEEYVHRIGRTGRVGHTGLATSFFNEKNKNVAK       570
XP_005168849.1      LVATAVAARGLDISNVKHVINFDLPSDIEEYVHRIGRTGRVGNLGLATSFFNDKNGNITK       602
NP_001180345.1      LVATAVAARGLDISNVKHVINFDLPSDIEEYVHRIGRTGRVGNLGLATSFFNERNINITK       554
NP_001080283.1      LVATAVAARGLDISNVKHVINFDLPSDIEEYVHRIGRTGRVGNLGLATSFFNEKNINITK       595
AJW08300.1          LVATAVAARGLDIPNVTHVINYDLPSDIDDYVHRIGRTGRAGNTGVATSFFNSNNQNIVK       524
NP_588033.1         MVATAVASRGLDIPNVTHVINYDLPTDIDDYVHRIGRTGRAGNTGQAVAFFNRNNKGIAK       543
                    :******:***** :* ****:***:*:::********** *: * *.:*** .* .:

NP_001254859.1      ELMDLIVEANQELPDWLEGMSGDMRSGGGYRGRGGRG----NGQRFGGRDHRYQGGSGNG       661
NP_001262379.1      DLLELLIETKQEIPSFMEDMSSDRGHGGAKRAGRGGG--GRYGGGFGSRDYRQSSGGGGG       730
XP_001747837.1      DLLN-------------------------------------------------------       532
XP_020899200.1      DLLSLVTETGQEVPSWLESIAYESNQNSKRG-----P--RRYGGFGGSRDYRQQRGNSAQ       623
XP_005168849.1      DLLDILVEAKQEVPSWLESLAYEHQHKSSSRG--RSK---RFSGGFGARDYRQNSSSGGG       657
NP_001180345.1      DLLDLLVEAKQEVPSWLENMAYEHHYKGSSRG--RSK--SRFSGGFGARDYRQSSGASSS       610
NP_001080283.1      DLLDLLVEAKQEVPSWLENMAYEQHHKSSSRG--RSK--SRFSGGFGAKDYRQSSSAGSS       651
AJW08300.1          GLMEILNEANQEVPTFLSDLSRQNSRGGRTRGGGGF----FNSRNNGSRDYRKHGGSGSF       580
NP_588033.1         ELIELLQEANQECPSFLIAMARESSFGGNGRGGRYSGRGGRGGNAYGARDFRRPTNSSSG       603
                     *:.

NP_001254859.1      GGG----NGGGGGFGGGG-----------------QRSGGGG---GFQSGG--GGGRQQQ       695
NP_001262379.1      GRSGPPPRSGGSGSGGGGGSYRS---NGNSYGKFGGNSGGGGYYGGGAGGGSYGGSYGGG       787
XP_001747837.1      -------------------------------------------------------------       532
XP_020899200.1      MNQM----HGYGGYGGGGGGYMHY----GGYSG----GGGGGGSGGRYH-----GGGGG-       665
XP_005168849.1      ------------GFGGRGG----RSTGGHGGNR----GFGGGGFGNFYSSDGYGGNY---       694
NP_001180345.1      ------------SFSSSRASSSRSGGGGHGSSR----GFGGGGYGGFYNSDGYGGNYNS-       653
```

```
NP_001080283.1      -------------FGSSRG----GRSSGHGGSR----AFG-GGYGGFYNSDGYGGNYGG-        688
AJW08300.1          GSTRP-RNTGTSNWGSIGGGFRND--------------------------NEKNGYG--        610
NP_588033.1         YSSGP-SYSG---Y----GGFESRT------------------------PHHGNTYN--        628


NP_001254859.1      QQQRAQPQQDWWS-      708
NP_001262379.1      SASHSSNAPDWWAQ      801
XP_001747837.1      --------------      532
XP_020899200.1      -----GGGQDWWN-      673
XP_005168849.1      ------SQVDWWGN      702
NP_001180345.1      ------QGVDWWGN      661
NP_001080283.1      -----SSQVDWWGN      697
AJW08300.1          -----SSNASWW--      617
NP_588033.1         ----SGSAQSWW--      636
```

1.4 High-throughput identification of contact-prone region

When conducting single-chain simulations of the many different deletion sequences, we note that many of them have a higher T_θ than the full-length RGG, counter to the expectation that longer chain length generally favors LLPS. We believe this effect in the simulation model can be attributed to a subtle balance between the changes in hydrophobicity, net charge, and SCD rather than a single sequence descriptor (SI Appendix Figure 1Bii). Given the simplicity of our simulation model and the errors associated with predicting phase separation based solely on T_θ, we can only distinguish sequences such as Δ21-30 which have more significant changes to LLPS behavior, but cannot capture smaller changes as with the other sequences without conducting additional extensive simulations on the phase behavior of these variants.

1.5 Calculation of minimum possible SCD for sequence with same composition as LAF-1 RGG

To obtain a sequence with the minimum possible SCD value, the charged amino acids must be clustered at the very ends of the sequence with positive charges at one end, negative charges at the other, and uncharged amino acids in between. Since we consider histidine in our model to have a +0.5 charge, the +1 charged amino acids should be at the very end with histidine residues following.

We also must consider that for in vitro studies, the initial methionine residue and the LEHHHHHH tag must be conserved. Thus a sequence with minimum possible SCD is:

```
MDDDDDDDDDDDDDDDDEEE......HRRRRRRRRRRRRRRRRRRRRRRRRLEHHHHHH
```

with all of the uncharged residues in between the negatively-charged N-terminal and the positively-charged C-terminal, and having an SCD of -28.032. Note that since D and E have the same charge, any permutation of residues 2-21 would not change the SCD value.

The probability of randomly sampling a sequence with the minimum SCD value can be calculated by considering the number of residues being shuffled as 176 – 9 = 167. Then one must consider the four regions that must be correct:

1. All D and E residues within 2-21
2. All R residues within 145-168
3. H residue at 144
4. All uncharged residues within 22-143

To account for these values and the degeneracies, we can calculate the probability of randomly sampling such a sequence as

$$\boldsymbol{p_{minSCD}} = \frac{\boldsymbol{n_{DE}}! \cdot \boldsymbol{n_{RK}}! \cdot \boldsymbol{n_H}! \cdot (\boldsymbol{n} - \boldsymbol{n_{DE}} - \boldsymbol{n_{RK}} - \boldsymbol{n_H})!}{\boldsymbol{n}!} = \frac{\boldsymbol{20}! \cdot \boldsymbol{24}! \cdot \boldsymbol{1}! \cdot \boldsymbol{122}!}{\boldsymbol{167}!} = \boldsymbol{9.914 \times 10^{-56}}$$

2. Supplementary Materials and Methods

2.1 Cloning

The WT, full-length LAF-1 gene was a gift of Shana Elbaum-Garfinkle and Clifford Brangwynne. WT RGG was amplified by PCR from LAF-1. All modified versions of the RGG domain were ordered as synthetic double-stranded DNA fragments (gBlocks; IDT). Plasmids were constructed using either In-Fusion cloning (Takara Bio) or NEBuilder HiFi DNA Assembly (New England BioLabs). For bacterial expression, genes were cloned into a pET vector in-frame with a C-terminal 6xHis-tag. For yeast expression, genes were cloned into the YIplac211 vector in frame with a C-terminal mEGFP (monomeric enhanced GFP) tag. YIplac211 is a yeast integrating plasmid with a URA3 marker[1]. Gene sequences were verified by Sanger sequencing (GENEWIZ).

2.2 Protein expression and purification

For bacterial expression, plasmids were transformed into BL21(DE3) competent *E. coli* (New England BioLabs). Colonies picked from fresh plates were grown for 8 h at 37 °C in 1 mL LB + 1% glucose while shaking at 250 rpm. This starter culture (0.5 mL) was then used to inoculate 0.5 L cultures. Cultures were grown overnight in 2L baffled flasks in Terrific Broth auto-induction medium (Formedium; supplemented with 4 g/L glycerol) at 37 °C while shaking at 250 rpm. The pET vectors used contained a kanamycin resistance gene; kanamycin was used at concentrations of 50 µg/mL in starter cultures and 100 µg/mL in the auto-induction medium[2]. After overnight expression, bacterial cells were pelleted by centrifugation. Pellets were resuspended in lysis buffer (1 M NaCl, 20 mM Tris, 20 mM imidazole, Roche EDTA-free protease inhibitor, pH 7.5) and lysed by sonication. Lysate was clarified by centrifugation at 15,000 × g for 30-60 minutes. Lysis was conducted on ice, but other steps were conducted at room temperature to prevent phase separation. Proteins were purified using an AKTA FPLC with 1 mL nickel-charged HisTrap columns (GE Healthcare Life Sciences) for affinity chromatography of the His-tagged proteins. The column was washed with 500 mM NaCl, 20 mM Tris, 20 mM imidazole, pH 7.5. Proteins were eluted with a linear gradient up to 500 mM NaCl, 20 mM Tris, 500 mM imidazole, pH 7.5. Proteins were dialyzed overnight using 7 kDa MWCO membranes (Slide-A-Lyzer G2, Thermo Fisher) into 500 mM NaCl, 20 mM Tris, pH 7.5 or 150 mM NaCl, 20 mM Tris, pH 7.5. Proteins were dialyzed at temperatures (25 °C -

42 °C) high enough to inhibit phase separation because phase-separated protein bound irreversibly to the dialysis membrane. Proteins were snap frozen in liquid N_2 in single-use aliquots and stored at -80 °C. For turbidity and microscopy experiments, protein samples were prepared as follows: Protein aliquots were thawed above the phase transition temperature. Proteins were then mixed with buffer (20 mM Tris, pH 7.5, 0 – 150 mM NaCl) to obtain solutions containing the desired protein and NaCl concentrations. Protein concentrations were measured based on their absorbance at 280 nm using a Nanodrop spectrophotometer (ThermoFisher). Proteins were mixed in a 1:1 ratio with 8 M urea to prevent phase separation during concentration measurements.

2.3 Turbidity assays

Temperature-dependent turbidity assays were conducted using a UV-Vis spectrophotometer (Cary 100 Bio; Agilent) equipped with a multicell Peltier temperature controller. Protein samples were assayed in quartz cuvettes with 1 cm path length (Thorlabs). Samples were first equilibrated above the phase transition temperature (25-60 °C depending on the sample) and blanked. Then, the samples were cooled at a rate of 1 °C per minute until reaching 2 °C. Absorbance was measured at λ = 600 nm every 0.5 °C throughout the temperature ramp. Upon cooling below the phase transition temperature, the samples changed from clear to turbid.

2.4 MALDI-TOF mass spectrometry

Molecular weights of purified proteins were measured by matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry on an Ultraflextreme mass spectrometer (Bruker). Protein samples were applied as spots to an MPT 384 polished steel target plate. Spots consisted of 1 µL protein solution (approximately 10 µM protein in 50 mM NaCl) plus 1 µL matrix solution (10 mg/mL sinapinic acid dissolved in a 50:50 acetonitrile:water mixture with 0.1% trifluoroacetic acid added).

2.5 SDS-PAGE and western blot

For chromatographically purified proteins, SDS-PAGE was run using NuPAGE 4-12% Bis-Tris gels (Invitrogen) and stained using a Coomassie stain (SimplyBlue SafeStain; Invitrogen). For western blotting, yeast cells were lysed as follows[3]: Cell cultures were pretreated with 2 M lithium acetate for 5 minutes on ice, then with 0.4 M NaOH for 5 minutes on ice. The cell cultures were then resuspended in SDS sample buffer, heated at 95 °C for 5 minutes, and centrifuged to remove cell debris. The supernatant was stored at -80 °C until use. The supernatant was run on a Novex 10% Tris-Glycine gel, WedgeWell format (Invitrogen), then transferred to a nitrocellulose membrane (0.2 µm pore size). The membrane was then incubated with two primary antibodies: rabbit polyclonal antibody to GFP (Invitrogen, catalog #A11122) for detection of the GFP-tagged LAF-1 constructs, and mouse monoclonal antibody to PGK1 (Invitrogen, catalog #459250) as a loading control. Secondary antibodies used for detection were IRDye 680RD goat anti-rabbit IgG (LI-COR, catalog #926-68071) and IRDye 800CW goat anti-mouse IgG (LI-COR, catalog #926-32210). Blots were visualized on a LI-COR Odyssey CLx infrared imaging system.

2.6 Yeast transformation and yeast cultures

YIplac211 plasmids were prepared for yeast chromosomal integration by restriction digest with EcoRV, which cuts in the URA3 marker. Linearized plasmids were transformed into *S. cerevisiae* YEF473A strain[4] using the Frozen-EZ Yeast Transformation II Kit (Zymo Research). Transformed yeast cells were cultured at 30°C in uracil dropout synthetic defined medium (-Ura dropout supplement was purchased from Takara Bio). To induce expression of genes under the control of the GAL1 promoter, yeast cultures were first grown overnight in dropout medium + 2% glucose, then grown for 8-10 hours in dropout medium + 2% raffinose, and finally grown overnight in dropout medium + 2% galactose with a target OD_{600} = 0.3 – 0.5 for imaging.

2.7 Microscopy: phase behavior, FRAP, and fusion

Imaging of temperature-dependent phase behavior in vitro and in yeast was performed on an Olympus IX81 inverted microscope equipped with a Yokogawa CSU-X1 spinning disk confocal unit and an iXon3 EMCCD camera (Andor). The microscope stage was outfitted with a Cherry Temp microfluidic temperature controller (Cherry Biotech), which enabled imaging samples over the temperature range 5 to

42 °C, with rapid switching (approximately 10 s) between temperature extremes. Imaging was conducted with a 100x/1.4 NA plan-apochromatic oil-immersion objective.

FRAP experiments were performed on a Zeiss Axio Observer 7 inverted microscope equipped with an LSM900 laser scanning confocal module and a 63x/1.4 NA plan-apochromatic oil-immersion objective. LAF-1 RGG and its variants were mixed with 5% of RGG-GFP-RGG, which partitions into the RGG droplets and serves as a FRAP probe[5]. GFP was imaged with a 488 nm laser and bleached with a 405 nm laser. Circular bleach regions of approximate radius R = 1.5 μm were drawn in the center of protein droplets whose radii were at least 2.5R. Recovery curves were fit to an infinite boundary model in three dimensions to calculate the recovery timescale τ[6]. The diffusion coefficient was calculated as $D = R^2/\tau$. The same Zeiss microscope was used for droplet fusion experiments, but using brightfield transillumination and imaging onto an Axiocam 702 sCMOS camera at a frame rate of approximately 62 Hz. Droplet fusion was analyzed by first fitting the image of the fusing droplets to an ellipse and calculating the aspect ratio of the ellipse. The aspect ratio was then plotted against time and the decreasing portion of the curve was fit to an exponential decay to calculate the relaxation time[7–9]. The droplet length scale was defined as the radius of the droplet after completion of fusion, when the merged droplet was circular (aspect ratio 1). FRAP and droplet fusion experiments were conducted at room temperature of 16-18 °C using protein concentrations above c_{sat} at that temperature. Image analysis and data processing were performed in MATLAB.

All other imaging was performed on a Leica DMi8 inverted microscope equipped with a spinning disk confocal unit (Spectral Applied Research) and an sCMOS camera (Orca Flash 4.0; Hamamatsu) using a 63x/1.4 NA or 100x/1.4 NA plan-apochromatic oil-immersion objective.

For imaging purified RGG proteins, the protein samples were placed in chambers on glass coverslips (#1.5 glass thickness) that had been passivated for >1 hr by incubation with 5% Pluronic F127 (for FRAP and droplet fusion experiments) or bovine serum albumin. Coated coverslips were thoroughly rinsed with buffer prior to the addition of RGG protein solutions. For imaging yeast, the glass surface was pretreated by incubation with 0.4 mg/mL concanavalin A (ConA; Sigma) for 5-10 minutes. After removing the ConA solution, yeast was pipetted into the imaging chamber and allowed to settle for several minutes before imaging.

2.8 Coarse-grained simulations

Coarse-grained simulations were conducted using an amino-acid-resolution model with 20 residue types to capture sequence specificity, having interactions based on relative hydropathies of each amino acid. Each system was simulated at a range of temperatures using constant volume and temperature using a Langevin thermostat, following similar protocols to our previous work[10]. Simulations of phase coexistence were conducted using HOOMD-Blue v2.1.5 software package[11].

2.9 All-atom simulations

Atomic-resolution simulations were conducted for systems containing either one or two copies of a 44-residue fragment of the LAF-1 RGG domain ($RGG_{106-149}$). Simulations were of 44-residue fragments as we have found this size to be computationally tractable for single- and two-chain simulations in the previous studies[12,13]. We selected residues 106-149 by calculating the overall sequence composition of all possible 44-residue fragments and comparing them with the total composition of the 168-residue RGG domain (SI Appendix Fig. S5A). The region having the overall composition most similar to that of the full RGG domain was residues 106-149. Notably, this fragment contains 6 arginine and 3 tyrosine residues constituting 13.6% and 6.8% of the 44-residue sequence, comparable to the 14.3% and 6.5% composition in the full RGG (SI Appendix Fig. S5B,C).

Simulations were conducted with either a single $RGG_{106-149}$ chain solvated in explicit water and ~100 mM NaCl or two chains at the same conditions. We used a modified version of the state-of-the-art Amber99SBws force field[14] with improved residue-specific dihedral corrections, tip4p/2005 water[15] and improved salt parameters from Luo and Roux[16]. To efficiently sample the configurational ensemble and contacts between amino acid residues, we employed enhanced sampling using parallel tempering in the well-tempered ensemble (PT-WTE) which couples replica exchange molecular dynamics (REMD)[17] and well-tempered metadynamics[18] applied to the total system energy to enhance fluctuations and reduces the number of replicas required for good replica exchanges[19]. For two-chain simulations, we also applied a well-tempered metadynamics bias on the intermolecular VDW contacts between heavy nonpolar atoms (i.e. $|q| < 0.25$) as we have done previously to improve sampling of binding and unbinding events[13]. Simulations were conducted using GROMACS 2016 software package[20] with PLUMED 2.4 plugin[21].

We calculated the free energy surface of the two-chain systems from the metadynamics bias using the built-in function (sum_hills) in PLUMED, and an alternative time-independent method from Tiwary and Parrinello[22], then subtract the difference between the two results to generate error bars for SI Appendix Fig. S6A. Contact propensities in all-atom two-chain PT-WTE simulations were reweighted based on free energy surface.

VDW contacts were considered as any two heavy atoms being within 6 Å of each other. Hydrogen bonds were considered as a donor atom and an acceptor atom being within 3 Å and the donor-hydrogen-acceptor angle being larger than 120°. The sp^2/π interactions were calculated as presented by Vernon et al[23] and considered as any two sp^2-hybridized groups having at least two pairs of atoms being within 4.9 Å and the angle between the normal axes of the two sp^2-planes being less than 60°. Cation-π interactions were considered as a cationic atom being within 7 Å of the center of an aromatic ring and less than 60° from the normal axis of the π face. Salt bridges are considered as a cationic atom and an anionic atom being within 6 Å of each other.

3. Supplementary Figures

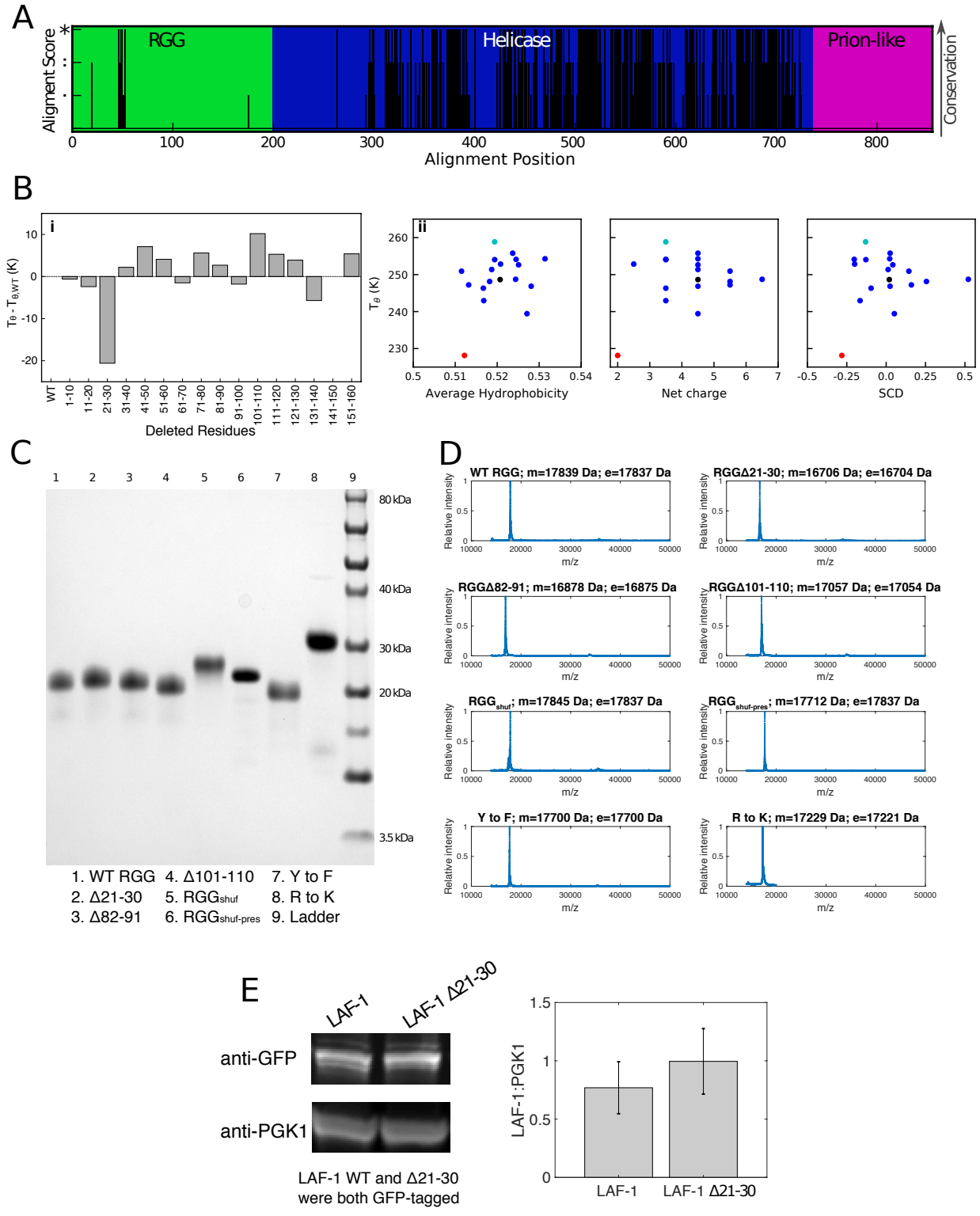


Figure S1: Characterization of deletion variants of LAF-1 RGG: A) Sequence conservation of full-

length LAF-1, showing a high degree of conservation in folded helicase domain, and poor conservation in disordered RGG and prion-like domains. B) i) T_θ calculated from single-chain simulations of the deletion series subtracted from the T_θ of WT LAF-1 RGG ii) T_θ values compared to sequence descriptors (SI Text 1.4). In general, higher T_θ is expected to be associated with higher average hydrophobicity, smaller absolute net charge, and more negative SCD. The symbol colors correspond to WT (black), Δ21-30 (red), Δ101-110 (cyan), with all other variants represented as blue. We note that many of the deletion sequences have a higher T_θ than the full-length RGG, counter to the expectation that longer chain length generally favors LLPS. We believe this effect in the simulation model can be attributed to a subtle balance between the changes in hydrophobicity, net charge, and SCD rather than a single sequence descriptor. Given the simplicity of our simulation model and the errors associated with predicting phase separation based solely on T_θ, it is possible our computational framework can distinguish sequences such as Δ21-30 which have more significant changes to LLPS behavior, but cannot capture smaller changes as with the other sequences. C) SDS-PAGE gel of purified RGG and its variants. D) MALDI-TOF mass spectra of RGG domain and its variants, where m denotes measured and e denotes expected molecular mass. The only discrepancy > 10 Da is $RGG_{shuf\text{-}pres}$ which is likely due to the loss of initiating methionine. E) Western blot shows a similar expression level of LAF-1 WT and LAF-1 Δ21-30 in yeast.

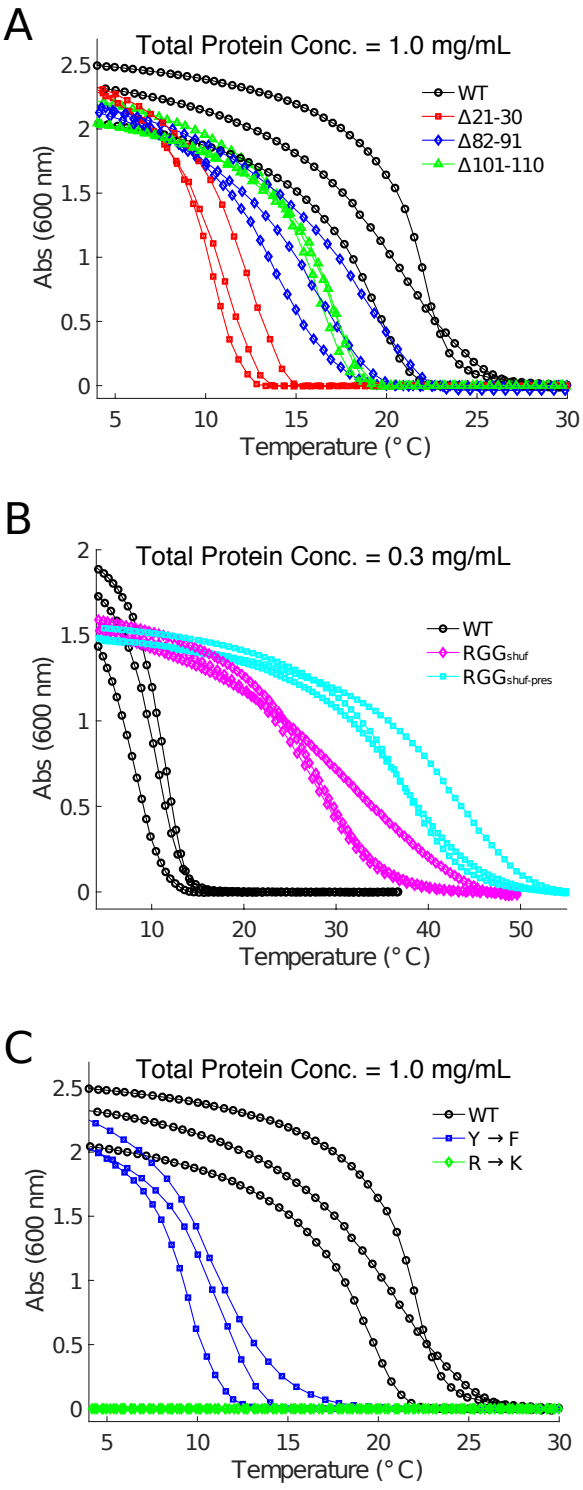


Figure S2: Replicates of turbidity experiments (corresponding to Fig. 1D, 2D, 3B): Turbidity curves of WT and A) deletion variants, B) shuffled sequences, and C) bulk mutations. Protein concentrations were 1 mg/mL for (A) and (C) and 0.3 mg/mL in (B). WT data is the same for (A) and (C). In all cases, proteins were in 150 mM NaCl, 20 mM Tris, pH 7.5.

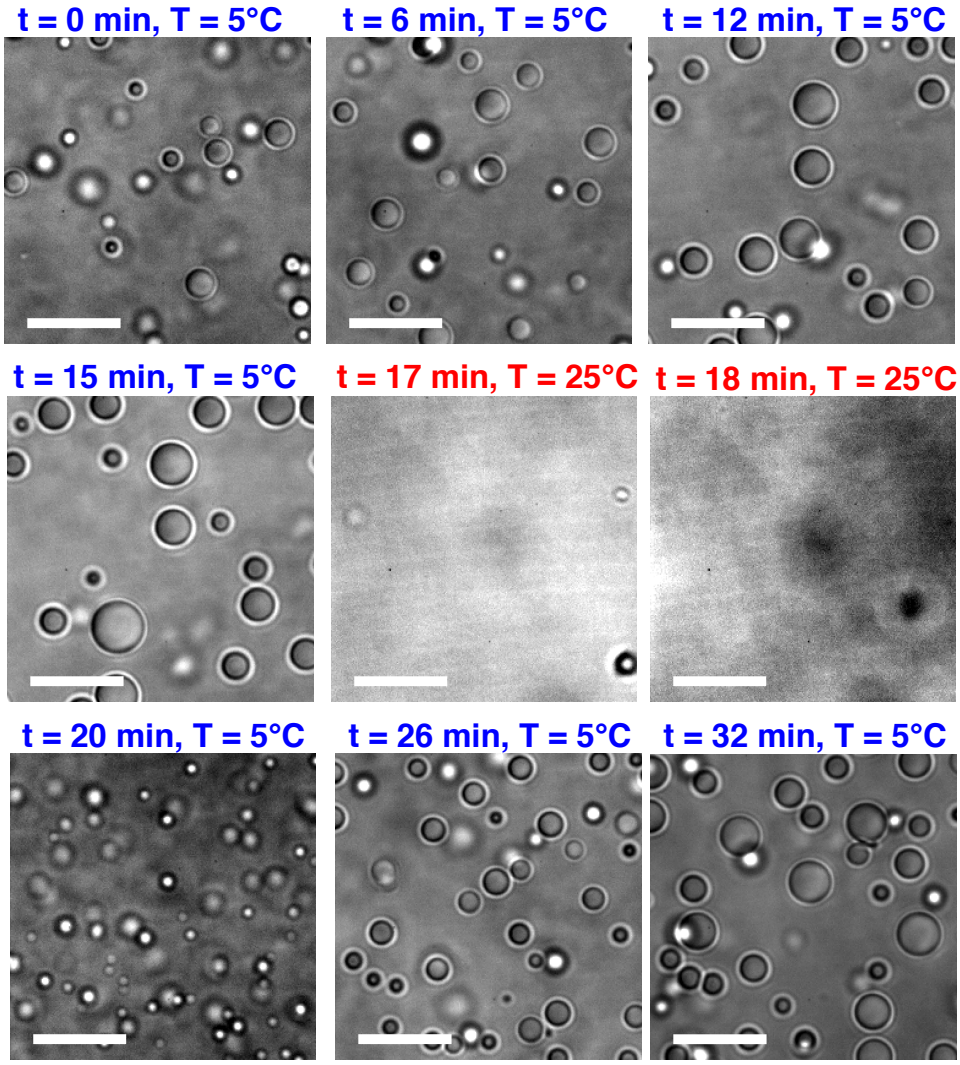


Figure S3: Reversible LLPS of Δ21-30 variant: The Δ21-30 variant of RGG undergoes reversible, temperature-dependent LLPS. Snapshots follow the formation of droplets over time starting at low temperature (5 °C), then rapidly increasing temperature from 5 °C to 25 °C to disperse the droplets, and then rapidly decreasing the temperature back to 5 °C to induce phase separation again. Scale bars: 10 μm.

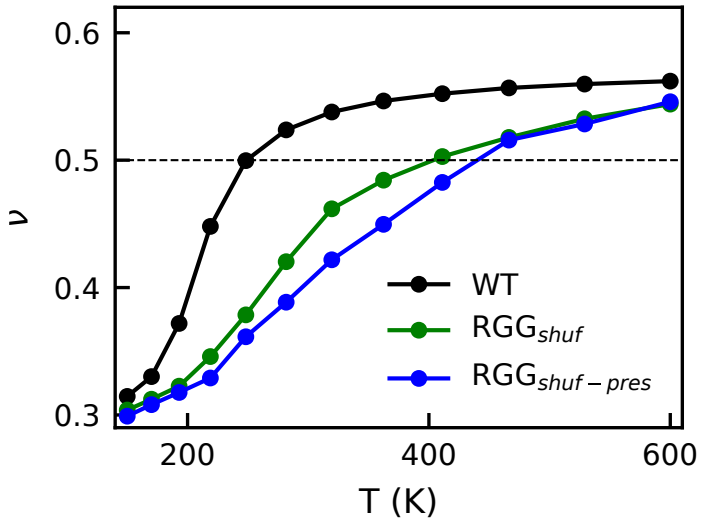


Figure S4: Single-chain compactness of RGG and shuffled variants: We calculate the Flory scaling exponent (v) of the three variants of RGG as in previous work[24,25] and see that the WT is significantly more extended than the shuffled variants at a wide range of temperatures. We also see that $RGG_{shuf\text{-}pres}$ is marginally more compact than RGG_{shuf}, consistent with our experimental results showing that $RGG_{shuf\text{-}pres}$ has the greatest LLPS propensity.

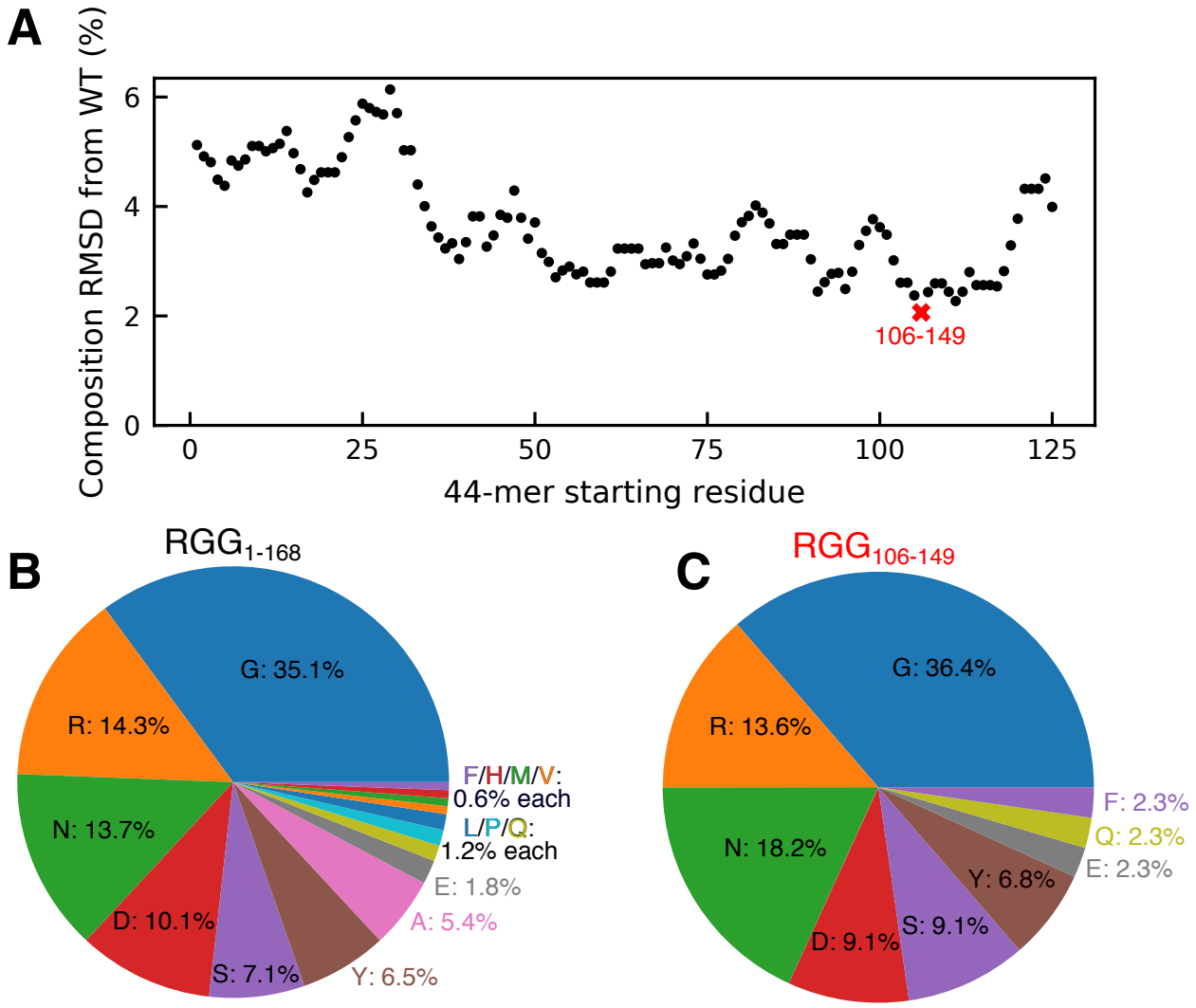


Fig S5: Sequence composition of WT RGG and 44-residue fragments: A) Composition-based RMSD is calculated for all continuous 44-residue fragments of LAF-1 RGG, showing the overall compositional similarity with the full 168-residue sequence. A total of 168 – 44 + 1 = 125 sequences of 44 residues were tested. B) Pie chart of amino acid composition of WT RGG is highly similar to C) pie chart of the lowest-RMSD 44-mer, $RGG_{106\text{-}149}$.

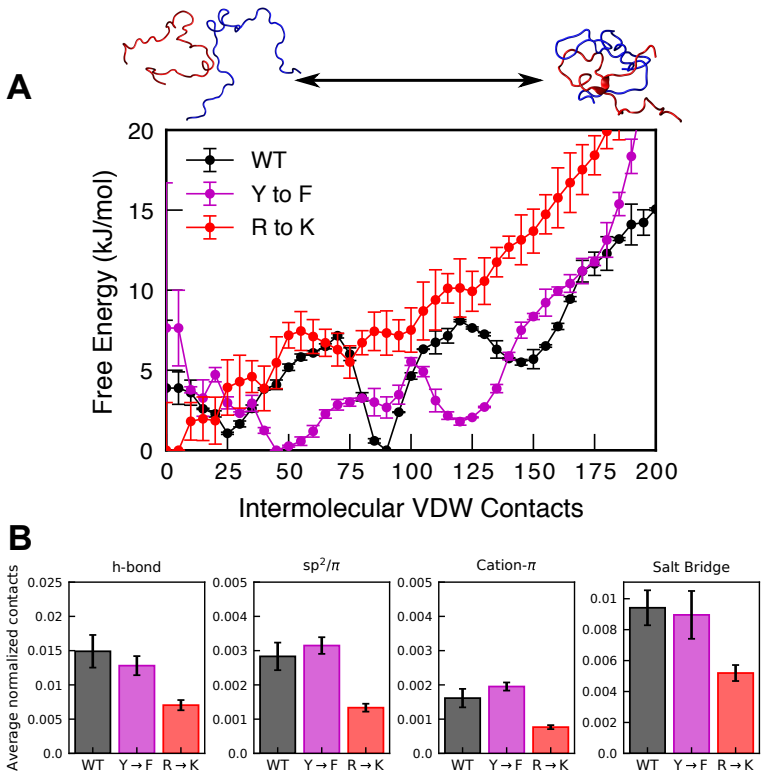


Figure S6: All-atom simulations of $RGG_{106\text{-}149}$ show R→K has lower self-association: A) Free energy profile of contact formation between two identical $RGG_{106\text{-}149}$ chains from simulations using well-tempered metadynamics. Both WT and Y→F show global minima at a finite number of contacts, while R→K has a global minimum at 0 contacts, indicating unfavorable self-interaction. B) Average total number of intermolecular contacts from two-chain simulations normalized by the average total number of VDW contacts for that system. Error bars for all plots are SEM with n = 2.

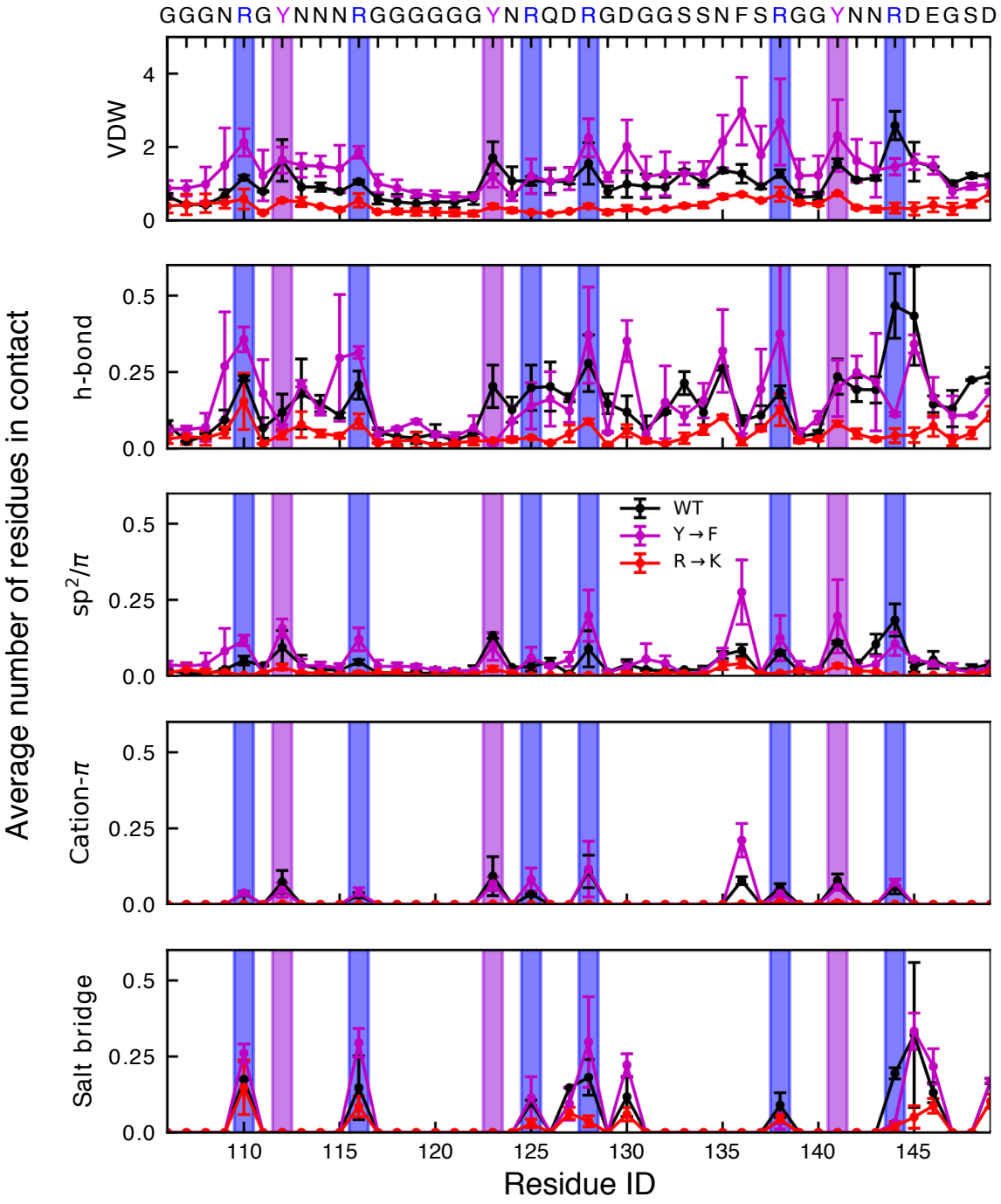


Figure S7: Per-residue contacts from all-atom simulations: Average number of intermolecular residue-residue pairs for each residue of the $RGG_{106\text{-}149}$ sequence. Two residues are considered to be in contact if there is at least one atom from each residue in contact (VDW) or at least one hydrogen bond, sp^2/π interaction, cation-π interaction, or salt bridge between the two residues. In the case of VDW, multiple residues may be in contact with a single residue as only one atom needs to be in contact, and

residues may have VDW interactions with many other amino acids on the other protein chain. Generally, we see that VDW interactions and hydrogen bonds are well-distributed throughout the sequence for all variants of $RGG_{106-149}$. Cation-pi, sp^2/π, and salt-bridge interactions are less well-distributed due to their dependence on certain amino acid side chains. To highlight the contribution of aromatic and cationic residues, we have highlighted the arginine and tyrosine residues in these plots. Error bars are SEM with n = 2.

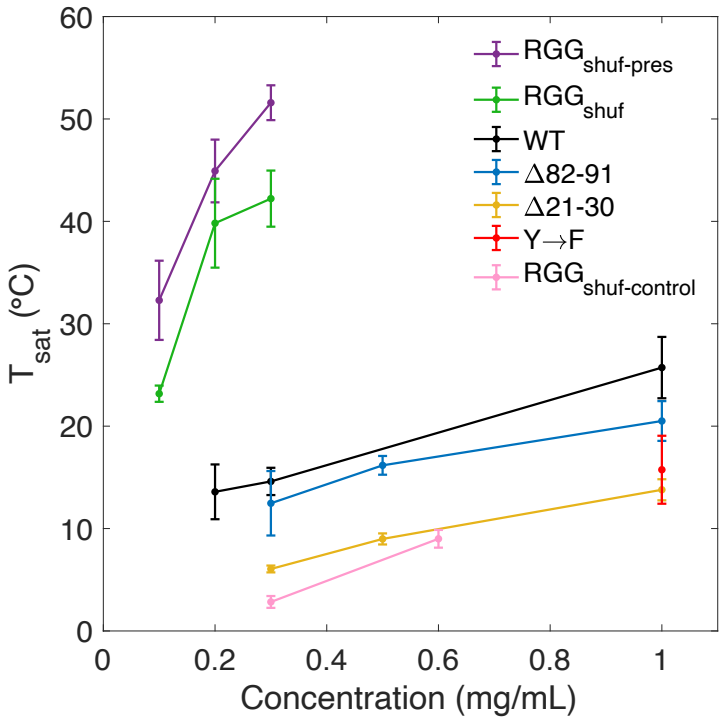


Figure S8: Phase diagrams of RGG mutants: Saturation temperature as a function of total protein concentration for RGG_{shuf}, $RGG_{shuf\text{-}pres}$, $RGG_{shuf\text{-}control}$, Δ21-30, Δ82-91, Y→F, and WT RGG. T_{sat} values are determined from turbidity curves where absorbance first exceeds 0.02. Error bars are STD with n = 3. T_{sat} of WT is significantly different than that of Δ21-30, RGG_{shuf}, $RGG_{shuf\text{-}pres}$, and $RGG_{shuf\text{-}control}$ ($p \leq 0.005$), but not significantly different than that of Δ82-91 ($p = 0.73$), based on one-way ANOVA followed by Tukey's post-hoc test at 0.3 mg/mL. T_{sat} of WT is significantly different than that of Δ21-30 and Y→F ($p \leq 0.005$), but not Δ82-91 ($p = 0.12$), based on one-way ANOVA followed by Tukey's post-hoc test at 1 mg/mL.

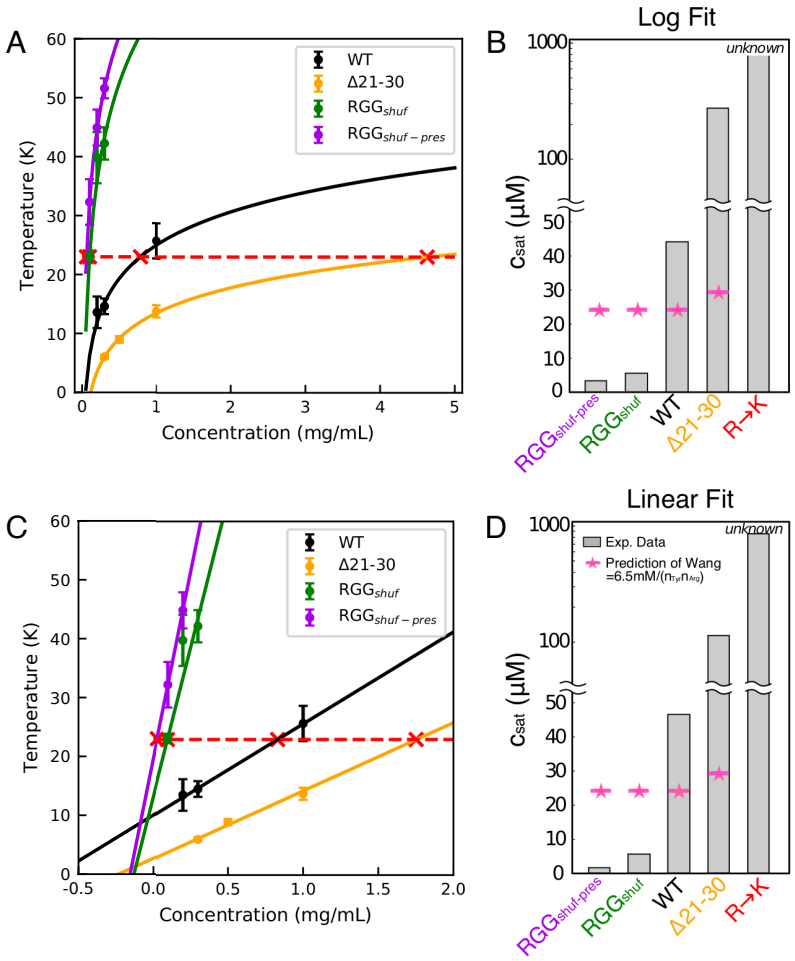
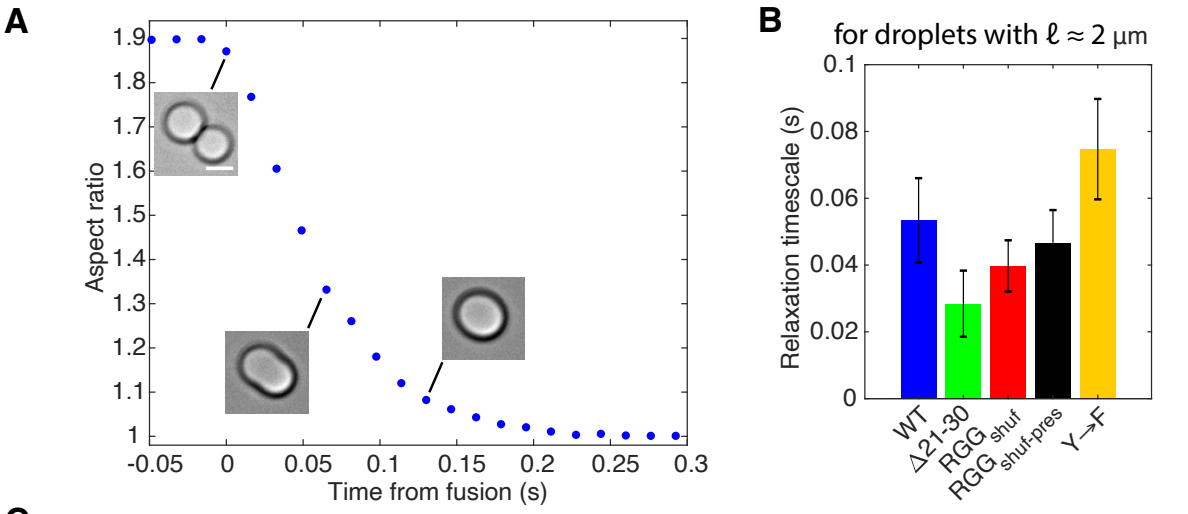


Figure S9: Fitting of c_{sat} from experimental data: A) Logarithmic fits to experimental data to calculate c_{sat} (red X's) at 23°C, (red dashed line). B) Bar plot of saturation concentrations for the different variants of RGG and comparison to empirical predictions using relationship from Wang et al.[26] C) Linear fits to experimental data to calculate c_{sat} as before. For $RGG_{shuf\text{-}pres}$, one data point was removed from the fitting such that the extrapolated c_{sat} value would be positive. D) Bar plot of saturation concentrations for the different variants of RGG using the linear fit and compared to empirical predictions using relationship from Wang et al. Error bars are STD with n = 3.



C

	Fusion: η/γ (s/μm)	FRAP: τ (s)	FRAP: D (μm²/s)
WT	0.018 ± 0.002	112 ± 45	0.025 ± 0.009
Δ21-30	0.033 ± 0.003	113 ± 20	0.022 ± 0.004
RGG_{shuf}	0.019 ± 0.004	238 ± 80	0.011 ± 0.003
$RGG_{shuf\text{-}pres}$	0.045 ± 0.006	244 ± 28	0.010 ± 0.001
Y→F	0.040 ± 0.006	142 ± 13	0.017 ± 0.002

Figure S10: Measurements of droplet material properties: A) Example trace showing aspect ratio of fusing droplets relaxing exponentially to a sphere, from which the relaxation timescale is calculated. The data shown corresponds to the Y→F droplet fusion event in Fig. 5A, several images of which are reproduced here as insets beside their corresponding data points (scale bar: 2 μm). B) Timescale of droplet fusion for droplets of lengthscale $\ell \approx 2\ \mu m$ (range 1.75-2.25 μm). Error bars represent STD (n ≥ 9). C) Table summarizing measurements of inverse capillary velocity η/γ from droplet fusion experiments, as well as recovery timescale τ and diffusivity D from FRAP.

4. Legends for Movies

Movie S1. Slab simulation of eIF4E binding domain, RYVPPHLR, in a simulation box of size 9.8 nm x 9.8 nm x 280 nm (total concentration 149.7 mg/ml) at 190 K.

Movie S2. Coexistence simulation of 100 chains of LAF-1 RGG_{shuf}, in a simulation box of size 250 nm x 250 nm x 250 nm at 250 K. One chain is highlighted by sequence composition where red is anionic, blue is cationic, white is hydrophobic, and green is polar. The highlighted chain diffuses freely through the protein-rich assembly.

Movie S3. Coexistence simulation of 100 chains of LAF-1 $RGG_{shuf\text{-}pres}$, in a simulation box of size 250 nm x 250 nm x 250 nm at 250 K. One chain is highlighted by sequence composition where red is anionic, blue is cationic, white is hydrophobic, and green is polar. The highlighted chain diffuses freely through the protein-rich assembly.

SI References

1. Gietz, R. D. & Sugino, A. New yeast-Escherichia coli shuttle vectors constructed with in vitro mutagenized yeast genes lacking six-base pair restriction sites. *Gene* **74**, 527–534 (1988).
2. Studier, F. W. Stable Expression Clones and Auto-Induction for Protein Production in E. coli. in *Methods in molecular biology (Clifton, N.J.)* **1091**, 17–32 (2014).
3. Zhang, T. *et al.* An improved method for whole protein extraction from yeast Saccharomyces cerevisiae. *Yeast* **28**, 795–798 (2011).
4. Bi, E. & Pringle, J. R. *ZDS1 and ZDS2, Genes Whose Products May Regulate Cdc42p in Saccharomyces cerevisiae*. *Molecular and Cellular Biology* **16**, (1996).
5. Schuster, B. S. *et al.* Controllable Protein Phase Separation and Modular Recruitment to Form Responsive, Membraneless Organelles. *Nat. Commun.* **9**, 2985 (2018).
6. Taylor, N. O., Wei, M. T., Stone, H. A. & Brangwynne, C. P. Quantifying Dynamics in Phase-Separated Condensates Using Fluorescence Recovery after Photobleaching. *Biophys. J.* **117**, 1285–1300 (2019).
7. Wei, M.-T. *et al.* Phase behaviour of disordered proteins underlying low density and high permeability of liquid organelles. *Nat Chem* **9**, 1118–1125 (2017).
8. Brangwynne, C. P., Mitchison, T. J. & Hyman, A. A. Active liquid-like behavior of nucleoli determines their size and shape in Xenopus laevis oocytes. *Proc. Natl. Acad. Sci.* **108**, 4334–4339 (2011).
9. Zhang, H. *et al.* RNA Controls PolyQ Protein Phase Transitions. *Mol. Cell* **60**, 220–230 (2015).
10. Dignon, G. L., Zheng, W., Kim, Y. C., Best, R. B. & Mittal, J. Sequence determinants of protein phase behavior from a coarse-grained model. *PLoS Comput. Biol.* **14**, e1005941 (2018).
11. Anderson, J. A., Lorenz, C. D. & Travesset, A. General purpose molecular dynamics simulations fully implemented on graphics processing units. *J. Comput. Phys.* **227**, 5342–5359 (2008).
12. Ryan, V. H. *et al.* Mechanistic View of hnRNPA2 Low-Complexity Domain Structure, Interactions, and Phase Separation Altered by Mutation and Arginine Methylation. *Mol. Cell* **39**, 465–479 (2018).
13. Murthy, A. C. *et al.* Molecular interactions underlying liquid−liquid phase separation of the FUS low-complexity domain. *Nat. Struct. Mol. Biol.* **26**, 637–648 (2019).
14. Best, R. B., Zheng, W. & Mittal, J. Balanced protein-water interactions improve properties of disordered proteins and non-specific protein association. *J. Chem. Theory Comput.* **10**, 5113–5124 (2014).
15. Abascal, J. L. & Vega, C. A general purpose model for the condensed phases of water: TIP4P/2005. *J. Chem. Phys.* **123**, 234505 (2005).
16. Luo, Y. & Roux, B. Simulation of osmotic pressure in concentrated aqueous salt solutions. *J. Phys. Chem. Lett.* **1**, 183–189 (2010).
17. Sugita, Y. & Okamoto, Y. Replica exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* **314**, 141–151 (1999).
18. Barducci, A., Bussi, G. & Parrinello, M. Well-tempered metadynamics: A smoothly converging and tunable free-energy method. *Phys. Rev. Lett.* **100**, 1–4 (2008).
19. Bonomi, M. & Parrinello, M. Enhanced sampling in the well-tempered ensemble. *Phys. Rev. Lett.* **104**, 1–4 (2010).
20. Hess, B., Kutzner, C., Van Der Spoel, D. & Lindahl, E. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **4**, 435–447 (2008).

21. Tribello, G. A., Bonomi, M., Branduardi, D., Camilloni, C. & Bussi, G. PLUMED 2: New feathers for an old bird. *Comput. Phys. Commun.* **185**, 604–613 (2014).

22. Tiwary, P. & Parrinello, M. A time-independent free energy estimator for metadynamics. *J. Phys. Chem. B* **119**, 736–742 (2015).

23. Vernon, R. M. *et al.* Pi-Pi contacts are an overlooked protein feature relevant to phase separation. *Elife* **7**, 1–48 (2018).

24. Dignon, G. L., Zheng, W., Best, R. B., Kim, Y. C. & Mittal, J. Relation between single-molecule properties and phase behavior of intrinsically disordered proteins. *Proc. Natl. Acad. Sci.* **115**, 9929–9934 (2018).

25. Dignon, G. L., Zheng, W., Kim, Y. C. & Mittal, J. Temperature-Controlled Liquid-Liquid Phase Separation of Disordered Proteins. *ACS Cent. Sci.* **5**, 821–830 (2019).

26. Wang, J. *et al.* A Molecular Grammar Governing the Driving Forces for Phase Separation of Prion-like RNA Binding Proteins. *Cell* **174**, 688–699 (2018).