

Supplemental Materials

Intrinsic disorder codes for leaps of protein expression

Chi-Ning Chuang¹, Tai-Ting Woo¹, Shih-Ying Tsai², Wan-Chen Li¹,
Chia-Ling Chen¹, Hou-Cheng Liu¹, Chiung-Ya Chen¹, Yi-Ping Hsueh¹, Ting-Fang Wang^{1,*}

¹Institute of Molecular Biology, Academia Sinica, Taipei 115, Taiwan

²Institute of Biological Chemistry, Academia Sinica, Taipei 115, Taiwan

*Corresponding author: E-mail: tfwang@gate.sinica.edu.tw

This supplemental information contains six tables (S1-S6) and three figures S1-S3.

Table S1. Sequences of IDRs used in this study

Name of SCDs or IDRs	Sequence
Hop1-SCD (258-324 a.a.)	MEKAGNTNFIRVDPFDLILQQQEENKLEESAPTK PQNFVTSQT TNVLGNLLNSSQASIQPTQFVSN
Rad51-NTD (1-66 a.a.)	MSQVQEQHISEQLQYGNGLSMSTVPADLSQSV VDGNGNGGSEDIEATNGSGDGGGLQEQAEOQE
Rad51-NTD-3SA	<u>M</u> AQVQEQHISE <u>A</u> QLQYGNGLSMSTVPADL <u>A</u> QSV VDGNGNGGSEDIEATNGSGDGGGLQEQAEOQE
Rad51-NTD-6SQA	<u>M</u> <u>A</u> <u>A</u> VQEQHISE <u>A</u> <u>A</u> LQYGNGLSMSTVPADL <u>A</u> <u>A</u> SV VDGNGNGGSEDIEATNGSGDGGGLQEQAEOQE
Rad51-NTD-9SQA	<u>M</u> <u>A</u> <u>A</u> <u>V</u> <u>A</u> <u>E</u> AHISE <u>A</u> <u>A</u> L <u>A</u> YGNGLSMSTVPADL <u>A</u> <u>A</u> SV VDGNGNGGSEDIEATNGSGDGGGLQEQAEOQE
Rad51-NTD-12SQA	<u>M</u> <u>A</u> <u>A</u> <u>V</u> <u>A</u> <u>E</u> AHISE <u>A</u> <u>A</u> L <u>A</u> YGNGLSMSTVPADL <u>A</u> <u>A</u> SV VDGNGNGGSEDIEATNGSGDGGGL <u>A</u> <u>E</u> <u>A</u> <u>E</u> <u>A</u> <u>A</u> <u>G</u> E
Rad53-SCD1 (1-29 a.a.)	MENITQPTQQSTQATQRFLIEKFSQEIQIG
Rad53-SCD1-5STA	MENI <u>A</u> Q <u>P</u> A <u>Q</u> Q <u>S</u> A <u>Q</u> A <u>A</u> QRFLIEK <u>F</u> AQEIQIG
Rad53-SCD1-7QA	MENIT <u>A</u> P <u>T</u> A <u>A</u> S <u>T</u> A <u>A</u> T <u>A</u> RFLIEK <u>F</u> S <u>A</u> E <u>A</u> I <u>G</u>
Rad53-SCD1-12STQA	MENI <u>A</u> <u>A</u> P <u>A</u> <u>A</u> <u>A</u> <u>S</u> <u>A</u> <u>A</u> <u>A</u> <u>A</u> <u>A</u> <u>A</u> RFLIEK <u>F</u> <u>A</u> <u>A</u> <u>E</u> <u>A</u> <u>I</u> <u>G</u>
Sml1-NTD (1-27 a.a.)	MQNSQDYFYAQNRSQQQQAPSTLRTVT
Sml1-NTD (1-50 a.a.)	MQNSQDYFYAQNRSQQQQAPSTLRTVTMAEFR VPLPPMAEVPMLSTQNS
Sup35-PND (1-39 a.a.)	MSDSNQGNNQQNYQQYSQNGNQQQGNNRYQG YQAYNAQA
Sup35-PND-S ¹⁷ A	MSDSNQGNNQQNYQQY <u>A</u> QNGNQQQGNNRYQG YQAYNAQA
Sup35-PND-3SA	<u>M</u> <u>A</u> <u>D</u> <u>A</u> NQGNNQQNYQQY <u>A</u> QNGNQQQGNNRYQ GYQAYNAQA
Sup35-PND-3QA	MSDSNQGNNQQNYQQYSQNGN <u>A</u> <u>A</u> <u>A</u> GNNRYQG YQAYNAQA
Sup35-PND-5QA	MSDSNQGNN <u>A</u> <u>A</u> <u>N</u> <u>Y</u> <u>A</u> <u>A</u> <u>Y</u> <u>S</u> ANGNQQQGNNRYQG YQAYNAQA
Sup35-PND-8QA	MSDSNQGNN <u>A</u> <u>A</u> <u>N</u> <u>Y</u> <u>A</u> <u>A</u> <u>Y</u> <u>S</u> ANGN <u>A</u> <u>A</u> <u>A</u> GNNRYQG YQAYNAQA
Sup35-PND-15SQA	<u>M</u> <u>A</u> <u>D</u> <u>A</u> <u>N</u> <u>A</u> GNN <u>A</u> <u>A</u> <u>N</u> <u>Y</u> <u>A</u> <u>A</u> <u>Y</u> <u>A</u> ANGN <u>A</u> <u>A</u> <u>A</u> GNNRY <u>A</u> GY <u>A</u> <u>A</u> <u>Y</u> <u>N</u> <u>A</u> <u>A</u> <u>A</u>
Sup35-PND-9NA	MSDS <u>A</u> Q <u>G</u> <u>A</u> <u>A</u> Q <u>Q</u> <u>A</u> YQQYSQ <u>A</u> <u>G</u> <u>A</u> Q <u>Q</u> <u>G</u> <u>A</u> <u>A</u> RYQG YQAY <u>A</u> <u>A</u> <u>Q</u> <u>A</u>
Sup35-PND-24SQNA	<u>M</u> <u>A</u> <u>D</u> <u>A</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>A</u> <u>A</u> <u>A</u> <u>Y</u> <u>A</u> <u>A</u> <u>Y</u> <u>A</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>A</u> <u>R</u> <u>Y</u> <u>A</u> GY <u>A</u> <u>A</u> <u>Y</u> <u>A</u> <u>A</u> <u>A</u> <u>A</u>
Sup35-PFD (1-114 a.a.)	MSDSNQGNNQQNYQQYSQNGNQQQGNNRYQG YQAYNAQAQPAGGYQNYQGYSGYQQGGYQQ YNPDAGYQQQYNPQGGYQQYNPQGGYQQQFNP QGGRGNYKNFNYNNSLQGYQ
New1-NPD (1-156 a.a.)	MPPKKFKDLNSFLDDQPKDPNLVASPFGGYFKNP AADAGSNNASKKSSYQQQRNWKQGGNYQQGG YQSYNSNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNK YNGQGYQKSTYKQSAVTPNQSGTPTPSASTTSLT SLNEKLSNLELTPISQFLSKIPECQS

Ure2-UPD (1-91 a.a.)	MMNNGNQNLSNALRQVNIGNRNSNTTDDQ SNINFEFSTGVNNNNNNSSNNNNNVQNNNSGR NGSQNNDNENNIKNTLEQHRQQQAF
Vps64-NTD (1-152 a.a.)	MVELEKRRRPPQLQHSPYVRDQSN SQGMTKTP ETSPPKRPMGRARSNSRSSGSRSNVDIDQYTIPPG LDLLPTASSPPSVHQVSQQQLSPILANKIRSPFEN QSQDQNDNSIDPTPAGQVTIPVEAVSPPALDELSK FQNGSTETLFRTG
Ssk2-NTD (1-191 a.a.)	MSHSDYFNYKPYGDSTEKPSSSKMRQSSSSSSSR LRSESLGRNSNTTQARVASSPISPLHSTQYFRSP NAVYSPGESPLNTVQLFNRLPGIPQGQFFHQNAIS GSSSSARSSRRPSNIGLPLPKNPQQSLPKLSTQPV PVHKKVEASKTESEIHKKPPVNSNQDPLLTPTL VISPELASLNTTNTSI
Ke11-NTD (1-69 a.a.)	MAGFSFAKKFTHKKHGKTPSDASISDQSREASLS TPPNEKFFTKQETPQKGRQFSQGYHSNVNKTSSP P

Table S2. Plasmids used in this study

Plasmid number	Description
pTFW2592	pYC2/NT-C, The mock control.
pTFW9957	<i>P_{RAD51}-LacZ-NVH</i>
pTFW9958	<i>P_{RAD51}-RAD51-NTD(1-66 a.a.)-LacZ-NVH</i>
pTFW9962	<i>P_{RAD51}-RAD53-SCD1(1-29 a.a.)-LacZ-NVH</i>
pTFW9963	<i>P_{RAD51}-HOPI1-SCD(258-324 a.a.)-LacZ-NVH</i>
pTFW10027	<i>P_{RAD51}-SML1-NTD(1-27 a.a.)-LacZ-NVH</i>
pTFW10028	<i>P_{RAD51}-SML1-NTD(1-50 a.a.)-LacZ-NVH</i>
pTFW10005	<i>P_{RAD51}-SUP35-PND(1-39 a.a.)-LacZ-NVH</i>
pTFW10034	<i>P_{RAD51}-sup35-PND-S¹⁷A-LacZ-NVH</i>
pTFW10221	<i>P_{RAD51}-SUP35-PFD(1-114 a.a.)-LacZ-NVH</i>
pTFW10169	<i>P_{RAD51}-NEW1-NPD(1-156 a.a.)-LacZ-NVH</i>
pTFW10177	<i>P_{RAD51}-URE2-UPD(1-91 a.a.)-LacZ-NVH</i>
pTFW10178	<i>P_{RAD51}-VPS64-NTD(1-152 a.a.)-LacZ-NVH</i>
pTFW10180	<i>P_{RAD51}-SSK2-NTD(1-191 a.a.)-LacZ-NVH</i>
pTFW10179	<i>P_{RAD51}-KEL1-NTD(1-69 a.a.)-LacZ-NVH</i>
pTFW9959	<i>P_{RAD51}-rad51-NTD-3SA-LacZ-NVH</i>
pTFW9993	<i>P_{RAD51}-rad51-NTD-6SQA-LacZ-NVH</i>
pTFW9994	<i>P_{RAD51}-rad51-NTD-9SQA-LacZ-NVH</i>
pTFW9995	<i>P_{RAD51}-rad51-NTD-12SQA-LacZ-NVH</i>
pTFW10000	<i>P_{RAD51}-rad53-SCD1-5STA-LacZ-NVH</i>
pTFW10001	<i>P_{RAD51}-rad53-SCD1-7QA-LacZ-NVH</i>
pTFW10002	<i>P_{RAD51}-rad53-SCD1-12STQA-LacZ-NVH</i>
pTFW10101	<i>P_{RAD51}-sup35-PND-3SA-LacZ-NVH</i>
pTFW10081	<i>P_{RAD51}-sup35-PND-3QA-LacZ-NVH</i>
pTFW10080	<i>P_{RAD51}-sup35-PND-5QA-LacZ-NVH</i>
pTFW10082	<i>P_{RAD51}-sup35-PND-8QA-LacZ-NVH</i>
pTFW10141	<i>P_{RAD51}-sup35-PND-15SQA-LacZ-NVH</i>
pTFW10142	<i>P_{RAD51}-sup35-PND-9NA-LacZ-NVH</i>
pTFW10143	<i>P_{RAD51}-sup35-PND-24SQNA-LacZ-NVH</i>
pTFW9974	<i>P_{RAD51}-GFP-NVH</i>
pTFW9975	<i>P_{RAD51}-RAD53-SCD1-GFP-NVH</i>
pTFW9976	<i>P_{RAD51}-GST-NVH</i>
pTFW9977	<i>P_{RAD51}-RAD53-SCD1-GST-NVH</i>
pTFW10003	<i>P_{RAD51}-GSTnd-NVH</i>
pTFW10004	<i>P_{RAD51}-RAD53-SCD1-GSTnd-NVH</i>

Table S3. Results of protein synthesis rate and IUPred prediction

Gene ID	Real protein synthesis rate [1]	Real protein synthesis per_mRNA [1]	Protein name	Protein length	Location of IDR [2, 3]
YFL014W	6.720136053	13.64122635	Hsp12	109	1-88 a.a.
YFR053C	7.99177396	10.89558219	Hxk1	485	1-30 a.a.
YER091C	35.72715777	5.714350831	Met6	767	0
YGR209C	21.52852161	2.143015389	Trx2	104	0
YPL061W	19.63688632	1.821993345	Adl6	500	0
YNL160W	1.975732553	1.81714069	Ygp1	354	0
YMR173W	5.051580741	1.3115736	Ddr48	430	1-430 a.a.
YLL024C	77.60251016	1.228181526	Ssa2	639	489-639 a.a.
YJR045C	20.8968953	1.223528285	Ssc1	654	513-654 a.a.
YLR259C	15.18869	1.047373969	Hsp60	572	551-572 a.a.
YGR254W	111.1911365	1.019383383	Eno1	437	0
YNL134C	1.485230878	0.930520671	YNL134C	376	0
YPL154C	5.414888063	0.882116345	Pep4	405	0
YOR184W	6.481168336	0.826375427	Ser1	395	1-22 a.a.
YDL229W	53.91880257	0.787528479	Ssb1	613	0
YDL185W	8.154896654	0.761063157	Vma1	1071	0
YGL253W	27.58023565	0.735193314	Hxk2	486	1-25 a.a.
YOR133W	27.65305609	0.732786289	Eft1	842	0
YDR385W	27.65305609	0.706330957	Eft2	842	0
YAL005C	22.23939511	0.704638842	Ssa1	642	489-642 a.a.
YML058W	0.552920114	0.028219389	Sml1	104	1-104 a.a.
YHR158C	0.063195461	0.037700276	Kel1	1164	1-113 a.a.
YPL226W	0.73846409	0.073673651	New1	1196	1-132 a.a.

Table S4. Predicted protein folding rates

Protein	Protein length	Protein folding rates $\ln(k_f)$ predicted by the “mixed (α -helices and β -sheets)” mode of FOLD-RATE [4]
LacZ-NVH	1060 a.a.	1.52/sec
Rad51-NTD-LacZ-NVH	1135 a.a.	7.79/sec
Rad51-NTD(3SA)-LacZ-NVH	1135 a.a.	14.10/sec
Rad51-NTD(6SQA)-LacZ-NVH	1135 a.a.	1.52/sec
Rad51-NTD(9SQA)-LacZ-NVH	1135 a.a.	19.00/sec
Rad51-NTD(12SQA)-LacZ-NVH	1135 a.a.	7.76/sec
Rad53-SCD1-LacZ-NVH	1092 a.a.	11.40/sec
Rad53-SCD1(5STA)-LacZ-NVH	1092 a.a.	19.60/sec
Rad53-SCD1(7QA)-LacZ-NVH	1092 a.a.	8.03/sec
Rad53-SCD1(12STQA)-LacZ-NVH	1092 a.a.	9.63/sec
Hop1-SCD-LacZ-NVH	1130 a.a.	2.62/sec
Sml1-NTD(1-27 a.a.)-LacZ-NVH	1090 a.a.	-3.56/sec
Sml1-NTD(1-50 a.a.)-LacZ-NVH	1113 a.a.	22.90/sec
Sup35-PND-LacZ-NVH	1102 a.a.	8.24/sec
Sup35-PND-S ¹⁷ A-LACZ-NVH	1102 a.a.	6.05/sec
Sup35-PND-3SA-LACZ-NVH	1102 a.a.	0.50/sec.
Sup35-PND-3QA-LACZ-NVH	1102 a.a.	18.00/sec
Sup35-PND-5QA-LACZ-NVH	1102 a.a.	0.13/sec
Sup35-PND-8QA-LACZ-NVH	1102 a.a.	9.79/sec
Sup35-PND-15SQA-LACZ-NVH	1102 a.a.	23.80/sec.
Sup35-PND-9NA-LACZ-NVH	1102 a.a.	21.60/sec
Sup35-PND-24SQNA-LACZ-NVH	1102 a.a.	9.98/sec
Sup35-PFD-LacZ-NVH	1177 a.a.	15/sec
Ure2-UPD-LacZ-NVH	1154 a.a.	19.90/sec
New1-NTD-LacZ-NVH	1219 a.a.	-0.21/sec
Vps64-NTD-LacZ-NVH	1215 a.a.	12.50/sec
Kel1-NTD-LacZ-NVH	1131 a.a.	17.20/sec
Ssk2-NTD-LacZ-NVH	1254 a.a.	13.70/sec
GFP-NVH	279 a.a.	4.98/sec
Rad53-SCD1-GFP-NVH	310 a.a.	18.50/sec
GST-NVH	264 a.a.	-9.14/sec
Rad53-SCD1-GST-NVH	295 a.a.	1.62/sec
GSTnd-NVH	264 a.a.	-5.23/sec
Rad53-SCD1-GSTnd-NVH	295 a.a.	1.45/sec
Rad51- Δ N	334 a.a.	1.09/sec
Rad51	400 a.a.	1.43/sec
Rad53-SCD1-Rad51- Δ N	365 a.a.	2.71/sec
Rad53-SCD1(5STA)-Rad51- Δ N	365 a.a.	2.70/sec

Sup35-PND-Rad51-ΔN	375 a.a.	0.84/sec
--------------------	----------	----------

Table S5. *S. cerevisiae* strains used in this study

Name	Genotype
WHY13008	<i>MATa, ho, leu2, ura3, his4-X::LEU2-(NgoMIV;+ori)-URA3, ERG1(SpeI), RAD51::hphMX4</i>
WHY13283	<i>MATa, ho, leu2, ura3, his4-X::LEU2-(NgoMIV;+ori)-URA3, ERG1(SpeI), rad51Δ::hphMX4</i>
WHY13416	<i>MATa, ho, leu2, ura3, his4-X::LEU2-(NgoMIV;+ori)-URA3, ERG1(SpeI), rad51ΔN::hphMX4</i>
WHY13744	<i>MATα, ho::LYS2, leu2, ura3, lys2, HIS4::LEU2-(BamHI;+ori), Erg1(Sall), dmc1::kanMX4, RAD53^{SCD1}-rad51ΔN::hphMX4</i>
WHY13743	<i>MATa, ho, leu2, ura3, his4-X::LEU2-(NgoMIV;+ori)-URA3, ERG1(SpeI), dmc1::kanMX4, rad53^{SCD1-5STA}-rad51ΔN::hphMX4</i>
WHY13741	<i>MATα, ho::LYS2, leu2, ura3, lys2, HIS4::LEU2-(BamHI;+ori), ERG1(Sall), SUP35^{PND}-rad51ΔN::hphMX4</i>
WHY10271	<i>MATa, ho::hisG, lys2, leu2::hisG, arg4-nsp, ura3</i>
WHY13970	<i>MATa his3Δ1, leu2Δ0, met15Δ0, ura3Δ0</i>
WHY13785	<i>MATa his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, hsp104::kanMX4</i>
WHY14126	<i>MATa his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, new1::kanMX4</i>
WHY14129	<i>MATa his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, doa4::kanMX4</i>
WHY14227	<i>MATa his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, doa1::kanMX4</i>
WHY13989	<i>MATa his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, san1::kanMX4</i>
WHY14132	<i>MATa his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, oaz1::kanMX4</i>

Table S6. Primers used for g-qPCR and RT-qPCR

Gene	Primer Polarity	Name	Sequence
<i>ACT1</i>	Sense	PA11240	5'-CCACCACTGCTGAAAGAGAAATTGT-3'
	Antisense	PA11241	5'-CTTGACCATCTGGAAGTTCGTAGGA-3'
<i>LacZ</i>	Sense	PA11246	5'-CCGCCGTTTGTGCCACGGA-3'
	Antisense	PA11247	5'-CCATCACCGCGAGGCGGTTT-3'

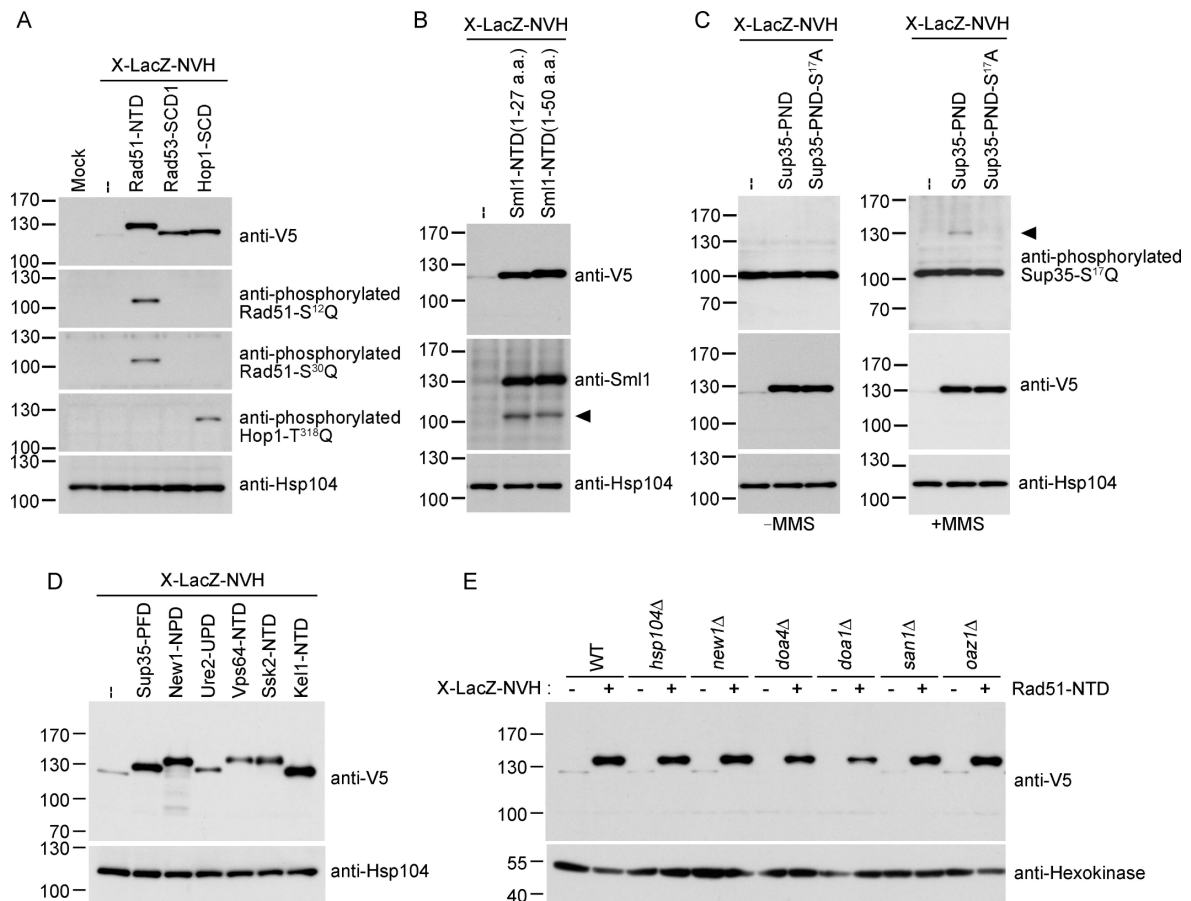


Figure S1. Western blots for visualization of LacZ-NVH fusion proteins. (A-D) N-terminal fusion of four DDR-SCDs, three prion domains and three N-IDRs promotes high-level expression of LacZ-NVH. Black arrowheads indicate degraded Sml1-NTD-LacZ-NVH (B) and phosphorylated Sup35-PND-S¹⁷Q-LacZ-NVH (C), respectively. (E) The steady-state protein levels of Rad51-NTD-LacZ-NVH and LacZ-NVH in WT and six protein homeostasis gene knockout mutants.

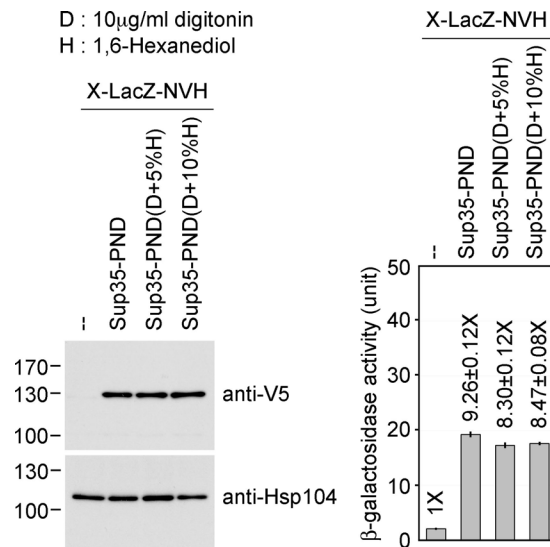


Figure S2. 1,6-HD does not affect the expression-promoting function of IDRs. Yeast cells expressing Sup35-PND-LacZ-NVH either lacked treatment (i.e., no drug) or were grown in the presence of 5% 1,6-HD and 10 μ g/ml digitonin or 10% 1,6-HD and 10 μ g/ml digitonin for 30 min. Digitonin makes yeast cells more permeable to 1,6-HD [5]. The Sup35-PND-LacZ-NVH proteins were visualized by immunoblotting with anti-V5 antisera. Hsp104 was used as a loading control. Quantitative yeast β -galactosidase (LacZ) assays were carried out as described previously [6, 7].

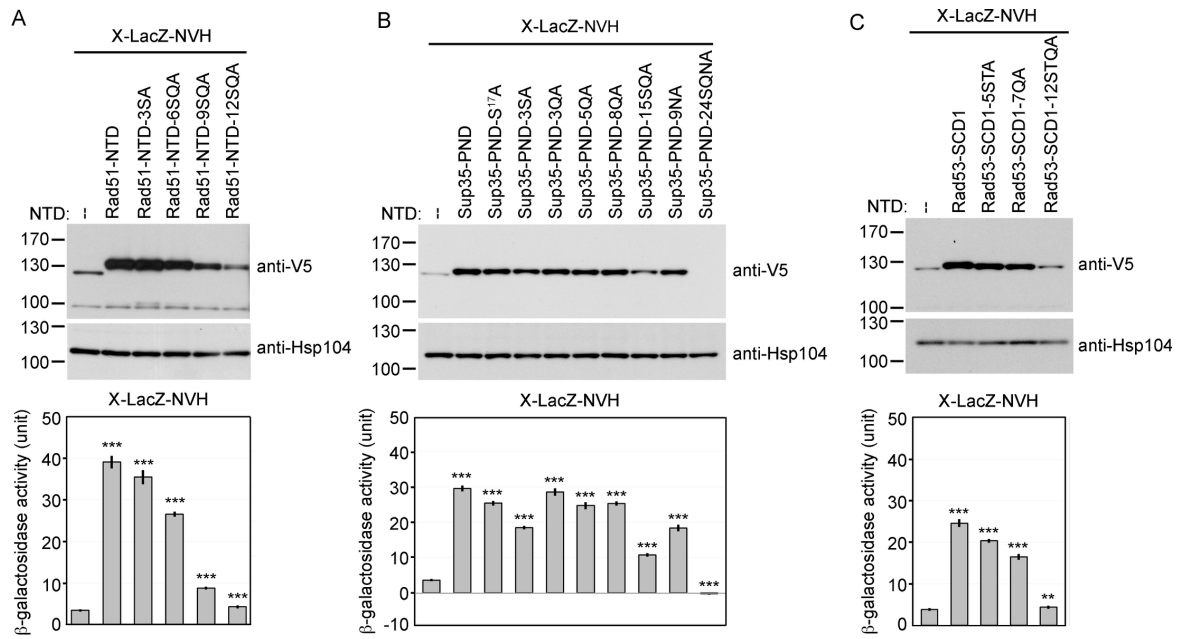


Figure S3. Alanine scanning mutagenesis of IDRs. The amino acid sequences of WT and mutant IDRs are listed in Table S1. Total protein lysates prepared from yeast cells expressing Rad51-NTD-LacZ-NVH (A), Sup35-PND-LacZ-NVH (B) or Rad53-SCD1-LacZ-NVH (C) were visualized by immunoblotting with anti-V5 antisera. Hsp104 was used as a loading control. Quantitative yeast β -galactosidase (LacZ) assays were carried out as described previously [6, 7]. Error bars indicate standard deviation between experiments ($n = 3$). Asterisks indicate significant differences when compared to LacZ-NVH with p values calculated with a two-tailed t-test (**, P value <0.01 and ***, P value <0.001).

References

1. Riba A, Di Nanni N, Mittal N, Arhne E, Schmidt A, Zavolan M. Protein synthesis rates and ribosome occupancies reveal determinants of translation elongation rates. *Proc Natl Acad Sci USA*. **116**,15023-15032 (2019).
2. Meszaros B, Erdos G, Dosztanyi Z. IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding. *Nucleic Acids Res*. 46 (W1), W329-W37 (2019).
3. Erdos G, Dosztanyi Z. Analyzing protein disorder with IUPred2A. *Curr Protoc Bioinformatics*. 70, e99 (2020).
4. Gromiha MM, Thangakani AM, Selvaraj S. FOLD-RATE: prediction of protein folding rates from amino acid sequence. *Nucleic Acids Res*. 34, W70-74 (2006).
5. A. E. Posey, A. S. Holehouse, R. V. Pappu, Phase separation of intrinsically disordered proteins. *Methods Enzymol* **611**, 1-30 (2018).
6. F. M. Lin, Y. J. Lai, H. J. Shen, Y. H. Cheng, T. F. Wang, Yeast axial-element protein, Red1, binds SUMO chains to promote meiotic interhomologue recombination and chromosome synapsis. *EMBO J* **29**, 586-596 (2010).
7. T. T. Woo, C. N. Chuang, M. Higashide, A. Shinohara, T. F. Wang, Dual roles of yeast Rad51 N-terminal domain in repairing DNA double-strand breaks. *Nucleic Acids Res* **48**, 8474-8489 (2020).