Structural proteomics defines a sequential priming mechanism for the progesterone receptor

Matthew D. Mann<sup>1,2</sup>, Min Wang<sup>3</sup>, Josephine C. Ferreon<sup>5</sup>, Michael P. Suess<sup>3</sup>, Antrix Jain<sup>4</sup> Anna Malovannaya<sup>5</sup>, Bruce D. Pascal<sup>6</sup>, Raj Kumar<sup>7</sup>, Dean P. Edwards<sup>3</sup>, and Patrick R. Griffin<sup>\*,1,2</sup>

<sup>1</sup>Skaggs Graduate School of Chemical and Biological Sciences, The Scripps Research Institute, Jupiter, FL, USA.

<sup>2</sup>Department of Molecular Medicine, The Herbert Wertheim UF Scripps Institute for Biomedical Innovation & Technology, Jupiter, FL, 33458

<sup>3</sup> Department of Molecular and Cellular Biology, Baylor College of Medicine, Houston, TX, 77030 USA.
<sup>4</sup> Mass Spectrometry Proteomics Core Facility. Advanced Technology Cores, Baylor College of Medicine, Houstonm TX, 77030

<sup>5</sup> Verna and Marrs McLean Department of Biochemistry and Molecular Pharmacology, Baylor College of Medicine, Houston, TX, 77030

<sup>6</sup> Omics Informatics LLC. 1050 Bishop Street #517, Honolulu, HI 96813

<sup>7</sup> Department of Pharmaceutical and Biomedical Sciences, Touro College of Pharmacy, Touro University, New York, NY, USA.

\* Correspondence should be addressed to <u>pgriffin2@ufl.edu</u> <u>Running title: Structural proteomics reveals a unique activation model for PR-A and PR-B</u>

<u>Acknowledgements:</u> This work was supported by NIH-NCI R01 (CA263574) to MPIs, PRG, RK, and DPE. The authors acknowledge the expert assistance of the Recombinant Protein Production and Characterization Core (RPPCC) for expression of recombinant proteins in the baculovirus insect cell system and the Mass Spectrometry Proteomcs (MSPC) Core for amino acid sequencing and phosphorylation analysis of recombinant proteins. The RPPCC and MSPC at BCM are supported by the NCI Cancer Center Support Grant (P30 CA125123) of the Dan L Duncan Comprehensive Cancer Center and the MSPC is additionally supported by a CPRIT (Cancer Prevention and Research Institute of Texas) Core Facility Support Award (RP210227). Additional support is from NIH S10 Shared Instrument Grant (1S100D030276) to JCF for the SEC-MALS instruments.

## Keywords:

Progesterone receptor; hydrogen-deuterium exchange; crosslinking; mass spectrometry; proteinprotein interactions; nuclear receptors, transcriptional co-regulatory proteins **Supplemental Figures** 

- 1. Affinity/SEC two-step purification of SII-tagged PR
- 2. Mass Spectrometry sequencing of purified major PR-A and PR-B bands
- 3. Mass Spectrometry sequencing of minor smaller molecular sized band product with PR-B
- 4. Affinity/SEC two-step purification for SRC3 tagged SII
- 5. Affinity/SEC two-step purification for p300 tagged SII
- 6. Mass Spectrometry of purified SRC3 and p300
- 7. SEC MALS of PR-A/SRC3/p300/DNA Complex
- 8. Differential scanning fluorimetry (DSF) analysis of PR-A with and without DNA
- 9. Intrinsic deuterium exchange for PR-A and PR-B
- 10. Workflow for PR:SRC3 model generation and selection
- 11. PR-B vs. PR-B:PRE consolidated HDX-MS data

## Supplemental Tables

- 1. SEC-MALS Experimental vs Theoretical MWs of individual purified proteins and various mixtures in the presence and absence of DNA
- 2. Transition melting temperatures (T  $_{M}$ ) of each purified protein as determined by DSF



**Fig S1. Affinity/SEC two-step purification of SII-tagged PR.** SII-tagged PR-A and PR-B(bound to progestin agonist R5020) were expressed in Sf9 cells and cell lysates were prepared and purification performed by two -step affinity Strep Trap XT column and size exclusion chromatography (SEC) S200 as described in Methods. A) SDS-PAGE analysis of PR-A affinity purification fractions T, total cell lysate, P pellet after centrifugation lysate, S supernatant after centrifugation, F/T flow through after binding to Strep TrapXT, M, protein molecular weight markers. Lanes A6-B12 (5uL) biotin eluted fractions with major enriched protein bands at expected size of ~83 kDa for PR-A. B) SEC fractionation of pooled fractions A6-B12 off the Strep Trap XT column. OD tracing at 280nM showsmajor peak 2 of monomeric size for PR-A and peak 1 containing high molecular weight aggregation. Peak 2 fractions were pooled, concentrated and analyzed SDS-PAGE gel showing a single band of >98% purity of the expected size of PR-A. C) A) SDS-PAGE analysis of PR-B affinity purification fractions and D) SEC second step purification by the same set-up and as PR-A except major band and purified product are expected size for monomeric PR-B (~100kDa).



**Fig S2. Mass Spectrometry sequencing of purified major PR-A and PR-B bands.** In-gel digestion of purified major bands for A) PR-A and B) PR-B (bound to progestin agonist R5020) was performed, followed by LC-MS/MS analysis. The peptide sequences highlighted in green were identified by MS, covering nearly the entire tryptic region



В



**Fig S3. Mass Spectrometry sequencing of minor smaller molecular sized band product with PR-B.** A) In-gel digestion of smaller sized PR-B band (bound progestin agonist R5020) was performed, followed by LC-MS/MS analysis. The peptide sequences highlighted in green were identified by MS, suggesting intact full-length PR-B protein. B) The peptide area under curve (AUC) for selected phosphorylated PR-B peptides from major and minor band suggests variable phosphorylation level. The peptide AUC was normalized to the total protein amount in that band.



**Fig S4.** Affinity/SEC two-step purification for SRC3 tagged SII. SII-tagged SRC3 was expressed in Sf9 insect cells and purification performed by two -step affinity by StrepTrap XT and size exclusion chromatography (SEC) S200 as described in Methods. A) SDS-PAGE analysis of affinity purification fractions T, total cell lysate, P pellet after centrifugation, S supernatant after centrifugation, F/T flow through after binding to Strep TrapXT M, protein molecular weight markers. Lanes E9-G3 are SDS-PAGE analysis (5uL) of biotin eluted fractions with a major enriched band of expected size of ~158kDa for SRC3. B) SEC (S200) fractionation of pooled fractions E11-G3 off the StrepTrap XT column. OD tracing at 280nM shows a major peak (peak 1) of the native size expected for monomeric SRC3 and a minor peak (peak 2) of lower molecular weight material. Peak 1 fractions were pooled, concentrated

A

and analyzed by SDS-PAGE(left panel) showing a single band of >98% purity of the expected size of monomeric SRC3



**FigS5 Affinity/SEC two-step purification for p300 tagged SII.** SII-tagged p300 was expressed in Sf9 insect cells and purification performed by two -step affinity by StrepTrap XT and size exclusion chromatography (SEC) S200 as described in Methods. *A)* SDS-PAGE analysis of affinity purification fractions T, total cell lysate, P pellet after centrifugation, S supernatant after centrifugation, F/T flow through after binding to Strep TrapXT M, protein molecular weight markers. Lanes A2-A11 are SDS-PAGE (5uL) of biotin eluted fractions with a major enriched band of the expected size for monomeric p300 (~267kDa). *B)* SEC (S200) fractionation of pooled fractions A3-A8 from Strep TrapXT shows a single major peak (OD tracing at 280nM) of expected size for native monomeric p300. The pooled SEC fractions from peak 1 on SDS-PAGE (left panel) analysis shows a single major band of the expected size for monomeric p300.



В

	1	11	21	31	41	51	61	71	81	91	101	111	121	131
1	мнининини	HGSGAENVVE	PGPPSAKRPK	LSSPALSASA	SDGTDFGSLF	DLEHDLPDEL	INSTELGLTN	GCDINOLOTS	LGMVQDAASK	HKQLSELLRS	GSSPNLNMGV	GGPGQVMASQ	AQOSSPGLGL	INSMVKSPMT
141	QAGLTSPNMG	MGTSGPNQGP	TOSTGMANSP	VNOPAMGMNT	GMNAGMNPGM	LAAGNGOGIM	PNOVMNGSIG	AGRGRONMOY	PNPGMGSAGN	LLTEPLOOGS	POMGGOTGLR	GPOPLKMCMM	NNPNPYGSPY	TONPGOOIGA
281	SGLGLQIQTK	TVLSNNLSPF	AMDKKAVPGG	GMPNMGQQPA	POVOOPGLVT	PVADGMGSGA	HTADPEKRKL	IQQQLVLLLH	AHKCORREGA	NGEVROCNLP	HCRTMANVLN	HMTHCQSGKS	COVAHCASSR	QIISHWKNCT
421	RHDCPVCLPL	KNAGDKRNOO	PILTGAPVGL	GNPSSLGVGQ	QSAPNLSTVS	QIDPSSIERA	YAALGLPYOV	NOMPTOPOVO	AKNOONOOPG	QSPQGMRPMS	NMSAS PMGVN	GOVOVOTPSL	LSDSMLHSAI	NSONPAMSEN
561	ASVPSLGPMP	TAAOPSTTGI	RKOWHEDITO	DLRNHLVHKL	VOAIFPTPDP	AALKORRMEN	LVAYARKVEG	DMYESANNRA	EYYHLLAEKI	AKIÖKETEEK	RRTRLOKONM	LPNAAGMVPV	SMNPGPNMGQ	POPCMTSNGP
701	LPDPSMIRGS	VPNOMMPRIT	POSGLNOFCO	MSMAQPPIVP	ROTPPLOHHG	QLAQPGALNP	PMGYGPRMQQ	PSNQGQFLPQ	TOFPSOGMNV	TNIPLAPSSG	QAPVSQAQMS	SSSCPVNSPI	MPPGSQGSHI	HCPQLPQPAL
841	HQNSPSPVPS	RTPTPHHTPP	SIGACOPPAT	TIPAPVPTPP	AMPPGPQSQA	LHPPPROTPT	PPTTQLPQQV	OPSLPAAPSA	DOPOCOPRSO	QSTAASVPTP	TAPLLPPOPA	TPLSOPAVSI	EGOVSNPPST	SSTEVNSOAL
981	AEKOPSOEVK	MEAKMEVDOP	EPADTOPEDI	SESKVEDCKM	ESTETEERST	ELKTEIKEEE	DOPSTSATOS	SPAPGOSKKK	IFKPEELROA	LMPTLEALYR	ODPESLPFRO	PVDPQLLGIP	DYFDIVKSPM	DISTIKRKLD
1121	TGOYOF PWOY	VDDIWLMENN	AWLYNDKTSD	VYKYCSKLSE	VEROETDRVM	OSLOVCCORK	LEESPOTLCC	YCKOLCTIPR	DATYYSYONE	VHECENCENE	TOCESVELCD	DESOPOTTIN	KEOFSKRKND	TLOPELEVEC
1261	TROCEMON	CULUNETTWD	ACEVEDCELV	VENDTOVENV	FEAVDLOSTD	I CTELENEUN	DELERONHER	SCRUTUPUTU	ACOVENENCE	CHICADEMOSC	PHARSEDVAT	VALEAFERID	CONTRACTOR	VORVESDERR
	- Theorem and		AGE VEDGEDA		- CHREATER		Di D			0100017000			B CONTRACT	
1401	PNURRVIISI	LDSVHFFRPR	CLRTAVIHEI	LIGILEIVKK	LGITTGHIWA	CPPSEGDDYI	FHCHPPDURI	PRPERCUEWI	KEMLDRAVSE	RIVHDIKDIF	RUATEDRETS	ARELPIFEGD	FWPNVLEESI	RELEUEEEER
1541	KREENTSNES	TDVTKGDSKN	AKKKNNKKTS	KNKSSLSRGN	KKKPGMPNVS	NDLSOKLYAT	MEKHKEVFFV	IRLIAGPAAN	SLPPIVDPDP	LIPCDLMDGR	DAFLTLARDK	HLEFSSLRRA	QWSTMCMLVE	LHTQSQDRFV
1681	YTCNECKHHV	ETRWHCTVCE	DYDLCITCYN	TKNHDHKMEK	LGLGLDDESN	NOCAAATOSP	GDSRRLSIQR	CIQSLVHACO	CRNANCSLPS	COKMKRVVQH	TKGCKRKTNG	GCPICKOLIA	LCCYHAKHCQ	ENKCPVPFCL
1821	NIKQKLR	LOHRLOOACM	LRRRMASMOR	TOVVODDOCL	PSPTPATPTT	PTGOOPTTPO	TPOPTSOPOP	TPPNSMPPYL	PRTOAACPVS	QCKAACQVTP	РТРРОТАОРР	LPGPPPAAVE	MAMOIORAAE	TOROMAHVOI
1961	FORPIOHOMP	PMTPMAPMGM	NPPPMTRGPS	GHLEPGMGPT	GMOOOPPWSO	GGLPOPOOLO	SCMPRPAMMS	VAQHGQPLNM	APOPGLOOVG	ISPLKPGTVS	QUALQNLLRT	LRSPSSPLOO	QQVLSILHAN	POLLAAFIKQ
2101	RAAKYANSNP	OPIPGOPCMP	QCOPGLOPPT	MPGOOGVHSN	PAMONIMPMO	AGVORAGLEO	OOPOOOLOPP	MGGMSPQAQQ	MANNHNTMPS	OFRDILRROO	18400000GA	GPGIGPGMAN	HNOFOOPOGV	GYPPODOCRM
2241		NMGQIGQLPQ	ALGAEAGASL	CAYOORLLOO	QMGSPVQPNP	MSPQQHMLPN	QAQSPHLQCQ	QIPNSLSNQV	RSPOPVPSPR	PQSQPPHSSP	SPRMQPQPSP	HHVSPQISSP	HPGLVAAQAN	PMEQGHFASP
2381	DQNSMLSQLA	SNPGMANLHG	ASATDLGLST	DNSDLNSNLS	QSTLDINGGS	AWSHPQFEK								

**Fig S6. Mass Spectrometry of purified SRC3 and p300.** In-gel digestion of purified A) SRC3 and B) p300 bands was performed, followed by LC-MS/MS analysis. The peptide sequences highlighted in green were identified by MS, covering nearly the entire tryptic region.



**Fig S7 SEC-MALS chromatograms with molar mass distribution for various proteins (PR-A, SRC3, p300) and DNA alone or in complexes.** Shown are double y-plots (left y-axis, molar mass; right y-axis, UV absorbance at 280 nm) vs time (x-axis). Molar mass distribution is displayed as gray dots across the peaks. PR-A (with agonist R5020) assembles only to a dimer in the presence of DNA. A multi-complex of p300:SRC3:PR-A:DNA at a 1:1:2:1 ratio is observed (top chromatogram). The presence of DNA in the complex is confirmed by deconvolution of the protein and DNA fractions in the peak. The full complex showed a heterogeneous distribution of molar mass due to overlap with the p300 monomeric peak.

Table S1. SEC-MALS Experimental vs Theoretical MWs of individual purified proteins and various mixtures in the presence and absence of DNA.

Table 1. SEC-MALS derived molecular weights for proteins, DNA and the complexes formed											
	Theoretical MW	Actual MW <sup>a</sup>									
Protein/DNA		byUV	by dRI								
DNA	19.6	18.2±0.9	20.1 ± 1.0								
PR-A(R5020)	83.6	94.9±4.8	91.2±4.6								
PR-A(RU486)	83.6	98.8±4.7	95.3±4.8								
PR-B(R5020)	100.4	116.4±5.8	101.6± 5.1								
PR-B(RU486)	100.4	96.3±4.8	97.7 ± 4.10								
p300	267	249.2 ± 12.5	200.5±9.2								
SRC3	158	170.3±8.5	159.7±8								

Complexes Theoretical MW Actual MW PR-A(R5020):DNA (2:1) 186.8  $161.6 \pm 8.1$ Protein component 167.2  $144.8 \pm 7.2$ **DNA** component 19.6  $16.8 \pm 0.8$ PR-A(RU486):DNA (2:1) 186.8  $189.3 \pm 9.5$ Protein component  $167.5 \pm 8.4$ 167.2 **DNA** component 19.6 21.8±1.1 PR-B(R5020):DNA (2:1) 224.0 ± 11.2 220.4 201.7±10.1 Protein component 200.8  $22.3 \pm 1.1$ **DNA** component 19.6 PR-B(RU486):DNA (2:1) 221.8 ± 11.1 220.4 197.7±9.9 Protein component 200.8  $24.1 \pm 1.2$ **DNA** component 19.6 p300:SRC3:PR-A(R5020):DNA (1:1:2:1)  $536.5 \pm 26.8^{b}$ 611.8 592.2 512.7 ± 25.6 Protein component **DNA** component 23.7 ± 1.2 19.6 PR-A(R5020):DNA (2:1)<sup>c</sup> 186.8  $190.7 \pm 9.5$ 161.4 ± 8.1 Protein component 167.2 **DNA** component 19.6 29.3 ± 1.5

<sup>a</sup> errors are 5% of calculated MW; dn/dc based off BSA values in similar buffer condition

<sup>b</sup> heterogeneous MW distribution, overlap with p300

<sup>c</sup> in same buffer conditions as above but with OM Urea



Fig S8. Differential scanning fluorimetry (DSF) analysis of PR-A with and without DNA. DSF analysis was performed with 4  $\mu$ M PR-A and 25× SYPRO Orange in 20 mM HEPES, pH 7.5, 200mM NaCl, 5% Glycerol, 1M Urea, 1mM TCEP with (purple) or without 3uM PRE DNA (yellow), compared with the same concentration of TLX-WT as a positive control (blue). Fluorescent intensities and the derivative of the fluorescent intensities (dF/dT) were plotted against temperature (Tm).

Table S2. Transition melting temperatures (T  $_{M}$ ) of each purified protein as determined by DSF. Samples were analyzed as three technical replicates and the average Tm was calculated with standard error of the mean.

Protein	T <sub>m</sub> (°C)
PR-A	40.7±0.3
PR-A + DNA	42.2±0.3
PR-B	43.5±0
PR-B + DNA	45.0±0
p300	49.1±0
SRC3	46.5±0.1



**Figure S9**. **Intrinsic deuterium exchange for PR-A and PR-B**. Woods plots showing the percent deuterium exchange (%D Uptake) for both PR-A (blue) and PR-B (red) across each residue across the protein in both non-DNA-bound (Apo) and DNA-bound (PRE) states. Graphs made using GraphPad Prism 10.



**Figure S10 Workflow for PR:SRC3 model generation and selection**. Workflow detailing the model generation for ternary complexes. AlphaFold3.0 webserver was used to generate PR:SRC3 models in quintuplicate. Models were ensemble reweighted using HDXer (see methods) to determine best representation based on HDX data.

		10		<b>20</b>		:	30		40		50		6	0		70		80 8		90	. 1	.00		110	120
MTEL	KAKG	PRA	рнуадо	, P P S E	EVGS	<b>P L L C</b>	RPAAG	PFPG	SQTSD	TLPEV	SAIP	ISL	DGLLF	PRPC	QGQD	PSDEK	тород	SLSDVE	GAY	SRAEATR	GAGGSS	SSPP	EKDSG	LLDSVL	DTLLAPSGP
				н II.																					
	130		14	0		150		160		170			180		190		200		210		220		230		240
GQSQ	9 S P P	ACE	<b>VTSSW</b>	LFGE	ELPE	DPPA	арато	RVLS	PLMSR	SGCKV	GDSS	GTA	ААНКУ	LPRG	LSPA	RQLLL	PASES	PHWSGA	<b>BAK</b>	PSPQAAA	<b>A E A E E E</b>	DGSE	SEESA	GPLLKG	KPRALGGAA
250			260		270		280		290	) <u> </u>	3	00		310		320		330		340		350		360	370
AGGG	AAAV	PPG	AAAGG	ALVE	KEDS	RFSA	PRVAL	V E Q D	армар	GRSPL	АТТУ	MDF	тнаьт	LPLN	HALL	AARTR	QLLED	ESYDGG	AGA	ASAFAPP	RSSPCA	SSTP	VAVGD	FPDCAY	PPDAEPKDD
	з	80		390		400		43	L0 .	4	20		430		44	0	42	50		460	470		480		490
AYPL	YSDF	QРР	ALKIKI	EEEG	AEAS	ARSP	RSYLV	AGAN	PAAFP	DFPLG		LPP	RATPS	RPGE	AAVT	AAPAS	ASVSS	ASSSGS	TLE	CILYKAE	GAPPQQ	GPFA	рррск	APGASG	CLLPRDGLP
50	0		510		520		5	30		540		550		5	60		570	, 5	80	, 59	0	6	00	610	) , 6:
STSA	SAAA	AGA	APALYI	ALGI	NGLP	D L C X	Q A A V I	KEGL	ЬÕАХБ	PYLNY	LRPD	SEA	sqspq	YSFE	SLPQ	KICLI	CGDEA	ѕҁснүс	VLT	CGSCKVF	FKRAME	СОНИ	YLCAG	RNDCIV	DKIRRKNCP
		630		64	0 .		<b>Ş50</b>		<b>Ģ60</b>		670		6	80		690		700		710	. 3	20		730	740
ACRL	RKCC	QAG	мүгссі	RKFKF	FNKV	RVVR	ALDAV	ALPQ	<b>BACAB</b>	NESQA	LSQR	RFTF	SPGQD	IQLI	PPLI	NLLMS	ΙΕΡΟΫ	IYAGHD	NTKI	PDTSSSL	LTSLNQ	LGER	QLLSV	vĸwsĸs	LPGFRNLHI
	750		76	0		770		780		790			800		810		<b>820</b>		830		840		850		860
DDQI	ггі	YSW	MSLMVI	GLGW	RSYK	HVSG	2 M L Y F	APDL	ILNEQ	RMKES	SFYS	LCL	тмжот	PQEF	AKTÖ	A S O E E I	FLCMK	VLLLN	ттр	LEGLRSQ	TQFEEM	RSSY	IRELI	KAIGLR	Q K G V V S S S Q
870			880		890		900		910	, o	9	20		930	_										
RFYQ	<b>ТКГ</b>	LDN	THDTAI	QLHI	YCLN	FFIQ	SRALS	VEFP	еммѕе	VIAAQ	LPKI	LAG	мчкрі	LFHK	К										
										20				•	10					50					

**Figure S11 PR-B vs. PR-B:PRE consolidated HDX-MS data.** Consolidated data output from HDXWorkbench for PR-B ± DNA HDX-MS experiments when RU-486 bound. This illustrates the concerted decrease in deuterium exchange when DNA bound, compared to non-DNA-bound protein. Each point represents an average across each peptide detected, with the differential HDX data following the color scheme shown at the bottom.