Supplementary Information

Divergence of multimodular polyketide synthases revealed by a didomain structure

Jianting Zheng[1], Darren C. Gay[1], Borries Demeler[2], Mark A. White[3],

Adrian T. Keatinge-Clay[1]*

[1]Department of Chemistry and Biochemistry, The University of Texas at Austin

[2]Department of Biochemistry, The University of Texas Health Science Center at San Antonio

[3]Sealy Center for Structural and Molecular Biophysics, UTMB Galveston

* Correspondence: adriankc@mail.utexas.edu

Supplementary Methods

Cloning, expression, and purification

Domain boundaries from previous KR structures guided the selection of the first and last residues of the Spn(KR+ER)2 didomain (SpnB residues 1216-1989)[1]. The DNA encoding Spn(KR+ER)2 was amplified using primers 5'-ATCGTAATC*CATATG*GCTCGGTTGTACCGGTTGAG-3' and 5'-TGATTCGAT*GAATTC*ACAACCGTTTCCGGAGGCTCT-3' (NdeI and EcoRI sites in italics; stop codon underlined) from *Saccharopolyspora spinosa* genomic DNA (extracted from *S. spinosa* using the DNeasy Blood & Tissue Kit from Qiagen), digested with NdeI and EcoRI, and ligated into pET28b (Novagen). The DNA encoding the SpnER2 domain was amplified using primers 5'-ATCGTAATC*CATATG*GGCACCGGGTGGCGATTGGA-3' and 5'-TGATTCGAT*GAATTC*ACACCGGAGGCATGGTGAGCA-3' and ligated into pET28b (Novagen) as described above. Reverse PCR was used to remove the ER domain from the Spn(KR+ER)2 didomain construct using 5'-phosphorylated primers 5'-CCTCCGGTGTGGGACGCCGCAGGCACGGTTCTGGTTAC-3' and 5'-CCCGGTGCCGTCCGGCACGGGCAGGACGTCATCGCT-3'. *E. coli* BL21(DE3) transformed with the expression plasmid was inoculated into LB media containing 50 mg/L kanamycin at 37 °C, grown to OD_{600} = 0.4, and induced with 0.5 mM IPTG. After 12 h at 15 °C, cells were collected by centrifugation and resuspended in lysis buffer (300 mM NaCl, 50 mM HEPES, pH 7.5). Following sonication, cell debris was removed by centrifugation (30,000 x g for 45 min). The supernatant was poured over a column of Nickel-NTA resin

(Qiagen), which was then washed with 50 mL lysis buffer containing 15 mM imidazole and eluted with 5 mL lysis buffer containing 300 mM imidazole. Proteins were further purified using a Superdex 200 gel filtration column (GE Healthcare Life Sciences) equilibrated with 150 mM NaCl, 10 mM HEPES, pH 7.5. Spn(KR+ER)2 was concentrated to 24 mg/mL in 25 mM NaCl, 1 mM DTT, 10 mM HEPES, pH 7.5. Selenomethionine-labeled Spn(KR+ER)2 was obtained by the pathway inhibition method[2]. Briefly, BL21(DE3) cells were grown in M9 minimal media containing 4 g/L glucose, 6 g/L Na_2HPO_4, 3 g/L KH_2PO_4, 1 g/L NH_4Cl, 0.5 g/L NaCl, 1 mM $MgSO_4$, 0.1 mM $CaCl_2$, and 50 mg/L kanamycin at 37 °C. Lysine, phenylalanine, threonine (each at 100 mg/L), isoleucine, leucine, valine (each at 50 mg/L), and selenomethionine (at 60 mg/L) were supplied at OD_{600} = 0.4. After 15 minutes at 37 °C, 0.5 mM IPTG was added, and the cells were induced for 16 hours at 15 ºC. The purification followed the protocol described for the native protein.

Site-directed mutagenesis

The GeneTailor Site-Directed Mutagenesis System (Invitrogen) was used to generate all the mutants (verified by sequencing). Purification of the point mutants followed the protocol described for the wild-type protein. The following oligonucleotides were used for mutagenesis: K422A, 5′-GCCGTTTCCTGGAGTTGGGGGCGACGGATGTTCGTGACC-3′ and 5′-CCCCAACTCCAGGAAACGGCCACCGCGCGGCAGCATTC-3′; D444A, 5′-CGTGTCTTACCAGGCTTTCGCTACCGTAGAGGCAGGCC-3′, and 5′-CGAAAGCCTGGTAAGACACGCCCGGATGCGCATCGGC-3′; Y241F,

5’-GCTCATCGCGCTCGGTATGTTTCCCGGTGTGGCATCGCT-3’ and 5’-ACATACCGAGCGCGATGAGCGCATCCCGGAAGTTCAC-3’.

Size-exclusion chromatography

Each sample (0.3 mL; 69 μM Spn(KR+ER)2; 132 μM SpnER2; 75 μM SpnKR2) was injected onto a Superdex 200 gel-filtration column equilibrated with 150 mM NaCl, 10 mM sodium phosphate, pH 7.4. The molecular weight was estimated by comparison with standards (Gel Filtration Standard, Biorad Laboratories) as previously described[7].

Analytical ultracentrifugation

All sedimentation experiments were performed with a Beckman Optima XL-I at the Center for Analytical Ultracentrifugation of Macromolecular Assemblies at the University of Texas Health Science Center at San Antonio (UTHSCSA). Sedimentation velocity data were analyzed with the UltraScan-II software version 9.9[3]. All calculations were performed on the Lonestar and Ranger clusters at the Texas Advanced Computing Center at the University of Texas at Austin, and on the Jacinto cluster at the Bioinformatics Core Facility at UTHSCSA. All measurements were made at 280 nm in a buffer containing 150 mM NaCl, 10 mM sodium phosphate, pH 7.4. The experimental data were collected at 20 °C and at 50 krpm using standard Epon 2-channel centerpieces. Hydrodynamic corrections for buffer density, viscosity, and partial specific volume were made according to methods implemented in UltraScan[41]. The partial specific volumes of Spn(KR+ER)2, SpnER2, and SpnKR2 were determined to be 0.737 cm^3/g, 0.736 cm^3/g, and 0.735 cm^3/g, respectively. All data were first analyzed by

2-dimensional spectrum analysis (2DSA) with simultaneous removal of time- and radial-invariant noise and fitting of the meniscus position, followed by van Holde-Weischet analysis and genetic algorithm (GA) refinement, followed by Monte Carlo analysis (MC)[4,5]. Each protein was measured at 280 nm at three loading concentrations (Spn(KR+ER)2, 1.2 μM, 4.1 μM, and 12.0 μM; SpnER2, 3.0 μM, 10.5 μM, and 31 μM; SpnKR2, 2.5 μM, 7.5 μM, and 23 μM).

Crystallization and structure determination

Crystals of Spn(KR+ER)2 grew over one week by sitting drop vapor diffusion at 22 °C. Drops were formed by mixing 2 μL protein solution (24 mg/mL Spn(KR+ER)2 in 25 mM NaCl, 1 mM DTT, 10 mM HEPES, pH 7.5) with 1 μL crystallization buffer (17% (w/v) poly(acrylic acid) 5100, 0.4 M $MgSO_4$, 0.1 M HEPES, pH 6.5). Crystals were frozen in liquid nitrogen. Three data sets (peak, inflection, and remote) from a crystal of selenomethioning-labeled Spn(KR+ER)2 were collected at ALS Beamline 5.0.2 and processed by HKL2000[6]. The structure was solved to 3.40 Å resolution by multiwavelength anomalous dispersion phasing using the program Phenix[7]. The program Molrep fit the resulting electron density map with AmpKR2 (PDB Code: 3MJS) and the ER from the porcine FAS (PDB Code: 2ZV8), and the solutions were combined to obtain the model of two Spn(KR+ER)2 molecules in the asymmetric unit[8-10]. A 3.0 Å-resolution data set collected from a crystal of native Spn(KR+ER)2 at ALS Beamline 8.2.1 was used to iteratively build and refine the model through the programs Coot and Refmac5[8,11].

SAXS data collection and analysis

Spn(KR+ER)2 was exchanged by gel filtration chromatography (Superdex 200) into 10% glycerol, 500 mM NaCl, 50mM $(NH_4)_2SO_4$, and 30 mM HEPES, pH 7.5. SAXS data were collected at 5 °C on a Rigaku BioSAXS-1000 (Kratky camera, with a 2.5kW FRE+ source) at the Sealy Center for Structural and Molecular Biophysics (SCSB) BioSAXS facility of the University of Texas Medical Branch at Galveston. A dilution series showed no difference between a 10 mg/mL and a 5 mg/mL sample; therefore, a sample concentration of 10 mg/mL was used for data collection. Five 30-minute exposures were collected. Radiation-induced aggregation effects slowly developed in the low-q regime (q<1.3/R_g); however, the scattering curves of all samples converged beyond this point, permitting all of the curves to be included in the averaging for the high-q regime (q>1.3/R_g). Five more 30-minute data collections were performed, each with fresh sample. Using the SAXSLab software (Rigaku), the high-q regions of these ten collections were averaged, and the low q-regime of the last five collections were also averaged. The high- and low-q curves were merged together and scattering from the buffer alone was subtracted (seven 30-minute scans were collected on the buffer alone) with the program PRIMUS[12]. The resulting SAXS curve was used for subsequent analysis. The radius of gyration (R_g) and $I(0)$ were evaluated using the Guinier approximation (0.0223 $Å^{-1}$ < q < 0.0437 $Å^{-1}$). The program GNOM was used to obtain the distance distribution function, $P(r)$, and to determine the value for the R_g from the entire scattering profile[13]. Good agreement of R_g(Guinier) and R_g(GNOM) indicated the internal consistency of the data[14]. A molecular weight was calculated from the SAXS curve[15]. The program CRYSOL predicted the scattering curves from two models (one model of

Spn(KR+ER)2 based on its crystal structure and one model of Spn(KR+ER)2 based on the superposition of SpnKR2 and SpnER2 on the KR and ER domains of the porcine FAS) and compared them to the experimentally-obtained SAXS curve[16]. *Ab initio* envelopes were generated by the program DAMMIF[17]. A developmental version of the pmb_dammif script within the Protein Model Builder (PMB) package was used to generate fifty envelopes, which were averaged and filtered with the program DAMAVER[18]. The resulting consensus envelope was then aligned and fit to each model using the program SUPCOMB[19].

Substrate synthesis

Pantethine (278 mg, 0.5 mmol, 1 eq.) was incubated with DTT (77 mg, 0.5 mmol, 1 eq.) in 1 mL water and stirred for 1 hour at room temperature. The water was removed under vacuum, and the clear oil was resuspended in THF (8 mL) and triethylamine (2 mL). Dimethylaminopyridine (6 mg, 0.05 mmol, 0.1 eq.) was added, and the solution was stirred on ice. Chilled crotonyl chloride (50 μL, 55 mg, 0.525 mmol, 1.05 eq.) was added, and the reaction was stirred for one hour. The reaction was allowed to warm to room temperature and was stirred for another hour. Solvent was removed under vacuum, and the remaining orange residue was resuspended in 50% ethyl acetate/acetone. Using the same solvent system as a mobile phase, crotonyl-pantetheine was purified by silica gel chromatography. Fractions containing crotonyl-pantetheine (R_f = 0.30, silica TLC plate, EtOAc) were dried under vacuum; the clear residue was resuspended in water and the solution was frozen until needed. *Crotonyl-pantetheine*. Electrospray ionization mass spectrometry (ESI-MS), expected mass 347.4; observed mass: $[M+H]^+$ 347.2 (also observed $[M+H-H_2O]^+$, 329.2; and $[M+Na]^+$,

369.4). A Varian 400-MHz DirectDrive instrument was used to obtain NMR spectra. ^{1}H NMR (D_2O, 400 MHz) δ 7.05-6.95 (dq, 1H, J=15.53, 6.91); 6.29-6.24 (dq, 1H, J=15.53, 1.68); 3.95 (s, 1H); 3.51-3.32 (m); 3.07 (t, 2H, J=6.30); 2.44 (t, 2H, J=6.49); 1.87 (dd, 3H, J=6.88, 1.69); 0.89 (s, 3H); 0.85 (s, 3H). These peaks match those previously reported[20].

HPLC and LC/MS characterization of Spn(KR+ER)2 functional assays

To assay the activity of isolated Spn(KR+ER)2 and its point mutants, the following reaction was initiated: 40 μM Spn(KR+ER)2, 2 mM crotonyl-pantetheine, 1 mM $NADP^+$, 150 mM glucose, 3 μM glucose dehydrogenase, 12% (v/v) glycerol, 150 mM HEPES, pH 7.0. After 48 hours, the reaction was injected on a C_{18} reversed-phase HPLC column, and monitored at 235 nm (100% water to 100% MeOH, 30 min.). For LC/MS analysis, an Agilent Technologies system (1200 Series HPLC, 6130 quadrupole mass spectrometer) was employed; a gradient of 5-95% B (solvent A, water with 0.1% formic acid; solvent B, acetonitrile with 0.1% formic acid) was run using a Gemini C_{18} column (5 μm, 2 x 50 mm) over 12 min. at a flow rate of 0.7 mL/min. Both the crotonyl-pantetheine substrate and the butyryl-pantetheine product have been previously synthesized and characterized[20,21].

Kinetic assays

A kinetic model derived from standard Michaelis–Menten kinetics was employed for the reduction and hydration reactions catalyzed by SpnER2.

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \overset{k_{cat(R)}}{\rightarrow} E + P_R \quad ES \overset{k_{cat(H)}}{\rightarrow} E + P_H \qquad (1)$$

$$K_m = \frac{k_{-1} + k_{cat(R)} + k_{cat(H)}}{k_1} \qquad (2)$$

$$V_{0(R)} = \frac{k_{cat(R)}[E_T][S]}{K_m + [S]} \qquad (3)$$

Each reaction was performed in duplicate (50 μL reaction volume: 2.5-160 mM crotonyl-pantetheine, 1 mM NADPH, 20 μM enzyme, 10% glycerol, 100 mM NaCl, 100 mM HEPES, pH 7.5 at 22 ºC). The consumption of NADPH was monitored at 340 nm, and readings were corrected for the nonenzymatic consumption of NADPH. Initial velocities were plotted against substrate concentration and fit to the kinetic model to obtain K_m and $k_{cat(R)}$ using the program Prism. To compare the hydration/reduction activities of SpnER and its point mutants the ratio of hydrated to reduced product generated by each SpnER2 mutant was compared to that of SpnER2. Reactions were stopped through the addition of an equal volume of methanol and analyzed by LC/MS (after removing precipitated protein by centrifugation). Extracted ion chromatograms for the hydration product (365.2 Da) and the reduction product (349.2 Da) were used to obtain the peak areas for the hydrated and reduced products, respectively.

References

1. Waldron, C. *et al.* Cloning and analysis of the spinosad biosynthetic gene cluster of Saccharopolyspora spinosa. *Chem. Biol.* **8**, 487-99 (2001).
2. Keatinge-Clay, A.T. & Stroud, R.M. The structure of a ketoreductase determines the organization of the beta-carbon processing enzymes of modular polyketide synthases. *Structure* **14**, 737-48 (2006).
3. Demeler, B. A comprehensive data analysis software package for analytical

ultracentrifugation experiments. in *Modern analytical ultracentrifugation: techniques and methods*. (Scott, D.J., Harding, S.E. & Rowe, A.J., Eds.) 210-229 (Royal Society of Chemistry, UK, 2005).

4. Brookes, E., Cao, W. & Demeler, B. A two-dimensional spectrum analysis for sedimentation velocity experiments of mixtures with heterogeneity in molecular weight and shape. *Eur. Biophys. J.* **39**, 405-14 (2009).

5. Demeler, B. & Brookes, E. Monte Carlo analysis of sedimentation experiments. *Colloid. Polym. Sci.* **286**, 129-37 (2008).

6. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. in *Methods in Enzymology, Volume 276: Macromolecular Crystallography, Part A* (Carter, C.W., Jr. and R.M. Sweet, R.M., eds) 307-26 (Academic Press, New York, NY, 1997).

7. Adams, P.D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213-21 (2010).

8. Potterton, E., Briggs, P., Turkenburg, M. & Dodson, E. A graphical user interface to the CCP4 program suite. *Acta Crystallogr. D Biol. Crystallogr.* **59**, 1131-7 (2003).

9. Maier, T., Leibundgut, M. & Ban, N. The crystal structure of a mammalian fatty acid synthase. *Science* **321**, 1315-22 (2008).

10. Zheng, J., Taylor, C.A., Piasecki, S.K. & Keatinge-Clay, A.T. Structural and functional analysis of A-type ketoreductases from the amphotericin modular polyketide synthase. *Structure* **18**, 913-922 (2010).

11. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr.*

D Biol. Crystallogr. **60**, 2126-32 (2004).

12. Konarev, P.V., Volkov, V.V., Sokolova, A.V., Koch, M.H.J. & Svergun, D.I. PRIMUS - a Windows-PC based system for small-angle scattering data analysis. *J. Appl. Cryst.* **36**, 1277-82. (2003).

13. Svergun, D.I. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* **25**, 495-503 (1992).

14. Putnam, C.D., Hammel, M., Hura, G.L. & Tainer, J.A. X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Q. Rev. Biophys.* **40**, 191-285 (2007).

15. Fischer, H., de Oliveira Neto, M., Napolitano, H.B., Polikarpov, I. & Craievich, A.F. Determination of the molecular weight of proteins in solution from a single small-angle x-ray scattering measurement on a relative scale. *J. Appl. Cryst.* **43**, 101-9 (2010).

16. Svergun, D.I., Barberato, C. & Koch, M.H.J. CRYSOL - a program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Cryst.* **28**, 768-773 (1995).

17. Franke, D. & Svergun, D.I. DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J. Appl. Cryst.* **42**, 342-6 (2009).

18. Volkov, V.V. & Svergun, D.I. Uniqueness of ab-initio shape determination in small-angle scattering. *J. Appl. Cryst.* **36**, 860-4. (2003).

19. Kozin, M. & Svergun, D. Automated matching of high- and low-resolution structural models. *J. Appl. Cryst.* **34**, 33-41 (2000).

20. Gomes, B., Fendrich, G., Abeles, R.H. Mechanism of action of glutaryl-CoA and butyryl-CoA

dehydrogenases. Purification of glutaryl-CoA dehydrogenase. *Biochemistry* **20**, 1481-90 (1981).

21. Tran, L., Tosin, M. Spencer, J.B., Leadlay, P.F., Weissman, K.J. Covalent linkage mediates communication between ACP and TE domains in modular polyketide synthases. *Chembiochem* **9**, 905-15 (2008).

Supplementary Results

Supplementary Table 1. Data collection and refinement statistics

Data Collection				
Dataset	Native	Remote (λ1)	Inflection (λ2)	Peak (λ3)
Wavelength (Å)	0.9999	0.9645	0.9797	0.9796
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions				
a, *b*, *c* (Å)	73.1, 110.2, 202.9	73.3, 110.7, 203.4	73.4, 110.7, 203.4	73.3, 110.7, 203.4
Resolution (Å)	50-3.0(3.05-3.0)	50-3.40(3.46-3.40)	50-3.4(3.46-3.40)	50-3.4(3.46-3.40)
R_{merge}	0.10(0.54)	0.15(0.69)	0.16(0.69)	0.15(0.61)
$I / \sigma(I)$	17.9(2.8)	12.0(2.1)	12.0(2.0)	13.0(2.5)
Completeness (%)	99.4(96.4)	98.0(88.2)	97.5(84.8)	98.3(88.4)
Redundancy	5.8(5.7)	12.1(10.4)	12.0(9.8)	12.3(8.7)
Refinement				
Resolution (Å)	50-3.00			
No. reflections	31,804			
R_{work} / R_{free}	0.210/0.256			
No. atoms				
Protein	11,014			
$NADP^+$	192			
SO_4	15			
B-factors (Å^2)				
Protein	64.8			
$NADP^+$	60.4			
SO_4	53.5			
R.m.s. deviations				
Bond lengths (Å)	0.007			
Bond angles (°)	1.014			

Supplementary Table 2. Steady-state kinetic parameters for reduction of crotonyl-pantetheine by SpnER2 and its point mutants. Data reported as mean ± SEM.

Enoylreductase	**K_m (mM)[a]**	**$k_{cat(R)}$ (s^{-1})[b]**	**$k_{cat(R)}/K_m$ ($M^{-1}s^{-1}$)**
SpnER2	74 ± 14	0.0085 ± 0.0007	0.11 ± 0.02
SpnER2(D444A)	6.3 ± 0.7	0.0019 ± 0.0001	0.30 ± 0.04
SpnER2(Y241F)	24 ± 3	0.0020 ± 0.0001	0.08 ± 0.01
SpnER2(K422A)	55 ± 11	0.0010 ± 0.0001	0.018 ± 0.004

[a] K_m is defined as $(k_{-1}+k_{cat(R)}+k_{cat(H)})/k_1$. ($k_{cat(H)}$ is the turnover number of the hydration side-reaction)

[b] $k_{cat(R)}$: turnover number of the reduction reaction

Supplementary Table 3. Comparison of hydrated and reduced products generated by SpnER2 and its point mutants.

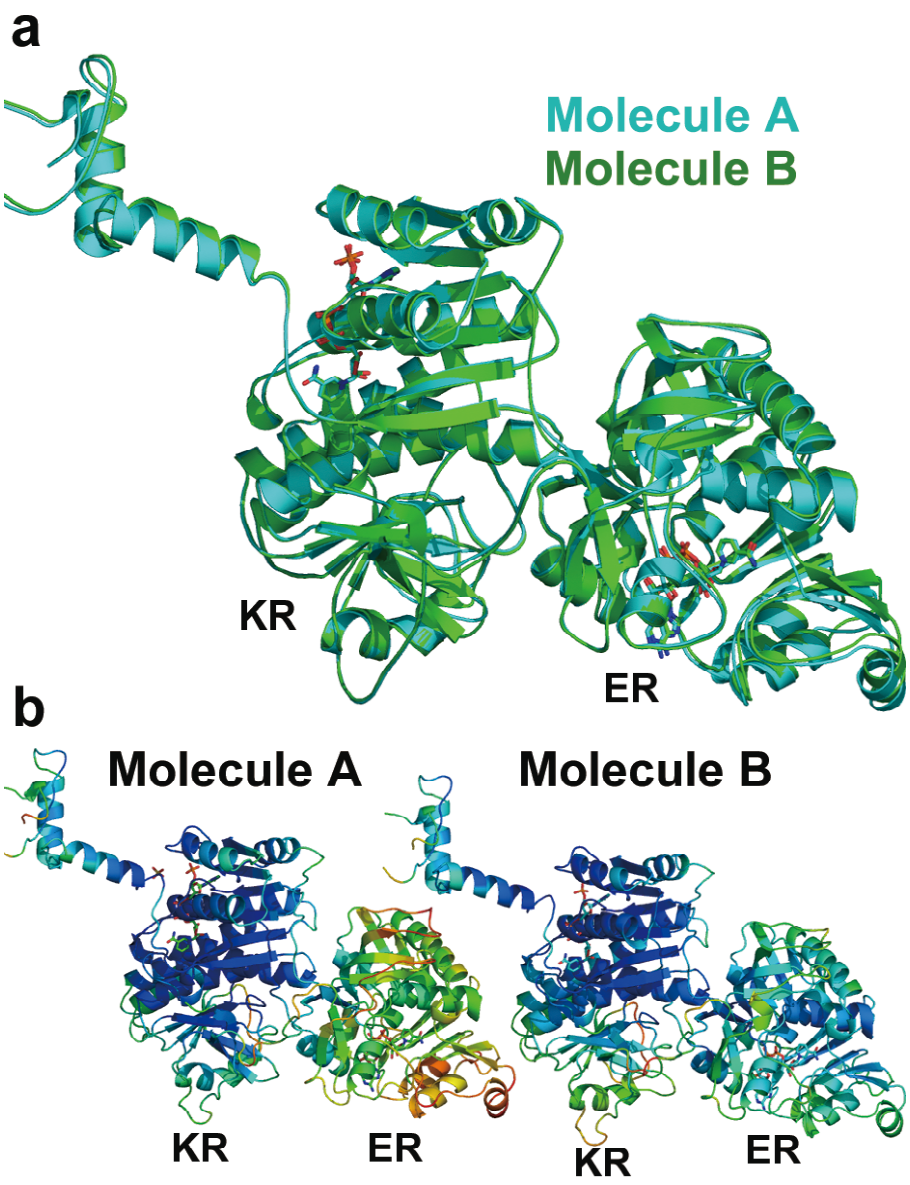
Enzyme	$A_{hydrated}$	$A_{reduced}$	$A_{hydrated}/A_{reduced}$
SpnER2	276533	8583817	0.03
SpnER2(D444A)	4916666	5759500	0.85
SpnER2(Y241F)	4582529	3647192	1.26
SpnER2(K422A)	5204331	749310	6.95

Reactions (20 mM substrate, two-hour incubation) were analyzed by LC/MS and the peak areas of the hydrated (365.2 Da) and reduced (349.2 Da) products were obtained form the extracted ion chromatograms respectively.

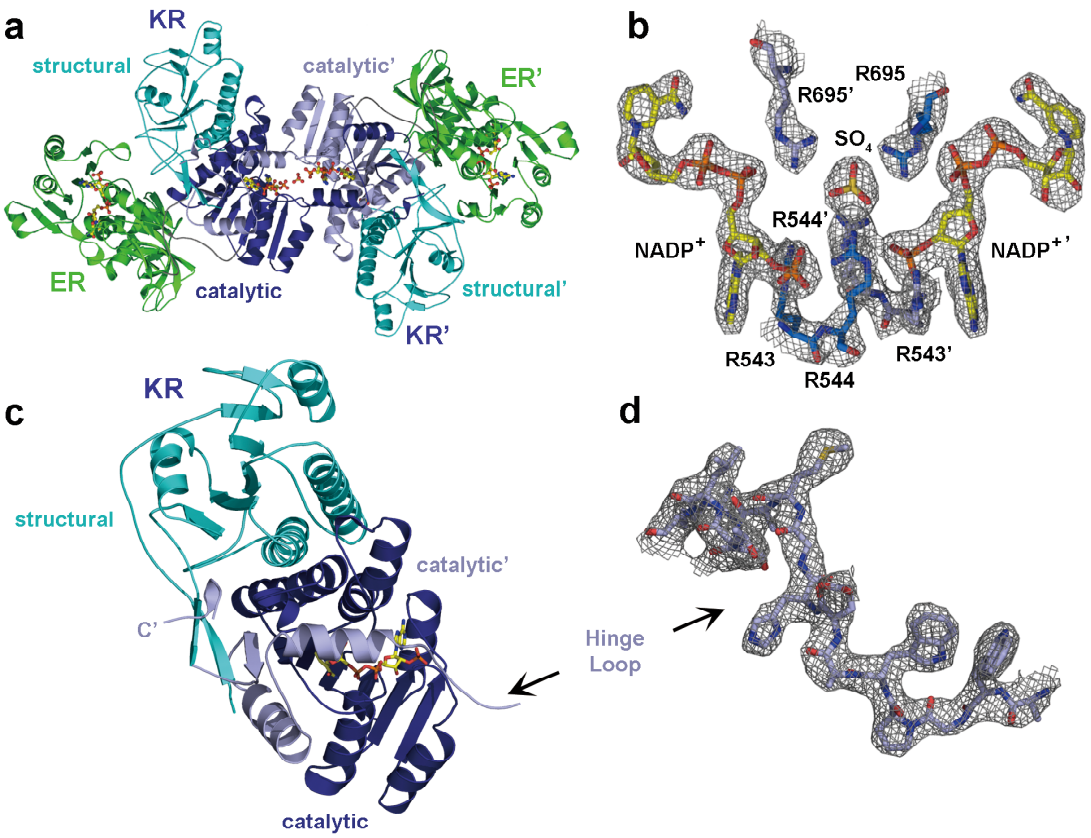
$A_{hydrated}$: peak area of hydrated product

$A_{reduced}$: peak area of reduced product

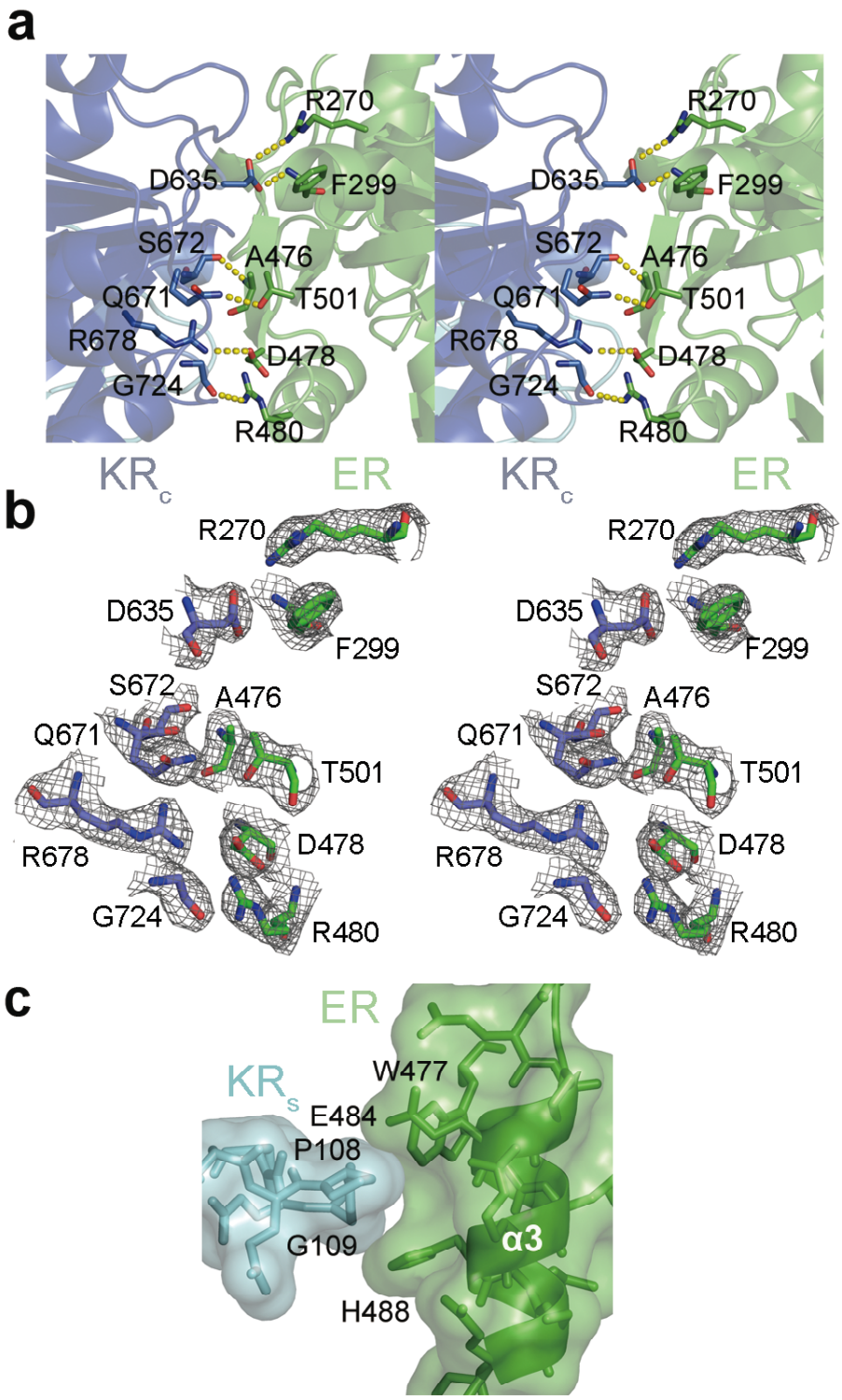
$A_{hydrated}/A_{reduced}$: peak area ratio of hydrated and reduced products



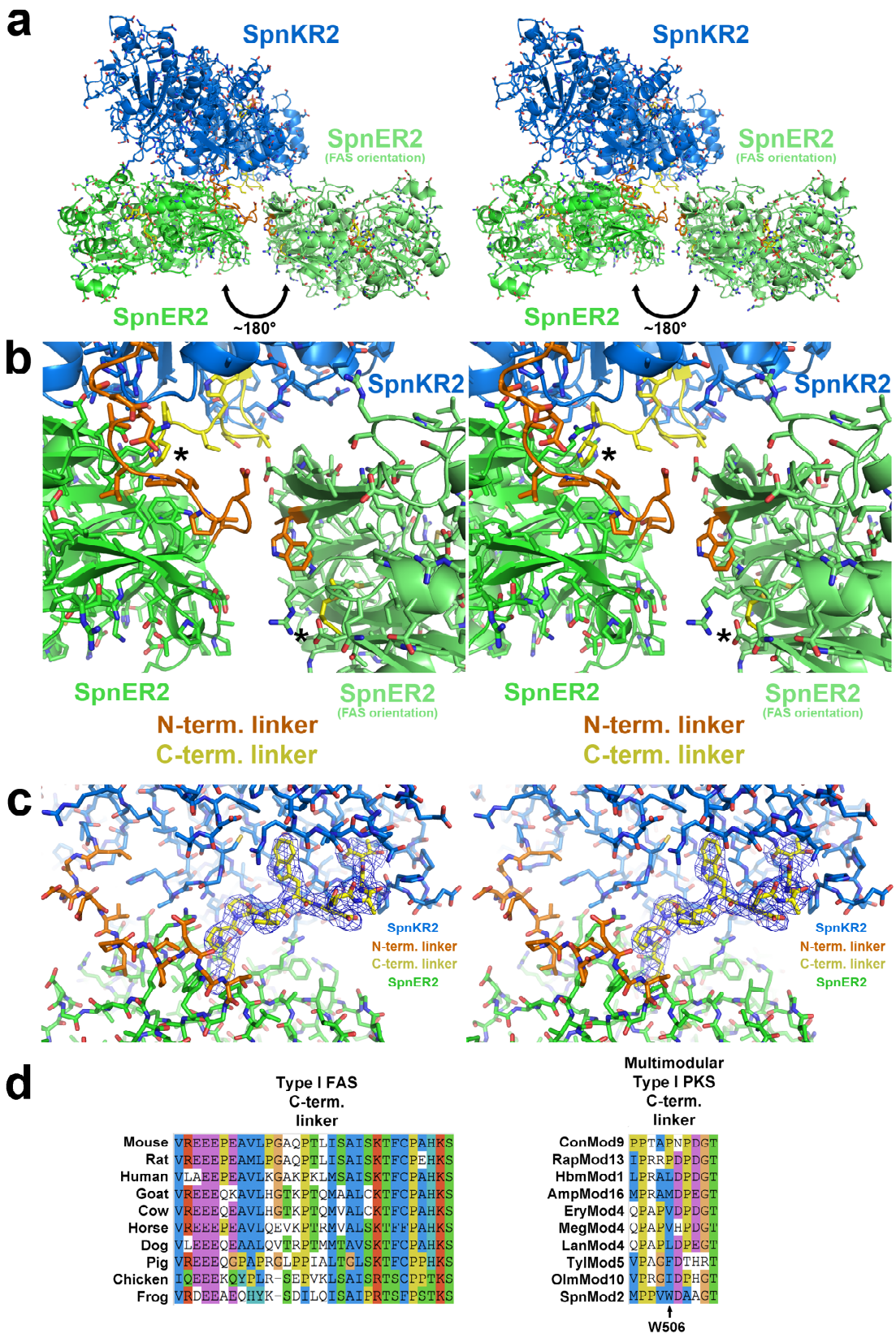
Supplementary Figure 1. Comparison of the two Spn(KR+ER)2 molecules in the asymmetric unit. (**a**) The superposition of the Spn(KR+ER)2 didomains shows their near-identical conformations (0.6 Å C_α r.m.s.d.). (**b**) The B-factors of ER in Molecule A are elevated compared to those of ER in Molecule B (96 $Å^2$ vs. 62 $Å^2$) (B-factors are indicated on a scale from blue to red).



Supplementary Figure 2. Domain-swapped, C-terminal segment. (**a**) Within the asymmetric unit, a KR/KR interface is formed between the two Spn(KR+ER)2 molecules. (**b**) The $2F_o$-F_c electron density map (contoured at 1.7 σ) shows the ionic interactions between arginines, $NADP^+$ molecules, and a sulfate at the pseudo-twofold axis of the asymmetric unit. (**c**) The 60 residues from the adjacent Spn(KR+ER)2 molecule are colored in grey. The positions of the swapped residues are as expected from related KR structures. (**d**) The $2F_o$-F_c electron density map (contoured at 1.7 σ) shows the well-ordered hinge loop (terminology from domain-swapping literature) that spans the Spn(KR+ER)2 molecules. The lid helix on the left and the tryptophans on the right are located within separate Spn(KR+ER)2 molecules, but where they would be expected from related KR structures.



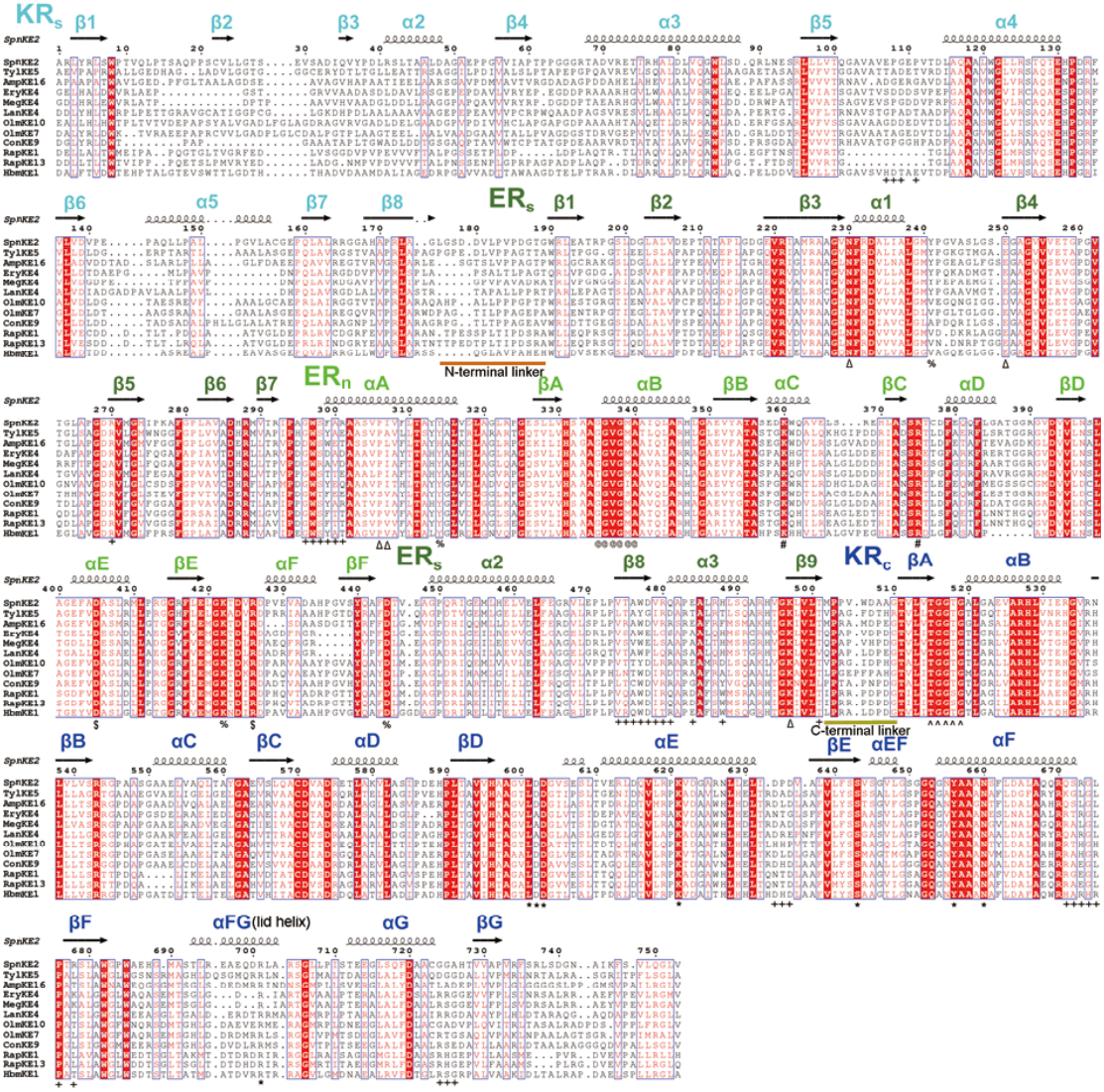
Supplementary Figure 3. The KR/ER interface. (**a**) A stereodiagram shows interactions between several hydrophilic residues at interface between ER and the catalytic subdomain of KR. (**b**) A stereodiagram of the $2F_o$-F_c electron density map (contoured at 1.0 σ) around these residues. (**c**) One contact was observed between ER and the structural subdomain of KR.



Supplementary Figure 4. ER orientation and the C-terminal linker. (**a**) A stereodiagram shows that when SpnKR2 and SpnER2 are superposed on the KR and ER domains of the

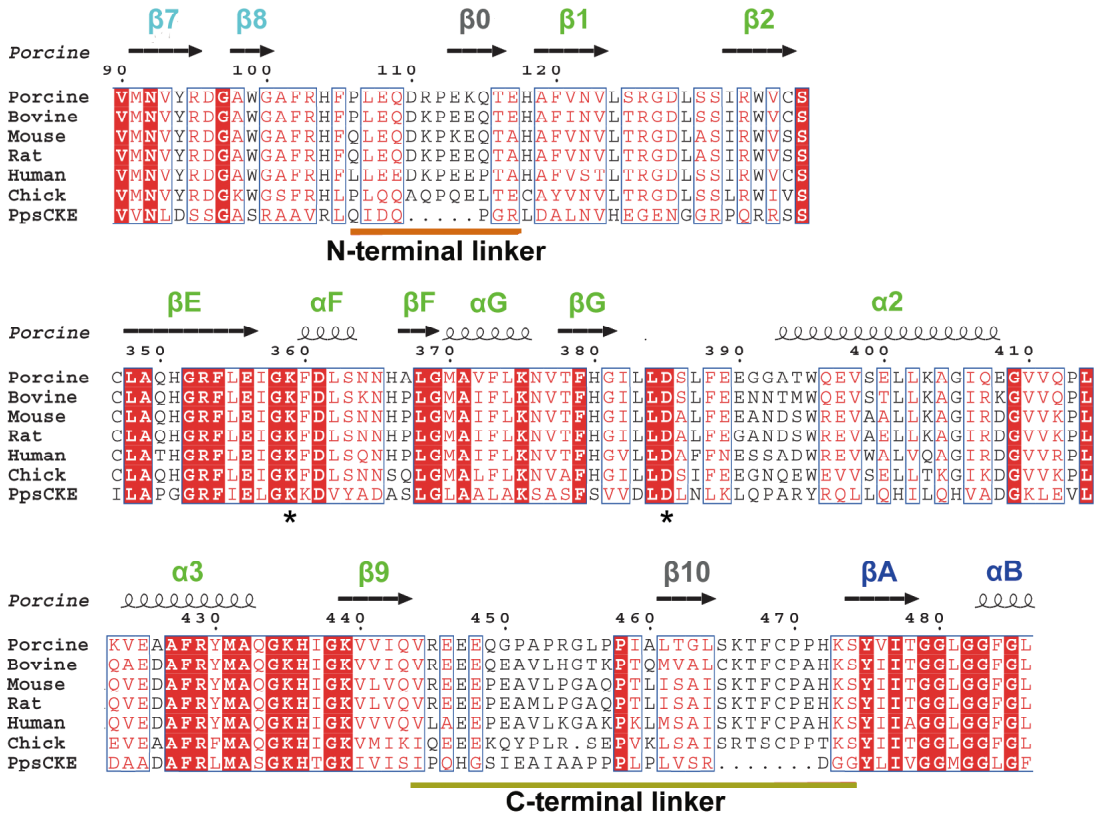
porcine FAS that SpnER2 is in an orientation ~180° from that observed in the Spn(KR+ER)2 crystal structure. (**b**) A stereodiagram suggests that Spn(KR+ER)2 cannot be in the FAS orientation since the C-terminal linker is not long enough to span the distance between ER and the catalytic subdomain of KR (the asterisks would need to connect). The insertion of W506 (middle of C-terminal linker) into a hydrophobic pocket of KR effectively reduces the length of the linker from 8 to 3 residues. (**c**) A stereodiagram illustrates the well-resolved electron density ($2F_o$-F_c, contoured at 1.7 σ) from the C-terminal linker of Spn(KR+ER)2. (**d**) A sequence alignment shows the conservation of the ~29-residue linker of FASs and the conservation of the 8-residue linker of polyketide synthases. GenBank accession numbers: Mouse (*Mus musculus*), NP_032014; Rat (*Rattus norvegicus*), AAA41145; Human (*Homo sapiens*), NP_004095; Goat (*Capra hircus*), ABI95140; Cow (*Bos taurus*), NP_001012687; Horse (*Equus caballus*), XP_001491342; Dog (*Canus lupus familiaris*), XP_540497; Pig (*Sus scrofa*), NP_001093400; Chicken (*Gallus gallus*), NP_990486; Frog (*Xenopus tropicalis*), XP_002937357; ConMod9 (module 9 of the concanamycin PKS from *Streptomyces neyagawensis*), AAG23265; RapMod13 (module 13 of the rapamycin PKS from *Streptomyces hygroscopicus*), CAA60462; HbmMod1 (module 1 of the herbimycin PKS from *Streptomyces hygroscopicus*), AAY28225; AmpMod16 (module 16 of the amphotericin PKS from *Streptomyces nodosus*), AAK73502; EryMod4 (module 4 of the erythromycin PKS from *Saccharopolyspora erythraea*), AAV51821; MegMod4 (module 4 of the megalomicin PKS from *Micromonospora megalomicea*), AAG13918, LanMod4 (module 4 of the lankamycin PKS from *Streptomyces avermitilis*), TylMod5 (module 5 of the tylosin PKS from *Streptomyces fradiae*), AAB66506; OlmMod10 (module 10 of the oligomycin PKS from

Streptomyces avermitilis), BAC70603; SpnMod2 (module 2 of the spinosyn PKS from *Saccharopolyspora spinosa*) AAG23265.

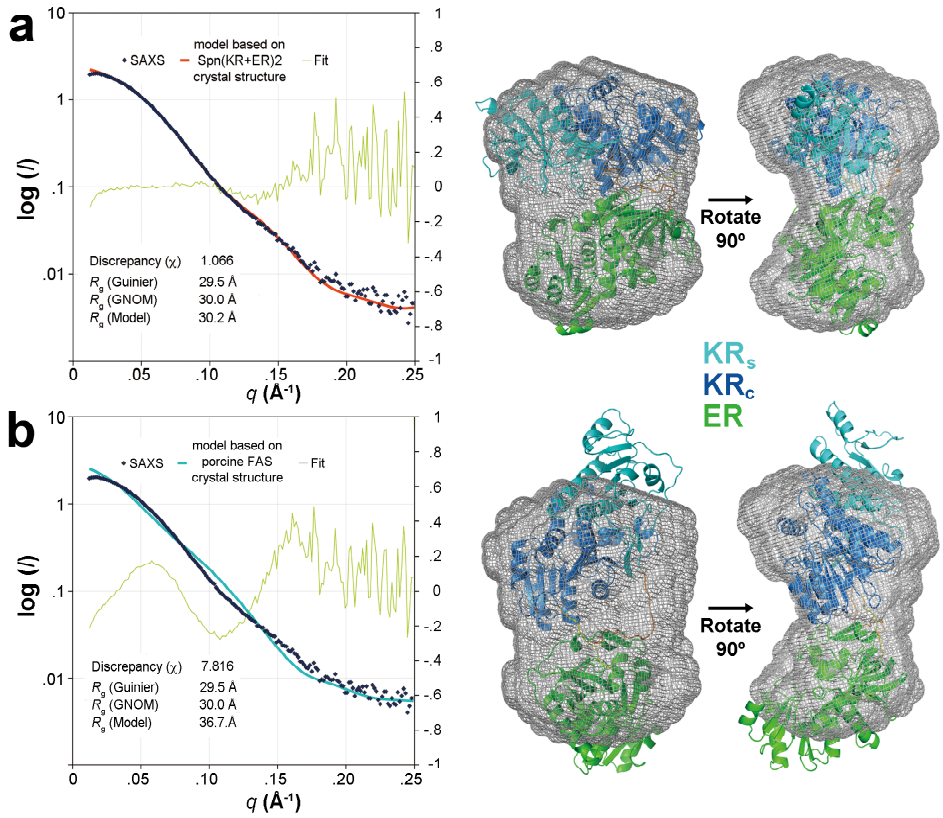


Supplementary Figure 5. Sequence alignment of KR+ER didomains from multimodular PKSs. The secondary structure shown above the alignment is from Spn(KR+ER)2. The N- and C-terminal linkers connecting KR and ER are underlined in orange and yellow, respectively. The residues at the KR/ER interface are labeled "+". Active site residues of KR are labeled "*". The GGVGMA sequence motif of ER for NADPH binding is labeled "@".

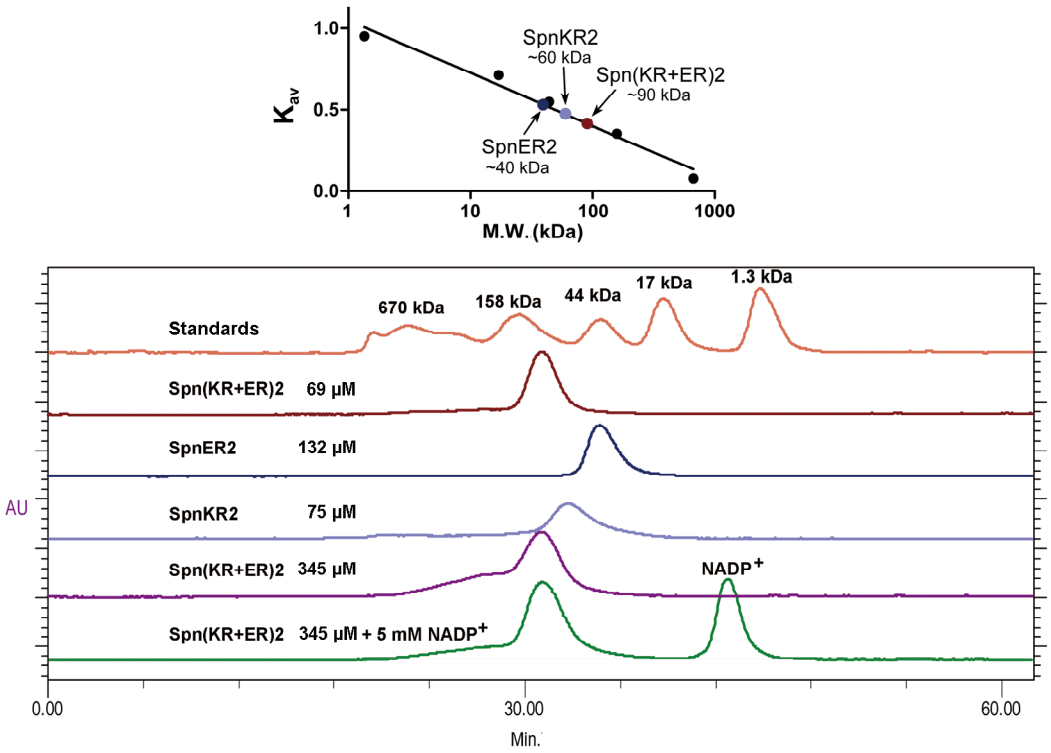
The TGGTG sequence motif of KR for NADPH binding is labeled "^". The conserved K360 and R375 that form salt bridges with the NADPH ribose phosphate are labeled "#". D405 and R426, which help orient αF, are labeled "$". The ER active site residues are labeled "%". The aligned KR+ER didomain sequences are named by their corresponding PKS and module of origin. GenBank accession numbers: SpnKE2 (module 2 of spinosyn PKS), AAG23265; TylKE5 (module 5 of tylosin PKS), AAB66506; AmpKE16 (module 16 of amphotericin PKS), AAK73502; EryKE4 (module 4 of amphotericin PKS), CAM00064; MegKE4 (module 4 of megalomicin PKS), AAG13918; LanKE4 (module 4 of lankamycin PKS), BAC76492; OlmKE10 (module 10 of oligomycin PKS), BAC70603; OlmKE7 (module 7 of oligomycin PKS), BAC70608; ConKE9 (module 9 of concanamycin PKS), AAZ94389; RapKE1 (module 1 of rapamycin PKS), CAA60460; RapKE13 (module 13 of rapamycin PKS), CAA60462; HbmKE1 (module 1 of herbimycin PKS), AAY28225.



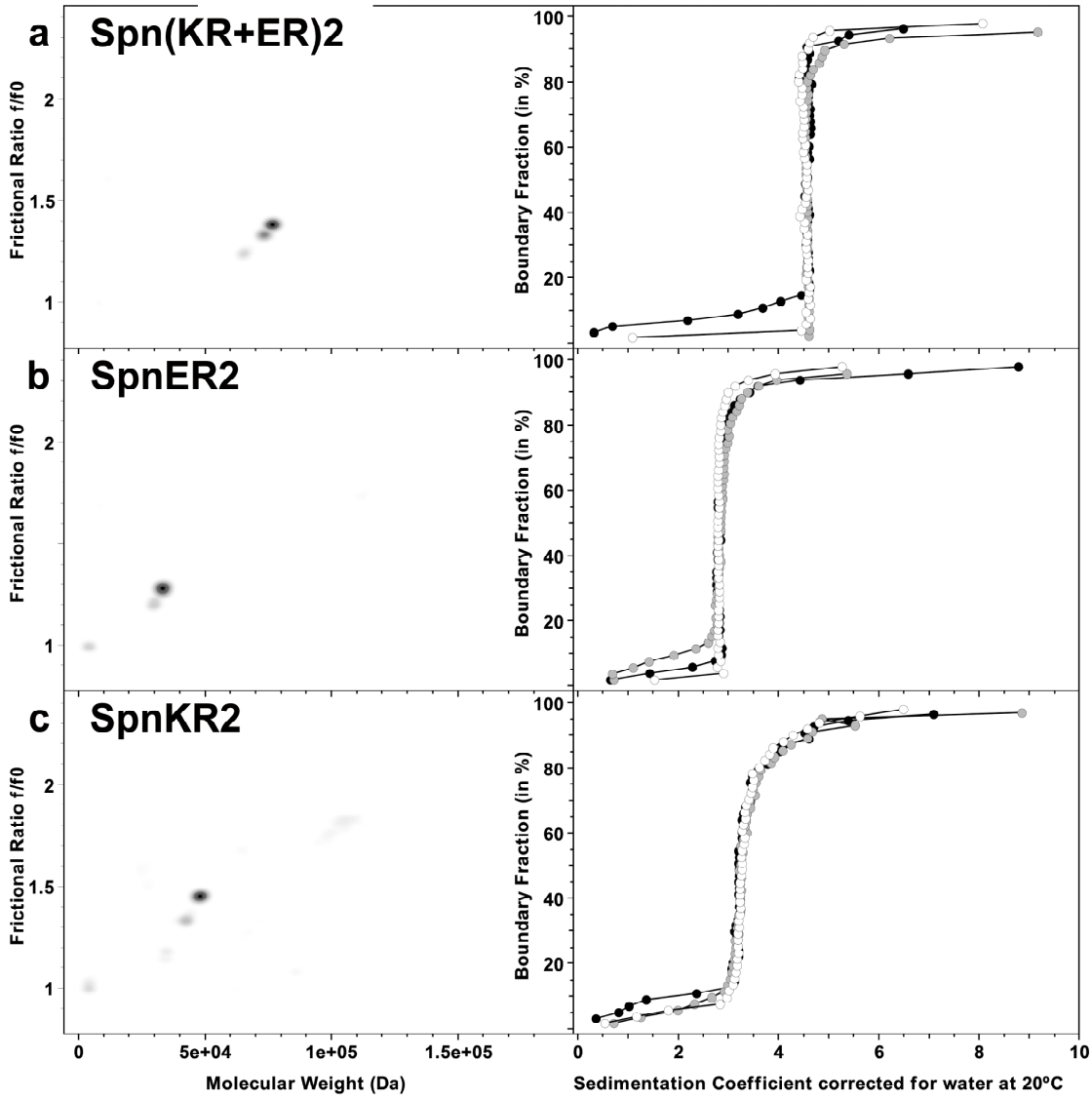
Supplementary Figure 6. Sequence alignments of regions within ERs of mammalian FASs and PpsC. The top alignment shows the N-terminal linkers connecting KR and ER (underlined in orange). The bottom alignment shows the region mediating ER dimerization and the close sequence similarity of PpsC to the mammalian FASs (compare with the sequences surrounding βF in the multimodular PKSs in Supplementary Fig. 5). ER active site residues are labeled by asterisks. GenBank accession numbers: Porcine (*Sus scrofa*), NP_001093400; Bovine (*Bos taurus*), Q71SP7; Mouse (*Mus musculus*), P19096; Rat (*Rattus norvegicus*), P12785; Human (*Homo sapiens*), P49327; Chick (*Gallus gallus*), P12276; PpsCKE (*Mycobacterium tuberculosis*), ZP_03433948.



Supplementary Figure 7. SAXS analysis of Spn(KR+ER)2. (**a**) The SAXS experimental scattering data is compared to a curve predicted by the program CRYSOL from a model of Spn(KR+ER)2 based on its crystal structure. The green line labelled "Fit" indicates the discrepancy between the SAXS data and the CRYSOL-predicted curve. The low discrepancy score indicates a good fit. An *ab initio* molecular envelope was also predicted by the program DAMMIF, and the Spn(KR+ER)2 model was fit to this envelope by the program SUPCOMB. (**b**) The curve predicted by CRYSOL from a model generated by superposing SpnKR2 and SpnER2 on the KR and ER domains of the porcine FAS is quite different from the experimental scattering data, as indicated by the high discrepancy score. This model does not fit the *ab initio* molecular envelope well. The experimental radius of gyration (R_g) determined from the Guinier plot and the program GNOM were also much closer to R_g predicted by CRYSOL for the model based on the Spn(KR+ER)2 crystal structure.



Supplementary Figure 8. Molecular weight estimation by size-exclusion chromatography. SpnER2 migrates at ~40 kDa (expected monomer mass: 36 kDa), SpnKR2 migrates at ~60 kDa (expected monomer mass: 51 kDa), and Spn(KR+ER)2 migrates at ~90 kDa (expected monomer mass: 83 kDa) compared to standards (running buffer: 150 mM NaCl, 10 mM sodium phosphate, pH 7.4). Domain-swapping does not occur in solution since dimerization was not observed with an injection of 345 µM Spn(KR+ER)2 (purple trace); supplementing the running buffer with 0.1 mM $NADP^+$ and pre-incubating 345 µM Spn(KR+ER)2 with 5 mM $NADP^+$ also did not elicit dimerization (green trace).



d

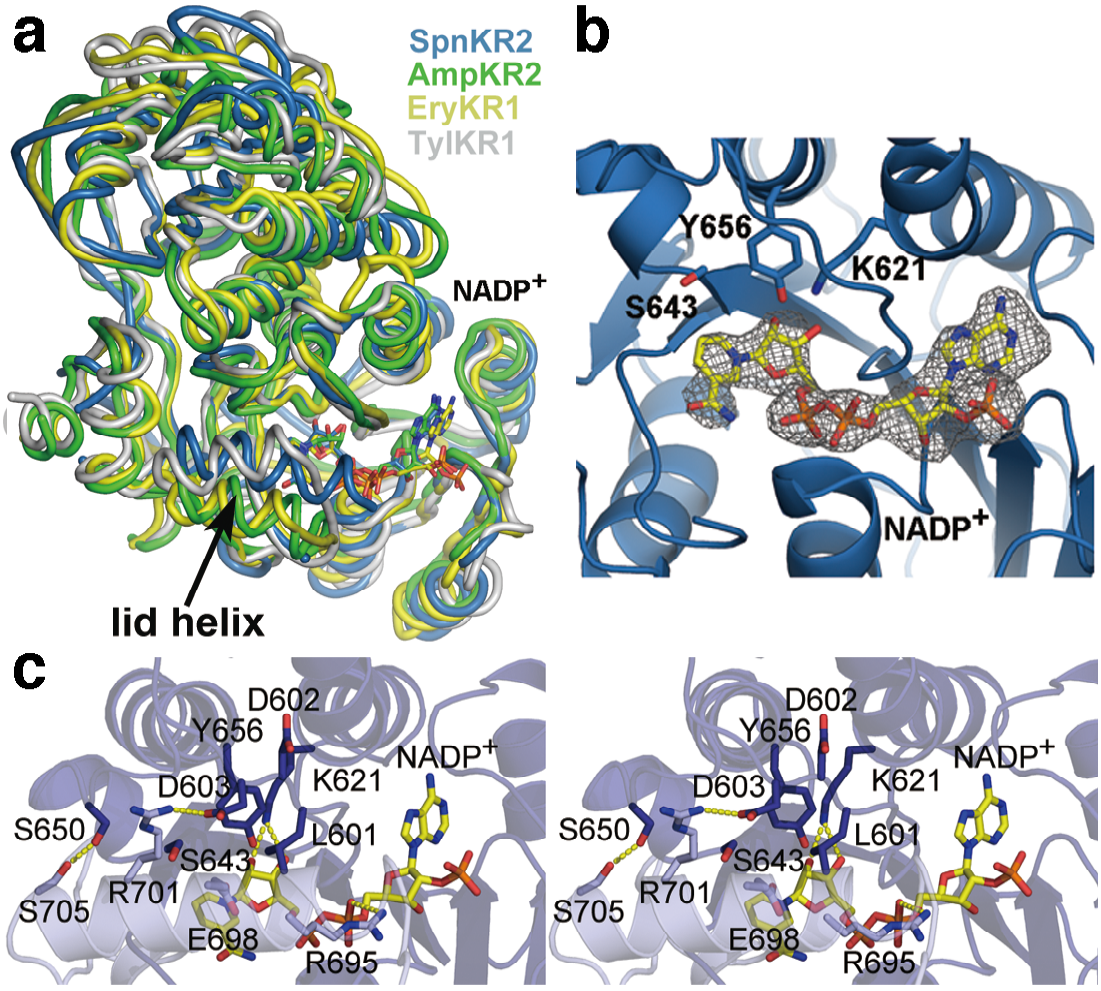
Genetic Algorithm-Monte Carlo Analysis Results

Protein	Frictional ratio f/f0	Molecular Weight (experimental)	Molecular Weight (theoretical)
Spn(KR+ER)2	1.36 (1.22, 1.50)	7.48×10^4 (6.36×10^4, 8.60×10^4)	8.34×10^4
SpnER2	1.28 (1.19, 1.36)	3.26×10^4 (2.94×10^4, 3.59×10^4)	3.58×10^4
SpnKR2	1.46 (1.44, 1.47)	4.78×10^4 (4.70×10^4, 4.86×10^4)	5.05×10^4

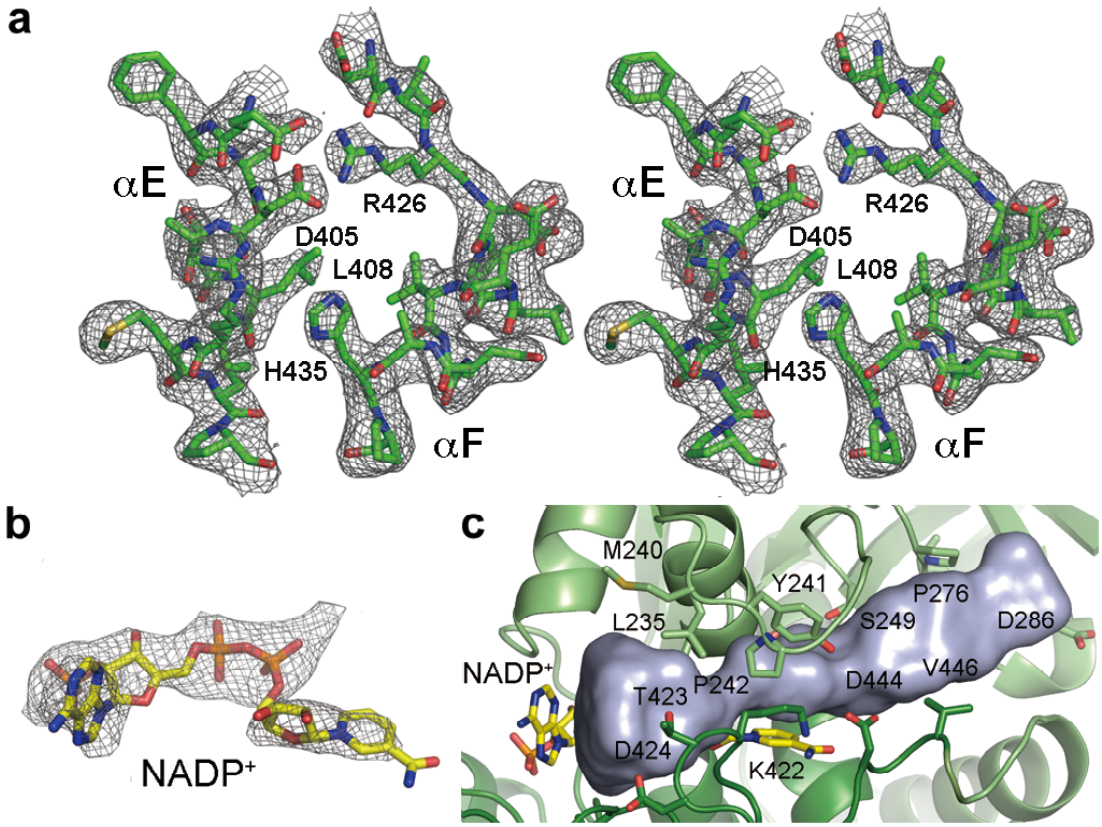
(Numbers in parentheses reflect 95% confidence intervals)

Supplementary Figure 9. Sedimentation velocity analysis of Spn(KR+ER)2, SpnER2, and SpnKR2. (**a**-**c**) On the right side are the combined van Holde-Weischet integral S-value distributions of three loading concentrations measured at 280 nm (**a**, right, Spn(KR+ER2): black: 1.2 μM, white: 4.1 μM, and gray: 12.0 μM; **b**, right, SpnER2: black: 3.0 μM, white: 10.5 μM, and gray: 31 μM; **c**, right, SpnKR2: black: 2.5 μM, white: 7.5 μM, and gray: 23 μM).

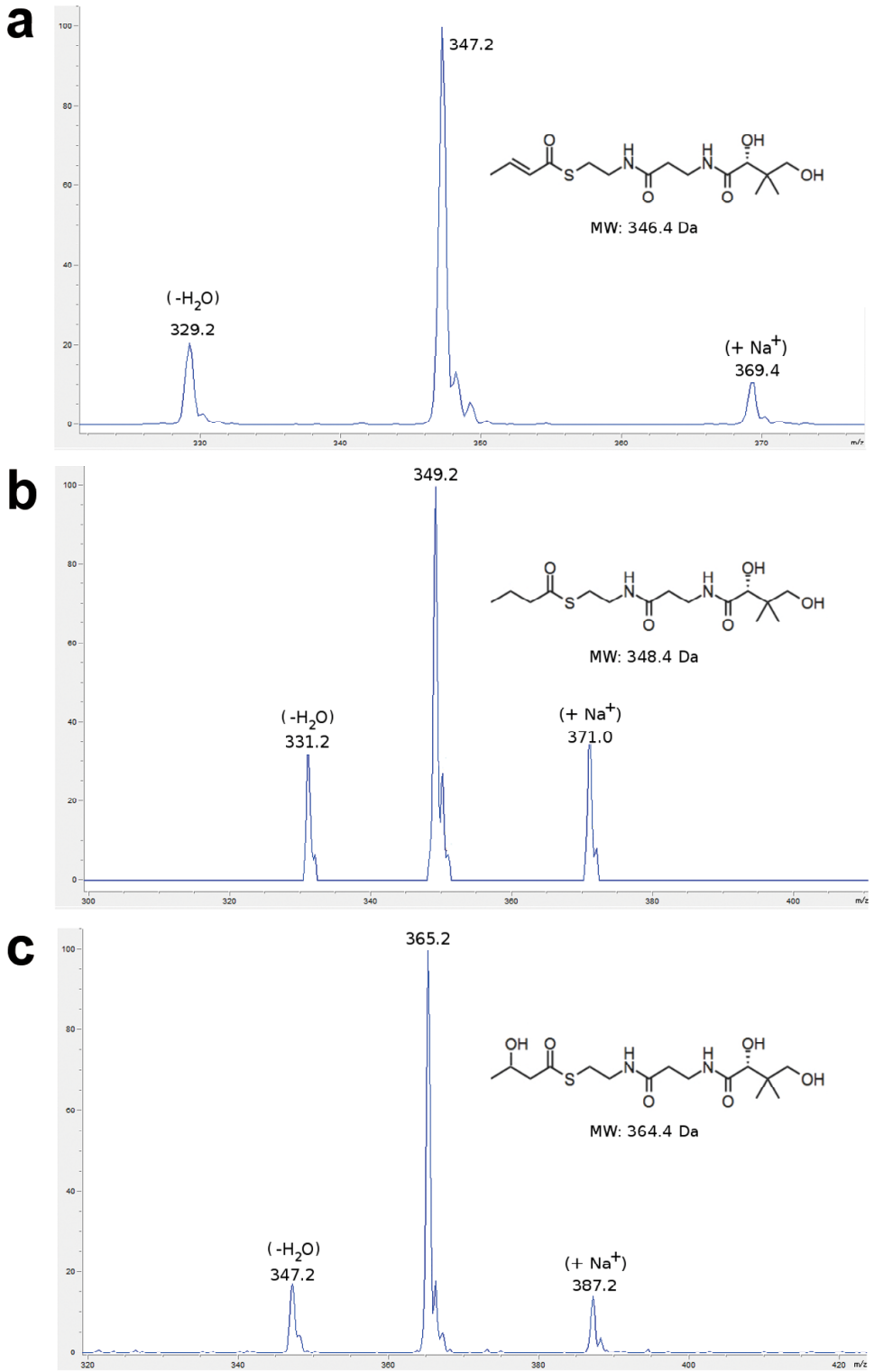
On the left side the combined genetic algorithm-Monte Carlo analysis of each protein is shown. All proteins have frictional ratios less than 1.5 and display one major species. The van Holde-Weischet analysis clearly shows no change in sedimentation coefficient and partial concentration upon change in loading concentration, indicating that in this concentration range there is no reversible association occurring and more than 80% of the sample corresponds to the monomeric form of each protein. **(d**) Molecular weights determined by the genetic algorithm analysis unambiguously show monomer only. Lower and higher molecular weight species are only minor contaminants and do not exhibit any mass action effects.



Supplementary Figure 10. SpnKR2 close-up. (**a**) Superposition of SpnKR2 (blue), AmpKR2 (green), EryKR1 (yellow), and TylKR1 (grey). (**b**) A simulated annealing F_o-F_c omit map (contoured at 3.0 σ) shows $NADP^+$ bound at the SpnKR2 active site next to the catalytic tyrosine, lysine, serine, and asparagine. (c) The interactions between the lid helix (from the adjacent molecule in the asymmetric unit) and the Leu-Asp-Asp motif. Of particular note is the salt bridge formed between the nearly invariant arginine (Arg701) and the second aspartic acid (Asp603) of the Leu-Asp-Asp motif, residues known to help enforce reduction stereochemistry.



Supplementary Figure 11. SpnER close-up. (**a**) A stereodiagram shows the $2F_o$-F_c electron density map (contoured at 1.0 σ) surrounding helices in the region usually employed by MDR enzymes for dimerization. (**b**) A simulated annealing F_o-F_c omit map (contoured at 2.5 σ) of $NADP^+$ bound at the SpnER2 active site. (**c**) The substrate binding tunnel (residues defining the tunnel are in sticks) that a polyketide intermediate would enter to be reduced by SpnER2.



Supplementary Figure 12. Characterization of ER reactions by mass spectrometry. (**a-c**) The crotonyl-pantetheine substrate, the butyryl-pantetheine product, and a side product that has the mass of β-hydroxybutryl-pantetheine (labeled **1**, **2**, and ***** in Figure 4) were analyzed by LC/MS.