SUPPORTING INFORMATION

Orthogonal synthetic zippers as protein scaffolds

George P. Anderson[1], Lisa C Shriver-Lake[1], Jinny L. Liu[1] and Ellen R. Goldman[1]*

[1] Center for BioMolecular Science and Engineering, U.S. Naval Research Laboratory, 4555 Overlook Avenue SW, Washington, District of Columbia 20375 USA

*corresponding author, ellen.goldman@nrl.navy.mil

Contents:

Figure S1. Protein sequences of constructs described in this work. Constructs are designated by abbreviating the domains in the order they occurred using the letter designation of the sdAb (D or A) and the number of the SYNZIP (1, 2, 3, 4, 5, 6, 17, 18) or designation E34m and R34m for the charged zipper pair. Protein sequences are given using single letter amino acid codes. Glycine-Serine (GS)-based linkers separate the domains as well as separate the final domain from the Histidine tail used for protein purification. The protein sequence, including the sdAb, linkers, and zipper are given for each construct that we produced and evaluated.

>D-1
EVQLQASGGGLVQDGGSLRLSCAVAGRPLSDYGVGWFRQASGKEREFVAVISGSGIV
TDYANSVKGRFTISRDVAKNVVHLQMNSLKPEDTAVYYCAALTNPVYAASRNSNDYG
YWGQGTQVTVSSAAAGGGGSGGGGSGSNLVAQLENEVASLENENETLKKKNLHKKDL
IAYLEKEIANLRKKIEEGGGGSGGGGSGSGLEHHHHHH

>D-2
EVQLQASGGGLVQDGGSLRLSCAVAGRPLSDYGVGWFRQASGKEREFVAVISGSGIV
TDYANSVKGRFTISRDVAKNVVHLQMNSLKPEDTAVYYCAALTNPVYAASRNSNDYG
YWGQGTQVTVSSAAAGGGGSGGGGSGSARNAYLRKKIARLKKDNLQLERDEQNLEKI
IANLRDEIARLENEVASHEQGGGGSGGGGSGSGLEHHHHHH

>D-3
EVQLQASGGGLVQDGGSLRLSCAVAGRPLSDYGVGWFRQASGKEREFVAVISGSGIV
TDYANSVKGRFTISRDVAKNVVHLQMNSLKPEDTAVYYCAALTNPVYAASRNSNDYG
YWGQGTQVTVSSAAAGGGGSGGGGSGSNEVTTLENDAAFIENENAYLEKEIARLRKE
KAALRNRLAHKKGGGGSGGGGSGSGLEHHHHHH

>D-4
EVQLQASGGGLVQDGGSLRLSCAVAGRPLSDYGVGWFRQASGKEREFVAVISGSGIV
TDYANSVKGRFTISRDVAKNVVHLQMNSLKPEDTAVYYCAALTNPVYAASRNSNDYG
YWGQGTQVTVSSAAAGGGGSGGGGSGSQKVAELKNRVAVKLNRNEQLKNKVEELKNR
NAYLKNELATLENEVARLENDVAEGGGGSGGGGSGSGLEHHHHHH

>D-1-3
EVQLQASGGGLVQDGGSLRLSCAVAGRPLSDYGVGWFRQASGKEREFVAVISGSGIV
TDYANSVKGRFTISRDVAKNVVHLQMNSLKPEDTAVYYCAALTNPVYAASRNSNDYG
YWGQGTQVTVSSAAAGGGGSGGGGSGSNLVAQLENEVASLENENETLKKKNLHKKDL
IAYLEKEIANLRKKIEEGGGGSGGGGSGSNEVTTLENDAAFIENENAYLEKEIARLR
KEKAALRNRLAHKKGGGGSGGGGSGSGLEHHHHHH

>D-3-1
EVQLQASGGGLVQDGGSLRLSCAVAGRPLSDYGVGWFRQASGKEREFVAVISGSGIV
TDYANSVKGRFTISRDVAKNVVHLQMNSLKPEDTAVYYCAALTNPVYAASRNSNDYG
YWGQGTQVTVSSAAAGGGGSGGGGSGSNEVTTLENDAAFIENENAYLEKEIARLRKE
KAALRNRLAHKKGGGGSGGGGSGSNLVAQLENEVASLENENETLKKKNLHKKDLIAY
LEKEIANLRKKIEEGGGGSGGGGSGSGLEHHHHHH

>D-5

EVQLQASGGGLVQDGGSLRLSCAVAGRPLSDYGVGWFRQASGKEREFVAVISGSGIV
TDYANSVKGRFTISRDVAKNVVHLQMNSLKPEDTAVYYCAALTNPVYAASRNSNDYG
YWGQGTQVTVSSAAAGGGGSGGGGSGSGSNTVKELKNYIQELEERNAELKNLKEHLK
FAKAELEFELAAHKFEGGGGSGGGGSGSGLEHHHHHH

>D-6
EVQLQASGGGLVQDGGSLRLSCAVAGRPLSDYGVGWFRQASGKEREFVAVISGSGIV
TDYANSVKGRFTISRDVAKNVVHLQMNSLKPEDTAVYYCAALTNPVYAASRNSNDYG
YWGQGTQVTVSSAAAGGGGSGGGGSGSGSQKVAQLKNRVAYKLKENAKLENIVARLE
NDNANLEKDIANLEKDIANLERDVARGGGGSGGGGSGSGLEHHHHHH

>D-17
EVQLQASGGGLVQDGGSLRLSCAVAGRPLSDYGVGWFRQASGKEREFVAVISGSGIV
TDYANSVKGRFTISRDVAKNVVHLQMNSLKPEDTAVYYCAALTNPVYAASRNSNDYG
YWGQGTQVTVSSAAAGGGGSGGGGSGSGSNEKEELKSKKAELRNRIEQLKQKREQLK
QKIANLRKEIEAYKGGGGSGGGGSGSGLEHHHHHH

>D-18
EVQLQASGGGLVQDGGSLRLSCAVAGRPLSDYGVGWFRQASGKEREFVAVISGSGIV
TDYANSVKGRFTISRDVAKNVVHLQMNSLKPEDTAVYYCAALTNPVYAASRNSNDYG
YWGQGTQVTVSSAAAGGGGSGGGGSGSGSSIAATLENDLARLENENARLEKDIANLE
RDLAKLEREEAYFGGGGSGGGGSGSGLEHHHHHH

>D-E34m
EVQLQASGGGLVQDGGSLRLSCAVAGRPLSDYGVGWFRQASGKEREFVAVISGSGIV
TDYANSVKGRFTISRDVAKNVVHLQMNSLKPEDTAVYYCAALTNPVYAASRNSNDYG
YWGQGTQVTVSSAAAGGGGSGGGGSGSITIRAAFLEKENTALRTEIAELEKEVGRSE
NIVSKYETRYGPLGSGGSGSGGSGSGLEHHHHHH

>D-R34m
EVQLQASGGGLVQDGGSLRLSCAVAGRPLSDYGVGWFRQASGKEREFVAVISGSGIV
TDYANSVKGRFTISRDVAKNVVHLQMNSLKPEDTAVYYCAALTNPVYAASRNSNDYG
YWGQGTQVTVSSAAAGGGGSGGGGSGSLEIRAAFLEKENTALRTRAAELRKVGRSRN
IVSKYETRYGPLGSGGSGSGGSGSGLEHHHHHH

>A-1 _
EFARSDVQLVESGGGLVQPGGSLRLTCAASGLIFGSYAMGWFRQAPGKAREFVAAIS
WSGGDTYADSVKGRFTISRDNAKNTVYLQMNSLEPEDTAVYSCAAVGSKYYISKDAK
DYGYWGQGTQVTVSSAAAGGGGSGGGGSGSNLVAQLENEVASLENENETLKKKNLHK
KDLIAYLEKEIANLRKKIEEGGGGSGGGGSGSGLEHHHHHH

>A-2
EFARSDVQLVESGGGLVQPGGSLRLTCAASGLIFGSYAMGWFRQAPGKAREFVAAIS
WSGGDTYADSVKGRFTISRDNAKNTVYLQMNSLEPEDTAVYSCAAVGSKYYISKDAK
DYGYWGQGTQVTVSSAAAGGGGSGGGGSGSARNAYLRKKIARLKKDNLQLERDEQNL
EKNIANLRDEIARLENEVASHEQGGGGSGGGGSGSGLEHHHHHH

>A-3
EFARSDVQLVESGGGLVQPGGSLRLTCAASGLIFGSYAMGWFRQAPGKAREFVAAIS
WSGGDTYADSVKGRFTISRDNAKNTVYLQMNSLEPEDTAVYSCAAVGSKYYISKDAK
DYGYWGQGTQVTVSSAAAGGGGSGGGGSGSNEVTTLENDAAFIENENAYLEKEIARL
RKEKAALRNRLAHKKGGGGSGGGGSGSGLEHHHHHH

>A-4
EFARSDVQLVESGGGLVQPGGSLRLTCAASGLIFGSYAMGWFRQAPGKAREFVAAIS
WSGGDTYADSVKGRFTISRDNAKNTVYLQMNSLEPEDTAVYSCAAVGSKYYISKDAK
DYGYWGQGTQVTVSSAAAGGGGSGGGGSGSQKVAELKNRVAVKLNRNEQLKNKVEEL
KNRNAYLKNELATLENEVARLENDVAEGGGGSGGGGSGSGLEHHHHHH

>1-A
EFARSNLVAQLENEVASLENENETLKKKNLHKKDLIAYLEKEIANLRKKIEVGGGGS
GGGGSGSDVQLVESGGGLVQPGGSLRLTCAASGLIFGSYAMGWFRQAPGKAREFVAA
VSWSGGDTYADSVKGRFTISRDNAKNTVYLQMNSLEPEDTAVYSCAAVGSKYYISKD
AKDYGYWGQGTQVTVSSAAAGGGGSGGGGSGSGLEHHHHHH

>3-A
EFARSNEVTTLENDAAFIENENAYLEKEIARLRKEKAALRNRLAHKKGGGGSGGGGS
GSDVQLVESGGGLVQPGGSLRLTCAASGLIFGSYAMGWFRQAPGKAREFVAAISWSG
GDTYADSVKGRFTISRDNAKNTVYLQMNSLEPEDTAVYSCAAVGSKYYISKDAKDYG
YWGQGTQVTVSSAAAGGGGSGGGGSGSGLEHHHHHH

>4-A
EFARSQKVAELKNRVAVKLNRNEQLKNKVEELKNRNAYLKNELATLENEVARLENDV
AEGGGGSGGGGSGSDVQLVESGGGLVQPGGSLRLTCAASGLIFGSYAMGWFRQAPGK
AREFVAAISWSGGDTYADSVKGRFTISRDNAKNTVYLQMNSLEPEDTAVYSCAAVGS
KYYISKDAKDYGYWGQGTQVTVSSAAAGGGGSGGGGSGSGLEHHHHHH
>A-1-A

EFARSDVQLVESGGGLVQPGGSLRLTCAASGLIFGSYAMGWFRQAPGKAREFVAAIS
WSGGDTYADSVKGRFTISRDNAKNTVYLQMNSLEPEDTAVYSCAAVGSKYYISKDAK
DYGYWGQGTQVTVSSAAAGGGGSGGGGSGSNLVAQLENEVASLENENETLKKKNLHK
KDLIAYLEKEIANLRKKIEEGGGGSGGGGSGSDVQLVESGGGLVQPGGSLRLTCAAS
GLIFGSYAMGWFRQAPGKAREFVAAISWSGGDTYADSVKGRFTISRDNAKNTVYLQM
NSLEPEDTAVYSCAAVGSKYYISKDAKDYGYWGQGTQVTVSSAAAGGGGSGGGGSGS
GLEHHHHHH

>A-5
EFARSDVQLVESGGGLVQPGGSLRLTCAASGLIFGSYAMGWFRQAPGKAREFVAAIS
WSGGDTYADSVKGRFTISRDNAKNTVYLQMNSLEPEDTAVYSCAAVGSKYYISKDAK
DYGYWGQGTQVTVSSAAAGGGGSGGGGSGSGSNTVKELKNYIQELEERNAELKNLKE
HLKFAKAELEFELAAHKFEGGGGSGGGGGGSGLEHHHHHH

>A-6
EFARSDVQLVESGGGLVQPGGSLRLTCAASGLIFGSYAMGWFRQAPGKAREFVAAIS
WSGGDTYADSVKGRFTISRDNAKNTVYLQMNSLEPEDTAVYSCAAVGSKYYISKDAK
DYGYWGQGTQVTVSSAAAGGGGSGGGGSGSGSQKVAQLKNRVAYKLKENAKLENIVA
RLENDNANLEKDIANLEKDIANLERDVARGGGGSGGGGSGSGLEHHHHHH

>A-17
EFARSDVQLVESGGGLVQPGGSLRLTCAASGLIFGSYAMGWFRQAPGKAREFVAAIS
WSGGDTYADSVKGRFTISRDNAKNTVYLQMNSLEPEDTAVYSCAAVGSKYYISKDAK
DYGYWGQGTQVTVSSAAAGGGGSGGGGSGSGSNEKEELKSKKAELRNRIEQLKQKRE
QLKQKIANLRKEIEAYKGGGGSGGGGSGSGLEHHHHHH

>A-18
EFARSDVQLVESGGGLVQPGGSLRLTCAASGLIFGSYAMGWFRQAPGKAREFVAAIS
WSGGDTYADSVKGRFTISRDNAKNTVYLQMNSLEPEDTAVYSCAAVGSKYYISKDAK
DYGYWGQGTQVTVSSAAAGGGGSGGGGSGSGSSIAATLENDLARLENENARLEKDIA
NLERDLAKLEREEAYFGGGGSGGGGSGSGLEHHHHHH

>A-E34m
EFARSDVQLVESGGGLVQPGGSLRLTCAASGLIFGSYAMGWFRQAPGKAREFVAAIS
WSGGDTYADSVKGRFTISRDNAKNTVYLQMNSLEPEDTAVYSCAAVGSKYYISKDAK
DYGYWGQGTQVTVSSAAAGGGGSGGGGSGSITIRAAFLEKENTALRTEIAELEKEVG
RSENIVSKYETRYGPLGSGGSGSGGSGSGLEHHHHHH

>A-R34m
EFARSDVQLVESGGGLVQPGGSLRLTCAASGLIFGSYAMGWFRQAPGKAREFVAAIS
WSGGDTYADSVKGRFTISRDNAKNTVYLQMNSLEPEDTAVYSCAAVGSKYYISKDAK
DYGYWGQGTQVTVSSAAAGGGGSGGGGSGSLEIRAAFLEKENTALRTRAAELRKRVG
RSRNIVSKYETRYGPLGSGGSGSGGSGSGLEHHHHHH

Figure S2. Cartoon of the steps involved in the SPR experiments. In step 1, ricin is covalently immobilized on all columns of the SPR chip. Step 2 consists of flowing six D-SYNZIP constructs, each in its own column (channel). The D portion of the construct binds tightly to the immobilized ricin, serving to orient the zipper domain. For step 3, serial dilutions of an A-SYNZIP construct are run across the channel in six rows; the dilutions always include a no-protein blank. This step enables the calculation of the on rate constants. In the final step (step4), buffer is flowed over each row. This enables the calculation of the off rate constant. The process enables the determination of both specificity and binding kinetics for each pair.

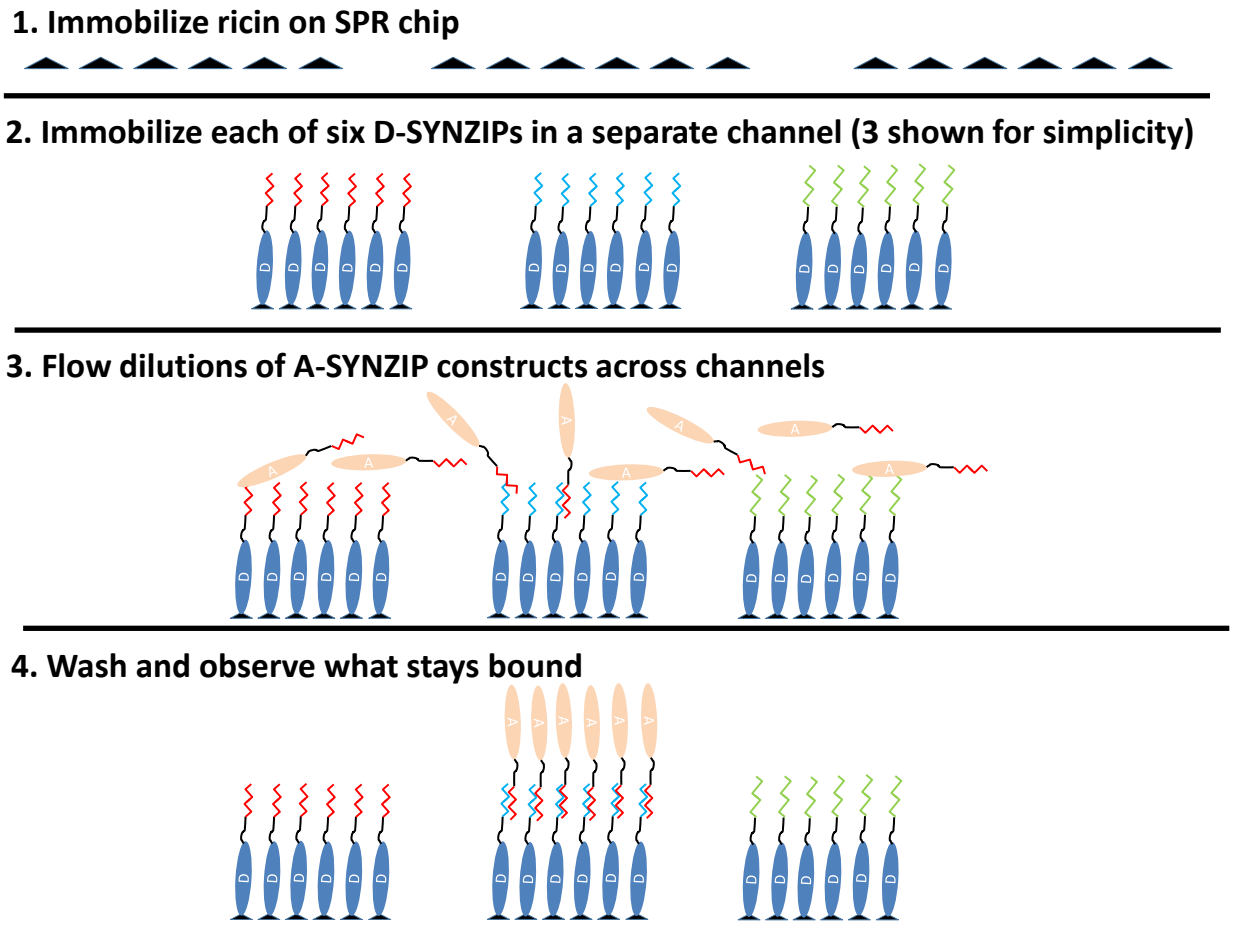
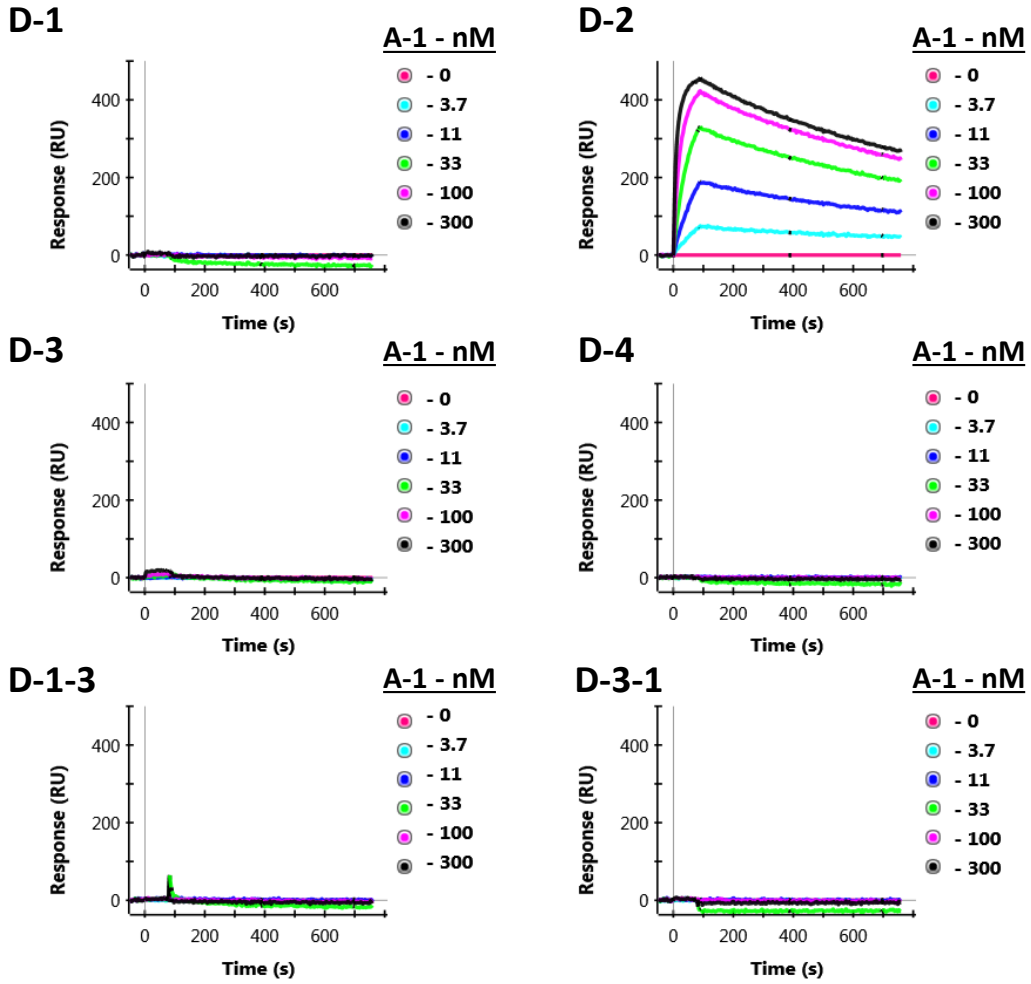


Figure S3. SPR data. Panels A-O follow (on pages S-10 through S24) showing representative SPR data collected in this work. Data was collected on two chips. One chip had the proteins D-1, D-2, D-3, D-4, D-1-3, D-3-1 immobilized each in its own column. The second chip had D-5, D-6, D-17, D-18, D-E34m, D-R34m immobilized in their own columns. Concentrations of the partner SYNZIP construct flowed over the chip are in nM and noted to the right of each trace. In panels A-N, on top we show the data for binding to each of the six immobilized D-SYNZIP constructs and on the bottom we show enlarged versions of the curve showing binding to the primary binding partner. Panel O shows 2 enlarged curves. All the enlarged curve shows the fit as well as the binding kinetics and dissociation constant. Each panel has a separate sub-legend describing the interactions that are shown on that page.

Figure S3-A: Top- A1 binding to D-1, D-1, D-3, D-4, D-1-3, D-3-1; Bottom – Enlarged view of A-1 binding to D-2 with affinity constants shown.



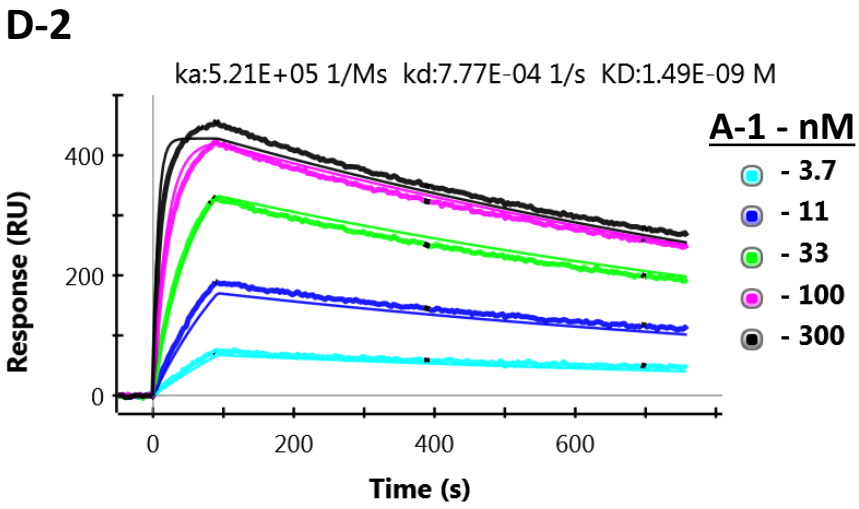
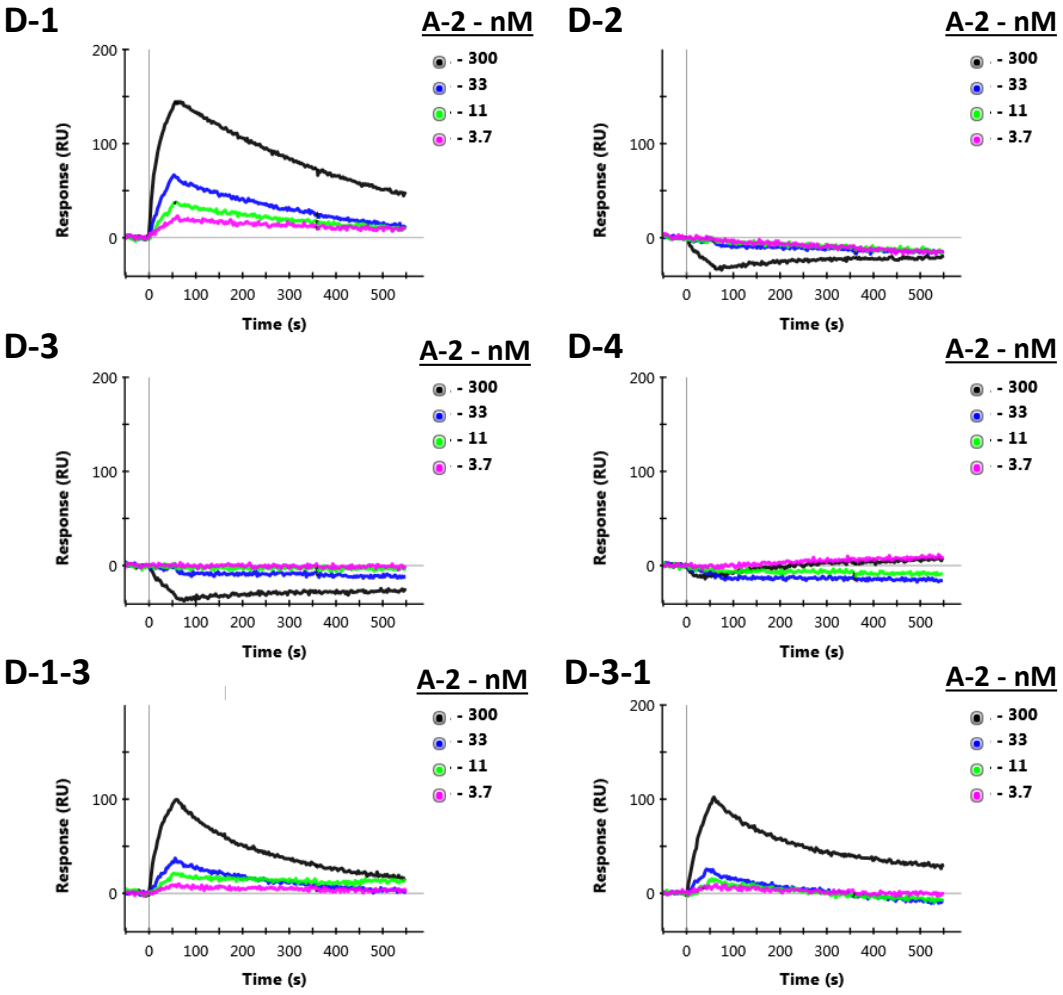


Figure S3-B: Top- A2 binding to D-1, D-1, D-3, D-4, D-1-3, D-3-1; Bottom – Enlarged view of A-2 binding to D-1 with affinity constants shown.



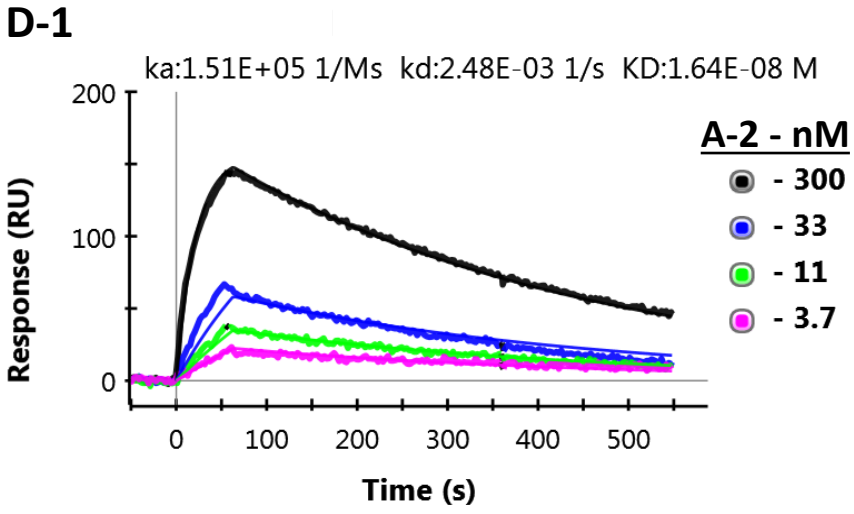
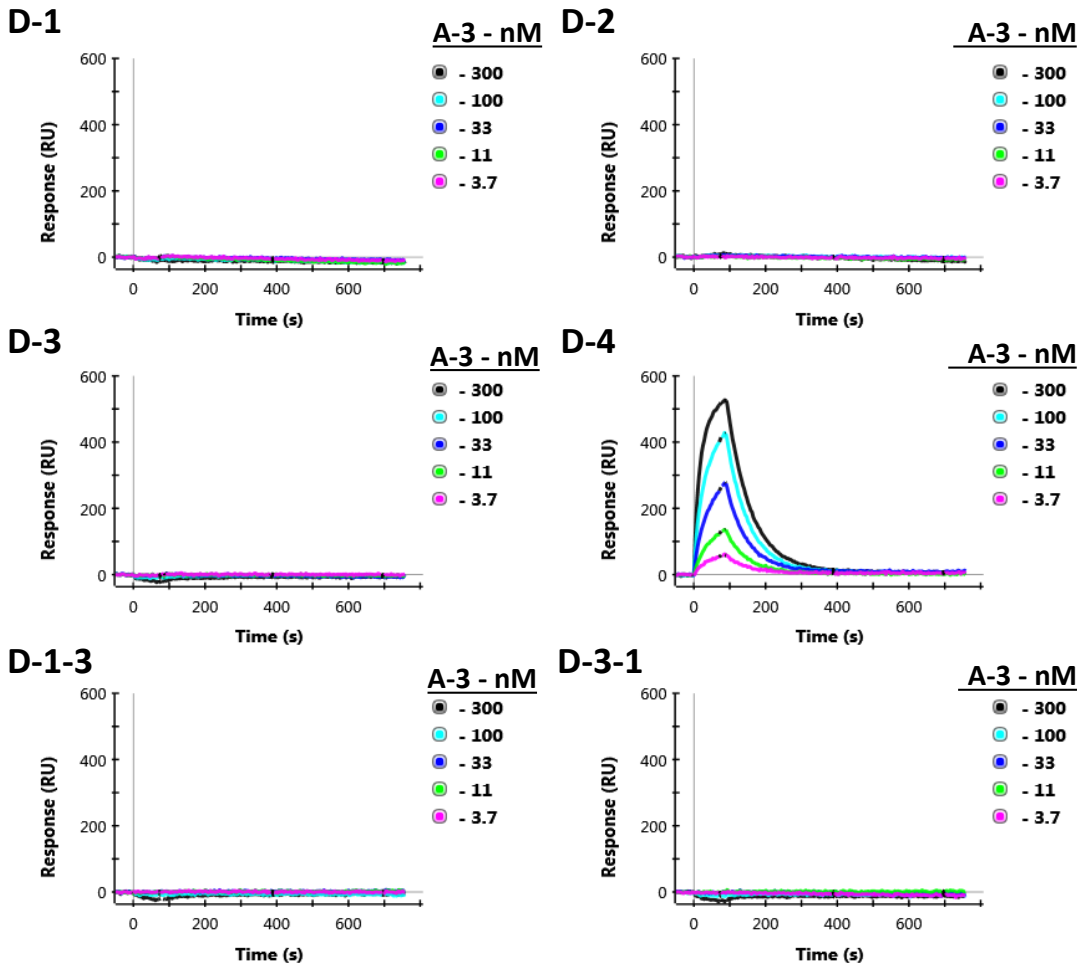


Figure S3-C: Top- A3 binding to D-1, D-1, D-3, D-4, D-1-3, D-3-1; Bottom – Enlarged view of A-3 binding to D-4 with affinity constants shown.



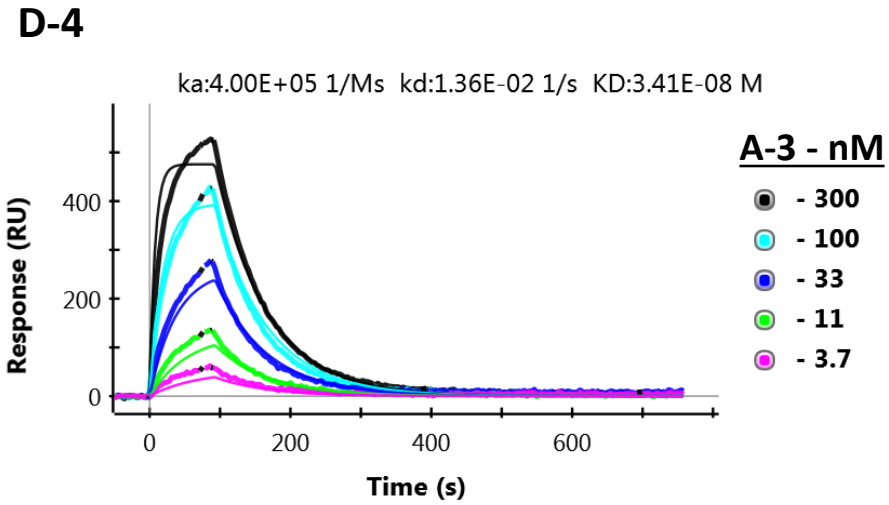
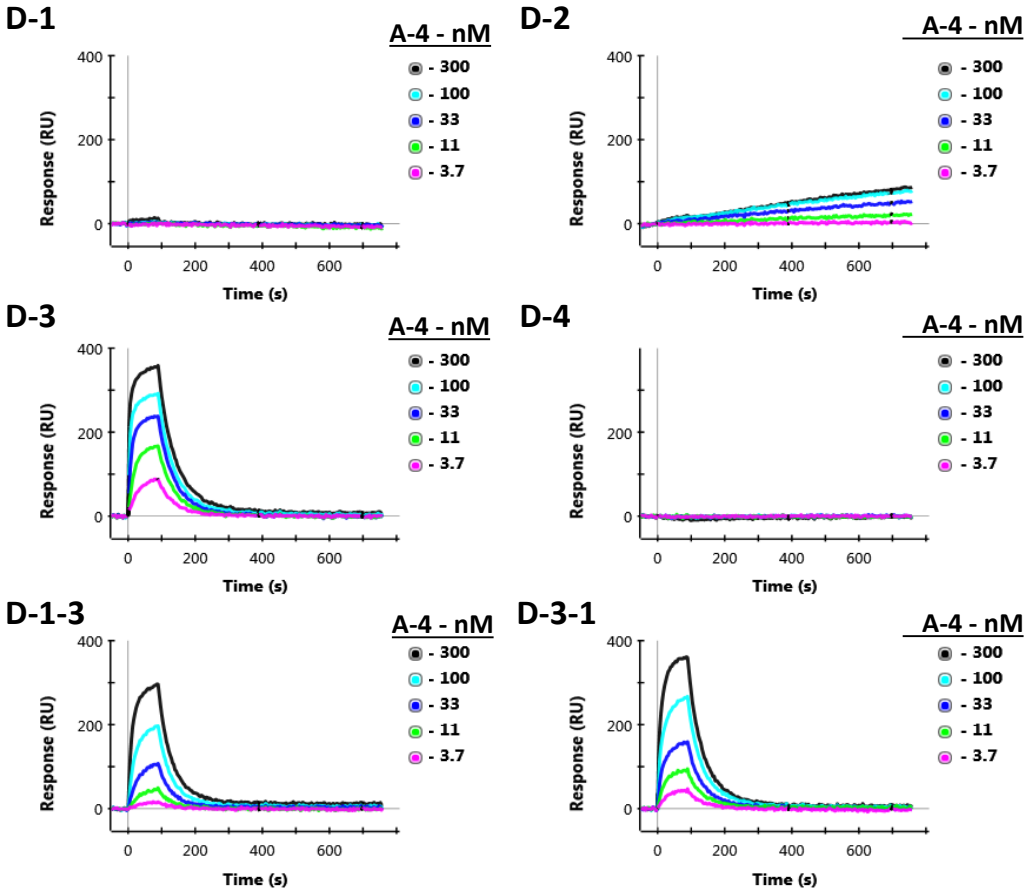


Figure S3-D: Top- A4 binding to D-1, D-1, D-3, D-4, D-1-3, D-3-1; Bottom – Enlarged view of A-3 binding to D-3 with affinity constants shown.



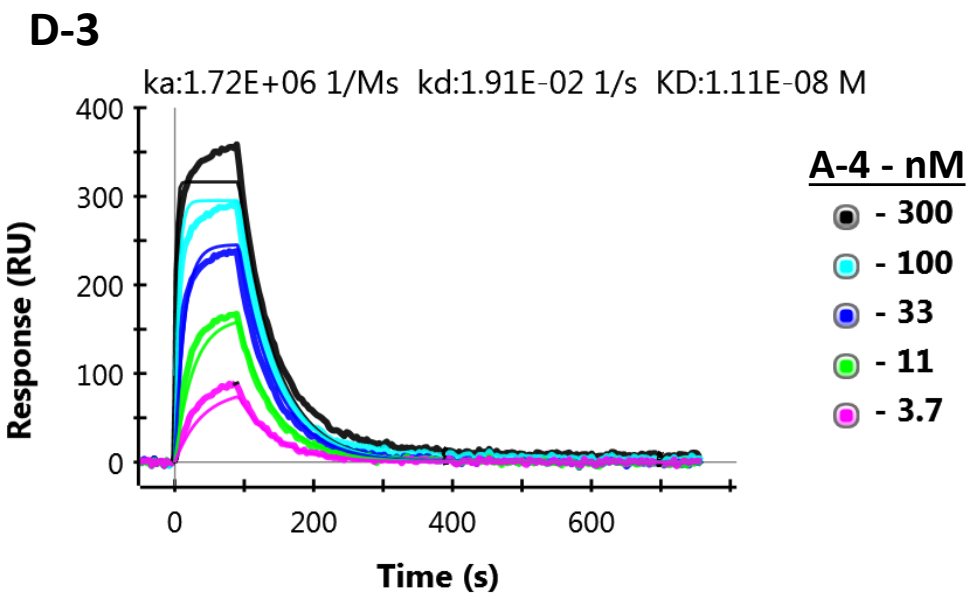
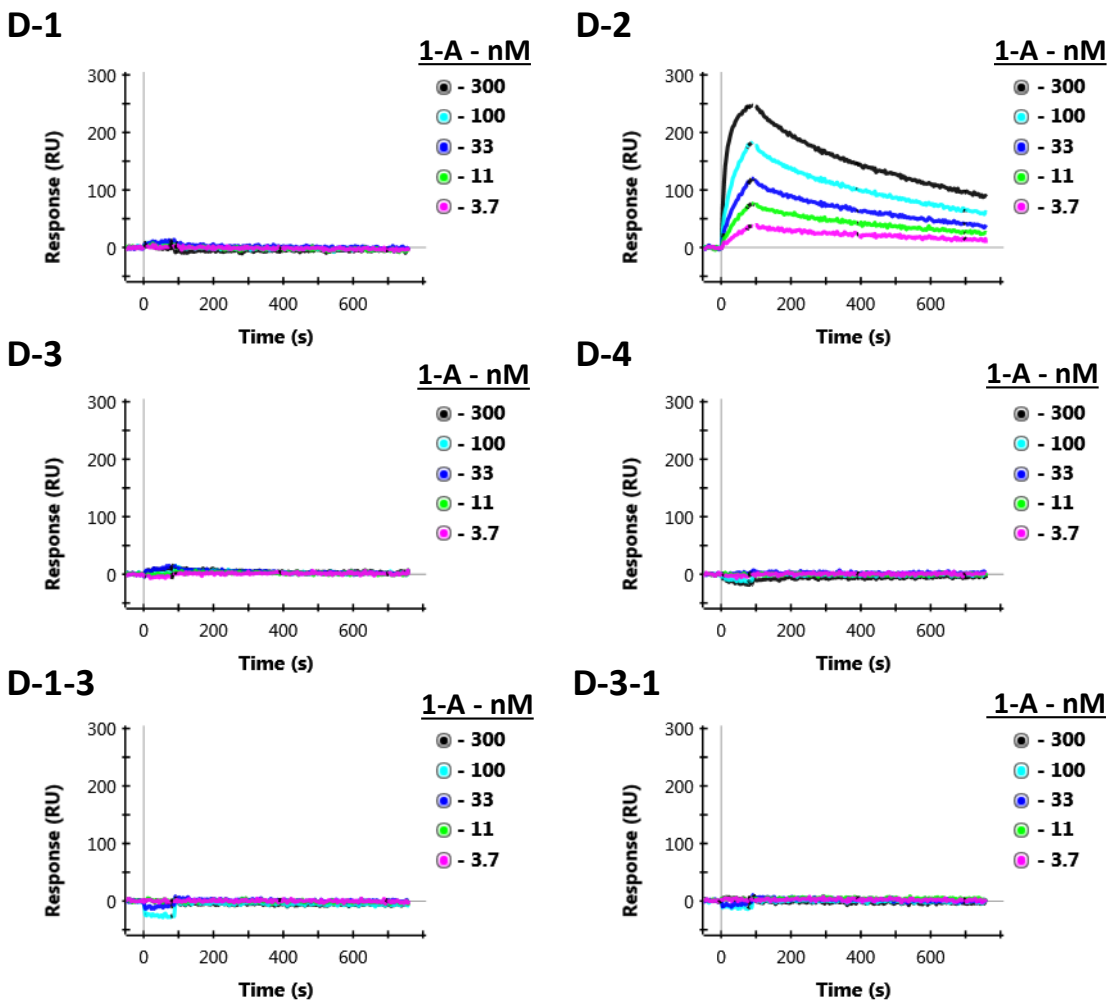


Figure S3-E: Top- 1-A binding to D-1, D-1, D-3, D-4, D-1-3, D-3-1; Bottom – Enlarged view of 1-A binding to D-2 with affinity constants shown.



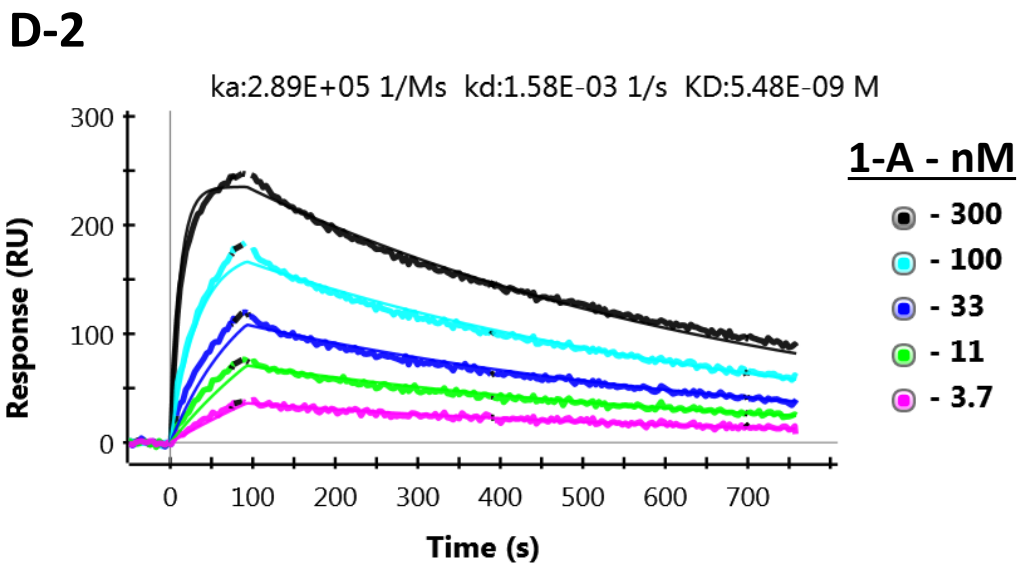
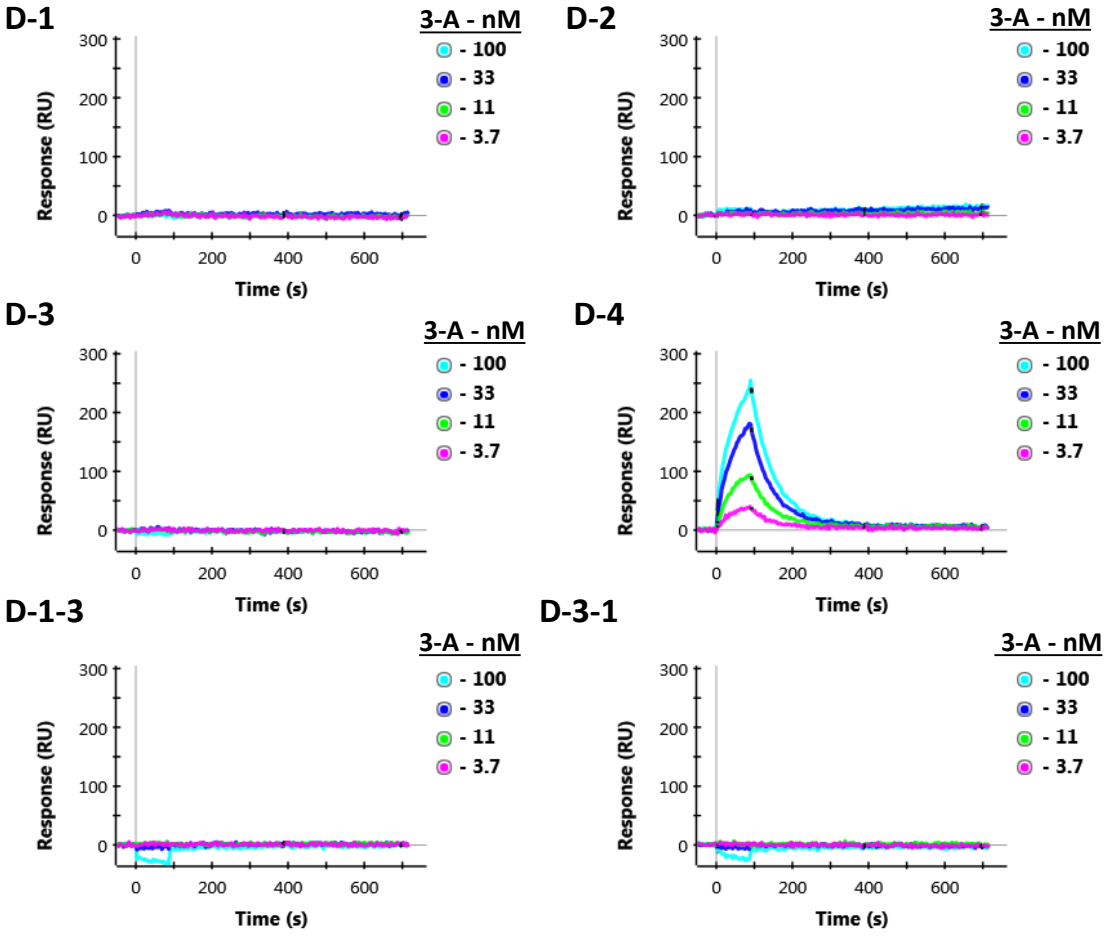


Figure S3-F Top- 3-A binding to D-1, D-1, D-3, D-4, D-1-3, D-3-1; Bottom – Enlarged view of 3-A binding to D-4 with affinity constants shown.



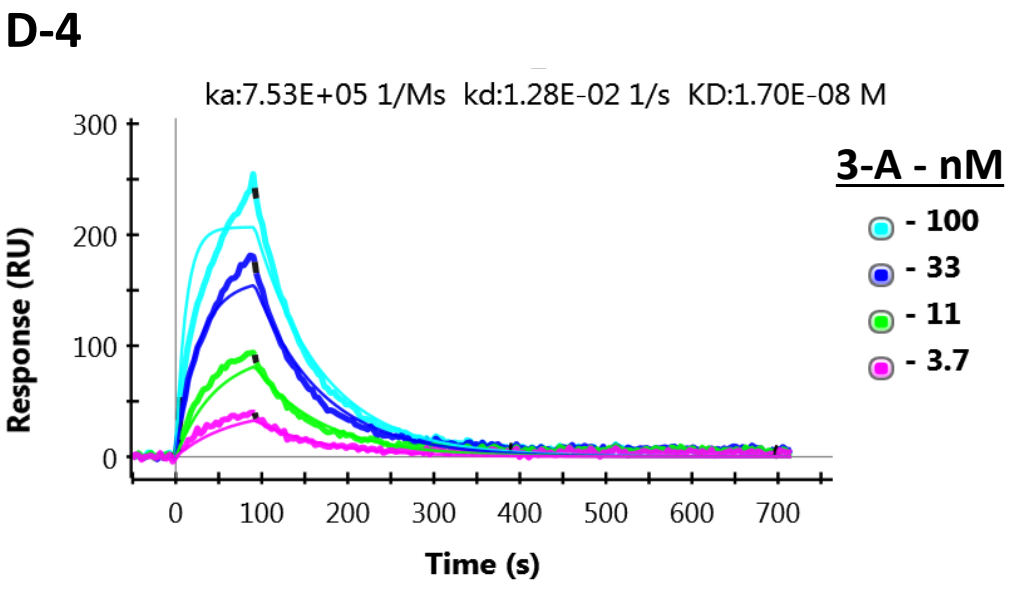
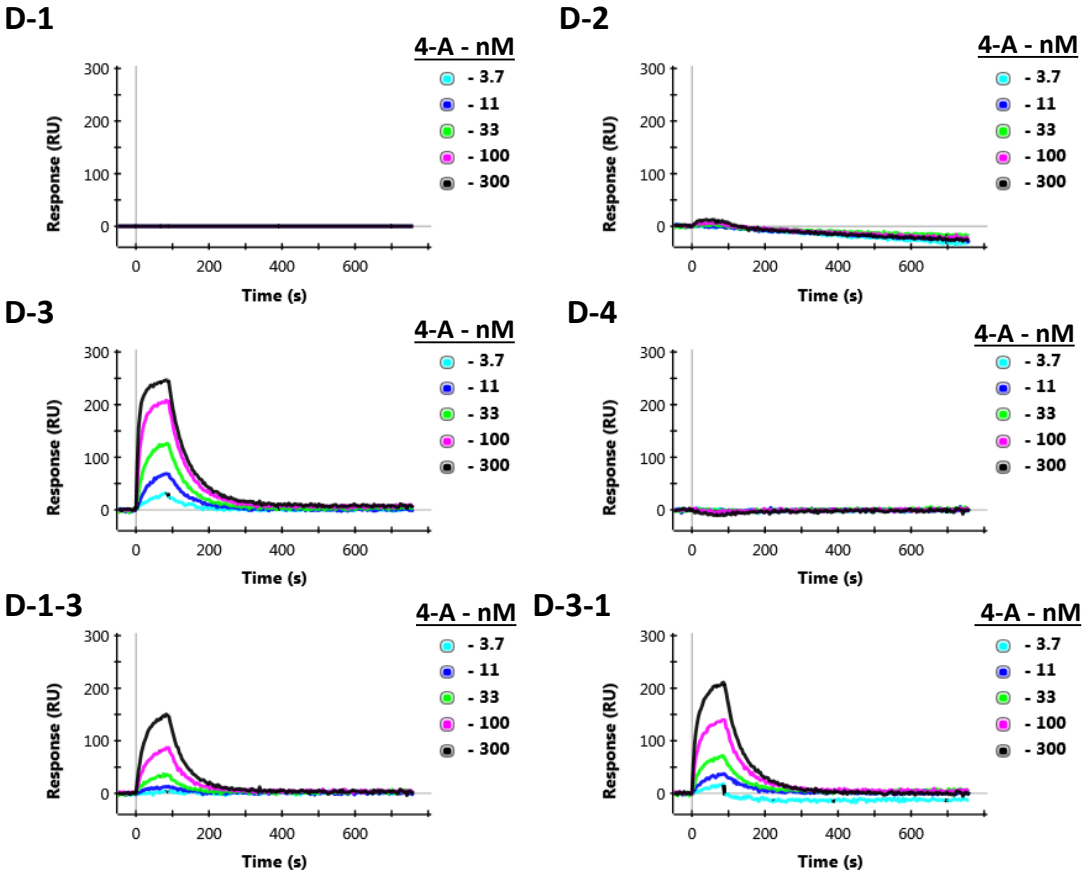


Figure S3-G: Top- 4-A binding to D-1, D-1, D-3, D-4, D-1-3, D-3-1; Bottom – Enlarged view of 4-A binding to D-3 with affinity constants shown.



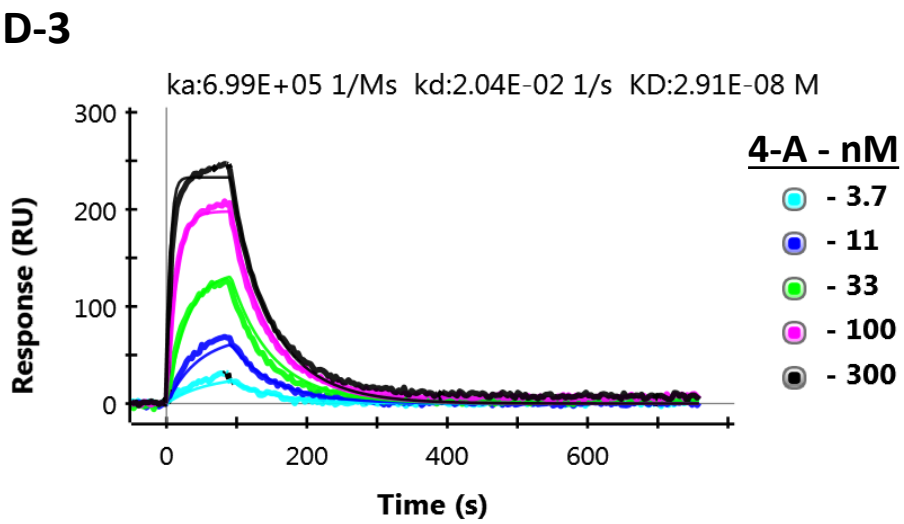
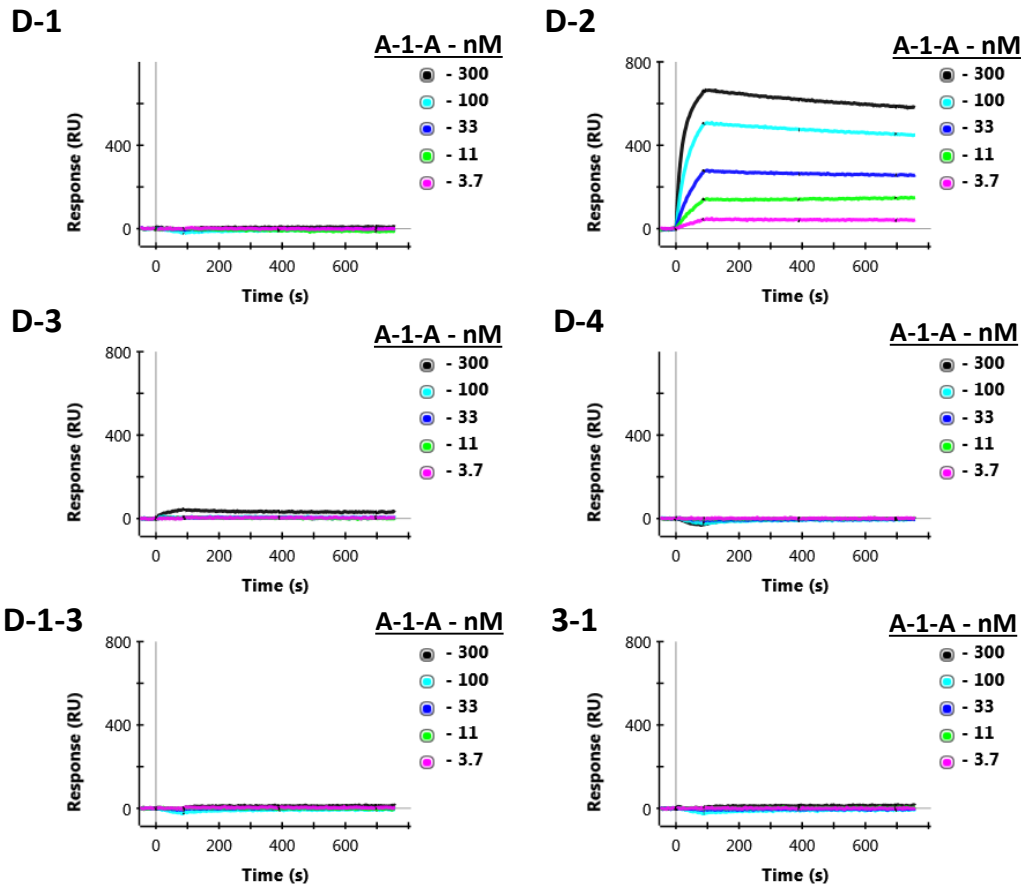


Figure S3-H: Top- A-1-A binding to D-1, D-1, D-3, D-4, D-1-3, D-3-1; Bottom – Enlarged view of A-1-A binding to D-2 with affinity constants shown.



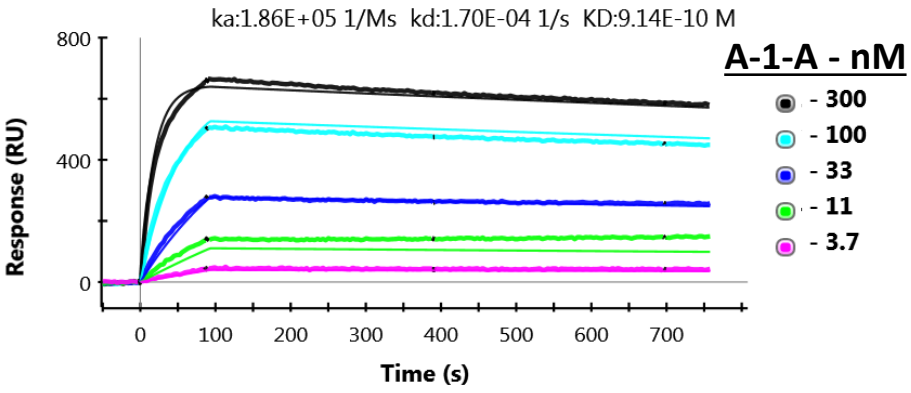
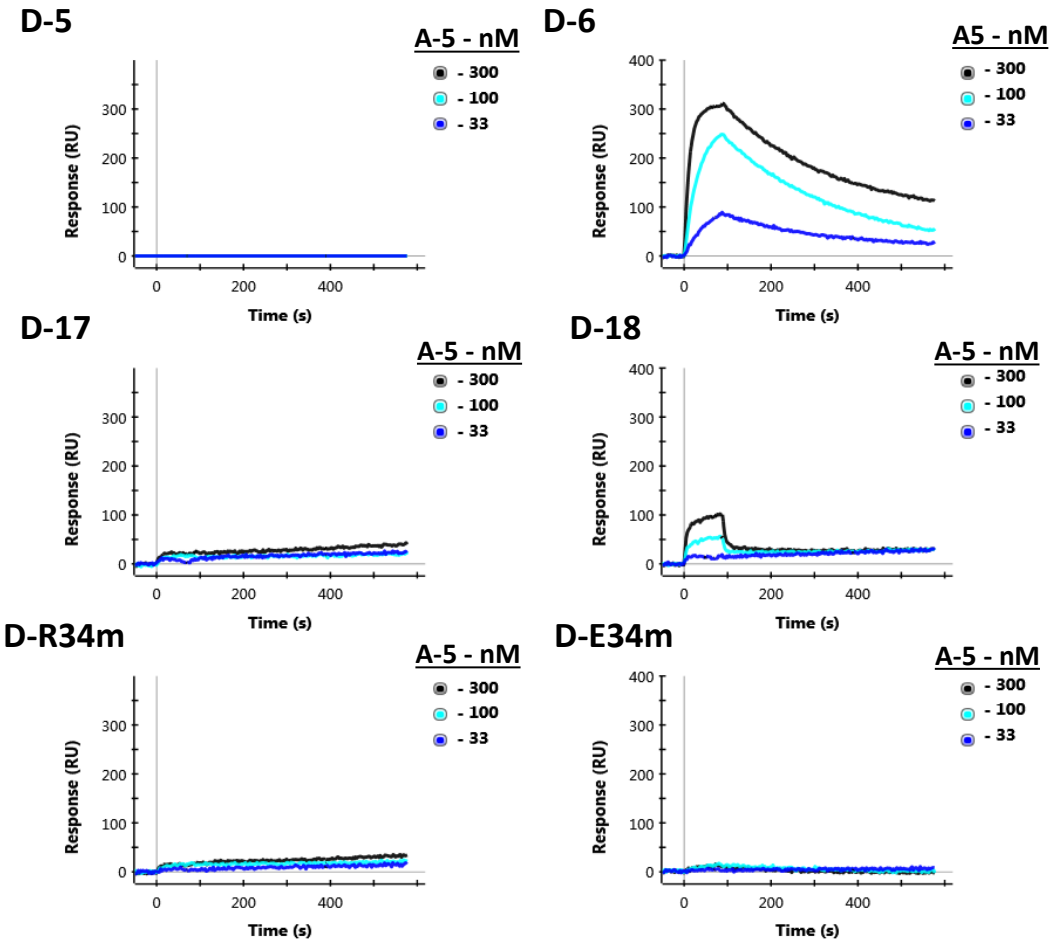


Figure S3-I: Top- A-5 binding to D-5, D-6, D-17, D-18, D-R34m, D-E34m; Bottom – Enlarged view of A-5 binding to D-6 with affinity constants shown.



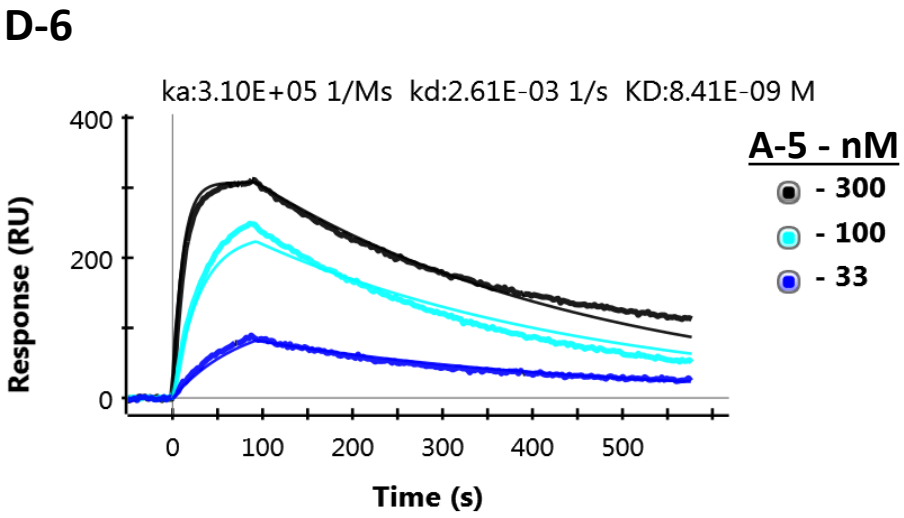
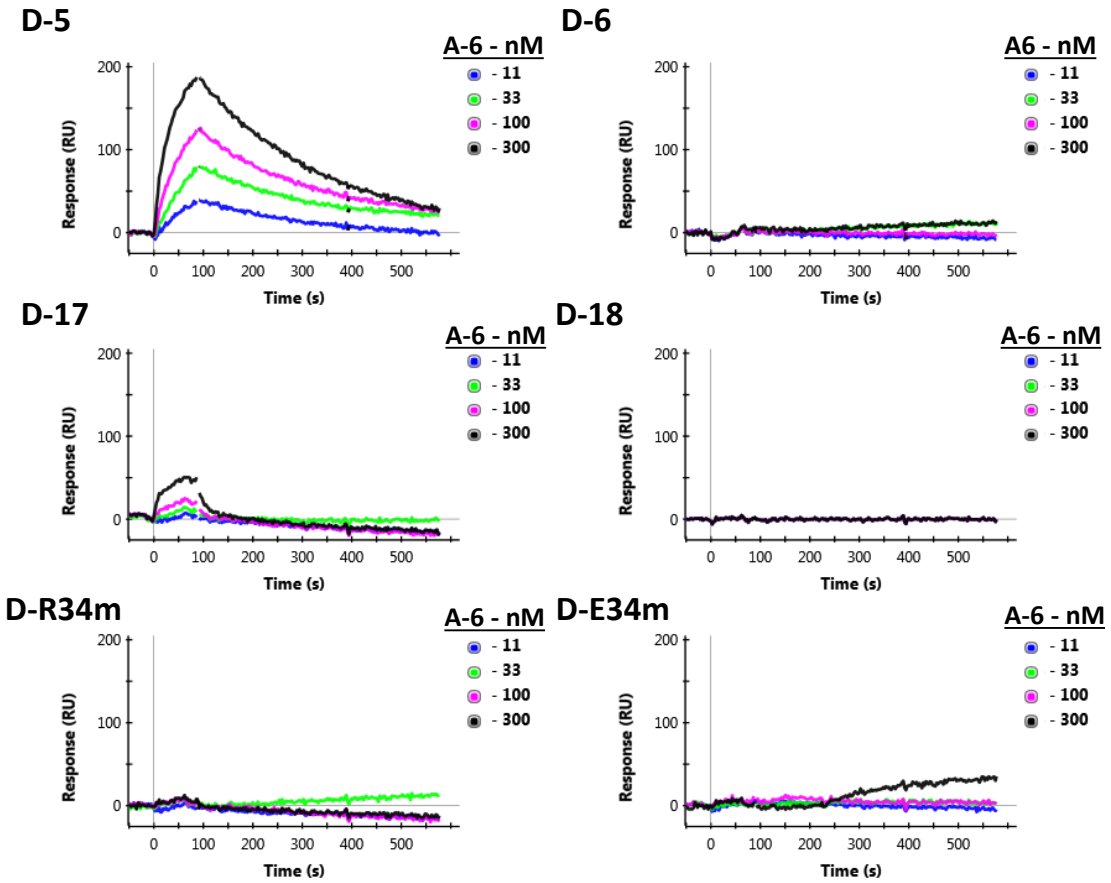


Figure S3-J: Top- A-6 binding to D-5, D-6, D-17, D-18, D-R34m, D-E34m; Bottom – Enlarged view of A-6 binding to D-5 with affinity constants shown.



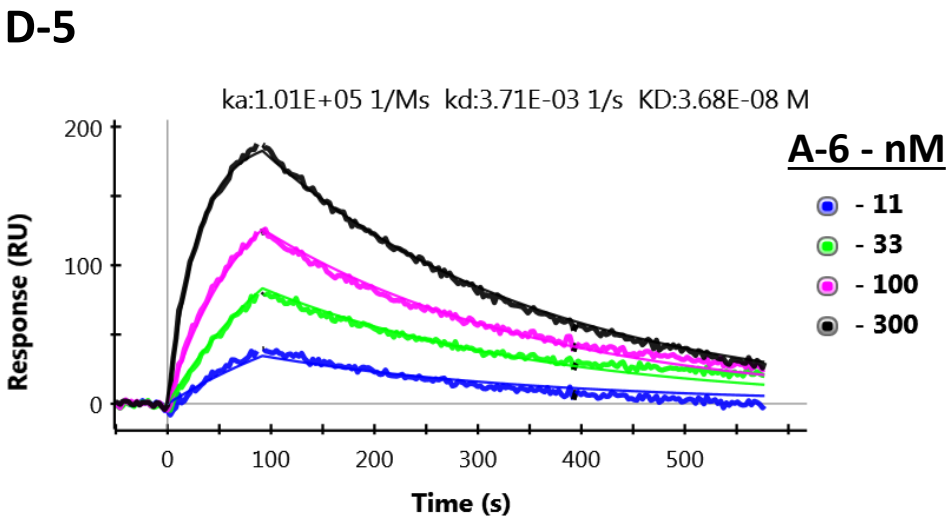
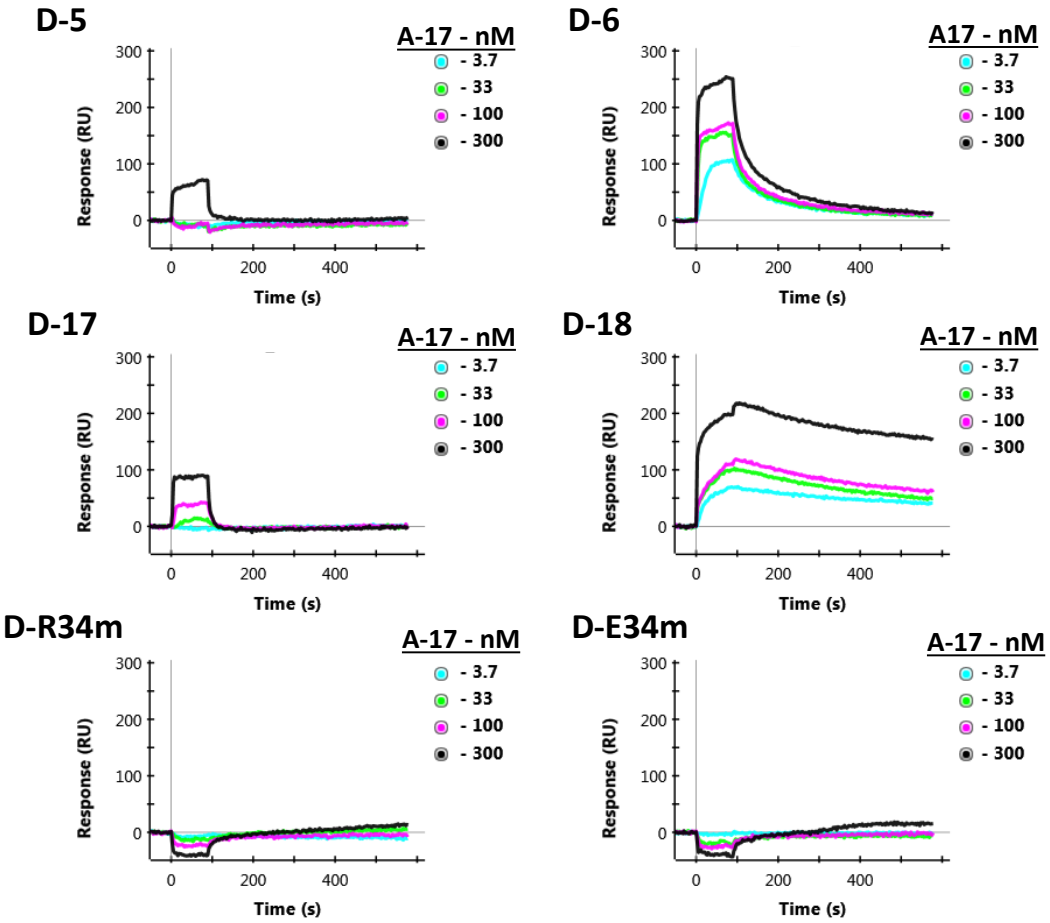


Figure S3-K: Top- A-17 binding to D-5, D-6, D-17, D-18, D-R34m, D-E34m; Bottom – Enlarged view of A-17 binding to D-18 with affinity constants shown.



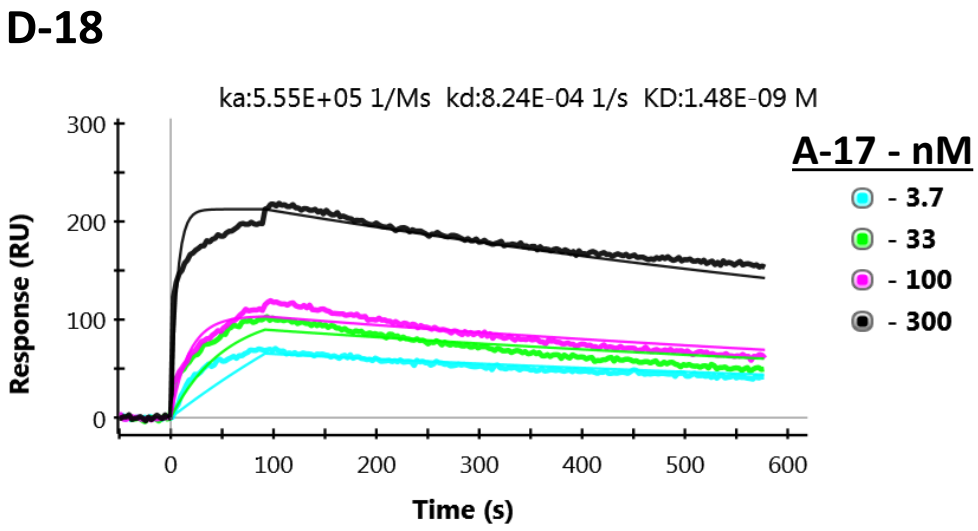
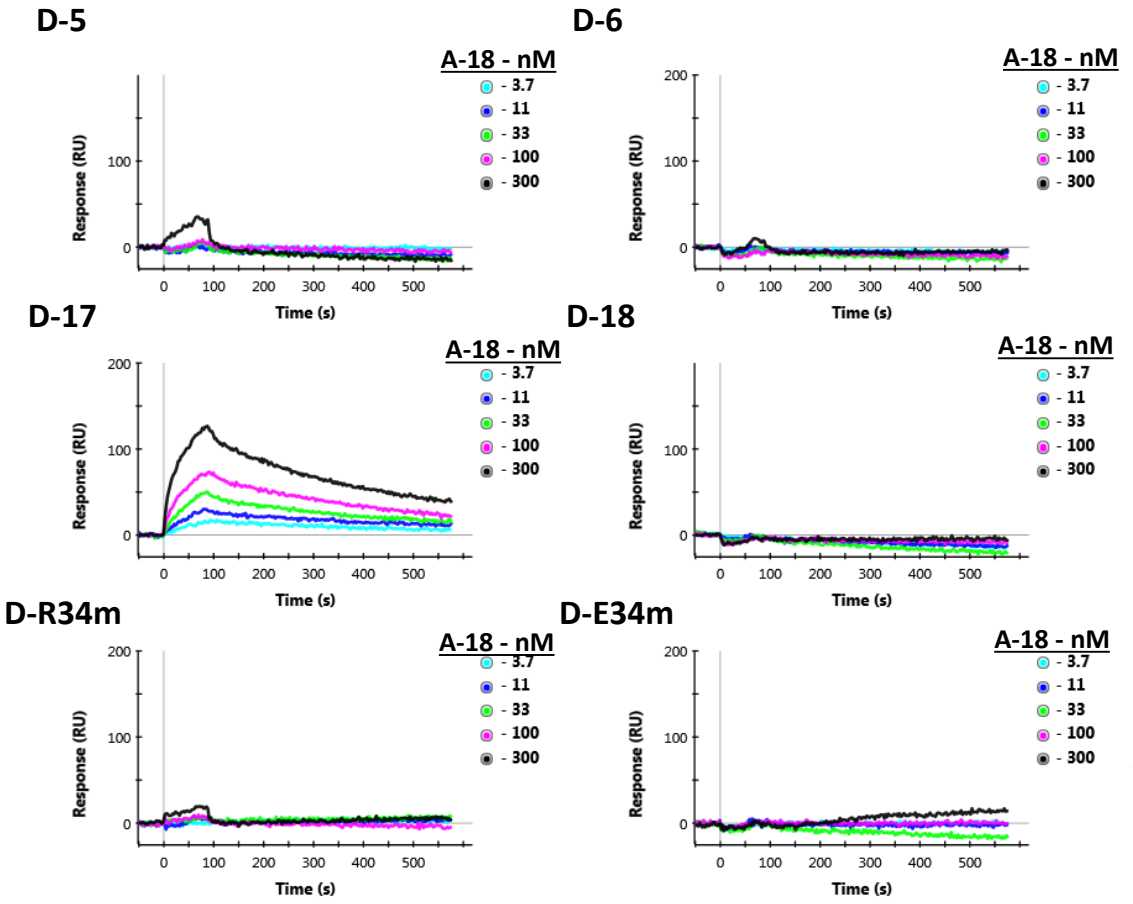


Figure S3-L Top- A-18 binding to D-5, D-6, D-17, D-18, D-R34m, D-E34m; Bottom – Enlarged view of A-18 binding to D-17 with affinity constants shown.



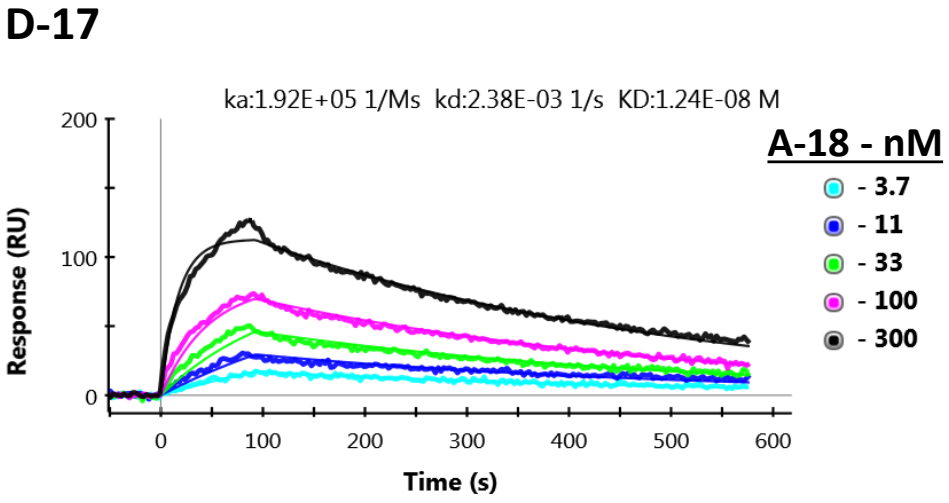
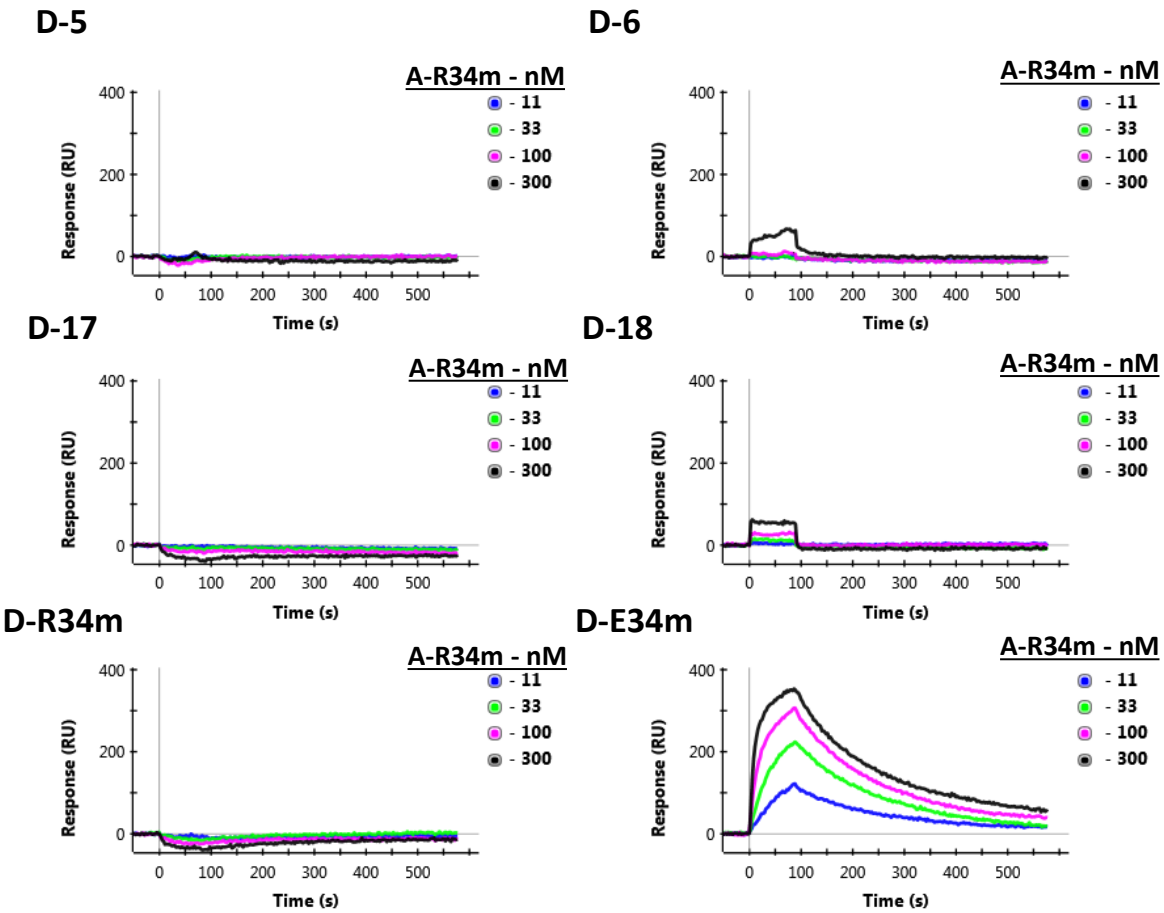


Figure S3-M: Top- A-R34m binding to D-5, D-6, D-17, D-18, D-R34m, D-E34m; Bottom – Enlarged view of A-R34m binding to D-E34m with affinity constants shown.



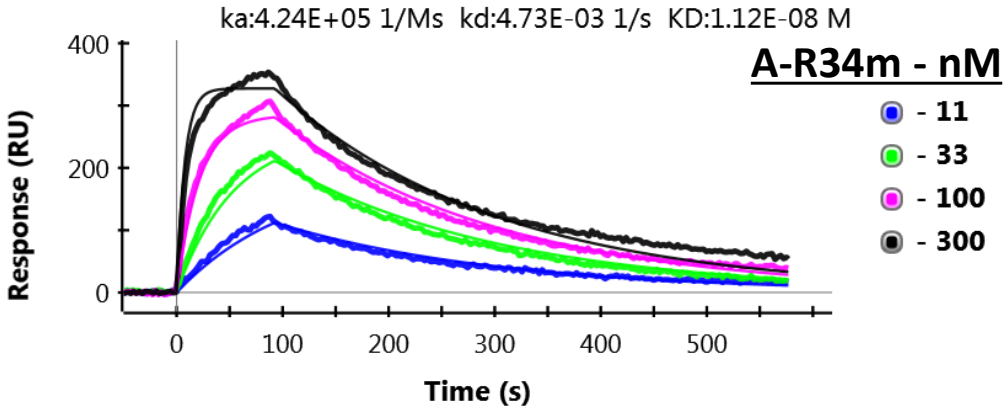
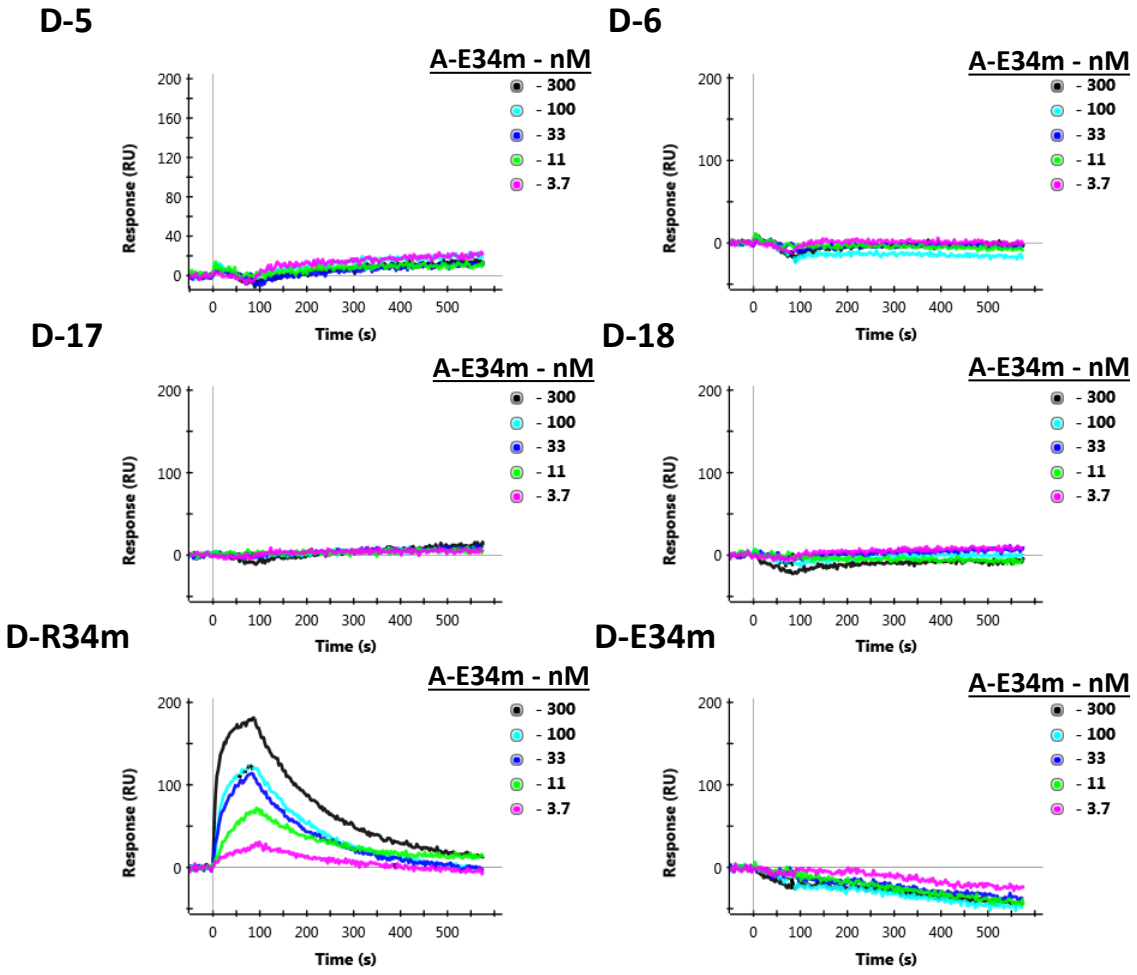


Figure S3-N: Top- A-18 binding to D-5, D-6, D-17, D-18, D-R34m, D-E34m; Bottom – Enlarged view of A-R34m binding to D-E34m with affinity constants shown



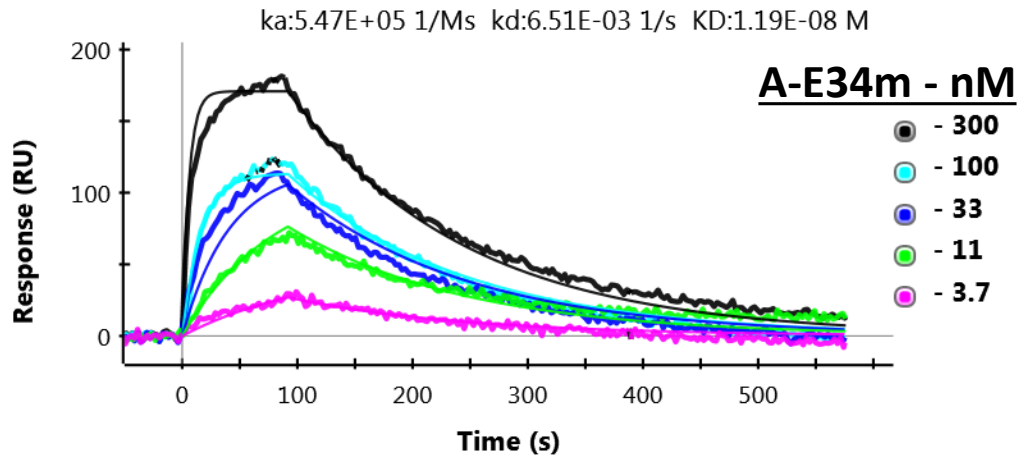


Figure S3-O: Top- A-3 binding to D-17; Bottom – A-4 binding to D-6 with affinity constants shown

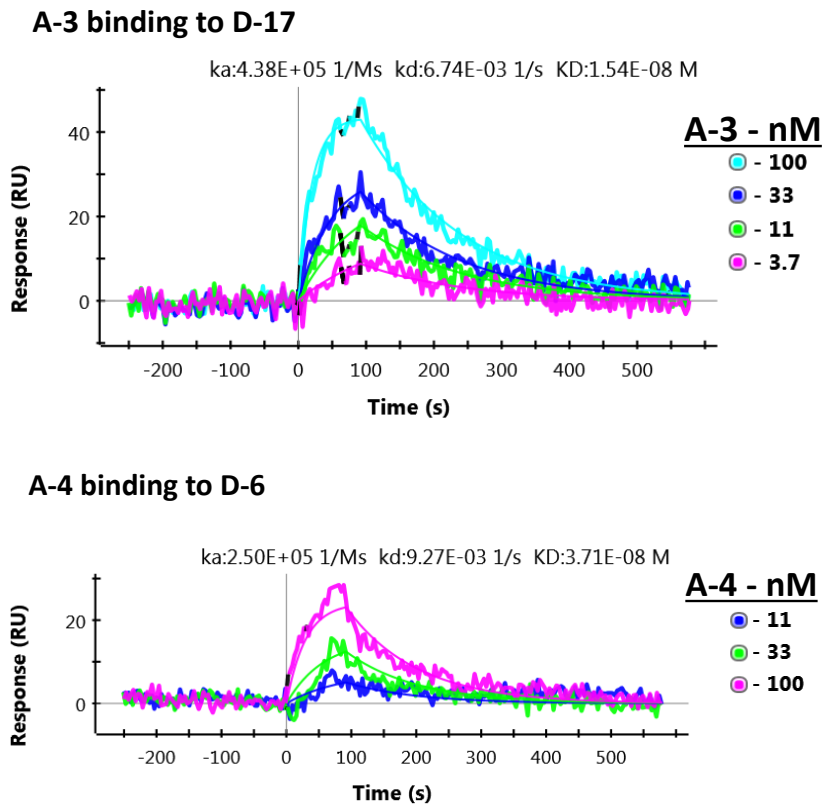


Figure S4. Size exclusion chromatography of individual components and mixes. The top panels show controls in which the D-3-1 (left) and D-1-3 (right) scaffolds were mixed with SYNZIPs A-1 and A-3. No assembly was observed with the control constructs (A-1 and A-3 with a D-1-3 or D-3-1 scaffold), as judged by the fact that the profile of the combinations was essentially the sum of the profile of the individual components. The bottom left panel shows that both A-2 and A-4 assembled on the D-1-3 scaffold as shown by the shift in size of the mixed samples. The combined peak for D-1-3, A-2, and A-4 elutes at 13.2026 mL versus an elution of 13.3159 mL for the D-1-3 and A-4 peak. The bottom right similarly shows A-2 and 4-A assembled on the D-1-3 scaffold. The combined D-1-3, A-2, 4-A peak elutes at 13.2179 mL versus 13.3999 mL for the combination of D-1-3 with 4-A. The concentration of all proteins was kept constant in each experiment.

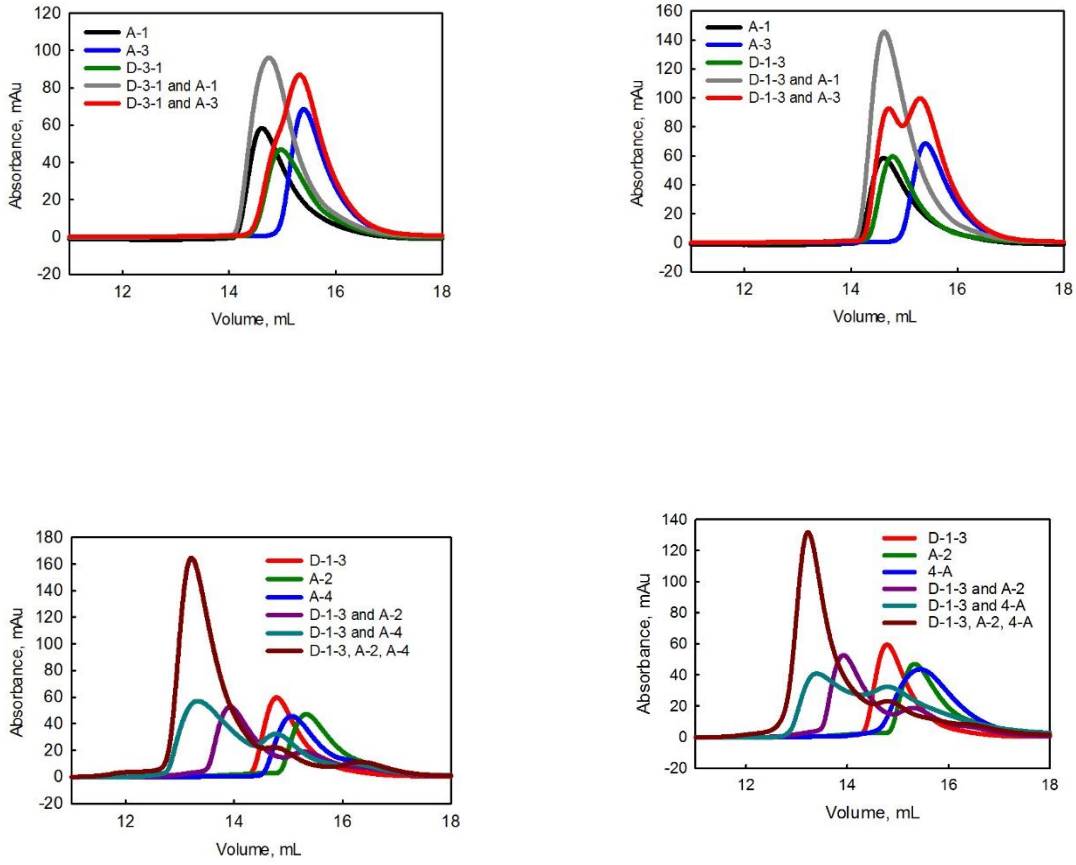


Table S1. Table of yields. Each of these constructs was produced and purified one or two times; typical yields are shown below.

Clone name	Protein yield, mg/L
D-5	25
D-6	17
D-17	16
D-18	4
D-E34m	23
D-R34m	1.1
A-5	5
A-6	12
A-17	7
A-18	9
A-E34m	13
A-R34m	1.3

Table S2. Table of zipper charge

Zipper	Number Arg/Lys	Number Glu/Asp	Net Charge
SYNZIP 1	9	10	-1
SYNZIP 2	11	10	+1
SYNZIP 3	9	8	+1
SYNZIP 4	11	10	+1
SYNZIP 5	8	10	-2
SYNZIP 6	11	10	+1
SYNZIP 17	14	8	+6
SYNZIP 18	6	11	-5
E34	7	8	-1
R34	10	5	+5

Table S3. Binding affinities as determined by SPR

In solution	Immoblized			
	D-1, nM	D-2, nM	D-3, nM	D-4, nM
A-5	--	--	--	--
A-6	--	--	--	--
A-17	--	10	1.3	--
A-18	--	--	--	--
A-R34m	--	--	--	--
A-E34m	--	--	--	--

Figure S5. SPR data corresponding to Table S3. The top curve shows A-17 binding to D-2 while the bottom shows A-17 binding to D-3. The SYNZIP 17 is positively charges (see table S2), and we observed the non-specific binding shown below.