Supporting Information

HER2 Targeting Peptides Screening and Applications in Tumor Imaging and Drug Delivery

Lingling Geng[†,1,3], Zihua Wang[†,1,2], Xiangqian Jia[1,4], Qiuju Han[1,4], Zhichu Xiang[1,2], Dan Li[1,2], Xiaoliang Yang[1,2], Di Zhang[1], Xiangli Bu[1,2], Weizhi Wang[*,1,2], Zhiyuan Hu[*,1,2], Qiaojun Fang[*,1,2]

[1]CAS Key Laboratory for Biomedical Effects of Nanomaterials & Nanosafety, National Center for Nanoscience and Technology, Beijing 100190, China

[2]CAS Center for Excellence in Nanoscience, National Center for Nanoscience and Technology, Beijing 100190, China

[3]National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China.

[4]Pharmacy College, Liaoning Medical University, Jinzhou, Liaoning 121001, China

[†]These authors contributed equally to this work.

*Correspondence authors:

Qiaojun Fang, National Center for Nanoscience and Technology, Beijing 100190, China.

Phone: +86-10-82545562; Fax: +86-10-82545643; E-mail: fangqj@nanoctr.cn.

Zhiyuan Hu, National Center for Nanoscience and Technology, Beijing 100190, China.

Phone: +86-10-82545643; Fax: +86-10-82545643; E-mail: huzy@nanoctr.cn.

Weizhi Wang, National Center for Nanoscience and Technology, Beijing 100190, China.

Phone: +86-10-82545643; Fax: +86-10-82545643; E-mail: wangwz@nanoctr.cn.

Contents:

Figure S11. Confocal fluorescence imaging analysis of peptides P51, P25 (labeled with FITC, green) binding toward HER2 in medium (MCF-7 and MDA-MB-231) and low (293A) expression of HER2 cell lines. Both peptides show weak signals in MCF-7 and MDA-MB-231 cells and no signals in 293A cells. Hoechst 33342-stained cell nuclei are in blue.

Figure S12. The MALDI-TOF-MS spectra of P25 and P25-Cy5.5 (a), P51 and P51-Cy5.5 (b). The molecular weights are consistent with the theoretical values.

Figure S13. *In vivo* and *ex vivo* fluorescence imaging of P51-Cy5.5 and P25-Cy5.5 to MDA-MB-468 cells (HER2 negative) xenografted breast tumors.

Figure S14. Transmission electron microscope (TEM) images of P51-LS-DOX (a), P25-LS-DOX (b) and LS-DOX (c).

Figure S15. Stability of peptides in plasma.

Section III. Supplementary Table Legends:

Table S1. Binding free energies and individual energy terms of HER1/HER1 dimer complex calculated by MM/GBSA (kcal/mol)

Table S2. Binding free energies and individual energy terms of complexes for HER2 and WP1 or its mutants calculated by MM/GBSA (kcal/mol)

Section I. Supplementary Materials and methods

MD simulation

The whole system was solvated in a truncated octahedron box filled with 8 Å TIP3P water molecules, with a minimum solute-wall distance of 12 Å [1], and the missing hydrogen atoms of each model were added using the *tleap* program. Then, counter-ions Na^+ were placed on the grids with the largest negative coulombic potentials around the protein.

The *sander* program was used before the MD simulations to minimize the structure via three steps: firstly, the whole protein was fixed and the water molecules and counter-ions were minimized (5000 cycles of steepest descent and 2500 cycles of conjugate gradient minimizations); secondly, the backbone atoms of the protein were fixed and the side chains of the protein were minimized (5000 cycles of steepest descent and 2500 cycles of conjugate gradient minimizations); thirdly, the whole system was minimized without any constrain (10000 cycles of steepest descent and 5000 cycles of conjugate gradient minimizations).The SHAKE procedure was applied, and the time step was set to 2.0 fs [2]. Particle Mesh Ewald (PME) was employed to deal with the long-range electrostatic interactions in the MD simulations [3]. Followed by minimization, the entire systems were gradually heated in the canonical ensemble from 0 to 310 K via seven steps in the NVT. Finally, 5 ns MD simulations were carried out under the constant temperature of 310 K. During the sampling process, the coordinates were saved every 0.2 ps, and the conformations generated from the simulations were used for further binding free energy calculations and decomposition analysis.

Binding free energy calculations

The binding free energy of each system was calculated using MM/GBSA technique

according to the following equation [4].

$$\Delta G_{bind} = G_{complex} - G_{protein} - G_{ligand}$$

$$= \Delta E_{MM} + \Delta G_{GB} + \Delta G_{SA} - T\Delta S \quad (1)$$

$$= \Delta E_{vdW} + \Delta E_{ele} + \Delta G_{GB} + \Delta G_{SA} - T\Delta S$$

where ΔE_{MM} is the interaction energy between protein and ligand in gas-phase, including the parts: the van der Waals energies (ΔE_{vdW}) and the electrostatic (ΔE_{ele}); ΔG_{GB} and ΔG_{SA} are the relative polar and nonpolar contributions to desolvation free energy, respectively, and $-T\Delta S$ represents the conformational entropic contribution at temperature T. In this study, the polar solvation free energy was calculated by the generalized born (GB) model [5]. In the GB calculations, the solvent and the solute dielectric constants were set to 80 and 4, respectively. The nonpolar solvation term was estimated based on the solvent accessible surface area (SASA) model by the LCPO method with a solvent-probe radius of 1.4 Å: ΔG_{SA} =0.0072×ΔSASA [6]**.** The binding free energy of each system was calculated based on 500 snapshots from 3 to 5 ns MD simulation trajectories of each complex by using the *mm_pbsa* program in AMBER12 [7]. Due to the low prediction accuracy and high computational demand of the conformational entropy (translation, rotation and vibration) upon the ligand binding ($-T\Delta S$), the entropic contribution was ignored [8, 9].

Free energy decomposition analysis

To have a clear understand of functional residues for interacting in all the modelling complexes, the MM/GBSA free energy decomposition process was applied by the *mm_pbsa* program in AMBER12 [10]. The binding free energy decomposition was to decompose the energy contribution of each residue from the association of the receptor with the ligand into four parts: van der Waals contribution (ΔE_{vdW}), electrostatic contribution (ΔE_{ele}), polar part, and nonpolar of

solvation ($\Delta G_{GB}+\Delta G_{SA}$) contribution. The polar contribution of desolvation free energy (ΔG_{GB}) was calculated using the generalized Born (GB) approximation model developed by Onufriev et al. [11], and the nonpolar solvation contribution (ΔG_{SA}) part was computed based on the SASA determined with the ICOSA method [12] . All energy components were calculated using 500 snapshots extracted from the MD trajectory from 3 to 5 ns.

Stability of peptides in plasma

Peripheral blood (5ml) from healthy volunteers was collected in tubes. The blood sample was centrifuged at 2100 g for 15 mins at 4°C for isolation of plasma. After centrifugation, the supernatant (plasma) was collected and the cell containing fraction (mainly red blood cells) was discarded. 5 μL peptide solution (1 mg/mL), 25 μL phosphate buffer pH 7.4 and 15 μL plasma solution were mixed in the tube. The mixture was gently shaken at 37°C incubator and 5μL aliquots were taken at time points of 0, 10, 15, 30, 60 and 120 mins. 100% ACN/0.1% TFA was added to each aliquot and centrifuged at 4000 rpm for 5min. The supernatants were analyzed using a monolithic C18 analytical column. Peptide concentrations in the serum were calculated by the integration of the peak areas of the chromatogram.

The detail of the OBOC screening process

During the OBOC screening, we employed three steps to eliminate the false positives: (1) In order to ensure an unbiased display for free molecular geometries, the interactions between the affinity peptide and HER2 protein is mediated by biotin-streptavidin conjugation, as shown in Figure S8a. HER2 was biotinylated with an average of two biotins per molecule. The biotinylated HER2 was incubated with the OBOC library and then the 1 μM streptavidin-coated magnetic beads were loaded to search the biotinylated peptide beads. It was expected that the magnetic

beads would associate with the peptide beads in an affinity-dependent manner. The ligands with the high affinity for the target protein could then be isolated using magnetic separation approaches in a continuous-flow microfluidic process (Figure S8b). (2) During the microfluidic screening. The positive magnetic based isolation was carried out in a three-cycle manner to eliminate the negative beads which could ensure the true positive rate to a certain extent. (3) The microfluidic chip system was fabricated on a silicon substrate. Each microwell was a cube shape which is a suitable size to trap individual peptide beads in a one-well-one-bead manner. After interaction, the mixture of the peptide library beads, biotinylated HER2 and the magnetic beads were introduced into the microfluidic chip. A magnetic field was applied upstream in the channel. Target-binding, or positive, beads were coated by the magnetic beads and trapped while the negative beads would flow through. A sheath flow configuration was used to sorting the positive beads and negative beads into different directions. The silicon chip was inserted into a modified MALDI target for *in situ* single peptide sequencing

After all, using the above the method, we discovered 72 positive target. Multiple sequence alignment and molecular simulation were introduced to find the ‘positive conserved sequence’. The affinities of the identified positive sequences have been confirmed by SPRi method and sequences of 72 peptides were found to cluster into three motifs.

Section II. Supplementary Figure Legends:

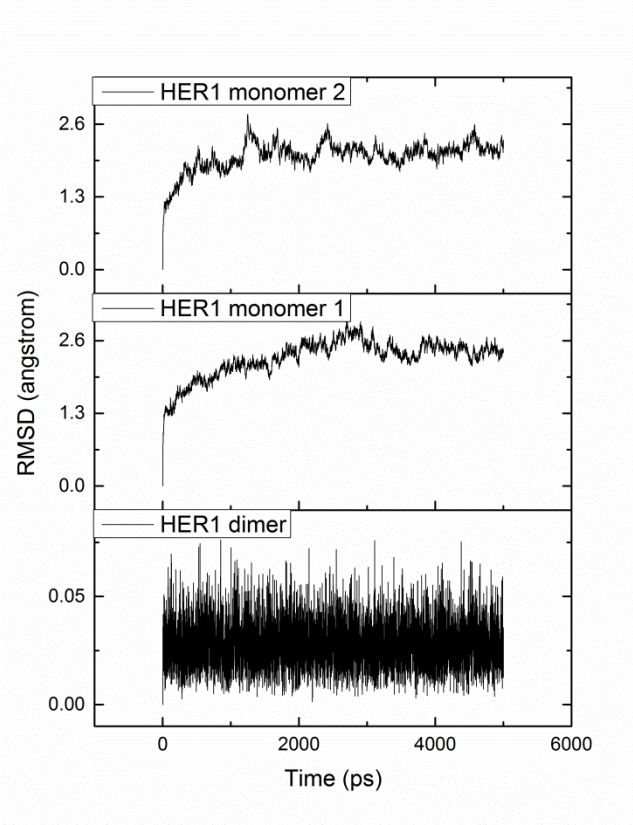

Figure S1. Backbone atoms root mean square deviation (RMSD) as a function of time for the initial and successive structures for HER1 dimer.

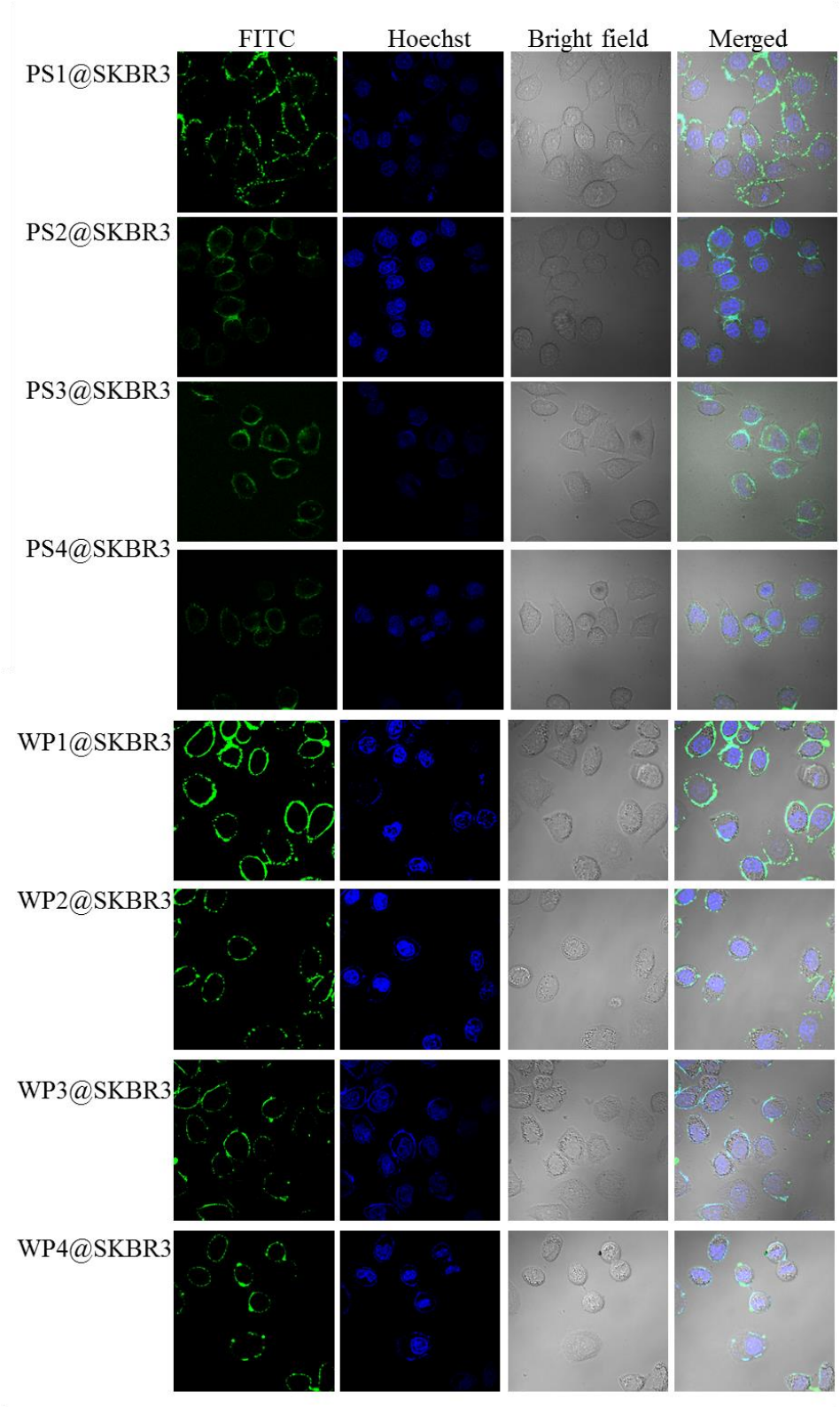

Figure S2. Confocal fluorescence imaging analysis of peptides PS1, PS2, PS3, PS4, WP1, WP2, WP3 and WP4 (labeled with FITC, green) in HER2 positive SKBR3 cells. Hoechst 33342-stained cell nuclei are in blue.

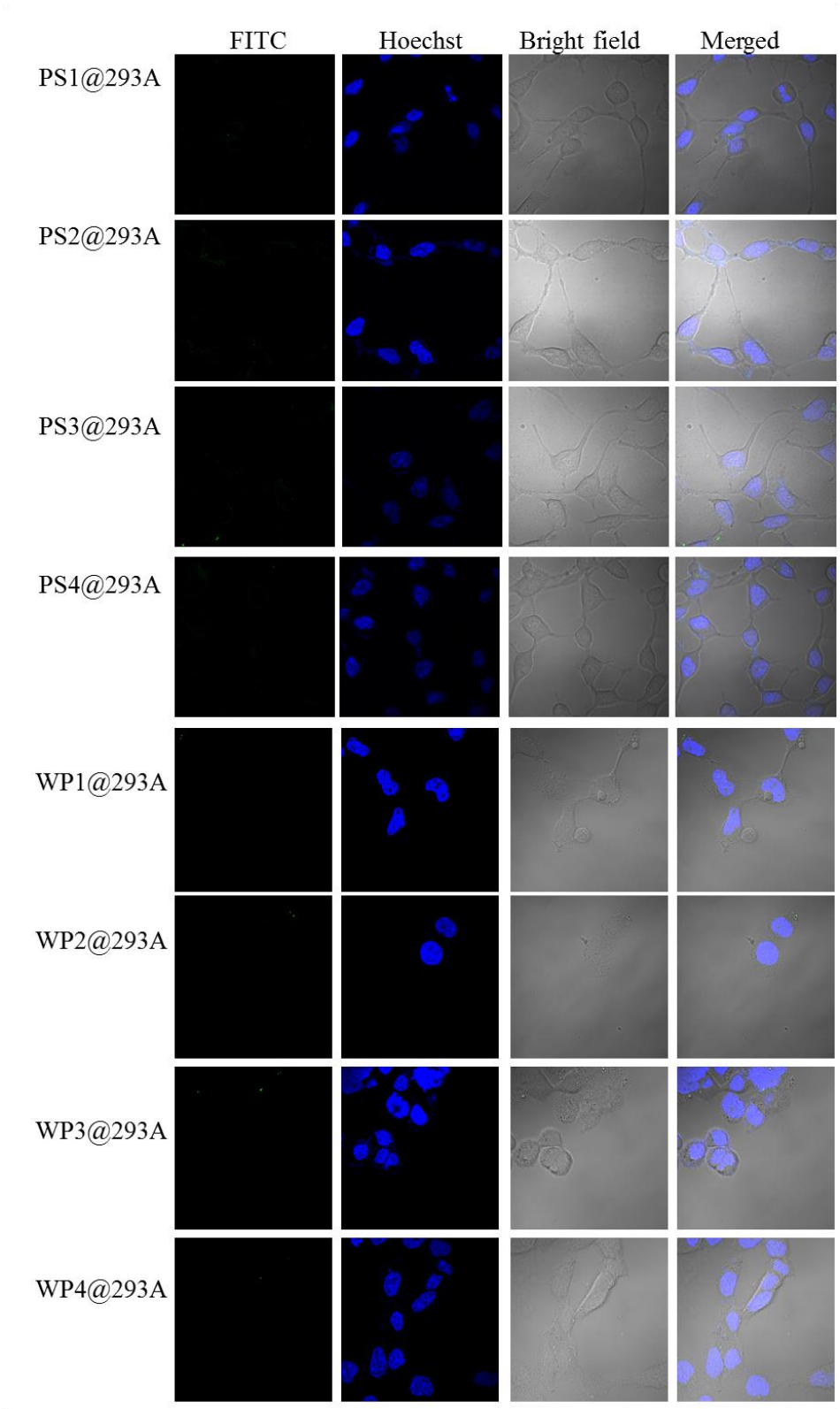

Figure S3. Confocal fluorescence imaging analysis of peptides PS1, PS2, PS3, PS4, WP1, WP2, WP3 and WP4 (labeled with FITC, green) in HER2 negative 293A cells. Hoechst 33342-stained cell nuclei are in blue.

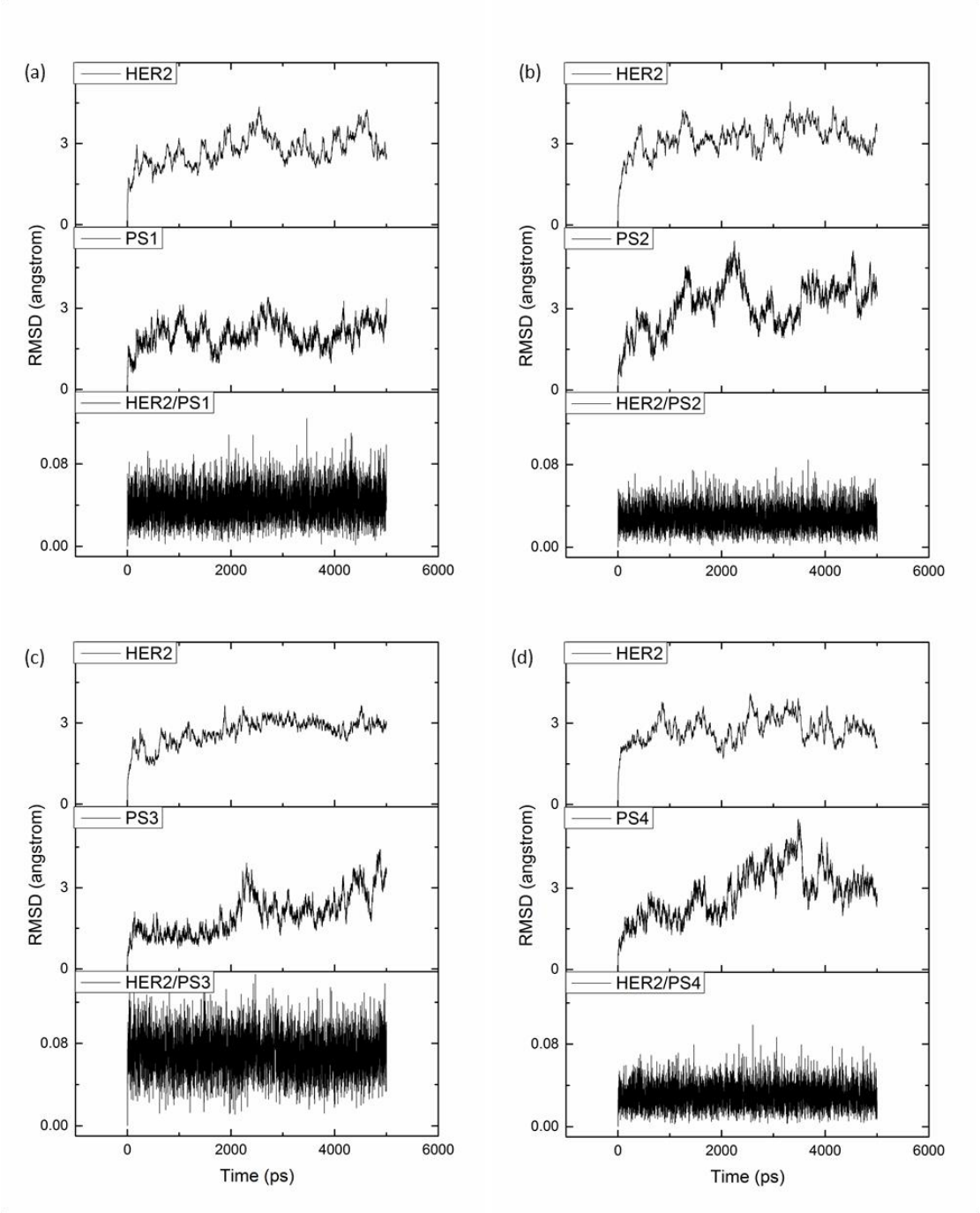

Figure S4. Backbone RMSD values as a function of time for the initial and successive structures of the receptor, ligand and the whole structure in the four complexes: HER2/PS1, HER2/PS2, HER2/PS3 and HER2/PS4.

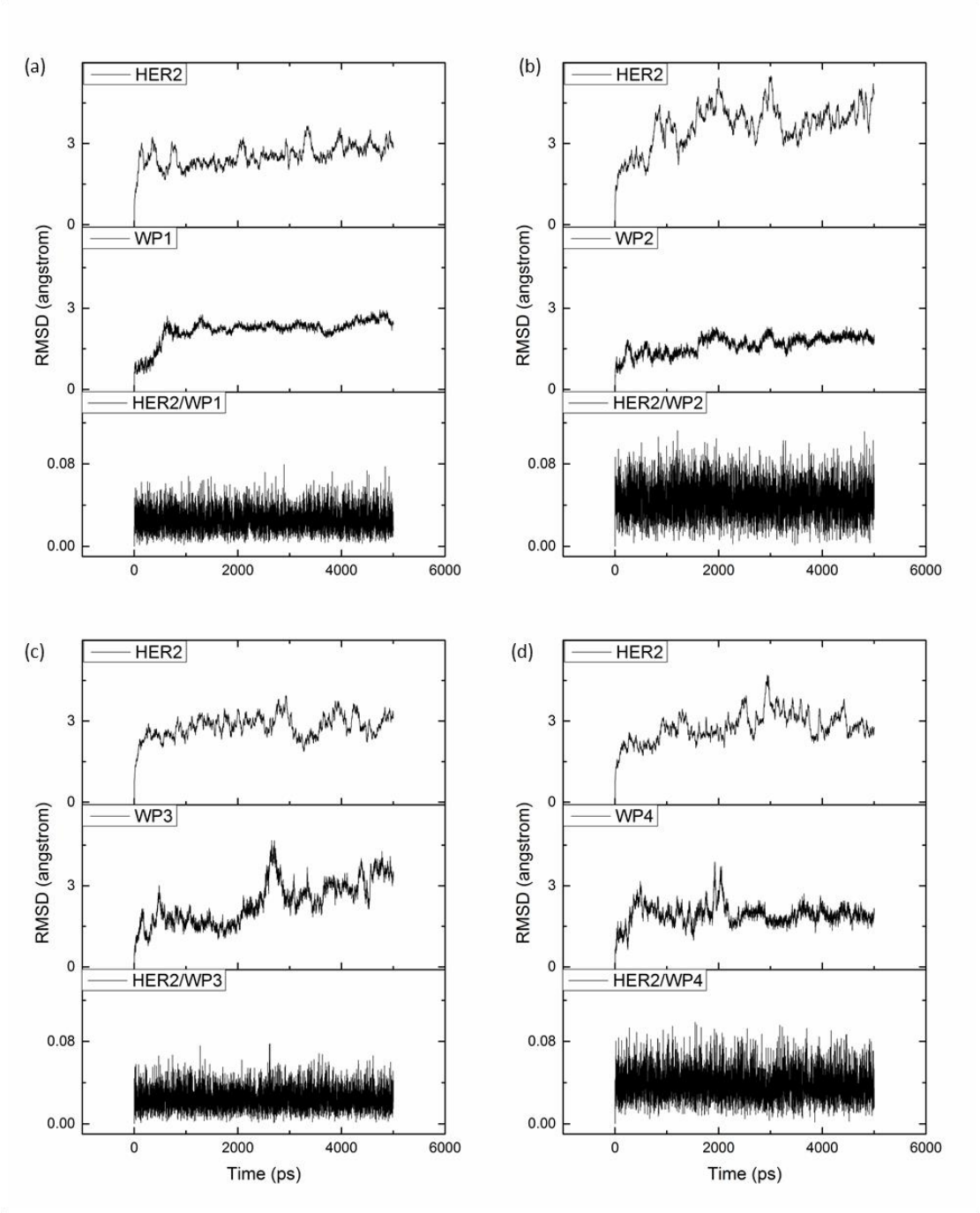

Figure S5. Backbone RMSD values as a function of time for the initial and successive structures of the receptor, ligand and the whole structure in the four complexes: HER2/WP1, HER2/WP2, HER2/WP3 and HER2/WP4.

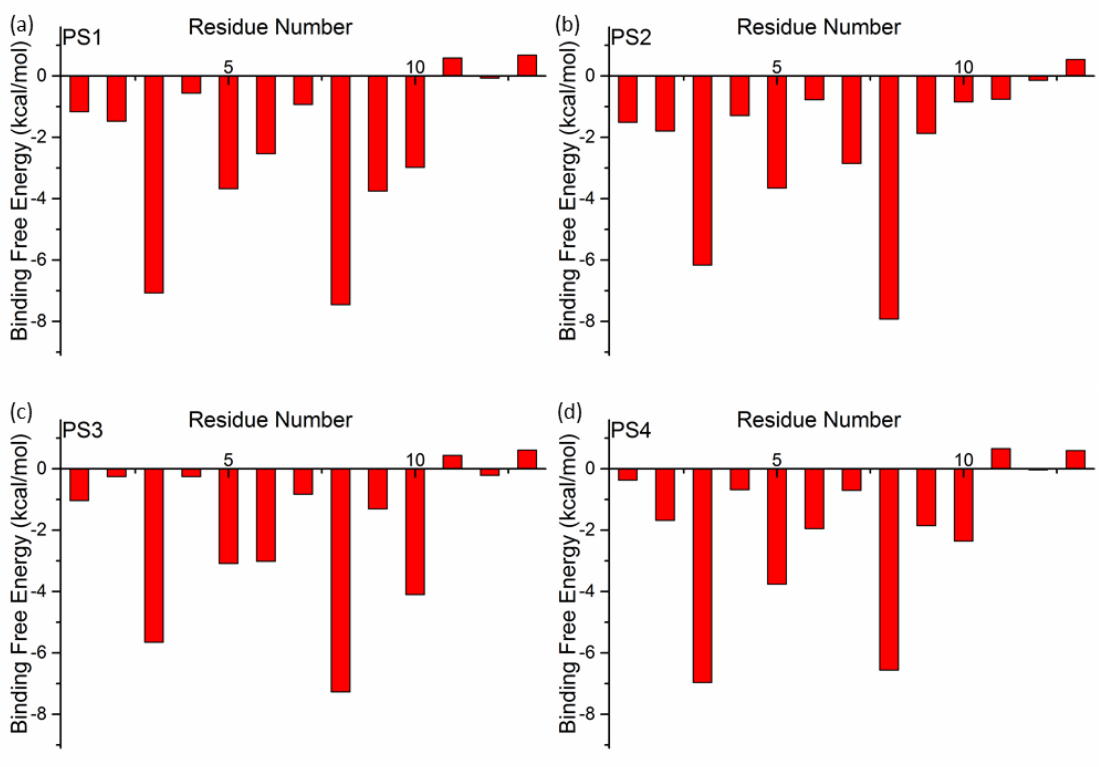

Figure S6. Comparison of the binding free energies of key residues in ligands for the four complexes: HER2/PS1, HER2/PS2, HER2/PS3 and HER2/PS4.

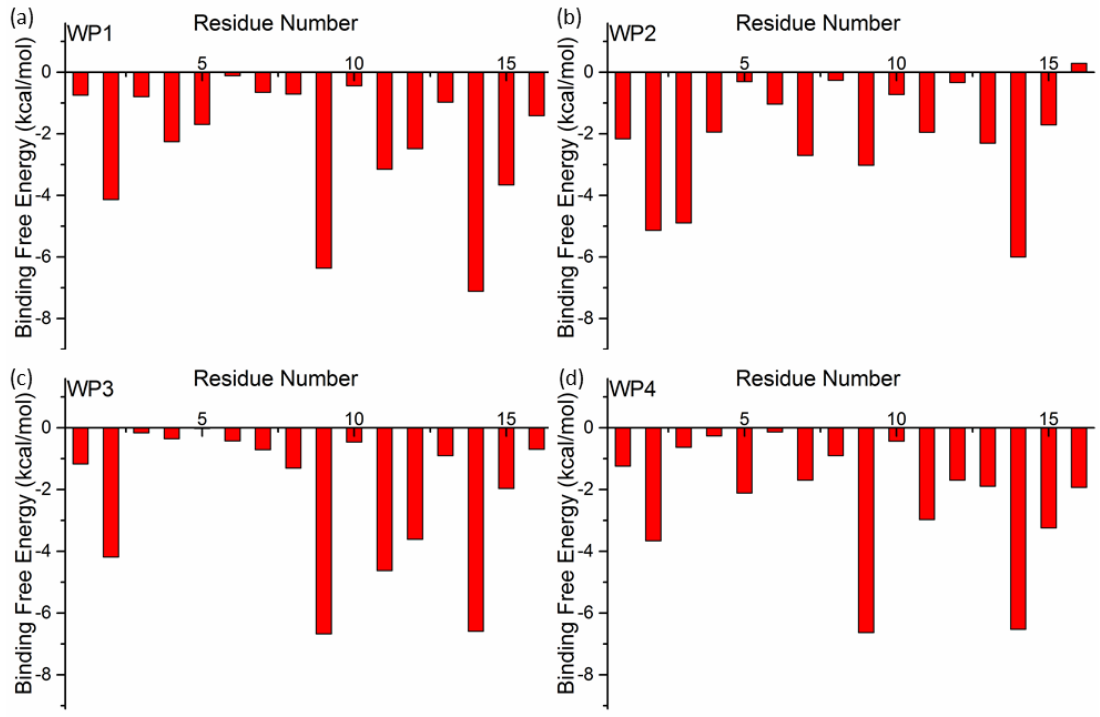

Figure S7. Comparison of the binding free energies of key residues in ligands for the four complexes: HER2/WP1, HER2/WP2, HER2/WP3 and HER2/WP4.

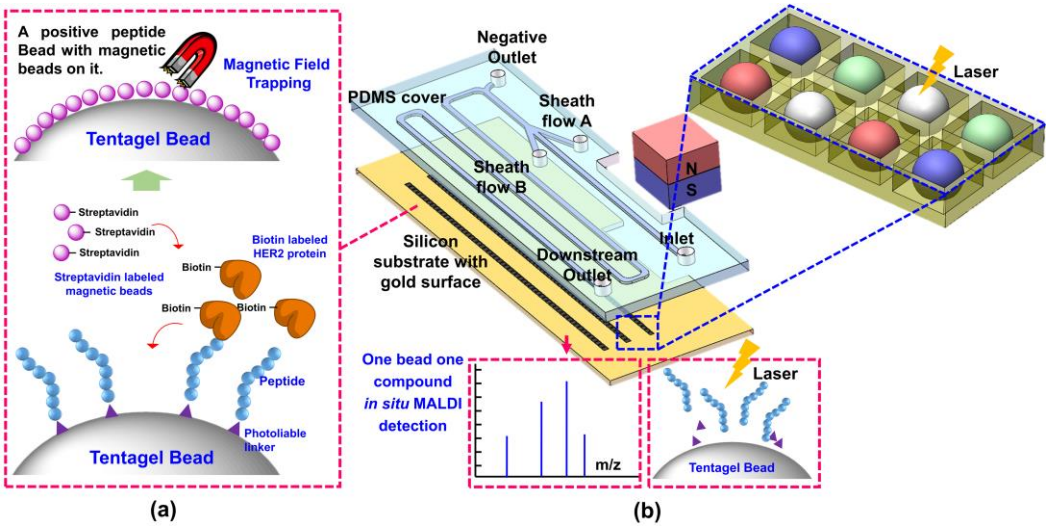

Figure S8. Principle of the screening process. (a) The positive peptide beads will be surrounded by the magnetic beads through the interaction bridge of peptide-HER2-biotin-streptavidin while native beads will remain naked. (b) The overview of chip system for beads trapping, sorting, and high-throughput *in situ* single bead sequencing.

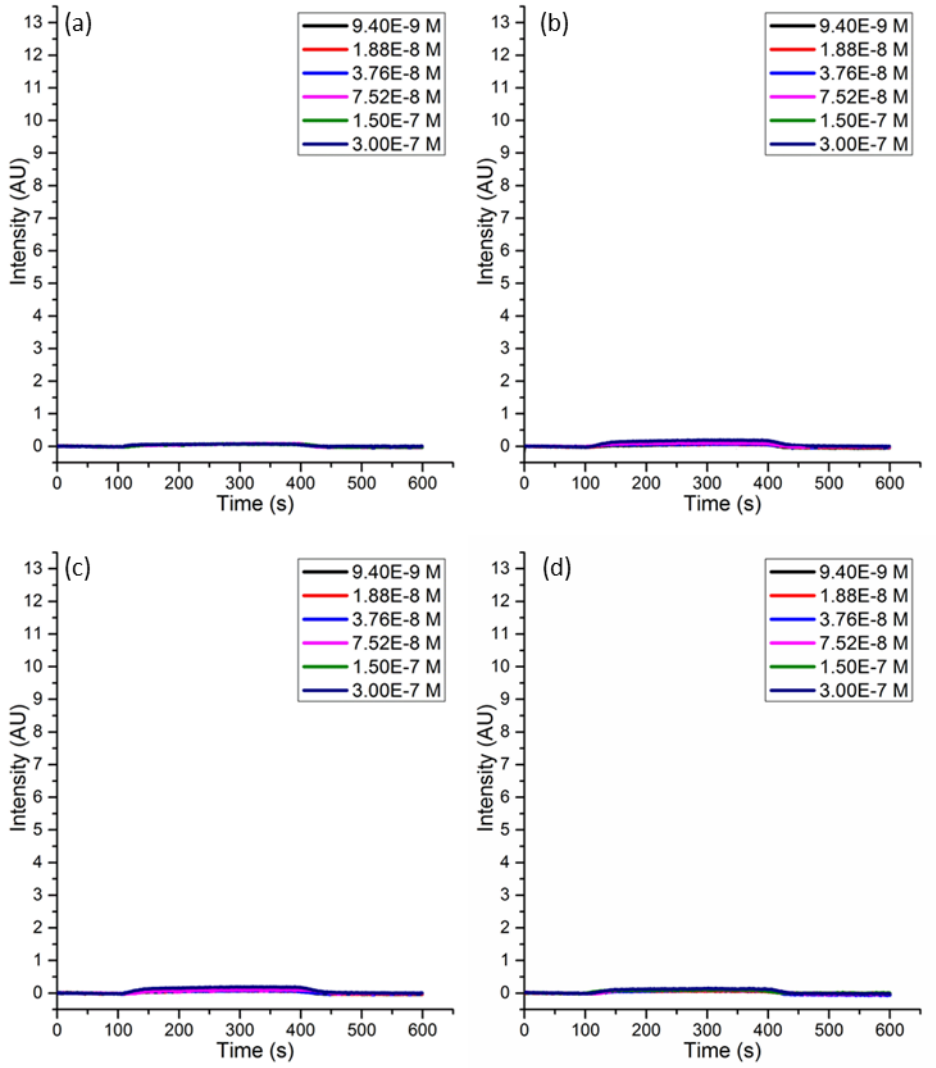

Figure S9. SPRi detection of the binding affinity of peptides P51 (a), P25 (b), P47 (c) and P40 (d) toward HSA protein.

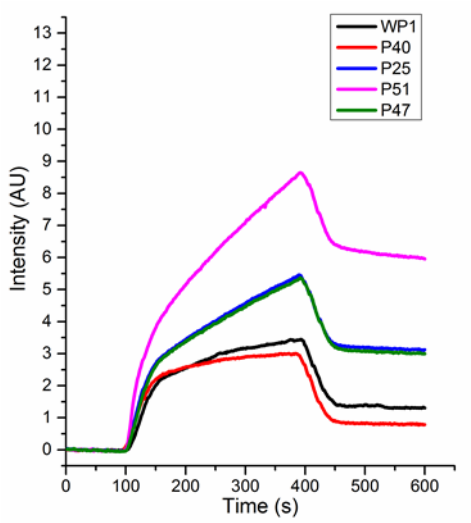

Figure S10. SPRi detection of the binding signals of peptides P51, P25, P47, P40 and WP1 toward HER2 protein at concentration of 9.1×10^{-8} M/L. All screened peptides except P40 bind stronger to HER2 than WP1.

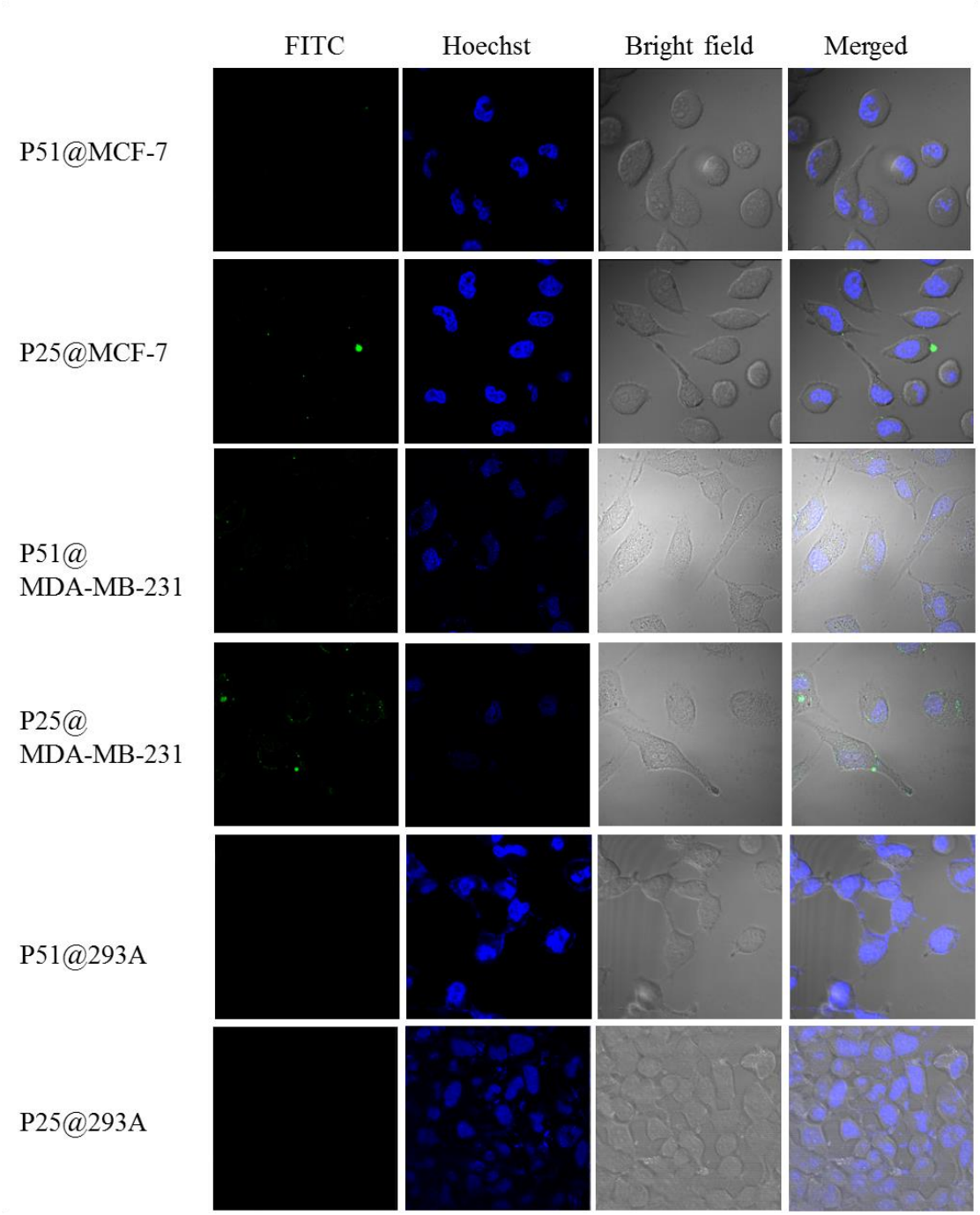

Figure S11. Confocal fluorescence imaging analysis of peptides P51, P25 (labeled with FITC, green) binding toward HER2 in medium (MCF-7 and MDA-MB-231) and low (293A) expression of HER2 cell lines. Both peptides show weak signals in MCF-7 and MDA-MB-231 cells and no signals in 293A cells. Hoechst 33342-stained cell nuclei are in blue.

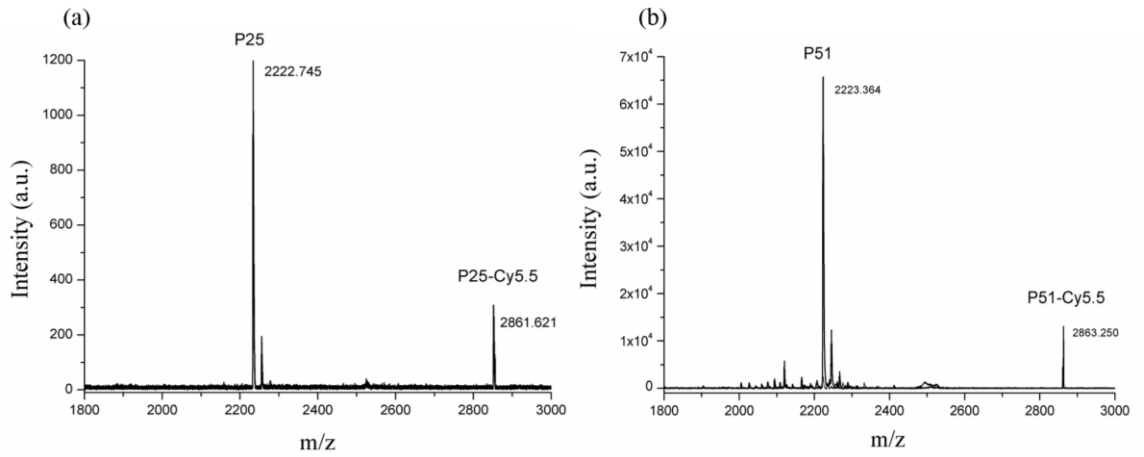

Figure S12. The MALDI-TOF-MS spectra of P25 and P25-Cy5.5 (a), P51 and P51-Cy5.5 (b). The molecular weights are consistent with the theoretical values.

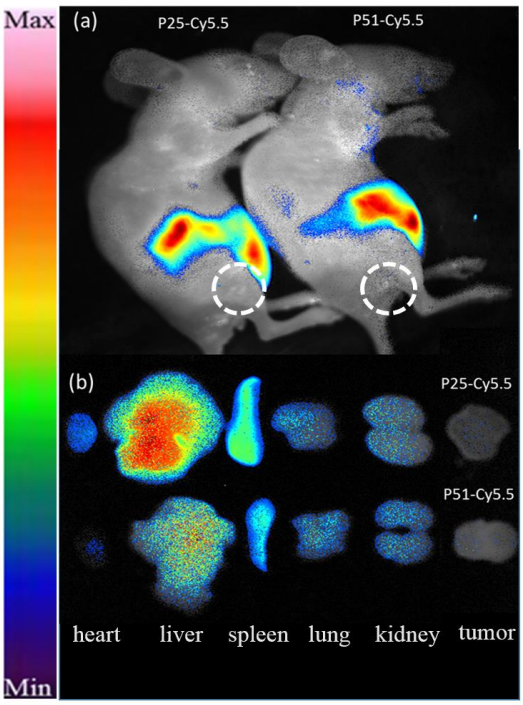

Figure S13. *In vivo* and *ex vivo* fluorescence imaging of P51-Cy5.5 and P25-Cy5.5 to MDA-MB-468 cells (HER2 negative) xenografted breast tumors.

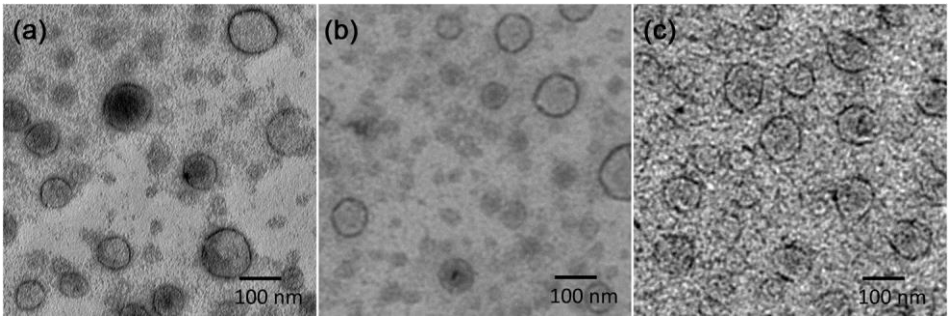

Figure S14. Transmission electron microscope (TEM) images of P51-LS-DOX (a), P25-LS-DOX (b) and LS-DOX (c).

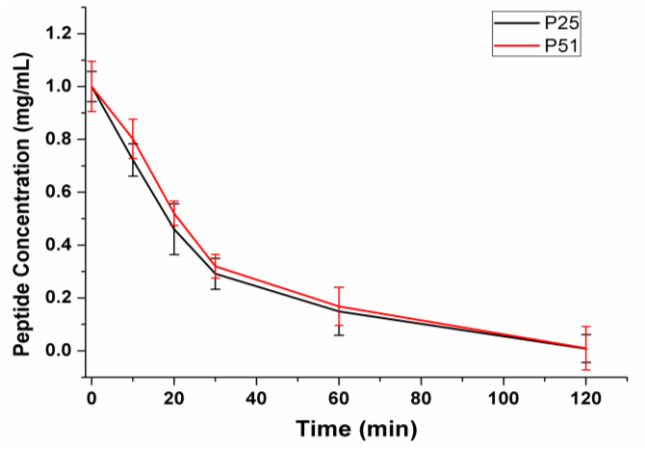

Figure S15. Stability of peptides in plasma.

Section III. Supplementary Table Legends:

Table S1. Binding free energies and individual energy terms of HER1 and HER1 dimer complex calculated by MM/GBSA (kcal/mol)

	ΔE_{vdw}	ΔE_{ele}	ΔG_{GB}	ΔG_{SA}	ΔG_{tot}
HER1/HER1	-220.71± 7.97	-91.17±9.17	122.73±8.29	-33.40±0.83	-222.56±7.41

ΔE_{vdw}, van der Waals contribution; ΔE_{ele}, electrostatic contribution; ΔG_{GB}, the polar contribution of desolvation; ΔG_{SA}, nonpolar contribution of desolvation; ΔG_{tot}, the total binding free energy without conformational entropy.

Table S2. Binding free energies and individual energy terms of complexes for HER2 and WP1 or its mutants calculated by MM/GBSA (kcal/mol)

	Peptide sequences	ΔE_{vdw}	ΔE_{ele}	ΔG_{GB}	ΔG_{SA}	ΔG_{tot}
HER2/WP1	DTCPPLMLYNPTTYQM	-71.11±5.14	1.57±3.68	6.93±2.91	-11.21±0.50	-73.83±5.10
HER2/WP1-D1K	KTCPPLMLYNPTTYQM	-74.23±8.37	-71.16±13.34	75.58±12.30	-11.49±1.10	-81.29±9.15
HER2/WP1-T2K	DKCPPLMLYNPTTYQM	-83.87±6.80	-33.86±6.30	41.81±5.97	-13.92±0.81	-89.85±6.88
HER2/WP1-T2M	DMCPPLMLYNPTTYQM	-71.67±5.43	8.77±2.98	0.12±2.75	-11.34±0.60	-74.11±5.34
HER2/WP1-T2Y	DYCPPLMLYNPTTYQM	-76.65±5.94	1.29±2.82	7.85±2.29	-11.78±0.63	-79.29±5.82
HER2/WP1-C3F	DTFPPLMLYNPTTYQM	-91.46±5.44	7.34±6.93	2.17±6.58	-13.14±0.51	-95.09±5.38
HER2/WP1-C3I	DTIPPLMLYNPTTYQM	-80.43±6.12	-0.81±4.32	7.75±3.98	-12.12±0.60	-85.61±6.24
HER2/WP1-P4W	DTCWPLMLYNPTTYQM	-80.52±7.41	13.89±5.02	-3.38±5.04	-13.17±0.73	-83.18±7.39
HER2/WP1-P4Y	DTCYPLMLYNPTTYQM	-100.66±5.53	9.00±5.85	3.44±5.28	-14.42±0.57	-102.64±5.43
HER2/WP1-P5W	DTCPWLMLYNPTTYQM	-98.92±6.24	8.60±3.52	1.64±2.87	-14.51±0.50	-103.18±5.88
HER2/WP1-P5Y	DTCPYLMLYNPTTYQM	-92.87±4.82	3.69±3.04	6.39±2.80	-14.00±0.48	-96.79±4.65
HER2/WP1-M7G	DTCPPLGLYNPTTYQM	-93.85±9.31	1.86±5.89	6.58±4.57	-14.18±0.82	-99.60±8.65
HER2/WP1-L8E	DTCPPLMEYNPTTYQM	-79.15±4.67	18.24±6.94	-8.70±5.90	-11.74±0.54	-81.35±4.08
HER2/WP1-L8Y	DTCPPLMYYNPTTYQM	-79.71±5.55	6.99±3.19	0.95±3.02	-12.33±0.58	-84.10±5.77
HER2/WP1-L8W	DTCPPLMWYNPTTYQM	-79.20±4.92	-3.29±4.24	12.29±3.71	-12.82±0.53	-83.01±4.62
HER2/WP1-Y9F	DTCPPLMLFNPTTYQM	-69.03±5.45	-3.52±5.07	10.80±4.24	-10.82±0.53	-72.58±5.30
HER2/WP1-Y9W	DTCPPLMLWNPTTYQM	-86.68±5.22	0.85±7.60	9.12±6.41	-14.05±0.68	-90.77±5.20
HER2/WP1-T12Q	DTCPPLMLYNPQTYQM	-81.64±4.35	1.19±5.01	11.99±3.98	-12.00±0.40	-80.46±3.96
HER2/WP1-T12N	DTCPPLMLYNPNTYQM	-76.49±5.34	10.03±6.34	0.34±5.49	-12.00±0.70	-78.13±4.89
HER2/WP1-T13E	DTCPPLMLYNPTEYQM	-76.98±6.70	32.23±6.72	-19.23±6.22	-12.08±1.18	-76.06±7.38
HER2/WP1-T13D	DTCPPLMLYNPTDYQM	-79.97±5.63	31.30±6.13	-21.47±5.35	-12.25±0.60	-82.39±5.75
HER2/WP1-Y14W	DTCPPLMLYNPTTWQM	-86.82±5.29	-2.98±4.87	11.99±4.24	-12.42±0.72	-90.23±5.35
HER2/WP1-Q15F	DTCPPLMLYNPTTYFM	-91.72±7.81	7.88±5.82	5.00±5.44	-13.38±0.80	-92.23±8.30
HER2/WP1-Q15R	DTCPPLMLYNPTTYRM	-87.79±5.19	-23.79±6.19	34.94±4.82	-13.56±0.42	-90.20±5.72
HER2/WP1-Q15W	DTCPPLMLYNPTTYWM	-69.20±6.09	2.67±6.17	7.13±5.27	-10.78±0.76	-70.18±5.98
HER2/WP1-M16K	DTCPPLMLYNPTTYQK	-84.26±5.55	-34.54±6.92	42.31±5.95	-13.99±0.67	-90.48±5.68
HER2/WP1-M16Q	DTCPPLMLYNPTTYQQ	-75.51±5.37	-10.47±4.44	19.07±4.01	-12.01±0.48	-78.92±5.14
HER2/WP1-M16Y	DTCPPLMLYNPTTYQY	-98.66±5.70	6.25±3.87	4.75±3.31	-14.90±0.67	-102.55±5.43

ΔE_{vdw}, van der Waals contribution; ΔE_{ele}, electrostatic contribution; ΔG_{GB}, the polar contribution of desolvation; ΔG_{SA}, nonpolar contribution of desolvation; ΔG_{tot}, the calculated total binding free energy.

References

1. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. J Chem Phys. 1983; 79: 926-35.
2. Ryckaert J-P, Ciccotti G, Berendsen HJC. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J Comput Phys. 1977; 23: 327-41.
3. Darden T, York D, Pedersen L. Particle mesh Ewald: An N [center-dot] log(N) method for Ewald sums in large systems. J Chem Phys. 1993; 98: 10089-92.
4. Kollman PA, Massova I, Reyes C, Kuhn B, Huo S, Chong L, et al. Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models. Accounts Chem Res. 2000; 33: 889-97.
5. Tsui V, Case DA. Theory and applications of the generalized Born solvation model in macromolecular simulations. Biopolymers. 2000; 56: 275-91.
6. Weiser J, Shenkin PS, Still WC. Approximate atomic surfaces from linear combinations of pairwise overlaps (LCPO). J Comput Chem. 1999; 20: 217-30.
7. Case DA, Cheatham TE, 3rd, Darden T, Gohlke H, Luo R, Merz KM, Jr., et al. The Amber biomolecular simulation programs. J Comput Chem. 2005; 26: 1668-88.
8. Sun H, Li Y, Shen M, Tian S, Xu L, Pan P, et al. Assessing the performance of MM/PBSA and MM/GBSA methods. 5. Improved docking performance using high solute dielectric constant MM/GBSA and MM/PBSA rescoring. Phys Chem Chem Phys. 2014; 16: 22035-45.
9. Hou T, Wang J, Li Y, Wang W. Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. J Chem Inf Model. 2011; 51: 69-82.
10. Gohlke H, Kiel C, Case DA. Insights into Protein–Protein Binding by Binding Free Energy Calculation and Free Energy Decomposition for the Ras–Raf and Ras–RalGDS Complexes. J Mol Biol. 2003; 330: 891-913.
11. Onufriev A, Bashford D, Case DA. Exploring protein native states and large-scale conformational changes with a modified generalized born model. Proteins. 2004; 55: 383-94.
12. Case DA, Darden TA, Cheatham, Simmerling CL, Wang J, Duke RE, et al. Amber 9; 2006.