SUPPORTING INFORMATION

DESIGNING FLUORESCENT PEPTIDE SENSORS WITH DUAL SPECIFICITY FOR THE DETECTION OF HIV-1 PROTEASE

Karla-Luise Herpoldt§, Arbel Artzy-Schnirman§, Andrew J. Christofferson†, Adam J. Makarucha†, Roberto de la Rica§#, Irene Yarovsky†*, Molly M. Stevens§*

§ Department of Materials, Department of Bioengineering, Institute of Biomedical Engineering, Imperial College London, Prince Consort Road, London, SW7 2AZ, UK

†Health Innovations Research Institute, RMIT University, GPO Box 2476, Victoria 3001, Australia

Correspondence should be addressed to (**Experimental**) Prof. Molly M. Stevens m.stevens@imperial.ac.uk (**Simulation**) Prof. Irene Yarovsky irene.yarovsky@rmit.edu.au

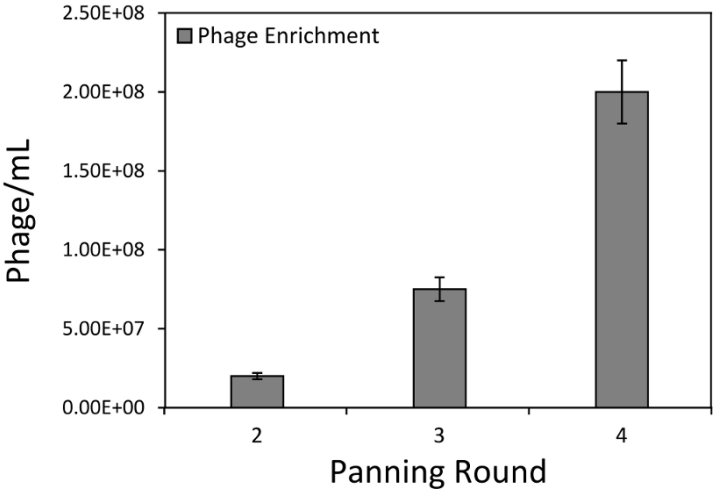

Figure S1. Phage enrichment data showing an increase of an order of magnitude. Error bars show 10% counting error.

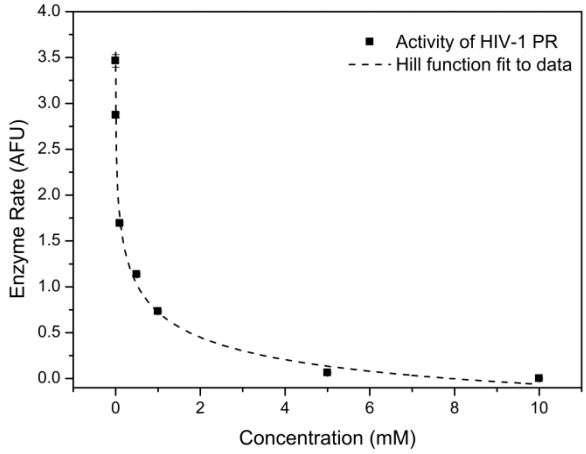

Figure S2. Inhibition data for WSRVGYW showing a K_i of 229 μM through a Hill fit to the data.

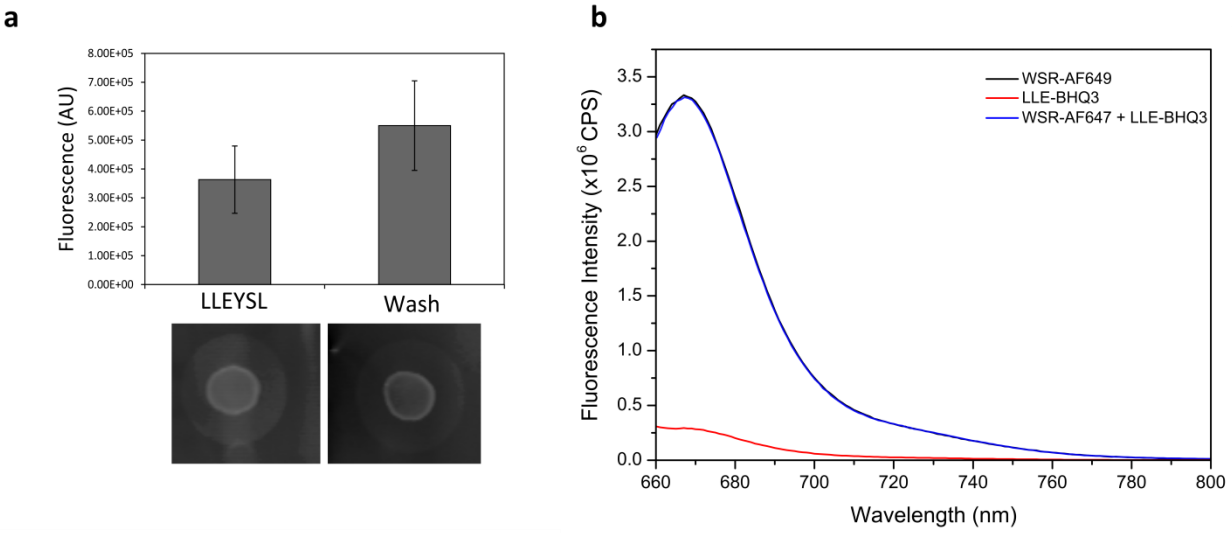

Figure S3. Surface FRET controls showing a) no decrease in signal compared to additional washing in PBST when the membrane is incubated with LLEYSL in the absence of quencher molecule, experiment is shown in duplicate. Images show representative dot blots indicating no loss in signal due to competition from non-dye labelled LLEYSL. b) No decrease in fluorescence intensity is observed when WSR-AF647 and LLE-BHQ-3 are incubated together in the absence of HIV-1 PR.

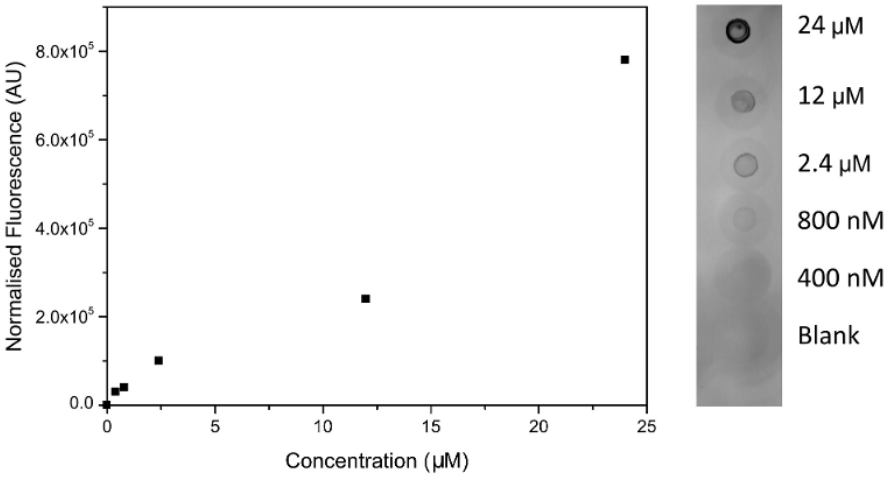

Figure S4. Fluorescent signal registered at different concentrations of immobilized HIV-1 PR incubated with WSR-AF647 at 210 nM.

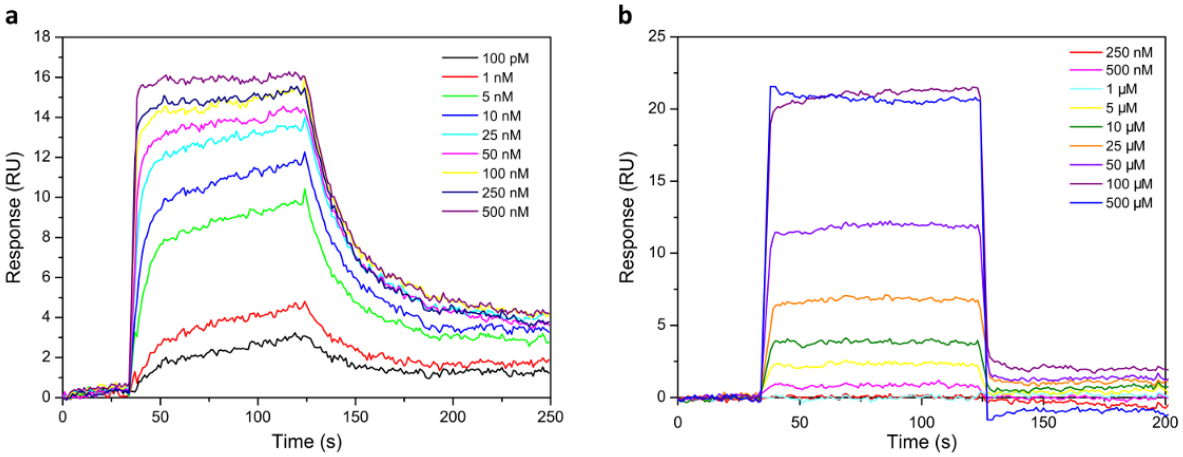

Figure S5. SPR showing binding affinity differences between LLEYSL (a) and WSRVGYW (b). The top hat function shown by WSRVGYW is due to the high k_{off} exhibited by this peptide which is faster than the sampling time of the Biacore 3000.

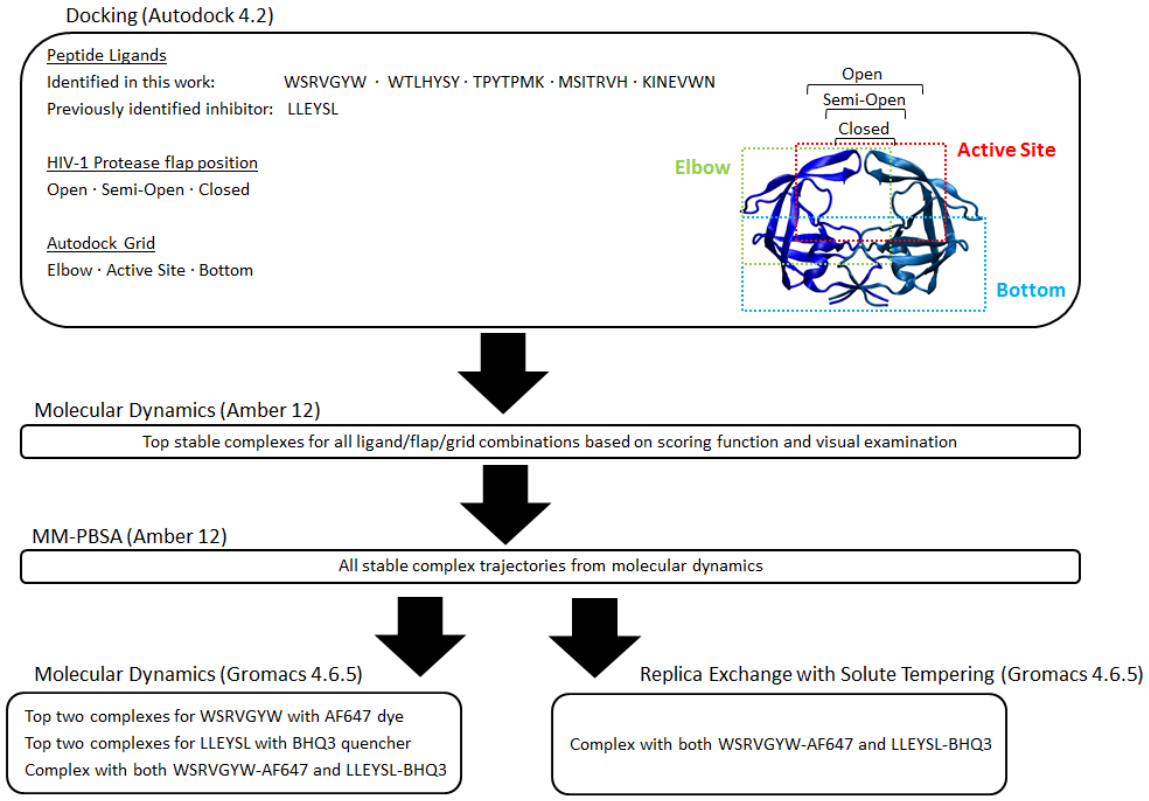

Figure S6. Flowchart schematic showing the sequence of docking and molecular dynamics simulations. Docking was carried out to three specified areas of the protease as shown.

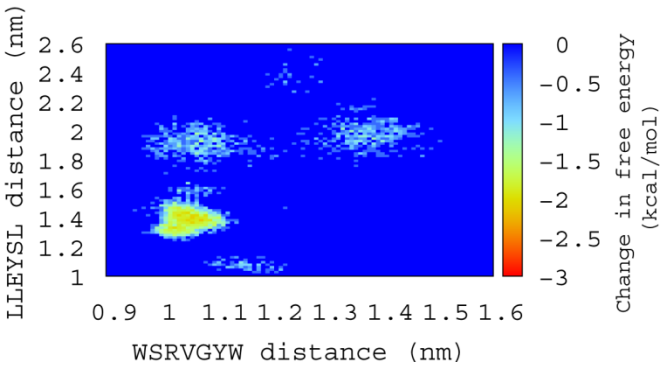

Figure S7. Change in free energy as a function of the radial distance of the centre of mass of each peptide backbone from the catalytic residues Asp25 and Asp25'.

REST Clustering Discussion

For the WSRVGYW peptide, the top cluster (W1) comprised 60% of all sampled structures, with the second most populated cluster (W2) comprising 27% and no other cluster above 5% of the ensemble. The structure is characterized by hydrogen bonding between the arginine side-chain of the peptide and the catalytic residue Asp25' of the HIV-1 PR. This feature is present in the two most populated clusters. The primary difference between the top two clusters was a shift in the N-terminal end of the peptide away from the active site and toward the flaps in cluster W2. Additionally, cluster W1 features hydrogen bonding between the peptide N-terminal Trp backbone and Gly48' backbone, the backbone of the dye and Lys45', and from the Tyr side-chain to the Glu21 side-chain. There are also persistent hydrophobic contacts between the peptide Val (residue 5, Figure 4b) and Ile50, Val82, Ile84, and Leu23, as well as the N-terminal Trp and Phe53' (Figure 4b).

For the LLEYSL peptide, the top cluster (L1) comprised 54% of all structures, with the second and third most populated clusters (L2 and L3) comprising 18% and 14%, respectively, and no other cluster above 5%. The primary difference between the top

two clusters was a shift in the peptide Tyr toward the active site. Cluster L1 features hydrogen bonding between the C-terminal and the Arg8' side-chain and between the Ser side-chain and Gly48 backbone. There are also persistent hydrophobic contacts between the C-terminal Leu, the second Leu from the N-terminus, and Ile50' and Val82' of the protease (Figure 4b).

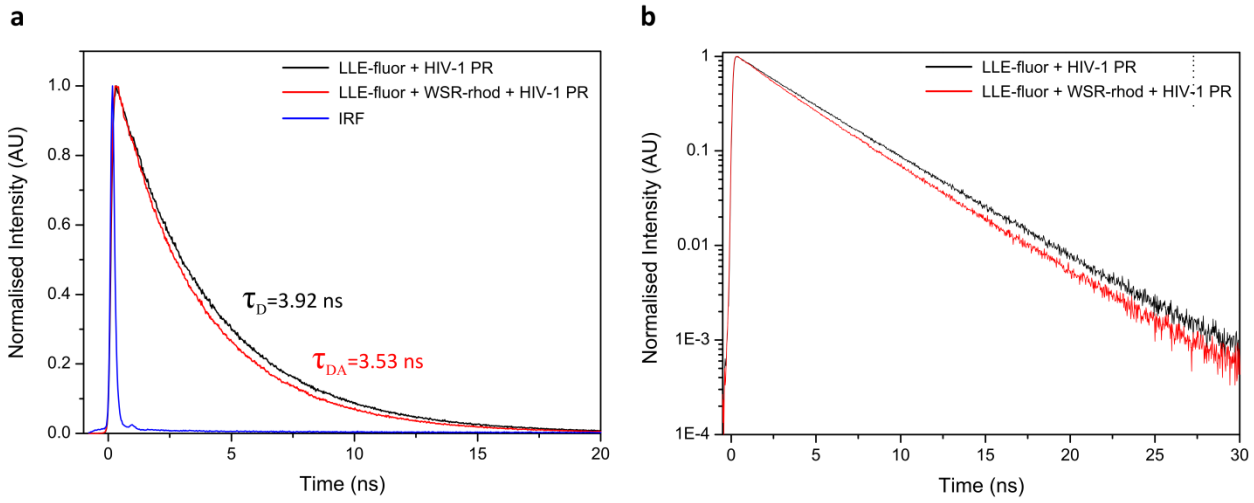

Figure S8. Lifetime fluorescence measurements of LLE-fluorescein interacting with HIV-1 PR in the presence and absence of WSR-rhodamine. a) Linear-linear scale showing exponential decay, b) data from (a) represented on a semi-log scale. IRF shows the instrument response function.

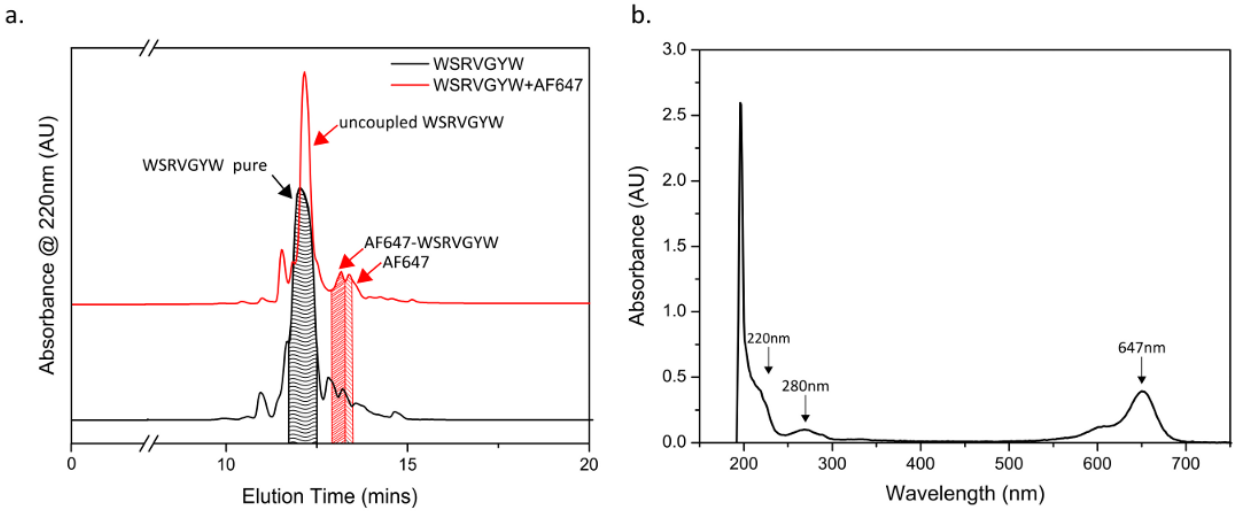

Figure S9. Characterization of fluorophore-labeled peptides. a) The peptide WSRVGYW was synthesized and purified by HPLC. The fraction corresponding to pure WSRVGYW (black trace) was isolated and its mass verified by MALDI. The fraction was then lyophilized to remove residual acetonitrile and TFA. The peptide was then conjugated to Alexa Fluor 647 and repurified by HPLC (red trace). Colored fractions were collected and the UV-Vis spectra analyzed (b). Fractions containing absorbance peaks corresponding to 220nm (peptide backbone), 280nm (aromatic side chain) and 647nm (dye) were assumed to contain fluorophore-labelled peptides.

MALDI-TOF was performed on all purified fractions that contained dye. A variety of matrices were used including sinapinic acid, alpha-cyano-4-hydroxycinnamic acid, nitroaniline, dihydroxy benzoic acid, dithranol, 2-[(2E)-3-(4-tert-butylphenyl)-2-methylprop-2-enylidene]malononitrile, picolinic acid, 2′,4′,6′-Trihydroxyacetophenone monohydrate, All trans retinol, aza thio thymine, 3-indoleacrylic acid and diamino benzophenone. MALDI-TOF was performed in both positive and negative mode. However no peaks corresponding to the expected molecular weight were identified. This has been reported before by Esteban et al. (2011) and is likely due to the highly charged nature of the dye molecule.

Supporting Methods

Surface Plasmon Resonance (SPR)
Surface plasmon resonance (SPR) experiments were carried out on a Biacore 3000 (GE Healthcare, Uppsala, Sweden) at 25°C using PBST, pH 7.4 (Tween at 0.005% (v/v)) as the running buffer. Dextran coated CM5 chips were purchased from GE Healthcare and 50 µg/mL of BSA and 50 µg/mL of HIV-1 PR were coated to flow cells 1 and 2 respectively. Coating was performed using NHS/EDC coupling at a flow rate of 10 µL/min and the surface was subsequently blocked with 1 M ethanolamine. LLEYSL and WSRVGYW were injected for 90 s over the surface of both flow cells at a flow rate of 40 µL/min in a concentration series. The injection was followed by a 150 s dissociation period. The surface of the CM5 chips was regenerated through a 30 s injection of 0.2 M glycine-HCl (pH 2.2) at 30 µL/min. Data were evaluated using Scrubber 2 (BioLogic Software, Australia) and OriginPro 8.5 and equilibrium dissociation constants were calculated from $K_d=k_{off}/k_{on}$.

Concentration Gradient Dot Blot
HIV-1 PR was diluted in PBST to create a concentration gradient. 5 µL of each solution was spotted onto a nitrocellulose membrane and allowed to dry. The membrane was then blocked in 2% (w/v) skim milk for 2 hours at room temperature before being washed in PBST (0.05% (v/v) Tween-20) to remove excess milk proteins. The fluorophore labelled peptide was then diluted 1:1000 in 1% (w/v) skim milk and the membrane incubated in this mixture for 1 hour. The membrane was washed well to remove non-specifically bound dye and then imaged on a Li-Cor Odyssey Infra-red scanner. The intensity of the spots was quantified using Li-Cor Image Studio.

Fluorescence Lifetime Measurements
The FRET assay was set-up as described in the main methods using a concentration of 4 µM of HIV-1 PR in a total volume of 100 µL. Measurements were made on a DeltaFlex Time-Correlated Single Photon Counting (TCSPC) system (Horiba) using a NanoLED with an excitation wavelength of 467 nm and 30,000 counts were recorded. Fluorescence lifetimes were extracted through an exponential fit to the normalized data in OriginPro 8.5.

FRET Distance Calculations
FRET efficiency can also be written in terms of the Förster radius, R_o – this is the characteristic length where energy transfer is 50% efficient. In this case efficiency is given by:

$$E = \frac{1}{1+\left(\frac{r}{R_0}\right)^6} \qquad \text{Equation S1}$$

In these calculations the R_o of fluorescein has been taken from the literature as being equal to 54 Å.[1] R_o can however be defined as:

$$R_0 = \frac{9000 Q_0 (\ln 10) \kappa^2}{128 \pi^5 n^4 N_A} \int f_D(\lambda) \varepsilon_A(\lambda) \lambda^4 d\lambda \qquad \text{Equation S2}$$

Where Q_o is the quantum yield of the donor in the absence of the acceptor, κ^2 describes the orientation factor between the two dipoles, which for the case of dye molecules free to rotate can be approximated as 2/3. n is the refractive index of the medium and N_A is Avogadro's number. The integral describes the spectral overlap between the fluorophores with respect to wavelength. f_D is the normalized donor emission spectrum, and ε_A is the acceptor molar extinction coefficient.

To find the distance between the donor and acceptor we can then equate equation 1 with equation S1 to give:

$$r = R_0 \sqrt[6]{\frac{\tau_{DA}}{\tau_D - \tau_{DA}}} \qquad \text{Equation S3}$$

MM-PBSA
An estimate of the free energy of binding was given by the Molecular Mechanics Poisson-Boltzmann Surface Area (MM-PBSA) method.[2] This analysis was performed on the final 10 ns of the simulation runs, with all frames (1000 total) used in the enthalpy calculation and every fifth frame (200 total) used in the entropy calculation, due to the greater computational cost of the latter. A value of 0.00542 kcal/(mol/Å²) was used for the surface tension and 0.15 M for the ionic strength, with an offset of 0.29 kcal/mol, and a dielectric constant of 4 and convergence root mean square gradient of 10^{-4} kcal•mol^{-1}•Å^{-1} for the entropy calculation.

Replica exchange with solute tempering
The top two binding orientations (based on MM-PBSA ranking) for the LLEYSL and WSRVGYW peptides were used as the starting configurations for simulations of the peptide-dye conjugates. Partial charges for BHQ-3 and AF647 were determined using Gaussian 09,[3] and missing parameters were taken from the general amber force field (GAFF)[4] using the antechamber function of the Amber MD code. Replica exchange with solute tempering (REST)[5,6] simulations were performed using the Gromacs 4.6.5 MD code[7], with 16 replicas and an effective temperature range of 300-600 K. The average probability of exchange was 0.4, with an

exchange attempt every 4 ps. Only the residues in the protease flaps (residues 46 to 56 and 46' to 56') and the peptides were perturbed. For the LLE-BHQ/WSR-AF657 combination, REST simulations were run for 200 ns. Additional simulations for each complex were run in Gromacs for 800 ns without REST as a control. For simulations using the Gromacs MD code, the v-rescale thermostat and Parrinello-Rahman barostat were implemented, with the LINCS[8] constraint on all bonds, and a 10 Å cutoff for electrostatics (PME) and non-bond interactions. All other parameters were identical to the Amber simulations.

SUPPORTING REFERENCES

(1) Lakowicz, J., editor. *Principles of Fluorescence Spectroscopy,* Springer US, Boston MA, 2006

(2) Miller, B.R.; McGee, T.D.; Swails, J. M.; Homeyer, N.; Gohlke, H.; Roitberg, A. E. MMPBSA.py: An Efficient Program for End-State Free Energy Calculations. *J. Chem. Theory Comput*. **2012**, 8, 3314-3321

(3) Frisch, M.J.; Trucks, G.W.; Schlegel, H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G.A. *et al.* Gaussian 09. Gaussian, Inc.: Wallingford, CT, USA, **2009**

(4) Wang, J.; Wolf, R.M.; Caldwell, J.W.; Kollman, P.A.; Case, D.A. Development And Testing Of A General Amber Force Field. *J. Comput. Chem*. **2004**, 25, 1157-1174

(5) Wright, L.B.; Walsh, T.R. Efficient Conformational Sampling Of Peptides Adsorbed Onto Inorganic Surfaces: Insights From A Quartz Binding Peptide. *Phys. Chem. Chem. Phys*. **2013**, 15, 4715-4726

(6) Terakawa, T.; Kameda, T.; Takada, S. On Easy Implementation Of A Variant Of The Replica Exchange With Solute Tempering In GROMACS. *J. Comput. Chem.* **2011**, 32, 1228-1234

(7) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput*. **2008**, 4, 435-447

(8) Hess, B.; Bekker, H.; Berendsen, H.J.C.; Fraaije, J.G.E.M. LINCS: A Linear Constraint Solver For Molecular Simulations. *J. Comput. Chem*. **1997**, 18, 1463-1472