# Supplementary Information

# Development of a Disaggregation-Induced Emission Probe for Detection of RecA inteins from *Mycobacterium Tuberculosis*

Xin Li, Hao Jiang, Zhaofeng Luo, Young Tae Chang and Liyun Zhang*

Institute of Technical Biology and Agriculture Engineering, Key Laboratory of Ion Beam Bioengineering, Hefei Institutes of Physical Science, Chinese Academy of Sciences, Hefei, Anhui 230031, P. R. China.

## Table of contents

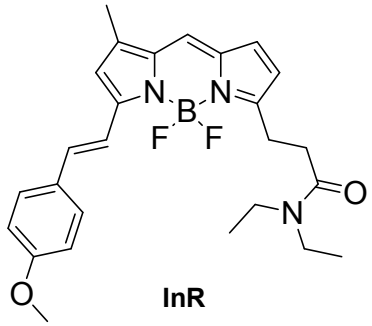
# A. Experimental

## Materials

All commercially available reagents and solvents were purchased from Sigma Aldrich, Alfa Aesar, Fluka, Merck or Acros, and used as received unless otherwise stated. $CH_2Cl_2$ (Fisher Scientific, analytical grade) was freshly distilled from $P_2O_5$ under nitrogen. Anhydrous THF was purchased from Alfa Aesar and used without further purification.

## Synthesis

**InR** was re-synthesized according to our previous published method[1], detail information is in below:




**InR**

**(*E*)-3-(5,5-difluoro-7-(4-methoxystyryl)-9-methyl-5*H*-4λ⁴,5λ⁴-dipyrrolo[1,2-*c*:2', 1'-*f*][1,3,2] diazaborinin-3-yl)-*N*,*N*-diethylpropanamide (InR):**

(E)-3-(5,5-difluoro-7-(4-methoxystyryl)-9-methyl-5H-5□4,6□4-dipyrrolo[1,2-c:2',1'-f][1,3,2]diazaborinin-3-yl)propanoic acid (**4**) was synthesized following the procedure described in reference[1].To a solution of **4** (10 mg, 24 μmol, 1 eq.) in 2 mL DMF/THF (1:1), was added HBTU (30 mg, 85 μmol, 3.5 eq.) and stirred at RT for 10 min. Diethylamine (4 μL, 37 μmol, 1.5 eq.) and triethylamine (11 μL, 85 μmol, 3.5 eq.) were subsequently added and the reaction mixture further stirred at RT for 3 h. Upon

completion, the reaction was quenched with water (20 mL), extracted with DCM (3 x 20 mL), dried over anhydrous $Na_2SO_4$ and evaporated *in vacuo*. The crude product was purified by flash column chromatography on silica gel (5:1 DCM/EA) to afford 2 as a deep purple solid (10 mg, 22 μmol, 92% yield); $^1$H NMR (500 MHz, DMSO-d6) δ 7.65 (d, *J* = 16.3 Hz, 1H), 7.61 (s, 1H), 7.59 (d, *J* = 8.8 Hz, 2H), 7.31 (d, *J* = 16.3 Hz, 1H), 7.07 (d, *J* = 4.0 Hz, 1H), 7.04 (s, 1H), 7.04 – 7.01 (m, 2H), 6.40 (d, *J* = 4.0 Hz, 1H), 3.81 (s, 3H), 3.30 – 3.27 (m, 4H), 3.12 (d, *J* = 7.4 Hz, 2H), 2.70 (dd, *J* = 8.9, 6.8 Hz, 2H), 2.31 (s, 3H), 1.08 (t, *J* = 7.1 Hz, 3H), 1.04 (t, *J* = 7.1 Hz, 3H); $^{13}$C NMR (125 MHz, DMSO-d6) δ 160.7, 157.2, 156.2, 143.2, 138.8, 135.9, 133.3, 129.6, 129.1, 128.5, 127.9, 123.1, 117.0, 116.7, 115.4, 114.7, 55.4, 41.2, 31.2, 22.0, 14.1, 13.1, 11.0; HRMS ($C_{26}H_{30}BFN_3O_2$): calc. $[M - F]^-$: 446.2414, found $[M - F]^-$: 446.2431.

**Protein expression and purification**

The inteins were overexpressed in E. coli host strain JM101 in Luria- Bertani medium as described previously [2]. The proteins were expressed in a fusion with chitin binding domain at the N terminus. The purification of the inteins was performed using affinity chromatography with chitin beads (New England Biolabs). After washing out impurities, the inteins were released from the chitin beads by 200 mM dithiothreitol (DTT).

**Absorbance**

UV/Vis absorption spectra of dyes and protein in buffer (20 mM Tris-HCl, 100 mM NaCl, pH 7.4) were recorded from 200 to 700 nm using a SpectraMax M2 spectrophotometer.

**Fluorescence**

Fluorescence measurements were carried out on a SpectraMax M2 spectrophotometer in 96-well plates by scanning the emission spectra between 540 and 700 nm ($\lambda_{ex}$ = 360 nm). All experiments were repeated three times. Data analysis was performed using Origin 8.0 (OriginLab Corporation, MA).

**The high throughput screening for $\Delta\Delta I_{hh}$-SM fluorescent sensor**

An unbiased screening was performed according to our previously approach[1, 3]. The preliminary screening concentration of $\Delta\Delta I_{hh}$-SM was set to 0.1 mg $mL^{-1}$ in buffer (20 mM Tris-HCl, 100 mM NaCl, pH 7.4). After screening of 5000 dyes generated from different fluorescent scaffolds, 23 compounds (data not shown) were selected as primary hits due to their fluorescence intensity change over 2 fold. Secondary screening was then carried out with a wide range of concentrations of $\Delta\Delta I_{hh}$-SM (i.e. 0.1, 0.2, 0.4 and 0.8 mg $mL^{-1}$) to eliminate the false positive results. Finally, 3 best responsive and reproducible hit compounds were rendered and molecular docking optimization was performed to validate the best final hit.

**The titration and competition experiments**

A standard procedure for the titration experiment, 1μL **InR** (2 mM) was added into 200 μL protein buffer solution (0-5 mg $mL^{-1}$) incubating at room temperature for 10 min. The fluorescence of the samples was measured on a SpectraMax M2 spectrophotometer in 96-well plates by scanning the emission spectra between 540 and 700 nm.

In the competition experiment, 1μL **InR** (2 mM) was added into 200 μL protein buffer solution (1 mg $mL^{-1}$$\Delta\Delta I_{hh}$-SM +1mg $mL^{-1}$ other protein) incubating at room temperature for 10 min. The fluorescence of the samples was measured.

**Quantum yield measurements**

Quantum yields were calculated by measuring the integrated emission area of the fluorescent spectra in its respective solvents and comparing to the area measured for Rhodamine B (reference) (10 μM, $\Phi_F$ = 0.31, $\lambda_{ex}$ = 500 nm) in water ($\eta$ = 1.333). Quantum yields were calculated using the equation:

$$\Phi_F^{sample} = \Phi_F^{reference}\left(\frac{F^{sample}}{F^{reference}}\right)\left(\frac{\eta^{sample}}{\eta^{reference}}\right)^2\left(\frac{Abs^{reference}}{Abs^{sample}}\right)$$

where *F* represents the area of fluorescent emission, $\eta$ is the refractive index of the solvent, and Abs is absorbance at the excitation wavelength. Emission was integrated between 540 and 700 nm.

**Dissociation Constant Measurements**

The $K_d$ of **InR** to the respective inteins was analyzed by Origin 8.0 (OriginLab Corporation, MA) using the following equation for a one site specific binding model:

$$y = \frac{y_{\max} x}{K_d + x}$$

Where $y$ represents the fluorescence fold change of **InR**, $y_{\max}$ the fluorescence fold change of **InR** when saturated with intein and $x$ the concentration of the intein in μM.

**Transmission Electron Microscope**

**InR** (10 μM) was first prepared in water and deposited on a thin copper-support film, followed by drying *in vacuo*. Images of the samples were obtained with JEOL JEM 3010 HRTEM microscope and operated at 100 kV without any contrast agent.

**Dynamic Light Scattering**

The dynamic light scattering of different concentration **InR** and other compounds were measured at 25°C in buffer using quartz cell. All measurements were performed in triplicate in Zetasizer Nano ZS.

**Molecular Modeling**

The crystallographic coordinates of inteins structures were obtained from the protein data bank (PDB ID 2IN0, 2IN8 and 2IMZ), processed using Reduce, and manually checked.[4, 5] The **InR** structure was optimized using the Gaussian03 program (B3LYP/6-31G* level).

AutoDock Vina[6] and AutoDock Tools were used for the molecular docking. A search space was chosen to include all the atoms of the protein dimer. The calculations with the exhaustiveness parameter of 256, 512 or 1024 were all performed. All 243 resulting conformations were analyzed in AutoDock Tools, and the conformation with the lowest energy was chosen for further analysis. All figures were rendered using PyMOL v0.99. (http://www.pymol.org).

## B. Supporting Figures

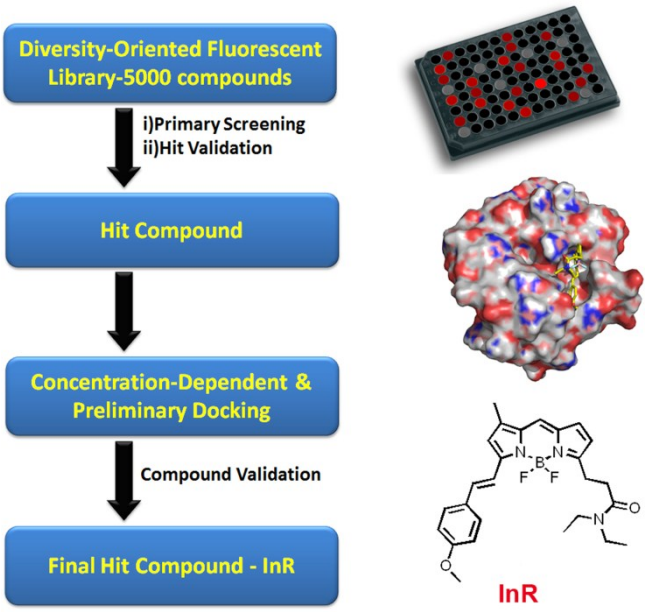



**Fig. S1** Schematic flowchart for the discovery of **InR**. Compound structures included in this screening can be found in the diversity-oriented fluorescent libraries reported.[7-13]

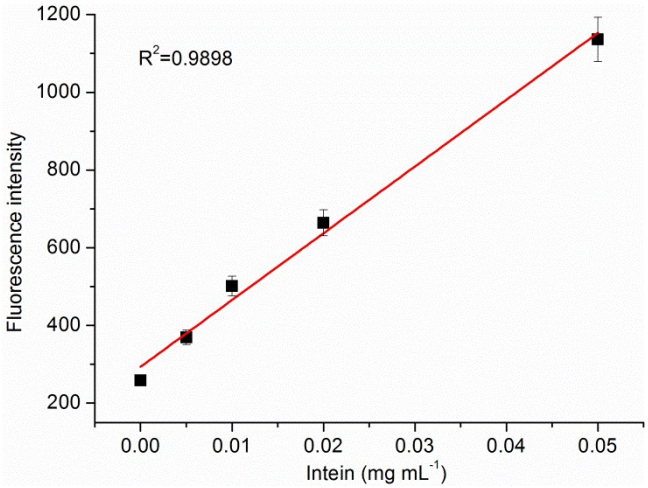



**Fig.S2** Fluorescence intensity changes of **InR** upon addition of $\Delta\Delta I_{hh}$-SM (0 – 0.05 mg $mL^{-1}$) in buffer (20 mM Tris-HCl,100 mM NaCl, pH 7.4).$\lambda_{ex}$: 360 nm. Values are represented as means (n = 3). Measurements were taken at RT.

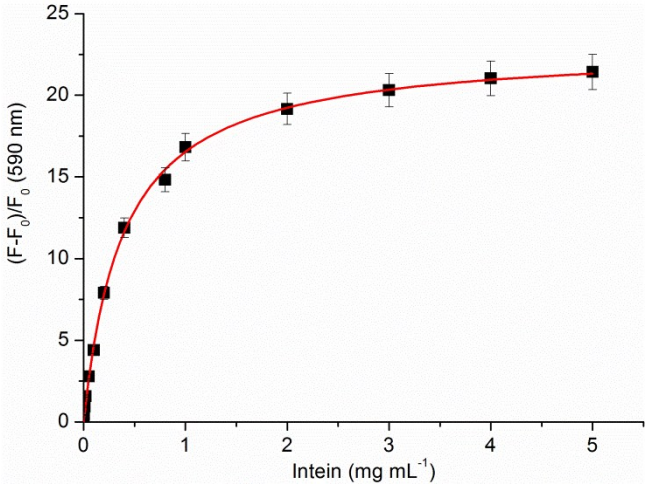



**Fig.S3** Fluorescence fold changes of **InR** upon addition of $\Delta\Delta I_{hh}$-SM (0 – 5 mg $mL^{-1}$) in buffer (20 mM Tris-HCl,100 mM NaCl, pH 7.4). $F_0$ and $F_{max}$ are the fluorescent maximum intensities of the **InR** in the absence and presence of $\Delta\Delta I_{hh}$-SM respectively. $\lambda_{ex}$: 360 nm. Values are represented as means (n = 3). Measurements were taken at RT.

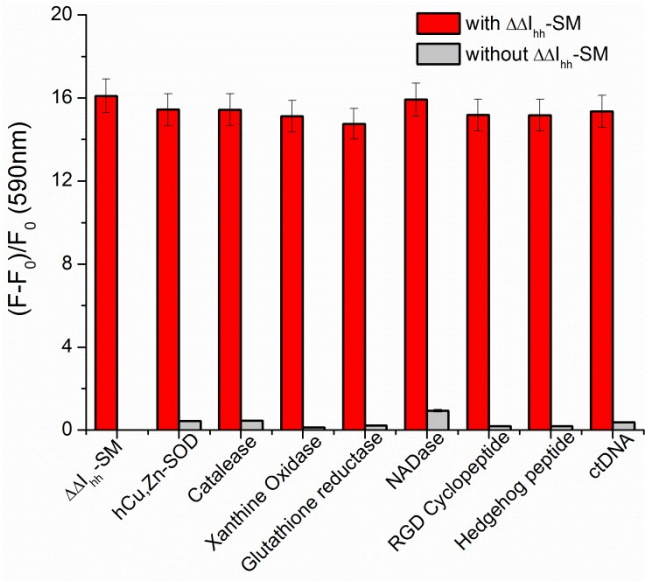



**Fig.S4**. Fluorescence response of **InR** (10 μM) upon interaction with various biomolecules (1mg $mL^{-1}$ together with and without $\Delta\Delta I_{hh}$-SM) in buffer(20 mM Tris-HCl,100 mM NaCl, pH 7.4). $\lambda_{ex}$: 360 nm. Values are represented as means (n = 3). Measurements were taken at RT.

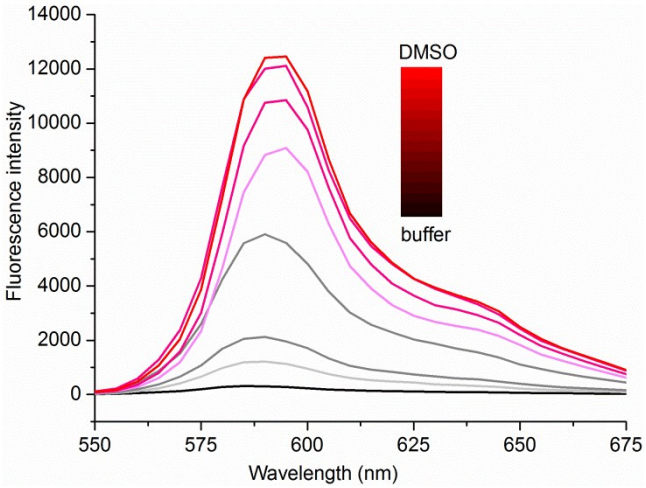



**Fig.S5**. Fluorescent emission spectra of **InR** (10 μM) in various DMSO-Buffer (20 mM Tris-HCl,100 mM NaCl, pH 7.4) mixtures. $\lambda_{ex}$: 360 nm. Values are represented as means (n = 3). Measurements were taken at RT.

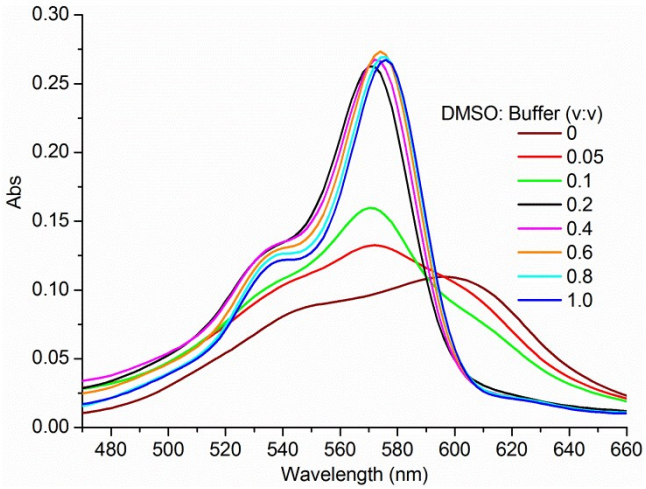



**Fig.S6**. Absorption spectra of **InR** (10 μM) in various DMSO-Buffer (20 mM Tris-HCl,100 mM NaCl, pH 7.4) mixtures. Values are represented as means (n = 3). Measurements were taken at RT.

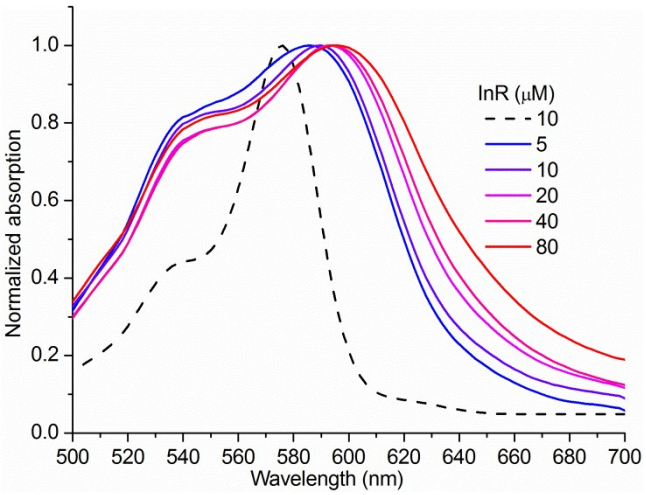



**Fig.S7** Absorption spectra of **InR** (5, 10, 20, 40 and 80 μM) in buffer (20 mM Tris-HCl,100 mM NaCl, pH 7.4). Dash line is absorption spectra of **InR** (10 μM) in DMSO. Values are represented as means (n = 3). Measurements were taken at room temperature (RT).

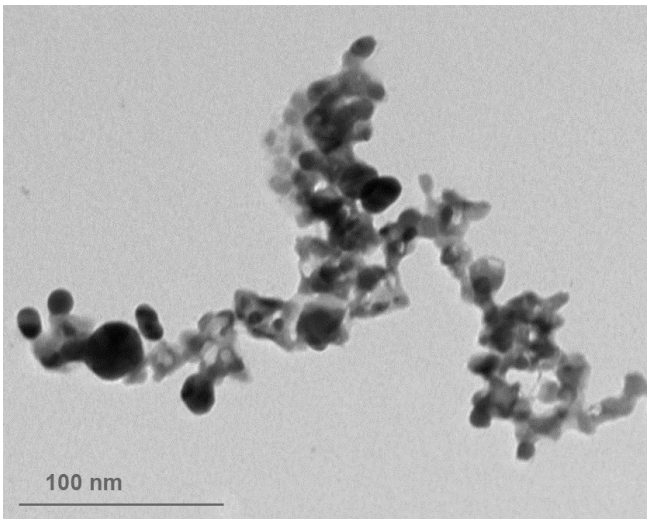



**Fig.S8** TEM image of **InR**-aggregates.

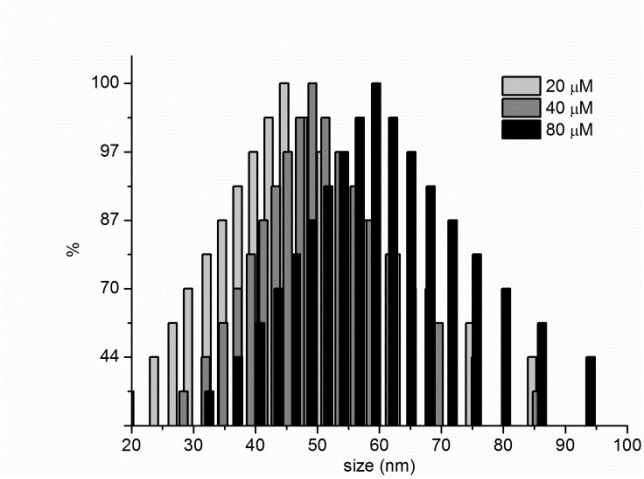



**Fig.S9** DLS measurement of **InR**-aggregates (20, 40, 80 μM) in buffer (20 mM Tris-HCl,100 mM NaCl, pH 7.4).

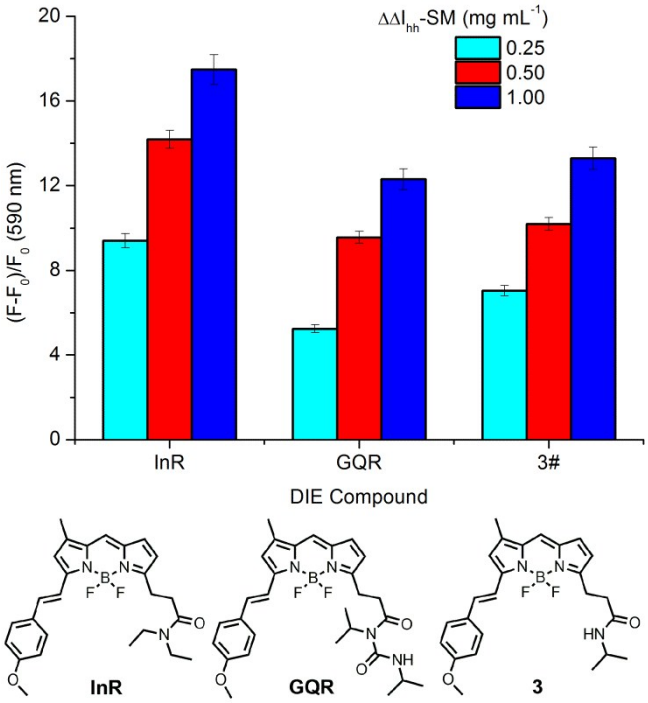



**Fig.S10** Fluorescence response of **InR** analogues (10 μM) upon incubation with serial concentrations of in 20 mM Tris-HCl buffer containing 100 mM (pH = 7.4). $\lambda_{ex}$: 360 nm. Values are represented as means and error bars as standard deviations (n = 3).

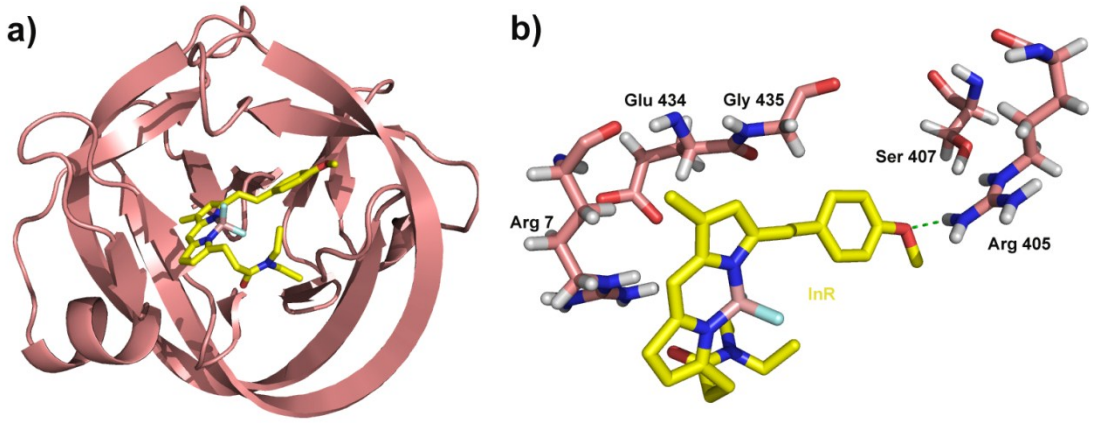



**Fig. S11** Molecular docking for the binding of **InR** at $\Delta\Delta I_{hh}$-CM (ID:2IN8). a) Illustration of the binding site of **InR** (yellow) at the surface of $\Delta\Delta I_{hh}$-CM. b) Suggested hydrogen bonding (green dash line) interactions between **InR** and different residues of $\Delta\Delta I_{hh}$-CM. Carbon atoms are colored in yellow for **InR** and salmon for $\Delta\Delta I_{hh}$-CM residues. Oxygen and nitrogen atoms are colored in red and blue respectively.

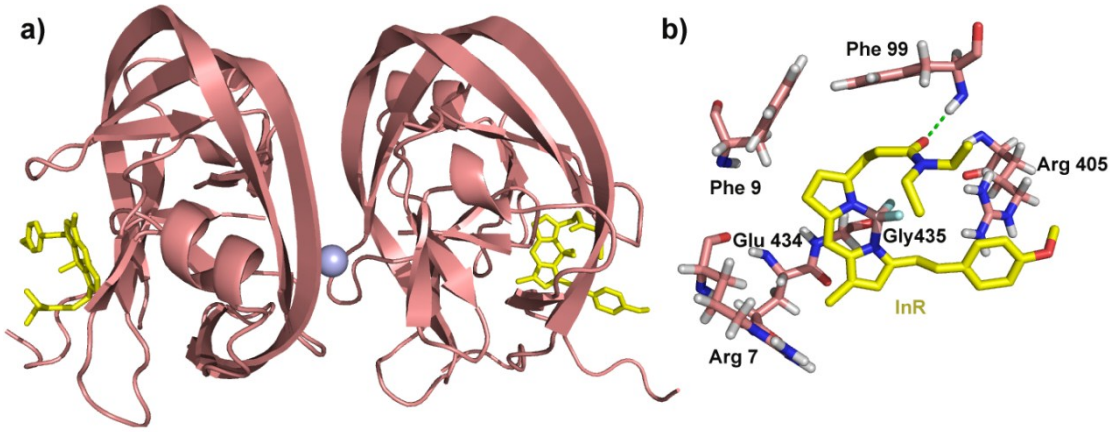



**Fig.S12** Molecular docking for the binding of **InR** at ΔI-SM(ID:2IMZ). a) Illustration of the binding site of **InR** (yellow) at the surface of ΔI-SM. b) Suggested hydrogen bonding (green dash line) interactions between **InR** and different residues ofΔI-SM. Carbon atoms are colored in yellow for **InR** and salmon for ΔI-SM residues. Oxygen and nitrogen atoms are colored in red and blue respectively.

## C. References


1. L. Zhang, J. C. Er, K. K. Ghosh, W. J. Chung, J. Yoo, W. Xu, W. Zhao, A. T. Phan and Y. T. Chang, *Sci. Rep.*, 2014, **4**, 3776.
2. L. Zhang, Y. Zheng, Z. Xi, Z. Luo, X. Xu, C. Wang and Y. Liu, *Mol. BioSyst.*, 2009, **5**, 644-650.
3. D. Zhai, B. K. Agrawalla, P. S. F. Eng, S.-C. Lee, W. Xu and Y.-T. Chang, *Chem. Commun.*, 2013, **49**, 6170-6172.
4. J. M. Word, S. C. Lovell, J. S. Richardson and D. C. Richardson, *J. Mol. Biol.*, 1999, **285**, 1735-1747.
5. R. W. Strange, S. Antonyuk, M. A. Hough, P. A. Doucette, J. A. Rodriguez, P. J. Hart, L. J. Hayward, J. S. Valentine and S. S. Hasnain, *J. Mol. Biol.*, 2003, **328**, 877-891.
6. O. Trott and A. J. Olson, *J. Comput. Chem.*, 2010, **31**, 455-461.
7. M. Vendrell, G. G. Krishna, K. K. Ghosh, D. Zhai, J.-S. Lee, Q. Zhu, Y. H. Yau, S. G. Shochat, H. Kim, J. Chung and Y.-T. Chang, *Chem. Commun.*, 2011, **47**, 8424.
8. D. Zhai, S.-C. Lee, M. Vendrell, L. P. Leong and Y.-T. Chang, *ACS Comb. Sci.*, 2012, **14**, 81-84.
9. J.-J. Lee, S.-C. Lee, D. Zhai, Y.-H. Ahn, H. Y. Yeo, Y. L. Tan and Y.-T. Chang, *Chem. Commun.*, 2011, **47**, 4508-4510.
10. J. C. Er, M. K. Tang, C. G. Chia, H. Liew, M. Vendrell and Y. T. Chang, *Chem. Sci.*, 2013, **4**, 2168-2176.
11. J. S. Lee, N. Y. Kang, Y. K. Kim, A. Samanta, S. H. Feng, H. K. Kim, M. Vendrell, J. H. Park and Y. T. Chang, *J. Am. Chem. Soc.*, 2009, **131**, 10077-10082.
12. Y.-H. Ahn, J.-S. Lee and Y.-T. Chang, *J. Am. Chem. Soc.*, 2007, **129**, 4510-4511.
13. N.-Y. Kang, S.-C. Lee, S.-J. Park, H.-H. Ha, S.-W. Yun, E. Kostromina, N. Gustavsson, Y. Ali, Y. Chandran, H.-S. Chun, M. Bae, J. H. Ahn, W. Han, G. K. Radda and Y.-T. Chang, *Angew. Chem., Int. Ed.*, 2013, **52**, 8557-8560.