Developing a DeSipher method for conformational change detection of membrane receptor complex at residue resolution

Qi Liu[1, 2, 3, #], Qing-tao He[1,3,4#], Xiao-xuan Lyu[1#], Fan Yang [2, 3, #], Zhong-liang Zhu[9#], Peng Xiao[3,4], Zhao Yang[2,3], Feng Zhang[1], Zhao-ya Yang[1,4], Xiao-yan Wang[1], Peng Sun[5], Qian-wen Wang[5], Chang-xiu Qu[3,4], Zheng Gong[4], Jing-Yu Lin[4], Zhen Xu[1], Shao-le Song[1], Shen-ming Huang[3], Sheng-chao Guo[3,4], Ming-jie Han[1,6], Kong-kai Zhu[10], Xin Chen[8], Alem W. Kahsai[11], Kun-Hong Xiao[12], Wei Kong[3], Xiao Yu[2], Ke Ruan[7], Fa-hui Li[1], Xiao-gang Niu[14], Chang-wen Jin[14], Jiangyun Wang[1,13]*, Jin-peng Sun[3,4*]

Affiliations:

[1]Institute of Biophysics, Chinese Academy of Sciences, 15 Datun Road, Chaoyang district, Beijing, 100101, China.

[2]Key Laboratory Experimental Teratology of the Ministry of Education and Department of Physiology, Shandong University School of Medicine, 44 Wenhua Xi Road, Jinan, Shandong, 250012, China.

[3]Key Laboratory of Molecular Cardiovascular Science, Ministry of Education, Peking University, 15 Xueyuan Road, Haidian District, Beijing, 100191, China.

[4]Key Laboratory Experimental Teratology of the Ministry of Education and Department of Biochemistry and Molecular Biology, Shandong University School of Medicine, 44 Wenhua Xi Road, Jinan, Shandong, 250012, China.

[5]Wuhan Institute of Physics and Mathematics, Chinese Academy of Sciences, 30 Xiaohongshan Road, Wuchang District, Wuhan, Hubei, 430071, China.

[6]Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, 32 Xiqi Road, Airport Economic Zone, Dongli District, Tianjin, 300308, China.

[7]Hefei National Laboratory for Physical Science at the Microscale, University of Science and Technology of China, 443 Huangshan Road, Hefei, Anhui 230027, China.

[8]Department of Medicinal Chemistry, School of Pharmaceutical Engineering and Life Science, Changzhou University, Changzhou, Jiangsu, 213164, China.

[9]School of Life Sciences, University of Science and Technology of China, 96 Jinzhai Road, Hefei, Anhui 230026, China.

[10] School of Biological Science and Technology, University of Jinan, 336 Nanxinzhuangxi Road, Shizhong District, Jinan 250022, China.

[11]Duke University, School of Medicine, Durham, North Carolina 27705, USA.

[12]Department of Pharmacology and Chemical Biology, School of Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania 15261, USA.

[13]College of Life Sciences and School of Future Technology, University of Chinese Academy of Sciences, Beijing 100049, China

[14]Beijing Nuclear Magnetic Resonance Center, College of Chemistry and Molecular Engineering, School of Life Sciences, Peking University, Beijing 100084, China

[#] These authors contributed equally to this work.

* Corresponding author: Jin-Peng Sun (corresponding author)
E-mail: sunjinpeng@sdu.edu.cn
Jiangyun Wang (corresponding author)
E-mail: jwang@ibp.ac.cn

Abstract

Characterization of the dynamic conformational changes in membrane protein signaling complexes by nuclear magnetic resonance (NMR) spectroscopy remains challenging. Here, we developed "DeSipher" technology to site-specifically incorporate 4-trimethylsilyl phenylalanine (TMSiPhe) into proteins and measured the conformational changes in the signaling complex by 1H NMR. Notably, crystallographic analysis revealed structural changes that reshaped the TMSiPhe-specific amino-acyl tRNA synthetase to accommodate the large trimethylsilyl (TMSi) group. The unique 1H NMR chemical shift of TMSiPhe and highly efficient incorporation of TMSiPhe enabled facile assignment of the associated NMR signal. We further showed that extracellular ligands induced conformational changes located in the polar core or ERK interaction site of β-arrestin-1 via direct receptor transmembrane core interactions in the phospho-β2AR/β-arrestin-1 signaling complex. These observations provided direct delineation and key mechanism insights that multiple receptor ligands were able to induce distinct arrestin functionally relevant conformational changes at residue levels.

Significance

Developing high sensitive and convenient structural characterization method for signaling protein complexes is important for understanding numerous physiological processes. Here we developed a new technology called DeSipher, through genetic code expansion to selectively and efficiently incorporate the unnatural element Silicon, as part of trimethylsilyl phenylalanine (TMSiPhe), at the desired positions of target proteins and utilizes the unique NMR spectrum to detect conformational changes in membrane protein signaling complexes. Application of this technology allowed us to decipher the conformational changes of arrestin at residue resolution, in a receptor signaling complex and in response to different ligands. These results provided new tools for structural study of arrestin and may facilitate rational drug design of arrestin-biased drugs downstream of receptor activation.

Keywords G protein-coupling receptor; β-arrestin-1; NMR; genetic code expansion; crystal structure.

All genetically coded natural proteins in living organisms consist of 5 main elements, namely, carbon (C), nitrogen (N), oxygen (O), sulfur (S) and hydrogen (H), but not silicon (Si), which is one of the most abundant elements on earth. Detection of proton NMR signals for the H atoms attached to amide nitrogen atoms or to carbon atoms in these proteins, as well

as multidimensional measurement or suppression of transverse relaxation, such as HSQC and TROSY (1-3), enables structural and dynamic characterization of protein complexes in important physiological processes. However, information regarding the structural dynamics of membrane proteins remains challenging to obtain using ^{1}H NMR spectroscopy due to the difficulty of spectrum assignment from overlapping proton signals in the large protein complexes, the low signal-to-noise ratio (S/N) and the need for large amounts of membrane proteins.

The chemical shifts of the H atoms attached to carbons are determined by the immediate environments. For instance, due to the high electronegativity of oxygen, nitrogen and unsaturated carbon, the ^{1}H NMR spectra of CH_3-O, CH_3-N and CH_3-$C_{sp2/sp}$ normally exhibit low field shifts between 1.5 ppm and 4.0 ppm. In contrast, the chemical shift of CH_3-C_{sp3} is between 0.7 ppm and 1.8 ppm (Fig.1*A* and Fig. S1*A*). In general, it is difficult to obtain structural information for large protein complexes at the residue level by one-dimensional proton NMR spectroscopy owing to the large number of overlapping signals. This difficulty has made characterization of the dynamic conformational changes in large membrane protein signaling complexes very challenging and has hampered rational drug design. To circumvent this obstacle, introduction of an element that is not found in native bio-macromolecules with distinct biophysical properties can provide unique advantage for structural characterization of large protein signaling complexes at the residue level (2, 4-6).

In the present study, we developed the "DeSipher" technology by incorporating the unnatural amino acid (UAA) 4-trimethylsilyl phenylalanine (TMSiPhe) at specific positions in protein complexes via genetic code expansion and using TMSiPhe-based ^{1}H-NMR to detect the conformational change at a specific site in a particular protein (7). The genetic code expansion enabled changes in the chemical properties of desired residues at both exposed and internal sites (8, 9), in contrast to traditional chemical labeling methods, which are available for only surface residues. Importantly, the low electronegativity of Si led to a uniquely low ^{1}H chemical shift for CH_3-Si (typically less than 0.55 ppm), which could be easily identified from the complicated ^{1}H spectra of macro-biomolecules (Fig. 1*A* and Fig. S1*A* and S1*B*). The three methyl groups of TMSiPhe provided nine degenerate ^{1}H spins, markedly increasing the S/N, thus decreasing the amount of sample required for NMR studies. We demonstrated that DeSipher is useful for characterization of the partner binding affinity and the activation procedure of arrestin, an important signal transducer downstream of most G-protein-coupled receptors (GPCRs) (5, 6, 10-17). Our crystallographic analysis provided a structural basis for recognition of UAA TMSiPhe by a specific tRNA synthetase. Moreover, application of the De**Si**pher method to study the β2-adrenergic receptor (β2AR)/arrestin signaling complex by ^{1}H-proton NMR characterization enabled us to not only monitor multiple conformational states of specific arrestin regions but also provide direct experimental evidence that the

distinct ligands on the same receptor drive functionally relevant conformations via direct receptor core/arrestin interactions, which laid an important foundation to direct future biased drug design for G-protein-coupled receptor superfamilies.

Results

Developing DeSipher technology by genetic code expansion

The genetic code expansion technique has been widely used recently to incorporate non-natural amino acids at desired positions in a protein to enable in-depth investigation of many important biological processes (18). Such a system normally includes a preferred UAA, an orthogonal aminoacyl-tRNA synthetase (aaRS)-tRNA pair normally derived from an in directed evolutionary selection system, and a host protein production organism (18). Therefore, we first selected a tyrosine analog, TMSiPhe, to incorporate the non-natural element Si into the protein, which was synthesized in five steps (Fig. S2). Trimethylchlorosilane, 1, and the Grignard reagent generated from 4-bromotoluene were coupled to form 4-trimethylsilyl toluene **2**, followed by free radical bromination to generate **3**. Reaction with a carbanion made from a glycine derivative followed by hydrolysis led to the production of TMSiPhe (19). The product TMSiPhe was then purified and used for selection of a tRNA synthetase that accommodated this UAA using a mutant library of the orthogonal *Methanococcus jannaschii* amber suppressor tRNA (*M*jtRNA)/tyrosyl tRNA synthetase (*M*j-TyrRS) pair (7, 20). The mutant library of the *M*j-TyrRS was designed by randomizing (NNK) six active site residues (Y32, L65, F108, Q109, D158 and L162) that were within 6.5 Å of the tyrosine substrate and by mutating any one of the six residues I63, A67, H70, Y114, I159, and V164 to G or keeping these residues unchanged as previously described (21). Both positive and negative selections were performed by using purified TMSiPhe. In the positive selection, cell survival was dependent on the suppression of an amber mutation in the chloramphenicol acetyltransferase gene in the presence of TMSiPhe. In contrast, cells were eliminated if amber codons in the barnase gene was suppressed by natural amino acids in the negative selection. Following three rounds of positive selection and two rounds of negative selection (7), a mutant *M. jannaschii* tyrosyl tRNA synthetase (*M*j-TyrRS) with specific activity toward TMSiPhe, termed TMSiPheRS, was identified. Sequence analysis revealed that the evolved TMSiPheRS harbors the mutations Tyr32His, Ile63Gly, Leu65Val, His70Gln, Asp158Gly, Ile159Gly and Val164Gly compared to wild-type *M*j-TyrRS (Fig. S3).

We next incorporated TMSiPhe into β-arrestin-1, a signaling protein used here as a model system for evaluation of the DeSipher method. In *E. coli*, only cells transformed with

the tRNA:aaRS pair for TMSiPhe and grown in the presence of TMSiPhe, but not other Si-containing amino acids, afforded full-length β-arrestin-1-H295TMSiPhe protein after His-tag affinity column purification, suggesting that the selected TMSiPheRS exhibited significant structural selectivity for TMSiPhe over other Si-containing unnatural amino acids (Fig. 1*B*). Mass spectrometric analysis unambiguously showed the incorporation of TMSiPhe at the desired position H295 in β-arrestin-1 with an approximately 100% efficiency after purification (Fig. 1*C*, Fig. S4 and Table S1). Moreover, the electron density of TMSiPhe was clearly observed in the crystal structure of a TMSiPhe-containing green fluorescence protein (GFP), an easily crystallized protein (Fig. 1*D* and Table S2). Crystallization of GFP-Y182TMSiPhe takes approximately 1 week at 16°C, thus confirming the stability of TMSiPhe after incorporation into a particular protein.

We then inspected the NMR properties of the TMSiPhe decorated protein. NMR spectroscopy of β-arrestin-1-H295TMSiPhe revealed a unique ^{1}H-NMR peak at 0.25 ppm, which provides a distinct NMR probe for examination of the structural dynamics of a specific site (Fig.1*E* and Fig. S5). Notably, the ^{1}H-NMR signal of TMSiPhe-incorporated arrestin can be detected at concentrations as low as 5 μM very rapidly (in less than 20 min) using a 950 MHz NMR spectrometer.

Molecular basis of the selective recognition of TMSiPhe by TMSiPheRS

To investigate the molecular basis of the selective recognition of TMSiPhe by TMSiPheRS, we crystallized TMSiPheRS and analyzed the structures by X-ray crystallography. The crystal structures of TMSiPheRS alone and the complex of TMSiPheRS with TMSiPhe were determined at 1.8 Å and 2.1 Å, respectively (Table S2). The 2Fo-Fc annealing omit map of the TMSiPheRS/TMSiPhe complex unambiguously assigned the electron density for TMSiPhe (Fig. 2*A*). Introduction of tri-methyl-Si at the parental tyrosine site significantly increased the volume at the site by approximately 60% (Fig. 2*B*). To compensate for this substantial change in volume, three residues, namely, Asp158, Ile159 and Val164, were replaced by the smallest amino acid Gly, and Tyr32 and Leu65 were substituted by the relatively small residues His32 and Val65, respectively (Fig. 2*B*-2*C* and Fig. S3). We then compared the crystal structure of the TMSiPheRS/TMSiPhe complex with that of apo TMSiPheRS. Compared to the structure of TMSiPheRS alone, we observed a dramatic 120-degree rotation of histidine 32 in response to TMSiPhe binding. Moreover, Leu162 rotated approximately 42 degrees to form hydrophobic interactions with the methyl groups of TMSiPhe (Fig. 2*D*). Altogether, Gly34, Val65, Gln70, Phe108, Gln109, Tyr151, Gln155, Gly158, Gly159, Gln173 and His177 defined a hydrophobic pocket for the accommodation of and specific interactions with the phenyl ring and trimethylsilyl group of TMSiPhe (Fig. 2*C*

194 and Fig. S3). These observations provided a structural basis for specific and efficient
195 incorporation of TMSiPhe using evolved TMSiPheRS.
196
197 **Incorporation of TMSiPhe at specific sites in β-arrestin-1 enabled characterization of**
198 **β-arrestin-1 activation**
199
200 We then incorporated TMSiPhe into functionally relevant structural motifs of β-arrestin-1,
201 the key signal transducer downstream of almost all 800 GPCRs encoded in the human
202 genome, which functions not only by desensitizing membrane receptors but also by mediating
203 independent downstream signaling after receptor activation (5, 6, 10-17, 22) (Fig. 3*A* and 3*B*).
204 Although the functions of many arrestin-mediated receptors have been identified and certain
205 motifs of arrestin are suspected to be involved in specific signaling pathways (Table S3), the
206 correlation between the conformational states of these arrestin motifs and selective receptor
207 functions remains to be elucidated. Incorporation of TMSiPhe into β-arrestin-1 at specific
208 positions, including the receptor-phosphate-binding site (Y21), the finger loop (Y63), the
209 hinge region (Y173), the β-sheet XVI (Y249), the loop between β-sheets XVIII and XVIIII
210 (R285), the lariat loop (H295) and the C-terminal swapping region (F388), led to
211 unambiguous assignment of NMR peaks between -0.3 ppm and 0.3 ppm in the ^{1}H-NMR
212 spectrum (Fig. 3*B*, Fig. S6-S7 and Table S4). These positions were proposed to be associated
213 with specific arrestin functions, including receptor or IP6 interactions, or the activation of
214 downstream ERK or AP2 but have never been fully characterized by biophysical methods
215 (Table S3). Therefore, TMSiPhe-containing β-arrestin-1 proteins provide a useful tool for
216 monitoring conformational changes in arrestin in response to receptor activation or other
217 stimuli. For example, the ^{1}H-NMR spectrum of native β-arrestin-1 F388-TMSiPhe exhibits a
218 peak at -0.05 ppm, which can be easily identified, in contrast to another UAA,
219 O-tert-butyltyrosine, the peaks for which cannot be assigned (Fig. S8). Notably, in response to
220 stimulation with increased concentrations of phospho-vasopressin-2 receptor-C-tail (V2Rpp),
221 the peak at -0.05 ppm gradually disappears, whereas a ^{1}H-NMR peak at 0.15 ppm appears,
222 reflecting the transition of the "inactive" arrestin conformation to an "active" arrestin
223 conformation at the F388 position by dislodgement of this specific C-terminal swapping
224 segment (Fig. 3*C* and 3*D*). Moreover, the Scatchard plot for the titration experiment
225 performed to examine the binding of V2R-phospho-C-tail to β-arrestin-1 exhibits a straight
226 line with a regression coefficient of 0.99. The calculated K_D value for the interaction of
227 V2Rpp with β-arrestin-1 was 6.9 ±0.2 μM (Fig. 3*E* and Fig. S9).
228
229 **Observation of the conformational change of the polar core of β-arrestin-1 by different**
230 **GPCR ligands through DeSipher**

231

232 Recent electron microscopic studies and computational simulations, as well as our
233 ^{19}F-NMR studies, have suggested that the receptor transmembrane core and the corresponding
234 phosphorylated C-tail were each able to independently stimulate arrestin activation (5, 6,
235 11-13, 23-29)(Fig. 4*A*). While the recent rhodopsin/visual arrestin complex structure has
236 provided a model of the interactions of the receptor core with arrestin at an atomic resolution
237 (29), there is little structural information regarding receptor core-induced structural
238 rearrangement of arrestin at the residue level due to technological difficulties in distinguishing
239 the contributions of the receptor core, the receptor-phospho-tail, or the linker and arrestin
240 mutant used in the crystal structures individually. This lack of data has hindered drug
241 development for the control of arrestin function by direct regulation of ligand-receptor
242 interactions.

243 Thus, incorporation of TMSiPhe at specific positions in arrestin might facilitate detection
244 of ligand-induced conformational changes in arrestin using DeSipher. One hallmark of
245 arrestin activation is the approximately 20-degree twist between the N- and C-domains of the
246 protein (Fig. 4*A*). In the inactive state, the N- and C-domains of β-arrestin-1 are tethered by
247 the polar core, which is composed of the extensive charged interactions of Asp26, Arg169,
248 Asp290, Asp297 and Arg393 (Fig. S10). Disruption of the salt bridge between Asp297 and
249 Arg393 and that between Asp304 and Arg382, as well as the equivalent rhodopsin-visual
250 arrestin interactions between Asp296 and Arg175 and Asp303 and Arg382, are known to
251 activate arrestin (29) (Fig. S10). Notably, the results of recent molecular dynamics studies
252 have indicated that the rotation of Asp296 of visual arrestin (Asp290 in β-arrestin-1) is closely
253 associated with interdomain twisting. We therefore incorporated TMSiPhe at the H295
254 position, which is close to both D290 and D297 of β-arrestin-1, to monitor the
255 receptor-induced conformational changes in the polar core (Fig. 4*A*-4*B* and Fig. S10).

256 Specific incorporation of TMSiPhe at the H295 position in β-arrestin-1 did not impair the
257 structural integrity of the protein, as H295-TMSiPhe-β-arrestin-1 exhibited normal activation
258 in response to the V2-receptor-phospho-tail interaction (Fig. S11). For structural validation of
259 the DeSipher study with H295-TMSiPhe-β-arrestin-1, we performed ^{1}H-NMR measurements
260 using the conditions for the crystal structures of β-arrestin-1 in both the inactive apo-arrestin
261 and in active arrestin stabilized by vasopressin 2 receptor phospho-tail (V2Rpp) and the
262 conformationally selective antibody Fab30 (26). Notably, in the "two-step arrestin recruitment
263 model of the receptor", V2Rpp/β-arrestin-1 mostly exhibited the "hanging" mode, whereas
264 the phospho-receptor/β-arrestin-1 complex encompassing the core interaction represented the
265 "snuggly" mode (23, 26) (Fig. 4*A*). Superimposition of the inactive and active arrestin
266 structures revealed that both the "hanging" and "snuggly" modes of active arrestin had similar
267 conformations at the H295 position, differing significantly from the modes of inactive arrestin,

268 which featured considerable movement of the lariat loop (Fig. 4*B*). In the inactive state, the
269 DeSipher spectrum of H295-TMSiPhe-β-arrestin-1 contained mainly one peak at 0.25 ppm,
270 which was designated S1 (Fig. 4*C*). Upon increasing the concentration of V2Rpp, the peak
271 volume of S1 gradually decreased, accompanied by the growth of a new peak at 0.15 ppm,
272 which was designated S2. The DeSipher spectrum obtained with a saturating concentration of
273 V2Rpp mainly exhibited an S2 peak, indicating that S2 represented an active state of
274 H295TMSiPhe, whereas S1 represented the inactive state of β-arrestin-1 H295TMSiPhe (Fig.
275 4*B* and 4*C*).

276 We then inspected the conformational change at the H295 site in response to occupation
277 of the receptor by a panel of ligands with the same phospho-receptor-tail by DeSipher using
278 the β2 adrenergic receptor (β2AR) as a prototypic model. As previously described, we
279 obtained the phospho-β2AR-V2-tail chimera (ppβ2V2R) by stimulating Sf9 cells with ISO
280 (Isoproterenol), a low-affinity ligand, before harvesting the cells and washing out the residual
281 ligands by affinity chromatography (30). The purified ppβ2V2R was then incubated with
282 various ligands and then used to form a stable receptor/arrestin complex by further incubation
283 with β-arrestin-1 and the conformationally selective antibody fragment Fab30 (Fig. S12).
284 Complex formation was verified by size-exclusion chromatography, and DeSipher was
285 performed to monitor changes in the NMR signal (Fig. 4*D*). Application of Fab30 alone had
286 no significant effect on the NMR spectrum of β-arrestin-1 H295TMSiPhe (Fig. S13). Notably,
287 upon incubation with ppβ2V2R alone, a new NMR signal appeared at 0.07 ppm (designated
288 S3), which was associated with the decrease in the S1 peak (Fig. 4*D* and Fig. S14-S15).
289 Therefore, S3 may represent the active arrestin state of H295TMSiPhe in the presence of
290 ppβ2V2R and the detergent MNG3, which was used to stabilize the receptor. The sharp S3
291 peak compared to S1 might indicate a highly solvent-exposed structure of the H295 state in
292 β-arrestin-1 after forming the complex with the ppβ2V2R, as observed in the crystal structure
293 of the rhodopsin/visual arrestin complex.

294 We next examined the DeSipher spectrum of the ppβ2V2R/β-arrestin-1 complex in the
295 presence of various β2AR ligands exhibiting different pharmacological activities. Importantly,
296 while the S1 state population of H295TMSiPhe decreased upon addition of different agonists,
297 including the full agonists ISO and BI-167107, or the partial agonists Clenbuterol (Clen) and
298 Salmaterol (Salm), the S3 state population increased (Fig. 4*E* and Fig. S15-S16). In particular,
299 the arrestin-biased β2AR agonist, Carvedilol (Carv), also promoted the increase in the S3
300 state with a corresponding decrease in the S1 state. In contrast, the neutral antagonist
301 Alprenolol (Alp) showed no effect on the S3 state, whereas the inverse antagonist ICI-118551
302 (ICI) reduced the population of the S3 state (Fig. 4*D* and Fig. S15). Moreover, the volume of
303 the S3 state corresponds to the function of the state in receptor internalization (Fig. 4*E* and
304 Fig. S17). These trends mirrored the ability of the ligand to promote torsion between helix VI

305 and helix III, a hallmark of the conformational changes in the receptor 7-transmembrane core
306 induced by agonists (Fig. 4*F* and Fig. S18). Overall, the DeSipher spectrum of β-arrestin-1
307 H295TMSiPhe indicated that the conformational changes in the polar core of the arrestin in
308 response to ligand properties are associated with the abilities of these ligands to promote
309 receptor internalization in the presence of the same receptor phosphorylation barcode
310 (Fig. 4*E*-4*F* and Fig. S17).
311
312 **The role of the receptor 7TM core in mediating the ligand-regulated conformational**
313 **change in the polar core of β-arrestin-1**
314
315 To confirm that the observed S3 signal in the DeSipher spectrum was dependent on the
316 interaction of β-arrestin-1 with the receptor core, we performed a competition assay using a
317 well-characterized binding partner of the receptor 7-transmembrane core, namely, the Gα
318 protein C-tail (Gα-CT) (Fig. 5*A*). Recently obtained crystal structures of the
319 rhodopsin/arrestin, rhodopsin/Gα-CT and rhodopsin/ArrFl (arrestin-finger loop peptide)
320 complexes have revealed a significant overlapping interface of the receptor 7-transmembrane
321 core with the finger loop of arrestin and Gα-CT of G-protein (29, 31). Moreover, direct
322 engagement of the β-arrestin-1 finger loop and Gα-CT of G-protein serves as a major
323 interaction interface with the receptor 7-transmembrane core, which was supported by recent
324 cross-linking and electron microscopic studies (26, 32-34). Therefore, we prepared the
325 ISO/ppβ2V2R/β-arrestin-1/Fab30 complex in the presence of Gα-CT. Incubation with Gα-CT
326 did not disrupt the ISO/ppβ2V2R/β-arrestin-1/Fab30 complex, as the presence of Gα-CT did
327 not alter the SEC (size-exclusion chromatography) profile (Fig. 5*B*). Notably, while
328 incubation of Gα-CT with His295-TMSiPhe-β-arrestin-1 led to no significant alteration in the
329 DeSipher spectrum, addition of Gα-CT with the ISO/ppβ2V2R/β-arrestin-1/Fab30 complex
330 significantly decreased the S3 state and increased the S1 state, suggesting that the observed
331 S1 state reduction was mainly due to elimination of the receptor core interaction with
332 β-arrestin-1 by the binding of Gα-CT (Fig. 5*C*). Because H295 is located in close proximity to
333 the polar core residues Asp290 and Asp297 of β-arrestin-1, the DeSipher spectra obtained
334 from Gα-CT competition experiments with H295-TMSiPhe-β-arrestin-1 confirmed that the
335 agonist ISO was able to induce conformational changes in the polar core of β-arrestin-1 via
336 direct transmembrane core interactions.
337 As both the receptor core and the phosphorylated receptor C-tail contributed to the
338 interaction between ppβ2V2R and arrestin, we next performed a V2R-phospho-C-tail
339 competition experiment (Fig. 5*D*). Incubation of the excess V2Rpp led to the dissociation of
340 ppβ2V2R from the ISO/ppβ2V2R/β-arrestin-1/Fab30 complex, as suggested by the SEC
341 results (Fig. 5*E*). However, the active conformation of H295 persisted even when the arrestin

342 dissociated from the receptor, as indicated by the maintenance of the amplitude of the S3 state
343 in the NMR spectrum (Fig. 5*F*). These data suggested that the S3 conformational state of
344 H295 in arrestin do not simply reflect the propensities of the ligands for stabilization of the
345 GPCR-arrestin complex, but is a consequence of the receptor-core-arrestin interaction.
346 Furthermore, maintenance of the S3 activation state of arrestin even after dissociation from
347 the receptor was consistent with the hypothesis proposed in recent cellular studies, which
348 suggested that an arrestin activation cycle occurred in response to activation by the receptor.
349
350 **Multiple conformational states observed at the ERK interaction site of β-arrestin-1**
351
352 We next extended the DeSipher technology to study the conformations of other
353 β-arrestin-1 sites associated with specific arrestin functions. We selected the R285 position of
354 β-arrestin-1, which was hypothesized to play important roles in interaction with ERK (35)
355 (Fig. 3*A* and Table S3). Notably, superimposition of the structures of inactive β-arrestin-1
356 structure and the rhodopsin-visual arrestin complex indicated that R285 assumed a highly
357 exposed and extended conformation (Fig. 6*A*), suggesting that receptor interaction may
358 regulate conformational change at this specific site.
359 We therefore incorporated TMSiPhe at the R285 site of β-arrestin-1 and monitored the
360 change in the ^{1}H-NMR spectrum in response to the binding of ppβ2V2R engaged with
361 different ligands (Fig. 6*B* and Fig. S19-S21). The functional integrity of
362 R285TMSiPhe-β-arrestin-1 was validated, and Fab30 was used to stabilize the
363 ppβ2V2R/β-arrestin-1 complex without perturbation in the NMR spectrum (Fig. S11 and S19).
364 Application of ppβ2V2R without or with different ligands eliminated the original NMR peak
365 at 0.158 ppm but broadened the conformational distributions from 0.03 ppm to 0.10 ppm (Fig.
366 6*B*-6*C* and Fig. S21). At least 5 different conformational states of the
367 ppβ2V2R/β-arrestin-1-R285TMSiPhe/Fab30 complex were discerned in the presence of
368 different ligands. Notably, β-arrestin-1 alone also has small but visible peaks in the 0.03
369 ppm-0.10 ppm region, indicating that a conformational selection model may also be suitable
370 for description of the receptor-induced conformational change at the β-arrestin-1-R285
371 position. In particular, addition of any receptor complexes without or with different ligands all
372 produced a similar peak at 0.09 ppm (R0 state), indicating that this conformational state may
373 be mainly due to the binding of the receptor-phospho-tail but is not significantly affected by
374 the receptor core interaction (Fig. 6*B* and Fig. S21).
375 The ligands mostly changed the distribution of NMR peaks from 0.04 ppm to 0.07 ppm,
376 which included 4 conformational states derived by simulation, namely, R1a-c (0.06-0.068
377 ppm) and R2 (0.05 ppm). Although application of the neutral antagonist Alp and the agonist
378 ISO had no significant effect on the NMR peak at R1a (0.065 ppm), application of the

379 long-term covalent agonist BI caused a small but significant low field shift of R1a to R1b
380 (0.068 ppm). In particular, engagement of β2V2R with the arrestin-biased ligand carvedilol
381 caused a shift of R1a to the R1c state (0.06 ppm), suggesting that a unique conformation of
382 the β-arrestin-1 R285 position may participate in carvedilol-induced ERK activation (Fig.
383 6*B*-6*C* and Fig. S21).

384 The application of partial and inverse agonists caused complex conformational changes.
385 Whereas Clen significantly diminished the distribution of the R1a state and promoted the
386 appearance of a high-field R2 state, the engagement of the receptor with the G-protein-biased
387 partial agonist Sal and the inverse agonist ICI almost completely eliminated the presence of
388 the R1 states and facilitated the emergence of the R2 states (Fig. 6*B*-6*C* and Fig. S21). As Sal
389 and ICI are not known for arrestin-dependent ERK signaling, the appearance of the R2
390 conformational states of the β-arrestin-1 R285 position may not contribute to ERK activation
391 in response to receptor/arrestin complex interactions. Taken together, multiple conformational
392 states of the β-arrestin-1 R285 position were detected by DeSipher in response to different
393 β2AR ligands, which was not strictly correlated with the ability of these ligands in either the
394 activation of G-protein (agonists vs. antagonists) or arrestin-mediated receptor internalization,
395 indicating that each specific receptor ligand may lead to a distinct conformational state at a
396 specific arrestin site, which contributes to the selective functions of these ligands.

397

398 **Discussion**

399

400 Despite the applications of NMR in determination of protein structure and
401 characterization of dynamic bio-macromolecular interactions, the use of this technique for the
402 study of large transmembrane protein complexes remains challenging. These complexes often
403 exhibit line broadening and overlapping resonance due to their large sizes as well as flexibility.
404 Recently, the introduction of non-natural elements (such as fluorine) that are not found in
405 bio-macromolecules has provided an exciting new method for the study of membrane protein
406 complexes due to the large chemical shift dispersion associated with these elements and the
407 minimal interference with background signals (2, 4, 27, 36, 37). One limitation of such a
408 method is that chemical labeling allows access to only the surface residues of proteins,
409 preventing observation of the important dynamic interactions that occur within protein
410 hydrophobic cores. As an alternative, UAA incorporation by genetic code expansion allows
411 labeling of desired residues at both exposed and internal sites. For example, by genetic code
412 expansion, we have developed a method to efficiently incorporate the UAA difluorotyrosine
413 (F2Y) into proteins of interest, enabling us to study how different receptor phospho-barcodes
414 localized in the receptor C-tail regulate distinct functionally selective arrestin conformations
415 (5, 6). Despite this significant progress, ^{19}F-NMR requires large amount of protein (usually

more than 100 μM), and each measurement generally takes more than 12 hours for production of a suitable S/N. Therefore, the development of a new chemical biological approach for examination of the conformational dynamics of transmembrane protein complexes using a low concentration of protein is urgently required.

In this context, we developed a new technique for expansion of the genetic code, namely, DeSipher, to incorporate the element Si into biomolecules and to utilize the unique proton NMR properties of the trimethylsilyl group. The efficient incorporation of TMSiPhe was verified by both mass spectrometry and crystallography. Parallel to our study, notably, TMSiPhe incorporation was also achieved by an in vitro cell-free translation system, but the characteristics and applications in biological studies have not been fully examined, due to the very limited amount of protein which can be produced using such a method (38). Here, by using our newly developed DeSipher method and a 950 MHz NMR instrument, we were able to detect the dynamic conformational changes in a transmembrane receptor supercomplex at the residue level at a low protein concentration ranging from 5 μM to 20 μM and with a short detection time of 9~20 min. Notably, a key to this advancement was the evolution of TMSiPheRS, a specific tRNA synthetase that selectively recognizes an artificial amino acid containing the element Si and incorporates this amino acid at the desired positions in the signaling molecule. Our newly solved crystal structures of TMSiPheRS and its complex with TMSiPhe revealed that substitution of 5 key residues in the amino acid-binding pocket, namely, Asp158, Ile159, Val164, Tyr32 and Leu65, as well as favorable hydrophobic interactions and unexpected ion-dipole interactions, are essential for incorporation of the inorganic atom Si into the protein. Thus, in addition to using evolved cytochrome to construct carbon-silicon bonds of micromolecules (39-42), our study provided a new method for genetic incorporation of the unnatural element silicon into proteins, allowing exploitation of the unique chemical properties of this element in different biochemical or biophysical processes.

By applying our newly developed DeSipher method, we were able to observe ligand-dependent conformational changes in arrestin via direct receptor core engagement. Notably, biased ligands with the ability to selectively perform specific arrestin or G-protein subtype functions are crucial for the development of the next generation of drugs targeting GPCRs (10). Recent studies by us as well as by others have provided important mechanistic insights, demonstrating that receptor-phospho-barcodes present in the receptor-C-tail play pivotal roles in the determination of selective arrestin functions (5, 6, 12, 13, 23, 24, 26). However, it remains very difficult to exploit this knowledge for the development of selective arrestin-biased GPCR ligands, partially due to the system bias generated by differential expression of GRKs in different cellular contexts (10). For example, by using the newly developed sortase-mediated ligation technology, recent experimental results have suggested

453 that the β2AR mutant, presumably with a core conformational defect and an arrestin-biased
454 function, did not acquire its biased function via direct receptor core interactions with arrestin
455 but instead acquired bias via selective GRK recruitment, indicating that receptor
456 phosphorylation plays a major role in biased signaling in this case (43). Therefore,
457 exploitation of the phospho-barcode to selectively regulate arrestin functions requires
458 understanding of the specific expression of GRK family members in a particular
459 physiological context. Moreover, structural characterization of receptor/GRK complex is in its
460 infancy (44), leading to further challenges in the development of biased ligands by utilization
461 of specific receptor-phospho-barcodes.

462 An alternative hypothesis for the development of arrestin-biased GPCR ligands is that
463 ligands for GPCRs can cause conformational changes in arrestin via direct receptor
464 core/arrestin interactions regardless of the C-terminal phosphorylation pattern. Notably, the
465 rhodopsin/visual arrestin complex crystal structure provided knowledge of receptor
466 core/arrestin interactions at the atomic level, and the FlAsH-BRET assays revealed that
467 different receptor activation resulted in diverse arrestin conformations in cells. However,
468 whether various ligands for a single receptor are able to direct selective conformational
469 changes of arrestins independent of receptor phosphorylation has not been demonstrated
470 because rhodopsin has only one covalent ligand, retinal, and there are distinct receptor
471 phosphorylation patterns in response to different agonist stimulations, as well as a lack of
472 effective methods for distinguishing the individual contributions of the receptor core and
473 phosphocodes. Recently, by developing the "finger loop" deletion mutant (FLD) or truncation
474 of the third loop of β2AR, the Shukla group, Lefkowitz group and our group have
475 demonstrated that functional characterization of the receptor core/arrestin interaction is
476 possible (6, 23, 30). In spite of these advancements, dynamic information and high-resolution
477 data regarding conformational changes in arrestin dictated by different receptor ligands via
478 the receptor core/arrestin interaction remains unclear, likely due to the low resolution of
479 cellular methods and the difficulty of the application of biophysical approaches for the study
480 of receptor complex systems. Here, by using the residue-specific conformational detection
481 method DeSipher, as well as fluorescence quenching experiments and cellular internalization
482 assays, our experimental results suggested that ligands directed structural alterations of the
483 7-transmembrane core of the receptor via interactions with arrestin; the ligands were
484 competed off by the Gs-peptide but not the V2Rpp tail and were able to cause conformational
485 change in the arrestin polar core at the specific residue H295, which is correlated with the
486 internalization ability of the receptor/arrestin complex. In this context, our work provided
487 direct experimental evidence that the extracellular ligand could induce a distinct conformation
488 and active state of arrestin via direct receptor core engagement at an explicit residue
489 resolution, indicating selective arrestin functions, without the need to consider the sources of

the system bias caused by differential GRK expression.

In addition to structural alterations in the polar core, we used DeSipher to examine the conformational changes that occurred at the R285 position of β-arrestin-1, a site associated with ERK activation (35). Importantly, R285TMSiPhe assumed multiple conformations in response to the engagement of different ligands with β2V2R harboring the same phosphorylated receptor C-tail. Importantly, the conformational states of the 285 site are not directly correlated to the functions of these β2AR ligands in either Gs activation or receptor internalization, indicating that different ligands of the same receptor were able to regulate distinct arrestin conformations at specific arrestin sites, which may be correlated with selective functions. Notably, the arrestin-ERK interaction may involve multiple interfaces. Therefore, the conformational changes in the R285 site observed by DeSipher likely contribute to, but are not the sole determinants of, arrestin-mediated ERK activation.

In summary, we have developed a new method, named DeSipher, for rapid detection of the dynamic conformational changes in transmembrane signaling supercomplexes using ^{1}H NMR within several minutes and for samples at micromolar concentrations. Using this approach, we identified the ligand-induced and functionally relevant arrestin conformational states via receptor core engagement, providing new clues for arrestin-biased GPCR ligand development (39).

Materials and Methods

Complete details and descriptions of reagents and constructs, synthesis of TMSiPhe, genetic selection of the mutant synthetase specific for TMSiPhe (TMSiPheRS), protein expression and purification, crystals preparation, data collection and structure determination, peptide synthesis, Superdex Exclusion Chromatography, NMR experiment, GST pull down assay, Bimane labeling of purified receptors and TRIQ experiment, BRET assay, Intra-gel digestion and LC-MS/MS analysis and database search, Q-TOF mass spectrometry spectrum analysis and database search are available in the *SI Materials and Methods*. The sfGFP Y182TMSiPhe, TMSiPheRS alone and TMSiPheRS-TMSiPhe complex crystal structures and associated diffraction data have been deposited in the Protein Data Bank with the accession codes 6KRG, 6J97 and 6J96, respectively.

Acknowledgements

NMR experiments were performed at the performed at the Beijing NMR Center and the NMR facility of National Center for Protein Sciences at Peking University, and Nuclear Magnetic Resonance Science Research Platform at Wuhan Institute of Physics and Mathematics, Chinese Academy of Sciences (CAS). We thank the staff of BL19U1 beamlines at National Center for Protein Sciences Shanghai and Shanghai Synchrotron Radiation

Factility for assistance during data collection. We thank Prof. Mai-li Liu and Prof. Xu zhang from Wuhan Institute of Physics and Mathematics, CAS, Hong-wei Li from Beijing nuclear magnetic resonance center, Ms Shan-shan Zang and Dr Xue-hui Liu from the Core Facility of Protein Research, Institute of Biophysics (IBP), CAS, for their help in the NMR data collection, analysis and valuable discussion. We thank Jun-ying Jia and Shuang Sun from Lab of Structural and Functional Analysis (IBP, CAS) for their technical assistance in flow cytometry analysis, and their colleague Jian-Hui Li for assistance in fluorescence measurement. We thank Xiang Ding and Zhen-sheng Xie from Lab of Proteomics (IBP, CAS) for LC-MS analysis.

We acknowledge support from the National Key Basic Research Program of China (2018YFC1003600 to J.-P.S; 2017YFA0505400, 2016YFA0501502 to J.Y.W. and F. H. L.), the National Natural Science Foundation of China (81773704 and 31611540337 to J.-P.S., 21837005, 21750003 to J.Y.W., U1532150 and 91640106 to F. H. L., 31700692 to P.X.), The National Science Fund for Distinguished Young Scholars (81825022 to J.-P.S.) and the Program for Changjiang Scholars and Innovative Research Team in University (IRT13028).

Author contributions: J.-P.S. conceives the idea for extracellular ligands induced conformational changes in β-arrestin-1 via direct receptor 7-transmembrane core interactions. J.Y.W. conceived the idea that TMSiPhe should be an excellent nuclear magnetic probe, and designed evolutionary strategy for TMSiPheRS. J.-P.S., J.Y.W. and X.Y. designed most of the experiments. Q.L., F.Y., P.S., Q.-W.W. and F. Z. collected and analyzed the ^{1}H NMR data. X.-X.L. and F.-H.L. realized insertion and verification of unnatural amino acids. Z.-L. Z. and Q.-T.H. Performed crystallization and structure solution and analysis of the synthetase and TMSiPhe decorated GFP. Q.L., X.-Y.W., F.Y., P. X., S.-M.H., S.-C.G. and M.-J.H. expressed and purified GPCR proteins. Q.-T.H. carried out Superdex Exclusion Chromatography . Q.-T. H.and C.-X.Q. supplied critical antibody Fab 30. Q.L., Z. X., and S.-L.S. Synthesized TMSiPhe. Prof. X.C. synthesized the compound BI-167107. Z. Y. performed BRET assay. Z.-Y.Y. performed GST pulldown assay. Q.L. performed fluorescence spectroscopy assay. Q.L., Z. G. and J.-Y. L. performed FSEC analysis for GPCR-arrestin complex (Data not shown). A.W. K., K.-H.X. and J.-P.S. decorated the GFP on the arrestin and developed the FSEC methods. K. R., X.-G.N. and C.-W. J. participated in the design and explanation of the NMR results. W. K., X.Y. supervised the fluorescence and BRET experiments.

J.-P.S., J.Y.W. and X.Y. supervised the overall project design and execution. J.-P.S. participated in data analysis and interpretation. Professor K.K.Z. and K.R. offered important advice and help especially in NMR experiment. J.-P.S. and J.Y. W wrote the manuscript. All of the authors have seen and commented on the manuscript.

Competing interests: The authors declare that they have no competing interests.

References

1. Zhang H & van Ingen H (2016) Isotope-labeling strategies for solution NMR studies of macromolecular assemblies. *Curr Opin Struct Biol* 38:75-82.
2. Didenko T, Liu JJ, Horst R, Stevens RC, & Wuthrich K (2013) Fluorine-19 NMR of integral membrane proteins illustrated with studies of GPCRs. *Curr Opin Struct Biol* 23(5):740-747.
3. Cattani G, Vogeley L, & Crowley PB (2015) Structure of a PEGylated protein reveals a highly porous double-helical assembly. *Nat Chem* 7(10):823-828.
4. Manglik A*, et al.* (2015) Structural Insights into the Dynamic Process of beta2-Adrenergic Receptor Signaling. *Cell* 161(5):1101-1111.
5. Yang F*, et al.* (2015) Phospho-selective mechanisms of arrestin conformations and functions revealed by unnatural amino acid incorporation and (19)F-NMR. *Nat Commun* 6:8202.
6. Yang F*, et al.* (2018) Allosteric mechanisms underlie GPCR signaling to SH3-domain proteins through arrestin. *Nat Chem Biol* 14(9):876-886.
7. Wang L, Brock A, Herberich B, & Schultz PG (2001) Expanding the genetic code of Escherichia coli. *Science* 292(5516):498-500.
8. Neumann-Staubitz P & Neumann H (2016) The use of unnatural amino acids to study and engineer protein function. *Curr Opin Struct Biol* 38:119-128.
9. Debelouchina GT & Muir TW (2017) A molecular engineering toolbox for the structural biologist. *Q Rev Biophys* 50:e7.
10. Smith JS, Lefkowitz RJ, & Rajagopal S (2018) Biased signalling: from simple switches to allosteric microprocessors. *Nat Rev Drug Discov* 17(4):243-260.
11. Sente A*, et al.* (2018) Molecular mechanism of modulating arrestin conformation by GPCR phosphorylation. *Nat Struct Mol Biol* 25(6):538-545.
12. Latorraca NR*, et al.* (2018) Molecular mechanism of GPCR-mediated arrestin activation. *Nature* 557(7705):452-456.
13. Eichel K*, et al.* (2018) Catalytic activation of beta-arrestin by GPCRs. *Nature* 557(7705):381-386.
14. Wang HM*, et al.* (2014) A stress response pathway in mice upregulates somatostatin level and transcription in pancreatic delta cells through Gs and beta-arrestin 1. *Diabetologia* 57(9):1899-1910.
15. Ning SL*, et al.* (2015) Different downstream signalling of CCK1 receptors regulates distinct functions of CCK in pancreatic beta cells. *Br J Pharmacol* 172(21):5050-5067.
16. Liu CH*, et al.* (2017) Arrestin-biased AT1R agonism induces acute catecholamine secretion through TRPC3 coupling. *Nat Commun* 8:14335.
17. Li F*, et al.* (2013) A genetically encoded 19F NMR probe for tyrosine phosphorylation. *Angew Chem Int Ed Engl* 52(14):3958-3962.
18. Chin JW (2017) Expanding and reprogramming the genetic code. *Nature* 550(7674):53-60.
19. Frankel M, Gertner D, Shenhar A, & Zilkha AJJotCS (1963) 971. Synthesis of DL-p-trimethylsilylphenylalanine.5049-5051.

608 20. Xie J & Schultz PG (2005) An expanding genetic code. *Methods* 36(3):227-238.
609 21. Wang J*, et al.* (2010) A biosynthetic route to photoclick chemistry on proteins. *J Am*
610 *Chem Soc* 132(42):14812-14818.
611 22. Peterson YK & Luttrell LM (2017) The Diverse Roles of Arrestin Scaffolds in G
612 Protein-Coupled Receptor Signaling. *Pharmacol Rev* 69(3):256-297.
613 23. Cahill TJ, 3rd*, et al.* (2017) Distinct conformations of GPCR-beta-arrestin complexes
614 mediate desensitization, signaling, and endocytosis. *Proc Natl Acad Sci U S A*
615 114(10):2562-2567.
616 24. Rahmeh R*, et al.* (2012) Structural insights into biased G protein-coupled receptor
617 signaling revealed by fluorescence spectroscopy. *Proc Natl Acad Sci U S A*
618 109(17):6733-6738.
619 25. Yang Z*, et al.* (2017) Phosphorylation of G Protein-Coupled Receptors: From the
620 Barcode Hypothesis to the Flute Model. *Mol Pharmacol* 92(3):201-210.
621 26. Shukla AK*, et al.* (2014) Visualization of arrestin recruitment by a G-protein-coupled
622 receptor. *Nature* 512(7513):218-222.
623 27. Liu JJ, Horst R, Katritch V, Stevens RC, & Wuthrich K (2012) Biased signaling
624 pathways in beta2-adrenergic receptor characterized by 19F-NMR. *Science*
625 335(6072):1106-1110.
626 28. Shukla AK*, et al.* (2013) Structure of active beta-arrestin-1 bound to a
627 G-protein-coupled receptor phosphopeptide. *Nature* 497(7447):137-141.
628 29. Kang Y*, et al.* (2015) Crystal structure of rhodopsin bound to arrestin by femtosecond
629 X-ray laser. *Nature* 523(7562):561-567.
630 30. Kumari P*, et al.* (2016) Functional competence of a partially engaged
631 GPCR-beta-arrestin complex. *Nat Commun* 7:13416.
632 31. Szczepek M*, et al.* (2014) Crystal structure of a common GPCR-binding interface for
633 G protein and arrestin. *Nat Commun* 5:4801.
634 32. Liang YL*, et al.* (2017) Phase-plate cryo-EM structure of a class B GPCR-G-protein
635 complex. *Nature* 546(7656):118-123.
636 33. Draper-Joyce CJ*, et al.* (2018) Structure of the adenosine-bound human adenosine A1
637 receptor-Gi complex. *Nature* 558(7711):559-563.
638 34. Koehl A*, et al.* (2018) Structure of the micro-opioid receptor-Gi protein complex.
639 *Nature* 558(7711):547-552.
640 35. Xu TR*, et al.* (2008) Mutations of beta-arrestin 2 that limit self-association also
641 interfere with interactions with the beta2-adrenoceptor and the ERK1/2 MAPKs:
642 implications for beta2-adrenoceptor signalling via the ERK1/2 MAPKs. *Biochem J*
643 413(1):51-60.
644 36. Ye L, Van Eps N, Zimmer M, Ernst OP, & Prosser RS (2016) Activation of the A2A
645 adenosine G-protein-coupled receptor by conformational selection. *Nature*
646 533(7602):265-268.
647 37. Brem J*, et al.* (2014) Rhodanine hydrolysis leads to potent thioenolate mediated
648 metallo-beta-lactamase inhibition. *Nat Chem* 6(12):1084-1090.
649 38. Loh CT, Adams LA, Graham B, & Otting G (2018) Genetically encoded amino acids
650 with tert-butyl and trimethylsilyl groups for site-selective studies of proteins by NMR
651 spectroscopy. *J Biomol NMR* 71(4):287-293.

39. Kan SB, Lewis RD, Chen K, & Arnold FH (2016) Directed evolution of cytochrome c for carbon-silicon bond formation: Bringing silicon to life. *Science* 354(6315):1048-1051.
40. Tacke R (1999) Milestones in the Biochemistry of Silicon: From Basic Research to Biotechnological Applications. *Angew Chem Int Ed Engl* 38(20):3015-3018.
41. Franz AK & Wilson SO (2013) Organosilicon molecules with medicinal applications. *J Med Chem* 56(2):388-405.
42. Remond E, Martin C, Martinez J, & Cavelier F (2016) Silicon-Containing Amino Acids: Synthetic Aspects, Conformational Studies, and Applications to Bioactive Peptides. *Chem Rev* 116(19):11654-11684.
43. Choi M*, et al.* (2018) G protein-coupled receptor kinases (GRKs) orchestrate biased agonism at the beta2-adrenergic receptor. *Sci Signal* 11(544).
44. Komolov KE*, et al.* (2017) Structural and Functional Analysis of a beta2-Adrenergic Receptor Complex with GRK5. *Cell* 169(3):407-421 e416.
45. Kenkichi, N.計測と制御 額 J (1962) 核磁気共鳴-とくに分析化学への応用. *Journal of the Society of & Instrument & Control Engineers* 1(2):89-94.
46. Arnold MR, Kremer W, Ludemann HD, & Kalbitzer HR (2002) 1H-NMR parameters of common amino acid residues measured in aqueous solutions of the linear tetrapeptides Gly-Gly-X-Ala at pressures between 0.1 and 200 MPa. *Biophys Chem* 96(2-3):129-140.
47. Yao X*, et al.* (2006) Coupling ligand structure to specific conformational switches in the beta2-adrenoceptor. *Nat Chem Biol* 2(8):417-422.

Figure. 1. Development of DeSipher by genetic code expansion and the selectivity of TMSiPheRS

(A) The ranges of the methyl ^{1}H chemical shifts (45) (shown by bidirectional arrows) and the distribution of random-coil aliphatic CH ^{1}H chemical shifts for the 20 genetically coded amino acids (46). The ^{1}H chemical shifts of methyl silicon group are specified in red.

(B) Coomassie-stained gel analysis of full-length β-arrestin-1 expression in *E. coli* cells that were cotransfected with the β-arrestin-1-H295TAG plasmid and the pEVOL-TMSiPheRS plasmid, encoding a specific *M. jannaschii* tyrosyl amber suppressor tRNA/tyrosyl-tRNA synthetase mutant grown in the presence or absence of different silicon-containing compounds. WT indicates wild-type arrestin without any change in genetic code. Full-length β-arrestin-1 protein was obtained only in the presence of TMSiPhe for TAG mutation of β-arrestin-1 or WT. These results suggested that the evolved TMSiPheRS exhibited significant structural selectivity for TMSiPhe over other silicon-containing chemicals. The chemical abbreviations are as follows:

(1) 4-(trimethylsilyl) phenylalanine, TMSiPhe;

(2) 3,5-dichloro-4-[(trimethylsilyl) methoxy]phenylalanine, TMSiM-dcTy;

(3) 2-amino-3-((trimethylsilyl)methylthio)propanoic acid, TMSiM-Cys;

(4) 2-amino-4-((trimethylsilyl)methylthio)butanoic acid, TMSiM-hCys;

(5) 4-[(trimethylsilyl)ethoxy]phenylalanine, TMSiM-Tyr;

(6) control, Ctl. There were no unnatural amino acids added to the culture.

(C) Schematic flowchart for the incorporation of TMSiPhe into β-arrestin-1 at the H295 site. Full-length β-arrestin-1 protein was obtained by cotransfection with the β-arrestin-1 H295 TAG mutant plasmid and the pEVOL-TMSiPheRS plasmid, with TMSiPhe supplementation of the culture medium. The purity of the protein was determined by electrophoresis. The protein was subjected to trypsin digestion and analyzed by MS/MS. These results unambiguously confirmed that TMSiPhe was selectively incorporated into β-arrestin-1 at the H295 position. m/z, mass/charge ratio.

(**D**) The 2Fo-Fc annealing omit map of sfGFP-Y182-TMSiPhe clearly shows the electron density of TMSiPhe. The map was contoured at 1.1 σ.

(**E**) 1D ^{1}H NMR spectra for the β-arrestin-1 H295 TMSiPhe mutant were compared with those for wild-type β-arrestin-1 cultured in the presence of TMSiPhe. The spectra were recorded in a buffer containing 50 mM Tris-HCl (pH 7.5) and 150 mM NaCl at 25°C using a Bruker 950 MHz NMR spectrometer. The β-arrestin-1 H295 TMSiPhe chemical shift at 0.26 ppm was consistent with the predicted chemical shift of the TMSi group. The ^{1}H NMR signals of methyl-silicon-substituted amino acids in a protein were generally located in the high-field region (<0.55 ppm, blue area).

Figure. 2. Structural basis for the selective recognition of TMSiPhe by TMSiPheRS

(**A**) Binding of TMSiPhe at the active site of TMSiPheRS. The 2Fo-Fc annealing omit electron density map of TMSiPhe was contoured at 1.0 σ.

(**B**) Comparison of the unnatural amino acid-binding pockets between TMSiPheRS (red dotted line) and the wild-type *M*j-TyrRS (black dotted line, PDB:1J1U). Five key mutations, indicated by arrows, increased the size of the TMSiPhe-binding pocket substantially.

(**C**) Interactions between TMSiPhe and TMSiPheRS. The specific interactions include hydrogen bonds (blue dotted line), π-cation interactions (red dotted line), ion-dipole interactions (magenta dotted line) and hydrophobic interactions with surrounding residues (left panel).

(**D**) The H32 residue in the β2 sheet was rotated approximately 120 degrees in the TMSiPhe/TMSiPheRS complex (green) compared with TMSiPheRS alone (magenta), leading to favorable charged interactions with TMSiPhe.

Figure. 3. Incorporation of TMSiPhe at different functionally relevant motifs of β-arrestin-1 and characterization of β-arrestin-1 activation by DeSipher.

(**A**) Frontal view of the TMSiPhe incorporation sites depicted by spheres in the inactive

β-arrestin-1 crystal structure (PDB: 1G4M). Orange, Y21 in the three elements; purple, Y63 in the finger loop; blue, Y173 in the hinge region; cyan, Y249 in β-sheet XVI; pink, R285, green, H295 in the lariat loop; red, F388 in the C-terminal swapping segment.

(**B**) 1D ^{1}H NMR spectra of β-arrestin-1 labeled as described in (3a). The spectra were recorded in a buffer containing 50 mM Tris-HCl (pH 7.5 and 150 mM NaCl at 25°C using a Bruker 950 MHz NMR spectrometer. The protein concentrations were 5~15 μM, and the total recording time per spectrum was 6~15 min. The chemical shift for the TMSiPhe protein was less than 0.55 ppm.

(**C**) Cartoon illustration of the activation of β-arrestin-1 and movement of the C-terminal swapping segment of β-arrestin-1. In response to the binding of an activator, such as the phospho-vasopressin receptor C-tail (V2Rpp), the originally embedded C-terminal swapping segment of β-arrestin-1 became highly solvent exposed, thus favoring binding to downstream signaling proteins, for example, clathrin or AP2 (adaptor protein 2). This conformational transition could be monitored by incorporation of TMSiPhe at the F388 position of β-arrestin-1, which is located in the C-terminal swapping segment.

(**D**) 1D ^{1}H NMR spectra of β-arrestin-1–F388-TMSiPhe in response to titration with V2Rpp. Two distinct peaks were observed. The peak (-0.05 ppm) representing the inactive state gradually decreased in intensity, while the peak representing the active state (0.15 ppm) steadily increased in intensity. The spectra were recorded in a buffer containing 50 mM Tris-HCl (pH 7.5) and 150 mM NaCl at 25°C using a Bruker 950 MHz NMR spectrometer.

(**E**) Analysis of the titration experiments monitored by 1D ^{1}H NMR spectroscopy of β-arrestin-1 F388TMSiPhe (3d). The curve was fitted to the nonlinear regression equation $y=B_{max}[X]/(K_D+[X]$, according to the scatchard plot analysis (fig. S9). The (R^2=0.99). The K_D value was calculated at 6.9±0.2 μM (R^2=0.99).

Figure. 4. Regulation of the conformational changes of the β-arrestin-1 polar core by different β2AR ligands.

(**A**) Cartoon illustration of two distinct interaction modes between GPCRs and β-arrestin-1 (hanging mode and snug mode). The phosphorylated C-tail and transmembrane core of the receptor were each able to independently stimulate arrestin activation. After activation, the N- and C-domains of β-arrestin-1 underwent approximately 20 degrees of rotation. The polar core of β-arrestin-1 is hypothesized to be a critical stabilizer of the inactive state, and introduction of the probe at H295, which is close to this key region, enables detection of conformational changes in the polar core in response to the GPCR activation.

(**B**). Structural comparison of the H295 position in inactive β-arrestin-1 (PDB: 1G4M), the V2Rpp/ β-arrestin-1 complex (PDB:4JQI) and the rhodopsin/arrestin complex (PDB: 5W0P).

The inactive β-arrestin-1 structure is depicted in gray; the V2Rpp/β-arrestin-1 complex is in green; and the rhodopsin-arrestin complex is in red. The two active arrestins have similar conformations at the H295 position, differing significantly from the pose in the inactive arrestin structure.

(**C**) 1D ^{1}H NMR spectra of β-arrestin-1–H295-TMSiPhe in response to titration with V2Rpp. Two distinct peaks were observed. With increasing concentrations of V2Rpp, the peak at 0.25 ppm decreased (representing the S1 state), whereas a new growing peak was observed at 0.15 ppm (representing the S2 state).

(**D**) 1D ^{1}H NMR spectra of β-arrestin-1 H295TMSiPhe alone or the ppβ2V2R/H295TMSiphe/Fab30 complex with or without different ligands and the chemical structures of the ligands used in the current study. After incubation with the phospho-β2AR-V2-tail (ppβ2V2R) and formation of the receptor-arrestin complex, a new NMR signal appeared at 0.07 ppm (designated S3), and the intensity of the S1 peak decreased. When incubated with different β2AR ligands before formation of the ppβ2V2R-β-arrestin-1/Fab30 complex, the S3 state signal intensity of the complex was positively correlated with effects of the ligands on the activation of downstream effectors, such as arrestin. BI, BI-167107; ISO, Isopreteronol; Clen, Clenbuterol; Salm, Salmeterol; Carv, Carvedilol, Alp, Alprenolol; ICI, ICI-118551. The buffer used for the experiment contained 20 mM HEPES, 150 mM NaCl, 0.01% LMNG, 0.002% CHS, and 10% D2O (pH 7.5 at 25°C).

(**E**) Best-fit linear correlation of the peak area representing the amount of the S3 state in the presence of different ligands, with the ligand efficacy for receptor internalization from the BRET experiment in vivo. See fig. S17 for details.

(**F**) Best-fit linear correlation of the peak area representing the amount of the S3 state in the presence of different ligands, with the ligand efficacy for separation of the receptor transmembrane III and VI from the TRIQ experiment in vitro (47). See fig. S18 for details.

Figure. 5. Gα-CT competition experiments confirmed the essential role of the 7TM core of the receptor in mediating the ligand-regulated conformational change in β-arrestin-1.

(**A**) Schematic diagram of the interaction between GPCRs and β-arrestin-1 in the presence of excess Gα C-terminus (Gα-CT), which has been described in previous reports (31). The interaction between β-arrestin-1 and the GPCR TM core was abolished via steric hindrance by Gα-CT. β-Arrestin-1 still interacts with the phosphorylated GPCR C-terminal tail and thus forms a complex with the receptor.

(**B**) ISO/ppβ2V2R/β-arrestin-1/Fab30 complex formation was not disrupted by Gα-CT in a size-exclusion assay. The similar SEC profile observed with or without Gα-CT suggests that Gα-CT did not disrupt the ISO/ppβ2V2R/β-arrestin-1/Fab30 complex. Size-exclusion

803 chromatography experiments were performed on an AKTA Purifier equipped with a Superdex
804 200 (10/300GL) column. Black: ISO/ppβ2V2R/β-arrestin-1-H295TMSiPhe/Fab30 complex,
805 green: ISO/ppβ2V2R/β-arrestin-1-H295TMSiPhe/Fab30 complex mixed with the 200 μM
806 Gα-CT.
807 (**C**) 1D ^{1}H NMR spectra of β-arrestin-1 H295TMSiPhe in the presence of Gα-CT. The
808 transformation from S1 to S3 induced by the ISO/ppβ2V2R/β-arrestin-1/Fab30 complex was
809 significantly weakened by the addition of Gα-CT, suggesting the observed S3 state reduction
810 was mainly due to the elimination of the receptor core interaction with β-arrestin-1 by the
811 binding of Gα-CT.
812 (**D**) Schematic diagram of the competing experiments of the ppβ2V2R/β-arrestin-1 complex
813 disrupted by the presence of excess V2Rpp. β-Arrestin-1 dissociated from the phosphorylated
814 β2V2R due to competition with V2Rpp.
815 (**E**) ISO/ppβ2V2R/β-arrestin-1/Fab30 complex formation was disrupted by incubation with
816 excess V2Rpp in a size-exclusion assay. Black:
817 ISO/ppβ2V2R/β-arrestin-1-H295TMSiPhe/Fab30. Red:
818 ISO/ppβ2V2R/β-arrestin-1-H295TMSiPhe/Fab30 complex mixed with 200 μM V2Rpp. The
819 red peak can be simulated by two components, which are as follows: the blue curve represents
820 the SEC of ppβ2V2R alone, and the pink curve represents the
821 V2Rpp/β-arrestin-1-H295TMSiPhe complex.
822 (**F**) 1D ^{1}H NMR spectra of β-arrestin-1 H295TMSiPhe in the presence of V2Rpp. The
823 incubation of the V2Rpp with the ISO/ppβ2V2R/β-arrestin-1-H295TMSiPhe/Fab30 complex
824 caused no significant change in the S3 state, but caused a shift of the S1 state to the S2 state.
825 c or f) The data collection buffer used for the experiments contained 20 mM HEPES, 150 mM
826 NaCl, 0.01% LMNG, 0.002 CHS, and 10% D2O (pH 7.5 at 25°C).
827
828 **Figure. 6. Monitoring the conformational states of site 285 of β-arrestin-1 in response to**
829 **different β2AR ligands.**
830
831 (**A**). Structural comparison of the R285 position in inactive β-arrestin-1 (PDB: 1G4M) and the
832 corresponding R291 position in the rhodopsin/arrestin complex (PDB: 5W0P). The active
833 β-arrestin-1 structure is depicted in gray, and the rhodopsin/arrestin complex is in red. The
834 activation of arrestin by a receptor led to a highly solvent-exposed configuration at the R285
835 position of β-arrestin-1, as suggested by the crystal structures.
836 (**B**) 1D ^{1}H NMR spectra of β-arrestin-1 R285TMSiPhe activated by ppβ2V2R with or without
837 different ligands. After incubation with ppβ2V2R, multiple new NMR signals appeared
838 between 0.04 ppm to 0.10 ppm, which are designated as R0 (0.09ppm), R1a (0.065ppm), R1b
839 (0.068 ppm), R1c (0.06 ppm), R2 (0.05 ppm), from low field to high field. The buffer used for

840 the experiment contained 20 mM HEPES, 150 mM NaCl, 0.01% LMNG, 0.002 CHS, and 10%
841 D2O (pH 7.5 at 25°C).
842 (**C**) Bar graph representing the population (simulated peak area) of each NMR peak for each
843 ligand condition. The values are also tabulated in fig. S21.

Fig 1

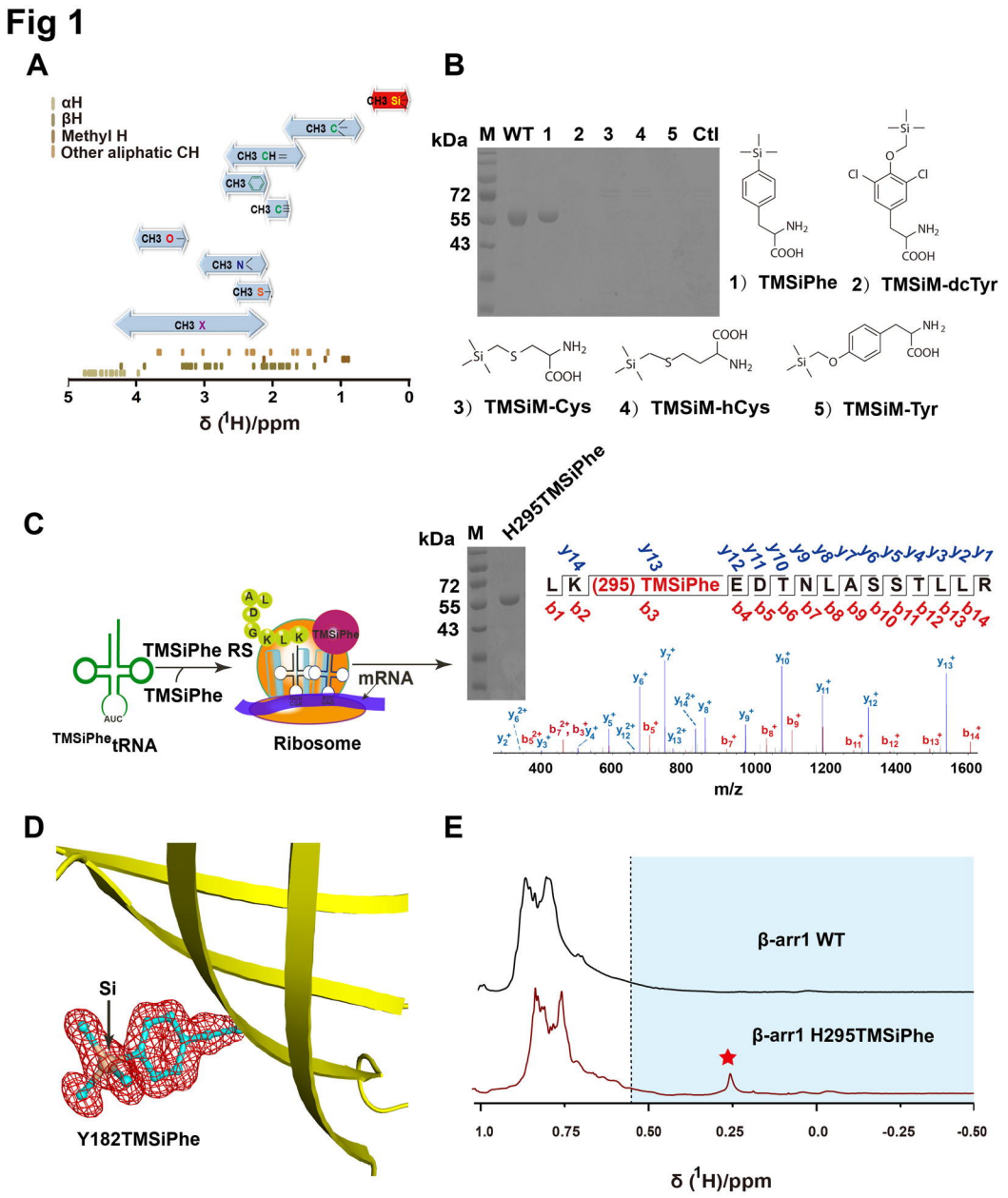

A
αH
βH
Methyl H
Other aliphatic CH
CH3 Si
CH3 C
CH3 CH=
CH3
CH3 C≡
CH3 O
CH3 N
CH3 S
CH3 X
5
4
3
2
1
0
δ (1H)/ppm
B
kDa
M WT 1 2 3 4 5 Ctl
72
55
43
1) TMSiPhe
2) TMSiM-dcTyr
3) TMSiM-Cys
4) TMSiM-hCys
5) TMSiM-Tyr
C
TMSiPhe RS
TMSiPhe
TMSiPhe tRNA
Ribosome
mRNA
kDa
M
H295TMSiPhe
72
55
43
L K (295) TMSiPhe E D T N L A S S T L L R
400
600
800
1000
1200
1400
1600
m/z
D
Si
Y182TMSiPhe
E
β-arr1 WT
β-arr1 H295TMSiPhe
1.0
0.75
0.50
0.25
0.0
-0.25
-0.50
δ (1H)/ppm

Fig 2

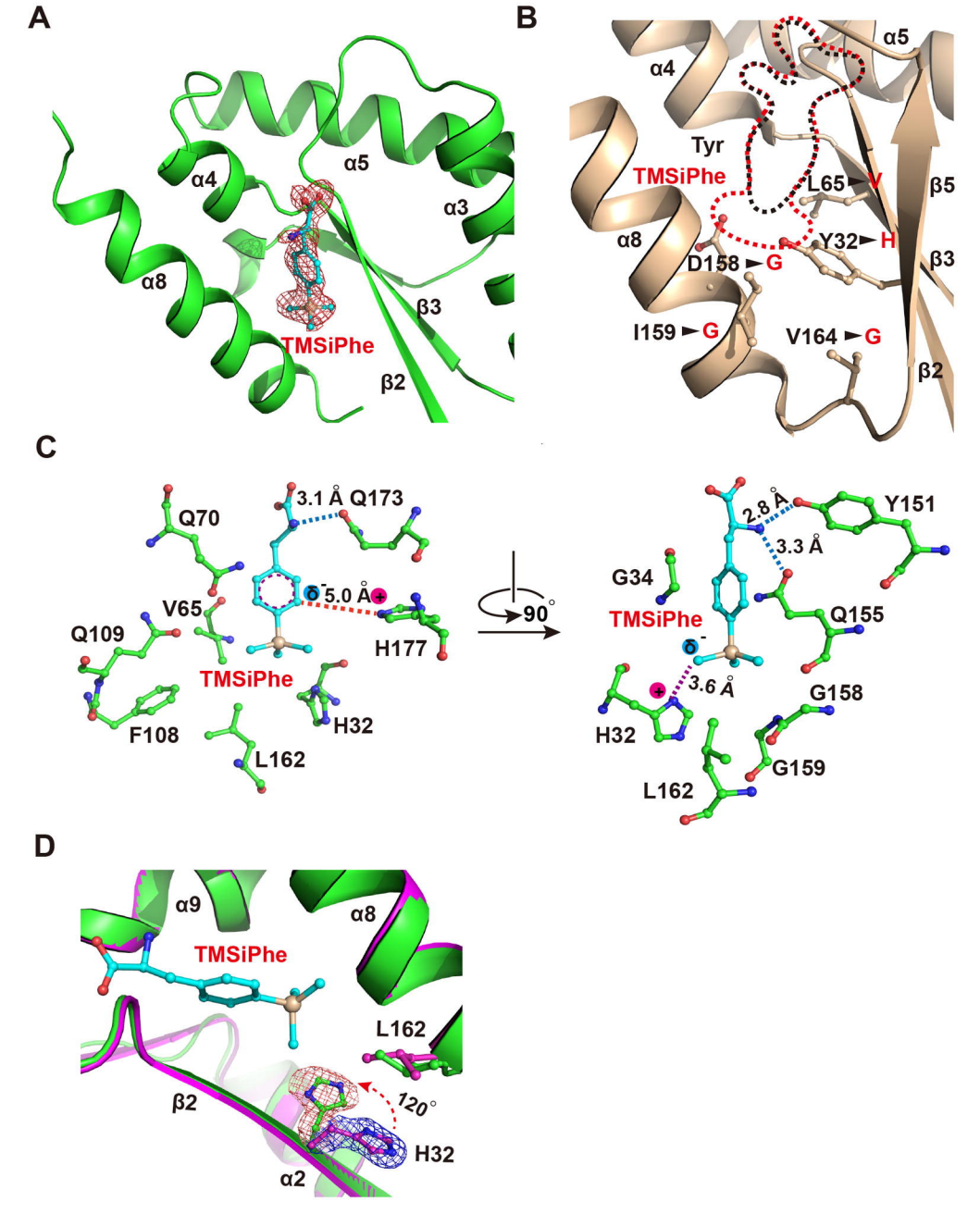

A
α4
α5
α3
α8
β3
TMSiPhe
β2
B
α5
α4
Tyr
TMSiPhe
L65 ► V
β5
α8
Y32 ► H
D158 ► G
β3
I159 ► G
V164 ► G
β2
C
3.1 Å Q173
Q70
δ- 5.0 Å +
V65
Q109
H177
TMSiPhe
F108
H32
L162
90°
2.8 Å
Y151
3.3 Å
G34
TMSiPhe
Q155
δ-
3.6 Å
G158
H32
G159
L162
D
α9
α8
TMSiPhe
L162
β2
120°
H32
α2

Fig 3

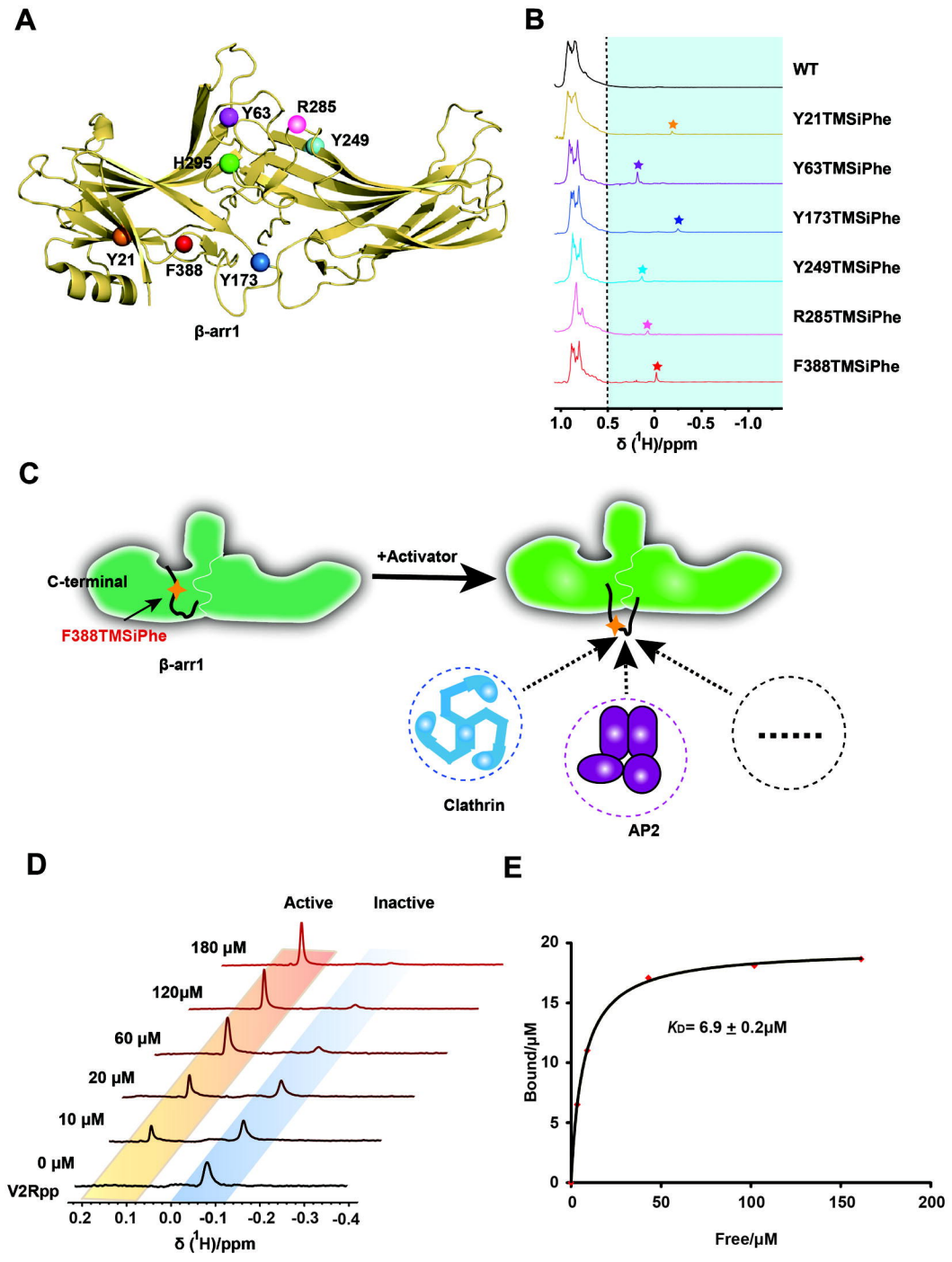

A
Y63
R285
Y249
H295
Y21
F388
Y173
β-arr1
B
WT
Y21TMSiPhe
Y63TMSiPhe
Y173TMSiPhe
Y249TMSiPhe
R285TMSiPhe
F388TMSiPhe
1.0 0.5 0 -0.5 -1.0
δ (1H)/ppm
C
C-terminal
F388TMSiPhe
β-arr1
+Activator
Clathrin
AP2
D
Active
Inactive
180 μM
120μM
60 μM
20 μM
10 μM
0 μM
V2Rpp
0.2 0.1 0.0 -0.1 -0.2 -0.3 -0.4
δ (1H)/ppm
E
Bound/μM
KD= 6.9 ± 0.2μM
0 5 10 15 20
0 50 100 150 200
Free/μM

Fig 4

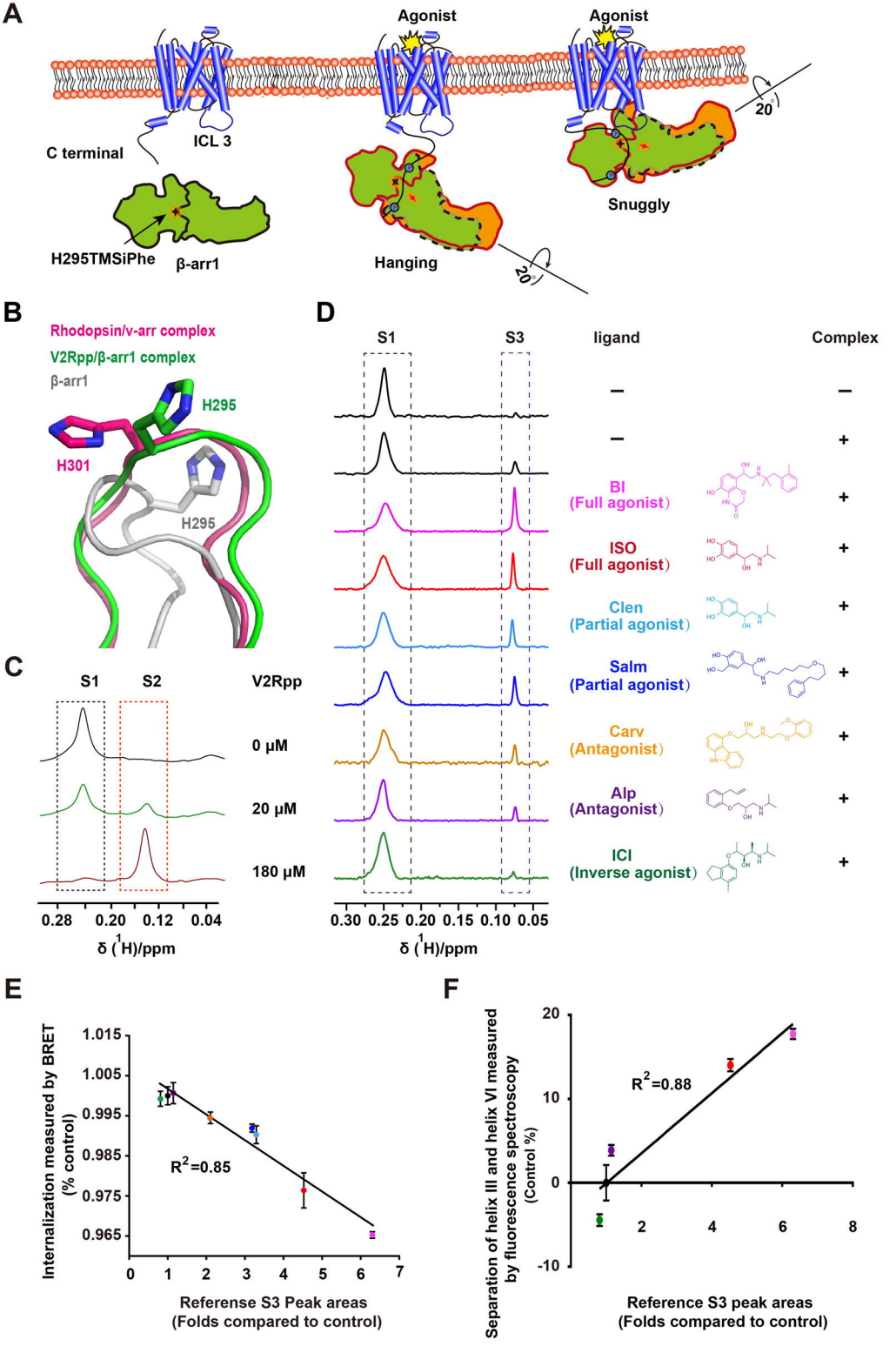

A
Agonist
Agonist
C terminal
ICL 3
20°
Snuggly
H295TMSiPhe
β-arr1
Hanging
20°
B
Rhodopsin/v-arr complex
V2Rpp/β-arr1 complex
β-arr1
H295
H301
H295
C
S1
S2
V2Rpp
0 μM
20 μM
180 μM
0.28
0.20
0.12
0.04
δ (1H)/ppm
D
S1
S3
ligand
Complex
BI
(Full agonist)
ISO
(Full agonist)
Clen
(Partial agonist)
Salm
(Partial agonist)
Carv
(Antagonist)
Alp
(Antagonist)
ICI
(Inverse agonist)
0.30
0.25
0.20
0.15
0.10
0.05
δ (1H)/ppm
E
Internalization measured by BRET
(% control)
1.015
1.005
0.995
0.985
0.975
0.965
R2=0.85
0
1
2
3
4
5
6
7
Referense S3 Peak areas
(Folds compared to control)
F
Separation of helix III and helix VI measured
by fluorescence spectroscopy
(Control %)
20
10
0
-10
R2=0.88
2
4
6
8
Reference S3 peak areas
(Folds compared to control)

Fig 5

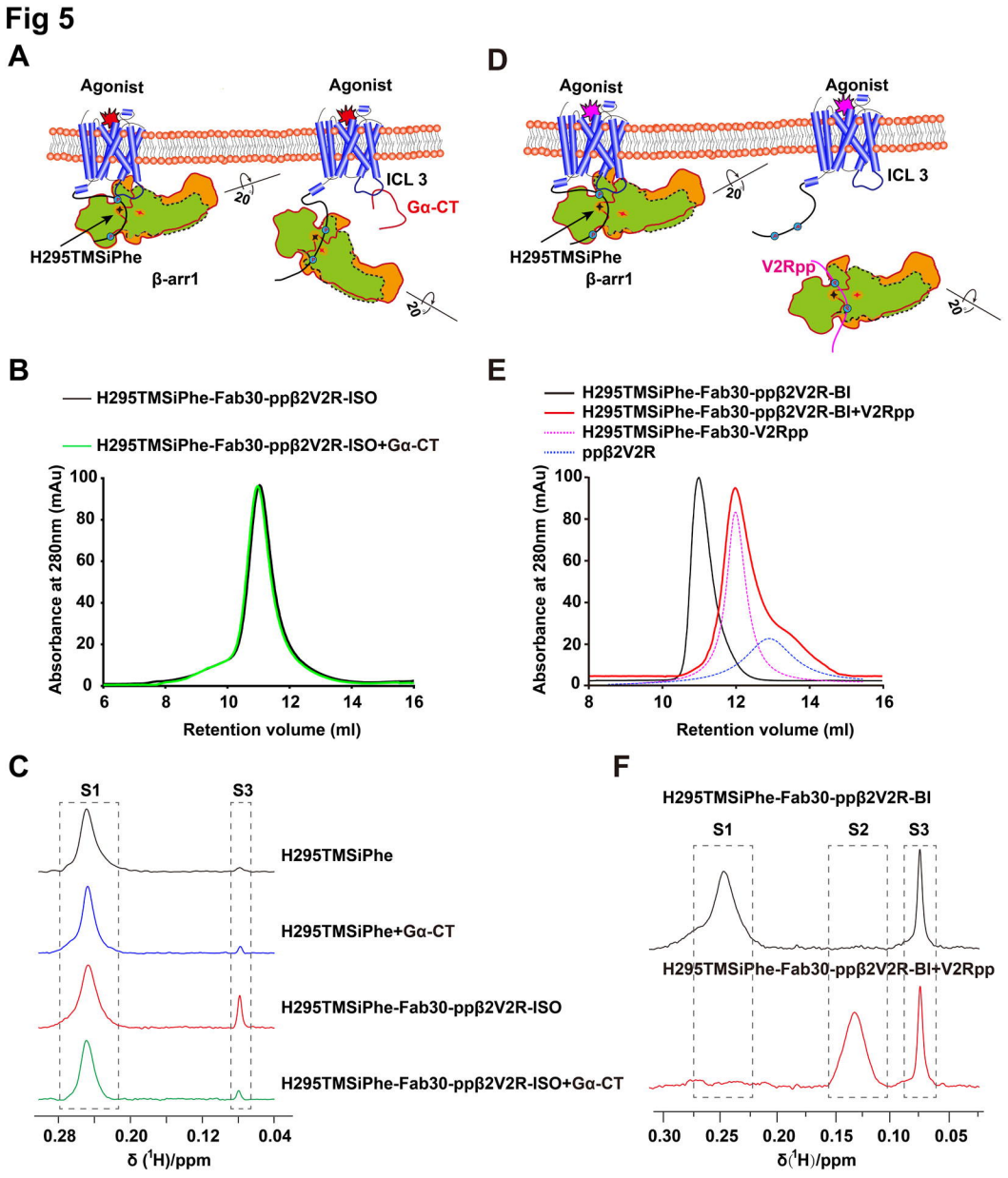
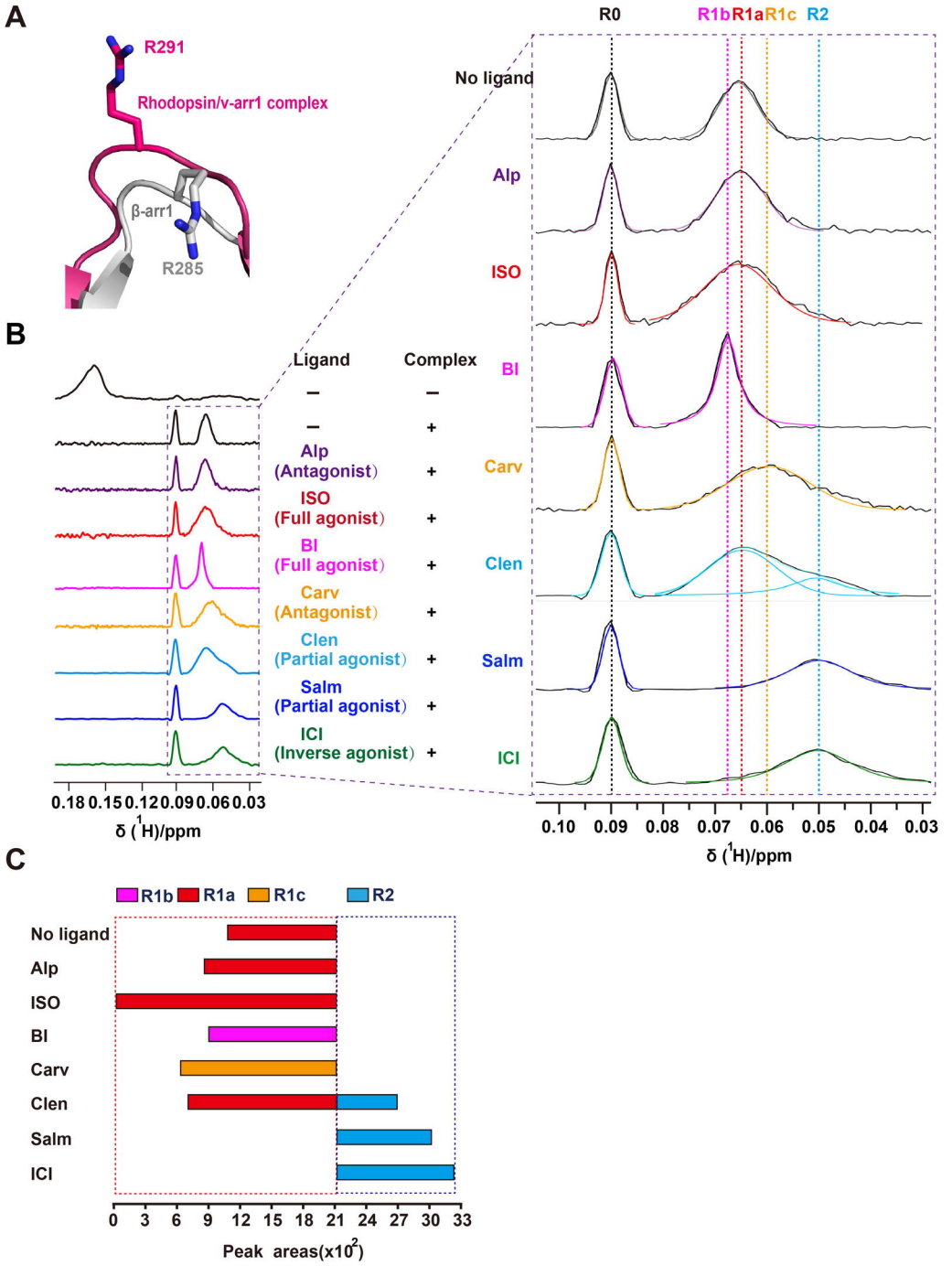

A
Agonist
Agonist
ICL 3
Gα-CT
20°
H295TMSiPhe
β-arr1
20°
D
Agonist
Agonist
ICL 3
20°
H295TMSiPhe
β-arr1
V2Rpp
20°
B
H295TMSiPhe-Fab30-ppβ2V2R-ISO
H295TMSiPhe-Fab30-ppβ2V2R-ISO+Gα-CT
Absorbance at 280nm (mAu)
Retention volume (ml)
E
H295TMSiPhe-Fab30-ppβ2V2R-BI
H295TMSiPhe-Fab30-ppβ2V2R-BI+V2Rpp
H295TMSiPhe-Fab30-V2Rpp
ppβ2V2R
Absorbance at 280nm (mAu)
Retention volume (ml)
C
S1
S3
H295TMSiPhe
H295TMSiPhe+Gα-CT
H295TMSiPhe-Fab30-ppβ2V2R-ISO
H295TMSiPhe-Fab30-ppβ2V2R-ISO+Gα-CT
δ (1H)/ppm
F
H295TMSiPhe-Fab30-ppβ2V2R-BI
S1
S2
S3
H295TMSiPhe-Fab30-ppβ2V2R-BI+V2Rpp
δ(1H)/ppm

Fig 6


A
R291
Rhodopsin/v-arr1 complex
β-arr1
R285
B
Ligand
Complex
–
–
–
+
Alp
(Antagonist)
+
ISO
(Full agonist)
+
BI
(Full agonist)
+
Carv
(Antagonist)
+
Clen
(Partial agonist)
+
Salm
(Partial agonist)
+
ICI
(Inverse agonist)
+
0.18 0.15 0.12 0.09 0.06 0.03
δ (1H)/ppm
R0
R1b R1a R1c
R2
No ligand
Alp
ISO
BI
Carv
Clen
Salm
ICI
0.10 0.09 0.08 0.07 0.06 0.05 0.04 0.03
δ (1H)/ppm
C
R1b R1a R1c
R2
No ligand
Alp
ISO
BI
Carv
Clen
Salm
ICI
0 3 6 9 12 15 18 21 24 27 30 33
Peak areas(x10^2)