Electronic Sculpting of ligand-GPCR subtype selectivity: the case of angiotensin II

Francesca Magnani, †Charalampos G. Pappas, ‡ Tim Crook, § Vassiliki Magafa, † Paul Cordopatis, ‡ Susumu Ishiguro, † Naomi Ohta, ‡ Jana Selent, ¥ Sanja Bosnyak, ‡ Emma S Jones, ‡ Ioannis P. Gerothanassis, ‡ Masaaki Tamura, § Robert E Widdop, # Andreas G. Tzakos‡†

† Medical Research council, Laboratory of Molecular Biology, Cambridge, UK (present address: Biological Crystallography Lab, Department of Biology and Biotechnology, University of Pavia, Italy),

‡ Department of chemistry, University of Ioannina, GR 45110, Greece, § Division of Cancer research, University of Dundee, Dundee, DD1 9SY, UK,

† Department of Pharmacy, University of Patras, Greece, ‡ Department of Anatomy & Physiology, Kansas State University College of Veterinary Medicine, Manhattan, Kansas 66506, USA,

¥ Computer-Assisted Drug Design Lab, research Programme on Biomedical Informatics (GRIB), Universitat Pompeu Fabra, Barcelona, Spain,

# Department of Pharmacology, Monash University, Clayton Victoria 3800, Australia

Contents

Supplementary Methods 2
Supplementary Tables 6
Supplementary Figures 14
Supplementary References 22
Supplementary Methods

Materials. AT2R receptor-specific blocker PD123319 was purchased from Sigma-Aldrich Chemical Co. (St. Louis, MO). Human AGTR2 pcDNA3.1+ was obtained from the UMR cDNA Resource Center (University of Missouri-Rolla, Rolla, MO). All other chemicals were of analytic grade. The AT1aR and AT2R constructs were a kind gift from Lazlo Hunyady (Semmelweis University, Budapest, Hungary).

Mutagenesis. AT2R mutants were generated as described elsewhere.

Cell culture and transient transfection. HEK-293T cells were maintained at 37 °C in Dulbecco’s modified Eagle’s medium (DMEM) supplemented with 10% (v/v) fetal bovine serum (FBS), 2 mM glutamine, 100 U/ml penicillin, and 100 µg/ml streptomycin. Cells were transfected with either AT1R (pcDNA3.1), AT2R (pcDNAI/Amp) or AT2R variants using GeneJuice transfection reagent according to the manufacturer’s instructions. Cells were harvested 48 hours after transfection, crude membranes were prepared and resuspended in binding buffer (50 mM Tris-HCl pH 7.4, 100 mM NaCl, 10 mM MgCl₂, 1mM EDTA, 0.2% BSA and 0.025% Bacitracin, protease inhibitors) at a protein concentration of 3.3 µg/µl as determined by the amido black protein assay. For whole cells assays, HEK-293 cells stably transfected with either AT₁ R or the AT₂ R (kindly supplied by Professor W. Thomas, University of Queensland) were used. These cells were grown to approximately 80% confluence and were then re-plated into 48 well plates at 1x10⁵ cells/well and grown for 48 hours at 37°C for a whole cell competition binding assay, as described previously. Minor differences between the binding assays performed on isolated membrane preparation and whole cells are probably due to the loss of G-protein in the former.

PC12W cells, a substrain from a clonal isolation of a rat adrenal chromaffin cell tumor, were cultured in DMEM supplemented with 10% FBS, 100 units/ml penicillin and 100µg/ml streptomycin (Invitrogen, Carlsbad, CA) as previously described. The cells were incubated in a 5% CO₂ humidified incubator at 37°C. Preparation of recombinant replication-deficient adenovirus containing the human AT2R coding region was carried out by VECTOR BIOLABS (Philadelphia, PA).

For gene transduction with adeno viral vectors cells were seeded at 1.2 x 10⁵ cells per well in a 6-well plate. After 24 hours, cells were incubated at 37°C for 6 hours in serum-free DMEM containing adeno viral vectors (Ad-AT2R 30 multiplicity of infection (MOI)), and shaken lightly every fifteen minutes. After 6 hr incubation, cells were cultured in 10% FBS containing DMEM at 37°C, 5% CO₂ for an additional 24 hours. Cells were then trypsinized and subcultured at a density of 2 x 10⁵ cells per well on a 24-well plate. For the evaluation of the effect of [Y]⁶-AII analogue on the neurite outgrowth of PC12W cells, at 24 hours after subculture, control cells and the AT2R over-expressing cells were treated with 1 nM AII or [Y]⁶-AII as indicated in Figure 3.

Radioligand binding assays. Saturation curves on cell membranes were obtained using 10 µg protein per data point and a range of [¹²⁵I]-AII (Amersham) concentration...
0-10 nM (8 data points in triplicate). Non-specific binding was determined in presence of 6 μM cold Ang II. Competition assays were performed using a concentration of \([^{[125]}\text{I}]-\text{Ang II}\) of 1 nM and various concentrations of unlabeled ligands, as indicated in the Figure 3. Samples were incubated for 2 hours at 4°C. Receptor-bound and free radioligand were separated by filtration through Whatman GF/B filters, pre-soaked with 0.3% polyethyamine. The filters were washed with 5 ml of ice-cold binding buffer and transferred to scintillation tubes. Radioactivity was counted on a Beckman LS6000 liquid scintillation counter and data were analyzed by non-linear regression using Prism software (GraphPad). \(K_i\) values were calculated according to the Cheng and Prusoff equation with a \(K_D\) for \([^{[125]}\text{I}]-\text{Ang II}\) of 1.8 nM (AT1R) and 2.3 nM (AT2R). Experiments on whole cells were performed in both AT1R- and AT2R-HEK-293 stable cell lines, as previously described in detail with minor modifications. Novel peptides and references compounds, CGP42112 and candesartan, were examined for their ability to inhibit specific binding of \([^{[125]}\text{I}]-\text{Ang II}\), and nonspecific binding was determined using 10μM cold Ang II. Following non-linear regression, IC50 values, representing the concentration at which each ligand displaced 50% binding of \([^{[125]}\text{I}]-\text{Ang II}\) in either AT1R- or AT2R- HEK-293 cell lines, were estimated. Log ratios of IC50 values for each ligand at AT1R:AT2R were determined as a measure of AT2R selectivity.

**Peptide synthesis and sample preparation.** The synthesis of the following peptides: \(\text{All} (D^-R^2-V^2-Y^4-i^3-H^6-P^7-F^8)\); \([Y]^6-\text{All}\) \((D^-R^2-V^2-Y^4-i^3-H^6-P^7-F^8)\); \([4-\text{OPO}_3\text{H}_2-F]^6-\text{All}\), \((D^-R^2-V^2-Y^4-i^3-H^6-P^7-F^8)\); \([F]^6-\text{All}\), \((D^-R^2-V^2-Y^4-i^3-H^6-P^7-F^8)\); \([4-\text{NO}_2-F]^6-\text{All}\), \((D^-R^2-V^2-Y^4-i^3-H^6-P^7-F^8)\) was achieved using Fmoc/tBu methodology. 2-Chlorotrityl chloride resin and N\(^2\)-Fmoc (9-fluorenylmethyloxycarbonyl) amino acids were used for the synthesis. Peptide purity was assessed by analytical HPLC, mass spectrometry (FABMS, ESIMS) and amino acid analysis. The samples were prepared for NMR spectroscopy by dissolving the peptide in 0.01 M KPi buffer (pH = 4), containing 0.02 M KCl. 2,2-dimethyl–2-sila-pentane sulfonate (DSS) was added to a concentration of 1 mM as an internal chemical shift reference. Peptide concentration was commonly 5 mM in 90% \(^1\text{H}_2\text{O} / 10\% \(^2\text{H}_2\text{O}. Trace amounts of NaN\(_3\) were added as a preservative.

**MCF7 breast carcinoma cell proliferation assay.** MCF7 breast carcinoma cells in log phase were treated with drug vehicle alone (which was not toxic to cells) or increasing concentrations of \([Y]^6-\text{All}\). Cells were counted daily and the effects of the analogue on proliferation expressed as cell numbers relative to controls which received only the drug vehicle (Figure S8). \([Y]^6-\text{All}\) efficiently inhibited proliferation at concentrations of 10\(^-5\) M and showed evidence of anti-proliferative effects at 10\(^-9\) M. The analogue showed an IC50 of \(~ 5 \times 10^-8\) M in MCF cells.

**NMR spectroscopy.** NMR spectra (2D \(^1\text{H}^-\text{H}^-\text{H} \text{TOCSY}\) and 2D \(^1\text{H}^-\text{H}^-\text{H} \text{NOESY}\) were acquired at 500 MHz using a Bruker Avance 500 spectrometer. For water suppression excitation sculpting with gradients was used. The mixing time for TOCSY spectra was 80 ms. Mixing times for NOESY experiments were set to 100, 200, 350 and 400 ms to determine NOE build – up rates. A mixing time of 350 ms provided sufficient cross-peak intensity without introducing spin-diffusion effects in the 2D – NOESY. Inter-
proton distances for were derived by measuring cross-peak intensities in the NOESY spectra. Structure calculations were performed with the program CYANA. The mole fraction of the peptide molecules in the cis isomeric form (Xcis) was obtained by measuring the areas of well-resolved peaks corresponding to the same proton resonance in the cis and trans forms in 1D spectra. To investigate solvent protection values for amide protons, the amide proton temperature coefficients (Δδ/ΔT) were measured in a range of temperatures from 283K to 308K. A plot of Δδ/ΔT vs. the chemical shift deviation (CSD) of the measured amide proton resonances at 283 K was constructed with appropriate random coil chemical shift correction. The dashed line (Δδ/ΔT = −7.8 (CSD) −4.4) represents the cut off of Δδ/ΔT between exposed and sequestered NHs of proteins. Gradients above the dashed line indicate exposed NHs, whereas those below indicate sequestered NHs. Diffusion ordered spectroscopy (DOSY) spectra were recorded using the Bruker microprogram stebpqps1s19 at 292 K. The eddy current delay (Te) was set to 5 ms. The diffusion time was adjusted to 100 ms. The duration of the pulse field gradient, δg, was optimized in order to obtain 5% residual signal with the maximum gradient strength with the resulting δ value of 3.6 ms. The pulse gradient was increased from 2 to 95% of the maximum gradient strength using a linear ramp 16k data points in the F2 dimension (20 ppm) and 16 data points in the F1 dimension were collected. Final data sizes were 16k×128.

**Modeling studies.** Generation of the homology model. A Blast search for the sequence of the human type-1 angiotensin II receptor (P30556) and human type-2 angiotensin II receptor (P50052) revealed as the closest template the structure of the human chemokine receptor CXCR1 (PDB ID 2LNL). However, as this structure is an Apo-form with a completely closed binding pocket and therefore not suitable for our purpose, we used instead the closely-related CXCR4 chemokine Receptor in complex with a cyclic peptide (PDB ID 3OE6). Importantly, the compounds of interest, All and the Y6 analogue, are agonists which likely bind to the active AT1R and AT2R. In order to represent well the active AT1R and AT2R, we used as a second template the closest-related active GPCR, which was found to be the beta2 adrenergic receptor (PDB ID 3SN6). Homology modeling was carried out using the MOE package. After sequence alignment (default settings) and manual corrections which ensured alignment of highly conserved residues as well as integral helices, 10 models were generated using the Amber12:EHT force field. The best models for the AT1R and AR2R were chosen for subsequent docking studies.

**Docking studies.** In order to obtain a reasonable starting structure, we used the GOLD software (Version 5.0) and docked All and the Y6-analouge into the AT1R and the AT2R applying a constraint between the alpha-carboxylate group of the Phe8 and the residues K199(AT1R)/K215(AT2R). In order to dock the peptide into the target, the binding site was defined as a large region of 30 Å around the NZ atom of Lys199. As scoring function, we applied the GoldScore. All other parameters were set to default.

**Conformational space analysis.** Subsequently, the conformational space of the obtained complexes was extensively explored using LowModeMD Search for the
ligand (completely free) and receptor residues within 12 Å of the ligand (fixed backbone atoms and completely free loop regions), as implemented in the MOE package. The LowModeMD Search method\textsuperscript{11} generates conformations using a short ~1 ps run of molecular dynamics which is followed by an energy minimization. Obtained conformations are stored in a database and ranked according to their potential energy. The LowModeMD Search was carried out using the Amber12:EHT force field\textsuperscript{9} and Born solvation.

**Final model selection.** Complexes with low energy obtained during the conformational space analysis were validated using available mutagenesis studies (see below). Final selected models were subjected to a comprehensive energetic minimization (Amber12:EHT force field\textsuperscript{9} and Born solvation) using the MOE package\textsuperscript{8}.

**Calculation of the binding pocket volume.** The volume was determined for the AT1R and AT2R in complex with [Y]\textsuperscript{6} analogue using the VMD plugin VOLSURF (default settings, [http://www2.fc.up.pt/PortoBioComp/Software/Volarea/Home.html](http://www2.fc.up.pt/PortoBioComp/Software/Volarea/Home.html)). The ligand was removed prior to volume calculations that were carried out using a box spanning the complete transmembrane binding site (yellow box, Figure S9).
Table S1. Complete resonance assignment for the trans isomer of the AII analogue.

| Amino Acid Residue | Proton assignment | $^1$H Chemical Shift (ppm) |
|--------------------|------------------|---------------------------|
| 1                  | HA               | 4.332                     |
| 1                  | HB2              | 3.030                     |
| 1                  | HB1              | 2.916                     |
| 2                  | HN               | 8.741                     |
| 2                  | HA               | 4.333                     |
| 2                  | HB2              | 1.699                     |
| 2                  | HB1              | 1.699                     |
| 2                  | HG2              | 1.501                     |
| 2                  | HG1              | 1.435                     |
| 2                  | HD2              | 3.106                     |
| 2                  | HD1              | 3.106                     |
| 2                  | HE               | 7.101                     |
| 2                  | HH21             | 6.877                     |
| 2                  | HH22             | 6.877                     |
| 2                  | HH11             | 6.429                     |
| 2                  | HH12             | 6.429                     |
| 3                  | HN               | 8.354                     |
| 3                  | HA               | 4.050                     |
| 3                  | HB               | 1.933                     |
| 3                  | HG21             | 0.886                     |
| 3                  | HG11             | 0.832                     |
| 4                  | HN               | 8.582                     |
| 4                  | HA               | 4.455                     |
| 4                  | HB2              | 2.765                     |
| 4                  | HB1              | 2.807                     |
| 4                  | HD1              | 6.948                     |
| 4                  | HE1              | 6.641                     |
| 5                  | HN               | 7.921                     |
| 5                  | HA               | 4.021                     |
| 5                  | HB               | 1.605                     |
| 5                  | HG12             | 1.309                     |
| 5                  | HD11             | 0.730                     |
| 5                  | HG21             | 1.022                     |
| 6                  | HN               | 8.362                     |
| 6                  | HA               | 4.632                     |
| 6                  | HB2              | 2.960                     |
| 6                  | HB1              | 2.806                     |
| 6                  | HD1              | 7.179                     |
| 6                  | HE1              | 6.817                     |
| 7                  | HA               | 4.324                     |
| 7                  | HB2              | 2.094                     |
| 7                  | HG2              | 1.897                     |
| 7                  | HG1              | 1.754                     |
| 7                  | HD2              | 3.484                     |
| 7                  | HD1              | 3.775                     |
| 8                  | HN               | 7.791                     |
| 8                  | HA               | 4.641                     |
| 8                  | HB2              | 3.011                     |
| 8                  | HB1              | 3.188                     |
| 8                  | HD1              | 7.233                     |
| 8                  | HE1              | 7.330                     |
Table S2. Complete resonance assignment for the cis isomer of the AII analogue.

| Amino Acid Residue | Proton assignment | 'H Chemical Shift (ppm) |
|--------------------|-------------------|------------------------|
| 1                  | HA                | 4.332                  |
| 1                  | HB2               | 3.030                  |
| 1                  | HB1               | 2.916                  |
| 2                  | HN                | 8.706                  |
| 2                  | HA                | 4.293                  |
| 2                  | HB2               | 1.648                  |
| 2                  | HG2               | 1.437                  |
| 2                  | HG1               | 1.349                  |
| 2                  | HD2               | 3.074                  |
| 2                  | HE                | 7.042                  |
| 3                  | HN                | 8.297                  |
| 3                  | HA                | 4.067                  |
| 3                  | HB                | 1.914                  |
| 3                  | HG21              | 0.865                  |
| 4                  | HN                | 8.635                  |
| 4                  | HA                | 4.625                  |
| 4                  | HB2               | 2.798                  |
| 4                  | HB1               | 2.965                  |
| 4                  | HD1               | 7.092                  |
| 4                  | HE1               | 6.752                  |
| 5                  | HN                | 8.093                  |
| 5                  | HA                | 4.239                  |
| 5                  | HB                | 1.793                  |
| 5                  | HD11              | 0.850                  |
| 5                  | HG21              | 0.910                  |
| 6                  | HN                | 8.394                  |
| 6                  | HA                | 3.997                  |
| 6                  | HB2               | 2.763                  |
| 6                  | HB1               | 2.922                  |
| 6                  | HD1               | 7.035                  |
| 6                  | HE1               | 6.815                  |
| 7                  | HA                | 3.304                  |
| 7                  | HD2               | 1.771                  |
| 8                  | HN                | 8.619                  |
| 8                  | HA                | 4.459                  |
| 8                  | HB2               | 3.129                  |
| 8                  | HB1               | 3.219                  |
| 8                  | HD1               | 7.198                  |
| 8                  | HE1               | 7.321                  |
Table S3. Summary of results for 125-I AII saturation assays on isolated membrane preparations and for analogues competition assays of wild-type and mutant AII receptors.

Assays were carried out as described in Supporting Information. Incubation of samples with ligands was for 2 hours on ice. The concentration of [125I]-AII used in competition binding assays was 2 nM. Kd data were fitted to the Michaelis-Menten equation using the non-linear regression equation of the software Prism. Ki values were calculated according to the Cheng and Prusoff equation with the Kd values in the first column. Values are representative of 2-3 independent experiments. Each data point was assayed in triplicate.

|             | Kd (nM) | Ki (M)  |
|-------------|---------|---------|
|              | [125I]-AII | [Y]-AII | [Y-PO₄]-AII |
| AT1R        | 1.8 ± 0.1 | - a     | - a         |
| AT2R        | 2.3 ± 0.3 | 3.4 x 10⁻⁹ | 0.9 x 10⁻⁹ |
| AT2R-Y189A  | 7.0 ± 0.9 | 1.2 x 10⁻⁸ | 5.1 x 10⁻⁸ |
| AT2R-Y189N  | 3.7 ± 0.5 | 1 x 10⁻⁸  | 2 x 10⁻⁷   |
| AT2R-F272A  | 4.2 ± 0.3 | 1.6 x 10⁻⁸ | 2.5 x 10⁻⁷ |
| AT2R-F272H  | 3.1 ± 0.6 | 1.6 x 10⁻⁸ | 2.1 x 10⁻⁸ |

* No detectable competitive binding was measured in the ligand concentration range used (6.4 x 10⁻¹²-2.5 x 10⁻⁶ M).
Table S4. Summary of results for 125-I AII saturation assays on whole cells and for analogues competition assays for the AII receptors.

| AT₁R  | Candesartan | CGP42112 | [F]⁶-AII | [Y]⁶-AII | [PO₄]⁶-AII | [NO₂]⁶-AII |
|-------|-------------|----------|----------|----------|------------|------------|
| IC₅₀  | 9.15x10⁻¹⁰ | 2.70x10⁻⁶ | 6.70x10⁻⁵ | 7.20x10⁻⁵ | 4.30x10⁻⁵ | 3.40x10⁻⁶ |
| Log IC₅₀ | -9.039 | -5.568 | -5.176 | -4.140 | -4.362 | -5.472 |
| Log S.E.M | 0.07053 | 0.09572 | 0.1159 | 0.2319 | 0.1323 | 0.1088 |
| 95% CI | 6.59x10⁻¹⁰ to 1.27x10⁻⁹ | 1.73x10⁻⁶ to 4.23x10⁻⁶ | 3.89x10⁻⁶ to 1.15x10⁻⁵ | 2.46x10⁻⁵ to 2.1x10⁻⁴ | 2.35x10⁻⁵ to 8.06x10⁻⁵ | 2.03x10⁻⁶ to 5.6x10⁻⁶ |

| AT₂R  | Candesartan | CGP42112 | F6-AII | Y6-AII | PO4-AII | NO2-AII |
|-------|-------------|----------|--------|--------|---------|---------|
| IC₅₀  | 1.40x10⁻⁵  | 5.60x10⁻¹⁰ | 1.40x10⁻⁹ | 3.40x10⁻⁹ | 9.6x10⁻⁹ | 4.90x10⁻⁹ |
| Log IC₅₀ | -4.870 | -9.253 | -8.853 | -8.400 | -8.019 | -8.308 |
| Log S.E.M | 0.09077 | 0.1238 | 0.05469 | 0.1464 | 0.1128 | 0.08018 |
| 95% CI | 8.80x10⁻⁶ to 2.07x10⁻⁵ | 3.12x10⁻¹⁰ to 9.98x10⁻¹⁰ | 1.09x10⁻⁹ to 1.82x10⁻⁹ | 2.00x10⁻⁹ to 7.93x10⁻⁹ | 5.64x10⁻⁹ to 1.63x10⁻⁸ | 3.38x10⁻⁹ to 7.18x10⁻⁹ |
Table S5. IC50 values and relative AT2R selectivity of the different (4-x-Phe)6-AII ligands.

| (4-x-Phe)6-AII | AT1R IC50 value (M) | AT2R IC50 value (M) | Fold selectivity AT1R/AT2R | AT2R selectivity | Stereoelectronic parameters |
|----------------|---------------------|---------------------|---------------------------|------------------|-----------------------------|
| 4-x-substitution |                      |                     |                           |                  |                             |
| -OH            | 7.2 $10^{-5}$       | 4.0 $10^{-9}$       | 18000                     | -0.37            | 40                          |
| -H             | 6.7 $10^{-6}$       | 1.4 $10^{-9}$       | 4786                      | +0.00            | 25                          |
| -OPO$_3$H$_2$  | 4.3 $10^{-5}$       | 9.6 $10^{-9}$       | 4479                      | +0.02            | 20                          |
| -NO$_2$        | 3.4 $10^{-6}$       | 4.9 $10^{-9}$       | 694                       | +0.78            | 5                           |
Table S6. Residue ID and corresponding Ballesteros-Weinstein numbering of residues that interact in the complexes of the [Y]$_6$-AII analogue with the AT1R and AT2R. Conserved residues are highlighted in bold.

| Residue ID | AT1R (P30556) | AT2R (P50052) | Ballesteros-Weinstein numbering |
|------------|----------------|----------------|-------------------------------|
| Y92 (ECL1) | Y108 (ECL1)    |                | -                             |
| R167       |                | -              | 4.64                          |
| I172 (ECL2)| I187 (ECL2)    |                | -                             |
| N174 (ECL2)| Y189 (ECL2)    |                | -                             |
| N188 (ECL2)| K203 (ECL2)    |                | -                             |
| S186 (ECL2)| Y204 (ECL2)    |                | -                             |
| P192       | S208           |                | 5.36                          |
| K199       | K215           |                | 5.43                          |
| H256       | F272           |                | 6.51                          |
| D278       |                | -              | 7.29                          |
| -          | D297           |                | 7.32                          |
Table S7. Examples of GPCR subtypes that contain proline in the sequence of their ligand*

| GPCR subtypes                          | Sequence of the peptide ligand |
|----------------------------------------|---------------------------------|
| VPAC1 receptor; VPAC2 receptor          | HSDAIFTEEYSKLLAKLALQKYLASILGSRTSPPP |
| AMY1 receptor; AMY2 receptor; AMY3 receptor; CT receptor | CGNLSTCLMLGTYTDFNKFHTFPQTAIGVGAP |
| GHSR1agrelin receptor                  | GSSFLSPEHQVRQKESKKPPAKLQPR       |
| PRRP receptor; Y1 receptor; Y2 receptor; Y4 receptor; Y5 receptor; Y6 receptor | YPSKPDNPGEAPAFDMARYSYALARHYINLITQRY |
| OX1 receptor; OX2 receptor              | RSGPPGLQGRLQLQASGNHAAGILTM       |
| MC1 receptor; MC3 receptor; MC4 receptor; MC5 receptor | SYSXEHRWGKPV                     |
| PAR1 receptor, PAR2 receptor, PAR3 receptor, PAR4 receptor | GYPQGV (human) GYPGKF (mouse)    |
| GnRH1, GNRH2                            | pGlu-His-Trp-Ser-His-Gly-Trp-Tyr-Pro-Gly-NH2 |

*Peptidergic GPCRs consist a major subfamily of GPCRs (IUPHAR/BPS). By searching in the GLIDA GPCR-Ligand Database (http://pharminfo.pharm.kyoto-u.ac.jp/services/glida/) there at least 68 peptide ligands that contain proline, 7 cases that contain the His-Pro motif (that can be substituted by 4-x-Phe-Pro), 10 entries containing the Tyr-Pro motif and 2 entries containing the Phe-Pro motif.
Table S8. Selected peptides and their PAR1 and PAR4 receptor activity; adopted from Faruqi T., et al. [12].

| Peptide   | PAR4 receptor activity | PAR1 receptor activity | 4-x-Phe Steroelectronic parameters |
|-----------|------------------------|------------------------|------------------------------------|
| GYPGKF    | 55                     | 0                      | -OH (electron-rich)                |
| GFPGKF    | 53                     | 35                     | -H (electron-neutral)              |
| G(F)PGKF* | 61                     | 111                    | -F (electron-deficient)            |

*[(F)=4-Fluoro-Phe]
Figure S1. Sequence alignment of the human AT1R (P30556) and the AT2R (P50052). Secondary structure is indicated by red (helix) or blue lines (turns). Conserved residues are highlighted in dark blue whereas residues with similarity are depicted in light blue. The sequence alignment was created using the MOE software (matrix blossom62).
**Figure S2.** Overlay of selected regions of 2D $^1$H-$^1$H NOESY spectrums of [Y$^\delta$]All (colored black) and native All (colored red) recorded under the same experimental conditions (0.01 M KPi buffer pH=5.7, 10% D$_2$O, 277K). The red and green lines denote the NOE connectivities for the trans and cis isomers respectively of [Y$^\delta$]All and the blue lines for the single trans isomer for the native All.
Figure S3. Region of the 750 MHz NOESY spectrum showing the intraresidue NOEs in the cis proline ring.

Figure S4. Region of the 750 MHz NOESY spectrum showing the intraresidue $C^\alpha H$-$C^\beta H$ cross peaks for cis Tyr$^6$ and cis Phe$^8$, as well as a number of characteristic NOEs of the folded conformation of the cis form of the peptide.
**Figure S5.** Indicative members of the families of clusters for the *cis* cases in the Tyr\(^{i-1}\)-Pro\(^i\)-Phe\(^{i+1}\) motif. From left to right are illustrated: the case of three ring clustering among Tyr\(^{i-1}\), Pro\(^i\) and Phe\(^{i+1}\); two ring clustering between Tyr\(^{i-1}\) and Pro\(^i\); and two ring clustering between Tyr\(^{i-1}\) and Phe\(^{i+1}\) respectively. The most populated cluster is the three ring clustering among Tyr\(^{i-1}\), Pro\(^i\) and Phe\(^{i+1}\).

**Figure S6.** (a) NH Δδ/ΔT vs CSD for [Y]\(^6\)-All. The dashed line corresponds to Δδ/ΔT = −7.8 (CSD) −4.4, which provides the optimum differentiation of sequestered NHs in the protein database. (b) \(^1\)H DOSY NMR spectrum of [Y]\(^6\)-All. Grey and black signals correspond to the *cis* and *trans* conformers, respectively.
**Figure S7.** Schematic view of ligand-receptor interaction in the Y6-AT1R (left) and the Y6-AT2R complex (right). Hydrogen bonds are illustrated in red dashed lines; hydrophobic contacts are shown in a red spike-like presentation. Completely conserved interactions between the Y6-AT1R and Y6-AT2R complex are highlighted in light green whereas homologue interactions are shown in dark green. We refer to homologue interactions, when a ligand functional group that is located in the same site interacts with a homologue residue at the same or nearby position.
Figure S8. Superposition of the \([Y]^6\)-AII in its complex state with AT1R (yellow stick & surface) and AT2R (blue stick & surface). The Tyr\(^6\)-Pro\(^7\)-Phe\(^8\) region of \([Y]^6\)-AII presents in these models a slightly deeper penetration into the binding pocket of AT2R with respect to AT1R.
Figure S9. In the X-ray structure of the complex between ubiquitin-protein ligase E3A and ubiquitin conjugating enzyme E2 (pdbid: 1C42), a Tyr-Pro-Phe motif (YPF), belonging to E2 (residues 61-63), is located in the interface of the interaction. In this motif (colored in red) there is a cis proline and its ring is packed against the aromatic rings of Tyr and Phe. The environment around the relevant motif was selected with a radius cut off of 6 Å (carbon colored in grey, nitrogen in blue and oxygen in red color). Interestingly, this environment closely resembles the environment existed near the ligand binding site of AT2 (residues colored in orange). The structure of AT2 used for this superposition was constructed based on the rhodopsin in its ligand-free state (pdbid: 3CAP). Homologues residues between AT2 and residues surrounding the environment of the Tyr-Pro-Phe motif are: W269/W105; K215/R96; Y189/Y694, Y51(A194)/F698, L97/L695, L124/L696, L305(I304)/L659, T276/S65, I196/L697, L124/L696, P271/P58, L190/Y694, H273/F66, F220/P68. Phe308, Phe129, Phe272 and Ile304 could assist the assembly of a similar motif in AT2.
Figure S10. Calculation of the binding pocket volume using the VMD plugin VOLSURF.
Figure S11. MCF-7 cells were grown without drug, with drug vehicle only and with various concentrations of the analogue as indicated. Cells were counted daily on days 2-5. Data are shown are cell numbers relative to control cells which received drug vehicle.

References

[1] L. Hunyady, A. J. Baukal, T. Balla, K. J. Catt, *J Biol Chem* 1994, 269, 24798.

[2] F. Magnani, Y. Shibata, M. J. Serrano-Vega, C. G. Tate, *Proc Natl Acad Sci U S A* 2008, 105, 10744.

[3] W. Schaffner, C. Weissmann, *Anal Biochem* 1973, 56, 502.

[4] S. Bosnyak, E. S. Jones, A. Christopoulos, M. I. Aguilar, W. G. Thomas, R. E. Widdop, *Clin Sci (Lond)* 2011, 121, 297.

[5] P. G. Strange, *Br J Pharmacol* 2008, 153, 1353.

[6] M. Tamura, Y. Wanaka, E. J. Landon, T. Inagami, *Hypertension* 1999, 33, 626.

[7] D. S. Wishart, C. G. Bigam, A. Holm, R. S. Hodges, B. D. Sykes, *J Biomol NMR* 1995, 5, 67.

[8] h. w. c. s. h. Molecular Operating Environment (MOE) software.

[9] P. R. Gerber, K. Muller, *J Comput Aided Mol Des* 1995, 9, 251.

[10] C. J. Verdonk ML, Hartshorn MJ, Murray CW, Taylor RD. Improved proteinligand docking using GOLD. Proteins. 2003 sep 1;52(4):609623.

[11] P. Labute, *J Chem Inf Model* 2010, 50, 792.
[12] T. R. Faruqi, E. J. Weiss, M. J. Shapiro, W. Huang, S. R. Coughlin, *J Biol Chem* **2000**, 275, 19728.