Supporting information

Synthesis, biological evaluation and automated docking of constrained analogs of the opioid peptide H-Dmt-D-Ala-Phe-Gly-NH_2 using the 4- or 5-methyl substituted 4-amino-1,2,4,5-tetrahydro-2-benzazepin-3-one scaffold.

Rien De Wachter,[1] Chris de Graaf,[2,3]* Atilla Keresztes,[4] Bart Vandormael,[1] Steven Ballet,[1] Géza Tóth,[4] Didier Rognan[2] and Dirk Tourwé.[1]*

[1]Organic Chemistry Department, Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussels, Belgium; [2]Structural Chemogenomics, UMR 7200 CNRS-UdS (Université de Strasbourg), Illkirch F-67401, France; [3]Leiden/Amsterdam Center for Drug Research, Division of Medicinal Chemistry, Faculty of Science, VU University Amsterdam, Amsterdam, The Netherlands; [4]Institute of Biochemistry, Biological Research Center, Hungarian Academy of Sciences, Szeged, Hungary;

Content:

***To whom correspondence should be addressed.**

Chris de Graaf, Leiden/Amsterdam Center for Drug Research, Division of Medicinal Chemistry, Faculty of Science, VU University Amsterdam, De Boelelaan 1083, 1081 HV Amsterdam, The Netherlands, phone +31 20 5987550, e-mail: C.de.Graaf@few.vu.nl;

Dirk Tourwé, Organic Chemistry Department, Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussels, Belgium, phone +32 2 6293295, e-mail: datourwe@vub.ac.be

Enantiopure β-methylphenylalanine isomers.

The diastereomeric *erythro*-(2*S*,3*S* and 2*R*,3*R*)- and *threo*-(2*S*,3*R* and 2*R*,3*S*)-β-MePhe racemates were obtained by fractional crystallization[1] of the isomeric mixture prepared by the method of Kataoka.[2] After Cbz-protection, resolution was performed by crystallization of the quinine or quinidine salts according to Kataoka.[2] The enantiomeric ratio (ee) was determined with FDAA derivatisation[3] (after small scale Cbz-deprotection) via RP-HPLC and UV detection at λ=340 nm. The retention times after derivatisation are: 16.3 min (for erythro-(2*S*,3*S*)-β-Me-Phe), 17.1 min (for erythro-(2*R*,3*R*)-β-MePhe), 16.33 min (for threo-(2*S*,3*R*)-β-Me-Phe) and 17.05 min (for threo-(2*R*,3*S*)-β-MePhe).

Enantiopure Boc-*erythro*-(4*S*,5*S*) (3), -*erythro*-(4*R*,5*R*) (4), -*threo*-(4*S*,5*R*) (5) and -*threo*-(4*R*,5*S*)-5-Me-Aba-Gly-OH (6) The enantiopure isomers of **3** to **6** were prepared using the previously described procedure for the racemates.[4] Briefly, the enantiopure β-methylphenylalanine isomers were phtaloyl protected using methyl-2-((succinimidooxy)carbonyl)benzoate, and coupled with methyl glycinate. The esters were subsequently hydrolyzed, followed by formation of the oxazolidinone by reaction with formaldehyde using p-TosOH catalysis. Ring closure was effected using trifluoromethanesulfonic acid (TFMSA) to provide the Phth-protected *erythro* and *threo*-5-Me-Aba-Gly-OH isomers. After phthaloyl deprotection by hydrazinolysis, final Boc protection was followed by crystallization of the *N,N*-dicyclohexylamine (DCHA) salts in Et_2O.[5] The crystals contained a small amount of di-Boc hydrazide which could not be removed. Before performing the peptide synthesis, the DCHA salts were converted into the free acid[5] and the Boc-protected enantiomers were purified by an additional acid-base extraction which removed the di-Boc-hydrazide. The optical rotation data and RP-HPLC retention times of the different stereoisomers are collected in Table 3:

Table S1: Optical rotations and HPLC retention times of **3-6**

Compound	Optical rotation α_D (solvent)	RP-HPLC t_{Ret}
Boc-*erythro*-(4*S*,5*S*)-5-Me-Aba-Gly-OH **3**	+90.6° (EtOAc)	15.4 min
Boc-*erythro*-(4*R*,5*R*)-5-Me-Aba-Gly-OH **4**	-93.0° (EtOAc)	15.3 min
Boc-*threo*-(4*S*,5*R*)-5-Me-Aba-Gly-OH **5**	Not determined*	14.6 min
Boc-*threo*-(4*R*,5*S*)-5-Me-Aba-Gly-OH **6**	Not determined*	14.6 min

* HPLC analysis showed that the Boc-*threo*-(4*S*,5*R*)-5-Me-Aba-Gly-OH.DCHA had epimerized to the erythro derivative [46% *threo*-(4*S*,5*R*) + 54% *erythro*-(4*S*,5*S*)] and the Boc-*threo*-(4*R*,5*S*)-5-Me-Aba-Gly-OH to the (4*R*,5*R*) erythro analogue [48% *threo*-(4*R*,5*S*) + 52% *erythro*-(4*R*,5*R*)]

Supplementary Figure S1. The amino acid sequence alignments of (boxed) transmembrane helices (TM1-7) and the intra- and extracellular loops (Ils en ELS) of bovine rhodopsin (OPSD_BOVIN), the human OPRM and OPRD. Conserved reference residues according to the Ballesteros-Weinstein numbering scheme[6] are indicated in red (N1.50 in TM 1, D2.50 in TM2, R3.50 in TM3, W4.50 in TM4, P5.50 in TM5, P6.50 in TM6, and P7.50 in TM7).

```
                              TM1                    IL1
OPRM_HUMAN    [ITAITIMALYSIVCVVGLFGNFLVMYVIVR] YTKMKT
OPRD_HUMAN    [ALAIAITALYSAVCAVGLLGNVLVMFGIVR] YTKMKT
ADRB2_HUMAN   [VWVVGMGIVMSLIVLAIVFGNVLVITAIAK] FERLQT
CCR5_HUMAN    [IAARLLPPLYSLVFIFGFVGNMLVILILIN] CKRLKS
OPSD_BOVIN    [WQFSMLAAYMFLLIMLGFPINFLTLYVTVQ] HKKLRT

                              TM2                    EL1
OPRM_HUMAN    [ATNIYIFNLALADALATSTLPFQSVNYLMG] -TWPFG
OPRD_HUMAN    [ATNIYIFNLALADALATSTLPFQSAKYLME] -TWPFG
ADRB2_HUMAN   [VTNYFITSLACADLVMGLAVVPFGAAHILM] KMWTFG
CCR5_HUMAN    [MTDIYLLNLAISDLFFLLTVPFWAHYAAAQ] --WDFG
OPSD_BOVIN    [PLNYILLNLAVADLFMVFGGFTTTLYTSLH] GYFVFG

                                 TM3                        IL2
OPRM_HUMAN    [TILCKIVISIDYYNMFTSIFTLCTMSVDRYIAV] CHPVKALDFRTP
OPRD_HUMAN    [ELLCKAVLSIDYYNMFTSIFTLTMMSVDRYIAV] CHPVKALDFRTP
ADRB2_HUMAN   [NFWCEFWTSIDVLCVTASIETLCVIAVDRYFAI] TSPFKYQSLLTK
CCR5_HUMAN    [NTMCQLLTGLYFIGFFSGIFFIILLTIDRYLAV] VHAVFALKARTV
OPSD_BOVIN    [PTGCNLEGFFATLGGEIALWSLVVLAIERYVVV] CKPMSNFRFGE

                          TM4                         EL2
OPRM_HUMAN    [RNAKIINVCNWILSSAIGLPVMF] MATTKYRQG--SIDCTLTFSHPTW-YW
OPRD_HUMAN    [AKAKLINICIWVLASGVGVPIMV] MAVTRPRDG--AVVCMLQFPSPSW-YW
ADRB2_HUMAN   [NKARVIILMVWIVSGLTSFLPIQ] MHWYRATHQ-EAINCYANETCCDFFT-
CCR5_HUMAN    [TFGVVTSVITWVVAVFASLPGII] --FTRSQKEGLHYTCSSHFPYSQYQFW
OPSD_BOVIN    [NHAIMGVAFTWVMALACAAPPLV] -GWSRYIPEGMQCSCGIDYYTPHEETN

                          TM5                         IL3
OPRM_HUMAN    [ENLLKICVFIFAFIMPVLIITVCYGL] MILRLK-SVRMLSGSKEKDRN
```

```
OPRD_HUMAN    [DTVTKICVFLFAFVVPILIITVCYGL] MLLRLR-SVRLLSGSKEKDRS
ADRB2_HUMAN   [NQAYAIASSIVSFYVPLVIMVFVYSR] VFQEAKRQLQKIDKSEGRFHVQNLSQVEQDGRTGHGLRRSSKFCLK
CCR5_HUMAN    [KNFQTLKIVILGLVLPLLVMVICYSG] ILKTLLRCRNEKK
OPSD_BOVIN    [NESFVIYMFVVHFIIPLIVIFFCYGQ] LVFTVKEAAAQQQESATTQKA

                          TM6                          EL3
OPRM_HUMAN    [LRRITRMVLVVVAVFIVCWTPIHIYV] IIKALVTIPETTFQT-
OPRD_HUMAN    [LRRITRMVLVVVGAFVVCWAPIHIFV] IVWTLVDIDRRDPLVV
ADRB2_HUMAN   [EHKALKTLGIIMGTFTLCWLPFFIVN] IVHVIQDNLIRK---
CCR5_HUMAN    [RHRAVRLIFTIMIVYFLFWAPYNIVL] LLNTFQEFFGLNNCSSSNRLD
OPSD_BOVIN    [EKEVTRMVIIMVIAFLICWLPYAGVA] FYIFTHQ-GSDFGP--

                          TM7                          helix8
OPRM_HUMAN    [VSWHFCIALGYTNSCLNPVLY] -AFLDENFKRCF-REFCIPTSS
OPRD_HUMAN    [AALHLCIALGYANSSLNPVLY] AFLDENFKRCF-RQLCRKPCG
ADRB2-HUMAN   [EVYILLNWIGYVNSGFNPLIY] CRSPDFRIAFQELLCLRRSS
CCR5_HUMAN    [QAMQVTETLGMTHCCINPIIY] -AFVGEKFRNYLLVFFQKHIAKRFCKCCSIFQQEAPERASS
OPSD_BOVIN    [IFMTIPAFFAKTSAVYNPVIY] IMMNKQFRNCMVTTLCCGK----NP
```

Table S2: Full Interaction Fingerprints (IFPs) of the docking poses of **JOM6**, compounds **9**, **11**, **12** and **14** (presented in Figure 4)

JOM6

D3.32	Y3.33	F3.37	G45.46	S45.47	D45.49	C45.50	T45.51	L45.52	T45.53	F45.54	D45.49	E5.35
1	0	0	1	0	1	0	0	0	0	0	0	0
K5.39	F5.43	F5.47	W6.48	I6.51	H6.52	V6.55	K6.58	T6.70	W7.35	H7.36	I7.39	Y7.43
1	0	0	0	0	0	0	0	0	0	0	0	0

Compound **12**

D3.32	Y3.33	F3.37	G45.46	S45.47	D45.49	C45.50	T45.51	L45.52	T45.53	F45.54	D45.4	E5.35
1	0	0	1	0	1	0	0	0	0	0	0	0
K5.39	F5.43	F5.47	W6.48	I6.51	H6.52	V6.55	K6.58	T6.70	W7.35	H7.36	I7.39	Y7.43
0	0	0	0	0	0	0	0	0	0	0	0	0

Compound **9**

D3.32	Y3.33	F3.37	G45.46	S45.47	D45.49	C45.50	T45.51	L45.52	T45.53	F45.54	D45.4	E5.35
1	0	0	1	0	1	0	0	0	0	0	0	0
K5.39	F5.43	F5.47	W6.48	I6.51	H6.52	V6.55	K6.58	T6.70	W7.35	H7.36	I7.39	Y7.43
0	0	0	0	0	0	0	0	0	0	0	0	0

Compound **11**

D3.32	Y3.33	F3.37	G45.46	S45.47	D45.49	C45.50	T45.51	L45.52	T45.53	F45.54	D45.4	E5.35
1	0	0	1	0	1	0	0	0	0	0	0	0
K5.39	F5.43	F5.47	W6.48	I6.51	H6.52	V6.55	K6.58	T6.70	W7.35	H7.36	I7.39	Y7.43
1	0	0	0	0	0	0	0	0	0	0	0	0

Compound **14**

D3.32	Y3.33	F3.37	G45.46	S45.47	D45.49	C45.50	T45.51	L45.52	T45.53	F45.54	D45.49	E5.35
1	0	0	1	0	1	0	0	0	0	0	0	0
K5.39	F5.43	F5.47	W6.48	I6.51	H6.52	V6.55	K6.58	T6.70	W7.35	H7.36	I7.39	Y7.43
1	0	0	0	0	0	0	0	0	0	0	0	0

IFP bit-strings:

1. hydrophobics
2. aromatic face to face
3. aromatic edge to face
4. H-bond (acceptor-donor)
5. H-bond (donor-acceptor)
6. Ionic (negative-positive)
7. Ionic (positive-negative)

References

1. Hruby, V.J.; Toth, G.; Gehrig, C.A.; Kao, L.-F.; Knapp, R.; Lui, G.K.; Yamamura, H.I.; Kramer, T.H.; Davis, P.; Burks, T.F. Topographically designed analogues of [D-Pen,D-Pen5]enkephalin. *J. Med. Chem.* **1991**, *34*,1823-1830.

2. Kataoka, Y.; Seto, Y, Yamamoto, M.; Studies of unusual amino-acids and their peptides. 6. Synthesis and optical resolutions of beta-methylphenylalanine and its dipeptide present in Bottromycin. *Bull. Chem. Soc. Japan* **1976**, *46*, 1081-1084.

3. Marfey, P. Determination of D-amino acids .2. Use of a bifunctional reagent, 1,5-difluoro-2,4-dinitrobenzene. *Carlsberg Res. Commun.* **1984,** *49*, 591-596.

4. De Wachter, R.; Brans, L.; Ballet, S.; Van den Eynde, I.; Urbanczyk-Lipkovska, Z.; Tourwe, D. Influence of ring substitutions on the conformation and β-turn mimicry of 4-amino-1,2,3,4-tetrahydro-2-benzazepin-3-one peptide mimetics. *Tetrahedron* **2009,** *65*, 2266-2278.

5. Stewart, J.M.; Young J.D. , Solid phase peptide synthesis, 2nd Ed. , Pierce Cemical Co, Rockford, 1984, pp. 63.

6. Ballesteros *et al*. Integrated Methods for the Construction of Three-Dimensional Models and Computational Probing of Structure-Function Relations in G Protein-Coupled Receptors. Ballesteros JA, Weinstein H. *Methods Neuroscience* **1995**, *25*, 366-428.