Supporting Information

Proposed bioactive conformations of Opiorphin, an endogenous dual APN/NEP inhibitor

Marta Pinto, Catherine Rougeot, Luis Gracia, Monica Rosa, Andres Garcia, Gemma Arsequell, Gregorio Valencia, Nuria B. Centeno[*]

Computer-Assisted Drug Design Laboratory, Research Programme on Biomedical Informatics (GRIB), IMIM-Universitat Pompeu Fabra, Dr. Aiguader 88, E-08003 Barcelona, Spain.

[*]*e-mail: nuria.centeno@upf.edu*

Table of contents

1. Experimental details

Peptide synthesis. Amino acid building blocks, coupling reagents and Wang resin were purchased from Novabiochem AG. All reagents used for synthesis were analytical grade. Peptides were synthesized manually following standard solid-phase methods[1] and Fmoc protocols on Wang resin using amino acids with orthogonal protections on lateral chains. Amide couplings were performed manually in a peptide synthesis column using DIC/HOBt in DMF under reciprocal oscillating agitation. Coupling efficiencies were monitored by Kaiser ninhydrin test. Fmoc groups were removed with a 20% piperidine in DMF solution. After assembling sequences, a cocktail of TFA:H2O:TIS (95:2.5:2.5) was used to remove side chain protecting groups and to cleave each peptide from the resin.

Peptide purification and analytical determinations. Crude peptides were precipitated on ice-cold t-butyldimethyl ether, redissolved in water and purified on a Reverse Phase (C-18) Column with VersaFlash^{TM} system using H2O-ACN gradient. Homogeneity and identity of final peptides was assured by analytical RP-HPLC and HRMS on a Waters UPLC-ESI/TOF system.

hAPN assays. Soluble recombinant hAPN and Ala-Mca[2], a fluorogenic substrate for measuring aminopeptidase activity, were purchased from R&D Systems and Sigma respectively. Using black half-area 96 well microplate the standard reaction consisted of enzyme (4 ng) in 100 mM Tris-HCl pH 7.0 (100 µl final volume). The Ala-Mca substrate (25 µM final concentration) was added after preincubation for 10 min at 28° C and the kinetics of appearance of the signal was monitored for 40 min at 28° C by using a fluorimeter reader at 380 nm excitation and 460 nm emission wavelengths. The intensity of the signal was directly proportional to the quantity of metabolites formed

during the 10-40 min time-period of the reaction. Under these conditions of initial velocity measurement, the rhAPN-mediated aminoproteolysis of Ala-Mca was directly calculated.

hNEP assays. Soluble recombinant hNEP was purchased from R&D Systems. NEP-endopeptidase activity was assayed by measuring the breakdown of the synthetic selective substrate Abz-dR-G-L-EDDnp[3] FRET-peptide synthesized by Thermo-Fisher Scientific (Germany). In a black half-area 96 well microplate, 12.5 ng of recombinant hNEP in 100 mM Tris-HCl pH 7 containing 200 mM NaCl (100 µl final volume) were preincubated for 10 min at 28º C with several concentrations of Opiorphin analogues in a buffer solution of 100 mM pH7 Tris-HCl, 150 mM NaCl and 0.05% Brij35. Specific substrate Abz-dR-G-L-EDDnp (15µM final concentration) was added and the kinetics of appearance of the fluorescent signal (RFU) was directly analyzed for 40 min at 28° C. (2.3 min-interval successive measures) by using a fluorimeter microplate reader (monochromator Infinite 200-Tecan) at 320 nm and 420 nm excitation and emission wavelengths, respectively. Under conditions of initial velocity measurement, the intensity of the signal was directly proportional to the quantity of metabolites formed during the 20-40 min time-period of the reaction. Thus, in absence of inhibitor, the initial velocity of rhNEP-mediated specific endoproteolysis of Abz-dR-G-L-EDDnp, was calculated from the linear regression.

2. Computational details

Conformational search. The conformational space of the peptides was sampled using an iterative simulated annealing (SA) protocol similar to that first described by Filizola et al.[4] The main features of this methodology is that it quickly reaches low-energy conformations, and in its recursive search the procedure is robust enough to obtain low-energy conformations situated in different valleys of the peptide landscape. For all peptides, starting zwitterionic structures were generated in their extended conformation using the AMBER (version 10) suite of programs[5] and the all-atom ff03 force field.[6] For each peptide, the initial extended structure was energy minimized and subsequently heated to 900 K using a fast heat rate of 100 K/ps to force the structure to jump among different regions of the conformational space. The resulting high energy structure was then slowly cooled to 200 K at a cool rate of 7 K/ps. The so obtained structure was used as a starting point for a new SA cycle. After each cycle, low temperature structures were minimized before being added to a library of low-energy conformations for further analysis.

Within the simulated annealing protocol, molecular dynamic simulations used a 2 fs time step. The SHAKE algorithm[7] was applied to all bonds involving hydrogen atoms. A modified generalized Born model[8] was used to describe the solvent. All calculations were carried out using the ALPB approximation[9,10] to handle electrostatic interactions within an implicit solvent model. No cutoff for the non-bonded interactions was used. The non-bonding list was updated every 10 steps. Minimizations consisted of 1,000 steps of steepest descent followed by conjugate gradient until the root-mean square deviation (RMSD) of the energy gradient was less than 10^{-3} kcal/mol Å.

This iterative protocol was repeated until no unique conformations (see below) appeared after 300 cycles in a range of 5 kcal/mol with respect to the lowest energy structure.

Previous studies[11] have shown the goodness of this criterion, since the minima distribution generated in such a way exhibit the same characteristics as the density of states derived from the rotational isomeric model for the low-energy subset of conformations.

A conformation was considered unique if at least one the backbone dihedral angles, excluding those situated at both termini, differs by more than 60º compared to any of the previous obtained structures.

Table S1. Data from the simulated annealing calculations performed on Opiorphin and the set of studied analogues. For each peptide, sampling cycles of SA were performed iteratively until the convergence criterion was achieved (see text for details). Conformations with relative energy to the lowest energy structure less or equal to 5 kcal/mol were kept for analysis (low-energy conformation library).

Name	**Number of SA cycles performed**	**Members of the low-energy conformation library**
Opiorphin	9,251	28
Opiorphin-(3-5)	7,969	25
Opiorphin-(2-5)	8,757	13
Opiorphin-(1-4)	8,158	18
$[Asn]^1$-Opiorphin	9,062	31
$[Thr]^4$-Opiorphin	8,933	14
$[Phe]^0$-Opiorphin	10,921	10
$[Cys]^0$-Opiorphin	11,895	36

All unique conformations with energy in the range of 5 kcal/mol with respect to the lowest energy structure were stored for further analysis. This set is referred in the text as a 'low-energy conformation library'. In each library, conformations were ranked by energy starting with the lowest energy structure. Conformations were named N.n, where N stands for the peptide and n for the position in the library.

Automated Simulated Annealing Interface (ASAI). An in-house program called *Automated Simulated Annealing Interface* (ASAI) was developed to carry out this iterative protocol in an efficient and automated way. The program serves as an interface to the AMBER software, constructing the initial extended conformations, sending dynamics and minimization jobs and reading the data back for clustering analysis.

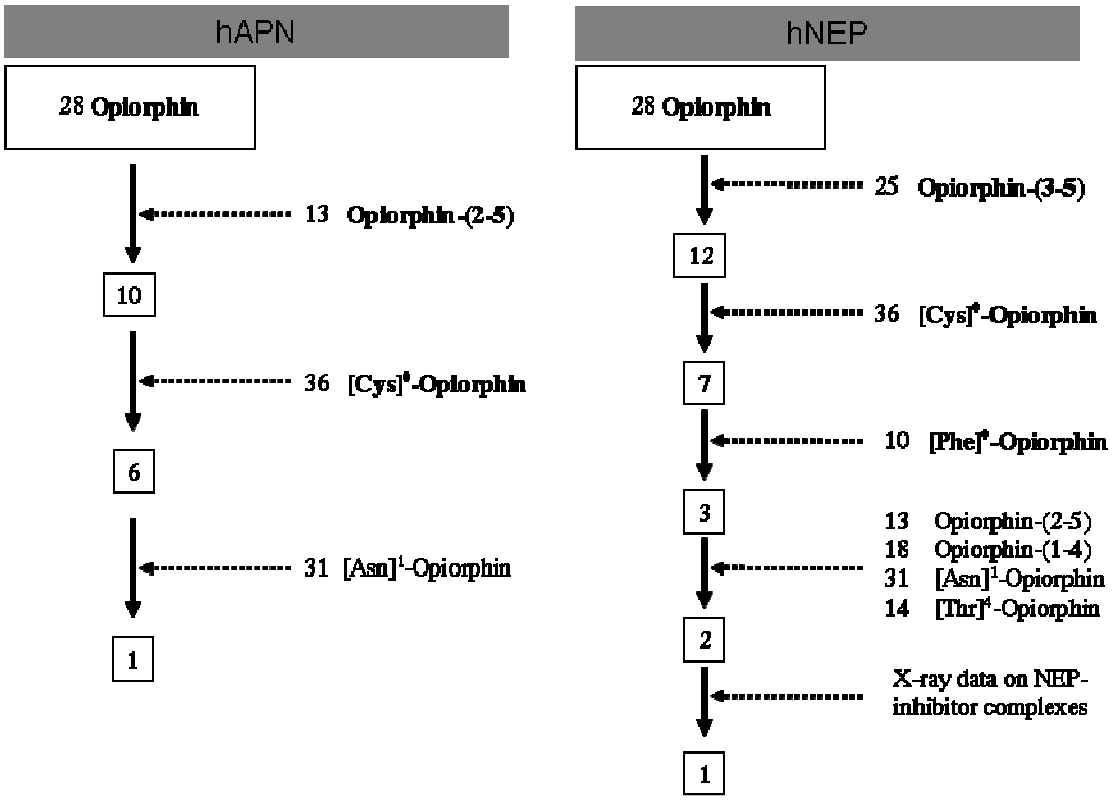
Previous algorithms[4] have generated a set of hundreds of conformations in a run, analyzed them, and then decide to continue the sampling based on the number of new unique conformations obtained. This method leads to an overhead in computational effort. First, unnecessary number of SA cycles is computed. Furthermore, comparisons among conformations are re-evaluated many times as each new set is added and analysis is started from scratch. We have developed an algorithm in which each new conformation is subjected to analysis after it is included in the library, therefore avoiding unnecessary computation and bringing the conformational search to an end as soon as the search criterion is met (see above).

After inserting each new minimized conformation *i* in the library, it is first sorted based on its relative energy to the lowest energy structure found. Now the uniqueness of conformation *i* is assessed against all other conformations in the library based on the criterion specified above (at least one of the backbone dihedral angles excluding those situated at both termini differs by more than 60º). To avoid recalculations and speed up the analysis, conformation *i* is first compared only to unique *j* conformations in the

library that are lower in energy. If *i* is different from all *j* conformations, it is considered unique. Otherwise, it is added as a member of cluster *j* and the comparison finishes here. However, if conformation *i* is found to be unique, conformations *k* of higher energy might now change their cluster membership and their uniqueness needs to be reassigned. Many repetitive comparisons can be avoided by keeping track of conformations *k* that have changed their uniqueness. This algorithm creates clusters based on the criterion established, in which the head of each cluster is the lowest energy conformation.

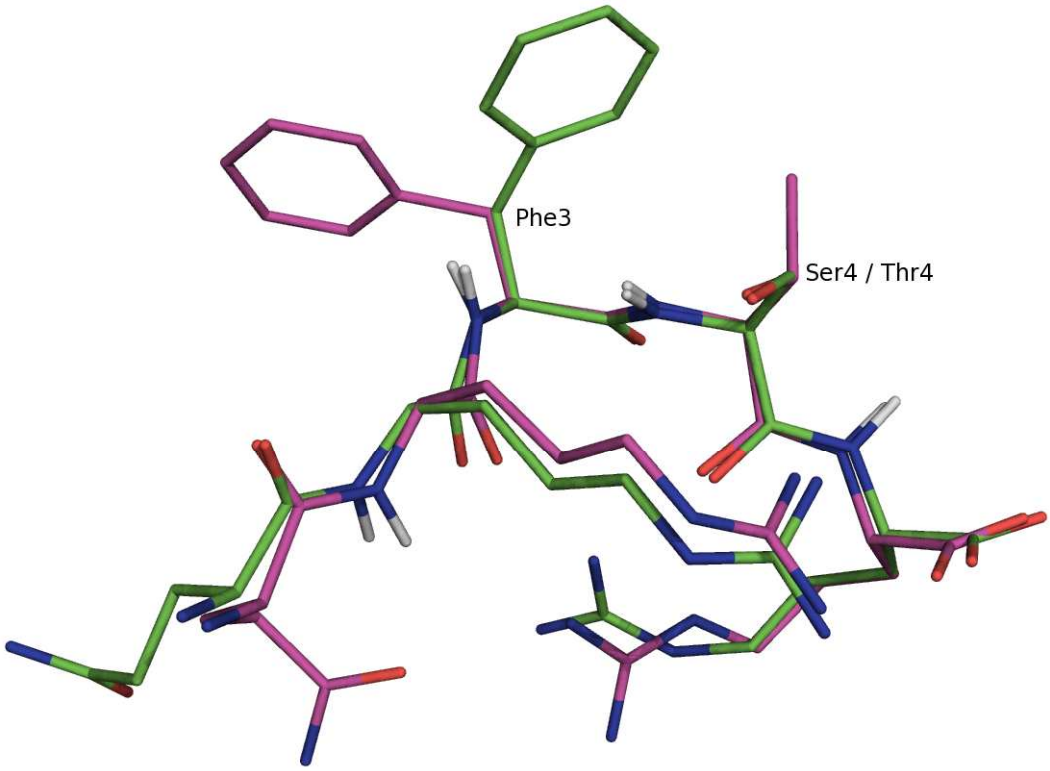
Comparative conformational analysis. In order to characterize the bioactive conformation of Opiorphin, pairwise cross-comparisons between the low-energy conformational libraries of this peptide and its analogues were performed. For this purpose, computation of the RMSD between all the backbone atoms of a subset of residues, depending on the peptide analogue compared, was performed to locate common conformations among them. The use of the RMSD as a similarity criterion requires selecting a threshold value that depends on the number of atoms involved in the comparison. For comparisons involving five residues, after testing different threshold values, two conformations were automatically considered similar if the RMSD value was lower than 0.40 Å and different if it was higher than 0.85 Å. For RMSDs between these two cutoffs, a careful visual inspection made by a trained user was required in order to determine if the structural differences were meaningful or not. The automated cuttofs were scaled linearly for comparisons of four (0.30 - 0.75 Å) and three residues (0.20 - 0.65 Å). Pymol[12] was used for visual inspection, as well as for the preparation of the figures.

Figure S1. Schematic diagram of the workflow followed in the comparative conformational analysis to propose the bioactive conformation of Opiorphin for hAPN (left) and hNEP (right) inhibition. Active peptides are shown in bold. Numbers prior to the peptide's names are the number of conformations in each library. Numbers in main-flow boxes indicate the number of Ophiorphin conformations remaining after comparison with the low-energy conformation library indicated in each step.



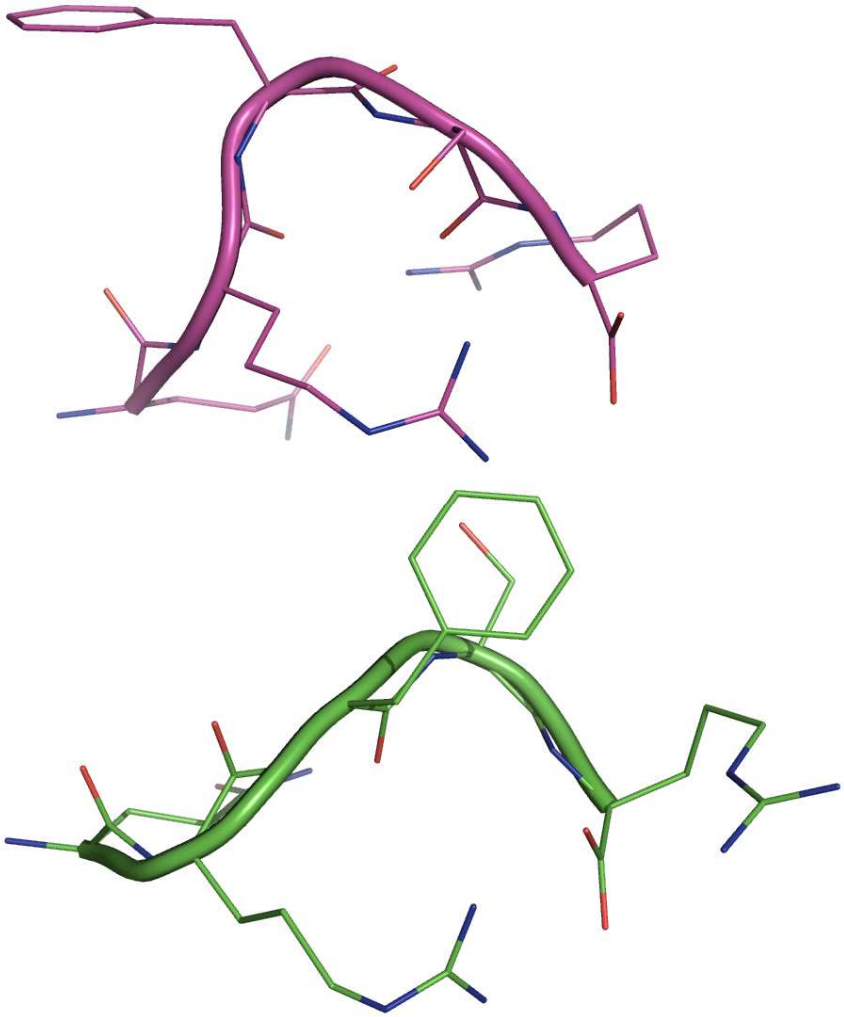
3. Comparison of bioactive conformation for APN common to Opiorphin and [Thr]^4-Opiorphin

Figure S2. Superimposition of the putative bioactive conformation of Opiorphin (in green) and the conformation of [Thr]^4-Opiorphin (in magenta) that exhibits the smallest RMSD with the Opiorphin structure. Differences in the position of the phenyl ring of Phe^3 are due to the presence of an additional methyl group in [Thr]4-Opiorphin. A possible explanation of the inactivity of [Thr]^4-Opiorphin may be the inability of the Phe^3 side chain to properly interact with hAPN.



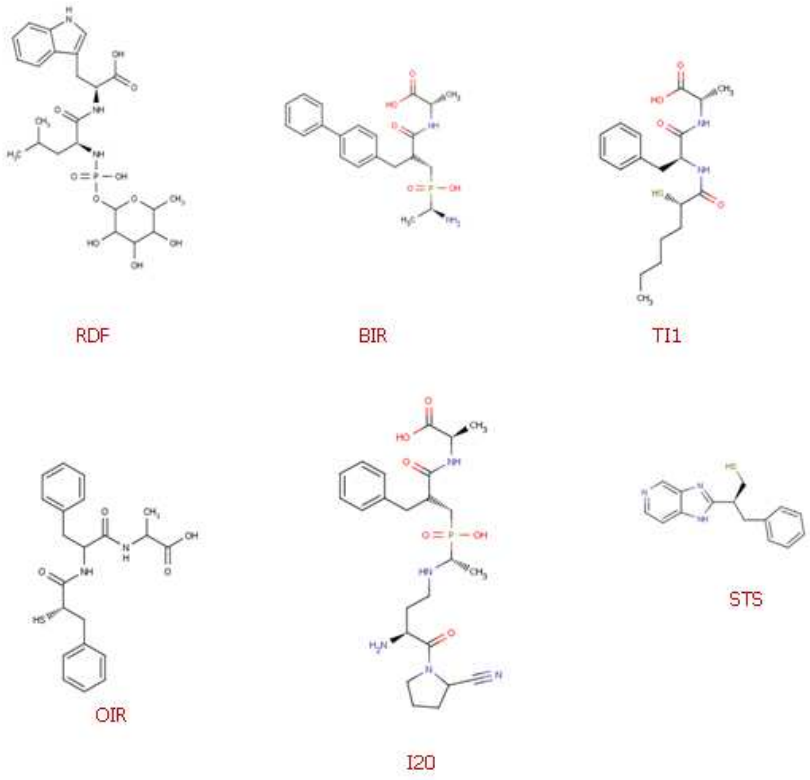
4. Candidates for hNEP bioactive conformation found by comparative conformational analysis

Figure S3. Representation of the 3D structure of the two hNEP bioactive Opiorphin conformation candidates found using the comparative conformational analysis method: 1.1 (in magenta) and 1.17 (in green). They exhibit distinct shapes mainly due to differences in the φ_4 dihedral angle value.



5. Comparison of inhibitor-hNEP complexes with the two hNEP bioactive Opiorphin conformation candidates

From the six available NEP-complexes[13-16] (see scheme S1), we selected the five which are dipeptide derivatives (PDB ligand identifiers: RDF, BIR, TI1, OIR and I20).



Scheme S1. Chemical structures and names of the ligands for six inhibitor-NEP complexes available in the PDB. All of them but STS are peptide derivatives.

hNEP structures were superimposed with Pymol[12] to analyze the bioactive conformation of the ligands and their binding mode. As can be seen in Figure S4, the ligands share the same binding mode and exhibit the same conformation at the C-terminal moiety. After superimposing our two candidates conformations (1.1 and 1.17) with them, we found that 1.17 adopts the same conformation as these ligands (backbone RMSD of the FSR fragment =0.26 Å) (Figure S5). For this reason, we selected 1.17 as the more likely bioactive conformation for hNEP inhibition.

Figure S4. Superimposition of the bioactive conformation of five dipeptide derivatives that have been obtained by X-ray crystallography: RDF (orange), BIR (blue), OIR (yellow), TI1 (grey) and I20 (magenta). It can be seen that all are arranged in the same manner and share the same conformation at the C-terminal moiety.

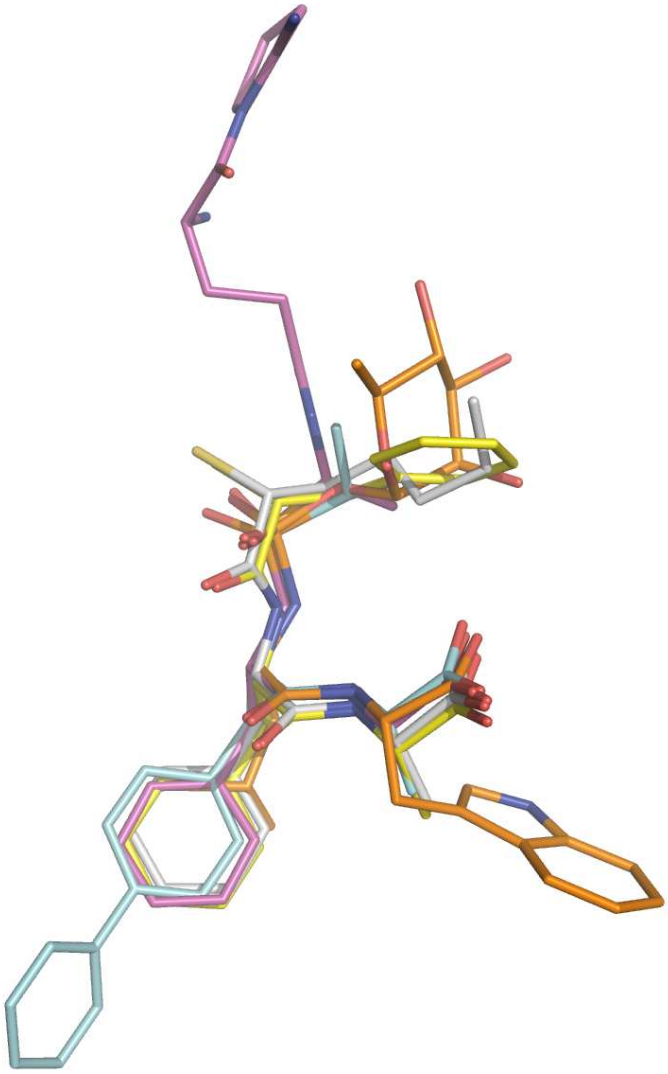
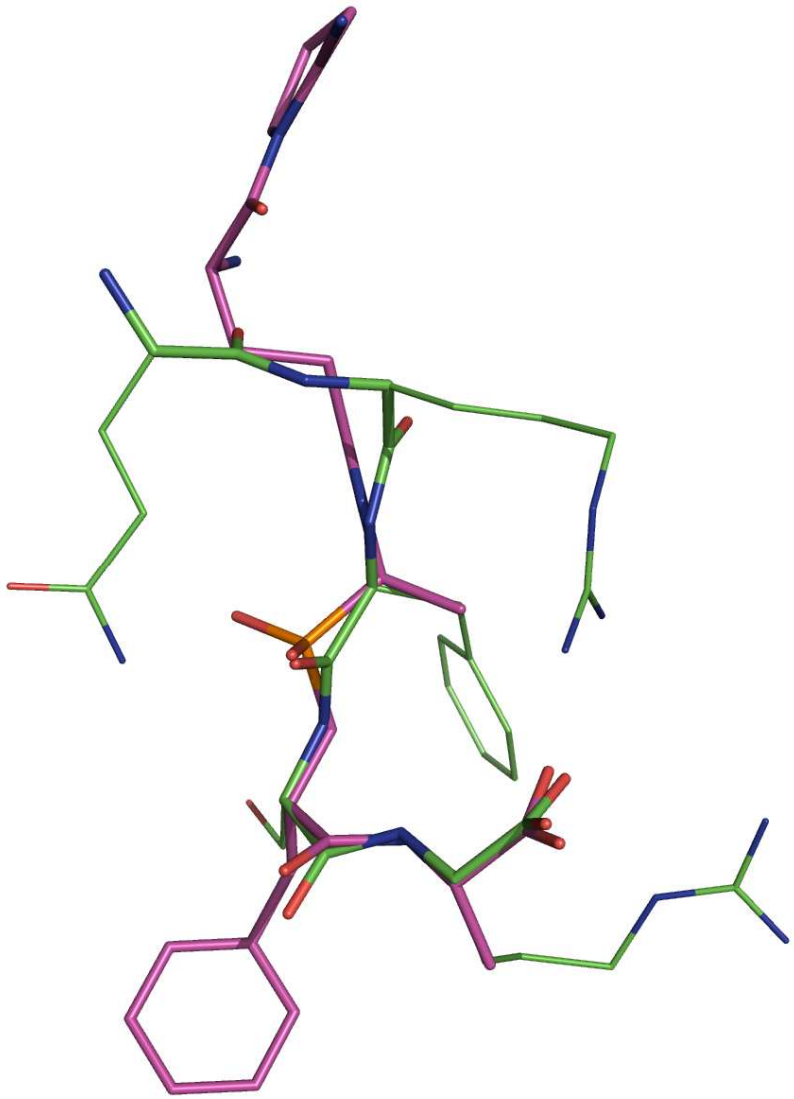


Figure S5. Structural comparison of Opiorphin's conformation 1.17 (in green) and I20 (in magenta), obtained by superposition of backbone atoms of the FSR fragment of conformation 1.17 and the C-terminal fragment of I20 (RMSD = 0.26 Å).



6. References

(1) Chan, W.C. and White, P.D. (eds) Fmoc solid phase peptide synthesis: a practical approach. Oxford University Press, Oxford, 2000.

(2) Mantle, D.; Hardy, M.F.; Lauffart, B.; McDermott, J.R.; Smith, A.I.; Pennington, R.J. Purification and characterization of the major aminopeptidase from human skeletal muscle. Biochem J. **1983**, 211, 567-73.

(3) Medeiros, M.A.; França, M.S.; Boileau, G.; Juliano, L.; Carvalho, K.M. Specific fluorogenic substrates for neprilysin (neutral endopeptidase, EC 3.4.24.11) which are highly resistant to serine- and metalloproteases. Braz J Med Biol Res. **1997**,30, 1157-1162.

(4) Filizola, M.; Centeno, N.B.; Perez, J.J. Computational Study of the Conformational Domains of Peptide T. *J. Pept. Sci.* **1997**, 3, 85-92.

(5) Case, D. A.; Darden, T. A.; Cheatham, T. E. III; Simmerling, C.L.; Wang, J.; Duke, R. E.; Luo, R.; Crowley, M.; Walker, R. C; Zhang, W.; Merz, K. M.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossvary, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Mathews, D. H. ; Seetin, M. G.; Sagui, C.; Babin, V.; Kollman P. A. **2008**, AMBER 10, University of California, San Francisco.

(6) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. A. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.* **2003**, *24*, 1999-2012.

(7) Ryckaert, J. P.; Bellemans, A. Molecular dynamics of liquid n-butane near its boiling point. *Chem. Phys. Lett.* **1975**, 30 (1), 123-125.

(8) Onufriev, A.; Bashford, D.; Case, D.A. Exploring protein native states and large-scale conformational changes with a modified generalized Born model. *Proteins*, **2004**, 55, 383–394.

(9) Sigalov, G.; Fenley, A.; Onufriev, A. Analytical electrostatics for biomolecules: beyond the generalized Born approximation. *J. Chem. Phys.* **2006**, *124*, 124902.

(10) Sigalov, G.; Scheffel, P.; Onufriev, A. Incorporating variable dielectric environments into the generalized Born model. *J. Chem. Phys.* **2005,** *122*, 094511.

(11) Perez, J.J.; Centeno, N.B. New insights into the multiple minima proble. *J. Mol. Struct. (TEOCHEM)* **1996**, 370, 173.

(12) The PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC.

(13) Oefner, C.; D'Arcy, A.; Hennig, M.; Winkler, F.K.; Dale, G.E. Structure of human neutral endopeptidase (Neprilysin) complexed with phosphoramidon. *J. Mol. Biol.* **2000**, 296, 341-349.

(14) Oefner, C.; Roques, B.P.; Fournié-Zaluski, M.C.; Dale, G.E. Structural analysis of neprilysin with various specific and potent inhibitors. *Acta Crystallogr. D Biol. Crystallogr.* **2004**, 60, 392-396.

(15) Oefner, C.; Pierau, S.; Schulz, H.; Dale, G.E.; Structural studies of a bifunctional inhibitor of neprilysin and DPP-IV. *Acta Crystallogr. D Biol. Crystallogr.* **2007**, 63, 975-981.

(16) Sahli, S.; Frank, B.; Schweizer, W.B.; Diederich, F.; Blum-Kaelin, D.; Aebi, J.D.; Böhm, H.J. Second-generation inhibitors for the metalloprotease neprilysin based on bicyclic heteroaromatic scaffolds: Synthesis, biological activity, and X-ray crystal-structure analysis. *Helv. Chem. Acta.* **2005**, 88, 731-50.