An *in silico* and chemical approach towards small protein production and application in phosphoproteomics

Ana M.G.C. Dias [a,b], Olga Iranzo [b]* and Ana C.A. Roque [a]*

Received (in XXX, XXX) Xth XXXXXXXXX 20XX, Accepted Xth XXXXXXXXX 20XX
DOI: 10.1039/b000000x

---

[a] *UCIBIO, REQUIMTE, Departamento de Química, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Campus Caparica, 2829-516 Caparica, Portugal, Fax: (+351) 21 294 8550; E-mail: cecilia.roque@fct.unl.pt*

[b] *Aix Marseille Université, Centrale Marseille, CNRS, iSm2 UMR 7313, 13397 Marseille, France; E-mail: olga.iranzo@univ-amu.fr, Previous address: Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, 2780-157 Oeiras, Portugal*

Supplementary Information:

Content List:

Pages

S1 - *In silico* studies

S1.1. - Molecular Dynamics Simulations - Stability studies - hPin1_WWdomain structures 2

S1.2. - Molecular Docking - Affinity Studies - hPin1_WWmutated 3

S2. - Peptide Solid-phase Chemical Synthesis and Characterization

S2.1. - Conditions for the standard and pseudoproline units protocol for peptide chemical synthesis 4

S2.2. - Reverse-Phase HPLC hPin1_WWmutated 5

S2.3 - Electrospray Ionization-Mass spectra (ESI-MS) 6

ESI-MS hPin1_WWmutated 6

S1 - *In silico* Studies:

S1.1. - Molecular Dynamics simulations - Stability studies - hPin1_WW domain structures:

The protocols for preparation of the structure for in *silico studies* can be found in the main text.

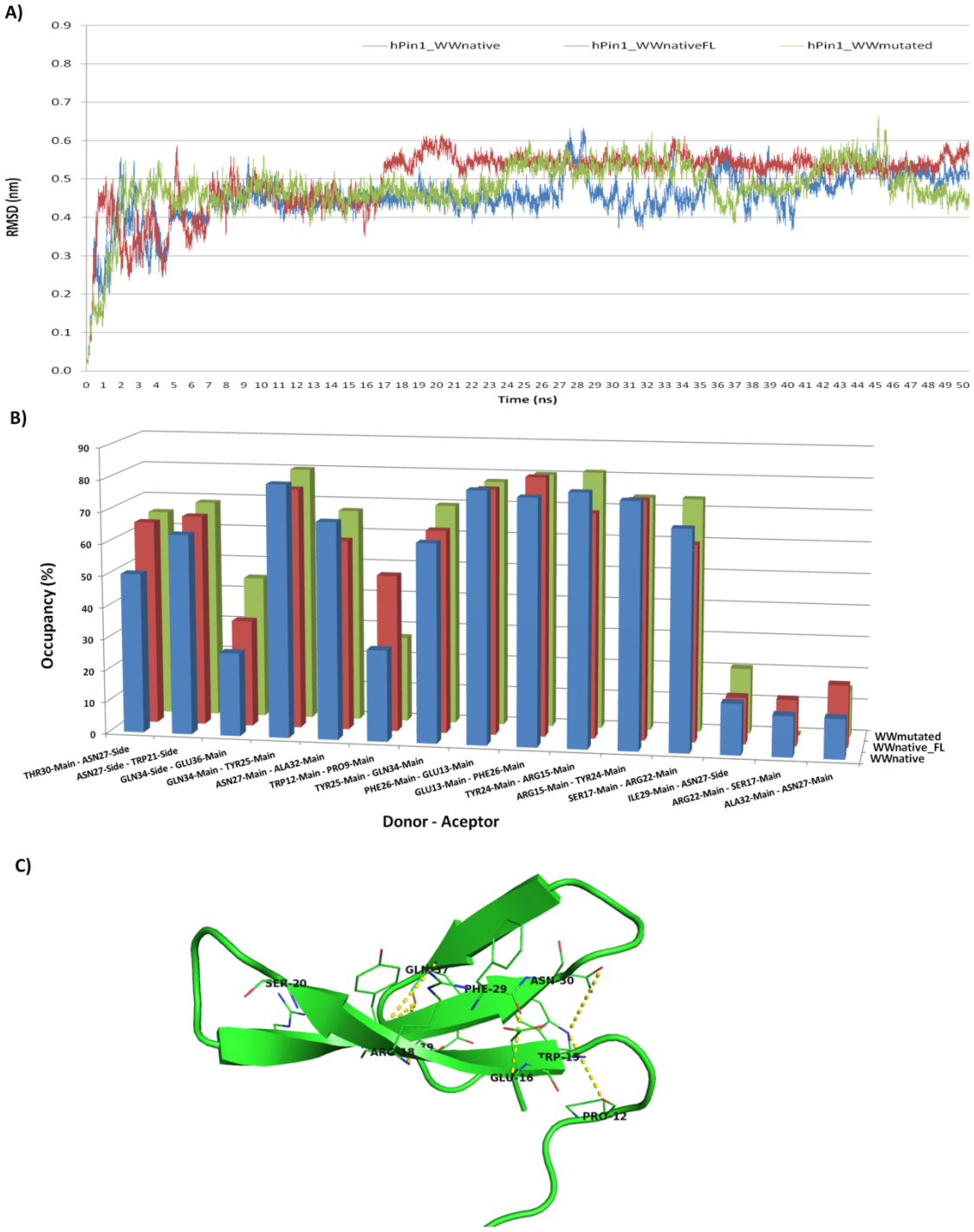

Figure S1. Molecular Dynamics results analysis for the 3 structures: hPin1_WWnative, hPin1_WWnativeFL and hPin1_WWmutated: A) RMSD (Root Mean Square Deviation); B) H-Bond interactions, only shown the most frequently observed over the time of the simulation; C) Identification of the most frequent H-Bond interactions observed in B (in hPin1_WWmutated structure). (Figure was produced using PyMol software).

S1.2. - Molecular Docking - Affinity studies - hPin1_WWmutated:

Table SI. Molecular Docking Results: hPin_WWmutated and different peptides.

Ligand	Docking Binding Energy (Kcal/mol)	Estimated Affinity Constant (M^{-1})	Literature Affinity Constant (M^{-1})	Source
PPPPYP	-10.09	2.49×10^{7}	n.d.	[a]
VPR **pT** PV	-7.31	2.28×10^{5}	2.04×10^{5}	Native ligand [b]
VPR **T** PV	-5.94	2.26×10^{4}	n.d.	-

[a] Native ligand shorter version of YAP65 WW domain (Group I) [1]
[b] Ligand described by Verdecia, M *et al* [2] with highest affinity constant with hPin1_WW.
n.d. - non determined

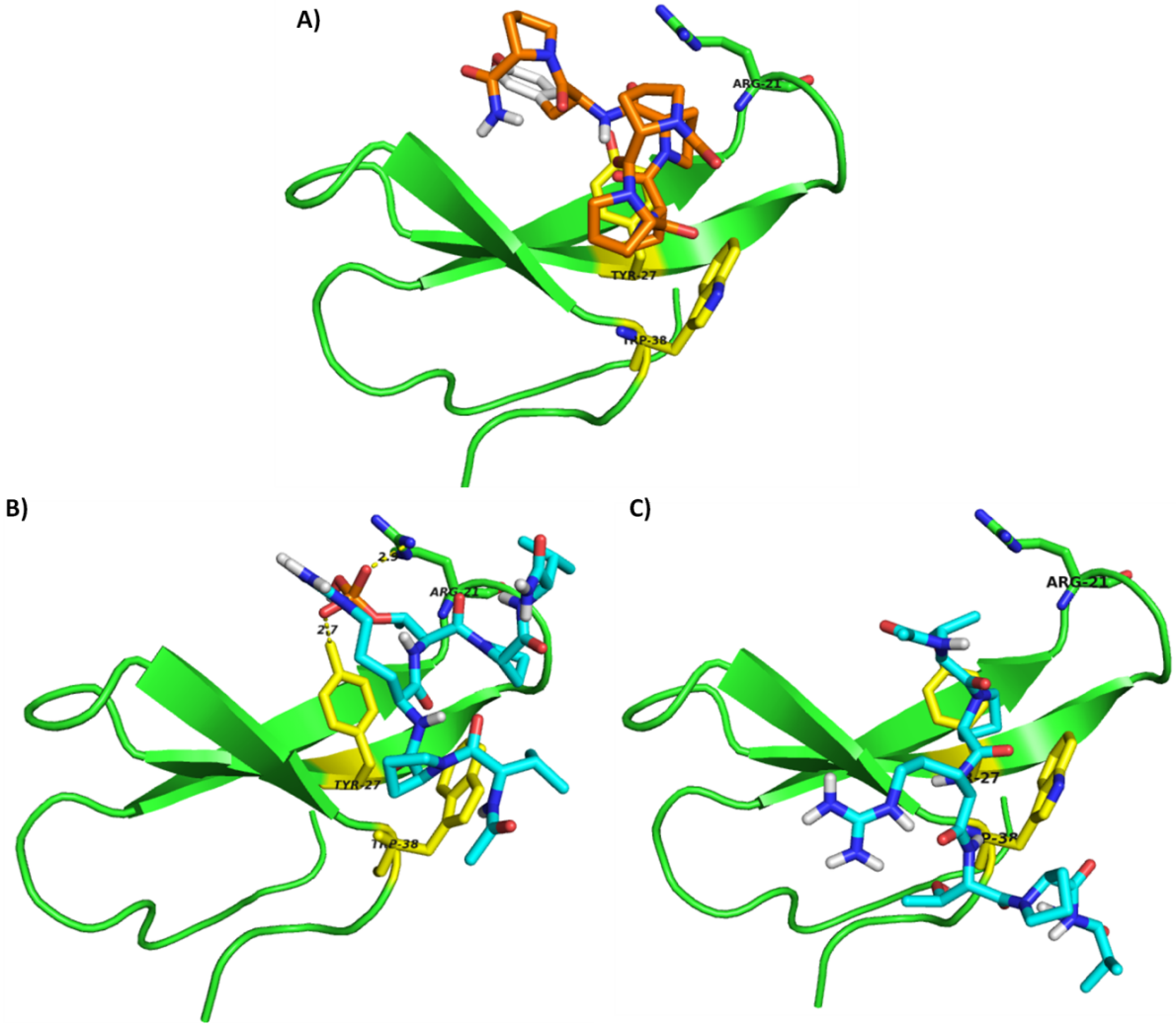

Figure S2. Molecular Docking results analysis, where hPin1_WWmutated was used as receptor and peptides, PPPPYP, VPRpTPV and VPRTPV were used as ligands. These conformations correspond to the values estimated in Table S1. In yellow the Tyr-27 and Trp-38 that form the "X-P-groove region". A) Best docking conformation of hPin1_WWmutated against PPPPYP, B) Best docking conformation of hPin1_WWmutated against VPRpTPV; and C) Best docking conformation of hPin1_WWmutated against VPRTPV. (Figures were produced using PyMol software).

S2. Peptide Solid-phase Chemical Synthesis and Characterization

S2.1 - Conditions for the standard and pseudoproline units protocol for peptide chemical synthesis

Table SII. Microwave Synthesis conditions for hPin1_WWmutated

Method		Reagents	Power (watts)	Temperature (ºC)	Reaction Time (sec)
Standard	Resin	Rink Amide MBHA resin 100-200 mesh	-	-	-
	Deprotection	20% piperidine/DMF	35	75	217
	Coupling	HOBT/HBTU/DIEA/DMF	25	75	304
	Coupling Arg (double coupling)	HOBT/HBTU/DIEA/DMF	0[a)] and 25[b)]	50	1805[a)] and 305[b)]
	Coupling His, Cys, Asp	HOBT/HBTU/DIEA/DMF	0	50	733
	Capping	10% Acetic Anhydrid	1	75	126
Pseudoproline Units	Resin	NovaPEG Rink amide resin	-	-	-
	Deprotection	20% piperidine/DMF	35	75	217
	Deprotection after 23 residue	20% piperidine/DMF	35	75	341
	Coupling	HBTU/DIEA/DMF	25	75	304
	Coupling Arg (double coupling)	HBTU/DIEA/DMF	0[a)] and 25[b)]	50	1950[a)] and 400[b)]
	Coupling after 23 residue	HBTU/DIEA/DMF	25	75	605
	Capping after 23 residue	10% Acetic Anhydrid	1	75	126
	Note: Arginine residue requires double coupling: a) 1st coupling and b) 2nd coupling.				

S2.2 Reverse-Phase HPLC hPin1_WWmutated:

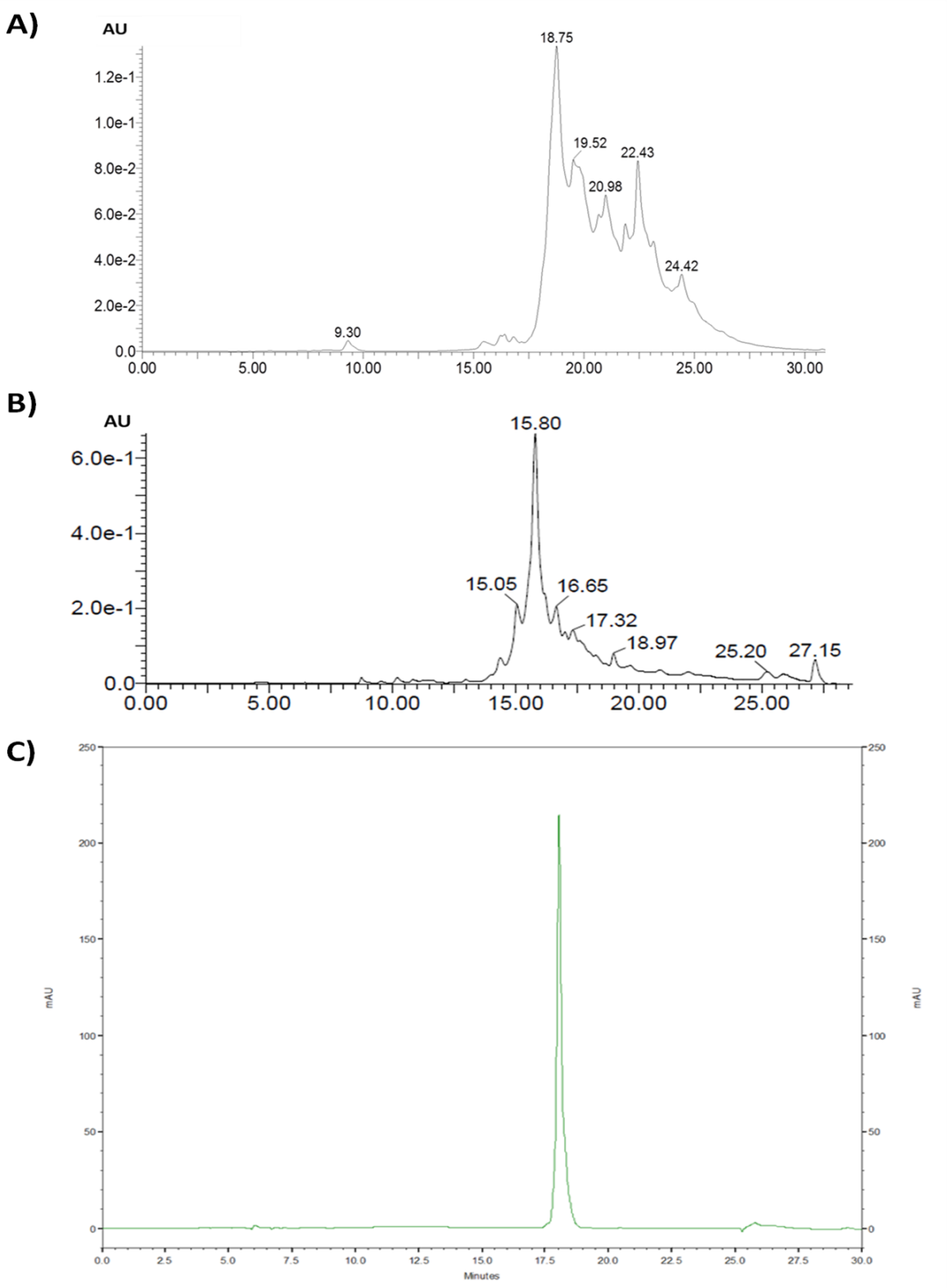

Figure S3. Reverse-Phase HPLC chromatograms: A) Preparative HPLC of the crude product of hPin1_WWmutated obtained by standard SPPS (without pseudoproline); B) Preparative HPLC of crude product of hPin1_WWmutated obtained by SPPS with pseudoproline strategy; C) Analytical HPLC chromatogram of purified hPin1_WWmutant. HPLC conditions used are described in the main text.

S2.3 - Electrospray Ionization-Mass spectrometry (ESI-MS)

The ESI-MS spectra of hPin1_WWmutated (Figure S3) were acquired in positive mode on a mass spectrometer API-linear Ion trap (PO03MS), model LTQ (Thermo-Finnigan) by direct infusion at a flow rate of 5 μL/min. All the spectra were obtained by the Mass Spectrometry Laboratory of the Analytical Services Unit at the Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa.

ESI-MS hPin1_WWmutated:

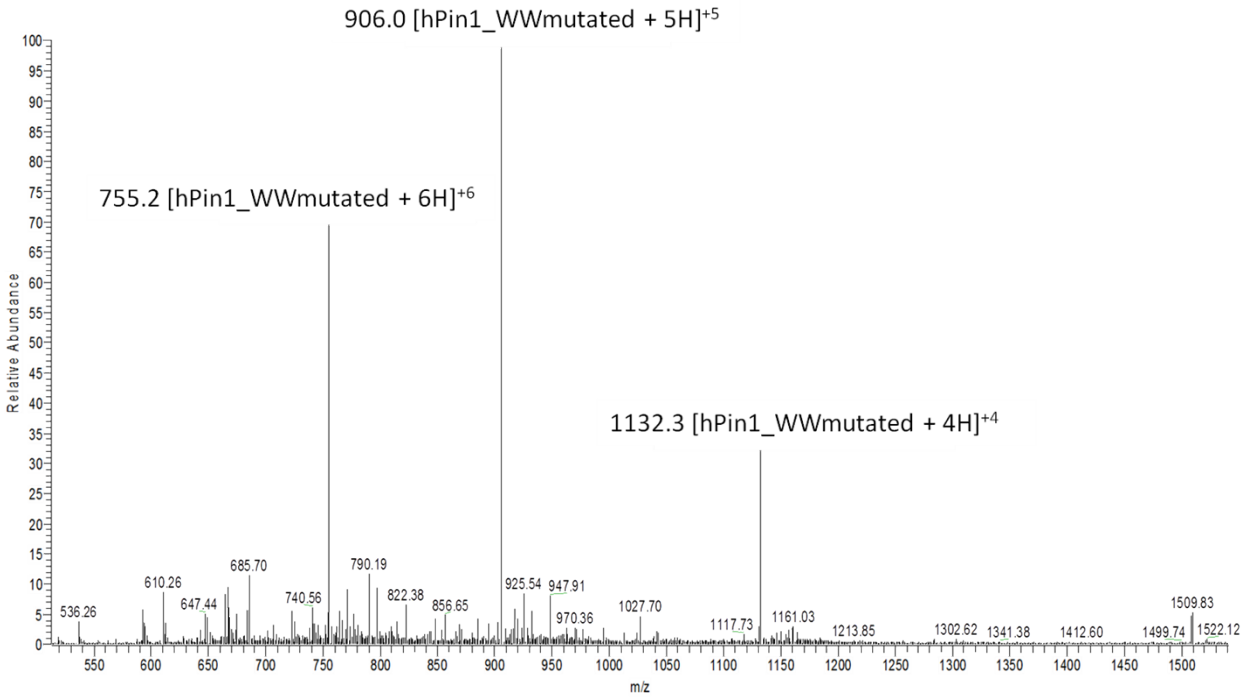

Figure S4. ESI-MS spectrum of hPin1_WWmutated (Acetyl - CADEEKLPPGWEKRMSRSSGRVYYFNHITNASQWE RPG-NH_2, Molecular Formula: $\mathbf{C_{198}H_{296}N_{60}O_{59}S_2}$) in positive mode: m/z calcd for [hPin1_WWmutated + 6H]$^{+6}$ = 755.18 Da, m/z found: 755.2; m/z calcd for [hPin1_WWmutated + 5H]$^{+5}$ = 906.01 Da, m/z found: 906.0 and m/z calcd for [hPin1_WWmutated + 4H]$^{+4}$ = 1132.26 Da, m/z found: 1132.3.

References:

1. E. K. Koepf, H. M. Petrassi, M. Sudol and J. W. Kelly, Protein Science, 1999, **8**, 841-853.

2. M. A. Verdecia, M. E. Bowman, K. P. Lu, T. Hunter and J. P. Noel, Nat Struct Mol Biol, 2000, **7**, 639-643.