Electronic Supplementary information (ESI)

Site-Selective modification of Tryptophan and Protein Tryptophan residues through PdNPs bionanohybrid-catalysed C-H Activation in Aqueous media

Carlos Perez-Rizquez,[a] Olga Abian,[b,c] and Jose M. Palomo[a]*,

[a] Department of Biocatalysis. Institute of Catalysis (CSIC). Marie Curie 2. Cantoblanco. Campus UAM, 28049 Madrid, Spain. Fax: +34-91-585-4760.

E-mail: josempalomo@icp.csic.es

[b]Instituto Aragonés de Ciencias de la Salud (IACS), Zaragoza, 50009, Spain

[c]Institute of Biocomputation and Physics of Complex Systems (BIFI), Joint Units IQFR-CSIC-BIFI, and GBsC-CSIC-BIFI, Universidad de Zaragoza, Zaragoza, 50018, Spain.

Experimental Section

General

Candida antartica B lipase (Cal-B) solution was from Novozymes (Denmark). 4-Methoxybenzenediazonium tetrafluoroborate (**2**), p-nitrophenylpropionate (pNPP), ethyl acetate and palladium acetate were from Sigma-Aldrich. N-Acetyl-L-tryptophan methyl ester (**1**) was bought from Alfa Aesar. The scanning electron microscopy (SEM) imaging was performed on a TM-1000 (Hitachi) microscope. The X-Ray diffraction (XRD) pattern was obtained using a Texture Analysis Diffractometer D8 Advance (Bruker) with Cu Kα radiation. The transmission electron microscopy (TEM) analysis were performed on a JEOL 2100F microscope (Oxford Instruments). Inductively coupled plasma atomic emission spectrometry (ICP-OES) was performed on a Perkin Elmer OPTIMA 2100 DV equipment to determine the amount of Pd on the biohybrid. To recover the biohybrids, a Biocen 22 R (Orto-Alresa, Spain) refrigerated centrifuge was used. HPLC spectrum P100 (Thermo Separation products) was used. Analyses were run at 25°C using an L-7300 column oven and a UV6000LP detector. The spectrophotometric analyses were run on a V-730 spectrophotometer (JASCO, Japan).

Synthesis of palladium nanoparticles enzyme (PdNPs-E) biohybrid

$Pd(OAc)_2$ (50 mg) were dissolved in 10 mL of MeOH. This solution was added to 40 mL of a distilled water solution containing 16 mg of Cal-B protein (1.6 mL of commercial Novozyme solution of 10 mg protein/mL. The final solution was kept under gentle magnetic stirring for 24 hours at room temperature. After that, the resulting suspension was separated by centrifugation (10,000 rpm; 4°C; 15 min). The recovered pellet was washed once with 10 mL of distilled water containing 20% (v/v) of MeOH and twice with distilled water (2 x 10 mL). After this, the suspension was directly lyophilized to obtain

catalyst in powder for later use**.** The solid was characterized by SEM, TEM and XRD. The content of Pd in the solid was 23%, determined by ICP-OES.

General procedure for C-H activation of 1

0.192 mmol (50 mg) of **1** and 0.192 mmol (42 mg) of **2** were added to a glass flask containing 5 mL of solvent. Solution was left under magnetic stirring until homogenization. Thereafter, a sample was taken and 2 mg of $Pd(OAc)_2$ or PdNPs-E biohybrid were added. The mixture was kept at room temperature for the indicated time. Also reactions were performed using PdNPs-E as catalyst reducing the initial amount of **1** from 5 to up to 20 times using methanol:water (1:1) as solvent. The reaction outgoing was monitored by HPLC analysis of the reaction's samples withdrawn at different times. Samples (100 μL) were centrifuged and then 50 μL were diluted 40 times in bi-distilled water before injection. The analysis conditions were performed with a Kromasil-C8 (150 × 4.6 mm and 5 μm Ø), at a flow of 1.0 mL / min; λ: 270 nm; and a mobile phase: 50% (v/v) ACN in MilliQ water.

Site-specific modification of proteins with 2

One mL of commercial Cal-B solution (10 mg) was dissolved in 4 mL of distilled water. Then, a 5 mL water solution of **2** (4.42 mg, 0.02 mol) was added to the enzyme solution. After that, 2 mg or 10 mg of PdNPs-E biohybrid was added to the solution. The mixture was kept under gentle stirring for 48 hours at room temperature. After that, the catalyst was separated by centrigugation and the solution was purified, washing 10 times with water to completely remove **2,** and concentrated using Ultracel®30kDa Centrifugal Filters (Merck) to finally recovering 1 mL of protein solution. The modified proteins were Cal-B-T-1 (using 2 mg catalyst) and Cal-B-T-2 (using 10 mg catalyst), respectively.

Control reactions were performed at the same conditions using Cal-B only in the presence of **2** (Cal-b*) or Cal-B only in the presence of PdNPs biohybrid, without **2** (Cal-B**). All proteins were characterized by MALDI-TOF.

Circular dichroism

Circular dichroism (CD) spectra of the different lipases were recorded in a Chirascan spectropolarimeter (Applied Photophysics) at 25(±1)°C. Near-UV spectra were recorded at wavelengths between 250 and 310 nm in a 1 cm path-length cuvette. Protein concentrations were 20 and 10 μM respectively in phosphate buffered saline, pH 7.2 (PBS; bioMerieux).

Fluorescence spectroscopy

Fluorescence measurements were performed in a Varian Cary Eclipse Fluorescence Spectrophotometer (Agilent Technologies) monitoring the intrinsic tryptophan fluorescence in 2 μM of different ROL variants and peptide conjugates solutions, using an excitation wavelength of 280 nm, with excitation and emission bandwidths of 5 nm, and recording fluorescence emission spectra between 300 and 400 nm. All spectroscopic measurements were made in water.

MALDI-TOF-MS spectra

MALDI-MS spectra of the different protein conjugates were recorded in a Bruker MicroFlex (Bruker Daltonics) by using 3-5-Dimethoxy-4-hydroxycinnamic acid (Sinapinic acid) as matrix. 1 μL of pure protein sample were mixed with 1 μl of a saturated sinapic solution (50:50 water/acetonitrile with 0.01% TFA final concentration). The mixture was spotted onto the surface of a stainless steel MALDI target and dried at ambient temperature.

Protein identification by MALDI and data-mining

MALDI-TOF MS analysis was performed using a 4800 Plus MALDI TOF/TOF MS analyzer (AB SCIEX). The measurements were carried out in linear positive ion mode. The instrument were equipped with a Nd:YAG laser operated a 325 nm and firing with 200 HZ repetition rate and up to 500 shots were accumulated for each spectrum. The laser intensity was set at 4900. A mixture of standard proteins (MSCAL3 from Sigma-Aldrich) was used to externally calibrate the instrument. The sample data were acquired using 4000 Series Explorer software (version 3.7.0 build 1; AB SCIEX) and visualized in a Data Explorer R software (version 3.7; build 126; AB SCIEX). Proteins were identified with the search engine Mascot using the SwissProt database. Search parameters used were: missed cleavage 1, fixed modifications carbamidomethyl (cysteines), variable modification oxidation (Met), peptide and fragment mass tolerance 0.2 Da and 0.3 Da, respectively.

Enzymatic activity assay

The activities of the soluble lipase, supernatant and enzyme suspension were analyzed spectrophotometrically measuring the increment in absorbance at 348 nm produced by the release of *p*-nitrophenol (pNP) ($\in$= 5.150 $M^{-1}cm^{-1}$) in the hydrolysis of 0.4 mM p-nitrophenylpropionate (*p*NPP) in 25 mM sodium phosphate at pH 7 and 25 °C. To initialize the reaction, 20 µL of protein solution was added to 2.5 mL of substrate solution in magnetic stirring. Enzymatic activity is given as µmol of hydrolyzed *p*NPP per minute per mg of enzyme (IU) under the conditions described above.

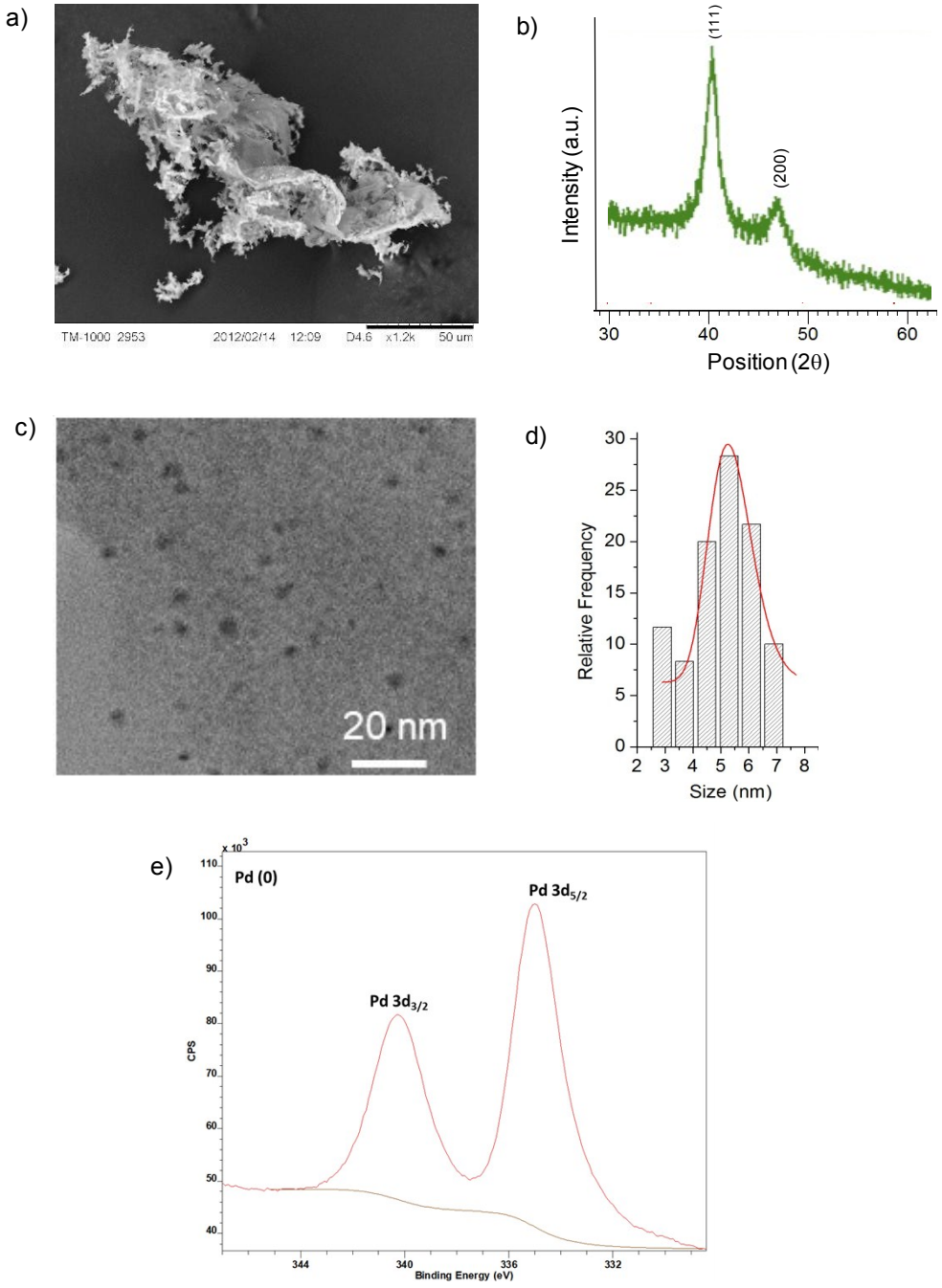

Figure S1. Characterization of PdNPs-E biohybrid. a) SEM, b) XRD pattern, c) TEM, d) Pd nanoparticles size distribution, e) XPS spectrum of Fe2p.

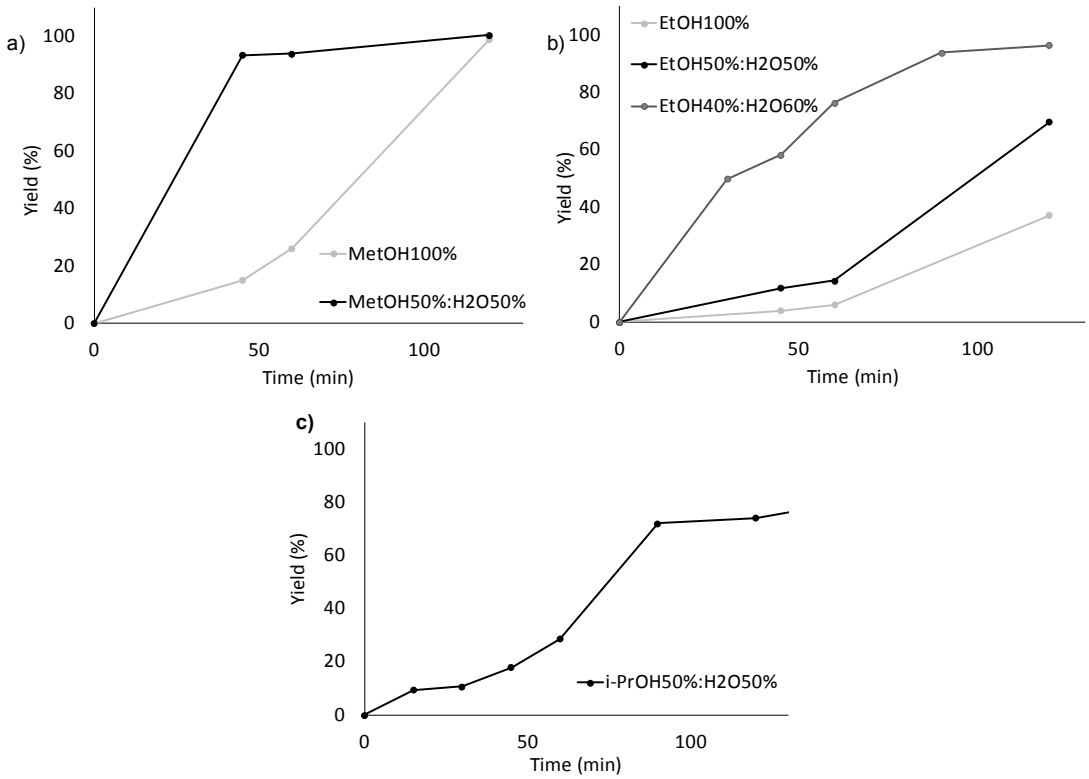

Figure S2. Reaction course of the arylation of 1 catalyzed by PdNPs-E biohybrid at different conditions at r.t. a) methanol solutions, b) ethanol solutions, c) isopropanol solution. Conditions: **1** (0.192 mmol), **2** (0.192 mmol), solvent (5 mL), PdNPs-E (2 mg), r.t (c.a 20°C). Yields of the product **3** were quantified by HPLC.

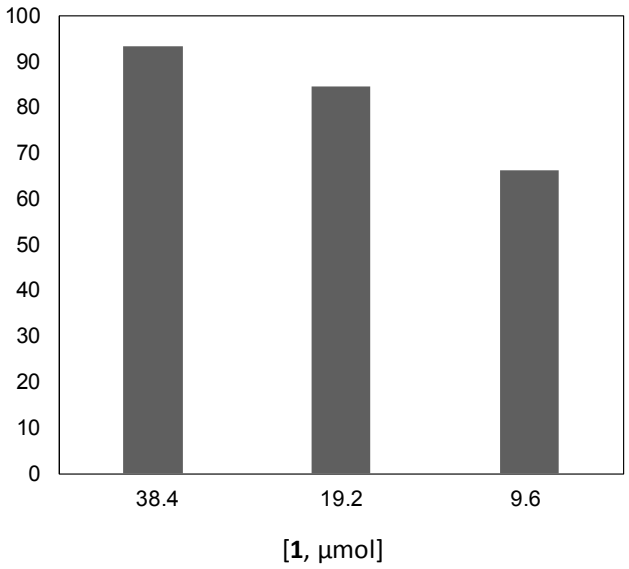

Figure S3. Effect of reducing the [1] in the C-H activation reaction. Conditions: **2** (0.192 mmol), MeOH:H_2O (5 mL), PdNPs-E (2 mg), r.t (c.a 20°C), 2h. Yields of the product **3** were quantified by HPLC.

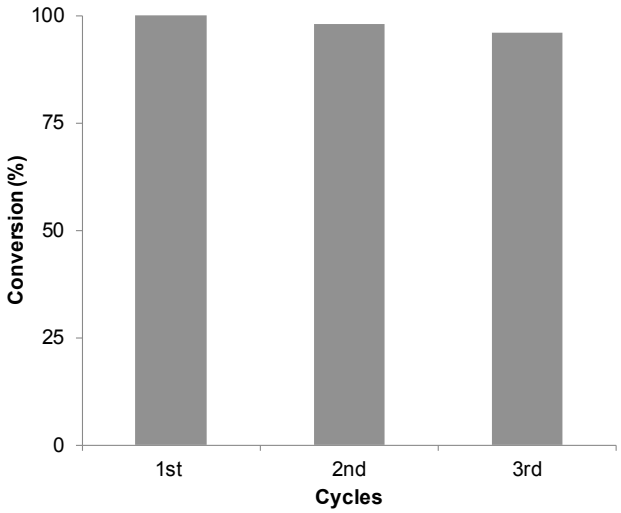

Figure S4. Reutilization studies of PdNPs-E biohybrid catalyst in the arylation of 1 in MeOH:water (1:1).

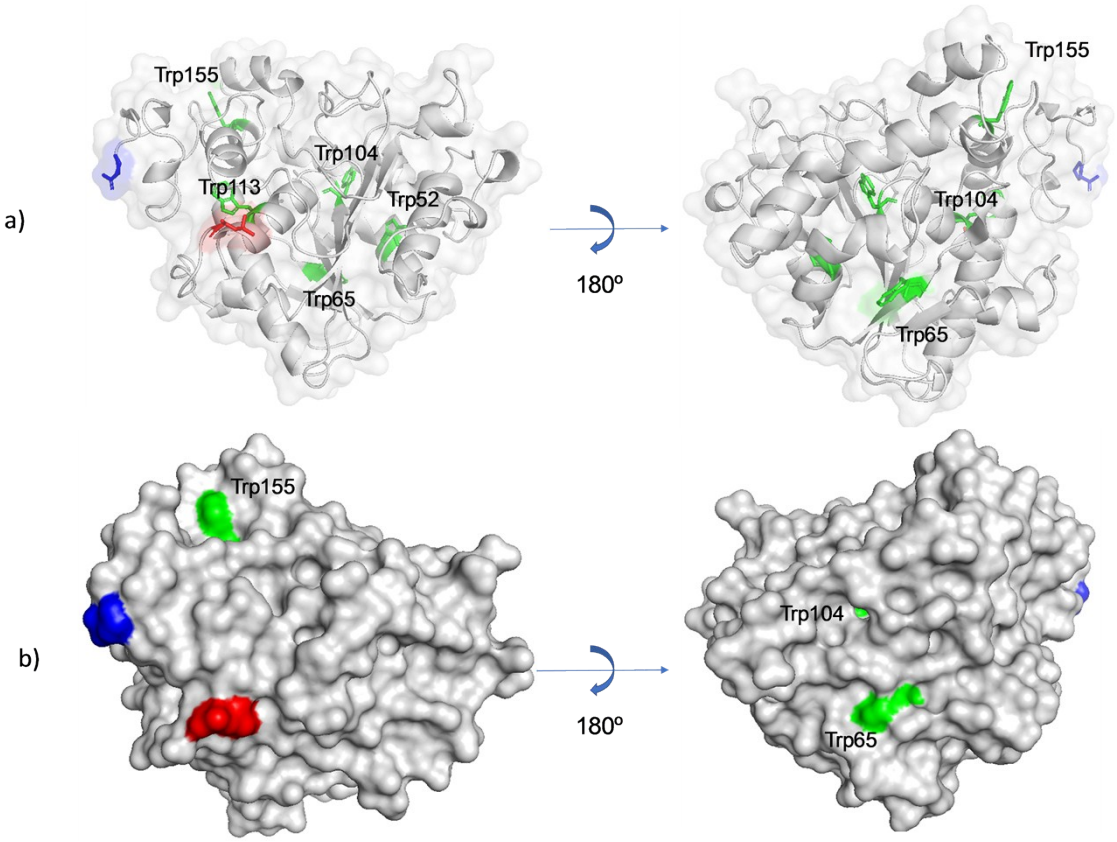

Figure S5. a) Ribbon representation of the structure of Cal-B with the Trps contained in the protein marked in green. N-terminus (red). C-terminus (blue). b) Tridimentional surface structure of Cal-B. The 3D structure was obtained from the Protein Data Bank (PDB) using Pymol vs 2.0.3. The pdb code for Cal-B is TCA.

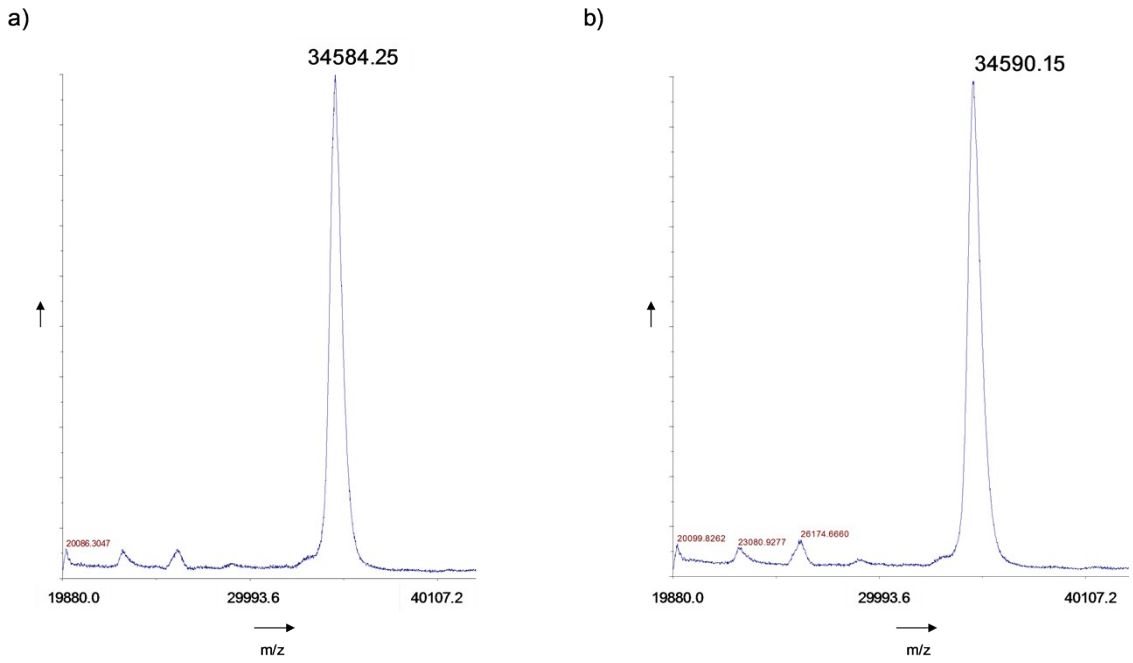

Figure S6. **MALDI-TOF spectra** of a) Cal-B* and b) Cal-B**

a) **Complete Cal-B aminoacids sequence:**

LPSGSDPAFSQPKSVLDAGLTCQGASPSSVSKPILLVPGTGTTGPQSFDSNWIPLSTQLGYTPCWISPPPFMLNDTQVNTEYMVNAITALYAGSGNNKLPV
LTWSQGGLVAQWGLTFFPSIRSKVDRLMAFAPDYK**GTVLAGPLDALAVSAPSVWQQTTGSALTTALR**NAGGLTQIVPTTNLYSATDEIVQPQVSNSPLDS
SYLFNGKNVQAQAVCGPLFVIDHAGSLTSQFSYVVGRSALRSTTGQARSADYGITDCNPLPANDLPEQKVAAAALLAPAAAAIVAGPKQNCEPDLMPYARP
FAVGKRTCSGIVTP

b) **Identified peptides of Cal-B**

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Expect	Rank	Unique	Peptide
20	3152.5547	3151.5474	3151.6877	-0.1403	0	183	4e-14	1	U	K.GTVLAGPLDALAVSAPSVWQQTTGSALTTALR.N
26	3220.4727	3219.4654	3219.6136	-0.1481	0	129	7.8e-09	1	U	K.NVQAQAVCGPLFVIDHAGSLTSQFSYVVGR.S

Trp155

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Expect	Rank	Unique	Peptide
8	2463.2126	2462.2053	2462.3857	-0.1803	0	37	26	1		GGSVQVAELISASQVLGLQQPLLR
15	2680.2480	2679.2407	2679.3214	-0.0807	1	32	68	1		ELKEETNLDAHAVSLIGVYEFMR + Oxidation (M)
10	2495.1814	2494.1741	2494.1006	0.0736	1	27	2.2e+02	1		GFHGYTCDVRSNIYTSSDMIR + Oxidation (M)
34	3933.8870	3932.8797	3932.8859	-0.0062	1	13	2.2e+03	1		QGQPIPEADEYFAQAMQSGLPDCAGVALGIDRLLMR + Oxidation (M)
18	3106.4333	3105.4260	3105.4713	-0.0453	1	13	3.6e+03	1		AEPLRDQNGEIVQWYVISVDIDDEMR + Oxidation (M)
30	3464.6606	3463.6533	3463.6278	0.0255	1	13	3.2e+03	1		AALDCLERGISVTHIISAMSEGMDICIDVCK
32	3647.6714	3646.6641	3646.8379	-0.1738	1	12	2.8e+03	1		RAAHGEQEGYEGLTINPVPLVAADSPYADLFVAAR
28	3372.4556	3371.4483	3371.5732	-0.1249	1	10	3e+03	1		APVALMPRMPCPWIQTPDATEMMLSPTSPV + 3 Oxidation (M)
23	3195.5249	3194.5176	3194.4840	0.0336	0	10	7.9e+03	1		AADYQPGGGGGMALFAGPDALDVTHTSSGRPR + Oxidation (M)

c) **Identified peptides of Cal-B-T1**

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Expect	Rank	Unique	Peptide
11	2673.3477	2672.3404	2672.4479	-0.1075	0	34	0.14	1	U	K.LPVLTWSQGGLVAQWGLTFFPSIR.S
17	3152.5508	3151.5435	3151.6877	-0.1442	0	170	2.7e-015	1	U	K.GTVLAGPLDALAVSAPSVWQQTTGSALTTALR.N
24	3220.4771	3219.4698	3219.6136	-0.1437	0	89	3e-007	1	U	K.NVQAQAVCGPLFVIDHAGSLTSQFSYVVGR.S

Trp104, Trp113

Trp155

Peptide matches not assigned to protein hits: (no details means no match)

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Expect	Rank	Unique	Peptide
8	2559.1587	2558.1514	2558.3162	-0.1648	1	20	3.3	1		QLQKNMQEAAELLDFELAAQLR
1	1720.6887	1719.6814	1719.8791	-0.1977	1	17	8.3	1		ADAPEEEDHVLVLRK
26	3372.4556	3371.4483	3371.5043	-0.0560	1	15	3.4	1		GEYTLVCPCSGKTVPDHYTLNCDCGSLIR
21	3180.5286	3179.5213	3179.6727	-0.1514	1	9	37	1		APLVTVDRLESYDIFALLHGGGPSGQAGALR
28	3647.7021	3646.6948	3646.7396	-0.0448	1	4	69	1		RPGKSHGAVLWMEYQLTPDSTVSTGLMNPAEDK + 2 Oxidation (M)

d) **Identified peptides of Cal-B-T2**

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Expect	Rank	Unique	Peptide
18	3152.5503	3151.5430	3151.6877	-0.1447	0	144	1.2e-012	1	U	K.GTVLAGPLDALAVSAPSVWQQTTGSALTTALR.N
25	3220.4673	3219.4600	3219.6136	-0.1535	0	100	2.2e-008	1	U	K.NVQAQAVCGPLFVIDHAGSLTSQFSYVVGR.S

Trp155

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Expect	Rank	Unique	Peptide
4	1476.5961	1475.5888	1475.6892	-0.1004	1	18	11	1		DFEAADEDPALRK
26	3248.4888	3247.4815	3247.5516	-0.0701	0	9	27	1		YLGEVDSTMTIIDVSMLTGFLPDTEDLTR + Oxidation (M)
16	3135.5798	3134.5725	3134.4444	0.1281	1	8	45	1		GPEDYLGGDGDRFLEIWNLVFMQYER + Oxidation (M)
22	3184.5415	3183.5342	3183.6021	-0.0679	1	8	50	1		SAFSINHVSRQVTSSGVSHGGTVSLQDAVTR
13	2705.3196	2704.3123	2704.2922	0.0202	0	6	95	1		FTEYLEDFHVWLSVYSRPSSSR
30	3647.7021	3646.6948	3646.7991	-0.1043	1	4	69	1		RPFGISCLIGGIDADGSARLFHTEPSGIFHEYK

Other peptides not identified, eg: Mw observed: 2808.2366

Expected [M+Na] of **LPVLTW(104)SQGGLVAQWGLTFFPSIRS**=2802.3404

Figure S7. **Peptide sequences profile from MALDI-TOF (MS/MS) obtained after the trypsin hydrolysis for Cal-B(b), Cal-B-T1(c) and Cal-B-T2(d)**. Mascot Score Histogram. Ions score is -10*Log(P), where P is the probability that the observed match is a random event. Individual ions scores > 38 indicate identity or extensive homology ($p<0.05$).

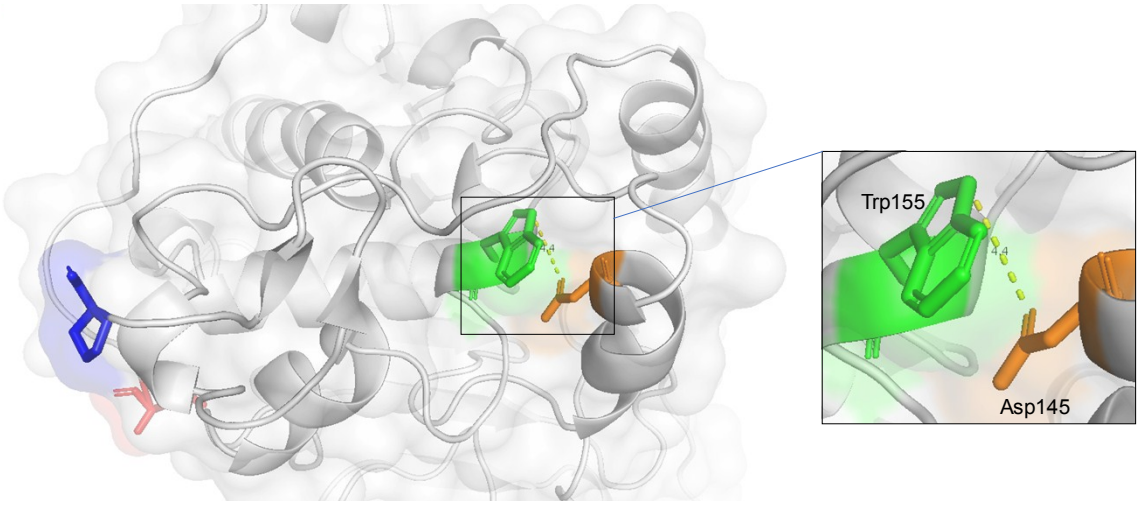

Figure S8. Representation of the Anion-π interaction between Trp155 and Asp145 in the structure of Cal-B. N-terminus (red). C-terminus (blue). The 3D structure was obtained from the Protein Data Bank (PDB) using Pymol vs 2.0.3. The pdb code for Cal-B is TCA.

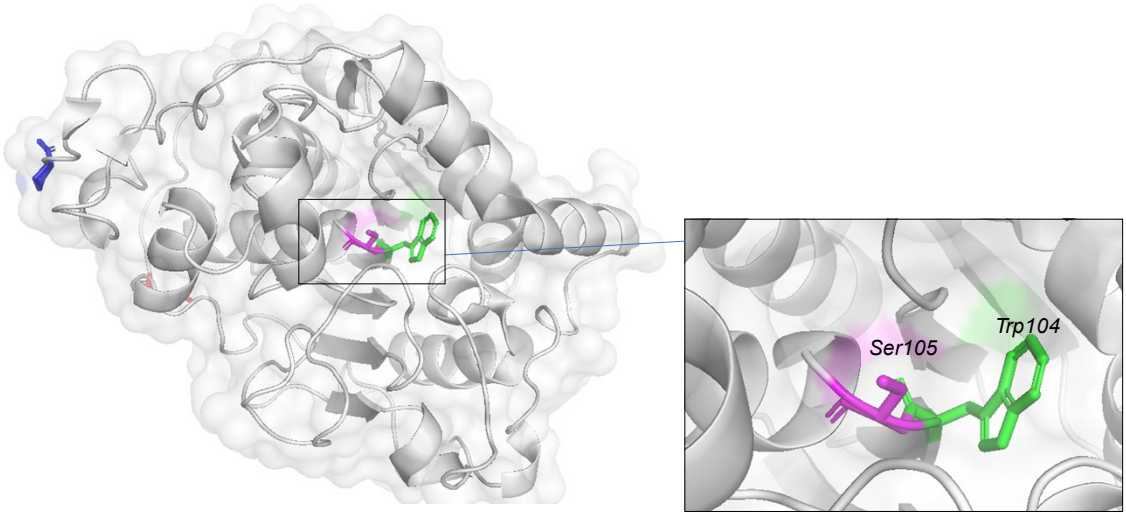

Figure S9. Representation of the Trp104 and catalytic Ser105 in the active site of structure of Cal-B. N-terminus (red). C-terminus (blue). The 3D structure was obtained from the Protein Data Bank (PDB) using Pymol vs 2.0.3. The pdb code for Cal-B is TCA.