Post-print of: (Paper) ChemCatChem 2014, 6, 1059-1065.

DOI: 10.1002/cctc.201300853

Aryl sulfotransferase from Haliangium ochraceum: a versatile tool for sulfation of small molecules

Iván Ayuso-Fernández,[a, b] Miquel A. Galmés,[a] Agatha Bastida,[a] and Eduardo García-Junceda*[a]

[a] *Departamento de Química Orgánica Biológica. Instituto de Química Orgánica General, CSIC*
Juan de la Cierva 3, 28006 Madrid (Spain)
Fax: (+34) 915-644-853
E-mail: eduardo.junceda@csic.es

[b] *Current address: Departamento de Biología Medioambiental. Centro de Investigaciones Biológicas, CSIC*
Ramiro de Maeztu 9, 28040 Madrid (Spain)

Received: October 8, 2013
Revised: December 13, 2013
Published online on March 26, 2014

Sulfation is an important molecular modification which regulates essential cellular processes and is also implicated in numerous pathological processes. The enzymes responsible for this reaction in living organisms are sulfotransferases. The gene Hoch_5094 from *Haliangium ochraceum* is annotated as a putative sulfotransferase. The protein codified by this gene —HocAST— was heterologous expressed in *E. coli* and showed arylsulfotransferase activity. CD analysis of HocAST showed a main α/β secondary structure that agrees with the overall structure of other cytosolic sulfotransferases. Interestingly, HocAST was able to use both *p*-NPS and PAPS as sulfuryl donors in a contrary manner to what happens with the arylsulfate sulfotransferase (ASST) that cannot use PAPS as donor. Regarding the specificity towards the acceptor, HocAST has shown quite a wide scope and was able to accept several mono- and dihydroxylated phenols and other phosphorylated compounds as substrates.

Introduction

Sulfation is an important molecular modification in all living organism which regulates essential cellular processes and is also implicated in numerous pathological processes.[1] The enzymes responsible for this reaction in living organisms are sulfotransferases (EC 2.8.2.-). These

enzymes catalyze the transfer of a sulfuryl moiety (SO_3^-) from the 3'-phosphoadenosine-5'-phosphosulfate (PAPS), to the hydroxyls and primary amines of a variety of acceptors. In eukaryotes two classes of sulfotransferases have been described: Golgi-membrane sulfotransferases (ST's) that sulfate larger molecules such as polysaccharides and proteins, and cytosolic sulfotransferases (SULT's) that transfer the sulfuryl group from PAPS to phenols, steroids, hormones, amines, and xenobiotics. This latter family is also known as phenol or aryl sulfotransferases (ArylST).[2] Besides these two families, a group of PAPS-independent sulfotransferases that use two aryl substrates, one as sulfuryl donor and the other one as acceptor, have been described in bacteria. Although the bacterial arylsulfate sulfotransferase (ASST) was first described in commensal intestinal bacteria as early as 1986, the identity of the natural donor has not been clearly established.[3]

From a synthetic point of view, STs are quite attractive because of their ability to sulphate, in a regio- and stereoselective manner, complex polysaccharides such as glycosaminoglycans (GAGs) that are polymers that play crucial roles in development and organogenesis and thus seem to be essential for multicellular life.[4] However, the synthetic use of ST is hampered due to the high cost and instability of PAPS and the problem of product inhibition caused by 3'-phosphoadenosine-5'-phosphate (PAP).[5] Besides their use for the sulfation of different flavonoids, steroids, peptides and aliphatic alcohols,[6] ArylST, in particular rat liver aryl sulfotransferase IV (Ast IV), has been used to develop a one-enzyme regeneration system for PAPS (Scheme 1).[7] This cycle is based in the reversibility of the reaction catalyzed by the ArylST.[8] Thus, when coupled to another PAPS-dependent sulfotransferase, the ArylST can transfer the sulfuryl group from *p*-nitrophenyl sulfate (*p*-NPS) to PAP regenerating PAPS *in situ* and overcoming the inhibition by PAP. On the other hand, scale-up of this regeneration system is hampered due to the poor stability of the pure Ast IV.[6c] Therefore, search for new ArylST with better properties as biocatalyst is still needed. In this sense, the gene Hoch_5094 from the bacteria *Haliangium ochraceum* codifies for a putative aryl sulfotransferase. Herein, we describe the heterologous expression of the gene Hoch_5094 in *Escherichia coli* and the biochemical characterization of the gene product as well as a preliminary analysis of its donor and acceptor specificity.

Results and Discussion

Sequence analysis of the aryl sulfotransferase from *Haliangium ochraceum*

In recent years the number of complete genomes, especially from bacteria which have been sequenced, has grown dramatically, providing a unique source for finding new enzymes. *Haliangium ochraceum* is the type species of the genus *Haliangium* in the myxococcal family *Haliangiaceae* and its genome has been sequenced recently.[9] The product of the gene Hoch_5094 is annotated as a putative sulfotransferase due to the sequence homology with the family SULT_1 (Pfam00685).[10] Since there is no experimental evidence of the activity of this enzyme we decided to make a comparative analysis of its sequence with the well characterized SULT from human (HsSULT1A1a), rat (RnSULT1A1) and mouse (MmSULT1D1). The domains that define the phenol sulfotransferase activity[11] are shown in figure 1:

Box A. This domain is known as the Phosphosulfurile Binding Site (PBS) and interacts with the 5'-phosphate from PAPS pointing the sulfuryl donor to the acceptor in the active site.

Box B. This highly conserved region includes the catalytic histidine (H105 in HocAST). It is expected that the amino acids in this domain contribute significantly to the structure of the active center.

Box C. In HsSULT1A1a R130 and S138 interact with the 3'-phosphate group of PAPS. In addition, S138 avoids the hydrolysis of PAPS in the absence of acceptor substrate.

Box D. This domain contains the motive GxxGxWK, highly conserved in this kind of sulfotransferases. This sequence interacts with the 3'-phosphate of PAPS and, together with domain C, stabilizes its binding to the active site. In HocAST there is an R instead of the K but since both amino acids are basic and have similar properties, it can be expected that the role of this domain is conserved in this enzyme.

We performed an automated protein structure homology modeling in the Swiss-Model[12] server using the protein SULT1D1 from M. musculus (PDB entry 2zyt as template).[13] The fold of HocAST is highly conserved when superposed to other sulfotransferases for which structures have been already solved. The topology of HocAST fold was □1, □1, □2, □3, □2 □4, □4, □5, □5, □6, □6, □7, □8, □9, □10, □11, □12, which was virtually identical to the topology of the SULT proteins (Figure 2A). In order to confirm that the catalytic residue of the HocAST could be the Hist105 (deduced by alignment) we carried out docking studies of the model protein with PAP and p-NPS obtaining a glide score of -8.59 kcal/mol (Figure 2B).[14] In the HocAST-PAP complex, the 5'-phosphate group of PAP is positioned in an analogous manner as does the phosphate group of PAPS in the crystallographic MmSULT1D1-PAPS complex (see SI, Figure 1S). In both cases the phosphate groups are at 3 Å from the catalytic Hist (residue 105 in HocAST and 108 in MmSULT1D1).

Therefore, since HocAST presented the main structural characteristics and an overall 3D structure in agreement with the cytosolic sulfotransferases, we decided to carry out its expression in E. coli to perform a detailed structural and functional analysis as well as to explore its synthetic applicability.

Expression and structural characterization of HocAST

The gene Hoch_5094 was amplified by PCR from the genomic DNA of *Haliangium ochraceum* using primers designed to complement specifically 25 bp at the 5' end of the codifying and complementary DNA strains. The recognition sequence for the restriction enzymes *Nde*I and *Xho*I were introduced in the amplification product during PCR. The resulting amplified band had the expected size (954 bp) and was digested and introduced in the expression vector pET-28b(+) which introduces a 6 histidines tag in the N-terminal of the recombinant protein.

E. coli BL21 (DE3) cells were transformed with this plasmid and the expression of recombinant protein was induced with 0.5 mM IPTG at OD_{600nm} 0.5-0.6. SDS-PAGE analysis of HocAST expression showed a band of the expected molecular mass (37 kDa) in the soluble fraction that represented 70% of the total protein. HocAST was purified by immobilized metal affinity chromatography (IMAC) through the His-tag.

Purified recombinant protein presented a mass of 37221 Da and the peptide mass fingerprint showed 18 peptides that cover the sequence of HocAST and identifies it unequivocally (Figure 3). The quaternary structure of the enzyme was analysed by size-exclusion chromatography. The recombinant protein in the chromatographic column was eluted in two peaks at 100 and 150 mL (see SI, Figure 2S). Peak 1 corresponds to the exclusion volume which contains oligomers of the protein. When the elution volume of peak 2 was interpolated into the calibration curve, weight of about 40 kDa corresponding to the monomer of HocAST was obtained (see SI, Figure 2S).

The spectroscopic characterization of purified HocAST is shown in Figure 4. The far-UV CD spectrum shows 2 minima in ellipticity at ~210 and ~220 nm, indicative of a main α/ □ secondary structure (Figure 4a). The tertiary structure of the protein in the environment of the fluorophores was analyzed by fluorescence emission spectroscopy of the 12 tryptophans and 12 tyrosines contained in the sequence of HocAST (Figure 4b). Upon excitation at 275 nm, the spectrum exhibited an emission maximum centered at 335 nm, shifted about 15 nm from the position of the maximum described for free tryptophan in aqueous solution (348–350 nm).[15] This indicates that the average microenvironment of the tryptophans in HocAST is significantly more hydrophobic than that of a residue completely exposed to the solvent. This blue-shift of the fluorescence, together with the circular dichroism spectrum, supports the idea that the protein is expressed properly folded. Moreover, emission spectra obtained upon excitation at 275 and 295 nm are practically identical, indicating an almost negligible contribution of tyrosines to the fluorescence spectrum of the protein. This contribution, lower than expected for a Trp:Tyr ratio of 12:12 is probably due to extensive resonance energy transfer from tyrosine fluorescence to nearby tryptophans and/or to quenching by other close side chains.

Biochemical characterization of HocAST

The time course of the reaction catalysed by the recombinant HocAST was followed in the sense of PAPS formation by monitoring the increase in absorbance at 405 nm due to the appearance of *p*-NP (**1**). The increase in absorbance was linear with enzyme concentration and reaction time. These results clearly indicated that HocAST was able to transfer the sulfuryl group from *p*-NPS to PAP (**2**) showing aryl sulfotransferase activity. The initial rate of the reaction measured over a range of substrate concentrations showed a Michaelis-Menten behavior (see SI, Figure 3S). Apparent kinetic parameters of HocAST enzyme were calculated independently for both substrates *p*-NPS and PAP (Table 1). It is difficult to compare these data with others reported in the literature since they are apparent parameters and then dependent on the conditions under which the assay was carried out. Nevertheless, they are of the same order as those reported for the sulfotransferase STF9 from *Mycobacterium avium*, another sulfotransferase which is also able to use PAPS and *p*-NPS as sulfuryl group donors.[16]

The recombinant HocAST exhibited activity between pH 6.0 and 7.5, with the maximum activity at pH 6.5. At pH's over 7.5 the activity drops drastically to virtually zero (see SI, Figure 4S). To study the thermal stability of the HocAST, the enzyme was incubated at 4, 25, 37 and 45 ºC over time and their remnant activities were evaluated at room temperature (Figure 5). We therefore determined in this assay the progressive loss of activity due to irreversible denaturation of the enzyme. The melting temperature of the recombinant HocAST was

determined by thermal shift experiment as 33 ºC (see Experimental section and Figure S5). Accordingly with this data, we observed a significant decreased of the half-life of the enzyme between 25 and 37 ºC (Figure 5). At 25 ºC the half-life of the sulfotransferase was over 3 h, but at 37 ºC it was only 36-37 min and at 45 ºC less than 6 min (Figure 5). On the other hand, at 4 ºC the enzyme kept over 90 % of activity for days.

Substrate specificity studies

For this preliminary study of the HocAST substrate specificity, we assayed different commercially available phenolic compounds, aliphatic alcohols and biologically relevant phosphorylated compounds that can be considered PAP analogous (Scheme 2).

A preliminary screening was carried out using a thermal shift assay.[17] The assay involves monitoring changes in the fluorescence signal of SYPRO orange dye as it interacts with the hydrophobic core of a protein undergoing thermal unfolding. The T_m value of the protein will increase in the presence of ligands that bind to the properly folded protein. In those cases where an increase of the T_m (ΔT_m) was detected, we carried out the sulfation reaction as described in the Experimental Section and the formation of the reaction product was confirmed by electrospray mass spectrometry (ESI-MS). The results are summarized in Table 2.

HocAST is able to transfer the sulfuryl group from *p*-NPS to the hydroxyl group of phenolic compounds with relatively good activity but aliphatic alcohols were not substrates of the enzyme. The best substrate was 4,4'-Biphenol (**11**) probably due to the symmetry of the molecule. In the case of catechol (**4**) we could detect not only the monosulfated product but the disulfated one as well ($[M+2]^+$ 191.2 and $[M+1]^+$ 269.8 respectively). The assayed nucleoside triphosphates were good acceptors for the enzyme and the obtained results indicate that the nucleoside moiety is needed to be recognized by the enzyme since triphosphate was neither ligand nor substrate. Finally, we assayed the capacity of HocAST to transfer the sulfuryl group from PAPS to three different nitrophenols (**1**, **8** and **9**). The activity with *p*-NP (**1**) was about half of the activity shown with *p*-NPS and PAP. With *o*-NP (**8**) the activity was quite similar but strongly decreased with dinitrophenol **9** with which showed a residual activity (about 2% of the activity showed with *p*-NPS).

Conclusion

In conclusion, we have shown experimentally that the gene Hoch_5094 from *Haliangium ochraceum* encodes for an arylsulfotransferase. CD analysis of HocAST showed a main α/ □ secondary structure that agrees with the overall structure of other cytosolic sulfotransferases. HocAST is a very versatile enzyme since it is capable of use both *p*-NPS and PAPS as sulfuryl donors. Regarding the specificity towards the acceptor, HocAST has shown a broad scope, being able to accept as substrates different phenolic compounds and nucleoside triphosphate as substrates. Therefore, HocAST is an interesting addition to the biocatalysts toolbox for enzymatic sulfation of compounds biologically relevant.

Experimental Section

Materials and General Procedures

E. coli BL21(DE3) competent cells were purchased from Stratagene Co. (San Diego, CA). Restriction enzymes, Taq polymerase and T4-DNA ligase were purchased from MBI Fermentas AB (Lithuania). PCR primers were purchased from Isogen Life Science (Spain) and the pET-28b(+) expression vector was purchased from Novagen. Isopropyl-□-D-thiogalactopyranoside (IPTG) was purchased from Applichem GmBH (Germany). Plasmids and PCR purification kits were from Promega (Madison, WI) and DNA purification kit from agarose gels was from Eppendorf (Hamburg, Germany). SDS-PAGE was performed using 10% and 5% acrylamide in the resolving and stacking gels, respectively. Gels were stained with Coomassie brilliant blue R-250 (Applichem GmBH, Germany). Electrophoresis was always run under reducing conditions, in the presence of 5% β-mercaptoethanol. Nickel-iminodiacetic acid (Ni^{2+}-IDA) agarose was supplied by Agarose Bead Technologies (Spain). All compounds were purchased from Sigma–Aldrich and were used directly. Solvents were of analytical grade. DNA manipulation was according to standard procedures.[18]

Expression and purification of the arylsulfotransferase from *Haliangium ochraceum*

Hoch_5094 gene was amplified by PCR on a iCycler Thermal Cycler (MiniOpticon, BioRad) using as template genomic DNA from *Haliangium ochraceum* obtained from DSMZ (German Collection of Microorganisms and Cell cultures, DSM nº 14365). Forward primer was 5'TATAACATATGAATTCTACTGACGAACAACACA3' (*Nde*I restriction site underlined) and the reverse primer was 5'CTCGAGTCAGTCCGGCAGCTCGCCG3' (*Xho*I restriction site underlined). PCR amplification was performed in 5μL reaction mixture containing DMSO 20% and subjected to 29 cycles of amplification. The cycle conditions were set as follows: denaturation at 94 ºC for 1 min, annealing at 50 ºC for 2 min and elongation at 72 ºC for 2 min. After digestion, the purified PCR product was ligated in the pET-28b(+) vector double digested with the same restriction enzymes.

For the expression of the Hoch_5094 gene, transformed *E. coli* BL21 (DE3) cells, carried on the pET-AST6xHis plasmid, were grown in LB medium supplemented with kanamycin (26 $\mu g \cdot mL^{-1}$) at 37 ºC. When the culture reached an OD_{600} of 0.5-0.6, protein expression was induced with isopropyl β-d-1-thiogalactopyranoside (IPTG, 0.4 mM) and the temperature was lowered to 30 ºC. After 18 hours, cells were harvested by centrifugation (3800 x g) for 30 min at 4 ºC. Cell pellets were re-suspended in monosodium phosphate buffer (25 mM; pH 7.4) and treated with lysozyme and DNase for protein extraction.[19] The solution was centrifuged for 40 min (8000 × g, 4 ºC) and collected for purification. The supernatant, containing the recombinant protein, was loaded onto a Ni^{+2}-IDA-agarose column pre-equilibrated with sodium phosphate buffer (20 mM, pH 7.5). The recombinant ArylST was eluted with the same buffer containing imidazole 500 mM. All the fractions containing protein were pooled together and dialyzed to remove the imidazole. Sodium dodecylsulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) showed a single band matching the expected molecular weight of the recombinant ArylST (37 kDa).

Sequence analysis and 3D modelization of the HochAST

To analyze the sequence of HocAST an alignment of 4 sequences was performed using ClustalX.[20] The modelization of HochAST was made in the SwissModel server with the best template option.

Protein analysis

Peptide-mass fingerprint analysis from the SDS-PAGE band corresponding to the hypothetical HocAST was performed at the Proteomic Unit of the Spanish National Center of Biotechnology (CNB-CSIC). Samples were digested with sequencing grade trypsin O/N at 37 ºC. The analysis by MALDI-TOF mass spectrometry produces peptide mass fingerprints and the peptides observed can be collated and represented as a list of monoisotopic molecular weights. Data were collected in the *m/z* range of 800-3600.

Quaternary structure of HocAST was determined by size-exclusion chromatography with no-denaturing conditions. To this extent, the protein was loaded into a HiLoad 26/60 SuperdexTM 75 PG column, controlled by the AKTA-FPLC system (GE-Healthcare Life Science). The column was developed in phosphate buffer (50 mM, pH 7.2) containing NaCl (0.15 M) at a constant flow rate of 1.0 mL·min^{-1}. A calibration curve was made with the retention times of well-known proteins with molecular weight (MW) between 15 kDa and 70 kDa.

Circular dichroism (CD) and fluorescence spectroscopy

Far-UV circular dichroism (CD) spectra were recorded in the wavelength range of 195-250 nm using a Jasco J-815 spectropolarimeter equipped with a constant temperature cell holder Jasco PTC423-S Peltier. The protein concentration was 14 □M. The optical p
The contribution of the buffer was always subtracted. For each sample, four spectra were accumulated at a scan speed of 20 nm/min with a bandwidth of 0.2 nm and averaged automatically. The mean residue ellipticity, [ ], is given in units of deg·cm^2·dmol^{-1}. A value of
110 g/mol was used as mean weight for residue.

Tryptophan fluorescence emission spectra of HocAST were recorded in a Jasco FP-8300 spectrofluorimeter and data were treated with the SpectraManager program. The contribution of the buffer was always subtracted. A quartz cell with a 1 cm path length in both the excitation (275 nm and 295 nm) and emission (295 to 400 nm) directions was used in all the experiments. The HocAST protein concentration was 10 □M in phosphate
pH=7.5) and experiments were carried out at 25 ºC.

Enzyme activity assays and steady-state kinetic analysis

HocAST activity measurement was carried out using a high throughput colorimetric 96-well plate assay. Sulfotransferase activity assays were run at room temperature following the increase of absorbance at 405 nm ($\varepsilon_{p\text{-NP}}$ = 18000 M^{-1} cm^{-1}) for 15 min. In a total volume of 250 □L of reaction mixture HEPES buffer (50 mM, pH 7.0), *p*-NPS (between 0 and 0.5 mM as indicated) and acceptor (between 0-0.02 mM as indicated) were added. The reactions were initiated upon addition of 50 µl of HocAST (14 □M). In all cases
addition of 50 □L of NaOH 1N. F
only one substrate), initial velocities (V_0) were fitted to the Michaelis-Menten equation.

Kinetics parameters were calculated using the built-in nonlinear regression tools in SigmaPlot 11.0

Substrate specificity studies

Thermal shift assays were carried out in a iQ5 Real Time Detection System (Bio-Rad,CA). To monitor the protein unfolding, the fluorescent dye Sypro orange was used. The dye's fluorescence signal is quenched in the aqueous environment of a properly folded protein in solution, but as the protein unfolds, it exposed its hydrophobic residues to the environment. The dye then binds to the hydrofobic regions and becomes unquenched. 5 □L of acceptor solutions (10 mM) were added to 45 □L of Tris buffer Sypro Orange and 4.4 □M of protein. Th 0.2 ºC·min^{-1}. The fluorescence intensity was measure with Ex/Em 490/530.

In those cases where a thermal shift was detected, sulfotransferase activity was confirmed by the colorimetric method described above. On the other hand, to determinate the activity of the HocAST using PAPS as sulfuryl donor, reaction mixtures contained 184 µl of phosphate buffer (50 mM, pH 7.0), 2.7 □L of PAPS (9 mM) and 7.5 □L were initiated upon addition of 50 µl of HocAST (14 □M).

In all cases, the identity of the reaction products was confirmed by Positive-ion ESI-MS spectrometry: m/z: 3'-phosphoadenosine-5'-phosphosulfate (PAPS) 506.9 [M+]; Adenosine-5'-triphosfate 5'-sulphate (ATPS) 587 [M+]; Guanosine-5'-triphosphate-5'-sulphate (GTPS) 603 [M+]; Citidine-5'-triphosphate-5'-sulphate (CTPS) 563.9 [M+]; Uridina-5'-triphosphate-5'-sulphate (UTPS) 564.7 [M+]; phenol sulfate 174.2 [M^{+}]; 4,4'-biphenol sulfate, 266.0 [M+]; Cathecol monosulphate 191.2 [M2+]; cathecol disulfate 269.8 [M]; *p*-Nitrophenolsulphate 245.0 [NaM+]; 2,4-Dinitrophenolsulphate 280.6 [NaM+]

Molecular docking studies

Crystal structure of the complex between human sulfotransferase SULT1A1 and PAP and from the mutant D249G with PAP and *p*-NP, were obtained from the PDB (entries 3u3j and 3u3r respectively). The receptor was prepared using the Wizard tool of the Schrödinger suit. The initial conformation of PAP and *p*-NP were taken from the crystal structures. Subsequently, they were prepared using LigPrep by modifying the torsions of the ligands and assigning its appropriate protonation states. In Glide, 32 stereochemical structures were generated per compound with possible states at target pH 7.0 ± 2.0 using Ionizer, tautomerized, desalted and optimized by producing low-energy 3D structure for the ligand under the OPLS 2005 force field[22] while retaining the specified chiralities of the input Maestro file. Then, receptor grid was calculated for the prepared receptor such that various ligand poses bind within the predicted site during docking. In Glide, grids were generated keeping the default parameters of van der Waals scaling factor 1.00 and charge cutoff 0.25 subjected to OPLS 2001 force field. A cubic box of specific dimensions (15 Å x 15 Å x 15 Å) center on the donor or acceptor was generated for the protein. After that, SP flexible ligand docking was carried out in Glide of Schrödinger-Maestro v9.2.7. Final scoring was performed on energy-minimized poses and displayed as Glide score. The best docked pose was recorded for the ligand.

Acknowledgements

We thank the Spanish Ministerio de Ciencia e Innovación (Grants PI11/01436 and CTQ2010-19073/BQU), Comunidad de Madrid (Grant S2009/PPQ-1752) and MAPFRE Foundation for financial support. We thank Dr. R. Wohlgemuth for the gift of PAPS and PAP. We thank the referees for their helpful comments.

References

[1] a) J. D. Mougous, R. E. Green, S. J. Williams, S. E. Brenner, C. R. Bertozzi, *Chem. Biol*., **2002**, *9*, 767-776; b) E. Chapman, M. D. Best, S. R. Hanson, C.-H. Wong, *Angew. Chem*., **2004**, *116*, 3610-3632; *Angew. Chem. Int. Ed.* **2004**, *43*, 3526-3548; c) V. L. Rath, D. Verdugo, S. Hemmerich, *Drug Discovery Today* **2004**, *9*, 1003-1011 , d) M. W. Schelle, C.R. Bertozzi, *ChemBiochem* **2006**, *7*, 1516-1524; d) P. Bojarova, S. J. Williams, *Curr. Opin. Chem. Biol.* **2008**, *12*, 573-581.

[2] a) K. G. Bowman, C. R. Bertozzi, *Chem. Biol*. **1999**, 6, R9-R22; b) M. Negishi, L. G. Pedersen, E. Petrotchenko, S. Shevtsov, A. Gorokhov, Y. Kakuta, L. C. Pedersen, *Arch. Biochem. Biophys*. **2001**, *390*, 149-157; c) M. W. Duffel, A. D. Marshal, P. McPhie, V. Sharma, W. B. Jakoby. *Drug Metab. Rev*. **2001**, *33*, 369-95; d) W. W. Cleland, A. C. Hengge, *Chem. Rev*. **2006**, *106*, 3252-3278; e) J. G. McCarthy, E. B. Eisman, S. Kulkarni, L. Gerwick, W. H. Gerwick, P. Wipf, D. H. Sherman, J. L. Smith, *ACS Chem. Biol*. **2012**, 7, 1994-2003.

[3] a) K. Kobashi, Y. Fukaya, D. H. Kim, T. Akao, S. Takebe, *Arch. Biochem. Biophys*. **1986**, *245*, 537–539; b) G. Malojcic, R. Glockshuber, *Antioxid. Redox Signaling*, **2010**, *13*, 1247-1259; c) X. Tang, K. Eitel, L. Kaysser, A. Kulik, S. Grond, B. Gust, *Nat. Chem. Biol*. **2013**, *9*, 610-615.

[4] a) Y. Xu, S. Masuko, M. Takieddin, H. Xu, R. Liu, J. Jing, S. A. Mousa, R. J. Linhardt, J. Liu, *Science* **2011**, *334*, 498-501; b) P. L. DeAngelis, J. Liu, R. J. Linhardt, *Glycobiology* **2013,** *23*, 764-777.

[5] C.-H. Lin, G.-J. Shen, E. García-Junceda, C.-H. Wong, *J. Am. Chem. Soc.* **1995**, *117*, 8031-8032.

[6] a) M. A. van der Horst, J. F. T. van Lieshout, A. Bury, A. F. Hartog, R. Wever, *Adv. Synth. Catal.* **2012**, *354*, 3501-3508; b) P. Marhol, A. F. Hartog, M. A. van der Horst, R. Wever, K. Purchartová, K. Fuksová, M. Kuzma, J. Cvačka, V. Křen, *J. Mol. Catal. B: Enzymat.* **2013**, *89*, 24-27; c) K. Purchartová, L. Engels, P. Marhol, M. Šulc, M. Kuzma, K. Slámová, L. Elling, V. Křen, *Appl. Microbiol. Biotechnol*. **2013**, DOI 10.1007/s00253-013-4794-0.

[7] M. D. Burkart, M. Izumi, C.-H. Wong, *Angew. Chem*., **1999**, *111*, 2912-2915; *Angew. Chem. Int. Ed.* **1999**, *38*, 2747-2750.

[8] a) F. C. Kauffman, M. Whittaker, I. Anundi, R. G. Thurman, *Mol. Pharmacol*. **1991**, *39*, 414-420; b) E. Tan, S. K. Pang, *Drug Metab. Dispos*. **2001**, *29*, 335-346.

[9] N. Ivanova, C. Daum, E. Lang, B. Abt, M. Kopitz, E. Saunders, A. Lapidus, S. Lucas, T. Glavina Del Rio, M. Nolan, H. Tice, A. Copeland, J. F. Cheng, F. Chen, D. Bruce, L. Goodwin, S. Pitluck, K. Mavromatis, A. Pati, N. Mikhailova, A. Chen, K. Palaniappan, M. Land, L. Hauser, Y. J. Chang, C. .D Jeffries, J. C. Detter, T. Brettin, M. Rohde, M. Göker, J. Bristow, V. Markowitz, J. A. Eisen, P. Hugenholtz, N. C. Kyrpides, H. P. Klenk, *Stand Genomic Sci*. **2010**, *2*, 96-106.

[10] Y. Kakuta, L. G. Pedersen, C. W. Carter, M. Negishi, L.C. Pedersen, *Nat. Struct. Biol*. **1997**, *11*, 904-908.

[11] N. Hempel, N. Gamage, J. L. Martin., M. E. McManus, Int. *J. Biochem. Cell Biol*. **2007**, *39*, 685-689.

[12] a) M. C. Peitsch, *Bio/Technology* **1995**, *13*, 658-660; b) K. Arnold, L. Bordoli, J. Kopp, T. Schwede, *Bioinformatics*, **2006**, *22*, 195-201; c) F. Kiefer, K. Arnold, M. Künzli, L. Bordoli, T. Schwede, *Nucleic Acids Res*. **2009**, *37*, D387-D392.

[13] T. Teramoto, Y. Sakakibara, M.-C. Liu, M. Suiko, M. Kimura, Y. Kakuta, *Biochem.Biophys.Res.Commun*. 2009, **383**, 83-87.

[14] a) R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis, P. S. Shenkin. *J. Med. Chem*. **2004**, *47*, 1739-49; b) T. A. Halgren, R. B. Murphy, R. A. Friesner, H. S. Beard, L. L. Frye, W. T. Pollard, J. L. Banks, *J. Med. Chem*., **2004**, *47*, 1750-1759; c) R. A. Friesner, R. B. Murphy, M. P. Repasky, L. L. Frye, J. R. Greenwood, T. A. Halgren, P. C. Sanschagrin, D. T. Mainz, *J. Med. Chem*., **2006**, *49*, 6177-6196.

[15] J. R. Lakowicz, *Principles of Fluorescence Spectroscopy*. Plenum Press, N. Y. **1999**.

[16] Md. M. Hossain, Y. Moriizumi, S. Tanaka, M. Kimura, Y. Kakuta, *Mol. Cell Biochem*. **2011**, *350*, 155-162.

[17] a) F. H. Niesen, H. Berglund, M. Vedadi, *Nat. Protoc*. **2007**, *2*, 2212-2221; b) J. K. Kranz, C. Schalk-Hihi, *Methods Enzymol*. **2011**, *493*, 277-298.

[18] J. Sambrook, E. F. Fritsch, T. Maniatis *Molecular cloning. A laboratory manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y, 2nd edn., **1989**.

[19] A. Bastida, A. Fernández-Mayoralas, R. Gómez, F. Iradier, J. C. Carretero, E. García-Junceda, *Chem. Eur. J*. **2001**, *7*, 2390-2397.

[20] M. A. Larkin, G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, J. D. Thompson, T. J. Gibson, D. G. Higgins, *Bioinformatics*, **2007**, *23*, 2947-2948.

[21] a) W.L. Jorgensen, J. Tirado-Rives, *J. Am. Chem. Soc*., **1988**, *110*, 1657-1666; b) W. L. Jorgensen, D. S. Maxwell, J. Tirado-Rives, *J. Am. Chem. Soc*. **1996**, *118*, 11225-11236; c) D. Shivakumar, J. Williams, Y. Wu, W. Damm, J. Shelley, W. Sherman, *J. Chem. Theory Comput*., **2010**, *6*, 1509-1519

Tables

Table 1. Apparent kinetic parameters of HocAST for *p*-NPS and PAP.[a]			
Substrate	K_M (μM)	V_{max} (μmol/min/mg)	V_{max}/K_M
p-NPS	45.9 ±2.2	0.84 ±0.008	0.018
PAP	4.0 ±0.9	0.76 ±0.015	0.19
[a] Kinetic parameters calculated at 25 ºC and pH 7.0			

Table 2. Screening of HocAST acceptor specificity.		
Compound	ΔT_m (ºC)	Activity (%)[a]
3	2.0	12.0
4	0.8	18.0
5	—[b]	N/A[c]
6	—[b]	N/A[c]
7	—[b]	N/A[c]
10	—[b]	N/A[c]
11	2.0	93.0
12	—[b]	N/A[c]
13	—[b]	N/A[c]
14	1.4	4.4
15	2.8	17.0
16	1.0	4.4
17	2.4	9.7
18	—[b]	N/A[c]
19	—[b]	N/A[c]
[a] 100 % of activity is the maximum reaction rate with PAP as acceptor and *p*-NPS as sulfuryl donor. [b] No increase of the T_m was detected. [c] N/A = Not applicable.		

Figure legends

Scheme 1. PAPS regeneration cycle based in the use of aryl sulfotransferases.

Figure 1. Alignment of the amino acid sequence of the hypothetical aryl sulfotransferase from *Haliangium ochraceum* versus three known sulfotransferases. The boxes (A, B, C and D) frame the conserved regions involved in the aryl sulfotransferase activity.

Figure 2. (A) Modelling of the tertiary structure of the protein HocAST; the reactive His105 is show in sticks and the conserved motifs in colours: Box A, density blue; Box B, cyan; Box C, deep blue and Box D, light blue; (B) Docking of *p*-NPS and PAP in the HocAST modelling. For details see text.

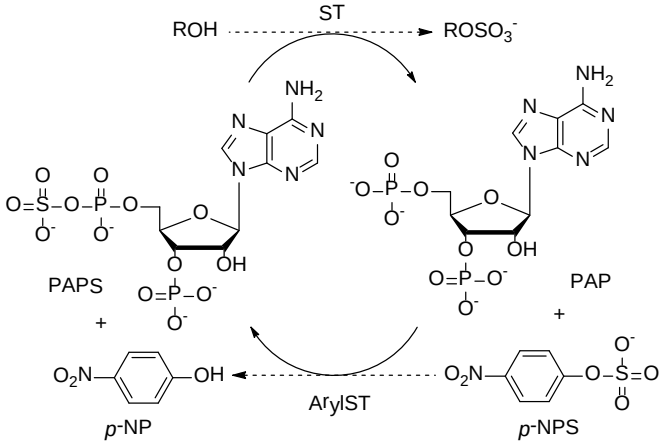
Figure 3. Peptide mass fingerprint of HocAST. The sequence of the identified peptides is shaded and underlined. The molecular mass of each peptide is indicated in Da.

Figure 4. Spectroscopic characterization of recombinant HocAST. (a) Far-UV circular dichroism spectrum. (b) Tryptophan fluorescence emission spectrum; the fluorescence intensity (FI) is presented in arbitrary units (au).

Figure 5. Thermal stability of the recombinant HocAST at 4 ºC (●), 25 ºC (▽), 37 ºC (■) and 45 ºC (◇).

Scheme 2. Structure of the compounds assayed in this study as possible substrates of HocAST.

Schemes and figures



Scheme 1.

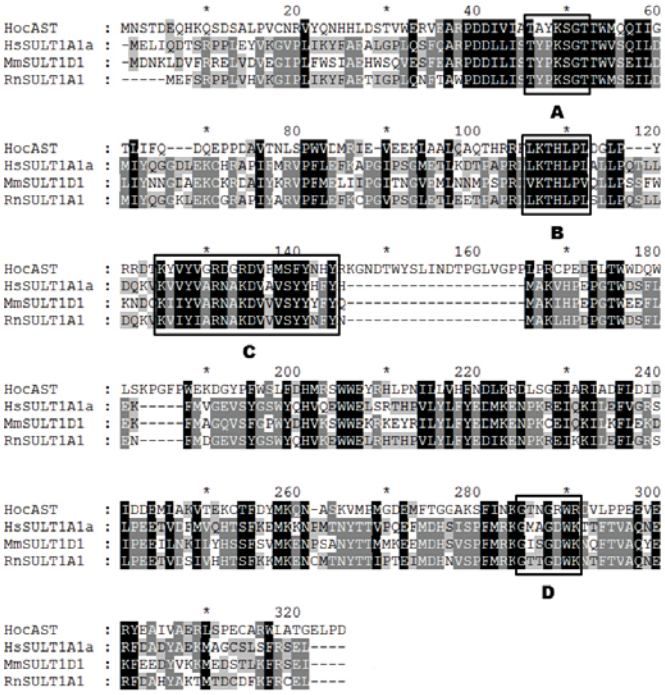


Figure 1.

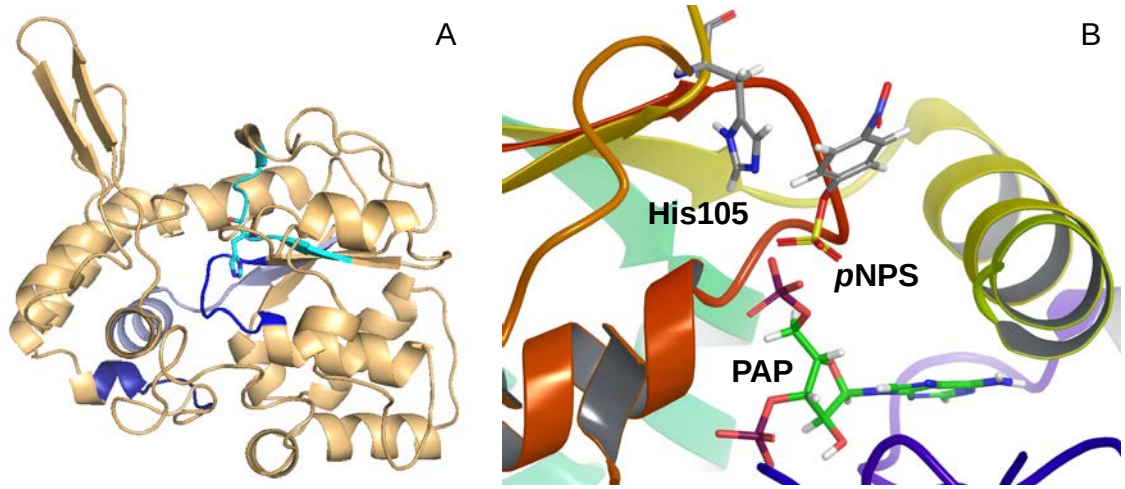


Figure 2.

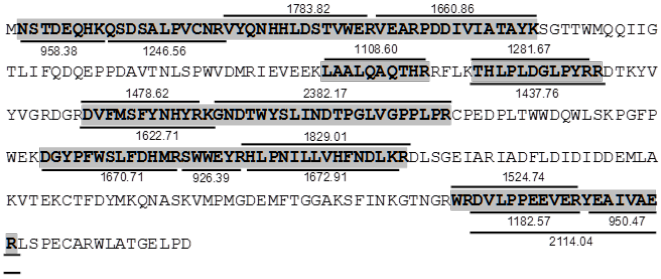


Figure 3.

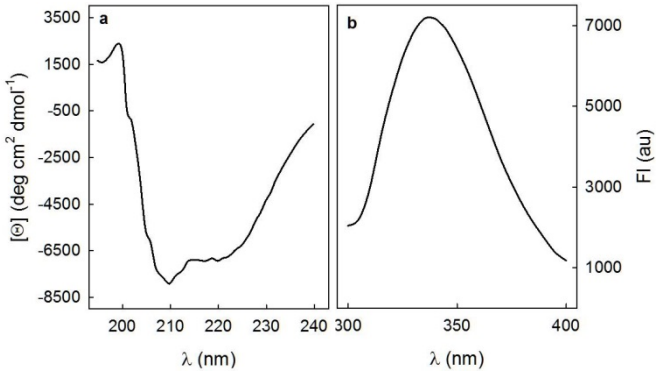


Figure 4.

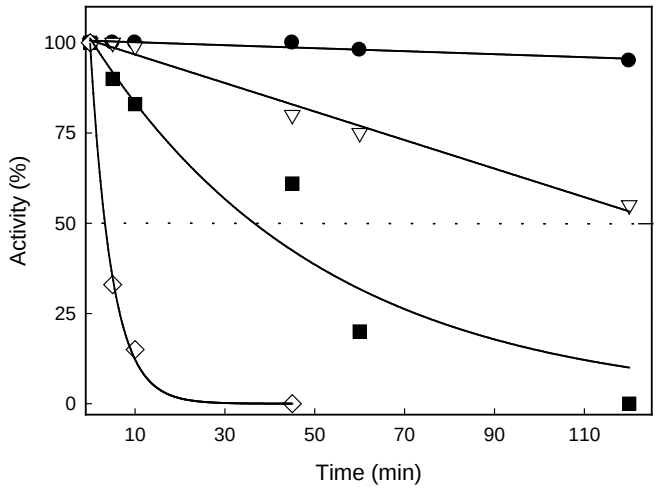
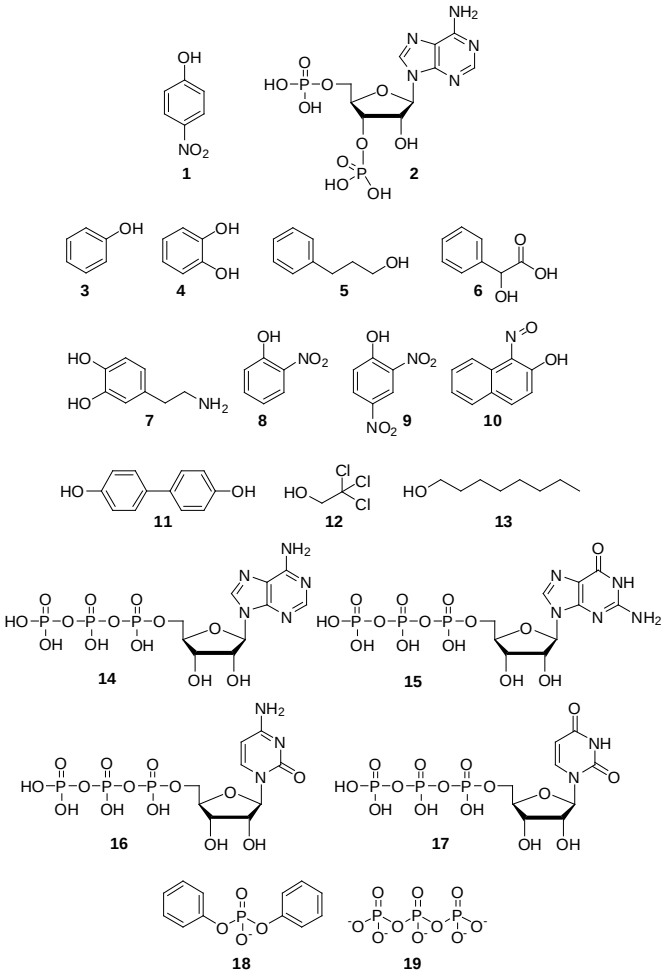


Figure 5.



Scheme 2.

Aryl sulfotransferase from *Haliangium ochraceum*: a versatile tool for sulfation of small molecules

*Iván Ayuso-Fernández, Miquel A. Galmés, Agatha Bastida and Eduardo García-Junceda**

Supporting Information

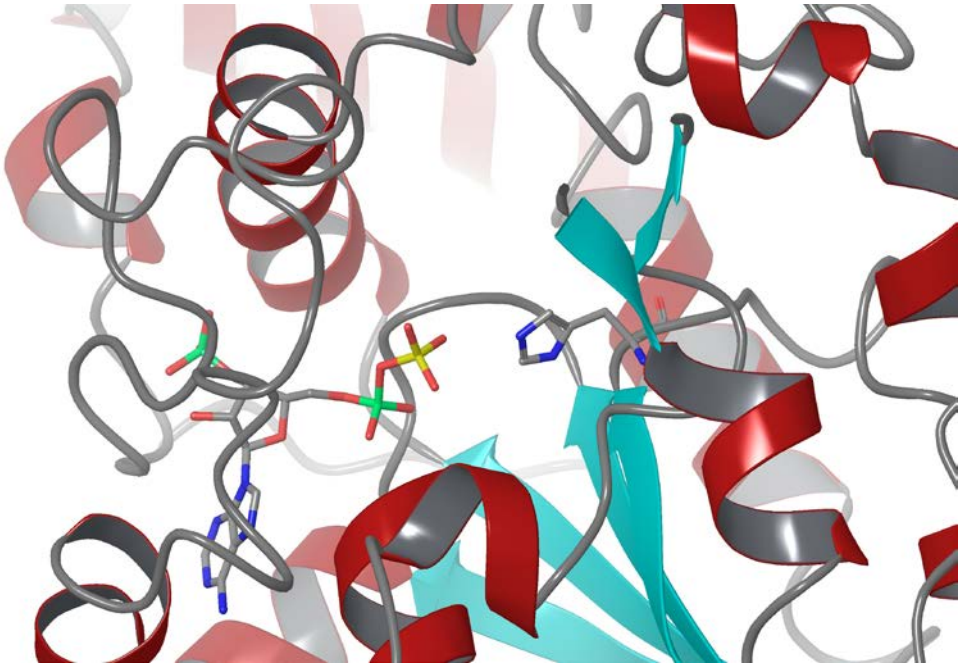


Figure S1. Crystallographic MmSULT1D1-PAPS complex (PDB entry 2zyt).[1] PAPS and His108 are shown in sticks.

[1] Berger I, Guttman C, Amar D, Zarivach R, Aharoni A (2011) The Molecular Basis for the Broad Substrate Specificity of Human Sulfotransferase 1A1. PLoS ONE 6(11): e26794. doi:10.1371/journal.pone.0026794

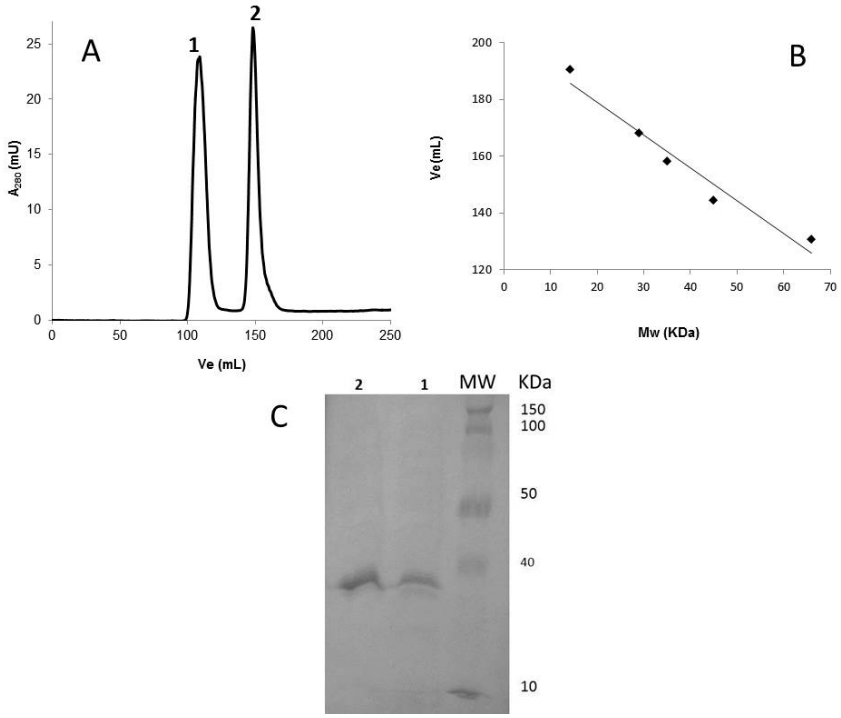


Figure S2. A) FPLC chromatogram of HocAST shows two main peaks with at elution volume of 110 mL (1) and 150 mL (2). B) Calibration curve prepared with proteins of known mass. C) SDS-PAGE of peaks 1 and 2.

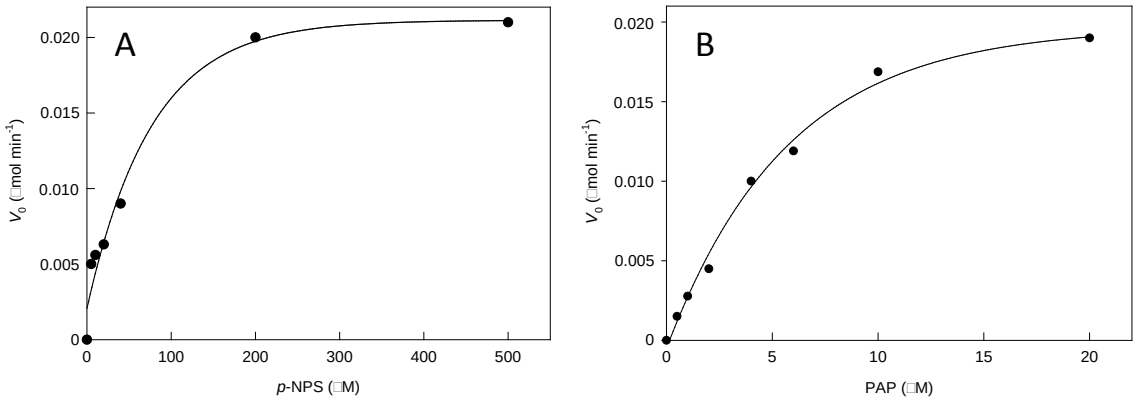


Figure S3. Substrate kinetics for HocAST at 25ºC and pH 7.0. (A) *p*-NPS concentration 0-500 μM with a fix concentration of PAP (20 μM); (B) PAP concentration 0-20 μM with a fix concentration of p-NPS (1.0 mM). In all cases reactions were stopped with NaOH 1N.

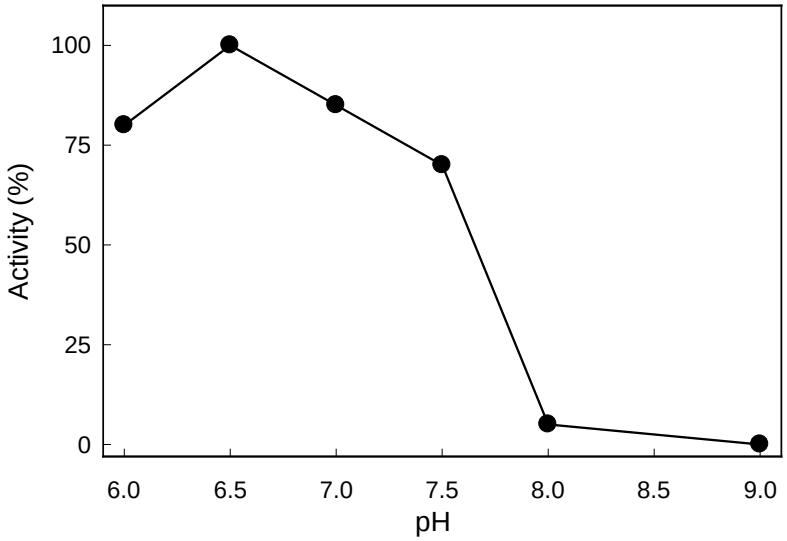


Figure S4. HocAST pH-activity profile.

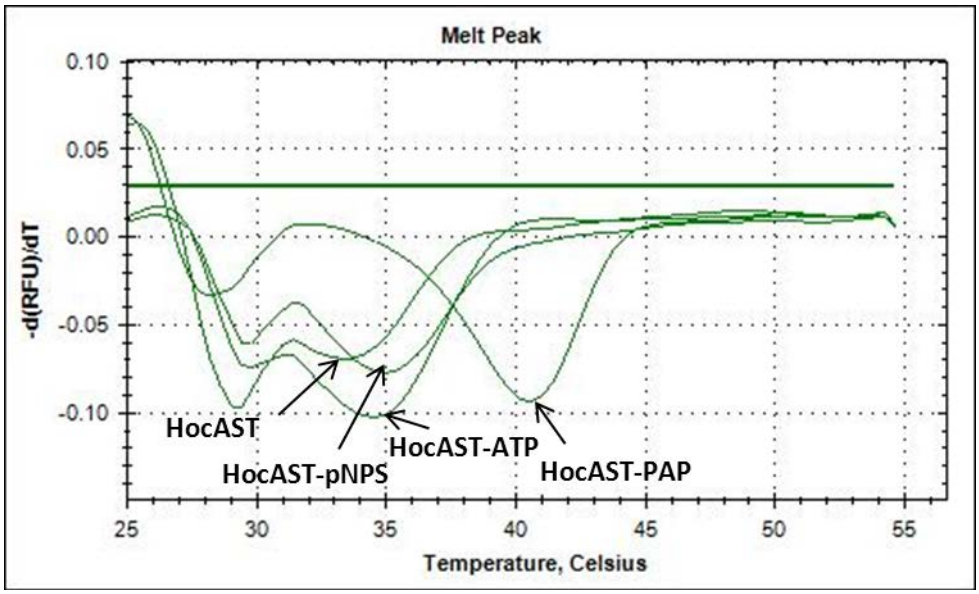


Figure S5. Thermal shift experiment of HocAST alone and with some ligands. In the case of the enzyme without any ligands a minimum at 33 ºC can be observed. This value corresponds to the melting temperature of the recombinant enzyme.