This is an Accepted Manuscript, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. This Accepted Manuscript will be replaced by the edited, formatted and paginated article as soon as this is available.

You can find more information about Accepted Manuscripts in the Information for Authors.

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal’s standard Terms & Conditions and the Ethical guidelines still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this Accepted Manuscript or any consequences arising from the use of any information it contains.
Phenylalanine Iminoboronates as New Phenylalanine Hydroxylase Modulators

Francesco Montalbano, João Leandro, Gonçalo D. V. F. Farias, Paulo R. Lino, Rita C. Guedes, João B. Vicente, Paula Leandro, Pedro M. P. Gois

Herein we report the discovery of new modulators of human phenylalanine hydroxylase (hPAH) inspired on the structure of its substrate and regulator L-Phenylalanine. These new hPAH modulators were simply prepared in good-to-excellent yields and excellent diastereoselectivities, based on a boron promoted assembly of L-Phenylalanine, salicylaldehyde and aryl boronic acids. Iminoboronate 8, prepared with L-Phenylalanine, para-methoxy-salicylaldehyde and phenyl boronic acid, was identified as the most efficient hPAH modulator, with an apparent binding affinity nearly identical to the natural allosteric activator L-Phenylalanine.

Introduction
Phenylketonuria (PKU) is the most frequent disorder of the amino acid metabolism with an overall incidence of one in 15,000 births and is characterized by intolerance to nutritional intake of L-Phenylalanine (L-Phe). Untreated, PKU leads to a severe psycho-motor impairment due to the toxic effect of increased L-Phe levels in the central nervous system that has been recently shown to self-assemble into toxic amyloid-like fibrils suggesting an amyloid etiology for PKU. The low levels of L-Phe-derived biosynthesized neurotransmitters also contributes to negative clinical outcome. In most cases, PKU is related with deficient activity of phenylalanine hydroxylase (PAH) exhibiting kinetic and conformational defects imposed by mutations in the PAH gene. Human PAH (hPAH) belongs to the family of aromatic amino acid hydroxylases and catalyzes the hydroxylation of L-Phe to L-tyrosine (L-Tyr) in the presence of the cofactors (6R)-L-erythro-5,6,7,8-tetrahydrobiopterin (BH₄, Scheme 1) and a non-heme mononuclear iron ion, with dioxygen as co-substrate, which is the first step of the catabolic degradation of L-Phe. To avoid L-Phe accumulation up to neurotoxic levels, PKU patients are forced to stringently hold on to an L-Phe-free diet, which often results in malnutrition and neurologic problems. Alternative strategies to treat PKU are now emerging, such as dietary supplementation with the natural cofactor BH₄, which is able to act as a pharmacological chaperone. Unfortunately, BH₄ is often used in high amounts per dose (20mg/Kg body weight) and patients with more severe phenotypes of PAH deficiency are unresponsive to this molecule. Therefore, the discovery of small molecule modulators of hPAH remains as a very important and challenging topic of research. In this context, compounds I and II (Scheme 1) were recently disclosed and shown to improve hPAH stability and the protein in vitro and in vivo steady-state levels.

Human PAH is a homotetrameric enzyme finely regulated by BH₄ and L-Phe. The cofactor inhibits the enzyme rendering a more stable form. Pre-incubation with the substrate L-Phe activates (~2.5-fold activation) the enzyme. Although this activation mechanism is not yet fully comprehended and there is no agreement on whether it is caused by L-Phe binding to an allosteric site or by homotropic activation.

Scheme 1. Boron promoted assembly of salicylaldehyde (SA), L-Phenylalanine (Phe) and boronic acids (PBA).

Regardless of its action mode, hPAH activation by L-Phe is the most important physiological mechanism to protect the organism from damage by increased levels of L-Phe. Therefore...
it is rather surprisingly that L-Phe has not been explored as the key structural motif to develop new modulators for this protein envisaging improving enzyme function.\textsuperscript{9,17} Recently we have demonstrated that boron may be efficiently used to assemble complex molecular structures that may be readily tuned for optimal interaction with a biological target (Scheme 1).\textsuperscript{18,19} Based on this, we envisioned that by incorporating L-Phe as one of the assembly components, we could simply generate useful structures to modulate hPAH activity and probe the enzyme active site.

Results
To test this idea, iminoboronates 1, 2, and 3 depicted in Scheme 2, were prepared using salicylaldehyde (SA), phenylboronic acid (PBA), L-Phe, L-leucine (L-Leu) and L-alanine (L-Ala) respectively as assembly components. The reaction was conducted in water at 90°C for 20h, and this simple protocol afforded the expected compounds in good-to-excellent yields and excellent diastereoselectivities. Once prepared, compounds 1-3 were readily evaluated for their effect on the activity of tetrameric wild-type hPAH, employing three experimental conditions (supplementary Scheme S1): i) condition I involved pre-incubation of hPAH with substrate and compound to evaluate the competition between the iminoboronate and L-Phe (‘substrate-activated’ condition); ii) condition II was performed with no pre-incubation step, the substrate and iminoboronate being added simultaneously at time zero (‘non-activated’ condition); and iii) condition III involved pre-incubation with the tested iminoboronate alone, to establish its ability to pre-activate the enzyme, mimicking L-Phe-promoted pre-activation (‘compound-activated’ condition). Pre-activation by each iminoboronate was thus calculated as the ratio between the activity in condition III and condition II. Blank assays with each iminoboronate alone and omitting L-Phe were performed to rule out L-Phe release from the iminoboronates bearing this moiety and consequent conversion to L-Tyr.

Very gratifyingly, as shown in Scheme 2, the iminoboronate 1 was able to modulate the activity of hPAH, activating the enzyme by 1.5-fold (P < 0.05) in the absence of L-Phe, while competing with the substrate L-Phe. Very differently, the iminoboronates 2 and 3, respectively prepared with L-Leu and L-Ala, failed to activate hPAH. Taken together with the fact that the individual components SA and PBA were also unable to activate the enzyme, these results clearly suggest that the observed effect was most probably due to the incorporation of L-Phe into the structure of heterocycle 1. Due to the high demands of the hPAH activity assay, we devised an experimental approach to determine the compounds’ binding kinetics, evaluating their effect on the protein’s thermal denaturation profiles, employing differential scanning fluorimetry (DSF, see supporting information). By fluorescence-monitored thermal denaturation profiles at different compound concentrations, it was determined that 1 binds to hPAH with an apparent binding affinity of \( K_D = 14.4 \pm 2.9 \) \( \mu \text{M} \), while the iminoboronates 3, prepared with L-Ala, did not reveal any measurable affinity to hPAH (Scheme 2 and supporting information).
Scheme 3. Biological activity of hPAH in the presence of iminoboronates containing 1-phenylalanine and para substituted phenyl boronic acids: methyl (4), fluoro (5) and methoxide (6). The enzymatic assays correspond to the substrate activated (●), non-activated (□) and compound activated conditions (●). Control assays were performed in the presence of DMSO. C_{0.5} concentration for half-maximal binding. Data are presented as mean ± SD. Statistical significance is given by *P < 0.05; **P < 0.01; ***P < 0.001; ****P < 0.0001 (n = 3), for the relative fold activation between compound-activated (II) and non-activated condition (III).

Based on these encouraging results, we embarked on a modification campaign to optimize the structure of compound 1. Therefore, following the aforementioned methodology, the iminoboronates 4, 5 and 6 were simply prepared combining 1-Phe, salicylaldehyde and para substituted phenyl boronic acids. As shown in Scheme 3, compound 4, bearing a methyl at the boronic acid aromatic para position, demonstrated a profile indicative of competition with the substrate as observed for 1, though with only marginal activation ability. Very differently, compound 5, prepared with 4-fluorophenylboronic acid, markedly activated hPAH (1.5-fold; P < 0.001), while exhibiting an apparent binding affinity (C_{0.5} of 14.4 ± 2.9 µM) comparable to the one observed when using 1. The introduction of a methoxide substituent in compound 6 resulted in a clear inhibition (Scheme 3).

Then, we evaluated the impact of substituents at the salicylaldehyde component with the synthesis of compounds 7 and 8 (Scheme 4). Very gratifyingly, although the introduction of a methyl at the para position of the salicylaldehyde had a marginal impact on the enzyme activation, the introduction of a methoxide in compound 8, clearly improved the hPAH activation by 1.8-fold (Conditions II and III in Scheme 4; P < 0.0001), maintaining a high apparent binding affinity (C_{0.5} of 14.8 ± 4.9 µM). Compound 8 displayed a 1.7h t_{1/2} in buffer (see ESI). As shown is Scheme 4, the combination of para-methoxy-salicylaldehyde with different aromatic and vinylboronic acids 9-13 (Scheme 4) did not improve the activation previously observed with compound 8, except for a slight increase (1.3-fold; P < 0.001) observed for the 2-bromophenylboronic acid substituent in compound 13.

Scheme 4. Biological activity of hPAH in the presence of iminoboronates containing 1-phenylalanine and substituents at the salicylaldehyde (7 and 8) and para-methoxy-salicylaldehyde in combination with aromatic and vinylboronic acids (9 to 13). The enzymatic assays correspond to the substrate activated (●), non-activated (□) and compound activated conditions (●). Control assays were performed in the presence of DMSO. C_{0.5} concentration for half-maximal binding. Data are presented as mean ± SD. Statistical significance is given by *P < 0.05; **P < 0.01; ***P < 0.001; ****P < 0.0001 (n = 3), for the relative fold activation between compound-activated (II) and non-activated condition (III).

Aiming at a rationalization of the observed biological activity when using substrate 1-Phe or compounds 1, 3, 6, 8, and to get insight into their potential binding and interactions with hPAH, we performed detailed in silico molecular docking studies in the active site of hPAH using the GOLD 5.1 software (Figure 1). Very interestingly, compounds 1 (Fig. 1B) and 8 (Fig. 1E and F), which elicit a ‘competitive-activator’ profile, adopt similar poses inside hPAH active site, with the phenyl ring in the same position of the phenyl ring of 1-Phe, showing only a small ring rotation. These compounds likely establish π-π interactions with the imidazole group of His285 at an average distance of 3.1 Å. This is particularly important as interactions of thiophyl and alkyl lateral chains of the substrate analogues 3-(2-thienyl)-1-alanine (THA; 3.8 Å) and norleucine (NLE; 3.9 Å) respectively, with His285 were reported as governing substrate binding affinity. In addition, the compound 8 methoxy moiety oxygen is predicted to form a hydrogen bond with Ser251 (1.9 Å), a residue that is also known to establish a hydrogen bond with the dihydroxypropyl side chain of the cofactor BH₄ (see Fig. S5 in SI). Other hydrophobic interactions are also observed between these compounds and Phe254, another residue essential for BH₄ binding. Differently, compound 6 which markedly inhibits hPAH, shows a completely different pose inside the active site (Fig. 1D), displaying no interactions with Ser251 and adopting a pose shifted to the right side of the pocket, though being unable to...
establish any interaction with the final part of the binding pocket. Finally compound 3 prepared with l-Ala, displayed no important interactions inside the binding pocket (Fig. 1C), corroborating the importance of the benzyl group for recognition. In fact, the benzyl group of compound 8 adopts a perfect position inside the binding pocket, positioning the phenyl ring in a conformation close to the iron (Fig. 1F). The superposition of the crystallographic structure of hPAH complexed with BH₄ and NLA (PDB code 1MMT) and the docking poses of compounds 1 and 8 revealed that these compounds are almost completely overlapped inside the binding site. The phenyl ring of compound 8 overlaps the alkyl chain of NLA and probably governs compound affinity and activation. The methoxy moiety of compound 8 occupies the BH₄ binding site explaining the slight decrease of hPAH activity observed for the l-Phe pre-activated conditions (see Fig. S6 in SI). It also points to a reversible binding of the compounds, as BH₄ and l-Phe are still able to bind to the protein, with concomitant production of l-Tyr (Scheme 1 and 4). The estimate of the binding energies (scores) of compounds 1 and 8 calculated with Goldscore fitness function revealed that these compounds show a strongest affinity to the hPAH (scores = 62 and 60) compared to l-Phe (score = 49). The binding parameters of the compounds determined by DSF (Fig. S1 and S2) are in the low micromolar range (10⁻⁴–14 µM), comparable with the apparent binding of l-Phe (16.3 ± 6.3 µM) determined by the same method.

In comparison with molecules I and II, which exhibit very promising results in the stabilization of hPAH, these iminoboronates, namely compound 8, are able to improve directly the enzyme activity by a pre-activation mechanism and to predispose the hPAH enzyme to accommodate the natural substrate.

**Conclusions**

In summary, in this study we have developed for the first time a series of iminoboronates that are able to directly increment hPAH activity by a pre-activation mechanism similar to the one induced by the substrate l-Phe. To this end, we have chosen to study the compounds on the fully active wild-type tetrameric hPAH, in the perspective of a future application of these iminoboronates in the activation of clinically relevant variants. Since missense mutations often originate functionally impaired variants, the most effective iminoboronates are expected to significantly increase the variant hPAH enzymatic activity to functionally relevant levels. The studied compounds were very efficiently prepared using a boron promoted assemblage of l-Phe with salicylaldehydes and aryl boronic acids. The resulting iminoboronates featuring the l-Phe (1, 4-6 and 8) structure were shown to interact with hPAH, while iminoboronates prepared with l-Leu (2) and l-Ala (3) did not. The l-Phe based iminoboronates (1, 5 and 8) exhibit apparent binding constants in the micromolar range (10⁻⁴–14 µM), which are comparable with the apparent binding constant of l-Phe (16.3 ± 6.3 µM). In this study, compound 8, prepared with l-Phe, para-methoxy-salicylaldehyde and phenyl boronic acid, was identified as the most effective activator of hPAH improving the enzyme activity by 1.8-fold (P< 0.0001), maintaining a high apparent binding affinity (C₅₀ of 14.8 ± 4.9 µM). Docking studies performed with these compounds, corroborated all experimental observations and revealed that iminoboronates 8 establishes important interactions either with the substrate and BH₄ recognition sites.

**Experimental Section**

*Escherichia coli* TOP 10 and the prokaryotic expression vector pTrcHis were obtained from Invitrogen (Carlsbad, CA). The cofactor (6R)-l-erythro-5,6,7,8-tetrahydrobiopterin (BH₄), l-Phe, Hepes were from Sigma (St. Louis, MO, USA). Ascorbic acid was obtained from Merck (Darmstadt, Germany). Unless stated otherwise, all reagents were of analytical grade.

**General procedure for preparation of boron heterocycles using water as a solvent**

A round bottom flask equipped with a magnetic stirrer was charged with amino acid (2.0 equiv.), aldehyde (1.5 equiv.) and distilled water (2.0 mL). This suspension was stirred at 90°C for 1 h after which the boronic acid (0.41 mmol) was added, the mixture was then stirred at 90°C for 20 h. The reaction mixture, which appears as a biphasic composition of precipitate and a supernatant liquid, was filtered and the solid retained in the filter was then washed with water followed by hexane. The desired compound was recovered with dichloromethane, which was subsequently removed under reduced pressure.

**Expression and purification of recombinant wild-type hPAH protein**

---

**Figure 1.** Best docking poses obtained for l-Phe (A), compounds 1 (B), 3 (C), 6 (D) and 8 (E). (F) Possible interactions between compound 8 (yellow) and hPAH residues at the active site (green), resulting from docking of the compound onto the hPAH structure (PDB code 1MMT).
Recombinant wild-type hPAH protein was expressed in E. coli as a fusion protein with the hexa-histidyl tag (6xHis-(pep)EK-hPAH) as described. Cells were grown at 37°C, protein expression was induced by addition of 1 mM isopropyl-β-D-thio galactoside (IPTG), and the cells were harvested after 3 h. After a first purification step by immobilized metal-affinity chromatography, the tetramers were isolated by size exclusion chromatography, using a HiLoad Superdex 200 HR column (1.6 cm x 60 cm, GE-Healthcare) and a mobile phase containing 20 mM NaHepes, 200 mM NaCl, pH 7 pumped at a flow rate of 0.7 mL·min⁻¹.

Enzymatic activity assays

The hPAH activity was measured essentially as previously described in a 200 µL final volume reaction mixture, containing 100 µM L-Phe, 0.1 M Na-Hepes, pH 7, 0.1 mg·mL⁻¹ catalase, 5 µg of recombinant wild-type hPAH tetramers, 100 µM of each compound or 1% DMSO (vehicle control). After 4 minutes of pre-incubation, 100 µM (NH₄)₂Fe(II)SO₄ was added and, unless otherwise stated, the reaction was started by addition of 75 µM BH₄ (together with 5 mM ascorbic acid) after 1 minute incubation with the iron (condition I in supplementary Scheme S1; ‘substrate-activated’ condition). To evaluate pre-activation of the enzyme by the specific activity of the non-activated hPAH, 100 µM L-Phe and 100 µM of each compound were added together with 75 µM BH₄ at the start of the hydroxylation reaction (condition II in supplementary Scheme S1; ‘non-activated’ condition). To evaluate pre-activation of the enzyme by the compound, hPAH was pre-incubated 4 minutes with each compound whereas the L-Phe substrate was only added at the start of the reaction, together with 75 µM BH₄ at the start of the reaction (condition III in supplementary Scheme S1; ‘compound-activated’ condition). Pre-activation by each iminoboronates was thus calculated as the ratio between the activity in condition III and condition II. Blank reactions where the substrate L-Phe were omitted were also made for each compound. The amount of L-Tyr produced after 1 min was quantified by HPLC method using a LiChroCART 250-4 LiChropher® 60 RP-select B (5 µm) column (Merck KGaA, Darmstadt, Germany), a 5% ethanol mobile phase pumped at 0.7 mL·min⁻¹ and fluorimetric detection (λexc= 274 nm and λem= 304 nm). Specific activities are presented as mean ± SEM obtained from three independent experiments. Tests for statistical significance were performed using 1-way ANOVA with *P < 0.05; **P < 0.01; ***P < 0.001 and ****P < 0.0001.

Differential Scanning Fluorimetry

Differential scanning fluorimetry (DSF) was performed in a C1000 Touch thermal cycler equipped with a CFX96 optical reaction module (Bio Rad). For all fluorescence measurements, samples containing purified recombinant wild-type hPAH tetramers at 100 µg·mL⁻¹ in 20 mM NaHepes, 200 mM NaCl, pH 7, 2.5-fold Sypro Orange (Invitrogen; 5000-fold commercial stock solution), 1% DMSO (unless otherwise stated) and 100 µM of each compound were incubated at 20 °C for 10 minutes. The PCR plate was sealed with Optical-Quality Sealing Tape (Bio-Rad) and centrifuged at 500xg for 1 min. The DSF assay was carried out by increasing the temperature from 20 to 90 °C, with a 1 s hold time every 0.2 °C and fluorescence acquisition using the FRET channel. Control experiments in the absence of DMSO and/or compounds were routinely performed in each microplate. Data were processed using CFX Manager Software V3.0 (Bio-Rad) and the GraphPad Prism 6. Temperature scan curves were fitted to a biphasic dose-response function and the Tm values were obtained from the midpoint of the first and second transitions. To monitor the binding properties of the regulatory and catalytic domain towards each compound, DSF assays were run in the presence of increasing compound concentrations (0-2.56 mM) or 1% DMSO (vehicle control). C₅₀ values provide apparent binding affinities and are best-fit parameters obtained from the effect of the compound on the contribution of the regulatory domain to the overall unfolded process (compounds 1, 5 and 8) or on the melting temperature of the first transition (Tm₁) (compound 3).

Docking Studies

All calculations were performed on iMed.ULisboa scientific cluster. GOLD (version 5.2)²⁴ was used for docking calculations and Molecular Operating Environment (MOE) software (version 2013.10)²⁵ were used to build and optimize the structures of iminoboronates molecules and for enzyme structure refinement. In the present study, the crystal structure of hPAH in complex with the physiological cofactor (6R)-L-erythro-5,6,7,8-tetrahydrobiopterin (BH₄) and the substrate analogue NLE at a resolution of 2.0 Å (PDB code: 1MMT)²⁶ was employed in the docking calculations. The ternary hPheOH-Fet(II)-BH₄-NLE structure comprises 308 residues, 149 structural water oxygen atoms, one NLE molecule, one BH₄ molecule and one sulphate ion. The physiological cofactor (6R)-L-erythro-5,6,7,8-tetrahydrobiopterin (BH₄), the substrate analogue NLE and all crystallographic water molecules were removed from the coordinate set using MOE software. Hydrogen atoms were added to this reduced crystal structure and the protein was protonated to pH = 7. The protein was then submitted to restrained molecular mechanics refinement using the AMBER99 force field implemented in MOE software. The final structure of the hPAH protein was used for the docking calculations. To establish our docking procedure, a preliminary validation study was carried out by redocking NLE to the refined protein structure. The performance of molecular docking protocol was evaluated by comparing the redocked binding poses of NLE with the experimental X-ray (PDB code 1MMT). The redocked pose agrees well with the X-ray crystallized pose (RMSD ≤ 2.0 Å). After this validation, the iminoboronates structures were docked into the active site of hPAH using the Gold Software with the goldscore scoring mode. The following GOLD parameters were employed: 1000 runs, population size of 100,000 genetic algorithm operations, and 5 islands at normal time speed up setting, were conducted for each compound. In the docking process, the protein conformation was fixed while the docked ligand was flexible. The best binding pose for each molecule was saved for further analysis. The 20 top-ranked compounds were visually inspected with PyMol.²⁶

Acknowledgements
Notes and references

* Research Institute for Medicines (iMed.ULisboa), Faculty of Pharmacy, Universidade de Lisboa, Av. Prof. Gama Pinto, 1649-003 Lisboa, Portugal.
Fax: +351 217946470; Tel: +351 217946400.
E-mail: pedrogois@ff.ulisboa.pt; aleandro@ff.ulisboa.pt
Homepage: http://www.ff.ul.pt/~pedrogois/index.html
# These authors contributed equally to this work
† Electronic Supplementary Information (ESI) available: Full description of compounds characterization, enzymatic activity assays, differential scanning fluorimetry and Docking Studies. See DOI: 10.1039/b000000x/
